4052–4066 *Nucleic Acids Research, 2010, Vol. 38, No. 12* *Published online 9 March 2010*
doi:10.1093/nar/gkq141

Heterodimerization of the human RNase P/MRP subunits Rpp20 and Rpp25 is a prerequisite for interaction with the P3 arm of RNase MRP RNA

Katherine L. D. Hands-Taylor[1], Luigi Martino[1], Renée Tata[1], Jeffrey J. Babon[2], Tam T. Bui[3], Alex F. Drake[3], Rebecca L. Beavil[1], Ger J. M. Pruijn[4], Paul R. Brown[1] and Maria R. Conte[1,*]

[1]Randall Division of Cell and Molecular Biophysics, King's College London, New Hunt's House, Guy's Campus, London SE1 1UL, UK, [2]Structural Biology Division, Walter and Eliza Hall Institute of Medical Research, 1G Royal Pde, Parkville 3052, VIC, Australia, [3]Pharmaceutical Science Division, King's College London, The Wolfson Wing, Hodgkin Building, Guy's Campus, London SE1 1UL, UK and [4]Department of Biomolecular Chemistry, Nijmegen Centre for Molecular Life Sciences, Institute for Molecules and Materials, Radboud University of Nijmegen, Nijmegen, The Netherlands

Received January 11, 2010; Revised February 16, 2010; Accepted February 17, 2010

ABSTRACT

Rpp20 and Rpp25 are two key subunits of the human endoribonucleases RNase P and MRP. Formation of an Rpp20–Rpp25 complex is critical for enzyme function and sub-cellular localization. We present the first detailed *in vitro* analysis of their conformational properties, and a biochemical and biophysical characterization of their mutual interaction and RNA recognition. This study specifically examines the role of the Rpp20/Rpp25 association in the formation of the ribonucleoprotein complex. The interaction of the individual subunits with the P3 arm of the RNase MRP RNA is revealed to be negligible whereas the 1:1 Rpp20:Rpp25 complex binds to the same target with an affinity of the order of nM. These results unambiguously demonstrate that Rpp20 and Rpp25 interact with the P3 RNA as a heterodimer, which is formed prior to RNA binding. This creates a platform for the design of future experiments aimed at a better understanding of the function and organization of RNase P and MRP. Finally, analyses of interactions with deletion mutant proteins constructed with successively shorter N- and C-terminal sequences indicate that the Alba-type core domain of both Rpp20 and Rpp25 contains most of the determinants for mutual association and P3 RNA recognition.

INTRODUCTION

Ribonucleoproteins (RNPs) are functional units formed by the association of protein-coding and non-coding RNAs with proteins. RNPs are involved in a large spectrum of molecular activities and govern key cellular functions such as gene expression and its regulation; the significance of their roles is emphasized by the many diseases caused by mutations that disrupt either the RNA or the protein component of the RNP, or the factors required for their correct assembly (1–3). Several RNPs are implicated in the biogenesis of RNA, including the related endoribonucleases RNase mitochondrial RNA processing (MRP) and RNase P, both composed of an RNA molecule and several protein subunits. RNase MRP, identified only in Eukarya, is involved in pre-rRNA processing, in particular in the formation of the mature 5′-end of the 5.8S rRNA (4,5). Very recently, however, a more prevalent role of RNase MRP in ribosome biogenesis has emerged, specifically for entry of 35S pre-rRNA processing into the canonical maturation pathway (6). Despite its predominant localization in the nucleolus (7–9), RNase MRP also functions in mitochondrial DNA replication, by cleaving an RNA primer required for this process (10), and it has been shown to partake in the degradation of the mRNA encoding the mitosis specific cyclin B2 in yeasts (11). The significance of RNase MRP's role in human growth and differentiation is substantiated by the link between mutations in the human RNA subunit and the

*To whom correspondence should be addressed. Tel: +44 20 7848 6194; Fax: +44 20 7848 6435; Email: sasi.conte@kcl.ac.uk

pleiotropic condition termed cartilage–hair hypoplasia (CHH), a severe form of dwarfism (12).

Unlike RNase MRP, RNase P is ubiquitous in all domains of life; in all cases, it processes precursor tRNA transcripts, removing the 5′ leader sequences to generate their 5′ mature termini (13). Moreover, the *Saccharomyces cerevisiae* enzyme has been shown to be involved in C/D small nucleolar RNA (snoRNA) processing (14).

The RNA components of RNase MRP and P are essential for enzymatic activity and a high degree of similarity has been found or predicted in their structural features across species (15–22). Such structural correspondence has prompted endorsement for an evolutionary relation between RNase P and MRP, and in support of this hypothesis both complexes have been found to share many protein subunits that co-purify with respective endoribonuclease activities (13,22,23). In particular, to date up to 10 subunits have been identified in both human enzymes (hPop1, Rpp38, Rpp21, Rpp29/hPop4, hPop5, Rpp25, Rpp20, Rpp14, Rpp30, Rpp40), of which nine have homologues in yeast; in addition, RNase MRP from *S. cerevisiae* contains two specific proteins: Smn1 and Rmp1, and reports are discordant as to whether yeast Rpp21 (Rpr2) is unique to RNase P (22,24–27). Although in eukaryotes the RNA moieties of RNase P and MRP are thought to embody the catalytic core of the respective enzymes, they are reliant on the protein subunits for function *in vitro* and *in vivo* (13,24). The specific reasons for this requirement remain however unknown; alleged roles of the protein components, that await to be demonstrated, include maintaining the correct and active 3D fold of the RNA molecule, determining sub-cellular localization and contributing to substrate RNA discrimination (13).

The elucidation of the overall composition and architecture of the eukaryotic RNase P and MRP holoenzymes has been stalled by the challenges encountered in the isolation of the RNP complexes from native sources on the one hand, and in obtaining pure, stable individual recombinant components for reconstitution studies on the other (13). As a first step towards a depiction of the spatial organization of the holoenzymes, binary interactions between protein subunits, and with the RNA component, have been investigated in both human and yeast systems (28–32). However, such studies posed the question as to whether the inconsistencies and lack of completeness encountered in the investigations were a reflection of experimental artefacts, leading to false positive or false negative results. First, numerous attempts foiled by technical difficulties in obtaining recombinant proteins or by their high susceptibility to misfold, aggregate, degrade or be altered by their fusion tag have been reported (13,28,31,32); second, these experiments suffer from the drawback that synergic actions are not taken into account (13). The detailed molecular interaction of some of the subunits has been well established by structural data, in particular concerning the archaeal proteins of RNase P (that have homologues in eukaryotic systems) namely Rpp29–Rpp21 (33,34) and Pop5–Rpp30 (35); however, to our knowledge no systematic biophysical analysis of eukaryotic subunits has been reported thus far. We have therefore undertaken a comprehensive biochemical and biophysical characterization of two essential subunits of human RNase P and MRP, namely Rpp20 and Rpp25, with the aim of understanding their molecular properties and identifying potential problems with the recombinant products that may encumber assembly/interaction analyses. Rpp20 and Rpp25 have been reported to associate stably with one another and to interact with the P3 arm of the RNase MRP RNA (25,31,36,37). We find that Rpp20/Rpp25 association is very strong and that the heterodimerization activity is located largely in the central Alba-type core domain of both proteins. Importantly our results clearly show that Rpp20 and Rpp25 operate in tandem and that the formation of a 1:1 heterodimer is an obligate prerequisite for RNA binding.

MATERIALS AND METHODS

Plasmid construction

The cDNAs of full-length human Rpp20 and Rpp25 in pCR4_TOPO were used as templates for further cloning. Rpp20 and Rpp25, and deletion mutants thereof were subcloned in many different expression vectors, but in many cases the recombinant products turned out to be insoluble, aggregated or affected by their fusion tag. We report here only the plasmids that have been utilized in this work.

Full-length Rpp25 (encompassing residues 1–199) was subcloned with an N-terminal hexahistidine tag into a pPROEX-HTb expression vector (Invitrogen) using Nco1/Not1 restriction sites. Full-length Rpp20 (residues 1–140) was subcloned by PCR into a pET-30 expression vector using the LIC methodology (Novagen); the pET-30 vector was modified to bear a TEV cleavage site to remove the N-terminal hexahistidine tag. Rpp20(35–140) was amplified from pET-30 Rpp20 by PCR to introduce 5′ Asc1 and 3′ Not1 sites and then ligated into a modified version of pET-15b vector (Novagen) with an Asc1 site inserted directly after the his tag.

Fragments subcloned in pETDuet-1 vector (Novagen) with an N-terminal hexahistidine tag were: Rpp20, Rpp20(16–140), Rpp25(25–170) and co-expressed Rpp20/Rpp25 (in which Rpp20 bears the his tag). Rpp25(25–170) was also subcloned in a pCDFDuet-1 vector (Novagen) also with an N-terminal hexahistidine tag (Supplementary Data). For cloning into pETDuet-1 and pCDFDuet-1, PCR primers were designed to complement appropriate regions of DNA in the pET-30 and pPROEX-htb clones encoding Rpp20 and Rpp25 proteins, respectively. Forward PCR primers used for producing his-tagged proteins encoded a TEV-cleavage site (ENLYFQG). Restriction sites incorporated in the forward/reverse PCR primers were as follows: EcoRI/NotI for Rpp20 and Rpp20(16–140); EcoRI/PstI for Rpp25(25–170); DNA encoding non-his-tagged proteins was amplified by PCR forward/reverse primers with restriction sites NdeI/XhoI. PCR products were cut with the appropriate restriction enzymes and ligated into similarly digested pETDuet-1 or pCDFDuet-1. *Escherichia*

coli KRX cells (Promega) were transformed with the ligation mixtures and transformants, obtained on LB amp plates (for pETDuet-1 clones) or LB streptomycin plates (for pCDFDuet-1 clones), were screened by colony PCR and recombinant clones were sequenced.

Protein expression and purification

Rpp25, Rpp20 (from pET-30 vector) and Rpp20(35–140) were expressed in *E. coli* strain BL21(DE3)pLysS at 37 and 18°C, respectively, induced with 1 mM IPTG (isopropyl β-D-thiogalactoside) at an OD_{600} of ~0.6 and incubated for a further 3 h. All the other proteins subcloned into pETDuet-1 or pCDFDuet-1 vectors were expressed in *E. coli* KRX strain at 25°C, induced using 1 mM IPTG and 0.1% (w/v) rhamnose at an OD_{600} of ~0.6 and incubated overnight.

Cell pellets were lysed by sonication in 20 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 7.4 and centrifuged to separate the soluble and insoluble fractions of the cell. All recombinant proteins listed above were soluble except Rpp20 expressed in pET-30. In this case, the insoluble fraction of the pellet was completely denatured in a buffer containing 8 M urea (see below).

