The stress sigma factor of RNA polymerase RpoS/σ S is a solvent-exposed open molecule in solution

Paola Cavaliere, Sebastien Brier, Petr Filipenko, Christina Sizun, Bertrand Raynal, Françoise Bonneté, Fabienne Levi-Acobas, Jacques Bellalou, Patrick England, Julia Chamot-Rooke, et al.

The stress sigma factor of RNA polymerase RpoS/σ^S is a solvent exposed open molecule in solution

Paola Cavaliere[1]+, Sébastien Brier[2,3]+, Petr Filipenko[4], Christina Sizun[5], Bertrand Raynal[6,7], Françoise Bonneté[8], Fabienne Levi-Acobas[7,9], Jacques Bellalou[7,9], Patrick England[6,7], Julia Chamot-Rooke[2,3], Claudine Mayer[7,10,11] and Françoise Norel[1,7,12]*.

[1] Institut Pasteur, Laboratoire Systèmes Macromoléculaires et Signalisation, Département de Microbiologie, 25 rue du Docteur Roux, 75015 Paris, France

[2] Institut Pasteur, Unité de Spectrométrie de Masse Structurale et Protéomique, Département de Biologie Structurale et Chimie, 28 rue du Docteur Roux, 75015 Paris, France

[3] CNRS USR 2000, 28 rue du Docteur Roux, 75015 Paris, France

[4] Hunter College-CUNY, Department of Computer Science, 695 Park Avenue, Hunter North Building, New York, NY 10065, USA

[5] Institut de Chimie des Substances Naturelles, CNRS UPR2301, Université Paris Saclay, 91190 Gif-sur-Yvette, France

[6] Institut Pasteur, Plateforme de biophysique moléculaire, Centre d'innovation et de recherche technologique, 25 rue du Docteur Roux, 75724 Paris

[7] CNRS UMR3528, 28 rue du Docteur Roux, 75015 Paris, France

[8] Institut des Biomolécules Max Mousseron (IBMM) UMR 5247 CNRS-UM-ENSCM, Université d'Avignon, 301, rue Baruch de Spinoza, F84000 Avignon, France

[9] Institut Pasteur, Plate-forme de Protéines Recombinantes, Département de Biologie Structurale et Chimie, 25 rue du Docteur Roux, 75015 Paris, France

[10] Institut Pasteur, Unité de Microbiologie Structurale, Département de Biologie Structurale et Chimie, 25 rue du Docteur Roux, 75015 Paris, France

[11] Université Paris Diderot, Sorbonne Paris Cité, Paris, France

[12] Institut Pasteur, Unité de Biochimie des Interactions Macromoléculaires, Département de Biologie structurale et Chimie, rue du Dr. Roux, 75015 Paris, France

*To whom correspondence should be addressed: Françoise Norel
Email: francoise.norel@pasteur.fr Tel: (33) 140613122

+ These authors contributed equally to the manuscript

Present address: Paola Cavaliere, Department of Biochemistry, Weill Cornell Medicine, New York, 10065, USA

ABSTRACT

In bacteria, one primary and multiple alternative σ factors associate with the RNA polymerase core enzyme (E) to form holoenzymes (Eσ) with different promoter recognition specificities. The alternative σ factor RpoS/σ^S is produced in stationary phase and under stress conditions and reprograms global gene expression to promote bacterial survival. To date, the three-dimensional structure of a full-length free σ factor remains elusive. The current model suggests that extensive interdomain contacts in a free σ factor result in a compact conformation that masks the DNA-binding determinants of σ, explaining why a free σ factor does not bind double stranded promoter DNA efficiently. Here, we explored the solution conformation of σ^S using amide hydrogen/deuterium exchange coupled with mass spectrometry, NMR, analytical ultracentrifugation and molecular dynamics. Our data strongly argue against a compact conformation of free σ^S. Instead, we show that σ^S adopts an open conformation in solution in which the folded σ_2 and σ_4 domains are interspersed by domains with a high degree of disorder. These findings suggest that E binding induces major changes in both the folding and domain arrangement of σ^S and provide insights into the possible mechanisms of regulation of σ^S activity by its chaperone Crl.

SHORT TITLE

σ^S is a solvent exposed open molecule

KEY WORDS

RpoS, sigma factor, RNA polymerase, conformation, structure, *Salmonella*

ABBREVIATIONS

RNAP, RNA polymerase; E, RNA polymerase core enzyme; Eσ, holoenzyme; ds, double strand; LRET, Luminescence Resonance Energy Transfer; MD, Molecular Dynamics; CD, Circular Dichroism; NMR, Nuclear Magnetic Resonance Spectroscopy; HDX-MS, Hydrogen/Deuterium eXchange coupled with Mass Spectrometry.

INTRODUCTION

The ability to adapt quickly to variations of environmental conditions is crucial for the growth and survival of bacteria in their natural environments. One major strategy used by bacteria, to coordinate the expression of genes needed for an adaptive response, is to use alternative sigma (σ) subunits of RNA polymerase (RNAP), allowing transcription initiation at different classes of promoters [1-3]. The σ factor associates with the catalytic core RNAP (E, $\alpha_2\beta\beta'\omega$) to guide the holoenzyme (Eσ) through important steps of transcription initiation, including promoter recognition and opening. All bacteria have a housekeeping sigma factor, which is responsible for the transcription of the majority of cellular genes that are essential for viability, and one or more alternative sigmas that direct RNAP to specific sets of genes in response to environmental conditions. The alternative sigma factor σ^S/RpoS is expressed in stationary phase of growth or under stress conditions and is required for bacterial resistance to multiple stress and starvation conditions [4-7].

Bacterial σ factors are classified into two structurally and functionally distinct families, named after the housekeeping σ^{70} and nitrogen-stress σ^{54} sigma factors of *Escherichia coli* [1,3,8]. σ^S belongs to the σ^{70} family, whose members consist of up to four structurally conserved domains that are connected by flexible linkers ($\sigma_{1.1}$, σ_2, σ_3 and σ_4, Figure 1A) [2-5,8]. The family is divided into four major phylogenetically and structurally distinct groups of σ factors [3,8]. All members of the four groups possess at least the σ_2 and σ_4 domains that include the major RNAP and promoter-binding determinants. σ_2 (regions 1.2 to 2.4) is the most conserved domain and binds the RNAP β' subunit coiled-coil and promoter -10 element, whereas σ_4 binds the RNAP β subunit flap and the promoter -35 element. σ_3 recognizes the extended -10 element of the promoter and is present only in σ factor groups 1-3. $\sigma_{1.1}$ is only found in σ factor group 1 and its sequence is poorly conserved, even though the presence of acidic amino-acids is preserved.

σ^{70} belongs to group 1 that contains primary σ factors, while σ^S belongs to group 2 that contains non-essential σ that are structurally related to group 1 sigmas [8]. The extensive sequence conservation between σ^S and σ^{70} in the DNA binding regions (Supplementary Figure S1A) is consistent with the finding that these two sigmas recognize similar promoter sequences [4-6]. The major differences in σ^S with respect to σ^{70} are its shorter domain $\sigma_{1.1}$ and the lack of a non-conserved region (NCR) in σ_2 (Figure 1A) [3,9].

