Refined solution structure of the 82-kDa enzyme Malate Synthase G from joint NMR and synchrotron SAXS restraints

ARTICLE *in* JOURNAL OF BIOMOLECULAR NMR · MARCH 2008

Impact Factor: 3.14 · DOI: 10.1007/s10858-007-9211-5 · Source: PubMed

J Biomol NMR (2008) 40:95–106
DOI 10.1007/s10858-007-9211-5

ARTICLE

Refined solution structure of the 82-kDa enzyme malate synthase G from joint NMR and synchrotron SAXS restraints

Alexander Grishaev · Vitali Tugarinov · Lewis E. Kay · Jill Trewhella · Ad Bax

Received: 14 June 2007 / Accepted: 2 November 2007 / Published online: 16 November 2007

Abstract Determination of the accurate three-dimensional structure of large proteins by NMR remains challenging due to a loss in the density of experimental restraints resulting from the often prerequisite perdeuteration. Solution small-angle scattering, which carries long-range translational information, presents an opportunity to enhance the structural accuracy of derived models when used in combination with global orientational NMR restraints such as residual dipolar couplings (RDCs) and residual chemical shift anisotropies (RCSAs). We have quantified the improvements in accuracy that can be obtained using this strategy for the 82 kDa enzyme Malate Synthase G (MSG), currently the largest single chain protein solved by solution NMR. Joint refinement against NMR and scattering data leads to an improvement in structural accuracy as evidenced by a decrease from ~4.5 to ~3.3 Å of the backbone rmsd between the derived model and the high-resolution X-ray structure, PDB code 1D8C. This improvement results primarily from medium-angle scattering data, which encode the overall molecular shape, rather than the lowest angle data that principally determine the radius of gyration and the maximum particle dimension. The effect of the higher angle data, which are dominated by internal density fluctuations, while beneficial, is also found to be relatively small. Our results demonstrate that joint NMR/SAXS refinement can yield significantly improved accuracy in solution structure determination and will be especially well suited for the study of systems with limited NMR restraints such as large proteins, oligonucleotides, or their complexes.

Keywords NMR · Protein structure · SAXS · RDC · Malate synthase G

Abbreviations

MSG	Malate synthase G
SAXS	Small-angle solution X-ray scattering
RDC	Residual dipolar coupling
RCSA	Residual chemical shift anisotropy
NOE	Nuclear Overhauser enhancement
SVD	Singular value decomposition
R_G	Gyration radius
d_{max}	Maximum particle dimension

Introduction

Small angle X-ray scattering (SAXS) data provide information on macromolecular shape and can be collected under solution conditions that are very similar to those commonly used in macromolecular NMR. It therefore has been long recognized that SAXS data, even at relatively low resolution, provide a powerful complement to structural data derived from NMR (Sunnerhagen et al. 1996;

Electronic supplementary material The online version of this article (doi:10.1007/s10858-007-9211-5) contains supplementary material, which is available to authorized users.

A. Grishaev (✉) · A. Bax
Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Building 5, Bethesda, MD 20892-0520, USA
e-mail: alexanderg@intra.niddk.nih.gov

V. Tugarinov · L. E. Kay
Departments of Medical Genetics, Biochemistry and Chemistry, University of Toronto, Toronto, ON, Canada M5S 1A8

J. Trewhella
School of Molecular and Microbial Bioscience, University of Sydney, Sydney, NSW 2006, Australia

Gabel et al. 2006; Bu et al. 1998; Garcia et al. 2001; Mattinen et al. 2002; Bernado et al. 2005; Marino et al. 2006; Schwieters and Clore 2007). We recently developed a method that allows direct incorporation of solution scattering data in the high-resolution structure determination protocol, leading to a better definition and increased accuracy of both individual domains and multi-domain protein geometry (Grishaev et al. 2005; Wu et al. 2005). The advantages of including long-range translational information provided by the SAXS data are most apparent when complemented by global orientational restraints such as residual dipolar couplings (RDCs) (Tolman et al. 1995; Tjandra and Bax 1997) or residual chemical shift anisotropies (RCSAs) (Lipsitz and Tjandra 2001; Choy et al. 2001; Wu et al. 2001). Higher-angle scattering data, which are sensitive to internal density fluctuations on the length scales of the individual amino acid residues, may offer additional advantages for enhancing the definition of the macromolecular geometry. X-ray scattering data are particularly useful when recorded for larger macromolecules, where they benefit from higher signal-to-noise due their quadratic dependence on molecular volume, and information content that increases with the maximum particle dimension.

In this study, we describe the effects of scattering data, measured over an angular range that corresponds to distances from ~230 to 8 Å, on the structural definition of the 82 kDa protein malate synthase G (MSG). MSG is an enzyme of the glyoxylate pathway that catalyzes the Claisen condensation of glyoxylate and acetyl-coenzyme A (acetyl-CoA) in bacteria to form malate in a magnesium-dependent manner. Crystal structures of MSG in complex with Mg^{2+} and glyoxylate at 2.0 Å resolution (Howard et al. 2000; PDB access code 1D8C) and in a ternary abortive complex with Mg^{2+}, pyruvate and acetyl-CoA at 2.95 Å resolution (Anstrom et al. 2003; PDB access code 1N8I), as well as an NMR model for the apo-MSG (Tugarinov et al. 2005; PDB access code 1Y8B) have been reported. The protein is composed of four domains that include a central TIM barrel β_8/α_8 core, an N-terminal α-helical clasp, an α/β domain, and a C-terminal 5-helix bundle that forms a plug for one of the ends of the β-barrel of the core domain and is connected to it by a flexible linker. The active site of the enzyme is located in the cleft between the C-terminal plug and the TIM-barrel core, while the function for the α/β domain remains unknown. The conformational changes associated with the binding of substrates to MSG are minor (Anstrom et al. 2003) and do not involve significant domain reorientations (Tugarinov and Kay 2003). The structure of the glyoxylate-bound MSG exhibits the best agreement with the NMR-derived RDCs and anisotropic shifts, consistent with its higher resolution, and will be used for structural comparisons throughout this study.

From the NMR point of view, MSG is the largest single-chain protein for which nearly complete backbone resonance assignments and a structural fold have been determined, therefore representing a prime example of state-of-the-art NMR methodology. Its spectral assignments and derivation of experimental restraints benefited from application of advanced TROSY-based pulse sequences (Pervushin et al. 1997; Salzmann et al. 1999; Yang and Kay 1999) and selective deuterium labeling techniques (Tugarinov et al. 2004). In addition to the commonly used NOEs and dihedral angles, its experimental restraints include $N–H^N$ RDCs and backbone carbonyl ^{13}C RCSAs, acquired in dilute liquid crystalline Pf1 phage medium at the combined level of ~1 per residue. The downside of perdeuteration, necessary for obtaining assignable spectra with sufficiently narrow resonance line widths, is that the density of the short-range NOE distance restraints becomes fairly low. In the case of MSG, these restraints only encompass $H^N–H^N$, H^N–methyl and methyl–methyl connectivities, at an overall level of only ~2 NOEs/residue. This low restraint density limits both precision and accuracy at which the structure could be determined, even when supplemented by $N–H^N$ RDC and $^{13}C'$ RCSA data. Comparison of the previous NMR-derived fold of *apo*-MSG with the crystal structure of the glyoxylate-bound enzyme (PDB code 1D8C) shows a ~4.5 Å root-mean-square difference (rmsd) for the backbone atoms, more likely resulting from the sparseness of the NMR data set than from real differences between protein geometries in solution and crystal. Here, we therefore investigate the improvements achievable by combining the NMR restraints with the SAXS data.

