Title:

Discovery of a SAR11 growth requirement for thiamin's pyrimidine precursor and its distribution in the Sargasso Sea

Running title: SAR11 requires HMP

Authors: Paul Carini[1,6], Emily O. Campbell[1], Jeff Morré[2], Sergio A. Sañudo-Wilhelmy[3], J. Cameron Thrash[1,7], Samuel E. Bennett[4], Ben Temperton[1,8], Tadhg Begley[5] and Stephen J. Giovannoni*[1]

[1]Department of Microbiology, Oregon State University Corvallis, OR 97331

[2]Department of Chemistry, Oregon State University, Corvallis, OR 97331

[3]Department of Biological Sciences, Marine Environmental Biology, and Earth Science, University of Southern California, Los Angeles, CA 90089

[4]Department of Environmental & Molecular Toxicology, Oregon State University, Corvallis, OR 97331

[5]Department of Chemistry, Texas A&M University, College Station, Texas 77843

[6]Current address: Horn Point Laboratory, University of Maryland Center for Environmental Science, Cambridge, MD 21613

[7]Current address: Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803

[8]Current address: Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth PL1 3DH, UK

*Corresponding author: steve.giovannoni@oregonstate.edu

1 **Abstract**

2   Vitamin traffic, the production of organic growth factors by some microbial

3 community members and their use by other taxa, is being scrutinized as a potential

4 explanation for variation and highly connected behavior observed in ocean plankton by

5 community network analysis. Thiamin (vitamin B_1), a cofactor in many essential

6 biochemical reactions that modify carbon-carbon bonds of organic compounds, is

7 distributed in complex patterns at sub-picomolar concentrations in the marine surface

8 layer (0-300 m). Sequenced genomes from organisms belonging to the abundant and

9 ubiquitous SAR11 clade of marine chemoheterotrophic bacteria contain genes coding for

10 a complete thiamin biosynthetic pathway, except for *thiC*, encoding the 4-amino-5-

11 hydroxymethyl-2-methylpyrimidine (HMP) synthase, which is required for *de novo*

12 synthesis of thiamin's pyrimidine moiety. Here we demonstrate that the SAR11 isolate

13 '*Candidatus* Pelagibacter ubique', strain HTCC1062, is auxotrophic for the thiamin

14 precursor HMP, and cannot use exogenous thiamin for growth. In culture, strain

15 HTCC1062 required 0.7 zeptomoles cell^{-1} (*ca*. 400 HMP molecules cell^{-1}). Measurements

16 of dissolved HMP in the Sargasso Sea surface layer, showed HMP ranged from

17 undetectable (detection limit: 2.4 pM) to 35.7 pM, with maximum concentrations

18 coincident with the deep chlorophyll maximum. In culture, some marine cyanobacteria,

19 microalgae and bacteria exuded HMP, and in the Western Sargasso Sea, HMP profiles

20 changed between the morning and evening, suggesting a dynamic biological flux from

21 producers to consumers.

26   **Introduction**

27   Thiamin (vitamin B_1) is an essential coenzyme found in proteins that catalyze

28   crucial transformations of carbon in all living systems. Specifically, thiamin is an

29   essential cofactor for enzymes of the TCA cycle, the non-oxidative portion of the

30   pentose-phosphate pathway, the Calvin cycle, and for enzymes required for the

31   biosynthesis of branched-chain amino acids and isoprenoids (via the non-mevalonate

32   pathway) (Lengeler et al., 1999). The pathways, enzymes, and regulation of *de novo*

33   thiamin synthesis and salvage have been the topic of extensive research in bacteria, yeasts

34   and some microalgae (Winkler & Breaker, 2005; Croft et al., 2007; Jurgenson et al.,

35   2009). In all organisms capable of *de novo* thiamin biosynthesis, the formation of thiamin

36   monophosphate (ThP) results from the enzyme-catalyzed linkage of two separately

37   synthesized moieties: 4-amino-5-hydroxymethyl-2-methylpyrimidine diphosphate (HMP-

38   PP) and 4-methyl-5-(2-phosphoethyl)-thiazole (THZ-P) (Fig. 1). Phosphorylation of ThP

39   yields the active thiamin coenzyme, thiamin diphosphate (ThPP) (Fig. 1) (reviewed in

40   Jurgenson et al., 2009).

41   Renewed interest in vitamin distributions in marine ecosystems has been driven

42   by the development of more sensitive analytical techniques to measure vitamin

43   concentrations (Sañudo-Wilhelmy et al., 2012) and a greater appreciation of the

44   importance of trace compounds to plankton productivity. Whereas the sources,

45   distributions and speciation of trace metals have been extensively researched as they

46   pertain to ocean productivity (reviewed in Morel & Price, 2003), relatively little is known

47   about vitamin biogeochemistry or the affect of vitamins on the structure and composition

48   of planktonic communities. Direct measurements of B-vitamin concentrations in coastal

49   ocean systems found picomolar concentrations and complex patterns in the distributions

50   of several vitamins, including thiamin (Sañudo-Wilhelmy et al., 2012; Barada et al.,

51   2013). In bottle experiments, iron and B-vitamins, particularly vitamin B_{12}, acted

52   synergistically to increase phytoplankton and bacterial productivity, suggesting co-

53   limitation (Panzeca et al., 2006; Bertrand et al., 2007). Supporting the view that the

54   exchange of vitamins between species is important, adaptive strategies for coping with

55   low vitamin concentrations have been identified in diatoms (Bertrand et al., 2012).

56   Furthermore, there is evidence that some marine bacteria produce vitamin B_{12} that is used

57   by phytoplankton (Croft et al., 2005).

58   Thiamin is a particularly interesting vitamin because the genomes of many

59   environmentally-abundant microorganisms do not encode for complete, canonical

60   thiamin biosynthetic pathways (Helliwell et al., 2013; Bertrand & Allen, 2012),

61   suggesting auxotrophy is common. The distribution of thiamin biosynthetic genes in algal

62   genomes does not correlate well with phylogeny, an indication that thiamin metabolism

63   has evolved and diversified in response to selective pressures that vary with habitat

64   (reviewed in Helliwell et al., 2013; Croft et al., 2006). The evolution of thiamin

65   metabolism in phytoplankton is likely complex, as evidenced by the ability of some

66   strains to use the thiamin moieties 4-methyl-5-thiazolethanol (THZ) or 4-amino-5-

67   aminomethyl-2-methylpyrimidine (AmMP), presumably natural thiamin degradation

68   products, in place of thiamin (Droop, 1958; Lewin, 1962). A specific requirement for the

69   thiamin pyrimidine precursor 4-amino-5-hydroxymethyl-2-methylpyrimidine (HMP) has

70   been described for the protist *Plasmodium falciparum* (Wrenger et al., 2006) and the

71   bacterium *Listeria monocytogenes* (Schauer et al., 2009). Moreover, thiamin is

72 exclusively obtained through salvage of thiamin moieties by the bacterium *Rhizobium*

73 *leguminosarum* bv. viciae str. 3841 (Karunakaran et al., 2006). Environmental

74 concentrations of these thiamin precursors or degradation products have not been

75 measured, and thiamin metabolism in marine bacteria is a relatively un-explored topic.

76   This study examines thiamin metabolism in the SAR11 clade of α-proteobacteria

77 (*Pelagibacterales*). These organisms are the most abundant chemoheterotrophic

78 bacterioplankton in the oceans, often comprising 25-50% of the cells in the euphotic zone

79 (Morris et al., 2002; Carlson et al., 2009). Both *in situ* studies and those with axenic

80 cultures show that the *Pelagibacterales* contribute significantly to the cycling of carbon

81 and sulfur in the ocean (reviewed in Tripp, 2013). The first cultivated *Pelagibacterales*

82 bacterium, '*Candidatus* Pelagibacter ubique' str. HTCC1062 (*Ca.* P. ubique), contains

83 one of the smallest genomes found in free-living organisms. The small genome of *Ca.* P.

84 ubique is attributed to streamlining selection (Giovannoni et al., 2005). Gene loss related

85 to streamlining selection has been proposed as an explanation for the unusual

86 combination of amino acids, reduced organosulfur compounds and organic acids required

87 for the growth of *Ca.* P. ubique (Tripp, 2013; Carini et al., 2013). Although the

88 macronutrient requirements of *Ca.* P. ubique have been identified, their requirements for

89 vitamins and other trace molecules have not been investigated.

90   We used comparative genomics to examine the distribution of genes for thiamin

91 metabolism among the *Pelagibacterales*, and studied the requirement for thiamin or its

92 precursors in *Ca.* P. ubique. Following up on the surprising finding that *Ca.* P. ubique

93 requires the thiamin precursor HMP, we applied high performance liquid

94 chromatography-coupled tandem mass spectrometry (LC-MS) to show that dissolved

95   HMP is present at pM concentrations in the oceans. These findings offer important new

96   insight into thiamin cycling, and identify HMP as a growth factor that is likely to play an

97   important role in vitamin-mediated interactions in the ocean.

