B-aggressive Lymphoma Family Proteins Have Unique Domains That Modulate Transcription and Exhibit Poly(ADP-ribose) Polymerase Activity*

Received for publication, May 17, 2005, and in revised form, June 30, 2005 Published, JBC Papers in Press, August 1, 2005, DOI 10.1074/jbc.M505408200

Ricardo C. T. Aguiar1, Kunihiko Takeyama, Chunyan He, Katherine Kreinbrink, and Margaret A. Shipp2

From the Division of Hematological Malignancies, Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115

BAL1 (B-aggressive lymphoma 1) was originally identified as a risk-related gene in diffuse large B-cell lymphoma. BAL1 encodes a nuclear protein with N-terminal macro domains and a putative C-terminal poly(ADP-ribose) polymerase (PARP) active site. Macro domains are sequences homologous to the non-histone region of histone macroH2A. Several lines of evidence suggest that these domains may modulate transcription, including a high concentration of histone macroH2A in the inactive X chromosome, direct interference with transcription factor binding in a positioned nucleosome, and structural similarity to DNA binding domains. Poly(ADP-ribosylation) is a critical post-translational modification that regulates chromatin configuration and transcription. In this report we describe two additional BAL family members, BAL2 and BAL3, with N-terminal macro domains and putative C-terminal PARP active sites and assess the function of these specific regions in BAL family members. Herein, we demonstrate that BAL macro domains repress transcription when tethered to a promoter. In addition, we show that BAL2 and BAL3, but not BAL1, exhibit PARP activity. In agreement with these data, BAL1 lacks several critical donor and acceptor residues that are conserved in the BAL2 and -3 PARP active sites. Of interest, BAL1 protein includes a C-terminal region with homology to the catalytic domain of PARP (poly [ADP-ribose] polymerase) proteins. PARP catalyzes the transfer of ADP-ribose onto acceptor proteins using NAD+ as a substrate (9, 10). The resulting poly(ADP-ribose)lation of acceptor proteins has been principally associated with functions of the prototype PARP-1 enzyme and its role in DNA repair (9). Recently, other PARP domain-containing proteins have been identified and implicated in additional nuclear processes (11–13). This diverse group of proteins contains a minimal PARP active site embedded within a larger polypeptide (9). In these proteins, termed the PARP superfamily, the functional consequences of PARP activity may be dependent on additional functional domains. For example, the PARP family member, Tankyrase 1, controls telomere physiology via poly(ADP-ribose)lation and inhibition of TRF-1, a negative regulator of telomere length (14).

In our initial description of BAL-1, we noted homology between the BAL-1 C terminus and the Tankyrase 1 PARP-like catalytic domain (1). However, BAL-1 did not exhibit PARP-like activity in preliminary functional assays. More recently, we identified additional BAL family members with potential critical differences in their C-terminal PARP-like regions. Given the emerging evidence that PARP family proteins represent possible cancer treatment targets (15–17) and the reported association between BAL-1 expression and chemoresistance in DLBCL, we performed detailed structure/function analyses of BAL family members.

The BAL1 (B-Agressive lymphoma 1) gene was originally identified in a genome-wide search for risk-related genes in diffuse large B-cell lymphoma (DLBCL)1 (1). In a pilot series of primary DLBCLs, BAL expression was significantly higher in chemoresistant tumors. In our initial report, we also showed that BAL1 encodes a nuclear protein with a duplicated N-terminal domain homologous to the non-histone region of histone macroH2A (2). This ~135-aa non-histone region, which contains a short stretch of basic residues and a putative leucine zipper motif, has been termed a macro domain (3, 4).

Several lines of evidence suggest that macro domains may modulate transcription. First, histone macroH2A is enriched in the inactive mammalian X chromosome, potentially implicating the macro domain-containing protein in gene silencing (5). More specifically, the macro domain of histone macroH2A represses transcription in vitro (6) and directly interferes with transcription factor binding in nucleosomes containing this variant histone (7). In addition, the crystallographic structure of a macro domain-only protein (AF1521, from Archaeoglobus fulgidus) includes a known DNA binding motif, further suggesting that macro domains might interact with nucleic acids (8).

In addition to containing N-terminal candidate macro domains, the BAL1 protein includes a C-terminal region with homology to the catalytic domain of PARP (poly [ADP-ribose] polymerase) proteins. PARPs catalyze the transfer of ADP-ribose onto acceptor proteins using NAD+ as a substrate (9, 10). The resulting poly(ADP-ribose)lation of acceptor proteins has been principally associated with functions of the prototype PARP-1 enzyme and its role in DNA repair (9). Recently, other PARP domain-containing proteins have been identified and implicated in additional nuclear processes (11–13). This diverse group of proteins contains a minimal PARP active site embedded within a larger polypeptide (9). In these proteins, termed the PARP superfamily, the functional consequences of PARP activity may be dependent on additional functional domains. For example, the PARP family member, Tankyrase 1, controls telomere physiology via poly(ADP-ribose)lation and inhibition of TRF-1, a negative regulator of telomere length (14).

