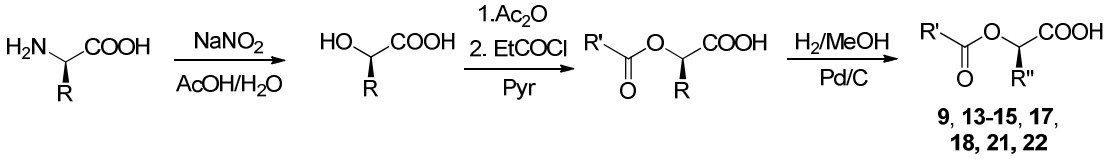
Supporting Information

Functional Annotation and Three-Dimensional Structure of an Incorrectly Annotated Dihydroorotase from cog3964 in the Amidohydrolase Superfamily.

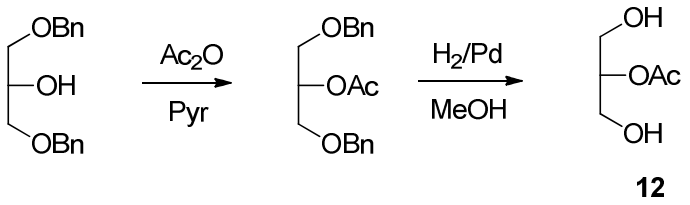
Synthesis of selected α-acetyl carboxylates

The syntheses of compounds **9, 13-15, 17-19, 21, 22** are outlined in **Scheme 1**. The D-α-hydroxy acids were synthesized from appropriate amino acids according literature procedure (*1*). Compounds **9**, **15**, **17**, **21** were synthesized from *O*-Bn protected D-amino acids. Alpha hydroxy acids were acetylated with Ac_2O or propionylated with EtCOCl in pyridine. To obtain the final products for compounds **9**, **15**, **17**, **21,** the benzyl protecting group was removed by catalytic hydrogenation in MeOH. (2*R*,3*S*)-2-acetoxy-3-hydroxybutanoic acid **(15):** ^{1}H NMR (MeOH-d_4): δ 2.27 (d, *J* = 6.6Hz, 3H), 2.18 (s, 3H), 4.18-4.34 (m, 1H), 4.89 (d, *J* = 3.3 Hz, 1H). ^{13}C NMR (MeOH-d_4): δ 18.1, 19.2, 66.5, 76.3, 170.6, 170.9. TOF MS/ESI $(M-H)^-$ 161.0577(found), 162.0528 (calculated). (*R*)-3-hydroxy-2-(propionyloxy)propanoic acid **(21):** ^{1}H NMR (MeOH-d_4): δ 1.14 (t, *J* = 7.6 Hz, 3H), 2.46 (q, *J* = 7.62 Hz, 2H), 3.90 (d, *J* = 4.5 Hz, 2H), 5.06 (t, *J* = 3.9 Hz, 1H). ^{13}C NMR (MeOH-d_4): δ 18.1, 19.2, 66.5, 76.3, 170.6, 170.9. TOF MS/ESI $(M-H)^-$ 161.0840 (found), 162.0528 (calculated).



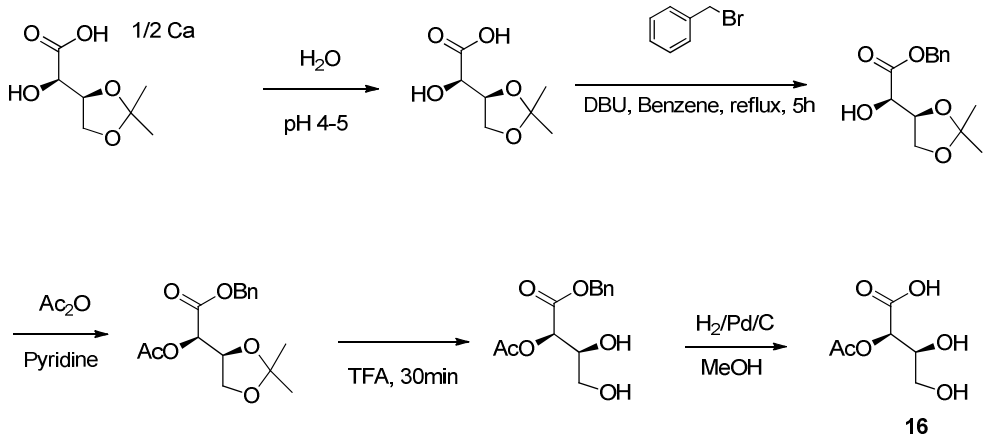
Scheme 1

The synthesis of **12** is outlined on **Scheme 2**.



