Helicobacter pylori RNA polymerase α -subunit C-terminal domain shows features unique to ϵ -proteobacteria and binds NikR/DNA complexes

Brendan N. Borin,¹ Wei Tang,² and Andrzej M. Krezel^{1,2,3*}

¹Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee 37232

²Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110

³Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, Missouri 63110

Received 12 November 2013; Revised 13 January 2014; Accepted 14 January 2014

DOI: 10.1002/pro.2427

Published online 17 January 2014 proteinscience.org

Abstract: Bacterial RNA polymerase is a large, multi-subunit enzyme responsible for transcription of genomic information. The C-terminal domain of the α subunit of RNA polymerase (α CTD) functions as a DNA and protein recognition element localizing the polymerase on certain promoter sequences and is essential in all bacteria. Although α CTD is part of RNA polymerase, it is thought to have once been a separate transcription factor, and its primary role is the recruitment of RNA polymerase to various promoters. Despite the conservation of the subunits of RNA polymerase among bacteria, the mechanisms of regulation of transcription vary significantly. We have determined the tertiary structure of *Helicobacter pylori* α CTD. It is larger than other structurally determined α CTDs due to an extra, highly amphipathic helix near the C-terminal end. Residues within this helix are highly conserved among ϵ -proteobacteria. The surface of the domain that binds A/T rich DNA sequences is conserved and showed binding to DNA similar to α CTDs of other bacteria. Using several NikR dependent promoter sequences, we observed cooperative binding of *H. pylori* α CTD to NikR:DNA complexes. We also produced α CTD lacking the 19 C-terminal residues, which showed greatly decreased stability, but maintained the core domain structure and binding affinity to NikR:DNA at low temperatures. The modeling of *H. pylori* α CTD into the context of transcriptional complexes suggests that the additional amphipathic helix mediates interactions with transcriptional regulators.

Keywords: RNA polymerase; α CTD; *Helicobacter pylori*; protein-DNA interaction; protein structure; NMR; NikR; EMSA

Abbreviations: α CTD, α subunit C-terminal domain; HSQC, heteronuclear single quantum correlation; NOESY, nuclear Overhauser effect spectroscopy; NTD, N-terminal domain; RMSD, root mean square deviation; TSS, transcriptional start site; PDB, protein databank; EMSA, electrophoretic mobility shift assay

Additional Supporting Information may be found in the online version of this article

Grant sponsor: NIH; Grant number: AI67628. Grant sponsor: NIH Molecular Biophysics Training Program; Grant number: 008320.

*Correspondence to: Andrzej M. Krezel, Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, 660 S. Euclid Avenue, St. Louis, MO 63110. E-mail: krezela@wustl.edu

Introduction

RNA polymerase is an essential protein complex that transcribes DNA into RNA. In bacteria, the core enzyme has a composition of $\alpha_2\beta\beta'\omega$, with a σ subunit completing the holoenzyme. The β and β' subunits are the largest and contain the catalytic site. Each is bound to one of the α subunits, which stabilize the complex by dimerizing through their N-terminal domains. The ω subunit has both structural and functional roles in the complex.¹ Promoter-specific initiation of transcription is facilitated by σ subunits (factors), which bind the core complex and recognize specific promoter elements.

Figure 1. Amino acid sequence alignment of several bacterial α CTDs including part of the interdomain linker. PFAM-defined conserved domain is boxed. The short 3_{10} helical segments are indicated by boxes above the sequences and are labeled ht1 and ht2. α -helices are indicated by cylinders above the sequences and are labeled H1-H5. Helix 5 is an additional α -helix found in *H. pylori*. The conservation in all sequences presented is indicated by symbols in the bottom line; * invariant, : very highly conserved, . conserved. Numbers correspond to positions in the native sequence of each protein. Hp *Helicobacter pylori*; Hh *Helicobacter hepaticus*; Ws *Wolinella succinogenes*; Cj *Campylobacter jejuni*; Ec *Escherichia coli*; Vc *Vibrio cholerae*; Bs *Bacillus subtilis*; Lm *Listeria monocytogenes*; Mt *Mycobacterium tuberculosis*.

Despite the high degree of evolutionary conservation among bacterial RNA polymerases in general, the *H. pylori* RNA polymerase exhibits several distinct features. Unlike in most species, *Helicobacter* and *Wolinella* β and β' subunits are found as a single fused gene product and this fusion likely confers a selective advantage.² Also, the *H. pylori* σ^{80} specificity subunit of the RNA polymerase has diverged significantly from the σ^{70} subunit of *E. coli* (32% identity, 51% similarity) and other bacteria.³ Even larger divergence is apparent between *H. pylori* and *E. coli* α subunit sequences showing 28% identity and 50% similarity.

The C-terminal domains of the α subunits (α CTDs) are separated from the dimerizing N-terminal domains by long, flexible linkers,⁴ and are essential for growth in *E. coli* and most likely all other bacteria.⁵ They are known to play an important role in transcription of many genes. A highly conserved region of the domain interacts with upstream elements, binding the minor groove of A/T rich sequences that are often found near the consensus -10 and -35 elements of bacterial promoters. Highly diverse examples in which the α CTDs stimulate transcription in a complex fashion were reported,^{6,7} using several areas on the surface of the domain to contact DNA and different transcription factors.⁸⁻¹² There are some examples of DNA-binding proteins that compete with α CTD for binding upstream elements to repress transcription,¹³ and although most interactions involving α CTD result in higher transcription rates, some lower transcription by positioning RNA polymerase in a suboptimal location with respect to the transcription start site.¹⁴

A sequence alignment of 11 bacterial α CTDs including flexible linkers between the N-terminal domain (NTD) and the CTD is shown in Figure 1. The 70-residue core of the RNA polymerase α subunit C-terminal domain (Pfam¹⁵ PF03118) is highly conserved in bacteria. The average identity between members of this domain family is 45%. The conservation does not extend to a number of C-terminal residues beyond the core domain. These C-termini diverge in length and sequence; however, among ϵ -proteobacteria, there is very good conservation in positions occupied by hydrophobic and charged residues.