In our initial description of BAL-1, we noted homology between the BAL-1 C terminus and the Tankyrase 1 PARP-like catalytic domain (1). However, BAL-1 did not exhibit PARP-like activity in preliminary functional assays. More recently, we identified additional BAL family members with potential critical differences in their C-terminal PARP-like regions. Given the emerging evidence that PARP family proteins represent possible cancer treatment targets (15–17) and the reported association between BAL-1 expression and chemoresistance in DLBCL, we performed detailed structure/function analyses of BAL family members.

MATERIALS AND METHODS

Cloning and Chromosomal Localization of BAL2 and BAL3—BAL1 cDNA and protein sequences were used to search human nucleotide and protein databases (www.ncbi.nlm.nih.gov/BLAST). Candidate homologue genes were further characterized by 5′-RACE-PCR and reverse transcription-PCR. For 5′-RACE-PCR, RNAs from normal mature B-cells and DLBCL cell lines were reverse-transcribed with a gene-specific primer (sequence available upon request) and submitted to successive rounds of nested amplification as previously described (18). To confirm the RACE-PCR findings, cDNAs from normal mature

<ref>THE JOURNAL OF BIOLOGICAL CHEMISTRY</ref>
B-cells and DLBCL cell lines were also used for reverse transcription-PCR with primers derived from the newly identified 5’ and 3’ sequences. To define the chromosomal location of BAL2 and BAL3, we initially searched the human genome data base with these gene sequences. To confirm these in-silica findings, we obtained a series of YAC (CEPH-Génétion, Paris, France) and BAC clones (BACPAC Resources Center, Children’s Hospital Oakland Research Institute, Oakland, California) with previously assigned chromosomal locations (www.ncbi.nlm.nih.gov/mapview). DNA from these clones was isolated, and PCR was performed as previously described (19).

**Analyses of BAL Family Sequences**—The full-length N- or C-terminal domains of BAL1-related proteins were analyzed in detail for conservation of critical residues using the MegAlign™ 5.8 software package (DNASTar, Inc., Madison, WI). These alignments were also displayed as phylogenetic trees and sequence distance matrices. Phylogenetic trees depict the evolutionary relationships predicted from the multiple sequence alignment where the length of each pair of branches represents the distance between sequence pairs. The distance matrix displays the divergence and percent identity values for each sequence pair in an alignment; divergence is calculated by comparing sequence pairs in relation to the phylogeny reconstructed by MegAlign™. Percent identity compares sequences directly without accounting for phylogenetic relationships. Note that divergence is not usually the inverse of percent identity (i.e. the sum of the percent identity and divergence values for a given pair is not usually 100).

**PARP Activity Assays**—The putative C-terminal PARP catalytic domains of the BAL family proteins (BAL1 aa 590–854, BAL2B aa 1455–1638, BAL3 aa 448–673) were fused to GST and the recombinant proteins generated using the pGEX-4T3 expression system (Amersham Biosciences). Poly(ADP-ribose) activity assays were carried out as previously described (14, 20). Reactions contained ~2 μg of purified BAL proteins or 0.1 μg of high purity recombinant PARP protein (Alexis Biochemicals, San Diego CA) as a positive control. Reactions were incubated at 25°C for 30 min in assay buffer (0.1 ml) containing 50 mM Tris-HCl, pH 8.0, 4 mM MgCl₂, 0.2 mM dithiothreitol, and with or without 200 ng of activated DNA (Sigma). In radioactive reactions 1.3 μM [³²P]NAD⁺ (4 μCi; PerkinElmer Life Sciences) was used, whereas 0.1–1 mM concentrations of unlabeled NAD⁺ (Roche Applied Science) served as substrate in nonradioactive assays. In certain assays the PARP inhibitor, 3-aminobenzamide (3-ABA) (Calbiochem), was included at 1 mM final concentration. Reactions were stopped by the addition of 20% trichloroacetic acid. Precipitated proteins were rinsed once in 5% trichloroacetic acid, suspended in SDS loading buffer, and fractionated by SDS-PAGE. In radioactive reactions, proteins were visualized by Coomassie Blue stain and exposed to autoradiography. For nonradioactive reactions, after size fractionation the proteins were transferred to polyvinylidene difluoride membranes (Millipore, Bedford, MA) and immunoblotted with a polyclonal antibody to poly(ADP-ribose) (LP98–10, Alexis Biochemicals).

