# NRC Publications Archive
# Archives des publications du CNRC

**Small-molecule inhibitors of the pseudaminic acid biosynthetic pathway : targeting motility as a key bacterial virulence factor**

Menard, Robert; Schoenhofen, Ian C.; Tao, Limei; Aubry, Annie; Bouchard, Patrice; Reid, Christopher W.; Lachance, Paul; Twine, Susan M.; Fulton, Kelly M.; Cui, Qizhi; Hogues, Hervé; Purisima, Enrico O.; Sulea, Traian; Logan, Susan M.

**NRC Publications Record / Notice d'Archives des publications de CNRC:**
https://nrc-publications.canada.ca/eng/view/object/?id=5bb3000c-2678-482b-8e38-2de222f03bc7
https://publications-cnrc.canada.ca/fra/voir/objet/?id=5bb3000c-2678-482b-8e38-2de222f03bc7

National Research Council Canada    Conseil national de recherches Canada

Canada

1   **Small-molecule inhibitors of the pseudaminic acid biosynthetic**

2   **pathway: targeting motility as a key bacterial virulence factor.**

3

4   Robert Ménard[1,†], Ian C. Schoenhofen[2], Limei Tao[1], Annie Aubry[2], Patrice Bouchard[1],

5   Christopher W. Reid[2,#], Paule Lachance[1], Susan M. Twine[2], Kelly M. Fulton[2], Qizhi Cui[1], Hervé

6   Hogues[1], Enrico O. Purisima[1], Traian Sulea[1,*] and Susan M. Logan[2,*]

7

8   [1] *Biologics Program, Human Health Therapeutics, National Research Council Canada, 6100*

9   *Royalmount Avenue, Montreal, QC H4P 2R2, Canada*

10  [2] *Vaccine Program, Human Health Therapeutics, National Research Council Canada, 100 Sussex*

11  *Drive, Ottawa, ON K1A 0R6, Canada*

12

13

14  [†] This work is dedicated to Robert Ménard who passed away on August 19, 2013.

15  [#] C.W. Reid: Current address: *Department of Science and Technology, Bryant University,*

16  *Smithfield, RI, USA*

17

18  [*] Corresponding authors: Traian.Sulea@nrc-cnrc.gc.ca, Susan.Logan@nrc-cnrc.gc.ca

19

20    **Abstract**

21    *Helicobacter pylori* is motile by means of polar flagella and this motility has been shown to play

22    a critical role in pathogenicity. The major structural flagellin proteins have been shown to be

23    glycosylated with the nonulosonate sugar, pseudaminic acid (Pse). This glycan is unique to

24    microorganisms and the process of flagellin glycosylation is required for *H. pylori* flagellar

25    assembly and consequent motility. As such, the Pse biosynthetic pathway offers considerable

26    potential as an anti-virulence drug target, especially since motility is required for *H. pylori*

27    colonization and persistence in the host. This study describes screening the five Pse

28    biosynthetic enzymes for small molecule inhibitors using both high throughput (HTS) and *in*

29    *silico* (VS) approaches. Using a 100,000 compound library, 1773 hits were identified by HTS that

30    exhibited a 40% threshold inhibition at 10 µM concentration. In addition, VS efforts using a 1.6

31    million compound library directed at two pathway enzymes identified 80 hits, 4 of which

32    exhibited reasonable inhibition at 10 µM concentration *in vitro*. Further secondary screening

33    was performed which identified 320 unique molecular structures or validated hits. Following

34    kinetic studies and SAR of selected inhibitors from our refined list of 320 compounds, we

35    demonstrated that three inhibitors with $IC_{50}$ values of approximately 14 µM, and which

36    belonged to a distinct chemical cluster, were able to penetrate the Gram negative cell

37    membrane and prevent flagella formation.

38

## Introduction

Infections caused by bacteria continue to represent major challenges to health care in both the hospital environment as well as the community setting. Microbial resistance to antibiotics has increased to the point where the current arsenal of antibacterial drugs are inadequate and on many occasions bacterial resistance to these drugs can lead to life threatening infection (1, 2). The development of new, effective antimicrobials is clearly needed and targeting bacterial virulence has gained considerable attention recently as an alternative approach to identify novel antibacterial therapeutics (3-6). Depriving pathogenic bacteria of their virulence functions could prevent the establishment of infection, and would allow the host immune system sufficient time to facilitate clearance of the organism. Virulence targeted drugs are organism-specific and are unlikely to be bactericidal; features which would have limited impact on host commensal flora and provide the additional benefit of reducing the risk of opportunistic infections. New anti-virulence drugs could also be used in combination with existing antibiotics to improve efficacy in current treatment strategies.

*Helicobacter pylori* is a significant gastrointestinal pathogen responsible for chronic active gastritis, peptic ulcers and related gastric cancers (7). The current established treatments for *H. pylori* infection are numerous and include triple or quadruple therapy both of which utilise two antibiotics (metronidazole, amoxicillin tetracycline or clarithromycin) in addition to either a proton pump inhibitor (PPI) (triple therapy) or PPI and bismuth (quadruple therapy). The efficacies of these treatment strategies have been severely hampered in recent years due to the rise in antibiotic resistance of *H. pylori* isolates and is now at the point where the current

60    rate of eradication has dropped below 70% in many countries (8). As such there is a clear need

61    to develop alternative therapeutic strategies for the management of *H. pylori* related disease.

62        It has been clearly established that motility is a critical virulence factor for *H. pylori*

63    infections (9-11). This motility, observed under conditions of elevated viscosity (as is found in

64    the gastric lumen), is due to a unipolar bundle of sheathed flagella of which the structural

65    filaments are composed of two flagellin protein species, FlaA and FlaB. To infect the stomach,

66    the bacteria must first transit the mucus layer from the gastric lumen, with the final destination

67    being the epithelial surface, which is the site of infection. The directed motility of cells is

68    essential to this process as *H. pylori* colonizes the interface of separate mucosa (antral and

69    fundic) in the stomach, and the organism must continually seek out this niche as conditions

70    vary between fasting and feeding (12). Importantly, in addition to being required for initial

71    colonization of the stomach it has also been shown that motility is also required for robust,

72    long-term, persistent infections (11, 13).

73        In previous studies we demonstrated that the structural flagellin proteins from *H. pylori*

74    and *Campylobacter jejuni* are glycosylated with the novel "sialic acid-like" nonulosonate sugar,

75    pseudaminic acid (Pse). Targeted gene disruption of the Pse biosynthetic pathway genes

76    showed that this glycosylation is essential for flagellar filament assembly and consequent

77    motility (9, 14). The *H. pylori* Pse pathway isogenic mutant strains were unable to colonise the

78    stomach in a mouse model of infection and *C. jejuni* Pse isogenic mutant strains were

79    attenuated in the ferret diarrhoeal disease model (9, 15). Pseudaminic acid derivatives are also

80    found in a number of other bacterial species as components of cell surface glycans such as LPS

81    O antigens, capsular polysaccharides and pili, and in many examples these surface glycans are

82    essential for bacterial virulence (16-19).

83         As a key virulence factor, as well as being a unique product made by microorganisms,

84    the Pse biosynthetic pathway offers potential as a novel therapeutic target. The Pse

85    biosynthetic pathways from *H. pylori* and *C. jejuni* have been elucidated and the function of

86    each of the pathway's five biosynthetic enzymes has been determined following recombinant

87    production and purification of each biosynthetic enzyme (20-23). In addition it has been

88    demonstrated that the all 5 Pse pathway enzymes can be combined in a single one-pot reaction

89    for the synthesis of Pse using UDP-GlcNAc as an initial substrate (20). Structural studies of three

90    of the biosynthetic enzymes have also been completed (24-26).

91         The observation that glycosylation of the flagellin structural proteins is required for

92    flagellar assembly and subsequent motility, in addition to the extensive body of work

93    characterizing  the novel bacterial pseudaminic acid biosynthetic pathway, has set the ground

94    work for small-molecule inhibitor screening of this key *H. pylori/C. jejuni* virulence factor. In this

95    study we have identified small-molecule hits from high-throughput screening (HTS) and virtual

96    screening (VS) campaigns. We disclose a subset of chemically-related small-molecule lead

97    compounds that inhibit *H. pylori* and *C. jejuni* Pse biosynthetic pathway enzymes and prevent

98    flagella formation in cell-based assays with *C. jejuni*.

