DOI: 10.1002/cbic.200800786

Probing Mutation-Induced Structural Perturbations by Refinement Against Residual Dipolar Couplings: Application to the U4 Spliceosomal RNP Complex

John P. Kirkpatrick,[a] Ping Li,[b] and Teresa Carlomagno*[a]

Confident interpretation of biochemical experiments performed with mutated proteins relies on verification of the integrity of the mutant structures. We present a simple and rapid refinement protocol for comparing the structures of mutated and wild-type proteins. Our approach involves measurement of residual dipolar couplings, and only requires assignment of the backbone resonances of the mutant species. We demonstrate application of the protocol to a mutant of the 15.5K protein, a core component of the U4 spliceosomal ribonucleoprotein (RNP) complex. Confirmation of the unperturbed structure of the mutated protein prompted re-examination of a previous mutagenesis study and indicated that the interpretation of mutant binding affinities in terms of direct interfacial contacts should be applied with caution.

Introduction

Protein mutagenesis experiments are a powerful biochemical tool and are frequently employed to probe protein–protein interactions. Measurement of binding affinities with specifically mutated binding partners should reveal those residues that contribute to the interaction energy. When the structures of the individual components are available, the identification of such residues enables mapping of the putative binding interface.

This interpretation of mutant binding affinities assumes that the mutations do not introduce structural changes within the individual components of the complex; however, mutation of residues not involved in the binding interaction could lead to reduced binding affinities as a result of perturbations to the global structure of the mutated protein. Full structure determinations for the mutated proteins would be a laborious task, and this highlights the need for a method by which the structural integrity of mutated proteins with respect to the known wild-type structure can be rapidly assessed. We propose the application of NMR-derived residual dipolar coupling restraints to rapidly derive the backbone structure of mutated proteins.

Residual dipolar couplings (RDCs) have become increasingly important in the field of solution-phase NMR spectroscopy since their first application over ten years ago.[1,2] They provide information on the orientation of internuclear vectors in a global molecular frame, and hence can be converted into distance-independent structural restraints, thus complementing the more traditional NOE information that yields only short-range restraints. In combination with NOE-derived restraints, RDCs are employed to improve the precision and accuracy of solution-state NMR structures. They are particularly important for the structure determination of large proteins in which the requirement for perdeuteration leads to a dramatically reduced NOE density[3] and for defining relative domain orientations in situations where few interdomain NOEs are measurable.[4] More recently, methods have been developed to determine ab initio structures by using solely RDC-derived information.[5,6]

In structure calculations that supplement traditional NOE data with RDC information, the RDC-derived restraints are typically introduced only during the latter stages of the calculation, when the global fold of the protein has already been defined from the NOEs. In this sense, the RDCs are used to improve an approximate structure. Similarly, RDCs can be used for refinement of structural models for a protein for which the models may be based on the structure of a related homologue, or, for mutants, on the structure of the wild-type protein. It is in this last context that we seek to apply RDCs. By refining a wild-type protein structure against RDCs measured for a mutant we can generate a structure for the mutant protein. Comparison of the wild-type and mutant structures will reveal if the mutations have introduced any structural perturbations. The strength of the method lies in the rapidity with which the model of the mutant structure can be generated. In contrast to a full solution-state NMR structure determination, only the backbone resonances need be assigned, and measurement of the RDCs themselves is becoming increasingly straightforward.

Here, we apply this approach to a mutant of the human 15.5K protein. This protein is a core component in the U4/U6–U5 ribonucleoprotein (RNP) intermediate of the mRNA splicing

[a] *Dr. J. P. Kirkpatrick, Dr. T. Carlomagno*
Structural and Computational Biology,
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg (Germany)
Fax: (+49) 6221-3878519
E-mail: teresa.carlomagno@embl.de

[b] *Dr. P. Li*
Department of NMR-based Structural Biology
Max Planck Institute for Biophysical Chemistry
Am Fassberg 11, 37077 Goettingen (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.200800786.

cycle.[7] In this complex, residues of the U6 snRNA that are crucial for catalysis are shielded by base pairing with the U4 snRNA.[8,9] The formation of the U4/U6 complex is initiated by the recognition of the 5′-stem–loop (SL) region of the U4 snRNA, located between the two U4/U6 base-paired regions, by the 15.5K protein (Figure 1).[7] This highly conserved protein

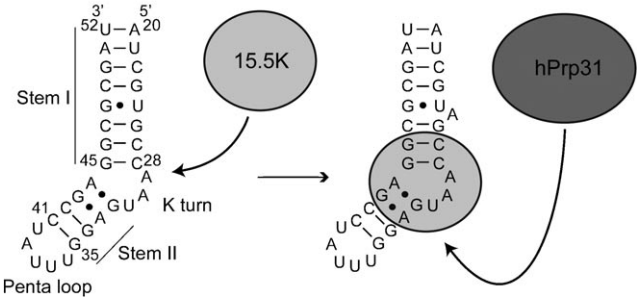

Figure 1. Schematic representation of the RNP complex examined in this study. In the first step of complex assembly, the 15.5K protein (light grey) binds to the K-turn region of the U4 5′-SL. The secondary binding protein hPrp31 (dark grey) is then able to bind to the preformed 15.5K–U4 5′-SL RNP.

recognizes and possibly stabilizes a particular sequence of the RNA that is known as the kink-turn (K-turn) motif.[10] In the hierarchical assembly pathway of the U4/U6 particle, the 15.5K–U4 5′-SL complex provides a binding platform for the human (h) Prp31 (U4/U6–61K) protein.[11] In the major spliceosome, the hPrp31–15.5K–U4 5′-SL ternary complex subsequently recruits the ternary hPrp3–hPrp4–CypH (U4/U6–90K–60K–20K) complex to the U4/U6 di-snRNP.[12]