X-ray scattering intensity from a macromolecule in solution depends upon the difference in electron density between the solute and the solvent, i.e. the contrast, and is typically expressed as a function of the scattering vector $q = 4\pi(\sin\theta)/\lambda$, where 2θ is the angle between the incident beam of wavelength λ and the elastically scattered radiation. Under conditions of dilute solution, i.e. in the absence of non-specific aggregation or correlations between the macromolecules resulting from weakly attractive or repulsive forces, the scattering profile, $I(q)$ versus q, is proportional to the particle form factor which depends on the macromolecular shape. Fourier transformation of $I(q)$ gives the inter-atomic distance distribution P(r) in the macromolecular volume, weighted by products of the electron density contrast of each volume element with respect to the bulk solvent. As the density of the solvent layer surrounding the protein is slightly higher than that of the bulk solvent, the measured scattering intensity

and the apparent P(r) extracted from it generally also reflect this phenomenon (Svergun et al. 1998).

Materials and methods

Protein sample preparation

Uniformly-[^{15}N, ^{2}H]; Ile δ1-[$^{13}CH_3$]-labeled MSG was prepared as described previously (Tugarinov et al. 2003) and intended originally for NMR data collection. As X-rays are scattered by electrons, isotopic labeling has no effect on the SAXS profile. The NMR sample (0.5 ml volume) was dialyzed overnight into an H_2O buffer at pH 7.1, containing 150 mM sodium chloride, 20 mM sodium phosphate and 5 mM dithiotreitol (DTT). After dialysis, the protein concentration was adjusted to 14 mg/ml. The protein concentration was measured on samples diluted 30-fold in a solution of 6 M GuHCl by UV absorption at 280 nm wavelength using an extinction coefficient calculated from the MSG sequence and free amino acid values by the ExPASy server (Gasteiger et al. 2003). A total of ~120 µl of protein solution was set aside for the SAXS data acquisition; ~90 µl was actually used.

Collection and processing of scattering data

As scattering signal decays by two to three orders of magnitude between near-zero scattering angles and $q \sim 0.3$–0.8 Å^{-1}, high intensity sources are needed for acquisition of precise scattering data within this range. We took advantage of the high flux X-ray beam available at the Stanford Synchrotron Radiation Laboratory (SSRL), a third generation synchrotron source. Beam Line 4–2, designated for biomolecular SAXS, was used for data collection. A ~12 µl polycarbonate container equipped with 25 µm thick mica windows, providing a 1 mm path-length cell, was precisely and reproducibly inserted into the sample holder and maintained at 20°C. Synchrotron X-ray radiation was focused with a bent cylindrical mirror and passed through a double crystal Si(111) monochromator. The 8 keV X-ray beam was incident in a pinhole configuration with the size of the beam spot on the sample of 0.2 × 1.5 mm. A one-dimensional wire position-sensitive proportional detector, filled with 80/20% Xe/CO_2, was used with a sample-to-detector distance of 1.25 m in a highly offset geometry to obtain the q range of 0.027–0.781 Å^{-1}, providing information on distances in the range 230–8 Å. The q-values at the individual detector channels were calibrated using the (100) reflection from a polycrystalline cholesterol myristate sample. Aluminum inserts of variable thickness were used to attenuate the intensity of the incident beam for evaluation of the radiation damage effects. The scattering due to the protein was obtained as a difference between the sample (protein in buffer) and the buffer-only scattered intensity profiles, normalized for the integrated incident beam intensity and scaled for the sample and buffer transmissions monitored on a separate detector channel. Buffer scattering was measured in the same cell immediately following sample measurements. All data were recorded automatically using the standard facilities at beam line 4–2 as series of consecutive 15 s frames. Pearson correlation coefficients to the data from the first frames were monitored as indicators of the sample integrity during experiment. Frames where the correlation coefficients dropped below 0.95 were set to be discarded by the data processing software. Measurement times for individual samples ranged from 7.5 to 10 min. Data reduction and evaluation were done on-site using the software packages SAPOKO, OTOKO and PRIMUS that have been customized for the SSRL (Koch 1990; Konarev et al. 2003).

NMR/SAXS structure refinement

The NMR experimental restraints used during structure refinement were the same as previously reported (Tugarinov et al. 2005). They consist of 1531 H^N–H^N, H^N–methyl and methyl–methyl NOEs, 533 pairs of backbone (ϕ,ψ) dihedral angles derived using the program TALOS (Cornilescu et al. 1999), 35 χ_1 angles of Val residues extracted from the $^{3}J_{C\gamma C}$ and $^{3}J_{C\gamma N}$ couplings, 415 H^N–N RDCs and 300 backbone carbonyl ^{13}C RCSAs observed upon weak alignment by the Pf1 phage liquid crystalline medium. Fixed values for the magnitude and rhombicity of the alignment tensor (−18.5 Hz and 0.45) were used, equal to those in the earlier NMR-only study (Tugarinov et al. 2005). Structure refinement with the SAXS data followed the approach outlined in our earlier study (Grishaev et al. 2005) with a pseudo-energy term coded into the CNS package (Brunger et al. 1998) that acted to minimize the χ^2_{SAXS} value between the experimental SAXS data and the predicted scattering profile. The latter was calculated from the current atomic coordinates via the previously described globbic approximation. Globbic correction factors computed over covalently linked groups of 3–4 heavy atoms were updated at the end of each successive refinement cycle, as in our previous study (Grishaev et al. 2005). A uniform layer of bound solvent with a thickness of 3 Å and a density 10% higher than bulk water was accounted for by an additional correction factor. All structure calculations, including those with SAXS data fitted, used a radius of gyration (R_G) restraint (Kuszewski et al. 1999), with the target value set at 26 Å. This target R_G value was estimated

from hydrodynamic radius (R_h) measurements (in the range of 33–38 Å) calculated from dynamic light-scattering experiments with protein concentrations ranging between 0.2 and 0.4 mM and the same buffer composition as used for NMR data collection (Tugarinov et al. 2005). A molecular radius of gyration was estimated from these measurements as $R_G = 0.774R_h$. The lowest possible value of R_h (33 Å) was used to account for its potential overestimation from slight deviations of the molecule from an ideal spherical shape. The terms for the carbonyl ^{13}C RCSAs (Lipsitz and Tjandra 2001) and the gyration radius restraint were ported into CNS from the Xplor-NIH (Schwieters et al. 2003) version 2.9.4. An effectively rigid-domain Cartesian refinement of MSG against NMR and SAXS data ($q \sim 0.027$–0.300 Å^{-1}) was performed by forcing the coordinates of its four constituent domains (N-terminal domain-residues 4–88; core domain-residues 116–132, 266–295, 334–550; α/β domain-residues 136–261, 317–329; and C-terminal domain-residues 589–722) to be identical to the corresponding coordinates in the 1D8C crystal structure. This was accomplished by using non-crystallographic-symmetry (NCS) restraint terms with force constants of 10,000 kcal Å^{-2}. In order to facilitate repositioning of the individual domains relative to each other, atoms of the side chains belonging to 149 residues situated at the inter-domain interfaces were excluded from such restraints and free to move.