As all the soluble recombinant proteins contained a hexahistidine-tag, or were co-expressed with hexahistidine-tag proteins, they were purified by affinity chromatography on 5 ml HisTrap columns (GE Healthcare) with a gradient of 20–500 mM imidazole. The His-affinity tag was removed by addition of TEV protease and incubation at 30°C overnight in 25 mM sodium phosphate, 100 mM NaCl, 1 mM DTT (Dithiothreitol), pH 7; the reaction mixture was subsequently applied to a Ni–NTA column (Qiagen) to remove the cleaved tags, the His-tagged TEV and any undigested product. The proteins were further purified on a 5 ml Hi-Trap heparin column, mainly to remove nucleic acids contaminants. The proteins eluted with a linear 0–2 M NaCl gradient were dialysed overnight in buffer A [phosphate buffered saline (PBS, Sigma) containing 10 mM phosphate, 2.7 mM KCl and 137 mM NaCl, adjusted to pH 7, plus 1 mM DTE (Dithioerythritol)] and loaded on a Superdex^{TM} 75 column in the same buffer (see below). The tag was not removable for Rpp20(35–140).

The cell pellet containing Rpp20 expressed from pET-30 vector was lysed by sonication in 20 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 7.4. After centrifugation, the pellet containing inclusion bodies was resuspended in 100 mM sodium phosphate, 20 mM imidazole, 8 M urea, pH 7 and then filtered through a 0.2 μm filter. The protein was purified on a 5 ml HisTrap column (GE Healthcare) under denaturing conditions with a 20–500 mM imidazole gradient. Rpp20 was then refolded by stepwise dialysis as follows: the eluted protein was placed in a dialysis membrane with molecular weight cut-off of 6000 Da and equilibrated for 24 h in a base buffer (0.4 M L-Arginine, 25 mM sodium phosphate, 100 mM NaCl, 1 mM DTE, pH 7) containing 4 M urea at 4°C. Denaturant was slowly removed by a series of overnight equilibrations with buffers of decreasing urea concentration (2, 1, 0 M). Finally, the sample was dialysed into TEV protease buffer for removal of the tag. Subsequent purification steps were as described for the other proteins. The correct refolding of Rpp20 was validated by CD analysis. When recombinant Rpp20 (in pET-30 vector) was co-expressed with Rpp25(25–170) (in pCDFDuet-1) for analytical ultracentrifugation (AUC) studies (see below), the complex was found in the soluble fraction and did not require refolding.

Protein concentrations were calculated based upon the near-UV absorption (ε_{280}) using theoretical extinction coefficients derived from ExPASy (38).

Size exclusion chromatography

Size exclusion chromatography (SEC) was performed using an AKTA Basic System (GE Healthcare) in combination with a Superdex^{TM} 75 column (10/300). All samples were exchanged into identical buffer conditions (buffer A). The column was calibrated using separate injections of aprotinin (6.5 kDa), ribonuclease A (13.7 kDa), carbonic anhydrase (29.0 kDa), ovalbumin (44.0 kDa) and conalbumin (77.0 kDa) from the Low Molecular Weight Gel Filtration Calibration Kit (GE Healthcare) under the same buffer conditions.

RNA sample preparation

The P3 arm of RNase MRP RNA was purchased from IBA GmbH (Göttingen, Germany) and dissolved in buffer A. The RNA were annealed by heating at 95°C for 5 min followed by slow cooling to room temperature as previously described (39). The concentration of the dissolved oligonucleotide was evaluated by UV measurement at 95°C, using a molar extinction coefficient at 260 nm calculated by the nearest-neighbour model (463 700 $M^{-1} cm^{-1}$) (40).

Isothermal titration calorimetry

The protein and RNA solutions were prepared in buffer A. Experiments titrating Rpp25 with Rpp20 (and deletion mutants) were performed at three different temperatures (293, 298 and 303 K) using a high-sensitivity ITC-200 microcalorimeter from Microcal (GE Healthcare). RNA–protein titrations were carried out at 298 K on an ITC-200 microcalorimeter. Before each isothermal titration calorimetry (ITC) experiment, the pH of each solution was checked, the reference cell was filled with deionized water, and the protein solutions were degassed for 2–5 min to eliminate air bubbles. The first addition was executed only after achieving baseline stability. Measurement from the first injection was discarded from the analysis of the integrated data, in order to avoid artefacts due to the diffusion through the injection port occurring during the long equilibration period, locally affecting the protein concentration near the syringe needle tip. To investigate protein–protein interactions, in each titration volumes of 2 μl of a solution of Rpp20 or mutants thereof at a concentration of 160–180 μM were injected into a solution of Rpp25 or mutants thereof (20 μM) in the same buffer, using a computer-controlled 40-μl microsyringe. For the RNA–protein studies, the P3

RNA at a concentration of 170 μM was injected into a solution containing individual proteins or their pre-assembled complexes at a concentration of 20 μM in buffer A. Notably, the titration of Rpp25 with Rpp20 was also repeated in perchlorate buffer (25 mM sodium phosphate, 10 mM sodium perchlorate, 1 mM DTE, pH 7) to check any effect of this buffer (used in the CD experiments, see below) on protein behaviour. A spacing of 200 s between each injection was applied to enable the system to reach the equilibrium. Heat produced by titrant dilution was verified to be negligible by performing a control experiment, titrating it into the buffer alone, under the same conditions. Integrated heat data obtained for the titrations were fitted using a non-linear least-squares minimization algorithm to a theoretical titration curve, using the MicroCal-Origin 7.0 software package. ΔH° (reaction enthalpy change in $kJ\,mol^{-1}$), K_b (binding constant in M^{-1}), and n (molar ratio between the two proteins in the complex) were the fitting parameters. The reaction entropy was calculated using the relationships $\Delta G^\circ = -RT\ln K_b$ ($R = 8.314\,J\,mol^{-1}\,K^{-1}$, $T = 298\,K$) and $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$. In addition, from the experiments at different temperatures (293, 298 and 303 K), the change in heat capacity ΔC_p° upon binding has been calculated as the resulting slope in a plot of ΔH° versus the experimental temperature.

The Rpp20 self-association was characterized following a dilution protocol (41), injecting 3 μl of a Rpp20 solution at a concentration of 6 μM into the calorimetric cell containing 200 μl of buffer A at 298 K. The resulting dissociation isotherm was interpolated using a non-linear regression fitting procedure based on a simple dissociation model (Supplementary Data).

Circular dichroism spectroscopy

Simultaneous UV and circular dichroism (CD) spectra of $0.2\,mg\,ml^{-1}$ protein solutions were recorded on the Applied Photophysics Ltd. Chirascan Plus Spectrometer (Leatherhead, UK), continuously flushed with nitrogen. The molar ellipticity per mean residue, [θ] ($deg\,cm^2\,dmol^{-1}$), was calculated from the equation: $[\theta] = [\theta]_{obs} \times (mrw)/10 \times l \times C$, where $[\theta]_{obs}$ is the ellipticity (deg), mrw is the mean residue molecular weight, C is the protein concentration ($g\,ml^{-1}$), and l is the optical path length of the cell (cm). Rectangular Suprasil cells with 1 and 0.05 cm path lengths were employed to record spectra in the regions between 400–230 and 260–180 nm, respectively. The parameters used to acquire the spectra were: spectral bandwidth of 1 nm, data step-size of 1 nm with a time-per-data-point of 1.5 and 3.0 s in the near-UV and the far-UV, respectively. All the spectra were baseline corrected by subtracting the buffer spectrum. Experiments were conducted in buffer A at 298 K. The sodium chloride content of many buffers absorbs light very strongly below 200 nm making CD measurements below 200 nm very difficult and secondary-structure analysis less reliable (42). A full far-UV CD spectrum down to 180 nm was only possible replacing NaCl with sodium perchlorate in the buffer medium (see above). To estimate the secondary structure content, curve fitting was performed using DICHROWEB (43).

Nuclear magnetic resonance spectroscopy

Recombinant ^{15}N-labelled Rpp25 and Rpp25(25–170) were prepared on minimal media as described previously (44) and dissolved in buffer A at concentrations of 0.1–0.3 mM. Nuclear magnetic resonance (NMR) spectra were recorded at 298 K on a Bruker Avance spectrometer operating at 16.4 T equipped with a triple resonance cryoprobe. Spectra were processed and analysed as previously reported (44).

Analytical ultracentrifugation

Sedimentation equilibrium experiments were performed using a Beckman Optima XL-A analytical ultracentrifuge as described previously (45). Samples were prepared in buffer A and data were acquired with an average of 25 absorbance measurements at a wavelength of 280 nm and a radial spacing of 0.001 cm. Sedimentation equilibrium experiments were performed at 4°C and rotor speeds of 17 000, 14 500 and 12 000 rpm for both Rpp20 and co-expressed Rpp20/Rpp25(25–170). Protein concentrations were in the range of 15–30 μM for Rpp20/Rpp25(25–170) and 90–180 μM for Rpp20. For data analysis, the partial specific volumes ($\bar{v}$) and monomeric molecular weights of the different proteins were calculated from amino acid composition using SEDNTERP (http://www.rasmb.bbri.org/) and gave the following values: Rpp20 from pETDuet-1 vector ($0.7269\,cm^3\,g^{-1}$; 15 706 Da); Rpp20/Rpp25(25–170) ($0.7253\,cm^3\,g^{-1}$; 16026 Da and $0.7303\,cm^3\,g^{-1}$; 15 435 Da, respectively). The solvent density (ρ) was calculated to be $1.00722\,g\,ml^{-1}$ at 4°C. The monomeric buoyant molecular mass $M(1 - \bar{v}\rho)$ was calculated to be: 4206 for Rpp20 from pETDuet-1, 4319 for Rpp20 from pET-30 and 4081 for Rpp25(25–170). Data for all concentrations and speeds were analysed simultaneously using a range of models in the Sigmaplot software package as described previously (45). Residuals were calculated by subtracting the best fit of the model from the experimental data.

RESULTS

Domain analysis of Rpp20 and Rpp25

Rpp20 and Rpp25 are postulated to function in tandem and to be components of both human RNase P and MRP complex (25,31,36,37). Comparative genomics and primary sequence profile analysis have unveiled an evolutionary connection for Rpp20, Rpp25 and the archaeal protein Alba; in particular, the phylogenetic tree derived from the alignment of the Alba domain has subdivided this superfamily into two eukaryotic families (containing Rpp20 and Rpp25 orthologues, respectively) and one archaeal Alba family (46). The primary function ascribed to Alba proteins relates to DNA packaging and chromosomal organization in archaea, in an equivalent role to that of eukaryotic histones (47). Structural studies have unveiled a compact α/β fold for the Alba

domain with a topology akin to the C-terminal domain of IF3 (48). Since the first structure of the archaeal Alba1 was published by Wardleworth *et al.* (48), several structures have been determined for Alba1 and Alba2 from different organisms, and the first structure available within the eukaryotic subgroups belongs to the At2g34160 protein from *Arabidopsis thaliana*, of unknown function, which has not been characterized beyond the deposition of the coordinates (PDB ID 2Q3V). The primary sequence alignment of Rpp20 and Rpp25 with proteins of known structure (Figure 1) complemented with their secondary-structure prediction analysis [PredictProtein (49), data not shown] would suggest the presence of an Alba-type core domain flanked by additional secondary-structure elements and/or unstructured regions. Guided by such analyses, we embarked on mutagenesis studies of Rpp20 and Rpp25 to identify regions required for their domain stability, their mutual interaction and RNA recognition.