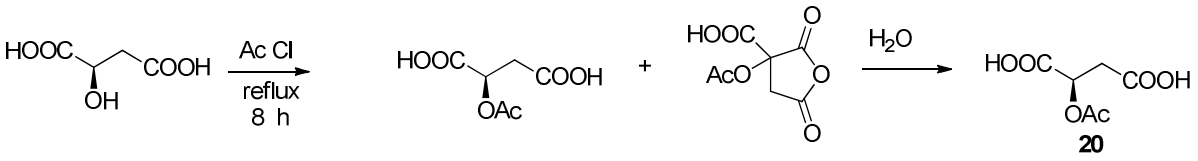
Scheme 2

The synthesis of **16** is outlined on **Scheme 3.** (*R*)-2-hydroxy-2-((*S*)-2,2-dimethyl-1,3-dioxolan-4-yl) acetic acid calcium salt was converted into free acid by acidifying solution to pH 4-5 (1 N HCl) and excessive extracting with EtOAc. Reactant was then converted to benzyl ester and subsequently acetylated with Ac_2O in pyridine (*2*). Deprotection of diacetonide protecting group in acidic conditions and benzyl group removal by catalytic hydrogenation in MeOH afforded **16**. (2*R*,3*S*)-2-acetoxy-3,4-dihydroxybutanoic acid **(16):** ^{1}H NMR (MeOH-d_4): δ 2.14 (s, 3H), 3.56 (dd, *J* =6.3, 2.7 Hz, 2H), 4.12-4.17 (m, 1H), 5.14 (d, 2.4 Hz, 1H), ^{13}C NMR (MeOH-d_4): δ 19.1, 61.5, 70.9, 72.3, 170.8. TOF MS/ESI $(M-H)^-$ 177.0908 (found), 178.0477 (calculated).



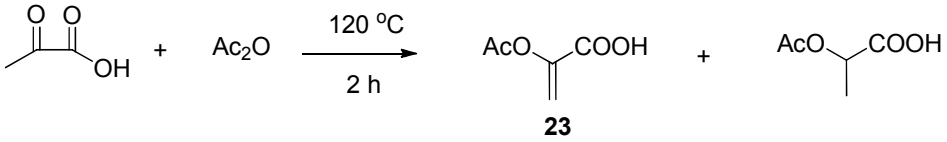
Scheme 3

The synthesis of compound **20** is outlined in **Scheme 4**. Malic acid was refluxed in acelyl chloride for several hours (5-6h). Evaporation of excess acetyl chloride resulted to the mixture of acid anhydride with the desired product. Side product was hydrolyzed with a small amount of water. Final product **20** was extracted from water with EtOAc.



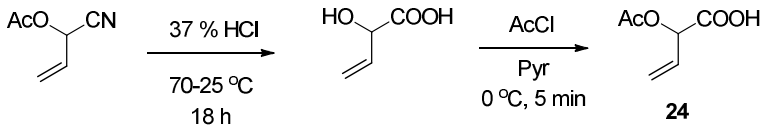
Scheme 4

The synthesis of compound **23** is outlined on **Scheme 5.** Synthesis of 2-acetoxyacrylic acid **23** was done according literature procedure (*3*). A mixture of pyruvic acid and acetyl anhydride was heated at 120 ^{o}C for 2 h and after cooling down to 23 ^{o}C was distilled under reduced pressure (11-15 mm). Fractions collected at 90-120 ^{o}C contained about 50 % of side product. Desired **23** as white crystalline compound was obtained after purification by column chromatography (toluene:EtOAc , 5:1). 2-acetoxyacrylic acid **23:** ^{1}H NMR ($CDCl_3$): δ 2.22 (s, 3H), 5.57 (d, *J* = 1.8 Hz, 1H), 6.17 (d, *J* = 2.1 Hz, 1H), 10.39 (s, 1H). ^{13}C NMR ($CDCl_3$): δ 20.3, 116.4, 143.9, 166.3, 169.2. TOF MS/ESI $(M-H)^-$ 129.0306 (found), 130.0266 (calculated).



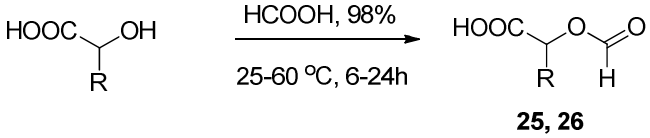
Scheme 5

The synthesis of compound **24** is outlined on **Scheme 6.** 2-hydroxybut-3-enoic acid was obtained by a published procedure (*4*). and then acetylated with AcCl in pyridine at 0 ^{o}C to obtain **24**.



