THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 27, pp. 20897–20903, July 2, 2010

Crystal Structure of Full-length *Mycobacterium tuberculosis* H37Rv Glycogen Branching Enzyme

*INSIGHTS OF N-TERMINAL β-SANDWICH IN SUBSTRATE SPECIFICITY AND ENZYMATIC ACTIVITY**[S]

Received for publication, March 9, 2010, and in revised form, April 26, 2010 Published, JBC Papers in Press, May 5, 2010, DOI 10.1074/jbc.M110.121707

Kuntal Pal[‡], Shiva Kumar[‡], Shikha Sharma[§1], Saurabh Kumar Garg[§2,3], Mohammad Suhail Alam[§2,4], H. Eric Xu[¶], Pushpa Agrawal[§5], and Kunchithapadam Swaminathan[‡6]

From the [‡]Department of Biological Sciences, National University of Singapore, Singapore 117543, the [§]Institute of Microbial Technology, Council of Scientific and Industrial Research, Chandigarh 160036, India, and the [¶]Laboratory of Structural Sciences, Van Andel Research Institute, Grand Rapids, Michigan 49503

The open reading frame Rv1326c of *Mycobacterium tuberculosis* (Mtb) H37Rv encodes for an α-1,4-glucan branching enzyme (MtbGlgB, EC 2.4.1.18, Uniprot entry Q10625). This enzyme belongs to glycoside hydrolase (GH) family 13 and catalyzes the branching of a linear glucose chain during glycogenesis by cleaving a 1→4 bond and making a new 1→6 bond. Here, we show the crystal structure of full-length MtbGlgB (MtbGlgBWT) at 2.33-Å resolution. MtbGlgBWT contains four domains: N1 β-sandwich, N2 β-sandwich, a central $(\beta/\alpha)_8$ domain that houses the catalytic site, and a C-terminal β-sandwich. We have assayed the amylase activity with amylose and starch as substrates and the glycogen branching activity using amylose as a substrate for MtbGlgBWT and the N1 domain-deleted (the first 108 residues deleted) MtbΔ108GlgB protein. The N1 β-sandwich, which is formed by the first 105 amino acids and superimposes well with the N2 β-sandwich, is shown to have an influence in substrate binding in the amylase assay. Also, we have checked and shown that several GH13 family inhibitors are ineffective against MtbGlgBWT and MtbΔ108GlgB. We propose a two-step reaction mechanism, for the amylase activity (1→4 bond breakage) and isomerization (1→6 bond formation), which occurs in the same catalytic pocket. The structural and functional properties of MtbGlgB and MtbΔ108GlgB are compared with those of the N-terminal 112-amino acid-deleted *Escherichia coli* GlgB (ECΔ112GlgB).

Tuberculosis is still a major killer infectious disease, at least in developing countries. *Mycobacterium tuberculosis* (Mtb),[7] the causative bacterial agent of tuberculosis, survives for a long period of time intracellularly and causes latent tuberculosis. If the organism is physiologically inactive for a long period of time, its storage sugars become very important for its survival. Therefore, understanding the nature of the enzymes that are involved in the metabolism of these storage sugars is very important. Glycogen is one of the most important storage sugars in the living world and provides nutrition to the host. Furthermore, the cell envelop of Mtb has a very important role during host-pathogen interactions. The outermost layer of the cell envelop of Mtb consists of a loosely bound structure, known as the capsule (1–3). It has been demonstrated that the major components of the capsular material are carbohydrates and proteins with a very small amount of lipids (4–6). The major carbohydrate constituent (~80%) of the Mtb capsule is a high molecular mass (>100,000 Da) α-glucan, which is composed of a (→4-α-D-Glc-1→) core and branched at position 6, every 5 or 6 residues, by (→4-α-D-Glc-1→) oligoglucosides (4, 7, 8). α-Glucan mediates non-opsonic binding of Mtb to CR3 (complement receptor3) (9) and is instrumental in blocking CD1 expression in Mtb (10). Stokes *et al.* (11) have shown that the capsular material of Mtb also displays antiphagocytic properties with certain types of macrophages.

Glycogen synthesis is an endergonic process. Glycogen, composed of branched polymer chains of glucose, is synthesized from monomers of UDP-glucose. The glycogen branching enzyme (GlgB), also known as amylo(α1→4–6)transglycosylase, catalyzes the transfer of a fragment of 6–7 glucose units from a non-reducing end to the hydroxyl group of C6 of a glucose unit, either on the same glucose chain or adjacent chains (Scheme 1). This enzyme belongs to carbohydrate binding module family 48 and glycoside hydrolase family 13 (GH13) (12). The details of a plausible reaction mechanism for GlgB are discussed later. Although GlgB is essential in several organisms, Sassati *et al.* (13) and Sambou *et al.* (14) have shown, in particular, that a

* This work was supported by grants from the Ministry of Education and Biomedical Research Council, Singapore (to K. S.) and Institute of Microbial Technology, India (to P. A.) and scholarships from National University of Singapore (to K. P. and S. K.). Use of the Advanced Photon Source was supported by the Office of Science of the U. S. Dept. of Energy.

