HMGB1 Interacts with Many Apparently Unrelated Proteins by Recognizing Short Amino Acid Sequences*

Received for publication, August 31, 2001, and in revised form, December 07, 2001
Published, JBC Papers in Press, December 17, 2001, DOI 10.1074/jbc.M108417200

Agnès Dintilhac and Jordi Bernues†‡

From the Departament de Biologia Molecular i Cel·lular, Institut de Biologia Molecular de Barcelona, CIB, Consell Superior d’Investigacions Científiques, Jordi Girona, 18-26, 08034 Barcelona, Spain

The chromatin high mobility group protein 1 (HMGB1) is a very abundant and conserved protein that is structured into two HMG box domains plus a highly acidic C-terminal domain. From the ability to bind DNA nonspecifically and to interact with various proteins, several functions in DNA-related processes have been assigned to HMGB1. Nevertheless, its functional role remains the subject of controversy. Using a phage display approach we have shown that HMGB1 can recognize several peptide motifs. A computer search of the protein data bases found peptide homologies with proteins already known to interact with HMGB1, like p53, and have allowed us to identify new potential candidates. Among them, transcriptional activators like the heterogeneous nuclear ribonucleoprotein K (hnRNPK), repressors like methyl-CpG binding protein 2 (MeCP2), and co-repressors like the retinoblastoma susceptibility protein (pRb) and Groucho-related gene proteins 1 (Grg1) and 5 (Grg5) can be found. A detailed analysis of the interaction of Grg1 with HMGB1 confirmed that the binding region contained the sequence homologous to one of the peptides identified. Our results have led us to propose that HMGB1 may play a central role in the stabilization and/or assembly of several multifunctional complexes through protein-protein interactions.

In the eukaryotic cell nucleus of all vertebrate cell types, HMGB1 (formerly named HMG1, see Ref. 1 for a revised nomenclature) is one of the most abundant non-histone proteins. HMGB1 has been shown to be essential because knockout mice die 24 h after birth (2). HMGB1 is highly conserved, particularly in mammals and to a lesser extent throughout the animal kingdom. HMGB1 is structured into three domains, two basic HMG boxes (HMGB domains A and B) and a highly acidic C-terminal domain, which confer an overall dipolar appearance to this protein (see Refs. 3–5 for reviews). Each of the HMG boxes is formed by two short and one long α-helix that upon folding produce an L- or V-shaped three-dimensional domain structure (6–8). Whereas the acidic C-terminal domain is presumably involved in the modulation of HMGB1 activity, the HMGB box domains allow the protein to bind to linear DNA with moderate affinity and to highly structured (3- and 4-way junction DNA, cruciform DNA) or distorted DNA (bent or kink DNA, bulged DNA, cisplatin-modified DNA) with higher affinity, but always without sequence specificity. The concave surface of the L- or V-shaped HMGB box domain contacts the DNA in the minor groove in two slightly different ways introducing important modifications in the structure of DNA, in particular a strong bend (reviewed in Ref. 5). Presumably, these features will be of relevance for the biological functions in which HMGB1 has been involved (DNA repair, recombination, replication, and transcription).

The activity of HMGB1 is not solely mediated by its ability to bind to DNA. Indeed, HMGB1 and the related HMGB2 protein can interact through their HMG box domains with a broad range of proteins ranging from nuclear cell proteins to viral proteins. Interactions of HMGB1 have been described with the recombination activation gene protein RAG1 (9), several transcription factors including the cellular tumor suppressor p53 (10), the octamer transcription factors Oct1, Oct2, Oct4, and Oct6 (11, 12), some homeotic HOX proteins (13), the steroid receptors (progesterone (PR), glucocorticoid (GR), estrogen (ER), and androgen (AR)) (14, 15), the general initiation factor human TATA-binding protein (hTBP) (16–18), and the viral replication proteins Rep78 and Rep68 (19). The consequences of these interactions are multiple. HMGB1 in general increases the DNA binding affinity of those factors and depending on the context and the assay conditions HMGB1 has been shown to have a positive or a negative effect on transcription (14–17, 20). In the case of RAG1 and Rep68/78 HMGB1 enhances the rate of the sequence-specific DNA cleavage reaction (19, 21). Interestingly, HMGB1 can also stimulate the ATPase activity of Rep78 (19).

The two HMG box domains of HMGB1 appear to have a similar but not identical behavior with respect to their protein-interacting features. Thus, HMGB box A is important for binding to hTBP and p53, whereas the binding to Oct factors, HOX factors, and hormone receptors can take place through boxes A or B (11, 13, 17, 22). However, the interaction with RAG requires both HMGB box domains (9).

To date, neither the HMGB box surface that is involved in the interaction with other proteins nor the required amino acids of HMGB1 are known. On the other hand, sequence analysis of the factors interacting with HMGB1 does not suggest any apparent homology or similarity. For instance, the interaction with RAG1, Oct, and HOX factors occurs at the homeodomain. In the case of hTBP, it is the H2’ α-helix of the core and with Rep78, two different regions are recognized. From these data, no consensus can be defined.

* This work was supported by a European Union Grant FMRX-CT97-0109 and Comisión Interdepartamental de Recerca i Innovació Tecnològica (CIRIT) of the Generalitat de Catalunya Grant SGR97-55. This work was carried out in the context of the Centro de Referencia en Biotecnologia of the CIRIT of the Generalitat de Catalunya. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Dept. Biologia Molecular i Cel·lular, Jordi Girona, 18-26, 08034 Barcelona, Spain. Tel.: 34 934 006 177; Fax: 34 932 095 904; E-mail: jbbmc@ibmb.csic.es.

