JMB

Structural Characterization of the RNase E S1 Domain and Identification of its Oligonucleotide-binding and Dimerization Interfaces

Mario Schubert[1,2], Robert E. Edge[1], Paula Lario[1], Michael A. Cook[1]
Natalie C. J. Strynadka[1,3], George A. Mackie[1] and
Lawrence P. McIntosh[1,2,3,4]*

[1]*Department of Biochemistry and Molecular Biology, The University of British Columbia Vancouver, BC, Canada V6T 1Z3*

[2]*The Protein Engineering Network of Centres of Excellence, The University of British Columbia, Vancouver BC, Canada V6T 1Z3*

[3]*The Biotechnology Laboratory The University of British Columbia, Vancouver, BC Canada V6T 1Z3*

[4]*Department of Chemistry, The University of British Columbia Vancouver, BC, Canada V6T 1Z3*

S1 domains occur in four of the major enzymes of mRNA decay in *Escherichia coli*: RNase E, PNPase, RNase II, and RNase G. Here, we report the structure of the S1 domain of RNase E, determined by both X-ray crystallography and NMR spectroscopy. The RNase E S1 domain adopts an OB-fold, very similar to that found with PNPase and the major cold shock proteins, in which flexible loops are appended to a well-ordered five-stranded β-barrel core. Within the crystal lattice, the protein forms a dimer stabilized primarily by intermolecular hydrophobic packing. Consistent with this observation, light-scattering, chemical crosslinking, and NMR spectroscopic measurements confirm that the isolated RNase E S1 domain undergoes a specific monomer–dimer equilibrium in solution with a K_D value in the millimolar range. The substitution of glycine 66 with serine dramatically destabilizes the folded structure of this domain, thereby providing an explanation for the temperature-sensitive phenotype associated with this mutation in full-length RNase E. Based on amide chemical shift perturbation mapping, the binding surface for a single-stranded DNA dodecamer ($K_D = 160(\pm 40)$ μM) was identified as a groove of positive electrostatic potential containing several exposed aromatic side-chains. This surface, which corresponds to the conserved ligand-binding cleft found in numerous OB-fold proteins, lies distal to the dimerization interface, such that two independent oligonucleotide-binding sites can exist in the dimeric form of the RNase E S1 domain. Based on these data, we propose that the S1 domain serves a dual role of dimerization to aid in the formation of the tetrameric quaternary structure of RNase E as described by Callaghan *et al*. in 2003 and of substrate binding to facilitate RNA hydrolysis by the adjacent catalytic domains within this multimeric enzyme.

Corresponding author

Keywords: RNase E; S1 domain; OB-fold; RNA binding; protein structure

Introduction

The metabolic instability of mRNA is an important, albeit infrequently appreciated, aspect of gene expression. mRNA lifetimes are short relative to cellular doubling times, whereas "stable RNAs" such as tRNA and rRNA remain intact and functional for at least several generations. The stability of an mRNA controls its rate of accumulation and its maximal steady-state level independently of promoter strength. In the current model of mRNA decay in *Escherichia coli*, the critical first step is

usually an endonucleolytic cleavage catalyzed by RNase E.[1−3] In addition to its crucial role in mRNA degradation, this enzyme also participates in the maturation of rRNA, tRNA, and other small RNAs. The specificity of RNase E has been addressed several times, yet a simple consensus sequence has never been found. Although RNase E seems to have a broad specificity with a preference for AU-rich sequences,[4] a G in position −2 relative to the cleavage site can significantly improve the efficiency of cleavage. More significantly, the rate of substrate hydrolysis by RNase E is strongly influenced by RNA elements that are distinct from the actual site of cleavage, such as adjacent stem−loops, phosphorylation at the 5′ terminus, and signals for translation initiation.

RNase E can be divided into three functionally distinct regions: an N-terminal segment (residues 1−498) containing the single-strand-specific endonuclease activity;[5] an arginine and proline-rich central region (residues ∼500−650); and a C-terminal portion (residues ∼650−1061) that provides a scaffold on which RhIB, enolase, and PNPase assemble.[6,7] The N-terminal segment of RNase E can process rRNA and is sufficient to support cell growth in the absence of the full-length protein.

Based on sequence comparisons, residues 37−119 in the N-terminal segment of RNase E are predicted to correspond to an S1 domain.[8] In particular, these residues show 30% sequence identity with the S1 domain of PNPase, for which a three-dimensional structure has been determined by NMR spectroscopy.[8] S1 domains are RNA-binding modules, originally identified in the ribosomal protein S1, that represent a subclass of the (oligonucleotide/oligosaccharide-binding fold) (OB-fold) family.[9] This fold, built on a distinctive β-sheet scaffold with appended loops or helices, is found in a wide variety of proteins that are involved principally in the recognition of carbohydrates and nucleic acids.[10,11]

Beyond an expected role in mediating RNA binding, the precise function of the S1 domain in RNase E is not understood. Two separate point mutations within this domain, *rne-3071* (L68F) and *ams-1* (G66S), are both lethal at elevated temperatures, indicating that the S1 domain is crucial for the activity of RNase E.[12] More recently, using homology modeling and site-directed mutagenesis, Diwa *et al.*[13] identified two distinct surface regions of the S1 domain. One is of general importance for its ribonuclease activity, while the other is entirely dispensable for catalysis *in vitro* yet important for feedback regulation of the expression of *rne*, the gene encoding this enzyme. Thus, based on even the limited information available to date, it would appear that the role of the RNase E S1 domain may be more complex than simply providing an RNA-binding surface.

So far there is no experimentally determined three-dimensional structural model available for any part of RNase E. A very recent biophysical study indicates that the N-terminal segment of RNase E (residues 1−529) forms a tetramer and that the S1 domains are readily susceptible to proteolytic attack. Although the quality of the crystals formed by this fragment were insufficient for solving its high-resolution structure, using X-ray solution scattering and crystallographic data, Callaghan *et al.*[14] proposed that four RNase E monomers pack with approximate D_2 point symmetry. A salient feature of their model involved the positioning of the S1 domains at the periphery of the tetramer; i.e. at maximal separation from each other. To better understand the role of the S1 domain in RNase E, we have determined its structure using both X-ray crystallography and NMR spectroscopy. In parallel, we have demonstrated that the isolated domain weakly dimerizes *in vitro*, and have identified its oligonucleotide and dimerization interfaces. Here, the implications of these results for understanding the quaternary structure of RNase E and the functional role played by this domain are presented.

Results

Using the SMART database,[15] residues 37−119 of RNase E were predicted to form an S1 domain. Allowing for possible variation in its precise boundaries, the DNA encoding residues 35−125 of RNase E was cloned into the expression vector pET15b. The resulting construct, denoted RneS1$^{35−125}$, encompasses the entire sequence of the predicted S1 domain, flanked by a small number of N and C-terminal residues. After expression, purification, and cleavage of a His$_6$-affinity tag, the construct produced a soluble, well-folded protein as judged by the data presented below. This confirms that residues 35−125 encompass a structural domain within RNase E.

Structure determination by X-ray crystallography

During the determination of the solution structure of the RNase E S1 domain, by serendipity, RneS1$^{35−125}$ crystallized in the buffer used for these NMR spectroscopic studies. Therefore, we undertook a parallel structural analysis by X-ray crystallography. Although the NMR-derived solution structure of RneS1$^{35−125}$ was calculated independently and before completion of the X-ray crystallographically derived structure, for simplicity we will begin by describing the structure of this protein in the crystalline state.

The crystal structure of RneS1$^{35−125}$ was determined to 2.0 Å resolution, using single anomalous dispersion (SAD) of a trimethyl lead(IV) acetate derivative. The crystallographic R and R_{free} values of 18.2% and 23.2%, respectively, and the stereochemical parameters are in the range expected for structures determined at a comparable resolution (Table 1). The asymmetric unit contains two copies of RneS1$^{35−125}$, denoted monomer A (residues 39−125)

Table 1. X-ray crystallographic statistics for RneS1$^{35-125}$

	Native	Pb derivative
A. *Data collection*		
Wavelength (λ, Å)	1.0781	0.947390
Space group	$P4_12_12$	$P4_12_12$
Cell axes *a, b, c* (Å)	70.4/70.4/ 87.9	70.6/70.6/ 87.9
Resolution (Å)^a	50–1.8 (1.86–1.80)	25–2.0 (2.07–2.00)
No. of observed reflections	162,333	106,555
No. of unique reflections	20,912 (1921)	15,387 (1309)
$\langle I/\sigma \rangle^a$	18.3 (4.65)	10.8 (2.6)
$R_{sym}{}^{a,b}$	0.050 (0.378)	0.091 (0.338)
Overall completeness (%)^a	99.0 (93.5)	98.5 (86.7)
B. *Refinement*		
$R/R_{free}{}^c$	0.199/ 0.231	0.182/ 0.232
rms deviations, bond lengths (Å)/ angles (deg.)	0.013/1.3	0.014/1.4
Average *B*-factor (Å²), main-chain/side-chain/water/lead	30.6/33.4/ 46.2/ –	23.0/25.4/ 35.2/31.5
Side-chains with multiple conformations	1 (Y82A)	2 (Y82A, E50B)
Protein atoms	1327	1327
Water molecules with occupancy 1.0 or 0.5	203/2	224/5
Pb ions	–	2
Ramachandran region population (%) for non-glycine, non-proline residues		
Most favored regions	94.2	92.1
Additionally allowed regions	5.8	7.9
Generously allowed regions	0	0
Forbidden regions	0	0

^{a} Values in parentheses correspond to the highest resolution shell.
^{b} $R_{sym} = \Sigma|I_j - \langle I \rangle|/\Sigma\langle I \rangle$, where I_j is the intensity for reflection *j*, and *I* is the mean intensity.
^{c} $R = ||F_o| - |F_c||/|F_c|$, calculated with the working set. R_{free} is similarly calculated but with 5% of the data excluded from the calculation of *R*.

and monomer B (residues 39–80 and 88–125), related by a non-crystallographic 2-fold axis. Residues N-terminal to A39 in both monomers, as well as residues 81–87 in monomer B, are apparently disordered, giving no detectable electron density. The two superimposed monomers are shown in Figure 1a, while their main-chain rms deviations and *B*-factors are presented in Figure 2.

RneS1$^{35-125}$ is composed of five β-strands, hydrogen bonded in a 1-2-3-5-4-1 topology diagnostic of the OB-fold (β1 residues 42–50; β2 55–59; β3 66–69; β4 99–106; β5 115–117; Figure 1c). The strands contain two bulges, located around residues 47/48 and 105/106, which allow β1 and β4 to kink and thereby to form a closed β-barrel. In comparison with the prototypical OB-fold of a five-stranded mixed β-barrel capped on one end with an α-helix, RneS1$^{35-125}$ shows some distinct features. The most notable deviation is the absence of an α-helix connecting strands β3 and β4. RneS1$^{35-125}$ uses instead three short 3_{10}-helices, or equivalently type III β-turns (70–72, 75–77, 90–

92), and two additional β-turns (83–86 type VIII, 95–98 type II) in an exposed loop denoted L_{34}. The position of strand β5 varies within the members of the OB-fold family, being hydrogen bonded with β3 in a parallel orientation to extend the β-sheet formed by β1–β2–β3, or pairing antiparallel with β4 to expand the β1–β4 sheet, or both. In RneS1,$^{35-125}$ β5 is very short, yet still hydrogen bonds with both β3 and β4. However, instead of pairing with β4 in a more extended fashion, the residues following β5 turn outwards to form a dimerization interface (see below).

Although monomers A and B superimpose well over their core β-strands (rmsd 0.24 Å on 116 main-chain atoms), there are several noteworthy differences between the two RneS1$^{35-125}$ molecules within the asymmetric unit (Figures 1 and 2). In addition to the dramatic disorder of residues 81–87 in monomer B, the flanking regions within the exposed L_{34} loop (residues 79, 80, 88, and 89) also adopt different conformations due to crystal packing. L_{34} in monomer A packs against another dimer, whereas in monomer B it faces towards a solvent channel. Furthermore, the single-turn 3_{10} helix formed by residues 90–92 at the end of this loop in monomer A is a distorted α-helix (residues 90–93) in monomer B. The position of the loop L_{12}, which also contains a short 3_{10} helix (51–53), is different between the monomers due to crystal packing. Finally, the precise conformations of residues in loops L_{23} and L_{45} differ, likely due to local dynamic behavior as evident by their elevated main-chain *B*-values, as well as their high rms deviations within the ensemble of NMR-derived structures and their distinctive ^{15}N relaxation properties (see below).

