

1 ORIGINAL PAPER

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

The *Acanthamoeba* Shikimate Pathway has a Unique Molecular Arrangement and is Essential for Aromatic Amino Acid Biosynthesis

Fiona L. Henriquez^a, Sara J. Campbell^{a,b}, Bharath K. Sundararaj^{a,b}, Antonella Cano^b,
Stephen P. Muench^c, and Craig W. Roberts^{b,1}

^aInstitute of Biomedical Science and Environmental Health Research, School of Science, University of the West of Scotland, Paisley, PA1 2BE, UK

^bStrathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, 161 Cathedral Street, Glasgow, G4 0NR, UK

^cSchool of Biomedical Sciences, University of Leeds, Leeds, LS2 9JT

Submitted June 22, 2014; Accepted October 29, 2014

Running title: *Acanthamoeba* Shikimate Pathway

¹Corresponding author; e-mail c.w.roberts@strath.ac.uk (C. W. Roberts).

23 **The shikimate pathway is the only known biosynthetic route for de novo synthesis of aromatic**
24 **compounds. It is described as an ancient eukaryotic innovation that has been retained in a**
25 **subset of eukaryotes, replaced in plants through the acquisition of the chloroplast, but lost in**
26 **many including humans. Herein, we demonstrate that *Acanthamoeba castellanii* possesses the**
27 **shikimate pathway by biochemical and a combination of bioinformatics and molecular**
28 **biological methods. The growth of *A. castellanii* (Neff strain and a recently isolated clinical**
29 **specimen, both T4 genotypes) is inhibited by glyphosate [N-(phosphonomethyl) glycine], an**
30 **inhibitor of EPSP synthase and the addition of phenylalanine and tryptophan, which are**
31 **dependent on the shikimate pathway, rescued *A. castellanii* from glyphosate indicating that**
32 **glyphosate was specific in action. *A. castellanii* has a novel complement of shikimate pathway**
33 **enzymes including unique gene fusions, two Type I and one Type II DAHP synthases (for**
34 **which their likely sensitivities to feedback inhibition by phenylalanine, tyrosine and**
35 **tryptophan has been modelled) and a canonical chorismate synthase. The shikimate pathway**
36 **in *A. castellanii* therefore has a novel molecular arrangement, is required for amino acid**
37 **biosynthesis and represents an attractive target for antimicrobials.**

38
39 **Key words:** *Acanthamoeba castellanii*; shikimate pathway; glyphosate inhibition; novel
40 molecular arrangement.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

46 Introduction

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Acanthamoeba (species) are facultative parasites that can cause a painful sight-threatening keratitis or fatal granulomatous encephalitis in humans. Although many options for the treatment of ocular *Acanthamoeba* infections exist, most are essentially disinfectants that have necessarily arduous regimens with limited efficacy (Marciano-Cabral and Cabral 2003). Furthermore, eye care solutions are not currently routinely tested against *Acanthamoeba* and no single-step contact lens solution has proven effective at decontaminating lenses from *Acanthamoeba*. Granulomatous Amoebic Encephalitis (GAE) is normally associated with immunosuppression and is almost always fatal as current treatments are generally ineffective (Seijo Martinez et al. 2000). Identification and exploitation of new antimicrobial targets against *Acanthamoeba* is therefore desirable.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

The shikimate pathway is a promising target for antimicrobial design since it is the only known route for de novo synthesis of aromatic compounds. The shikimate pathway was once thought to be restricted to prokaryotes, plants (where it is associated with the chloroplast) and fungi. It has more recently been demonstrated to be present in some chromalveolates including the medically important apicomplexans *Toxoplasma gondii* and *Plasmodium* species, at least some ciliates and the agriculturally important oomycete, *Phytophthora ramorum* (Richards et al. 2006; Roberts et al. 1998). The pathway is now known to be present in phylogenetically divergent extant eukaryotic taxa and it has been proposed that it was therefore present in their last common ancestor (Campbell et al. 2004; Richards et al. 2006). Some additional weight to this is given by the molecular arrangement of the enzymes responsible and their phylogeny. Thus with the exception of phototrophic organisms that have a chloroplast-located, prokaryotic derived pathway, all eukaryotes examined to date have a characteristic pentafunctional gene fusion (arom) encoding a single polypeptide responsible for 5 (3-dehydroquinase, 3-dehydroquinase dehydratase, shikimate 5-dehydrogenase, shikimate kinase and EPSP synthase) out of the seven enzyme activities (Richards et al. 2006). Phylogenetic studies support that this molecular arrangement is an ancient eukaryotic innovation that has been inherited in diverse eukaryotes through vertical descent

72 (Richards et al. 2006). In contrast, bacterial shikimate pathway enzymes are encoded on individual
73 polypeptides (Roberts et al. 1998). This arrangement is conserved in plants with the exception of 3-
74 dehydroquinate dehydratase and shikimate 5-dehydrogenase, which form a bifunctional protein
75 (Campbell et al. 2004). In addition, two non-homologous types of DAHP synthase (designated
76 class I and II) have been found across a variety of organisms (Butler et al. 1974). Class I proteins
77 (often occurring as multiple paralogs in a single organism) were originally identified in prokaryotes,
78 but are also found in many fungi and the chromalveolates, *Phytophthora ramorum* and
79 *Phytophthora infestans* (Coggins et al. 1987; Herrmann and Weaver 1999; Richards et al 2006).
80 Although class II DAHP synthases were first identified in plants, they are now known to be in
81 bacteria, fungi and the chromalveolate *Toxoplasma gondii* (Dyer et al. 1990; Gosset et al. 2001;
82 Richards et al. 2006; Shumilin et al. 1996). A number of prokaryotes and some fungi including
83 *Neurospora crassa* contain both classes of DAHP synthases (Walker et al. 1996). Class I DAHP
84 synthases often existing as multiple paralogs and class II DAHP synthases are frequently regulated
85 by one or more aromatic amino acids (Jensen et al. 2002).

86 The shikimate pathway is still relatively rare in sampled eukaryotes and is absent in
87 mammals that acquire their aromatic compounds from diet. This makes it an attractive target for
88 antimicrobial agents (Bentley 1990; Jensen 1996; Roberts et al. 2002). Herein, we demonstrate that
89 the shikimate pathway is essential for *Acanthamoeba* growth in absence of exogenous aromatic
90 amino acids, describe its molecular organisation and demonstrate its potential as a target for
91 antimicrobials.

92 **Results**

93 **Glyphosate Restricts *Acanthamoeba castellanii* Growth in vitro**

94 The effectiveness of glyphosate as an inhibitor of *A. castellanii* growth was assessed by a modified
95 version of the alamarBlue microtiter plate assay. *A. castellanii* trophozoites were grown in defined
96

97 medium lacking aromatic compounds (folate and the aromatic amino acids tyrosine, phenylalanine
98 and tryptophan). Glyphosate significantly restricted *A. castellanii* Neff strain and the *A. castellanii*
2
3
99 clinical strain growth in a dose dependent manner (Fig. 1). The IC₅₀ for the Neff strain was
4
5
100 determined to be between 17.5μM and 35μM for the Neff strain and 70μM to 140μM for the
6
7
101 clinical strain. Importantly, it did not induce encystment (Fig. 1D).
8
9

10
11
102

12
13

14
103

Glyphosate is not Toxic to Rabbit Corneal Epithelial Cells

16

17
104

18

19
205

21

206

23

24

25
107

26

27

28
208

***Acanthamoeba* Relies on the Shikimate Pathway for Aromatic Amino Acids**

30

31
109

32

33
110

34

35
111

37

38
112

39

40

41
113

42

43
114

44

45
115

46

47
116

48

49
117

50

51
118

52

53
119

54

55
120

56

57

58

59
60
61
62
63
64
65

To confirm that glyphosate was specifically acting on EPSP synthase and thus the shikimate
pathway and to determine which products derived from the shikimate pathway, precursors were
necessary for *A. castellanii*, PABA and aromatic amino acids were added alone and in combination
to glyphosate treated trophozoites. The addition of tyrosine (up to 0.66mM), phenylalanine (up to
10mM) and tryptophan (up to 1.96mM) separately did not negate the effects of glyphosate (Fig.
2A). However, simultaneous addition of phenylalanine and tryptophan ablated any inhibitory effect
of glyphosate (Fig. 2B). Tyrosine, in combination with either phenylalanine or tryptophan, and
PABA were neither necessary nor sufficient to ablate glyphosate inhibition (Fig. 2B, C).
Bioinformatic analyses have identified all enzymes necessary for the production of phenylalanine
(Fig. 3), tryptophan and folates from chorismate (Supplementary Material Fig. S1). Enzymes
involved in tyrosine biosynthesis directly from chorismate are not present, but those necessary for

120 tyrosine biosynthesis from phenylalanine via phenylalanine-4-hydroxylase are present (Genbank
121 ELR14932.1). Notably, the enzymes involved in ubiquinone biosynthesis are not found.

2
3
4
5
6
7

123 ***Acanthamoeba* have both Type I and II DAHP Synthases**

9
1024 The transcriptome data (Clarke et al. 2013) was interrogated and two type I DAHP synthases
11
125 (DAHPI) were identified.(Accession numbers ELR16577.1 and ELR11971.1). Both DAHP
13
14
1526 synthases were successfully amplified from *A. castellanii* Neff strain and clinical specimen cDNAs
16
1727 (Accession numbers KC471625 and KC471626, respectively, Fig. 4A).

18
19
2028 The Phyre2 program produced models for each enzyme and permitted an examination of the
21
229 Phe/Tyr/Trp inhibition site conservation for AcDAHPIa and AcDAHPIb. A model for the type I
23
24
2530 DAHP synthase was generated from both *A.castellanii* sequences based on the *E. coli* crystal
26
2731 structure to which both sequences share very high sequence identity (Fig. 4A). The Phe binding
28
29
3032 pocket is formed by 16 residues, of which only 1 differs in the *E. coli* structure with Ser being
31
3233 replaced by Gly in the AcDAHPIa, but not AcDAHPIb sequence (Fig. 4A). The predicted position
33
34
3534 of Ser allows it to pack against the the Phe substrate with the replacement by Gly in the AcDAHPIa
36
3735 sequence possibly accommodation of the bulkier hydroxyl group of a Tyr residue (Fig. 4C, D). This
38
39
4036 Ser/Gly change may allow for the type I DAHP synthase to be regulated by Tyr or Phe in the
41
4237 AcDAHPIa and AcDAHPIb sequences, respectively.

43
4438 DAHP synthase type II (DAHPII) was found in the *A. castellanii* transcriptome (Accession
45
46
4739 number ELR24167) (Clark et al. 2013). The putative *A. castellanii* DAHPII shares a high degree of
48
4940 similarity to a number of type II DAHP synthases. Modelling of the *A. castellanii* DAHP type II
50
51
5241 (AcDAHPII) synthase was based on *M. tuberculosis* DAHP typeII (MtDAHPII), which has been
53
5442 solved in complex with both Phe and Trp inhibitors (Webby et al. 2010). The MtDAHPII contains a
55
56
5743 primary Phe binding site located in the heart of the dimer interface, formed by residues from each
58
5944 monomer. Of the 10 residues which make up the binding pocket only 2 are conserved with a further

60
61
62
63
64
65

145 2 showing close similarity (Fig. 5A), with 2 residues found on the C-terminal β -strand which is
146 predicted to be structurally absent in the AcDAHPII structure (Fig. 4A). Moreover, those residues
147 responsible for forming a hydrogen bond network in MtDAHPII (Arg171 and Asn175) are replaced
148 by Gln and His, respectively (Fig. 5B). The replacement of Ala in MtDAHPII to Phe in AcDAHPII
149 in the predicted Phe binding site, may also create a steric clashes and would make dimer formation
150 in the presence of Phe unlikely (Fig. 5C). The secondary Phe binding site within MtDAHPII
151 involves 9 residues of which Arg23, Arg256 and Glu53 have been shown to be critical in forming a
152 hydrogen bond network to the α -carboxylate group of the Phe inhibitor, none of these are conserved
153 in AcDAHPII but are replaced by Thr, Ala and Gln, respectively. This significant change would
154 make it unlikely that the AcDAHPII enzyme is inhibited by Phe in a manner similar to that seen for
155 MtDAHPII. In addition to Phe, MtDAHPII is also inhibited by Trp which binds in a site formed by
156 17 residues of which 6 are identical and 4 are similar in the AcDAHPII enzyme. Modelling suggests
157 that AcDAHPII would have a similar hydrophobic binding pocket of the same approximate shape.
158 Interestingly, MtDAHPII is inhibited only by the presence of both Phe/Tyr and Trp at biologically
159 relevant concentrations. The predicted lack of a conserved Phe/Tyr binding site may imply that
160 AcDAHPII is inhibited only by Trp and does not require the allosteric binding of Phe or Tyr.
161 Alternatively Trp does not regulate the enzyme, or an alternative binding site is present which has
162 so far not been characterised. Further biochemical and structural investigation is required to validate
163 these models which give only a prediction of the *A. castellanii* DAHP family.

