Biochem. J. (2012) **446**, 395–404 (Printed in Great Britain) doi:10.1042/BJ20120520

Crystal structure of the dopamine *N*-acetyltransferase–acetyl-CoA complex provides insights into the catalytic mechanism

Kuo-Chang CHENG*, Jhen-Ni LIAO* and Ping-Chiang LYU*†‡[1]

*Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu 30013, Taiwan, †Department of Medical Sciences, National Tsing Hua University, Hsinchu 30013, Taiwan, and ‡Graduate Institute of Molecular Systems Biomedicine, China Medical University, Taichung 40402, Taiwan

The daily cycle of melatonin biosynthesis in mammals is regulated by AANAT (arylalkylamine *N*-acetyltransferase; EC 2.3.1.87), making it an attractive target for therapeutic control of abnormal melatonin production in mood and sleep disorders. *Drosophila melanogaster* Dat (dopamine *N*-acetyltransferase) is an AANAT. Until the present study, no insect Dat structure had been solved, and, consequently, the structural basis for its acetyl-transfer activity was not well understood. We report in the present paper the high-resolution crystal structure for a *D. melanogaster* Dat–AcCoA (acetyl-CoA) complex obtained using one-edge (selenium) single-wavelength anomalous diffraction. A binding study using isothermal titration calorimetry suggested that the cofactor bound to Dat first before substrate. Examination of the complex structure and a substrate-docked model indicated that Dat contains a novel AANAT catalytic triad. Site-directed mutagenesis, kinetic studies and pH-rate profiles confirmed that Glu^{47}, Ser^{182} and Ser^{186} were critical for catalysis. Collectively, the results of the present study suggest that Dat possesses a specialized active site structure dedicated to a catalytic mechanism.

Key words: arylalkylamine *N*-acetyltransferase (AANAT), catalytic triad, dopamine *N*-acetyltransferase (Dat), melatonin, single-wavelength anomalous diffraction.

INTRODUCTION

Melatonin is a major hormonal mediator of light-induced photoperiodic changes in circadian biological events [1–3] and is found in bacteria, protozoa, macroalgae, plants, fungi, invertebrates and vertebrates. In vertebrates, melatonin, a pineal hormone, regulates seasonal and circadian cycles, including the sleep/wake cycle. In invertebrates, melatonin appears to initiate such physiological processes as reproduction and diapause [4,5]. In insects, amine acetylation by AANATs (arylalkylamine *N*-acetyltransferases) occurs during melatonin formation [6], sclerotization [7] and neurotransmitter inactivation [8–11]. These many different activities suggest that several insect AANAT isoforms exist. *Drosophila melanogaster* has two AANAT isoforms encoded by distinct genes [12,13], with Dat (dopamine *N*-acetyltransferase; EC 2.3.1.87) encoded by *aaNAT1a*.

Dat, as an AANAT, is a member of the GNAT (GCN5-related *N*-acetyltransferase) superfamily. GNATs are universally distributed in Nature and use an acyl-CoA to acylate their cognate substrates. Dat catalyses the transfer of the acetyl group in AcCoA (acetyl-CoA) to various arylalkylamines, including dopamine, serotonin, phenylethylamine and tryptamine. GNATs, including Dat, appear to have evolved from a common ancestor, as their sequences all contain four moderately conserved motifs denoted C, D, A and B (listed from the N- to the C-terminus) [14]. Up to the present study, a three-dimensional structure for an insect AANAT had not been solved. Notably, only one other AANAT structure has been solved, that of the *Ovis aries* (sheep) AANAT, SNAT (serotonin *N*-acetyltransferase) in its apo form and complexed with a bisubstrate analogue [15,16]. In SNAT, His^{120}, His^{122} and Tyr^{168} were suggested to participate in the catalytic reaction [17]. The Dat sequence is not similar to that of SNAT, suggesting that the Dat catalytic residues and those of SNAT might differ, as might their catalytic mechanisms. In the present study, we describe the 1.46 Å (1 Å = 0.1 nm) resolution crystal structure of a truncated *D. melanogaster* Dat ($Dat_{21–230}$)–AcCoA complex obtained from one-edge (selenium) single-wavelength anomalous diffraction data [18]. By combining examination of the crystal structure and a ternary ($Dat_{21–230}$)–AcCoA–docked-substrate (tryptamine) model, and site-directed mutagenesis in conjunction with a kinetic study, we identified a probable catalytic triad and could propose a catalytic mechanism. This is the first structural study of an insect AANAT and has therefore provided insights into the catalytic mechanism and the structure–function relationships of GNATs.

EXPERIMENTAL

Protein expression and purification

The *Dat* gene was obtained from the DGRC (*Drosophila* Genomics Resource Center). PCR-based mutagenesis techniques were employed to make the N-terminal and C-terminal truncation mutant [19]. DNA encoding $Dat_{21–230}$ was ligated into a pGEX-6P-3 vector and transformed into *Escherichia coli* BL21 (DE3) cells for subsequent expression of GST (glutathione transferase)-tagged $Dat_{21–230}$. The Dat mutants were achieved using the QuikChange® site-directed mutagenesis kit (Stratagene). The cells were grown in 2-litre flasks that contained 400 ml of Luria–Bertani broth and ampicillin (100 μg/ml), which were then incubated at 37 °C with shaking until the D_{600} of the culture reached 0.5–0.6. The flasks were cooled to room temperature (25 °C) and then treated with isopropyl β-D-thiogalactopyranoside (final concentration, 1 mM). The cultures were then incubated at 24 °C for 7 h.

Abbreviations used: AANAT, arylalkylamine *N*-acetyltransferase; AcCoA, acetyl-CoA; Dat, dopamine *N*-acetyltransferase; DTT, dithiothreitol; GNAT, GCN5-related *N*-acetyltransferase; GST, glutathione transferase; ITC, isothermal titration calorimetry; RMSD, root mean square deviation; SNAT, serotonin *N*-acetyltransferase.

The atomic co-ordinates and structural factors have been deposited in the PDB under accession code 3TE4 for the $Dat_{21–230}$–AcCoA complex structure.

[1] To whom correspondence should be addressed (email pclyu@mx.nthu.edu.tw).

The cell pastes were suspended in 30 ml of lysis buffer [PBS, pH 6.9, 10 mM DTT (dithiothreitol), 10 % (v/v) glycerol and 10 mM EDTA], and the cells were lysed by passage through a High-Pressure Homogenizer EmulsiFlex-C3 at 15000 psi (1 psi = 6.9 kPa). Insoluble protein and cell debris were removed by centrifugation (4 °C, 48000 ***g***, 30 min), and the supernatant was applied to a 5 ml column of glutathione–Sepharose 4B (GE Healthcare). After elution from the glutathione–Sepharose 4B column, fractions that contained the fusion protein were combined and dialysed overnight against 20 mM Tris/HCl, pH 7.0, 150 mM NaCl, 1 mM DTT, 0.5 mM EDTA and 0.6 % Triton X-100. The GST tag was removed by PreScission protease (2 units/mg of fusion protein, Sigma), and the reaction was monitored by SDS/PAGE. Then the GST tag and the protease were removed by passage through 5 ml of glutathione–Sepharose 4B equilibrated with lysis buffer. The recovered protein was dialysed against 50 mM Tris/HCl, pH 7.0, 10 mM DTT, 1 mM EDTA and 10 % (w/v) glycerol, and then concentrated using a 5000 Da molecular mass cut-off (PES) Vivaflow 50 ultrafiltration apparatus (Sartorius Stedim Biotech). The concentrated protein was then subjected to anion-exchange chromatography through 5 ml of HiTrap Q HP column (GE Healthcare); fractions containing $Dat_{21–230}$ were collected and the buffer exchanged to lysis buffer by dialysis. The protein, which had been concentrated to 10–15 mg/ml, was then subjected to size-exclusion chromatography (Sephacryl S-100 HR, GE Healthcare). The purified protein was subjected to SDS/PAGE and quadrupole time-of-flight MS (Quattro Ultima, Micromass). The same protocol was used to purify wild-type Dat, Dat mutant and selenomethionine-labelled $Dat_{21–230}$ from an M9-medium expression culture [20].

Crystallization and data collection

The purified selenomethionine-labelled $Dat_{21–230}$ was mixed with a 4-fold molar excess of AcCoA. Crystallization trials were performed at 15 °C using the hanging-drop vapour-diffusion method, and crystals were obtained using the Wizard I random sparse matrix crystallization screen (Emerald Biosystems). The crystals most suitable for X-ray diffraction were obtained after 7 days at 4 °C using a mixture of 1 μl of concentrated protein in 50 mM sodium phosphate, pH 6.8, 5 mM DTT, 1 mM EDTA and 1 μl of reservoir solution (0.1 M imidazole, pH 8.5, 1.0 M NaH_2PO_4, 1.6 M K_2HPO_4 and 0.2 M NaCl), which was equilibrated against 200 μl of reservoir solution. Diffraction data (1.46 Å resolution) were collected at 100 K at the SPXF beamline BL13B1, National Synchrotron Radiation Research Center (Taiwan) using an ADSC Quantum-315r CCD Area Detector. A complete single-wavelength anomalous dispersion dataset was collected at the selenium-absorption high-remote wavelength (0.96369 Å, determined by fluorescence scan). The datasets were processed and scaled using HKL2000 [21]. Crystallographic and refinement statistics are provided in Table 1.

Structure resolution and refinement

The seven selenium sites in the asymmetric unit of the complex were located using SHELX C/D/E [22]. The overall figure of merit was 0.75. The electron density map was of good quality, such that 99 % of the structure could be built automatically by wARP [23]. The remaining 1 % was manually built with the aid of Coot [24]. Alternating model-building and refinement cycles using REFMAC5 [25] were performed with appropriate entries made in the dictionary for AcCoA. The final *R*-factor was 0.140 (R_{free} = 0.198). The overall geometry of the final structure was assessed by PROCHECK [26]. Structural representations and models were generated using PyMol (Schrödinger).

Table 1 Data collection and refinement statistics

Values in parentheses are for the highest-resolution shell. $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - <I(hkl)>|/\Sigma_{hkl}\Sigma_i I_i(hkl)$. R_{work} and $R_{free} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$, where R_{free} was calculated over 5 % of amplitudes that were chosen at random and not used in refinement. Ramachandran analysis was from PROCHECK [26].

Parameter	Selenomethionine derivative
PDB code	3TE4
Data collection	
Protein	$Dat_{21–230}$–AcCoA
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	44.03, 56.62, 83.94
α, β, γ (°)	90, 90, 90
Wavelength (Å)	0.96369
Resolution (Å)	30–1.46 (1.51–1.46)
R_{merge}	0.058 (0.209)
$I/\sigma I$	23.4 (9.25)
Completeness (%)	98.7 (98.7)
Redundancy	5.5 (5.3)
Refinement	
Resolution (Å)	23.6–1.46 (1.50–1.46)
Number of reflections	36626
R_{work}/R_{free}	0.140/0.198
Number of atoms	
Protein	1687
AcCoA	51
Water	345
B-factors	
Protein	13.7
AcCoA	8.9
Water	30.5
RMSD	
Bond lengths (Å)	0.019
Bond angles (°)	1.494
Ramachandran plot statistics (%)	
Residues in most favoured regions (%)	97.6
Residues in additional allowed regions (%)	2.4
Residues in generously allowed regions (%)	0

CD

CD spectra were acquired using an Aviv 202 spectropolarimeter (AVIV Biomedical) and a 1 mm pathlength cuvette. The far-UV CD spectra of the proteins, each corrected for the contribution of the buffer (7.5 μM protein, 5 mM sodium phosphate, 0.5 mM DTT and 0.1 mM EDTA, pH 6.7) are reported as mean residue ellipticity ([θ], deg·cm^2·dmol^{-1}). Ellipticities were measured between 190 nm and 260 nm at 1 nm intervals at 25 °C. The CD data were displayed using GraphPad Prism (GraphPad Software).

