

# Archaeal MBF1 binds to 30S and 70S ribosomes via its helix–turn–helix domain

Fabian BLOMBACH<sup>\*1,2</sup>, Helene LAUNAY<sup>†1,3</sup>, Ambrosius P. L. SNIJDERS<sup>‡</sup>, Violeta ZORRAQUINO<sup>\*</sup>, Hao WU<sup>\*4</sup>, Bart DE KONING<sup>\*</sup>, Stan J. J. BROUNS<sup>\*</sup>, Thijs J. G. ETTEMA<sup>§</sup>, Carlo CAMILLONI<sup>||</sup>, Andrea CAVALLI<sup>||</sup>, Michele VENDRUSCOLO<sup>||</sup>, Mark J. DICKMAN<sup>‡</sup>, Lisa D. CABRITA<sup>†</sup>, Anna LA TEANA<sup>¶</sup>, Dario BENELLI<sup>\*\*</sup>, Paola LONDEI<sup>\*\*</sup>, John CHRISTODOULOU<sup>+5</sup> and John VAN DER OOST<sup>\*5</sup>

<sup>\*</sup>Laboratory of Microbiology, Wageningen University, Wageningen 6703HB, The Netherlands

<sup>†</sup>Institute of Structural and Molecular Biology, University College London (UCL) and Birkbeck College, University of London, London WC1E 6BT, U.K.

<sup>‡</sup>Department of Chemical and Process Engineering, University of Sheffield, Sheffield S1 3JD, U.K.

<sup>§</sup>Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, Uppsala SE-75137, Sweden

<sup>||</sup>Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.

<sup>¶</sup>Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, Ancona 60131, Italy

<sup>\*\*</sup>Dipartimento Biotecnologie Cellulari ed Ematologia, Università degli Studi di Roma "La Sapienza", Roma 00161, Italy

MBF1 (multi-protein bridging factor 1) is a protein containing a conserved HTH (helix–turn–helix) domain in both eukaryotes and archaea. Eukaryotic MBF1 has been reported to function as a transcriptional co-activator that physically bridges transcription regulators with the core transcription initiation machinery of RNA polymerase II. In addition, MBF1 has been found to be associated with polyadenylated mRNA in yeast as well as in mammalian cells. aMBF1 (archaeal MBF1) is very well conserved among most archaeal lineages; however, its function has so far remained elusive. To address this, we have conducted a molecular characterization of this aMBF1. Affinity purification of interacting proteins indicates that aMBF1 binds to ribosomal subunits. On sucrose density gradients, aMBF1 co-fractionates with free 30S ribosomal subunits as well as with 70S ribosomes engaged in translation. Binding of aMBF1 to ribosomes does not inhibit

translation. Using NMR spectroscopy, we show that aMBF1 contains a long intrinsically disordered linker connecting the predicted N-terminal zinc-ribbon domain with the C-terminal HTH domain. The HTH domain, which is conserved in all archaeal and eukaryotic MBF1 homologues, is directly involved in the association of aMBF1 with ribosomes. The disordered linker of the ribosome-bound aMBF1 provides the N-terminal domain with high flexibility in the aMBF1–ribosome complex. Overall, our findings suggest a role for aMBF1 in the archaeal translation process.

**Key words:** helix–turn–helix domain (HTH domain), multi-protein bridging factor 1 (MBF1), ribosome, *Sulfolobus*, transcription, translation, translation fidelity.

## INTRODUCTION

Archaea and eukaryotes share a common set of proteins involved in genetic information processing (transcription, translation and replication), including several proteins containing HTH (helix–turn–helix) domains [1–3]. Most of these proteins carry out functions within the core transcription machinery in both archaea and eukaryotes. This includes the eukaryotic protein MBF1 (multi-protein bridging factor 1) that has been shown to act as a transcriptional co-activator, transmitting the signal from eukaryote-specific transcription factors to the core transcription machinery by physically bridging these factors with the TBP (TATA-box-binding protein) via the HTH domain of MBF1 [4–7].

Besides its characterized function as a transcriptional co-activator, previous studies suggest that eukaryotic MBF1 might be a moonlighting protein. In yeast, a frameshift mutation in the *mbf1* sequence as well as deletion of the entire *mbf1* gene have been

shown to alter the rate of ribosomal frameshifting as well as the sensitivity of the strains to aminoglycoside antibiotics including paromomycin [8–10]. In addition, yeast MBF1 has been recently shown to co-purify with Pab1 [poly(A)-binding protein 1]. The interaction is sensitive to RNase treatment, suggesting that MBF1 is associated with polyadenylated mRNA [11]. Furthermore, yeast MBF1 binds directly to RNA via the less-conserved N-terminal domain [11]. Similarly, human MBF1 was also identified as an mRNA-binding protein in embryonic stem cells and HEK (human embryonic kidney)-293 cells [12,13].

When the first archaeal genomes became available, aMBF1 (archaeal MBF1) orthologues were identified on the basis of sequence homology encompassing the HTH domain [2]; however, aMBF1 has remained functionally uncharacterized ever since. The evolutionary conservation of TBP across all eukaryotes and archaea might suggest that TBP and aMBF1 also interact in archaea [5]. However, the fact that experimental investigations using chimaeric constructs bearing HTH domains originating

Abbreviations: aIF6, archaeal translation initiation factor 6; aMBF1, archaeal MBF1; aMBF1-C, C-terminal domain of aMBF1; aMBF1-N, N-terminal domain of aMBF1; HTH, helix–turn–helix; MBF1, multi-protein bridging factor 1; PFG, pulsed-field gradient; TBP, TATA-box-binding protein.

<sup>1</sup> These authors contributed equally to this work.

<sup>2</sup> Present address: Research Department of Structural and Molecular Biology, Institute of Structural and Molecular Biology, University College London (UCL), London WC1E 6BT, U.K.

<sup>3</sup> Present address: CNRS-UMR 8576-University of Lille1, Villeneuve d'Ascq, Lille FR-59655, France.

<sup>4</sup> Present address: Cardiovascular Biology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, U.S.A.

<sup>5</sup> Correspondence may be addressed to either of these authors (email j.christodoulou@ucl.ac.uk or john.vanderost@wur.nl).

Co-ordinates and structure factors of the aMBF1 (archaeal multi-protein bridging factor 1) helix–turn–helix domain have been deposited in the PDB under code 2MEZ. The NMR assignment data have been deposited in the BMRB under accession number 19028.

from different archaeal species were unable to functionally replace the endogenous HTH domain of MBF1 in yeast hints at a functional difference between archaeal and eukaryotic MBF1 [14]. This is corroborated by the observation that the well-conserved C-terminal HTH domain of archaeal and eukaryotic MBF1 is linked to an N-terminal MBF1-specific domain in eukaryotic MBF1, and to a distinct N-terminal domain (predicted zinc-ribbon fold) in aMBF1 orthologues [2,15].

In the present study, we make an important step forward towards elucidating the function of aMBF1 by presenting a biochemical and functional characterization of an aMBF1 orthologue. The results show that aMBF1 from *Sulfolobus solfataricus* binds to the small ribosomal subunit during translation via its conserved HTH motif. These results suggest an unexpected physiological function for aMBF1 linked to translation.

## EXPERIMENTAL

### Molecular cloning of *mbf1* and *in vitro* translation templates

The *mbf1* gene from *S. solfataricus* P2 (GeneID 1455418) [16] was PCR-amplified from genomic DNA using forward primer 5'-GCGCGCATATGCAAGCTAATAGTGAAGAATAC-3' and reverse primer 5'-GCGCGCTCGAGCTTCTTCCCTCTTTAA-TATTTACC-3' and cloned into vector pET26b via NdeI and XhoI restriction sites (underlined), resulting in plasmid pWUR298. For molecular cloning of the isolated N-terminal domain of aMBF1 (aMBF1-N) encompassing amino acids 1–58, the reverse primer was exchanged for 5'-GCGCGGCC-CTCGAGCTTACGTGTTTCGCTTTTCTTAC-3' (resulting in plasmid pWUR557). For molecular cloning of the isolated C-terminal domain of aMBF1 (aMBF1-C) encompassing amino acids 57–165, the forward primer was exchanged for 5'-GCGCGGCCCATATGCGTAAGAAAGCCACTCTTAAACCACC-3' (resulting in plasmid pWUR300). As no complementary stop codons were included in the reverse primers, all three proteins were designed to have a C-terminal His<sub>6</sub> tag.

### Heterologous expression and purification of aMBF1

Plasmids pWUR298, pWUR300 and pWUR557 were transformed into *Escherichia coli* Rosetta (DE3) cells (Novagen), and heterologous expression was carried out using standard procedures. Full-length aMBF1, aMBF1-N and aMBF1-C were produced as follows. Cells were resuspended in buffer TK300 (20 mM Tris/HCl, pH 7.5, 300 mM KCl and 1 mM DTT) and passed through a French pressure cell (Aminco) three times at 16000 psi (1 psi = 6.9 kPa). Cell debris was removed by centrifugation at 37000 g for 30 min at 4 °C. The cell-free extract was incubated at 75 °C for 15 min and the heat-unstable proteins from the expression host were then removed by centrifugation. The resulting supernatant was then purified further by nickel-affinity chromatography using His-Select Nickel affinity gel (Sigma) and size-exclusion chromatography using a HiLoad 16/60 Superdex 75 column (GE Healthcare) equilibrated in TK300. Elution fractions containing aMBF1 were combined and the buffer was exchanged by ultrafiltration to 20 mM Tris/HCl (pH 7.2), 50 mM NH<sub>4</sub>Cl, 10 mM magnesium acetate, 1 mM DTT and 10% (v/v) glycerol. Aliquots of the proteins were snap-frozen in liquid nitrogen and stored at –80 °C. All proteins were quantified on the basis of their absorption at 280 nm using calculated molar absorption coefficients for the respective protein [17].

For isotopic labelling of recombinant proteins for NMR studies, the heterologous expression was carried out in M9 medium (6.78 g/l Na<sub>2</sub>HPO<sub>4</sub>, 3 g/l KH<sub>2</sub>PO<sub>4</sub>, 0.5 g/l NaCl, 2 mM MgSO<sub>4</sub>, 100 μM CaCl<sub>2</sub>, 1 × MEM vitamin solution), with <sup>15</sup>NH<sub>4</sub>Cl (1 g/l) and 0.4% [<sup>13</sup>C]glucose (Cambridge Isotope Laboratories) as the sole sources of nitrogen and carbon respectively.

### Immunodetection of aMBF1, aMBF1-C, aMBF1-N and Alba

Rabbit antiserum against recombinant aMBF1-C was produced at Eurogentec. Antiserum from the final bleed was purified over Protein A-agarose (Sigma–Aldrich), and antibodies were allowed to react with digoxigenin-3-*O*-methylcarbonyl-ε-aminocaproic acid-*N*-hydroxysuccinimide ester (Roche) at a 1:10 molar ratio according to the manufacturer's protocol. Immunodetection of aMBF1 and aMBF1-C was performed as described previously for *S. solfataricus* HfiX [18] using digoxigenin-labelled primary antibodies and alkaline-phosphatase-conjugated anti-digoxigenin Fab fragments (Roche) as secondary antibodies. For immunodetection of aMBF1-N, samples were resolved by Tris/Tricine SDS/PAGE [19]. Immunodetection was performed with anti-His<sub>6</sub> antibody (Roche) following the manufacturer's recommendation. For the immunodetection of Alba, a 1:3000 dilution of Alba antiserum (a gift from Malcolm White, University of St. Andrews, St. Andrews, U.K.) was used with alkaline-phosphatase-conjugated donkey anti-(goat IgG) (Promega) (1:10000 dilution). For the detection, CDP-Star<sup>®</sup> reagent (NEB) was used and signals were captured on Kodak Biomax films.

### Cell growth of *S. solfataricus* strains

*S. solfataricus* strain P2 was obtained from the DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen; Braunschweig, Germany), and strain PBL2025 was a gift from Paul Blum (University of Nebraska–Lincoln, Lincoln, NE, U.S.A.). The generation of the *Δmbf1* strain was achieved generally according to [20] (exact details available from B.d.K. on request). For the preparation of cell lysates for *in vitro* translation experiments and ribosome co-purification, cell cultures were grown in shake flasks at 80 °C on modified Brock medium [21] supplemented with 0.4% sucrose and 0.1% tryptone to a *D*<sub>600</sub> of 0.7–1.2 and 0.4 respectively. For all other cell cultures of *S. solfataricus* P2 strain, modified Brock medium was supplemented with 0.3% glucose. For isotope labelling of proteins, <sup>14</sup>NH<sub>4</sub>Cl was replaced by <sup>15</sup>NH<sub>4</sub>Cl (Cambridge Isotope Laboratories) as described previously [22].

### Affinity purification and identification of interacting proteins

Recombinant aMBF1 was purified as described above, except that Tris was replaced by triethanolamine and DTT was omitted. A 600 μg amount of aMBF1 or His<sub>6</sub> peptide (Innovagen) was coupled to 200 μl of pre-washed *N*-hydroxysuccinimide-activated Sepharose resin (GE Healthcare) according to the manufacturer's protocol. *S. solfataricus* S30 cell lysates were produced as described previously [23] in 20 mM Tris/HCl (pH 7.4), 50 mM KCl, 10 mM magnesium acetate, 5 mM CaCl<sub>2</sub>, 20% (v/v) glycerol, 0.5% Nonidet P40 and 1 mM DTT, and were diluted to 5 mg/ml. Then, 25 mg of lysate was treated with 300 μg of micrococcal nuclease (Fermentas) at 30 °C for 10 min to degrade nucleic acids that can lead to bridging effects and overall increased background [24].

Volumes of 25 μl of aMBF1-coupled beads were added to 2 ml of further diluted cell lysate (0.5–4 mg of protein/ml) and incubated for 2 h at 4 °C. The cell lysates were then transferred

to a spin column, and beads were washed four times with 500  $\mu$ l of 20 mM Tris/HCl (pH 7.4), 105 mM KCl, 20% (v/v) glycerol and 0.1% Nonidet P40. Proteins were eluted with 100  $\mu$ l of 2 $\times$  SDS/PAGE sample buffer. As a negative control, the same experiment was performed using His<sub>6</sub> peptide-coupled beads and <sup>15</sup>N-labelled cell lysate. Equal volumes of <sup>14</sup>N- and <sup>15</sup>N-labelled elution fractions from aMBF1-coupled or His<sub>6</sub> peptide-coupled beads respectively were mixed and resolved by SDS/PAGE [8% (30:2) gel].

