Solution structure of domain 1.1 of the $\sigma^A$ factor from *Bacillus subtilis* is preformed for binding to the RNA polymerase core

Received for publication, March 13, 2017, and in revised form, May 19, 2017. Published, Papers in Press, May 24, 2017. DOI 10.1074/jbc.M117.784074

Milan Zachrdla‡§, Petr Padra‡§, Alžbeta Rabatinová, Hana Šanderová, Ivan Barvík, Libor Krásný†, and Lukáš Žídek‡§

From the ‡Central European Institute of Technology and the §National Centre for Biomolecular Research (NCBR), Faculty of Science, Masaryk University, CZ-62500 Brno, Czech Republic, the ¶Laboratory of Microbial Genetics and Gene Expression, Institute of Microbiology, The Czech Academy of Sciences, CZ-14220 Prague 4, Czech Republic, and the ¶¶Division of Biomolecular Physics, Institute of Physics, Faculty of Mathematics and Physics, Charles University, CZ-12116 Prague 2, Czech Republic

Edited by Norma Allmell

Bacterial RNA polymerase (RNAP) requires $\sigma$ factors to recognize promoter sequences. Domain 1.1 of primary $\sigma$ factors ($\sigma 1.1$) prevents their binding to promoter DNA in the absence of RNAP, and when in complex with RNAP, it occupies the DNA-binding channel of RNAP. Currently, two 3D structures of $\sigma 1.1$ are available: from *Escherichia coli* in complex with RNAP and from *Thermoplasma maritima* solved free in solution. However, these two structures significantly differ, and it is unclear whether this difference is due to an altered conformation upon RNAP binding or to differences in intrinsic properties between the proteins from these two distantly related species. Here, we report the solution structure of $\sigma 1.1$ from the Gram-positive bacterium *Bacillus subtilis*. We found that *B. subtilis* $\sigma 1.1$ is highly compact because of additional stabilization not present in $\sigma 1.1$ from the other two species and that it is more similar to *E. coli* $\sigma 1.1$. Moreover, modeling studies suggested that *B. subtilis* $\sigma 1.1$ requires minimal conformational changes for accommodating RNAP in the DNA channel, whereas *T. maritima* $\sigma 1.1$ must be rearranged to fit therein. Thus, the mesophilic species *B. subtilis* and *E. coli* share the same $\sigma 1.1$ fold, whereas the fold of $\sigma 1.1$ from the thermophile *T. maritima* is distinctly different. Finally, we describe an intriguing similarity between $\sigma 1.1$ and $\delta$, an RNAP-associated protein in *B. subtilis*, bearing implications for the so-far unknown binding site of $\delta$ on RNAP. In conclusion, our results shed light on the conformational changes of $\sigma 1.1$ required for its accommodation within bacterial RNAP.

Transcription of DNA into RNA is an essential cellular process. It is mediated by DNA-dependent RNA polymerase (RNAP). RNAP is a multisubunit enzyme, and in bacteria the RNAP core is composed of five subunits (2 $\alpha$, $\beta$, $\beta^\prime$, $\omega$). Gram-positive Firmicutes contain two additional subunits, $\delta$ and $\epsilon$ (1). A number of functions have been ascribed to $\delta$, including affecting affinity of RNAP for DNA (2) and for initiating nucleotide triphosphates (3) or functioning as a transcription factor (4). The absence of $\delta$ decreases competitive fitness and virulence (3, 5). The effects of $\delta$ are augmented by HsdD, a helicase-like protein associating with RNAP (6). The function of $\epsilon$ is less clear (6, 7). The RNAP core is capable of transcription elongation, but it is not able to bind to promoter DNA, where transcription begins. For RNAP to bind to promoter DNA, the presence of a $\sigma$ factor is necessary. Upon a $\sigma$ factor binding to the RNAP core, the resulting RNAP holoenzyme is capable of recognizing promoter sequences and subsequently initiating transcription (8). Preventing the binding of $\sigma$ factors to the RNAP core is also used as a strategy for development of novel antibacterial compounds (9).

Different bacterial species contain different numbers of different $\sigma$ factors, ranging from one per species to more than 100 (10). According to their structure, $\sigma$ factors are divided into two principally distinct groups: the $\sigma^{70}$ and $\sigma^{44}$ families. Factors from the $\sigma^{44}$ family have no sequence similarity with $\sigma^{70}$ factors, and they require the binding of ATP-dependent activators (11). Factors from the $\sigma^{70}$ family are present in all bacterial species, and no ATP-dependent activators are needed. The $\sigma^{70}$ family is further subdivided into four groups (groups 1–4), based on domain composition. Group 1 contains vegetative $\sigma$ factors ($\sigma^{70}$ in *Escherichia coli*, $\sigma^A$ in *Bacillus subtilis*) essential for transcription of housekeeping genes. Groups 2–4 contain structurally related alternative $\sigma$ factors responsible for transcription of genes whose expression is important during various environmental stresses (10, 11).

The vegetative $\sigma$ factors (group 1) contain four domains: domain 1.1, domain 2 (regions 1.2–2.4), domain 3 (regions 3.0–
3.2), and domain 4 (regions 4.1–4.2). Regions 2.4 (domain 2) and 4.2 (domain 4) recognize the −10 and −35 promoter consensus hexamer sequences, respectively, that are critical for the initial RNAP-DNA binding (closed complex) and subsequent formation of the transcription bubble, the so-called open complex. Region 1.2 (domain 2) interacts with the DNA region between the transcription start site (+1) and the −10 hexamer and affects the stability of the open complex. Domain 3 binds to the −10 extended motif (TGx). This motif precedes the −10 hexamer, and it is not present in all promoters. When it is present, however, it increases the promoter affinity for RNAP and boosts transcription (10).

