BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

# Structure of MMACHC Reveals an Arginine-Rich Pocket and a Domain-Swapped Dimer for Its $B_{12}$ Processing Function

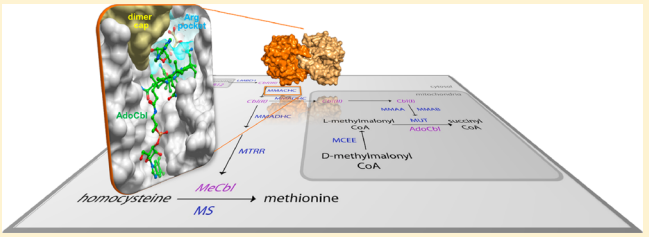
D. Sean Froese,[†,⊥] Tobias Krojer,[†,⊥] Xuchu Wu,[§] Roshi Shrestha,[†] Wasim Kiyani,[†] Frank von Delft,[†] Roy A. Gravel,[§] Udo Oppermann,[†,‡] and Wyatt W. Yue*[,†]

[†]Structural Genomics Consortium and [‡]Botnar Research Centre, NIHR Oxford Biomedical Research Unit, University of Oxford, Oxford, U.K.

[§]Department of Biochemistry & Molecular Biology, University of Calgary, Calgary, Canada

Supporting Information

**ABSTRACT:** Defects in the MMACHC gene represent the most common disorder of cobalamin (Cbl) metabolism, affecting synthesis of the enzyme cofactors adenosyl-Cbl and methyl-Cbl. The encoded MMACHC protein binds intracellular Cbl derivatives with different upper axial ligands and exhibits flavin mononucleotide (FMN)-dependent decyanase activity toward cyano-Cbl as well as glutathione (GSH)-dependent dealkylase activity toward alkyl-Cbls. We determined the structure of human MMACHC·adenosyl-Cbl complex, revealing a tailor-made nitroreductase scaffold which binds adenosyl-Cbl in a "base-off, five-coordinate" configuration for catalysis. We further identified an arginine-rich pocket close to the Cbl binding site responsible for GSH binding and dealkylation activity. Mutation of these highly conserved arginines, including a replication of the prevalent MMACHC missense mutation, Arg161Gln, disrupts GSH binding and dealkylation. We further showed that two Cbl-binding monomers dimerize to mediate the reciprocal exchange of a conserved "PNRRP" loop from both subunits, serving as a protein cap for the upper axial ligand *in trans* and required for proper dealkylation activity. Our dimeric structure is supported by solution studies, where dimerization is triggered upon binding its substrate adenosyl-Cbl or cofactor FMN. Together our data provide a structural framework to understanding catalytic function and disease mechanism for this multifunctional enzyme.




Vitamin $B_{12}$ (cobalamin, Cbl) is essential to the function of two human enzymes, methionine synthase and methylmalonyl-CoA mutase, utilizing methylcobalamin (MeCbl) and adenosylcobalamin (AdoCbl) as cobalt (Co)-containing cofactors, respectively.[1] The intracellular delivery of exogenous Cbl from its point of entry to the destination enzymes, in its usable cofactor forms, requires an exquisite pathway of uptake, processing, and subcellular trafficking proteins.[2] Eight different defects in the Cbl metabolic pathway, defined by the complementation groups *cblA-G* and *mut*, give rise to the rare inherited disorders of methylmalonic aciduria and homocystinuria that affect children in the newborn period or early childhood.[1] While the genes corresponding to these defects have been identified, the underlying protein dysfunctions remain poorly understood.

The most common intracellular Cbl disorder is caused by genetic defects in the *cblC* complementation group (OMIM 277400), which includes early-onset and late-onset forms.[3] The affected gene MMACHC (methylmalonic aciduria type C with homocystinuria) encodes the 282 amino acid (aa) MMACHC protein (also known as CblC) lacking overall sequence homology with other protein families.[4] The affected *cblC* patients (~400 to date) have combined homocystinuria and methylmalonic aciduria[5] due to the inability to synthesize both MeCbl and AdoCbl, suggesting an early functional block in the Cbl pathway. The exact function of MMACHC, which exhibits a versatile capacity to bind diverse Cbl forms containing different upper axial ligands (e.g., AdoCbl, MeCbl, cyanocobalamin CNCbl, hydroxycobalamin OHCbl),[6,7] is only beginning to be unravelled. MMACHC has recently been shown to reductively decyanate CNCbl to cob(II)alamin[6] in the presence of NADPH and FMN/FAD[7] and to dealkylate Cbl derivatives with C2−C6 alkanes, adenosyl- or methyl- as the upper axial ligand in the presence of reduced glutathione (GSH).[8,9] Together, these data demonstrate a role for MMACHC to bind internalized Cbl derivatives in any form available to the cell and process them to a common intermediate by removing the upper axial ligand. This reactivity of MMACHC involves a bond breakage between the Co atom and the upper axial ligand, a common mechanism in Cbl-binding enzymes that can proceed in either a homolytic or heterolytic manner.[10] Subsequently, MMACHC, in concert with its downstream interacting partner MMADHC,[11] partitions the processed Cbl intermediate to the ensuing AdoCbl and MeCbl synthetic pathways.

**Received:** February 2, 2012
**Revised:** April 26, 2012
**Published:** May 29, 2012

**Table 1. Summary of Diffraction and Refinement Statistics[a]**

| data collection | peak | remote |
|---|---|---|
| space group | *P*1 | |
| *a*, *b*, *c* (Å) | 71.7, 72.0, 300.1 | |
| *α*, *β*, *γ* (deg) | 88.6, 85.2, 83.8 | |
| wavelength (Å) | 0.9788 | 0.9686 |
| resolution (Å) | 20−2.70 (2.85−2.70) | 20−2.40 (2.53−2.40) |
| $R_{merge}$ (%) | 0.277 (1.179) | 0.193 (0.915) |
| $I/\sigma I$ | 11.5 (3.5) | 8.6 (2.1) |
| completeness (%) | 98.7 (98.4) | 98.6 (98.0) |
| redundancy | 14.5 (14.4) | 7.2 (7.2) |
| *phasing* | | |
| resolution (Å) | 20−2.4 | |
| $R_{cullis}$ | 0.7 | |
| phasing power | | |
| anomalous/isomorphous | 0.425/0.197 | |
| figure of merit (acentric) | 0.203 | |
| *refinement* | | |
| resolution (Å) | 20−2.4 | |
| no. reflections | 228734 | |
| $R_{work}/R_{free}$ (%) | 18.2/20.6 | |
| no. atoms | | |
| protein/ligand/water | 29534/2056/1646 | |
| *B*-factors (Å$^2$) | | |
| protein/ligand/water | 34.1/30.2/35.5 | |
| rms deviations | | |
| bond lengths (Å)/bond angles (deg) | 0.010/1.27 | |
| PDB code | 3SOM | |

