ELSEVIER

NMR analysis of protein interactions

Alexandre MJJ Bonvin, Rolf Boelens and Robert Kaptein

Recent technological advances in NMR spectroscopy have alleviated the size limitations for the determination of biomolecular structures in solution. At the same time, novel NMR parameters such as residual dipolar couplings are providing greater accuracy. As this review shows, the structures of protein–protein and protein–nucleic acid complexes up to 50 kDa can now be accurately determined. Although *de novo* structure determination still requires considerable effort, information on interaction surfaces from chemical shift perturbations is much easier to obtain. Advances in modelling and data-driven docking procedures allow this information to be used for determining approximate structures of biomolecular complexes. As a result, a wealth of information has become available on the way in which proteins interact with other biomolecules. Of particular interest is the fact that these NMR-based methods can be applied to weak and transient protein–protein complexes that are difficult to study by other structural methods.

Addresses
Bijvoet Center for Biomolecular Research, Utrecht University, NL-3584 CH Utrecht, The Netherlands

Corresponding author: Kaptein, Robert (r.kaptein@chem.uu.nl)

Current Opinion in Chemical Biology 2005, **9**:501–508

This review comes from a themed issue on
Analytical techniques
Edited by Chris D Geddes and Ramachandram Badugu

Available online 24th August 2005

1367-5931/$ – see front matter

DOI 10.1016/j.cbpa.2005.08.011

Introduction

In recent years, we have seen great improvements in NMR spectroscopy as a tool for the study of biomolecular interactions. The sensitivity of the technique has been enhanced by the advent of high-field spectrometers (900 MHz) and cryogenic probes. Novel methodologies (transverse relaxation-optimized spectroscopy (TROSY) [1]; residual dipolar couplings [2,3]) have enabled the structural analysis of larger molecules and complexes. Also, better use can now be made of spectral information such as chemical shift perturbations (CSPs) through improved modelling and docking procedures. Consequently, it is possible to characterize larger and biologically more relevant biomolecular complexes with higher accuracy. In this review, we discuss recent advances in NMR structural studies of the interactions of proteins with other proteins, RNA and DNA. Interactions with small molecules and peptides is not included, as these have been adequately covered in other reviews [4,5].

Biomolecular complexes reported since 2003 range from full *de novo* determined structures calculated from NMR-derived restraints (such as nuclear Overhauser effects (NOEs) and residual dipolar couplings (RDCs)) to models obtained by docking individual components based on knowledge of the interface from chemical shift perturbations. Although in the latter case the coordinate accuracy is necessarily lower, the information on the interaction gained is often quite important in guiding subsequent research.

Methodological developments

Next to the 'classical' approach based on the use of intermolecular NOEs, in combination with RDCs when available [6], the characterization of protein interactions has greatly benefited from the incorporation of interface mapping information in the computational modelling of complexes. NMR is particularly powerful in mapping interfaces, allowing the study of weak and transient complexes that can be very difficult to study by other experimental techniques.

The use of CSP data obtained from NMR titration experiments as structural restraints for solving the structure of protein complexes was first demonstrated by McCoy and Wyss [7]: RDC data were first introduced to orient the complex and the solutions were then optimized by a grid search and back calculation of chemical shift perturbations with the SHIFTS software [8]. In the HADDOCK approach [9••], CSP data are translated into ambiguous interaction restraints to drive the docking process. The CSP interaction restraints can also be combined with RDC data allowing a better definition of the relative orientation of the components [10•,11]. Note that CSP data have also been combined with docking methods to filter *a posteriori* the docking solutions [12], also in combination with RDC data [13], or to provide anchor points to start the search [14]. New chemical shift-based methods relying on amino-acid specific labeling have also been developed to map interfaces [15,16]. These methods no longer require resonance assignment provided the 3D structure of the unbound protein is known.

Several other NMR techniques can provide information on interfaces. In cross-saturation or saturation transfer (SAT) experiments [17], the observed protein is perdeuterated and ^{15}N-labeled with its amide deuterons

exchanged back to protons, whereas the other 'donating' partner protein is unlabeled. Saturation of the unlabeled protein leads by cross-relaxation mechanisms to signal attenuation (typically monitored by ^{15}N-HSQC spectra) of those residues in the labeled protein that are in close proximity. Deuteration here is a requisite. Cross-saturation experiments are believed to give a more reliable picture of the interface than CSP data, which can suffer from 'false positives' due to conformational changes. Experimental intensity changes from SAT have been used as restraints in a docking procedure in combination with RDC data [18]. For paramagnetic systems, another useful NMR parameter is the pseudocontact shift, a long range effect that results from electron–nuclei dipolar interactions. The use of paramagnetic tags attached to a protein can also induce this phenomenon. Pseudocontact shifts have been used to study transient redox complexes [19] and recently to break the symmetry in symmetrical complexes [20]. It is also possible to use paramagnetic ions as probes, using the line broadening effect for the residues they contact. In a complex, the interface residues will be protected from such effects allowing identification of the interface [21,22]. Finally, orientational information similar to that provided by RDCs can also be extracted from relaxation experiments in the case of rotational diffusion anisotropy [23].

Protein–protein interactions

NMR studies of protein–protein complexes have varied from full structure determination to NMR-filtered docking and modeling using interface information. We limit our discussion mainly to complexes for which the atomic coordinates have been deposited into the Protein Data Bank (PDB; http://www.rcsb.org) (see Table 1).

Since 2003, the structures of 14 protein–protein complexes have been solved using intermolecular NOEs detected from ^{13}C/^{15}N-filtered NOESY experiments in combination with RDCs when available. From these, 12 were determined completely *de novo* by NMR [24•,

Table 1

Overview of protein–protein, protein–RNA and protein–DNA complexes solved by NMR and deposited in the Protein Data Bank (http://www.rcsb.org) since 2003.

