Extremophiles
https://doi.org/10.1007/s00792-020-01154-4

ORIGINAL PAPER

Consensus protein engineering on the thermostable histone-like bacterial protein HUs significantly improves stability and DNA binding affinity

Anastasios Georgoulis[1] · Maria Louka[1] · Stratos Mylonas[1] · Philemon Stavros[2] · George Nounesis[2] · Constantinos E. Vorgias[1]

Received: 12 April 2019 / Accepted: 6 January 2020

Abstract

Consensus-based protein engineering strategy has been applied to various proteins and it can lead to the design of proteins with enhanced biological performance. Histone-like HUs comprise a protein family with sequence variety within a highly conserved 3D-fold. HU function includes compacting and regulating bacterial DNA in a wide range of biological conditions in bacteria. To explore the possible impact of consensus-based design in the thermodynamic stability of HU proteins, the approach was applied using a dataset of sequences derived from a group of 40 mesostable, thermostable, and hyperthermostable HUs. The consensus-derived HU protein was named HU*Best*, since it is expected to perform best. The synthetic HU gene was overexpressed in *E. coli* and the recombinant protein was purified. Subsequently, HU*Best* was characterized concerning its correct folding and thermodynamic stability, as well as its ability to interact with plasmid DNA. A substantial increase in HU*Best* stability at high temperatures is observed. HU*Best* has significantly improved biological performance at ambience temperature, presenting very low K_d values for binding plasmid DNA as indicated from the Gibbs energy profile of HU*Best*. This K_d may be associated to conformational changes leading to decreased thermodynamic stability and, therefore, higher flexibility at ambient temperature.

Keywords HU histone-like protein · Consensus design · Thermostability · DNA binding activity

Abbreviations

IPTG	Isopropyl thio-β-D-galactoside
E. coli	*Escherichia coli*
SDS–PAGE	Sodium dodecyl sulphate–polyacrylamide gel electrophoresis
EMSA	Electrophoretic mobility shift assay
MSA	Multiple sequence alignment
MST	Microscale thermophoresis
Tris	Tris(hydroxymethyl)aminomethane
T_m	Protein melting temperature
DBD	DNA binding domain
HTH	Helix–turn–helix
DS	Dimerization signal

Communicated by L. Huang.

Anastasios Georgoulis and Maria Louka equally contributed to this study.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00792-020-01154-4) contains supplementary material, which is available to authorized users.

✉ Constantinos E. Vorgias
cvorgias@biol.uoa.gr

[1] Department of Biochemistry and Molecular Biology, National and Kapodistrian University of Athens, 157 01, Zografou, Greece

[2] Biomolecular Physics Laboratory, INRASTES, National Centre for Scientific Research "Demokritos", 153 10, Agia Paraskevi, Greece

Introduction

As a concept, consensus-based protein engineering relies on a semi-rational approach. It comprises a strategic design, whereupon stability-dependent as well as function-related protein properties can be altered to meet the needs and requirements of specific research and technological goals. The basic hypothesis is that at any given position in an amino acid MSA of homologous proteins, the most frequently

Published online: 24 January 2020

occurring residue contributes substantially more towards the structural integrity and the functional activity of the protein than any of the less common ones. (Lehmann et al. 2000a, b; Lehmann and Wyss 2001; Cole and Gaucher 2011). The consensus-based design is thus emerging as an appealing strategy for engineering proteins towards enhanced biological activity and increased thermal and chemical stability under a diversity of environmental conditions. So far, the application of this approach extending through the entire sequence of a protein molecule has only been carried out for a handful of successful cases with significant results. (Lehmann et al. 2002; Mosavi et al. 2001; Viader-Salvadó et al. 2010; Jochens et al. 2010; Anbar et al. 2012; Aerts et al. 2013; Silva et al. 2013; Ito et al. 2013; Wu et al. 2014; Dror et al. 2014; Magliery and Regan 2004; Larson et al. 2000). Additionally, a combination of consensus and rational design has been widely used with positive outcome in many other cases (Lehmann and Wyss 2001; Cole and Gaucher 2011). However, a major difficulty to determine consensus sequences in a protein set is the presence of highly variable regions (Nikolova et al. 1998; Wang et al. 1999).

The HU molecule is one of the most ubiquitous prokaryotic DNA binding protein in both Gram-positive and Gram-negative bacteria with global regulatory functions. The HU proteins are typically ca 20 kDa, basic and usually heat-stable heterodimers in Gram-negative, or homodimers in Gram-positive bacteria comprising essential structural elements of the bacterial nucleoids (Rouviere-Yaniv et al. 1979; Drlica and Rouviere-Yaniv 1987; Dame and Goosen 2002; Grove 2011). Briefly, the major functions of HU proteins related to DNA are: (1) DNA compaction, (2) global regulation of gene expression (Dame and Goosen 2002; Grove 2011; Berger et al. 2010), (3) involvement in DNA repair by homologous recombination (Dri et al. 1992), (4) regulation of chromosomal supercoiling by activating gyrase and by decreasing topoisomerase I activity (Drlica and Rouviere-Yaniv 1987; Malik et al. 1996), (5) regulation of the DNA replication process and support of the DnaA protein in initiating DNA replication (Roth et al. 1994; Miller and Jackson 2012) and finally (6) involvement in recA gene-dependent DNA repair and SOS induction pathways (Miyabe et al. 2000), since HUs can displace LexA from its binding sites on SOS-regulated operators (Preobrajenskaya et al. 1994).

HU interacts with the DNA minor groove in a non-specific manner inducing bending and stabilization (White et al. 1989). HU preferentially binds to four-way junctions (Pontiggia et al. 1993; Bonnefoy et al. 1994) or to particular conformations of duplex DNA (Castaing et al. 1995). HU also binds with high affinity to DNA intermediate structures during repair or recombination processes (Kamashev and Rouviere-Yaniv 2000). Interestingly, HU is also involved in osmolality/supercoiling responses and functions as a global regulator in the environmental programming of the cellular response during aerobic and acidic stress (Oberto et al. 2009).

Naturally occurring HUs attain the biologically required level of thermal stability that is imposed by the conditions of the growth environment of the bacterium of origin and especially by the range of temperature variations. Over two decades, HU proteins have been used as model molecules for thermostability studies. In most of the cases, the role of single or multiple point mutations has been explored based upon comparisons between HUs with high homology and structural identity (Wilson et al. 1990; Padas et al. 1992; Christodoulou and Vorgias 1998, 2002; Christodoulou et al. 2003; Kawamura et al. 1996, 1998; Boyko et al. 2015, 2016; Welfle et al. 1992). Several structures of HU and HU-DNA complexes have so far been determined by either X-ray crystallography or NMR spectroscopy. The first structure was described for HU from *B. stearothermophilus* (HU*Bst*) currently named *Geobacillus stearothermophilus* (Tominaga et al. 1999) and was then refined at 2 Å resolution (Tanaka et al. 1984). The solution structure of HU*Bst* has also been determined by NMR (Vis et al. 1995; Boelens et al. 1996; White et al. 1999). The crystal structures of the heterodimeric HU (HUαβ) (Guo and Adhya 2007) and homodimeric ($HU\alpha_2$) forms of HU from *E. coli* have also been solved (Coste et al. 1999; Ramstein et al. 2003), as well as the crystal structure of HU*Tmar* from the hyperthermophilic *Thermotoga maritima* (Christodoulou et al. 2003; Christodoulou and Vorgias 1998). HU-DNA co-crystal structure from *Anabaena* (Swinger et al. 2003), HU*bb*-DNA complex crystal structure from *Borrelia burgdorferi* (Mouw and Rice 2007) as well as HU from *Staphylococcus aureus* complexed with DNA (Kim et al. 2014) have been reported. Recently, in an interesting publication, HU has been suggested as a potential target for the development of therapies against tuberculosis (Bhowmick et al. 2014).

Our presented work is to show a nice application of the consensus-based protein engineering approach focused on a functionally important bacterial protein. We hope to contribute to the solution of the basic question concerning protein engineering i.e. can we design and produce novel protein molecules with properties beyond the natural limits exhibited by the members of a family, by straightforwardly applying the consensus approach?

To resolve this, we have designed a consensus HU protein based on primary sequence alignments derived from a set of 40 mesostable, thermostable, and hyperthermostable HUs and retrieved from the Uniprot. The consensus protein was named HU*Best*, the corresponding gene synthesized, the recombinant protein was produced, purified, and thoroughly studied. Since the bacterial HU protein family displays a conserved 3D-fold, a direct link between sequence conservation and stability may be anticipated. Interestingly, we find that the consensus HU exhibits significant gains in

stability at high temperatures, measured as the difference in the Gibbs free energy, compared to previously published results for other HUs. In addition, even though the DNA binding domain of HU protein is highly conserved, the consensus HU protein revealed a remarkable enhancement of the DNA binding affinity at room temperature since an unexpected lowering of K_d has been measured. The obtained results are discussed based on currently available structural and sequence data. They strongly support the notion that substitution of non-consensus by consensus amino acids is a feasible approach for enhancing stability-related protein properties and possibly other biological functions in extreme environments.

