4482–4497 *Nucleic Acids Research, 2009, Vol. 37, No. 13* *Published online 27 May 2009*
doi:10.1093/nar/gkp419

Construction and functional analyses of a comprehensive σ^{54} site-directed mutant library using alanine–cysteine mutagenesis

Yan Xiao[1], Siva R. Wigneshweraraj[2], Robert Weinzierl[3], Yi-Ping Wang[1,*] and Martin Buck[3]

[1]National Laboratory of Protein Engineering and Plant Genetic Engineering, College of Life Sciences, Peking University, Beijing 100871, China, [2]Division of Investigative Sciences, Faculty of Medicine, Flowers Building and [3]Division of Biology, Faculty of Natural Sciences, Sir Alexander Fleming Building, Imperial College London, London SW7 2AZ, UK

Received March 9, 2009; Revised May 5, 2009; Accepted May 6, 2009

ABSTRACT

The σ^{54} factor associates with core RNA polymerase (RNAP) to form a holoenzyme that is unable to initiate transcription unless acted on by an activator protein. σ^{54} is closely involved in many steps of activator-dependent transcription, such as core RNAP binding, promoter recognition, activator interaction and open complex formation. To systematically define σ^{54} residues that contribute to each of these functions and to generate a resource for site specific protein labeling, a complete mutant library of σ^{54} was constructed by alanine–cysteine scanning mutagenesis. Amino acid residues from 3 to 476 of Cys(-)σ^{54} were systematically mutated to alanine and cysteine in groups of two adjacent residues at a time. The influences of each substitution pair upon the functions of σ^{54} were analyzed *in vivo* and *in vitro* and the functions of many residues were revealed for the first time. Increased σ^{54} isomerization activity seldom corresponded with an increased transcription activity of the holoenzyme, suggesting the steps after σ^{54} isomerization, likely to be changes in core RNAP structure, are also strictly regulated or rate limiting to open complex formation. A linkage between core RNAP-binding activity and activator responsiveness indicates that the σ^{54}-core RNAP interface changes upon activation.

INTRODUCTION

Major regulation of gene expression mainly depends upon the control of transcription initiation. In particular, the activity of RNA polymerase is regulated to control the level of gene expression. In bacteria, RNA polymerase holoenzyme is composed of core RNAP and a σ factor, which is responsible for the specific recognition and melting of promoter sequence (1). The σ factors in bacteria can be divided into two classes: the σ^{70} and σ^{54} class (2). Although they bind to the same core RNAP, these two classes of σ factors have no similarity in primary sequence and the modes of their transcription regulation are very different.

Unlike σ^{70}-holoenzyme, the initiation rates of σ^{54}-holoenzyme are mainly controlled via regulation of the DNA melting step (3,4). The σ^{54} -holoenzyme first binds promoter DNA in a transcriptionally silent closed complex (5–7). Conversion of the closed complex to a transcriptionally proficient open complex requires a specialized activator protein which interacts with σ^{54}-holoenzyme through DNA looping and brings about a conformational change in the closed promoter complex in a nucleotide hydrolysis dependent way, leading to σ^{54}-holoenzyme isomerization and promoter melting (7–11).

σ^{54} plays an important role in the process of open complex formation and interacts extensively with activator protein, promoter DNA and core RNAP to regulate the proper formation of the open complex. Open complex formation by σ^{54}-holoenzyme appears to involve the activator dependent relocation of three structurally distinct σ^{54} domains within the holoenzyme to (i) allow positioning of promoter DNA for entry into the active site of RNAP and (ii) removal of a σ^{54} associated protein density that would otherwise prevent DNA delivery (12). Based on biochemical and genetic characterization and fragmentation analyses, σ^{54} can be divided into three regions (2,13–16) (Figure 1). Region I comprises the amino terminal 56 residues and is important for activator responsiveness (10,17,18). Region II comprises residues 57–106 and is characterized by a predominance of acidic residues. Region II is not essential for the function of σ^{54}. In the carboxyl terminal (residues 107–477) lies Region III which

*To whom correspondence should be addressed. Tel: +86 10 6275 8490; Fax: +86 10 6275 6325; Email: wangyp@pku.edu.cn

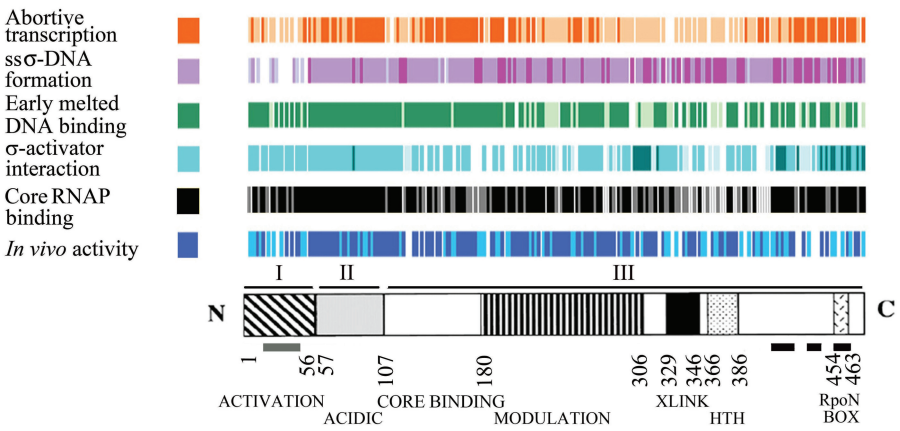

Figure 1. Diagram of the functional regions of σ^{54} and summary of the results. Grey bar below the diagram of σ^{54} indicates the predicted helical region in Region I. Black bars below the diagram of σ^{54} indicate the helices in the carboxyl terminal DNA binding domain based on the partial structure of σ^{54} from *Aquifex aeolicus* (20,21). Darker color is used to represent higher activity. The function each color represents is indicated in the left of the figure. The squares at the left of the colored bars indicate the color corresponding to the activity of Cys(−)σ^{54}. In the bar representing the activity of σ^{54}-activator interaction, the dark blue color indicates the variants that form additional complexes in the presence of $PspF_{1-275}$ and ADP·AlFx. The data are from Table 1.

contains a primary core RNAP-binding domain located between residues 120 and 215 (16), and a DNA-binding domain located between residues 329 and 477 (13,19,20). A modulation domain lies between these two domains and influences DNA-binding activity indirectly (14).

Due to the lack of a high-resolution structure of σ^{54}, it is still not clear which residues in σ^{54} are responsible for the interactions with itself or other elements such as core RNAP and promoter DNA in the process of transcription initiation. Nor are the primary σ^{54} sequences of densities evident in electron microscopy studies precisely delineated (12). To systematically define the functions of each residue in σ^{54}, a site-directed mutant library of σ^{54} was constructed using alanine–cysteine scanning mutagenesis. Mutation of residues other than glycine to alanine can be viewed as side-chain deletion and alanine scanning has been widely used in the research of protein functions (21,22), while mutation to cysteine can facilitate site specific protein modification (23–25). Alanine–cysteine scanning mutagenesis used in this article combines the advantages of these two kinds of mutations and facilitates both structural and functional analyses of a protein.

Here, residues 3–476 of a cysteine free form of *Klebsiella pneumoniae* σ^{54} were changed to alanine and cysteine in groups of two consecutive residues, producing a total of 237 variants of σ^{54}. This single cysteine mutant library is very helpful in both the structural and functional analyses of σ^{54}, and this paper focuses on the latter.

The results indicate that: (i) two parts of the Region I are the main targets for activator interaction; (ii) the carboxyl terminal of modulation domain can affect σ^{54} isomerization, probably through it influencing the −12 promoter sequence interaction; (iii) the carboxyl terminal part of DNA-binding domain is important for maintain the *in vivo* integrity of σ^{54}, but less important in the regulation of σ^{54} activity. The linkage between different functions of σ^{54} and possible mechanisms are discussed. The single cysteine mutant library of σ^{54} provides a powerful tool for further structural analyses using ensemble and single molecule Foster Resonant Energy Transfer (FRET) approaches (26).

MATERIALS AND METHODS

Site-directed mutagenesis

The pET28 based plasmid pSRWCys(−) (25), which directs the synthesis of the amino-terminal 6-His-tagged Cys(−)σ^{54}, was used as the template to create the single cysteine variants of σ^{54} using the Quickchange mutagenesis kit (Stratagene). The whole *rpoN* gene directing the synthesis of σ^{54} protein in each mutated plasmid was sequenced to ensure that only the desired substitutions were present.

β-Galactosidase assays

Plasmids carrying mutated *rpoN* genes were transformed to *Escherichia coli rpoN* knockout strain TH1/pMB221 (27). pMB221 contains both the reporter gene *lacZ* and the activator gene *nifA*. The reporter gene *lacZ* is under the control of *K. pneumoniae nifH* promoter, and the activator NifA is expressed from *bla* promoter. Transformants were grown at 37°C and then 100 µl of the overnight culture was used to inoculate 10 ml of LB with 50 µg/ml kanamycin and 34 µg/ml chloromycetin. Cells were grown at 37°C until A_{600} reaches 0.8. σ^{54}-dependent gene expression in this system depend on a basal T7-RNAP independent expression of *rpoN* from pET28. The average β-galactosidase results from three independent colonies are presented.

Immunoblotting

Mutated plasmids were transformed to TH1/pMB221 and were cultured under the same condition used in β-galactosidase assays. Cells (1 ml) were then collected by centrifugation and resuspended in 60 µl of 10 mM Tris, 0.1 mM

EDTA, pH 7.9. Ten microliters of the concentrated cell solution was lysed with 10 µl of 2× SDS sample buffer, heated at 95°C and 10 µl of each were loaded. Proteins were separated on denaturing 8% SDS–PAGE gels and blotted onto PVDF membranes. Anti-σ^{54} (prepared by Institute of Genetics and Developmental Biology, Chinese Academy of Sciences) and alkaline phosphatase conjugated anti-rabbit IgG (Pierce) were used for detection.