165 ***Acanthamoeba* have a Novel Tetrafunctional Protein with Tetrafunctional Gene Fusion**
166 **Comprising 3-dehydroquinate Synthase (DHQS), 5-enolpyruvylshikimate-3-phosphate**
167 **Synthase (EPSPS), Shikimate Kinase (SK) and 3-dehydroquinate Dehydratase (DH)**

168 Bioinformatical analysis of the transcriptome data (Clarke et al. 2013) reveals that four out of five
169 of the AROM enzymes are present in a tetrameric fusion and that the shikimate dehydrogenase is

170 missing from the sequence. Accession number ELR24870 is named shikimate kinase but it actually
171 contains 3-hydroquinase synthase (DHQS), 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS),
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

missing from the sequence. Accession number ELR24870 is named shikimate kinase but it actually contains 3-hydroquinase synthase (DHQS), 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), shikimate kinase (SK) and 3-dehydroquinase dehydratase (DH). This unusual and novel arrangement was confirmed by PCR from cDNA from both *A. castellanii* Neff strain and the clinical strain (KC479580 and KC471623, respectively; Fig. 6). The AROM-like sequence (aromN –Novel-) is 3927 nucleotides in length, and codes for protein with a predicted molecular weight of 143.3kDa and shares identity and similarity with a number of AROM proteins from other species (Supplementary Material Fig. S2).

279 ***Acanthamoeba castellanii* have a Novel Trifunctional Gene Fusion Comprising Shikimate**
280 **Dehydrogenase, Phosphoribosylanthranilate Isomerase and Indole-3-glycerol-phosphate**
281 **Synthase**

302 Analysis of the transcriptome data (Clarke et al. 2013) confirmed the existence of a novel
303 trifunctional gene fusion comprising of indole-3-glycerol-phosphate synthase,
304 phosphoribosylanthranilate isomerase and shikimate dehydrogenase (Accession number
305 ELR21144). This fusion was confirmed by PCR in both strains studied (Accession numbers
306 KC479581 and KC471624, Neff and clinical isolates, respectively; Supplementary Material Fig.
307 S3). Modelling of the trimeric fusion proved unreliable due to the absence of a current complex
308 structure. However, two models could be reliably generated of the bi-functional 2
309 phosphoribosylanthranilate isomerase:3 indoleglycerolphosphate (RCSB ID. 1pii) and bi-functional
310 3–dehydroquinase dehydratase/shikimate dehydrogenase (RCSB ID 2o7q). The common fold
311 shared between the bi-functional complexes and their subsequent superposition permits for an
312 approximate model of the trimeric fusion to be generated.

194 *Acanthamoeba* have a Canonical Chorismate Synthase

1
195 Based on these sequences primers (chorismate synthase For 5'-GGTGCAGGCGACAAGGCAAC-
3
196 3' and chorismate synthase REV 5'-AAGTACTTTTTCTGGAGCTG-3') were designed and hot
5
6
197 start, which spans both contigs identified. *A. castellanii* chorismate synthase is present in the
8
198 transcriptome (Accession number ELR21143) (Clarke et al. 2013). PCR (annealing at 58 °C)
10
199 amplified the gene encoding chorismate synthase from *A. castellanii* Neff strain cDNA (Accession
12
200 number KC471628) and subsequently from the clinical isolate (Accession number KC471627).
15
201 AcCS shares identity and similarity with a number of type I chorismate synthases from other
17
18
202 species (Supplementary Material Fig. S4).
20
21
22
23
24

204 **Discussion**

25
26
27
28
295 Our work has demonstrated that the shikimate pathway has a wider eukaryotic phylogenetic
30
31
206 distribution than previously thought (Campbell et al. 2004; Richards et al. 2006). Thus, the
32
33
207 shikimate pathway would appear to have been present in the last universal common ancestor of
34
35
208 eukaryotes and has been selectively lost in the vast majority of extant eukaryotes sampled to date,
36
37
209 replaced in plants predominantly through the acquisition of the chloroplast and retained by fungi,
38
39
40
210 and a select few of sampled chromalveolates including some apicomplexans, ciliates and oomycetes
41
42
211 (Campbell et al. 2004; Richards et al. 2006). The current studies now demonstrate the shikimate
43
44
212 pathway in at least one member of the Amoebozoa. This would support the notion that the
45
46
213 shikimate pathway is an ancient eukaryotic trait. The ability of glyphosate to inhibit the growth of a
47
48
214 laboratory strain of *A. castellanii* (Neff) and a recently isolated clinical strain demonstrates the
49
50
215 potential utility of targeting this pathway with antimicrobials.
51
52
53
54
55

216 The shikimate pathway catalyses the sequential conversion of erythrose 4-phosphate and
56
57
217 phosphoenol pyruvate to chorismate in seven steps. The studies described here fully characterise
58
59
60
61
62
63
64
65

218 and report the existence of the shikimate pathway in *A. castellanii*. The pathway has the potential
219 to be an antimicrobial target used to limit infection and the genes encoding the enzymes of this
220 pathway have a novel molecular arrangement.

221 The potential for the shikimate pathway to be an antimicrobial target is exemplified by
222 glyphosate which restricts *A. castellanii* growth in a dose-dependent manner, albeit in the absence
223 of the pathway's end products (aromatic amino acids, folate and ubiquinone). Interestingly, only a
224 combination of phenylalanine and tryptophan or all three aromatic amino acids together can rescue
225 *A. castellanii* growth from the effects of glyphosate. Bioinformatic analysis reveals that *A.*
226 *castellanii* can create phenylalanine and tryptophan directly from chorismate, whereas the important
227 enzyme, prephenate dehydrogenase in tyrosine synthesis from chorismate is not present. However,
228 tyrosine can be synthesised from phenylalanine via the enzyme phenylalanine-4-hydroxylase. *A.*
229 *castellanii* also possess the ability to synthesise folates. However, the addition of folates alone did
230 not ablate the inhibitory effect of glyphosate. A plausible explanation to the reason that aromatic
231 amino acids can rescue *A. castellanii* growth in the absence of folate but the reverse is not possible
232 may be due to the fact that folates can be recycled through gammaglutamyl hydrolase (EC
233 3.4.19.9), which has been identified in *A. castellanii* (ELR19434) (Clarke et al. 2013). In our
234 experiments, it is likely that *A. castellanii* are recycling folate, since they are initially maintained in
235 a rich mycological peptone medium before a step by step passage to a defined minimal medium.
236 Enzymes involved in ubiquinone biosynthesis have not been found despite extensive
237 bioinformatical searches. Future studies will be required to determine how *A. castellanii* acquires
238 ubiquinone.

239 We have for the first time identified two class I DAHP synthases and a class II DAHP
240 synthase in *A. castellanii*. This is similar to the situation in *Neurospora crassa* where it has both a
241 tyrosine sensitive and a phenylalanine sensitive class I DAHP synthase along with a class II DAHP
242 synthase that is sensitive to tryptophan (Chaleff et al. 1974). Modelling of the two *A. castellanii*
243 class I DAHP sequences has suggested the apparent conservation of the Phe binding site with the

244 replacement of Ser to Gly in the binding site possibly permitting the two *A. castellanii* class I
245 DAHP synthases to be regulated by both Phe and Tyr (Fig. 3), although further biochemical studies
246 are required to validate this. Unlike bacteria, which contain an AroF, AroG and AroH class I
247 DAHP synthase, which can be regulated by Tyr, Phe and Trp, respectively, *A. castellanii* has only
248 the AroG homologue. The two different AroG sequence, may allow for a Phe or Tyr regulation
249 mechanism although no clear Trp site can be identified. Further biochemical characterisation will be
250 required to show that both Tyr and Phe can regulate the shikimate pathway through the two class I
251 enzymes.

252 In addition to the class I DAHP enzymes a second class of DAHP enzyme has been found
253 which is distinct in architecture form the class I family. This second class has been shown to be
254 regulated by Phe, Tyr and Trp with either Phe/Tyr being required in the presence of Trp to cause
255 inhibition. The structure of the *M. tuberculosis* class II DAHP enzyme (MtDAHPII) has
256 characterised the architecture of both the Phe/Tyr and Trp binding pockets. Modelling of *A.*
257 *castellanii* homologues suggests it may lack the pre-requisite amino acids, which form the binding
258 site with an Ala/Phe mutation causing a severe steric clash with the proposed bound Phe. The Trp
259 binding site appears well conserved within the *A. castellanii* class II DAHP (AcDAHPII) although
260 whether this binding requires the cooperative binding of Phe/Tyr as well is as yet undetermined.
261 The putative modelling has predicted that *A. castellanii* has 3 DAHP enzymes, (2 class I and 1 class
262 II), of these the class I enzymes can be regulated by Phe/Tyr and the class II by Trp. Other binding
263 sites may exist which have not yet been characterised. However, these results show that the
264 shikimate pathway in *A. castellanii* could be feedback regulated by 3 different aromatic amino acids
265 as described in *Neurospora crassa* (Chaleff et al. 1974).

266 The shikimate pathway enzymes have distinct molecular organisations in different groups of
267 species (Chaleff et al. 1974). Each of the shikimate pathway enzymes is encoded on individual
268 polypeptides in bacteria (Butler et al. 1974). Plants have discrete polypeptides with each enzyme
269 activity with the exception of 3-dehydroquinate dehydratase and shikimate 5-dehydrogenase which

270 are fused as separate domains on a single polypeptide (Coggins et al. 1987). In all eukaryotes
271 previously examined, 3-dehydroquinate synthase, 3-dehydroquinate dehydratase, shikimate 5-
272 dehydrogenase, shikimate kinase and EPSP synthase are encoded on a single transcript to form a
273 pentafunctional polypeptide (termed AROM) (Richards et al. 2006). The current studies now
274 describe a further novel arrangement as *A. castellanii* consisting of two previously undescribed gene
275 fusions. Thus, *A. castellanii* have a novel tetrafunctional protein comprising 3-dehydroquinate
276 synthase, 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase, 3-dehydroquinate dehydratase and
277 shikimate kinase. This fusion has the four enzyme activities ordered in the same sequence as the
278 first 4 activities of AROM. We have designated this tetrafunctional protein AROMn ('n' is not only
279 next in the alphabet, but to also denotes this novel arrangement). The enzyme activity missing from
280 the AROM-like protein, shikimate dehydrogenase is instead part of a novel trifunctional protein
281 comprising shikimate dehydrogenase, phosphoribosylanthranilate isomerase and indole-3-glycerol-
282 phosphate synthase. Although a gene fusion comprising shikimate dehydrogenase with these 2
283 tryptophan biosynthesis enzymes is novel, certain eukaryotes including *Phaeodactylum tricornutum*
284 (a diatom), *P. ramorum* (an oomycete) and *Ustilago maydis* (a fungus) have a fused
285 phosphoribosylanthranilate isomerase and indole-3-glycerol-phosphate synthase (Jiroutová et al.
286 2007). The functional implications, if any of this fusion is yet to be determined, but its unique
287 arrangement could prove to be a novel target for therapeutic development. These novel gene fusions
288 will require phylogenetic analyses to determine their likely evolutionary origins.

289 Chorismate synthase is responsible for the final reaction of the shikimate pathway and is
290 dependent on reduced flavin for its function. Chorismate synthase is known to have an endogenous
291 flavin reductase capability in fungi, apicomplexans and ciliates examined to date (Ehammer et al.
292 2007). However, in plants and bacteria chorismate synthase has been found to lack flavin reductase
293 ability and is thus dependent on an alternative source. From chorismate a number of aromatic
294 products are derived, including folate, ubiquinone and aromatic amino acids. All enzymes

295 belonging to these pathways have been identified through bioinformatical analysis (Clarke et al.
296 2013) (Supplementary Material Fig. S1).

297 The shikimate pathway has been exploited in plants as a target for herbicides (Bentley,
298 1990) and proposed as an antimicrobial target in a number of pathogens including bacteria, fungi
299 and protozoans. In addition to glyphosate and number of compounds targeting EPSP synthase have
300 been experimentally tested as antimicrobials (Bentley 1990). Other enzymes in the shikimate
301 pathway or downstream of the shikimate pathway have been inhibited by substrate analogues. Thus
302 for example, (6S)-6-fluoroshikimate and (6R)-6-fluoroshikimate through a process of
303 biotransformation ultimately inhibit 4-aminobenzoic acid synthesis and chorismate synthase
304 respectively (Bentley 1990). The work described here suggests that the current antimicrobials
305 designed against the shikimate pathway have potential against *A. castellanii* and may provide
306 further motivation to develop new inhibitors of this pathway. These inhibitors could be used to
307 improve the treatment of *A. castellanii* and improve the ability of contact lens solutions to kill this
308 potential pathogen. It is also likely that the shikimate pathway is also present in other
309 *Acanthamoeba* species based on genetic similarities (Henriquez et al. 2008).

310 In conclusion we demonstrate that *A. castellanii* have a novel shikimate pathway with as yet
311 unique gene fusions. This pathway is essential for the production of aromatic amino acids and the
312 survival of *A. castellanii* and may be exploitable as an antimicrobial agent target.

