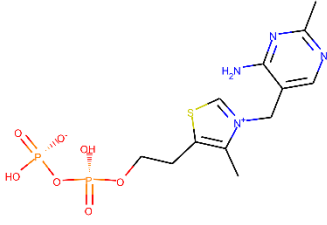
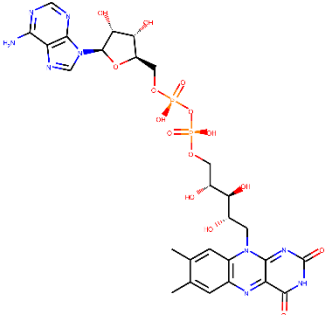
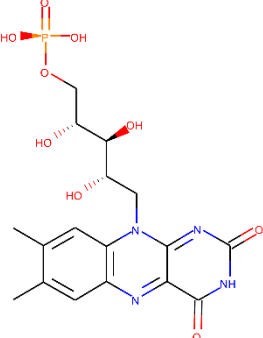
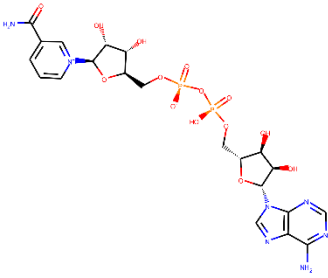
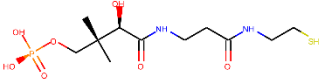
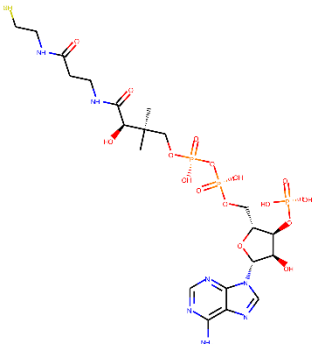
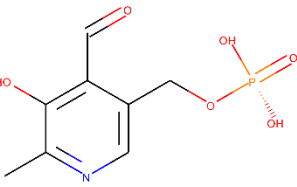
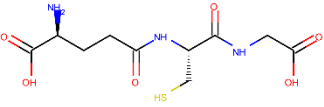


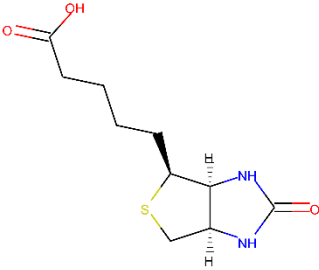
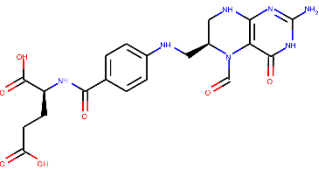
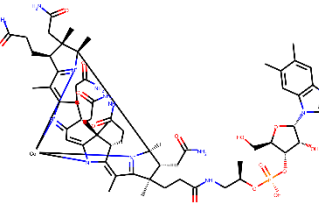
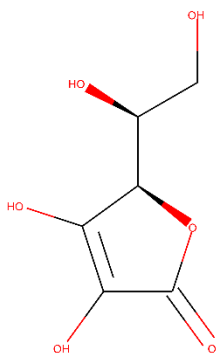
## Supplementary Information

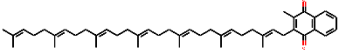
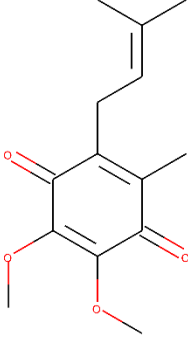
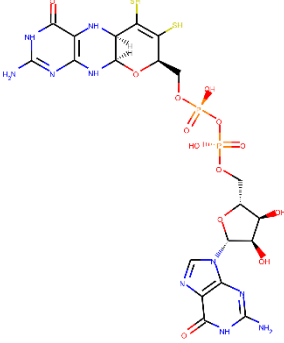
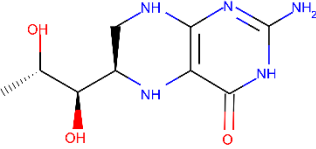
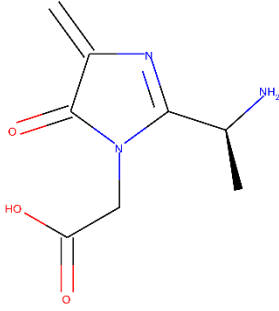
### Tables

