NMR characterization of an engineered domain fusion between maltose binding protein and TEM1 β-lactamase provides insight into its structure and allosteric mechanism

Chapman M. Wright,1 Ananya Majumdar,2 Joel R. Tolman,3 and Marc Ostermeier1*

1 Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, Maryland 21218
2 Biomolecular NMR Center, Johns Hopkins University, Baltimore, Maryland 21218
3 Department of Chemistry, Johns Hopkins University, Baltimore, Maryland 21218

INTRODUCTION

The complex biology of the cell is mediated by specific interactions between proteins and other biomolecules. For example, allosteric proteins use cellular signals to regulate activities such as gene transcription, metabolism, and apoptosis. The design of novel allosteric proteins whose activity can be under the control of the designer is an attractive goal for applications such as biosensing, gene regulation, and therapeutics.1 One method for creating novel allosteric proteins is to combine the functions of two independent proteins into a hybrid protein such that their functions are coupled.2 We have previously demonstrated this approach by fusing TEM-1 β-lactamase (BLA) and E. coli maltose binding protein (MBP) to create maltose-activated BLAs.3,4 RG13 is one such engineered allosteric enzyme whose level of BLA activity increases in the presence of maltose.3 RG13 was identified from a combinatorial library comprised of random circularly permuted variants of the bla gene that were randomly inserted into the gene encoding MBP. The BLA domain of RG13 (cpBLA) has the original N- and C- termini joined by a GSGGG peptide linker, and the cpBLA domain is circularly permuted between residues 226 and 227. These two residues serve as the sites of connection to the MBP domain. The cpBLA domain is inserted into MBP in place of MBP residue 317.

Purified RG13 has essentially wild-type BLA activity in the presence of saturating amounts of maltose, but its catalytic efficiency is reduced 25-fold in the absence of maltose.3,5 This implies that the overall structure of cpBLA is similar to BLA in the presence of maltose but perturbed in the absence of maltose. In the presence of saturating amounts of substrate, RG13 has wild-type MBP affinity for maltose (Kd = 1 μM), implying that the overall structure of the MBP domain is similar to wild-type MBP. Affinity for maltose is compromised about fivefold in the absence of substrate, which indicates that maltose affects both substrate binding and the catalytic step.

The binding of maltose to MBP induces a 35° conformational change about the hinge region.6,7 Biochemical and biophysical studies of RG13 suggest that its MBP domain undergoes a similar conformational change and that this conformational change is required for the allosteric effect.8 For example, sugars that induce a large (>30°) conformational change in MBP (such as maltose and maltotriose) cause a 25-fold increase in RG13 catalytic activity, whereas β-cyclodextrin, which shifts MBP only 10°,
has minor effects on activity (a twofold increase). In addition, mutations in the hinge region of MBP that are known to induce partially closed states in the apo-form of MBP decrease the allosteric effect in RG13 by raising the catalytic activity of the maltose-unbound state.

RG13 was identified from a combinatorial library and not designed, so the reasons for the reduced catalytic activity of cpBLA in the absence of maltose are not known. However, biochemical studies have provided some clues. As discussed previously, we know that substrate binding and the catalytic steps are compromised. Penicillin and cephalosporin hydrolysis by BLA is well-known to follow Scheme 1, where E, S, ES, EAc, and P represent enzyme, substrate, noncovalent enzyme-substrate complex, covalent enzyme-substrate acyl intermediate and hydrolyzed product, respectively. Kinetic studies of RG13 indicate that maltose modulates the rate of β-lactam hydrolysis by modulating the relative values of $k_2$ and $k_3$. Most likely the rate of deacylation is compromised in the absence of maltose and the binding of maltose relieves whatever constraints are limiting deacylation. There is good agreement that a water molecule activated by E166 of BLA hydrolyzes the EAc intermediate during catalysis. This suggests that a displacement of BLA-E166 of RG13 in the absence of maltose may be the cause of RG13’s reduced hydrolysis rate.

Studies on the switching mechanism of RG13 would greatly benefit from structural data of RG13 in the presence and absence of maltose. However, crystals of RG13 have proved difficult to obtain. In addition, if the maltose-induced changes are subtle, these may not be readily apparent in a crystallographic study. Thus, we have explored the use of protein NMR to characterize RG13.

