Different Roles for Basic and Aromatic Amino Acids in Conserved Region 2 of *Escherichia coli* σ^{70} in the Nucleation and Maintenance of the Single-stranded DNA Bubble in Open RNA Polymerase-Promoter Complexes*

Received for publication, June 1, 2001
Published, JBC Papers in Press, July 6, 2001, DOI 10.1074/jbc.M105027200

Mark Tomsič, Laura Tsuchikawa‡, Gianina Panaghie, Yang Wang, Joseph Azok, and Pieter L. deHaseth§

From the Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106-4935

Amino acid residues in region 2 of σ^{70} have been shown to play an important role in the strand separation step that is necessary for formation of the functional or open RNA polymerase-promoter complex. Here we present a comparison of the roles of basic and aromatic amino acids in the accomplishment of this process, using RNA polymerase bearing alanine substitutions for both types of amino acids in region 2. We determined the effects of the substitutions on the kinetics of open complex formation, as well as on the ability of the RNA polymerase to form complexes with single-stranded DNA, and with forked DNA duplexes carrying a single-stranded overhang consisting of bases in the −10 region. We concluded that two basic amino acids (Lys^{414} and Lys^{418}) are important for promoter binding and demonstrated distinct roles, at a subsequent step, for two aromatic amino acids (Tyr^{430} and Trp^{433}). It is likely that these four amino acids, which are close to each other in the structure of σ^{70}, together are involved in the nucleation of the strand separation process.

RNA synthesis in prokaryotes is carried out by a multi-subunit RNA polymerase commonly referred to as the core enzyme (E). For promoter recognition, a sigma (initiation) factor is required; it interacts with the core polymerase to yield the holoenzyme (E\(\sigma\)), which is able to form an initiation-competent complex at promoter sequences in a multistep process involving conformational changes in both the protein and the DNA (1–3). A striking feature of such a complex is a region of strand separation that spans about 14 base pairs from the upstream edge of the conserved −10 promoter element to just beyond the start site of transcription initiation (4). It is thought that, kinetically, strand separation initiates in the −10 region and proceeds in a downstream direction. Measurement of the size and location of the region of strand separation as a function of temperature shows that at low temperatures a small single-stranded region can be detected that, as the temperature is increased, expands toward the start site (3, 5–7). In addition, the introduction of nicks and mismatches in the −10 region is more effective in the acceleration of open complex formation than if such distortions are introduced at a more downstream position (8, 9).

The predominant sigma factor in *Escherichia coli*, which enables recognition of promoters of housekeeping genes, is referred to as σ^{70}. Sequence comparison has shown that a large group of sigma factors shows significant homology to σ^{70}. Four regions of sequence conservation have been identified, of which some have been subdivided to reflect the most extensive sequence conservation (10). A large body of data has implicated region 2.3 of the main sigma factors of *E. coli*, *Bacillus subtilis*, and other prokaryotes in the nucleation of the strand separation. This process eventually results in the formation of the active or open complex, possibly by facilitating base flipping of the highly conserved A at −11 of the nontemplate strand. The supporting experimental evidence has been derived from analysis of the effects of alanine substitution for aromatic amino acids on open complex formation (5, 11) and on the ability of RNA polymerase to interact with model substrates such as single-stranded DNA (12, 13) and duplexes carrying regions of unpaired DNA (14), also referred to as forked templates (15). The former are thought to model the unpaired regions of the strand-separated bubble, the latter the junction between double- and single-stranded DNA of the bubble. Based on the use of forked templates, Gralla and co-workers (14) have concluded that Tyr^{430} and Trp^{433} of region 2.3, which jut out of the body of the protein (16), are particularly important for the initiation of the strand separation process.

In this paper we demonstrate that alanine substitutions at the positions of Tyr^{430} and Trp^{433} lead to different effects on the interaction of *E\(\sigma\)^{70} with forked DNA. We provide new results in support of the idea that multiple aromatic amino acids jointly interact with single-stranded DNA downstream of the region where strand separation is initiated. Our results allow us to single out two basic amino acid residues in region 2 (lysines 414 and 418) as being particularly important for the interaction of RNA polymerase with DNA. Finally, we demonstrate that the substitutions of alanines for basic amino acids in region 2 have effects that are fundamentally different from those of substitutions for aromatic amino acids in the same region.

