An Extended RNA Binding Site for the Yeast Branch Point-binding Protein and the Role of Its Zinc Knuckle Domains in RNA Binding*

Received for publication, April 3, 2006, and in revised form, July 10, 2006 Published, JBC Papers in Press, July 21, 2006, DOI 10.1074/jbc.M603137200

Stephen M. Garrey1, Rodger Voelker2, and J. Andrew Berglund3
From the Department of Chemistry and Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403-1229

The highly conserved branch point sequence (BPS) of UAC-UAAC in Saccharomyces cerevisiae is initially recognized by the branch point-binding protein (BBP). Using systematic evolution of ligands by exponential enrichment we have determined that yeast BBP binds the branch point sequence UACUAAC with highest affinity and prefers an additional adenosine downstream of the BPS. Furthermore, we also found that a stem-loop upstream of the BPS enhances binding both to an artificially designed RNA (30-fold effect) and to an RNA from a yeast intron (3-fold effect). The zinc knuckles of BBP are partially responsible for the enhanced binding to the stem-loop but do not appear to have a significant role in the binding of BBP to single-strand RNA substrates. C-terminal deletions of BBP reveal that the linker regions between the two zinc knuckles and between the N-terminal RNA binding domains (KH and QUA2 domains) and the first zinc knuckle are important for binding to RNA. The lack of involvement of the second highly conserved zinc knuckle in RNA binding suggests that this zinc knuckle plays a different role in RNA processing than enhancing the binding of BBP to the BPS.

The branch point-binding protein (BBP)4 in Saccharomyces cerevisiae is an RNA-binding protein that is thought to be the first protein to bind the branch point sequence (BPS) during splicing. In yeast BBP binds to the highly invariant branch point sequence UACUAAC (branch point A is underlined), recognizing every nucleotide (1). BBP binds the BPS in the second commitment complex (CC2) and is necessary for stable formation of this complex (2). Splicing factor 1 (SF1), the human orthologue of BBP, is a more promiscuous RNA-binding protein, recognizing a degenerate sequence, CURAY (R = purine, Y = pyrimidine), in which only the U and A are critical (1). When discussing the yeast branch point-binding protein we will use the term BBP, when discussing the human protein we will use the term SF1, and when discussing both proteins we will use the term BBP/SF1. In the mammalian system, the equivalent complex to the yeast CC2 is the E complex in which SF1 binds the BPS (3, 4). In addition to recognizing the BPS, BBP/SF1 plays an important role in bridging the 5′ end and 3′ end of the intron by interacting with U1 snRNP at the 5′ splice site and Mud2p at the 3′ end of yeast introns and U2AF65 at the 3′ end of human introns (5–8). BBP/SF1 is replaced at the BPS by U2 snRNP in an ATP-dependent step that leads to the formation of A complex (9, 10).

BBP/SF1 contains multiple domains, many of which are conserved between yeast and human (Fig. 1). The BBP/SF1 N-terminal domain interacts with Mud2p in yeast and U2AF65 in humans (5, 6, 8). The C terminus contains a proline-rich domain, which in yeast interacts with Prp40, a component of the U1 snRNP (7). FBP11, a potential vertebrate homologue of Prp40, binds to the proline-rich domain of human SF1 (11).

One of the most highly conserved regions of BBP/SF1 encompasses the RNA binding domains (Fig. 1). The region of BBP/SF1 responsible for binding the BPS contains a heterogeneous nuclear ribonucleoprotein K homology (KH) domain, a QUA2 domain, and one or two zinc knuckles (1, 6, 12). Most KH proteins are RNA-binding proteins that have multiple KH domains arranged in multiple repeats (13–16). However, SF1/BBP is a member of a subset of KH proteins that contain a single extended KH domain, often referred to as a “maxi-KH” domain. This subclass of KH-containing proteins is also called the STAR/GSG family (15, 17). STAR refers to signal transduction and activation of RNA because members of this family are linked to the signal transduction pathways, whereas GSG refers to the original members of this family, GRP33, Sam68, and GLD-1. These STAR/GSG proteins contain two conserved regions called QUA1 and QUA2 that flank the maxi-KH domain. The QUA1 region of the STAR/GSG proteins is necessary for homodimerization (18–21). This dimerization greatly increases the RNA binding affinity of these proteins (20). Unlike the other members of this group, BBP/SF1 lacks the QUA1 domain and shows no ability to dimerize but still binds RNA (1, 6, 12). The QUA2 region is important for the stability of the maxi-KH domain (12). Its contact with the RNA increases the binding site size for these proteins, specifically recognizing the first adenosine in the BPS in the case of SF1, and...
has been shown to provide important contacts with the RNA (22).

Orthologs of BBP/SF1 are found in nearly every eukaryotic organism, and nearly all of them contain one or two zinc knuckles of the form $\text{CX}_2\text{CX}_2\text{H}_2\text{C}_2$, situated C-terminal to the KH/QUA2 domains (Fig. 1). The zinc knuckles of BBP have previously been implicated in general RNA binding affinity based on truncation studies. Specific RNA binding of BBP/SF1 to the BPS was not changed by replacing the zinc knuckles with an RNA-binding peptide derived from the HIV-1 nucleocapsid protein or by replacement of a short arginine-serine (RS)_n peptide (12). These results suggest that the BBP zinc knuckles are important for nonspecific binding to the phosphate backbone, whereas the KH and QUA2 domains are responsible for sequence specific binding to the BPS. Zinc knuckles have been shown to be directly involved in RNA binding in other proteins. The best studied example is the HIV-1 nucleocapsid protein, which contains two zinc knuckles and binds single-stranded and stem-loop RNAs that are important for HIV genome recognition and encapsidation (23–30).