99    **Materials and methods**

100   **HTS library**

101    The screened HTS library consists of 96,116 pure chemical compounds acquired from the

102    commercial HTS libraries of ChemDiv Inc. (San Diego, CA) and Tripos Inc. (St. Louis, CA), which

103    were selected based on their chemical diversity and drug-like properties.

**Pse-pathway enzymes**

105    Recombinant production and purification of *H. pylori* and *C. jejuni* Pse biosynthetic enzymes

106    were as previously described (20, 22) and recombinant plasmids are listed in **Table S5**. Purified

107    proteins were dialyzed against 20 mM HEPES pH 7.2, 50 mM NaCl prior to assays.

**Phosphate based primary screening assay**

109    For HTS in 384-well plates, the reaction volume was 10 µL per well. A substrate master mix of

110    7.26 µL (containing 0.5 mM UDP-GlcNAc, 0.5 mM PLP, 7 mM L-Glu, 0.5 mM acetyl-Co A and 0.5

111    mM PEP) was combined with 2.74 µL of enzyme mix at concentrations indicated in **Table S1**.

112    Each well contained 10-20 µM of library compound and reaction was incubated for 60 min at

113    37°C, 95% RH. Pi detection was done by adding 40 ALS reagent (Pi Colorlok assay kit from

114    Innova Biosciences), incubated 5 min, then reaction stopped by addition of 4 µL of stabilizer.

115    Final $OD_{595}$ reading was done on PerkinElmer Envision plate reader.

**RapidFire HTMS secondary screening**

117    Reactions were carried out in 384-well plates with final reaction volume of 10 µL per well. The

118    reaction conditions for each assay are provided as supplemental data (**Table S3**). Reactions

119    were stopped by adding 30 µL of stop solution (20 : 80 : 0.2 = acetonitrile : water : formic acid),

120    then 25 µL of stopped reaction mixture was transferred from 384-well assay plates into 4 X 96-

121    well plates containing 175 μL of stop solution. HTMS plates were sealed, centrifuged and kept

122    at 4°C until processed by the RapidFire 200 (Agilent) coupled to Agilent QQQ6410B mass

123    spectrometer. Sample (10 μL) was delivered directly from 96-well plates to the Graphitic

124    Carbon cartridge to replace the non-volatile buffer with $H_2O$ in a 2.5-s wash cycle at a flow rate

125    of 1.5 mL/min. The substrate and product were co-eluted to the mass spectrometer in 4 to 8 s

126    using 70 : 30 : 0.02 = acetonitrile: water: ammonium hydroxide at a flow rate of 0.9 mL/min.

127    The chromatograph system produced baseline-resolved peaks. The eluted sample passed

128    directly into the MS ion source under negative ion mode for quantitative analysis. The other

129    parameters for each analyte are given as supplemental data (**Table S2**). The concentration of

130    each inhibitor in HTMS screening reactions was 30 μM.

131    **Chemical clustering**

132    Chemical clustering of compounds was carried out in Sybyl 8.1.1 (Tripos Inc., St. Louis, MO) by

133    hierarchical clustering. 2D-FINGERPRINTS, which are binary variables for the presence (1) or

134    absence (0) of specific fragments, and ATOM_PAIR_FP, which are fingerprints describing the

135    minimum path lengths between atoms in molecules, were used as clustering descriptors of

136    molecular structure at 2D level. Hierarchical clustering was complete (**Fig. S2**), that is, taking the

137    inter-cluster distance to be the greatest separation between their elements, thus producing

138    dendrograms with multiple, compact root clusters and minimizing the generation of singletons.

139    **PseB kinetic assay**

140    Reactions were carried out in 384-well plate in 20 mM HEPES, 20 mM NaCl pH 7.2, 1% DMSO,

141    with final reaction volume of 50 μL per well. PseB at a concentration of 2μM was preincubated

142     with library compounds at 100, 66.7, 44.4, 29.6, 19.8, 13.2, 8.8, 5.9, 3.9, 2.6, 1.7 or 1.2 $\mu$M for

143     10 min at 25$^{\circ}$C. The reaction was started by addition of 200 $\mu$M UDP-GlcNAc. After incubation

144     for 30 min incubation, the reactions were stopped by transferring 4 $\mu$l of each reaction mixture

145     into 195 $\mu$L stop solution (20 : 80 : 0.2 = acetonitrile : water : formic acid) in 96-well plates.

146     Then samples were analyzed by RapidFire coupled to mass spectrometer.

147     **Cell-based assays**

148     *C. jejuni* 81-176 and *C. jejuni* 81-176 *PseB::Cm* were grown overnight on Mueller Hinton agar

149     under microaerophilic conditions at 37$^{\circ}$C. Cells were harvested from an agar plate into Mueller

150     Hinton broth and this suspension was used to inoculate a well with 1 mL of Mueller Hinton

151     broth containing 0.01% DMSO in a Falcon Multiwell$^{TM}$ 6 well plate to $OD_{600}$ of 0.1. Inhibitors (in

152     DMSO) were added to wells to final concentration as specified. The multiwell plates were

153     incubated with shaking (200 rpm) at 37$^{\circ}$C under microaerophilic conditions for 7 h, the $OD_{600}$

154     measured and cells harvested by centrifugation. Cells were fixed overnight in 3% formalin in

155     PBS, washed in PBS and $OD_{600}$ adjusted to 0.08 for coating on 96 well plates for ELISA. *C. jejuni*

156     strain 81-176 and the 81-176 *PseB ::Cm* grown in multiwell plates in identical fashion to

157     inhibitor test samples were used as positive and negative controls for ELISA assay. Assays were

158     completed on two independent occasions.

159     **ELISA assays**

160     Nunc MaxisorpPlates were coated with 100 uL of formalin-fixed cells overnight at 37$^{\circ}$C. Plates

161     were blocked (1% BSA in PBS) and then washed 3x with PBS/Tween 0.05% (PBS-T). A His-tagged

162     sdAb specific for 81-176 flagellin (a gift from M. Arbabi, NRC) was then added (1:1000 in PBS-

163    BSA, 0.85 mg/ml) and plates incubated for 2 h at RT. Plates were washed 3x (PBS-T) and then

164    incubated with rabbit anti-His HRP conjugated antibody (1:10,000 in PBS-BSA, 1 mg/ml) for 1h

165    at RT. Following washing, the antibody was detected with TMB for 10 min and reaction stopped

166    with 1M $H_3PO_4$. Samples were analyzed in triplicate and absorbance was measured at 450 nm.

167    **Inhibitor docking and virtual screening**

168    The 1.85-Å resolution crystal structure of PseG in complex with UDP (PDB code 3HBN) and the

169    1.9-Å resolution crystal structure of PseB in complex with UDP-GlcNAc (PDB code 2GN4) were

170    used for virtual screening and inhibitor docking after removal of water, substrate/product

171    molecules and co-solvent/buffer molecules, and addition of hydrogen atoms according to

172    standard ionization states. In the case of PseB, the $NADP^+$ cofactor and two subunits of the

173    hexamer were retained from the crystal structure as they are essential for shaping the PseB

174    substrate-binding cleft.

175        Virtual screening was carried out on a library of 1.6 million commercially-available drug-

176    like compounds from the ZINC database (27). We used a high-throughput VS docking-scoring

177    pipeline (28, 29). The exhaustive docking program Wilma (28) within the VS pipeline was used

178    with default increment parameters and the WilmaScore1 energy function (29). The ligand

179    conformations were generated by Omega (OpenEye, Inc., Santa Fe, NM) and controlled by

180    setting the internal energy cutoff to 20 kcal/mol and adjusting the pose clustering parameter to

181    produce at most 5000 conformations. Wilma-generated poses were refined by constrained

182    energy minimization as described previously (28, 30) prior to binding affinity scoring with the

183    solvated interaction energy (SIE) function (31) using default parameters for the electrostatic

184 and non-polar contributions to interaction and desolvation energies (30, 32). Pose selection for

185 a given inhibitor against a given enzyme variant was based on the lowest SIE score. These SIE

186 scores were then used for binding affinity ranking among various inhibitor-enzyme complexes.