When σ^A binds to RNAP that is not in complex with DNA, σ1.1 occupies the DNA-binding channel (12, 13). Further, σ1.1 plays a specific role in autoregulation of the σ factor. It inhibits the binding of the σ factor to promoter DNA sequence alone. Trans-binding experiments with the E. coli σ1.1 region and truncated σ^70 variants suggested that σ1.1 binds to domain 4 (region 4.2) in free state (14). Furthermore, examination of possible interdomain interactions of Thermotoga maritima σ factor as studied by an interdomain cross-linking approach suggested that σ1.1 is in a close proximity to domains 2 and 4 (15). Molecular details of these interactions, however, are still elusive.

Currently, two structures of σ1.1 are available: one from T. maritima solved by NMR (15) and the other from E. coli solved by crystallography in complex with RNAP (12). Despite sequence similarities in σ1.1 in these two organisms, their 3D structures differ. In both organisms, this domain consists of three helices (HI–HIII) connected by two loops. However, although in T. maritima HI and HIII are roughly anti-parallel to one another and pack perpendicularly against HI, in E. coli the three helices show anti-parallel packing, leading to a distinctly different morphology. Here, to provide a basis for a better understanding of the structural diversity among σ domains 1.1 from different species that could have implications for their binding to RNAP, we solved the solution structure of σ^A domain 1.1 from the model soil-dwelling Gram-positive bacterium B. subtilis.

## Results

### σ1.1 structure determination

We decided to solve the structure of σ1.1 by NMR because of its small size (9.4 kDa) and the benefit of getting additional information on flexibility of the protein. Despite the high occurrence of Glu and Gln residues in the σ1.1 sequence, an almost complete backbone and side-chain assignment was obtained. As expected, overlapped peaks from the His tag and missing peaks from the N-terminal residue were not assigned; otherwise only two backbone and four side-chain chemical shifts remained unassigned. Spectra measured on a diluted sample confirmed the monomeric state of the sample. NOE assignment yielded 1886 unambiguous 1H–1H distances, including 486 long-range NOEs. Additional restraints, 21 3J_{HNHA} and 86 RDC values, were used for the residues that were predicted to form α-helices. The calculated structure is in agreement with the secondary structure prediction from chemical shifts. Statistics of the structure calculation are presented in Table 1.

### Structure of σ1.1

The first 71 amino acids of B. subtilis σ1.1 form three helices (helix I, Phe12–Arg26; helix II, Tyr31–Phe41; and helix III, Ser45–Glu57) that are connected by two short loops (Fig. 1). The HI–HII loop is formed by amino acids Gly27–Thr30, and the HII–HIII loop is formed by amino acids Glu42–Glu44. The N terminus (amino acids Ala1–Thr11) and the C terminus (amino acids Gln58–Asp71) are mainly unstructured. An important exception is a small portion of the C terminus (amino acids Gln58–Asp71) which together with the HI–HII loop (residues 28–VLT30) forms a β-sheet motif that is stabilized via hydrogen bonding between the respective parts of the σ1.1 main chain. The total charge of the 71-amino acid B. subtilis σ1.1 is −15, i.e. −21 × (Glu or Asp) + 6 × (Lys or Arg). The frequency of negatively charged amino acids is increasing from the N terminus to the C terminus. Although Helix I is still slightly positively charged (+1), the remaining structural motifs display progressively increasing negative charge: the N terminus (−1), helix II (−2), HII–HIII (−2), helix III (−5), and the C terminus (−6). It is therefore clear that the B. subtilis σ1.1 can mimic a portion of downstream duplex DNA that is also strongly negatively charged because of the presence of phosphate groups in the phosphodiester bonds between nucleotides.

### 15N relaxation

NMR relaxation was used to probe the dynamics of the σ1.1 domain. A set of 15N relaxation rates, including R1 and R2 autorelaxation rates, steady-state [1H,15N] heteronuclear Overhauser enhancement (ssNOE), and longitudinal and transverse cross-correlated relaxation rates, was measured (16). The software relax was used to analyze the relaxation data in the model-free manner (17–20). The analysis showed that

---

**Table 1**

| Parameters | Number/value |
|------------|--------------|
| NMR assignment (%) | >99% |
| Backbone nuclei | >99% |
| Side-chain nuclei | |
| Restraints | |
| Total NOE | 1886 |
| Short range, | 925 |
| Medium range, 1 < | 475 |
| Long range, | 486 |
| 3J_{HNHA} | 21 |
| RDC | 86 |
| Average violations | 0.0 |
| NOE > 0.5 Å | 0.0 |
| 3J_{HNHA} > 3σ | 0.0 |
| RDC > 3σ | 0.0 |
| Root mean square deviation to the mean structure | 0.38 ± 0.10 Å |
| Backbone atoms | 1.01 ± 0.14 Å |
| Heavy atoms | |
| Ramachandran diagram statistics | |
| Most favourite region | 96.3% |
| Additional favored region | 3.6% |
| Generously allowed region | 0.1% |
| Disallowed region | 0.0% |

* Assignment statistics were obtained without including His tag.

* Restraints part only (residues 10–63) was analyzed for the final ensemble of 20 structures.