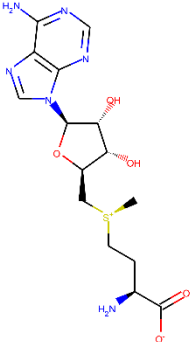
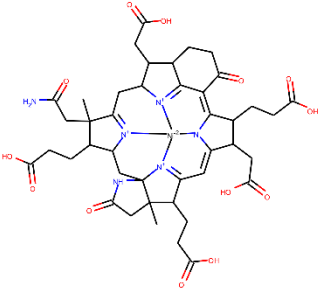
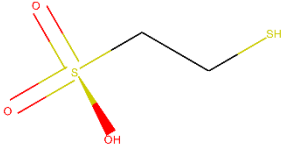
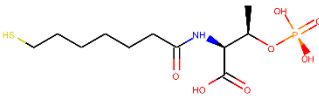
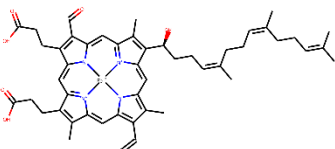
**Table 1: Details of the 27 cofactor classes in the CoFactor database and the template molecule identified for each cofactor class**

Cofactor class	Cofactor type	Reference chemical component ID from PDB chemical component dictionary	Chemical structure	Smiles code
1	Thiamine diphosphate	TDP		<chem>Cc1c(sc[n+]1Cc2cnc(nc2N)C)CCOP(=O)(O)OP(=O)(O)O</chem>
2	Flavin adenine dinucleotide	FAD		<chem>[P@](=O)(O)OC[C@H]4[C@H]([C@H]([C@@H](O4)n5cnc6c5ncnc6N)O)O)O)O</chem>
3	Flavin Mononucleotide	FMN		<chem>Cc1cc2c(cc1C)N(C3=NC(=O)NC(=O)C3=N2)C[C@@H]([C@@H]([C@@H](COP(=O)(O)O)O)O)O</chem>

4	Nicotinamide-adenine dinucleotide	NAD		<chem>c1cc(c[n+](c1)[C@H]2[C@@H]([C@@H]([C@H](O2)CO[P@@](=O)([O-])O[P@@](=O)(O)OC[C@H]3[C@H]([C@H]([C@@H](O3)n4cnc5c4ncnc5N)O)O)O)C(=O)N</chem>
5	Phosphopantetheine	PNS		<chem>CC(C)(COP(=O)(O)O)[C@H](C(=O)NCCC(=O)NCCS)O</chem>
6	Coenzyme A	COA		<chem>CC(C)(CO[P@](=O)(O)O[P@@](=O)(O)OC[C@H]1[C@H]([C@H]([C@@H](O1)n2cnc3c2ncnc3N)O)OP(=O)(O)O)[C@H](C(=O)NCCC(=O)NCCS)O</chem>
7	Pyridoxal 5'-phosphate	PLP		<chem>Cc1c(c(c(cn1)COP(=O)(O)O)C=O)O</chem>
8	Glutathione	GSH		<chem>c1cc(ccc1CSC[C@@H](C(=O)NCC(=O)O)NC(=O)CC[C@@H](C(=O)O)[NH3+])</chem>

9	Biotin	BTN		<chem>C1[C@H]2[C@@H]([C@@H](S1)CCCC(=O)O)NC(=O)N2</chem>
10	Tetrahydrofolic acid	FFO		<chem>O=C(O)CC[C@@H](C(=O)O)NC(=O)c1ccc(cc1)NC[C@H](N2)CNC(c23)nc(N)[nH]c3=O</chem>
11	Adenosylcobalamin	B12		<chem>Cc1cc2c(cc1C)n(cn2)C3C(C(C(O3)CO)OP(=O)(O)OC(C)CNC(=O)CC4(C(C5C6(C(C(C7=[N]6[Co+2]89[N]5=C4C(=C1[NH]8C(=CC2=[N]9C(=C7C)C(C2CCC(=O)N)(C)CC(=O)N)C(C1CCC(=O)N)(C)C)CC(=O)N)(C)CC(=O)N)C)CC(=O)N)C)O</chem>
12	Ascorbic acid	ASC		<chem>C([C@@H]([C@@H]1C(=C(C(=O)O1)O)O)O)O</chem>