Free σ factors are unable to efficiently recognize double stranded (ds) promoter DNA [2,3,9]. In the case of group 1 σ factors, it has been proposed that $\sigma_{1.1}$ auto-inhibits DNA binding [10-13]. However, even in the absence of $\sigma_{1.1}$, promoter recognition by σ alone is

weak [12,14,15] and σ factors that are naturally devoid of $\sigma_{1.1}$ also display a low DNA binding capacity [3,9]. Three-dimensional structures have been solved for σ factors in complexes with RNAP and anti–σ factors, or for stable domains after σ proteolysis, but there is no structure available for a full-length σ factor in the absence of a binding partner [3,9]. The σ–E interaction seems to induce conformational changes of E and σ leading to an active conformation of σ in which the DNA binding determinants in σ_2 and σ_4 are exposed and appropriately spaced [14,15]. Indeed, luminescence resonance energy transfer measurements (LRET) revealed that the distance between σ_2 and σ_4 is about 35 Å in free σ^{70}, a distance too small to allow for the simultaneous interaction of these regions with their corresponding promoter elements [15]. Upon binding of σ^{70} to E, the distance between σ_2 and σ_4 was increased by 15 Å and there was a displacement of $\sigma_{1.1}$ of about 20 Å with respect to σ_2 [15]. In addition, the chemical reactivity of cysteines at several positions in σ^{70} showed that E binding increases solvent exposure of the DNA binding domains of σ^{70} [14]. Group 3 flagellar σ, σ^{28}, entirely lacks $\sigma_{1.1}$, but aligns well with σ_2, σ_3 and σ_4 of group 1 σ, although σ^{28} does not contain the first helix of region $\sigma_{1.2}$ [8,16] (Figure 1A). In the crystal structure of σ^{28} bound to its inhibitory anti- σ factor FlgM, σ^{28} adopts a compact conformation with buried DNA-binding determinants, very different from the elongated conformation of σ^{70} in complex with E [16]. However, secondary structures of σ_2, σ_3 and σ_4 of σ^{28} bound to FlgM are very similar to those of σ^{70} in the holoenzyme, with the exception of the linker between domains σ_3 and σ_4 (Figure 1A). The σ_3-σ_4 linker in σ^{28} forms a bent tightly packed α-helix that is pivotal to maintain the closed structure of σ^{28} through interactions with each of the other domains [16]. In contrast, the corresponding linker region ($\sigma_{3.2}$ loop) is unstructured in Eσ^{70} [17]. Cross-linking experiments suggested that free σ^{28} assumes a similar compact conformation as the one bound to FlgM [16,18]. These findings led Sorenson *et al.* to propose that a helix-coil transition of the σ_3-σ_4 linker of σ^{28} occurs upon E binding and that this conformational change extends to all σ^{70}-family members [16,18,19]. However, since there is no structure available for Eσ^{28} and free σ^{70}, the folds of the σ_3-σ_4 linker in the holoenzyme and of $\sigma_{3.2}$ in free σ^{70} remains unknown.

Altogether, these results have led to the idea that the DNA binding determinants for all σ^{70} family members are masked by inter-domain contacts in a compact free σ conformation, even though the nature and strength of these contacts may differ between σ factors [18,19]. To investigate the validity of this hypothesis, we aimed to solve the solution conformation of σ^S from the human pathogen *Salmonella enterica* serovar Typhimurium (hereafter referred as *Salmonella*). σ^S is a widely studied σ factor due to its key role in survival and general stress resistance of many Gram-negative bacteria and to the large size

of the set of genes it controls [4-6,20,21]. In *Salmonella*, σ^S contributes to virulence and biofilm formation [22-24]. Our earlier structural model of *Salmonella* σ^S, based on the crystal structures of $E\sigma^{70}$ [25], is consistent with the recently released crystal structure of the DNA associated *E. coli* holoenzyme $E\sigma^S$ [26]. These studies show a structure and arrangement of σ^S in the RNAP holoenzyme similar to that of σ^{70}. In particular, the $\sigma_{3.2}$ linker is unstructured in $E\sigma^S$, as it is in $E\sigma^{70}$ [26].

Here, we combined biophysical and computational techniques to probe the ensemble structure of σ^S in solution. We used analytical ultracentrifugation to calculate the hydrodynamic radius of free σ^S and hydrogen/deuterium exchange to evaluate the solvent accessibility and secondary structure of σ^S. Together with molecular dynamics simulations, the data reveal that free σ^S adopts a solvent exposed open conformation that is distinct from its RNAP-bound conformation. These findings are discussed in light of the structural rearrangements required to form a complex with RNAP and the unique regulation of σ^S activity by the Crl protein [27,28].

MATERIALS AND METHODS

Protein production

The N-terminal $(his)_6$-tagged σ^S and σ^S_{1-162} from *S*. Typhimurium were produced in *E. coli* strain BL21 (DE3) harbouring plasmid derivatives of pETM11 and p-MCN-EAVNH, respectively (Supplementary Table S1). Production and purification of the proteins were carried out as previously described [28,29]. Proteins with isotope labeling were produced by replacing rich culture medium by M9 medium supplemented with 1 g/L $^{15}NH_4Cl$ (Eurisotop) and 3 g/L ^{13}C-labeled glucose (Eurisotop).

Molecular modeling

The open structural model of σ^S from *Salmonella* described in Monteil *et al.* [25] was used and is similar to the recently released structure of *E. coli* σ^S in $E\sigma^S$ [26]. The closed conformation of σ^S was initially generated using as a template the structure of σ^{28} in the σ^{28}-FlgM complex (PDB code 1RP3,[16]). When the closed model was compared to the open conformation of σ^S, it appeared that region comprising residues 219 to 249 undergoes important conformational rearrangements and this region was therefore modelled using Phyre2 [30]. The other domains of σ^S were placed by superposition on the corresponding domains of σ^{28}. The complete built closed model of σ^S was then subjected to energy minimization with the NAMD2 program [31] using CHARMM27 force field. The system was minimized by 3000 steps of conjugate gradient minimization.

Analytical Ultracentrifugation (AUC) and Circular dichroism (CD)

Sedimentation velocity experiments were carried out at 20°C in an XL-I analytical ultracentrifuge (Beckman-Coulter) equipped with double UV and Rayleigh interference detection. σ^S samples (10-100 µM) were spun at 42,000 rpm using an An50Ti rotor with double-sector Epon centerpieces (3 mm, 12 mm). Absorbance and interference profiles were recorded every 6 min. All experiments were performed in 50 mM sodium-phosphate pH 8.0, 300 mM NaCl. Standard deviations were calculated from the analysis of the integrated peaks. The partial specific volume of σ^S (0.734 ml/g) was estimated from its amino acid sequence using the software Sednterp (available on-line http://sednterp.unh.edu/). The same software was used to estimate the buffer viscosity (η = 1.056 cP) and density (ρ = 1.0145 $g.ml^{-1}$). Sedimentation values were extrapolated to zero concentration in standard condition $S_{20,w.}$ Theoretical hydrodynamic radii, radii of gyration and sedimentation coefficients of both open and close conformational σ^S models were calculated using Winhydropro [32] with a

radius size per amino acid allowing to maintain a computed volume of the model compatible with σ^S theoretical hydrated volume (typically 6.3 Å).

