Investigation of higher-order self-assembly of the DNA binding protein H-NS

by

Arsen Petrović

Department of Biochemistry and Molecular Biology

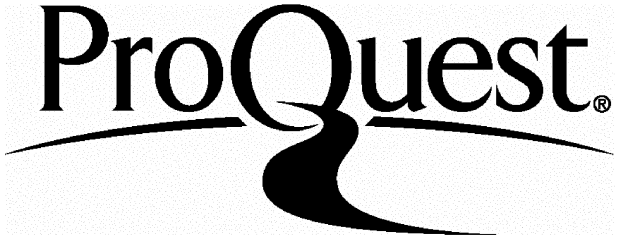
University College London

Gower Street

London WC1E 6BT

This thesis is submitted to the University of London in fulfilment of the requirements for the degree of Doctor of Philosophy in the Faculty of Science

September 2003



To Kosta and Ubavka

"However happy you may be with a solution, never think of it as final. There are great solutions, but a final solution does not exist. All our solutions are fallible."

K. Popper

Abstract

H-NS, a DNA binding protein widely distributed in Gram-negative bacteria, acts as a global transcriptional regulator for the expression of a large number of genes. The 136 residue protein has a modular organisation with two structurally independent domains responsible for DNA binding and protein-protein interactions, respectively. Self-association of the H-NS polypeptide is intimately linked to the formation of a specific nucleoprotein structure capable of inhibiting transcription. Attempts at deciphering the structure/function relationship of the H-NS polypeptide have highlighted the convoluted relationship between the various regions in the primary sequence.

A combination of complementary techniques was used to investigate self-assembly of the *Salmonella typhimurium* H-NS polypeptide. The results thus obtained suggest a possible model for the H-NS oligomerisation and highlight the importance of secondary structure in mediating a higher-order structure formation. The assembly of H-NS oligomers occurs through the association of the dimeric intermediate via two protein interfaces. The initial dimer formation occurs through the coiled-coil motif residing in the amino terminal domain (residues 1-49). The dimers thus formed associate in a head-to-tail fashion through the region encompassing residues 64-89 and 1-20. Mutagenesis studies underline the importance of the short amino-terminal helix H2 in supporting the higher-order oligomer formation. Hydrodynamic and modelling studies of the H-NS homologue, StpA, are also presented.

Statement

The work in this thesis was conducted in the Department of Biochemistry and Molecular Biology, University College London between 1999 and 2002. The AUC experiments were performed at the National Institute for Medical Research (NIMR) in collaboration with Dr. J. Eccleston. Except for the NMR spectra, which were collected by Dr. M. Williams, this work is entirely that of the author.

Some of the work in this thesis has been published elsewhere:

Esposito D., **Petrovic A.**, Harris R., Ono S., Eccleston J.F., Mbabaali A., Haq I., Higgins C.F., Hinton J.C., Driscoll P.C. and Ladbury J.E. (2002) H-NS oligomerization domain structure reveals the mechanism for high order self-association of the intact protein. *J.Mol.Biol.* **324,** 841-850.

Acknowledgements

This thesis would not appear in its present form without the kind assistance and support of many people.

Firstly, I wish to thank my supervisor, Prof. John Ladbury, for his continual guidance and support throughout this project. Without his useful suggestions I would no doubt still be searching for an overall focus for this thesis. To Dr. John Eccleston I owe my appreciation for his assistance, without which this thesis would not have been possible. I would also like to thank the other members of the biophysics lab, past and present, for their guidance and wealth of expertise, especially Mark Williams and Roger George. Thanks are also due to Dr. Paul Driscoll and Dr. Snezana Djordjevic for their encouragement with my thesis.

My deepest admiration goes to certain people in the department. To Reza, who has kept my mind occupied with a sea of new curiosities and for teaching me the importance of all things Persian, and to Gillian, who supplied me with vital doses of friendship and who is an amazing dancing and drinking partner! I would also like to acknowledge Reg, my personal editor, with an amazing set of eyes constantly searching for the gaps in sentences, and for being there from the start. My friend Diya, likewise, was a great pillar of support during writing of this thesis. Last, but not least, I would like to thank my family for their constant support and encouragement. Although they will probably never read this thesis, they are the happiest people to see it finished.

To each of the above, I extend my deepest appreciation.

Contents

List of Tables

Chapter 1.

Chapter 2.

Chapter 4.

List of Figures

Chapter 1.

Chapter 2.

Chapter 3.

Chapter 4.

Appendix

Abbreviations

A	adenine
Å	angstrom
AFM	atomic force microscopy
APS	ammonium persulphate
ATP	adenosine triphosphate
AUC	analytical ultracentrifuge
°C	degrees Celsius
C	cytosine
$CaCl_2$	calcium chloride
CD	circular dichroism
cPHB	poly-(R)-3-hydroxybutyrate
Da	dalton
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
DTT	dithiothreitol
EDTA	ethylenediaminetetra-acetic acid
FPLC	fast protein liquid chromatography
G	guanine
HSQC	heteronuclear single quantum coherence
IPTG	isopropyl-β-D-thiogalactopyranoside
K_d	dissociation constant
KB	kilo base pair
M_r	relative molecular mass
mRNA	messenger RNA
NMR	nuclear magnetic resonance
nm	nanometre
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	protein data bank
PMSF	phenyl-methyl-sulfoxide
RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
2D	two-dimensional
T	thymine
T_m	melting temperature
Tris	tris(hydroxymethyl)aminoethane
θ	molar ellipticity
TAE	tris-acetate
TEMED	N, N, N', N'-tetramethylethylenediamine
U	unit
UV	ultraviolet

1. Introduction

1.1 Eukaryotic and archaeal histones and nucleosomes

Results from DNA sequence comparisons of genes encoding the small subunit rRNA component of ribosomes (16S rRNA in prokaryotes and 18S rRNA in eukaryotes) have overturned the older prokaryote-eukaryote classification in favour of the three-domain (bacteria, archea and eukarya) structuring of the living world (Woese *et al.*, 1990; Woese, 1998). Archea are morphologically and metabolically similar to bacteria, but their genes involved in the information processing systems, including transcription, translation, DNA replication and recombination, most closely resemble eukaryal orthologues (Musgrave *et al.*, 2002). As well as bacteria, archea lack a nuclear membrane and thus have their genomes organised into a nucleoid-type structure (see 1.2). However, the architectural proteins, which are involved in the chromosome organisation, are largely different (Sandman et al., 1998, 2001; Sandman and Reeve, 2000). In spite of the significant discrepancy in genome size, both archea and eukarya exploit similar mechanisms, which facilitate DNA packaging based on an evolutionary conserved histone motif (Sandman *et al.*, 1998). Archeal histones were first identified in the hyperthermophile *M. fervidus* and the members of this family have since been recognized in both mesophilic and thermophilic archeal species (Sandman *et al.*, 1990). Apart from minor differences, archeal and eukaryal histones have very similar three-dimensional organisation ("histone fold"), consisting of three amphipathic α helices separated by short loops and β-strand regions (Starich *et al.*, 1996; Decanniere *et al.*, 2000). It is well known that the eukaryotic histone proteins are relatively small, basic proteins that have been grouped in five classes: H1, H2A, H2B, H3 and H4 (Kornberg and Lorch, 1999). Furthermore, a histone octamer formed by an $(H3+H4)_2$ tetramer flanked by two (H2A+H2B) dimers is associated with ~146 bp of DNA

to form the fundamental, nucleoprotein repeating unit of eukaryotic chromatin known as a nucleosome (Luger *et al.*, 1997). It has emerged that the archeal nucleosome is based on histone tetramer circumscribed by ~80bp of DNA resembling the structure formed by the central $(H3+H4)_2$ tetramer within a eukaryal nucleosome (Sandman *et al.*, 1998, 2001). The novel features found in the eukaryotic nucleosome, namely histone tails (N- and C-terminal amino acid extensions flanking the histone fold) and the (H2A+H2B) dimers, lead to an increase in the length of DNA packaged and contribute to the nucleosome remodelling events (Sandman and Reeve, 2000). It has been argued that the eukaryal nucleosome may have evolved from a structure that is similar to the present day archeal nucleosome as the expansion of the genome size and the development of the cellular differentiation necessitated archiving of nonexpressed genes during early eukaryal evolution (Sandman and Reeve, 2001). In line with its near universal conservation in archeal (euryarcheota) and eukaryal species, the histone-mediated mechanism of DNA compaction may represent the most efficient system in terms of DNA base-pairs packaged per unit protein (Sandman *et al.*, 1998, 2001).

1.2 Nucleoid structure and nucleoid-associated proteins – a brief overview

The genetic material in bacteria is generally organised in one or few compact structures known as nucleoids, despite the absence of both nucleosome-like structures and a nuclear membrane. Notwithstanding many years of research, the elucidation and characterisation of the nucleoid structure is progressing slowly. Nevertheless, light and fluorescence microscopic images suggest the presence of some underlying organisation (Kellenberger *et al.*, 1986; Zimmerman, 2003). Several factors have been put forward to explain DNA compaction observed *in vivo* (these factors need not be independent):

(a) DNA supercoiling

(b) participation of nucleoid-associated protein in chromosome compaction

(c) stabilisation of the nucleoid by the surrounding highly concentrated cytoplasm ("macromolecular crowding")

It is well established that the protein composition of the nucleoid is dominated by a small group of abundant polypeptides, commonly referred to as 'chromatin-associated' or nucleoid proteins (Drlica and Rouviere-Yaniv, 1987; Hayat and Mancarella, 1995; Pettijohn, 1998; McLeod and Johnson, 2001; Dorman and Deighan, 2003). The term 'histone-like' is also in use in the literature, although it has been pointed out that this can be misleading since these proteins do not represent true homologues of eukaryotic and archeal histones but only possess superficial similarity (DNA binding capacity, abundance, small size and the overall charge) to them (Ussery *et al.*, 1994; Dorman and Deighan, 2003). Nucleoid proteins can exert strong influence on both the local and global structure of the bacterial chromosome and they are thought to contribute to the structural organisation of the nucleoid (Schmid, 1990). Despite this, it has been difficult to assign the precise architectural role of each protein in chromosome condensation. Furthermore, new experimental evidence has challenged the established beliefs as to the structural roles of these proteins (Brunetti *et al.*, 2001; Dame and Goosen, 2002). On the other hand, numerous studies have demonstrated their involvement in a variety of physiological processes, such as DNA replication, recombination and transcription. The best-characterised nucleoid proteins in *E.coli* are HU (heat unstable), IHF (integration host factor), Fis (factor for inversion stimulation) and H-NS (histone-like nucleoid structuring). For more comprehensive overviews of nucleoid proteins and their role in the regulation of gene expression refer to those by McLeod and Johnson (2001) and Dorman and Deighan

(2003). The remaining part of the introduction will focus on the analysis of the structural and functional aspects of H-NS.

1.3 Histone-like nucleoid structuring protein (H-NS)

1.3.1 General aspects

As well as HU, the H-NS (histone-like nucleoid structuring protein) DNA binding protein is a major component of the enteric bacterial chromosome with profound roles in the regulation of bacterial physiology. Numerous studies over the last three decades aimed at understanding structural and functional aspects of this elusive protein are now starting to provide insights into its three-dimensional organisation and the mechanistic details responsible for its specific regulatory activities.

The *E.coli* H-NS protein was originally identified as a heat-stable transcription factor (Jacquet *et al.*, 1971) that represents one of the major components of the bacterial nucleoid (Varshavsky *et al.*, 1977). Its abundance and association with isolated nucleoids led to the recognition of the possible role of H-NS in the structural organisation of the *E.coli* chromosome. In support of this hypothesis it has been shown that H-NS is able to compact and constrain DNA supercoils *in vitro* (Spassky *et al.*, 1984; Tupper *et al.*, 1994). Morphological changes in cells overproducing H-NS are associated with highly condensed nucleoids (Spurio *et al.*, 1992). Furthermore, direct *in vivo* evidence has shown that *hns* mutants have increased levels of negative supercoiling of both plasmid and chromosomal DNA (Mojica and Higgins, 1997). This complements studies in which plasmids isolated from *hns* mutants show altered levels of supercoiling (Higgins *et al.*, 1988; Hulton *et al.*, 1990; Hinton *et al.*, 1992). Consistent with its role in chromosome organisation is the co-localisation of H-NS within the *E.coli* nucleoid (Dürrenberger *et al.*, 1991). Supporting

evidence for the intracellular localisation of H-NS comes from immunofluorescence experiments that showed that in exponential and stationary phase cells H-NS is uniformly distributed throughout the nucleoid (Azam *et al.*, 2000).

Atomic force microscopy (AFM) studies of H-NS/DNA complexes have yielded insights into the mechanism of DNA compaction (Dame *et al.*, 2000). In the presence of H-NS, structural changes in the circular DNA molecule are evident; namely, a lateral condensation of large regions of the plasmids in addition to the formation of globular structures that incorporate considerable amounts of DNA (Dame *et al.*, 2000). Force-extension measurements of single λ-DNA/H-NS complexes have suggested that H-NS polymerises along DNA, thereby increasing the end-to-end extension of the complex compared with that of naked DNA (Amit *et al.*, 2003). The results from this study challenge the assumption that H-NS alone is able to influence DNA compaction. Significantly, as shown in a recent *in vivo* study, major histone-like proteins including H-NS may not represent important factors required for the organisation of the bacterial nucleoids since their depletion does not appear to influence the supercoiled loop-domain arrangement within the *E.coli* chromosome (Brunetti *et al.*, 2001).

There is ample evidence to suggest that in addition to its role in influencing nucleoid architecture, H-NS acts as a direct regulator of transcription for a large number of genes. Mutations in *hns* result in various phenotypes since H-NS is involved in the modulation of synthesis of a large number of apparently unlinked gene products (Atlung and Ingmer, 1997). Large-scale proteomic and transcriptomic analyses of H-NS-deficient and wild-type strains of *E.coli* have demonstrated that as much as 5% of all genes analysed were directly or indirectly altered in the *hns* mutant (Hommais *et al.*, 2001; Schröder and Wagner, 2002).

Those genes that have their expression regulated through H-NS are often linked to bacterial adaptation to stressful environments. Furthermore, H-NS is known to participate in a plethora of other cellular processes such as recombination and translation (Schröder and Wagner, 2002).

A single-copy gene (*hns*) encoding H-NS polypeptide has been mapped to the 27 and 34-min region on the *E.coli* and *S.typhimurim* chromosomes, respectively (Görannson *et al.*, 1990; Hulton *et al.*, 1990; May *et al.*, 1990). H-NS is a small, neutral (pI~7.5) DNA binding protein with an approximate molecular mass of 15 500 Da. Although initially referred to as H1 (Jacquet *et al.*, 1971) and B1 (Varshavsky *et al.*, 1977), its DNA structuring activities led to the name H-NS, which has now gained general acceptance (Ussery *et al.*, 1994; Atlung and Ingmer, 1997; Schröder and Wagner, 2002). It has become apparent, however, that structural and functional characteristics of H-NS are clearly distinct from eukaryotic and archeal histones.

Contrary to the initial belief of the restricted species distribution of the *hns* gene, mounting phylogenetic and experimental evidence suggest that H-NS and H-NS-like proteins represent a large family of functionally and structurally conserved DNA binding proteins in Gram-negative bacteria (Ussery *et al.*, 1994; Bertin *et al.*, 1999).

It will become apparent in the forthcoming sections that the ability of H-NS to modulate the expression of a large number of genes stems from its atypical DNA binding properties, which results in the formation of a nucleoprotein complex, capable of a dynamic response to the presence of various environmental stimuli and/or other regulatory proteins.

1.3.2 Primary structure

The primary sequence of H-NS has been elucidated for a number of bacterial species (e.g. Hulton *et al.*, 1990; Marsh and Hillyard, 1990; Falconi *et al.*, 1998). It is highly conserved among enteric bacteria and is comprised of between 135 and 137 amino acids (Fig.1.1). The amino acid sequence of H-NS is distinct from most other DNA binding proteins since the majority of charged residues are predominantly acidic rather than basic, which is reflected in its low pI value of 5.4. However, under native conditions, H-NS is neutral, and migrates with a pI of approximately 7.5 (Spassky *et al.*, 1984). It should be pointed out that the putative DNA binding region (see 1.4.3) is predicted to have pI of 9.7. Furthermore, H-NS bears no sequence similarity to any other known DNA binding proteins (Ussery *et al.*, 1994). In line with these findings, the DNA binding domain shows no sequence or structural conservation to any classical DNA binding motifs (Ussery *et al.*, 1994; Shindo *et al.*, 1995; Schröder and Wagner, 2002).

The occurrence of three isoforms of the H-NS polypeptide differing in their isoelectric points (pI) has been described (Spassky *et al.*, 1984). The presence of multiple isoforms was clearly demonstrated in a later study, which confirmed that these are *hns* gene products (Ussey et al., 1994). Until recently the nature of the biochemical differences leading to the different isoforms was unknown (Ussery *et al.*, 1994). No evidence was found for the incorporation of phosphate (Ussery *et al.*, 1994), or for binding of cofactors like nucleotides or cAMP (Schröder and Wagner, 2002). Furthermore, no differences in the ratio of different isoforms were found under different growth condition regimes (e.g. low temperature, high osmolarity) or growth phase (Ussery *et al.*, 1994). There has been a finding regarding the possible nature of H-NS protein post-translational modifications (Reusch *et al.*, 2002). Large amounts of short-chain poly-(R)-3-hydroxybutyrate (cPHB)

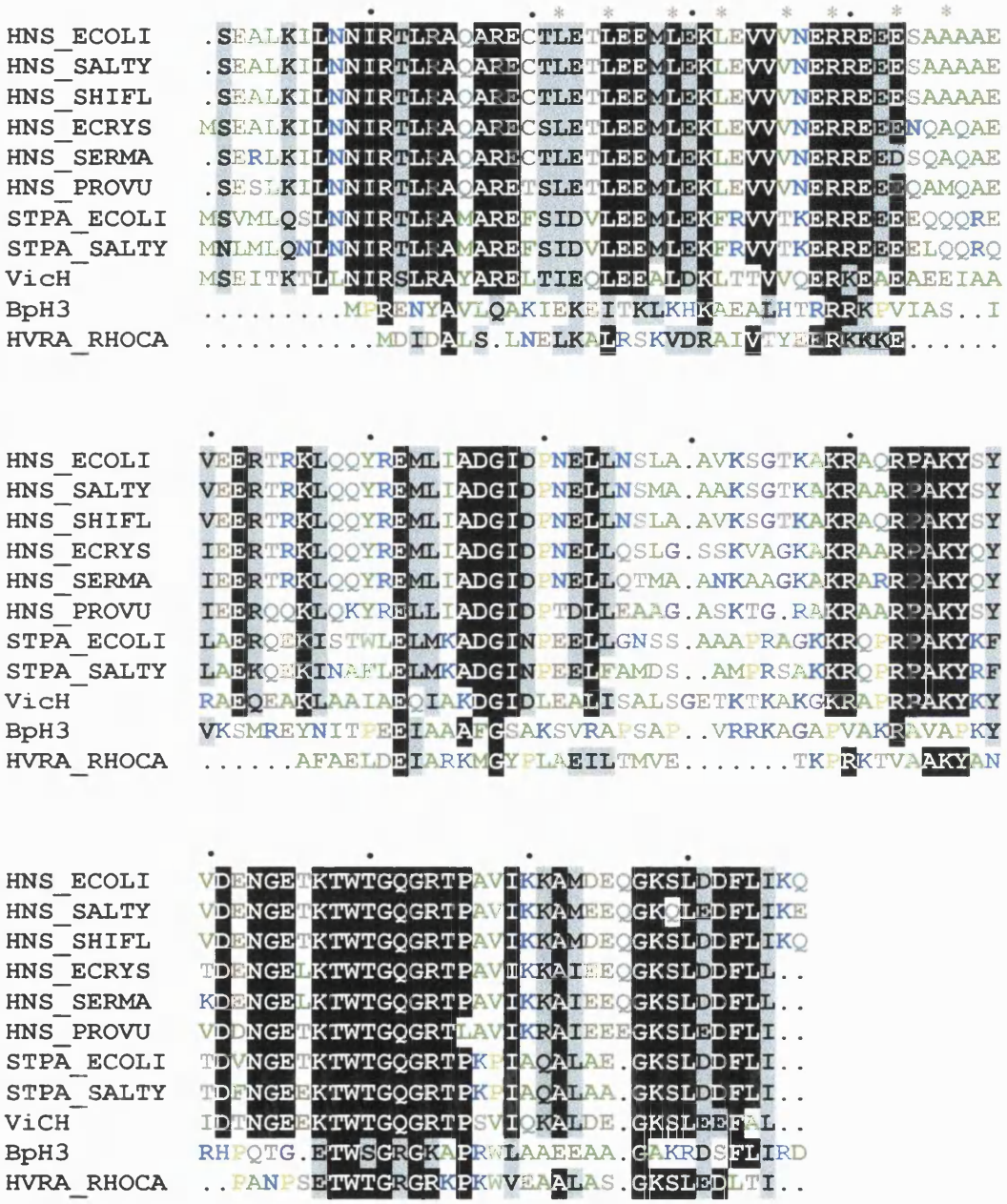


Figure 1.1 Multiple sequence alignment of H-NS homologues. The numbering scheme has been taken from Renzoni *et al.* (2000). Every tenth residue is indicated by a dot. The accession numbers for the sequences are listed in the Material and Methods chapter. Conserved residues of the coiled-coil heptad repeat are indicated by an asterix.

were shown to be associated with H-NS preparations. The chemical characteristic of cPHB including its amphiphilic nature with alternating methyl and carbonyl ester groups allows it to associate simultaneously with both polar (e.g. the phosphate backbone of DNA) and nonpolar surfaces (e.g. the hydrophobic regions in polypeptides) (Huang and Reusch, 1996; Schröder and Wagner, 2002). The importance of this post-translational modification in modulating structure and/or function of H-NS is, at present, unknown but it may play a role in DNA binding (Schröder and Wagner, 2002). It is interesting to note that the majority of cPHB present in the *E.coli* cells are associated with proteins in the cell cytoplasm, particularly in the ribosomal fraction (>70%), which includes the nucleoid (Huang and Reusch, 1996).

The C-terminally truncated form of H-NS (residues 1-67) is apparently capable of supporting post-translational modifications characteristic of the wild-type protein (Donato and Kawula, 1999). Moreover, the N-terminal point mutations (A18E and L26P) yield only one H-NS isoform, which suggests that potential modification sites are located near residues 18 and 26. A single isoform is normally obtained under standard conditions of H-NS overexpresion (Ussery *et al.*, 1994).

1.3.3 Tertiary structure

Several different approaches have been used to map functional domains onto the primary structure (Shindo *et al.*, 1995; Ueguchi *et al.*, 1996; Williams *et al.*, 1996; Ueguchi *et al.*, 1997; Smyth *et al.*, 2000). Both genetic and biochemical studies indicate that the H-NS protein has a modular domain organisation (Fig. 1.2) (Dorman *et al.*, 1999). Initially, the DNA binding region was delineated by showing that the C-terminal end of the molecule (residues 91-137), in isolation, is able to bind DNA (Shindo *et al.*, 1995). The first

structural insights were gained from the NMR studies of the DNA binding domain that revealed unusual and novel features in terms of its secondary structure organisation and the elements involved in the DNA recognition (Shindo *et al.*, 1995, 1999). Subsequent studies based on the *in vivo* and *in vitro* analysis of *hns* mutants have shown that H-NS has at least two functional domains (Ueguchi *et al.*, 1996, 1997). Mutations in the C-terminal end of the protein caused a reduced DNA binding activity, thus complementing earlier observations and suggesting its functional role in DNA recognition. On the other hand, the effect of mutations located in the N-terminus caused the loss of H-NS repressor function whilst retaining DNA binding. Based on cross-linking analysis it has been suggested that H-NS monomers can self-associate to form dimers and oligomers (Spassky and Buc, 1977; Gualerzi *et al.*, 1986; Falconi *et al.*, 1988).

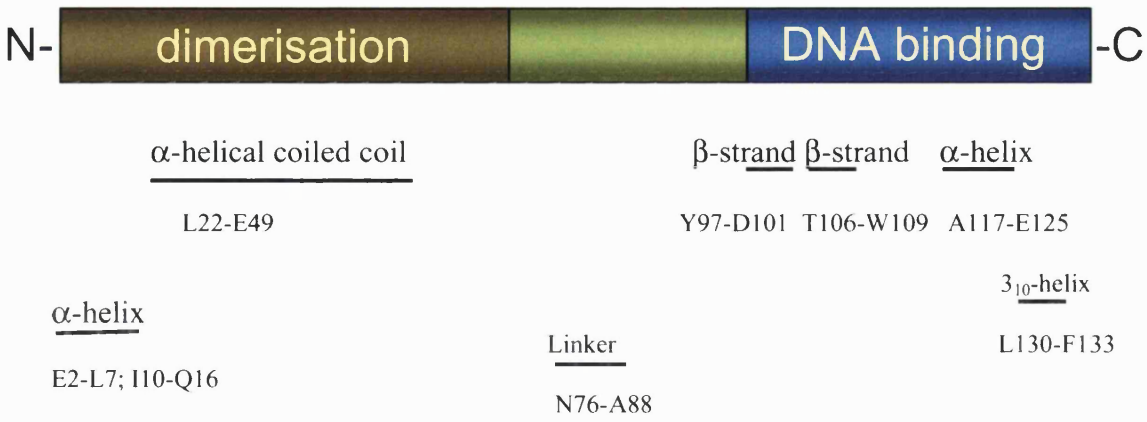


Figure 1.2 Modular structure of H-NS protein. The amino-terminal dimerisation domain, the carboxy-terminal DNA binding domain and the linker region are illustrated (various regions are not drawn to scale). Secondary structure and linker regions are indicated below the figure.

These findings were substantiated by gel filtration analysis of wild-type H-NS (Ueguchi *et al.*, 1996). The analysis of dominant negative and missense mutants have illuminated the role of the N-terminal region in mediating protein-protein interactions and provided functional correlation between dimer formation and the ability of H-NS to function as a transcriptional repressor (Williams *et al.*, 1996; Ueguchi *et al.*, 1997).

1.3.3.1 The N-terminal domain – structural perspective

H-NS sequence analysis by Higgins and colleagues (Ussery *et al.*, 1994) revealed the similarity between the N-terminal region and characteristic motifs found in α-helical coiled-coils. Since the α-helical coiled-coil is a common assembly motif found in a diverse array of structural and regulatory proteins in nature (Lupas, 1996) a possible structural explanation relating to the ability of H-NS to oligomerise was suggested. *In silico* analysis of the sequences of mutant H-NS proteins known to be impaired in dimer formation showed a strong correlation between self-association behaviour and the propensity for coiled-coil formation (Dorman *et al.*, 1999). In agreement with these results a detailed biophysical investigation on the N-terminal dimeric fragment encompassing first 64 residues has corroborated the hypothesis that protein-protein interaction is mediated via a coiled-coil formation (Smyth *et al.*, 2000).

Until recently, no information pertaining to the three-dimensional organisation of the H-NS N-terminal domain was available. Multidimensional NMR studies on the N-terminal fragment (1-57) gave a first insight into the three-dimensional structure of the H-NS dimerisation domain (Esposito *et al.*, 2002). Within the context of a dimeric domain each protomer is composed of three helices (Fig. 1.3). The dimer interface is formed by the longest helix (H3, residues 22-49) that adopts a left-handed parallel coiled-coil

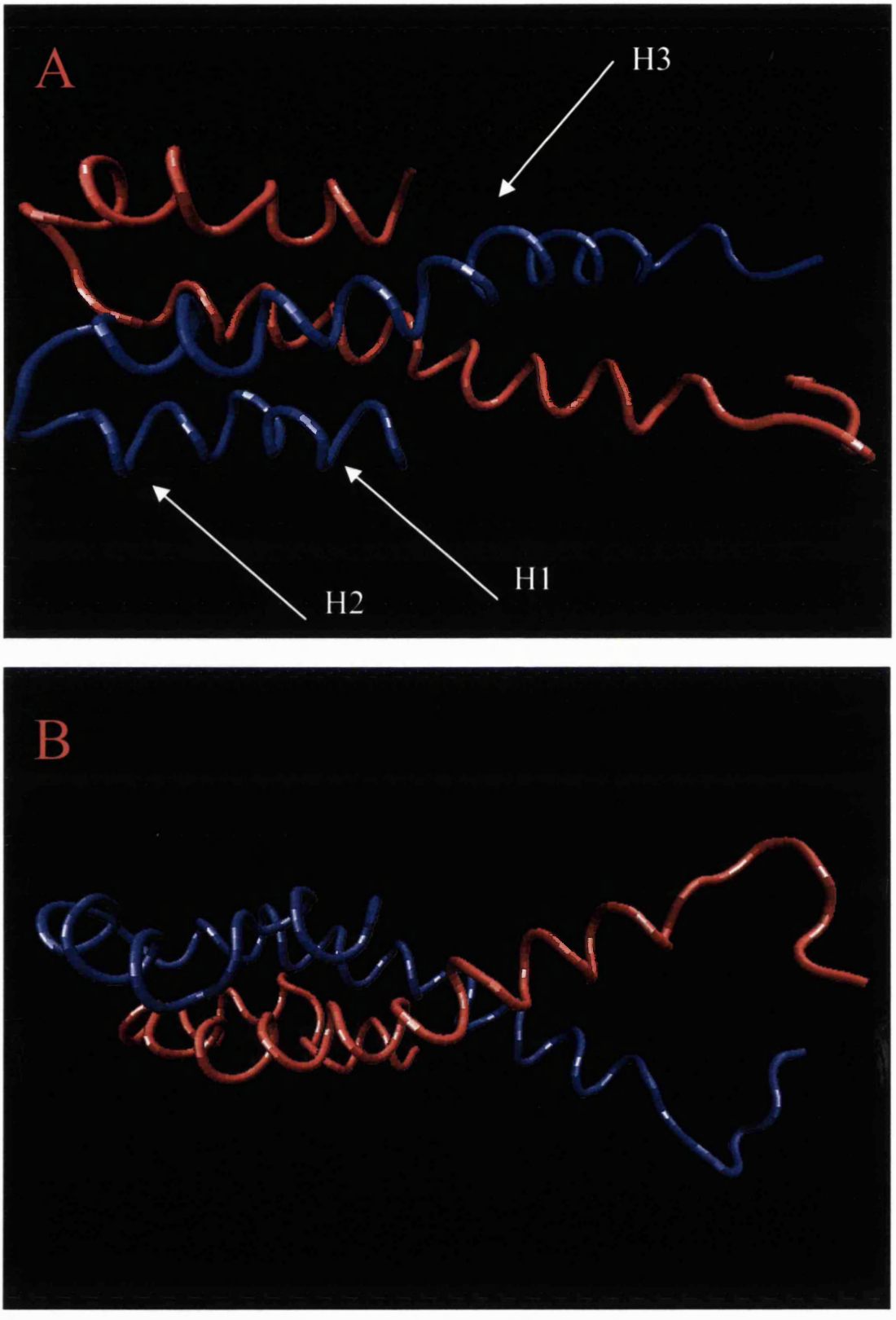


Figure 1.3 Ribbon illustrations of the three-dimensional structure of the H-NS 1-57 dimerisation domain in top view (A) and side view (B). The VMD program (Humphrey *et al.*, 1996) and the H-NS N-terminal domain coordinates deposited in the Protein Data Bank (1LR1; Esposito *et al.*, 2002) were used. The two protomers are shown in red and blue, respectively. Helices 1, 2 and 3 are indicated by arrows in (A).

configuration. Two short N-terminal helices (H1, residues 2-7; and H2, residues 10-16) fold back onto the coiled-coil core from the same protomer. The characteristic heptad repeat, commonly denoted $(abcdefg)_n$, is occupied by residues Leu22, Leu29, Val36 and Glu43 at the *a* and Leu25, Leu32, Arg 39 and Ala46 at the *d* positions respectively. Previous studies have identified several mutants that are distributed throughout the N-terminal region that interfere with the function of the protein (Ueguchi *et al.*, 1996; 1997; Williams *et al.*, 1996). In the light of the available three-dimensional structure, placement of the several mutants can be rationalised in terms of the structural disruption of the dimeric interface. For example, residues L25 and L29 form crucial interactions in the *a* and *d* positions of the heptad repeat, suggesting that the substitution of leucine side-chain in the hydrophobic interior compromises coiled-coil a formation.

Recently, the structure of the N-terminal domain (residues1-46) derived from the *E.coli* H-NS polypeptide has been determined (Bloch *et al.*, 2003). In accordance with the findings presented by Ladbury and co-workers (Esposito *et al.*, 2002), the H-NS 1-46 fragment adopts dimeric configuration with similar secondary structural arrangements composed of the three helical segments: α1 (residues1-7), α2 (residues11-18) and α3 (residues 22-46) (Fig. 1.4). However, there is a lack of agreement with regards to the topological arrangements of the protomers within the dimer. The protein-protein interface is formed by the association of the α3 helices that adopt an antiparallel coiled-coil structure. The α2 and α3 helices from each protomer form a U-shaped conformation, which is capped by the α1 helix lying perpendicular to them. Furthermore, two N-terminal helices (α1 and α2) from each monomer pack against the coiled-coil helix of the other protomer, thus inducing the conformation whereby they face one side of the structure.

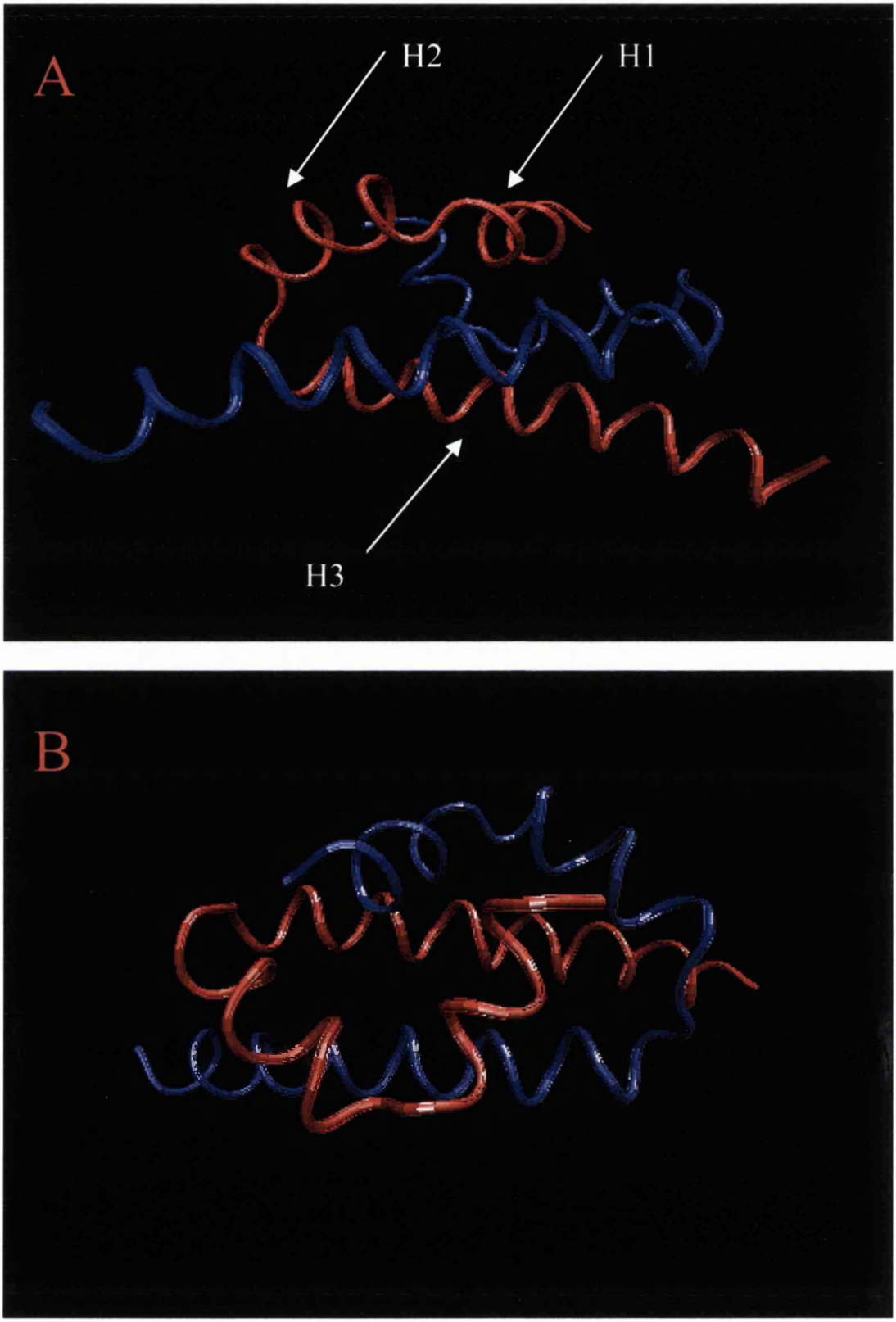


Figure 1.4 Ribbon illustrations of the three-dimensional structure of the H-NS 1-46 dimerisation domain in side view (A) and top view (B). The VMD program (Humphrey *et al.*, 1996) and the H-NS N-terminal coordinates deposited in the Protein Data Bank (1NI8; Bloch et al., 2003) were used. The two protomers are shown in red and blue, respectively. Helices 1, 2 and 3 are indicated by arrows in (A).

One can speculate what the functional implications of the recent investigations into the structural organisation of the H-NS dimerisation domain may hold. The prospect that the two seemingly contradictory views regarding the three-dimensional structure of the H-NS N-terminal domain represent alternative conformations of the same fold seems an exciting idea. In light of the recent evidence regarding the temperature-induced structural changes in the H-NS polypeptide this remains a testable hypothesis (Amit *et al.*, 2003) (see Chapter 5).

1.3.3.2 The C-terminal domain –structural perspective

The structure of the C-terminal domain of H-NS (residues 91-137) implicated in DNA binding has been solved by NMR spectroscopy (Shindo *et al.*, 1995). The arrangement of the secondary structural elements creates a unique DNA binding motif (Fig. 1.5). As shown by these studies, the 47 residue C-terminal fragment of *E.coli* H-NS polypeptide is composed of three flexible loop regions connecting an antiparallel β-sheet (residues 97-101), a α-helix (residues 117-125) and short 3_{10}-helix (residues 130-133). The native structure is stabilised by the formation of a hydrophobic core composed of the side chains of four aromatic and two aliphatic residues (Y97, Y99, W109, F133, V118 and I119). As can be seen in Figure 1.5 the short α-helical segment is positioned at right angles relative to the β-sheet with the 3_{10}-helix situated next to the first β-strand. Subsequent examinations of the longer C-terminal fragments have extended the region involved in DNA binding to include residues 80-90 that lack a defined secondary structure within the context of the H-NS 60-137 polypeptide (Shindo *et al.*, 1999). Elucidation of the regions involved in DNA binding by chemical shift perturbation studies highlighted the contribution of two segments involving positively charged residues spanning T107-A117 and R90-K96. Interestingly, these residues lie in the unstructured regions occupied by loop 2 and the N-terminal end

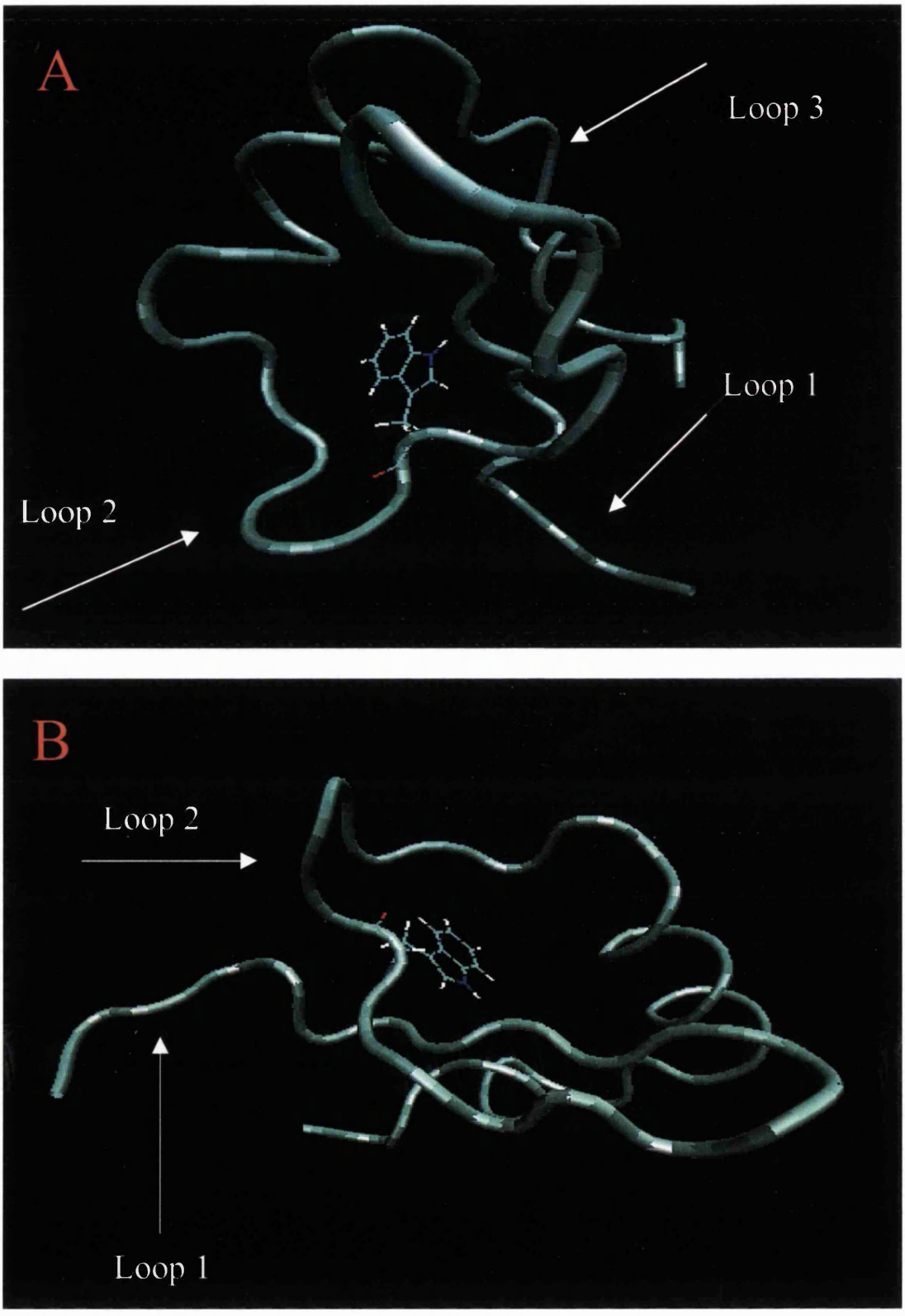


Figure 1.5 Ribbon illustrations of the three-dimensional structure of a C-terminal DNA binding domain of H-NS in two different orientations (A and B). The VMD program (Humphrey *et al.*, 1996) and the C-terminal domain (residues 91-137) coordinates deposited in the Protein Data Bank (1HNR; Shindo *et al.*, 1995) were used. The conserved tryptophan residue (W109) is shown as a side chain. Loop regions are indicated by arrows (loop 1, 2 and 3 in (A) and loop 1 and 2 in (B).