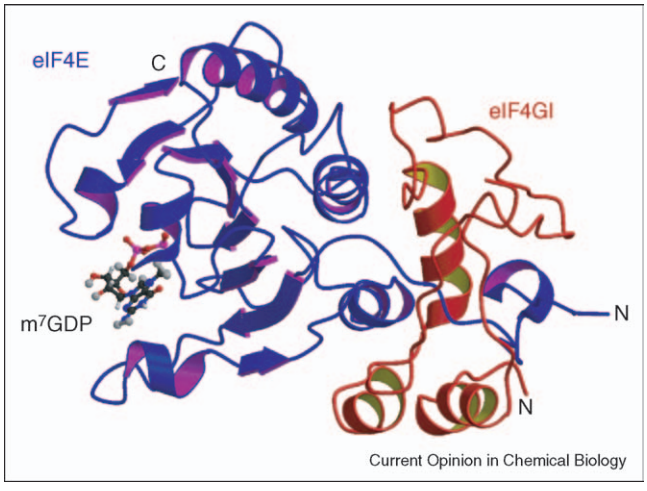
Complex name	Size[a] (aa/bp/nt)	PDB ID	Reference
Protein–protein complexes			
NOE-based complexes			
Enzyme IIA^{Mannose}–HPr	2 × (129 + 85)	1VRC	[36•]
Nck-2 SH3 domain–PINCH-1 LIM4 domain	71 + 66	1U5S	[24•]
CD3t–CD3δ ectodomains	178	1XMW	[25]
Ubiquitin–Npl4 zinc finger	76 + 31	1Q5W	[26]
HP1 chromo domain–PXVXL motif of CAF-1	2 × 75 + 30	1S4Z	[27]
ThKaiA108C–KaiC peptide	2 × 107 + 34	1SUY	[28]
CPB TAZ1–CITED2	50 + 100	1R8U	[30]
EIF4Ecap–eIF4G	213 + 100	1RF8	[31••]
Enzyme IIA^{Glc}–Enzyme IICB^{Glc}	168 + 90	1O2F	[32••]
HHR23A Ubl–S5a-UIM-2	45 + 78	1P9D	[34]
CUE–Ubiquitin	49 + 76	1OTR	[33]
CITED2 TAD–p300 CH1	52 + 101	1P4Q	[35]
MMP-3–N-TIMP-1	168 + 126	1OO9	[21]
Chemical shift mapping-based complexes			
RPA32 C domain–SV40 T antigen domain	105 + 132	1Z1D	[40•]
Plastocyanin–cytochrome *f*	105 + 254	1TU2	[38•]
UbcH5B–CNOT4	147 + 52	1UR6	[39•]
Atx1–Ccc2 copper transporting ATPases	73 + 72	1UV1/2	[37]
Protein–RNA complexes			
Tandem zinc finger domain of TIS11d–AU-rich ssRNA	70 + 9 nt	1RGO	[49••]
Nucleocapsid protein–Mlv encapsidation signal RNA	56 + 101nt	1U6P	[50]
Nucleocapsid protein–AACAGU	56 + 6 nt	1WWD1WW	[51]
Nucleocapsid protein–UUUUGCU	56 + 7 nt	E1WWF	
Nucleocapsid protein–CCUCCGU	56 + 7 nt	1WWG	
Nucleocapsid protein–UAUCUG	56 + 6 nt		
RNase III (Rnt1P)–dsRNA	90 + 32 nt	1T4L	[53•]
Protein–DNA complexes			
Oct1–Sox2–Hoxb1 DNA (ternary complex)	167 + 88 + 19 bp	1O4X	[56••]
MarA–α-CTD RNAP–DNA (ternary complex)	132 + 84 + 20 bp	1XS9	[41•]
Cdc13 CTD–ssDNA	199 + 11 nt	1S40	[60•]
Dimeric lac headpiece–non-specific DNA	2 × 62 + 18 bp	1OSL	[57••]

[a] The size is expressed in amino acid residues (aa) unless otherwise specified with bp (base pairs) or nt (nucleotides).

25–30,31•,32••,33–35] while the remaining two [21,36•] were solved starting from the known 3D structures of the unbound components or from parts of other complexes. Only five exceed 200 amino acid residues, indicating that the structure determination of large protein complexes by NMR remains challenging. Three of these were obtained by including RDCs to properly define the relative orientation of the molecules [21,32••,36•]. In particular, the 258 amino acid complex between the signal-transducing Enzyme $IIA^{Glucose}$ and the cytoplasmic domain of the glucose transporter Enzyme IICB [32••] provides a good illustration of the power of the combination of NOE and RDC for *de novo* NMR structure determination. One of the largest (313 amino acids) complexes solved by NMR was the one between eIF4G and eIF4E [31••] (Figure 1). The NOE information was supplemented in this case by restraints derived from paramagnetic broadening effects. The recently published 48 kDa Enzyme $IIA^{Mannose}$–HPr complex [36•] is the largest NMR complex solved on the basis of intermolecular NOEs and RDCs. An ultraweak complex with a K_d in the millimolar range was solved using a combination of intermolecular NOEs and RDCs, illustrating the power of NMR when it comes to studying weak interactions [24•].