**Transcription Repression Assays**—Human embryonic kidney 293 cells were cultured in Dulbecco’s modified Eagle’s medium (Invitrogen) containing 10% fetal calf serum (Invitrogen). For these experiments human embryonic kidney 293 cells were seeded on 6-well plates at 1 x 10⁵ cells/well 1 day before transfection. Cells were triple-transfected (Effectene Transfection Reagent, Qiagen, Valencia, CA) with combinations of plasmid DNAs including 100 ng of the report plasmid (Gal4-TK-Luc), 100 ng of the specific effector plasmid, and 100 ng of pCMVβ (Clontech, Palo Alto, CA) as previously described (6).

The effector plasmids were generated by PCR and included the Gal4 DNA binding domain (DBD) fused to full-length BAL1 (aa 2–854), BAL1 N terminus complete, including the macro domains (aa 2–464), BAL1 N terminus proximal (aa 2–119), BAL3 full-length (aa 2–673), BAL3 macro domains (aa 58–454), and BAL3 PARP domain (aa 450–673). All these fragments were cloned in-frame with amino acids 1–147 of Gal4 protein in a pcDNA3 backbone. Gal4 DBD-HDAC5 (histone deacetylase domain of HDAC5) served as a positive control. The latter construct as well as the reporter plasmid were gifts of Saadi Kochbin (Grenoble, France) and were described previously (6). All constructs were verified by sequencing.

Transfected cells were harvested after 24 h. Luciferase and β-galactosidase activities were measured using the luciferase assay system and β-galactosidase enzyme assay, respectively (Promega, Madison, WI). The β-galactosidase activity from pCMVβ and Gal4 DBD were used as control for transfection efficiency. All experiments were performed in triplicate. In dose-dependent experiments, the total amounts of plasmid DNA were kept constant by adding appropriate amounts of empty vectors. Western blot with an anti-Gal4 antibody indicated that all fusion proteins were expressed (data not shown).

**RESULTS**

The BAL Gene Family Is Composed of Three Genes—We previously identified an additional highly conserved BAL1-like sequence, suggesting that BAL1 is a member of a larger gene family (1). For this reason, we used BAL1 cDNA and protein sequences to search human EST, nucleotide, and protein databases (www.ncbi.nlm.nih.gov) and identify two related genes termed BAL2 and BAL3 (Fig. 1).

BAL2 corresponds to the partially characterized KIAA1268 sequence (Entrez GeneID 54625) (1). Using 5’-RACE-PCR, we identified additional BAL2 5’ sequences, generating a ~ 4.5-kilobase open reading frame that is 120 aa longer than the previously described N terminus (locus NM_017554, updated on December 2004). However, the extended BAL2 cDNA lacks an in-frame stop codon upstream of the most 5’ ATG, raising the possibility that BAL2 encodes a protein with additional N-terminal sequences. The existence of BAL2 rat and chicken orthologues with longer N termini supports this hypothesis (data not shown). Of interest, BAL2 includes three N-terminal macro domains (Fig. 1).

In addition, we used 3’-RACE-PCR to fully characterize the 3’ end of BAL2 and identify 2 alternatively spliced transcripts that encode 33- or 154-aa unique C termini (BAL2A and BAL2B, respectively, Fig. 1; GenBank™ accession numbers DQ063585 and DQ063584). Whereas BAL2B encodes a protein containing a putative PARP catalytic domain, BAL2A lacks these sequences, suggesting that these C-terminal differences might be functionally relevant. In Northern blot analysis of multiple human tissues, we found similar patterns of BAL1 and BAL2 expression with the highest transcript levels in lymphocyte-rich tissues (spleen, lymph node, and peripheral blood leukocytes) and the lowest in total brain (data not shown and Ref. 1).

An additional group of incompletely characterized BAL-like cDNA clones were identified that did not correspond to BAL2 (RefSeq accession NM_152615.1; mRNA for predicted protein FLJ40597, recently reassigned to Entrez GeneID 165631). Using reverse transcription-PCR, these sequences were organized into an open reading frame of ~2025 nucleotides, which encoded a 656-amino acid protein termed BAL3 (Fig. 1) (GenBank™ accession number DQ063586). The BAL3 N terminus is substantially longer than that of the predicted protein encoded by GeneID 165631. However, there are still unidentified coding sequences at the 5’ end of BAL3, whereas cloning of 3′-BAL3 is complete. Using electronic resources (UniGene EST ProfileViewer and Gene Expression Omnibus, www.ncbi.nlm.nih.gov/geo), the expression
BAL Transcriptional Modulation and PARP Activity

