Structural and functional analysis of RNA and TAP binding to SF2/ASF

Aura M. Tintaru, Guillaume M. Hautbergue, Andrea M. Hounslow, Ming-Lung Hung,

Lu-Yun Lian, C. Jeremy Craven and Stuart A. Wilson

Supplementary Information

Methods

NMR analysis Because of the intrinsic insolubility of SF2/ASF RRM2, NMR spectra for structural calculations were carried out in a buffer containing 50mM L-Arg and 50mM L-Glu which allowed concentration of the protein to 1mM as described previously (Tintaru, 2007) Spectra were acquired on a Bruker Avance 800MHz spectrometer at 303K. Backbone assignment was achieved using data from 3D HNCO, HNCACO, HNCA, CBCACONH, HNCACB spectra and a monte carlo assignment program (Reed et al., 2003). Sidechain assignment was achieved using HCCH- and CCH-TOCSY experiments and using information from the 3D NOESY spectrum. Structural constraints were obtained from a 3D simultaneously 15N/13C-edited NOESY experiment (100ms mixing time). 15N T2 values were measured as described elsewhere (Farrow et al., 1994), with relaxation delays in the range 20-240ms.

Structure calculation. NOEs were assigned in the NOESY spectrum using FELIX, and in-house macros and UNIX scripts. NOE intensities were converted to upper distance bounds using a smooth interpolation function based on the observed intensities for a range of covalently fixed distances. Dihedral constraints were obtained from the program TALOS. Structures were calculated using the standard anneal protocol in CNS vsn 1.1. 100 structures were calculated and the 20 lowest energy structures were chosen to represent the ensemble.

Chemical shift mapping. RNA (Dharmacon) was resuspended in 20mM Na phosphate pH 6.0, 50 mM NaCl, 5 mM EDTA. The RNA was titrated into a 15N-labelled protein samples (0.3mM initial concentration) lacking L-Arg + L-Glu and 15N HSQC spectra were recorded after each addition (Fig. 2). The weighted amide chemical shift change, $w\Delta\delta$, was calculated as $w\Delta\delta = \sqrt{\{\Delta\delta H2 + (\Delta\delta N/10)^2\}}$. For the closest comparison with the UV-crosslinking data, additional HSQC based titrations (Fig S4) were performed with the same shortened form of the SF2/ASF RRM2 comprising residues 107-196. Due to solubility limitations, these spectra (Fig S4) were acquired in the same buffer as for the structure determination, *i.e.* including 50 mM L-Glu, 50 mM L-Arg. The protein concentration was *ca.*0.2mM (except as noted below) and HSQC spectra were collected at 0, 0.25, 0.5, 0.75 and 1 equivalents of added RNA for the wild type sequence, H183A, and R117/118A proteins, and at 0, 0.25, 0.5, and 1 equivalents of added RNA for the W134A protein.

References

Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D. and Kay, L.E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry*, 33, 5984-6003.

Reed, M.A., Hounslow, A.M., Sze, K.H., Barsukov, I.G., Hosszu, L.L., Clarke, A.R., Craven, C.J. and Waltho, J.P. (2003) Effects of domain dissection on the folding and stability of the 43 kDa protein PGK probed by NMR. *J Mol Biol*, 330, 1189-1201.

Tintaru, A.M., Hautbergue, G.M., Hounslow, A.M., Lian, L.Y., Wilson, S.A and Craven, C.J. (2007) Assignment of 1H, 13C and 15N resonances for ASF/SF2/ASF RNA recognition motif 2. *J Biomol NMR*, In press.

Table I

Statistics for an ensemble of the 20 lowest energy structures of SF2/ASF amino acids 107-215

NOE distances	
Intraresidue	349
Sequential	248
Medium Range	111
Long range (Δ>4)	254
Dihedral restraints	76
Hydrogen bond restraints	48
NOE violations > 0.2Å	3
Dihedral restraint violation > 2.5°	0
R.m.s. deviations from mean structure (Å)*	
Backbone heavy atoms	0.6
All heavy atoms	1.3
Ramachandran analysis*	
Most favoured regions (%)	84
Additionally allowed regions (%)	16
Generously allowed regions (%)	0.6
Disallowed regions (%)	0.4