Next to the NOE-based structures, an increasing number of complexes are being reported that have been solved by (data-driven) docking using mainly CSP data, often in combination with additional information to either generate or validate the resulting structures. Coordinates for a few of these (see Table 1) have been deposited into the PDB. These structures are the direct result of recent methodological development to make use of CSP data [9••,10•] and include weak and transient complexes [37,38•]. Most structural models obtained in this way were validated using independent data. They often provide a starting point for site-directed mutagenesis in combination with binding assays [39•–41•]. The 95 kDa complex of the acyltransferase protein with the acyl carrier protein was modeled based on NMR (CSP + RDCs) and mutagenesis data [42•]. That chemical shifts perturbations alone can lead to reliable models was shown for the RPA32–SV40 complex [40•]: the CSP-based model was validated using RDC data. Finally, the structure of plastocyanin and cytochrome *f* complex [38•] illustrates the great promise of paramagnetic NMR in characterizing transient complexes by a combination of CSP data with pseudo-contact shifts providing long-range distance and angular information.

Figure 1



NMR solution structure of the complex of the eukaryotic initiation factors 4E (eIF4E) (in blue) and 4G (eIF4G) (in red) (PDB entry 1RF8) [31••]. The complex is formed by a coupled folding transition of eIF4G and the N-terminus of eIF4E.

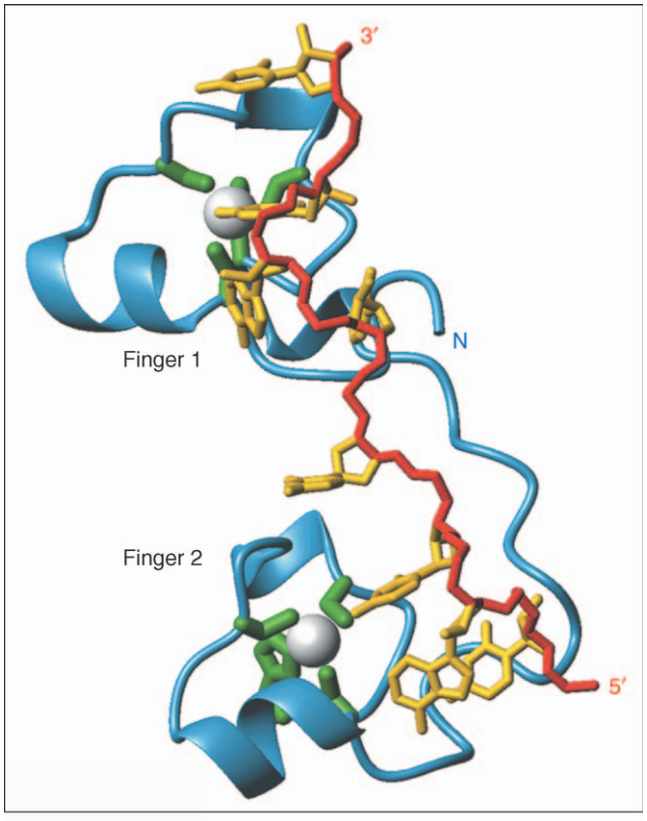
Protein–RNA interactions

One of the most abundant RNA binding motifs with an occurrence of 1.5–2% in the human genome is the RNA recognition motif (RRM) recently reviewed in [43,44]. RRMs are often found as tandem repeats within a protein together with other domains. Specificity mainly comes from direct interactions between the RNA bases and the protein side chains and main chains. Recently, the complex of the two N-terminal domains RRM1 and RRM2 of nucleolin with the nucleolin responsive element (b2NRE) was solved by NMR [45•]. The higher stability of an *in-vitro* selected pre-rRNA [46] is accounted for by additional contacts with the double-stranded RNA stem. Not all RRM motifs, however, are involved in RNA binding; interactions with protein have been observed as well [44].

Another abundant single-stranded RNA binding motif is the KH domain, reviewed in [43]. A solution structure has been determined for the KH domain of splicing factor SF1 in complex with an 11-mer RNA containing a 7 nucleotide branchpoint sequence [47].

The most abundant eukaryotic nucleic acid binding motif (with 3% of the genes in the human genome) is the CCHH-type zinc-finger (ZF) motif. Although this motif is mainly associated with DNA regulatory proteins such as transcription factors, it has been found to bind RNA as well. The X-ray structure of the complex of a triple ZF domain of TFIIIA with a truncated 5S RNA shows how the α-helices of the different ZFs interact with the phosphate backbone and the 5S RNA grooves [48]. The solution structure of the tandem zinc finger domain of the protein TIS11d in complex with an AU-rich single-stranded mRNA element is very different [49••] (Figure 2). It shows two compact CCCH-type zinc-fingers structures in a novel fold with the mRNA along the surface of the two ZF motifs. The NMR structures of several CCHC knuckles of the nucleocapsid (NC) domains of retroviral Gap polyproteins in complex with RNA packaging signals have been determined over the

Figure 2



NMR solution structure of the tandem zinc-finger domain of the protein TIS11d bound to the class II AU-rich element (PDB entry 1RGO) [49••]. Only ordered regions of TIS11d (residues 153–217) (blue) and RNA bases U2–U9 (red and yellow) are shown. The location and coordination of the zinc atom in each finger is indicated. Reprinted with the permission from [49••]. Copyright 2004, *Nature Structural Biology*, http://www.nature.com.

past few years. The NC protein has a compact and globular structure that interacts with unpaired 4nt RNA sequences, as noted in the various complexes of Moloney murine leukaemia virus NC [50,51] (see Table 1). The high affinity for such sequences causes significant RNA secondary structure rearrangements that are implicated in viral packaging and encapsidation.

Double-stranded-RNA binding domains (dsRBDs) are found in a large number of RNA regulatory proteins, recently reviewed in [52]. The NMR structure of the dsRBD of Rnt1p RNase III in complex with a dsRNA fragment showed that the AGNN tetraloop is recognized in a structure-specific manner [53•].