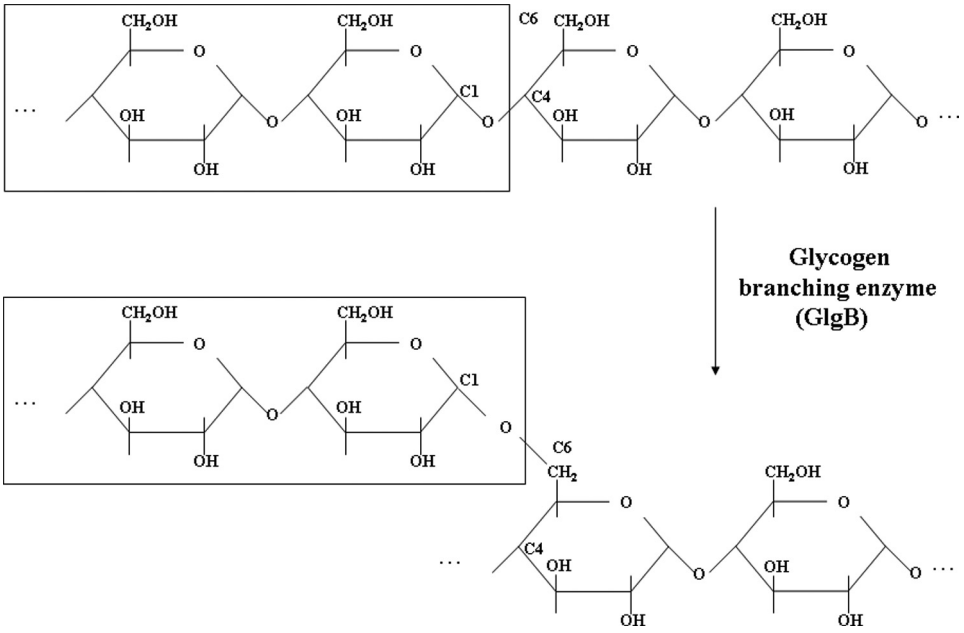
The atomic coordinates and structure factors (code 3K1D) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S3.

[1] Supported by the CSIR-India Network (project SIP-10).
[2] Supported by a fellowship from CSIR, India.

[3] Present address: Vollum Inst., Oregon Health and Science University, Portland, OR 97239.
[4] Present address: Dept. of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556.
[5] To whom correspondence may be addressed. Tel.: 91-172-2695-3264; Fax: 91-172-269-5585; E-mail: pushpa@imtech.res.in.
[6] To whom correspondence may be addressed. Tel.: 65-6516-7932; Fax: 65-6779-2486; E-mail: dbsks@nus.edu.sg.

[7] The abbreviations used are: Mtb, *M. tuberculosis*; GlgB, glycogen branching enzyme; GH13, glycoside hydrolase family 13; MtbGlgB, α(1–4)glucan branching enzyme of *M. tuberculosis* H37Rv; MtbGlgBWT, full-length GlgB of *M. tuberculosis* H37Rv; MES, 4-morpholineethanesulfonic acid; r.m.s.d., root mean square deviation; WT, wild type.



SCHEME 1. **Formation of a glycogen branch.** GlgB catalyzes the transfer of a fragment of 6–7 glucose units from a non-reducing end to the hydroxyl group at C6 of another glucose unit, either in the same chain or an adjacent chain.

functional GlgB, encoded by the ORF Rv1326c, is essential for normal growth of Mtb.

The GlgB is classified under the α-amylase family of enzymes, which includes α-amylase, isoamylase, pullulanase, and cyclodextrin glucanotransferase. The x-ray crystal structures of several α-amylase family enzymes show that they all have a common catalytic $(\beta/\alpha)_8$ barrel (also known as the TIM barrel) core, an N-terminal domain, and a C-terminal domain. Devillers *et al.* (15) have demonstrated that the N-terminal domain of *Escherichia coli* GlgB provides support for glucan substrate during cleavage and transfer of α(1–4)glucan chains. The *M. tuberculosis* H37Rv strain harbors open reading frame Rv1326c, which has been annotated (16) to encode for an α(1–4)glucan branching enzyme. Our previous study (17) has verified that the recombinant α(1–4)glucan branching enzyme of *M. tuberculosis* H37Rv (MtbGlgB) utilizes amylose as substrate.

Any defect or deletion of the GlgB enzyme leads to glycogen storage diseases and GlgB deficiency, a genetic disorder mainly in the American quarter horse. The lack of GlgB is directly linked to type IV glycogenesis (eponym: Andersen disease) (18), which is a rare but fatal genetic disorder in infants (19) and horses (20).

Sequence alignment and structural prediction results have shown that GlgB and isoamylase are structurally very close members of the GH13 family. These enzymes bind to sugars at the 1–6 position. The crystal structure of N-terminal 112-amino acid-deleted *E. coli* GlgB (ECΔ112GlgB) has been reported (21). Even though sequence alignment shows good homology between Mtb and *E. coli* GlgB proteins, there are marked insertions and deletions in both sequences (supplemental Fig. S1). Furthermore, the *E. coli* crystal structure lacks the N-terminal N1 domain (residues 1–105, MtbGlgB numbering). This domain, a β-sandwich, is observed in several GH13 members and is implicated in substrate specificity, recognition, and binding (22).

Along with this background, we still have several unanswered questions like: what are the structural features of the full-length GlgB protein that contribute to substrate specificity and influence its amylase and branching enzymatic activities? What are the structural similarities and differences between the GlgB enzyme from *E. coli* and Mtb? What are the structural features of the GlgB enzyme that could be potentially exploited toward successful therapeutic applications, especially against tuberculosis? To address these questions, we have determined the crystal structure of the full-length protein (MtbGlgBWT) at 2.33-Å resolution. Our present study adds more light to our current knowledge and understanding of this metabolically important enzyme.

EXPERIMENTAL PROCEDURES

Cloning of GlgB the 108-Amino Acid N1 Domain Deletion Mutant—The cloning details of the full-length Mtb *glgB* gene have already been published (17). The gene encoding the N-terminal 108-amino acid deletion mutant of Mtb *glgB* was PCR-amplified using the following primers: forward primer, 5′-AAT TAA TTA GGA TCC ATG ACC CTG GGC GAG GTC GAC CTG-3′ and reverse primer, 5′-ATA TAT CTC GAG CTA GGC GGG CGT CAG CCA CAG C-3′ (BamHI and XhoI restriction sites in the forward and reverse primers, respectively, are underlined) and cloned between the BamHI and XhoI sites of the pET29a vector (Novagen).