[a]Numbers in parentheses refer to the highest resolution shell. $R_{merge}$, unweighted *R*-value on *I* between symmetry mates. $R_{work} = \sum hkl||F_{obs}(hkl)| - k|F_{calc}(hkl)||/\sum hkl|F_{obs}(hkl)|$ for the working set of reflections; $R_{free}$ is the *R*-value for 5% of the reflections excluded from refinement. $R_{cullis} = \sum hkl||FPH| \mid P|| - |FH(calc)|/\sum hkl(cent)||FPH| \pm |FP||$.

To provide insight into structure−function relationships and disease mechanism, we have determined the structure of human MMACHC in complex with AdoCbl, which corresponds to the largest upper axial ligand (5′-deoxyadenosine, Ado) among the physiological Cbl forms and is a substrate for the GSH-dependent dealkylation reaction.[8,9] During the preparation of our manuscript, Koutmos et al. reported the structure of MMACHC in its *apo* and MeCbl-bound forms.[12] Our MMACHC·AdoCbl structure agrees with Koutmos et al. on the overall protein fold and Cbl binding site but further shows a highly conserved dimerization cap for the upper axial ligand Ado, as well as an arginine-rich GSH-binding pocket that includes the site of the most prevalent missense mutation, at Arg161.[7]

## EXPERIMENTAL PROCEDURES

**Cloning, Expression, and Purification of MMACHC.** Constructs of full-length human MMACHC (aa 1−282; $MMACHC_{FL}$) and Arg161Gln mutant ($MMACHC_{R161Q}$) were prepared previously.[13] Other substitution ($MMACHC_{R206Q}$, $MMACHC_{R230Q}$, $MMACHC_{R111G}$) and deletion ($MMACHC_{\Delta PNR}$) mutants were constructed using the QuikChange site-directed mutagenesis kit (Strategene). A DNA fragment encoding C-terminally truncated (aa 1−235; $MMACHC_{\Delta 236-282}$) MMACHC was generated from the $MMACHC_{FL}$ construct and subcloned into the vector pNIC28-Bsa4 (GenBank ID: EF198106). The various MMACHC proteins were expressed and purified as described previously.[13] Selenomethione (SeMet)-derivatized proteins were expressed using standard procedures and purified as with wild-type proteins.

**Crystallization and Structure Determination.** SeMet-$MMACHC_{FL}$ crystals were grown by vapor diffusion at 20 °C in sitting drops containing 50 nL of protein (17 mg/mL, preincubated with 1 mM AdoCbl in buffer A: 10 mM HEPES pH 7.5; 500 mM NaCl; 5% glycerol), 100 nL of mother liquor (0.25 M ammonium citrate pH 5.0, 25% v/v PEG3350), and 25 nL of an additive mixture from Silver Bullets HT well E12 (Hampton Research). Crystals were cryo-protected in mother liquor containing 25% (v/v) ethylene glycol and flash-cooled in liquid nitrogen. $MMACHC_{FL}$ crystallized in space group *P*1 with unit cell dimensions $a$ = 71.7 Å, $b$ = 72.0 Å, $c$ = 300.1 Å, $\alpha$ = 88.5°, $\beta$ = 85.3°, and $\gamma$ = 83.8°. Two highly redundant data sets at different wavelengths were collected on beamline I24 at the Diamond Light Source. All data sets were integrated with XDS[14] and scaled with SCALA.[15] SHELXD[16] identified 40 selenium sites which were used for phase refinement with autoSHARP.[17] The initial electron density map after solvent flattening with SOLOMON[18] did not enable us to build a complete model, but the secondary structure elements identified with BUCANEER[19] allowed us to assemble the strands and helices of one region of the map into an approximate monomer model. This model was superimposed on the remaining fragments with COOT[20] to establish the 16-fold NCS in the asymmetric unit. A new round of solvent flattening with PARROT,[21] using initial phases from SHARP and 16-fold NCS, produced an electron density map of excellent quality. Subsequent automated model building was performed with BUCANEER. After several cycles of manual