In a previous study, it was reported that the 15.5K-2–U4 snRNA complex containing a mutant of 15.5K (15.5K-2) with a four-amino-acid mutation in the C-terminal region of helix α3 (E74R/D75K/K76M/N77T) binds hPrp31 with reduced affinity, as detected by pull-down assays.[13] However, in a recently solved crystal structure of the hPrp31–15.5K–U4 RNP,[14] the C terminus of helix α3 in 15.5K does not make direct structural contacts with hPrp31 (Figure 2). To investigate the basis of the putative functional deficiency of the 15.5K-2 mutant, we first generated a structural model for 15.5K-2 by refining the wild-type crystal structure against RDCs measured on the mutant. Our aim was to verify that the four-amino-acid mutation does not induce significant structural perturbations in 15.5K-2, particularly in regions distant from the mutation site that could distort the hPrp31-binding interface. Second, we conducted a detailed band-shift analysis of the interaction of hPrp31 with single and double-point mutants of the 15.5K–U4 snRNA complex to pin down the mutation site that is most responsible for the weakened binding to hPrp31. The results from the RDC refinement indicated that the structure of 15.5K is unperturbed by the four-residue mutation in 15.5K-2, while the band-shift experiments confirmed the weakened interaction between the 15K-2–U4 RNP and hPrp31. Furthermore, we were able to attribute the reduced affinity of the 15.5K-2–U4 snRNA complex for hPrp31 to a long-range electrostatic repulsion between K75 of

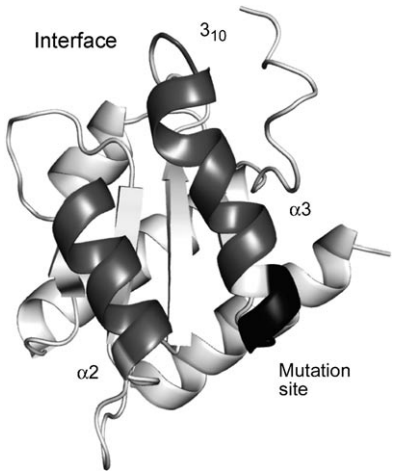

Figure 2. Ribbon diagram of 15.5K highlighting the interfacial region with hPrp31 (dark grey) and the site of mutation in 15.5K-2 (black). The mutation site does not lie within the 15.5K–hPrp31 interface.

15.5K-2 and K243 of hPrp31, rather than to disruption of an important intermolecular contact or to an indirect structural effect on the α2/α3 surface of 15.5K. This result suggests prudence in interpreting mutagenesis data in terms of direct structural contacts.

Results

Designing the refinement protocol

Significant changes in the local conformation of a mutant protein are mostly expected at the mutation site. However, in the case of the 15.5K-2 mutant, whose complex with the U4 RNA has a reduced affinity for hPrp31, the mutation site is not located at the protein–protein interface, so that any changes in the local structure (if they occur) are not expected to be responsible for the reduced binding affinity. Therefore, we decided to address the possibility that the mutations might introduce small but significant structural perturbations elsewhere in the protein structure, and that such perturbations might be responsible for the weakened binding to hPrp31. Long-range structural distortions would be expected to principally involve orientational changes in secondary structure elements, as well as possible changes in the local geometry. The refinement protocol is constructed in such a way as to explore both sources of structural differences, namely the rearrangement of secondary structure elements as well as changes in the local conformation at both the mutation and distant sites. In the early stages of the refinement protocol we seek to correctly adjust the orientations of the secondary structure elements in which there are differences between these orientations in the wild-type and mutant structures, while in the latter stages we allow for adjustments of the local geometry.

When RDCs are employed in traditional NMR structure determinations, they are supplemented by NOE data and, frequently, backbone dihedral restraints obtained from chemical shifts and coupling constants. In contrast, for the purpose described

in this paper, the RDCs are the only available experimental data. Therefore, in the first stages of the protocol, it is necessary to apply nonexperimental restraints to maintain the secondary structure conformation and prevent the translation of the RDC restraints into backbone conformational distortions that would leave the possibility of overall reorientations of the secondary structure elements unexplored. In the latter stages of the protocol, these non-experimental restraints are relaxed to allow backbone adjustments, thereby correcting possible changes in the local geometry between the wild-type and mutant protein. In our protocol, we use non-crystallographic symmetry (NCS) restraints to maintain the backbone conformation in the secondary structure elements.

The refinement protocol was implemented in the Xplor-NIH structure determination package,[15,16] and validated as described below. The program incorporates the orientational information encoded in RDCs in the form of the SANI potential, E_{SANI}, given by:

$$E_{SANI} = k_{SANI}(D_{exp} - D_{calcd})^2 \quad (1)$$

where D_{exp} is the experimentally observed RDC, D_{calcd} the RDC back-calculated from the current structure, and k_{SANI} the associated force constant.[1] The back-calculated RDC depends on both the orientation and the magnitude of the alignment tensor. In the refinement protocol, the orientation of the alignment tensor is represented by a pseudomolecule, OXYZ, whose three bonds, O–X, O–Y and O–Z, are mutually orthogonal and which is free to rotate during the refinement. The magnitude of the alignment tensor is expressed in terms of the magnitude, D_a, and rhombicity, R_h, which are defined through the equation that describes the dependence of the dipolar coupling, D_{IS}, on the orientation of the I–S bond relative to the alignment tensor:

$$D_{IS} = D_a\left[\left(3cos^2\theta - 1\right) + \frac{3}{2}R_h\left(\sin^2\theta\ \cos 2\phi\right)\right] \quad (2)$$

where θ and ϕ are the polar angles of the I–S internuclear vector in the principal axis frame of the alignment tensor.