Proteins involved in RNA interference show a conserved domain involved in recognition of small interfering RNA (siRNA) called the PAZ domain. The solution structure of the Argonaute2 PAZ domain bound to short single-stranded RNA and single-stranded DNA oligonucleotides shows how the two-nucleotide 3′ ssRNA overhang of an siRNA is recognized [54••].

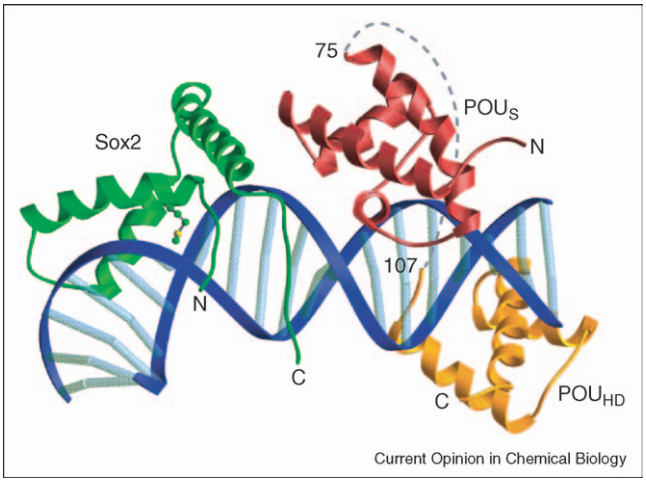
A common observation in protein–RNA complex formation is a significant adaptation of protein and RNA structure to make an optimal complex. Examples are found in the structure of the complex of the Jembrana virus Tat protein and HIV TAR RNA. The unstructured peptide folds into a stable β-hairpin in the complex and stabilizes an RNA conformation with a bulged-out uridine [55].

Protein–DNA interactions

A structure has been reported for the ternary complex of the POU domain of Oct1, Sox2 and the Hoxb1-DNA regulatory element [56••]. A modular approach was used to build the structure based on the binary DNA complexes of POU and Sox2 (or rather the homologous SRY). The structure (shown in Figure 3) was refined by extensive use of RDCs. Both transcription factors co-interact while bound to adjacent sites at the DNA. Comparison with various other regulatory sites sheds light on the mechanism of cell-specific transcription regulation [56••].

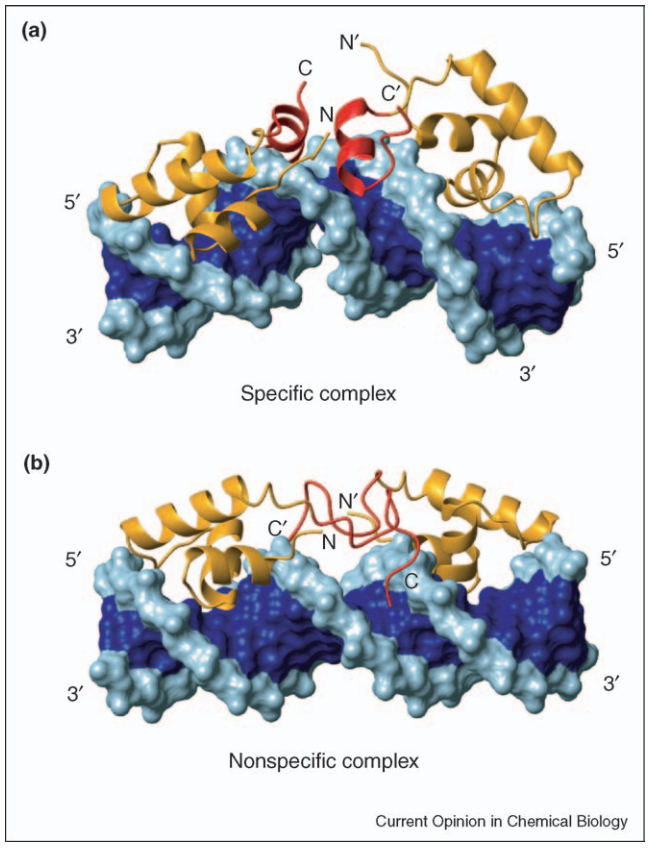
Gene regulatory proteins that bind DNA in a sequence-specific manner usually also interact non-specifically. This non-specific binding allows a much faster search for the specific target site by one-dimensional diffusion along the DNA. For the *Escherichia coli* lac repressor, the NMR structure of a complex between a covalent lac headpiece dimer and non-operator DNA has given detailed information about the non-specific binding mode [57••]. The structure revealed many differences with that of the complex with lac operator (Figure 4). It accounts for

Figure 3



NMR solution structure of the ternary Oct1–Sox2–Hoxb1 DNA complex (PDB entry 1O4X), solved using a combination of intermolecular NOEs and RDCs [56••].

Figure 4



Comparison of the NMR solution structures of the **(a)** specific and **(b)** non-specific dimeric lac headpiece–DNA complexes. The specific complex (PBD entry 1L1M) was solved using the 22 bp natural O1 operator sequence, whereas for the non-specific complex (PDB entry 1OSL) [57••] an 18 bp non-operator DNA sequence was used. In the specific complex, the hinge helices are formed leading to DNA bending whereas these helices are unstructured in the non-specific complex. A large difference in dynamics was noted: the protein in the non-specific complex displayed extensive mobility in the microsecond to millisecond range, whereas all motions are frozen in the specific complex.

long-standing thermodynamic data and, for instance, the fact that the affinity towards random DNA is much more salt dependent than that towards lac operators.

Another approach to the study of non-specific DNA binding makes use of paramagnetic relaxation enhancement [58,59]. A paramagnetic tag (Mn^{2+} chelated by EDTA) is attached to deoxythymidine in a DNA duplex. The relaxation enhancement of protons on the bound protein yields long-range distance information. Combined with RDCs and limited NOE information, this data improves the accuracy with which the structures of protein–DNA complexes can be determined. Applications were reported for the HMG box proteins SRY and HMGB-1 [58,59].