The first determined structure of an α CTD, from *E. coli*, showed that these C-terminal residues are, in fact, well-ordered and contribute key hydrophobic core interactions.¹⁶ Additional examples of isolated and bound α CTD supported this observation.^{8,16,17} In *H. pylori*, the α CTD C-terminal segment (residues 314-344) is eleven residues longer than in *E. coli*, and the two species' sequences are highly dissimilar. We determined the solution structure of α CTD from *H. pylori*. The structure shows that not only are the C-terminal residues ordered, but that they form a fifth helix. We analyzed the contribution of this fifth helix to the domain stability by producing a 19-residue truncated form of the domain containing only residues 231-325, hereafter referred to as α CTD325. Both full length and truncated α CTD domains were used in binding assays with the *H. pylori* NikR transcriptional regulator and several NikR promoter sequences. These experiments showed that high affinity binding of *H. pylori* α CTD to DNA depends critically on the presence of A/T-rich sequences and interactions with NikR. We

Table I. Structural Statistics for Ensemble of 15 Structures of *H. pylori* α CTD

Nonredundant NOE restraints	957
Intraresidue	316
Short	252
Medium	169
Long	220
Hydrogen bond restraints	36
TALOS dihedral angle restraints	119
Average CYANA target function	1.52
Number of violations > 0.2 Å	0
Average AMBER energies ($\pm$ standard deviation)	
Input structures	3953 ($\pm$ 2135)
Energy minimized structures	–4354 ($\pm$ 24)
Average Ramachandran statistics from PROCHECK (residues 254–337)	
Most favored (%)	90.0
Additionally allowed (%)	9.4
Generously allowed (%)	0.6
Disallowed (%)	0
Average RMSD from mean structure (Å, residues 254–337)	
Backbone (N,C $^{\alpha}$,C',O)	0.41
Heavy atoms	0.93

have modeled *H. pylori* α CTD into known three-dimensional structures of α CTD-containing complexes. Although much work has been done to describe α CTD interactions both structurally and functionally in *E. coli* and *Bacillus subtilis*, this type of analysis has not yet been done for *H. pylori*. Our models show that the *H. pylori* α CTD would not be able to fit within the *E. coli* protein-protein interactions. The inability of *H. pylori* and *E. coli* RNA polymerases to utilize each other's promoters due to differences in the σ subunits has been documented.^{18,19} We also present a model of an interaction between α CTD and NikR, a transcription factor that positively regulates the urease operon in *H. pylori*.

Results

Structure of the full length *H. pylori* α CTD (residues 231–344)

To determine the structure of the α CTD from *H. pylori*, we produced a 114-residue fragment (residues 231–344) consisting of the interdomain linker, the core domain, and the C-terminal residues to the end of the native protein sequence. We used NMR methods to determine a solution structure. The final structural ensemble (residues 256–337) consists of 15 models with an average backbone RMSD of 0.41 Å and heavy atom RMSD of 0.93 Å (Table I). Similar to other bacterial α CTDs, the well-conserved domain consists of four helices with short connecting loops, and a longer N-terminal, ordered loop (A254–S267) containing two isolated, short 3_{10} helical segments that contribute residues critical to the hydrophobic core. Analyzing the final ensemble, we noticed that

the backbone carbonyl oxygen atoms of N-terminal loop residues A259 and L257 act as hydrogen bonding partners for the backbone amide hydrogen atoms of helix 2 terminal residues V282 and G283, respectively. During the course of resonance assignments, the backbone amide proton of S267, also in the N-terminal loop, was found far downfield shifted at 10.43 ppm, suggesting a strong hydrogen bond. Initial structure calculations showed the amide proton to be in the vicinity of the sidechain oxygens of E306. This glutamate is highly conserved, and crystal structures of homologous domains exhibit these hydrogen bonds, so we included them as structural restraints in the remaining calculations. Helices 2 and 3 are rather short, while helices 1 and 4 are longer and contribute residues that are required for DNA-binding. Chemical shift perturbation experiments were performed to confirm that α CTD binds A/T rich DNA in a similar manner as other α CTDs do (data not shown).

The structure of the C-terminal segment was quite different from that of *E. coli* α CTD residues 310–329 [Fig. 2(b)], which is well ordered without taking on regular secondary structure, and makes important hydrophobic contacts.^{16,17} In the case of the *H. pylori* α CTD, the C-terminal segment forms a short loop and the additional fifth helix [Fig. 2(a)]. Helix 5 is amphipathic, contributing three leucine sidechains that make many contacts with other hydrophobic residues in the hydrophobic core. Nearly all other residues in this helix have charged sidechains. To accommodate the C-terminal helix, *H. pylori* α CTD also contains additional hydrophobic residues in helix 3. L287 in *H. pylori* replaces Q283 in *E. coli* and makes contact with L330 of helix 5.

