1	Structural basis of transcription activation by the global regulator Spx
2	Jing Shi ^{1,#,*} , Fangfang Li ^{1,#} , Aijia Wen ^{3,4,#} , Libing Yu ^{5,#} , Lu Wang ¹ , Fulin Wang ¹ , Yuanling Jin ¹ ,
3	Sha Jin ^{3,4} , Yu Feng ^{3,4,*} , Wei Lin ^{1,2,6,*}
4	
5 6	¹ Department of Pathogen Biology, School of Medicine & Holistic Integrative Medicine, Nanjing
7	² State Key Laboratory of Natural Medicines, China Pharmaceutical University, Nanjing, China
8	³ Department of Biophysics, Zheijang University School of Medicine, Hangzhou, China,
9	⁴ Department of Pathology of Sir Run Run Shaw Hospital, Zhejiang University School of
10	Medicine, Hangzhou, China.
11	⁵ Institute of Materials, China Academy of Engineering Physics, Mianyang, China
12	⁶ Jiangsu Collaborative Innovation Center of Chinese Medicinal Resources Industrialization,
13	Nanjing 210023, China.
14	
15	#, Equal contribution.
16	
17	*, Correspondence: <u>weilin@njucm.edu.cn OR yufengjay@zju.edu.cn OR</u>
18	shijing301@njucm.edu.cn
19	
20	Running title: Structural insights into the physiological function of Spx in bacteria
21	
22	
23	
24	
25	
26	
27	
28	
29	





49



51 (A) SDS-PAGE of the purified complex. 20 pmoles of complex was loaded.

52 (B) Relative transcription activity of *B. subtilis* RNAP-promoter open complex with 53 or without Spx by using Mango-based transcription assay. Data for *in vitro* transcription 54 assays are means of three technical replicates. Error bars represent mean \pm SEM of n = 3 55 experiments.

56 (C) Chromatogram map of gel filtration.

(D) Native gel analysis of the purified complex. The gel was stained with 4S Red Plus
Nucleic Acid Stain (Sangon Biotech, Inc.) according to the procedure of the
manufacturer.





94 (B) Gold-standard FSC. The gold-standard FSC was calculated by comparing the two

95 independently determined half-maps from RELION. The dashed line represents the

96 0.143 FSC cutoff, which indicates a nominal resolution of 4.22 Å.

- 97 (C) Cryo-EM density map colored by local resolution. Local resolution calculation
- 98 was performed using blocres (1). View orientation as in **Figure 1B**.
- 99 (**D**) Angular distribution of particle projections. View orientations as in (C).
- 100 (E) FSC calculated between the model and the half map used for refinement (work),
- 101 the other half map (free), and the full map.



116 Figure. S5. The SPA quality assessment of *B. subtilis* Spx-TAC from remote

- 117 **3DFSC** processing server (<u>https://3dfsc.salk.edu/</u>) followed the procedures as (2).
- 118 (A) Histogram and directional FSC Plot for *B. subtilis* Spx-TAC.
- (B) FSC plots for *B. subtilis* Spx-TAC. (C) FT power plot for *B. subtilis* Spx-TAC.



Figure S6. Representative cryo-EM densities of superimposed models in
Spx-TAC.

123 (A) Cryo-EM density map (blue mesh) and the superimposed model (yellow cartoon) 124 of σ^{A} . (B) Cryo-EM density map (blue mesh) and the superimposed model of β' 125 coiled-coil. (C) Cryo-EM density map (blue mesh) and the superimposed model of 126 Spx-RNAP α CTD subunit. (D) Cryo-EM density map (blue mesh) and the 127 superimposed model of Spx-DNA. Other colors are shown as in Figure 1.

128

- 129
- 130
- 131
- 132
- 133



134

Figure S7. Extra Cryo-EM density map for β CTD. (A-C) Cryo-EM map and model for

136 Spx-TAC; (**D**) Extra Cryo-EM density map for β CTD is shown in coot. Other colors

are shown as in **Figure 1C**.



