

## Positive Selection Scanning Reveals Decoupling of Enzymatic Activities of Carbamoyl Phosphate Synthetase in *Helicobacter pylori*

Wagied Davids,<sup>1</sup> Junaid Gamielidien,<sup>2</sup> David A. Liberles,<sup>1</sup> Winston Hide<sup>2</sup>

<sup>1</sup>Department of Biochemistry and Biophysics and Stockholm Bioinformatics Center, Stockholm University, 10691 Stockholm, Sweden

<sup>2</sup>South African National Bioinformatics Institute, University of the Western Cape, Bellville, South Africa

Received: 11 June 2001 / Accepted: 12 September 2001

**Abstract.** In an effort to detect factors which may be under positive selection, a survey for such genes in two pathogenic strains of *Helicobacter pylori* (J99 and 26695) was performed. Based on an analysis of synonymous and nonsynonymous substitutions, we identified 19 candidate genes under positive selection. A search for homologues with known crystallographic structures revealed *Escherichia coli* carbamoyl phosphate synthetase as a homologue of *H. pylori* carbamoyl phosphate synthetase. Carbamoyl phosphate synthetase as isolated from *E. coli* is a heterodimeric enzyme that possesses two different but coupled functionalities and is involved in the first committed step in the separate biosynthetic pathways for arginine and pyrimidine nucleotides. In this study, we provide evidence indicating that one of these functionalities appears to be under selective pressure. Reports from previously published site-directed mutagenesis studies point to a decoupling of amidotransferase and synthetase activities. Implications of these findings for a metabolic enzyme under positive selection are discussed in terms of the mechanisms of *H. pylori* pathogenesis.

**Key words:** Bioinformatics — Adaptation — Structure–function correlation — Human pathogen — Molecular evolution — Pathoadaptive mutations

### Introduction

The availability of complete genome sequences for pathogenic organisms enables us to approach the understanding of pathogenesis at a new level, that of the whole genome. Genome-scale studies enable more detailed and comparative analyses between virulent and avirulent strains of the same species or even between different isolates of the same strain for identification of candidate genes which are involved in pathogenesis. By comparing the genomic sequences of two virulent strains we may get an indication of which genes are under selective pressure to maintain virulence [previous work in this field has been reviewed by Lan and Reeves (2000)].

Pathogenicity of microorganisms is frequently considered a secondary property of adaptation to a foreign environment, namely, the host. The evolution of bacteria toward a more pathogenic phenotype could potentially involve the acquisition of additional genes that encode for specific virulence factors via horizontal transfer and appropriate functional modification or through modification or loss of preexisting genetic material (Sokurenko et al. 1999). Mechanisms involving a change of function depend on the occurrence of random mutations (pathoadaptive mutations) of preexisting genes which enable a pathogen to reach optimal fitness through a process of fixation within its novel environment.

According to the neutralist perspective on molecular evolution, the vast majority of mutations affecting gene function are neutral or deleterious and become fixed by random genetic drift in a population (Kimura 1983). In contrast, the selectionist viewpoint maintains that advantageous mutations occur at a higher frequency in more

genes than is thought by neutralists or near-neutralists. Theory predicts (i) a great rate variation in sequence evolution caused by directional selection of advantageous mutations and by hitchhiking effects for linked sites and (ii) an elevated rate of nonsynonymous substitutions in rapidly evolving genes. Genes on which positive selection operate are considered to have an evolutionary history where the nonsynonymous substitution rate ( $K_a$ ) is higher than the synonymous substitution rate ( $K_s$ ). Nonsynonymous nucleotide substitutions that alter the physicochemical properties of amino acids are known to occur at lower rates than those that leave the amino acid properties unaltered (Zhang 2000). Amino acid changes that result in radical changes in physicochemical properties of amino acids are therefore more likely to be subject to purifying selection than conservative ones and also most likely to contribute to evolutionary changes in protein function if fixed. This knowledge can therefore be used to study candidate genes which are under selective pressure in pathogenic bacteria in the hope of identifying factors responsible for increased host adaptivity and virulence.

Large-scale systematic searches for genes under positive selection have revealed well-documented cases of adaptive evolution (Endo et al. 1996; Benner et al. 2000; Liberles et al. 2001). However, sensitive methodology for large-scale study of adaptive amino acid changes responsible for protein structure and function has not been developed. Identifying from the myriad amino acid replacements those few most likely to be responsible for observed adaptive changes/functional differences appears to be an elusive task (Golding and Dean 1998; Suzuki and Gojobori 1999; Gaucher et al. 2001; Gu 2001). Sequence comparisons alone are usually considered insufficient to identify replacements of functional consequence. Site-directed mutagenesis, particularly when guided by phylogeny, can be useful to search among a limited number of replacements. However, the best way remains identifying likely replacements in the context of three-dimensional structures.