(loop 1) of the DNA binding domain. In agreement with these studies, H-NS mutants defective in DNA binding identified by Mizuno and colleagues (Ueguchi *et al.*, 1996, 1997) map to the positions occupied by these loops.

Recent biochemical studies have suggested that N-terminal dimerisation domain also contributes to DNA binding, since a mutation of the surface exposed residues (R11E/R14A) previously shown to affect gene repression, causes a loss in the ability to preferentially recognise intrinsically curved DNA sequences (Bloch *et al.*, 2003).

1.3.4 Quaternary structure

The conclusions emanating from investigations into the functional mechanisms of H-NS underline the linkage of its self-associative behaviour to its ability to exert regulatory properties through DNA binding (Spurio *et al.*, 1997; Badaut *et al.*, 2002; Dame *et al.*, 2001; Amit *et al.*, 2003). There have been a number of studies that have shown that H-NS exists in solution in various oligomeric forms. No satisfactory model has been attained despite the intensive efforts aimed at delineating the nature of the lowest oligomeric form and the mechanisms responsible for its interconversion into higher-order states (Ussery *et al.*, 1994; Schröder and Wagner, 2002).

Based on early cross-linking studies, it was estimated that H-NS could form dimers and higher oligomers (tetramers) in a concentration-dependant fashion, irrespective of the presence of DNA (Falconi *et al.*, 1988). In support of this, gel-filtration experiments (Ueguchi *et al.*, 1996) have revealed that the H-NS dimer is capable of forming higher-order oligomers of even-numbered subunits (tetramers and hexamers). In conjunction with

H-NS mutant studies, it was hypothesised that two different types of protein interfaces may be responsible for the formation of initial dimers and higher-order oligomers, respectively.

It is believed that H-NS interacts with target DNA in an oligomeric form since its ability to repress transcription is severely impaired in a mutant defective in oligomerisation (Spurio *et al.*, 1997). Similarly, the disruption of the dimerisation domain, and presumably higher-order oligomers, has been shown to have detrimental effects on H-NS function *in vivo* (Ueguchi *et al.*, 1997). Loss of the cooperative DNA binding characteristic of wild-type H-NS has been observed in several N-terminally truncated mutants, thereby suggesting the involvement of oligomerisation in the formation of H-NS/DNA complexes (Ueguchi *et al.*, 1997). The central role of H-NS polymerisation on the DNA template in the formation of complexes capable of inhibiting transcription has been firmly established from recent footprinting studies (Badaut *et al.*, 2002).

In vitro analysis of a series of C-terminally truncated H-NS polypeptides highlighted the inherent complexity underlying H-NS quaternary structure (Smyth *et al.*, 2000). Biophysical examinations of the full-length H-NS polypeptide from *S.typhimurium* have demonstrated a concentration-dependent formation of heterogeneous populations of higher-order oligomers (2-20mers). In accordance with a previous study (Ueguchi *et al.*, 1996), the N-terminal fragment lacking a DNA binding domain (residues 1-89) shows oligomerisation behaviour equivalent to wild-type H-NS. A shorter N-terminal fragment (residues 1-64), on the other hand, results in the formation of a discrete dimer (see Chapter 3), which is incapable of further self-association. The central region of H-NS has eluded the investigations into its structural organisation and the possible role in oligomerisation (see 1.4.4.1). Interestingly, the N-terminally truncated H-NS polypeptide (residues 60-137)

appears to behave as a dimer, implying the possibility that some other as yet unrecognised structural motif may also be involved in inter-monomer interactions (Ueguchi *et al.*, 1997). Intriguingly, proline residue 115 appears to play a crucial role in oligomer formation despite its location in the DNA binding domain (Spurio *et al.*, 1997).

Opposed to the previous study (Smyth *et al.*, 2000), H-NS monomers were shown to undergo dimer-tetramer equilibrium, which is influenced by the presence of different ions and temperature (Ceschini *et al.*, 2000). Dimer formation has a noticeable hydrophobic character, whereas the interactions leading to tetramerisation are of electrostatic origin (Ceschini *et al.*, 2000).

Quantitative cross-linking experiments have substantiated the notion that wild-type H-NS is able to adopt a range of different oligomeric states (Badaut *et al.*, 2002). Finally, in line with previous mutagenesis evidence (Ueguchi *et al.*, 1997), Rimsky and colleagues have shown that mutations in the coiled-coil region (L26P), that affect its stability, have a major impact on higher-order oligomer formation, thereby underscoring the structural and functional linkage between native quaternary configuration and transcriptional repression (Badaut *et al.*, 2002).

1.3.4.1 Formation of the full-length H-NS – molecular 'lego' approach

Several studies have suggested that the region connecting the N-terminal dimerisation domain (residues 1-50) to the C-terminal DNA binding region consist of a poorly structured flexible linker that occupies a position between residues 76 and 88 (Cusik and Belfort, 1998; Dorman *et al.*, 1999). Initial NMR studies on the full-length H-NS polypeptide were consistent with a model where the C-terminal DNA binding domain is

structurally independent and isolated from the N-terminal region by virtue of a flexible linker (Smyth *et al.*, 2000). Using fluorescence resonance energy-transfer measurements, a distance of 45Å between the C- and N-terminal domains of H-NS was deduced, which supports the view that within the context of the full-length polypeptide these two structured regions are located at opposing protein surfaces (Schröder *et al.*, 2001). The precise location of the linker region and its functional role remains to be determined (Schröder and Wagner, 2002). Certainly, part of this region appears to actively contribute to the DNA binding and is an essential part of the DNA binding domain (Shindo *et al.*, 1999). In accordance with biochemical evidence, NMR solution studies have shown that the residues encompassing the linker region form a stretch of disordered chain (Shindo *et al.*, 1999). Indeed, the linker region appears to be the most divergent part of the protein and, significantly, no mutations influencing H-NS functions *in vivo* have been discovered within this sequence to date (reviewed in Dorman *et al.*, 1999). Interestingly, an analysis of H-NS homologues has revealed an important pattern of domain conservation emphasising that the independent function of each domain is structurally established by the presence of a linker region (see 1.4.5).

As previously stated, a number of studies have contributed to our understanding of the functional role and structural organisation of the N-terminal dimerisation domain (Ueguchi *et al.*, 1997; Smyth *et al.*, 2000; Esposito *et al.*, 2002). In spite of the wealth of evidence demonstrating the ability of the H-NS polypeptide to adopt a range of higher-order oligomeric states the structural basis for H-NS self-association is not properly understood.

It was hypothesised early on that the H-NS dimer could serve as a basic core unit, which undergoes further self-association (Ueguchi *et al.*, 1996, 1997). The addition of 25

residues to the core region encompassing the N-terminal dimeric fragment (residues 1-64) results in further self-association suggesting that the central region (residues 65-89) contains structural determinants responsible for the higher-order oligomer formation (Smyth *et al.*, 2000). It is likely that the general principles governing the formation of the H-NS quaternary architecture are more complex as evident from the observations that the integrity of the coiled-coil region is essential for the establishment of higher-order oligomers (Badaut *et al.*, 2002). The contribution of the various elements in the H-NS primary sequence to the native quaternary fold remains to be established.

1.3.5 Regulation of the *hns* gene

H-NS has often been described as one of the most abundant nucleoid-associated proteins, and a major constituent of the bacterial nucleoid (e.g. Higgins *et al.*, 1990). Regulation of the level of H-NS is critical for the maintenance of intracellular homeostasis since variations in its concentration have been shown to have severe effects upon gene expression and genome stability (Lejeune and Danchin, 1990; Spurio *et al.*, 1992; McGovern *et al.*, 1994; Palchaudhuri *et al.*, 1998; Hommais *et al.*, 2001). Notwithstanding its important regulatory roles at many functional levels during cell metabolism, there is conflicting evidence as to the amount of H-NS protein present *in vivo* and its growth phase-dependant variation (Schröder and Wagner, 2002).

The *hns* gene is regulated at both transcriptional and post-transcriptional levels. At the transcriptional level, the *hns* gene is subject to negative feedback control by its own product, which is consistent with the presence of the curved high-affinity binding sites for H-NS in the upstream promoter regions of the *hns* gene (Falconi *et al.*, 1993; Ueguchi *et al.*, 1993; Dersch *et al.*, 1993). These findings are supported by the observations that

disruption of the *hns* gene or induction of H-NS synthesis leads to an approximately fourfold increase in the expression of *hns* reporter gene fusions and repression respectively (Dersch *et al.*, 1993; Ueguchi *et al.*, 1993).

In addition to the negative feedback regulatory loop, the antagonistic role of FIS protein in the transcriptional control of *hns* has been recognized (Falconi *et al.*, 1996). *In vitro* studies have established that FIS binds to at least seven sites in the promoter region of *hns*; thus, to some extent, overlapping the H-NS binding sites involved in the transcriptional autorepression of the *hns* gene (Falconi *et al.*, 1996). Despite the presence of partially overlapping binding sites, it is likely that both H-NS and FIS can simultaneously bind to the same region of DNA. Pertaining to the functional significance of the nucleoprotein complex formation, it was shown that FIS alleviates the H-NS-mediated repressive influence on the *hns* promoter (Falconi *et al.*, 1996). Since FIS protein is known to undergo dramatic changes in its intracellular concentration as a function of the growth phase and/or nutrient availability it could influence the regulation (in combination with H-NS) of the expression of not only *hns* but also other genes (e.g. rRNA) in a manner dependant on external stimuli (Falconi *et al.*, 1996; Schröder and Wagner, 2002; Schneider *et al.*, 2003).

The involvement of the small (87 nucleotides), untranslated RNA molecule DsrA in the post-transcriptional regulation of H-NS has been reported since it was shown that DsrA reduces H-NS protein levels (Lease *et al.*, 1998; Lease and Belfort, 2000a). DsrA is a member of the small RNAs of *E.coli*, which exert its regulatory activities (riboregulation) through sequence-specific RNA-RNA interactions with different mRNAs, thus influencing their translation and turnover rates (Lease and Belfort, 2000b). *In silico* analysis of DsrA

has revealed regions of complementarity to several different genes including *hns* and *rpo*S, and ensuing analysis suggested that DsrA acts in trans by complementary RNA-RNA basepairing with these two target mRNAs (Lease and Belfort, 2000a). Binding of DsrA to the *hns* mRNA blocks H-NS synthesis by strongly enhancing the mRNA turnover (eightfold) rate by several possible mechanisms. Interestingly, the levels of DsrA are influenced both by environmental conditions and by several proteins including H-NS (Lease and Belfort, 2000b and references therein). Thus, the control of DsrA levels, coupled with its differential effects on H-NS and RpoS synthesis, could indirectly modulate cellular transcriptional status.

Although a relatively clear picture has emerged with respect to the regulation of *hns* expression, inconsistent information has been presented about the pattern of growth-dependent changes in the H-NS levels (Hinton *et al.*, 1992; Azam *et al.*, 1999). Several studies have reported an increase in H-NS levels as cells enter the stationary phase (Spassky *et al.*, 1984; Dersch *et al.*, 1993; Ueguchi *et al.*, 1993; Azam *et al.*, 1999), which is in contrast to other reports, which have not detected such changes (Hinton *et al.*, 1992; Free and Dorman, 1995; Deighan *et al.*, 2003). It has been suggested that the autoregulation of H-NS synthesis provides a mechanistic link required for matching H-NS production to the demands of DNA synthesis under different growth conditions (Free and Dorman, 1995). In support of this, the inhibition of DNA synthesis causes down-regulation of *hns* transcription and this effect is dependent upon continued H-NS synthesis. These findings clearly parallel *in vivo* observations of the low levels of *hns* mRNA in stationary phase and its increased synthesis in log phase (Free and Dorman, 1995; Deighan *et al.*, 2003). Moreover, cellular levels of H-NS protein do not vary during different growth

stages as a consequence of autoregulation (Hinton *et al.*, 1992; Free and Dorman, 1995; Deighan *et al.*, 2003).

Available evidence points to the absence of environmental regulation of *hns* transcription (Hulton *et al.*, 1990; Hinton *et al.*, 1992; Göransson *et al.*, 1990; Sonden and Uhlin, 1996). Thus, it appears that changes in the intracellular levels of H-NS do not play any significant role in mediating its regulatory functions. It should be noted, however, that there is an approximately fourfold increase in *hns* gene transcription after a temperature shift from 37°C to 10°C, which is dependent on CspA, the cold-shock transcriptional enhancer (La Teana *et al.*, 1991).

1.3.6 Phylogenetic distribution

In contrast to the HU protein, which is conserved in all eubacteria, the distribution of the H-NS seems to be restricted to Gram-negative bacteria (Ussery *et al.*, 1994; Atlung and Ingmer, 1997; Bertin *et al.*, 1999). Although the *hns* gene was originally identified in several enterobacterial species and in addition to the increasing output from several sequencing projects, a number of studies have shown that H-NS and H-NS-like proteins are distributed among the large number of Gram-negative bacteria (Cusick and Belfort, 1998; Dorman *et al.*, 1999; Bertin *et al.*, 2001; Tendeng *et al.*, 2003).

On the basis of multiple alignments, several plasmid and chromosome-encoded H-NS homologues have been identified (Cusick and Belfort, 1998; Dorman *et al.*, 1999). Interestingly, the plasmid-encoded KorB protein, identified in these studies, consists solely of two H-NS carboxy-terminal domains fused together (Dorman *et al.*, 1999 and references therein). Another study indicated that the Hha regulatory protein from *E.coli* could be

considered a specialised homologue of the amino-terminal domain of H-NS (Nieto *et al.*, 2002). As pointed out, as with H-NS-like proteins identified so far, the nucleic acid-binding domain of H-NS is linked to the amino-terminal domain, which exhibits a low amino acid identity to the same region in H-NS (Bertin *et al.*, 1999; Dorman *et al.*, 1999;). For instance, several of the H-NS homologues such as BpH3 (from *B. pertussis*), HvrA (from *R. capsulatus*) or VicH (from *V. cholerae*) share a highly conserved H-NS C-terminal domain (>40%), which is in contrast to the less prominent conservation within the amino-terminal region (Bertin *et al.*, 1999; Tendeng *et al.*, 2000). On the basis of theoretical analysis and functional complementation studies it was shown that these proteins belong to the family of H-NS-related DNA binding proteins (Bertin *et al.*, 1999). Furthermore, database searches using an amino acid 'motif' from the H-NS C-terminal DNA binding domain identified more than twenty different proteins belonging to the same class of H-NS DNA binding proteins (Bertin *et al.*, 2001).

Despite the low amino acid conservation in the N-terminal domain of H-NS-like proteins, this region is predicted to be mainly α-helical with a propensity for a coiled-coil conformation (Bertin *et al.*, 1999; Dorman *et al.*, 1999; Bertin *et al.*, 2001). Regardless of the sequence divergence in the N-terminal domain of BpH3 and HvrA and differences in length (about 10 amino acids shorter than H-NS N-terminal region), both proteins are still able to form dimers (Bertin *et al.*, 1999). Similarly, the VicH protein, which is reported to play a role in the pathogenicity of *V. cholerae*, has the ability to form oligomers *in vitro* (Tendeng *et al.*, 2000). The presence of a putative coiled-coil region in the H-NS-like proteins presumably allows them to engage in protein-protein interactions. In support of this, the three-dimensional high-resolution crystal structure of the N-terminal region of

VicH reveals the presence of a coiled-coil dimerisation motif (S. Arold, personal communication).

Recently, Bertin and co-workers have characterised new members of the H-NS family isolated from extreme environments (Tendeng *et al.*, 2003). Interestingly, an H-NS-like polypeptide isolated from psychrophilic bacterium collected from Antarctica was shown to be able to form higher-order oligomers at 4°C but not at 37°C. Moreover, differences in the N-terminal domain organisation compared with the mesophilic counterparts was hypothesised to represent a structural adaptation to a low temperature environment.

Combined theoretical and experimental efforts have given rise to the notion that the ability of various H-NS-like proteins to oligomerise is mediated by the amino-terminal coiled-coil region. As the structural conservation in the H-NS C-terminal domain is high among the recognised homologues (Bertin *et al.*, 1999; J. Ladbury, personal communication), variations in the amino-terminal region may cause functional differences and generate a novel regulatory DNA binding protein (Dorman *et al.*, 1999). Could the H-NS homologue with its different expression profiles (e.g. it is environmentally regulated) interfere with the H-NS function through complex formation (e.g. inhibit oligomer formation)?

1.3.7 The H-NS homologue, StpA

In addition to H-NS, *E.coli* and its near relatives possess a second protein, StpA, which has a high level of amino acid sequence identity (58%) with H-NS, and has been recognised as its paralogue (i.e. an intraspecies homologue). The StpA protein is closely related to H-NS both in terms of the structural organisation and functional characteristics (Sondén and Uhlin, 1996; Zhang *et al.*, 1996; Dorman *et al.*, 1999; Sonnenfield *et al.*, 2001). The *stpA*

gene was originally identified and cloned as a suppressor of the Td⁻ phenotype of a splicing-defective phage T4 *td* intron mutant (Zhang and Belfort, 1992). Subsequently, the *stpA* gene was identified based on its ability to function in complementing the phenotypes of *hns* mutants in expression of the *adi* (arginine decarboxylase) gene in *E.coli* (Shi and Bennett, 1994). Based on *in vitro* studies, RNA chaperone activity has been attributed to StpA, as it was shown to facilitate association and resolve the misfolding of RNA (Zhang and Belfort, 1995; Zhang *et al.*, 1996). Recent *in vivo* evidence suggests that StpA destabilises RNA tertiary interactions (*td* group I intron), thus allowing the misfolded molecule to reach a native splicing competent state (Waldsich *et al.*, 2002). The ability of StpA to promote RNA-specific activities (annealing and trans-splicing) is dependant on the presence of the C-terminal domain (Cusick and Belfort, 1998). Interestingly, chaperone activity was also observed for H-NS, albeit with inferior RNA-specific effects compared with StpA (Zhang *et al.*, 1996).

Several ensuing observations highlighted many functional features in common with the H-NS polypeptide, such as the ability to act as a transcriptional repressor, recognise curved DNA sequence motifs and constrain DNA supercoils *in vitro* as efficiently as H-NS (Zhang *et al.*, 1996). Although both proteins exhibit preference for curved DNA sequences, StpA has a four-to six-fold greater affinity for DNA than H-NS (Sonnenfield *et al.*, 2001).

In line with their functional, overlaps the expression of the *stpA* and *hns* genes is down-regulated by their products in a reciprocal manner in addition to being able to exert negative autorepression (i.e. both H-NS and StpA proteins repress the promoter of its own gene) (Sondén and Uhlin, 1996; Zhang *et al.*, 1996). Furthermore, the transdominant StpA

protein can compromise the activity of H-NS at the level of protein-protein interactions (Williams *et al.*, 1996).

The expression of StpA is strongly influenced by H-NS as the levels of stpA mRNA are increased in an *hns* mutant (Sondén and Uhlin, 1996; Zhang *et al.*, 1996; Free and Dorman, 1997). It is clear that the levels of StpA polypeptide are significantly lower compared with H-NS (Sondén and Uhlin, 1996; Zhang *et al.*, 1996). Even in the absence of H-NS, the concentration of StpA represents only 10% of the amount of H-NS present in the wild-type strain (Sonnenfield *et al.*, 2001). Similarly to H-NS (see 1.3.4), there are several conflicting reports in the literature as to the growth-phase regulation of StpA (Atlung and Ingmer, 1997;Schröder and Wagner, 2002). Unlike H-NS, whose levels are maintained at approximately constant level during each stage of the growth cycle and/or different environmental conditions such as osmolarity and temperature (see 1.3.4), *stpA* transcription is strongly regulated by the latter (Göransson *et al.*, 1990; Free and Dorman, 1995; Free and Dorman, 1997; Dorman *et al.*, 1999; Deighan *et al.*, 2003). It has been established that the *stpA* gene is transiently expressed during a short period in the exponential phase in rich growth media even in the presence of H-NS in a FIS-independent manner (Free and Dorman, 1997; Deighan *et al.*, 2003). Importantly, *stpA* transcription can be induced by osmotic shock and thermal upshift (Sonden and Uhlin, 1996; Free and Dorman, 1997). Prolonged transcription of *stpA* has been observed in M9 minimal medium, which is dependent on trans-activation by leucine-responsive regulatory protein (Lrp).

It has been suggested that StpA might serve as a molecular back-up of H-NS, based on its ability to substitute for H-NS in regulation of some, but not all, genes (Sondén and Uhlin, 1996; Zhang *et al.*, 1996). As pointed out by Dorman and co-workers, it is doubtful

whether StpA plays any significant role in compensating for the loss of functional H-NS since *hns* knock-out mutants have clear phenotypes in the presence of intact *stpA* gene (Dorman *et al.*, 1999; Deighan *et al.*, 2000; Deighan *et al.*, 2003).

The failure of StpA to compensate for phenotypes associated with *hns* mutants is likely to be directly correlated with the cellular levels of StpA protein (Sonnenfield *et al.*, 2001). The comparison of the expressed proteins by 2D PAGE analysis demonstrates that for certain genes, the simultaneous presence of both H-NS and StpA is required for normal regulation (Zhang *et al.*, 1996; Deighan *et al.*, 2000). However, it has been difficult to decipher promoters regulated by StpA alone. A clear example in which the relative contribution of H-NS and StpA to the wild-type pattern of gene expression has been distinguished concerns the outer membrane porin protein, OmpF (Deighan *et al.*, 2000). It appears that both proteins cooperate in the regulation of the amount of *micF* RNA (antisense RNA that blocks translation of the *ompF* RNA) at both transcriptional (H-NS) and post-transcriptional levels (StpA).

It is clear that the overall domain organisation of the StpA polypeptide is comparable with that of H-NS (Williams *et al.*, 1996; Ueguchi *et al.*, 1996,1997; Cusick and Belfort, 1998;; Dorman *et al.*, 1999; Smyth *et al.*, 2000). In accordance with their similar domain structuring and high level of sequence identity, several studies have shown that H-NS and StpA are able to form heteromeric complexes (see 1.3.11).

1.3.8 H-NS DNA binding interaction

The dual role of H-NS, both in terms of the structural organisation of the nucleoid in prokaryotic cells and the regulation of the gene expression, presents a challenge in terms of

understanding the nature of different H-NS nucleic acid interactions. *In vitro* studies aimed at delineating the mechanisms responsible for the establishment of H-NS/DNA complexes have been hampered by several factors, including a lack of specificity for any particular sequence, a relatively low affinity for DNA and the convoluted relationship between oligomerisation and DNA binding. Despite having the knowledge of the region involved in DNA binding and its three-dimensional organisation, molecular details of the interaction are not properly understood (Schröder and Wagner, 2002).

The available biochemical data clearly suggest that H-NS has a broad specificity and affinity for all types of nucleic acids (Schröder and Wagner, 2002). H-NS binds to double-stranded DNA of random sequence and, with a lower affinity, to single-stranded DNA or RNA (Friedrich *et al.*, 1998). In accordance with the general affinity for nucleic acids is the absence of DNA consensus motif usually associated with the sequence-specific recognition of DNA (Ussery *et al.*, 1994; Schröder and Wagner, 2002). It has been shown, however, that H-NS recognises, and preferentially binds, to intrinsically curved DNA sequences generated by phased A tracts (Bracco *et al.*, 1989; Yamada *et al.*, 1990; Owen-Hughes *et al.*, 1992; Tupper *et al.*, 1994; Zuber *et al.*, 1994; Dame *et al.*, 2001; Rimsky *et al.*, 2001). Comparative binding studies demonstrated that DNA curvature (and specific architecture) rather than the presence of A tracts is responsible for the preferential binding, as H-NS showed similar affinities towards GC curves (Zuber *et al.*, 1994; Jordi *et al.*, 1997). However, the difference in affinity between 'specific' (i.e. curved) sequence is only a few-fold higher than for 'non-specific' (i.e. generic) DNA. This suggests that the nature of H-NS/DNA interaction and its regulatory mechanism is distinct from the classical repressor protein (Ussery *et al.*, 1994). In addition, H-NS binding has been shown to alter the curvature of the target DNA sequence (Tipnner *et al.*, 1994; Spurio *et al.*, 1997;

Afflerbach *et al.*, 1999). Thus, deformability of the DNA in addition to the intrinsic curvature may be an important factor in determining specific DNA recognition. While it is clear that DNA curvature plays a role in H-NS action, there is no clear relationship between the degree of curvature and H-NS affinity (Schröder and Wagner, 2002). Further uncertainty remains as to whether H-NS binds to the major or minor DNA groove (Tippner *et al.*, 1994; Afflerbach *et al.*, 1999; Schröder and Wagner, 2002).

A number of genes, known to be regulated by H-NS have been shown to contain regions either upstream or downstream of the transcription start site that are intrinsically curved (Tippner *et al.*, 1994; Ussery *et al.*, 1994).

Importantly, H-NS can form two distinct complexes with DNA, depending on H-NS: DNA ratio (Tupper *et al.*, 1994). In their examination of the interaction between H-NS and plasmid DNA, Higgins and co-workers (Tupper *et al.*, 1994) noticed that at high concentrations, H-NS binds non-specifically along the entire length of a DNA molecule. Specific binding, on the other hand, occurs at much lower H-NS: DNA ratios, as observed from a series of discrete sites protected from Dnase I digestion (Tupper *et al.*, 1994; Ussery *et al.*, 1994). *In vitro* evidence for the preferential binding of H-NS to the promoter regions of several genes, which are known to be H-NS responsive, has been found (Lucht *et al.*, 1994; Tippner *et al.*, 1994; Tupper *et al.*, 1994; Falconi *et al.*, 1998; Afflerbach *et al.*, 1999; Govantes *et al.*, 2000; Jordi *et al.*, 2000; Badaut *et al.*, 2002). For example, binding of H-NS to *cydAB* (encoding the subunits of the cytochrome d oxidase) promoter region was shown to protect specific regions from DnaseI digestion at low concentrations (Govantes *et al.*, 2000). Significantly, a small increase in H-NS concentration was sufficient to cause full protection. A similar pattern of H-NS binding has been observed

previously in other systems, since only a two- to six-fold increase in H-NS concentration was sufficient to saturate DNA with H-NS, relative to the amount which caused preferential interaction with the high-affinity sites (Falconi *et al*; 1993; Lucht *et al.*, 1994; Rimsky *et al.*, 2001). Consistent with the peculiarities of H-NS/DNA interactions addressed above is the observation denoting the importance of the oligomerisation to the capacity to recognise curved DNA sequences and to actively bend non-curved DNA (Spurio *et al.*, 1997). These findings led to the suggestion that multiple H-NS molecules bind co-operatively and polymerise on the DNA fragment (Rimsky and Spassky, 1990; Falconi *et al.*, 1993; Lucht *et al.*, 1994). Spectroscopic investigation of the specific and non-specific H-NS/DNA complexes revealed two different complex conformations highlighting dissimilar binding modes (Tippner and Wagner, 1995). A specific H-NS/DNA interaction has a marked hydrophobic character (and possibly hydrogen bonding), whereas non-specific binding is of electrostatic origin (Tippner and Wagner, 1995).

Detailed *in vitro* studies of H-NS interaction with semi-synthetic promoters containing curved (phased A_n tracts) or non-curved upstream sequences led to the initial indication that H-NS binding to DNA may occur in several distinct steps (Zuber *et al.*, 1994; Rimsky *et al.*, 2001). DNaseI footprint titration experiments of H-NS with different DNA fragments highlighted the importance of the curved upstream sequences in the formation of different nucleoprotein complexes capable of "specific" and "non-specific" transcriptional repression (Rimsky *et al.*, 2001). A specific mode of repression necessitates the presence of high-affinity H-NS binding sites presented by the curved DNA template and the establishment of a H-NS/DNA complex at these discrete positions (the nucleation step). The occupancy of high-affinity sites (i.e. nucleation) strongly influences the binding of H-NS at other specific sites, which results in the formation of an early nucleoprotein structure

by cooperative recruitment of H-NS molecules (the propagation step). This, in turn, leads to polymerisation of H-NS along the DNA fragment and the inhibition of transcription (Rimsky *et al.*, 2001). In the absence of the curved upstream insert, transcriptional repression (non-specific) by H-NS involves only the polymerisation step but this requires a higher protein concentration. The nature of the different nucleoprotein complexes involved in specific and non-specific transcriptional repression is unknown but is likely to involve protein conformational changes and has been observed elsewhere (Tippner and Wagner, 1995). Subsequent *in vitro* examinations of the interaction between various dominant-negative H-NS mutants and the DNA template have confirmed the central role of H-NS oligomerisation in the formation of the specific nucleoprotein complex capable of transcriptional repression (Badaut *et al.*, 2002).

Recently, several biophysical approaches have been used to study the structure of H-NS/DNA complexes (Dame *et al.*, 2000, 2001, 2002; Schneider *et al.*, 2001; Amit *et al* 2003). AFM experiments have shown that specific recognition of DNA by H-NS occurs at regions of intrinsic curvature, thereby inducing the formation of an intramolecular bridge by bringing together two DNA helices flanking the apex of the curve as a result of H-NS oligomerisation (Dame *et al.*, 2001). Similarly, zipper-like propagation of H-NS molecules along the DNA template following the initial nucleation event has been proposed based on the microscopic evidence (Schneider *et al.*, 2001). Binding of H-NS induces a larger end-to-end extension of DNA compared with bare DNA as a result of extensive polymerisation along the double-stranded tracts (Amit *et al.*, 2003). The results from these studies are consistent with biochemical evidence obtained so far (Rimsky *et al.*, 2001; Badaut *et al.*, 2002).

It is clear that the interplay between specific structural features of the DNA molecule and the oligomeric properties of H-NS polypeptide leads to the establishment of specific nucleoprotein complexes capable of repressing transcription. Importantly, conformational changes in both the DNA template and the H-NS polypeptide itself can lead to the modulation of the specific nucleoprotein structures. Compared with its regulatory roles, the effects of H-NS on the structural organisation of the bacterial nucleoid remain poorly defined.

1.3.9 H-NS and its target genes

Although H-NS is able to exert its biological effects on different cellular processes, one of its major roles concerns the modulation of the expression of a large number of plasmid and chromosomal genes at the level of transcription (Ussery *et al.*, 1994; Atlung and Ingmer, 1997; Schröder and Wagner, 2002). New light has been shed on the relationship between the control of gene expression by H-NS and the maintenance of the intracellular homeostasis through the application of large-scale experiments (Laurent-Winter *et al.*, 1997; Hommais *et al.*, 2001).

Comparative expression profiling of wild-type and H-NS-deficient *E.coli* strains using DNA arrays suggest that as much as 5% (~250) of the genes were directly or indirectly altered in the *hns* mutant strain (Hommais *et al.*, 2001). Additionaly, these experiments have highlighted the role of H-NS as a repressor of gene expression since the level of more than 80% of the target genes was induced in the *hns* mutant strain (Atlung and Ingmer, 1997; Hommais *et al.*, 2001). The genes regulated by H-NS are very diverse (Table 1.1) and their products have been implicated in many essential biological processes. However, many H-NS-regulated genes share the common feature of being involved in bacterial

Gene	Function	Regulatory level	Effect
appY	Anaerobic growth phase activator	Transcription	Repression
bgl	β-glucosidase	Transcription	Repression
bolA	Cell shape regulator	?	Repression
cfaD	Pili formation	Transcription	Repression
chiA	Endochitinase	Transcription	Repression
clyA	CytolysinA	Transcription	Repression
csgA	Curli formation	Transcription	Repression
csiD	C-stravation indiced genes	Transcription	Activation
cspA	Major cold-shock protein	Transcription	Repression
cydAB	Cytochrome c oxidase	Transcription	Repression
espADB	Enterohemorrhagic *E.coli*	Transcription	Thermoregulation
fimA	Type 1 fimbrial subunit	Transcription	Activation
fimB	Recombinase for *fimA* promoter flipping	Transcription	Repression
flhD	Flagellar control protein	Transcription	Activation
fliA	Flagella-specific σ factor	Transcription	Activation
gadABC	Glutamic acid decarboxylase	Transcription	Repression
hdfR	Transcriptional regulator of the flagellar master operon *flhDC*	Transcription	Repression
hlyCABD	Haemolysin	Transcription	Thermoregulation
kpsF	Capsule gene	Transcription	Repression
lrp	Leucine-responsive regulatory protein	Transcription	Repression
luxC	*luxC*, *luxRI* region of luminous bacteria	Transcription	Repression
lysU	Lysyl-tRNA synthetase	Transcription	Repression
malEFG	Maltose regulon	Post-transcriptional	Activation
malT	Activator of maltose regulon	Translation	Activation
micF	Antisense RNA	Transcription	Repression
nirB	Nitrate reductase	Transcription	Repression
ompC	Outer membrane porin	Transcription	Repression
pap	Pili formation	Transcription	Repression
proU	Proline/glycine betaine transport system	Transcription	Repression
rcsA	Activator of capsular polysaccharide synthesis	Transcription	Repression
rpoS	Stationary σ factor	Post-transcriptional	Repression
rrn	Ribosomal RNA	Transcription	Repression
virB	Invasion regulatory gene (*Shigella flexneri*)	Transcription	Thermoregulation
toxR	Virulence gene regulatory protein (*Vibrio cholerae*)	Transcription	Repression

Table 1.1 H-NS regulated genes. Adapted from Schröder and Wagner, 2002.

responses to various environmental shifts and stress conditions including changes in osmolarity, pH, oxygen availability and temperature (Atlung and Ingmer, 1997; Schröder and Wagner, 2002). Importantly, H-NS has been described as a central regulator of many bacterial virulence functions, which are often thermoregulated (Dorman and Porter, 1998; Falconi *et al.*, 1998; Hurme and Rhen, 1998; White-Ziegler *et al.*, 1998). In line with its important regulatory functions is the involvement of H-NS in the control of genes whose products are implicated in global bacterial networks such as stringent or growth rate control (Afflerbach *et al.*, 1998; Wagner, 2000; Schneider *et al.*, 2003).

Functional classification of H-NS-regulated genes identified by expression profiling (Fig. 1.6) on DNA arrays suggest that as many as 30% of these genes are coding for cell envelope components or are involved in responses to multiple environmental stimuli. Similarly, approximately 20% of the gene products have been implicated in gene expression (transcription and translation) processes (Hommais *et al.*, 2001). The greatest decrease in the mRNA levels corresponding to seven of the flagellar cascade genes was observed in the mutant strain, which is consistent with the previous study highlighting the key role of H-NS in the expression of the flagellar master operon (Soutourina *et al.*, 1999). An addition set of genes whose expression is modulated in response to multiple environmental parameters (such as anaerobiosis or temperature) includes those encoding porins (*ompC*, *ompF*), colanic acid (gmd, wzc) and lipopolysaccharide (LPS) biosynthesis. Interestingly, 34% of the target genes shown to encode proteins of unknown functions have been predicted to localise primarily at the cell surface.

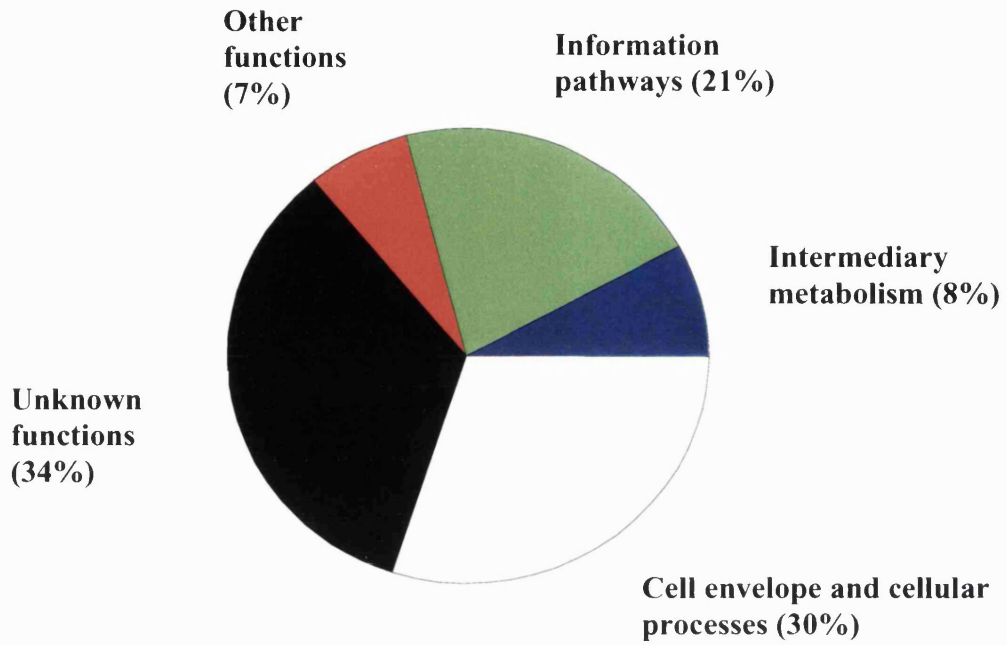


Figure 1.6 Functional classification of the H-NS regulated genes. Adapted from Hommais *et al.*, 2001.

In summary, several studies have brought into focus the complexity of H-NS-mediated regulation of gene expression and its effects on global changes in bacterial physiology.

1.3.10 H-NS and its role in diverse cellular processes

Although H-NS is known to exert its regulatory activities mainly at the level of transcription, it is clear that it additionally plays important roles in diverse cellular processes such as transposition, recombination, DNA repair and translation (Schr der and Wagner, 2002). In several instances, H-NS was implicated in the post-transcriptional regulation of the expression of a number of genes. In the case of the *rpoS* gene (encoding stationary phase-specific sigma factor, σ^S), H-NS has been shown to affect both the translational efficiency of *rpoS* mRNA and the stability of σ^S (Barth *et al.*, 1995;

translational efficiency of *rpoS* mRNA and the stability of σ^S (Barth *et al.*, 1995; Yamashino *et al.*, 1995). Similarly, an involvement of H-NS in the translational stimulation of the MalT (activator of maltose regulon) has been reported (Johansson *et al.*, 1998). As well as influencing translational events, H-NS is known to affect a number of recombination processes (Schröder and Wagner, 2002). For example, *hns* mutations are associated with aberrant transpositional recombination events mediated by IS1 (Shiga *et al.*, 2001). H-NS also contributes to the modulation of a site-specific recombination event of the *fim* switch promoted by the FimB recombinase (O'Gara and Dorman, 2000). Depletion of H-NS causes an increase in the frequency of the Mu transposition (Falconi *et al.*, 1991) and the rate of spontaneous chromosomal deletions (Lejeune and Danchin, 1990). Intriguingly, H-NS is known to suppress γ-ray-induced illegitimate recombination in *E.coli* and to regulate DNA repair in *Shigella* (Palchaaudhuri *et al.*, 1998; Shanando *et al.*, 2001).

1.3.11 Interaction of H-NS with other proteins

There are a number of reports, suggesting that H-NS may specifically interact with other proteins (Williams *et al.*, 1996; Johansson *et al.*, 2001; Nieto *et al.*, 2002; Deighan *et al.*, 2003). The details of such interactions have been described in some detail in very few instances (Johansson *et al.*, 2001). The functional significance of such interactions is unclear but is likely to have important modulatory effects on the *in vivo* processes for which these proteins are responsible. Changes in the levels of the H-NS interacting partners as a function of the growth phase and environmental conditions may offer an increased repertoire for homo- and heteromer formation. This may have important implications in diversifying bacterial responses under the conditions experienced during the infection process (Dorman *et al.*, 1999). As the number of potential interacting partners

increase, the prospective for the establishment of a highly dynamic network of different heteromeric complexes has been highlighted recently (Deighan *et al.*, 2003).

It has been shown that H-NS interacts with the bacteriophage T7 gene 5.5 protein product, both *in vivo* and *in vitro* (Liu and Richardson, 1993). The interaction relieves H-NS-mediated repression, suggesting a role in aiding the reproduction of bacteriophage T7. Interestingly, H-NS also binds to the flagellar torque-generating protein, FliG, which modulates flagellelar rotational speed (Donato and Kawula, 1998). Single H-NS amino acid mutants caused an increase in H-NS/FliG binding affinity, which resulted in a 50% increase of flagellum rotational speed and hypermotility (Donato and Kawula, 1998; Schröder and Wagner, 2002). The significance of the effects of H-NS on flagella rotation is unclear and further studies are needed to provide suitable mechanistic models.

A number of studies have focused on the elucidation of the nature of the complex formation between H-NS and StpA (reviewed in Williams and Rimsky, 1997; Schröder and Wagner, 2002). The extensive sequence identity between H-NS and StpA and the existence of many functional and mechanistic parallels led Williams and co-workers (Williams *et al.*, 1996) to demonstrate that the H-NS N-terminal dimerisation domain (residues 1-64) can be cross-linked to StpA. Turnover studies have established that the StpA protein is unstable in the absence of H-NS and that it is rapidly degraded by the Lon protease *in vivo* (Johansson and Uhlin, 1999). This led to the conclusion that StpA mainly exists in heteromeric complexes with H-NS in *E.coli*. To what extent H-NS-StpA complexes are functionally different from their homomeric counterparts is a subject of intense investigation. Environmental regulation of StpA expression (see 1.3.7) has the potential to lead to the redistribution of the cellular levels of homo- and heteroligomeric

complexes, which in turn may subtly influence gene expression profile and thus allow adaptation to physiological changes (Free *et al.*, 1997; Dorman *et al.*, 1999; Johansson and Uhlin, 1999).

An interesting parallel between H-NS/StpA complexes and the different forms of HU ($HU_{\alpha 2}$, $HU_{\beta 2}$ and $HU_{\alpha}HU_{\beta}$) has been drawn (Dorman *et al.*, 1999). Interaction between H-NS and StpA is mediated via two regions of H-NS, one situated between amino acids 39 and 60 and the other residing between amino acid 80 and the C-terminal domain (Johansson *et al.*, 2001). Truncated H-NS polypeptide lacking one of the regions implicated in the protein-protein interaction is still able to interact with and prevent the degradation of StpA.

Recently, it has been found that Hha protein from *E.*coli specifically interacts with H-NS (Nieto *et al.*, 2002). The Hha protein is a member of a new family of proteins that modulate bacterial virulence gene expression in response to environmental signals that includes YmoA (from *Y. enterolitica*) and possibly proteins involved in conjugative plasmid transfer (Nieto and Juàrez, 1999; Madrid *et al.*, 2002). It has been suggested that both Hha and YmoA represent specialised homologues of the amino-terminal domain of H-NS (Nieto *et al.*, 2002). In contrast to other studies, a relatively clear picture has emerged as to the functional importance of heterooligomer formation between H-NS and Hha in the thermoregulation of α-haemolysin toxin expression (Nieto *et al.*, 2000, 2002). It has been suggested that during the nucleoprotein assembly at the upstream region of the *hly* operon H-NS imparts DNA binding specificity whereas Hha acts through direct interaction with H-NS to generate a hetero-oligomeric complex capable of more efficient repression than the H-NS molecule *per se* (Madrid *et al.*, 2002).