* Violations of scalar and residual dipolar couplings were evaluated using the standard deviation of the experimental data.
the molecule tumbles as a prolate spheroid, with the global correlation time \((1/6D_{iso})\) of 6.13 ns at 25 °C and with \(D_{||}/D_{\perp}\) of 2.23. Internal dynamics of individual residues was characterized by the order parameter \(S^2\) and the effective correlation time \(\tau_e\) (Fig. 2). The decrease of the order parameter \(S^2\) values, indicating more flexible regions, was observed in the loop between \(\alpha\)-helices II and III, as well as in the \(\beta\)-sheets. In many cases, introduction of a second mode of internal motion, described by an additional order parameter and correlation time, resulted in a statistically significant improvement of the fit. The correlation time of the second mode was well defined in the terminal regions (1-3 ns) but determined with a large uncertainty in the well ordered regions (Fig. 2). Also, the analysis provided significant exchange contribution (\(R_{ex}\)) for many residues (Fig. 2C), indicating a presence of a slow conformational exchange, sufficiently well described by a two-state model with a relatively uniform exchange rate of \(\sim 3,500\) Hz and with the minor state being populated by 1.5% (Fig. 3). The \(^{15}\)N chemical shift differences between both states varied from 2 to 6 ppm, with the most significant changes at the end of helix I, in helix II, and in \(\beta\)-sheet II.

**Position of \(\sigma1.1\) within the DNA-binding channel of RNAP**

In RNAP that is not bound to DNA, \(\sigma1.1\) is positioned inside the DNA-binding channel (12, 13). To gain insights into the position of \(B.\) subtilis \(\sigma1.1\) within \(B.\) subtilis RNAP, we carried out structural alignments using the software package Molsoft (<www.molsoft.com>). As a template we used the crystal structure of \(E.\) coli RNAP (PDB code 4LK1; Ref. 12). Our NMR structure of \(B.\) subtilis \(\sigma1.1\) and a previously published homology model of \(B.\) subtilis RNAP core (21) were structurally aligned with the template.

In this model, \(B.\) subtilis \(\sigma1.1\) occupies the downstream duplex DNA-binding channel with its center of gravity at...
Structure of domain σ1.1 from B. subtilis

Figure 3. Parameters of slow exchange calculated for a two-state model from CPMG relaxation dispersion data obtained at 600, 850, and 950 MHz spectrometers. A, the exchange rate, $k_{ex}$. B, the population of the minor state, $p_B$. C, the chemical shift difference between the states, $\Delta \omega$. Residues with standard deviation exceeding 100% are shown in gray. Sequence and secondary structure elements are shown above the graph.

Discussion

We have determined the solution structure of σ1.1, the N-terminal domain of the primary σ factor, σ70, from B. subtilis. For a long time, the structure of this domain had not been available because of its flexibility until it was solved by NMR for T. maritima (15), and, several years later, also by crystallography for E. coli (12). In the crystal structure, σ1.1 is a part of σ70, in a context of a complex with RNAP. The two known structures (E. coli and T. maritima) differ significantly. In the following text, we provide detailed comparisons of B. subtilis σ1.1 with these two structures. The comparisons shed light on interactions and conformational changes of σ1.1 required for its accommodation within RNAP.

σ1.1 sequence comparisons

The σ1.1 domains from B. subtilis, E. coli, and T. maritima displayed a modest degree of sequence similarity except for the non-conserved N terminus (residues 1–30) of T. maritima (B. subtilis and E. coli lack this fragment). Pairwise alignments of respective sequences from B. subtilis, E. coli, and T. maritima revealed a sequence identity of ~25%. Multiple alignment of these sequences then yielded a sequence identity of only ~10%. Nevertheless, it should be noted that in the case of Glu/Asp amino acid residues, which we believe are critical for the binding of σ1.1 to the RNAP core, there are numerous point substitutions that do not change charge (either Glu to Asp or Asp to Glu, seven occasions for T. maritima versus B. subtilis and E. coli; Fig. 1). Moreover, these amino acids are often shifted just by one position in the σ1.1 sequences (five occasions for T. maritima versus B. subtilis and E. coli; Fig. 1). These evolutionary differences mean that the sequence similarity is significantly greater than

approximately +8 (+1 is the transcription start site), where it must be displaced by the DNA upon formation of the open promoter complex. It is wedged in the RNAP channel, interacting with the β subunit (i.e. amino acids Asp151, Arg183, Arg188, Arg241, and Arg248) and with structural elements of the β′ subunit, namely the β′ clamp (Ile110 and Arg127), the rudder around residue 301, and two amino acid residues of the β′-pincer (Lys125 and Arg144). Many salt bridges stabilizing B. subtilis σ1.1 in the downstream DNA-binding channel of the RNAP core can be predicted based on our model, involving helices I and II and especially helix III (core RNAP/σ1.1): Asp151/Arg246, Arg183/Glu36, Arg188/Glu42, Arg241/Glu42, Arg498/Glu55, Ile110/Phε44, Arg127/Glu50, Arg200/Arg303/Glu44, Glu50/Asp46, Lys1125, and Arg144/Glu57 (Fig. 4A). (The spaces between some residues indicate that the residues have more than one possible salt bridge partner.)