Scheme 6

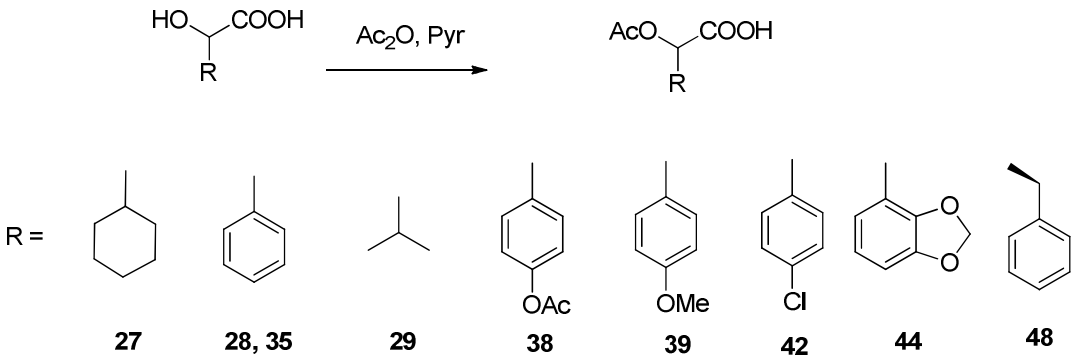
The syntheses of compounds **25** and **26** are outlined on **Scheme 7**. 2-Formyloxy acids **25** and **26** were synthesized from *DL*-lactic acid and glycolic acid respectively. Formylation was achieved in concentrated (98 %) formic acid at 50-60^{o}C for 12 h.



(**25**) R = H, (**26**) R = Me

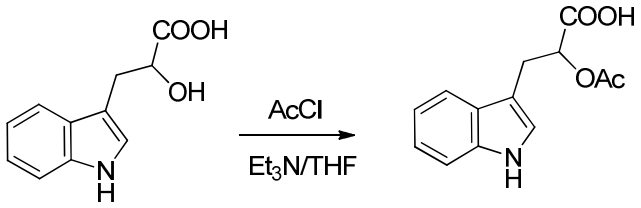
Scheme 7

The syntheses of **27-29**, **35**, **38**, **39**, **42**, **44**, and **48** are outlined in **Scheme 8**. *DL* (**27-29**, **38**, **39**, **42**, **44**), *L*- (**35**) and *D*- (**48**) were obtained by acetylation with Ac_2O of alpha hydroxyl acids in pyridine.



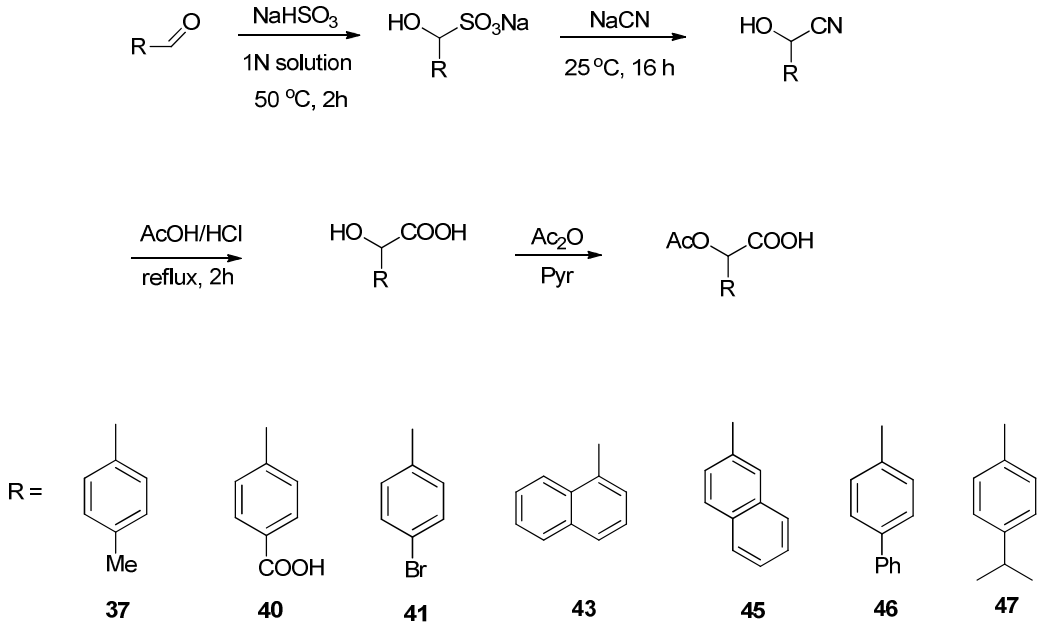
Scheme 8

The synthesis of **50** was synthesized according to **Scheme 9**. 2-acetoxy-3-(1H-indol-3-yl)propanoic acid **(50)** was synthesized from DL-1*H*-indole-3-lactic acid according described procedure (*5*).