Structure determination by NMR spectroscopy

Nearly complete assignments of the ^{1}H, ^{13}C and ^{15}N resonances of nuclei in the backbone and side-chains of RneS1$^{35-125}$ were obtained through a standard set of ^{1}H/^{13}C/^{15}N scalar correlation experiments. This process was greatly facilitated by using a suite of MUSIC experiments that provided amino acid-selective identification of many peaks within the ^{1}H–^{15}N HSQC spectrum of the protein.[16,17] Under the experimental conditions the N-terminal residues GSHM, as well as S52 and N111, likely display rapid hydrogen exchange and thus do not have detectable ^{1}H^N resonances. Residues H65, H85, G86, and G110 also show weak broad ^{1}H^N signals.

The tertiary structure of RneS1$^{35-125}$ in solution was calculated independently with ARIA/CNS version 1.2[18] using distance restraints from an extensive set of 3D ^{15}N and ^{13}C-resolved NOESY spectra, combined with TALOS-derived[19] backbone dihedral angle restraints, *J* coupling-derived χ_1 dihedral angle restraints, and a limited set of hydrogen bond restraints. A summary of the structural refinement statistics, corresponding to the final energetically best ten structures, is listed
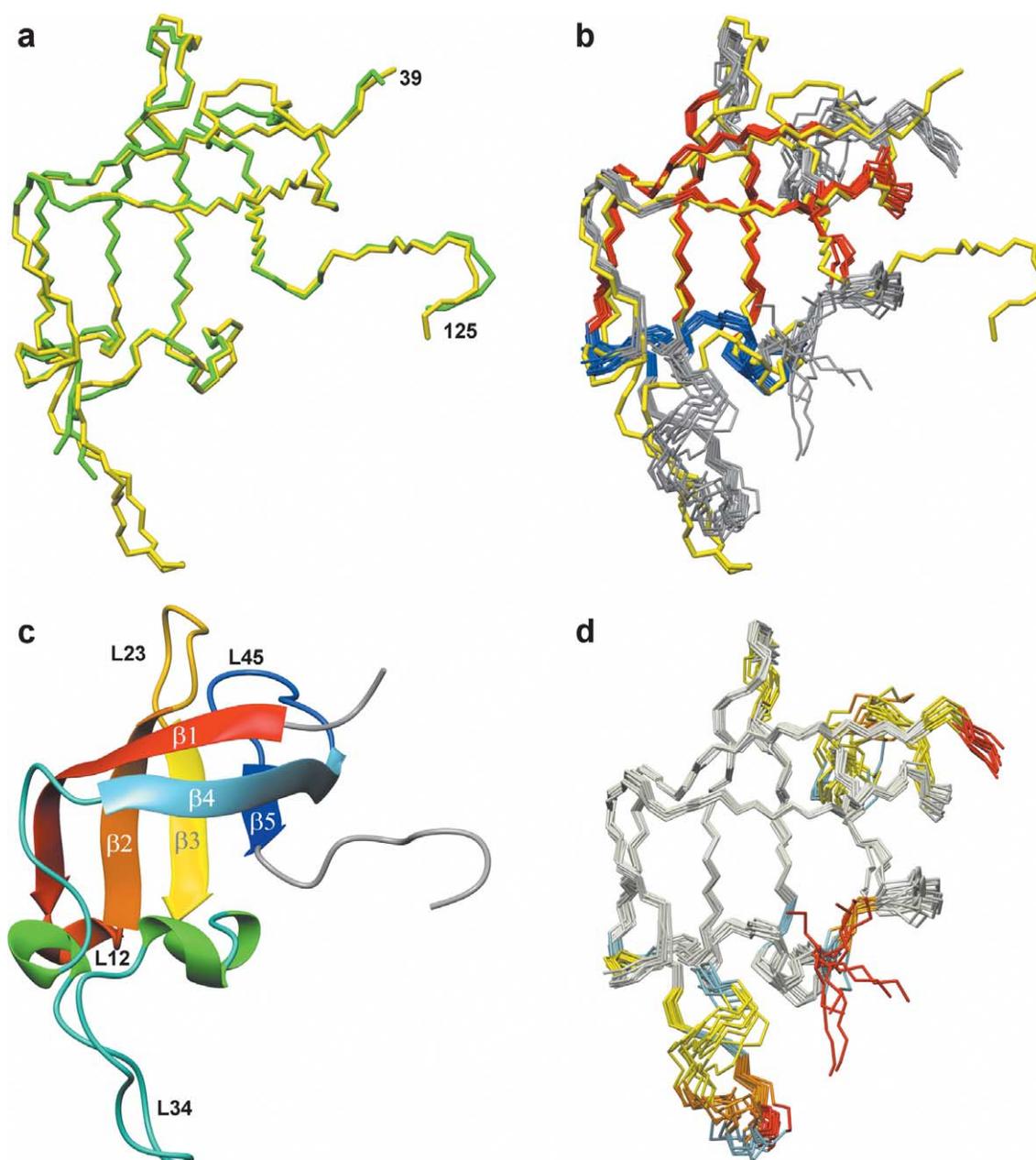
Figure 1. Tertiary structure of RneS1.[35−125] a, A backbone representation of the crystal structure of the Pb-derivative of RneS1[35−125] showing a superimposition of monomers A (yellow) and B (green) from the asymmetric unit. b, A backbone representation of the ensemble of ten NMR-derived structures of RneS1,[35−125] superimposed on monomer A (yellow) from the crystal structure. The secondary structure of the ensemble is highlighted with β-strands in red and helices in blue. c, A cartoon representation, using chain A, of the secondary structure of RneS1[35−125]. The β-strands (arrows) are named consecutively β1 to β5, while the intervening loops are L_{12} through L_{45}. As drawn, the RNA-binding interface lies on the back of the molecule and the dimerization interface in front. d The amides of RneS1[35−125] showing enhanced mobility on a sub-nanosecond timescale, as evident by reduced ^{1}H{^{15}N}-NOE values, are highlighted in color on the backbone representation of the ensemble of NMR-derived structures. Residues with ^{1}H{^{15}N}-NOE values of <0.2 (highest mobility) are shown in red, between 0.2 and 0.4 in orange, between 0.4 and 0.6 in yellow, and over 0.6 in gray. Proline residues or residues where data are not available (e.g. due to spectral overlap) are identified in light blue.

in Table 2. The structural ensemble is shown in Figure 1b and d.

Overall, the NMR-derived secondary and tertiary structure of RneS1[35−125] is very similar to that determined by X-ray crystallography. The rms deviation of the core β-strands between the average NMR structure and the monomer A of the crystal structure is 0.67 Å for the 116 main-chain and 1.95 Å for the 110 side-chain heavy atoms. However, variations both within the ensemble of solution structures and between the solution-state and crystal-state structures are evident for residues
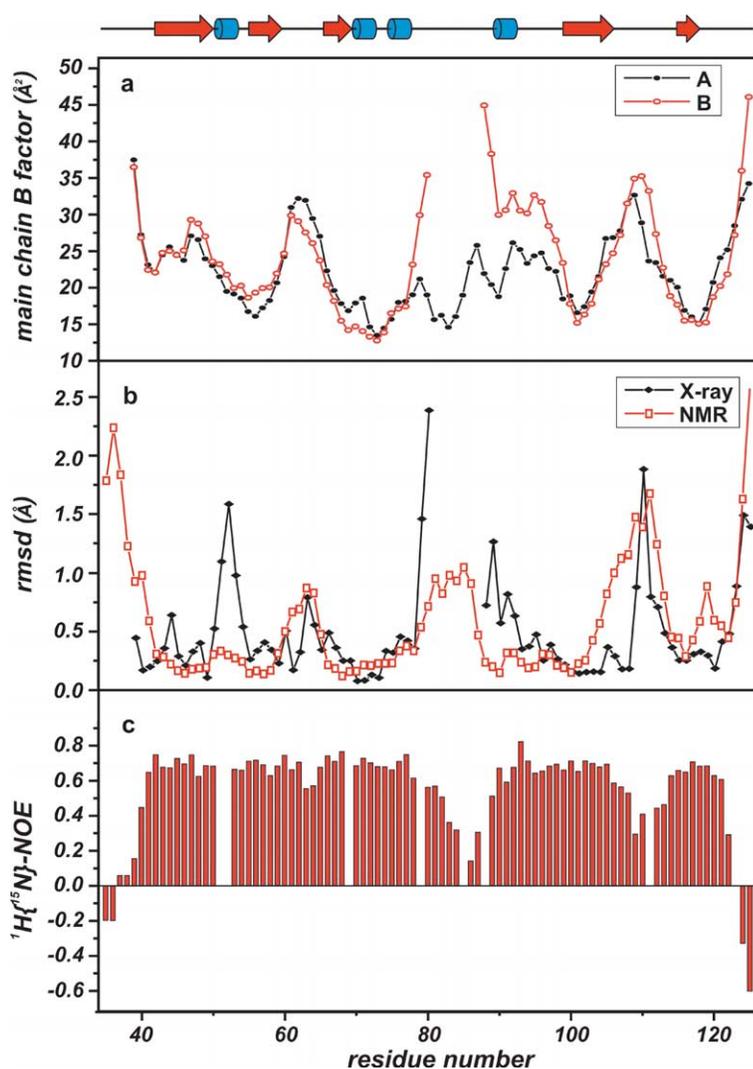
Figure 2. The termini and loop regions of RneS1$^{35-125}$ exhibit conformational mobility as evident by both X-ray crystallography and NMR spectroscopy. a, The main-chain (C^α,C',N) B-factors for monomers A and B in the crystal structure of Pb-derivative are plotted *versus* residue number. b, The C^α rms deviations between chains A and B in the X-ray crystallographically determined structure of the Pb-derivative are plotted *versus* residue number, along with the main-chain (C^α,C',N) rms deviations between members of the NMR-derived ensemble and the average solution structure of this protein. c, The backbone heteronuclear ^{1}H{^{15}N}-NOE relaxation values of RneS1$^{35-125}$ acquired on a 600 MHz spectrometer at a protein concentration of 1.5 mM and 30 °C. A cartoon representation of the secondary structure of RneS1$^{35-125}$ is shown on the top with β-strands as arrows and helices as cylinders.

at the termini of the protein, as well as in the loops L_{23}, L_{45} and especially L_{34}. As discussed below, NMR relaxation measurements indicate that this reflects the conformational mobility of these exposed amino acid residues (Figure 2). In addition, the C-terminal residues of RneS1$^{35-125}$ fold back to the third 3_{10} helix (75–77), rather than forming intermolecular contacts within the crystal dimer.

Although the structures of RneS1$^{35-125}$ determined by NMR spectroscopy and X-ray crystallography are similar, the NMR-derived coordinates could not be used successfully for molecular replacement to overcome the crystallographic phase problem. In addition to the recognized difficulties in using NMR-based structures for molecular replacement,[20] the uniform nature of the barrel-shaped RneS1$^{35-125}$ is especially problematic in allowing for the unique determination of the correct rotation and translation functions in this phasing method.