* amino acids 121-195

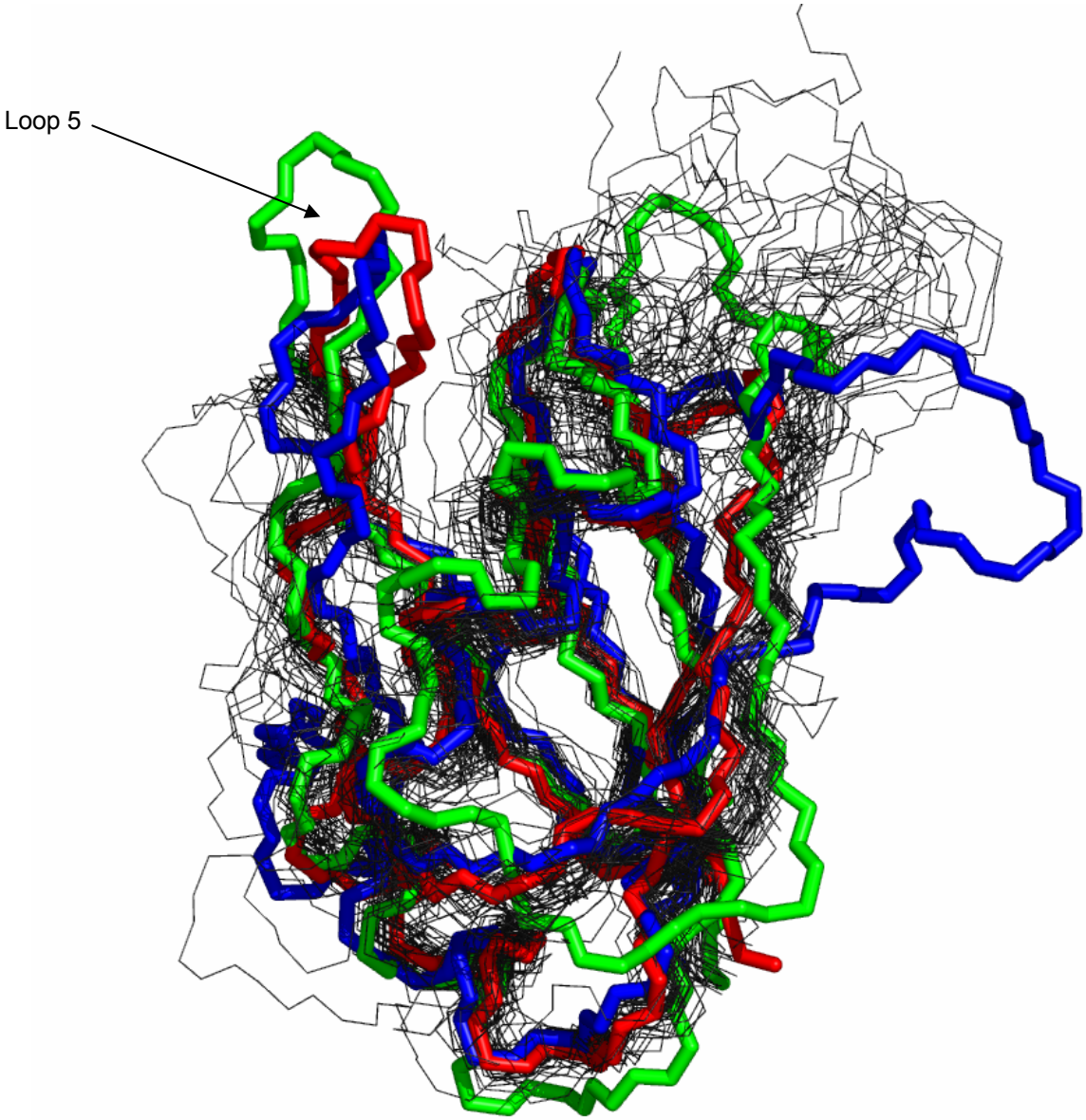


Figure S1. Overlay of structures from DALI. The structured part of a representative structure from the ensemble calculated for SF2/ASF was submitted to the DALI server (www.ebi.ac.uk/dali). Hits with a Z score >6 (44 hits, and representing the vast majority of proteins annotated as RNA-BINDING) are represented above, overlaid on the SF2/ASF structure (in red). Loop 5 is marked. Structures with a similarly extended loop 5 are shown in green (1uw4) and blue (2dnl). Overlay translations and rotations were as determined by DALI. The figure is in the same orientation as Figures 1A and 1B