‡ To whom correspondence should be addressed: Dept. Biologia Molecular i Cel·lular, Jordi Girona, 18-26, 08034 Barcelona, Spain. Tel.: 34 934 006 177; Fax: 34 932 095 904; E-mail: jbbmc@ibmb.csic.es.

† The abbreviations used are: HMGB1, high mobility group protein 1; pRb, retinoblastoma protein; hnRNPK, heterogeneous nuclear ribonucleoprotein K; MeCP2, methyl-CpG binding protein 2; Grg1, Groucho-related gene protein 1; Grg5, Groucho-related gene protein 5; GST, glutathione S-transferase.
The fast progress of genomics and proteomics has made it obvious that an important focus in understanding biological processes is to characterize how proteins interact in macromolecular complexes. Attempts to define general rules for predicting specific recognition between proteins have been unsuccessful because each protein-protein interaction has its own properties. Nevertheless, some indications are emerging. The development of powerful tools has led to the discovery that one type of recognition involves asymmetric interactions that occur between a particular domain and a short region, often less than 10 amino acids in length, within the other protein. A recent review (23) recapitulated several examples of protein domains involved in these kinds of interactions like SH3 (Src homology 3), phosphotyrosine-binding WW, EH (Eps15 homology), PDZ modules (PSD-95/dlg/ZO1), as well as pHb and the ER. Some of them, like the ER, could interact in several modes with different peptides, in a manner that depends on the bond ligand.

In the present study we have explored the molecular recognition properties of HMGB1 by ligand selection from a large library of heptapeptides displayed on phages. Our results do not give support to one unique strong consensus sequence but rather to a few different kinds of peptide sequences. A BLAST search enabled us to predict new proteins that may interact with HMGB1. We have tested and confirmed the interactions with a few of these and have shown that HMGB1 can interact not only with transcriptional activators but also with repressors and co-repressors. Taken together the data suggest a complex network of protein-protein interactions that will clarify the biological function(s) of HMGB1.

EXPERIMENTAL PROCEDURES

Expression and Purification of Recombinant HMGB1 Box A, Box B, and Full-length HMGB1—The plasmids pT7-HMGB1bA, pT7-HMGB1bB, and pET4b-HMGB1 used for the expression of rat HMGB1 box A, box B, and full-length HMGB1, respectively, have already been described. The procedure for the expression and purification of the three recombinant proteins was as described previously (17). Native calf thymus HMGB1 was purified as described previously (24).

Peptide Phage Display Analysis with HMGB1 Boxes A and B—A phage display heptapeptide library kit (Novagen, Madison, WI) was used to screen for peptides binding to HMGB1 box A and B. The kit contained a random combinatorial collection of heptapeptides fused via a flexible linker sequence to the N terminus of protein PIII of bacteriophage M13. Each phage expressed at the tip of the cover 3–5 copies of the unique peptide it encoded. The library complexity contained all the possible combinations of the 20 natural amino acid residues, except glycine. We followed the instructions as indicated by the manufacturer. Four independent experiments were run in which 30°/° copies of the unique peptide it encoded. The library of heptapeptides displayed on phages. Our results do not give support to one unique strong consensus sequence but rather to a few different kinds of peptide sequences. A BLAST search enabled us to predict new proteins that may interact with HMGB1. We have tested and confirmed the interactions with a few of these and have shown that HMGB1 can interact not only with transcriptional activators but also with repressors and co-repressors. Taken together the data suggest a complex network of protein-protein interactions that will clarify the biological function(s) of HMGB1.

RESULTS

Preferential Interaction of HMGB1 Box Domains A and B with Several Peptides—In an attempt to identify targets that could be potentially recognized by HMGB1, and because other general approaches like yeast two-hybrid analysis were not possible (likely due to toxic effects of the expression of HMGB1 boxes in yeast, results not shown) a peptide library screening approach was carried out. A highly complex library containing the whole collection of natural heptapeptide sequences displayed on phage M13 was used. Four rounds of selection were carried out for each of the four independent experiments that were performed using highly purified recombinant HMGB1 boxes A or B as bait proteins. During the biopanning process the stringency was increased by using higher detergent concentrations in the washing buffer. Bound phages were recovered by either a nonspecific acid elution (experiment 1) or by affinity competition using HMGB1 boxes A or B free in solution (experiments 2–4). Several peptides were selected as potentially interacting with HMGB1 with some specificity (Table I). From these results, it became clear from their sequences that they...
were not related to a single strong consensus sequence, suggesting that either the interactions between HMGB1 and the peptides were weak or that HMGB1 boxes could interact with several unrelated motifs. We noted that the frequency of appearance of some amino acids in the selected peptides clearly deviated from the random theoretical level indicating that the selection process was successful. That is, if interactions of HMGB1 were specific they should be independent of the particular growth features of the phages and their statistics should clearly differ from those of the unselected phages in the biopanning assays and be due to their ligand-binding requirements. For example, positively and negatively charged amino acids usually presented frequencies lower than expected in the selected peptides, suggesting that the interactions did not mainly rely on electrostatic forces despite the highly basic character of the HMGB1 boxes. Also, a high rate of aromatic residues was observed, in particular a high level of tryptophan in experiment number 4, which contrasts with the observation that this amino acid tends to decrease naturally without selection. A remarkable level of proline was also obtained in experiment 4, which contrasts with the observation in the sequence of HMGB1. Because HMGB1 can only very weakly interact with the HMGB1 molecule (28), this suggested that the peptides selected were representative of reasonable rather than very weak interactions. Despite this variability in the peptide sequences it was also clear that amino acids usually presented frequencies lower than expected in the selected peptides, suggesting that the interactions did not mainly rely on electrostatic forces despite the highly basic character of the HMGB1 boxes. Also, a high rate of aromatic residues was observed, in particular a high level of tryptophan in experiment number 4, which contrasts with the observation that this amino acid tends to decrease naturally without selection. A remarkable level of proline was also obtained in experiment 4, which contrasts with the observation in the sequence of HMGB1. Because HMGB1 can only very weakly interact with the HMGB1 molecule (28), this suggested that the peptides selected were representative of reasonable rather than very weak interactions.