Previous studies have shown that the KH and QUA2 domains of BBP specifically recognize the conserved yeast BPS of UACUAAC (1, 12). However, it has not been determined if the BPS is indeed the high affinity site for the KH/QUA2 domains and if BBP binds additional RNA elements through the zinc knuckles. To determine whether the BPS is the high affinity RNA site for BBP and to identify additional nucleotides or motifs outside of the BPS, we performed an in vitro selection experiment, also known as systematic evolution of ligands by exponential enrichment (SELEX), with a recombinant construct of S. cerevisiae BBP. To better define how the zinc knuckles of BBP bind RNA, we performed mutation and truncation studies on the zinc knuckle region of BBP. The results of this analysis are presented here.

**EXPERIMENTAL PROCEDURES**

**Plasmid Construction**—Our previously described BBP (193) construct (12), here known as BBP-331, was used as the template for subsequent BBP constructs. This construct contains amino acids 147–331, not 145–330 as previously stated in Bergrund et al. (1). The BBP-269 construct (amino acids 147–269) was previously described (yBBP125 (12)).

The same forward primer was used to amplify the PCR fragments for constructs BBP-296, BBP-287, and BBP-279, which contained the sequence $5'-\text{GGGACCCTGGGATCCAAATTTTACGACAAATGATTATAATCCCG-3'}$. The reverse primers to amplify PCR fragments for constructs BBP-296, BBP-287, and BBP-279 were $5'-\text{CCCGGAATTCGCTGTCGATGTGTAATAGCTTCTTTAACC-3'}$, $5'-\text{CCGGAATTCGCTGTCGATGTGTAATAGCTTCTTTAACC-3'}$, and $5'-\text{CCGGGAATTCGCTGTCGATGTGTAATAGCTTCTTTAACC-3'}$, respectively. All BBB PCR fragments were cut with BamHI at the N terminus and EcoRI at the C terminus and cloned into pGEX-6P-1 (GE Healthcare).

Both the first and the second zinc knuckles were mutated using the QuikChange PCR-based site-directed mutagenesis kit (Stratagene). The forward and reverse primers to mutate the first zinc knuckle (mutations C273A and C276A) were $5'-\text{GAAGATATGGGATCCAAATTTTACGACAAATGATTATAATCCCG-3'}$ and $5'-\text{GCCTTTATGATCTCTTTTTAACCAGCGATGCAGGCTTCTTTAACC-3'}$ and $5'-\text{CCTTTATGATCTCTTTTTAACCAGCGATGCAGGCTTCTTTAACC-3'}$.

The human SF1 (amino acids 134–297) fragment was PCR-amplified using the pGEX6p-mBBP/SF1 (1–361) as template (5). Primers used were $5'-\text{CTGGGATCCACACTAGTGAGTGAAGATATGGGATCCAAATTTTACGACAAATGATTATAATCCCG-3'}$ and $5'-\text{GGGGAATTCGCTGTCGATGTGTAATAGCTTCTTTAACC-3'}$. PCR fragments were flanked with BamHI and EcoRI cut sites and inserted into pGEX-6P-1.

**Protein Purification and Expression**—All BBP and SF1 constructs were transformed into BL21(Star) cells (Novagen). Cells were resuspended in 25 mM Tris (pH 7.5), 300 mM NaCl, 1 mM EDTA, 1 mM DTT, and 0.1% Triton X-100 (Sigma). Cells were sonicated and spun to pellet insoluble debris at 19,800 $\times g$ for 30 min.

The supernatant was added to glutathione-agarose beads (Sigma #G4510) for 15 min and washed 3 times with resuspension buffer (above). Beads were washed once with 1× cleavage buffer (50 mM Tris (pH 7.0), 150 mM NaCl, 1 mM EDTA, 1 mM DTT, and 0.1% Triton X-100). The protein of interest was cleaved from the GST tag using Precision protease (GE Healthcare). For the GST-BBP-331 construct the Precision protease cleavage was skipped and instead eluted with reduced glutathione (Sigma-Aldrich). BBP and SF1 constructs were purified over an ion exchange column (Mono S HR 10/10), concentrated, and dialyzed into protein storage buffer (25 mM Tris (pH 7.5), 150 mM NaCl, 1 mM DTT, and 15% glycerol).

**SELEX**—The protocol used for the BBP SELEX was modeled after other SELEX experiments previously described (32, 33), with our experimental details described below. First, 25 pmol (1.5 $\times 10^{13}$ molecules) of the template oligonucleotide, 5’-GGGAATTTGATCCACACTAGTGAGTGAAGATATGGGATCCAAATTTTACGACAAATGATTATAATCCCG-3’ was amplified with 8 rounds of PCR with primers, 5’-GATAATCAGCTACTATAGGAGTAGTCCACACATCTGAGGCTTCTTTAACC-3’ and 5’-GACTTTACGACAGTGTCGCTGTCGATGTGTAATAGCTTCTTTAACC-3’. The PCR reaction was then phenol/chloroform-extracted and ethanol-purified. The DNA template was resuspended in 50 $\mu$L of water and spun through a Bio-Spin 6 size exclusion cartridge. The DNA template concentration was calculated using the A$_{260}$. A library of $1 \times 10^{13}$ molecules of DNA template was used for the in vitro transcription reaction. The RNA pool was transcribed as previously described (34). The RNA pool was cleaned up by phenol/chloroform extraction and ethanol precipitation. The RNA pellet was resuspended in 50 $\mu$L of water and spun through a Bio-Spin 6 size exclusion cartridge. The DNA template was removed by adding DNase (Promega) and incubated at 37°C for 1 h. The
RNA pool was again cleaned with another round of phenol/chloroform extraction and ethanol precipitation, resuspended, and spun through a Bio-Spin 6 size exclusion cartridge. GST-BBP-331 was incubated with glutathione-agarose beads (Sigma) at 4 °C for 30 min and washed three times with binding buffer (25 mM Tris (pH 7.5), 50 mM NaCl, 2 mM DTT, 0.1% Triton X-100, 0.07 mg/ml tRNA from bakers' yeast, Sigma) to remove any unbound protein. The purified RNA pool was added to the GST-BBP-331 protein bound to the glutathione-agarose beads in binding buffer for 15 min at room temperature. Unbound RNAs were removed by washing three times with binding buffer. Bound RNAs were eluted by phenol/chloroform extraction and ethanol precipitation. The eluted RNAs were resuspended in 20 μl of water. The selected RNAs were reverse-transcribed with avian myeloblastosis virus reverse transcriptase and the reverse primer, 5′-AACCTCGTAGTTCGA-3′, at 4 °C for 30 min. A new round of SELEX was begun again with PCR amplification. A total of seven rounds were done.