dx.doi.org/10.1021/bi300150y | *Biochemistry* 2012, 51, 5083−5090

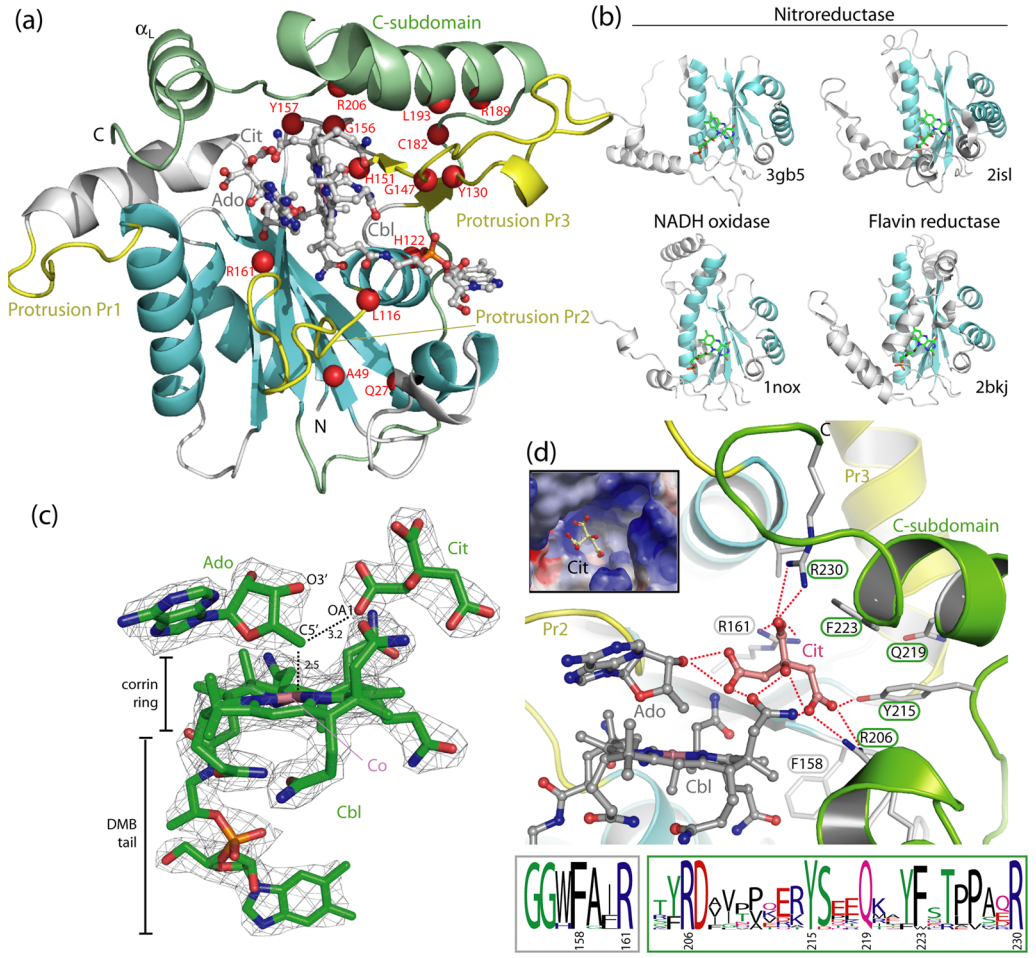



**Figure 1.** Structure of the MMACHC·AdoCbl complex. (a) The MMACHC protomer aligns with the core $\alpha + \beta$ fold (cyan) in other NTR homologues but contains unique loop protrusions (yellow) and a C-terminal subdomain (green). The bound AdoCbl and citrate (Cit) are in sticks. Missense mutations are labeled as red spheres. (b) Structural comparison of the NTR superfamily members, including *M. musculus* iodotyrosine deiodinase (PDB ID: 3GB5), *S. meliloti* BluB gene product (PDB ID: 2ISl), *T. Thermophilus* NADH oxidase (PDB ID: 1NOX), and *V. harveyi* flavin reductase P (PDB ID: 2BKJ), each labeled with their PDB codes. The bound FMN in each structure is shown in sticks. (c) SigmaA-weighted electron density map for the AdoCbl (Ado, Cbl) and citrate molecule (Cit) in the MMACHC active site. A number of atoms mentioned in the main text are labeled. Interatomic distances (dotted lines) are in angstroms (Å). (d) The arginine-rich pocket close to the Cbl binding site features positive electrostatic potentials (inset). Hydrogen bonds are shown in red dotted lines. The sequence logos (bottom), generated from an alignment of 30 MMACHC homologous sequences (WebLogo tool), highlight the invariant residues in this pocket, including Arg161, Arg206, and Arg230.

rebuilding in COOT and refinement with BUSTER,[22] the final model converged to a final $R_{work}/R_{free}$ of 18.2%/20.6%. Statistics for data collection and refinement are summarized in Table 1.

**Size Exclusion Chromatography.** Analytical gel filtration was performed on a Superdex 200 HiLoad 10/30 column (GE Healthcare) pre-equilibrated with 10 mM HEPES pH 7.5, 150 mM NaCl and 5% glycerol. The column was calibrated using carbonic anhydrase (29 kDa), bovine serum albumin (66 kDa), alcohol dehydrogenase (150 kDa), and apoferritin (443 kDa) as protein standards. Experiments were performed with 25 $\mu$M protein either in the *apo*-state or in the presence of 250 $\mu$M AdoCbl or MeCbl, or 50, 100, or 500 $\mu$M FMN.

**Differential Scanning Fluorimetry.** MMACHC was assayed for shifts in melting temperature caused by ligand binding in a 96-well PCR plate using an Mx3005p RT-PCR machine (Stratagene). Each well (20 $\mu$L) consisted of protein (3 $\mu$M in buffer A), SYPRO-Orange (Invitrogen) diluted 1000×, and various concentrations of reduced glutathione (GSH) in the presence or absence of 20 $\mu$M AdoCbl or 20 $\mu$M MeCbl. Fluorescence intensities were measured from 25 to 96 °C with a ramp rate of 1 °C/min. The temperature shifts, $\Delta T_m^{obs}$, for each ligand were determined as described.[13] $AC_{50}$ values (half-maximal effective ligand concentration) and final graphs were generated using GraphPad Prism (v.5.01; GraphPad Software) as described.[13]

**Dealkylation Assay.** Dealkylation of MeCbl was initiated by adding 1 mM GSH to a 100 $\mu$L reaction mixture containing 50 $\mu$M MMACHC and 5 $\mu$M MeCbl in buffer A in the dark under aerobic conditions at 20 °C. At various time points (0–120 min) the absorbance spectrum (300–750 nm) was read using a POLARstar Omega plate reader (BMG Labtech), and formation of aquocobalamin was followed at 355 nm. Rate constants were determined by GraphPad Prism for MMACHC wild-type and $MMACHC_{R161Q}$ following incubation with various concentrations of GSH (0–1000 $\mu$M), measuring absorbance changes at 355 nm over 60 min and determination of linear velocity.