Materials and methods

Gene synthesis, cloning and protein overproduction, and purification

The amino acid primary structure of the calculated consensus was named HU*Best* protein, considering that the protein will show the "best" performance. The primary structure of HU*Best* was converted to gene, *hubest*, with simultaneous codon optimization for *E. coli* using the OptimumGene™ Codon Optimization Analysis. The synthetic gene was cloned in pUC57 and verified by sequencing (GenScript, USA), (Fig. 2, supplement). Following, the open reading frame of *hubest* was cloned into the pET-11a vector using *Nde*I and *Bam*HI cloning sites at the 5′-end and the 3′-end of the gene, respectively (Novagen). *E. coli* BL21(DE3) cells were used as expression host cells. BL21(DE3) cells harboring the pET-11a-*hubest* were grown in LB medium at 37 °C in the presence of 100 µg/ml ampicillin. The overproduction of the recombinant HU*Best* was obtained by adding 1 mM IPTG at middle log phase of bacterial culture (0.7 $OD600_{nm}$) at 37 °C for 3 h. *E. coli* cells overproducing the HU*Best* were collected by low-speed centrifugation, washed once with PBS and the cell pellet was sonicated in lysis buffer (20 mM Tris–HCl pH 7.0, 100 mM NaCl, 1 mM EDTA, 0.5% Triton X-100 and 0.5 mM phenylmethanesulfonyl fluoride, PMSF). The extractable protein fraction was obtained by centrifugation at 13,000 r.p.m. at 4 °C for 20 min. The clear supernatant was subjected to ammonium sulphate fractionation. The post 0–80% saturation in ammonium sulphate supernatant was collected, diluted ten times in buffer A (20 mM sodium phosphate, pH 8.0, 1 mM EDTA) and loaded onto a HiTrap Heparin column, previously equilibrated in buffer A. The protein was eluted with a 0–2 M linear NaCl gradient. HU*Best* was eluted between 600 and 800 mM NaCl. HU*Best* containing fractions were pooled, diluted ten times, and the pH was adjusted to 6.0 and passed through an SP Sepharose column and eluted in one step with 500 mM NaCl for protein concentration. The overall yield of a routine preparation varied between 15 and 20 mg of highly purified HU*Best* protein per 1 L bacterial culture. The purified protein was dialyzed against 10 mM Tris–HCl pH 7.0, 50 mM NaCl and stored at 4 °C.

Protein analysis by SDS–PAGE

0.1%SDS–15%PAGE analysis was run according to the Laemmli procedure (Laemmli 1970). The gels were run at constant current of 30–40 mA at room temperature and stained with Coomassie Brilliant Blue G-250.

Determination of the protein concentration

The concentration of purified HU*Best* was determined at 205 nm using the Scopes method which is suitable and very precise for small basic proteins containing no Trp or Tyr residues and cannot be measured by the Bradford method (Scopes 1974; Bradford 1976).

Protein characterization

Dynamic light scattering (DLS) and mass spectroscopy (MS) measurements were carried out at the Sample Preparation and Characterization facility at EMBL Hamburg within the framework of Biostruct-X, an EU-funded FP7 infrastructure. For the DLS measurements, the DynaPro Nanostar (Wyatt Technology Corporation) was used and the data were processed using Dynamics v.7 software. The MS measurements have been carried out using MALDI-TOF.

Circular dichroism spectroscopy

Circular dichroism (CD) measurements of HU*Best* were conducted with a JASCO J-715 spectropolarimeter (JASCO, Easton, MD), equipped with a PTC 348 WI temperature controller. Wavelength scans in the FUV region (190–260 nm) were carried out using Quartz SUPRASIL (HELMA) cells of 1-mm path length. Spectra at 25 °C were obtained by averaging five to seven successive accumulations with a wavelength step 0.2 nm at a rate 50 nm/min, a response time of 2 s and a bandwidth of 2 nm. The spectrum at 95 °C was obtained only once to prevent aggregation phenomena likely occurring after prolonged stay in the denatured state, using the above setting parameters for the spectropolarimeter. Buffer spectra, obtained at identical conditions, have been directly subtracted from the ones for sample solutions. Thermal denaturation scans (Greenfield 1996) were carried out at a heating rate of 1.5 K/min by monitoring changes of the ellipticity at 222 nm. For all CD experiments the protein concentration was 0.15 mg/mL.

High-accuracy adiabatic differential scanning calorimetry

The VP-DSC differential scanning calorimeter (MicroCal Inc., Northampton, MA, USA) was employed to explore the thermal unfolding of HU*Best*. High-accuracy measurements of the heat capacity at constant pressure ($<\Delta C_p>$) *vs.* temperature (T) were collected as T was raised continuously from room temperature to above the denaturation temperature. A protein concentration of $C_t = 2.5$ mg/mL was used while ten reference scans with only buffer-filled cells (volume of 0.523 mL) preceded each data-acquisition run to eliminate device thermal history effects and achieve maximum baseline repeatability. Protein and buffer (10 mM Tris–HCl, pH 7.0, 50 mM NaCl) solutions were thoroughly degassed under vacuum prior to loading to the calorimeter. The heating scanning rate was selected at 1.5 K/min. The reversibility of the calorimetric data was tested by performing a second consecutive heating scan and comparing the results for the total enthalpy. At the heating rate of 1.5 K/min the calorimetric reversibility, measured as the ratio of the overall enthalpy change between the two consecutive scans, was found at 87%. The temperature dependence of pre- and post-transition baselines was not substantial, so the assumption that $\Delta C_p = C_{pD} - C_{pN}$ is temperature-independent was made, where C_{pN} and C_{pD} are the heat capacity values in the native and the denature state, respectively. The calorimetric data have been analyzed via non-linear least square fitting procedures of ORIGIN 7.0 software using a two-state with dimer dissociation model, which is described in the results section. (Privalov and Potekhin 1986; Freier 1995; Biltonen and Freire 1978). The thermodynamic stability expressed as the change in the Gibbs free energy vs. temperature is calculated based on the following equation (Ruiz-Sanz et al. 2004; Steif et al. 1993):

$$\Delta G(T) = \Delta H\left(1 - \frac{T}{T_{1/2}}\right) + \Delta C_p\left(T - T_{1/2} - T\ln\frac{T}{T_{1/2}}\right) - RT_{1/2}\ln C_t.$$

Microscale thermophoresis (MST)

The method, used for the determination of binding constants, is based on the directed movement of proteins along a temperature gradient, an effect termed "thermophoresis": a locally applied temperature difference, ΔT, leads to a site-dependent change in the molecule's concentration, which can be quantified by the Soret coefficient S_T, which is the ratio of the thermal diffusion coefficient to the normal diffusion coefficient (Zhao et al. 2014) For performing experiments with the HU*Best* protein binding to DNA, a fluorescent label (NT-647) was covalently attached to the protein (NHS coupling). In a typical MST experiment, the concentration of NT-647-labeled HU*Best* was kept constant at 10 nM, while the concentration of the non-labeled pUC18 was varied between 50 nM and 1 pM. The assay was performed in buffer 20 mM sodium phosphate pH: 7.0, 1 mM EDTA, 100 mM NaCl supplemented with 0.05% Tween-20. After a short incubation, the samples were loaded into MST NT.115 hydrophilic glass capillaries and the MST analysis was performed using the Monolith NT.115. The measurement was performed in standard capillaries at 80% LED and 20% MST power. Laser on time was set at 30 s and laser off time at 5 s. For evaluating MST data, a standard routine was set in the "Thermophoresis + T-Jump" settings, according to the manufacturer instructions so that both changes in fluorescence intensity due to thermophoresis and also the temperature jump are used to determine binding constants. (Wienken et al. 2010; Zillner et al. 2012).

Bioinformatics

Determination of the consensus HU protein sequence

A set of 40 HU proteins was collected from the non-redundant protein bank Uniprot (https://www.uniprot.org). The data set is comprised from the following 40 HU proteins (Table 1 in supplement) with Uniprot accession numbers, the organisms of origin and their optimal growth temperature.

The sequences were aligned using the ClustalO (McWilliam et al. 2013) software, available on-line and the consensus sequence of HUs was determined from the most commonly occurring amino acid in each position of the MSA (Fig. 1, Supplement).

The sequences of various HUs were retrieved from the database Uniprot (https://www.uniprot.org). MSA was performed with ClustalO (https://www.ebi.ac.uk/tools/msa/clustalo) (McWilliam et al. 2013). Homology-based molecular modeling was performed using the server of WHATIF (https://swift.cmbi.ru.nl/whatif, (Vriend 1990). The comparison of the molecular interactions between the HU*Best* (model), HU from *Staphylococcus aureus* HU*Staam* (4QJU), HU from *Bacillus stearothermophilus* HU*Bst* (1HUU) and HU from *Thermotoga maritima* HU*Tmar* (1B8Z) was performed according the default parameters of the server: (https://pic.mbu.iisc.ernet.in/) (Tina et al. 2007). The computations of the surfaces were carried out with PISA (https://www.ebi.ac.uk/pdbe/pisa/) (Krissinel and Henrick 2007) and the computations of the cavities were carried out with POCASA (https://altair.sci.hokudai.ac.jp/g6/service/pocasa/) (Yu et al. 2010). The Shannon variability index H of the MSA of the 40 bacterial HU was calculated according the defaults parameters of the Protein Variability Server (PVS) (https://imed.med.ucm.es/PVS/pvs) (Shannon 1948; Kabat et al. 1977; Garcia-Boronat et al. 2008).

The homology modeling was performed in the webserver of What IF https://swift.cmbi.ru.nl/whatif (Vriend 1990; Hekkelman et al. 2010) and the model was verified using the program PROCHECK (v.3.5.4) (https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/) as well as the Protein Model Check from the WHAT IF (https://swift.cmbi.ru.nl/servers/html/index.html).

Results

Determination of the consensus HU sequence (HUBest)

As described in the experimental part, the 40 sequences for HU from mesophilic, thermophilic, and hyperthermophilic bacteria have been selected from the Uniprot. Since a crucial point to calculate the consensus sequence is the quality of the protein group, special care was taken to avoid under or overrepresentation of certain species. Our data set contains HU without internal insertions or deletions. A limited number of small extensions at the N- and C-termini of the sequences have not been omitted. Further addition of HUs in the data set in many cases introduces insertions. By deleting these insertions, the amino acid sequence of the consensus sequence is not affected. Therefore, we consider this data set of adequate quality to determine the consensus sequence of HUs.