Gel lanes were cut into eight blocks, destained and treated with trypsin for in-gel digestion. Peptides were analysed by LC–MS/MS using an Ultimate 3000 HPLC and a MaXis UHR-Q-TOF (ultra-high-resolution quadrupole time-of-flight) tandem mass spectrometer. All MS data were acquired in profile mode. Bruker .baf files were converted into mzXML files by CompassExport. Mascot Distiller then used mzXML files for peak detection and quantification. Mascot version 2.1 was used to search the peak lists against a database containing the *S. solfataricus* P2 proteome sequences in fasta format concatenated with a randomized version of the same database. The Mascot Distiller precursor quantification protocol was performed using a <sup>15</sup>N metabolic incorporation percentage of 98%.

### **In vitro translation and formaldehyde cross-linking**

For the template used in the *in vitro* translation reactions, a synthetic gene was generated by GENEART (Germany) (pWUR560) based on the *orf104* gene [25] (encoding ribosomal protein L30ae) that has been well-characterized in the *S. solfataricus* *in vitro* translation systems [23,26,27] (Supplementary Figure S1 at <http://www.biochemj.org/bj/462/bj4620373add.htm>). *In vitro* transcription with the Megascript T7 kit (Ambion) was carried out on SacI-linearized plasmid pWUR560 and transcripts were purified on RNeasy spin columns (Qiagen).

The aIF6 (archaeal translation initiation factor 6) ORF was PCR-amplified from a pET-based expression plasmid [26] using forward primer 5'-GCGCGCGGTACCGCCTAATGAGGTGAAATGTAATGAATCTGCAAAGGTTATCAGTTTTTGG-3' and reverse primer 5'-GCGCGCTCTAGATCATTACC-TAATGCTTTTTGAATTC-3' (start codon in bold) and cloned via KpnI and XbaI restriction sites (underlined). The forward primer contains a leader sequence identical with that found in the *orf104* template [27]. The aIF6 mRNA was generated by *in vitro* transcription similarly to as described above.

The preparation of *S. solfataricus* cell lysates and *in vitro* translation reactions as well as cross-linking of 70S ribosomes were conducted as described previously [18], with a pre-incubation of the cell lysates for 10 min at 73°C to uncharge all ribosomes.

To localize aMBF1 in the cell lysates, samples were loaded on 10.5 ml 10–30% sucrose gradients in buffer A (20 mM Tris/HCl, pH 7.4, 40 mM NH<sub>4</sub>Cl, 10 mM magnesium acetate and 1 mM DTT) and centrifuged for 4 h at 220 000 g in a TST41.14 rotor (Kontron Instruments) at 4°C. Fractions of 500  $\mu$ l were subjected to sodium deoxycholate/trichloroacetic acid precipitation. Protein pellets were resuspended in 20  $\mu$ l of 3 $\times$  SDS/PAGE sample buffer, resolved by Bis-Tris SDS/PAGE [8% (30:2) gel] using Mes running buffer, and aMBF1-containing fractions were identified by immunodetection.

For detection of newly synthesized proteins during *in vitro* translation, the assays were modified as follows. The lysates were supplemented with micrococcal nuclease and 1 mM CaCl<sub>2</sub> for the degradation of endogenous mRNA for 20 min at 20°C. The reaction was stopped by the addition of 2 mM sodium EGTA

(pH 7.4). A 25  $\mu$ l volume of *in vitro* translation assay mixture contained 17.3  $\mu$ Ci of L-[<sup>35</sup>S]methionine (specific radioactivity >1000 Ci/mmol) (PerkinElmer) and 700 ng of mRNA. After 50 min of incubation at 72°C, the reaction was stopped by the addition of 0.5 volume of 3 $\times$  SDS/PAGE sample buffer. Then, 15  $\mu$ l of each sample was separated by Tris/Tricine SDS/PAGE (15% gel). Gels were dried and exposed to autoradiography.

### **Determination of aMBF1 expression during different growth phases**

Cells were grown as described above. Aliquots of 50  $\mu$ l were withdrawn at the given time points. Cells were harvested by centrifugation at 4°C. After resuspension in buffer A, cells were lysed by sonication. After removal of cell debris by centrifugation at 16 100 g for 40 min at 4°C, the protein concentration of the lysates was determined using the Bradford assay and adjusted to 4 mg/ml for all lysates. Lysates were supplemented with 0.5 volumes of 3 $\times$  SDS/PAGE sample buffer and proteins were resolved by Bis-Tris SDS/PAGE using Mes running buffer. Immunodetection of aMBF1 and Alba was carried out as described above.

### **Isolation of ribosomes and ribosomal subunits**

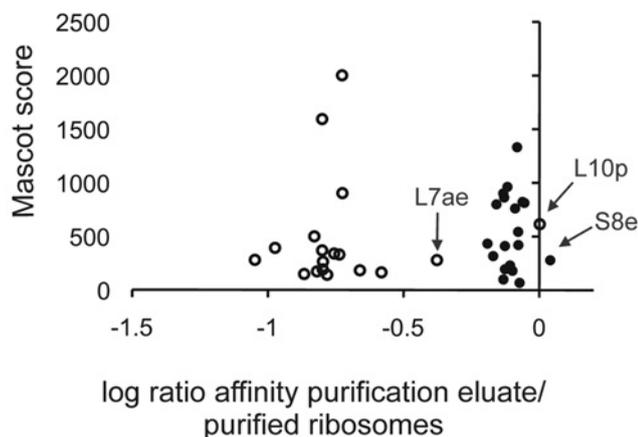
Ribosomes were purified in buffer A as described previously [23,28] and quantified on the basis of absorbance measurements at 260 nm (1 A<sub>260</sub> unit corresponds to 40 pmol). Isolated 30S ribosomal subunits were recovered from sucrose gradients by ultrafiltration. Each A<sub>260</sub> unit was assumed to correspond to 70 pmol of 30S ribosomal subunit on the basis of a concentration of 70  $\mu$ g per A<sub>260</sub> unit and a calculated molecular mass of ~1 MDa on the basis of the genome sequence.

### **Ribosome-binding assays**

For ribosome-binding assays, 100 pmol of recombinant aMBF1 or aMBF1-C was incubated with 100 pmol of ribosomes in 100  $\mu$ l of buffer A for 30 min on ice. Samples were loaded on 10.5 ml 10–30% sucrose gradients in buffer A and processed further as described above.

### **NMR spectroscopy**

All aMBF1 and ribosome preparations were buffer-exchanged to 10 mM Hepes/KOH (pH 7.3), 40 mM NH<sub>4</sub>Cl, 10 mM MgCl<sub>2</sub> and 1 mM DTT. A 190  $\mu$ M solution of <sup>15</sup>N-labelled aMBF1 was supplemented with 10% <sup>2</sup>H<sub>2</sub>O and 0.033% 3-(trimethylsilyl)propane-1-sulfonic acid. NMR spectra were recorded on a Bruker Avance III 700 MHz spectrometer equipped with a TXI cryogenic probe. The Fast-HSQC pulse sequence [29] was used for heteronuclear spectra of aMBF1 constructs, with a spectral width of 30 p.p.m. in the indirect dimension and 128 complex points acquired. The direct dimension was recorded with 2048 complex points and a spectral width of 12 p.p.m. All spectra were recorded at 25°C. <sup>15</sup>N-heteronuclear relaxation rates were measured using standard procedures [30], and were analysed using the Lipari–Szabo formalism [31]. PFG (pulsed-field gradient) diffusion experiments were performed using the stimulated echo-PFG [32] and the heteronuclear Xstep-PFG sequence [33] with a bipolar 1 ms square gradient for isolated protein and 2 ms square gradient for ribosomal subunits with the gradient strength varying from 5% to 95% of the maximum strength (0.563 T·m<sup>-1</sup>) together with a 200 ms diffusion delay. The PFG diffusion data were analysed using the Stejskal–Tanner equation [32]. The diffusion coefficient obtained from the PFG



**Figure 1** Relative enrichment of individual ribosomal proteins in the aMBF1 affinity purification experiment from *S. solfataricus* strain P2 cell lysate

Half of the  $^{15}\text{N}$ -labelled eluate from an aMBF1 affinity purification experiment was mixed with 4 pmol of purified  $^{14}\text{N}$ -labelled ribosomes. Proteins of the 30S ribosomal subunits are represented by closed circles, and proteins of the 50S ribosomal subunits by open circles. Mascot probability scores are calculated as  $-10 \times \log_{10}(P)$ , where  $P$  is the probability that the match was a random event.

data and the rotational correlation time obtained from  $^{15}\text{N}$ -relaxation data were used to approximate an apparent molecular mass using the Stokes–Einstein and Stokes–Einstein–Debye equations respectively, using a sphere model and a 2 Å (0.2 nm) hydration layer [34]. All spectra were processed in nmrPipe [35]. The backbone resonance assignment of the C-terminal aMBF1 resonances was performed via standard triple-resonance (HNCACB, HNCOCACB, HNCO and HNCACO) experiments, processed in nmrPipe and analysed using CCPN software [36]. The secondary-structure populations were calculated from the chemical shifts using the  $\delta 2\text{D}$  method [37]. A structural model was determined using CA, CB, CO, N and HN chemical shift with the CHESHIRE fragment replacement protocol [38]. The ensemble has been generated using the chemical shifts as restraints in a replica-averaged MD simulation [39,40] using the Amber03W force field [41] (see the Supplementary Online Data at <http://www.biochemj.org/bj/462/bj4620373add.htm> for details). Residual dipolar couplings were measured using the difference in the  $^{15}\text{N}$ -coupling of HSQC-IPAP (in-phase/antiphase) spectra [42] recorded on an anisotropic sample and a sample of MBF1 partially aligned in 5% penta(ethylene glycol) dodecyl ether/hexanol [43]. The correlation spectra of [ $^{15}\text{N}$ ]aMBF1 in complex with archaeal ribosomal subunits were detected via a SOFAST-HMQC acquisition scheme [44].

## RESULTS AND DISCUSSION

### Heterologous production of *S. solfataricus* aMBF1

An aMBF1 orthologue was identified in *S. solfataricus*, comprising the archaea-specific N-terminal putative zinc-ribbon domain and the C-terminal HTH domain [2,15]. Recombinant *S. solfataricus* aMBF1 was successfully produced as a His-tag fusion protein in *E. coli* Rosetta (DE3) cells, but expression in *E. coli* strain BL21(DE3) was found to yield an apparently uniformly truncated product; analysis of this truncated product by tryptic digest and MS identified peptides covering the entire C-terminal domain from Lys<sup>59</sup> onwards (results not shown).

Constructs for the expression of the isolated N-terminal (amino acids 1–58) and C-terminal (amino acids 57–165) domains (aMBF1-N and aMBF1-C respectively) were then designed according to these partial proteolysis data and the mutant proteins were stably produced in *E. coli* Rosetta (DE3) cells (Supplementary Figure S2 at <http://www.biochemj.org/bj/462/bj4620373add.htm>). MS determined the mass of aMBF1-C to be  $13161 \pm 1$  Da, corresponding precisely to the calculated mass of the protein with an N-terminal proteolytic processing of the first three amino acids (the start methionine residue, Arg<sup>57</sup> and Lys<sup>58</sup>; numbering according to full-length MBF1) (Supplementary Figure S3 at <http://www.biochemj.org/bj/462/bj4620373add.htm>). N-terminal processing of aMBF1-C was also observed by NMR spectroscopy (see below).

### Affinity purification of interacting proteins

High-throughput screens for protein complexes by tandem affinity purification have not previously identified any complex that included MBF1 in yeast [45]. We screened for possible protein interactors of aMBF1 in an attempt to gain information on its possible physiological function. To identify transient protein–protein interactions in our screen, we combined a single-step affinity chromatography procedure with quantitative MS. *S. solfataricus* P2 cells were grown in  $^{14}\text{N}$ - or  $^{15}\text{N}$ -containing medium yielding isotopically labelled proteins. P2 cell lysate was mixed with immobilized aMBF1 and affinity-purified proteins were compared with a negative control experiment carried out using cell lysate with different isotopic labelling and immobilized His<sub>6</sub> peptide instead of aMBF1. The eluate fractions from the two different experiments were mixed, resolved by SDS/PAGE and subjected to MS, allowing the determination of the relative levels of each identified protein in the two experiments. Proteins that were enriched at least 10-fold using immobilized aMBF1 when compared with the control experiment were considered to be possible interactors of aMBF1. Intriguingly, the majority of the potential interactors (nine of the ten) were proteins from the 30S ribosomal subunit, and, in addition, a single protein from the 50S ribosomal subunit could be confirmed to bind to aMBF1 (Table 1). Comparison between experiments using the reverse  $^{14}\text{N}$ - and  $^{15}\text{N}$ -labelling of *S. solfataricus* P2 cells confirmed the overall trend of the experiment, with an S.D. of  $\pm 10\%$  for the normalized ratios.

The affinity purification indicated binding of aMBF1 to ribosomes or specific ribosomal proteins. In order to investigate further whether specific ribosomal proteins were enriched during the aMBF1 affinity purification, we compared the level of individual ribosomal proteins in the aMBF1 affinity purification with those within purified *S. solfataricus* ribosomes. To do this, the  $^{15}\text{N}$ -labelled affinity purification eluate was mixed with 4 pmol of ribosomes and analysed by MS. Overall, the results suggested a strong preference of aMBF1 for the 30S ribosomal subunit over the 50S ribosomal subunit (Figure 1). The relative proportions of ribosomal proteins of each individual ribosomal subunit were similar for both the affinity-purified material and the purified ribosomes, with three exceptions. Ribosomal proteins S8e, L10p and L7ae were significantly overrepresented in the affinity-purification fractions compared with the purified ribosomes. L7ae has been shown to be part of not only the 50S, but also of the 30S ribosomal subunits in Archaea [46,47] explaining why it was enriched to a similar extent as the other 30S ribosomal proteins. The data suggested that, in the experiment, intact 30S ribosomal subunits interacted directly with the immobilized aMBF1. There appeared to be no evidence to suggest that an additional extrinsic factor mediates this interaction nor that there was a subset of ribosomal proteins enriched that maps to a specific part of the

**Table 1** Proteins identified in the aMBF1 affinity purification from *S. solfataricus* strain P2 cell lysate

Mascot probability scores are calculated as  $-10 \times \log_{10}(P)$ , where  $P$  is the probability that the match was a random event. Specific interactors (sp.) were defined as those proteins being at least 10-fold enriched in the aMBF1 affinity purification compared with the control experiment. non-sp. indicates that the enrichment was below threshold, n.q. indicates that a quantification was not possible.