A heteronuclear ^1H - ^{15}N NOE experiment (Fig. 3) was recorded to determine the position at which the flexible linker segment (connecting α NTD and α CTD) ends and the stable, structured domain begins, and also to determine the dynamics of the C-terminal residues. The first residue with a steady-state NOE greater than 0.5 is L257, which also displays many hydrophobic contacts to the core of the domain. Residue A254 has an NOE of 0.5, on the borderline between the disordered and stable, well-defined conformations. This is consistent with our observation of just a few interactions between A254 and the structured domain. The heteronuclear steady-state NOEs also showed that the sequence corresponding to the well-conserved core domain (residues 255–313) up to the end of helix 4 has a stable backbone conformation. There is a steady decline in the heteronuclear NOEs from Y316 to L322. The structural ensemble shows that this region has a greater range of conformations than the rest of the structured domain. On the other hand, the backbone of residues S323 to L337 appears well ordered based on the measured heteronuclear NOEs. This is supported by the

Figure 2. Structure of *H. pylori* RNA polymerase α CTD residues 256–337. A: Ribbon diagram with rainbow coloring, N-terminus blue and C-terminus red. All α -helical segments are labeled, the fifth, C-terminal helix is visible at the top in orange and red. Sidechains of the three leucines from the helix 5 and other hydrophobic residues that they contact are shown. Well-conserved DNA-binding residues (R269, N298, K302) are also indicated. B: Superimposition of *H. pylori* α CTD (orange) with those of *E. coli* (purple, PDB ID 3K4G) and *B. subtilis* (blue, PDB ID 3GFK).

positioning of these residues in an α -helix that contributes side chains to the hydrophobic core.

Stability of the truncated *H. pylori* α CTD (residues 231–325)

We produced α CTD325, a truncated version of *H. pylori* α CTD lacking the last 19 residues, corresponding to the 5th helix (residues S323–L337) and

Figure 3. Heteronuclear ^1H - ^{15}N NOE values for backbone amides of α CTD. NOE values were calculated as a ratio of intensities of peaks in the spectra with and without saturation, and are plotted against the residue numbers. The two short 3_{10} helical segments and the five α -helices are indicated by boxes above the graph.

the last seven C-terminal unstructured residues (EDKGGND). We assessed the stability of α CTD325 by circular dichroism (CD) and NMR methods. The stability of α -helices 1 to 4 in α CTD325 decreased dramatically, as measured by a decrease in ellipticity with increasing temperature [Fig. 4(a)]. The CD derived²⁰ midpoint temperature of unfolding changed from 54°C to 32°C, and the corresponding folding enthalpy from –51 kcal/mol to –27 kcal/mol. However, the patterns of peak dispersion in ^1H - ^{15}N correlated spectra (HSQC) of ^{15}N labeled α CTD325 and α CTD were very similar [Fig. 4(b)] at temperatures up to 28°C, where loss of tertiary structure was evident. Above 28°C, the sample showed ^1H - ^{15}N HSQC peak pattern indicative of unstructured polypeptide. Similarly, the NOESY spectra of α CTD325 showed multiple, well resolved through-space cross peaks at temperatures below 28°C (data not shown). The similarity of patterns of NOE interactions and ^1H - ^{15}N correlations between the full length and truncated α CTD below 28°C strongly suggests similar secondary and tertiary structures for residues 231–325 in both full length and truncated domain.

α CTD interaction with NikR transcriptional regulator bound to DNA

To gain more insight into potential protein–protein interactions, we produced electrostatic surface potentials for both the *H. pylori* and *E. coli* α CTDs

Figure 4. Comparison of full-length and truncated forms (α CTD325) of α CTD. A: Thermal denaturation by measurement of circular dichroism at 222 nm. The calculated T_m values are 54°C for the full-length α CTD (red) and 32°C for α CTD325 (black). The solid lines show nonlinear fits to the data points. B: Superimposed ^1H - ^{15}N HSQC spectra of *H. pylori* α CTD (red) and α CTD325 (black).

(Fig. 5). Overall, the two proteins have very similar theoretical isoelectric points (5.31 for *H. pylori* vs. 5.41 for *E. coli*); however, there are significant differences in surface charge distribution. The surface of the C-terminal segments in the *H. pylori* α CTD is much more charged than that of *E. coli*, which is fairly hydrophobic. We attempted to model the *H. pylori* α CTD into known α CTD complexes from *E. coli* (Supporting Information Fig. S2). Our modeling suggests that the *H. pylori* α CTD interacts with cognate transcription factors using a larger and more charged interface than *E. coli* α CTD.

We chose to test the nickel dependent transcription factor NikR from *H. pylori* to determine whether it can interact directly with the α CTD during recruitment of the RNA polymerase.^{21–23} NikR is a member of the ribbon-helix-helix family of transcriptional regulators that binds specific sequences of DNA in response to elevated concentrations of nickel.^{20–22} In *E. coli*, NikR acts only as a repressor, but in *H. pylori*, it has been found to be a repressor or an activator depending on the promoter. NikR activates the transcription of *H. pylori* urease, whose promoter has been studied in detail. Nickel-bound NikR recognizes two weakly conserved 6 bp imper-

fect inverted repeat half-sites separated by a 15 bp spacer. All known NikR promoters are A/T-rich and thus contain potential sites recognized by the α CTD of the RNA polymerase. The binding site of NikR within the urease promoter has been mapped by footprinting analysis to between 91 bp and 56 bp upstream of the transcriptional start site (TSS),²⁴ and its imperfect palindromic half-sites have been identified.²¹ This region is approximately two turns of the DNA helix upstream of the putative σ factor binding site, making a direct interaction between NikR and σ unlikely. On the other hand, α CTD often binds upstream of the σ factor at A/T-rich sequences, several of which can be found throughout the NikR footprinting region, making a NikR and α CTD interaction possible (Fig. 6).