Protein Expression and Purification—The full-length MtbGlgBWT protein was expressed and purified as described previously (17) but with some modifications. In brief, the protein was overexpressed in *E. coli* BL21(DE3). Protein solubility was enhanced by the coexpression of the GroEL/GroES chaperons (in the pKY206 vector) under a constitutive promoter. The pKY206 transformed cells with Mtb*glgB* were grown in LB medium, containing tetracycline (12.5 μg/ml) and kanamycin (30 μg/ml), to an A_{600} of 0.6 at 37 °C. Protein expression was induced by 0.3 mM isopropyl β-D-1-thiogalactopyranoside. After induction, the cells were grown initially at 30 °C for 2 h to get enough GroEL/GroES expressed and later the culture was shifted to 16 °C for overnight. The harvested cell pellet was re-suspended in lysis buffer containing 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 25 mM imidazole, 5% glycerol, 1 mM phenylmethylsulfonyl fluoride, and subjected to French-press. The crude lysate was centrifuged at 38,000 × *g* using the JA 25.15 rotor (Beckman) for 30 min at 4 °C, and the supernatant was incubated with Co^{2+} containing Talon resin (Clontech) for affinity purification. The column was washed with wash buffer

(50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 25 mM imidazole), followed by protein elution using buffer containing 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 250 mM imidazole, 5% glycerol. The eluted protein was further purified using gel-filtration chromatography on a HiLoad 26/60 Superdex-200 column (Amersham Biosciences). Column equilibration and protein purification were performed in buffer containing 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 2 mM β-mercaptoethanol, 5% glycerol. MtbGlgBWT was eluted as a monomer in gel filtration. The purified protein was analyzed by SDS-PAGE and native gels for purity and homogeneity. Dynamic light scattering data also indicate a monomeric state of the MtbGlgB protein.

X-ray Crystallography—The MtbGlgBWT protein was concentrated to 12 mg/ml and setup for crystallization using the hanging drop, vapor-diffusion method at room temperature with various commercially available crystallization kits. Crystals first appear after 4 weeks in JB Screen Classic 2 (Jena Bioscience), condition D1 (100 mM MES (pH 6.5), 30% (w/v) polyethylene glycol 4000). Further optimization of this condition (using a matrix of pH 6, 6.5, and 7 and the precipitant concentrations 25%, 30, and 35%) produced crystals of approximate dimensions $0.3 \times 0.1 \times 0.05$ mm in 2 weeks.

A crystal of suitable size was transferred to a cryoprotectant solution containing 100 mM MES (pH 6.5), 25% (v/v) glycerol, 33% (w/v) polyethylene glycol 4000. X-ray diffraction data were collected from a single crystal at temperature 100 K at beamline LS-CAT 21-ID-G of Advanced Photon Source (Chicago, IL) on an MAR Mosaic225 charge-coupled device detector. The diffraction images were processed using HKL2000 (23). The asymmetric unit contains one full MtbGlgB WT molecule (731 amino acids, Mathew's coefficient of 2.56 $Å^3\ Da^{-1}$, solvent content 51.9%). The structure was determined by the molecular replacement method using the Molrep program (24) with the ECΔ112GlgB structure (PDB ID: 1M7X) (21) as the input search model. Model building was carried out with the help of Arp/wARP (25) and O (26). A model for the N1 domain was predicted using the Geno3D server (27) and was subsequently adjusted in electron density. The structure was refined with the CNS program suite (28), and the geometry of the molecule was checked with PROCHECK (29). The correctness of the model, including the N1 domain structure, was verified with the simulated annealing omit maps option of the CNS program. Electron density for the first nine N-terminal amino acids was not observed in the map. The electron density along residues 367–376 and 422–433 is very poorly defined, and some of these residues occupy the disallowed region on the Ramachandran plot. All drawings were prepared with the MOLSCRIPT (30) and RASTER3D (31) programs. Table 1 shows the data and refinement statistics.

Enzyme Activity and Kinetics—The amylase activity of the purified MtbGlgBWT and MtbΔ108GlgB proteins on amylose and starch (as substrates) was assayed as described earlier (17). In brief, a 100-μl reaction mixture containing the required amount of substrate in 50 mM citrate buffer (pH 7.0) and 1.5 μM purified MtbGlgB protein. After 30 min of incubation, at 25 °C, the reaction was stopped using KI/I_2 solution, and the change in optical density was recorded at 660 nm.

TABLE 1

Data collection and refinement statistics

Values in parentheses are for the last shell (2.41-2.33 Å).

Data collection	
Wavelength (Å)	0.98972
Unit-cell parameters (Å)	a = 109.37, b = 156.86, c = 48.02
Space group	$P2_12_12$
Resolution range (Å)	50–2.33
Total no. of observed reflections	524,723
Total no. of unique reflections	36,272
Redundancy	14.4 (15.6)
Completeness (%)	99.9 (100.0)
R_{sym} [a]	0.175 (0.76)
Overall $I/\sigma(I)$	12.9 (2.5)
Refinement	
Resolution range (Å)	20–2.33
R_{work} [b]	0.193
R_{free} [c]	0.236
r.m.s.d. bond lengths (Å)	0.006
r.m.s.d. bond angles (°)	1.3
Average B-factors ($Å^2$)	
Protein atoms (5724 atoms)	35.37
Water molecules (323 atoms)	35.35
Ramachandran plot	
Most favored regions (%)	82.2
Additional allowed regions (%)	15.2
Generously allowed regions (%)	1.0
Disallowed regions (%)	1.8 (from the 367 and 422 loops)

[a] $R_{sym} = \Sigma_{hkl}\Sigma_i[|I_i\ (hkl) - \langle I(hkl)\rangle|]/\Sigma_{hkl}I(hkl)$.

[b] $R_{work} = \Sigma||F_{obs} - F_{calc}||/\Sigma|F_{obs}|$, where F_{calc} and F_{obs} are the calculated and observed structure factor amplitudes, respectively.

