DR. TOMÁŠ KOVAĽ (Orcid ID : 0000-0002-2836-3765)

DR. JAN DOHNALEK (Orcid ID : 0000-0002-8083-7980)

Received Date : 14-Dec-2018

Revised Date : 29-Mar-2019

Accepted Date : 01-Apr-2019

Article type : Research Letter

Domain structure of HelD, an interaction partner of *Bacillus subtilis* RNA polymerase

Tomáš Kovaľ[1*], Petra Sudzinová[2*], Terézia Perháčová[1†], Mária Trundová[1], Tereza Skálová[1], Karla Fejfarová[1], Hana Šanderová[2], Libor Krásný[2], Jarmila Dušková[1], Jan Dohnálek[1]

[1]Laboratory of Structure and Function of Biomolecules, Institute of Biotechnology of the Czech Academy of Sciences, v. v. i., Biocev, Průmyslová 595, 25250 Vestec, Czech Republic

[2]Laboratory of Microbial Genetics and Gene Expression, Institute of Microbiology of the Czech Academy of Sciences, v. v. i., Vídeňská 1083, 14220, Praha 4, Czech Republic

*These authors contributed equally.
† Died on 27 August 2017

Corresponding author

Jan Dohnálek, Laboratory of Structure and Function of Biomolecules, Institute of Biotechnology of the Czech Academy of Sciences, v. v. i., Biocev, Průmyslová 595, 25250 Vestec, Czech Republic, tel. +420 325 873 758, fax +420 325 873 710, e-mail: dohnalek@ibt.cas.cz, http://www.ibt.cas.cz/vyzkum/laboratore/laborator-struktury-a-funkce-biomolekul/

Running title
***B. subtilis* HelD domains**

Keywords
RNAP; HelD; SAXS; *Bacillus subtilis*

Abbreviations
AMP-PNP, adenosine 5′-(β,γ-imido)triphosphate
MBP, maltose binding protein
Rg, radius of gyration
SAXS, small-angle X-ray scattering
SEC, size-exclusion chromatography
SEC-SAXS, size-exclusion chromatography in line with small-angle X-ray scattering
TCEP, tris(2-carboxyethyl)phosphine

Abstract

HelD is a helicase-like protein binding to *Bacillus subtilis* RNA polymerase (RNAP), stimulating transcription in an ATP-dependent manner. Here, our small angle X-ray scattering data bring the first insights into the HelD structure: HelD is compact in shape and undergoes a conformational change upon substrate analog binding. Furthermore, the HelD domain structure is delineated, and a partial model of HelD is presented. In addition, the unique N-terminal domain of HelD is characterized as essential for its transcription-related function but not for ATPase activity, DNA binding, or binding to RNAP. The study provides a topological basis for further studies of the role of HelD in transcription.

Introduction

Bacteria are the dominant form of life on Earth. They inhabit every possible niche and excel at adaptation. Adaptation depends on changes in gene expression. The first step in gene expression is transcription of genetic information from DNA into RNA. The key enzyme of this process in bacteria is a multi-subunit enzyme – DNA dependent RNA polymerase (RNAP). The activity of RNAP has to be tightly regulated. This is mediated by various factors, such as small molecules including nucleoside triphosphates (NTPs) [1,2] and ppGpp [3], by small RNAs, like 6S and Ms1 RNA [4,5], and by numerous proteins [6-8].

Previously, we and others identified a new interaction partner of RNAP in *Bacillus subtilis*, a helicase-like protein termed HelD (~ 90 kDa) [9,10]. HelD belongs to the UvrD helicase family and, based on a previous bioinformatics analysis, consists of three domains: (i) the N-terminal domain, which bears no homology to other known helicases, (ii) the ATPase domain, and (iii) the C-terminal domain. Other helicases from this protein family (superfamily 1, SF1) are capable of unwinding DNA in either 3'-5' (SF1A subfamily) or 5'-3' (SF1B subfamily) translocation direction. Structural experiments with the SF1 family helicases indicate that these enzymes are monomers [11]. The most thoroughly studied helicases are PcrA from gram-positive bacteria and UvrD from gram-negative bacteria that share the same domain organization [12-15]. To the contrary, the predicted domain structure of HelD is different and the sequence identity is low (12% between HelD and UvrD). Nevertheless, UvrD is a protein with known 3D structure that is most related to HelD.

UvrD's ATP-ase and "C-terminal" domains are similar to HelD but with different topology; importantly, the DNA-binding domain of UvrD is not present in HelD.

The *E. coli* protein RapA, involved in the release of stalled transcription complexes by backward translocation, represents the only related protein (sequence identity to HelD 21.0%), for which coordinates of the complex with RNAP are available (PDB ID 4S20) [16]. The domain structure of RapA resembles that of UvrD (N-terminal domain, ATPase domains) but with different chain topology [17].

