

Northumbria Research Link

Citation: Karagianni, Eleni-Pavlina, Kontomina, Evanthia, Lowe, Edward D., Athanasopoulos, Konstantinos, Papanikolaou, Georgia, Garefalaki, Vasiliki, Kotseli, Varvara, Zaliou, Sofia, Grimaud, Tom, Arvaniti, Konstantina, Tsatiri, Maria-Aggeliki, Fakis, Giannoulis, Glenn, Anthony E., Roversi, Pietro, Abuhammad, Areej, Ryan, Ali, Sim, Robert B., Sim, Edith and Boukouvala, Sotiria (2023) Fusarium verticillioides NAT1 (FDB2) N-malonyltransferase is structurally, functionally and phylogenetically distinct from its N-acetyltransferase (NAT) homologues. FEBS Journal, 290 (9). pp. 2412-2436. ISSN 1742-464X

Published by: Wiley-Blackwell

URL: <https://doi.org/10.1111/febs.16642> <<https://doi.org/10.1111/febs.16642>>

This version was downloaded from Northumbria Research Link:
<https://nrl.northumbria.ac.uk/id/eprint/50765/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

1 ***Fusarium verticillioides* NAT1 (FDB2) *N*-malonyltransferase is structurally and**
2 **functionally distinct from its *N*-acetyltransferase (NAT) homologues**

3

4 Eleni-Pavlina Karagianni^{§1,2}, Evanthia Kontomina^{§1,2}, Edward D. Lowe^{§3},
5 Konstantinos Athanasopoulos^{§1}, Georgia Papanikolaou¹, Vasiliki Garefalaki¹, Varvara
6 Kotseli¹, Sofia Zaliou^{1,4}, Tom Grimaud¹, Konstantina Arvaniti¹, Maria-Aggeliki
7 Tsatiri¹, Giannoulis Fakis^{1,2}, Anthony E. Glenn⁵, Pietro Roversi^{6,7}, Areej
8 Abuhammad^{2,8}, Ali Ryan^{‡2,4}, Robert B. Sim^{†2}, Edith Sim^{2,4}, Sotiria Boukouvala^{*1,2}

9

10 ¹Democritus University of Thrace, Department of Molecular Biology and Genetics,
11 Alexandroupolis, Greece

12 ²University of Oxford, Department of Pharmacology, Oxford, U.K.

13 ³University of Oxford, Department of Biochemistry, Oxford, U.K.

14 ⁴Kingston University London, Faculty of Science, Engineering and Computing,
15 Kingston-upon-Thames, U.K.

16 ⁵U.S. Department of Agriculture, Agricultural Research Service, National Poultry
17 Research Center, Toxicology & Mycotoxin Research Unit, Athens, GA, U.S.A.

18 ⁶Institute of Agricultural Biology and Biotechnology, IBBA-CNR Unit of Milan,
19 Milan, Italy

20 ⁷Leicester Institute of Structural and Chemical Biology, Department of Molecular and
21 Cell Biology, University of Leicester, U.K.

22 ⁸The University of Jordan, School of Pharmacy, Amman, Jordan

23 **§Equal contribution of authors**

24 **‡Current address:** Northumbria University, Department of Applied Sciences,
25 Newcastle-upon-Tyne, U.K. (A.R.)

26 **†Deceased**

27 ***Corresponding author:** Sotiria Boukouvala, Assoc. Prof. in Molecular Genetics,
28 Democritus University of Thrace,
29 Department of Molecular Biology and Genetics,
30 Building 10, University Campus,
31 Alexandroupolis 68100, Greece.
32 Tel./Fax.: +30-25510-30632
33 E-mail: sboukouv@mbg.duth.gr
34 ORCID: 0000-0002-3162-5375

35 **Running title:** *Fusarium* NAT1 (FDB2) *N*-malonyltransferase

36 **Abstract**

37 *Fusarium* endophytes damage cereal crops and contaminate produce with mycotoxins.
38 Those fungi overcome the main chemical defence of host via detoxification by a
39 malonyl-CoA dependent enzyme homologous to xenobiotic metabolizing arylamine
40 *N*-acetyltransferase (NAT). In *Fusarium verticillioides* (teleomorph *Gibberella*
41 *moniliformis*, GIBMO), this *N*-malonyltransferase activity is attributed to
42 (GIBMO)NAT1, and the fungus has two additional isoenzymes, (GIBMO)NAT3 (*N*-
43 acetyltransferase) and (GIBMO)NAT2 (unknown function). We present the
44 crystallographic structure of (GIBMO)NAT1, also modelling other fungal NAT
45 homologues. Monomeric (GIBMO)NAT1 is distinctive, with access to the catalytic
46 core through two "tunnel-like" entries separated by a "bridge-like" helix. In the
47 quaternary arrangement, (GIBMO)NAT1 monomers interact in pairs along an
48 extensive interface whereby one entry of each monomer is covered by the *N*-terminus
49 of the other monomer. Although monomeric (GIBMO)NAT1 apparently
50 accommodates acetyl-CoA better than malonyl-CoA, dimerization changes the active
51 site to allow malonyl-CoA to reach the catalytic triad (Cys110, His158, Asp173) via
52 the single uncovered entry, and anchor its terminal carboxyl-group via hydrogen
53 bonds to Arg109, Asn157 and Thr261. Lacking a terminal carboxyl-group, acetyl-
54 CoA cannot form such stabilizing interactions, while longer acyl-CoAs enter the
55 active site but cannot reach catalytic Cys. Other NAT isoenzymes lack such structural
56 features, with (GIBMO)NAT3 resembling bacterial NATs and (GIBMO)NAT2
57 adopting a structure intermediate between (GIBMO)NAT1 and (GIBMO)NAT3.
58 Biochemical assays confirmed differential donor substrate selectivity of
59 (GIBMO)NAT isoenzymes, with phylogenetic analysis demonstrating evolutionary
60 separation. Given the role of (GIBMO)NAT1 in enhancing *Fusarium* pathogenicity,
61 unravelling the structure and function of this enzyme may benefit research into more
62 targeted strategies for pathogen control.

63 **Keywords:** *Fusarium verticillioides*; *Gibberella moniliformis*; *N*-malonyltransferase;
64 *N*-acetyltransferase; protein crystal structure

65 **Abbreviations:** 2AP, 2-aminophenol; 3,4-DCA, 3,4-dichloroaniline; 5AS, 5-
66 aminosalicylate; BOA, 2-benzoxazinone; BX, benzoxazinoid; DIBOA, 2,4-
67 dihydroxy-1,4-benzoxazin-3-one; DSF, differential scanning fluorimetry; ESI-MS,
68 electrospray ionization mass spectrometry; MALS, multi-angle light scattering;
69 PABA, *p*-aminobenzoate.

70 **Introduction**

71 The NAT enzyme family (Pfam ID: PF00797) includes homologues across a broad
72 taxonomic range of prokaryotes and eukaryotes, except plants. Considerable
73 variability is observed in *NAT* gene sequence composition and number of sister loci
74 per genome (Boukouvala and Fakis 2005; Vagena et al. 2008; Glenn et al. 2010), and
75 functional diversity has been reported for certain microbial NAT homologues (Floss
76 and Yu 1999; Bhakta et al. 2004; Karagianni et al. 2015; Garefalaki et al. 2019). The
77 NAT enzymes have been investigated primarily as *N*-acetyltransferases (E.C.: 2.3.1.5)
78 that use acetyl-coenzyme A (CoA) to *N*-acetylate xenobiotic arylamines in ping-pong
79 Bi-Bi reactions catalyzed by a Cys-His-Asp triad similar to that of cysteine proteases
80 (Riddle and Jencks 1971; Sinclair et al. 2000). Although this conjugation reaction
81 usually contributes to xenobiotic detoxification, NAT enzymes can also bioactivate *N*-
82 hydroxaryl amines via *O*-acetylation (E.C.: 2.3.1.118), leading to the generation of
83 mutagenic metabolites. Aspects of NAT research and its history have been presented
84 in review articles, e.g. (Weber and Hein 1985; Grant 1993; Boukouvala and Fakis
85 2005; Sim et al. 2008; Butcher and Minchin 2012; Kubiak et al. 2013a; McDonagh et
86 al. 2014), and in a recent collective book edited by (Laurieri and Sim 2018).

87 Although the literature has described *NAT* genes and their recombinant
88 enzymatic products for several bacteria, e.g. (Watanabe et al. 1992; Sinclair et al.
89 2000; Sandy et al. 2002; Bhakta et al. 2004; Westwood et al. 2005; Rodrigues-Lima et
90 al. 2006; Suzuki et al. 2007; Pluvinage et al. 2007; Fullam et al. 2008; Martins et al.
91 2008; Takenaka et al. 2009; Abuhammad et al. 2013; Kubiak et al. 2013b; Cocaign et
92 al. 2014; Garefalaki et al. 2019, 2021), much less is known about NAT function in
93 eukaryotic microorganisms, such as fungi. A few studies have investigated fungal
94 NAT relative to the metabolism of xenobiotic arylamines, including certain by-

95 products of agrochemicals (Martins et al. 2009; Coccagna et al. 2013; Chan Ho Tong et
96 al. 2015). However, the most comprehensive knowledge about the biological role of
97 fungal *NAT* genes comes from studies of plant pathogens, implicating one particular
98 *Fusarium* homologue in the detoxification of naturally occurring benzoxazinoids
99 (BXs) (Glenn and Bacon 2009; Karagianni et al. 2015; Kettle et al. 2015).

100 BXs are phytoanticipins produced by corn, wheat, rye, wild barley and cane
101 (Zúñiga et al. 1983; Niemeyer 2009). The major role of BXs is plant protection
102 against competitors or pathogens (Niemeyer 2009; Bednarek 2012), such as weeds
103 (Barnes and Putnam 1987; Sicker and Schulz 2002; Tabaglio et al. 2008), insects
104 (Bohidar et al. 1986), nematodes and endophytic microorganisms (Zasada et al. 2005;
105 Niemeyer 2009). The first BX biosynthetic gene cluster was identified in maize (Frey
106 et al. 1997, 2009), generating the hydroxamic acid 2,4-dihydroxy-1,4-benzoxazin-3-
107 one (DIBOA) that is stored in the plant cell vacuole as a glucoside (Gierl and Frey
108 2001; Frey et al. 2009). When released, the aglucone of DIBOA is quickly degraded
109 to 2-benzoxazolinone (BOA) (Woodward et al. 1978; Hashimoto and Shudo 1996).
110 Although BOA is effective against many fungal pathogens, certain *Fusarium* species
111 have evolved mechanisms to overcome its toxicity (Glenn et al. 2001).

112 Study of BOA detoxification by the corn pathogen *Fusarium verticillioides*
113 (teleomorph *Gibberella moniliformis*, GIBMO) led to the mapping of two genetic
114 loci, *FDB1* and *FDB2*, responsible for the observed phenotypic tolerance of the
115 fungus (Glenn et al. 2002). Each of those two loci contains a cluster of co-regulated
116 genes, encoding enzymes involved in BOA detoxification (Glenn and Bacon 2009;
117 Glenn et al. 2016). In the first step of this pathway, *FDB1*-mediated decarbonylation
118 of BOA produces the intermediate metabolite 2-aminophenol (2AP), a toxic
119 arylamine. In the second step, an *FDB2*-encoded enzyme undertakes conjugation of

120 2AP to malonate, generating non-toxic *N*-(2-hydroxyphenyl)-malonamic acid (Glenn
121 et al. 2003). This *N*-malonyltransferase activity is malonyl-CoA dependent and is
122 attributed to the (GIBMO)NAT1 homologue of *F. verticillioides* (Glenn and Bacon
123 2009; Karagianni et al. 2015). The fungus has two additional *NAT* genes, encoding for
124 isoenzymes named (GIBMO)NAT2 and (GIBMO)NAT3 according to consensus
125 *NAT* nomenclature (Hein et al. 2008). The enzymatic function of (GIBMO)NAT2 is
126 unknown, while (GIBMO)NAT3 demonstrates characteristics typical of *NAT*
127 enzymes that *N*-acetylate arylamine xenobiotics (Karagianni et al. 2015).

128 In view of the unique *N*-malonyltransferase activity and important biological
129 role of (GIBMO)NAT1, here we present an investigation of its molecular structure,
130 biochemistry and phylogeny. Solving the structure of (GIBMO)NAT1 required the
131 high-yield production and crystallography grade purification of recombinant protein.
132 It was during this endeavour that we had the fortune of day-to-day interaction with
133 Professor Robert B. Sim at the Department of Pharmacology in Oxford. Bob, to
134 whom this article is dedicated, was a keen scientist who generously offered his
135 expertise and patiently guided us through the complexities of protein chromatography,
136 always ready to leave his desk for the bench.

137 **Results and discussion**

138 **Production and characterization of (GIBMO)NAT1 recombinant protein**

139 The (GIBMO)NAT1 protein of *F. verticillioides* was expressed in *Escherichia*
140 *coli* and purified through a series of chromatographic procedures. First, affinity
141 chromatography was used to extract the recombinant protein from soluble bacterial
142 lysate, providing highly pure protein in good yields. Ion exchange chromatography
143 was then performed through a positively charged column, based on the isoelectric

144 point of 5.9 calculated for (GIBMO)NAT1 (indicating a negative net charge at buffer
145 pH 7.5). Although the protein was eluted at low salt concentrations, suggesting loose
146 binding to the anion exchange column, remaining contaminants were removed at this
147 stage. Finally, the protein was fully purified by gel filtration chromatography
148 providing ~20 mg of recombinant (GIBMO)NAT1 per litre of bacterial culture (Fig.
149 1A and Suppl. Fig. S1). Under similar expression-purification conditions, production
150 of pure (GIBMO)NAT1 was much more efficient compared with human NAT1 (1.6
151 mg/l) and NAT2 (0.5 mg/l), the only other eukaryotic NAT proteins
152 crystallographically analyzed to date (Wu et al. 2007).

153 Samples retrieved at different stages during the purification process were
154 analyzed by SDS-PAGE, indicating a recombinant protein of about 40 kDa (Fig. 1A),
155 matching the molecular weight calculated for (GIBMO)NAT1 (40.548 kDa without
156 the *N*-terminal tail). Electrospray ionization mass spectrometry (ESI-MS) accurately
157 determined a molecular mass of 40.552 kDa for the protein (Suppl. Fig. S2A).
158 However, gel-filtration chromatography relative to appropriate standards (Suppl. Fig.
159 S2B) estimated 72.4 kDa as the molecular weight of the eluted recombinant protein,
160 i.e. almost twice the size determined by SDS-PAGE and ESI-MS. This finding
161 suggested that the protein may either form a dimer and/or have a non-spherical
162 molecular shape affecting its chromatographic mobility through the size-exclusion
163 column. Multi-angle light scattering (MALS) analysis also detected oligomers and
164 peaks with masses of 84, 89 and 170 kDa, i.e. about 2- and 4-fold the molecular mass
165 estimated for the monomeric protein by ESI-MS and SDS-PAGE. Finally, when
166 concentrated recombinant protein was tested by SDS-PAGE after storage on ice for
167 two weeks, a band of about twice the size of the monomeric protein appeared on gels.
168 A sample of the protein was reduced with dithiothreitol (DTT), while a second sample

169 was alkylated with iodoacetamide. Upon SDS-PAGE, the DTT-treated sample
170 provided only the monomeric band, while the iodoacetamide-treated sample provided
171 both the monomeric and the dimeric band (Suppl. Fig. S2C).

172 The fully purified recombinant protein was functionally validated by assaying
173 its enzymatic activity with malonyl- or acetyl-CoA as donor substrate, and with 5-
174 aminosalicylate (5AS) or *p*-aminobenzoate (PABA) as acceptor substrate. The protein
175 demonstrated the expected enzymatic activity of (GIBMO)NAT1, with malonyl-CoA
176 and 5AS being the preferred donor and acceptor substrate, respectively (Fig. 1B).

177 The (GIBMO)NAT1 monomer is about 3-10 kDa larger than the studied NAT
178 homologues of bacteria and mammals, including human and other primates (Sinclair
179 et al. 2000; Sandy et al. 2002; Westwood et al. 2005; Martins et al. 2008; Pluvinaige et
180 al. 2011; Abuhammad et al. 2013; Kubiak et al. 2013b; Cocaign et al. 2014; Tsirka et
181 al. 2018; Garefalaki et al. 2019, 2021). Indeed, the overwhelming majority of
182 annotated prokaryotic NAT proteins comprise polypeptide chains shorter than 300
183 amino acids, while mammalian NATs are 290 amino acids in length. In contrast,
184 eukaryotic microbes, including fungi, typically bear polypeptide chains longer than
185 300 amino acids, suggesting possible evolutionary and functional divergence (see
186 NAT website, <http://nat.mbg.duth.gr/>). Formation of dimers (and oligomers), that may
187 be enzymatically active, has been reported for certain prokaryotic NAT proteins
188 (Sinclair et al. 2000; Sandy et al. 2005; Fullam et al. 2008; Abuhammad et al. 2011,
189 2013; Cocaign et al. 2014). However, the implications of dimerization on enzyme
190 function have not been explored to date.

191 **Crystallographic determination of (GIBMO)NAT1 protein structure**

192 Initial crystallization screens of (GIBMO)NAT1 protein were performed with five
193 commercial 96-well blocks (see materials and methods). After one week, only a few

194 very flat (almost two-dimensional) rhomboid protein crystals were seen in one well of
195 the Morpheus block (Gorrec 2009) (Fig. 1C). After repeating for a longer incubation
196 period (> 2 weeks), Morpheus conditions (wells) D5, D9, E5, E9, F5, F9, H5 and H9
197 provided additional crystals with similar morphology, and three of them were
198 collected from well H5 (0.1 M amino acids, 0.1 M sodium HEPES/MOPS buffer pH
199 7.5, 30% PEGMME 550/PEG 20K precipitant stock) and taken to the synchrotron.