Acetyltransferase kinetic assay

The assay detected the reaction of Ellman's reagent [5,5′-dithiobis-(2-nitrobenzoic acid)] [27] with the thiolate anion of CoA that was generated by the enzymatic acetyl-transfer reaction. Reactions were carried out in 0.3 ml of 50 mM sodium phosphate, pH 7.2, 500 mM NaCl, 2 mM EDTA, 0.05 mg/ml BSA and 2 mM AcCoA, with a variable tryptamine concentration (0–2 mM) at 30 °C in 1.5 ml microfuge tubes. Reactions were initiated by addition of enzyme (30 nM final concentration) that had been diluted 1000-fold with reaction buffer immediately before use and had been kept on ice prior to use in the assay. Reactions were

Table 2 Kinetic parameters for Dat *in vitro*

The activities of the mutants were normalized to that of the wild-type enzyme and all activities were expressed as percentages. The assays were performed at 30 °C in 50 mM sodium phosphate buffer, pH 7.2, with 500 mM NaCl, 2 mM EDTA, 0.05 mg/ml BSA and 2 mM AcCoA using tryptamine as the substrate with different concentrations (0, 0.05, 0.1, 0.2, 0.4, 0.8, 1.5 or 2 mM).

Dat form	k_{cat} (s^{-1})	Tryptamine K_m (mM)	AcCoA K_m (mM)	Tryptamine k_{cat}/K_m ($mM^{-1}\cdot s^{-1}$)	Enzyme activity (%)
Wild-type Dat	21.9 ± 1.4	0.47 ± 0.09	0.43 ± 0.09	47 ± 9	100.0
S182A	1.8 ± 0.1	0.65 ± 0.09	0.36 ± 0.05	2.7 ± 0.4	5.7
S186A	4.8 ± 0.4	0.55 ± 0.07	0.39 ± 0.05	8.7 ± 1.2	18
E47A	0.13 ± 0.01	1.2 ± 0.1	0.04 ± 0.01	0.11 ± 0.01	0.23
Y185F	22.0 ± 1.3	0.47 ± 0.08	0.43 ± 0.02	47 ± 8	100.0
Y185A	22.0 ± 1.2	0.49 ± 0.07	0.40 ± 0.03	45 ± 6	96

quenched for 3 min with 0.6 ml of 3.2 M guanidinium/HCl and 0.1 M sodium phosphate, pH 7.2. Then, 0.1 ml of 2 mM Ellman's reagent, 0.1 M sodium phosphate, pH 6.8, and 10 mM EDTA was added to each solution, which were vortex-mixed, and allowed to stand for 5 min before absorbance readings were performed at 412 nm. A background correction was also made for AcCoA, because the AcCoA preparation was contaminated with a small amount of a thiol compound. The rate of conversion of AcCoA into CoA in the absence of tryptamine was negligible. Rates were measured under initial conditions (<10 % turnover of the limiting substrate). The concentration of CoA produced was calculated on the basis of a standard curve of CoA. All assays were performed in triplicate, and rate values generally agreed within 10 %. Kinetic constants were calculated by fitting the data to the Michaelis–Menten equation using a non-linear curve fit (GraphPad Prism), and the values are shown as means ± S.D. in Table 2. The error bars in the substrate–velocity plot represent the S.D. of each data point. The kinetic constants in the pH range 7–11 were measured as described above, with the exception that the reactions were carried out in sodium phosphate buffer at the indicated pH values (5.0, 6.0, 7.0, 8.0, 9.0 and 10.0). The results plotted were the average of three separate experiments. Curve fits for k_{cat} and k_{cat}/K_m data were calculated as described previously [28].

ITC (isothermal titration calorimetry)

Binding of dopamine and AcCoA to Dat was measured by ITC using an Micro VP-ITC microcalorimeter (MicroCal). Aliquots of 6 μl of 5 mM AcCoA or dopamine were titrated at 25 °C by injection into protein (40 μM in 1.4 ml) in 50 mM sodium phosphate (pH 6.8), 5 mM DTT and 1 mM EDTA. The titration with 5 mM dopamine was performed in the presence of 1 mM CoA to avoid enzyme turnover. Background heats from ligand to buffer titrations were subtracted, and the corrected heats from the binding reaction were used to derive values for the stoichiometry of the binding (n), the association constant (K_a), the apparent enthalpy of binding (ΔH), the change in Gibbs free energy occurring upon binding (ΔG) and the entropy changes ($\Delta S = \Delta H - \Delta G\ /T$) at 298 K. Data were fitted using a non-linear least-squares routine using a single-site binding model with Origin for ITC v5.0 (MicroCal).

Molecular docking

The $Dat_{21–230}$–AcCoA structure was docked with tryptamine using GOLD [29]. Hydrogen atoms were added to the tryptamine using Discovery Studio 2.0 (Accelrys). The program automatically docked tryptamine into the active site of the $Dat_{21–230}$–AcCoA complex in all possible orientations. GoldScore was used as the fitness function with standard default settings employed for all calculations. The configuration file was defined as follows. The docking site was defined as all atoms within 10 Å of a specified centroid (x, y, z co-ordinates: 80.43, 84.41, 90.48). The best-ranked 25-docked poses were computationally evaluated and visually inspected. The pose with the most favourable energy was selected and further refined by steepest descendent minimization. Substrate–protein interactions were analysed by Discovery Studio 2.0 and Ligplot v.4.0 [30].

RESULTS AND DISCUSSION

Characterization of a truncated Dat

We truncated the *D. melanogaster Dat* gene because the PredictProtein server [31] had predicted that residues 1–20 and 231–240 would be disordered in the native structure, which might have interfered with crystallization. The truncated Dat (called $Dat_{21–230}$) contains the four highly conserved motifs and has a molecular mass of 24411 Da, which is consistent with its theoretical molecular mass. Moreover, $Dat_{21–230}$ and full-length Dat have very similar CD spectra and identical enzymatic activities. The ITC experiments revealed the similar affinity of substrate and cofactor for $Dat_{21–230}$ and full-length Dat (Supplementary Figure S1 at http://www.BiochemJ.org/bj/446/bj4460395add.htm). In summary, these results showed that the truncation form of Dat is similar to native Dat both functionally and structurally, without detectably having an impact on the structural integrity of the enzyme.

Crystal structure of the $Dat_{21–230}$–AcCoA complex

The crystal structure of the $Dat_{21–230}$–AcCoA complex was solved at 1.46 Å resolution; the diffraction data having been acquired by one-edge (selenium) single-wavelength anomalous diffraction. Each asymmetry unit contains one $Dat_{21–230}$–AcCoA complex. $Dat_{21–230}$ is a globular protein containing a seven-strand mixed β-sheet surrounded by nine α-helices (Figure 1A). The β-sheet is highly twisted and can be considered as two β-sheets ($\beta1$–$\beta4$ and $\beta5$–$\beta7$), which are partially separated at $\beta4$ and $\beta5$. The strands in each neighbouring pair are orientated in opposite directions (antiparallel), except for $\beta4$ and $\beta5$, which are parallel with each other. The β-sheet separates the α-helices into two groups that reside on opposite sides of the β-sheet. Motif C (Tyr^{22}–Phe^{43}) is located in $\beta1$ and $\alpha1$. Motif D is located in $\beta2$ and $\beta3$ (Lys^{76}–Asn^{90}). Motif A (Leu^{139}–Gly^{173}) is located in $\beta4$, $\alpha5$ and $\alpha6$. Motif B (Val^{179}–Phe^{199}) is located in $\beta5$, $\alpha7$ and $\beta6$.

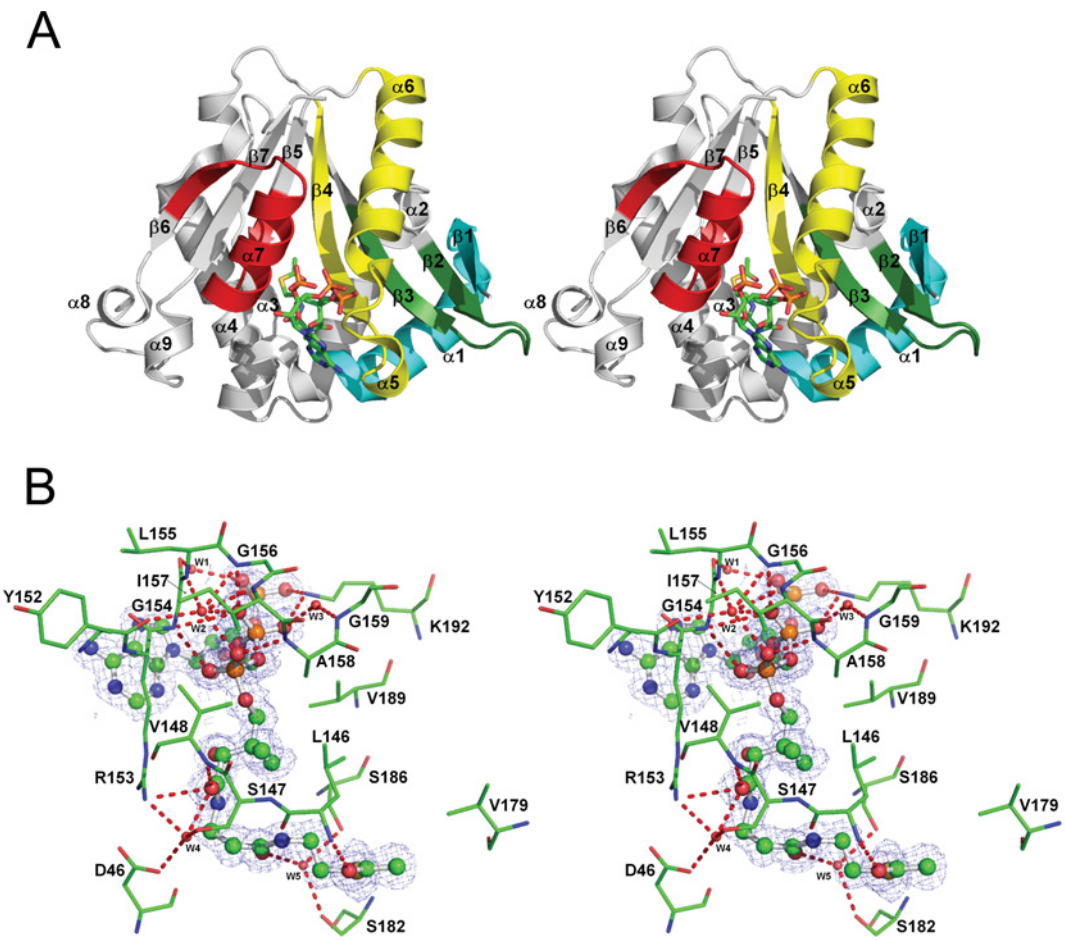

Figure 1 The crystal structure of the $Dat_{21–230}$–AcCoA complex and interactions between $Dat_{21–230}$ and AcCoA

(**A**) A stereo ribbon diagram of the $Dat_{21–230}$–AcCoA complex. Dat contains nine α-helices and seven β-strands. The conserved GNAT motifs C, D, A and B are coloured cyan, green, yellow and red respectively. AcCoA is shown as a stick model and is situated in its binding cavity. (**B**) A stereo $2F_o - F_c$ electron density map for AcCoA contoured at 1 σ within the stick model of the AcCoA-binding site. Hydrogen bonds are shown as broken lines. A certain number of the $Dat_{21–230}$–AcCoA interactions are mediated by hydrogen bonding with water molecules (W1–W5, small red balls).

The AcCoA-binding site

The position of AcCoA was clearly defined in the electron density map and is located in a cleft that appears to be part of the funnel-shaped active site. The AcCoA conformation is approximately S-shaped. The cleft is almost completely covered by motifs A and B, which include the aforementioned secondary structural elements and the loop between α5 and α6 (Figure 1A). The AcCoA pantetheine is bound in this cleft at a position where β4 and β5 diverge. AcCoA forms hydrogen bonds with the backbone atoms of Leu[146], Val[148], Arg[153], Gly[154], Gly[156] and Ala[158] and one salt bridge with Lys[192]. Additionally, it also forms water-mediated hydrogen bonds with Asp[46], Ser[147], Tyr[152], Leu[155], Ile[157], Gly[159], Ser[182] and Ser[186], and van der Waals interactions with Val[148], Arg[153], Gly[154], Ala[158], Val[179] and Val[189] (Figure 1B and Supplementary Table S1 at http://www.BiochemJ.org/bj/446/bj4460395add.htm). Residues in the turn that joins α5 and α6 of motif A form a positively charged pocket that contains the pyrophosphate (Supplementary Figure S2A at http://www.BiochemJ.org/bj/446/bj4460395add.htm). Notably, the residues in the turn are the most highly conserved within motif A of insect AANATs (Figure 2). The AcCoA acetyl is located between β5 and a β-bulge in β4 and appears to be stabilized by contacts with the protein. Interestingly, in this β-bulge, the Lys[144] and Ile[145] side chains are on the same face of the β-sheet (Supplementary Figure S3 at http://www.BiochemJ.org/bj/446/bj4460395add.htm). The carbonyl oxygen of the AcCoA acetyl is within hydrogen-bonding distance of the Leu[146] backbone nitrogen, which is positioned distinctively as a result of the β-bulge. Additionally, the main-chain carbonyl oxygen of Leu[146] is within hydrogen-bonding distance of the nearest pantetheine nitrogen (N4P). By constraining AcCoA at two points via these two possible Leu[146] backbone hydrogen bonds, the AcCoA sulfur atom is within hydrogen-bonding distance of the Ser[186] hydroxy group (discussed below). Of note, Lys[192] can form a distinct salt bridge with 3′-phosphate or α-phosphate of the ADP moiety. This feature has not been observed in other AANAT complex structures to date.

