The Structure and Specificity of *Escherichia coli* Maltose Acetyltransferase Give New Insight into the LacA Family of Acyltransferases[†,‡]

Leila Lo Leggio,[◇,§] Florence Dal Degan,[◇,∥,⊥] Peter Poulsen,[∥,#] Søren Møller Andersen,[▽,△] and Sine Larsen*[,§]

Centre for Crystallographic Studies, Department of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark, Danisco Innovation Copenhagen, Langebrogade 1, 1011 Copenhagen K, Denmark, and Department of Organic Chemistry, Building 201, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

Received November 8, 2002; Revised Manuscript Received February 14, 2003

ABSTRACT: The crystallographic three-dimensional structure of the *Escherichia coli maa* gene product, previously identified as a maltose *O*-acetyltransferase (MAT) [Brand, B., and Boos, W. (1991) *J. Biol. Chem. 266*, 14113−14118] has been determined to 2.15 Å resolution by the single anomalous dispersion method using data from a crystal cocrystallized with trimethyllead acetate. It is shown here that MAT acetylates glucose exclusively at the C6 position and maltose at the C6 position of the nonreducing end glucosyl moiety. Furthermore, MAT shows higher affinity toward artificial substrates containing an alkyl or hydrophobic chain as well as a glucosyl unit. The presence of a long hydrophobic patch near the acceptor site provides the structural explanation for this preference. The three-dimensional structure reveals the expected trimeric left-handed parallel β-helix structure found in all other known hexapeptide repeat enzymes. In particular, the structure shows similarities both overall and at the putative active site to the recently determined structure of galactoside acetyltransferase (GAT), the *lacA* gene product [Wang, X.-G., Olsen, L. R., and Roderick, S. L. (2002) *Structure 10*, 581−588]. The structure, together with the new biochemical data, suggests that GAT and MAT are more closely related than previously thought and might have similar cellular functions. However, while GAT is specific for acetylation of galactosyl units, MAT is specific for glucosyl units and is able to acetylate maltooligosaccharides, an important property for biotechnological applications. Structural differences at the acceptor site reflect the differences in substrate specificity.

The *maa* gene product was first identified in a study of mutants in the maltose transport system of *Escherichia coli*. Boos and co-workers (*1*) showed that the compound leaking from amylomaltase-deficient cells accumulating maltose was acetylmaltose. They further showed that the enzymatic activity responsible for maltose acetylation was distinct from the *lacA* gene product, galactoside *O*-acetyltransferase (GAT). Later, the same group cloned the *maa* gene and purified its gene product to homogeneity, leading to partial biochemical characterization (*2*). Acetyl-CoA[1] was identified as an efficient donor, while it was established that glucose, maltose, and to a much lesser extent maltooligosaccharides could function as acceptors, although with rather elevated K_m, in the tens of millimolar range. Monosaccharide specificity was also established, which showed that glucose was by far the best substrate.

Our initial interest in MAT emerged from a desire to produce acetylated starch *in planta*. Acetylated starch types, which have wide applications in both food and nonfood industries, are produced by chemical acetylation of postharvest starches in the traditional starch industry. Therefore, it would be desirable to produce enzymatically acetylated starch types directly in the plants, since the costs for downstream processing would be reduced. MAT seemed to be a good starting point for studies of a glucan modifying enzyme of this type, and this initiated the biochemical and structural characterization reported here.

† This research has been funded by the Danish National Research Foundation and EMBO through a long-term fellowship to L.L.L.

‡ The coordinates and structure factors for the MAT structure have been deposited in the Protein Data Bank with codes 1OCX and R1OCXSF.

* Corresponding author. Telephone: (+45)35320282. Fax: (+45)-35320299. E-mail: sine@ccs.ki.ku.dk.

◇ These authors contributed equally to this work.

§ University of Copenhagen.

∥ Danisco Innovation Copenhagen.

⊥ Present address: Pharmexa, Kogle Allé 6, DK-2970 Hørsholm, Denmark.

Present address: Department of Physiology, Carlsberg Laboratory, Gamle Carlsberg Vej 10, 2500 Valby, Denmark.

▽ Technical University of Denmark.

△ Present address: Department of Chemistry, University of Alberta, Edmonton, Alberta T6E 2G2, Canada.

[1] Abbreviations: MAT, maltose *O*-acetyltransferase; GAT, galactoside *O*-acetyltransferase; SAD, single-wavelength anomalous dispersion; LβH, left-handed parallel β-helix; EDTA, ethylenediaminetetraacetic acid; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); RP-HPLC, reverse-phase high-performance liquid chromatography; IEF, isoelectric focusing; CAPS, 3-(cyclohexylamino)-1-propanesulfonic acid; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; MOPS, 3-(*N*-morpholino)propanesulfonic acid; MAD, multiwavelength anomalous dispersion; NCS, noncrystallographic symmetry; *R*-factor, crystallographic *R*-factor for the working set; *R*-free, crystallographic *R*-factor for the test set; rmsd, root mean square deviation; TDT, tetrahydrodipicolinate succinyltransferase; PaXAT, *Pseudomonas aeruginosa* xenobiotic acetyltransferase; GlmU, *N*-acetylglucosamine-1-phosphate uridyltransferase; LpxA, UDP-*N*-acetylglucosamine acetyltransferase; IPT, isopropyl β-D-thio; CoA, coenzyme A; pNP, *p*-nitrophenyl; NodL, nodulation protein L, a chitin oligosaccharide *O*-acetyltransferase.

10.1021/bi0271446 CCC: $25.00
Published on Web 04/12/2003

The *maa* sequence strongly supports that this enzyme belongs to the hexapeptide repeat family of proteins (*3*). Discovered about 10 years ago, this family comprises proteins with regions characterized by a repeated hexapeptide motif containing isoleucine, leucine, or valine at its first position. Although most members of the hexapeptide repeat family are acyltransferases, large variation is found in substrate specificity, and some members of the family are carbonic anhydrases. Several structures are now available for this family of enzymes (*4*−*13*), and they all share the same overall architecture, a trimer mediated by a left-handed parallel β-helix (LβH) domain (for a recent review on parallel-β helix structures see ref *14*). The functional differences are reflected by large structural differences in the domains external to the LβH, in the number of β-helix turns present, and in the nature and length of loops protruding from the LβH domain. We have previously reported the trimeric organization of MAT in solution, its crystallization, and evidence that it forms a β-helix based on X-ray diffraction data (*15*). Here we report the full structure determination of MAT.

Among the LβH enzymes for which the structure is known, the one with highest sequence identity to MAT is GAT (*12*) (41% sequence identity after structure-based sequence alignment). This is reflected by very high structural similarity also in the non-LβH regions, whereas clear differences prevail at the acceptor site. A structural comparison of MAT with GAT, combined with new biochemical evidence, suggests that the *lacA* and *maa* gene products might be more similar in function and catalytic mechanism than previously thought but clearly differ in acceptor specificity. This difference confers MAT its desirable biotechnological property of acetylating maltooligosaccharides.

MATERIALS AND METHODS

Cloning, Overexpression, and Purification of MAT. Cloning of the *E. coli maa* gene, expression, and purification of recombinant MAT were described previously (*15*). The purification is similar to the one described by Boos and coworkers (*2*).

Determination of Enzyme Concentration and Activity. The concentration of solutions of pure MAT was estimated spectrophotometrically at 280 nm using an extinction coefficient of 13370 M^{-1} cm^{-1} as determined from the amino acid composition of MAT according to Gill and von Hippel (*16*). The acetyltransferase activity of MAT was assayed spectrophotometrically according to a modified Alpers' assay (*17*). The assay mixture, with a total volume of 1 mL, contained 50 mM potassium phosphate, 2 mM EDTA buffer at pH 7.5, 100 μL of 1 M maltose, 100 μL of 0.4 mM acetyl-CoA, 10 μL of 40 mM DTNB dissolved in methanol, and 10 μL of enzyme. The reaction was initiated by the addition of enzyme or maltose and was monitored at 412 nm at 25 °C using a Perkin-Elmer Lambda 18 UV−vis spectrophotometer. One activity unit was defined as the amount of enzyme that produced an increase in absorbance of 1 per minute at 25 °C. An extinction coefficient of 13600 M^{-1} cm^{-1} at 412 nm was used for DTNB in order to calculate the consumption of acetyl-CoA.