200 Diffraction data was collected and processed to a resolution of 1.8 Å and the
201 molecular structure of (GIBMO)NAT1 protein was determined by molecular
202 replacement (Table 1). The three-dimensional structure of (GIBMO)NAT1 (Fig. 2)
203 displays the typical NAT polypeptide fold consisting of a *N*-terminal α -helical bundle
204 (domain I, residues 1-126), a central β -barrel (domain II, residues 127-250), an
205 interdomain region (residues 251-273) and a *C*-terminal α/β lid (domain III, residues
206 274-345). The protein also carries the characteristic catalytic triad of NAT enzymes
207 (Sinclair et al. 2000), composed by residues Cys110, His158 and Asp173 (Fig. 2A,B).
208 Moreover, in the quaternary structure of (GIBMO)NAT1, the monomers were found
209 to strongly interact in pairs, with two homodimers forming a tetrameric arrangement
210 in the asymmetric unit (Fig. 2C,D). The interface between paired monomers is
211 extensive, involving several residues from domain I (amino acids 5-13, 16, 19, 55 and
212 122), domain II (amino acids 127-130, 134, 137, 140, 142, 144-145, 148-149, 152-
213 156, 164, 169, 208-215, 218-220, 222 and 245-250) and the interdomain region
214 (amino acids 252-253, 255-257, 259-260, 263-266 and 271-272), but not domain III
215 (except residue 295) (Fig. 2C). Computational PDBePISA analysis predicted a
216 thermodynamically solution-stable dimeric structure ($\Delta G^{\text{diss}} = 20.9$ kcal/mol) with an
217 interface of 2740.5 Å² between the two essentially identical monomers. This interface
218 is apparently formed through hydrophobic interactions and positive protein affinity

219 ($\Delta^{\ddagger}G = -11.1$ kcal/mol), and is further stabilized via hydrogen bonds ($N_{\text{HB}} = 48$) and
220 salt bridges ($N_{\text{SB}} = 18$), but not disulfide bonds ($N_{\text{DS}} = 0$). The same analysis ruled out
221 a physiological role for the dimer:dimer interface in the tetrameric arrangement,
222 therefore, the dimer is likely to be the biologically relevant assembly.

223 The catalytic residues of (GIBMO)NAT1 are located at the bottom of a deep
224 cavity. Access to this cavity is restricted by a "bridge-like" helical structure (amino
225 acids 138-150) separating two "tunnel-like" entries leading to the catalytic core (Fig.
226 3). Alignment of (GIBMO)NAT1 structure to other NAT proteins (two bacterial and
227 one human), co-crystallized with CoA, suggested that access of the compound to the
228 catalytic core could be through either of those two entries (Fig. 3). Interestingly, in the
229 crystallized dimeric form of (GIBMO)NAT1, one of those two entries of each subunit
230 is covered by the *N*-terminus of the other subunit. In fact, the *N*-terminus of each
231 monomer appears to be essential for tight subunit interaction in the dimer (Fig. 2D).

232 Although (GIBMO)NAT1 appears to combine features of both bacterial and
233 human NATs, it is also distinctively different (Suppl. Fig. S3). Unlike bacterial NATs
234 (Sinclair et al. 2000; Fullam et al. 2008; Xu et al. 2015), access to (GIBMO)NAT1
235 catalytic core is not through an exposed cleft, and is also entirely different from
236 (BACAN)NAT1 which has been reported as more eukaryotic-like (Pluvinage et al.
237 2011). Of the two "tunnel-like" entries of (GIBMO)NAT1, the first aligns with the
238 single entry of human NAT (Wu et al. 2007), but this is the entry covered in the
239 (GIBMO)NAT1 dimer. Conversely, the second "uncovered" entry of (GIBMO)NAT1
240 aligns with a part of human NAT where the catalytic pocket is covered by the *C*-
241 terminus and the so-called "eukaryotic-like" loop (Wu et al. 2007).

242

243

244 **Substrate binding to (GIBMO)NAT1 protein**

245 The interaction of monomeric (GIBMO)NAT1 with acetyl-, propionyl-,
246 malonyl- and succinyl-CoA, as well as with CoA and 2AP (i.e. the natural acceptor
247 substrate of the enzyme), was investigated by molecular docking analysis. The CoA
248 thiol group and the 2AP amino group were docked at 3.7 and 4.5 Å, respectively,
249 from the thiol group of catalytic Cys110. The 2AP additionally formed hydrogen
250 bonds with catalytic residue His158 and its adjacent Asn157 (Fig. 4). When studied as
251 a monomer in isolation, the protein bound all four acyl-CoAs via either of its two
252 "tunnel-like" entries (one covered and one uncovered in the dimeric form), but the
253 distance between the acyl-CoA reactive sulphur and the thiol group of catalytic
254 Cys110 varied considerably across different donor substrates. The distance was 4.0-
255 5.9 Å, 4.7-8.6 Å, 6.1-8.2 Å and 6.9-10.1 Å for the determined binding conformations
256 of acetyl-, propionyl-, malonyl- and succinyl-CoA, respectively. The majority of
257 conformations accessed the catalytic site through the "uncovered" entry, with the
258 exception of malonyl-CoA which appeared to orient itself closer to catalytic Cys110
259 through the "covered" entry (Suppl. Fig. S4). LigPlot⁺ analyses further allowed a
260 more detailed view of polar and non-polar contacts between determined acyl-CoA
261 conformations and the amino acids of (GIBMO)NAT1 monomer (Suppl. Fig. S5).
262 Although a number of binding conformations were assessed as plausible in terms of
263 spatial fitting within the catalytic core through the two tunnels of (GIBMO)NAT1
264 monomer, the distance and positioning of acyl-CoA reactive sulphur relative to
265 Cys110 thiol group could not explain why this particular enzyme prefers malonyl-
266 CoA versus other acyl-group donor substrates. In fact, acetyl- and propionyl-CoA
267 seemed to approach Cys110 more intimately (< 5 Å) than malonyl-CoA (> 6 Å).

268 Docking was thus repeated for the dimeric form of (GIBMO)NAT1, using all
269 four donor substrates and CoA as ligands (Figs. 5 and 6). It was readily evident that
270 the *N*-terminus of one monomer (completely covering the first tunnel to the active site
271 of the other monomer) changed the ability of different donor substrates to approach
272 catalytic Cys110 through the second (uncovered) tunnel. Fitting of CoA in the active
273 site was essentially the same as in the monomer (Fig. 4), maintaining a thiol group
274 distance of 3.7 Å with Cys110 (Fig. 5B). In the case of acetyl-CoA (Fig. 5C), the
275 distance between its reactive sulphur and the thiol group of Cys110 was shorter (3.2
276 Å) than in the monomer (4 Å), and LigPlot⁺ analysis predicted the formation of a
277 hydrogen bond between the two moieties (Fig. 6B). A second hydrogen bond was
278 predicted between the side chain of Thr261 and the carbonyl group of acetyl-CoA
279 (Fig. 6B). In contrast, due to its longer side chain, propionyl-CoA failed to approach
280 Cys110 closer than 6.9 Å (Fig. 5D), and no stabilizing hydrogen bonds were predicted
281 by LigPlot⁺ analysis between the compound and the active site (Fig. 6C).

282 When docked to the dimeric form of (GIBMO)NAT1, the dicarboxylic
283 substrate malonyl-CoA readily oriented its reactive sulphur within 4.4 Å from the
284 thiol group of catalytic Cys110, in a plausible conformation relative to the catalytic
285 triad (Fig. 5E). This was stabilized through hydrogen bonds between the terminal
286 carboxyl group of the compound and the active site residues Arg109 and Asn157
287 (adjacent to catalytic Cys110 and His158, respectively). A third hydrogen bond was
288 formed between the terminal carboxyl group of malonyl-CoA and Thr261 (Fig. 6D).
289 The same hydrogen bond arrangement was also observed for succinyl-CoA, i.e. the
290 second dicarboxylic acyl-CoA compound investigated, but its longer side chain
291 prevented the thioester sulphur from approaching Cys110 closer than 5.8 Å (Figs. 5F

292 and 6E). Overall, those results support malonyl-CoA to be the selective donor
293 substrate reacting with (GIBMO)NAT1 catalytic triad.

294 The results also demonstrate the dimeric form as the functional unit of
295 (GIBMO)NAT1 protein. Protein-protein docking for the two interacting monomers
296 was performed with ClusPro, PatchDock and pyDOCK software (Fig. 7), and all three
297 programmes predicted dimerization along roughly the same interface determined by
298 crystallography (Fig. 2). As expected, the *N*-terminus of one monomer extended well
299 onto the surface of the other monomer, usually covering one of the two entries to the
300 active site. However, some models also demonstrated a dimeric conformation with
301 both entries open, suggesting functional flexibility of the *N*-terminal tail (Fig. 7).

302 **Comparison of (GIBMO)NAT1 with other fungal NAT homologues**

303 The crystallographic structure of (GIBMO)NAT1 was used to model the secondary
304 (Suppl. Fig. S6) and tertiary (Fig. 8) structure of *F. verticillioides* (GIBMO)NAT2
305 and (GIMBO)NAT3 isoenzymes. The template protein was differentiated mainly by
306 its characteristically elongated *N*-terminal extension (residues 1-18), an expanded α -
307 helical region (α 2- η 2- α 3, residues 41-60) in domain I, the distinctive "bridge-like"
308 protruding helical structure (η 4- α 6, residues 138-150) separating the two "tunnel-like"
309 entries of the active site, the shorter β 6 strand of the second β -sheet in domain II
310 (residues 190-196), two protruding loops between strands β 7- β 8 (residues 206-221,
311 involved in dimer formation) and β 8- β 9 (residues 228-238) of the same β -sheet, as
312 well as the shorter loop (residues 280-285) connecting β 10- β 11 strands of domain III
313 β -sheet, and the longer *C*-terminus (η 8, residues 339-345) (Fig. 8 and Suppl. Fig. S6).
314 The "bridge-like" structure between the two active site entries of (GIBMO)NAT1 was
315 much slimmer in (GIBMO)NAT2 and completely absent in (GIBMO)NAT3, where
316 access to the active site is through an exposed cleft along the surface of the molecule

317 (Fig. 8), resembling bacterial NATs (Sinclair et al. 2000), but not human NATs (Wu
318 et al. 2007). The (GIBMO)NAT1 dimerization surface was also less evident in
319 (GIBMO)NAT2 and (GIBMO)NAT3, which lacked a *N*-terminal tail and displayed a
320 shorter β 7- β 8 loop (Fig. 8 and Suppl. Fig. S6).

321 Among fungal NAT proteins functionally investigated before (Karagianni et
322 al. 2015), the distinctive structural features of (GIBMO)NAT1 from *F. verticillioides*
323 were all conserved in the homologous (GIBZE)NAT1 of *F. graminearum* and
324 (FUSOX)NAT1 of *F. oxysporum* f.sp. *lycopersici*. Structural conservation was also
325 evident among the NAT2 and among the NAT3 homologues of those three fungi
326 (Suppl. Fig. S7). (FUSOX)NAT4 demonstrated features more similar to the NAT2
327 homologues and so did (ASPFN)NAT3 of *Aspergillus flavus*. Remarkably, the
328 acetyltransferases (EMENI)NAT1 of *Aspergillus nidulans* and (ASPFN)NAT2 of *A.*
329 *flavus* showed some similarity with *Fusarium* NAT1 malonyltransferases, evidenced
330 by the two separated "tunnel-like" entries to the active site and the divulging loop
331 involved in dimerization of (GIBMO)NAT1 (Suppl. Fig. S7).

332 Sequence alignment (Suppl. Fig. S6) of those thirteen fungal NATs revealed
333 key differences in the active site residues shown by docking analysis to selectively
334 interact with malonyl-CoA via hydrogen bonding (Fig. 6). Specifically, Arg109,
335 Asn157 and Thr261 were found in all *Fusarium* NAT1 malonyltransferases, but
336 differed in the *Fusarium* NAT3 acetyltransferases. *Fusarium* NAT2 homologues had
337 only Thr261, while (FUSOX)NAT4 had only Asn157. Among *Aspergillus* NATs, the
338 acetyltransferases (EMENI)NAT1 and (ASPFN)NAT2 had only Thr261, while
339 (ASPFN)NAT3 differed at all three amino acid positions (Suppl. Fig. S6).

340

341 **Changes in protein thermal stability upon interaction with components of the**
342 **enzymatic reaction**

343 Differential scanning fluorimetry (DSF) has been proposed as an easy method for
344 inferring functionality of NAT homologues (Garefalaki et al. 2019, 2021) and was
345 previously employed to determine selective acyl-CoA binding to the three NAT
346 isoenzymes of *F. verticillioides* (Karagianni et al. 2015). Here, those proteins were
347 assayed with different components of the NAT enzymatic reaction as ligands.
348 Specifically, recombinant protein was tested with donor only (acetyl-, propionyl- or
349 malonyl-CoA), acceptor only (3,4-dichloroaniline; 3,4-DCA), or the two substrates
350 combined, as well as with CoA only, *N*-acylated arylamine only or the two products
351 combined (Fig. 9). (GIBMO)NAT1 was tested with malonyl-CoA as preferred donor
352 substrate and *N*-malonylated 3,4-DCA as the corresponding product (Fig. 9A).
353 (GIBMO)NAT3 was tested with its preferred acetyl-CoA substrate and *N*-acetylated
354 3,4-DCA (Fig. 9B). Although apparently inactive enzymatically, (GIBMO)NAT2 has
355 previously been demonstrated to non-selectively bind acetyl-, malonyl- and
356 propionyl-CoA (Karagianni et al. 2015); therefore, DSF was performed with all three
357 donor compounds and their respective conjugated 3,4-DCA derivatives (Fig. 9C).

358 The functional isoenzymes (GIBMO)NAT1 and (GIBMO)NAT3 interacted
359 with their selective donor substrates malonyl- and acetyl-CoA, respectively. This was
360 evident by the marked increase (by 5.3 and 21.8 °C, respectively) in the denaturation
361 midpoint transition temperature (T_m) of the two proteins (Fig. 9A,B). Consistent with
362 the ping-pong Bi-Bi reaction mechanism of NAT enzymes, no interaction was evident
363 with 3,4-DCA alone, but the two substrates together increased the T_m . Moreover, CoA
364 alone or with acylated arylamine stabilized the two isoenzymes (Fig. 9A,B). Unlike
365 bacterial NATs (Garefalaki et al. 2019, 2021), (GIBMO)NAT1 and (GIBMO)NAT3

366 were also stabilized by their respective terminal reaction product, i.e. the conjugated
367 arylamine (Fig. 9A,B), suggesting that they may not release it as effectively.

368 A more complex picture was observed for (GIBMO)NAT2. As expected, the
369 isoenzyme bound all three acyl-CoA compounds tested, when those were used alone
370 or with 3,4-DCA. In contrast, neither 3,4-DCA, nor any of its acylated derivatives
371 appeared to interact with the protein (Fig. 9C). An interesting pattern was observed
372 for CoA (alone or with acyl-3,4-DCA), as this produced protein thermal denaturation
373 curves that were biphasic. Both phases demonstrated increased T_m values, with shifts
374 more moderate (3.2-4.6 °C) for the lower temperature range (T_{m1}) and much more
375 pronounced (21.9-22.9 °C) for the higher temperature range (T_{m2}) (Fig. 9C).

376 **Enzymatic preference for donor substrates**

377 Enzymatically active (GIBMO)NAT1 and (GIBMO)NAT3 were assayed over a
378 concentration range (0-300 μ M) of donor substrates acetyl-, propionyl- or malonyl-
379 CoA, using 5AS as acceptor substrate. Enzymatic activity of (GIBMO)NAT1 with
380 malonyl-CoA was substantially higher than with acetyl- or propionyl-CoA, and the
381 opposite pattern was evident for (GIBMO)NAT3 (Fig. 10). Although highly selective
382 for malonyl-CoA, (GIBMO)NAT1 also produced 5-10 fold lower activity with acetyl-
383 and propionyl-CoA (Fig. 10A). In contrast, the activity generated by (GIBMO)NAT3
384 with malonyl-CoA was marginal compared with acetyl- and propionyl-CoA (Fig.
385 10B). Overall, with selective substrates, activity was substantially higher (up to 130-
386 fold) for (GIBMO)NAT3, compared with (GIBMO)NAT1.

387 The enzymatic assays were repeated over a broader concentration range (0-
388 5000 μ M) of malonyl-CoA for (GIBMO)NAT1 (Fig. 10C) and acetyl-CoA for
389 (GIBMO)NAT3 (Fig. 10D), using excess 5AS to saturate the second step of the
390 reaction. Compared with the previous experiments, a substantial decrease in activity

391 was observed, potentially due to enzyme inhibition by the amount of generated
392 products. This is in line with the DSF experiments above, that demonstrated
393 stabilization of both isoenzymes by CoA and their respective conjugated arylamine.

394 **Phylogeny of fungal NAT homologues**

395 Previous studies have demonstrated the monophyletic origin of fungal *NAT* genes,
396 encountered mainly in filamentous ascomycetes (subphylum *Pezizomycotina*) (Glenn
397 et al. 2010; Martins et al. 2010). Early phylogenetic analyses (Glenn et al. 2010)
398 indicated distinct lineages of *NAT* orthologues in the ascomycetes, particularly within
399 the clades of plant pathogens. As the number of sequenced fungal genomes has
400 increased considerably (9500 genomes in the Genome database in November 2021),
401 we have updated the list of annotated *NAT* genes, dissecting their phylogeny within
402 the kingdom (Suppl. Fig. S8). The results indicate *NAT* genes to be absent in
403 ascomycetes other than *Pezizomycotina*, although they are found in basidiomycetes
404 and in the lower taxa of *Chytridiomycetes* (e.g. *Batrachochytrium*, *Rhizoclosmatium*),
405 *Zoopagomycota* (e.g. *Conidiobolus*, *Entomophthora*, *Basidiobolus*) and
406 *Mucoromycota* (e.g. *Jimgerdemannia*, *Umbelopsis* and several species in the family
407 of *Mortierellaceae*). The *NAT* homologues of filamentous ascomycetes form a clearly
408 distinct clade, unlike the homologues of lower fungi where mixed phylogeny of
409 sequences is apparent in chytrids and the two phyla (*Zoopagomycota*, *Mucoromycota*)
410 of zygomycetes. Monophyly is also less evident for the *NAT* sequences of
411 basidiomycetes that exhibit low basal resolution and are placed closer to the chytrids.