13	Menaquinone	MQ7		<chem>CC(C)=CCC/C(C)=C/CC/C(C)=C/CC/C(C)=C/CC/C(C)=C/CC/C(C)=C/CC/C(C)=C/CC(C1=O)=C(C)C(=O)c12)cccc2</chem>
14	Ubiquinone	UQ1		<chem>CC1=C(C(=O)C(=C(C1=O)OC)OC)CC=C(C)C</chem>
15	Molybdopterin	MGD		<chem>c1nc2c(n1[C@H]3[C@@H]([C@@H]([C@H](O3)CO[P@@](=O)(O)O[P@@](=O)(O)OC[C@H]4C(=C([C@H]5[C@@H](O4)NC6=C(N5)C(=O)NC(=N6)N)S)S)O)O)N=C(NC2=O)N</chem>
16	Bioterin	H4B		<chem>C[C@@H]([C@@H]([C@H]1CNC2=C(N1)C(=O)NC(=N2)N)O)O</chem>
17	MIO (4-methylideneimidazole-5-one)	MDO		<chem>C[C@@H]([C1=NC(=C)C(=O)N1C(=O)O)N</chem>

18	S-adenosylmethionine	SAM		<chem>C[S@@+](CC[C@@H](C(=O)[O-])N)C[C@@H]1[C@H]([C@H]([C@@H](O1)n2cnc3c2ncnc3N)O)O</chem>
19	Factor F430	F43		<chem>NC(=O)C[C@@]1(C)[C@H](CCC(=O)O)[C@H](C2)[N+]3=C1Cc4c(CC(=O)O)c(c5n4[Ni]36N7C8[C@@H](CCC(=O)O)[C@]([C@@]279)(C)CC(=O)N9)CCC(=O)C5=C1N6C(=C8)[C@@H](CC(=O)O)[C@H]1CCC(=O)O</chem>
20	Coenzyme M	COM		<chem>SCCS(=O)(=O)O</chem>
21	Coenzyme B	TP7		<chem>O=P(O)(O)O[C@H](C)[C@@H](C(=O)O)NC(=O)CCCCCS</chem>
22	Heme	HEA		<chem>CC(C)=CCC/C(C)=C/CC/C(C)=C/CC[C@H](O)C1=C(C)C(=[N+]2C1=Cc3c(C)c(C=C)c4n3[Fe]256)C=c7n5c(c(c7C=O)CCC(=O)O)=CC8=[N+]6C(=C4)C(C)=C8CCC(=O)O</chem>

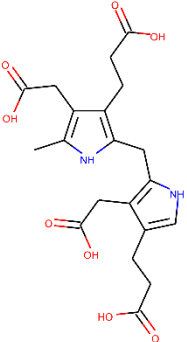
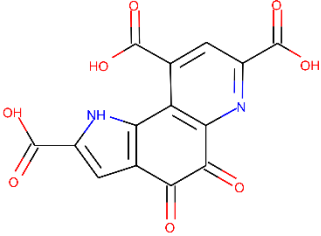
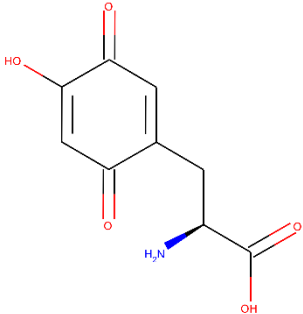
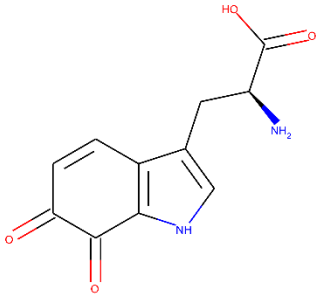
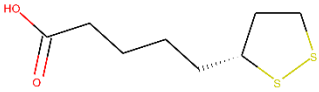
24	Dipyrromethane	DPM		<chem>O=C(O)CCC(=C1CC(=O)O)[C@@H](N=C1C)CC2=NCC(CCC(=O)O)=C2CC(=O)O</chem>
25	Pyrroloquinoline Quinone	PQQ		<chem>O=C(O)c(c1)[nH]c(c12)c3c(C(=O)O)C2=O)nc(C(=O)O)cc3C(=O)O</chem>
26	Topaquinone	TPQ		<chem>C1=C(C(=O)C=C(C1=O)O)C[C@@H](C(=O)O)N</chem>
27	Orthoquinone	TRQ		<chem>O=C(O)[C@@H](N)Cc(c[nH]1)c(c12)C=CC(=O)C2=O</chem>
28	Lipoic acid	LPA		<chem>C1CSS[C@@H]1CCCCC(=O)O</chem>

Table 2. The complete set of EC numbers identified using information from CoFactor and BRENDA databases for each cofactor class