**EXPERIMENTAL PROCEDURES**

**Materials**

Oligonucleotides (oligos) were synthesized by Life Technologies, Inc. or Genset. Nonradioactive NTPs and dNTPs were purchased from...
Different Roles for Amino Acids in Region 2 of E. coli σ^70

Roche Molecular Biochemicals. Cpa was from Sigma. [γ-^32P]ATP, [γ-^32P]ATP, and [α-^32P]UTP were from PerkinElmer Life Sciences. DNA-modifying enzymes were purchased from either New England Biolabs or Roche Molecular Biochemicals. Er^70 core enzyme was prepared as described (17) or purchased from Epicentre.

Methods

Deoxyoligonucleotide Labeling and Annealing—5'-end-labeled DNA oligonucleotides were generated by incubation with [γ-^32P]ATP and T4 polynucleotide kinase (New England Biolabs) using established procedures. The two strands of forked DNA templates were re-annealed at concentrations of 100 nM (5'-end-labeled non-template strand) and 150 nM (unlabeled template strand) in a buffer containing 25 mM Tris (pH 7.9 at 25°C) and 50 mM NaCl by heating to 90°C and slow cooling.

σ^70 Mutagenesis and Purification—All manipulations of the σ^70 codon region were performed on the plHIN12-His expression plasmid, an isopropyl-1-thio-galactopyranoside-inducible version of Pet11a vector from Novagen exactly as previously described (11). Site-directed single mutations were introduced using the QuickChange site-directed mutagenesis kit (Stratagene) and the appropriate primers according to the manufacturer’s instructions. The mutagenized fragments were subcloned into the wild type plHIN12-His expression vector using the PacI and BamHI restriction enzymes, and the recombinant plasmids were reconstituted in storage buffer (10 mM Tris-HCl (pH 7.9), 100 mM KCl, 1 mM dithiothreitol, 50% glycerol) from purified E. coli core enzyme (Epicsentech) and purified σ^70 in a molar ratio of 1.5 at 4°C for 1 h. Proper interaction with core was verified by a gel mobility shift assay (11, 18).

Electroelphoretic Mobility Shift Assays (EMSA)—In experiments to determine the k_{off} for heparin resistant (open) complex formation, we used a template in which the AP_{70} promoter was inactivated by two mutations in its −35 region, whereas the AP_{70} promoter had an “up” mutation in its −10 region (TAGATT to TAGAAT). The experiments were carried out as described previously (11). Briefly, about 2–5 nM [γ-^32P]ATP was incubated with 200 nM Er^70 at 37 or 20°C in 20 mM HEPES buffer (pH 7.5, 100 mM KCl, 1 mM dithiothreitol), containing 50 μg/ml bovine serum albumin, for various amounts of time. Each reaction was challenged with heparin to 100 μg/ml for 1 min prior to the addition of the loading solution and loading onto a 5% nondenaturating gel (29.1 acrylamide/bisacrylamide) run at room temperature. Frozen gels were autoradiographed and quantified by PhosphorImager (Molecular Dynamics). The radioactivity in each band, as the percentage of the total in the lane, was plotted versus the time of incubation with the Er^70. The k_{off} for the wild type and mutant RNA polymerases were determined by fitting the data to the equation y = Yf \cdot (1 - e^{-t \cdot k_{off}}) + Yo, where y is the percent of open complexes formed, t = time of DNA/Er^70 incubation, Yf and Yo are the limiting values for y, and k_{off} is the pseudo-first order rate constant.

Interaction of Er^70 with Single-stranded and Forked DNAs—All reactions were performed in HEPES buffer (see above). Binding to single-stranded (ss) DNA was studied by incubating 5'-β-^32P-labeled DNA oligo (10 nM) and Er^70 (65 nM) for 30 min at 25°C followed by the addition of loading solution and loading onto a 5% nondenaturating gel, which was run at 4°C (13). The intensity of bands corresponding to free and Er^70-bound oligo was determined by PhosphorImager (Molecular Dynamics). Binding to forked templates was determined similarly, except that the reactions were subjected to a 10-min challenge with 100 μg/ml heparin prior to loading onto the gel at 25°C (15). To determine the stability of Er^70 forked DNA complexes, 60-μl solutions were prepared containing 10 nM [γ-^32P]ATP labeled forked DNA template and 65 nM Er^70 in HEPES buffer. After a 30-min incubation at 25°C, heparin was added to 100 μg/ml, and at regular time intervals 4.5-ml aliquots were removed for analysis on a 5% nondenaturing gel as described above.

