CMLS, Cell. Mol. Life Sci. 54 (1998) 359–362
1420-682X/98/040359-04 $ 1.50 + 0.20/0

CMLS Cellular and Molecular Life Sciences

NMR structural studies of glutathione S-transferase

L.-Y. Lian

Biological NMR Centre, P.O. Box 138, Maurice Shock Medical Sciences Building, University of Leicester, Leicester LE1 9HN (UK), Fax +44 116 2523013, e-mail: yun@leicester.ac.uk

Abstract. The use of nuclear magnetic resonance (NMR) spectroscopy for the structure determination of small proteins is now widely recognized; what is less frequently reported is the application of NMR techniques for high-resolution studies of large proteins (M_r larger than 30 kD). We demonstrate here how an integrated approach, using heteronuclear NMR and X-ray crystallography, can provide useful and biologically important information for large protein systems. The dynamic features of the human A1-1 glutathione S-tranferase and the role of the C-terminal region are being probed by NMR; in the X-ray crystal structure, the electron densities for this region of the protein are uninterpretable.

Key words. Glutathione S-transferase; NMR; high resolution; phenylalanines; isotopic labelling; conformational changes; dynamics.

Introduction

Over the last 15 years, nuclear magnetic resonance (NMR) spectroscopy has become one of the most important tools for structure characterization [1, 2]. Apart from being able to provide high-resolution three-dimensional (3D) structures, NMR has been increasingly used to obtain details of dynamics, of the ionization states of groups in proteins and nucleic acids, and of either local or global conformational changes on ligand binding, on the formation of transition states in an enzymic reaction or during protein folding and unfolding. NMR is now perceived as an essential tool in structural biology rather than a luxury. Many of the experiments that were at one time seen as advanced techniques are now used routinely in most NMR laboratories. Two main advances have contributed to this tranformation: molecular biology – for making isotopically labelled recombinant proteins and mutants – and NMR technology, where larger magnets, better electronics and hardware, and faster computers have made more sophisticated experiments easier to perform, giving data with higher sensitivity and fewer artefacts. Due to the characteristics of an NMR resonance (relaxation decay and chemical shift dispersion) there is still an upper limit to the size of a macromolecular system for which high-resolution 3D information can be obtained. Nevertheless, when used in combination with X-ray crystallography (and molecular biology) the normal limits of molecular masses for structural studies by NMR can be pushed upwards; very detailed and useful information can now be obtained for systems of molecular masses up to 80 kD. One of the main advantages in being able to express proteins in *Escherichia coli* is that it is now possible to make isotopically labelled proteins in a variety of ways – uniform $^{13}C/^{15}N$, amino acid-specific $^{13}C/^{15}N$, labelling at specific sites of amino acids, perdeuteration in combination with uniform or selective $^{13}C/^{15}N$ labelling, perdeuteration with selective protonation, and so on [3, 4]. All these labelling strategies are geared towards simplifying the complicated NMR spectrum of large molecules and for reducing the linewidths of the NMR signals to make them detectable at good signal-to-noise ratios.

Our study of glutathione S-transferases (GST) is an example of how NMR is used in a complementary manner to X-ray crystallography and other biophysical techniques. GSTs are a group of detoxification enzymes