We have investigated α CTD and α CTD325 binding to the *ureA*, *nixA* and *frpB4* promoters in the presence and absence of NikR (Supporting Information Fig. S1). We also tested their binding to a modified *ureA* promoter, which had the A/T-rich sequences flanking and/or separating the inverted repeats replaced by G/C-rich sequences, *ureA-32mod*. The molecular mass of a tetrameric NikR complexed with 49 bp DNA (total molecular mass 101 kDa) makes the high resolution NMR studies of this complex rather impractical. Instead, we used polyacrylamide gel electrophoretic mobility shift assays (EMSA) to estimate the affinity of the observed interactions. The EMSA approach has been used extensively to analyze NikR interactions with its promoters.^{21,25,26}

Without the presence of DNA, in a native polyacrylamide gel electrophoresis, NikR, and α CTD proteins showed no observable interaction up to the highest concentrations tested, 12 μM (data not shown). Free α CTD and α CTD325 showed low affinity to the wild type 49-bp *ureA* promoter sequence (*ureA-49*, Supporting Information Fig. S1B(i)) with an estimated 2–3 μM dissociation constant (K_d). NikR binding of *ureA-49* showed a K_d of 9 nM, (based on the protein monomer concentration, or 2.25 nM if tetramerization of NikR is factored in, (Supporting Information Fig. S1B(ii)). The complex of NikR:*ureA-49* showed binding of α CTD with K_d below 1 μM , (Supporting Information Fig. S1B(ii)). The α CTD325 showed somewhat weaker affinity for the NikR:*ureA-49*, with the $K_d \sim 1.5 \mu\text{M}$, (Supporting Information Fig. S1B(iii)). This difference may in part be the effect of lower isoelectric point of α CTD325 (theoretical pI = 4.79). Similar affinities for NikR and α CTD were observed in the presence of a shorter 32 bp *ureA-32* promoter lacking the flanking A/T-rich sequences but still possessing an A/T-rich region in the spacer between inverted repeats, [Supporting Information Fig. S1B(iv and v)]. However, the *ureA-32mod* DNA duplex, lacking any A/T-rich potential binding sites, either flanking

Figure 5. Electrostatic surface potentials calculated using Delphi. A: *H. pylori* α CTD. B: *E. coli* α CTD taken from PDB entry 1LB2. Blue indicates positive charge; red indicates negative charge. The structurally divergent C-terminal fragments are indicated by black coloring of the backbone ribbons. The C-terminal helix from *H. pylori* α CTD presents a larger and more charged surface to the solvent than the ordered hydrophobic terminus of the *E. coli* α CTD.

inverted repeats or in the spacer between them, showed weaker binding to NikR with a $K_d \sim 250$ nM, [Supporting Information Fig. S1B(vi and vii)]. Also, we were unable to observe any α CTD binding to NikR:*ureA-32mod*, indicating $K_d > 2$ μ M, [Supporting Information Fig. S1B(vi and vii)]. We also tested NikR:*nixA-49* ($K_d \sim 20$ nM) binding of α CTD and α CTD325, where both showed K_d in 1–2 μ M range, [Supporting Information Fig. S1B(viii)]. Similarly, NikR:*frpB4-49* ($K_d \sim 10$ nM) complex bound α CTD with a K_d in 1–2 μ M range, [Supporting Information Fig. S1B(ix)].

Overall, the EMSA experiments with α CTD showed that it interacts cooperatively with the NikR and the DNA of NikR promoters provided at least the spacer between inverted repeats contains an A/T rich sequence.

Discussion

The relationship between infection and transcription in *H. pylori* has been interrogated in a large number of studies²⁷; however, very few of them reached the structural level. There are no reports of specific

interactions between *H. pylori* transcription factors and α CTD or other subunits of RNA polymerase. The tertiary structure of the conserved core domain of *H. pylori* α CTD is similar to that of *E. coli* α CTD (PDB accession code 3K4G), with a backbone RMSD of 1 Å. The backbone of *H. pylori* α CTD also superimposes closely on the α CTD from *Bacillus subtilis*²⁸ over the core domain (residues 257–318), with a backbone RMSD of 1 Å [Fig. 2(b)]. The C-terminal ends of the *H. pylori* and *B. subtilis* α CTDs, however, are even more divergent than *H. pylori* and *E. coli*. The *B. subtilis* α CTD terminus does not fold back to make hydrophobic contacts with the N-terminal loop or the beginning of helix 3. This can be explained by the fact that the *B. subtilis* α CTD helix 3 is more hydrophilic than that of *H. pylori* α CTD, and *B. subtilis* E282 forms a salt bridge to K300 of helix 4, covering L310. In *H. pylori*, Y281 at the N-terminal end of helix 3 makes contact with L334 and L337 from helix 5. In *B. subtilis*, the corresponding residue is a threonine (T273). Hydrophobic residues L287 and M288 at the end of helix 3 in *H. pylori* correspond to residues N279 and K280 in *B. subtilis*.

Figure 6. Urease promoter region showing potential α CTD binding sites. The -10 and -35 σ factor binding elements are highlighted in red. The NikR footprint is indicated with two half-sites shown in blue. Two α CTDs are shown bound to DNA and interacting with NikR; however, we do not know how many or at which sites the α CTDs actually bind.