Physicochemical Characterization of MAT. N-Terminal sequencing of pure MAT was performed using an Applied Biosystems 476A protein sequencer. One nanomole of protein was desalted by RP-HPLC on a C2 column (4.6/30) prior to its loading onto the sequencer.

The purity and molecular mass of MAT were determined by SDS−PAGE under reducing conditions in 15% homogeneous gels and by MALDI-TOF mass spectrometry on a Voyager DE instrument from PerSeptive Biosystems. The isoelectric point of MAT was determined by focusing on a PhastGel IEF 4-6.5 (Amersham Bioscience). The stability of MAT at various pH values was investigated by incubating the enzyme at a concentration of 15.3×10^{-6} M in 50 mM buffer containing 100 mM NaCl. The buffers used were sodium acetate (pH 3.0−5.5), Bis-Tris, potassium phosphate, sodium phosphate, HEPES, MOPS, Tricine or Trizma (pH 5.0−9.5), and CAPS (pH 9.0−10.0). The pH of the enzyme mixture was always checked with a pH meter at the relevant temperature and found to be satisfactory. Aliquots for activity determination were taken after different incubation times. The residual activity of MAT was measured as described above. The stability of MAT at various temperatures was investigated in a similar way by incubating the enzyme at a concentration of 15.3×10^{-6} M in 50 mM potassium phosphate, pH 7.5, containing 2 mM EDTA at temperatures ranging from 40 to 70 °C.

Enzymatic Properties of MAT. The pH profile of MAT was determined using the assay described above but varying the pH between 4.0 and 9.0. The same buffer series as above was used. To determine kinetic constants, the concentrations of both the acetyl acceptor and acetyl-CoA were varied in the standard assay procedure. Kinetic constants were derived from double reciprocal plots.

NMR Structure Determination of the Acetylated Reaction Product. The nature of the acetylated products obtained from the MAT-catalyzed reaction were investigated using either glucose or maltose as the acetyl acceptor by ^{1}H NMR on a Bruker AM 500 MHz spectrometer. Ten milligrams of acetyl acceptor, 17.2 nmol of [^{14}C]acetyl-CoA, and MAT at a concentration of 0.15×10^{-6} M were incubated in 50 mM potassium phosphate, pH 7.5, containing 2 mM EDTA for 48 h at 4 °C. A total amount of 10 mg of acetyl-CoA was added in 1 mg aliquots at different time intervals. Extra enzyme aliquots were added as well. The reactants were then separated by thin-layer chromatography on silica gel plates, 60 F254 (Merck), developed with 1-propanol/water (in a ratio of 7:1). The plates were autoradiographed, and acetylated products were eluted from the plates with 80% ethanol. The samples (2 mg of each) were lyophilized twice, redissolved in D_2O, and analyzed.

Crystallization. Cocrystals of MAT with trimethyllead were used for structure determination. This crystal form was previously described (*15*) as crystal form IV. These crystals grow after dissolution of crystals of different habit (previously referred to as crystal form III) grown in hanging drops consisting of 3 μL of reservoir (0.1 M sodium citrate, pH 5.6, 3% ethanol, 15% PEG 4000) and 1 μL of protein solution (15 mg mL^{-1}). To a drop containing fully formed form III crystals, 6 μL of reservoir and 1 μL of 50 mM trimethyllead acetate solution were added. This caused the form III crystals to dissolve and the form IV crystals to grow. Form IV crystals belong to space group $C222_1$ with a = 64.7 Å, b = 106.6 Å, and c = 176.0 Å.

Table 1: Data Collection and Phasing Statistics[a]

	ESRF		ELETTRA
resolution range (Å)	15.00−2.15 (2.23−2.15)		20.0−2.40 (2.44−2.40)
measurements	120930		117029
unique reflections	33084	62694[b]	45220[b]
completeness (%)	99.7 (99.8)	99.0 (98.7)[b]	99.9 (99.9)[b]
R_{merge}	0.061 (0.396)	0.049 (0.350)[b]	0.050 (0.206)[b]
reflections with $I/\sigma(I) > 2$ (%)	84.6 (58.9)	79.9 (48.8)[b]	90.2 (74.2)[b]
no. of sites identified by SOLVE (15−3.0 Å)		4	5
SOLVE *Z*-score (15−3.0 Å)		10.46	11.26
SOLVE figure of merit (15−3.0 Å)		0.24	0.27

[a] The ELETTRA data set was used for initial phasing and refinement, while the ESRF data set was used as native in the later stages of refinement. [b] Indicates statistics for data processing where Friedel mates were kept separate.

Data Collection and Structure Determination. Two independent three-wavelength MAD data sets were collected at the LIII edge of lead (0.95 Å) from two form IV crystals, one at beamline ID14-4, ESRF, Grenoble, France, and one at the X-ray crystallography beamline, Elettra Synchrotron, Trieste, Italy. Difficulties in obtaining good quality fluorescence scans prior to data collections were experienced, in one case also due to interference from lead components of the beamline. Selection of the appropriate wavelengths was therefore problematic. Both crystals were cryocooled at 100 K after they were transferred to cryoprotectant (a mixture of reservoir and ethylene glycol in a 17:3 ratio). The data were processed using DENZO/SCALEPACK (*18*), scaling the Friedel-related reflections independently to best preserve the anomalous signal. Data collection statistics are shown in Table 1.

Initial attempts to phase the structure by MAD using SOLVE (*19*) were unpromising. Refined values of f' and f'' in SOLVE were inconsistent with the intended choices of wavelengths for both data sets. Using a variety of resolution ranges, some solutions with *Z*-scores between 10 and 30 were identified, but the figures of merit were low, often below 0.3. Typically, several solutions were identified in each run, with some shared sites between them. Lack of a promising set of sites and the disturbing refined values of f' and f'' led us to pursue the structure determination by the SAD method instead. The data sets for the intended peak wavelengths for the two crystals were input in SOLVE to search for the lead sites and SAD phasing. For each data set SOLVE easily identified one single solution, and the two solutions were consistent, with *Z*-scores of around 10, although the figures of merit were still below 0.3. Five heavy atom sites were identified in the ELETTRA data set and four in the ESRF data set. SOLVE phasing statistics (Table 1) and map quality were superior in the lower resolution data set collected at ELETTRA, probably due to its higher anomalous redundancy (2.6 as opposed to 1.9 for the ESRF data set at the full resolution). The ELETTRA data set was used for initial tracing and refinement of the model, and later the model was refined against the data set collected at the ESRF. For this latter refinement, the ESRF data set was used as a native after reprocessing of the data merging the Friedel-related reflections (Table 1).