In our protocol, D_a and R_h are fixed and are calculated by global fitting of the experimental RDCs to the template structure by using singular value decomposition (PALES[17,18]). As it is presumed that the template structure is only an approximation to the true structure, the values of D_a and R_h thus computed are themselves approximations to the correct values. To overcome this limitation, the values of D_a and R_h were optimised in an iterative manner by repeating the protocol with D_a and R_h recalculated for each round of refinement by using the refined structure from the previous round as a template. This process was continued until D_a and R_h converged (within twelve iterations—see Table S1 and Figure S1 in the Supporting Information).

The refinement protocol itself was designed by using solely conjugate–gradient-based minimisation of the RDC energy. Such an approach is suitable in situations in which the template structure is a good approximation to the true structure. In addition to the geometric force field and SANI potential, the secondary structure of the protein was preserved in the first stages of the protocol by applying a NCS-potential term, in which the relative positions of the atoms within a user-defined secondary-structure element are restrained according to their positions in a reference molecule. While the NCS potential acts to maintain local conformation within the defined group, it places no restrictions on the global orientation.

The protocol consisted of several rounds of conjugate–gradient minimisation, with successive reductions in the weight of the NCS potential and redefinition of the NCS reference molecule at the start of each stage as the refined structure from the previous round (in the first round, the NCS reference molecule is the template structure). In the early stages, the local backbone conformation in the regions of defined secondary structure is strongly restrained, so that reduction in the SANI energy proceeds initially through optimisation of the orientation of the secondary-structure elements. The strongly reduced weight of the NCS potential in the latter stages permits adjustment of the backbone conformation, and allows for the possibility of significant local structural changes at both the mutation and distant sites.

Protocol validation

To validate the refinement protocol, an artificial RDC data set was simulated (PALES[18]) by using the template structure for refinement (the alignment tensor for simulation of these data was defined to have the magnitude and orientation of the tensor obtained by fitting the experimental data to the template structure). For this data set, the template structure is the "correct" structure, and the performance of the refinement protocol can then be assessed by examining its ability to return an artificially distorted 15.5K-2 molecule to the correct structure.

The distorted test structure was generated by allowing the 15.5K-2 template structure to relax in a constant-temperature molecular dynamics simulation.[19] The secondary structure was maintained through the application of strong NCS restraints. This simulation produced a test structure with backbone conformations in regions of secondary structure that are close to the starting structure, but in which the orientations of the secondary structure elements were significantly tilted relative to the template. The degree of distortion was quantified by calculating all pairs of interhelix angles for the five helices in 15.5K-2. The RMS deviation of the interhelix angles in the distorted structure relative to the correct (template) structure was 30.0°.

The data content of the simulated data used for refining the distorted test structure was equal to the available experimental data (84 N–H^N and 48 N–C′ RDCs). In addition, Gaussian noise was added to give errors in the simulated data that were approximately equal to the corresponding uncertainties in the experimental data. Fitting these data to the distorted test structure gave a quality factor (Q-factor) of 0.72. The Q-factor is a measure of the agreement of the experimental—simulated, in this case—RDCs, D_{exp}, with those back-calculated from a trial structure, D_{calc}, and is defined as:

$$Q = \sqrt{\frac{\sum_i \left(D^i_{exp} - D^i_{calcd} \right)^2}{\sum_i \left(D^i_{exp} \right)^2}} \quad (3)$$

The refinement protocol produced a structure with a Q-factor of 0.16 and an interhelix angular root-mean-square deviation (RMSD) relative to the correct structure of 4.9°; this demonstrates the efficiency of the protocol in correcting orientations of misaligned secondary-structure elements. A control refinement in which RDC data were omitted gave a structure with an interhelix RMSD of 20.8° (Q-factor of 0.70). Results of similar quality were obtained when testing the ability of the protocol to find the correct local conformation by starting from a locally distorted structure by using N–H^N, N–C′ and H^N–C′ RDCs.

Application to the 15.5K-2–U4 5′-SL (33nt) complex

Previous studies that were aimed at mapping the interaction surface between 15.5K and hPrp31 indicated that the 15.5K-2 mutant (E74R/D75K/K76M/N77T) exhibits a fourfold-reduced binding affinity to hPrp31 in pull-down assays.[13] However, in a recently solved crystal structure of the hPrp31–15.5K–U4 RNP,[14] these residues do not make direct structural contacts with hPrp31. Hence, any structural changes at the mutation site should not affect the binding interaction with hPrp31. However, it is possible that in addition to local perturbations, the mutations could also introduce global changes in the structure of the 15.5K protein, such that the interfacial surface with hPrp31 is perturbed, thereby disrupting the interaction between the two proteins. To confirm or discount this hypothesis, we applied our RDC refinement protocol to generate a structural model for the mutated 15.5K-2 protein from the crystal structure of the wild-type protein in the 15.5K–U4 5′-SL binary complex.[10]

The N–H^N correlation maps suggest there is strong similarity between the structures of the wild-type and mutant 15.5K–U4 5′-SL binary complexes, with the largest chemical shift differences located primarily at the mutation site (see Figure S2). This indicates that the crystal structure of 15.5K in the 15.5K–U4 5′-SL binary complex is an appropriate starting point for RDC refinement. Initial fitting of the N–H^N and N–C′ RDCs measured on the mutant binary complex of this template structure also demonstrates that the solution structure of the 15.5K-2 mutant does not differ greatly from the crystal structure of the wild-type protein and yields Q-factors of 0.43–0.48 for the three available data sets (one N–C′ and two N–H^N data sets; Figure 3A). A similar Q-factor of 0.4 is found when fitting the wild-type N–H^N data to the wild-type crystal structure (Figure 3B). The orientation tensors obtained by fitting the wild-type and mutant N–H^N data to the respective template structures are also similar (see Figure S1A).