Concentration of RNA and protein for the different rounds of SELEX are as shown in Table 1. After the last round the PCR fragments were ligated into TOPO TA-2.1 vector and then transformed into TG1 competent cells (Stratagene). 48 individual colonies were picked and sequenced.

**Table 1**

| Round | RNA | Protein |
|-------|-----|---------|
| 1     | 18  | 6       |
| 2     | 20  | 2       |
| 3     | 10  | 1       |
| 4     | 5   | 0.5     |
| 5     | 2.5 | 0.25    |
| 6     | 1.0 | 0.1     |
| 7     | 0.5 | 0.05    |

Gels were dried and exposed on a Storm 860 PhosphorImager (Amersham Biosciences). Quantification of RNA substrate. Gels were analyzed using ImageQuant. Data were plotted in Kaleidagraph, and data points were fit to the following equation to calculate an approximate value for Kd for these RNAs because the RNA concentration could not be reduced sufficiently for an accurate measurement to be made. RNA #12 and #14 from group II bind with approximate Kd values of 300 and 740 pm, respectively (Fig. 4B). As predicted, we observed two complexes forming with BBP-331, indicating that BBP is able to bind both sites simultaneously. A distinct possibility for RNA #14 had been taken using a speed of 5 nm/min in a 1-cm cell at 4 °C. All spectra represent the average of four separate scans and are plotted in molar ellipticity (degree cm² dmol⁻¹) after the subtraction of buffer scans.

**RESULTS**

Selection of High Affinity RNA Sites for BBP—We made a GST-tagged BBP construct containing the RNA binding domains (KH, QUA2, and both zinc knuckles, amino acids 147–331; Fig. 1). We refer to this construct as GST-BBP-331. We then performed seven rounds of SELEX with GST-BBP-331 and a library of RNAs containing a 30-nucleotide randomized region. Selection for higher affinity targets was observed; pool 7 RNA bound BBP-331 (lacking GST) with 50-fold higher affinity than pool zero RNA (Fig. 2). Reverse transcription-PCR was performed on pool 7 RNA, and 48 individual clones were picked and sequenced. Fig. 3A shows all 48 sequences are unique and that all of these clones except one contain the highly conserved yeast branch point sequence UACUAAC. The one clone lacking the yeast BPS has a very similar sequence of UAUUAAC (see Fig. 3A, sequence 48).

Analysis of the selected sequences reveals that BBP has a preference for an adenosine downstream of the BPS sequence. This observation is consistent with two bioinformatic analyses which showed that an adenosine is overrepresented in yeast introns at the nucleotide immediately 3′ of the BPS as well as a biochemical study showing an adenosine at this position enhances splicing efficiency (35–37). Our data suggest that BBP recognizes this adenosine in yeast introns. Our data also reveal a preference for guanosines upstream of the BPS (Fig. 3B).

We divided the 48 sequences into three groups based on sequence and structural differences. The nine group I sequences contain a conserved motif GRUG (R = purine) directly upstream of the BPS. The first four nucleotides of this motif, GRUG, can base pair with CAUC in the 5′ constant terminal sequence of the SELEX RNA to form a potential stem-loop containing different loop sizes. All sequences contain a single guanosine between the stem-loop and BPS (Fig. 3C). Secondary structure predictions were done using mfold (38, 39). The seven group II sequences contain UACUAAC repeated twice, and one of these sequences (#12) is also predicted to have a stem-loop structure present in the RNA in a similar position as seen in group I. The linker between the stem-loop and the UACUAAC motif in sequence #12 is too nucleotides instead of one compared with group I sequences. The other 32 sequences fall into group III and contain the UACUAAC motif and lack any common secondary structure.

We picked two sequences from each group for gel shift assays to compare binding affinities between the three groups (Fig. 4). RNAs #1 and #7 from group I bind BBP-331 with the highest affinity. Kd values are well below 1 nm (Fig. 4A). It was not possible to determine a Kd for these RNAs because the RNA concentration could not be reduced sufficiently for an accurate measurement to be made. RNAs #12 and #14 from group II bind with approximate Kd values of 300 and 740 pm, respectively (Fig. 4B). As predicted, we observed two complexes forming with BBP-331, indicating that BBP is able to bind both sites simultaneously. A distinct possibility for RNA #14 had been...
that only one BBP would be able to bind because the two UACUAAC sites lack a linker (Fig. 3A), but clearly BBP is able to bind both UACUAAC motifs simultaneously. RNA #12 binds with 2.5-fold stronger affinity than RNA #14 possibly because RNA #12 contains a potential stem-loop two nucleotides upstream of the first UACUAAC site. From group III, RNAs #33 and #43 were used for analysis, and their $K_d$ values for BBP-331 are 4 and 1.4 nM (Fig. 4C). In general, we found that group I binds with the
highest affinity to BBP, possibly due to the upstream stem-loop, whereas group II binds with the next highest level of affinity, likely due to the duplicate UACUAAC sites, and group III binds BBP with the lowest relative affinity for the SELEX RNAs. To test the role of the upstream stem-loop in high affinity binding to BBP, we designed two 24-nucleotide RNAs, yBP24-SL and yBP24-SS, and compared their binding.

