NMR Approaches to Large Proteins:

trp Repressor and Chloramphenicol Acetyltransferase

L.-Y. Lian, J. P. Derrick, V. Ramesh, R. O. Frederick,

S. E. H. Syed and G. C. K. Roberts

Biological NMR Centre, University of Leicester,
PO Box 138, Medical Sciences Bldg.,
University Road, Leicester LE1 9HN, U.K.

1. Introduction

The current upper limit for protein structure determination by ^1H nmr [1] is in the region of M_r 12-15,000. The use of stable isotope labelling, with ^2H, ^{13}C or ^{15}N, can substantially extend this limit, perhaps to M_r 30,000 (for methodological reviews, see [2]-[4]). However, there remain many interesting and important proteins which are much larger than this, and we have been trying to establish what information nmr can provide in such cases. Two systems we have been studying in this context are the *E. coli trp* repressor and the enzyme chloramphenicol acetyltransferase.

The *trp* repressor is a dimer of total M_r 25,000, and is a member of the "helix-turn-helix" family of DNA binding proteins. A number of high resolution crystal structures of the protein are available, but the mechanisms of its activation and DNA-binding specificity remain poorly understood. Our main priority is therefore to determine the solution structure of the protein-oligonucleotide complex, and the effects on this of changes in protein and oligonucleotide structure; our work to date has largely been devoted to developing the necessary methods for this. We have, in collaboration with the group of Jardetzky at Stanford, achieved a virtually complete sequence-specific assignment of the ^1H nmr spectrum of the protein [5]-[7], and Jardetzky's group have used this data to determine the solution structure of the repressor-tryptophan complex [8]. Nmr studies of the binding of corepressors such as L-tryptophan and the inducer, indole-3-propionic acid, show that the environments of the two classes of ligand in the protein differ, and strongly suggest that this arises from a difference in the orientation of their indole rings [9]. Even the repressor alone is of such a size that these studies required the extensive use of isotope labelling, and we have developed appropriate expression systems for efficient biosynthetic labelling of the protein. The overall molecular mass of the repressor-tryptophan-operator oligonucleotide complex is *ca.* 38,000, and for this complex rather little useful information can be obtained by nmr without labelling.

Chloramphenicol acetyltransferase is responsible for resistance to the antibiotic chloramphenicol in bacteria. Its crystal structure is known to high resolution, and it is the subject of a substantial programme of protein engineering in the laboratory of Prof.

W.V. Shaw in Leicester [10]. Since it is a trimer of total molecular mass 75,000, it is very large for study by nmr, and we are using it as a test system for the development of nmr methods suitable for large proteins, focussing particularly on complexes, such as that with the product, diacetyl-chloramphenicol, which have not proved amenable to crystallographic study.

2. Deuteration

Selective deuteration has been used to simplify the ^{1}H nmr spectra of proteins for many years [11], and more recently has been combined with 2D nmr spectroscopy (*e.g.*, [12]). Selective deuteration of carefully chosen combinations of residues has been combined with the sequential assignment method of Wüthrich (see [1]) to yield a considerable number of resonance assignments in the *trp* repressor, both alone and in its complex with an operator oligonucleotide [6], [7]. With the improved resolution of the 3D spectra, we have recently assigned all the intermolecular NOEs between protons of the protein and those of bound tryptophan, permitting the kind of 'ligand-docking' experiments which we have already carried out on phospholipase A_2 [13]. In a species as large as the repressor-operator complex, the substitution of most of the protons in the protein by deuterons leads to a notable decrease in the resonance linewidth of the remaning protons; similarly, we have employed random fractional deuteration [14] to good effect in these large systems. In particular, deuteration of chloramphenicol acetyl-transferase to the level of about 85% had three beneficial effects: (a) it allowed the unambiguous observation and assignment of resonances of the bound substrate, (b) it allowed the observation of intra-molecular NOEs in the bound substrate, thus defining its conformation, and (c) it allowed the observation of inter-molecular NOEs between protons of the (isotopically normal) substrate and nearby residues of the protein [15].