Substrate and cofactor binding study by ITC

To investigate the binding property of Dat for its substrate and cofactor, ITC measurements were used to accurately determine the binding parameter and thermodynamic characterizations of Dat with dopamine and AcCoA respectively (Figure 3). The ITC data revealed that the interaction of Dat with dopamine appeared to be an exothermic reaction with enthalpy (ΔH) = − 7.85 kcal/mol, and entropy (ΔS) = − 5.32 cal/mol per K. Fitting to one set of the site-binding model, the K_a of dopamine binding to Dat was

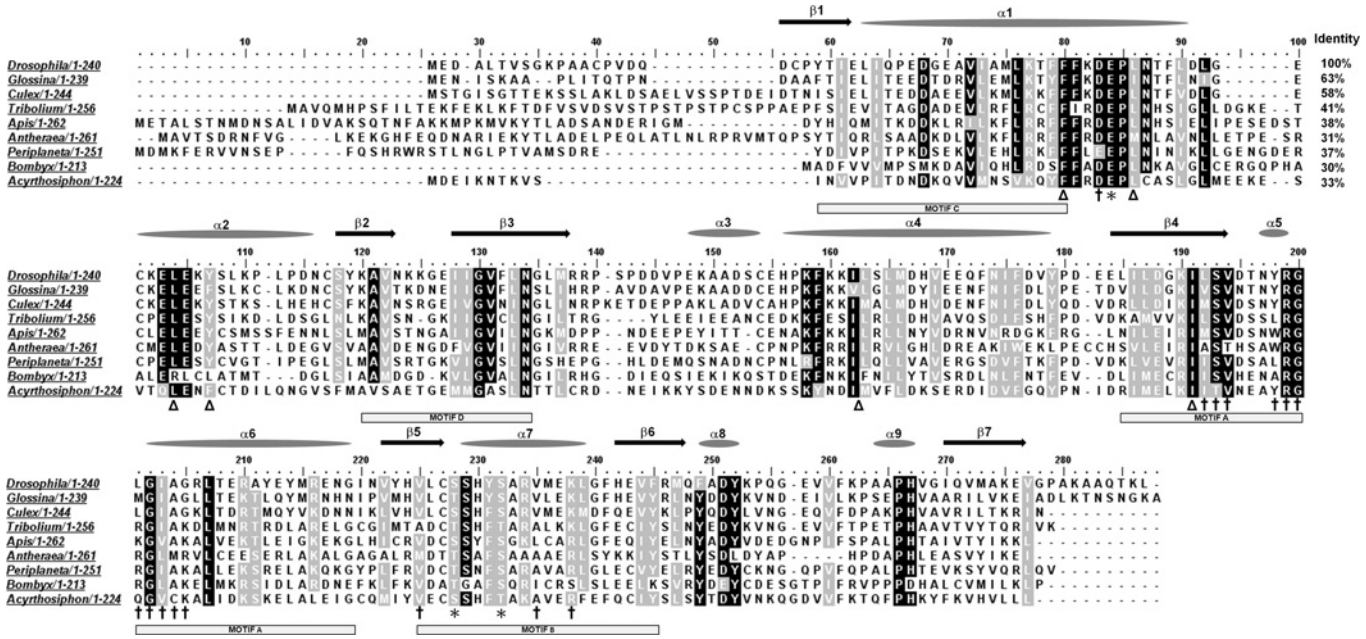

Figure 2 Sequence alignment of insect AANATs

Alignment of Dat and eight other insect AANAT sequences. Identical and conserved residues found in ⩾80 % of the protein sequences at a given position are highlighted in black and grey respectively. $Dat_{21–230}$ secondary structural elements are shown above the sequence alignment. *, Dat catalytic-triad residues; †, residues implicated in the binding of $Dat_{21–230}$ and AcCoA; and Δ, non-polar residues predicted to be involved in tryptamine binding according to the ternary-docking model. The conserved motifs C, D, A and B are shown as boxes. Motif C (Tyr^{22}–Phe^{43}) is located in $\beta1$ and $\alpha1$. Motif D is located in $\beta2$ and $\beta3$ (Lys^{76}–Asn^{90}). Motif A (Leu^{139}–Gly^{173}) is located in $\beta4$, $\alpha5$ and $\alpha6$. Motif B (Val^{179}–Phe^{199}) is located in $\beta5$, $\alpha7$ and $\beta6$. Species: *Acyrthosiphon pisum*, *Antheraea pernyi*, *Apis mellifera*, *Bombyx mori*, *Culex quinquefasciatus*, *Drosophila melanogaster*, *Glossina morsitans*, *Periplaneta americana* and *Tribolium castaneum*.

calculated to be 3.88×10^4 M^{-1}, and the binding constant K_d was 25.9 μM. The interaction of Dat with AcCoA had a higher affinity, with a K_a of 1.88×10^5 M^{-1}. As shown in Table 3, a significant amount of heat was released when Dat associated with AcCoA, indicating that the binding interactions had significant enthalpic contributions ($\Delta H = -11.5$ kcal/mol), which were dictated by hydrogen bonding and van der Waals interactions (Figure 1B and Supplementary Table S1). Analysis of the ITC data revealed a slightly unfavourable entropic contribution ($\Delta S = -14.4$ cal/mol per K), possibly indicating that the Dat structure was slightly stabilized upon binding to AcCoA and that very few solvent molecules were freed from the binding pocket (solvent entropy). These results were in agreement with crystallographic data, as indicated by: (i) the significant reduction of the *B* values for three regions upon formation of the Dat–AcCoA complex (conformational entropy) (Supplementary Figure S4 at http://www.BiochemJ.org/bj/446/bj4460395add.htm); and (ii) there were not only a few water molecules in contact with AcCoA through hydrogen bonding, but also there were some water molecules in the binding pocket. To determine the binding order for the substrate and cofactor towards Dat, the dopamine titration experiment indicated that there was no significant binding to Dat in the absence of CoA (results not shown). This ITC result suggested that the cofactor binds first to Dat followed by substrate binding.

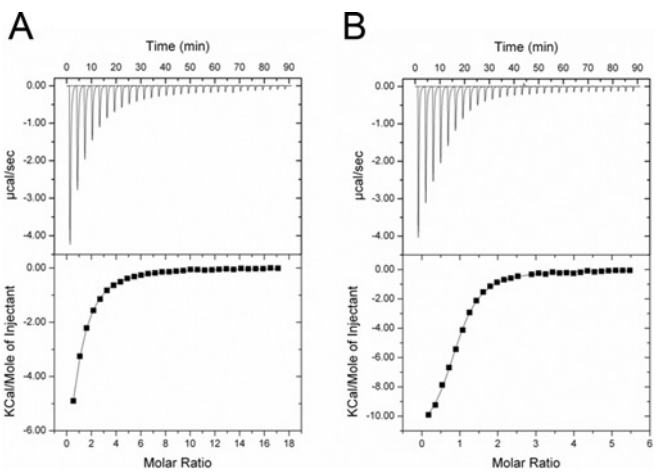

Figure 3 ITC studies of the substrate and cofactor binding to Dat

Titrations of dopamine (**A**) and AcCoA (**B**) against Dat were performed. The isothermal calorimetric enthalpy changes (upper panels) and the integrated heat data (lower panels) for the dopamine and AcCoA titrating to Dat show a direct binding fitted to a single-site binding model. The protein concentration in the sample cell was 40 μM, and ligand concentrations in the syringe were 5 mM for dopamine and AcCoA. Each of the spikes on the upper panels corresponds to a 6 μl injection of ligand into the calorimeter cell. Thermodynamic values are shown in Table 3.

Comparison of $Dat_{21–230}$ with other *N*-acetyltransferase domains revealed a structurally conserved core domain

As shown by iSARST [32], the $Dat_{21–230}$ tertiary structure is similar to those of *N*-acetyltransferase domains found in other GNATs (Supplementary Figure S5 at http://www.BiochemJ.org/bj/446/bj4460395add.htm), including SNAT [16], tabtoxin-resistance protein [33], aminoglycoside 6′-*N*-acetyltransferase [34], histone acetyltransferase [35], and glucosamine-6-phosphate *N*-acetyltransferase 1 [36], which had structure diversity scores in comparison with the $Dat_{21–230}$ structure of 3.47, 4.16, 4.58, 5.75 and 6.63 respectively (Table 4). The pairwise alignments of the C, D, A and B motifs in $Dat_{21–230}$ with those

Table 3 Thermodynamic analysis of binding of the substrate and cofactor to Dat

1 cal ≈ 4.184 J.

Molecule	K_a (10^4 M^{-1})	K_d (μM)	ΔH (kcal/mol)	ΔS (cal/mol per K)	ΔG (kcal/mol)	n
Dopamine	3.88 ± 0.26	25.9 ± 2.5	− 7.85 ± 0.17	− 5.32	− 6.26	1.03 ± 0.03
AcCoA	18.8 ± 2.6	5.43 ± 1.09	− 11.5 ± 0.3	− 14.4	− 7.20	0.88 ± 0.01

Table 4 Pairwise comparisons of the $Dat_{21–230}$–AcCoA structure with five *N*-acetyltransferases by the iSARST server [32]

AAC (6′)-Ii, aminoglycoside 6′-*N*-acetyltransferase type Ii; GNA1, glucosamine-6-phosphate *N*-acetyltransferase 1; HPA2, histone acetyltransferase 2; TTR, tabtoxin-resistance protein.

N-acetyltransferase	Aligned residues	Identity*	Similarity†	RMSD for the Cα carbons (Å)	Structure diversity‡ iSARST score	PDB code
SNAT	125	12 % (20/166)	23 % (38/166)	1.89	3.47	1CJW
TTR	128	8.4 % (14/166)	26 % (43/166)	2.33	4.16	1GHE
AAC (6′)-Ii	120	12 % (22/181)	25 % (45/181)	2.20	4.58	1B87
HPA2	127	15 % (22/150)	29 % (44/150)	3.41	5.75	1QSM
GNA1	126	15 % (23/154)	31 % (47/154)	3.82	6.63	1I12

*Values in parentheses: (number of identical residues in the aligned regions/inquiry protein sequence length) × 100.
†Values in parentheses: (number of conservatively exchanged residues in the aligned regions/inquiry protein sequence length) × 100.
‡Structure diversity: RMSD/(number of residues in the aligned sequence/total number of residues in the shorter-length protein)$^{1.5}$.

in the other five *N*-acetyltransferase domains [37] are generally consistent, although some loop structures are not well aligned. Comparison of the values for the Cα RMSDs (root mean square deviations) of structurally equivalent residues in $Dat_{21–230}$ and in each of the other proteins indicates that the core fold structures are very similar despite a low level of sequence identity and similarity (Table 4). The conformation of the bound AcCoA in $Dat_{21–230}$ and its orientation compared with those of the other five GNAT proteins is virtually indistinguishable. Additionally, each structure possesses a β-bulge at a position equivalent to Lys^{144} and Ile^{145} in $Dat_{21–230}$. The presence of the β-bulge is remarkably well conserved in active sites among GNATs, suggesting a critical role in the formation of the AcCoA-binding site [38,39]. However, one striking difference between $Dat_{21–230}$ and the other five structures exists: the tyrosine residue found in each of those five structures, whose phenolic hydroxy group is within hydrogen-bonding distance of the AcCoA sulfur atom and may therefore stabilize the thiolate anion of the departing CoA molecule, is replaced by Ser^{186} (see below and Supplementary Figures S5 and S6 at http://www.BiochemJ.org/bj/446/bj4460395add.htm). The critical role of this tyrosine residue in catalysis was also reported in these acetyltransferases of GNATs [17,40]. A structure-based alignment of the Dat and SNAT sequences shows that the sequences of the conserved GNAT motifs were unambiguously aligned (Supplementary Figure S7 at http://www.BiochemJ.org/bj/446/bj4460395add.htm). Inspection of their structures revealed that Dat has a deeper funnel leading into its active site (Supplementary Figure S2). A remarkable feature that distinguishes the Dat structure from that of SNAT is that the Dat substrate-binding site is more deeply buried. The geometry and non-polar nature of the funnel interior provide a basis for the enzymatic properties of Dat. The ITC results revealed Dat had a much higher affinity for substrate than that of SNAT (Dat, $K_d = 222.5$ μM for serotonin, results not shown; SNAT, $K_d = 1299$ μM for serotonin) [41]. The deeply embedded cavity provides a non-polar environment that can bind the substrate more avidly. The shape and size of the cavity must therefore dictate substrate selectivity and specificity. Some common protein–acetyl and protein–pantetheine interactions were found in both structures. Nevertheless, $Dat_{21–230}$ has more residues involved in the interactions with AcCoA than those of SNAT (17 residues in Dat; nine residues in SNAT) [16]. The ITC data also provide evidence of an increased binding affinity of approximately 45-fold between Dat and AcCoA compared with that of SNAT (Dat, $K_d = 5.43$ μM; SNAT, $K_d = 242$ μM) [41].