DNase I Footprinting—DNase I footprinting experiments were carried out as described (11) on DNA fragments with wt sequence, for the observed footprint is at the P_{E}' promoter. The DNA substrates were made by polymerase chain reaction using 5'-β-^32P-labeled primers and purified on a 6% native polyacrylamide gel. RNA polymerase was incubated with labeled promoter DNA in 20 μl of HEPES buffer at 37°C for 30 min. At the conclusion of the incubation period, heparin was added to a final concentration of 100 μg/ml. After the MgCl2 concentration was adjusted to 10 mM the DNA was cut with 0.4 units of DNase I (Ambion) for 30 s at 37°C. The reactions were terminated and analyzed on denaturing gels as described previously (11).

Abortive Initiation—Abortive initiation assays were done essentially as described (19). For analysis of open complex formation at the AA promoter, the unlabeled DNA fragment (2–5 nM) obtained by polymerase chain reaction) and Er^70 (200 nM) were incubated in transcription buffer at 37°C for 30 min. After the incubation period heparin was added to a final concentration of 100 μg/ml, and reactions were performed for an additional 5 min. To initiate transcription, Cpa and [γ-^32P]UTP were added (providing final concentrations of 0.5 μM Cpa, 50 μM UTP, and 1 μM of [α-^32P]UTP). The reactions were incubated for 15 min at 37°C after which 5 μl of transcription stop solution (7 nM urea, 0.1 mM EDTA, 0.4% (v/v) SDS, 40 mM Tris-HCl (pH 8.0), 0.2 mM spermidine, and 0.5% xylene cyanol FF) was added. The CpaU was separated from the substrates on a 20% denaturing polyacrylamide gel. Gels were frozen and exposed to film for 2–3 h. The intensities of the bands corresponding to UTP and CpaU were determined by densitometry.

RESULTS

The σ^70 mutant set used in this study is shown in Fig. 1a. It consists of alanine substitutions for basic and aromatic amino acids and one nonpolar residue (1435A; also 1435L). All basic amino acids in region 2.3 as well as two residues in regions 2.2 and 1.2 are included. The aromatic amino acids (described in Ref. 11) are all in region 2.3; only the buried residue Phe^419 and residue Tyr^421 from this region were not substituted. Studies from Gralla and co-workers (14) have made use of random overlapping set of substitutions extending from Tyr^424 in the C-terminal direction through Ile^152. The results presented here extend those of the latter two studies.

With the exception of σ^70 with alanine or leucine substitutions for Ile^152, all were able to bind core as determined by a gel mobility shift assay (data not shown), which was carried out exactly as described (11). The Ile^152 residue is highly conserved, probably for structural reasons, as it is buried in the σ^70 structure. Even a conservative substitution of this residue with

FIG. 1. Sequences of the σ^70 variants and the DNA fragments used in this study. a, the sites of single alanine substitutions in σ^70 are indicated. b, ssDNA oligos. Both oligos have the sequence of the non-template strand of the P_{E}' promoter, except in the −10 sequences (underlined), whereas the σ^70 core enzyme has the TATAAT sequence, whereas the −11C oligo is CATAAT. c, forked templates. The sequences are based on that of the P_{E}' promoter of bacteriophage λ. The positions of the −10 and −35 elements are indicated by boxes. The short fork has just the overlapping A at position −11; the long fork has an entire consensus −10 sequence in the overhang.
leucine apparently leads to major structural defects that interfere with $\sigma^{70}$ core interactions. DNase I footprinting of complexes of the $E_{70}$ holoenzyme, reconstituted with wt $\sigma^{70}$ or $\sigma^{70}$ containing substitutions in basic amino acids of region 2 (see Fig. 1), and the $\Phi_R$ promoter is shown in Fig. 2. The $E_{70}$ and the promoter were incubated for 30 min prior to a heparin challenge for 1 min and exposure to DNase I for 30 s. With the exception of $E_{70}$ reconstituted with K414A $\sigma^{70}$, all holoenzymes afforded complete protection over a region of DNA between $-50$ and $+20$. At $37^\circ C$, for the $\Phi_R$ promoter, such an extended footprint is characteristic of the open RNA polymerase-promoter complex. Holoenzyme containing the Lys414 $\sigma^{70}$ afforded only partial protection, consistent with the observation that this mutant $E_{70}$ forms open complexes with only about 25% of the DNA even after long incubation times, as compared with over 60% for the others (data not shown). No evidence was obtained for a heparin-sensitive complex with a “short” footprint, characteristic of a closed complex, with any of the $E_{70}$ tested bearing alanine substitutions for basic amino acids.