314 **Methods**

315
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

316 **Maintenance of *Acanthamoeba castellanii* trophozoites:** *Acanthamoeba castellanii* (*A.*
317 *castellanii*) (Neff strain) was originally obtained from Keith Vickerman (Glasgow, United
2
318 Kingdom). *A. castellanii* T4 (Clinical isolate), isolated from a patient with keratitis was obtained
3
4
5
319 from Antonella Mattana (University of Sassari, Italy) *A. castellanii* trophozoites were routinely
6
7
320 grown in 2% mycological peptone w/v (Sigma, Poole, United Kingdom) and 0.9% w/v maltose
8
9
10
321 (Sigma), or M11 media as previously described by Shukla *et al.*, (1990) modified by removal of
11
12
322 aromatic compounds and augmentation of glucose levels to 36% w/v. Media were supplemented
13
14
15
323 with 125µg penicillin/ streptomycin (Sigma). They were incubated until confluent at room
16
17
324 temperature, unless otherwise stated, in 75-cm² tissue culture flasks when they were either
18
19
20
325 subcultured or harvested following mechanically induced detachment.
21
22

23 **Rabbit corneal epithelial cells:** Rabbit corneal epithelial (RCE) cells (ECACC No.
24
25
327 950810146) were cultured in Dulbecco's Modified Eagle's Medium (DMEM) Ham's F-12 (1:1)
26
27
328 containing 15mM HEPES and L-glutamine, supplemented with human corneal growth supplement
28
29
30
329 (HCGS) (Life Technologies Paisley, United Kingdom), 10% heat-inactivated foetal calf serum
31
32
330 (Sigma), 125µg penicillin/ streptomycin and 125µg amphotericin B (Sigma). RCE cells were
33
34
331 incubated at 37 °C, 5% CO₂ in 75-cm² tissue culture flasks until 90-95% confluent, then harvested
35
36
37
332 or sub-cultured by using TrypLE™ Express enzyme cell detaching medium (Life Technologies).
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

333 ***A. castellanii* trophozoite growth inhibition assays:** *A. castellanii* trophozoites were
334 routinely cultured as described above. Confluent flasks were subcultured by mechanical detachment
335 to increasing concentrations of modified M11 media (50%, 80%, 100%), supplemented with 2%
336 mycological peptone and 0.9% maltose and with 125µg penicillin/ streptomycin and 125µg
337 amphotericin B (Sigma, Poole, United Kingdom). The effectiveness of N-(phosphonomethyl)
338 glycine (glyphosate) as an inhibitor of *A. castellanii* trophozoites was determined by a modified
339 version of the colorimetric microtiter plate assay described by McBride *et al.*, (2005). *A. castellanii*
340 cultured in 100% modified M11 media were seeded in triplicate at a concentration of either 4 x
341 10⁵/ml (Neff Strain) or 1 x 10⁶ (Clinical isolate) in a 96 well tissue culture plate (TPP, Switzerland).
342 Glyphosate (Sigma) was dissolved in modified M11 media and 50µl was added to each well in
343 serial dilutions from 1130µM to 1.1µM. Control wells were supplemented with 50µl of modified
344 M11 medium. Where appropriate, aromatic amino acids and/or para-aminobenzoic acid (PABA)
345 were added separately at different concentrations as stated to determine if such treatment could
346 ablate the effects of glyphosate. The assay was then incubated for 48 hours at room temperature and
347 10µl of AlamarBlue reagent (Biosource, Europe) was added to each well and further incubated at
348 room temperature, in the dark, for 24 hours. Absorbances were then read on spectromax (Molecular
349 Devices) at OD₅₇₀ and OD₆₀₀.

350 The percentage inhibition of AlamarBlue™ was calculated by the following formula:

$$351 \quad \{[(\epsilon_{ox}\lambda_2)(A \lambda_1)] - [(\epsilon_{ox}\lambda_1)(A \lambda_2)] / [(\epsilon_{ox}\lambda_2)(A {}^o\lambda_1)] - [(\epsilon_{ox}\lambda_1)(A {}^o\lambda_2)]\} \times 100$$

352 Where $\epsilon_{ox}\lambda_1$ is 80,586 (molar extinction coefficient of oxidised AlamarBlue at 570nm); $\epsilon_{ox}\lambda_2$ is
353 117,216 (molar extinction coefficient of oxidised AlamarBlue at 600nm); $A \lambda_1$ is the absorbance of
354 the treated wells at 570nm; $A \lambda_2$ is the absorbance of the treated wells at 600nm; $A {}^o\lambda_1$ is the
355 absorbance of the untreated control wells at 570nm; $A {}^o\lambda_2$ is the absorbance of the untreated control
356 wells at 600nm. These absorbance values were subtracted from 100 to give percentage of

357 AlamarBlue™ reduction in comparison to untreated controls. The results were expressed as a mean
358 for each triplicate ± the standard error (SE) and student *T*-test analyses were performed.

359 **Assessment of inhibitor-induced encystment:** *A. castellanii* were cultured, harvested and
360 seeded as described previously with glyphosate (Sigma) dissolved in modified M11 media and
361 incubated for 72 hrs at room temperature. Sodium dodecyl sulphate (SDS), which causes
362 immediate lysis of *A. castellanii* trophozoites while mature cysts remain intact, was then used to
363 assess if the presence of glyphosate had caused *A. castellanii* to form cysts. The medium was
364 carefully removed and replaced with 100µl of 5% SDS then cells were counted manually on a
365 haemocytometer. The effect of 5% SDS was also assessed on cyst controls. In brief 4 x 10⁵/ml *A.*
366 *castellanii* were resuspended in encystment medium (20 mM Tris-HCl [pH 8.8], 100 mM KCl, 8
367 mM MgSO₄, 0.4 mM CaCl₂, 1 mM NaHCO₃) (Campbell et al. 2008) and incubated at room
368 temperature until cysts had formed. Cells were then resuspended in 5% SDS and counted manually
369 on a haemocytometer.

370 **Susceptibility of rabbit corneal epithelial (RCE) cells to glyphosate:** RCE were cultured
371 and harvested as described. The effect of glyphosate was determined as described by McBride et al.
372 (2007). Cells were seeded in triplicate at 280 cells per well in 50µl of RCE medium in a 96 well
373 tissue culture plate (TPP, Switzerland) and allowed to adhere for 3 hours at 37 °C, 5% CO₂.
374 Glyphosate (Sigma) was dissolved in 10ml of RCE medium and 50µl was added to each well to
375 give the final concentrations from 1130µM to 17.5µM. Control wells were supplemented with 50µl
376 of RCE medium and plates were incubated for 96 hours at 37°C, 5% CO₂. 10µl of alamarBlue
377 reagent was added to wells 6 hours prior to the end of the incubation. Absorbance was then read on
378 spectromax (Molecular Devices) at OD₅₇₀ and OD₆₀₀ and the percentage inhibition of alamarBlue
379 reduction was calculated as previously described.

380 **RNA extraction:** *A. castellanii* were cultured and harvested as normal from modified M11
381 medium. Cells were collected by a 5 minute centrifugation at 21000g at 4 °C and the pellet was
382 then suspended in 1.0ml of Trizol® reagent (Life Technologies). Total RNA was isolated using a

383 method based on a single-step acid guanidinium thiocyanate-phenol-chloroform protocol described
384 by Chomczynski & Sacchi (1987). The concentration was determined by measuring absorbance at
2
385 260 nm on a spectrophotometer (GeneQuant pro, Amersham Biosciences, and United Kingdom).
4
5
386 Extracted RNA was then stored at -80 °C until required for cDNA synthesis and the integrity of the
7
387 RNA was assessed via PCR and or gel electrophoresis on a 2% agarose gel.
9

10
388 ***A. castellanii* genomic DNA extraction:** Genomic DNA (gDNA) was obtained by a
12
389 modification of the method of Johnson et al. (1986). In brief, 10⁷ cells were collected by 10-min
14
390 centrifugation at 7000g and 4 °C. The pellet was resuspended in 10ml of 0.2M NaCl/10mM
15
391 EDTA/10mM Tris-HCl (pH 8.0), 1% SDS, 200µg/ml Proteinase K (Sigma, Poole, United
19
392 Kingdom), and incubated 3 hours at 50 °C with gentle agitation. 10ml of
21
22
393 phenol/chloroform/isoamyl alcohol (50:48:2) was added before centrifugation for 10 minutes at
24
394 3000g and 4 °C. The supernatant was removed and this was repeated until the interface was clear.
26
27
395 The DNA solution was extracted with 10ml of chloroform and centrifugation for 10 minutes at
28
396 3000g at 4 °C before the addition of 20µg/ml of RNase A and a further incubation at 4 °C for 15
31
397 minutes. DNA was precipitated with 10M ammonium acetate and 100% ethanol at -20 °C for 1hr
34
398 followed by a 10 minute centrifugation at 10,000g and 4 °C. The pellet was resuspended in 0.5 ml
36
399 of TE buffer (10mM Tris-HCL, 1mM EDTA, pH 8.0) and the concentration determined by
38
400 measuring absorbance at 260 nm on a spectrophotometer (GeneQuant pro, Amersham Biosciences,
41
401 United Kingdom). The integrity of the DNA was assessed via gel electrophoresis on a 0.8%
43
44
402 agarose gel. All samples were stored at 4 °C until required.
46

47
403 **Complementary DNA (cDNA) synthesis and Polymerase chain reaction (PCR):**
48
49
404 Complementary DNA (cDNA) was synthesised from *A. castellanii* total RNA, using AffinityScript
51
405 as per manufacturer instructions (Stratagene, Cambridge, UK). All samples were incubated 55 °C
53
406 for 60 minutes and 70 °C for 15 minutes with the exception of random primers, which were pre
55
407 incubated at 25 °C for 10 minutes. Synthesised cDNA was then stored at -20°C until required.
58
59
408 Oligonucleotides were designed using MacVector™ (Oxford, Biomolecular, UK) and synthesized
60
61
62
63
64
65

409 commercially by Life Technologies Ltd. (Paisley, UK). All standard PCR amplification reactions
410 were performed in 25µl reactions. Each reaction contained 1µl 10X High Fidelity PCR Buffer, 1µl
411 10 mM dNTP mixture, 1µl 50 mM MgSO₄, 25 pmol forward and reverse oligonucleotide primers,
412 18µl molecular H₂O, and 1µl of *A. castellanii* cDNA or gDNA. The samples were then incubated at
413 94 °C for 1 minute before the addition of 0.5µl Platinum® Taq High Fidelity (Life Technologies).
414 25pmol forward and reverse oligonucleotide primers, 10.5µl molecular H₂O (Life Technologies),
415 1µl of *A. castellanii* cDNA or gDNA. The reactions were performed with initial denaturing at 95°C
416 for 3-5 minutes followed by 35 or 40 cycles of denaturing at 95 °C for 30 seconds, annealing at 52
417 °C-64 °C for 45 seconds and extension at 72 °C for 1-2 minutes. These reactions were then
418 completed with a final extension at 72 °C for 10 minutes. PCR products were visualised in a 0.8 -
419 2% agarose gel on a transilluminator following ethidium bromide staining.

420 **Cloning and sequencing of PCR products:** PCR-amplified DNA fragments were isolated
421 from ethidium bromide stained agarose gels via MinElute Gel Extraction Kit (Qiagen, Crawley,
2 United Kingdom). The purified PCR amplified products were ligated into the pDRIVE vector using
422 the Qiagen PCR Cloning Kit (Qiagen) according to the manufacturer's instructions. Competent
5 DH5 α were transformed with 5 μ l of the ligation reaction using the heat shock method (Cohen et al.
7 1972). Transformed cells were then spread evenly onto LB agar (Sigma, Poole, United Kingdom),
425 which had been previously coated with 100 μ g/ml ampicillin, 500 μ M/ml IPTG and 50 μ g/ml X-gal
14 (Sigma, Poole, United Kingdom) and incubated overnight at 37 °C. Successful transformants were
15 screened by blue/white screening and plasmid purification using the QIAprep Spin Miniprep Kit
16 (Qiagen) according to the manufacturer's instructions. Restriction enzyme digest with EcoR1
17 Digested fragments were assessed via agarose gel electrophoresis. Automated sequencing of PCR
430 amplified products was achieved using the M13 forward and reverse primers and performed
24 commercially by Geneservice Ltd., Cambridge, UK or GATC in accordance with company
25 guidelines. Sequences were then assembled using SequencherTM 4.0 (GeneCodes, USA).
26
27
28
29
30
31
32

33 **Modelling:** Modelling of the *A. castellanii* shikimate pathway proteins was done using the
34 Phyre2 server (Kelley and Sternberg 2009) and proposed ligand binding sites detected using the
35 3DLigand server (Wass et al. 2010).
36
37
38
39
40
41
42

438 **Acknowledgements**

439 This work was supported by the WH Ross Foundation for Research to Prevent Blindness; The
49 University of Strathclyde Research and Development Fund; The University of the West of Scotland,
50 Innovation and Research Office Support Grant; and the Carnegie Trust for Universities of Scotland.
51 SPM is supported by an MRC-career development award (G100567).
52
53
54
55
56
57
58
59
60
61
62
63
64
65