dx.doi.org/10.1021/bi300150y | *Biochemistry* 2012, 51, 5083–5090

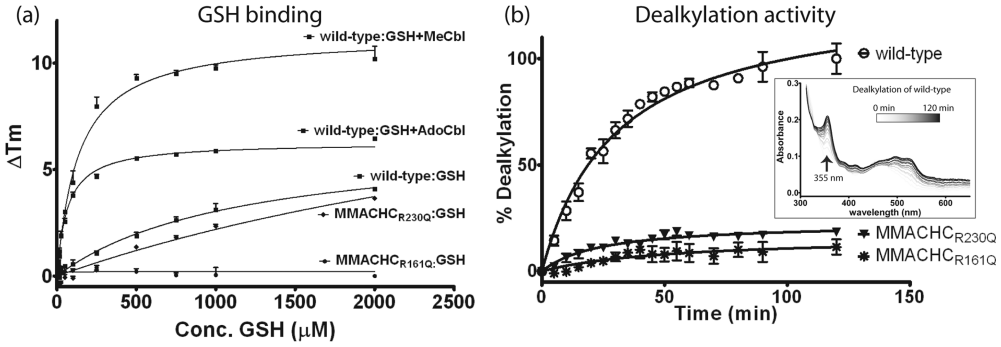



**Figure 2.** The arginine cluster is involved in GSH binding and dealkylation. (a) A plot of $T_m$ shift as a function of GSH concentration for wild-type MMACHC (wild-type:GSH), MMACHC$_{R161Q}$ (MMACHC$_{R161Q}$:GSH), and MMACHC$_{R230Q}$ (MMACHC$_{R230Q}$:GSH) in the presence of increasing concentrations of GSH alone and in the presence of 5 $\mu$M MeCbl (wild-type:GSH+MeCbl) and 5 $\mu$M AdoCbl (wild-type:GSH+AdoCbl). (b) Dealkylation activity, monitored by absorption at 355 nm, of MeCbl-bound MMACHC$_{R161Q}$ and MMACHC$_{R230Q}$ as compared to wild-type MMACHC$_{FL}$. Inset shows the dealkylation reaction over time (increasing gray shade) for the wild-type protein as a function of wavelength.

## RESULTS AND DISCUSSION

**MMACHC Modifies the Nitroreductase Scaffold with $B_{12}$ Binding Features.** We determined the MMACHC·AdoCbl structure at 2.4 Å resolution by MAD phasing using SeMet-derivatized protein, in the crystal space group *P*1 with 16 protomers in the asymmetric unit (rmsd ~ 0.2 Å among protomers). The refined monomeric model comprises aa 5−230 (Figure 1a), while the C-terminal 52 residues (aa 231−282), encompassing a Pro-rich region absent in many MMACHC orthologues (Figure S1 of the Supporting Information), are disordered. Despite a lack of sequence conservation to known protein families, MMACHC exhibits remote structural homology to several bacterial nitroreductases (NTRs), with DALI search *Z*-scores of 5.0−5.9 and rmsd of ~3.0 Å (Figure 1b, top). The NTRs, together with NADH oxidases and flavin reductases (Figure 1b, bottom),[23] form a superfamily of flavin mononucleotide (FMN)-binding proteins, characterized by a core $\alpha + \beta$ scaffold and a domain-swapping dimer with the substrate and FMN binding sites at the dimer interface (Figure S2a).

Compared to NTRs, nevertheless, MMACHC displays significant structural modifications. First, only 90−100 $C^{\alpha}$ positions of MMACHC superimpose with the typical ~200 aa core NTR scaffold (cyan region in Figure 1a,b). Second, although MMACHC has been shown to decyanate CNCbl in the presence of FMN,[12] its binding site cannot be inferred from the FMN binding site of the NTR homologues because the equivalent region in MMACHC is constricted by additional secondary structure elements and is too distant from the Cbl binding site for catalysis (Figure S2b). Third, MMACHC appends to the core NTR scaffold three unique protrusions (Pr1: aa 69−77; Pr2: 104−116; Pr3: 128−150) (Figure 1a, yellow) as well as a C-terminal subdomain (aa 166−230; Figure 1a, green). Significantly, the longest protrusion Pr3 and the C-terminal subdomain together construct a binding pocket for the Cbl corrin ring and 5,6-dimethylbenzimidazole (DMB) base. Pr3 forms an extended linker connecting a short $\beta$-hairpin and creates a platform to bind Cbl. The C-terminal subdomain has a long linker wrapping around the core scaffold and a helical bundle sitting atop Pr3. The last helix $\alpha_L$ in the C-terminal subdomain, immediately before the disordered Pro-rich region, exhibits positional flexibility and flanks the Cbl binding site.

Bound in each of the 16 MMACHC protomers is an AdoCbl molecule, with 5′-deoxyadenosine (Ado) as the upper axial ligand for the Cbl moiety (Figure 1c). The Cbl moiety is bound such that the DMB tail is not ligated to the Co atom at the lower axial position. This "base-off" configuration is consistent with spectroscopic studies in solution[6,7] and also observed in the reported MMACHC·MeCbl structure.[12] The Co atom is tetracoordinated equatorially by the corrin ring and further coordinated at its upper axial position by the Ado C5′ atom. The observed $Cbl_{Co}$:$Ado_{C5'}$ distance (2.5 Å) is longer than expected for an unbroken Co(III)−C bond (2.0 Å),[24] and has previously been suggested to represent an artifactual bond breakage during X-ray analysis.[25] MMACHC lacks a conserved His residue employed by many $B_{12}$-utilizing enzymes as a lower axial ligand to replace the DMB tail and complete the six-coordinate configuration for the Co(III) oxidation state. Hence, MMACHC mediates a "base-off, five-coordinate" mode to bind AdoCbl and presumably other alkyl-Cbls. This binding adaptation of MMACHC provides a means to facilitate the catalytic cleavage of the upper axial ligand (e.g., Ado, Me, CN moieties) from the Cbl moiety, an important functionality for the enzyme. The base-off, five-coordinate mode weakens the Co−C bond compared to Cbl forms containing a lower axial ligand (e.g., "base-on" or "base-off/His-on")[26] and has been shown to reduce the energy barrier in favor of a heterolytic reaction,[27] which MMACHC purportedly utilizes for its dealkylation activity.[9]

**A Highly Conserved Arginine Cluster near the Cbl Binding Site.** The available MMACHC structures allow the three-dimensional mapping of *cblC* patient mutations in order to correlate mutation and disease phenotype. Twenty missense mutations have been reported to date (Table S1), a majority of which are clustered into the region that interacts with the corrin ring and DMB tail (Figure 1a, red spheres) and hence are expected to compromise Cbl binding to varying degrees. The site of the most prevalent missense mutation Arg161Gln,[5] nevertheless, does not interact with Cbl directly and is situated in a cavity near the corrin ring and Ado moieties. This cavity harbors several strictly invariant residues (Figure 1d, Figure S1), among which the guanidinium side chains of Arg161, Arg206, and Arg230 cluster into an electropositive pocket (Figure 1d, inset). In each of the 16 MMACHC·AdoCbl protomers, this arginine-rich pocket is bound by a citrate ion, presumably derived from the crystallization solution. The citrate ion is fixed by many charged interactions with the arginine cluster (Figure 1d, red dashed lines) using its central carboxylate and one of the terminal carboxylate groups. The other terminal carboxylate forms hydrogen bonds to the Ado