Previously, we showed that HelD binds to RNAP and stimulates its activity in an ATP-dependent manner by stimulating transcriptional cycling and elongation [9]. Also, HelD was reported to be involved in DNA repair and homologous recombination [18] and amyloid-like fibrils formation [19]. A strain lacking the HelD encoding gene displays prolonged lag phase [9]. Nevertheless, the specific role(s) and structure of HelD are still unknown.

Here, we extend the knowledge on HelD by characterizing its shape by small angle X-ray scattering (SAXS), by analysis of its domain structure, and by studies of the importance of the unique N-terminal domain of HelD for binding to DNA and to RNAP and for the activity of the protein.

Materials and methods

Preparation of protein samples and complexes

Bacillus subtilis HelD

HelD from *Bacillus subtilis* was expressed from pHelD-His6 (LK800, Table 1) in *E. coli* BL21 (DE3). The construct contained a 6xHis tag at the N-terminus cleavable by TEV protease. Details of the cloning, expression, and purification procedures are in [9].

Bacillus subtilis HelDΔN

The truncated version of HelD (HelDΔN, residues 204-774) lacking the N-terminal domain was expressed in *E. coli* using pET28-MBP-TEV, a gift from Zita Balklava & Thomas Wassmer (Addgene plasmid # 69929; http://n2t.net/addgene:69929 ; RRID:Addgene_69929) [20]. HelDΔN was prepared in fusion with maltose binding protein (MBP) and His-tag at the N-terminus. MBP is cleavable from the construct using TEV protease. HelDΔN expression was induced with 1mM IPTG in *E. coli* Lemo21 (DE3) cells (New England Biolabs) grown in Power broth (Molecular Dimensions). HelDΔN was purified using Ni-NTA affinity chromatography and size exclusion chromatography. For details see Supplementary material.

TEV digestion

Except for samples for SAXS experiments, both HelD and HelDΔN were treated with TEV protease. TEV cleavage was performed in 50 mM Tris/HCl, pH 7.5, 100 mM NaCl, 0.5 mM TCEP, 0.5 mM dithiothreitol (DTT), 1 mM EDTA, and 5% (v/v) glycerol at 37 °C for 1 hour. Both samples were then run on charged HisTrap^{TM} FF (1 mL) columns (GE-Healthcare) using an ÄKTA purifier and the proteins were in the flow-through fraction. Samples were analyzed using SDS-PAGE (Fig. S1).

Bacillus subtilis RNA polymerase

RNAPΔHelD was produced using a strain lacking HelD (LK782, Table 1). Expression and purification were done according to [9] with the addition of size–exclusion chromatography (SEC) performed using an ÄKTA purifier, a Superose 6 10/300 GL column (GE Healthcare), and 100 mM Na/K phosphate buffer, pH 7.5 supplemented with 50 mM NaCl, 0.5 mM TCEP, and 3% (v/v) glycerol.

SAXS data collection

Size-exclusion chromatography in line with small-angle X-ray scattering (SEC-SAXS) were used to analyze HelD in solution -/+ an ATP analogue. HelD (10 mg/mL) in the presence of 10 mM adenosine 5′-(β,γ-imido)triphosphate (AMP-PNP), a non-hydrolysable analogue of ATP, was applied to a Superdex 200 10/300 GL column (GE Healthcare) at a flow rate of 0.5 mL/min. HelD without AMP-PNP (19 mg/mL) was applied to a Superose 6 10/300 GL column (GE Healthcare) at a flow rate of 0.4 mL/min. Both samples were applied using an FPLC Agilent BioInert system with an autosampler. Buffer containing 100 mM Na/K phosphate pH 7.5, 50 mM NaCl, 0.5 mM TCEP, and 3% (v/v) glycerol was used for HelD without AMP-PNP. For HelD with bound AMP-PNP, 50 mM Tris pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 0.5 mM TCEP, and 3% (v/v) glycerol was used. SAXS data were collected on the eluted fractions at beam line P12 of the synchrotron radiation source PetraIII in Hamburg [23] using a Pilatus 2M detector (Dectris) at sample to detector distance 3 m, λ = 1.24 Å, exposure time per image 0.995 s, at 20 °C. Quality of SAXS data and the SEC profiles are shown in Supplementary data (Figs. S2-S4).

SAXS data analysis and ab initio model calculations

The SEC-SAXS data were processed using ATSAS version 2.8.2 with help of versions 2.7.2 (stable version of DAMMIF) and 2.8.3 [24]. Analysis of the resulting *ab initio* models, generation of three-dimensional molecular envelopes, fitting with protein structures, and generation of the graphics were performed with the UCSF CHIMERA package [25].