In addition to possessing H-NS and StpA, a novel member of the H-NS-like protein family (Sfh) has been identified in *S. flexneri*, the causative agent of bacillary dysentery in humans (Deighan *et al.*, 2003). The presence of conserved H-NS-like N-terminal oligomerisation domain allows the Sfh polypeptide to engage in protein-protein interactions with itself and H-NS and StpA, respectively. Growth-phase-dependant variation in the cellular levels of H-NS-like proteins in *S.flexneri* may lead to the formation of different heteromeric complexes, which could exert different roles on the cellular processes (Deighan *et al.*, 2003).

1.4 H-NS mode of action - functional and mechanistic perspectives

It is clear that the mechanism by which H-NS affects transcription differs from the 'classical' sequence-specific regulatory proteins (Ussery *et* al., 1994; Schröder and Wagner, 2002). Several different models have been proposed to explain function of H-NS as a transcriptional repressor (Atlung and Ingmer, 1997; Schröder and Wagner, 2002). It has been suggested that H-NS is able to indirectly influence transcription by modulating the supercoiling of the DNA, and thereby influencing promoter activity (Higgins *et al.*, 1990; Hulton *et al.*, 1990; Ussery *et al.*, 1994). Although it has been established that H-NS can influence DNA topology, a number of observations have been made, which are not consistent with the notion H-NS is able to affect promoter activity via alternations in the supercoiled state of DNA (reviewed in Atlung and Ingmer, 1997; Brunetti *et al.*, 2001). Alternatively, H-NS can also act as a direct transcriptional repressor through inhibition of RNA polymerase binding to the promoter regions (Yarmolinsky, 2000; Rojo, 2001). The involvement of H-NS in mediating strong negative control ("silencing") of the *bgl* operon is well-known example (Yarmolinsky, 2000). The term silencing has been used to describe H-NS-dependent reversible transcriptional repression via the formation of nucleoprotein

complexes through protein polymerisation that obscure promoter regions (Yarmolinsky, 2000; Schröder and Wagner, 2002). For example, sequences involved in H-NS-mediated silencing of the *bgl* promoter are ~200 bp away from the promoter start site (Caramel and Schnetz, 1998; Rimsky *et al.*, 2001). In line with recognised DNA binding properties and the capacity to polymerise, H-NS could sequester DNA, thus effectively inhibiting interaction with RNA polymerase or other transcriptional activators (Schröder and Wagner, 2002). The extensive polymerisation of H-NS on DNA recently observed is consistent with the role of H-NS as a transcriptional silencer (Amit *et al.*, 2003).

Since H-NS binding sites often extend into the promoter regions of a number of genes, a simple model involving physical competition between RNA polymerase and H-NS can be envisaged to operate in the transcriptional repression. However, evidence for a direct involvement of H-NS in transcriptional repression following productive initiation complex formations has been presented in several cases (Spasky *et al.*, 1984; Schröder and Wagner, 2000). It is well established that H-NS plays important role in the down-regulation of rRNA synthesis (Tippner *et al.*, 1994). Although both H-NS and RNA polymerase have overlapping binding sites in the *rrnB* P1 (ribosomal RNA) promoter region, their binding is not mutually exclusive. Detailed biochemical evidence supports the model in which bound H-NS physically traps RNA polymerase in an open complex, which then becomes unavailable for productive RNA synthesis (Schröder and Wagner, 2000). Structural explanation for the H-NS-mediated trapping of RNA polymerase has been obtained (Dame *et al.*, 2002). Microscopic studies show that the addition of H-NS to preformed RNA polymerase/DNA complexes leads to ternary complex formation in which DNA on both sides of the RNA polymerase is bridged by H-NS. It has been suggested that a similar

mechanism may operate in the H-NS-mediated transcriptional regulation of the *proU* operon (Schröder and Wagner, 2002).

The involvement of H-NS polypeptide in the expression of virulence genes has been well documented (Hurme and Rhen, 1998; Nye *et al.*, 2000; Beloin and Dorman, 2003). The cooperative binding of H-NS to specific promoter regions is often seen as an important factor in transcriptional repression of virulence genes. Importantly, various environmental stimuli (e.g. temperature) are able to influence nucleoprotein complexes in such a way as to relieve repression. H-NS is involved in the temperature-dependent repression of *virF* (positive activator of a multi-step virulence regulatory cascade) only below a critical temperature of 32°C. It was proposed that the *virF* promoter fragment undergoes specific temperature-dependent conformational transition at ~32°C, which, in turn, influences cooperative interactions between bound H-NS molecules, thereby resulting in transcriptional derepresion (Falconi *et al.*, 1998). Similarly, temperature-dependent structural modification of the LEE1 (locus of enterocyte effacement) promoter region has been proposed as a physical basis for the thermoregulated expression of LEE1 (Umanski *et al.*, 2002). The antagonistic interplay between FIS and H-NS in the regulation of ribosomal RNA synthesis is mediated through protein-induced changes in the DNA geometry that favours binding of either one of the two transcription factors (Afflerbach *et al.*, 1999). Moreover, H-NS polypeptide itself may act as a 'thermal sensor' and establish a mechanistic link between environmental changes and gene regulation. Recent *in vitro* studies have shown that H-NS undergoes conformational changes in response to small alterations in temperature (change of 5-7°C) and osmolarity, which, in turn, prevents polymerisation of H-NS along DNA (Amit *et al.*, 2003). Temperature-dependant changes in the H-NS quaternary structure have also been proposed to explain *in vivo* transcriptional

derepression of the *pap* (encoding pap pili) operon (White-Ziegler *et al.*, 1998). Furthermore, the characteristic polymerisation of H-NS on the *proU* DNA fragment is severely inhibited by a temperature increase (Badaut *et al.*, 2002).

There are a number of examples where nucleoprotein structure formation involving several global regulators (e.g. FIS, IHF, Fnr) has been implicated in the H-NS-mediated control of gene expression (Schröder and Wagner, 2002; Browning *et al.*, 2000). For example, the regulation of *csgD* gene (encoding the positive transcriptional regulator of extracelullar matrix components, curli fimbriae and cellulose) expression in response to oxygen tension is regulated through a nucleoprotein complex comprising OmpR, IHF and H-NS polypeptides (Gerstel *et al.*, 2003). In addition to conformational changes in the DNA topology, the mechanism governing *virF* expression also involves FIS protein in facilitating a switch between repression and stimulation (Falconi *et al.*, 2001). The assembly of such complexes may be facilitated by specific protein-protein interactions, although no clear evidence for interaction between H-NS and any of these regulatory proteins has been presented so far (Schröder and Wagner, 2002). Similarly, a structural and functional dissection of such complex nucleoprotein structures has been limited.

In conclusion, the peculiar structural organisation of H-NS polypeptide in addition to its promiscuous DNA binding characteristics endows it with outstanding functional features in terms of its regulatory properties (Schröder and Wagner, 2002).

1.5 Aim of thesis

The work presented in this thesis aimed to provide information on the self-assembly of the *Salmonella typhimurium* H-NS polypeptide. A range of biochemical and biophysical techniques were employed to facilitate the following studies:

1. Characterization of the oligomeric state of the basic structural unit capable of undergoing self-association
2. Delineation of the protein-protein interfaces leading to the oligomer formation
3. Determination of the secondary structure elements capable of supporting higher-order oligomer formation.

2. Materials and Methods

2. Materials

Chemicals were from Sigma or BDH unless otherwise stated. Media components were purchased from Difco, DNA miniprep and PCR purification kits from Qiagen, Vent DNA polymerase from NEB. *E.coli* DH5α (Promega) or XL-1 blue (Promega) strains were used for all cloning experiments. All primers used were purchased from Pharmacia or MWG. The source of other reagents and facilities are stated in the text as necessary.

2.1 Nucleic acid manipulation

2.1.1 Preparation of DNA

2.1.1.1 Manipulation of DNA by polymerase chain reaction

PCR (polymerase chain reaction) amplification were performed in 50 μl reaction volumes containing 1X Thermopol Buffer (NEB), 200 μM each of dNTP, 10 ng template, 10 pmol primers and 0.5 units Vent DNA polymerase (NEB). Amplifications were performed in a Techne PROGENE THERMOCYCLER. Following a 2-minute 'hot start' at 92°C, samples were denatured at 92°C for 30 seconds, and extended at 72°C for 1 minute/Kbp. Annealing temperatures were calculated for each primer using the equation:

$$T_m\,(°C) = 2(A+T) + 4(G+C) \qquad (1)$$

For each pair of primers, the lowest value was used as the annealing temperature. The completed PCR reaction was removed from the tube, and the small volume of the extension products were resolved by TAE/agarose gel electrophoresis to check for the presence of

band of the right size. The PCR product was purified using the QIAquick PCR purification kit.

2.1.1.2 DNA digestion with restriction enzymes

All restriction enzymes were purchased from New England Biolabs. Restriction enzyme digests were performed generally in either 20 µl or 50 µl reaction volumes in NEB buffers as described in the NEB catalogue buffer chart. Reactions were incubated at 37°C for 90 minutes.

2.1.1.3 Electrophoresis of DNA fragments

Agarose was dissolved in TAE buffer (40 mM Tris-Acetate, 1 mM EDTA) by heating, then cooled and ethidium bromide was added to a concentration of 1 µg/ml before gel casting. DNA fragments of less than 1 KB in size were resolved on 1.5% (w/v) agarose gels. 6X gel loading buffer (0.25% (w/v) xylene cyanol, 0.25% (w/v) bromophenol blue, 30% glycerol dissolved in water) was added to DNA samples, which were electrophoresed in TAE buffer at 60 mA. DNA molecular weight markers (100 bp DNA ladder, NEB) were electrophoresed simultaneously. Resolved DNA bands were visualised by illuminating the gel on a UV light box.

2.1.1.4 DNA extraction and purification from agarose gels

DNA fragments of interest were excised using a clean sharp scalpel. DNA was purified from the slice using the QIAquick gel extraction kit (Qiagen). In brief, the gel slice was solubilised by adding 3 volumes of buffer QC, and incubated at 50°C for 10 minutes. For small DNA fragments (less than 500 bp) the next step generally involved the addition of 1 gel volume of isopropanol. The sample was then loaded onto a silica-gel membrane, and

impurities (salts, primers, ethidium bromide) were washed away. An additional wash in ethanol containing PE buffer completely removed salt contaminants. The fragment was then eluted in 50 µl water.

2.1.1.5 Miniprep plasmid purification

Small-scale preparations of plasmid DNA were performed using Qiaprep Spin Miniprep Kit (Qiagen) following the manufactures instructions. Preparations were made from host cells grown overnight in 4 ml LB, which were pelleted by centrifugation at 2500 rpm for 15 minutes at 4°C. The plasmid purification protocol is based on the alkaline lysis of bacterial cells and neutralization followed by adsorption of DNA onto silica. Endonucleases and salts are then removed in two washing steps followed by elution in 50 µl distilled water. Purified plasmid is suitable for downstream applications, such as DNA sequencing and transformation.

2.1.2 DNA ligation and transformation

2.1.2.1 Ligation of DNA fragments

Ligation reactions contained between 3-fold and 6-fold molar excess of insert DNA relative to the plasmid as estimated from the band intensity on an agarose gels. Ligations were performed in a 10 µl reaction volume containing 1X ligation buffer (120 mM Tris-HCl (pH 7.6), 12 mM $MgCl_2$, 15 mM DTT and 1.2 mM ATP) and 400 U of T4 DNA Ligase (NEB). Reactions were incubated at room temperature overnight. A reaction with no insert DNA was performed in parallel as a control.

2.1.2.2 Preparation of competent cells ($CaCl_2$ method)

A single colony of BL21(DE3)pLys cells freshly streaked from a glycerol stock (Novagen) were inoculated into 5 ml LB containing 34 μg/ml chloramphenicol and incubated overnight at 37°C with shaking. The overnight culture was then diluted 1 in 20 into 50 ml LB containing chloramphenicol (34 μg/ml) and the growth was then allowed until the mid-log phase. The cells were then placed on ice for 20 minutes after which they were spun down at 4°C for 15 minutes at 2500 rpm. The cells pellet was gently resuspended in ice cold 0.1M $CaCl_2$ and left at 4°C for 20 min. The cell suspension was then spun down as described above, resuspended in 1 ml 0.1M $CaCl_2$ containing 10%(v/v) glycerol and aliquoted as 100 μL fractions, which were then stored at -80°C.

2.1.2.3 Preparation of competent cells ($RuCl_2$ method)

A single colony of DH5α or XL-1 blue cells were inoculated into 5 ml LB media and incubated overnight at 37°C with shaking. Overnight culture was diluted 1 in 10 in 100 ml prewarmed LB broth and incubated with shaking at 37°C until an OD_{600} of 0.5 was reached. The cell culture was then cooled on ice for 5 minutes and the cells were collected by centrifugation for 15 minutes at 2500rpm. The cell pellet was gently resuspended in 30 ml cold buffer A (100 mM RbCl, 50 mM $MnCl_2$, 30 mM potassium acetate, 10 mM $CaCl_2$, 15% glycerol, pH 5.8) and was kept on ice for additional 90 minutes. Cells were collected by centrifugation as above, and the cell pellet was resuspended in 4 ml ice-cold bufferB (10 mM MOPS, 10 mM RbCl, 75 mM $CaCl_2$, 15% glycerol pH 6.8). Cells were aliquoted in 100 μl and stored at -80°C.

2.1.2.4 Transformation

100 µL aliquots of competent cells (BL21(DE3)pLysS, Xl-1 blue or DH5α) were allowed to thaw on ice. These were then mixed with either 1 µL of appropriate plasmid DNA or 2 µL ligation reaction and left on ice for 30 minutes. The cells were then heat shocked at 42°C for 45 seconds and transferred to ice for further 2 minutes. The cells were allowed to recover for 1 hour by the addition of 900 µL LB and incubation at 37°C with shaking before streaking on agar plates containing appropriate antibiotics. The plates were incubated overnight at 37°C.

2.1.3 Plasmid manipulation for expression in bacterial cells

2.1.3.1 Construction of histidine-tagged H-NS and StpA polypeptides

Plasmids encoding N- and C-terminally truncated H-NS polypeptides (residues 12-64, 12-89, 17-89 and 20-89) were generated by PCR. Plasmid pJSH6, which encodes full length $H\text{-}NS_{C20S}$ was used as a template. All primers used are shown in Table 2.1. The PCR fragments with engineered 5' *NdeI* and 3' *BamHI* restriction sites were cloned into the pET14(b) vector (Novagen) resulting in the N-terminal fusion to cleavable (thrombin) histidine tag. The plasmid pT7-7 harbouring *E.coli stpA* gene was a generous gift from Dr. J. Hinton (Institute of Food Research). The coding region was isolated by restriction digest with *BamHI* and *NdeI* enzymes and subcloned into pET14(b) vector. Recombinant plasmids were obtained from transformed *E.coli* DH5α or XL-1 and purified using the Qiagen Miniprep kit. The presence of the correct insert was verified by sequencing (Oswel).

Primer name	Sequence (5'-3')
H-NS 12f	GGAATTCCATATGACTCTTCGTGCGCAGGCAAGAG
H-NS 64r	CGCGGATCCTTATAACATTTCACGATACTGTTGCAG
H-NS 89r	CGCGGATCCTTAGCGTTTAGCTTTGGTACCGGATTTAGCGGC
H-NS 17f	CTAATGGCTAATGGCATATGGCAAGAGAAAGCACTCT
H-NS 20f	CTAATGGCTAATGGCATATGAGCACTCTGGAAACGC

Table 2.1 Summary of the primers used in cloning experiments. H-NS 12f, 17f and 20f are forward primers. H-NS 64r and 89r are reverse primers.

2.1.3.2 Site-directed mutagenesis

Point mutations were created in plasmids by site-directed mutagenesis using the QuickChange™ Site-Directed Mutagenesis Kit (Stratagene, LaJolla, CA, USA). Primers containing the desired mutation (Table 2.2) were used in a PCR reaction using *PfuTurbo* DNA polymerase to generate nicked circular plasmid DNA containing the desired mutation. PCR amplifications were performed in 50 µl reaction volumes containing 1X reaction buffer (20 mM Tris-HCl pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2$ SO4, 0.2 mM $MgSO_4$, 0.1% Triton X-100, 0.1 mg/ml BSA), 125 ng primers, 200 µM dNTP mix, 2.5U *PfuTurbo* DNA polymerase and a range of plasmid concentrations (5-50 ng). The enzyme Dpn 1 (10 U) was used to digest parental methylated DNA, by adding it to the PCR reaction and incubating at 37 °C for 1 hour. 1 µl of Dpn 1 treated PCR reaction was used to transform *E.coli* XL1-Blue supercompetent cells. The cells were transformed by incubating on ice with the DNA for 30 minutes, before heat shocking at 42°C for 45 seconds. 0.5 ml of NZY^+ broth (10 g/L casin hydolysate, 5g/L yeast extract, 5g/L NaCl, 20%, 20 mM glucose, 12.5 mM $MgCl_2$, 12.5 mM $MgSO_4$) heated to 42°C was then added to the cells, which were then incubated at 37°C with shaking for 1 hour.

Primer name	Sequence (5'-3')
K5P[a]	GCCATATGAGCGAAGCACTTCCAATTCTGAACAACATCCG
Q16P[a]	CCGTACTCTTCGTGCGCCGGCAAGAGAAAGCACTCTGG
K5P[b]	CGGATGTTGTTCAGAATTGGAAGTGCTTCGCTCATATGGC
Q16P[b]	CCAGAGTGCTTTCTCTTGCCGGCGCACGAAGAGTACGG

Table 2.2 Summary of the primers used in mutagenesis experiments. a and b refer to the two oligonucleotide primers, each complementary to opposite strands of the vector.

The entire volume of transformation reaction was streaked onto agar plates containing 100 μg/ml ampicilin. DNA was purified from colonies and sequenced.

2.2 Large-scale protein expression

2.2.1 Media

All liquid media was made using distilled, de-ionised water and sterilised by autoclaving.

Luria-Bertani (LB) Medium:

Tryptone 10 g/l

NaCl 10 g/l

Yeast Extract 5 g/l

For the preparation of agar plates the medium was supplemented with 15 g/l agar. Antibiotics were added after the medium has been cooled to <50°C.

Minimal Medium (MM) for the uniform enrichment with ^{15}N:

Na_2HPO_4 6 g/l

KH_2PO_4 3 g/l

NaCl 0.5 g/l

$(^{15}NH_4)_2SO_4$ 1 g/l

pH value was adjusted to 7-7.4 and autoclaved. The following were then added:

1M $MgCl_2$	2 ml/l
0.1M $CaCl_2$	100 µl/l
0.1M $FeSO_4$	100 µl/l
20%(v/v) glucose	10 ml/l
10x stock Vitamins	100 µl/l
10x stock Micronutrients	100 µl/l

2.2.2 Large-scale protein expression in LB media

Large-scale protein expression was routinely performed in 4L Luria-Bertani (LB) rich media. Freshly transformed BL21(DE3)pLysS cells were inoculated in 200 ml LB supplemented with carbenicillin (200 µg/ml) and chloramphenicol (34 µg/ml) and incubated overnight at 37°C with a moderate shaking. Following overnight growth, cultures were diluted 1:10 into 500 ml LB supplemented with antibiotics and incubated at 37°C with shaking until the OD at 600 nm reached 0.5. Protein expression was then induced with 1 mM IPTG (Melford Laboratories Ltd.) and the cells were allowed to grow for further 3 hours at the same temperature. The cells were pelleted at 4°C, then spun at 5000 rpm for 15 minutes using a Sorvall GS3 rotor in an RC5B centrifuge. The pellets were pooled together and transferred into ice-cold 50 ml Falcon tubes, centrifuged using a benchtop Labofuge400R at 4°C, and spun at 3000 rpm for 30 minutes with the resulting pellets stored at -20°C.

2.2.3 Large-scale protein expression in MM

In order to over-express ^{15}N labelled H-NS truncated polypeptides and the corresponding mutants for use in the 2D ^{15}N HSQC experiments, a procedure similar to that described in Section 2.2.2 was followed.

A single colony of freshly transformed BL21(DE3)pLysS cells was inoculated into 5 ml of LB supplemented with carbenicillin (200 μg/ml) and chloramphenicol (34 μg/ml). These were incubated at 37°C until an OD_{600} value between 0.4 and 0.6 had been attained. Subsequently, 1 ml of this culture was inoculated into 200 ml M9 minimal media supplemented with appropriate vitamins, nutrients (see section 2.2.1) and antibiotics. Following overnight growth, the cultures were diluted 1 in 20 into 500 ml M9 minimal media supplemented with appropriate vitamins, nutrients and antibiotics. The cultures were allowed to grow at 37°C with shaking until an OD_{600} of 0.4-0.6 had been reached, following which IPTG (Melford Laboratories Ltd) was added to each flask to an overall concentration of 1 mM. The growth was continued for 3 hours at 37°C. The cells were pelleted at 4°C, then spun at 5000 rpm for 15 minutes using a Sorvall GS3 rotor in an RC5B centrifuge. The pellets were pooled together and transferred into ice-cold 50 ml Falcon tubes, centrifuged using a benchtop Labofuge400R at 4°C, and spun at 3000 rpm for 30 minutes with the resulting pellets stored at -20°C.

2.3 Protein purification

2.3.1 Purification of truncated and mutant H-NS polypeptides

Initially, DNA fragments encoding H-NS 12-64 and H-NS 12-89 polypeptides were amplified by PCR using primers designed to engineer *BamHI* and *EcoRI* restriction sites, and cloned in an expression vector pRSETB (Invitrogen) resulting in the addition of the N-

terminal histidine tag and an enterokinase cleavage site. Attempts to overexpress H-NS 12-64 and 12-89 polypeptides were unsuccessful under the range of conditions. Subsequently, as stated above (section 2.1.3.1), all PCR fragments encoding truncated H-NS polypeptides were cloned into pET14(b) expression vector.

Frozen cells were resuspended in 50 ml lysis buffer (20 mM Tris-HCL, pH8, 500 mM NaCl, 5% glycerol, 1% Triton X-100) containing 1 mM PMSF and 2 EDTA-free protease inhibitor cocktail tablets (Boehringer Manheim). Cells were lysed by sonication at 4°C using 10 second pulses 10 times with a 20 second rest between pulses. The lysates were clarified by centrifugation (Sorvall SS34 rotor) at 20 000 rpm and 4°C for 60 minutes.

Full-length H-NS_{C20S}, H-NS 1-89 and 1-89 K5P polypeptides were purified by two successive chromatography procedures. The cleared lysate was loaded onto a Talon metal (Clontech) affinity column (20 ml) equilibrated with 20 mM Tris-HCl (pH8), 500 mM NaCl, 5% glycerol at a flow rate of 1 ml/min. Contaminant wash was accomplished with a step gradient of 10 mM and 50 mM imidazole, respectively. The full-length and 1-89 H-NS polypeptides were eluted in 2 column volumes of 20 mM Tris-HCl (pH 7), 500 mM NaCl, 500 mM imidazole and 5% glycerol. The elution fraction was dialysed (3.5 kDa cut-off membrane, Spectrapor) against 2X 2L 20 mM Tris-HCl (pH8), 500 mM NaCl, 2.5 mM $CaCl_2$ and 5% glycerol at 4°C. The N-terminal histidine tag was removed by incubation with thrombin (Calbiochem) at 1 U/mg of recombinant protein for 3 hours at room temperature. The cleaved histidine tag and any uncleaved H-NS polypeptides were removed by the application of reaction products onto Talon metal affinity resin (20 ml) equilibrated with 20 mM Tris-HCl (pH 8), 500 mM NaCl, and 5 % glycerol. Samples were always taken before and after cleavage reaction to check for the extent of cleavage.

The H-NS 1-64, 12-64, 1-64 K5P and 1-64 Q16P polypeptides were purified as above with the following changes. Subsequent to second Talon affinity chromatography step, various truncated polypeptides and corresponding mutants were dialysed against 20mM Tris-HCl (pH8), 300 mM NaCl and 5% glycerol. This were then concentrated (section 2.3.3) and loaded on a Superdex 75 gel filtration column (Pharmacia) equilibrated in the same buffer.

The H-NS 12-89, 17-89 and 1-89 Q16P polypeptides were purified in the same manner as full-length H-NS. Due to their extremely low solubility, cell pellets were always resuspended in 200 ml lysis buffer (20 mM Tris-HCL, pH8, 500 mM NaCl, 10 % glycerol, 1 % Triton X-100) as this was found to prevent precipitation to a certain extent. Furthermore, all buffers used contained 10% glycerol. In addition, H-NS 12-89, 17-89 and 1-89Q16P polypeptides were always diluted 10 times with the buffer lacking imidazole (20 mM Tris-HCl (pH 7), 500 mM NaCl, 10% glycerol) prior to dialysis. This reduced H-NS aggregation during overnight dialysis at 4°C. Thrombin cleavage reaction was performed at 4°C for 24 hours in 20 mM Tris-HCl (pH8), 500 mM NaCl, 2.5 mM $CaCl_2$, 10 % glycerol.

The H-NS 20-89 and 1-89 K5P Q16P polypeptides were completely insoluble. Typical results from the SDS-PAGE analysis of various H-NS polypeptides are shown in Fig. 2.1.

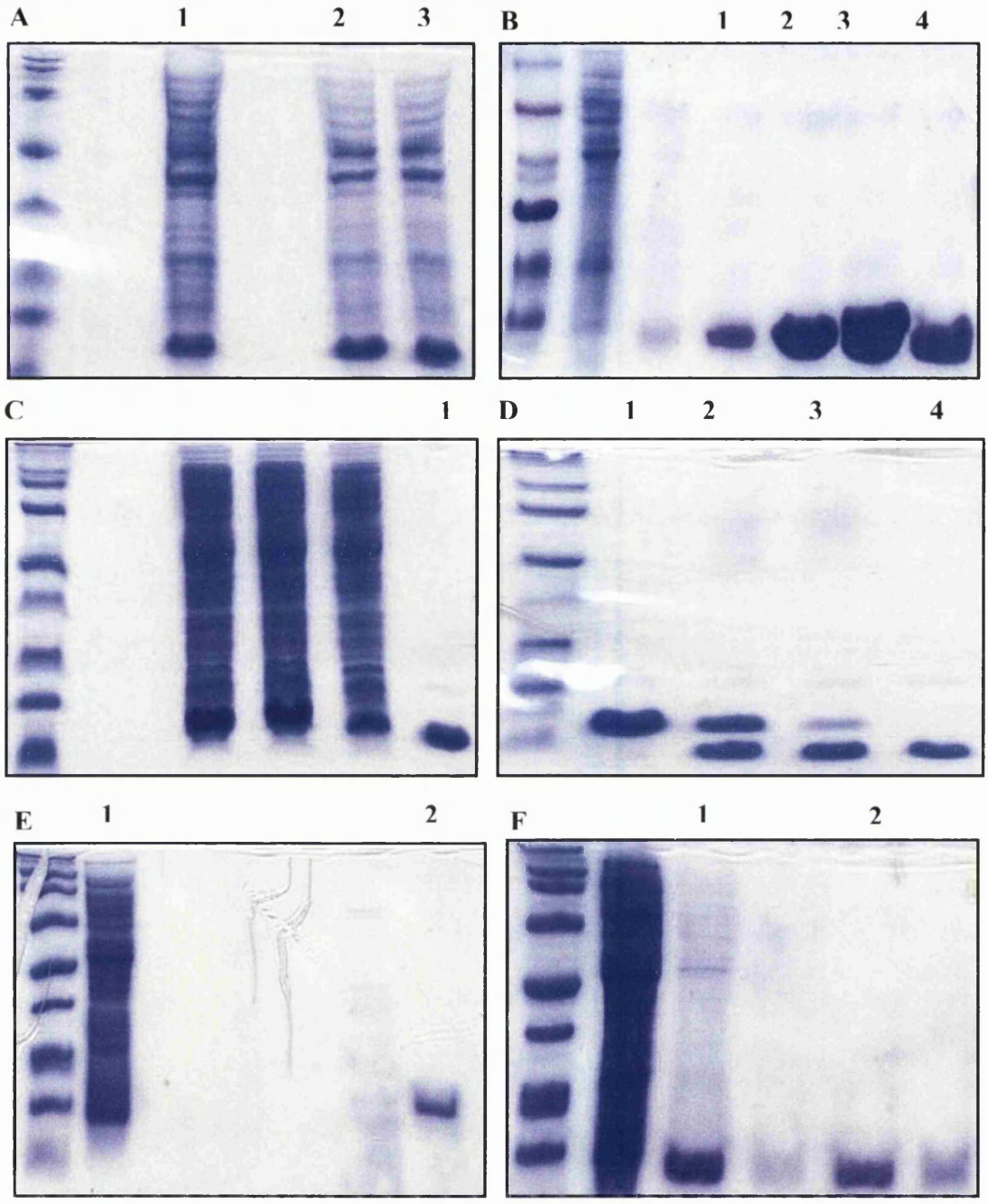


Figure 2.1 Typical results from the SDS-PAGE analysis of various H-NS polypeptides. (A) lanes 1-3: cleared lysates of H-NS 1-64, H-NS 1-64 K5P and H-NS Q16P respectively. (B) lanes 1-4: purified H-NS 12-64 polypeptide. (C) lane 1: purified H-NS 1-64. (D) lane1: H-NS 1-64 with histidine tag; lanes 2-4: thrombin cleavage at 1, 2 and 3 hours, respectively. (E) lane 1: H-NS 12-89 cleared lysate; lane 2: purified 12-89. (F) lane 1: H-NS 17-89 cleared lysate; lane 2: purified H-NS 17-89.

2.3.2 Purification of StpA polypeptide

Initial expression tests showed low StpA levels (from pT7-7 expression vector) as judged by the intensity of the band on the SDS-PAGE (section 2.3.5) gels of cleared cell lysates. In order to obtain better overproduction, the StpA gene was subcloned from pT7-7 plasmid into pET- 14b expression vector as described above. StpA expression was tested in BL21(DE3), BL21(DE3)pLysS and C41 cells, using different inducer concentrations (0.1-1 mM IPTG) and for different periods of induction.

A two-step procedure was developed which resulted in >95% purified protein preparation (Fig. 2.2 A, B). Initial stages of purification using TALON metal affinity chromatography were essentially as described for the full-length H-NS polypeptide with minor changes in the buffer composition. Frozen cells were resuspended in 50 ml lysis buffer (20 mM Tris-HCL, pH8, 500 mM NaCl, 5% glycerol, 1% Triton X-100) containing 1 mM PMSF and 2 EDTA-free protease inhibitor cocktail tablets (Boehringer Manheim). Cells were lysed by sonication at 4°C using 10 second pulses 10 times with a 20 second rest between pulses. The lysates were clarified by centrifugation (Sorvall SS34 rotor) at 20 000 rpm and 4°C for 60 minutes. The cleared lysate was loaded onto a Talon metal (Clontech) affinity column (20ml) equilibrated with 20 mM Tris-HCl (pH8), 500 mM NaCl, 10% glycerol at a flow rate of 1 ml/min. Contaminant wash was accomplished with a step gradient of 10 mM and 50 mM imidazole, respectively. The StpA polypeptide was eluted in 2 column volumes of 20 mM Tris-HCl (pH 7), 500 mM NaCl, 500 mM imidazole, 10% glycerol. The elution fraction was diluted 1 in 10 with 20 mM Tris-HCl (pH8), 500 mM NaCl, 10% glycerol. This was dialysed (3.5 kDa cut-off membrane, Spectrapor) against 2X 2L 20 mM Tris-HCl (pH8), 500 mM NaCl, 5% glycerol at 4°C.

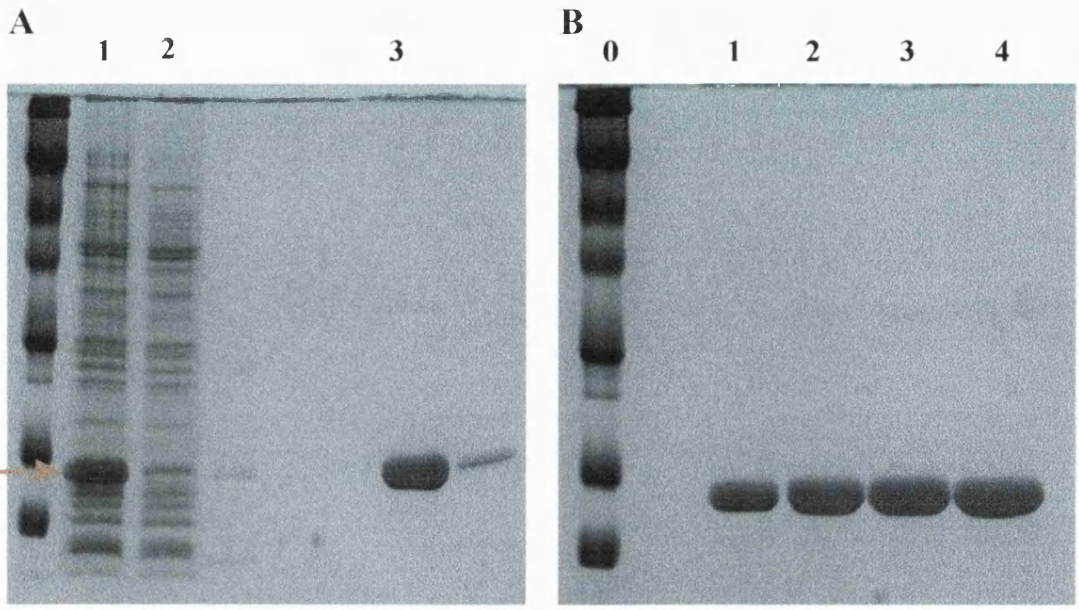


Figure 2.2 SDS-PAGE analysis of the Stp-A purification. Result shown in (A): lane 1: cleared lysate loaded onto Talon column (red arrow indicates over-expresses StpA polypeptide); lane 2: Talon flow-through; lane 3: 0.5 M imidazole elution fraction containing Stp-A polypeptide. Results shown in (B): lanes 1-4: DNA cellulose 0.5-1M NaCl elution of StpA polypeptide; lane 0: molecular weight markers in descending order: 97, 66, 45, 30, 20.1 and 14.3 kDa.

The removal of the N-terminal histidine tag caused immediate aggregation of the StpA polypeptide. This step was therefore omitted in subsequent purifications.

Second purifications step was employed in order to remove several contaminants of higher molecular weight present in small amounts. StpA polypeptide was applied to a 25 ml DNA cellulose column (Sigma) equilibrated in buffer A (20 mM Tris-HCl (pH8), 500 mM NaCl, 5% glycerol) and the protein was eluted with a 0.5-1 M NaCl gradient. StpA preparations were judged pure from the absence of detectable contaminating bands in overloaded gels (e.g. Fig. 2.2 B, lane 4).

2.3.3 Protein concentration

Both StpA and various H-NS polypeptide samples were concentrated using a Centriprep concentrator (Amicon) with a 3000 Da molecular weight (MW) cut-off at 4°C and then dialysed into required buffer using Spectra/Por (Pierce) dialysis membrane.

2.3.4 Protein storage

Purified protein samples were stored at 4° in 20mM Tris-HCl (pH8), 0.5M NaCl, 10% glycerol. In order to prevent any microbial growth, sodium azide (NaN_3) was added to a final concentration of 0.05 %(v/v) to protein samples.

2.3.5 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Tris-Tricine SDS PAGE was used to analyse H-NS and StpA protein samples. A MiniProtean II gel apparatus (BioRad) was used throughout. The stacking (4% acrylamide) and resolving (12% acrylamide) gel components are detailed in Table 2.3. Protein samples were boiled in SDS sample buffer (5X stock containing 50 % glycerol, 10 % SDS, 0.5 M Tris-HCl, pH 6.8) for 5 minutes, micro-centrifuged at 13000 rpm for 20 seconds and loaded using a 25 µl Hamilton syringe (Supelco). Acrylamide gels were electrophoresed at 100V.

Component	12% resolving gel	4% stacking gel
Gel buffer (3M Tris-HCl, pH 8.4 0.3% SDS)	5.6 ml	2 ml
30% acrylamide/bisacrylamide	5 ml	0.8 ml
50% glycerol	3.4 ml	-
10% APS	60 µl	50 µl
TEMED	20 µl	20 µl
H_20	-	3.2 ml

Table 2.3 Summary of the constituents of the 12% Tris-Tricine SDS-PAGE gels.

2.3.6 Coomassie blue staining of proteins in acrylamide gels

Following electrophoresis, the stacking gel was cut away and the gel was placed in a plastic container. The gel was submerged in an aqueous solution containing 0.2% (w/v) coomassie brilliant blue R (Sigma), 45% (v/v) methanol and 10% (v/v) acetic acid for 1 hour. Protein bands were then visualised by washing the gel in a de-staining solution (20% (v/v) methanol and 10% (v/v) acetic acid).

2.4 Characterisation of purified proteins

2.4.1 Measurement of protein concentration by UV spectroscopy

Protein concentrations were measured using Cary 50 UV-Visible spectrophotometer (Varian) in a 1 cm pathlength quartz cuvette (Hellma). Extinction coefficients of H-NS polypeptides were calculated from amino acid sequences by the method of Gill and von Hippel (1989) shown in Table 2.4. The protein concentration was calculated using Beer-Lambert law (equation 2):

$$A_{280} = \varepsilon \cdot c \cdot l \qquad (2)$$

where A_{280} is the absorbance at 280 nm, ε is the molar extinction coefficient (M^{-1} cm^{-1}), c is the protein concentration (M) and l is the cell pathlength (cm). For H-NS and StpA, calculated values stated in the text refer to the protein monomer concentrations.

Protein	Extinction coefficient ($M^{-1}cm^{-1}$)
H-NS full length	8250
N-NS 1-89 wt, K5P, Q16P	1280
H-NS 1-64 wt, K5P, Q16P	1280
StpA	12660

Table 2.4 Summary of the extinction coefficients used.

2.4.2 Analytical ultracentrifuge (AUC)

Analytical ultracentrifugation experiments (sedimentation equilibrium) were performed using an Optima XL-A analytical ultracentrifuge using An60Ti rotor (Beckman Instruments, Inc., Fullerton, CA). Samples used for sedimentation equilibrium were loaded into Epon charcoal-filled six-channel centrepieces, and were sedimented to equilibrium at the indicated rotor speeds, with the experiments at the slowest rotor speeds being performed first, followed by the experiments at increasing rotor speeds, with 120 μl of sample loaded per chamber for all sedimentation equilibrium experiments. Absorbance data were collected by scanning sample cells at a wavelength of 220 nm or 280 nm at intervals of 0.003 cm with five averages in a continuous scan mode.

Data analysis was carried out using a nonlinear least square analysis in the ORIGIN (v 4.1) package (Beckman Instruments). Data were fit to single-species model both as individual data sets and as a global fit of all data sets for each of the H-NS polypeptide studied using equation:

$$C_r = C_M(r_0)\exp[M\phi(r^2 - r_0^2)] \quad (3)$$

where C_r is the protein concentration at position r, r_0 is arbitrary radius point and $C_M(r_0)$ is the monomer concentration at r_0, M is the molecular mass of the monomer, $\phi = \omega^2(1-\nu\varphi)/(2RT)$, and T, R, ω, ν and φ are the absolute temperature, gas constant, angular speed of the rotor, partial specific volume of the monomer and solvent density, respectively. A modified version of the equation (3) including terms for self-association was also used. The validity of different models was evaluated by the distribution of residuals and the variance. The specific volume (ν) was estimated from the amino acid composition of the wild-type H-NS and the respective deletion and mutant polypeptides using the program

SEDNTERP (John Filo, Amgen). The solution density (φ) was calculated for each buffer using SEDNTERP.

2.4.3 Analytical gel filtration

The protein size and thus the oligomerisation state were determined using a size-exclusion chromatography column with Superose 12 HR 10/30 (Amersham Pharmacia Biotech) on an ÄKTA FPLC system (Pharmacia, Sweden) at 25°C. The protein samples (100-200µL) were injected onto the pre-equilibrated column and eluted under isocratic conditions at a flow rate of 0.3 ml/min. Sample was detected by measuring absorption at 280 nm.

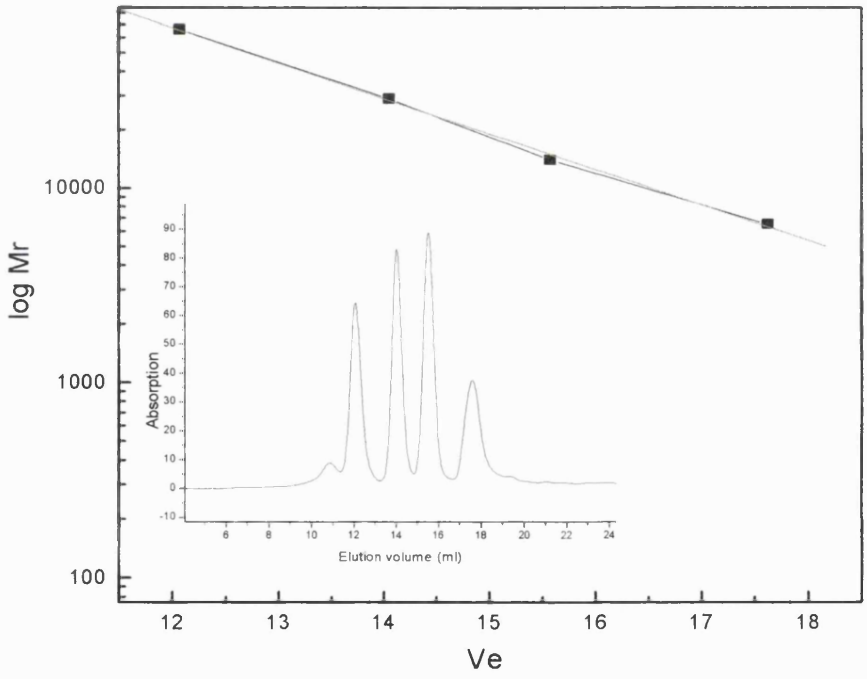


Figure 2.3 Linear fit of the of the logarithm of the molecular weight versus the elution volume. The apparent molecular weight of the sample protein was then obtained by comparing the elution volumes with the standard curve. Molecular weight standards are: BSA (66 kDa), carbonic anhydrase (29 kDa), cytochrome c (15 kDa) and apoferritin (6.5 kDa).

The column was calibrated by running molecular weight standards spanning a range of sizes (6.5-66 kDa) (BioRad). The plot of the logarithm of molecular weight of selected standards versus elution volume is shown in Fig. 2.3. The elution volumes were estimated using the Pharmacia UNICORN Evaluation package. Samples of appropriate H-NS polypeptides were equilibrated in the same buffer as present in the column, and allowed to equilibrate for 10 minutes before they were loaded onto the column.