Dimerization of RneS1$^{35-125}$

Crystalline RneS1$^{35-125}$ forms a homodimer within the asymmetric unit. As illustrated in Figure 3, the dimer interface is formed by the packing of surface side-chains from strands β1 and β4 (A39, I41, K43, E99, I101), the short 3_{10} helix (E76, Y77) in L_{34}, and the C terminus (I120, L122). In addition, three pairs of intermolecular hydrogen bonds (K43 N$^\zeta$/S120 O; E99 O$^\varepsilon$/A123 N; E99 N/A123 O), and a bridging water molecule between the side-chains of E76 in both monomers, was observed. The formation of this interface leads to the burial of a total $(A + B)$ surface area of 1380 Å^2, of which 1020 Å^2 is associated with hydrophobic residues. Given that this value, which represents 12.2% of the total surface area of the two monomers, is at the lower end of the range observed for many homodimeric protein–protein interfaces,[21] we undertook further studies to determine the oligomerization state of RneS1$^{35-125}$ in solution. Initially, light-scattering measurements with protein at concentrations ranging from 50 μM to 300 μM revealed a monodisperse species with an apparent molecular mass of approximately 12 kDa. These data (not shown) indicate that RneS1$^{35-125}$ adopts a predominantly monomeric state when in dilute solution. However, upon mild

Table 2. NMR spectroscopic statistics for RneS1$^{35-125}$

Summary of restraints	
Unambiguous (ambiguous) distance restraints assigned/used by ARIA	2113 (597)
Average number of unambiguous distance restraints per residue	22.0
Dihedral restraints $\phi/\psi/\chi_1$	67/67/25
Hydrogen bond restraints	15
Deviation from restraints	
Averaged rmsd from distance restraints (Å)	0.026 ± 0.0006
rmsd from experimental torsion angle restraints (deg.)	0.93 ± 0.084
Number of distance restraint violations >0.5 Å	0 ± 0
Number of torsion angle restraint violations >0.5°	1.4 ± 0.685
Average rms deviations from idealized covalent geometry	
Bonds (Å)	0.0046 ± 0.0001
Angles (deg.)	0.589 ± 0.018
Impropers (deg.)	2.011 ± 0.17
Atomic rms deviations versus average structure (Å)	
Heavy atoms of core residues[a]	0.538 ± 0.065
Backbone atom of core residues[a]	0.20 ± 0.027
Heavy atoms of all residues	1.85 ± 0.23
Backbone atoms of all residues	1.57 ± 0.33
Ramachandran region population (%) for non-glycine, non-proline residues	
Most favored regions	72.4
Additionally allowed regions	24.6
Generously allowed regions	1.7
Forbidden regions	1.2

[a] Core residues (13–21, 26–30, 37–43, 70–77, 86–88) were used to superimpose the structure ensemble.

treatment with glutaraldehyde, 45 μM RneS1$^{35-125}$ readily, although not quantitatively, formed cross-linked homodimers (Figure 4, lanes 2 and 6). In contrast, under the same conditions, the isolated S1 domain from PNPase underwent crosslinking to itself with much less efficiency (lanes 3 and 7) and did not form heterodimers with RneS1$^{35-125}$ (lanes 4 and 8). These data reveal that RneS1$^{35-125}$ has the propensity to specifically self-associate *in vitro*.

To further investigate the oligomerization of RneS1$^{35-125}$, ^{1}H–^{15}N HSQC spectra of the protein were measured at sample concentrations ranging from 0.1 mM to 1.2 mM. Upon dilution, the signals from several amides showed clear, albeit small (e.g. ~12 Hz in ^{1}H), changes in chemical shift (Figure 5a and c). When mapped onto the crystal structure of RneS1$^{35-125}$, the amides exhibiting such chemical shift perturbations all cluster within the dimer interface (Figure 3b and c). This strongly indicates that, in solution, RneS1$^{35-125}$ exits in an equilibrium between monomeric and dimeric forms, the latter adopting a structure represented closely by that observed within the crystal lattice. Furthermore, since the chemical shift perturbations occur across the entire concentration range examined, the K_D for this equilibrium must be of the order of millimolar. Additional support for this conclusion is provided below by ^{15}N relaxation measurements, as discussed below.

Dynamics of RneS1$^{35-125}$ from ^{15}N relaxation measurements

Amide ^{15}N T_1 and T_2 lifetimes and steady-state ^{1}H{^{15}N}-NOE values were measured at concentrations of 0.6 mM and 1.2 mM for 81 backbone amide nitrogen atoms of RneS1$^{35-125}$. Excluding amides exhibiting anomalous T_1/T_2 ratios or ^{1}H{^{15}N}-NOE values indicative of internal mobility, the average T_1 and T_2 lifetimes, with standard deviations, were $0.48(\pm 0.03)$ second and $0.097(\pm 0.01)$ second at 0.6 mM and $0.52(\pm 0.02)$ second and $0.087(\pm 0.01)$ second at 1.2 mM, respectively. These lifetimes correspond to isotropic correlation times of $7.4(\pm 0.03)$ ns and $8.3(\pm 0.02)$ ns for global tumbling at 0.6 mM and 1.2 mM, respectively. The values of these correlation times, which are higher than expected for a monomeric 10.6 kDa protein, yet lower than expected for a dimeric protein,[22] combined with their concentration-dependence provide further evidence of an equilibrium between monomeric and dimeric forms of RneS1$^{35-125}$ with a K_D in the millimolar range.

In addition to yielding information about the global hydrodynamic properties of a protein, ^{15}N relaxation measurements also provide insights into local backbone mobility. Due to the equilibrium between the monomeric and dimeric states of RneS1$^{35-125}$, an analysis of these data by the usual Lipari–Szabo model-free formalism to obtain local order parameters cannot be applied.[23] However, the steady-state ^{1}H{^{15}N}-NOE is a sensitive indicator of motions on a sub-nanosecond time-scale. As shown in Figure 2c and mapped onto the structure of RneS1$^{35-125}$ in Figure 1d, residues at the N and C termini of the protein, as well as in the L$_{34}$ (residues 82–89) and L$_{45}$ (residues 108–133) loops, and to a lesser extent, the L$_{23}$ loop, exhibit reduced ^{1}H{^{15}N}-NOE values indicative of local flexibility. Excluding the L$_{12}$ loop, which appears relatively well-ordered in solution yet differs in conformation between monomers A and B due to crystal packing, there is a close correlation between the residues with reduced ^{1}H{^{15}N}-NOE values and those with elevated crystallographic B-factors, with high rms deviations between the monomers in the crystal dimer, and with high rms deviations between members of the NMR-derived structural ensemble (Figure 2). This strongly indicates that the ^{1}H{^{15}N}-NOE values, crystallographic B-factors, and structural rms deviations all reflect the local mobility of the backbone of RneS1$^{35-125}$. A similar correlation has been observed between the crystallographically and spectroscopically determined structures of the major *E. coli* cold-shock protein CspA,[24] which adopts an OB-fold closely related to that of RneS1$^{35-125}$.

Temperature-sensitive mutants of RNase E

Previous studies indicated that the mutations G66S and L68F impart a temperature-sensitive
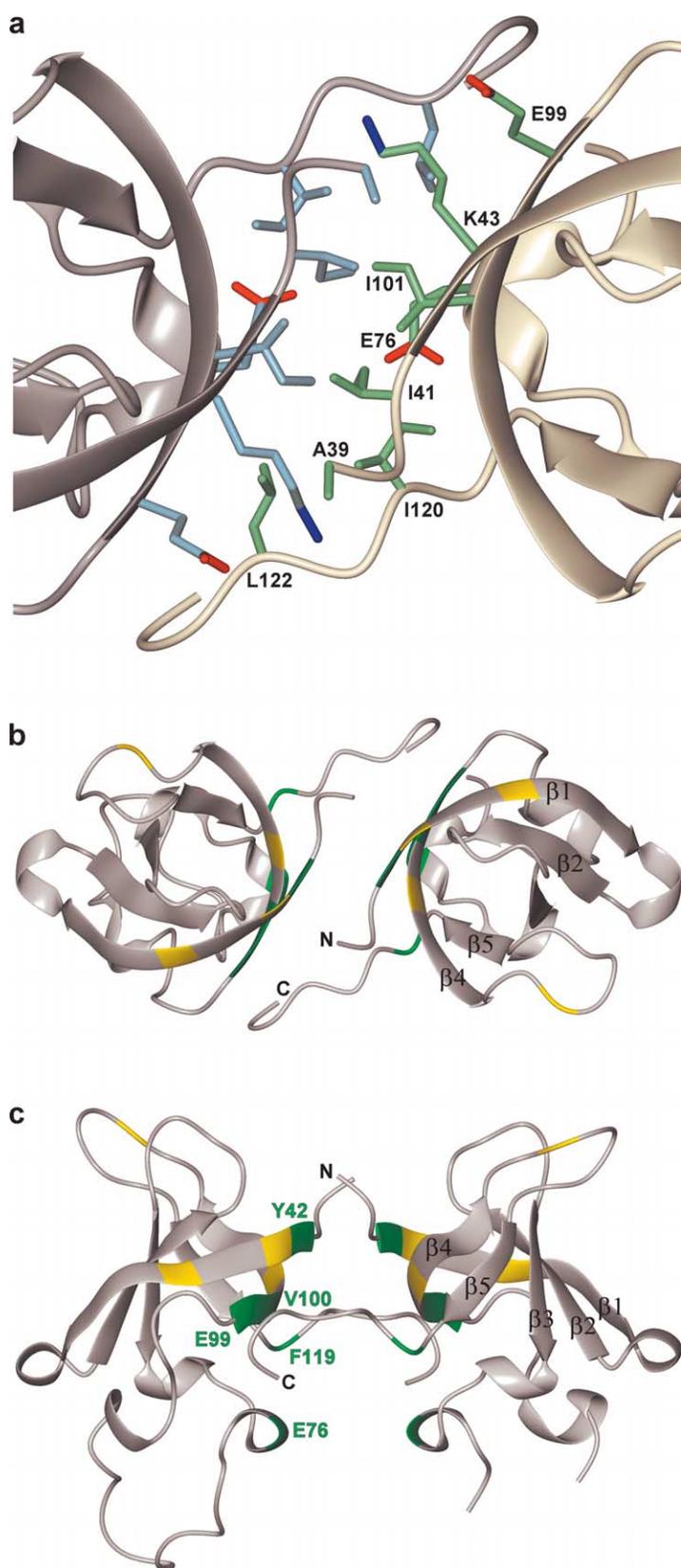
Figure 3. Quaternary structure of the crystallographic dimer of RneS1[35−125] presented as a ribbon diagram. a, An expanded view of side-chains contributing to the dimerization interface with monomer A on the left and B on the right. As shown by the top b, and side c, views of the dimer, amides experiencing significant ^{1}H–^{15}N chemical shift perturbations upon changing the protein concentration from 0.1 mM to 1.2 mM (see Figure 6) map to this interface (green, $\Delta\delta > 10$ Hz; yellow, 7 Hz $< \Delta\delta > 10$ Hz). The nucleic acid-binding interface lie along the shallow groove opposite to the dimerization interface (see Figure 6).

phenotype on the activity of RNase E.[12] From the structure of RneS1[35−125], residues 66 and 68 are located with strand β3, such that substitution with Ser or Phe, respectively, would likely disrupt the packing of the hydrophobic core of the protein

(Figure 7b). To test this hypothesis, the G66S substitution was introduced into RneS1[35−125] and CD spectropolarimetry was used to monitor the conformation of wild-type (WT) and mutant proteins as a function of temperature. Whereas
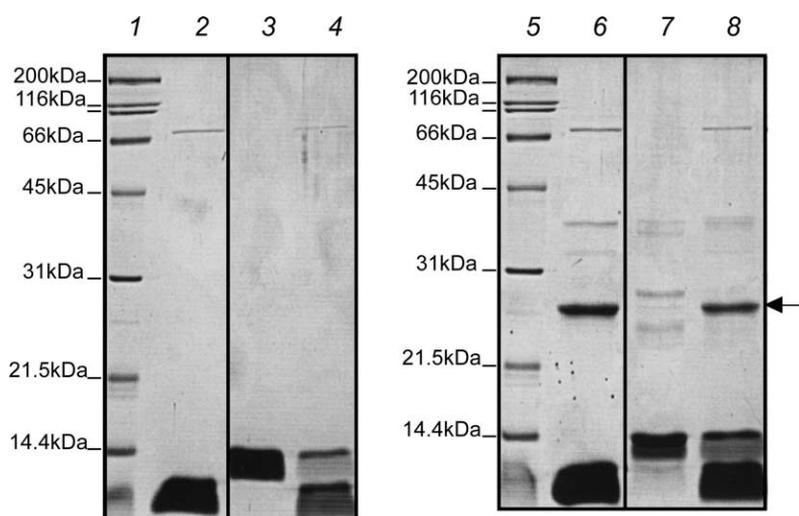
Figure 4. The S1 domain from RNase E, but not PNPase, can form a glutaraldehyde-crosslinked homodimer. Samples (45 μM) of RneS1$^{35-125}$ and the PNPase S1 domain (residues 617–700) were treated with 0.01% glutaraldehyde, separated on an SDS-15% PAGE gel, and stained with Coomassie brilliant blue. Lanes 1 and 5, standard molecular mass markers; lanes 2–4, the S1 domains without glutaraldehyde; and lanes 6–8, the same proteins treated with glutaraldehyde. Lanes 2 and 6, RneS1$^{35-125}$; lanes 3 and 7, the PNPase S1 domain; and lanes 4 and 8, a 1:1 mixture of these two S1 domains. The arrow (←) in the right margin denotes the position of the putative RneS1$^{35-125}$ homodimer.