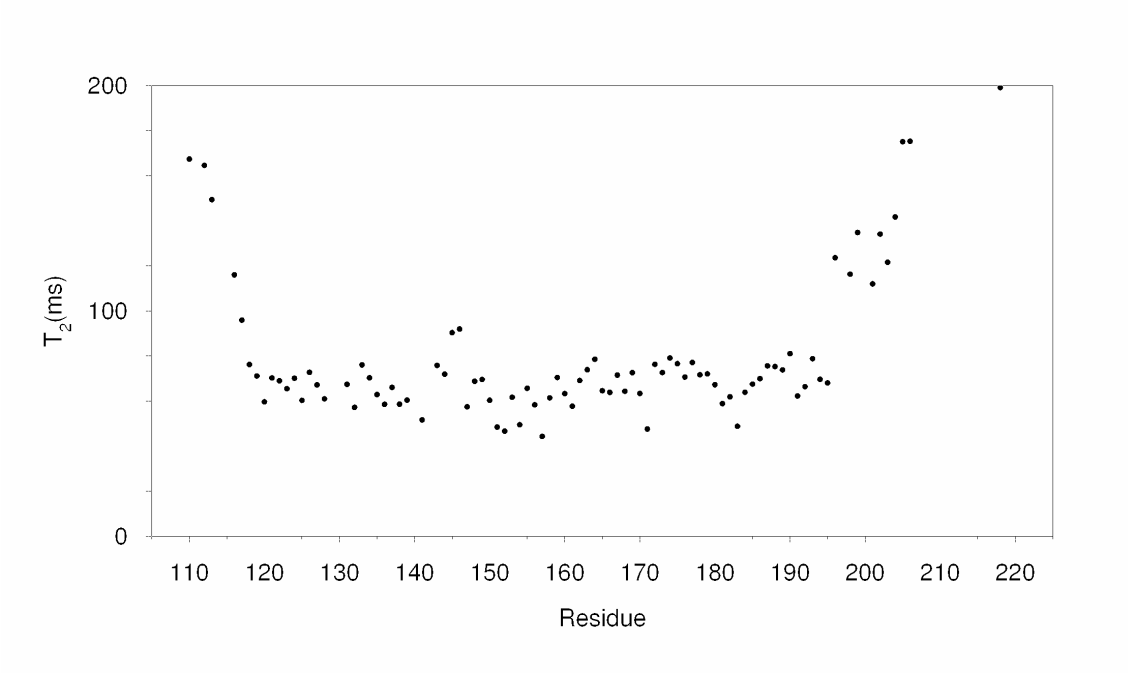
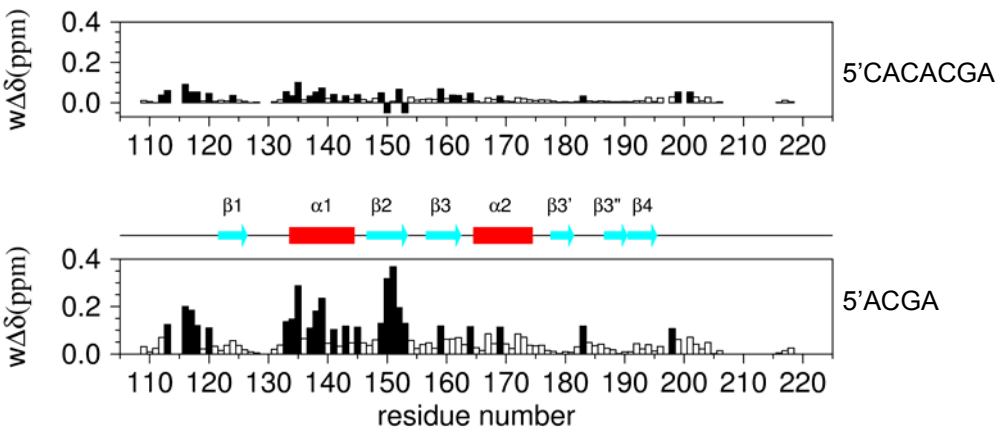


Figure S2. ^{15}N T2 data for SF2/ASF.

A



B

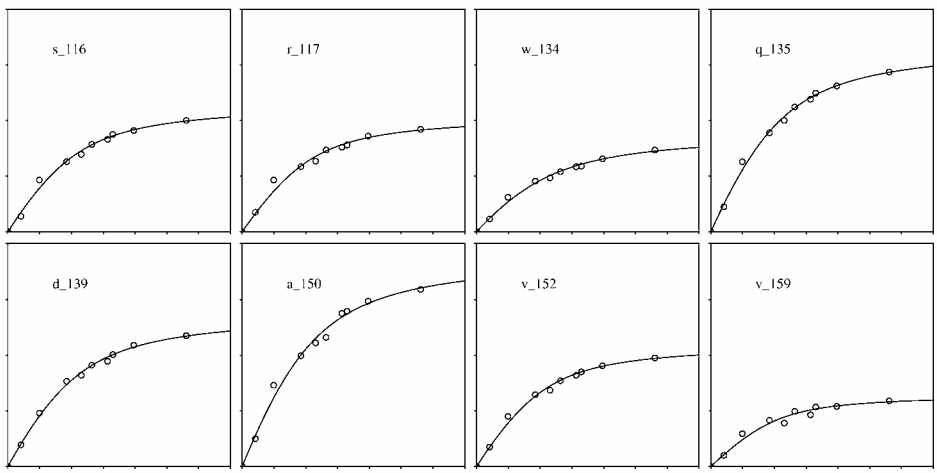


Figure S3.

(A) Comparison of chemical shift changes in SF2/ASF upon addition of the 4 nucleotide RNA ACGA (1.7 molar equivalents, lower plot), and the consensus sequence CACACGA (0.7 molar equivalents, upper plot). The data for ACGA is plotted as in Figure 2B. The data for CACACGA is plotted similarly, with filled bars corresponding to shift changes greater than 0.03 ppm. Small negative bars indicate residues which became attenuated upon addition of RNA to such an extent that shift changes could not be measured. The direction of the changes in the 2D HSQC spectra are also well correlated between the two ligands (data not shown).