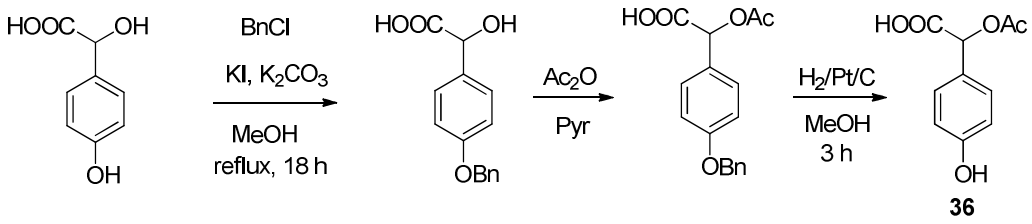
Scheme 9

The syntheses of compounds **37**, **40**, **41**, **43**, **45**, **46** and **47** are outlined on **Scheme 10**. *O*-Acetylated alpha *DL*-hydroxyl acids were synthesized from appropriate aryl aldehydes according **scheme 10**. Aryl aldehydes were converted to aryl nitriles which further were hydrolyzed to acids and acetylated with Ac_2O in pyridine (*6*).



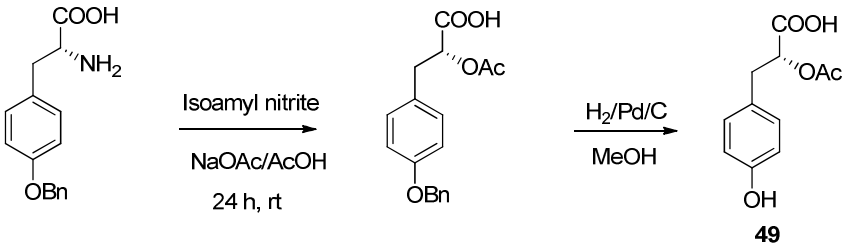
Scheme 10

The synthesis of 2-acetoxy-2-(4-hydroxyphenyl)acetic acid **36** is outlined in **Scheme 11**. 4-Hydroxy group of 4-hydroxymandelic acid was protected with benzyl group (*7*) and product was acetylated and then p-OH group was recovered by 3 h catalytic hydrogenation (by controlling TLC and NMR).



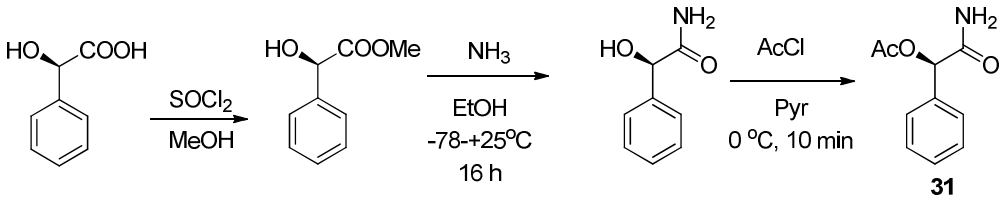
Scheme 11

The synthesis of (*R*)-2-acetoxy-3-(4-hydroxyphenyl)propanoic acid **49** is outlined in **Scheme 12.** *O*-Benzyl protected D-tyrosine was converted to *O*-acetylated product (*8*) and later deprotected by catalytic hydrogenation to afford **49**.



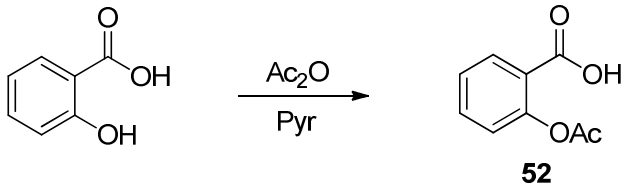
Scheme 12

Synthesis of (*R*)-2-amino-2-oxo-1-phenylethyl acetate **31** was obtained starting from R-mandelic acid according to **Scheme 13**.



Scheme 13

Compound **52** was synthesized by acetylation of salicylic acid with Ac_2O in pyridine according to (**Scheme 14**).



Scheme 14

Experimental

General procedure for acetylation of α-hydroxy acids with Ac_2O. *D*-, *L*-or *DL*-α-hydroxy acid (1 mmol) was dissolved in 5 mL of pyridine and 1.5 equiv. (1.5 mmol) of Ac_2O was slowly added. Reaction mixture was stirred at 25 ^{o}C overnight; this mixture was then poured on ice-water (20 mL) and acidified with 37% HCl to pH 1-2. Product was extracted with EtOAc (4 X 25 mL). The combined organic layers were washed with brine, dried over $NaSO_4$ and evaporated to dryness affording the title compounds in 72 -77 % yield.

General procedure for acetylation with AcCl and propionylation with EtCOCl. To a stirred solution of the α-hydroxy acid (1mmol) in pyridine (5 mL) at 0 ^{o}C was added acetyl chloride or propionyl chloride (3-4 mmol) and stirring was continued another 5 min. The reaction mixture was then diluted with EtOAc (30 mL) and washed with 5 % HCl solution and then with brine, dried over $NaSO_4$ and evaporation to dryness affording the title compounds in 68 -70 % yield. In some cases the crude products were subjected to flash column chromatography for further purification.