Characterization of Rpp20

Human Rpp20 is a 16 kDa protein predicted to consist of an Alba-like domain and an N-terminal tail. To characterize the molecular properties of Rpp20, far-UV CD analysis was applied to full-length protein (encompassing residues 1–140) and to two truncated C-terminal fragments, spanning amino acids 16–140 and 35–140,

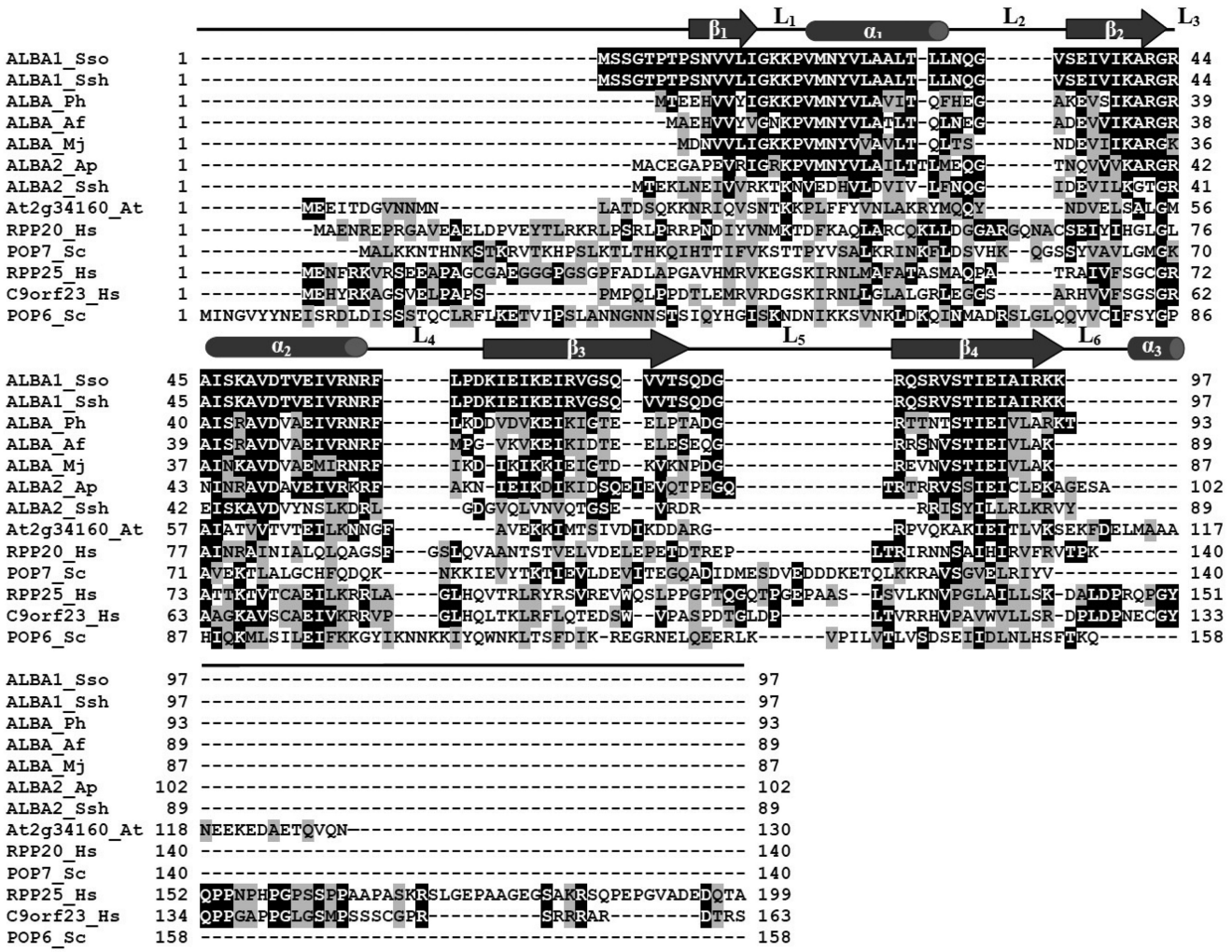

Figure 1. Sequence alignment for the Alba superfamily of proteins. For the archaeal subgroup, all Alba proteins whose structures have been solved to date are displayed; At2g34160 is an eukaryotic protein of this superfamily whose structure has been determined; human Rpp20 and Rpp25 are shown together with their respective yeast homologues Pop7 and Pop6; human C9orf23, belonging to the subgroup of Rpp25, is also included. The alignment was obtained using ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Invariant residues are boxed in black and conserved residues are in grey. The secondary-structure elements found in the archaeal Alba proteins are superposed on the amino acid sequence. The topology of all structures is $\beta_1\alpha_1\beta_2\alpha_2\beta_3\beta_4$, but the plant At2g34160 protein contains an additional α-helix in the C-terminal region (labelled α_3). The loop regions are labelled L_1–L_5. The protein species along with their gene identifier are: ALBA1_Sso, *Sulfolobus solfataricus*, gi46397340; ALBA1_Ssh, *Sulfolobus shibatae*, gi46397339; ALBA_Ph, *Pyrococcus horikoshii*, gi34582290; ALBA_Af, *Archaeoglobus fulgidus*, gi34810782; ALBA_Mj, *Methanocaldococcus jannaschii*, gi40889143; ALBA2_Ap, *Aeropyrum pernix*, gi34582348; ALBA2_Ssh, *Sulfolobus shibatae*, gi157881229; At2g34160_At, *Arabidopsis thaliana*, gi52696237; RPP20_Hs, *Homo sapiens*, gi153791431; POP7_Sc, *Saccharomyces cerevisiae*, gi6319644; RPP25_Hs, *Homo sapiens*, gi74733233; C9orf23_Hs, *Homo sapiens*, gi55958109; POP6_Sc, *Saccharomyces cerevisiae*, gi1723656.

henceforth referred to as Rpp20(16–140) and Rpp20(35–140), respectively. The deletion mutants were designed to retain the regions conserved within the superfamily (46) (Figure 1), and presumably an intact Alba-type domain. It is noteworthy that extensive molecular biology and biochemistry work had to be devoted to obtain reproducibly soluble and stable proteins (see 'Materials and Methods' section and below).

Far-UV CD spectral analyses indicated that Rpp20 adopts an α/β fold, compatible with the Alba structure, alongside flexible regions, and consists of ~18% α-helix, 25% β-strand, 24% turns and 33% of disordered/other structures (Figure 2). As the shape of the curves is very similar for Rpp20 and its deletion mutants (Figure 2A), the comparative CD study supports the hypothesis that the folded core of the protein is retained in all the constructs examined, and that the most N-terminal 34 residues are likely to be predominantly unstructured. The secondary-structure content for the deletion mutants of Rpp20, however, could not be estimated because of the instability and aggregation of Rpp20(16–140) and Rpp20(35–140) in perchlorate and other buffers tested which would be transparent below 200 nm (see 'Materials and Methods' section).

During purification of recombinant proteins, we noticed that full-length Rpp20 and its deletion mutants Rpp20(16–140) and Rpp20(35–140) eluted in SEC at apparent molecular weights of about 47, 43 and 40 kDa, respectively (Figure 3 and Table 1), indicating the formation of oligomers in solution. Subsequent AUC sedimentation equilibrium experiments on full-length Rpp20 could not confirm the presence of a single ideal species, as systematic residuals were observed (Figure 3); nevertheless, a good fit was obtained with a model of self-association and non-interacting mixtures, signifying that under the conditions used Rpp20 exists primarily in a homodimeric state, alongside small amounts of contaminating aggregates not in equilibrium with the dimer itself. This may be due to part of the protein irreversibly unfolding and associating during the AUC experiments. Interestingly, most of the Alba proteins studied also form dimers in solution.

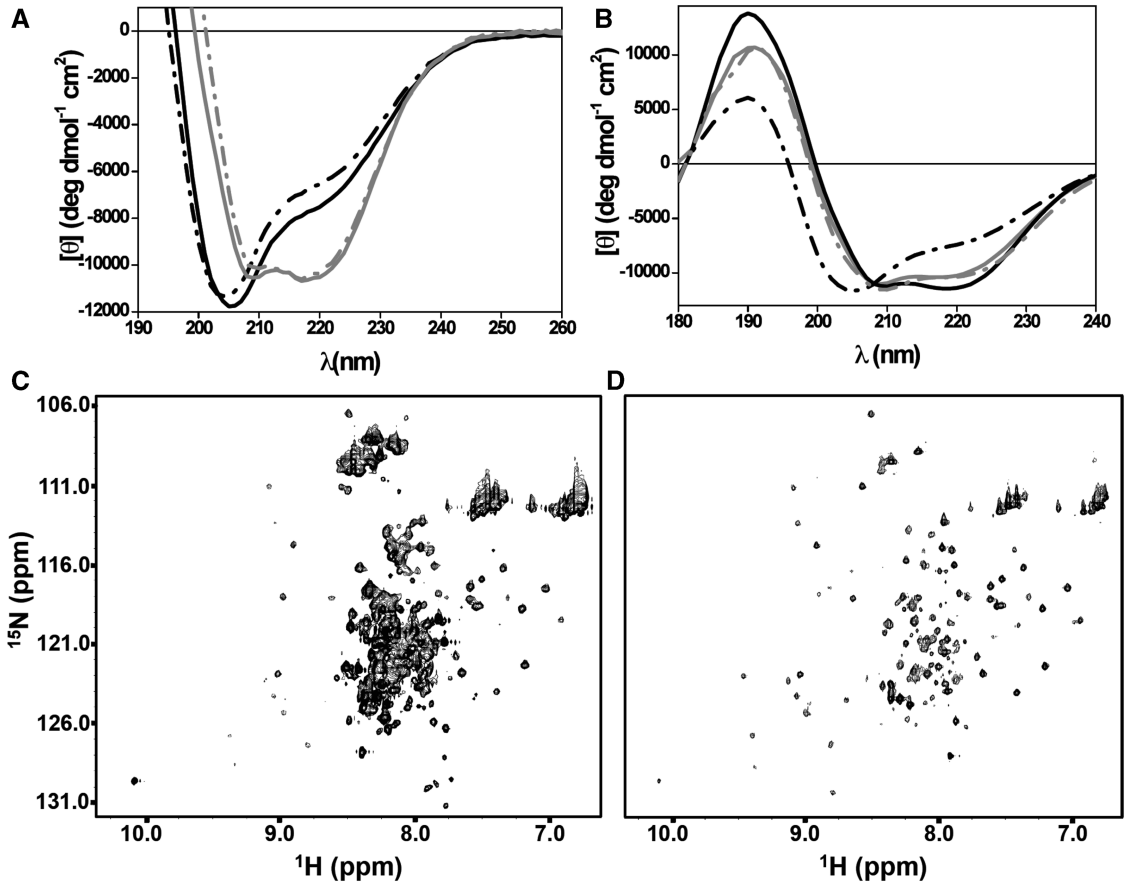

Figure 2. Molecular characterization of Rpp20 and Rpp25. (**A**) Far-UV CD spectra of Rpp20 grey, straight line; Rpp20(35–140) grey, dotted line; Rpp25 black, straight line; and Rpp25(25–170) black, dotted line. The spectra were acquired in buffer A (see 'Materials and Methods' section), pH 7 at 298 K. Rpp20(35–140) contained a 20-residue N-terminal histidine tag. (**B**) Far-UV CD spectra of apo Rpp20 black, straight line; apo Rpp25 black, dotted line; Rpp20/Rpp25 complex at 1:1 molar ratio, grey, straight line; theoretical curve for the Rpp20/Rpp25 complex as a weighted sum of the isolated curves, grey, dotted line. The spectra were recorded in 20 mM sodium phosphate, 100 mM sodium perchlorate, pH 7 at 298 K. The secondary-structure content estimated by DICROWEB CD spectra analysis gave the following values: Rpp20 (18% α; 25% β; 24% turn; 33% irregular); Rpp25 (18% α; 22% β; 25% turn; 34% irregular). (**C**) [^{1}H–^{15}N] HSQC NMR spectra of Rpp25 and (**D**) Rpp25(25–170) recorded in buffer A, at 16.4 T and 298 K.