SAXS data processing. SEC-SAXS frames were merged in interval selected by R_g value, omitting frames of low quality (Fig. S2). The selected intervals were: for HelD:AMP-PNP, frames 1268-1331 as body and 1420-1569 as solvent; for HelD, frames 2443-2448 and 895-1027, 2807-2980 as solvent. Solvent-subtracted merged SAXS curves, Guinier plot, Kratky plot, and distance distribution function are shown in Figures S3 and S4. R_g, D_{max} and molecular weight estimated from SAXS data are shown in Table 2.

Sequence analysis

Multiple and pairwise sequence alignments were performed using the Clustal Omega server [27], MUSCLE v3.8.31 [28], and GeneDoc [29]. Manual editing of sequence alignments was done in GeneDoc. Structure-based sequence alignments were calculated using the Secondary structure matching algorithm of the PDBeFold server (http://www.ebi.ac.uk/msd-srv/ssm/) [30] and processed using GeneDoc and MUSCLE.

Molecular Modeling

Generation of the partial HelD model was based on the coordinates of the C-terminal domain of the putative DNA helicase from *Lactobacillus plantarum* (LpCter, PDB ID 3DMN, unpublished). Alternative conformations and ligands were removed from the model. Two positions of the domain in each of the structures of helicase UvrD from *E. coli* (PDB ID 2IS4) [15] and of protein RapA from *E. coli* (PDB ID 4S20) [16] were found using the PDBeFold server with multiple hits option and standard parameters. For each target structure the first 20 matches were analyzed and assigned to two categories: (i) putative 1A domain, (ii) C-terminal domain (putative 2A domain). Alignments of the full sequences of UvrD, RapA, and HelD, and of sequences of individual domains of these proteins with the addition of LpCter provided information on localization of both domains in the HelD sequence. These results together with structural alignments of 3D structures of the full proteins and of individual domains were used as a basis for assembling the model of the recurring domains in HelD. On the basis of a higher sequence similarity between HelD and UvrD (as opposed to RapA) in the interdomain region, the mutual position of the putative 1A and C-terminal (putative 2A) domains was defined as in UvrD.

The C-terminal domain (putative 2A) of HelD could be reliably modeled with LpCter (100% coverage, 40% identity). Unreliable parts of the 3D model of the putative 1A domain similar to LpCter (as judged by agreement between UvrD, RapA, and LpCter) were deleted. Secondary structure prediction for HelD was calculated using Jpred4 [31]. The multiple structure-based sequence alignment of the putative 1A and C-terminal (putative 2A) domains of the known structures was aligned against the sequence of HelD to satisfy the results of the secondary structure prediction. A satisfactory alignment was used to assign sequence to the model of the putative 1A domain of HelD. The assembled model of both domains of HelD was energy minimized using the Yasara server [32].

In vitro transcription assays

Transcription experiments were performed with the *B. subtilis* RNAP core lacking HelD (isolated from the strain LK782) reconstituted with saturating concentration of σ^A (ratio 1:5, strain LK22, [21]) and HelD/HelDΔN (ratio 1:4) in storage buffer (50 mM Tris-HCl, pH 8.0, 0.1 M NaCl, 50% (v/v) glycerol, 3 mM β-mercaptoethanol) for 15 min at 30 °C. Multiple round transcription reactions were carried out in 10 µL reaction volumes with 30 nM RNAP holoenzyme and 100 ng of supercoiled plasmid DNA containing *B. subtilis rrn*B P1 promoter [22]. The transcription buffer contained 40 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 1 mM DTT, 0.1 mg/mL BSA, 150 mM KCl, and NTPs (ATP, CTP were at 200 µM; GTP 1300 µM; UTP was 10 µM plus 2 µM radiolabeled [α-^{32}P]UTP). All transcription reactions were allowed to proceed for 15 min at 30 °C and were stopped with equal volumes of formamide stop solution (95% formamide, 20 mM EDTA, pH 8.0). Samples were loaded onto 7 M urea-7% polyacrylamide gels and electrophoresed. The dried gels were scanned with a Molecular Imager FX (Bio-Rad) and visualized and analyzed using the Quantity One software (Bio-Rad).