Although the position of the structured part of B. subtilis σ1.1 within the DNA channel can be well predicted, the position of the unstructured C terminus of the B. subtilis σ1.1 (the linker to σ1.2) can be only roughly approximated. Nevertheless, it is apparent that the C terminus of σ1.1 would interact with mobile and functionally important parts of the RNAP core, namely with the bridge helix (responsible for the DNA–RNA translocation), and with the trigger loop (opening or closing access of NTP into the active site of RNAP through the secondary channel). The C terminus of the B. subtilis σ1.1 contains many negatively charged amino acids (for example Glu67, Glu68, Glu72, Asp71, and Glu73) that could create additional salt bridges with a number of positively charged amino acid residues in the bridge helix (Arg784, Lys785, Lys793, Arg802, and Arg803) and in the trigger loop (Arg937, Arg953, and Arg963; Fig. 4A).

Discussion

We have determined the solution structure of σ1.1, the N-terminal domain of the primary σ factor, σ70, from B. subtilis. For a long time, the structure of this domain had not been available because of its flexibility until it was solved by NMR for T. maritima (15), and, several years later, also by crystallography for E. coli (12). In the crystal structure, σ1.1 is a part of σ70, in a context of a complex with RNAP. The two known structures (E. coli and T. maritima) differ significantly. In the following text, we provide detailed comparisons of B. subtilis σ1.1 with these two structures. The comparisons shed light on interactions and conformational changes of σ1.1 required for its accommodation within RNAP.

σ1.1 sequence comparisons

The σ1.1 domains from B. subtilis, E. coli, and T. maritima displayed a modest degree of sequence similarity except for the non-conserved N terminus (residues ~1–30) of T. maritima (B. subtilis and E. coli lack this fragment). Pairwise alignments of respective sequences from B. subtilis, E. coli, and T. maritima revealed a sequence identity of ~25%. Multiple alignment of these sequences then yielded a sequence identity of only ~10%. Nevertheless, it should be noted that in the case of Glu/Asp amino acid residues, which we believe are critical for the binding of σ1.1 to the RNAP core, there are numerous point substitutions that do not change charge (either Glu to Asp or Asp to Glu, seven occasions for B. subtilis versus E. coli; Fig. 1). Moreover, these amino acids are often shifted just by one position in the σ1.1 sequences (five occasions for T. maritima versus B. subtilis and E. coli; Fig. 1). These evolutionary differences mean that the sequence similarity is significantly greater than

approximately +8 (+1 is the transcription start site), where it must be displaced by the DNA upon formation of the open promoter complex. It is wedged in the RNAP channel, interacting with the β subunit (i.e. amino acids Asp151, Arg183, Arg188, Arg241, and Arg248) and with structural elements of the β′ subunit, namely the β′ clamp (Ile110 and Arg127), the rudder around residue 301, and two amino acid residues of the β′-pincer (Lys125 and Arg144). Many salt bridges stabilizing B. subtilis σ1.1 in the downstream DNA-binding channel of the RNAP core can be predicted based on our model, involving helices I and II and especially helix III (core RNAP/σ1.1): Asp151/Arg246, Arg183/Glu36, Arg188/Glu42, Arg241/Glu42, Arg498/Glu55, Ile110/Phε44, Arg127/Glu50, Arg200/Arg303/Glu44, Glu50/Asp46, Lys1125, and Arg144/Glu57 (Fig. 4A). (The spaces between some residues indicate that the residues have more than one possible salt bridge partner.)

Although the position of the structured part of B. subtilis σ1.1 within the DNA channel can be well predicted, the position of the unstructured C terminus of the B. subtilis σ1.1 (the linker to σ1.2) can be only roughly approximated. Nevertheless, it is apparent that the C terminus of σ1.1 would interact with mobile and functionally important parts of the RNAP core, namely with the bridge helix (responsible for the DNA–RNA translocation), and with the trigger loop (opening or closing access of NTP into the active site of RNAP through the secondary channel). The C terminus of the B. subtilis σ1.1 contains many negatively charged amino acids (for example Glu67, Glu68, Glu72, Asp71, and Glu73) that could create additional salt bridges with a number of positively charged amino acid residues in the bridge helix (Arg784, Lys785, Lys793, Arg802, and Arg803) and in the trigger loop (Arg937, Arg953, and Arg963; Fig. 4A).

Discussion

We have determined the solution structure of σ1.1, the N-terminal domain of the primary σ factor, σ70, from B. subtilis. For a long time, the structure of this domain had not been available because of its flexibility until it was solved by NMR for T. maritima (15), and, several years later, also by crystallography for E. coli (12). In the crystal structure, σ1.1 is a part of σ70, in a context of a complex with RNAP. The two known structures (E. coli and T. maritima) differ significantly. In the following text, we provide detailed comparisons of B. subtilis σ1.1 with these two structures. The comparisons shed light on interactions and conformational changes of σ1.1 required for its accommodation within RNAP.

σ1.1 sequence comparisons

The σ1.1 domains from B. subtilis, E. coli, and T. maritima displayed a modest degree of sequence similarity except for the non-conserved N terminus (residues ~1–30) of T. maritima (B. subtilis and E. coli lack this fragment). Pairwise alignments of respective sequences from B. subtilis, E. coli, and T. maritima revealed a sequence identity of ~25%. Multiple alignment of these sequences then yielded a sequence identity of only ~10%. Nevertheless, it should be noted that in the case of Glu/Asp amino acid residues, which we believe are critical for the binding of σ1.1 to the RNAP core, there are numerous point substitutions that do not change charge (either Glu to Asp or Asp to Glu, seven occasions for B. subtilis versus E. coli; Fig. 1). Moreover, these amino acids are often shifted just by one position in the σ1.1 sequences (five occasions for T. maritima versus B. subtilis and E. coli; Fig. 1). These evolutionary differences mean that the sequence similarity is significantly greater than

approximately +8 (+1 is the transcription start site), where it must be displaced by the DNA upon formation of the open promoter complex. It is wedged in the RNAP channel, interacting with the β subunit (i.e. amino acids Asp151, Arg183, Arg188, Arg241, and Arg248) and with structural elements of the β′ subunit, namely the β′ clamp (Ile110 and Arg127), the rudder
the sequence identity. Taking into account that their interaction counterparts in salt bridges are positively charged amino acids with long flexible side chains (either Lys or Arg), such subtle variations are unlikely to affect the $\alpha1.1$ function(s). The highest sequence conservation is around the HI–HII loop region (Leu$^{19}$, Gly$^{23}$, Lys$^{24}$, Gly$^{27}$, Thr$^{30}$, and Tyr$^{31}$).