In Fig. 1 in supplement, the total MSA and the sequence of the HU*Best* is presented. The functional domains as well as secondary structure elements of the HU are depicted in Fig. 1. It is obvious that the helix–turn–helix (HTH) domain of the HU protein molecule, which is strongly involved in the thermal adaptation of the molecule, is more variable (variability index *H*: 1.61 ± 0.76) compared to the highly conserved DNA binding bomain (DBD) (variability index *H*: 1.03 ± 0.71). The Shannon variability index *H* ranges from 0 (only one amino acid type is present at that position) to 4.322 (all 20 amino acids are equally represented at that position) (Shannon 1948). The dimerization signal (DS), responsible for the dimer formation of the HU molecules is conserved (Fig. 1; Figs. 1 and 3 in Supplement).

Figure 1 illustrates the protein variability index *H* of each amino acid according to Shannon entropy analysis to estimate the protein variability of the MSA of our HU data set (Fig. 1 supplement) (Shannon 1948). The primary structure of the consensus HU protein, named "HU*Best*" was obtained by selecting, at each position, the most frequently encountered amino acid in the 40 sequences (Figs. 2 and 3 in Supplement).

Comparing the consensus sequence and each individual sequence of the 40 HUs, the range of the sequence identity varied between 47.78% for HU from *Bifidobacterium*

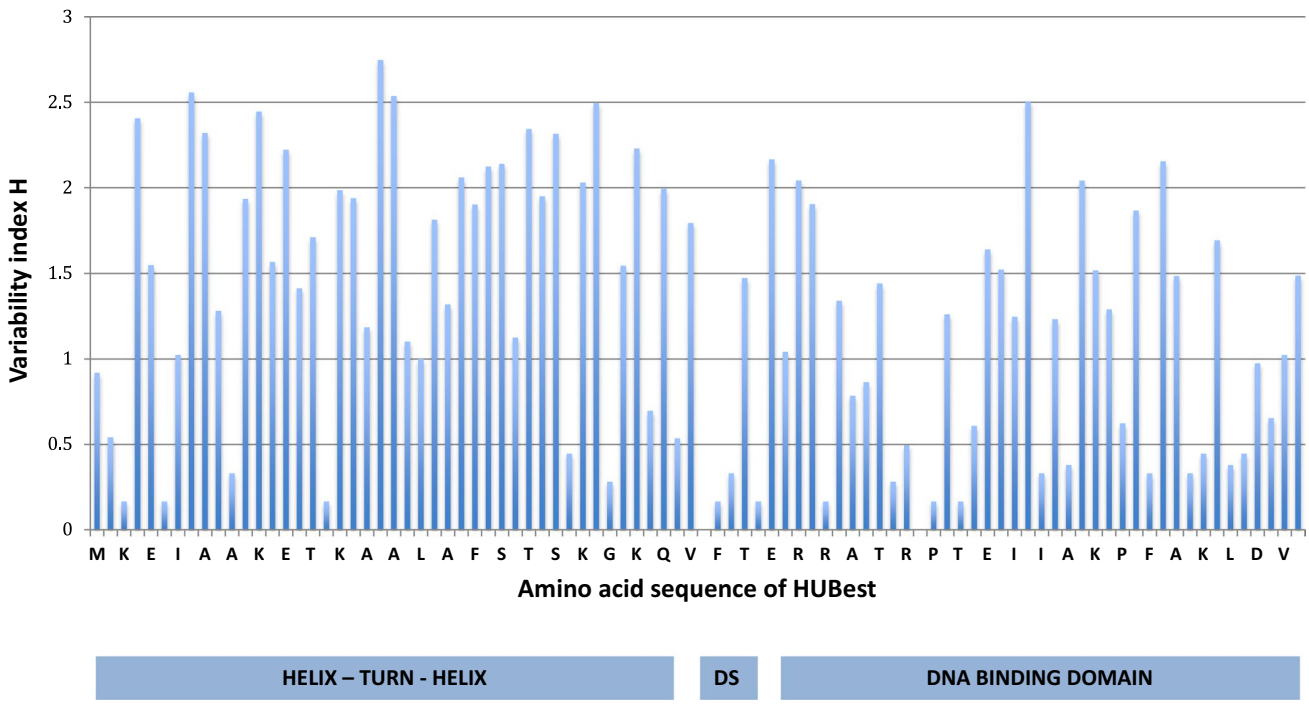

Fig. 1 Variability of the aligned HU proteins used to calculate the consensus sequence of HU. The variability index H (*y*-axis) for each sequence position (*x*-axis) was determined as described in the methods section. The functional domains as well as secondary structure elements of the HU monomer are also displayed (bottom panel)

longum (P17615) and 81.11% for HU from *Geobacillus* sp. (P0A3H0, P0A3H1, P0A3H2). The sequence identity scores of each HU in the data set against HU*Best* are summarized in Table 1 in supplement.

In Fig. 2 in supplement, the calculated consensus sequence and the corresponding codon optimized gene for *E. coli* are also presented. As expected, the sequence Prosite signature "pF00216" (GSK)-F-x(2)-(LIVMF)-x(4)-(RKEQA)-x(2)-(RST)-x(1,2)-(GA)-x-(KN)-P-x-(TN) was strictly retained in the consensus sequence. It has to be mentioned that at amino acid position 15, although the consensus residue was *E* was replaced with *G* since our studies have clearly shown that G15 plays a significant role in the thermostability of the thermostable and hyperthermostable HUs (Christodoulou and Vorgias 2002).

Cloning, overproduction and purification of HUBest

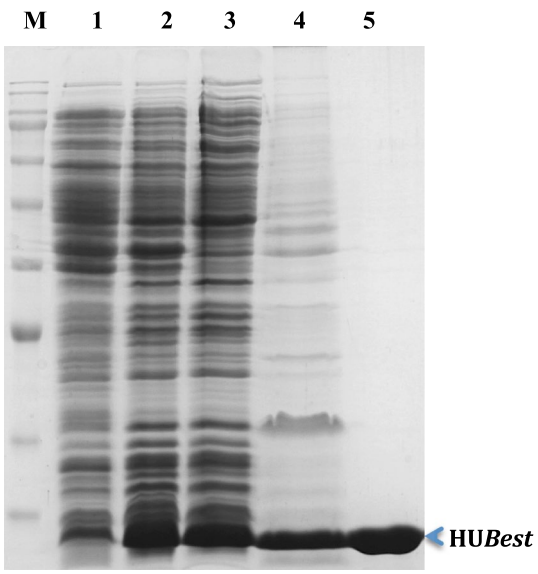

Fig. 2 The 0.1%SDS–15%PAGE analysis of the major stages during the overproduction and purification of HU*Best* in BL21(DE3) *E. coli* cells. Lane M: molecular weight marker from the top: 180, 130, 100, 75, 63, 48, 35, 28, 17, 10 kDa. Lane 1: Total extract of non-induced *E. coli* cells harbouring the pET-11a-*hubest*. Lane 2: total extract of induced *E. coli* cells harbouring the pET-11a-*hubest* with 1 mM IPTG for 3 h. Lane 3: total extractable proteins of induced *E. coli* cells. Lane 4: post 80% saturation ammonium sulphate protein fraction. Lane 5: HiTrap Heparin peak fraction of highly purified HUBest

The *hubest* gene for the HU*Best* consensus protein was synthesized, cloned, overexpressed and the recombinant protein was purifed with a very fast and efficient method producing over 15 mg of HU*Best* per L of bacterial culture within 2 days. Figure 2 presents SDS–PAGE analysis of protein samples at the major steps of the production and purification procedure.

Biochemical characterization of the recombinant HUBest

The recombinant and highly purified HU*Best* was further characterized by MALDI-TOF and the experimentally determined molecular weight was found to be 9629 Da for the monomer, which is very close to the theoretically calculated value of 9627.1 Da for the monomer, based on the amino acid sequence. A serious problem that can occur when working with small, basic proteins is that they are often aggregation-prone molecules. Using dynamic light scattering, we have precisely determined that over 98% of the HU*Best* molecular population in solution, at various concentrations and pHs, does not appear to be in any aggregated state. Moreover, during the purification procedure, neglected amounts of endogenous *E. coli* HUαβ are easily removed, since their chromatographic behavior is quite different from HU*Best* and, therefore, we consider our protein free of endogenous HU protein. This is also supported from MALDI-TOF results.

Determination of thermal stability of HUBest using CD and DSC

We have used CD and high-accuracy adiabatic DSC to investigate the thermodynamic stability of HU*Best*. The CD results are displayed in Fig. 3. As it can be seen in the top panel of Fig. 3, upon heating to 95 °C, the far-UV CD spectrum of HU*Best* is dramatically altered. A CDNN (Bohm et al. 1992) analysis of the spectra reveals a loss of up to 77% in the helical content and gains in random coil of approximately 5%. These drastic changes are analogous in magnitude to previous findings for HU*Tvo* where a 60% loss in the helical content of the thermally denatured state was recorded (Orfaniotou et al. 2009). Similar drastic changes in the far-UV CD spectra have also been recorded for HU*Tmar* (Ruiz-Sanz et al. 2004). Remarkably, upon cooling back to room temperature, the native spectrum is fully recovered, revealing a reversible thermal unfolding for HU*Best*. Once again the observed reversibility of the thermal unfolding is quite analogous to HU*Tvo* and HU*Tmar,* the latter though exhibited reversible thermal behavior only at low pH (pH < 4). The thermal CD data presented in Fig. 3 (bottom panel) show the changes in the ellipticity at $\lambda = 222$ nm for the sample with concentration $C_t = 13$ μM. A "melting" temperature T_m of 72.6 ± 0.2 °C can be estimated using a two-state model. This thermal denaturation temperature is considerably higher than what was found for HU*Tvo* at pH 7.0 (Orfaniotou et al. 2009) and is comparable

