Identification of a candidate CO2 sensor from Helicobacter pylori

CHARLOTTE GRACE ROUGHTON

Identification of a candidate CO_2 sensor from *Helicobacter pylori*

Charlotte Grace Roughton

Thesis submitted for the degree of Master of Science by Research

Department of Biosciences

Durham University

2021

Contents

List of Figures 3
List of Tables 4
Abstract 5
Abbreviations 7
Statement of Copyright 8
Acknowledgements 9
Chapter 1: Introduction 11
1.1 CO_2 in Biology 11
1.2 Bacterial CO_2 Sensing Mechanisms 15
1.3 CO_2 in *H. pylori* Physiology 24
1.4 *H. pylori* Nickel-Responsive Regulator 27
Chapter 2: Materials and Methods 33
2.1 Reagents and Buffers 33
2.2 *Hp*NikR Expression 34
2.3 *Hp*NikR Purification 34
2.4 SDS-PAGE 35
2.5 DNA Duplex Preparation 36
2.6 *Hp*NikR Carbamate Trapping 37
2.7 Mass Spectrometry 38
2.8 Bioinformatics 38
2.9 Fluorescence Anisotropy 39
2.10 Electrophoretic Mobility Shift Assays 39
Chapter 3: Results 42
3.1 Lys6 is a site of carbamylation 42
3.2 Lys6 is completely conserved across all *H. pylori* NikR homologues 45
3.3 Lys6 conservation is lost among *Helicobacter spp.* NikR homologues 51
3.4 Wild type *Hp*NikR exhibits a decrease in affinity for *ureA* and *nixA* promoters in response to CO_2/HCO_3^- 56
3.5 Mutation of Lys6 to Arg reduces *Hp*NikR sensitivity to CO_2 ... 62

Chapter 4: Discussion..... 65
Chapter 5: References 83

List of Figures

Figure 1. Simplified organisation of bacterial environmental sensing mechanisms 16
Figure 2. Mechanism of protein carbamate formation 17
Figure 3. Proposed trapping mechanism of a protein carbamate with TEO 21
Figure 4. *Hp*NikR tetramer complexed with DNA 29
Figure 5. EMSA quantification process 41
Figure 6. Structure of y_6 and y_7 ions 43
Figure 7. ESI-MS showing the trapped carbamate at Lys6 44
Figure 8. Sequence logo for *Hp*NikR homologues 48
Figure 9. Sequence logo for *Helicobacter spp.* NikR homologues . 54
Figure 10. Change in anisotropy upon the addition of wild type *Hp*NikR to promoter fragments 57
Figure 11. EMSAs with wild type *Hp*NikR, 1.5 mM $MgCl_2$, and 25 mM $NaHCO_3$ 58
Figure 12. Quantified EMSAs (Figure 11) 58
Figure 13. EMSAs with wild type *Hp*NikR and 750 µM $MgCl_2$ 60
Figure 14. EMSAs with wild type *Hp*NikR, 750 µM $MgCl_2$, and 25 mM $NaHCO_3$ 60
Figure 15. Quantified EMSAs (Figure 13, Figure 14) 61
Figure 16. EMSAs with Lys6Arg *Hp*NikR, 750 µM $MgCl_2$, and 25 mM $NaHCO_3$ 63
Figure 17. Quantified EMSAs (Figure 16) 63
Figure 18. Predicted *in vivo* regulation by *Hp*NikR under low CO_2 . 71
Figure 19. Predicted *in vivo* regulation by *Hp*NikR under high CO_2 72
Figure 20. Non-cooperativity versus positive cooperativity 73
Figure 21. Structure of Flu-dT 75

List of Tables

Table 1. Buffers used for *Hp*NikR purification and DNA binding assays 33

Table 2. Oligonucleotide sequences used in DNA duplex preparation 36

Table 3. Alignment of *Hp*NikR N-terminal arm sequences 46

Table 4. Conservation scores of N-terminal arm residues 48

Table 5. Alignment of *Helicobacter spp.* N-terminal arm sequences 51

Table 6. Apparent wild type *Hp*NikR binding affinities with 1.5 mM $MgCl_2$ 59

Table 7. Apparent wild type and Lys6Arg *Hp*NikR binding affinities with 750 μM $MgCl_2$ 61

Abstract

Carbon dioxide (CO_2) is a ubiquitous metabolite. It is fundamental in pathogenesis, involved in the activation of signalling pathways. Several species of human pathogens harbour CO_2 sensing mechanisms that allow for the coupling of virulence factor expression to host CO_2 levels, facilitating colonisation and persistence. *Helicobacter pylori* is an example of a human pathogen for which CO_2 is especially important in its biology – requiring CO_2 for growth and generating CO_2 as a result of urease activity. However, what remains unclear are the details of the mechanisms underpinning CO_2 sensing and the identity of CO_2 sensors.

A candidate mechanism by which CO_2 sensing may occur is the direct modification of regulatory proteins, termed carbamylation, entailing the nucleophilic attack of a neutral N-terminal α-amino group or ε-amino group of Lys onto CO_2, forming a carbamate. While carbamates observed to date, such as those in haemoglobin and RuBisCO, have significant functional roles, their systematic identification has been limited by their lability and the lack of a tool for their direct investigation, until recently.

This thesis sought to assess the validity of the *H. pylori* nickel-responsive regulator (*Hp*NikR) as a CO_2 sensor, using carbamylation as a mechanism to achieve this, given its pleiotropic regulatory activity, the importance of Lys residues in its DNA binding, and its responsiveness to the environmental cues of Ni(II) availability and pH. Carbamylation would affect DNA binding, which in turn would affect gene expression and allow *H. pylori* to adapt to fluctuating CO_2 levels, potentially in accordance with urease activity. Modification of a conserved Lys was detected, DNA binding assays showed this modification

reduced DNA affinity, and this effect was reduced in a mutant insensitive to carbamate formation. This suggests that *Hp*NikR may indeed behave as a CO_2 sensor, potentially coupling gene transcription to this additional environmental cue.

Abbreviations

ESI-MS	Electrospray ionisation mass spectrometry
GuHCl	Guanidine hydrochloride
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulphonic acid
HPLC	High performance liquid chromatography
HPSF	High purity salt free
IPTG	Isopropyl β-d-1-thiogalactopyranoside
MALDI-MS	Matrix-assisted laser desorption/ionization mass spectrometry
MALT	Mucosa-associated lymphoid tissue
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
PMSF	Phenylmethylsulphonyl fluoride
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
S-Trap	Suspension trapping
TEO	Triethyloxonium tetrafluoroborate
Tris	Tris(hydroxymethyl)aminomethane

Statement of Copyright

Acknowledgements

There is a large number of people without whom the completion of this thesis would not have been possible. Firstly, I owe many thanks to my excellent supervisors: Martin Cann and Peter Chivers. Martin was the person who helped plant the initial seed that maybe doing academic research isn't such a wild idea after all. He invited me to continue my studies at Durham University as a postgraduate, and I am grateful for his supervision and the opportunity to undertake such a challenging, albeit interesting, research project. I am thankful for Peter's kind and patient mentorship. He has helped to foster my growth as a scientist, offering wisdom, encouragement, and unwavering faith in me when I lacked it. His contagious enthusiasm for science has inspired me to keep asking questions and searching for answers – unless that question happens to be what kimchi pancakes are made of.

Secondly, I would like to thank Vicki Linthwaite for generously offering her time and expertise to assist with carbamate trapping experiments, analysis of mass spectrometry data, and for patiently answering my incessant questions.

I would also like to thank Steve Chivasa for kindly allowing me to use his Typhoon gel imager, an integral piece of equipment this research would not have been possible without.

I also owe thanks to my thesis committee, Tim Blower and Paul Denny, for their encouragement and insight.

I consider myself incredibly lucky to have been part of such a wonderful community of brilliant researchers over the past year. I would like to thank Sam Firth, Claudia Driscoll, Matthew Boutflower, Jack Bolton, Will Michaels, Lydia

Newton, Joy Paterson, Louisa Stewart, Karrera Djoko, Tessa Young, Andrew Foster, and Deenah Morton for their friendship, support, and for welcoming me with open arms. My time here would not have been nearly as enjoyable were it not for such an inspirational, talented, and kind group of people. They help to foster an environment that all research labs should strive for – one of support, collaboration, and companionship. It truly has been a pleasure to work with and learn from them all – academia is a brighter place with them in it.

Finally, I would like to express my deepest gratitude to my family – Mum, Jess, Jonny, Nana, and Ken. I am thankful for their unfaltering belief in me, for nurturing my love for science, for always supporting me and my decisions, and of course for always humouring me while I try and convince them that “tummy bugs” are definitely incredibly interesting. I would not be where I am today without them.

Chapter 1: Introduction

1.1 CO_2 in Biology

CO_2 is usually regarded as a waste product of oxidative metabolism. However, it has an important role as a stimulus to elicit the activation of several cellular signalling pathways in both eukaryotic and prokaryotic organisms (Cummins et al., 2014). The mechanisms underpinning the ability of organisms to sense and respond to fluctuating CO_2 levels have implications in physiological and pathophysiological processes, though these mechanisms are poorly understood (Cummins et al., 2014; Taylor and Cummins, 2011).

In higher eukaryotic organisms such as mammals, CO_2 plays a vital role in pH homeostasis (Cummins et al., 2020). This arises from its potential to form carbonic acid (H_2CO_3) when in solution, which can then dissociate into bicarbonate ions (HCO_3^-) and protons [Equation 1, (Berg et al., 2002)]. While this series of reversible reactions occurs spontaneously and at a moderate pace in the absence of a catalyst, the hydration of CO_2 is too slow at physiological pH (Khalifah, 1973; Smith and Ferry, 2000). Therefore, in almost all organisms this reversible combination of H_2O and CO_2 is further aided by the highly efficient catalytic activity of the zinc-containing metalloenzyme carbonic anhydrase (Berg et al., 2002; Lindskog, 1997).

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons H^+ + HCO_3^- \qquad (1)$$

Given that CO_2 is primarily transported via the circulatory system from respiring tissues to the lungs dissolved in H_2O and plasma, an increase in CO_2 partial

pressure beyond the physiological range (38 – 42 mm Hg; 5%), termed hypercapnia, results in a right shift in the above equilibria (Patel and Majmundar, 2019). Consequently, as proton concentration increases, pH decreases, and a state of acidosis is introduced (Patel and Majmundar, 2019). On the contrary, a decrease in CO_2 partial pressure below the physiological range, termed hypocapnia, results in the opposite scenario, introducing a state of alkalosis (Patel and Majmundar, 2019). These pathological disturbances in acid-base homeostasis prove highly detrimental to the organism, potentially leading to cell and tissue damage, organ failure, and death (Cummins et al., 2020). States of hyper- and hypocapnia are commonly encountered in both healthy and diseased organisms, with conditions such as sleep apnoea, chronic obstructive pulmonary disease, cystic fibrosis, and subjection to extremes of altitude and exercise all potentially disturbing acid-base homeostasis by virtue of alterations in CO_2 levels beyond their normal physiological range (Cummins et al., 2020). Therefore, the ability to sense and respond to these frequent fluctuations is paramount in maintaining homeostasis (Cummins et al., 2020).

In addition to CO_2 sensing requirements in higher organisms such that internal acid-base homeostasis can be maintained, the formation of external CO_2 concentration gradients due to its exhalation by respiring organisms also play an important role in biology and interactions between species (Luo et al., 2009). Elevated CO_2 concentrations relative to that of low levels of atmospheric CO_2 (0.03%) serve as environmental cues that are utilised by several species of insects and nematodes to initiate innate behavioural responses such as prey seeking (in *Culex quinquefasciatus*) and stress avoidance [in *Caenorhabditis*

elegans and *Drosophila melanogaster* (Bretscher et al., 2008; Turner and Ray, 2009)].

However, CO_2 sensing is not restricted to eukaryotic multicellular organisms – simple unicellular organisms can also sense and respond to altered levels of CO_2 (Cummins et al., 2014). Thus, CO_2 sensing mechanisms, and the corresponding molecular response, are pertinent to further understanding the physiology, and in particular the virulence, of human pathogens. In several species of pathogenic bacteria, the expression of virulence factors is enhanced when grown *in vitro* under 5% CO_2 conditions, which matches the CO_2 levels available in the host (Cummins et al., 2014). By sensing and responding to changes in CO_2 levels, these pathogens can synchronise their gene expression to this environmental cue, facilitating colonisation and persistence within the host. For example, the Gram negative human respiratory tract pathogen *Bordetella bronchiseptica* exhibits elevated transcription of genes encoding virulence factors such as adenylate cyclase toxin, members of the type III secretion system locus, among others, when grown in 5% CO_2 conditions *in vitro* (Hester et al., 2012). In addition to the observed transcriptional changes, phenotypic changes in adherence and cytotoxicity further evidence the ability of *B. bronchiseptica* to sense and respond to changes in CO_2 concentration, tailoring their virulence factor gene expression specifically to different microenvironments (Hester et al., 2012). The zoonotic aetiological agent of Lyme disease, *Borrelia burgdorferi*, also exhibits modulation of its gene expression and protein production in accordance with CO_2 availability (Hyde et al., 2007). This environmental sensing is particularly important given the complexity of the enzootic *B. burgdorferi* life cycle and diverse environmental

conditions encountered inside both its arthropod transmission vector and mammalian reservoir hosts (Hyde et al., 2007). In the presence of 5% CO_2 *in vitro*, translation of lipoprotein genes involved in adherence to the host and genes involved in the oxidative stress response, all important in adaptation to the host environment, is increased (Hyde et al., 2007). This indicates that mammalian physiological CO_2 levels serve as a cue for host adaptation. The aetiological agent of cholera, *Vibrio cholerae*, also exhibits modulation of the expression of its virulence factor genes in response increasing CO_2 levels (Shimamura et al., 1985). Cholera toxin (CT) production is enhanced in cultures grown in 5% CO_2 *in vitro*, with maximal yield occurring in a hypercapnic 10% CO_2 atmosphere (Shimamura et al., 1985). HCO_3^- (Equation 1) has also been found to stimulate virulence gene expression, with CT and toxin-coregulated pilus genes, both major virulence factors required for *V. cholerae* to cause disease, induced in the presence of 0.3 – 0.5% $NaHCO_3$ (Abuaita and Withey, 2009; Iwanaga and Kuyyakanond, 1987). This HCO_3^--associated virulence is attenuated by the addition of the carbonic anhydrase inhibitor ethoxzolamide. This highlights the importance of carbonic anhydrase-mediated interconversion of CO_2 and HCO_3^- and is an additional example of the modulation of pathogen virulence by changing CO_2 levels (Abuaita and Withey, 2009; Iwanaga and Kuyyakanond, 1987)

These examples represent just a few of the many species of pathogenic bacteria that exhibit enhanced growth and virulence in response to the elevated CO_2 levels encountered within the mammalian host environment. As such, there evidently exists widespread CO_2-sensing mechanisms expressed by pathogenic bacteria, allowing them to sense and adapt to the host environment.

1.2 Bacterial CO_2 Sensing Mechanisms

CO_2 sensing is widespread across virtually all species and spans all ranges of organismal complexity, from multicellular animals to the unicellular pathogens that colonise them (Cummins et al., 2014; Taylor and Cummins, 2011). The capacity to sense and respond to fluctuations in CO_2 levels has been an essential evolutionary development, with the reasons for this requirement varying from species to species. In the case of bacterial pathogens, the ability to coordinate gene expression in response to environmental sensing is a crucial feature of their relationship with eukaryotic hosts (Cummins et al., 2014). By coupling the regulation of the production and secretion of toxins and virulence factors to mammalian environmental cues, such as elevated CO_2 levels, successful colonisation of host tissues and evasion of the host immune response can be facilitated (Miller et al., 2007).

Of particular interest with regard to further understanding microbial physiology are the details of the mechanisms underpinning bacterial CO_2 sensing, and the identity of CO_2 sensors. Generally, bacterial sensing of variations in physicochemical parameters is accomplished via two different systems (Figure 1). The first, and more complex of the two, is a two-component system comprised of a membrane-bound sensor histidine kinase and response regulator, usually a transcriptional regulator (Gunn and Richards, 2007; Krell et al. 2010). Typically, a specific signal initiates a phosphorylation cascade beginning with the autophosphorylation of a conserved histidine residue at the sensor kinase followed by transfer of this phosphoryl group to a conserved aspartate residue at the response regulator. The phosphorylated state of the response regulator then confers an alteration in its activity, usually activation,

through modification of its affinity for promoters (Gunn and Richards, 2007; Krell et al., 2010). The second, simpler system entails the diffusion or transport of a small ligand molecule into the cell where it then interacts directly with a regulatory protein to either stimulate or inhibit its activity (Miller et al., 2007; Yang et al., 2008). Though they differ, central to both of these systems is the alteration of DNA transcription such that gene expression can be coordinated with changes in environmental signals.

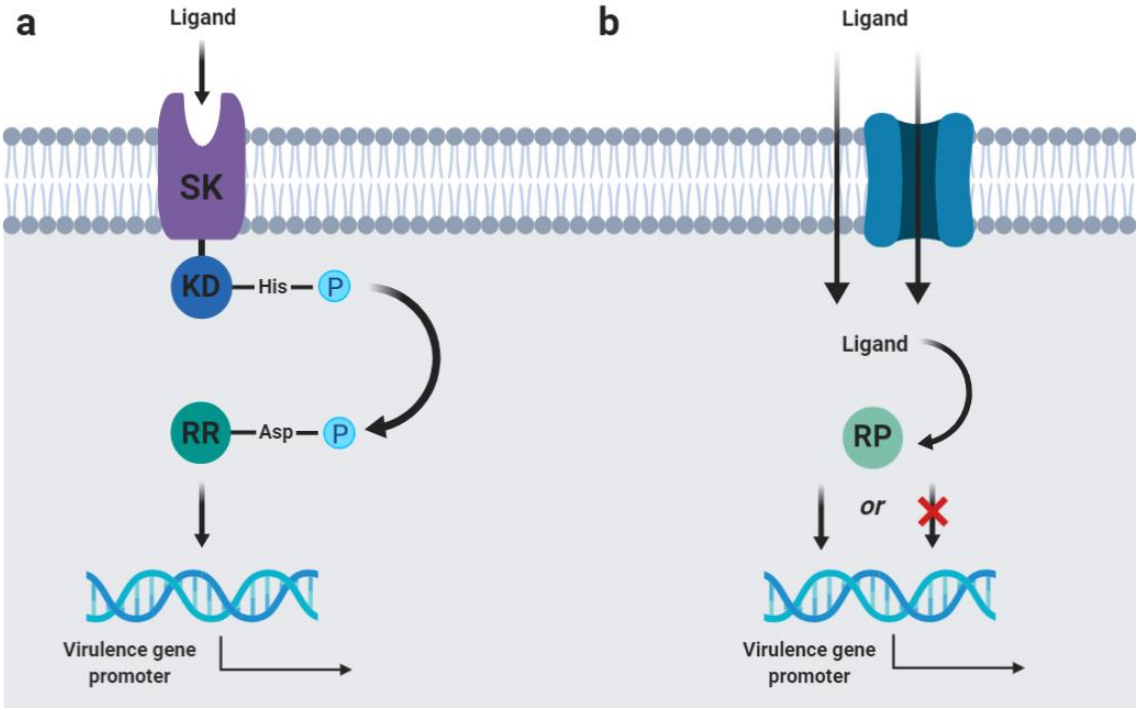


Figure 1. Schematic demonstrating the simplified organisation of the two-component system (a) and direct interaction between ligand molecules and regulatory proteins (b) used by bacteria to sense environmental signals. SK = sensor kinase; KD = kinase domain; RR = response regulator; RP = regulatory protein. Figure created using BioRender.

A candidate mechanism by which environmental CO_2 may interact directly with regulatory proteins to alter their function is the carbamylation of amine groups. This post-translational modification (Figure 2) entails the nucleophilic attack of a

neutral N-terminal α-amino group or the ε-amino group of a Lys side chain onto CO_2, generating an anionic carbamate (Lorimer, 1983).

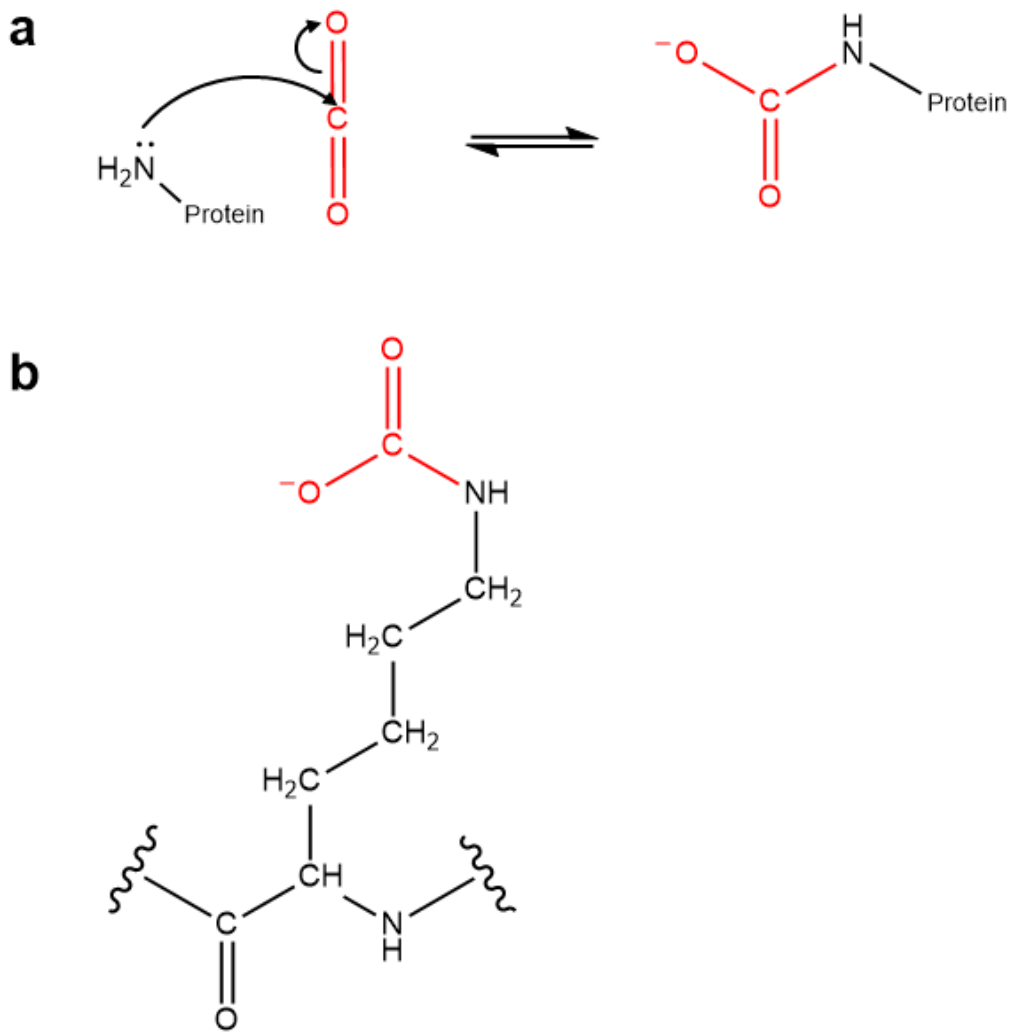


Figure 2. Mechanism of protein carbamate formation (a) (Linthwaite et al., 2018) and structure of a carbamylated Lys ε-amino group (b). Carbamate groups are shown in red. Figure created using ChemDraw 19.1 (Perkin Elmer).