444 **References**

- 1
445 **Ahn HJ, Yoon HJ, Lee B 2nd, Suh SW** (2004) Crystal structure of chorismate synthase: a novel
3
446 FMN-binding protein fold and functional insights. *J Mol Biol* **336**:903-915
5
6
447
8
448 **Bentley R** (1990) The shikimate pathway--a metabolic tree with many branches. *Crit Rev Biochem*
10
449 *Mol Biol* **25**:307-384
11
12
13
450
15
16
451 **Butler, JR, Alworth, WL, Nugent, MJ** (1974) Mechanism of dehydroquinase catalysed
17
18
452 dehydration 1. Formation of a Schiff base intermediate. *J Am Chem Soc* **96**:1617-1618
20
21
453
22
23
454 **Campbell SJ, Ingram PR, Roberts CW, Henriquez FL** (2008) Induced encystment improves
25
455 resistance to preservation and storage of *Acanthamoeba castellanii*. *Parasitology* **135**:1401-1405
27
28
456
29
30
457 **Campbell SA, Richards TA, Mui EJ, Samuel BU, Coggins JR, McLeod R, Roberts CW** (2004)
32
33
458 A complete shikimate pathway in *Toxoplasma gondii*: an ancient eukaryotic innovation. *Int J*
34
35
459 *Parasitol* **34**:5-13
37
38
460
39
40
461 **Chaleff RS** (1974) The inducible quinate-shikimate catabolic pathway in *Neurospora crassa*:
42
462 induction and regulation of enzyme synthesis. *J Gen Microbiol* **81**:357-372
44
45
463
46
47
48
464 **Chomczynski P, Sacchi N** (1987) Single-step method of RNA isolation by acid guanidinium
49
50
465 thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162**:156-159
51
52
466
54
55
467 **Clarke M, Lohan AJ, Liu B, Lagkouvardos I, Roy S, Zafar N, Bertelli C, Schilde C,**
56
57
468 **Kianianmomeni A, Bürglin TR, Frech C, Turcotte B, Kopec KO, Synnott JM, Choo C,**
59
60
469 **Paponov I, Finkler A, Heng Tan CS, Hutchins AP, Weinmeier T, Rattei T, Chu JS, Gimenez**
61

470 **G, Irimia M, Rigden DJ, Fitzpatrick DA, Lorenzo-Morales J, Bateman A, Chiu CH, Tang P,**
471 **Hegemann P, Fromm H, Raoult D, Greub G, Miranda-Saavedra D, Chen N, Nash P, Ginger**
2
3
472 **ML, Horn M, Schaap P, Caler L, Loftus BJ** (2013). Genome of *Acanthamoeba castellanii*
4
5
473 highlights extensive lateral gene transfer and early evolution of tyrosine kinase signaling. Genome
7
474 Biol. **14**:R11
9
10
475
11
12
476 **Coggins JR, Duncan K, Anton IA, Boocock MR, Chaudhuri S, Lambert JM, Lewendon A,**
14
15
477 **Millar G, Mousdale DM, Smith DD** (1987). The anatomy of a multifunctional enzyme. Biochem
16
17
478 Soc Trans **15**:754-759
19
20
479
21
22
480 **Cohen SN, Chang AC, Hsu L** (1972) Nonchromosomal antibiotic resistance in bacteria: genetic
24
25
481 transformation of *Escherichia coli* by R-factor DNA. Proc Natl Acad Sci USA **69**:2110-2114
26
27
482
28
29
483 **Dyer WE, Weaver LM, Zhao JM, Kuhn DN, Weller SC, Herrmann KM** (1990) A cDNA
31
32
484 encoding 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase from *Solanum tuberosum* L. J
34
485 Biol Chem. **265**:1608-1614
36
37
486
38
39
487 **Ehammer H, Rauch G, Prem A, Kappes B, Macheroux P** (2007) Conservation of NADPH
41
42
488 utilization by chorismate synthase and its implications for the evolution of the shikimate pathway.
43
44
489 Mol Microbiol **65**:1249-1257
46
47
490
48
49
491 **Gosset G, Bonner CA, Jensen RA** (2001) Microbial origin of plant-type 2-keto-3-deoxy-D-
51
492 arabino-heptulosonate 7-phosphate synthases, exemplified by the chorismate- and tryptophan-
53
54
493 regulated enzyme from *Xanthomonas campestris*. J Bacteriol **183**:4061-4070
55
56
494
58
59
60
61
62
63
64
65

495 **Henriquez, FL, Ingram, PR, Muench SP, Rice DW, Roberts CW** (2008) Molecular basis for
496 resistance of *Acanthamoeba* tubulins to all major classes of antitubulin compounds. *Antimicrob*
2
3
497 *Agents Chemother* **52**:1133-1135
4
5
498
6
499 **Herrmann KM, Weaver LM** (1999) The shikimate pathway. *Annu Rev Plant Physiol Plant Mol*
9
10
500 *Biol* **50**:473-503
11
12
501
14
15
502 **Jensen RA** (1996) Tyrosine and phenylalanine biosynthesis: Relationship between alternative
16
17
503 pathways, regulation and subcellular location. In Conn EE (ed), *Recent Advances in*
19
20
504 *Phytochemistry*, vol. 20, *The Shikimic Acid Pathway*. Plenum Press, New York and London, pp
21
22
505 57–81
24
25
506
26
27
507 **Jensen RA, Xie G, Calhoun DH, Bonner CA** (2002) The correct phylogenetic relationship of
28
29
508 KdsA (3-deoxy-d-manno-octulosonate 8-phosphate synthase) with one of two independently
30
31
509 evolved classes of AroA (3-deoxy-d-arabino-heptulosonate 7-phosphate synthase). *J Mol Evol*
32
33
510 **54**:416-423
34
35
511
36
37
512 **Jiroutová K, Horák A, Bowler C, Oborník M** (2007) Tryptophan biosynthesis in stramenopiles:
38
39
513 eukaryotic winners in the diatom complex chloroplast. *J Mol Evol* **65**:496-511
40
41
42
514
43
44
515 **Johnson AM, Dubey JP, Dame JB** (1986) Purification and characterization of *Toxoplasma gondii*
46
47
516 tachyzoite DNA. *Aust J Exp Biol Med Sci* **64**:351-355
48
49
50
51
517
52
53
518 **Kelley LA, Sternberg MJ** (2009) Protein structure prediction on the Web: a case study using the
54
55
519 Phyre server. *Nat Protoc* **4**:363-371
56
57
58
59
520
60
61
62
63
64
65

521 **Marciano-Cabral F, Cabral G** (2003) *Acanthamoeba* spp. as agents of disease in humans. Clin
522 Microbiol Rev **16**:273-307
523
524 **McBride J, Ingram PR, Henriquez FL, Roberts CW** (2005) Development of colorimetric
525 microtiter plate assay for assessment of antimicrobials against *Acanthamoeba*. J Clin Microbiol
526 **43**:629-634
527
528 **McBride J, Mullen AB, Carter KC, Roberts CW** (2007) Differential cytotoxicity of phospholipid
529 analogues to pathogenic *Acanthamoeba* species and mammalian cells. J Antimicrob Chemother
530 **60**:521-525
531
532 **Richards TA, Dacks JB, Campbell SA, Blanchard JL, Foster PG, McLeod R, Roberts CW**
533 (2006) Evolutionary origins of the eukaryotic shikimate pathway: gene fusions, horizontal gene
534 transfer, and endosymbiotic replacements. Eukaryot Cell **5**:1517-1531
535
536 **Roberts CW, Roberts F, Lyons RE, Kirisits MJ, Mui EJ, Finnerty J, Johnston JJ, Ferguson**
537 **DJ, Coggins JR, Krell T, Coombs GH, Milhous WK, Kyle DE, Tzipori S, Barnwell J, Dame**
538 **JB, Carlton J, McLeod R** (2002) The shikimate pathway and its branches in apicomplexan
539 parasites. J Infect Dis **185** Suppl 1:S25-S36
540
541 **Roberts F, Roberts CW, Johnson JJ, Kyle DE, Krell T, Coggins, JR, Coombs GH, Milhous**
542 **WK, Tzipori S, Ferguson DJ, Chakrabarti, D, McLeod R** (1998). Evidence for the shikimate
543 pathway in apicomplexan parasites. Nature **393**:801-805
544
545 **Seijo Martinez M, Gonzalez-Mediero G, Santiago P, Rodriguez De Lope A, Diz J, Conde, C.**
546 **Visvesvara GS** (2000) Granulomatous amebic encephalitis in a patient with AIDS: Isolation of

547 *Acanthamoeba* sp. Group II from brain tissue and successful treatment with sulfadiazine and
548 fluconazole. J Clin Microbiol **38**:3892-3895
2
3
549
4
5
550 **Shukla OP, Kaul SM, Mehlotra RK** (1990) Nutritional studies on *Acanthamoeba culbertsoni* and
7
551 development of chemically defined medium. J Protozool **37**:237-242
9
10
552
11
12
553 **Shumilin IA, Kretsinger RH, Bauerle R** (1996) Purification, crystallization, and preliminary
14
554 crystallographic analysis of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from
15
16
17
555 *Escherichia coli*. Proteins **24**:404-406
19
20
556
21
22
557 **Shumilin IA, Bauerle R, Wu J, Woodard RW, Kretsinger RH** (2004) Crystal structure of the
24
558 reaction complex of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Thermotoga*
25
26
559 *maritima* refines the catalytic mechanism and indicates a new mechanism of allosteric regulation. J
27
28
560 Mol Biol **341**:455-466
31
32
561
34
562 **Walker GE, Dunbar B, Hunter IS, Nimmo HG, Coggins JR** (1996) Evidence for a novel class of
36
563 microbial 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase in *Streptomyces coelicolor*
37
38
39
564 A3(2), *Streptomyces rimosus* and *Neurospora crassa*. Microbiology **142**:1973-1982
41
42
565
43
44
566 **Wass MN, Kelley LA, Sternberg MJ** (2010) 3DLigandSite: predicting ligand-binding sites using
46
567 similar structures. Nucleic Acids Res **38**: W469-W473
47
48
49
568
51
569 **Webby CJ, Jiao W, Hutton RD, Blackmore NJ, Baker HM, Jameson GB, Parker EJ** (2010)
53
570 Synergistic allostery, a sophisticated regulatory network for the control of aromatic amino acid
54
55
56
571 biosynthesis in *Mycobacterium tuberculosis*. J Biol Chem **285**:30567-30576
58
59
60
61
62
63
64
65

Figure Legends

Figure 1. Glyphosate restricts *A. castellanii* growth at concentrations which are not toxic to mammalian cells. The percentage of alamarBlue reduction was determined for *A. castellanii* Neff strain (**A**), Clinical strain (**B**) and RCE (**C**) by comparing test cultures to those of untreated controls. . Glyphosate significantly inhibited growth of the Neff strain, in a dose dependent manner from 17.5 μ M or above ($p < 0.05$) (**A**). The IC_{50} is between 17.5 μ M and 35 μ M. Glyphosate significantly inhibited growth of a Clinical isolate, from 17.5 μ M or above ($p < 0.05$) (**B**). Glyphosate did not inhibit RCE growth (**C**). *A. castellanii* inhibited by glyphosate was washed in 5% SDS to determine cyst formation (**D**). 2.52 $\times 10^5$ /ml cysts were found in the encystment media (C2). 8.52 $\times 10^3$ /ml cysts were counted in the untreated controls (C3). Glyphosate does not induce cyst formation. There is a, non-dose dependent, increase in the number of cysts in the presence of glyphosate to that of untreated controls. In the presence of glyphosate (280 μ M) 9.17 $\times 10^3$ /ml to (70 μ M) 1.25 $\times 10^4$ / ml cysts were counted. The results are expressed as means for triplicate cultures \pm SE.

Figure 2. Rescue assays were performed to determine specificity of glyphosate to EPSP synthase of the shikimate pathway. Rescue was determined by measuring the percentage alamarBlue reduction of test cultures compared to those of untreated controls. The addition of single aromatic amino acids (Phe, Trp and Tyr) did not ablate the effects of glyphosate (**A**). However, by adding either phenylalanine and tryptophan or all three amino acids together the effect of glyphosate is significantly diminished ($p < 0.0001$ ***) (**B**). The presence of exogenous PABA did not rescue glyphosate inhibition of trophozoites (**C**).