**1.1 structure comparisons**

The 3D solution structure of *B. subtilis* $\alpha1.1$ resembles the crystal structure of *E. coli* $\alpha1.1$, which was obtained in the context of RNAP (Fig. 5). Despite the similarities, helix I in *B. subtilis* $\alpha1.1$ is slightly longer than helix I in *E. coli* $\alpha1.1$ (12). Furthermore, several interactions that contribute to anchoring Helix I to the rest of the structure of *B. subtilis* $\alpha1.1$ are missing in *E. coli* $\alpha1.1$. These interactions are mediated by Phe$^{12}$, whose bulky side chain is nestled between side chains of Met$^{38}$, Phe$^{41}$, and Ile$^{43}$ of the HII–HIII loop (Fig. 1).

The *B. subtilis* and *E. coli* structures then markedly differ from that of one of *T. maritima* where HI packs perpendicularly to HII and HIII. This contrasts with the all-anti-parallel packing of helices in *B. subtilis* and *E. coli* $\alpha1.1$ (Fig. 5). It should be noted that HI from *T. maritima* $\alpha1.1$ is by far the longest one, and it is also preceded by the non-conserved and unstructured N terminus.

**$\alpha1.1$ interactions with RNAP**

Predicted salt bridges between *B. subtilis* $\alpha1.1$ and the DNA channel have corresponding analogous interactions in the structure of *E. coli* RNAP (PDB code 4LK1; Ref. 12; compare Fig. 4, A and B). The compact structure of *B. subtilis* $\alpha1.1$ is undoubtedly important for its optimal interactions with the RNAP core in the downstream DNA-binding channel (especially salt bridges formed by Glu$^{57}$, Glu$^{65}$, Arg$^{36}$, and Glu$^{42}$; Fig. 4A). The compact structure likely compensates for the absence of some additional parts of RNAP (including those of $\alpha1.1$) that are found uniquely in either *E. coli* or *T. maritima* and that participate in positioning of $\alpha1.1$ in the DNA channel by reducing the breathing movements of RNAP cleft arms. The “extra” part of *E. coli* RNAP consists of a large insertion (i.e. amino acids Gly$^{938}$–Thr$^{1131}$; see PDB code 4YLN; Ref. 22) in the trigger loop that regulates access of NTPs into the RNAP active site via the secondary channel. The extra part of *T. maritima* RNAP is the relatively long N terminus of its $\alpha1.1$, which was unstructured in the solution structure of isolated $\alpha1.1$ (15).

Remarkably, HI of *T. maritima* $\alpha1.1$ apparently does not fit into the downstream DNA channel of RNAP (supplemental Fig. S1), whereas HI of *B. subtilis* $\alpha1.1$ fits this space smoothly (Fig. 4C and supplemental Fig. S1). It indicates that HI of
T. maritima σ1.1 likely undergoes a conformational change to be accommodated into the RNAP core. Interestingly, the structure of T. maritima σ1.1 is very similar to the structured N-terminal part of the σ subunit of B. subtilis RNAP, consisting of four α-helices (helices Ia, Ib, II, and III, formed by residues Gln8–Lys12, Leu16–His27, Phe33–Leu44, and Gly52–Asn63, respectively). In fact, helices Ia and Ib, II, and III of δ correspond to HI, HII, and HIII of T. maritima σ1.1, respectively (Fig. 5 and supplemental Fig. S2). In addition, σ contains a short anti-parallel β-sheet composed of three short β-strands (residues Val31–Pro32, Phe68–Ala70, and Thr75–Leu78) at the top of a “twisted tripod” formed by helices Ib, II, and III; Refs. 23–25; Fig. 5). This motif structurally overlaps with the short two-strand β-sheet found in our NMR structure of B. subtilis σ1.1. Future experiments will have to reveal whether the apparent structure similarity between σ1.1 and σ may provide clues for identifications of the so-far unknown binding site of σ on RNAP.

In conclusion, the determined solution structure of B. subtilis σ1.1 showed for the first time a preformed 3D conformation that requires minimal, if any, conformational changes to be accommodated within the DNA-binding channel. Moreover, the NMR relaxation revealed that the determined structure of σ1.1 is in a slow exchange with a minor state, differing mostly in the helix HI and its proximity. One can speculate that the minor state may resemble the solved structure of σ1.1 from T. maritima. However, further experiments are needed to test this hypothesis.

Experimental procedures

Sample preparation

σ1.1 was prepared using a standard protocol including cloning, expression, and purification methods to produce a 13C,15N-uniformly labeled sample in a sufficient concentration. The truncated gene of the σ1.1 coding only its σ1.1 part (amino acids 1–71) was cloned into a pET28b vector between NcoI and XhoI sites. The additional six histidine residues at the C terminus served as the His tag facilitating the protein purification process, and the two residues preceding the His tag (LE) were inserted because of the restriction enzyme (XhoI) used for cloning.