FIGURE 2. Comparison of BBP binding pool 0 and pool 7. A, gel shift assay comparing beginning RNA pool to round 7 RNA pool. BBP-331 concentrations used were 15, 30, 60, 125, 250, 500, 1000, and 2000 nM. B, quantification of the results of the gel shift assay.

FIGURE 3. SELEX results for BBP. A, the alignment of SELEX sequences for BBP (147–331) is separated into three groups. The UACUAAC motif is highlighted in red. Group I contains the GRUGG (R = purine) motif before the UACUAAC sequence, which is marked with an underline. Group II consists of the UACUAAC motif repeated twice within the randomized region. Group III contains just the UACUAAC motif and no common structural motif. B, the Web-logo consensus sequence of the 48 SELEX sequences was generated at the weblogo website. The size of the letters indicates degree of conservation. C, a representation of the common secondary structure and sequence motif of the group I RNAs. The gray portion represents the constant terminal region, and the black indicates the randomized sequence of the SELEX RNAs.
The yBP24-SL RNA contains an artificially designed stem-loop positioned a single guanosine nucleotide upstream of a UACUAAC motif, whereas yBP24-SS lacks the stem-loop (see Fig. 5A). To ensure the stability of the stem-loop in yBP24-SL, a stable UUCG tetraloop was used to cap the stem (40–42). The yBP24-SS RNA is similar to yBP24-SL except for point mutations that disrupt the stem-loop (Fig. 5A). The difference in binding affinity between yBP24-SL (Fig. 5, lanes 1–8) and yBP24-SS (Fig. 5, lanes 9–16) is more than 30-fold under stringent conditions (see "Experimental Procedures"). The stem-loop RNA has a \(K_d\) of 4 nM, and the single-stranded UACUAAC RNA binds with a \(K_d\) of 156 nM. Under less stringent conditions (see "Experimental Procedures"), the difference in binding between yBP24-SL and yBP24-SS is 10-fold (data not shown).

To determine whether a stem-loop upstream of the BPS in a yeast intron enhances BBP binding we tested the binding of an RNA from a yeast intron containing a putative stem-loop upstream of the BPS. The region upstream of several yeast introns was folded with mfold (38, 39) to identify candidates that contain a stem-loop upstream of the BPS that might function to enhance BBP RNA binding as we observed with the SELEX RNAs. From our search, we picked one good candidate for additional studies, the intron of the LSm gene (accession number NC_001134.7). The LSm intron contains a putative stem-loop directly upstream of the BPS (Fig. 6).

We found mutating the yeast LSm stem-loop has a more than 3-fold effect on binding (lanes 9–16). The \(K_d\) to the LSm WT RNA is 110 nM and increases to 367 nM for the LSm mutant RNA. Compensatory mutations that restore the stem, bring BBP binding back to wild type levels with a \(K_d\) of 103 nM supporting the formation of the stem-loop and indicating its presence enhances the binding of BBP to this RNA. These studies in combination with those in Fig. 5 clearly indicate BBP binding is enhanced by the presence of a stem-loop upstream of the BPS. The different effects the UUCG and LSm stem-loops exert on BBP binding suggest the location, stability, and possibly the sequence of the stem-loop all have a role in affecting BBP binding to RNA.

The Presence of an Upstream Stem-loop Does Not Enhance RNA Binding of Human SF1—One explanation for the enhanced binding of BBP to an RNA containing a stem-
loop upstream of the BPS is that this stem-loop presents the BPS in a favorable conformation for binding by BBP and that the stem-loop is not bound directly by BBP. This model suggests that human SF1 might also bind with higher affinity to an RNA with a stem-loop upstream of the BPS. To test this model we used a human SF1 construct containing the KH, QUA2, and one zinc knuckle (amino acids 134–297). Using the stem-loop containing RNA (yBP24-SL) and two different single-stranded RNAs (BP-22 and yBP24-SS), we found that a stem-loop upstream of the BPS did not enhance the binding of SF1 to RNA (Fig. 7). SF1 binds the single-stranded RNAs BP-22 and yBP24-SS with a $K_d$ of 310 and 908 nM, respectively. An enhancement in binding is not seen with yBP24-SL (Fig. 7, lanes 9–16), which binds with a $K_d$ of 340 nM. SF1 has similar affinities for the single-stranded RNA BP-22 and the stem-loop RNA yBP24-SL. This suggests that the strong interaction of BBP to the stem-loop UACUAAC RNA is not due to a purely entropic effect (the BPS sequence being presented in some manner by the stem-loop) but probably is the result of additional contacts between yeast BBP and the stem-loop.

**Zinc Knuckle Mutations Affect BBP Binding to the Stem-loop RNA**—To test the role of the zinc knuckles in RNA binding we performed mutational analysis of BBP. Mutating either of the zinc knuckles individually by changing the first two cysteines to alanines resulted in no observable effects on RNA binding to either the single-stranded BPS RNA substrates or to the stem-loop BPS RNA substrate (data not shown). Meanwhile, BBP-331 with both zinc knuckles mutated showed slightly more than a 1.5-fold weaker binding to the single-stranded yBP24-SS RNA (compare panels A and B in Fig. 8, lanes 1–6) and 10-fold weaker binding to the stem-loop yBP24-SL RNA (compare lanes 1–6 in panels A and B of Fig. 8). For the wild type BBP-331 protein binding to the yBP24-SS RNA, the $K_d$ is 156 nM, and for the double zinc knuckle mutant protein, the $K_d$ is 254 nM. For the stem-loop RNA, yBP24-SL, the BBP-331 binds with a $K_d$ of 4 nM, and for the double zinc knuckle mutant, the $K_d$ increases to 39 nM.