[c] R_{free} is same as R_{work}, but 5.0% of the total reflections, chosen at random, were omitted during refinement.

We have also assayed the branching activity of MtbGlgBWT and MtbΔ108GlgB by measuring the amount of glycogen formed. The principle of the assay is: the formed glycogen is converted to glucose (by glucoamylase), which is then specifically oxidized to produce a product that reacts with the OxiRed probe to generate a specific color (570 nm). The glycogen assay kit was purchased from BioVision. To make sure that both the proteins were active, a parallel iodine assay was also performed. 2 μg of amylose was treated with 1.5 μM MtbGlgBWT and MtbΔ108GlgB, for 30 min at 30 °C, in duplicate. A similar reaction was maintained that underwent the iodine test as a control to confirm the utilization of amylose by the proteins. After 30 min, 2 μl of the glucoamylase mixture, provided in the kit, was added, and the sample was incubated again for 30 min at 25 °C. After that, 50 μl of the developing reagent was added to the reaction mixture, and the mixture was incubated at 25 °C for another 30 min before measuring the optical density at 570 nm. To estimate the amount of glycogen formed during this reaction, a glycogen standard curve was prepared by using the glycogen standard provided in the kit. Suitable controls were used throughout. The experiment was repeated with a fresh batch of proteins, and the average value of two different sets of experimental results is provided.

The effects of different known inhibitors were also studied. Stock solutions of ADP, ADP-glucose, castenospermine, nojirimycin, tunicamycin, and acarbose (some of the known inhibitors of the GH13 family of enzymes) were prepared in 50 mM citrate buffer (pH 7.0). To study the effect of these inhibitors, a 100-μl reaction was set up, which contained 0.2 mg/ml amylose in 50 mM citrate buffer (pH 7.0) and then the required amount of the inhibitor. The reaction was started by adding 0.3 μM purified MtbGlgB proteins. After 30 min of incubation, the reaction was stopped using KI/I_2 solution, and the change in

optical density was recorded as described earlier. A suitable buffer control was maintained throughout the study. To study enzyme kinetics, 0.15 μM protein was used, and data were analyzed using SigmaPlot 8.0.

RESULTS

Structural Overview—The crystal structure of the full-length GlgB of *M. tuberculosis* H37Rv (MtbGlgBWT) has been determined at 2.33-Å resolution. As shown in Fig. 1*A*, the GlgB molecule consists of four domains: N-terminal β-sandwich domain N1 (residues 1–105, colored *brown*), N-terminal β-sandwich N2 (residues 106–226, *magenta*), the central $(\beta/\alpha)_8$ domain (residues 227–631, *green*), and the C-terminal β-sandwich domain (residues 631–731, colored *red*). The $(\beta/\alpha)_8$ domain, as the main body of the structure, is sandwiched between the N2 and C-terminal β-sandwiches. It consists of the well characterized $(\beta/\alpha)_8$ barrel with eight alternating β-strands and surrounded by eight α-helices. This domain contains the residues that are involved in substrate binding and catalysis (supplemental Fig. S2). The central tunnel has an approximate minimum diameter of ~13.5 Å (measured between the 254 Cα and 459 Cα atoms). The loops that connect the α-helices and β-strands in this domain do not obstruct the tunnel. This clearance makes the entry of a long linear glucose chain substrate into the active site and the exit of branched glucose chains very easy. A DALI (32) search with the central $(\beta/\alpha)_8$ domain (residues 226–631) identified maltooligosyltrehalose trehalohydrolase (PDB: 2BY2), glycosyltrehalose trehalohydrolase (1EHA), and maltogenic amylase (1GVI) as the three top superimposing GH13 family structures with MtbGlgB (apart from *E. coli* GlgB), with *Z* values 36.7, 35.0, and 34.2, respectively. Their corresponding r.m.s.d. vales were 2.3 Å for 239 Cα atoms, 2.5 Å for 223 Cα atoms, and 2.6 Å for 265 Cα atoms, respectively.

Abad *et al.* (21) have given a clear description of the active site of ECΔ112GlgB. In short, the catalytic active site (Fig. 1*B*) superimposes well with that of other GH13 family members. It contains the seven residues (Asp-341, His-346, Arg-409, Asp-411, Glu-464, His-531, and Asp-532, MtbGlgB numbering) that are highly conserved in all members of α-amylase family 13. The overall surface around the active site pocket in ECΔ112GlgB, similar to that of isomaltulose synthase (also known as PalI, Uniprot entry: Q8KR84, another GH13 family member) (33), is highly negatively charged. This highly negative character of the $(\beta/\alpha)_8$ barrel domain is important for sugar-protein interactions.

The N-terminal β-sandwich N1 domain (Fig. 1*C*) is very similar to the N2 domain. These two domains superimpose with an r.m.s.d. of 1.5 Å for 95 Cα atoms. The topology of this domain is a typical immunoglobulin fold. The two β sheets of the sandwich are formed by strands β1, β2, β5, and β4 and strands β3, β6, and β7, respectively. The top two hits in a DALI search, using the MtbGlgB N1 and N2 domains, are pullulanase (PDB: 2FH8, Z = 11.6 and r.m.s.d. = 2.7 for 217 Cα) and the pullulanase type I protein (2E8Z/11.0/3.7 for 205 residues).

Even though the ECGlgB (residues 117–728) and MtbGlgB (residues 106–731) structures superimpose in their corresponding domains with an overall r.m.s.d. of 1.12 Å for 553 Cα atoms (Fig. 1*D*), there are some local misalignments, especially in the coil region between residues Thr-230 and Pro-250. Also, due to unsuccessful attempts to crystallize the full protein, the *E. coli* GlgB structure represents only the N1 domain truncated form. Based on the significant level of sequence homology between the N1 domains (29% identity and 49% similarity) of the two proteins, we can expect that the N1 domain of *E. coli* GlgB should also have the same β-sandwich topology.