The telomere binding protein Cdc13 of budding yeast binds specifically to telomeric single-stranded DNA. A complex was solved for the DNA binding of Cdc13 with cognate single-stranded DNA [60•]. The structure provides a rationale for the high affinity and high specificity towards GT sequences.

In several reports, protein–DNA interactions were studied by CSP mapping using protein structures obtained from NMR or X-ray crystallography [61–64]. The ternary complex consisting of the transcriptional activator MarA, its DNA target site and the C-terminal domain of the α-subunit of RNA polymerase (α-CTD RNAP) was studied by CSP mapping from a titration of the MarA–DNA complex with α-CTD RNAP [41•]. The structure of the ternary complex was obtained by data-driven computational docking.

The human replication protein A (RPA) has received considerable attention. The protein plays a crucial role in DNA replication, recombination and repair by interacting with single-stranded DNA. It is a heterotrimer consisting of three subunits of 70, 32 and 14 kDa. In the context of DNA repair, RPA together with the exclusion repair protein XPA forms a complex with the damaged DNA site. The interaction of RPA with a DNA decamer, containing a cyclobutane thymine dimer lesion, was studied using CSP mapping [65]. Addition of XPA displaces RPA to the undamaged strand, whereas XPA prefers binding to the lesion. It was also shown that single-stranded DNA and XPA compete for the same binding site on RPA [66]. These studies shed light on the role of RPA and XPA in recognition and verification of DNA damage. Interaction surfaces of RPA were determined for a variety of protein–protein interactions. These include the interaction with the large T antigen (Tag) of SV40 [40•], with Rad 51 and its role in homologous recombination [67], and with the P53 transactivation domain [68].

Conclusions

It is clear that NMR has become a powerful and versatile method for the study of biomolecular interactions. Several spectacular protein–protein and protein–nucleic acid complexes in the molecular weight rang 40–50 kDa have been solved, and larger assemblies are within reach. The identification of interaction surfaces by relatively simple NMR experiments such as ^{15}N-HSQC is becoming very popular. The resulting chemical shift perturbations (or any information related to the interaction) can now be used as restraints in data-driven docking algorithms to produce approximate structures for the complexes. Thus, NMR appears to be ready for the structural characterization of biomolecular complexes, an area that is often viewed as the next phase in structural proteomics.

Acknowledgements

We thank Dr Peter Wright (Scripps Institute) and Dr Marius Clore (NIH) for providing the material for Figures 2 and 3, respectively.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Pervushin K, Riek R, Wider G, Wüthrich K: **Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution**. *Proc Natl Acad Sci USA* 1997, **94**:12366-12371.

2. Prestegard JH, Kishore AI: **Partial alignment of biomolecules: an aid to NMR characterization**. *Curr Opin Chem Biol* 2001, **5**:584-590.

3. Blackledge M: **Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings**. *Prog Nucl Magn Reson Spectrosc* 2005, **46**:23-61.

4. Diercks T, Coles M, Kessler H: **Applications of NMR in drug discovery**. *Curr Opin Chem Biol* 2001, **5**:285-291.

5. Villar HO, Yan J, Hansen MR: **Using NMR for ligand discovery and optimization**. *Curr Opin Chem Biol* 2004, **8**:387-391.

6. Clore GM: **Accurate and rapid docking of protein-protein complexes on the basis of intermolecular nuclear Overhauser enhancement data and dipolar couplings by rigid body minimization**. *Proc Natl Acad Sci USA* 2000, **97**:9021-9025.

7. McCoy MA, Wyss DF: **Structures of protein-protein complexes are docked using only NMR restraints from residual dipolar coupling and chemical shift perturbations**. *J Am Chem Soc* 2002, **124**:2104-2105.

8. Xu XP, Case DA: **Automated prediction of N-15, C-13(alpha), C-13(beta) and C-13′ chemical shifts in proteins using a density functional database**. *J Biomol NMR* 2001, **21**:321-333.

9. •• Dominguez C, Boelens R, Bonvin AMJJ: **HADDOCK: a protein-protein docking approach based on biochemical or biophysical information**. *J Am Chem Soc* 2003, **125**:1731-1737.

The use of CSP and mutagenesis data to model biomolecular interactions is demonstrated for three protein–protein complexes. Interface mapping information is introduced as ambiguous interaction restraints to drive the docking process using a combination of rigid body minimization and semi-flexible simulated annealing refinement.

10. • Clore GM, Schwieters CD: **Docking of protein-protein complexes on the basis of highly ambiguous intermolecular distance restraints derived from 1H/15N chemical shift mapping and backbone 15N-1H residual dipolar couplings using conjoined rigid body/torsion angle dynamics**. *J Am Chem Soc* 2003, **125**:2902-2912.

The use of CSP data as ambiguous interaction restraints as introduced in [9••] is complemented by RDC restraints which can resolve possible orientational ambiguities.

11. van Dijk ADJ, Fushman D, Bonvin AMJJ: **Various strategies of using residual dipolar couplings in NMR-driven protein docking: application to Lys48-linked di-ubiquitin and validation against 15N-relaxation data**. *Proteins* 2005, **60**:367-381.

12. Morelli XJ, Palma PN, Guerlesquin F, Rigby AC: **A novel approach for assessing macromolecular complexes combining soft-docking calculations with NMR data**. *Protein Sci* 2001, **10**:2131-2137.

13. Dobrodumov A, Gronenborn AM: **Filtering and selection of structural models: combining docking and NMR**. *Proteins* 2003, **53**:18-32.