Protein purification

Mutated σ^{54} proteins were overexpressed in *E. coli* B834 (DE3). Two hundred microliters of overnight cultures were used to inoculate 20 ml of LB with 50 µg/ml kanamycin and cells were grown at 37°C until A_{600} reached 0.6. The cultures were then shifted to 25°C and IPTG was added to 1 mM. Incubation continued for 2 h at 25°C. Cells were collected by centrifugation, and resuspended in 500 µl of Buffer A [25 mM sodium phosphate (pH 7.0), 50 mM NaCl, 5% (v/v) glycerol]. The concentrated cells were lysed by adding 55 µl of FastBreak™ Reagent (Promega) into the buffer and were shaken for 15 min at room temperature. The lysates were then incubated for 3 min with 450 µl of Ni–NTA resin (Qiagen) in the columns, which were then washed with 500 µl of Buffer B (25 mM sodium phosphate (pH 7.0), 50 mM NaCl, 30 mM imidazole, 5% (v/v) glycerol). Mutated σ^{54} proteins were eluted in 400 µl of Buffer C [25 mM sodium phosphate (pH 7.0), 50 mM NaCl, 1 M imidazole, 5% (v/v) glycerol]. Purified σ^{54} proteins were dialysed against 10 mM Tris (pH 7.5), 0.1 mM EDTA, 100 mM NaCl, 50% glycerol, and stored at −20°C.

Core RNAP-binding assays

E. coli core RNAP (Epicentre, 250 nM) and different amounts of σ^{54} variants were mixed together in Tris–NaCl buffer (40 mM Tris–HCl (pH 8.0), 10% (v/v) glycerol, 0.1 mM EDTA, 1 mM DTT, 100 mM NaCl) and incubated for 10 min at 30°C, followed by the addition of glycerol bromophenol blue loading dye. Samples were loaded onto native 4.5% PAGE gels and run at 50 V for 2 h at room temperature in TG buffer (25 mM Tris, 200 mM glycine, pH 8.6). Proteins were visualized by Coomasie Blue staining.

Gel mobility shift assay

^{32}P-labeled 88-mer top strand −12/−11 mismatched heteroduplex DNA fragment consisting of the −60 to +28 *Sinorhizobium meliloti nifH* promoter sequence was used as a template for gel mobility shift assays. The reactions contained 20 nM DNA, 1 µM σ^{54} in STA buffer [25 mM Tris–acetate (pH 8.0), 8 mM magnesium acetate, 10 mM KCl, 3.5% (w/v) PEG 8000] and were incubated for 10 min at 30°C. Where indicated, $PspF_{1-275}$ (5 µM), dGTP (4 mM) were added for 5 min prior to gel loading. Free DNA and σ^{54} bound DNA were separated on 4.5% native PAGE gels and run at 80 V for 70 min at room temperature in TG buffer (25 mM Tris, 200 mM glycine, pH 8.6). Quantitative data were from phosphorimager analyses of the gels.

σ^{54}-activator interaction

Twenty micromolar $PspF_{1-275}$ and 1 µM of σ^{54} were incubated at 30°C for 5 min in the above STA buffer with ADP (0.2 mM) and NaF (5.0 mM). After addition of $AlCl_3$ (0.2 mM) the samples were incubated for a further 10 min and loaded onto 4.5% native PAGE gels and run at 120 V for 50 min at room temperature in the above TG buffer. Proteins were visualized by Coomassie Blue staining.

Abortive transcription assay

Open complexes were formed on plasmid pMKC28 (28) (10 nM final concentration) in the presence of 100 nM σ^{54}-holoenzyme (constituted using 2.5:1 molar ratio of σ^{54} over core RNAP), 20 U RNase inhibitor, 4 mM dATP and 5 µM of $PspF_{1-275}$. A mix containing 100 µg/ml heparin, 0.5 mM UpG, 4 µCi [α-^{32}P] GTP was added to the reaction for the abortive synthesis of a four nucleotide transcript (UGGG). The reactions were separated on a 20% denaturing gel and quantified as in gel mobility shift assay. Replicate assays were conducted to confirm the low activities of the variants defective in abortive transcription.

RESULTS

Mutagenesis

σ^{54} from *K. pneumoniae* contains two cysteine residues at positions 198 and 346 (29). Substitution of these two cysteine residues with alanine does not affect the function of σ^{54} significantly (25). Therefore, in order to construct variants carrying a single cysteine, the pET28b based plasmid pSRWCys(–), which directs the synthesis of the Cys (–)σ^{54}, was used as the template (25). Residues of Cys (–)σ^{54} from 3 to 476 were systematically substituted with alanine–cysteine in groups of two adjacent residues at a time, for a total of 237 variants of σ^{54}. Substitution with alanine can be viewed as a side-chain deletion, whereas substitution with cysteine facilitates side-chain modification at specific positions. Each mutated *rpoN* gene was sequenced in its entirety to establish the presence of only the expected substitutions.

Region II is not essential for the *in vivo* activity of σ^{54}

β-Galactosidase assays were used to test the *in vivo* activities of the variants. The plasmid carrying the mutated *rpoN* gene and another plasmid, pMB221 carrying the reporter and activator genes (27), were transformed to *E. coli* TH1, which lacks functional σ^{54}. The expression of mutated σ^{54} depends on the low level of basal expression of *rpoN* gene from the pET28 based vector. The β-galactosidase gene is under the control of *K. pneumoniae nifH* promoter and the activator NifA is constitutively expressed from the *bla* promoter. NifA acts with σ^{54} to activate the expression of β-galactosidase gene from σ^{54} dependent *nifH* promoter. Therefore, the activity of β-galactosidase can be used to indicate the *in vivo* activity of σ^{54} variants.

The mutations that affected the *in vivo* activity of σ^{54} are distributed extensively in Region I and III, and are

relatively concentrated in the carboxyl terminal of σ^{54} that includes the −24 DNA-binding bihelical structure (30) (Table 1). Consistent with the fact that Region II is not well conserved in sequence, none of the Region II variants exhibited poor activity.

Low activities of the variants could be caused by protein instability. Immunoblotting was therefore carried out to measure the expression levels of the variants which retained <20% activity compared to Cys(–)σ^{54}. σ^{54} variants were expressed under the same condition used in the *in vivo* activity assay. Variants found at lower than wild-type expression levels have substitutions mainly located in the DNA-binding domain [AAs 329–477, defined previously by proteolysis experiments and including the −24 promoter recognition structure (30,31)], especially in the carboxyl terminal, between residues 431 and 476 (Figure 2), indicating that many residues in this region are important for maintaining the structural integrity of σ^{54} *in vivo*, and that the instabilities of these variants may be one of the reasons for their low *in vivo* activities. However, the series of purified σ^{54} variants changed in residues 457–476 migrated as discrete bands in native gels, each with a mobility similar to the Cys(–)σ^{54} (when loaded at higher concentrations for detection), suggesting that their gross structural integrity is not seriously affected *in vitro* (Figure S1A). These variants were detected in native gels at concentrations where the Cys(–)σ^{54} was not easily detected, indicating a less diffuse behavior, and by inference some differences in conformation compared to the control protein Cys(–)σ^{54}.

Residues contributing to core RNAP binding are extensively distributed throughout Region I and III

Next, all the variants as well as Cys(–)σ^{54} were purified by Ni affinity chromatography for further functional analyses *in vitro*. An essential and first step en route to open

Table 1. Summary of the activities of σ^{54} variants[a]

σ^{54} mutant	Functional region	*In vivo* trxn[b]	Core binding	Activator contact	Early melted DNA binding	Isomerization	Abortive trxn
Cys(–)		100	wt-like	+	100	100	100
Region I							
AC3/4	Activation	31	wt-like	+	86	34	39
AC5/6	Activation	59	weak	+	93	83	83
AC7/8	Activation	51	wt-like	+	85	22	48
AC9/10	Activation	66	wt-like	+	88	41	49
AC11/12	Activation	35	wt-like	+/–	91	0	18
AC13/14	Activation	76	wt-like	+	107	0	98
AC15/16	Activation	18	wt-like	+	89	0	44
AC17/18	Activation	23	wt-like	–	66	0	16
AC19/20	Activation	9	weak	+	46	59	42
AC21/22	Activation	27	wt-like	+	29	78	40
AC23/24	Activation	1	wt-like	+	63	52	19
AC25/26	Activation	6	wt-like	+	21	0	25
AC27/28	Activation	47	wt-like	+	77	0	37
AC29/30	Activation	16	wt-like	–	19	0	14
AC31/32	Activation	72	wt-like	+	95	0	42
AC33/34	Activation	16	wt-like	+	7	0	15
AC35/36	Activation	75	wt-like	+	90	0	34
AC37/38	Activation	8	wt-like	–	15	144	21
AC39/40	Activation	75	wt-like	+	92	52	29
AC41/42	Activation	109	wt-like	+	72	0	47
AC43/44	Activation	49	wt-like	+	58	55	31
AC45/46	Activation	30	wt-like	+	95	26	71
AC47/48	Activation	4	wt-like	–	14	0	23
AC49/50	Activation	84	wt-like	+	92	168	109
AC51/52	Activation	72	wt-like	+	102	110	59
AC53/54	Activation	81	wt-like	+	98	109	87
AC55/56	Activation	71	wt-like	+	92	105	57
Region II							
AC57/58	Acidic	72	wt-like	+	95	72	46
AC59/60	Acidic	81	wt-like	+	97	101	128
AC61/62	Acidic	71	wt-like	+	78	98	78
AC63/64	Acidic	83	wt-like	+	94	89	80
AC65/66	Acidic	117	wt-like	+	90	93	45
AC67/68	Acidic	108	wt-like	+	89	94	76
AC69/70	Acidic	81	wt-like	+	93	96	87
AC71/72	Acidic	92	wt-like	+	88	84	38
AC73/74	Acidic	44	wt-like	+	88	98	75
AC75/76	Acidic	128	wt-like	+	90	99	46
AC77/78	Acidic	133	wt-like	+	86	97	51
AC79/80	Acidic	37	wt-like	+	88	91	74

(continued)