Analysis of the formation of RNAP:HelD/HelDΔN complexes

To analyze the association of HelDΔN to RNAP *in vitro* we performed gel-shift experiments under native conditions in non-denaturing PAGE gel. Full-length HelD was used as a positive control. Both HelD and HelDΔN were treated using TEV protease prior to experiments (Fig. S1). For the gel-shift assay, RNAP was mixed with either HelD or HelDΔN at molar ratio 1:1 and 1:2 and left in the dark at room temperature for 1 hour. This incubation was performed in 100 mM Na/K phosphate buffer, pH 7.5, with 100 mM NaCl, 0.5 mM TCEP, and 3% (v/v) glycerol. The gel-shift assay was performed using an XCell SureLock™ mini–cell electrophoresis system, polyacrylamide NuPAGE® 7% Tris-acetate gel, Novex™ Tris-Glycine Native Sample Buffer, and Novex™ Tris-Glycine Native Running Buffer (ThermoFisher Scientific). Electrophoresis was done according to the manufacturer's instructions.

ATPase activity of HelD and HelDΔN

ATPase activity of both full-length HelD and of the truncated construct HelDΔN was tested using samples treated with TEV protease. Reactions were carried out at 25 °C for 30 min. Mixtures (100 µL) contained 10 µg of protein and 10 mM of ATP in 50 mM Tris pH 7.5, 50 mM NaCl, and 5 mM $MgCl_2$. The amount of released phosphate was analyzed according to the modified molybdenum blue method [33] by spectrophotometry at λ = 850 nm using a microplate reader CLARIOStar (BMG LABTECH).

DNA binding assay

Preparation of radiolabeled DNA

The DNA primer LK999 5´-GCGCTACGGCGTTTCACTTC-3´was radiolabeled at the 5´end by T4 polynucleotide kinase and purified with Nucleotide removal kit (Qiagene). A DNA fragment was prepared with PCR from the plasmid containing P*veg* (pRLG7558) using the ^{32}P-labeled primer LK999 and primer LK1000 5´-CCACCTGACGTCTAAGAAACC-3´. The DNA fragment started at –118 relative to the transcription start site and ended at +55.

Native PAGE assays

0.5 pmol of 5´-radiolabeled DNA and HelD/HelDΔN at ratios 1:10, 1:20, 1:100, 1:400 and 1:800 were incubated for 15 min at 30 °C in buffer containing 50 mM Tris-HCl, pH 8.0, 0.1 M NaCl, 50% (v/v) glycerol, 3 mM β-mercaptoethanol. Final concentrations in the 10μl reactions were 0.05 μM for DNA and up to 40 μM for proteins, respectively. As controls, denatured HelD/HelDΔN (5 min at 95 °C) at ratios 1:10, 1:100, 1:400, and 1:800 were used. After incubation samples were mixed with Native PAGE 4X Sample buffer (Invitrogen), loaded onto the Native PAGE 4-16% Bis-Tris Gel (Invitrogen), and electrophoresed. The dried gels were scanned with a Molecular Imager_FX (BioRad).

Results

HelD shape

To provide structural information for HelD, we performed numerous crystallization experiments with the protein and its complex with RNAP, all without success (data not shown). Therefore, we performed SAXS experiments. SAXS data were obtained for HelD and HelD in complex with AMP-PNP (an ATP analogue used to mimick the ATP-bound state) in the SEC-SAXS mode. The average R_g values corresponded to the size of a monomer of HelD (Table 2).

Domain structure and modeling of HelD fragment

To get better insights into the domain structure of HelD, we performed sequence analysis and 3D modeling of (or parts of) HelD. We used sequences and structures of related proteins RapA, UvrD, and the C-terminal fragment of the *Lactobacillus plantarum* HelD homologue (LpCter), displaying the Rossman fold (Pfam, http://pfam.xfam.org) [34]. SF1 helicases, such as UvrD, typically consist of 2 domains divided into 4 subdomains: 1A, 1B, 2A, 2B. The domain structure of UvrD is shown in Fig. 1A. The following analysis starts with the C-terminal domain (of HelD) and continues towards the N-terminus.

As the starting point, we used LpCter and created its 3D superpositions with RapA and UvrD (Fig. 1B and C) using SSM (secondary structure matching). The 3D structure of LpCter superposed with 2A domains in both RapA and UvrD all showing Rossman fold (r.m.s.d. 2.8 Å on 109 C^{α} atoms of RapA and 1.8 Å on 152 C^{α} atoms of UvrD, sequence identity 11.9% and 23.0%, respectively). This Rossman fold was also identified at a second location, in domains 1A (r.m.s.d of C^{α} coordinates 3.0 Å for 113 residues of RapA and 2.5 Å for 109

residues of UvrD, sequence identity 11.5% and 9.2%, respectively). The fit of LpCter with domain 2A was better than the fit with domain 1A. Therefore, we propose to name the C-terminal domain of HelD as putative 2A and an internal part of HelD as putative 1A (Fig. 2).