FIGURE 1. Diagrammatic representation of the BAL family of proteins. The macro domains and putative PARP domains of BAL1, -2, and -3 are highlighted. The numbering corresponds to the amino acid positions at the beginning and end of these specific domains. The first and last amino acid positions of each BAL family member are indicated. For BAL2 and BAL3, the first methionine in the current sequence is designated as amino acid 1 (both sequences may have longer N termini, see “Results” for details). The BAL2A isoform results from premature termination; BAL2B encodes an additional 154 amino acids including a putative PARP domain.

pattern of BAL3 in human tissues was found to be similar to that of BAL1 and BAL2 (data not shown).

BAL1 was previously mapped to chromosome 3q21 (1). By searching the human genome data base with the BAL2 and -3 cDNA sequences, we found that these genes also mapped to chromosome band 3q21. In fact, the three genes are localized in tandem within ~200 kilobases (contig NT 005612), reinforcing the notion that they are functionally (and evolutionarily) related. These in-silica findings were confirmed by PCR amplification of BAC/YAC clones assigned to this chromosomal region (data not shown).

BAL2 and BAL3, but Not BAL1, Exhibit PARP Activity—All three BAL family members have C termini that include sequences with partial identity to the PARP-1 catalytic domain (Fig. 1 and 2A). The putative PARP active site (the minimal PARP-1 fragment retaining catalytic activity) (9) of BAL2B and BAL3 is ~20% identical to that of the prototype enzyme, PARP-1, and ~26% identical to that of another PARP family member, Tankyrase 1 (14) (Fig. 2B). Of note, BAL2B and BAL3 are ~72% identical at their PARP active site; these sequences also include the requisite catalytic glutamic acid residue (aa 1631 in BAL2B and 649 in BAL3) and additional highly conserved donor (NAD+) and acceptor (polymer binding) sites (Fig. 2A) (10, 21–23). In contrast, the BAL1 C terminus is only ~14% identical with the PARP-1 catalytic domain; in addition, this BAL1 sequence lacks the required catalytic glutamic acid (position 988 in PARP-1) and additional donor and acceptor sites conserved in BAL2B and -3 (Figs. 2A). Further analyses of the divergence and evolutionary relationships between these PARP active sites indicate that the BAL1 sequence is more closely related to that of BAL2 and 3 than that of PARP-1 or Tankyrase 1 (Figs. 2B and C).

To determine whether any of the BAL family members exhibit PARP activity, we generated recombinant GST-tagged BAL1, BAL2B, and BAL3 C termini for use in PARP functional assays (8, 13). In these experiments, recombinant PARP-1 was included as a positive control. We first measured the ability of these proteins to add radiolabeled ADP-ribose moieties to protein acceptors using [32P]NAD+ as a substrate. Coomassie Blue staining of the gel before exposure to x-ray film confirmed that comparable amounts of BAL1, 2B, and 3 were used in these assays (Fig. 3A, left panel). The C-terminal domains of BAL2B and BAL3, but not BAL1, were capable of auto(ADP-ribosyl)ation (Fig. 3A, right panel). Inclusion of the PARP-specific inhibitor, 3-ABA, markedly decreased the auto(ADP-ribosyl)ation of BAL2B and -3, confirming the specificity of these reactions (Fig. 3A, right panel). In similar reactions carried out with cold NAD+, auto(ADP-ribosyl)ated PARP-1, BAL2B, and -3, but not BAL1, were also identified with an antiserum directed against poly(ADP-ribose) (anti-PAR) (Fig. 3B). The observed auto(ADP-ribosyl)ation of BAL2B, BAL3, and PARP-1 was inhibited by 3-ABA, again confirming the specificity of the assay (Fig. 3B). Of interest, the PARP activity of BAL3 was DNA-independent; the absence of DNA in the reaction buffer did not change the extent of BAL3 auto(ADP-ribosyl)ation, whereas it markedly decreased PARP-1 activity (Fig. 3C).

The BAL Proteins Contain Multiple N-terminal Macro Domains—BAL1, -2 (isoforms A and B), and -3 include N-terminal sequences with partial homology to previously described macro domains (Figs. 1 and 4). BAL1 and BAL3 sequences include two macro domains, whereas BAL2 A and B contain three similar regions (Figs. 1 and 4A). These seven macro domains are 26–37% identical to the consensus macro sequence (identified by reciprocal PSI-BLAST searches in multiple species and described as the A1pp domain (conserved domains data base, CDD www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=ccd) (Figs. 4, A and B).