Table 1. Continued

σ^{54} mutant	Functional region	*In vivo* trxn[b]	Core binding	Activator contact	Early melted DNA binding	Isomerization	Abortive trxn
AC81/82	Acidic	102	wt-like	+	89	104	76
AC83/84	Acidic	70	wt-like	+[c]	99	192	51
AC85/86	Acidic	97	wt-like	+	92	119	100
AC87/88	Acidic	47	wt-like	+	92	95	145
AC89/90	Acidic	76	wt-like	+	87	164	92
AC91/92	Acidic	38	wt-like	+	87	90	117
AC93/94	Acidic	39	wt-like	+	74	117	117
AC95/96	Acidic	78	wt-like	+	81	89	120
AC97/98	Acidic	28	wt-like	+	86	102	148
AC99/100	Acidic	122	wt-like	+	79	108	111
AC101/102	Acidic	104	wt-like	+	87	97	117
AC103/104	Acidic	108	wt-like	+	88	93	89
AC105/106	Acidic	109	wt-like	+	95	98	72
Region III							
AC107/108	Core binding	110	wt-like	+	94	110	29
AC109/110	Core binding	66	wt-like	+	100	60	34
AC111/112	Core binding	111	wt-like	+	92	78	23
AC113/114	Core binding	72	wt-like	+	90	81	25
AC115/116	Core binding	97	wt-like	+	81	85	57
AC117/118	Core binding	79	wt-like	+	79	66	85
AC119/120	Core binding	33	wt-like	+	101	189	46
AC121/122	Core binding	88	weak	–	53	181	38
AC123/124	Core binding	8	wt-like	+/–	93	225	51
AC125/126	Core binding	14	wt-like	+/–	107	131	20
AC127/128	Core binding	8	wt-like	+/–	96	89	37
AC129/130	Core binding	60	wt-like	+	93	66	148
AC131/132	Core binding	63	wt-like	+	95	83	33
AC133/134	Core binding	16	wt-like	–	65	138	46
AC135/136	Core binding	65	wt-like	+	92	69	127
AC137/138	Core binding	19	wt-like	–	87	95	36
AC139/140	Core binding	61	weak	–	88	113	152
AC141/142	Core binding	100	wt-like	+	88	75	174
AC143/144	Core binding	107	wt-like	+	97	102	114
AC145/146	Core binding	16	wt-like	+/–	92	78	24
AC147/148	Core binding	89	wt-like	+	102	67	124
AC149/150	Core binding	96	wt-like	–	75	27	77
AC151/152	Core binding	6	weak	–	81	83	64
AC153/154	Core binding	24	wt-like	+	107	0	37
AC155/156	Core binding	83	wt-like	+	98	55	185
AC157/158	Core binding	102	wt-like	+	100	75	153
AC159/160	Core binding	32	wt-like	–	44	85	115
AC161/162	Core binding	27	wt-like	–	90	113	86
AC163/164	Core binding	100	wt-like	+	85	77	117
AC165/166	Core binding	83	wt-like	–	102	95	87
AC167/168	Core binding	87	wt-like	+	72	85	104
AC169/170	Core binding	79	wt-like	+	97	98	142
AC171/172	Core binding	106	wt-like	+	103	97	172
AC173/174	Core binding	65	wt-like	–	82	87	67
AC175/176	Core binding	65	weak	–	91	106	126
AC177/178	Core binding	74	wt-like	–	80	168	103
AC179/180	Core binding	23	weak	–	80	218	29
AC181/182	Core binding	52	wt-like	–	78	169	136
AC183/184	Core binding	47	wt-like	+	81	123	26
AC185/186	Core binding	2	wt-like	–	83	143	43
AC187/188	Core binding	9	wt-like	–	90	115	27
AC189/190	Core binding	115	wt-like	–	69	71	154
AC191/192	Core binding	92	wt-like	+	87	90	156
AC193/194	Core binding	21	wt-like	+	90	147	190
AC195/196	Core binding	30	wt-like	+/–	81	153	113
AC197/198	Core binding	99	wt-like	+	88	181	104
AC199/200	Core binding	12	wt-like	–	57	201	26
AC201/202	Core binding	27	wt-like	–	64	131	27
AC203/204	Core binding	55	wt-like	–	64	169	64
AC205/206	Core binding	98	wt-like	–	95	134	138
AC207/208	Core binding	85	wt-like	+	28	188	77
AC209/210	Core binding	69	wt-like	–	21	148	56
AC211/212	Core binding	87	wt-like	+	93	107	69
AC213/214	Core binding	96	wt-like	–	86	122	47

(continued)

Table 1. Continued

σ^{54} mutant	Functional region	*In vivo* trxn[b]	Core binding	Activator contact	Early melted DNA binding	Isomerization	Abortive trxn
AC215/216	Core binding	71	wt-like	–	44	136	38
AC217/218	Modulation	97	wt-like	+	69	180	61
AC219/220	Modulation	23	wt-like	–	38	269	26
AC221/222	Modulation	89	wt-like	+	77	126	74
AC223/224	Modulation	25	wt-like	–	43	138	16
AC225/226	Modulation	108	wt-like	+	79	144	35
AC227/228	Modulation	5	wt-like	+	91	140	46
AC229/230	Modulation	59	wt-like	+	72	144	61
AC231/232	Modulation	61	weak	–	57	122	42
AC233/234	Modulation	62	wt-like	+	33	230	56
AC235/236	Modulation	30	wt-like	–	46	175	33
AC237/238	Modulation	69	wt-like	+	47	131	40
AC239/240	Modulation	90	wt-like	+/–	56	63	43
AC241/242	Modulation	74	weak	–	20	143	32
AC243/244	Modulation	77	wt-like	–	78	109	57
AC245/246	Modulation	70	wt-like	+	75	171	56
AC247/248	Modulation	72	wt-like	–	66	85	49
AC249/250	Modulation	89	wt-like	+	92	115	76
AC251/252	Modulation	52	wt-like	+/–	47	124	56
AC253/254	Modulation	23	weak	–	87	139	62
AC255/256	Modulation	74	wt-like	–	53	135	60
AC257/258	Modulation	89	wt-like	+	71	145	75
AC259/260	Modulation	24	wt-like	+	84	187	20
AC261/262	Modulation	86	wt-like	+	80	145	44
AC263/264	Modulation	102	wt-like	+	85	125	60
AC265/266	Modulation	107	wt-like	+	98	114	76
AC267/268	Modulation	98	wt-like	+	98	146	82
AC269/270	Modulation	47	wt-like	+	96	118	50
AC271/272	Modulation	86	wt-like	+	97	92	74
AC273/274	Modulation	71	wt-like	+	82	102	50
AC275/276	Modulation	102	weak	–	69	153	18
AC277/278	Modulation	78	weak	–	59	128	31
AC279/280	Modulation	91	weak	+/–	51	153	33
AC281/282	Modulation	106	wt-like	+	67	136	51
AC283/284	Modulation	102	wt-like	+/–	73	103	43
AC285/286	Modulation	7	weak	–	88	95	27
AC287/288	Modulation	61	wt-like	+	106	100	63
AC289/290	Modulation	49	wt-like	+/–	81	68	43
AC291/292	Modulation	78	wt-like	+	79	89	35
AC293/294	Modulation	92	wt-like	+	88	103	50
AC295/296	Modulation	103	wt-like	+	17	0	36
AC297/298	Modulation	28	weak	+/–	8	153	30
AC299/300	Modulation	8	wt-like	+[c]	18	0	27
AC301/302	Modulation	96	wt-like	+[c]	50	142	38
AC303/304	Modulation	20	wt-like	+[c]	62	96	32
AC305/306	Modulation	55	wt-like	+[c]	51	133	60
AC307/308		92	wt-like	+[c]	58	185	51
AC309/310		119	wt-like	+[c]	37	163	42
AC311/312		68	wt-like	+[c]	39	173	31
AC313/314		103	wt-like	+/–	50	122	51
AC315/316		95	wt-like	+/–	61	237	34
AC317/318		17	weak	+[c]	67	141	42
AC319/320		6	wt-like	+	64	54	20
AC321/322		124	wt-like	+	90	181	36
AC323/324		52	wt-like	+	83	0	22
AC325/326		48	wt-like	+	48	92	21
AC327/328		12	wt-like	+	39	59	15
AC329/330	Xlink	7	wt-like	+	4	451	16
AC331/332	Xlink	39	wt-like	+	81	72	32
AC333/334	Xlink	5	wt-like	+/–	13	294	21
AC335/336	Xlink	9	weak	+	7	0	45
AC337/338	Xlink	89	wt-like	+/–	80	71	26
AC339/340	Xlink	34	wt-like	+/–	13	265	30
AC341/342	Xlink	73	weak	–	46	152	36
AC343/344	Xlink	93	weak	–	85	73	22
AC345/346	Xlink	86	wt-like	+	93	107	40
AC347/348		12[d]	weak	–	10	290	13
AC349/350		99	wt-like	+	101	158	58

(continued)

Table 1. Continued

σ^{54} mutant	Functional region	*In vivo* trxn[b]	Core binding	Activator contact	Early melted DNA binding	Isomerization	Abortive trxn
AC351/352		33	weak	+	11	77	41
AC353/354		41	weak	–	61	144	39
AC355/356		4	wt-like	–	3	193	15
AC357/358		16	weak	+/–[c]	20	504	27
AC359/360		93	weak	–	58	148	52
AC361/362		37	weak	+/–[c]	50	175	54
AC363/364		49	wt-like	+/–[c]	7	306	42
AC365/366	HTH	95	weak	–	41	235	49
AC367/368	HTH	75	weak	–	17	343	43
AC369/370	HTH	29	weak	–	12	246	16
AC371/372	HTH	69	wt-like	+	95	109	29
AC373/374	HTH	79	wt-like	+	97	94	43
AC375/376	HTH	91	weak	+	92	78	20
AC377/378	HTH	68	weak	+/–	88	124	33
AC379/380	HTH	2	wt-like	+	10	1654	19
AC381/382	HTH	45	weak	–	10	600	32
AC383/384	HTH	0	wt-like	+	1	637	16
AC385/386	HTH	10	wt-like	+	78	66	66
AC387/388		23	wt-like	+	56	27	30
AC389/390		16	weak	–	67	111	31
AC391/392		98	wt-like	+	104	120	39
AC393/394		50	weak	–	71	146	28
AC395/396		42	weak	–	50	146	60
AC397/398		85	weak	–	84	118	81
AC399/400		23	weak	–	14	142	30
AC401/402		25	weak	+/–	30	122	69
AC403/404		7	weak	+/–	30	131	25
AC405/406		45	wt-like	+	95	123	41
AC407/408		20	wt-like	+/–	74	179	31
AC409/410		61	wt-like	+[c]	71	153	55
AC411/412		39	wt-like	+[c]	59	205	70
AC413/414		87	wt-like	+[c]	94	123	52
AC415/416		53	wt-like	+[c]	9	929	54
AC417/418		27	wt-like	+	81	133	34
AC419/420		72	wt-like	+	90	113	67
AC421/422		122	wt-like	+	98	128	44
AC423/424		10	wt-like	–	28	369	115
AC425/426		12	weak	+[c]	18	574	68
AC427/428		16	wt-like	+/–	39	296	107
AC429/430		90	wt-like	+/–	91	119	85
AC431/432		0[d]	wt-like	+/–	3	202	41
AC433/434		46	wt-like	+/–	59	177	79
AC435/436		18	wt-like	–	62	190	96
AC437/438		8[d]	wt-like	+	78	124	77
AC439/440		0[d]	wt-like	+/–	10	339	34
AC441/442		15	wt-like	+	17	356	135
AC443/444		96	wt-like	+[c]	89	138	82
AC445/446		13	wt-like	+	84	103	99
AC447/448		96	wt-like	+[c]	89	116	92
AC449/450		8[d]	wt-like	+	69	171	59
AC451/452		21	wt-like	+[c]	58	150	112
AC453/454	RpoN box	34	wt-like	+[c]	60	210	67
AC455/456	RpoN box	0	wt-like	+	67	0	23
AC457/458	RpoN box	0[d]	wt-like	+	2	251	65
AC459/460	RpoN box	23	wt-like	+[c]	35	210	121
AC461/462	RpoN box	0[d]	wt-like	+	42	156	28
AC463/464	RpoN box	86	wt-like	+[c]	104	141	71
AC465/466		52	wt-like	+[c]	74	109	99
AC467/468		1[d]	wt-like	+	24	303	62
AC469/470		99	wt-like	+	93	146	61
AC471/472		66	wt-like	+[c]	10	132	58
AC473/474		0[d]	wt-like	+	19	174	23
AC475/476		32	wt-like	+[c]	32	134	70