140 Figure S8. Comparisons of Spx-TAC, TAP-TAC and CAP-TAC.

141 (A-C) Spx-TAC, CAP-TAC and TAP-TAC (cartoon in left panel; surface in right

142 panel). Cyan, CAP or TAP; blue, Spx; yellow, σ ; white, pink, gray and dark gray,

- 143 RNAP α NTD, α CTD, β , and β '.
- 144
- 145
- 146



148 Figure S9. Spx derivatives show defects on Spx-TAC formation.

(A) Electrophoretic mobility shift assay for wild-type Spx and its mutants 149 (Spx-D108R, Spx-C10S, Spx-T11A, Spx-C13S and Spx-K15A). (B) Electrophoretic 150 mobility shift assay for wild-type Spx and its mutants (Spx-R60E, Spx-K66A, 151 152 Spx-R91A, Spx-R92A and Spx-N106A). (C) Electrophoretic mobility shift assay for wild-type Spx and its mutants (Spx-R111A, Spx-R128X, Spx-R100A, Spx-P115A and 153 Spx-S58A). (D) Electrophoretic mobility shift assay for wild-type Spx and its mutants 154 (Spx-E50A, Spx-D51R, Spx-E109R, Spx-G52R and Spx-D54R). Reaction conditions 155 are described in detail in Materials and Methods. Bands of Spx-TAC, RPo and free 156 157 DNA (trxA-mango promoter DNA) are showed on the left, respectively.

- 158
- 159
- 160 161
- ____
- 162
- 163





Figure S10 Thermal stability assays to detect the stabilities of wild-type Spx and
its mutants and whether folded properly as wild-type Spx.

(A) Thermal stability assays for wild-type Spx and its mutants (Spx-R100A,
Spx-C10S, Spx-T11A, Spx-C13S and Spx-K15A). (B) Thermal stability assays for
wild-type Spx and its mutants (Spx-R60E, Spx-S58A, Spx-P115A, Spx-D51R and
Spx-K66A). (C) Thermal stability assays for wild-type Spx and its mutants
(Spx-E109R, Spx-N106A, Spx-D108R, Spx-R91A and Spx-R92A). (D) Thermal

172	stability assays for wild-type Spx and its mutants (Spx-R128X, Spx-G52R,
173	Spx-R111A, Spx-E50A and Spx-D54R). If the thermostability curve of Spx mutants is
174	similar to that of wild-type Spx, the stabilities of these mutants will be similar to
175	wild-type Spx and it also means the mutant proteins folded properly.
176	
177	
178	
179	
180	
181	
182	
183	
184	
185	
186	
187	
188	
189	
190	
191	
192	
193	
194	
195	
196	
197	
198	
199	
200	
201	