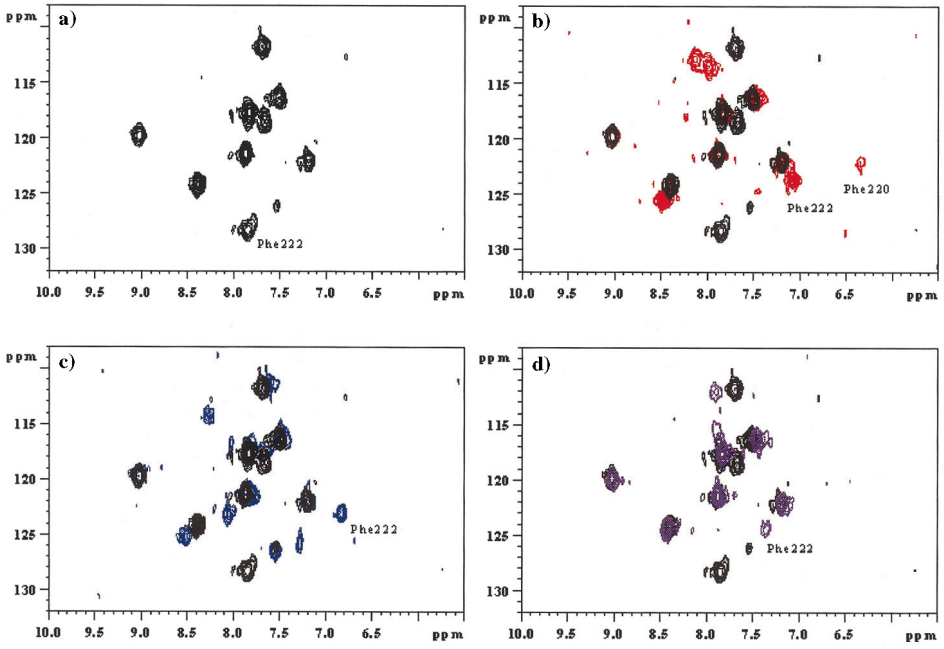

Figure 1. ^{13}N-^{1}H heteronuclear single quantum coherence (HSQC) spectrum of 0.5 mM human A1-1 C112S mutant glutathione S-transferase at 25 °C, pH 6.5. The mutant was used to avoid the complications of having to add reductants such as β-mercaptoethanol or dithiothreitol to the sample, both of which bind to the protein in the G site. The resonance assignments for Phe 220 and Phe 222 are indicated. The spectrum of the apo-protein is shown in (*a*); the ligand-bound spectra in the presence of 3 mM dinitrophenylGSH (red), 3 mM CDNB (blue) and 1.5 mM GSH (magenta) are shown in (*b*), (*c*) and (*d*), respectively. The complex spectra are all plotted as overlays, with the apo-protein spectrum (in black).

that catalyse the conjugation of glutathione to electrophilic subtrates, many of which are toxic chemicals [5 and references therein]. All eukaryotic species possess multiple cytosolic and membrane-bound GST isozymes, each of which shows distinct catalytic properties. There are five main classes of cytosolic enzymes, designated alpha, mu, pi, sigma and theta [6]. All the cytosolic GST are dimers of subunit molecular masses ranging from 23 kD to 27 kD. Representative crystal structures are available for each of the five structures, and in the presence and absence of substrates and ligands. The overall polypeptide fold is very similar between the crystal structures, but each class exhibits unique features, in particular about the active site and at the C-terminus. In general the glutathione-binding (G) site exhibits a high specificity for glutathione, whereas the substrate binding (H) site displays a broad specificity towards hydrophobic compounds.

Despite enormous progress in obtaining crystal structures of the different proteins in the presence and absence of substrates and inhibitors, many of the fundamental details of the enzyme mechanism are still poorly understood. An examination of the temperature factors of the various crystal structures reveals that in all the GST structures, several common regions of high temperature factors are found; these regions generally include the substrate-binding sites and regions where significant structural changes have been shown to occur on binding to substrates. These results suggest that flexibility and dynamics are important characteristics of the protein and that, when a substrate binds, 'movement' of sections of the protein is involved. Whether there is cooperativity between the GSH- and substrate-binding sites remains to be established structurally. The flexibility of the hydrophobic substrate binding site is not surprising considering the broad specificity of GSTs towards hydrophobic substrates.

It is the dynamic characteristics of the protein which disadvantage X-ray crystallography as a tool for obtaining detailed structural information, and at the same time make NMR a more suitable technique for detecting and monitoring the important changes in dynamics and structure when the protein interacts with substrates and inhibitors. We shall show below how NMR is used for these kinds of studies, using the human alpha A1-1 GST ($M_r = 52$ kD) as an example.

The first problem to overcome is spectral simplification; this is conveniently done by expressing the plasmid harbouring the human alpha A1-1 GST gene in auxotrophic strains of *E. coli*, using a defined growth medium which contains the desired isotopically labelled amino acid [4]. Auxotrophic strains were used to give a high level of isotopic incorporation without the risk of isotopic dilution and/or cross-labelling, both of which are disadvantageous and complicate the resulting spectrum. Figure 1 shows the spectra of GST with only the phenylalanine residues labelled with ^{15}N. Sequence-specific assignments were made by performing the same NMR experiment on a phenylalanine mutant. Examination of the crystal structures of human A1-1 GST reveals that the phenylalanine residues are located in both the glutathione (G)- and the hydrophobic substrate (H)-binding sites. Hence, the resonances from phenylalanines make good NMR probes for structural studies of both the G and H sites. In all the different complexes, the same five phenylalanine amide resonances are virtually unaffected by the presence of any ligands; these resonances must be from phenylalanine residues that are found in neither the G nor H sites. Figure 2 shows a superposition of the backbone atoms of two human alpha A1-1 GST crystal structures; the yellow and blue structure is, respectively, that of the apo-protein in the absence [7] and in the presence of S-benzylGSH (coloured magenta) [8]. It can be seen from this figure that at least five phenylalanine residues are indeed situated far from the substrate-binding sites.

In the absence of any substrates [GSH or chlorodinitrobenzene (CDNB)], only 9 out of the 10 expected phenylalanine NH cross-peaks are observed (fig. 1a). Only when both the G and H sites are occupied is the amide resonance of the 10th phenylalanine observed, albeit at a much lower intensity when compared with the other phenyalanine resonances. This tenth resonance has now been assigned to Phe 220 (fig. 1b). The chemical shift of the resonance from Phe 222 is significantly affected when either GSH or CDNB binds. There are differences in the spectra of the two complexes (figs 1c, d). For example, in the presence of GSH ($K_d = 0.25$ μM) only one set of Phe 222 is present. This is not the case for the CDNB complex ($K_d = 2.1$ μM); multiple resonances are observed for the Phe 222 amide group and several other unassigned Phe residues. Combined with other NMR studies, the presence of multiple peaks is explained by the existence of several bound conformations (orientations) of the CDNB molecule in slow exchange on the NMR time scale.