+: σ^{54} variants bind $PspF_{1-275}$ normally. +/–: σ^{54} variants bind $PspF_{1-275}$ weakly. –: σ^{54} variants cannot bind $PspF_{1-275}$.
[a]Activities in all the assays are measured relative to Cys(–)σ^{54}. Isomerization ability of σ^{54} is calculated as a percentage of initially bound DNA converted to ssσ–DNA in an activator and dGTP dependent way.
[b]Tested using β-galactosidase assay.
[c]New complexes were formed when $PspF_{1-275}$ was trapped with σ^{54} in the presence of ADP·AlFx.
[d]Low level of *in vivo* expression.

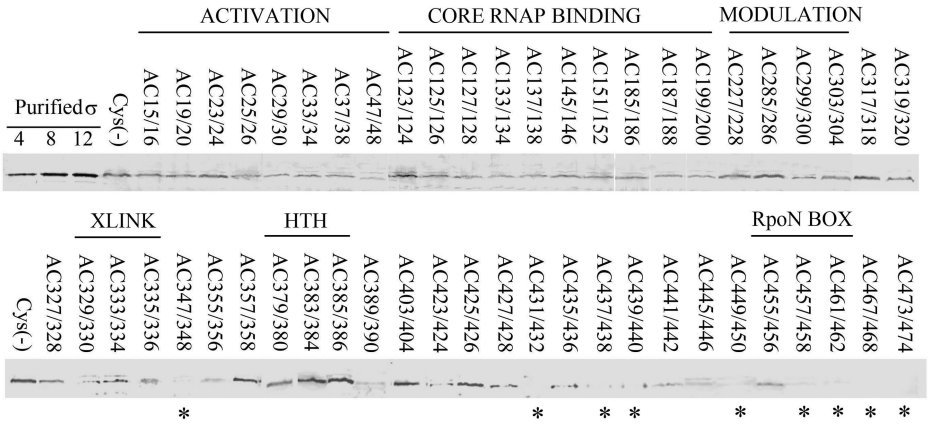

Figure 2. *In vivo* stability of σ^{54} variants which retained <20% *in vivo* activity compared to Cys(-)σ^{54}. The functional regions in which these variants are located are also shown. The asterisks indicate the variants with very low expression levels. Immunoblots were of lysates of TH1 cells used in *in vivo* activity assays, containing pMB221 and mutant *rpoN* genes. Equal amounts of each cell extract were loaded. Immunoblots of purified σ^{54} samples are also shown presented at 4, 8 and 12 ng.

complex formation is holoenzyme forming between σ and core RNAP. Native gel mobility shift assay was used to measure the binding of σ^{54} variants to core RNAP. Holoenzyme can be detected in Coomassie blue stained gels based on its different mobility versus core RNAP.

Consistent with the view that the core RNAP interface of σ^{54} comprises at least two functionally important but distinct sequences: 1–56 and 120–215 (16), many mutations located in these two sequences altered the interaction of σ^{54} with core RNAP: Variants AC 5/6, AC 19/20, AC 121/123, AC 139/140, AC 151/152, AC 175/176 and AC 179/180 shifted all core RNAP at a molar ratio of 4:1 (σ^{54}: core RNAP), whereas Cys(–)σ^{54} can shift all core RNAP at a molar ratio of 1:1 (Figure 3). In Region I, mutations that affected binding of σ^{54} to core RNAP mainly locate in its amino terminus, while mutations between residues 39–56 had no effect upon holoenzyme formation (Figure S1A). Furthermore, in this assay, many residues in modulation domain or DNA-binding domain were found to be important for core RNAP binding (Figure 3), indicating residues affecting interaction of σ^{54} with core RNAP are widely distributed across Regions I and III. Although Region II is reported to play a role in the binding of σ^{54} to core RNAP (32), none of the mutations in Region II affected the core RNAP-binding activities in our assays (Figure S1B).

Regions I and III cooperate for activator interaction

σ^{54}-Holoenzyme initiates transcription in response to an interaction with an activator protein, and σ^{54} is the major and direct target for this interaction (7,33,34). To identify the residues in σ^{54} responsible for the stable interaction with activator protein, an assay to measure the interaction between the σ^{54} and the activator protein by native-PAGE of the reaction was used. A DNA-binding domain deletion form of the *E. coli* phage shock protein F ($PspF_{1-275}$) served as a model activator in these assays (35). Since the interaction between σ^{54} and its activators is normally very transient, stable complex formation between σ^{54} and $PspF_{1-275}$ depends on ADP·AlFx—an ATP analog that mimics the state of ATP at the point of hydrolysis, which 'traps' the transient complex between $PspF_{1-275}$ and σ^{54} so that it can be resolved and detected on native gels (34).

Results show that four Region I variants (AC17/18, AC29/30, AC37/38 and AC47/48.) failed to form the ADP·AlFx dependent stable complex with $PspF_{1-275}$, indicating the substituted residues contribute to part of the interface σ^{54} makes with activator protein (Figure 4). Mutations of these residues may alter this interface and prevent the variants from making stable interaction with activator protein. Low *in vivo* activities of these four variants may also be attributed to their altered interaction with activator protein.

Many Region III variants also cannot form a detectable stable complex with $PspF_{1-275}$, implying that determinants contributing to the interface between σ^{54} and activator protein are widely distributed. Considering that Region I constitute the primary target for activator interaction (34), it is probable that substitution of these Region III residues indirectly affect the stable σ^{54}-activator complex formation through the interaction with Region I.

Interestingly, some σ^{54} variants in the trapping reactions showed formation of a new complex which has not been observed before. This new complex has a slower mobility compared to the normal σ^{54}-activator complex (Figure 4). We speculate that this new complex might be a σ^{54}-activator complex with a new conformation, caused by the altered interaction of σ^{54} with $PspF_{1-275}$. The new complexes were detected under conditions supporting PspF–ADP·AlFx formation, but we did not determine if all the reaction components (AlCl, NaF, ADP) were required. However, the presence of ADP alone was not sufficient for formation of new complexes (data not shown). Among the variants with this phenotype, most have substitutions located in Region III and only one variant was in Region II, residues 83–84.

Promoter DNA binding

σ^{54} is a DNA-binding protein that can *in vitro* occupy some of its cognate promoters in the absence of core

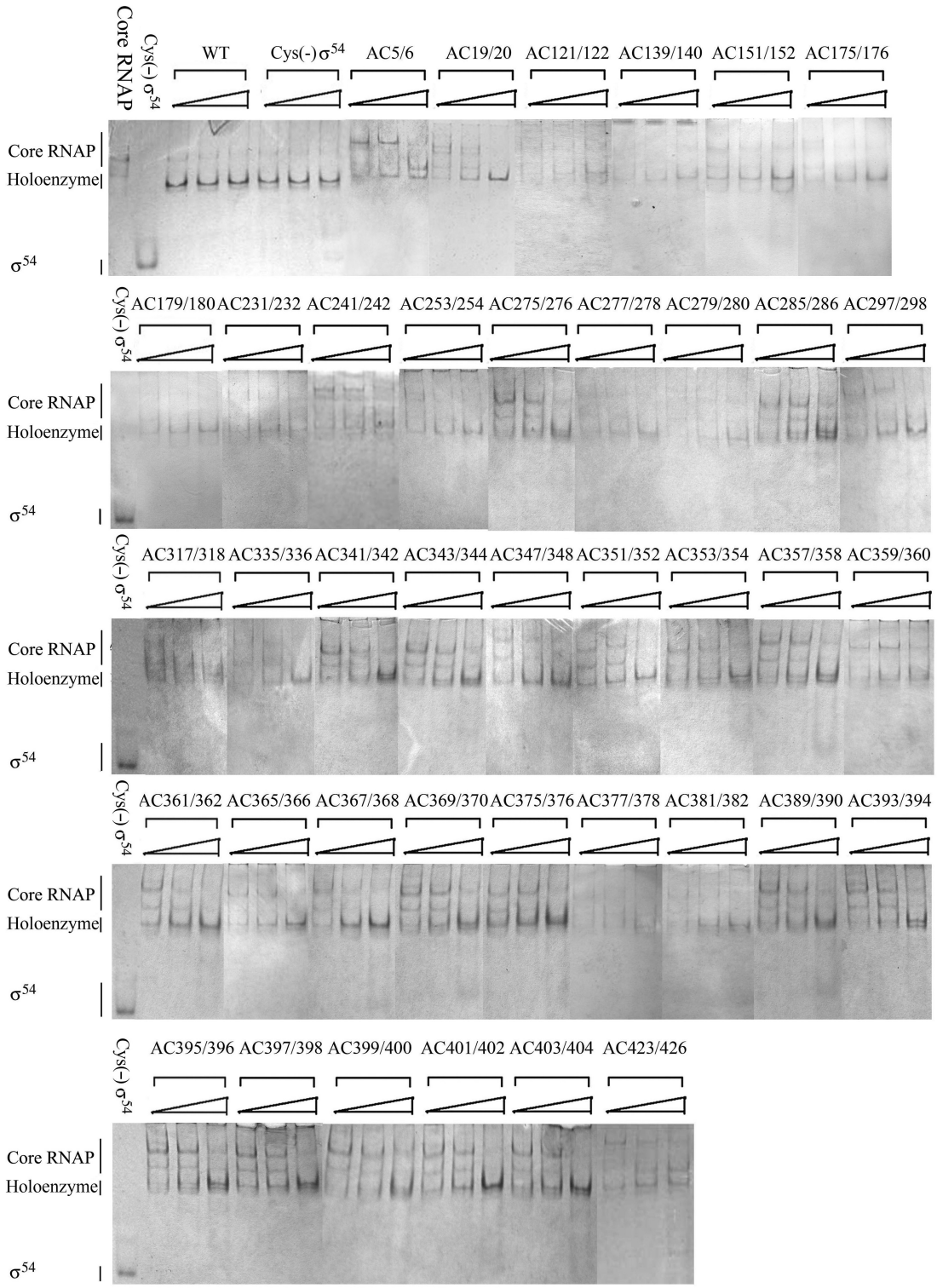

Figure 3. Formation of holoenzyme by σ^{54} variants which bind core RNAP with low affinity. Holoenzyme were formed at 1:1, 1:2 and 1:4 ratios of core RNAP (250 nM) to σ. Migration positions of core RNAP, holoenzyme and free σ are indicated.