Given the sequence similarity between RapA, UvrD, *L. plantarum* HelD and *B. subtilis* HelD (illustrated in Fig. S5), it was possible to create a limited model of HelD consisting of two Rossman fold domains: the putative domain 1A and the C-terminal domain which corresponded to domain 2A (Fig. 1D). This partial HelD model, based on the coordinates of UvrD (PDB ID 2IS4) and LpCter (PDB ID 3DMN), was further used for interpretation of the SAXS data.

The remaining parts of HelD are the HelD-specific and N-terminal domains. In the linear aa sequence, the HelD-specific domain is inserted into the putative domain 1A. No 3D structures of these domains or their homologs exist.

The above described results suggested division of the HelD protein into four structurally compact parts of which one (putative domain 1A) is divided in the sequence in two parts (by the HelD-specific domain) and the other domains are continuous: N-terminal (residues 1–203, MNQQ...HHSD), putative 1A - part 1 (residues 204– ~ 292, TQMK... EQAT), HelD-specific domain (~ 293– ~539, FQEY...KNTK), putative 1A - part 2 (~540 – 606, IKHL...LKRT), and the C-terminal (putative 2A) domain (607–774, YRST...QIAE). The suggested domain structure of HelD and its comparison with the domain structures of RapA and UvrD are shown in Fig. 2.

Subsequently, the SAXS-based *ab initio* shape calculations provided sets of three-dimensional models, which could be represented by average models of HelD shown in Fig. 3A, B, and C. The optimized alignments between the SAXS models and monomers of UvrD and RapA (Fig. 3D and E) revealed that the AMP-PNP-bound HelD acquired a different conformation with the overall shape similar to its distant homologs. Based on the fit of UvrD and RapA in the HelD envelopes, an approximate position of the partial HelD model can be estimated (Fig. 3F).

N-terminal domain of HelD

Motivated by the sequence analysis, we decided to provide data for the unique N-terminal domain. We designed a construct of HelD lacking this domain, hereafter called HelDΔN (residues 204-774). This truncated version of HelD was recombinantly produced in, and purified from *E. coli*, and subsequently used to address the role(s) of this part of the protein in the activity of HelD and its interactions with RNAP. First, we tested whether the enzymatic activity of HelD was still intact: HelDΔN maintained its ATPase activity at a level comparable to full-length HelD (Fig. 4A). Second, we investigated the importance of the N-terminal domain for the interaction with RNAP. Fig. 4B shows that truncated HelD associated with the RNAP core similarly as HelD. As HelD was previously shown to interact rather non-specifically with DNA [9], we compared HelDΔN with full-length HelD in their

ability to bind DNA. Fig. 4C shows that even in the absence of the N-terminal domain HelD still interacted with DNA although less strongly than full-length HelD. Finally, we tested the ability of the truncated protein to enhance transcription *in vitro*. Multiple round transcription assays revealed that the N-terminal domain was essential for the stimulatory effect of HelD (Fig. 4D).

Discussion

In this study, we obtained the first insights into the structure-function relationship of the RNAP-associated protein HelD. We characterized its overall 3D shape and a conformational change induced by ATP. Importantly, we described the domain structure of HelD and partially characterized the role of its unique N-terminal domain.

The shape of AMP-PNP-bound HelD differs from that of the unliganded form and approximately resembles the observed crystal structures of *E. coli* UvrD and RNAP-bound RapA. HelD without AMP-PNP likely represents a structurally distinct form. Even if limited by the low resolution of the current SAXS results, the observed changes suggest that, upon ATP binding/cleavage, HelD undergoes a pronounced conformational change (data for isolated HelD in solution). This conformational change is supported by the observed substantial increase in the thermal stabilization of HelD upon AMP-PNP binding (change of T_m from 51 °C to 62 °C, Fig. S6). Similarly to HelD, a conformational change (rotation of the 2B subdomain) was observed for UvrD upon binding of nucleotide or/and DNA [35]. The observed change of HelD conformation can be also reminiscent of e. g. the translation elongation factor Tu (EF-Tu) where the binding of GTP or GDP radically affects the conformation, and in the absence of either molecule the protein collapses [36].

The structure of HelD can be newly dissected into the N-terminal domain (residues 1-203), the putative 1A domain, topologically split in two parts (204-~292 and ~540-606), the HelD-specific domain inserted in between them (293-~539), and the C-terminal (putative 2A) domain (607-774). The putative 1A domain and the C-terminal domain together form the ATPase unit. The overall organization of the protein is different from UvrD and RapA. The role of the ATPase unit most likely lies in securing the transfer of the ATP-driven conformational changes (confirmed by our SAXS results). Our results show that the N-terminal domain of HelD appears to be essential for its transcription-related function but not for ATPase activity, DNA binding, or binding to RNAP. This domain is instrumental for the HelD function by a yet unknown molecular mechanism.