202	Bsu_Spx		α1 00000000	β2 20 30		α2	α3 2020
203	Bsu_Spx Bsu_MgsR Bsu_YusI	MVTLYTSPSC MEQQLTFYSYPSC MSLTFYWYPKC	TSCRKAR TSCRKTKI GTCRKAKI	AW <mark>I</mark> EEHEIPFVEF HWIKAHQIEFNEF KWIEEHGKEINEJ	RNIFSEPLSIDI RHLFRETPTREI IHIAEQPPSKEI	LIKQILRMT LKYILSLT LKALYEKS	EDGTDEII TEGIDEIL GLDLKKFF
204	Sau_Spx Efa_Spx Ban_Spx1 Ban_Spx2	MVTLFTSPSC MLTLYTSPSC MVTLYSSPSC MVVLYTTASC	TSCRKAK TSCRKAR TSCRKAK ASCRKAK	A W J. Q E H D I P Y T E F A W J. Q E H E I P F K E F . W J. E E N H I P Y T E F A W J. E E N Q I D Y T E F	RNIFSEHLTIDI RNIFSEPLNIEI RNIFSDPLTIEI KNIVSNSMTVDI	EIKQILKMT ELKAILIMT EIKEILRMT ELKSILRLT	E D G TD E I I E D G TE E I I E S G TD E I I E E G A TE I I
205	Bme_Spx Bam_Spx Bce_Spx Bth_Spx	MVTLYTSPSC .MKELIFYSYPSC MVTLYSSPSC	TSCRKAK TSCRKTKI TSCRKAKI TSCRKAKI	A W LEENDIGYTEF HWLKAHNIDFHEF JWLEENHIPYTEF	RNIFSEPLSIDI RHLFRETPTIDI RNIFSDPLTIEI	EIKEILRMT ELKQILSLT EIKEILRMT	EDGTDEII TEGIDEIL ESGTDEII
206	Spn_ArsC Smu_Spx Spy_Spx	MLEFIEYPKC MVTLFLSPSC MVTLFLSPSC	STCKKAK TSCRKAR TSCRKAR	QELNQLGVDYKA AWLNRHDVVFQEH AWLVKHEVDFQEH	HIVEETPSQET INIMTSPLSRDI INIITSPLSRDI	/ILNWLETS SLLKILSYT SLMSILSFT	GFELKQFF ENGTEDII ENGTEDII
207	Ssa_Spx Str_Spx Lis_Spx Lmo_Spx	MLQFIEYPKC MITLFLSPSC MVTLYTSPSC MIWFYWYPKC	STCRKAR TSCRKAR TSCRKAR STCKKAR	ADJ, NQLGVDFEA AWJ, LNHEVPFQEH AWJ, EEHDIPYKEF AWJ, EEHDIPYKEF	/ D I V Q N T P S R D (H N I M T S P L S A P H R N I F S E P L S L D H / D I K T E T P T A E H	QLLDWIQNS ELQHILSLT EIKEILRMT ELQQWHEAS	DFELKSFF ENGTDDII EDGTDEII GLPVRRFF
208	Lla_Spxl Lla_Spx2 Lla_YhjF Lla_YnfG	MIDLYLSPSC MITIYTAPSC MYDFYWYPKC	TSCRKAR TSCKKAK STCRKAK SSSKKMK	AWDQSHKVPFVEH FWDSYHHIPFNEH AKDDELGIDYQA JWDITNKLEFEE	INILTQPMTTNI RNLIADPLSTTI IDLKETPPSQKI	DLRHILTKT EISQILQKC NFEKWFKEA	ENGTEDII DDGVEGLI DFPIKKFF ETGVEEII
209	Lla_YigC Lku_Spx Mca_Spx	MIKIYTVASC MVNLYTSPSC MVTLYTSPSC	SSCKKAKI TSCRKAR TSCRKAR	EWLEKQNLAYQEI AWLEEHEIPFEE SWLEEHDIPYQEF	INLVTSRICKEI RNIFSEPLSVDI RNIFSEPLSIEI	DILEILALT SIKQILRMT SIKEILRMT	E E G T G D I I E D G T D E I I E D G T D E I I