<b>Cofactor class ID</b>	<b>Representative chemical component id from PDB for each Cofactor class</b>	<b>EC numbers for each Cofactor class from the CoFactor Database</b>	<b>Additional EC numbers identified for each cofactor class from the BRENDA Database</b>	<b>EC numbers identified for a Cofactor class</b>
<b>1</b>	<b>TDP</b>	<b>30</b>	<b>19</b>	<b>49</b>
<b>2</b>	<b>FAD</b>	<b>235</b>	<b>316</b>	<b>551</b>
<b>3</b>	<b>FMN</b>	<b>43</b>	<b>154</b>	<b>197</b>
<b>4</b>	<b>NAD</b>	<b>745</b>	<b>365</b>	<b>1110</b>
<b>5</b>	<b>PNS</b>	<b>8</b>	<b>5</b>	<b>13</b>
<b>6</b>	<b>COA</b>	<b>237</b>	<b>17</b>	<b>254</b>
<b>7</b>	<b>PLP</b>	<b>138</b>	<b>87</b>	<b>225</b>
<b>8</b>	<b>GSH</b>	<b>35</b>	<b>10</b>	<b>45</b>
<b>9</b>	<b>BTN</b>	<b>11</b>	<b>4</b>	<b>15</b>
<b>10</b>	<b>FFO</b>	<b>6</b>	<b>22</b>	<b>28</b>
<b>11</b>	<b>B12</b>	<b>15</b>	<b>33</b>	<b>48</b>
<b>12</b>	<b>ASC</b>	<b>20</b>	<b>16</b>	<b>36</b>



13	MQ7	1	10	11
14	UQ1	7	16	23
15	MGD	14	27	41
16	H4B	6	4	10
17	MDO	4		4
18	SAM	170	105	275
19	F43	1		1
20	COM	1	1	2
21	TP7	1		1
22	HEA	127	217	344
24	DPM		1	1
25	PQQ	9		9
26	TPQ		3	3
27	TRQ	1		1
28	LPA	6	2	8
<b>Grand Total</b>		<b>1871</b>	<b>1434</b>	<b>3305</b>

Table 3: Cofactor class description, match score limits to the template and total number of chemical components

Cofactor Class ID	Cofactor class name	Cofactor representative chemical component id	Number of chemical components	Match score limit
-------------------	---------------------	---	-------------------------------	-------------------

1	Thiamine diphosphate	TDP	38	0.6
2	Flavin adenine dinucleotide	FAD	15	0.87
3	Flavin Mononucleotide	FMN	7	0.86
4	Nicotinamide-adenine dinucleotide	NAD	44	0.68
5	Phosphopantetheine	PNS	1	1
6	Coenzyme A	COA	75	0.67
7	Pyridoxal 5'-phosphate	PLP	10	0.8
8	Glutathione	GSH	53	0.42
9	Biotin	BTN	5	0.58
10	Tetrahydrofollic acid	FFO	8	0.94
11	Adenosylcobalamin	B12	4	0.84
12	Ascorbic acid	ASC	1	1
13	Menaquinone	MQ7	1	1
14	Ubiquinone	UQ1	6	0.53
15	Molybdopterin	MGD	9	0.43
16	Biopterin	H4B	8	0.78

17	MIO	MDO	1	1
18	S-adenosylmethionine	SAM	24	0.6
19	Factor F430	F43	2	0.94
20	Coenzyme M	COM	1	1
21	Coenzyme B	TP7	6	0.75
22	Heme	HEA	25	0.52
24	Dipyrromethane	DPM	3	0.97
25	Pyrroloquinoline Quinone	PQQ	1	1
26	Topaquinone	TPQ	10	0.55
27	Orthoquinone	TRQ	4	0.88
28	Lipoic acid	LPA	2	1
	Grand total		364	

Table 4: Summary of the cofactor API endpoints

Method	Endpoint	Description
GET	<a href="https://www.ebi.ac.uk/pdbe/api/pdb/compound/cofactors">https://www.ebi.ac.uk/pdbe/api/pdb/compound/cofactors</a>	Summary of the cofactor annotation in the PDBe
GET	<a href="https://www.ebi.ac.uk/pdbe/api/pdb/entry/cofactor/:pdbid">https://www.ebi.ac.uk/pdbe/api/pdb/entry/cofactor/:pdbid</a>	Cofactors present in a PDB entry.
GET	<a href="https://www.ebi.ac.uk/pdbe/api/pdb/compound/cofactors/het/:id">https://www.ebi.ac.uk/pdbe/api/pdb/compound/cofactors/het/:id</a>	List of cofactor-like molecules similar to the query and their similarity to the template.