The other single-stranded RNA, BP-22, which was first used in Fig. 7, contains different sequences adjacent to the UACUAAC and needed here to determine whether these sequences differentially affect the binding of the BBP zinc knuckle mutant compared with sequences in yBP24-SS (compare lanes 1–6 in panels C and D of Fig. 8). Lanes 7–12 of panels C and D in Fig. 8 were a negative control, using a BP-22 with the branch point adenosine mutated. For the BP-22 RNA the difference in affinity between wild type BBP and the double zinc knuckle mutant BBP was on average little more than 3-fold ($K_d$ of 25 nM for wild type and 91 nM for the mutant). The single-stranded RNAs tested in Fig. 8 show a 2–3-fold difference, whereas the stem-loop RNA yBP24-SL shows a much larger 10-fold difference. The weaker binding of BBP to yBP24-SS compared with BP-22

**FIGURE 7.** Human SF1 binding to the BBP SELEX motif. A, a gel shift assay comparing human SF1 binding to an RNA containing the BBP high affinity stem-loop + UACUAAC motif (yBP24-SL, lanes 9–16) with two different single-stranded UACUAAC RNAs (lanes 1–8, BP-22; lanes 17–24, yBP24-SS). The yBP24-SL RNA will be represented from this point forward as the stem-loop UACUAAC picture depicted below lanes 17–24. SF1 concentrations used were 0.075, 0.15, 0.3, 0.6, 1.25, 2.5, and 5 μM. B, quantification of the binding in A.

**FIGURE 8.** Comparison of the RNA binding of wild type BBP-331 to the double zinc knuckle mutant BBP. WT BBP-331 (A) was bound to the stem-loop RNA, yBP24-SL (lanes 1–6), and single-stranded RNA, yBP24-SS RNA (lanes 7–12), and compared with the double zinc knuckle mutant BBP (B) protein (lanes 1–6, yBP24-SS; lanes 7–12, yBP24-SS). C and D, two more RNAs were tested with the WT BBP-331 (C) and double zinc knuckle mutant (D) BBP proteins, BP-22 RNA (lanes 1–6) and mutant BP-22 RNA (lanes 7–12, negative control). Protein concentrations used were 1.25, 2.5, 5, 10, and 20 nM for A and B and 25, 50, 100, 200, and 400 nM for C and D.
Yeast BBP SELEX

is the result of more stringent binding conditions used with yBP24-SS and yBP24-SL ("Experimental Procedures") to help reveal the difference in binding between BBP-331 and the double zinc knuckle mutant with the stem-loop RNA.

For us to properly compare the binding of the double zinc knuckle mutant and BBP-331 constructs, we later went back and tested the protein activities using the standard protocol (34) by titrating in increasing BP-22 RNA to a constant amount of protein. We found that the double zinc knuckle mutant was approximately half as active as BBP-331, which was ~100% active. Even after adjusting the previously mentioned double zinc knuckle mutants $K_d$ values by half, we concluded that these results suggest that the zinc knuckles may be partially responsible for the high affinity binding to the stem-loop RNA but play little, if any, role in the binding to the single-stranded BPS. Other researchers using NMR or deletion analysis also did not detect any interactions between the one zinc knuckle in the human protein and the single-stranded BPS RNA (6, 22). Our results suggest that the zinc knuckles bind the stem-loop but that other regions of the protein also likely bind the stem-loop because the double zinc knuckle mutation still binds the stem-loop RNA with higher affinity compared with the single-stranded RNAs.

**BBP Zinc Knuckle Linker Regions Help in RNA Binding, whereas the Second Zinc Knuckle Does Not**—Previously, we showed that the region of BBP containing the zinc knuckles is necessary for high affinity binding to the BPS but that the zinc knuckle domains worked through nonspecific binding because they could be replaced with the short basic peptide (RS)$_7$, which rescued the binding affinity. RNA binding specificity with the BBP-RS chimera was maintained, indicating that the KH-QUA2 domains contain the determinants for specific binding to the BPS (12). To better understand the role of the zinc knuckles and the linkers, we made a series of deletions of BBP to determine which amino acids are important for the RNA binding affinity of BBP. We first repeated our previous published work showing that deleting amino acids 270–331 abolished RNA binding (Fig. 9E). This construct contains amino acids 147–269 and only consists of the KH and QUA2 domains and is referred to as BBP-125 in Berglund et al. (12) and as BBP-269 here. This construct did not bind to any of the single-stranded or double-stranded RNAs tested.

Puzzled by the results in which deleting the zinc knuckles eliminates binding whereas the double zinc knuckle mutant has only a slight effect on binding, we asked which parts of the zinc knuckles are required for binding. We made three new truncations of the zinc knuckles (see Fig. 9, B–D) and compared the binding of these three proteins to the wild type BBP-331 using the RNAs BP-22 and mutant BP-22 (Fig. 9). Each of the zinc knuckle truncations except BBP-269 was confirmed to have ~100% activity by titrating in BP-22 RNA with a constant amount of protein (data not shown), as described in the standard protocol (34). BBP-269 could not be tested for activity because it does not bind RNA. The first construct tested was amino acids 147–296 (BBP-296), which deletes the second zinc knuckle and contains an intact first zinc knuckle and the linker region between the two zinc knuckles. Surprisingly, the binding of BBP-296 to the yBP22-SS RNA has a $K_d$ of 3 nM. This is an 8-fold increase in binding affinity compared with BBP-331, which binds with a $K_d$ of 25 nM. These results suggest that the second zinc knuckle is not involved in RNA binding and may potentially even inhibit RNA binding. We also happen to notice that it is the second zinc knuckle that is the most conserved between the two. As shown in the alignment in Fig. 1B, vertebrates are missing the first zinc knuckle, and their lone zinc knuckle is most similar to the second of those species that contain two. Deletion of this conserved zinc knuckle in humans also had little effect on RNA binding (6).