The cell volume was consistent with the presence of two to five molecules in the asymmetric unit, so a trimer per asymmetric unit was expected. The CCP4 program GETAX (*20*) easily identified the expected noncrystallographic 3-fold axis relating the three monomers in the asymmetric unit in a DM (*21*) density-modified map. Phased molecular replacement in MOLREP (*22*) could successfully position three monomers of tetrahydrodipicolinate succinyltransferase [TDT (*6*)], the closest model available at the time (about 28% sequence identity). Only the LβH domain of TDT was used as a search model (residues 102−166 and 176−234). The positions of the monomers were consistent with the axis identified by GETAX. NCS (noncrystallographic symmetry) parameters and an initial mask were created on the basis of the positioned TDT monomers, and density modification was repeated including density averaging. An initial model was built by cutting and pasting the best fitting parts of the positioned TDT monomers and of the model automatically obtained by the SOLVE/RESOLVE autotracing routine (*23*) and by manual building in O (*24*). CNS (*25*) was used for initial refinement against the ELETTRA data set. The initial model consisted of 113 residues in each of three identical monomers and had an *R*-factor of 46.7% and an *R*-free of 51.2% after one cycle of simulated annealing in CNS to 2.8 Å resolution. Several cycles of manual rebuilding and refinement in CNS were carried out. When the *R*-free reached 49.5%, a rigid body refinement was carried out against the higher resolution ESRF data set, which was used from then on for refinement and map calculation (maintaining the same free *R* set of reflections). Final refinement cycles in CNS included insertion of five trimethyllead molecules and water molecules in the model. The *R*-factor and *R*-free were at this stage 22.4% and 26.3%, respectively. The *R*-factor and *R*-free were a little higher than could be expected at this resolution and the Ramachandran quality a little lower (87.3% of the residues in the core region, with some residues in the generously allowed regions). Also, the number of water molecules found was somewhat low (233 in total). This raised concerns that refinement might be hampered by the large differences in *B*-factors between the three monomers and between different regions in each monomer (Table 2), which meant NCS restraints had to be fairly relaxed. To overcome this problem, refinement was completed using REFMAC (version 5 in CCP4 v. 4.1.1) (*26*), which provides the option of separate TLS refinement of different domains. The N-terminal region, loop region, and LβH region of each monomer were treated as separate TLS groups, while individual atomic *B*-factors were refined for water and trimethyllead molecules. After TLS refinement, relatively stringent NCS restraints could be applied, with consequent improvement of *R*-factor, *R*-free, and Ramachandran plot quality (Table 2). Some extra water molecules could also be identified either by automatic water picking in ARP (*27*) or by manual means.

The occupancies of the trimethyllead molecules were initially set to the values refined by SOLVE against the ELETTRA data set and then adjusted according to the difference Fourier maps to values between 0.4 and 0.6. NCS restraints were also used for 47 NCS-related waters identified with the CCP4 program WATNCS (*28*). The final model

Table 2: Refinement Statistics

		molecule		
	all	A	B	C
R-factor (%)	19.4			
R-free (%)	23.5			
no. of non-H protein atoms	4230	1410	1410	1410
no. of trimethyllead atoms	20			
no. of water molecules[a]	47 × 3 + 163[a]	47[a]	47[a]	47[a]
rmsd bonds (Å)	0.016			
rmsd angles (deg)	1.55			
Ramachandran plot[b] (% of residues in core and additionally allowed regions)	90.1/9.9	89.7/10.3	90.3/9.7	90.3/9.7
B-factors[c] for protein atoms (Å^2) (main chain/side chain)	36.5/41.2	40.8/46.1	31.4/35.7	37.1/41.7
B-factors[c] for protein atoms (Å^2) in N-terminal (residues 2−55), loop (residues 111−128), and β-helix (all remaining residues)	49.7/50.4/31.1	60.1/61.8/31.6	40.5/47.6/27.6	48.6/41.7/34.0
B-factors for water molecules (Å^2)[a]	43.4	37.4[a]	37.2[a]	37.5[a]
B-factors for trimethyllead (Å^2)	54.4			

[a] Forty-seven NCS-related water molecules (obeying full NCS) were identified in each monomer. Seventy-two waters obey only one NCS operation while the remaining are not NCS-related using a 1 Å cutoff. The *B*-factors given for the three separate monomers are only for these NCS-related water molecules, while the *B*-factor for all is the average *B*-factor for both NCS-related and non-NCS-related. [b]From PROCHECK (*44*). [c]These *B*-factors are B_{eq} including both TLS contribution and residual individual *B*-factors.

includes all amino acid residues predicted from the DNA sequence (Accession Number P77791) excluding the initiation methionine, five trimethyllead molecules, and 304 water molecules. It has an *R*-factor for the working set of 19.4% and an *R*-free of 23.5%. Other refinement statistics are shown in Table 2.

RESULTS AND DISCUSSION

Cloning and Sequence Analysis of the maa Gene. The *maa* gene was mapped between 490 and 495 kb on the physical map of *E. coli* (*2*). Consequently, we isolated a 4.3 kb *Eco*RI fragment from λ phage 8C4(151) from the Kohara Collection and inserted this fragment in pBluescript II SK (+). The resulting plasmid gave rise to highly elevated levels of maltose acetyltransferase activity in *E. coli*. Further analysis of the plasmid indicated that the *maa* gene was harbored on a 3.2 kb *Eco*RI−*Pst*I fragment, and the nucleotide sequence of this fragment was determined. The sequence revealed an open reading frame potentially encoding a protein of 183 amino acids. This ORF had 39% identical amino acids with GAT and a predicted molecular weight of 20096, which was in good correspondence with the estimated molecular weight of 20000 for MAT. Thus, it was concluded that the 183 ORF is the *maa* gene.

The sequence has up to 79% sequence identity with bacterial and yeast sequences resulting from genome projects, as determined by BLAST searches (*29*). Among sequences for which the gene product has been at least partly characterized, MAT had 64% and 46% sequence identity, respectively, with an acetyltransferase from *Bacillus subtilis* and the chitin oligosaccharide *O*-acetyltransferase NodL from *Rhizobium meliloti*, besides its 39% sequence identity with GAT (*12*, *30*, *31*).

Overexpression and Purification of MAT. Purification to homogeneity was achieved after three chromatographic steps using a combination of anion-exchange and affinity chromatography. Following the procedure described in ref *15*, 30 mg of pure recombinant MAT could be obtained from 1 L of culture, corresponding to a recovery after purification of 43% and a 1280-fold purification. The purity of the enzyme was assessed by SDS−PAGE and mass spectrometry. The latter revealed a molecular weight of 19983, in good agreement with the mass expected from the DNA-deduced amino acid sequence excluding the initiation methionine of 19964.8. The N-terminal sequence (48 residues) of recombinant MAT is identical to the deduced amino acid sequence after initiation methionine and confirmed the homogeneity of the protein.

Physicochemical Properties of MAT. The isoelectric point of MAT was determined by isoelectric focusing. A sharp band corresponding to a p*I* of 5.7 was obtained, confirming the absence of microheterogeneity. MAT was stable for at least 6 h between pH 4.0 and pH 10.0 at 25 °C but showed a significantly reduced lifetime at low pH, with a $t_{1/2}$ of 7 min at pH 3.0. The thermostability of MAT was examined at pH 7.5. MAT was stable at 40 °C for at least 4 h and exhibited $t_{1/2}$ values of >4 h, 70 min, and 20 min at 50, 60, and 70 °C, respectively, according to activity measurements.

pH Optimum. The pH dependence of MAT was investigated using acetyl-CoA and maltose as the acetyl donor and acceptor, respectively. MAT showed a pH optimum of 7.8. The presence of 100 mM NaCl in the solution did not affect the activity of the enzyme, but the nature of the buffer had a tremendous effect on the activity of MAT. At pH 7.5 the highest activity was observed in potassium or sodium phosphate, with decreasing activity in HEPES, MOPS, and Tricine in the given order. No activity could be detected in Trizma or Bis-Tris in the useful pH ranges of these buffers. Previously Boos and co-workers (*2*) showed that triethanolamine and Tris could act as acceptors for MAT but with very low V_{max}; i.e., they can act as competitive inhibitors. At the high concentrations of Tris and the related compound Bis-Tris used in the buffer system in our assay condition, negligible MAT activity remains.