RNAP (36). Interaction of σ^{54} with repressive fork junction promoter sequences around −12 plays an important role in the regulation of open complex formation (12, 37–39). To determine the σ^{54} sequences that contribute to promoter DNA binding, we conducted DNA-binding assay using heteroduplex promoter DNA (opened at −12 and −11, termed early melted DNA) to mimic the state of DNA in the closed complex (Figure 5A). Mutations that affected early melted DNA binding mainly located in two regions: the carboxyl terminal of Region I (residues from 25

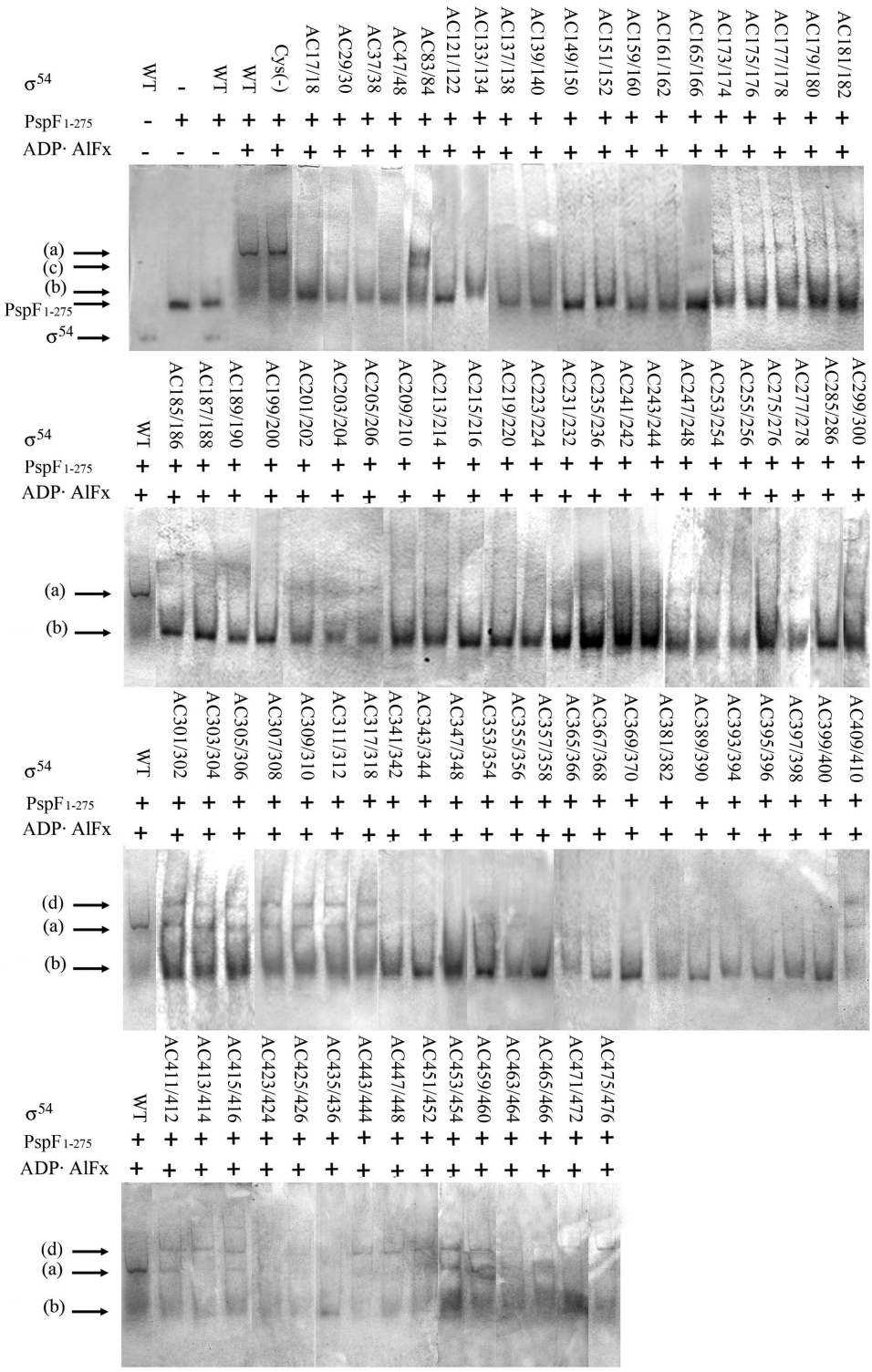

Figure 4. σ^{54} variants which cannot bind $PspF_{1–275}$ normally in the presence of ADP·AlFx. ADP·AlFx dependent complexes formed between σ^{54} and $PspF_{1–275}$ were detected by Coomassie staining. σ^{54} and $PspF_{1–275}$ were presented at 1 and 20 μM separately. Arrow (a) indicates the complex formed between σ^{54} and $PspF_{1–275}$ in the presence of ADP·AlFx. Arrow (b) indicates the changed position of $PspF_{1–275}$ complex in the presence of ADP·AlFx. Arrow (c) and (d) indicates the new bands formed in this assay.

to 48) and the DNA-binding domain (residues from 329 to 477). In Region I (residues from 25 to 38), it is interesting to note the periodicity of the DNA-binding defective phenotype (Figure 5B). Some other mutations located in modulation domain (residues 207–210, 241–242 and 295–300) also affect early melted DNA binding, consistent with the view that modulation domain functions to assist DNA binding (14).

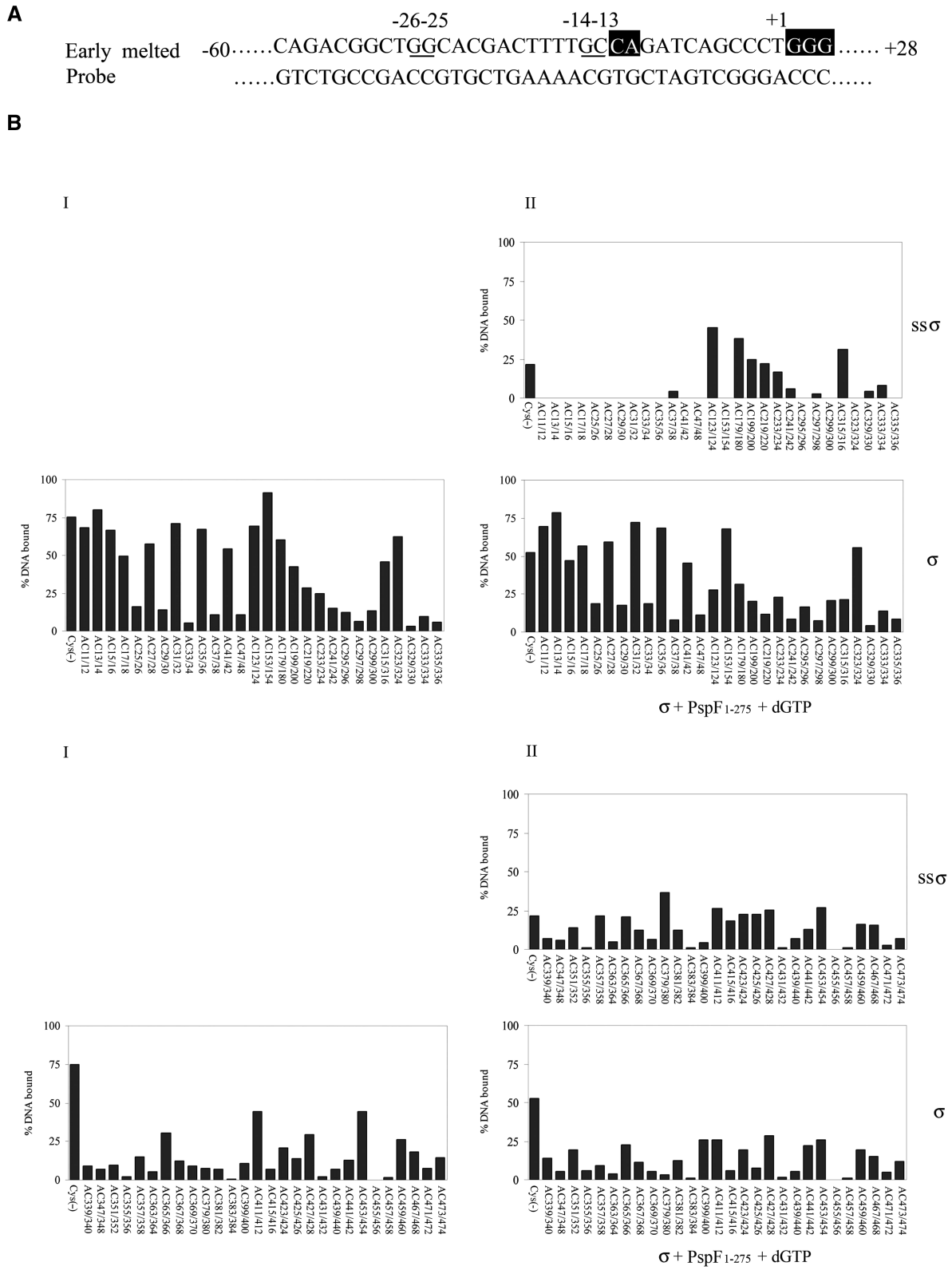

Figure 5. (**A**) *S. meliloti nifH* heteroduplex promoter probe used for DNA binding and σ^{54} isomerization assay. The consensus GG and GC elements are underlined. The mismatched region and transcription start sequence are boxed in black. (**B**) The amounts of σ–DNA and ssσ–DNA complex resulting from the action of $PspF_{1–275}$ and dGTP are plotted. Only the examples of variants that fail to bind DNA, form ssσ–DNA complex, or variants that more readily form the ssσ–DNA complex are shown. I. Sigma binding to early melted DNA. II. ssσ–DNA complex formation in the presence of $PspF_{1–275}$ and dGTP.