Solution NMR experiments showed that α CTD binds DNA with only modest affinity. The EMSA experiments with α CTD and α CTD325 showed that they interact cooperatively with the NikR:promoter DNA binary complexes provided that at least the spacer between NikR-binding inverted repeats contains an A/T-rich sequence. To what extent this cooperativity involves direct physical α CTD:DNA and/or α CTD:NikR interactions remains to be investigated. Both DNA and NikR change their conformations significantly upon binary complex formation, and such induced conformations may bind α CTD with higher affinity than any single isolated component. It is also worth noting that *H. pylori* NikR-binding inverted repeats are A/T-rich and may also serve as α CTD binding sites. Based on the crystal structure of *E. coli* NikR:DNA complex, both inverted repeats in a bent promoter are occupied by NikR, however NikR does not wrap fully around the DNA. The bending ability of DNA depends on its sequence, and we observed large changes in the NikR binding affinity when the spacer was made more G/C-rich in the *ureA-32mod* DNA. However, even when this DNA was fully complexed at high NikR concentration, we could not observe α CTD binding, suggesting either α CTD recognizes A/T-rich DNA directly as it forms the ternary complex or that only a particular DNA-bound conformation of NikR is capable of binding α CTD. The underlying complex allostery of NikR with respect to nickel and other ion concentrations, as well as variations in promoter sequences fine tune NikR regulatory function in *H. pylori*, and the interactions with α CTD can serve as a physical connection to the RNA polymerase. A significant result of the α CTD binding experiments was that NikR-occupied promoters from all three tested *H. pylori* genes showed very similar affinities for α CTD. The centers of 32-bp core NikR binding sequences in *nixA*, *frpB4*, and *ureA* promoters are positioned at -21, -1, and -72 bp relative to the TSS, respectively. In *H. pylori*, two of these genes *nixA* and *frpB4* are repressed by NikR, whereas *ureA* is activated, suggesting that the promoter positioning relative to the TSS differentiates between activating and repressing functionality of the NikR. Only on the *ureA* promoter could α CTDs bind to NikR and position the polymerase upstream of the TSS. By contrast, for both *nixA* and *frpB4*, NikR would block the polymerase from binding the promoter. Although this conclusion is consistent with known activities and positioning of NikR promoters, we cannot exclude that NikR also interacts with additional domains of RNA polymerase or other mediating proteins.

The C-terminally truncated version (α CTD325) of the α CTD showed similar affinity to NikR:DNA complexes compared with the full size α CTD, indicating that the C-terminal residues 326–344 are not

critical in this ternary complex. The C-terminal deletions of *B. subtilis* α CTD also retained certain functions, while other were lost.¹¹ The EMSA binding experiments are typically performed at 5°C, where truncated domains can retain folded structure similar to the intact domain. However, the lack of 5th helix leucine residues' (L330, L334, L337) contributions to the hydrophobic core destabilizes this domain. Consequently, we do not expect *H. pylori* α CTD325 mutant to be viable, but less severe truncations may produce temperature sensitive mutants.

We used modeling to explore the function of the unique C-terminal segment of *H. pylori* α CTD (Fig. 6, Supporting Information Fig. S4). Supporting Information Fig. S3 shows the faces of α CTD and NikR involved in such possible interaction. A glutamate at position 291 in α CTD, which is highly conserved only in *Helicobacter* species, may interact with a small, positively charged patch on NikR, formed by K48. This lysine (K48) is also a residue that interacts with E47 within the N-terminal domain of NikR, and that interaction increases affinity of NikR for *ureA* promoters, but not *nixA* promoters.²⁶ *H. pylori* α CTD L322 in the flexible loop preceding helix 5 can interact with a hydrophobic patch on NikR that includes V50. Given the sequences and spacing of regulatory elements within the urease promoter (*ureA*) of *H. pylori*, we believe that this interaction can occur *in vivo*. Both proteins from *H. pylori* have significantly more charged surfaces than their *E. coli* counterparts. For example, the residues at positions 290–292 vary significantly among ϵ -proteobacteria (Supporting Information Fig. S4). All known genus *Helicobacter* sequences have EEE at this position, except for *Helicobacter hepaticus*, which has ENE. We propose that the middle glutamate interacts with NikR, and its conservation among *Helicobacter* species supports this suggestion. Comparative studies of the regulation of urease activity in *H. pylori* and *H. hepaticus* have shown that elevated nickel concentrations in *H. hepaticus* do not result in higher transcript levels of urease, unlike in *H. pylori*.²⁹ In *E. coli*, which does not use NikR for positive regulation, the α CTD contains a valine at this middle position.

The α CTD from *H. pylori* and other ϵ -proteobacteria features a fifth helix with a well-conserved pattern of hydrophobic and charged residues, even though these bacteria have some of the most rapidly evolving genomes. *H. pylori* has one of the highest rates of mutation due to its apparent lack of several DNA repair enzymes³⁰ and also one of the highest rates of exogenous DNA uptake and homologous recombination of all bacteria.³¹ The fact that the features of α CTDs are highly conserved on both the tertiary and primary levels strongly supports their functional importance. Absent this, we would expect to see greater sequence divergence.³²

We suggest that the fifth helix forms a required interaction site for other transcriptional regulators. An intriguing possibility is that the highly positively charged surface of the *H. pylori* α CTD is another site regulated by the starvation-signaling polyphosphate, using a mechanism analogous to, or even synergistic with, the binding of σ^{80} subunit of RNA polymerase.³³

Determining the details of transcriptional regulation in *H. pylori* is important because of the possibility of developing compounds to specifically target *H. pylori*. RNA polymerase is a promising target for antibiotic drug development. Significant structural differences in the essential α -subunit CTD open up the possibility of targeting specific bacteria with novel antibiotics.