One striking feature of these plots is the outlying point that corresponds to H68 in helix α3. The isolated nature of this outlier suggests that the underlying difference between the true structure and the template is a local conformational change rather than a re-orientation of helix α3. Furthermore, fitting the N–H^N RDCs measured on the wild-type complex to the wild-type crystal structure also gives a similar discrepancy for H68 (Figure 3B and Figure S3); this indicates a local difference between the solution and crystal structures of the wild-type 15.5K that is preserved in the 15.5K-2 mutant. The pattern of carbon chemical shifts in this region for wild-type 15.5K is also suggestive of a noncanonical backbone conformation, with a large negative secondary chemical shift difference $\Delta\delta_{C\alpha}-\Delta\delta_{C\beta}$ for H68 (top panel, Figure S4). The $\Delta\delta_{C\alpha}$ shift differences for 15.5K, which also exhibit an unusual alternating negative–positive pattern from I66 to L69, are exactly mirrored in the mutant 15.5K-2 (middle and bottom panels, Figure S4); this further demonstrates that while there are differences between the crystal and solution structures of 15.5K, these differences are reproduced very closely in 15.5K-2.

To confirm the expected similarity between the structures of the mutant and wild-type 15.5K proteins, a structural model of 15.5K-2 was generated by RDC refinement using the crystal

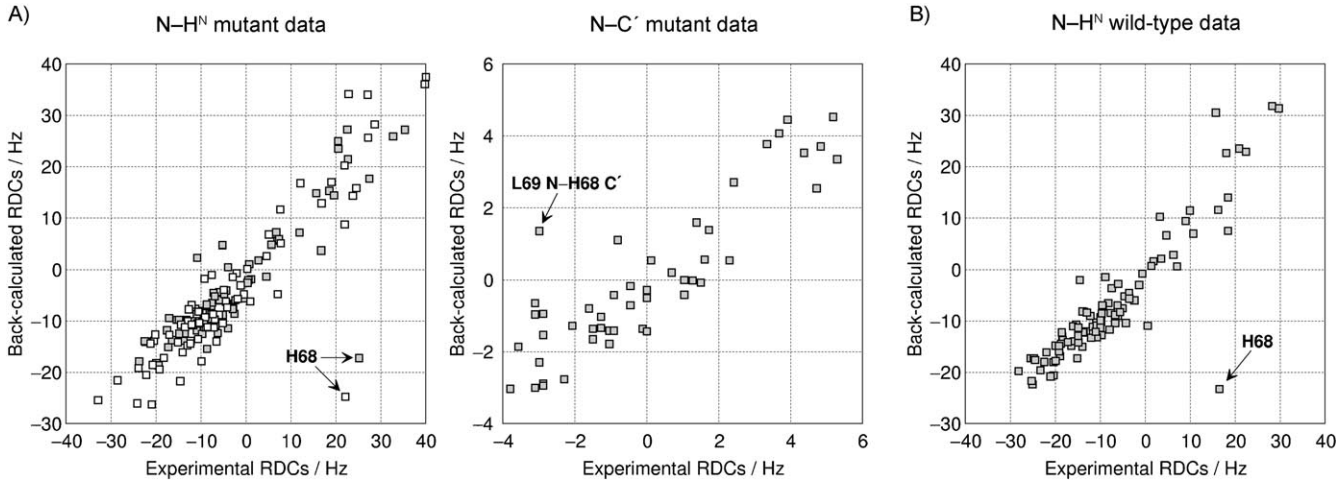

Figure 3. A) Correlations between experimental and back-calculated RDCs for the 15.5K-2 mutant. Left: the two N–H^N data sets, with Q-factors of 0.48 (grey boxes) and 0.43 (white boxes). Right: N–C′ data (Q-factor of 0.48). B) Correlation between experimental and back-calculated N–H^N RDCs for wild-type 15.5K (Q-factor of 0.40). The N–H^N data are plotted as a function of residue number in Figure S3 (Supporting Information).

structure of 15.5K as a template. After refinement, the Q-factor representing the agreement between the experimental RDCs and those that were back-calculated from the refined structure dropped to 0.1 and 0.08 for the two N–H^N RDC data sets and to 0.19 for the N–C′ data; this confirms the convergence of the refinement procedure. The higher Q-value observed for the N–C′ data reflects the lower weight that is given to the N–C′ restraints during the refinement. The corresponding RMSDs are approximately equal to the RMS uncertainties in the experimental data (1.3–1.6 Hz and 0.5 Hz for the N–H^N and N–C′ data sets, respectively).