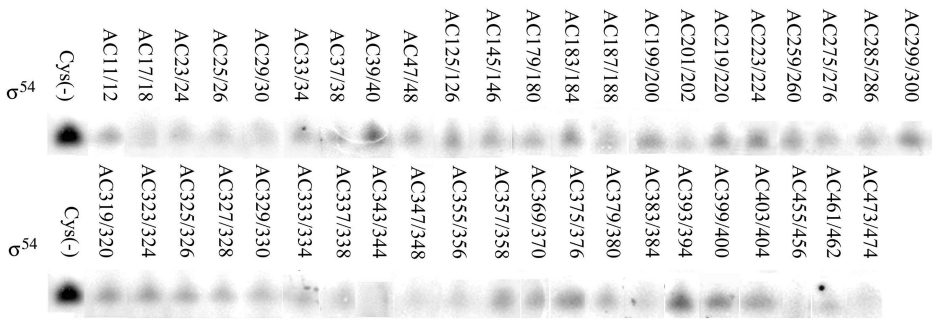

Figure 6. Mutations in σ^{54} that affect transcription initiation. Autoradiograph of the denaturing gel showing the synthesis of abortive transcripts from holoenzyme formed with σ^{54} variants which affect transcription initiation.

Variants defective in early melted DNA-binding function abnormally in isomerization assay

σ^{54} bound to early melted DNA responds to activator protein in the presence of an hydrolysable NTP source and isomerizes independently of core RNAP to form a new slower migrating 'supershifted (ss)' DNA complex (33). To characterize the sequences in σ^{54} that contribute to activator responsiveness, we attempted to form the isomerized supershifted σ–DNA complex (ssσ–DNA) using activator $PspF_{1-275}$ and hydrolysable nucleotide dGTP. For each variant, a comparison of the fraction of initially bound DNA converted, in an activator and hydrolysable nucleotide dependent way, to the new isomerized ssσ–DNA species was made. Reactions without hydrolysable nucleotide were also conducted to potentially identify any variant that can response to activator protein in a hydrolysable nucleotide independent way. However, we failed to detect any variant with this phenotype, indicating the importance of hydrolysable nucleotide to the normal function of activator protein (data not shown).

The σ^{54} isomerization assay revealed two classes of variants that behave differently from Cys(–)σ^{54}. The first class of variants is represented by some Region I variants which failed to form the isomerized ssσ–DNA species. This phenotype is also shared with some other variants that have substitutions located in core RNAP-binding domain (153–154), carboxyl terminal of modulation domain (295–296, 299–300), and DNA-binding domain (323–324, 335–336 and 455–456) (Figure 5B). However, residues that contribute to σ^{54} isomerization mainly located in two parts of Region I (11–18 and 25–36), indicating that Region I comprises the major elements influencing activator responsiveness, but likely interacts with other parts of σ^{54}. The second class of variants is represented by some Region III variants. In striking contrast to the first class, this class of variants can form isomerized ssσ–DNA species much more efficiently than Cys(–)σ^{54} (Figure 5B). As expected, early melted DNA binding seems to be important to the normal formation of ssσ–DNA, since most variants defective in early melted DNA binding either failed to form ssσ–DNA (represented by Region I variants), or formed ssσ–DNA much more efficiently (represented by Region III variants) (Figure 5B and Table 1). The latter group of variants seem to have a reduced barrier to being isomerized, implying Region III residues keep isomerization in check and that σ^{54} could function through concerted changes in the organization of its domains.

Transcription activities of the variants in carboxyl terminal of DNA-binding domain are restored *in vitro*

To further define the elements in σ^{54} that contribute to open complex formation and transcription initiation, abortive transcription assays were conducted from a supercoiled *S. meliloti nifH* promoter template (pMKC28) (28), whose sequence from −1 to +3 is TGGG (Figure 5A). If the activator $PspF_{1-275}$ and dATP (needed for the ATPase activity of the activator protein) were provided, the tetranucleotide UGGG can be synthesized by σ^{54}-holoenzyme in the presence of the priming dinucleotide UpG, and transcription substrate GTP. Heparin was also added together with GTP and after dATP plus PspF to measure transcription from a single round of activation.

The *in vitro* transcription activities of σ^{54} variants were largely in accordance with their *in vivo* activity, but in the carboxyl terminal of DNA-binding domain, many variants with low *in vivo* activity functioned normally in the abortive transcription assay (Table 1 and Figure S3). Considering the low *in vivo* expression levels of these variants detected by immunoblotting, the low *in vivo* activities of these variants could be attributed to their *in vivo* instability or reduced abilities to compete with other σ factors *in vivo*, and these defects are not operational *in vitro*. Notably for many variants that had improved activities in forming the isomerized ssσ–DNA species, the abilities of the holoenzyme they formed to initiate transcription were seldom improved, sometimes even impaired (Figure 6 and Table 1), demonstrating the influence of steps after σ^{54} isomerization to open complex formation and suggesting coordinated changes in σ^{54} and core RNAP structure are required for making activator dependent open complexes.

DISCUSSION

Cryo-EM reconstruction of σ^{54}–holoenzyme reveals four density regions (D1, Db, D2 and D3) attributable to σ^{54}. Region I sequence, including the −12 recognition region,

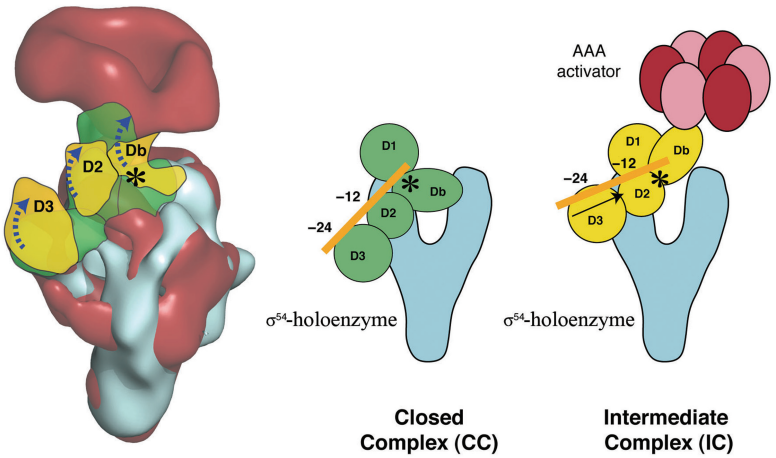

Figure 7. The left panel shows the cryo-EM reconstruction of σ^{54}-holoenzyme bound to the activator protein in the absence of promoter DNA (12). Shown in green and yellow are the positions of the σ^{54} densities before and after binding to the activator protein, respectively. The σ^{54}-holoenzyme is shown alone (light blue) and in complex with the activator $PspF_{1-275}$ (red). The middle and right panels show cartoons depicting the structural changes inferred from cryo-EM reconstruction of the σ^{54}-holoenzyme and σ^{54}-holoenzyme bound by the activator protein. The promoter DNA is shown as an orange line: In the 'CC conformation' σ^{54} domains D2 and D3 contact the −24 and −12 consensus promoter sequences upstream of the +1 site, respectively, thereby distorting DNA next to −12. The interaction with the activator protein (red/pink) in a mixed nucleotide state is thought to enable σ^{54}-holoenzyme to adopt a conformation similar to that in the intermediate complex (stabilized experimentally with ADP·AlFx).

is within Db density which is located at the position of DNA loading into the RNAP active centre, and is suggested to prevent the active centre from accessing the template DNA strand, therefore inhibiting the spontaneously formation of the open complex from double stranded DNA. Interaction of activator protein and σ^{54} likely causes a conformational change of Region I which removes the steric obstruction and modifies the interaction with −12 promoter to allow the loading of the template strand into the active centre, leading to open complex formation (Figure 7) (12). Our results fully support this model, indicating that Region I is the major element interacting with activator protein and plays an important role in coordinating activator responsiveness. Although several distinct activities reside in discrete parts of σ^{54}, residues of σ^{54} contributing to core RNAP binding or activator contacting are widely distributed across Regions I and III of σ^{54}, consistent with the observation that multiple interactions with core RNAP occur with several domains of σ^{54} and that the main connecting density between $PspF_{1-275}$ and σ^{54}–holoenzyme occurs not only at the interface of Db, but also at D1 and D2 which contain major core RNAP-binding element of σ^{54} and part of Region III of σ^{54} (12). These residues probably constitute the surfaces of the density regions through which σ^{54} exerts its influence upon the reorganization of RNAP.

Roles of Region I sequences

The amino terminal Region I plays a central role in activator responsiveness and may act as an organizing centre that brings the key σ^{54}, core RNAP and DNA component together (10,40,41). Mutations affecting σ^{54} function located between residues 15 and 47 have been described previously (18,21), and we identified one additional sequence affecting activator responsiveness of σ^{54} residues 11–14 (Figure 1). Most mutations affecting σ^{54} isomerization were located between residues 11–18 and 25–36, suggesting these two sequences are the primary sequences for activator responsiveness. Since the −12 promoter sequence was shown to have a role in contributing to supershifted ssσ–DNA complex formation (33), for some variants defective in early melted DNA binding (AC25/26, 29/30, 33/34, 47/48), the failure to form isomerized ssσ–DNA species may be attributed to the altered interaction of σ^{54} with −12 promoter sequence. Although some variants in DNA-binding domain also bound DNA poorly and less ssσ–DNA complexes were formed, the percentage of the initial bound DNA converted to ssσ–DNA is much higher compared to Cys(–)σ^{54}, suggesting that for these Region I variants, poor DNA binding *per se* does not limit the isomerization of sigma–DNA complex. Indeed tight DNA binding at −12 may limit isomerization. It has been shown before that interaction of σ^{54} and −12 promoter sequence influences the regulation of open complex formation (37,39,42). Consistent with this, our results show a very clear correlation between Region I variants that have a defect in early melted DNA binding and their poor activities in activated transcription, contrasting the result that many other variants function normally in activated transcription, even if they have the same overall defect in early melted DNA binding (Figure 1). It has been suggested that an α-helix in Region I could be required for normal σ^{54} function (21), and in our assay, it is interesting to note that mutations impairing DNA binding distribute periodically between residues 25–38 (Figures 1 and 5B).