(B) Titration curves for selected residues of SF2/ASF upon addition of the RNA ACGA. Each graph shows the weighted chemical shift change versus concentration of RNA. The x axis in all cases spans 0 to 700 μM, and the y axis spans 0 to 0.4 ppm. The curves are fits to a simple binding equation, with the total chemical shift change and Kd as fitted parameters.

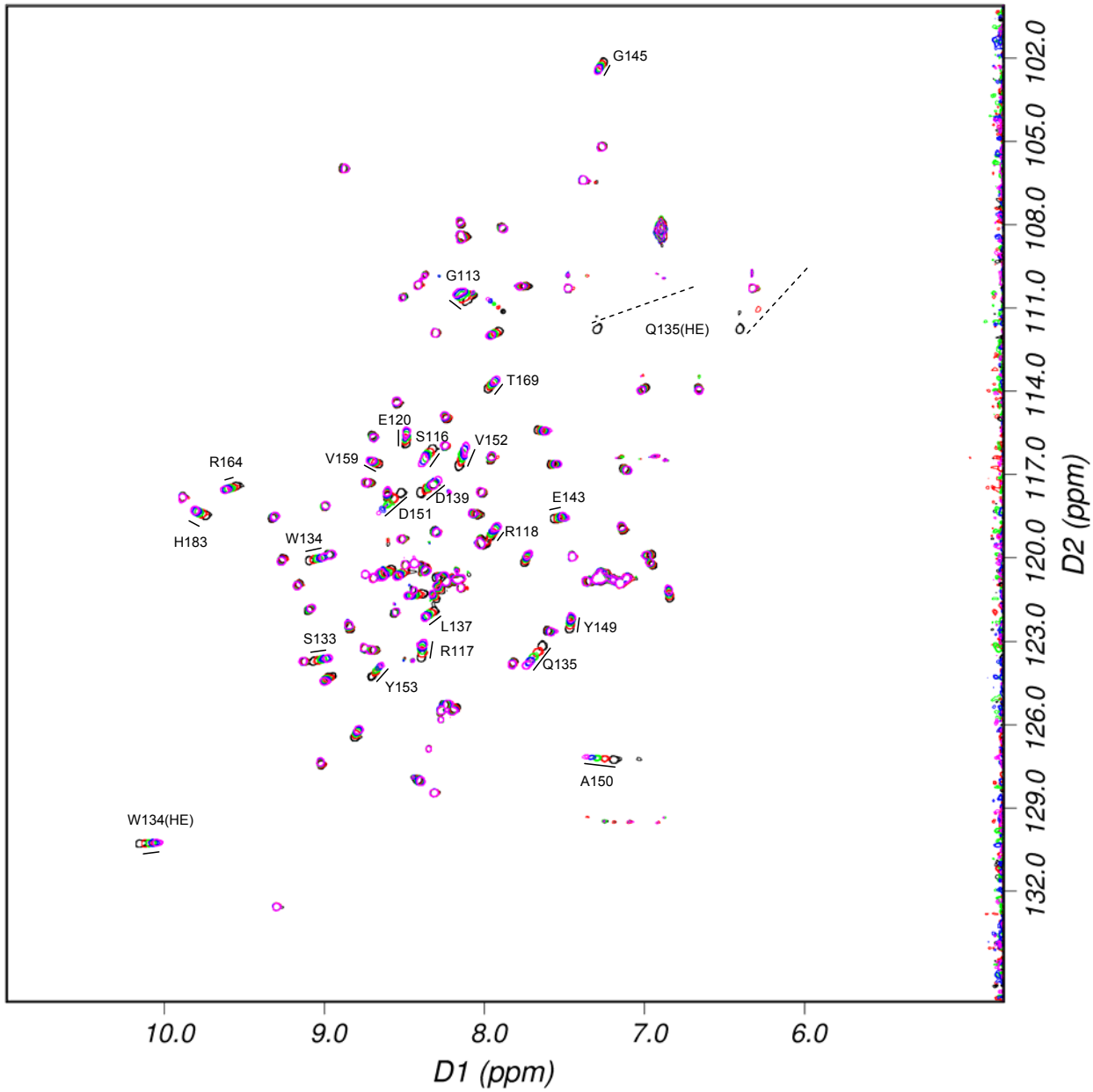


Figure S4(a)

The 5 HSQC spectra acquired for the wild type SF2/ASF RRM2 (aa107-196) are shown overlaid. The spectra are plotted in black (0 equivalents RNA added), red (0.25 equivalents), green (0.5 equivalents), blue (0.75 equivalents) and magenta (1 equivalent). The chemical shift changes observed are very similar to those observed in the titration reported in the main body of the text (see Figure S4(b)). The exchange kinetics are slightly more rapid, as it is possible to observe the crosspeaks for A150 across the entire titration, and also to observe the crosspeak for one of the sidechain amide protons of Q135 after the first addition of RNA. This is presumably due to the addition of arginine and glutamate to the buffer. This titration acts as a control that the reagents behave equivalently to those used in the titration reported in the main body of the text.