98

99   **Methods**

100   *Metabolic reconstruction of thiamin biosynthesis in Ca.* P. ubique *and other*

101   *Pelagibacterales:* To identify putative protein domains involved in thiamin biosynthesis,

102   amino acid sequences of known *E. coli* (ThiC, ThiD, ThiE, ThiS, ThiG, ThiL, ThiF, IscS

103   and ThiH), *Bacillus subtilis* (ThiO) and *Saccharomyces cerevisiae* (NMT1) thiamin

104   biosynthesis proteins were used as query sequences in a HMMER search against the

105   Pfam database (v27.0), using the PFam website (http://pfam.sanger.ac.uk/search) with

106   default settings. Identified PFam domains were extracted from the Pfam-A database and

107   prepared as an hmmscan (v3.1b) compliant database. This database was used to search

108   predicted amino acid sequences of *Ca.* P. ubique ORFs for putative protein domains

109   involved in thiamin biosynthesis using hmmscan (http://hmmer.janelia.org; v3.1b)

110   (Supplementary Dataset 1). A similar approach was used to identify *Ca.* P. ubique genes

111   involved in thiamin biosynthesis using the Sifting Families HMM database (Sfam)

112   (Sharpton et al., 2012) in place of Pfam (Supplementary Dataset 2). When an ORF from

113   *Ca.* P. ubique was predicted to match a Pfam and/or Sfam identified from a Thi_ query

114   (*e*-value $\leq 1.0 \times 10^{-35}$), it was assumed that the *Ca.* P. ubique gene was a homolog of the

115   query. The best-hit for *E. coli* ThiL in the Pfam database (PF00586) is the N-terminal

116   domain of aminoimidazole ribonucleotide synthase related proteins – a putative ATP

117   binding domain. Proteins associated with this Pfam model are numerous and functionally

118   diverse. Therefore, ThiL homologs in *Ca.* P. ubique were assigned on the strength of their

119   best-hit Sfam model alone.

120   The Hal pipeline (Robbertse et al., 2011) was used to identify genes encoding Thi

121   biosynthesis homologs, in seven additional *Pelagibacterales* genomes (HTCC1002,

122   HTCC9565, HTCC7211, HIMB5, HIMB114, HIMB59 and IMCC9063). Orthologous

123   groups were established using the pipeline Hal, as described in Thrash et al., (2014). The

124   Hal pipeline connects the programs BLASTP, MCL, user specified alignment programs,

125   GBlocks, ProtTest and user specified phylogenetic programs. Hal uses an all-versus-all

126   blastP search and MCL clustering to identify orthologs, as described in detail in

127   Robbertse et al., (2011).

128   *Construction of ThiV phylogenetic trees:* RAxML (Stamatakis, 2006) was used

129   for phylogenetic inference, after alignment with MUSCLE (Edgar, 2004), curation with

130   Gblocks (Castresana, 2000), and amino acid substitution modeling with ProtTest

131   (Abascal et al., 2005). SAR11_0811 was initially identified as a ThiV homolog by

132   searching the amino acid sequence against others at MicrobesOnline

133   (http://microbesonline.org/). This search identified SAR11_0811 as a member of the

134   COG591 gene family, which had orthologs in the genomes of eight additional organisms:

135   *Methylobacillus flagellatus* KT, *Marinobacter* sp. ELB17, *Clostridium* sp. OhILAs,

136   *Haloquadratum walsbyi* DSM 16790, *Haloarcula marismortui* ATCC 43049,

137   *Halorhabdus utahensis* DSM 12940, *Haloferax volcanii* DS2 and *Halogeometricum*

138   *borinquense* PR3, DSM 11551. Eight SAR11_0811 orthologs in other SAR11 genomes

139   (HTCC1002, HTCC9565, HTCC7211, HIMB5, AAA240-E13, AAA288-G21,

140   HIMB114, and IMCC9063) were identified with the Hal pipeline (Robbertse et al., 2011;

141   Thrash et al., 2014). To provide a fuller phylogenetic context for the trees, additional

142   homologs to ThiV amino acid sequences from the genomes above were searched against

143   the SFam Hidden Markov Model (HMM) database (Sharpton, et al., 2012). Further

144   details are provided in supplementary documentation.

145   *Organism source and cultivation details: Ca.* P. ubique was revived from 10%

146   glycerol stocks and propagated in AMS1, without added vitamins, amended with

147   oxaloacetate (1 mM), glycine (50 μM), methionine (50 μM) and $FeCl_3$ (1 μM) (Carini et

148   al., 2013). Thiamin or precursors were added as indicated in figure legends and text. All

149   cultures were grown in acid-washed and autoclaved polycarbonate flasks and incubated

150   at 20°C with shaking at 60 RPM in the dark, unless noted otherwise. Cells for counts

151   were stained with SYBR green I and counted with a Guava Technologies flow cytometer

152   at 48-72 h intervals as described elsewhere (Tripp et al., 2008).

153   Cultures tested for HMP exudation were grown in acid-washed and autoclaved

154   polycarbonate flasks, incubated at 20°C with shaking at 60 RPM on a 14-h/10-h light

155   (140-180 μmol photons m^{-2} s^{-1})/dark cycle and monitored by flow cytometry as described

156   for *Ca*. P. ubique. For HMP exudation assays, axenic batch cultures of *Synechococcus sp*.

157   WH8102 and *Prochlorococcus sp*. MED4 (CCMP2389) were grown in PCRS-11 Red

158   Sea medium (Rippka et al., 2000). *Dunaliella tertiolecta* (CCMP1320) was grown in

159   AMS1 medium without vitamins (Carini et al., 2013). The OM43-clade isolate, sp.

160   HTCC2181, was grown in natural seawater with no added vitamins as described

161   elsewhere (Giovannoni et al., 2008).

162   All AMS1 constituents, reagents and vitamins were of the highest available

163   quality (labeled 'ultrapure' when possible). To minimize unintended traces of vitamins

164   from glassware, all nutrient and vitamin stocks were prepared in combusted glassware

165   (450°C for 4 h) with nanopure water, 0.1 μm-filter-sterilized and frozen in amber tubes

166   immediately after preparation. HMP was synthesized as described in Reddick et al.,

167   (2001). AmMP was synthesized as described in Zhao et al., (2012). HMP was purified by

168   chromatography and then recrystallized. It was characterized by ^{1}H and ^{13}C NMR

169   spectroscopy and by mass spectrometry. AMP was purified by crystallization and was

170   characterized by ^{1}H NMR and ^{13}C NMR spectroscopy. No impurities were detected.

171   *HMP and thiamin concentrations in seawater:* Seawater for vitamin analysis was

172   collected from Hydrostation S (32°10'N, 64°30'W) from casts at 20:00 (local time) on 19

173   September 2012, and 08:00 (local time) on 20 September 2012. At the time of collection,

174   samples were filtered through nanopure water-rinsed 0.2 μm pore-size supor filters into

175   acid-washed amber polypropylene bottles and frozen immediately.

176   HMP and thiamin were extracted from 300 mL seawater to a reverse-phase C18

177   silica bead solid phase (Agilent HF-Bondesil) as described in Sañudo-Wilhelmy et al.,

178   (2012). For quantification purposes, standard curves were constructed from aged

179   seawater (collected from Hydrostation S in July of 2009) spiked with known amounts of

180   HMP and thiamin (ranging from 0 pM to 100 pM). These standard curves

181   (Supplementary Figures S1 and S2) were extracted alongside samples using identical

182   procedures.

183   Extracts were reconstituted in 125 μL HPLC-grade water. Samples were

184   centrifuged to pellet insoluble matter and the supernatant was transferred to sampling

185   vials. HMP was quantified using an Applied Biosystems MDS Sciex 4000 Q TRAP mass

186   spectrometer coupled to a Shimadzu HPLC system. An Agilent Zorbax SB-Aq (2.1 × 100

187   mm, 3.5-micron) HPLC column was used for separation over a 10-minute gradient flow

188   with mobile phases of pH 4 (formic acid) methanol (MeOH) and pH 4 (formic acid) 5

189   mM ammonium formate (AmF). The flow rate was 0.4 mL min^{-1} and a gradient starting

190   at 98% AmF: 2% MeOH for 1 minute changing to 75% AmF: 25% MeOH over 3

191   minutes, 50% AmF: 50% MeOH over 0.2 minutes, and finally to 10% AmF:90% MeOH

192   over 0.8 minutes. The retention time of HMP was appx. 1.8 minutes.