Regiospecificity of the MAT-Catalyzed Acetylation. To investigate the regiospecificity of MAT with respect to which −OH groups of the acceptor substrate are acetylated by the enzyme, milligram amounts of MAT-catalyzed acetylated glucose and maltose were prepared. These compounds were consequently purified, and their structure was determined by ^{1}H NMR (Figure 1). Glucose was found to be acetylated at the C6 position only, and maltose was acetylated at the

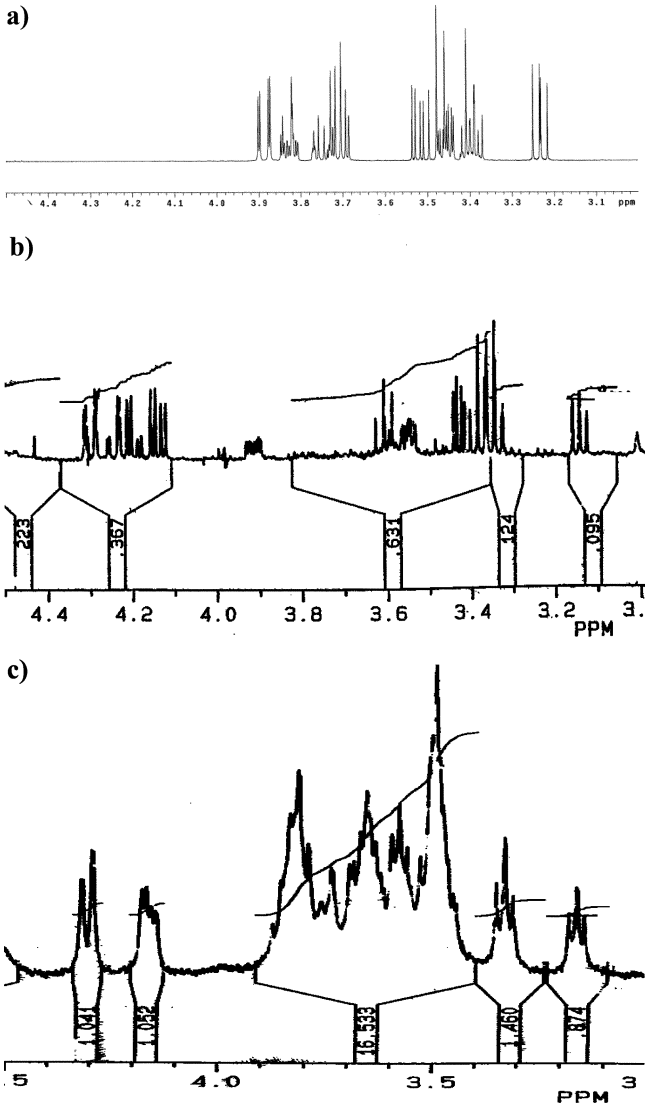

FIGURE 1: NMR spectra of glucose and products of MAT-catalyzed acetyl transfer from acetyl-CoA to glucose and maltose. Panels: (a) glucose; (b) product from glucose; (c) product from maltose. By the relative downfield shift of the hydrogens on the acetylated carbon, the position of the acetyl group was assigned by ^{1}H NMR to be the C6 position for glucose and maltose (shift from 3.6−3.9 ppm to 4.1−4.4 ppm). For glucose there were four H-6 protons, due to the α/*β* mixture. For maltose there were only two acetylated protons, despite the α/*β* mixture, suggesting that acetylation occurred on the nonreducing glucose moiety.

same position on its nonreducing glucose moiety only. The regiospecificity of another *O*-acetyltransferase of this family, NodL, has previously been investigated using acetyl-CoA and chitin pentaose as donor and acceptor substrates, respectively (*32*). NodL was found to catalyze the transfer of an acetyl group to the C6 position of the nonreducing terminal hexose residue only. MAT and NodL thus exhibit the same regiospecificity and are both exo-enzymes, acting solely on the nonreducing hexose residue of their substrate. To our knowledge regiospecificity has been experimentally addressed for MAT and NodL only, and the results confirm the widely accepted postulate that all hexose *O*-acetyltransferases of this family exhibit regiospecificity for the C6 position of their acceptor substrate.

Acceptor Specificity. The acceptor specificity of MAT has previously been investigated using a limited number of substrates (*2*). Here, initial reaction rates of MAT were measured at three arbitrary concentrations of a wider panel of potential acceptor substrates (Table 3). As previously reported, MAT exhibits a marked preference for glucose among the different hexoses tested (furanose/pyranose). Fucose, which lacks a free hydroxyl group at C6, was as expected not a substrate. *myo*-Inositol, a structural analogue of glucose, was not a substrate either. Galactose, which differs from glucose and mannose in the orientation of its C4 hydroxyl group, was the poorest substrate among these hexoses. Among the glucose derivatives tested, IPTgalactose, the best substrate for GAT, was a poor acceptor for MAT, whereas IPTglucose was among the best substrates for MAT. This illustrates well the difference in acceptor substrate preference of MAT and GAT. MAT showed a marked preference for derivatives of the *β*-anomeric form of glucose compared to those of the α anomer, as seen for methyl, *p*-nitrophenyl, and *n*-octyl glucopyranoside. *β*-Alkyl derivatives of glucose were among the best acceptor substrates, especially the ones with long or bulky alkyl chains. Glucose 1-phosphate, UDP-glucose, and ADP-glucose, which are biosynthetic precursors of polysaccharides such as starch or cellulose, were very poor substrates or nonsubstrates for MAT. Derivatives of glucose at C atoms other than C1 were also tested as acceptor substrates: glucosamine, bearing an amino group at C2, was a good substrate whereas a C3 derivative (3-*O*-methyl-D-glucose) was a very poor substrate. C6 glucose derivatives were not substrates, as expected.

As far as disaccharides are concerned, MAT exhibited an exclusive specificity for α-glycosidic linkages. However, the position of the linkage was of a lesser importance, since glucose disaccharides linked in α-1,4, α-1,6, and α-1,3, but not α-1,1, could be acetylated by MAT. Derivation on the C1 atom of maltose had little influence on the reaction rates as compared to derivation of glucose.

Since glucose and maltose were among the best substrates, we also investigated the effect of the length of oligomaltosides on the initial reaction rate. The length of the oligomaltoside was found to be inversely proportional with the reaction rate; in particular substrates longer than maltotriose were acetylated at a very slow, although not negligible rate. This is in contrast with the closely related NodL, a chitin oligosaccharide acetyltransferase for which substrate preference increases with the length of the acceptor substrate, at least up to a pentameric substrate (*33*).

Kinetic parameters were determined for IPTglucose, glucose, maltose, and maltotriose as the acceptor substrate and acetyl-CoA as the donor substrate (Table 4). As previously reported (*2*), K_m values for acetyl-CoA were at least 1000-fold lower than for the acceptor substrates. The nature of the acceptor had a minor effect on the K_m value of acetyl-CoA, possibly reflecting a variable overlap in the binding modes of the two substrates or different values for the dissociation constants of the different acetylated reaction products. As a general trend, both K_m and V_{max} contribute in combination to the substrate specificity ranking, the best substrate exhibiting both the best $K_{m(acceptor)}$ and the best V_{max} and the poorest substrate exhibiting both the poorest $K_{m(acceptor)}$ and the poorest V_{max}. However, there is one exception to this trend, since the $K_{m(acceptor)}$ of maltose is better than that of glucose, whereas V_{max} and $V_{max}/K_{m(acceptor)}$ are poorer for maltose than for glucose. The examination of the ratio $V_{max}/K_{m(acceptor)}$ reveals that IPTglucose is the best substrate,

Table 3: Effect of the Nature of the Acceptor Substrate on the Initial Rate of MAT-Catalyzed Acetyl Transfer Reactions[a]