A ternary docking model

To begin to understand how a substrate binds in Dat, tryptamine was docked into the $Dat_{21–230}$–AcCoA crystal structure (Figure 4A). Our docking strategy was to bring the tryptamine amino group in close proximity to the acetyl carbonyl group of AcCoA as would be expected during the reaction. A non-polar pocket, denoted the dopamine-binding pocket, that is mainly formed by Phe^{43}, Leu^{49}, Leu^{61}, Tyr^{64}, Ile^{116} and Ile^{145}, was found inside the funnel-shaped active site and contained the indole ring of the docked tryptamine, which made extensive non-polar contacts with the side chains of these residues. Given the geometry and non-polar nature of the pocket, steric factors and non-polar interactions should dominant substrate–Dat binding. The six residues are highly conserved in insect AANATs (Figure 2). Van der Waals interactions and packing of non-polar residues therefore are probably responsible for the binding of arylalkylamine substrates in insect AANATs. The dopamine-binding site is solvent-accessible via the funnel, so that a substrate can enter and an N-acetylated product can leave, and it is adjacent to the acetyl in AcCoA.

Potential catalytic residues

Alignment of nine insect AANAT sequences revealed that the sequence identity among them is substantial, between 30 and 63 % (Figure 2). There are many regions in which > 80 % of the residues are conserved. Examination of the $Dat_{21–230}$–AcCoA structure and the substrate-docked model showed that certain residues are important for substrate and AcCoA binding, and others may be involved in catalysis (Figures 2B and 4A). The conserved insect AANAT Glu^{47}, Ser^{182} and Ser^{186} may be the catalytic residues, because they are situated near the leaving and acceptor groups

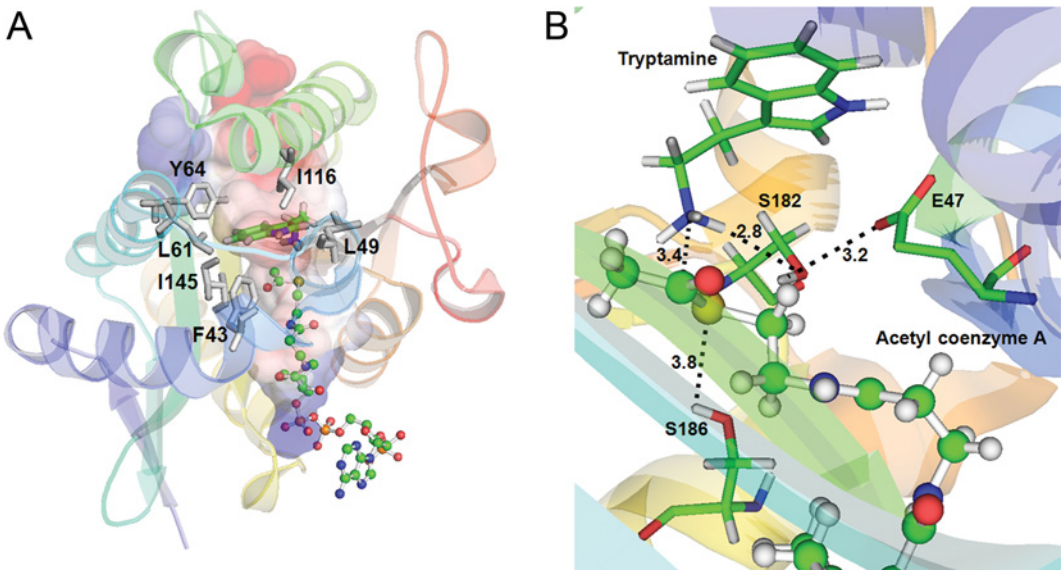

Figure 4 A ternary-docking model

(**A**) Tryptamine was docked into the substrate-binding cavity of the $Dat_{21–230}$–AcCoA complex by GOLD [29,50]. The enzyme is shown as a ribbon diagram. The surface electrostatic potential representation of the cavity is shown. Tryptamine is shown as a stick model. Non-polar residues are shown as grey stick models. AcCoA is shown as a ball-and-stick model. Both substrate and cofactor are within the cavity. (**B**) The catalytic triad residues of Dat. The tryptamine and catalytic triad residues are represented as stick models. AcCoA is shown as a ball-and-stick model. The broken lines identify interactions between the catalytic residues, tryptamine and AcCoA. The numbers represent distance separations between various atoms in angstroms.

in AcCoA and tryptamine respectively (Figure 4B). Besides, the Dat sequence contains a Tyr^{185} just before Ser^{186}. As tyrosine is essential for the catalysis in some acetyltransferases of the GNAT superfamily, the potential involvement of Tyr^{185} in catalysis also requires elucidation. Therefore we individually mutated each of these residues to either alanine or phenylalanine (only for Tyr^{185}) and examined their effects on Dat catalysis.

Comparison of the kinetic parameters of E47A, S182A, Y185F/Y185A, S186A and wild-type Dat

The secondary structures of all mutants and wild-type Dat were identical according to far-UV CD spectroscopy (Figure 5A). Given the highly variable effects on both the turnover and the Michaelis–Menten constants for the enzymatic activity of the mutants except for Y185F/Y185A (Figure 5B and Table 2), Glu^{47}, Ser^{182} and Ser^{186} probably function as a catalytic triad. Interestingly, the relative impact of each mutation on the activity of Dat was quite different. Compared with the value for Dat, mutation of Ser^{186} and Ser^{182} caused an apparent decrease in the k_{cat}/K_m value for tryptamine, by 82 % and 94 % respectively. E47A had a k_{cat}/K_m value for tryptamine that is only 0.23 % of that found for Dat, and is the only mutation of a Dat active-site residue identified to date that has such a dramatic effect on enzymatic activity. Furthermore, E47A reduced the k_{cat} value 168-fold and increased the tryptamine K_m value 2.5-fold. A smaller decrease in the k_{cat} value was observed for S186A (4.5-fold reduction), whereas S182A exhibited a greater decrease in the k_{cat} value (12-fold reduction). It is clear that these three mutants showed increases in the K_m for tryptamine and decreases in the K_m for AcCoA in the reaction. Notably, the drastically reduced activity for E47A also suggests that Glu^{47} initiates catalysis (see below). Of note, the Y185F and Y185A variants show no significant effect on Dat enzyme activity. Structural analysis (Supplementary Figure S6) combined with kinetic studies exclude the possibility of Tyr^{185} involved in catalysis.

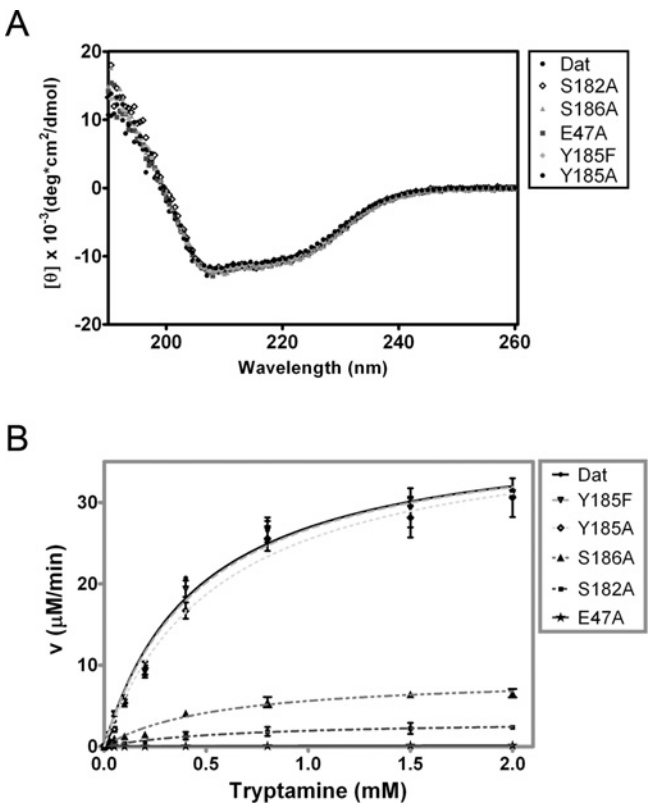

Figure 5 CD spectra and enzymatic activities of Dat and five variants each mutated at a catalytic-site position

(**A**) The far-UV CD spectra of Dat, E47A, S182A, Y185F/Y185A and S186A. (**B**) Plots of velocity as a function of substrate concentration used to determine the enzyme activities of Dat and its mutants. The enzyme kinetic data were fitted to the Michaelis–Menten equation in non-linear regression using GraphPad Prism. The order of Dat mutants in the box reflects the enzyme activity from high activity to low activity.

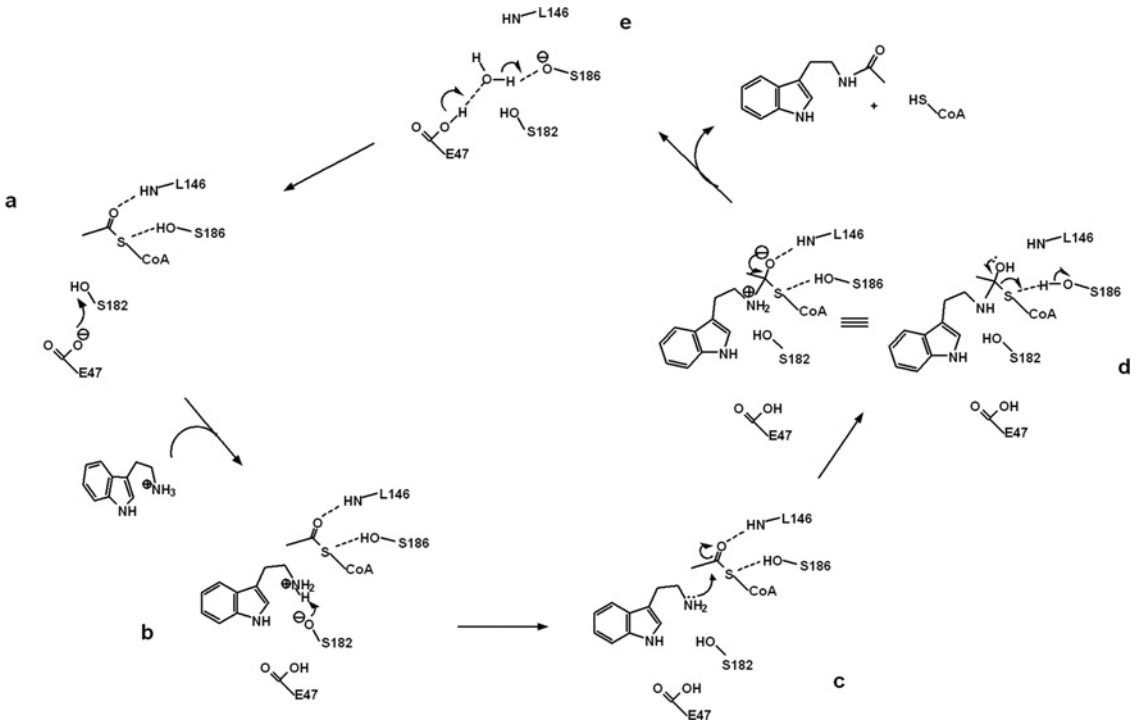

Figure 6 Proposed catalytic mechanism for Dat

The residues in Dat, tryptamine and AcCoA that are directly involved in catalysis are included in the Figure and are labelled. (**a**) The proposed mechanism involves Glu[47], acting as a base catalyst, to initiate the reaction by abstracting a proton from the hydroxy group of Ser[182]. (**b** and **c**) The hydroxylate of Ser[182] can then act as a general base catalyst for deprotonation of the substrate amino, thereby forming a tetrahedral intermediate. (**d**) The hydroxy group of Ser[186], which is separated by 3.8 Å from the AcCoA sulfur, could then serve as a general acid catalyst to protonate the thiolate anion of the leaving CoA. (**e**) The proton is transferred from the carboxylate group of Glu[47] to the alkoxide group on Ser[186] through a water molecule (see Supplementary Figure S10 at http://www.BiochemJ.org/bj/446/bj4460395add.htm). The backbone amide of Leu[146] can form a hydrogen bond with the carbonyl oxygen of the acetyl group to impose the appropriate stereochemistry on the tetrahedral intermediate.