187 Top-scored 200 complexes in the case of PseG and 357 in the case of PseB were subjected to

188 SIE averaging on molecular dynamics trajectories using SIETRAJ as described previously (30, 32,

189 33) to obtain the final binding affinity scores.

## Results

**Development of an optimized HTS assay for Pse biosynthetic enzymes**

192 The Pse biosynthetic pathway is outlined in **Figure 1**. Previous work had demonstrated that the

193 synthesis of Pse could be accomplished by combining all five enzymes of the pathway in a single

194 reaction with UDP-GlcNAc and necessary cofactors. Since the last enzymatic step results in the

195 release of inorganic phosphate, the entire pathway can be screened for inhibitors

196 simultaneously using a phosphate detection based assay. As such we first developed a 384-well

197 plate assay using the 5 enzymes of the Pse pathway and measuring the release of free

198 phosphate with the PiColorLock ALS reagents (Innova Biosciences) according to the

199 manufacturer's instructions. The assay conditions (relative enzyme concentrations,

200 stoichiometry, reagents concentrations, incubation times, etc.) were optimized to allow

201 identification of hits that target any one of the five Pse enzymes (see Materials and Methods).

202 The final concentrations of each *H. pylori* enzyme are provided as supplemental data (**Table S1**).

203 In a preliminary screening of over 20,000 compounds these conditions led to a Z'-factor of 0.88

204 indicating that the assay was robust and suitable for HTS (supplemental data, **Fig. S1**).

205        Using this assay, we have completed in duplicate an HTS of a library comprising

206   approximately 100,000 small-molecule compounds in order to identify inhibitors for the 5

207   enzymes of the pathway (**Fig. 2**). This screening was carried out at the compound concentration

208   of 10 µM. The quality of the screen allowed us to establish 40% inhibition as a threshold for the

209   identification of hits. This led to a hit-list of 1,773 unique compounds from two independent

210   primary screens (within the red square in **Fig. 2**). Chemical similarity based clustering of these

211   hits revealed that a certain fraction of them comprises derivatives of a small number of

212   chemical scaffolds. For example, even at a deep level of clustering corresponding to 500

213   clusters, there were 28 clusters each containing at least 10 close analogs, with 7 clusters

214   represented by at least 20 congeners (supplemental data, **Fig. S2**).

215   **Secondary screening hit validation**

216   It was important to confirm the primary hits obtained with a secondary assay using a different

217   technique. To this end, we next used label-free high-throughput mass spectroscopy (HTMS)

218   screening with RapidFire technology to both confirm the primary hits and to identify which of

219   the 5 enzymes was the respective target of each of those primary hits. The RapidFire HTMS

220   incorporates automated sample handling, in-line solid-phase extraction (SPE) cartridge system

221   for sample clean-up and analyte concentration, and an injection system coupled to a triple-

222   quadrupole mass spectrometer. The typical throughput of the instrument is approximately 500

223   samples per hour. The secondary HTMS screens were carried out against each of the four

224   enzymes from the Pse pathway, separately or in combination.

225         RapidFire HTMS methods were established for the simultaneous detection of multiple

226     reaction monitoring (MRM) signals of substrate and product for individual enzyme, giving no

227     signal crosstalk between the channels, thus allowing the percentage conversion to be

228     calculated. Since all the substrates and products in the Pse pathway are hydrophilic (**Fig. 1**),

229     solid-phase cartridge of HILIC (type Z) and Graphitic Carbon (type D) were evaluated, and

230     Graphitic Carbon cartridge gave the better performance and was chosen for the analysis. The

231     sampling and running parameters were optimized so as to give the best throughput without

232     compromising sensitivity or reproducibility of detection. Mass spectrometric parameters for

233     each compound and the linear ranges of all substrates and products are provided as

234     supplemental data (**Table S2**).

235         Enzymatic characterization of recombinant purified enzyme PseC, PseH or PseG was

236     performed to develop optimal assay conditions for inhibitors. For enzyme PseB or PseI, the

237     conditions developed in HTS assay were used. The reaction conditions are given as

238     supplemental data (**Table S3**).

239         The HTMS assay was used for screening of 1,773 unique hits identified from the primary

240     screening. Single-concentration screening was performed at 30 µM compound concentration.

241     The compounds were screened in total of 155 96-well plates for each of the 5 enzymes. The Z'

242     scores for individual plates were calculated using the formula (34):

243     
$$Z' = 1 - \left[ \frac{3 \times (SD_{PosCtrl} + SD_{NegCtrl})}{Avg_{PosCtrl} - Avg_{NegCtrl}} \right]$$

244   where Avg and SD are the averages and standard deviations, respectively, of the inhibition

245   signals produced by the positive and negative controls (PosCtrl and NegCtrl, respectively) on

246   each plate. Thus, Z' is reflective of both the assay signal dynamic range and the data variation

247   associated with the signal measurements and is used for comparison and evaluation of the

248   quality of assays.  A high throughput assay with $1 > Z' > 0.5$ score is considered an excellent

249   assay (35). Therefore, results from the plates with $Z' > 0.5$ were analyzed. Hit selection for each

250   enzyme was based on the Z' scores obtained for that particular enzyme. Therefore, hit selection

251   was not necessarily based on the level of inhibition by a particular compound but by how many

252   standard deviations (SD) that inhibition level deviated from the mean inhibition across all

253   tested hits (**Fig. 3**). The following thresholds for hit selection were established: more than 1 SD

254   for PseB and coupled PseB/PseC assays, more than 2 SD for PseG, PseH and PseI, and more than

255   3 SD for PseC. This led to 169 hits for PseB, 25 hits for PseC, 89 hits from the coupled PseB/PseC

256   assay, 118 hits for PseG, 89 hits for PseH and 100 hits for PseI. Because some of these hits were

257   found to inhibit more than one enzyme, the validated set of Pse-pathway inhibitors thus

258   comprises a total of 320 compounds that inhibit at least one of the five enzymes in the

259   pathway.

260       We repeated chemical clustering for these 320 unique molecular structures and

261   obtained 146 singletons and 174 compounds clustered into 22 clusters with at least 2

262   compounds, with the most populated cluster reaching 40 compounds (**Fig. 4**). The fraction of

263   compounds that were found to inhibit more than one enzyme in the pathway varies with the

264   chemical nature of the scaffolds representing the clusters. Overall, 53.4% of clustered

265   compounds inhibit more than one enzyme, which is similar to the fraction of 52.7% for multi-

266    enzyme inhibitors found among the remaining unclustered compounds (singletons).

267    **Structure based virtual screening**

268    In parallel with the HTS screening of an actual compound library, we took advantage of some of

269    the structural information available for enzymes in the Pse biosynthetic pathway. We carried

270    out VS of a virtual library of 1.6-million compounds against the crystal structures of PseB and

271    PseG, using a proprietary VS pipeline that demonstrated some of the better performances in

272    blind tests (28, 29). We acquired 38 virtual hits from the PseG screening and 42 virtual hits from

273    the PseB screening and tested them in the 5-enzyme assay described earlier. Four of the PseB

274    virtual hits showed a reasonable inhibition of the pathway at 10 μM of inhibitor concentration

275    (supplemental data, **Table S4**). The PseB inhibition was further confirmed in the secondary

276    HTMS assay, with degrees of inhibition above the noise level established from the HTMS based

277    secondary screening experiments. Interestingly, three of these compounds were also found to

278    significantly inhibit PseG, possibly suggesting certain mimicry by these compounds of the

279    similar UDP-sugar substrates of PseB and PseG. The comparison of the docked poses of these

280    inhibitors with the actual binding mode of UDP-GlcNAc in the PseB active site appears to

281    support this hypothesis (**Fig. 5**, and supplemental data, **Fig. S3**).