Region I sequence binds core RNAP weakly and can be footprinted by core RNAP (16,31), and our results identified the sequences in Region I contributing to this function, associated with inhibiting the RNA polymerase from making open complexes (12). It has been demonstrated previously that residue 36 is proximal to core

RNAP (25). Consistent with this, our results show that mutations around residue 36 (residues 33–34 and 37–38) affect core RNAP binding slightly (Figure S2). The variants with reduced core RNAP-binding activity are also impaired in σ^{54} isomerization and/or activator responsiveness, suggesting that the substituted residues contribute to the interface σ^{54} makes with core RNAP through which Region I exerts its influence upon holoenzyme to regulate function.

Roles of Region II sequences

Region II is very variable in sequence. Groups of acidic residues in Region II are involved in promoter melting and deletion of Region II affects holoenzyme formation and activator responsiveness of σ^{54} (32,43). In our activator interaction assays, one Region II variant AC83/84, behaves differently from Cys(–)σ^{54} (Figure 4), suggesting that an altered interaction with activator protein may occur. The activity of this variant in isomerized ss σ–DNA formation is also improved compared to Cys (–)σ^{54}, implying the substituted residues may influence activator responsiveness (Table 1). However, we failed to identify any Region II variants significantly affecting the transcription activity of the assembled σ^{54}-holoenzyme both *in vivo* and *in vitro*.

Roles of Region III sequences

Region III includes a major determinant for core RNAP binding at its amino terminal and sequences for DNA binding located at the carboxyl terminal (13,15,16). In the cryo-EM reconstruction of σ^{54}–holoenzyme, the primary core RNAP-binding domain and carboxyl terminal DNA-binding domain are assigned to D1 and D3, respectively, while some other Region III residues located in D2 with Region I contribute to maintaining the stable closed complex (Figure 7) (28,41,44,45).

Our data show that although the primary core RNAP-binding sequence is located between residues 120 and 215, many other residues outside this region also affect core RNAP binding (Figure 3). This is consistent with the fact that sequences footprinted by core RNAP lie outside the primary core RNAP-binding domain, and are near or within the DNA-binding domain (31). On the basis of the model of the complex formed between the carboxyl terminal DNA-binding domain of σ^{54} (AAs 396–465) and promoter −24 sequence, it is proposed that the first loop of this domain may interact with the β-flap region of core RNAP (20,30). Interestingly, we find that mutations between residues 423–426, which locate in the first loop of the carboxyl terminal DNA-binding domain based on the structure of σ^{54} from *A. aeolicus* (Figure 1), indeed affect the core RNAP-binding activity (Figure 3).

Residues affecting activator interaction are also distributed widely in Region III (Figure 4). Since Region I is the major determinant for activator interaction (33), these Region III variants may influence activator interaction through the interaction between Regions III and I (28,41). Region III variants affecting activator interaction can be divided into two classes. Class one variants are somewhat defective in activator interaction, but unlike Region I variants with this same phenotype, many of these Region III variants can still respond to activator protein. Class two variants can interact with activator protein, but an additional complex with slower mobility was formed (Figure 4). This new complex has not been observed before and we suggest that this new complex might reflect a conformational change in the σ^{54}-activator complex, implying an altered interaction of σ^{54} with activator protein. Residues around 307 have been proposed to be a site for activator interaction (46). Interestingly, our data show that mutations between residues 297 and 318 either lead to the failure of σ^{54} variants to form stable complex with activator protein, or the formation of an additional complex (Table 1), both cases reflecting the altered interaction with activator protein. Interactions that activator protein mediates between Regions III and I can be used to achieve the apparently concerted domain movements of σ^{54} seen in structural reconstructions of the $PspF_{1-275}$ bound σ^{54}-holoenzyme (Figure 7) (12).

In contrast to the results of core RNAP-binding activity and σ^{54}-activator interaction, residues affecting DNA binding are mainly located in the previously delineated DNA-binding domain, residues 329–477, fully consistent with the reconstruction model of σ^{54}-holoenzyme–DNA complex where D2 and D3 densities cover −12 and −24 promoter sequences, with D1 containing the primary core RNAP-binding domain (residues 120–215). Based on the structure of the carboxyl terminal DNA-binding domain of σ^{54} from *A. aeolicus*, the residues D439, T457 and N471 of σ^{54} from *K. pneumoniae* can interact with −24 promoter sequence through hydrogen bonding (30). In accordance with this, our results show that substitutions of these three residues all affect the DNA-binding ability significantly (Table 1). Specific residues in modulation domain affecting DNA binding were also identified, confirming the view that modulation domain functions to assist DNA binding (14).

Unlike Region I variants, there is no close linkage between the defect in early melted DNA binding and ssσ–DNA formation. Among the DNA-binding defective variants in Region III, only three of them (AC295/296, 299/300 and 355/356) are defective for early melted DNA binding (Table 1). It has been suggested that residue 336 is proximal to −12 promoter sequence and interacts with sequences downstream to −12 GC element to regulate RNAP function together with Region I (28,41). Thus, like the Region I variants, the defect of variant AC335/336 in ssσ–DNA formation could also result from the altered interaction of σ^{54} and the −12 promoter sequence. In contrast to the properties of these variants, in the carboxyl terminal DNA-binding domain variants defective in DNA binding (AC439/440, AC441/442, AC457/458 etc.) are not impaired in ssσ–DNA formation, indicating that it is an interaction with −12 promoter sequence that is crucial for σ^{54} isomerization, with the rpoN box providing the −24 promoter sequence contact that is not directly involved in the isomerization process. Variants AC295/296 and AC299/300 share the phenotype of AC335/336 and some Region I variants, implying that residues 295–300 located in the carboxyl terminal of modulation

domain may also have a role in interactions with −12 promoter sequence, which may have a basis in communication with Region I.

Although most Region III variants are not impaired in ssσ–DNA formation, a distinctive class of variants is revealed on the basis of their greatly improved activities in ssσ–DNA formation. This class of variants can be roughly divided into two groups based on the linkage with other functional defects. Group I isomerization changed variants are represented by the variants in core RNAP-binding domain whose core RNAP-binding activities are impaired. Mutations leading to this phenotype are mainly located in two patches around residues 120 and 180 (Figure 1). It has been suggested that some interactions between σ^{54} and core RNAP are changed upon isomerization of the σ^{54}–DNA complex and recent structural analyses of the activator bound σ^{54}-holoenzyme are in agreement with the activator causing a reorganization of σ^{54} domains within the holoenzyme (12,47). The impaired core RNAP-binding activities of these variants described here may reflect an altered core RNAP-binding surface which is closer to the conformation of σ^{54} in the isomerized state, hence leading to the improved activity in isomerized ss σ–DNA formation. The activities of these σ^{54} variants in transcription assay are somehow impaired, suggesting that the substituted residues may contribute to the communication with core RNAP and that when they are substituted, σ^{54} isomerization is uncoupled from the changes that must occur in the core RNAP, required for the later stages of the DNA opening process and stable open complex formation (48,49).

Group II isomerization changed variants are in the DNA-binding domain. For these variants there is a clear linkage between the defect in DNA binding and an improved activity in ssσ–DNA formation (Figure 1). The weak contact these variants make with DNA may facilitate the changes in σ–DNA interactions which are important for ssσ–DNA complex formation (33,47). However, increased formation of the ssσ–DNA complex has little overall effect on the normal function of holoenzyme in the context of *in vitro* transcription assays, arguing for the existence of other steps limiting open complex formation. For most of the variants in carboxyl terminal of σ^{54} (residues 431–476), the normal functions of the holoenzyme in *in vitro* transcription assay, together with their low expression levels *in vivo*, strongly indicates the carboxyl terminal of σ^{54} plays an important role in maintaining the structural integrity perhaps related to turnover of σ^{54} *in vivo*.

In conclusion, by constructing the single cysteine mutant library of σ^{54} and analyzing the functions of σ^{54} variants both *in vivo* and *in vitro*, σ^{54} sequences interacting with core RNAP, DNA and activator protein are systematically identified. Linkages between different functions are revealed and the possible mechanisms for those linkages are also discussed. Furthermore, the characterized σ^{54} library provides a powerful resource for functional analyses using site specific probes, and for interpretations of molecular structures.

SUPPLEMETARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank D. Bose and X. Zhang for providing the model of the transcription complex.

FUNDING

Human Frontier Science Program [RGP22/2007 to Y.P.W.]; Natural Science Foundation of China [No. 30830005 to Y.P.W.]; the Program of Introducing Talents of Discipline to Universities [No. B06001 to Y.P.W.]; 973 program on nitrogen fixation to Y.P.W.; BBSRC project grant to M.B. and BBSRC David Phillips Fellowship [BB/E023703 to S.R.W.].

Conflict of interest statement. None declared.