This modification is reversible and non-catalysed, with the rate limiting step of its formation apparently linked to the charge of the amine group, given the observation that an increase in pH increases the amount of carbamate formed (Stadie and O'Brien, 1936). The charge of an amine group is linked to its pK_a, which for Lys side chains has been determined to typically be around 10.5

within proteins (Takayama et al., 2008). This means that at physiological pH (7.4), most Lys side chains are in the basic, protonated state (Abraham et al., 2009). However, the pK_a of individual Lys side chains within proteins can vary significantly depending on the specific environment the residue is in as the polypeptide chain folds and results in several interactions between ionisable groups (Kesvatera et al., 1996). For instance, in calmodulin and calbindin D9k, the pK_a of Lys side chains have been found to vary between 9.3 and 13.2 depending on the specific environment (Abraham et al., 2009; André et al., 2007; Kesvatera et al., 1996; Zhang and Vogel, 1993). Therefore, carbamates are thought to form on amines in what are referred to as “privileged” sites, created when the protein folds in such a way as to create restricted hydration spaces that lead to a reduction in the pK_a of the Lys side chain, and as a result facilitate an uncharged amine at physiological pH (Linthwaite, 2017).

The formation of a carbamate at a Lys side chain introduces a negative charge (Figure 2) and alters its functionality from basic/neutral to acidic (Linthwaite, 2017). This has great potential in modulating protein activity, for instance through altering the involvement of the residue in electrostatic interactions (Zhang and Vogel, 1993). There exist several characterised examples that affirm the importance of carbamates in proteins, playing a diverse range of structural and functional roles (Golemi et al., 2001; Hall et al., 2004; Kilmartin and Rossi-Bernardi, 1971; Lorimer et al., 1976; Morollo et al., 1999; Pearson et al., 1998).

The most widely known examples of confirmed sites of carbamylation include the CO_2-mediated stabilisation of the deoxygenated form of haemoglobin via carbamylation of the α-amino group of the N-terminal valine, and the reliance of the activity of ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) on

the rate-limiting addition of CO_2 to ensure carbamylation of the Lys201 side chain within the catalytic subunit, followed by stabilisation of the carbamate by a Mg(II) ion (Kilmartin and Rossi-Bernardi, 1971; Lorimer et al., 1976; Lorimer and Miziorko, 1980; Matthew et al., 1977).

In addition to animals and plants, several examples of functionally significant carbamates have also been discovered in bacteria, largely through crystallographic analysis. For example, the X-ray crystal structure of alanine racemase from *Bacillus stearothermophilus* revealed the formation of a carbamate at the side chain of Lys129 (Morollo et al., 1999). This carbamate is stabilised through interactions with Arg136, which in turn ensures the correct positioning of Arg136 and is proposed to facilitate catalysis and substrate binding (Morollo et al., 1999). The crystal structure of the *Propionibacterium shermanii* transcarboxylase 5S subunit revealed the presence of a carbamate at the side chain of Lys184 in the active site, which coordinates a Co(II) ion and forms hydrogen bonds with several conserved residues involved in the catalysis of biotin carboxylation (Hall et al., 2004). Lys184Glu and Lys184Ala mutants are catalytically inactive, demonstrating the crucial role of the carbamate in catalysis (Hall et al., 2004). In *Klebsiella aerogenes* urease, X-ray crystallography revealed the presence of a carbamate at the Lys217 side chain that serves to bridge the two active site Ni(II) ions comprising the bi-nuclear metal centre, providing an oxygen ligand to each ion (Jabri et al., 1995). Incubation with Ni(II) and HCO_3^- as a source of CO_2 is required for *in vitro* activation of the apoprotein, with this CO_2-dependence likely due to the carbamylated Lys217 (Yamaguchi and Hausinger, 1997). Lys217Ala, Lys217Glu, and Lys217Cys mutants all demonstrate catalytic inactivity which can be structurally and functionally rescued with the addition of formate, which can successfully

substitute for the carbamate and bridge the Ni(II) centre (Pearson et al., 1998). Therefore, the role of the Lys217 carbamate is one of a metallocentre ligand rather than participating directly in catalysis (Pearson et al., 1998). X-ray crystallography and ^{13}C NMR spectroscopy of the *Pseudomonas aeruginosa* class D OXA-10 β-lactamase revealed the formation of a carbamate at Lys70 within the active site (Golemi et al., 2001; Maveyraud et al., 2000). This carbamate is essential in both stages of hydrolysis of the β-lactam ring, with the enzyme showing a lack of activity in degassed buffer that was restored upon supplementation with $NaHCO_3$ (Golemi et al., 2001). Unusually, this carbamate does not coordinate a metal cation and is instead stabilised through the formation of hydrogen bonds with a H_2O molecule and other residues of the active site (Golemi et al., 2001; Maveyraud et al., 2000). It behaves as a catalytic base, forming a mechanism of catalysis distinct from that of other classes of β-lactamases (Golemi et al., 2001; Maveyraud et al., 2000).

Evidently, carbamylation is a biologically significant process in eukaryotes and prokaryotes, with carbamates identified to date all having clear functional roles (Golemi et al., 2001; Hall et al., 2004; Kilmartin and Rossi-Bernardi, 1971; Lorimer et al., 1976; Morollo et al., 1999; Pearson et al., 1998). However, the systematic identification of carbamates has been limited by the absence of a tool for their direct investigation, attributed to the lability of the modification [Figure 2, (Linthwaite, 2017)]. Therefore, prior studies investigating carbamates have discovered these modifications fortuitously and under conditions that do not accurately reflect those which would be encountered within the cellular environment, such as those required for X-ray crystallography. This means carbamates have remained understudied and are likely far more widespread than previously appreciated, potentially forming the basis of a mechanism of

regulation by which protein activity is coupled to the availability of CO_2. Further elucidating the functional relevance of this unexplored modification is required to gain additional insight into the molecular mechanisms of CO_2 sensing within biological systems.

Recently, a novel methodology was developed that uses chemical trapping to irreversibly modify carbamates and allow for their identification under physiologically relevant conditions (Linthwaite et al., 2018). This methodology uses triethyloxonium (TEO) tetrafluoroborate, a water-soluble electrophilic alkylating agent with a favourable experimental half-life of 7.4 minutes (Linthwaite, 2017). The nucleophilic attack of the carbamate upon TEO results in the transfer of an ethyl group to the modification, chemically trapping it (Figure 3).

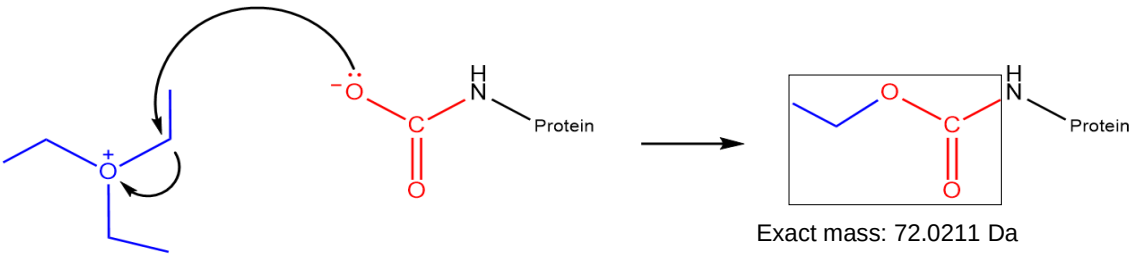


Figure 3. Proposed trapping mechanism of a protein carbamate (red) with TEO (blue), including the exact mass of a carbamate plus an ethyl group (Linthwaite, 2017). Figure generated using ChemDraw 19.1 (Perkin Elmer).

This reduces the lability of the carbamate and creates a robust modification amenable to downstream analysis, in this instance tandem mass spectrometry [MS-MS, (Linthwaite et al., 2018)]. As peptides are preferable to whole proteins due to their mass being in the more favourable range for fragmentation by MS-MS, proteases are used (Olsen et al., 2004). The protease most commonly

employed for this purpose is trypsin, favoured for its superior stability, high proteolytic activity, and high cleavage specificity for the C-termini of Lys and Arg residues (Olsen et al., 2004). The peptides are then subjected to MS-MS, using soft ionisation techniques that result in minimal fragmentation of the peptide of interest and are therefore appropriate for protein mass spectrometry (Fenn et al., 1989; Karas and Hillenkamp, 1988). The two main soft ionisation techniques used in protein mass spectrometry are MALDI and ESI, the former of which uses a solid phase, laser energy absorbing matrix to assist in ionisation, whereas ESI forces the liquid sample out of the tip of a capillary needle to which a high voltage is applied, nebulising the sample and forming a fine aerosol of charged droplets (Fenn et al., 1989; Han et al., 2008; Karas and Hillenkamp, 1988). ESI can easily be coupled to an MS-MS system, with precursor ions selected and fragmented to obtain sequence information and to determine the specific location of the trapped carbamates (Linthwaite, 2017). The presence of a carbamate plus an ethyl group adds an additional mass of 72.0211 Da to the amine (Figure 3), with this shift in mass used to identify trapped carbamates (Linthwaite, 2017). The presence of a missed trypsin cleavage at a potential carbamate site adds confidence to the hit and reduces the chance of a false positive (Linthwaite, 2017). As the activity of trypsin is highly efficient and specific, missed cleavages are a rare occurrence (Olsen et al., 2004). However, the addition of a trapped carbamate to a Lys side chain increases its bulk and it fails to stabilise the Asp residue within the substrate binding pocket, as would occur with a positively charged unmodified Lys or Arg side chain (Olsen et al., 2004). Therefore, trypsin cannot cleave sites where trapped carbamates are present, and missed cleavage sites are used to increase confidence in carbamate hits (Linthwaite, 2017).

This methodology has been successfully used to verify the presence of the known site of carbamate formation in haemoglobin, in addition to a small-scale screen of the *Arabidopsis thaliana* proteome to identify sites of carbamylation (Linthwaite et al., 2018). It operates under conditions that reflect those within the cellular environment, a feature that prior techniques used in the identification of carbamates have lacked (Linthwaite et al., 2018). It therefore provides the capability for the systematic identification of sites of carbamate formation under physiologically relevant conditions, founding the basis of future scientific questions pertaining to the functional relevance of these modifications. Protein carbamate formation is likely more biologically significant than is appreciated, potentially forming the basis of a widespread mechanism for regulation by which protein function is directly coupled to environmental CO_2. Identifying novel sites of carbamylation and characterising the functional relevance of these carbamates will aid in further expanding current understanding of CO_2-mediated protein regulation and how biological systems detect and respond to CO_2.

1.3 CO_2 in *H. pylori* Physiology

H. pylori is a Gram negative, microaerophilic, spiral bacterium that colonises the gastric mucosa of the human stomach, selectively overlaying gastric epithelial cells (Blaser, 1990; Marshall and Warren, 1984; Marshall et al., 1985). Originally isolated from biopsies of the antral mucosa of patients suffering gastritis, this bacterium has been found to be the aetiological agent of gastritis, peptic ulceration, and in the case of chronic infections, predisposition to the development of gastric adenocarcinoma, MALT lymphoma, and non-Hodgkin's lymphoma (Marshall et al., 1985; Parsonnet et al., 1994; Sipponen and Marshall, 2000; Stolte et al., 2002).

Treatment of *H. pylori* infection currently comprises of complex multi-drug regimes combining several antibiotics with proton pump inhibitors and/or bismuth compounds (Argueta and Moss, 2019). However, failure in eradicating *H. pylori* infection is a growing problem attributed to poor medication compliance, the development of antibiotic resistance, the slow growth rate of this bacterium which inhibits the uptake of bactericidal agents, and the acid-labile nature of several antibiotics (Loughlin, 2003). In addition to this, *H. pylori* infects approximately half of the world's population, has been classified as a Class 1 carcinogen, and demonstrates propensity to result in chronic infections (Eusebi et al., 2014; IARC, 1994; Peleteiro et al., 2014). In 2017, the World Health Organisation published a list of antibiotic resistant "priority pathogens" deemed to pose the most significant threat to human health and for which the research and development of new antibiotics is crucial – on this list, *H. pylori* was categorised as a high priority pathogen (WHO, 2017). Therefore, the disease burden of *H. pylori* is significant, meaning there is a great need to

further understand the mechanisms underpinning its virulence and thus develop novel anti-*Helicobacter* compounds (Eusebi et al., 2014; Peleteiro et al., 2014; IARC, 1994).

Similar to the previously mentioned examples of bacterial pathogens, CO_2 is also significant in *H. pylori* virulence. Elevated CO_2 concentrations (5 – 10%) are essential for optimal *H. pylori* growth *in vitro* and long-term survival, with a lack of CO_2 inducing the stringent response (Park et al., 2011). Some research groups have argued that this pathogen is in fact a capnophile rather than a microaerophile, given that high (atmospheric) O_2 concentrations that would otherwise be toxic are tolerated so long as CO_2 concentrations remain high (Bury-Moné et al., 2006; Park et al., 2011). Maintenance of pH homeostasis in the cytoplasm and periplasm is not the sole purpose of these CO_2 requirements either, meaning there exists additional reasons for an elevated CO_2 growth atmosphere being essential (Bury-Moné et al., 2006; Park et al., 2011). For example, CO_2 and HCO_3^- are utilised by *H. pylori* in the synthesis of organic molecules such as pyrimidine nucleotides, with several enzymes involved in CO_2 fixation (Mendz et al., 1994). These include acetyl-CoA carboxylase, pyruvate ferredoxin oxidoreductase, and oxoglutarate ferredoxin oxidoreductase (Kuhns et al., 2016).

In addition to the high levels of CO_2 required for its growth, *H. pylori* generates CO_2 as a result of the activity of urease, the most abundant protein produced by this bacterium (Benoit and Maier, 2003). Urease is a metalloenzyme central to *H. pylori* virulence, key in its colonisation of the gastric mucosa by catalysing the hydrolysis of urea [$NH_2(CO)NH_2$] to produce ammonia (NH_3) and CO_2

(Mobley, 2001). These products consume protons entering the cytoplasm and periplasm, and can exit the cell to neutralise and buffer gastric acid (Mobley, 2001; Sachs et al., 2005). Also essential in this acid acclimation process is the periplasmic α carbonic anhydrase, which catalyses the conversion of CO_2 to HCO_3^-, required as an essential buffer in the periplasm [Equation 1, (Marcus et al., 2005)]. This serves to increase the pH of the microenvironment surrounding the bacterium, giving rise to its acid tolerance. Given the discrepancy between the ecological niche *H. pylori* occupies, with a slightly acidic pH of 5 within the gastric mucus, and its neutral bioenergetic profile, urease and the periplasmic α carbonic anhydrase are essential factors in its success as a human pathogen (Bury-Moné et al., 2004; Meyer-Rosberg et al., 1996; van Vliet et al., 2004).

Given the clear importance of CO_2 in *H. pylori* pathophysiology, it is likely that this organism harbours mechanisms of sensing and responding to the levels of CO_2 in its environment. However, the mechanisms by which *H. pylori* achieves this remain unknown.

1.4 *H. pylori* Nickel-Responsive Regulator

The *H. pylori* nickel-responsive regulator (*Hp*NikR) is a key protein in *H. pylori*'s ability to colonise and thrive within such a hostile niche. *Hp*NikR is a pleiotropic activator and repressor, recognising poorly conserved, imperfect inverted repeats within the promoters of a plethora genes in accordance with the environmental cues of Ni(II) availability and pH – an important feature given the comparatively small size of the *H. pylori* genome (Bury-Moné et al., 2004; Dosanjh and Michel, 2006; Li and Zamble, 2009; Tomb et al., 1997; van Vliet et al., 2004, 2002). These genes are involved in several *H. pylori* functions, including acid adaptation, Ni(II) metabolism, Fe(III) uptake and storage, motility, and stress responses, all requirements for surviving and thriving within the host (Abraham et al., 2006; Benanti and Chivers, 2007; Delany et al., 2005; Ernst et al., 2006, 2005; van Vliet et al., 2002). Examples include *ureA*, a member of the gene cluster required for the synthesis and activation of urease and transcriptionally activated by *Hp*NikR (Ernst et al., 2005; McGee et al., 1999), and *nixA*, which encodes the nickel transport protein NixA, responsible for Ni(II) uptake through the inner membrane and transcriptionally repressed by *Hp*NikR (Bauerfeind et al., 1996; Ernst et al., 2005).

The promiscuous regulatory activity of *Hp*NikR distinguishes it from other NikR homologues, with the main function of NikR in other species of bacteria being to regulate Ni(II) import (Benanti and Chivers, 2010; Chivers and Sauer, 2000, 1999; Dosanjh and Michel, 2006). This expanded NikR regulon in *H. pylori* raises questions regarding whether *Hp*NikR responds to environmental signals additional to Ni(II) availability and pH.

The quaternary structure of *Hp*NikR consists of a tetrameric C-terminal metal binding domain which binds four Ni(II) ions with high affinity [$K_d \leq 5$ pM (Benanti and Chivers, 2007)], and two dimeric ribbon-helix-helix N-terminal DNA binding domains (Figure 4). In addition, there is another cation binding site [Mg(II)] that stabilises the protein-DNA complex [Figure 4b, (Benanti and Chivers, 2007; Dosanjh et al., 2007)]. Based on the most recent crystal structure available (Figure 4), when *Hp*NikR is complexed with DNA, Ni(II) ions are bound in a four-coordinate, square planar coordination geometry mediated by Cys107, His99, and His101 residues from one monomer, and His88 from an adjacent monomer [Figure 4a (Abraham et al., 2006; Benanti and Chivers, 2007; Dosanjh et al., 2007)]. However, alternative coordination geometries have also been observed where two Ni(II) ions are instead bound in either a five- or six-coordinate, square pyramidal or octahedral geometry mediated by His74 and His101 from one subunit, His88 from an adjacent subunit, and two or three H_2O molecules (Figure 4c). The ability of Ni(II) to coordinate to different sites within *Hp*NikR may be related to changes in pH or other environmental cues (Dian et al., 2006; West et al., 2012, 2010).

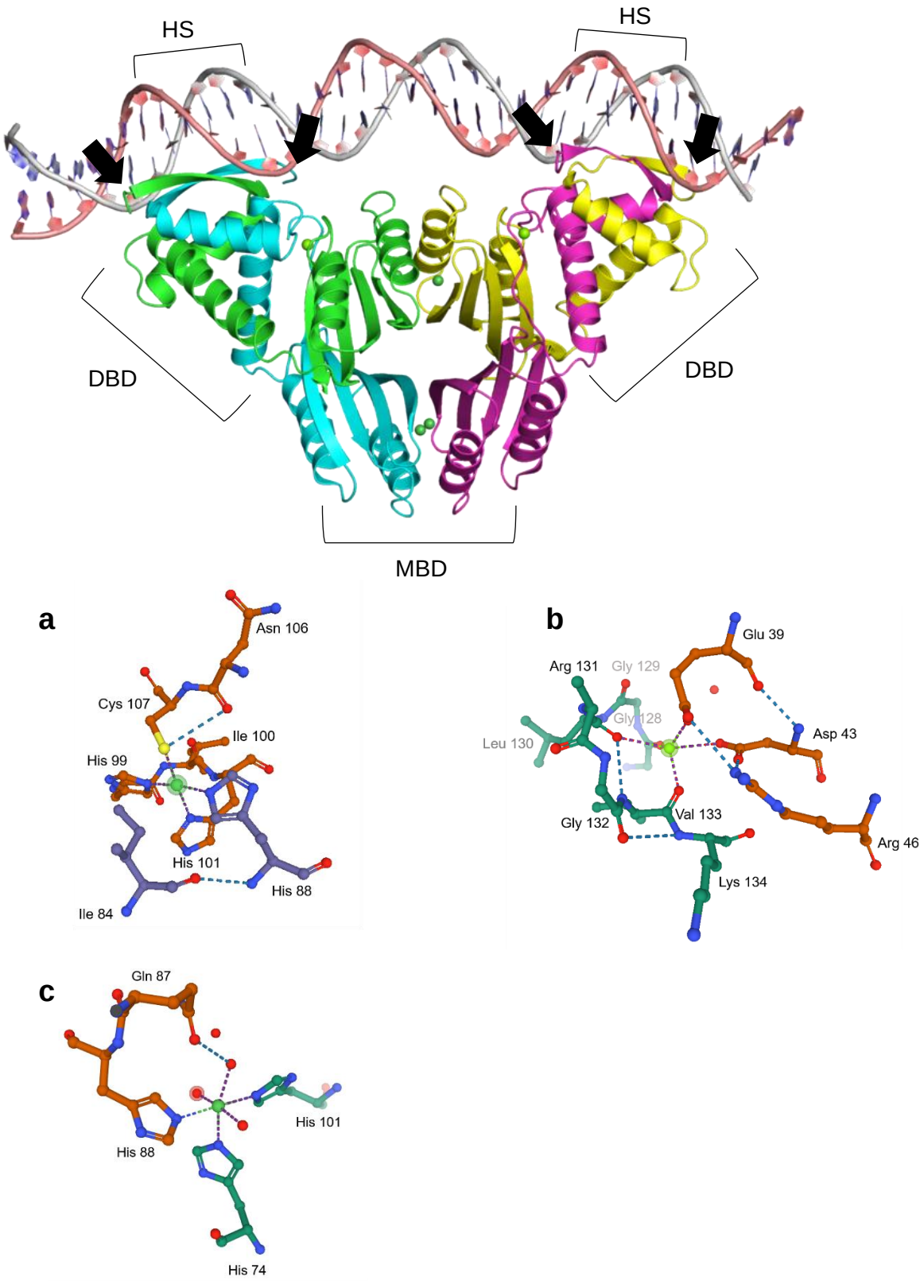


Figure 4. *Hp*NikR tetramer complexed with DNA, represented as a ribbon diagram (PDB ID 6MRJ). Each monomer is coloured respectively. DNA binding domains (DBD), the metal binding domain (MBD), and imperfect inverted repeat half-sites (HS) are labelled respectively. The four Ni(II) ions are coloured dark green, and the two Mg(II) ions coloured pale green. Locations of the N-terminal

arms are indicated by arrows, however it is important to note that in this crystal structure, residues are missing from each arm. The far left arm begins at Asp8, the centre left arm begins at Asp7, the centre right arm begins at Asp7, and the far right arm begins at Lys6. (a) represents the square planar Ni(II) binding site, (b) the Mg(II) binding site, and (c) the octahedral Ni(II) binding site. (a) and (b) were derived from PDB 6MRJ, and (c) from PDB 2CAD.