Figure 3. Diagram of predicted phenylalanine and tyrosine biosynthesis in *A. castellanii* species. Enzymes highlighted in bold have been identified are present in Genbank. Pathways in grey are not present. (Chorismate mutase ELR23397, prephenate dehydratase ELR22681, aminotransferase ELR13846, phenylalanine-4-hydroxylase ELR14932, tyrosine aminotransferase ELR16760)

Figure 4. (A) Multiple sequence alignment of the two class I *A. castellanii* DAHP synthase sequences (AcDAHPIa, AcDAHPIb) and the *E. coli* DHAP synthase (EcDAHPI). Those residues conserved and similar are highlighted by red and blue boxes, respectively. Those residues involved in binding the Phe inhibitor are labelled by and F above the alignment. **(B)** Superposition of the EcDAHPI structure (yellow) and AcDAHPIa model structure (Magenta). Those residues which form the Phe binding site

1 are shown in stick format and coloured yellow, red, and blue for carbon, oxygen and nitrogen,
2 respectively. (C) Surface view of the Phe binding site from EcDAHPI. (D) Surface view of the
3 EcDAHPI structure with the Ser/Gly mutation as seen in the AcDAHPIb sequence which produces an
4 increase in space for the bound Tyr residue shown in stick format.
5
6
7

8 **Figure 5.** (A) Multiple sequence alignment of the AcDAHPII and MtDAHPII enzymes with those
9 residues conserved and similar highlighted by red and blue boxes, respectively. Those residues
10 involved in binding an amino acid inhibitor are shown by the label F1 (Primary Phe binding site), F2
11 (Secondary Phe binding site) and W (Trp binding site). The primary Phe binding site (B) and Trp
12 binding site (C) in MtDAHPII with the AcDAHPII model structure superposed. The MtDAHPII structure
13 is shown in yellow and AcDAHPII magenta. Those residues which have been shown to play a role in
14 binding are shown in stick format and coloured Yellow (MtDH2) or Magenta (AcDAHPII) for carbon,
15 red for oxygen and blue for nitrogen. The Lys residue which has been proposed to stabilise the α -
16 carboxylate group of the Trp inhibitor in MtDAHPII is highlighted by a star (C).
17
18
19
20
21
22
23
24
25
26
27

28 **Figure 6.** Arrangement of the shikimate pathway enzymes. In Bacteria and plants DAHP synthase
29 (red), 3-dehydroquinate synthase (orange), 5-enolpyruvylshikimate-3-phosphate synthase (purple),
30 shikimate kinase (blue), 3-dehydroquinate dehydratase (yellow), shikimate 5-dehydrogenase (green)
31 and chorismate synthase (pink) are encoded on individual polypeptides. In plants, 3-dehydroquinate
32 dehydratase and shikimate 5-dehydrogenase are found on separate domains of the same
33 polypeptide. Plant enzymes have n-terminal transit sequences that target the enzymes to the
34 chloroplast. In fungi, alveolates and oomycetes, 3-dehydroquinate synthase, 3-dehydroquinate
35 dehydratase, shikimate 5-dehydrogenase, shikimate kinase and EPSP synthase are encoded on a
36 single transcript to form a pentafunctional polypeptide (AROM). Two evolutionary unrelated forms of
37 DAHP synthase have been found and designated DAHP synthase I and DAHP synthase II. Bacteria
38 such as *E. coli* have 3 isoenzymes of DAHP synthase I, but plants such as *Arabidopsis thaliana* and
39 the apicomplexan *T. gondii* have a single DAHP synthase II. Certain fungi including *N. crassa* have 2
40 isoenzymes of DAHP synthase I and a DAHP synthase II. In *A. castellanii*, 3-dehydroquinate
41 synthase, 3-dehydroquinate dehydratase, shikimate 5-dehydrogenase, shikimate kinase and EPSP
42 synthase are part of a tetra-functional polypeptide (ARON). DAHP synthase and chorismate synthase
43 are found as individual polypeptides. *A. castellanii*, has 2 isoenzymes of DAHP synthase I and a
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

DAHP synthase II. Shikimate 5-dehydrogenase is found fused to PRAI-IGPS, an arrangement that is not found in any other organism to date.

Supplementary Figures

Figure S1. Enzymes involved in the folate and tryptophan pathways are present in the *A. castellanii* transcriptome²⁶

Figure S2. Multiple alignment of AROMn with other known AROM sequences from fungal species

Figure S3. Multiple alignment of *A.castellanii* IGPS-SD with other fungal species. The tri-protein fusion of IGPS-SD in *Acanthamoeba* is novel in nature

Figure S4. Multiple alignment of *A. castellanii* chorismate synthase with other fungal species

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

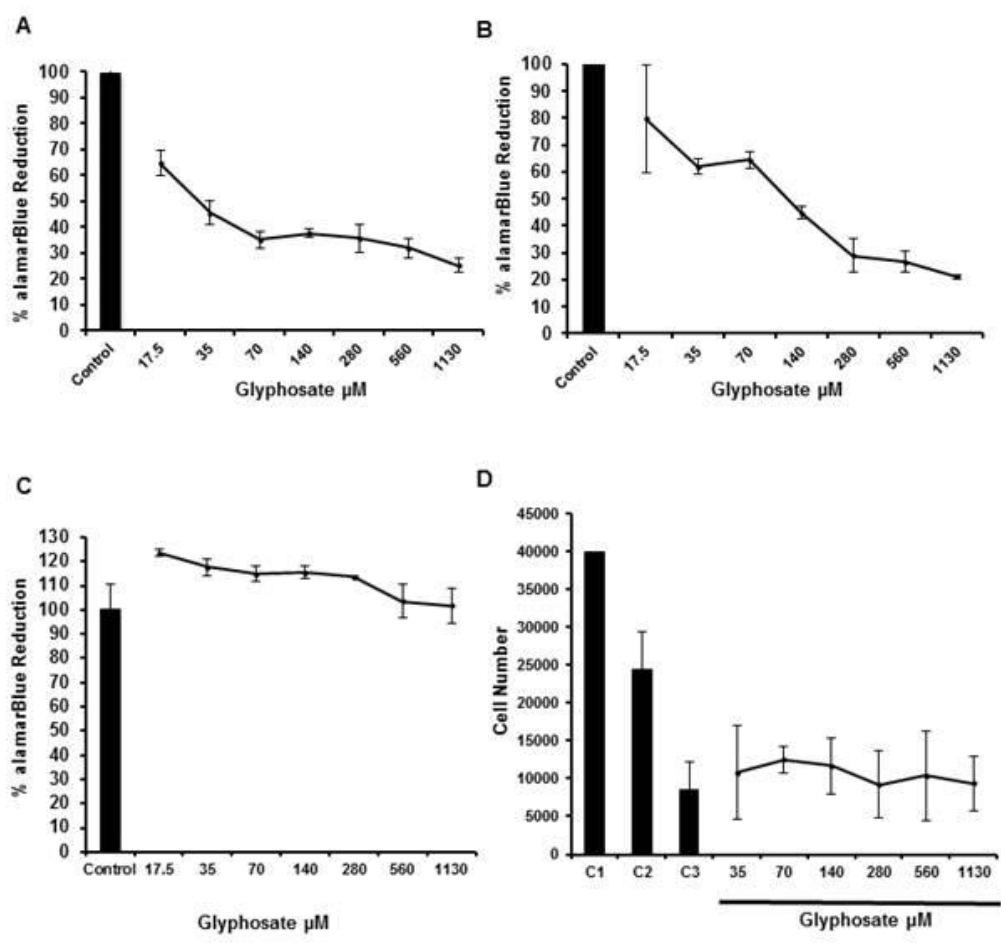


Figure 1

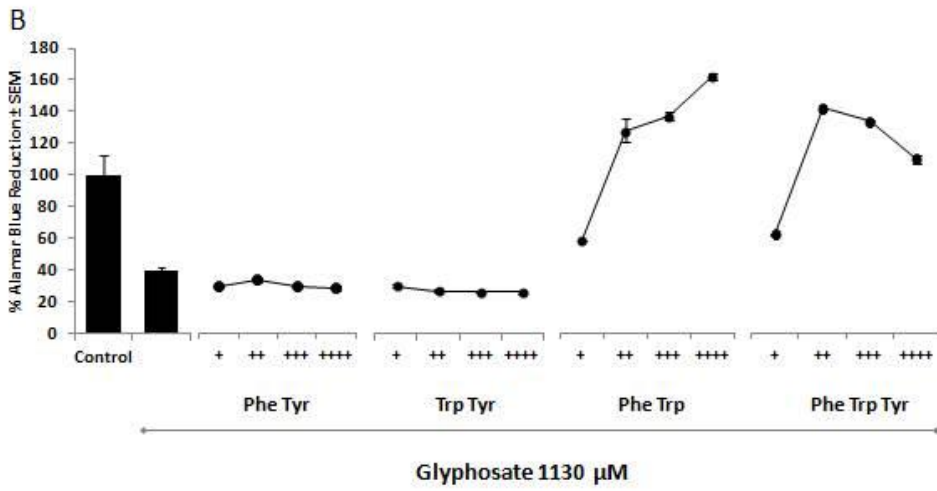
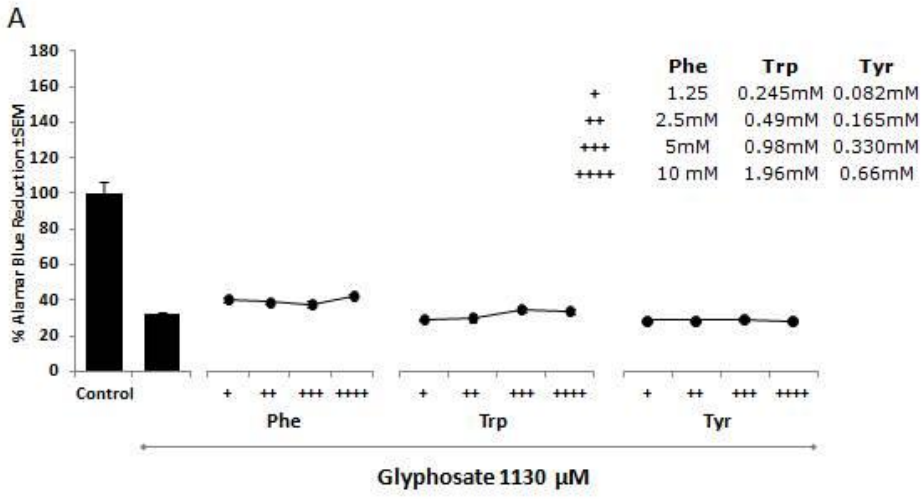


Figure 2

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

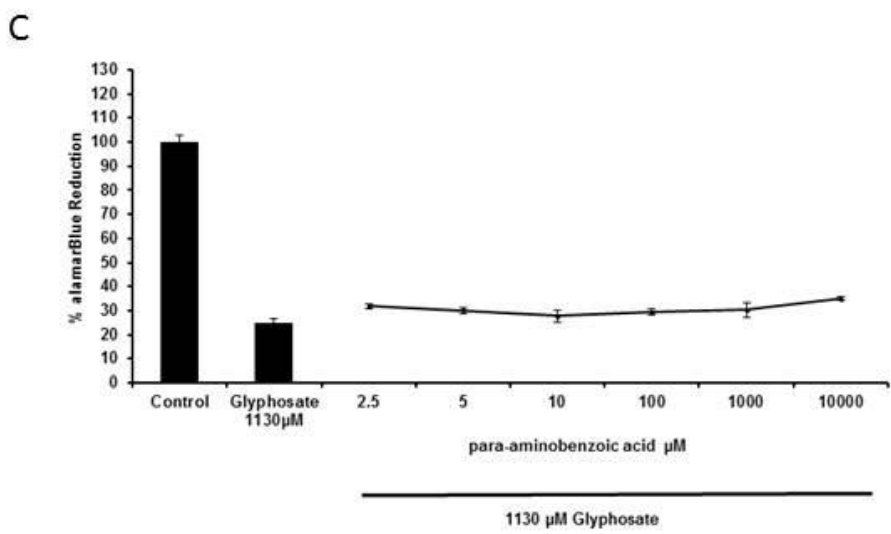


Figure 2

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Phenylalanine and Tyrosine synthesis

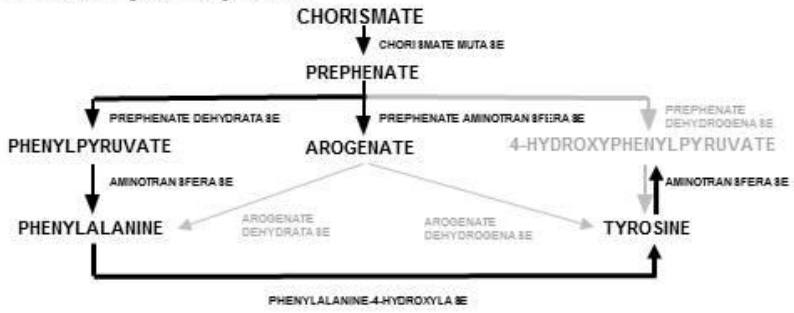


Figure 3

A

AcDAHP1a	M S A A A	D N N E G L I	D D V V V T	L I L P	P L I P P S I	L E L V	P N I A A	T R S A R T Q A
AcDAHP1b	M P R V S S T D F	K A T K N G V	D D V V I A I L L	P I I P P Q I	L L E L L	L I P A S T T	T V L E R R R E A	
EcDAHP1	M N	Q N	D D I L L K	L I E	L L P V A	L Q K F	A E N A A N T V A H A R K A I	

AcDAHP1a	A R	V R R R P G E L	D D R I L V V V	C P C S I N D P	A A A I	E Y A K	K L K E	E A E R I K E D I C I I M R V Y
AcDAHP1b	E A	I V R K Q	D D R I L V I V	C P C S I N D P	Q A A L	E Y A M	R L K E	Q V E R I X D D I C I I M R V Y
EcDAHP1	H K	L K G N	D D R I L V V I	C P C S I N D P	V A A K	E Y A T	R L L A	E R I K D D I C I I M R V Y