The pH-dependence of steady-state kinetic measurement for wt Dat, S182A, S186A and E47A

Investigations of the pH rate profile can sometimes provide information about the properties of ionizable groups that participate in enzymatic catalysis. Analyses of $\log(k_{cat}/K_m)$ compared with pH should reflect pK_a values of ionizable groups present in the free enzyme and/or substrate [42]. The pH-dependence of $\log(k_{cat}/K_m)$ of the wild-type Dat with tryptamine and AcCoA, an indicator of the state of the free enzyme or the substrate, revealed two ionizable groups with pK_a values of 6.9 ± 0.1 (pK_1) and 9.0 ± 0.2 (pK_2) (Supplementary Figure 8A at http://www.BiochemJ.org/bj/446/bj4460395add.htm). The $\log(k_{cat}/K_m)$ compared with pH profiles of S182A and S186A revealed a distinct change in the higher pH values (Supplementary Figures 8B and 8C). Although the acidic limb with a pK_a of ~ 6.9 was maintained, the basic limb of the wild-type profile was absent, suggesting that the higher pK_a value (pK_2) reflects the ionization of the hydroxy group of Ser[182]/Ser[186]. Therefore we could reasonably refer the ionizable groups to Ser[182] and Ser[186], which could act as an acid/base and nucleophile for catalysis in the deprotonated form. Such a pK_a would be close to that of a previously proposed serine residue in catalysis of fatty acid amide hydrolase [43,44]. Even though Ser[182]/Ser[186] would have a pK_a value of approximately 9.0, which is 4.0 units lower than a free serine residue in solution ($pK_a = 13$), it is well known that protein active sites can change the pK_a values of ionizable groups on catalytic residues compared with their values in free solution [45]. The pK_a can be altered immensely by the microenvironment, including electrostatic effects, hydrogen bonding or a hydrophobic surrounding provided by an enzyme, to the extent that residues which are basic in solution may act as proton donors, and vice versa. The high resolution of the Dat complex structure has allowed us to define a hydrogen-bonded network that includes water-mediated hydrogen bonds between Ser[182] and Ser[186], which would probably be one of the factors required to depress the pK_a value significantly (W5 in Figure 1B). There are also many examples of catalytic residues that have significantly lower pK_a values as compared with the native pK_a of the amino acid [45–47]. Moreover, the pH-dependence of $\log(k_{cat})$ for the E47A mutant showed a linear dependence on $[OH^-]$ with a slope of 0.3 (Supplementary Figure 8D). The E47A mutation resulted in a pH-dependence profile of $\log(k_{cat}/K_m)$, which abolished the acid limb seen with the wild-type enzyme. This suggests that the pK_1 value reflects the ionization of the carboxylate group of Glu[47]. The greatly reduced enzyme activity of the E47A mutant coupled with its altered pH-rate profile was consistent with a role for Glu[47] to act as a base in the catalysis reaction. There are also several examples of glutamate residues that play a catalytic role as a general base [48,49].

Proposed catalytic mechanism for Dat

The structural and kinetic activity data that are available for GNATs support a mechanism that proceeds by direct nucleophilic attack on the AcCoA acetyl by an acceptor substrate [28,38]. For Dat, the crystallographic evidence in conjunction with the altered pH-rate profile support the simplest acetyl-transfer mechanism involving one of the two serine residues acting as a general

base during deprotonation of the arylalkylamine amino group and the other acting as a general acid during protonation of the CoA thiolate-leaving group. Moreover, the carboxylate anion of Glu^{47} in close proximity with Ser^{182} makes sense for its role as a base catalyst, required for enhancing the basicity of Ser^{182}-O to abstract a proton from the amino group of the substrate arylalkylamine ($pK_a \sim 10$) (Figure 4B), because deprotonation of the arylalkylamine amino group must occur before the acetyl group can be transferred. This suggestion is plausible, because Ser^{182} is positioned proximally to the amino group of the arylalkylamine substrate**,** which is above the AcCoA-binding site. Protonation of the CoA thiolate is probably accomplished by Ser^{186}, which is located beneath the AcCoA-binding site and may therefore act as a general acid catalyst. Its hydroxy group makes a direct hydrogen bond to the sulfur atom of AcCoA. Notably, given the effect of the β-bulge on the position of Leu^{146}, its backbone amide nitrogen is within hydrogen-bonding distance of the acetyl carbonyl oxygen. Such a hydrogen bond would help impose the appropriate stereochemistry on the tetrahedral intermediate by co-ordinating the carbonyl oxygen of the acetyl group. We therefore propose a catalytic mechanism for Dat that involves the following steps: (i) the Glu^{47} carboxylate anion accepts a proton from the Ser^{182} hydroxy group; (ii) the hydroxylate of Ser^{182} accepts a proton from the nucleophilic amino group of the substrate; (iii) the neutral amino group then attacks the carbon atom of the acetyl carbonyl to form a tetrahedral intermediate; and (iv) the hydroxy group of Ser^{186} protonates the leaving thiolate anion of CoA as the tetrahedral intermediate collapses into products (Figure 6). Most members of GNAT superfamily possess a tyrosine residue that is well positioned to act as a general acid, which is in sharp contrast with the serine residue of Dat. Besides, the catalytic residues identified as involved in Dat and SNAT catalysis differ in kind and in position (Supplementary Figure S9 at http://www.BiochemJ.org/bj/446/bj4460395add.htm). Dat has a novel Glu-Ser-Ser catalytic triad, which is completely different from that found in SNAT. This novel catalytic triad, to our knowledge, has not been observed in any other known enzymes. Additionally, our proposed mechanism for formation of the tetrahedral intermediate differs from the one proposed for SNAT [16]. The Dat Ser^{182} directly deprotonates the substrate amino instead of deprotonating it through a chain of well-ordered water molecules, i.e. the so-called proton wire, that connects the acceptor substrate amino group to a base, which may be involved in removal of the proton [16]. For SNAT, His^{120} was proposed to act as such a base. In conclusion, homologous AANATs found in mammals and insects may have different catalytic mechanisms.

Conclusions

The present study represents the first structure of an insect AANAT. We have also confirmed that Dat is a member of the GNAT superfamily and suggested that Dat from *D. melanogaster* utilizes a Glu-Ser-Ser catalytic triad mechanism. For members of a protein superfamily to be able to catalyse the same chemical reaction but target different substrates requires a common catalytic domain, but also structural diversity within the domain. It is interesting to research how specificity for many different substrate types can be generated given this common central domain. Such studies should increase our knowledge of Dat in relation to other GNAT superfamily members and should allow for the development of specific insecticides and potential lead compounds that might treat mood and sleep disorders induced by abnormal melatonin production.

AUTHOR CONTRIBUTION

Kuo-Chang Cheng and Ping-Chiang Lyu designed the research, analysed the data and wrote the paper. Kuo-Chang Cheng performed all of the experiments. Jhen-Ni Liao purified the recombinant proteins and carried out enzyme kinetics and ITC experiments. Kuo-Chang Cheng determined the structure. Ping-Chiang Lyu commented on the paper.

ACKNOWLEDGEMENTS

Portions of this research were performed at the National Synchrotron Radiation Research Center, a national user facility supported by the National Science Council of Taiwan, ROC. The Synchrotron Radiation Protein Crystallography Facility is supported by the National Core Facility Program for Biotechnology. We thank Cheng-Hung Chiang (beam manager) and the facility staff for excellent technical assistance.

FUNDING

This work was supported by the National Science Council, Taiwan [grant numbers NSC 101-2311-B-007-003, NSC 101-2319-B-400-001 and NSC 100-2627-B-007-005].

REFERENCES

1 Hardeland, R., Pandi-Perumal, S. R. and Cardinali, D. P. (2006) Melatonin. Int. J. Biochem. Cell Biol. **38**, 313–316
2 Hardeland, R. and Poeggeler, B. (2003) Non-vertebrate melatonin. J. Pineal Res. **34**, 233–241
3 Hardeland, R. and Fuhrberg, B. (1996) Ubiquitous melatonin. Presence and effects in unicells, plants and animals. Trends Comp. Biochem. Physiol. **2**, 25–45
4 Finocchiaro, L., Callebert, J., Launay, J. M. and Jallon, J. M. (1988) Melatonin biosynthesis in *Drosophila*: its nature and its effects. J. Neurochem. **50**, 382–387
5 Vivien-Roels, B. and Pévet, P. (1993) Melatonin: presence and formation in invertebrates. Experientia **49**, 642–647
6 Reiter, R. J. (1993) The melatonin rhythm: both a clock and a calendar. Experientia **49**, 654–664
7 Wright, T. R. (1987) The genetics of biogenic amine metabolism, sclerotization, and melanization in *Drosophila melanogaster*. Adv. Genet. **24**, 127–222
8 Dewhurst, S. A., Croker, S. G., Ikeda, K. and McCaman, R. E. (1972) Metabolism of biogenic amines in *Drosophila* nervous tissue. Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol. **43**, 975–981
9 Evans, P. and Michaelfox, P. (1975) Enzymatic N-acetylation of indolealkylamines by brain homogenates of the honeybee, *Apis mellifera*. J. Insect Physiol. **21**, 343–353
10 Nishimura, K. (1975) Catabolism of tryptamine by cockroach head enzyme preparation. Pestic. Biochem. Physiol. **5**, 557–565
11 Sekeris, C. E. and Karlson, P. (1966) Biosynthesis of catecholamines in insects. Pharmacol. Rev. **18**, 89–94
12 Amherd, R., Hintermann, E., Walz, D., Affolter, M. and Meyer, U. A. (2000) Purification, cloning, and characterization of a second arylalkylamine N-acetyltransferase from *Drosophila melanogaster*. DNA Cell Biol. **19**, 697–705
13 Hintermann, E., Grieder, N. C., Amherd, R., Brodbeck, D. and Meyer, U. A. (1996) Cloning of an arylalkylamine N-acetyltransferase (aaNAT1) from *Drosophila melanogaster* expressed in the nervous system and the gut. Proc. Natl. Acad. Sci. U.S.A. **93**, 12315–12320
14 Neuwald, A. F. and Landsman, D. (1997) GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. Trends Biochem. Sci. **22**, 154–155
15 Hickman, A. B., Klein, D. C. and Dyda, F. (1999) Melatonin biosynthesis: the structure of serotonin N-acetyltransferase at 2.5 Å resolution suggests a catalytic mechanism. Mol. Cell **3**, 23–32
16 Hickman, A. B., Namboodiri, M. A., Klein, D. C. and Dyda, F. (1999) The structural basis of ordered substrate binding by serotonin N-acetyltransferase: enzyme complex at 1.8 Å resolution with a bisubstrate analog. Cell **97**, 361–369
17 Scheibner, K. A., De Angelis, J., Burley, S. K. and Cole, P. A. (2002) Investigation of the roles of catalytic residues in serotonin N-acetyltransferase. J. Biol. Chem. **277**, 18118–18126
18 Hendrickson, W. A. (1991) Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science **254**, 51–58
19 Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**, 487–491
20 Doublie, S. (1997) Preparation of selenomethionyl proteins for phase determination. Methods Enzymol. **276**, 523–530