In this paper, we explore a methodology to screen for candidate genes involved in pathogenesis by selecting those genes for which nonsynonymous substitution rates are higher than synonymous substitutions, followed by analysis of known structural homologues using reported mutagenesis studies for inferring the correlation between structural and functional changes. This approach is applied to *H. pylori*, one of the few pathogens to have complete genome sequences determined for two strains, *H. pylori* J99 and 26995 (Tomb et al. 1997; Alm et al. 1999). *H. pylori* is probably the most common cause of acute and chronic bacterial infection of the stomach in humans, present in almost half of the world's population (Covacci et al. 1999).

Our approach, which can readily be automated, was used to screen genomes of two strains of *H. pylori* for

genes under selective pressure and subsequently to visualize regions which may be responsible for these adaptive changes using a structural homologue. The methodology employed in this study, similar to one developed for the analysis of the evolution of specificity in metazoan deoxyribonucleoside kinases (Almgren 2001) and in MAP kinases (Caffrey et al. 2000), potentially holds promise for aiding a screening program to discover new candidate genes involved in pathogenesis.

## Materials and Methods

### *DNA and Protein Sequences*

Predicted protein coding sequences were downloaded from the Genbank sequence database for *H. pylori* (J99 and 26695) and were translated to amino acid sequences. Protein structures were obtained from the Protein Data Bank (Berman et al. 2000).

### *Generation of In-Frame Coding Sequence Alignments*

Protein alignments were generated for sequence pairs having less than 100% identity in the FASTA (Pearson and Lipman 1988) search, using the Clustal W (Thompson et al. 1994) program. The alignments were used as "scaffolds" for making accurate alignments of the coding sequences, thus avoiding possible frameshifts.

### *Analysis of Synonymous and Nonsynonymous Substitutions to Identify Genes Under Positive Selection*

To identify genes containing regions under positive selection, we used the *wina* program (Endo et al. 1996), which calculates synonymous ( $K_a$ ) and nonsynonymous ( $K_s$ ) substitutions in 20 codon windows of a gene according to the Nei and Gojobori (1986) method. This windowing approach was used to identify gene regions under positive selection, where  $K_a$  for the whole gene would still be less than  $K_s$ , thus not missing interesting genes. This was especially significant in our study, as the two isolates tested are evolutionarily very close, and very few substitutions, especially nonsynonymous ones, were expected. Only those regions where  $K_s$  is greater than 0 were considered, to exclude possibly random, or nearly neutral, nonsynonymous point mutations. Similarly, those regions where  $K_s$  is greater than 1 were also excluded, to avoid the saturation effect of nucleotide substitution [see Liberles (2001) and Siltberg, Lagergren, and Liberles (manuscript in preparation) for a discussion of methods to identify genes under positive selection].

### *Identification of Homologues with a Known Tertiary Structure*

To identify homologues with known structures, a BLAST (Altschul et al. 1990) search was performed against protein sequences in the Protein Data Bank with an *E* value of 1.00 using standard BLAST settings. Pairwise PAM distances were calculated between BLAST hits to estimate the evolutionary distances between protein sequences, with a PAM 120 cutoff to identify close homologues. A multiple-sequence alignment was then generated between the two *H. pylori* sequences and the *E. coli* sequence.

### Visualization and Mapping Mutations onto Homologues with a Known Tertiary Structure

The Grantham (1974) physicochemical matrix was used to score amino acid substitutions. At each position, the Grantham matrix value representing the physicochemical distance between the *H. pylori* 26995 and the *H. pylori* J99 carbamoyl phosphate synthetase (CPS) sequences was utilized. The values were normalized by dividing by the maximal value. These normalized values were used to color each residue in a gradient from blue to red in the tertiary structure. This procedure was also performed for *H. pylori* 26995 and *Escherichia coli* CPS sequence differences. The obtained protein structures were then visualized using Rasmol (Sayle et al. 1995), with the amino acid residues colored by the Grantham factor.

## Results

### Identification of Genes Which Are Under Positive Selection in *H. pylori*

Results from a comparative analysis of two independent isolates, *H. pylori* J99 and 26995, revealed a high level of conservation, with only 7% of the proteins being strain specific according to Alm et al. 1999 (there may be some variability in this number). We performed an intraspecies search between *H. pylori* strains to ascertain whether new virulence genes may be identified from genes under positive selection.