WT RneS1$^{35-125}$ showed a cooperative unfolding transition with a $t_m \sim 51\,°C$ and a ΔH of approximately 100 kJ/mol, the G66S mutant exhibited a very broad transition with no clear midpoint unfolding temperature (Supplementary Material). Furthermore, in contrast to the $^1H-^{15}N$ HSQC spectrum of the WT protein, with well-dispersed peaks indicative of a folded species, the spectra of the G66S mutant recorded between 10 °C and 35 °C show a predominance of amides with ^1H chemical shifts characteristic of a predominantly unstructured random coil polypeptide (Supplementary Material). Together these data reveal that the introduction of a serine residue at position 66 of RneS1$^{35-125}$ severely destabilizes the folded conformation of this protein fragment, and strongly suggest that the temperature-sensitive phenotype of the G66S mutation in RNase E results from the disruption of the structure and hence function of its S1 domain.

Identification of the oligonucleotide-binding interface

To test the hypothesis that the RNase E S1 domain is involved in oligonucleotide binding, a filter-binding assay was used to screen a variety of RNAs and DNAs for association with RneS1$^{35-125}$ (R.E.E. & G.A.M, unpublished results). The studies revealed that, at neutral pH and in low ionic strength buffer, RneS1$^{35-125}$ binds RNA polynucleotides (80–180 nt) including *rps*T(268–447)-polyA$_{30}$,[25] SL9A RNA,[26] poly(A), poly(U), and poly(C). The apparent K_D values for all five RNA ligands ranged from 2 μM to 4 μM. DNA oligonucleotides also bind to RneS1$^{35-125}$ in this assay, but with affinities at least an order of magnitude weaker.

To identify the oligonucleotide-binding interface of RneS1$^{35-125}$, $^1H-^{15}N$ HSQC NMR spectroscopy was used to monitor the titration of the protein with a single-stranded DNA (ssDNA) decamer, 5′-d(ACAGTATTTG)-3′. This corresponds to the RNase E cleavage site of RNAI.[27] ssDNA, rather than ssRNA, was utilized to avoid degradation by trace contaminating ribonucleases, including PNPase, whose presence was confirmed by Western blotting (data not shown). As shown in Figure 5, several amides exhibit progressive chemical shift perturbations upon addition of the ssDNA. When mapped on the structure of RneS1$^{35-125}$, these amides cluster to a surface region of positive electrostatic potential that is rich in exposed basic (K106, R109, K112) and aromatic (F57, F67) residues (Figure 6). Since amide chemical shifts are exquisitely sensitive to structural changes, this perturbation mapping provides a qualitative identification of the oligonucleotide-binding surface as being formed primarily by residues in β-strands β2 and β3 and the mobile loops L$_{23}$ and L$_{45}$. In parallel, quantitative analyses of titration data for five residues with the largest shift changes (R64, L68, R109, A115, L116) yields a K_D of 160($\pm$40) μM for the binding of this ssDNA by RneS1$^{35-125}$.

Discussion

Structural features of RneS1$^{35-125}$

S1 domains have been found in or are predicted to occur in four of the major enzymes of RNA processing and/or degradation in *E. coli*: RNase E, PNPase, RNase II, and RNase G.[8] Here, we report experimentally determined crystalline and solution state structures for the isolated S1 domain of RNase E. Together, these provide the first high-resolution conformational and dynamic information available for this important enzyme, and verify the major features of this domain that were predicted by threading.[13] In particular, a high-quality
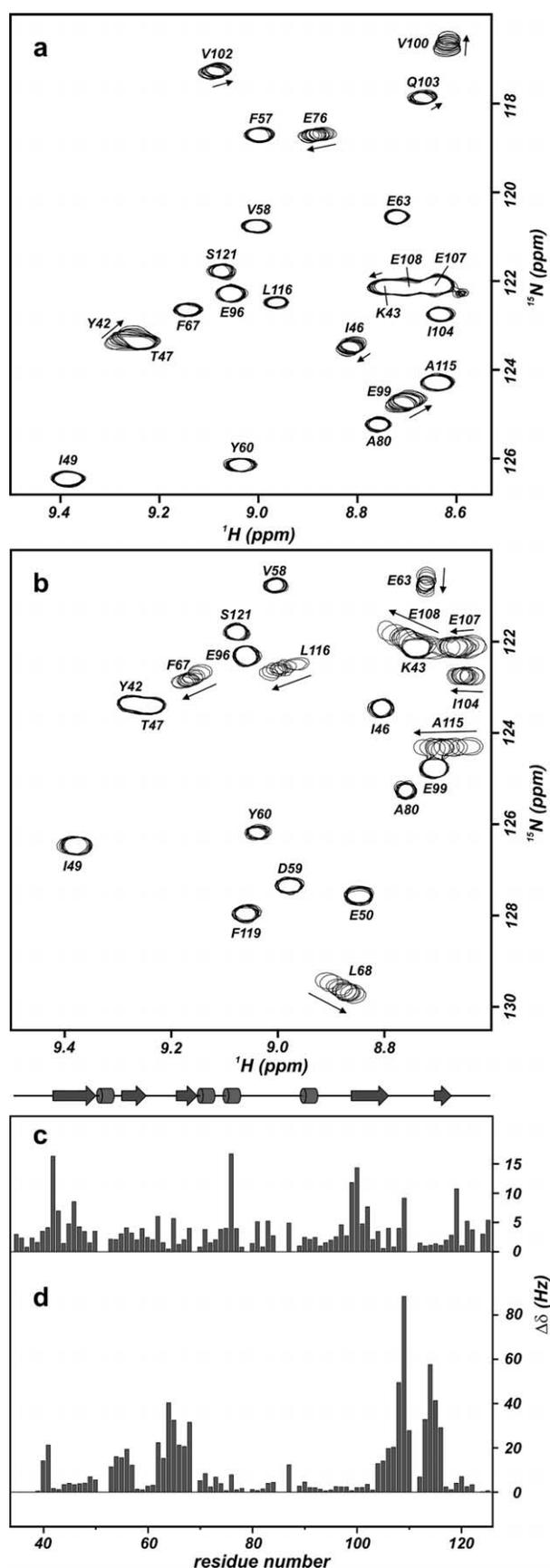
Figure 5. Identifying the dimerization and oligo-nucleotide-binding interfaces of RneS1$^{35-125}$ from amide chemical shift perturbations. a, RneS1$^{35-125}$ exists in an equilibrium between monomer and dimer in solution.

structure of the RneS1$^{35-125}$ dimer was solved crystallographically, whereas a structural ensemble of the monomeric species in solution was calculated using spectroscopic data. Both techniques revealed that RneS1$^{35-125}$ adopts an OB-fold with conformationally mobile loops linking successive β-strands of a well-ordered β-barrel core. Furthermore, NMR spectroscopy was used to define a previously unrecognized equilibrium between monomeric and dimeric forms of RneS1,$^{35-125}$ to identify the oligonucleotide-binding interface of this S1 domain, and to demonstrate that the G66S substitution found in the temperature-sensitive *ams-1* mutant disrupts the folded state of RneS1$^{35-125}$.

Tertiary structure comparisons using DALI[28] reveal a very close relationship between RneS1$^{35-125}$ and a variety of OB-fold proteins including the *E. coli* major cold shock protein CspA (1MJC; rmsd 2.2 Å; 21% sequence identity over A39-G61, A62-E76, V93-E107 and K112-F119), the S1 domain of PNPase (1SRO; rmsd 2.4 Å; 30% sequence identity over A39-P51, L53-A56, V58-E76, P88-E107 and R109-F119), and an archaeal homolog of the eukaryotic RNA polymerase II RPB4/RPB7 complex (1GO3, rmsd 2.8 Å; 20% sequence identity over N40-P51, E54-G61, E63-E76, F78-N81 and D92-L122). This similarity places the S1 domain of RNase E within the "cold-shock DNA-binding" family of OB-folds, as annotated in the structural classification of proteins (SCOP) database.[10,29] All of these proteins adopt a closed β-barrel structure, yet lack the pronounced α-helix cap between β-strands β3 and β4 found in the prototypical OB-fold.[9] Although the position and length of strand β5 and loop L_{34} varies somewhat between these proteins, many of them share a type II β-turn preceding β4 and have a short α or 3_{10} helix following β3. In Figure 7a, the sequence of the RNase E S1 domain is aligned against selected members of this cold-shock DNA-binding family in order to highlight

Portions of 12 ^{1}H–^{15}N HSQC spectra of uniformly ^{15}N-labeled protein at concentrations of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1 and 1.2 mM are overlaid. The arrows indicate the directions in which the ^{1}H–^{15}N peaks shift with increasing protein concentration. b, NMR titration of RneS1$^{35-125}$ with the ssDNA 10-mer 5′-d(ACAGTATTTG)-3′. Portions of ten ^{1}H–^{15}N HSQC spectra of uniformly ^{15}N-labeled protein at an initial concentration of 0.3 mM in the presence of 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.1, 1.5, 2.0 and 3.0 equivalents of DNA are overlaid. The arrows indicate the directions in which the ^{1}H–^{15}N peaks shift with added DNA. c, A histogram of the amide ^{1}H–^{15}N chemical shift changes ($\Delta\delta = (\Delta\delta_N^2 + \Delta\delta_{HN}^2)^{0.5}$) of RneS1$^{35-125}$ accompanying the concentration of the protein from 0.1 mM to 1.2 mM is shown, along with a cartoon representation of the secondary structure of this protein. d, A histogram of the amide ^{1}H–^{15}N chemical shift changes RneS1$^{35-125}$ resulting from the addition of 3 eq. of ssDNA. These data are mapped on the tertiary structure of RneS1$^{35-125}$ in Figure 6.
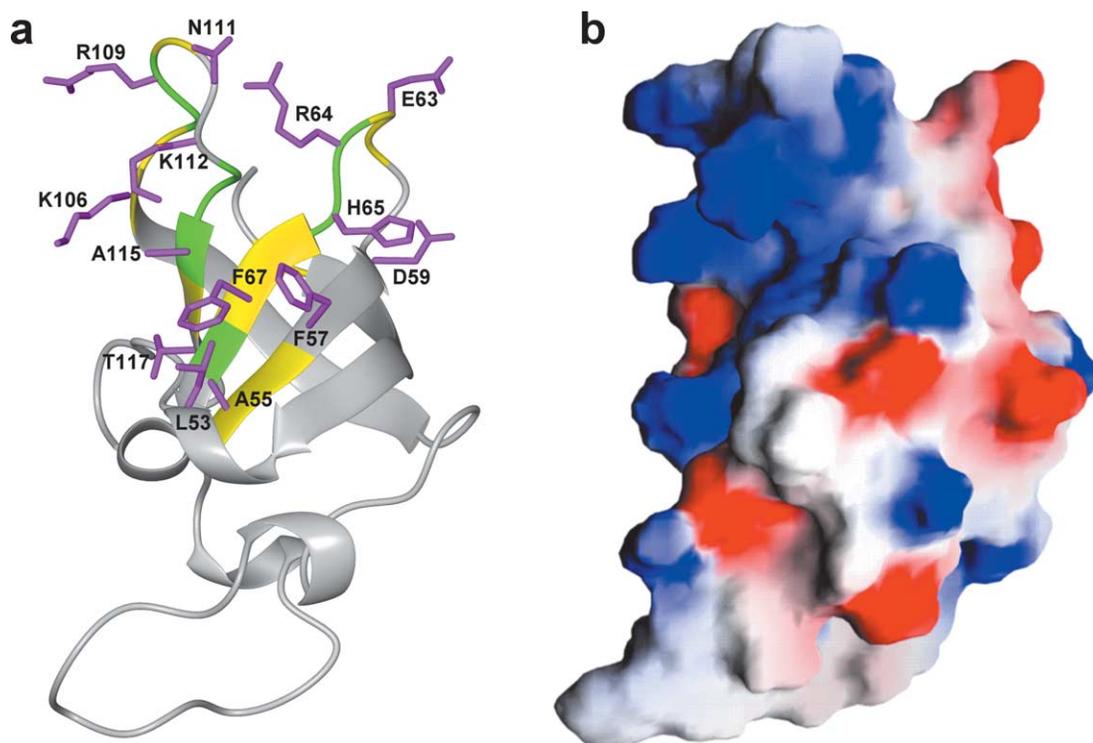
Figure 6. The oligonucleotide-binding site of RneS1[35−125] was identified by NMR chemical shift perturbation mapping. a, The positions of amides experiencing significant ^{1}H−^{15}N chemical shift changes upon titration with 5′-d(ACAGTATTTG)-3′ are mapped on the crystal structure of RneS1[35−125] (green, $\Delta\delta > 30$ Hz; yellow, 15 Hz $< \Delta\delta <$ 30 Hz; see Figure 5d). The exposed side-chains along this binding site are shown in magenta. b, A surface representation of RneS1[35−125], in the same orientation as in a, is shown with regions of positive and negative electrostatic potential color coded in blue and red, respectively.