NMR measurements

All NMR experiments were performed at 25 °C using 0.8 mM 13C,15N-labeled sample or 0.6 mM 15N-labeled sample. Temperature was calibrated according to the chemical shift differences of pure methanol peaks.

Resonance assignment was done using experiments acquired on 700 MHz Bruker Avance III spectrometer equipped with the TXO cryogenic probehead with z axis gradients and 850 MHz Bruker Avance III spectrometer equipped with the TCI cryogenic probehead with z axis gradients. Standard set of 3D triple-resonance experiments HNCA, HN(CO)CA, HNCO, HNCACB, and CBCA(CO)NH (26) was used for backbone assignment. Multiple experiments for side-chain assignment, HCCH-TOCSY, TOCSY-HSQC,13C-edited aromatic NOESY-HSQC, H(CC)(CO)NH, and (H)CC(CO)NH (26) were employed to overcome chemical shift degeneracies caused by a high occurrence of Glu and Gln residues in the sequence. Assignment of all obtained chemical shifts was deposited in the BioMagResBank (http://www.bmrb.wisc.edu; Ref. 27) under accession code 34089.

Protein structure calculation was based on 1H-1H distance restraints obtained from 15N-edited NOESY-HSQC and 13C-edited NOESY-HSQC (both aromatic and aliphatic spectral regions) experiments (26). In addition, three-bond scalar couplings obtained from HNHA experiment (28) were used to determine φ torsion angles of the protein backbone. 1D(HN¹), 1D(Cα²), 1D(NC³), and 2D(H⁵C) RDCs were obtained from 1H,15N-IPAP, HN[C]-3E, and 13C-detected (H)CACO-IPAP experiments run on the isotropic sample and on a sample partially oriented in 5% polyacrylamide gel. Program S3EPY (29) was used to evaluate RDCs obtained from the measured spectra. The (non-uniform) RDC errors were estimated from 1D lineshape using Cramer–Rao lower bound theory (30). The spectra were processed using the program NMRPipe 8.1 (31).
and analyzed using the program Sparky 3.115 (T. D. Goddard and D. G. Kneller, University of California, San Francisco, CA). Secondary structure prediction utilizing \(^{15}\text{C}^\text{N}, \, ^{13}\text{C}^\text{P}, \, ^{13}\text{C}^\text{K}, \, ^1\text{H}, \, ^{15}\text{N}\) chemical shifts was done using program ssp 1.0 (32). Automated assignment of NOESY spectra was performed using CANDID (33), an algorithm included in the program CYANA (34). CNS 1.2 (35) was used for refinement of structures in water using RECOORD scripts modified for our needs (36). Module TENSO (37) was used to include RDCs into structure calculation. Both \(^{3}\text{J}_{\text{HNHA}}\) and RDCs were used only for residues in \(\alpha\)-helices. 300 initial structures were calculated, and 150 structures with the lowest energy were further refined using an explicit water model. The final ensemble of 20 structures with the lowest energy was validated using program CING (38, 39) and deposited in the Protein Data Bank (www.rcsb.org; Ref. 40) under code 5MWV.

NMR relaxation experiments were performed at a 600 MHz Bruker Avance III spectrometer equipped with the QCI cryogenic probehead with \(z\) axis gradients, at a 850 MHz Bruker Avance III spectrometer equipped with the TCI cryogenic probehead with \(z\) axis gradients, and at a 950 MHz Bruker Avance III spectrometer equipped with the TCI cryogenic probehead with \(z\) axis gradients. The overall number of 2048 complex points was acquired in the acquisition dimension and 320 complex points were acquired in the indirect dimension for autorelaxation rates \(R_1, R_2\) and steady-state \(^{15}\text{N}^{-1}\text{H}\) nuclear Overhauser effect. Standard experiments (16) were used for the measurement of \(R_1\) (relaxation delays 11.1, 55.5, 133.2, 233.1, 377.4, 555, 888, and 1942.5 ms) and \(R_2\) with the delay between the 180 pulses in the CPMG train equal to 0.83 ms (relaxation delays 0, 14.4, 28.8, 43.2, 57.6, 72, and 86.4 ms). The asterisks denote spectra recorded twice to estimate experimental error. The ssNOE values were measured under a steady-state condition, achieved by a 5-ms \(^1\text{H}\) irradiation with 226 repeats of 2000-\(\mu s\) 180°\(^1\text{H}\) pulses (41), separated by 22.22-ms delays, and with a 30-s interscan relaxation delay. Reference spectra and the spectra measured under steady-state conditions were measured in an interleaved manner.

Transverse cross-correlated relaxation rates \(\Gamma_x\) (relaxation delays 30, 50, and 70 ms) and longitudinal cross-correlated relaxation rates \(\Gamma^*_x\) (relaxation delays 100, 175, and 250 ms) were obtained using experiments based on symmetrical reconstruction (42, 43). Programs relax (44–47) and Octave (48) were used to obtain relaxation rates by fitting peak intensities to a monoexponential decay.

Consistency of the relaxation data were tested as proposed by Morin and Gagné (49). The field independent \(J(0)\) values obtained from the autocorrelated and cross-correlated relaxation rates by the spectral density mapping show that the data were not biased by imperfect temperature calibration during the measurements at different magnetic fields. The model-free analysis (17–19) of autorelaxation rates from 600 MHz, 850 MHz, and 950 MHz was performed using the d’Auvergne protocol in the software relax. The prolate diffusion model was selected based on the lowest value of Akaike’s information criterion (50). The selected model was then used to estimate errors of calculated parameters by performing 500 Monte Carlo simulations. Relaxation rates and model-free order parameters were deposited in the BioMagResBank under accession number 27011. Relaxation dispersion CPMG experiments with CPMG frequency ranging from 111 to 2000 Hz (51) were acquired and analyzed using the relaxation dispersion autoanalysis in the software relax (20) to study the slow exchange.