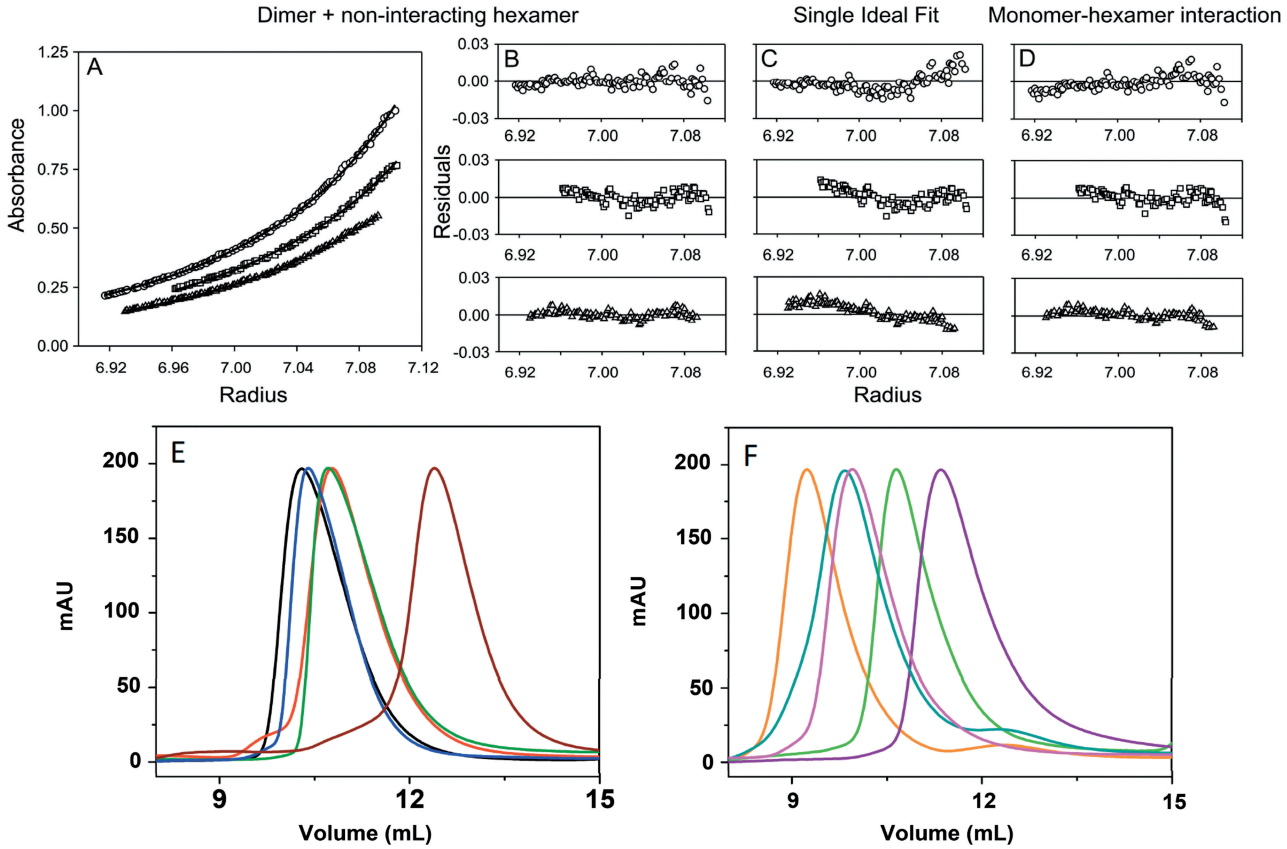

Figure 3. Analysis of the oligomeric state of Rpp20, Rpp25 and their complexes. (**A**) Sedimentation equilibrium analytical ultracentrifugation data for Rpp20, collected at 17 000 r.p.m. using protein concentrations of 90 μM (circles), 135 μM (squares) and 180 μM (triangles). The data curves were simultaneously fitted to a model of homodimer + non-interacting hexamer, with the residuals (**B**) randomly distributed around zero. Attempts to fit the data curves to a single species or to other models, for example to a monomer in equilibrium with a hexamer, gave poorer fits as shown by the residuals in (**C**) and (**D**), respectively. (**E**) Size exclusion chromatography elution profiles for apo proteins: Rpp20 black; Rpp20(16–140) blue; Rpp20(35–140) olive; Rpp25 red; and Rpp25(25–170) wine. (**F**) Size exclusion chromatography elution profiles for complexes: [Rpp20-Rpp25]/P3 RNA orange; [Rpp20(35–140)–Rpp25(25–170)]/P3 RNA cyan; Rpp20/Rpp25 magenta; co-expressed Rpp20/Rpp25(25–170) green; Rpp20(35–140)/Rpp25(25–170) purple. Formula and calculated molecular masses are shown in Table 1.

Table 1. Size exclusion chromatography elution values for Rpp20, Rpp25, their mutants and their complexes

Protein	Formula MW (kDa)	Calculated MW (kDa)	Elution volume (ml)
Rpp20[a]	32.1 (16.0)[a]	47.1	10.30
Rpp20(16–140)[a]	28.1 (14.1)[a]	45.3	10.40
HisRpp20(35–140)[a,b]	26.4 (13.2)[a]	39.9	10.72
Rpp25	20.6	39.2	10.77
Rpp25(25–170)	15.4	20.7	12.39
Rpp20/Rpp25[c]	36.6	53.8	9.96
Rpp20/Rpp25(25–170)[d]	31.5	41.1	10.65
Rpp20(35–140)/Rpp25(25–170)	28.6	31.1	11.35
[Rpp20-Rpp25]/P3 RNA	52.5	71.4	9.24
[Rpp20(35–140)-Rpp25(25–170)]/P3 RNA	44.4	56.4	9.84

The calculated molecular weights were obtained as reported in the 'Materials and Methods' section. The formula molecular weights of the complexes was reported considering 1:1 heterodimers for Rpp20:Rpp25 and mutants thereof, and 1:1:1 for Rpp20:Rpp25:P3 RNA (see text).
[a]Formula molecular weight of the homodimer, with the monomeric value in brackets.
[b]His-tag was not removed (see 'Materials and Methods' section).
[c]Identical values were found for complexes prepared by mixing separately made proteins and for complexes obtained by co-expressing recombinant proteins (see 'Materials and Methods' section).
[d]Complex obtained by co-expression of recombinant subunits (see 'Materials and Methods' section).

The process of association of Rpp20 was characterized further by conducting dilution experiments using ITC, where the amount of heat measured upon each injection is governed by the enthalpy change of dissociation and the monomer–dimer equilibrium constant (Supplementary Data) (41). Analysis of the data

shown in Supplementary Figure S1 revealed an enthalpic change of 20 kJ mol^{-1} and a derived association constant of 8×10^5 M^{-1} (Supplementary Data).

Characterization of Rpp25

Human Rpp25 is a 20.6 kDa protein that belongs to the eukaryotic group most divergent from the archaeal Alba subfamily. A distinctive trait of Rpp25 and its orthologues is the occurrence of a long C-terminal extension bearing the conserved GYQXP signature (at position 150–154 in human Rpp25), whose function remains unidentified. Bioinformatic analyses of Rpp25 (Figure 1 and see above) favour the presence of a putative Alba-type core domain flanked by N- and C-terminal extensions and to characterize this further a number of deletion mutants were designed accordingly. Nonetheless, for several, the lack of expression or solubility of the recombinant products, or else instability of the purified proteins could not be overcome (data not shown), hence efforts converged towards the attainment of the full-length protein and one deletion mutant truncated at both N- and C-terminal tails, specifically Rpp25(25–170).

To evaluate the effects of such N- and C-terminal truncations on the secondary structure of Rpp25, far-UV CD spectra were recorded for the wild-type and Rpp25(25–170) mutant, revealing comparable curves for the two proteins and bearing the hallmarks of α/β domains. Nonetheless, the molar ellipticities are significantly less negative for the full-length Rpp25, in concomitance with the negative peak shifted towards 200 nm. It is therefore conceivable that longer Rpp25 has a lower amount of structured content, supporting the notion of partially disordered N- and C-terminal tails. Although a full secondary structure content prediction for Rpp25(25–170) was impeded by the poor behaviour of the protein in perchlorate buffer, the qualitative CD results are in excellent agreement with the NMR analysis: here the comparison of ^{1}H–^{15}N HSQC spectra of full-length Rpp25 and its mutant Rpp25(25–170) demonstrated that the well-dispersed resonances belonging to the folded portion of the protein remain unaltered in the shorter construct, plus the sharp signals exclusive to the intact Rpp25 spectrum manifest a very narrow range of chemical shifts, symptomatic of flexible conformations (50) (Figure 2).

On a size exclusion column, Rpp25 migrated with an apparent molecular weight of ~39 kDa; however, the partially unstructured nature of the N- and C-terminal appendices (see above) is expected to influence the effective hydrodynamic radius of the macromolecule, and may therefore result in an overestimation of the protein mass by this technique. Unfortunately, accurate determination of the molecular weight using shape-independent AUC equilibrium sedimentation was precluded by the inherent instability of the full-length product at concentrations required for this experiment over a period of 2–3 days. Nonetheless, the presumably more globular mutant Rpp25(25–170) (15.4 kDa), behaved plausibly as a monomer on the SEC, with an apparent molecular mass of ~21 kDa (Figure 3 and Table 1). On the basis that full-length Rpp25 elutes later than full-length Rpp20 from the size exclusion column in the same experimental conditions (Figure 3), despite a greater formula molecular weight and conceivably a less globular shape, we concluded that Rpp25 in solution exists largely in a monomeric state.