Both the NMR and kinetics data suggest the following. First, when both the G and H sites are occupied at the same time, the C-terminus is less mobile; when only one site is occupied, the C-terminus is still mobile, although

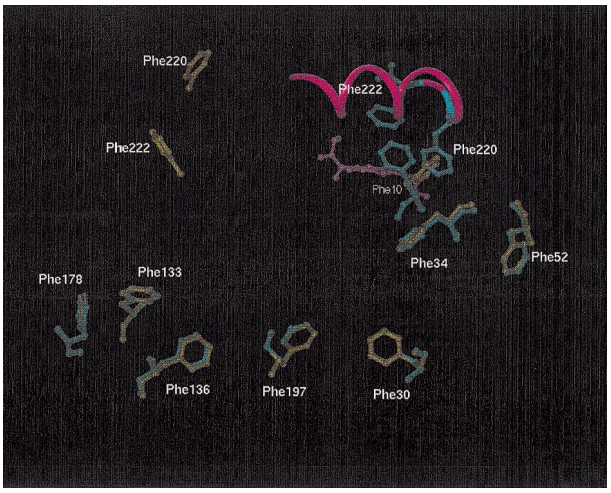

Figure 2. Superposition of the backbone of X-ray structures of human A1-1 GST in the absence (yellow) and presence of S-benzylGSH (blue) illustrating the locations of the phenylalanine side chains.

structural changes from the apo-protein occur, as reflected in the significant chemical shift changes seen for Phe 222, and possibly Phe 10 and Phe 34. Second, multiple conformations are present when a small hydrophobic substrate such as CDNB binds to the G site; when a larger substrate such as ethacrynic acid is used ($K_d = 0.9$ μM), only one set of resonances is observed, suggesting the existence of one bound conformation of the complexed protein (data not shown).

The X-ray structures of this protein in the various forms have shown that in the absence of any ligand, the electron densities for the C-terminal region (210–222) are very weak and uninterpretable. Only in the presence of substrates or inhibitors (ethacrynic acid or GSH conjugates) is the C-terminal helix interpreted as being present, although the temperature factors for this region are still higher than the rest of the molecule. It has been suggested that the C-terminal helix plays an important role in substrate activation, mainly through providing the proper environment for substrate to bind in a productive manner. One of the main questions is whether a helix is only formed when there is a bound ligand, or whether a helix is in fact present but that the poor electron densities are due to mobility in this region. The observation of NMR resonances from the C-terminus region even in the absence of a ligand means that the behaviour of this region can now be probed by NMR. The significant chemical shift changes in the presence of a ligand, and the similarities in the characteristics of these changes in the presence of the different ligands, suggest that the resulting overall conformational

changes are very similar, regardless of which binary complex is formed. In other words, the state of the C-terminal region is similar when either GSH or CDNB or ethacryanic acids binds. However, when both the G and H sites are simultaneously occupied, the C-terminus region appears to be less mobile when compared with the apo form and binary complexes of the protein.

In conclusion, we have demonstrated how NMR can be used in the study of large proteins and how the technique can be used in a complementary manner to X-ray crystallography. Where X-ray crystallography has failed to provide information for the C-terminus region of the apo-protein, NMR data are available for this region. While the NMR data are not at high resolution, it is possible to study many H- and G-site complexes in a very short period of time (a few days). When these data are interpreted together with known crystal structures, useful conclusions can be drawn.

1 NMR Supplement (1997) Nature Struc. Biol. **4:** 841–865
2 Roberts G. C. K. (ed.) (1993) NMR of Macromolecules: A Practical Approach, IRL Press, Oxford
3 LeMaster D. M. (1992) Deuteration in protein proton magnetic resonance. Meth. Enzymol. **177:** 23–43
4 Muchmore D. C., McIntosh L. P., Russell C. B., Anderson D. E. and Dahlquist F. W. (1992) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. Meth. Enzymol. **177:** 44–73
5 Armstrong R. N. (1997) Catalytic mechanism and evolution of the glutathione transferases. Chem. Res. Toxicol. **10:** 2–18
6 Mannervik B., Awasthi Y. C., Board P. G., Hayes J. D., Di Ilio C., Ketterer B. et al. (1992) Nomenclature for human glutathione transferases. Biochem. J. **282:** 305–308
7 Cameron A. D., Sinning I., L'Hermite G., Olin B., Board P. G., Mannervik B. et al. (1995) Structural analysis of human alpha-class glutathione transferaser A1-1 in the apo-form and in complexes with ethacrynic acid and its gluthathione conjugate. Structure **3:** 717–727
8 Sinning I., Kleywegt G. J., Cowan S. W., Reinemer P., Dirr H. W., Huber R. et al. (1993) Structure determination and refinement of human alpha class glutathione transferase A1-1, and a comparison with the mu and pi class enzymes. J. Mol. Biol. **232:** 192–212