The Structure and Mechanism of Serine Acetyltransferase from *Escherichia coli*

Serine acetyltransferase (SAT) catalyzes the first step of cysteine synthesis in microorganisms and higher plants. Here we present the 2.2 Å crystal structure of SAT from *Escherichia coli*, which is a dimer of trimers, in complex with cysteine. The SAT monomer consists of an amino-terminal α-helical domain and a carboxyl-terminal left-handed β-helix. We identify His109 and Asp143 as essential residues that form a catalytic triad with the substrate for acetyl transfer. This structure shows the mechanism by which cysteine inhibits SAT activity and thus controls its own synthesis. Cysteine is found to bind at the cysteine substrate site and not at the acetyl-CoA site that had been reported previously. On the basis of the geometry around the cysteine binding site, we are able to suggest a mechanism for the O-acetylation of serine by SAT. We also compare the structure of SAT with other left-handed β-helical structures.

Serine acetyltransferase (SAT) participates in the dual-step process of sulfur assimilation of microorganisms (1–3) and higher plants (4, 5). First, SAT catalyzes the production of O-acetyl-L-serine from acetyl-CoA and L-serine, and then O-acetyl serine (thiol) lyase (OAS-TL) converts O-acetyl-L-serine into L-cysteine in the presence of sulfide. Kredich and co-workers (2, 6) reported that SAT (from *Salmonella typhimurium*) represents the rate-limiting component and is reversibly associated with ~5% of the total cellular OAS-TL to form the multi-enzyme complex referred to as “cysteine synthase.” Cysteine constitutes the almost exclusive metabolic entrance for the biosynthesis of essential compounds including methionine, several vitamins, and metal clusters (7). The production of cysteine is therefore of biotechnological interest for pharmacological processes and as a nutritional supplement for food and feed.

SAT is known to be a member of the bacterial O-acetyltransferases subfamily of O-acetyltransferases (8, 9), where amino acid sequence, tertiary structures, and mechanisms are known. The folding pattern that dominates the O-acetyltransferase family is the left-handed β-helix, which is recognized by a hexapeptide repeating signature in which residue i is aliphatic, i + 1 is usually glycine, and i + 4 is a small residue, thus [LIV]-[GAED]-X2-[STAV]-X. Structures are triangular in cross-section and are formed by parallel β-strands folding into a helix with three strands per turn. The first of such proteins to be studied by X-ray crystallography was the *lpxA* gene product that was shown to be a trimeric protein with this left-handed β-helical fold (10). A folding pattern arises from a hexapeptide repeat, which occurs, albeit to varying degrees, in other members of the O-acetyltransferase family. SAT has such hexapeptide repeats at its carboxyl-terminal region. The three clefts between the three subunits form the catalytic centers in which a histidine residue is essential for transfer of the acetyl or succinyl moiety from CoA to the second substrate. From the sequence alignment in Fig. 1, it can be seen that the amino-terminal region of SAT is more varied in length and sequence than the carboxyl-terminal region. The carboxyl-terminal region of SATs is highly conserved and is of special interest because it is responsible for the hetero-oligomerization with OAS-TL (11). Left-handed β-helices pack together in a trimeric structure with the 3-fold axis parallel to the helical axis; the main interactions between monomers are hydrophobic ones.