A unique structural feature of *Hp*NikR is the presence of a nine-residue extension at its N-terminus (Figure 4), which plays an important role in modulating DNA binding (Benanti and Chivers, 2011, 2007). In its entirety, the N-terminal arm is essential for the inhibition of non-specific DNA binding and the maintenance of a hierarchy of affinities for several promoters, with experiments involving site-directed mutagenesis and *Hp*NikR variants with truncated N-terminal sequences demonstrating this in addition to the differential roles of the residues Asp7, Asp8, Pro4, and Lys6 in *Hp*NikR function (Benanti and Chivers, 2007). Asp7 and Asp8 mediate a cation requirement for binding specifically to the promoter of *nixA*, potentially required for the prevention of electrostatic repulsion between the Asp residues and the negatively charged DNA backbone and/or other negatively charged amino acids (Benanti and Chivers, 2007). Pro4 and Lys6, on the other hand, are required for maximal DNA binding affinity, with Pro4Ala and Lys6Met variants demonstrating decreased affinity for the promoters of *nixA* and *ureA* (Benanti and Chivers, 2007). The N-terminal arm is important in the ability of *Hp*NikR to differentiate specific DNA sequences comprising its expanded regulon, with differential contributions from numerous residues and conformational differences upon binding to different sequences promoting this unique functionality (Benanti and Chivers, 2011, 2007). Additionally, great variability exists in both the length and the composition of this arm among NikR homologues in different species of bacteria and even different strains of *H. pylori* (Benanti and Chivers, 2010, 2007). This structural feature

appears to be adaptable, potentially having evolved to accommodate the unique physiologies among different microorganisms. In addition to the N-terminal DNA binding domain, conformational differences upon promoter binding have been demonstrated to occur beyond the *Hp*NikR-DNA interface too, with a Glu47-Lys48 salt bridge evidenced to be essential for high affinity binding to the promoter of *ureA*, but not *nixA* (Benanti and Chivers, 2011).

H. pylori is exquisitely adapted to occupy a hostile niche, requiring the efficient sensing of several environmental stressors and the rapid coordination of appropriate responses in order to survive. Given the unusually small size of the *H. pylori* genome relative to that of other species of bacteria such as *Escherichia coli*, with this comes a restricted number of transcriptional regulators and therefore a requirement for these limited transcriptional regulators to adopt a pleiotropic role in order to fulfil the regulatory demands of this pathogen, such is the case with *Hp*NikR (Contreras et al., 2003; Dosanjh and Michel, 2006; Tomb et al., 1997). A question that remains, however, is the extent to which this pleiotropic regulator behaves as a sensor that couples the transcription of genes important in host adaptation to environmental cues, and whether *Hp*NikR responds to cues in addition to Ni(II) availability and pH.

Given the significant role of CO_2 in *H. pylori* pathophysiology, essential for growth and a product of urease-mediated acid adaptation; the likely presence of but as yet undiscovered CO_2 sensing mechanisms within *H. pylori*; the role of *Hp*NikR as a pleiotropic transcriptional regulator already established to behave as a sensor that couples its activity to the environmental cues of Ni(II) availability and pH, and the evidenced importance of two Lys residues, Lys6

and Lys48, in *Hp*NikR-DNA binding (Benanti and Chivers, 2011, 2007; Bury-Moné et al., 2006, 2004; Dosanjh and Michel, 2006; Li and Zamble, 2009; Mendz et al., 1994; Mobley, 2001; van Vliet et al., 2004, 2002), it was hypothesised that *Hp*NikR may also behave as a CO_2 sensor that couples its regulatory activity to the availability of CO_2, using carbamylation as a mechanism to achieve this. The research comprising this thesis sought to investigate this, first using the carbamate trapping methodology developed by Linthwaite et al. in order probe the entire sequence of *Hp*NikR for evidence of carbamate formation, followed by a bioinformatic approach to investigate the conservation of any sites of carbamate formation by sequence alignment, and finalised with an assessment of the functional consequences of carbamate formation using DNA binding assays.

Chapter 2: Materials and Methods

2.1 Reagents and Buffers

Buffers were prepared with reagent grade chemicals, H_2O deionised using a Milli-Q water system, and pH adjusted using dropwise additions of either HCl or NaOH.

Table 1. Buffers used for *Hp*NikR purification and DNA binding assays.

Buffer	Purpose	Composition
A	Nickel affinity equilibration buffer	20 mM Tris (pH 7.5) 500 mM NaCl 10 mM imidazole
B	Nickel affinity elution buffer	20 mM Tris (pH 7.5) 500 mM NaCl 300 mM imidazole
C	Anion exchange equilibration buffer	20 mM Tris (pH 7.5) 200 mM NaCl
D	Anion exchange elution buffer	20 mM Tris (pH 7.5) 1 M NaCl
E	Size exclusion buffer	20 mM Na_2HPO_4 (pH 7.5) 50 mM NaCl
F	SDS-PAGE running buffer	25 mM Tris 192 mM glycine 0.1% SDS
G	Oligonucleotide annealing buffer	10 mM HEPES (pH 7.5) 300 mM NaCl
H	EMSA running buffer	50 mM Tris 25 mM boric acid 1.5 mM $MgCl_2$
I	EMSA binding buffer	10 mM K_2HPO_4 (pH 7.5) 100 mM KCl 1.5 mM $MgCl_2$
J	Reduced Mg EMSA binding buffer	10 mM K_2HPO_4 (pH 7.5) 100 mM KCl 750 µM $MgCl_2$

2.2 *Hp*NikR Expression

E. coli BL21 cells were transformed with pEB116 full-length wild type or pEB116 full-length Lys6Arg *Hp*NikR, with pEB116 created using a pET-21b backbone, containing a T7 promoter, and the *Hp*NikR stop codon present 5' to the Xho I site to prevent translation of the plasmid-encoded His-tag [constructs and transformed cells were made and provided by P. T. Chivers (Benanti and Chivers, 2007)]. Cells were grown in LB media containing 200 μg/mL ampicillin, 1 mM $MgSO_4$, and 0.1% glucose. Following induction with 0.4 mM IPTG, cells were grown for 4 hr and then harvested using centrifugation (4,000 rpm, 25 min, 4 °C). The media supernatant was discarded, and the pellets resuspended in approximately 20 mL Buffer A per L of culture. The resuspended pellets were stored at -80 °C until purification.

2.3 *Hp*NikR Purification

All purification steps were completed at 20 °C. Cells were thawed, PMSF was added to a final concentration of 0.5 mM to inhibit protease activity, and then lysed using sonication with 4 × 1-minute cycles at approximately 55% power. The lysate was centrifuged to remove cell debris (17,000 rpm, 30 min, 4 °C), and the supernatant retained for purification.

The supernatant (40 – 50 mL) was applied to a pre-packed Ni-NTA column (5 mL, HisTrap™, GE Healthcare) for nickel affinity chromatography that had been equilibrated with 4 volumes of Buffer A, assisted by a peristaltic pump at a flow rate of 1 mL/min. Flow-through was collected. The column was then washed with an additional 4 volumes of Buffer A and the wash collected. Bound proteins were eluted using a 50 mL gradient from Buffer A to Buffer B. Fractions were collected in 2 mL volumes.

Fractions containing *Hp*NikR were identified using SDS-PAGE (Section 2.4), pooled, and then diluted 5-fold using ddH_2O to reduce NaCl concentration for further purification using anion exchange chromatography. Fractions were applied to a Q-Sepharose column (5 mL, GE Healthcare) equilibrated with 4 volumes of Buffer C, assisted by a peristaltic pump at a flow rate of 1 mL/min. Flow-through was collected. Following application of the sample, the column was washed with a further 4 volumes of Buffer C and the wash collected. Bound proteins were eluted using a 50 mL gradient from Buffer C to Buffer D. Fractions were collected in 2 mL volumes, and those containing *Hp*NikR identified using SDS-PAGE (Section 2.4).

A *Hp*NikR-containing sample totalling 5 mL was applied to a size exclusion column (320 mL, HiLoad™ 26/600 Superdex™ 75 pg) for the final purification step, equilibrated with Buffer E. Fractions were collected in 5 mL volumes after allowing 100 mL to flow through, prior to which *Hp*NikR does not elute. *Hp*NikR-containing fractions were identified using SDS-PAGE (Section 2.4). Final concentration was determined in 6 M GuHCl using $\varepsilon = 9860\ M^{-1}\ cm^{-1}$. Molecular weight (expected: 17,147 Da; observed: 17,147 Da) was confirmed using MALDI-MS (Department of Chemistry, Durham University).

2.4 SDS-PAGE

SDS-PAGE was used to identify fractions containing *Hp*NikR. Samples containing 8 µL fraction, 7 µL ddH_2O, and 3 µL 6x loading buffer (0.2 M Tris-HCl (pH 6.8), 10% (v/v) SDS, 30% (v/v) glycerol, 10 mM β-Mercapthoethanol, 0.05% (w/v) Bromophenol Blue) were incubated for 10 min at 95 °C to enable protein denaturation. Samples were run on 12% polyacrylamide separating gels with 5% polyacrylamide stacking gels in Buffer F for 1 hr at 150 V. A PageRuler™

pre-stained protein molecular weight ladder (Thermo Scientific) was used to estimate protein size. Gels were stained by incubating with InstantBlue™ (Expedeon) for a minimum of 1 hr, but usually overnight (16 – 18 hrs).

2.5 DNA Duplex Preparation

HPLC- and HPSF-purified oligonucleotides for the *ureA* and *nixA* promoter fragments (Eurofins Genomics), one strand Flu-dT or Cy3-labelled ([F]) and a complementary unlabelled strand, were resuspended in DNase-free H_2O (100 µM final concentration) upon receipt in preparation for annealing. The central Flu-dT tag was chosen due to its successful use in prior studies (Dosanjh and Michel, 2007; Dosanjh et al., 2009), whereas the terminal Cy3 tag was chosen when attempting to optimise the fluorescence anisotropy assay (Section 3.4).

Table 2. Oligonucleotide sequences used in DNA duplex preparation. Underlined sequences indicate *Hp*NikR recognition sites.

Gene	Fluorescent label	Sequence (5' → 3')
ureA +	Flu-dT	CTTCAAAGATATAACACTAAT**[F]**TCATTTTAAATAATAATTAGTTAATGAAC
ureA +	Cy3	**[F]**CTTCAAAGATATAACACTAATTCATTTTAAATAATAATTAGTTAATGAAC
ureA -	-	GTTCATTAACTAATTATTATTTAAAATGAATTAGTGTTATATCTTTGAAG
nixA +	Flu-dT	TTAACAAAATATATTACAAT**[F]**TACCAAAAAAGTATTATTTTTCTTAAAAGG
nixA +	Cy3	**[F]**TTAACAAAATATATTACAATTACCAAAAAAGTATTATTTTTCTTAAAAGG
nixA -	-	CCTTTTAAGAAAAATAATACTTTTTTGGTAATTGTAATATATTTTGTTAA

+ = sense sequence; - = anti-sense sequence.

Double stranded probes were produced by mixing forward and reverse primers in a 1:1 ratio (10 μM final concentration) in Buffer G, totalling 50 μL. Annealing reactions were heated to 95 °C using a thermal cycler, followed by switching the machine off and allowing the reactions to cool slowly to 20 °C for 16 – 18 hrs. Duplexes were then wrapped in aluminium foil to protect the fluorescent probes from light exposure and stored at -20 °C.

2.6 *Hp*NikR Carbamate Trapping

*Hp*NikR carbamate trapping was performed as described previously (Linthwaite et al., 2018). 8.4 mg $NaHCO_3$ was dissolved in 3 mL 50 mM phosphate buffer (pH 7.5) and added to 1 mg *Hp*NikR stored in 1 mL Buffer E, chosen to prevent the interference of Tris with the trapping reaction (final $NaHCO_3$ concentration 25 mM). The mixture was then transferred to a TIM856 Titration Manager (Radiometer Analytical; Cann lab, Department of Chemistry, Durham University). 280 mg Et_3OBF_4, dissolved in 1 mL ddH_2O, was added in 3 portion-wise steps at 2 sec intervals, with a constant pH (7.5) maintained through the addition of 1 M NaOH solution via the automatic burette to counteract the hydrolysis of Et_3OBF_4. The reaction mixture was incubated at 25 °C for 1 hr with stirring. The 5 mL reaction mixture was then transferred into dialysis tubing [35 mm, 3.5K MW-cut off SnakeSkin™ (Thermo Scientific)] and dialysed against 1 L ddH_2O (16 – 18 hr, 5 °C). The reaction mixture was then transferred into 2 mL microcentrifuge tubes, dried down using vacuum concentration at 30 °C (Eppendorf Concentrator plus; Cann lab, Department of Chemistry, Durham University), and stored at -20 °C until preparation for mass spectrometry (Section 2.7).

2.7 Mass Spectrometry

Dried *Hp*NikR (Section 2.6) was trypsin-digested using S-Trap™ micro spin columns according to the manufacturer's protocol, without modifications (ProtiFi). The resulting peptide mixture was analysed using ESI-MS (Department of Biosciences, Durham University) as described previously (Linthwaite et al., 2018) and the trapped carbamate at Lys6 verified using PEAKS Studio (Tran et al., 2019) by searching for the mass of a carbamate plus an ethyl group (72.0211 Da, Figure 3).

2.8 Bioinformatics

The *Hp*NikR protein sequence of *H. pylori* 26695 (HP1338) was used in a protein BLAST search of the NCBI non-redundant protein sequence database [performed in April 2020 (NCBI Resource Coordinators, 2018)]. The search set was limited to only include sequences from *H. pylori* strains (taxid: 210) or the *Helicobacter* genus (taxid: 209). Sequences of significant similarity (E value < 1e-50) were selected, automatically filtered for redundant sequences (100% identity) using Jalview 2.11.1.0 (Livingstone and Barton, 1993; Waterhouse et al., 2009), aligned using ClustalWS (Thompson et al., 1994), and then manually screened for the presence of the Ni(II)-binding site ligands (Cys107, His99, His101, His88, and His74). The sequences N-terminal to Arg12, a highly conserved β1 residue that precedes the N-terminal arm, of the remaining sequences were selected and further automatically filtered for redundancy (100% identity). Conservation scores were calculated using Jalview 2.11.1.0 (Livingstone and Barton, 1993; Waterhouse et al., 2009). Sequence logos were generated using WebLogo (Crooks et al., 2004).

2.9 Fluorescence Anisotropy

Fluorescence anisotropy measurements were taken using a Cary Eclipse Fluorescence Spectrophotometer at 20 °C. 1 mL binding reactions containing 10 nM DNA labelled with Flu-dT or Cy3 in Buffer I were added to a 1 cm pathlength quartz cuvette (Hellma) that was pre-treated with Rain-X to prevent DNA or protein from sticking to the cuvette walls. Excitation and emission wavelengths were 495 nm and 519 nm respectively for assays containing Flu-dT labelled DNA, and 552 nm and 570 nm respectively for assays containing Cy3 labelled DNA. The anisotropy (r_{obvs}) of free DNA was measured, followed by the stepwise addition of *Hp*NikR to the binding reactions until a final concentration of 1 μM was reached. Binding reactions were allowed to equilibrate for 3 min following each addition of *Hp*NikR, and the change in r_{obvs} then recorded. Two replicate measurements were taken per addition. Graphs plotting r_{obvs} against *Hp*NikR concentration were generated using GraphPad Prism version 5.0.1 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com.

2.10 Electrophoretic Mobility Shift Assays

EMSAs were adapted from a previously described method (Benanti and Chivers, 2007). Native 7% polyacrylamide gel wells were flushed with ddH_2O and pre-run at a constant voltage (100 V) with Buffer H until the current no longer decreased with time (45 – 60 min; ~ 25 → 8 mA). 15 μL binding reactions contained 25 nM DNA incubated in Buffer I and 5% glycerol with a dilution series of at least 15 *Hp*NikR concentrations. To minimise the effects of CO_2 evaporation from solution, in experiments where CO_2-dependent effects were investigated, 10 mL aliquots of Buffer H and Buffer I were removed from

the stock solutions immediately prior to setting up the experiment and 210 mg $NaHCO_3$ added to Buffer H (final concentration 250 mM) and 63 mg $NaHCO_3$ added to Buffer I (final concentration 75 mM). pH was continually monitored with stirring while the added $NaHCO_3$ dissolved, with HCl added dropwise to counteract the pH increase and ensure maintenance at 7.5. The Buffer H aliquot was then used to make the 7% native polyacrylamide gels (final $NaHCO_3$ concentration 25 mM), and the Buffer I aliquot used in binding reactions (final $NaHCO_3$ concentration 25 mM). Binding reactions were incubated for 30 min at 25 °C followed by loading of the total volume directly onto the running gel.

After running (30 – 45 min), gels were stored in ddH_2O and covered in aluminium foil to protect the fluorescent probes from light exposure. Approximately 15 min after running, due to travel time between departments, gels were imaged using a Typhoon 94000 (Chivasa lab, Department of Biosciences, Durham University). Gels were quantified using ImageJ (Schindelin et al., 2012). The fluorescent signal in each band was determined by selecting each lane of the gel, generating an intensity histogram for each lane, manually defining the histogram for individual bound and unbound bands, and then using the Magic Wand tool to select each histogram and generate a numerical value corresponding to signal intensity (Figure 5).

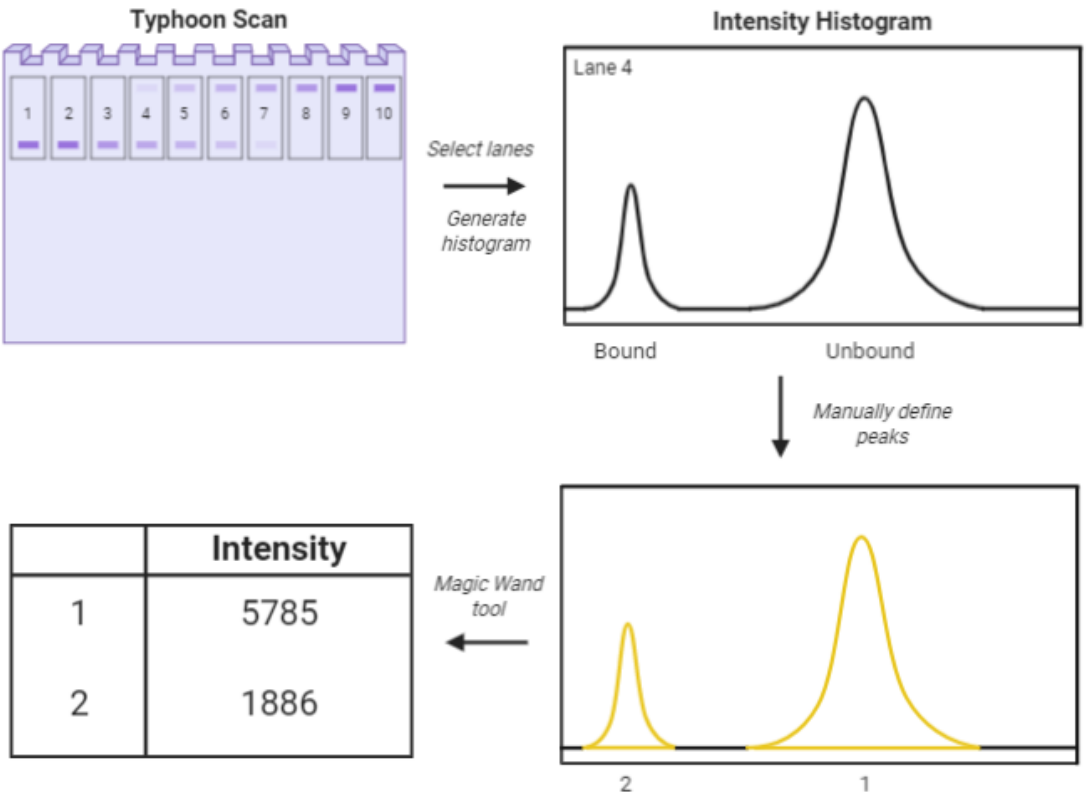


Figure 5. Schematic demonstrating the quantification of EMSAs using ImageJ. Figure created using BioRender.

Fraction bound values were calculated using Equation 2, where F_{bound} is the signal intensity of bound bands and $F_{unbound}$ the signal intensity of unbound bands.

$$\text{Fraction bound} = \frac{F_{bound}}{F_{bound} + F_{unbound}} \qquad (2)$$

Binding curves were generated and apparent affinities calculated using GraphPad Prism version 5.0.1. The data were fit using a non-linear regression analysis and Equation 3, where y is the fraction of DNA bound; K_d is the *Hp*NikR concentration required for half-maximal DNA binding; x is the concentration of *Hp*NikR, and *n* is the Hill coefficient.

$$y = \frac{1}{1 + \left(\frac{K_d}{x}\right)^n} \qquad (3)$$

Chapter 3: Results

The restrictions imposed by lockdown as a result of COVID-19 have considerably affected the work presented in this thesis. A reasonable set of experiments to complete this thesis were agreed upon with thesis committee members Dr. Tim Blower and Dr. Paul Denny in order to accommodate the substantial loss of time that would otherwise have been dedicated to conducting experiments.

3.1 Lys6 is a site of carbamylation

Given the significant role of CO_2 in *H. pylori* pathophysiology, the role of *Hp*NikR as a pleiotropic transcriptional regulator known to modulate its activity in accordance with the environment signals pH and Ni(II) availability, and the evidenced importance of two Lys residues, Lys6 and Lys48, in DNA binding (Benanti and Chivers, 2011, 2007; Bury-Moné et al., 2006, 2004; Dosanjh and Michel, 2006; Li and Zamble, 2009; Mendz et al., 1994; Mobley, 2001; van Vliet et al., 2004, 2002), it was hypothesised that *Hp*NikR may also respond the environmental signal CO_2, using carbamylation as a mechanism to achieve this. To investigate this, the carbamate trapping methodology developed by Linthwaite et al. was employed (Linthwaite et al., 2018). Purified *Hp*NikR was subjected to CO_2 treatment, followed by TEO trapping, digestion using trypsin, and the peptides analysed by ESI-MS. The data were analysed in PEAKS Studio, where Lys modifications with masses of 72.0211 Da consistent with the mass of a carbamate with the addition of an ethyl group were searched for. It was revealed that Lys6 is a site of carbamylation, with the difference in mass between the y_6 and y_7 ions 72.0211 Da greater than would be expected for an

unmodified Lys (Figures 6 & 7), consistent with carbamate formation. Confidence in the presence of this carbamate was improved by the location of the modified residue within the peptide, indicating a missed trypsin cleavage, the presence of the modification in multiple peptides, and the surrounding of the modification by the ESI fragmentation pattern [Figure 7, (Linthwaite, 2017)], reducing the likelihood of the result being a false positive. While the presence of the Lys6 carbamate is significant, it is not known exactly what proportion of *Hp*NikR is carbamylated, nor how many monomers per tetramer contain carbamates at Lys6.