Deletion of the linker between the first and second zinc knuckle results in a fragment of BBP containing amino acids 147–287 (BBP-287). The loss of the linker reduces the binding affinity of BBP-287 to BP-22 RNA. The $K_d$ for BBP-287 is 69 nM compared with 3 nM for BBP-296. The binding complexes for BBP-287 are smeared, indicating the complex dissociates while migrating through the gel. The difference in binding between BBP-296 and BBP-287 is fairly significant. This suggests that the linker amino acids, 288–296, between the two zinc knuckles are important for RNA binding, and they either contact the RNA directly to stabilize the complex or do not contact the RNA but, instead, stabilize the complex indirectly in some manner.

The third truncation, in which amino acids 280–287 were deleted (BBP-279), results in a fragment of BBP that has only half the first zinc knuckle. This deletion removes the histidine and last cysteine of the first zinc knuckle that coordinates to the zinc ion. Fig. 9 shows BBP-279 binds with similar affinity compared with BBP-287. The $K_d$ for BBP-279 is 103 nM, whereas for BBP-287 it is 69 nM. The BBP-279-RNA complexes are also smeared, as with BBP-287. There is not a major difference in binding between the BBP-279 and BBP-287, suggesting this deletion (amino acids 280–287) does not play a major role in
RNA binding. However, deleting the next 10 amino acids (280–270, forming construct BBP-269) eliminates binding. Therefore, we conclude that the linker between the QUA2 domain and the first zinc knuckle in addition to the linker between the two zinc knuckles as mentioned above are important for RNA binding either by direct interactions with the RNA or through an indirect interaction.

These results are consistent with the zinc knuckle mutation experiments. If the key amino acids for binding are in the zinc knuckle linker regions, then disrupting the zinc knuckles themselves with point mutations would not be expected to affect RNA binding. We attempted to determine which amino acids in the linker region were important in RNA binding, but mutating both arginine 189 and lysine 190 to alanines had little effect on RNA binding (data not shown).

**The Folding and Stability of the KH-QUA2 Domains Are Not Affected by C-terminal Deletions**—To determine whether the zinc knuckles and linker regions affect the folding of the KH-QUA2 domains and their ability to bind RNA, we used circular dichroism (CD) to investigate the secondary structures of the various BBP constructs. BBP-296, BBP-279, and BBP-269 were all assayed using CD, and all of them show a typical curve in the far-UV CD for a protein containing α-helices (Fig. 10A).

The NMR structure of the human SF1 KH-QUA2 domains contains four α-helices and a β sheet (22), similar to the NMR structure of the *Xenopus* laevis Quaking protein, which also has the same four homologous α-helices plus two small α-helices in the variable loop region of the KH domain (43). The removal of the zinc knuckles and linkers does not affect the stability of the KH-QUA2 domains, as shown by the $T_m$ measurements, which differ very little among the various BBP fragments tested. All of the BBP fragments show a single melting transition at $\sim$55°C (Fig. 10B). The difference in the CD spectrum (218–226 nm) between BBP-296 and the two shorter fragments (BBP-279 and BBP-269) suggests that the linker between the zinc knuckles may contain an additional α-helix. The sequence of this linker is compatible with the formation of an α-helix because there are hydrophobic amino acids spaced 2–3 amino acids apart in this region in BBP. This region is not conserved among branch point-binding proteins; therefore, this α-helix might be specific to BBP and related to its ability to bind to RNAs containing stem-loops upstream of the branch point sequence.

**DISCUSSION**

In this report we show that the optimal binding site for BBP matches the highly conserved yeast branch point sequence UACUAAC. This is an expected result and confirms the power of SELEX to determine the binding sites of nucleic acid-binding proteins. The SELEX experiment also reveals that BBP interacts with nucleotides outside of the conserved BPS site. 30 of the 48 sequences contain an adenosine directly downstream of the UACUAAC sequence, indicating BBP prefers an adenosine at this position. Biochemical and bioinformatic studies have shown that adenosine is the preferred nucleotide at this position in yeast introns (35–37); indicating the preference of BBP for an adenosine at this position is biologically relevant. Previous studies have shown that stem-loops that bring the 5’ splice site and BPS sequence closer together in yeast introns are important for efficient spliceosome formation and splicing both in vitro and in vivo (44–47). We have shown that stem-loops upstream of the BPS enhance BBP binding, and we propose stem-loops upstream of the BPS may function to enhance BBP binding; therefore, we propose that RNA secondary structure affects the binding of BBP and U1 snRNP as well as later steps of yeast splicing. Finally, if the nucleotides upstream of the BPS are involved in base-pairing it means these nucleotides are not going to base pair with the BPS and inhibit BBP and U2 snRNP binding.

**Specificity of BBP Compared with Other KH Domain Proteins**—The STAR/GSG protein, QKI, was found to have a similar specificity to BBP with a consensus site of 5’-NA(A>C)U(A>C)/U(A>C) (48) and recently found to have an identical BBP core SELEX motif of UACUAAC (49). QKI is a mammalian RNA-binding protein that regulates translational control (50), alternative splicing (51), RNA export (52), and RNA stability (53, 54). The identification of the same binding site for BBP and QKI is not surprising since there is 70% sequence identity (14 of 20 amino acids) and 85% sequence similarity (17 of 20 amino acids) in the residues contacting the core UACUAAC sequence. This comparison was made by using the SF1-BPS structure as a model for the other STAR/GSG proteins similar to the analysis done by Ryder et al. (55) and Lehmann-Blount and Williamson (56). There appears to be a correlation between the conservation of STAR/GSG KH-QUA2 domains and their SELEX motifs or known binding sites. GLD-1, another STAR/GSG protein,
also has a binding site (UACU(A/C)A) that is almost identical to the BPS (55, 56). GLD-1 is a Caenorhabditis elegans protein involved in germ line development and apoptosis that represses translation of such mRNAs as tra-2 and cep-1/p53 (57–60). GLD-1 and BBP also have a 70% sequence identity and 85% sequence similarity in their RNA binding amino acids.