**In silico modeling**

Structural alignments were produced using the ICM Molsoft software package (www.molsoft.com).\(^4\) Amino acids sequences were aligned using the Clustal Omega web server (http://www.ebi.ac.uk/Tools/msa/clustalo/).

**Author contributions**—M. Z., I. K., and L. Z. conceived and designed the research; A. R., H. S., and L. K. prepared the samples; M. Z., P. P., and L. Z. acquired and analyzed the NMR data and solved the structure; I. B. did the *in silico* modeling; and M. Z., I. B., L. K., and L. Z. wrote the manuscript. All authors reviewed and contributed to the manuscript.

**Acknowledgments**—CIISB research infrastructure project LM2015043 funded by the Ministry of Education, Youth and Sports of the Czech Republic is gratefully acknowledged for a partial financial support of the measurements at the Josef Dadok National NMR Centre, CEITEC, Masaryk University.

**References**

1. Weiss, A., and Shaw, L. N. (2015) Small things considered: the small accessory subunits of RNA polymerase in Gram-positive bacteria. *FEMS Microbiol. Rev.* 39, 541–554
2. Lópezde Sario, F. J., Woody, A.-Y., and Helmann, J. D. (1995) Structural analysis of the *Bacillus subtilis* \(\delta\) factor: a protein polyanion which displaces RNA from RNA polymerase. *J. Mol. Biol.* 252, 189–202
3. Rabatínová, A., Šanderová, H., Jiráň Matejčková, J., Korelusová, J., Sojka, L., Barvík, I., Papousková, V., Šlenkař, V., Žídek, L., and Krásný, L. (2013) The \(\delta\) subunit of RNA polymerase is required for rapid changes in gene expression and competitive fitness of the cell. *J. Bacteriol.* 195, 2603–2611
4. Prajapati, R. K., Sengupta, S., Rudra, P., and Mukhopadhyay, J. (2016) *Bacillus subtilis* \(\delta\) function as a transcriptional regulator by facilitating the open complex formation. *J. Biol. Chem.* 291, 1064–1075
5. Weiss, A., Ibarra, J. A., Paletti, J., Carroll, R. K., and Shaw, L. N. (2014) The \(\delta\) subunit of RNA polymerase guides promoter selectivity and virulence in *Staphylococcus aureus*. *Infect. Immun.* 82, 1424–1435
6. Wiedermannová, J., Sudzínová, P., Koval', T., Rabatínová, A., Šanderová, H., Ramaník, O., Rittich, S., Dohnálek, J., Fu, Z., Halada, P., Lewis, P., and Krásný, L. (2014) Characterization of helD, an interacting partner of RNA polymerase from *Bacillus subtilis*. *Nucleic Acids Res.* 42, 5151–5163
7. Keller, A. N., Yang, X., Wiedermannová, J., Delumeau, O., Krásný, L., and Lewis, P. J. (2014) \(\epsilon\), a new subunit of RNA polymerase found in Gram-positive bacteria. *J. Bacteriol.* 196, 3622–3632
8. Murakami, K. S., and Darst, S. A. (2003) Bacterial RNA polymerases: the whole story. *Curr. Opin. Struct. Biol.* 13, 31–39
9. Ma, C., Yang, X., Kandemir, H., Mielczarek, M., Johnston, E. B., Griffith, R., Kumar, N., and Lewis, P. J. (2013) Inhibitors of bacterial transcription initiation complex formation. *ACS Chem. Biol.* 8, 1972–1980
10. Paget, M. S. (2015) Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biomolecules* 5, 1245–1265
11. Österberg, S., del Paso-Santos, T., and Shingler, V. (2011) Regulation of alternative sigma factor use. *Proc. Natl. Acad. Sci. U.S.A.* 110, 19772–19777
12. Murakami, K. (2013) X-ray crystal structure of *Escherichia coli* RNA polymerase \(\sigma^{32}\) holoenzyme. *J. Biol. Chem.* 288, 9126–9134
14. Dombroski, A. J., Walter, W. A., and Gross, C. A. (1993) Amino-terminal amino acids modulate α-factor DNA-binding activity. Genes Dev. 7, 2446–2455

15. Schwartz, E. C., Shekhtman, A., Dutta, K., Pratt, M. R., Cowburn, D., Darst, S., and Muir, T. W. (2008) A full-length group 1 bacterial sigma factor adopts a compact structure incompatible with DNA binding. Chem. Biol. 15, 1091–1103

16. Korzhnev, D., Billette, M., Arseniev, A., and Orekhov, V. (2001) NMR studies of brownian tumbling and internal motions in proteins. Prog. Nucleic Magn. Reson. Spectrosc. 38, 197–266

17. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules: 1. theory and range of validity. J. Am. Chem. Soc. 104, 4546–4559

18. Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C., and Gronenborn, A. M. (1990) Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. J. Am. Chem. Soc. 112, 4989–4991

19. Morin, S., Linnet, T. E., Lescanne, M., Schanda, P., Cot, M., Lescanne, M., MacDougall, I. J., Lewis, P. J., and Griffith, R. (2005) Homology modeling of RNA polymerase and associated transcription factors from Bacillus subtilis. J. Mol. Graphics Model. 23, 297–303