412 The phylogeny of *NAT* homologues in filamentous ascomycetes exhibits
413 separate orthologous lineages, particularly in species (mostly plant pathogens) with
414 multiple *NAT* sister loci in their genome (Fig. 11 and Suppl. Fig. S8). The analysis
415 demonstrates four such lineages, defined by the 13 functionally investigated *NAT*

416 homologues described above. The first lineage, for which no specific function is yet
417 known, is split into one sub-lineage that includes (GIBMO)NAT2 of *F. verticillioides*,
418 (GIBZE)NAT2 of *F. graminearum* and (FUSOX)NAT2 of *F. oxysporum* f.sp.
419 *lycopersici*, plus another sub-lineage with (FUSOX)NAT4 and (ASPFN)NAT3 of *A.*
420 *flavus*. The second lineage includes the acetyltransferases (ASPFN)NAT2 of *A. flavus*
421 and (EMENI)NAT1 of *A. nidulans*. The third and fourth lineages arise from the same
422 main clade and include the NAT3 acetyltransferases and NAT1 malonyltransferases
423 of *Fusarium*. Lineage separation precedes speciation, indicating evolutionary
424 diversification of fungal NATs to serve new functions. NAT1 malonyltransferases
425 have diverged from the acetyltransferases, consistent with our experimental findings.

426 **Concluding remarks**

427 The vast repertoire and functional variability of xenobiotic metabolizing enzymes is
428 the outcome of billions of years of co-evolution and adaptation of organisms, in
429 constant need of protection against the chemical arsenal of competitors and under the
430 stress of survival in fluctuating chemical and/or nutritional environments (Nebert
431 1997; Nebert and Dieter 2000). Although the evolutionary history of NAT enzyme
432 family is much older than the advent of manmade chemicals, investigators have
433 focused their interest almost exclusively on the role of NAT in the metabolism and
434 detoxification (or bioactivation) of drugs, carcinogens, pollutants, agrochemicals,
435 industrial materials, etc. (Laurieri and Sim 2018). The NAT1 isoenzyme of *F.*
436 *verticillioides* and other *Fusarium* species associated with cereal plants represents the
437 only well-understood example of how NAT may enhance the natural potential of
438 endophytic fungi for survival and infection, via annihilation of host chemical defence
439 (Glenn and Bacon 2009; Kettle et al. 2015; Baldwin et al. 2019). The NAT1
440 homologue of *Fusarium* is also the only NAT enzyme known to function as a *N*-

441 malonyltransferase both *in vitro* and *in vivo* (Glenn and Bacon 2009; Karagianni et al.
442 2015), clearly diverging from *Fusarium* NAT2 and NAT3, but also from other known
443 NAT homologues, as the present study demonstrates. A prokaryotic NAT with
444 preference for malonyl-CoA *in vitro* was recently found in the actinobacterium
445 *Tsukamurella paurometabola* (Garefalaki et al. 2019), but its biological role and *in*
446 *vivo* substrate selectivity remains elusive (Kontomina et al. 2021).

447 As this study concludes, the (GIBMO)NAT1 isoenzyme of *F. verticillioides*
448 has evolved to selectively employ malonyl-CoA, instead of acetyl-CoA, via a
449 remarkable adaptation of its functional structural unit and catalytic mechanism,
450 involving dimerization of the protein and interaction of specific active site residues
451 with the terminal carboxyl-group of malonate. This is the first description of NAT
452 substrate selectivity and enzyme function relying on dimer formation, and the first
453 demonstration by docking analysis of selective acyl-CoA binding to the NAT active
454 site.

455 *F. verticillioides* and other related endophytes can be devastating to maize and
456 wheat, causing crop disease and contaminating produce with harmful mycotoxins.
457 Those fungi are, thus, associated with severe economic loss and social hardship,
458 especially in the developing world (Robens and Cardwell 2003; Wu 2004). Given the
459 role of NAT1 in enhancing *Fusarium* virulence and resistance to BOA, unravelling
460 the molecular, structural, biochemical and physiological aspects of this important
461 enzyme may benefit research into more targeted strategies for pathogen control,
462 limiting harvest loss and mycotoxin exposure of domestic animals and humans.

463

464

465 **Materials and methods**

466 **Recombinant protein expression**

467 A frozen stock (-80 °C, 25% v/v glycerol) of *E. coli* BL21(DE3)pLysS cells
468 (Promega) carrying the (GIBMO)*NAT1* open reading frame in pET28b(+) (Novagen)
469 vector (Karagianni et al. 2015) was used to inoculate 10 ml of LB medium (10 g/l
470 tryptone, 5 g/l yeast extract, 5 g/l NaCl, 50 µg/ml kanamycin, pH 7.0 ± 0.2), followed
471 by overnight incubation (37 °C, 180 rpm). The next day, 500 ml of Terrific Broth
472 (TB) medium were inoculated with 5 ml (1% v/v) of starter culture and incubated
473 under the same conditions. The TB medium was prepared with BD Difco reagents, as
474 follows: two solutions were made and autoclaved separately, the first (0.9 l)
475 containing 12 g tryptone, 24 g yeast extract and 4 ml glycerol, and the second (0.1 l)
476 containing 2.31 g KH₂PO₄ and 12.54 g K₂HPO₄. After cooling to 37 °C, the two
477 solutions were mixed and kanamycin (50 µg/ml) was added.

478 When culture optical density (OD) at 600 nm reached 0.8-1.5 units,
479 recombinant protein expression was induced by addition of isopropyl β-D-1-
480 thiogalactopyranoside (IPTG) at 1 mM. After overnight incubation (15 ± 1 °C, 180
481 rpm), induced cells were harvested by centrifugation at 6,000 g (4 °C, 20 min) and
482 their pellet was stored at -80 °C, at least overnight, to facilitate lysis. The frozen cell
483 paste was slowly thawed on ice and resuspended in lysis buffer [20 mM Tris-HCl pH
484 7.5, 300 mM NaCl, 5 mM imidazole, 2 mM β-mercaptoethanol, 5% v/v glycerol, 1x
485 ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor cocktail (Roche) and
486 0.1% w/v 3-((3-cholamidopropyl)dimethyl-ammonium)-1-propanesulfonate
487 (CHAPS)]. Approximately 30 ml of lysis buffer were used per 1 l of culture. The
488 suspension was sonicated on ice, performing 20 or more cycles of vibration at 10 kHz
489 (45 s of vibration, 30 s of pause), as necessary. Deoxyribonuclease I (DNaseI) was

490 then added at 10 µg/ml (10 min, room temperature), to fully digest the bacterial DNA.
491 After centrifugation at 20,000 g (20 min, 4 °C) in a Sorvall centrifuge, the soluble
492 fraction (supernatant) of each bacterial lysate was recovered and maintained on ice at
493 4 °C for subsequent purification. Samples of uninduced cells, as well as of the
494 insoluble fraction of bacterial lysate (pellet), were retained for comparison.

495 **Recombinant protein purification**

496 *Affinity chromatography*

497 Immobilized metal affinity chromatography was performed to purify hexa-His tagged
498 recombinant NAT proteins from the soluble fraction of bacterial lysates, using
499 commercially available (Thermo Scientific) nitrilotriacetic acid (NTA) agarose resin
500 positively charged with Nickel (Ni^{2+}) and maintained in 50% v/v ethanol. The
501 procedure took place at 4 °C, using 2-4 ml of resin per soluble lysate derived from 1 l
502 of bacterial culture. Before use, the ethanol was eluted and the resin was equilibrated
503 with 10 volumes of washing buffer (20 mM Tris HCl pH 7.5, 300 mM NaCl). The
504 soluble bacterial lysate was then loaded, followed by gentle mixing (45 min, orbital
505 shaker), before allowing flow through the resin. Ni^{2+} -bound protein was eluted by
506 successive washes with increasing concentrations of imidazole (10, 25, 50, 100, 200
507 and 250 mM) in buffer with 20 mM Tris HCl pH 7.5 and 300 mM NaCl (5-10 times
508 the resin volume). The purified protein was kept on ice at 4 °C for immediate use,
509 avoiding freeze-thawing. A sample of each chromatographic fraction was examined
510 for purity of recombinant protein by sodium dodecyl sulphate-polyacrylamide gel
511 electrophoresis (SDS-PAGE) and the yield was assessed at 280 nm.

512 *Ion exchange chromatography*

513 Before further purification, the affinity chromatography fractions (10-20 ml) were
514 transferred to low-salt buffer (20 mM Tris-HCl pH 7.5, 20 mM NaCl), following

515 overnight dialysis with gentle magnetic stirring (4 °C) through a 10 kDa cut-off
516 porous membrane (Sigma). The *N*-terminal hexa-His tag was then removed, after
517 overnight incubation (gently rotating at 4 °C) with human plasma thrombin (Sigma)
518 added at 1 unit per mg of affinity chromatography purified protein. The preparations
519 were spun at maximum speed in a micro-centrifuge (4 °C) to remove any precipitated
520 protein debris, and high performance ion exchange chromatography followed through
521 a positively charged MonoQ 5/50 GL column on an ÄKTA purifier core system (GE
522 Healthcare). Increase of ionic strength was achieved by gradual mixing of buffer A
523 (20 mM Tris-HCl pH 7.5, 50 mM ϵ -aminocaproic acid, 5 mM EDTA) with buffer B
524 (20 mM Tris-HCl pH 7.5, 50 mM ϵ -aminocaproic acid, 5 mM EDTA, 1 M NaCl). All
525 buffer solutions were vacuum filtered before use. The chromatography flow rate was
526 set at 1.5 ml/min with 1 ml collected per fraction, and the process was monitored
527 using UNICORN 5.11 software (GE Healthcare). Elution of recombinant protein was
528 detected at 280 nm and the corresponding eluted fractions from each run were
529 analyzed by SDS-PAGE and pooled together.

530 *Gel filtration chromatography*

531 High performance gel filtration chromatography was performed on the ÄKTA purifier
532 core system (GE Healthcare), using a Superdex 75 10/300 GL column to purify
533 proteins with molecular weight of up to 75 kDa. The vacuum-filtered
534 chromatographic buffer solution consisted of 20 mM Tris-HCl (pH 7.5) and 300 mM
535 NaCl. The flow rate was set at 0.5 ml/min with 0.5 ml collected per fraction. Elution
536 of recombinant protein was monitored at 280 nm and the pooled fractions of
537 successive runs were inspected by SDS-PAGE.

538

539 **Molecular weight determination of recombinant protein**

540 Gel filtration chromatography was additionally used to estimate the molecular weight
541 of purified NAT protein. Marker proteins (Sigma) were loaded onto the column and
542 their elution volume was plotted against the logarithm of their known molecular
543 weight. Those markers were 500 µg of 66 kDa albumin from bovine serum, 500 µg of
544 44.3 kDa albumin from chicken egg white, 250 µg of 30 kDa carbonic anhydrase
545 from bovine erythrocytes and 250 µg of 12.3 kDa of cytochrome-c from bovine heart
546 (which absorbs at 550 nm). From the elution volume of recombinant NAT protein, it
547 was possible to calculate its approximate molecular weight relative to the markers.

548 ESI-MS was further used to determine the molecular weight of pure
549 recombinant NAT protein (1 mg/ml in 20 mM Tris-HCl pH 7.5 and 20mM NaCl),
550 followed by MALS analysis. Those services were provided by the Biophysical
551 Instrument Facility of the Department of Biochemistry, University of Oxford, U.K.

552 **Protein crystallization and structure determination**

553 Prior to crystallization experiments, the preparation of fully purified NAT protein was
554 concentrated to ~10 mg/ml in buffer with 20 mM Tris-HCl pH 7.5 and NaCl adjusted
555 to 20 mM. This was performed by centrifugation (2,000 g, 4 °C) through a 3 kDa cut-
556 off pore size filter concentrator (Microcon), until the appropriate reduction in protein
557 solution volume and salt concentration was achieved. DTT (1 mM) was then added.

558 For crystallographic screens, an automated robotic system (Tecan Genesis
559 ProTeam 150) was used to prepare 96-well crystallization plates (MA 96 Innovaplate
560 SD-2), dispensing 50 µl/well of screen solution from commercial blocks Morpheus,
561 JCSG Plus, Stura I+II_Natrix, Structure Screen 1+2 HT-96 (Molecular Dimensions)
562 and Wizard Full I+II (Emerald BioSystems). The sitting-drop vapour diffusion
563 method was used, employing a Mosquito (TTP Labtech) automated system to

564 dispense 0.1 μ l drops of the protein sample and of the screen solution to the
565 crystallization plates. The plates were then sealed with optically clear tape and placed
566 into a cold (4 °C) robotic incubator (CrystalMation) equipped with an automated
567 imaging system. Crystallization plates were incubated for at least one week at 4 °C.

568 Under a stereomicroscope (Leica), crystals were inspected and recovered from
569 the crystallization drops with a mesh, followed by flash-cooling in liquid nitrogen.
570 Diffraction data was collected on beamline ID14.EH4 at the European Synchrotron
571 Radiation Facility (ESRF), Grenoble, France. Data was collected and processed to a
572 resolution of 1.8 Å using DIALS (Beilsten-Edmands et al. 2020) and the CCP4 (Winn
573 et al. 2011) software suite. The structure was determined by molecular replacement in
574 space group P1 with four molecules in the asymmetric unit, using PHASER (McCoy
575 et al. 2007). The search model used was an ensemble constructed from chains A-D of
576 PDB ID: 1W6F and chains A-D of PDB ID: 1QX3 using MRBUMP (Keegan et al.
577 2018) and CCP4MG (McNicholas et al. 2011) to align the chains and remove regions
578 of high variance. A model was built using Coot (Emsley et al. 2010) and refined using
579 REFMAC5 (Murshudov et al. 2011) and phenix.refine (Afonine et al. 2012). The
580 refined structure was deposited in the Protein Data Bank (PDB) with the accession
581 number 7QI3.

582 **Computational analysis of protein structure**

583 ProtParam (Expasy) was used to calculate the molecular weight, pI and extinction
584 coefficient of (GIBMO)NAT1 protein. Other thermodynamic parameters were
585 determined by PDBePISA (Krissinel and Henrick 2007). T-Coffee Expresso
586 (Notredame et al. 2000) was used for protein sequence structural alignment,
587 visualized with ESPript 3.0 (Robert and Gouet 2014). Visualization of modelled and
588 crystallographic protein structures was performed on PyMOL (Schrödinger) and

589 UCSF Chimera (Pettersen et al. 2004). Homology modelling was performed on
590 Swiss-Model (Waterhouse et al. 2018). Docking experiments employed
591 AutoDockTools v.1.5.6 and AutoDock Vina (Trott and Olson 2010; Forli et al. 2016).
592 Workflow optimization was initially performed using the available crystallographic
593 structure of human NAT2 co-crystallized with CoA (PDB ID: 2PFR), and the results
594 were validated against the published experimental data (Wu et al. 2007). Docking
595 analysis was then implemented for (GIBMO)NAT1, as described below.

596 UCSF Chimera was used to prepare the protein molecule (".pdb" file), by
597 removing possible miscellaneous co-crystallized elements (e.g. ions, metals,
598 substrates, water molecules etc.). Further processing took place on AutoDockTools,
599 which was used to incorporate polar hydrogens and distribute electrostatic charges to
600 the protein molecule. The structure was then saved in format compatible with
601 AutoDock Vina (".pdbqt" file). A similar approach was used to prepare ligands for
602 docking, retrieved from the DrugBank (<https://go.drugbank.com/>) or ChemSpider
603 (<https://www.chemspider.com/>). Those compounds were: CoA (DrugBank ID:
604 DB01992), acetyl-CoA (ChemSpider ID: 392413), malonyl-CoA (ChemSpider ID:
605 559121), propionyl-CoA (ChemSpider ID: 83731), succinyl-CoA (ChemSpider ID:
606 83179) and 2-AP (ChemSpider ID: 5596). Ligand structure files were first prepared in
607 ".pdb" format on UCSF Chimera, as necessary. AutoDockTools was then used to add
608 polar hydrogens and charges, as well as to specify the torsional degrees of freedom in
609 ligand molecules that allowed bond flexibility. A ".pdbqt" file was then generated.
610 The suitable grid box for docking was finally defined for the protein, and this
611 enclosed the whole active site with its two "tunnel-like" entries.

612 Docking experiments were executed in AutoDock Vina. Each round produced
613 nine different ligand conformations, and thirty rounds were implemented per protein-

614 ligand pair (270 conformations in total). Those binding conformations were visually
615 inspected one-by-one on PyMOL, discarding those that were not plausible (i.e. they
616 either did not bind to the active site at all, or bound in the wrong orientation).
617 Plausible conformations were further examined for their positioning (i.e. distance and
618 orientation of their reactive sulphur atom) relative to the thiol group of catalytic
619 Cys110. Binding affinity and other spatial features were also assessed, as well as
620 specific ligand-protein interactions. LigPlot⁺ (Laskowski and Swindells 2011) was
621 additionally used to determine and visualize hydrogen bonds and hydrophobic
622 interactions with specific amino acid residues, finally leading to the selection of
623 optimal ligand conformations.