CD experiments were performed using Aviv CD spectrometer model 215 equipped with a water-cooled Peltier unit. Experiments were performed at concentrations 0.3-0.5 mg/mL in 50 mM Na-phosphate pH 8.0, 300 mM NaCl. Spectra were recorded at 20°C in cells with a path length of 0.1-mm (region from 200 to 260 nm) (121.QS, Hellma). Three consecutive scans were merged to produce an average spectrum. The spectra were corrected using buffer baseline measured under the same conditions. The molar ellipticity per mean residue, [θ] in deg cm^2 $dmol^{-1}$, was calculated from the equation $[\theta]=[\theta]_{obs}$ mrw $(10\ l\ C)^{-1}$,where $[\theta]_{obs}$ is the ellipticity measured in degrees, mrw is the mean residue molecular weight (111.5 Da), C is the protein concentration in g/mL, and l is the optical path length of the cell in cm. Percentage of secondary structure was obtained using Dichroweb (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml).

Nuclear magnetic resonance (NMR) spectroscopy

NMR experiments were recorded on 600 and 800 MHz Bruker Avance III NMR spectrometers equipped with TCN cryoprobes. NMR data were processed using Topspin 3.2 (Bruker) and analyzed in CCPNMR [33]. ^{15}N- or $^{13}C^{15}N$-labeled $\sigma^S_{1\text{-}162}$ samples (30-70 µM) in NMR buffer (50 mM sodium phosphate, 300 mM NaCl, pH 7.5) contained 7,5% D_2O to lock the spectrometer frequency. ^1H-^{15}N HSQC and triple resonance HNCA, HN(CO)CA and HNCO spectra of 45 µM $\sigma^S_{1\text{-}162}$ in NMR buffer supplemented with 50 mM arginine and 5 mM potassium glutamate were used for sequential assignment of $\sigma^S_{1\text{-}162}$ backbone chemical shifts at 298K (Supplementary Table S2). ^1H chemical shifts were referenced to DSS.

Hydrogen/Deuterium eXchange coupled with Mass Spectrometry (HDX-MS)

Triplicate analyses were performed for each time point for all HDX-MS analyses. Prior to addition of the deuterated buffer, purified free σ^S protein solutions (27 µM) were equilibrated for 1 h at 20°C. For intact protein analysis, continuous labeling was performed at 20°C for *t* = 0.16, 0.5, 1, 2, 5, 10, 30, 60 and 180 min. The labeling was initiated by mixing 45 µL of 99.9% D_2O in 20 mM sodium phosphate, 150 mM NaCl, pH 8.0 with 5 µL σ^S protein solution. Aliquots of 13.5 pmol were withdrawn at each experimental time point and quenched upon mixing with ice-cold 0.5% formic acid solution to achieve a final pH of 2.5. Quenched samples were immediately snap-frozen in liquid nitrogen and stored at −80 °C for approximately 12 h. Undeuterated controls were treated using an identical procedure. For local HDX analysis, intact protein was labeled for *t* = 0.16, 0.5, 1, 5, 10, 30, 60, 120 and 240, quenched and stored as described above. Fully deuterated samples of σ^S were achieved

after 6 h incubation at 60°C (final D_2O content of 90%). Prior to mass spectrometric analysis, samples were rapidly thawed and immediately injected into a cooled nanoACQUITY UPLC HDX system (Waters) maintained at 0°C. For intact mass measurement, intact proteins (5 pmol final) were loaded onto a Vanguard C4 pre-column (BEH, 1.7 µm, 2.1 x 5 mm, Waters) and desalted for 2 min at 100 µL/min with 95% solvent A (0.15% formic acid, pH 2.5) and 5% solvent B (100% acetonitrile, 0.15% formic acid, pH 2.5). Proteins were eluted from the trap column over a 2 min gradient of 5-90% solvent B at 100 µL/min. For peptide analysis, intact proteins (6.6 pmoles) were on-line digested using an in-house prepared cartridge of immobilized pepsin beads (Thermo Scientific, Rockford, IL) for 2 min at 100 µL/min and 20°C. Peptic peptides were trapped and desalted on a Vanguard C18 pre-column (BEH 1.7 µm, 2.1 x 5 mm, Waters), and separated using an analytical ACQUITY UPLC™ BEH C18 column (BEH, 1.7 µm, 1 x 100 mm, Waters) over a 6 min gradient of 5-40% solvent B at 40 µL/min. After each run, the pepsin column was manually cleaned with two washes of 0.8% formic acid, 5% acetonitrile, 1.5 M guanidinium chloride, pH 2.5. Blank injections were performed between each sample to confirm the absence of carry-over. The LC flow was directed to a Synapt™ G2-Si HDMS™ mass spectrometer (Waters) equipped with a standard electrospray ionization (ESI) source. Mass accuracy was ensured by continuously infusing a Glu-1-Fibrinogen solution (100 fmol/µL in 50% acetonitrile) through the reference probe of the ESI source. Mass spectra were acquired in positive-ion and resolution mode over the *m/z* range of 50–2000. Peptides were identified in undeuterated samples using a data-independent acquisition scheme (MS^E). Peptide identifications were made by database searching in ProteinLynx Global Server 3.0 (Waters) and each fragmentation spectrum was manually inspected for assignment validation. DynamX 3.0 HDX software (Waters) was used to extract the centroid masses of all peptides selected for local HDX-analyses; only one charge state was considered per peptide. No adjustment was made for back-exchange and the results are reported as relative deuterium exchange levels expressed in either mass unit or fractional exchange. Fractional exchange data was calculated by dividing the experimentally measured uptake by the theoretically maximum number of exchangeable backbone amide hydrogens that could be replaced into each peptide.

Molecular Dynamics (MD) simulation

MD simulations were carried out at 300 K and 300 mM NaCl. For molecular dynamics, we used the crystal structure of σ^S in the holoenzyme (PDB 5IPL, [26]). The protein structure was refined using the Protein Preparation Wizard (Maestro 10.4). Bond order and formal charges were assigned and hydrogen atoms were added. To further refine the structure, an OPLS3 force field parameter was used to alleviate steric clashes and the minimization was terminated when heavy atoms RMSD reached a maximum cutoff value of 0.30 Å.

Protonation states were assigned to residues according to the pKa based on pH = 7.0 using the Epik v3.4 module. Each structure was placed in a cubic cell, using DESMOND v 4.4 System Builder workflow, with size adjusted to maintain a minimum distance of 10 Å to the cell boundary, TIP3P water was added with an appropriate number of ions to establish 300 mM NaCl concentration. Simulations using the TIP4P water model equilibrated to the same end state. Molecular dynamic simulations were completed using DESMOND 4.4 package. The equations of motion were integrated using the multistep RESPA integrator with an inner time step of 2.0 fs for bonded interactions and non-bonded interactions within the short range cutoff. An outer time step of 6.0 fs was used for non-bonded interactions beyond the cutoff. A Nose–Hoover thermostat with relaxation time of 1.0 ps and the Martina–Tobias–Klein method with relaxation time of 2.0 ps were used to maintain the constant simulation temperature and to control the pressure, respectively. Smooth particle-mesh Ewald method with tolerance of $1e^{-09}$ was used to calculate long-range electrostatic interactions. Short range electrostatic interactions were calculated with cutoff radius at 9.0 Å. The system was equilibrated with the default protocol provided in DESMOND v4.4, which consists of steepest descent (SD) minimization with a maximum of 2000 steps and a gradient threshold of 50 kcal/mol/A, followed by 12 ps of Berendsen NVT (constant number of particles, volume, and temperature) simulation at 10 K, followed by 24 ps of Berendsen NVT (constant number of particles, pressure, and temperature) equilibration at 300 K. Default equilibration was followed by a 150ns production simulation using NTP ensemble at 300K. Energy was saved in 1.2 ps intervals and trajectory was saved in 4.8 ps intervals. Radius of gyration was calculated as

$$\left(\frac{1}{2N^2}\sum_{i,j}\left\|r_i - r_j\right\|^2\right)^{1/2}$$

where r_i and r_j are the positions of particles i and j from the given list. RMSD clustering was done on 1000 trajectory frames using Schrodinger Maestro 11.2 RMSD trajectory clustering applet. Hierarchical cluster linkage was performed selecting four clusters based on the average distance between all pairs of frames in clusters. Geometric clustering was performed to select the most representative structure from each of the four clusters. Resulting structures were aligned with simulation starting frame using Schrodinger Maestro 11.2 Structure Alignment tool.