Full structure determination using NMR is normally limited to proteins of <40 kDa, which is well below the size of RG13 (72 kDa). However, substantially larger proteins can be studied by exploiting recent methodological advances in NMR. NMR studies are now feasible for proteins in the 100 kDa range using ultra-high magnetic fields, special isotope labeling schemes, and optimized pulse sequences. An NMR study of RG13 is attractive because chemical shift assignments of BLA and MBP with and without ligand are published. Based on RG13’s conservation of near wild-type affinity for maltose and near wild-type BLA activity in the presence maltose, we expected that the structure of the two domains in RG13 would be substantially conserved from MBP and BLA. Thus, many of the chemical assignments may be simply transferred to our RG13 data from the published chemical shift assignments for the individual proteins. Spectral differences between that of RG13 and the parental proteins will highlight regions of the domains in RG13 that differ from MBP and BLA. RG13 spectral differences between the presence and absence of maltose will highlight regions that may be important for the allosteric mechanism, generating hypotheses for future studies. An understanding of the allosteric mechanism of RG13 would aid in the design of synthetic allosteric enzymes for applications and may provide insight into natural allosteric mechanisms.

MATERIALS AND METHODS

Materials

All chemicals used were purchased from Sigma (St. Louis, MO) unless otherwise noted. Maltose was purchased from Fisher Scientific (Pittsburg, PA). NMR tubes were purchased from Shigemi (Allison Park, PA).

Expression and purification of RG13

Competent BL21 E. coli cells were purchased from Novagen (La Jolla, CA). The expression plasmid was pDIM-C8, which contains a tac promoter and chloramphenicol antibiotic resistance. For RG13 expression in M9/D2O minimal media, the method of Tugarinov et al. was followed with slight variations. A 10 mL overnight culture of BL21 cells harboring the pDIM-C8 plasmid encoding the RG13 gene was grown in LB rich media at 37°C. The cells were then pelleted and washed with 10 mL of 1 × M9 salts in H2O and pelleted again. The washed pellet was resuspended with 10 mL of 1 × M9 salts and the culture to a final concentration of 1 mM IPTG was added to induce RG13. M9/D2O minimal media was 1×, containing 0.6. At this point, IPTG was added to the culture to a final concentration of 1 mM to induce RG13. The 500 mL culture was switched to a 25°C shaker for

\[
E + S \rightleftharpoons ES \rightarrow EAc \rightarrow E + P
\]

**Scheme 1**

Kinetic scheme of penicillin and cephalosporin hydrolysis by BLA.
18–20 hrs. RG13 expression in M9/D2O minimal media took a total of between 48 and 52 hrs to complete.

After expression, the cells were harvested and resuspended in column buffer (50 mM phosphate buffer, pH 7.0, containing 250 mM NaCl) and lysed using a French Press. Cleared lysates were passed over an amylose column (New England Biolabs, Ipswich, MA) equilibrated with column buffer. The resin was washed with 10 mL of column buffer and then RG13 was eluted in 3 mL of elution buffer (50 mM phosphate buffer, containing 250 mM NaCl and 10 mM maltose). The purity of RG13 was >95% as judged by SDS-PAGE. Pure RG13 was then dialyzed against 10 L of NMR buffer (25 mM phosphate buffer, pH 7.0) to remove salts and maltose. NaN₃ was added to 0.05% to a concentrated RG13 sample (250–300 μM), which was stored at 4°C.

NMR spectral acquisition

D₂O was added to 8% to a 200 μM sample of RG13. The 350 μL sample was then loaded into a Shigemi tube. Data were acquired on an 800 MHz (1H) Varian INOVA spectrometer equipped with a cryogenic probe and z-axis gradients, at 25°C. A sensitivity enhanced TROSY sequence14 was used, with 16 scans/FID and an interscan delay of 1.9 s. Spectral widths of 16 kHz (1H) and 2836 Hz (15N) were employed, with an acquisition time of 60 ms in both 1H and 15N dimensions. The total data collection time was 3.25 hours per spectrum. Quadrature detection in the 15N dimension was achieved using the echo-antiecho method.15 The data were apodized with 60° phase-shifted cosine squared bell window functions in both dimensions and zero-filled to a final digital resolution of 4 Hz/pt in 1H and 1.4 Hz/pt in 15N. Data were processed and analyzed using NMRRpipe and NMRDraw software.