Preliminary crystallographic analysis (12) revealed that SAT is likely to associate in a hexameric form with 3-2 symmetry; this was supported by chemical cross-linking and gel filtration experiments. Because only preliminary crystallographic analysis (12) and the sequential mechanism as proposed by Leu and Cook (13) were available, Wirtz et al. (14) proposed a structural model for SAT and aimed to validate it by site-directed mutagenesis. Virtually nothing was known about the tertiary structure of the SAT hexapeptide-repeat domain or its oligomerization with OAS-TL. Wirtz et al. (14) proposed that the carboxyl-terminal domain was likely to comprise a β-helix and was involved not only in SAT-OAS-TL interactions but also in homotrimeric interactions of SAT. The amino terminus, proposed to be an α-helical domain (11), was suggested to be involved in SAT-SAT interactions, thus forming a homodimer of homotrimers. Here we present the three-dimensional structure of SAT from *Escherichia coli*, solved by X-ray crystallography at 2.2 Å resolution, and confirm the quaternary arrange-
Diagram 1. Sequence comparison of SATs. The SAT from the following divergent species (see Ref. 40 for phylogenetic tree) are aligned with *E. coli* SAT (percentage of sequence identity with *E. coli* SAT): *Helicobacter pylori* (43% over 162 aa), *Nostoc* sp. (41.7% over 168 aa), *Bacillus cereus* (46% over 161 aa), *Geobacter sulfurreducens* (39.2% over 181 aa), *Wolinella succinogenes* (44.7% over 170 aa), *Azotobacter chroococcum* (42.2% over 161 aa), *Mycobacterium tuberculosis* (45.3% over 170 aa), *Streptococcus pneumoniae* (43.1% over 160 aa), *Halobacterium* sp. (39.3% over 168 aa), *Entamoeba histolytica* (37.3% over 169 aa), *Thermotoga maritima* (41.6% over 161 aa), *Salmonella enterica* subspecies (95.6% over 273 aa), *Arabidopsis thaliana* (50.4% over 248 aa), *Citrullus lanatus* (54.7% over 258 aa), and *Schizosaccharomyces pombe* (40.2% over 241 aa). Amino- and carboxyl-terminal regions (according to *E. coli* SAT) are denoted N and C, respectively, at the bottom of the sequences. Important residues identified in the active site of *E. coli* SAT are clearly conserved and are highlighted in white; other conserved residues are colored gray.
ment of this enzyme. In light of this high-resolution structure, we are also able to discuss the mechanism of SAT from a structural viewpoint.

Previous kinetic studies of SAT (13) suggested a ping-pong mechanism involving an acetyl-enzyme intermediate; recently, Hindson and Shaw (15) have strongly argued that SAT of *E. coli* (like each of the microbial *O*-acetyltransferases studied thus far) is likely to have a steady-state random-order mechanism one that involves a productive ternary complex of substrates and enzyme without a covalent enzyme-substrate in-

| Structure and Mechanism of SAT | 40731 |
|--------------------------------|-------|
| Helicobacter                   | TEIGDVYTIYHGTTL----------CTGKFGRHPLGNRVVGAARLQCVDDLVR 142 |
| Nostoc                        | AIVGDYIYQIVFQL----------CGYKQKGRPLGENVYVAGVLRGLIQIENVR 145 |
| Bacillus                      | CEIGDVYTIYHGGTTL----------CAGLEKRYHPLGNRVVGAARLQCVDDLVR 143 |
| Geobacter                     | AEIGDVYTIYHGTTL----------CVRLEKRYHPLGNRVVGAARLQCVDDLVR 144 |
| Wolinella                     | AIVGDVYTIYQIVL----------CQSERKRRHPLGNRVVGAARLQCVDDLVR 145 |
| Azotobacter                   | AEIGDVYTIYQIVL----------CAGLEKRYHPLGNRVVGAARLQCVDDLVR 143 |
| Mycobacterium                 | AEIGDVYTIYQIVL----------CAGLEKRYHPLGNRVVGAARLQCVDDLVR 143 |
| Streptococcus                 | AEIGDVYTIYQIVL----------CAGLEKRYHPLGNRVVGAARLQCVDDLVR 143 |
| Helobacter                    | AEIGDVYTIYQIVL----------CAGLEKRYHPLGNRVVGAARLQCVDDLVR 143 |
| Entamoeba                     | ALGQKHYQIVL----------CAGLEKRYHPLGNRVVGAARLQCVDDLVR 143 |
| Thermotoga                    | ASVGRGTLLLLH----------CTKRCPLVHPLGNRVVGAARLQCVDDLVR 143 |
| Escherichia                   | AVIEQDSYQVLG----------CQVKKLYHPLGNRVVGAARLQCVDDLVR 143 |
| Salmonella                    | AVIEQDSYQVLG----------CQVKKLYHPLGNRVVGAARLQCVDDLVR 143 |
| Arabidopsis                   | AEVGSYQVLG----------CQVKKLYHPLGNRVVGAARLQCVDDLVR 143 |
| Citrullus                     | AEVGSYQVLG----------CQVKKLYHPLGNRVVGAARLQCVDDLVR 143 |
| Schizosaccharomyces           | AEVGSYQVLG----------CQVKKLYHPLGNRVVGAARLQCVDDLVR 143 |

FIG. 1—continued

| Structure and Mechanism of SAT | 40731 |
|--------------------------------|-------|
| Helicobacter                   | IGANVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Nostoc                        | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Bacillus                      | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Geobacter                     | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Wolinella                     | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Azotobacter                   | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Mycobacterium                 | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Streptococcus                 | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Helobacter                    | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Entamoeba                     | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Thermotoga                    | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Escherichia                   | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Salmonella                    | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Arabidopsis                   | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Citrullus                     | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Schizosaccharomyces           | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |

FIG. 1—continued

| Structure and Mechanism of SAT | 40731 |
|--------------------------------|-------|
| Helicobacter                   | ILSFQGQIANQGGAQIAKPLVA----------SSQERHSREDSQWNLKDAQIEPLDAGI-253 |
| Nostoc                        | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Bacillus                      | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Geobacter                     | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Wolinella                     | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Azotobacter                   | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Mycobacterium                 | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Streptococcus                 | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Helobacter                    | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Entamoeba                     | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Thermotoga                    | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Escherichia                   | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Salmonella                    | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Arabidopsis                   | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Citrullus                     | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |
| Schizosaccharomyces           | IGASVDSLAPGGTAVGAKAATIKT----------DD-171 |

FIG. 1—continued

| Structure and Mechanism of SAT | 40731 |
|--------------------------------|-------|
| Helicobacter                   | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Nostoc                        | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Bacillus                      | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Geobacter                     | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Wolinella                     | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Azotobacter                   | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Mycobacterium                 | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Streptococcus                 | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Helobacter                    | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Entamoeba                     | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Thermotoga                    | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Escherichia                   | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Salmonella                    | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Arabidopsis                   | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Citrullus                     | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |
| Schizosaccharomyces           | KMEVDELWKLKEGLEVE----------AKLEDKKNYH-220 |

FIG. 1—continued

| Structure and Mechanism of SAT | 40731 |
|--------------------------------|-------|
| Helicobacter                   | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Nostoc                        | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Bacillus                      | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Geobacter                     | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Wolinella                     | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Azotobacter                   | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Mycobacterium                 | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Streptococcus                 | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Helobacter                    | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Entamoeba                     | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Thermotoga                    | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Escherichia                   | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Salmonella                    | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Arabidopsis                   | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Citrullus                     | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
| Schizosaccharomyces           | ARLE-10GQGQK2QGVRIR----------EAGWQIHDFSI-229 |
termediate. Furthermore, Hindson (16) has shown that cysteine competes with serine and not with acetyl-CoA as previously proposed (1, 2) and also suggested that binding of cysteine to the serine binding site of SAT may give rise to a reduction in affinity for acetyl-CoA, thus explaining the apparent phenomenon of competitive inhibition with respect to acetyl-CoA observed by steady-state kinetics. The structure of SAT presented here is in complex with cysteine, permitting us to discuss the competitive nature of cysteine with serine and not with acetyl-CoA binding and appraise the manner by which cysteine negatively regulates the first step in its own synthesis.

EXPERIMENTAL PROCEDURES

Expression, Purification, and Crystallization—The E. coli CysE gene (17) was cloned between the NdeI and HindIII restriction sites in expression vector pET28a (Novagen), yielding plasmid pCOCe3. This construct engineers a polyhistidinyl tag at the amino terminus of the recombinant protein. Seleno-L-methionine substituted (SeMet) SAT was expressed and purified from E. coli strain B834(DE3) harboring plasmid pCOCe3 using a method adapted from Budisa et al. (18). Cells from a 40-ml overnight culture in LB were pelleted by centrifugation and resuspended in 2 ml of minimal media, and this was used to inoculate 1.2 liters of minimal media containing trace compounds (CuCl2·2H2O, MnCl2·4H2O, ZnCl2, and Na2MoO4·2H2O), thiamine (10 mg liter−1), biotin (10 mg liter−1), CaCl2 (10 mg liter−1), glucose (3 g liter−1), kanamycin (50 mg liter−1), and seleno-methionine (50 mg liter−1). Suppression of methionine biosynthesis was achieved by adding lysine, phenylalanine, and threonine at 80 mg liter−1 and isoleucine, leucine, and valine at 40 mg liter−1. After 15 h of growth at 37 °C (A600 = 0.7), the cultures were induced with 1 mM isopropyl-β-thiogalactopyranoside, and additional thiamine (10 mg liter−1), biotin (10 mg liter−1), and glucose (3 g liter−1) were also added at this stage. Cells were harvested, after an additional 12 h, by centrifugation at 4 °C and stored frozen (−20 °C) overnight. Cell pellets were defrosted and resuspended in 25 ml of an ice-cold solution comprising 10 mM Tris/HCl buffer, pH 7.5, 2 mM dithiothreitol, and 15 mM phenylmethylsulfonyl fluoride; lysed with lysozyme (20 µg ml−1) and DNase I (10 µg ml−1); and sonicated. Cellular debris was removed by centrifugation, solid ammonium sulfate was stirred slowly into the ice-cold supernatant to give ~45% saturation, and the solution was left on ice to equilibrate for 20 min. Precipitated protein was pelleted by centrifugation, resuspended in 25 ml of wash buffer (10 mM imidazole, 0.5 mM NaCl, and 10 mM Tris/HCl, pH 7.5), and dialyzed overnight against wash buffer. The sample was loaded onto a HiTrap Chelating column (Sigma) precharged with Ni2+ and equilibrated with wash buffer. SAT was eluted from the column with an imidazole gradient as per the manufacturer’s guidelines. Leading fractions were assessed for purity by SDS-PAGE, fractions containing SAT were pooled, and buffer was changed to 10 mM Tris/HCl, 50 mM NaCl, and 10 mM EDTA, pH 7.5, and concentrated to ~10 mg ml−1 in centrifuge modules (Mr, 10,000 cutoff; Amicon). Protein concentration was assessed by serial dilutions on SDS-PAGE; total yield of soluble SAT was 30 mg/liter culture. SAT was stored in the above-mentioned buffer, without additives, at −80 °C. Crystals grew from 0.1 M MES, pH 6.6, 10% (+)-2-methyl-2,4-pentanediol, 0.5 M sodium thiocyanate, 5.5 mM cysteine, and SeMet SAT (10 mg ml−1) sitting drops, 20 °C, within 7 days. Crystals grew with approximate dimensions of 1.0 × 0.8 × 0.4 mm. Crystals were transferred to a solution containing 0.1 M MES, pH 6.6, 40% (+)-2-methyl-2,4-pentanediol, and 5.5 mM cysteine and flash-frozen.