14. Fahmy A, Wagner G: **TreeDock: a tool for protein docking based on minimizing van der Waals energies**. *J Am Chem Soc* 2002, **124**:1241-1250.

15. Hajduk PJ, Mack JC, Olejniczak ET, Park C, Dandliker PJ, Beutel BA: **SOS-NMR: A saturation transfer NMR-based method for determining the structures of protein-ligand complexes**. *J Am Chem Soc* 2004, **126**:2390-2398.

16. Reese ML, Dotsch V: **Fast mapping of protein-protein interfaces by NMR spectroscopy**. *J Am Chem Soc* 2003, **125**:14250-14251.

17. Takahashi H, Nakanishi T, Kami K, Arata Y, Shimada I: **A novel NMR method for determining the interfaces of large protein-protein complexes**. *Nat Struct Biol* 2000, **7**:220-223.

18. Matsuda T, Ikegami T, Nakajima N, Yamazaki T, Nakamura H: **Model building of a protein-protein complexed structure using saturation transfer and residual dipolar coupling without paired intermolecular NOE**. *J Biomol NMR* 2004, **29**:325-338.

19. Prudêncio M, Ubbink M: **Transient complexes of redox proteins: structural and dynamic details from NMR studies**. *J Mol Recognit* 2004, **17**:524-539.

20. Gaponenko V, Sarma SP, Altieri AS, Horita DA, Li J, Byrd RA: **Improving the accuracy of NMR structures of large proteins using pseudocontact shifts as long-range restraints**. *J Biomol NMR* 2004, **28**:205-212.

21. Arumugam S, Van Doren SR: **Global orientation of bound MMP-3 and N-TIMP-1 in solution via residual dipolar couplings**. *Biochemistry* 2003, **42**:7950-7958.

22. Sakakura M, Noba S, Luchette PA, Shimada I, Prosser RS: **An NMR method for the determination of protein-binding interfaces using dioxygen-induced spin-lattice relaxation enhancement**. *J Am Chem Soc* 2005, **127**:5826-5832.

23. Fushman D, Varadan R, Assfalg M, Walker O: **Determining domain orientation in macromolecules by using spin-relaxation and residual dipolar coupling measurements**. *Prog Nucl Magn Reson Spectrosc* 2004, **44**:189-214.

24. • Vaynberg J, Fukuda T, Chen K, Vinogradova O, Velyvis A, Tu Y, Ng L, Wu C, Qin J: **Structure of an ultraweak protein-protein complex and its crucial role in regulation of cell morphology and motility**. *Mol Cell* 2005, **17**:513-523.

This ultraweak complex ($K_D \sim 3 \times 10^{-3}$ M) was solved *de novo* using a combination of NOEs and RDCs, nicely illustrating the power of NMR when it comes to studying weak interactions.

25. Sun Z-YJ, Kim ST, Kim IC, Fahmy A, Reinherz EL, Wagner G: **Solution structure of the CD3epsilondelta ectodomain and comparison with CD3epsilongamma as a basis for modeling T cell receptor topology and signaling**. *Proc Natl Acad Sci USA* 2004, **101**:16867-16872.

26. Alam SL, Sun J, Payne M, Welch BD, Blake BK, Davis DR, Meyer HH, Emr SD, Sundquist WI: **Ubiquitin interactions of NZF zinc fingers**. *EMBO J* 2004, **23**:1411-1421.

27. Thiru A, Nietlispach D, Mott HR, Okuwaki M, Lyon D, Nielsen PR, Hirshberg M, Verreault A, Murzina NV, Laue ED: **Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin**. *EMBO J* 2004, **23**:489-499.

28. Vakonakis I: **LiWang AC: Structure of the C-terminal domain of the clock protein KaiA in complex with a KaiC-derived peptide: implications for KaiC regulation**. *Proc Natl Acad Sci USA* 2004, **101**:10925-10930.

29. Feng W, Long J-F, Fan J-S, Suetake T, Zhang M: **The tetrameric L27 domain complex as an organization platform for supramolecular assemblies**. *Nat Struct Mol Biol* 2004, **11**:475-480.

30. De Guzman RN, Martinez-Yamout MA, Dyson HJ, Wright PE: **Interaction of the TAZ1 domain of the CREB-binding protein with the activation domain of CITED2: regulation by competition between intrinsically unstructured ligands for non-identical binding sites**. *J Biol Chem* 2004, **279**:3042-3049.

31. •• Gross JD, Moerke NJ, von der Haar T, Lugovskoy AA, Sachs AB, McCarthy JEG, Wagner G: **Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E**. *Cell* 2003, **115**:739-750.

The solution structure of the large (312 amino acid) complex between the eukaryotic initiation factors 4G and 4E is described. The NOE information was supplemented by restraints derived from paramagnetic broadening effects. An interesting aspect of this complex is that it is the result of a coupled folding transition of its components.

32. Cai M, Williams DC, Wang G, Lee BR, Peterkofsky A, Clore GM:
•• **Solution structure of the phosphoryl transfer complex between the signal-transducing protein IIAGlucose and the cytoplasmic domain of the glucose transporter IICBGlucose of the *Escherichia coli* glucose phosphotransferase system**. *J Biol Chem* 2003, **278**:25191-25206.
This complex was solved *de novo* to high precision using a combination of NOEs, several sets of RDCs and $C\alpha C\beta$ chemical shift restraints. This work illustrates state-of-the art NMR methodology for the structural studies of protein–protein complexes.

33. Kang RS, Daniels CM, Francis SA, Shih SC, Salerno WJ, Hicke L, Radhakrishnan I: **Solution structure of a CUE-ubiquitin complex reveals a conserved mode of ubiquitin binding**. *Cell* 2003, **113**:621-630.