RESULTS

Free σ^S in solution does not adopt a closed conformation

We previously built a structural model for *Salmonella* σ^S based on the crystal structures of Eσ^{70} [25] (Figure 1B). This model is in agreement with the crystal structure of *E. coli* σ^S-transcription initiation complex that was released during the completion of this work [26]. In the RNAP holoenzyme Eσ^S structure, $\sigma_{1.1}$ is not visible, but the structures of σ^S and σ^{70} are similar from domain $\sigma_{1.2}$ to the C-terminus [26] and have an elongated shape, spacing σ_2 and σ_4 for interaction with the promoter regions [9,26]. In stark contrast, the proposed structure of free σ^{28} is compact, and the σ_3-σ_4 linker forms a bent tightly packed α-helix that maintains the closed structure of σ^{28} through interactions with each of the other domains [16,18]. These observations led the authors to propose that a helix-coil transition of the σ_3-σ_4 linker of σ^{28} occurs upon E binding and that this conformational change is conserved over all σ^{70}-family members [16,18,19]. However, unwinding of the σ_3-σ_4 linker in σ^{28} remains hypothetical in the absence of an Eσ^{28}-structure. Vice versa, folding of the corresponding linker region ($\sigma_{3.2}$) in σ^S upon its release from Eσ^S [26] remains speculative in the absence of a free σ^S structure.

To assess whether free σ^S adopts a compact conformation similar to that proposed for σ^{28}, we carried out AUC experiments at increasing σ^S concentrations and results were extrapolated to zero concentration. At concentrations higher than 20 μM, interactions between σ^S monomers resulted in the formation of dimers and low amounts of higher order species, as previously reported for other σ factors [18,34]. Therefore, the extrapolation was performed using only the first peak (see Figure 2A) corresponding to the lowest sedimentation value (around 2.7 S) compatible with the monomeric form of σ^S. The sedimentation coefficient ($S_{20,\,w}$) obtained for the σ^S monomer was 2.7 ± 0.1 S with a calculated frictional ratio (f/f_0) of 1.5, which is indicative of a hydrated/elongated protein. This experimental value was slightly higher than the theoretical sedimentation coefficient of 2.5 S calculated for the open σ^S model (Figure 1B), while it was much lower than the 3.1 S value calculated for the structural model of σ^S based on the closed conformation of σ^{28} bound to FlgM (Figure 1C). These data suggest that the conformation of free σ^S is not as compact as that in the closed model, but that it is not as elongated as that in the open model. However, it cannot be ruled out that the experimental sedimentation coefficient value might reflect the presence of two or more σ^S conformations at equilibrium in solution, with at least one elongated form.

Global HDX-MS analysis does not support the existence of an equilibrium between open and closed structures of σ^S

We assessed the possibility that open and closed σ^S conformations coexist in solution at equilibrium by using Hydrogen/Deuterium eXchange (HDX) combined with Mass Spectrometry (MS). HDX-MS measures the isotopic exchange of backbone amide hydrogens with deuteriums along the entire protein backbone, allowing for the simultaneous analysis of protein conformation and dynamics. The rate of amide hydrogen exchange is directly influenced by both the solvent accessibility and the presence of secondary structural elements (*i.e.,* hydrogen bonding) [35-39]. Amide hydrogens fully exposed to the solvent and located in unstructured parts of the protein exchange rapidly, while those present in secondary structural elements (*i.e.,* α-helices or β-sheets) exchange at much slower rates due to hydrogen bonding. Therefore, backbone amide hydrogens represent excellent structural probes to monitor changes in protein structure, conformation and dynamics.

The presence of distinct stable σ^S conformers, such as open and closed species, could result in a difference in solvent accessibility observable on the time scale of our experiments, only if the interconversion rate of the conformers is slow compared to the intrinsic HDX rate. In such a scenario, the more open conformers (*i.e.,* more solvent exposed) would incorporate more deuterium than the closed species, leading to a multimodal isotopic distribution of protein ions. On the contrary, if the interconversion is much faster than the H/D exchange, a unimodal pattern will be observed, as for a single conformer. When full-length σ^S was exposed to deuterium under native conditions, a unique binomial isotopic pattern was observed for all exchange times (Figure 3A), which is more in favor of this second hypothesis. This result was further confirmed by performing peak width analysis of full-length σ^S ions [38,40,41]. In some cases, the measured MS signals of coexisting conformers might be too close to be resolved. Therefore, the observed isotopic pattern remains binomial but the width of the distribution undergoes detectable widening during the time course of the experiment [40]. As depicted in Figure 3B, the extracted peak width of the +44 charge state of σ^S did not experience any significant widening over the time course of the experiment, and remained centered around an average value of 77 ± 5 Da. This was also observed for other charge states (Supplementary Figure S2). Altogether these data indicate that only one conformational state of σ^S, with respect to secondary structure and solvent accessibility, can be detected in solution by HDX-MS. We cannot exclude however the existence of σ^S conformers interconverting much faster than the H/D exchange.

σ^S adopts a solvent-exposed conformation with many regions of structural disorder

To further characterize the structure of free σ^S in solution, the full-length σ^S protein was exposed to deuterium, subsequently digested with pepsin, and kinetics of deuterium uptake of 34 (of 67) peptides covering 90.6% of the protein sequence were followed (Supplementary Figure S3). The relative fractional uptake values calculated for each selected peptide were plotted as a function of peptide position (Figure 4A). An increase of deuterium uptake over time was observed in domains σ_2 (residues 72 to 164) and σ_4 (residues 256 to 316) only, confirming the presence of stable secondary structural elements (Figure 4A; Supplementary Figure S4). In contrast, the σ^S peptides covering residues 1 to 71 ($\sigma_{1.1}$ and $\sigma_{1.2}$) and 165 to 255 (σ_3) were fully exchanged since the first time point. These data provide direct evidence that, in the context of full-length free σ^S, region 1-71 and 165-255 do not contain stable secondary structural elements (Figure 4A; Supplementary Figure S4). Of note, the $\sigma_{3.2}$ linker (peptides 213-228, 216-235 and 218-235, Figure 4A; Supplementary Figure S4) appears structurally disordered (as observed in the E-bound σ^S crystal structure [26]), in sharp contrast with the helical fold adopted by the σ_3-σ_4 linker in the σ^{28}/FlgM crystal [16].