In Fig. 3, the rate of open complex formation at $P_{60} -10up$ for $E_{70}$ containing substitutions in basic amino acids is compared with the rates previously reported for $E_{70}$ bearing substitutions in aromatic amino acids (data from Ref. 11; note that YW bears alanine substitutions at positions 430 and 433; YYW, at 425, 430, and 433; FYW at 427, 430, and 433; FYYW at 427, 430, 433, and 434. Just as with the aromatic amino acids (5, 20, 21), the $E_{70}$ bearing Ala substitutions at basic amino acid residues Arg422, Arg423, Lys426, and Arg436 confer cold sensitivity compared with $E_{70}$ reconstituted with wt $\sigma^{70}$. However, at $37^\circ C$ these same four substitutions have relatively minor effects on $E_{70}$ function. In contrast, substitutions at Lys414 and Lys418 lead, even at $37^\circ C$, to very slow formation of open complexes. The W433A substitution has an equally large effect. Only the $E_{70}$ reconstituted with triply or quadruply substituted $\sigma^{70}$ had lower rates of formation of open complexes (this pattern is confirmed by abortive initiation data shown in Fig. 4; see below). Because of the low precision in the data for K414A, we were unable to assess whether the temperature dependence data obtained for this substitution were reliable.

Analogous results were obtained in experiments in which the formation of $E_{70}$-promoter complexes, competent for formation of the first phosphodiester bond, was monitored by determination of the synthesis of the trinucleotide CpApU from CpA and UTP. In these experiments, $E_{70}$ and promoter DNA are incubated for 30 min prior to a 1-min heparin challenge and addition of the RNA synthesis substrates. The amount of trinucleotide synthesis reflects the formation, during the 30-min incubation period, of complexes that were competent to initiate RNA synthesis. The ranking of the mutant $E_{70}$ in this experiment roughly mirrors that established by their relative $k_{obs}$ of
open complex formation, with K414A and K418A each being about as detrimental as multiple substitutions for aromatic amino acid residues. However, in contrast to the data shown in Fig. 3, here the K418A substitution is significantly less damaging than K414A. The lack of observed effects of the R422A and R436A substitutions in this experiment stands in contrast to their effect on the rate of open complex formation (Fig. 3). This likely reflects the relatively long preincubation between $E\sigma^{70}$ and the DNA in the abortive initiation experiments, which permits detection of only severe effects.

To better pinpoint the nature of the defects in the substituted sigma factors that are responsible for their impaired function, we employed the model templates shown in Fig. 1, b and c. To assess the extent to which the substitutions might impede the ability of $E\sigma^{70}$ to interact with the single-stranded DNA, and thereby render it less competent in propagating strand separation during open complex formation, we determined the ability of the mutant $E\sigma^{70}$ to interact with ssDNA. It had been shown previously that $E\sigma^{70}$ holoenzyme, but not free $\sigma^{70}$ or core enzyme, is able to sequence-specifically interact with single-stranded DNA spanning the −10 promoter element and having the sequence of the non-template strand of promoter DNA (12, 13, 22). Two oligos were employed, bearing either the consensus TATAAT −10 sequence or a non-consensus variant, CATAAT (Fig. 1b). The results are shown in Fig. 5. $E\sigma^{70}$ variants containing the multiply substitutions. Thus it is not possible to conclude that the multiple substitutions do not affect the binding of the short fork and to infer from these data alone that, for the long fork, these substitutions mostly affect the interaction with the ssDNA tail downstream of −11A. However the data for the interaction of $E\sigma^{70}$ with single-stranded DNA (Fig. 5), where the binding affinity decreased in a similar order, $YW > YYW > FYW > FYWW$, indicating that substitutions at positions 425–434 affect the interaction of the long fork with $E\sigma^{70}$. The binding of the short fork may have reached the other extreme (background) with the wt $\sigma^{70}$. All $E\sigma^{70}$ variants bind the long fork more tightly than the short fork, in most cases so tightly that the ability to discriminate differences in binding affinities is likely to be outside the useful window of this experiment. However, with the long fork, the multiply substituted $\sigma^{70}$ variants could be differentiated; the extent of binding decreases in the order $YW > YYW > FYW > FYWW$, indicating that substitutions at positions 425–434 affect the interaction of the long fork with $E\sigma^{70}$. The binding of the short fork could not be concluded that the multiple substitutions do not affect the binding of the short fork and to infer from these data alone that, for the long fork, these substitutions mostly affect the interaction with the ssDNA tail downstream of −11A. However the data for the interaction of $E\sigma^{70}$ with single-stranded DNA (Fig. 5), where the binding affinity decreased in a similar order, $YW > YYW > FYW > FYWW$, would seem to lend support to such a conclusion.