Materials and Methods

Protein expression and purification

Residues 231–344 encoded by the *rpoA* gene from *H. pylori* strain J99 (JHP1213) were PCR amplified from genomic DNA using the following oligonucleotide primers: forward- 5'-gacggatccctgggcgtttttggcgaaag, reverse- 5'-gacggataccgtttgtgtctcatcagtcgttacctcc. The PCR product was cloned into a modified pET vector that introduced an N-terminal, 12-residue His₆ tag (MRGSHHHHHHGS). Transformed *E. coli* BL21 (DE3) cells were grown in LB media to OD₆₀₀=1 and induced with 0.4 mM IPTG for 3 h. Cells were collected and lysed by sonication in binding buffer (20 mM Tris-HCl, 5 mM imidazole, 0.5M NaCl, 8M urea, pH 7.9). Soluble proteins were loaded onto Ni-NTA resin, washed (20 mM Tris-HCl, 30 mM imidazole, 0.5M NaCl, 8M urea, pH 7.9), and eluted (20 mM Tris-HCl, 0.5M NaCl, 100 mM EDTA, 8M urea, pH 7.9). Refolding was achieved by extensive dialysis against NMR sample buffer (25 mM KH₂PO₄, 225 mM KCl, 1 mM TCEP, pH 7.3). TCEP (Tris(2-carboxyethyl) phosphine) was used to keep the two cysteines in a reduced state. ¹⁵N and ¹³C isotope-labeled samples were prepared by growing cells in M9 minimal media with ¹⁵NH₄Cl and/or ¹³C-u-glucose (CIL, Andover, MA) as sole sources of nitrogen and/or carbon.

The untagged native sequence of *H. pylori* NikR (HP1338, *H. pylori* strain 26695) was expressed in *E. coli* BL21(DE3) using the pEB116 plasmid,²⁵ a generous gift from Dr. Peter T. Chivers. Cells were grown and induced as described above for the α CTD expression. Collected cells were lysed by sonication and loaded on Ni-NTA column in 50 mM Tris, 0.2M NaCl, 5 mM imidazole, 50 μ M NiSO₄, pH 7.5. The column was washed with lysis buffer, and NikR was eluted using 5–500 mM imidazole gradient, with the majority of NikR eluting below 50 mM imidazole. Protein containing fractions were pooled and dialyzed to the low salt anion exchange chromatogra-

phy buffer, 50 mM Tris, 0.1M NaCl, 50 μ M NiSO₄, pH 8.0. NikR was eluted from Q Sepharose XL column using 0.1 to 1.0M NaCl gradient, with the majority of NikR eluting around 0.3M NaCl. Protein was concentrated and exchanged into EMSA buffer by pressure ultrafiltration.

NMR experiments

NMR experiments were performed on Bruker Avance 600 and 800 MHz spectrometers equipped with cryoprobes. Samples were prepared at ~1 mM protein concentration in 25 mM KH₂PO₄, 225 mM KCl, 1 mM TCEP, pH 7.3, and placed in 3 mm NMR tubes to reduce the amount of salt in the detection volume. All experiments used for structural studies of α CTD were performed at 25°C. Singly labeled ¹⁵N samples were used to acquire 2D HSQC experiments. Doubly labeled ¹⁵N, ¹³C samples were used to acquire 3D CBCANH, CCCONH, HCCCONH, and HCCH-TOCSY experiments³⁴ used for backbone and sidechain assignments. 3D ¹⁵N and ¹³C NOESY-HSQC experiments³⁴ were used for assigning NOESY crosspeaks used in structure calculations.

A heteronuclear ¹H-¹⁵N NOE experiment was recorded to estimate the backbone dynamics. Peak assignments from a ¹H-¹⁵N HSQC were transferred to the spectra with and without saturation, and for each residue, the ratio of the intensities of the peak in the two spectra was taken as a measure of the steady-state heteronuclear NOE. Residues with peaks that overlapped in the spectrum (I239, Y244, D253, K255, D256, L293, K301, Y304, E309, D344) were excluded.

Structure determination

NMR data were processed using TOPSPIN (Bruker, Billerica, MA) and analyzed using SPARKY (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). Structure calculations were performed using CYANA³⁵ version 2.1 with 25,000 steps for each structure. Distance restraints were calibrated automatically using CYANA routines. Hydrogen bond restraints were included only in later stages of calculations when they could be identified in a majority of structures. A total of 119 backbone ϕ and ψ dihedral angle restraints calculated using TALOS³⁶ were employed in the calculations. For the final round of calculations, 500 structures were calculated in CYANA, and the 50 with the lowest target function were energy-minimized using AMBER 9³⁷ with 3,000 steps of steepest descent energy minimization. Energy-minimized structures were analyzed with AQUA and PROCHECK-NMR.³⁸ The final ensemble consists of the 15 structures with the lowest energies and the best non-bonded backbone geometry. Chimera³⁹ was used for visualization of structures and for producing figures. Delphi⁴⁰ was used for

calculating the electrostatic surface potentials. The *E. coli* α CTD model used for electrostatic surface potential calculations came from PDB ID 1LB2. The PDB accession code for the ensemble of *Helicobacter pylori* α CTD is 2MAX and the BMRB accession number is 19380.