Protein name	Identifier	Mascot score	Specificity
Proteins of the 30S ribosomal subunit			
S2p	15897033	206	sp.
S3p	15897617	47	n.q.
S4p	15897040	116	sp.
S5p	15897604	306	sp.
S7p	15897165	94	n.q.
S8p	15897609	76	sp.
S9p	15897035	149	sp.
S10p	15897163	71	non-sp.
S11p	15897039	67	n.q.
S12p	15897167	104	n.q.
S13p	15897041	84	n.q.
S17p	15897615	32	n.q.
S19p	15897619	34	n.q.
S4e	15897612	130	sp.
S6e	15897344	269	sp.
S8e	15897114	80	n.q.
S19e	15897290	48	sp.
S24e	15897365	174	sp.
S25e	15897354	48	n.q.
Proteins of the 50S ribosomal subunit			
L1p	15897279	51	n.q.
L4p	15897622	140	sp.
L5p	15897611	31	n.q.
L6p	15897608	99	non-sp.
L11p	15897280	38	n.q.
L18p	15897605	30	n.q.
L30	15897603	39	n.q.
Non-ribosomal proteins			
Thermosome $\beta$	15897225	69	non-sp.
Alba	15897841	91	non-sp.

ribosome. It should be noted that the experiment was carried out under buffer conditions in which the ribosomal particles are intact. The C-terminal HTH of yeast MBF1 has been shown to interact with the general transcription factor TBP in GST-pull-down assays with purified recombinant proteins [4,5]. Interestingly, we did not detect TBP as a possible interactor for aMBF1 in our screen.

### Expression and cellular localization of aMBF1

The expression levels of aMBF1 during different growth phases of *S. solfataricus* were determined by immunodetection. aMBF1 was expressed during all growth phases, but expression was observed to be at its highest during exponential growth (Figure 2A). On the basis of a concentration calibration using recombinant aMBF1, the expression levels were calculated to range from  $320 \pm 60$  ng of aMBF1/mg of cytosolic proteins in the exponential growth phase to  $90 \pm 30$  ng of aMBF1/mg of cytosolic proteins in the stationary growth phase.

We also investigated the interaction of endogenous aMBF1 with ribosomal subunits by isolating ribosomes and testing co-purification of aMBF1. When ribosomes were pelleted by ultracentrifugation from an S30 extract prepared from cells in late-exponential growth phase, a significant fraction of aMBF1 was depleted from the S100 supernatant (Figure 2B), suggesting that

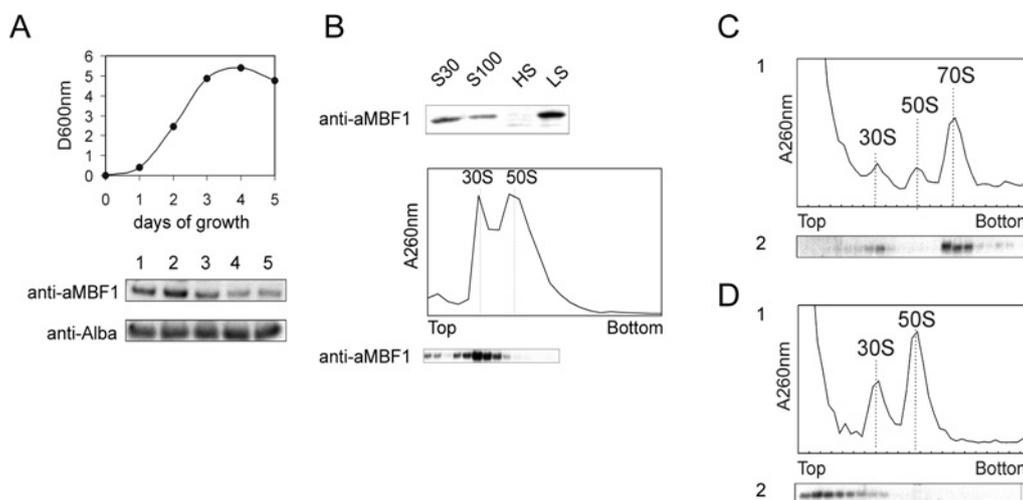
it co-purifies with the ribosomal pellet. Further purification of the crude ribosome fraction on sucrose cushions under low (100 mM) and high (500 mM) salt conditions revealed that the association of MBF1 is salt-sensitive. Loading of crude ribosomes on sucrose density gradients revealed that aMBF1 co-migrated specifically with the 30S, but not with the 50S ribosomal subunit.

Another question was whether aMBF1 also associates with 70S ribosomes. 70S ribosomes of *S. solfataricus* are relatively unstable and the intact particle readily dissociates into 30S and 50S ribosomal subunits during sucrose density gradient centrifugation. The entire ribosomal complex can be reassembled by programming a cell lysate for translation in an *in vitro* translation assay using the well-characterized *orf104* mRNA template (Supplementary Figure S1) and subsequently stabilizing the 70S ribosomes by chemical cross-linking [23]. This procedure can potentially also cross-link aMBF1 directly to the ribosomes. Under these conditions, only a minor fraction of free aMBF1 was found, and most aMBF1 co-migrated with 70S ribosomes and isolated 30S ribosomal subunits (Figure 2C). Similar results were obtained using *aIF6* mRNA as template, indicating that aMBF1–30S ribosomal subunit interaction is not limited to specific mRNAs (results not shown).

The cell lysates used in these *in vitro* translation assays were pre-incubated at high temperature before their use in the translation assay in order to unload the ribosomes from endogenous mRNA and to increase the specificity of the translation reaction for the recombinant mRNA provided. We also tested whether aMBF1 is associated with ribosomal subunits in the pre-incubated lysates without the subsequent *in vitro* translation reaction and chemical cross-linking. Under these conditions, aMBF1 was spread over many fractions in the upper third of the gradient down to the position of the 30S ribosomal subunit. This possibly indicates that aMBF1 dissociated during centrifugation from the 30S ribosomal subunit (Figure 2D).

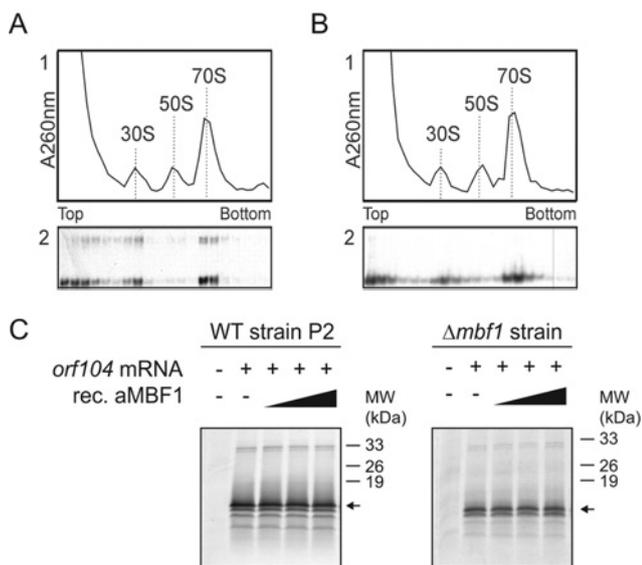
Similarly to endogenous aMBF1, recombinant aMBF1 co-migrated with 70S ribosomes and free 30S ribosomal subunit when 50 pmol of recombinant aMBF1 was added to a cell lysate prepared from a  $\Delta mbf1$  strain before programming for translation (Figure 3A). The *S. solfataricus*  $\Delta mbf1$  strain did not reveal any significant differences in growth kinetics from wild-type strain P2 under standard laboratory conditions (B. de Koning and J. van der Oost, unpublished work). Immunodetection of recombinant aMBF1 gave the expected signals at  $\sim 20$  kDa corresponding to aMBF1 monomer and a minor fraction at  $\sim 40$  kDa. The reason for this apparent dimerization observed only for recombinant full-length aMBF1 is unknown, but it might involve (re-)oxidation of the cysteine residues in the N-terminal domain despite the presence of reducing agents throughout all experiments. Supplementation of up to 300 nM recombinant aMBF1 to cell lysate from *S. solfataricus* strain P2 (exceeding more than three times the concentration of endogenous aMBF1 being present in the assay) or the  $\Delta mbf1$  strain in *in vitro* translation assays did not affect protein synthesis from the *orf104* mRNA template (Figure 3C).

To investigate which domain of aMBF1 is responsible for the binding to ribosomes, we complemented a cell lysate of a  $\Delta mbf1$  strain with the domain deletion variants aMBF1-C and aMBF1-N. Deletion of the predicted N-terminal zinc-ribbon domain of aMBF1 did not affect the ribosome interaction (Figure 3B), suggesting that the HTH domain (aMBF1-C) is sufficient to mediate the interaction with the 30S ribosomal subunit. In experiments carried out under the same conditions with aMBF1-N, the protein was not detectable (results not shown), probably due to degradation during the high-temperature *in vitro* translation reaction.



**Figure 2** Expression of endogenous aMBF1 during different growth phases of *S. solfataricus* strain P2, co-purification with ribosomal subunits and binding of aMBF1 to 70S ribosomes in cross-linked cell lysates after activation for translation

(A) Expression of endogenous aMBF1 during different growth phases of *S. solfataricus* P2 as detected by immunodetection. Upper panel: representative growth curve of a *S. solfataricus* P2 culture. Lower panel: immunodetection of aMBF1 and the abundant nucleic acid-binding protein Alba as control for samples taken at the given time points. Equal total soluble protein content was loaded on each lane. (B) Co-purification of aMBF1 with *S. solfataricus* P2 ribosomes. Upper panel: immunodetection of aMBF1 in different fractions obtained from ribosome isolation. S30 and S100 extracts (5% of total fraction respectively) and ribosomes purified on sucrose cushions under low (100 mM) (LS) and high (500 mM) (HS) salt conditions were tested. Lower panel: fractionation of 400 pmol of crude ribosomes (on the basis of absorption at 260 nm) on a 15–30% sucrose gradient (30 mM KCl) to separate the ribosomal subunits and to verify the co-fractionation of aMBF1 with ribosomes. (C and D) Localization of endogenous aMBF1 in *S. solfataricus* P2 cell lysate. (C) Cell lysate (480  $\mu$ g of protein content) was pre-incubated at 73 °C to unload the ribosomes from any endogenous mRNA and subsequently the cell lysate was programmed for translation with *orf104* mRNA as template and followed by formaldehyde cross-linking to stabilize 70S ribosomes. The cross-linked *in vitro* translation reaction was fractionated on a 10–30% sucrose density gradient, and immunodetection was used to localize aMBF1. (D) A control experiment with the same amount of cell lysate, but without subsequent *in vitro* translation and cross-linking. (C and D) Upper panels (1): absorption at 260 nm of the fractions after sucrose density gradient centrifugation to identify the position of the ribosomal subunits. Lower panels (2): immunodetection of aMBF1.



**Figure 3** Localization of recombinant aMBF1 and aMBF1-C in cell lysates programmed for translation

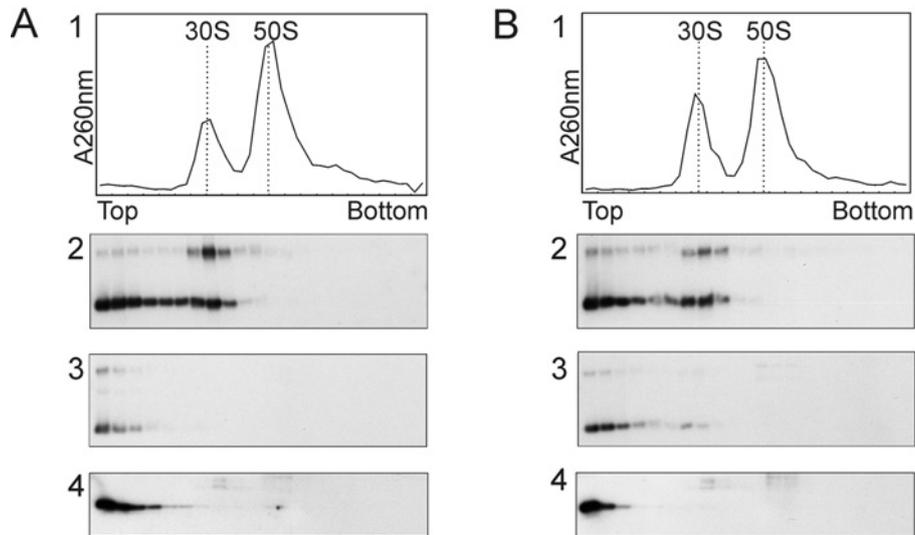
(A and B) *In vitro* translation assays with cell lysate from the  $\Delta mbf1$  strain (480  $\mu$ g of protein content) were supplemented with 50 pmol of recombinant aMBF1 (A) and aMBF1-C (B). The lysates were pre-incubated, programmed for translation, chemically cross-linked and fractionated on sucrose density gradient as described in Figure 2. Upper panels (1): absorption profile at 260 nm of the sucrose density gradient was measured to identify the position of the ribosomal subunits. Lower panels (2): immunodetection of recombinant aMBF1 in the fractions obtained from the sucrose density gradient. (C) Effect of increasing amounts of recombinant aMBF1 (0, 3, 30 and 300 nM) on cell-free translation. *In vitro* translation was carried out in the presence of [<sup>35</sup>S]methionine using lysate from *S. solfataricus* P2 or the  $\Delta mbf1$  strain and production of ORF104 was detected by autoradiography after resolving the samples by Tris/Tricine SDS/PAGE. The first lane shows an *in vitro* translation reaction without an mRNA template. Molecular masses (MW) are indicated in kDa.