dx.doi.org/10.1021/bi300150y | *Biochemistry* 2012, 51, 5083−5090

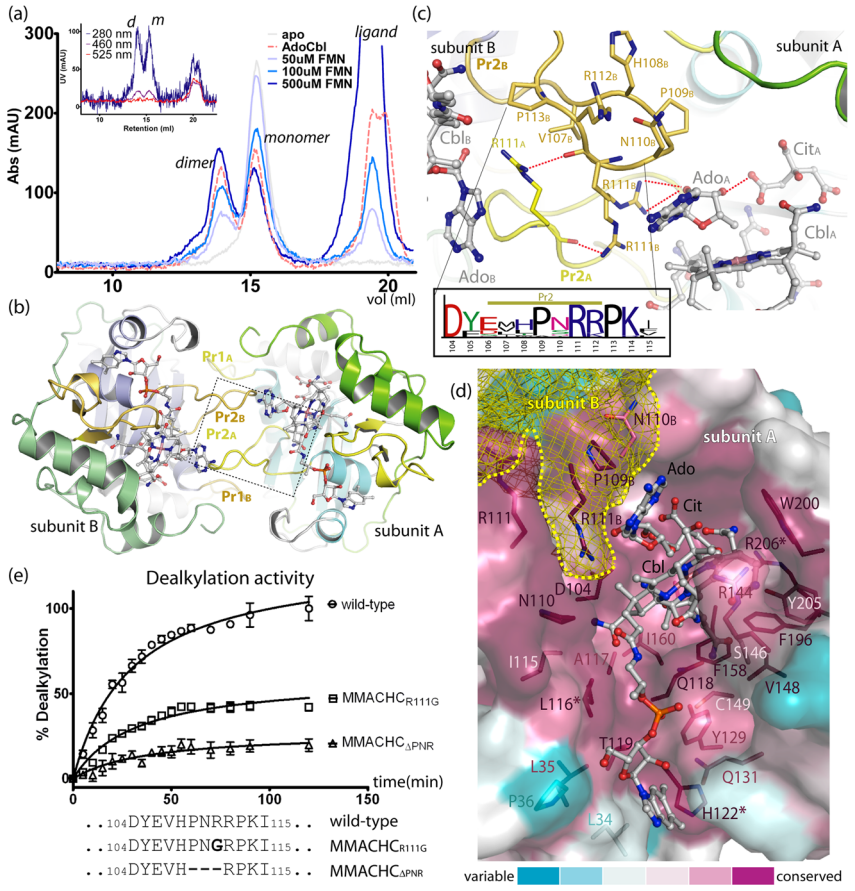



**Figure 3.** MMACHC forms a dimer in solution and in the crystal structure. (a) Gel filtration profile for $MMACHC_{FL}$ in the *apo* form (gray line), bound with 50, 100, or 500 $\mu$M FMN (lines of increasing blue shade) or bound with 250 $\mu$M AdoCbl (red dashed line). Inset: the $A_{460}/A_{525}$ absorbance ratio for the monomer (*m*) and dimer (*d*) fractions reveals a "base-off" configuration for the AdoCbl ligand. (b) Structure of a MMACHC dimer, highlighting the interface formed by protrusions Pr1 and Pr2 from both subunits A and B. (c) Close-up view of the dimer interface (boxed in panel B), showing the protrusion Pr2 in subunit B ($Pr2_B$) interacts with its counterpart in subunit A ($Pr2_A$) and with the 5′-deoxyadenosine moiety of subunit A ($Ado_A$). Two alternative side-chain conformations of $R111_B$ are shown. Inset: sequence logo of the Pr2 PNRRP motif showing the invariant amino acids therein. (d) Surface representation of the Cbl binding site in subunit A, showing that the Pr2 PNRRP motif from subunit B forms a cap (yellow meshes) for the Ado moiety in subunit A. Residues that interact with AdoCbl are shown in sticks and colored on a scale of sequence conservation. *cblC* missense mutations identified in the Cbl binding site are marked with asterisks. (e) Dealkylation activity, monitored by absorption at 355 nm, of MeCbl-bound cap mutants $MMACHC_{R111G}$ and $MMACHC_{\Delta PNR}$ as compared to the wild-type. The sequences for the domain-swapped cap of the three proteins are shown.

O3 ribose, with one of its oxygen (OA1) atoms 3.2 Å from the Ado C5′ (Figure 1c).

The proximity of the arginine-rich pocket to the upper axial ligand (e.g., Ado) of the Cbl binding site suggests a physiological role in binding an enzyme cofactor. To this end, the binding of GSH, required for the dealkylation reaction of MMACHC, was considered because the role of arginines in binding GSH is well documented for such GSH-utilizing enzymes as glyoxalase,[28] glutathione reductase,[29] and glutathione-*S*-transferase.[30] To probe the function of the arginine cluster, we constructed three site-directed single mutants (Arg161Gln, Arg206Gln, and Arg230Gln) to replace each arginine in the cluster with an uncharged amino acid. The resulting $MMACHC_{R206Q}$ protein was insoluble when expressed recombinantly (data not shown), suggesting that Arg206, also a site harboring *cblC* patient mutations,[5] has a structural role contributing to overall protein stability. The $MMACHC_{R161Q}$ and $MMACHC_{R230Q}$ proteins, by contrast, were expressed in the soluble fraction in *E. coli* (data not shown) and then compared to the wild-type protein in their ability to bind GSH and catalyze the dealkylation reaction.