Determination of the stability of the secondary structure by circular dichroism

Samples of α CTD and α CTD325 were prepared at 50 μ M in 0.15M NaCl, 25 mM NaH₂PO₄, pH 7.5. Thermal melting experiments were performed on Applied Photophysics Chirascan CD spectrometer in 1 mm cuvette. The temperature was incremented by 1°C per 2 min with 280–200 nm wavelength scan at each temperature point between 5 and 90°C. The data were fitted to 4 or 6 parameter equations²⁰ using XLSTAT nonlinear regression routine.

Electrophoretic mobility shift assays

The binding reactions were carried out in 10 mM Tris, 100 mM KCl, 3 mM MgCl₂, 50 μ M NiSO₄, pH 7.9. DNA fragments were purchased (IDT, USA) as duplexes at 1 μ mole scale and used without further purification. DNA fragments (2.5 ng per reaction, $\sim$ 8 nM) were incubated with proteins in 10 μ L at 22°C for 30 min in 96-well plates. Because the operation of electrodes of an electrophoresis apparatus is rapidly degraded by Ni²⁺ ions, we ran the apparatus with reversed polarity of electrodes using dummy polyacrylamide gel in 2 mM EDTA containing Ni²⁺-free buffer for 2 h before using the apparatus. Electrophoresis of binding reactions was performed at 5°C using 9- or 18-well 14% polyacrylamide gels (10 $\times$ 8 cm, acrylamide:bis-acrylamide 29:1) polymerized in and resolved for 2 h with a buffer containing 50 mM Tris, 25 mM boric acid, and 50 μ M NiSO₄, pH 8.45. Before loading the samples, gels were prerun for 10 min 120 V, and the electrophoresis buffers were replaced. Immediately before loading, 10 μ L of 30% glycerol in binding buffer was added to each sample, mixed and the 20- μ L samples were loaded simultaneously onto a running gel (120 V) using an 8-channel pipettor. After electrophoresis, gels were fixed for 20 min, at 5°C in 10% acetic acid, 40% ethanol, and stained for 1 h with GRGreen nucleic acid stain diluted 20,000 times (GRGreen 10,000 stock, Excellgen). Gels were photographed in a dark-box from a 25 cm distance through a 590 nm orange bandpass filter using a digital video-camera (Sentech 152USB, 20 FPS, 1/2" sensor 1360x1024, 1:1.4/12.5mm Fujinon lens) and a blue light transilluminator (Dark Reader, Clare Chemical, USA) with exposure times of 30 s. Apparent affinities were estimated based on protein concentrations present in reactions showing equal amounts of unshifted and shifted DNA.

References

1. Mathew R, Chatterji D (2006) The evolving story of the omega subunit of bacterial RNA polymerase. *Trends Microbiol* 14:450–455.
2. Dailidienė D, Tan S, Ogura K, Zhang M, Lee AH, Severinov K, Berg DE (2007) Urea sensitization caused by separation of *Helicobacter pylori* RNA polymerase beta and beta' subunits. *Helicobacter* 12:103–111.
3. Solnick JV, Hansen LM, Syvanen M (1997) The major sigma factor (RpoD) from *Helicobacter pylori* and other gram-negative bacteria shows an enhanced rate of divergence. *J Bacteriol* 179:6196–6200.
4. Jeon YH, Yamazaki T, Otomo T, Ishihama A, Kyogoku Y (1997) Flexible linker in the RNA polymerase alpha subunit facilitates the independent motion of the C-terminal activator contact domain. *J Mol Biol* 267: 953–962.
5. Hayward RS, Igarashi K, Ishihama A (1991) Functional specialization within the alpha-subunit of *Escherichia coli* RNA polymerase. *J Mol Biol* 221:23–29.
6. Mencia M, Monsalve M, Rojo F, Salas M (1996) Transcription activation by phage phi29 protein p4 is mediated by interaction with the alpha subunit of *Bacillus subtilis* RNA polymerase. *Proc Natl Acad Sci USA* 93: 6616–6620.
7. Mencia M, Monsalve M, Rojo F, Salas M (1998) Substitution of the C-terminal domain of the *Escherichia coli* RNA polymerase alpha subunit by that from *Bacillus subtilis* makes the enzyme responsive to a *Bacillus subtilis* transcriptional activator. *J Mol Biol* 275:177–185.
8. Benoff B, Yang H, Lawson CL, Parkinson G, Liu J, Blatter E, Ebright YW, Berman HM, Ebright RH (2002) Structural basis of transcription activation: the CAP-alpha CTD-DNA complex. *Science* 297:1562–1566.
9. McLeod SM, Aiyar SE, Gourse RL, Johnson RC (2002) The C-terminal domains of the RNA polymerase alpha subunits: contact site with Fis and localization during co-activation with CRP at the *Escherichia coli* proP P2 promoter. *J Mol Biol* 316:517–529.
10. Dangi B, Gronenborn AM, Rosner JL, Martin RG (2004) Versatility of the carboxy-terminal domain of the alpha subunit of RNA polymerase in transcriptional activation: use of the DNA contact site as a protein contact site for MarA. *Mol Microbiol* 54:45–59.
11. Calles B, Monsalve M, Rojo F, Salas M (2001) A mutation in the C-terminal domain of the RNA polymerase alpha subunit that destabilizes the open complexes formed at the phage phi 29 late A3 promoter. *J Mol Biol* 307:487–497.
12. Schweimer K, Prasch S, Sujatha PS, Bubunenkov M, Gottesman ME, Rosch P (2011) NusA interaction with the alpha subunit of *E. coli* RNA polymerase is via the UP element site and releases autoinhibition. *Structure* 19:945–954.
13. Quinones M, Kimsey HH, Ross W, Gourse RL, Waldor MK (2006) LexA represses CTXphi transcription by blocking access of the alpha C-terminal domain of RNA polymerase to promoter DNA. *J Biol Chem* 281:39407–39412.
14. Yamamoto K, Ishihama A (2003) Two different modes of transcription repression of the *Escherichia coli* acetate operon by IclR. *Mol Microbiol* 47:183–194.
15. Finn RD, Tate J, Mistry J, Coghill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A (2008) The Pfam protein families database. *Nucleic Acids Res* 36:D281–D288.
16. Jeon YH, Negishi T, Shirakawa M, Yamazaki T, Fujita N, Ishihama A, Kyogoku Y (1995) Solution structure of