### Reconstituted complexes

To determine whether aMBF1 directly interacts with the small ribosomal subunit, we purified ribosomes at different salt concentrations from cells grown to early stationary phase. Levels of co-purified endogenous aMBF1 were below the detection limit for ribosomes purified both under low (100 mM  $\text{NH}_4\text{Cl}$ ) and high (500 mM  $\text{NH}_4\text{Cl}$ ) salt conditions. Recombinant aMBF1 was added to the purified 30S subunits and the formation of the 30S–aMBF1 complex was observed using sucrose density gradient and immunodetection of aMBF1 in the 30S elution fractions. When incubated with both low- and high-salt-washed ribosomes, recombinant aMBF1 co-fractionated with 30S, indicating that no cofactor was apparently required to form the complex (Figure 4). A significant proportion of aMBF1 eluted in the low-molecular-mass fractions of the sucrose gradient, suggesting that a fraction of the aMBF1–30S ribosomal subunit complex dissociated during the centrifugation (Figures 4A and 4B, panels 2). Surprisingly, when the experiment was repeated using only the C-terminal HTH domain of aMBF1 (aMBF1-C), only a small fraction of aMBF1-C remained bound to the 30S ribosomal subunits after sucrose density gradient centrifugation (Figures 4A and 4B, panels 3). This might be due to dissociation of the aMBF1-C small ribosomal subunits complex during sucrose density gradient centrifugation. The lower stability of the aMBF1-C small ribosomal subunits complex compared with that formed with full-length aMBF1 might indicate a role for the predicted zinc-ribbon domain or parts of the linker region to increase the complex affinity.

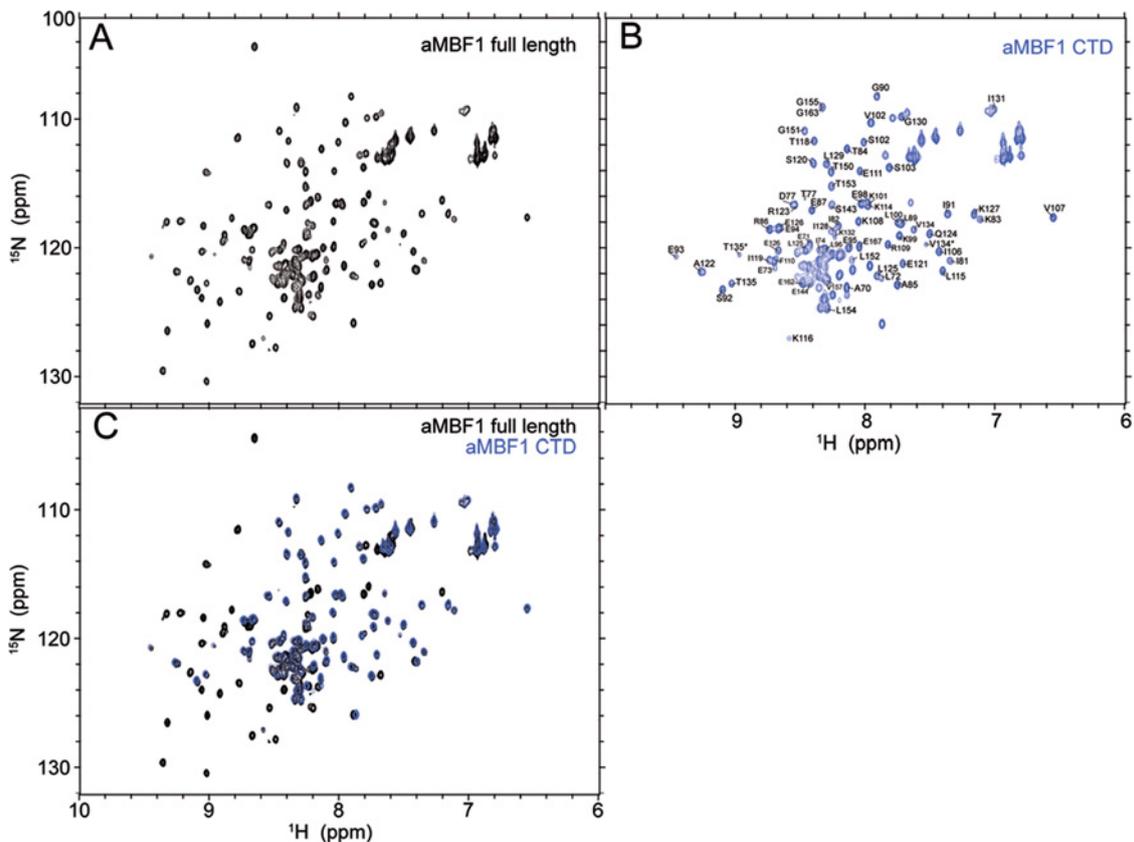
### Structural analysis of isolated aMBF1 and aMBF1-C

Having established that aMBF1 binds directly to the small ribosomal subunits without cofactor, we aimed to characterize the structural and dynamic features of this interaction using



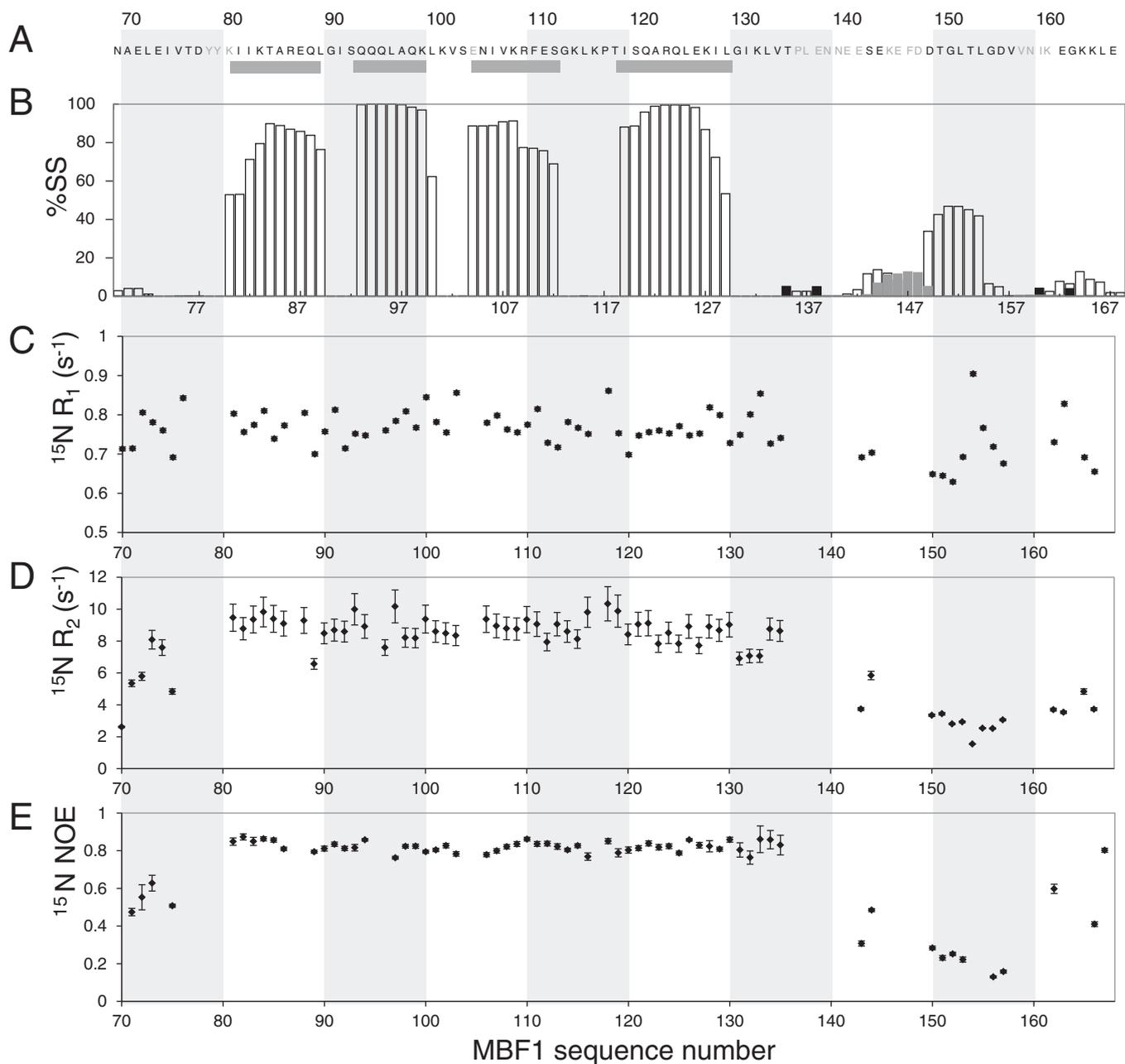
**Figure 4** Binding of aMBF1, aMBF1-N and aMBF1-C to purified ribosomal subunits

Ribosomes (100 pmol) purified under low-salt (100 mM) (A) or high-salt (500 mM) (B) conditions were incubated with 100 pmol of recombinant aMBF1 or mutant proteins and incubated on ice for 30 min. The samples were subsequently resolved on 10–30% sucrose gradients and SDS/PAGE, and immunodetection was used to localize aMBF1 and the mutant proteins. Panels 1: representative  $A_{260}$  profile with the position of the ribosomal subunits. Panels 2–4: immunodetection of binding assays for aMBF1 (panels 2), aMBF1-C (panels 3) and aMBF1-N (panels 4). The immunodetection of recombinant aMBF1 gave a second minor band approximately twice the apparent molecular mass of monomeric aMBF1. The reason for this apparent aMBF1 dimerization observed only for recombinant full-length aMBF1 is unknown, but it might involve (re-)oxidation of the cysteine residues in the N-terminal domain of aMBF1 despite the presence of reducing agents throughout all experiments.



**Figure 5**  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC of full-length aMBF1 (A) and aMBF1-C (B) and overlay of the spectra (C)

The resonances associated with the *cis* isomeric state of the Thr<sup>135</sup>–Pro<sup>136</sup> peptide bond are labelled with asterisks (\*). The assignment of the aMBF1-C residues is deposited in the BMRB under entry accession number 19028. (C) Overlay of the spectra of (A) full-length aMBF1 (black) and (B) aMBF1-C (blue).



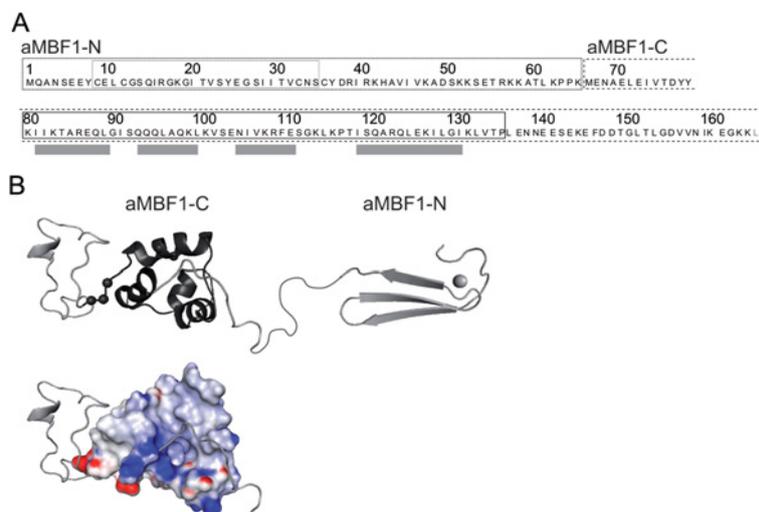
**Figure 6** NMR analysis of the C-terminal HTH domain of aMBF1

(A) Sequence of aMBF1 with the  $\alpha$ -helices as defined by the  $C_{\alpha}$  and  $C'$  chemical shift shown by the grey bars. (B) Secondary-structure populations estimated from the chemical shifts using the  $\delta 2\text{D}$  method [37] ( $\alpha$ -helices are shown as white bars,  $\beta$ -sheet as black bars and polyproline II as grey bars). (C–E)  $^{15}\text{N}$  relaxation parameters:  $R_1$  (C),  $R_2$  (D) and the heteronuclear NOE (E) (see the Experimental section for details).

NMR spectroscopy. A structural analysis of aMBF1 resulted in a well-resolved 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum in which almost all of the expected 169 resonances are clearly discerned: 91 well-dispersed resonances as well as  $\sim 78$  in the central region of the spectrum ( $^1\text{H}$  frequency lying between 7.5 and 8.5 p.p.m.) were observed (Figure 5A).