210	Cae_Spx Efa_Spx Eav_Spx	MIVLEVEYPKC IFTQKKLVLENEC MIKLYTSASC	STCKKAKI SVEREAKI TSCRKAKI	AW <mark>L</mark> DEHGVEYIDF RWLWKRBRRFIE AW <mark>L</mark> RENGLAFEEF	RDIVLDNPTADI 2NLIRQPLTKNI RNIMSNPLTKDI	ELATWIARS EFIHLLKLT EIKEILTMT	GLPVRRFF NNGVTEML ETGTDEII
211		cr4	-	<i>a</i> 5	m2 B3	84	<i>n</i> ³
212	Bsu_Spx				$\eta_2 \qquad \mu_3$ $200 \qquad \longrightarrow$ 90		
213	Bsu_Spx Bsu_MgsR Bsu_YusI Sau_Spx Efa_Spx	ATRSQTFKNIN. NTSGMKYREINLK STRSKTYQKIN.	LNIEE.M EKLYH.M VDIDS.LH	IVNEVLELLIEKE SEDEQLELLASDO LQDLYSIIQDNE	KLLRRPILVDI MLIKRPLTT.I GLLRRPIILDI	N.KKLVIGY DGEKVTVGF N.KRLQVGY	NPGELLKL KEDQFEEN NEDEIRRF
214	Ban_Spx1 Ban_Spx2 Bme_Spx	STRSKVFQELN STRSKTFQDLN STRSKTFQKLD	VNLES.LI INIEE.LS VQVDA.MI	PLQDLYKMIRDYE SLNEFYKLIIEHE PLQDLYELIQQNE	GILRRPIMID LMLRRPIMLD GLLRR <mark>P</mark> IIID	. KRLQV <mark>G</mark> Y . KRLQI <mark>G</mark> F . KRLQV <mark>G</mark> Y	NEDEIRRF NDEEIRKF NEDEIRRF
215	Bam_Spx Bce_Spx Bth_Spx Spn_ArsC	ATRSQTFKNIN STRSKVFQELN STRSKVFQELN NTSGIKYRELGLK	LNLEE.M VNLES.L VNLES.L DKVGS.L	IVNEVLKLLTEK LQDLYKMIRDY LQDLYKMIRDY SNQEAAELLASDO	PKLLRRPILID GILRRPIMID GILRRPIMID GILRRPIMID GMLLKRPILVE	H.KKLVI <mark>G</mark> Y E.KRLQV <mark>G</mark> Y E.KRLQV <mark>G</mark> Y NGTVKQI G Y	NPGELMKL NEDEIRRF NEDEIRRF RKS.YEEL
216	Smu_Spx Spy_Spx Ssa_Spx	STRSKVFQKLD STRSKVFQKLD NTSGLKYRELGLK	IDVDE.LS IDVEE.LS EKVPH.LT	VSELINLISKNE SISDLIDLIAKNE FAQEAADLLSTD	SLLRR <mark>P</mark> IIMDI SLLRRPIIMD(MLIKRPLLVRI	N. KRMQI <mark>G</mark> F Q. KRMQIGF DNQILQIGY	NEDEIRAF NEDEIRAF RKT.YEEL
217	Lis_Spx Lmo_Spx Lla_Spx1	STRSKIFQKLN NTSGIKYRELGLK STRSKVFQKLA	VDLDS.LH DKLDT.LS VDVDN.L	LQQLFELIQKNE PEEAYKLLASD INELLDLVTEFE	GLLRRPIIID MLIKRPLTT.I NLLRRPIITD	E.KRLQVGY NGKEVTLGF KHLQI <mark>G</mark> F	NEDEIRRF KEEEFEAT NEDEIRAF