Further studies will focus on explaining the role of the N-terminal domain, the structure and function of the HelD-specific domain, and elucidation of the exact mechanistic details of the ATP-driven action of HelD.

Acknowledgments

This work was supported by CSF (grants no. 15-05228S and 19-12956S), by the projects CIISB4HEALTH (CZ.02.1.01/0.0/0.0/16_013/0001776), BIOCEV (CZ.1.05/1.1.00/02.0109), and ELIBIO (CZ.02.1.01/0.0/0.0/15_003/0000447) from the ERDF, institutional support (RVO 86652036), and MEYS support for CMS-Biocev (LM2015043). We would like to thank J. Vohradsky for providing access to Molecular Imager FX Scanner, supported by the Czech research infrastructure for systems biology C4SYS (LM2015055).

Author contributions

LK and JDo designed research, TK, PS, HS, TP, and MT prepared samples and performed experiments, TK designed, performed, and analyzed SAXS experiments, TS performed computational modeling and analysis, KF performed sequence-structure analysis and stability calculations, JDu secured laboratory background, JDo performed sequence-structure analysis, modeling, and data interpretation, TK, PS, HS, LK, and JDo discussed results and wrote the manuscript.

References

1 Krásný L and Gourse RL (2004) An alternative strategy for bacterial ribosome synthesis: *Bacillus subtilis* rRNA transcription regulation. *EMBO J* **23**, 4473–4483.

2 Murray HD, Schneider D and Gourse RL (2003) Control of rRNA expression by small molecules is dynamic and nonredundant. *Mol Cell* **12**, 125–134.

3 Hauryliuk V, Atkinson GC, Murakami KS, Tenson T and Gerdes K (2015) Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat Rev Microbiol* **13**, 298–309.

4 Steuten B, Hoch PG, Damm K, Schneider S, Köhler K, Wagner R and Hartmann RK (2014) Regulation of transcription by 6S RNAs. *RNA Biol* **11**, 508–521.

5 Hnilicová J, Jirát Matějčková J, Šiková M, Pospíšil J, Halada P, Pánek J and Krásný L (2014) Ms1, a novel sRNA interacting with the RNA polymerase core in mycobacteria. *Nucleic Acids Res* **42**, 11763-11776.

6 Borukhov S, Lee J and Laptenko O (2004) Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol Microbiol* **55**, 1315–1324.

7 Brantl S and Licht A (2010) Characterisation of *Bacillus subtilis* Transcriptional Regulators Involved in Metabolic Processes. *Curr Protein Pept Sci* **11**, 274–291.

8 Weiss A and Shaw LN (2015) Small things considered: the small accessory subunits of RNA polymerase in Gram-positive bacteria. *FEMS Microbiol Rev* **5**, 541–554.

9 Wiedermannová J, Sudzinová P, Kovaľ T, Rabatinová A, Šanderová H, Ramaniuk O, Rittich Š, Dohnálek J, Fu Z, Halada P, Lewis P and Krasny L (2014) Characterization of HelD, an

interacting partner of RNA polymerase from *Bacillus subtilis*. *Nucleic Acids Res* **42**, 5151–5163.

10 Delumeau O, Lecointe F, Muntel J, Guillot A, Guédon E, Monnet V, Hecker M, Becher D, Polard P and Noirot P. (2011) The dynamic protein partnership of RNA polymerase in *Bacillus subtilis*. *Proteomics* **11,** 2992-3001.

11 Singleton MR, Dillingham MS and Wigley DB (2007) Structure and Mechanism of Helicases and Nucleic Acid Translocases. *Annu Rev Biochem* **76**, 23-50.

12 Petit MA, Dervyn E, Rose M, Entian KD, McGovern S, Ehrlich SD and Bruand C (1998) PcrA is an essential DNA helicase of *Bacillus subtilis* fulfilling functions both in repair and rolling-circle replication. *Mol Microbiol* **29**, 261–273.

13 Epshtein V, Kamarthapu V, McGary K, Svetlov V, Ueberheide B, Proshkin S, Mironov A and Nudler E (2014) UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature* **505**, 372–377.

14 Gwynn EJ, Smith AJ, Guy CP, Savery NJ, McGlynn P and Dillingham MS (2013) The conserved C-terminus of the PcrA/UvrD helicase interacts directly with RNA polymerase. *PLoS One* **8**, e78141.

15 Lee JY and Yang W (2006) UvrD Helicase Unwinds DNA One Base Pair at a Time by a Two-Part Power Stroke. *Cell* **127**, 1379–1360.

16 Liu B, Zuo Y and Steitz TA (2015) Structural basis for transcription reactivation by RapA. *Proc Natl Acad Sci USA* **112**, 2006-2010.