As the KH-QUA2 domains of the STAR/GSG proteins become less similar to BBP, their SELEX binding motifs or consensus binding sites become less like the UACUAAC BBP binding motif. For example, SAM-68 is only 40% identical and 55% similar to BBP in the RNA contact residues. The SAM-68 SELEX sequence is U/A-rich with the tightest binding motif of UAAA (61). This binding site is shorter than the other STAR/GSG proteins by appearing to lack the upstream nucleotides of UAC that the QUA2 domain binds. The SAM-68 QUA2 domain is significantly different from that of BBP and SF1 and may explain why there is no extended upstream UAC binding. KH domain proteins that are not part of the STAR/GSG family tend to have shorter core binding motifs. For example, NOVA with three KH domains has been found to bind RNAs containing multiple UCAY repeats (62).

Although yeast BBP and human SF1 are orthologs, they have different RNA binding specificities and significantly different RNA binding affinities. The difference in affinity is most obvious when comparing the binding of BBP-296 to that of SF1 (amino acids 134–297). Both constructs contain one zinc knuckle and the KH and QUA2 domains, yet BBP-296 binds the yeast BBP-22 RNA with a $K_d$ of 3 nM, whereas SF1 (134–297) binds with a $K_d$ of 310 nM. The lower binding constants we have observed compared with those previously published (1, 5) are likely due to the use of more highly purified proteins as well as optimized binding conditions. We propose that BBP makes several additional contacts with the BPS and other regions of the RNA, increasing BBP binding affinity compared with SF1. These additional contacts could also potentially explain why BBP displays more specificity than SF1. For example, mutational analysis showed the first uridine of UACUAAC is not recognized by SF1 (1), and the NMR structure of SF1 with the BPS also shows no contacts between SF1 and the uridine (22). For BBP binding, this position is clearly important; every BBP binding (1). The determinants for BPS recognition are similar to BBP in the RNA contact residues. The SAM-68 QUA2 domain is significantly different from that of BBP and SF1 and may explain why there is no extended upstream UAC binding. KH domain proteins that are not part of the STAR/GSG family tend to have shorter core binding motifs. For example, NOVA with three KH domains has been found to bind RNAs containing multiple UCAY repeats (62).

**The Role of the Zinc Knuckles and Linker Regions**—Our deletion studies of BBP reveal that amino acids 270–296 are important for RNA binding. Mutating the first two cysteines (amino acids 273 and 276) to alanines has no effect on RNA binding, indicating that coordinating the zinc ion is not important for RNA binding but that the amino acids within this region are important for binding. A zinc knuckle is only found in this region of BBP in non-vertebrates; SF1 and the branch point-binding protein in other vertebrates contain only the more conserved second zinc knuckle. Our finding that this zinc knuckle and its surrounding amino acids have no role in RNA binding and may even inhibit binding indicates this domain is not conserved for RNA binding but is likely conserved for another purpose, possibly involved in protein-protein interactions. This zinc knuckle may also be important for the removal of BBP from the BPS when U2 snRNP binds the BPS. In support of this model, we have found that expressing BBP containing a mutated second zinc knuckle in yeast cells is lethal. In the future we will determine whether this mutant BBP blocks the splicing of essential genes by not being properly removed from pre-mRNAs.

**Acknowledgments**—We thank members of the Berglund laboratory for experimental advice and helpful comments on the manuscript.

**REFERENCES**

1. Berglund, J. A., Chua, K., Abovich, N., Reed, R., and Rosbash, M. (1997) Cell 89, 781–787
2. Rutz, B., and Seraphin, B. (1999) RNA 5, 819–831
3. Seraphin, B., and Rosbash, M. (1989) Cell 59, 349–358
4. Michaud, S., and Reed, R. (1991) Genes Dev. 5, 2534–2546
5. Berglund, J. A., Abovich, N., and Rosbash, M. (1998) Genes Dev. 12, 858–867
6. Rain, J. C., Rafi, Z., Rhani, Z., Legrain, P., and Kramer, A. (1998) RNA 4, 551–565
7. Abovich, N., and Rosbash, M. (1997) Cell 89, 403–412
8. Fromont-Racine, M., Rain, J. C., and Legrain, P. (1997) Nat. Genet. 16, 277–282
9. Valcarcel, J., Gaur, R. K., Singh, R., and Green, M. R. (1996) Science 273, 1706–1709
10. Fleckner, J., Zhang, M., Valcarcel, J., and Green, M. R. (1997) Genes Dev. 11, 1864–1872
11. Lin, K. T., Lu, R. M., and Tarn, W. Y. (2004) Mol. Cell. Biol. 24, 9176–9185
12. Berglund, J. A., Fleming, M. L., and Rosbash, M. (1998) RNA 4, 998–1006
13. Burd, C. G., and Dreyfuss, G. (1994) Science 265, 615–621
14. Gibson, T. J., Thompson, J. D., and Heringa, J. (1993) FEBS Lett. 324, 361–366
15. Lukong, K. E., and Richard, S. (2003) Biochim. Biophys. Acta 1653, 73–86
16. Siomi, H., Matunis, M. I., Michael, W. M., and Dreyfuss, G. (1993) Nucleic Acids Res. 21, 1193–1198
17. Vernet, C., and Artzt, K. (1997) Trends Genet. 13, 479–484
18. Wu, J., Zhou, L., Tonissen, K., Tee, R., and Artzt, K. (1999) J. Biol. Chem. 274, 29202–29210
19. Chen, T., and Richard, S. (1998) Mol. Cell. Biol. 18, 4863–4871
20. Chen, T., Damaj, B. B., Herrera, C., Lasko, P., and Richard, S. (1997) Mol. Cell. Biol. 17, 5707–5718
21. Di Fruscio, M., Chen, T., Bonyadi, S., Lasko, P., and Richard, S. (1998) J. Biol. Chem. 273, 30122–30130
22. Liu, Z., Luyten, I., Bottomley, M. J., Messias, A. C., Hougninou-Molango, S., Spranger, R., Zanier, K., Kramer, A., and Sattler, M. (2001) Science 294, 1098–1102
23. Berglund, J. A., Charpentier, B., and Rosbash, M. (1997) Nucleic Acids Res. 25, 1042–1049
24. Berkowitz, R. D., Luban, J., and Goff, S. P. (1993) J. Virol. 67, 7190–7200
25. Berkowitz, R. D., and Goff, S. P. (1994) Virology 202, 233–246
26. Berkowitz, R. D., Ohagen, A., Hoglund, S., and Goff, S. P. (1995) J. Virol. 69, 6445–6456
27. De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F. (1998) Science 279, 384–388
28. Dannull, J., Surovoy, A., Jung, G., and Moelling, K. (1994) EMBO J. 13, 1525–1533
29. Clever, J., Sassetti, C., and Parslow, T. G. (1995) J. Virol. 69, 2101–2109
30. Morellet, N., Demene, H., Teilleux, V., Huynh-Dinh, T., de Rocquigny, H., Fournie-Zaluski, M. C., and Roques, B. P. (1998) J. Mol. Biol. 283, 5707–5718
Cavaloc, Y., Bourgeois, C. F., Kister, L., and Stevenin, J. (1999) RNA 5, 468–483.