2.4.4. Circular dichroism (CD)

Measurements were performed with Aviv 202SF spectropolarimeter (Aviv Instruments Inc.) connected to a thermostated water bath at 25°C. The wavelength dependence of [θ] between 190 and 260 nm was monitored using five scans in 1 nm increments with a sampling time of 5 seconds. Results are expressed as mean residue molar ellipticity [θ] ($deg \cdot cm^2 \cdot dmol^{-1}$) calculated from equation (4)

$$[\theta] = (\theta_{obs} \times MRM)/(10 \times l \times c) \quad (4)$$

where θ_{obs} is the observed ellipticity expressed in millidegrees, MRM is the mean residue molar mass (molecular mass of the polypeptide divided by the number of amino acid residues), l is the optical path length in cm, and c is the final polypeptide concentration in mol/L.

An analysis of protein secondary structures from CD spectra was performed with DICHROWEB accessible via a server located at http://www.cryst.bbk.ac.uk/cdweb (Lobley *et al.*, 2002). Percent helicity was calculated according to the following equation:

$$^{max}\text{Content } (\alpha\text{-helix}) = \theta_{MRE222} / -40.000 \cdot [1-(2.5/n)] \quad (5)$$

where ^{max}Content (α-helix) represents maximal helical content, θ_{MRE222} mean residue molar ellipticity at 222 nm and n number of amino acid residues in the polypeptide.

2.4.5 Nuclear magnetic resonance (NMR)

NMR spectra were acquired on Varian Unityplus 500MHz spectrometer at a temperature of 25°C. The experiment recorded on ^{15}N labelled H-NS 12-64, 1-64 K5P and Q16P polypeptides was 2D ^{15}N HSQC. Spectra were manually assigned upon comparison with 2D ^{15}N HSQC spectrum of wild-type H-NS 1-64 polypeptide (Renzoni *et al.*, 2000).

2.4.6 Sequence analysis

Multiple sequence alignment of H-NS homologues was performed with CLUSTALW method (Thomson *et al.*, 1996). Aligned sequences were edited using BioEdit (v5.0.9) software (Hall, 1999). The accession numbers for the sequences are: H-NS *E.coli*, P08936; H-NS *Salmonella typhimurium*, P17428; H-NS *Shigella flexneri*, P09120; H-NS *Erwinia chrysanthemi*, X89444; H-NS *Serratia marcescens*, P18955; H-NS *Proteus vulgaris*, P18818; StpA *E.coli*, P30017; StpA *S.typhimurium*, AF009363; VicH *Vibrio cholerae*, AJ010791; BpH3 *Bordetella pertussis*, U82566; HvrA *Rhodobacter capsulatus*;

Prediction of coiled-coil structure in the H-NS polypeptide was performed using COILS program (http://www.ch.embnet.org/software/COILS_form.html) (Lupas *et al.*, 1991). Three-dimensional structure of the N-terminal region of *E.coli* StpA polypeptide was calculated by homology modelling with the NMR structure of the N-terminal dimerisation domain of *S.typhimurium* H-NS (Esposito *et al.*, 2002) using SWISS-MODEL (http://swissmodel.expasy.org) (Schwede *et al.*, 2003). Swissmodel project files contained

atomic coordinates contained in the 1LR1 PDB file (H-NS 1-57) and the sequence alignment between first 57 residues of H-NS and StpA.

3. Results I: The effect of the N-terminal truncations on the oligomeric properties of H-NS protein

3.1 Introduction

It is now well established that the ability of the H-NS polypeptide to recognise DNA is intimately linked to its self-associative properties via the formation of a precisely defined nucleoprotein complex (Schröder and Wagner, 2002). Within the context of the H-NS polypeptide chain, these two functions are separated by the presence of the flexible linker (see Chapter1). The formation of such assemblies has an effect on various DNA transactions, most notably on gene transcription and the structural organisation of the nucleoid. On the more "global" scale, the interplay between environmental signals, cellular levels of H-NS, interactions with other protein factors, changes in the DNA structure and possibly conformation(s) of the H-NS polypeptide itself results in a finely orchestrated cellular response.

Attempts at deciphering the structure/function relationship of the H-NS polypeptide highlighted the unusual and complicated connection between the various regions in the primary sequence (Ueguchi *et al.*, 1996). Based on the experimental evidence presented by Smyth *et al.* (2000), a possible model for the H-NS self-assembly has emerged. It is hypothesised here that the region encompassing H-NS 1-64 polypeptide constitutes a basic structural unit that in the presence of additional C-terminal residues (65-89) assembles to higher-order oligomers characteristic of the wild-type H-NS. The exact oligomeric state of

the basic structural unit has been a matter of controversy (Smyth *et al.*, 2000; Badaut *et al.*, 2002). Here the H-NS 1-64 truncated polypeptide from *S.typhimurium* is shown to adopt a dimeric configuration.

It is further hypothesised in this thesis that the dimeric basic structural unit could thus serve as an important intermediate in the assembly process since the formation of a large H-NS macromolecular scaffold could be based on the association of isolated dimers. The monomer-dimer-oligomer association would therefore require the existence of two protein-protein interfaces, one responsible for the formation of initial dimers (residing in the first 64 residues) and the other responsible for higher-order oligomers.

A deletion mutagenesis strategy was adopted, in order to gain further insights in the mechanism of H-NS self-assembly and to identify the interfaces leading to the oligomer formation. An attempt has been made here to prevent the self-associative behaviour characteristic of the full-length H-NS polypeptide by deleting one of the hypothetical interacting surfaces. The truncated polypeptides were analysed using an array of biophysical techniques to gain insights into their hydrodynamic and structural properties. The results thus obtained are used to build a model, which reveals how the dimeric units self-associate in a head-to-tail manner and highlights some of the possible secondary structure elements necessary to achieve this.

3.2 The H-NS full-length protein and its C-terminal derivative H-NS 1-89 form polydisperse oligomers in a concentration-dependent manner

The C-terminally deleted H-NS polypeptide lacking the DNA binding domain has been subjected to size-exclusion chromatography to highlight the similarity in the hydrodynamic behaviour to the full-length H-NS. The Superose 12 gel-filtration traces obtained for H-NS 1-89 polypeptide are shown in Fig. 3.1. The H-NS 1-89 elution profiles at a range of different concentrations are characterised by a sharp leading edge and an extended trailing edge. This elution behaviour is expected for rapidly equilibrating solutes and has been observed experimentally elsewhere (Stevens, 1986, 1989). The characteristic elution profile in terms of broad and asymmetric peak shapes and the distinctive concentration dependence of the peak maxima are reminiscent of the solution behaviour exhibited by the full-length H-NS polypeptide (Smyth *et al.*, 2000). Given a monomer molecular mass of 10.5 kDa, the H-NS 1-89 oligomer composition ranges from six ($Mr/M_{r\,(calc.)}$=5.7) to eleven ($Mr/M_{r\,(calc.)}$=11.6) subunits (obtained by a comparison with the elution volumes of a mixture of standard proteins of known molecular mass – see Fig. 2.3). It should be noted that calculated molecular masses must be considered as approximate estimations, as they are based on comparison with that of the globular protein standards. The gel-filtration traces obtained for H-NS 1-89 polypeptide indicate that no single defined state was reached within the concentration range tested (120 μM-270 μM).

In accordance with other biochemical data these results substantiate the notion that the determinants of the self-associative behaviour of the H-NS polypeptide reside in the N-terminal portion, whereas the residues in the C-terminal region of the protein do not make contribution to the H-NS oligomer formation (Ueguchi *et al.*, 1996; Smyth *et al.*, 2000).

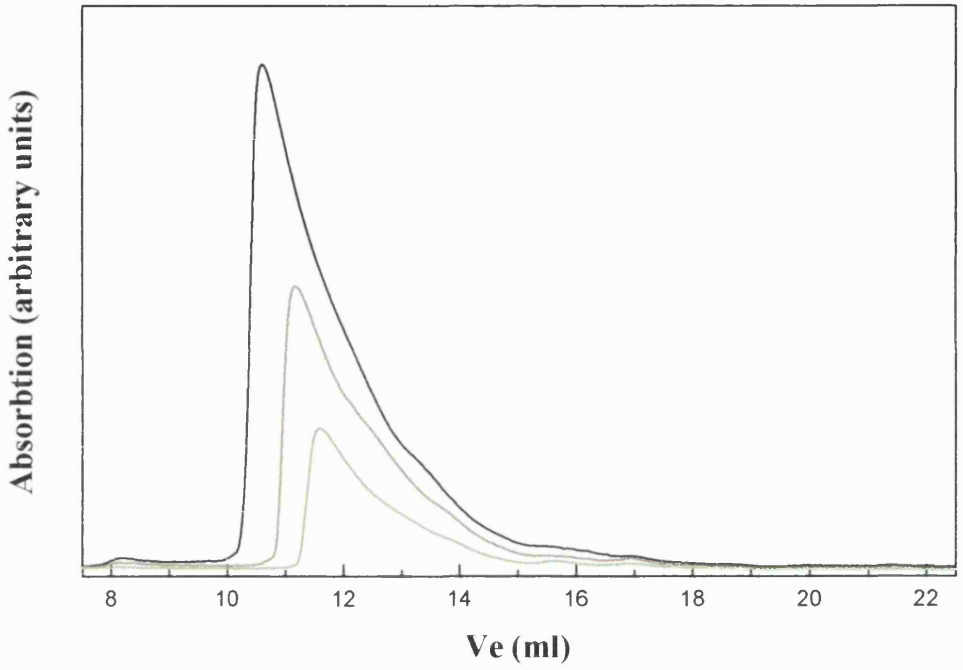


Figure 3.1 Size-exclusion chromatography profile of the H-NS 1-89 polypeptide. H-NS 1-89 at 270 (–), 170 (–) and 120 (–) μM. All experiments were performed in 20 mM KPi (pH 7), 300 mM NaCl and 5% (v/v) glycerol at 25°C. The elution profile shows absorption at 280 nm as a function of elution volume (V_e).

Furthermore, these results rationalise the use of H-NS 1-89 polypeptide as a substrate to probe mechanisms responsible for self-assembly.

3.3 The N-terminal H-NS 1-64 fragment forms a dimer

The N-terminal polypeptide encompassing first 64 amino acid residues (H-NS 1-64) has been shown previously, as demonstrated by size-exclusion chromatography to associate to a single, defined, oligomeric state (Smyth *et al.*, 2000). A comparison with molecular-weight standards indicated an approximate molecular weight of 26 kDa, which corresponds to an oligomeric state that is between trimer and tetramer. A previous sedimentation equilibrium study reported that the H-NS 1-64 polypeptide chain from *S.typhimurium*

forms trimers (Smyth *et al.*, 2000). The study was performed at total H-NS concentrations of 430 and 570 μM and at two rotor speeds (8000 and 15000 rpm).

In this section, experimental evidence on the self-associative properties of H-NS 1-64 polypeptide has been re-evaluated since the original data were collected using lower rotor speeds in expectation of oligomeric states in the range of trimers and tetramers based on the gel-filtration experiments. Furthermore, sedimentation equilibrium runs were performed at 20 °C (compared with 4°C in the previous study), as this allows direct comparison with gel-filtration experiments, which were performed at room temperature.

Representative sedimentation equilibrium data for H-NS 1-64 are shown in Fig. 3.2. The H-NS 1-64 samples were loaded at three different concentrations (160, 300 and 450 μM) and allowed to reach equilibrium. As a first step in the analysis, data sets were fit individually to an ideal, single species model in which the molecular mass was allowed to vary. In each case, the apparent molecular mass was close to that expected for a dimer. Other possible models were excluded by fitting data sets with the molecular mass held constant at the known monomeric value to single species, monomer-dimer and monomer-trimer models. In each case, the data were fit best by a single species dimer model, as indicated by the random distribution of residuals and variances close to 1. Significant deviations of residuals are observed for fits to models other than the ideal dimer (Fig. 3.2), most notably monomer (Fig. 3.2. (c)) and trimer (Fig. 3.2 (d)). The apparent molecular mass of the H-NS 1-64 polypeptide determined from the sedimentation equilibrium experiment is 15.7 kDa, which is within 1% of the theoretical molecular mass expected for a dimer.

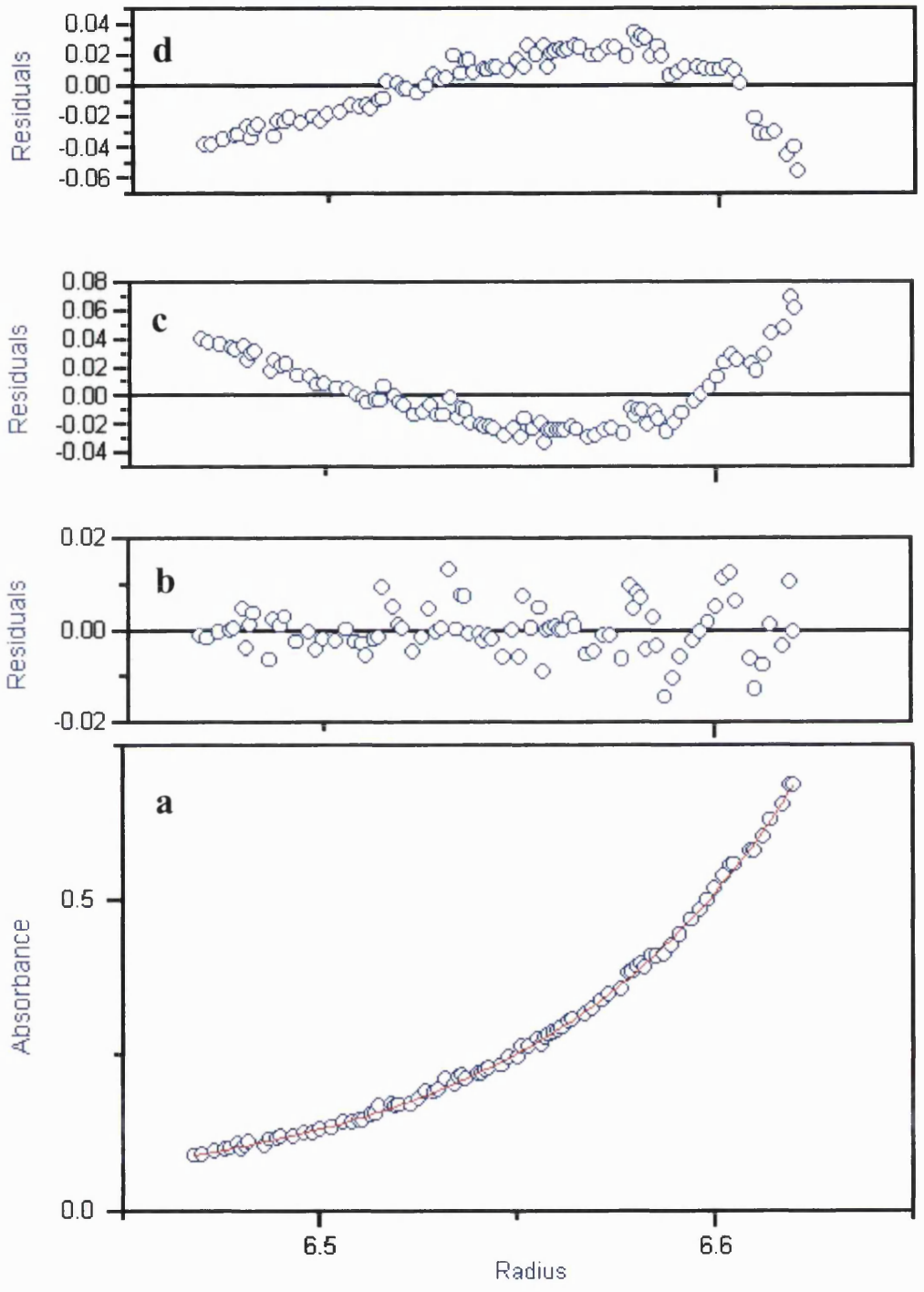


Figure 3.2 Representative sedimentation equilibrium data for the H-NS 1-64 polypeptide. (a) H-NS 1-64 centrifuged to equilibrium at 20 °C in 20 mM KPi (pH7) and 300 mM NaCl. Open points were obtained at 36,000rpm. The continuous line represents the best fit of equation (3), corresponding with the dimer model. (b) Residuals between measured data obtained in (a) and data fitted for a species with the molecular weight of 15735 g/mol. (c,d) Residuals between measured data obtained in (a) and data fitted as monomer (c) and trimer (d). Of various models, the distribution of residuals indicate that the data are best accounted for by a single species dimer model.

Equivalent results were obtained for H-NS 1-64 from several independent preparations. These results agree with those obtained from study by Badaut *et al.* (2002) in which the *E.coli* H-NS 1-64 polypeptide chain was shown to behave as a dimer by both sedimentation equilibrium and cross-linking studies.

Indeed, the first structural characterisation of the H-NS N-terminal domain fragment (see Introduction) comprising first 57 residues was modelled as a dimer (Esposito *et al.*, 2002). Subsequent solution NMR studies by Bloch *et al.* (2003) on the *E.coli* N-terminal 46 residues confirmed these findings. Recent studies on the crystal structure of the N-terminal oligomerisation domain of VicH (H-NS protein of *Vibrio cholerae*) also established the dimeric nature of this fragment (S. Arold, personal communication). Thus, it seems that dimerisation of the N-terminal region of H-NS through coiled-coil structural motif represents a common theme.

In the light of the three-dimensional structure of the H-NS 1-64 polypeptide chain, the anomalous elution behaviour of this fragment can be explained by its elongated, non-globular shape. Based on the evidence obtained from the sedimentation equilibrium and gel-filtration experiments (data not shown), it can be concluded that the oligomeric state of the H-NS 1-64 fragment is stable across the concentration range examined, as there is no evidence of the formation of higher-order species. Coiled-coil regions *per se* are capable of promoting only a limited number of oligomeric states including two, three, four and five-stranded structures and are therefore incompatible with the higher oligomers seen for H-NS full-length polypeptide (O'Shea *et al.*, 1991;Harbury *et al.*, 1993, 1994; Malashkevich *et al.*, 1996).

Thus, the structural elements that lie outside the 1-64 region could be responsible for higher order oligomer formation either alone or in conjunction with the 1-64 area.

A model is presented below to explain how the presence of two oligomerisation interfaces leads to the self-associative behaviour characteristic of the full-length H-NS (Fig 3.3):

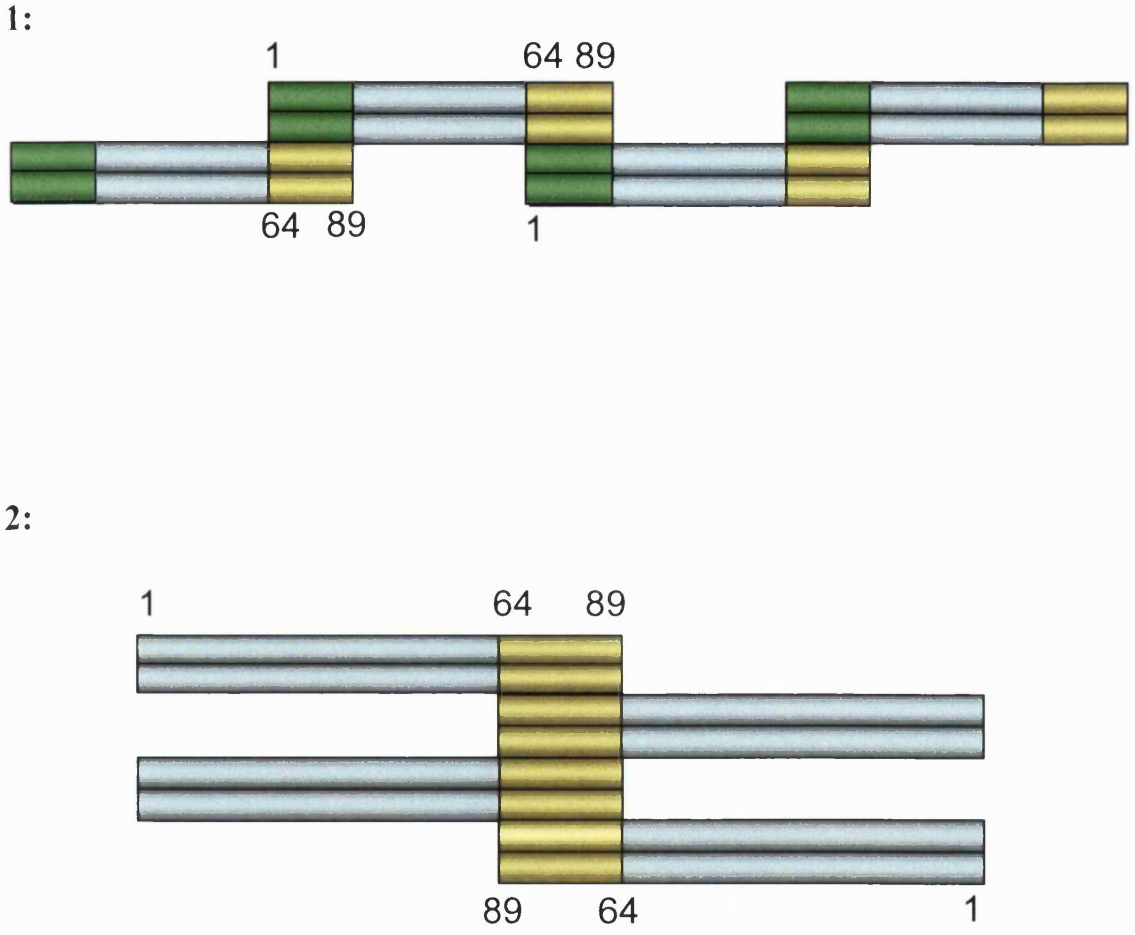


Figure 3.3 Schematic representation of the two possible models used to explain the self-associative behaviour of the H-NS polypeptide. (1) head-to-tail model (2) tail-to-tail model.

The dimeric H-NS 1-64 fragment constitutes the basic structural module (**dimerisation** domain), which on its own is not sufficient to form high-order oligomers. In the presence of additional 25 residues (65-89) a second interface is created (**oligomerisation** interface) which causes the protein to polymerise in a fashion characteristic of the wild-type H-NS.

In the first model, self-assembly of the H-NS protein involves the interaction between the oligomerisation interface (i.e. residues 65-89) and the N-terminal region of the dimerisation domain (so called "head-to-tail model"). In principle, this model agrees with the experimental evidence, which demonstrates that the deletion of residues 65-89 within the context of the H-NS 1-89 polypeptide precludes self-assembly of the H-NS dimer. In model 2 (so called "tail-to-tail model") self-assembly is achieved solely through the oligomerisation interface (65-89). The available experimental evidence is again consistent with model 2. To differentiate between these two models a deletion mutagenesis strategy was adopted. Deletion of the residues from the N-terminus of the H-NS 1-89 polypeptide would be expected to prevent the formation of higher order oligomers (i.e produce a dimer) if the head-to-tail model is correct (model 1). On the other hand, the N-terminal truncations have no effect on the self-association behaviour if the dimer-dimer interactions were mediated through the residues 65-89 alone (model 2).

3.4 Oligomeric properties of the H-NS 12-89 polypeptide

The effect of the N-terminal deletions have been examined in the context of the H-NS 1-89 polypeptide chain as this was shown to have indistinguishable self-association behaviour from the full-length H-NS (see section 3.2). In the absence of any structural information on the N-terminal part of the H-NS molecule, secondary structure prediction programs have been used to define the region (Fig. 3.4) with the highest propensity for coiled-coil formation.

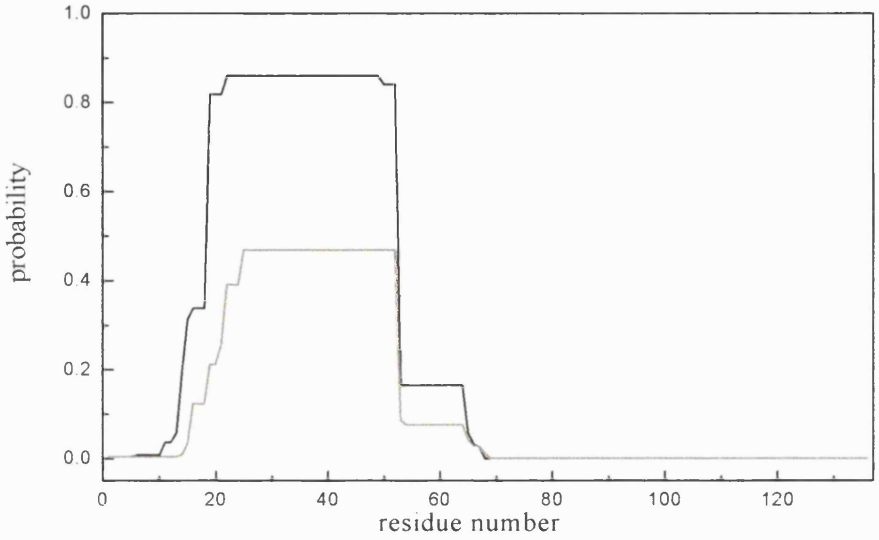


Figure 3.4 Prediction of propensity for coiled-coil formation in the H-NS full-length polypeptide. The plot shows probability for coiled-coil formation as a function of an amino acid sequence (1-136) using two different scoring matrices MTDK (–) and MTK (–).

As already noted before (Dorman *et al.*, 1999), the region in the amino-terminal between amino acids 15-60/70 has the highest probability for coiled-coil formation. The truncated version of the H-NS 1-89 polypeptide chain, in which first 11 residues were deleted, was examined since this region was predicted to lie outside the region most likely involved in coiled-coil interactions. Analytical size-exclusion chromatography has been used to ascertain the nature of the oligomeric state adopted by the H-NS 12-89 polypeptide. The Superose 12 gel filtration traces of H-NS 12-89 are shown in Figure 3.5. The drastic change in elution profile is apparent upon comparison with that of the H-NS 1-89 (Fig. 3.1). Unlike the broad and asymmetric peaks, which are characteristic of the H-NS 1-89 polypeptide, the behaviour of H-NS 12-89 clearly shows different oligomerisation properties. At all concentrations tested H-NS 12-89 elutes as two peaks, which merge into each other (one major peak with a shoulder). Molecular mass standards showed that the

major peak corresponded to an apparent molecular mass of approximately 25 kDa, while the second one corresponded to 38.5 kDa.

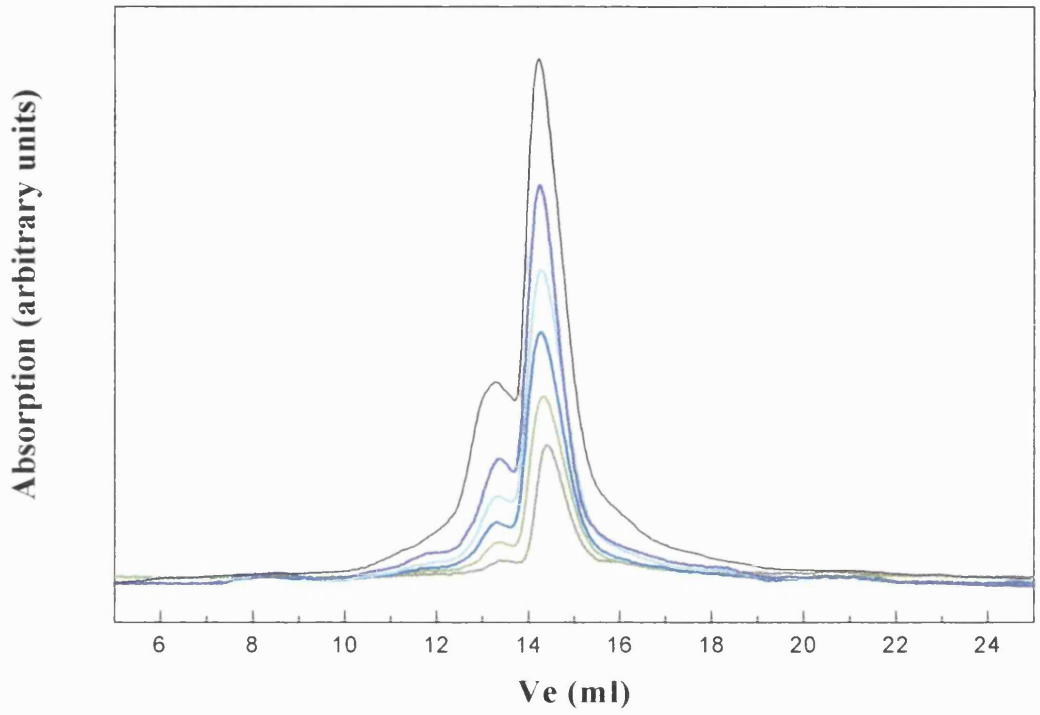


Figure 3.5 Size-exclusion chromatography profile of the H-NS 12-89 polypeptide. H-NS 12-89 at 30 (–), 24 (–), 21 (–), 18 (–), 16 (–) and 13 (–) μM. All experiments were performed in 20 mM KPi (pH 7) and 300 mM NaCl at 25°C. The elution profile shows absorption at 280 nm as a function of elution volume (V_e).

Upon comparison with the monomer molecular weight (9.2 kDa) the main peak can be assigned a value that lies between the dimeric and trimeric state, based on globular protein standards ($Mr/M_{r\ (calc.)}$ ~2.7). The second merging peak eluting at smaller values corresponds to tetramer ($Mr/M_{r\ (calc.)}$ ~4.2). Successive experiments in which H-NS 12-89 loading concentrations were reduced from 30 μM to 13 μM showed that the relative elution times of both species changes relatively little over the range tested. This is in stark contrast to the chromatographic behaviour of the H-NS full-length and H-NS 1-89 polypeptides. The concentration range of H-NS 12-89 polypeptide examined was small due to the

previously mentioned solubility problems. Nevertheless, it is apparent that the relative proportion of the two eluting species changes with the progressive dilution of the sample (compare the ratio of peak heights at different concentrations). These results are expected for mass-action law, because if the reaction is at equilibrium (e.g. dimer $\rightleftharpoons$ tetramer), the proportion of tetramer to dimer should depend on the concentration.

In the light of the recent structural information on the N-terminal fragment (residues 1-57), truncation of the N-terminal region in the H-NS 12-89 polypeptide is likely to compromise the integrity of the H2 (see 1.3.3.1), besides completely deleting H1. This could favour opening up of the structure, thus exposing hydrophobic residues, which line the bottom of the groove formed between the two major helices (H3) from each monomer that are otherwise protected by "folding over" of the H1 and H2 (Fig. 1.3). The exposure of the hydrophobic "patch" could enhance the tendency for aggregation. Thus, the hydrodynamic properties of the H-NS 12-89 polypeptide observed during the chromatographic run could suggest that the H-NS 12-89 dimer (see below) is in slowly interconverting form with the tetramer (i.e. non-specific dimer-dimer interaction)

This conclusion is supported by the observation that the addition of glycerol (in order to reduce hydrophobic interactions) leads to changes in the elution behaviour of the H-NS 12-89 polypeptide. The Superose 12 gel-filtration traces for H-NS 12-89 at concentrations of 25 μM and 50 μM in 20mM KPi (pH 7), 300 mM NaCl and in the presence of 5%(v/v) glycerol are shown in Figure 3.6. At both concentrations tested H-NS 12-89 elutes as a single peak at a volume similar to the main peak shown in Fig. 3.5 (14.1 ml compared with 14.2 ml), which corresponds to an oligomeric state that is close to that of a trimer. There is lack of concentration dependence of the peak maxima, and both peaks exhibit near

symmetrical nature. The difference in oligomeric behaviour between H-NS 1-89 and its truncated version H-NS 12-89 is further highlighted in the same figure (Fig. 3.6)

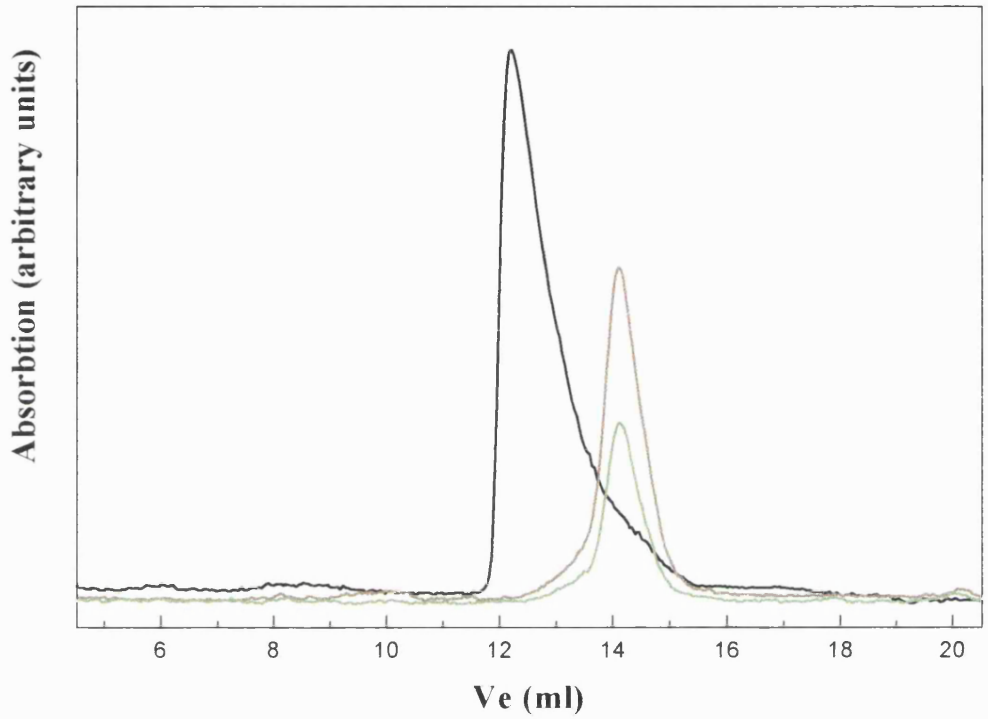


Figure 3.6 Size-exclusion chromatography profile of the H-NS 12-89 polypeptide in the presence of 5%(v/v) glycerol. H-NS 12-89 at 50 (–) and 25 (–) μM. All experiments were performed in 20 mM KPi (pH 7), 300 mM NaCl and 5% (v/v) glycerol at 25°C. Also shown is the elution profile for H-NS 1-89 polypeptide at 80 (–) μM. The elution profile shows absorption at 280 nm as a function of elution volume (V_e).

3.5 Sedimentation equilibrium studies on the H-NS 12-89 polypeptide

Attempts have been made to exploit the sedimentation equilibrium AUC technique so as to provide an independent measurement of the molecular weight adopted by the H-NS 12-89 polypeptide. These measurements have an advantage over the chromatographic techniques used in Section 3.4 since the shape-dependant approximations stemming from the non-globular fold of the protein do not have to be taken into account when calculating the apparent molecular weight.

The sedimentation equilibrium experiments could not be performed, under the conditions used in section 3.4 for H-NS 12-89 and previously used for H-NS 1-64 since the H-NS 12-89 polypeptide exhibited time-dependant aggregation indicating the poor solubility of this construct.

However, the addition of 5% (v/v) glycerol to the buffer established the conditions in which the H-NS 12-89 polypeptide was stable over the course of the AUC experiments. The representative results of the sedimentation equilibrium experiments performed on the H-NS 12-89 polypeptide are shown in Fig. 3.7. The absorbance (220 nm) profiles from sedimentation equilibrium experiments shown in Fig. 3.7 were collected at four different speeds (20K, 32K, 36K and 42K) and two loading concentrations of H-NS 12-89 (50 and 100 μM). In order to determine the oligomeric state adopted by the H-NS 12-89 polypeptide, these data were analysed individually by non-linear, least-square (NLLS) methods using a single ideal species model (equation (3); see 2.4.2). The fit to a single species gave a molecular mass of 18.7 kDa, which is significantly bigger than that of a monomeric protein (9.26 kDa). In fact, this value lies close to the theoretical molecular mass calculated for a dimeric protein (18.5 kDa). Figure 3.7 (c,d) show fits of the data along with residuals, assuming a monomeric and trimeric species respectively. The non-random distribution of the residuals of both fits clearly indicates that none of these models represent the correct association behaviour. Only a model, which corresponds to a dimer species, gave a satisfactory fit with a random distribution of residuals as shown in Fig. 3.7 (b).

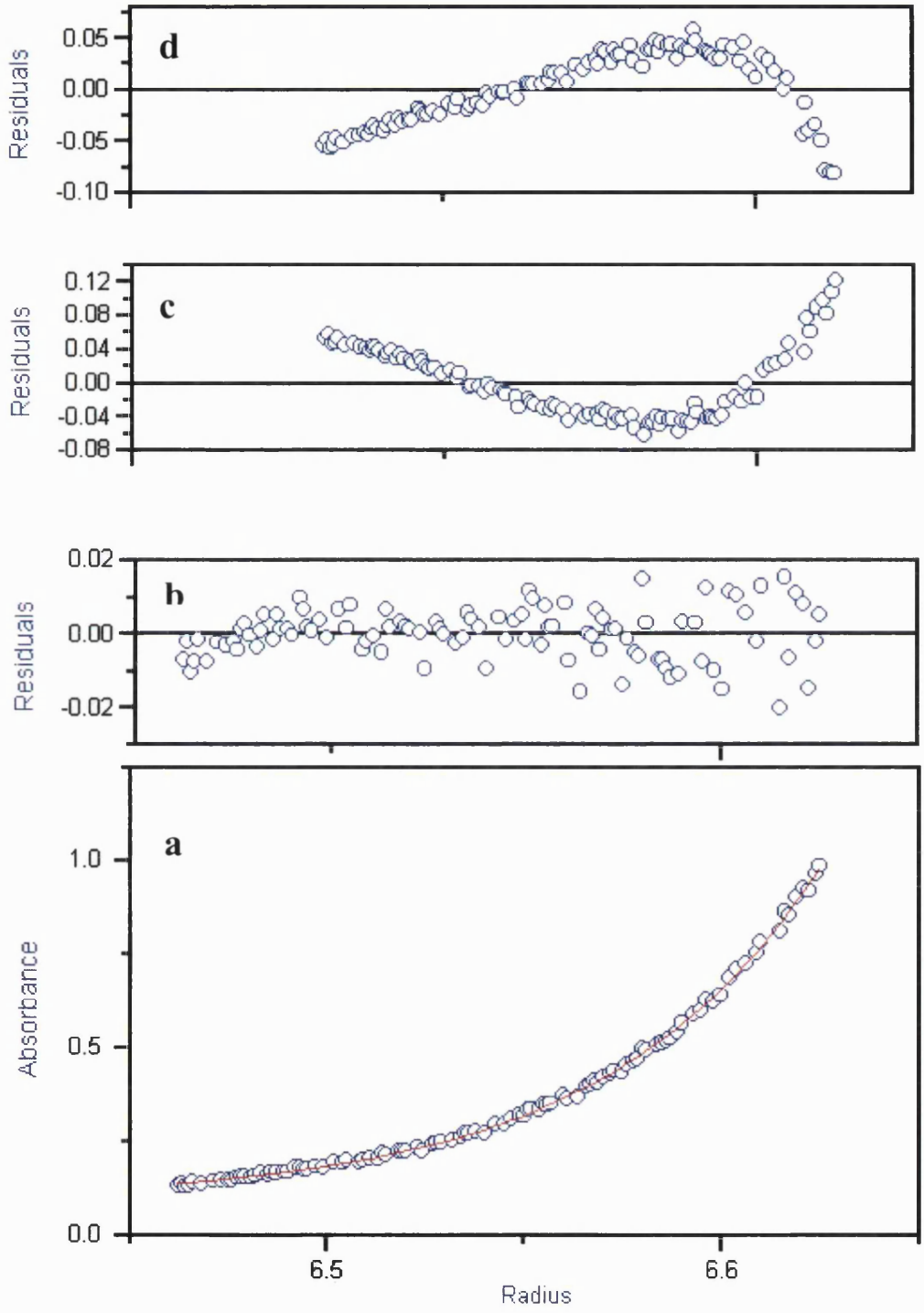


Figure 3.7 Representative sedimentation equilibrium data for the H-NS 12-89 polypeptide. (a) H-NS 12-89 centrifuged to equilibrium at 20 °C in 20 mM KPi (pH7), 300 mM NaCl and 5% (v/v) glycerol. Open points were obtained at 36,000 rpm. The continuous line represents the best fit of equation (3), corresponding with the dimer model. (b) Residuals between measured data obtained in (a) and data fitted for a species with the molecular weight of 18764 g/mol. (c,d) Residuals between measured data obtained in (a) and data fitted as monomer (c) and trimer (d). The distribution of residuals indicate that the data are best accounted for by a single species dimer model.

The size-exclusion chromatographic experiments run in the presence of 5% (v/v) glycerol gave a single peak at 14.2 ml. From these AUC experiments it can be concluded that the species shown in Fig. 3.6 corresponds with a dimer.

Taken together, these results are consistent with model 1 (Fig. 3.3), demonstrating that the higher-order self-assembly of the H-NS polypeptide occurs via a head-to-tail association of dimers. In the light of the previous studies, which suggest that the C-terminal domain of H-NS is separated from the N-terminal half by a flexible linker (Cusik and Belfort, 1998; Smyth *et al.*, 2000), model 1 assumes that the amino-terminal 89 residues act independently of the C-terminal DNA binding domain. The N-terminal 89 residues associate to form a dimer, which then self-assembles, through interaction of the N- and C-terminal residues of the core dimeric unit.

3.6 Size-exclusion chromatography of the H-NS 12-64 polypeptide

Recent investigation into the proteins coded by dominant negative mutants of the *hns* gene have shown that H-NS L26P polypeptide is unable to form higher order oligomers in solution in the manner characteristic of the wild-type H-NS (Badaut *et al.*, 2002). The introduction of a proline residue at this position is likely to compromise the stability of the coiled-coil (based on the analysis by Coils software; data not shown). Furthermore, it is known that this residue is part of a heptad repeat (see Fig. 1.1; residue L25) that contributes to the stability of the dimeric coiled-coil interface (Esposito *et al.*, 2002; Bloch *et al.*, 2003). In spite of the poor solubility of the H-NS L26P mutant polypeptide, it was shown to adopt a dimeric configuration at low concentrations (Badaut *et al.*, 2002).

With the aim of delineating the effects of the N-terminal truncation on the stability of coiled-coil interface, studies presented in this section have been undertaken.

Despite the dramatic change in the oligomeric behaviour of the H-NS 12-89 polypeptide, further structural characterisation proved difficult as the protein formed visible aggregates at concentrations above approximately 120 μM, even in the presence of 20% glycerol (which is below the requirements for standard NMR experiments). In order to examine the structural perturbation caused by the deletion of the N-terminal 11 residues, an alternative strategy was adopted in which the corresponding truncation was introduced in the H-NS 1-64 polypeptide. It was assumed that the structural changes in the N-terminal region lacking the first 11 residues are equivalent in the context of both H-NS 1-89 and H-NS 1-64 polypeptide chains.

The effect of the N-terminal truncation on the observed change in the oligomeric behaviour of the H-NS 12-89 polypeptide could have arisen in principle from two different structural perturbations: (1) *bona fide* disruption of the oligomerisation interface (at the N-terminal region) responsible for the high-order assembly in a head-to-tail manner or (2) "non-specific" disruption of the coiled coil region leading to the loss of one of the interacting regions responsible for self-association. The H-NS 12-64 construct was produced to discern between these and to serve as a control, which could indicate effects on the coiled-coil region.