17 Shaw G, Gan J, Zhou YN, Zhi H, Subburaman P, Zhang R, Joachimiak A, Jin DJ and Ji W (2008) Structure of RapA, a Swi2/Snf2 Protein That Recycles RNA Polymerase during Transcription. *Structure* **16**, 1417-1427.

18 Carrasco BA, Ndez SF, Petit M-A and Alonso JC (2001) Genetic Recombination in *Bacillus subtilis* 168: Effect of Δ*helD* on DNA Repair and Homologous Recombination. *J Bacteriol* **183**, 5772–5777.

19 Kaur G, Kapoor S and Thakur KG (2018) *Bacillus subtilis* HelD, an RNA Polymerase Interacting Helicase, Forms Amyloid-Like Fibrils. *Front. Microbiol*. **9**, 1934.

20 Currinn H, Guscott B, Balklava Z, Rothnie A, Wassmer T. (2015) APP controls the formation of PI(3,5)P vesicles through its binding of the PIKfyve complex. *Cell. Mol. Life Sci.* **73**, 393–408.

21 Chang BY and Doi RH (1990) Overproduction, purification, and characterization of *Bacillus subtilis* RNA polymerase sigma A factor. *J Bacteriol* **172**, 3257–3263.

22 Krásný L, Tiserová H, Jonák J, Rejman D and Sanderová H (2008) The identity of the transcription +1 position is crucial for changes in gene expression in response to amino acid starvation in *Bacillus subtilis*. *Mol Microbiol* **69**, 42-54.

23 Blanchet CE, Spilotros A, Schwemmer F, Graewert MA, Kikhney A, Jeffries CM, Franke D,

Mark D, Zengerle R, Cipriani F, Fiedler S, Roessle M and Svergun DI (2015) Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J Appl Crystallogr* **48**, 431-443.

24 Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, Kikhney AG, Hajizadeh NR, Franklin JM, Jeffries CM and Svergun DI (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Cryst* **50**, 1212-1225.

25 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC and Ferrin TE (2004) UCSF Chimera - a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612.

26 Konarev VP, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Cryst* **36**, 1277-1282.

27 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD and Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539.

28 Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797.

29 Nicholas KB, Nicholas HB Jr and Deerfield DW II (1997) GeneDoc: Analysis and Visualization of Genetic Variation. *EMBNEW.NEWS* **4**, 14.

30 Krissinel E and Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D***60**, 2256-2268.

31 Drozdetskiy A, Cole C, Procter J and Barton GJ (2015) JPred4: a protein secondary structure prediction server. *Nucl Acids Res* **43**, W389-394.

32 Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D and Karplus K (2009) Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* **77**, Suppl 9, 114-122.

33 He Z and Honeycutt WC (2005) A Modified Molybdenum Blue Method for Orthophosphate Determination Suitable for Investigating Enzymatic Hydrolysis of Organic Phosphates, *Communications in Soil Science and Plant Analysis* **36**, 1373-1383.

34 Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J and Bateman A (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* **44**, D279-D285.

35 Jia H, Korolev S, Niedziela-Majka A, Maluf NK, Gauss GH, Myong S, Ha T, Waksman G and Lohman TM (2011) Rotation of the 2B sub-domain of *E. coli* UvrD helicase/translocase coupled to nucleotide and DNA binding. *J Mol. Biol.* **411**, 633-648.

36 Lai J, Ghaemi Z and Luthey-Schulten Z (2017) The Conformational Change in Elongation Factor Tu Involves Separation of Its Domains. *Biochemistry* **56**, 5972-5979.

Tables

Table 1. Bacterial strains and plasmids used in this study.

Strain	Description	Source
LK782	Bsu RNAP *rpoC*-10xHis, *helD::MLS*	Wiedermannová *et al.*, 2014 [9]
LK22	pCD2/Bsu_sigA; BL21 (DE3)	Chang and Doi, 1990 [21]
LK800	pHelD-His6; BL21 (DE3)	Wiedermannová *et al.*, 2014 [9]
HelDΔN	MBP-HelDΔN, Lemo21 (DE3)	This work
Plasmid		
pRLG7558	p770 with P*veg* (-38/+1, +1G)	Krasny and Gourse, 2004 [1]
pLK28	p770 with *rrn*B P1 (-248/+8)	Krasný *et al.*, 2008 [22]

Table 2. Results of SAXS data processing. The individual columns correspond to samples of HelD and complexes HelD:AMP-PNP. The parameters (excluding DAMMIF values) were calculated using PRIMUS [26]. R_g is radius of gyration calculated from the Guinier plot (Figs. S2-S4). D_{max} is the estimated maximum size of the particle (Fig. S2).