RESULTS AND DISCUSSION

We sought to acquire NMR spectra of RG13 in the presence of absence of maltose and compare these resonances to those of BLA and MBP. This would provide tentative backbone assignments for RG13. Spectral differences between RG13 and the parental proteins and differences between the absence and presence of maltose would highlight regions of interest for future studies. High-resolution spectra of large proteins such as RG13 (72 kDa) can be difficult to obtain due to significant peak broadening. However, the use of transverse relaxation optimized spectroscopy (TROSY) in the presence of deuterium improves spectral resolution though peak sharpening.16 Furthermore, TEM-1 BLA has been described as one of the most ordered proteins studied by solution-state NMR,11 giving us encouragement that informative spectra might be acquired from the BLA domain of RG13. MBP is also amenable to solution state NMR investigations as multiple NMR studies of MBP in its "open" or "closed" states have been published.7,12

Spectral acquisition

We acquired an initial HSQC spectrum of a 15N labeled RG13 to determine if further characterization was merited (data not shown). This spectrum indicated that RG13 was stable during data acquisition. However, broad resonances stemming from this 72 kDa protein led to poor signal to noise and severe spectral overlap, preventing the data from being usable for further analysis. Thus, a deuterated form of RG13 was produced using a protocol described by Tugarinov et al.1 The deuterium incorporation using the M9/D₂O protocol was estimated at 60–70% and an 1H-15N TROSY-HSQC spectrum was acquired, resulting in substantial improvement in the RG13 spectra. Approximately 58% of the backbone amides of RG13 were accounted for in these spectra using NMRview17 to determine the peak count. To increase the deuteration level and sharpen the peaks even further, RG13 was produced in M9/D₂O media using deuterated glucose as the sole carbon source.13 Approximately 90% deuteration was achieved and the TROSY spectrum in the absence of maltose improved with a peak count of ~70% (see Fig. 1). The addition of maltose to the same NMR tube resulted in a spectra with many overlapping peaks and, importantly, many shifted peaks (see Fig. 1). This is strong evidence that the two spectra are aligned correctly and that the resonance changes observed after the addition of maltose are due to conformational changes within RG13.

Tentative backbone assignments of RG13

The locations of peaks in the two RG13 spectra were each compared to published chemical shift data of BLA11 (BMRB 6357) and MBP12 (BMRB 4986 ligand free and BMRB 4987 containing maltotriose). A spectrum in the presence of maltotriose was used because peak assignments for a spectrum in the presence of maltose are not present in the BMRB database. Binding of maltose and maltotriose produces almost identical structural changes in MBP except for a few changes locally around the sugar-binding site.18

The 15N and 1H chemical shifts of BLA and MBP were used to locate residues within RG13 that were in the same chemical environment as in the wild-type protein. Peaks and their corresponding amino acid residue in RG13 were labeled as "assigned" if a peak in the RG13 spectrum appeared at the same location as in the published peak assignments of BLA and MBP. A tolerance of ±0.2 ppm on the 15N axis and ±0.05 ppm on the 1H axis was allowed in this assignment. This tolerance was chosen as roughly the radius of a peak on both the 15N and 1H axes. If no peak was present at the chemical shifts...
of an assigned peak in BLA or MBP, that amino acid residue in RG13 was classified as "unassigned." Unassigned peaks may result from the amino acid being in a different chemical environment and thus in a different location in the spectrum. Indeed, some peaks in the spectrum RG13 did not have a corresponding peak in spectra of MBP or BLA. Unassigned peaks can also result from the fact that our peak count is 70%.

Of the 639 residues of RG13, 92% (587 with maltose and 585 without maltose) had published peak assignments in the spectra of BLA\textsuperscript{11} and MBP\textsuperscript{12}. Of those potentially assignable residues, 70% (411 with maltose and 404 without maltose) were assigned by our analysis (i.e., 64% of all residues in RG13 and 91% of the peaks in RG13's spectra) (Table 1). The results of comparisons between the two RG13 spectra resonances and those of BLA and MBP can be seen in Figures 2 and 3. Figure 2 highlights all unassigned residues. This includes residues that are unassigned in the NMR studies of BLA\textsuperscript{11} and MBP\textsuperscript{12} (i.e., residues we could not possibility assign by the methods used here) as well as additional residues unassigned in RG13 because a peak at the expected resonance is not observed. In Figure 3, the unassigned residues that occur in both RG13 structures are stripped away, so that only the maltose-dependent differences in unassigned residues can be seen. Unassigned residues are highlighted in these figures because they represent locations at which the structure in RG13 may be different than the structure of MBP and BLA. Since RG13 has properties that MBP and BLA lack (i.e., allosteric), such conformational differences may be mechanistically important for these properties.