193   For HMP quantification, the mass spectrometer was run in 'Multiple Reaction

194   Monitoring' (MRM) mode. The HMP parent ion m/z was 140.2, and ion transitions of

195   81.1 and 54.1 were used for quantification and qualification, respectively. Peaks were

196   analyzed using the Analyst software package v 1.5.2 (AB SCIEX; Concord, ON,

197   Canada). Measured HMP values are the average of technical LC-MS replicates. The

198   greatest standard deviation of replicate measurements was 3.5 pM (CV = 10%) in the 120

199   m 08:00 sample, and the lowest was 0.22 pM (CV=3.5%) in the 200 m 20:00 sample.

200   Thiamin was detected and quantified as described in Sañudo -Wilhelmy et al., (2012).

201   The limit of detection (LOD) is defined as 3 times the standard deviation of the

202   procedural controls and the limit of quantification (LOQ) as 10 times the standard

203   deviation of the procedural controls. The LOD for HMP was 2.4 pM (LOQ: 8.0 pM) and,

204   for thiamin, 0.81 pM (LOQ: 2.7 pM; from Sañudo-Wilhelmy, et al., (2012)).

205   *Cell harvesting of marine microbes for HMP exudation assays and detection of*

206   *HMP background in AMS1:* During mid-logarithmic growth (appx 1.0×10^7 cells ml^{-1}),

207   100 ml of culture was gently filtered (to prevent cell lysis) through 0.1 or 0.2 μm pore-

208   size supor filters to remove cells. The filtrate was collected in an acid-washed amber

209   polypropylene bottle and frozen immediately. Uninoculated media (negative control) for

210   each media type (AMS1, PCRS-11 Red Sea medium and natural seawater medium for

211   HTCC2181) was extracted alongside spent medium treatments for comparison. HMP

212   extraction and detection by LC-MS were performed as described for natural seawater

213   samples.

214

215   **Results**

216   Thiamin biosynthetic pathways were incomplete in all eight *Pelagibacterales*

217   genomes we studied (Table 1). Despite the apparent inability to synthesize thiamin *de*

218   *novo*, multiple genes encoding ThPP-dependent enzymes were identified in *Ca*. P.

219   ubique, indicating thiamin is necessary for normal metabolism (Supplementary Fig. S3).

220   Four *Pelagibacterales* strains contained the same thiamin biosynthesis and transport

221   genes as *Ca*. P. ubique (Table 1). Two additional *Pelagibacterales* strains, IMCC9063

222   and HIMB114, have complements of thiamin biosynthesis and transport genes similar to

223   *Ca*. P. ubique, except both are missing *thiL* (Table 1). Additionally, IMCC9063 encodes

224   the AmMP salvage enzyme, *tenA* (Table 1). In *Pelagibacterales* str. HIMB59, *thiC*, *thiD*,

225   *thiG*, *thiE* & *thiE2*, and *thiS* are absent. However, a gene encoding a thiamin-specific

226   periplasmic binding protein (*thiB*) (Webb et al., 1998) was identified in HIMB59 (Table

227   1).

228   Genes encoding the HMP synthase (*thiC*) are absent from all *Pelagibacterales*

229   genomes (Table 1). ThiC catalyzes the molecular rearrangement of the purine nucleotide

230   biosynthetic intermediate 5-aminoimidazole ribotide to form HMP (Fig. 1) (Martinez-

231   Gomez & Downs, 2008) and is essential for *de novo* thiamin biosynthesis in bacteria,

232   archaea and plants. Genes that encode alternate HMP synthesis or salvage proteins were

233   not identified in the *Ca*. P. ubique genome. For example, *Ca*. P. ubique lacks genes

234   encoding for NMT1, which synthesizes HMP from vitamin B_6 and histidine in

235   *Saccharomyces cerevisiae* (Fig. 1) (Wightman & Meacock, 2003). Genes encoding TenA

236   homologs, which catalyze the hydrolysis of AmMP to yield HMP (Jenkins et al., 2007),

237   were also not present in *Ca*. P. ubique (Fig. 1). Some organisms can transport thiamin

238   intact with the thiamin-specific ABC transporter encoded by *thiBPQ*. No homologs of the

239   thiamin-specific binding protein, ThiB, were identified in *Ca*. P. ubique genomes (Fig. 1).

240   Further, *Ca*. P. ubique does not encode homologs of the predicted bacterial HMP/AmMP

241   ABC transport complexes ThiXYZ (Jenkins et al., 2007) and YkoEDC, or for the

242   putative HMP/AmMP permeases HmpT and CytX (Rodionov et al., 2002; 2008).

243   A single predicted ThPP-activated RNA riboswitch was identified in the *Ca*. P.

244   ubique genome (Meyer et al., 2009) in an unusual configuration upstream of a coding

245   sequence annotated as a sodium:solute symporter family protein (encoded by *Ca*. P.

246   ubique ORF SAR11_0811). A similarly configured riboswitch was previously identified

247   in the genome of *Methylobacillus flagellatus,* upstream of a coding sequence for an

248   uncharacterized putative transporter named *thiV* (Rodionov et al., 2002). Maximum-

249   likelihood phylogenetic analysis of the *Pelagibacterales thiV* homologs showed that they

250   form a monophyletic group with the *thiV* sequences from of *M. flagellatus* and a diverse

251   group of microbes, including, Haloarchaea, Gram-positive bacteria and β- and γ-

252   proteobacteria (Fig. 2A and Supplementary Fig. S4). Genes orthologous to *thiV* in all

253   organisms (except for *Marinobacter algicola*) are either i) in an operon with genes

254   encoding enzymes that enable the salvage of HMP and THZ moieties for thiamin

255   synthesis (*thiD*, *thiM* and *thiE*; Fig. 2B); ii) in an operon with one or two copies of the

256   *tenA* gene (encoding an AmMP salvage enzyme; Fig. 2B); or iii) are preceded by a ThPP-

257   riboswitch motif (Fig. 2B, C).

258   We hypothesized that *Ca*. P. ubique is auxotrophic for HMP because genes

259   coding for known HMP synthesis pathways (*thiC* and *NMT1*) and AmMP salvage

260   mechanisms (*tenA*), were absent (Fig. 1; Table 1). To test this hypothesis, the growth

261   responses of *Ca*. P. ubique to HMP, AmMP, and thiamin, were investigated across seven

262   orders of magnitude (Fig. 3). Cultures grown in medium containing no added HMP,

263   without additional thiamin or precursors, attained maximum cell densities of

264    $3.09 \pm 0.75 \times 10^7$ cells ml^{-1} (mean $\pm$ s.d., n=3) (Fig. 3). Cell yields responded linearly to

265   HMP additions between 1 and 100 pM (Supplementary Fig. S5) and reached maximal

266   cell yields (ca. 3.5×10^8 cells ml^{-1}) at HMP concentrations $\geq$1 nM (Fig. 3). The cellular

267   HMP requirement was calculated to be 0.66 zeptomoles (396 molecules) cell^{-1} from the

268   slope of the linear regression between 1 pM and 100 pM (Supplementary Fig. S5).

269   Thiamin and AmMP were ineffective at restoring thiamin-limited growth at pico- or

270   nanomolar concentrations; these compounds restored growth only when supplied at 1.0

271   μM (Fig. 3). The average growth rate of *Ca*. P. ubique was 0.29 ± 0.03 d^{-1} (mean $\pm$ s.d.,

272   n=123) and did not vary with vitamin or precursor treatments (for example, see

273   Supplementary Fig. S6).

274   To rule out NMT1 activity as a potential source of HMP, thiamin was replaced

275   with histidine and vitamin B$_6$ (NMT1's substrates (Ishida et al., 2008)). Consistent with

276   the prediction that *Ca*. P. ubique lacks the ability to synthesize HMP through NMT1

277   activity, histidine + vitamin B$_6$ did not alleviate thiamin-limited growth (Supplementary

278   Fig. S7). Thiamin-limited growth was not relieved by pantothenate or THZ addition

279   (Supplementary Fig. S7) as has been reported previously for other organisms (Downs,

280   1992; Droop, 1958).

281   To date, measurements of HMP or AmMP concentrations in the environment

282   have not been reported. To determine if HMP is present in an environment where

283   *Pelagibacterales* bacteria are also found, thiamin and HMP were extracted from Sargasso

284   Sea seawater collected at two different times of day (20:00 and 08:00 local time,

285   approximately 1 hour after sunset and sunrise, respectively) and quantitatively measured

286   by LC-MS. HMP ranged from undetectable (detection limit: 2.4 pM) to 35.7 pM (Fig. 4).