The refined 15.5K-2 structure is shown in Figure 4. As predicted, it is very similar to the crystal structure of the wild-type protein, with a backbone RMSD of 0.64 Å between the two

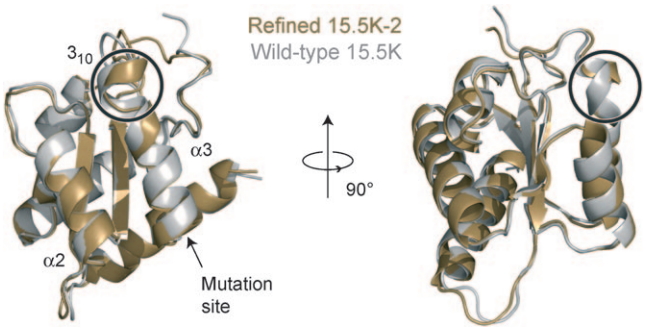

Figure 4. Structural model for 15.5K-2. Two orthogonal views showing overlay of the refined 15.5K-2 structure (gold) with the wild-type 15.5K crystal structure (grey). The principal difference between the two structures is the partial loss of the 3_{10} helix in the solution mutant structure (highlighted by the circles).

structures (residues 8–126). There are only minimal differences in the relative orientations of the secondary structure elements, as shown by the low RMSD value of all interhelix angles (3.1°). The changes in the backbone conformation are also relatively small, with RMSDs in the ϕ- and ψ-backbone dihedral angles of approximately 10°. As expected from the initial analysis on the unrefined template structure (Figure 3), there is a slight conformational change in the region of the 3_{10} helix located at the N terminus of α3, with loss of regular helical structure for I66 (as defined by STRIDE[20]). Closer inspection of the backbone dihedral angles for this stretch of residues reveals that there are appreciable differences (20–30°) in one or both of ϕ and ψ for the four residues 66–69 between the template and refined structures. The loss of the recognised 3_{10} helical structure for I66 can be attributed to both the change in the ϕ and ψ angles and the concomitant lengthening of the hydrogen bond from the amide group of I66 to the carbonyl group of L69 (from 1.9 Å to 3.0 Å). Interestingly, a large deviation is observed between the experimental and back-calculated values for the N–H^N RDC of H68 and for the L69 N–H68 C′ coupling, but not for the L67 N–I66 C′ coupling (Figure 3A). Overall, these observations indicate that the conformational change in the 3_{10} helical region of α3 is driven by a reorientation of the peptide plane between H68 and L69.

These local conformational changes are located in the interfacial region between 15.5K and hPrp31, and it may be proposed that they could contribute to the apparent reduced binding affinity between the 15.5K-2 mutant and hPrp31. However, the similarity between the N–H^N RDCs recorded on wild-type 15.5K and those from the mutant in the region under consideration (Figure 2B), and the matching $\Delta\delta_{C\alpha}$ shift differences (Figure S4) both indicate that the conformational differences between the generated structure of the 15.5K-2 mutant and the wild-type crystal structure are also present in the wild-type solution structure. As such, these conformational rearrangements cannot be invoked to explain the reduction in the strength of the interaction between 15.5K-2 and hPrp31.

Electrophoretic gel mobility shift assays

To pinpoint the residue responsible for the reported dysfunction of the 15.5K-2 mutant, four further mutants were created: E74R, D75K, K76M and E74R/D75K. Gel mobility shift assays were performed with these mutants to investigate their ability to form ternary complexes with hPrp31 and U4 5′-SL RNA. The four mutants and wild-type 15.5K were incubated with radiolabelled U4 5′-SL-33nt, titrated with increasing amounts of MBP–hPrp31, and the resulting mixtures of protein–RNA complexes were separated on a native gel (Figure 5). All of the 15.5K mutants were capable of forming ternary complexes with hPrp31. 15.5K-2 showed a slight reduction in binding affinity compared to the wild-type, as did the D75K and E74R/D75K mutants, with D75K showing the most noticeable change. There was no detectable reduction in the binding affinity for the K76M mutant (data not shown). Taken together, these results indicate that the decreased binding affinity of the 15.5K-2–U4 5′-SL complex for hPrp31 can be attributed to the D75K mutation.

Discussion

We have presented a simple RDC-based structure refinement protocol for assessing the structural integrity of mutated proteins, and demonstrated its application to the 15.5K-2 protein. Previous experiments had indicated that the 15.5K-2–U4 RNP exhibited reduced binding to hPrp31, but subsequent structural studies on the ternary complex showed that the mutated residues in 15.5K-2 are not directly involved in the protein–protein interface. To investigate the possibility that the attenuated binding affinity could be the result of global structural changes that are induced by the mutations, we generated a structural model for the 15.5K-2 mutant by RDC refinement of the 15.5K crystal structure. The structure thus determined was very similar to the template structure, with no identifiable structural distortions that would interfere with the binding to hPrp31. A slight change in the local backbone conformation was revealed in the 3_{10} helix at the N terminus of α3, but RDC and carbon chemical shift data collected on the wild-type protein indicated that a similar change is expected in the solution-state structure of 15.5K; this discounts the possibility that this change is responsible for the reduced binding affinity of 15.5K-2.