Peak misassignments are possible in our analysis. This would result from the serendipitous change in chemical environment of two residues in which the peak for one residue in RG13 has moved into the location previously occupied by a different residue in MBP and BLA. However, given that there is a one-to-one correspondence for 91% of the peaks in the spectra of RG13 and the spectra of the two parental proteins, we are confident that the majority of peak assignments are correct; therefore, general analysis of structural difference between RG13 and the parental proteins and of the affect of maltose on RG13 is warranted. Further confidence in the validity of our assignment process comes from the agreement of our analysis with specific, reasonable expectations as discussed in the following sections.

**Analysis of the MBP domain of RG13**

The pattern of unassigned residues in the MBP domain of RG13 appear randomly dispersed throughout the structure [Fig. 2(C,D)], which suggests that most unassigned residues arise due to experimental limitations. This, combined with RG13's high affinity for maltose, lead us to conclude that the structure of the MBP domain in RG13 is largely conserved from MBP. Unassigned residues rarely appear sequentially in groups of three or more. One exception makes sense from the fusion geometry of the two domains. The cpBLA domain is inserted in place of MBP residue 317, which is in a helix of MBP. Thus, one would expect that the residues near the insertion site would be unassigned in our analysis because the insertion would disrupt the structure and change the chemical environment of these residues. The residues of this helix were indeed unassigned in our analysis both in the presence and absence of maltose as expected [Fig. 2(C,D)]. The overall slightly larger number of unassigned residues in the presence of maltose (compared to the absence; see Table 1) may arise from the fact that we are comparing a maltose-bound RG13 spectra with a maltotriose-bound MBP structure. The conformations of MBP residues Glu44, Trp62, Arg66, Glu153, and Tyr341 in the sugar-binding site are known to differ depending on whether maltose or maltotriose is bound.\textsuperscript{18} These residues are unassigned in RG13, but only in the presence of maltose—lending further credence to our assignment process.

The fraction of residues in the MBP domain that are assigned in only one of the two states (20% with maltose and 18% without maltose) is greater than the corresponding fraction in the cpBLA domain (4% with maltose).
ose and 12% without maltose). The reason for the higher fraction in the MBP domain is not clear. Part of the reason could be that the spectra are interpreted by comparison to two different datasets of the parental MBP protein. The strength of NMR signals for a particular residue will vary with its chemical environment and maltose-binding changes the environment of many residues in the MBP domain. Hence, many residues may have a peak present in one state and not the other because of experimental limitations on the number of peaks observed. The fact that such maltose-dependent unassigned residues are scattered throughout the MBP domain [Fig. 3(CD)] supports this argument. This effect would be expected to be less prevalent in the cpBLA domain if maltose affected the chemical environment of fewer residues in the cpBLA domain.

Analysis of the cpBLA domain of RG13

Our analysis of the pattern of assigned and unassigned residues of the cpBLA leads to the following observations and conclusions.

1. There are more assigned peaks in the presence of maltose than in the absence (75% vs. 67%; Table I). This is consistent with the expectation that the cpBLA domain of RG13 in the absence of maltose is in a more perturbed state relative to BLA.

2. Near the fusion site to MBP, many residues in the cpBLA domain are perturbed relative to BLA, particularly without maltose [Fig. 2(A)]. For example, residues in the α-helix containing residues BLA-Leu220 through BLA-Leu225 were unassigned both in the presence and absence of maltose. The occurrence of many unassigned residues in this region is consistent with the expectation that fusion to MBP will locally alter the structure of the BLA domain in the vicinity of the fusion site.

3. Whereas residues that have unassigned peaks both in the presence and absence of maltose appear randomly scattered throughout the structure [Fig. 2(A,B)], unassigned peaks unique to the spectra without maltose tend to cluster in certain regions [Fig. 3(A)]. These regions include the original C-terminal α-helix of BLA, the segment of the cpBLA domain fused to the
N-terminal fragment of MBP and the area near the disulfide bond between BLA-C77 and BLA-C123. Such patterning makes stochastic explanations for maltose-dependent difference in unassigned residues unlikely. Thus, these regions may hold importance for understanding how the cpBLA domain is compromised in activity in the absence of maltose. For example, the segment of the cpBLA domain fused to the N-terminus of MBP is relatively undisturbed in the presence of maltose [Fig. 2(B)], but disturbed in the absence of maltose (i.e., there are many unassigned residues in the absence of maltose) [Fig. 2(A)]. Potentially, conformational change propagation through this segment facilitates changes in the cpBLA active site upon maltose binding. The area around a disulfide bond between BLA-C77 and BLA-C123 is another region of the cpBLA domain of RG13 that was perturbed only in the absence of maltose. This region is proximal to the active site of BLA; however, removal of these cysteines by mutagenesis does not affect the enzymatic or allosteric properties of RG13, suggesting that this region may not hold much importance for the allosteric mechanism.