acceptor substrate	initial rate[b] 10 mM acceptor	50 mM acceptor	100 mM acceptor
hexoses			
glucose (S, B)[c]	**0.876**	**2.952**	**4.304**
mannose (M)	*d*	0.839	1.243
fructose (M)		0.279	0.552
galactose (M)		0.030	0.052
fucose (S)		0	0
myo-inositol (S)		0	0
α-glucose derivatives			
α-D-glucopyranose 1-phosphate (S)		0	0.001
UDP-glucose (S)		0.002	0.003
methyl α-D-glucopyranoside (S)		0.720	1.320
p-nitrophenyl α-D-glucopyranoside (S)	0.690	2.007	
n-octyl α-D-glucopyranoside (S)	1.159		
β-glucose and *β*-thioglucose derivatives			
ADP-glucose (S)		0	
methyl *β*-D-glucopyranoside (S)		1.060	1.980
n-hexyl *β*-D-glucopyranoside (C)	6.240	15.273	17.124
n-heptyl *β*-D-glucopyranoside (C)	7.711	15.306	14.943
n-octyl *β*-D-glucopyranoside (S)	11.744	16.075	14.749
n-nonyl *β*-D-glucopyranoside (C)	13.723	13.683	13.833
n-decyl *β*-D-glucopyranoside (C)	2.073	0.227	
p-nitrophenyl *β*-D-glucopyranoside (S)	13.054	13.863	10.378
isopropyl *β*-D-thioglucopyranoside (S)	**7.940**	**14.656**	**17.496**
n-heptyl *β*-D-thioglucopyranoside (C)	10.226	16.274	17.426
n-octyl *β*-D-thioglucopyranoside (C)	15.534	15.199	16.294
isopropyl *β*-D-thiogalactoside (S)		0.157	0.349
other derivatives			
glucosamine (S)		1.326	1.306
3-*O*-methyl-D-glucopyranoside (S)		0.028	0.075
D-glucopyranose 6-phosphate (S)		0.001	0.001
α-D-glucose 1,6-diphosphate (S)		0	0
disaccharides and maltose derivatives			
maltose (α-1,4 linkage) (S)	**0.483**	**1.986**	**2.944**
isomaltose (α-1,6 linkage) (S)		1.885	2.297
nigerose (α-1,3 linkage) (S)		0.486	1.360
trehalose (α-1,1 linkage) (S)		0	0.001
cellobiose (*β*-1,4 linkage) (S)			0
lactose (4-*O*-*β*-galactopyranosyl-D-glucose) (S)			0.009
sucrose (α-D-glucopyranosyl *β*-D-fructofuranoside) (S)		0.006	0.008
p-nitrophenyl α-D-maltoside (S)	0.852	2.196	2.548
p-nitrophenyl *β*-D-maltoside (S)	3.490	4.071	2.962
n-decyl *β*-D-maltoside (C)	4.780	6.960	6.418
n-dodecyl *β*-D-maltoside (C)	1.671	1.614	1.431
oligomaltosides			
maltotriose (S)			**0.156**
maltotetraose (S)			0.010
maltopentaose (S)			0.018
maltohexaose (S)			0.024
maltoheptaose (S)			0.023
maltodextrin[e]			0.011

[a] Reactions were performed with 40 μM acetyl-CoA as the acetyl donor and were started by the addition of the acceptor substrate. Kinetic parameters for acceptors indicated in bold type were further investigated. [b] One unit corresponds to an increase in A_{412} of 1 per minute at a MAT monomer concentration of 5 μg/mL. [c] The letter in parentheses indicates the supplier: S, Sigma; B, BioSource International; M, Merck; C, Calbiochem. [d] A blank indicates not determined. [e] This substrate is a mixture of maltodextrins ranging in size from 10 to 30 glucose residues. The concentration used in the assay was 20 mg/mL.

followed by glucose, maltose, and maltotriose, thus confirming that MAT has higher affinity for the alkyl-substituted derivative than glucose itself. Increasing the length of the acceptor substrate results in a dramatic decrease in the catalytic efficiency, as suggested by the initial reaction rate on oligomaltosides of increasing length (Table 3). The primary plots of $1/v$ against 1/(substrate) show a series of straight lines that intersect at a unique point at the left of the $1/v$ axis. This is consistent with a sequential bi-bi kinetic mechanism. Such a mechanism was also demonstrated for GAT (*34*), chloramphenicol acetyltransferase (*35*), and NodL (*36*), indicating that this family of *O*-acetyltransferases may share a similar catalytic machinery and mechanism. However, it is not possible to differentiate between a random or ordered binding of substrates from the measurement of initial rates in the absence of products.

Structure of MAT. The structure of MAT has been refined to 2.15 Å resolution with data collected for a trimethyllead cocrystal (crystal form IV). The final *R*-factor was 19.4%, and the *R*-free was 23.5%. Other refinement statistics are shown in Table 2. The structure has good geometrical and packing quality as judged by WHAT_CHECK (*37*).

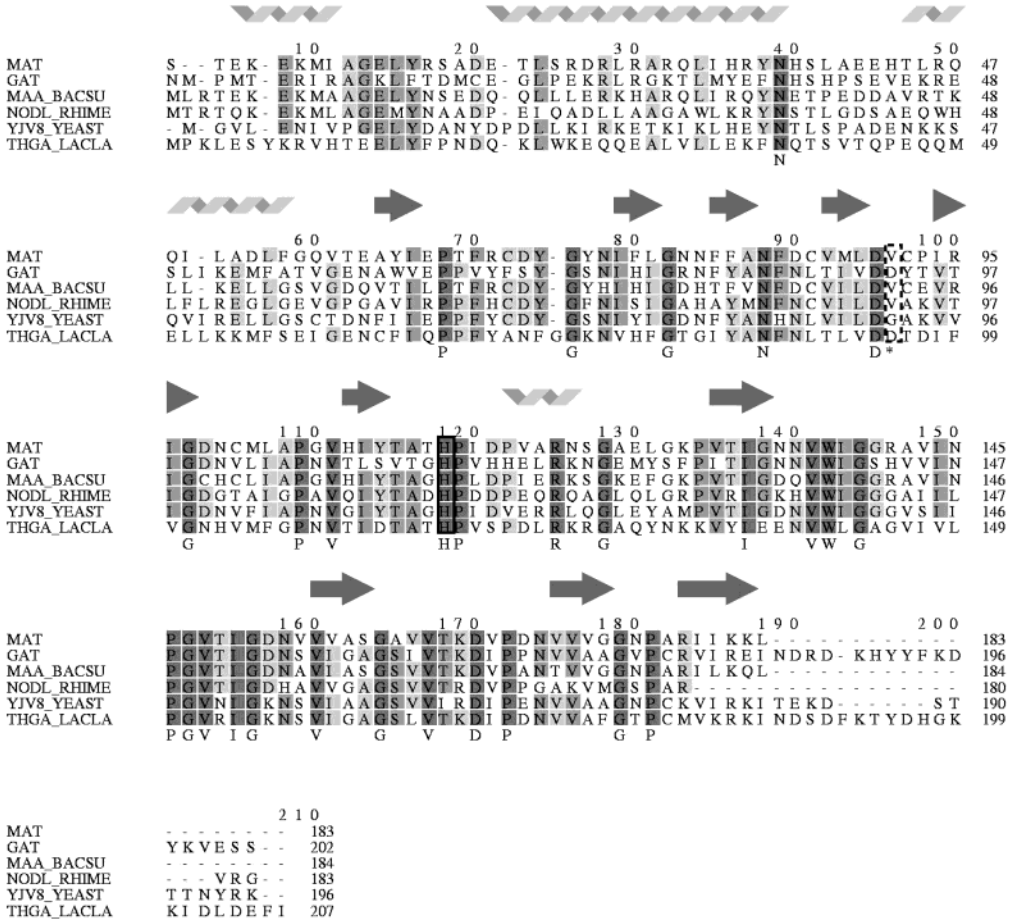

FIGURE 2: Structure-based sequence alignment of MAT (PDB code 1OCX) and GAT (PDB code 1KRU). The alignment was performed with INDONESIA (*45*). Secondary structure elements as identified by PROMOTIF (*38*) are shown. After the structure-based sequence alignment, four highly related sequences in *lacA* family were aligned to it while keeping the structurally conserved regions in GAT and MAT locked with respect to each other. The sequences are MAA_BACSU (Accession Number P37515), *B. subtilis O*-acetyltransferase (*30*); NODL_RHIME (Accession Number P28266), *R. meliloti* NodL (*31*); YJV8_YEAST (Accession Number P40892), putative acetyltransferase from *Saccharomyces cerevisiae* (*46*); THGA_LACLA (Accession Number P52984), probable galactoside acetyltransferase from *Lactococcus lactis* (*47*). The sequence alignment is shaded according to conservation (the more conserved residues are in darker gray). A consensus is shown under the alignment. The putative active site histidine is boxed with a solid line. Residues corresponding to Asp93 in GAT/Val 91 in MAT are indicated by an asterisk under the sequence and are boxed with a dashed line.