Data Collection—A three-wavelength MAD data set was collected from a single SeMet SAT crystal on station ID29, European Synchrotron Radiation Facility (Grenoble, France). X-ray fluorescence was used to select the three optimal wavelengths around the K absorption edge of selenium. Data were collected to 2.2 Å at 100 K using a Quantum4 charge-coupled device detector, R0.5% of data were collected using a 0.75° oscillation per image, and MOSFLM (19) was used to predict a data collection strategy to ensure complete data sets. Diffraction intensities from the data sets were processed using DENZO and SCALEPACK (20). Subsequently, only the peak wavelength data set was required to solve the structure; data collection statistics are shown in Table I. The crystals belong to space group P321 with unit cell dimensions a = b = 122.0 Å c = 127.5 Å.

Structure Determination and Refinement—Phases were determined using SOLVE (21) with 30–2.2 Å data from the peak wavelength data set. The mean figure of merit from SOLVE was only 0.28, with a Z-score of 94.0 for one solution containing 27 sites, indicating one trimer per asymmetric unit. After density modification using the CCP4 package (22), the figure of merit increased to 0.75, and a readily interpretable map was calculated for the P321 enantiomer. The 27 sites and phases found by SOLVE were put into ARP_WARP (23), mode wapnTrace, option H, to automatically build the protein chain, and this resulted in 39 chains, 242 residues, and a connectivity index of 0.59. The output from this was used as input for further auto-building, mode wapnTrace, option R; this resulted in 16 chains, 716 residues, and a connectivity index of 0.85, providing a solid base from which to start manual building. Initial free-R and R values were 36.5% and 34.0%, respectively. One trimer is present per asymmetric unit. Manual rebuilding and refinement were performed using packages CNS (24) and XtaView (25). Eleven residues from each of the 273-residue monomers are not visible at the carboxyl termini and the polyhistidine tag is not visible at the amino termini in the electron density maps (maps calculated with oA-weighted Fourier coefficients). The active site regions of the SAT trimer are located between monomers, and cysteine was manually built into the positive Fo-Fc density that was present at each site. Water molecules were added gradually during further rounds of side chain adjustment coupled with positional and B-factor refinement. The refinement statistics for the final model are given in Table I. Figures were made using Bobscript (26) and Raster3D (27). Coordinates and diffraction data have been deposited with the Protein Data Bank (code 1T3D).

Table I

| Data collection statistics | Wavelength (Å) | 0.9787 |
|---------------------------|----------------|--------|
| Resolution (Å)            | 2.2            |
| Unique reflections        | 55,977         |
| Redundancy                | 5.5            |
| Completeness (%)          | 99.5 (96.0)    |
| Rmerge (%)                | 0.045 (0.20)   |
| I/σ(I)                    | 13.1           |

Refinement statistics

| Rmerge (%) | 17.5 (18.2) |
| Rfree (%)  | 17.7 (19.5) |
| r.m.s.d.: main chain bonds (Å) angles (°) | 0.019/1.8 |
| No. of protein atoms/water molecules | 5910/415 |
| Average B factor (Å²) (main chain/side chain + water) | 29.5/33.2 |
| % residues: most favored/additionally allowed | 90.4/99.6 |

Values in parentheses are for resolution shell 2.25–2.20 Å.