Enzyme Activity/Kinetics—The GlgB enzyme catalyzed two reactions, concurrently, in the same active site. A long glucose chain was cleaved at the 1→4 glycosidic bond (amylase activity) and a branching glucose chain was formed by making a new 1→6 glycosidic bond (glycogen formation branching reaction). We measured the amylase as well as glycogen-forming efficiencies of MtbGlgBWT and its N1 domain deletion mutant (MtbΔ108GlgB). In the amylase activity study, when amylose was used as the substrate, the specific activity of MtbGlgB WT was 63.75 units/mg, whereas that for MtbΔ108GlgB was 42 units/mg of protein. Similar to the specific activity, the rate of reaction for MtbGlgBWT protein was also faster: V_{max} = 1.35 mg ml^{-1} min^{-1}, where V_{max} is expressed in milligrams of product formed by 1.5 μM enzyme; K_m = 0.56 mg/ml; K_{cat} = 9 min^{-1}. Also, MtbGlgB has lower amylase activity when compared with that reported for *E. coli* GlgB. The corresponding values for MtbΔ108GlgB were V_{max} = 0.62 mg ml^{-1} min^{-1}, K_m = 0.33 mg/ml, and K_{cat} = 4.13 min^{-1} (Fig. 2, *A* and *B*). However, the enzymatic assay using ECGlgB (21) reports equal K_m values for the WT and Δ112 proteins. Even though the amylase activity of MtbΔ108GlgB was lower when amylose was used as substrate, both proteins showed almost equal amylase activity when soluble starch was used as substrate.

Next we tested the branching activity of MtbGlgB using a glycogen assay kit. The amount of glycogen formed by MtbGlgBWT was estimated from a glycogen standard curve to be 1.46 μg, whereas the value for MtbΔ108GlgB was 1.44 μg. 2 μg of amylose and 1.5 μM protein was used in the reaction.

BAY e4609 has been shown to be an effective inhibitor for the *E. coli* GlgB enzyme (21, 34, 35). Nonetheless, we wanted to test the effect of some of the known inhibitors of the GH13 family proteins on MtbGlgB. Inhibitors such as ADP, ADP glucose, tunicamycin, castenospermine, nojirimycin, or acarbose had no effect on the enzymatic activity of either MtbGlgBWT or MtbΔ108GlgB, when amylose was used as substrate (supplemental Fig. S3). This clearly suggests the need for more inhibitor screening experiments to identify additional inhibitors, apart from BAY e4609.

DISCUSSION

Influence of the N1 Domain—Recently, Palomo *et al.* (36) have studied the influence of the N-terminal domains of the *Deinococcus* glycogen branching enzymes on the unique glycogen-branching patterns. The present study shows that the N1 β-sandwich has a differential preference in substrate recognition and binding during amylase activity, when amylose is used as a substrate. The N1-truncated MtbΔ108GlgB protein is ~50% less active. However, to our surprise, there was no difference in substrate binding or amylase activity when starch was used as a substrate or in the amount of glycogen formed during