A flexible-domain refinement was also performed using a Cartesian simulated annealing protocol of 40,000 steps of 2 fs with the temperature linearly decreased from 2,000 to 0 K. Multiple structure calculation stages were performed in succession until convergence. In order to establish a "reference point", initial calculations were done with only NMR data and light scattering-based R_G enforced, mirroring the previous NMR-only structure determination. These runs were followed by the refinement with the SAXS data term included, with the calculation done in three separate stages. During the first stage, SAXS data up to q_{max} of 0.22 Å^{-1} were added to the NMR dataset and the R_G restraint. During the second stage of refinement, an empirical pseudo-potential describing backbone-backbone hydrogen bonding (Grishaev and Bax 2004) was added with the H-bonding partners identified in a fully automated manner without any user input. In order to investigate the effect of the order of including these extra restraints on the structural accuracy, calculations were also performed with the switched order of the first two stages, that is, the H-bonding PMF was added during the first stage, and the low-resolution SAXS data during the subsequent stage. Finally, the complete experimental SAXS data set, up to a q_{max} of 0.78 Å^{-1}, was fitted.

Evaluation of the uncertainty in the fitted orientational and translational positioning of MSG domains

The uncertainties in the relative orientations of the four domains that constitute MSG were determined by adding random noise to the input RDC and RCSA data, as described in the Supplementary Information section. The precision of the translational positioning of the peripheral domains of MSG with respect to the protein core of the NMR/SAXS-refined structure was determined by subjecting these domains to random translations of a preset magnitude, followed by evaluation of the χ^2_{SAXS} difference up to $q_{max} = 0.30$ Å^{-1} between the scattering curves calculated from the original and the translated geometries, with relative errors equal to those of the experimental data. Ten such translations that did not include any non-bonded inter-atomic contacts closer than 0.9 Å were used to calculate the average and rms values for each displacement magnitude.

Results and discussion

MSG scattering data are free from radiation damage, inter-particle interference, and aggregation

One of the primary concerns with a synchrotron measurement is potential damage to the macromolecule as a result of the exposure to the intense incident radiation. Since the mechanism of radiation damage involves bond breakage due to highly reactive OH free radicals (Fischetti et al. 2003), usage of agents that capture these species is a well-warranted precaution. Radiation sensitivity concerns are particularly valid for MSG since its enzymatic activity is known to decrease upon exposure to X-rays (Howard et al. 2000). Increased protein aggregation was also observed at the conditions of the previously reported SAXS experiments (Durchschlag and Zipper 1985). We therefore included 5 mM DTT in the buffer and performed data acquisitions at $\sim 2\times$ and $\sim 10\times$ attenuated incident beam intensities. Surprisingly, the MSG sample proved to be very robust under our experimental conditions. The data taken at both attenuation levels showed no differences outside of what was expected from the photon count statistics up to 10 min of exposure, with all individual 15 s frames exhibiting correlation coefficients of >0.99 to the first frames of the experiments. Sample degradation was, however, observed beyond that point at the $2\times$ beam attenuation, accompanied by an increase in the apparent radius of gyration. The scattering data of MSG averaged over 3 exposures of 7.5 min each at $2\times$ beam attenuation on separate samples are shown in Fig. 1a along with the corresponding P(r) curves in Fig. 1b.

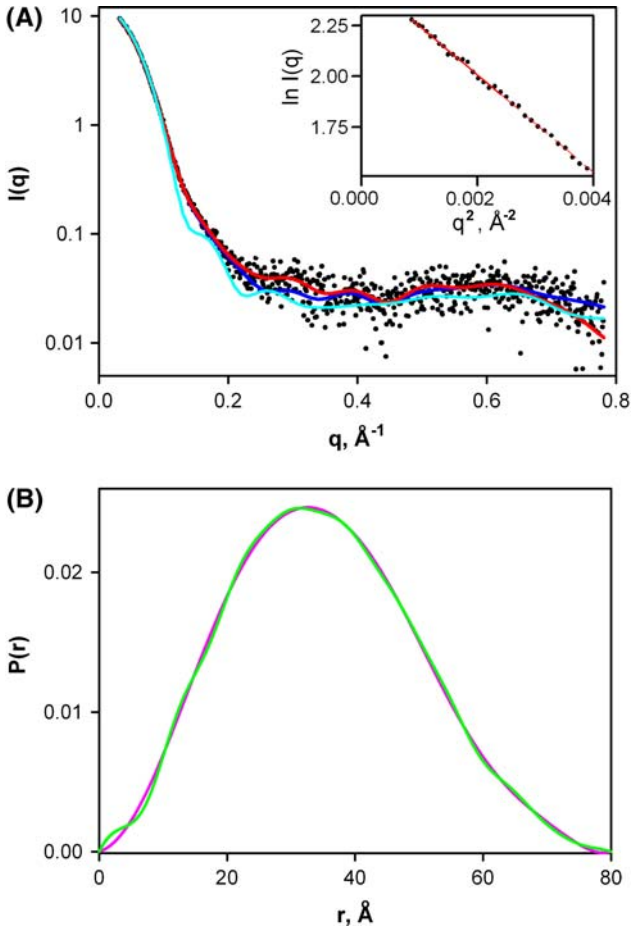

Fig. 1 (**a**) Experimental X-ray scattering data for MSG (black dots) with the curves fitted via CRYSOL (Svergun et al., 1995) from the X-ray structure 1D8C (blue line, $\chi_{SAXS} = 1.010$), the NMR model 1Y8B (cyan line, $\chi_{SAXS} = 3.045$), and the SAXS/NMR structure (red line, $\chi_{SAXS} = 0.974$). Inset: Guinier plot with the solid red line indicating the region of applicability $qR_G < 1.3$). (**b**) The inter-atomic distance distribution curves generated using GNOM (Svergun 1992), fitting data up to 0.220 $Å^{-1}$ (magenta line) and up to 0.781 $Å^{-1}$ (green line)

Inter-particle spatial correlation (interference) effects between proximal macromolecules can complicate scattering data analysis as they lead to modulation of the scattering curve by an unknown structure factor. Such effects are proportional to the number density of molecules in the solution (i.e. concentration dependent) and typically manifest themselves as a decrease of the scattering intensity at q values approaching zero and an associated downward curvature of the data in Guinier plots (ln I(q) vs. q^{-2}). Due to the electrostatic nature of these long-range correlations, addition of salt to the buffer, as done in this study (150 mM NaCl), is often used to minimize their effects. Protein aggregation is the ultimate potential complication of the SAXS measurement, as it can render the data completely uninterpretable. Since the intensity of the scattered radiation increases with the square of the molecular volume, aggregation can be much more pronounced in the SAXS data than it is with NMR, where the higher-MW species become effectively invisible due to their slower rotational diffusion. A signature of aggregation in the scattering data is an upwards curvature at the lowest angles in the Guinier plot. However, due to their opposite effects on the Guinier curve, inter-particle interference and aggregation effects can mask each other at a particular concentration. Therefore, a large number of criteria has to be used to rule out their presence. Some of the criteria we used are R_G-based: Guinier plots showed no deviation from linearity (inset for Fig. 1a); no changes in the R_G were observed when the protein concentration was decreased by a factor of 2; the Guinier-derived R_G (26.2 ± 0.3 Å) agreed well with the previous light scattering measurements (26 ± 1 Å, Tugarinov et al. 2005), and with the second moment of the P(r) distribution obtained from the entire scattering curve (26.8 ± 0.2 Å). The P(r) distribution also showed no evidence of aggregation as the long vector-length tail smoothly decays to zero at ~80 Å (Fig. 1b). Simultaneous extraction of the form and structure factors from the experimental data also indicated a very small magnitude for the inter-particle interference effects within the entire recorded q-range (see the Supplementary information section).