conserved residues important for the OB-fold, as well as those involved in the dimerization and nucleic acid-binding interfaces of RneS1[35−125]. This information is also summarized in a cartoon representation of the RneS1[35−125] structure, shown in Figure 7b.

In addition to tertiary structure, the similarity of RNase E S1 domain to other cold-shock DNA-binding OB-fold family members extends to dynamics. For example, as with RneS1[35−125] (Figure 2), high rms deviations are observed for the loop regions within the ensemble of NMR-derived structures of the *E. coli* major cold shock protein CspA[24] and the human Y-box protein YB-1.[30] The electron densities of residues within the L_{34} loop of the former protein are also ill-defined by X-ray crystallography.[31] Furthermore, despite their low level of sequence identity, corresponding residues in RneS1[35−125] and CspA display strikingly similar patterns of amide ^{15}N relaxation rates (Figure 2; and see Feng *et al.*[24]). In particular, the backbone forming the long loop L_{34} and the oligonucleotide-binding loop L_{45} of both proteins exhibit considerable conformational flexibility on a sub-nanosecond timescale. These features may provide a degree of plasticity necessary for functions such as the binding of single-stranded nucleic acids, common to the cold-shock DNA-binding family members.

Dimerization of RneS1[35−125]

RneS1[35−125] dimerizes within the crystallographic asymmetric unit through intermolecular hydrogen bonding and van der Waals packing of hydrophobic surface side-chains contributed primarily from strands β1 and β4 and its C-terminal sequence. Our structural analysis thus expands the diverse repertoire of recognized mechanisms by which the surface of the OB-fold has evolved to mediate macromolecular association.[10,11] Several lines of evidence demonstrate that the protein undergoes a monomer–dimer equilibrium in solution. Interestingly, similar to RneS1,[35−125] NMR spectroscopic studies of the major cold shock protein CspB from *Bacillus subtilis* yielded a monomeric structure,[32] whereas X-ray crystallographic studies demonstrated a dimer formed *via* intermolecular hydrogen bonding between residues along strand β4 of its OB-fold.[33] Subsequent biophysical studies confirmed that CspB exhibits a phosphate-dependent monomer–dimer equilibrium in solution.[34]

Isolated RneS1[35−125] dimerizes in solution with a K_D in the millimolar range. This is consistent with the modest amount of hydrophobic surface area buried within the dimer interface, combined with a limited number of intermolecular hydrogen bonding interactions (Figure 3). We recognize that

such a K_D value is too weak to be of physiological significance on its own. Nevertheless, based on three observations, we hypothesize that S1 domain dimerization is important in the context of native RNase E. First, this endonuclease is active as an oligomer. Thus, the S1 domains may form stable dimers in synergy with additional self-associating domains of the protein. Second, the side-chains involved in the dimerization interface of RneS1[35−125] are conserved among RNase E homologs from a variety of organisms, but not among other cold-shock DNA binding OB-fold family members such as PNPase (Figure 7). In particular, in RNase G, a closely related paralog of RNase E,[35] highly conserved residues include those important for the OB-fold as well as those involved in the dimerization interface of RneS1[35−125]. These data suggest that the dimerization surface adopted by RneS1[35−125] is a distinct feature of the S1 domain of RNase E and its functional homologs. Third, Diwa *et al.*[13] observed that the substitutions K37A and Y60A within the RNase E S1 domain disrupted the feedback inhibition of *rne* expression, but not the catalytic activity of the native enzyme. These two residues lie along the dimerization interface of RneS1[35−125], near its C_2 symmetry axis, such that K37 of one monomer would be proximal to Y60 of the other (Supplementary Material). We thus speculate that these residues form a surface of the S1 domain dimer important for feedback inhibition of RNase E synthesis.

Nucleic acid binding

Using both filter-binding assays and NMR spectroscopy, we demonstrated that RneS1[35−125] associates with a variety of oligomeric and polymeric single-stranded RNA and DNA ligands. In general agreement with predictions made by Bycroft *et al.*[8] and Diwa *et al.*,[13] these results confirm that the S1 domain of RNase E indeed functions as a nucleic acid-binding module with affinities similar to those of other cold-shock OB-fold family members.[36] The binding surface of RneS1[35−125] was identified as a groove of positive electrostatic potential extending across strands β2 and β3 (Figures 5b and d, and 6a and b). Part of the binding surface is also formed by the flexible loops L23 and L45, possibly allowing an induced-fit oligonucleotide-binding mechanism with a broad specificity.[11] Importantly, this surface is located opposite the dimerization interface of RneS1[35−125], indicating that two independent oligonucleotide-binding sites can exist in the dimeric form of this protein (Figure 3c and Supplementary Material).

As seen in Figures 6 and 7, RneS1[35−125] maintains the conserved nucleic acid-binding features documented for the OB-fold,[10,11] both in the location of its oligonucleotide-binding interface and by the presence of exposed aromatic (F57, H65, F67), hydrophobic (L53, A55, A115), and positively charged side-chains (R64, K106, R109, K112) along this interface. Mutations made to key residues
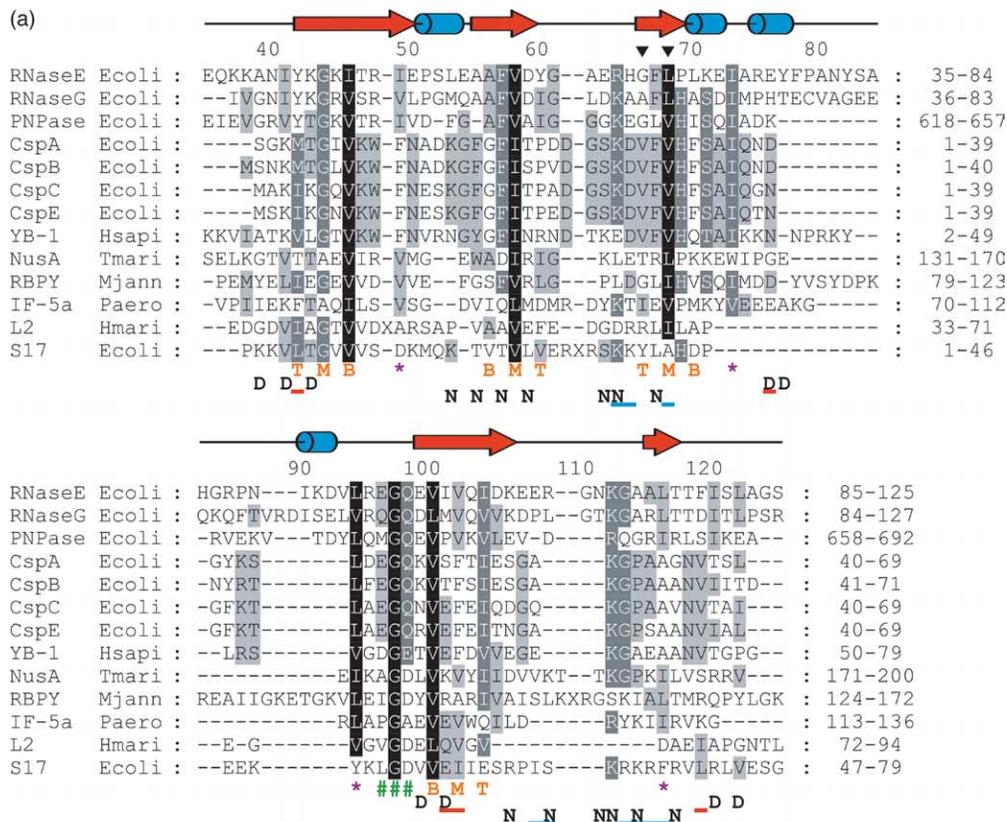
within this site show a significant effect on RNase E feedback regulation *in vivo* and/or on ribonuclease activity *in vitro*.[13] In particular, alanine substitutions of F57, F67, or K112 located on strands β2, β3, and in the flexible loop L45, respectively, strongly impair both functions of this enzyme. Likewise, alanine substitutions of R64 or R109 in the flexible loops L23 and L45 partially impair feedback regulation.

Temperature-sensitive mutations of RNase E

Early studies of RNase E identified two conditional mutations, *rne-3071* (L68F) and *ams-1* (G66S), each leading to a lethal phenotype at elevated temperatures. Both of these residues lie within strand β3 of the S1 domain (Figure 7a and b), and thus the amino acid substitutions, which would introduce bulkier groups into the hydrophobic core of this domain, could potentially disrupt RNA binding by perturbing the precise conformation of its oligonucleotide-binding interface. However, our analyses of WT and mutant RneS1[35−125] indicate that the G66S, and by inference, the L68F mutations act more globally by destabilizing the folded structure of the RNase E S1 domain. Positions 66 and 68 are generally occupied by hydrophobic residues that contribute to the interior of the OB-fold (Figure 7(a)). However, in the case of RneS1[35−125], G66 packs against the side-chains of Y60 (in loop L23), A114 (loop L45), and L116 (strand β5). Replacing an H^α of this glycine with a polar serine side-chain leads to the observed unfolding of isolated G66S-RneS1[35−125] (Supplementary Material), most likely due to unfavorable steric clashes with these three residues, as well as the introduction of a hydroxyl group with an unsatisfied hydrogen bond. Within the context of the full-length RNase E, the S1 domain may be stabilized by additional interactions, allowing the mutant to be sufficiently functional at permissive temperatures. Similarly, L68 is involved in an extensive set of van der Waals interactions with V58 (strand β2), I73 (loop L34), V100 (strand β4), and T118 (after strand β5). The substitution of a planar phenylalanine aromatic ring at this position would likely also destabilize the folded S1 domain by disrupting its core packing.