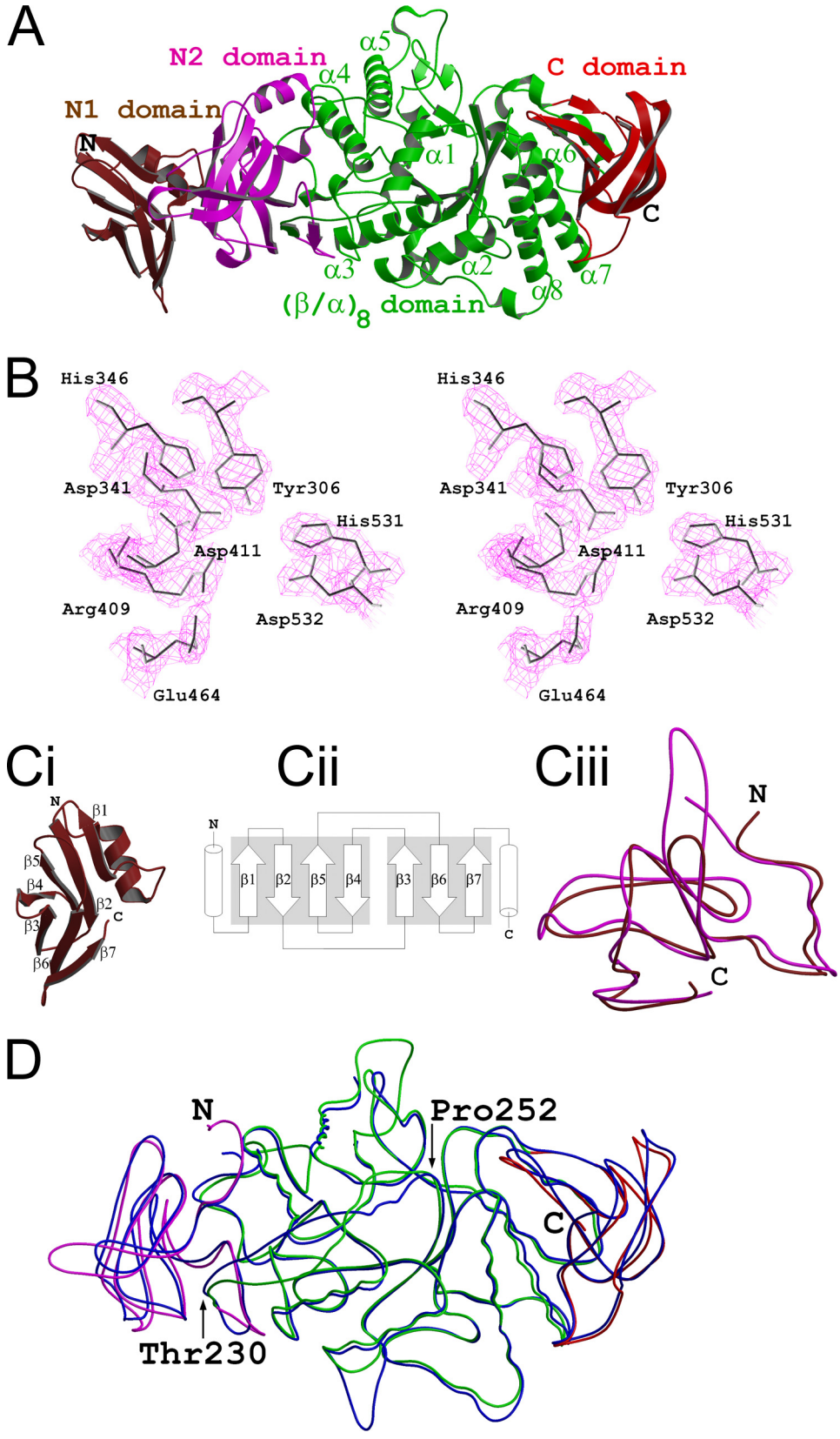


FIGURE 1. **Overall structure.** *A*, the MtbGlgBWT protein contains four domains: the N1 β-sandwich domain (residues 1–105, colored *brown*), the N2 β-sandwich domain (106–226, *magenta*), the central $(\beta/\alpha)_8$ catalytic domain (227–630, *green*) and the C-terminal β-sandwich (632–731, *red*). *B*, electron density around the catalytic pocket formed by the residues that are conserved in the GH13 family (MtbGlgB numbering). The $2F_o - F_c$ map is contoured at the 1.5 σ level. In *C*: *I*, the N-terminal N1 domain has the immunoglobulin sandwich fold. *ii*, the topological arrangement of the seven β-strands. The two sheets of the sandwich are made by strands β1, β2, β5, and β4 and strands β3, β6, and β7, respectively. Each of the flanking helices comes from the N1 and N2 domains, respectively. *iii*, superimposition of the N1 (*brown*) and N2 (*magenta*) domains of MtbGlgB. The r.m.s.d. is 1.5 Å for 95 Cα atoms. *D*, superimposition of MtbGlgB (residues 106–731 and *domain colors* as in *A*) and ECΔ112GlgB (113–528, *blue*). The r.m.s.d. is 1.12 Å for 553 Cα atoms. Note the region between residues 230 and 250 (MtbGlgB numbering) is misaligned very much.

branching activity. Thus it is difficult to draw the same conclusion as reported in the ECΔ112GlgB study about substrate preference. The existence of an additional N-terminal β-barrel domain in an α-amylase has been reported for AmyB from *Halothermothrix orenii* (37). However, there are marked structural differences between the N-terminal domains of AmyB and MtbGlgB.

Catalytic Mechanism—The substrate recognition scheme and binding sites for several amylase family enzymes, like PalI (33), AmyA from *Halothermothrix orenii* (38), α-amylase (39), CGTase (40), TAKA-amylase with substrate analogs (41), amylosucrase with D-glucose and mutated amylosucrase with sucrose (42, 43), have been identified. A substrate binding model has been proposed for ECΔ112GlgB (21). GlgB enzymes from different organisms transfer chains of different lengths (the number of glucose units). We propose a two-step reaction mechanism, for the amylase reaction (breaking a 1→4 glycosidic bond) and isomerization (the branching reaction, by making a new 1→6 glycosidic bond), which happens in the same catalytic pocket (Scheme 2). The potential catalytic triad (Asp^{411}, Glu^{464}, and Asp^{532}) and two histidine residues (His^{346} and His^{531}) of GlgB are highly conserved in almost all α-amylase and glycosyltransferase enzymes. These residues form a catalytic pocket (Fig. 1*B*) that binds amylose and breaks the 1→4 glycosidic bond. The similarity of the active site architecture strongly suggests that the amylase reaction occurs via a general acid catalysis, as in all glucoside hydrolases (43). Glu^{464} acts as the general acid catalyst to protonate the oxygen of the 1→4 glycosidic linkage for the amylase reaction. Asp^{411}, the attacking nucleophile, forms a bond with C1 to form a β-glucosyl-enzyme intermediate, which is presented to the OH group at C6 of another glucose unit, either on the same or another chain to form a 1→6 glycosidic bond.

Truncation Mutants—Three MtbGlgB N-terminal deletion mutants (at residues 112, 121, and 171) were overexpressed, purified, and tested for enzymatic activity. As expected (data not shown), these proteins were less active in proportion to the number of amino acids truncated. Interestingly, the deletion also affects the amount of protein expression. In PalI (33), we