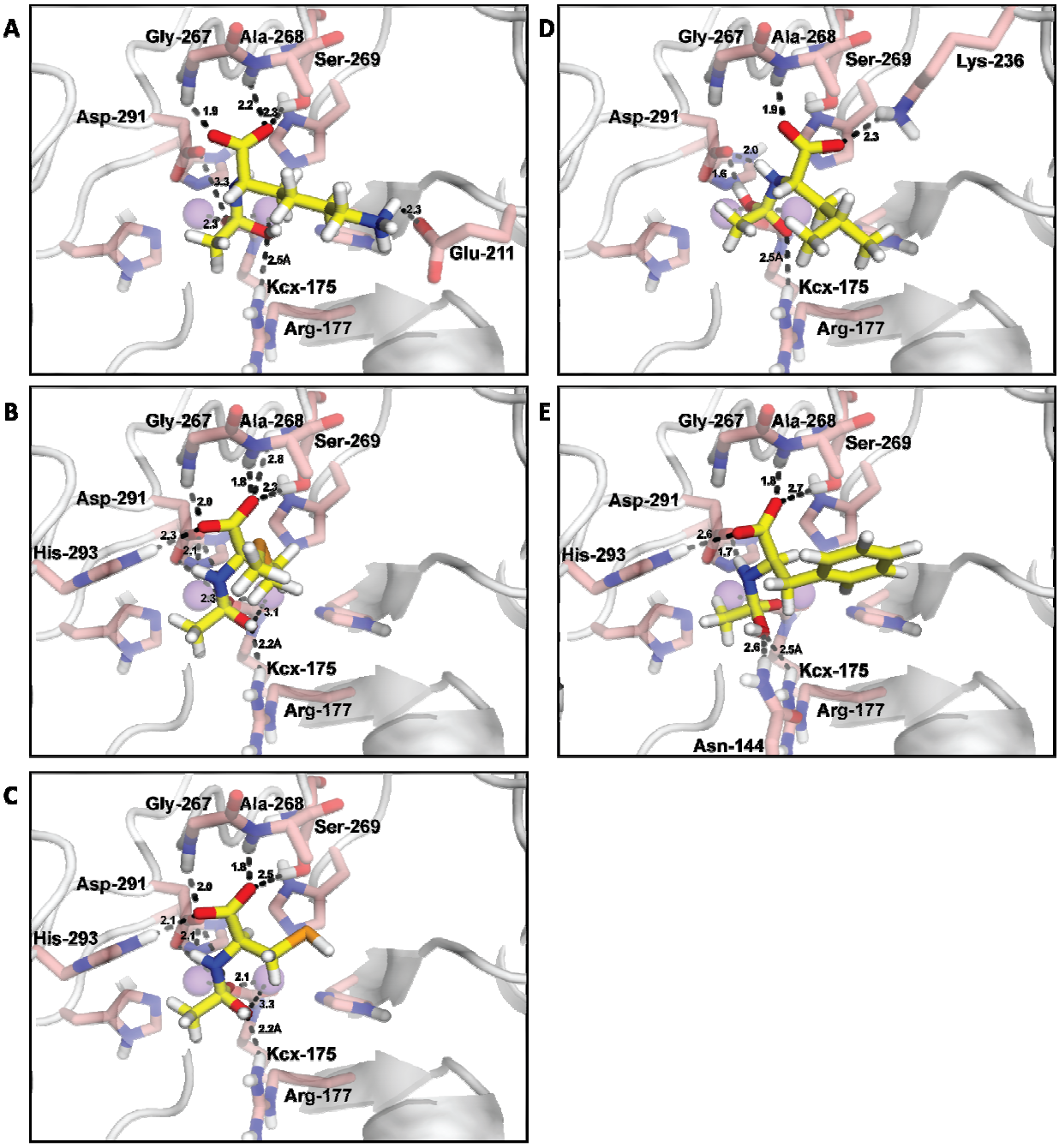


Figure S1: Five well-posed acetylated amino acids docked well to Atu3266 out of 10 present in the KEGG HEI library. All five metabolites were found in the top 0.76% of the docking list, where acetylated amino acids: (**A**) L-lysine is ranked 40, (**B**) D-methionine is ranked 171, (**C**) D-cysteine is ranked 218, (**D**) L-leucine is ranked 313 and (**E**) D-phenylalanine is ranked 440.