624 Additional protein-protein docking experiments were used to simulate
625 interaction of (GIBMO)NAT1 monomers in their dimeric form, using ClusPro
626 (Comeau et al. 2004), PatchDock (Schneidman-Duhovny et al. 2005) and
627 pyDockWEB (Jiménez-García et al. 2013). Reconstruction of the *N*-terminus was
628 performed using PyMOL Builder, with ModLoop (Modeller) used for refinement
629 (Fiser and Sali 2003). Default parameters were used to run those algorithms.

630 **Enzymatic characterization of fungal NAT recombinant proteins**

631 Affinity chromatography purified (GIBMO)NAT1, (GIBMO)NAT2 and
632 (GIBMO)NAT3 recombinant proteins were enzymatically investigated by
633 performing: i) DSF analysis to determine changes in protein midpoint transition
634 temperature (T_m) upon interaction with ligands (substrates or products of the
635 enzymatic reaction, alone or combined); ii) colourimetric assays to measure enzyme
636 specific activity with various substrates; iii) Enzymatic Michaelis-Menten
637 experiments to determine apparent kinetic parameters against donor substrates. Those
638 experiments were performed following previously published procedures of our

639 laboratory, described for NAT enzymes of bacteria (Garefalaki et al. 2019, 2021),
640 fungi (including *F. verticillioides*) (Karagianni et al. 2015) and primates (Tsirka et al.
641 2014, 2018). Details are provided in the supplementary materials and methods.

642 **Phylogenetic analyses of fungal NAT sequences**

643 Genomic database searches for fungal *NAT* sequences were carried out as previously
644 described (Glenn et al. 2010; Garefalaki et al. 2019) and details are provided in the
645 supplementary materials and methods.

646 **Acknowledgements**

647 The authors would like to acknowledge the support of international mobility
648 programmes providing opportunities for collaboration, as follows: a) Democritus
649 University of Thrace internship to E.K. to visit Oxford (U.K.) in 2011; b) European
650 Union FP7 “Capacities” Coordination & Support Action REGPOT-2008-1
651 programme (acronym BioStrength) grant to G.F. and S.B. to visit Oxford (U.K.), and
652 to A.E.G. to visit Alexandroupolis (Greece), in 2011; c) Fulbright-Schuman research
653 scholarship to S.B. to visit Athens (GA, U.S.A.), in 2012; d) Democritus University
654 of Thrace Leonardo da Vinci scholarship to S.Z. to visit Kingston-upon-Thames
655 (U.K.) in 2014; e) Erasmus+ International Credit Mobility grant to S.B. to visit
656 Athens (GA, U.S.A.) in 2018; f) University of Angers internship to T.G. in 2021 to
657 work with the Greek team. Additionally, P.R. was the recipient of a LISCB Wellcome
658 Trust ISSF award, grant reference 204801/Z/16/Z, and of a Wellcome Trust Seed
659 Award in Science, grant reference 214090/Z/18/Z; A.A. was the recipient of a
660 University of Jordan Ph.D. scholarship, while at Oxford; E.K. was the recipient of a
661 Ph.D. scholarship (2016-2019) co-financed by Greece and the European Union
662 (European Social Fund-ESF) through Operational Programme “Human Resources

663 Development, Education and Lifelong Learning” in the context of project
664 “Strengthening Human Resources Research Potential via Doctorate Research” (MIS-
665 5000432), implemented by the State Scholarships Foundation (IKY). The funding
666 sources had no involvement in the conduct of the research, the preparation of the
667 article or the decision to submit the article for publication.

668 **Conflicts of interest/Competing interests**

669 The authors have no conflicts of interest to declare.

670 **Availability of data and material**

671 All data generated and/or analyzed during the study are included in this article and its
672 supplementary information. Other materials are available from the corresponding
673 author on reasonable request. The crystallographic structure of (GIBMO)NAT1
674 protein is deposited in the Protein Data Bank (PDB ID: 7QI3).

675 **Authors' contributions**

676 E.P.K., E.K., E.D.L. and K.A. (presented according to the chronological order of their
677 participation in the project) performed the main experimental work relevant to
678 crystallography and the ensuing structural and functional analyses, with equal
679 contributions; G.P., V.G., V.K. and S.Z. carried out additional work with recombinant
680 proteins; T.G. performed homology modelling; E.K., T.G., K.A. and M.A.T.
681 contributed to genomic searches and phylogenetic analyses; G.F., A.E.G., P.R., A.A.,
682 A.R., R.B.S., E.S. and S.B. guided the work of students and/or supported project
683 implementation with specialized expertise and resources; S.B. conceived of and
684 coordinated the project, contributed to the experimental procedures and wrote the
685 manuscript with input from other authors. The bulk of work was carried out at
686 Democritus University of Thrace, Greece, and at the University of Oxford, U.K.

687 **References**

- 688 Abuhammad A, Lack N, Schweichler J, et al (2011) Improvement of the expression
689 and purification of *Mycobacterium tuberculosis* arylamine N-acetyltransferase
690 (TBNAT) a potential target for novel anti-tubercular agents. *Protein Expr Purif*
691 80:246–252. <https://doi.org/10.1016/j.pep.2011.06.021>
- 692 Abuhammad A, Lowe ED, McDonough MA, et al (2013) Structure of arylamine N-
693 acetyltransferase from *Mycobacterium tuberculosis* determined by cross-seeding
694 with the homologous protein from *M. marinum*: triumph over adversity. *Acta*
695 *Crystallogr D Biol Crystallogr* 69:1433–1446.
696 <https://doi.org/10.1107/S0907444913015126>
- 697 Afonine P V, Grosse-Kunstleve RW, Echols N, et al (2012) Towards automated
698 crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol*
699 *Crystallogr* 68:352–367. <https://doi.org/10.1107/S0907444912001308>
- 700 Baldwin T, Baldwin S, Klos K, et al (2019) Deletion of the benzoxazinoid
701 detoxification gene NAT1 in *Fusarium graminearum* reduces deoxynivalenol in
702 spring wheat. *PLoS One* 14:e0214230.
703 <https://doi.org/10.1371/journal.pone.0214230>
- 704 Barnes JP, Putnam AR (1987) Role of benzoxazinones in allelopathy by rye (*Secale*
705 *cereale* L.). *J Chem Ecol* 13:889–906. <https://doi.org/10.1007/BF01020168>
- 706 Bednarek P (2012) Chemical warfare or modulators of defence responses - the
707 function of secondary metabolites in plant immunity. *Curr. Opin. Plant Biol.*
708 15:407–414
- 709 Beilsten-Edmands J, Winter G, Gildea R, et al (2020) Scaling diffraction data in the
710 DIALS software package: algorithms and new approaches for multi-crystal
711 scaling. *Acta Crystallogr Sect D, Struct Biol* 76:385–399.
712 <https://doi.org/10.1107/S2059798320003198>
- 713 Bhakta S, Besra GS, Upton AM, et al (2004) Arylamine N-acetyltransferase is
714 required for synthesis of mycolic acids and complex lipids in *Mycobacterium*
715 *bovis* BCG and represents a novel drug target. *J Exp Med* 199:1191–1199.
716 <https://doi.org/10.1084/jem.20031956>
- 717 BOHIDAR K, WRATTEN SD, NIEMEYER HM (1986) Effects of hydroxamic acids
718 on the resistance of wheat to the aphid *Sitobion avenae*. *Ann Appl Biol* 109:193–
719 198. <https://doi.org/10.1111/j.1744-7348.1986.tb03199.x>
- 720 Boukouvala S, Fakis G (2005) Arylamine N-acetyltransferases: what we learn from
721 genes and genomes. *Drug Metab Rev* 37:511–564.
722 <https://doi.org/10.1080/03602530500251204>
- 723 Butcher NJ, Minchin RF (2012) Arylamine N-acetyltransferase 1: a novel drug target
724 in cancer development. *Pharmacol Rev* 64:147–165.
725 <https://doi.org/10.1124/pr.110.004275>
- 726 Chan Ho Tong L, Dairou J, Bui LC, et al (2015) Screen for soil fungi highly resistant
727 to dichloroaniline uncovers mostly *Fusarium* species. *Fungal Genet Biol* 81:82–
728 87. <https://doi.org/10.1016/j.fgb.2015.05.011>
- 729 Cocaign A, Bui LC, Silar P, et al (2013) Biotransformation of trichoderma spp. and
730 their tolerance to aromatic amines, a major class of pollutants. *Appl Environ*
731 *Microbiol* 79:4719–4726. <https://doi.org/10.1128/AEM.00989-13>
- 732 Cocaign A, Kubiak X, Xu X, et al (2014) Structural and functional characterization of
733 an arylamine N-acetyltransferase from the pathogen *Mycobacterium abscessus*:
734 differences from other mycobacterial isoforms and implications for selective
735 inhibition. *Acta Crystallogr D Biol Crystallogr* 70:3066–3079.

736 <https://doi.org/10.1107/S1399004714021282>

737 Comeau SR, Gatchell DW, Vajda S, Camacho CJ (2004) ClusPro: a fully automated
738 algorithm for protein-protein docking. *Nucleic Acids Res* 32:W96-9.
739 <https://doi.org/10.1093/nar/gkh354>

740 Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of
741 Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501.
742 <https://doi.org/10.1107/S0907444910007493>

743 Fiser A, Sali A (2003) ModLoop: automated modeling of loops in protein structures.
744 *Bioinformatics* 19:2500–2501. <https://doi.org/10.1093/bioinformatics/btg362>

745 Floss HG, Yu TW (1999) Lessons from the rifamycin biosynthetic gene cluster. *Curr*
746 *Opin Chem Biol* 3:592–597

747 Forli S, Huey R, Pique ME, et al (2016) Computational protein-ligand docking and
748 virtual drug screening with the AutoDock suite. *Nat Protoc* 11:905–919.
749 <https://doi.org/10.1038/nprot.2016.051>

750 Frey M, Chomet P, Glawischnig E, et al (1997) Analysis of a chemical plant defense
751 mechanism in grasses. *Science* (80-) 277:696–699.
752 <https://doi.org/10.1126/science.277.5326.696>

753 Frey M, Schullehner K, Dick R, et al (2009) Benzoxazinoid biosynthesis, a model for
754 evolution of secondary metabolic pathways in plants. *Phytochemistry* 70:1645–
755 1651

756 Fullam E, Westwood IM, Anderton MC, et al (2008) Divergence of cofactor
757 recognition across evolution: coenzyme A binding in a prokaryotic arylamine N-
758 acetyltransferase. *J Mol Biol* 375:178–191.
759 <https://doi.org/10.1016/j.jmb.2007.10.019>

760 Garefalaki V, Kontomina E, Ioannidis C, et al (2019) The actinobacterium
761 *Tsukamurella paurometabola* has a functionally divergent arylamine N-
762 acetyltransferase (NAT) homolog. *World J Microbiol Biotechnol* 35:174.
763 <https://doi.org/10.1007/s11274-019-2755-1>

764 Garefalaki V, Papavergi M-G, Savvidou O, et al (2021) Comparative Investigation of
765 15 Xenobiotic-Metabolizing N-Acetyltransferase (NAT) Homologs from
766 Bacteria. *Appl Environ Microbiol* 87:e0081921.
767 <https://doi.org/10.1128/AEM.00819-21>

768 Gierl A, Frey M (2001) Evolution of benzoxazinone biosynthesis and indole
769 production in maize. *Planta* 213:493–498

770 Glenn AE, Bacon CW (2009) FDB2 encodes a member of the arylamine N-
771 acetyltransferase family and is necessary for biotransformation of
772 benzoxazolinones by *Fusarium verticillioides*. *J Appl Microbiol* 107:657–671.
773 <https://doi.org/10.1111/j.1365-2672.2009.04246.x>

774 Glenn AE, Davis CB, Gao M, et al (2016) Two horizontally transferred xenobiotic
775 resistance gene clusters associated with detoxification of benzoxazolinones by
776 *Fusarium* species. *PLoS One* 11:. <https://doi.org/10.1371/journal.pone.0147486>

777 Glenn AE, Gold SE, Bacon CW (2002) Fdb1 and Fdb2, *Fusarium verticillioides* loci
778 necessary for detoxification of preformed antimicrobials from corn. *Mol Plant-
779 Microbe Interact* 15:91–101. <https://doi.org/10.1094/MPMI.2002.15.2.91>

780 Glenn AE, Hinton DM, Yates IE, Bacon CW (2001) Detoxification of corn
781 antimicrobial compounds as the basis for isolating *Fusarium verticillioides* and
782 some other *Fusarium* species from corn. *Appl Environ Microbiol* 67:2973–2981.
783 <https://doi.org/10.1128/AEM.67.7.2973-2981.2001>

784 Glenn AE, Karagianni EP, Ulndreaj A, Boukouvala S (2010) Comparative genomic
785 and phylogenetic investigation of the xenobiotic metabolizing arylamine N-

786 acetyltransferase enzyme family. FEBS Lett 584:3158–3164.
787 <https://doi.org/10.1016/j.febslet.2010.05.063>

788 Glenn AE, Meredith FI, Morrison WH, Bacon CW (2003) Identification of
789 intermediate and branch metabolites resulting from biotransformation of 2-
790 benzoxazolinone by *Fusarium verticillioides*. Appl Environ Microbiol 69:3165–
791 9. <https://doi.org/10.1128/aem.69.6.3165-3169.2003>

792 Gorrec F (2009) The MORPHEUS protein crystallization screen. J Appl Crystallogr
793 42:1035–1042. <https://doi.org/10.1107/S0021889809042022>

794 Grant DM (1993) Molecular genetics of the N-acetyltransferases. Pharmacogenetics
795 3:45–50. <https://doi.org/10.1097/00008571-199302000-00005>

796 Hashimoto Y, Shudo K (1996) Chemistry of biologically active benzoxazinoids.
797 Phytochemistry 43:551–559

798 Hein DW, Boukouvala S, Grant DM, et al (2008) Changes in consensus arylamine N-
799 acetyltransferase gene nomenclature. Pharmacogenet Genomics 18:367–368.
800 <https://doi.org/10.1097/FPC.0b013e3282f60db0>

801 Jiménez-García B, Pons C, Fernández-Recio J (2013) pyDockWEB: a web server for
802 rigid-body protein–protein docking using electrostatics and desolvation scoring.
803 Bioinformatics 29:1698–1699. <https://doi.org/10.1093/bioinformatics/btt262>

804 Karagianni EP, Kontomina E, Davis B, et al (2015) Homologues of xenobiotic
805 metabolizing N-acetyltransferases in plant-associated fungi: Novel functions for
806 an old enzyme family. Sci Rep 5:12900. <https://doi.org/10.1038/srep12900>

807 Keegan RM, McNicholas SJ, Thomas JMH, et al (2018) Recent developments in
808 MrBUMP: better search-model preparation, graphical interaction with search
809 models, and solution improvement and assessment. Acta Crystallogr Sect D,
810 Struct Biol 74:167–182. <https://doi.org/10.1107/S2059798318003455>

811 Kettle AJ, Batley J, Benfield AH, et al (2015) Degradation of the benzoxazolinone
812 class of phytoalexins is important for virulence of *Fusarium pseudograminearum*
813 towards wheat. Mol Plant Pathol 16:946–962. <https://doi.org/10.1111/mpp.12250>

814 Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from
815 crystalline state. J Mol Biol 372:774–797.
816 <https://doi.org/10.1016/j.jmb.2007.05.022>

817 Kubiak X, Dairou J, Dupret J-M, Rodrigues-Lima F (2013a) Crystal structure of
818 arylamine N-acetyltransferases: insights into the mechanisms of action and
819 substrate selectivity. Expert Opin Drug Metab Toxicol 9:349–362.
820 <https://doi.org/10.1517/17425255.2013.742505>

821 Kubiak X, Li de la Sierra-Gallay I, Chaffotte AF, et al (2013b) Structural and
822 biochemical characterization of an active arylamine N-acetyltransferase
823 possessing a non-canonical Cys-His-Glu catalytic triad. J Biol Chem 288:22493–
824 22505. <https://doi.org/10.1074/jbc.M113.468595>

825 Laskowski RA, Swindells MB (2011) LigPlot+: multiple ligand-protein interaction
826 diagrams for drug discovery. J Chem Inf Model 51:2778–2786.
827 <https://doi.org/10.1021/ci200227u>

828 Laurieri N, Sim E (2018) Arylamine N-Acetyltransferases in Health and Disease.
829 WORLD SCIENTIFIC

830 Martins M, Dairou J, Rodrigues-Lima F, et al (2010) Insights into the phylogeny or
831 arylamine N-acetyltransferases in fungi. J Mol Evol 71:141–152.
832 <https://doi.org/10.1007/s00239-010-9371-x>

833 Martins M, Pluvinage B, Li de la Sierra-Gallay I, et al (2008) Functional and
834 structural characterization of the arylamine N-acetyltransferase from the
835 opportunistic pathogen *Nocardia farcinica*. J Mol Biol 383:549–560.