Since the aMBF1–30S interaction seems to be mediated mainly by the C-terminal HTH domain, the NMR spectrum of aMBF1-C was compared with that of the full-length protein; its spectrum (Figure 5B) was found to overlay very well with the spectrum of the full-length aMBF1 (Figure 5C) indicating that the HTH and

zinc-ribbon domains are structurally independent. The backbone resonances of aMBF1-C were assigned via standard triple-resonance strategies (see the Experimental section) to  $\sim 82\%$  completion (Figure 6A). The secondary-structure populations calculated by the  $\delta 2\text{D}$  method using the CA, CB,  $C'$ , N and HN chemical shifts (Figure 6, see the Experimental section) indicate the presence of four  $\alpha$ -helices at Ile $^{81}$ –Gln $^{88}$ , Gln $^{93}$ –Lys $^{99}$ , Glu $^{104}$ –Glu $^{111}$  and Ile $^{119}$ –Gly $^{130}$  in line with bioinformatics prediction of a tetrahelical bundle HTH domain [1]. Leu $^{133}$  and Val $^{134}$  in the fourth  $\alpha$ -helix as well as residues that are structurally close to that  $\alpha$ -helix (Leu $^{125}$  and Glu $^{126}$ ) give rise to two sets of



**Figure 7** Structural characterization of *S. solfataricus* aMBF1

(A) Amino acid sequence with the aMBF1-N and aMBF1-C domain boundaries shown as broken-line boxes, the structured domains shown by boxes and the four  $\alpha$ -helices of the HTH domain as defined by the  $\delta$ 2D method [37] shown by grey bars (see Figure 2). The C-terminal two residues in grey constitute the linker to the His<sub>6</sub> tag. (B) Top: structural model of aMBF1. A model for the zinc-ribbon motif of aMBF1 was built using homology modelling and the zinc-ribbon structure of *Methanococcus jannaschii* translation initiation factor 2 $\beta$  (PDB code 1K81) as template. The zinc atom is represented as a sphere. A structural model for the HTH domain of aMBF1 was built on the basis of the information contained in the chemical shifts following the procedure implemented in [38]. The intrinsically disordered sequences are represented as lines; an ensemble of structures representing the dynamics of the protein determined using chemical shift restrained MD [39,40] is shown in Supplementary Figure S5 (<http://www.biochemj.org/bj/462/bj4620373add.htm>). The residues that give two sets of resonances associated with the isomerization of the Thr<sup>135</sup>–Pro<sup>136</sup> peptide bond are shown as spheres. Bottom: the electrostatic surface is shown ranging from  $-20$  kT/e (red) to  $+20$  kT/e (blue).

resonances in NMR spectra. The intensity ratio for those two sets of resonances is  $\sim 70\%/30\%$  and probably results from the *cis*–*trans* isomerization of Pro<sup>136</sup> on a long timescale for NMR spectroscopy ( $> 100$  ms). Interestingly, Pro<sup>136</sup> is widely conserved among crenarchaeal aMBF1 sequences (Supplementary Figure S4 at <http://www.biochemj.org/bj/462/bj4620373add.htm>). The consequence of this putative *cis*–*trans* isomerization has not been investigated in detail, but, as shown below, does not seem to affect ribosome binding.

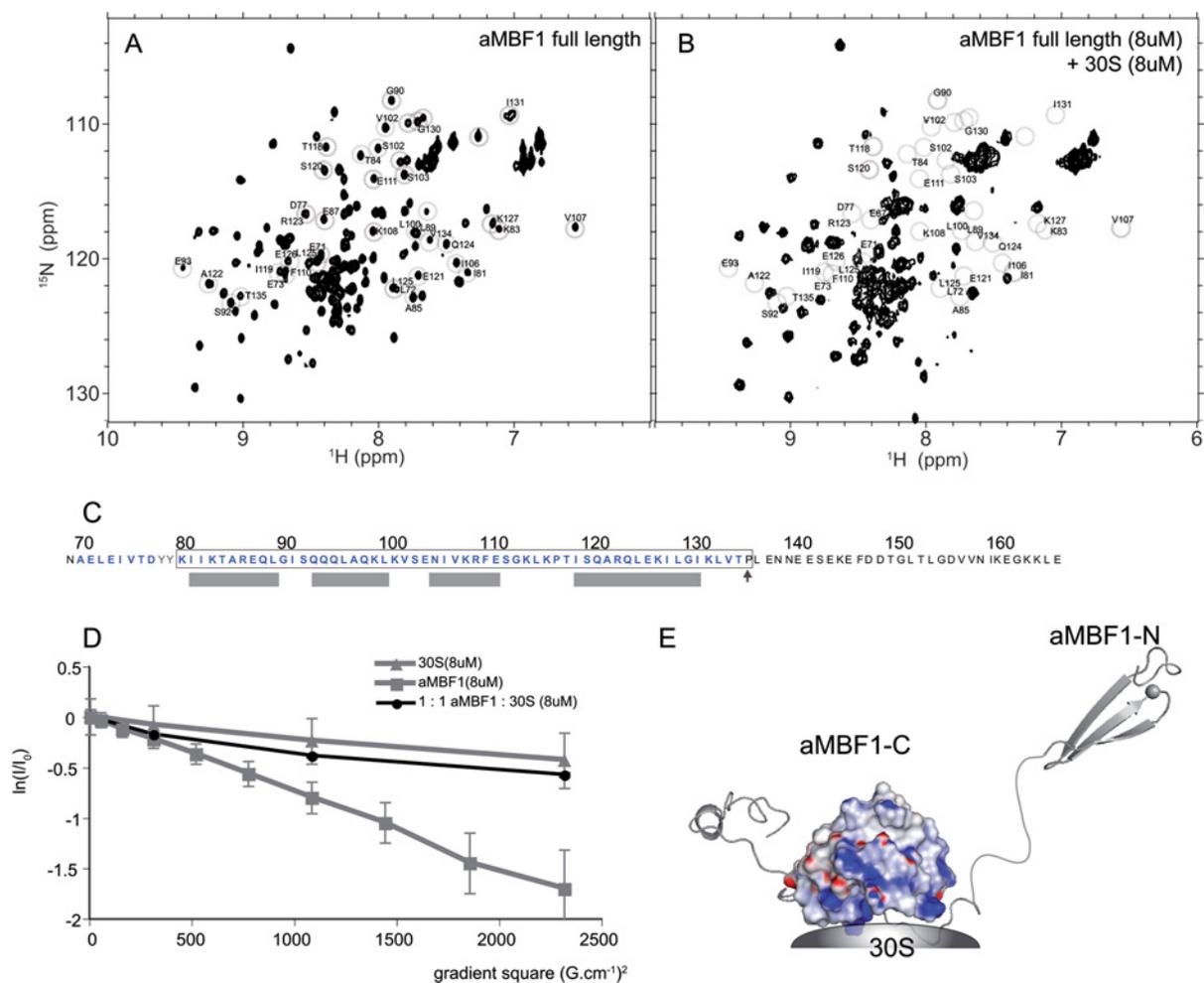
The boundaries of the HTH motif within aMBF1 were investigated further via a series of <sup>15</sup>N NMR measurements of the relaxation parameters ( $T_1$ ,  $T_2$  and  $\{^1\text{H}-^{15}\text{N}\}$ -NOE, see the Experimental section). Relaxation NMR data were obtained for 82 of the 100 residues in aMBF-C (Figure 6). The data are characterized by uniform  $R_1$ ,  $R_2$  and  $\{^1\text{H}-^{15}\text{N}\}$ -NOE values from Ile<sup>81</sup> to Thr<sup>135</sup>. The absence of regions of the polypeptide showing higher  $R_2$  values indicates the likely absence of significant conformational exchange processes on the millisecond timescale. aMBF1-C appears to be a compact domain with a rotational correlation time ( $\tau_c$ ) of  $5.8 \pm 0.2$  ns. Flanking this domain are regions (before Ile<sup>81</sup> and after Thr<sup>135</sup>) characterized by reduced  $\{^1\text{H}-^{15}\text{N}\}$ -NOE values ( $< 0.5$ ), indicating that these regions undergo sub-nanosecond motions. This is characteristic of structurally intrinsically disordered regions. Notably, the C-terminal disordered region is highly conserved in sequence and specific to aMBF1 orthologues [15].

An ensemble of structures representing the structure and dynamics of the HTH domain was then determined using NMR chemical shifts as restraints in MD simulations using the CamShift-MD approach [39,40] (Supplementary Figure S5 at <http://www.biochemj.org/bj/462/bj4620373add.htm>). The resulting structures (PDB code 2MEZ) show the disordered regions flanking the HTH domain, with the electrostatic surfaces indicating the presence of a positively charged region at the N-terminus of the HTH domain (Figure 7, lower panel). In comparison, eukaryotic MBF1 exhibits

a significantly less pronounced positively charged surface [48] (Supplementary Figure S6 at <http://www.biochemj.org/bj/462/bj4620373add.htm>).

### Structural investigation of the interaction of aMBF1 and 30S

In order to clarify the role of the HTH domain in the interaction of aMBF1 with the 30S ribosomal subunit, <sup>1</sup>H–<sup>15</sup>N correlation NMR spectra of <sup>15</sup>N-labelled aMBF1 in the absence and presence of unlabelled 30S ribosomal subunit were recorded. These showed the selective broadening of a highly discrete subset of 38 cross-peaks occurring in the presence of the 30S ribosomal subunit (Figures 8A and 8B), and further analysis showed that these resonances arose from the C-terminal HTH motif (Figure 8C). More specifically, all resonances assigned to residues within Ala<sup>70</sup> to Ile<sup>135</sup> were broadened beyond detection, indicating that the ribosomal interaction is specific and mediated by the HTH domain. The doubled set of resonances for residues neighbouring Pro<sup>136</sup> that were assigned to two isoforms of the isomerization of Pro<sup>136</sup> were both broadened beyond detection, indicating that both isomers bind to the ribosome to the same extent. The remaining resonances observable in the complex were probed using X-STE diffusion NMR methods. A diffusion coefficient of  $(4 \pm 0.2) \times 10^{-11}$  m<sup>2</sup>·s<sup>-1</sup> was determined from these data, a value significantly lower than that observed for the isolated protein [ $(1.3 \pm 0.05) \times 10^{-10}$  m<sup>2</sup>·s<sup>-1</sup>]. Moreover, as the former value is identical with that measured for the <sup>1</sup>H resonances observed for the 30S ribosomal subunit alone (Figure 8D), this appears to reflect the association of aMBF1 to the 30S ribosomal subunit. An NMR titration of <sup>15</sup>N-labelled aMBF1-C to the 30S subunit was then undertaken (Supplementary Figure S7 at <http://www.biochemj.org/bj/462/bj4620373add.htm>) and showed that resonances of free aMBF1-C could be detected only with a 10-fold excess of aMBF1 to 30S (at a concentration of 5  $\mu$ M). This finding suggests that the complex lifetime is at least 10-fold



**Figure 8**  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NMR spectra of aMBF1 in the absence and presence of 30S ribosomal subunits

(A) Full-length aMBF1 ( $^1\text{H}$ - $^{15}\text{N}$  HSQC). (B) Full-length aMBF1 and 30S ribosomal subunits at a 1:1 molar ratio (8  $\mu\text{M}$ ) ( $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC). The cross-peaks marked with open circles were broadened in the presence of 30S ribosomal subunits; these were all assignable to aMBF1-C. (C) Sequence of the C-terminal domain of aMBF1. The folded region determined by NMR is boxed and residues whose resonances are broadened due to the ribosome interaction are shown in blue. The grey residues are not assigned. (D) Translational diffusion NMR measurements of the interactions. Stejskal-Tanner plot [relative NMR signal intensities of the aMBF1-C resonances gradient strengths ( $\text{G}^2\cdot\text{cm}^{-2}$ )] for aMBF1 (grey squares), 30S (grey circles) and aMBF1 in the presence of 30S (black circles). In the presence of 30S particle, the resonances of aMBF1-N are associated with a diffusion coefficient identical with that of the ribosomal subunit, indicating an interaction with the complex. (E) Schematic diagram of the aMBF1-30S complex, in which the interaction is mediated by the positively charged surface of the aMBF1-C domain, and both the C-terminal disordered end of aMBF1-C and the N-terminal domain remains flexible enough to tumble independently from the aMBF1-30S complex.

lower than the NMR acquisition time (50 ms), confirming a highly transient complex (Supplementary Figure S7).

The flanking disordered C-terminal region (from Ser<sup>143</sup> onwards) is, however, observed in the spectrum of 30S-bound aMBF1-C, whereas the linker region at the N-terminal end of the HTH domain is broadened beyond detection. The positively charged surface at the N-terminal hemisphere of the model structure of the HTH domain (Figure 7B, blue) could mediate the interaction with rRNA.

The resonances of the predicted zinc-ribbon domain are only partially broadened in the presence of 30S ribosomal subunit, indicating that the flexible linker remains disordered, resulting in sufficient mobility of the N-terminal domain to tumble independently from the MDa ribosomal complex (Figure 8E). Although the zinc-ribbon domain of *S. solfataricus* aMBF1 does not participate directly in the interaction with 30S ribosomal subunit, sucrose density gradient centrifugation of reconstituted complexes alluded to a contribution of the zinc-ribbon to

30S ribosomal subunit binding (Figure 4). The effect may be indirect through the influence of the zinc-ribbon domain on the conformation of the linker region; however, we cannot rule out some bridging function with an unidentified partner (in analogy to eukaryotic MBF1). The interaction of aMBF1 with 30S ribosomal subunits is thus restricted to the HTH domain and the adjacent part of the linker that are shared with eukaryotic MBF1, whereas the Archaea-specific N-terminal zinc-ribbon domain and C-terminal extension are not involved.

#### aMBF1 deletion does not influence misreading in translation fidelity

Deletion of *mbf1* in yeast affects the rate of ribosomal frameshifting, as well as the sensitivity to the antibiotic paramomycin that targets the 30S ribosomal subunit. The molecular basis for this phenomenon is unknown, but, given

that all other identified suppressors of frameshift mutations in yeast map to core factors of the translation apparatus [tRNAs, EF1 $\alpha$  (translation elongation factor 1 $\alpha$ ) and ribosomal protein S3] [8,49–52], we reasoned that yeast MBF1 possibly interacts directly with the translation machinery as well. To test whether aMBF1 might influence translation fidelity in a similar manner, we made use of the fact that paramomycin induces misreading in *S. solfataricus* cell-free translation systems, whereas, in general, the archaeal translation apparatus is rather insensitive to antibiotics [53–55]. Misreading was measured using *in vitro* translation assays with a synthetic poly(U) RNA template coding for polyphenylalanine. The use of near-cognate tRNA<sup>Leu</sup> instead of cognate tRNA<sup>Phe</sup> was measured as the misincorporation rate of leucine into polyphenylalanine using radiolabelled amino acids. We observed no significant difference in the basic rate of misreading for lysate of the  $\Delta mbf1$  strain or its parental strain PBL2025, and misreading increased to the same extent in response to paramomycin for both strains (Supplementary Figure 8A at <http://www.biochemj.org/bj/462/bj4620373add.htm>). Furthermore, when a cell lysate programmed for translation was supplemented with 100  $\mu$ M paramomycin, aMBF1 was still mostly associated with 70S ribosomes, suggesting that paramomycin does not compete with aMBF1 for its ribosome-binding site (Supplementary Figure 8B).