Despite the fact that the sequences selected were very variable (as shown in Table I), the appearance of several copies of the same peptide in each experiment indicated a high enrichment and specificity in the screening. Note also that some peptides (e.g. HWGMWSY, HAIPYRH) were selected in independent experiments with both HMGB1 domains. Despite this variability in the peptide sequences it was also clear that amino acid distribution in the peptides was not random (Table I). Thus, peptides could be grouped into at least two classes: proline-rich and tryptophan-rich. Two peptides enriched in tryptophan, HWGMWSY and HSWLWLP, accounted for 50% of the clones in experiment 4, and HWGMWSY appeared also in experiments 1 and 2. A minimal consensus WXW motif could be a potential site for interaction. In the case of proline-enriched peptides it was more difficult to define a consensus given the diversity of these sequences.

Because of the high complexity in the peptide sequences retrieved in the phage display experiments we were concerned about the potential existence of false positives in the selected set of peptides. As a second approach to confirm the bona fide association of HMGB1 with those peptides, we performed in vitro assays in which peptide sequences were fused to GST. The twelve peptides were used among those selected with HMGB1 boxes A and B and were representative of the different kinds of peptides obtained. They were fused to the C terminus of GST to facilitate their expression and purification in E. coli. As a control a GST with an extended, unrelated, and never selected peptide was included (GST-KG). Because direct GST pull-down assays did not work, likely because of steric hindrance of the GST moiety to the very small peptide (not shown), Far-Western experiments were performed (Fig. 1). In contrast to the negative control, all the other peptides showed interaction with HMGB1 box B (Fig. 1B). The relative intensity of each band varied from experiment to experiment indicating that the results were not quantitative. Nevertheless, these interactions were always detectable, whereas interaction with the control was never detected (Fig. 1B, lane GST-KG). Moreover, the same assays were also done using either HMGB1 box A or HMGB1 full-length and the results were the same (not shown).

We noted that none of the peptides selected was represented in the sequence of HMGB1. Because HMGB1 can only very inefficiently interact with itself forming a homodimer (27) and HMGB1 boxes do not interact with each other inside the HMGB1 molecule (28), this suggested that the peptides selected were representative of reasonable rather than very weak interactions.

**Discovery of New Partners for HMGB1**—Once our results were confirmed it was of interest to look for proteins containing regions of homology to the peptides selected in the phage display assay in order to perform a survey. This point was addressed by searching the available data bases with the BLAST...
We noted a potential WER, a factor previously identified to interact with HMGB1 (15). The program for nuclear proteins. HWGMSWY/HSSLWPP (B). The name of the candidate protein is indicated in the left column, the sequence of homology to the query peptide with indication of the residue numbers for each case is shown in the middle column, and the accession number is quoted on the right column.

| Putative interactors | Motifs of homology | Accession number |
|----------------------|---------------------|-----------------|
| Grg5                 | 140 LPLTLPLP 146    | Q61695          |
| Grg1/TL1             | 141 VPLTPLP 147    | Q62440          |
| TLE2                 | 160 VPLTPLP 166    | NP 062699       |
| p53                  | 76 GPAPAP 82       | P02340          |
| hnRNP K              | 307 LPLPP 313      | Q60577          |
| MeCP2                | 380 MLLRSP 386     | NP 034918       |
| pRB                  | 780 PSHHP 774      | A33718          |
| PI(2)                | 956 PPLSP 962      | Q64701          |
| PI(3)                | 1028 PPLSP 1034    | Q64700          |
| DP2                  | 19 NPYTAP 25       | Q64163          |
| RelB                 | 75 GAPAAP 81       | Q40863          |
| PR                   | 415 FLAAP 421      | Q00175          |
| AR                   | 280 LEPAP 274      | P19061          |
| HAIRLESS             | 179 WPLAPNP 185    | Q61645          |
| CDP                  | 913 OPTTP 919      | P53564          |
| RBP-Jx               | 375 APTTAP 381     | P31286          |
| ERF                  | 243 ELSFP 249      | P70459          |
| IRF2                 | 315 APPTP 321      | P22986          |
| MNT                  | 511 LPLYP 517      | Q60879          |
| MCMT                 | 1327 LPLFP 1333    | P13864          |
| RFX1                 | 80 QATAP 86        | P43877          |
| Sox18                | 288 NLPSP 294      | P43680          |
| TTP                  | 69 APLAP 75        | P22983          |
| Six5                 | 535 LVPAP 541      | P50178          |
| FREAC 3              | 65 GTPCTP 72       | CAA11239        |
| TFE3                 | 312 LPVFPP 318     | Q64092          |
| SREBP-1              | 93 APLSP 99        | Q89713          |
| TGF1                 | 154 ESPSF 160      | AAH05724        |
| CTCF                 | 648 QVPAP 654      | NP 031820       |
| AF-4                 | 1103 SPLSP 1109    | AA882427        |
| HNF-3G               | 26 SVPNVP 32       | CAAM2892        |
| C-EBPβ               | 12 LPLPP 18        | P17876          |
| TAP250               | 1640 GTVP 1646     | P21675          |
| HES1                 | 187 PLLVP 193      | P35428          |
| MATH1                | 29 PLLT 35         | P48985          |
| MATH2                | 258 GPLSP 265      | P48986          |
| Consensus             | XPLTPAP            |                |

| Putative interactors | Motifs of homology | Accession number |
|----------------------|---------------------|-----------------|
| ER α                 | 200 BYGWSWCS 206   | P19785          |
| ER β                 | 160 BYGWSWCS 166   | O08537          |
| TIA1                 | 309 QWGWSQ 315     | P70318          |
| ETS1                 | 79 DWHPWAV 85      | P27577          |
| T2A                  | 123 GWWHMP 129     | Q3651           |
| pRB                  | 67 RGMCC 73        | P13405          |
| TOP1                 | 203 QWGWSWGE 209   | BA060950        |
| ERGβ                 | 139 RGWM 139       | CAA42055        |