However, the percent identity among the BAL-derived macro domains varies from 18 to 61% (Fig. 4B). Analyses of the divergence and evolutionary relationships of BAL macro domains identify specific pairings and four separate subgroups (Figs. 4, B and C). As shown, the first
(most N-terminal) macro domains of BAL1 and BAL2 form one cluster, the second macro domain of BAL2 and the first of BAL3 define another pair, and the third macro domain of BAL2 segregates with the second macro domain of BAL3 (Fig. 4C). Although it is most related to the latter pair, the second macro domain of BAL1 branches independently from the others (Fig. 4C). These patterns suggest that the individual macro domains within a given BAL protein might have subtle functional distinctions that are conserved in all three BAL proteins.

The Macro Domain Superfamily—Analyses of the most recent human genome assemblies revealed the existence of at least five additional human genes encoding macro domain proteins (Figs. 4, A–C). Two of these genes are the well characterized and related histone macroH2A1 and macroH2A2. Histone macroH2A1 (H2AFY), the original member of this family (2), is expressed as two functionally distinct alternatively spliced transcripts, macroH2A1.1 and 1.2 (H2AFY isoform 1 and 2). The more recently described histone macroH2A2 (H2AFY2) (24, 25) encodes a single protein that is more closely related to macroH2A1.2 (Fig. 4C).

Much less is known about the other three human genes encoding single macro domain proteins, GDAP2, LRP16, and c20orf113. The first two, which are conserved from plants to higher mammals, have been implicated in ganglioside-induced neural differentiation (GDAP2) (26) and increased proliferation of MCF-7 cells (LRP16) (27). c20orf113 (GeneID 170486) was cloned as part of a human cDNA sequencing project and does not yet have an ascribed function.

In summary, there are 8 human genes encoding 9 proteins with 13 distinct macro domains. Additional structural and functional features, such as the presence of histone or PARP domains, delineate specific gene families (histone macroH2A and BAL). A third cluster, GDAP2, LRP16, and c20orf113, is defined by the evolutionary relationship of their respective macro domains (Fig. 4C). Of note, BAL family members are the only known genes with multiple macro domains in the same protein (Fig. 1).

The BAL Proteins Mediate Transcription Repression; Role of the Macro Domains—Given the described association of the macro domain of histone macroH2A with transcriptional repression (6), we next assessed the possibility that the prototype BAL family member, BAL1, might have a similar function. Initially, we investigated the ability of BAL1 to repress transcription when tethered (via a Gal4 DNA binding domain) to the thymidine kinase (TK) promoter (Fig. 5A). This experimental approach, which is often used to characterize the repressive activity of chromatin-associated proteins, previously defined the repressive activities of the macro domain of histone macroH2A1 (6).

The Gal4DBD alone markedly increased TK-driven luciferase activity as previously described (28) (Figs. 5, B and C). For this reason, we used a more stringent approach to assess BAL-mediated repression, comparing TK-driven luciferase activity of Gal4DBD-BAL1-transfected cells with that of cells transfected with reporter alone (no effector) (21). As shown, Gal4DBD-BAL1 decreased TK-driven transcription in a dose-dependent manner (Fig. 5B); the magnitude of transcription repression was comparable to that reported for the macro domain of histone macroH2A (6).

To define the role of BAL1 macro domains in the observed transcriptional repression, an additional series of BAL1 deletion constructs was generated; N terminus proximal (aa 2–119) and N terminus complete (aa 2–464, including macro domains aa 119–464) (Fig. 5B). As expected, a BAL1 proximal N terminus construct which lacks the macro domains had no effect on TK-driven luciferase activity (Fig. 5C). In contrast, the macro domain-containing BAL1 N terminus and the full-length BAL1 protein were equally effective in repressing transcription (Fig. 5C).

BAL1-mediated transcription repression was not promoter-dependent as similar results were obtained using the TATA-containing Gal4-MLP-Luc reporter (data not shown). In addition, the Gal4DBD-BAL fusion proteins did not repress the activity of promoter constructs (TK-luc or MLP-luc), which lacked the five Gal4-binding sites (data not...
FIGURE 3. PARP activity of BAL proteins. A, radioactive assays. C-terminal sequences encompassing the putative PARP active site domains of the BAL proteins (BAL1 aa 590–854, BAL2 aa 1455–1638, BAL3 aa 433–656 (Fig. 1)) were fused to GST, and the recombinant proteins were generated using the pGEX expression system. Reactions were performed with \[^{32}P\]NAD~\(^{+}\) in the presence or absence of the PARP inhibitor 3-ABA (designated + versus – in the figure). Samples contained 2 μg of purified GST (alone) or GST-BAL proteins or 0.1 μg of high purity recombinant PARP protein. The products were analyzed by Coomassie Blue staining (left), which confirmed equal loading, and autoradiography (right) of SDS-PAGE gels. In these assays, BAL2B, BAL3, and PARP-1 ADP-ribosylate themselves, whereas no activity was detected for BAL1 or GST alone. These enzymatic activities were clearly inhibited by 3-ABA, confirming their specificity.