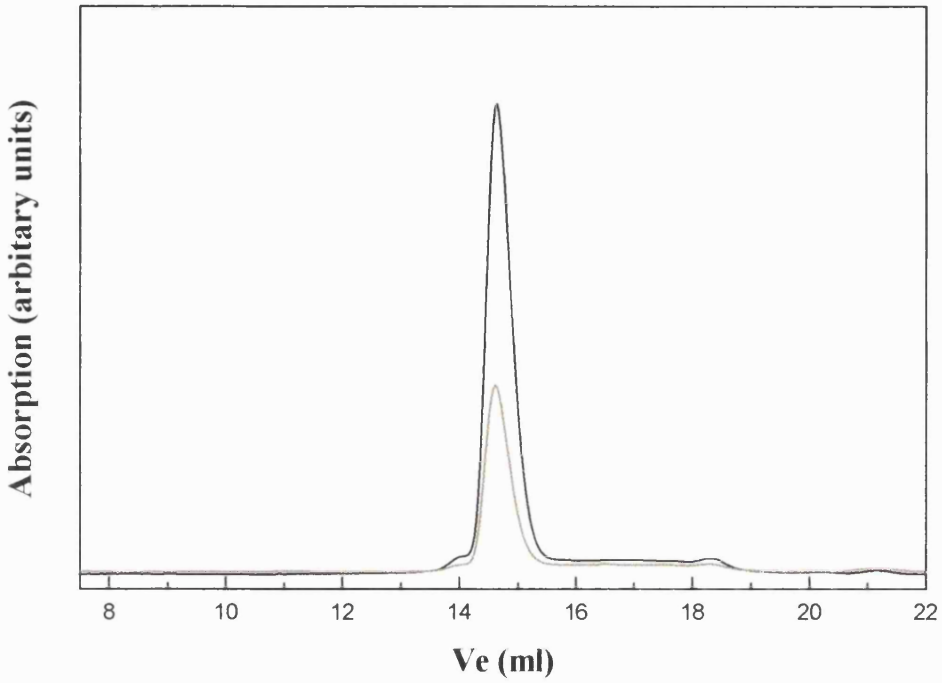


Figure 3.8 Size-exclusion chromatography profile of the H-NS 12-64 polypeptide. H-NS 12-89 at 139 (–) and 50 (–) μM. All experiments were performed in 20 mM KPi (pH 7), 300 mM NaCl and 5% (v/v) glycerol at 25 °C. The elution profile shows absorption at 280 nm as a function of elution volume (V_e).

The Superose 12 gel filtration traces obtained for H-NS 12-64 are shown in Fig. 3.8. The position and shape of the peaks are similar to that of the H-NS 1-64 polypeptide. There is no concentration dependence of the peak maxima, which suggest that H-NS 12-64 polypeptide adopts a single defined oligomeric state under the experimental conditions. Both peaks elute at 14.6 ml which, upon comparison with molecular weight standards, gives an apparent molecular weight of 21.8 kDa.

Upon comparison with the monomer molecular weight (6.7 kDa), the obtained value corresponds to an oligomeric state, which is slightly higher than a trimer ($Mr/M_{r\,(calc.)}$ ~3.2). This apparent molecular weight was also observed on size-exclusion studies of the 1-64 truncated version of H-NS (Smyth *et al.*, 2000). It is unlikely that the species observed in the gel filtration experiment represents trimer since the H-NS 1-64 polypeptide is only capable of adopting dimeric state. The most probable cause of anomalous partitioning in the gel filtration matrix lies in the elongated shape of the polypeptide. Indeed, the same explanation has been used to clarify the discrepancy between the estimated oligomeric states adopted by the H-NS 1-64 polypeptide obtained from two different experimental approaches (see 3.3).

The first 11 residues truncated in the H-NS 12-89 polypeptide extend into H2 (residues 10-16). It is possible that this truncation affects stability of the rest of the molecule. To assess whether the truncation leads to unfolding of the coiled-coil of H3 (difficult to ascertain by size-exclusion chromatography), CD spectroscopy was used to determine possible changes in the secondary structural content (Fig. 3.9).

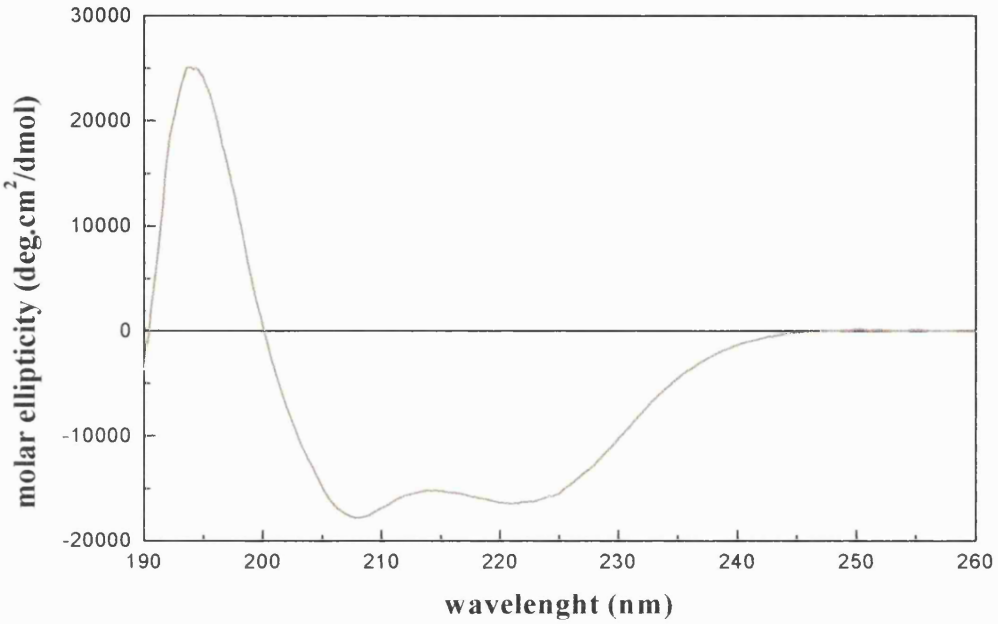


Figure 3.9 Far-UV CD scan of the H-NS 12-64 polypeptide. CD spectra of H-NS 12-64 polypeptide at 80 μM in 20 mM KPi, 10 mM NaCl at pH 7 and 25°C. Results are expressed as mean residue molar ellipticity calculated from equation (4).

The effect of N-terminal truncation on the stability of global fold of protein (coiled-coil) can be ruled out since the CD spectrum of H-NS 12-64 polypeptide clearly indicates that it is folded and shows a characteristic α-helical structure (Fig. 3.9). However, destabilisation of the part of coiled-coil region cannot be excluded (see 3.7 and chapter 4).

3.7 Structural characterisation of the N-terminal H-NS 12-64 polypeptide

NMR spectroscopy has been used here to assess the effect of the N-terminal deletion on the structural features of H-NS 1-64 polypeptide. The 2D ^{1}H-^{15}N HSQC spectrum for a uniformly labelled sample of H-NS 12-64 polypeptide is shown in Figure 3.10. This experiment gives the backbone amide resonances of all the residues in the protein, except proline. The equivalent spectrum of the H-NS 1-64 wild-type protein is shown in Fig. 3.11.

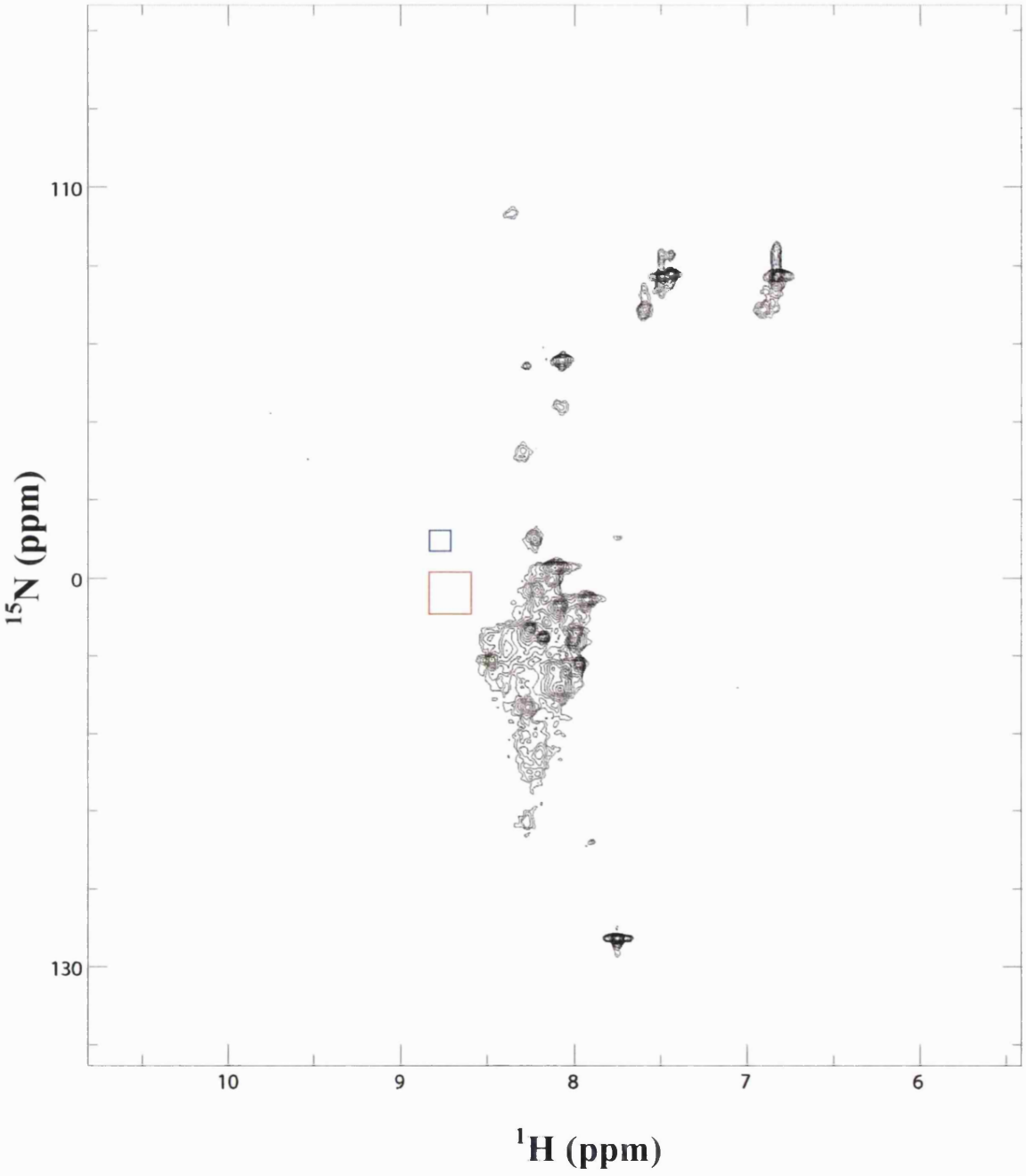


Figure 3.10 2D ^{1}H-^{15}N HSQC spectrum of H-NS 12-64 recorded at 500 MHz. The buffer is 20 mM KPi (pH 7), 300 mM NaCl and 1 mM EDTA, 10% $^{2}H_2O$ and 25°C. The protein concentration is 500 µM (referred to monomer).

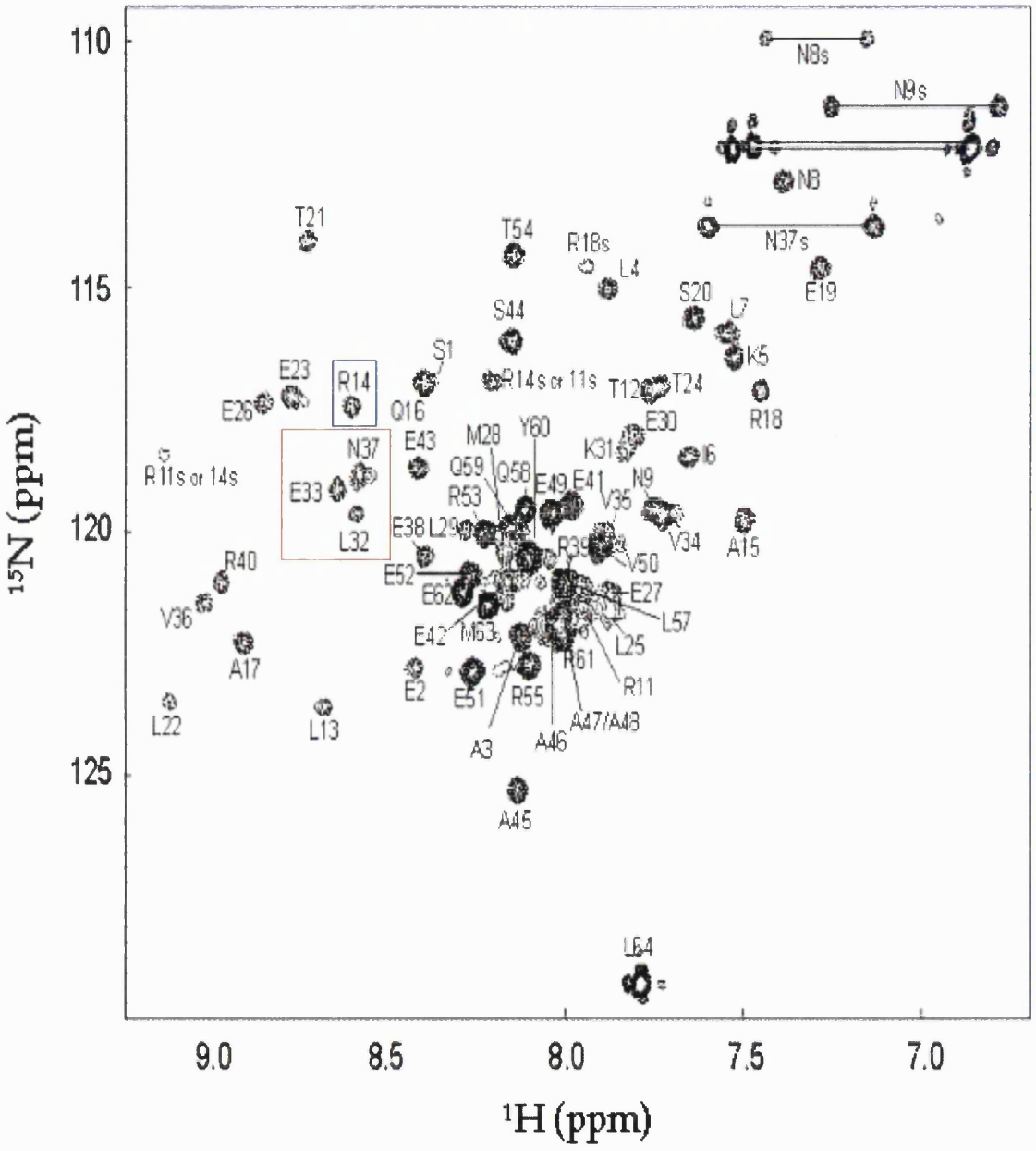


Figure 3.11 2D ^{1}H-^{15}N HSQC spectrum of H-NS 1-64 recorded at 600 MHz. The buffer is 20 mM KPi (pH 7), 300 mM NaCl and 1 mM EDTA, 10% $^{2}H_2O$ and 25°C. The protein concentration is 1.5 mM (referred to monomer).

Dramatic changes in the appearance of the spectrum are apparent upon comparison with that of the H-NS 1-64 wild type. As expected, the cross-peaks corresponding to the region encompassing first 11 amino acid residues are absent. In addition to this, cross-peaks matching the H2 region (e.g. R14 highlighted in blue rectangle) are missing from the spectra. An absence of cross-peaks corresponding to the coiled-coil interface (e.g. L32, E33, N37 highlighted in red rectangle) is evident upon the examination of the spectrum.

The 2D ^{1}H-^{15}N HSQC spectrum of the wild type H-NS 1-64 polypeptide is characterised by the limited resonance dispersion, as can be seen by a substantial cross-peak overlap in the central region of the spectrum. This is a consequence of the entirely α-helical secondary structural elements and the nature of the amino acid sequence, which contains several identical amino acid repeats and the presence of uninterrupted identical amino acids. These characteristic features of the wild type H-NS 1-64 spectrum and the three-dimensional structure of the H-NS 1-57 polypeptide can be used to rationalise the appearance of the H-NS 12-64 spectrum. In the H-NS 1-57 structure, two short N-terminal helices (H1 and H2) fold back on to the H3 helix from the same protomer, which results in the establishment of a contiguous hydrophobic ridge (Fig. 3.12). The isoleucines (residue numbers 6, 7 and 10) and leucine (13) side chains from helices H1 and H2 participate in the formation of a hydrophobic interface along with the leucine residues (22, 25, 29 and 32) of the main chain H3 helix, which line the bottom of this surface. Truncation of the first 11 residues will affect three isoleucine side chains (6,7 and 10) and is likely to compromise the stability of the short H2 helix (10-16), which includes hydrophobic interaction formed by the leucine residue at position 13. It is possible that the N-terminal helices are stabilised by the interaction with the main chain coiled-coil helix through the set of conserved hydrophobic residues.

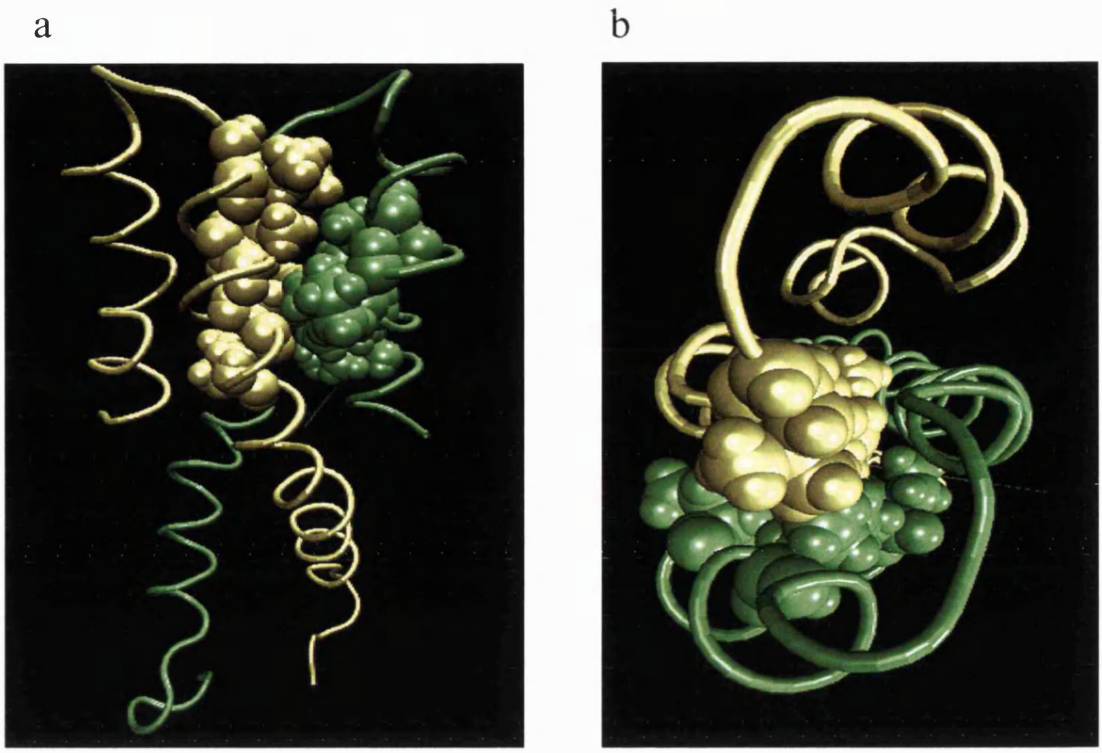


Figure 3.12 Ribbon model of the H-NS 1-57 dimerisation domain. The VMD program (Humphrey *et al.*, 1996) and the H-NS N-terminal domain (residues 1-57) coordinates deposited in the Protein Data Bank (1LR1; Esposito *et al.*, 2002) were used. The two protomers are shown in yellow and green, respectively. Both protomers are displayed as a ribbon structure while the hydrophobic residues are shown with a van der Waals surface representation. The hydrophobic residues from the H1 and H2 (6,7,10 and 13) from one protomer are shown in green. The hydrophobic residues from the H3 (22,25,29 and 32) of the other protomer are shown in yellow. Ribbon diagrams are shown in side view (a) and top view (b).

In support of this, CD spectroscopy of synthetic peptide comprising first 20 residues shows no folded structure (data not shown).

Additionally, inter-molecular electrostatic interactions between arginine side chains (14) and the H3 glutamate residues (23 and 26) will be affected in the H-NS 12-64 polypeptide. The disruption of this interaction, in addition to the interference with a hydrophobic interface, could lead to the destabilisation of the region at the beginning of the coiled-coil dimerisation motif (20-30).

This would lead to even more pronounced resonance overlap than that observed in the H-NS 1-64 wild type spectra. On the basis of CD experiments (Fig. 3.9), it is argued that the 2D ^{1}H-^{15}N HSQC spectrum of the H-NS 12-64 does not represent globally unfolded polypeptide chain. Additionally, size-exclusion chromatography results indicate that the truncated H-NS 12-64 polypeptide is capable of forming dimers. It is likely that the 2D ^{1}H-^{15}N HSQC spectrum of the coiled-coil region would have similar features to that shown in Fig. 3.10.

3.8 Oligomeric properties of the H-NS 17-89 polypeptide

The determination of the three-dimensional structure of the H-NS 1-57 dimerisation domain during the course of this research allowed us a rational design of the truncated versions of the H-NS 1-89 polypeptide. The H-NS 17-89 and H-NS 20-89 constructs have been produced to corroborate the data on the oligomeric behaviour of the H-NS 12-89 polypeptide. Both H-NS 17-89 and 20-89 deletion mutants lack two N-terminal helices H1 and H2 but encompass the whole coiled-coil region.

The Superose 12 gel filtration traces obtained for H-NS 17-89 polypeptide chain are shown in Fig. 3.13. Overlaid on the same graph are the elution profiles for the non-truncated H-NS 1-89 polypeptide. The elution profiles are reminiscent of that of the H-NS 12-89 polypeptide (Fig. 3.6), both in terms of the elution volumes and the symmetrical shape of the peaks. There is a negligible shift of the peak maxima over the concentration range tested and the elution profile is consistent with the notion that the H-NS 17-89 polypeptide exist as a single discrete species with an apparent molecular weight of 26.7 kDa. This value gives an estimate of the oligomerisation state of the H-NS 17-89 as a homotrimer ($Mr/M_{r\,(calc.)}$ ~3). The discrepancy between the estimated molecular weight examined with.

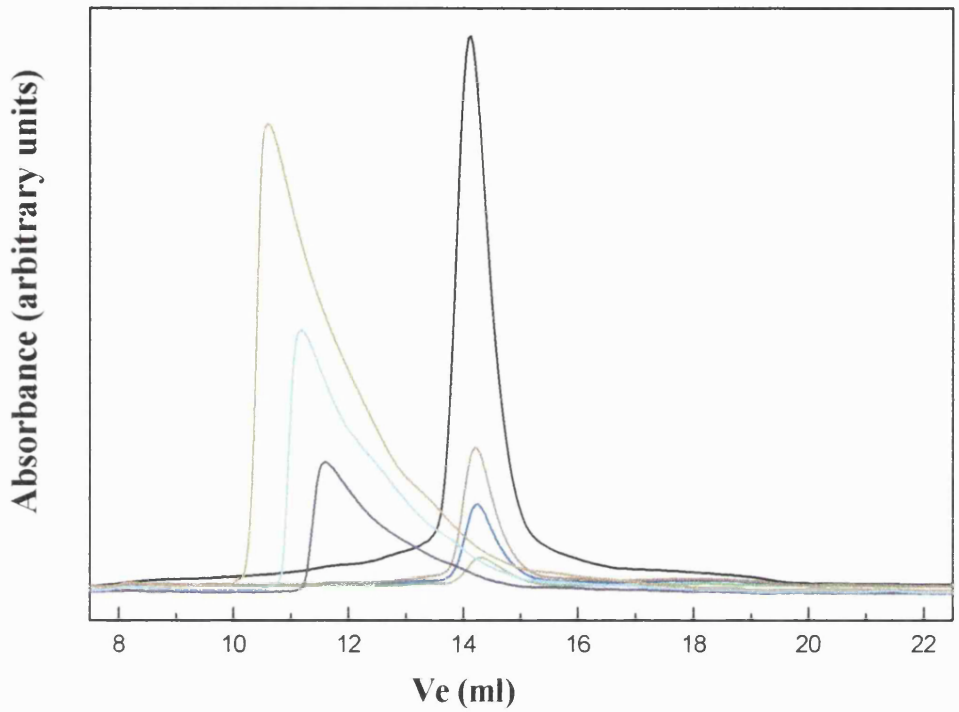


Figure 3.13 Size-exclusion chromatography profile of the H-NS 17-89 polypeptide. H-NS 17-89 at 200 (–), 50 (–), 30 (–) and 4 (–) μM. All experiments were performed in 20 mM KPi (pH 7), 300 mM NaCl and 5%(v/v) glycerol at 25 °C. The elution profile shows absorption at 280 nm as a function of elution volume (V_e). Also shown are the data for H-NS 1-89 polypeptide examined under same conditions at 270 (–), 170 (–) and 120 (–) μM.

two different experimental approaches has already been observed for the H-NS 12-89 polypeptide. Equivalent explanation based on the non-globular shape will be used here to account for the anomalous elution characteristic of the H-NS 17-89 polypeptide. Henceforth, H-NS 17-89 polypeptide is assumed to adopt a dimeric configuration.

This result for the truncated mutant is in contrast to the elution profile of the H-NS 1-89 polypeptide (and H-NS full-length), which exhibits a characteristic asymmetrical shape.

The distinctive nature of the H-NS 17-89 elution profile is clearly exemplified in the overlay graph (Fig. 3.13), as both polypeptides at similar concentrations partition differently on the gel filtration column. The asymmetrical feature of the H-NS 1-89 polypeptide is characteristic of self-association of molecules of equal molecular weight. First described by Winzor and Scheraga (1963), this asymmetry takes the form of a sharply ascending leading edge and a trailing edge extending to the trailing edge position obtained in the case of nonassociation. The position of the trailing edge follows from the fact that in any system at equilibrium, a fraction of the interactive species is in free, unassociated state. It is interesting to note that the trailing edge of the H-NS 1-89 elution profile extends to the position at which H-NS 17-89 polypeptide elutes.

Although the concentration range examined was higher than for the H-NS 12-89 construct, similar problems of low solubility were encountered. The second truncation mutant H-NS 20-89 was not amenable for further study since it exhibited an extremely low solubility at a range of solution conditions.

The combined results from the analytical size-exclusion chromatography and sedimentation equilibrium experiments on the N-terminally truncated H-NS 1-89 polypeptides showed a dramatic change in their oligomeric behaviour. The deletion of one or both of the N-terminal helices (H1 and H2) causes the H-NS 1-89 polypeptide to adopt a dimeric configuration (Fig. 3.6 and Fig. 3.13). Structural and biochemical examination of the H-NS 1-64 polypeptide, which forms dimers, shows that the monomer interfaces are stabilised through a coiled-coil type interaction. This type of interaction is also maintained in the deletion mutants of the H-NS 1-89 polypeptide and is responsible for them forming dimers. Indeed, a truncated version of the H-NS 1-64 polypeptide lacking the first 11 amino acids forms dimers and exhibits features of the α-helical structure, thus excluding

the possibility of the significant destabilising effects of the N-terminal truncations on the dimerisation interface. These results would argue against a model in which the oligomerisation interface is located in the residues 65-89 as the deletion of the N-terminal region (H1 and H2) would have no effect on the oligomerisation. Results obtained by Rimsky and co-workers (Badaut *et al.*, 2002) are in fair agreement with the work presented in this chapter. They have shown that the destabilisation of the coiled-coil interface affects H-NS oligomerisation in solution. More specifically, as shown in this work, removal of the two short N-terminal helices disrupts higher-order oligomer formation without globally affecting coiled-coil stability. Thus, the association of the H-NS dimers occurs through two different regions located at N-terminus (1-20) and the central part of the polypeptide (65-89), respectively.

3.9 Oligomeric and structural properties of the StpA polypeptide

Despite high sequence identity between H-NS and StpA and the ability of StpA to functionally substitute for H-NS in several cellular processes (see Chapter 1), the StpA polypeptide displays unique properties of its own. Here, the oligomerisation properties of StpA were investigated using the techniques previously applied for H-NS.

The Superose 12 elution profile of StpA full-length polypeptide chain at a range of concentrations is shown in Fig. 3.14. The elution behaviour of StpA clearly resembles that of the H-NS full length (and H-NS 1-89). The peaks obtained are broad and asymmetric and there is clear shift in the peak maxima toward higher order species with increasing concentrations signifying self-associative behaviour. The elution behaviour of the StpA polypeptide characterised by a sharp leading edge and an extended trailing edge is expected for rapidly equilibrating solutes (Stevens, 1986, 1989).

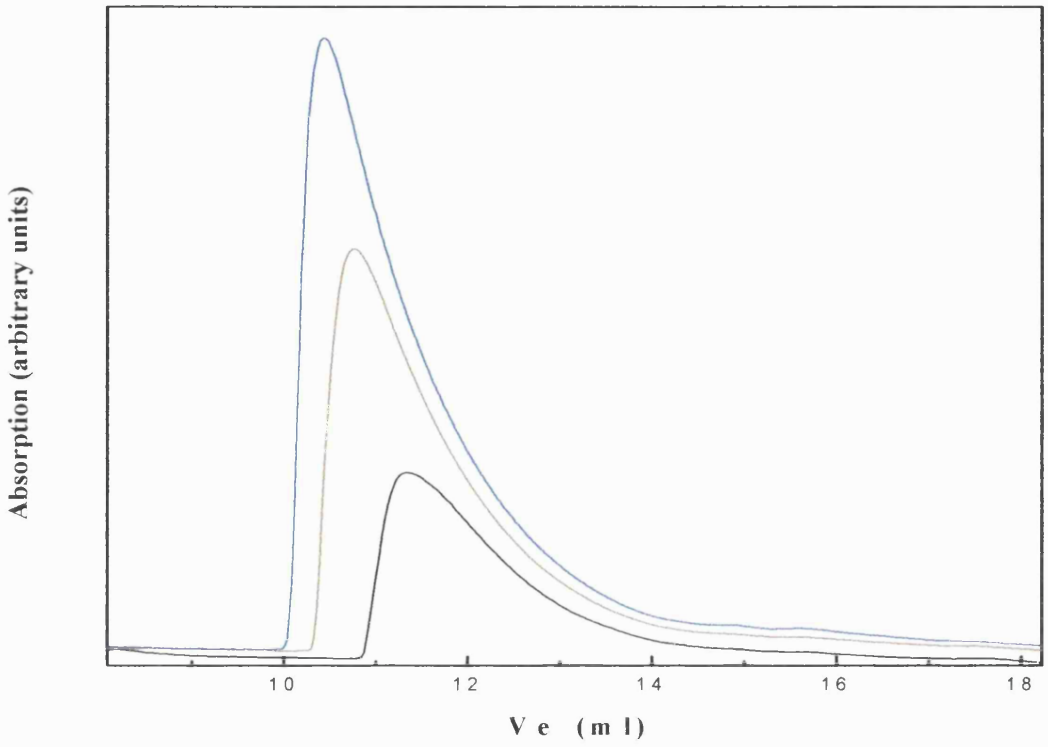


Figure 3.14 Size-exclusion chromatography profile of the StpA polypeptide. StpA at 180 (–), 150 (–) and 80 (–) μM. All experiments were performed in 20 mM KPi (pH 7), 500 mM NaCl and 5% (v/v) glycerol at 25 °C. The elution profile shows absorption at 280 nm as a function of elution volume (V_e).

A comparison with molecular weight standards results in apparent molecular weights between 130 kDa and 88 kDa within the concentration range tested. These values correspond to a change in the oligomeric state from hexamer ($Mr/M_{r\,(calc.)}$ ~5.8) to octamer ($Mr/M_{r\,(calc.)}$ ~8.3) assuming globular shape of the polypeptide.

Preliminary structural studies of the H-NS full-length polypeptide chain by NMR spectroscopy have been hampered by the fact that only a limited number of cross-peaks were observed in the standard 2D ^{1}H-^{15}N HSQC experiment, all of which could be accounted for by the residues in the C-terminal DNA binding domain (Smyth *et al.*, 2000).

Despite the fact that the structural examination of the full-length StpA protein was fraught by its solution properties, in particular, its requirements for a high-salt content, 1D ^{1}H NMR spectra has been recorded (Fig. 3.15 A). A qualitative distinction between folded and unfolded protein is possible with the 1D ^{1}H NMR experiment. The 1D ^{1}H NMR spectrum of full-length StpA exhibited wide chemical shift dispersion characteristic of a fully folded protein notably in the amide (6 ppm-11 ppm) and methyl (below 0.5 ppm) regions. The 1D ^{1}H NMR spectrum gave an alternative experimental indication of folded StpA conformation since the CD spectra proved difficult to interpret due to the considerable noise arising from salt absorption in the far-UV region (data not shown).

A two-domain structural organisation comparable with the H-NS protein has been proposed for the StpA polypeptide (Williams *et al.*, 1996; Cusick and Belfort, 1998; Smyth *et al.*, 2000). Based on both the hydrodynamic properties presented (see Fig. 3.14) and the available biochemical data, it is likely that the 2D ^{1}H-^{15}N HSQC spectrum of the StpA polypeptide would exhibit similar features to that of the H-NS, notably, the limited number of the cross-peaks all of which could be accounted for by the residues present in the C-terminal domain. To aid to this conclusion, the 1D ^{1}H NMR spectra of the full-length StpA and its corresponding C-terminal DNA binding domain (90-133) have been compared (Fig 3.15 B). Indeed, the spectra are almost identical despite the fact that the StpA C-terminal domain represents only one-third of the protein. As with H-NS, it is plausible that the presence of the flexible linker between the N- and C-terminal domains allows unrestricted mobility of the DNA binding region within the context of the StpA oligomers. This would thus allow a spectrum of the C-terminus to be observable.

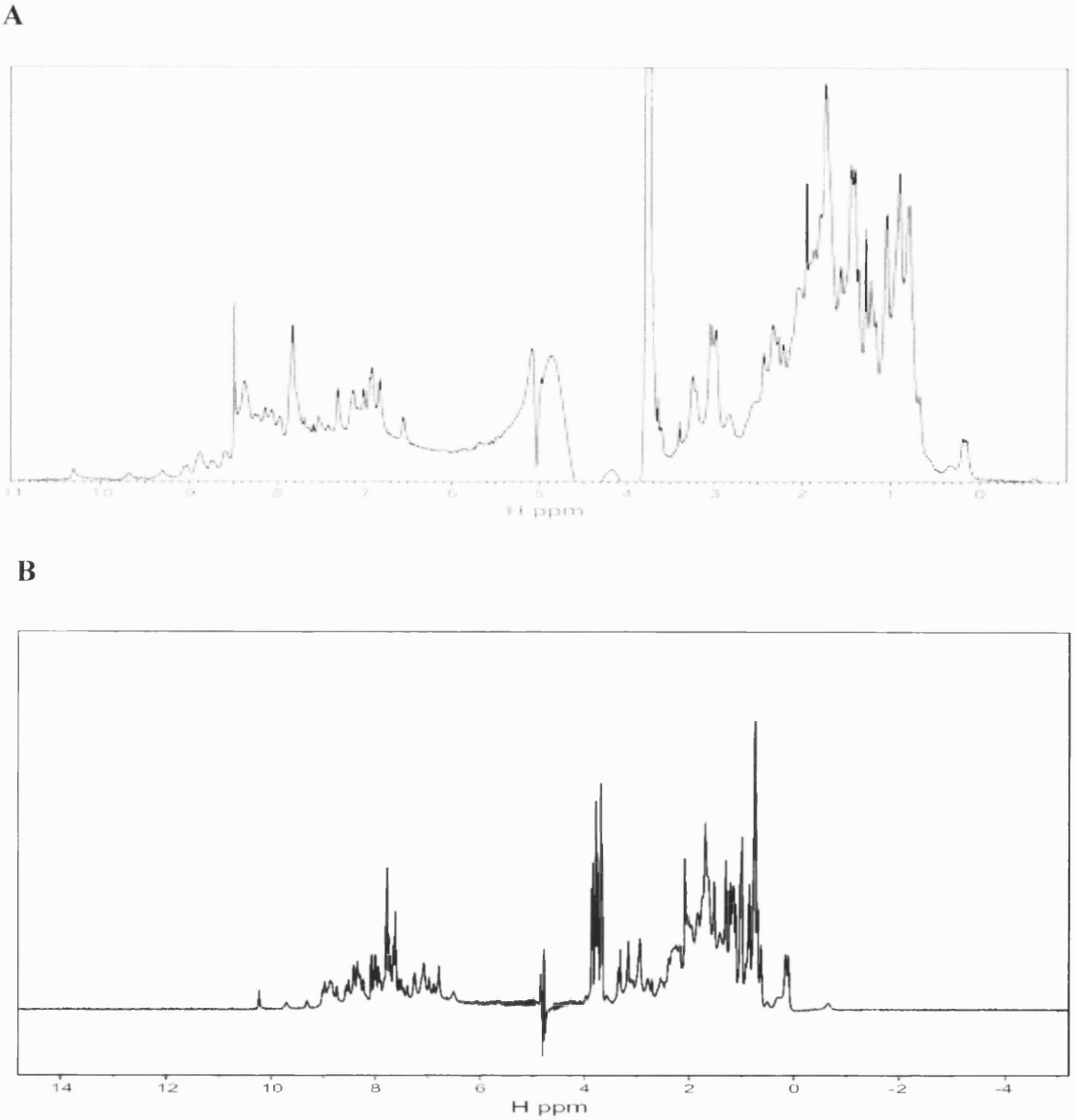


Figure 3.15 A comparison of the one-dimensional (1D) proton spectrum of full-length StpA (A) with its C-terminal domain (residues 90-133) (B). (A) StpA full-length at 100 μM in 20 mM KPi (pH 7), 500 mM NaCl and 5% (v/v) glycerol at 25°C. (B) StpA C-terminal domain at 400 μM in 20 mM KPi (pH 7) and 100 mM NaCl at 25°C.

The oligomeric state of the N-terminus, however, results in a large molecular weight complex, which is invisible to NMR due to extensive line-broadening effects. Together with the observed self-associative properties of the StpA polypeptide, these results further substantiate the notion of the similar structural organisation of the two polypeptides.

3.10 Homology modelling of the N-terminal StpA fragment

In the absence of any structural information on the StpA polypeptide, modelling based on the known three-dimensional structure of a homologous protein provides the only reliable method to obtain structural data. The feasibility of homology modelling relies on the fact that the three-dimensional structures of homologous proteins have been better conserved through evolution compared with their primary amino acid sequences.

The position of residues involved in coiled-coil formation is conserved between N-terminal regions of H-NS and StpA polypeptides (Fig. 1.1). In addition (see Chapter 1) it has been shown that H-NS and StpA proteins form heteromeric complexes *in vitro* and *in vivo*. Homology modelling has been used here to gain structural insight into the N-terminal region (residues 1-57) of StpA polypeptide. The results obtained (Fig. 3.16) substantiate the notion of the extensive structural similarity between the N-terminal regions of the H-NS and StpA polypeptides. The position of the N-terminal helices (H1, H2 and H3) is compatible with the StpA sequence and it is likely that StpA polypeptide dimerises in a manner analogous to H-NS, via the formation of the coiled-coil interface through H3 helix (Fig. 3.16). Furthermore, the H-NS/StpA heterodimer structure is expected to exhibit similar features to the H-NS N-terminal dimerisation domain (Esposito *et al.*, 2002). As shown in other studies, StpA protein is usually expressed at low levels and is rapidly degraded by the Lon protease in the absence of H-NS, which led to the conclusion that StpA is present mainly in complex with H-NS polypeptide *in vivo* (Johansson and Uhlin, 1999). It is not immediately apparent from the homology model of the StpA N-terminal domain as to the structural basis of increased susceptibility towards Lon-mediated proteolysis.

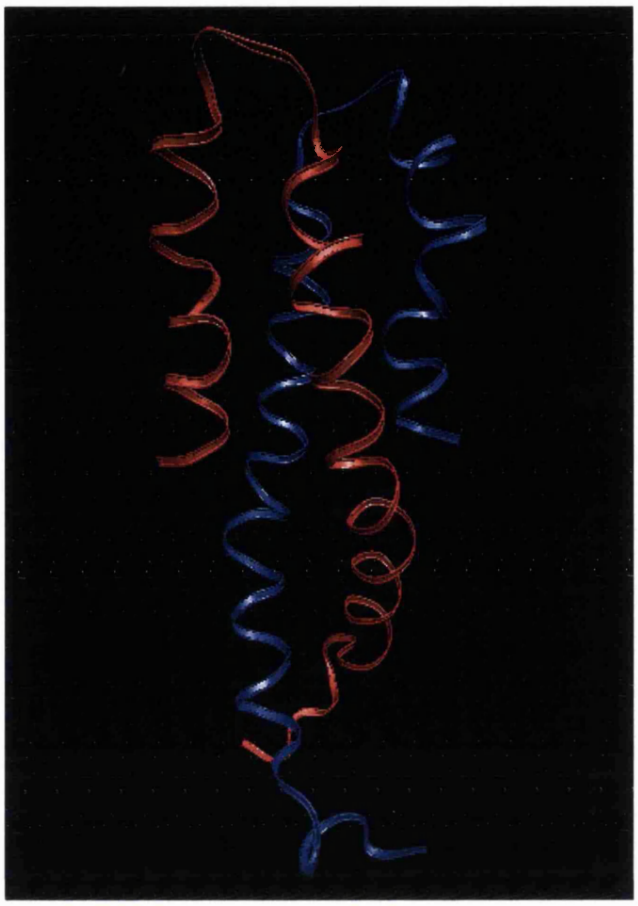


Figure 3.16 Ribbon diagram of the homology model of the StpA N-terminal domain. The H-NS N-terminal domain (residues 1-57) coordinates deposited in the Protein Data Bank (1LR1; Esposito *et al.*, 2002) were used for homology modelling. The two protomers are shown in red and blue, respectively.

Within the region encompassing two short N-terminal helices, substitution of several amino acid residues (e.g. E2V, K5Q, Q16M and C20F) creates a more hydrophobic surface compared with the H-NS N-terminal dimerisation domain. This alteration may favour the interaction between StpA and Lon proteins. This conclusion is supported by the observation that the replacement of a phenylalanine to a cysteine in StpA F20C polypeptide (i.e. introducing polar instead of a hydrophobic residue in the Lon consensus sequence) reduces Lon-dependant proteolysis (Johansson and Uhlin, 1999). The H-NS/StpA complex may weaken its interaction with Lon protease through similar effects.

3.11 Summary

Despite its modular domain organisation (Fig. 1.2), delineation of the structure-function relationships of the H-NS polypeptide has been hampered by the observation that various regions of the amino acid sequence act cooperatively in order to create a platform for diverse H-NS activities including higher-order oligomer formation and DNA binding. For instance, recent studies have suggested that the N-terminal dimerisation domain also contributes to DNA binding in addition to the already established C-terminal region (Bloch *et al.*, 2003).