The overall HDX-MS behavior monitored with the free full-length σ^S correlates well with the recently solved crystal structure of the RNAP-bound σ^S factor ($E\sigma^S$) [26]. In this regard, most part of domains σ_4 and σ_2 (from $\sigma_{2.1}$ to $\sigma_{2.4}$) are folded in both the free and RNAP-bound states (hence explaining the dynamic exchange behavior observed in these regions). However, the absence of dynamic events in region $\sigma_{1.2}$ of domain σ_2 (peptides 46-60 and 61-71), domain σ_3 (regions $\sigma_{3.0}$ and $\sigma_{3.1}$; peptides 165-178 and 198-212), and the N-terminal region $\sigma_{4.1}$ of domain σ_4 (peptide 239-255), was not in agreement with the α-helical content observed in those regions in the crystal (Figure 4B). This data suggests that the binding of E to σ^S may either induce the folding of these regions or stabilize transient secondary structures. Finally, our HDX-MS data also reveal the unstructured and solvent-exposed nature of the first N-terminal residues (residues 1 to 52 containing the $\sigma_{1.1}$ domain) that were not resolved in the crystal structure of $E\sigma^S$ [26]. In contrast to the solution structures of free $\sigma_{1.1}$ [42,43] and of $E\sigma^{70}$ [17,44], $\sigma_{1.1}$ in the free full-length σ^S factor does not appear to form a globular α-helical domain.

In view of the differences observed between our HDX-MS data and the structure of several regions of σ^s in the $E\sigma^S$ complex, we decided to complement our HDX-MS approach using circular dichroism to assess the secondary structure of full-length σ^S and NMR spectroscopy to analyze the structure of a truncated version of σ^S comprising $\sigma_{1.1}$ and σ_2 (σ^S_{1-162}). The CD spectrum of full-length σ^S indicated that ca. 45% of the protein forms α-helices (Figure 2B), a value lower than the 58% calculated for the σ^s crystal structure with STRIDE

(for this calculation, the first 52 missing residues of σ^S were included and considered as unstructured) [45]. The CD value agrees well with the amount of the protein sequence covered by peptides that reach full deuteration within the first time point (Figure 4A) and suggests that the high solvent accessibility observed in HDX-MS indeed stems from a high amount of non structured regions in σ^S. NMR provided insight into the structural heterogeneity of $^{13}C^{15}N$ labeled $\sigma^S_{1\text{-}162}$ fragment (supplementary Figure S5). Only chemical shifts of the N-terminus, i.e. domain $\sigma_{1.1}$, could be assigned in a residue-specific manner, due to the low concentration required to keep the protein from oligomerization and/or aggregation. Backbone chemical shift analysis was indicative of an intrinsically disordered $\sigma_{1.1}$. The $\sigma^S_{1\text{-}162}$ fragment also contains ca. 50 residues that are strongly protected from exchange with water, even at high temperature and high pH. This strongly suggests the presence of a stable protein core domain, which, by extrapolation, must be located in domain σ_2.

Taken together, our HDX-MS, CD and NMR results are in favor of a stable and solvent-exposed conformation of free σ^S in solution, where the well-folded domains σ_2 and σ_4 are flanked by regions with high structural disorder ($\sigma_{1.1}$, $\sigma_{1.2}$ and σ_3).

Free σ^S adopts a distinct conformation from that in the holoenzyme

The AUC data argue against a compact conformation of free σ^S like the one proposed for free σ^{28}. In combination with HDX, these experiments reveal that free σ^S is a solvent-exposed molecule with an open conformation (which can nevertheless fluctuate between different open conformational states due to the presence of disordered regions in the middle part of the σ^S protein). To explore the conformational landscape of free σ^S, we performed molecular dynamics simulations over 600 ns using the open conformation of σ^S excised from the Eσ^S crystal structure (PDB= 5IPL), as a starting point. Simulations were performed in the presence of 300 mM NaCl, to be consistent with the experimental conditions used in AUC and HDX experiments. The root mean square deviation (RMSD) of the alpha carbons (Cα), which reflect the overall conformational changes of the protein through the simulation, increased rapidly to a considerable 12.5 Å after 20 ns and continued to increase until reaching a stable plateau at ca. 17.5 Å after 100 ns (Figure 5A). The secondary structure of the protein remained essentially the same throughout the simulation, with the exception of the region encompassing residues 175-190 (σ_3) and the N- and C-terminal regions. In those regions, the α-helical content was lower in the final frame than in the initial frame, suggesting that the helices present in the crystal structure were not fully stable when σ^S is dissociated from E (Supplementary Figure S6A, B). A convenient measure to quantify how much amino acids move from a pre-defined position in a structure is the root mean square fluctuation (RMSF). The N- and C-termini of proteins typically exhibit big RMS fluctuations, as it is also

the case in σ^S (Supplementary Fig. S7). Little RMS fluctuations were observed for amino-acids in well-defined regular secondary structure elements, like α-helices (Supplementary Fig. S7, pink bars), while fluctuations increased in unstructured regions (Supplementary Fig. S7, white gaps). For example, unstructured regions in σ_2 and σ_4 showed increased RMS fluctuations (residues 130-145). Moreover, in agreement with the rapid exchange behavior observed for peptides 165-178 and 179-197 in the HDX experiments (Figure 4; Supplementary Figure S4), the RMS fluctuations of amino acids in region σ_3 (residues 168-188) were significantly bigger than what would be expected for a folded α-helix. Hence, molecular dynamics simulation further support the disorder in free σ^S in this region.

Structure compactness was calculated by the radius of gyration for σ^S for each frame along the 600 ns. Ultimately, the radius of gyration allows for the simulated structure to be compared with the experimentally obtained sedimentation value by AUC (Figure 2A). The radius of gyration decreased from 35.3 Å in the starting frame to 24.1 Å in the final frame, consistent with the decreasing distance between σ_2 and σ_4 from the initial to the final frame from 70 Å to 40 Å. The radius of gyration calculated for the closed model of σ^S is 21.5 Å, which is smaller than the value obtained for the last frame after MD simulations (24.1 Å). Clearly, these calculations yet again rule out a closed conformation for free σ^S. Note, however, that the theoretical sedimentation coefficient calculated for the final frame obtained by MD is 2.9 S, which still differs from the experimental one (2.7 S), suggesting that the actual σ^S conformation is more open than the conformation obtained in the final frame of the MD simulations or that the protein structure samples an ensemble of more or less open conformations due to flexible hinges between domains. Interestingly, in the MD final frame, domain σ_4 rotated from its original position in the initial frame with respect to domain σ_2 (Figure 5B). This movement might be involved in the conformational changes that σ^S undergoes upon E binding, as previously suggested [14,15].

DISCUSSION

The lack of 3D structures of free σ factors has led to speculations about their conformations. The knowledge of the actual conformation of free σ is of great importance to understand (i) how σ factors are regulated by their binding partners such as anti-sigmas, chaperones, proteases, (ii) why they bind the RNAP core enzyme E with variable affinities, and (iii) why they do not bind dsDNA in the absence of E. In *E. coli* and *Salmonella*, σ^S is a tightly regulated molecule, at the levels of transcription, translation, stability and activity [4-6,27]. The last two regulatory steps likely involve conformational changes in σ^S. This study integrated a number of biophysical techniques that collectively show that free σ^S adopts a solvent-exposed open conformational state in solution and argue against the concept of a compact conformation of σ similar to the one proposed for σ^{28} [16,18]. Moreover, our data indicate that the tertiary structure of free σ^S is distinct from the open conformation adopted by σ^S in complex with RNAP and pinpoint regions of σ^S that are disordered in free σ^S, while being folded into α-helices in the crystal structure of the Eσ^S-DNA complex [26].