21 Otwinowski, Z. and Minor, W. (1997) [20] Processing of X-ray diffraction data collected in oscillation mode. **276**, 307–326
22 Sheldrick, G. M. (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Crystallogr., Sect. D: Biol. Crystallogr. **66**, 479–485
23 Hattne, J. and Lamzin, V. S. (2008) Pattern-recognition-based detection of planar objects in three-dimensional electron-density maps. Acta Crystallogr., Sect. D: Biol. Crystallogr. **64**, 834–842
24 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr., Sect. D: Biol. Crystallogr. **60**, 2126–2132
25 Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr., Sect. D: Biol. Crystallogr. **53**, 240–255
26 Laskowski, R. A., Moss, D. S. and Thornton, J. M. (1993) Main-chain bond lengths and bond angles in protein structures. J. Mol. Biol. **231**, 1049–1067
27 Riddles, P. W., Blakeley, R. L. and Zerner, B. (1983) Reassessment of Ellman's reagent. Methods Enzymol. **91**, 49–60
28 Khalil, E. M., De Angelis, J. and Cole, P. A. (1998) Indoleamine analogs as probes of the substrate selectivity and catalytic mechanism of serotonin N-acetyltransferase. J. Biol. Chem. **273**, 30321–30327
29 Jones, G., Willett, P., Glen, R. C., Leach, A. R. and Taylor, R. (1997) Development and validation of a genetic algorithm for flexible docking. J. Mol. Biol. **267**, 727–748
30 Wallace, A. C., Laskowski, R. A. and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. **8**, 127–134
31 Rost, B., Yachdav, G. and Liu, J. (2004) The PredictProtein server. Nucleic Acids Res. **32**, W321–W326
32 Lo, W. C., Lee, C. Y., Lee, C. C. and Lyu, P. C. (2009) iSARST: an integrated SARST web server for rapid protein structural similarity searches. Nucleic Acids Res. **37**, W545–W551
33 He, H., Ding, Y., Bartlam, M., Sun, F., Le, Y., Qin, X., Tang, H., Zhang, R., Joachimiak, A. and Liu, J. (2003) Crystal structure of tabtoxin resistance protein complexed with acetyl coenzyme A reveals the mechanism for β-lactam acetylation. J. Mol. Biol. **325**, 1019–1030
34 Wybenga-Groot, L. E., Draker, K., Wright, G. D. and Berghuis, A. M. (1999) Crystal structure of an aminoglycoside 6′-N-acetyltransferase: defining the GCN5-related N-acetyltransferase superfamily fold. Structure **7**, 497–507
35 Angus-Hill, M. L., Dutnall, R. N., Tafrov, S. T., Sternglanz, R. and Ramakrishnan, V. (1999) Crystal structure of the histone acetyltransferase Hpa2: a tetrameric member of the Gcn5-related N-acetyltransferase superfamily1. J. Mol. Biol. **294**, 1311–1325
36 Peneff, C. (2001) The crystal structures of Apo and complexed *Saccharomyces cerevisiae* GNA1 shed light on the catalytic mechanism of an amino-sugar N-acetyltransferase. J. Biol. Chem. **276**, 16328–16334
37 Vetting, M. W., S. de Carvalho, L. P., Yu, M., Hegde, S. S., Magnet, S., Roderick, S. L. and Blanchard, J. S. (2005) Structure and functions of the GNAT superfamily of acetyltransferases. Arch. Biochem. Biophys. **433**, 212–226
38 Dyda, F., Klein, D. C. and Hickman, A. B. (2000) GCN5-related N-acetyltransferases: a structural overview. Annu. Rev. Biophys. Biomol. Struct. **29**, 81–103
39 Richardson, J. S. (1981) The anatomy and taxonomy of protein structure. Adv. Protein Chem. **34**, 167–339
40 Mio, T., Yamada-Okabe, T., Arisawa, M. and Yamada-Okabe, H. (1999) *Saccharomyces cerevisiae* GNA1, an essential gene encoding a novel acetyltransferase involved in UDP-N-acetylglucosamine synthesis. J. Biol. Chem. **274**, 424–429
41 Obsil, T., Ghirlando, R., Klein, D. C., Ganguly, S. and Dyda, F. (2001) Crystal structure of the 14-3-3ζ:serotonin N-acetyltransferase complex. a role for scaffolding in enzyme regulation. Cell **105**, 257–267
42 Dixon, M. and Webb, E. C. (1964) Enzymes. Academic Press, New York
43 Patricelli, M. P. and Cravatt, B. F. (2000) Clarifying the catalytic roles of conserved residues in the amidase signature family. J. Biol. Chem. **275**, 19177–19184
44 McKinney, M. K. and Cravatt, B. F. (2003) Evidence for distinct roles in catalysis for residues of the serine-serine-lysine catalytic triad of fatty acid amide hydrolase. J. Biol. Chem. **278**, 37393–37399
45 Harris, T. K. and Turner, G. J. (2002) Structural basis of perturbed pKa values of catalytic groups in enzyme active sites. IUBMB Life **53**, 85–98
46 Dao-pin, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W. and Matthews, B. W. (1991) Structural and thermodynamic consequences of burying a charged residue within the hydrophobic core of T4 lysozyme. Biochemistry **30**, 11521–11529
47 Planas, A. and Kirsch, J. F. (1991) Reengineering the catalytic lysine of aspartate aminotransferase by chemical elaboration of a genetically introduced cysteine. Biochemistry **30**, 8268–8276
48 Kondo, M. Y., Okamoto, D. N., Santos, J. A., Juliano, M. A., Oda, K., Pillai, B., James, M. N., Juliano, L. and Gouvea, I. E. (2010) Studies on the catalytic mechanism of a glutamic peptidase. J. Biol. Chem. **285**, 21437–21445
49 Zito, C. R., Antony, E., Hunt, J. F., Oliver, D. B. and Hingorani, M. M. (2005) Role of a conserved glutamate residue in the *Escherichia coli* SecA ATPase mechanism. J. Biol. Chem. **280**, 14611–14619
50 Jones, G., Willett, P. and Glen, R. C. (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J. Mol. Biol. **245**, 43–53

Received 26 March 2012/11 June 2012; accepted 21 June 2012
Published as BJ Immediate Publication 21 June 2012, doi:10.1042/BJ20120520

Biochem. J. (2012) **446**, 395–404 (Printed in Great Britain) doi:10.1042/BJ20120520

SUPPLEMENTARY ONLINE DATA

Crystal structure of the dopamine *N*-acetyltransferase–acetyl CoA complex provides insights into the catalytic mechanism

Kuo-Chang CHENG*, Jhen-Ni LIAO* and Ping-Chiang LYU*†‡[1]

*Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu, 30013, Taiwan, †Department of Medical Sciences, National Tsing Hua University, Hsinchu, 30013, Taiwan, and ‡Graduate Institute of Molecular Systems Biomedicine, China Medical University, Taichung, 40402, Taiwan

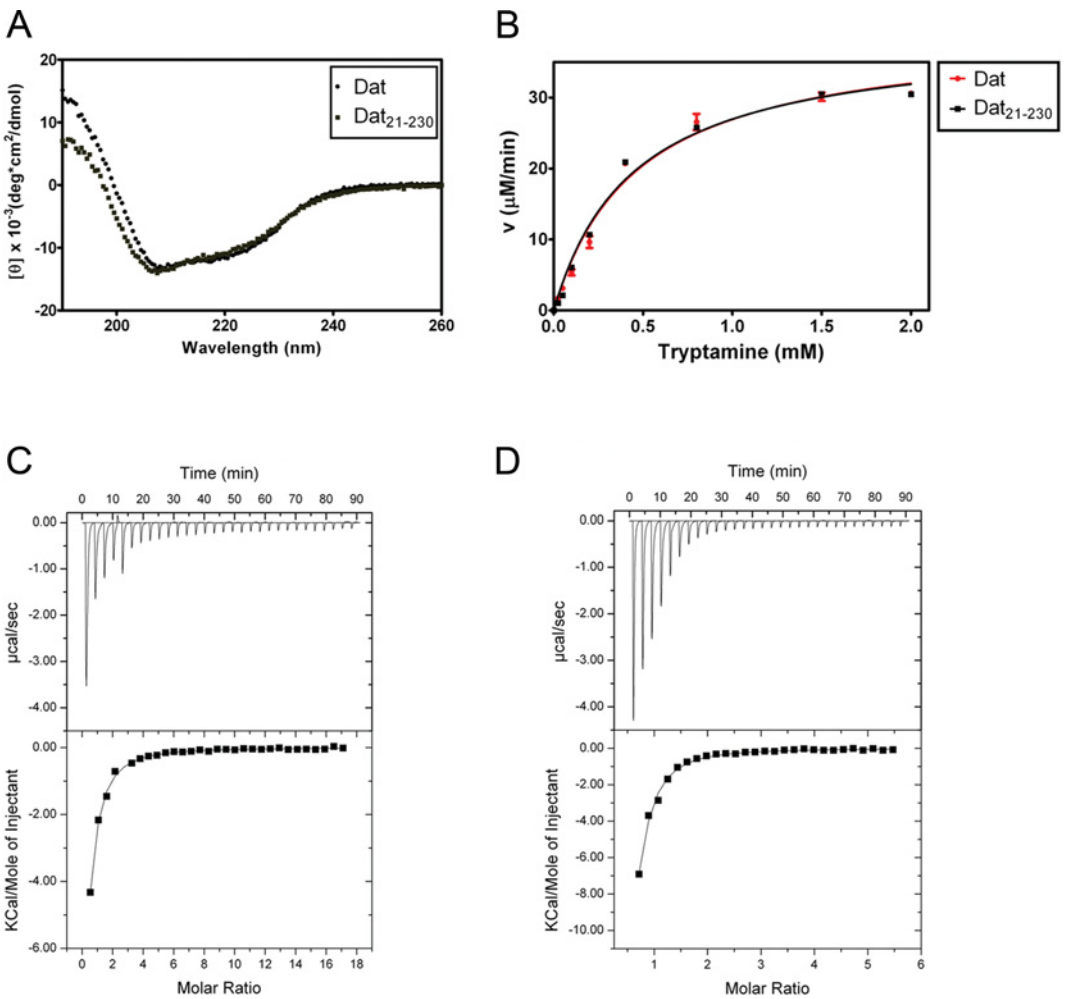

Figure S1 Comparisons of Dat and $Dat_{21–230}$

(**A**) CD spectra and (**B**) enzymatic kinetic profiles of Dat and $Dat_{21–230}$. Release of the CoA was used to measure the enzymatic activity. Titrations of dopamine (**C**) and AcCoA (**D**) against $Dat_{21–230}$ were performed. Results are shown for heat change (upper panels) and peak integration (lower panels). The continuous line represents the best fits to a single-site binding model. The binding constants (K_d) for dopamine and AcCoA were 12.6 μM and 2.82 μM respectively.

The atomic co-ordinates and structural factors have been deposited in the PDB under accession code 3TE4 for the $Dat_{21–230}$–AcCoA complex structure.

[1] To whom correspondence should be addressed (email pclyu@mx.nthu.edu.tw).