Solution scattering data for MSG are generally consistent with the 1D8C structure

The agreement between the R_G values determined the SAXS data with those calculated from the coordinates of the X-ray and NMR structures (26.2 and 26.3 Å, respectively, no hydration layer) indicates the absence of major conformational changes involving re-distribution of the molecular mass within MSG between the crystal and solution environments. Overall, the entire experimental scattering curve provides an excellent fit ($\chi_{SAXS} = 1.010$) to the previously determined X-ray structure of the glyoxylate-bound MSG (Fig. 1a). The P(r) function for MSG (Fig. 1b), typical for a globular protein, was obtained using the indirect regularized Fourier transform as implemented in the program GNOM (Svergun 1992). The ~80 Å maximum particle dimension (d_{max}) value from the SAXS data agrees well with the values obtained from the coordinates of the X-ray and NMR models (82 and 83 Å, respectively, for 1D8C and 1Y8B structures).

Rigid-domain refinement suggests *apo*-MSG is very close to the ligated X-ray structure

RDCs and $^{13}C'$ RCSAs indicate that relative orientations of the four major domains of MSG closely match those in the

X-ray structure of the glyoxylate-bound MSG, (Tugarinov and Kay 2003). Together with localization of the backbone chemical shift changes to the vicinity of the binding site upon binding of acetyl-CoA, pyruvate and glyoxylate, these results provide strong evidence against major structural rearrangements within MSG upon ligand binding, as well as against differential mobility of its individual domains. However, the RDCs and RCSAs are sensitive only to the orientation of the individual bond vectors, and not to the translational positions of the individual domains. A subtle domain translation would not be inconsistent with the backbone chemical shifts being largely unperturbed upon ligand binding. Therefore, the possibility of a translational shift of the individual domains of MSG in solution with respect to the substrate-bound X-ray structure could not be excluded a priori on the basis of the available NMR data. That holds in particular for the C-terminal domain, due to elevated main chain B-factors for the linker residues 577–588 observed in the X-ray model and the increased mobility of these residues observed from the NMR relaxation data (Tugarinov et al. 2002). An earlier SAXS study of the trimeric malate synthase from baker's yeast (Zipper and Durchschlag 1977) indicated a decrease of the R_G upon substrate binding. The position of the C-terminal domain relative to the core is important because its residue D631 is part of the glyoxylate binding site. Therefore, below we estimate the magnitude of the minimal solution/crystal structural differences taking into account the uncertainties within our experimental data.

To this extent, we performed a rigid four-domain NMR/SAXS data refinement for MSG starting from structures where the positions of the three peripheral domains were randomized relative to the core by extensive unrestrained rigid body dynamics simulations. Such a calculation could have resulted in a structure identical to 1D8C since the four individual domains were kept identical to those in the X-ray model. Following this refinement, the full-length backbone rmsd between the obtained models and the X-ray structure decreases to ~1.4 Å. Superimposition of the core domains of the 1D8C and the NMR/SAXS structure obtained in this manner indicates that N-terminal helix is rotated by only 3.2° and translated by 0.4 Å, the α/β domain is rotated by 2.4° and translated by 0.9 Å, and the C-terminal domain rotated by 3.0° and translated by 1.2 Å with respect to their positions in the X-ray structure. As these apparent rotation angles are only slightly above the estimated uncertainties in the relative domain orientations (see the Supplementary information), they are not likely to be statistically significant.

We have also investigated whether small relative domain translations observed in the rigid-body refinement are statistically significant, given the measurement errors in our scattering data. Evaluation of the difference between the two predicted scattering curves as a function of the magnitude of the domain displacement (see Fig. 2) indicates that translations up to ~1.0 Å, where χ_{SAXS} reaches unity, would not be detectable from data such as ours. The small magnitude of the discrepancy between the scattering curves predicted from the 1D8C X-ray model and rigid-domain NMR/SAXS models that include domain rotations as well as translations ($\chi_{SAXS} = 0.736$) indicates that the residual translational displacements of 0.4–1.2 Å seen in the rigid-domain refinement are dominated by the uncertainties in our data and are not statistically significant. In particular, the position of residue D631 in the rigid-domain NMR/SAXS model is less than 1 Å away from its position in the ligated structure, confirming the absence of major structural rearrangements between the glyoxylate binding sites of the *apo*- and ligand-bound forms of the enzyme. Based on these results we can rule out both rotational and translational global rearrangements of the peripheral domains of *apo*-MSG in solution with respect to those seen in the crystal structure of the glyoxylate-bound enzyme.

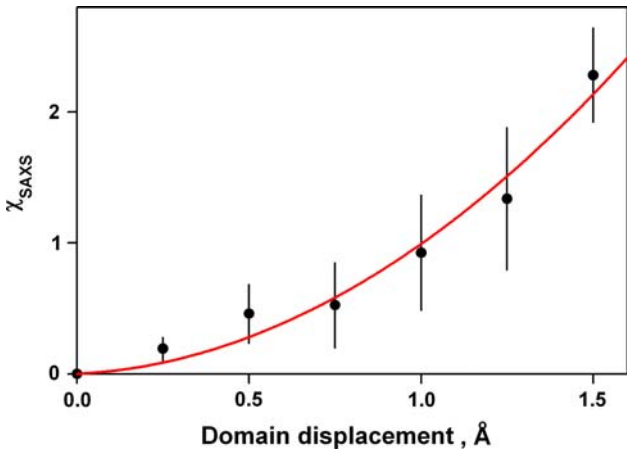

Fig. 2 Discrepancy between the scattering curves predicted from the structures that exhibit a given translation of the peripheral domains with respect to the core. Averages and standard deviations over 10 random displacements of a given magnitude are shown as dots and error bars. Calculation of the χ_{SAXS} was weighted by the q-dependent relative errors that matched those in the experimental SAXS data. As plotted χ_{SAXS} values depict the difference between the scattering curves for the translated and the original geometries, the coordinate plane origin was added as an extra point. The red line corresponds to a quadratic fit of the average χ_{SAXS} values

Joint NMR/SAXS structure refinement improves the solution model of MSG

The results presented above indicate that the crystal structure of the ligand-bound MSG (PDB code 1D8C) is a suitable reference for evaluating the accuracy of the structural models of the enzyme in solution. In order to determine the impact of the SAXS data in the usual

structure refinement, where no fixed models exist for the individual domains, we have performed a flexible-domain structure refinement with both NMR and SAXS data as restraints. Reference refinement runs, with only NMR and light scattering data fitted, produced a family of structures with statistics very similar to those reported for the original NMR structure determination (Table 1). As described in the Methods section, structure refinement with SAXS data was done fitting medium-q data up to q_{max} = 0.22 Å^{-1} first, and then expanding the fitted q-range up to the maximum observed. Thus, our main strategy is adding data of increasing resolution as the calculation progresses. It is worth pointing out that the lowest-angle scattering data up to $q_{max} \sim 0.05$ Å^{-1}, dominated by the R_G of the protein, are largely silent throughout our calculations as they are redundant with respect to the explicit 26 Å R_G restraint from the earlier light scattering data. By itself, the R_G restraint, or, alternatively, the Guinier region scattering data are important for definition of the MSG structure, given the low density of the NOE distance restraints. Its removal from the NMR-only "Model 0" calculation increases the whole-chain rmsd to the 1D8C structure by ~0.4 Å (see Table 1), accompanied by expansion of the protein's size to R_G ~28–29 Å.