We also determined the stabilities of the heparin-resistant complexes formed with $E\sigma^{70}$ containing the wt and mutant sigma factors. We show elsewhere$^2$ that formation of a stable complex between $E\sigma^{70}$ and the short fork DNA proceeds kinetically through a heparin-sensitive intermediate. At equilibrium, a substantial fraction of the heparin-sensitive complexes persist; upon addition of heparin these complexes dissociate with a rate that is too fast to measure by the manual mixing methods employed here.$^2$ Our experiments address the stability of the fraction of the complexes that dissociate with a slower rate. All substitutions that were assayed in the course of this study affect the stability of these complexes (see Fig. 7); the position −12. It serves to define a potential path of the ssDNA on $\sigma^{70}$ in an open complex.

Guo and Grailla (15) have established “forked” DNAs, modeling the junction between double- and single-stranded regions, as particularly useful model templates, able to form heparin-resistant complexes with $E\sigma^{70}$. We have employed here the “short fork” containing a minimal ss region consisting of just the A at −11, as well as the “long fork,” where the ss extension covers the entire −10 sequence (see Fig. 1c). The results are shown in Fig. 6. $E\sigma^{70}$ containing $\sigma^{70}$ with the Y430A and K418A single substitutions, as well as those containing multiple substitutions including Y430A, are particularly deficient in their ability to bind the short fork. On the other hand, the Y425A, R422A, and K426A substitutions, which drastically affect the ability of $E\sigma^{70}$ to bind ssDNA, have very small effects, if at all, as compared with the wt $\sigma^{70}$. All $E\sigma^{70}$ variants bind the long fork more tightly than the short fork, in most cases so tightly that the ability to discriminate differences in binding affinities is likely to be outside the useful window of this experiment. However, with the long fork, the multiply substituted $\sigma^{70}$ variants could be differentiated; the extent of binding decreases in the order $YW > YYW > FYW > FYWW$, indicating that substitutions at positions 425–434 affect the interaction of the long fork with $E\sigma^{70}$. The binding of the short fork may have reached the other extreme (background) with the $E\sigma^{70}$ containing the multiple substitutions. Thus it is not possible to conclude that the multiple substitutions do not affect the binding of the short fork and to infer from these data alone that, for the long fork, these substitutions mostly affect the interaction with the ssDNA tail downstream of −11A. However the data for the interaction of $E\sigma^{70}$ with single-stranded DNA (Fig. 5), where the binding affinity decreased in a similar order, $YW > YYW > FYW > FYWW$, would seem to lend support to such a conclusion.

We also determined the stabilities of the heparin-resistant complexes formed with $E\sigma^{70}$ containing the wt and mutant sigma factors. We show elsewhere$^2$ that formation of a stable complex between $E\sigma^{70}$ and the short fork DNA proceeds kinetically through a heparin-sensitive intermediate. At equilibrium, a substantial fraction of the heparin-sensitive complexes persist; upon addition of heparin these complexes dissociate with a rate that is too fast to measure by the manual mixing methods employed here.$^2$ Our experiments address the stability of the fraction of the complexes that dissociate with a slower rate. All substitutions that were assayed in the course of this study affect the stability of these complexes (see Fig. 7); the

$^2$ L. Tsujikawa, O. Tsendikov, and P. L. deHaseth, manuscript in preparation.
forked DNA. Eσ70 was incubated with the forked DNAs as indicated (“Materials and Methods”), then the complexes were challenged with heparin for 1 min, after which the extent of complex formation was determined by EMSA. The sequences of the short (dark bars) and long (light bars) forks are shown in Fig. 1c.