To assess binding of GSH by MMACHC, we used differential scanning fluorimetry to detect ligand-induced changes in protein thermostability, a method previously employed to characterize the binding of Cbl derivatives to wild-type and mutant MMACHC.[13] $MMACHC_{FL}$ incubated with GSH alone (1 mM) showed a flat concentration dependent increase in $T_m$, yielding an $AC_{50}$ of ∼1 mM (Table S2) and a moderate $T_m$ shift (3 °C) (Figure 2a). In the presence of 20 $\mu$M AdoCbl or MeCbl, however, GSH binding and stabilization are markedly increased, with $AC_{50}$ values reduced by ∼17-fold (AdoCbl) and 8-fold (MeCbl) (Table S2), and stabilization of 6 °C (AdoCbl) or 10 °C (MeCbl) at 1 mM GSH (Figure 2a). This increased affinity for and stabilization by GSH in the presence of alkyl-Cbls suggests that binding of these cofactors primes MMACHC to bind GSH, possibly to initiate the dealkylation reaction. Similar upward $T_m$ shifts were obtained when GSH and AdoCbl/MeCbl were added to $MMACHC_{\Delta 236-282}$ (data not shown), indicating that the Pro-rich region is not involved in GSH binding. These $T_m$ shifts were diminished, however, in the $MMACHC_{R161Q}$ mutant and, to a much lesser extent, the


dx.doi.org/10.1021/bi300150y | *Biochemistry* 2012, 51, 5083−5090

$MMACHC_{R230Q}$ mutant, when bound to GSH alone (Figure 2a), suggesting that residues in the arginine cluster are involved in GSH binding. The physiological relevance of the interaction between MMACHC and GSH is demonstrated by the dealkylation of AdoCbl or MeCbl in the presence of GSH. We found that dealkylation of MeCbl by 1 mM GSH was severely disrupted in both $MMACHC_{R161Q}$ and $MMACHC_{R230Q}$, with only 11.5 ± 6.1% and 18.3 ± 4.2% of wild-type activity after 120 min (Figure 2b, Figure S3), in spite of the fact that all MeCbl was bound to the protein (Figure S4).

Together, our data for the first time demonstrate a direct role of the invariant arginine cluster in the structure and catalysis of MMACHC and attribute the deleterious effect of the common Arg161Gln disease mutation to weakened GSH binding and abolished dealkylation activity. This is consistent with previous data[13] showing that $B_{12}$ binding is unlikely to be the primary dysfunction caused by the mutation. Although crystallization of the functional ternary complex (e.g., MMACHC·AdoCbl·GSH) has remained elusive, the binding mode of GSH can be inferred from our citrate-bound MMACHC·AdoCbl structure. Residues in the arginine-rich pocket are in position to coordinate the two terminal carboxylates of GSH and position its central thiolate group toward the scissile Co−C bond in the alkyl-Cbl substrate. The thiolate group may adopt a similar position as the citrate OA1 atom (shown in Figure 1c) to provide a good trajectory (~3.2 Å) for the nucleophilic dealkylation. The arginine cluster could also function as counterions for the ionization of the thiolate nucleophile. Furthermore, the relatively larger GSH molecule (with respect to citrate) can be accommodated by means of minor side-chain shifts at the GSH pocket, as well as a moderate $C_\alpha$ movement of the C-terminal helix ($\alpha_L$) that harbors Arg230 as the last ordered residue in our structure. Consistent with this, we observed conformational flexibility of helix $\alpha_L$ among the protomers in our structure (data not shown), and as reported,[12] reflecting the ability to accommodate a GSH-mediated induced-fit rearrangement.

**MMACHC Dimerizes in Solution and in the Crystal.** Although MMACHC has previously been shown to exist as a monomer in the *apo* state,[6] we set out to determine whether ligand binding results in functional oligomer formation, using size exclusion chromatography (Figure 3a). By this method, *apo*-$MMACHC_{FL}$ exists predominantly as a monomer (gray line). However, upon incubation with FMN, it elutes with both monomer and dimer peaks, indicating the existence of a monomer−dimer equilibrium. The dimer peak height is commensurate with the concentration of FMN administered (lines of increasing blue shade), demonstrating that the binding of FMN favors the equilibrium toward dimer formation. A similar phenomenon is observed upon binding of AdoCbl (red dashed line), and the $A_{460}/A_{525}$ absorbance ratio is consistent with the presence of bound AdoCbl in the base-off configuration for both monomeric and dimeric fractions (inset). Dimerization was also induced by the presence of MeCbl, albeit to a lesser extent (Figure S5). The above monomer−dimer profile for the $MMACHC_{FL}$ protein is also observed with the truncated $MMACHC_{\Delta 236-282}$ protein lacking the Pro-rich region (data not shown), indicating that the first 235 aa are sufficient for dimer formation in solution.

Our solution data demonstrate that while a MMACHC monomer, equipped with the corrin ring and DMB binding pockets, is sufficient to bind AdoCbl, two ligand-binding monomers can assemble into a dimer. Inspection of the intermolecular packing of the MMACHC·AdoCbl structure, which involves the three unique protrusions Pr1−Pr3, provides an explanation for the observed solution dimer. The 16 protomers in the crystal are arranged as a string of four tetramers, each tetramer consisting of two dimers related by a 90° screw rotation (Figure S6a). Each dimer buries 1050 $Å^2$ of solvent accessible surface area per monomer, above the statistical average for biological dimers (~800 $Å^2$).[31] The dimeric shape and interface differ markedly from those in the NTR homologues (Figure S2a) and are largely mediated by protrusions Pr1 and Pr2 from both subunits (Figure 3b), thus forming the base for a U-shaped structure with the two tips comprised of protrusion Pr3 (Figure S6b). The tips are further involved in tetramerization, allowing two dimers to wedge in between their U-shaped cavities (Figure S6b). Although tetramerization was not detected in solution, this solvent-exposed intermolecular surface can provide the protein−protein interaction platform for binding partners such as MMADHC, a protein downstream of the Cbl pathway.[11]

Close inspection of the dimer interface reveals that protrusion Pr2 from each subunit, encompassing a highly conserved "$P_{109}NRRP_{113}$" sequence motif, crosses over, and interacts with the Ado moiety of AdoCbl in the opposing dimer (Figure 3c), essentially forming a domain-swapped "cap" for the upper axial ligand *in trans* (Figure 3d, yellow mesh). Contacts between the cap and the upper axial ligand involve van der Waals interactions contributed by Pro109, Asn110, and Arg111 as well as hydrogen bonds contributed by Arg111 (Figure 3c). The invariant Arg111 also adopts an alternative guanidinium side-chain conformation that forms head-to-tail arginine−arginine interactions at the intermolecular interface. This dimeric packing mode, independently observed eight times in the *P*1 asymmetric unit, was also adopted by NCS-related molecules in two other crystal lattices that we characterized (monoclinic *P*2 and tetragonal *P*4 space groups; data not shown). In addition, it is also present in the reported MMACHC·MeCbl structure (PDB ID: 3SC0) when applying crystallographic symmetry to its asymmetric unit monomer. Together, the MMACHC dimer in the MMACHC·AdoCbl and MMACHC·MeCbl complexes reveals a general function of the PNRRP cap to accommodate different upper axial ligands. Relevant to this, the PNRRP cap extends across the intermolecular interface to bind *in trans* the upper axial ligand only in the MMACHC·AdoCbl/MeCbl complexes, but in the reported *apo*-MMACHC structure the PNRRP cap (labeled $L_{3F}$ in ref 12) is held more in place within its own monomer,[12] further demonstrating that the function of the dimeric cap is dependent upon the presence of the ligand.