As described in the Methods section, the first two stages of the structure calculation beyond the NMR-only model were done in two sequences, with H-bonding PMF introduced first, followed by the low-angle SAXS data, and vice versa. With the latter order, the first round of refinement, which included scattering data up to a q_{max} of 0.220 Å^{-1}, produced a considerable improvement of the structural quality, as manifested by a 0.85 Å decrease of the backbone rmsd to the full-length (residues 3–722) X-ray structure of MSG (PDB code 1D8C). This improvement is to be expected as the intermediate-angle data are dominated by the low-resolution particle shape, which reflects the relative positions of the individual domains. In order to achieve a comparable improvement with NOEs only, ~1,500 additional distance restraints modeled from the 1D8C structure would have to be introduced (V. Tugarinov, unpublished results). A subsequent refinement stage with the H-bonding pseudo-potential resulted in a further ~0.1 Å decrease of the backbone rmsd to 1D8C accompanied by an improved definition of the secondary structure elements. As expected, the resulting structures exhibit good agreement with the SAXS data in the fitted range ($q \sim 0.02$–0.22 Å^{-1}). On the other hand, the SAXS data within the higher resolution range ($q \sim 0.22$–0.78 Å^{-1}), not fitted at this point, can also be used for validation purposes. The comparison of the predicted scattering curves before and after the initial SAXS data fit (Fig. 1 and Table 1), shows a decrease of χ_{SAXS} in the 0.22–0.78 Å^{-1} range from ~1.48 to ~0.93, which also points to an improvement in the accuracy of the solution structure.

Table 1 Structural refinement statistics

Refinement stage	χ_{SAXS} for medium-q/high-q/all[a]	χ_{RDCs} (Hz[b])	$\chi_{\Delta\delta}$ (ppb[b])	rmsd to 1D8C[c]				
				All	Core	N-term	α/β	C-term
Model 0′[d]	12.448/1.599/6.327	2.82 ± 0.09	8.8 ± 0.3	4.92 ± 0.78	4.21 ± 0.48	2.16 ± 0.16	2.59 ± 0.30	3.78 ± 0.42
Model 0[d]	7.547/1.478/4.017	2.86 ± 0.07	8.6 ± 0.2	4.50 ± 0.57	3.71 ± 0.30	1.63 ± 0.15	2.31 ± 0.28	3.49 ± 0.43
Model I[d]	1.154/0.940/1.062	2.76 ± 0.05	9.6 ± 0.2	3.65 ± 0.28	3.33 ± 0.24	2.11 ± 0.17	2.36 ± 0.26	3.08 ± 0.34
Model II[d]	1.155/0.934/1.060	2.70 ± 0.04	10.0 ± 0.2	3.55 ± 0.18	3.08 ± 0.17	1.49 ± 0.13	2.29 ± 0.23	2.97 ± 0.29
ModelIII[d]	1.158/0.944/0.979	3.78 ± 0.06	12.4 ± 0.6	3.31 ± 0.16	3.13 ± 0.15	1.64 ± 0.14	2.23 ± 0.19	3.06 ± 0.26
Model IV[d]	1.081/1.088/0.977	5.03 ± 0.04	22.4 ± 0.4	1.39 ± 0.10	0	0	0	0
X-ray model 1D8C	1.198/1.704/1.033	5.79[e]	22.9[e]			–		

[a] All fits were done with CRYSOL 2.5 (Svergun et al., 1995), using 50 as the maximum order of harmonics and 18 as the Fibonacci grid order. For models 0′, 0, I, II, III and 1D8C, medium-q corresponds to 0.027–0.220 Å^{-1} and high-q to 0.220–0.781 Å^{-1}. For model IV, medium-q corresponds to 0.027–0.300 Å^{-1} and high-q to 0.300–0.781 Å^{-1} reflecting q_{max} of 0.300 Å^{-1} that it was refined against. The medium-q and the entire-range (0.027–0.781 Å^{-1}) data fits were obtained leaving all other adjustable parameters within their default ranges; the high-q data fits were obtained fixing these adjustable parameters at the values best-fitted to the medium-q data

[b] As reported by the CNS program utilities (Brunger et al. 1998)

[c] rmsd values in Å are calculated over the backbone C/N/C$^{\alpha}$ atoms of residues 3–722 (all residues); 116–132,266–295,334–550 (core domain); 3–88 (N-terminal domain); 135–262,296–333 (α/β domain); 589–722 (C-terminal domain); calculated with MolMol v 2.1K (Koradi et al. 1996)

[d] Model 0′ is obtained by fitting NMR-derived restraints only; Model 0 is obtained by fitting NMR data and the light-scattering-derived R_G, matching the previous structural study (Tugarinov et al. 2005); Model I is obtained by adding to the data for Model 0 the X-ray scattering data within the q-range of 0.027–0.220 Å^{-1}; Model II is obtained by adding to the data for Model I the H-bonding pseudo-potential; Model III is obtained by adding to the data of Model II the scattering data within the q-range from 0.220–0.781 Å^{-1} and deposited into the Protein Data Bank, accession number 2JQX; Model IV is obtained by a rigid-body 4-domain refinement with NMR data, R_G, and the scattering data in the q-range of 0.027–0.300 Å^{-1}

[e] Obtained by a joint SVD fit of the N–H^N RDC and ^{13}C′ RCSA data as described in the Section "Materials and methods"

As the accuracy of automated detection of the interacting partners by the H-bonding PMF depends on the accuracy of the input structures, it is possible that its performance will deteriorate as the quality of the input structures deteriorates. In order to investigate the magnitude of this effect, we have switched the order of the first two stages (Model I and Model II in the Table 1) with the results of these calculations reported in the Supplementary Information Table 1. During the first stage of such structure refinement with only the H-bonding PMF added (Model I', Supplementary Table 1), we observe improvement of the backbone geometry of all individual domains of the MSG, as expected from the local action of the potential on linked networks of hydrogen bonds that constitute the secondary structure elements of the individual domains. However, the resulting structures show no improvement of the overall (whole-chain) structural accuracy, which also makes sense as this statistic is dominated by the relative positions of the MSG individual domains. Since the inter-domain contacts in MSG involve only sidechain/sidechain interactions, little improvement is indeed to be expected from the backbone H-bonding PMF. The results of the second stage of this calculation with the intermediate angle scattering data added to the H-bonding PMF (Model II', Supplementary Table 1) are very similar to the results of the calculations with H-bonding PMF turned on first and SAXS data subsequently (Model II, Table 1). This equivalence indicates that the H-bonding PMF is capable of establishing largely correct H-bond pairings even as the input structural quality varies (Model 0 vs. Model I in Table 1).

The final family of structures, generated by fitting the entire experimental range of the scattering data, shows a further 0.24 Å decrease of the backbone rmsd to the 1D8C coordinates. The higher angle data, which in practice require the high brilliance of a synchrotron source, are increasingly influenced by the fluctuations of the internal density of the protein, reflecting smaller length-scale aspects of its architecture. Considering the small number of experimental restraints per residue, the final structures exhibit fairly good Ramachandran statistics as evaluated by PROCHECK (Laskowski et al. 1993), with 80.2% in most favored, 16.1% in allowed, 2.6% in generously allowed and 1.1% in disallowed areas, virtually unchanged throughout refinement. The fit quality of RDCs and $^{13}C'$ RCSAs shows only a slight decrease when more SAXS data are being fitted (see Table 1 and Fig. 3), with the bulk of the effect occurring at the last stage of the refinement when the highest-q data are being included.