4. The active site is not grossly perturbed either in the absence or presence of maltose. Among key active site residues, BLA-S70 is unassigned in the published BLA spectra and BLA-S130 is unassigned in both spectra’s of RG13. Thus, little can be said about the chemical environment of these serines. However, both BLA-K73 and BLA-K234 are assigned in both spectra of RG13. This, taken together with the fact that many residues adjacent to these active site residues are assigned lead us to conclude that the active site is not grossly perturbed either in the presence or absence of maltose. This is in line with kinetic data on the rate at which RG13 hydrolyzes a β-lactam substrate. In the presence of maltose, enzymatic activity is near wild-type BLA levels and thus the active site should be relatively unperturbed. In the absence of maltose RG13’s catalytic activity is only 4% of that in the presence of maltose but is still substantial compared to many enzymes (kcat/Km = 0.4 μM−1 s−1). One would not expect a large perturbation in the active site of RG13 in the absence of maltose based on this high catalytic activity and one is not observed.

Figure 3
The solely maltose-dependent differences in unassigned residues in RG13. Residues unassigned only in the absence of maltose (A, C) or only in the presence of maltose (B, D) are shown in red. Unassigned residues were traced onto the structures of BLA (A, B) and MBP in the absence (C) and presence (D) of maltotriose. The BLA structures (A, B) have the active site BLA-E166 represented as a stick model, as well as the MBP fusion sites and GSGGG linker labeled. The MBP structures (C, D) have the cpBLA domain insertion site featured (MBP317). In panel D, maltotriose is displayed in stick form in orange. The PDB IDs are 1BTL (A,B), 1JW4 (C) and 3MBP (D).
5. The active site of the cpBLA domain has more unassigned residues in the absence of maltose [Fig. 4(A)] than in the presence [Fig. 4(B)] and thus is more perturbed in the absence of maltose relative to BLA. In particular, BLA-E166 is unassigned only in the absence of maltose. In the presence of maltose, a peak is observed at the expected resonance of BLA-E166 [Fig. 4(C)]. The maltose-dependent difference in the chemical environment of BLA-E166 may reflect changes in the active site responsible for reduced activity in the absence of maltose. This provides evidence that BLA-E166 is properly positioned in the presence of maltose, but is displaced in its absence. There is good agreement on a role of BLA-E166 in activating water to deacylate the acyl-enzyme intermediate.9 Thus, a displaced BLA-E166 would hamper the deacylation step of β-lactam hydrolysis during RG13 turnover in the absence of maltose and explains why the activity is down, but not completely off. Our previous kinetic analysis of the rates of acylation and deacylation in RG13 led us to suggest that in the absence of maltose BLA-E166 may be subtly displaced resulting in a slower deacylation rate and that maltose-induced conformational changes alleviated this defect.8 RG13’s NMR spectra suggest that maltose-dependent changes in the conformation of the MBP domain somehow alleviate the perturbation in the BLA domain and allow BLA-E166 to regain its wild-type conformation.

CONCLUSIONS

Using advanced NMR techniques combined with data from previous NMR studies of MBP12 and BLA,11 we have gained insight into the conformations of RG13, an engineered allostERIC enzyme consisting of a fusion of MBP and BLA. We identified regions of RG13 that are structurally conserved relative to MBP and BLA and regions that may undergo maltose-dependent conformational changes. These studies support previous kinetic studies8 implicating perturbations of E166 in the cpBLA domain in the maltose-dependent differences in catalytic activity. Our results highlight the advantage of using NMR to study RG13 and the power of NMR to report on very small changes in conformation—the kind expected for the maltose-dependent differences in enzyme activity found in RG13.

REFERENCES

4. Guntas G, Mansell TJ, Kim JR, Ostermeier M. Directed evolution of protein switches and their application to the creation of