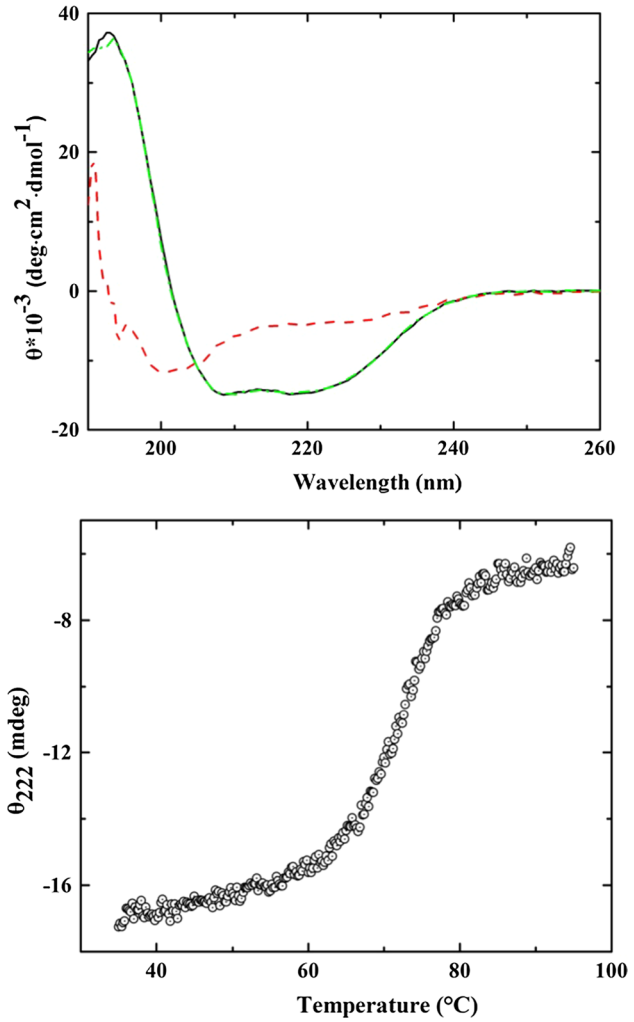

Fig. 3 CD spectra for the heat-induced denaturation of HU*Best*. Top panel, far-UV spectra at room temperature (solid black line), above the unfolding transition, at 95 °C (red dashed line) and after cooling to room temperature from the denatured state (green dot-dashed line), demonstrating an excellent reproducibility of the thermal unfolding process. Bottom panel, thermal CD data for the unfolding transition recording the ellipticity at $\lambda = 222$ nm. Based on a two-state model T_m is estimated at 72.6 ± 0.2 °C

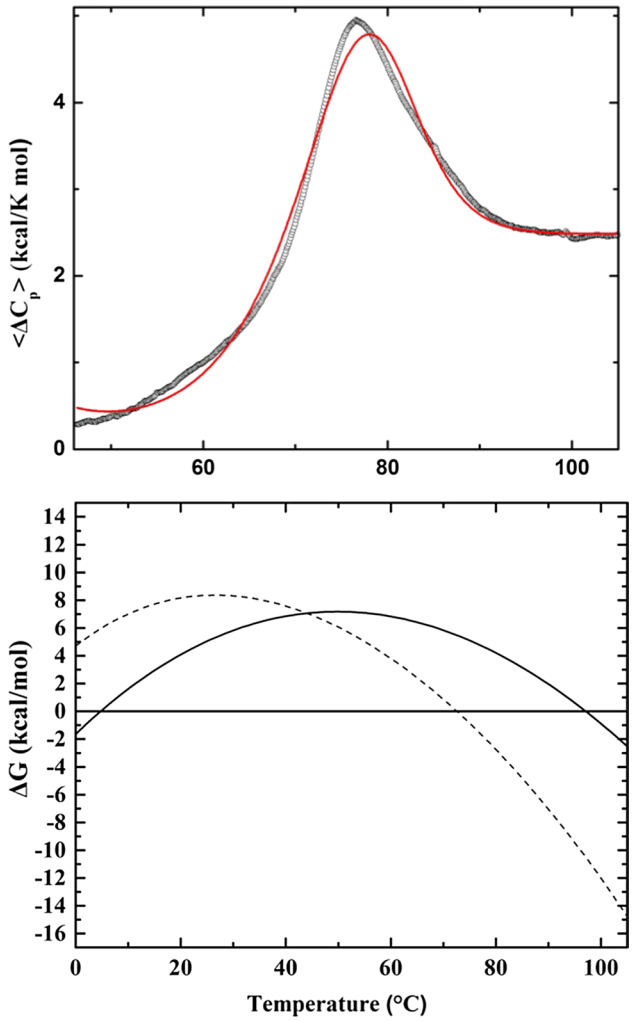

Fig. 4 High-accuracy DSC data for the thermal unfolding of HU*Best*. Top panel, normalized DSC $<\Delta C_p>$ vs. T data after subtraction of regular instrumentals baseline. The solid red line represents the fir of single calorimetric peak to a two-state model with dimer dissociation, as described in the text. Bottom panel: the thermodynamic stability ΔG vs T for HU*Best* as derived based on the $N_2 \leftrightarrow 2U$ model. For comparison, also shown in the same diagram are the results obtained for HU*Tvo* (dashed line)

to HU*Tmar* at pH 4.0 (Ruiz-Sanz et al. 2004), for analogous concentrations.

The reversibility of the thermal transition of HU*Best* provides a good assurance for applying an equilibrium thermodynamic analysis, based on calorimetric data. Thus, we have performed high-accuracy adiabatic DSC measurements of HU*Best* at a concentration $C_t = 0.26$ mM. A characteristic $<\Delta C_p>$ vs. T profile for a heating scan at a rate of 1.5 K/min is presented in Fig. 4. The endothermic calorimetric peak is characterized by near-perfect (> 87%) reversibility between consecutive heating scans and shows no dependence upon the applied heating rate, excluding thus "slow" kinetic effects affecting the thermal transition. A straightforward analysis of the calorimetric peak leads to $T_m = 78.0 \pm 0.2$ °C and enthalpy change $(\Delta H)_{cal} = 54 \pm 6$ kcal/mol.

Compared to previous results for HU*Tmar* and HU*Tvo*, HU*Best* appears to be thermostable, characterized as all other HUs by a small magnitude of $(\Delta H)_{cal}$. The observed difference in T_m as function of C_t, between the CD and the DSC measurements, is indicative of the fact that the native state of HU*Best* is dimeric, as was the case of HU*Tvo* and HU*Tmar*. Calorimetric studies of *E. coli* HUβ_2 homodimer protein have revealed a complex double-peaked heat capacity profile that could only be analyzed via a three-state model with dissociation: $D_2 \leftrightarrow I_2 \leftrightarrow 2U$. Here I_2 stands for a stable intermediate dimeric state whose structure has been solved

by NMR experiments (Garnier et al. 2011). A similar denaturation process has recently been published for HU from *Spriroplasma melliferum* KC3 (Boyko et al. 2016). Since our calorimetric data present absolutely no evidence for two calorimetric anomalies, we have proceeded by analyzing them using a two-state model with dimer dissociation: $N_2 \leftrightarrow 2U$ (Ruiz-Sanz et al. 2004; Privalov and Potekhin 1986; Steif et al. 1993).

The fitting of the calorimetric data is shown as a solid line in the top panel of Fig. 4. No chemical baseline has been subtracted allowing for the parameters do be determined by the fitted two-state model. The quality of the fit is relatively poor. HU*Best* is characterized by a relatively low per-residue change of the calorimetric enthalpy. Indeed, $(\Delta H)_{cal}$ is calculated to be 0.60 ± 0.07 kcal/mol-residue, which is smaller than the average value of 0.77 kcal/mol-residue expected for small globular proteins (Benitez-Cardoza et al. 2001). On the other hand, ΔC_p is found to be 24.1 ± 2.4 cal/K mol-residue, which is larger than typical bibliographical values of 12–18 cal/K mol-residue (Becktel and Schellman 1987). This value for ΔC_p is similar to what was found for H*Tvo* but it is higher than for HU*Tmar*, at pH = 4 (Orfaniotou et al. 2009; Ruiz-Sanz et al. 2004). It has been argued that amino acid residues at the flexible, already exposed to the solvent, DNA-binding regions of HU*Best* have only small contribution to $(\Delta H)_{cal}$. Indeed 35 of the 90 residues of HU*Best* are located in these flexible regions. Considering their contribution as "minimal", the per-residue value of $(\Delta H)_{cal}$ can be recalculated to 0.98 k ± 0.07 cal/mol-residue, which is larger than the bibliographic average. This is indicative of the fact that a considerable number of amino acids at the flexible regions of HU*Best* do contribute to the enthalpic content of the thermal transition.

Here $T_{1/2}$ is the concentration-dependent temperature where the molar fractions of the native dimer are equal to the molar fraction of the unfolded monomers. The results for the temperature dependence of ΔG are plotted in Fig. 4 bottom panel. Comparing the present results with the ones obtained for HU*Tvo*, also plotted in Fig. 4, (although the two refer to different protein concentrations, the outcome is not substantially altered), it becomes evident that while both molecules are characterized by similar maximum stability, remarkably, in the case of HU*Best* the $\Delta G(T)$ plot is shifted to the right. Characteristically, the temperature T_g at which $\Delta G = 0$ shifts upwards in temperature from 72.15 to 100.27 °C. Compared to HU*Tvo*, HU*Best* gains stability at high temperatures while losing stability at room temperature. The published results for HU*Tmar* also refer to a maximum ΔG analogous to HU*Best* although measured at pH 4.0. A T_g of ~ 100 °C is reported in this case (Christodoulou et al. 2003). In addition, an already published thermal analysis for HU*Bst* has yielded a T_m of ~ 62 °C; significantly lower than the values reported here (Christodoulou and Vorgias 2002).