ITC experiments reveal a strong interaction between Rpp25 and Rpp20

Despite earlier reports that Rpp20 and Rpp25 interact strongly *in vivo* and *in vitro*, no quantitative measure of this interaction has ever been reported. This would be important to determine, with the aim of understanding whether the two proteins are likely to operate as a single working pair or act individually; also the analysis of the energetics of the interaction might provide insights into their mechanism of action. To this aim, we employed ITC, which is largely used to investigate molecular binding reactions by measuring the heat generated or absorbed in the binding event and thereby providing the binding constant, the molar ratio of the two proteins in the complex and the enthalpy change (ΔH°) of the interaction. For the Rpp25–Rpp20 system the integrated heat data, at each investigated temperature, showed that the binding process is composed of one clear event centred on a molar ratio of one (Figure 4). The binding isotherm curves corresponding to this reaction have been interpolated using an independent-sites model, revealing that at 298 K full-length proteins interact with each other with an association constant (K_b) of 5.3×10^7 M^{-1}, with an enthalpy change (ΔH°) of −94 kJ mol^{-1} and an entropic contribution (defined as $-T\Delta S^{\circ}$) of 50 kJ mol^{-1}. The thermodynamics of the interaction indicate that the binding event is enthalpically driven.

The N- and C-terminal regions of Rpp25 and the N-terminal tail of Rpp20 are not involved in mutual recognition

To delineate the regions of Rpp20 responsible for Rpp25 recognition, ITC measurements were carried out on the deletion mutants Rpp20(16–140) and Rpp20(35–140) truncated in the N-terminal tail. Interestingly, in both cases the global thermodynamics of the interaction with Rpp25 is seemingly unaffected (Table 2), suggesting that the N-terminal region of Rpp20 does not participate in this recognition.

A second set of experiments was performed monitoring complex formation between the Rpp25(25–170) mutant and all the available versions of Rpp20, to evaluate the contribution to binding of the N- and C-terminal regions of Rpp25 (Figure 4 and Table 2). The thermodynamic signature of the association is fully preserved in all cases (Table 2), thereby demonstrating that the elements required for the mutual interaction remain intact in the deletion mutants and that both the N- and C-terminal tails of Rpp25 do not partake in the recognition process. Furthermore, the K_b values obtained in the reactions with the truncated Rpp25 mutant are slightly higher than the ones found with full-length Rpp25, indicative of an improved efficiency of the interaction and

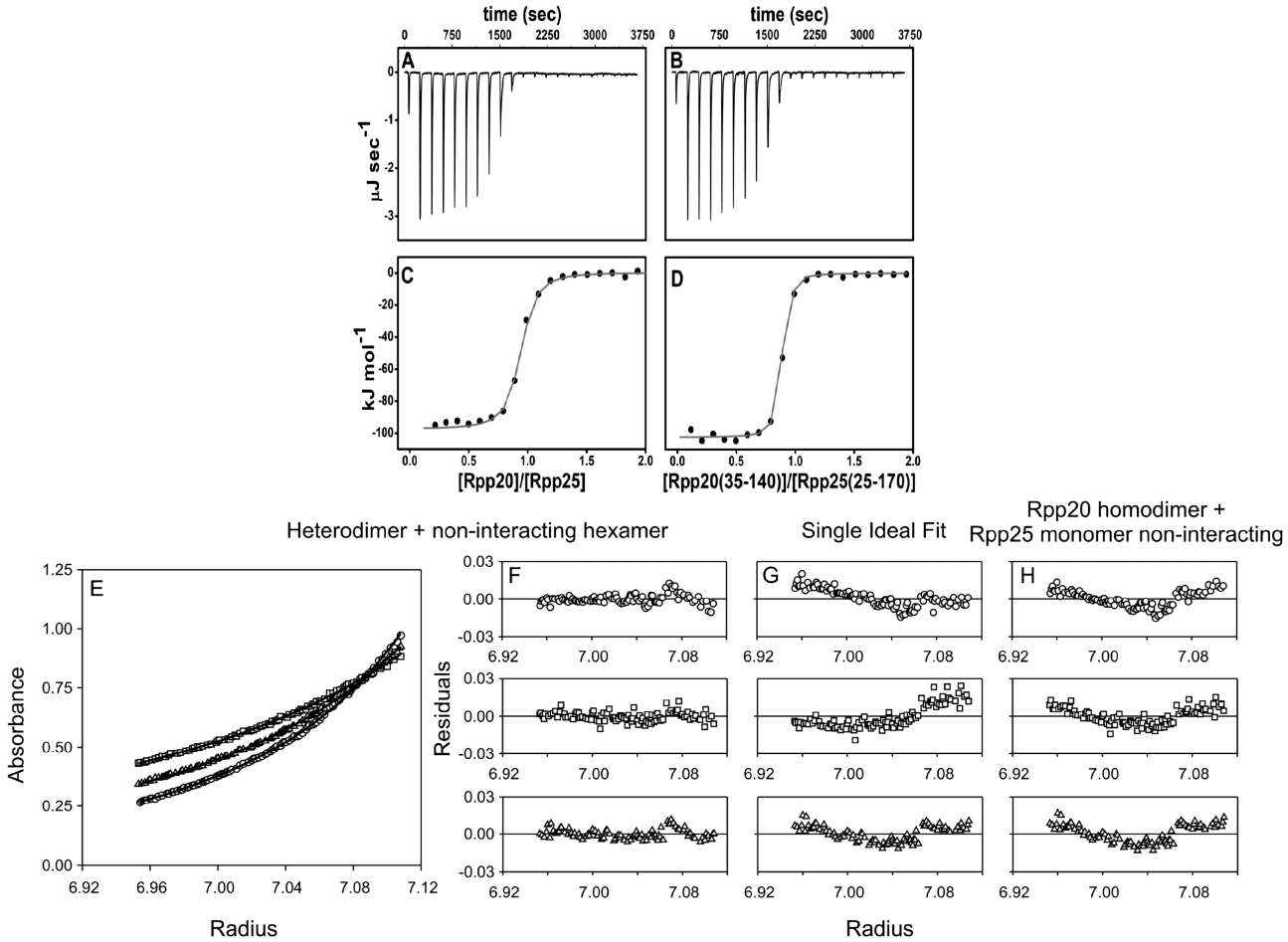

Figure 4. Analysis of Rpp20–Rpp25 complex formation. Raw titration data showing the thermal effect of (**A**) injecting Rpp20 into a calorimetric cell containing Rpp25 and (**B**) injecting Rpp20(35–140) into a calorimetric cell containing Rpp25(25–170). All the proteins were dissolved in buffer A. (**C** and **D**) The normalized heat of interaction, for the titrations shown in (A) and (B), respectively, was obtained by integrating the raw data and subtracting the heat of ligand dilution into the buffer alone. The grey line represents the best fit obtained by a non-linear least-squares procedure based on an independent binding sites model. (**E**) Sedimentation equilibrium analytical ultracentrifugation data for co-expressed Rpp20/Rpp25(25–170), collected at protein concentration of 24 μM and three speeds: 17 000 r.p.m. (circles), 12 000 r.p.m. (squares) and 14 500 r.p.m. (triangles). The data curves were simultaneously fitted to a heterodimer + non-interacting hexamer model, with the residuals (**F**) randomly distributed around zero. Attempts to fit the data curves to a single species or to other models, for example to a mixture of non-interacting Rpp20 homodimers and Rpp25 monomers, gave poorer fits as shown by the residuals in (**G**) and (**H**), respectively.

Table 2. Thermodynamic signature of the interactions between Rpp20 and Rpp25, their truncation mutants and the P3 RNA

Interaction	n	K_b (M^{-1})	ΔH° (kJ mol^{-1})	$-T\Delta S^\circ$ (kJ mol^{-1})	ΔG° (kJ mol^{-1})
Rpp25/Rpp20	0.8	5.3×10^7	−94.1	50.0	−44.1
Rpp25/Rpp20(16–140)	0.9	7.9×10^7	−105.8	60.8	−45.0
Rpp25/Rpp20(35–140)	0.9	3.2×10^7	−98.7	55.9	−42.8
Rpp25(25–170)/Rpp20	0.8	5.8×10^7	−105.0	60.7	−44.3
Rpp25(25–170)/Rpp20(16–140)	0.8	4.9×10^7	−103.8	59.9	−43.9
Rpp25(25–170)/Rpp20(35–140)	0.8	4.9×10^7	−102.1	58.2	−43.9
[Rpp25-Rpp20]/P3 RNA	0.9	1.3×10^7	−27.6	−13.0	−40.6
[Rpp25(25–170)–Rpp20(35–140)]/P3 RNA	0.8	6.3×10^6	−24.3	−14.5	−38.8

The reported values represent the average over three independent measurements and the error was found to be <5%.

probably linked to the absence of the mainly unstructured tails.

Taken together, the ITC experiments reveal that the regions within the mutants Rpp20(35–140) and Rpp25(25–170) are sufficient for mutual interaction, indicating that this recognition could be mediated largely, if not exclusively, by the Alba-type core domains. In support, CD spectroscopy showed that the association between Rpp20 and Rpp25 has no detectable influence on their secondary/tertiary structure (Figure 2); in other

words, this interaction is not accompanied by structural rearrangements, thus ruling out refolding events in the N- and/or C-terminal extensions upon binding.

The association reaction results in a large loss of solvent-accessible area

To investigate the Rpp20–Rpp25 interaction in more depth, we measured the values for the change in heat capacity ΔC_p° upon binding; these are all negative and centred around $-3\,kJ\,mol^{-1}\,K^{-1}$ (Supplementary Table S1). From the ΔC_p° values, an estimation of the solvent-accessible surface area buried upon association could in principle be derived; this however requires an accurate determination of the ionization and protonation contributions to the ΔH°, achieved by performing titration experiments in different buffer/salt conditions (51). Unfortunately, we could not find a sufficient number of experimental conditions for such measurements because of protein misbehaviour. Nonetheless, the large negative ΔC_p° values point towards a large loss of solvent-accessible surface area upon complex formation, suggesting that each protein participates in the association through a large interacting surface. This finding is consistent with the high values measured for the K_b.

Rpp20 and Rpp25 form a stable 1:1 heterodimer

ITC experiments indicate that Rpp20 and Rpp25 interact with one another in a 1:1 molar ratio. Nonetheless, the homodimeric nature of Rpp20 invites the question whether the complex consists of a 1:1 (Rpp20/Rpp25) heterodimer or a 2:2 (Rpp20/Rpp25) heterotetramer. Understanding the exact stoichiometry of the Rpp20/Rpp25 complex is a matter of key importance: since many of the RNase P/MRP subunits do indeed self-associate, it has been suggested that the holoenzyme(s) might contain multiple copies of its constituents (13,26).