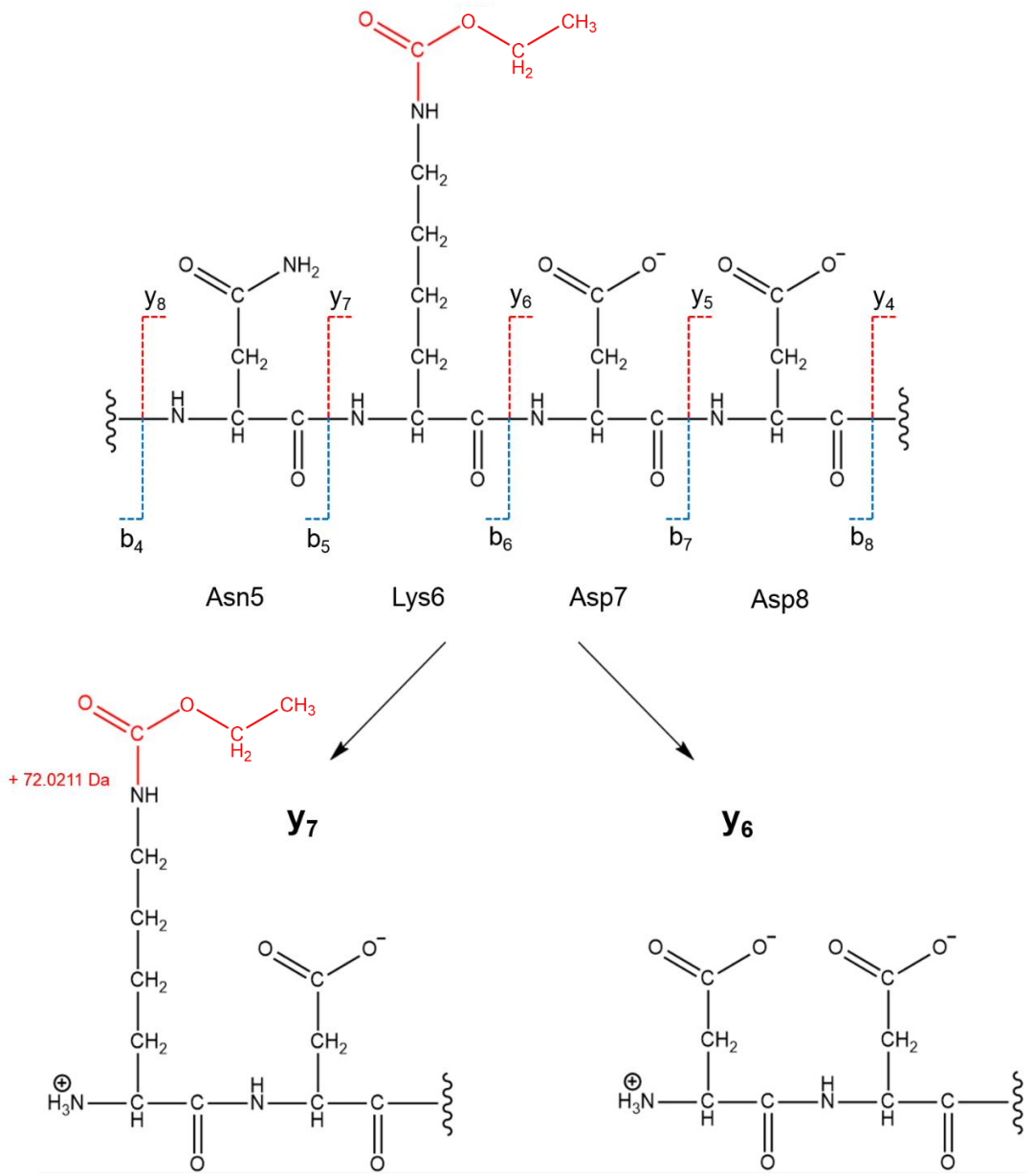


Figure 6. Peptide fragmentation resulting in the formation of b and y ions. The structures the of y_6 and y_7 ions are shown, demonstrating the 72.0211 Da mass addition resulting from the trapping of a carbamate with an ethyl group (red).

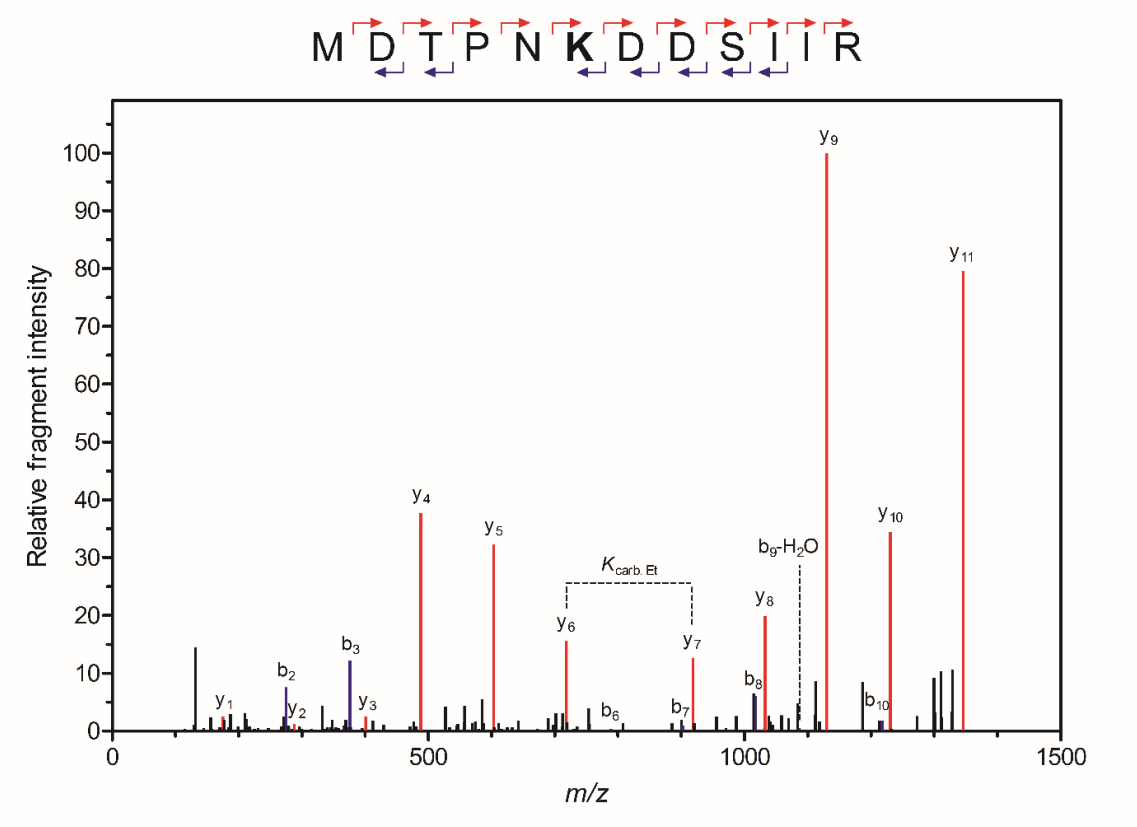


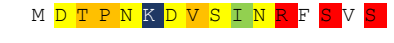
Figure 7. Relative fragment intensity versus mass/charge ratio (m/z) for ESI-MS data identifying an ethyl-trapped carbamate on the ε-amino group of Lys6. The peptide sequence signifies the y (red) and b (blue) ions present in the plot, with the site of carbamylation shown in bold.

3.2 Lys6 is completely conserved across all *H. pylori* NikR homologues

Given that Lys6 is a site of carbamylation in the *H. pylori* 26695 NikR homologue, and the significant variability in the length and the composition of the N-terminal arm that exists among NikR homologues (Benanti and Chivers, 2010, 2007), a bioinformatic approach was adopted to explore the N-terminal arm sequence variation among *Hp*NikR homologues in greater detail across different strains of *H. pylori* and to determine the conservation of the site of carbamate formation.

Table 3. Alignment of *Hp*NikR N-terminal arm sequences.

H. pylori strain	Accession number	MLST lineage	*cag* PAI status	N-terminal sequence
26695	WP_000380787.1	HpEurope	+	MDTPNKDDSIIRFSVS
SJM180	WP_000380308.1	HpEurope	+	MDTLNKDDSIIRFSVS
PZ5056	EQD99900.1	HpEurope	+	MTAFNELMPKKDFSQVDTPNKDDSIIRFSVS
HPAG1	WP_000379504.1	HpEurope	+	MDTHNKGDSIIRFSVS
GAM264Ai	EMH23368.1	HpAfrica1	-	MPKKDFSQMDTHHKDDSIIRFSVS
GAM83Bi	EMH37885.1	HpAfrica1	-	MPKKDFSQMDTHNKDDSIIRFSVS
Hp P-4c	WP_000379459.1	HpAfrica1	+	MDTHHKDDLIIRFSVS
SouthAfrica20	WP_020972932.1	HpAfrica2	-	MDTHNKDNKDNKDDSIIRFSVS
SouthAfrica50	WP_021435452.1	HpAfrica2	-	MDTHNKDNKDDSIIRFSVS
FD577	WP_000339837.1	HspEAsia	+	MDANNKDDSIIRFSVS
52	WP_000354193.1	HspEAsia	+	MDIPNKDDSIIRFSVS
HP15012	WP_120834227.1			MDTPNKDESIIRFSVS
B659-C2	WP_120877193.1			METPNKDDSIIRFSVS
MHP41	WP_156594252.1			MDTPSKDDSIIRFSVS
HP13029	WP_120960577.1			MDTPNKDDLIIRFSVS
55:1	WP_108378236.1			MDTPNKDNSIIRFSVS
G-Mx-2003-108	WP_079333917.1			MDTPNKGDSIIRFSVS
HP15001	WP_131127447.1			MGTHNKDDSIIRFSVS

MKF8	WP_096401031.1	
25-A-EK9	WP_139519125.1	
22095	WP_078268302.1	
MHP19	WP_156557893.1	
C-Mx-2006-677	WP_165560623.1	
CHL29	WP_154410287.1	
HP12064	WP_120819234.1	
Nic47	WP_165515711.1	
CHL20	WP_154479160.1	
UM400AS	WP_058904968.1	
CHL41	WP_154506487.1	
HP12020	WP_120892346.1	
HP15059	WP_120902988.1	

Red = DNA-contacting β1 residues.

Orange = conservation scores 0 – 3; yellow = conservation scores 4 – 6; green = conservation scores 7 – 10; dark blue = conservation score 11

cag PAI = *cag* pathogenicity island; MLST = multilocus sequence typing. *cag* PAI and MLST data obtained from van Vliet and Kusters (van Vliet and Kusters, 2015; van Vliet, 2016).

Table 4. Conservation scores of the N-terminal arm residues calculated using Jalview from a multiple sequence alignment of 31 unique *Hp*NikR sequences.

N-terminal arm residue	Conservation score
Asp2	4
Thr3	2
Pro4	0
Asn5	4
Lys6	11
Asp7	5
Asp8	3
Ser9	5
Ile10	7
Ile11	6

Figure 8. N-terminal arm sequence logo generated using WebLogo from a multiple sequence alignment of 31 unique *Hp*NikR sequences (Table 3). Overall stack height is proportional to the degree of conservation at that position and symbol height within stacks indicates the relative frequency of each amino acid at that position. Sequence numbering starts at residue 16 for clarity, as only six sequences extend beyond this range.

A total of 31 non-identical sequences N-terminal to the β-sheet were obtained following a protein BLAST search of the NCBI non-redundant protein sequences database, which showed a high degree of variability in both length and composition (Table 3). By comparing the number of residues present between the N-terminal Met and the β1 Arg12, it was revealed that the length of the N-terminal arm in *Hp*NikR homologues among different strains of *H. pylori* ranged from 10 to 25 amino acids. 25 of the 31 sequences contained 10 amino acids between Met and Arg; 1 sequence contained 11 amino acids; 1 sequence contained 13 amino acids; 1 sequence contained 16 amino acids; 2 sequences contained 18 amino acids, and 1 sequence contained 25 amino acids (Table 3). An important feature, however, is the presence of an internal Met and Val within the N-terminal arms 18 and 25 amino acids in length respectively. Given that the sequences obtained are based on annotations of the *nikR* gene, it is a possibility that these unusually long N-terminal arms are a result of sequencing or annotation error, as suggested by E. L. Benanti (Benanti, 2009). In an attempt to determine any relationships between strain lineage, pathogenicity, and *Hp*NikR regulation, the multilocus sequence typing (MLST) lineages and *cag* PAI (*cag* pathogenicity island) status were sought after in the literature (van Vliet and Kusters, 2015; van Vliet, 2016). However, this information was only available for 11 of the listed strains in Table 3, with very few representative strains from each MLST lineage from which robust conclusions could be drawn. There appears to be no patterns between N-terminal arm sequence and length and MLST lineage, although all *cag* negative strains contain unusually long N-terminal arms (Table 3). However, the *cag* positive strain PZ5056 also contains an unusually long N-terminal arm, and as previously noted, the unusually long

N-terminal arms may be a result of sequencing or annotation error rather than of biological significance (Benanti, 2009).

In addition to variability in length, great variability in the composition of the N-terminal arms among *Hp*NikR homologues was observed. Conservation scores were determined for each N-terminal arm residue, measured as a numerical index with scores ranging from 0 to 11 corresponding to ascending conservation (Tables 3 & 4). These scores reflect the conservation of physico-chemical properties in the alignment, with a score of 11 indicating complete residue conservation, and a score of 10 indicating the presence of mutations but total conservation of physico-chemical properties. The corresponding WebLogo (Figure 8) serves to visually represent this conservation. While there exists significant variability in the amino acid composition of the *Hp*NikR N-terminal sequences, there is a notable complete conservation of the Lys6 residue (Figure 8, Tables 3 & 4). Lys6 is the only residue of the entire N-terminal arm that is completely conserved, with a total lack of mutations or substitutions for residues of similar physico-chemical properties as has been observed with other residues, such as the substitution of Asp for Glu. Evidently, this residue is of great importance in *Hp*NikR function, having already been demonstrated to be required for maximal DNA binding affinity (Benanti and Chivers, 2007).

3.3 Lys6 conservation is lost among *Helicobacter spp.* NikR homologues

To further explore the conservation of Lys6, the BLAST search was expanded to include NikR homologues within all species of the *Helicobacter* genus.

Table 5. Alignment of *Helicobacter spp.* NikR N-terminal arm sequences. Species are organised based on tissue tropism.

Helicobacter spp.	Urease	Accession number	N-terminal sequence
Gastric			
pylori 26695	+	WP_000380787.1	MDTPNKDDSIIRFSVS
acinonychis Sheeba	+	WP_011577228.1	MDTNNKDDSIIRFSVS
cetorum MIT 99-5656	+	WP_014659486.1	MDTHQKEDSTIRFSVS
cetorum	+	WP_104748176.1	MDTHYKEDSTIRFSVS
cetorum	+	WP_104712242.1	MDTFKEESTIRFSVS
ailurogastricus	+	WP_053941568.1	MENKEDSIIRFSVS
heilmannii ASB1.4	+	WP_015107465.1	MENKEDAIIRFSVS
felis	+	WP_104577964.1	MEHKDDTIIRFSVS
baculiformis	+	WP_104752615.1	MEHKDDAIIRFSVS
sp. 11S02596-1	+	WP_095332026.1	MHPKEQDNIIRFSVS
sp. 12S02232-10	+	WP_095296397.1	MYPKEQDSIIRFSVS
sp. 11S03491-1	+	WP_095274287.1	MQKDQDFTIRFSVS
sp. L2	?	WP_121020462.1	MEHKDDSIIRFSVS
salomonis	+	WP_104696041.1	MDHKDDIIRFSVS

sp. 12S02634-8	+	WP_095225626.1	MHTKDLDPTIRFSVS
bizzozeronii	+	WP_104638666.1	MENRDDTIIRFSVS
suis HSMm R04052c	+	WP_163533880.1	MEHREDAIIRFSVS
sp. 13S00482-2	+	WP_095295080.1	MQLKTHTKENDCTIRFSVS
sp. 13S00477-4	+	WP_095271709.1	MYKKEEDLSIRFSVS
mustelae NCTC 12198	+	WP_013023516.1	MRTMEKEKNSLMRFSVS
Enterohepatic			
anseris	+	WP_115578549.1	MENSKNELIRFSVS
sp. MIT 05-5293	?	WP_034369306.1	MKHHNIIRFSVS
magdeburgensis MIT 96-1001	-	WP_034586346.1	MKQNIIRFSVS
sp. MIT 01-3238	-	WP_115520014.1	MATINTKSKNHTSHTENSLKSQNPAKKLTKNPAQNPASNHAENTQNPKHSHKQHSPKQHSPKQHSKEESVTRLSIS
fennelliae MRY12-0050	-	WP_023946458.1	MKIKKSITNPNHTHKTHTHTSKNEGVIRFSIS
sp. MIT 03-1616	?	WP_034327098.1	MKPNIIRFSVS
sp. MIT 16-1353	?	WP_112059938.1	MDKNTLTRFSIS
japonicus MIT 01-6451	-	WP_034360409.1	MKSNIIRFSVS
canis NCTC 12740	-	WP_023930163.1	MRDHKVVRFSVS
canis NCTC 12410	-	WP_115012060.1	MSNHKVVRFSIS
canis CCUG 32756T	-	WP_150337514.1	MRDRKVVRFSVS
trogontum	+	WP_034346470.1	MRQVRNEKLQDPKDSIIRFSVS
bilis ATCC 51630	+	WP_004083936.1	MRQVKSEKTQENKDSIIRFSVS
muridarum	+	WP_034557661.1	MLDNNQIIRFSVS
aurati	+	WP_104762656.1	MPASKNTNTVIRFSVS
sp. MIT 17-337	?	WP_115543716.1	MQDKVKNNKDSIIRFSVS
sp. CLO-3	-	WP_066457018.1	MKKDENKVVRFSIS
valdiviensis WBE14	+	WP_111229746.1	MKDLIRFCVS

burdigaliensis CNRCH 2005/566H	+	WP_121764537.1	MKELIRFCVS
bilis	+	WP_104746840.1	MRQVKSEKVQENKDSIIRFSVS
hepaticus ATCC 51449	+	WP_011115195.1	MKSSIIRFSVS
macacae MIT 99-5501	-	WP_023928564.1	MATIKTKSKNHTSYTKNSLKNQNLAEKLTKNLTQNPASNHAENTQNPKHSPKQHSPKQHSKEDSVTRLSIS
Unclassified			
sp. house sparrow 1	?	WP_104696819.1	MDLKNEVIRFSVS

Red = DNA-contacting β1 residues. ? = urease status could not be identified in the literature. Unclassified = site of colonisation could not be identified in the literature.

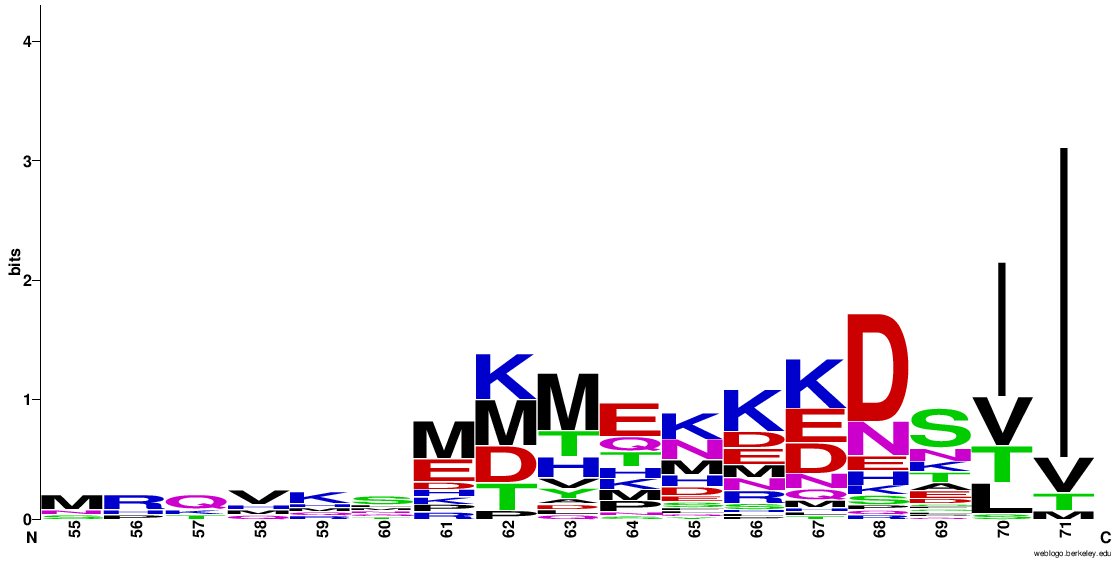


Figure 9. N-terminal arm sequence logo generated using WebLogo from a multiple sequence alignment of 43 unique *Helicobacter spp.* NikR sequences. Overall stack height is proportional to the degree of conservation at that position and symbol height within stacks indicates the relative frequency of each amino acid at that position. Sequence numbering starts at residue 55 for clarity, as only 7 sequences extend beyond this range

A total of 43 non-identical sequences N-terminal to the β-sheet were obtained, which similarly to the *Hp*NikR sequences showed a high degree of variability in both length and composition (Table 5, Figure 9). The length of the N-terminal arm in NikR homologues among different *Helicobacter* species ranged from 4 to 70 amino acids. 2 of the 43 sequences contained 4 amino acids between Met and Arg; 4 sequences contained 5 amino acids; 5 sequences contained 6 amino acids; 3 sequences contained 7 amino acids; 10 sequences contained 8 amino acids; 5 sequences contained 9 amino acids; 5 sequences contained 10 amino acids; 1 sequence contained 11 amino acids; 1 sequence contained 12 amino acids; 1 sequence contained 13 amino acids; 3 sequences contained 16 amino acids; 1 sequence contained 26 amino acids; 1 sequence contained 65 amino acids, and 1 sequence contained 70 amino acids (Table 5). However, internal Met, Val, and Leu residues are present within several N-terminal arms >

11 amino acids in length, which similarly to the *Hp*NikR sequences may have resulted from sequencing or annotation errors.