Interaction of HUBest with plasmid DNA

Concerning the DNA binding properties of the HU histone-like proteins it has become clear that, although the HU-DNA interaction is not sequence specific, it is nevertheless dependent upon the length of DNA (Pontiggia et al. 1993; Bonnefoy et al. 1994; Castaing et al. 1995; Kamashev et al. Kamashev and Rouviere-Yaniv 2000). At the molecular level, the strength of the interaction is not determined solely by contribution from the amino acids that create the initial positively charged electrostatic scaffold. Instead, subtle variations in the flexibility and mobility of adjacent amino acids may play an important role in the binding affinity to DNA. Crystallographic evidence in combination with studies involving point mutations and precise K_d determination are required to support this idea. Usually the K_d values for binding of various HU proteins to DNA have been determined via EMSA using synthetic oligonucleotides.

In this work, we studied, for the first time, the interaction of HU*Best* with pUC18 plasmid DNA, in solution, by employing MST as described in the experimental procedures section. A typical dissociation plot for HU*Best* is presented in Fig. 5. A K_d value of 1.07 nM ± 0.05 nM was calculated for HU*Best*, this signifies a high DNA binding affinity. We have also measured the K_d of three additional HU proteins (HU*Bst*) from *B. stearothermophilus*, HU*Dr* from *Deinococcus radiodurans* and HU*Tth* from *Thermus thermophilus* with pUC18 under identical conditions. The K_d for HU*Bst* was determined to be 1100.0 ± 149.0 nM, for HU*Dr* 3.6 ± 0.5 nM and for HU*Tth* 1100.0 ± 150 nM. It has to be pointed out that the stoichiometry of the HU with DNA cannot be precisely defined since HU can form dimer, tetramers etc. depends on the concentration (Sagi et al. 2004). The thermal effect during the thermophoresis in combination with the aggregation could explain the variations in standard deviation.

By reviewing published results, it is becoming increasingly apparent that HU orthologs encoded by different eubacteria exhibit significantly different DNA-binding properties in vitro, depending on the length of bound DNA. Briefly, for non-specific binding of duplex DNA, the DNA length range between ~ 9 bp for *E. coli* HU and 12 bp for *B. subtilis* HU to 17–19 bp for HU from *Anabaena* and *Helicobacter pylori* and > 35 bp for HU*Tmar* from *T. maritima* and HU*Dr* from *Deinococcus radiodurans*. DNA binding affinities range between K_d ~ 5 nM for HU*Tmar*, > 200 nM for *E. coli* HU, and 350 nM for HU*Staam* from *Staphylococcus aureus* (Grove 2011, p. 30, Bonnefoy et al. 1994; Castaing et al. 1995; Broyles and Pettijohn 1986; Lavoie et al. 1996; Li and Waters 1998; Azam and Ishihama 1999; Kobryn et al. 1999; Grove et al. 1996; Esser et al. 1999; Chen et al. 2004; Ghosh and Grove 2004; Kamau et al. 2005; Koh et al. 2008; Kim et al. 2014). Noteworthy at this point is also the

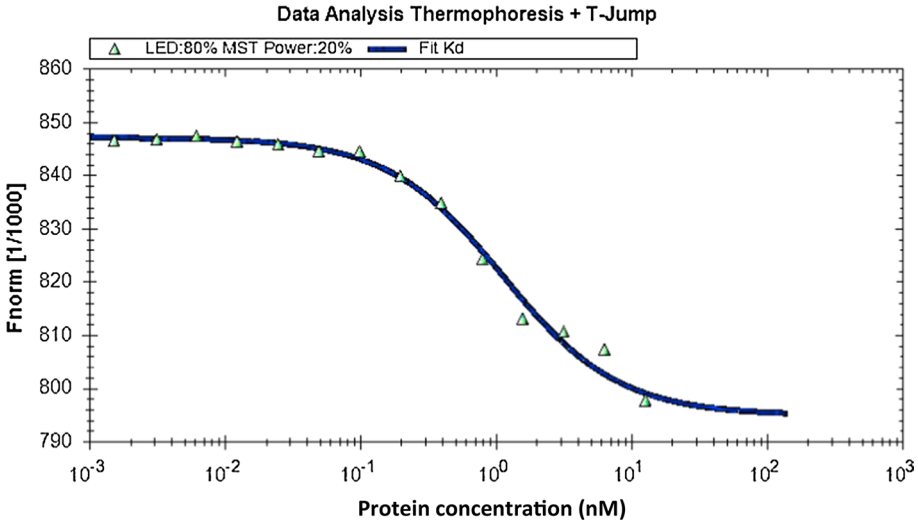

Fig. 5 Determination of binding constant K_d of HU*Best* to plasmid pUC18 by MST as described in "Materials and methods". Binding of NT-647-labeled HU*Best* protein to pUC18. The concentration of the fluorescently labelled HU*Best* was kept constant, while the concentration of the pUC18 was varied from 50 nM–1 pM. After a 10 min equilibration, MST analysis was performed ($n = 3$). Concentrations on the *x*-axis are plotted in nM. A K_d of 1.07 nM ± 0.05 nM was determined for this interaction as described in "Materials and methods"

non-specific relatively high affinity value ($K_d = \sim 2$ nM) of HU for binding to distorted substrates reported by Kamashev and Rouviere-Yaniv (2000). Long DNA strands, like plasmids, have not been used in any of the above-mentioned measurements. However, our measurement may reflect the natural conditions for the HU proteins, since it is known that DNA structure (like negative supercoiling) affect the DNA binding properties. Therefore we make a simple comparison and not a general statement.

Previously it was reported that the flexibility of DNA binding domain of HU dictates both binding and recognition in *Staphylococcus aureus* (HU*Staam*). It has been shown that arginine residues are particularly important for the flexibility of the β-arms directly affecting the binding ability of HU to DNA and thus the function of the protein. Especially, residues Arg55 and Arg61 have been discussed to be the most energetically stable residues in the HU-DNA complex. Significant residues for DNA binding are also Arg58 and Lys59. By mutating each of them to alanine, the K_d of 370 nM was increased up to 1.2 and 5 times, respectively compared to the wild type (Kim et al. 2014). The DBD of HU*Best* is nearly identical to the DBD of the HU*Staam* and since the X-ray structure of HU*Staam*-DNA complex (4QJU) has been solved (Kim et al. 2014), some modeling-based comparisons are possible. A sequence alignment of the HU*Best* with HU*Staam* (Fig. 3 in supplement) reveals that the binding site of the HU*Best* completely retains the electrostatic properties from HU*Staam*, since none of the core amino acids responsible for interaction with DNA has been changed. The only sequence changes are encountered at positions: T59, E67, and K70 in HU*Best* and correspond to amino acid residues K59, K67, and D70 in HU*Staam*. These amino acids are located at the β_3, β_4-sheets, at the tip of the flexible arm (Kim et al. 2014) and as mentioned earlier, only Lys59 has been shown to play a significant role in DNA binding. In the case of the Lys (HU*Staam*) to Glu (HU*Best*) change at position 67, located on the linker between β_3 and β_4 sheets, either of the two amino acids is facing towards the solvent and is, therefore, unlikely to directly interact with DNA. The change in the charge may nevertheless influence the flexibility of the β-arm locally. Finally, the D70 in HU*Staam*, changed to K70 in HU*Best*, located at the end of the β_4 sheet and faced to the solvent is not expected to influence the HU-DNA interaction either. Away from DBD, Lys3, the highly conserved residue among HU homologs, essential for determining the length of the binding DNA (Grove and Saavedra 2002) is also present in HU*Best*.

Homology modeling and structural comparison of HUBest with HUBst, HUTmar, and HUStaam

The model structure of HU*Best* was calculated using homology modeling. Two alternatives could be selected as templates. Either the X-ray structure of HU*Bst* (1HUU) with 81.11% or of that of HU*Staam* bound to DNA (4QJU) with 78.89% sequence identity to HU*Best*, respectively. We have chosen to use the 4QJU structure as a template in the homology modeling since the HU molecule bound to DNA is fully resolved (Kim et al. 2014) while in the structure of HU*Bst*, the flexible arms are not visible.

The model structure of HU*Best* is presented in Fig. 6. Figure 6a presents the secondary structure elements of the monomer and Fig. 6b presents the variability index on the structure.

In an attempt to correlate the thermostability differences with structural elements at the level of intra- and inter-polypeptide chain interactions, we have performed a comprehensive comparison between the structures of HU*Bst* from *B. stearothermophilus* (1HUE), HU*Staam* from *S. aureus* (4QJU), HU*Tmar* from *T. maritima* (1B8Z)

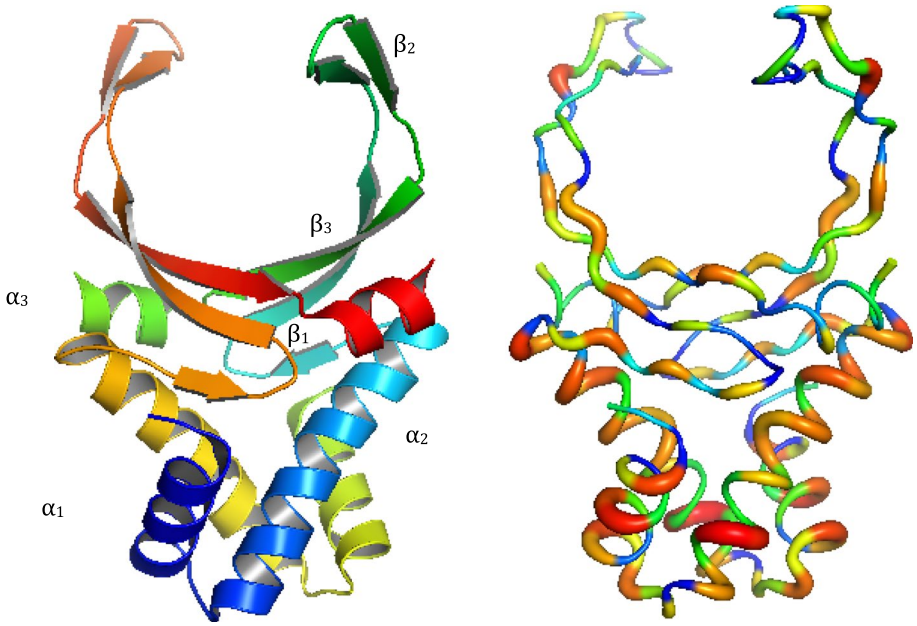

Fig. 6 Homology-based model structure of HU*Best.* The model structure of HU*Best* calculated using homology modeling as described in "Materials and methods". **a** The secondary structure elements and **b** the variability index *H* mapped on the structure (thick red lines high *H*, thin blue lines low *H*)

and the model structure of HU*Best* using the freely available web servers (https://swift.cmbi.ru.nl/servers/html/index.html) and https://pic.mbu.iisc.ernet.in/.