Table 1 represents a list of genes broadly categorized according to metabolic vs antigenic selection depending on whether the protein products interact directly with the host immune system or participate in metabolism. A large portion of *H. pylori* genes appears to be under metabolic selection. This may indicate an effort to increase metabolic function rather than antigenic variation. Further speculation about the potential role of these mutations in maintaining virulence will be described in greater detail elsewhere (Gamielidien and Hide, manuscript in preparation).

To identify homologues with known structures, a search for protein sequences contained in the Protein Data Bank was performed. One protein, carbamoyl phosphate synthetase, appears in both Table 1 and the Protein Data Bank, with sequences separated by 102 PAM units.

### Structural Analysis of Carbamoyl Phosphate Synthetase

Using a novel methodology described under Materials and Methods, analyzed using a combination of phylogenetic, structural, and functional methods. The *E. coli* CPS amidotransferase subunit is shown in Fig. 1A, with mutations colored according to *H. pylori* 26995–J99 Grantham physicochemical matrix properties. Figure 1B depicts the *E. coli* CPS amidotransferase colored according to Grantham values between *H. pylori* 26995 and *E. coli* amidotransferases. Figure 2 shows a multiple-sequence alignment of this protein from the three organ-

**Table 1.** Gene groups on which positive selection may operate in *H. pylori*<sup>a</sup>

Designation	Functional annotation
<b>A. Antigenic selection</b>	
HP0818	Osmoprotection protein (proWX)
HP0522	cag pathogenicity island protein (cag3)
HP0547	cag pathogenicity island protein (cag26)
HP1274	Paralyzed flagella protein (pflA)
HP0459	virB4 homologue (virB4)
<b>B. Metabolic selection</b>	
HP0640	pol(A) polymerase (papS)
HP0279	Lipopolysaccharide heptosyltransferase-1 (rfaC)
Hp1191	ADP-heptose-lps heptosyltransferase II (rfaf)
HP1266	NADH-ubiquinone oxidoreductase (NQO3) subunit
HP0290	Diaminopemilate decarboxylase dap decarboxylase
HP1237	Carbamoyl phosphate synthetase (pyrA)
HP0329	NH <sub>3</sub> -dependent NAD <sup>+</sup> synthetase (nadE)
HP0006	Pantoate-β-alanine ligase (panC)
HP0366	Spore coat polysaccharide biosynthesis protein
HP1232	Dihydropteroate synthetase (folP)
HP0293	Paraaminobenzoate synthetase (pabB)
HP0004	Carbonic anhydrase (icfA)
HP0717	DNA polymerase III γ and τ subunits (dnaX)
HP0731	Hypothetical protein

<sup>a</sup> This analysis is based upon pairwise alignment of the genes in the genomes of the two strains, followed by calculation of  $K_a/K_s$  rate ratios using the windowing approach of Endo et al. (1996). Genes were divided into those selected by antigenic/immune pressures and those under metabolic selection and listed below.

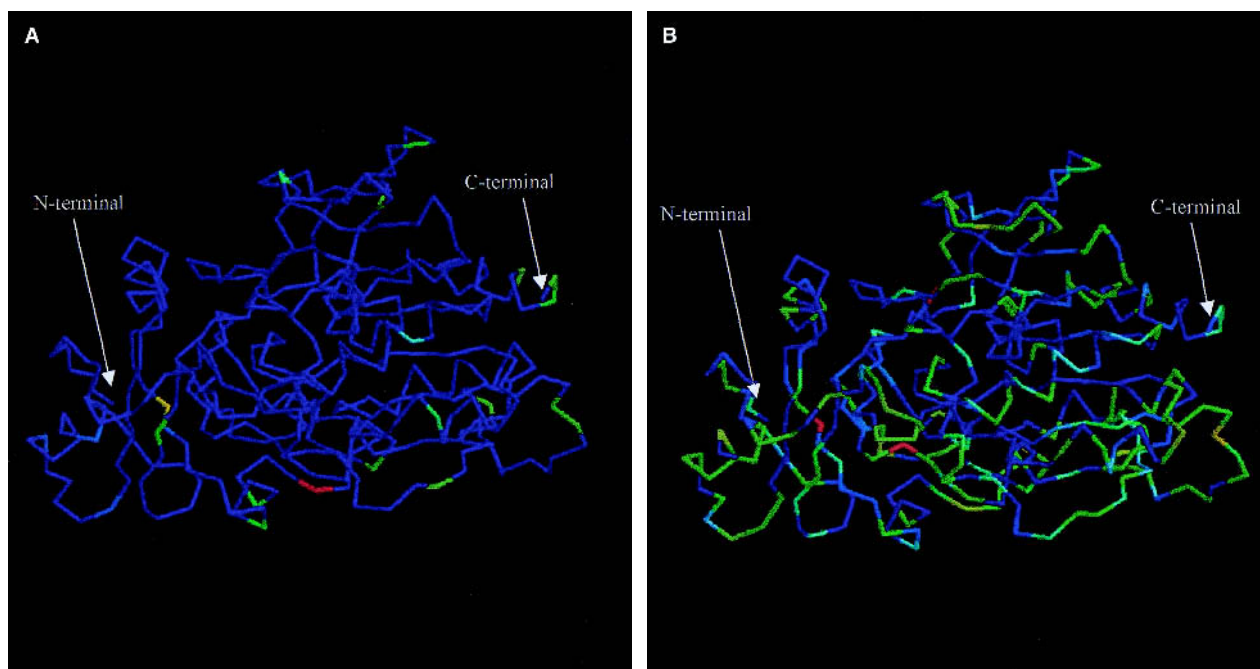
isms, *E. coli* and the two strains of *H. pylori*. An analysis of the individual mutations and their roles structurally and functionally is presented below.