Implications for the quaternary structure and function of RNase E

Here, we both confirmed previous predictions that Rne[35−125] encodes an independently folded nucleic acid binding domain and discovered that it dimerizes with a K_D in the millimolar range. These two findings may be key to understanding the role of the S1 domain in RNase E. Recent data have shown that the N-terminal 529 residues of RNase E can form a tetramer;[14] in addition, RNase G is at least a dimer.[37] Moreover, multimerization is required for the activity of both enzymes.[14,37] Accordingly, we propose that in native RNase E:
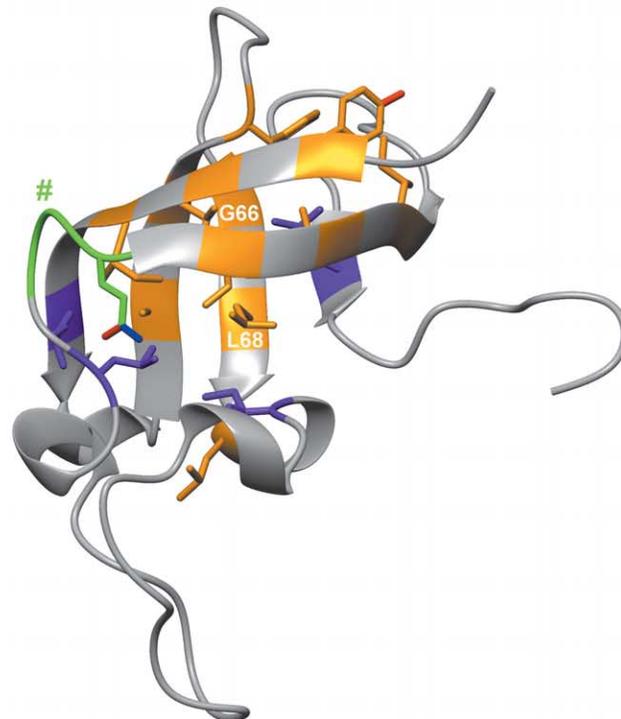
(a)

```
                        40          50          60              70          80
RNaseE Ecoli : EQKKANIYKGKITR-IEPSLEAAFVDYG--AERHGFLPLKELAREYFPANYSA :  35-84
RNaseG Ecoli : --IVGNIYKGRVSR-VLPGMQAAFVDIG--LDKAAFLHASDIMPHTECVAGEE :  36-83
PNPase Ecoli : EIEVGRVYTGKVTR-IVD-FG-AFVAIG--GGKEGLVHISQLADK-------- : 618-657
CspA   Ecoli : ----SGKMTGIVKW-FNADKGFGFITPDD-GSKDVFVHFSAIQND-------- :   1-39
CspB   Ecoli : ---MSNKMTGLVKW-FNADKGFGFISPVD-GSKDVFVHFSAIQND-------- :   1-40
CspC   Ecoli : ----MAKIKGQVKW-FNESKGFGFITPAD-GSKDVFVHFSAIQGN-------- :   1-39
CspE   Ecoli : ----MSKIKGNVKW-FNESKGFGFITPED-GSKDVFVHFSAIQTN-------- :   1-39
YB-1   Hsapi : KKVIATKVLGTVKW-FNVRNGYGFINRND-TKEDVFVHQTAIKKN-NPRKY-- :   2-49
NusA   Tmari : SELKGTVTTAEVIR-VMG--EWADIRIG---KLETRLPKKEWIPGE------- : 131-170
RBPY   Mjann : -PEMYELIEGEVVD-VVE--FGSFVRLG---PLDGLIHVSQLMDD-YVSYDPK :  79-123
IF-5a  Paero : -VPIIEKFTAQILS-VSG--DVIQLMDMR-DYKTIEVPMKYVEEEAKG----- :  70-112
L2     Hmari : --EDGDVIAGTVVDXARSAP-VAAVEFE--DGDRRLILAP------------- :  33-71
S17    Ecoli : ---PKKVLTGVVVS-DKMQK-TVTVLVERXRSKKYLAHDP------------- :   1-46
```

        T M B   *   B M T   T M B   *
     D D D   N N N   NN N   D D

```
                  90         100         110         120
RNaseE Ecoli : HGRPN---IKDVLREGQEVIVQIDKEER--GNKGAALTTFISLAGS :  85-125
RNaseG Ecoli : QKQFTVRDISELVRQGQDLMVQVVKDPL--GTKGARLTTDITLPSR :  84-127
PNPase Ecoli : -RVEKV---TDYLQMQQEVPVKVLEV-D----RQGRIRLSIKEA-- : 658-692
CspA   Ecoli : -GYKS-------LDEGQKVSFTIESGA-----KGPAAGNVTSL--- :  40-69
CspB   Ecoli : -NYRT-------LFBGQKVTFSIESGA-----KGPAAANVIITD-- :  41-71
CspC   Ecoli : -GFKT-------LAEGQNVEFEIQDGQ-----KGPAAVNVTAI--- :  40-69
CspE   Ecoli : -GFKT-------LAEGQRVEFEITNGA-----KGPSAANVIAL--- :  40-69
YB-1   Hsapi : --LRS-------VGDGETVEFDVVEGE-----KGAEAANVTGPG-- :  50-79
NusA   Tmari : ----------EIKAGDLVKVYIIDVVKT--TKGPKILVSRRV--- : 171-200
RBPY   Mjann : REAIIGKETGKVLEIGDYVRARIVAISLKXRGSKIALTMRQPYLGK : 124-172
IF-5a  Paero : -----------RLAPGAEVEVWQILD------RYKIIRVKG----- : 113-136
L2     Hmari : --E-G-------VGVGDELQVGV-----------DAEIAPGNTL : 72-94
S17    Ecoli : --EEK-------YKLGDVVEIIESRPIS----KRKRFRVLRLVESG :  47-79
```

         * ### B M T   *
          D D   N N   NN N N   D D

(b)

Figure 7. Sequence analyses of residues crucial for the OB-fold and the dimerization and oligonucleotide-binding interfaces of RneS1.[35–125] (a) Alignments based on sequence (CLUSTAL W)[65] and three-dimensional structure comparisons (DALI)[28] of a subset of cold-shock DNA-binding family OB-fold proteins, including: RNase E, RNase G and PNPase of *E. coli*; major cold shock proteins CspA, CspB, CspC, and CspE of *E. coli*; cold shock domain of the Y-Box protein YB-1 of *Homo sapiens*; NusA of *Thermotoga maritima*; archaeal RNA polymerase II RPB4/RPB7 complex of *Methanococcus jannaschii*; initiation translation factor IF-5a of *Pyrobaculum aerophilum*; ribosomal protein L2 of *Haloarcula marismortui*; and the ribosomal protein S17 of *E. coli*. Non-conserved insertions are indicated by an X. Numbers at the top correspond to amino acid residues in *E. coli* RNase E, and those to the right indicate the actual range for each sequence. Residues highlighted in black, dark gray and light gray are conserved in over 90%, in 60–90%, and in 40–60% of the proteins, respectively. The secondary structure of RNase E is indicated on the top. G66 and L68 are marked

(i) dimerization of two S1 domains occurs co-operatively along with the self-association of other domains in the protein to produce an initial stable dimer and subsequent higher-order oligomer; and (ii) this structural arrangement positions the S1 domains, which are required for substrate binding, adjacent to the catalytic domains. Two cartoon models of possible quaternary structures of RNase E: are presented in Figure 8. Consistent with the crystallographic data reported by Callaghan *et al.*,[14] each of these models exhibits D_2 point symmetry. However, in contrast to the hypothesis by these authors that the monomeric S1 domains are maximally separated from one another on the periphery of the tetramer, we propose that these domains self-associate to form at least pairs of dimers. Furthermore, based on the crystal structure of Rne^{35-125}, the S1 domain of one monomer may lie in proximity to the catalytic domain of its partner monomer. In addition to contributing to the tetrameric structure of RNase E, the close positioning of these S1 domains to one another and to two catalytic domains could increase the overall affinity and specificity of the enzyme for its substrates. The models in Figure 8 also provide a novel explanation for the temperature-sensitive phenotypes of the *rne-3071* (L68F) and *ams-1* (G66S) mutations. Namely, disruption of the structure of the S1 domain leads to the disruption of the quaternary structure of the entire RNase E enzyme. Finally, we note that Callaghan *et al.*[14] cite their observation that Y82 of RNase E is sensitive to cleavage by chymotrypsin as support for the placement of the S1 domains at the exposed edges of the tetramer. However, our structural analysis of Rne^{35-125} shows that this residue lies in the large, flexible loop L_{34}, and thus could be accessible to proteolysis even within a dimeric state.

The combined X-ray crystallographic and NMR spectroscopic analyses of Rne^{35-125} provide a much needed structural and dynamic framework for understanding the precise functions of the S1 domain in RNase E. More importantly, these data prompt readily testable hypotheses regarding the roles of dimerization and oligonucleotide binding by this OB-fold domain in establishing the quaternary structure and substrate specificity of RNase E, a model enzyme for RNA metabolism.
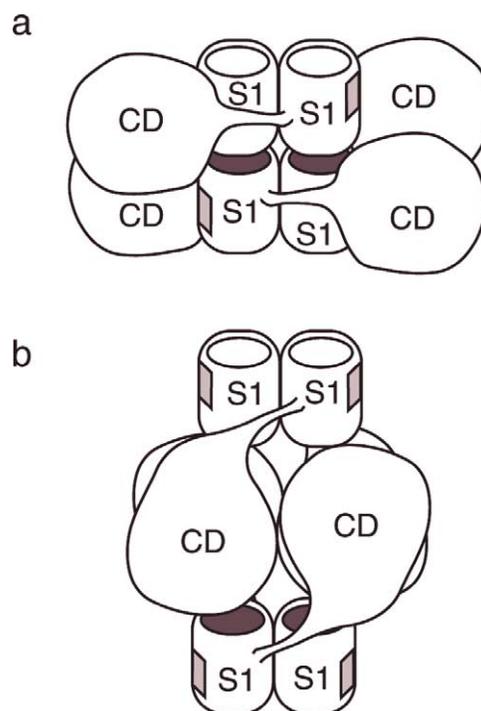

Figure 8. Models of the quaternary structure of an RNase tetramer incorporating the dimeric association of S1 domains. Each RNase E monomer is drawn with an N-terminal S1 domain, represented by a barrel, and the remaining fraction of the protein, including the catalytic domain, represented by an oval. In model b, this latter fraction forms the core of the tetramer, whereas in model a, this role is played by a hypothetic dimer of S1 domain dimers. The S1 domain RNA-binding sites are indicated in light gray along the sides of the barrels.

Materials and Methods

Expression and purification

A DNA fragment encoding residues 35–125 of *E. coli* RNase E (RneS1$^{35-125}$) was ligated into the bacterial expression vector pET15b (Novagen) as a fusion protein with an N-terminal His$_6$ sequence and thrombin cleavage site. The protein encoded by the ensuing plasmid had the sequence MGSS-(H)$_6$-SSGLVPRG SHML-(^{35}EQKKANIYKGKITRIEPSLEAAFVDYGAERH GFLPLKEIAREYFPANYSAHGRPNIKDVLREGGQEVIVQ

by black triangles. Hydrophobic residues crucial for the OB-fold as described by Theobald *et al.*[11] are indicated at the bottom by orange letters T, M or B, which correspond to the top, middle and bottom part of the hydrophobic core, respectively. Additional residues that extend the hydrophobic core in RneS1$^{35-125}$ are highlighted (purple asterisk). The characteristic type II β-turn preceding strand β4 is indicated (green #). Residues which are part of the dimerization interface in the crystallographic dimer of RneS1$^{35-125}$ are indicated by D, and exposed residues on the RNA/DNA-binding interface of RneS1$^{35-125}$ by N. The red and blue lines indicate residues in RneS1$^{35-125}$ whose ^{1}H–^{15}N chemical shifts change together more than 10 Hz or 30 Hz upon dimerization or DNA binding, respectively. Note the variable length of loop L_{34}. (b) Ribbon diagram of the crystal structure (monomer A) of RneS1,$^{35-125}$ with hydrophobic side-chains crucial for the OB-fold as described by Theobald *et al.*,[11] shown in orange. Additional residues that extend the hydrophobic core in RneS1$^{35-125}$ are highlighted in purple. These residues are conserved in a subset of OB-fold proteins including major cold shock proteins and some S1 domains. A characteristic feature of this subset of S1 domains is a type II β-turn before strand β4 (indicated in green) that is formed by the sequence EGQ. The conserved Gly98 within this turn is marked by #. The positions of the temperature-sensitive mutations L68F and G66S are also marked.