34. Mueller TD, Feigon J: **Structural determinants for the binding of ubiquitin-like domains to the proteasome**. *EMBO J* 2003, **22**:4634-4645.

35. Freedman SJ, Sun Z-YJ, Kung AL, France DS, Wagner G, Eck MJ: **Structural basis for negative regulation of hypoxia-inducible factor-1alpha by CITED2**. *Nat Struct Biol* 2003, **10**:504-512.

36. Williams DC, Cai M, Suh JY, Peterkofsky A, Clore GM: **Solution**
• **NMR structure of the 48 kDa IIAmannose-HPr complex of the *Escherichia coli* mannose phosphotransferase system**. *J Biol Chem* 2005, **280**:20775-20784.
This large complex of two HPr molecules bound to dimeric $IIA^{mannose}$ was assembled from the unbound components using a conjoint rigid body/torsion angle simulated annealing procedure on the basis of intermolecular NOEs and RDCs.

37. Arnesano F, Banci L, Bertini I, Bonvin AMJJ: **A docking approach to the study of copper trafficking proteins; interaction between metallochaperones and soluble domains of copper ATPases**. *Structure (Camb)* 2004, **12**:669-676.

38. Díaz-Moreno I, Díaz-Quintana A, De la Rosa MA, Ubbink M:
• **Structure of the complex between plastocyanin and cytochrome f from the cyanobacterium *Nostoc* sp. PCC 7119 as determined by paramagnetic NMR: the balance between electrostatic and hydrophobic interactions within the transient complex determines the relative orientation of the two proteins**. *J Biol Chem* 2005, **280**:18908-18915.
The potential of paramagnetic NMR for the study of weak and transient complexes is demonstrated: CSP data and intermolecular pseudo contact shifts are introduced in a rigid-body calculation to determine the relative orientation of the two molecules.

39. Dominguez C, Bonvin AMJJ, Winkler GS, van Schaik FMA,
• Timmers HTM, Boelens R: **Structural model of the UbcH5B/CNOT4 complex revealed by combining NMR, mutagenesis, and docking approaches**. *Structure (Camb)* 2004, **12**:633-644.
CSP data in combination with site-directed mutagenesis were used in HADDOCK to model the UbcH5B/CNOT4 complex. Charge reversal mutants could be generated based on the docking models in which the identity of a Lys and Glu were switched across the interface. These swapped mutations were shown to restore binding in a two-hybrid binding assay.

40. Arunkumar AI, Klimovich V, Jiang X, Ott RD, Mizoue L, Fanning E,
• Chazin WJ: **Insights into hRPA32 C-terminal domain-mediated assembly of the simian virus 40 replisome**. *Nat Struct Mol Biol* 2005, **12**:332-339.
The interaction between hRPA32C and the Tag origin DBD is studied by affinity chromatography, NMR and site-directed mutagenesis. The latter was guided by a structural model of the complex generated from CSP data and independently validated with RDC data.

41. Dangi B, Gronenborn AM, Rosner JL, Martin RG: **Versatility of the**
• **carboxy-terminal domain of the alpha subunit of RNA polymerase in transcriptional activation: use of the DNA contact site as a protein contact site for MarA**. *Mol Microbiol* 2004, **54**:45-59.
The ternary complex of the C-terminal domain of the α-subunit of RNA polymerase with the transcription factor MarA and DNA was modeled with HADDOCK using NMR-based chemical shift mapping.

42. Jain NU, Wyckoff TJO, Raetz CRH, Prestegard JH: **Rapid analysis**
• **of large protein-protein complexes using NMR-derived orientational constraints: the 95 kDa complex of LpxA with acyl carrier protein**. *J Mol Biol* 2004, **343**:1379-1389.
This article demonstrates that NMR can play a role in the analysis of very large complexes: by combining RDC and CSP data measured for the small (9 kDa) acyl carrier protein with mutagenesis data for LpxA and exploiting the internal C3 symmetry for alignment, the 95kDa homotrimeric complex could be modelled.

43. Messias A, Sattler M: **Structural basis of single-stranded RNA recognition**. *Acc Chem Res* 2004, **37**:279-287.

44. Maris C, Dominguez C, Allain F: **The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression**. *FEBS J.* 2005, **272**:2118-2131.

45. Johansson C, Finger L, Trantirek L, Mueller T, Kim S, Laird-Offringa
• I, Feigon J: **Solution structure of the complex formed by the two N-terminal RNA-binding domains of nucleolin and a pre-rRNA target**. *J Mol Biol* 2004, **337**:799-816.
The first two RNA-binding domains of nucleolin are responsible for RNA-binding specificity and affinity. The linker between the domains, flexible in free nucleolin, becomes structured in the complex due to extensive contacts to the b2NRE major groove.

46. Allain F, Yen Y, Masse J, Schultze P: **T D, J F: Molecular basis of sequence-specific recognition of pre-ribosomal RNA by nucleolin**. *EMBO J* 2000, **19**:6870-6881.

47. Liu Z, Luyten I, Bottomley M, Messias A, Houngninou-Molango S, Sprangers R, Zanier K, Kramer A, Sattler M: **Structural basis for recognition of the intron branch site RNA by splicing factor 1**. *Science* 2001, **294**:1098-1102.

48. Lu D, Searles M, Klug A: **Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition**. *Nature* 2003, **426**:96-100.

49. Hudson B, Martinez-Yamout M, Dyson H, Wright P: **Recognition**
•• **of the mRNA AU-rich element by the zinc finger domain of TIS11d**. *Nat Struct Mol Biol* 2004, **11**:257-264.
Two CCCH zinc finger motifs, that represent a large family of eukaryotic proteins, present a platform for highly specific recognition of an AU-rich mRNA sequence. Sequence specificity comes mostly from TIS11d main chain hydrogen bond contacts with the Watson–Crick edges of the bases along the extended mRNA chain.