The physiological relevance of assembling two Cbl-binding monomers to form a MMACHC dimer may be manifested in how it facilitates and extends the functional capability of the protein. The PNRRP motif in Pr2, together with residues interacting with the corrin ring and DMB base at the Cbl binding site, represents the most evolutionarily conserved residues among MMACHC orthologues (Figure 3d, magenta). The intradimeric crossover of the PNRRP cap creates a cavity that accommodates the upper axial ligand of varying sizes (e.g., the bulky Ado group, and other alkyl- or CN-substituents). The relaxed specificity of this cavity takes the function of MMACHC to the realm of utilizing any Cbl forms that a cell can encounter extracellularly. A protein domain capping the upper axial ligand is a common architectural feature in Cbl-utilizing enzymes,[32] to assist catalysis by sequestering its active

site and providing a complete shielding of the Cbl ligand, as well as proper positioning of the upper axial ligand to mediate catalysis. The essential nature of the PNRRP cap to MMACHC function is succinctly illustrated by the reduced dealkylation activity, relative to wild-type proteins, of cap mutants (Figure 3e) that either substituted Arg111 ($MMACHC_{R111G}$) or removed part of the domain-swapped residues (Pro109-Asn110-Arg111) in the PNRRP cap ($MMACHC_{\Delta PNR}$).

## CONCLUDING REMARKS

Positioned early in the Cbl metabolic pathway, MMACHC is tasked with multiple binding and enzymatic roles, a feat unprecedented compared to other $B_{12}$ binding/utilizing proteins. To accomplish this, MMACHC seems to have evolved from a known structural NTR binding core and tailor-made to include unique ligand-binding features: a Cbl pocket to bind and prepare various Cbl substrates for Co−C bond cleavage and a nearby arginine-rich pocket for GSH coenzyme binding. These functions are likely to be further facilitated by dimer formation, which creates a substrate-binding cap for the active site as well as additional interaction platforms for the Cbl substrates and functional partners. Given our new structural understanding of this protein, it is easier to appreciate how mutations targeted to the Cbl and arginine-rich pockets are deleterious in *cblC* patients. What remains unresolved, however, is how MMACHC accommodates the FMN/FADH-dependent decyanation reaction. In particular, it is not apparent why MMACHC retains the core NTR fold known for its FMN binding, yet projection of the FMN binding site from structural homologues leaves it too distant from the core $B_{12}$- and GSH-binding actions to be catalytically effective. The observations that FMN induces MMACHC dimerization and that the MMACHC dimer is structurally distinct from the NTR homologues, together imply the possibility of an alternative FMN binding mode in MMACHC—which could be in the arginine-rich coenzyme pocket, or a site formed upon MMACHC dimerization, although these remain to be determined.

## ASSOCIATED CONTENT

### Supporting Information

Missense mutations reported for the MMACHC gene (Table S1), $AC_{50}$ values for GSH binding (Table S2), sequence alignment of MMACHC homologues (Figure S1), FMN binding site in nitroreductases (Figure S2), dealkylation rate of MMACHC wild-type and $MMACHC_{R161Q}$ (Figure S3), evaluation of MeCbl binding to MMACHC before the dealkylation reaction (Figure S4), size exclusion chromatography of MMACHC (Figure S5), subunit organization in the *P*1 asymmetric unit (Figure S6). This material is available free of charge via the Internet at http://pubs.acs.org.

### Accession Codes

Coordinates and structure factors for the human MMACHC·AdoCbl structure have been deposited in the Protein Data Bank as entry 3SOM.

## AUTHOR INFORMATION


### Corresponding Author

*Phone +44 (1865) 617757; e-mail wyatt.yue@sgc.ox.ac.uk.

### Author Contributions

⊥D.S.F. and T.K. contributed equally to this work.



### Funding

The Structural Genomics Consortium is a registered charity (Number 1097737) funded by the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institutet, the Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust, the Ontario Ministry for Research and Innovation, Merck and Co., Inc., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research, and the Wellcome Trust. R.A.G. acknowledges funding support by the Canadian Institutes for Health Research Grant MOP-44353.

### Notes

The authors declare no competing financial interest.


## ACKNOWLEDGMENTS


We thank the SGC protein crystallography group and the staff at Diamond Light Source for help in synchrotron data collection.


## ABBREVIATIONS

aa, amino acids; MMACHC, methylmalonic aciduria type C with homocystinuria; Cbl, cobalamin; Co, cobalt; AdoCbl, adenosylcobalamin; MeCbl, methylcobalamin; CNCbl, cyanocobalamin; OHCbl, hydroxycobalamin; Ado, 5′-deoxyadenosine; DMB, 5,6-dimethylbenzimidazole; GSH, reduced glutathione; MAD, multiwavelength anomalous dispersion; SeMet, selenomethionine; NTR, nitroreductase; FMN, flavin mononucleotide.

## REFERENCES


(1) Rosenblatt, D. S., Fenton, W. A., Scriver, C. R., Beaudet, A. L., Valle, D., and Sly, W. S. (2001) Inherited disorders of folate and cobalamin transport and metabolism, in *The Metabolic and Molecular Bases of Inherited Disease*, pp 3897−3933, McGraw-Hill, New York.

(2) Froese, D. S., and Gravel, R. A. (2010) Genetic disorders of vitamin B metabolism: eight complementation groups--eight genes. *Expert Rev. Mol. Med. 12*, e37.

(3) Rosenblatt, D. S., Aspler, A. L., Shevell, M. I., Pletcher, B. A., Fenton, W. A., and Seashore, M. R. (1997) Clinical heterogeneity and prognosis in combined methylmalonic aciduria and homocystinuria (cblC). *J. Inherited Metab. Dis. 20*, 528−538.