287   The maximum concentration of HMP was observed in samples collected at 08:00 near

288   the deep chlorophyll maximum (Fig. 4). HMP concentrations at 20:00 were substantially

289   higher at 0 meters depth, but lower at depths of 40, 80, 120, 160 and 200 m, compared to

290   samples collected at 08:00 (Fig. 4). HMP was not detected in the 250 m sample collected

291   at 08:00. Thiamin was measured in the same samples and ranged from undetectable

292   (detection limit: 0.81 pM (Sañudo-Wilhelmy et al., 2012)) to 23 pM, and was present in

293   samples from 0 to 160 m, but not detected in samples from 250 and 300 m (Fig. 4).

294   To determine whether marine microbes exude HMP, we measured HMP

295   concentrations in growth media before and after cell growth in strains known to have a

296   complete complement of thiamin biosynthetic genes (Table 2). The marine

297   cyanobacterium *Synechococcus sp*. WH8102 and the marine chlorophyte *Dunaliella*

298   *tertiolecta* exuded nanomolar amounts of HMP during growth (Table 2). Moderate

299   amounts of excess HMP were also detected in spent media from cyanobacterium

300   *Prochlorococcus* MED4 and the OM43-clade of marine β-proteobacteria isolate, str.

301   HTCC2181 (Giovannoni, et al., 2008). Two *Pelagibacterales* cultures were also tested:

302   *Ca*. P. ubique and *Pelagibacterales sp*. str. HTCC7211. In both cases, HMP was not

303   detected after cell growth (Table 2).

304

305   **Discussion**

306   Thiamin has long been recognized as an important vitamin for microalgal growth

307   (reviewed in Croft et al., 2006). The physiological requirement for thiamin led to the

308   hypothesis that environmental concentrations of thiamin may exert control over some

309   phytoplankton populations (Natarajan, 1968; Panzeca, et al., 2006). Environmental

310   distributions of thiamin, as determined by bioassay, were variable, and in some cases,

311   coupled to productivity (Natarajan, 1968; 1970; Natarajan & Dugdale, 1966). Studies of

312   thiamin auxotrophy in the laboratory showed that thiamin moieties or degradation

313   products were able to satisfy the thiamin requirement of some microalgae (Lewin, 1962;

314   Droop, 1958). However, research pursuing the ecological importance of these findings

315   tapered off. The experimental results presented here reintroduce the idea that thiamin

316   pyrimidines are important growth determinants in marine ecosystems. We show that the

317   thiamin pyrimidine precursor, HMP, is required for growth of the marine

318   chemoheterotrophic bacterium *Ca*. P. ubique (Fig. 3), a representative isolate of one of

319   the most abundant groups of organisms on the planet. Surprisingly, neither thiamin itself,

320   nor AmMP satisfied this requirement (Fig. 3). Comparative genomics extended the

321   significance of this requirement to multiple members the *Pelagibacterales* clade (Table

322   1). The importance of these findings were further supported by the detection of dissolved

323   HMP in the Sargasso Sea (Fig. 4), one of the most oligotrophic ocean systems on earth, at

324   concentrations often exceeding those of thiamin. This discovery shows that fundamental

325   information needed to understand thiamin biogeochemistry in marine ecosystems is

326   incomplete – specifically, that environmental measurements of thiamin alone may only

327   partially explain interactions related to the thiamin requirements of planktonic cells.

328   The inability of *Ca*. P. ubique to utilize thiamin or its degradation product

329   AmMP was surprising given that many algal species are able to utilize these compounds

330   (Lewin, 1962; Droop, 1958). The *Ca*. P. ubique genome encodes no thiamin transporter

331   (Fig. 1 & Table 1); consistent with the observation that exogenous thiamin does not

332   support growth (Fig. 3). Likewise, we propose that the absence of the *tenA* gene (Fig. 1 &

333   Table 1), necessary for the conversion of AmMP to HMP, explains why AmMP does not

334   substitute for HMP in thiamin biosynthesis. However, genome analysis of

335   *Pelagibacterales* str. HIMB59 indicates that this strain lacks genes required for *de novo*

336   synthesis of thiamin (*thiC*, *thiD*, *thiG*, *thiE* and *thiS*), as well as the AmMP salvage

337   enzyme (*tenA*; Table 1) and *thiV*; therefore, we postulate that this strain requires

338   exogenous thiamin. Supporting this idea, *thiB*, encoding the periplasmic subunit of a

339   thiamin-specific thiamin ABC transporter, was identified in HIMB59 (Table 1).

340   The new data reported here indicate that thiamin cycling in the oceans may follow

341   complex patterns and involve multiple processes and intermediates. Whereas we show

342   that marine microbes can release HMP into the surrounding environment (Table 2), some

343   phytoplankton exude thiamin (Carlucci & Bowes, 1970a; 1970b). Although thiamin is

344   labile in seawater (Gold, 1968; Gold et al., 1966), its decomposition products in seawater

345   have not been fully characterized and the effect of various environmental factors on

346   degradation are poorly understood. For example, thiamin is a light sensitive molecule that

347   is readily cleaved by UV-B radiation to AmMP and other products (Machlin, 1984;

348 Okumura, 1961). Although no measurements of AmMP concentrations in the

349 environment have been reported, the physiological responses of phytoplankton to AmMP

350 (Lewin, 1962; Droop, 1958) and the presence of *tenA* genes in some bacterial genomes

351 that lack the *thiC* gene (Supplementary Table S1), including *Pelagibacterales sp.* str.

352 IMCC9063 (Table 1), suggest that environmental AmMP is present, and might also be an

353 important growth determinant in marine ecosystems.

354 Light-mediated decay of thiamin may be an important factor in thiamin

355 geochemistry and influence HMP production patterns in marine surface waters. The

356 depth profiles showing that the dissolved HMP maximum coincides with the deep

357 chlorophyll maximum (Fig. 4) suggest that marine phytoplankton may be important HMP

358 producers. Intriguingly, previous studies reported diel periodicity in the transcription and

359 translation of *thiC* (the HMP synthase) in laboratory cultures of *Prochlorococcus* MED4

360 (Waldbauer et al., 2012). Similarly, the abundance of environmental transcripts mapping

361 to *thiC* of *Synechococcus sp*. followed a diel pattern (Ottesen et al., 2013). In both

362 reports, maximum *thiC* transcript levels were observed in the mid afternoon, shortly after

363 the periods of highest light intensity. We speculate that the large differences in dissolved

364 HMP concentrations from profiles collected at different times (Fig. 4) may be an

365 indication that HMP exudation by *thiC*-containing cyanobacteria also follows a diel

366 pattern. Although measurements of dissolved vitamins (and precursors) reflect

367 equilibrium concentrations, not fluxes, reports of rapid rates of ^{3}H-thiamin uptake by

368 plankton communities (Koch et al., 2012) suggest that rapid water column vitamin

369 depletion due to biological scavenging is feasible. The notable production of HMP by

370 *Synechococcus sp*. WH8102 and modest exudation by *Prochlorococcus* MED4 batch

371   cultures (Table 2) is consistent with the idea that cyanobacteria are important HMP

372   producers, however diel patterns of HMP production were not tested in our experiments.

373   The absence of *thiC*, and thus the requirement for exogenous thiamin pyrimidines,

374   is not unique to the *Pelagibacterales*, but is broadly and unevenly distributed among

375   diverse microbial taxa inhabiting marine waters. Incomplete thiamin biosynthetic gene

376   complements were previously reported in the genomes of the uncultivated SAR86-clade

377   of marine γ-proteobacteria (Dupont et al., 2012) and in some phytoplankton (reviewed in

378   Bertrand & Allen, 2012; Helliwell et al., 2013). Genes for ThiC are also absent from the

379   genomes of many other ecologically important marine bacteria and archaea

380   (Supplementary Table S1). The observation that canonical thiamin biosynthetic pathways

381   are incomplete in sequenced organisms was further mirrored in metagenomic datasets.