Little is known in detail of the manifold protein-protein interactions responsible for the formation of higher-order H-NS oligomers. A comparison of the oligomeric properties obtained for the two C-terminally truncated polypeptides (H-NS 1-64 and 1-89, respectively) has suggested that residues 64-89 represent the oligomerisation interface responsible for the formation of higher-order multimers (Smyth *et al.*, 2000). As shown in this chapter and other work (Badaut *et al.*, 2002), residues in the N-terminal region also represent determinants, in part contributing to the formation of higher-order H-NS oligomers in solution. The formation of H-NS oligomers is based on association of the dimeric core unit (dimerisation motif residing in the N-terminal 1-64 region) in a head-to-tail manner (Fig. 3.3).

The *S. typhimurium* H-NS 1-64 polypeptide, as demonstrated by AUC experiments (Fig. 3.2), behaves as a dimer. This is in agreement with work by Rimsky and co-workers (Badaut *et al.*, 2002) in which the *E.coli* H-NS 1-64 fragment was shown to adopt a dimeric configuration.

The formation of the core dimeric fragment may indeed represent a common feature amongst H-NS-like proteins (Arold and Ladbury, personal communication). The interface between two monomers appears to be very stable (K_d<5 nM) as shown by fluorescence anisotropy measurements on the H-NS 1-46 polypeptide (Bloch *et al.*, 2003). The contribution of the conserved hydrophobic residues in the coiled-coil interface is likely to have a major impact on the stability of the N-terminal dimerisation domain. The short N-terminal helix H2 may also affect dimer stability to some extent (Chapter 4).

When observed by size-exclusion chromatography and AUC (Figs. 3.6 and 3.7, respectively), the N-terminally truncated H-NS 12-89 polypeptide clearly exhibits different self-association behaviour to that of the H-NS 1-89 protein (Fig. 3.1). The size-exclusion chromatography experiments (Fig. 3.6), providing a qualitative assessment of the changes, shows a single peak that displays a lack of concentration-dependence phenomena, which is a characteristic for the H-NS 1-89 (and full-length) polypeptides. Sedimentation equilibrium studies demonstrate that H-NS 12-89 polypeptide self-associates to form dimers (Fig. 3.7). These results clearly indicate that the residues at the N-terminal region are also required for the H-NS 1-89 polypeptide to display typical oligomerisation features. The effect of the N-terminal truncations on the stability of the dimerisation domain has been examined within the context of the H-NS 1-64 polypeptide using a number of complementary techniques. The results confirm that the removal of the first 11 amino acid residues had minimal effects on the global stability of the N-terminal dimerisation domain. The examination of both tertiary (as measured by hydrodynamic radius on size-exclusion chromatography) and secondary structure (by CD) (Figs. 3.8 and 3.9, respectively) show that H-NS 12-64 polypeptide forms dimers and is folded, excluding any possible effect of dimer instability on the oligomeric properties of H-NS 12-89 truncation mutant.

It can be concluded that H-NS 12-89 dimers are maintained through the same coiled-coil interface responsible for the formation of the N-terminal H-NS 1-64 fragment (assuming no other dimerisation interface is present outside the 1-64 region). Thus, the truncated H-NS 12-89 polypeptides retain the ability to form coiled-coil interactions but lose their capacity to self-assemble to form high-molecular-weight oligomeric structures. In terms of the possible self-association models shown in figure 3.3, the data presented in this chapter are compatible with the head-to-tail mechanism. In line with the evidence presented, the capability to self-associate is lost in the truncation polypeptide lacking helices H1 and H2, as shown in figure 3.13.

Based on the three-dimensional structure of the H-NS 1-57 polypeptide, one can speculate that the N-terminal truncation, which completely removes H1 and two residues of H2, will expose a number of hydrophobic residues (Fig. 3.12) and may have an effect on the stability of H2. The exposure of hydrophobic residues in the H-NS 12-89 polypeptide could cause non-specific association of the respective dimers and may have been observed in the size-exclusion studies on the H-NS 12-89 truncation mutant in the absence of glycerol, as shown in Figure 3.5. The 2D NMR studies on the H-NS 12-64 truncated mutant show global changes in the appearance of the spectrum compared with the H-NS 1-64 polypeptide (Figs. 3.10 and 3.11, respectively). Although spectra thus obtained suggest a global unfolding of the H-NS 12-64 mutant, these can be argued against in terms of the very poor chemical shift dispersion of the NMR spectra characteristic of the H-NS 1-64 polypeptide, which together with the absence of the first 11 residues and possible structural disruption of H2 helix would lead to an even more pronounced lack of resonance dispersion.

StpA polypeptide exhibits oligomerisation behaviour similar to that of the full-length H-NS and H-NS 1-89 polypeptides (Fig. 3.14): size-exclusion chromatography experiments indicate a concentration-dependant formation of large oligomeric structures. The structure formed by the N-terminal region of the StpA polypeptide (residues 1-57) obtained by homology modelling (Fig. 3.16) is almost identical to that of the H-NS 1-57 fragment. These data support previous genetic and biochemical studies and suggests that the H-NS/StpA heterodimer is likely to occur via the formation of the coiled-coil helix H3.

4. Results II: The effect of N-terminal point mutations on the oligomeric behaviour of H-NS

4.1 Introduction

The C-terminally deleted H-NS polypeptide chain comprising residues 1-89 (the **oligomerisation** domain) has been shown to form high-order oligomers, the size of which is concentration dependent. This shows the equivalent behaviour in terms of oligomeric properties to the full-length H-NS (see Section 3.2). The distinctive asymmetric profile observed in gel filtration experiments results from the rapid dissociation without reassociation experienced by complexes that migrate in advance of the bulk of the solute zone, and from diminishing rates of reassociation with decreasing concentration of reactants in the trailing edge of the solute zone (Stevens, 1989). Such self-associative behaviour has precluded the resolution of distinct H-NS oligomers and has complicated the search for plausible mechanisms responsible for higher-order assembly in isolation and on the DNA template. Studies presented in Chapter 3 have assessed the effect of deletions of the residues at the N-terminus of the H-NS 1-89 polypeptide on its ability to form higher-order oligomers. The findings revealed that the oligomerisaton process could be accounted for by using a model based on the association of dimers of the oligomerisation domain in a head-to-tail manner. This model implies the existence of two interacting surfaces, one located at the N-terminus of the dimeric oligomerisation domain, and the second lying at the C-terminus of this domain. Structural studies on the truncated polypeptide H-NS 1-64 (Renzoni *et al.*, 2000) and H-NS 1-57 (Esposito *et al.*, 2002) revealed the three-dimensional organisation of this dimeric fragment. The N-terminus of H-NS is comprised

of two short helices, H1 (residues 2-7) and H2 (residues 10-16). These helices are arranged in an anti-parallel fashion (see 1.3.3.1) relative to the main helix H3 (residues 20-49) that forms the basis of the dimeric coiled-coil. In this chapter, a site-directed point mutagenesis strategy was employed to assess the relative importance of helices H1 and H2 in the formation of the high-order oligomers and their potential role in stabilising the H3-based dimers. The effects of the potential disruption of H1 and H2 on the oligomeric properties of H-NS 1-89 have been examined.

Within the framework of the head-to-tail model, a disruption of the structural elements at the N-terminus of the oligomerisation domain, which are important for self-association, would prevent high-order oligomer formation. On the other hand, these would have no effect on the self-assembly properties if the interactions required for high-order oligomer formation were mediated through the interface involving residues 65-89 alone. Using high-resolution information of the amino acid residues involved in the formation of H1 and H2, sites have been selected for the substitution of proline residues in the two N-terminal helices in the H-NS 1-89 polypeptide. The insertion of proline residues within the α-helix has been used in this work in an attempt to disrupt the secondary structure of helices H1 and H2. The structure of proline residues and their restricted conformational freedom results in them having a 'helix breaking' effect when they are inserted into α-helical secondary structure elements. Since the 1-89 polypeptide does not provide useful NMR data, the underlying structural effects of the mutants have been examined by the introduction of proline residues at the equivalent positions in the H-NS 1-64 polypeptide fragment. These have been examined using 2D NMR and CD spectroscopy. This allows us to correlate the potential change in the oligomeric properties of the H-NS 1-89 mutants with the specific disruption of H1 or H2.

The restricted torsion angle for the N-Cα-bond of proline allows only a limited number of conformations and imposes stress on α-helical secondary structure. In addition to this, the pyrrolidine ring structure prevents proline engaging in the usual hydrogen bonding observed between NH and CO groups of another amino acids. Due to these properties, proline residues are likely to be observed at the beginning of the helices, at the edges of the β-sheets and most frequently within loops, unordered structures or turn regions of proteins (Creighton, 1993).

4.2 Oligomeric properties of the H-NS 1-89 K5P mutant polypeptide

Detailed structural information of the N-terminal oligomerisation domain allows us to define the structural features, which have been affected in the H-NS 1-89 N-terminal deletion constructs. The H-NS 12-89 polypeptide chain completely lacks H1 and the first two residues of the H2, whilst the H-NS 17-89 construct has both H1 and H2 missing. Comparative analysis of the H-NS 12-64 and 12-89 polypeptides suggests that the N-terminal truncation prevents higher-order oligomer formation without significantly affecting global coiled-coil stability (Chapter 3). On the basis of these considerations the relative importance of the structural integrity of either H1 or H2 on the higher-order oligomer formation is lacking. A more targeted approach involving the specific disruption of H1 and H2 allows us to discern the structural features of importance for self-association. Mutants of H-NS 1-89 were made to initially assess the role of helices H1 and H2 on the oligomerisation. An NMR study on the H-NS 1-64 polypeptide by Renzoni *et al.* (2000) was used to select residues for site-directed mutagenesis. The lysine residue in the middle of the H1 has been replaced by a proline side-chain, as the introduction of this residue at this position was expected to cause the most severe effect on the structure of H1.

The Superose 12 gel filtration traces obtained for the H-NS 1-89 K5P mutant at concentrations between 25 and 225 μM, in 20mM KPi (pH 7), 300 mM NaCl and 5%(v/v) glycerol at 25°C are shown in Figure 4.1. The elution profile of the H-NS 1-89 K5P mutant is reminiscent of the behaviour observed for the H-NS full length and the H-NS 1-89 polypeptides. The peaks obtained are asymmetric at all protein concentrations and there is a clear shift to a higher molecular weight at increased loading concentrations. The H-NS 1-89 K5P mutant is capable of adopting a range of oligomeric states, which cannot be resolved.

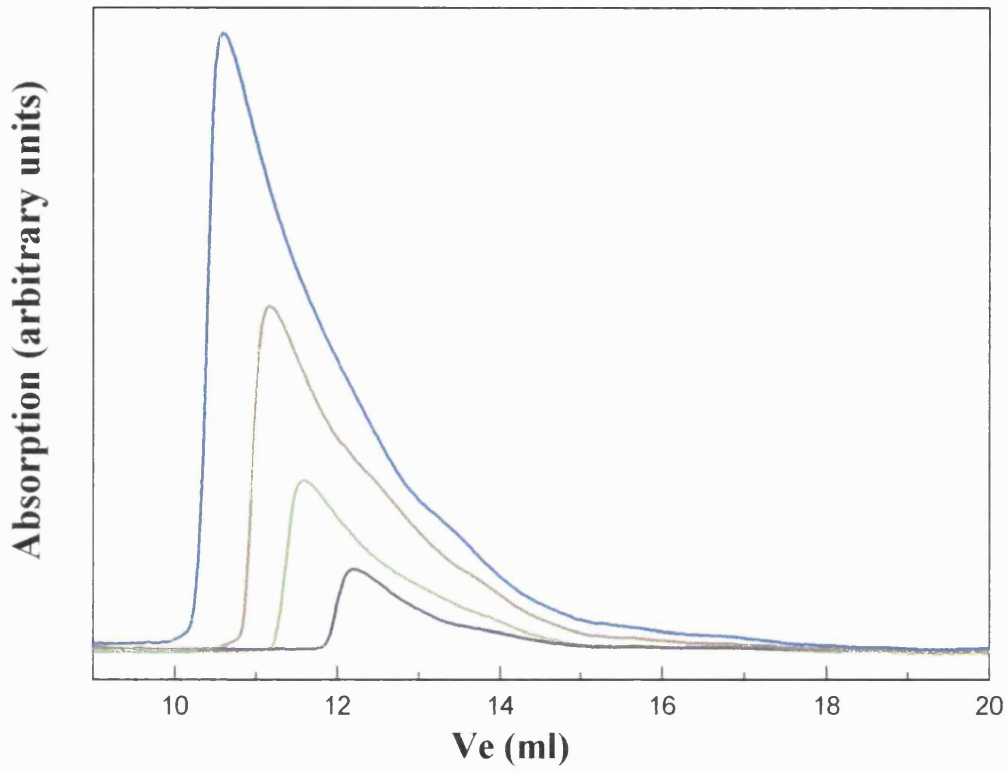


Figure 4.1 Size-exclusion chromatography profile of the H-NS 1-89 K5P polypeptide. H-NS 1-89 K5P at 265 (–), 170 (–), 115 (–) and 80 (–) μM. All experiments were performed in 20 mM KPi (pH 7), 300 mM NaCl and 5% (v/v) glycerol at 25 °C. The elution profile shows absorption at 280 nm as a function of elution volume (V_e).

Upon comparison with molecular weight standards, the apparent molecular weight of the H-NS 1-89 K5P mutant shifts from that corresponding to hexamer to that of a 12-mer within the concentration range studied. It is apparent that the lysine to proline substitution had no effect on the self-association behaviour of the H-NS 1-89 polypeptide. In the absence of any structural information of the effects of proline substitution, a simple interpretation of the results shown in Fig. 4.1 would suggest that the structural integrity of H1 is not important for the formation of the oligomerisation interface.

4.3 Structural characterisation of the N-terminal H-NS 1-64 K5P mutant

Since H-NS 1-89 polypeptide forms higher-order oligomers it is impossible to determine the structural detail using NMR spectroscopy. Therefore identical mutants were made in the H-NS 1-64 polypeptide, which is amenable to NMR studies.

The H-NS 1-89 K5P mutant made to investigate the effects of disruption of the H1 on its oligomeric properties behaves in a manner that is apparently indistinguishable from its wild-type counterpart (Fig. 4.1). Fig. 4.2 displays the 2D ^{1}H-^{15}N HSQC spectrum for a uniformly labelled sample of H-NS 1-64 K5P polypeptide. This experiment gives the backbone amide resonances of all the residues in the protein, with proline as the only exception. Specific residue peaks have been manually assigned upon comparison with the 2D ^{1}H-^{15}N HSQC spectrum for wild-type H-NS 1-64 (Renzoni *et al.*, 2000). Similarly to the H-NS 1-64 wild type (Fig. 3.11), there is a substantial resonance overlap in the spectrum stemming from the entirely α-helical secondary structural elements and the redundant nature of the primary amino acid sequence containing several identical amino acid repeats and the presence of uninterrupted identical amino acids. Several of the cross-peaks corresponding to the H2 can be identified (e.g. R14, A15, Q16). The cross-peaks of most of the residues involved in the dimeric interface are unchanged and can readily be recognised. On the other hand, the cross-peaks matching the residues around the site of the mutation have disappeared. The cross-peaks around serine 20 corresponding to the amino acids in the H1 (L7 and K5) are absent from the spectra. Likewise, residues 4 and 8 are missing from the spectra. The localised nature of missing peaks is consistent with the specific disruption of the region involving H1. The position of the altered residue in the H-NS 1-57 three-dimensional structure is highlighted in Figure 4.3.

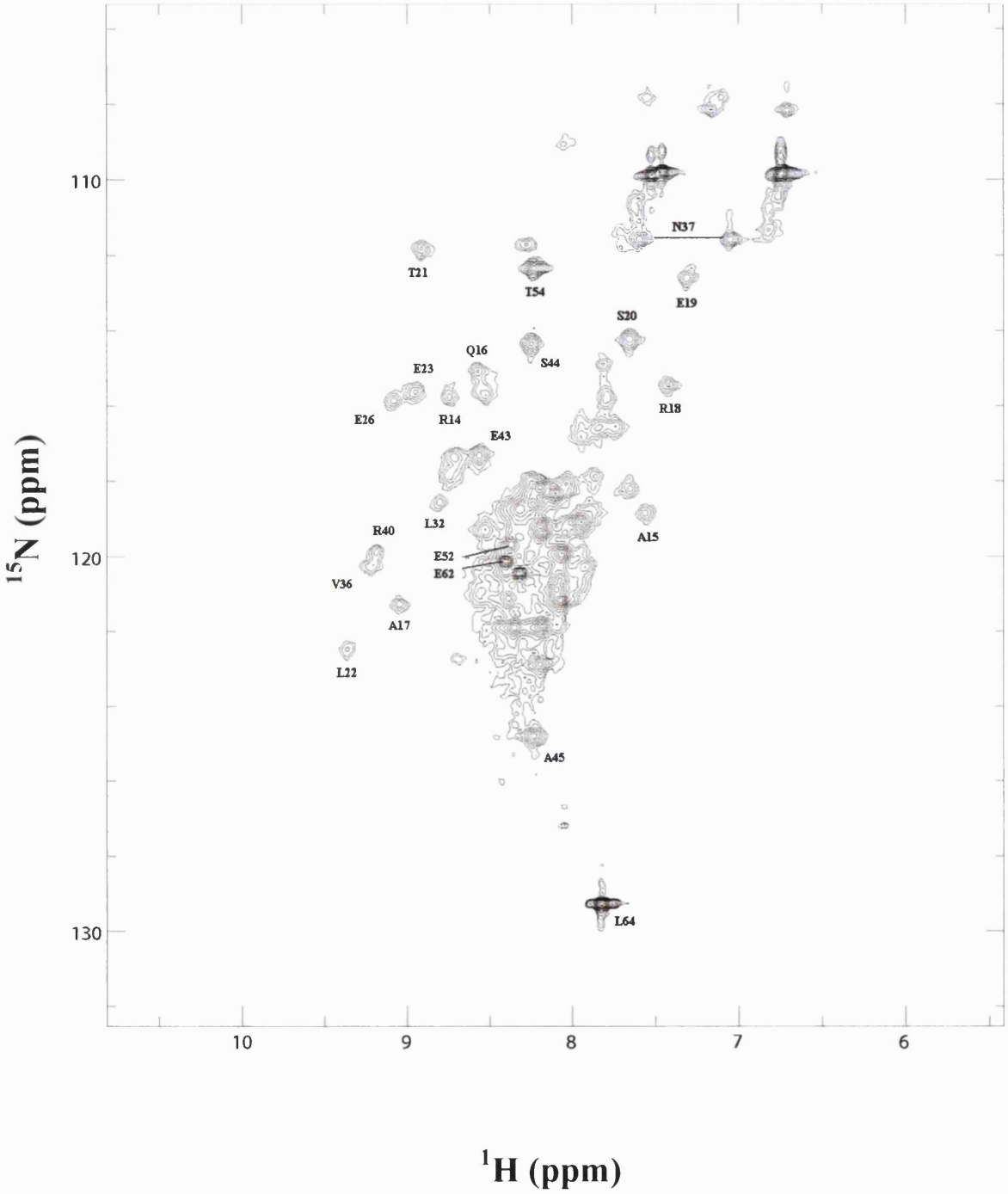


Figure 4.2 2D ^{1}H-^{15}N HSQC spectrum of the H-NS 1-64 K5P recorded at 500 MHz. The buffer is 20 mM KPi, 300 mM NaCl at pH 7.0, 10 % $^{2}H_2O$ and 25°C. The protein concentration is 500 µM (referred to monomer).

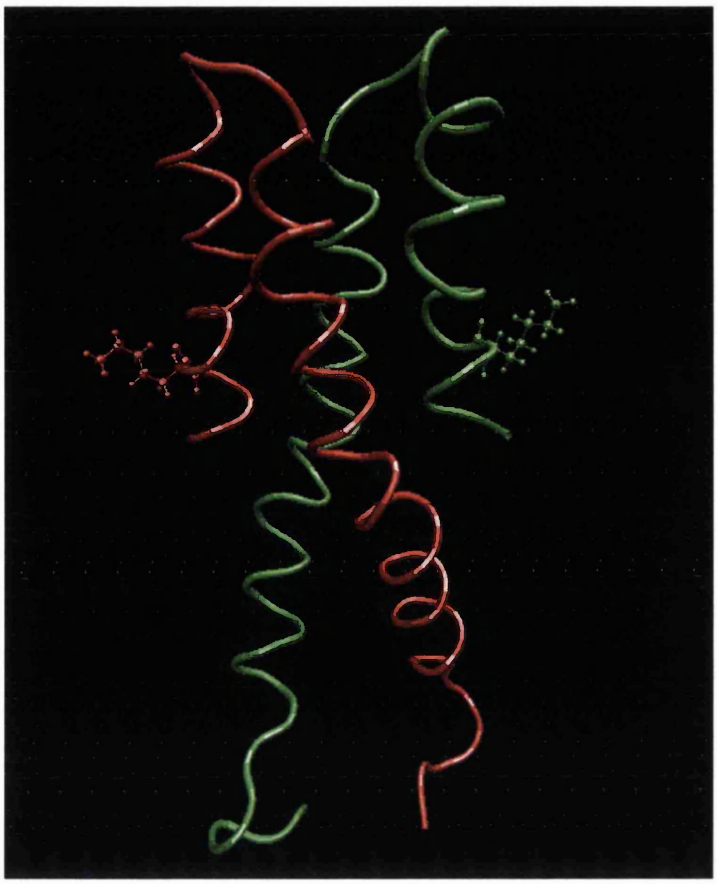


Figure 4.3 Ribbon model of the H-NS 1-57 structure. The VMD program (Humphrey *et al.*, 1996) and the H-NS N-terminal domain (residues 1-57) coordinates deposited in the Protein Data Bank (1LR1; Esposito *et al.*, 2002) were used. The two monomers are shown in red and green, respectively. The side-chain orientations of K5 residues from each protomer are displayed in stick format.

4.4 The effect of the K5P mutation on the on the helical content of H-NS 1-64

The effect of the K5P mutation within the context of the H-NS 1-89 polypeptide shows no discernable manifestation, in terms of its self-association properties compared with its wild-type counterpart. To corroborate NMR data on the H-NS 1-64 K5P mutant, which shows a highly localised disruption of helix H1, CD experiments were performed in order to obtain an independent measure of the overall changes in the α-helical content. Figure 4.4 shows the CD spectrum of H-NS 1-64 K5P mutant in the far-UV region. The shape of the spectra in particular the two negative maxima close to 208 and 222 nm and positive

maximum around 192 signifies the presence of folded structure with a considerable amount of α-helical structure. Comparison with the H-NS 1-64 wild-type CD spectra (Fig. 4.4) clearly shows a lack of any significant structural perturbation brought about by K5P mutation. This result is in agreement with the NMR data. Similarly deconvolution of the H-NS 1-64 K5P CD spectra (see 2.4.4) confirms these findings, since the observed percentage distribution of secondary structures (data not shown) is very similar to that of the H-NS 1-64 wild type. The ellipticity at 222 nm (Table 4.3) obtained for H-NS 1-64 K5P mutant and analysed according to equation (5) (see 4.9) is indicative of the small changes in the amount of α-helical structure compared with the H-NS 1-64 polypeptide.

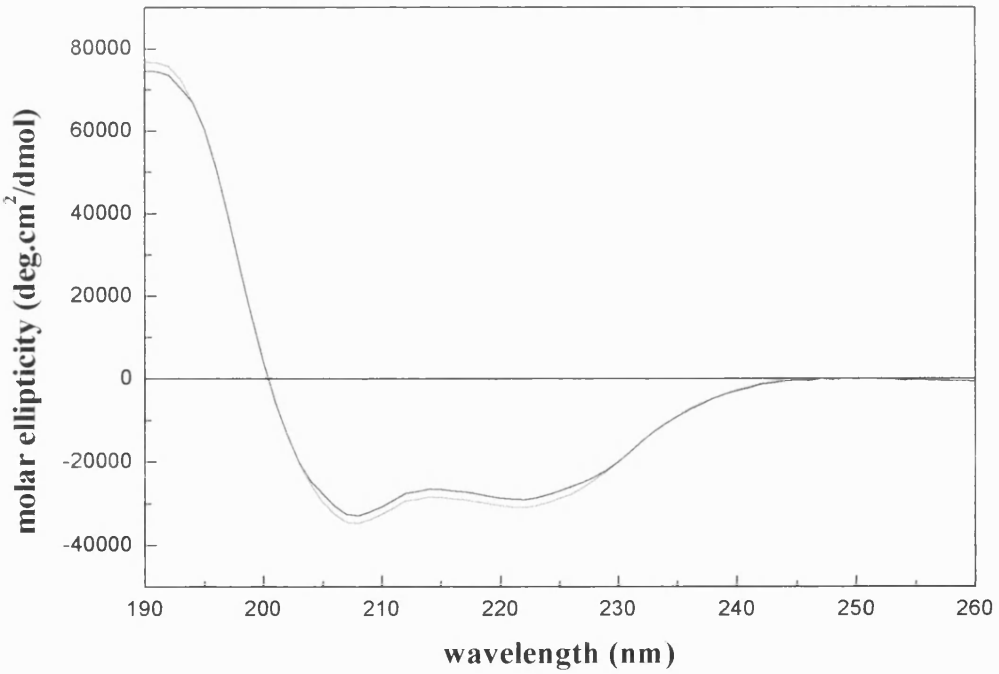


Figure 4.4 Far-UV CD spectra of H-NS 1-64 and 1-64 K5P mutants. CD spectra of H-NS 1-64 (–) at 60 and H-NS 1-64 K5P (–) at 80 μM, respectively, in 20 mM KPi, 10 mM NaCl at pH 7 and 25°C. Results are expressed as mean residue molar ellipticity calculated from equation (4).

4.5 The effect of the K5P mutation on the oligomeric properties of H-NS 1-64

Several lines of evidence suggest subtle effects of the K5P mutation on the conformational properties of H-NS 1-64 polypeptide. Beside residue specific information obtained from the NMR experiments, far-UV CD experiments (Fig. 4.4) show a minor loss of helical structure. Nonetheless, these studies provide no information on the assembly state of H-NS 1-64 K5P mutant. To establish the nature of the oligomeric state adopted by H-NS 1-64 K5P mutant, two complementary techniques have been used.

Size-exclusion studies performed with a Superose 12 fast protein liquid chromatography are shown in Fig. 4.5. Loaded onto this column at initial concentrations ranging from 16 μM to 54 μM, the H-NS 1-64 K5P mutant polypeptide eluted consistently as a single peak with no tailing. This behaviour suggests that under the conditions described, the H-NS 1-64 K5P polypeptide is present in a single well-defined state. H-NS 1-64 K5P elutes at a volume of 14 ml on an analytical Superose 12 column, which corresponds to a molecular mass of 28 kDa. This estimate of the apparent molecular weight leads to values corresponding to an oligomeric state that is between trimer and tetramer ($Mr/M_{r\,(calc.)}$ ~3.5). Indeed, an apparent molecular weight of H-NS 1-64 wild type derived from gel-filtration experiments gives a similar value (Smyth *et al.*, 2000). Series of sedimentation equilibrium experiments (see Section 3.3) led to the re-evaluation of the previous results and firmly established that H-NS 1-64 is a homodimer. The non-globular, elongated shape of the protein can explain this discrepancy. In line with the spectroscopic evidence, the elution behaviour of the H-NS 1-64 K5P polypeptide is consistent with the idea of limited disruption of tertiary and quaternary structure. Sedimentation equilibrium experiments on H-NS 1-64 K5P are presented in Section 4.6).

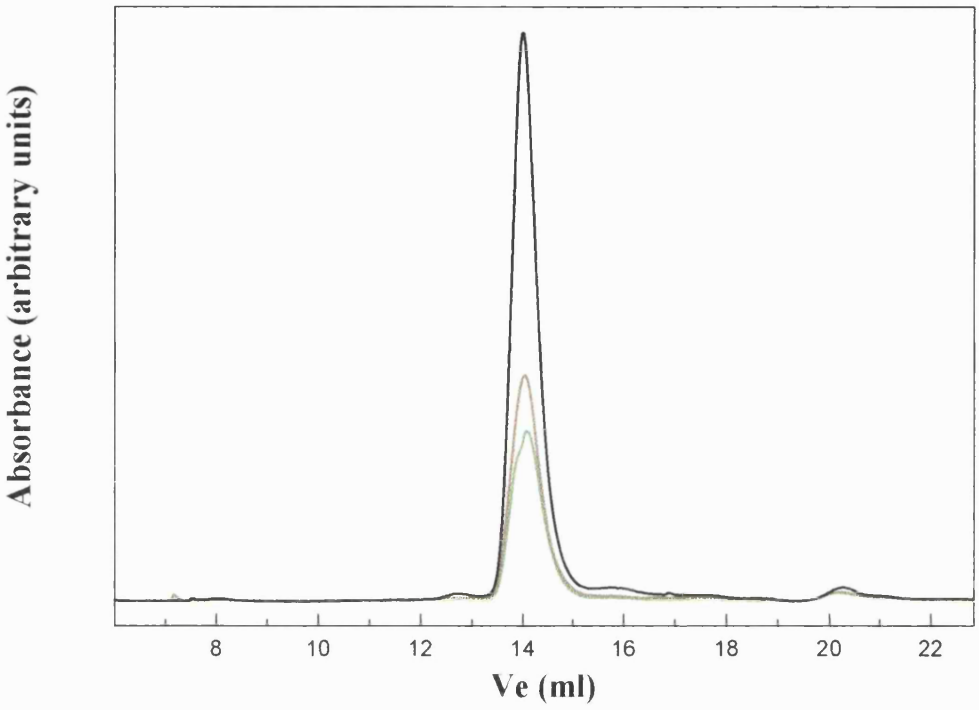


Figure 4.5 Size-exclusion chromatography profile of the H-NS 1-64 K5P polypeptide. H-NS 1-64 K5P at 54 (–), 22 (–) and 16 (–) μM. All experiments were performed in 20 mM KPi (pH 7), 300 mM NaCl and 5% (v/v) glycerol at 25°C. The elution profile shows absorption at 280 nm as a function of elution volume (V_e).

4.6 Sedimentation equilibrium of the H-NS 1-64 K5P polypeptide

Sedimentation equilibrium runs of H-NS 1-64 K5P shown in Fig. 4.6 have been carried out at 20°C at a speed of 42000 rpm. To aid comparison, AUC experiments have been performed using identical solution conditions to size-exclusion chromatography runs namely 20 mM KPi (pH7), 300 mM NaCl and 5% (v/v) glycerol. Samples were loaded at initial concentrations of 160, 300 and 450 μM, respectively and allowed to reach equilibrium at different rotor speeds. Equilibrium distribution of the system was judged by no further change in the overlays of the absorbance profile taken at 2 hourly intervals. Figure 4.6 D, C and B show fits of the data, along with residuals, assuming a trimeric, monomeric and dimeric species, respectively. Of various models, the sedimentation equilibrium data are best accounted for by a single-species dimer (Fig 4.6 B) as judged by

the random distribution of the residuals. Furthermore, the non-random distribution of the residuals of both fits (C and D) clearly indicates that none of these models represent the correct oligomeric state.

The apparent molecular mass of the K5P variant of the H-NS 1-64 polypeptide chain determined from individual fits of the sedimentation equilibrium data (six data sets) is 15.9 kDa, which is close to the value expected for a dimer (15.85 kDa). Furthermore, the mass obtained from a global fit of the six data sets is 14678 kDa, which is approximately 7% lower than expected for a dimer. In addition, the variance, which approaches 1 as the fit improves, of a global fit of the six data sets is lowest for the single-species dimer model (variance=1.07) and much higher variances and non-random distribution of residuals are seen for global fits to other possible models such as single-species trimer, monomer-dimer and monomer-trimer.

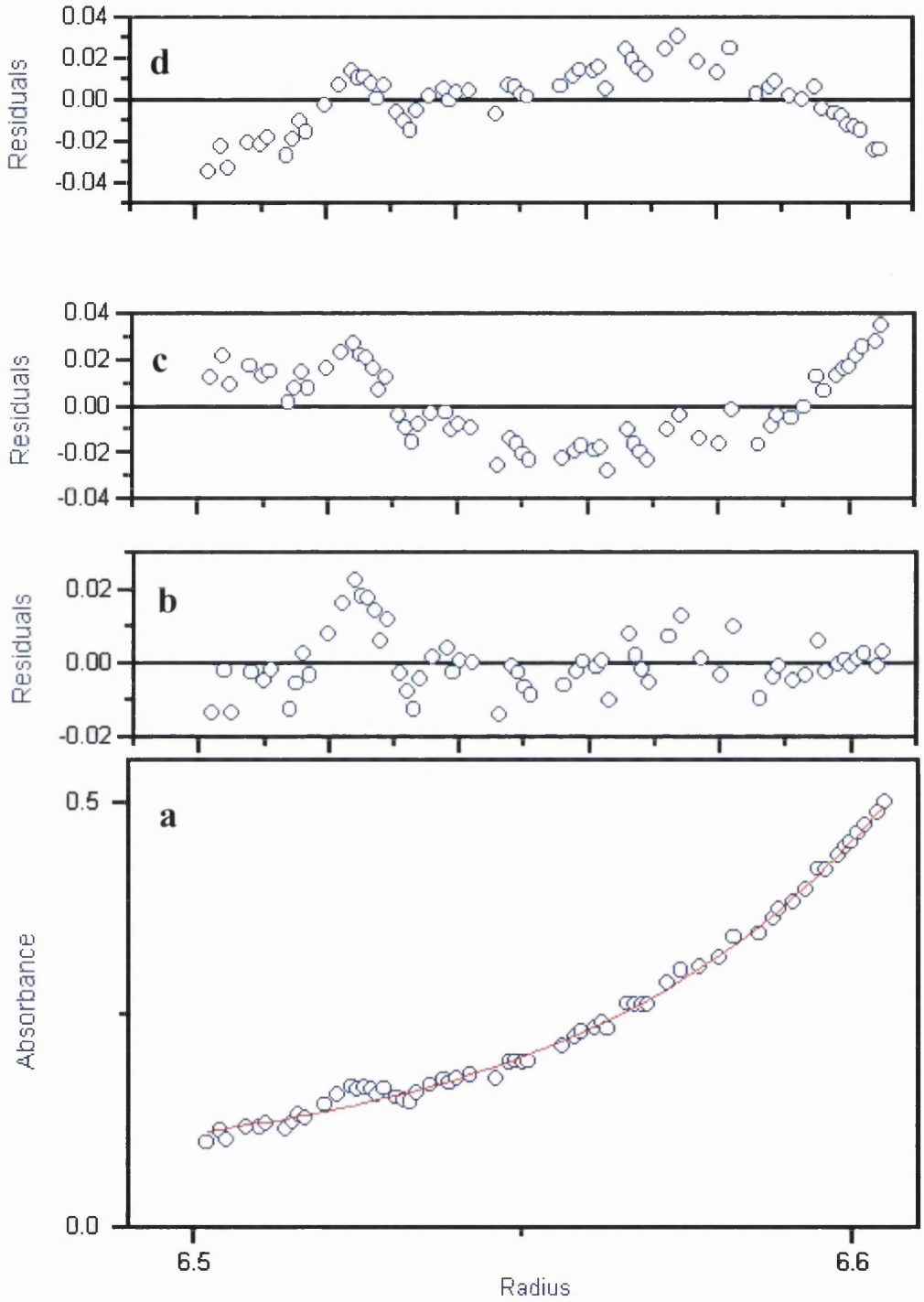


Figure 4.6 Representative sedimentation equilibrium data for the H-NS 1-64 K5P polypeptide. (a) H-NS 1-64 K5P centrifuged to equilibrium at 20°C. Open points were obtained at 42,000 rpm. The continuous line represents the best fit of equation (3), corresponding with the dimer model. (b) Residuals between measured data obtained in (a) and data fitted for a species with the molecular weight of 15908 g/mol. (c,d) Residuals between measured data obtained in (a) and data fitted as monomer (c) and trimer (d). Of various models, the distribution of residuals indicate that the data are best accounted for by a single-species dimer model.

4.7 Oligomeric properties of the H-NS 1-89 Q16P mutant

The H-NS 1-89 Q16P mutant was made to investigate the relative importance of the region encompassing H2 on the self-association properties of H-NS 1-89 polypeptide. It is now known that the coiled-coil residues (E23 and 26) are involved in an inter-molecular salt bridge interaction with the H2 arginine 14 side-chain from each protomer (Esposito *et al.*, 2002). The residue Q16 at the end of H2 was chosen as a target for proline substitution in order to prevent the potential destabilising effects of the salt-bridge disruption on the coiled-coil fold. The H-NS 1-89 Q16P mutant had a greatly reduced solubility compared with the H-NS 1-89 wild-type and H-NS 1-89 K5P polypeptide chains. This resulted in extremely low yields of the purified protein preparation. The Superose 12 gel filtration traces of the H-NS 1-89 Q16P mutant at several different concentrations are shown in Fig. 4.7. Dramatically different behaviour of H-NS 1-89 Q16P mutant was apparent upon comparison with the elution profile of the H-NS 1-89 K5P polypeptide (Fig. 4.1). H-NS 1-89 Q16P elutes in two peaks. The two peaks are not separated, as the absorbance value does not reach the baseline value between them. In contrast to the H-NS 1-89 K5P mutant, there was no shift towards higher molecular weights as the protein concentration is increased, although the range of concentration examined is much smaller due to the above mentioned solubility problems. The retention volume of one of the peaks is ~12 ml, whereas the other peak elutes at ~14.4 ml. A comparison with protein standards gives an apparent molecular weight of approximately 43 kDa for the peak eluting at ~12 ml. The second peak eluting at around 14.4 ml is estimated to give rise to 23.5 kDa species. These apparent molecular weights correspond to a tetramer ($Mr/M_{r\ (calc.)}$= 4.12) and a dimer ($Mr/M_{r\ (calc.)}$= 2.24), respectively.

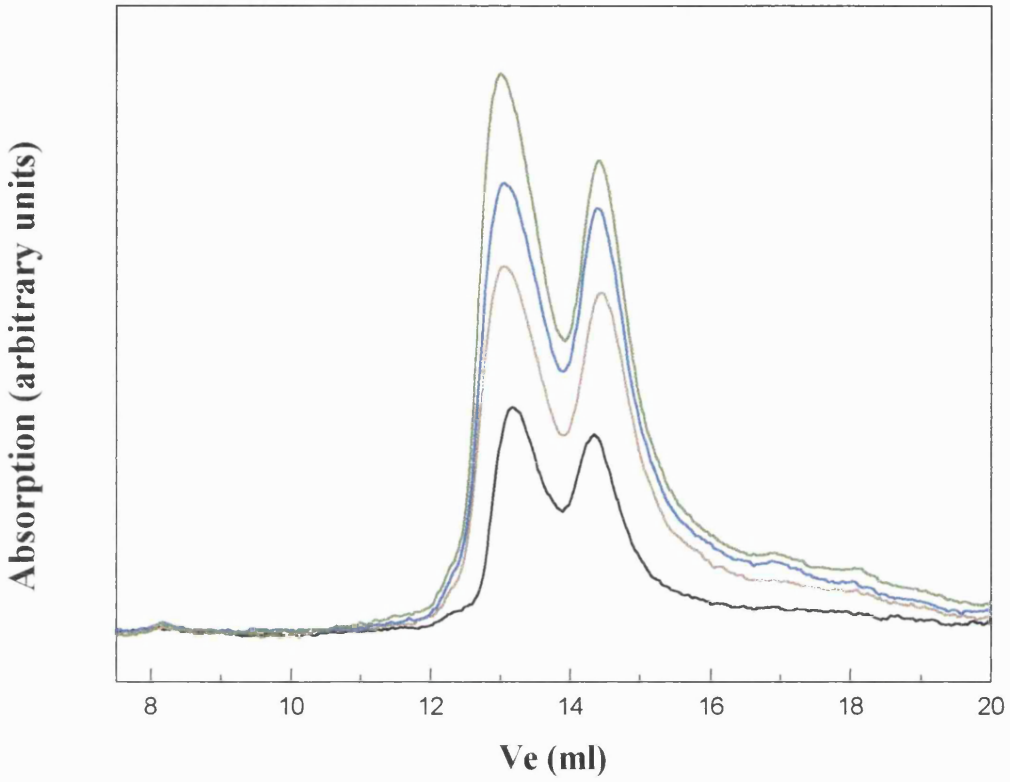


Figure 4.7 Size-exclusion chromatography profile of the H-NS 1-89 Q16P polypeptide. H-NS 1-89 Q16P at 7 (–), 5.7 (–) and 4.5 (–) μM. All experiments were performed in 20 mM KPi (pH 7), 300 mM NaCl and 5% (v/v) glycerol at 25°C. The elution profile shows absorption at 280 nm as a function of elution volume (V_e).

It was not feasible to assess whether the respective dimeric and tetrameric forms were in equilibrium with each other because it was not possible to separate two different species. Furthermore, the poor solubility of H-NS 1-89 Q16P prevented measurements of the relative proportions of the two oligomeric species at different loading concentrations, as these were expected to change according to the law of mass action.

On the other hand, the presence of two distinct peaks could result from the existence of two different conformers and their differential interaction with the gel matrix, although the salt concentration used (300 mM) should minimise this possibility. This explanation is

unlikely, however, if the two different conformational states had drastically different shapes these would separate differently, as the partitioning in gel permeation is a function of molecular shape as well as size.

In the head-to-tail model, residues at both the N and C terminus of the H-NS 1-89 are hypothesised to be involved in establishing higher-order oligomers through the interaction of the dimeric units. The structural integrity of both the dimeric (i.e. coiled-coil) and oligomerisation interfaces are necessary for the establishment of the oligomeric behaviour characteristic of the wild-type H-NS. Since one of the regions of the interface implicated in the oligomer formation lies at the N-terminus (i.e. the region encompassing H1 and H2) and at the same time forms part of the dimeric coiled-coil region (see Fig. 1.3), it is crucial to correlate the structural effects of the mutants on this region to their functional effects in terms of different oligomeric properties when present in the H-NS 1-89 polypeptide.

4.8 Structural characterisation of the N-terminal H-NS 1-64 Q16P mutant

In order to correlate drastically different oligomeric behaviour of the H-NS 1-89 Q16P polypeptide (Fig. 4.7) with the underlying structural changes, a proline residue has been introduced at position 16 in the H-NS 1-64 construct, thus replacing the glutamine side-chain. The 2D ^{1}H-^{15}N HSQC spectrum of the H-NS 1-64 Q16P mutant is shown in Fig. 4.8. Dramatic changes in the appearance of the spectrum are apparent upon comparison with the spectra obtained for both H-NS 1-64 and H-NS 1-64 K5P polypeptides. Unlike the very localised and subtle changes observed in the H-NS 1-64 K5P mutant, the 2D ^{1}H-^{15}N HSQC spectrum of the H-NS 1-64 Q16P polypeptide resembles that of the H-NS 12-64 (Fig. 3.10).