20. Zuo, Y., and Steitz, T. (2015) Crystal structures of the E. coli transcription initiation complex with a complete bubble. Mol. Cell 58, 534–540

21. Chen, X., Miguez, J., Komárek, J., Kaderávek, P., Otrusinová, O., Rabatinová, A., Krásný, L., Zídek, Sklenár, L. V., and Wimmerová, M. (2014) X-ray vs. NMR structure of N-terminal domain of δ-subunit of RNA polymerase. J. Struct. Biol. 187, 174–186

22. Papoušková, V., Kaderávek, P., Otrusinová, O., Rabatinová, A., Šanderová, Nováček, H. J., Krásný, L., Sklenár, V., and Žídek, L. (2013) Structural study of the partially disordered full-length δ-subunit of RNA polymerase from Bacillus subtilis. ChemBioChem 14, 1772–1779

23. Motáčková, V., Šanderová, H., Žídek, L., Nováček, J., Padrta, P., Švenková, A., Korelusová, J., Jonáčí, J., Krásný, L., and Švenková, V. (2010) Solution structure of the N-terminal domain of Bacillus subtilis δ subunit of RNA polymerase and its classification based on structural homologs. Proteins 78, 1807–1810

24. Sattler, M., Schleucher, J., and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog. Nucleic Magn. Reson. Spectrosc. 34, 93–158

25. Ulrich, E. L., Akutsu, H., Doreleijers, J. F., Harano, Y., Ioannidis, Y. E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z., Nakatani, E., Schulte, C. F., Tolmie, D. E., Kent Wenger, R., Yao, H., et al. (2008) BioMagResBank. Nucleic Acids Res. 36, D462–D468

26. Vuister, G. W., and Bax, A. (1993) Quantitative J correlation: a new approach for measuring homonuclear three-bond (H{H}) coupling constants in N2-enriched proteins. J. Am. Chem. Soc. 115, 7772–7777

27. Novák, P., Žídek, L., Motáčková, V., Padrta, P., Švenková, A., Nuzillard, J.-M., Krásný, L., and Sklenár, V. (2010) S3EPY: A Sparky extension for determination of small scalar couplings from spin-state-selective excitation NMR experiments. J. Biomol. NMR 46, 191–197

28. Hagen, N., Kupinski, M., and Deneriak, E. L. (2007) Gaussian profile estimation in one dimension. Applied Optics 46, 5374–5383

29. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6, 277–293

30. Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006) Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: implications for fibrillation. Protein Sci. 15, 2795–2804

31. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J. Mol. Biol. 319, 209–227

32. Güntert, P., and Buchner, L. (2015) Combined automated NOE assignment and structure calculation with CYANA. J. Biomol. NMR 62, 453–471

33. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D 54, 905–921

34. Jansen, S., Chmelík, J., Žídek, L., Padrta, P., Novák, V., Zdráhal, Z., Píckmon, J. F., Löfstedt, C., and Sklenár, V. (2007) Structure of Bombyx mori chemosensory protein 1 in solution. Arch. Insect Biochem. Physiol. 66, 135–145

35. Sattler, M., Schleucher, J., and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog. Nucleic Magn. Reson. Spectrosc. 34, 93–158

36. Pelupessy, P., Espallargas, G. M., and Bodenhausen, G. (2003) Symmetrical reconvension: measuring cross-correlation rates with enhanced accuracy. J. Magn. Reson. 161, 258–264

37. Herrmann, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. Nucleic Acids Res. 28, 235–242

38. Ferrage, F., Cowburn, D., and Ghose, R. (2009) Accurate sampling of high-frequency motions in proteins by steady-state {15}N-{1}H nuclear Overhauser effect measurements in the presence of cross-correlated relaxation. J. Am. Chem. Soc. 131, 6048–6049

39. LeNer, D., Liu, M., and Yang, D. (2008) Accurately probing slow motions on millisecond timescales with a robust NMR relaxation experiment. J. Biomol. NMR 41, 905–921

40. d’Avergne, E. J., Gooley, P. R. (2006) Model-free model elimination: a new step in the model-free dynamic analysis of NMR relaxation data. J. Biomol. NMR 35, 117–135

41. d’Avergne, E. J., Gooley, P. R. (2007) Set theory formulation of the model-free problem and the diffusion seeded model-free paradigm, Mol. BioSyst. 3, 483–494

42. d’Avergne, E. J., Gooley, P. R. (2008) Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. J. Biomol. NMR 40, 107–119

43. d’Avergne, E. J., Gooley, P. R. (2008) Optimisation of NMR dynamic models II. A new methodology for the dual optimisation of the model-free parameters and the Brownian rotational diffusion tensor. J. Biomol. NMR 40, 121–133

44. Eaton, J. W., Bateman, D., and Hauberg, S. (2008) GNU Octave Manual, Version 3, Network Theory Ltd., Bristol, United Kingdom

45. Morin, S., and Gagné, M. S. (2009) Simple tests for the validation of multiple field spin relaxation data. J. Biomol. NMR 45, 361–372

46. d’Avergne, E. J., and Gooley, P. R. (2003) The use of model selection in the model-free analysis of protein dynamics. J. Biomol. NMR 25, 25–39

47. Long, D., Liu, M., and Yang, D. (2008) Accurately probing slow motions on millisecond time scales with a robust NMR relaxation experiment. J. Am. Chem. Soc. 130, 2432–2433