To address this point, AUC sedimentation equilibrium experiments were attempted on a number of complexes with various combinations of full-length and deletion mutant proteins, including co-expressed recombinant fragments. Some of the results, however, were ambiguous and deemed inconclusive, once more thwarted by the high propensity of the complexes to aggregate and in some cases by their susceptibility to degrade over a short period of time. The best results were obtained for the co-expressed Rpp20/Rpp25(25–170), though yet again the data did not correspond to a single species but gave a reasonable fit using a model of self-association with non-interacting aggregates. The best fit was found for a 1:1 heterodimer in the presence of, but not in equilibrium with, other species of larger molecular weight (Figure 4). Consistent with this, the pattern of the SEC profiles clearly supports the Rpp20/Rpp25 heterodimer model, with apparent molecular masses of their complexes (including a variety of combinations with truncation mutants) well below the expected formula weights for putative tetramers (Figure 3 and Table 1). This not only implies a significant overlap in the surfaces of Rpp20 implicated in homodimerization and Rpp25 binding, but also that the two subunits are optimized to maximize heterotypic complementarity. A quantitative treatment of the equilibrium between the Rpp20 homodimer and the Rpp20/Rpp25 heterodimer revealed that the preferential heterodimer formation arise from a higher association constant of Rpp20/Rpp25 with respect to the self-association of Rpp20 (Supplementary Data). Interestingly, this closely resembles the molecular behaviour of Alba1/Alba2 proteins, which form an obligate heterodimer though existing individually as tight homodimers (52). A 1:1 stoichiometry is also in agreement with what was found for the yeast homologue Pop6/Pop7 complex (53), although the oligomeric state of Pop7 (homologous to Rpp20) was not characterized beyond observations of self-association in GST-pull down assays (32).

Rpp20 and Rpp25 interact with the P3 arm of RNase MRP RNA in a highly synergic fashion

The Rpp20/Rpp25 complex and the isolated Rpp25 have been reported to interact with the P3 arm of the RNase MRP RNA (hereafter referred to as P3 RNA) (36,37). To gain a better insight into the nature of these interactions, we performed ITC titrations between the P3 RNA and an array of Rpp20 and Rpp25 proteins and complexes. To begin with, the RNA binding ability of the individual full-length proteins was appraised: as shown in Figure 5, the ITC results indicated that isolated Rpp20 was unable to interact with P3 RNA, whereas Rpp25 displayed a weak association, with a dissociation constant lower than the threshold of binding that could be fully characterized by ITC, i.e. around 0.1 mM. Next, under the same experimental conditions, we measured the binding of a previously assembled Rpp20/Rpp25 complex with the P3 RNA (Figure 5 and Table 2): in this case a much stronger interaction was detected, with a binding affinity of $1.3 \times 10^7\,M^{-1}$ and favourable changes in enthalpy and entropy (as $-T\Delta S^\circ$) of $-28\,kJ\,mol^{-1}$ and $-13\,kJ\,mol^{-1}$, respectively. The negative enthalpy change denotes the formation of new intermolecular contacts between the protein complex and the RNA molecule, whereas the negative entropic contribution suggests an increased degree of freedom of the system upon binding, probably because of displacement of bound water molecules from interacting surfaces. It is noteworthy that, as the protein concentrations used in the ITC measurements are a 1000-fold higher than the K_d value describing protein–protein binding, we anticipate that the Rpp20/Rpp25 mixture would exist solely in the heterodimeric form prior to RNA interaction, as confirmed by gel filtration analysis (Figure 5). Our results therefore reveal a truly synergic mechanism by which heterodimer formation augments the affinity of the individual subunits for the P3 RNA by at least a 1000-fold, thereby conferring P3 RNA binding proficiency to the Rpp20 and Rpp25 subunits. The strong specific binding to the P3 stem–loop is in complete agreement with previously reported GST pull-down studies (36).

Thus it appears that the generation of the required RNA-binding surface for this RNA target strictly hinges on the heterodimerization of the protein subunits, and to determine whether the partially disordered extensions of Rpp20 and Rpp25 would participate in RNA interaction,

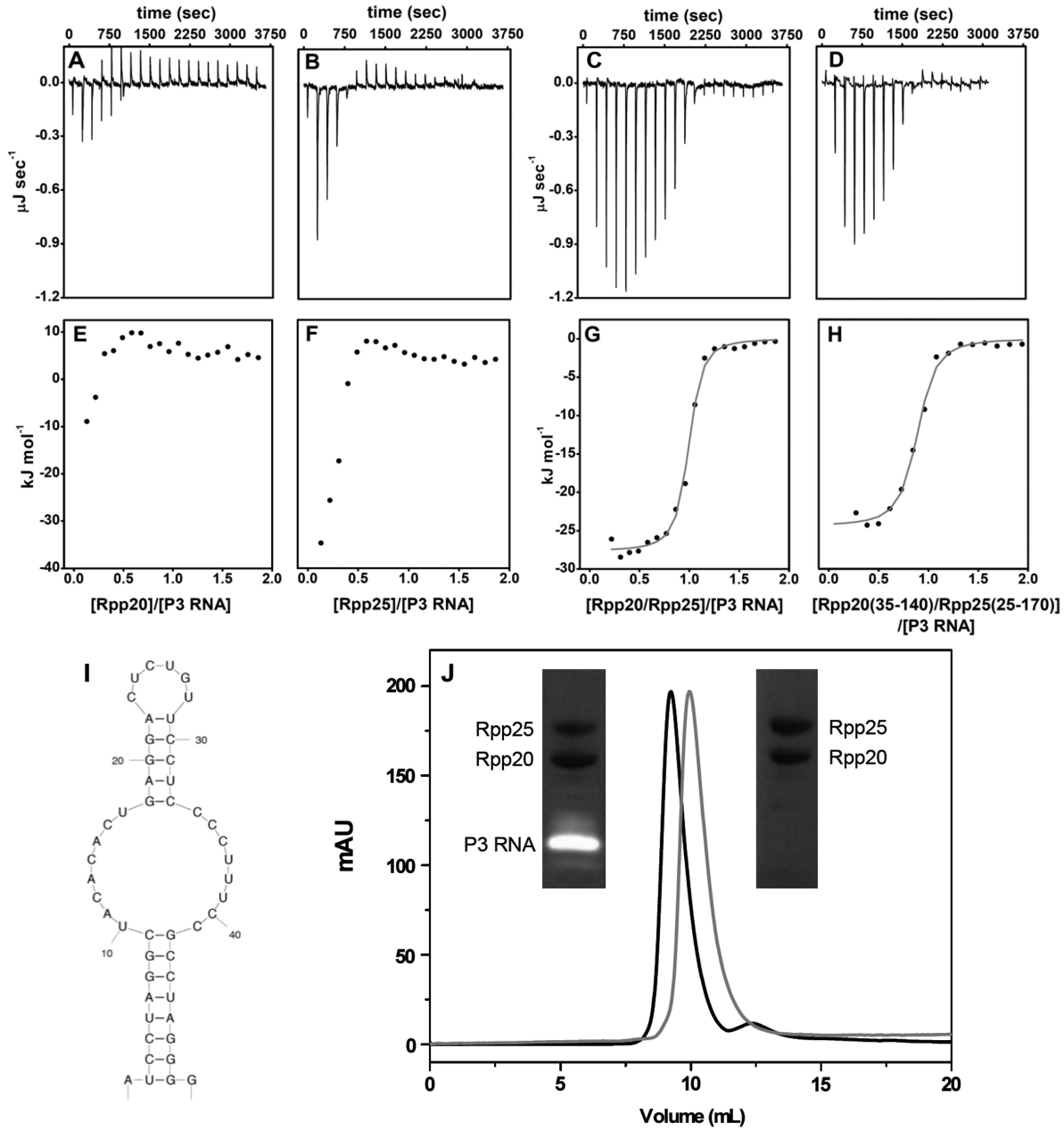

Figure 5. Rpp20/Rpp25 interaction with P3 RNA. Raw titration data showing the thermal effect of 2 μl injections of 180 μM P3 RNA into a calorimetric cell filled with (**A**) 20 μM Rpp20; (**B**) 20 μM Rpp25; (**C**) 20 μM Rpp20/Rpp25 complex; and (**D**) 20 μM Rpp20(35–140)/Rpp25(25–170) complex. All the molecules were dissolved in buffer A. The heat effect reveals null to weak interactions in (A) and (B) and a strong association in (C) and (D). (**E–H**) The normalized heat of interaction, for the titrations shown in (A–D), respectively, was obtained by integrating the raw data and subtracting the heat of P3 RNA dilution into the buffer alone. The grey lines in G and H represent the best fit derived by a non-linear least-squares procedure based on an independent binding sites model. (**I**) Expected secondary structure of the P3 arm of the RNase MRP RNA obtained with the software mfold (http://mobyle.pasteur.fr/cgi-bin/portal.py?form=mfold). (**J**) Gel filtration elution profiles and SDS–PAGE analysis of Rpp20/Rpp25 (grey) and [Rpp20-Rpp25]/P3 RNA (black) complexes from the ITC experiment samples (see text).

a second set of ITC experiments using the truncated heterodimer Rpp20(35–140)/Rpp25(25–170) was undertaken. The results in Figure 5 and Table 2 show that this complex binds to P3 RNA with slightly lower affinity than the full-length counterparts, suggesting that, whereas the extensions beyond the domain core of Rpp20 and Rpp25 had essentially no effect on protein–protein interaction, they may well play a role in RNA recognition. Nevertheless, since the decrease in RNA binding affinity exhibited by the deletion mutant complex is relatively small, the core of the Rpp20/Rpp25 proteins plausibly dominates the RNA interaction, with the tails only yielding a minor contribution to stabilization of the RNA complex.

Contrary to archaeal Alba proteins, which interact with double stranded DNA at stoichiometries of 12 bp or 6 bp per dimer (48,52), our isothermal binding curves clearly indicate that each Rpp20/Rpp25 heterodimer binds to one

molecule of 50 nt P3 RNA. A stoichiometry of 1:1:1 (Rpp20:Rpp25:P3) agrees well with the size exclusion profiles (Figure 3 and Table 1).

DISCUSSION

The purpose of this study was to shed light on the organization and function of the RNase P and MRP, in particular focussing on the essential Rpp20 and Rpp25 subunits. The recombinant proteins were subjected to an array of biochemical and biophysical methodologies, with the aim of understanding their molecular behaviour and the details of molecular recognition within the RNP complex. The value of these investigations is 2-fold: first, such molecular characterizations are *per se* key to elucidate RNP architecture and mode of action, and they are generally not afforded by studies on purified native particles; secondly, a detailed scrutiny of the molecular behaviour of recombinant products could provide a solid platform for the design of subsequent experiments. It is noteworthy that a further dimension of complexity and interest of these systems is the nearly identical composition of the protein subunit element for the RNase P and MRP of a given organism, despite their distinctive functions; furthermore, it was very recently reported that the human RNase MRP RNA associates with the telomerase reverse transcriptase to generate an RNP with RNA-dependent RNA polymerase activity (54). This adds conviction to the view that the determinants of the function for a given RNP particle reside in its entirety, underscoring therefore the importance of elucidating the molecular basis of recognition between the different components within the RNP systems.