Program for nuclear proteins. HWGMSWY was a peptide that generated interesting candidates and among them appeared ER, a factor previously identified to interact with HMGB1 (15). We noted a potential WXXW consensus motif (Table II B) for the factors belonging to this group. These factors were not studied in detail because in comparison another peptide, LPLTPLP, generated the most interesting homologies and allowed the identification of a new set of nuclear proteins that could potentially interact with HMGB1. These putative factors are summarized in Table II A. Once again, two other factors, p53 and PR, already described to interact with HMGB1 appeared (10, 14) and interestingly, the region homologous to our peptide in p53 was highly conserved in mammals. Remarkably, all the proteins listed presented a potential PXVXPX consensus motif, and among them components for many transcription factor complexes can be found. In this list we had only included the proteins containing the sequence motifs homologous to the peptide that were conserved among mammalian species. Note that in some cases two conserved motifs could be found (e.g., Grg1, p53). Also, the two orientations of the motif were considered because similar proline-rich motifs were reported to be recognized by SH3 domains independently of their orientation (Ref. 29 and references therein).

To confirm that HMGB1 can physically interact with the surveyed proteins, we used a pull-down approach employing GST-p53 as a positive control and proteins fused to GST, which were selected as representative of the several classes described above: the RNA-binding protein and positive regulator hnRNP K, the negative regulator MeCP2, and co-repressors pRB, Grg1, and Grg5. Because HMGB1 is a protein that presents many post-translational modifications, including at least acetylation, ADP-ribosylation, and methylation (reviewed in Ref. 30), we have taken this fact into account and have analyzed in parallel the interaction with recombinant and native calf thymus purified HMGB1 (i.e., presenting many post-translational modifications, not shown). Fig. 2 shows that HMGB1 can effectively interact with all the proteins tested with some notable differences between the recombinant and the purified forms. Whereas either recombinant or purified HMGB1 interacted with Grg1 to a similar extent, recombinant HMGB1 showed some preference for Grg5, p53, hnRNP K, and pRB. Purified HMGB1, in contrast, interacted with MeCP2 more efficiently than recombinant HMGB1. Neither recombinant nor purified HMGB1 can interact with GST. We wanted to emphasize, however, that because those results were semiquantitative, only large differences should be taken into consideration. Thus, it was tempting to suggest that whereas post-translational modification did not seem to have a major effect on HMGB1 interaction with some (mainly Grg1, but also Grg5 to some extent), it appeared to negatively affect interaction with others (especially pRB and p53) and in one case was slightly preferred (MeCP2). In this assay the efficiency of the interaction was

**TABLE II**

Some potential candidates for interaction with HMGB1

| Putative interactors | Motifs of homology | Accession number |
|----------------------|---------------------|-----------------|
| ER α                 | 200 BYGWSWCS 206   | P19785          |
| ER β                 | 160 BYGWSWCS 166   | O08537          |
| TIA1                 | 309 QWGWSQ 315     | P70318          |
| ETS1                 | 79 DWHPWAV 85      | P27577          |
| T2A                  | 123 GWWHMP 129     | Q3651           |
| pRB                  | 67 RGMCC 73        | P13405          |
| TOP1                 | 203 QWGWSWGE 209   | BA060950        |
| ERGβ                 | 139 RGWM 139       | CAA42055        |

**FIG. 2.** HMGB1 interacts with several transcriptional regulators. Pull-down assay of recombinant (r) and purified (p, highly modified) HMGB1 with GST fusions of p53, pRB, hnRNP K, MeCP2, Grg1, and Grg5 (see “Experimental Procedures” for details). GST alone was used as negative control. Lanes containing 10% of the input material and corresponding to recombinant (Input r) and purified (Input p) HMGB1 are shown on the left and the right, respectively. In order to analyze the effect of post-translational modifications on the interactions of HMGB1, equal amounts of GST fusions were used with the two HMGB1 preparations. Recombinant HMGB1 migrates slightly above purified HMGB1 because of the addition of a ~1-kDa histidine tag.

Downloaded from http://www.jbc.org/ by guest on July 26, 2018
dependent on the fusion protein. Preliminary data obtained using recombinant HMG boxes A and B showed the same results (not shown).

Mapping of the Interacting Regions on Grg1—We further studied the potential of the PXXPXP motif by using Grg1 as a model target. Grg1 is structured into five domains, each having a particular function (Fig. 3A). A highly conserved N-terminal glutamine-rich domain (Q) is involved in protein oligomerization, a C-terminal WD-repeat domain is used for interaction with other proteins, and a central domain encompassing the CcN motif allows the nuclear localization of Grg1. The two other domains, GP and SP, are poorly conserved and may play a direct role in repression of transcription (31). By using different deletions of Grg1 the involvement of the PXXPXP motif in interaction with HMGB1 was tested (Fig. 3A). As shown in Fig. 3B splitting of the Grg1 molecule into two moieties, N- and C-terminal, showed that HMGB1 was able to interact with both, the one containing domains Q-GP-CcN (lane 4) and the other containing domains SP-WD (lane 8), respectively.