218	Lla_Spx2 Lla_YhjF Lla_YnfG Lla_YigC	SSRNRFVKTIG NTSGLVYRDINLK SKRSSAYKKISKI SRRSQAYQRIN	VDFED.IS DKLAT.LS IDFDA.L IDFET.IF	SLSQAIKIISEN SEKEMAELLSSN CLNSLVELIVQN KLNDLIQIIEEN	QIMRR <mark>P</mark> IIMDI GMLIKRPLLVKI 2KLLRRPLVVDI TLLRRPLIVDI	. KRLHV <mark>G</mark> Y NGQVQQI <mark>G</mark> F . HRLQV <mark>G</mark> Y H. KRLQV <mark>G</mark> Y	NEEEIRAF KESSYEKI NEDDIRKF NDDEIRKF
219	Lku_Spx Mca_Spx Cae_Spx Efa_Spx	STRSKVFQKLN STRSKTFQKLN NSSGMKYRELGLK	VQVES.LS VNLES.MH ARLDAGM	LQELFELMSEH LQDLYEVIKEN DRECYELLATD	GLLRRPIIMDI GLLRRPIIIDI GMLVKRPLLV.	E.KRLQV <mark>G</mark> Y E.KRLQVGY GDDFAIPGF	NEDEIRRF NEDEIRRF RESAWVEA
220	Eav_Spx Eav_Spx	STRSKVYEKLD.	VDFDE.LS	SLSELVDI IEKYE	SLLRRPLVFD	E. TKFQV <mark>G</mark> Y	NEDEIHQF
221	Bsu Spx	α6 000000000000000000000000000000000000	00				
222	Bsu_Spx Bsu_MgsR	120 LPRKVRSFQLREA SKKKTVHQSA	130 QRLAN				
223	Bsu_YusI Sau_Spx Efa_Spx Ban_Spx1	WA LPRKVRTFQLQEA LPRDVRQLELRQA LPRTVRTFQLREA	QRMVD QLMAGL. QRLVN				· · · · · · · · · · · · · · · · · · ·
224	Ban_Spx2 Bme_Spx Bam_Spx	LPRSVRTFLNIEL LPRKVRTFQLREA TKKKTVHQSVS.	QKLAN QRLVNG.				· · · · · · · · · · · · · · · · · · ·
225	Bce_spx Bth_spx Spn_ArsC Smu_spx	LPRIVRIFQLREA GLK LPRDYRKOELROA	QRLVN QRLVN TIRAEVE(GEDD		· · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·
226	Spy_Spx Ssa_Spx Str_Spx Lis_Spx	LSRDYRKQELRQA GLLPRSYRKEELRSA	TIKAEIEC TMRADIQ	3			· · · · · · · · · · · ·
227	Lla_Spx1 Lla_Spx2	WKA LPREYRRA LPRTVRVLENGGA	EMLSTID RLRSAI.			· · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·
228	Lla_YhjF Lla_YnfG Lla_YigC Lku_Sow	LPRKVRQLGLITA LPRKKREIQIKVA	IENVI TEASYNLI	RIWDQAK DIEEQFHEG			· · · · · · · · · · ·
229	Mca_Spx Cae_Spx Efa_Spx Eav_Spx	LPRKVRTYQLREA LL LPKTKRKQELNEL LPREVRRVTSKK	QRLVN	DFKENNLCVPVIS	SCFSNMLVNKNI	VPGKATTSP	GNTALSNE
230		THE ARM AND A TO UNIT					
231 Figure S	11 Prote	in sequence a	lignme	nt of Spx fr	om non-re	dundan	t 28