382   Comparisons of the abundances of *thiC*, *thiD* and *thiG* across a metagenomic depth

383   profile from the Sargasso Sea found that *thiC* genes were depleted relative to *thiD* and

384   *thiG* genes at 0, 40 and 80 m, but near the deep chlorophyll maximum, copies of *thiC*

385   exceeded those of *thiD* (Supplementary Fig. S8). The relative deficiency of *thiC* to other

386   essential thiamin biosynthesis genes in shallow waters is consistent with the idea that

387   HMP salvage is important for thiamin synthesis at those depths.

388   We postulate that ThiV sodium:solute symporters constitute a new family of

389   thiamin pyrimidine transport proteins. Previously it was hypothesized that ThPP-

390   regulated sodium:solute symporters, like ThiV, might transport thiamin moieties in

391   eukaryotes (Worden et al., 2009). A phylogeny of bacterial and archaeal ThiV orthologs

392   supports this interpretation by showing that *thiV* genes co-localize with genes encoding

393   for thiamin pyrimidine salvage enzymes (*tenA* in archaeal genomes and with *thiD*, *thiM*

and *thiE* in the *Alkaliphilus oremlandii* genome) (Fig. 2), implying that ThiV orthologs transport thiamin pyrimidines (HMP or AmMP). We speculate that *Ca.* P. ubique regulates the acquisition of HMP from the environment by controlling the expression of ThiV with a ThPP-binding riboswitch, in a manner akin to the ThPP-riboswitch regulation of *de novo* HMP synthesis (via ThiC) in other organisms (Winkler et al., 2002). When thiamin is bound to ThPP-riboswitches, transcription and translation of the downstream coding sequence is repressed, thus the detection of ThiV and other ThPP-regulated gene products in metaproteomes may be useful indicators of thiamin deprivation in the environment. For example, peptides mapping to *Pelagibacter* ThiV orthologs were detected in environmental metaproteomes from the Sargasso Sea (Sowell et al., 2009), but not the Southern Ocean (Williams et al., 2012), perhaps indicating differences in the thiamin status of the two biomes.

The dependence of *Ca.* P. ubique, and likely other *Pelagibacterales*, on HMP implies that these cells gain an advantage by outsourcing HMP production to other plankton, in essence relying on HMP as a publically available commodity. This perspective is consistent with genome streamlining theory, and previous reports of unusual nutrient requirements associated with genome reduction in *Pelagibacterales* (Tripp et al., 2008; Carini et al., 2013). Streamlining theory predicts that atypical nutrient requirements can arise in microorganisms that have large effective population sizes in response to selection favoring small cell size and the efficient use of limiting nutrient resources (Giovannoni et al., 2005). The 'Black Queen Hypothesis' explored the co-evolutionary implications of genome streamlining theory, examining the broader context of adaptive gene loss in a framework that considered competition for public goods

417 (Morris et al., 2012). In this context, because the *Pelagibacterales* depend on

418 environmental HMP, there is potential for *Pelagibacter* growth limitation by HMP,

419 intimately tying the success of these organisms to HMP producers.

420   Because *Ca*. P. ubique cells are among the smallest known, and replicate

421 efficiently at very low nutrient concentrations, elucidating the trace nutrient requirements

422 of these cells is technically challenging. Even in a defined minimal medium, when

423 precautions were taken to minimize trace vitamin background, *Ca*. P. ubique reached 2-3

424 $\times 10^7$ cells ml^{-1} in the absence of added vitamins or precursors (Fig. 3 and Supplementary

425 Figures S6 & S7). These yields are within a factor of two of theoretical yields (1.8×10^7

426 cells ml^{-1}) based on the cellular HMP requirement (Supplementary Fig. S5) and the

427 amount of "background" HMP measured in the medium (12 pM). This "background"

428 HMP disappeared in the presence of *Ca*. P. ubique, implying consumption of the nutrient

429 (Table 2). Previously, background levels of vitamins in heterotrophic growth medium

430 were proposed to underlie scant growth of vitamin auxotrophs in the absence of added

431 vitamins (Wu et al., 2005; Norman et al., 1981), and the difficulty associated with

432 thiamin removal from growth medium has been noted (Button, 1968). The number of

433 HMP molecules required per *Ca*. P. ubique cell is on the order of 400 molecules cell^{-1}

434 (Supplementary Fig. S5). Assuming each HMP molecule is used to make one thiamin

435 molecule, and an estimate of 6 fg carbon cell^{-1} (unpublished data), the thiamin/carbon

436 ratio of *Ca*. P. ubique was calculated to be 25 ng thiamin/mg carbon - similar to the

437 values measured for marine phytoplankton (5-100 ng thiamin/mg carbon (Carlucci &

438 Bowes, 1972; Brown et al., 1999)). Thus, the cell titers we observed in the absence of

439 added HMP are consistent with the explanation that even pure reagents (e.g. 98-99%) and

440   water from reverse osmosis purifiers can contain very small concentrations of vitamins

441   and vitamin precursors – enough to support the growth of cells that require miniscule

442   amounts of vitamins.

443   Contaminating HMP was detected in the thiamin stock solution that was added to

444   thiamin-amended treatments. The level of HMP "contamination" in the concentrated

445   thiamin stock was measured (via LC-MS) to be ~2.6 nmoles HMP per 1 μmole thiamin

446   (=0.0012 g HMP per g thiamin) (Supplementary Fig. S9). The unintended addition of

447   approximately 2.6 nM HMP as a contaminant of the thiamin stock is the probable

448   explanation for the growth restoration by thiamin at culture concentrations of 1 μM (Fig.

449   3). The source of contaminating HMP appears to be the result of the commercial thiamin

450   manufacturing process. Contaminating amounts of HMP in the AmMP stock could not be

451   determined because HMP and AmMP have similar liquid chromatography retention

452   times, thus the application of large amounts of AmMP to the chromatography column

453   obscured the detection of possible traces of HMP. We propose that HMP contamination

454   in the AmMP preparation is also a plausible explanation for the slightly elevated yields at

455   high AmMP concentrations.

456   This investigation illustrates the value of combining metabolic reconstruction

457   from genomes with experimentation in the laboratory and field measurements of specific

458   compounds to explore biogeochemical cycles. The demonstration that HMP exclusively

459   satisfies the thiamin requirement of a highly abundant marine organism (Fig. 3), is found

460   in the ocean (Fig. 4), and is exuded by some marine organisms (Table 2), identifies this

461   compound as an important, previously unknown growth factor in marine systems. It is

462   particularly surprising that thiamin and AmMP were not used by *Ca*. P. ubique, implying

463   that HMP-producing organisms potentially could exert control over *Pelagibacterales*

464   populations. Extending these findings outside of the *Pelagibacterales*, multiple genomes

465   of cosmopolitan marine bacteria display incomplete thiamin synthesis pathways

466   (Supplementary Table S1), suggesting thiamin moiety scavenging may be a common

467   strategy in marine waters. The specific mechanism of HMP exudation by marine

468   phytoplankton is unknown. It is possible that in high light environments, intracellular

469   thiamin is relatively unstable, preventing repression of the ThPP-regulated HMP synthase

470   gene (*thiC*), and resulting in HMP overproduction. But, HMP might also partition to the

471   membrane and from there to the extracellular environment because it is relatively

472   hydrophobic, or its exudation could be driven by co-evolutionary interactions. As yet,

473   there is no evidence that favors one of these alternatives over another. A more complete

474   understanding of HMP production patterns, as they pertain to vitamin cycling, will likely

475   be important for understanding turnover and connectedness in plankton communities

476   (Steele et al., 2011; Fuhrman et al., 2006) .

477

478   Supplementary information is available at ISMEJ's website.

479

480

481   **Acknowledgements**

486   Kimberly Halsey for phytoplankton cultures and manuscript revisions, Brateen Shome

487   for synthesizing and characterizing HMP and AmMP and the crew of the R/V Atlantic

488   Explorer for assistance during seawater collection. Mass spectrometry was performed at

489   the Oregon State University Mass Spectrometry Facility.

490

491

492

493

494   References:

495   Abascal F, Zardoya R, Posada D. (2005). ProtTest: selection of best-fit models of protein
496   evolution. Bioinformatics 21:2104–2105.

497   Barada LP, Cutter L, Montoya JP, Webb EA, Capone DG, Sañudo-Wilhelmy SA. (2013).
498   The distribution of thiamin and pyridoxine in the western tropical North Atlantic Amazon
499   River plume. Front Microbiol 4:25.

500   Bertrand EM, Allen AE. (2012). Influence of vitamin B auxotrophy on nitrogen
501   metabolism in eukaryotic phytoplankton. Front Microbiol 3:375.

502   Bertrand EM, Allen AE, Dupont CL, Norden-Krichmar TM, Bai J, Valas RE, et al.
503   (2012). Influence of cobalamin scarcity on diatom molecular physiology and
504   identification of a cobalamin acquisition protein. Proceedings of the National Academy
505   of Sciences 109:E1762–E1771.

506   Bertrand EM, Saito MA, Rose JM, Riesselman CR, Lohan MC, Noble AE, et al. (2007).
507   Vitamin B_{12} and iron colimitation of phytoplankton growth in the Ross Sea. Limnol.
508   Oceanogr 52:1079–1093.

509   Brown MR, Mular M, Miller I, Farmer C, Trenerry C. (1999). The vitamin content of
510   microalgae used in aquaculture. Journal of Applied Phycology 11:247–255.

511   Button D. (1968). Selective Thiamine removal from culture media by ultraviolet
512   irradiation. Appl Microbiol 16:530–531.