836 <https://doi.org/10.1016/j.jmb.2008.08.035>

837 Martins M, Rodrigues-Lima F, Dairou J, et al (2009) An acetyltransferase conferring
838 tolerance to toxic aromatic amine chemicals: Molecular and functional studies. *J*
839 *Biol Chem* 284:18726–18733. <https://doi.org/10.1074/jbc.M109.015230>

840 McCoy AJ, Grosse-Kunstleve RW, Adams PD, et al (2007) Phaser crystallographic
841 software. *J Appl Crystallogr* 40:658–674.
842 <https://doi.org/10.1107/S0021889807021206>

843 McDonagh EM, Boukouvala S, Aklillu E, et al (2014) PharmGKB summary: very
844 important pharmacogene information for N-acetyltransferase 2. *Pharmacogenet*
845 *Genomics* 24:409–425. <https://doi.org/10.1097/FPC.0000000000000062>

846 McNicholas S, Potterton E, Wilson KS, Noble MEM (2011) Presenting your
847 structures: the CCP4mg molecular-graphics software. *Acta Crystallogr D Biol*
848 *Crystallogr* 67:386–394. <https://doi.org/10.1107/S0907444911007281>

849 Murshudov GN, Skubák P, Lebedev AA, et al (2011) REFMAC5 for the refinement
850 of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr*
851 *67:355–367*. <https://doi.org/10.1107/S0907444911001314>

852 Nebert DW (1997) Polymorphisms in drug-metabolizing enzymes: what is their
853 clinical relevance and why do they exist? *Am. J. Hum. Genet.* 60:265–271

854 Nebert DW, Dieter MZ (2000) The evolution of drug metabolism. *Pharmacology*
855 *61:124–135*. <https://doi.org/10.1159/000028393>

856 Niemeyer HM (2009) Hydroxamic Acids Derived from 2-Hydroxy-2 *H* -1,4-
857 Benzoxazin-3(4 *H*)-one: Key Defense Chemicals of Cereals. *J Agric Food Chem*
858 *57:1677–1696*. <https://doi.org/10.1021/jf8034034>

859 Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for fast and
860 accurate multiple sequence alignment. *J Mol Biol* 302:205–217.
861 <https://doi.org/10.1006/jmbi.2000.4042>

862 Pettersen EF, Goddard TD, Huang CC, et al (2004) UCSF Chimera--a visualization
863 system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.
864 <https://doi.org/10.1002/jcc.20084>

865 Pluvinage B, Dairou J, Possot OM, et al (2007) Cloning and molecular
866 characterization of three arylamine N-acetyltransferase genes from *Bacillus*
867 *anthracis*: identification of unusual enzymatic properties and their contribution to
868 sulfamethoxazole resistance. *Biochemistry* 46:7069–7078.
869 <https://doi.org/10.1021/bi700351w>

870 Pluvinage B, Li de la Sierra-Gallay I, Kubiak X, et al (2011) The *Bacillus anthracis*
871 arylamine N-acetyltransferase ((BACAN)NAT1) that inactivates
872 sulfamethoxazole, reveals unusual structural features compared with the other
873 NAT isoenzymes. *FEBS Lett* 585:3947–3952.
874 <https://doi.org/10.1016/j.febslet.2011.10.041>

875 Riddle B, Jencks WP (1971) Acetyl-coenzyme A: arylamine N-acetyltransferase.
876 Role of the acetyl-enzyme intermediate and the effects of substituents on the
877 rate. *J Biol Chem* 246:3250–3258

878 Robens J, Cardwell K (2003) The Costs of Mycotoxin Management to the USA:
879 Management of Aflatoxins in the United States. *J Toxicol Toxin Rev* 22:139–
880 152. <https://doi.org/10.1081/TXR-120024089>

881 Robert X, Gouet P (2014) Deciphering key features in protein structures with the new
882 ENDscript server. *Nucleic Acids Res* 42:W320-4.
883 <https://doi.org/10.1093/nar/gku316>

884 Rodrigues-Lima F, Dairou J, Diaz CL, et al (2006) Cloning, functional expression and
885 characterization of *Mesorhizobium loti* arylamine N-acetyltransferases: rhizobial

886 symbiosis supplies leguminous plants with the xenobiotic N-acetylation
887 pathway. *Mol Microbiol* 60:505–512. <https://doi.org/10.1111/j.1365->
888 2958.2006.05114.x

889 Sandy J, Mushtaq A, Holton SJ, et al (2005) Investigation of the catalytic triad of
890 arylamine N-acetyltransferases: essential residues required for acetyl transfer to
891 arylamines. *Biochem J* 390:115–123. <https://doi.org/10.1042/BJ20050277>

892 Sandy J, Mushtaq A, Kawamura A, et al (2002) The structure of arylamine N-
893 acetyltransferase from *Mycobacterium smegmatis*--an enzyme which inactivates
894 the anti-tubercular drug, isoniazid. *J Mol Biol* 318:1071–1083.
895 [https://doi.org/10.1016/S0022-2836\(02\)00141-9](https://doi.org/10.1016/S0022-2836(02)00141-9)

896 Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ (2005) PatchDock and
897 SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res*
898 33:W363-7. <https://doi.org/10.1093/nar/gki481>

899 Sicker D, Schulz M (2002) Benzoxazinones in plants: Occurrence, synthetic access,
900 and biological activity. *Stud Nat Prod Chem* 27:185–232.
901 [https://doi.org/10.1016/S1572-5995\(02\)80037-0](https://doi.org/10.1016/S1572-5995(02)80037-0)

902 Sim E, Walters K, Boukouvala S (2008) Arylamine N-acetyltransferases: from
903 structure to function. *Drug Metab Rev* 40:479–510.
904 <https://doi.org/10.1080/03602530802186603>

905 Sinclair JC, Sandy J, Delgoda R, et al (2000) Structure of arylamine N-
906 acetyltransferase reveals a catalytic triad. *Nat Struct Biol* 7:560–564.
907 <https://doi.org/10.1038/76783>

908 Suzuki H, Ohnishi Y, Horinouchi S (2007) Arylamine N-acetyltransferase responsible
909 for acetylation of 2-aminophenols in *Streptomyces griseus*. *J Bacteriol*
910 189:2155–2159. <https://doi.org/10.1128/JB.01708-06>

911 Tabaglio V, Gavazzi C, Schulz M, Marocco A (2008) Alternative weed control using
912 the allelopathic effect of natural benzoxazinoids from rye mulch. *Agron Sustain*
913 *Dev* 28:397–401. <https://doi.org/10.1051/agro:2008004>

914 Takenaka S, Cheng M, Mulyono, et al (2009) Gene cloning and characterization of
915 arylamine N-acetyltransferase from *Bacillus cereus* strain 10-L-2. *J Biosci*
916 *Bioeng* 107:27–32. <https://doi.org/10.1016/j.jbiosc.2008.09.012>

917 Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of
918 docking with a new scoring function, efficient optimization, and multithreading.
919 *J Comput Chem* 31:455–461. <https://doi.org/10.1002/jcc.21334>

920 Tsirka T, Boukouvala S, Agianian B, Fakis G (2014) Polymorphism p.Val231Ile
921 alters substrate selectivity of drug-metabolizing arylamine N-acetyltransferase 2
922 (NAT2) isoenzyme of rhesus macaque and human. *Gene* 536:65–73.
923 <https://doi.org/10.1016/j.gene.2013.11.085>

924 Tsirka T, Konstantopoulou M, Sabbagh A, et al (2018) Comparative analysis of
925 xenobiotic metabolising N-acetyltransferases from ten non-human primates as in
926 vitro models of human homologues. *Sci Rep* 8:9759.
927 <https://doi.org/10.1038/s41598-018-28094-6>

928 Vagena E, Fakis G, Boukouvala S (2008) Arylamine N-acetyltransferases in
929 prokaryotic and eukaryotic genomes: a survey of public databases. *Curr Drug*
930 *Metab* 9:628–660

931 Watanabe M, Sofuni T, Nohmi T (1992) Involvement of Cys69 residue in the
932 catalytic mechanism of N-hydroxyarylamine O-acetyltransferase of *Salmonella*
933 *typhimurium*. Sequence similarity at the amino acid level suggests a common
934 catalytic mechanism of acetyltransferase for *S. typhimurium* and high. *J Biol*
935 *Chem* 267:8429–8436

936 Waterhouse A, Bertoni M, Bienert S, et al (2018) SWISS-MODEL: homology
937 modelling of protein structures and complexes. *Nucleic Acids Res* 46:W296–
938 W303. <https://doi.org/10.1093/nar/gky427>

939 Weber WW, Hein DW (1985) N-acetylation pharmacogenetics. *Pharmacol Rev*
940 37:25–79

941 Westwood IM, Holton SJ, Rodrigues-Lima F, et al (2005) Expression, purification,
942 characterization and structure of *Pseudomonas aeruginosa* arylamine N-
943 acetyltransferase. *Biochem J* 385:605–612. <https://doi.org/10.1042/BJ20041330>

944 Winn MD, Ballard CC, Cowtan KD, et al (2011) Overview of the CCP4 suite and
945 current developments. *Acta Crystallogr D Biol Crystallogr* 67:235–242.
946 <https://doi.org/10.1107/S0907444910045749>

947 Woodward MD, Corcuera LJ, Helgeson JP, Upper CD (1978) Decomposition of 2,4-
948 Dihydroxy-7-methoxy-2 H -1,4-benzoxazin-3(4 H)-one in Aqueous Solutions .
949 *Plant Physiol* 61:796–802. <https://doi.org/10.1104/pp.61.5.796>

950 Wu F (2004) Mycotoxin risk assessment for the purpose of setting international
951 regulatory standards. *Environ Sci Technol* 38:4049–4055.
952 <https://doi.org/10.1021/es035353n>

953 Wu H, Dombrovsky L, Tempel W, et al (2007) Structural basis of substrate-binding
954 specificity of human arylamine N-acetyltransferases. *J Biol Chem* 282:30189–
955 30197. <https://doi.org/10.1074/jbc.M704138200>

956 Xu X, Li de la Sierra-Gallay I, Kubiak X, et al (2015) Insight into cofactor
957 recognition in arylamine N-acetyltransferase enzymes: structure of
958 *Mesorhizobium loti* arylamine N-acetyltransferase in complex with coenzyme A.
959 *Acta Crystallogr D Biol Crystallogr* 71:266–273.
960 <https://doi.org/10.1107/S139900471402522X>

961 Zasada IA, Meyer SLF, Halbrendt JM, Rice C (2005) Activity of hydroxamic acids
962 from *Secale cereale* against the plant-parasitic nematodes *Meloidogyne incognita*
963 and *Xiphinema americanum*. *Phytopathology* 95:1116–1121.
964 <https://doi.org/10.1094/PHYTO-95-1116>

965 Zúñiga GE, Argandoña VH, Niemeyer HM, Corcuera LJ (1983) Hydroxamic acid
966 content in wild and cultivated gramineae. *Phytochemistry* 22:2665–2668.
967 [https://doi.org/10.1016/S0031-9422\(00\)97669-6](https://doi.org/10.1016/S0031-9422(00)97669-6)

968 **Supporting Information**

969 **Supplementary materials and methods**

970 **Suppl. Fig. S1:** Purification of *Fusarium verticillioides* (GIMBO)NAT1 recombinant
971 protein for crystallography.

972 **Suppl. Fig. S2:** Determination of the molecular weight of recombinant
973 (GIBMO)NAT1 protein.

974 **Suppl. Fig. S3:** Comparison of (GIBMO)NAT1 with one bacterial and one human
975 NAT.

976 **Suppl. Fig. S4:** Docking of various donor substrates to the (GIBMO)NAT1
977 monomer.

978 **Suppl. Fig. S5:** LigPlot+ analyses for different donor substrates docked to the
979 (GIBMO)NAT1 monomer.

980 **Suppl. Fig. S6:** Structural alignment of the amino acid sequences of 13 fungal NAT
981 homologues functionally investigated before.

982 **Suppl. Fig. S7:** Comparison of thirteen fungal NAT homologues functionally
983 investigated before.

984 **Suppl. Fig. S8:** Phylogenetic analysis of fungal NAT sequences.

985 **Table 1: Statistics for crystallographic data collection and structure refinement**

Data collection statistics	
Wavelength (Å)	0.9393
Resolution range (Å)	87.18-1.80 (1.83-1.80)
Space group	P1
<i>Unit cell dimensions</i>	
a, b, c (Å)	72.48, 76.77, 92.38
α , β , γ (°)	76.24, 73.08, 72.50
Unique reflections	160393 (6978)
Completeness (%)	96.5 (84.5)
Mean I/sigma(I)	6.4 (1.1)
Multiplicity	1.9 (1.5)
Rmerge (%)	0.039 (0.633)
Rpim (%)	0.039 (0.633)
CC ½	0.998 (0.644)
Wilson B-factor	23.58
Structure refinement	
Resolution (Å)	87.3-1.8
R _{work} / R _{free} (%)	17.84 / 22.16
<i>No. atoms</i>	
Protein	11430
Ligands	100
Water	1415
<i>RMSD</i>	
Bond lengths (Å)	0.007
Bond angles (°)	0.829
Average B-factors	31.9
<i>Ramachandran plot</i>	
Favoured (%)	97.56
Allowed (%)	2.4
Outliers (%)	0.3

986

987 **Figure legends**

988 **Fig. 1:** Production, enzymatic assay and crystallization of (GIBMO)NAT1
989 recombinant protein of *Fusarium verticillioides*. A: Pure recombinant protein (10
990 mg/ml, 15 μ l loaded in lane 1) was recovered after affinity, anion exchange and gel
991 filtration chromatography. A protein band with the expected size is visible by SDS-
992 PAGE/Coomassie blue staining. Lane M is the High-Range Rainbow MW protein
993 marker (GE Healthcare). B: Enzymatic activity assay with four different combinations
994 of donor (malonyl- or acetyl-CoA) and acceptor (5AS or PABA) substrates. Pure
995 recombinant protein was assayed in triplicate and the generated enzymatic curves are
996 shown. Specific activities (nmol of produced CoA per minute per mg of protein) were
997 calculated for 5 min time. C: Representative protein crystals, grown (one week, 4 $^{\circ}$ C)
998 in well D5 (0.12 M alcohols, 0.1 M sodium HEPES/MOPS buffer pH 7.5 and 30%
999 PEGMME 550/PEG 20K precipitant stock) of the commercial Morpheus
1000 crystallographic screen block (Molecular Dimensions).

1001 **Fig. 2:** Molecular structure of (GIBMO)NAT1 protein. A: Cartoon representation of
1002 the protein monomer with the typical NAT-fold organized in domain I (blue), II
1003 (raspberry) and III (olive green), with the last two domains connected via an
1004 interdomain region (grey). B: Alignment of (GIBMO)NAT1 (magenta) with
1005 (PSEAE)NAT1 (lime green) of *Pseudomonas aeruginosa* (PDB ID: 1W4T). C: The
1006 crystallized homodimeric structure of (GIBMO)NAT1, formed by monomers shown
1007 in magenta and blue. Coloured pink and cyan, respectively, are the protein surfaces
1008 forming the interface between the two interacting monomers. Specific molecular
1009 contacts are indicated with amino acids coloured warm pink and sky blue,
1010 respectively. D: Surface representation of the homotetrameric arrangement in the

1011 crystal asymmetric unit, formed by two pairs of tightly interacting monomers (each
1012 dimer consists of one monomer in blue and one monomer in magenta colour). In A-C,
1013 the catalytic triad residues are shown in stick format, yellow for (GIMBO)NAT1 and
1014 orange for (PSEAE)NAT1. Protein *N*- and *C*-termini are also indicated.

1015 **Fig. 3:** Surface representations of (GIBMO)NAT1, illustrating access to the catalytic
1016 core of the enzyme. The top panel shows the "bridge-like" structure, coloured light
1017 blue, separating the two "tunnel-like" entries leading to the active site (the catalytic
1018 triad is coloured yellow). In the dimeric form (middle panel), one entry of each
1019 monomer is tightly sealed by the protruding *N*-terminus of the other monomer, while
1020 the second entry remains uncovered. The bottom panel shows a partial view of
1021 (GIBMO)NAT1 surface aligned with bound CoA conformations previously
1022 determined via co-crystallization of the ligand with one human (NAT2; PDB ID:
1023 2PFR) and two bacterial (PDB IDs: 4NV7 of *Mesorhizobium loti* NAT1 and 2VFC
1024 of *Mycobacterium marinum* NAT1) proteins. The binding conformation of CoA from
1025 human NAT2 (cyan) aligns through the "covered entry" of (GIBMO)NAT1 (bottom-
1026 left), while the binding conformations of CoA from the two bacterial NATs (green
1027 and pink) align through the "uncovered entry" of (GIBMO)NAT1 (bottom-right).

1028 **Fig. 4:** Docking of CoA and 2AP to the (GIBMO)NAT1 monomer. The top panel
1029 shows detailed views of two conformations of CoA (left) and one conformation of
1030 2AP (right), docked to the active site of (GIBMO)NAT1 protein. The bottom panel
1031 shows the results of LigPlot⁺ analysis for CoA (the cyan conformation in the top
1032 panel) and 2AP. The ligands form hydrogen bonds or hydrophobic interactions with
1033 residues labelled in green or black font, respectively.

1034 **Fig. 5:** Docking experiments performed with the (GIBMO)NAT1 dimer. Surface
1035 representation of dimeric (GIBMO)NAT1 (A), with different acyl-CoA compounds
1036 docked to the blue monomer via the single uncovered "tunnel-like" entry. CoA (B),
1037 acetyl-CoA (C), propionyl-CoA (D), malonyl-CoA (E) and succinyl-CoA (F) were
1038 used as ligands, and their orientation relative to the catalytic triad residues (shown in
1039 yellow stick format) is demonstrated. Red boxes enclose the enzyme-ligand reactive
1040 moieties, with distances between them shown in Å.

1041 **Fig. 6:** Interaction of ligands with the (GIBMO)NAT1 dimer. The best-fitting binding
1042 conformation determined during docking experiments for CoA (A), acetyl-CoA (B),
1043 propionyl-CoA (C), malonyl-CoA (D) or succinyl-CoA (E) is illustrated after
1044 LigPlot⁺ analysis. Docked ligands form hydrogen bonds or hydrophobic interactions
1045 with residues labelled in green or black font, respectively. Red circles indicate
1046 interactions also observed for those ligands, when docked to the monomeric protein.

1047 **Fig. 7:** Results of protein-protein docking experiments simulating dimerization of
1048 (GIBMO)NAT1. Surface representations of homodimers predicted using ClusPro
1049 (left), PatchDock (middle) and pyDock (right) algorithms to generate models that
1050 were then superimposed in PyMOL. In the first experiment, one monomer (grey) was
1051 retained exactly as crystallographically determined and docking of the second
1052 monomer produced a series of models illustrated with different colours. In the
1053 remaining two experiments, more flexibility was allowed for both monomers (grey)
1054 and the *N*-terminal extensions of generated models are illustrated in various colours.
1055 In all three illustrations, the monomer on the right-hand side displays its "tunnel-like"
1056 entries to the active site (bright yellow) and the "bridge-like" structure separating
1057 them (white). The red boxes indicate the interface between docked monomers.