The 2 HU proteins were selected to be compared with HU*Best* because (a) HU*Bst* ($T_m = 61.6$ °C) has the highest identity to HU*Best,* 81.11%, and (b) HU*Tmar* is HU with the highest $T_m = 80.5$ °C and 63.33% identity.

The obtained results are presented in Table 2 in supplement and several interesting observations can be drawn at the level of intra- and inter-subunit interactions. It appears that HU*Best* is characterized by an increased number of hydrophobic interactions for intra-subunit as well as inter-monomer interactions. On the other hand, it is also characterized by a significantly low number of intra-subunit interactions concerning side-chain—side-chain hydrogen bonds. Finally, the number of side-chain—main chain hydrogen bonds appears to be on the elevated side for intra-subunit interactions. According to previous studies using MD simulations and NMR experiments suggest that the volume of the dimer interfacial cavity constitutes another parameter to be considering for intrinsic stability (Garnier et al. 2011). We have tested this hypothesis in few HUs with known T_m and compared with the model structure of HU*Best*. The *E. coli* HU was not included since we consider it as a particular case, due to formation of heterodimer. The results in Table 2 in supplement, show indeed that in all cases the interfacial cavity volume is zero (even a single water molecule does not fit) due to the nature of the highly hydrophobic DS, strongly suggesting that minimum interfacial cavity has an effect on the thermostabilization of the dimeric HUs.

The results indicate that using a consensus primary structure, a new balanced rearrangement is established between various intra- and inter-molecular interactions leading to enhance stability at high temperatures.

Discussion

HU proteins belong to a large family of histone-like DNA binding proteins in prokaryotes and show sequence diversity, highly conserved fold, non-specific DNA binding activity, folding-refolding reversibility in various environments and thermal adaptability. The stability of HUs at high temperatures is solely based on primary structure differences among the various bacteria, grown at various temperatures. Currently a considerable number of HU genes and their corresponding proteins have been deposited in various data banks. On the structural level over 50 structures of free or bound to DNA HU proteins are appearing in the PDB bank.

At the amino acid sequence level HU proteins display a significant diversity. This diversity is greater at the HTH domain while the DS and DBD are by far highly conserved, as presented in the results section. The DS is a hydrophobic pocket created by (G)-F-x(2)-(F) from each monomer and it is highly conserved, since it is responsible for the fast homo-dimerization of the HU monomers (Christodoulou and Vorgias 2002; Christodoulou et al. 2003; Vis et al. 1995, 1998). The DBD consists of basic amino acids and forms a flexible arm wrapping the DNA upon binding (White et al. 1989, p. 48, Swinger et al. 2003; Kim et al. 2014). It would thus appear that, stability issues are mainly originating from the HTH domain and its interactions upon homo-dimerization.

ΔC_p, the heat capacity difference between the thermally denatured and the native states has been proposed to be an important parameter in thermostability. The magnitude of ΔC_p directly affects the curvature of the $\Delta G(T)$ curve. The smaller the ΔC_p, the wider the $\Delta G(T)$ curve for a given ΔG maximum and thus the higher the melting point T_g (Kumar et al. 2003; Alexander et al. 1992). The small size of ΔC_p has been associated to enriched polar interactions (Zhou 2002; Murphy and Freire 1992) and thus changes in the polar and apolar solvent accessible surface areas, directly contributing to ΔG (Hilser et al. 1997). Comparing the present results for ΔC_p with previously published data though, no such conclusions can be deduced since ΔC_p for HU*Best* and HU*Tvo* are characterized by similar ΔC_p, yet HU*Best* is significantly more thermostable. Analogously, HU*Best* has comparable thermostability to HU*Tmar* even though it exhibits a larger ΔC_p. Hilser et al. (1997) have also proposed that the interatomic differences between different atom types, which directly determine packing density, may be a crucial parameter in thermostability. It can thus be hypothesized that the observed ΔG changes at high temperatures between HU*Best* and other thermophilic HUs are likely originating from the differences in the compactness between the native and the structure-less denatured states of each molecule. Another interesting finding of the present study is the high affinity for binding to DNA that is exhibited by HU*Best* compared to other HUs reported in the literature. Of course, the current data refer to long DNA molecules compared to the previously published results. Nevertheless this observed tendency might be directly related to the stability decrease recorded for HU*Best* at room temperature, also supported from the Gibbs energy curve (Figs. 3, 4). Binding to DNA will likely result to stable complexes thus enhancing the strength of the interaction.

In this work, we have focused on designing the consensus sequence of HU*Best* protein, which was produced and purified for further biochemical, biophysical and structural studies to elucidate structure function relationships in HU proteins. The experimental data strongly suggest that the information for correct protein structure and function is predominantly encoded positionally, at the level of consensus, and second in higher-orders correlations. This work can be considered as a typical example of "entire" molecule consensus design supporting the "dogma" that each consensus amino acid of a protein family plays a particular important role in the biological and the structural integrity for each of the family members.

In conclusion, for bacterial HU proteins sharing a common fold, a direct link between protein sequence conservation and protein stability has been demonstrated. Interestingly, the consensus sequence based on the 40 bacterial HU sequences results in a significant elevation of the unfolding temperature of the consensus HU*Best* and an unexpected lowering of K_d for plasmid DNA. Both these findings can be understood by the observed properties of HU*Best* where stability is elevated at high temperatures and lowered at room temperature. Mechanisms that are likely originating from changes in the packing density and thus the compactness of the native and the denatured states may be responsible for these observations.

Closing, it is widely accepted that protein folding and stability is one of the most difficult biological questions. Large scale statistical analyses did not provide certain rules except than indicating some preferences for certain amino acids. We suggest that an alternative way to get closer to a possible solution to the problem is to compare small proteins with experimentally determined stability and focus on distinct domains trying to extract information in combination with evolutionary data.

Acknowledgements We acknowledge technical support by the SPC facility at EMBL Hamburg in the frame of Biostruct-X (EE FP7). We would also like to acknowledge technical assistance from A. Tsoka and N Papandreou for biocomputing in our department. F. S. has been supported in part by the ARISTEIA I program (Grant Number 1125), administered by the General Secretariat of Research and Technology of Greece, co-financed by the European Social Fund and the State of Greece. The financial support of "The National Research Infrastructures on Integrated Structural Biology, Drug Screening Efforts and Drug target functional characterization" with the acronym "INSPIRED" and code (MIS) 5002550 from the Greek Government is also acknowledged.

References

Aerts D, Verhaeghe T, Joosten HJ, Vriend G, Soetaert W, Desmet T (2013) Consensus engineering of sucrose phosphorylase, the outcome reflects the sequence input. Biotechnol Bioeng 110:2563–2572. https://doi.org/10.1002/bit.24940

Alexander P, Fahnestock S, Lee T, Orban J, Bryan P (1992) Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2: why small proteins tend to have high denaturation temperatures. Biochemistry 31:3597–3603

Anbar M, Gul O, Lamed R, Sezerman UO, Bayer EA (2012) Improved thermostability of Clostridium thermocellum endoglucanase Cel8A by using consensus-guided mutagenesis. Appl Environ Microbiol 78:3458–3464. https://doi.org/10.1128/AEM.07985-11

Azam TA, Ishihama A (1999) Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. J Biol Chem 274:33105–33113

Becktel WJ, Schellman JA (1987) Protein stability curves. Biopolymers 26:1859–1877

Benitez-Cardoza CG, Rojo-Dominguez A, Hernadez-Arana A (2001) Temperature-induced denaturation and renaturation of triosephosphate isomerase from *Saccharomyces cerevisiae*: evidence of dimerization coupled to refolding of the thermally unfolded protein. Biochemistry 40:9049–9058

Berger M, Farcas A, Geertz M, Zhelyazkova P, Brix K, Travers A, Muskhelishvili G (2010) Coordination of genomic structure and transcription by the main bacterial nucleoid-associated