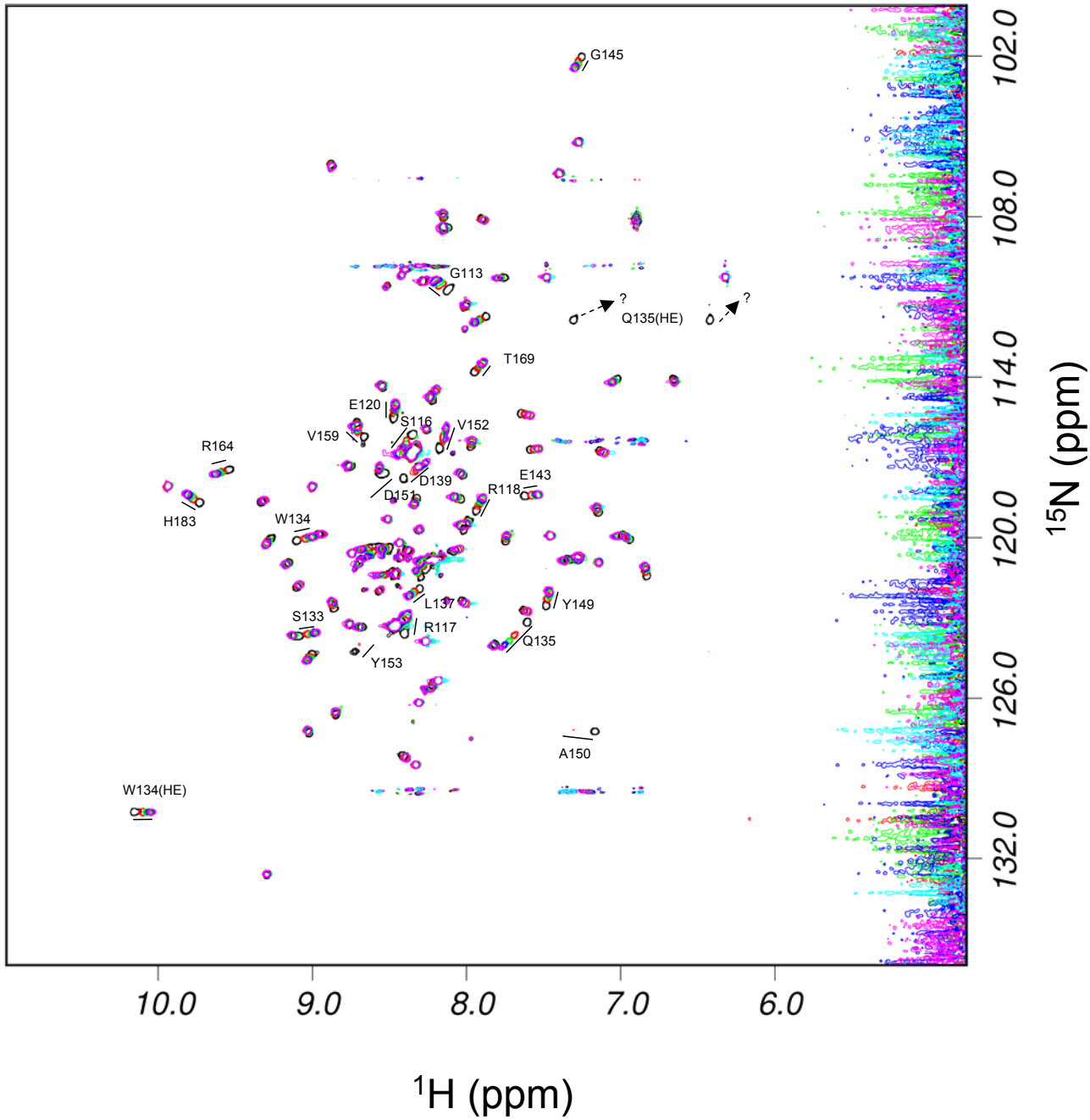


Figure S4(b)

5 HSQC spectra from the titration reported in the main body of the paper (SF2/ASF residues 107-215) are replotted for comparison with the other supplementary titration figures. This titration was performed at a different set of added RNA equivalents; the five most similar points to the supplementary titrations have been selected here. The spectra are plotted in black (0 equivalents RNA added), red (0.3 equivalents), green (0.56 equivalents), blue (0.8 equivalents) and magenta (1 equivalent). Following addition of RNA, the crosspeaks for A150, D151 and Y153 were not observable at this contour level (see Fig 2 in the main body of the paper).