	HelD	HelD:AMP-PNP
$I(0)$ [a.u.] (from Guinier)	3920.3	1588.9
R_g [Å] (from Guinier)	35.0	36.2
Guinier analysis fidelity [%]	72	87
$I(0)$ [a.u.] (from P(r))	3913.0	1590
R_g [Å] (from P(r))	35.0	36.3
D_{max} [Å]	113.8	115.1
Porod volume [Å^3] (from P(r))	144	167
MW from sequence [kDa]	90	90
MW from DAMMIF [kDa]	107	90
MW from Porod [kDa]	90	104

Figures legends

Fig. 1. Typical domain organization of SF1 helicases and structural alignments of UvrD and RapA with partial model of HelD. All protein chains are shown as cartoons with secondary structure elements. (A) Domain structure of UvrD from *E. coli.* Domain 1A is colored dark blue, domain 1B light blue, domain 2A orange, and domain 2B pale yellow. Domains are marked. The N- and C-termini of UvrD in the structure are marked (protein chain continues with the C-terminal extension). (B) *E. coli* RapA (grey) with the best superposition of LpCter on the internal ATPase domain (hot pink), on the C-terminal domain (magenta), and the partial model of HelD – internal ATPase domain in green and the C-terminal domain in yellow. (C) *E. coli* UvrD with the best superposition of LpCter on the internal ATPase domain (hot pink), on the C-terminal domain (magenta), and the partial model of HelD – internal ATPase domain (putative 1A) in green and the C-terminal domain in yellow. (D) Partial HelD model of the internal ATPase domain (putative 1A, green) and of the C-terminal domain (yellow) based on the coordinates of UvrD and RapA. Black dashed line – residues missing from the model in the linker region between the two domains. The typical ATP-binding site residues in the ATPase domain are colored in magenta (residues GSGK of the ATP-binding box, see Supplementary material Fig. S5, residues 236-239 and Fig. 2 for the placement of the ATP-binding box).

Fig. 2. Comparison of domain structure of RapA and UvrD from *E. coli*, and HelD from *Bacillus subtilis*. Residue numbers delimiting selected domains are shown. The black bar represents the estimated ATP-binding box, residues 220-258 in HelD.

Fig. 3. SAXS-based molecular envelopes for *B. subtilis* HelD. (A) HelD with AMP-PNP. (B) HelD – without AMP-PNP in the presence of 100 mM PO_4 ions. (C) Optimized fit of envelopes of HelD in the presence (pink) and absence (cyan) of AMP-PNP. (D) Fit of the solvated structure of *E. coli* UvrD in the envelope of HelD:AMP-PNP. (E) Fit of the solvated structure of RapA in the envelope of HelD: AMP-PNP. (F) Partial model of HelD fitted into the SAXS envelope of HelD:AMP-PNP; the approximate position is based on the fit of UvrD and RapA in panels D and E, respectively.

Fig. 4. Truncated HelDΔN lacking the N-terminal domain has still the ATPase activity and ability to form complex with RNAP and bind DNA but cannot stimulate transcription. (A) HelDΔN has the same ATPase activity as full-length HelD. The ATPase assay was done in the absence (indicated by -) and presence (indicated by +) of AMP-PNP as competitive inhibitor, in triplicates. Activity is quantified using the amount of PO_4 (in μmol) released from ATP by 1 μmol of the enzyme at 25 °C in 1 min. (B) Gel-shift assay under native conditions using non-denaturing PAGE. Lanes 1 and 5, RNAP expressed in the strain of *B. subtilis* lacking the ability to produce HelD (LK782, Table 1); lane 2, HelD; lanes 3 and 4, RNAP mixed and incubated with HelD at molar ratios of 1:1 and 1:2, respectively; lane 6, HelDΔN; lanes 7 and 8, RNAP mixed and incubated with HelDΔN at molar ratios 1:1 and 1:2. (C) HelD and HelDΔN interact with radiolabeled DNA. DNA was prepared by PCR and the ratios of DNA: HelD/HelDΔN were 1:10, 1:20, 1:100, 1:400, and 1:800. As a control, the same ratios of heat-denaturated HelD/HelDΔN (with the exception of 1:20) were used (marked by symbol °C). (D) The N-terminal domain of HelD is required for stimulation of

transcription. Multiple round transcriptions were performed with *B. subtilis* RNAPσ^{A} from a plasmid containing the *B. subtilis rrnB* P1 promoter. Transcription in the presence of HelD or HelDΔN is indicated; transcription in the absence of any factor is marked as (-) and was set as 1. The experiment was performed three times, the bars show the average values and the error bars ±SD. While the full-length HelD showed stimulation of transcription as we reported previously [9], the stimulation level with the HelDΔN product was significantly lower – close to the background level without HelD.