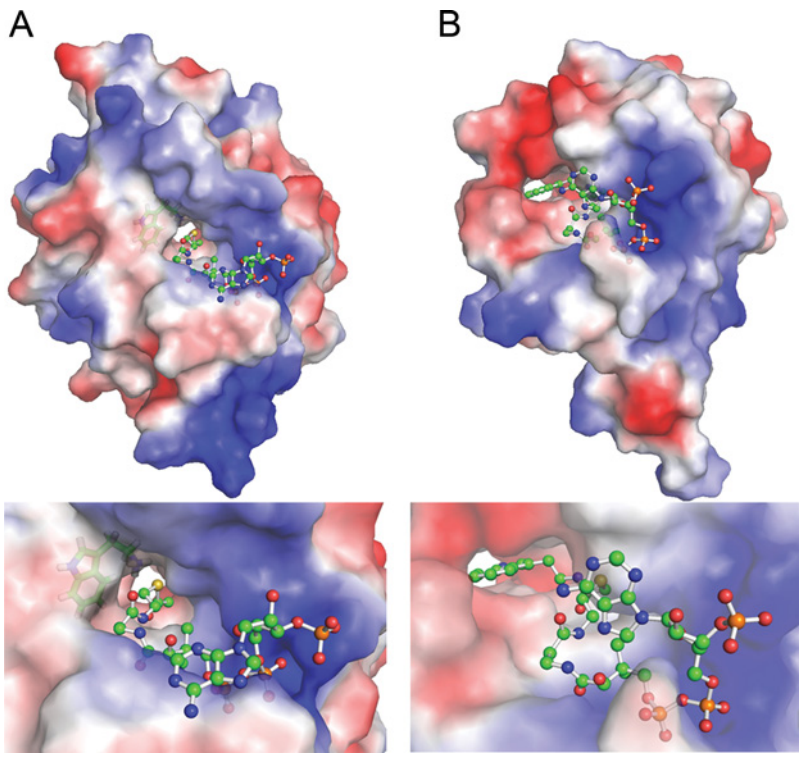

Figure S2 Substrate- and AcCoA-binding pockets in $Dat_{21–230}$ and in SNAT

(**A** and **B**) Surface electrostatic potential representations of the active sites in $Dat_{21–230}$ and SNAT respectively. Tryptamine and AcCoA in their respective binding pockets in $Dat_{21–230}$ are shown as stick and ball-and-stick models respectively. In SNAT, the bisubstrate analogue, CoA-S-acetyltryptamine, is shown as ball-and-stick model. The atoms of AcCoA are coloured green (carbon), blue (nitrogen), yellow (sulfur), red (oxygen) and orange (phosphorous). The colour code for the electrostatic potential is blue for positive charges and red for negative charges. The bottom panels are close-ups of the active sites.

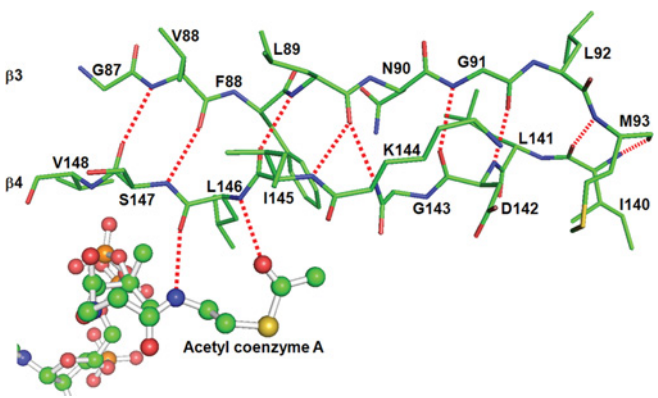

Figure S3 The Dat_{21-230} β-bulge

The residues of β3 and β4 are shown as stick models with the β-bulge formed by Lys^{144} and Ile^{145}, which positions their side chains on the same face of the β-sheet. AcCoA is shown as a ball-and-stick model. The atoms of AcCoA are coloured green (carbon), blue (nitrogen), yellow (sulfur), red (oxygen) and orange (phosphorous). The broken lines indicate hydrogen bonds.

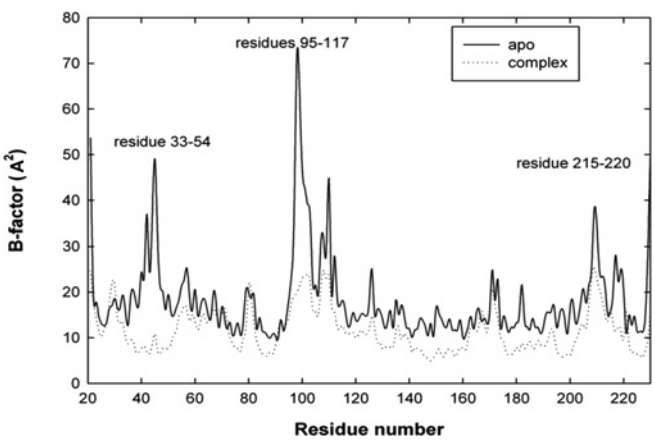

Figure S4 Temperature factor plots of Dat crystal structures in apo form and complex form

Plot of main-chain temperature factors for the apo form (continuous line) and complex form (broken line) of Dat crystal structures. In the complex structure, three regions, residues 35–54, 95–117 and 215–220, show significantly lower temperature factors compared with that of the apo form.

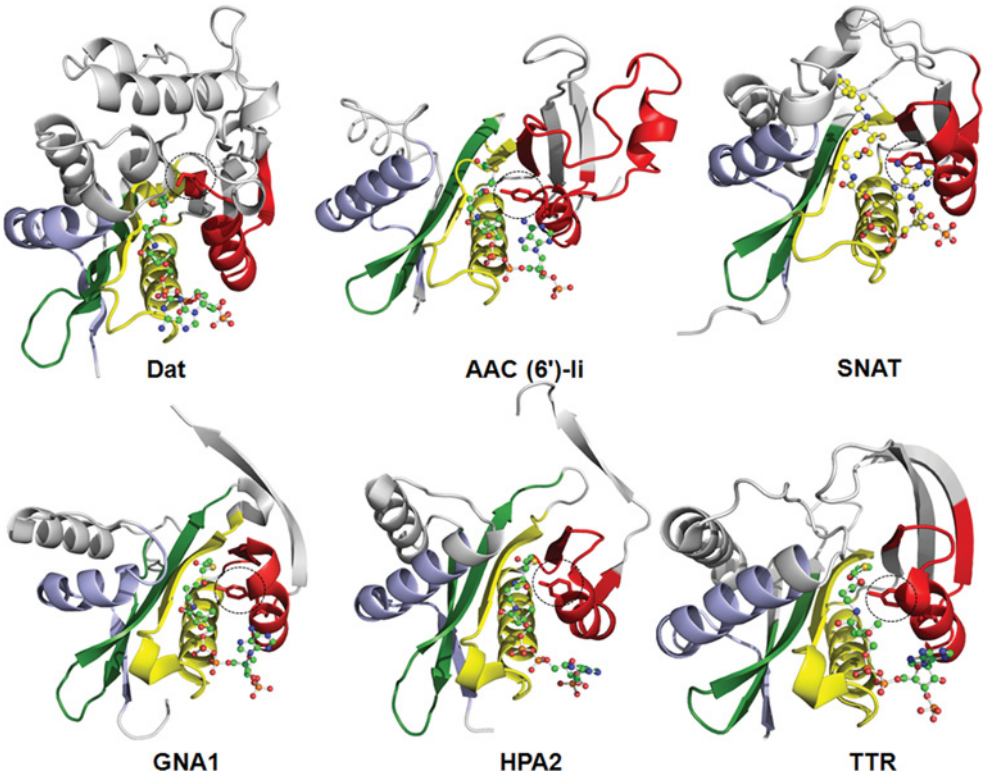

Figure S5 Structures of $Dat_{21–230}$ and other GCN5 *N*-acetyltransferase domains

The conserved GNAT motifs C, D, A and B are coloured light blue, green, yellow and red respectively. AcCoA is shown as a ball-and-stick model. The conserved catalytic tyrosine residues (serine in Dat) are shown as red stick models (highlighted in the broken circle). AAC (6′)-Ii, aminoglycoside 6′-*N*-acetyltransferase type Ii, *Enterococcus faecium*; GNA1, glucosamine-6-phosphate *N*-acetyltransferase 1, *Saccharomyces cerevisiae*; HPA2, histone acetyltransferase, *Saccharomyces cerevisiae*; SNAT, serotonin *N*-acetyltransferase, *Ovis aries*; TTR, tabtoxin-resistance protein, *Pseudomonas syringae*.

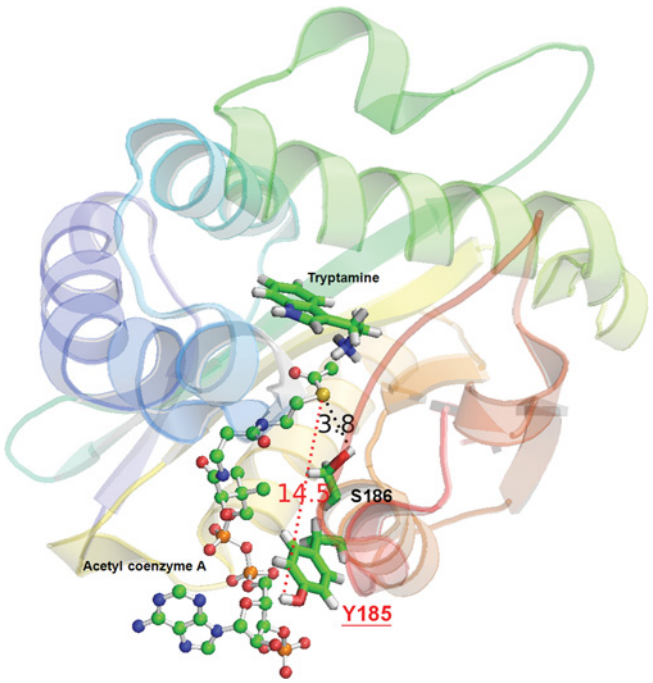

Figure S6 Structure of $Dat_{21–230}$–AcCoA complex with a docked tryptamine: spatial relationships between potential residues and AcCoA at the active site

The Dat sequence contains a Tyr^{185} residue just before Ser^{186}. The residues Tyr^{185} (red) and Ser^{186} (black) are represented as stick models. The broken lines identify distance between the AcCoA sulfur atom and target residues. The numbers represent distance separations between various atoms in angstroms. Tryptamine is shown as a stick model and AcCoA is shown as a ball-and-stick model.

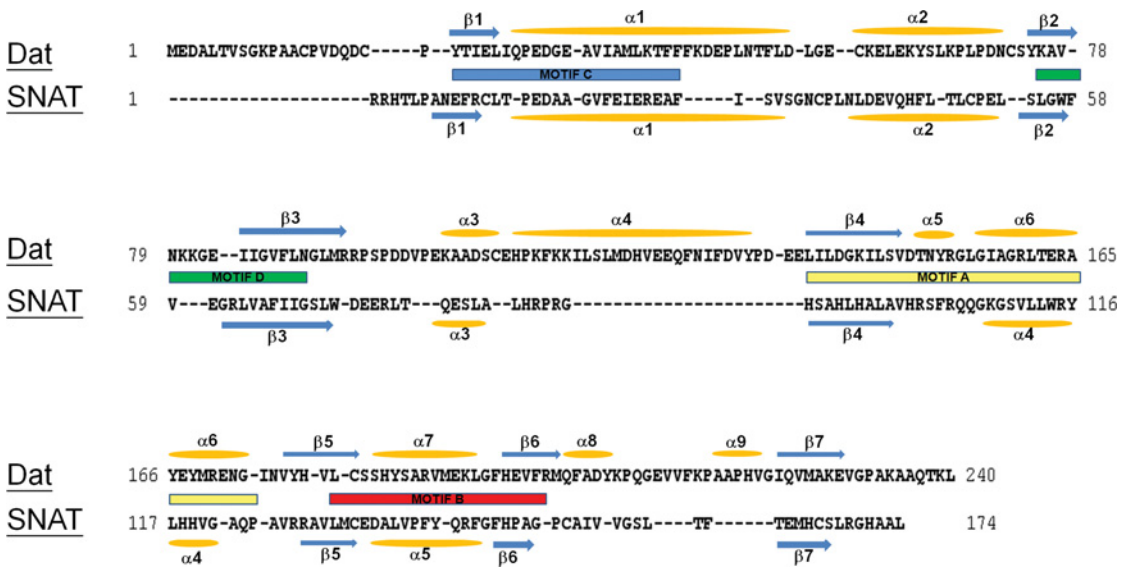

Figure S7 Structure-based sequence alignment of Dat and *O. aries* SNAT

To structurally align the sequences, first their three-dimensional structures were manually aligned with gaps introduced as needed. Then, the sequences of the common secondary structural elements were manually aligned with the positions of those in SNAT adjusted so that they were aligned with the corresponding secondary structures in $Dat_{21–230}$. The secondary structure elements in Dat and SNAT are indicated above and below the sequences respectively. The conserved motifs C, D, A and B are shown as boxes coloured cyan, green, yellow and red respectively.