Rpp20 and Rpp25 have previously been shown *in vivo* and *in vitro* to associate with one another, with functional relevance for holoenzyme association and localization (36). Nonetheless, it remained unclear whether the Rpp20/Rpp25 complex would constitute a single working unit, in other words whether this complex would represent the minimal protein and RNA-binding unit within the RNP particle. Such a conclusion can in fact only be derived from systematic biophysical and biochemical analyses with multiple components, beyond the largely qualitative screening of binary Rpp25–RNA and Rpp20–RNA interactions hitherto reported (31,36,37,55,56). Although Rpp20 has been said to have the ability to enhance the interaction between Rpp25 and the RNA (36,37), the results presented here firmly indicate that Rpp20 and Rpp25 operate as a single working unit, whereby their assembly is an obligate prerequisite for P3 RNA interaction. In fact, Rpp20–Rpp25 nanomolar dissociation constant would strongly argue for a preponderance of the heterodimeric form also *in vivo*, which agrees well with particle composition analysis studies (31); in addition, the ITC results unambiguously prove that the binding to the P3 arm of the RNase MRP RNA is a fully synergic event, with an association constant for the complex at least a 1000-fold higher than for Rpp25 alone. Given the similarity between the P3 domain of RNase P and RNase MRP RNA, it is likely that the same molecular recognition occurs in both holoenzymes. A similar conclusion was also proposed for the yeast homologues Pop6/Pop7, although the suggested synergic mechanism here stopped short of full characterization because of difficulties in obtaining both individual proteins (53).

The detailed molecular explanation for such synergic behaviour remains to be uncovered; nonetheless, our current results point to the likelihood that the interface required for P3 RNA binding is generated upon the formation of the heterodimer. The bioinformatic analysis in conjunction with the CD and ITC investigations presented here show that Rpp20 and Rpp25 contain a core Alba-type domain flanked by extensions (N-terminal for Rpp20 and N- and C-terminal for Rpp25) which do not contribute to core domain stabilization or to protein–protein interaction, but may play a minor role in RNA recognition. Specifically, our study explored whether truncation of these extensions affected the structural and functional properties relative to the wild-type. As none of the mutants analysed elicit structural alterations or exhibited differences in mutual recognition compared to the full-length proteins, we reason that the extensions are probably largely unstructured and that Rpp20 and Rpp25 interact with one another mainly through the Alba-type core domain without undergoing substantial conformational rearrangements, forming a tight 1:1 heterodimer. Within the structure catalogue of archaeal Alba proteins, the heterodimeric organization is exemplified by the Alba1/Alba2 complex from *Sulfolobus solfataricus* (52) (Figure 6).

This quaternary arrangement exhibits a high degree of conservation amongst the homodimeric archaeal Alba structures and, interestingly, is extremely similar to the very recent structure of the yeast Pop6/Pop7 heterodimer bound to the P3 arm of the yeast RNase MRP (57). The protein–protein interface in these complexes is formed by helix α2 and the last two β-strands (β3 and β4) from both monomers, engaging in an extensive network of hydrophobic interactions and specific hydrogen bond contacts (Figure 6). Our results would support an overall structural resemblance between Rpp20/Rpp25 and the yeast homologue system: akin to Pop6/Pop7, Rpp20 and Rpp25 in fact interact with one another *via* the Alba-like domain with a very extensive solvent-accessible surface area buried by this molecular association, as suggested by our ITC experiments (specifically the large negative ΔC_p° values). Interestingly, Pop6 and Pop7 displayed an Alba-type domain fold with additional short secondary-structure elements in the N-terminal region (in particular a β-strand and an α-helix) (57); whether this is also the case for Rpp20 and Rpp25 remains to be established. Bioinformatics analysis, however, shows that for Pop6 these elements are comprised in a region not present at all in Rpp25 [Figure 1 and ref. (57)], whereas a very low degree of conservation is displayed in this amino acid stretch for Pop7 and Rpp20 [Figure 1 and ref. (57)].

The Rpp20/Rpp25 heterodimer has been shown to interact with the P3 RNA with a 1:1 stoichiometry, in contrast with the situation in archaeal proteins where

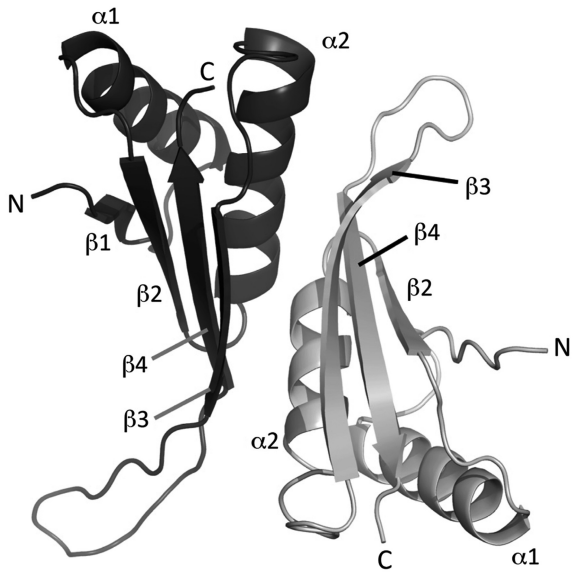

Figure 6. Heterodimer interface analysis for the Alba superfamily of proteins. Ribbon diagram showing the crystal structure of the Alba1/Alba2 complex (PDB ID 2BKY). The secondary-structure elements are indicated. The structure has been displayed using PyMol (http://www.pymol.org/). The solvent-accessible surface area buried by heterodimer formation was calculated by PyMol to be 1384 $Å^2$.

the binding to DNA entails a much higher protein:nucleic acid ratio. This divergence though is not entirely unexpected, and conforms to Alba's main role in DNA packaging and organization. One may speculate that the Rpp20/Rpp25 heterodimer would exhibit a nucleic acid binding mechanism distinct from that of archaeal homologues, following modifications in the course of evolution in response to different functional requirements. Consistent with this idea, the recent evidence for the yeast system showed that Pop6/Pop7 recognize specifically a stem-bulge element in yeast P3 RNase MRP RNA (57) compared to the non-specific double-helix DNA recognition which is characteristic of Alba.

Although the binding to the P3 RNA appears to be the key factor in mediating the association of Rpp20 and Rpp25 to the holoenzyme, previous work has suggested that other protein subunits might further stabilize their assembly into the RNP complex (31). Nonetheless, these investigations through binary screening might not provide a faithful reflection of the molecular recognition process in the context of the holoenzymes, where synergic interactions are likely to take place. In the case of Rpp20 and Rpp25, our results clearly indicate that the outcome of some of the previous investigations might have been perturbed because the interacting entities used were the individual proteins instead of the working unit defined as a pair, and the isolated proteins may be unlikely to retain authentic binding behaviour. Crucially, it may also appear that this is not a prerogative of Rpp20–Rpp25, as other pairs which are likely to perform as a working unit have already been identified, namely Rpp29–Rpp21 (33,34) and Pop5–Rpp30 (35), although this remains to be confirmed by quantitative measurements of their association. The identification of pair-wise interactions between subunits represents an important step forward in the study of the assembly, structure and dynamics of the RNase P and MRP holoenzymes.

In conclusion, the biochemical and biophysical characterization of Rpp20 and Rpp25, and their interaction with RNA will help elucidate the mechanistic basis of their actions and will be an invaluable contribution to understanding the architecture and function of the eukaryotic RNase P and MRP.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to the MRC Biomedical NMR Centre, Mill Hill and its staff, for a generous allocation of NMR time and for expert technical assistance. We thank Drs Tim Welting and Sandy Mattijssen for helpful discussions.

FUNDING

The Wellcome Trust for the Biomolecular Centre for Molecular Spectroscopy (to M.R.C., A.F.D.); EPSRC-case PhD studentship (to K.L.D.H-T.). Funding for open access charge: The Wellcome Trust.

Conflict of interest statement. None declared.

REFERENCES

1. Lukong,K.E., Chang,K.W., Khandjian,E.W. and Richard,S. (2008) RNA-binding proteins in human genetic disease. *Trends Genet.*, **24**, 416–425.
2. Cooper,T.A., Wan,L. and Dreyfuss,G. (2009) RNA and disease. *Cell*, **136**, 777–793.
3. Glisovic,T., Bachorik,J.L., Yong,J. and Dreyfuss,G. (2008) RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett.*, **582**, 1977–1986.
4. Schmitt,M.E. and Clayton,D.A. (1993) Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in *Saccharomyces cerevisiae*. *Mol. Cell Biol.*, **13**, 7935–7941.
5. Lygerou,Z., Allmang,C., Tollervey,D. and Seraphin,B. (1996) Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP *in vitro*. *Science*, **272**, 268–270.
6. Lindahl,L., Bommankanti,A., Li,X., Hayden,L., Jones,A., Khan,M., Oni,T. and Zengel,J.M. (2009) RNase MRP is required for entry of 35S precursor rRNA into the canonical processing pathway. *RNA*, **15**, 1407–1416.
7. Reimer,G., Raska,I., Scheer,U. and Tan,E.M. (1988) Immunolocalization of 7-2-ribonucleoprotein in the granular component of the nucleolus. *Exp. Cell Res.*, **176**, 117–128.
8. Gill,T., Aulds,J. and Schmitt,M.E. (2006) A specialized processing body that is temporally and asymmetrically regulated during the cell cycle in *Saccharomyces cerevisiae*. *J. Cell Biol.*, **173**, 35–45.
9. Kiss,T., Marshallsay,C. and Filipowicz,W. (1992) 7-2/MRP RNAs in plant and mammalian cells: association with higher order structures in the nucleolus. *EMBO J.*, **11**, 3737–3746.
10. Chang,D.D. and Clayton,D.A. (1987) A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication. *EMBO J.*, **6**, 409–417.
11. Gill,T., Cai,T., Aulds,J., Wierzbicki,S. and Schmitt,M.E. (2004) RNase MRP cleaves the CLB2 mRNA to promote cell cycle progression: novel method of mRNA degradation. *Mol. Cell Biol.*, **24**, 945–953.