1058 **Fig. 8:** Comparison of the three NAT isoenzymes of *F. verticillioides*. The
1059 crystallographic structure of (GIBMO)NAT1 (blue models, top panel) was used as
1060 template on Swiss-Model to predict the structure of (GIBMO)NAT2 (red models,
1061 middle panel) and (GIBMO)NAT3 (green models, bottom panel). The models on the
1062 left display the side of protein molecules accessible to acyl-CoA substrates (docked).
1063 The models on the right display the opposite view of the protein molecules. Light
1064 colouring of the protein surface indicates areas substantially differentiated between
1065 the three isoenzymes. The residues of those areas are numbered for (GIBMO)NAT1
1066 (top), and a structural sequence alignment with (GIBMO)NAT2 and (GIBMO)NAT3
1067 is provided in Suppl. Fig. S6.

1068 **Fig. 9:** Changes to the thermal denaturation midpoint transition temperature (T_m) of
1069 *F. verticillioides* NAT isoenzymes, upon interaction with different components of the
1070 enzymatic reaction. Shifts in T_m ($\Delta T_m \pm$ standard deviation, in $^{\circ}\text{C}$) were recorded in
1071 duplicate for the *N*-malonyltransferase (GIBMO)NAT1 (A), the *N*-acetyltransferase
1072 (GIBMO)NAT3 (B) and the non-selective (GIBMO)NAT2 (C), upon addition of
1073 ligands separately or in various combinations. All reactions contained 0.5% v/v
1074 DMSO, except for the first one of each set where the protein was assayed alone.

1075 **Fig. 10:** Activity assays for (GIBMO)NAT1 and (GIBMO)NAT3 isoenzymes of *F.*
1076 *verticillioides*. To determine donor substrate selectivity, (GIBMO)NAT1 (A) and
1077 (GIBMO)NAT3 (B) were initially assayed over a lower concentration range (0-300
1078 μM) of malonyl-, acetyl- or propionyl-CoA, using 500 μM of 5AS as acceptor
1079 substrate. The two proteins were then assayed over a higher concentration range (0-
1080 5000 μM) of their preferred donor substrate: (GIBMO)NAT1 (C) was assayed with
1081 malonyl-CoA and 1000 μM 5AS, while (GIBMO)NAT3 (D) was assayed with acetyl-

1082 CoA and 1500 μ M 5AS. In the graphs showing specific activity (μ M of enzymatically
1083 produced CoA per second per μ M of protein) vs. concentration (μ M) of the donor
1084 substrate, each scatter plot indicates the observed (obs) experimental measurements
1085 per assay set, while the corresponding calculated (calc) fitted curves were generated
1086 by non-linear regression analysis to provide optimal Michaelis-Menten curves.
1087 Calculated apparent (app) kinetic values are provided in boxes.

1088 **Fig. 11:** Phylogenetic analysis of 65 NAT sequences, demonstrating distinct lineages
1089 of orthologues in *Fusarium*. The variably coloured lineages are defined by 10
1090 previously annotated and functionally investigated NAT homologues (Glenn et al.
1091 2010; Karagianni et al. 2015), shown in bold ("BN" accession numbers). Grey circles
1092 indicate bootstrap values above 50%.

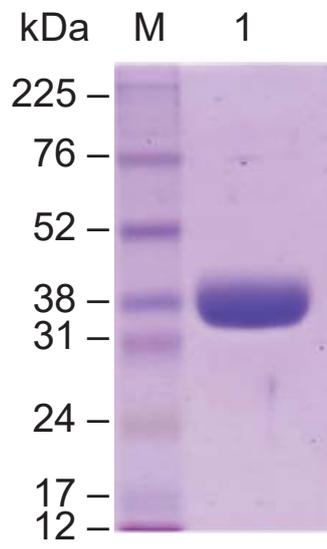
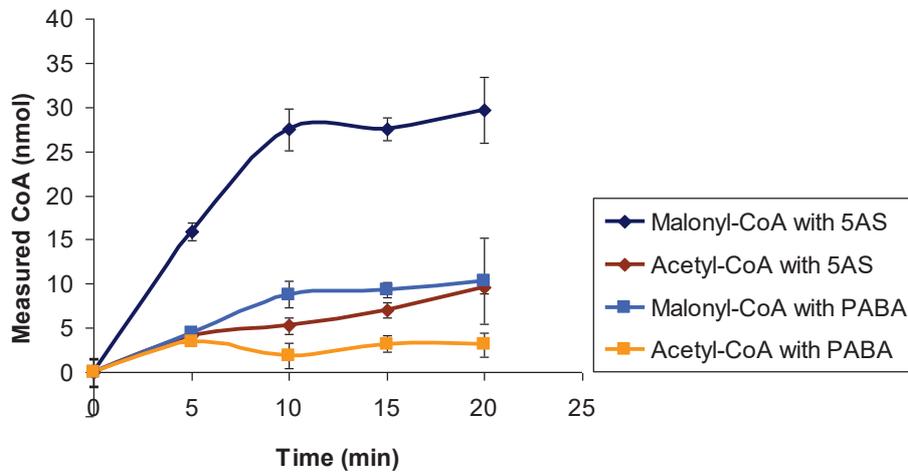
A

Fig. 1

B

Substrate combination	Specific activity (nmol/min/mg)
Malonyl-CoA with 5AS	3189
Acetyl-CoA with 5AS	842
Malonyl-CoA with PABA	880
Acetyl-CoA with PABA	704

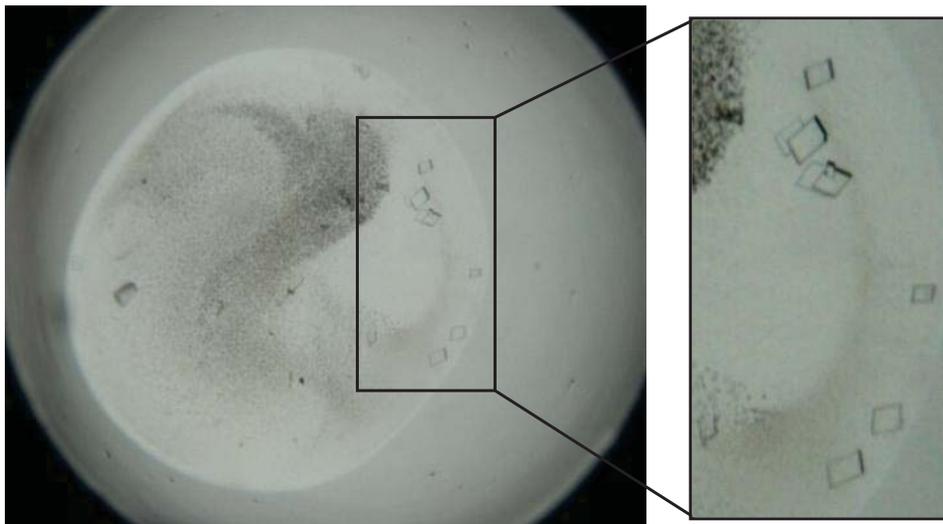
C

Fig. 2

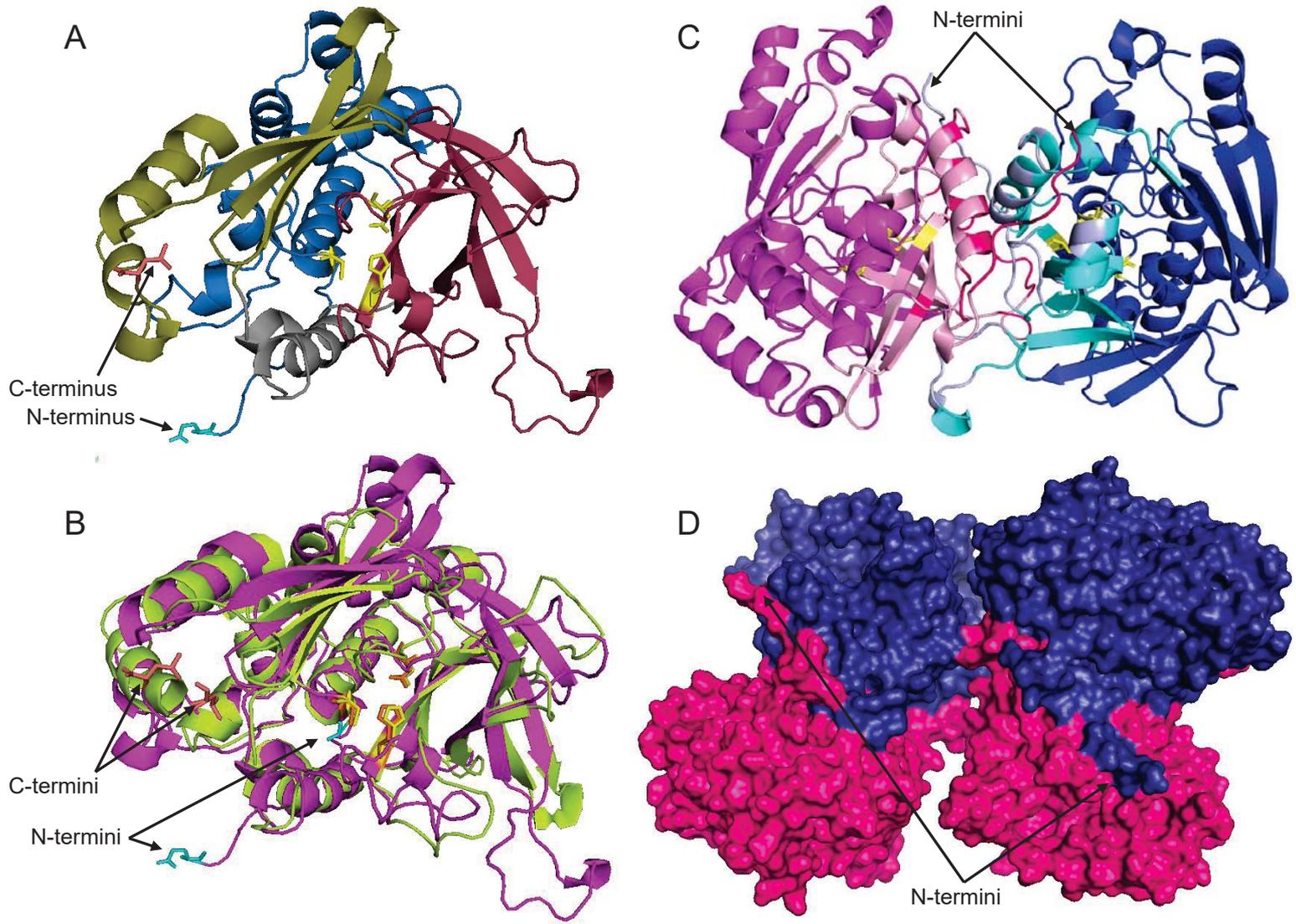
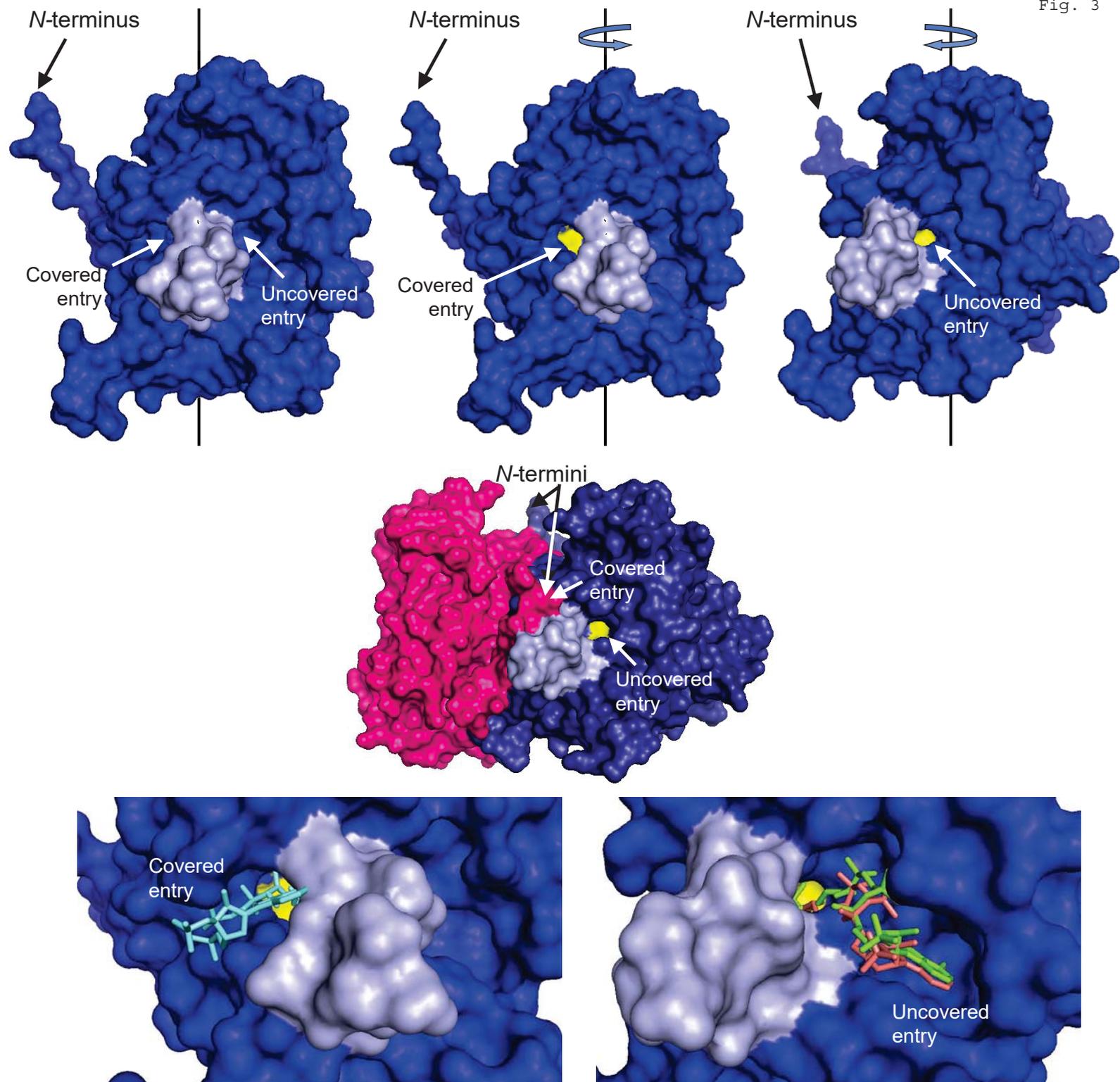