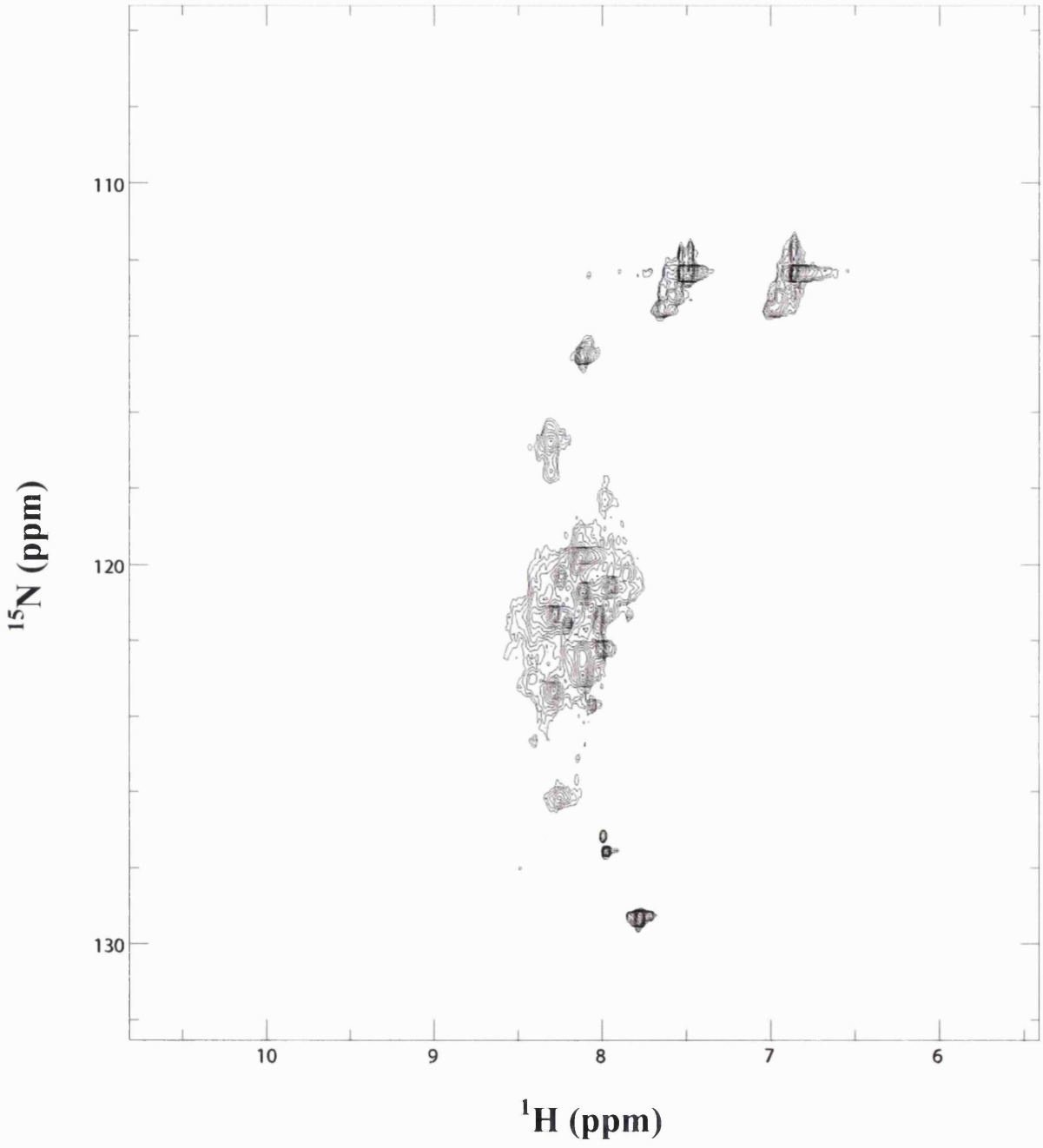


Figure 4.8 2D ^{1}H-^{15}N HSQC spectrum of H-NS 1-64 Q16P recorded at 500 MHz. The buffer is 20 mM KPi (pH 7), 300 mM NaCl, 10% $^{2}H_2O$ and 25°C. The protein concentration is 400 µM (referred to monomer).

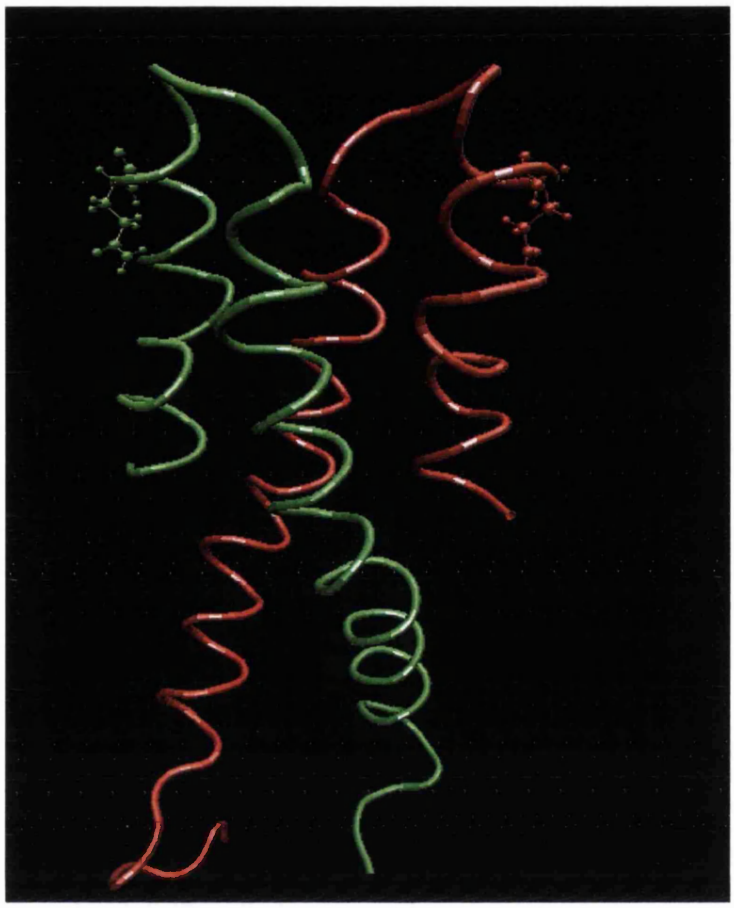


Figure 4.9 Ribbon model of the H-NS 1-57 structure. The VMD program (Humphrey *et al.*, 1996) and the H-NS N-terminal domain (residues 1-57) coordinates deposited in the Protein Data Bank (1LR1; Esposito *et al.*, 2002) were used. The two monomers are shown in red and green, respectively. The side-chain orientations of Q16 residues from each protomer are displayed in stick format.

The cross-peaks corresponding to residues that encompass regions comprising H1 and H2 are absent. Furthermore, many of the cross-peaks belonging to H3 are not present. The three-dimensional structure of the H-NS 1-57 polypeptide allows us to rationalise some of the findings relating to the effects of the Q16P mutation on the conformational properties of the H-NS 1-64 polypeptide. Both H1 and H2 α-helices at the N-terminus of the protein fold back onto the coiled-coil core (Fig. 4.9) burying a large hydrophobic patch, which would be otherwise left exposed.

The disruption of such arrangement is likely to have a destabilising effect on the dimeric interface. Furthermore, it is possible that the inter-molecular salt bridge interaction between arginine at position 14 and glutamines 23 and 26 will also be compromised. Based on the 2D NMR results presented in Fig. 4.8, it appears that the single amino acid substitution (Q16P) had equivalent effects to the removal of first 11 residues observed in the H-NS 12-64 deletion mutant, in terms of the structural disruption. Although the H-NS 12-64 polypeptide had a 2D ^{1}H-^{15}N HSQC spectrum that suggested an unfolded structure, size-exclusion and CD studies (Fig. 3.8 and 3.9) ruled out global destabilisation (i.e. unfolding) as a result of the N-terminal truncation. In line with the partial disruption of the N-terminal region of H-NS 1-64 Q16P polypeptide is the CD data presented in section 4.9.

A combination of site-directed mutagenesis and structural work suggests that H2 appears to be necessary for maintaining the integrity of both H1 and part of the coiled-coil interface. It is likely that the mutual stabilisation of H1 and H2 and the long coiled-coil helix is maintained through their hydrophobic interface. The exposure of hydrophobic patch as a result of disruption of H1 and H2 may account for the low solubility of the H-NS Q16P and H-NS 12-89 and 17-89 polypeptides. The H-NS 1-89 K5P Q16P double mutant, although expressing at normal levels, was not amenable for further study due to an extremely low solubility.

When present in the H-NS 1-89 polypeptide, the Q16P mutation has a serious effect on self-association properties, leading to the loss of the characteristic higher-order oligomer formation. The importance of H2 is further emphasised by the inability of H-NS 12-89 and H-NS 17-89 deletion mutants to self-associate in a manner characteristic of the H-NS 1-89 and H-NS full-length polypeptide chains.

The disruption of H1, on the other hand, has no effects on the oligomeric properties of H-NS 1-89 polypeptide. The behaviour of the H-NS 1-89 Q16P polypeptide on the size-exclusion chromatography requires further explanation. In contrast to the H-NS 12-89 and 17-89 polypeptides, which, as shown in figures 3.6 and 3.13, elute as a single peak, the H-NS 1-89 Q16P deletion mutant partitions in terms of two species on a Superose 12 column. This behaviour is observed in spite of the presence of similar structural disruption in all three polypeptides. Although difficult to explain, it is possible that the different hydrodynamic properties of the H-NS 1-89 Q16P polypeptide arise as a consequence of a longer unfolded N-terminal region compared with the H-NS 12-89 and 17-89 deletion mutants, which would expose a number of hydrophobic residues (isoleucine 6, 7, 10; leucine 13) in contrast to 12-89 (only leucine 13) and 17-89 (no residues) polypeptides. As already used to explain chromatographic behaviour of the H-NS 12-89 polypeptide in the absence of glycerol, the elution properties of the H-NS 1-89 Q16P polypeptide could arise from the non-specific interactions between the respective H-NS 1-89 Q16P dimers. The presence of the 5% glycerol in the running buffer used for size-exclusion studies of the H-NS 1-89 Q16P deletion mutant did not lead to the presence of a single peak in a manner observed for the H-NS 12-89 polypeptide. This may be due to the larger number of exposed hydrophobic residues. It the interesting to note that the solubility properties of the H-NS 1-89 truncation and single-point mutants parallel the number of exposed hydrophobic chains.

4.9 The effect of the Q16P mutation on the helical content of the H-NS 1-64 polypeptide

Conformational changes in the polypeptide due to solvent perturbation, melting or variations brought about by mutation(s) are easily followed by CD spectroscopy. Typical protein secondary structures have different CD spectra and characteristics of the α-helix are a negative band at 222 nm and a negative and positive peak at about 208 nm and 190 nm, respectively. The CD spectrum of the H-NS 1-64 Q5P mutant in the far-UV region (190-260nm) is shown in Figure 4.10. The spectrum has been overlaid with that of the H-NS 1-64 wild type and both spectra have been normalised, which allows a direct comparison of the changes brought about by the mutation. The H-NS 1-64 Q16P mutant exhibits clearly different spectroscopic properties.

It is apparent that there is a dramatic decrease in the overall helical content (i.e band peaks have significantly lower magnitudes) and there is slight shift of the spectrum toward lower wavelengths (Table 4.1). This indicates changes in the structural amount of H-NS 1-64 Q16P due to the substitution of glutamine 16 with a proline residue. This spectrum differs from the characteristic spectra of unstructured peptides and proteins, which are dominated by a large negative maximum at about 200 nm. The magnitude of ellipticity at 222 nm is subject to the largest changes caused by the perturbations in the α-helical content. This result is consistent with the NMR studies (see Fig. 4.8), which show structural changes. The CD data further show that the H-NS 1-64 Q16P mutant is not unfolded but instead has undergone a substantial structural disruption.

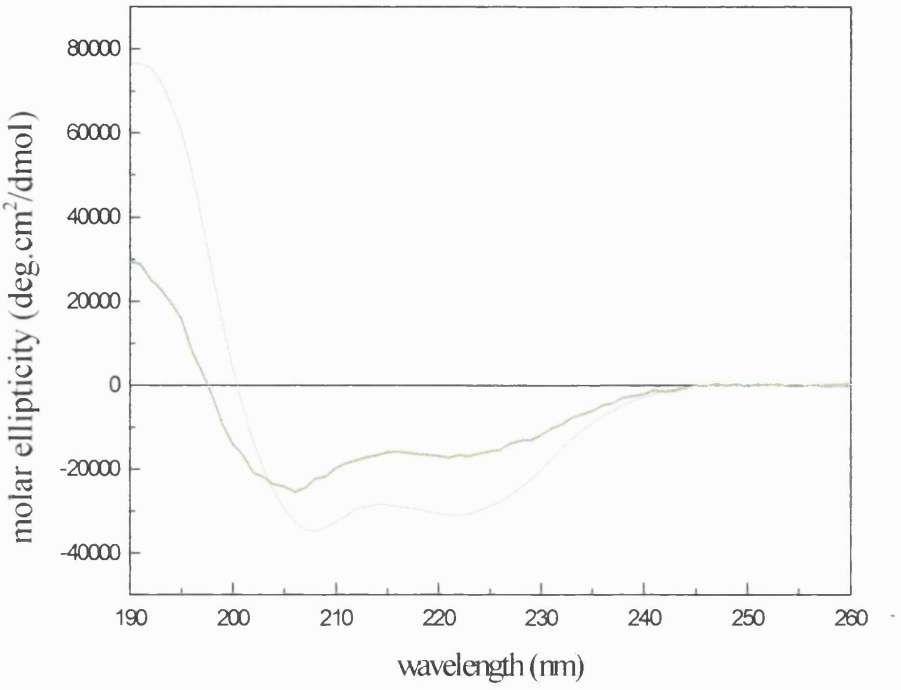


Figure 4.10 Far-UV CD spectra of H-NS 1-64 and H-NS Q16P mutants. CD spectra of H-NS 1-64 (–) at 60 and H-NS 1-64 Q16P (–) at 50 μM, respectively in 20 mM KPi, 10 mM NaCl at pH 7 and 25°C. The ellipticity values have been normalised.

Analysis of the CD data in terms of secondary structure content (spectra deconvolution) provides another means to assess the changes brought about by mutations in terms of the distribution of secondary structure elements. For the detailed empirical analysis of percentages of α-helix, random coil and other types of secondary structures from CD spectroscopic data, DICHROWEB (see section 2.4.4) was used.

Several approaches employing a number of the most popular secondary structure determination algorithms corroborate the explanation inferred from both the NMR and CD experiments (Table 4.2). Another independent measurement of the α-helical content of the respective H-NS 1-64 mutant polypeptide spectra was calculated according to the equation (5) (Forood *et al.*, 1993) (see Section 2.4.4):

Protein	max/θ	min I/θ	min II/θ
H-NS 1-64	190/76578	208/-34824	222/-31815
H-NS 1-64 Q16P	190/29500	206/-25562	221/-17304
H-NS 1-64 K5P	190/75548	208/-32855	222/-29045

Table 4.1 Properties of the CD spectra of H-NS 1-64, 1-64 Q16P and 1-64 K5P polypeptides. The table summarizes the maxima (max) and minima (min I and min II) of the CD spectra and the respective molar ellipticities (θ) at these wavelengths.

	Secondary structure content (%)*		
Protein	α	rc	others
H-NS 1-64 Q16P**	0.216	0.661	0.122
H-NS 1-64 **	0.611	0.311	0.08
H-NS 1-64 Q16P***	0.6	0.32	0.07
H-NS 1-64 ***	1	0	0

* analysis of protein secondary structures from CD spectra accessible via a server located at http://www.cryst.bbk.ac.uk/cdweb

** results from CONTIN algorithm

*** results from K2D algorithm

Table 4.2 Secondary structure composition analysis of H-NS 1-64 and 1-64 Q16P polypeptides. The data were obtained from the spectra shown in Fig. 4.10. The data were analysed in terms of the fractions of residues in α-helical (α), random-coil (rc) and others (β-sheet and turns).

Protein	$-(\theta)_{222}$(deg·cm^2·dmol^{-1})	Helicity, %*
H-NS 1-64	-31015	80.5
H-NS 1-64 Q16P	-16735	43.4
H-NS 1-64 K5P	-29045	75

* Percentage helicity has been calculated as $100x([\theta]_{222}/^{max}[\theta]_{222})$.

Table 4.3 α-helical content of H-NS 1-64, H-NS 1-64 K5P and H-NS 1-64 Q16P polypeptides calculated according to equation (5). The values of θ_{222} represents the value of mean residue molar ellipticity at 222 nm.

$$^{max}\text{Content } (\alpha\text{-helix}) = \theta_{MRE222} / -40.000 \cdot [1-(2.5/n)] \quad (5)$$

100% helicity was estimated by the above formula. According to this method, the mean residue ellipticity at 222 nm for a 68-residue polypeptide (i.e. H-NS 1-64 plus 4 amino acids after thrombin cleavage) chain in a completely helical conformation is –38529 deg·cm^2·dmol^{-1}. Table 4.3 shows the α-helical content of the H-NS 1-64 wild type and the two corresponding mutants calculated according to the above formula.

4.10 The effect of the Q16P mutation on the oligomeric properties of the H-NS 1-64 polypeptide

In order to examine whether the Q16P mutation had any effect on the oligomeric state of H-NS 1-64 Q16P, samples were loaded onto the Superose 12 column spanning several different concentrations. Elution profiles of H-NS 1-64 Q16P polypeptide are shown in Fig. 4.11. The H-NS 1-64 Q16P elution profile shows a slight concentration-dependence of the peak position. At the lowest concentration (16.5 μM), an apparent molecular weight of 23.3 kDa was estimated. The highest concentration tested (100 μM) gives rise to an apparent molecular weight of 26.77 kDa.

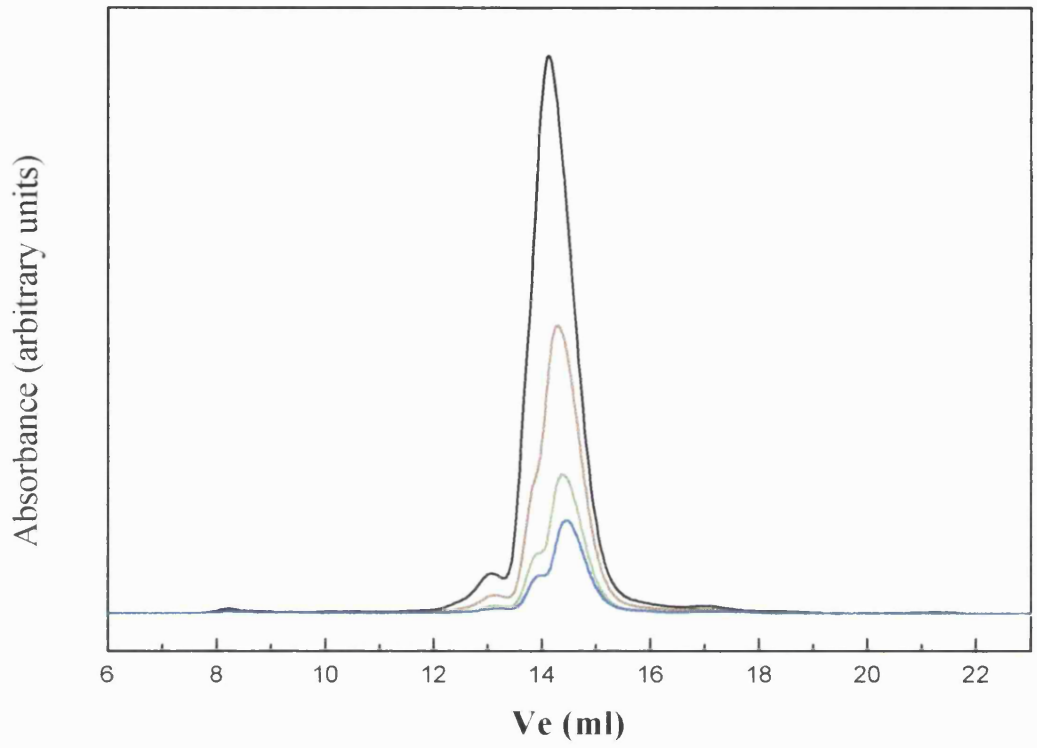


Figure 4.11 Size-exclusion chromatography profile of the H-NS 1-64 Q16P polypeptide. H-NS 1-64 Q16P at 100 (–), 51 (–), 25 (–) and 16.5 (–) μM. All experiments were performed in 20 mM KPi (pH 7), 300 mM NaCl and 5% (v/v) glycerol at 25°C. The elution profile shows absorption at 280 nm as a function of elution volume (V_e).

These values correspond to change in oligomeric state from somewhat smaller than a trimer ($Mr/M_{r\ (calc.)}$=2.91) to that of slightly higher than a trimer ($Mr/M_{r\ (calc.)}$=3.35) based on comparison with globular protein standards. These results indicate that the change in the apparent molecular weight within the concentration range tested is smaller than a theoretical monomer molecular weight. Nevertheless, a shift could be observed from an elution volume of 14.1 ml to 14.45 ml. The elution behaviour of the H-NS 1-64 Q16P mutant is clearly different from H-NS 12-64 and 1-64 K5P polypeptides which partition as a single peak on the Superose 12 column.

Additionally, a shoulder peak is apparent at three lowest concentrations tested. The presence of the second shoulder peak (most pronounced at the highest concentration) is also evident. It thus seems that the H-NS 1-64 Q16P polypeptide exhibits concentration-dependent aggregation phenomena. The exposure of the unfolded N-terminal region revealing a number of side-chains could cause this association through hydrophobic interactions. The structural disruption in the H-NS 12-64 and 1-64 K5P polypeptides may not be sufficient for such behaviour to be observed

4.11 Sedimentation equilibrium of the H-NS 1-64 Q16 polypeptide

Equilibrium sedimentation experiments were performed with the H-NS 1-64 Q16P mutant polypeptide chain at a range of concentration and rotor speeds at 20°C (Fig. 4.12). The sample was not amenable for proper data analysis since the protein exhibited both time- and concentration-dependant aggregation in the AUC cell.

In Figure 4.12, an absorbance of the three H-NS 1-64Q16P samples (a, b and c) each at different loading concentrations is plotted versus radial distance. Differently coloured curves within each cell denote equilibrium distribution of H-NS 1-64 Q16P at four different speeds (see figure legend). It is apparent from the examination of the AUC results presented in Fig. 4.12 that, apart from the data collected at the lowest speed (20 K) and lowest concentration in which the distribution of the sample from the meniscus to the bottom of the cell is even, absorption profiles at higher speeds and concentrations show an abrupt increase in absorption toward the bottom of the cell (points b and c) due to aggregation. As a consequence, whole data points could not be collected for these sets.

In spite of the H-NS 1-64 Q16P aggregation, data sets collected at different speeds and lowest concentration have been analysed as these provided accurate data across the whole cell. Diagnostic graphs provide a qualitative characterisation of the solution behaviour. The apparent H-NS 1-64 Q16P molecular weights extracted from the sample at the lowest concentration (**a**) as a function of the rotor speed have been plotted in Fig. 4.13. There is a pronounced and systematic decrease in the apparent molecular weight with increased rotor speed. The apparent molecular mass of H-NS 1-64 Q16P ranges from 27 kDa at 20000 rpm to 10700 kDa at 42000 rpm for a protein concentration of 156 μM, which, in turn, corresponds to change in stoichiometry from 3.3 to 1.3.

This type of behaviour is diagnostic for sample heterogeneity (Laue, 1995). This has been clearly distinguished from other conditions which may be observed such as: (1) increase in molecular weight with increasing concentration; (2) lack of change in molecular weight with increasing concentration and (3) absence of change in molecular weight with increase in rotor speed all of which are indicative of other types of behaviour. The behaviour of the H-NS 1-64 Q16P polypeptide observed by AUC does not resemble that of either H-NS 1-64 wild type or the K5P polypeptide.

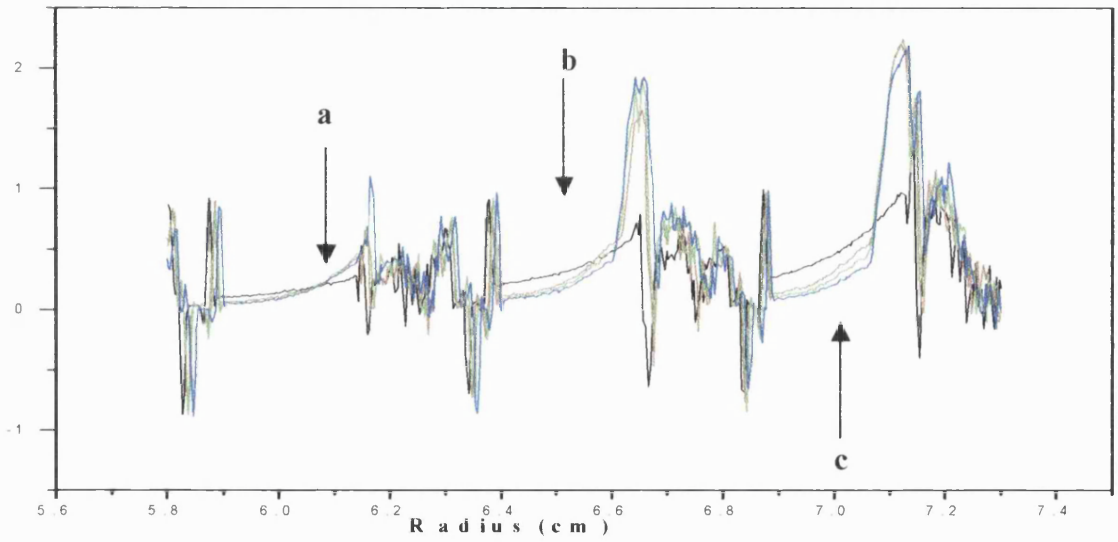


Figure 4.12 Absorbance versus radial distance profiles of the H-NS 1-64 Q16P polypeptide. a, b and c refer to three different loading concentrations of 156, 310 and 498 μM, respectively.

It is clear from both size exclusion and AUC experiments that the H-NS 1-64 Q16P polypeptide has the propensity for aggregation. This has precluded its analysis in terms of the oligomeric state adopted. Based on the CD experiments, global unfolding of the coiled-coil motif can be excluded. It is unlikely that the H-NS 1-64 Q16P mutant is a monomer. Recent studies have shown that the H-NS L26P mutant polypeptide lacks the ability to form higher-order oligomers but adopts a dimeric configuration in spite of the destabilising effects of the proline mutation on the coiled-coil domain (Badaut *et al.*, 2002).

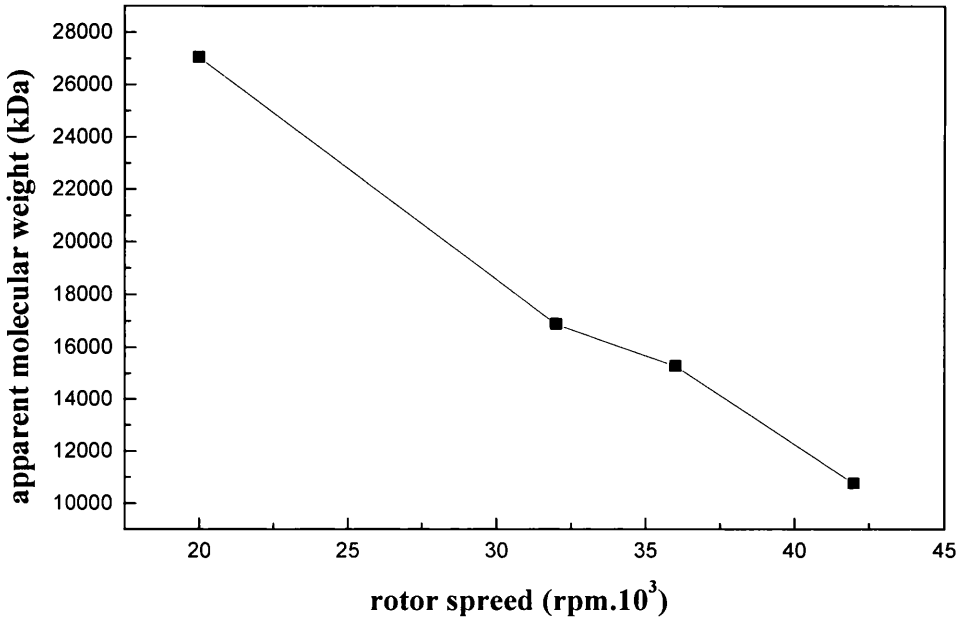


Figure 4.13 Decrease in the apparent molecular weight for the H-NS 1-64 Q16P polypeptide. The apparent molecular weights extracted from Fig. 4.12 (a) are plotted against rotor speed.

4.12 Summary

The investigations described in Chapter 3 aimed to address the importance of the N-terminal region (encompassing H1 and H2 helices) in mediating the self-assembly of the H-NS 1-89 polypeptide. The information thus obtained suggests that the integrity of this region is essential for the formation of the native quaternary structure. The biophysical properties of the H-NS 12-64 polypeptide imply that the N-terminal truncation does not significantly alter coiled-coil dimer stability. In accordance with other data presented, it can be concluded that the self-assembly of H-NS 1-89 dimers (and H-NS full-length) involves the interaction between amino acid residues at the N-terminal end (H1 and H2) and the region encompassing residues 65-89. The underlying molecular mechanism of such interaction remains elusive.

In this chapter, an attempt was made to assign the structural role of the N-terminal helices H1 and H2 to specific functions in terms of their involvement in supporting higher-order oligomer formation. The 2D NMR studies on the H-NS 1-64 K5P mutant (Fig. 4.2) demonstrate specific disruption of the helical structure of H1. When present in the H-NS 1-89 polypeptide, the equivalent substitution has no effects on its oligomeric properties. The concentration dependence observed from size-exclusion experiments on H-NS 1-89 K5P mutant polypeptide (Fig. 4.1) offers clear, though qualitative, evidence for an associative behaviour. Taken together, gel-filtration and NMR studies show that the H-NS 1-89 K5P mutant is capable of self-assembly to form oligomeric structure reminiscent of the H-NS 1-89 polypeptide. Additional CD spectroscopic experiments of the mutant H-NS 1-64 K5P polypeptide demonstrate that the proline substitution had no major effect on the structure of the N-terminal domain. As shown in Figure 4.4, H-NS 1-64 K5P yields a CD spectrum that is almost identical to that of the H-NS 1-64 polypeptide, demonstrating that no major changes in the overall structure occur upon lysine 5 to proline substitution. The

analysis of the oligomeric properties of the mutant H-NS 1-64 K5P polypeptide by size-exclusion chromatography and AUC shows that the ability to self-associate to form dimers is not disrupted (Fig. 4.5 and 4.6, respectively). The evidence presented suggests that the H1 helix is not required for the preservation of a proper oligomerisation interface at the N-terminal region of the H-NS 1-89 polypeptide.

The analysis of the second H-NS mutant, engineered in an attempt to disrupt the helical structure of H2, clearly shows different hydrodynamic and spectroscopic properties. Unlike the subtle structural alterations observed in the H-NS 1-64 K5P mutant polypeptide, substitution of the glutamine 16 with a proline residue induces significant changes in the structural composition (i.e. α-helical content) as observed by CD spectroscopy (Fig. 4.10). The analysis of the obtained CD data substantiates these observations, as it shows that the calculated distribution of secondary structures contains less α-helical content (e.g. Table 4.3) compared with the H-NS 1-64 polypeptide. Contrasting the K5P substitution, which causes a small decrease in the helical content (75 % versus 80.5 % in the wild-type), the introduction of a proline side-chain at position 16 causes a significant loss of the α-helical structure (43.4 % versus 80.5 % in the wild-type). The three-dimensional structure of the H-NS 1-57 polypeptide (Esposito *et al.*, 2002) shows the first 49 residues involved in the structure formation. In terms of the decrease in the overall helical content in the H-NS 1-64 Q16P mutant polypeptide, this could correspond to the unfolding of H1, H2 and residues at the beginning of the coiled coil.

2D NMR studies of the H-NS 1-64 Q16P mutant polypeptide show global changes in the appearance of the spectrum (Fig. 4.8), and this has already been observed in the NMR studies of the N-terminally truncated H-NS 1-64 derivative (Fig. 3.10).

The appearance of the spectrum shown in Figure 4.8 can be rationalised in similar terms used to explain the dramatic changes observed during 2D NMR studies of H-NS 12-64 polypeptide. It is evident that the Q16P mutation impairs the ability of the H-NS 1-64 polypeptide to attain the wild-type structure by interfering with the formation of a α-helical structure at the N-terminal region. It is likely that the Q16P mutation causes structural defects similar to the removal of the N-terminal 11 residues (i.e. H2 and residues at the beginning of the coiled-coil). The appropriate conformation and the structural integrity of the H2 may therefore be an important factor in stabilising both H1 and the N-terminal region of the coiled-coil motif. The introduction of deletion mutation (e.g. alanine) that removes an interaction without causing a disruption of structure (Fersht, 1998) would potentially provide more quantitative information as to the stabilising effects of H2.

When introduced in the H-NS 1-89 polypeptide, proline substitution strongly affects its self-associative behaviour. As shown in Figure 4.7, size-exclusion chromatography experiments show dramatic changes in the elution behaviour of the H-NS 1-89 Q16P mutant polypeptide. Unlike H-NS 12-89 and 17-89 truncation mutants, which show a single peak on size-exclusion column, H-NS 1-89 Q16P polypeptide elutes as a doublet. Nevertheless, the elution volume of the second peak (~14.4 ml) is similar to that of the H-NS 12-89 (Fig. 3.6) polypeptide shown to adopt dimeric configuration. It is probable that the presence of the unfolded N-terminal tail corresponding to the region encompassing H1 and H2 complicates the elution behaviour due to the non-specific hydrophobic interactions. The unfolding of the N-terminal region in the H-NS 1-64 Q16P mutant polypeptide may also be responsible for the size-exclusion chromatographic behaviour shown in Figure 4.11.

It is apparent that the H-NS 1-64 Q16P mutant exhibits concentration-dependent aggregation phenomena (Fig. 4.11). This is in stark contrast to the chromatographic properties of the H-NS 12-64 and 1-64 K5P polypeptides. In the case of the H-NS 12-64 polypeptide, the lack of the 11 amino acid region may reduce such non-specific interactions. Similarly, the H-NS 1-64 K5P mutant polypeptide with its structured H2 helix and the coiled-coil region may not provide enough of a hydrophobic surface that is favourable for association.

In summary, studies on both truncation and point-mutant derivatives of H-NS 1-89 and 1-64 polypeptides corroborate the notion that the residues at the N-terminus of the H-NS protein are important determinants of the native quaternary structure. Furthermore, changes in the H2 region induce destabilisation of the H1 and part of the coiled-coil region. In the context of the H-NS 1-46 polypeptide, the R11E/R14A double mutant does not affect global stability of the fold (Bloch *et al.*, 2003). The conservative mutations of these surface exposed residues within H2 helix are unlikely to disturb hydrophobic packing (I10 and L13) of this region against the coiled-coil core. The importance of the N-terminal coiled-coil region in the H-NS self-assembly process was also suggested by Rimsky and co-workers (Badaut *et al.*, 2002). In their studies, the destabilisation of the coiled-coil region by L26P substitution was shown to strongly affect self-association of H-NS polypeptide. More specifically, the H-NS L26P mutant polypeptide was shown to adopt a dimeric configuration. The H-NS L26P dimer recognises specific DNA sequences but fails to undergo a polymerisation step and thus lead to the formation of a nucleoprotein complex capable of transcriptional repression (Badaut *et al.*, 2002).

In the light of the structural information on the N-terminal domain organisation, besides affecting the heptad positions of the coiled-coil core, a L26P mutation is likely to have a destabilising influence on the N-terminal helices H1 and H2 (via disruption of the hydrophobic ridge formed between H1 and H2 and the groove between H3 helices). Indeed, the H-NS 1-64 L26P mutant shows an NMR spectrum typical of a random coil (Bloch *et al.*, 2003) that is reminiscent of the spectroscopic data observed for the H-NS 12-64 and 1-64 Q16P polypeptides in this study.

It is suggested here that H2 helix, directly or indirectly, supports a conformation in which H-NS self-association is more favourable. Furthermore, residues within H2 (R11 and R14) increase affinity and specificity for intrinsically curved DNA (Bloch *et al.*, 2003). The structural basis for such functional plurality remains to be determined.

5. General discussion

Many of the structural and functional details of the H-NS protein have been identified, although only a fraction of these have yielded meaningful clues as to the mechanism(s) of H-NS-dependent transcriptional repression. A consistent feature that has emerged from a number of studies suggests that the specific regulatory activities of H-NS are mediated via the formation of a nucleoprotein structure that interferes with the establishment of a transcriptionaly competent complex between RNA polymerase and DNA sequences at promoter regions (Schröder and Wagner, 2000; Rimsky *et al.*, 2001; Badaut *et al.*, 2002; Dame *et al.*, 2002). The formation of such nucleoprotein assemblies often involves additional regulatory factors, which may directly or indirectly modulate H-NS function (Schröder and Wagner, 2002, and references therein). Moreover, environmental changes such as temperature have a profound effect on the extent of repression (Falconi *et al.*, 1998; White-Ziegler *et al.*, 1998). The H-NS/DNA nucleoprotein structure appears to be responsive to external stimuli *per se* and as such may represent a focal point between environmental input signals and changes in the gene expression (Badaut *et al.*, 2002; Amit *et al.*, 2003). It has been established that specific binding of H-NS to DNA is a multi-step process, which critically depends on the presence of a curved DNA sequence and subsequent polymerisation of H-NS along the DNA fragment, thus leading to the formation of a nucleoprotein complex (Rimsky *et al.*, 2001; Badaut *et al.*, 2002). Previous biochemical and biophysical studies have shown that H-NS forms higher-order oligomers both in solution and on the DNA template (Smyth *et al.*, 2000; Dame *et al.*, 2000, 2001; Badaut *et al.*, 2002; Amit *et al.*, 2003). The structural basis of H-NS self-association equilibrium remained poorly defined.

The investigations described in this thesis aimed to further our understanding of the mechanisms of oligomerisation of H-NS in solution. In particular, several questions were addressed: 1) what is the stoichiometry of the lowest oligomeric form adopted by H-NS polypeptide, (2) what effect does removal of the N-terminal region have on the oligomeric properties of H-NS polypeptide, (3) what is the contribution of the N-terminal helices H1 and H2 to the formation of quaternary structure by H-NS polypeptide, (4) do helices H1 and H2 contribute towards stability of the coiled-coiled region and (5) what is the structural basis for the interaction between H-NS and StpA polypeptides. These questions have been approached experimentally using a combination of deletion and site-directed mutagenesis strategies with subsequent analysis of mutant polypeptides by a number of complementary biophysical techniques. Results thus obtained have been used to suggest a possible mechanism of H-NS oligomerisation.

In agreement with previous studies, removal of the C-terminal DNA binding domain does not affect the oligomerisation of H-NS, highlighting the notion that the structural determinants responsible for higher-order oligomer formation reside within first 89 residues. This is in contrast to the N-terminal H-NS 1-64 polypeptide, which completely loses the ability to form higher-order oligomers. The detailed examination of the oligomeric properties of H-NS 1-64 polypeptide by sedimentation equilibrium shows that it adopts a dimeric configuration. Results obtained hitherto suggest that the N-terminal regions of other H-NS proteins and its homologues also self-associate to form dimers (Badaut *et al.*, 2002; Ladbury and Arold, personal communication).

Although there is some confusion in the literature as to the three-dimensional organisation of the N-terminal region of H-NS, it is clear that the dimerisation interface is formed through a coiled-coil motif (Esposito *et al.*, 2002; Bloch *et al.*, 2003).

It was hypothesised that formation of H-NS dimers via a coiled-coil motif residing in the N-terminal region constitutes the core event, which allows the respective dimers to polymerise through the second interface to form higher-order oligomers characteristic of the full-length H-NS. In an attempt to delineate regions encompassing second interface, it was shown that truncation of the N-terminal 11 residues, predicted to lie outside the coiled-coil region, significantly weakens H-NS self-assembly. Furthermore, AUC experiments clearly indicate that the H-NS 12-89 polypeptide behaves as a dimer. Based on the spectroscopic and hydrodynamic analysis of the H-NS 12-64 polypeptide, it is shown that despite the structural disruption of the H2 helix and part of the coiled-coil region, the N-terminal truncation did not affect dimer formation. Taken together, these results imply that self-assembly of H-NS dimers involves residues at the N-terminus in addition to the region encompassing central part of the polypeptide (residues 64-89). As noted before (Chapter 1) there is a lack of structural information available for the region connecting the N-terminal dimerisation and the C-terminal DNA binding domain. Solution studies have shown that residues 60-89 are unfolded within the context of the H-NS 60-137 polypeptide (Shindo *et al.*, 1999). It has been suggested that the interaction of H-NS and DNA induces a structuring of the central region, which allows H-NS to polymerise along the DNA template (Bloch *et al.*, 2003). However, it is known that even in the absence of DNA, the H-NS polypeptide is capable of self-association to form oligomeric structures larger than dimer (Smyth *et al.*, 2000; Badaut *et al.*, 2002).

Furthermore, spectroscopic studies of the H-NS R54C mutant polypeptide, known to be defective in oligomerisation, have indicated a significant disruption of the structure, which is predicted to adopt an α-helical conformation (Schröder *et al.*, 2001). In accordance with the above evidence, the region encompassing residues 50-80 is predicted to form a helix-turn-helix motif (S. Arold, personal communication).

The importance of the structural integrity of the N-terminal dimerisation domain in the formation of the higher-order oligomers has also been established in studies by Rimsky and co-workers (Badaut *et al.*, 2002). It has been shown that dominant-negative mutants of the H-NS polypeptide modified in the N-terminal region are unable to form higher-order oligomers, but instead adopt a dimeric configuration. These H-NS mutant dimers have been hypothesised to represent initial structures, and are able to recognise specific sites on DNA, prior to self-association and the formation of a nucleoprotein complex capable of transcriptional repression (Badaut *et al.*, 2002).

The results presented in this thesis further indicate the differential role of the two N-terminal helices H1 and H2 in supporting higher-order oligomer formation. Both hydrodynamic and spectroscopic evidence suggest that disruption of the helix H1 has no effects on the H-NS oligomerisation. Equally, as shown by NMR studies, disruption of the H1 helix has little effect on the structural stability of the N-terminal dimerisation domain. This is in contrast to the H2 helix whose integrity was shown to be an important determinant of the H-NS quaternary structure. As well as compromising the stability of the H1 and H2, the Q16P mutation also leads to the destabilisation of the part of the coiled-coil region.

Based on the comparative spectroscopic evidence, it is suggested that the disruption of H2 has equivalent structural effects to the removal of the N-terminal 11 residues. These results indicate that the structure and/or position of the H2 helix is intimately involved in (1) partially contributing to the stability of the N-terminal dimerisation domain and (2) supporting higher-order oligomer formation. Recent studies indicate that residues within H2 helix may also provide structural platform for the specific recognition of the intrinsically curved DNA sequences (Bloch *et al.*, 2003).