Rmerge = Rfree calculated on 10% of data excluded from refinement. A final round of refinement using all data the CCP4 program REFMAC5 reduced the R-factor to 15.1% (17.9%).

According to Ramachandran plot.

His316, the catalytic residue in the cysteine/serine binding site, has distorted angles of φ = 70° and ψ = −50° in each protomer.

Structure and Mechanism of SAT
RESULTS

The Structure of SAT—The crystal structure of SAT from E. coli was solved by SAD phasing from SeMet SAT to 2.2 Å. The SAT monomer (Fig. 2a) is composed of two domains: residues 1–140 form an α-helical domain, and residues 141–262 form a left-handed β-helical domain. The α-helical domain is comprised of eight α-helices (α1–α8). From α1 (Cys8–Cys23) there is a short turn, leading to α2 (Pro25–Thr34), which runs anti-parallel to α1; the amino-terminal coil of α2 has β10 helical geometry. α3 (Leu41–Leu53) deviates from the α1-α2 hairpin at an angle of ~80°, following α3 there are two β-turns (type II and type VIII), which then run into α4 (Ala60–Ala73); α4 and α3 run in an anti-parallel fashion. At the carboxyl-terminal end of α4 there is another change in direction, at an angle of ~60°, into α5 (Glu76–Arg91); α5 is in close proximity to α1. There is a β-turn meander, again type II and type VIII, to α6 (Ser99–Leu103). From α6 runs α7 (Lys106–Gln123), which is almost perpendicular in direction, an inverse γ-turn leads to the final helix, α8 (Arg126–Phe140), which in turn leads into the β-helical domain. The β-helical domain is a typical left-handed β-helix comprising of fourteen β-strands forming five coils of the helix. The apex of the β-helix is at residue Pro130. Residues Ala241–Gly262 form a meandering loop that covers one side of the β-helix prism, reaching almost as far as the α-helical domain before it meanders back toward the apex of the β-helix, terminating in a short α-helix. This carboxyl-terminal loop has a helical propensity and forms some H-bonding interactions with itself, but very few with the rest of the monomer. Apart from the final carboxyl-terminal loop there is only one break from the β-helix, this is a loop from residue Gly164 to His193 that has random coil topology. α5 and α8 of the α-helical domain form H-bonding interactions with the first coil of the β-helix.