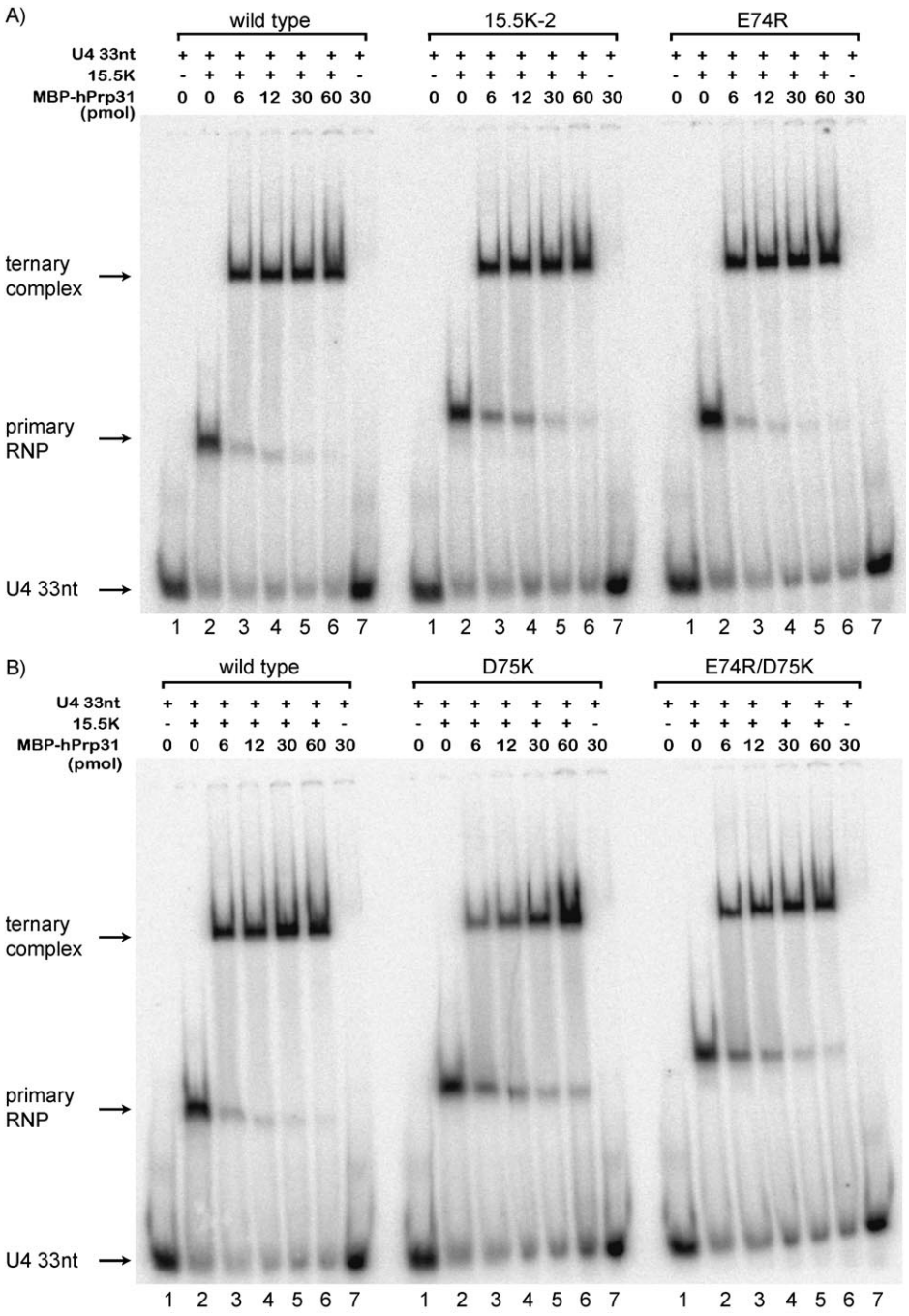

Figure 5. Autoradiograms of gel mobility shift assays performed on the wild-type and mutant 15.5K–U4 5′-SL primary RNPs titrated with MBP–hPrp31. The "+" and "−" signs indicate the presence or absence, respectively, of the corresponding component. For MBP–hPrp31, the increasing amounts are listed in pmol. In each case, lane 1 contains U4 5′-SL-33nt alone, lane 2 contains the primary RNP, lanes 3–6 contain the primary RNP titrated with increasing amounts of MBP–hPrp31, and lane 7 represents a negative control without 15.5K protein (U4 5′-SL and MBP–hPrp31).

Next, we performed gel mobility shift assays with 15.5K-2 and four related mutants. The results for 15.5K-2 confirmed the previous findings from pull-down assays, and indicated that this mutant exhibits weaker binding (ca. two- to fourfold reduction) to hPrp31 than wild-type 15.5K. However, the crystal structure of the complex indicates that the mutated residues are not directly contacting the hPrp31. An explanation for these observations was revealed from the gel-shift assays with the four related mutants. The mutants D75K and E74R/D75K showed similar reductions in binding affinity to 15.5K-2, whereas E74R and K76M both exhibited near wild-type affinity for hPrp31. In fact, the single mutation D75K is sufficient to account for all the corresponding binding affinity reduction shown by 15.5K-2. Inspection of the crystal structure of hPrp31$^{78–333}$–15.5K–U4 5′-SL[14] reveals that the negatively charged carboxylate group of D75 in 15.5K is only ~4–6 Å away from the positively charged amino group of K243 in hPrp31 (Figure 6). This attractive electrostatic interaction is changed into repulsion in the mutant D75K and lowers the binding affinity of the two proteins, despite the fact that D75 does not belong to the protein–protein interface. This indicates that binding studies can be sensitive to non-direct interactions, and suggests that a cautious approach is advisable when translating mutagenesis data into direct structural contacts.