282    **Kinetic studies and SAR of select PseB inhibitors**

283    We next focused our attention on inhibitors of PseB, the first enzyme in the pathway. Of the

284    169 compounds targeting PseB as confirmed by HTMS based secondary screening and the four

285    VS based PseB inhibitors, we selected the top 20 compounds for further kinetic evaluation.

286    Kinetic studies were completed for each of the 20 inhibitors with the PseB enzyme using the

287    RapidFire HTMS PseB assay (see above) where substrate and product were quantified. Of the

288    20 compounds tested, 5 showed good dose-dependent inhibition behaviours, with IC50 values

289    ranging from 12 µM to 72 µM (**Fig. 6**).

290            Following chemical clustering analysis of the top 320 confirmed hits against PseB (**Fig.**

291    **4**), we found that these five PseB inhibitors belonged to two distinct chemical clusters but

292    nevertheless shared a related core substructure (**Fig. 6**). Inhibitors CD09463 and CD23703

293    belong to cluster 6 and are both substituted *N*-phenyl-2,5-dimethyl-pyrrole differing only in the

294    position of the hydroxyl and carboxylate substituents of the phenyl ring. Inhibitors CD26389,

295    CD24868, CD36508 belong to cluster 5 and are almost double in size relative to the two cluster-

296    6 inhibitors. One can easily notice that these larger inhibitors bear the core substructure *N*-

297    phenyl-2-pyrrolidone which is structurally similar to the *N*-phenyl-2,5-dimethyl-pyrrole

298    structure of the smaller inhibitors. The substitution pattern on the phenyl ring of the core

299    structure is also similar, with CD23703 from cluster 6 matching CD24868 and CD26389 from

300    cluster 5, and with CD09463 from cluster 6 matching CD36508 from cluster 5. Hence, one can

301    expect that the core substructure affords most of the binding affinity to PseB, while the larger

302    substituents of the cluster-5 compounds (at position 3 and 5 of the unsaturated 2-pyrrolidone

303    ring) have a smaller albeit favourable contribution to binding affinity and hence enzyme

304    inhibition. Molecular docking of these inhibitors in the PseB substrate-binding cleft supports

305    this structure-activity relationship data (**Fig. 6**, and supplemental data, **Fig. S4**).

306    **Inhibition of *Campylobacter jejuni* PseB enzyme**

307    As *H. pylori* flagella produced at the cell surface are covered with a membranous sheath,

15

308      detection of the assembled flagellin filaments at the cell surface presents a challenge. In

309      contrast, the assembled flagella filaments on Campylobacter cells are not covered with this

310      membranous sheath, which permits the detection of assembled filaments at the cell surface

311      with flagellin-specific immunological reagents. To allow us to examine the top five PseB

312      compounds identified above in a cell based assay it was necessary to confirm that these

313      inhibitors selected were active for the PseB enzyme from the related organism *C. jejuni*. Four of

314      the compounds indeed inhibited *C. jejuni* PseB, albeit at levels slightly lower than inhibition of

315      *H. pylori* PseB (**Fig. 7**). However, one of the cluster-5 compounds (CD36508) appeared not to

316      inhibit *C. jejuni* PseB *in vitro* at the tested concentration, which indicates weaker inhibitory

317      potency at best.

318      **Cell based assay**

319      To demonstrate efficacy of Pse inhibitors which were identified by secondary, we next

320      developed a cell based assay (ELISA) to measure flagella production on bacterial cells when

321      grown in the presence of inhibitor. We focused these assays on compounds identified as PseB

322      inhibitors. As indicated above, the assay uses *C. jejuni* 81-176, which has been shown to also

323      glycosylate the flagellin structural proteins with pseudaminic acid derivatives (14). The flagellin

324      proteins are detected on the cell surface using a single-domain antibody (sdAb) specific for the

325      *C. jejuni* 81-176 flagellin protein (36). The three compounds of the same structural cluster 5

326      (**Fig. 6**) were found to inhibit flagellin production in a dose-dependent manner. Inhibition was

327      observed at compound concentrations of 100 µM or higher (**Fig. 8**).

328      **Discussion**

329    Due to significant drug resistance it is becoming increasingly difficult to eradicate *H. pylori* and

330    as such the need for new, effective therapeutic regimes to treat *H. pylori* infections is globally

331    recognised (37). Here we present the results of a combined HTS and *in silico* approach to

332    identify inhibitors of the pseudaminic acid biosynthetic pathway enzymes. Development of a

333    pathway screening assay offered considerable advantage over parallel screenings of individual

334    enzymes. The efficiency of screening multiple enzymes simultaneously reduced the total

335    number of assays required in the HTS and also permitted the use of a commercially available

336    substrate (UDP-GlcNAc) rather than the more unique sugar pathway intermediates which are

337    not available commercially and would require more costly enzymatic or chemical synthesis

338    strategies. The optimised HTS 5-enzyme assay was used to identify 1773 unique compounds

339    and secondary HTMS was used to confirm the primary hits as well as identify the respective

340    inhibitor enzyme target. From these analyses, a validated set of 320 compounds have been

341    identified as inhibitors of at least one Pse biosynthetic enzyme.

342          Analysis of these validated hits indicated the presence of several chemical clusters with

343    close congeners, which can be used to generate structure activity relationships (SAR) data

344    valuable for further optimization efforts. Interestingly, with two exceptions (clusters 14 and 20,

345    see **Fig. 4**), these clusters include analogs that inhibit more than one enzyme in the pathway.

346    Inhibiting the biosynthetic pathway simultaneously at several points can be beneficial not only

347    for increased efficacy and specificity, but also to circumvent eventual bacterial resistance. One

348    may speculate that the discovery of multi-enzyme inhibitors is a reflection of the substrate

349    similarities shared by the component enzymes in the pathway. It is also possible that multi-

350    enzyme inhibition is a direct consequence of the HTS strategy developed here, as it originates in

351     the 5 enzyme one-pot assay used during primary screening.

352         Kinetic analysis combined with molecular docking of five validated PseB inhibitors

353     provided valuable SAR data (**Fig. 6**, **Fig. S4**). The common core substructure of these analogs

354     (two smaller cluster-6 inhibitors and three larger cluster-5 inhibitors) interacts tightly and

355     complementarily into the deep pocket of the binding site that accommodates the

356     pyrophosphate and the sugar moieties of the sugar-UDP substrate. However, the large

357     substituent at position 3 of the unsaturated 2-pyrrolidone ring of the larger cluster-5 analogs is

358     solvent-exposed and makes minimal contact with the enzyme, and is not predicted to bind in

359     the enzyme pocket used for UDP binding. Interestingly, the size of the substituent at position 5

360     of the pyrrol(e/idine) ring is predicted to override the substitution pattern on the N-phenyl ring.

361     Consequently, the predicted binding modes of the inhibitors from the two clusters correspond

362     to flipped orientations of the N-phenyl ring and its substituents within the binding pocket. In

363     the case of cluster-6 inhibitors, the smaller 5-methyl substituent allows the pyrrole ring to

364     interact snuggly with the enzyme sub-pocket used for binding the ribose ring of the UDP-sugar

365     substrate. In contrast, the cluster-5 inhibitors are predicted to have the larger 5-phenyl

366     substituent interacting in this sub-pocket. This would displace the pyrrolidone ring toward a

367     more central position in the binding cleft.

368         The subsequent finding that three out of these five analogs can inhibit flagellin

369     production in cellular assay is also informative from several viewpoints. First, one of the three

370     analogs that inhibited flagellin production in *C. jejuni* was found not to inhibit *in vitro* the PseB

371     enzyme at the tested concentration, which is indicative at best of weaker inhibition, although

372 all three inhibited similarly PseB from *H. pylori* (**Fig. 7**). However, we noted that all three

373 compounds are also *in vitro* inhibitors of PseG in *H. pylori* (52-70% inhibition at the

374 concentration in HTMS assay). Although we have not tested their activity against PseG from *C.*

375 *jenuni*, one can speculate that the multi-enzyme inhibitory characteristic of these and other

376 compounds discovered in this study may provide a mechanism for achieving robust efficacy in

377 the biological environment. Secondly, we found that the two smaller cluster-6 analogs do not

378 possess cellular activity despite showing inhibition of both PseB and PseG enzymes (in both

379 *H.pylori* and *C. jeuni*). Given that these analogs are negatively charged with a carboxylate

380 moiety, their membrane permeability is expected to be quite poor. However, the derivatization

381 of the larger analogs from cluster-5 leads to a significant increase in their hydrophobicity

382 (calculated octanol-water partition coefficient of 3.5-6.5) relative to the smaller inactive

383 analogs from cluster-6 (only 2.3). Hence, increased hydrophobicity and membrane permeability

384 of these pathway inhibitors is another key property required for achieving cellular efficacy.