The closed conformation of σ^{28} is dependent on the ability of the σ_3-σ_4 linker (Figure 1) to fold into a α-helix and to establish intramolecular interactions with the other domains of σ^{28} [16,18]. Our HDX data showed that this region is disordered in free σ^S (Figure 4), like in E-bound σ^S [26]. Nevertheless, these data also show that the conformation of free σ^S substantially differs from the open folded structure adopted within the holoenzyme. The model obtained after MD simulations clearly points toward a conformation in which domains σ_2 and σ_4 are separated by ca. 40 Å, in good agreement with previous LRET analysis of free σ^{70} that revealed a σ_2-σ_4 distance of 35 Å [15]. In this conformation, domains σ_2 and σ_4 are not properly spaced to interact with promoter regions, fully explaining why free σ^S and σ^{70} cannot bind efficiently dsDNA without E [2,3,9]. Moreover, the final model obtained by MD simulations showed a relative rotation of domain σ_4, as compared to the position of this domain in the Eσ^S crystal structure. Interestingly, LRET data [15] also suggested that translational and/or rotational motions of domains σ_1 and σ_4 could result in the final core-induced rearrangement of σ^{70}, in agreement with our model. Therefore, it is possible that the lack of σ dsDNA binding in the absence of E is due, not only to the incorrect spacing of the DNA binding domains, but also to the relative position of the domains so that they do not expose their DNA binding faces for promoter binding properly. The N-terminal portion of domain σ_2 ($\sigma_{1.2}$; 46-60 and 61-71), domain σ_3 ($\sigma_{3.0}$ and $\sigma_{3.1}$; 165-178 and 198-212) and the N-terminal portion of domain σ_4 ($\sigma_{4.1}$; 239-255) are unstructured in free σ^S (Figure 4) and folded in the Eσ^S-DNA crystal structure [26]. We suggest that E binding induces the folding or

stabilizes these regions, or parts of them, providing a way to open and rearrange domains of σ^S to allow subsequent DNA binding. Indeed, domain σ_3 has also been shown to participate in DNA binding in *E. coli* Eσ^S [26]. Thus, it is possible that the structural disorder observed in the N-terminal portion of domain σ_3 in the absence of RNAP also contributes to the lack of efficient dsDNA recognition of free σ^S.

$\sigma_{1.1}$ is proposed to limit the ability of free housekeeping σ to bind dsDNA [10,11,13]. Cross-linking studies performed on the *Thermotoga maritima* housekeeping σ, σ^A, showed that regions $\sigma_{1.1}$, σ_2 and σ_4 are situated close in space [12]. $\sigma_{1.1}$ might stabilize a compact conformation of free housekeeping σ through electrostatic interactions between the negatively charged surface of $\sigma_{1.1}$ and the positively charged DNA binding regions of σ_2 and σ_4 [12]. However, the interactions observed by the cross-linking approach between $\sigma_{1.1}$, σ_2 and σ_4 in free housekeeping σ are likely weak and transient [12,13] and NMR spectra of $\sigma_{4.2}$ are not perturbed by $\sigma_{1.1}$ [10]. In housekeeping sigmas, $\sigma_{1.1}$ forms a globular α-helical structural unit [12,17,43,44]. In contrast, our HDX and NMR experiments with σ^S provide evidence for a lack of secondary structural content in $\sigma_{1.1}$. Consistently, no secondary structure was predicted for residues 1-40 of σ^S, in contrast to σ^{70} (Supplementary Fig. S8). As in housekeeping sigmas, $\sigma_{1.1}$ in σ^S is negatively charged. However, $\sigma_{1.1}$ in σ^S is smaller and its sequence is not conserved, compared to $\sigma_{1.1}$ from housekeeping sigmas. Thus, it is possible that the nature of $\sigma_{1.1}$ differentially affects the conformation of this domain in different σ factors. In σ^S, $\sigma_{1.1}$ might play a role in σ^S degradation by the ClpXP protease and the adaptor protein RssB. It has been proposed that a ClpX binding site close to the N-terminus of σ^S [46] is exposed following conformational changes of σ^S induced by interaction between RssB and the N-terminal part of σ^S domain σ_3 (residue K173). It remains to be determined whether the absence of secondary structure revealed here in $\sigma_{1.1}$ of free σ^S plays a role in this mechanism by preventing ClpX binding and whether the $\sigma_{1.1}$ fold is affected by RssB binding to K173.

In *E. coli* and *Salmonella*, the small protein Crl binds to domain σ_2 of σ^S and increases σ^S activity by enhancing its affinity for E [27,47]. One possible model postulates that σ^S exists as an equilibrium between open and closed conformations and that Crl interaction favors binding to an open conformer that is more prone to E binding. This is inconsistent with our observation that σ^S only adopts an open conformation. It remains to be determined whether Crl binding induces some of the conformational changes in σ^S that are required for its accommodation within RNAP. Interestingly, one Crl binding region on σ^S corresponds to the second α-helix of $\sigma_{1.2}$ that includes R82, a σ^S residue key for Crl binding [28]. This region is close to the N-terminus of $\sigma_{1.2}$, which appears to be unfolded in σ^S but folded in the Eσ^S-DNA

complex. In housekeeping RNAP, this region interacts with the promoter discriminator element and is involved in open complex formation [3,9]. Hence, it is plausible that Crl binding induces folding in this region of σ^S, thus accounting for the role for Crl in the formation of Eσ^S-DNA open complexes [47,48].

The mechanism by which σ^S activity is regulated by Crl is unique. Indeed, regulation of the activity of σ factors is classically achieved by anti-σ factors, which bind to and inhibit their cognate σ by preventing their interaction with E [3]. For σ^{28}, the anti-σ FlgM wraps around σ_2 and σ_4 [16], a binding mode that might be well adapted to the compact structure of free σ^{28} in solution. In enterobacteria such as *E. coli* and *Salmonella*, there are no known anti-σ^S proteins. However, it has been recently reported that in the aquatic proteobacterium *Shewanella oneidensis,* which does not contain a *crl* gene, σ^S can be bound by an anti-σ molecule CrsR [49]. The sequence of the σ^S protein from *Shewanella* is very similar to that of σ^S from *Salmonella* and *E. coli*, except in domain $\sigma_{1.1}$ (Supplementary Figure S1B). Moreover, α-helical secondary structures were predicted for residues 1-40 of *Shewanella* σ^S, in contrast to *Salmonella* σ^S (Supplementary Figure S8). It would be interesting to determine whether the conformation of *Shewanella* σ^S in solution is similar to that of *Salmonella* σ^S or more compact, like in σ^{28}, potentially favoring CrsR binding.

In conclusion, we have provided evidence for a highly solvent exposed open conformation of free σ^S, well different from the compact conformation proposed for free σ^{28}. Clearly further comparison of the conformations of free σ factors is required to evaluate whether σ factor groups adopt distinct solution structures and to come to a better understanding of the role of these conformations in the regulation of σ. However, it is tempting to speculate that the conformational diversity among σ factors is adapted to the mechanism of σ regulation.

ACKNOWLEDGEMENTS

We thank Daniel Ladant and members of the LSMS and BIM laboratories for their kind support.

DECLARATION OF INTEREST

The Authors declare that there are no competing interests associated with the manuscript.

FUNDING

This work was supported by the French National Research Agency [grant ANR- 11-BSV3-009] and by grants from the Institut Pasteur and the Centre National de la Recherche Scientifique. The MS work was supported by the Investissements d'Avenir through the CACSICE project [grant ANR- 11-EQPX-008]. Access to high field NMR spectrometers financially supported by IR-RMN-THC FR3050 CNRS is gratefully acknowledged.