Quaternary Structure of SAT—The asymmetric unit of the SAT crystals is a trimer with 3-fold symmetry, independent of the crystallographic 3-fold symmetry, forming a three-sided pyramid shape with approximate dimensions of 65 Å (amino-terminal base), 40 Å (carboxyl-terminal apex), and 50 Å in height (Fig. 2b). The monomer-monomer interactions within the trimer are ~70% hydrophobic and involve a vast network of hydrogen bonds, many via water molecules. The main interacting regions between monomers are as follows: α1-α2 loop with α3-α4 loop; β-turns between α5 and α6 with loop Thr185–Lys187 from the β-helix; α6-α7 loop with α8; β-helix corner including residues Asp157–Thr160, Gln178, Ala204, and Ala222 with the side of the adjacent β-helix comprising residues Asp143–Val163, Thr181, Lys207, and Leu227; residue Gly238 with Val239 at the top of the β-helices; and the carboxyl-terminal loop Val250–Gln258 with the loop Thr185–Lys187 from the β-helix. The interacting surface between two monomers in the trimer is 2720 Å². Observing the crystal packing, it seems evident that the SAT trimer interacts with another SAT trimer at the amino-terminal ends (Fig. 2c). The total buried surface between the two trimers is 4450 Å², consistent with an oligomer interaction rather than a crystallization artifact, and has a similar hydrophobic nature to the monomer-monomer interactions (~65%). The dimer of trimers also concurs well with previous preliminary crystallographic analysis, in which SAT crystallized in space group P212121, and exhibited 32 symmetry; in addition, chemical cross-linking and gel filtration studies are consistent with the same quaternary arrangement (12).
The Cysteine Binding Site—Between each monomer in the trimeric interaction, a cysteine molecule is observed (Fig. 2b). The cysteine binding site is situated in a small cleft between adjacent subunits and is formed by $\beta$-strand (Asp 157-Thr 160) and the $\beta$-turn meander from 5-6 of one subunit and the extended loop (Gly 184-His 193) from the left-handed $\beta$-helix of the neighboring subunit. The Gly 184-His 193 loop is rich in conserved small residues and thus provides ample space for substrate/inhibitor binding. The cysteine ligand is flanked by two histidine residues (His 158 and His 193), one from each subunit, which are able to form hydrogen bonds with the cysteine ligand sulfur (His 158 N 2-Cys S 3.2 Å, His 193 N 2-Cys S 3.4 Å). Pro 93 flanks the carboxyl group of the cysteine. Three other charged residues direct their side chains toward the cysteine: two aspartic acid residues (Asp 92 and Asp 157), one from each subunit, and arginine (Arg 192) from the adjacent subunit. Asp 92 and Asp 157 are attracted into place by the amide of the cysteine ligand. The close proximity of these negative side chains is neutralized by local main chain amides (Asp 92 with Pro 93 and Ala 94, Asp 157 with His 158); nevertheless, the closeness of these aspartic acids is reminiscent of aspartic proteases (28, 29). These residues, important in cysteine coordination, are conserved throughout species (Fig. 1). All other residues around the binding pocket are small and uncharged. Many ordered water molecules are observed in the binding pocket and help to form a network of hydrogen bonds (Fig. 3a). The carboxyl terminus of the cysteine ligand forms direct hydrogen bonds with the side chain of Arg 192 (N$\eta$1 and N$\epsilon$) and with the main chain Glu 166 nitrogen via a water molecule. The amino terminus of the cysteine forms direct hydrogen bonds with the carboxylate side chains of Asp 157 and Asp 92 and to the same side chains via two water molecules.

Ternary Intermediate Modeling—Based on the cysteine binding site and structure of SAT presented here, we have manually modeled a ternary intermediate for this reaction. The carboxylate and amine of the serine moiety would be able to form the same hydrogen bonding pattern as the cysteine ligand, with the side chains of Arg 192 (N$\eta$1 and N$\epsilon$) and with the main chain Glu 166 nitrogen via a water molecule. The O$\gamma$ would form a hydrogen bond with the imidazole of His 158. This allows us to suggest a mechanism for acyl transfer similar to that described for chloramphenicol acetyltransferase (30) shown in Fig. 3c.
DISCUSSION

Previous preliminary quaternary structure analysis of SAT by chemical cross-linking, gel filtration, and analysis of crystal packing (12) had indicated that SAT exists as a dimer of trimers, which was in contrast to the tetrameric arrangement proposed originally (31). The structure presented here concludes that the quaternary arrangement of SAT is a dimer of trimers. Each monomer is comprised of an amino-terminal α-helical domain and a carboxyl-terminal left-handed β-helical domain. The trimer formation is approximately parallel with the axis of the β-helical domain, and the dimer of trimers interface is at the amino-terminal α-helical domain.

Between adjacent subunits in each trimer, a small cleft accommodates the cysteine ligand. Upon examination of the structure, a solvent-filled channel is observed extending away from the cysteine ligand binding site. Inspection of the cysteine binding site suggests a mechanism of O-acetylation of serine. When serine is bound in place of cysteine, its Oγ would form a classical “catalytic triad” with the imidazole of His158 from one subunit of the trimer and the carboxylic acid of Asp143 of the adjacent subunit. This provides a mechanism for His158 to act as a base to activate the serine hydroxyl for nucleophilic attack on the carbonyl carbon of the acetyl group of acetyl-CoA. This would form an oxanion tetrahedral intermediate. The imidazole of His193 is in a prime position to stabilize the negative charge, especially if protonated, although nearby main chain amides would also be candidates for this role. The collapse of the tetrahedral intermediate forms CoA and O-acetylserylne (Fig. 3c). The competitive binding of cysteine and serine for the catalytic site provides an explanation for the competitive inhibition (and therefore control) by cysteine described by Hindson (16) but does not explain the apparent lack of S-acetylation of cysteine by acetyl-CoA. Although sulfur is not as electronegative as oxygen, it would still be expected to make an excellent nucleophile for attack on the acetyl carbonyl carbon. The slightly longer sulfur-carbon bond pushes the sulfur atom into a position such that it can make potential hydrogen bonds with both His158 and His193 that are significantly longer than the corresponding oxygen hydrogen bonds (32). Furthermore, the formation of a shorter hydrogen bond interaction between