385      The identification of three inhibitors which show activity in bacterial cell-based assays

386 provides a robust starting point for further hit-to-lead evaluation. Further characterization of

387 the binding modes of these inhibitors will provide a guide for structure-based design of more

388 potent and selective inhibitors towards the Pse biosynthetic enzymes which can then be tested

389 as novel anti-virulence therapeutics targeting *H. pylori*. Furthermore, the strategies utilized

390 here may be useful for identifying inhibitors of other related nonulosonate pathways, such as

391 legionaminic acid from *Legionella pneumophila*.

392
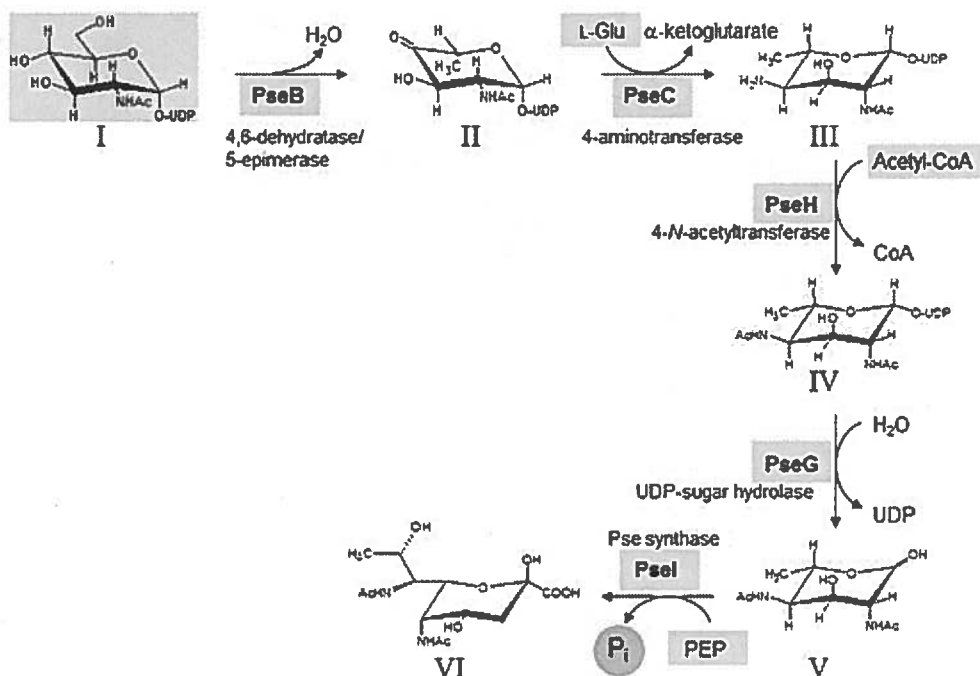
**Figure 1. The pseudaminic acid biosynthetic pathway in *H. pylori* and *C. jejuni*.** Biosynthetic intermediates are: (I) UDP-GlcNAc; (II) UDP-2-acetamido-2,6-dideoxy-β-L-*arabino*-hexos-4-ulose; (III) UDP-4-amino-4,6-dideoxy-β-L-AltNAc; (IV) UDP-2,4-diacetamido-2,4,6-trideoxy-β-L-altropyranose; (V) 2,4-diacetamido-2,4,6-trideoxy-L-altropyranose; (VI) pseudaminic acid.
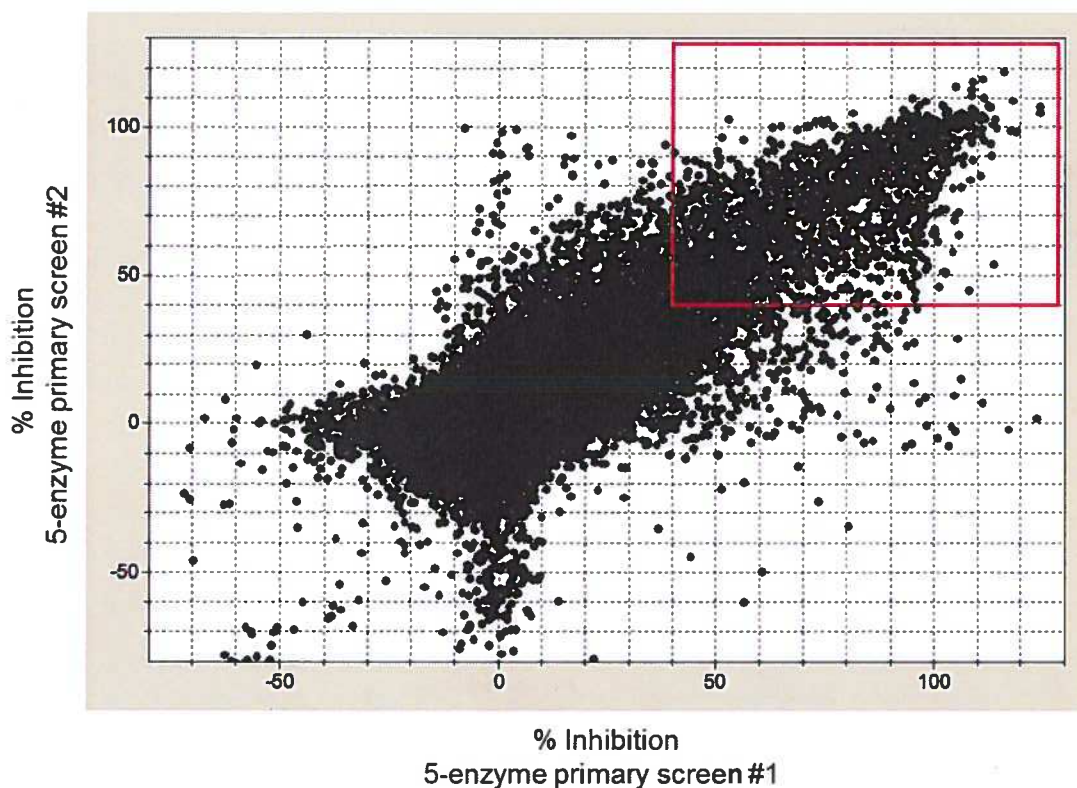
400

401

402

403

404

405

406

**Figure 2**. **HTS screening using 5-enzyme primary assay.** The HTS library of 96,116 compounds

was screened twice against a mixture of *H. pylori* Pse pathway enzymes with concentrations

and assay conditions as shown in supplemental Table S1. The area enclosed in the red rectangle

focuses on 1,847 hits (corresponding to 1,773 unique compounds) showing more than 40%

inhibition in both screens. Note that 3 standard deviations from mean inhibition corresponds to