B, anti-PAR (poly(ADP-ribose)) immunoblots. Reactions were performed as above but supplemented with 1 mM unlabeled NAD~\(^{+}\) instead of \[^{32}P\]NAD~\(^{+}\). After size fractionation, the proteins were transferred to polyvinylidene difluoride membranes (Millipore) and immunoblotted with a polyclonal antibody to poly(ADP-ribose). As in the radioactive reactions, BAL2B and BAL3 PARP activities were readily detectable and inhibited by 3-ABA. The experiments shown are representative examples of multiple similarly performed assays.

C, DNA-independent BAL3 PARP activity. Increasing amounts of BAL3 recombinant protein or 0.1 μg of PARP-1 protein were analyzed for PARP activity (as above) in the presence or absence of activated DNA. Samples were immunoblotted with an antibody to poly(ADP-ribose). The poly(ADP-ribose)ylation of BAL3 was unchanged by the presence or absence of DNA, whereas the PARP-1 activity was DNA-dependent.
shown). These data confirmed that BAL1-mediated repression was due to physical association with the promoter area rather than an indirect (general) inhibitory effect.

**PARP Activity and Transcription Repression**—The BAL1 protein, which represses transcription via its N-terminal macro domain, lacks C-terminal PARP activity. However, BAL2 and -3 contain macro domains and active...
PARP sites, of interest given the known association between PARP-1 function and transcriptionally active chromatin (9, 11–13). For these reasons, we also investigated the ability of full-length BAL3 and the N-terminal BAL3 macro domain to repress transcription (Fig. 5D). Full-length BAL3 was significantly less effective than full-length BAL1 in repressing transcription (p = 0.004), likely due to the PARP activity of BAL3 (Fig. 5D).
**DISCUSSION**

The BAL gene family is composed of three members mapping in tandem on chromosome band 3q21. BAL proteins have multiple N-terminal macro domains capable of mediating transcription repression; BAL2B and BAL3 also exhibit PARP activity.

The macro domain is an evolutionary conserved sequence of ~135
**BAL Transcriptional Modulation and PARP Activity**

amino acids initially described as the non-histone region of the variant histone macroH2A (2). Some proteins consist almost entirely of the macro domain (notably in lower species); in other instances this sequence is embedded within much larger polypeptides and associated with a variety of unrelated motifs such as histone domain in the macroH2A family, the PARP domain in the BAL family, and the sec14 region (a putative lipid binding domain) in GDAP2. In our genome-wide search, we identified 13 human macro domains encoded by 8 unique genes including the 3 BAL family members. Although BAL macro domains have highly similar sequences, the divergence analysis and the evolutionary tree indicates that the two or three macro domains within a given BAL protein do not cluster with each other; instead, they cluster with corresponding macro domains in other BAL family members. This structural organization likely ensures the representation of specific macro regions in all three BAL proteins, prompting speculation regarding subtle functional differences between these macro sequences. Consistent with this possibility, recent studies suggest that the alternatively spliced macro domains in macroH2A (macroH2A.1 and macro H2A1.2) differ in their ability to bind NAD metabolites (29).

Although recent progress has been made, the role of the macro domains is still incompletely characterized. Histone macroH2A is enriched on the inactive mammalian X chromosome (5), and the macro domain from this variant histone represses transcription in vitro (6). A molecular mechanism was recently proposed for the domain-specific transcriptional repression of histone macro2A (7). The macro domain of histone macroH2A in a positioned nucleosome was found to interfere with transcription factor binding, whereas the histone region disrupted SWI/SNF nucleosome remodeling. This observation was further substantiated by crystallographic studies that identified similarities between macro domains and DNA binding structures (8). Recent studies also suggest that certain macro domains may bind monomeric or polymeric ADP-ribos moieties (29, 30). Although the functional consequences of these interactions remain to be defined, they are of particular interest given the demonstrated PARP activity of BAL2B and -3.