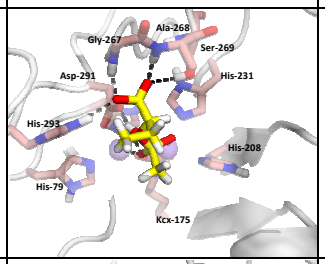
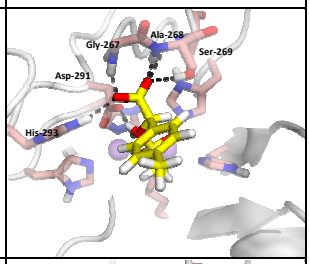
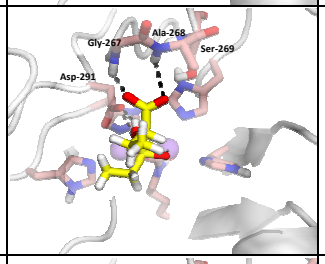
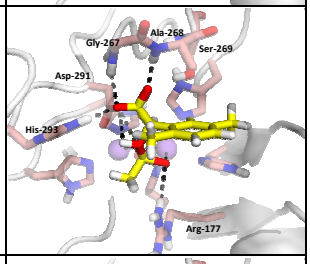
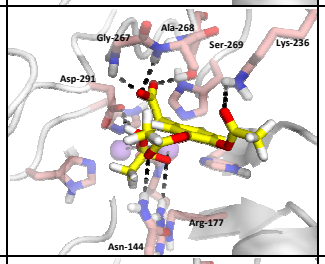
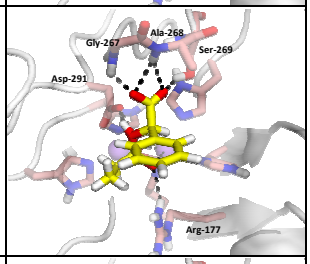
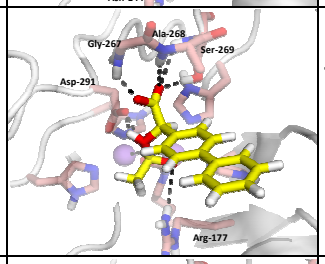
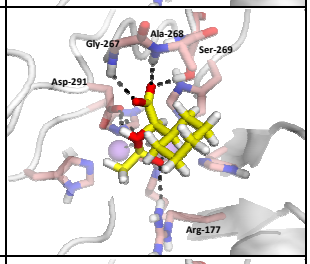
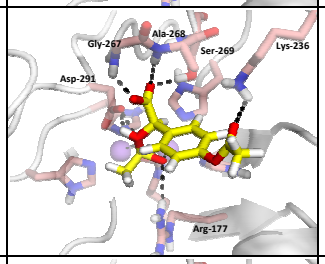
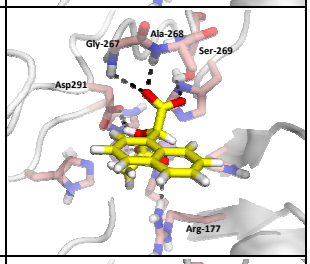
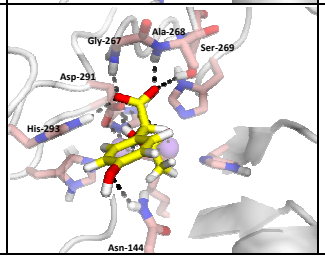
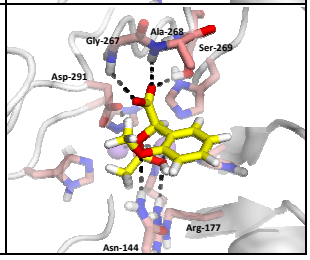
	Substrate Name	Substrate #	Docking Rank	Docking Pose		Substrate Name	Substrate #	Docking Rank	Docking Pose
a	*R*-2-acetoxybutanoic acid	**18**	**77**		**g**	*R*-4-methyl acetyl mandelate	**37**	**160**	
b	*R*-2-(propionyloxy)-butanoate	**22**	**78**		**h**	*R*-2,5-dimethyl acetyl mandelate	n/a	**206**	
c	*R*-2-acetoxy-2-(3,4-diacetoxyphenyl) acetic acid	n/a	**91**		**i**	R-propionyloxy mandelate	n/a	**241**	
d	R-2-([1,1'-biphenyl]-4-yl)-2-acetoxyacetic acid	**46**	**100**		**j**	R-2-acetoxy-2-cyclohexylacetic acid	**27**	**330**	
e	R-2-acetoxy-2-(4-acetoxyphenyl) acetic acid	**38**	**109**		**k**	R-2-acetoxy-2-(naphthalen-1-yl) acetic acid	**43**	**431**	
f	R-2-acetoxy-2-(4-hydroxyphenyl) acetic acid	**36**	**150**		**l**	R-2-acetoxy-2-(2-methoxyphenyl) acetic acid	n/a	**452**	

Figure S2: The docked poses from the dedicated HEI library, for compounds that contain the acetylated α-hydroxy carboxylate core that can be turned over by Atu3266 and Oant2987. Substituents represent (**A-B**) branched alkanes, (**J**) cycloalkanes, and (**B-I, K-L)** phenyl ring containing compounds. The phenyl ring can be substituted in the para (**D, G**) or ortho and meta positions (**H, K**). Other arene substitution patterns allow for charged interactions with (**F, L**) Asn-144 or (**C,E**) Lys-236.