## Conclusions

We have shown that aMBF1 from *S. solfataricus* interacts with the 30S ribosomal subunit. aMBF1 is composed structurally of two domains, which have independent mobility. The ribosome-binding interface is potentially the positively charged surface that we identified at the N-terminal hemisphere of the HTH domain of aMBF1. This binding interface is likely to be affected by the conformational sampling of the linker region in the presence of the zinc-ribbon aMBF1-N domain. Within the aMBF1–ribosome complex, the aMBF1 zinc-ribbon-binding domain maintains very high mobility, suggesting that it remains accessible for potential interaction with a third binding partner. Our data suggest a function for aMBF1 related to translation, perhaps as a recruitment factor of the translation apparatus by bridging different ligands, analogous to the function of eukaryotic MBF1 in transcription regulation. aMBF1 does not inhibit translation and its expression is highest during exponential growth. This may indicate that the role of aMBF1 in protein synthesis is not related to the stress response. Our understanding of the molecular function of aMBF1 will be advanced further by the identification of potential ligands that may bind to the N-terminal zinc-ribbon domain. Archaeal MBF1 and eukaryotic MBF1 are orthologous proteins and hence it has been proposed that aMBF1 functions as a core transcription factor in archaea as well [2]. We provide evidence that aMBF1 interacts physically with the translation machinery and it is likely that aMBF1 carries out a function related to the translation process. Interestingly, there is accumulating evidence that also the eukaryotic MBF1 carries out additional function(s) beyond transcription initiation. It is worth noting that, since the conserved HTH of aMBF1 mediates the interaction with the 30S ribosomal subunit, eukaryotic MBF1 might bind similarly to the 40S ribosomal subunit. The recent findings that MBF1 associates with polyadenylated mRNA in different eukaryotic species [11–13] cannot be explained by the known function of MBF1 in transcription initiation, but it would be compatible with a conserved interaction of MBF1 and ribosomes in archaea and eukaryotes.

## AUTHOR CONTRIBUTION

Fabian Blomach, Helene Launay, Ambrosius Snijders, Hao Wu, Bart de Koning, Michele Vendruscolo, Martin Dickman, Lisa Cabrera, Anna La Teana, Dario Benelli, Paola Londei, John Christodoulou and John van der Oost designed the experiments. Fabian Blomach, Helene Launay, Ambrosius Snijders, Violeta Zorraquino, Hao Wu, Bart de Koning, Carlo Camilloni, Aandrea Cavalli and Anna La Teana performed the experiments. Fabian Blomach, Helene Launay, Stan Brouns, Thijs Ettema, John Christodoulou and John van der Oost wrote the paper.

## ACKNOWLEDGEMENTS

We acknowledge the assistance of the staff and the use of the MRC Biomedical NMR Centre, Mill Hill, for the collection of the NMR data on isolated C-terminal aMBF1 as well as Dr John Kirkpatrick of the UCL Biological NMR Facility for technical support.

## FUNDING

This work was supported by the Netherlands Organization for Scientific Research (NWO) (VICI-grant number 865.05.001 to (J.v.d.O.) and VIDI-grant number 864.11.005 (to S.J.J.B.)), a Wellcome Trust New Investigator Award (to J.C.), the Biotechnology and Biological Sciences Research Council [grant numbers 9015651/JC (to J.C.) and BB/D011795/1 (to M.J.D.)], the Engineering and Physical Sciences Research Council [grant number EP/E036252/1 (to M.J.D.)], Istituto Pasteur-Fondazione Cenci Bolognietti [‘From leaderless to leadered mRNA: mRNA features modulating ribosome binding and translation initiation in Archaea’ project (to P.L.)], the European Molecular Biology Organization [short-term fellowship ASTF304-2008 (to F.B.)] and a UCL Countess of Lisburne Scholarship for Doctoral Research (to H.L.).

## REFERENCES

- Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M. and Iyer, L. M. (2005) The many faces of the helix–turn–helix domain: transcription regulation and beyond. *FEMS Microbiol. Rev.* **29**, 231–262 [PubMed](#)
- Aravind, L. and Koonin, E. V. (1999) DNA-binding proteins and evolution of transcription regulation in the archaea. *Nucleic Acids Res.* **27**, 4658–4670 [CrossRef PubMed](#)
- Blombach, F., Makarova, K. S., Marrero, J., Siebers, B., Koonin, E. V. and van der Oost, J. (2009) Identification of an ortholog of the eukaryotic RNA polymerase III subunit RPC34 in Crenarchaeota and Thaumarchaeota suggests specialization of RNA polymerases for coding and non-coding RNAs in Archaea. *Biol. Direct* **4**, 39 [CrossRef PubMed](#)
- Takemaru, K., Harashima, S., Ueda, H. and Hirose, S. (1998) Yeast coactivator MBF1 mediates GCN4-dependent transcriptional activation. *Mol. Cell. Biol.* **18**, 4971–4976 [PubMed](#)
- Liu, Q. X., Nakashima-Kamimura, N., Ikeo, K., Hirose, S. and Gojobori, T. (2007) Compensatory change of interacting amino acids in the coevolution of transcriptional coactivator MBF1 and TATA-box-binding protein. *Mol. Biol. Evol.* **24**, 1458–1463 [CrossRef PubMed](#)
- Li, F. Q., Ueda, H. and Hirose, S. (1994) Mediators of activation of fushi tarazu gene transcription by BmFTZ-F1. *Mol. Cell. Biol.* **14**, 3013–3021 [CrossRef PubMed](#)
- Kabe, Y., Goto, M., Shima, D., Imai, T., Wada, T., Morohashi, K., Shirakawa, M., Hirose, S. and Handa, H. (1999) The role of human MBF1 as a transcriptional coactivator. *J. Biol. Chem.* **274**, 34196–34202 [CrossRef PubMed](#)
- Hendrick, J. L., Wilson, P. G., Edelman, II, Sandbaken, M. G., Ursic, D. and Culbertson, M. R. (2001) Yeast frameshift suppressor mutations in the genes coding for transcription factor Mbf1p and ribosomal protein S3: evidence for autoregulation of S3 synthesis. *Genetics* **157**, 1141–1158 [PubMed](#)
- Costanzo, M. C., Mueller, P. P., Strick, C. A. and Fox, T. D. (1986) Primary structure of wild-type and mutant alleles of the PET494 gene of *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **202**, 294–301 [CrossRef PubMed](#)
- Culbertson, M. R., Gaber, R. F. and Cummins, C. M. (1982) Frameshift suppression in *Saccharomyces cerevisiae*. V. Isolation and genetic properties of nongroup-specific suppressors. *Genetics* **102**, 361–378 [PubMed](#)
- Klass, D. M., Scheibe, M., Butter, F., Hogan, G. J., Mann, M. and Brown, P. O. (2013) Quantitative proteomic analysis reveals concurrent RNA–protein interactions and identifies new RNA-binding proteins in *Saccharomyces cerevisiae*. *Genome Res.* **23**, 1028–1038 [CrossRef PubMed](#)
- Baltz, A. G., Munschauer, M., Schwanhausser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M. et al. (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* **46**, 674–690 [CrossRef PubMed](#)

- 13 Kwon, S. C., Yi, H., Eichelbaum, K., Fohr, S., Fischer, B., You, K. T., Castello, A., Krijgsveld, J., Hentze, M. W. and Kim, V. N. (2013) The RNA-binding protein repertoire of embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1122–1130 [CrossRef PubMed](#)
- 14 Marrero Coto, J., Ehrenhofer-Murray, A. E., Pons, T. and Siebers, B. (2011) Functional analysis of archaeal MBF1 by complementation studies in yeast. *Biol. Direct* **6**, 18 [CrossRef PubMed](#)
- 15 de Koning, B., Blombach, F., Wu, H., Brouns, S. J. and van der Oost, J. (2009) Role of multiprotein bridging factor 1 in archaea: bridging the domains? *Biochem. Soc. Trans.* **37**, 52–57 [CrossRef PubMed](#)
- 16 She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C., Clausen, I. G., Curtis, B. A., De Moors, A. et al. (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7835–7840 [CrossRef PubMed](#)
- 17 Gill, S. C. and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319–326 [CrossRef PubMed](#)
- 18 Blombach, F., Launay, H., Zorraquino, V., Swarts, D. C., Cabrera, L. D., Benelli, D., Christodoulou, J., Londei, P. and van der Oost, J. (2011) An HflX-type GTPase from *Sulfolobus solfataricus* binds to the 50S ribosomal subunit in all nucleotide-bound states. *J. Bacteriol.* **193**, 2861–2867 [CrossRef PubMed](#)
- 19 Schagger, H. and von Jagow, G. (1987) Tricine–sodium dodecyl sulfate–polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368–379 [CrossRef PubMed](#)
- 20 Schelert, J., Dixit, V., Hoang, V., Simbahan, J., Drozda, M. and Blum, P. (2004) Occurrence and characterization of mercury resistance in the hyperthermophilic archaeon *Sulfolobus solfataricus* by use of gene disruption. *J. Bacteriol.* **186**, 427–437 [CrossRef PubMed](#)
- 21 Zaparty, M., Esser, D., Gertig, S., Haferkamp, P., Kouril, T., Manica, A., Pham, T. K., Reimann, J., Schreiber, K., Sierocinski, P. et al. (2009) “Hot standards” for the thermoacidophilic archaeon *Sulfolobus solfataricus*. *Extremophiles* **14**, 119–142 [CrossRef PubMed](#)
- 22 Snijders, A. P., Walther, J., Peter, S., Kinnman, I., de Vos, M. G., van de Werken, H. J., Brouns, S. J., van der Oost, J. and Wright, P. C. (2006) Reconstruction of central carbon metabolism in *Sulfolobus solfataricus* using a two-dimensional gel electrophoresis map, stable isotope labelling and DNA microarray analysis. *Proteomics* **6**, 1518–1529 [CrossRef PubMed](#)
- 23 Benelli, D. and Londei, P. (2007) *In vitro* studies of archaeal translational initiation. *Methods Enzymol.* **430**, 79–109 [CrossRef PubMed](#)
- 24 Nguyen, T. N. and Goodrich, J. A. (2006) Protein–protein interaction assays: eliminating false positive interactions. *Nat. Methods* **3**, 135–139 [CrossRef PubMed](#)
- 25 Puhler, G., Lottspeich, F. and Zillig, W. (1989) Organization and nucleotide sequence of the genes encoding the large subunits A, B and C of the DNA-dependent RNA polymerase of the archaeobacterium *Sulfolobus acidocaldarius*. *Nucleic Acids Res.* **17**, 4517–4534 [CrossRef PubMed](#)
- 26 Benelli, D., Marzi, S., Mancone, C., Alonzi, T., la Teana, A. and Londei, P. (2009) Function and ribosomal localization of aIF6, a translational regulator shared by archaea and eukarya. *Nucleic Acids Res.* **37**, 256–267 [CrossRef PubMed](#)
- 27 Condo, I., Ciammaruconi, A., Benelli, D., Ruggero, D. and Londei, P. (1999) *Cis*-acting signals controlling translational initiation in the thermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.* **34**, 377–384 [CrossRef PubMed](#)
- 28 Londei, P., Teixido, J., Acca, M., Cammarano, P. and Amils, R. (1986) Total reconstitution of active large ribosomal subunits of the thermoacidophilic archaeobacterium *Sulfolobus solfataricus*. *Nucleic Acids Res.* **14**, 2269–2285 [CrossRef PubMed](#)
- 29 Mori, S., Abeygunawardana, C., Johnson, M. O. and van Zijl, P. C. (1995) Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. *J. Magn. Reson. B* **108**, 94–98 [CrossRef PubMed](#)
- 30 Peng, J. W. and Wagner, G. (1994) Investigation of protein motions via relaxation measurements. *Methods Enzymol.* **239**, 563–596 [CrossRef PubMed](#)
- 31 Lipari, G. and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.* **104**, 4559–4570 [CrossRef](#)
- 32 Stejskal, E. O. and Tanner, J. E. (1965) Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J. Chem. Phys.* **42**, 288–293 [CrossRef](#)
- 33 Ferrage, F., Zoonens, M., Warschawski, D. E., Popot, J. L. and Bodenhausen, G. (2003) Slow diffusion of macromolecular assemblies by a new pulsed field gradient NMR method. *J. Am. Chem. Soc.* **125**, 2541–2545 [CrossRef PubMed](#)
- 34 Cavanagh, J., Fairbrother, W. J., Palmer III, A. G., Rance, M. and Skelton, N. J. (2007) *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, London
- 35 Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 [PubMed](#)
- 36 Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J. and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 [CrossRef PubMed](#)
- 37 Camilloni, C., De Simone, A., Vranken, W. F. and Vendruscolo, M. (2012) Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry* **51**, 2224–2231 [CrossRef PubMed](#)
- 38 Cavalli, A., Salvatella, X., Dobson, C. M. and Vendruscolo, M. (2007) Protein structure determination from NMR chemical shifts. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 9615–9620 [CrossRef PubMed](#)
- 39 Camilloni, C., Robustelli, P., De Simone, A., Cavalli, A. and Vendruscolo, M. (2012) Characterization of the conformational equilibrium between the two major substates of RNase A using NMR chemical shifts. *J. Am. Chem. Soc.* **134**, 3968–3971 [CrossRef PubMed](#)
- 40 Camilloni, C., Cavalli, A. and Vendruscolo, M. (2013) Assessment of the use of NMR chemical shifts as replica-averaged structural restraints in molecular dynamics simulations to characterize the dynamics of proteins. *J. Phys. Chem. B* **117**, 1838–1843 [CrossRef PubMed](#)
- 41 Best, R. B. and Mittal, J. (2010) Protein simulations with an optimized water model: cooperative helix formation and temperature-induced unfolded state collapse. *J. Phys. Chem. B* **114**, 14916–14923 [CrossRef PubMed](#)
- 42 Ottiger, M., Delaglio, F. and Bax, A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.* **131**, 373–378 [CrossRef PubMed](#)
- 43 Rückert, M. and Otting, G. (2000) Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J. Am. Chem. Soc.* **122**, 7793–7797 [CrossRef](#)
- 44 Schanda, P., Kupce, E. and Brutscher, B. (2005) SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J. Biomol. NMR* **33**, 199–211 [CrossRef PubMed](#)
- 45 Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A. P. et al. (2006) Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **440**, 637–643 [CrossRef PubMed](#)
- 46 Armache, J. P., Anger, A. M., Marquez, V., Franckenberg, S., Frohlich, T., Villa, E., Berninghausen, O., Thomm, M., Arnold, G. J., Beckmann, R. et al. (2013) Promiscuous behaviour of archaeal ribosomal proteins: implications for eukaryotic ribosome evolution. *Nucleic Acids Res.* **41**, 1284–1293 [CrossRef PubMed](#)
- 47 Marquez, V., Frohlich, T., Armache, J. P., Sohmen, D., Donhofer, A., Mikolajka, A., Berninghausen, O., Thomm, M., Beckmann, R., Arnold, G. J. et al. (2011) Proteomic characterization of archaeal ribosomes reveals the presence of novel archaeal-specific ribosomal proteins. *J. Mol. Biol.* **405**, 1215–1232 [CrossRef PubMed](#)
- 48 Salinas, R. K., Camilo, C. M., Tomaselli, S., Valencia, E. Y., Farah, C. S., El-Dorry, H. and Chamberg, F. S. (2008) Solution structure of the C-terminal domain of multiprotein bridging factor 1 (MBF1) of *Trichoderma reesei*. *Proteins* **75**, 518–523 [CrossRef](#)
- 49 Winey, M., Mathison, L., Sorel, C. M. and Culbertson, M. R. (1989) Distribution of introns in frameshift-suppressor proline-tRNA genes of *Saccharomyces cerevisiae*. *Gene* **76**, 89–97 [CrossRef PubMed](#)
- 50 Sandbaken, M. G. and Culbertson, M. R. (1988) Mutations in elongation factor EF-1 $\alpha$  affect the frequency of frameshifting and amino acid misincorporation in *Saccharomyces cerevisiae*. *Genetics* **120**, 923–934 [PubMed](#)
- 51 Mendenhall, M. D. and Culbertson, M. R. (1988) The yeast *SUF3* frameshift suppressor encodes a mutant glycine tRNA<sup>CCC</sup>. *Nucleic Acids Res.* **16**, 8713 [CrossRef PubMed](#)
- 52 Ball, C. B., Mendenhall, M. D., Sandbaken, M. G. and Culbertson, M. R. (1988) The yeast *SUF5* frameshift suppressor encodes a mutant glycine tRNA<sup>CCC</sup>. *Nucleic Acids Res.* **16**, 8712 [CrossRef PubMed](#)
- 53 Cammarano, P., Teichner, A., Londei, P., Acca, M., Nicolaus, B., Sanz, J. L. and Amils, R. (1985) Insensitivity of archaeobacterial ribosomes to protein synthesis inhibitors: evolutionary implications. *EMBO J.* **4**, 811–816 [PubMed](#)
- 54 Ruggero, D. and Londei, P. (1996) Differential antibiotic sensitivity determined by the large ribosomal subunit in thermophilic archaea. *J. Bacteriol.* **178**, 3396–3398 [PubMed](#)
- 55 Londei, P., Altamura, S., Sanz, J. L. and Amils, R. (1988) Aminoglycoside-induced mistranslation in thermophilic archaeobacteria. *Mol. Gen. Genet.* **214**, 48–54 [CrossRef PubMed](#)