The binding site on the N-terminal moiety was located at the N-terminal region of the GP domain because deletion encompassing residues 131–155 (Fig. 3B, compare lanes 6 and 7, GST-GP-CcN and GST-ΔGP-CcN, respectively) completely abolished the interaction. Additionally, the Q domain clearly showed no interaction with HMGB1 (lane 5). The sequence, which upon deletion abolished binding of HMGB1 on this side of the protein, precisely corresponded to the region of homology to the LPLTPLP peptide (Fig. 3C, shaded). Note that in the closely related Grg5 factor this region is also present and is likely used for interaction with HMGB1 as well (Fig. 2).

The site of interaction in the C-terminal moiety is likely located at the SP domain because deletion of the WD domain did not abolish the interaction. Nevertheless, the existence of still another site on the WD domain cannot be ruled out. A careful examination of the amino acid sequence of the SP domain revealed another region of weaker homology (vPfp-PmP) to the LPLTPLP peptide (Fig. 3D). In this case, this site is absent in the related Grg5 factor.

DISCUSSION

On the HMGB1 Side—HMGB1 was first isolated almost 30 years ago and since then much effort has been put into characterizing this protein and finding its functional role (30). The latter point has been shown to be difficult because of controversial results, and it is still a matter of discussion. Because HMGB1 does not present any enzymatic activity on its own or any sequence-specific DNA or RNA binding, it is clear that the functions in which it could be involved would require the assistance of other factors, likely proteins. Several reports have clearly shown that HMGB1 interacts with a set of different factors that present very little, if any, sequence or structure in common. The question of how HMGB1 can recognize its partners is one we have addressed here.

As yet, the rules that govern the interactions between two given proteins are unknown, and in any case interactions are impossible to predict. A hypothesis that will reconcile all the data on HMGB1 interactions could be that the interaction surface recognized by HMGB1 on these proteins might be very small, so that only a few amino acids would be important. Eventually, not only one of such motifs could be recognized by HMGB1. This idea is based on recent findings showing that many different proteins can recognize motifs as short as 4–5 residues long such as for example, LXCXE for pRb (32) or WRPyW for Groucho in Drosophila (33).

As an approach that occasionally helped in determining the residues that might be required for interactions to take place, the screening of peptide repertoires has been successfully used by several groups to shed light on the mechanisms of ligand recognition of macromolecular complexes. The use of a peptide phage display approach allowed us to begin understanding some previously reported interactions, to predict others, and to begin testing them. In our hands, HMGB1 can recognize sev-
eral short motifs through the HMG boxes. However, the lack of a unique sequence motif or a single consensus sequence derived from them complicated the interpretation of the results and clearly suggested that HMGB1 is a protein that does not have a highly preferred interaction sequence but can establish interactions with many peptides having rather different sequences. Therefore, no strong enrichment is observed, and this heterogeneity makes the interpretation of the results difficult. As far as we know, this seems to be a particular behavior of HMGB1. Some of the peptides were isolated with both HMG boxes A and B (HWGMSWY and HAIYPRH) suggesting that either recognition with these peptides is due to the three-dimensional structure of the HMG box or it uses some of the most highly conserved regions between the two domains. The other peptides seem to be different for each HMG box. However, the complexity of the patterns obtained suggests that the relative affinities for the peptides are similar and weak. This may in turn enrich a particular experiment in a certain set of peptides because of slight changes (in temperature for instance). But it could also be attributable to the differences in sequence of the two HMG boxes because their three-dimensional structures are rather similar (6, 8). On the other hand, taking into account sequence identity in HMG boxes A and B is very high when compared with the corresponding domains in HMGB2, it is also very likely that many if not all of the peptide motifs will also be recognized by HMGB2, as well. In fact, the apparent redundancy of HMGB1 and HMGB2 is a general observation for all of the interactions described so far for these proteins, including Oct factors (11, 12), steroid receptors (15), hTBP, and RAG1 (9) among others.

Our results do not discount the highly acidic C-terminal domain as potentially interacting with other factors. In fact, the C-terminal domain is the one involved in the interaction with histone H1 (34) and with the histone dimer H2A-H2B as well (24), and recently it has been claimed to interact with the glutamine-rich N-terminal domain of human and Drosophila TBP (35). However, the highly acidic nature of this domain makes its interactions highly electrostatic and in general weakly specific.

New HMGB1-interacting Proteins Come to Light—The use of peptides in our assay is likely limiting the set of potential partners for interaction with HMGB1 to those that are recognized only by sequence, because little structure can be expected from heptapeptides. Therefore, it is reasonable to expect some of the interactions previously described not to appear in this assay. This may be the case, for instance, of the interaction to the H2’ α-helix of hTBP previously described (17). Moreover, the fact that a particular protein contains one of the motifs in its sequence does not necessarily imply that this will be the site used for interaction or that another motif can be used. In addition, the contribution of peptidic sequences around the motifs identified here can be determinant, because in some cases they have been shown to play a major role either in modulating or even in affecting the specificity of the interactions. Additionally, the accessibility of the motifs embedded in some protein contexts might be rather limited. Finally, there is always a potential for HMGB1 to be actively recognized by the other partner (see below), and then motifs can be completely different.

When searching protein data bases, a point that must be taken into account is that the relative representation of the different peptides and motifs is not homogeneous. Thus, some peptide motifs produce long lists of proteins, like for PXXP1, whereas other motifs do not. Nevertheless, some motifs not studied here in detail also generated very interesting candidates. For instance, a sequence closely related to the potential WXXW motif, and in particular to peptide HWGMSW, can be found in conserved sequences of the α- and β-ER of mouse, rat, and human, of ETS-1, and also in a conserved sequence of the mouse and human RNA-binding protein nucleolysin TIAR among others (Table IIB). Remarkably, the interaction with α-ER was previously described. Moreover, the proposed HMGB1-binding sequence lies very close to the zinc finger of the ER that interacts with DNA and is clearly accessible in the co-crystal structure (15, 36). These features might help to explain how HMGB1 can stabilize the interaction of ER with DNA.