Fig. 3



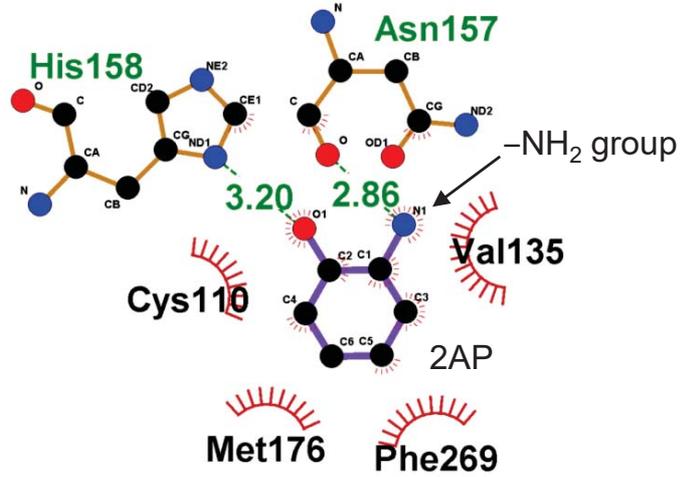
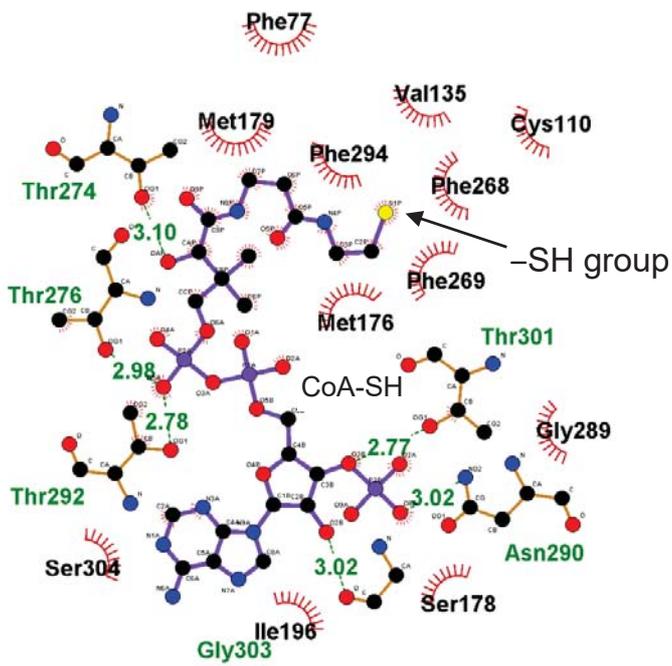
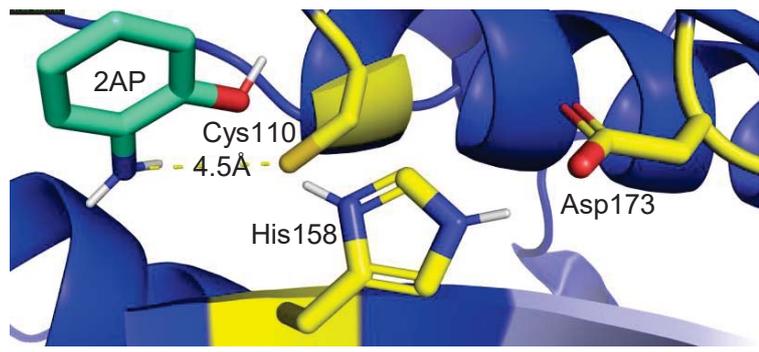
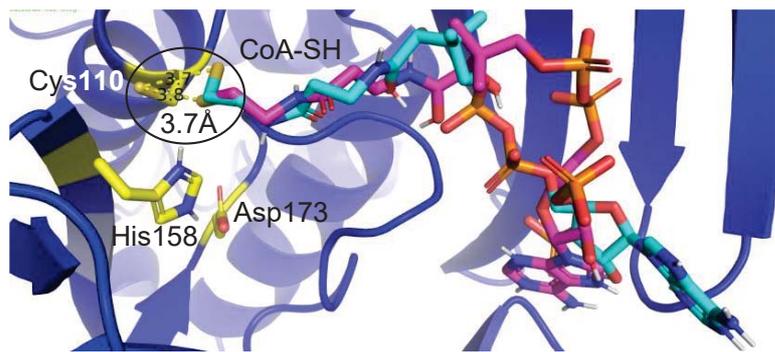
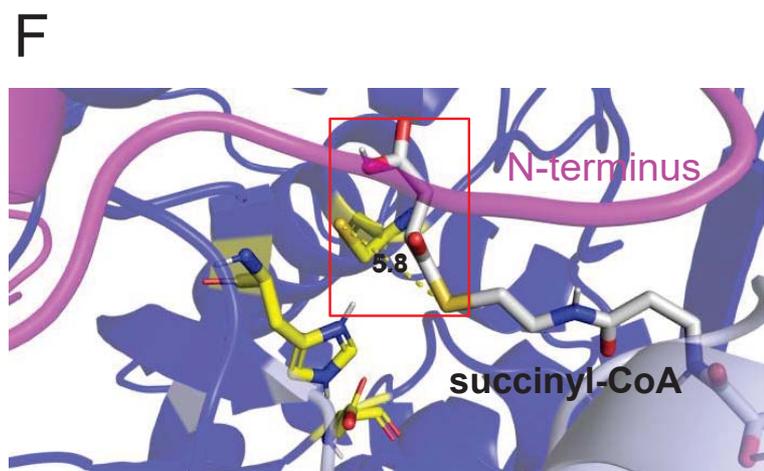
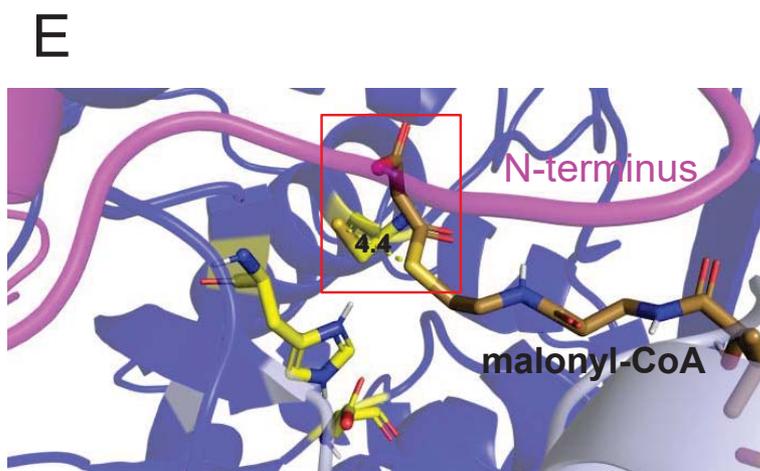
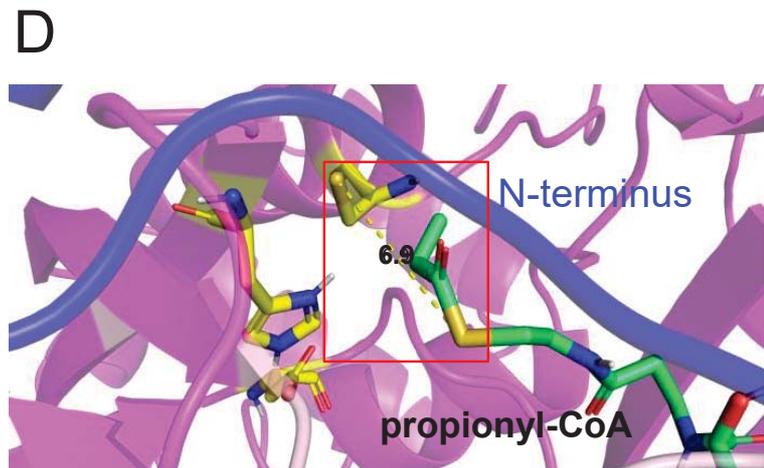
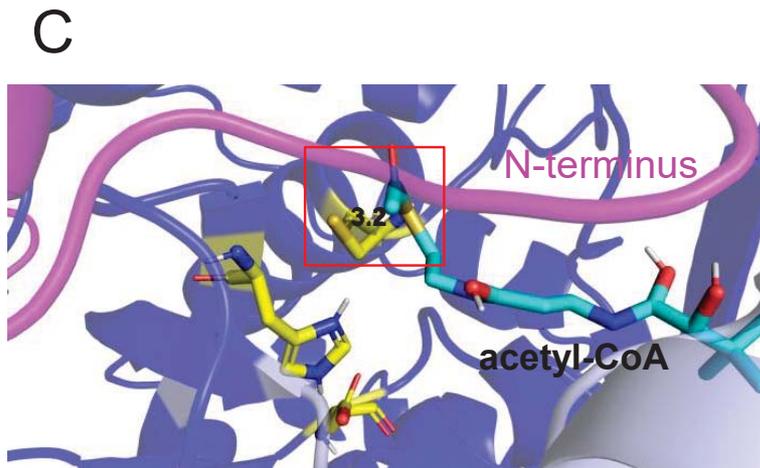
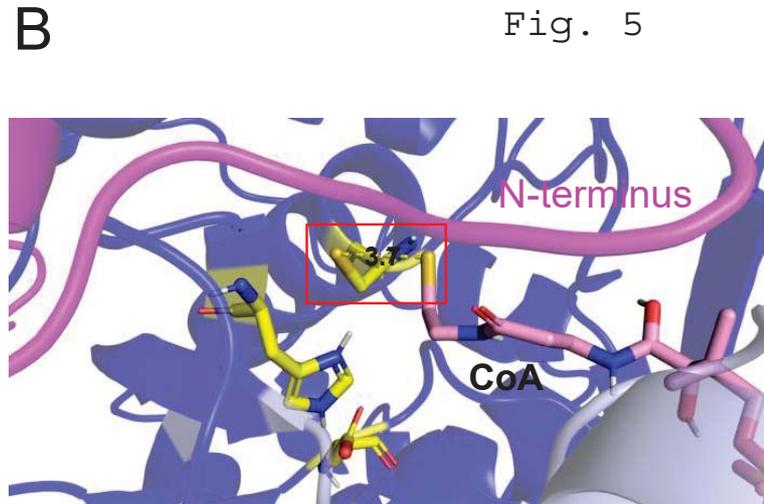
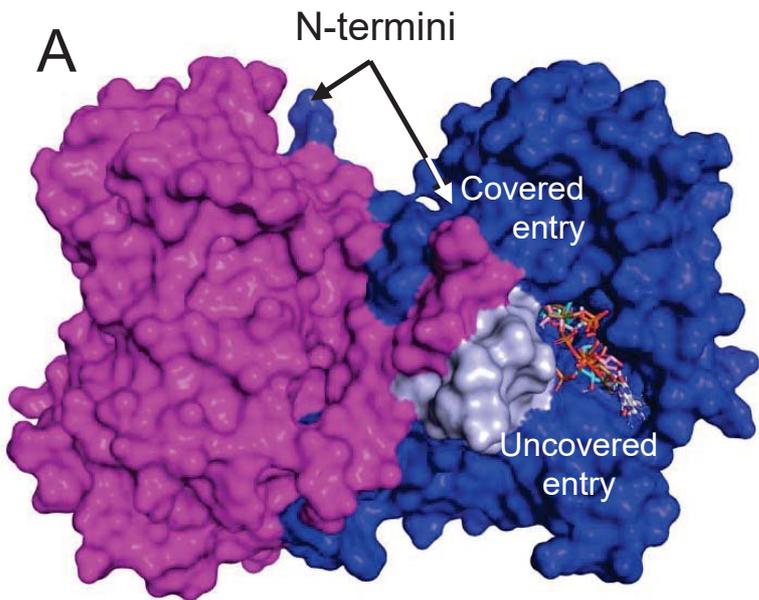