protein HU. EMBO Rep 11:59–64. https://doi.org/10.1038/embor.2009.232
Bhowmick T, Ghosh S, Dixit K, Ganesan V, Ramagopal UA, Dey D, Sarma PS, Ramakumar S, Nagarajia V (2014) Targeting *Mycobacterium tuberculosis* nucleoid-associated protein HU with structure-based inhibitors. Nat Commun 5:4124. https://doi.org/10.1038/ncomms5124
Biltonen RL, Freire E (1978) Thermodynamic characterization of conformational states of biological macromolecules using differential scanning calorimetry. CRC Crit Rev Biochem 5:85–124
Boelens R, Vis H, Vorgias CE, Wilson KS, Kaptein R (1996) Structure and dynamics of the DNA binding protein HU from *Bacillus stearothermophilus* by NMR spectroscopy. Biopolymers 40:553–559
Bohm G, Muhr R, Jaenicke R (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks. Protein Eng 5:191–195
Bonnefoy E, Takahashi M, Yaniv JR (1994) DNA-binding parameters of the HU protein of *Escherichia coli* to cruciform DNA. J Mol Biol 242:116–129. https://doi.org/10.1006/jmbi.1994.1563
Boyko KM, Gorbacheva M, Rakitina T, Korzhenevskiy D, Vanyushkina A, Kamashev D, Lipkin A, Popov V (2015) Expression, purification, crystallization and preliminary X-ray crystallographic analysis of the histone-like HU protein from *Spiroplasma melliferum* KC3. Acta Crystallogr F Struct Biol Commun 71:24–27. https://doi.org/10.1107/S2053230X14025333
Boyko KM, Rakitina TV, Korzhenevskiy DA, Vlaskina AV, Agapova YK, Kamashev DE, Kleymenov SY, Popov VO (2016) Structural basis of the high thermal stability of the histone-like HU protein from the mollicute *Spiroplasma melliferum* KC3. Sci Rep 6:36366. https://doi.org/10.1038/srep36366
Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254
Broyles SS, Pettijohn DE (1986) Interaction of the *Escherichia coli* HU protein with DNA. Evidence for formation of nucleosome-like structures with altered DNA helical pitch. J Mol Biol 87:47–60
Castaing B, Zelwer C, Laval J, Boiteux S (1995) HU protein of *Escherichia coli* binds specifically to DNA that contains single-strand breaks or gaps. J Biol Chem 270:10291–10296
Chen C, Ghosh S, Grove A (2004) Substrate specificity of helicobacter pylori histone-like HU protein is determined by insufficient stabilization of DNA flexure points. Biochem J 383:343–351. https://doi.org/10.1042/BJ20040938
Christodoulou E, Vorgias CE (1998) Cloning, overproduction, purification and crystallization of the DNA binding protein HU from the hyperthermophilic eubacterium *Thermotoga maritima*. Acta Crystallogr D Biol Crystallogr 54:1043–1045
Christodoulou E, Vorgias CE (2002) The thermostability of DNA-binding protein HU from mesophilic, thermophilic, and extreme thermophilic bacteria. Extremophiles 6:21–31
Christodoulou E, Rypniewski WR, Vorgias CE (2003) High-resolution X-ray structure of the DNA-binding protein HU from the hyperthermophilic *Thermotoga maritima* and the determinants of its thermostability. Extremophiles 7:111–122. https://doi.org/10.1007/s00792-002-0302-7
Cole MF, Gaucher EA (2011) Utilizing natural diversity to evolve protein function: applications towards thermostability. Curr Opin Chem Biol 15:399–406. https://doi.org/10.1016/j.cbpa.2011.03.005
Coste F, Hervouet N, Oberto J, Zelwer C, Castaing B (1999) Crystallization and preliminary X-ray diffraction analysis of the homodimeric form alpha2 of the HU protein from *Escherichia coli*. Acta Crystallogr D Biol Crystallogr 55:1952–1954
Dame RT, Goosen N (2002) HU, promoting or counteracting DNA compaction? FEBS Lett 529:151–156
Dri AM, Moreau PL, Rouvière-Yaniv J (1992) Role of the histone-like proteins OsmZ and HU in homologous recombination. Gene 120:11–16
Drlica K, Rouviere-Yaniv J (1987) Histonelike proteins of bacteria. Microbiol Rev 51:301–319
Dror A, Shemesh E, Dayan N, Fishman A (2014) Protein engineering by random mutagenesis and structure-guided consensus of *Geobacillus stearothermophilus* Lipase T6 for enhanced stability in methanol. Appl Environ Microbiol 80:1515–1527. https://doi.org/10.1128/AEM.03371-13
Esser D, Rudolph R, Jaenicke R, Bohm G (1999) The HU protein from *Thermotoga maritima*: recombinant expression, purification and physicochemical characterization of an extremely hyperthermophilic DNA-binding protein. J Mol Biol 291:1135–1146. https://doi.org/10.1006/jmbi.1999.3022
Freire E (1995) Differential scanning calorimetry. Methods Mol Biol 40:191–218. https://doi.org/10.1385/0-89603-301-5:191
Garcia-Boronat M, Diez-Rivero CM, Reinherz EL, Reche PA (2008) PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. Nucleic Acids Res 36:W35–41. https://doi.org/10.1093/nar/gkn211
Garnier N, Loth K, Coste F, Augustyniak R, Nadan V, Damblon C, Castaing B (2011) An alternative flexible conformation of the *E. coli* HUβ_2 protein: structural, dynamics, and functional aspects. Eur Biophys J 40:117–129. https://doi.org/10.1007/s00249-010-0630-y
Ghosh S, Grove A (2004) Histone-like protein HU from *Deinococcus radiodurans* binds preferentially to four-way DNA junctions. J Mol Biol 337:561–571. https://doi.org/10.1016/j.jmb.2004.02.010
Greenfield NJ (1996) Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. Anal Biochem 235:1–10. https://doi.org/10.1006/abio.1996.0084
Grove A (2011) Functional evolution of bacterial histone-like HU proteins. Curr Issues Mol Biol 13:1–12
Grove A, Saavedra TC (2002) The role of surface-exposed lysines in wrapping DNA about the bacterial histone-like protein HU. Biochemistry 41:7597–7603
Grove A, Galeone A, Mayol L, Geiduschek EP (1996) Localized DNA flexibility contributes to target site selection by DNA-bending proteins. J Mol Biol 260:120–125
Guo F, Adhya S (2007) Spiral structure of *Escherichia coli* HUalpha-beta provides foundation for DNA supercoiling. Proc Natl Acad Sci USA 13:4309–4314. https://doi.org/10.1073/pnas.0611686104
Hekkelman ML, Te Beek TA, Pettifer SR, Thorne D, Attwood TK, Vriend G (2010) WIWS: a protein structure bioinformatics. Nucleic Acids Res (Web Server Issue). https://doi.org/10.1093/nar/gkq453
Hilser VJ, Townsend BD, Freire E (1997) Structure-based statistical thermodynamic analysis of T4 lysozyme mutants: structural mapping of cooperative interactions. Biophys Chem 64:69–79
Ito Y, Ikeuchi A, Imamura C (2013) Advanced evolutionary molecular engineering to produce thermostable cellulase by using a small but efficient library. Protein Eng Des Sel 26:73–79. https://doi.org/10.1093/protein/gzs072
Jochens H, Aerts D, Bornscheuer UT (2010) Thermostabilization of an esterase by alignment-guided focussed directed evolution. Protein Eng Des Sel 23:903–909. https://doi.org/10.1093/protein/gzq071
Kabat EA, Wu TT, Bilofsky H (1977) Unusual distributions of amino acids in complementarity-determining (hypervariable) segments of heavy and light chains of immunoglobulins and their