inhibition thresholds of at least 41% and 42% in screens #1 and #2, respectively.
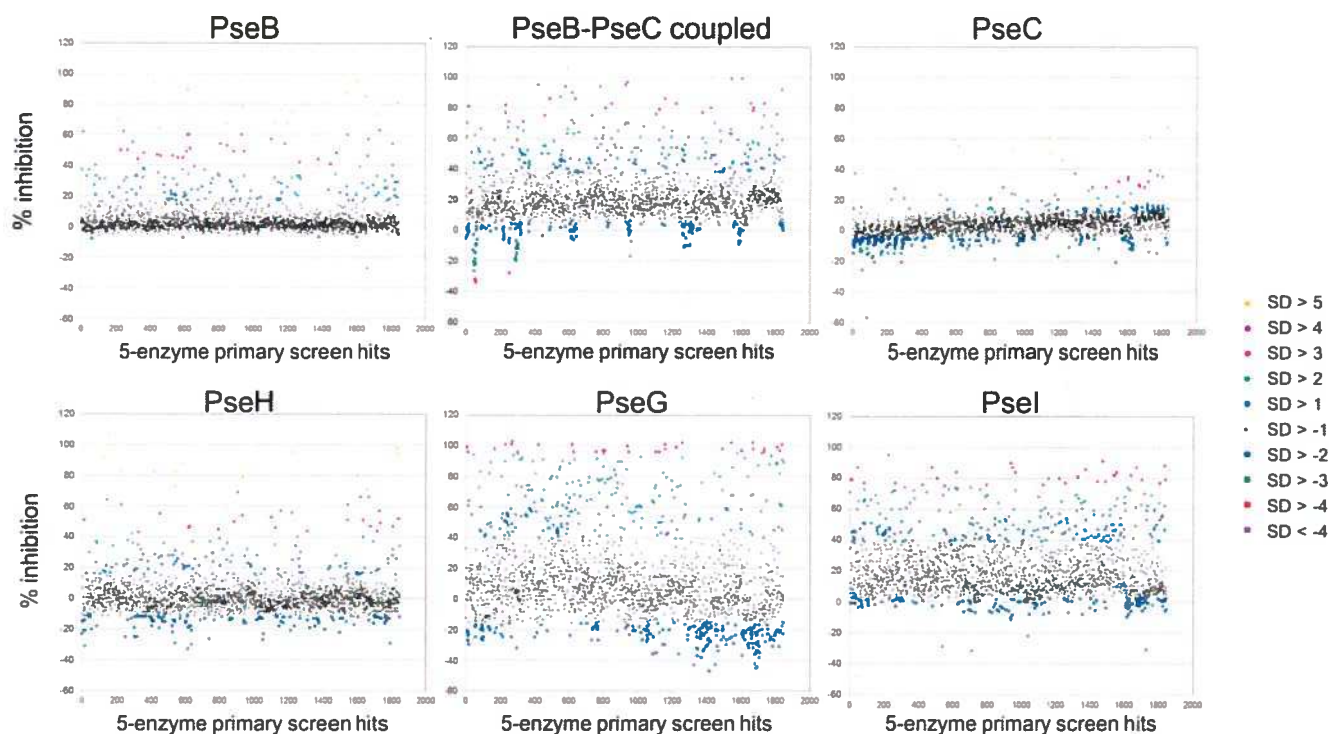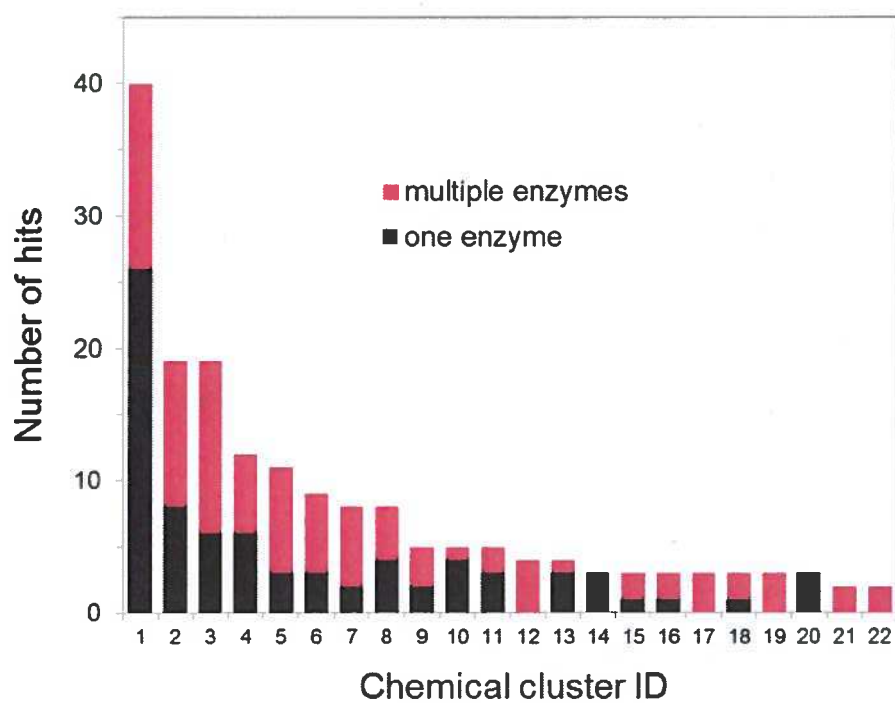
413

414

415

**Figure 3. HTMS-based enzyme-specific secondary screening.** The level of inhibition of each *H. pylori* Pse-pathway enzyme is plotted for the 1,847 hits (corresponding to 1,773 unique compounds) from the primary screens, and color coded according to the number of standard deviations (SD) from the mean for each screened enzyme (see legend).

420

421 **Figure 4. Clustering of validated hits from secondary screening.** The 320 compounds identified

422 after HTMS-based enzyme-specific screening were clustered by 2D-chemical similarity. With

423 two exceptions, these clusters contain at least a fraction of compounds that inhibit more than

424 one enzyme in the pathway.

425
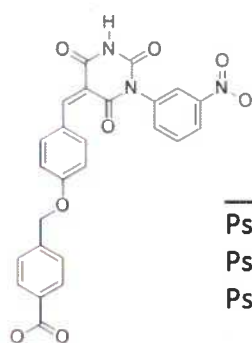
|  | % Inhibition |
| --- | --- |
| Pse-pathway | 50 |
| PseB | 32 |
| PseG | 55 |

ZINC06974903

426

427

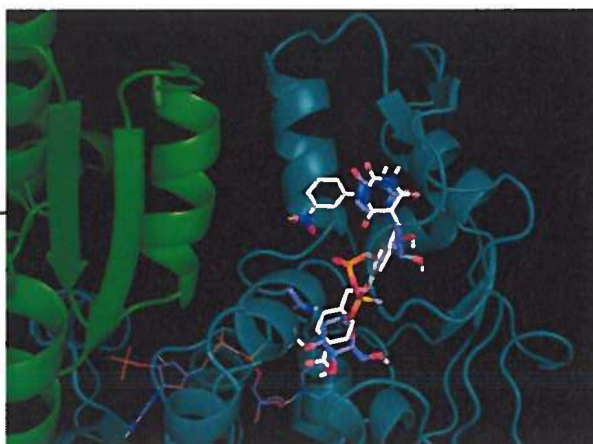428 **Figure 5. Example of confirmed virtual screening hit.** Chemical identification is given as the ID

429 in the ZINC database (http://zinc.docking.org). *In vitro* data is from 5-enzyme pathway assay

430 and HTMS enzyme-specific assays. The image on the right depicts a docked pose of the inhibitor

431 (sticks model with C atoms in white) overlaid with the actual binding mode of the UDP-GlcNAc

432 substrate (sticks model with C atoms in purple) in the PseB active site located at the interface of

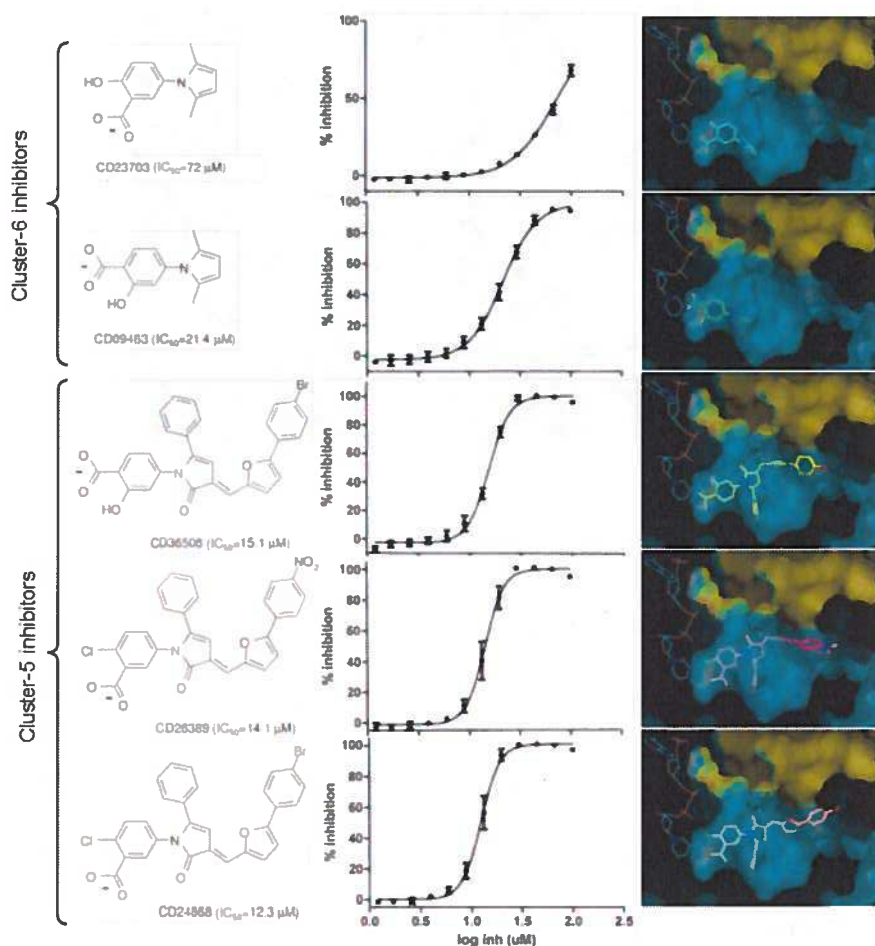433 two enzyme molecules (shown as ribbons of different colors, PDB entry 1GN4).