## Figure

Figure 1 - Distribution of cofactor-like chemical components as observed in the PDB chemical component dictionary across Enzyme classes

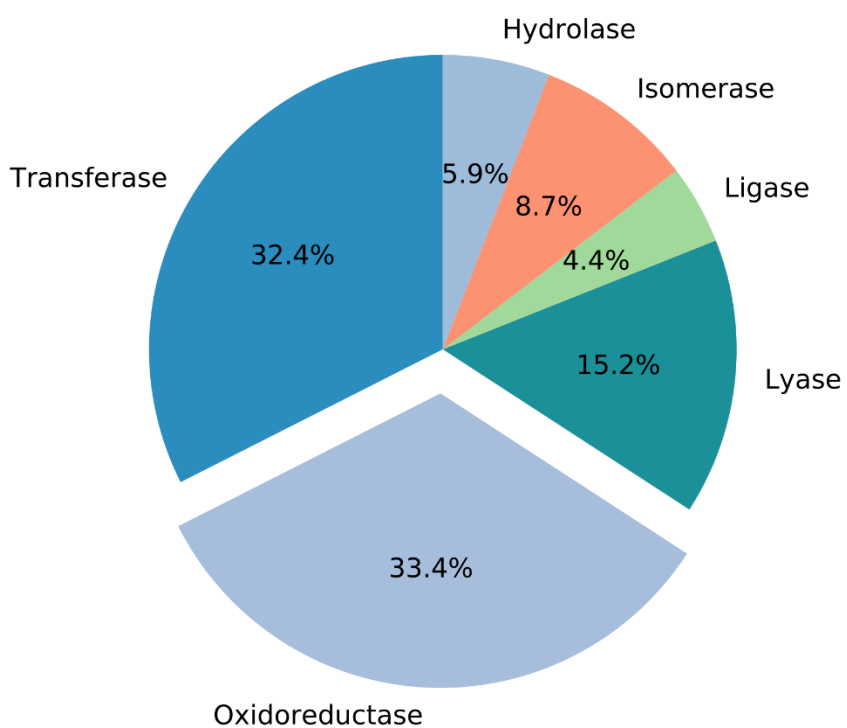


Figure 2 the PDBe entry page (pdbe.org/5hht) where the newly identified cofactor molecules are displayed under the “Ligand and Environment” section.

## PDBe > 5hht

Crystal structure of E. coli transketolase triple variant  
 Ser385Tyr/Asp469Thr/Arg520Gln

Source organism: *Escherichia coli* K-12

### Primary publication:

Structural Analysis of an Evolved Transketolase Reveals Divergent Binding Modes.

Affaticati PE, Dai SB, Payongsri P, Hailes HC, Tittmann K, Dalby PA

Sci Rep 6 35716 (2016)

PMID: 27767080

## Function and Biology

Details

### Reaction catalysed:

Sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate = D-ribose 5-phosphate + D-xylulose 5-phosphate

### Biochemical function:

- transferase activity

### Biological process:

- cellular response to antibiotic

### Cellular component:

- protein-containing complex

### Sequence domains:

- Transketolase-like, pyrimidine-binding domain
- Transketolase C-terminal/Pyruvate-ferredoxin oxidoreductase domain II
- Transketolase, N-terminal
- Thiamin diphosphate-binding fold
- Transketolase, bacterial-like
- Transketolase family
- Transketolase, C-terminal domain

## Ligands and Environments

### Cofactor:



2 x TDP

### 2 bound ligands:



36 x EDO

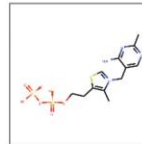


2 x CA

No modified residues

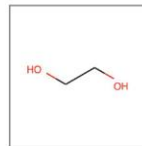
## Ligands and Environments

### Cofactor:

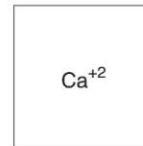


2 x TDP

### 2 bound ligands:



36 x EDO



2 x CA

Figure 3 Cofactor class and the similarity to the template molecule are shown on the individual ligand page (<http://www.ebi.ac.uk/pdbe/entry/pdb/5hht/bound/TDP>).

### 5hht > TDP

#### THIAMIN DIPHOSPHATE

**Formula:**  $C_{12} H_{18} N_4 O_7 P_2 S$

**Molecular weight:** 424 Da

**Putative function:** Cofactor

Cofactor class: *Thiamine diphosphate*

Similarity to cofactor template (TDP): 1.0

#### THIAMIN DIPHOSPHATE

**Formula:**  $C_{12} H_{18} N_4 O_7 P_2 S$

**Molecular weight:** 424 Da

**Putative function:** Cofactor

Cofactor class: *Thiamine diphosphate*

Similarity to cofactor template (TDP): 1.0

#### Environment details

TDP 711 bound to chain B