AUTHOR CONTRIBUTION

FN and PC designed research; PC, SB, PF, CS, BR, FLA, CM performed research; PC, SB, PF, CS, BR, FB, JB, PE, DL, CM, and FN analyzed data; FN, PC, SB, CS, CM wrote the paper. PC SB, PF, CS, BR, FB, FLA, JB, PE, JCR, CM and FN corrected the paper.

SUPPLEMENTARY DATA

Supplementary Data are available online.

REFERENCES

1. Osterberg, S., del Peso-Santos, T. and Shingler, V. (2011) Regulation of alternative sigma factor use. *Annu. Rev. Microbiol.* **65**, 37-55.
2. Feklistov, A., Sharon, B.D., Darst, S.A. and Gross, C.A. (2014) Bacterial sigma factors: a historical, structural, and genomic perspective. *Annu. Rev. Microbiol.* **68**, 357-376.
3. Paget, M.S. (2015) Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. *Biomolecules* **5**, 1245-1265.
4. Battesti, A., Majdalani, N. and Gottesman, S. (2011) The RpoS-mediated general stress response in *Escherichia coli*. *Annu. Rev. Microbiol.* **65**, 189-213.
5. Hengge, R. (2011) The general stress response in Gram-negative bacteria. *Bacterial Stress Responses (Second Edition, eds. G. Storz and R. Hengge), ASM Press, Washington D.C.*, 251-289.
6. Landini, P., Egli, T., Wolf, J. and Lacour, S. (2014) sigmaS, a major player in the response to environmental stresses in *Escherichia coli*: role, regulation and mechanisms of promoter recognition. *Environ. Microbiol. Rep.* **6**, 1-13.
7. Schellhorn, H.E. (2014) Elucidating the function of the RpoS regulon. *Future Microbiol.* **9**, 497-507.
8. Lonetto, M., Gribskov, M. and Gross, C.A. (1992) The sigma 70 family: sequence conservation and evolutionary relationships. *J. Bacteriol.* **174**, 3843-3849.
9. Murakami, K.S. (2015) Structural biology of bacterial RNA polymerase. *Biomolecules* **5**, 848-864.
10. Camarero, J.A., Shekhtman, A., Campbell, E.A., Chlenov, M., Gruber, T.M., Bryant, D.A., Darst, S.A., Cowburn, D. and Muir, T.W. (2002) Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR. *Proc. Natl. Acad. Sci. USA* **99**, 8536-8541.
11. Dombroski, A.J., Walter, W.A., Record, M.T., Jr., Siegele, D.A. and Gross, C.A. (1992) Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. *Cell* **70**, 501-512.
12. Schwartz, E.C., Shekhtman, A., Dutta, K., Pratt, M.R., Cowburn, D., Darst, S.A. 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.
13. Dombroski, A.J., Walter, W.A. and Gross, C.A. (1993) Amino-terminal amino acids modulate sigma-factor DNA-binding activity. *Genes Dev.* **7**, 2446-2455.

14. Callaci, S., Heyduk, E. and Heyduk, T. (1998) Conformational changes of *Escherichia coli* RNA polymerase sigma70 factor induced by binding to the core enzyme. *J. Biol. Chem*. **273**, 32995-33001.
15. Callaci, S., Heyduk, E. and Heyduk, T. (1999) Core RNA polymerase from *E. coli* induces a major change in the domain arrangement of the sigma 70 subunit. *Mol. Cell* **3**, 229-238.
16. Sorenson, M.K., Ray, S.S. and Darst, S.A. (2004) Crystal structure of the flagellar sigma/anti-sigma complex sigma(28)/FlgM reveals an intact sigma factor in an inactive conformation. *Mol. Cell* **14**, 127-138.
17. Murakami, K.S. (2013) X-ray crystal structure of *Escherichia coli* RNA polymerase sigma70 holoenzyme. *J. Biol. Chem*. **288**, 9126-9134.
18. Sorenson, M.K. and Darst, S.A. (2006) Disulfide cross-linking indicates that FlgM-bound and free sigma28 adopt similar conformations. *Proc. Natl. Acad. Sci. USA* **103**, 16722-16727.
19. Campbell, E.A., Westblade, L.F. and Darst, S.A. (2008) Regulation of bacterial RNA polymerase sigma factor activity: a structural perspective. *Curr. Opin. Microbiol.* **11**, 121-127.
20. Levi-Meyrueis, C., Monteil, V., Sismeiro, O., Dillies, M.A., Monot, M., Jagla, B., Coppee, J.Y., Dupuy, B. and Norel, F. (2014) Expanding the RpoS/sigmaS-network by RNA sequencing and identification of sigmaS-controlled small RNAs in *Salmonella. PloS one* **9**, e96918.
21. Lago, M., Monteil, V., Douche, T., Guglielmini, J., Criscuolo, A., Maufrais, C., Matondo, M. and Norel, F. (2017) Proteome remodelling by the stress sigma factor RpoS/sigmaS in *Salmonella*: identification of small proteins and evidence for post-transcriptional regulation. *Sci. Rep.* **7**, 2127.
22. Hengge, R. (2011) Stationary-Phase Gene Regulation in *Escherichia coli* section sign. *EcoSal Plus* **4**.
23. Robbe-Saule, V., Jaumouille, V., Prevost, M.C., Guadagnini, S., Talhouarne, C., Mathout, H., Kolb, A. and Norel, F. (2006) Crl activates transcription initiation of RpoS-regulated genes involved in the multicellular behavior of *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* **188**, 3983-3994.
24. Dong, T. and Schellhorn, H.E. (2010) Role of RpoS in virulence of pathogens. *Infect. Immun.* **78**, 887-897.
25. Monteil, V., Kolb, A., Mayer, C., Hoos, S., England, P. and Norel, F. (2010) Crl binds to domain 2 of sigma(S) and confers a competitive advantage on a natural *rpoS* mutant of *Salmonella enterica* serovar Typhi. *J. Bacteriol.* **192**, 6401-6410.

26. Liu, B., Zuo, Y. and Steitz, T.A. (2016) Structures of *E. coli* sigmaS-transcription initiation complexes provide new insights into polymerase mechanism. *Proc. Natl. Acad. Sci. USA* **113**, 4051-4056.
27. Cavaliere, P. and Norel, F. (2016) Recent advances in the characterization of Crl, the unconventional activator of the stress sigma factor sigmaS/RpoS. *Biomol. Concepts* **7**, 197-204.
28. Cavaliere, P., Sizun, C., Levi-Acobas, F., Nowakowski, M., Monteil, V., Bontems, F., Bellalou, J., Mayer, C. and Norel, F. (2015) Binding interface between the *Salmonella* sigma(S)/RpoS subunit of RNA polymerase and Crl: hints from bacterial species lacking *crl. Sci. Rep.* **5**, 13564.
29. Cavaliere, P., Levi-Acobas, F., Mayer, C., Saul, F.A., England, P., Weber, P., Raynal, B., Monteil, V., Bellalou, J., Haouz, A. *et al.* (2014) Structural and functional features of Crl proteins and identification of conserved surface residues required for interaction with the RpoS/sigmaS subunit of RNA polymerase. *Biochem. J.* **463**, 215-224.
30. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845-858.
31. Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L. and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781-1802.
32. Garcia de la Torre, J., Huertas, M.L. and Carrasco, B. (2000) HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.* **147**, 138-146.
33. Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J. and Laue, E.D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-696.
34. Rist, W., Jorgensen, T.J., Roepstorff, P., Bukau, B. and Mayer, M.P. (2003) Mapping temperature-induced conformational changes in the *Escherichia coli* heat shock transcription factor sigma 32 by amide hydrogen exchange. *J. Biol. Chem.* **278**, 51415-51421.
35. Miranker, A., Robinson, C.V., Radford, S.E., Aplin, R.T. and Dobson, C.M. (1993) Detection of transient protein folding populations by mass spectrometry. *Science* **262**, 896-900.
36. Eyles, S.J. and Kaltashov, I.A. (2004) Methods to study protein dynamics and folding by mass spectrometry. *Methods* **34**, 88-99.
37. Krishna, M.M., Hoang, L., Lin, Y. and Englander, S.W. (2004) Hydrogen exchange methods to study protein folding. *Methods* **34**, 51-64.