434

435

436

Figure 6. Structure-activity data for select PseB inhibitors. Chemicals structures and IC50

values for *H. pylori* PseB inhibition are shown in the left column. Corresponding dose

dependencies are presented in the middle column. Docked binding modes in the PseB

substrate-binding cleft are depicted in the right column, with the inhibitor in sticks model, the

cofactor in lines model, and the enzyme represented by molecular surface (differently coloured
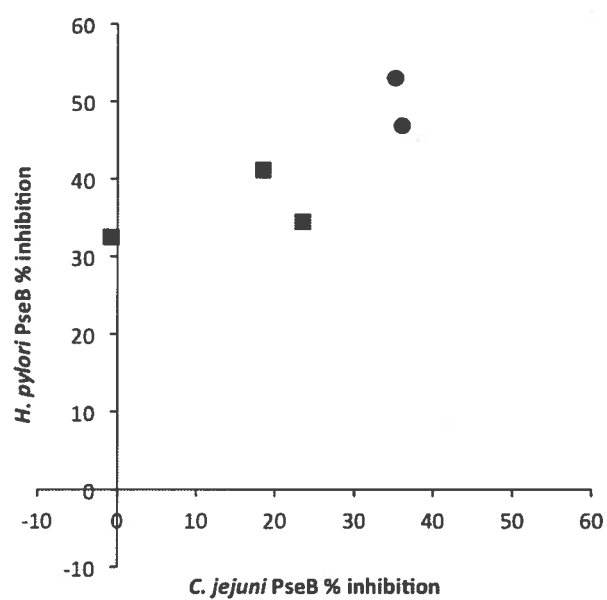
monomers).

443

444

445

446

447 **Figure 7. Correlation between inhibitions the PseB enzymes from _H. pylori_ and _C. jejuni_.** Data

448 shown are for the 5 inhibitors selected for kinetic analysis and molecular docking (Fig. 6).

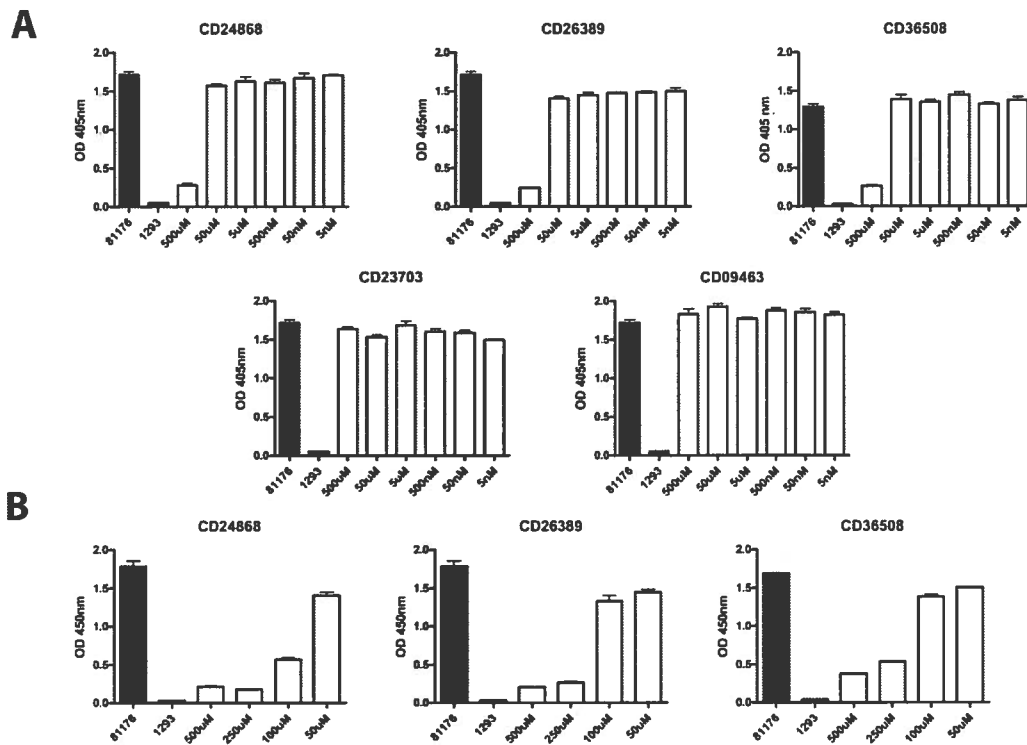449 Squares denote cluster-5 inhibitors and circles are for cluster-6 analogs.

450

**A**

CD24868

CD26389

CD36508

CD23703

CD09463

**B**

CD24868

CD26389

CD36508

451

452

453

454 **Figure 8**. **Inhibition of *C. jejuni* flagella production**. Whole cell ELISA assay using flagellin

455 specific antibody FlagV1. Black bars are *C. jejuni* 81-176 and *C jejuni* 81-176 *pseB::Cm* controls

456 grown in MH broth with DMSO and white bars are *C. jejuni* 81-176 grown in MH broth with

457 variable concentration of inhibitor as indicated. A. Inhibitor concentration 500 μM, 50 μM, 500

458 nM, 50 nM, 5 nM.  B. Inhibitor concentration 500 μM, 250 μM, 100 μM, 50 μM.

459

461

462    1.    **Boucher HW, Talbot GH, Bradley JS, Edwards JE, Gilbert D, Rice LB, Scheld M, Spellberg B,**
463         **Bartlett J.** 2009. Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society
464         of America. Clin Infect Dis **48**:1-12.

465    2.    2010. The 10 x '20 Initiative: pursuing a global commitment to develop 10 new antibacterial
466         drugs by 2020. Clin Infect Dis **50**:1081-1083.

467    3.    **Barczak AK, Hung DT.** 2009. Productive steps toward an antimicrobial targeting virulence. Curr
468         Opin Microbiol **12**:490-496.

469    4.    **Cegelski L, Marshall GR, Eldridge GR, Hultgren SJ.** 2008. The biology and future prospects of
470         antivirulence therapies. Nat Rev Microbiol **6**:17-27.

471    5.    **Rasko DA, Sperandio V.** 2010. Anti-virulence strategies to combat bacteria-mediated disease.
472         Nat Rev Drug Discov **9**:117-128.

473    6.    **Allen RC, Popat R, Diggle SP, Brown SP.** 2014. Targeting virulence: can we make evolution-proof
474         drugs? Nat Rev Microbiol **12**:300-308.

475    7.    **Suerbaum S, Michetti P.** 2002. Helicobacter pylori infection. N Engl J Med **347**:1175-1186.

476    8.    **Urgesi R, Cianci R, Riccioni ME.** 2012. Update on triple therapy for eradication of Helicobacter
477         pylori: current status of the art. Clin Exp Gastroenterol **5**:151-157.

478    9.    **Schirm M, Soo EC, Aubry AJ, Austin J, Thibault P, Logan SM.** 2003. Structural, genetic and
479         functional characterization of the flagellin glycosylation process in Helicobacter pylori. Mol
480         Microbiol **48**:1579-1592.