50. D'Souza V, Summers M: **Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus**. *Nature* 2004, **431**:586-590.

51. Dey A, York D: **A S-M, MF S: Composition and sequence-dependent binding of RNA to the nucleocapsid protein of Moloney murine leukemia virus**. *Biochemistry* 2005, **44**:3735-3744.

52. Chang K, Ramos A: **The double-stranded RNA-binding motif, a versatile macromolecular docking platform**. *FEBS J.* 2005, **272**:2109-2117.

53. Wu H, Henras A, Chanfreau G, Feigon J: **Structural basis for**
• **recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III**. *Proc Natl Acad Sci USA* 2004, **101**:8307-8312.
Recognition of the dsRNA by the dsRBD is mostly structure specific.

54. Lingel A, Simon B, Izaurralde E, Sattler M: **Nucleic acid 3′-end**
•• **recognition by the Argonaute2 PAZ domain**. *Nat Struct Mol Biol* 2004, **11**:576-577.
This structure reveals a unique mode of single-stranded nucleic acid binding. Specific recognition is mediated by binding of the two-nucleotide 3′ overhang characteristic of siRNAs to a hydrophobic cleft in the PAZ domain.

55. Calabro V, Daugherty M, Frankel A: **A single intermolecular contact mediates intramolecular stabilization of both RNA and protein**. *Proc Natl Acad Sci USA* 2005, **102**:6849-6854.

56. Williams DC, Cai M, Clore GM: **Molecular basis for synergistic**
•• **transcriptional activation by Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1.Sox2.Hoxb1-DNA ternary transcription factor complex**. *J Biol Chem* 2004, **279**:1449-1457.
This large ternary protein–DNA complex was solved using state of the art NMR methodology. The interaction of two transcription factor domains (Oct1-POU and Sox2) bound to adjacent sites on the DNA suggests a mechanism for cell-specific transcription regulation.

57. Kalodimos CG, Biris N, Bonvin AMJJ, Levandoski MM,
•• Guennuegues M, Boelens R, Kaptein R: **Structure and flexibility**

adaptation in nonspecific and specific protein-DNA complexes. *Science* 2004, **305**:386-389.
The structure of a complex of a covalently linked lac repressor headpiece with non-operator DNA is described (see Figure 4). Large differences in structure and dynamics between non-specific and specific complexes provide insight into the way gene regulatory proteins locate their DNA target sites.

58. Iwahara J, Schwieters CD, Clore GM: **Ensemble approach for NMR structure refinement against (1)H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule**. *J Am Chem Soc* 2004, **126**:5879-5896.

59. Iwahara J, Schwieters CD, Clore GM: **Characterization of nonspecific protein-DNA interactions by 1H paramagnetic relaxation enhancement**. *J Am Chem Soc* 2004, **126**:12800-12808.

60. • Mitton-Fry RM, Anderson EM, Theobald DL, Glustrom LW, Wuttke DS: **Structural basis for telemoric single-stranded DNA recognition by Yeast Cdc13**. *J Mol Biol* 2004, **338**:241-255.
Cdc13 binds to single-stranded GT-rich telomere ends and protects them from degradation. The complex of the DNA-binding domain of Cdc13 with telomeric ssDNA explains the high affinity and specificity towards GT-rich single-stranded DNA.

61. Rumpel S, Razeto A, Pillar CM, Vijayan V, Taylor A, Giller K, Gilmore MS, Becker S, Zweckstetter M: **Structure DNA-binding properties of the cytolysin regulator CylR2 from *Enterococcus faecalis***. *EMBO J* 2004, **23**:3632-3642.

62. Surdo PL, Bottomley MJ, Sattler M, Scheffzek K: **Crystal structure and nuclear magnetic resonance analyses of the SAND domain from glucocorticoid modulatory element binding protein-1 reveals deoxyribonucleic acid and zinc binding regions**. *Mol Endocrinol* 2003, **17**:1283-1295.

63. Dupureur CM: **NMR studies of restriction enzyme-DNA interactions: role of conformation in sequence specificity**. *Biochemistry* 2005, **44**:5065-5074.

64. Pristovsek P, Sengupta K, Löhr F, Schäfer B, von Trebra MW, Rüterjans H, Bernhard F: **Structural analysis of the DNA-binding domain of the *Erwinia amylovora* RcsB protein and its interaction with the RcsAB box**. *J Biol Chem* 2003, **278**:17752-17759.

65. Lee J-H, Park C-J, Arunkumar AI, Chazin WJ, Choi B-S: **NMR study on the interaction between RPA and DNA decamer containing cis-syn cyclobutane pyrimidine dimer in the presence of XPA: implication for damage verification and strand-specific dual incision in nucleotide excision repair**. *Nucleic Acids Res* 2003, **31**:4747-4754.

66. Daughdrill GW, Buchko GW, Botuyan MV, Arrowsmith C, Wold MS, Kennedy MA, Lowry DF: **Chemical shift changes provide evidence for overlapping single-stranded DNA- and XPA-binding sites on the 70 kDa subunit of human replication protein A**. *Nucleic Acids Res* 2003, **31**:4176-4183.

67. Stauffer ME, Chazin WJ: **Physical interaction between replication protein A and Rad51 promotes exchange on single-stranded DNA**. *J Biol Chem* 2004, **279**:25638-25645.

68. Vise PD, Baral B, Latos AJ, Daughdrill GW: **NMR chemical shift and relaxation measurements provide evidence for the coupled folding and binding of the p53 transactivation domain**. *Nucleic Acids Res* 2005, **33**:2061-2077.