## Discussion

### A Case Study: Comparative Structural–Functional Analysis of Carbamoyl Phosphate Synthetase

Carbamoyl phosphate synthetase (CPS) is a well-characterized enzyme in terms of its biochemistry (Anderson and Meister 1965), phylogeny (Lawson et al. 1996; Nyunoya et al. 1985), and structural biology (Thoden et al. 1999a, b). CPS is a metabolic enzyme involved in the formation of carbamoyl phosphate from one molecule of bicarbonate, two molecules of Mg<sup>2+</sup>-ATP, and one molecule of glutamine or ammonia and catalyzes the first committed step in the separate biosynthetic pathways for arginine and pyrimidine nucleotides. Carbamoyl phosphate (CP), a high-energy biological compound, plays a key role in the introduction of both ammonia and single carbon units into the metabolic pool and may serve either as the precursor for synthesis of pyrimidine nucleotides or in an alternate pathway, as the precursor for arginine.

Most bacteria studied to date contain a heterodimeric



**Fig. 1.** A visualization of the  $\alpha$ -carbon backbone model of *E. coli* carbamoyl phosphate synthetase using (A) *H. pylori* J99–26995 mutational differences colored by Grantham matrix values and (B) *H. pylori*–*E. coli* mutational differences colored by Grantham matrix values. A full description of the methodology is described under Materials and Methods.

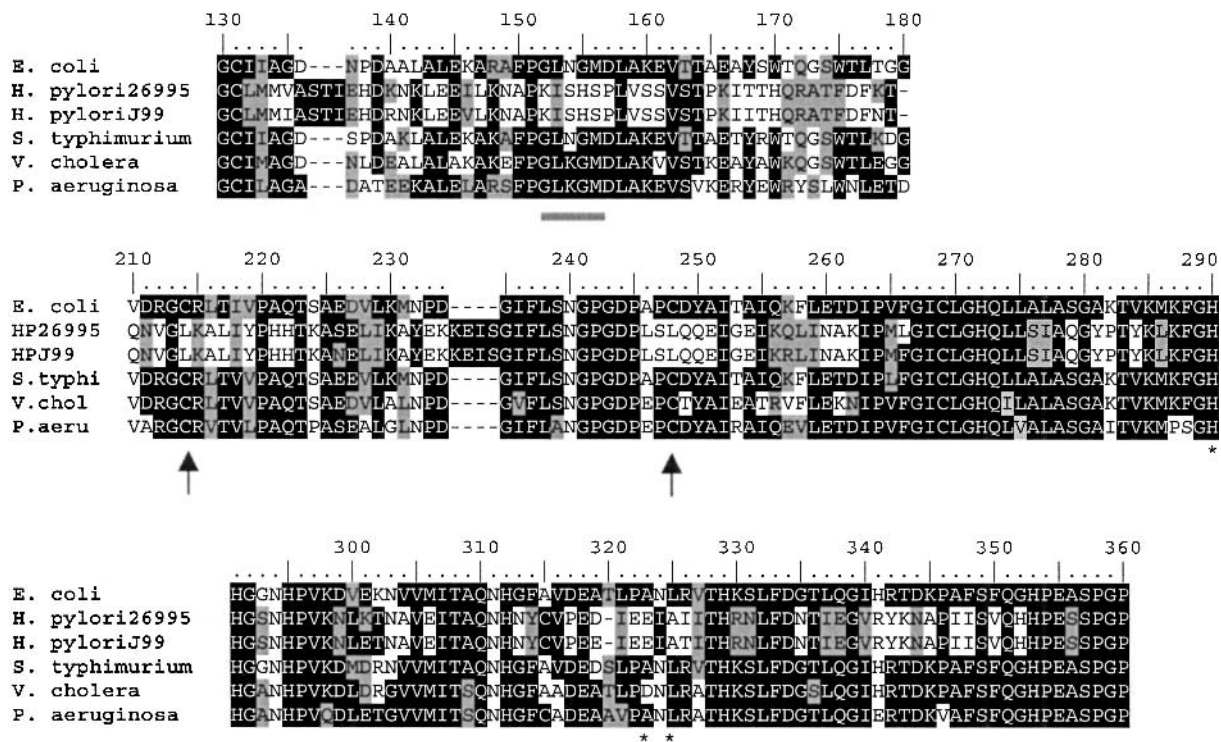
form of the enzyme which is composed of a small amidotransferase subunit of molecular weight  $\sim 42$  kDa complexed to a larger synthetase subunit of molecular weight  $\sim 118$  kDa linked via a hinge-like loop containing a type II  $\beta$  turn. The small subunit belongs to the triad or type I class of amidotransferases, which contain a cysteine–histidine (Cys269–His353) couple essential for activity. The recently obtained crystallographic structure of CPS from *E. coli* confirmed the presence of separate sites for the phosphorylation of bicarbonate and carbamate on the large subunit, while the small subunit contains the active site for the hydrolysis of glutamine to ammonia (Thoden et al. 1999a). These three active sites are separated by nearly 100 Å via an intramolecular tunnel that runs through the interior of the entire protein.