IDKEERGNKGAALTTFISLAGS[125]). The G66S mutation was subsequently introduced into this plasmid using PCR-based site-directed mutagenesis. The BL21(λDE3) *E. coli* host strain containing either recombinant plasmid was grown at 37 °C in LB medium for unlabelled protein and M9 minimal medium supplemented with the following: 1 g/l (^{15}N, 99%)–NH_4Cl for uniform ^{15}N labeling; 1 g/l (^{15}N, 99%)–NH_4Cl and 3 g/l ($^{13}C_6$, 99%)–glucose for uniform ^{13}C–^{15}N labeling; and 0.3 g/l ($^{13}C_6$, 99%)–glucose and 2.7 g/l ($^{12}C_6$)–glucose for non-random 10% ^{13}C-labeling. At an A_{600} of 0.75, protein expression was induced by the addition of 1 mM isopropyl-β-thiogalactopyranoside, followed by growth for five hours at 37 °C.

After harvesting by centrifugation, the cell pellet was suspended in a buffer of 50 mM Hepes (pH 7.5), 150 mM NaCl, 10 mM imidazole, and 5% (v/v) glycerol, and lyzed by freeze-thawing (-80 °C), passage through a French press, and sonication. Lysates were cleared by centrifugation at 15,000*g* for 60 minutes, followed by filtration through a 0.8 μm cut-off membrane. RneS1$^{35-125}$ was purified from the lysate by Ni^{2+}-affinity chromatography (Amersham Biosciences) using a buffer of 50 mM Hepes (pH 7.5), 150 mM NaCl, and 5% glycerol with 10, 30 and 100 mM imidazole for loading, washing, and elution, respectively. After dialysis into a loading buffer of 50 mM Hepes (pH 7.5), 150 mM NaCl, and 5% glycerol without imidazole, the His_6-tag was removed by incubation with thrombin. The extent of proteolysis was monitored by a reduction in apparent molecular mass by using SDS-PAGE, and upon completion (two days), was terminated with a 30 minute incubation within 300 μl of *p*-aminobenzamidine beads (Sigma). One milliliter of water-rinsed Talon beads (Clontech) was added to the samples to remove the cleaved His_6-tag and any uncleaved full-length protein.

Electrospray ionization mass spectroscopy yielded a mass of 10,687 Da for the unlabeled WT protein, which is consistent with the expected value of 10,687 Da for the RneS1$^{35-125}$ construct with five additional residues (Gly-Ser-His-Met-Leu) remaining at its N terminus after proteolytic cleavage. Protein concentrations were determined by absorbance spectroscopy using a predicted $ε_{280} = 5120$ M^{-1} cm^{-1} (ProtPara†).

Crystallization and diffraction data

Large crystals (up to 400 μm diameter) of RneS1$^{35-125}$ initially grew in the NMR buffer (1.65 mM protein concentration, 20 mM phosphate (pH 6.5), 50 mM NaCl, and 0.05% (w/v) NaN_3) during four weeks of storage at 4 °C. However, difficulties with cryo-protecting these large crystals led us to pursue different crystallization conditions. Subsequently, isomorphous crystals were obtained by the hanging-drops vapor-diffusion method at 18 °C using a 1.3 mM protein solution (in 20 mM Hepes (pH 6.5), 50 mM NaCl) and a well solution consisting of 0.17 M sodium acetate (pH 6.5), 85 mM sodium cacodylate, 50% (w/v) PEG8000, 15% glycerol. Crystals had a typical size of 60 μm × 60 μm × 130 μm. No additional cryo-protection was necessary. A Pb-derivative was obtained by soaking a crystal for 30 minutes in a solution of 20 mM trimethyl lead(IV) acetate, 0.14 M sodium acetate, 68 mM sodium cacodylate, 40% PEG8000, 12% glycerol.

Diffraction data for the native crystals were collected at 100 K on an ADSC Quantum 210 CCD detector at

† http://www.expasy.org/tools/protparam.html

beamline 8.2.1 of the Advanced Light Source, Lawrence Berkeley National Laboratory. A single anomalous dispersion (SAD) dataset of the Pb-derivative was collected at 100 K on an ADSC Quantum 4 CCD detector at beamline X8-C of the National Synchrotron Light Source, Brookhaven National Laboratory. All diffraction patterns were indexed and scaled with the HKL program package,[38] which revealed a tetragonal $P4_12_12$ space group in both cases. The asymmetric unit contains two protein molecules in accordance with a calculated Matthew's coefficient of 2.6 $Å^3$/Da and assuming 51.6% solvent.

Crystal structure analysis and refinement

Using SAD phasing methods, the Pb sites were located and refined with SOLVE.[39] Phase extension including solvent flattening and non-crystallographic symmetry averaging with RESOLVE[40] produced a clearly interpretable electron density map, from which an initial model was built using XtalView[41] and improved by iterative rounds of manual fitting in Xtalview and refinement in Refmac5.[42] Non-crystallographic symmetry averaging did not improve the model. Initial phases for the native model were obtained from the Pb-derivative. The refined structures of the native and derivative crystals yielded comparable *R* and R_{free} values, and superimposed with an rms deviation of 0.187 Å on the 672 main-chain and 0.573 Å on all 1327 observed atoms. However, the data from the Pb-derivative refined to a significantly more ordered structure with lower overall and individual *B*-factors, and was therefore chosen for subsequent analyses. Statistics on the X-ray diffraction data and resulting structures are summarized in Table 1.

NMR data collection and analysis

Prior to data collection, RneS1$^{35-125}$ was dialyzed at 4 °C in 20 mM phosphate buffer (pH 6.5), containing 50 mM NaCl and 0.05% NaN_3, and concentrated using a 1K filter (Pall Filtron). The final samples had a protein concentration of 1.4–1.65 mM, and contained 8% or 99% (v/v) ^2H_2O for signal lock. Except where noted, all NMR spectra were recorded with the ^{13}C/^{15}N-labeled RneS1$^{35-125}$ at 30 °C on a 600 MHz Varian INOVA spectrometer equipped with an inverse triple resonance probe and pulsed-field gradient accessory. NMR data were processed using NMRPipe[43] and analyzed using the program SPARKY.[44] ^{1}H and ^{13}C chemical shifts were referenced to an external sample of sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) and ^{15}N referenced indirectly *via* magnetogyric constant ratios.[45]

Assignments of the resonances from the ^{1}H, ^{13}C, and ^{15}N nuclei of RneS1$^{35-125}$ were obtained using a conventional set of sensitivity-enhanced heteronuclear NMR spectra, including the ^{13}C- and ^{15}N-HSQC, HNCACB, CBCA(CO)NH, HNCO, H(CCO)NH, C(CO)NH, HCCH-TOCSY and HACAN experiments, as described.[46,47] In addition, a suite of MUSIC experiments provided amino acid-type selective amide resonance assignments.[16,17] The assignments of resonances from aromatic side-chains were obtained using ^{1}H–^{13}C $C^βH^δ$ and $C^βH^ε$ experiments,[48,49] combined with a non-sensitivity-enhanced ^{1}H–^{15}N HSQC experiment with delays ($\frac{1}{4}J \sim 11$ ms) optimized for long-range couplings within histidine imidazole rings.[50] Stereospecific assignments of the diasterotopic methyl groups of the valine and leucine residues were determined using a constant time ^{1}H–^{13}C HSQC spectrum acquired on a non-randomly

10% ^{13}C-labeled sample, combined with long-range C′–C$^\gamma$ and N–C$^\gamma$ spin echo spectra, as described.[46] Stereospecific assignments of non-degenerate β-methylene protons were based upon HNHB and short mixing time ^{15}N-TOCSY-HSQC (τ_m = 32 ms) experiments recorded on a 500 MHz Varian UNITY spectrometer with a ^{15}N-labeled sample, as described.[46] Stereospecific assignments of side-chain Gln and Asn amide protons were obtained from an EZ-HMQC-NH$_2$ spectrum.[51]

NMR-based structure calculations

Structural calculations were completed using ARIA/CNS version 1.2.[18] Distance restraints were acquired from three-dimensional ^{15}N-HSQC-NOESY, aliphatic ^{13}C-HSQC-NOESY, simultaneous ^{13}C- and ^{15}N-HSQC-NOESY, aromatic ^{13}C-HSQC- and constant time methyl–methyl NOESY (simultaneous ^{13}C/^{13}C/^{1}H and ^{13}C/^{15}N/^{1}H)[52] spectra, all with 100 ms mixing times. NOE lists from each of these individual NOESY spectra, consisting in total of 1044 manually assigned and of 3407 unassigned peaks, were merged by ARIA to yield 2113 unambiguously and 597 ambiguously assigned distances restraints for the calculation of the final structural ensemble. Backbone dihedral angles were determined using TALOS.[19] The χ_1 side-chain dihedral angles for residues with β-methylene protons were determined from the analysis of the short mixing time ^{15}N TOCSY-HSQC (τ_m = 32 ms) and HNHB spectra, χ_1 dihedral angles for Ile, Val, and Thr residues on the bases of $^3J_{NC'}$ and $^3J_{C'C'}$ coupling constants, as described.[46] All Xaa-Pro amides were constrained to the *trans* conformation based upon the proline chemical shifts using the program POP.[53] Hydrogen bond restraints were included for amides located in elements of regular secondary structure. Using these restraints, a total of 200 structures were calculated, of which the ten energetically best were further refined in a water box using Lennard–Jones potentials. Statistics on the NMR spectroscopic data and resulting structures are summarized in Table 2.

NMR relaxation analysis

Amide ^{15}N T_1, T_2 and steady-state heteronuclear ^{1}H{^{15}N}NOE relaxation parameters for 0.6 mM and 1.2 mM ^{15}N-labeled RneS1$^{35-125}$ were acquired on a 500 MHz NMR spectrometer at 30 °C, as described.[54] In addition, steady-state heteronuclear ^{1}H{^{15}N}NOE relaxation parameters were measured with a 1.5 mM ^{13}C/^{15}N sample of this protein on a 600 MHz NMR spectrometer. Relaxation rates and isotropic correlation times were calculated using SPARKY[44] and TENSOR2,[55] respectively.

Structural analysis, comparison and graphical representation

Analyses of the completed structures were performed using WHATCHECK,[56] VADAR,[57] PROCHECK[58] and PROCHECK_NMR.[59] Secondary structure boundaries for the ensemble were defined according to PROMOTIF.[60] Surface areas were calculated with VADAR.[57] Structural comparisons were done by ARIA[18] and LSQKAP within the CCP4 program suite.[61] Figures were prepared using MOLMOL[62] and GRASP.[63]

Static light-scattering

The oligomerization state of RneS1$^{35-125}$ was analyzed using a Superdex 75 HPLC gel exclusion column interfaced to a WYATT Technology MiniDawn light-scattering unit with an in-line Optilab DSP interferometric refractometer. The varying concentrations of protein were suspended in a solution containing 20 mM Hepes (pH 7.5), and 150 mM NaCl.

Glutaraldehyde crosslinking

Chemical crosslinking reactions were carried out by incubating 50 μl of 45 μM total protein in 20 mM phosphate buffer (pH 6.5), 50 mM NaCl with 0.01% (v/v) glutaraldehyde at 37 °C for two minutes. The reactions were quenched with 10 μl of 300 mM Tris–HCl (pH 8.0) and stored on ice. The proteins were subsequently concentrated by the addition of 40 μl of 17.5% (v/v) trichloroacetic acid, storage on ice for 30 minutes, centrifugation at 13,000 rpm for 15 minutes at 4 °C (Sorvall S534 rotor), and washing of the precipitate twice with 200 μl of 80% (v/v) acetone. Samples were dissolved in 50 mM Tris–HCl (pH 6.8) containing 0.1% (w/v) SDS, 50 mM DTT, 5% glycerol and boiled prior the separation on a SDS-15% PAGE gel.

Oligonucleotide binding studies

The titration of 300 μM ^{15}N-labeled RneS1$^{35-125}$ with 0, 0.2, 0.4, 0.6, 0.8, 1, 1.1, 1.5, 2.0 and 3.0 equivalents of single-stranded 5′-d(ACAGTATTTG)-3′ DNA (Sigma/Genosys) was monitored using ^{1}H–^{15}N HSQC spectra recorded at 30 °C on a 500 MHz spectrometer. Equilibrium K_D values were obtained by non-linear least-squares fitting of the data to the Langmuir isotherm describing the binding of one DNA molecule to a single protein site in the fast exchange limit.[64]

Protein Data Bank accession numbers

The atomic coordinates of RneS1$^{35-125}$ (accession codes 1SLJ, 1SMX, 1SN8) have been deposited in the Protein Data Bank of the Research Collaboratory for Structural Bioinformatics†, and the NMR chemical shifts (entry number 6122) in the BioMagResBank‡.