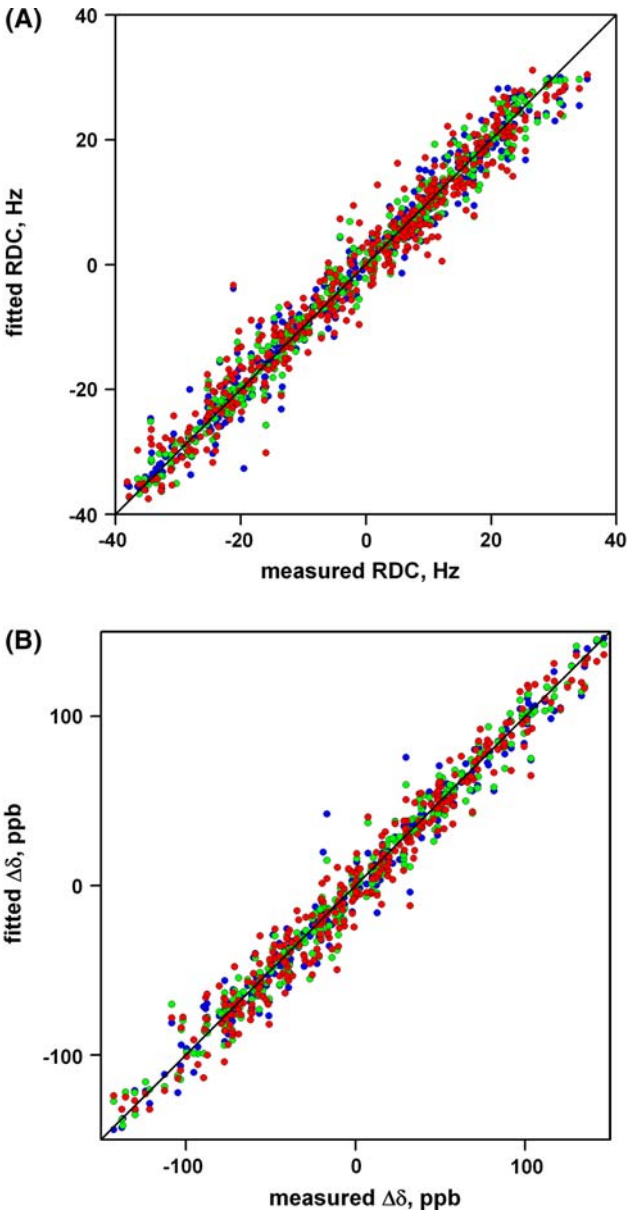

Fig. 3 Quality of the fit for the experimental N-H^N RDCs (panel A) and ^{13}C RCSAs (panel B) is shown for the structures without SAXS data fitted (model 0, blue symbols), with low-q data fitted (model II, green symbols), and with high-q data fitted (model III, red symbols)

Protein representation chosen for the structure refinement is adequate for the problem at hand

Our globbic approximation, which uses small clusters of 3–4 heavy atoms to represent the MSG amino acid sequence, is computationally expensive as it leads to representation of MSG by ~2,500 globs. The computational burden is particularly severe for the very last stage of structure calculation in which the number of fitted SAXS data points is increased by a factor of ~4, with each of the ~400,000 steps of MD calculation taking ~20 s on a Pentium 4-class processor. However, such representation allows very accurate modeling of the scattering data within the fairly wide experimental q-range. The ~15% maximum value of our globbic correction at $q \sim 0.72$ Å^{-1}, where the signal to noise in the data is among the worst, is safely below the ~40% experimental data

uncertainty at this angle. Since the systematic errors remaining after this correction is applied are smaller than the correction itself, their effect throughout the entire calculation remains negligible. The correction for the surface solvent layer involves estimating the value of the bound solvent density, the only adjustable parameter in our approach. Structure calculations done with the bound solvent density set at levels of 1.015 and 1.030 times the bulk solvent density ($\rho_o = 0.334$ e/Å^3) for the shell with a thickness of 3.0 Å showed no discernable structural differences with respect to the calculations done with $1.100\rho_o$, used for the majority of the results reported here.

In this study, we represent the structure of MSG in terms of a single static model, in agreement with the prior NMR data that indicate the absence of significant domain mobility. However, for other systems where domain mobility is possible, an ensemble representation when fitting SAXA data would also be feasible, as recently demonstrated (Bernado et al. 2007) with a caveat that it would further decrease the number of observables vs. experimental degrees of freedom. For a case such as MSG, which is characterized by a very low density of experimental restraints, such treatment is not recommended unless the individual domains are held rigidly and/or additional restraints such as order parameters are used to limit the spread of the structural ensemble (Lindorff-Larsen et al. 2005; Clore and Schwieters 2006).

MSG models obtained via NMR-SAXS refinement exhibit translational shifts of the peripheral domains and improved individual domain definition with respect to the NMR-only models

The final structures determined by refinement against SAXS and NMR data exhibit primarily translational domain rearrangements with respect to the NMR-only structure1Y8B, as evident from an approximately parallel orientation of the major secondary structure elements when the two are superimposed (Fig. 4). The largest difference between the NMR-only and NMR/SAXS-refined structures, which exhibit ca. 4 Å backbone rmsd relative to one another, corresponds to repositioning of the α/β domain with respect to the rest of the protein. The observed translations ($\sim$4–5 Å for the structural elements within the α/β domain and $\sim$3–4 Å for the C-terminal domain) are well outside the estimated $\sim$1 Å translational positioning uncertainty based on the SAXS data fit (Fig. 2). Overall, the MSG structure refined with the SAXS data included shows a substantially lower backbone rmsd to the X-ray reference structure, indicating an improvement of the structural accuracy of the backbone coordinates. As in our previous study (Grishaev et al. 2005), the improvement of the overall agreement between the crystal structure and the NMR/SAXS-refined model is generally accompanied by an improvement of most of the individual domain geometries too (see Table 1), in particular for the core TIM barrel β_8/α_8 domain, the least well defined of the four domains of the enzyme when calculated from the NMR restraints only. The difficulty with the core domain as defined by the NMR data is a direct consequence of the labeling scheme used in the NMR study, where highly deuterated, methyl protonated protein is employed. This eliminates the bulk of the distance restraints that would normally be obtained in the form of NOEs between proximal H^α protons across strands that form parallel β-sheets of the core. Refinement against the low-to-medium-q SAXS data improves the agreement with 1D8C model for the core and C-terminal domains by about 0.4 Å, but worsens the rmsd by 0.05 Å for the α/β domain and by *ca* 0.5 Å for the N-terminal domain. The

