© 1997 Nature Publishing Group http://www.nature.com/nsmb



# Large modular proteins by NMR

Sir—Due to increased overall rotational correlation times and the resulting relaxation properties, large proteins generally have essentially unobservable high-resolution solution NMR signals. Therefore the technique is usually restricted to the study of proteins of less than 30,000–40,000 $M_r$. We have recently obtained high-resolution heteronuclear NMR data for the *Escherichia coli* chemotaxis kinase CheA, having a dimer molecular weight of 142,000 $M_r$, which is much higher than the current sizes of proteins studied by solution NMR. In this case, the modular domain structure of CheA, having domains connected by flexible linkers, has made it possible to obtain excellent quality high-resolution NMR spectra of this protein dimer.

The chemotaxis autokinase CheA consists of five functional domains responsible for phosphotransfer, phosphoacceptor binding, ATP binding and catalysis, receptor coupling and CheW coupling as determined by sequence analysis and biochemical characterization[1–5]. The full-length protein consists of 654 amino acids and exists as a dimer at biologically relevant concentrations[6]. The NMR chemical shift assignments and structures of the isolated phosphotransfer[7] (P1) and phosphoacceptor binding[8,9] (P2) domains have been determined by solution NMR studies.

The observation that the P1 and P2 domains are connected by a flexible linker of approximately 24 amino acid residues and behave independently[10] led us to try to observe the properties of the linkers and domains in the context of the full-length protein dimer. Surprisingly, chemical shifts characteristic of the folded P2 domain[8] were present in the spectrum of the full-length protein (Fig. 1). The fact that resonances are observable at all in a protein dimer of this size is remarkable, though resonances from terminal domains or unfolded regions of high molecular weight proteins have been observed in other cases[11]. The chemical shifts, and therefore likely the structure, of this domain are little affected by the rest of the kinase. We conclude that this domain behaves essentially as an independent unit tethered by flexible linkers with little or no stable interactions with the rest of the protein. Also observed in this spectrum are a number of resonances that are not identifiable as peaks from the P2 domain. These peaks, many of which occur in the central, highly overlapped region of the spectrum, may be from other disordered regions of the full-length protein that also have motions independent from those of the bulk of the dimer.

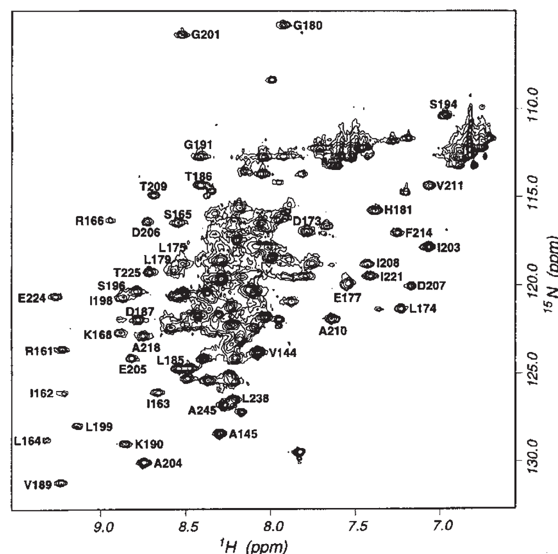


**Fig. 1** $^1$H–$^{15}$N correlation spectrum[12] of $^{15}$N-labelled CheA at 500 MHz and 30 °C using pulsed field gradients for solvent suppression and coherence-pathway rejection. The spectrum was acquired as a 256 (real)×1024 (complex) matrix using 256 scans per increment. Protein concentration is approximately 0.8 mM in 50 mM Tris pH 7.5, 1 mM EDTA, 10% glycerol, 2 mM DTT, 0.2 mM PMSF and 0.02% sodium azide. The peaks that are characteristic of the isolated P2 domain[8] are labelled in this spectrum. The proton peak linewidths are approximately 30 Hz, as compared to 23 Hz for the isolated 14,000 $M_r$ P2-containing fragment.

Generally, for kinases of the two-component family, the phosphotransfer domain containing the histidine that is the site of phosphorylation is adjacent to the regions of ATP binding and catalysis[1]. The P2 domain is not conserved in other kinases of two-component signal transduction pathways. For the autokinase CheA, the P2 domain may serve to enhance specificity in the phosphotransfer reaction to limit crosstalk among a number of homologous systems which may be present within a single cell. In order to localize the downstream phosphoacceptor, the P2 domain does not need to be an integral structural part of the kinase mechanism. The apparent independence of the P2 domain from the rest of the kinase molecule may reflect a later evolutionary addition to the kinase held by flexible linkers in order to minimize disturbance of the existing catalytic functions.

We suggest that the modular domain structure of many proteins may make solution NMR a generally useful technique for the study of proteins with much higher molecular weights than has been appreciated.

**Megan M. McEvoy, Aida Flor A. de la Cruz and Frederick W. Dahlquist**

Institute of Molecular Biology, University of Oregon
Eugene, Oregon 97403 USA

Correspondence should be addressed to F.W.D
*fwd@nmr.uoregon.edu*

Received 19 September; accepted 9 October, 1996.

1. Parkinson, J.S. & Kofoid, E.C. *Annu. Rev. Genet.* **26**, 71–112 (1992).
2. Morrison, T.B. & Parkinson, J.S. *Proc. Natl. Acad. Sci. USA* **91**, 5485–5489 (1994).
3. Swanson, R.V. *et al. Biochemistry* **32**, 7623–7629 (1993).
4. Bourret, R.B. *et al. J. Bacteriol.* **175**, 2097–2101 (1993).
5. Stock, J.B. *et al.* in Two-Component Signal Transduction (Eds J.A. Hoch & T.J. Silhavy) 25–51 (American Society for Microbiology, 1995).
6. Gegner, J.A. & Dahlquist, F.W. *Proc. Natl. Acad. Sci. USA* **88**, 750–754 (1991).
7. Zhou, H. *et al. Biochemistry* **34**, 13858–13870 (1995).
8. McEvoy, M.M. *et al. Biochemistry* **34**, 13871–13880 (1995).
9. McEvoy, M.M. *et al. Biochemistry* **35**, 5633–5640 (1996).
10. Zhou, H. *et al. Biochemistry* **35**, 433–443 (1996).
11. Gettins, P. & Cunningham, L.W. *Biochemistry* **25**, 5011–5017 (1986), and references therein.
12. Wider, G. & Wüthrich, K. *J. Magn. Reson.* B**102**, 239–241 (1993).