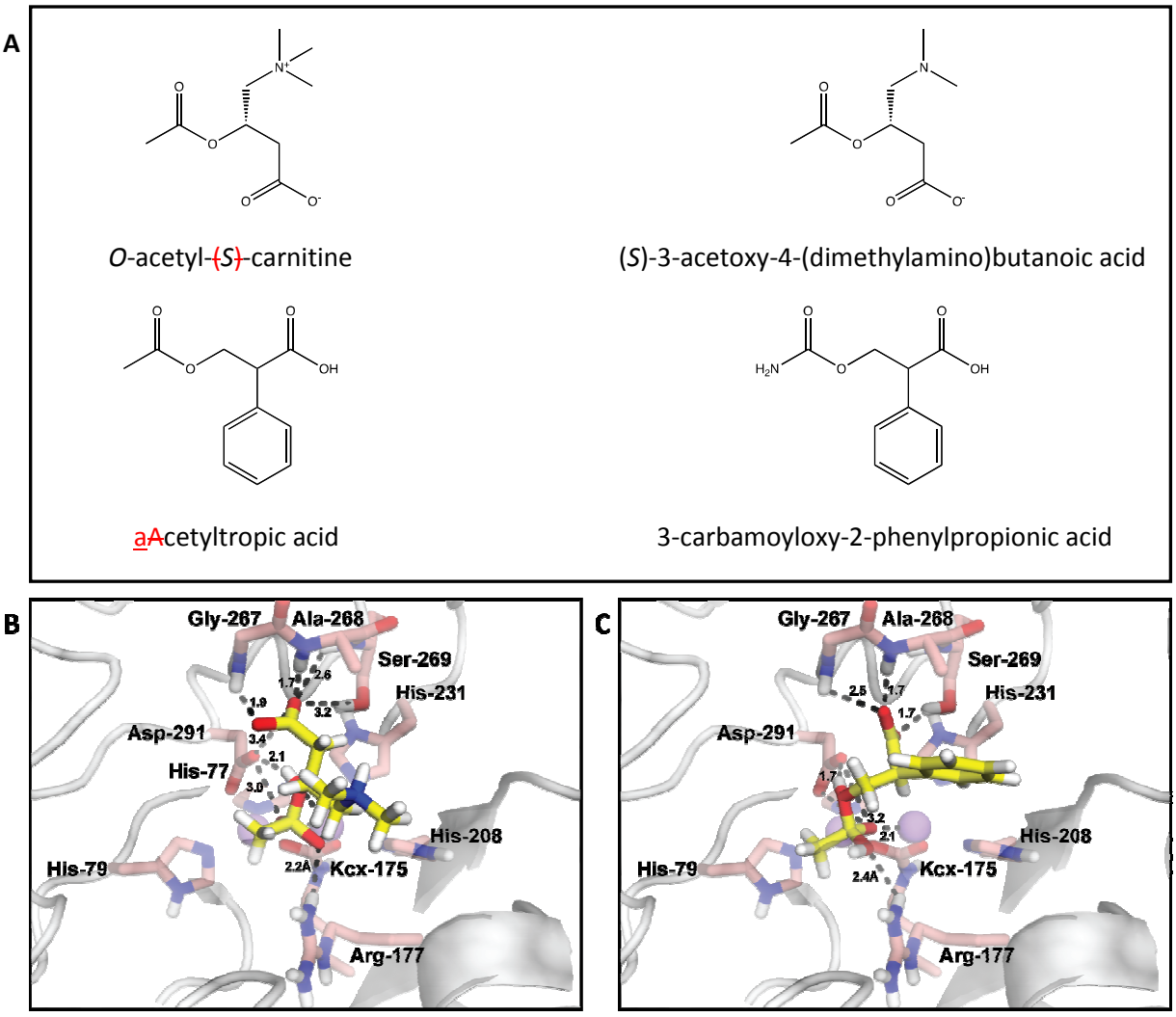


Figure S3: (**A**) Molecules with one carbon insertions around the chiral carbon of the acetylated α-hydroxy carboxylate backbone can dock into the active site of Atu3266. Two compounds were tested to see if substitutions on the acetyl or carboxy side are tolerated. (**B**) Shows the non active (*S*)-3-acetoxy-4-(dimethylamino)butanoic acid and (**C**) is acetyltropic acid that points the extra carbon extension to the top of the active site and display low turnover with Atu3266.