32. Tuerk, C., and Gold, L. (1990) Science 249, 505–510.

33. Ellington, A. D., and Szostak, J. W. (1990) Nature 346, 818–822.

34. Smith, C. W. J. (ed) (1998) RNA:Protein Interactions, A Practical Approach. The Practical Approach Series (Hames, B. D., ed) pp. 40–43, 128, and 129, Oxford University Press, Oxford.

35. Rain, J. C., and Legrain, P. (1997) EMBO J. 16, 1759–1771.

36. Spingola, M., Grate, L., Haussler, D., and Ares, M., Jr. (1999) RNA 5, 221–234.

37. Lopez, P. J., and Seraphin, B. (1999) RNA 5, 1135–1137.

38. Smith, C. W. J. (ed) (1998) RNA:Protein Interactions, A Practical Approach. The Practical Approach Series (Hames, B. D., ed) pp. 40–43, 128, and 129, Oxford University Press, Oxford.

39. Rain, J. C., and Legrain, P. (1997) EMBO J. 16, 1759–1771.

40. Spingola, M., Grate, L., Haussler, D., and Ares, M., Jr. (1999) RNA 5, 221–234.

41. Tuerk, C., Gauss, P., Thermes, C., Groebe, D. R., Gayle, M., Guild, N., Stormo, G., d’Aubenton-Carafa, Y., Uhlenbeck, O. C., Tinoco, I., Jr., Brody, E. N., and Gold, L. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 5901–5905.

42. Antao, V. P., and Tinoco, I., Jr. (1992) Nucleic Acids Res. 20, 819–824.

43. Maguire, M. L., Guler-Gane, G., Nietlispach, D., Raine, A. R., Zorn, A. M., Standart, N., and Brodhurst, R. W. (2005) J. Mol. Biol. 348, 265–279.

44. Charpentier, B., and Rosbash, M. (1996) RNA 2, 509–522.

45. Libri, D., Stutz, F., McCarthy, T., and Rosbash, M. (1995) RNA 1, 425–436.

46. Newman, A. (1987) EMBO J. 6, 3833–3839.

47. Howe, K. J., and Ares, M., Jr. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 12467–12472.

48. Ryder, S. P., and Williamson, J. R. (2004) RNA 10, 1449–1458.

49. Galarneau, A., and Richard, S. (2005) Nat. Struct. Mol. Biol. 12, 691–698.

50. Saccomanno, L., Loushin, C., Jan, E., Punek, E., Artzt, K., and Goodwin, E. B. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 12605–12610.

51. Wu, I. L., Reed, R. B., Grabowski, P. J., and Artzt, K. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 4233–4238.

52. Larocque, D., Pilotte, J., Chen, T., Cloutier, F., Massie, B., Pedraza, L., Couture, R., Lasko, P., Almazan, G., and Richard, S. (2002) Neuron 36, 815–829.

53. Larocque, D., Galarneau, A., Liu, H. N., Scott, M., Almazan, G., and Richard, S. (2005) Nat. Neurosci. 8, 27–33.

54. Li, Z., Zhang, Y., Li, D., and Feng, Y. (2000) J. Neurosci. 20, 4944–4953.

55. Ryder, S. P., Frater, L. A., Abramovitz, D. L., Goodwin, E. B., and Williamson, J. R. (2004) Nat. Struct. Mol. Biol. 11, 20–28.

56. Lehmann-Blount, K. A., and Williamson, J. R. (2005) J. Mol. Biol. 346, 91–104.

57. Schumacher, B., Hanazawa, M., Lee, M. H., Nayak, S., Volkmann, K., Hofmann, E. R., Hengartner, M., Schedl, T., and Gartner, A. (2005) Cell 120, 357–368.

58. Clifford, R., Lee, M. H., Nayak, S., Ohmachi, M., Giorgini, F., and Schedl, T. (2000) Development 127, 5265–5276.

59. Jan, E., Motzny, C. K., Graves, L. E., and Goodwin, E. B. (1999) EMBO J. 18, 258–269.

60. Kuwabara, P. E., and Perry, M. D. (2001) BioEssays 23, 596–604.

61. Lin, Q., Taylor, S. J., and Shalloway, D. (1997) J. Biol. Chem. 272, 27274–27280.

62. Jensen, K. B., Musunuru, K., Lewis, H. A., Burley, S. K., and Darnell, R. B. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 5740–5745.

63. Manceau, V., Swenson, M., Le Caer, J. P., Sobel, A., Kielkopf, C. L., and Maucuer, A. (2006) FEBS J. 273, 577–587.