481   10.    **Eaton KA, Suerbaum S, Josenhans C, Krakowka S.** 1996. Colonization of gnotobiotic piglets by
482         Helicobacter pylori deficient in two flagellin genes. Infect Immun **64**:2445-2448.

483   11.    **Ottemann KM, Lowenthal AC.** 2002. Helicobacter pylori uses motility for initial colonization and
484         to attain robust infection. Infect Immun **70**:1984-1990.

485   12.    **Van Zanten SJ, Dixon MF, Lee A.** 1999. The gastric transitional zones: neglected links between
486         gastroduodenal pathology and helicobacter ecology. Gastroenterology **116**:1217-1229.

487   13.    **Boonjakuakul JK, Canfield DR, Solnick JV.** 2005. Comparison of Helicobacter pylori virulence
488         gene expression in vitro and in the Rhesus macaque. Infect Immun **73**:4895-4904.

489   14.    **Thibault P, Logan SM, Kelly JF, Brisson JR, Ewing CP, Trust TJ, Guerry P.** 2001. Identification of
490         the carbohydrate moieties and glycosylation motifs in Campylobacter jejuni flagellin. J Biol Chem
491         **276**:34862-34870.

492   15.    **Guerry P, Ewing CP, Schirm M, Lorenzo M, Kelly J, Pattarini D, Majam G, Thibault P, Logan S.**
493         2006. Changes in flagellin glycosylation affect Campylobacter autoagglutination and virulence.
494         Mol Microbiol **60**:299-311.

495   16.    **Knirel YA, Vinogradov EV, L'Vov V L, Kocharova NA, Shashkov AS, Dmitriev BA, Kochetkov NK.**
496         1984. Sialic acids of a new type from the lipopolysaccharides of Pseudomonas aeruginosa and
497         Shigella boydii. Carbohydr Res **133**:C5-8.

498   17.    **Kenyon JJ, Marzaioli AM, Hall RM, De Castro C.** 2014. Structure of the K2 capsule associated
499         with the KL2 gene cluster of Acinetobacter baumannii. Glycobiology **24**:554-563.

500   18.    **Kiss E, Reuhs BL, Kim JS, Kereszt A, Petrovics G, Putnoky P, Dusha I, Carlson RW, Kondorosi A.**
501         1997. The rkpGHI and -J genes are involved in capsular polysaccharide production by Rhizobium
502         meliloti. J Bacteriol **179**:2132-2140.

503   19.    **Horzempa J, Dean CR, Goldberg JB, Castric P.** 2006. Pseudomonas aeruginosa 1244 pilin
504         glycosylation: glycan substrate recognition. J Bacteriol **188**:4244-4252.

505  20.  **Schoenhofen IC, McNally DJ, Brisson JR, Logan SM.** 2006. Elucidation of the CMP-pseudaminic
506       acid pathway in Helicobacter pylori: synthesis from UDP-N-acetylglucosamine by a single
507       enzymatic reaction. Glycobiology **16:**8C-14C.
508  21.  **Liu F, Tanner ME.** 2006. PseG of pseudaminic acid biosynthesis: a UDP-sugar hydrolase as a
509       masked glycosyltransferase. J Biol Chem **281:**20902-20909.
510  22.  **Schoenhofen IC, McNally DJ, Vinogradov E, Whitfield D, Young NM, Dick S, Wakarchuk WW,**
511       **Brisson JR, Logan SM.** 2006. Functional characterization of dehydratase/aminotransferase pairs
512       from Helicobacter and Campylobacter: enzymes distinguishing the pseudaminic acid and
513       bacillosamine biosynthetic pathways. J Biol Chem **281:**723-732.
514  23.  **Creuzenet C, Schur MJ, Li J, Wakarchuk WW, Lam JS.** 2000. FlaA1, a new bifunctional UDP-
515       GlcNAc C6 Dehydratase/ C4 reductase from Helicobacter pylori. J Biol Chem **275:**34873-34880.
516  24.  **Schoenhofen IC, Lunin VV, Julien JP, Li Y, Ajamian E, Matte A, Cygler M, Brisson JR, Aubry A,**
517       **Logan SM, Bhatia S, Wakarchuk WW, Young NM.** 2006. Structural and functional
518       characterization of PseC, an aminotransferase involved in the biosynthesis of pseudaminic acid,
519       an essential flagellar modification in Helicobacter pylori. J Biol Chem **281:**8907-8916.
520  25.  **Rangarajan ES, Proteau A, Cui Q, Logan SM, Potetinova Z, Whitfield D, Purisima EO, Cygler M,**
521       **Matte A, Sulea T, Schoenhofen IC.** 2009. Structural and functional analysis of Campylobacter
522       jejuni PseG: a udp-sugar hydrolase from the pseudaminic acid biosynthetic pathway. J Biol Chem
523       **284:**20989-21000.
524  26.  **Ishiyama N, Creuzenet C, Miller WL, Demendi M, Anderson EM, Harauz G, Lam JS, Berghuis**
525       **AM.** 2006. Structural studies of FlaA1 from Helicobacter pylori reveal the mechanism for
526       inverting 4,6-dehydratase activity. J Biol Chem **281:**24489-24495.
527  27.  **Irwin JJ, Shoichet BK.** 2005. ZINC--a free database of commercially available compounds for
528       virtual screening. Journal of Chemical Information and Modeling **45:**177-182.
529  28.  **Sulea T, Hogues H, Purisima EO.** 2012. Exhaustive search and solvated interaction energy (SIE)
530       for virtual screening and affinity prediction. Journal of Computer-Aided Molecular Design
531       **26:**617-633.
532  29.  **Hogues H, Sulea T, Purisima EO.** 2014. Exhaustive docking and solvated interaction energy
533       scoring: lessons learned from the SAMPL4 challenge. Journal of Computer-Aided Molecular
534       Design.
535  30.  **Sulea T, Cui Q, Purisima EO.** 2011. Solvated interaction energy (SIE) for scoring protein-ligand
536       binding affinities. 2. Benchmark in the CSAR-2010 scoring exercise. Journal of Chemical
537       Information and Modeling **51:**2066-2081.
538  31.  **Naim M, Bhat S, Rankin KN, Dennis S, Chowdhury SF, Siddiqi I, Drabik P, Sulea T, Bayly CI,**
539       **Jakalian A, Purisima EO.** 2007. Solvated Interaction Energy (SIE) for Scoring Protein-Ligand
540       Binding Affinities. 1. Exploring the Parameter Space. Journal of Chemical Information and
541       Modeling **47:**122-133.
542  32.  **Sulea T, Purisima EO.** 2012. The solvated interaction energy method for scoring binding
543       affinities. Methods in molecular biology **819:**295-303.
544  33.  **Cui Q, Sulea T, Schrag JD, Munger C, Hung MN, Naim M, Cygler M, Purisima EO.** 2008.
545       Molecular dynamics-solvated interaction energy studies of protein-protein interactions: the
546       MP1-p14 scaffolding complex. J Mol Biol **379:**787-802.
547  34.  **Ozbal CC, LaMarr WA, Linton JR, Green DF, Katz A, Morrison TB, Brenan CJ.** 2004. High
548       throughput screening via mass spectrometry: a case study using acetylcholinesterase. Assay
549       Drug Dev Technol **2:**373-381.
550  35.  **Zhang JH, Chung TD, Oldenburg KR.** 1999. A Simple Statistical Parameter for Use in Evaluation
551       and Validation of High Throughput Screening Assays. J Biomol Screen **4:**67-73.

552    36.    **Riazi A, Strong PC, Coleman R, Chen W, Hirama T, van Faassen H, Henry M, Logan SM,**
553         **Szymanski CM, Mackenzie R, Ghahroudi MA.** 2013. Pentavalent single-domain antibodies
554         reduce Campylobacter jejuni motility and colonization in chickens. PLoS One **8:**e83928.
555    37.    **Gatta L, Vakil N, Vaira D, Scarpignato C.** 2013. Global eradication rates for Helicobacter pylori
556         infection: systematic review and meta-analysis of sequential therapy. BMJ **347:**f4587.
557
558