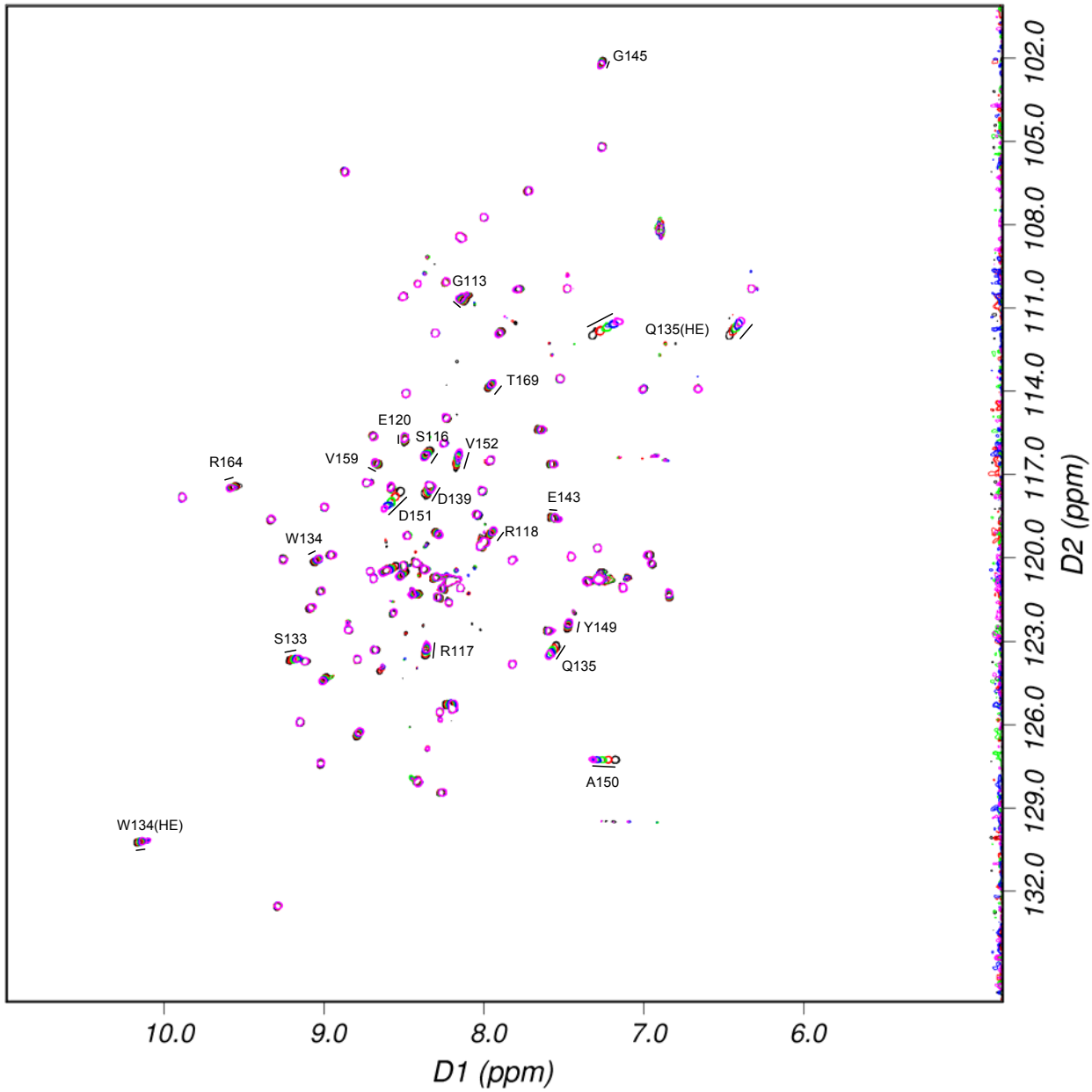


Figure S4(c)

The 5 HSQC spectra acquired for the H183A SF2/ASF RRM2 are shown overlaid, plotted as for Figure S4(a). The spectrum for the apo protein is very similar to that for the wild type SF2/ASF RRM2, except in the immediate vicinity of the mutation, indicating that the mutation does not seriously disrupt the structure. Notably, the resonances of W134 and of the sidechain of Q135 are not strongly affected by the mutation.

For those crosspeaks that can be readily assigned, the directions of movement with increasing RNA concentration are very closely similar to those for the wild type SF2/ASF RRM2, indicating that the mode of binding is not significantly altered. The magnitudes of the movements are, in general, smaller than for the wild type protein which is consistent with a reduced loading of the protein due to a reduced binding affinity. The binding dynamics are also more rapid, as both the crosspeaks for sidechain amide resonances of Q135 are observable across the titration. This is also consistent with a reduced binding affinity for this mutant.