possible roles in specificity of antibody-combining sites. J Biol Chem 252:6609–6616
Kamashev D, Rouviere-Yaniv J (2000) The histone-like protein HU binds specifically to DNA recombination and repair intermediates. EMBO J 19:6527–6535. https://doi.org/10.1093/emboj/19.23.6527
Kamau E, Tsihlis ND, Simmons LA, Grove A (2005) Surface salt bridges modulate the DNA site size of bacterial histone-like HU proteins. Biochem J 390:49–55. https://doi.org/10.1042/BJ20050274
Kawamura S, Kakuta Y, Tanaka I, Hikichi K, Kuhara S, Yamasaki N, Kimura M (1996) Glycine-15 in the bend between two alpha-helices can explain the thermostability of DNA binding protein HU from *Bacillus stearothermophilus*. Biochemistry 35:1195–1200. https://doi.org/10.1021/bi9515811
Kawamura S, Abe Y, Ueda T, Masumoto K, Imoto T, Yamasaki N, Kimura M (1998) Investigation of the structural basis for thermostability of DNA-binding protein HU from *Bacillus stearothermophilus*. J Biol Chem 273:19982–19987. https://doi.org/10.1074/jbc.273.32.19982
Kim DH, Im H, Jee JG, Jang SB, Yoon HJ, Kwon AR, Kang SM, Lee BJ (2014) β-Arm flexibility of HU from *Staphylococcus aureus* dictates the DNA-binding and recognition mechanism. Acta Crystallogr D Biol Crystallogr 70:3273–3289. https://doi.org/10.1107/S1399004714023931
Kobryn K, Lavoie BD, Chaconas G (1999) Supercoiling-dependent site-specific binding of HU to naked Mu DNA. J Mol Biol 289:777–784. https://doi.org/10.1006/jmbi.1999.2805
Koh J, Saecker RM, Record MT Jr (2008) DNA binding mode transitions of *Escherichia coli* HU(alphabeta): evidence for formation of a bent DNA–protein complex on intact, linear duplex DNA. J Mol Biol 383:324–346. https://doi.org/10.1016/j.jmb.2008.07.024
Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372:774–797. https://doi.org/10.1016/j.jmb.2007.05.022
Kumar S, Tsai CJ, Nussinov R (2003) Temperature range of thermodynamic stability for the native state of reversible two-state proteins. Biochemistry 42:4864–4873. https://doi.org/10.1021/bi027184+
Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685
Larson SM, DiNardo AA, Davidson AR (2000) Analysis of covariation in an SH3 domain sequence alignment: applications in tertiary contact prediction and the design of compensating hydrophobic core substitutions. J Mol Biol 303:433–446. https://doi.org/10.1006/jmbi.2000.4146
Lavoie BD, Shaw GS, Millner A, Chaconas G (1996) Anatomy of a flexer-DNA complex inside a higher-order transposition intermediate. Cell 85:761–771
Lehmann M, Wyss M (2001) Engineering proteins for thermostability, the use of sequence alignments versus rational design and directed evolution. Curr Opin Biotechnol 12:371–375
Lehmann M, Kostrewa D, Wyss M, Brugger R, D'Arcy A, Pasamontes L, van Loon AP (2000a) From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase. Protein Eng 13:49–57
Lehmann M, Pasamontes L, Lassen SF, Wyss M (2000b) The consensus concept for thermostability engineering of proteins. Biochim Biophys Acta 1543:408–415
Lehmann M, Loch C, Middendorf A, Studer D, Lassen SF, Pasamontes L, van Loon AP, Wyss M (2002) The consensus concept for thermostability engineering of proteins, further proof of concept. Protein Eng 15:403–411
Li S, Waters R (1998) Escherichia coli strains lacking protein HU are UV sensitive due to a role for HU in homologous recombination. J Bacteriol 180:3750–3756
Magliery TJ, Regan L (2004) Beyond consensus: statistical free energies reveal hidden interactions in the design of a TPR motif. J Mol Biol 343:731–745. https://doi.org/10.1016/j.jmb.2004.08.026
Malik M, Bensaid A, Rouviere-Yaniv J, Drlica K (1996) Histone-like protein HU and bacterial DNA topology, suppression of an HU deficiency by gyrase mutations. J Mol Biol 256:66–76. https://doi.org/10.1006/jmbi.1996.0068
McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R (2013) Analysis tool web services from the EMBL-EBI. Nucleic Acids Res 41:W597–W600. https://doi.org/10.1093/nar/gkt376
Miller KM, Jackson SP (2012) Histone marks: repairing DNA breaks within the context of chromatin. Biochem Soc Trans 40:370–376. https://doi.org/10.1042/BST20110747
Miyabe I, Zhang QM, Kano Y, Yonei S (2000) Histone-like protein HU is required for recA gene-dependent DNA repair and SOS induction pathways in UV-irradiated *Escherichia coli*. Int J Radiat Biol 76:43–49
Mosavi LK, Minor DL Jr, Peng ZY (2001) Consensus-derived structural determinants of the ankyrin repeat motif. Proc Natl Acad Sci USA 99:16029–16034. https://doi.org/10.1073/pnas.252537899
Mouw KW, Rice PA (2007) Shaping the Borrelia burgdorferi genome: crystal structure and binding properties of the DNA-bending protein Hbb. Mol Microbiol 63:1319–1330. https://doi.org/10.1111/j.1365-2958.2007.05586.x
Murphy KP, Freire E (1992) Thermodynamics of structural stability and cooperative folding behavior in proteins. Adv Protein Chem 43:313–361
Nikolova PV, Henckel J, Lane DP, Fersht AR (1998) Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability. Proc Natl Acad Sci USA 95:14675–14680
Oberto J, Nabti S, Jooste V, Mignot H, Rouviere-Yaniv J (2009) The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. PLoS ONE 4:e4367. https://doi.org/10.1371/journal.pone.0004367
Orfaniotou F, Tzamalis P, Thanassoulas A, Stefanidi E, Zees A, Boutou E, Vlassi M, Nounesis G, Vorgias CE (2009) The stability of the archaeal HU histone-like DNA-binding protein from *Thermoplasma volcanium*. Extremophiles 13:1–10. https://doi.org/10.1007/s00792-008-0190-6
Padas PM, Wilson KS, Vorgias CE (1992) The DNA-binding protein HU from mesophilic and thermophilic bacilli, gene cloning, overproduction and purification. Gene 117:39–44
Pontiggia A, Negri A, Beltrame M, Bianchi ME (1993) Protein HU binds specifically to kinked DNA. Mol Microbiol 7:343–350
Preobrajenskaya O, Boullard A, Boubrik F, Schnarr M, Rouvière-Yaniv J (1994) The protein HU can displace the LexA repressor from its DNA-binding sites. Mol Microbiol 13:459–467
Privalov PL, Potekhin SA (1986) Scanning microcalorimetry in studying temperature-induced changes in proteins. Methods Enzymol 131:4–51
Ramstein J, Hervouet N, Coste F, Zelwer C, Oberto J, Castaing B (2003) Evidence of a thermal unfolding dimeric intermediate for the *Escherichia coli* histone-like HU proteins: thermodynamics and structure. J Mol Biol 331:101–121
Roth A, Urmoneit B, Messer W (1994) Functions of histone-like proteins in the initiation of DNA replication at oriC of *Escherichia coli*. Biochimie 76:917–923
Rouviere-Yaniv J, Yaniv M, Germond JE (1979) E. coli DNA binding protein HU forms nucleosome like structure with circular double-stranded DNA. Cell 17:265–274

Ruiz-Sanz J, Filimonov VV, Christodoulou E, Vorgias CE, Mateo PL (2004) Thermodynamic analysis of the unfolding and stability of the dimeric DNA-binding protein HU from the hyperthermophilic eubacterium *Thermotoga maritima* and its E34D mutant. Eur J Biochem 271:1497–1507. https://doi.org/10.1111/j.1432-1033.2004.04057.x

Sagi D, Friedman N, Vorgias CE, Oppenheim AB, Stavans J (2004) Modulation of DNA conformations through the formation of alternative high-order HU-DNA complexes. J Mol Biol 341:419–428. https://doi.org/10.1016/j.jmb.2004.06.023

Scopes RK (1974) Measurement of protein by spectrophotometry at 205 nm. Anal Biochem 59:277–282

Shannon CE (1948) The mathematical theory of communication. Bell Syst Tech J 27:379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Silva IR, Larsen DM, Jers C, Derkx P, Meyer AS, Mikkelsen JD (2013) Enhancing RGI lyase thermostability by targeted single point mutations. Appl Microbiol Biotechnol 97:9727–9735. https://doi.org/10.1007/s00253-013-5184-3

Steif C, Weber P, Hinz HJ, Flossdorf J, Cesareni G, Kokkinidis M (1993) Subunit interactions provide a significant contribution to the stability of the dimeric four-alpha-helical-bundle protein ROP. Biochemistry 32:3867–3876

Swinger KK, Lemberg KM, Zhang Y, Rice PA (2003) Flexible DNA bending in HU-DNA cocrystal structures. EMBO J 22:3749–3760. https://doi.org/10.1093/emboj/cdg351

Tanaka I, Appelt K, Dijk J, White SW, Wilson KS (1984) 3-Å resolution structure of a protein with histone-like properties in prokaryotes. Nature 310:376–381

Tina KG, Bhadra R, Srinivasan N (2007) PIC: protein interactions calculator. Nucleic Acids Res 35:W473–W476. https://doi.org/10.1093/nar/gkm423

Tominaga T, Nakagawa A, Tanaka I, Kawamura S, Kimura M (1999) High-resolution crystals of the HU mutant K38N from *Bacillus stearothermophilus*. J Struct Biol 125:86–89. https://doi.org/10.1006/jsbi.1998.4076

Viader-Salvadó JM, Gallegos-López JA, Carreón-Treviño JG, Castillo-Galván M, Rojo-Domínguez A, Guerrero-Olazarán M (2010) Design of thermostable beta-propeller phytases with activity over a broad range of pHs and their overproduction by *Pichia pastoris*. Appl Environ Microbiol 76:6423–6430. https://doi.org/10.1128/AEM.00253-10

Vis H, Mariani M, Vorgias CE, Wilson KS, Kaptein R, Boelens R (1995) Solution structure of the HU protein from *Bacillus stearothermophilus*. J Mol Biol 254:692–703. https://doi.org/10.1006/jmbi.1995.0648

Vis H, Vorgias CE, Wilson KS, Kaptein R, Boelens R (1998) Mobility of NH bonds in DNA-binding protein HU of *Bacillus stearothermophilus* from reduced spectral density mapping analysis at multiple NMR fields. J Biomol NMR 11:265–277

Vriend G (1990) WHAT IF, a molecular modeling and drug design program. J Mol Graph 8:52–56

Wang Q, Buckle AM, Foster NW, Johnson CM, Fersht AR (1999) Design of highly stable functional GroEL minichaperones. Protein Sci 8:2186–2193. https://doi.org/10.1110/ps.8.10.2186

Welfle H, Misselwitz R, Welfle K, Groch N, Heinemann U (1992) Salt-dependent and protein-concentration-dependent changes in the solution structure of the DNA-binding histone-like protein, HBsu, from *Bacillus subtilis*. Eur J Biochem 204:1049–1055

White SW, Appelt K, Wilson KS, Tanaka I (1989) A protein structural motif that bends DNA. Proteins 5:281–288. https://doi.org/10.1002/prot.340050405

White SW, Wilson KS, Appelt K, Tanaka I (1999) The high-resolution structure of DNA-binding protein HU from *Bacillus stearothermophilus*. Acta Crystallogr D Biol Crystallogr 55:801–809. https://doi.org/10.1107/s0907444999000578

Wienken CJ, Baaske P, Rothbauer U, Braun D, Duhr S (2010) Protein-binding assays in biological liquids using microscale thermophoresis. Nat Commun 1:100. https://doi.org/10.1038/ncomms1093

Wilson KS, Vorgias CE, Tanaka I, White SW, Kimura M (1990) The thermostability of DNA-binding protein HU from Bacilli. Protein Eng Des Sel. https://doi.org/10.1093/protein/4.1.11

Wu TH, Chen CC, Cheng YS, Ko TP, Lin CY, Lai HL, Huang TY, Liu JR, Guo RT (2014) Improving specific activity and thermostability of *Escherichia coli* phytase by structure-based rational design. J Biotechnol 175:1–6. https://doi.org/10.1016/j.jbiotec.2014.01.034

Yu J, Zhou Y, Tanaka I, Yao M (2010) Roll: a new algorithm for the detection of protein pockets and cavities with a rolling probe sphere. Bioinformatics 26:46–52. https://doi.org/10.1093/bioinformatics/btp599

Zhao C, Fu J, Oztekin A, Cheng X (2014) Measuring the Soret coefficient of nanoparticles in a dilute suspension. J Nanopart Res 16(10):2625–2642. https://doi.org/10.1007/s11051-014-2625-6

Zhou HX (2002) Toward the physical basis of thermophilic proteins: linking of enriched polar interactions and reduced heat capacity of unfolding. Biophys J 83:3126–3133. https://doi.org/10.1016/S0006-3495(02)75316-2

Zillner K, Jerabek-Willemsen M, Duhr S, Braun D, Längst G, Baaske P (2012) Microscale thermophoresis as a sensitive method to quantify protein, nucleic acid interactions in solution. Methods Mol Biol 815:241–252. https://doi.org/10.1007/978-1-61779-424-7_18