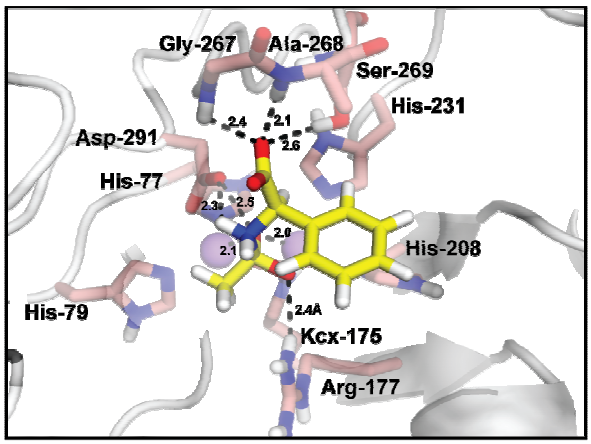


Figure S4: Well-posed acetyl-D-phenyl glycine (**33**) from the docking run of the dedicated HEI library showing the complementarity of the Atu3266 active site and the *N*-acetylated amino acid backbone chemotype.

Table S1: Additional enzymes found in group 6. Locus tag and respective organism encoding other proteins in group 6 of COG3964.

Locus tag	Organism
Atu3266	*Agrobacterium tumefaciens* C58
Oant2987	*Ochrobactrum anthrop*i ATCC 49188
RHE_PE00295	*Rhizobium etli* CFN 42
pRL110419	*Rhizobium leguminosarum* bv. Viciae 3841
RHECIAT_PA0000241	*Rhizobium etli* CIAT 658
Rleg2_5803	*Rhizobium leguminosarum* bv. Trifolii WSM2304
Arad_7942	*Agrobacterium radiobacter* K84
Avi_5090	*Agrobacterium vitis* S4
Rleg_6703	*Rhizobium leguminosarum* bv. Trifolii WSM 1325

References

1. Deechongkit, S., You, S.-L., Kelly J. W. (2004) Synthesis of all nineteen appropriately protected chiral alpha-hydroxy acid equivalents of the alpha-amino acids for boc solid-phase depsi-peptide synthesis. *Org Lett***,** *6* (4), 497-500.
2. Gallo, E. A., and Gellman, S. H. (1993) Hydrogen-bond-mediated folding in depsipeptide models of beta -turns and alpha -helical turns. *J. Am. Chem. Soc.***,** *115* (21), 9774-9788.
3. Seifert, P., Vogel, E., Rossi, A., and Schinz, H. (1950) Some reactions with derivatives of alpha -keto acids and alpha –ketoesters. *Helvetica Chimica Acta*, *33*, 725-736.
4. Li, R., Powers, V. M., Kozarich, J. W., and Kenyon, G. L. (1995) Racemization of Vinylglycolate Catalyzed by Mandelate Racemase. *J. Org. Chem. 60* (11), 3347-3351.
5. Thompson, A. M., Fry, D. W., Kraker, A. J., and Denny, W. A. (1994) Tyrosine kinase inhibitors. 2. Synthesis of 2,2'-dithiobis(1H-indole-3-alkanamides) and investigation of their inhibitory activity against epidermal growth factor receptor and pp60v-src protein tyrosine kinases. *Journal of Medicinal Chemistry*, *37*, 598-609.
6. Weist, S., Kittel, C., Bischoff, D., Bister, B., Pfeifer, V., Nicholson, G. J., Wohlleben, W., and Suessmuth, R. D. (2004) Mutasynthesis of Glycopeptide Antibiotics: Variations of Vancomycin's AB-Ring Amino Acid 3,5-Dihydroxyphenylglycine. *Journal of the American Chemical Society*, *126*, 5942-5943.
7. Massolini, G., Fracchiolla, G., Calleri, E., Carbonara, G., Temporini, C., Lavecchia, A., Cosconati, S., Novellino, E., and Loiodice, F. (2006) Elucidation of the enantioselective recognition mechanism of a penicillin G acylase-based chiral stationary phase towards a series of 2-aryloxy-2-arylacetic acids. *Chirality*, *18*, 633-643.

8. Maerki, H. P., Crameri, Y., Eigenmann, R., Krasso, A., Ramuz, H., Bernauer, K., Goodman, M., and Melmon, K. L. (1998) Optically pure isoproterenol analogs with side chains containing an amide bond. Synthesis and biological properties. *Helvetica Chimica Acta*, *71*, 320-336.

9. Wilchek, M., Oka, T., and Topper, Y. J. (1975) Structure of a soluble superactive insulin is revealed by the nature of the complex between cyanogen bromide-activated Sepharose and amines. *Proc. Nat. Acad. Sci. U.S.A. 72*, 1055-1058.