In conclusion, we have presented a simple protocol that is based on easily accessible residual dipolar coupling data for rapid verification of the structural integrity of mutated proteins; this allows increased confidence in the interpretation of results from biochemical experiments performed by using the mutant proteins. We have applied the protocol to the 15.5K-2 mutant protein, and demonstrated that its structure is not significantly altered from wild-type 15.5K. This result in combination with the recently solved crystal structure of the hPrp31–15.5K–U4 complex prompted us to re-examine a previous pull-down study in which the 15.5K-2–U4 RNP was shown to exhibit reduced binding to hPrp31. Gel-shift assays with 15.5K-2 and related single and double mutants enabled us to pinpoint the single amino acid mutation that is principally responsible for the weaker binding,

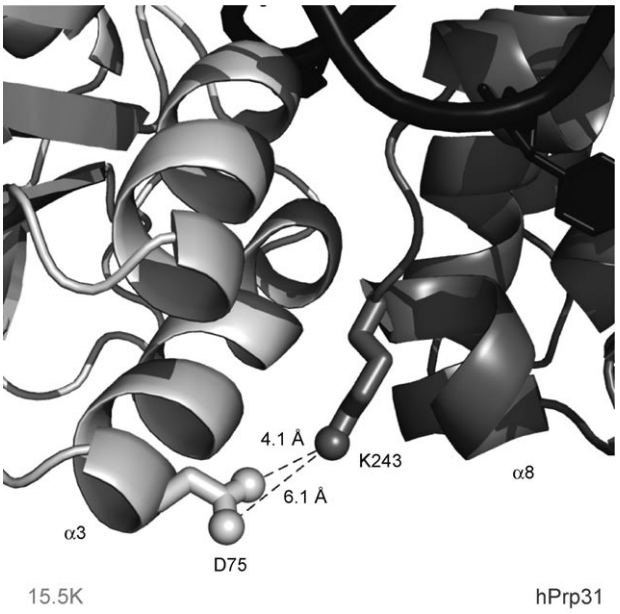

Figure 6. Detail of 15.5K–U4 5′-SL–hPrp31 crystal structure (15.5K in light grey, U4 5′-SL in black and hPrp31 in dark grey) highlighting the electrostatic interaction between the D75 (15.5K) carboxylate group and the K243 (hPrp31) amino group.

and led to the conclusion that the reduction in binding affinity is due to a non-direct interaction rather than a direct interfacial contact.

Experimental Section

Site-directed mutagenesis: Single and double amino acid changes were generated in 15.5K by using the QuikChange site-directed mutagenesis kit (Qiagen). The plasmid DNA was obtained by using the Miniprep kit (Qiagen).

NMR spectroscopy experiments on 15.5K–U4 5′-SL–33nt and 15.5K-2–U4 5′-SL–33nt: 15.5K and its mutants were expressed as GST fusion proteins and were purified as described for wild-type 15.5K in Liu et al.[14] 15.5K and 15.5K-2 were uniformly labelled with ^{15}N and ^{13}C by expression in minimal medium containing $^{15}NH_4Cl$ and ^{13}C-glucose. In addition, 15.5K was 70% deuterated. The unlabelled U4 5′-SL 33nt oligonucleotide was purchased from IBA GmbH (Goettingen, Germany), dialysed into the NMR sample buffer (10 mM Tris–HCl, pH 7.6, 120 mM NaCl, 2 mM DTT) and annealed at 65 °C for 90 s before combination with the proteins. The backbone assignments of the 15.5K–U4 complex were obtained from TROSY versions of HNCA, HN(CO)CA, HNCACB, and HN(CO)CACB spectra,[21,22] which were recorded on a Bruker 900 MHz spectrometer. The backbone assignments of the 15.5K-2–U4 complex were obtained from non-TROSY versions of HNCA and HN(CO)CA spectra, which were recorded on a Bruker 600 MHz spectrometer. All experiments were recorded on ~0.3 mM samples at 308 K.

Residual dipolar couplings (RDCs) were measured for both wild-type and 15.5K-2 proteins in complex with the U4 5′-SL RNA (33nt). Anisotropic samples were prepared by addition of filamentous phage Pf1 (ASLA Biotech, Riga, Latvia) to a concentration of ~10 mg mL^{-1}. N–H^N couplings were measured by using a modified IPAP-HSQC experiment.[23] N–C′ couplings were determined by using a double-interleaved E.COSY-type ^{15}N HSQC experiment,[24] for which the two components of the reduced multiplets are offset by the N–C′ and C′–H^N couplings in the ^{15}N and ^{1}H dimensions, respectively. RDCs were determined for two samples of the 15.5K-2–U4 5′-SL complex; for the first sample, only N–H^N RDCs were recorded, whereas for the second sample, both N–H^N and N–C′ RDCs were measured. For the wild-type complex, a single set of N–H^N RDCs were acquired. All experiments were carried out at 308 K and 600 MHz. The extracted data were trimmed to exclude values from flexible residues (residues 1–7, 127 and 128) and residues with overlapped peaks in the ^{15}N HSQC spectrum. In addition, some N–C′ RDCs were excluded due to insufficient signal-to-noise ratios of the corresponding resonances. The final data set for refinement consisted of 80 N–H^N couplings from the first sample, and 73 N–H^N and 48 N–C′ couplings from the second sample (in total, 84 residues are represented by at least one N–H^N RDC). The root-mean-square (RMS) uncertainties in the N–H^N and N–C′ RDCs were calculated as 1.3–1.6 Hz and 0.5 Hz, respectively.

Structure refinement of the 15.5K-2–U4 5′-SL-33nt complex with RDCs: The template structure for refinement was derived from the 15.5K–U4 5′-SL crystal structure (PDB ID: 1E7K[10]). The side chains of residues 74–78 were replaced according to the mutations in 15.5K-2 (Chimera[25]), hydrogen atoms were added, and the resulting structure was regularised by using a short minimisation.