**DISCUSSION**

The experiments presented here were carried out to study nucleation of strand separation as orchestrated by region 2 of *E. coli* σ70. From the findings we conclude that both basic and aromatic residues are important for open complex formation: All basic residues examined, with the possible exception of ArgC22, play a role in open complex formation, as is most evident at 20 °C, but substitutions for LysA14 and LysA18 have the most prominent effects. Substitutions at LysA14, ArgC23, LysC28, and ArgC36 affect ssDNA binding and at LysA14 and LysA18 formation of a heparin-resistant complex with forked DNA. We previously established the involvement of TyrB25, PheD27, TyrB30, and TrpB33 in open complex formation. Here we show that TyrB25, PheD27, and TyrB30 are important for interaction with ssDNA, PheD27, TyrB30, and TrpB33 for formation of a heparin-resistant complex with forked DNA and TyrB30 for its stabilization. Because PheD27 is buried (16), the effects of its substitution could be indirect. TrpB33 may be involved in forked DNA binding (Ref. 14; see also Fig. 6) but we have insufficient data for this residue. Together with the data from Gralla’s group (14), we conclude that residues of σ70 from 414 to 452, thus including helix 13 (16), the connecting loop, and helix 14, are involved in formation of an open complex.

We attempted to correlate the deficiency in either ssDNA binding or formation of a heparin-resistant Eσ70-forked complex, with the kobs of open complex formation. No significant correlation was observed between the fraction of ssDNA bound (Fig. 5) and kobs (Fig. 3) (R2 = 0.03). However, in searching for a correlation between kobs for the various mutant Eσ70 and their ability to form a heparin-resistant complex with the short fork, a peculiar but striking difference in the behavior of Eσ70-bearing alanine substitutions for basic and aromatic amino acids in region 2 came to light, as shown in Fig. 8. If the entire data set is considered, a reasonable (R2 = 0.7) correlation is observed between the kobs and the % forked DNA bound in a heparin-resistant complex. However, the correlation was much better (R2 = 0.94) when only substitutions of alanine for basic residues were considered; the corresponding linear least squares fit is displayed in Fig. 8. The aromatic residues are clearly off the line. The points for Y433A, F427A, and Y425A (e, f, and g, respectively) lie above the plot (i.e. compared with the substitutions for basic amino acids, they bind better to the forked DNA than expected based on their value of kobs for open complex formation). The points for the multiply substituted sigma factors (a, b, and c) as well as for Y430A (d) show the opposite behavior. There was no a priori reason for expecting a correlation between the kinetic data for open complex formation at promoters and the equilibrium data for the extent of short fork binding. However, the fact that one is observed indicates that the differences in dissociation rates (or half-lives) we observe for the forked DNAs (see Fig. 7) might be small compared with the differences that exist in the association rate constants. The fork binding data would essentially reflect relative rates of formation of complexes.

Some time ago, Record and colleagues (2) as well as Buc and McClure (1) showed that formation of an open complex proceeds through at least two intermediates.
Here R represents the RNA polymerase holoenzyme and P the promoter. R-P₁ is the first sequence-specific complex formed, where the strands have not yet separated; it is often referred to as the closed complex. R-P₂ is the second intermediate, where $E\sigma^{70}$ may have undergone a conformational change and the nucleation of strand separation may have taken place, perhaps by the flipping of the $-11A$ out of the helix (14). We interpret our results in terms of the basic and aromatic residues being involved in the different steps of Scheme 1. We have shown previously that the aromatic amino acids of $\sigma^{70}$ region 2.3 were probably involved in the second step of Scheme 1, the interconversion between R-P₁ and R-P₂ (11). In agreement with this interpretation, the results obtained here show that $E\sigma^{70}$ bearing multiple substitutions for aromatic amino acids, is deficient in binding ss and short fork DNA as well as in stabilizing $E\sigma^{70}$-short fork DNA complexes. Based on the results presented here, we conclude that the basic amino acid residues would facilitate the formation of R-P₁. This conclusion is supported by the comparison of affinities with which $E\sigma^{70}$ bearing alanine substitutions for these residues bound short fork DNA.² It also is consistent with the failure to observe a heparin-sensitive short footprint for these mutant $E\sigma^{70}$ (data not shown) as was observed for the YYW and FYW $E\sigma^{70}$ (11).