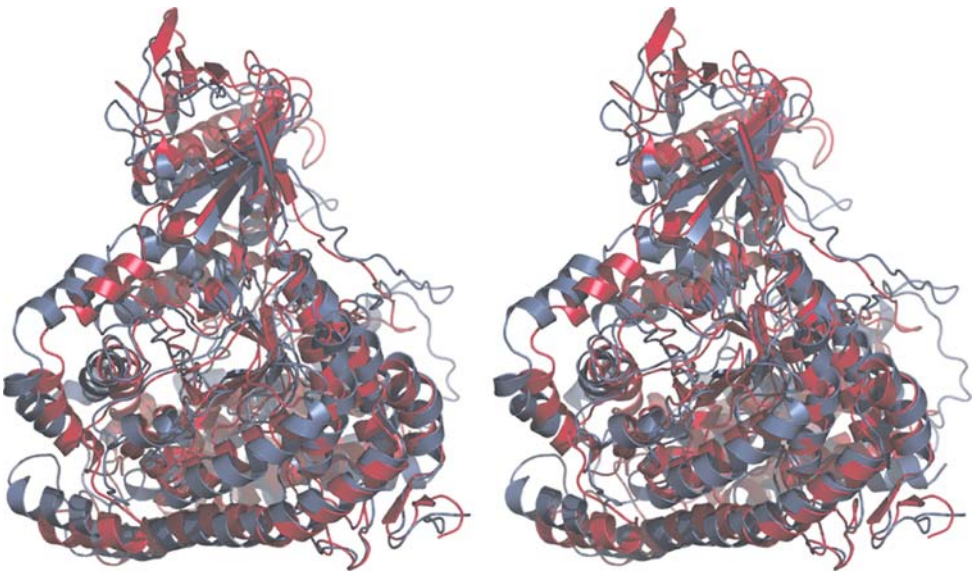

Fig. 4 Stereo representation of the superimposition of the structure of MSG obtained by the joint fit of SAXS and NMR data (red) and the NMR-only model (blue). The figure is generated with PyMOL (DeLano 2002)

increase of the backbone rmsd to 1D8C for the N-terminal domain is not accompanied by worsening of the fit of the corresponding RDCs and ^{13}C shifts. Instead, residues 47–73 that are part of the long helix, which constitutes the bulk of the N-terminal domain, are translated by ~ 2 Å towards the core domain as a result of the SAXS data fit. This translation is largely negated during the subsequent step when the local helical geometry is stiffened by the H-bonding potential (see Table 1). Inclusion of the H-bonding PMF decreases the rmsd to 1D8C for all 4 individual domains with the improvements ranging from 0.62 Å (N-terminal domain) to 0.07 Å (α/β domain). Subsequent refinement against higher-q data results in very small changes in the accuracy of the individual domain geometries, with rmsds to 1D8C slightly increasing for all domains except the α/β.

Medium-angle scattering data have the largest impact on structural accuracy

The refinement procedure followed in this study allows us to evaluate the relative effects of different q-ranges of the scattering data on structural accuracy, a subject for which there is no general consensus at this time. A number of programs are widely used for reconstructing the molecular shapes from both small-angle (Chacon et al. 1998; Svergun 1999; Walther et al. 2000) and higher-angle (Svergun et al. 2001) scattering data. Purely from the number of the fitted data points, restricting ourselves to q_{max} of 0.220 Å^{-1} compared to the complete set with q_{max} of 0.781 Å^{-1} would seem to decrease the information content of such data nearly four-fold, ignoring the larger uncertainty of the higher-q data, which is due to a decrease in the scattering amplitude and the reduced solid angle captured by the one-dimensional detector. In contrast, our results indicate that the bulk of the structural accuracy improvement comes from fitting the medium-q data whereas the effect of the higher-q data, while beneficial, is substantially smaller. A possible explanation is that the higher-resolution, smaller length-scale aspects of the structure are already more adequately defined by the NMR restraints and the all-atom model, and it is the longer length-scale structural aspects for which the contribution of the scattering data is most complementary. This observation is encouraging since acquisition of higher-angle scattering data that are both precise and accurate can be quite challenging. It requires very intense sources which are more likely to cause radiation damage, higher protein concentrations which may make them more prone to aggregation and inter-particle correlations, and higher precision of the relative scaling factors for the sample and buffer data. The latter necessitates a very low lateral drift of the incident beam over the transmitted intensity sensor throughout the measurement, which becomes more difficult to achieve at high X-ray flux values. Thus, it is reassuring that in a refinement scheme such as the one used in the present study, lower-to-medium-q data are sufficient to produce the bulk of the structural improvement that we are observing, particularly considering the increased computational burden of the final refinement stage.

This observation also holds true when obtaining low-resolution molecular shape from the scattering data alone using tools such as DAMMIN (Svergun 1999) or GASBOR (Petoukhov and Svergun 2003), as no benefits of fitting the expanded q-range are observed in this mode (see the Supplementary information section). The advantage of higher-angle scattering data is more apparent when it is used for fold recognition from a structure database, as implemented in the DARA server (Sokolova et al. 2003). In this case, scattering data at q above 0.15 Å^{-1} were necessary for correct identification of the protein (see the Supplementary information section).

Conclusions

In this study we observe a substantial improvement in structural accuracy for an 82 kDa protein when orientational NMR restraints obtained in a weakly aligning liquid crystalline medium are supplemented by SAXS data in a direct refinement procedure. It is clear that all three major components employed in our refinement, global orientational restraints from RDCs and $^{13}C'$ RCSA, long-range translational information from the SAXS data, and local orientational and translational restraints from the H-bonding PMF, are beneficial for a successful structure refinement. Unlike our earlier application (Grishaev et al. 2005), here we observe that even in the absence of changes in such simple parameters as R_G or d_{max}, SAXS data can improve the structural accuracy by influencing the relative positioning of the individual domains within the overall model. The bulk of this improvement does not result from either the inner-most part of the scattering curve, which is sensitive to the overall particle dimension, or from the higher-angle data, which are influenced by variation of the protein's internal density. Instead, the intermediate angle data, which reflect the low-resolution particle shape, are the driving force behind the rearrangement of the MSG domains. A joint analysis of NMR and scattering data indicates that the magnitude of the differences between the MSG domain orientations in solution and in the crystal is only slightly higher than the experimental uncertainty of the relative domain orientations. No statistically significant translational displacements of the peripheral domains with respect to those in the X-ray structure of the glyoxylate-

bound MSG were detected from our scattering data. In this study, immediately analyzable SAXS data were acquired in less than 1 hour, using a small fraction of the NMR sample at a moderate concentration. Data analysis tools that we employed are user-friendly and straightforward even for a novice. We thus expect that the application of the joint NMR/SAXS structure refinement methodology will become routine for other larger macromolecular systems, where acquisition of NMR structural restraints is particularly challenging.

Acknowledgements This work was supported by the Intramural Research Program of the NIDDK, NIH, and the Intramural Antiviral Target Program of the Office of the Director, NIH (A. G. and A. B.) an ARC Federation Fellowship and a grant from the Office of Science (BER), U.S. Department of Energy, Grant No. DE-FG02-05ER64026, NAAR #843 (J. T.). This work utilized facilities supported in part by the National Science Foundation under Agreement No. DMR-0454672. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program. We thank Dr. Hiro Tsuruta (SSRL) for assistance with beamline instrumentation.