The refinement protocol was written in the structure calculation program Xplor-NIH.[15,16] The SANI potential[1] for the RDC restraints was implemented as a harmonic potential, with force constants, k_{SANI}, of 0.5 kcal mol^{-1} Hz^{-2} and 0.05 kcal mol^{-1} Hz^{-2} for N–H^N and N–C′ restraints, respectively. N–C′ couplings were premultiplied by a scaling factor, given by $(\gamma_H/\gamma_C)\cdot(r_{N-C'}/r_{N-H})^3 = 8.3$. The force constants were empirically adjusted such that the RMSD of the back-calculated RDCs after refinement was approximately equal to the estimated uncertainty in the experimental data. Two alignment tensors corresponding to the two samples were defined in the refinement protocol, with distinct magnitudes, D_a and R_h, and orientations. Six rounds of minimisation were performed, with the weight of the NCS potential reduced by a factor of ten at each stage, from an initial value of 100 000 kcal mol^{-1} Å^{-2} to 1 kcal mol^{-1} Å^{-2} in the final round. Each round consisted of 5000 steps of conjugate–gradient minimisation. NCS restraints were applied to the nine secondary-structure elements of the 15.5K–U4 5′-SL structure as determined by using the program STRIDE.[20] The NCS reference molecule was re-defined at each stage as the refined structure from the previous round.

The distorted test structure used for protocol validation was generated by allowing the 15.5K-2 template structure to relax in 10 ps of constant-temperature (2000 K) molecular dynamics. The secondary structure was maintained through application of strong NCS restraints (force constant of 10 000 kcal mol^{-1} Å^{-2}).

Electrophoretic gel mobility shift assays: Radioactively 5′-labelled RNA oligonucleotides (30–50 fmol) were used in the band-shift assays. In vitro reconstitution of protein–RNA complexes was carried out by adding 100-fold molar excesses of 15.5K (wild-type and mutants) and hPrp31 in the quantities indicated in Figure 5 to U4 5′-SL–33nt. The protein concentration was determined by the Bradford assay. The buffer contained *E. coli* tRNA (10 μg), Triton X-100 (0.2%) and DTT (1 mM) in an end volume of 20 μL. The samples were incubated at 4 °C for 30–60 min and were loaded onto native polyacrylamide gels (6–9%). Gel electrophoresis was performed at 4 °C in tris-borate-EDTA (TBE) buffer (0.5×) at 9 W for 1.5 h.

Keywords: mutagenesis · NMR spectroscopy · protein structures · residual dipolar couplings · spliceosome

[1] N. Tjandra, J. G. Omichinski, A. M. Gronenborn, G. M. Clore, A. Bax, *Nat. Struct. Biol.* **1997**, *4*, 732.
[2] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9279.
[3] G. A. Mueller, W. Y. Choy, D. W. Yang, J. D. Forman-Kay, R. A. Venters, L. E. Kay, *J. Mol. Biol.* **2000**, *300*, 197.
[4] N. R. Skrynnikov, N. K. Goto, D. W. Yang, W. Y. Choy, J. R. Tolman, G. A. Mueller, L. E. Kay, *J. Mol. Biol.* **2000**, *295*, 1265.
[5] J.-C. Hus, D. Marion, M. Blackledge, *J. Am. Chem. Soc.* **2001**, *123*, 1541.
[6] J.-C. Hus, D. Marion, M. Blackledge, *J. Mol. Biol.* **2000**, *298*, 927.
[7] S. Nottrott, K. Hartmuth, P. Fabrizio, H. Urlaub, I. Vidovic, R. Ficner, R. Luhrmann, *EMBO J.* **1999**, *18*, 6119.
[8] P. Bringmann, B. Appel, J. Rinke, R. Reuter, H. Theissen, R. Luhrmann, *EMBO J.* **1984**, *3*, 1357.
[9] C. Hashimoto, J. A. Steitz, *Nucleic Acids Res.* **1984**, *12*, 3283.
[10] I. Vidovic, S. Nottrott, K. Hartmuth, R. Luhrmann, R. Ficner, *Mol. Cell* **2000**, *6*, 1331.
[11] S. Nottrott, H. Urlaub, R. Luhrmann, *EMBO J.* **2002**, *21*, 5527.
[12] C. Schneider, C. L. Will, O. V. Makarova, E. M. Makarov, R. Luhrmann, *Mol. Cell. Biol.* **2002**, *22*, 3219.
[13] A. Schultz, S. Nottrott, N. J. Watkins, R. Luhrmann, *Mol. Cell. Biol.* **2006**, *26*, 5146.
[14] S. Liu, P. Li, O. Dybkov, S. Nottrott, K. Hartmuth, R. Luhrmann, T. Carlomagno, M. C. Wahl, *Science* **2007**, *316*, 115.
[15] C. D. Schwieters, J. J. Kuszewski, G. M. Clore, *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 47.
[16] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, *J. Magn. Reson.* **2003**, *160*, 65.
[17] J. A. Losonczi, M. Andrec, M. W. F. Fischer, J. H. Prestegard, *J. Magn. Reson.* **1999**, *138*, 334.
[18] M. Zweckstetter, A. Bax, *J. Am. Chem. Soc.* **2000**, *122*, 3791.
[19] R. D. Seidel, J. C. Amor, R. A. Kahn, J. H. Prestegard, *J. Biol. Chem.* **2004**, *279*, 48307.
[20] D. Frishman, P. Argos, *Proteins Struct. Funct. Genet.* **1995**, *23*, 566.
[21] S. Grzesiek, A. Bax, *J. Magn. Reson.* **1992**, *96*, 432.
[22] S. Grzesiek, A. Bax, *J. Biomol. NMR* **1993**, *3*, 185.
[23] M. Ottiger, F. Delaglio, A. Bax, *J. Magn. Reson.* **1998**, *131*, 373.
[24] K. Y. Ding, A. M. Gronenborn, *J. Magn. Reson.* **2002**, *158*, 173.
[25] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605.

Received: November 27, 2008
Published online on March 23, 2009