![Comparison of left-handed β-helical structures with SAT.](image-url)

Fig. 4. Comparison of left-handed β-helical structures with SAT. a, serine acetyltransferase; b, xenobiotic O-acetyltransferase (39% identity over 51 aa; Protein Data Bank code 2XAT); c, galactoside O-acetyltransferase (30% identity over 49 aa; Protein Data Bank code 1KRR); d, carbonic anhydrase (34.8% identity over 23 aa; Protein Data Bank code 1QRE); e, tetrahydrodipicolinate N-succinyltransferase (26% identity over 80 aa; Protein Data Bank code 3TDT); f, UDP-N-acetylglucosamine acetyltransferase (32.1% identity over 53 aa; Protein Data Bank code 1J2Z). Structures are illustrated as Ca traces colored from amino terminus (red) to carboxyl terminus (blue).
The most similar structure to SAT is the serine binding pocket and not the acetyl-CoA binding site, such that the affinity for the acetyl-CoA binding channel is occupied by cysteine to inhibit the channel that is occupied by cysteine to inhibit the O-acetylation of serine reaction performed by SAT. The tight binding of cysteine, particularly the interactions of the two histidine side chains from adjacent monomers, may involve a slight distortion of the acetyl-CoA binding site, such that the affinity for the coenzyme is reduced. This may explain why binding of cysteine appears to inhibit acetyl-CoA binding (16).

The left-handed β-helical fold observed in the SAT structure, which arises from the hexapeptide repeat, is common in the O-acetyltransferase family. The most similar structure to SAT is found to be xenobiotic acetyltransferase (34), which superimposes with a 2.1 Å r.m.s.d. over 105 aa (Co). Other similar structures are galactoside O-acetyltransferase (35), which superimposes with 2.7 Å r.m.s.d. over 119 aa; carbonic anhydrase (36), which superimposes with 2.9 Å r.m.s.d. over 106 aa; tetrahydropyridilinate N-succinyltransferase (37), which superimposes with 3.0 Å r.m.s.d. over 115 aa; and, finally, UDP-N-acetylglucosamine acetyltransferase (10, 38), which superimposes with a 3.3 Å r.m.s.d. over 114 aa (Fig. 4). The overall sequence identities of these structures are quite low, but the important hexapeptide repeat is present. It is observed that the main conserved region in these structures is the left-handed β-helical domain, which does vary in length from ~5 to ~8 coils. Each of these structures oligomerizes to form similar β-helical domains, and is in a prime position to receive the O-acetylsereine product form of SAT to complete the cysteine synthesis. Consequently, six OAS-TL molecules could bind on to the SAT dimer of trimers, resulting in a very efficient cysteine synthesis machinery requiring tight control by negative feedback inhibition. During the process of review of this article, Olsen et al. (39) published the x-ray structure of the homologous enzyme from Haemophilus influenzae in complexes with CoA and cysteine. The results are in agreement with those presented here.

Acknowledgments—We thank Gordon Leonard for help during data collection and the European Synchrotron Radiation Facility for the provision of synchrotron radiation. We are grateful to Jonas Emsley for technical assistance and to Geoff Briggs, Alexandre Gingras, Jane Wiley, Adrian Lloyd, Iain Murray, and Jackie Ellis for fruitful discussions.