Our initial comparative analysis of CPS concentrated on mutations between *H. pylori* 26995 and *H. pylori* J99, as shown in Fig. 1A. Amino acid residues forming part of the catalytic triad (Cys269–His353–Glu355) of the active site of CPS are known to be evolutionary conserved in all bacterial CPS enzymes (Zalkin et al. 1998). These residues are conserved in both *H. pylori* strains, indicating their functional importance in catalysis. Adaptive mutations between *H. pylori* 26995 and *H. pylori* J99 appear to be novel and map to regions located on the surface of the enzyme. It is known that surface exposed amino acid residues are among the most rapidly evolving codons within a protein, frequently behaving neutrally (Miyamoto and Fitch 1995). If these residues are playing an adaptive role, rather than behaving neutrally, the apparent lack of function of these residues related to bind-

ing may suggest some allosteric role in affecting protein structure or function.

Comparative analysis between both *H. pylori* strains and *E. coli* CPS amidotransferase subunits revealed substitutions in regions with known functional importance from reported site-directed mutagenesis studies, as shown in Fig. 1B. The three histidine residues His272, His312, and His353 are conserved in both *H. pylori* strains. His341 is, however, replaced with arginine, representing a relatively conservative substitution. The role of these four conserved histidine residues, His272, His312, His341, and His353, within the amidotransferase subunit have been probed using site-directed mutagenesis studies (Miran et al. 1991). His353 forms part of the catalytic triad and also acts as a general acid–base catalyst for proton transfer. His312 serves a critical role for the binding of glutamine to the active site. Mutations of His272 and His341 have been shown to result in a two-fold reduction in the rate of CP synthesis relative to the wild-type enzyme. Arginine mutation at position 341 was shown to have smaller but real effects.

Studies by Mareya and co-workers have shown a good correlation between the extent of enhancement of the partial glutaminase activity and uncoupling of the phosphorylation reactions that occur on the large subunit (Mareya et al. 1994). Mutations aimed at identifying the reactive cysteine residues by mutating Cys131, Cys214, and Cys248 within the small subunit resulted in a mutant at position Cys248 with increased partial glutaminase activity relative to the wild-type enzyme, with the formation of carbamoyl phosphate using glutamine com-

**Note:**

- \* -active site residue
- -hinge- like loop

↑ -substitutions involved in decoupling of enzyme activity

**Fig. 2.** A multiple-sequence alignment of the carbamoyl phosphate synthetase amidotransferase subunit from Gram-negative pathogenic bacteria is presented. Important functional residues are indicated underneath the alignment following the Discussion.

pletely abolished. Mutations at Glu355 are also known to uncouple the two partial reactions such that no carbamoyl phosphate is produced (Huang et al. 1999).