Table 4: Kinetic Parameters of MAT[a]

acceptor substrate	$K_{m(acceptor)}$ (mM)	$K_{m(acetyl-CoA)}$ (μM)	V_{max}[b]	$V_{max}/K_{m(acceptor)}$
IPTglucose	17.2	10.7	38.5	2.260
glucose	71.0	6.1	9.2	0.130
maltose	42.9	1.8	4.2	0.097
maltotriose	285.8	3.2	1.7	0.006

[a] Four different acetyl-CoA concentrations were tested, and seven to nine different concentrations of the acceptor were tested for each of the acetyl-CoA concentrations and vice versa. The kinetic parameters were then derived from the corresponding secondary plots and are average values from duplicates or triplicates. [b] One unit is defined as in Table 3.

Secondary structure elements as identified by PROMOTIF (*38*) are shown above the sequence in Figure 2. As expected from the presence of the hexapeptide repeat motif, MAT forms a trimer mediated by the LβH domain (Figure 3a). The LβH domain contains 16 full hexapeptide repeats, which form $5^1/_3$ β-helix coils (Figure 4). The MAT monomer has only one long loop (residues 111−128) protruding from its LβH domain. Close to the β-helix the loop forms a β-hairpin involving residues 111−112 and 126−128, and it has a six-residue helical stretch in the middle. The N-terminus of the LβH is capped by a domain consisting of three helices (residues 2−55). The C-terminus is capped by a β-hairpin (residues 169−183).

Thr65 and Pro176 are involved in *cis*-peptide bonds. Asn79 and Cys86 are found to adopt two alternate conformations in all three monomers. One trimethyllead per monomer is bound to His113 in a hydrophobic area at the interface between two monomers (Figure 3d), situated in the long loop protruding from the LβH domain. These three trimethyllead molecules are related by the same 3-fold NCS as the protein monomers and are modeled in the pyramidal geometry found in small molecules for tetracoordinated trimethyllead. Two other trimethyllead molecules are bound by Glu6 of one monomer (A or B) and Glu12 from another monomer (B or C) belonging to a trimer related by crystallographic symmetry. These trimethyllead molecules are related by the same NCS operation as monomers A and B; however, Glu6 of monomer C is far from Glu12 of crystallographically related molecules and therefore cannot bind trimethyllead, so that the NCS is not perfectly observed. These two trimethyllead molecules are consistent with the flat geometry observed in small molecules for pentacoordi-

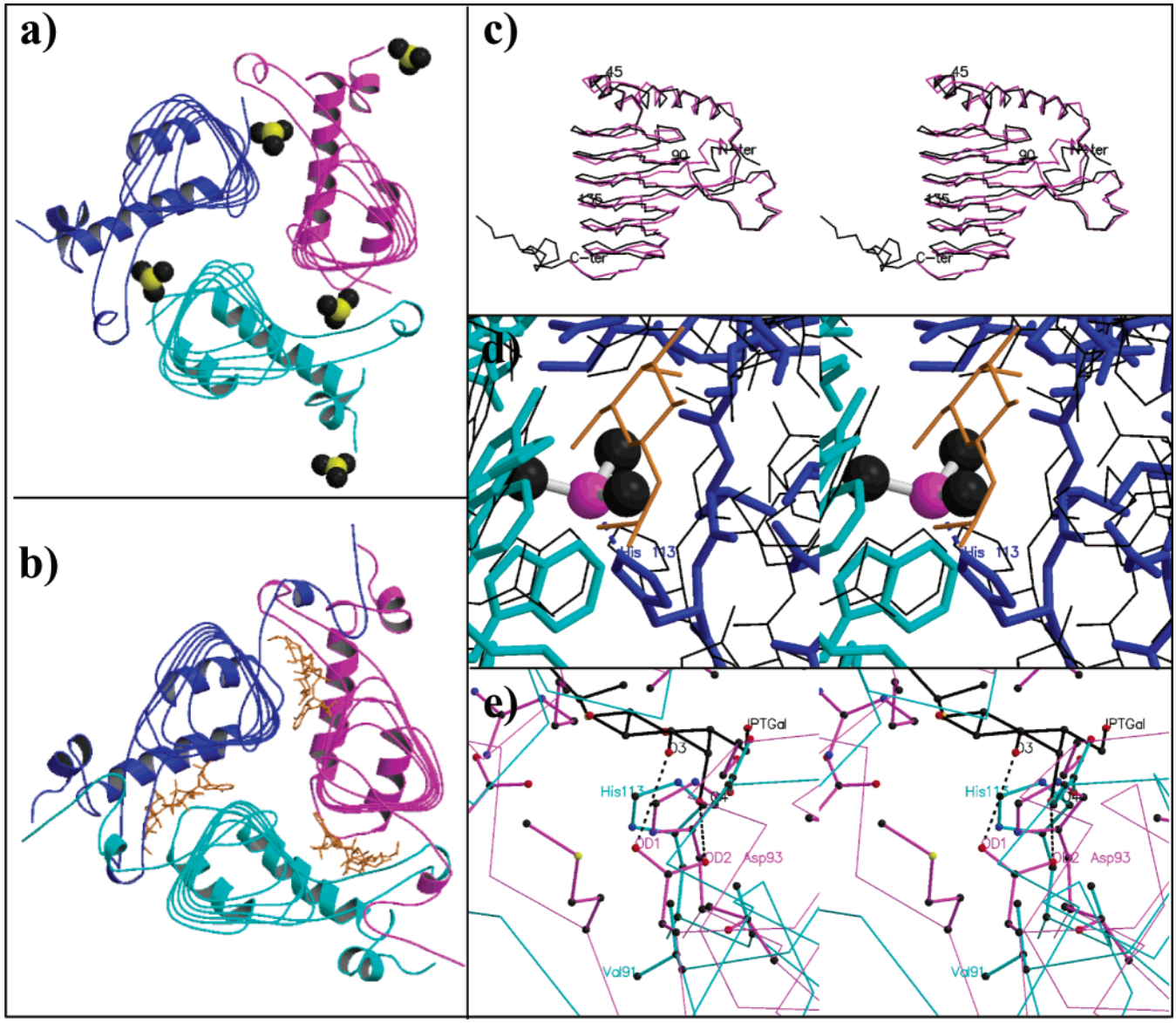

FIGURE 3: View of the structures of MAT and GAT after superposition. The ternary complex of GAT with CoA and IPTgalactose is used (PDB code 1KRU). (a) and (b) show top views of the MAT and GAT trimers, respectively, with bound trimethyllead (spheres) and CoA/IPTgalactose (orange sticks), respectively. (c) Stereoview of superposed MAT and GAT monomers as Cα traces. The N- and C-terminus as well as occasional residue numbers are marked for MAT. (d) Stereoview of the putative active site histidine (His113) binding to trimethyllead (in ball-and-stick representation) in MAT. Two MAT monomers are shown in blue and cyan. GAT is shown with a thin black line for comparison with the vicinal IPTgalactose bound in orange. (e) Stereoview of the interaction of Asp93 in GAT (magenta) with O3 and O4 of IPTgalactose (black). In MAT (cyan) these interactions are lost, as Asp93 is substituted by Val91. Only residues within 4 Å of IPTgalactose are shown with side chains. This figure was produced using MOLSCRIPT (*48*) and Raster3D (*49*).

nated trimethyllead, although the coordination geometry is not clearly defined in the electron density.

The rmsd for Cα atoms was 0.15 Å between molecules A and B, 0.19 Å between molecules A and C, and 0.22 Å between molecules B and C. The rmsd for all protein atoms was 0.35 Å between molecules A and B, 0.45 Å between molecules A and C, and 0.47 Å between molecules B and C. Differences between the MAT monomers are more pronounced in the N-terminus and in the loop region, which might reflect the fact that these are also the most disordered regions (see *B*-factors in Table 2).

Comparison of MAT with Other LβH Enzymes. One of the differences between the parallel β-helical domains found in different enzymes is in the number of coils, as the structure of the coils themselves is highly conserved even in the absence of high sequence identity. With its $5^1/_3$ coils (Figure 4), MAT is one of the shorter LβH, the shortest being PaXAT (*7*) with $4^1/_3$ coils. In MAT, the axis of each LβH forms an angle of around 10° with the axis of the trimer, consistent with the proposal that shorter β-helix monomers in the trimer tend to be less parallel to each other, because of reduced interactions over the helix length (*7*). However, it must be noted that the truncated version of GlmU, only containing five coils of the β-helix, has nonetheless a parallel arrangement of monomers (*8*). In MAT only one loop protrudes from the LβH domain, as compared, for example, to two in UDP-*N*-acetylglucosamine acyltransferase (LpxA) (*4*) and three in carbonic anhydrase (*5*).