(4) Lerner-Ellis, J. P., Tirone, J. C., Pawelek, P. D., Dore, C., Atkinson, J. L., Watkins, D., Morel, C. F., Fujiwara, T. M., Moras, E., Hosack, A. R., Dunbar, G. V., Antonicka, H., Forgetta, V., Dobson, C. M., Leclerc, D., Gravel, R. A., Shoubridge, E. A., Coulton, J. W., Lepage, P., Rommens, J. M., Morgan, K., and Rosenblatt, D. S. (2006) Identification of the gene responsible for methylmalonic aciduria and homocystinuria, cblC type. *Nat. Genet. 38*, 93−100.

(5) Lerner-Ellis, J. P., Anastasio, N., Liu, J., Coelho, D., Suormala, T., Stucki, M., Loewy, A. D., Gurd, S., Grundberg, E., Morel, C. F., Watkins, D., Baumgartner, M. R., Pastinen, T., Rosenblatt, D. S., and Fowler, B. (2009) Spectrum of mutations in MMACHC, allelic expression, and evidence for genotype-phenotype correlations. *Hum. Mutat. 30*, 1072−1081.

(6) Kim, J., Gherasim, C., and Banerjee, R. (2008) Decyanation of vitamin B12 by a trafficking chaperone. *Proc. Natl. Acad. Sci. U. S. A. 105*, 14551−14554.

(7) Froese, D. S., Zhang, J., Healy, S., and Gravel, R. A. (2009) Mechanism of vitamin B12-responsiveness in cblC methylmalonic aciduria with homocystinuria. *Mol. Genet. Metab. 98*, 338−343.

(8) Hannibal, L., Kim, J., Brasch, N. E., Wang, S., Rosenblatt, D. S., Banerjee, R., and Jacobsen, D. W. (2009) Processing of alkylcobala-

mins in mammalian cells: A role for the MMACHC (cblC) gene product. *Mol. Genet. Metab. 97*, 260−266.
(9) Kim, J., Hannibal, L., Gherasim, C., Jacobsen, D. W., and Banerjee, R. (2009) A human vitamin B12 trafficking protein uses glutathione transferase activity for processing alkylcobalamins. *J. Biol. Chem. 284*, 33418−33424.
(10) Krautler, B. (2005) Vitamin B12: chemistry and biochemistry. *Biochem. Soc. Trans. 33*, 806−810.
(11) Plesa, M., Kim, J., Paquette, S. G., Gagnon, H., Ng-Thow-Hing, C., Gibbs, B. F., Hancock, M. A., Rosenblatt, D. S., and Coulton, J. W. (2011) Interaction between MMACHC and MMADHC, two human proteins participating in intracellular vitamin B metabolism. *Mol. Genet. Metab. 102*, 139−148.
(12) Koutmos, M., Gherasim, C., Smith, J. L., and Banerjee, R. (2011) The structural basis of multifunctionality in a B12 processing enzyme. *J. Biol. Chem. 286*, 29780−29787.
(13) Froese, D. S., Healy, S., McDonald, M., Kochan, G., Oppermann, U., Niesen, F. H., and Gravel, R. A. (2010) Thermolability of mutant MMACHC protein in the vitamin B12-responsive cblC disorder. *Mol. Genet. Metab. 100*, 29−36.
(14) Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 133−144.
(15) CCP4. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 50*, 760−763.
(16) Sheldrick, G. M. (2008) A short history of SHELX. *Acta Crystallogr., Sect. A 64*, 112−122.
(17) Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007) Automated structure solution with autoSHARP. *Methods Mol. Biol. (N. Y., NY, U. S.) 364*, 215−230.
(18) Abrahams, J. P., and Leslie, A. G. W. (1996) Methods used in the structure determination of bovine mitochondrial F-1 ATPase. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 52*, 30−42.
(19) Cowtan, K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 62*, 1002−1011.
(20) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2126−2132.
(21) Zhang, K. Y., Cowtan, K., and Main, P. (1997) Combining constraints for electron-density modification. *Methods Enzymol. 277*, 53−64.
(22) Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vonrhein, C., and Womack, T. O. (2011) BUSTER, 2.8.0 ed., Global Phasing Ltd., Cambridge, UK.
(23) Brenner, S. E., Chothia, C., Hubbard, T. J., and Murzin, A. G. (1996) Understanding protein structure: using scop for fold interpretation. *Methods Enzymol. 266*, 635−643.
(24) Ouyang, L., Rulis, P., Ching, W. Y., Nardin, G., and Randaccio, L. (2004) Accurate redetermination of the X-ray structure and electronic bonding in adenosylcobalamin. *Inorg. Chem. 43*, 1235−1241.
(25) Larsson, K. M., Logan, D. T., and Nordlund, P. (2010) Structural basis for adenosylcobalamin activation in AdoCbl-dependent ribonucleotide reductases. *ACS Chem. Biol. 5*, 933−942.
(26) Lexa, D., and Saveant, J. M. (1983) The electrochemistry of vitamin B12. *Acc. Chem. Res. 16*, 235−243.
(27) Wirt, M. D., Kumar, M., Wu, J. J., Scheuring, E. M., Ragsdale, S. W., and Chance, M. R. (1995) Structural and electronic factors in heterolytic cleavage: formation of the Co(I) intermediate in the corrinoid/iron-sulfur protein from Clostridium thermoaceticum. *Biochemistry 34*, 5269−5273.
(28) Schasteen, C. S., and Reed, D. J. (1983) Involvement of arginine residues in glutathione binding to yeast glyoxalase I. *Biochim. Biophys. Acta 742*, 419−425.
(29) Janes, W., and Schulz, G. E. (1990) Role of the charged groups of glutathione disulfide in the catalysis of glutathione reductase: crystallographic and kinetic studies with synthetic analogues. *Biochemistry 29*, 4022−4030.
(30) Patskovsky, Y. V., Patskovska, L. N., and Listowsky, I. (2000) The enhanced affinity for thiolate anion and activation of enzyme-bound glutathione is governed by an arginine residue of human Mu class glutathione S-transferases. *J. Biol. Chem. 275*, 3296−3304.
(31) Lo Conte, L., Chothia, C., and Janin, J. (1999) The atomic structure of protein-protein recognition sites. *J. Mol. Biol. 285*, 2177−2198.
(32) Gruber, K., Puffer, B., and Krautler, B. (2011) Vitamin B(12)-derivatives-enzyme cofactors and ligands of proteins and nucleic acids. *Chem. Soc. Rev. 40*, 4346−4363.