513   Carini P, Steindler L, Beszteri S, Giovannoni SJ. (2013). Nutrient requirements for
514   growth of the extreme oligotroph 'Candidatus Pelagibacter ubique' HTCC1062 on a
515   defined medium. ISME J 7:592–602.

516 Carlson CA, Morris R, Parsons R, Treusch AH, Giovannoni SJ, Vergin K. (2009).
517 Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the
518 northwestern Sargasso Sea. ISME J 3:283–295.

519 Carlucci A, Bowes PM. (1970a). Production of vitamin B_{12}, thiamine, and biotin by
520 phytoplankton. Journal of Phycology 6:351–357.

521 Carlucci A, Bowes PM. (1972). Vitamin B_{12}, thiamine, and biotin contents of marine
522 phytoplankton. Journal of Phycology 8:133–137.

523 Carlucci A, Bowes PM. (1970b). Vitamin production and utilization by phytoplankton in
524 mixed culture. Journal of Phycology 6:393–400.

525 Castresana J. (2000). Selection of conserved blocks from multiple alignments for their
526 use in phylogenetic analysis. Mol Biol Evol 17:540–552.

527 Croft M, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG. (2005). Algae acquire
528 vitamin B_{12} through a symbiotic relationship with bacteria. Nature 438:90–93.

529 Croft M, Warren M, Smith A. (2006). Algae need their vitamins. Eukaryotic Cell
530 5:1175–1183.

531 Croft MT, Moulin M, Webb ME, Smith AG. (2007). Thiamine biosynthesis in algae is
532 regulated by riboswitches. Proceedings of the National Academy of Sciences 104:20770–
533 20775.

534 Downs DM. (1992). Evidence for a new, oxygen-regulated biosynthetic pathway for the
535 pyrimidine moiety of thiamine in *Salmonella typhimurium*. J Bacteriol 174:1515–1521.

536 Droop MR. (1958). Requirement for thiamine among some marine and supra-littoral
537 protista. Journal of the Marine Biological Association of the United Kingdom 37:323–
538 329.

539 Dupont CL, Rusch DB, Yooseph S, Lombardo M-J, Richter RA, Valas R, et al. (2012).
540 Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage.
541 ISME J 6:1186–1199.

542 Eddy SR. (2011). Accelerated profile HMM Searches. PLoS Comput Biol 7:e1002195.

543 Edgar RC. (2004). MUSCLE: multiple sequence alignment with high accuracy and high
544 throughput. Nucleic Acids Res 32:1792–1797.

545 Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown MV, Naeem S. (2006).
546 Annually reoccurring bacterial communities are predictable from ocean conditions.
547 Proceedings of the National Academy of Sciences 103:13104–13109.

548 Giovannoni SJ, Hayakawa DH, Tripp HJ, Stingl U, Givan SA, Cho J-C, et al. (2008). The
549 small genome of an abundant coastal ocean methylotroph. Environ Microbiol 10:1771–

550   1782.

551   Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, et al. (2005).
552   Genome streamlining in a cosmopolitan oceanic bacterium. Science (New York, NY)
553   309:1242–1245.

554   Gold K. (1968). Some factors affecting the stability of thiamine. Limnology and
555   Oceanography 13:185–188.

556   Gold K, Roels OA, Bank H. (1966). Temperature dependent destruction of thiamine in
557   seawater. Limnology and Oceanography 11:410–413.

558   Helliwell KE, Wheeler GL, Smith AG. (2013). Widespread decay of vitamin-related
559   pathways: coincidence or consequence? Trends Genet 29:469–478.

560   Ishida S, Tazuya-Murayama K, Kijima Y, Yamada K. (2008). The direct precursor of the
561   pyrimidine moiety of thiamin is not urocanic acid but histidine in *Saccharomyces*
562   *cerevisiae*. J Nutr Sci Vitaminol (Tokyo) 54:7–10.

563   Jenkins AH, Schyns G, Potot S, Sun G, Begley TP. (2007). A new thiamin salvage
564   pathway. Nat Chem Biol 3:492–497.

565   Jurgenson CT, Begley TP, Ealick SE. (2009). The structural and biochemical foundations
566   of thiamin biosynthesis. Annu. Rev. Biochem. 78:569–603.

567   Karunakaran R, Ebert K, Harvey S, Leonard ME, Ramachandran V, Poole PS. (2006).
568   Thiamine is synthesized by a salvage pathway in *Rhizobium leguminosarum* bv. viciae
569   strain 3841. J Bacteriol 188:6661–6668.

570   Koch F, Hattenrath-Lehmann TK, Goleski JA, Sanudo-Wilhelmy S, Fisher NS, Gobler
571   CJ. (2012). Vitamin B_1 and B_{12} uptake and cycling by plankton communities in coastal
572   ecosystems. Front Microbiol 3:363.

573   Lengeler JW, Drews G, Schlegel HG. (1999). Biology of the prokaryotes. Wiley-
574   Blackwell: Malden, MA.

575   Lewin RA, ed. (1962). Organic micronutrients. In: Physiology and biochemistry of algae,
576   Academic press: New York, pp. 141–156.

577   Machlin LJ, ed. (1984). Handbook of vitamins : nutritional, biochemical, and clinical
578   aspects. Marcel Dekker, Inc.: New York.

579   Martinez-Gomez N, Downs D. (2008). ThiC is an [Fe-S] cluster protein that requires
580   AdoMet to generate the 4-amino-5-hydroxymethyl-2-methylpyrimidine moiety in
581   thiamin synthesis. Biochemistry 47:9054–9056.

582   Meyer M, Ames T, DP S, Weinberg Z, Schwalbach M, Giovannoni S, et al. (2009).
583   Identification of candidate structured RNAs in the marine organism '*Candidatus*

584   Pelagibacter ubique'. BMC Genomics 10:268.

585   Morel FMM, Price NM. (2003). The biogeochemical cycles of trace metals in the oceans.
586   Science (New York, NY) 300:944–947.

587   Morris JJ, Lenski RE, Zinser ER. (2012). The Black Queen Hypothesis: Evolution of
588   dependencies through adaptive gene loss. mBio 3:e00036–12.

589   Morris R, Rappé MS, Connon SA, Vergin KL, Siebold WA, Carlson CA, et al. (2002).
590   SAR11 clade dominates ocean surface bacterioplankton communities. Nature 420:806–
591   810.

592   Natarajan K. (1970). Distribution and significance of vitamin B_{12} and thiamine in the
593   subarctic Pacific Ocean. Limnology and Oceanography 15:655–659.

594   Natarajan K, Dugdale R. (1966). Bioassay and distribution of thiamine in the sea.
595   Limnology and Oceanography 11:621–629.

596   Natarajan KV. (1968). Distribution of thiamine, biotin, and niacin in the sea. Appl
597   Microbiol 16:366–369.

598   Norman SM, Maier VP, Echols LC. (1981). Development of a defined medium for
599   growth of *Cercospora-rosicola* Passerini. Appl Environ Microbiol 41:334–336.

600   Okumura K. (1961). Decomposition of thiamine and its derivatives by ultraviolet
601   radiation. J Vitaminol (Kyoto) 24:158–163.

602   Ottesen EA, Young CR, Eppley JM, Ryan JP, Chavez FP, Scholin CA, et al. (2013).
603   Pattern and synchrony of gene expression among sympatric marine microbial populations.
604   Proceedings of the National Academy of Sciences 110:E488–97.

605   Panzeca C, Tovar-Sanchez A, Agusti S, Reche I, Duarte C, Taylor G, et al. (2006). B
606   vitamins as regulators of phytoplankton dynamics. EOS Transactions 87:593–596.

607   Reddick JJ, Nicewonger R, Begley TP. (2001). Mechanistic studies on thiamin phosphate
608   synthase: evidence for a dissociative mechanism. Biochemistry 40:10095–10102.

609   Rippka R, Coursin T, Hess W, Lichtlé C, Scanlan DJ, Palinska KA, et al. (2000).
610   *Prochlorococcus marinus* Chisholm *et al*. 1992 subsp. *pastoris* subsp. nov. strain PCC
611   9511, the first axenic chlorophyll a2/b2-containing cyanobacterium (Oxyphotobacteria).
612   Int J Syst Evol Microbiol 50 Pt 5:1833–1847.

613   Robbertse B, Yoder RJ, Boyd A, Reeves J, Spatafora JW. (2011). Hal: an automated
614   pipeline for phylogenetic analyses of genomic data. PLoS Curr 3:RRN1213.

615   Rodionov DA, Hebbeln P, Eudes A, Beek ter J, Rodionova IA, Erkens GB, et al. (2008).
616   A novel class of modular transporters for vitamins in prokaryotes. J Bacteriol 191:42–51.

617 Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. (2002). Comparative
618 genomics of thiamin biosynthesis in procaryotes. J Biol Chem 277:48949–48959.