Other major differences between hexapeptide repeat proteins occur at the N- and C-terminus outside the LβH region. For example, while MAT has an N-terminal α-helical domain and a β-hairpin at its C-terminus, LpxA (*4*) and

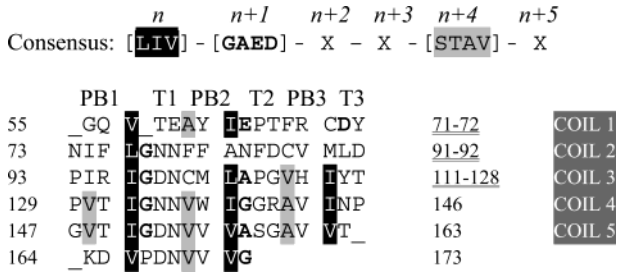

FIGURE 4: The hexapeptide repeat in MAT according to the determined three-dimensional structure. Residues which conform to the consensus hexapeptide at the first, second, and fifth positions are shown in reverse shading, bold, and gray shading, respectively. Deletions in the consensus are indicated by an underscore. The structural role of residues in the β-helix is indicated. PB1, PB2, and PB3 refer to strands forming the three parallel β-sheets, while T1, T2, and T3 refer to the turns between them. Residue numbers of inserted loops after T3 are double underlined.

PaXAT (*7*) have helical C-terminal domains. TDT (*6*) has a mostly helical N-terminal domain, which is completely different from the one found in MAT.

In MAT, the monomers are close together at the C-terminus, and they are more distant toward the N-terminus of the β-helix. Important interactions between the monomers are provided by the loop protruding from one β-helix and embracing the neighboring monomer. The N-terminal domain of each monomer interacts mostly with the long loop from the same monomer.

Among all the parallel β-helix structures available to date, GAT (*12*) is the most similar to MAT, with a Cα rmsd of 1.0 Å over 169 residues aligned between two selected monomers. The angle between monomers and the trimer axis is reported to be 12°, very similar to MAT. MAT and GAT share 41% sequence identity after structure-based sequence alignment (Figure 2). Figure 3c shows an overlay of a MAT and a GAT monomer. Apart from a C-terminal extension in GAT, the backbones of the two structures are very similar, also in the long loop region and the N-terminal and C-terminal caps. The similarity in the loop region is particularly significant, as the loops protruding from the LβH domains are usually involved in active site formation in LβH enzymes and vary considerably in size and number between different members of the family.

The structural and sequence similarities are, as we have seen, correlated with functional similarity. MAT acetylates, like GAT, the O6 of hexose sugars, and both enzymes prefer substrates with hydrophobic moieties at the reducing end. The main difference between GAT and MAT is in the specificity. MAT prefers glucose to galactose (*2*; this study), while GAT is reported to have the opposite preference (*17*, *34*, *39*). The crystallization of NodL has been reported (*40*), although the structure is not yet available. MAT and NodL share 48% sequence identity and display similar specificities. NodL also acetylates the O6 of glucose-derived hexose sugars but prefers oligosaccharides whereas MAT prefers mono- to trisaccharides.

The Putative Active Site. Catalytically important histidine residues have been identified at the active sites of several structures of LβH enzymes (*5*, *9*, *11*, *13*). In GAT, the involvement of His115 in catalysis has been supported both by a crystallographic study of the enzyme in the ternary complex with IPTgalactose and CoA (*12*) and by mutagenesis and kinetic studies showing a 1800-fold decrease in k_{cat} for the H115A enzyme variant (*39*). His115 has been proposed to abstract the proton from the C6 hydroxyl of the acceptor prior to acetyl transfer.

Given the high structural and sequence conservation of MAT and GAT and similarity in their steady-state kinetics, it is likely that they operate by the same catalytic mechanism. His113 in MAT is in an equivalent position and could fulfill an analogous function to His115 in GAT.

Panels a and b of Figure 3 show top views of the MAT and GAT trimers, respectively. The ternary complex of GAT with CoA and IPTgalactose is shown. The active site is formed in a tunnel open at both ends, situated at each of the three dimer interfaces. In the MAT structure, trimethyllead is bound in the tunnel to the putative active site His113 (Figure 3d). It should be noted that His113 is situated on the long loop protruding from the β-helix, around the middle of the tunnel.

In the GAT structure, many residues belonging to the parallel β-helix are involved in interactions with the substrates. In this sense, the LβH seems to be divided in two parts, with the N-terminal involved in acceptor binding and the C-terminal involved in acetyl-CoA binding. Lys166, Arg180, and Arg183 interacting with the phosphate/pyrophosphate groups of CoA in GAT are conserved in MAT though the latter Arg is replaced by a Lys. Pro178 interacting with the adenine moiety of CoA in GAT (Pro176 in MAT) is also conserved and so are Trp139, Val157, and Val173 (Trp137, Val155, and Val171 in MAT), which interact in GAT with the pantotheine moiety of CoA. These residues are also highly conserved in the related sequences shown in Figure 2. Sequence identity is in general very high at the C-terminal end of the β-helix binding to the common acetyl-CoA donor but less pronounced at the N-terminal end of the helix, reflecting the differences in acceptor specificities.

The Putative Acceptor Site. Figure 5 shows a surface view of MAT with superimposed IPTgalactose molecules as bound to GAT. An extended hydrophobic surface is found near the isopropyl moiety of IPTgalactose, explaining the preference of MAT for acceptors with long alkyl chains. The hydrophobic nature of this patch gives also a rationale for the selection against charged substrates such as glucose phosphates.

However, it is more difficult to explain why MAT prefers β-linked to α-linked hydrophobic substituents at the anomeric carbon but α-linked to β-linked oligosaccharides. An explanation could be that the hydrophobic patch cannot accommodate the hydroxyl groups of the carbohydrate moieties, so that α-linked oligosaccharides bind in a slightly different orientation, allowing binding of a sugar moiety to the distal IPTgalactose binding site identified in GAT (*12*).

The tunnel shape of the active site, together with the presence of a hydrophobic patch inaccessible to sugar moieties, explains the exo specificity of MAT toward oligosaccharides, as there is only room for a terminal glucose unit to come in close proximity of the donor for acetyl transfer.

One important difference in the specificity of MAT and GAT is the preference of MAT for glucosyl rather than galactosyl moieties. This difference in specificity is what makes MAT particularly biotechnologically attractive for the modification of starch and maltooligosaccharides.

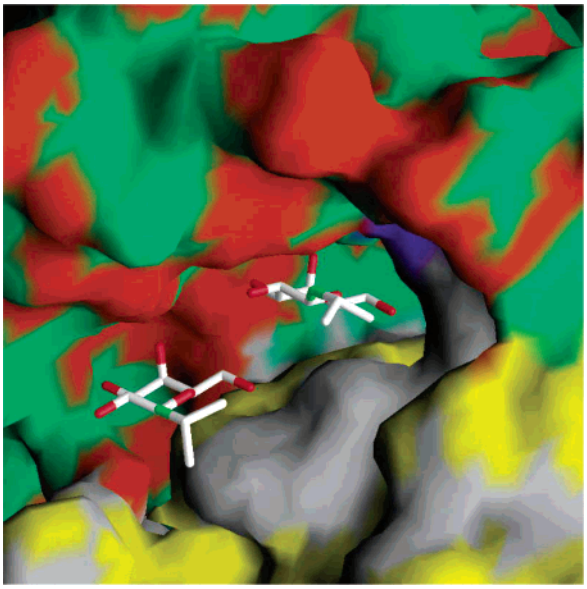

FIGURE 5: MAT surface representation produced using GRASP (*50*). Orange and green represent O and N (polar) and C and S (hydrophobic) atoms in one monomer; yellow and gray represent the same, respectively, in the other monomer shown. The putative active site His113 is in purple. IPTgalactose molecules from the GAT trinary complex are shown.