AcDAHP1a	R K P R T T V G W K G L I	S D R F L C S	S Y Q I R K	Q L R I A R	G L L C F L W E M G V	P A A V	R P L O T	I S
AcDAHP1b	R K P R T T V G W K G L I	N D P N L G E	T Y Q I R K	Q L R I A R	G L L C F L W E M G V	P A A V	R P L O T	I S
EcDAHP1	R K P R T T V G W K G L I	N D P H M G N	S F Q I R D	G L R I A R	K L L D I N D	S G L P A A G	R P L O M	I T

AcDAHP1a	D I V A D V	S W S A I G A R T T T	S Q V H R T L A S G L S	V P Y G F R N D Y D G	G V Q V A I M A	K R A S
AcDAHP1b	R Q I A D L V	S W S A I G A R T T T	S Q V H R T L A S G L S	V P Y G F R N D Y D G	G V Q V A I M A	K R A S
EcDAHP1	R A Y L A D M	S W S A I G A R T T T	S Q V H R T L A S G L S	C P Y G F R N D Y D G	T K V A I M A I N A A G	

AcDAHP1a	N R R Q F L S	G V T H Q G L T A	V R T L D R P A C H V I	L R G G R G G	P N F D E T H V	Q K T V	L A R R K K I
AcDAHP1b	N R R Q F L S	V S H Q G L T A	V R T L D R P A C H V I	L R G G R G G	P N F D E T H V	Q Q V M E	Q L E R K K I
EcDAHP1	A D R C F L S	V T H W G H S A R V N T S	D R G A C H I I	L R G G R K E	P N Y S A K H V	A E V K E	G L N Y A G I

AcDAHP1a	S T S I M I R C S H D N S R R	N H A N G P I V S A A T	A E G V A A G C A D I I	B V M I P S N I V F R R Q D P
AcDAHP1b	P T S I M I R C S H D N S R R	N H L N G P T V A Q S I	A D G I A A G N D S I T	B V M I P S F I V E R R Q D P
EcDAHP1	P A Q V M I D F S H A M S S	K Q F K K C M D V C A D V	C Q Q I A G H E K A I I	B V M V P S H I V E R N O S E

AcDAHP1a	A D G P Q H	Q Y G K S I Y R A C	S F K D T I P V	E S A A A V R E R A	K K A G S A A T T N G S A
AcDAHP1b	K E G P R Y	Q Y G K S I Y R A C	L S W Q Q T V P V L	Q S A C A V R R R R	G T P E E A S A N T Q
EcDAHP1	S G E P	A Y G K S I Y R A C I G W E	D I D A L I R Q	A N A V K A R R G	

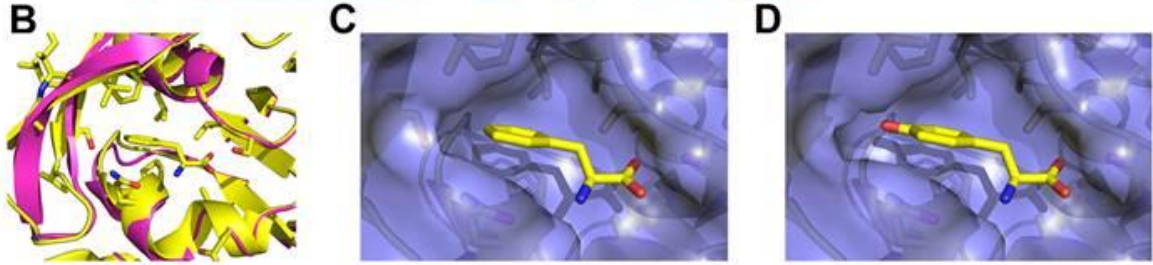


Figure 4

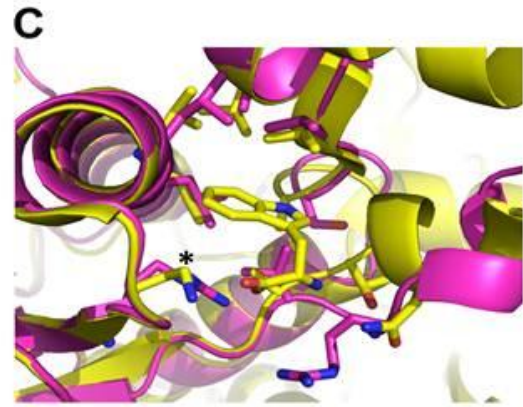
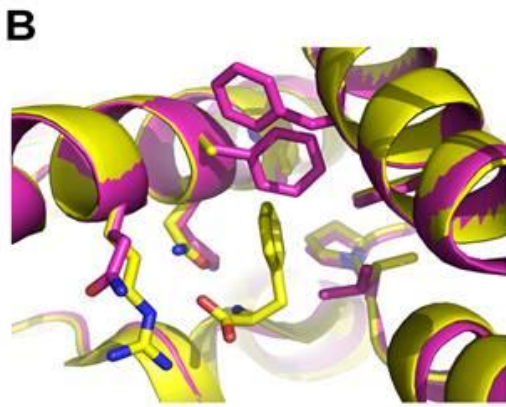
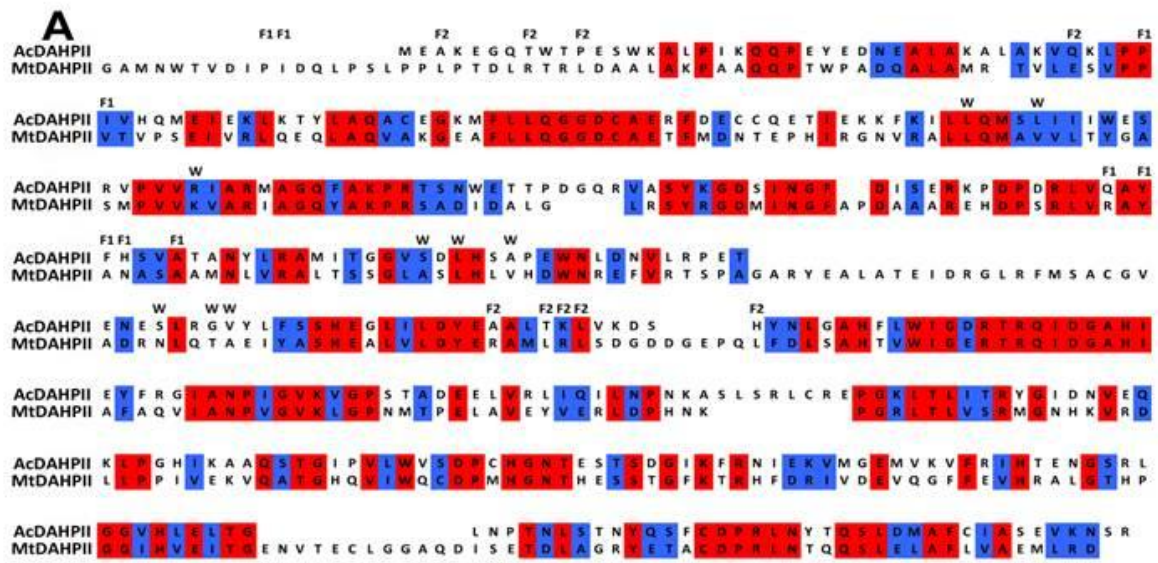


Figure 5

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

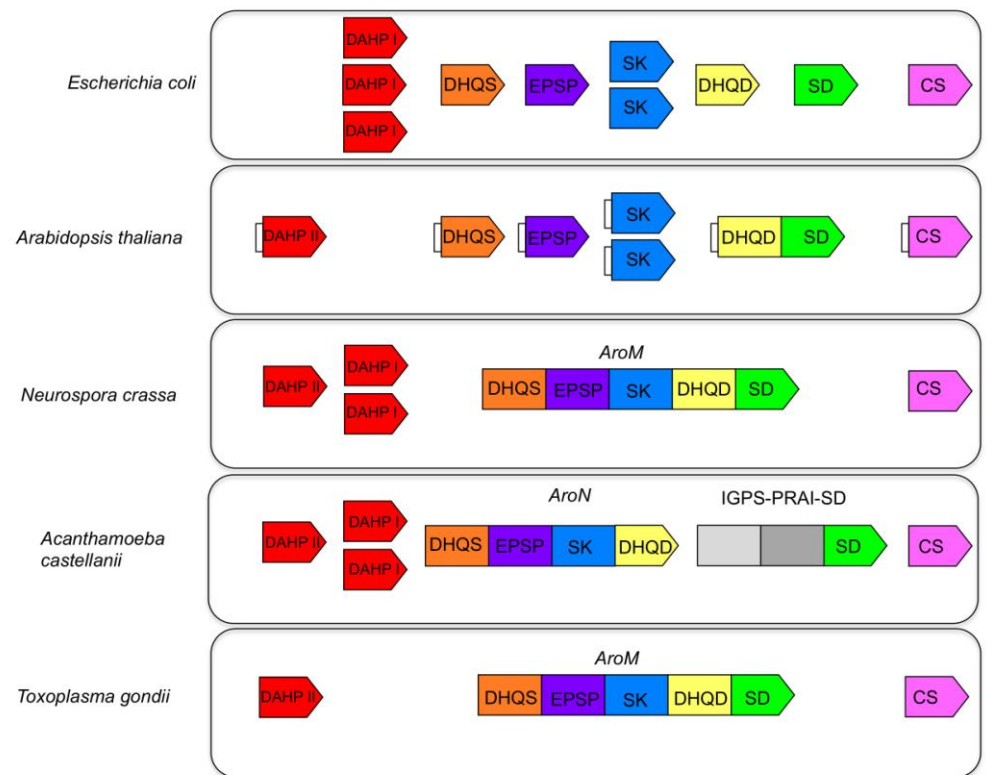
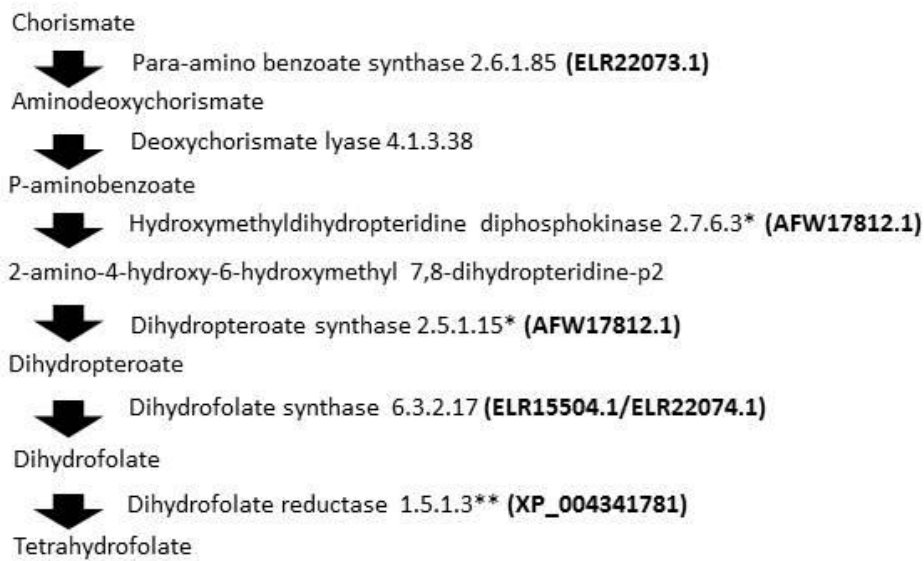


Figure6

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

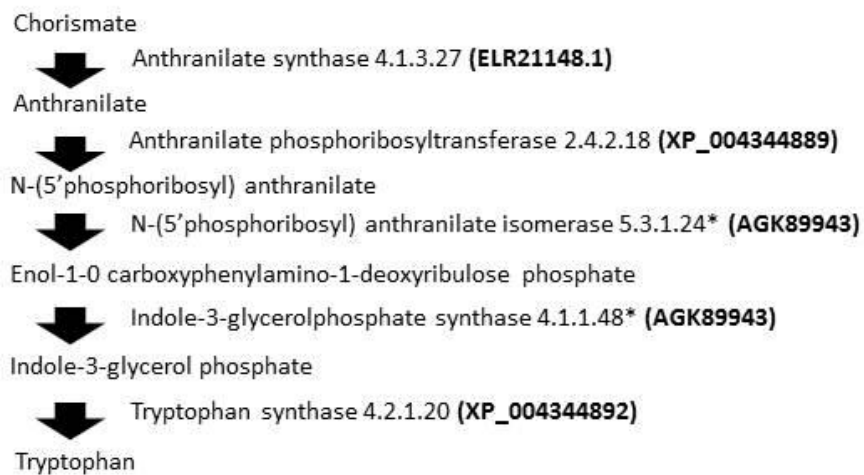


Figure 7



14 *trifunctional gene product with dihydroneopterin aldolase

15 **Bifunctional gene product with thymilidate synthase



27 *trifunctional gene product indole-3-glycerol phosphate synthase,
 28 phosphoribosylanthranilate isomerase and shikimate dehydrogenase