The PXXP1 consensus motif has predicted the most interesting partners upon BLAST search and for that reason has been studied in detail. These include pRb, MeCP2, hnRNP K, Grp1, Grp5, and p53. Remarkably, p53, which was previously shown to interact with HMGB1, appeared, giving credit to this linear sequence as a potential HMGB1 recognition motif. In fact, all the PXXP1-containing proteins tested here interacted with HMGB1. The reason for this clear result is likely due to the high proline content of the motif, which makes the adoption of a defined secondary structure difficult even if embedded in a protein sequence. Thus, it is very likely that other proteins containing this motif and listed in Table IIA will interact with HMGB1 as well.

HMGB1 has already been shown to interact with p53 and be a unique activator of this factor (10). The PXXP1 motif is always found at the N terminus of p53 around position 80–90 in different species and forming part of a proline-rich domain that negatively affects p53 interaction with DNA (37). The interaction of HMGB1 at this domain could explain the stimulatory effects observed on DNA binding and transactivation (10). Because the proline-rich domain is dispensable for transactivation (38) the role of HMGB1 might be restricted to stabilize the p53-DNA complex. We must mention, however, that residues 363–376 in the basic domain of p53 have been recently reported to recognize HMG box A on HMGB1 and that p53 conversely enhanced HMGB1 interaction with cisplatin-modified DNA (22). In this region of p53 no homology to any of the peptides identified in this work was found. This might be in apparent contradiction to the data discussed above but could be reconciled if there is mutual recognition between p53 and HMGB1. If so, HMGB box A would be recognized by the C-terminal region of p53, and the proline-rich domain of p53 would be recognized by the HMG boxes of HMGB1. Depending on the experimental conditions one or the other interaction would prevail and be selected.

This may also be the case with pRb because in addition to having a sequence showing homology to the PXXP1 motif, the pocket region of pRb can recognize the LXXCXXE motif in other proteins (32). This motif (as the sequence LFCE) is present in the HMG box B of both HMGB1 and HMGB2 and is absolutely conserved in vertebrates. We have not analyzed here in detail whether recognition is via the LXXCXXE motif (i.e., pRb recognizes HMGB1) or the PXXP1 motif in pRb (i.e., HMGB1 recognizes pRb) and simply described this interaction. Although it is possible that the interaction may work in both directions we reasoned that it may be difficult since the LFCE sequence is structured into an α-helix facing the concave side of the L-shaped domain and is partially buried in HMG box B (6).

The interaction between HMGB1 and Grp1 has been studied in more detail because of the high homology to the peptide and also because this is the first interaction of HMGB1 with a co-repressor. In Grp1 two regions of interaction with HMGB1 have been uncovered. The one corresponding to the GP domain

---

2 J. Bernues, unpublished results.
fits nicely with high homology to the LPLTPLP peptide isolated in the phage display analysis and very likely accounts for that interaction as predicted. In the second domain interacting with HMGB1, the SP domain, also a region with the PXXPXP motif is present albeit homology to the original peptide sequence is weaker. These results support the PXXPXP motif as a good recognition sequence for HMGB1. Brantjes et al. (51) have recently shown that all Tcf HMGB box transcription factors interact with Groucho-related co-repressors and even more, all the long members containing five domains (here represented by Grg1) mediated repression of the Tefs, whereas the short member (Grg5) mediated de-repression. All Tefs are HMGB box transcription factors but so far the region that interacts with the Grgs has never included the HMGB box domain. In contrast HMGB1 uses the HMGB box domains to interact with Grg proteins, suggesting that recognition of the PXXPXP motif is not a feature common to all HMGB box domains. On the other hand, transcription mediated by AR is inhibited by Grg5 (39). Because the PXXPXP motif is also present in the AR sequence (Table IIA), and the interaction of HMGB1 with AR stabilizes the AR-DNA complex as shown by others, the interaction of HMGB1 with Grg5 could cooperate by displacing it from AR and helping to relieve inhibition (15).

MeCP2 is a protein involved in the recognition of methylated DNA at CpG islands and is closely related to gene silencing (see Ref. 40 for a review). The HMGB1 binding motif is located at the C terminus of this protein (residues 380–386) in a region where no other factor has been shown to interact. Therefore, it is possible that HMGB1 interaction does not interfere with the binding of the many factors known to interact with MeCP2, (among others: mSin3A, NcoR, and c-Ski), and it remains to be analyzed whether in this case interaction could also stabilize MeCP2-DNA complexes similar to HMGB1 in other complexes.

hnRNP K presents a remarkable variety of protein interactions with some factors involved in signal transduction and others involved in several aspects of gene expression (reviewed in Refs. 41 and 42). This diversity suggests that the hnRNP K protein may act as a docking platform or as a scaffold protein within multiple functional modules. Then, HMGB1-hnRNP K interaction may connect two systems, which have a large protein-protein interaction potential, in order to extend their respective domains of action even further. For example, hnRNP K was shown to be a transcription-activating factor for the c-myc promoter (43, 44). Although we have not been able to establish a direct connection, it is remarkable that c-myc mRNA levels in HMGB1+/− mouse cells are about 5–10-fold lower than those in wild type cells.