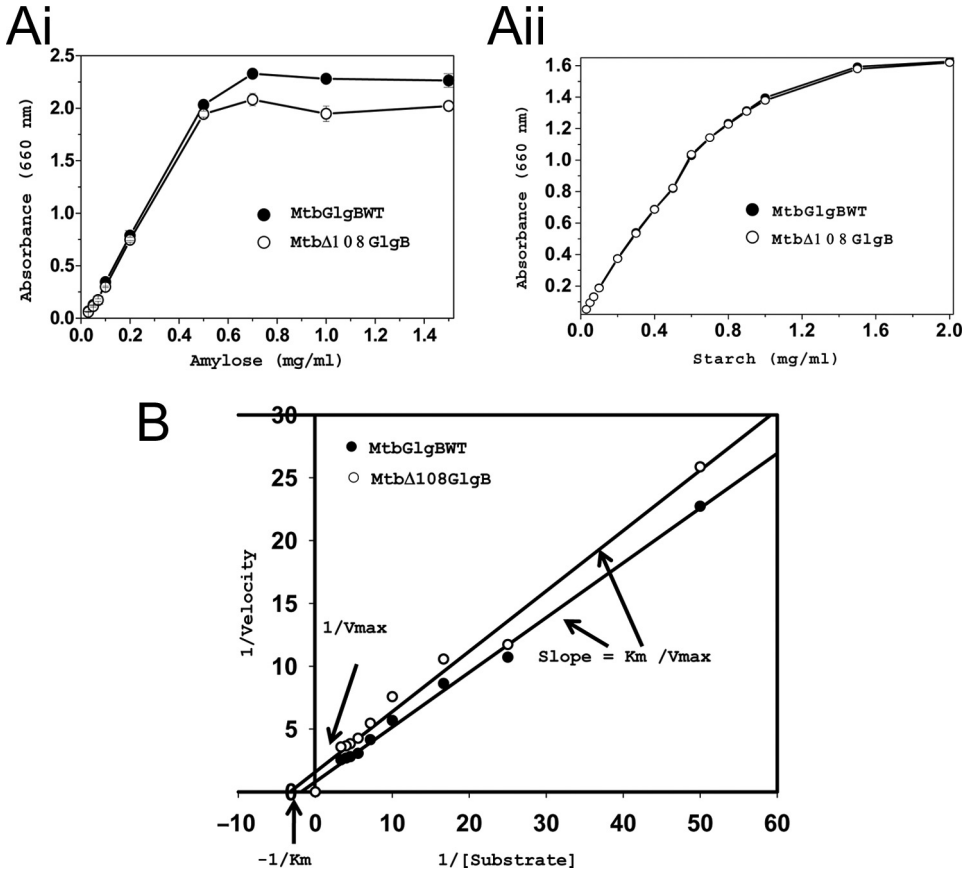
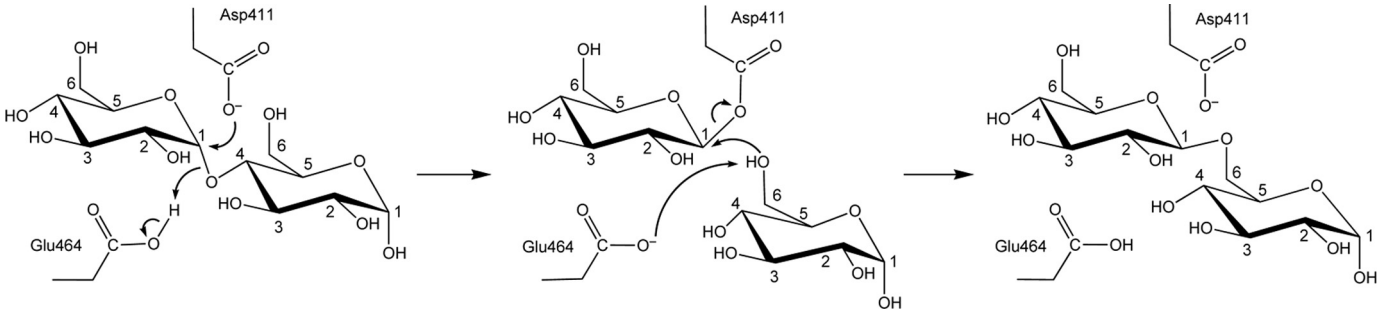


FIGURE 2. **Enzyme assays and kinetic studies.** *A*, substrate utilization curves for MtbGlgBWT and MtbΔ108GlgB proteins with amylose (*i*) and starch (*ii*) as substrates. Substrates were prepared in 50 mM citrate buffer, pH 7.0. In a 100-μl reaction mixture, the required amount of substrate was incubated with 1.5 μM purified protein with different concentrations of substrate and incubated at 25 °C for 30 min. The reaction was stopped by adding KI/I_2 solution, and absorbance was recorded at 660 nm. *B*, Lineweaver-Burke plots for MtbGlgBWT and MtbΔ108GlgB. The protein concentration was 0.15 μM, whereas the substrate concentration was varied between 0.02 and 0.3 mg/ml at regular intervals.



SCHEME 2. **Mechanism of glycogen branching.** *Panel 1*, Glu^{464} acts as a general acid catalyst to protonate approximately the 7th glycosidic oxygen from a non-reducing end. Asp^{411}, the attacking nucleophile, forms a glucosyl-enzyme intermediate. *Panel 2*, The substrate fragment can be transferred to the hydroxyl group at C6 of a glucose unit on the same or another chain for the 1→6 bond formation step. *Panel 3*, view of a glycogen branch point.

had experimentally shown that selected truncations at the C-terminal domain showed reduction in activity (again, proportional to the length deleted). However, our attempt to prepare three C-terminal truncation mutants for MtbGlgB did not express any protein. Based on our PalI results, we predict that in MtbGlgB also the central catalytic domain might undergo conformational changes, and the active site gradually becomes inactive when the N- and C-terminal domains are truncated and the presumable structural pressure is relieved.

Clinical Significance—The evolution of drug-resistant strains of Mtb has created a significant concern. Mutational studies using the H37RvΔ*glgB*/pVV*glgBtb* strain have shown the inability of the bacterium to survive, and this promising result should be further exploited to make use of GlgB as an attractive therapeutic target against Mtb. Also, comparison of the sequences of GlgB from several pathogenic and non-pathogenic bacterial genera clearly presents a case of domain closeness among them (supplemental Fig. S1). This sequence similarity indicates a good possibility of potential drugs, based on the MtbGlgB structure, for treatment against a wide spectrum of bacterial infection. A recent screening study for inhibitors against human pancreatic α-amylase (34) has shown that the two potential lead compounds are very ineffective against other closely related GH13 members. Binderup *et al.* (35) have used BAY e4609, a pseudo oligosaccharide in a branching enzyme inhibition study. However, many known GH13 family inhibitors have no effect on either MtbGlgBWT or MtbΔ108GlgB. This warrants a more thorough drug screening against this protein, which is very likely a potential and important drug target in the fight against Mtb.

Acknowledgments—We thank J. S. Brunzelle for assistance in data collection at the beamlines of sector 21 (LS-CAT). We also thank Chu-Young Kim and Kinya Hotta for their help in redrawing Scheme 2.