Received 12 November 2013/13 May 2014; accepted 14 May 2014

Published as BJ Immediate Publication 14 May 2014, doi:10.1042/BJ20131474

## SUPPLEMENTARY ONLINE DATA

# Archaeal MBF1 binds to 30S and 70S ribosomes via its helix–turn–helix domain

Fabian BLOMBACH<sup>\*1,2</sup>, Helene LAUNAY<sup>†1,3</sup>, Ambrosius P. L. SNIJDERS<sup>‡</sup>, Violeta ZORRAQUINO<sup>\*</sup>, Hao WU<sup>\*4</sup>, Bart DE KONING<sup>\*</sup>, Stan J. J. BROUNS<sup>\*</sup>, Thijs J. G. ETTEMA<sup>§</sup>, Carlo CAMILLONI<sup>||</sup>, Andrea CAVALLI<sup>||</sup>, Michele VENDRUSCOLO<sup>||</sup>, Mark J. DICKMAN<sup>‡</sup>, Lisa D. CABRITA<sup>†</sup>, Anna LA TEANA<sup>¶</sup>, Dario BENELLI<sup>\*\*</sup>, Paola LONDEI<sup>\*5</sup>, John CHRISTODOULOU<sup>†5</sup> and John VAN DER OOST<sup>\*5</sup>

<sup>\*</sup>Laboratory of Microbiology, Wageningen University, Wageningen 6703HB, The Netherlands

<sup>†</sup>Institute of Structural and Molecular Biology, University College London (UCL) and Birkbeck College, University of London, London WC1E 6BT, U.K.

<sup>‡</sup>Department of Chemical and Process Engineering, University of Sheffield, Sheffield S1 3JD, U.K.

<sup>§</sup>Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, Uppsala SE-75137, Sweden

<sup>||</sup>Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.

<sup>¶</sup>Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, Ancona 60131, Italy

<sup>\*\*</sup>Dipartimento Biotecnologie Cellulari ed Ematologia, Università degli Studi di Roma "La Sapienza", Roma 00161, Italy

## STRUCTURAL MODELLING OF THE aMBF1 HTH DOMAIN

An initial model of the aMBF1-C HTH structure was derived from the chemical shift data using the CHESHIRE protocol [1] as follows. Fragments of the protein (from three to nine residues) are generated with main-chain dihedral angles and secondary structure compatible with the information contained in the chemical shifts. The fragments are then assembled in a combinatorial manner (molecular fragment replacement) to produce an ensemble of trial structures that are subsequently refined by exploiting the information about tertiary structure contained in the chemical shifts. In order to obtain the ensemble of aMBF1-C structures, replica-averaged chemical shift-restrained MD simulations were performed using GROMACS and PLUMED as described previously [2,3] using the Amber03W

force field [4]. The starting conformation was built starting from the initial HTH structures derived from chemical shifts by adding the disordered segments using PyMOL (<http://www.pymol.org>). The structure was protonated and solvated with 21 000 water molecules in a dodecahedron box of 666 nm<sup>3</sup> volume. The final structure from four 1 ns simulations were selected as starting structures for four replicas. Each replica was evolved through a series of annealing cycles between 300 and 380 K, each cycle being composed of 100 ps at 300 K, 100 ps of linear increase in the temperature to 380 K, 100 ps of constant temperature at 380 K and 300 ps of linear decrease in the temperature to 300 K. Each replica was evolved for 100 ns. Only structures from the 300 K portions of the simulations were taken into account for analysis. The resulting ensemble is composed by all the structures sampled at 300 K in the four replicas after discarding the first 10 ns.

<sup>1</sup> These authors contributed equally to this work.

<sup>2</sup> Present address: Research Department of Structural and Molecular Biology, Institute of Structural and Molecular Biology, University College London (UCL), London WC1E 6BT, U.K.

<sup>3</sup> Present address: CNRS-UMR 8576-University of Lille1, Villeneuve d'Ascq, Lille FR-59655, France.

<sup>4</sup> Present address: Protein Studies Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, U.S.A.

<sup>5</sup> Correspondence may be addressed to either of these authors (email [j.christodoulou@ucl.ac.uk](mailto:j.christodoulou@ucl.ac.uk) or [john.vanderoost@wur.nl](mailto:john.vanderoost@wur.nl)).

Co-ordinates and structure factors of the aMBF1 (archaeal multi-protein bridging factor 1) helix–turn–helix domain have been deposited in the PDB under code 2MEZ. The NMR assignment data have been deposited in the BMRB under accession number 19028.

```

taatacgcactcactatagggagaGAGGTGAAATATAAATGTCTCAAAGCTTTGAGGGAGA
      T7 promoter          SD motif          M S Q S F E G E

ATTAAAACAATTCTCAGAAGTGGCAAAGTTATTTTGGGAACAAGGAAGACATGAAATT
  L K T I L R S G K V I L G T R K T L K L

ATTAAAGACAGGAAGGTAAAGGGAGTAGTAGTTTCTTCTACATTAAGGCAGGATCTAAA
  L K T G K V K G V V V S S T L R Q D L K

AGACGATATAATGACATTTTCAAATTTTCTGATATTCCAATTTATCTCTATAAAGGTAG
  D D I M T F S K F S D I P I Y L Y K G S

TGGATATGAATTAGGGACATTATGCGGTAAACCTTTTATGGTATCTGTTATAGGTATAGT
  G Y E L G T L C G K P F M V S V I G I V

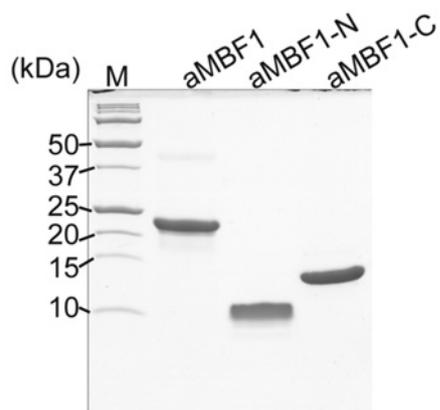
TGATGAAGGGGAATCAAAAATTTTGGAGTTTATTAAGAGGTGAAGCAATGAGTGCCAGA
  D E G E S K I L E F I K E V K Q *

AATTAAATTAAC

```

**Figure S1** Sequence of the *in vitro* translation template used in the present study based on the *orf104* mRNA

Mutations in the nucleotide and protein sequence are indicated as bold letters. The T7 promoter sequence used for *in vitro* transcription and the Shine–Dalgarno (SD) motif are underlined. Synthetic DNA of this sequence was cloned via KpnI and SacI into a pBluescript-derived vector and the plasmid was linearized with SacI before *in vitro* transcription.



**Figure S2** Heterologous expression of *S. solfataricus* aMBF1

Coomassie Blue-stained Tris/Tricine SDS/PAGE gel with 3.4  $\mu$ g of purified recombinant aMBF1 and the isolated N- and C-terminal domains. Molecular masses (M) are indicate in kDa.

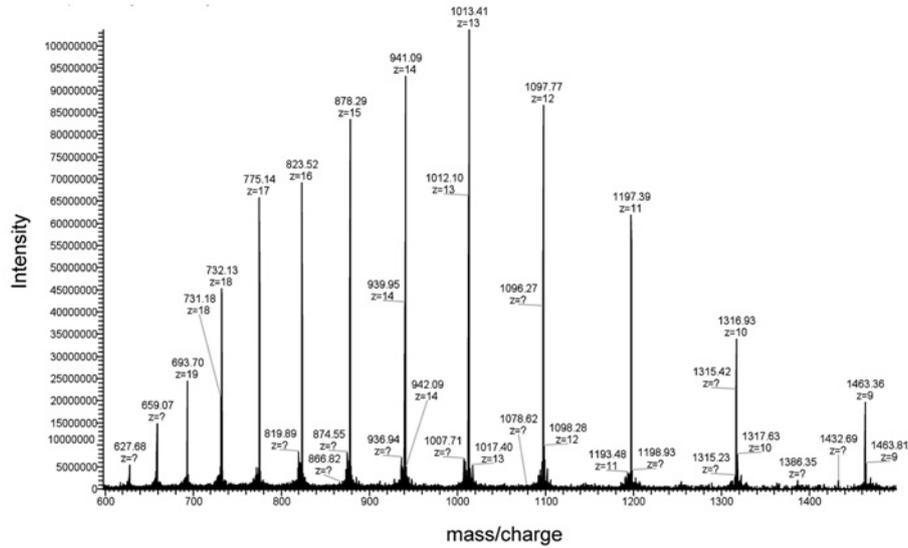


Figure S3 Mass spectrum of intact MBF1-C indicates a molecular mass of 13.161 kDa

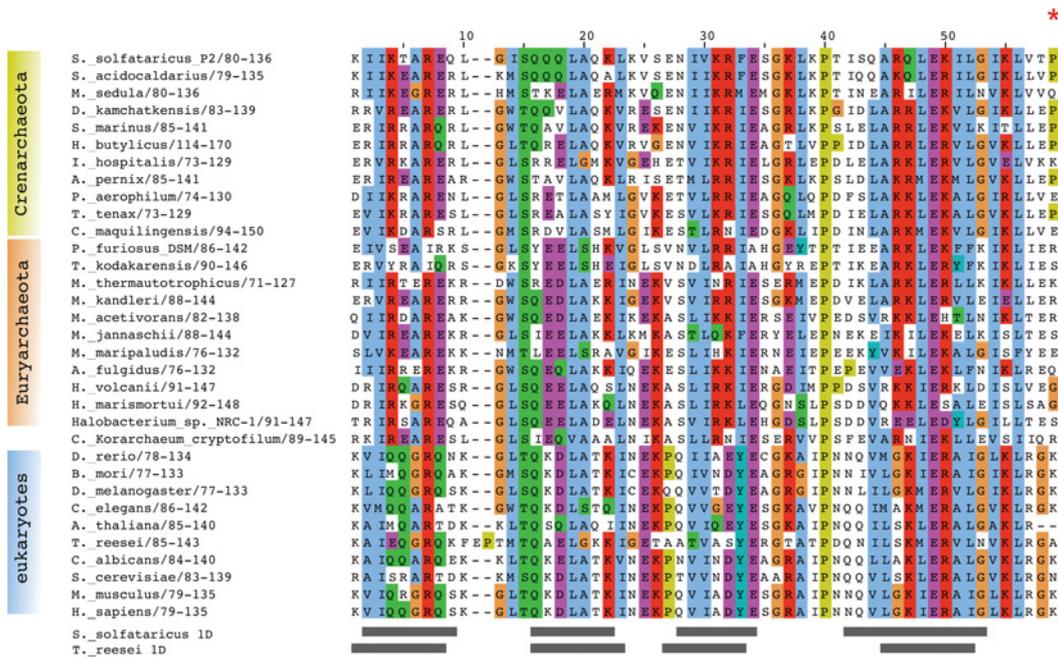
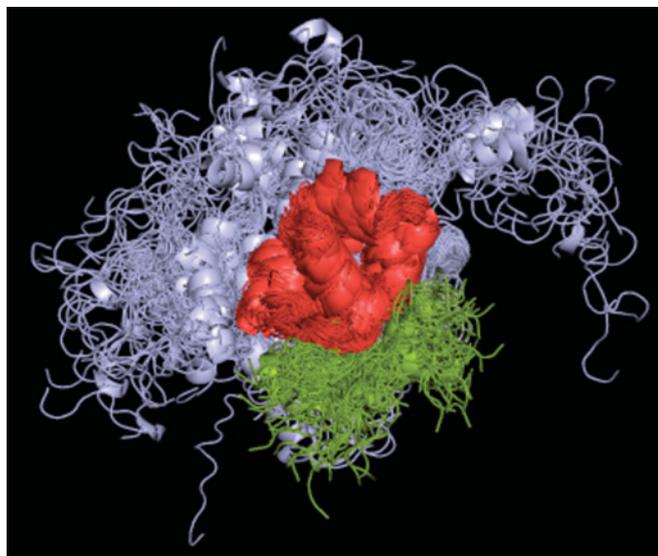