REFERENCES

1. Gross,C.A., Lonetto,M. and Losich,R. (1992) Sigma factors. In Yamamoto,K.R. and McKnight,S.L. (eds), *Transcriptional regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 129–176.
2. Merrick,M.J. (1993) In a class of its own – the RNA polymerase sigma factor sigma 54 (sigma N). *Mol. Microbiol.*, **10**, 903–909.
3. Popham,D., Keener,J. and Kustu,S. (1991) Purification of the alternative sigma factor, sigma 54, from *Salmonella typhimurium* and characterization of sigma 54-holoenzyme. *J. Biol. Chem.*, **266**, 19510–19518.
4. Wang,J.T., Syed,A. and Gralla,J.D. (1997) Multiple pathways to bypass the enhancer requirement of sigma 54 RNA polymerase: roles for DNA and protein determinants. *Proc. Natl Acad. Sci. USA*, **94**, 9538–9543.
5. Sasse-Dwight,S. and Gralla,J.D. (1988) Probing the *Escherichia coli glnALG* upstream activation mechanism in vivo. *Proc. Natl Acad. Sci. USA*, **85**, 8934–8938.
6. Popham,D.L., Szeto,D., Keener,J. and Kustu,S. (1989) Function of a bacterial activator protein that binds to transcriptional enhancers. *Science*, **243**, 629–635.
7. Wedel,A. and Kustu,S. (1995) The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation. *Genes Dev.*, **9**, 2042–2052.
8. Wedel,A., Weiss,D.S., Popham,D., Droge,P. and Kustu,S. (1990) A bacterial enhancer functions to tether a transcriptional activator near a promoter. *Science*, **248**, 486–490.
9. Buck,M., Miller,S., Drummond,M. and Dixon,R. (1986) Upstream activator sequences are present in the promoters of nitrogen-fixation genes. *Nature*, **320**, 374–378.
10. Cannon,W., Gallegos,M.T., Casaz,P. and Buck,M. (1999) Amino-terminal sequences of sigmaN (sigma54) inhibit RNA polymerase isomerization. *Genes Dev.*, **13**, 357–370.
11. Huo,Y.X., Tian,Z.X., Rappas,M., Wen,J., Chen,Y.C., You,C.H., Zhang,X., Buck,M., Wang,Y.P. and Kolb,A. (2006) Protein-induced DNA bending clarifies the architectural organization of the sigma54-dependent glnAp2 promoter. *Mol. Microbiol.*, **59**, 168–180.
12. Bose,D., Pape,T., Burrows,P.C., Rappas,M., Wigneshweraraj,S.R., Buck,M. and Zhang,X. (2008) Organization of an activator-bound RNA polymerase holoenzyme. *Mol. Cell*, **32**, 337–346.
13. Cannon,W., Missailidis,S., Smith,C., Cottier,A., Austin,S., Moore,M. and Buck,M. (1995) Core RNA polymerase and promoter DNA interactions of purified domains of sigma N: bipartite functions. *J. Mol. Biol.*, **248**, 781–803.

14. Cannon,W.V., Chaney,M.K., Wang,X. and Buck,M. (1997) Two domains within sigmaN (sigma54) cooperate for DNA binding. *Proc. Natl Acad. Sci. USA*, **94**, 5006–5011.
15. Wong,C., Tintut,Y. and Gralla,J.D. (1994) The domain structure of sigma 54 as determined by analysis of a set of deletion mutants. *J. Mol. Biol.*, **236**, 81–90.
16. Gallegos,M.T. and Buck,M. (1999) Sequences in sigmaN determining holoenzyme formation and properties. *J. Mol. Biol.*, **288**, 539–553.
17. Sasse-Dwight,S. and Gralla,J.D. (1990) Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor sigma 54. *Cell*, **62**, 945–954.
18. Syed,A. and Gralla,J.D. (1998) Identification of an N-terminal region of sigma 54 required for enhancer responsiveness. *J. Bacteriol.*, **180**, 5619–5625.
19. Guo,Y. and Gralla,J.D. (1997) DNA-binding determinants of sigma 54 as deduced from libraries of mutations. *J. Bacteriol.*, **179**, 1239–1245.
20. Doucleff,M., Malak,L.T., Pelton,J.G. and Wemmer,D.E. (2005) The C-terminal RpoN domain of sigma54 forms an unpredicted helix-turn-helix motif similar to domains of sigma70. *J. Biol. Chem.*, **280**, 41530–41536.
21. Casaz,P., Gallegos,M.T. and Buck,M. (1999) Systematic analysis of sigma54 N-terminal sequences identifies regions involved in positive and negative regulation of transcription. *J. Mol. Biol.*, **292**, 229–239.
22. Sanderson,A., Mitchell,J.E., Minchin,S.D. and Busby,S.J. (2003) Substitutions in the *Escherichia coli* RNA polymerase sigma70 factor that affect recognition of extended -10 elements at promoters. *FEBS Lett.*, **544**, 199–205.
23. Burrows,P.C., Severinov,K., Ishihama,A., Buck,M. and Wigneshweraraj,S.R. (2003) Mapping sigma 54-RNA polymerase interactions at the -24 consensus promoter element. *J. Biol. Chem.*, **278**, 29728–29743.
24. Rappas,M., Schumacher,J., Beuron,F., Niwa,H., Bordes,P., Wigneshweraraj,S., Keetch,C.A., Robinson,C.V., Buck,M. and Zhang,X. (2005) Structural insights into the activity of enhancer-binding proteins. *Science*, **307**, 1972–1975.
25. Wigneshweraraj,S.R., Fujita,N., Ishihama,A. and Buck,M. (2000) Conservation of sigma-core RNA polymerase proximity relationships between the enhancer-independent and enhancer-dependent sigma classes. *Embo. J.*, **19**, 3038–3048.
26. Leach,R.N., Gell,C., Wigneshweraraj,S., Buck,M., Smith,A. and Stockley,P.G. (2006) Mapping ATP-dependent activation at a sigma54 promoter. *J. Biol. Chem.*, **281**, 33717–33726.
27. Cannon,W. and Buck,M. (1992) Central domain of the positive control protein NifA and its role in transcriptional activation. *J. Mol. Biol.*, **225**, 271–286.
28. Chaney,M. and Buck,M. (1999) The sigma 54 DNA-binding domain includes a determinant of enhancer responsiveness. *Mol. Microbiol.*, **33**, 1200–1209.
29. Fouts,D.E., Tyler,H.L., DeBoy,R.T., Daugherty,S., Ren,Q., Badger,J.H., Durkin,A.S., Huot,H., Shrivastava,S., Kothari,S. *et al.* (2008) Complete genome sequence of the N2-fixing broad host range endophyte *Klebsiella pneumoniae* 342 and virulence predictions verified in mice. *PLoS Genet.*, **4**, e1000141.
30. Doucleff,M., Pelton,J.G., Lee,P.S., Nixon,B.T. and Wemmer,D.E. (2007) Structural basis of DNA recognition by the alternative sigma-factor, sigma54. *J. Mol. Biol.*, **369**, 1070–1078.
31. Casaz,P. and Buck,M. (1997) Probing the assembly of transcription initiation complexes through changes in sigmaN protease sensitivity. *Proc. Natl Acad. Sci. USA*, **94**, 12145–12150.
32. Southern,E. and Merrick,M. (2000) The role of region II in the RNA polymerase sigma factor sigma(N) (sigma(54)). *Nucleic Acids Res.*, **28**, 2563–2570.
33. Cannon,W.V., Gallegos,M.T. and Buck,M. (2000) Isomerization of a binary sigma-promoter DNA complex by transcription activators. *Nat. Struct. Biol.*, **7**, 594–601.
34. Chaney,M., Grande,R., Wigneshweraraj,S.R., Cannon,W., Casaz,P., Gallegos,M.T., Schumacher,J., Jones,S., Elderkin,S., Dago,A.E. *et al.* (2001) Binding of transcriptional activators to sigma 54 in the presence of the transition state analog ADP-aluminum fluoride: insights into activator mechanochemical action. *Genes Dev.*, **15**, 2282–2294.
35. Bordes,P., Wigneshweraraj,S.R., Schumacher,J., Zhang,X., Chaney,M. and Buck,M. (2003) The ATP hydrolyzing transcription activator phage shock protein F of *Escherichia coli*: identifying a surface that binds sigma 54. *Proc. Natl Acad. Sci. USA*, **100**, 2278–2283.
36. Buck,M. and Cannon,W. (1992) Specific binding of the transcription factor sigma-54 to promoter DNA. *Nature*, **358**, 422–424.
37. Guo,Y., Wang,L. and Gralla,J.D. (1999) A fork junction DNA-protein switch that controls promoter melting by the bacterial enhancer-dependent sigma factor. *EMBO J.*, **18**, 3736–3745.
38. Wang,L. and Gralla,J.D. (1998) Multiple in vivo roles for the -12-region elements of sigma 54 promoters. *J. Bacteriol.*, **180**, 5626–5631.
39. Gallegos,M.T. and Buck,M. (2000) Sequences in sigma(54) region I required for binding to early melted DNA and their involvement in sigma-DNA isomerisation. *J. Mol. Biol.*, **297**, 849–859.
40. Wang,J.T., Syed,A., Hsieh,M. and Gralla,J.D. (1995) Converting *Escherichia coli* RNA polymerase into an enhancer-responsive enzyme: role of an NH2-terminal leucine patch in sigma 54. *Science*, **270**, 992–994.
41. Wigneshweraraj,S.R., Chaney,M.K., Ishihama,A. and Buck,M. (2001) Regulatory sequences in sigma 54 localise near the start of DNA melting. *J. Mol. Biol.*, **306**, 681–701.
42. Wang,L., Guo,Y. and Gralla,J.D. (1999) Regulation of sigma 54-dependent transcription by core promoter sequences: role of −12 region nucleotides. *J. Bacteriol.*, **181**, 7558–7565.
43. Wong,C. and Gralla,J.D. (1992) A role for the acidic trimer repeat region of transcription factor sigma 54 in setting the rate and temperature dependence of promoter melting in vivo. *J. Biol. Chem.*, **267**, 24762–24768.
44. Chaney,M., Pitt,M. and Buck,M. (2000) Sequences within the DNA cross-linking patch of sigma 54 involved in promoter recognition, sigma isomerization, and open complex formation. *J. Biol. Chem.*, **275**, 22104–22113.
45. Wang,L. and Gralla,J.D. (2001) Roles for the C-terminal region of sigma 54 in transcriptional silencing and DNA binding. *J. Biol. Chem.*, **276**, 8979–8986.
46. Lee,J.H. and Hoover,T.R. (1995) Protein crosslinking studies suggest that *Rhizobium meliloti* C4-dicarboxylic acid transport protein D, a sigma 54-dependent transcriptional activator, interacts with sigma 54 and the beta subunit of RNA polymerase. *Proc. Natl Acad. Sci. USA*, **92**, 9702–9706.
47. Cannon,W., Gallegos,M.T. and Buck,M. (2001) DNA melting within a binary sigma(54)-promoter DNA complex. *J. Biol. Chem.*, **276**, 386–394.
48. Wigneshweraraj,S.R., Burrows,P.C., Severinov,K. and Buck,M. (2005) Stable DNA opening within open promoter complexes is mediated by the RNA polymerase beta'-jaw domain. *J. Biol. Chem.*, **280**, 36176–36184.
49. Wigneshweraraj,S.R., Savalia,D., Severinov,K. and Buck,M. (2006) Interplay between the beta' clamp and the beta' jaw domains during DNA opening by the bacterial RNA polymerase at sigma54-dependent promoters. *J. Mol. Biol.*, **359**, 1182–1195.