619 Sañudo-Wilhelmy SA, Cutter LS, Durazo R, Smail EA, Gomez- Consarnau L, Webb EA,
620 et al. (2012). Multiple B-vitamin depletion in large areas of the coastal ocean.
621 Proceedings of the National Academy of Sciences 109:14041–14045.

622 Schauer K, Stolz J, Scherer S, Fuchs TM. (2009). Both thiamine uptake and biosynthesis
623 of thiamine precursors are required for intracellular replication of *Listeria monocytogenes*.
624 J Bacteriol 191:2218–2227.

625 Sharpton TJ, Jospin G, Wu D, Langille MG, Pollard KS, Eisen JA. (2012). Sifting
626 through genomes with iterative-sequence clustering produces a large, phylogenetically
627 diverse protein-family resource. BMC bioinformatics 13:264–264.

628 Sowell SM, Wilhelm LJ, Norbeck AD, Lipton MS, Nicora CD, Barofsky DF, et al.
629 (2009). Transport functions dominate the SAR11 metaproteome at low-nutrient extremes
630 in the Sargasso Sea. ISME J 3:93–105.

631 Stamatakis A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic
632 analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.

633 Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY, et al. (2011). Marine
634 bacterial, archaeal and protistan association networks reveal ecological linkages. ISME J
635 5:1414–1425.

636 Thrash JC, Temperton B, Swan BK, Landry ZC, Woyke T, Delong EF, et al. (2014).
637 Single-cell enabled comparative genomics of a deep ocean SAR11 bathytype. ISME J
638 doi: 10.1038/ismej.2013.243
639
640 Tripp HJ. (2013). The unique metabolism of SAR11 aquatic bacteria. J Microbiol
641 51:147–153.
642
643 Tripp HJ, Kitner JB, Schwalbach MS, Dacey JWH, Wilhelm LJ, Giovannoni SJ. (2008).
644 SAR11 marine bacteria require exogenous reduced sulphur for growth. Nature 452:741–
645 744.

646 Waldbauer JR, Rodrigue S, Coleman ML, Chisholm SW. (2012). Transcriptome and
647 proteome dynamics of a light-dark synchronized bacterial cell cycle. PLoS ONE
648 7:e43432.

649 Webb E, Claas K, Downs D. (1998). *thiBPQ* encodes an ABC transporter required for
650 transport of thiamine and thiamine pyrophosphate in *Salmonella typhimurium*. J Biol
651 Chem 273:8946–8950.

652 Wightman R, Meacock PA. (2003). The *THI5* gene family of *Saccharomyces cerevisiae*:
653 distribution of homologues among the hemiascomycetes and functional redundancy in the
654 aerobic biosynthesis of thiamin from pyridoxine. Microbiology 149:1447–1460.

655 Williams TJT, Long EE, Evans FF, Demaere MZM, Lauro FMF, Raftery MJM, et al.
656 (2012). A metaproteomic assessment of winter and summer bacterioplankton from
657 Antarctic peninsula coastal surface waters. ISME J 6:1883–1900.

658 Winkler W, Nahvi A, Breaker RR. (2002). Thiamine derivatives bind messenger RNAs
659 directly to regulate bacterial gene expression. Nature 419:952–956.

660 Winkler WC, Breaker RR. (2005). Regulation of bacterial gene expression by
661 riboswitches. Annu Rev Microbiol 59:487–517.

662 Worden AZ, Lee J-H, Mock T, Rouze P, Simmons MP, Aerts AL, et al. (2009). Green
663 evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes
664 *Micromonas*. Science (New York, NY) 324:268–272.

665 Wrenger C, Eschbach M-L, Müller IB, Laun NP, Begley TP, Walter RD. (2006). Vitamin
666 B_1 *de novo* synthesis in the human malaria parasite *Plasmodium falciparum* depends on
667 external provision of 4-amino-5-hydroxymethyl-2-methylpyrimidine. Biol Chem 387:41–
668 51.

669 Wu H, Ito K, Shimoi H. (2005). Identification and characterization of a novel biotin
670 biosynthesis gene in *Saccharomyces cerevisiae*. Appl Environ Microbiol 71:6845–6855.

671 Zhao L, Ma X-D, Chen F-E. (2012). Development of two scalable syntheses of 4-Amino-
672 5-aminomethyl-2-methylpyrimidine: key intermediate for vitamin B_1. Org. Process Res.
673 Dev. 16:57–60.

674

675 **Figure Legends:**

676 Figure 1: Simplified illustration of thiamin metabolism in *Ca*. P. ubique. Black colored

677 lines and enzyme abbreviations represent reactions and enzymes encoded by the *Ca*. P.

678 ubique genome. Red colored lines and enzyme abbreviations represent reactions and

679 enzymes that are absent from the *Ca*. P. ubique genome. Abbreviations: AIR -

680 aminoimidazole ribotide; his – histidine; HMP(-P) - 4-amino-5-hydroxymethyl-2-

681 methylpyrimidine (-phosphate); HMP-PP - 4-amino-5-hydroxymethyl-2-

682 methylpyrimidine diphosphate; THZ-P - 4-methyl-5-(2-phosphoethyl)-thiazole; AmMP -

683 4-amino-5-aminomethyl-2-methylpyrimidine; ThP – thiamin monophosphate; ThPP –

684   thiamin diphosphate; dDXP - 1-deoxy-D-xylulose 5-phosphate; gly – glycine; cys -

685   cysteine.

686   Figure 2: Gene phylogeny, synteny and conservation of riboswitch structure for the

687   *Pelagibacterales* ThiV-family sodium:solute symporter. *Pelagibacterales* genome

688   elements are highlighted in red. A) Maximum Likelihood phylogenetic tree showing a

689   subset of amino acid sequences extracted from a complete tree (Supplementary Fig. S4).

690   B) For the same taxa shown in "A", the chromosomal co-localization of *thiV* genes with

691   putative ThPP-binding riboswitches (red stem-loop structure) and genes encoding

692   thiamin salvage enzymes (*thiDME* or *tenA*). Dashed line indicates no ThPP-riboswitch or

693   associated salvage genes were identified. C) Nucleotide sequences of predicted ThPP-

694   binding riboswitches depicted in (B). Dashed box encapsulates the riboswitch sequences

695   from nine *Pelagibacterales* genomes and their consensus sequence (illustrated at the top).

696   Sequences that are marked with (*) were predicted to contain ThPP-binding motifs using

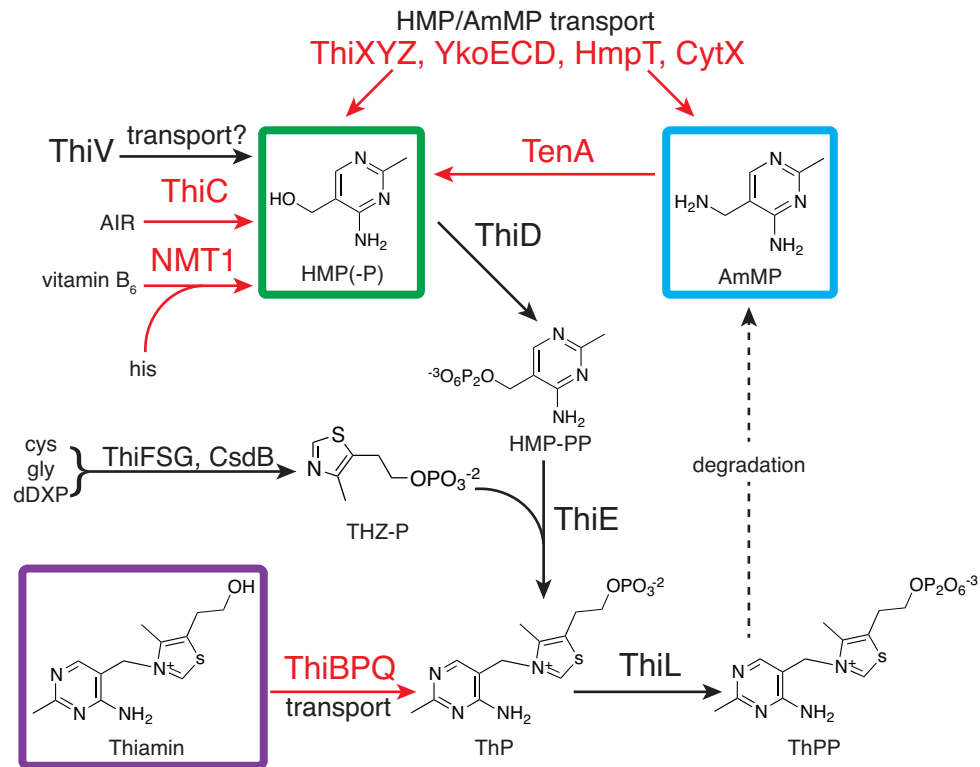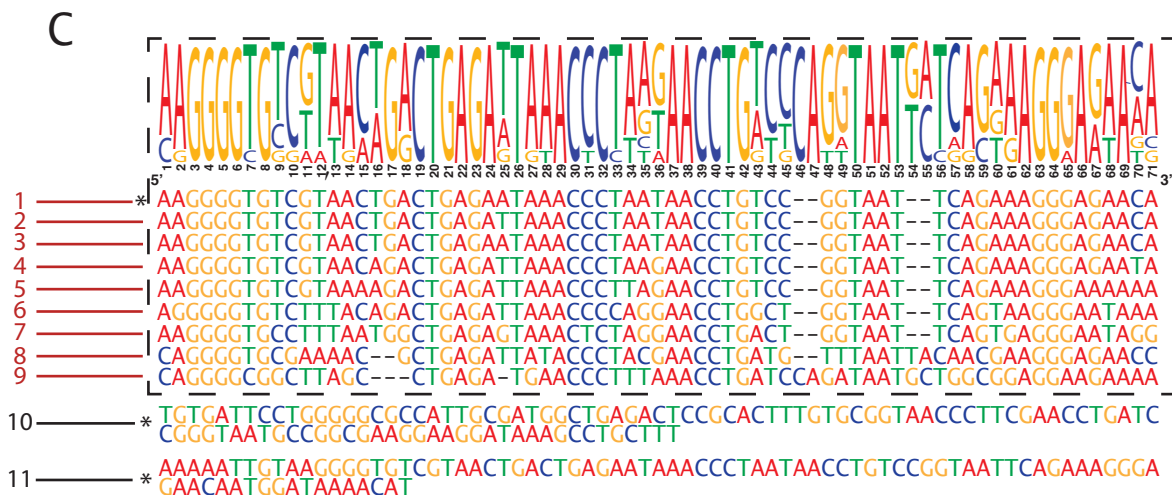697   the rfam (http://rfam.sanger.ac.uk) sequence search tool.