In addition to variability in length, great variability in the composition of the N-terminal arms among *Helicobacter spp.* NikR homologues was observed. All residues are poorly conserved (Figure 9), including a total loss of Lys6 conservation. Notable, however, is the presence of at least one Lys residue in almost every sequence, with no apparent correlation between the tissue tropism of *Helicobacter* species (gastric or enterohepatic) and the number or location of Lys residues within the N-terminal arm, with the exceptions of *H. bizzozeronii*, *H. suis*, and *H. muridarum* (Table 5). The *H. bizzozeronii* and *H. suis* NikR N-terminal arm Lys residues are substituted for Arg, whereas *H. muridarum* NikR contains no basic residues (Table 5). These substitutions could be attributed to a sequencing error, or could indicate that the role of Lys in these NikR homologues is one requiring its basic (pK_a) properties that could be interchangeable with Arg. *H. muridarum* NikR appears to completely lack the requirement for a basic residue within the N-terminal arm. However, the vast majority of *Helicobacter spp.* NikR homologues do contain at least one Lys residue, indicative of a generally widespread requirement for this residue across the *Helicobacter* genus.

3.4 Wild type *Hp*NikR exhibits a decrease in affinity for *ureA* and *nixA* promoters in response to CO_2/HCO_3^-

Having evidenced the formation of a carbamate at Lys6 and investigated the conservation of the residue, the next step was to determine the functional effect of this modification on *Hp*NikR activity. Given the introduction of a negative charge that occurs as a result of carbamate formation, it was hypothesised that the presence of this charge may weaken DNA binding affinity as a result of electrostatic forces of repulsion between the carbamate and the negatively charged phosphate groups comprising the DNA backbone. To investigate this, DNA binding assays were conducted with and without the addition of $NaHCO_3$ as a means to introduce CO_2 (Equation 1) in order to quantify the binding activity of *Hp*NikR in the presence and absence of CO_2 and calculate approximate affinities for the promoters of *ureA* and *nixA*. *ureA* and *nixA* were chosen due to their being among the best characterised *Hp*NikR targets, in addition to their opposing regulatory outcomes upon *Hp*NikR binding. Fluorescence anisotropy assays were originally chosen due to being an efficient, solution-based, and quantifiable method that had already been successfully used by Dosanjh et al. to quantify the binding of *Hp*NikR to the promoter of *ureA* (Dosanjh et al., 2007). However, the addition of *Hp*NikR to binding reactions failed to generate a significant change in anisotropy, indicating a lack of DNA binding (Figure 10).

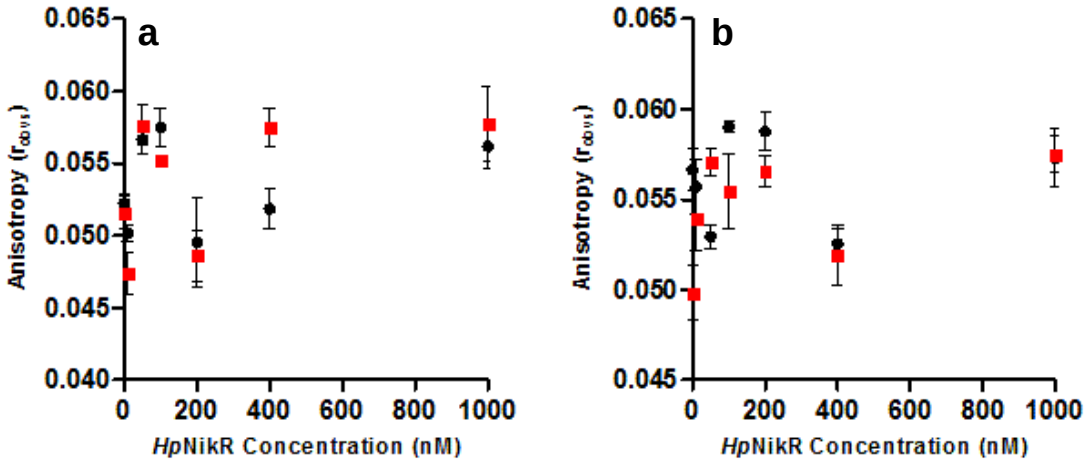


Figure 10. Change in anisotropy (r_{obvs}) following the titration of wild type *Hp*NikR into 20 nM *nixA* (●) and *ureA* (■) promoter fragments labelled with Flu-dT (a) and Cy3 (b). Error bars represent standard error.

This was unexpected, given that DNA probes of the same length, same concentration, with the same fluorescent tag conjugated to the same thymine, and *Hp*NikR of the same concentration as Dosanjh et al. were used. DNA duplexes labelled with an alternative fluorescent probe (Cy3) were tried, but these also failed to generate a change in anisotropy (Figure 10). One way in which the methods differed was in the pre-treating of the cuvette with Rain-X instead of 175 Bloom porcine gelatin. Perhaps Rain-X treatment was insufficient in preventing protein or DNA adherence to the cuvette walls. Attempts to determine the reason for the lack of detectable HpNikR-DNA binding were not continued. Instead, EMSAs were chosen as an alternative method of quantifying HpNikR binding activity due to their reliability and more rapid optimisation.

In the presence of 1.5 mM $MgCl_2$, the addition of 25 mM $NaHCO_3$ did not elicit any significant changes in affinity for *ureA* and *nixA* (Figures 11 & 12, Table 6). Additionally, calculated affinities were approximately 6 – 8-fold weaker than

expected, with prior studies yielding K_d values ranging from 49 – 60 nM for *nixA* and 45 – 67 nM for *ureA* (Benanti and Chivers, 2007; Dosanjh et al., 2009).

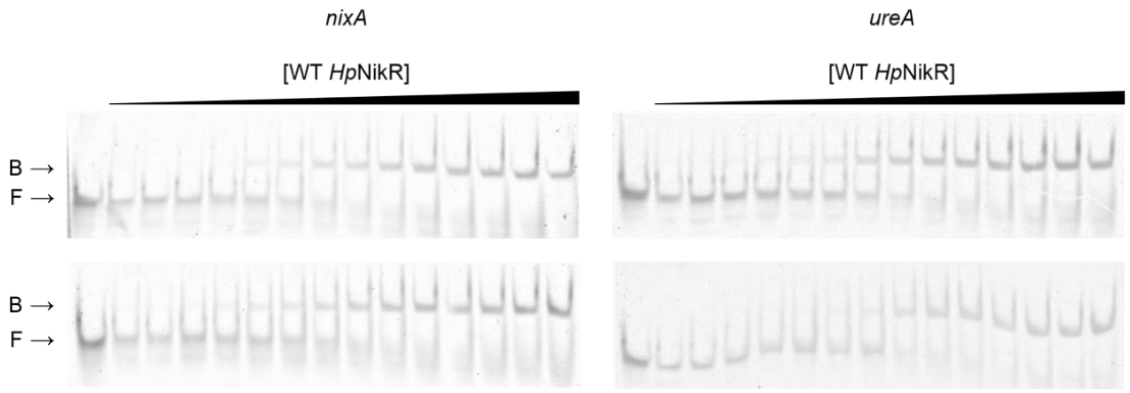


Figure 11. EMSAs of wild type (WT) *Hp*NikR serially diluted 1.18-fold from 1 μM to 120.9 nM. Binding reactions contained 25 nM *nixA* (left) or *ureA* (right) promoter fragments and 1.5 mM $MgCl_2$ with 25 mM $NaHCO_3$ (bottom) or no additional salt (top). The left lane in each titration contains DNA alone. Each titration represents one of two independent replicates. F = free DNA; B = bound DNA.

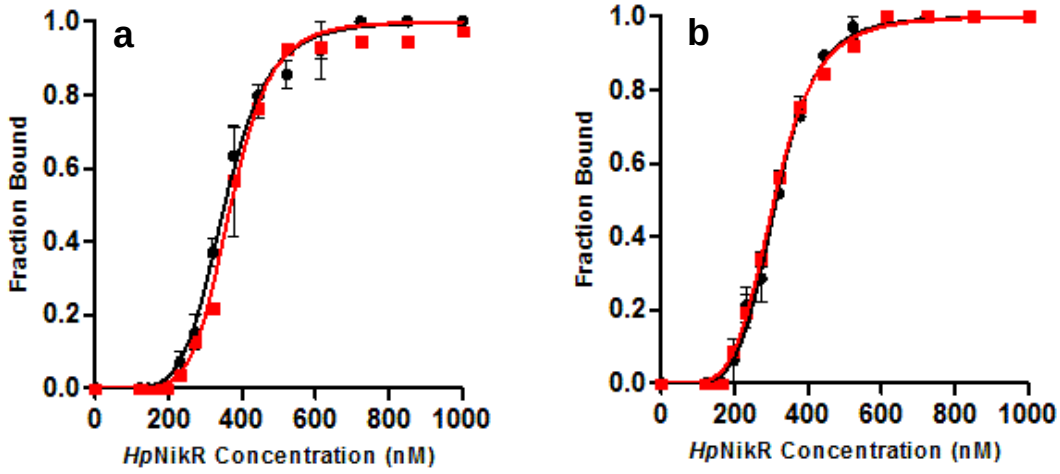


Figure 12. Change in fluorescence intensity (as fraction bound) following the addition of wild type *Hp*NikR to 25 nM *ureA* (a) and *nixA* (b) with 1.5 mM $MgCl_2$ and either 25 mM $NaHCO_3$ (■) or no additional salt (●). An average of two independent replicates is shown, error bars represent standard error, and the data are fit using a non-linear regression analysis and Equation 3.

Table 6. Apparent wild type *Hp*NikR DNA binding affinities (K_d), ΔK_d [K_d (+ $NaHCO_3$) – K_d (- $NaHCO_3$)], and Hill coefficients (n) calculated from the best fits of EMSAs containing 1.5 mM $MgCl_2$ using Equation 3. Values are reported in nM with the standard deviation from the averages of two independent experiments.

	K_d - $NaHCO_3$	K_d + $NaHCO_3$	ΔK_d	n - $NaHCO_3$	n + $NaHCO_3$
nixA	313.8 ± 13.4	307.8 ± 6.3	-6	5.70	5.28
ureA	352.7 ± 23.3	369.2 ± 20.9	16.5	5.88	6.70

Given the negative charge that is introduced as a result of Lys6 carbamylation and the ability of carbamates to coordinate metal ions, it was next hypothesised that the amount of $MgCl_2$ in the binding reactions could potentially be interfering with the visualisation of any effect of the modification on *Hp*NikR binding. Mg(II) plays a crucial role in *Hp*NikR-DNA binding, requiring a concentration of 3 mM for maximal DNA binding affinity attributed to the presence of a low affinity metal binding site that coordinates Mg(II) (Figure 4), in addition to the proposed second Mg(II)-binding site at the N-terminus mediated by Asp7 and Asp8 (Benanti and Chivers, 2007; Dosanjh et al., 2007). To investigate this, $MgCl_2$ concentration was reduced from 1.5 mM to 750 µM. As expected, given the Mg(II) requirement for maximal DNA binding affinity, the reduction in $MgCl_2$ concentration resulted in a 1.77-fold and 2-fold decrease in affinity for *nixA* and *ureA* respectively (Table 7).

Intriguingly, the addition of $NaHCO_3$ under these conditions of reduced $MgCl_2$ concentration resulted in an additional reduction in affinity for both *nixA* and *ureA* (Figures 13 – 15, Table 7). Affinity for *nixA* decreased 1.55-fold, while

affinity for *ureA* decreased 1.80-fold (Table 7). Notably, there was a differential effect, with the decrease in affinity greater for *ureA* and a failure to achieve saturation (Figure 15), meaning the affinity for *ureA* calculated in the presence of HCO_3^-/CO_2 is a lower limit estimate.

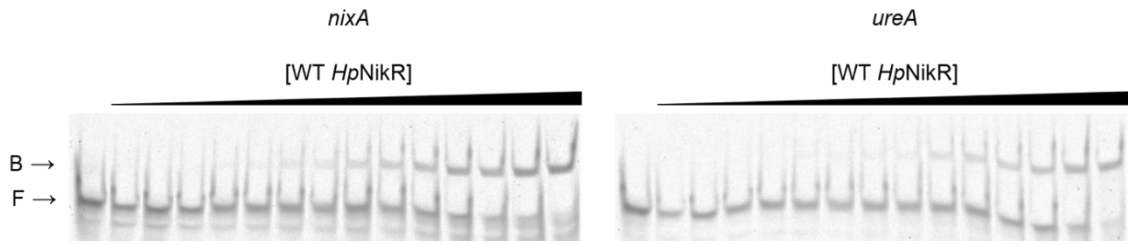


Figure 13. EMSAs of wild type (WT) *Hp*NikR serially diluted 1.18-fold from 1 μM to 120.9 nM. Binding reactions contained 25 nM *nixA* (left) or *ureA* (right) promoter fragments and 750 μM $MgCl_2$. The left lane in each titration contains DNA alone. Each titration represents one of three independent replicates. F = free DNA; B = bound DNA.

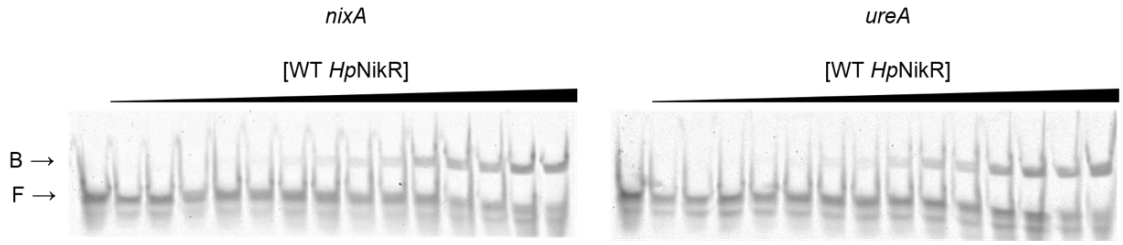


Figure 14. EMSAs of wild type (WT) *Hp*NikR serially diluted 1.18-fold from 1.5 μM to 184.4 nM. Binding reactions contained 25 nM *nixA* (left) or *ureA* (right) promoter fragments, 750 μM $MgCl_2$, and 25 mM $NaHCO_3$. The left lane in each titration contains DNA alone. Each titration represents one of three independent replicates. F = free DNA; B = bound DNA.

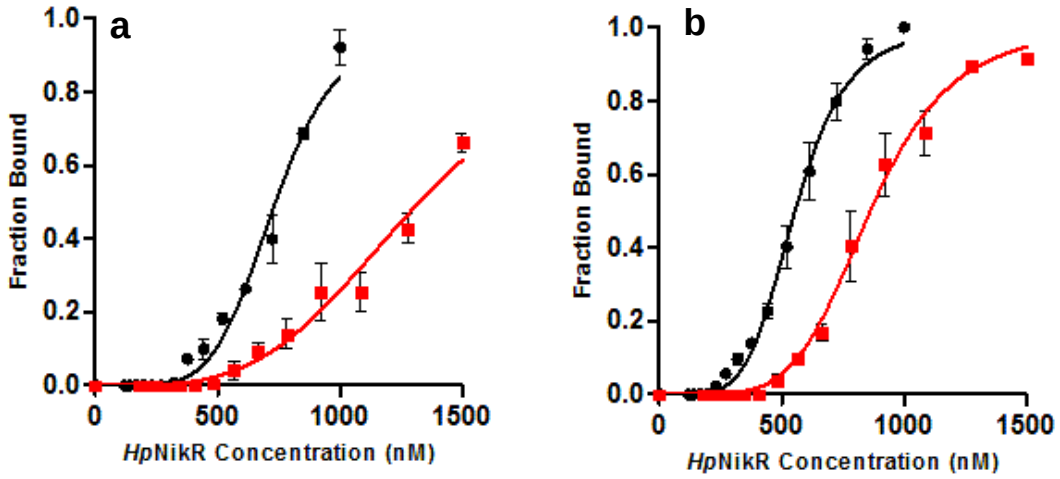


Figure 15. Change in fluorescence intensity (as fraction bound) following the addition of wild type *Hp*NikR to 25 nM *ureA* (a) and *nixA* (b) with 750 μM $MgCl_2$ and either 25 mM $NaHCO_3$ (■) or no additional salt (●). An average of three independent replicates is shown, error bars represent standard error, and the data are fit using a non-linear regression analysis and Equation 3.

Table 7. Apparent wild type (WT) and Lys6Arg (K6R) DNA binding affinities (K_d), ΔK_d [K_d (+ $NaHCO_3$) – K_d (- $NaHCO_3$)], $\Delta\Delta K_d$ (Wild type ΔK_d – Lys6Arg ΔK_d), and Hill coefficients (n) calculated from the best fits of EMSAs containing 750 μM $MgCl_2$ using Equation 3. Values are reported in nM with the standard deviation from the averages of three independent experiments.

	K_d - $NaHCO_3$	K_d + $NaHCO_3$	ΔK_d	$\Delta\Delta K_d$	n - $NaHCO_3$	n + $NaHCO_3$
WT *nixA*	555.2 ± 42.3	862.8 ± 73.6	307.6	-	5.20	5.18
WT *ureA*	736.9 ± 20.6	1325 ± 35.3*	588.1	-	5.47	3.74
K6R *nixA*	824.8 ± 25.6	1035 ± 76.9	210.2	97.4	5.48	5.69
K6R *ureA*	1089 ± 105.4	1308 ± 23.5	219	369.1	4.49	8.07

* = Lower limit estimate.

3.5 Mutation of Lys6 to Arg reduces *Hp*NikR sensitivity to CO_2

To determine whether the observed reduction in DNA binding affinity in response to $NaHCO_3$ addition can be specifically attributed to the CO_2-mediated formation of a carbamate at Lys6 as opposed to a detrimental effect imposed by anion (HCO_3^-) addition, a Lys6Arg mutant was generated. Given the higher pK_a of Arg relative to Lys, the Arg residue in this mutant is expected to exist in the protonated state and therefore be unable to undergo carbamylation. Lys6Arg *Hp*NikR was subjected to the same series of experiments under conditions of reduced Mg(II) as with the wild type.

The affinity of Lys6Arg *Hp*NikR for the tested promoters was 1.49-fold and 1.48-fold weaker for *nixA* and *ureA* respectively, relative to the wild type in the absence of CO_2. This could be attributed to a requirement for a Lys residue in the deprotonated state for maximal DNA binding, or the introduction of noise due to the low concentration of the Lys6Arg *Hp*NikR stock solution, leading to inaccuracies when measuring A_{280} to determine concentration.

Following the addition of $NaHCO_3$, there was a 1.25-fold and 1.20-fold reduction in affinity for *nixA* and *ureA* respectively (Figures 16 & 17, Table 7). While there is still an effect of $NaHCO_3$ addition on K6R *Hp*NikR function, the fold-change in affinity for both promoters is lower in comparison to the wild type. This indicates that while the addition of $NaHCO_3$ appears to still have a detrimental effect on Lys6Arg *Hp*NikR function that could be attributed to the introduction of HCO_3^- anions, there seems to be an additional reduction in affinity exclusive to the wild type that appears to be specific to the CO_2-mediated formation of a carbamate at the intact Lys6. In an attempt to elucidate this, $\Delta\Delta K_d$ was calculated (Table

7). These yielded above-zero values of 97.4 for binding to *nixA* and 369.1 for binding to *ureA* (Table 7), indicative of a reduction in binding affinity that can be attributed specifically to the Lys6 carbamate, and affects binding to the *ureA* promoter more than to the *nixA* promoter. Therefore, the formation of a carbamate at Lys6 appears to be functionally significant, resulting in a differential reduction in binding affinity for both promoters.

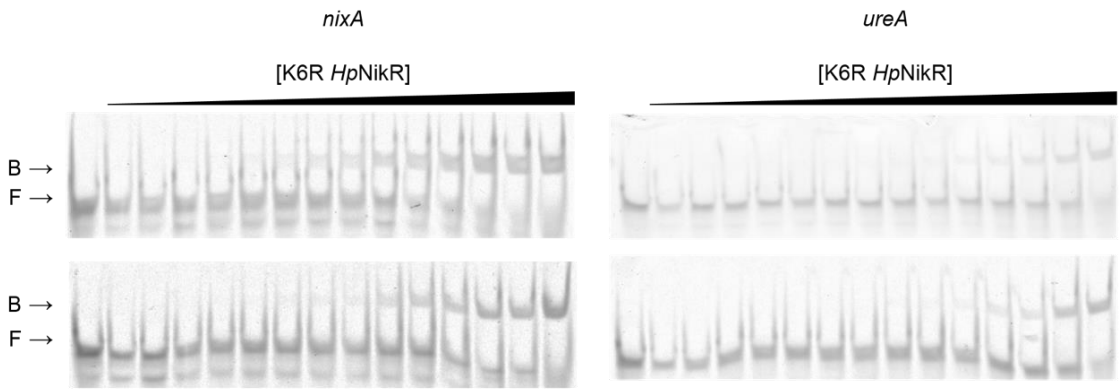


Figure 16. EMSAs of Lys6Arg (K6R) *Hp*NikR serially diluted 1.18-fold from 1.5 μM to 184.4 nM. Binding reactions contained 25 nM *nixA* (left) or *ureA* (right) promoter fragments and 750 μM $MgCl_2$ with 25 mM $NaHCO_3$ (bottom) or no additional salt (top). The left lane in each titration contains DNA alone. Each titration represents one of three independent replicates. F = free DNA; B = bound DNA.

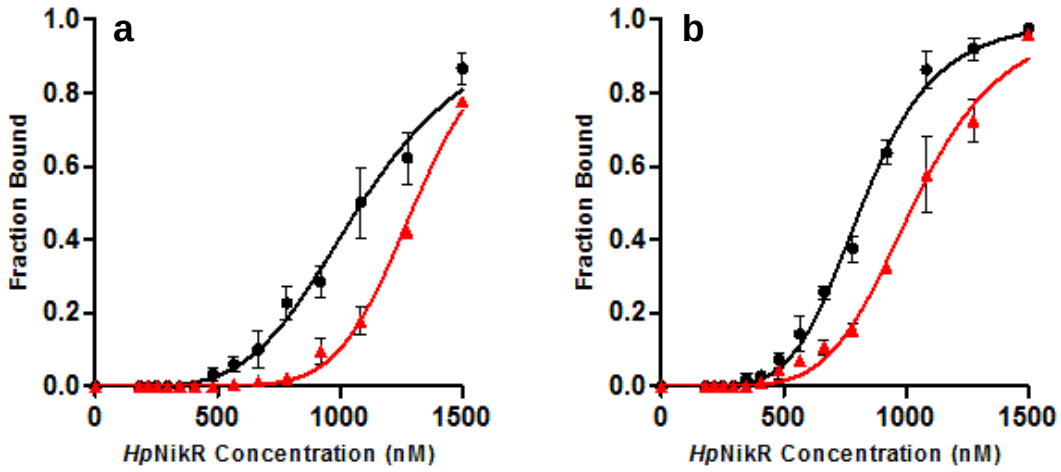


Figure 17. Change in fluorescence intensity (as fraction bound) following the addition of Lys6Arg *Hp*NikR to 25 nM *ureA* (a) and *nixA* (b) with 750 μM $MgCl_2$ and either 25 mM $NaHCO_3$ (▲) or no additional salt (●). An average of three independent replicates is shown, error bars represent standard error, and the data are fit using a non-linear regression analysis and Equation 3.