- the activator contact domain of the RNA polymerase alpha subunit. *Science* 270:1495–1497.
17. Lara-Gonzalez S, Birktoft JJ, Lawson CL (2010) Structure of the *Escherichia coli* RNA polymerase alpha subunit C-terminal domain. *Acta Cryst D* 66:806–812.
 18. Beier D, Spohn G, Rappuoli R, Scarlato V (1998) Functional analysis of the *Helicobacter pylori* principal sigma subunit of RNA polymerase reveals that the spacer region is important for efficient transcription. *Mol Microbiol* 30:121–134.
 19. Shirai M, Fujinaga R, Akada JK, Nakazawa T (1999) Activation of *Helicobacter pylori* ureA promoter by a hybrid *Escherichia coli*-*H. pylori* rpoD gene in *E. coli*. *Gene* 239:351–359.
 20. Greenfield NJ (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat Protoc* 1:2527–2535.
 21. Dosanjh NS, West AL, Michel SL (2009) *Helicobacter pylori* NikR's interaction with DNA: a two-tiered mode of recognition. *Biochemistry* 48:527–536.
 22. Ernst FD, Stoof J, Horrevoets WM, Kuipers EJ, Kusters JG, van Vliet AH (2006) NikR mediates nickel-responsive transcriptional repression of the *Helicobacter pylori* outer membrane proteins FecA3 (HP1400) and FrpB4 (HP1512). *Infect Immun* 74:6821–6828.
 23. van Vliet AH, Kuipers EJ, Stoof J, Poppelaars SW, Kusters JG (2004) Acid-responsive gene induction of ammonia-producing enzymes in *Helicobacter pylori* is mediated via a metal-responsive repressor cascade. *Infect Immun* 72:766–773.
 24. Ernst FD, Kuipers EJ, Heijens A, Sarwari R, Stoof J, Penn CW, Kusters JG, van Vliet AH (2005) The nickel-responsive regulator NikR controls activation and repression of gene transcription in *Helicobacter pylori*. *Infect Immun* 73:7252–7258.
 25. Benanti EL, Chivers PT (2007) The N-terminal arm of the *Helicobacter pylori* Ni²⁺-dependent transcription factor NikR is required for specific DNA binding. *J Biol Chem* 282:20365–20375.
 26. Benanti EL, Chivers PT (2011) *Helicobacter pylori* NikR protein exhibits distinct conformations when bound to different promoters. *J Biol Chem* 286:15728–15737.
 27. Danielli A, Amore G, Scarlato V (2010) Built shallow to maintain homeostasis and persistent infection: insight into the transcriptional regulatory network of the gastric human pathogen *Helicobacter pylori*. *PLoS Pathog* 6:e1000938.
 28. Newberry KJ, Nakano S, Zuber P, Brennan RG (2005) Crystal structure of the *Bacillus subtilis* anti-alpha, global transcriptional regulator, Spx, in complex with the alpha C-terminal domain of RNA polymerase. *Proc Natl Acad Sci USA* 102:15839–15844.
 29. Belzer C, Stoof J, Beckwith CS, Kuipers EJ, Kusters JG, van Vliet AH (2005) Differential regulation of urease activity in *Helicobacter hepaticus* and *Helicobacter pylori*. *Microbiology* 151:3989–3995.
 30. Wang G, Humayun MZ, Taylor DE (1999) Mutation as an origin of genetic variability in *Helicobacter pylori*. *Trends Microbiol* 7:488–493.
 31. Suerbaum S, Josenhans C (2007) *Helicobacter pylori* evolution and phenotypic diversification in a changing host. *Nat Rev Microbiol* 5:441–452.
 32. Swapna LS, Rekha N, Srinivasan N (2012) Accommodation of profound sequence differences at the interfaces of eubacterial RNA polymerase multi-protein assembly. *Bioinformation* 8:6–12.
 33. Yang ZX, Zhou YN, Yang Y, Jin DJ (2010) Polyphosphate binds to the principal sigma factor of RNA polymerase during starvation response in *Helicobacter pylori*. *Mol Microbiol* 77:618–627.
 34. Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog Nucl Magn Reson Spectrosc* 34:93–158.
 35. Guntert P, Mumenthaler C, Wuthrich K (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol* 273:283–298.
 36. Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* 13:289–302.
 37. Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. *J Comp Chem* 26:1668–1688.
 38. Laskowski RA, Rullmann JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8:477–486.
 39. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comp Chem* 25:1605–1612.
 40. Gilson MK, Sharp K, Honig B (1987) Calculating the electrostatic potential of molecules in solution: method and error assessment. *J Comp Chem* 9:327–335.