REFERENCES

1. Chapman, G. B., Hanks, J. H., and Wallace, J. H. (1959) *J. Bacteriol.* **77,** 205–211
2. Hanks, J. H. (1961) *Int. J. Leprosy* **29,** 74–83
3. Daffé, M., and Draper, P. (1998) *Adv. Microb. Physiol.* **39,** 131–203
4. Ortalo-Magné, A., Dupont, M. A., Lemassu, A., Andersen, A. B., Gounon, P., and Daffé, M. (1995) *Microbiology* **141,** 1609–1620
5. Lemassu, A., Ortalo-Magné, A., Bardou, F., Silve, G., Laneélle, M. A., and Daffé, M. (1996) *Microbiology* **142,** 1513–1520
6. Besra, G. S., and Chatterjee, D. (1994) in *Lipids and Carbohydrates of M. tuberculosis* (Bloom, B. R., ed) pp. 285–306, American Society for Microbiology, Washington, D. C.
7. Lemassu, A., and Daffé, M. (1994) *Biochem. J.* **297,** 351–357
8. Dinadayala, P., Lemassu, A., Granovski, P., Cérantola, S., Winter, N., and Daffé, M. (2004) *J. Biol. Chem.* **279,** 12369–12378
9. Cywes, C., Hoppe, H. C., Daffé, M., and Ehlers, M. R. (1997) *Infect. Immun.* **65,** 4258–4266
10. Gagliardi, M. C., Lemassu, A., Teloni, R., Mariotti, S., Sargentini, V., Pardini, M., Daffé, M., and Nisini, R. (2007) *Cell Microbiol.* **9,** 2081–2092
11. Stokes, R. W., Norris-Jones, R., Brooks, D. E., Beveridge, T. J., Doxsee, D., and Thorson, L. M. (2004) *Infect. Immun.* **72,** 5676–5686
12. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) *Nucleic Acids Res.* **37,** D233–238
13. Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2003) *Mol. Microbiol.* **48,** 77–84
14. Sambou, T., Dinadayala, P., Stadthagen, G., Barilone, N., Bordat, Y., Constant, P., Levillain, F., Neyrolles, O., Gicquel, B., Lemassu, A., Daffé, M., and Jackson, M. (2008) *Mol. Microbiol.* **70,** 762–774
15. Devillers, C. H., Piper, M. E., Ballicora, M. A., and Preiss, J. (2003) *Arch. Biochem. Biophys.* **418,** 34–38
16. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* **393,** 537–544
17. Garg, S. K., Alam, M. S., Kishan, R. K., and Agrawal, P. (2007) *Prot. Express Purif.* **51,** 198–208
18. Brown, B. I., and Brown, D. H. (1966) *Proc. Natl. Acad. Sci. U.S.A.* **56,** 725–729
19. Andersen, D. H. (1956) *Lab. Invest.* **5,** 11–20
20. Valberg, S. J., Ward, T. L., Rush, B., Kinde, H., Hiraragi, H., Nahey, D., Fyfe, J., and Mickelson, J. R. (2001) *J. Vet. Intern. Med.* **15,** 572–580
21. Abad, M. C., Binderup, K., Rios-Steiner, J., Arni, R. K., Preiss, J., and Geiger, J. H. (2002) *J. Biol. Chem.* **277,** 42164–42170
22. Machovic, M., and Janecek, S. (2006) *Cell. Mol. Life Sci.* **63,** 2710–2724
23. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
24. Vagin, A., and Teplyakov, A. (2000) *Acta Crystallogr. D Biol. Crystallogr.* **56,** 1622–1624
25. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463
26. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
27. Combet, C., Jambon, M., Deléage, G., and Geourjon, C. (2002) *Bioinformatics* **18,** 213–214
28. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54,** 905–921
29. Laskowsi, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
30. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
31. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol.* **277,** 505–524
32. Holm, L., Kääriämen, S., Rosenström, P., and Schenkel, A. (2008) *Bioinformatics* **24,** 2780–2781
33. Zhang, D., Li, N., Lok, S. M. Zhang, L. H., and Swaminathan, K. (2003) *J. Biol. Chem.* **278,** 35428–35434
34. Tarling, C. A., Woods, K., Zhang, R., Brastianos, H. C., Brayer, G. D., Andersen, R. J., and Withers, S. G. (2008) *Chembiochem* **9,** 433–438
35. Binderup, K., Libessart, N., and Preiss, J. (2000) *Arch. Biochem. Biophys.* **374,** 73–78
36. Palomo, M., Kralj, S., van der Maarel, M. J., and Dijkhuizen, L. (2009) *Appl. Environ. Microbiol.* **75,** 1355–1362
37. Tan, T. C., Mijts, B. N., Swaminathan, K., Patel, B. K., and Divne, C. (2008) *J. Mol. Biol.* **378,** 852–870
38. Sivakumar, N., Li, N., Tang, J. W., Patel, B. K., and Swaminathan, K. (2006) *FEBS Lett.* **580,** 2646–2652
39. Uitdehaag, J. C., Mosi, R., Kalk, K. H., van der Veen, B. A., Dijkhuizen, L., Withers, S. G., and Dijkstra, B. W. (1999) *Nat. Struct. Biol.* **6,** 432–436
40. Brzozowski, A. M., and Davies, G. J. (1997) *Biochemistry* **36,** 10837–10845
41. Mirza, O., Skov, L. K., Remaud-Simeon, M., Potocki de Montalk, G., Albenne, C., Monsan, P., and Gajhede, M. (2001) *Biochemistry* **40,** 9032–9039
42. Skov, L. K., Mirza, O., Sprogøe, D., Dar, I., Remaud-Simeon, M., Albenne, C., Monsan, P., and Gajhede, M. (2002) *J. Biol. Chem.* **277,** 47741–47747
43. Koshland, D. E. (1953) *Biol. Rev. Camb. Philos Soc.* **28,** 416–436