Residue Cys269 is conserved in *H. pylori* 26995 and J99. It forms part of the catalytic triad and directly participates in the chemical reaction mechanism of glutamine hydrolysis via the formation of a thioester intermediate (Wellner et al. 1975; Rubino et al. 1986; Thoden et al. 1998). Cysteine residues at positions Cys214 and Cys248 are replaced with leucine residues in both *H. pylori* strains, representing radical substitutions. Most notable is the mutation at Cys248, which NEM-labeling studies have shown to be the reactive cysteine residue on the small subunit (Mareya et al. 1994). Mutations at residue Cys248 are known to result in an increase in partial glutaminase activity and a decrease in synthetase activity. The severity of these substitutions may indicate an important role in maintaining a coordinated functional linkage between the small amidotransferase and the large synthetase subunit in the *E. coli* CPS (Hewagama et al. 1999). This linkage is mediated allosterically, where binding of substrate to the synthetase domain is necessary for glutamine hydrolysis in the glutaminase domain.

Position Glu355 is conserved in both *E. coli* and *H. pylori* strains, indicating its conserved function in interacting with the  $\gamma$ -glutamyl thioester intermediate formed during glutamine hydrolysis (Thoden et al. 1998).

A fully functional tunnel is considered essential for the structural integrity and catalytic coupling of the sequential and parallel reactions occurring within the three active sites contained within the small and large subunits of CPS. The role of the four residues, Leu153, Asn154, Gly155, and Met156, making up the type II  $\beta$  turn in maintaining critical catalytic coupling interactions at the molecular interface has been investigated (Huang et al. 2000a-c). An effort was made to disrupt interactions at the interface between the two domains by replacing residue Leu153 (the first residue of the type II turn) with bulkier residues, resulting in a Leu153 mutant with increased glutaminase catalytic activity but also a partial uncoupling between the hydrolysis of ATP and glutamine during the overall synthesis of CP. Mutations aimed at rendering the hinge-loop more conformationally flexible by mutating flanking residues Gly152 and Gly155 resulted in increased glutaminase activity with reduced ATPase activity.

A multiple sequence alignment of amidotransferase subunits of Gram-negative pathogenic bacteria inhabiting the more neutral-pH gastrointestinal tract (*E. coli*, *S. typhimurium*, *V. cholera*, and *P. aeruginosa*) revealed an apparent conservation in the hinge-like region (Fig. 2). In contrast, *H. pylori* strains display substitutions with varying Grantham values from residues 152 to residue 156 forming the type II  $\beta$  turn. A conservative substitution at Leu153, replaced by isoleucine, indicates the importance of this position as being the first residue in the type II turn.

Our comparative intraspecies analyses between two pathogenic *H. pylori* strains reveal ongoing adaptive changes between similar strains through novel mutations and, thus, seem to highlight differentially regions that are functionally diverging under selective pressure in the persistence of *H. pylori*. Further investigation of *H. pylori* 26995 CPS with its *E. coli* homologue reveals evidence suggestive of the effect of adaptive evolution on the decoupling of enzymatic activities present in *H. pylori* CPS. Our results indicate (i) conservation of residues forming the catalytic triad required for proper glutaminase activity of the amidotransferase subunit, (ii) mutations of residues involved in forming the hinge-like region of the type II  $\beta$  turn required for proper interface interactions and catalytic coupling of partial reactions between two separate subunits of CPS, and (iii) mutations of reactive cysteine residues Cys214 and Cys248, notably Cys248, which is responsible for increased glutaminase activity and decoupling of enzyme activities. Taken together, these findings indicate the effect of positive selection on the decoupling of amidotransferase from synthetase activities in the emergence of *H. pylori* CPS, with further evidence for continued selection in at least one of the *H. pylori* strains. This may correlate with differences in the isolates, which occupy different regions of the gastrointestinal tract with different clinical manifestations.

#### *Extending from the Protein to the Organism*

A study of the evolution of bacteria indicates that they have evolved many mechanisms to survive and prosper in unfavorable environments. Survival in the gastric environment seems to be the most complex aspect of the phenotype specific to *H. pylori*. Analogous to acidophilic bacteria, *H. pylori* apparently maintains a constant cytoplasmic pH (7.0–7.3) and membrane potential (–85 to –240 mV), which are required to drive ATP synthesis and nutrient transport over a wide pH range (Berg et al. 1997). Outer membrane proteins are proposed to play a critical role in the ability of *H. pylori* to colonize host cells in an acidic environment (Alm et al. 2000). One of *H. pylori*'s most characteristic enzymes is known to be a potent multisubunit urease which metabolizes urea to carbon dioxide and ammonia and is crucial for its sur-

vival and successful colonization of the gastric environment (Weeks et al. 2000). Although intracellular metabolic enzymes have also been found in other studies to be important in *H. pylori* species-specific phenotype, their roles have not been analyzed in detail (Huynen et al. 1998).