Fig. 4

Fig. 5



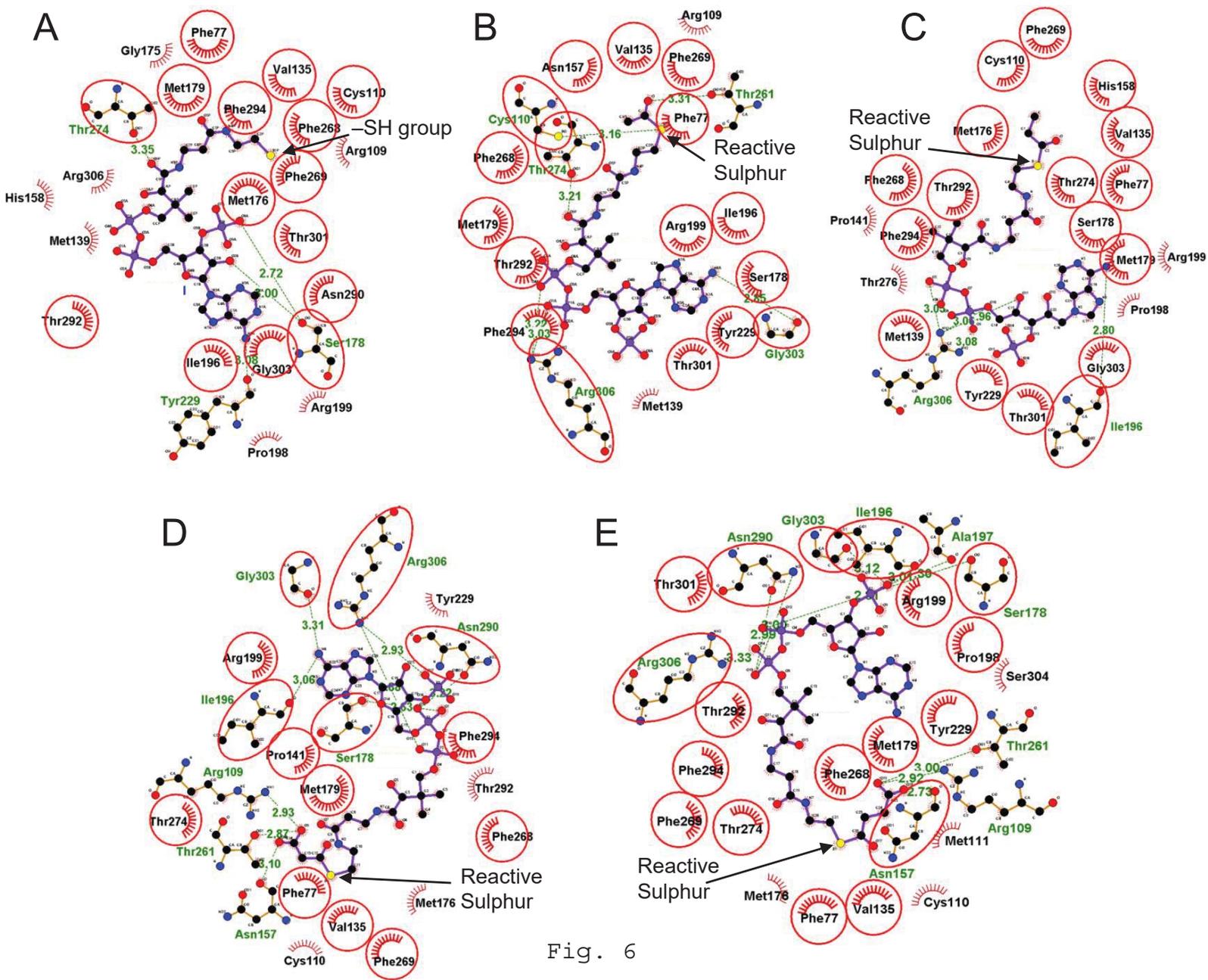


Fig. 6

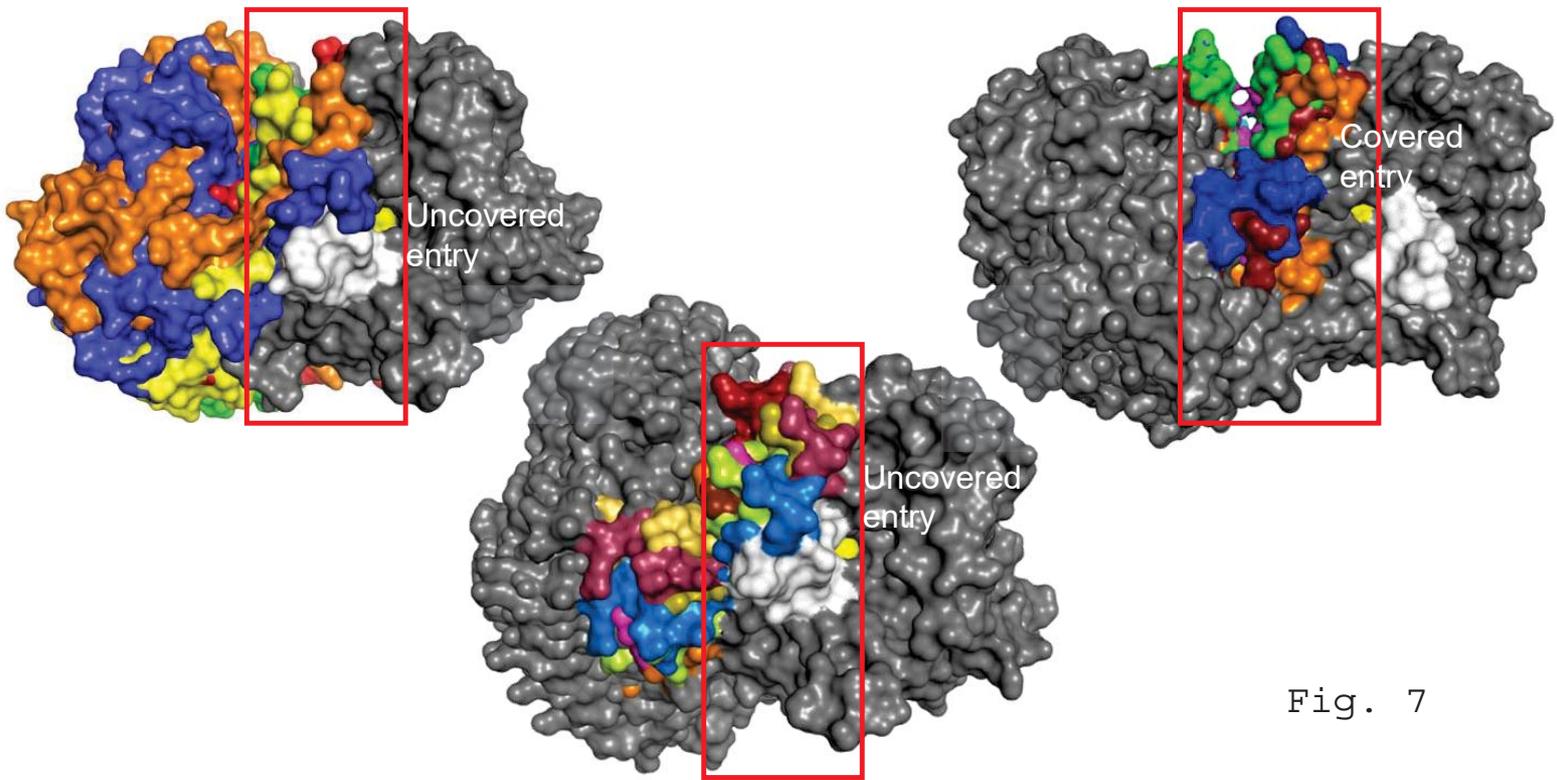
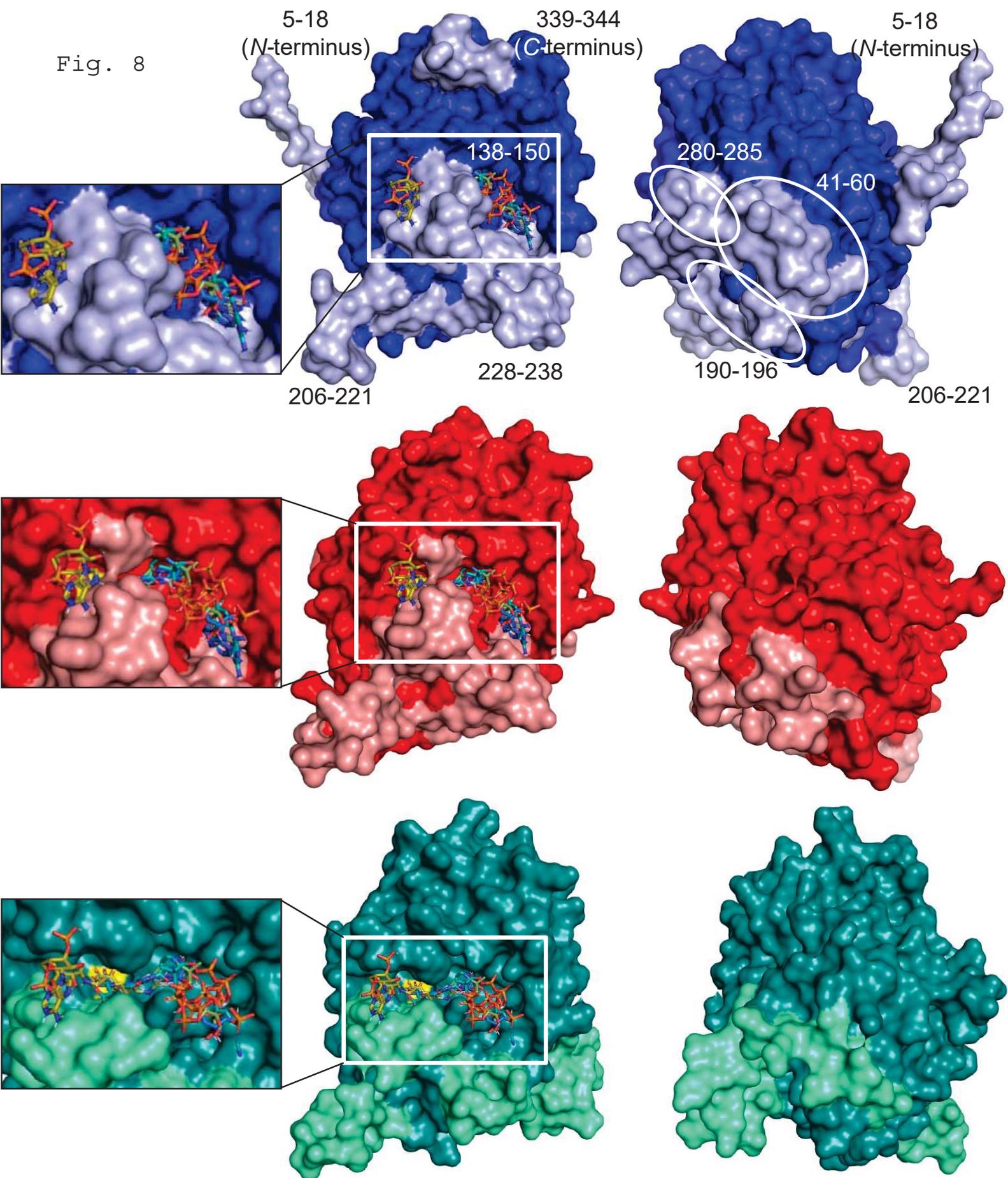
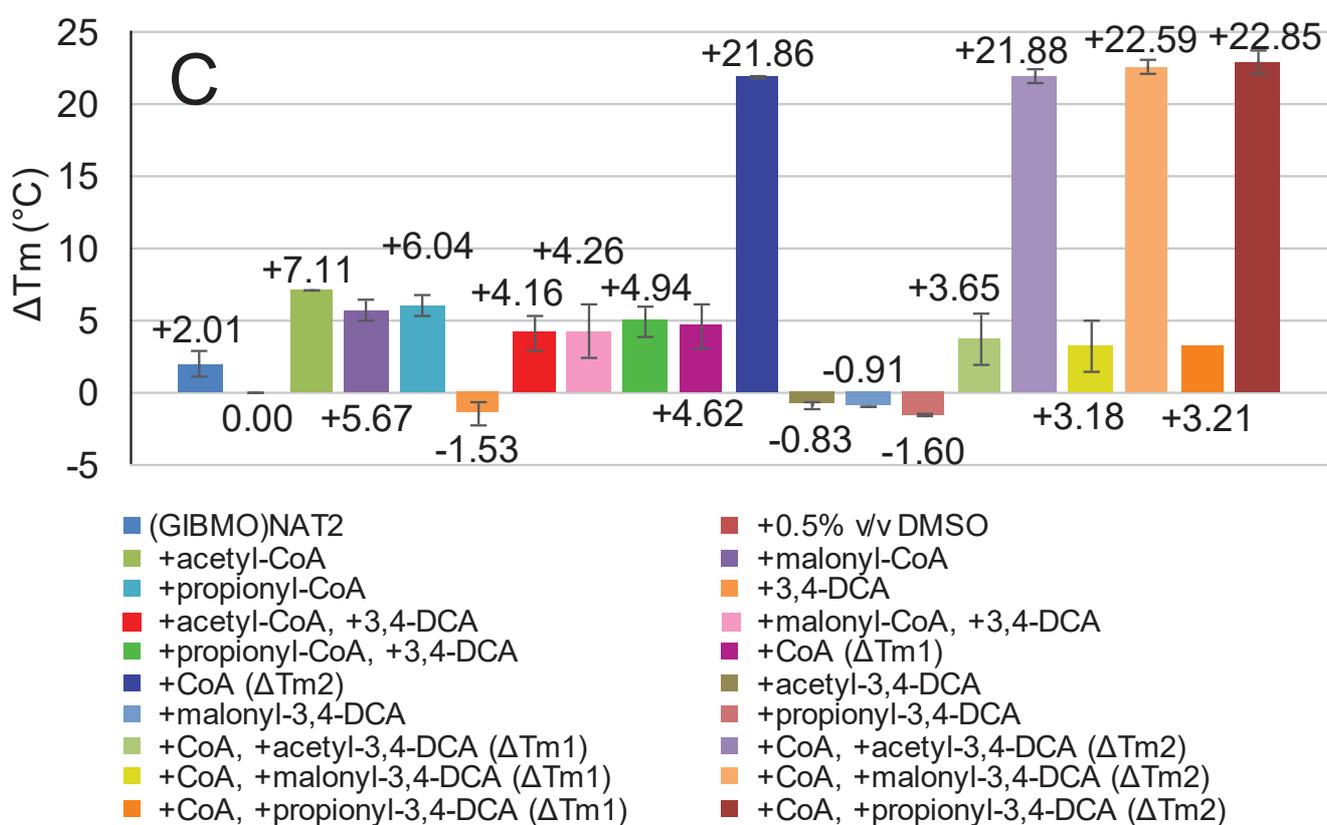
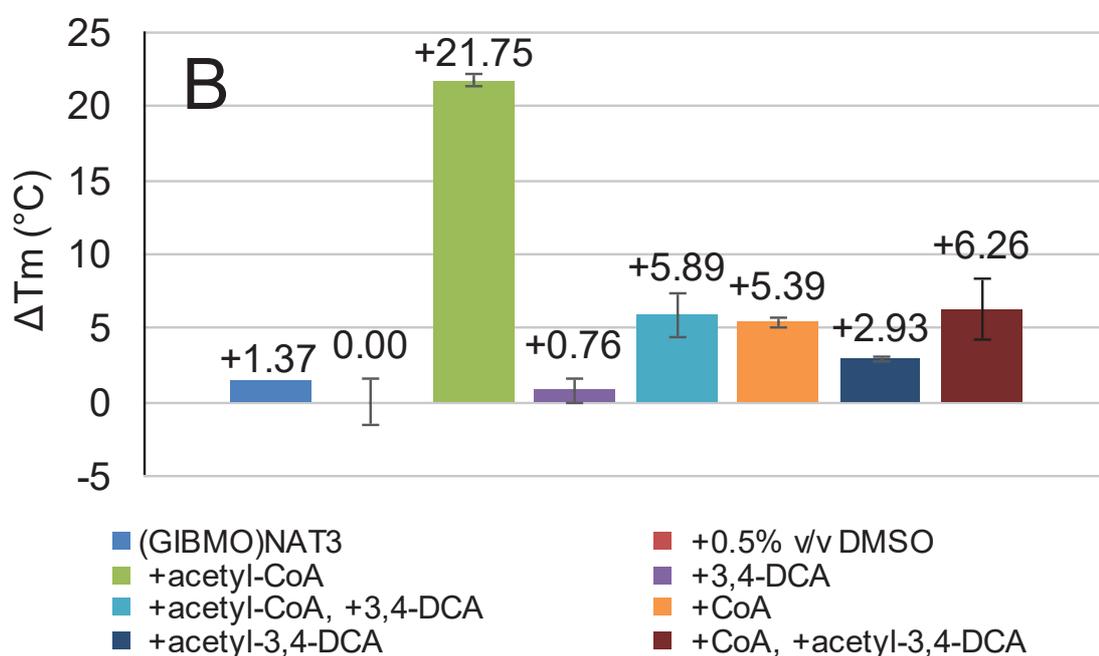
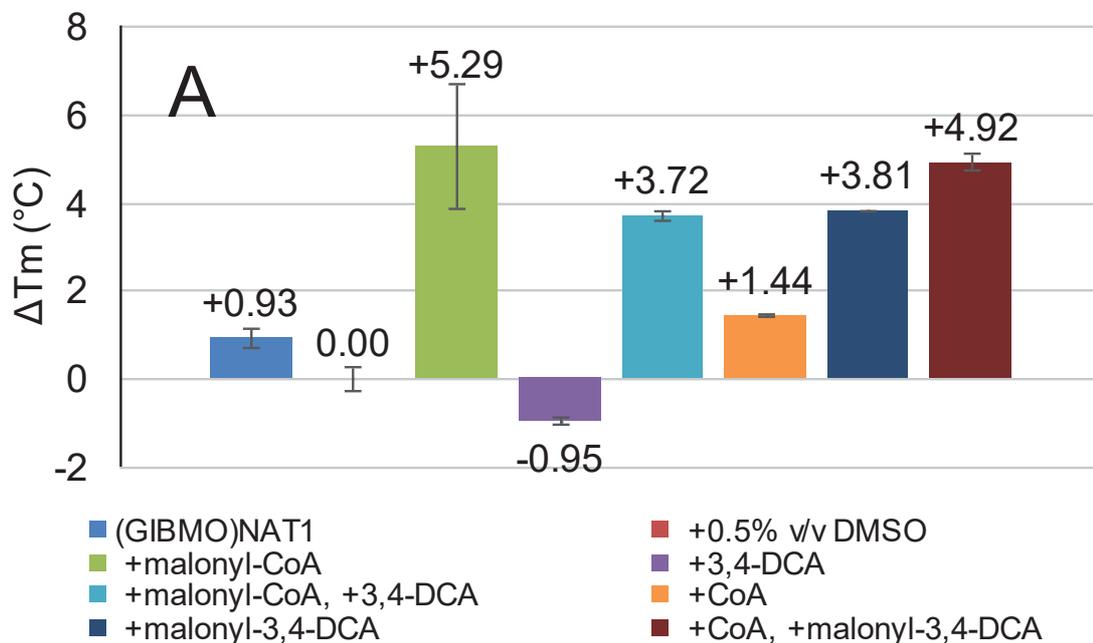


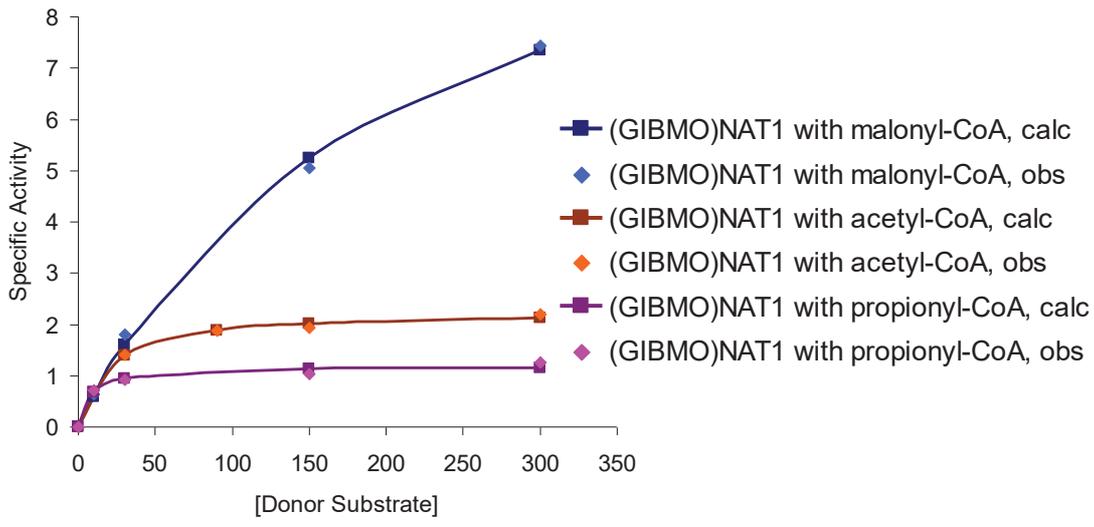
Fig. 7

Fig. 8



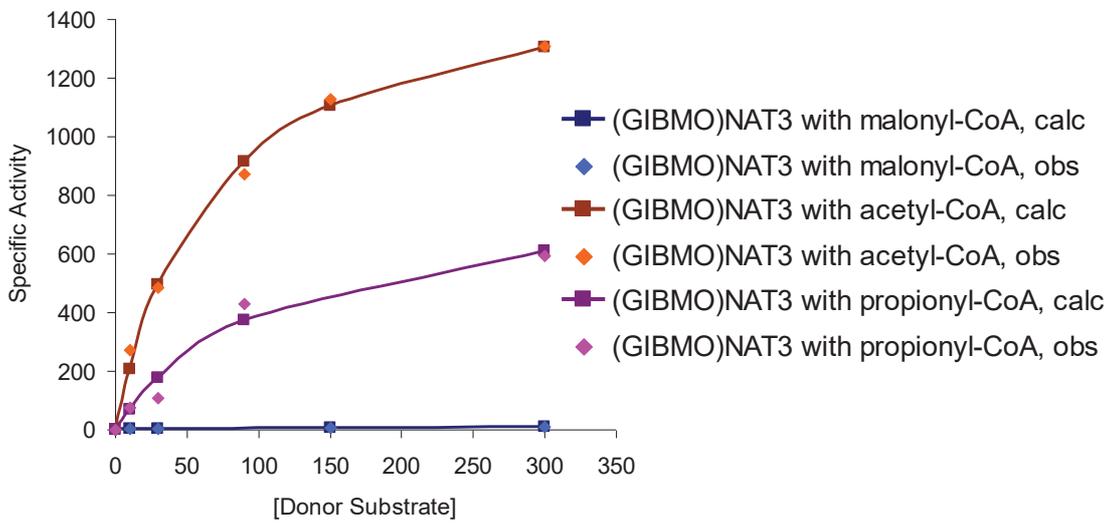


A



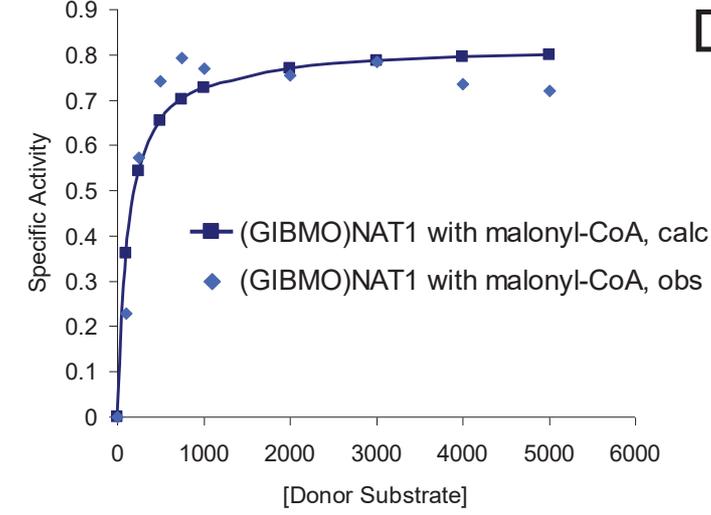
(GIBMO)NAT1	
Malonyl-CoA	
app <i>K</i> _{cat} (1/s)	12.3
app <i>K</i> _m (μM)	201.6
Acetyl-CoA	
app <i>K</i> _{cat} (1/s)	2.3
app <i>K</i> _m (μM)	18.7
Propionyl-CoA	
app <i>K</i> _{cat} (1/s)	1.2
app <i>K</i> _m (μM)	7.3

B



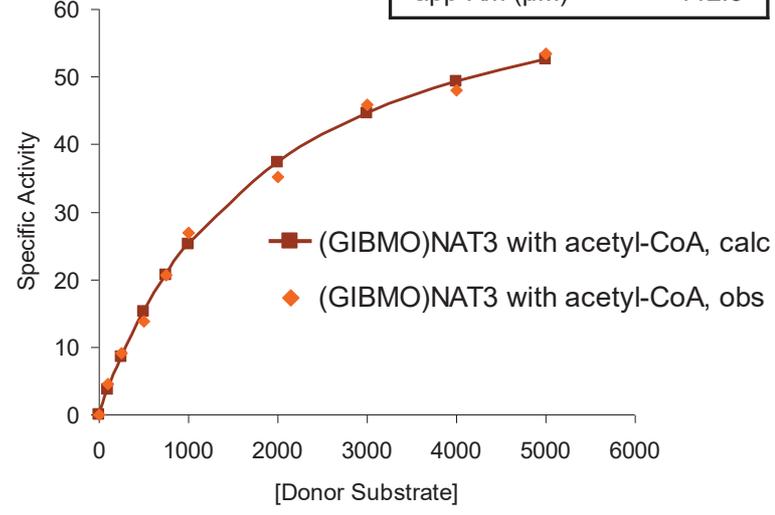
(GIBMO)NAT3	
Malonyl-CoA	
app <i>K</i> _{cat} (1/s)	8.9
app <i>K</i> _m (μM)	24.6
Acetyl-CoA	
app <i>K</i> _{cat} (1/s)	1593.6
app <i>K</i> _m (μM)	66.6
Propionyl-CoA	
app <i>K</i> _{cat} (1/s)	838.1
app <i>K</i> _m (μM)	112.3

C



(GIBMO)NAT1	
Malonyl-CoA	
app <i>K</i> _{cat} (1/s)	0.8
app <i>K</i> _m (μM)	127.2

D



(GIBMO)NAT3	
Acetyl-CoA	
app <i>K</i> _{cat} (1/s)	72.4
app <i>K</i> _m (μM)	1876.9

Fig.10

Fig. 11