698

699   Figure 3: Maximum cell yields of *Ca.* P. ubique batch cultures in response to AmMP,

700   thiamin and HMP additions. Cells were grown in AMS1 amended with thiamin, HMP or

701   AmMP as indicated. Bar heights are the average densities of biological replicates ± s.d.

702   (n=3). The dashed line represents the calculated maximum density expected ($\sim 1.8 \times 10^7$

703   cells ml^{-1}) from the "background" level of HMP (see text for details). We attribute the

704   growth with 1 μM thiamin or AmMP to "contaminating" HMP (see text for details).

705

706   Figure 4: Depth distribution of dissolved 4-amino-5-hydroxymethyl-2-methylpyrimidine

707   (HMP) and thiamin in the Sargasso Sea. Times of collection are presented in local time.

708   HMP values are the average of technical replicate analyses for each sample. There was

709   no technical replication for the thiamin measurements due to insufficient sample. HMP

710   was not detected in the 250 m sample collected at 08:00. Thiamin was not detected in

711   samples collected from 200 m at 20:00 or at 250 m and 300 m at either time. LOQ: limit

712   of quantification. M-L: Mixed Layer. DCM: deep chlorophyll maximum.

713

A

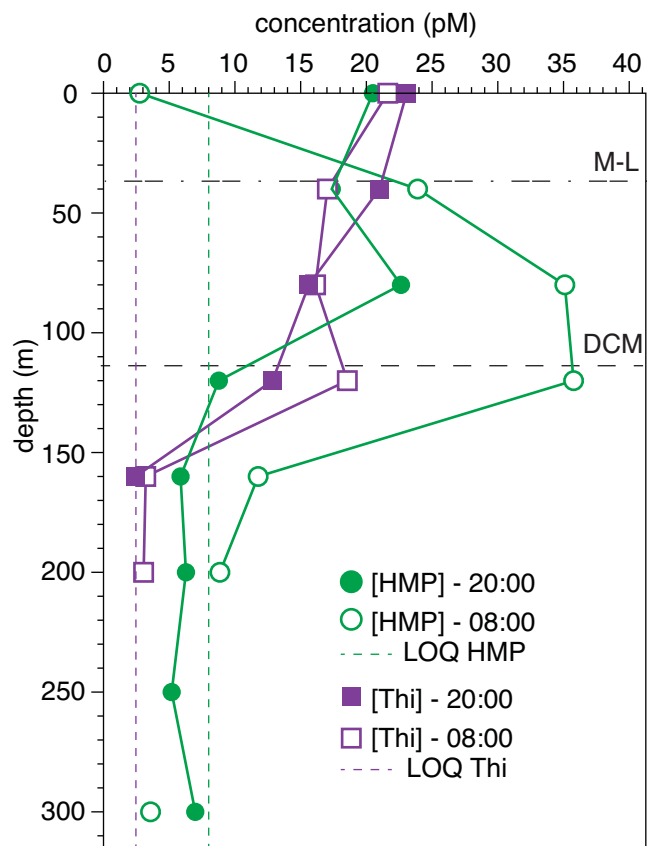| 91 | SAR11_0811 Candidatus Pelagibacter ubique SAR11 HTCC1062 | 1 |
99	PU1002_02346 Candidatus Pelagibacter ubique SAR11 HTCC1002	2
86	HTCC9565_00013660 Candidatus Pelagibacter ubique HTCC9565	3
98	HTCC7211_00007120 Candidatus Pelagibacter sp HTCC7211	4
100	HIMB5b_00008600 SAR11 sp HIMB5	5
100	IBGYDRAFT_00002740 alpha proteobacterium AAA288-G21	6
90	A240E13DRAFT_00011540 alpha proteobacterium AAA240-E13	7
98	HIMB114_00011070 SAR11 sp HIMB114	8
86	IMCC9063_000100005200 Candidatus Pelagibacter sp IMCC9063	9
35 100	Mfla_1786 Methylobacillus flagellatus KT	10
61	MELB17_01450 Marinobacter sp ELB17	11
	MDG893_05384 Marinobacter algicola DG893	12
38	Clos_2692 Alkaliphilus oremlandii OhILAs	13
46	YP_0040442561 Halogeometricum borinquense DSM 11551	14
35	Huta_0686 Halorhabdus utahensis AX-2 DSM 12940	15
98	HQ2370A Haloquadratum walsbyi HBSQ001 DSM 16790	16
100	YP_0035334541 Haloferax volcanii DS2	17
	rrnAC1769 Haloarcula marismortui ATCC 43049	18

Table 1: Comparative genomics of thiamin biosynthesis in the *Pelagibacterales*

Strain	thiC	thiD	thiE_0583 [†]	thiE_0360 [†]	thiF	thiS	thiG
Ca. P. ubique	absent	637671479	637671458	637671224	637671266	637671603	637671604
HTCC1002	absent	639129819	639129840	639130075	639130033	639129702	639129701
HTCC7211	absent	2503353714	2503353735	2503352435	2503352394	2503352877	2503352878
HTCC9565	absent	2503364149	2503364170	2503364413	2503364372	2503364883	2503364884
HIMB5	absent	2504109247	2504109269	2504109551	2504109508	2504108506	2504108507
HIMB114	absent	2503356000	2503356022	2503356319	2503356274	2503355884	2503355883
IMCC9063	absent	2505688345	2505688367	2505687345	2505687250	2505687878	2505687879
HIMB59	absent	absent	absent	absent	2504110146	absent	absent

Gene numbers are IMG/ER Gene ID's (https://img.jgi.doe.gov/er)

[†]There are two copies of *thiE* in *Ca.* P. ubique: *SAR11_0583* and *SAR11_0360.*

[1]*csdB* is predicted to encode the cysteine desulfurase activity necessary for thiazole biosynthesis (see supplementary methods)

csdB [1]	thiL	thiB	tenA
637671616	637671913	absent	absent
639130662, 639129689	639130810	absent	absent
2503352890	2503353193	absent	absent
2503364896	2503365124	absent	absent
2504108519, 2504109389	2504108893	absent	absent
2503355872	absent	absent	absent
2505688259	absent	absent	2505687352
2504110964	2504110802	2504111022	absent

Table 2: HMP concentrations in uninoculated and partially spent media

Organism	HMP (pM)	
	uninoculated	partially spent
Synechococcus sp. WH8102	N/D	2,909.6
Dunaliella tertiolecta	11.6	1,584.3
Prochlorococcus sp. MED4	N/D	32.8
OM43 isolate HTCC2181	12.9	33.0
Ca. P. ubique str. HTCC1062	11.6	N/D
Pelagibacterales sp. str. HTCC7211	11.6	N/D

N/D: not detected. Limit of detection = 2.4 pM