In this study, we have performed a search for genes under positive selection between two strains of *H. pylori* to ascertain whether new virulence genes may be identified this way. We have identified 19 candidate genes in *H. pylori* that are under positive selection based on an increased ratio of synonymous/nonsynonymous substitutions. Some of these genes may mediate specific adaptations of *H. pylori* to its environment. We have analyzed one such candidate gene, an *E. coli* structural homologue to *H. pylori* CPS. Based on reported mutagenesis studies performed on the *E. coli* CPS, we have mapped these mutations to regions which are implicated in decoupling of enzyme activities in CPS. Our use of a physicochemical distance matrix to weight the severity of substitutions between two strains of *H. pylori* CPS coincides with reported site-directed mutagenesis studies of functional changes and also highlight its usefulness in our structure–function analysis for probing putatively adaptive mutational events. The availability of structure–function information obtained from published site-directed mutagenesis studies enabled us to perform a more detailed analysis that could be used for analyzing candidate genes under positive selection and also provide rigorous tests of the structural, functional, and, by implication, adaptive hypotheses in our study. However, the strategy employed during our study is dependent on a structural homologue and previous mutagenesis studies. Although the case study exemplifies this approach, its potential promises to yield profound effects in bringing candidate genes to the surface when new protein structures are determined.

It is interesting to speculate about the amino acid changes seen in *H. pylori* and their contribution to pathogenesis. Genes coding for products that do not interact directly with the host immune system are particularly interesting since they may represent selection for improved function rather than antigenic variation. Substitutions in *H. pylori* CPS (from its ancestor with *E. coli* CPS) map to regions which are known to decouple amidotransferase from synthetase activities, leading to increased amidotransferase activity and resultant ammonia production. It is thought that the increased ammonia production acts as a buffer to increase the local pH required for successful colonization and growth in an acidic gastric environment. Positive selection of CPS could therefore represent an alternative mechanism of adaptation to the gastric environment affording a selective advantage in adapting to its human host.

*Acknowledgments.* This work was supported by SIDA Grant SRP-2000-065P and by SSF (Sweden). Wagied Davids is the recipient of a

fellowship from the National Research Foundation (South Africa). This work is dedicated to the memory of Ralston Muller.