In our studies, the BAL N-terminal macro domain repressed transcription when brought into the close proximity of a promoter. The magnitude of BAL macro domain-mediated repression was similar to that seen with the macro domains of histone macroH2A (6), highlighting the functional relationship between macro domains from distinct proteins. Like histone macroH2A, BAL proteins may sterically block the access of transcription factors and coactivators to specific chromatin regions (8). The macro domains of BAL2B and -3 might also serve as guides to direct BAL2B and -3 PARP activity to specialized compartments in the nucleosome. In our in vitro studies BAL family members with functional PARP domains exhibited significantly less transcription repression, suggesting that BAL activity is determined by family member-specific macro domains and PARP sites. In this regard, it is of interest that enzymatically active or inactive PARP-1 can have opposite effects on transcription (9, 11–13).

It is not yet clear whether BAL proteins modulate transcription independently or in association with a multiprotein co-repressor complex. However, considering the recently identified role of ubiquitination in transcription modulation (31), it is of interest that BAL1 binds to BBAP (B-lymphoma and BAL-associated protein), a DELTEX family member with ubiquitin E3 ligase activity (32).

Based on C-terminal sequence homologies, BAL proteins were predicted to be members of the PARP superfamily and termed PARP-9 (BAL1), PARP-14 (BAL2), and PARP-15 (BAL3) (9). However, several residues critical for PARP catalytic activity, elongation, and branching are missing in BAL1, including glutamic residue 988 of PARP-1 (22, 23, 33, 34). Consistent with these structural differences, BAL1 lacks PARP activity, whereas BAL2B and -3 exhibit poly(ADP-ribose) polymerase activity. Although BAL2B and -3 have highly similar PARP active sites and a higher level of conservation with the prototype PARP-1 catalytic site, less than half of the identified acceptor and donor sites (21) is present in these new BAL family members. These findings suggest that the requirements for a functional PARP domain might be less stringent than initially described.

We specifically found that BAL2B and BAL3 are capable of poly-(ADP-ribosyl)ating themselves in a DNA-independent manner. In PARP-1 and Tankyrase, automodification serves as a regulatory mechanism that limits the activities of these enzymes toward their specific targets (9, 35). It is likely that automodification of the BAL proteins serves a similar objective. Identification of additional BAL binding partners and putative targets for this post-translational modification will help clarify the role of BAL-mediated poly(ADP-ribosyl)ation in cellular systems.

Interestingly, recent studies also implicate PARP activity in DNA-repair independent nuclear processes that regulate chromatin configuration and transcription (11–13). In general, these studies point to an association between enzymatically active PARP-1 and high transcription activity via decondensation of chromatin (13), poly(ADP-ribosyl)ation of co-repressor complexes (11), or automodification (12). Conversely, by binding to the nucleosome (12) or co-repressor complex (11), inactive PARP-1 plays a structural role that adversely influences gene expression. These recent studies and the data presented here suggest a model whereby the transcription repression mediated by N-terminal macro-domains of the BAL proteins might be curtailed by C-terminal PARP activity via automodification and/or poly(ADP-ribosyl)ation of nucleosomal histones. Alternatively, poly(ADP-ribosyl)ation might be the primary functional attribute of the BAL2B and BAL3 proteins, and the macro domains would promote the critical interaction with specific compartments of the chromatin. The noted differences between BAL1 and BAL2/BAL3 will provide the ideal tools to test these models and to define the signals necessary to trigger BAL PARP activity in vivo. Given emerging evidence that PARP proteins represent possible cancer treatment targets (15–17), functional characterization of BAL family members may have therapeutic applications in lymphoma.

Acknowledgment—We thank Susan Smith (New York University, New York) for input regarding PARP activity assays.