3. ^{13}C and ^{15}N labelling

Notwithstanding the usefulness of selective deuteration, complete assignments of the backbone resonances of the repressor have required the use of ^{15}N-labelled protein, in combination with 3D nmr spectroscopy [7]. The same approach has led to a substantial number of resonance assignments in the repressor-operator complex, and to the clear observation of inter-molecular NOEs [7], [16]. In addition to the standard NOESY-HMQC and TOCSY-HMQC experiments, we have found the HMQC-NOESY-HMQC experiment [17] valuable in this α-helical protein in which there is significant overlap of both ^{1}H and ^{15}N chemical shifts of the backbone amides. In some cases, there is a considerable benefit in selective labelling; in the *trp* repressor we have used this successfully for ^{15}N-leucine and for [amide-^{15}N]-asparagine and glutamine [16], and in chloramphenicol acetyltransferase for [imidazole 2-^{13}C]-histidine [18]. The latter experiment permitted the detection of the histidine C2-^{1}H signals, and the use of 2D ^{1}H-^{13}C correlation spectra allowed overlapping signals to be resolved, even in a protein of M_r 75,000. We are also exploring the usefulness of 50-65% perdeuteration in combination with ^{15}N-labelling as a means of improving the quality of the 3D spectra.

In chloramphenicol acetyltransferase we have used nmr to study the binding of diacetyl[^{13}C]-chloramphenicol by means of ^{13}C-edited ^{1}H-^{1}H NOESY experiments. A number of clear intermolecular NOEs were observed, in particular to aromatic protons of the protein. Candidate aromatic residues were identified by model-building on the basis of the crystal structure, and were replaced in turn by isoleucine residues. ^{13}C-edited NOESY spectra of the complexes with these mutants allowed the two aromatic residues with which the acetyl groups of the ligand made contact to be identified unambiguously, thus allowing the orientation of the product in the binding site to be

defined [15], and assisting in the modelling of the transition-state complex.

4. References

[1] Wüthrich, K., *Science* **243** (1989) 45.

[2] Bax, A. (1989) *Ann. Rev. Biochem.* **58**, 223.

[3] Oppenheimer, N.J., and James, T.L. (1989) Methods in Enzymology, **176**, Academic Press, New York.

[4] Oppenheimer, N.J., and James, T.L. (1989) Methods in Enzymology, **177**, Academic Press, New York.

[5] Hyde, E.I., Ramesh. V., Roberts. G.C.K., Arrowsmith, C.H., Treat-Clemons, L., Klaic, B., and Jardetzky, O. (1989) Eur. J. Biochem. **183**,545.

[6] Arrowsmith, C.H., Pachter, R., Altman, R.B., Iyer, S.B., and Jardetzky, O. (1990) Biochemistry **29**, 6332.

[7] Ramesh, V., Frederick, R.O., Syed, S.E.H., and Roberts, G.C.K., unpublished work; Arrowsmith, C.H., and Jardetzky, O., unpublished work.

[8] Arrowsmith, C.H., Pachter, R., Altman, R.B., and Jardetzky, O. (1991) Eur. J. Biochem., **202**, 53.

[9] Hyde, E.I., Ramesh, V., Frederick, R., and Roberts, G.C.K. (1991) Eur. J. Biochem., **201**, 569.

[10] Shaw, W.V., and Leslie, A.G.W. (1991) Ann. Rev. Biophys. Biophys. Chem., **20**, 363.

[11] Jardetzky, O. & Roberts, G.C.K. (1981) *NMR in Molecular Biology*, Academic Press, New York.

[12] Feeney, J., Birdsall, B., Akiboye, J., Tendler, S.J.B., Barbero, J.J., Ostler, G., Arnold, J.R.P., Roberts, G.C.K., Kuhn, A., & Roth, K. (1989) FEBS Lett., **248**, 57.

[13] Bennion, C., Connolly, S., Gensmantel, N.P., Hallam, C., Jackson, C.G., Primrose, W.U., Roberts, G.C.K., Robinson, D.H., and Slaich, P.K. (1992) J. Med. Chem., in press.

[14] LeMaster, D.M. (1990) Quart. Rev. Biophys., **23**, 133.

[15] Derrick, J.P., Lian, L.-Y., Roberts, G.C.K., and Shaw, W.V. (1992) Biochemistry, in press.

[16] Ramesh, V., Frederick, R.O., Syed, S.E.H., and Roberts, G.C.K. (1992) manuscript in preparation.

[17] Frenkiel, T., Bauer, C., Carr, M.D., Birdsall, B., and Feeney, J. (1990) J. Magn. Reson., **90**, 420.

[18] Derrick, J.P., Lian, L.-Y., Roberts, G.C.K., and Shaw, W.V. (1991) FEBS Lett., **280**, 125.