An additional unusual component of the data collected were the unexpectedly high *n* values for all tested conditions, ranging from 3.74 to 8.07 (Tables 6 & 7). The DNA binding behaviour of *Hp*NikR would usually be expected to result in *n* values approximately equal to one. In an attempt to account for these erroneously high *n* values, the quantification methodology was altered such that individual free and bound bands were selected and the Measure tool used to generate integrated density values that correspond to the sum of all the pixel intensities in each band, rather than manually defining the peaks corresponding to each band from an intensity histogram of the entire lane. Integrated density values were then substituted into Equation 2 to generate a series of fraction bound values. However, the resulting plots were poor representations of what could be visually observed on the corresponding gels, generating erroneously high fraction bound values at low *Hp*NikR concentrations. While it is important that the cause of these unexpectedly high *n* values is determined, they cannot be attributed to the original quantification method.

Chapter 4: Discussion

The research comprising this thesis sought to investigate the role of *Hp*NikR as a candidate CO_2 sensor in *H. pylori*. Three lines of evidence suggest that *Hp*NikR may indeed behave as sensor that couples the transcription of *ureA* and *nixA* to the environmental cue CO_2, using the formation of a carbamate at Lys6 as a mechanism to achieve this. Carbamate formation occurs exclusively at Lys6 (Figure 7), with a strict conservation of the Lys6 residue among *Hp*NikR homologues across different strains of *H. pylori*, despite the significant variability that exists in both length and composition within the N-terminal arm (Figure 8, Tables 3 & 4). While the N-terminal arm may be an adaptable structural feature that is tailored to the unique physiologies of different microorganisms (Benanti and Chivers, 2007), the role that Lys6 plays at the *Hp*NikR-DNA interface is evidently one that cannot be compromised. Furthermore, the lack of conservative replacements for an amino acid of similar physicochemical properties, such as Arg, indicates that the role of this residue likely extends beyond the basic (pK_a) properties of Lys, which is consistent with the verified formation of a carbamate at this site and therefore the propensity of this residue to exist in the neutral rather than protonated state. While the strict conservation of Lys6 is lost in NikR homologues among gastric, enterohepatic, human colonising and non-human colonising *Helicobacter* species, the majority of N-terminal arms do contain at least one Lys residue (Figure 9, Table 5). However, there were exceptions for the N-terminal arm sequences of *H. bizzozeronii*, *H. suis*, and *H. muridarum* (Table 5). *H. bizzozeronii* is a zoonotic *Helicobacter* species that colonises the gastric mucosa of a broad range of hosts in addition to their natural canine host, including humans, cats, cheetahs, non-human primates, and wild rats (Schott et al., 2011; Whary and Fox, 2004).

H. suis is also a zoonotic *Helicobacter* species that colonises the gastric mucosa of humans in addition to their natural porcine host (Padra et al., 2019; Solnick and Schauer, 2001; Whary and Fox, 2004). *H. muridarum*, on the other hand, is an enterohepatic *Helicobacter* species that colonises the intestinal mucosa of rodents (Lee et al., 1992; Solnick and Schauer, 2001; Whary and Fox, 2004). The Lys-Arg substitutions observed in the N-terminal arm sequences of *H. bizzozeronii* and *H. suis* (Table 5) may reflect a lack of a requirement for a CO_2-responsive residue at the DNA binding interface, perhaps evolved as an adaptation to the specific physiology of their natural hosts. Alternatively, these substitutions may have simply resulted from sequencing errors, for example a single adenine to guanine change which could result in a change from a Lys codon (AAA, AAG) to an Arg one (AGA, AGG). The total lack of Lys or Arg residues in the N-terminal arm sequence of the *H. muridarum* (Table 5) may similarly reflect an adaptation to the specific environment within the host it colonises, in this case the rodent intestinal mucosa. *H. muridarum* is urease-positive, although the role of urease in enterohepatic *Helicobacter* pathogenesis remains poorly understood (Ge et al., 2008; Lee et al., 1992). It is possible that, although urease-positive, the upregulation of the production of catalytically active urease in *H. muridarum in vivo* is not essential for survival, given the likelihood that pH is less of a significant environmental stressor encountered by this pathogen in the less acidic intestine. This may account for the loss of a CO_2-responsive residue at the DNA binding interface, given the reduction in CO_2 production that may occur in this organism should basal urease levels be significantly lower in comparison to gastric *Helicobacter* species that require high urease expression for survival. Notably, in the enterohepatic *Helicobacter hepaticus*, which colonises the

intestine, bile ducts, and liver of its murine host, urease expression was found to be regulated independently of nickel availability, pH, and NikR. Instead, expression is regulated in accordance with iron availability, mediated by the ferric uptake regulator, Fur (Belzer et al., 2006). The physiological purpose of this regulatory mechanism is unknown, but proposed to be a means by which *H. hepaticus* may indirectly sense an unfavourable environment (via low iron availability) that requires an upregulation of urease activity to aid in colonisation, albeit in the context of using ammonia production to assimilate nitrogen rather than acid adaptation (Belzer et al., 2006). On the other hand, the gastric *Helicobacter mustelae*, *Helicobacter acinonychis*, and *Helicobacter felis* which colonise ferrets, big cats, and cats, respectively, contain two urease gene clusters (Stoof et al., 2009). In *H. mustelae*, these two urease systems, UreAB and UreA2B2, are inversely regulated by nickel, with UreA2B2 also induced upon supplementation of iron (Stoof et al., 2009). NikR interacts directly with both promoters to mediate this regulation, with the additional iron-responsive urease gene cluster proposed to be an adaptation to the nickel-poor, iron-rich environment expected within a host that consumes a carnivorous diet (Stoof et al., 2009). This evidences the differential roles of urease in host adaptation, with its function extending beyond its involvement in nickel metabolism and acid tolerance that is essential in the survival of gastric *Helicobacter* species. Instead, urease activity in enterohepatic *Helicobacter* species may be more closely linked to nitrogen metabolism, tailored to niches that, in the intestine for instance, are flourishing with the host's native flora that serve as competitors for nitrogen sources, in comparison to the stomach that is inhospitable to such sources of competition (Belzer et al., 2006; Stoof et al., 2009). It therefore

follows that with these differential roles of urease will come differential environmental cues by which the expression of this enzyme is regulated.

As a whole, Lys is among the few amino acids that are consistently present in almost all of the N-terminal arm sequences of NikR homologues across the *Helicobacter* genus (Table 5), unusual given the significant variability in amino acid composition that exists in this sequence and possibly indicative of a widespread conservation of the potential role of Lys in CO_2 sensing. Normally, the involvement of Lys and Arg residues in DNA binding affinity entails the formation of contacts with the negatively charged phosphate backbone mediated by the positive charge of Arg and the positive charge of Lys when located in an environment that does not promote deprotonation. The evolution of structurally privileged sites that give rise to neutral, CO_2-responsive Lys residues that may underpin CO_2 sensing is an adaptation that appears to have been developed by some, but not all, *Helicobacter* NikR variants that may have evolved in conjunction with the specific physiologies and urease requirements of each organism. Given the lack of crystal structure to date demonstrating a well-resolved N-terminal arm (Figure 4), the specific structural features within the microenvironment surrounding Lys6 that would lead to sufficient lowering of its pK_a to promote deprotonation are unknown. In other proteins within which carbamates have been characterised, containment of the Lys within a hydrophobic environment or close proximity of the Lys an Arg residue appears to facilitate the neutral form of the side chain (Golemi et al., 2001; Morollo et al., 1999; Takayama et al., 2008). For example, the side chain of Lys70 in *P. aeruginosa* β-lactamase is sequestered by five residues within the active site, creating a hydrophobic environment that favours the uncharged amine (Golemi

et al., 2001). The carbamylated Lys129 of *B. stearothermophilus* alanine racemase, on the other hand, is positioned in close proximity to Arg136, which is proposed to facilitate deprotonation (Morollo et al., 1999). In *Hp*NikR, Lys6 is unlikely to be positioned in close enough proximity to Arg12 to promote deprotonation, though the N-terminal arm may fold in such a way as to create a hydrophobic environment that facilitates a neutral amine, especially given the presence of Pro4 in the arm which may introduce unusual peptide bond angles (Benanti and Chivers, 2011). However, a crystal structure of *Hp*NikR including a fully-resolved N-terminal arm would be required to elucidate the exact structural features that create this privileged environment.

EMSAs with wild type and Lys6Arg *Hp*NikR, conducted as a means to investigate the functional significance of the Lys6 carbamate, revealed that the introduction of CO_2 via the addition of $NaHCO_3$ leads to a Mg(II)-dependent, differential reduction in binding affinity for the promoters of both *ureA* and *nixA* that can be attributed to the formation of a carbamate at Lys6 (Figures 11 – 17, Tables 6 & 7). As carbamylation results in the introduction of a negative charge, the observed reduction in binding affinity may be a consequence of electrostatic forces of repulsion between the carbamate and the negatively charged phosphodiester backbone and/or adjacent negatively charged Asp residues. Carbamates can behave as ligands for metal ions, as has been observed in the structures of RuBisCO [Mg(II)], *K. aerogenes* urease [Ni(II)] and *P. shermanii* transcarboxylase 5S [Co(II)] (Cleland et al., 1998; Hall et al., 2004; Jabri et al., 1995; Lorimer et al., 1976). This may explain why conducting EMSAs at a reduced Mg(II) concentration resulted in an observable effect of the Lys6

carbamate on *Hp*NikR binding affinity, with the higher Mg(II) concentration potentially leading to total masking of the effects of the negative charge.

The differential effect of the Lys6 carbamate on *Hp*NikR binding affinity could potentially reflect a sophisticated mechanism of transcriptional regulation, whereby the degree to which genes are regulated in response to CO_2 depends on the role the expressed proteins play in the context of CO_2 in *H. pylori* physiology. Given that *ureA* is required for the synthesis and activation of urease, it makes sense that the transcription of this gene may be tightly coupled to CO_2 availability, potentially forming a feedback mechanism linking urease expression to the CO_2 produced as a result of its activity. As CO_2 concentration within the cytoplasm increases, indicative of increasing urease activity, carbamylation of Lys6 may serve as a mechanism to inhibit the *Hp*NikR-mediated activation of *ureA* transcription (Figures 18 & 19). While this may result in a loss of *ureA* activation and a subsequent decrease in urease activity, there are also other environmental signals, pH and Ni(II) availability, in effect that also contribute to urease regulation. How Lys6 carbamylation modulates *Hp*NikR binding activity in the context of these additional environmental signals needs to be tested. On the other hand, while NixA, encoded by *nixA*, is not directly linked to CO_2 physiology, it is indirectly involved in that it is required to supply Ni(II) ions to ensure the catalytic activity of urease (Bauerfeind et al., 1996). This could potentially explain the greater effect of CO_2 on affinity for *ureA* relative to *nixA*, given the more direct role urease plays in CO_2 physiology as a likely major contributor to cytoplasmic CO_2 levels. Additionally, given the de-repression of *nixA* transcription that may arise as a result of Lys6 carbamylation, the differential effect on DNA binding may reflect a greater requirement to reduce urease production and subsequent activity

in response to rising CO_2 levels compared to de-repressing Ni(II) uptake, which is less important in the context of reducing urease activity given the requirement for Ni(II) in catalytically active urease. As previously noted, though, other environmental signals additional to CO_2 will likely also be contributing to the overall *Hp*NikR transcriptional response.

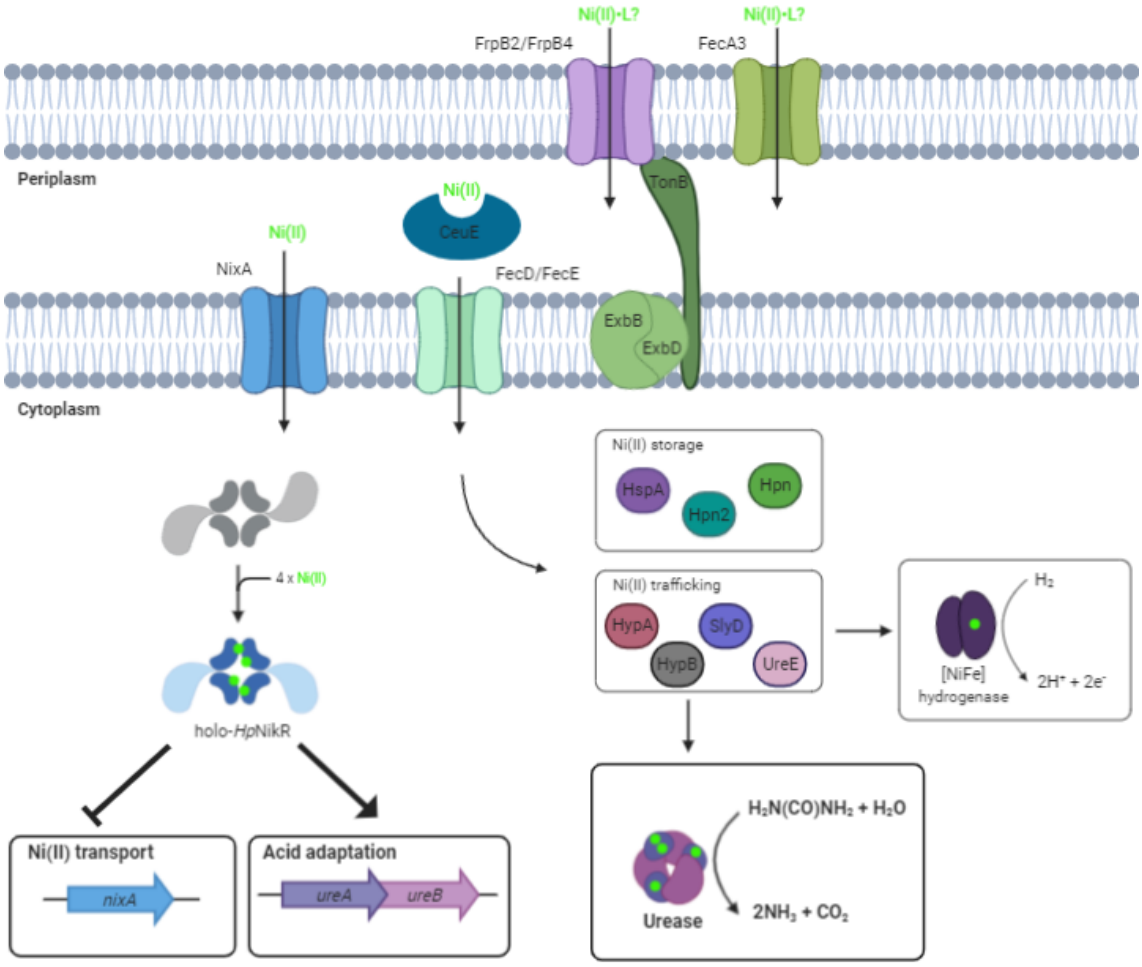


Figure 18. Schematic demonstrating predicted *in vivo* transcriptional regulation by holo-*Hp*NikR under conditions of low CO_2. L represents a potential ligand that may complex Ni(II) for entry into the periplasm via FrpB2/B4 or FecA3, although its existence and identity is not known.

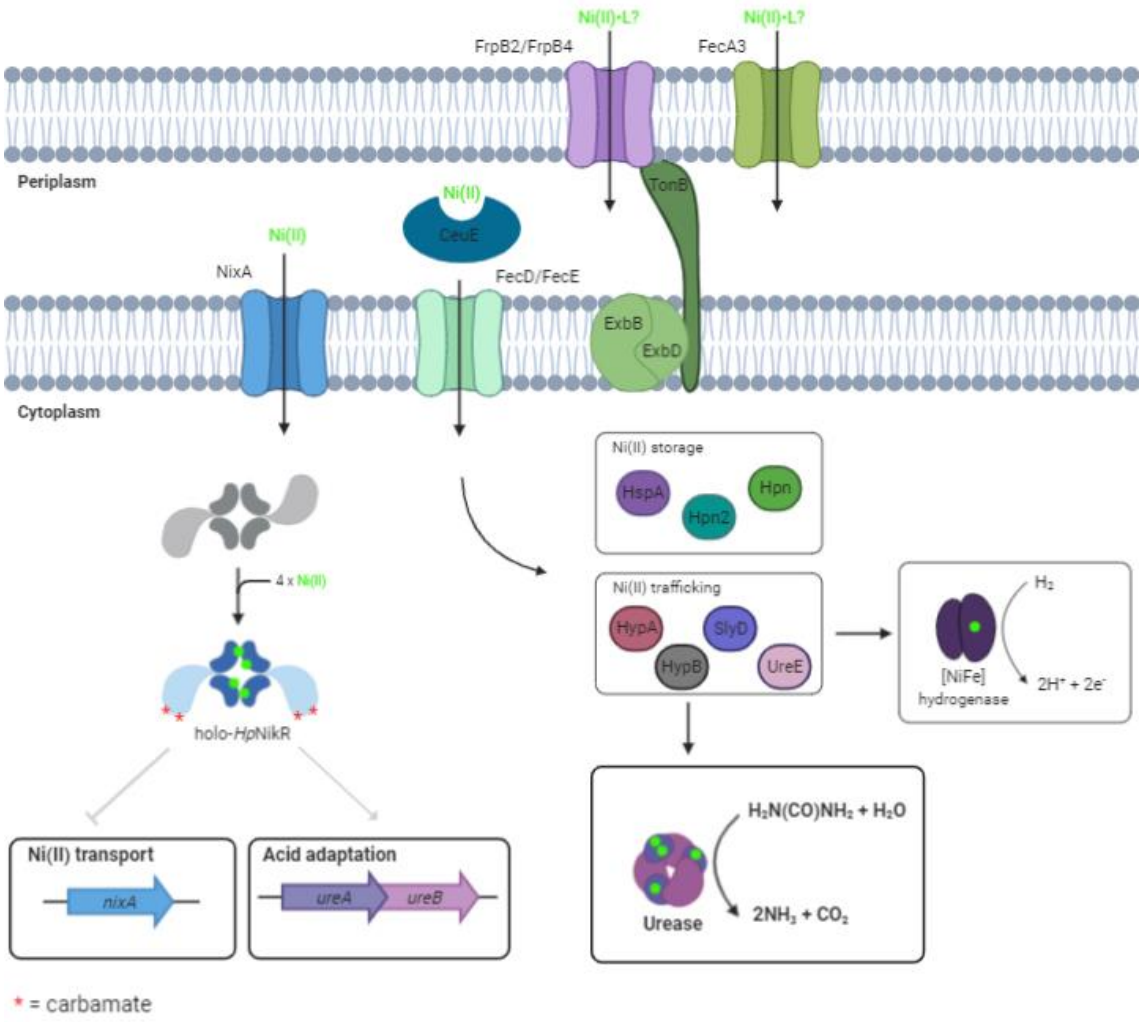


Figure 19. Schematic demonstrating predicted *in vivo* transcriptional regulation by holo-*Hp*NikR under conditions of high CO_2. L represents a potential ligand that may complex Ni(II) for entry into the periplasm via FrpB2/B4 or FecA3, although its existence and identity is not known.

While there is an evident effect of the Lys6 carbamate on *Hp*NikR function that suggests that *Hp*NikR may use this modification as a mechanism to couple CO_2 availability to gene transcription, substantial experimental work and additional assay optimisation is required to further elucidate the validity of this protein as a CO_2 sensor in *H. pylori*.

The cause of the unexpectedly high *n* values requires investigation. Non-cooperative binding behaviour, as is the case with *Hp*NikR binding to DNA in a 1:1 ratio, would result in *n* values approximately equal to one. However, the data

gathered during this research project yielded *n* values substantially greater than one (Tables 6 & 7), suggesting positive cooperative binding (Figure 20).

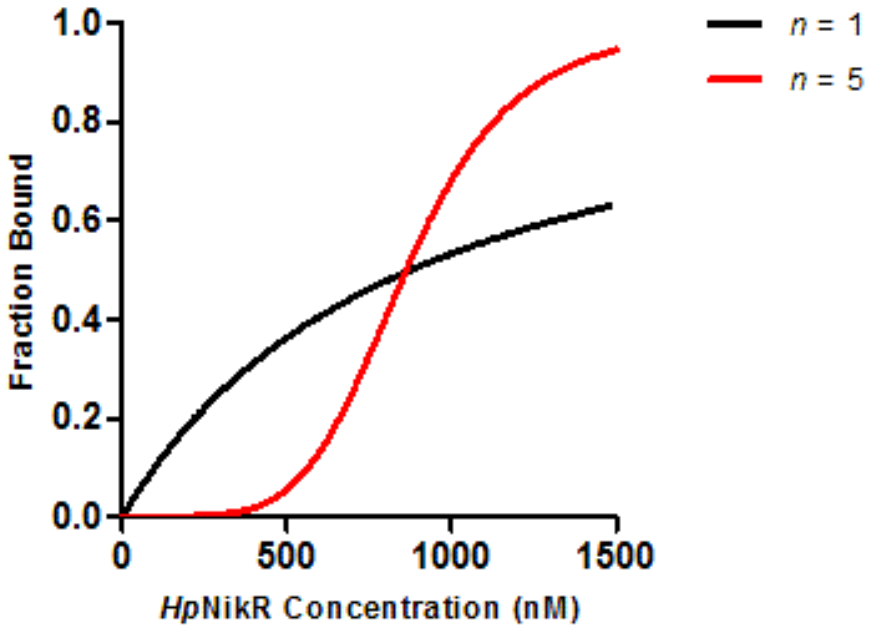


Figure 20. Non-cooperative binding behaviour (*n* = 1) compared to positive cooperative binding behaviour (*n* > 1).

This apparent positive cooperativity is unexpected and cannot be attributed to the binding of multiple *Hp*NikR tetramers to one DNA duplex, given the presence of only one species of reduced mobility in the gels. Were more than one tetramer binding, several shifted species of decreasing mobility depending on the number of tetramers bound would be expected. EMSA binding reactions were set up such that the concentrations of all variables, with the exception of *Hp*NikR, remained equal, so the apparent positive cooperativity cannot be attributed to the titration of an additional variable such as Mg(II), DNA or $NaHCO_3$. Altering the quantification method also failed to account for the unusual *n* values (Section 3.5). The use of EMSAs for this experimental work is limited in its requirement for high DNA concentrations in the nM – µM range [compared to concentrations in

the pM – μM range used in previous work (Benanti and Chivers, 2007)] in order to visualise the bands, which may be a potential cause of the apparent positive cooperativity. Nonetheless, the observed mobility shifts can certainly be attributed to specific interactions between *Hp*NikR and the *ureA* and *nixA* promoters as opposed to aggregation or non-specific binding. Negative control experiments conducted previously using a fragment of the *H. pylori rpoD* gene demonstrated specific binding of *Hp*NikR to its target promoters, with no detectable binding to the *rpoD* fragment (Benanti and Chivers, 2007). Aggregation would lead to the appearance of smears, ladders, or sticking of the DNA in wells as opposed to discrete bands (Hellman and Fried, 2007).