REFERENCES

1. Aguiar, R. C., Yakuhasijin, Y., Kharbanda, S., Salgia, R., Fletcher, J. A., and Shipp, M. A. (2000) Blood 96, 4328–4334
2. Peluhn, J. R., and Fried, V. A. (1992) Science 257, 1398–1400
3. Ladurner, A. G. (2003) Mol. Cell 12, 1–3
4. Aravind, L. (2003) Trends Biotech. Sci. 26, 273–275
5. Costanzo, C., and Peprson, J., R. J. (1998) Nature 393, 599–601
6. Perche, P. Y., Vourc’h, C., Konczy, L., Souchier, C., Robert-Nicoud, M., Dimitrov, S., and Khochbin, S. (2000) Curri. Bioi. 16, 1531–1534
7. Angelov, D., Molla, A., Perche, P. Y., Hans, F., Cote, J., Khochbin, S., Bouvet, P., and Dimitrov, S. (2003) Mol. Cell 11, 1033–1041
8. Allen, M. D., Buckle, A. M., Cordell, S., Dow, J., and Bycroft, M. (2003) J. Mol. Biol. 330, 503–511
9. Ame, J. C., Speleman, C., and de Murcia, G. (2004) BioEssays, 26, 882–893
10. Smith, S. (2001) Trends Biochem. Sci. 26, 174–179
11. Ju, B. G., Song, D. E., and Lee, J. K., Rose, D. W., Glass, C. K., and Rosenfeld, M. G. (2004) Cell 119, 815–829
12. Kim, M. Y., Mauro, S., Gevry, N., Lisi, J., and Kraus, W. L. (2004) Cell 119, 803–814
13. Tulin, A., and Spradling, A. (2003) Science 299, 560–562
14. Smith, S., Giriat, I., Schmitt, A., and de Lange, T. (1998) Science 282, 1484–1487
15. Seimiya, H., Muramatsu, Y., Ohishi, T., and Tsuruo, T. (2005) Cancer Cell 7, 25–37
16. Farmer, H., McCabe, N., Lord, C. J., Tutt, A. N., Johnson, D. A., Richardson, T. B., and Smith, S. (2001) \"Trends in Molecular and Cell Biology, 119, 815–829\"
Santarosa, M., Dillon, K. J., Hickson, I., Knights, C., Martin, N. M., Jackson, S. P., Smith, G. C., and Ashworth, A. (2005) Nature 434, 917–921
17. Bryant, H. E., Schultz, N., Thomas, H. D., Parker, K. M., Flower, D., Lopez, E., Kyle, S., Meuth, M., Curtin, N. J., and Helleday, T. (2005) Nature 434, 913–917
18. Aguiar, R. C., Yakushijin, Y., Kharbanda, S., Tiwari, S., Freeman, G. J., and Shipp, M. A. (1999) Blood 94, 2403–2413
19. Aguiar, R. C., Chase, A., Coulthard, S., Macdonald, D. H., Carapeti, M., Reiter, A., Sohal, J., Lennard, A., Goldman, J. M., and Cross, N. C. (1997) Blood 90, 3130–3135
20. Simonin, F., Poch, O., Delarue, M., and de Murcia, G. (1993) J. Biol. Chem. 268, 8529–8535
21. Augustin, A., Spenlehauer, C., Dumond, H., Menissier-De Murcia, J., Piel, M., Schmit, A. C., Apiou, F., Vonesch, J. L., Kock, M., Bornens, M., and De Murcia, G. (2003) J. Cell Sci. 116, 1551–1562
22. Ruf, A., Rolli, V., de Murcia, G., and Schulz, G. E. (1998) J. Biol. Chem. 273, 57–65
23. Ruf, A., de Murcia, G., and Schulz, G. E. (1998) Biochemistry 37, 3893–3900
24. Costanzi, C., and Pehrson, J. R. (2001) J. Biol. Chem. 276, 21776–21784
25. Chadwick, B. P., and Willard, H. F. (2001) Hum. Mol. Genet. 10, 1101–1113
26. Liu, H., Nakagawa, T., Kanematsu, T., Uchida, T., and Tsuji, S. (1999) J. Neurochem. 72, 1781–1790
27. Han, W. D., Mu, Y. M., Lu, X. C., Xu, Z. M., Li, X. J., Yu, L., Song, H. J., Li, M., Lu, J. M., Zhao, Y. L., and Pan, C. Y. (2003) Endocr. Relat. Cancer 10, 217–224
28. Wang, A. H., Bertos, N. R., Vezmar, M., Pelletier, N., Crosato, M., Heng, H. H., Thi ng, J., Han, J., and Yang, X. J. (1999) Mol. Cell. Biol. 19, 7816–7827
29. Kustatscher, G., Hothorn, M., Pugieux, C., Schefler, K., and Ladurner, A. G. (2005) Nat. Struct. Mol. Biol. 12, 624–625
30. Karras, G. I., Kustatscher, G., Buhecha, H. R., Allen, M. D., Pugieux, C., Sait, F., Bycroft, M., and Ladurner, A. G. (2005) EMBO J. 24, 1911–1920
31. Muratani, M., and Tansey, W. P. (2003) Nat. Rev. Mol. Cell Biol. 4, 192–201
32. Takeyama, K., Aguiar, R. C., Gu, L., He, C., Freeman, G. J., Kutok, J. L., Aster, J. C., and Shipp, M. A. (2003) J. Biol. Chem. 278, 21930–21937
33. Rolli, V., O’Farrell, M., Menissier-de Murcia, J., and de Murcia, G. (1997) Biochemistry 36, 12147–12154
34. Ruf, A., Menissier de Murcia, J., de Murcia, G., and Schulz, G. E. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 7481–7485
35. Smith, S., and de Lange, T. (2000) Curr. Biol. 10, 1299–1302