REFERENCES

1. Kredich, N. M., and Tomkins, G. M. (1966) J. Biol. Chem. 241, 4955–4963
2. Kredich, N. M., Becker, M. A., and Tomkins, G. M. (1969) J. Biol. Chem. 244, 2428–2433
3. Wiebers, J. L., and Grow, H. R. (1967) J. Biol. Chem. 242, 12–23
4. Smith, I. K., and Thompson, J. F. (1969) Biochem. Biophys. Res. Commun. 35, 929–945
5. Nakamori, S., Kobayashi, S., Kobayashi, C., and Takagi, H. (1998) Appl. Environ. Microbiol. 64, 1607–1611
6. Kredich, N. M. (1996) in Escherichia coli and Salmonella—Cellular and Molecular Biology (Neidhardt, F. C., Curtiss, R., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Núñez, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., eds) 2nd Ed., pp. 514–527, American Society for Microbiology, Washington, D.C.
7. Saito, K. (1999) in Plant Amino Acids (Sing, B. K., ed) pp. 267–292, Dekker Inc., New York.
8. Downie, J. A. (1989) Mol. Microbiol. 3, 1649–1651
9. Murray, I. A., and Shaw, W. V. (1997) Antimicrob. Agents Chemother. 41, 1–6
10. Raetz, C. R. H., and Roderick, S. L. (1995) Science 270, 997–1000
11. Bogdanova, N., and Hell, R. (1997) Plant J. 11, 251–262
12. Hindson, V. J., Moody, P. C., Rowe, A. J., and Shaw, W. V. (2000) J. Biol. Chem. 275, 461–466
13. Leu, L. E., and Cook, F. P. (1994) Biochemistry 33, 2667–2671
14. Wirtz, M., Berkowitz, O., Droux, M., and Hell, R. (2001) Eur. J. Biochem. 288, 686–693
15. Hindson, V. J., and Shaw, W. V. (2000) Biochemistry 42, 3113–3119
16. Hindson, V. J. (2003) Biochem. J. 375, 745–752
17. Denk, D., and Bock, A. (1987) J. Gen. Microbiol. 133, 515–525
18. Budisa, N., Stiepe, B., Demange, P., Eckerskorn, C., Kellerman, J., and Huber, R. (1995) Eur. J. Biochem. 229, 786–796
19. Leslie, A. G. W. (1994) Mof sill Users Guide, MRC-LMB Cambridge, United Kingdom
20. Owczarski, Z., and Minor, W. (2000) in International Tables for Crystallography (Rossman, G. M., and Arnold, E., eds) Vol. F, Kluwer, Dordrecht, The Netherlands
21. Terwilliger, T. C., and Berendzen, J. (1999) Acta Crystallogr. Sect. D Biol. Crystallogr. 55, 849–861
22. CCP4. (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 760–763
23. Penczek, A., Morris, R. J. H., and Luzzati, V. (1999) Nat. Struct. Biol. 6, 458–463
24. Brunger, 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., Simonov, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–921
25. Meele, D. B. (1999) J. Struct. Biol. 125, 156–165
26. Rozans, R. M. (1999) Acta Crystallogr. Sect. D Biol. Crystallogr. 55, 938–940
27. Merrit, E. A., and Murphy, M. P. E. (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 869–873
28. Blundell, T. L., Jenkins, J. A., Sewell, B. T., Pearl, L. H., Cooper, J. B., Tickle, I. J., Veerapandian, B., and Wood, S. P. (1990) J. Mol. Biol. 211, 919–941
29. Aguilar, C. F., Cronin, N. B., Badasso, M., Dreyer, T., Newman, M. P., Cooper, J. B., Hoover, D. J., Wood, S. P., Johnson, M. S., and Blundell, T. L. (1997) J. Mol. Biol. 267, 899–915
30. Moody, P. C. E., and Wilkins, A. J. (1999) in "Protein Engineering" in Focus (Rickwood, D., ed) pp. 50–52, IRL Press, Oxford, United Kingdom
31. Wigley, B. D., Derrick, J. P., and Shaw, W. V. (1990) FEBS Lett. 277, 267–271
32. Gregoret, L. M., Rader, S. D., Fletterick, R. J., and Cohen, F. E. (1991) Protein 9, 9–107
33. Maier, T. H. P. (2003) Nat. Biotechnol. 21, 422–427
34. Beaman, T. W., Sugantins, M., and Roderick, S. L. (1998) Biochemistry 34, 6678–6686
35. Wang, X.-G., Olsen, L. R., and Roderick, S. L. (2002) Structure (Lond.) 10, in press
36. Iversen, T. M., Alber, B. E., Kisker, C., Ferrry, J. G., and Rees, D. (2000) Biochemistry 8, 10363–10370
37. Beaman, T. W., Blanchard, J. S., and Roderick, S. L. (1998) Biochemistry 37, 10363–10370
38. Lee, B., and Suh, S. E. (2003) Proteins 53, 772–774
39. Olsen, R., Huang, B., Vetting, M. W., and Roderick, S. L. (2004) Biochemistry 43, 6013–6019
40. Parker, G., Walshaw, D., O’Hourke, K., Broad, S., Tingey, A., Poole, P. S., and Robson, R. L. (2001) Microbiology 147, 25530–25560
The Structure and Mechanism of Serine Acetyltransferase from *Escherichia coli*
Valerie E. Pye, Andrew P. Tingey, Robert L. Robson and Peter C. E. Moody

*J. Biol. Chem.* 2004, 279:40729-40736.
doi: 10.1074/jbc.M403751200 originally published online July 1, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M403751200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 35 references, 7 of which can be accessed free at http://www.jbc.org/content/279/39/40729.full.html#ref-list-1