As shown in this thesis and other work (Badaut *et al.*, 2002), oligomerisation of the H-NS polypeptide occurs through the association of the core dimeric units in a head-to-tail manner. A likely mechanism for specific H-NS/DNA interaction could involve an association of the dimeric core unit at the intrinsically curved sites, through the simultaneous binding of the N- and C-terminal domains. Subsequent association between the respective dimers might be based on the interaction between helix H2 and the structures predicted to adopt a α- helical conformation in a central region.

It has been shown recently that changes in the temperature and osmolarity prevent polymerisation of H-NS along the DNA (Amit *et al.*, 2003). It is tempting to speculate that conformational changes in the N-terminal dimerisation domain may prevent polymerisation of H-NS by negating the postulated head-to-tail assembly. Similar temperature-induced conformational changes in the coiled-coil domain have been shown to regulate repressor function of TlpA polypeptide (Hurme *et al.*, 1996, 1997). In the light of the differences in the structure of the N-terminal dimerisation domain, the presence of two different conformations cannot be excluded (Esposito *et al.*, 2002; Bloch *et al.*, 2003).

Appendix

Circular dichroism (CD) is a valuable spectroscopic technique and important analytical tool for studying protein structure in solution (Greenfield, 1999). It is a phenomenon that results from the differential absorption of the left and right circularly polarized components of plane-polarized light. CD is exquisitely sensitive to changes of protein conformation, which is a property that has been extensively exploited in modern biochemical and biophysical research. The major optically active groups that contribute to the CD spectra of proteins are the amide bonds of the peptide backbone and the aromatic amino acid side chains. In case of the amide bonds, their placement in different regular secondary structure elements (such as α-helices, β-sheets and β-turns) gives rise to distinct and characteristic CD patterns. For example, the typical CD spectrum of protein, which is rich in α-helices, displays large CD peaks with negative ellipticity at 222 and 208nm, and positive ellipticity at 193nm (Fig. A1). On the other hand β-sheets exhibit a broad negative band near 218nm and a large positive band near 195nm. Far-UV CD (CD in the wavelength region of 170-260 nm) is sensitive to changes in the secondary structure, whereas the near-UV CD signal (260-330 nm) is diagnostic of the chiral environment of aromatic residues and can serve as useful probe of tertiary structure of proteins. Despite the fact the resolution is not as high as with NMR, CD studies are simpler to perform and allow a more rapid quantitative estimate of the secondary structure content of proteins with relatively modest amounts of material. The principal information available from CD studies is the quantitative estimate (see below) of the secondary structure content of proteins (i.e. the analysis of the CD spectrum of a protein in the far-UV region to estimate the percentage of different types of secondary structure present).

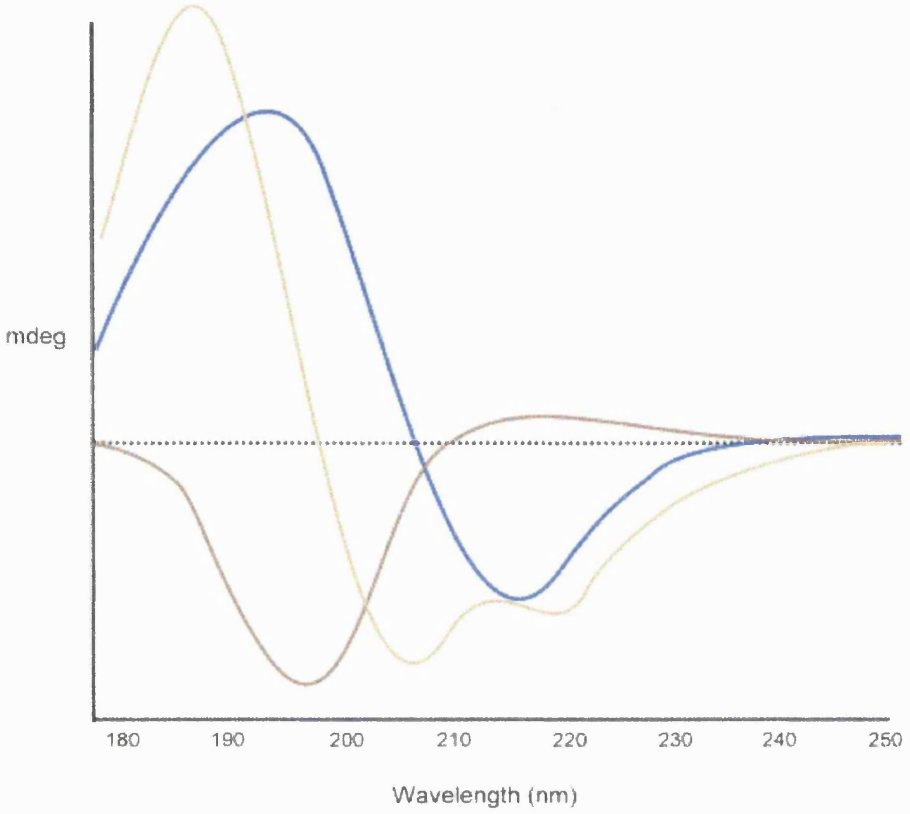


Figure A1. Far-UV CD spectra associated with α-helix (-), β-sheet (-) and unfolded polypeptide (-).

In addition to secondary structure analysis, CD spectroscopy can also be used to 1) monitor the rates of the conformational transitions in proteins (e.g. protein folding), 2) explore structural changes associated with protein-ligand and protein-nucleic acid interaction, thereby allowing the elucidation of the dissociation constant and the stoichiometry of interaction and 3) changes in the CD signal can be used to monitor structural perturbation induced by chaotropic agents (e.g. urea) or heat, thus permitting estimates of the conformational stability.

The problems associated with meaningful CD analysis have both practical and theoretical origins. Like other optical measurements, CD studies require samples, which must be free of contaminating proteins or other optically active impurities. Major problems can arise from uncertainties in the measurement of protein concentration and the ability to collect reliable data over a large range of wavelengths. It is often extremely difficult to obtain data below 190 nm due to the absorption of other components present in the buffer and sometimes it is possible to gather data to only 195 or 200 nm. This greatly diminishes the precision of the structural estimates and necessitates the use of "transparent" buffers and cuvettes with short path-lengths (0.05-0.1 mm). The analysis of secondary structure of proteins is most reliable when structure of the protein of interest has been well characterized (by X-ray crystallography or NMR) and therefore can be used as a reference. Far-UV CD spectra can then be used with high precision to assess possible structural changes of the corresponding mutants.

The basic principle underlying the analysis of protein CD spectra in terms of the secondary structure content relies on the assumption that the spectrum of a protein can be represented by a linear combination of the spectra of the secondary structure elements including a noise term arising from the contribution of aromatic chromophores according to equation (6):

$$\theta_\lambda = \Sigma F_i S_{\lambda i} + \text{noise} \qquad (6)$$

where θ_λ is the CD of the proteins as a function of the wavelength, F_i is the fraction of each secondary structure, *i*, and $S_{\lambda i}$ is the ellipticity at each wavelength of each *i*th secondary structure element.

Numerous empirical methods have been developed for quantitative analysis of protein CD spectra in terms of secondary structure content. The major methods used for extracting information on secondary structure content are:

1) multilinear regression

2) singular value decomposition

3) ridge regression

4) convex constraint analysis

5) neural network analysis

6) self-consistent methods

Generally all of the methods listed above give a reliable estimate of the helical content. An excellent review by Norma Greenfield (Greenfield, 1996) should be consulted for thorough analysis of the advantages and pitfalls of the various methods used to obtain structural information from the CD data.

References

Afflerbach H., Schroder O. and Wagner R. (1998) Effects of the *Escherichia coli* DNA-binding protein H-NS on rRNA synthesis *in vivo*. *Mol.Microbiol.* **28,** 641-653.

Afflerbach H., Schroder O. and Wagner R. (1999) Conformational changes of the upstream DNA mediated by H-NS and FIS regulate *E. coli rrnB* P1 promoter activity. *J.Mol.Biol.* **286,** 339-353.

Amit R., Oppenheim A.B. and Stavans J. (2003) Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. *Biophys.J.* **84,** 2467-2473.

Atlung T. and Ingmer H. (1997) H-NS: a modulator of environmentally regulated gene expression. *Mol.Microbiol.* **24,** 7-17.

Azam T.A., Iwata A., Nishimura A., Ueda S. and Ishihama A. (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J.Bacteriol.* **181,** 6361-6370.

Azam T.A., Hiraga S. and Ishihama A. (2000) Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells* **5,** 613-626.

Badaut C., Williams R., Arluison V., Bouffartigues E., Robert B., Buc H. and Rimsky S. (2002) The degree of oligomerization of the H-NS nucleoid structuring protein is related to specific binding to DNA. *J.Biol.Chem.* **277,** 41657-41666.

Barth M., Marschall C., Muffler A., Fischer D. and Hengge-Aronis R. (1995) Role for the histone-like protein H-NS in growth phase-dependent and osmotic regulation of sigma S and many sigma S-dependent genes in *Escherichia coli*. *J.Bacteriol.* **177,** 3455-3464.

Beloin C. and Dorman C.J. (2003) An extended role for the nucleoid structuring protein H-NS in the virulence gene regulatory cascade of *Shigella flexneri*. *Mol.Microbiol.* **47,** 825-838.

Bertin P., Benhabiles N., Krin E., Laurent-Winter C., Tendeng C., Turlin E., Thomas A., Danchin A. and Brasseur R. (1999) The structural and functional organization of H-NS-like proteins is evolutionarily conserved in gram-negative bacteria. *Mol.Microbiol.* **31,** 319-329.

Bertin P., Hommais F., Krin E., Soutourina O., Tendeng C., Derzelle S. and Danchin A. (2001) H-NS and H-NS-like proteins in Gram-negative bacteria and their multiple role in the regulation of bacterial metabolism. *Biochimie* **83,** 235-241.

Bloch V., Yang Y., Margeat E., Chavanieu A., Auge M.T., Robert B., Arold S., Rimsky S. and Kochoyan M. (2003) The H-NS dimerization domain defines a new fold contributing to DNA recognition. *Nat.Struct.Biol.* **10,** 212-218.

Bracco L., Kotlarz D., Kolb A., Diekmann S. and Buc H. (1998) Synthetic curved DNA sequences can act as transcriptional activators in *Escherichia coli*. *EMBO J.* **8,** 4289-4296.

Browning D.F., Cole J.A. and Busby S.J. (2000) Suppression of FNR-dependent transcription activation at the *Escherichia coli nir* promoter by Fis, IHF and H-NS: modulation of transcription initiation by a complex nucleo-protein assembly. *Mol.Microbiol.* **37,** 1258-1269.

Brunetti R., Prosseda G., Beghetto E., Colonna B. and Micheli G. (2001) The looped domain organization of the nucleoid in histone-like protein defective *Escherichia coli* strains. *Biochimie* **83,** 873-882.

Caramel A. and Schnetz K. (1998) Lac and lambda repressors relieve silencing of the *Escherichia coli bgl* promoter. Activation by alteration of a repressing nucleoprotein complex. *J.Mol.Biol.* **284,** 875-883.

Ceschini S., Lupidi G., Coletta M., Pon C.L., Fioretti E. and Angeletti M. (2000) Multimeric self-assembly equilibria involving the histone-like protein H-NS. A thermodynamic study. *J.Biol.Chem.* **275,** 729-734.

Creighton T.E. (1993) *Proteins: Structures and Molecular Properties.* W.H. Freeman and Company, New York, NY, USA.

Cusick M.E. and Belfort M. (1998) Domain structure and RNA annealing activity of the *Escherichia coli* regulatory protein StpA. *Mol.Microbiol.* **28,** 847-857.

Dame R.T., Wyman C. and Goosen N. (2000) H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res.* **28,** 3504-3510.

Dame R.T., Wyman C. and Goosen N. (2001) Structural basis for preferential binding of H-NS to curved DNA. *Biochimie* **83,** 231-234.

Dame R.T. and Goosen N. (2002) HU: promoting or counteracting DNA compaction? *FEBS Lett.* **529,** 151-156.

Dame R.T., Wyman C., Wurm R., Wagner R. and Goosen N. (2002) Structural basis for H-NS-mediated trapping of RNA polymerase in the open initiation complex at the *rrnB* P1. *J.Biol.Chem.* **277,** 2146-2150.

Decanniere K., Babu A.M., Sandman K., Reeve J.N. and Heinemann U. (2000) Crystal structure of recombinant histones HmfA and HmfB from the hyperthermophilic archaeon *Methanothermus fervidus*. *J.Mol.Biol.* **303,** 35-47.

Deighan P., Free A. and Dorman C.J. (2000) A role of the *Escherichia coli* H-NS-like protein StpA in OmpF porin expression through modulation of *micF* RNA stability. *Mol.Microbiol.* **38,** 126-139.

Deighan P., Beloin C. and Dorman C.J. (2003) Three-way interactions among the Sfh, StpA and H-NS nucleoid-structuring proteins of *Shigella flexneri* 2a strain 2457T. *Mol.Microbiol.* **48,** 1401-1416.

Dersch P., Schmidt K. and Bremer E. (1993) Synthesis of the *Escherichia coli* K-12 nucleoid-associated DNA-binding protein H-NS is subjected to growth-phase control and autoregulation. *Mol.Microbiol.* **8,** 875-889.

Donato G.M. and Kawula T.H. (1998) Enhanced binding of altered H-NS protein to flagellar rotor protein FliG causes increased flagellar rotational speed and hypermotility in *Escherichia coli*. *J.Biol.Chem.* **273,** 24030-24036.

Donato G.M. and Kawula T.H. (1999) Phenotypic analysis of random *hns* mutations differentiate DNA-binding activity from properties of *fimA* promoter inversion modulation and bacterial motility. *J.Bacteriol.* **181,** 941-948.

Dorman C.J. and Porter M.E. (1998) The *Shigella* virulence gene regulatory cascade: a paradigm of bacterial gene control mechanisms. *Mol.Microbiol.* **29,** 677-684.

Dorman C.J., Hinton J.C. and Free A. (1999) Domain organization and oligomerization among H-NS-like nucleoid-associated proteins in bacteria. *Trends Microbiol.* **7,** 124-128.

Dorman C.J. and Deighan P. (2003) Regulation of gene expression by histone-like proteins in bacteria. *Curr.Opin.Genet.Dev.* **13,** 179-184.

Drlica K. and Rouviere-Yaniv J. (1987) Histonelike proteins of bacteria. *Microbiol.Rev.* **51,** 301-319

Durrenberger M., La Teana A., Citro G., Venanzi F., Gualerzi C.O. and Pon C.L. (1991) *Escherichia coli* DNA-binding protein H-NS is localized in the nucleoid. *Res.Microbiol.* **142,** 373-380.

Esposito D., Petrovic A., Harris R., Ono S., Eccleston J.F., Mbabaali A., Haq I., Higgins C.F., Hinton J.C., Driscoll P.C. and Ladbury J.E. (2002) H-NS oligomerization domain structure reveals the mechanism for high order self-association of the intact protein. *J.Mol.Biol.* **324,** 841-850.

Falconi M., Gualtieri M.T., La Teana A., Losso M.A. and Pon C.L. (1988) Proteins from the prokaryotic nucleoid: primary and quaternary structure of the 15-kD Escherichia coli DNA binding protein H-NS. *Mol.Microbiol.* **2,** 323-329.

Falconi M., McGovern V., Gualerzi C., Hillyard D. and Higgins N.P. (1991) Mutations altering chromosomal protein H-NS induce mini-Mu transposition. *New Biol.* **3,** 615-625.

Falconi M., Higgins N.P., Spurio R., Pon C.L. and Gualerzi C.O. (1993) Expression of the gene encoding the major bacterial nucleotide protein H-NS is subject to transcriptional auto-repression. *Mol.Microbiol.* **10,** 273-282.

Falconi M., Brandi A., La Teana A., Gualerzi C.O. and Pon C.L. (1996) Antagonistic involvement of FIS and H-NS proteins in the transcriptional control of *hns* expression. *Mol.Microbiol.* **19,** 965-975.

Falconi M., Colonna B., Prosseda G., Micheli G. and Gualerzi C.O. (1998) Thermoregulation of *Shigella* and *Escherichia coli* EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of *virF* promoter to transcriptional repressor H-NS. *EMBO J.* **17,** 7033-7043.

Falconi M., Prosseda G., Giangrossi M., Beghetto E. and Colonna B. (2001) Involvement of FIS in the H-NS-mediated regulation of *virF* gene of *Shigella* and enteroinvasive *Escherichia coli*. *Mol.Microbiol.* **42,** 439-452.

Fersht A. (1998) *Structure and Mechanism in Protein Science*. W. H. Freeman and Company, New York, NY, USA.

Forood B., Feliciano E.J. and Nambiar K.P. (1993) Stabilisation of α-helical structures in short peptides via end capping. *Proc.Natl.Acad.Sci.U.S.A* **90,** 838-842.

Free A. and Dorman C.J. (1995) Coupling of Escherichia coli *hns* mRNA levels to DNA synthesis by autoregulation: implications for growth phase control. *Mol.Microbiol.* **18,** 101-113.

Free A. and Dorman C.J. (1997) The *Escherichia coli stpA* gene is transiently expressed during growth in rich medium and is induced in minimal medium and by stress conditions. *J.Bacteriol.* **179,** 909-918.

Friedrich K., Gualerzi C.O., Lammi M., Losso M.A. and Pon C.L. (1988) Proteins from the prokaryotic nucleoid. Interaction of nucleic acids with the 15 kDa *Escherichia coli* histone-like protein H-NS. *FEBS Lett.* **229,** 197-202.

Gerstel U., Park C. and Romling U. (2003) Complex regulation of *csgD* promoter activity by global regulatory proteins. *Mol.Microbiol.* **49,** 639-654.

Gill S.C. and von Hippel P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal.Biochem.* **182,** 319-326. Erratum in: *Anal.Biochem.* (1990), **189,** 283.

Göransson M., Sonden B., Nilsson P., Dagberg B., Forsman K., Emanuelsson K. and Uhlin B.E. (1990) Transcriptional silencing and thermoregulation of gene expression in *Escherichia coli. Nature.* **344,** 682-685.

Govantes F., Orjalo A.V. and Gunsalus R.P. (2000) Interplay between three global regulatory proteins mediates oxygen regulation of the *Escherichia coli* cytochrome d oxidase (*cydAB*) operon. *Mol.Microbiol.* **38,** 1061-1073.

Greenfield N.J. (1996) Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Anal. Biochem.* **235,** 1-10.

Greenfield N.J. (1999) Application of circular dichroism in protein and peptide analysis. *Trends Anal. Chem.* **18,** 236-244.

Gualerzi C.O., Losso M.A., Lammi M., Friedrich K., Pawlik R.T., Canonaco M.A., Gianfranceschi G., Pingoud A. and Pon C.L. (1986) Proteins from the prokaryotic nucleoid. Structural and functional characterization of the *Escherichia coli* DNA-binding proteins NS (HU) and H-NS. In *Bacterial Chromatin*, pp 101-134. Edited by Gualerzi C.O. and Pon C.L. Heidelberg: Springer, Berlin.

Hall T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl.Acids.Symp.Ser.* **41,** 95-98.

Harbury P.B., Zhang T., Kim P.S. and Alber T. (1993) A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262,** 1401-1407.

Harbury P.B., Kim P.S. and Alber T. (1994) Crystal structure of an isoleucine-zipper trimer. *Nature* **371,** 80-83.

Hayat M.A. and Mancarella D.A. (1996) Nucleoid proteins. *Micron* **26,** 461-480.

Higgins C.F., Hinton J.C., Hulton C.S., Owen-Hughes T., Pavitt G.D. and Seirafi A. (1990) Protein H1: a role for chromatin structure in the regulation of bacterial gene expression and virulence? *Mol.Microbiol.* **4,** 2007-2012.

Higgins C.F., Dorman C.J., Stirling D.A., Waddell L., Booth I.R., May G. and Bremer E. (1988) A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *S.Typhimurium* and *E.coli. Cell* **52,** 569-584.

Hinton J.C., Santos D.S., Seirafi A., Hulton C.S., Pavitt G.D. and Higgins C.F. (1992) Expression and mutational analysis of the nucleoid-associated protein H-NS of *Salmonella typhimurium. Mol.Microbiol.* **6,** 2327-2337.

Hommais F., Krin E., Laurent-Winter C., Soutourina O., Malpertuy A., Le Caer J.P., Danchin A. and Bertin P. (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol.Microbiol.* **40,** 20-36.

Huang R. and Reusch R.N. (1996) Poly(3-hydroxybutyrate) is associated with specific proteins in the cytoplasm and membranes of *Escherichia coli. J.Biol.Chem.* **271,** 22196-22202.

Hulton C.S., Seirafi A., Hinton J.C., Sidebotham J.M., Waddell L., Pavitt G.D., Owen-Hughes T., Spassky A., Buc H. and Higgins C.F. (1990) Histone-like protein H1 (H-NS), DNA supercoiling, and gene expression in bacteria. *Cell* **63,** 631-642.

Humphrey W., Dalke A. and Schulten A. (1996) VMD – Visual Molecular Dynamics. *J.Molec.Graphics.* **14.1,** 33-38.

Hurme R., Berndt K.D., Namork E. and Rhen M (1996) DNA binding exerted by a bacterial gene regulator with an extensive coiled-coil domain. *J.Biol.Chem.* **271,** 12626-12631.

Hurme R., Berndt K.D., Normark S.J. and Rhen M (1997) A proteinaceous gene regulatory thermometer in *Salmonella. Cell.* **90,** 55-64.

Hurme R. and Rhen M. (1998) Temperature sensing in bacterial gene regulation: what it all boils down to. *Mol.Microbiol.* **30,** 1-6.

Jacquet M., Cukier-Kahn R., Pla J. and Gros F. (1971) A thermostable protein factor acting on *in vitro* DNA transcription. *Biochem.Biophys.Res.Commun.* **45,** 1597-1607.

Johansson J., Dagberg B., Richet E. and Uhlin B.E. (1998) H-NS and StpA proteins stimulate expression of the maltose regulon in *Escherichia coli. J.Bacteriol.* **180,** 6117-6125.

Johansson J. and Uhlin B.E. (1999) Differential protease-mediated turnover of H-NS and StpA revealed by a mutation altering protein stability and stationary-phase survival of *Escherichia coli. Proc.Natl.Acad.Sci.U.S.A* **96,** 10776-10781.

Johansson J., Eriksson S., Sonden B., Wai S.N. and Uhlin B.E. (2001) Heteromeric interactions among nucleoid-associated bacterial proteins: localization of StpA-stabilizing regions in H-NS of *Escherichia coli. J.Bacteriol.* **183,** 2343-2347.

Jordi B.J., Fielder A.E., Burns C.M., Hinton J.C., Dover N., Ussery D.W. and Higgins C.F. (1997) DNA binding is not sufficient for H-NS-mediated repression of *proU* expression. *J.Biol.Chem.* **272,** 12083-12090.

Jordi B.J. and Higgins C.F. (2000) The downstream regulatory element of the *proU* operon of *Salmonella typhimurium* inhibits open complex formation by RNA polymerase at a distance. *J.Biol.Chem.* **275,** 12123-12128.

Kellenberger E., Carlemalm E., Sechaud J., Ryter A. and de Haller G. (1986) Considerations on the condensation and the degree of compactness in non-eukaryotic DNA-containing plasmas. In *Bacterial Chromatin*, pp 11-25. Edited by Gualerzi C.O. and Pon C.L. Heidelberg: Springer, Berlin.

Kornberg R.D. and Lorch Y. (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell.* **98,** 285-294.

La Teana A., Brandi A., Falconi M., Spurio R., Pon C.L. and Gualerzi C.O. (1991) Identification of a cold shock transcriptional enhancer of the *Escherichia coli* gene encoding nucleoid protein H-NS. *Proc.Natl.Acad.Sci.U.S.A* **88,** 10907-10911.

Laue T. (1995) Sedimentation equilibrium as thermodynamic tool. *Methods Enzymol.* **259,** 427-452.

Laurent-Winter C., Ngo S., Danchin A. and Bertin P. (1997) Role of *Escherichia coli* histone-like nucleoid-structuring protein in bacterial metabolism and stress response: identification of targets by two-dimensional electrophoresis. *Eur.J.Biochem.* **244,** 767-773.

Lease R.A., Cusick M.E. and Belfort M. (1998) Riboregulation in *Escherichia coli*: DsrA RNA acts by RNA:RNA interactions at multiple loci. *Proc.Natl.Acad.Sci.U.S.A* **95,** 12456-12461.

Lease R.A. and Belfort M. (2000a) A trans-acting RNA as a control switch in *Escherichia coli*: DsrA modulates function by forming alternative structures. *Proc.Natl.Acad.Sci.U.S.A* **97,** 9919-9924.

Lease R.A. and Belfort M. (2000b) Riboregulation by DsrA RNA: trans-actions for global economy. *Mol.Microbiol.* **38,** 667-672.

Lejeune P. and Danchin A. (1990) Mutations in the *bglY* gene increase the frequency of spontaneous deletions in *Escherichia coli* K-12. *Proc.Natl.Acad.Sci.U.S.A* **87,** 360-363.

Liu Q. and Richardson C.C. (1993) Gene 5.5 protein of bacteriophage T7 inhibits the nucleoid protein H-NS of *Escherichia coli*. *Proc.Natl.Acad.Sci.U.S.A* **90,** 1761-1765.

Lobley A., Whitmore L. and Wallace B.A. (2002) DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* **18,** 211-212.

Lucht J.M., Dersch P., Kempf B. and Bremer E. (1994) Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled *proU* operon of *Escherichia coli*. *J.Biol.Chem.* **269,** 6578.

Luger K., Mäder A.W., Richmond R.K., Sargent D.F. and Richmond T.J. (1997) Crystal structure of the nucleosome core particle at 2.8Å resolution. *Nature.* **389,** 251-260.

Lupas A., van Dyke M. and Stock J. (1991) Predicting coiled coils from protein sequences. *Science* **252,** 1162-1164.

Lupas A. (1996) Coiled coils: new structures and new function. *Trends Biochem. Sci.* **21,** 375-382.

Madrid C., Nieto J.M., Paytubi S., Falconi M., Gualerzi C.O. and Juarez A. (2002) Temperature- and H-NS-dependent regulation of a plasmid-encoded virulence operon expressing *Escherichia coli* hemolysin. *J.Bacteriol.* **184,** 5058-5066.

Malashkevich V.N., Kammerer R.A., Efimov V.P., Schulthess T. and Engel J. (1996) The crystal structure of a five-stranded coiled coil in COMP: a prototype ion channel? *Science* **274,** 761-765.

Marsh M. and Hillyard D.R. (1990) Nucleotide sequence of hns encoding the DNA-binding protein H-NS of *Salmonella typhimurium*. *Nucleic Acids Res.* **18,** 3397.

May G., Dersch P., Haardt M., Middendorf A. and Bremer E. (1990) The *osmZ* (*bglY*) gene encodes the DNA-binding protein H-NS (H1a), a component of the *Escherichia coli* K12 nucleoid. *Mol.Gen.Genet.* **224,** 81-90.

McGovern V., Higgins N.P., Chiz R.S. and Jaworski A. (1994) H-NS over-expression induces an artificial stationary phase by silencing global transcription. *Biochimie* **76,** 1019-1029.

McLeod S. and Johnson R.C. (2001) Control of transcription by nucleoid proteins. *Curr.Opin.Microbiol.* **4,** 152-159.

Mojica F.J. and Higgins C.F. (1997) *In vivo* supercoiling of plasmid and chromosomal DNA in an *Escherichia coli hns* mutant. *J.Bacteriol.* **179,** 3528-3533.

Musgrave D., Zhang X. and Dinger M. (2002) Archeal genome organisation and stress responses: implications for the origin and evolution of cellular life. *Astrobiology.* **2,** 241-253.

Nieto J.M. and Juarez A. (1999) The putative Orf4 of broad-host-range conjugative plasmid R446 could be related to the H-NS family of bacterial nucleoid-associated proteins. *Plasmid* **41,** 125-127.

Nieto J.M., Madrid C., Prenafeta A., Miquelay E., Balsalobre C., Carrascal M. and Juarez A. (2000) Expression of the hemolysin operon in *Escherichia coli* is modulated by a nucleoid-protein complex that includes the proteins Hha and H-NS. *Mol.Gen.Genet.* **263,** 349-358.

Nieto J.M., Madrid C., Miquelay E., Parra J.L., Rodriguez S. and Juarez A. (2002) Evidence for direct protein-protein interaction between members of the enterobacterial Hha/YmoA and H-NS families of proteins. *J.Bacteriol.* **184,** 629-635.

Nye M.B., Pfau J.D., Skorupski K. and Taylor R.K. (2000) *Vibrio cholerae* H-NS silences virulence gene expression at multiple steps in the ToxR regulatory cascade. *J.Bacteriol.* **182,** 4295-4303.

O'Gara J.P. and Dorman C.J. (2000) Effects of local transcription and H-NS on inversion of the *fim* switch of *Escherichia coli*. *Mol.Microbiol.* **36,** 457-466.

O'Shea E.K., Klemm J.D., Kim P.S. and Alber T. (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled-coil. *Science* **254,** 539-544.

Owen-Hughes T.A., Pavitt G.D., Santos D.S., Sidebotham J.M., Hulton C.S., Hinton J.C. and Higgins C.F. (1992) The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. *Cell* **71,** 255-265.

Palchaudhuri S., Tominna B. and Leon M.A. (1998) H-NS regulates DNA repair in *Shigella*. *J.Bacteriol.* **180,** 5260-5262.

Pettijohn, D.E. (1998) Histone-like proteins and bacterial chromosome structure. *J.Biol.Chem.* **263,** 12793-12796.

Renzoni D., Esposito D., Pfuhl M., Hinton J.C., Higgins C.F., Driscoll P.C. and Ladbury J.E. (2000) Structural characterization of the N-terminal oligomerisation domain of the bacterial chromatin-structuring protein, H-NS. *J.Mol.Biol.* **306,** 1127-1137.

Reusch R.N., Shabalin O., Crumbaugh A., Wagner R., Schroder O. and Wurm R. (2002) Posttranslational modification of *E. coli* histone-like protein H-NS and bovine histones by short-chain poly-(R)-3-hydroxybutyrate (cPHB). *FEBS Lett.* **527,** 319-322.

Rimsky S. and Spassky A. (1990) Sequence determinats for H1 binding on *Escherichia coli lac* and *gal* promoters. *Biochemistry* **29,** 3765-3771.

Rimsky S., Zuber F., Buckle M. and Buc H. (2001) A molecular mechanism for the repression of transcription by the H-NS protein. *Mol.Microbiol.* **42,** 1311-1323.

Rojo F. (2001) Mechanisms of transcriptional repression. *Curr.Opin.Microbiol.* **4,** 145-151.

Sandman K., Krzycki J.A., Dobrinski B., Lurz R. and Reeve J.N. (1990) Hmf, a DNA-binding protein isolated from the hyperthermophilic archaeon *Methanothermus fervidus*, is most closely related to histones. *Proc.Natl.Acad.Sci.U.S.A* **87,** 5788-5791.

Sandman K., Pereira S.L. and Reeve J.N. (1998) Diversity of prokaryotic chromosomal proteins and the origin of nucleosome. *Cell.Mol.Life.Sci.* **54,** 1350-1364.

Sandman K. and Reeve J.N. (2000) Structure and functional relationships of archeal and eukaryal histones and nucleosomes. *Arch.Microbiol.* **173,** 165-169.

Sandman K., Soares D. and Reeve J.N. (2001) Molecular components of the archeal nucleosome. *Biochimie.* **83,** 277-281.

Schmid M.B. (1990) More than just "histone-like" proteins. *Cell.* **63,** 451-453.

Schneider D.A., Ross W. and Gourse R.L. (2003) Control of *rRNA* expression in *Escherichia coli. Curr.Opin.Microbiol.* **6,** 151-156.

Schneider R., Lurz R., Lüder G., Tolksdorf C., Travers A. and Muskhelishvili G. (2001) An architectural role of the *Escherichia coli* chromatin protein FIS in organizing DNA. *Nucleic Acids Res.* **29,** 5107-5114.

Schroder O. and Wagner R. (2000) The bacterial DNA-binding protein H-NS represses ribosomal RNA transcription by trapping RNA polymerase in the initiation complex. *J.Mol.Biol.* **298,** 737-748.

Schroder O., Tippner D. and Wagner R. (2001) Toward the three-dimensional structure of the *Escherichia coli* DNA-binding protein H-NS: A CD and fluorescence study. *Biochem.Biophys.Res.Commun.* **282,** 219-227.

Schröder O. and Wagner R. (2002) The bacterial regulatory protein H-NS--a versatile modulator of nucleic acid structures. *Biol.Chem.* **383,** 945-960.

Schwede T., Kopp J., Guex N. and Peitsch M.C. (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **31,** 3381-3385.

Shanado Y., Hanada K. and Ikeda H. (2001) Suppression of gamma ray-induced illegitimate recombination in *Escherichia coli* by the DNA-binding protein H-NS. *Mol.Genet.Genomics* **265,** 242-248.

Shi X. and Bennett G.N. (1994) Plasmids bearing *hfq* and the *hns*-like gene *stpA* complement *hns* mutants in modulating arginine decarboxylase gene expression in *Escherichia coli*. *J.Bacteriol.* **176,** 6769-6775.

Shiga Y., Sekine Y., Kano Y. and Ohtsubo E. (2001) Involvement of H-NS in transpositional recombination mediated by IS1. *J.Bacteriol.* **183,** 2476-2484.

Shindo H., Iwaki T., Ieda R., Kurumizaka H., Ueguchi C., Mizuno T., Morikawa S., Nakamura H. and Kuboniwa H. (1995) Solution structure of the DNA binding domain of a nucleoid-associated protein, H-NS, from *Escherichia coli*. *FEBS Lett.* **360,** 125-131.

Shindo H., Ohnuki A., Ginba H., Katoh E., Ueguchi C., Mizuno T. and Yamazaki T. (1999) Identification of the DNA binding surface of H-NS protein from *Escherichia coli* by heteronuclear NMR spectroscopy. *FEBS Lett.* **455,** 63-69.

Smyth C.P., Lundback T., Renzoni D., Siligardi G., Beavil R., Layton M., Sidebotham J.M., Hinton J.C., Driscoll P.C., Higgins C.F. and Ladbury J.E. (2000) Oligomerization of the chromatin-structuring protein H-NS. *Mol.Microbiol.* **36,** 962-972.

Sonden B. and Uhlin B.E. (1996) Coordinated and differential expression of histone-like proteins in *Escherichia coli*: regulation and function of the H-NS analog StpA. *EMBO J.* **15,** 4970-4980.

Sonnenfield J.M., Burns C.M., Higgins C.F. and Hinton J.C. (2001) The nucleoid-associated protein StpA binds curved DNA, has a greater DNA-binding affinity than H-NS and is present in significant levels in *hns* mutants. *Biochimie* **83,** 243-249.

Soutourina O., Kolb A., Krin E., Laurent-Winter C., Rimsky S., Danchin A. and Bertin P. (1999) Multiple control of flagellum biosynthesis in *Escherichia coli*: role of H-NS protein and the cyclic AMP-catabolite activator protein complex in transcription of the *flhDC* master operon. *J.Bacteriol.* **181,** 7500-7508.

Spassky A. and Buc H. (1977) Physico-chemical properties of a DNA binding protein: *Escherichia coli* factor H1. *Eur.J.Biochem.* **81,** 79-90.

Spassky A., Rimsky S., Garreau H . and Buc H. (1984) H1a, an *E.*coli DNA-binding protein which accumulates in stationary phase, strongly compacts DNA *in vitro*. *Nucleic Acids Res.* **12,** 5231-5340.

Spurio R., Durrenberger M., Falconi M., La Teana A., Pon C.L. and Gualerzi C.O. (1992) Lethal overproduction of the *Escherichia coli* nucleoid protein H-NS: ultramicroscopic and molecular autopsy. *Mol.Gen.Genet.* **231,** 201-211.

Spurio R., Falconi M., Brandi A., Pon C.L. and Gualerzi C.O. (1997) The oligomeric structure of nucleoid protein H-NS is necessary for recognition of intrinsically curved DNA and for DNA bending. *EMBO J.* **16,** 1795-1805.

Starich M.R., Sandman K., Reeve J.N. and Summers M.F. (1996) NMR structure of HmfB from the hyperthermophile, *Methanothermus fervidus*, confirms that this archeal protein is a histone. *J.Mol.Biol.* **255,** 187-203.

Stevens F.J. (1986) Analysis of protein-protein interaction by simulation of small-zone size-exclusion chromatography: Application to an antibody-antigen association. *Biochemistry* **25,** 981-993.

Stevens F.J. (1989) Analysis of protein-protein interaction by simulation of small-zone size-exclusion chromatography. *Biophys.J.* **55,** 1155-1167.

Tendeng C., Badaut C., Krin E., Gounon P., Ngo S., Danchin A., Rimsky S. and Bertin P. (2000) Isolation and characterization of *vicH*, encoding a new pleiotropic regulator in *Vibrio cholerae*. *J.Bacteriol.* **182,** 2026-2032.

Tendeng C., Krin E., Soutourina O.A., Marin A., Danchin A. and Bertin P.N. (2003) A novel H-NS-like protein from an antarctic psychrophilic bacterium reveals a crucial role for the N-terminal domain in thermal stability. *J.Biol.Chem.* **278,** 18754-18760.

Thomson J.D., Higgins D.G. and Gibson T.J. (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighing, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673-4680.

Tippner D., Afflerbach H., Bradaczek C. and Wagner R. (1994) Evidence for a regulatory function of the histone-like *Escherichia coli* protein H-NS in ribosomal RNA synthesis. *Mol.Microbiol.* **11,** 589-604.

Tippner D. and Wagner R. (1995) Fluorescence analysis of the *Escherichia coli* transcription regulator H-NS reveals two distinguishable complexes dependent on binding to specific or nonspecific DNA sites. *J.Biol.Chem.* **270,** 22243-22247.

Tupper A.E., Owen-Hughes T.A., Ussery D.W., Santos D.S., Ferguson D.J., Sidebotham J.M., Hinton J.C. and Higgins C.F. (1994) The chromatin-associated protein H-NS alters DNA topology in vitro. *EMBO J.* **13,** 258-268.

Ueguchi C. and Mizuno T. (1993) The *Escherichia coli* nucleoid protein H-NS functions directly as a transcriptional repressor. *EMBO J.* **12,** 1039-1046.

Ueguchi C., Suzuki T., Yoshida T., Tanaka K. and Mizuno T. (1996) Systematic mutational analysis revealing the functional domain organization of *Escherichia coli* nucleoid protein H-NS. *J.Mol.Biol.* **263,** 149-162.

Ueguchi C., Seto C., Suzuki T. and Mizuno T. (1997) Clarification of the dimerization domain and its functional significance for the *Escherichia coli* nucleoid protein H-NS. *J.Mol.Biol.* **274,** 145-151.

Umanski T., Rosenshine I. and Friedberg D. (2002) Thermoregulated expression of virulence genes in enteropathogenic *Escherichia coli*. *Microbiology* **148,** 2735-2744.

Ussery D.W., Hinton J.C., Jordi B.J., Granum P.E., Seirafi A., Stephen R.J., Tupper A.E., Berridge G., Sidebotham J.M. and Higgins C.F. (1994) The chromatin-associated protein H-NS. *Biochimie* **76,** 968-980.

Varshavsky A.J., Nedospasov S.A., Bakayev V.V., Bakayeva T.G. and Georgiev G.P. (1977) Histone-like proteins in the purified *Escherichia coli* deoxyribonucleoprotein. *Nucleic Acids Res.* **4,** 2725-2745.

Wagner R. (2000) Regulatory networks. In *Transcription Regulation in Prokaryotes*, pp 264-335. Oxford University Press, Oxford, UK.

Waldsich C., Grossberger A. and Schroeder R. (2002) RNA chaperone StpA loosens interactions of the tertiary structure in the td group I intron *in vivo*. *Genes Dev.* **16,** 2300-2312.

White-Ziegler C.A., Angus Hill M.L., Braaten B.A., van der Woude M.W. and Low D.A. (1998) Thermoregulation of *Escherichia coli pap* transcription: H-NS is a temperature-dependent DNA methylation blocking factor. *Mol.Microbiol.* **28,** 1121-1137.

Williams R.M., Rimsky S. and Buc H. (1996) Probing the structure, function, and interactions of the *Escherichia coli* H-NS and StpA proteins by using dominant negative derivatives. *J.Bacteriol.* **178,** 4335-4343.

Williams R.M. and Rimsky S. (1997) Molecular aspects of the *E. coli* nucleoid protein, H-NS: a central controller of gene regulatory networks. *FEMS Microbiol.Lett.* **156,** 175-185.

Winzor D.J. and Sheraga H.A. (1963) Studies of chemically reacting systems on sephadex. I. Chromatographic demonstration of the Gilbert theory. *Biochemistry* **2,** 1263-1267.

Woese C.R. (1998) Default taxonomy: Ernst Mayr's view of microbial world. *Proc.Natl.Acad.Sci.U.S.A* **95,** 11043-11046.

Woese C.R., Kandler O. and Wheelis M.L. (1990) Towards a natural system of organisms: proposal for the domains Archea, Bacteria, and Eukarya. *Proc.Natl.Acad.Sci.U.S.A* **87,** 4576-4579.

Yamada H., Muramatsu S. and Mizuno T. (1990) An *Escherichia coli* protein that preferentially binds to sharply curved DNA. *J.Biochem.(Tokyo)* **108,** 420-425.

Yamashino T., Ueguchi C. and Mizuno T. (1995) Quantitative control of the stationary phase-specific sigma factor, sigma S, in *Escherichia coli*: involvement of the nucleoid protein H-NS. *EMBO J.* **14,** 594-602.

Yarmolinsky M. (2000) Transcriptional silencing in bacteria. *Curr.Opin.Microbiol.* **3,** 138-143.

Zhang A. and Belfort M. (1992) Nucleotide sequence of a newly-identified *Escherichia coli* gene, *stpA*, encoding an H-NS-like protein. *Nucleic Acids Res.* **20,** 6735.

Zhang A., Derbyshire V., Salvo J.L. and Belfort M. (1995) *Escherichia coli* protein StpA stimulates self-splicing by promoting RNA assembly *in vitro*. *RNA.* **1,** 783-793

Zhang A., Rimsky S., Reaban M.E., Buc H. and Belfort M. (1996) *Escherichia coli* protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. *EMBO J.* **15,** 1340-1349.

Zimmerman S.B. (2003) Underlying regularity in the shapes of the nucleoids of *Escherichia coli*: Implications for nucleoid organization and partition. *J.Struct.Biol.* **142,** 256-265

Zuber F., Kotlarz D., Rimsky S. and Buc H. (1994) Modulated expression of promoters containing upstream curved DNA sequences by the *Escherichia coli* nucleoid protein H-NS. *Mol.Microbiol.* **12,** 231-240.