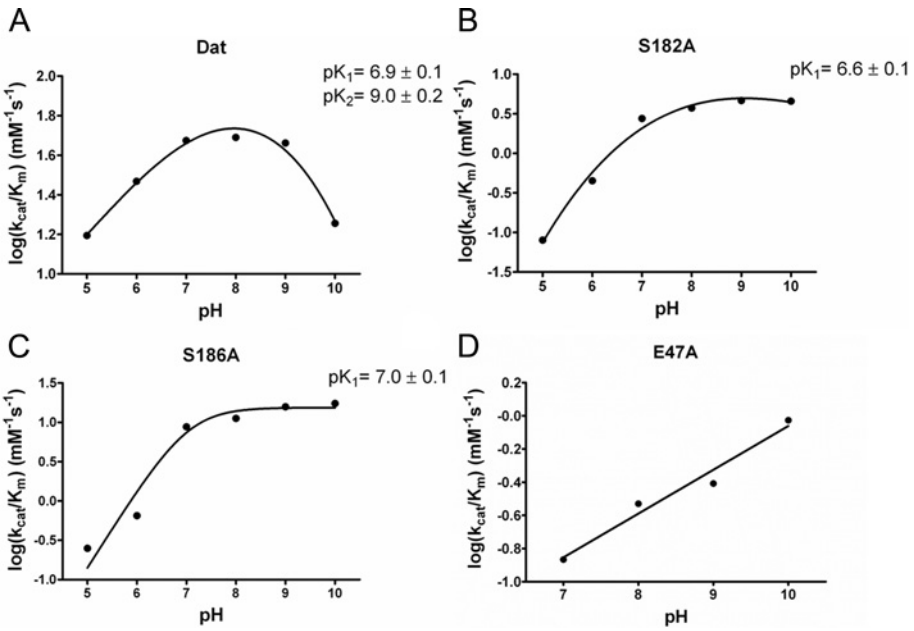

Figure S8 The pH profiles of log(k_{cat}/K_m) for wild-type Dat, S182A, S186A and E47A

The pH-dependence of log(k_{cat}/K_m) observed for Dat (**A**) are shown with curve fits obtained by non-linear least-squares regression. The pK_a values are given on the graphs. pK_1 (6.9±0.1) refers to the acidic limb of the profile, and pK_2 (9.0±0.2) represents the basic limb. pH values compared with log(k_{cat}/K_m) profiles of S182A (**B**) and S186A(**C**) indicate a ionizable residue which is active only in the deprotonated form involved in catalysis with a pK_a of approximately 6.9. The pH profile for the E47A mutant (**D**) was fitted to a line with a slope of 0.3.

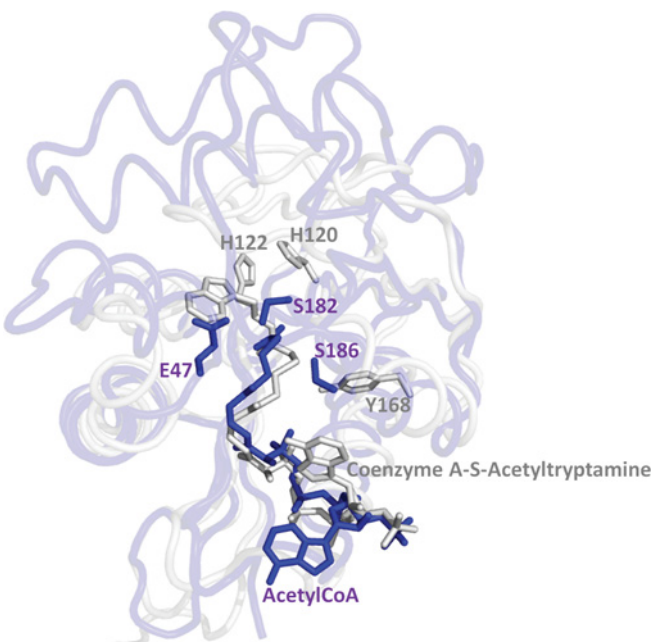

Figure S9 Comparison of the catalytic triads of AANATs from *D. melanogaster* and *O. aries*

The structural superimposition of Dat (blue) and SNAT (grey). The ligands and catalytic residues for Dat and SNAT are labelled and shown in stick representation in blue and grey respectively. The two structures share a very similar structural fold despite low sequence identity (12 %), with an RMSD value of 1.89 Å.

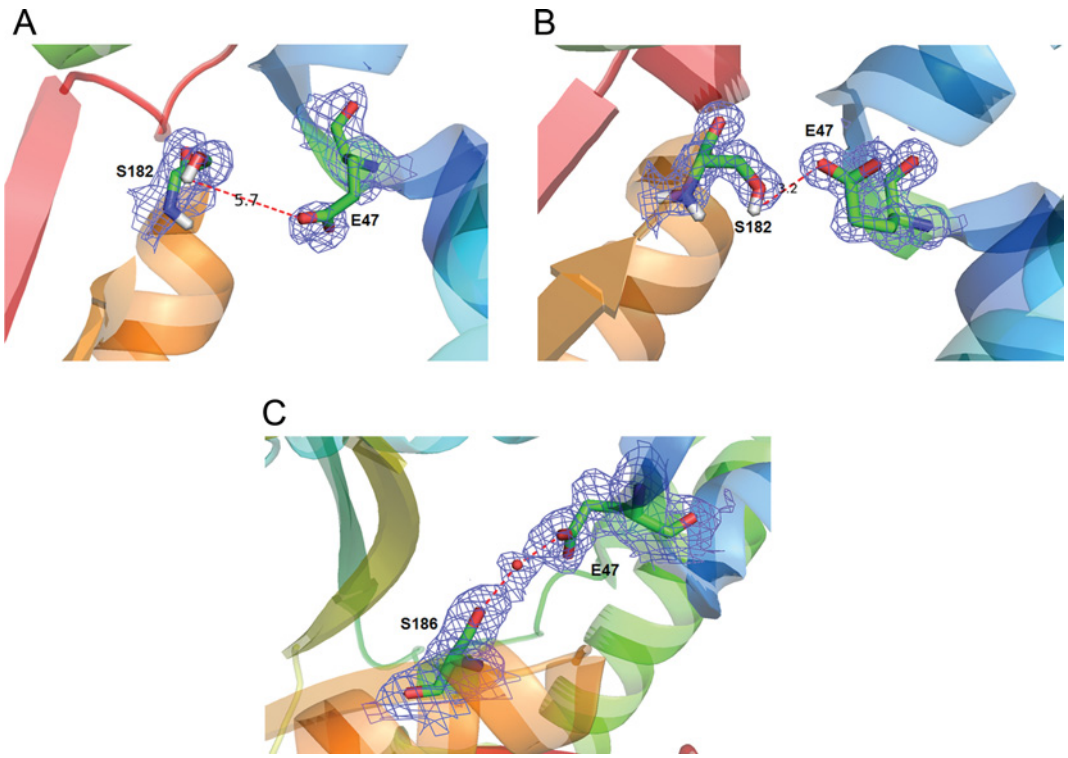

Figure S10 Comparison of the catalytic triad residues of Dat in apo form and complex form

The catalytic triad residues Glu^{47} and Ser^{182} are represented as stick models in the apo form (**A**) and complex form (**B**) of the Dat structure. The broken lines identify the distance between the catalytic residues. The numbers represent distance separations between various atoms in angstroms. (**C**) The catalytic triad residues Glu^{47} and Ser^{186} form a hydrogen bond network with each other by a proton-conducting water molecule (shown as a red ball). The density, shown as a blue mesh, is a $2F_o - F_c$ map, contoured at 1.1 σ. Note that, by comparing the Dat structure between apo form and complex form, two crystallographic facts could implicate the catalytic role of Glu^{47}: (i) the distance between the Glu^{47} carboxylate anion and Ser^{182} hydroxy group decreased after formation of the Dat–AcCoA complex structure, which enabled Glu^{47} to abstract a proton from the hydroxy group of Ser^{182}. Moreover, the carboxylate of Glu^{47} forms a hydrogen bond to a proton-conducting water molecule. This water molecule is co-ordinated by hydrogen bonding and is 2.8 Å away from the hydroxy oxygen of Ser^{186} in an appropriate position to transfer a proton (crystallographic and refinement statistics of the apo form structure are not shown. K.-C. Cheng and P.-C. Lyu, unpublished work).

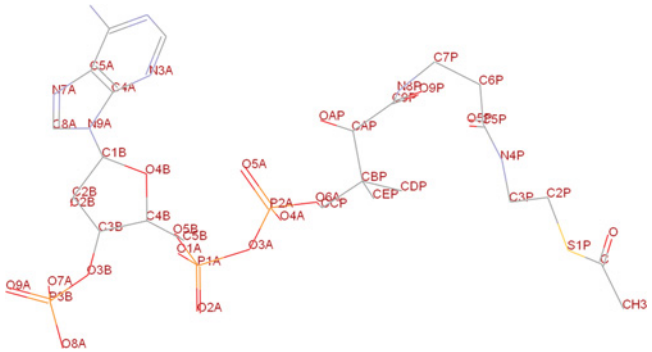

Figure S11 AcCoA with the atoms labelled as in Table S1

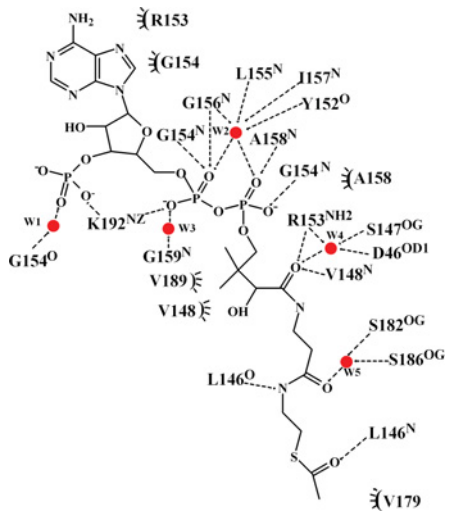

Figure S12 Schematic diagram of the AcCoA–protein contacts

Table S1 Detailed protein–cofactor contacts

The numbers represent distance separations between cofactor atoms and protein residue atoms in angstroms. See Figures S11 and S12 for more details.

(i) AcCoA salt bridge

Atom	Salt bridge
O8A	Lys^{192} NZ 2.7
O2A	Lys^{192} NZ 2.7

(ii) AcCoA hydrogen bonds

Atom	Hydrogen bond	Water-mediated hydrogen bond
O7A	W1 2.7	Gly^{154} O 2.9
O1A	Gly^{154} N 3.3	–
	Gly^{156} N 2.8	–
	W2 3.4	Gly^{156} N 3.1
	W2 3.4	Leu^{155} N 2.8
	W2 3.4	Ile^{157} N 2.9
	W2 3.4	Tyr^{152} O 2.7
O2A	W3 2.7	Gly^{159} N 2.8
O4A	Ala^{158} N 2.8	–
	W2 2.7	Gly^{156} N 3.1
	W2 2.7	Leu^{155} N 2.8
	W2 2.7	Ile^{157} N 2.9
	W2 2.7	Tyr^{152} O 2.7
O5A	Gly^{154} N 2.9	–
O9P	Arg^{153} NH2 3.3	–
	Val^{148} N 3.1	–
	W4 2.9	Arg^{153} NH2 3.2
	W4 2.9	Ser^{147} OG 2.7
	W4 2.9	Asp^{46} OD1 2.8
O5P	W5 2.9	Ser^{182} OG 2.8
	W5 2.9	Ser^{186} OG 2.7
N4P	Leu^{146} O 2.9	
O	Leu^{146} N 3.0	

(iii) AcCoA van der Waals interactions

Atom	Van der Waals interaction
C2A	Arg^{153} CB 3.5
	Arg^{153} CG 4.3
C5A	Gly^{154} CA 3.8
	Arg^{153} C 4.0
C6A	Arg^{153} CB 3.8
	Arg^{153} CA 4.4
C8A	Gly^{154} CA 3.7
C5B	Val^{189} CG2 3.7
CCP	Val^{189} CG2 4.2
	Ala^{158} CB 3.8
CDP	Leu^{146} CG 3.6
	Val^{148} CG2 3.8
CEP	Val^{189} CG2 4.4
C9P	Arg^{153} CZ 3.6
CH3	Val^{179} CG1 3.8

Received 26 March 2012/11 June 2012; accepted 21 June 2012
Published as BJ Immediate Publication 21 June 2012, doi:10.1042/BJ20120520