Residues Lys⁴¹⁴, Lys⁴¹⁸, Tyr⁴³⁰, and Trp⁴³³ appear particularly important for progressing through the individual steps shown in Scheme 1. Our results indicate that Lys⁴¹⁴, Lys⁴¹⁸, and Tyr⁴³⁰ are vital in the formation of heparin-resistant $E\sigma^{70}$. forked DNA complexes; in addition, Trp⁴³³ plays a crucial role in the stabilization of these complexes. The effects of alanine substitutions for Lys⁴¹⁴, Lys⁴¹⁸, and Trp⁴³³ on open complex formation are evident at both 20 and 37 °C, whereas the effect of substitution for Tyr⁴³⁰ apparently can be masked partially at 37 °C but not at 20 °C. These four amino acid residues are close to each other in $\sigma^{70}$; they are located within a sphere with a radius of about 5 Å in the structure of $\sigma^{70}$ (16). We propose that, together, they participate in the nucleation of strand separation. The roles of the two basic amino acids would be to hold the promoter DNA in the proper orientation and allow the aromatic amino acids to nucleate the strand separation process, likely by flipping the $-11A$ out of the helix by a mechanism that is not yet fully understood. It has been proposed that the aromatic rings of residues 430 and 433 would sandwich the $-11A$ in between them (14). This model is unlikely to be entirely correct, as our findings clearly indicate that alanine substitutions for the two residues do not behave similarly in all assays. The substitution for Tyr⁴³⁰ has a much more pronounced effect on forked DNA binding to $E\sigma^{70}$, whereas substitution for Trp⁴³³ has a greater effect on $k_{obs}$ for open complex formation at 37 °C. Also, it is evident from the results presented in Fig. 8 that Y430A behaves differently from the other single substitutions for aromatic amino acids we examined. One possibility for reconciling the findings is that Trp⁴³³ participates in “forcing” the flipped base out of a DNA duplex, whereas Tyr⁴³⁰ would interact with the flipped out base as provided either by the forked DNA or subsequent to the action of Trp⁴³³ on duplex DNA.

It is also unlikely that three electron-rich rings, such as those of Tyr⁴³⁰, −11A, and Trp⁴³³ would engage in the formation of a sandwich. Such a sandwich has been proposed for the human 3-methyladenine DNA glycosylase (25), but the DNA base in that case is alkylated and electron-deficient. It is more likely that the putatively flipped out $-11A$ would partially overlap one of the aromatic residues, perhaps the tyrosine at 430, so that electron-deficient regions of one would be over the electron cloud of the other. This would also be more consistent with the structural and mutagenesis data that have been reported for flipped out bases, which are found in close proximity to just one aromatic amino acid residue. This is a tryptophan in the case of the E. coli repair enzyme AlkA (26) and the Tn5 transposon (27), a tyrosine for human 3-methyladenine DNA glycosylase and B. subtilis DNA polymerase I (25, 28), and a phenylalanine for N6-adenine DNA methyltransferase (29) and E. coli and human uracil DNA glycosylase (30, 31). Another aromatic residue could be involved in the above mentioned forcing out of the flipped base or could be interacting edge-on with it, like tyrosines 162 and 159, respectively, of human 3-methyladenine DNA glycosylase (25) and here possibly Trp⁴³³ and/or Trp⁴³⁰.

REFERENCES

Different Roles for Basic and Aromatic Amino Acids in Conserved Region 2 of \textit{Escherichia coli} in the Nucleation and Maintenance of the Single-stranded DNA Bubble in Open RNA Polymerase-Promoter Complexes

Mark Tomsic, Laura Tsujikawa, Gianina Panaghie, Yang Wang, Joseph Azok and Pieter L. deHaseth

doi: 10.1074/jbc.M105027200 originally published online July 6, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M105027200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 31 references, 10 of which can be accessed free at http://www.jbc.org/content/276/34/31891.full.html#ref-list-1