12. Ridanpää,M., van Eenennaam,H., Pelin,K., Chadwick,R., Johnson,C., Yuan,B., vanVenrooij,W., Pruijn,G., Salmela,R., Rockas,S. *et al*. (2001) Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia. *Cell*, **104**, 195–203.
13. Walker,S.C. and Engelke,D.R. (2006) Ribonuclease P: the evolution of an ancient RNA enzyme. *Crit. Rev. Biochem. Mol. Biol.*, **41**, 77–102.
14. Coughlin,D.J., Pleiss,J.A., Walker,S.C., Whitworth,G.B. and Engelke,D.R. (2008) Genome-wide search for yeast RNase P substrates reveals role in maturation of intron-encoded box C/D small nucleolar RNAs. *Proc. Natl Acad. Sci. USA*, **105**, 12218–12223.
15. Li,X., Frank,D.N., Pace,N., Zengel,J.M. and Lindahl,L. (2002) Phylogenetic analysis of the structure of RNase MRP RNA in yeasts. *RNA*, **8**, 740–751.
16. Walker,S.C. and Avis,J.M. (2005) Secondary structure probing of the human RNase MRP RNA reveals the potential for MRP RNA subsets. *Biochem. Biophys. Res. Commun.*, **335**, 314–321.
17. Walker,S.C., Aspinall,T.V., Gordon,J.M. and Avis,J.M. (2005) Probing the structure of *Saccharomyces cerevisiae* RNase MRP. *Biochem. Soc. Trans.*, **33**, 479–481.
18. Reddy,R. and Shimba,S. (1995) Structural and functional similarities between MRP and RNase P. *Mol. Biol. Rep.*, **22**, 81–85.
19. Forster,A.C. and Altman,S. (1990) Similar cage-shaped structures for the RNA components of all ribonuclease P and ribonuclease MRP enzymes. *Cell*, **62**, 407–409.
20. van Eenennaam,H., Jarrous,N., van Venrooij,W.J. and Pruijn,G.J. (2000) Architecture and function of the human endonucleases RNase P and RNase MRP. *IUBMB Life*, **49**, 265–272.
21. Jarrous,N. (2002) Human ribonuclease P: subunits, function, and intranuclear localization. *RNA*, **8**, 1–7.
22. Lopez,M.D., Rosenblad,M.A. and Samuelsson,T. (2009) Conserved and variable domains of RNase MRP RNA. *RNA Biol.*, **6**, 208–220.
23. Woodhams,M.D., Stadler,P.F., Penny,D. and Collins,L.J. (2007) RNase MRP and the RNA processing cascade in the eukaryotic ancestor. *BMC Evol. Biol.*, **7**, S13.
24. Chamberlain,J.R., Lee,Y., Lane,W.S. and Engelke,D.R. (1998) Purification and characterization of the nuclear RNase P holoenzyme complex reveals extensive subunit overlap with RNase MRP. *Genes Dev.*, **12**, 1678–1690.
25. Welting,T.J., Kikkert,B.J., van Venrooij,W.J. and Pruijn,G.J. (2006) Differential association of protein subunits with the human RNase MRP and RNase P complexes. *RNA*, **12**, 1373–1382.
26. Salinas,K., Wierzbicki,S., Zhou,L. and Schmitt,M.E. (2005) Characterization and purification of *Saccharomyces cerevisiae* RNase MRP reveals a new unique protein component. *J. Biol. Chem.*, **280**, 11352–11360.
27. Schmitt,M.E. and Clayton,D.A. (1994) Characterization of a unique protein component of yeast RNase MRP: an RNA-binding protein with a zinc-cluster domain. *Genes Dev.*, **8**, 2617–2628.
28. Houser-Scott,F., Xiao,S., Millikin,C.E., Zengel,J.M., Lindahl,L. and Engelke,D.R. (2002) Interactions among the protein and RNA subunits of *Saccharomyces cerevisiae* nuclear RNase P. *Proc. Natl Acad. Sci. USA*, **99**, 2684–2689.
29. Jiang,T. and Altman,S. (2001) Protein-protein interactions with subunits of human nuclear RNase P. *Proc. Natl Acad. Sci. USA*, **98**, 920–925.
30. Jiang,T., Guerrier-Takada,C. and Altman,S. (2001) Protein-RNA interactions in the subunits of human nuclear RNase P. *RNA*, **7**, 937–941.
31. Welting,T.J., van Venrooij,W.J. and Pruijn,G.J. (2004) Mutual interactions between subunits of the human RNase MRP ribonucleoprotein complex. *Nucleic Acids Res.*, **32**, 2138–2146.
32. Aspinall,T.V., Gordon,J.M., Bennett,H.J., Karahalios,P., Bukowski,J.P., Walker,S.C., Engelke,D.R. and Avis,J.M. (2007) Interactions between subunits of *Saccharomyces cerevisiae* RNase MRP support a conserved eukaryotic RNase P/MRP architecture. *Nucleic Acids Res.*, **35**, 6439–6450.
33. Honda,T., Kakuta,Y., Kimura,K., Saho,J. and Kimura,M. (2008) Structure of an archaeal homolog of the human protein complex Rpp21-Rpp29 that is a key core component for the assembly of active ribonuclease P. *J. Mol. Biol.*, **384**, 652–662.
34. Xu,Y., Amero,C.D., Pulukkunat,D.K., Gopalan,V. and Foster,M.P. (2009) Solution structure of an archaeal RNase P binary protein complex: formation of the 30-kDa complex between *Pyrococcus furiosus* RPP21 and RPP29 is accompanied by coupled protein folding and highlights critical features for protein-protein and protein-RNA interactions. *J. Mol. Biol.*, **393**, 1043–1055.
35. Kawano,S., Nakashima,T., Kakuta,Y., Tanaka,I. and Kimura,M. (2006) Crystal structure of protein Ph1481p in complex with protein Ph1877p of archaeal RNase P from *Pyrococcus horikoshii* OT3: implication of dimer formation of the holoenzyme. *J. Mol. Biol.*, **357**, 583–591.
36. Welting,T.J., Peters,F.M., Hensen,S.M., van Doorn,N.L., Kikkert,B.J., Raats,J.M., van Venrooij,W.J. and Pruijn,G.J. (2007) Heterodimerization regulates RNase MRP/RNase P association, localization, and expression of Rpp20 and Rpp25. *RNA*, **13**, 65–75.
37. Welting,T.J., Mattijssen,S., Peters,F.M., van Doorn,N.L., Dekkers,L., van Venrooij,W.J., Heus,H.A., Bonafe,L. and Pruijn,G.J. (2008) Cartilage-hair hypoplasia-associated mutations in the RNase MRP P3 domain affect RNA folding and ribonucleoprotein assembly. *Biochim. Biophys. Acta*, **1783**, 455–466.
38. Gasteiger,E., Gattiker,A., Hoogland,C., Ivanyi,I., Appel,R.D. and Bairoch,A. (2005) In Walker,J.M. (ed.), *The Proteomics Protocols Handbook*. Humana Press, NJ, pp. 571–607.
39. Conte,M.R., Conn,G.L., Brown,T. and Lane,A.N. (1997) Conformational properties and thermodynamics of the RNA duplex r(CGCAAAUUUGCG)2: comparison with the DNA analogue d(CGCAAATTTGCG)2. *Nucleic Acids Res.*, **25**, 2627–2634.
40. SantaLucia,J. Jr, Allawi,H.T. and Seneviratne,P.A. (1996) Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry*, **35**, 3555–3562.
41. Burrows,S.D., Doyle,M.L., Murphy,K.P., Franklin,S.G., White,J.R., Brooks,I., McNulty,D.E., Scott,M.O., Knutson,J.R., Porter,D. *et al*. (1994) Determination of the monomer-dimer equilibrium of interleukin-8 reveals it is a monomer at physiological concentrations. *Biochemistry*, **33**, 12741–12745.
42. Fox,M.F., Barker,B.E. and Hayon,E. (1978) Far-ultraviolet solution spectroscopy of chloride ion. *J. Chem. Soc., Faraday Trans. 1*, **74**, 1776–1785.
43. Whitmore,L. and Wallace,B.A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.*, **32**, W668–W673.
44. Martino,L., He,Y., Hands-Taylor,K.L., Valentine,E.R., Kelly,G., Giancola,C. and Conte,M.R. (2009) The interaction of the *Escherichia coli* protein SlyD with nickel ions illuminates the mechanism of regulation of its peptidyl-prolyl isomerase activity. *FEBS J.*, **276**, 4529–4544.
45. Shi,J., Ghirlando,R., Beavil,R.L., Beavil,A.J., Keown,M.B., Young,R.J., Owens,R.J., Sutton,B.J. and Gould,H.J. (1997) Interaction of the low-affinity receptor CD23/Fc epsilonRII lectin domain with the Fc epsilon3-4 fragment of human immunoglobulin E. *Biochemistry*, **36**, 2112–2122.
46. Aravind,L., Iyer,L.M. and Anantharaman,V. (2003) The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism. *Genome Biol.*, **4**, R64.
47. Sandman,K. and Reeve,J.N. (2000) Structure and functional relationships of archaeal and eukaryal histones and nucleosomes. *Arch. Microbiol.*, **173**, 165–169.
48. Wardleworth,B.N., Russell,R.J., Bell,S.D., Taylor,G.L. and White,M.F. (2002) Structure of Alba: an archaeal chromatin protein modulated by acetylation. *EMBO J.*, **21**, 4654–4662.
49. Rost,B., Yachdav,G. and Liu,J. (2004) The PredictProtein server. *Nucleic Acids Res.*, **32**, W321–W326.
50. Jacks,A., Babon,J., Kelly,G., Manolaridis,I., Cary,P.D., Curry,S. and Conte,M.R. (2003) Structure of the C-terminal domain of human La protein reveals a novel RNA recognition motif

coupled to a helical nuclear retention element. *Structure*, **11**, 833–843.
51. Xie,D. and Freire,E. (1994) Structure based prediction of protein folding intermediates. *J. Mol. Biol.*, **242**, 62–80.
52. Jelinska,C., Conroy,M.J., Craven,C.J., Hounslow,A.M., Bullough,P.A., Waltho,J.P., Taylor,G.L. and White,M.F. (2005) Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a mechanism for control of DNA packaging. *Structure*, **13**, 963–971.
53. Perederina,A., Esakova,O., Koc,H., Schmitt,M.E. and Krasilnikov,A.S. (2007) Specific binding of a Pop6/Pop7 heterodimer to the P3 stem of the yeast RNase MRP and RNase P RNAs. *RNA*, **13**, 1648–1655.
54. Maida,Y., Yasukawa,M., Furuuchi,M., Lassmann,T., Possemato,R., Okamoto,N., Kasim,V., Hayashizaki,Y., Hahn,W.C. and Masutomi,K. (2009) An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. *Nature*, **461**, 230–235.
55. Guerrier-Takada,C., Eder,P.S., Gopalan,V. and Altman,S. (2002) Purification and characterization of Rpp25, an RNA-binding protein subunit of human ribonuclease P. *RNA*, **8**, 290–295.
56. Yuan,Y., Tan,E. and Reddy,R. (1991) The 40-kilodalton to autoantigen associates with nucleotides 21 to 64 of human mitochondrial RNA processing/7-2 RNA *in vitro*. *Mol. Cell Biol.*, **11**, 5266–5274.
57. Perederina,A., Esakova,O., Quan,C., Khanova,E. and Krasilnikov,A.S. (2010) Eukaryotic ribonucleases P/MRP: the crystal structure of the P3 domain. *EMBO J.*, **29**, 761–769.