232	low-GC-content Gram-positive bacterial species. The sequences were extracted
233	from UniProt Database by BLAST. The alignment was performed by Clustal Omega
234	and the figure was prepared using ESPript 3.0 (3). The invariant residues among Spx
235	from different low-GC-content Gram-positive bacteria are highlighted in red, and
236	conserved amino acids are boxed.
237	
238	
239	
240	
241	
242	
243	
244	
245	
246	
247	
248	
249	
250	
251	
252	
253	

254 Table S1. Primers and sequences used in this study.

Primer name	Sequence (5 to 3)
trxA-mango_F1	cggctgtgatcaggaaaaaataatttgtaagcattaaaatagcgtgaacgaatgggag
trxA-mango_F2	atagcgtgaacgaatgggagatgctatactaaaaatcatcatttcacattggagg
trxA-mango_R1	ggcacgtacgaatataccacataccaatccttccttcgtacg
trxA-mango_R2	accaatccttccttcgtacgtgccgccattattgaattcctccaatgtgaaatgatg
rrnJ P1-mango_F1	tagggaaaggatgccgctctttttaaatcccttagtatttcttcaaaaaaaa
rrnJ P1-mango_F2	cttcaaaaaaactattgcactattatttactaggtggtatattattattcgttgccgc
rrnJ P1-mango_R1	ggcacgtacgaatataccacataccaatccttccttcgtacgtgcc
rrnJ P1-mango_R2	tccttccttcgtacgtgcccgttatcgccttgtttagcggcaacgaataataatatacc
G52R sense	ctgcgtatgaccgaagatcgcaccgatgaaatc
G52R antisense	gatttcatcggtgcgatcttcggtcatacgcag
D54R sense	gaccgaagatggcacccgtgaaatcatctctacccg
D54R antisense	cgggtagagatgatttcacgggtgccatcttcggtc
C10S sense	gtacaccagcccgagcagcaccagctgcc
C10S antisense	ggcagctggtgctgctcgggctggtgtac
T11A sense	gcccgagctgcgccagctgccgt
T11A antisense	acggcagctggcgcagctcgggc
C13S sense	gagetgeaceageegtaaagegegtg
C13S antisense	cacgcgctttacggctgctggtgcagctc
K15A sense	tgcaccagctgccgtgcagcgcgtgcgtggc
K15A antisense	acgcacgcgctgcacggcagctggtgcagc
S58A sense	gcaccgatgaaatcatcgctacccgtagcaaagt
S58A antisense	actttgctacgggtagcgatgatttcatcggtgc
R60E sense	caccgatgaaatcatctctaccgagagcaaagttttccagaaactg
R60E antisense	cagtttctggaaaactttgctctcggtagagatgatttcatcggtg
K66A sense	cgtagcaaagttttccaggagctgaacgttaacgttgaaagc
K66A antisense	gctttcaacgttaacgttcagctcctggaaaactttgctacg
R91A sense	catccgggtctgctggctcgtccgatcatcatc
R91A antisense	gatgatcggacgagccagcagacccggatg
R92A sense	catccgggtctgctgcgtgctccgatcatcatcgatg
R92A antisense	catggatgatgatcggagcacgcagcagacccggatg
N106A sense	tctgcaggttggttacgccgaagatgaaatccgtc
N106A antisense	acggatttcatcttcggcgtaaccaacctgcagac
D108R sense	gttggttacaacgaagctgaaatccgtcgtttcc
D108R antisense	gaaacgacggatttcagcttcgttgtaaccaacc
E50A sense	cctgcgtatgaccgcagatggcaccgatga
E50A antisense	catcggtgccatctgcggtcatacgcaggatc
D51R sense	tcctgcgtatgaccgaacgtggcaccgatgaaatc
D51R antisense	gatttcatcggtgccacgttcggtcatacgcagg
R100A sense	ccgatcatcatcgatgaaaaagctctgcaggttggttac
R100A antisense	accaacctgcagagctttttcatcgatgatgatcggac

E109R sense	tggttacaacgaagatgcaatccgtcgtttcctgc
E109R antisense	gcaggaaacgacggattgcatcttcgttgtaacca
R111A sense	acaacgaagatgaaatcgctcgtttcctgccgcgtaaag
R111A antisense	acgcggcaggaaacgagcgatttcatcttcgttgtaac
P115A sense	gaaatccgtcgtttcctggcgcgtaaagttcgtag
P115A antisense	ctacgaactttacgcgccaggaaacgacggatttc
R128X sense	ccagctgcgtgaagcgcagtagctggcgaactaaaagcttg
R128X antisense	caagettttagttegeeagetaetgegetteaegeagetgg
pET28a_F	caccaccaccaccactg
pET28a_R	gctgccgcgggcaccag
Bs_a_F	gcctggtgccgcggcagcatgatcgagattgaaaaaaccaaaa
Bs_a_R	cagtggtggtggtggtggtgtcaatcgtctttgcgaagtcc
Bs_ σ^{A}_{F}	gcctggtgccgcggcagcatggctgataaacaaacccacg
Bs_ σ^{A}_{R}	ctcagtggtggtggtggtggtgttattcaaggaaatctttcaaacg

256

257

258

259

260

261 Supplementary References

- 262 1. Cardone, G., Heymann, J.B. and Steven, A.C. (2013) One number does not fit all: mapping
- local variations in resolution in cryo-EM reconstructions. J. Struct. Biol., 184, 226-236.
- 264 2. Tan, Y.Z., Baldwin, P.R., Davis, J.H., Williamson, J.R., Potter, C.S., Carragher, B. and
- 265 Lyumkis, D. (2017) Addressing preferred specimen orientation in single-particle cryo-EM
- through tilting. *Nature methods.*, **14(8)**, p.793.
- 267 3. Robert, X. and Gouet, P. (2014) Deciphering key features in protein structures with the new
- 268 ENDscript server. *Nucleic Acids Res.*, **42**, W320-W3244.