† http://www.rcsb.org/
‡ http://www.bmrb.wisc.edu/

References

1. Coburn, G. A. & Mackie, G. A. (1999). Degradation of mRNA in *Escherichia coli*: an old problem with some new twists. *Prog. Nucl. Acid Res. Mol. Biol.* **62**, 55–108.

2. Steege, D. A. (2000). Emerging features of mRNA decay in bacteria. *RNA*, **6**, 1079–1090.

3. Regnier, P. & Arraiano, C. M. (2000). Degradation of mRNA in bacteria: emergence of ubiquitous features. *BioEssays*, **22**, 235–244.

4. McDowall, K. J., Lin-Chao, S. & Cohen, S. N. (1994). A + U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *J. Biol. Chem.* **269**, 10790–10796.

5. McDowall, K. J. & Cohen, S. N. (1996). The N-terminal domain of the rne gene product has RNase E activity and is non-overlapping with the arginine-rich RNA-binding site. *J. Mol. Biol.* **255**, 349–355.

6. Ow, M. C. & Kushner, S. R. (2002). Initiation of tRNA maturation by RNase E is essential for cell viability in *E. coli*. *Genes Dev.* **16**, 1102–1115.

7. Vanzo, N. F., Li, Y. S., Py, B., Blum, E., Higgins, C. F., Raynal, L. C. *et al.* (1998). Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. *Genes Dev.* **12**, 2770–2781.

8. Bycroft, M., Hubbard, T. J., Proctor, M., Freund, S. M. & Murzin, A. G. (1997). The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell*, **88**, 235–242.

9. Murzin, A. G. (1993). OB(oligonucleotide/oligo-saccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* **12**, 861–867.

10. Arcus, V. (2002). OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr. Opin. Struct. Biol.* **12**, 794–801.

11. Theobald, D. L., Mitton-Fry, R. M. & Wuttke, D. S. (2003). Nucleic acid recognition by OB-fold proteins. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 115–133.

12. McDowall, K. J., Hernandez, R. G., Lin-Chao, S. & Cohen, S. N. (1993). The ams-1 and rne-3071 temperature-sensitive mutations in the ams gene are in close proximity to each other and cause substitutions within a domain that resembles a product of the *Escherichia coli* mre locus. *J. Bacteriol.* **175**, 4245–4249.

13. Diwa, A. A., Jiang, X., Schapira, M. & Belasco, J. G. (2002). Two distinct regions on the surface of an RNA-binding domain are crucial for RNase E function. *Mol. Microbiol.* **46**, 959–969.

14. Callaghan, A. J., Grossmann, J. G., Redko, Y. U., Ilag, L. L., Moncrieffe, M. C., Symmons, M. F. *et al.* (2003). Quaternary structure and catalytic activity of the *Escherichia coli* Ribonuclease E amino-terminal catalytic domain. *Biochemistry*, **42**, 13848–13855.

15. Letunic, I., Goodstadt, L., Dickens, N. J., Doerks, T., Schultz, J., Mott, R. *et al.* (2002). Recent improvements to the SMART domain-based sequence annotation resource. *Nucl. Acids Res.* **30**, 242–244.

16. Schubert, M., Smalla, M., Schmieder, P. & Oschkinat, H. (1999). MUSIC in triple-resonance experiments: amino acid type-selective ${}^{1}H–{}^{15}N$ correlations. *J. Magn. Reson.* **141**, 34–43.

17. Schubert, M., Oschkinat, H. & Schmieder, P. (2001). MUSIC, selective pulses, and tuned delays: amino acid type-selective ${}^{1}H–{}^{15}N$ correlations, II. *J. Magn. Reson.* **148**, 61–72.

18. Linge, J. P., O'Donoghue, S. I. & Nilges, M. (2001). Automated assignment of ambiguous nuclear Overhauser effects with ARIA. *Methods Enzymol.* **339**, 71–90.

19. Cornilescu, G., Delaglio, F. & Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, **13**, 289–302.

20. Chen, Y. W. (2001). Solution solution: using NMR models for molecular replacement. *Acta Crystallog. sect. D*, **57**, 1457–1461.

21. Bahadur, R. P., Chakrabarti, P., Rodier, F. & Janin, J. (2003). Dissecting subunit interfaces in homodimeric proteins. *Proteins: Struct. Funct. Genet.* **53**, 708–719.

22. Wagner, G. (1997). An account of NMR in structural biology. *Nature Struct. Biol.* **4**, 841–844.

23. Schurr, J. M., Babcock, H. P. & Fujimoto, B. S. (1994). A test of the model-free formulas. Effects of anisotropic rotational diffusion and dimerization. *J. Magn. Reson. ser. B*, **105**, 211–224.

24. Feng, W., Tejero, R., Zimmerman, D. E., Inouye, M. & Montelione, G. T. (1998). Solution NMR structure and backbone dynamics of the major cold-shock protein (CspA) from *Escherichia coli*: evidence for conformational dynamics in the single-stranded RNA-binding site. *Biochemistry*, **37**, 10881–10896.

25. Coburn, G. A., Miao, X., Briant, D. J. & Mackie, G. A. (1999). Reconstitution of a minimal RNA degradosome demonstrates functional coordination between a 3′ exonuclease and a DEAD-box RNA helicase. *Genes Dev.* **13**, 2594–2603.

26. Spickler, C. & Mackie, G. A. (2000). Action of RNase II and polynucleotide phosphorylase against RNAs containing stem-loops of defined structure. *J. Bacteriol.* **182**, 2422–2427.

27. McDowall, K. J., Kaberdin, V. R., Wu, S. W., Cohen, S. N. & Lin-Chao, S. (1995). Site-specific RNase E cleavage of oligonucleotides and inhibition by stem-loops. *Nature*, **374**, 287–290.

28. Holm, L. & Sander, C. (1995). Dali–a network tool for protein-structure comparison. *Trends Biochem. Sci.* **20**, 478–480.

29. Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536–540.

30. Kloks, C. P., Spronk, C. A., Lasonder, E., Hoffmann, A., Vuister, G. W., Grzesiek, S. & Hilbers, C. W. (2002). The solution structure and DNA-binding properties of the cold-shock domain of the human Y-box protein YB-1. *J. Mol. Biol.* **316**, 317–326.

31. Schindelin, H., Jiang, W., Inouye, M. & Heinemann, U. (1994). Crystal structure of CspA, the major cold shock protein of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **91**, 5119–5123.

32. Schnuchel, A., Wiltscheck, R., Czisch, M., Herrler, M., Willimsky, G., Graumann, P. *et al.* (1993). Structure in solution of the major cold-shock protein from *Bacillus subtilis*. *Nature*, **364**, 169–171.

33. Schindelin, H., Marahiel, M. A. & Heinemann, U. (1993). Universal nucleic acid-binding domain revealed by crystal structure of the *B. subtilis* major cold-shock protein. *Nature*, **364**, 164–168.

34. Makhatadze, G. I. & Marahiel, M. A. (1994). Effect of pH and phosphate ions on self-association properties of the major cold-shock protein from *Bacillus subtilis*. *Protein Sci.* **3**, 2144–2147.

35. Condon, C. & Putzer, H. (2002). The phylogenetic distribution of bacterial ribonucleases. *Nucl. Acids Res.* **30**, 5339–5346.

36. Phadtare, S. & Inouye, M. (1999). Sequence-selective interactions with RNA by CspB, CspC and CspE, members of the CspA family of *Escherichia coli*. *Mol. Microbiol.* **33**, 1004–1014.

37. Briant, D. J., Hankins, J. S., Cook, M. A. & Mackie, G. A. (2003). The quaternary structure of RNase G from *Escherichia coli*. *Mol. Microbiol.* **50**, 1381–1390.

38. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Macromol. Crystallog. sect. A*, **276**, 307–326.

39. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849–861.

40. Terwilliger, T. C. (2000). Maximum-likelihood density modification. *Acta Crystallog. sect. D*, **56**, 965–972.

41. McRee, D. E. (1999). XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165.

42. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.

43. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multi-dimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.

44. Goddard, T. D. & Kneeler, D. G. (1999). *Sparky 3*, 3rd edit., University of California, San Franscisco.

45. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E. *et al.* (1995). ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR*, **6**, 135–140.

46. Johnson, P. E., Joshi, M. D., Tomme, P., Kilburn, D. G. & McIntosh, L. P. (1996). Structure of the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC determined by nuclear magnetic resonance spectroscopy. *Biochemistry*, **35**, 14381–14394.

47. Kanelis, V., Donaldson, L., Muhandiram, D. R., Rotin, D., Forman-Kay, J. D. & Kay, L. E. (2000). Sequential assignment of proline-rich regions in proteins: application to modular binding domain complexes. *J. Biomol. NMR*, **16**, 253–259.

48. Yamazaki, T., Forman-Kay, J. D. & Kay, L. E. (1993). Two-dimensional NMR experiments for correlating ^{13}Cβ and ^{1}Hδ/ε chemical shifts of aromatic residues in ^{13}C-labeled proteins *via* scalar couplings. *J. Am. Chem. Soc.* **115**, 11054–11055.

49. Slupsky, C. M., Gentile, L. N. & McIntosh, L. P. (1998). Assigning the NMR spectra of aromatic amino acids in proteins: analysis of two Ets pointed domains. *Biochem. Cell. Biol.* **76**, 379–390.

50. Pelton, J. G., Torchia, D. A., Meadow, N. D. & Roseman, S. (1993). Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIGlc, a signal-transducing protein from *Escherichia coli*, using two-dimen-

51. McIntosh, L. P., Brun, E. & Kay, L. E. (1997). Stereo-specific assignment of the NH$_2$ resonances from the primary amides of asparagine and glutamine side chains in isotopically labeled proteins. *J. Biomol. NMR*, **9**, 306–312.

52. Zwahlen, C., Gardner, K. H., Sarma, S. P., Horita, D. A., Byrd, R. A. & Kay, L. E. (1998). An NMR experiment for measuring methyl–methyl NOEs in ^{13}C-labeled proteins with high resolution. *J. Am. Chem. Soc.* **120**, 7617–7625.

53. Schubert, M., Labudde, D., Oschkinat, H. & Schmieder, P. (2002). A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on ^{13}C chemical shift statistics. *J. Biomol. NMR*, **24**, 149–154.

54. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G. *et al.* (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry*, **33**, 5984–6003.

55. Dosset, P., Hus, J. C., Blackledge, M. & Marion, D. (2000). Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR*, **16**, 23–28.

56. Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. (1996). Errors in protein structures. *Nature*, **381**, 272.

57. Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R. F., Sykes, B. D. & Wishart, D. S. (2003). VADAR: a web server for quantitative evaluation of protein structure quality. *Nucl. Acids Res.* **31**, 3316–3319.

58. Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. (1993). Procheck—a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.

59. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**, 477–486.

60. Hutchinson, E. G. & Thornton, J. M. (1996). PROMOTIF–a program to identify and analyze structural motifs in proteins. *Protein Sci.* **5**, 212–220.

61. Bailey, S. (1994). The CCP4 suite–programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.

62. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55. see also pp. 29–32..

63. Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281–296.

64. Johnson, P. E., Tomme, P., Joshi, M. D. & McIntosh, L. P. (1996). Interaction of soluble cellooligosaccharides with the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC 2. NMR and ultra-violet absorption spectroscopy. *Biochemistry*, **35**, 13895–13906.

65. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673–4680.

sional heteronuclear NMR techniques. *Protein Sci.* **2**, 543–558.

SCIENCE *d* DIRECT ®

www.sciencedirect.com

Supplementary Material for this paper comprising four Figures illustrating the ^{1}H–^{15}N HSQC spectra and CD-monitored thermal denaturation profiles of WT and G66S RneS1$^{35-125}$, as well as the location of the nucleic acid binding sites within the protein dimer, is available on Science Direct