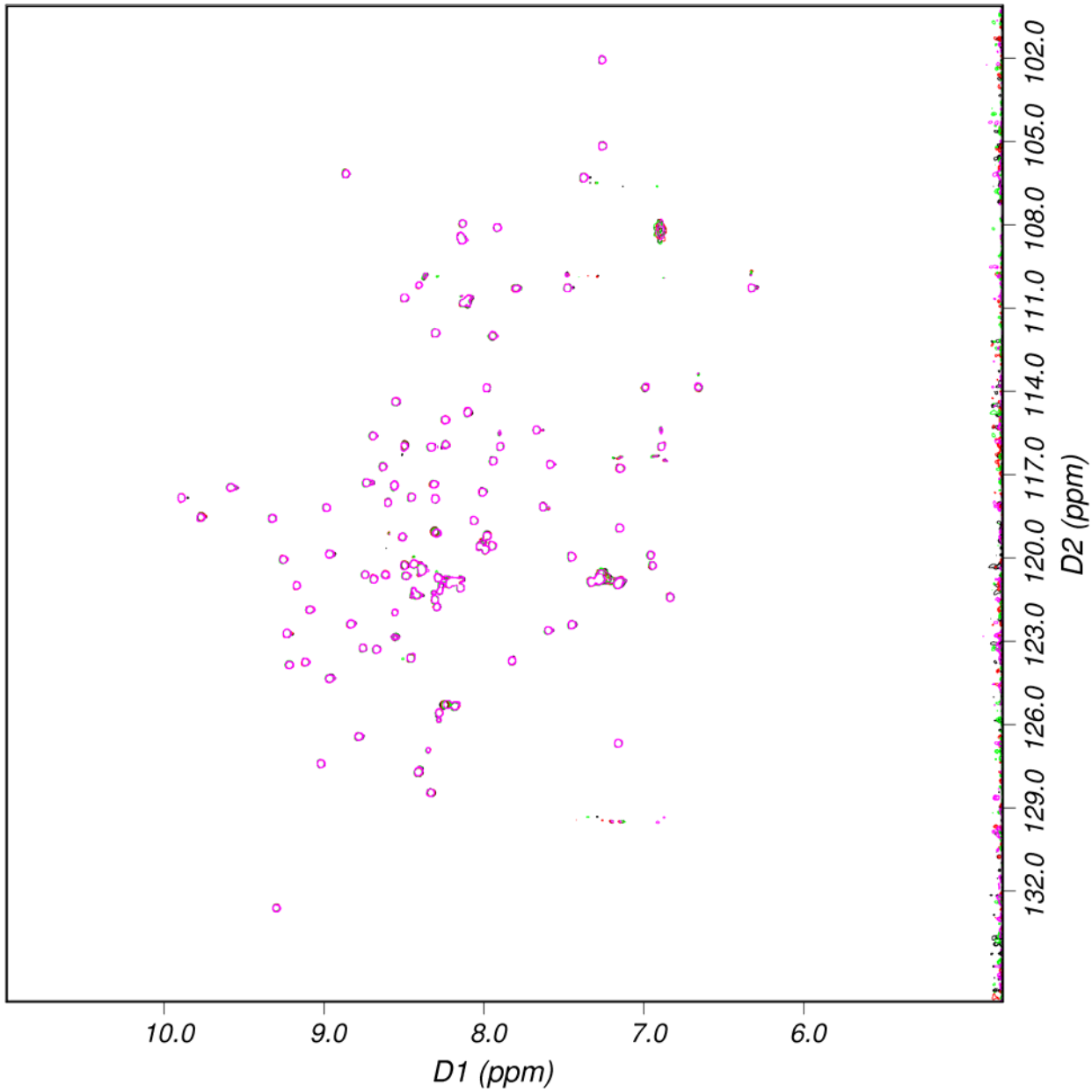


Figure S4(d)

The 4 HSQC spectra acquired for the W134A SF2/ASF RRM2 are shown overlaid, plotted essentially as for Figure S4(a) (except that no spectrum was acquired for 0.75 equivalents, as negligible shift changes were observed in the earlier spectra). The spectrum for the apo protein is in the main very similar to that for the apo wild type SF2/ASF RRM2, indicating that mutation of W134 does not significantly perturb the overall fold of the protein, consistent with its surface exposed position. More particularly, the pattern of sequential dNN NOEs running through helix 1 can be still be clearly traced in a 3D 15N edited NOESY spectrum of the mutant protein, indicating that helix 1 is intact. The spectra are almost perfectly superimposable across the entire titration indicating that the binding is essentially abolished, consistent with the central importance of this residue in RNA binding.