It is possible that the Flu-dT probe used to fluorescently label the DNA duplexes is responsible for the apparent positive cooperativity. In both promoters, the Flu-dT probe is conjugated to a thymine within the central spacer region of the promoter (Table 2). The spacer region separates the two poorly conserved, imperfect inverted repeat half-sites and has been established to be critical in sequence-specific DNA binding as well as the most significant contributor to DNA affinity (Benanti and Chivers, 2011). Unfortunately, the only crystal structure to date depicting *Hp*NikR complexed with DNA does not show the N-terminal arms in their entirety (Figure 4), therefore the precise location of the N-terminal arm residues with respect to the central spacer region cannot be determined. However, previous experiments with Thr3Cys, Asn5Cys, and Asn20Cys *Hp*NikR covalently modified with Fe-BABE to bestow these variants with nuclease activity revealed cleavage sites, indicative of close proximity (within 12 angstroms) of the Fe-BABE binding site to DNA such that cleavage can occur, which were mapped onto the promoters of *ureA* and *nixA* (Benanti and Chivers, 2011). The Flu-dT-

modified thymine residues of both promoters are adjacent to adenine and thymine residues that are included within these previously established cleavage sites mediated by Fe-BABE-modified residues both within the N-terminal arm (Thr3Cys, Asn5Cys) and the α1 helix of the DNA binding domain (Asn20Cys). Therefore, if the Flu-dT probe were capable of folding into the centre of the *Hp*NikR-DNA complex, this may result in unfavourable interactions with the N-terminal arm and/or α1 helix that could account for the unusual binding behaviour (Figure 21).

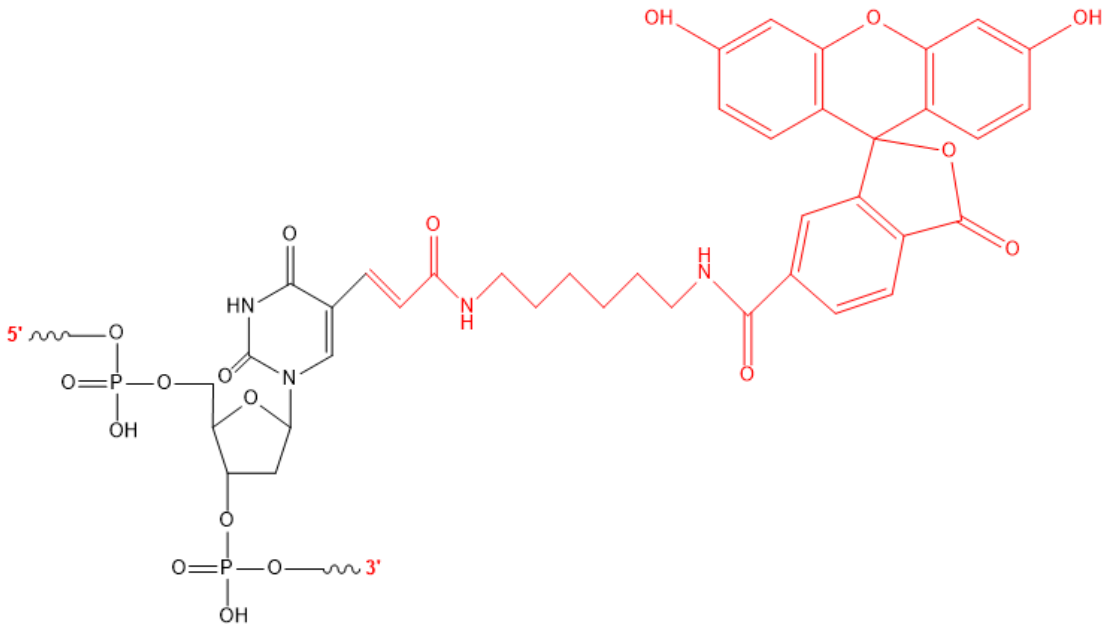


Figure 21. The structure of Flu-dT. Figure created using ChemDraw 19.1 (Perkin Elmer).

To test this, DNA binding assays could be conducted using DNA labelled with fluorescent probes conjugated to a base located at the end of the sequence. Should the central probe be interfering with DNA binding, the expectation would be that assays conducted with end-labelled DNA would reveal non-cooperative binding activity and stronger binding affinity.

An additional possible explanation for the apparent positive cooperativity could be the presence of insufficiently metallated *Hp*NikR tetramers within the protein stocks. Insufficient metallation would result in weaker affinity for DNA, however as metallated *Hp*NikR begins to bind DNA, this would result in an increase in the affinity of unoccupied Ni(II) sites for Ni(II), resulting in their subsequent saturation and an increase in the affinity of the tetramer for DNA. To verify this, Ni(II) could be added to protein stocks prior to conducting DNA binding assays. If the presence of under-metallated *Hp*NikR tetramers is responsible for the unexpected *n* values, the addition of Ni(II) would ensure saturation, resulting in a decrease in K_d, perhaps closer to values that have been obtained in previous studies (Benanti and Chivers, 2007; Dosanjh et al., 2009), as DNA binding affinity increases and a decrease in *n* towards one, consistent with non-cooperative binding.

It is also important in future work that a more reliable and robust fluorescence anisotropy assay is established. While EMSAs were chosen for this experimental work for their reliability and rapid optimisation, they are limited by the inability to conduct these assays within an anaerobic chamber where atmospheric CO_2 concentration can be controlled. This makes the maintenance of CO_2 concentrations within binding reactions a challenge, as dissolved CO_2 will escape from solution as an equilibrium is established between the CO_2 dissolved in solution and the CO_2 in the gas phase as per Henry's law (Avishay and Tenny, 2020), though attempts were made to minimise the loss of CO_2 from solution by adding $NaHCO_3$ to buffers immediately prior to setting up experiments (Section 2.10). Additionally, blemishes on gels are capable of introducing noise to intensity histograms, which can impede accurately defining the peaks corresponding to

each band. DNA/*Hp*NikR-DNA complex sticking to the sides of the wells created trails that introduced noise, in addition to smearing at the bottom of the lanes and the presence of an additional band of low molecular weight relative to all species present in the gel. As this band was also present in the DNA only lane, it is possible that the DNA stock was contaminated with DNase, resulting in the generation of a smaller DNA fragment. As a result of this, intensity histograms were subject to noise, which introduced some difficulty in confidently defining peaks and potentially impeding the accuracy of the subsequent fraction bound values.

Fluorescence anisotropy, on the other hand, is a solution-based, more efficient assay alternative that produces inherently quantitative data and can be conducted within an anaerobic chamber, meaning the partial pressure and therefore the amount of CO_2 dissolved in solution can be controlled. It is limited in comparison to EMSAs in that cannot determine the number or nature of different complexes that may be forming, however for the purposes of this experimental work it is the superior assay. It is uncertain why *Hp*NikR-DNA binding could not be detected using this method, however testing DNA labelled with a Cy3 probe conjugated to the end of the duplexes failed to elicit any change in anisotropy indicative of binding (Figure 10). Therefore, the lack of detectable binding cannot be attributed to the use of a Flu-dT, centrally located probe.

Beyond assay optimisation, there are several additional *in vitro* experiments that are required to further assess the validity of *Hp*NikR as a CO_2 sensor in *H. pylori*. It is important to determine the precise sensitivity of the observed effect to CO_2, given that this experimental work only tested two conditions – no additional CO_2,

and additional CO_2 introduced via 25 mM $NaHCO_3$. To do this, an extensive titration testing *Hp*NikR binding over a range of $NaHCO_3$ concentrations would be required. Additionally, it would be useful to conduct carbamate trapping over a range of $NaHCO_3$ concentrations, including a $NaHCO_3$-free control, to determine the minimum concentration at which carbamate formation occurs. Structure determination across a range of $NaHCO_3$ concentrations could also be useful in determining what proportion of *Hp*NikR monomers within each tetramer forms carbamates under these conditions, similarly to Golemi et al., who determined that varying proportions of *P. aeruginosa* β-lactamase subunits contained Lys70 carbamates under differing pH conditions (Golemi et al., 2001). It is also likely that 25 mM HCO_3^- is beyond that which is physiologically relevant within the *H. pylori* cytoplasm. Although the precise concentrations of HCO_3^- and CO_2 within the *H. pylori* cytoplasm are not known, assuming the concentration of $NH_2(CO)NH_2$ encountered by *H. pylori in vivo* is the same as that which has been found in serum [1.7 – 3.4 mM, (Mobley, 2001)], this means the concentration of HCO_3^- and subsequent CO_2 produced by urease [one molecule of $NH_2(CO)NH_2$ produces one molecule of CO_2 in the urease reaction (Ansari and Yamaoka, 2017)] in the cytoplasm will probably be much lower than 25 mM (Blusiewicz et al., 2005; Jones et al., 2018). This may have contributed to the decreases in binding affinity observed in Lys6Arg *Hp*NikR upon the addition of $NaHCO_3$ that could be attributed to the presence of an excess of HCO_3^- anions having a detrimental effect on *Hp*NikR function.

An alternative possibility is that that the conformational changes elicited within *Hp*NikR upon DNA binding may create additional privileged Lys sites that are susceptible to carbamylation, for example Lys48, which may contribute to the

observed effect of CO_2 on DNA binding (Benanti and Chivers, 2011). If this were the case, the absence of DNA in the carbamate trapping reactions means that these privileged sites cannot be created, and the subsequent carbamates therefore not detected. The possible presence of additional carbamates would account for the decrease in DNA binding affinity that still occurs with Lys6Arg *Hp*NikR, however this would require verification, potentially using the carbamate trapping methodology but with the addition of DNA to encourage the formation of any additional privileged Lys sites. Testing lower $NaHCO_3$ concentrations in trapping reactions and DNA binding assays will also help to determine sensitivity and may improve the analysis of the specific effect of the Lys6 carbamate on binding activity by minimising the potential detrimental effects of excess HCO_3^- anions if a lower concentration is sufficient to induce carbamylation.

Similarly, conducting an extensive Mg(II) titration will be useful in determining the exact dependence of the observed effect on the availability of Mg(II). Furthermore, given the promiscuity of *Hp*NikR binding activity and the implication from this experimental work that binding to different promoters is differentially affected by the Lys6 carbamate, it is essential that a wider selection of *Hp*NikR binding targets are tested to gain a broader insight into the extent to which the Lys6 carbamate serves to couple CO_2 availability to transcription across the regulon. In addition to the pleiotropic regulatory activity of *Hp*NikR, this protein has been established to be responsive to the environmental cues of Ni(II) availability and pH. *In vitro*, the likelihood is that these signals will be simultaneously providing input into the regulatory activity of *Hp*NikR. Therefore, it would be useful to determine the CO_2-responsive activity of *Hp*NikR in the

context of these additional environmental cues, and whether these signals behave synergistically or otherwise with respect to gene regulation.

More broadly, the effect of the Lys6 carbamate on the wider physiology of *H. pylori* could be determined *in vivo* with the use of the Lys6Arg *Hp*NikR variant, although it is important to first verify that this mutant retains normal Ni(II)-dependent regulatory properties. Once this has been established, Lys6Arg *Hp*NikR would then be a useful tool to test *in vivo* as a carbamylation insensitive mutant to determine whether the observed biochemical variation in DNA binding leads to physiological variation. As a starting point, Lys6Arg *Hp*NikR could be introduced into *H. pylori*, and the transcription levels of *nixA* and *ureA* measured using quantitative reverse transcription-PCR. These quantified transcription levels could then be compared between wild type and Lys6Arg *H. pylori* under varying CO_2 conditions to determine whether the apparent reduction in CO_2-responsive binding activity observed *in vitro* is reflected *in vivo*. Should this be the case, when CO_2 is increased, wild type *ureA* transcription would be expected to be reduced relative to conditions of decreased CO_2, in concert with a relative increase in *nixA* transcription, as Lys6 undergoes carbamylation, inhibits binding of *Hp*NikR to its target promoters, and subsequently inhibits *ureA* upregulation and *nixA* repression. In the Lys6Arg mutant, on the other hand, the CO_2-mediated changes in *ureA* transcription and increase in *nixA* transcription would be expected to be diminished, although to a greater extent with *ureA* transcription. Western blots could also be conducted and densitometric data generated to quantify any changes in protein expression and determine whether the observed changes in *nixA* and *ureA* transcription are reflected downstream in changes in the levels of NixA and urease. Beyond transcriptional effects, wild type and

Lys6Arg *H. pylori* could be grown as liquid cultures under varying CO_2 conditions, optical density measured over time, and growth curves plotted to determine the effects, if any, of de-coupling CO_2 responsivity from *Hp*NikR on growth rate. Additionally, the two strains could be co-cultured under varying CO_2 conditions to elucidate any competitive advantages granted by CO_2 responsivity. To determine any effects on virulence, Mongolian gerbils [robust and cost-effective rodent models that mimic several features of human *H. pylori* infections, including gastritis and carcinogenesis (Noto et al., 2018)] could be inoculated with wild type and Lys6Arg variants of *H. pylori*, followed by measurement and comparison of colonisation loads to determine any differences in colonisation efficiency, as described previously (Bury-Moné et al., 2004). Furthermore, gastric pathology could be compared using gastric histologic activity indexes, generated by combining scores for criteria which include inflammation, epithelial defects, oxyntic atrophy, hyperplasia and dysplasia (Ge et al., 2018).

A reduction in the ability of *H. pylori* to regulate *ureA* expression and subsequent urease activity in conjunction with the key indicator of its activity could lead to a pathological over-production of urease beyond that which is required for acid adaptation, potentially resulting in detrimental increases in pH, uncontrolled production of ammonia, and subsequent impairments in growth rate and virulence. However, *Hp*NikR is not the sole transcriptional regulator underpinning acid adaptation in *H. pylori*, nor is CO_2 the primary environmental cue *Hp*NikR coordinates its activity in conjunction with, so it's possible that other transcriptional regulators and/or the responsiveness of *Hp*NikR to pH could compensate and minimise any potential deleterious effects. Furthermore, a concurrent reduction in CO_2-mediated *nixA* de-repression may in turn lead to

inhibition of Ni(II) uptake, which may serve to prevent the metallation and subsequent catalytic activity of a pathological excess of urease. Considerable *in vivo* experimental work will be vital in fully elucidating the physiological significance of the observed biochemical variations in *Hp*NikR-DNA binding, and whether the loss of CO_2 responsiveness proves detrimental in the context of the whole organism.

In summary, Lys6 of *Hp*NikR is a site of carbamate formation. The formation of this carbamate appears to differentially affect the binding of *Hp*NikR to the promoters of *ureA* and *nixA*. This implies that *Hp*NikR may behave as a CO_2 sensor in *H. pylori*, coupling the availability of CO_2 to gene transcription and using the carbamylation of Lys6 as a mechanism to achieve this. However, substantial experimental work is required to further investigate the validity of this potential novel role. The restrictions imposed by lockdown due to COVID-19 have resulted in a considerable loss of time that would otherwise have been dedicated to addressing these important questions.

Chapter 5: References

Abraham, L.O., Li, Y., and Zamble, D.B. (2006). The metal- and DNA-binding activities of *Helicobacter pylori* NikR. *Journal of Inorganic Biochemistry, Special Issue Containing Contributions from the 12th International Conference on Biological Inorganic Chemistry.* **100**, 1005–1014. https://doi.org/10.1016/j.jinorgbio.2005.10.014

Abraham, S.J., Kobayashi, T., Solaro, R.J., and Gaponenko, V. (2009). Differences in Lysine pK_a Values May Be Used to Improve NMR Signal Dispersion in Reductively Methylated Proteins. *Journal of Biomolecular NMR.* **43**, 239–246. https://doi.org/10.1007/s10858-009-9306-2

Abuaita, B.H., and Withey, J.H. (2009). Bicarbonate Induces *Vibrio cholerae* Virulence Gene Expression by Enhancing ToxT Activity. *Infection and Immunity.* **77**, 4111–4120. https://doi.org/10.1128/IAI.00409-09

André, I., Linse, S., and Mulder, F.A.A. (2007). Residue-Specific pK_a Determination of Lysine and Arginine Side Chains by Indirect 15N and 13C NMR Spectroscopy: Application to apo Calmodulin. *Journal of the American Chemical Society.* **129**, 15805–15813. https://doi.org/10.1021/ja0721824

Ansari, S., and Yamaoka, Y. (2017). Survival of *Helicobacter pylori* in gastric acidic territory. *Helicobacter.* **22**. https://doi.org/10.1111/hel.12386

Argueta, E.A., and Moss, S.F. (2019). Treatment of *Helicobacter pylori*. *Current Opinion in Gastroenterology.* **35**, 544–550. https://doi.org/10.1097/MOG.0000000000000578

Avishay, D.M., and Tenny, K.M. (2020). Henry's Law. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing.

Bauerfeind, P., Garner, R.M., and Mobley, L.T. (1996). Allelic exchange mutagenesis of nixA in *Helicobacter pylori* results in reduced nickel transport and urease activity. *Infection and Immunity.* **64**, 2877–2880. https://doi.org/10.1128/IAI.64.7.2877-2880.1996

Belzer, C., van Schendel, B. A. M., Kuipers, E. J., Kusters, J. G., and van Vliet, A. H. M. (2006). Iron-Responsive Repression of Urease Expression in *Helicobacter hepaticus* Is Mediated by the Transcriptional Regulator Fur. *Infection and Immunity*. **75**, 745–752. doi:10.1128/iai.01163-06

Benanti, E.L., and Chivers, P.T. (2011). *Helicobacter pylori* NikR protein exhibits distinct conformations when bound to different promoters. *Journal of Biological Chemistry.* **286**, 15728–15737. https://doi.org/10.1074/jbc.M110.196055

Benanti, E.L., and Chivers, P.T. (2010). *Geobacter uraniireducens* NikR Displays a DNA Binding Mode Distinct from Other Members of the NikR Family. *Journal of Bacteriology*. **192**, 4327–4336. https://doi.org/10.1128/JB.00152-10

Benanti, E.L., and Chivers, P.T. (2007). The N-terminal Arm of the *Helicobacter pylori* Ni2+-dependent Transcription Factor NikR Is Required for Specific DNA Binding. *Journal of Biological Chemistry*. **282**, 20365–20375. https://doi.org/10.1074/jbc.M702982200

Benanti, E.L. (2009). *The evolution and regulation of DNA-binding by the nickel-dependent transcription factor NikR*. All Theses and Dissertations (ETDs), Washington University. **35**. Available at: https://openscholarship.wustl.edu/etd/35

Benoit, S., and Maier, R.J. (2003). Dependence of *Helicobacter pylori* Urease Activity on the Nickel-Sequestering Ability of the UreE Accessory Protein. *Journal of Bacteriology*. **185**, 4787–4795. https://doi.org/10.1128/JB.185.16.4787-4795.2003

Berg, J.M., Tymoczko, J.L., and Stryer, L. (2002). Making a Fast Reaction Faster: Carbonic Anhydrases. In: *Biochemistry*. 5th edition. New York: W H Freeman.

Blaser, M.J. (1990). *Helicobacter pylori* and the Pathogenesis of Gastroduodenal Inflammation. *Journal of Infectious Diseases*. **161**, 626–633. https://doi.org/10.1093/infdis/161.4.626

Blusiewicz, K., Rydzewska, G., and Rydzewski, A. (2005). Gastric juice ammonia and urea concentrations and their relation to gastric mucosa injury in patients maintained on chronic hemodialysis. *Roczniki Akademii Medycznej Bialymstoku*. **50**, 188–192.

Bury-Moné, S., Kaakoush, N.O., Asencio, C., Mégraud, F., Thibonnier, M., De Reuse, H., and Mendz, G.L. (2006). Is *Helicobacter pylori* a true microaerophile? *Helicobacter*. **11**, 296–303. https://doi.org/10.1111/j.1523-5378.2006.00413.x

Bury-Moné, S., Thiberge, J.M., Contreras, M., Maitournam, A., Labigne, A., and Reuse, H.D. (2004). Responsiveness to acidity via metal ion regulators

mediates virulence in the gastric pathogen *Helicobacter pylori. Molecular Microbiology*. **53**, 623–638. https://doi.org/10.1111/j.1365-2958.2004.04137.x

Chivers, P.T., and Sauer, R.T. (2000). Regulation of High Affinity Nickel Uptake in Bacteria. Ni2+-Dependent Interaction of NikR with Wild-Type and Mutant Operator Sites. *Journal of Biological Chemistry*. **275**, 19735–19741. https://doi.org/10.1074/jbc.M002232200

Chivers, P.T., and Sauer, R.T. (1999). NikR is a ribbon-helix-helix DNA-binding protein. *Protein Science*. **8**, 2494–2500. https://doi.org/10.1110/ps.8.11.2494

Cleland, W.W., Andrews, T.J., Gutteridge, S., Hartman, F.C., and Lorimer, G.H. (1998). Mechanism of Rubisco: The Carbamate as General Base. *Chemical Reviews*. **98**, 549–562. https://doi.org/10.1021/cr970010r

Contreras, M., Thiberge, J.M., Mandrand-Berthelot, M.-A., and Labigne, A. (2003). Characterization of the roles of NikR, a nickel-responsive pleiotropic autoregulator of *Helicobacter pylori*. *Molecular Microbiology.* **49**, 947–963. https://doi.org/10.1046/j.1365-2958.2003.03621.x

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: A Sequence Logo Generator. *Genome Research*. **14**, 1188–1190. https://doi.org/10.1101/gr.849004

Cummins, E.P., Selfridge, A.C., Sporn, P.H., Sznajder, J.I., and Taylor, C.T. (2014). Carbon dioxide-sensing in organisms and its implications for human disease. *Cellular and Molecular Life Sciences*. **71**, 831–845. https://doi.org/10.1007/s00018-013-1470-6

Cummins, E.P., Strowitzki, and M.J., Taylor, C.T. (2020). Mechanisms and consequences of oxygen- and carbon dioxide-sensing in mammals. *Physiological Reviews*. **100**, 463–488. https://doi.org/10.1152/physrev.00003.2019

Delany, I., Ieva, R., Soragni, A., Hilleringmann, M., Rappuoli, R., and Scarlato, V. (2005). *In vitro* analysis of protein-operator interactions of the NikR and fur metal-responsive regulators of coregulated genes in *Helicobacter pylori*. *Journal of Bacteriology*. **187**, 7703–7715. https://doi.org/10.1128/JB.187.22.7703-7715.2005

Dian, C., Schauer, K., Kapp, U., McSweeney, S.M., Labigne, A., and Terradot, L. (2006). Structural Basis of the Nickel Response in *Helicobacter pylori*: Crystal Structures of *Hp*NikR in Apo and Nickel-bound States. *Journal of*

Molecular Biology. **361**, 715–730. https://doi.org/10.1016/j.jmb.2006.06.058

Dosanjh, N.S., and Michel, S.L.J. (2006). Microbial nickel metalloregulation: NikRs for nickel ions. *Current Opinion in Chemical Biology, Bioinorganic Chemistry / Biocatalysis and Biotransformation*. **10**, 123–130. https://doi.org/10.1016/j.cbpa.2006.02.011