38. Wales, T.E. and Engen, J.R. (2006) Partial unfolding of diverse SH3 domains on a wide timescale. *J. Mol. Biol.* **357**, 1592-1604.
39. Konermann, L., Pan, J. and Liu, Y.H. (2011) Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* **40**, 1224-1234.
40. Weis, D.D., Wales, T.E., Engen, J.R., Hotchko, M. and Ten Eyck, L.F. (2006) Identification and characterization of EX1 kinetics in H/D exchange mass spectrometry by peak width analysis. *J. Am. Soc. Mass Spectrom.* **17**, 1498-1509.
41. Chen, S., Brier, S., Smithgall, T.E. and Engen, J.R. (2007) The Abl SH2-kinase linker naturally adopts a conformation competent for SH3 domain binding. *Protein Sci.* **16**, 572-581.
42. 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.
43. Zachrdla, M., Padrta, P., Rabatinova, A., Sanderova, H., Barvik, I., Krasny, L. and Zidek, L. (2017) Solution structure of domain 1.1 of the sigmaA factor from *Bacillus subtilis* is preformed for binding to the RNA polymerase core. *J. Biol. Chem.* **292**, 11610-11617.
44. Bae, B., Davis, E., Brown, D., Campbell, E.A., Wigneshweraraj, S. and Darst, S.A. (2013) Phage T7 Gp2 inhibition of *Escherichia coli* RNA polymerase involves misappropriation of sigma70 domain 1.1. *Proc. Natl. Acad. Sci. USA* **110**, 19772-19777.
45. Heinig, M. and Frishman, D. (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic acids Res.* **32**, W500-502.
46. Studemann, A., Noirclerc-Savoye, M., Klauck, E., Becker, G., Schneider, D. and Hengge, R. (2003) Sequential recognition of two distinct sites in sigma(S) by the proteolytic targeting factor RssB and ClpX. *EMBO J.* **22**, 4111-4120.
47. England, P., Westblade, L.F., Karimova, G., Robbe-Saule, V., Norel, F. and Kolb, A. (2008) Binding of the unorthodox transcription activator, Crl, to the components of the transcription machinery. *J. Biol. Chem.* **283**, 33455-33464.
48. Bougdour, A., Lelong, C. and Geiselmann, J. (2004) Crl, a low temperature-induced protein in *Escherichia coli* that binds directly to the stationary phase sigma subunit of RNA polymerase. *J. Biol. Chem.* **279**, 19540-19550.
49. Bouillet, S., Genest, O., Jourlin-Castelli, C., Fons, M., Mejean, V. and Iobbi-Nivol, C. (2016) The General Stress Response sigmaS Is Regulated by a Partner Switch in the Gram-negative Bacterium *Shewanella oneidensis*. *J. Biol. Chem.* **291**, 26151-26163.

FIGURE LEGENDS

Figure 1. Structural models for σ^S. A) Schematic representation of domains in σ^{70}, σ^S and σ^{28}. The linker that connects σ_3 and σ_4 is called $\sigma_{3.2}$ for σ^{70} and σ^S [26] and σ_3-σ_4 linker for σ^{28} [16]. Cartoon representation of the *Salmonella* σ^S (B) open model [25] based on available Eσ 3D structures and (C) closed model based on the crystal structure of σ^{28}/FlgM [16]. Residues of σ^S forming a loop in the structure of Eσ^S [26] are shown in a box (the $\sigma_{3.2}$ loop) (see also Supplementary Figure S4).

Figure 2. Conformation of free σ^S in solution assessed by AUC and CD. A) Experimental sedimentation coefficient $S_{20,W}$ of σ^S carried out at different σ^S concentrations and extrapolated at zero concentration (inset) using the sedimentation value compatible to the monomeric form of σ^S (dashed lines). B) Far-UV (180-260 nm) CD spectra of σ^S at 0.3 mg/mL. The percentage of secondary content calculated is ca. 45% of α-helices, with 17% folded in turns and the remain of the protein is unstructured.

Figure 3. HDX-MS data for the intact σ^S protein. (A) ESI raw mass spectra of the +44 charge state for selected time points. σ^S incorporates deuterium following a binomial distribution, suggesting the presence of a unique population in solution (Und, Undeuterated control). (B) Peak width analysis of σ^S. The peak width of the +44 charge state was measured at a relative intensity of 50% (see the solid bar in panel A which is set at 50% peak intensity) with HX Express [40] and plotted as a function of labeling time. The peak width in Da has been calculated by multiplying the peak width in m/z by the charge (+44). The average peak width value is reported as a dashed line.

Figure 4. Local HDX exchange behavior of the σ^S protein. (A) Exchange profile of free σ^S. Deuterium uptake was calculated relative to the peptide mass and as a fraction of the theoretical maximum of incorporable deuteriums and plotted as a function of peptide position. The black to red lines correspond to data acquired at 10 sec up to 4 h deuteration (see color legend). Each dot corresponds to an average of three independent experiments. The average standard deviation calculated for all peptides and all time points is of +/- 0.06 Da. The dynamic events observed in domains σ_2 and σ_4 are characteristic of secondary structural elements. (B) Ribbon representation of the σ^S structure in the Eσ^S-DNA crystals (PDB = 5IPL) [26]. The N-terminal domain ($\sigma_{1.1}$, residues 1 to 52) has not been traced in these σ^S -transcription initiation complexes. The three σ^S regions that appeared folded differentially in

free σ^S and in the Eσ^S holoenzyme crystal structure are colored in red. The deuterium uptake curves of five selected peptides covering those regions are also reported. The full deuterium level measure for each peptide is shown in each graph as a dotted line. Deuterium uptake curves for all peptides are shown in Supplementary Figure S4.

Figure 5. Molecular dynamics using the crystal structure of σ^S in the RNAP holoenzyme (PDB 5IPL). A) RMSD of Cα *versus* simulation time of σ^S using the crystal structure of σ^S in complex with E [26] as the starting point. The color-code is by clusters (blue, green, yellow and red for cluster 1, 2, 3 and 4, respectively) obtained from hierarchical cluster linkage based on the average distance between all pairs of frames in clusters. Superposition of the initial σ^S frame (in orange) and the representative structures from each cluster (darker grey scales at increasing times) are shown. B) Superposition of the initial frame and the representative structure obtained form cluster 4 in which is highlighted the rotation of domain σ_4 in the MD structure, with respect to the initial frame.