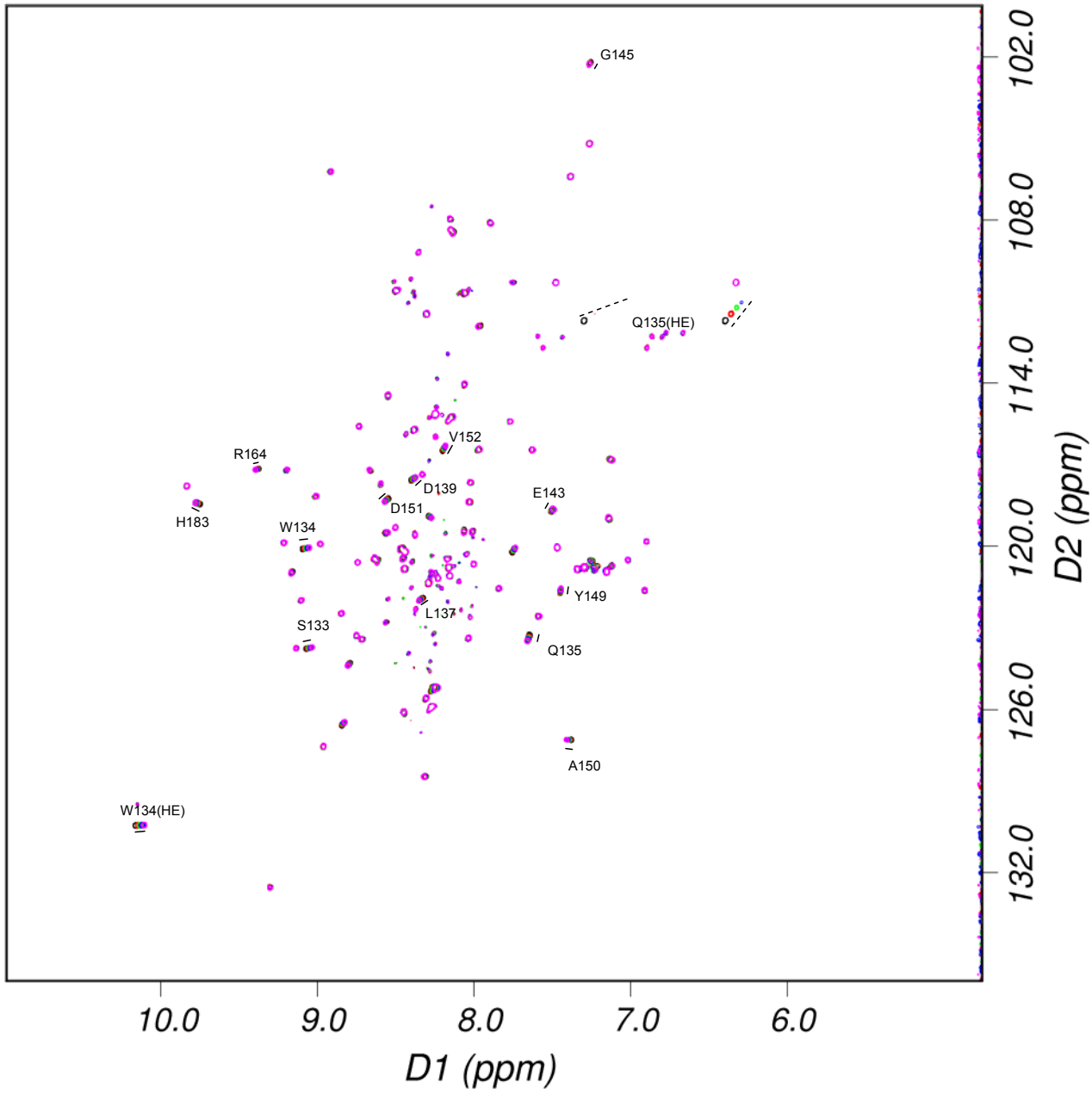


Figure S4(e) – legend overleaf

Figure S4(e)

The 5 HSQC spectra acquired for the R117/118A SF2/ASF RRM2 are shown overlaid, plotted as for Figure S4(a). This mutant expressed poorly, the concentration of sample being only 55% of that of the sample in Figure S4(a), and the sample contained a proportion (*ca*. 20%) of unfolded protein.

Surprisingly, although the overall HSQC spectrum of the apo protein is similar to that for the wild type SF2/ASF RRM2, there are some significant shift changes in regions distant from the mutation. This suggests that there are transient interactions between the R117/118 region of the otherwise unstructured N-terminal region of the protein and the folded domain. Residues showing moderate shift changes ($w\Delta\delta$~0.2-0.4 ppm) are R122, V123, V147, Y149, A150, E161, R164. Residues showing smaller shift changes ($w\Delta\delta$ ~0.1-0.2 ppm) are S116, S119, E120, V124, V125, E143, V159, K165, E166, D167, M168, T169, V172, R173 and D176. The affected residues do not suggest a single mode of interaction, however the larger shift changes mostly lie in a cluster including Y149 and A150 which are implicated in RNA binding (see main text). Transient interactions with the folded domain are also consistent with the observation that whilst the structure calculations indicate the N terminal domain to the largely unstructured up to residue 120, the 15N relaxation measurements suggest that the residues from approximately 117 onwards have reduced mobility.

The shift changes arising from the mutation reduce the number of crosspeaks that can be simply reassigned. Nevertheless, the shift changes are much smaller for this mutant, and (similarly to H183A) the binding dynamics are more rapid. It is not possible from this data to say whether the reduction in affinity is directly due to loss of interactions of the mutated residues with the RNA, or whether it is a secondary affect arising from the loss of transient interactions with the folded domain.