Dosanjh, N.S., Hammerbacher, N.A., and Michel, S.L.J. (2007). Characterization of the *Helicobacter pylori* NikR−PureA DNA Interaction: Metal Ion Requirements and Sequence Specificity. *Biochemistry*. **46**, 2520–2529. https://doi.org/10.1021/bi062092w

Dosanjh, N. S., West, A. L., and Michel, S. L. J. (2009). *Helicobacter pylori* NikR's Interaction with DNA: A Two-Tiered Mode of Recognition. *Biochemistry*. **48**, 527–536. https://doi.org/10.1074/jbc.M702982200

Ernst, F.D., Kuipers, E.J., Heijens, A., Sarwari, R., Stoof, J., Penn, C.W., Kusters, J.G., and van Vliet, A.H.M. (2005). The Nickel-Responsive Regulator NikR Controls Activation and Repression of Gene Transcription in *Helicobacter pylori. Infection and Immunity*. **73**, 7252–7258. https://doi.org/10.1128/IAI.73.11.7252-7258.2005

Ernst, F.D., Stoof, J., Horrevoets, W.M., Kuipers, E.J., Kusters, J.G., and van Vliet, A.H.M. (2006). NikR mediates nickel-responsive transcriptional repression of the *Helicobacter pylori* outer membrane proteins FecA3 (HP1400) and FrpB4 (HP1512). *Infection and Immunity*. **74**, 6821–6828. https://doi.org/10.1128/IAI.01196-06

Eusebi, L.H., Zagari, R.M., and Bazzoli, F. (2014). Epidemiology of *Helicobacter pylori* Infection. *Helicobacter*. **19**, 1–5. https://doi.org/10.1111/hel.12165

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science*. **246**, 64–71. https://doi.org/10.1126/science.2675315

Ge, Z., Lee, A., Whary, M.T., Rogers, A.B., Maurer, K.J., Taylor, N.S., Schauer, D.B., and Fox, J.G. (2008). *Helicobacter hepaticus* urease is not required for intestinal colonization but promotes hepatic inflammation in male A/JCr mice. *Microbial Pathogenesis*. **45**, 18–24. https://doi.org/10.1016/j.micpath.2008.02.003

Ge, Z., Sheh, A., Feng, Y., Muthupalani, S., Ge, L., Wang, C., Kurnick, S., Mannion, A., Whary, M. and Fox, J. (2018). *Helicobacter pylori*-infected C57BL/6 mice with different gastrointestinal microbiota have contrasting

gastric pathology, microbial and host immune responses. *Scientific Reports*, **8** (1). https://doi.org/10.1038/s41598-018-25927-2

Golemi, D., Maveyraud, L., Vakulenko, S., Samama, J.P., and Mobashery, S. (2001). Critical involvement of a carbamylated lysine in catalytic function of class D beta-lactamases. *Proceedings of the National Academy of Sciences of the United States of America*. **98**, 14280–14285. https://doi.org/10.1073/pnas.241442898

Gunn, J.S., and Richards, S.M. (2007). Recognition and Integration of Multiple Environmental Signals by the Bacterial Sensor Kinase PhoQ. *Cell Host & Microbe*. **1**, 163–165. https://doi.org/10.1016/j.chom.2007.05.001

Hall, P.R., Zheng, R., Antony, L., Pusztai-Carey, M., Carey, P.R., and Yee, V.C. (2004). Transcarboxylase 5S structures: assembly and catalytic mechanism of a multienzyme complex subunit. *EMBO Journal.* **23**, 3621–3631. https://doi.org/10.1038/sj.emboj.7600373

Han, X., Aslanian, A., and Yates, J.R. (2008). Mass Spectrometry for Proteomics. *Current Opinion in Chemical Biology*. **12**, 483–490. https://doi.org/10.1016/j.cbpa.2008.07.024

Hester, S.E., Lui, M., Nicholson, T., Nowacki, D., and Harvill, E.T. (2012). Identification of a CO_2 Responsive Regulon in *Bordetella*. *PLoS ONE*. **7**, e47635. https://doi.org/10.1371/journal.pone.0047635

Hellman, L. M., and Fried, M. G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. *Nature Protocols*. **2**, 1849–1861. doi:10.1038/nprot.2007.249

Hyde, J.A., Trzeciakowski, J.P., and Skare, J.T. (2007). *Borrelia burgdorferi* Alters Its Gene Expression and Antigenic Profile in Response to CO_2 Levels. *Journal of Bacteriology.* **189**, 437–445. https://doi.org/10.1128/JB.01109-06

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. (1994). Schistosomes, Liver Flukes and *Helicobacter pylori*. In: *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*. Volume. 61. Lyon: International Agency for Research on Cancer.

Iwanaga, M., and Kuyyakanond, T. (1987). Large production of cholera toxin by *Vibrio cholerae* O1 in yeast extract peptone water. *Journal of Clinical Microbiology*. **25**, 2314–2316.

Jabri, E., Carr, M.B., Hausinger, R.P., and Karplus, P.A. (1995). The crystal structure of urease from *Klebsiella aerogenes. Science*. **268**, 998–1004.

Jones, M.D., Li, Y., and Zamble, D.B. (2018). Acid-responsive activity of the *Helicobacter pylori* metalloregulator NikR. *Proceedings of the National Academy of Sciences of the United States of America*. **115**, 8966–8971. https://doi.org/10.1073/pnas.1808393115

Karas, M., and Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry*. **60**, 2299–2301. https://doi.org/10.1021/ac00171a028

Kesvatera, T., Jönsson, B., Thulin, E., and Linse, S. (1996). Measurement and modelling of sequence-specific pK_a values of lysine residues in calbindin D9k. *Journal of Molecular Biology*. **259**, 828–839. https://doi.org/10.1006/jmbi.1996.0361

Khalifah, R.G. (1973). Carbon dioxide hydration activity of carbonic anhydrase: paradoxical consequences of the unusually rapid catalysis. *Proceedings of the National Academy of Sciences of the United States of America*. **70**, 1986–1989. https://doi.org/10.1073/pnas.70.7.1986

Kilmartin, J.V., and Rossi-Bernardi, L. (1971). The binding of carbon dioxide by horse haemoglobin. *Biochemical Journal*. **124**, 31–45.

Krell, T., Lacal, J., Busch, A., Silva-Jiménez, H., Guazzaroni, M.-E., and Ramos, J.L. (2010). Bacterial Sensor Kinases: Diversity in the Recognition of Environmental Signals. *Annual Review of Microbiology.* **64**, 539–559. https://doi.org/10.1146/annurev.micro.112408.134054

Kuhns, L.G., Benoit, S.L., Bayyareddy, K., Johnson, D., Orlando, R., Evans, A.L., Waldrop, G.L., and Maier, R.J. (2016). Carbon Fixation Driven by Molecular Hydrogen Results in Chemolithoautotrophically Enhanced Growth of *Helicobacter pylori*. *Journal of Bacteriology*. **198**, 1423–1428. https://doi.org/10.1128/JB.00041-16

Lee, A., Phillips, M.W., O'Rourke, J.L., Paster, B.J., Dewhirst, F.E., Fraser, G.J., Fox, J.G., Sly, L.I., Romaniuk, P.J., and Trust, T.J. (1992). *Helicobacter muridarum* sp. nov., a microaerophilic helical bacterium with a novel ultrastructure isolated from the intestinal mucosa of rodents. *International Journal of Systematic Bacteriology*. **42**, 27–36. https://doi.org/10.1099/00207713-42-1-27

Li, Y., and Zamble, D.B. (2009). pH-responsive DNA-binding activity of *Helicobacter pylori* NikR. *Biochemistry*. **48**, 2486–2496. https://doi.org/10.1021/bi801742r

Lindskog, S. (1997). Structure and mechanism of carbonic anhydrase. *Pharmacology & Therapeutics*. **74**, 1–20. https://doi.org/10.1016/s0163-7258(96)00198-2

Linthwaite, V.L. (2017). *Development of a Technology for the Discovery of Protein Carbamates.* Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/12029/

Linthwaite, V.L., Janus, J.M., Brown, A.P., Wong-Pascua, D., O'Donoghue, A.C., Porter, A., Treumann, A., Hodgson, D.R.W., and Cann, M.J. (2018). The identification of carbon dioxide mediated protein post-translational modifications. *Nature Communications*. **9**, 3092. https://doi.org/10.1038/s41467-018-05475-z

Livingstone, C.D., and Barton, G.J. (1993). Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. *Computer Applications in the Biosciences*. **9**, 745–756. https://doi.org/10.1093/bioinformatics/9.6.745

Lorimer, G.H. (1983). Carbon dioxide and carbamate formation: the makings of a biochemical control system. *Trends in Biochemical Sciences*. **8**, 65–68. https://doi.org/10.1016/0968-0004(83)90393-6

Lorimer, G.H., Badger, M.R., and Andrews, T.J. (1976). The activation of ribulose-1,5-bisphosphate carboxylase by carbon dioxide and magnesium ions. Equilibria, kinetics, a suggested mechanism, and physiological implications. *Biochemistry*. **15**, 529–536. https://doi.org/10.1021/bi00648a012

Lorimer, G.H., and Miziorko, H.M. (1980). Carbamate formation on the .epsilon.-amino group of a lysyl residue as the basis for the activation of ribulosebisphosphate carboxylase by carbon dioxide and magnesium(2+). *Biochemistry*. **19**, 5321–5328. https://doi.org/10.1021/bi00564a027

Loughlin, M.F. (2003). Novel therapeutic targets in *Helicobacter pylori*. *Expert Opinion on Therapeutic Targets*. **7**, 725–735. https://doi.org/10.1517/14728222.7.6.725

Luo, M., Sun, L., and Hu, J. (2009). Neural detection of gases—carbon dioxide, oxygen—in vertebrates and invertebrates. *Current Opinion in Neurobiology*. **19**, 354–361. https://doi.org/10.1016/j.conb.2009.06.010

Marcus, E.A., Moshfegh, A.P., Sachs, G., and Scott, D.R. (2005). The Periplasmic α-Carbonic Anhydrase Activity of *Helicobacter pylori* Is

Essential for Acid Acclimation. *Journal of Bacteriology*. **187**, 729–738. https://doi.org/10.1128/JB.187.2.729-738.2005

Marshall, B.J, and Warren, J.R. (1984). Unidentified Curved Bacilli In The Stomach Of Patients With Gastritis And Peptic Ulceration. *The Lancet*. **1**, 1311–1315. https://doi.org/10.1016/S0140-6736(84)91816-6

Marshall, B.J., Armstrong, J.A., McGechie, D.B., and Glancy, R.J. (1985). Attempt to fulfil Koch's postulates for pyloric *Campylobacter*. *Medical Journal of Australia*. **142**, 436–439. https://doi.org/10.5694/j.1326-5377.1985.tb113443.x

Marshall, B.J., McGechie, D.B., Rogers, P.A., Glancy, R.J. (1985). Pyloric *Campylobacter* infection and gastroduodenal disease. *Medical Journal of Australia*. **142**, 439–444.

Matthew, J.B., Morrow, J.S., Wittebort, R.J., and Gurd, F.R. (1977). Quantitative determination of carbamino adducts of alpha and beta chains in human adult hemoglobin in presence and absence of carbon monoxide and 2,3-diphosphoglycerate. *Journal of Biological Chemistry*. **252**, 2234–2244.

Maveyraud, L., Golemi, D., Kotra, L.P., Tranier, S., Vakulenko, S., Mobashery, S., and Samama, J.P. (2000). Insights into Class D β-Lactamases Are Revealed by the Crystal Structure of the OXA10 Enzyme from *Pseudomonas aeruginosa*. *Structure*. **8**, 1289–1298. https://doi.org/10.1016/S0969-2126(00)00534-7

McGee, D.J., May, C.A., Garner, R.M., Himpsl, J.M., and Mobley, H.L.T. (1999). Isolation of *Helicobacter pylori* Genes That Modulate Urease Activity. *Journal of Bacteriology*. **181**, 2477–2484.

Mendz, G.L., Jimenez, B.M., Hazell, S.L., Gero, A.M., and O'Sullivan, W.J. (1994). De novo synthesis of pyrimidine nucleotides by *Helicobacter pylori*. *Journal of Applied Bacteriology*. **77**, 1–8. https://doi.org/10.1111/j.1365-2672.1994.tb03036.x

Meyer-Rosberg, K., Scott, D.R., Rex, D., Melchers, K., and Sachs, G. (1996). The effect of environmental pH on the proton motive force of *Helicobacter pylori*. *Gastroenterology*. **111**, 886–900. https://doi.org/10.1016/s0016-5085(96)70056-2

Miller, S.I., Hoffman, L.R., and Sanowar, S. (2007). Did Bacterial Sensing of Host Environments Evolve from Sensing within Microbial Communities? *Cell Host & Microbe*. **1**, 85–87. https://doi.org/10.1016/j.chom.2007.04.002

Mobley, H.L.T. (2001). Urease. In: Mobley, H.L., Mendz, G.L., and Hazell, S.L. (Eds.), *Helicobacter Pylori: Physiology and Genetics.* Washington (DC): ASM Press.

Morollo, A.A., Petsko, G.A., and Ringe, D. (1999). Structure of a Michaelis Complex Analogue: Propionate Binds in the Substrate Carboxylate Site of Alanine Racemase. *Biochemistry*. **38**, 3293–3301. https://doi.org/10.1021/bi9822729

NCBI Resource Coordinators (2018). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research.* **46**, D8–D13. https://doi.org/10.1093/nar/gkx1095

Noto, J.M., Romero-Gallo, J., Piazuelo, M. B., and Peek, R. M. (2016). The Mongolian Gerbil: A Robust Model of *Helicobacter pylori*-Induced Gastric Inflammation and Cancer. *Methods in Molecular Biology (Clifton, N.J).* **1422**, 263–280. https://doi.org/10.1007/978-1-4939-3603-8_24

Olsen, J.V., Ong, S.-E., and Mann, M. (2004). Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues. *Molecular & Cellular Proteomics*. **3**, 608–614. https://doi.org/10.1074/mcp.T400003-MCP200

Padra, M., Adamczyk, B., Flahou, B., Erhardsson, M., Chahal, G., Smet, A., Jin, C., Thorell, A., Ducatelle, R., Haesebrouck, F., Karlsson, N.G., and Lindén, S.K. (2019). *Helicobacter suis* infection alters glycosylation and decreases the pathogen growth inhibiting effect and binding avidity of gastric mucins. *Mucosal Immunology*. **12**, 784–794. https://doi.org/10.1038/s41385-019-0154-4

Park, S.A., Ko, A., and Lee, N.G. (2011). Stimulation of growth of the human gastric pathogen *Helicobacter pylori* by atmospheric level of oxygen under high carbon dioxide tension. *BMC Microbiology*. **11**, 96. https://doi.org/10.1186/1471-2180-11-96

Parsonnet, J., Hansen, S., Rodriguez, L., Gelb, A.B., Warnke, R.A., Jellum, E., Orentreich, N., Vogelman, J.H., and Friedman, G.D. (1994). *Helicobacter pylori* infection and gastric lymphoma. *New England Journal of Medicine*. **330**, 1267–1271. https://doi.org/10.1056/NEJM199405053301803

Patel, S., and Majmundar, S.H. (2019). Physiology, Carbon Dioxide Retention, In: *StatPearls*. Treasure Island (FL): StatPearls Publishing.

Pearson, M.A., Schaller, R.A., Michel, L.O., Karplus, P.A., and Hausinger, R.P. (1998). Chemical Rescue of *Klebsiella aerogenes* Urease Variants

Lacking the Carbamylated-Lysine Nickel Ligand. *Biochemistry*. **37**, 6214–6220. https://doi.org/10.1021/bi980021u

Peleteiro, B., Bastos, A., Ferro, A., and Lunet, N. (2014). Prevalence of *Helicobacter pylori* Infection Worldwide: A Systematic Review of Studies with National Coverage. *Digestive Diseases and Sciences.* **59**, 1698–1709. https://doi.org/10.1007/s10620-014-3063-0

Sachs, G., Weeks, D.L., Wen, Y., Marcus, E.A., Scott, D.R., and Melchers, K. (2005). Acid acclimation by *Helicobacter pylor*i. *Physiology (Bethesda)*. **20**, 429–438. https://doi.org/10.1152/physiol.00032.2005

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*. **9**, 676–682. https://doi.org/10.1038/nmeth.2019

Schott, T., Kondadi, P.K., Hänninen, M.L., and Rossi, M. (2011). Comparative Genomics of *Helicobacter pylori* and the human-derived *Helicobacter bizzozeronii* CIII-1 strain reveal the molecular basis of the zoonotic nature of non-pylori gastric *Helicobacter* infections in humans. *BMC Genomics.* **12**, 534. https://doi.org/10.1186/1471-2164-12-534

Shimamura, T., Watanabe, S., and Sasaki, S. (1985). Enhancement of enterotoxin production by carbon dioxide in *Vibrio cholerae*. *Infection and Immunity*. **49**, 455–456.

Sipponen, P., and Marshall, B.J. (2000). Gastritis and Gastric Cancer: Western Countries. *Gastroenterology Clinics of North America*. **29**, 579–592. https://doi.org/10.1016/S0889-8553(05)70131-X

Smith, K.S., and Ferry, J.G. (2000). Prokaryotic carbonic anhydrases. *FEMS Microbiology Reviews*. **24**, 335–366. https://doi.org/10.1111/j.1574-6976.2000.tb00546.x

Solnick, J.V., and Schauer, D.B. (2001). Emergence of Diverse *Helicobacter* Species in the Pathogenesis of Gastric and Enterohepatic Diseases. *Clinical Microbiology Reviews*. **14**, 59–97. https://doi.org/10.1128/CMR.14.1.59-97.2001

Stadie, W.C., and O'Brien, H. (1936). The Carbamate Equilibrium I. the Equilibrium of Amino Acids, Carbon Dioxide, and Carbamates in Aqueous Solution; with a Note on the Ferguson-Roughton Carbamate Method. *Journal of Biological Chemistry*. **112**, 723–758.

Stolte, M., Bayerdorffer, E., Morgner, A., Alpen, B., Wundisch, T., Thiede, C., and Neubauer, A. (2002). *Helicobacter* and gastric MALT lymphoma. *Gut*. **50**, iii19–iii24. https://doi.org/10.1136/gut.50.suppl_3.iii19

Stoof, J., Kuipers, E. J., and van Vliet, A. H. M. (2009). Characterization of NikR-responsive promoters of urease and metal transport genes of *Helicobacter mustelae*. *BioMetals*. **23**, 145–159. doi:10.1007/s10534-009-9275-7

Takayama, Y., Castañeda, C.A., Chimenti, M., García-Moreno, B., and Iwahara, J. (2008). Direct evidence for deprotonation of a lysine side chain buried in the hydrophobic core of a protein. *Journal of the American Chemical Society.* **130**, 6714–6715. https://doi.org/10.1021/ja801731g

Taylor, C.T., and Cummins, E.P. (2011). Regulation of gene expression by carbon dioxide. *Journal of Physiology*. **589**, 797–803. https://doi.org/10.1113/jphysiol.2010.201467

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*. **22**, 4673–4680.

Tomb, J.F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzegerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Watthey, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M., and Venter, J.C. (1997). The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature*. **388**, 539–547. https://doi.org/10.1038/41483

Tran, N.H., Qiao, R., Xin, L., Chen, X., Liu, C., Zhang, X., Shan, B., Ghodsi, A., and Li, M. (2019). Deep learning enables de novo peptide sequencing from data-independent-acquisition mass spectrometry. *Nature Methods.* **16**, 63–66. https://doi.org/10.1038/s41592-018-0260-3

van Vliet, A.H.M. (2016). Use of pan-genome analysis for the identification of lineage-specific genes of *Helicobacter pylori*. *FEMS Microbiology Letters*. **364**, fnw296. doi:10.1093/femsle/fnw296

van Vliet, A.H.M., and Kusters, J.G. (2015). Use of Alignment-Free Phylogenetics for Rapid Genome Sequence-Based Typing of *Helicobacter pylori* Virulence Markers and Antibiotic Susceptibility. *Journal of Clinical Microbiology.* **53**, 2877–2888. doi:10.1128/jcm.01357-15

van Vliet, A.H.M., Kuipers, E.J., Stoof, J., Poppelaars, S.W., and Kusters, J.G. (2004). Acid-Responsive Gene Induction of Ammonia-Producing Enzymes in *Helicobacter pylori* Is Mediated via a Metal-Responsive Repressor Cascade. *Infection and Immunity*. **72**, 766–773. https://doi.org/10.1128/IAI.72.2.766-773.2004

van Vliet, A.H.M., Poppelaars, S.W., Davies, B.J., Stoof, J., Bereswill, S., Kist, M., Penn, C.W., Kuipers, E.J., and Kusters, J.G. (2002). NikR Mediates Nickel-Responsive Transcriptional Induction of Urease Expression in *Helicobacter pylori*. *Infection and Immunity*. **70**, 2846–2852. https://doi.org/10.1128/IAI.70.6.2846-2852.2002

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. **25**, 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

West, A.L., Evans, S.E., González, J.M., Carter, L.G., Tsuruta, H., Pozharski, E., and Michel, S.L.J. (2012). Ni(II) coordination to mixed sites modulates DNA binding of *Hp*NikR via a long-range effect. *Proceedings of the National Academy of Sciences of the United States of America*. **109**, 5633–5638. https://doi.org/10.1073/pnas.1120283109

West, A.L., John, F.St., Lopes, P.E.M., MacKerell, A.D., Pozharski, E., and Michel, S.L.J. (2010). Holo-Ni(II)HpNikR is an asymmetric tetramer containing two different nickel binding sites. *Journal of the American Chemical Society*. **132**, 14447–14456. https://doi.org/10.1021/ja104118r

Whary, M.T., and Fox, J.G. (2004). Natural and experimental *Helicobacter* infections. *Comparative Medicine*. **54**, 128–158.

World Health Organisation. (2017). WHO publishes list of bacteria for which new antibiotics are urgently needed. *World Health Organisation.* Available from: https://www.who.int/news-room/detail/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed

Yamaguchi, K., and Hausinger, R.P. (1997). Substitution of the Urease Active Site Carbamate by Dithiocarbamate and Vanadate. *Biochemistry*. **36**, 15118–15122. https://doi.org/10.1021/bi971767f

Yang, J., Hart, E., Tauschek, M., Price, G.D., Hartland, E.L., Strugnell, R.A., and Robins-Browne, R.M. (2008). Bicarbonate-mediated transcriptional activation of divergent operons by the virulence regulatory protein, RegA, from *Citrobacter rodentium*. *Molecular Microbiology*. **68**, 314–327. https://doi.org/10.1111/j.1365-2958.2008.06171.x

Zhang, M., and Vogel, H.J. (1993). Determination of the side chain pK_a values of the lysine residues in calmodulin. *Journal of Biological Chemistry*. **268**, 22420–22428.