References

Anstrom DM, Kallio K, Remington SJ (2003) Structure of the Escherichia coli malate synthase G: Pyruvate: acetyl-coenzyme A abortive ternary complex at 1.95 angstrom resolution. Protein Sci 12:1822–1832

Bernado P, Blanchard L, Timmins P, Marion D, Ruigrok RWH, Blackledge M (2005) A structural model for unfolded proteins from residual dipolar couplings and small-angle X-ray scattering. Proc Natl Acad Sci USA 102:17002–17007

Bernado P, Mylonas V, Petoukhov MV, Blackledge M, Svergun DI (2007) Structural characterization of flexible proteins using small-angle X-ray scattering studies of biological macromolecules in solution. J Am Chem Soc 129:5656–5664

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Cryst D 54:905–921

Bu ZM, Koide S, Engelman DM (1998) A solution SAXS study of Borrelia burgdorferi OspA, a protein containing a single-layer beta-sheet. Protein Sci 7:2681–2683

Chacon P, Moran F, Diaz JF, Pantos E, Andreu JM (1998) Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm. Biophys J 74:2760–2775

Choy WY, Tollinger M, Mueller GA, Kay LE (2001) Direct structure refinement of high molecular weight proteins against residual dipolar couplings and carbonyl chemical shift changes upon alignment: an application to maltose binding protein. J Biomol NMR 21:31–40

Clore GM, Schwieters CD (2006) Concordance of residual dipolar couplings, backbone order parameters and crystallographic B-factors for a small alpha/beta protein: a unified picture of high probability, fast atomic motions in proteins. J Mol Biol 355: 879–886

Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 13:289–302

DeLano WL (2002) The PyMOL molecular graphics system. DeLano Scientific, Palo Alto, CA, USA

Durchschlag H, Zipper P (1985) Post-irradiation inactivation, protection, and repair of the sulfhydryl enzyme malate synthase—effects of formate, superoxide-dismutase, catalase, and dithiothreitol. Radiat Environm Biophys 24:99–111

Fischetti RF, Rodi DJ, Mirza A, Irving TC, Kondrashkina E, Makowski L (2003) High-resolution wide-angle X-ray scattering of protein solutions: effect of beam dose on protein integrity. J Synchr Rad 10:398–404

Gabel F, Simon B, Sattler M (2006) A target function for quaternary structural refinement from small angle scattering and NMR orientational restraints. Eur Biophys J Biophys Lett 35:313–327

Garcia P, Serrano L, Durand D, Rico M, Bruix M (2001) NMR and SAXS characterization of the denatured state of the chemotactic protein CheY: implications for protein folding initiation. Protein Sci 10:1100–1112

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPaSy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31:3784–3788

Grishaev A, Bax A (2004) An empirical backbone–backbone hydrogen-bonding potential in proteins and its applications to NMR structure refinement and validation. J Am Chem Soc 126:7281–7292

Grishaev A, Wu J, Trewhella J, Bax A (2005) Refinement of multidomain protein structures by combination of solution small-angle X-ray scattering and NMR data. J Am Chem Soc 127:16621–16628

Howard BR, Endrizzi JA, Remington SJ (2000) Crystal structure of Escherichia coli malate synthase G complexed with magnesium and glyoxylate at 2.0 angstrom resolution: mechanistic implications. Biochemistry 39:3156–3168

Koch MHJ (1990) Otoko, program package, release 01–90

Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI (2003) Primus: a Windows PC-based system for small-angle scattering data analysis. J Appl Cryst 36:1277–1282

Koradi R, Billeter M, Wuthrich K (1996) MolMol: A program for display and analysis of macromolecular structures. J Mol Graph 14:51–55

Kuszewski J, Gronenborn AM, Clore GM (1999) Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration. J Am Chem Soc 121:2337–2338

Laskowski RA, Macarthur MW, Moss DS, Thornton JM (1993) Procheck—a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283–291

Lindorff-Larsen K, Best RB, DePristo MA, Dobson CM, Vendruscolo M (2005) Simultaneous determination of protein structure and dynamics. Nature 433:128–132

Lipsitz RS, Tjandra N (2001) Carbonyl CSA restraints from solution NMR for protein structure refinement. J Am Chem Soc 123:11065–11066

Marino M, Zou PJ, Svergun D, Garcia P, Edlich C, Simon B, Wilmanns M, Muhle-Gol C, Mayans O (2006) The Ig doublet Z1Z2: a model system for the hybrid analysis of conformational dynamics in Ig tandems from titin. Structure 14:1437–1447

Mattinen ML, Paakkonen K, Ikonen T, Craven J, Drakenberg T, Serimaa R, Waltho J, Annila A (2002) Quaternary structure built from subunits combining NMR and small-angle X-ray scattering data. Biophys J 83:1177–1183

Pervushin K, Riek R, Wider G, Wuthrich K (1997) Attenuated T-2 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 94:12366–12371

Petoukhov MV, Svergun DI (2003) New methods for domain structure determination of proteins from solution scattering data. J Appl Cryst 36:540–544

Salzmann M, Wider G, Pervushin K, Senn H, Wuthrich K (1999) TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. J Am Chem Soc 121:844–848

Schwieters C, Clore G (2007) A physical picture of atomic motions within the Dickerson DNA dodecamer in solution derived from joint ensemble refinement against NMR and large-angle X-ray scattering data. Biochemistry 46:1152–1166

Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The XPLOR-NIH NMR molecular structure determination package. J Magn Reson 160:65–73

Sokolova AV, Volkov VV, Svergun DI (2003) Prototype of a database for rapid protein classification based on solution scattering data. J Appl Cryst 36:865–868

Sunnerhagen M, Olah GA, Stenflo J, Forsen S, Drakenberg T, Trewhella J (1996) The relative orientation of Gla and EGF domains in coagulation factor X is altered by Ca^{2+} binding to the first EGF domain. A combined NMR small angle X-ray scattering study. Biochemistry 35:11547–11559

Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Cryst 25:495–503

Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys J 76:2879–2886

Svergun DI, Petoukhov MV, Koch MHJ (2001) Determination of domain structure of proteins from X-ray solution scattering. Biophys J 80:2946–2953

Svergun DI, Richard S, Koch MHJ, Sayers Z, Kuprin S, Zaccai G (1998) Protein hydration in solution: experimental observation by X-ray and neutron scattering. Proc Natl Acad Sci USA 95:2267–2272

Tjandra N, Bax A (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science 278:1111–1114

Tolman JR, Flanagan JM, Kennedy MA, Prestegard JH (1995) Nuclear magnetic dipole interactions in field-oriented proteins—information for structure determination in solution. Proc Natl Acad Sci USA 92:9279–9283

Tugarinov V, Choy WY, Orekhov VY, Kay LE (2005) Solution NMR-derived global fold of a monomeric 82-kDa enzyme. Proc Natl Acad Sci USA 102:622–627

Tugarinov V, Hwang PM, Kay LE (2004) Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. Annu Rev Biochem 73:107–146

Tugarinov V, Hwang PM, Ollerenshaw JE, Kay LE (2003) Cross-correlated relaxation enhanced 1H–13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. J Am Chem Soc 125:10420–10428

Tugarinov V, Kay LE (2003) Quantitative NMR studies of high molecular weight proteins: application to domain orientation and ligand binding in the 723 residue enzyme malate synthase G. J Mol Biol 327:1121–1133

Tugarinov V, Muhandiram R, Ayed A, Kay LE (2002) Four-dimensional NMR spectroscopy of a 723-residue protein: chemical shift assignments and secondary structure of malate synthase G. J Am Chem Soc 124:10025–10035

Walther D, Cohen FE, Doniach S (2000) Reconstruction of low-resolution three-dimensional density maps from one-dimensional small-angle X-ray solution scattering data for biomolecules. J Appl Cryst 33:350–363

Wu ZR, Delaglio F, Wyatt K, Wistow G, Bax A (2005) Solution structure of gamma s-crystallin by molecular fragment replacement NMR. Protein Sci 14:3101–3114

Wu ZR, Tjandra N, Bax A (2001) P-31 chemical shift anisotropy as an aid in determining nucleic acid structure in liquid crystals. J Am Chem Soc 123:3617–3618

Yang DW, Kay LE (1999) Trosy triple-resonance four-dimensional NMR spectroscopy of a 46 ns tumbling protein. J Am Chem Soc 121:2571–2575

Zipper P, Durchschlag H (1977) Small-angle X-ray studies on malate synthase from bakers-yeast. Biochem Biophys Res Commun 75:394–400