29 Figure S1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

```
Acanthamoeba MCKKDDPVEDVVERVEGQKDFITDNEIEKQWLAIFIBERNPFAVLEITITDNLGGLVGGTFFIFPQVLIH---TKPFVKVIFPQERKSRQRETEEDNIAK--ASDID  
Neospora MAEIL-----ENFRINILQKSNITIDGQWLNFAVQDILLONIKSSTYIITDNLNLTVPVQSVFENAA--PQDVRLLIYIAPPQKSRRTKARIEDWMLSL--GSTRD  
Fusarium MAQASD-----QDPTRIISLQGFNITVDHQLWLNFAVQDILLONITATSTYVITDNLNLSYVFPQSVFENAA--GKQTRLLIYIAPPQKSRRTKARIEDWMLSL--GSTRD  
Aspergillus -----M-----TEPTRIISLQGFNITVDHQLWLNFAVQDILLONIKSSTYIITDNLNLTVPVQSVFENAA--PQDVRLLIYIAPPQKSRRTKARIEDWMLSL--GSTRD  
  
Acanthamoeba CYVALGGVIGDLAGVAATLRGVYVQVPTLLAMVDSSIGGKTAIDTFMGKNIQAPWQPKRIYIDLTFLPVPREFINGMAEVKTAALWDBEFTALBENAKALBAVRSNK-  
Neospora VVIALGGVIGDLAGVAATLRGVYVQVPTLLAMVDSSIGGKTAIDTFMGKNIQAPWQPKRIYIDLTFLPVPREFINGMAEVKTAALWDBEFTALBENAKALBAVRSNK-  
Fusarium VVIALGGVIGDLAGVAATLRGVYVQVPTLLAMVDSSIGGKTAIDTFMGKNIQAPWQPKRIYIDLTFLPVPREFINGMAEVKTAALWDBEFTALBENAKALBAVRSNK-  
Aspergillus VVIALGGVIGDLAGVAATLRGVYVQVPTLLAMVDSSIGGKTAIDTFMGKNIQAPWQPKRIYIDLTFLPVPREFINGMAEVKTAALWDBEFTALBENAKALBAVRSNK-  
  
Acanthamoeba -----ERLEIEIVHSARKAAEYVVSDDEREGGRLNLLNFQHSIGHARAILTPQLHGEVAIGMVKEARLARFQVLRGAVARLYKCLAVYDLPTSLQDREK---LTA  
Neospora SADLLEIDILKRIIVGDSRYKAEVVSDDEREGGRLNLLNFQHSIGHARAILTPQLHGEVAIGMVKEARLARFQVLRGAVARLYKCLAVYDLPTSLQDREK---LTA  
Fusarium EGRLEIEIVHSARKAAEYVVSDDEREGGRLNLLNFQHSIGHARAILTPQLHGEVAIGMVKEARLARFQVLRGAVARLYKCLAVYDLPTSLQDREK---LTA  
Aspergillus GQRLEIEIVHSARKAAEYVVSDDEREGGRLNLLNFQHSIGHARAILTPQLHGEVAIGMVKEARLARFQVLRGAVARLYKCLAVYDLPTSLQDREK---LTA  
  
Acanthamoeba TLEETPEDLNKMLDKKNANSNIRGVKDFDQVRRHPVYVDDKQTRFLVLPNVAQVLOSGLVSGVITVPSKSIENRLLLAALGGDQHKGKLAADDTVMLAALRRLLGVQYK  
Neospora QKECFVDVLLKMAVDKKNLQKKKIVLLSAIQKTYEASVVDQRAIRVLSFCIRVAVQVPKDLNVYVTPQSKSISNRALLAALGGDQTRIKNLLSDDTQVMLNAVAQLQDASIS  
Fusarium QKRCFVDVLLKMAVDKKNLQKKKIVLLSAIQKTYEASVVDQRAIRVLSFCIRVAVQVPKDLNVYVTPQSKSISNRALLAALGGDQTRIKNLLSDDTQVMLNAVAQLQDASIS  
Aspergillus QKRCFVDVLLKMAVDKKNLQKKKIVLLSAIQKTYEASVVDQRAIRVLSFCIRVAVQVPKDLNVYVTPQSKSISNRALLAALGGDQTRIKNLLSDDTQVMLNAVAQLQDASIS  
  
Acanthamoeba WLDDQITLVKNGQDQAPPKPEIPIIGNAGTASRPLTCTEIVTPAGSEKSVTITLGVNLRERPSIDLVVALRNGCVNTELRKRGVPIEIVVGGKALNGGVIPLAKISSQYK  
Neospora WELDDQITLVKNGQDQAPPKPEIPIIGNAGTASRPLTCTEIVTPAGSEKSVTITLGVNLRERPSIDLVVALRNGCVNTELRKRGVPIEIVVGGKALNGGVIPLAKISSQYK  
Fusarium WELDDQITLVKNGQDQAPPKPEIPIIGNAGTASRPLTCTEIVTPAGSEKSVTITLGVNLRERPSIDLVVALRNGCVNTELRKRGVPIEIVVGGKALNGGVIPLAKISSQYK  
Aspergillus WELDDQITLVKNGQDQAPPKPEIPIIGNAGTASRPLTCTEIVTPAGSEKSVTITLGVNLRERPSIDLVVALRNGCVNTELRKRGVPIEIVVGGKALNGGVIPLAKISSQYK  
  
Acanthamoeba SSIILMAAFYANPVYTLRLVGGKPIISQYIDMTIAMMASFGIDVNEVSKETPETHYHIPGQYVNPPEYVSDASSATYPLAAAIITOTCTVFNIGSKSLQDARPAVDVLRPMGCTVEG  
Neospora SSIILMAAFYANPVYTLRLVGGKPIISQYIDMTIAMMASFGIDVNEVSKETPETHYHIPGQYVNPPEYVSDASSATYPLAAAIITOTCTVFNIGSKSLQDARPAVDVLRPMGCTVEG  
Fusarium SSIILMAAFYANPVYTLRLVGGKPIISQYIDMTIAMMASFGIDVNEVSKETPETHYHIPGQYVNPPEYVSDASSATYPLAAAIITOTCTVFNIGSKSLQDARPAVDVLRPMGCTVEG  
Aspergillus SSIILMAAFYANPVYTLRLVGGKPIISQYIDMTIAMMASFGIDVNEVSKETPETHYHIPGQYVNPPEYVSDASSATYPLAAAIITOTCTVFNIGSKSLQDARPAVDVLRPMGCTVEG  
  
Acanthamoeba FLEETTVKIPARQLEKAMDIIDMSMTDIPVAVLAAVASG-----ETLIVANQRVKECNRITAMKDLAKFGVVCBLEDGIVVITPTYTENLNDRHRVAMSPFVVC  
Neospora FLEETTVKIPARQLEKAMDIIDMSMTDIPVAVLAAVASG-----ETLIVANQRVKECNRITAMKDLAKFGVVCBLEDGIVVITPTYTENLNDRHRVAMSPFVVC  
Fusarium FLEETTVKIPARQLEKAMDIIDMSMTDIPVAVLAAVASG-----ETLIVANQRVKECNRITAMKDLAKFGVVCBLEDGIVVITPTYTENLNDRHRVAMSPFVVC  
Aspergillus FLEETTVKIPARQLEKAMDIIDMSMTDIPVAVLAAVASG-----ETLIVANQRVKECNRITAMKDLAKFGVVCBLEDGIVVITPTYTENLNDRHRVAMSPFVVC  
  
Acanthamoeba IYVYDGLITIDRGGVSTLIPVYDIFRFGVYVVSFTENEQKQFATAAAALVYVYDGGPQDSDRSVLEKMRGQSRFLGKALADGNWRTIDIVLRAEADHLEASND-ALGWP  
Neospora IYVYDGLITIDRGGVSTLIPVYDIFRFGVYVVSFTENEQKQFATAAAALVYVYDGGPQDSDRSVLEKMRGQSRFLGKALADGNWRTIDIVLRAEADHLEASND-ALGWP  
Fusarium IYVYDGLITIDRGGVSTLIPVYDIFRFGVYVVSFTENEQKQFATAAAALVYVYDGGPQDSDRSVLEKMRGQSRFLGKALADGNWRTIDIVLRAEADHLEASND-ALGWP  
Aspergillus IYVYDGLITIDRGGVSTLIPVYDIFRFGVYVVSFTENEQKQFATAAAALVYVYDGGPQDSDRSVLEKMRGQSRFLGKALADGNWRTIDIVLRAEADHLEASND-ALGWP  
  
Acanthamoeba FRABEILRQVGNHQPQVVISIGGGIVETFAFEYLLAKKED-VTEVQRNPNVEDVIEYLSADVADLOREIKTIWERRRFLYQVQDDEHETIGRGEDWA-ALBRGLVDLQWRV  
Neospora FRABEILRQVGNHQPQVVISIGGGIVETFAFEYLLAKKED-VTEVQRNPNVEDVIEYLSADVADLOREIKTIWERRRFLYQVQDDEHETIGRGEDWA-ALBRGLVDLQWRV  
Fusarium FRABEILRQVGNHQPQVVISIGGGIVETFAFEYLLAKKED-VTEVQRNPNVEDVIEYLSADVADLOREIKTIWERRRFLYQVQDDEHETIGRGEDWA-ALBRGLVDLQWRV  
Aspergillus FRABEILRQVGNHQPQVVISIGGGIVETFAFEYLLAKKED-VTEVQRNPNVEDVIEYLSADVADLOREIKTIWERRRFLYQVQDDEHETIGRGEDWA-ALBRGLVDLQWRV  
  
Acanthamoeba RDRAPLPLPGLDRDTGESHVLSLTPVVEECIFFADEEFADVDLIDRQVDLIDSDQ-----DDELEQALRRRTAKTELLFLKTRGGGGRYBETGRADBIYRIGCALRAGVVEI  
Neospora RDRAPLPLPGLDRDTGESHVLSLTPVVEECIFFADEEFADVDLIDRQVDLIDSDQ-----DDELEQALRRRTAKTELLFLKTRGGGGRYBETGRADBIYRIGCALRAGVVEI  
Fusarium RDRAPLPLPGLDRDTGESHVLSLTPVVEECIFFADEEFADVDLIDRQVDLIDSDQ-----DDELEQALRRRTAKTELLFLKTRGGGGRYBETGRADBIYRIGCALRAGVVEI  
Aspergillus RDRAPLPLPGLDRDTGESHVLSLTPVVEECIFFADEEFADVDLIDRQVDLIDSDQ-----DDELEQALRRRTAKTELLFLKTRGGGGRYBETGRADBIYRIGCALRAGVVEI  
  
Acanthamoeba DLEITLQDPIEAVLDEKRRALGALTEIENLQNDQPKSEE-DVYVYVLCCHALGGTGAALSASSLVQYKVNIFAFADQVMTLRNGVVRQAEIVANLOGEPIIALAALNIGKLSRVL  
Neospora DLEITLQDPIEAVLDEKRRALGALTEIENLQNDQPKSEE-DVYVYVLCCHALGGTGAALSASSLVQYKVNIFAFADQVMTLRNGVVRQAEIVANLOGEPIIALAALNIGKLSRVL  
Fusarium DLEITLQDPIEAVLDEKRRALGALTEIENLQNDQPKSEE-DVYVYVLCCHALGGTGAALSASSLVQYKVNIFAFADQVMTLRNGVVRQAEIVANLOGEPIIALAALNIGKLSRVL  
Aspergillus DLEITLQDPIEAVLDEKRRALGALTEIENLQNDQPKSEE-DVYVYVLCCHALGGTGAALSASSLVQYKVNIFAFADQVMTLRNGVVRQAEIVANLOGEPIIALAALNIGKLSRVL  
  
Acanthamoeba KPLTFVIRHPKAAAPQQLSAAIKSRLGQVIGQKKAIVFGTPISSSSPALHNTLAFANGLFHYSELSEPAKDVEDVITSPQOASVTIPKLDIMPDLDDSAAMVIGAVN  
Neospora KPLTFVIRHPKAAAPQQLSAAIKSRLGQVIGQKKAIVFGTPISSSSPALHNTLAFANGLFHYSELSEPAKDVEDVITSPQOASVTIPKLDIMPDLDDSAAMVIGAVN  
Fusarium KPLTFVIRHPKAAAPQQLSAAIKSRLGQVIGQKKAIVFGTPISSSSPALHNTLAFANGLFHYSELSEPAKDVEDVITSPQOASVTIPKLDIMPDLDDSAAMVIGAVN  
Aspergillus KPLTFVIRHPKAAAPQQLSAAIKSRLGQVIGQKKAIVFGTPISSSSPALHNTLAFANGLFHYSELSEPAKDVEDVITSPQOASVTIPKLDIMPDLDDSAAMVIGAVN  
  
Acanthamoeba TLEIVRNDCK--VITIGDNTDWMVMVALRNAGVVKCSKESPTAQMVGVGOTTRAAVALHLDLGFAPYVARNADRVKAAEAPFAVYDIESESTPEEVAAESDAQSVVISTIPAL  
Neospora TLEIVRNDCK--VITIGDNTDWMVMVALRNAGVVKCSKESPTAQMVGVGOTTRAAVALHLDLGFAPYVARNADRVKAAEAPFAVYDIESESTPEEVAAESDAQSVVISTIPAL  
Fusarium TLEIVRNDCK--VITIGDNTDWMVMVALRNAGVVKCSKESPTAQMVGVGOTTRAAVALHLDLGFAPYVARNADRVKAAEAPFAVYDIESESTPEEVAAESDAQSVVISTIPAL  
Aspergillus TLEIVRNDCK--VITIGDNTDWMVMVALRNAGVVKCSKESPTAQMVGVGOTTRAAVALHLDLGFAPYVARNADRVKAAEAPFAVYDIESESTPEEVAAESDAQSVVISTIPAL  
  
Acanthamoeba KPIDSMRNVVIRHPII-----EYADQKHVLEEMAYIPHIFPLMQLAALDQVQDIFOLEVLAAGGYQDQWGTQITFIYDAAAVVQK  
Neospora KPIDSMRNVVIRHPII-----EYADQKHVLEEMAYIPHIFPLMQLAALDQVQDIFOLEVLAAGGYQDQWGTQITFIYDAAAVVQK  
Fusarium KPIDSMRNVVIRHPII-----EYADQKHVLEEMAYIPHIFPLMQLAALDQVQDIFOLEVLAAGGYQDQWGTQITFIYDAAAVVQK  
Aspergillus KPIDSMRNVVIRHPII-----EYADQKHVLEEMAYIPHIFPLMQLAALDQVQDIFOLEVLAAGGYQDQWGTQITFIYDAAAVVQK
```