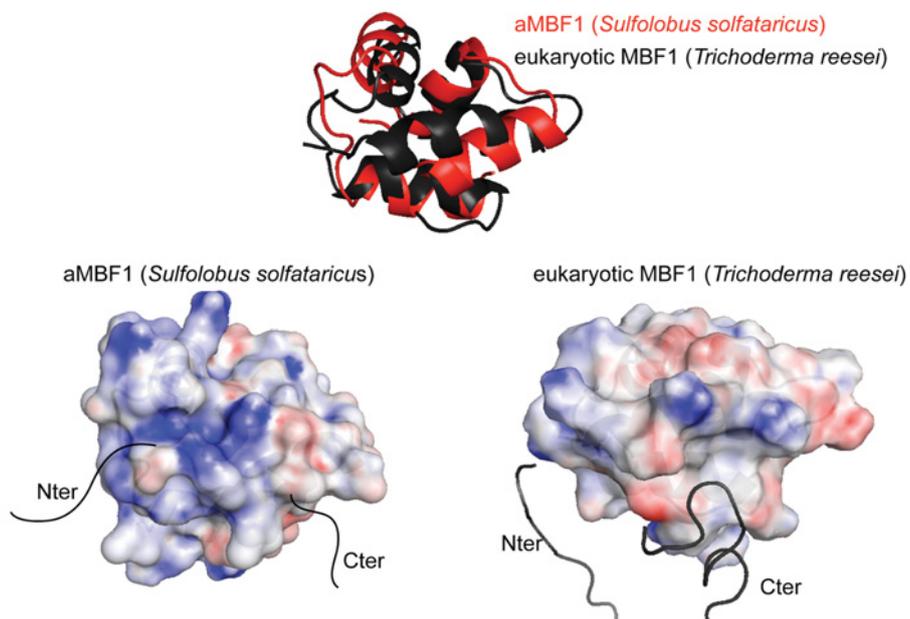
Figure S4 Sequence alignment of the HTH domains of archaeal and eukaryotic MBF1 orthologues

The red asterisk denotes the position of Pro<sup>136</sup> that undergoes *cis*–*trans* isomerization. Key to species names and GI numbers of sequences: *Sulfolobus solfataricus* strain P2, 15896971; *Sulfolobus acidocaldarius*, 70605853; *Metallosphaera sedula*, 146302785; *Desulfurococcus kamchatkensis*, 218883314; *Staphylothermus marinus*, 126464913; *Hyperthermus butylicus*, 124026906; *Ignicoccus hospitalis*, 156936795; *Aeropyrum pernix*, 118430835; *Pyrobaculum aerophilum*, 18311643; *Thermoproteus tenax*, 352681234; *Caldivirga maquilingensis*, 159040592; *Pyrococcus furiosus*, 18976372; *Thermococcus kodakarensis*, 57639935; *Methanothermobacter thermoautotrophicus*, 15678031; *Methanopyrus kandleri*, 20093440; *Methanosarcina acetivorans*, 20088899; *Methanocaldococcus jannaschii*, 15668172; *Methanococcus maripaludis*, 45357563; *Archaeoglobus fulgidus*, 11497621; *Haloferax volcanii*, 292654178; *Haloarcula marismortui*, 55376942; *Halobacterium* sp. NRC-1, 15789340; *Candidatus Korarchaeum cryptofilum*, 170289627; *Danio rerio*, 312144725; *Bombyx mori*, 112984061; *Drosophila melanogaster*, 116010443; *Caenorhabditis elegans*, 392973747; *Arabidopsis thaliana*, 240254678; *Trichoderma reesei*, 340517347; *Saccharomyces cerevisiae*, 330443743; *Mus musculus*, 372099108; *Homo sapiens*, 224589821. The location of  $\alpha$ -helices in *T. reesei* MBF1 and *S. solfataricus* aMBF1 are according to Salinas et al. [5] and the present study respectively.



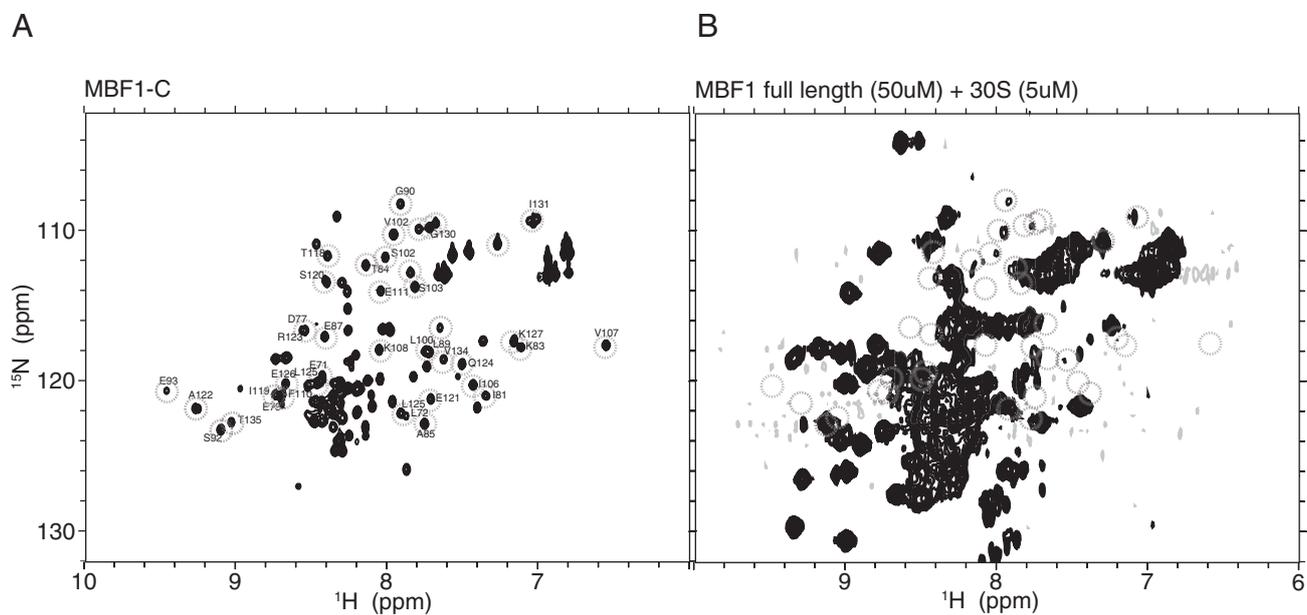
**Figure S5** Overlay of the ensemble of aMBF1-C structures determined using the CamShift-MD approach

See the Experimental section of the main text for a detailed description. The structures are aligned on the HTH motif (red). The flexible C-terminus is depicted in grey; the truncated part of the flexible linker is depicted in green.



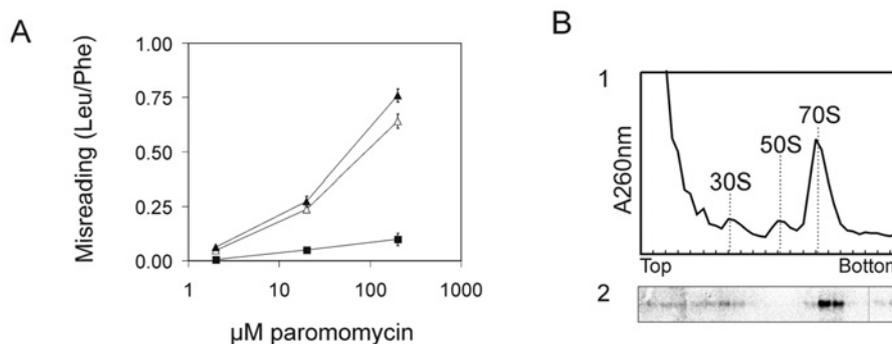
**Figure S6** Comparison of the electrostatic surfaces of the HTH domains of aMBF1 from *S. solfataricus* and eukaryotic MBF1 from *Trichoderma reesei* (PDB code 2JVL)

Upper panel: overlay of the two HTH domains. Lower panel: comparison of the electrostatic surfaces with the proteins in the same orientation as above. A range from  $-20$  kT/e (red) to  $+20$  kT/e (blue) is shown.



**Figure S7**  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of an excess of aMBF1 in the presence of 30S ribosomal subunits

(A)  $^1\text{H}$ - $^{15}\text{N}$  HSQC of isolated aMBF1-C. The cross-peaks marked with open circles were broadened in the presence of 30S ribosomal subunits. (B)  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC of full-length aMBF1 and 30S ribosomal subunits at a 10:1 molar ratio (50 to 5  $\mu\text{M}$ ). The cross-peaks marked with open circles were broadened in the presence of 30S ribosomal subunits.



**Figure S8** Effect of aMBF1 on paromomycin-induced misreading in poly(U)-directed translation assays

(A) *In vitro* misreading rate as detected by the incorporation of leucine into polyphenylalanine by poly(U)-programmed ribosomes. Poly(U) encodes polyphenylalanine; upon misreading of  $\text{tRNA}^{\text{leu}}$  by the ribosomes, leucine will be incorporated. The ratio between phenylalanine and leucine incorporated into the polypeptide is measured. *In vitro* translation reactions were set up as in the other experiments, but the mRNA template was replaced by 20  $\mu\text{g}$  of poly(U) per 25  $\mu\text{l}$  assay volume. In addition, assays contained 3 mM spermine to increase translation fidelity [6] and 20  $\mu\text{M}$  phenylalanine and leucine replacing the amino acid mixture. In each assay, one of the amino acids was replaced by either  $\text{L}$ -[ $^{14}\text{C}$ ]phenylalanine or  $\text{L}$ -[ $^{14}\text{C}$ ]leucine (PerkinElmer) respectively. After incubation at 70  $^{\circ}\text{C}$  for 30 min, 18  $\mu\text{l}$  was spotted on a 1 cm  $\times$  1.5 cm sheet of 3 mm Whatman chromatography filter paper and precipitated overnight at 4  $^{\circ}\text{C}$  in 10% trichloroacetic acid. Filters were washed at 95  $^{\circ}\text{C}$  in 5% trichloroacetic acid for 5 min, three times in ice-cold 5% trichloroacetic acid and finally, filters in 96% ethanol before scintillation counting. Each assay was carried out in duplicate plus two additional samples lacking poly(U) for background subtraction. Different cell lysates were tested:  $\Delta\text{mbf1}$  (open triangle), parental strain PBL2025 (closed triangle) and wild-type strain P2 (closed square). The rate of leucine residues incorporated per phenylalanine residue was approximately 0.004 for the wild-type strain P2 in the absence of antibiotic, in good agreement with the published value of 0.003 [6]. Both the  $\Delta\text{mbf1}$  strain and its parental strain PBL2025 exhibited misreading rates of approximately 0.020 leucine residues incorporated per phenylalanine residue in the absence of antibiotic. The  $\sim$ 5-fold higher misreading rates observed in PBL2025 and the  $\Delta\text{mbf1}$  strain might be due to the deletion of  $\sim$ 50 genes in the parental PBL2025 strain [7]. Alternatively, it could be a phenotype of the *S. solfataricus* 98/2 strain from which PBL2025 was derived. (B) Effect of 100  $\mu\text{M}$  paromomycin on the co-migration of endogenous aMBF1 with 30S ribosomal subunits and 70S ribosomes in cell lysate programmed for translation.

## REFERENCES

- 1 Cavalli, A., Salvatella, X., Dobson, C. M. and Vendruscolo, M. (2007) Protein structure determination from NMR chemical shifts. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 9615–9620 [CrossRef](#) [PubMed](#)
- 2 Camilloni, C., Cavalli, A. and Vendruscolo, M. (2013) Assessment of the use of NMR chemical shifts as replica-averaged structural restraints in molecular dynamics simulations to characterize the dynamics of proteins. *J. Phys. Chem. B* **117**, 1838–1843 [CrossRef](#) [PubMed](#)
- 3 Best, R. B. and Mittal, J. (2010) Protein simulations with an optimized water model: cooperative helix formation and temperature-induced unfolded state collapse. *J. Phys. Chem. B* **114**, 14916–14923 [CrossRef](#) [PubMed](#)
- 4 Ottiger, M., Delaglio, F. and Bax, A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.* **131**, 373–378 [CrossRef](#) [PubMed](#)
- 5 Salinas, R. K., Camilo, C. M., Tomaselli, S., Valencia, E. Y., Farah, C. S., El-Dorry, H. and Chambergo, F. S. (2008) Solution structure of the C-terminal domain of multiprotein bridging factor 1 (MBF1) of *Trichoderma reesei*. *Proteins* **75**, 518–523 [CrossRef](#)
- 6 Londei, P., Teixido, J., Acca, M., Cammarano, P. and Amils, R. (1986) Total reconstitution of active large ribosomal subunits of the thermoacidophilic archaeobacterium *Sulfolobus solfataricus*. *Nucleic Acids Res.* **14**, 2269–2285 [CrossRef](#) [PubMed](#)
- 7 Schelert, J., Dixit, V., Hoang, V., Simbahan, J., Drozda, M. and Blum, P. (2004) Occurrence and characterization of mercury resistance in the hyperthermophilic archaeon *Sulfolobus solfataricus* by use of gene disruption. *J. Bacteriol.* **186**, 427–437 [CrossRef](#) [PubMed](#)

---

Received 12 November 2013/13 May 2014; accepted 14 May 2014

Published as BJ Immediate Publication 14 May 2014, doi:10.1042/BJ20131474