We have focused so far on the nuclear environment where HMGB1 is expected to play a role in gene expression. However, HMGB1 is an extraordinary protein and besides being a nuclear protein in most cases it is also true that for more than 10 years HMGB1, under the name of amphoterin, was found on the outer cell membrane of some cell types (45). It is now clear that HMGB1 can be actively released outside the cell in response to tumor necrosis factor and interleukin 1 (46) and passively by necrosis or cell damage in a variety of cell types, mainly immature and transformed cells (47). Release of HMGB1 induces some pathological processes as a potent late mediator of endotoxin lethality and inflammation. Many of these phenomena require the interaction of HMGB1 with the receptor for advanced glycation products (RAGE) (reviewed in Ref. 48). Examination of the mouse RAGE amino acid sequence shows a region highly homologous to the LPLTPLP motif in the second Ig-like C2-type domain that is highly conserved among several species. Similarly, HMGB1 was also shown to interact with Syndecan-1, a cell surface heparan sulfate-rich proteoglycan (49) in which sequence a PXXPXP motif can also be found. These and other data suggest that HMGB1 might use the same motifs to interact with cell surface proteins and with nuclear factors.

Final Considerations—Prior to this work, a long list of proteins were reported to interact with HMGB1. Along with the new candidates reported here, this may give the impression that HMGB1 interacts with almost every protein in the cell. Despite the appearance that HMGB1 is in fact a “sticky” protein (48), it is by no means true. For example, we have been trying to extend the initial interaction of HMGB1 with human TBP to other factors of the general transcription machinery, and we failed to observe any interaction of HMGB1 with TFIIF, TFIIB, TFIIF, the CTD of RNA polymerase II, TFIIE (despite a clear interaction that can be observed with the p56 subunit, no interaction can be observed with the native tetramer), and a TBP-related factor among others,4 suggesting that the list of factors interacting with HMGB1 may be long but by no means indiscriminate. This particular behavior makes it difficult to attribute a defined role for HMGB1 in the organism but may explain the general weakness observed in the HMGB1 knock-out mice that lead them to death 24 h after birth (2). HMGB1 seems to have developed a high potential for protein-protein interaction with multiple partners, always taking part in a macromolecular complex and either assisting in the assembly or stabilizing the assembled factors. This feature along with the remarkable abundance of HMGB1 in the nucleus and the many post-translational modifications that it undergoes could explain its general involvement in nuclear processes and its modulation.

On the one hand, HMGB1 is a protein that can interact with angled DNA on its own. On the other hand, a general observation is that upon interaction HMGB1 can help stabilize many proteins previously bound to DNA. These functions may be compatible; however, although the concave region of the L-shaped HMGB box domain clearly is the DNA interaction site, there is no data about the region of the HMGB box domain involved in the interaction with other protein factors. Some evidence at the enhanceosome of BHIF-1 suggests that HMGB1 can interact with both the narrow groove of DNA and with protein factors (50). However, more detailed work is required in order to discover whether HMGB1 simultaneously interacts with proteins and DNA, the stoichiometry of the interactions, and how its binding can stabilize protein-DNA complexes. A potential role as a co-factor is emerging for HMGB1 in many different processes.

Acknowledgments—We thank Dr. C. Lobe (Molecular Biology Institute, McMaster University, Ontario) for generously providing cDNAs for Grg1 and Grg5, Dr. A. Bird (Institute of Cell and Molecular Biology, University of Edinburgh) for the pGEX/Mcp2 plasmids, Drs. S. K. Jang and J. H. Kim (Department of Life Science, Pohang University of Science and Technology, Korea) for the pGEX-KG/hnRNP K plasmid, and Dr. M. A. Martinez-Balbás (Institut de Biologia Molecular de Barcelona, CSIC) for the pGEX53 and pGEX/pRb plasmids. We also thank J. Font for creating Fig. 3.

REFERENCES

1. Bustin, M. (2001) Trends Biochem. Sci. 26, 152–153
2. Calogero, S., Grassi, F., Aguzzi, A., Voitla, T., Ferrier, P., Ferrari, S., and Bianchi, M. E. (1999) Nat. Genet. 22, 276–280
3. Bustin, M., and Reeves, R. (1996) Prog. Nucleic Acids Res. Mol. Biol. 54, 35–100
4. Bustin, M. (1999) Mol. Cell. Biol. 19, 5237–5246
5. Thomas, J. O., and Travers, A. A. (2001) Trends Biochem. Sci. 26, 167–174
6. Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C., and Norman, D. G. (1993) Nucleic Acids Res. 21, 3427–3436

3 A. Dintilhac and J. Bernues, unpublished results.
4 M. Sutrias-Gran, J. Font, and J. Bernues, unpublished results.
New Factors Interacting with HMGB1

7. Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A. R., Laue, E. D., and Thomas, J. O. (1993) *EMBO J.* 12, 1311–1319
8. Hardman, C. H., Broadhurst, R. W., Raine, A. R. C., Grasser, K. D., Thomas, J. O., and Laue, E. D. (1995) *Biochemistry* 34, 16596–16607
9. Aidinis, V., Bonaldi, T., Beltram, M., Santagata, S., Bianchi, M. E., and Spansopoulou, E. (1999) *Mol. Cell. Biol.* 19, 6532–6541
10. Jayaraman, L., Moorthy, N. C., Murthy, K. G. K., Manley, J. L., Bustin, M., and Prives, C. (1998) *Genes Dev.* 12, 462–472
11. Zwillling, S., Koning, H., and Wirth, T. (1995) *EMBO J.* 14, 1198–1208
12. Butteroni, C., De Felici, M., Scholer, H. R., and Pesce, M. (2000) *J. Mol. Biol.* 304, 529–540
13. Zappavigna, V., Faciola, L., Citterich, M. H., Mavilio, F., and Bianchi, M. E. (1996) *EMBO J.* 15, 4981–4991
14. Ohate, S. A., Prendergast, P., Wagner, J. P., Nissen, M., Reeves, R., Pettijohn, D. E., and Edwards, D. P. (1994) *Mol. Cell. Biol.* 14, 3376–3391
15. Boonyaratanakornkit, V., Melvin, V., Prendergast, P., Altman, M., Ronfani, L., Bianchi, M. E., Taraseviciene, L., Nordeen, S. K., Allegretto, E. A., and Edwards, D. P. (1998) *Mol. Cell. Biol.* 18, 4471–4487
16. Ge, H., and Roeder, R. G. (1994) *J. Biol. Chem.* 269, 17136–17140
17. Sutrias-Grau, M., Bianchi, M. E., and Bernaes, J. (1999) *J. Biol. Chem.* 274, 1628–1634
18. Lu, W., Peterson, R., and Dasgupta, A., Scovell, W. M. (2000) *J. Biol. Chem.* 275, 35006–35012
19. Costello, E., Saudan, P., Winocour, E., Pizer, L., and Beard, P. (1997) *EMBO J.* 16, 5943–5954
20. Shykoff, B. M., Kim, J., and Sharp, P. A. (1995) *Genes Dev.* 9, 1354–1365
21. Van Gent, D. C., Hiom, K., Pauli, T. T., and Gellert, M. (1997) *EMBO J.* 16, 2655–2670
22. Imamura, T., Inumi, H., Nagatani, G., Ise, T., Nomoto, M., Iwamoto, Y., and Kahan, E. (1998) *J. Biol. Chem.* 273, 7534–7540
23. Kay, B. K., Kasanov, J., Knight, S., and Kurakin, A. (2000) *FEBS Lett.* 480, 55–62
24. Bernues, J., Espel, E., and Querol, E. (1986) *Biochim. Biophys. Acta* 866, 242–251
25. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* 25, 3383–3402
26. Carpent, F. (1998) *Nucleic Acids Res.* 16, 10881–10890
27. Ranatunga, W., Lebowitz, J., Ake, B., Pavlik, P., Kar, S. R., and Scovell, W. M. (1999) *Biochim. Biophys. Acta* 1432, 1–12
28. Ramstein, J., Locker, D., Bianchi, M. E., and Leng, M. (1999) *Eur. J. Biochem.* 260, 692–700
29. Kay, B. K., Williamson, M. P., and Sudol, M. (2000) *FASEB J.* 14, 231–241
30. Johns, E. W. (1982) *The HMG Chromosomal Proteins,* Academic Press, London
31. Chen, G., and Courey, A. J. (2000) *Gene (Amst.)* 249, 1–16
32. Dahiya, A., Gavvin, M. H., Luo, R. X., and Dean, D. C. (2000) *Mol. Cell. Biol.* 20, 6799–6805
33. Fisher, A. L., and Caudy, M. (1998) *Genes Dev.* 12, 1931–1940
34. Carballo, M., Puigdomenech, P., and Palau, J. (1983) *EMBO J.* 2, 1759–1764
35. Das, D. M., and Scovell, W. M. (2001) *J. Biol. Chem.* 276, 32489–32494
36. Schwabe, J. W. R., Chapman, L., Finch, J. T., and Rhodes, D. (1993) *Cell* 75, 567–578
37. Muller-Tiemann, B. F., Halazonetis, T. D., and Elting, J. J. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 6079–6084
38. Walker, K. K., and Levine, A. J. (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 15335–15340
39. Yu, X., Li, P., Roeder, R. G., and Wang, Z. (2001) *Mol. Cell. Biol.* 21, 4614–4625
40. Bird, A., and Wolff, A. P. (1999) *Cell* 99, 451–454
41. Sioni, H., and Dreyfuss, G. (1997) *Carr. Opin. Genet. Dev.* 7, 345–353
42. Shibuya, A.-B., and Wilkinson, M. F. (2000) *Cell* 102, 135–138
43. Tomonaga, T., and Levens, D. (1995) *J. Biol. Chem.* 270, 4875–4881
44. Takimoto, M., Tomonaga, T., Matunis, M., Avigan, M., Krutzsch, H., Dreyfuss, G., and Levens, D. (1993) *J. Biol. Chem.* 268, 18249–18258
45. Parkin, K., Hau, E., Merenmis, J., Nolo, R., Olavi Kajander, E., Rauan, R. (1993) *J. Biol. Chem.* 268, 19726–19738
46. Wang, H., Bloom, O., Zhang, M., Vishnubhat, J. M., Ombrellino, M., Che, J., Frazier, A., Yang, H., Ivanova, S., Berwilkova, L., Manogue, K. R., E., F., Abraham, E., Anderson, J., Anderson, U., Molina, P. E., Abumrad, N. N., Sama, A., and Tracey, K. J. (1999) *Science* 285, 248–251
47. Degryse, B., Bonaldi, T., Scafidi, P., Muller, S., Resnati, M., Sanvito, F., Arrigoni, G., and Bianchi, M. E. (2001) *J. Cell. Biol.* 152, 1197–1206
48. Muller, S., Scafidi, P., Degryse, B., Bonaldi, T., Ronfani, L., Agresti, A., Beltrame, M., and Bianchi, M. E. (2001) *EMBO J.* 20, 4337–4340
49. Salvemirita, M., Rauvala, H., Helenius, K., and Jalkanen, M. (1992) *Exp. Cell Res.* 200, 444–451
50. Elwood, K. B., Yen, Y.-M., Johnson, R. C., and Carey, M. (2000) *Mol. Cell. Biol.* 20, 4337–4340
51. Brantjes, H., Roose, J., van de Wetering, M., and Clevers, H. (2001) *Nucleic Acids Res.* 29, 1410–1419
HMGB1 Interacts with Many Apparently Unrelated Proteins by Recognizing Short Amino Acid Sequences
Agnès Dintilhac and Jordi Bernués

J. Biol. Chem. 2002, 277:7021-7028. doi: 10.1074/jbc.M108417200 originally published online December 17, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M108417200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 50 references, 25 of which can be accessed free at http://www.jbc.org/content/277/9/7021.full.html#ref-list-1