Figure S2

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

```
Acantamoeba .....
Schizosaccharomyces MS EKMDVDSK ..... QDAS ..... ENAVKEVAERPVMQDNYDSFTNIVYVLSANRRKYPIVMSFNSITIDDELKLNQ - LKLVSPGPHPARQMLCNELSRFAGKIFLIVG
Rhodospiridium MAAPIGNQSGPLTVNDGKASASSSSLA PKPFI DPNNHVMQDNYDSFTNIVYVLSANRRKYPIVMSFNSITIDDELRAQHPTMTLHVSPGPHPARQMLCNELSRFAGKIFLIVG
Mycoplasma

Acantamoeba .....
Schizosaccharomyces CMGLGCIHETMGGQDSAGHIIHGRKSNINDDGCGVGHDPNLSVTRYHSLAGQIISLPPCDVTSWTF ..... NQIMGQRHRVYIHDVQVHPESILSIVGKLYIQNPFNEITAGTW
Rhodospiridium CMGLGCIITVKGGVIRAGIETASIKRSTSAIIDPKQGFNIAQVAZTRYHSLAGQIGPFIIVTSIDEL ..... GQIMGRHRVYIHDVQVHPESILSIVGKLYIQNPFNEITAGTW
Mycoplasma MGFSCIVAVFGQVSTGIIHGKTSILSDGKGTGHTGVAATRYHSLAGTHIFTEIEMSPVAVGPDGQGVIMGRHRKRYIHDVQVHPESILSIVGKLYIQNPFNEITAGTW

Acantamoeba .....
Schizosaccharomyces ..... MASITLQKIQGRIGDVAISLSPFALIERRLVDEKQVNRGQVIFENITPRLVSA ..... FALINLADEVKRASPSKGDIAF
Rhodospiridium FIKSIIDPPTKKVAVPMEESVNSVSKIRKQETLISAGQITDINPSSVGGI ..... SS ..... GDIOTYENSATFETITPRLVSA ..... FALINLADEVKRASPSKGDIAF
Mycoplasma FIFPFIIEPKKPSASPLA - - - NSAVS FNAAVTELOHIGRIGDVAISLSPFALIERRLVDEKQVNRGQVIFENITPRLVSA ..... FALINLADEVKRASPSKGDIAF

Acantamoeba .....
Schizosaccharomyces G - - - IDAQDIAVAGAAVNSITETPFWPK - - - DCGIILNIGAMGIEPRPALRKKDFLEIYOIVEARVYGGADTVLLIVAAEEDVNRLEISGEGMEPLVETANABEMENALRV
Rhodospiridium D - - - IAAIAIADIAVAGAAVNSITETPFWPK - - - DCGIILNIGAMGIEPRPALRKKDFLEIYOIVEARVYGGADTVLLIVAAEEDVNRLEISGEGMEPLVETANABEMENALRV
Mycoplasma D - - - IAAIAIADIAVAGAAVNSITETPFWPK - - - DCGIILNIGAMGIEPRPALRKKDFLEIYOIVEARVYGGADTVLLIVAAEEDVNRLEISGEGMEPLVETANABEMENALRV

Acantamoeba .....
Schizosaccharomyces GSNVIGVNNRNRIEPEVDIETFEIAADVA - - - QSIRGVTFALSOITREIIVVIEKAVSVGLVGSLSRAADPAATFEELODEET
Rhodospiridium GSNVIGVNNRNRIEPEVDIETFEIAADVA - - - DDVITLALSGISPADAVIISGQVSVLVGESLMRASPAAPARELLNSNSHIS - - - N
Mycoplasma GSNVIGVNNRNRIEPEVDIETFEIAADVA - - - KTITLIALSGIADGVEIENQVAVLVEPLMRAADITRITLLELSISGSDI - - - I

Acantamoeba .....
Schizosaccharomyces - - - RTPALVKTGLEDVPTALAVGAGADTGLPABS - - - RRKVTIIVAREIIVADKWITAGGQAFSAHEFMAENRRAQESFGSIVKL - - - SVAVESLPTLVGVHADNDFANNEV
Rhodospiridium GRKTSIFANQCGIRSLIIRKILYSGDDEIQLIFPESSTKAVSYAREIITPTEFRRRI - - - LI - - - EKANG - - - SPTIRRYENIPSTFELVGVKQDEIETLSI
Mycoplasma SSSPRDPLVKTCGIRTFEAAATAAGADILQLIFAPSKRNYSIQALEISAV - - - RARIR - - - QD - - - KRATADAADT - - - SDWPSIASRISAPREPLVGVFNPSSIEITVIST
N - SQKIDLVKTCGIRTFEAAATAAGADILQLIFAPSKRNYSIQALEISAV - - - RARIR - - - QD - - - KRATADAADT - - - SDWPSIASRISAPREPLVGVFNPSSIEITVIST

Acantamoeba .....
Schizosaccharomyces ATAAHLDIQLDSSPAVAGQEVKCEKALVYGGPBEINSTAGPAGG - - - NVVGLIIGYDPH - - - ALGQIRAHDSIASW - - - OEKIFVLAGGLFQNNAAVAV
Rhodospiridium EAEVNLDIQLHQSPEIWAHIDPVPVIRAFVYDEKADADAEAAQLREAFRFQHALRELDTKSKDQALSAGQIFVDMANRREVSRRVQGHRLFTVLAGGLDAANKEEAVT
Mycoplasma HRLISQDITLHQSPEIWAHIDPVPVIRAFVYDEKADADAEAAQLREAFRFQHALRELDTKSKDQALSAGQIFVDMANRREVSRRVQGHRLFTVLAGGLDAANKEEAVT

Acantamoeba .....
Schizosaccharomyces - - - SIFLDVDSGQVETDVKVAKIRAFVYHAKITTDQVADGSQFYLQSPFIGNSPSPLLHNTGFETLGIADQHRYLCDTKDKVQVVYLRDRRTGGGVTMPHKQTVMPFLDQIS
Rhodospiridium - - - SIFLDVDSGQVETDVKVAKIRAFVYHAKITTDQVADGSQFYLQSPFIGNSPSPLLHNTGFETLGIADQHRYLCDTKDKVQVVYLRDRRTGGGVTMPHKQTVMPFLDQIS
Mycoplasma GDARQVAVDVDSGQVETDVKVAKIRAFVYHAKITTDQVADGSQFYLQSPFIGNSPSPLLHNTGFETLGIADQHRYLCDTKDKVQVVYLRDRRTGGGVTMPHKQTVMPFLDQIS

Acantamoeba .....
Schizosaccharomyces PGAWAIGAVNTVSKDPAGHLLGDNTDWLAIHQLTKQRLLAALOKTGAKFVGLVIGAGGTAHAAACYALQLDAEYIYNRTPARAGHLADRFAGVRLPSTPDELAALAGRVDDVIVSTVPPT
Rhodospiridium .....
Mycoplasma .....

Acantamoeba .....
Schizosaccharomyces AGPTLPDPAFFRRRPSGADAA5AAAGLVVVELVRQCAKNDENVVRVIVVEGIEILLAQGLAQPIWTGREAAPRAAIVDKIVATFKDGLYASPLPTSFQ
Rhodospiridium .....
Mycoplasma .....
```

Figure S3

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

```

Acanthamoeba  M S S F G R A F R V T T E G E S H G G V G C I I D G V P P C L P L T E A D I Q P Q L S R R R P G Q S S L T T P R N E A D Q V K I Q A G T E N
Saccharomyces M S T F G K L F R V T T Y G E S H C K S V G C I V D G V P P G M S L T E A D I Q P Q L T R R R P G Q S K L S T P R D E K D R V E I Q S G T E F
Rhodosporidian M S T F G A L F R V T T Y G E S H C A S V G A I I D G C P P G M P L T D E D I Q P Q M T R R R P G Q S N L T T P R N E A D A V Q I Q S G V E A
Aspergillus      M S T W G E Y F R V T T Y G E S H C R S V G C I V D G C P P G M E L T E E D I Q P Q M T R R R P G Q S A L T T P R N E K D R V E I Q S G T E F

Acanthamoeba  G Y T L G S P I S L F V A N Q D Q R P V D Y S - - D M S K I P R P S H A D Y T Y A A K Y N I K S S S G G R S S A R E T I G R V C A G A V A E
Saccharomyces G K T L G T P I A M M I K N E D Q R P H D Y S - - D M D K F P R P S H A D F T Y S E K Y G I K A S S G G R A S A R E T I G R V A S G A I A E
Rhodosporidian G V T L G T P I G L L V R N K D Q R P H D Y T - - E T D H Y P R P S H A D W T Y L L K Y G L K A S S G G R S S A R E T I G R V A A G T I A E
Aspergillus      G I T L G T P I G M M V R N E D Q R P R D Y G G S T M D L Y P R P S H A D Y T Y L E K Y G V K A S S G G R S S A R E T I G R V A A G A I A E

Acanthamoeba  K W L K L K Y G V E I V A W V S S I G D Q E I E R E - - - - - L D L D T I S R E D V D K - S L V R C P D E A A T A K M I E I
Saccharomyces K F L A Q N S N V E I V A F V T Q I G E I K M N - - - - - R D S F D P E F Q H L L N T I T R E K V D S M G P I R C P D A S V A G L M V K E
Rhodosporidian R Y L K L A Y G V E I V A F V S S V G K V H M P E S S - - T D S D L L S E D Y L K L L S T I T R E K V D E - N T I R C P H A E T A Q A M E E R
Aspergillus      K Y L R L S H G V E I V A F V S S V G N E H L F P P T P E H P S P S T N P E F L K L I E T I D R K T V D A F V P T R C P N E E A A A R M T K V

Acanthamoeba  I E A A K A D K D S I G G V V T C V C R N V P T G L G E P C F D K L E A M L A H A M L S I P A T K G F E I G S G F A G T R M R G S K H N D P F
Saccharomyces I E K Y R G N K D S I G G V V T C V V R N L P T G L G E P C F D K L E A M L A H A M L S I P A S K G F E I G S G F Q G V S V P G S K H N D P F
Rhodosporidian I M A A K A K N D S I G G T V T C V I R R S P V G L G E P V F D K L E A K L A H A M L S I P A T K G F E I G S G F R G T E V P G S B H N D A F
Aspergillus      I E T F R D N Q D S I G G T V T C V I R N V P V G L G E P C F D K L E A K L A H A M L S I P A T K G F E I G S G F G C E V P G S I H N D P F

Acanthamoeba  V V K T - - - G A D G K K R L G T T N H S G G I Q G G I T N G E H V V F K V A F K P P A T I S Q A Q K T S E Y G G G E A V L E A Q G R H D P
Saccharomyces Y F E K - - - - - E T N R L R L T K T N N S G G V Q G G I S N G E N I Y F S V P P K S V A T I S Q E Q K T A T Y D G E E G I L A A K G R H D P
Rhodosporidian V K K - - - - - A D G S L G T K T N R S G G I Q G G I T N G E D I Y F K I A F K S P A T I S Q E Q A T A K Y D G E S G V L A T R G R H D P
Aspergillus      T V S E V Q T R T G S T Q R L T T K T N N S G G I Q G G I S N G A P I Y F R V A F K P P A T I G Q A Q T T A S Y S F E E G I L E A K G R H D P

Acanthamoeba  C V V P R A I P I V E A M A A L V I A D A A L L Q L A R Q G S L V S E P I V A W Q I - - - - -
Saccharomyces A V T P R A I P I V E A M T A L V L A D A L L I Q A R D F S R S V V H - - - - -
Rhodosporidian C V V P R A V P I V E A M A A L V I M D A V L I Q A R Q T A A S L L P P L E T P L P P S M S L P P K K V E E K M V K E I V Q E Q Q A
Aspergillus      C V T P R A V P I V E A M S A L V M D A L M A Q A R E S A K N L L P P L P S T L P T K P T L G S S G A P A S S - - - - -

```

Figure S4