## References

- Alm RA, et al. (1999) Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* 397:176–180
- Almgren MAE (2001) Genomic approaches to understanding protein function, MSc thesis. Stockholm University, Stockholm
- Anderson PM, Meister A (1965) Evidence for an activated form of carbon dioxide in the reaction catalysed by *Escherichia coli* carbamoyl phosphate synthetase. *Biochemistry* 4:2803–2809
- Altschul SF, Stephen FF, Gish W, Miller W, Meyers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Benner SA, Chamberlin SG, Liberles DA, Govindarajan S, Knecht L (2000) Functional inferences from reconstructed evolutionary biology involving rectified databases—An evolutionarily grounded approach to functional genomics. *Res Microbiol* 151:97–106
- Berg DE, Hoffman PS, Appelmelk BJ, Kusters J (1997) The *Helicobacter pylori* genome sequence: Genetic factors for long life in the gastric mucosa. *Trends Microbiol* 5:468–474
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235–242
- Caffrey DR, O'Neill LAJ, Shields DC (2000) A method to predict residues conferring functional differences between related proteins: Application to MAP kinase pathways. *Protein Sci* 9:655–670
- Covacci A, Telford JL, Del Giudice G, Parsonnet J, Rappuoli R (1999) *Helicobacter pylori* virulence and genetic geography. *Science* 284:1328–1333
- Endo T, Ikeo K, Gojobori T (1996) Large-scale search for genes on which positive selection may operate. *Mol Biol Evol* 13:685–690
- Gaucher EA, Miyamoto MM, Benner SA (2001) Function-structure analysis of proteins using covarion-based evolutionary approaches: Elongation factors. *Proc Natl Acad Sci USA* 98:548–552
- Grantham R (1974) Amino acid difference formula to help explain protein evolution. *Science* 85:862–864
- Golding GB, Dean A (1998) The structural basis of molecular adaptation. *Mol Biol Evol* 15:355–369
- Gu X (2001) Maximum likelihood approach for gene family evolution under functional divergence. *Mol Biol Evol* 18:453–464
- Hewagama A, Guy HI, Vickrey JF, Evans DR (1999) Functional linkage between the glutaminase and synthetase domains of carbamoyl-phosphate synthetase. *J Biol Chem* 274:28240–28245
- Huang X, Raushel FM (1999) Deconstruction of the catalytic array within the amidotransferase subunit of carbamoyl phosphate synthetase. *Biochemistry* 30:15909–15914
- Huang X, Raushel FM (2000a) Role of the hinge loop linking the N- and C-terminal domains of the amidotransferase subunit of carbamoyl phosphate synthetase. *Arch Biochem Biophys* 380:174–180
- Huang X, Raushel FM (2000b) An engineered blockage within the ammonia tunnel of carbamoyl phosphate synthetase prevents the use of glutamine as a substrate but not ammonia. *Biochemistry* 39:3240–3247
- Huang X, Raushel FM (2000c) Restricted passage of reaction intermediates through the ammonia tunnel of carbamoyl phosphate synthetase. *J Biol Chem* 275:26233–26240
- Huynen M, Dandekar T, Bork P (1998) Differential genome analysis applied to the species-specific features of *Helicobacter pylori*. *FEBS Lett* 426:1–5
- Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, London
- Lan R, Reeves PR (2000) Intraspecific variation in bacterial genomes: The need for a species genome concept. *Trends Microbiol* 8:396–401
- Lawson FS, Charlesbois RL, Dillon JAR (1996) Phylogenetic analysis of carbamoylphosphate synthetase genes: Complex evolutionary history includes an internal duplication within a gene which can root the tree of life. *Mol Biol Evol* 13(7):970–977
- Liberles DA (2001) Evaluation of methods for determination of a reconstructed history of gene sequence evolution. *Mol Biol Evol* 18(11):2040–2047
- Liberles DA, Schreiber DR, Govindarajan S, Chamberlin SG, Benner SA (2001) The adaptive evolution database (taed). *Genome Biol* 2(8):research0028.1–research0028.6
- Mareya SM, Raushel FM (1994) A molecular wedge for triggering the amidotransferase activity of carbamoyl phosphate synthetase. *Biochemistry* 33:2945–2950
- Miran SG, Chang SH, Raushel FM (1991) Role of four conserved histidine residues in the amidotransferase domain of carbamoyl phosphate synthetase. *Biochemistry* 30:7901–7907
- Miyamoto M, Fitch WM (1995) Testing the covarion hypothesis of molecular evolution. *Mol Biol Evol* 12:503–513
- Nei M, Gojobori T (1986) Simple methods for the estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3:418–426
- Nyunoya H, Broglie KE, Lusty CJ (1985) The gene coding for carbamoyl phosphate synthetase I was formed by fusion of an ancestral glutaminase gene and synthetase gene. *Proc Natl Acad Sci USA* 82:2244–2246
- Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. *Proc Natl Acad Sci USA* 85:2444–2448
- Rubino SE, Nyunoya H, Lusty CJ (1986) Catalytic domains of carbamoyl phosphate synthetase. *J Biol Chem* 261:11320–11327
- Sayle RA, Milner-White EJ (1995) RASMOL: Biomolecular graphics for all. *Trends Biochem Sci* 20:374
- Sokurenko EK, Hasty DL, Dykhuizen DE (1999) Pathoadaptive mutations: Gene loss and variation in bacterial pathogens. *Trends Microbiol* 7:191–195
- Suzuki Y, Gojobori T (1999) A method for detecting positive selection at single amino acid sites. *Mol Biol Evol* 16:1315–1328
- Thoden JB, Miran SG, Phillips JC, Howard A, Raushel FM, Holden HM (1998) Carbamoyl phosphate synthetase: Caught in the act of glutamine hydrolysis. *Biochemistry* 37:8825–8831
- Thoden JB, Raushel FM, Benning MM, Rayment I, Holden HM (1999a) The structure of carbamoyl phosphate synthetase determined to 2.1 Å resolution. *Acta Crystallogr D Biol Crystallogr* 55:8–24
- Thoden JB, Huang X, Raushel FM, Holden HM (1999b) The small subunit of carbamoyl phosphate synthetase: Snapshots along the reaction pathway. *Biochemistry* 38:16158–16166
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Tomb JF, et al. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388:539–547
- Weeks DL, Eskanadri S, Scott DR, Sachs G (2000) A H<sup>+</sup>-gated urea channel: The link between *Helicobacter pylori* and gastric colonization. *Science* 287:482–486
- Wellner VP, Meister A (1975) Enhancement of glutaminase activity of carbamoyl phosphate synthetase by alterations in the interaction between the heavy and light subunits. *J Biol Chem* 250:3261–3266
- Zalkin H, Smith JL (1998) Enzymes utilizing glutamine as an amide donor. *Adv Enzymol Relat Areas Mol Biol* 72:87–145
- Zhang J (2000) Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. *J Mol Evol* 50:56–68