In the complexes of GAT with IPTgalactose and pNPβ-galactose at the acceptor site, Asp93 makes an interaction with both C3 and C4 hydroxyls (Figure 3e). This is the only interaction made by GAT with the axial C4 hydroxyl. In MAT the residue corresponding to Asp93 is Val91, which suggests that this substitution is at least partly responsible for the different hexose specificity of GAT and MAT. An alignment with closely related sequences to MAT and GAT shows that this residue is variable. NodL in *Rhizobia* has been shown to transfer an acetyl group from acetyl-CoA to the O6 of *N*-acetylglucosamine and the nonreducing moiety of chitin and lipochitin oligosaccharides (*32*). NodL has like MAT a valine at the corresponding position of GAT Asp93. This is consistent with its substrate specificity, as *N*-acetylglucosamine possesses O3 and O4 in the same orientations as glucose.

Conclusions. The biochemical and structural characterization of MAT presented here highlights the similarities and differences between MAT and GAT, two members of the *lacA* family of acyltransferases. Despite extensive biochemical characterizations, the cellular functions for MAT and GAT are not clear, although roles in detoxification have been proposed for both (*1*, *41*). The *lacA* gene encoding for GAT is part of the classical *E. coli* lactose operon, while the *maa* genes encoding for MAT in *E. coli* and its *B. subtilis* homologue have rather uncharacterized gene neighbors and do not seem to be part of an operon according to our analysis with Colibri and Subtilist (*42*, *43*).

Both enzymes have a preference for acetyl-CoA as donor, both acetylate hexoses at the C6 hydroxyl, and both have a marked preference for sugar moieties containing hydrophobic substituents. The preference for hydrophobic substituents can be rationalized by the presence of an extended hydrophobic surface near the acceptor site.

However, GAT and MAT differ in the specificity at the C4 hydroxyl of the hexose, GAT having a strong preference for galactose as opposed to glucose and vice versa for MAT. This difference makes MAT more attractive for modification of starch-derived maltooligosaccharides.

Comparison of the two structures makes us suggest that the substitution of Asp93 in GAT to Val91 in MAT might be at least partly responsible for the specificity at the C4 hydroxyl of the hexose acceptor. This might help to predict the specificity of new highly related sequences resulting from genome projects.

ACKNOWLEDGMENT

We thank Clive Phipps Walter (Danisco Innovation Copenhagen) for protein N-terminal sequencing and gratefully acknowledge the excellent technical assistance provided by Vibeke Larsen, Britta Hamborg Nielsen, and Maj-Britt Eicke. We thank the beamline staff at the ESRF and ELETTRA synchrotrons for assistance and the EU for travel support through the program Access to Research Infrastructures. We also thank Pernille Harris, Jens Bukrinsky, and Flemming Hansen for help with the data collections.

REFERENCES

1. Boos, W., Ferenci, T., and Shuman, H. A. (1981) *J. Bacteriol. 146*, 725−732.
2. Brand, B., and Boos, W. (1991) *J. Biol. Chem. 266*, 14113−14118.
3. Vuorio, V., Härkönen, Tolvanen, M., and Vaara, M. (1994) *FEBS Lett. 337*, 289−292.
4. Raetz, C. R. H., and Roderick, S. L. (1995) *Science 270*, 997−1000.
5. Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G., and Rees, D. C. (1996) *EMBO J. 15*, 2323−2330.
6. Beaman, T. W., Binder, D. A., Blanchard, J. S., and Roderick, S. L. (1997) *Biochemistry 36*, 489−494.
7. Beaman, T. W., Sugantino, M., and Roderick, S. L. (1998) *Biochemistry 37*, 6689−6696.
8. Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreulx, D., Camnbillau, C., and Bourne, Y. (1999) *EMBO J. 18*, 4096−4107.
9. Olsen, L. R., and Roderick, S. L. (2001) *Biochemistry 40*, 1913−1921.
10. Kostrewa, D., D'Arcy, A., Takacs, B., and Kamber, M. (2001) *J. Mol. Biol. 305*, 279−289.
11. Sulzenbacher, G., Gal, L., Peneff, C., Fassy, F., and Bourne, Y. (2001) *J. Biol. Chem. 276*, 11844−11851.
12. Wang, X.-G., Olsen, L. R., and Roderick, S. L. (2002) *Structure 10*, 581−588.
13. Sugantino, M., and Roderick, S. L. (2002) *Biochemistry 41*, 2209−2216.
14. Jenkins, J., and Pickersgill, R. (2001) *Prog. Biophys. Mol. Biol. 77*, 111−175.
15. Lo Leggio, L., Dal Degan, F., Sørensen, S. O., Harlow, K., Harris, P., and Larsen, S. (2001) *Acta Crystallogr. D57*, 1915−1918.
16. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
17. Alpers, D. H., Appel, S. H., and Tomkrins, G. M. (1965) *J. Biol. Chem. 240*, 10−13.
18. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
19. Terwillinger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. D55*, 849−861.
20. Vonrhein, C. (2000) *CCP4 Program Manuals*, version 4.1.
21. Cowtan, K. (1994) *Jt. CCP4 ESF-EACBM Newsl. Protein Crystallogr. 31*, 34−38.
22. Vagin, A., and Teplyakov, A. (1997) *J. Appl. Crystallogr. 30*, 1022−1025.
23. Terwillinger, T. C. (2001) *Acta Crystallogr. D57*, 1755−1762.
24. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
25. 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. D54*, 905−921.

26. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D53*, 240−255.
27. Lamzin, V., and Wilson, K. (1993) *Acta Crystallogr. D49*, 129−147.
28. Lu, G. (2001) *CCP4 Program Manuals*, version 4.1.
29. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res. 25*, 3389−3402.
30. Quirk, P. G., Guffanti, A. A., Clejan, S., Cheng, J., and Krulwich, T. A. (1994) *Biochim. Biophys. Acta 1186*, 27−34.
31. Baev, N., and Kondorosi, A. (1992) *Plant Mol. Biol. 18*, 843−846.
32. Bloemberg, G. V., Thomas-Oates, J. E., Lugtenberg, B. J. J., and Spaink, H. P. (1994) *Mol. Microbiol. 11*, 793−804.
33. Bloemberg, G. V., Lagas, R. M., van Leeuwen, S., Van der Marel, G. A., Van Boom, J. H., Lugtenberg, B. J. J., and Spaink, H. P. (1995) *Biochemistry 34*, 12712−12720.
34. Musso, R. E., and Zabin, I. (1973) *Biochemistry 12*, 553−557.
35. Kleanthous, C., and Shaw, W. V. (1984) *Biochem. J. 223*, 211−220.
36. Hindson, V. J., Duun, S. O., Rowe, A. J., and Shaw, W. V. (2000) *Biochim. Biophys. Acta 1479*, 203−213.
37. Hooft, R. W. W., Vriend, G., Sander, C. Abola, E. E. (1996) *Nature 381*, 272−272.
38. Hutchinson, E. G., and Thornton, J. M. (1996) *Protein Sci. 5*, 212−220.
39. Lewendon, A., Ellis, J., and Shaw, W. V. (1995) *J. Biol. Chem. 270*, 26327−26331.
40. Dunn, S. M., Moody, P. C. E., Downie, J. A., and Shaw, W. V. (1996) *Protein Sci. 5*, 538−541.
41. Zabin, I., Kepes, A., and Monod, J. (1962) *J. Biol. Chem. 237*, 253−257.
42. http://genolist.Pasteur.fr/Colibri/.
43. http://genolist.Pasteur.fr/Subtilist/.
44. Laskowski, R. A., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
45. Madsen, D. http://xray.bmc.uu.se/dennis/manual/.
46. Vandenbol, M., Durand, P., Bolle, P. A., Dion, C., Portetelle, D., and Hilger, F. (1994) *Yeast 10*, 1657−1662.
47. Griffin, H. G., and Gasson, M. J. (1994) *Biotechnol. Lett. 16*, 1125−1130.
48. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
49. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
50. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins 11*, 281−296.

BI0271446