Discrimination against Deoxyribonucleotide Substrates by Bacterial RNA Polymerase

Received for publication, July 6, 2004, and in revised form, July 15, 2004
Published, JBC Papers in Press, July 15, 2004
DOI 10.1074/jbc.C400316200

Vladimir Svetlov§, Dmitry V. Vassylyev§, and Irina Artisimovitch¶
From the §Department of Microbiology, The Ohio State University, Columbus, Ohio 43210 and the ¶Structural and Molecular Biology Laboratory and §Structurome Research Group, RIKEN Harima Institute at SPring-8, Hyogo 679-5148, Japan

Nucleic acid polymerases have evolved elaborate mechanisms that prevent incorporation of the non-cognate substrates, which are distinguished by both the base and the sugar moieties. While the mechanisms of substrate selection have been studied in single-subunit DNA and RNA polymerases (DNAPs and RNAPs, respectively), the determinants of substrate binding in the multisubunit RNAPs are not yet known. Molecular modeling of Thermus thermophilus RNAP-substrate NTP complex identified a conserved β′ subunit Asn724 residue in the active site that could play an essential role in selection of the substrate ribose. We utilized the Escherichia coli RNAP model system to assess this prediction. Functional in vitro analysis demonstrates that the substitutions of the corresponding β′ Asn458 residue lead to the loss of discrimination between ribo- and deoxyribonucleotide substrates as well as to defects in RNA chain extension. Thus, in contrast to the mechanism utilized by the single-subunit T7 RNAP where substrate selection commences in the inactive pre-insertion site prior to its delivery to the catalytic center, the bacterial RNAPs likely recognize the sugar moiety in the active (insertion) site.

In Vitro Transcription Reactions—The wild-type and altered RNAPs were expressed in pIA299 background and purified as described previously. All templates for transcription reactions were generated by PCR amplification. To form halted transcription elongation complexes (TECs), linear DNA template (40 nM), RNAP (50 nM), ApU (100 μM), and starting NTPs (2.5 μM ATP and GTP, 1 μM CTP, 10 μCi of [α-32P]CTP (3000 Ci/mmol)) were mixed on ice in 50 μl of TGA buffer (20 mM Tris acetate, 20 mM sodium acetate, 10 mM magnesium acetate, 5% glycerol, 2 mM 2-mercaptoethanol, 0.1 mM EDTA, pH 8.0). Reactions were incubated at 37 °C for 15 min and purified from nucleotides by gel filtration through G50 spin columns (GE Health) equilibrated in TGA buffer. Transcription was reinitiated after the addition of substrates indicated in figure legends, and samples were removed at selected times and quenched by the addition of an equal volume of STOP buffer (10 x urea, 20 mM EDTA, 45 mM Tris borate, pH 8.3).
RESULTS

Experimental Set-up—To determine whether β' Asn458 residue is critical for the interaction with the 2'-OH, we changed Asn458 to Asp and Ser; these are the conservative changes that would not be expected to substantially alter the structure of the protein but would change either the chemical properties (Asp) or the size (Ser) of the discriminating Asn side chain. The altered enzymes were overexpressed from a polycistronic vector that allows assembly of the core αβ' RNAP in vivo (19), purified, and tested for the ability to discriminate between the cognate and non-cognate substrates. Similarly purified wild-type (WT) RNAP was used as a control. We assembled TECs on pIA349 template that encodes a T7A1 promoter (21). On this template, TECs can be initially halted at position 37 by withholding UTP from the reaction mix (Fig. 2A), purified by gel filtration to remove the unincorporated substrates and then “walked” to the next template position during addition of a subset of NTPs. We used a similar approach to measure rC/dC selectivity using pIA171 template (20) instead.

Substitution of Asn458 Leads to Increased dNTP Utilization—We formed a halted TEC on an appropriate template and tested extension of 32P-labeled RNA upon addition of different unlabeled substrates. For each r/dNTP pair assayed, selected substrate was added to a concentration ranging between 0.1–62.5 μM for rNTPs and 2–1250 μM for dNTPs, reactions were allowed to proceed for 2 min at 37 °C and quenched with STOP buffer (Fig. 2B). The WT RNAP exhibits strong preference toward rNTP substrates, while in the N458D variant these preferences are relaxed. For example, a 2-min incubation with 0.1 μM rGTP led to efficient extension of the nascent RNA by WT RNAP, whereas an NT, not tested. In E. coli RNAP, all substitutions were in the β' subunit; the effects of selected substitutions in T7 RNAP (6) are presented for comparison.

![Image](https://via.placeholder.com/150)

**Fig. 2.** Assay for the incorporation of the r/dNMPs. A, linear pIA349 template used to generate radiolabeled TECs halted at position G37 with the start site indicated by an arrow (top) and the schematic representation of the assays used to measure the utilization of r/dNTP substrates (bottom). B, representative gel panels illustrating the extension of the nascent RNA in complexes halted at A39 or U38 upon addition of the increasing concentrations of r/dGTP and r/dATP substrates. Selection of r/dCTP was assayed on a similar T7A1 promoter template, pIA171. C, comparison of the discrimination efficiencies between four r/d nucleotide combinations. The assays were repeated two to four times for each enzyme-r/dNTP combination; the discrimination efficiencies varied within 20%; NT, not tested. In E. coli RNAP, all substitutions were in the β' subunit; the effects of selected substitutions in T7 RNAP (6) are presented for comparison.
As expected, substitution of Asn<sup>458</sup> for Asp led to substantially relaxed sugar discrimination ranging from 4.8-fold loss of discrimination for rG/dG pair to 26-fold loss in the case of the rU/dT pair. Substitution of Asn<sup>458</sup> for Ser led to a smaller loss of discrimination for rG/dG pair to 26-fold loss in the case of the partially relaxed sugar discrimination ranging from 4.8-fold loss of the bacterial and T7 enzymes (Fig. 2). The direct comparison of these data is not straightforward, however, as we have used pre-steady state assays of substrate incorporation by the TEC, whereas Sousa and co-workers (6, 8) used multiround assays that include both the initiation and elongation steps, which resulted in different discrimination efficiencies for the same RNAP depending on the assay design. In T7 RNAP Tyr to Phe substitution likely completely disrupts polar discriminative interactions with the substrate ribose, whereas Asn for Asp and Ser substitutions in the bacterial enzyme might still maintain specific hydrogen bonding with rNTP sugar moiety. In addition, Asn<sup>458</sup> might be not the only residue participating in the ribose recognition in the active site. Indeed, according to the modeling guanidinium group of a highly conserved in bacteria and eukaryotes ␦␤␥Arg<sup>425</sup> appears proximal to the substrate sugar (Fig. 1B). The experiments are now under way to test the role of ␦␤␥Arg<sup>425</sup> in substrate selection.

Molecular modeling of the multisubunit enzyme pre-insertion site (10) suggested that <i>E. coli</i> ␤␥Thr<sup>790</sup> residue could play a role in ribose selection analogous to that of the T7 Tyr<sup>639</sup> residue. However, the T790V RNAP not only did not loose preference for rNTP substrates, it actually exhibited more stringency in substrate selection relative to the WT, ranging from 1.1-fold effect in case of rC/dC to 4.8-fold effect for rU/dT (Fig. 2C). These data underscore the importance of the Thr<sup>790</sup> residue in substrate selection but argue against involvement of its hydroxyl group in positive selection of the substrate ribose; less conservative substitutions of Thr to Ala and Leu led to gross defects in catalysis (data not shown).

A greater loss of discrimination upon the Asn to Asp substitution suggests that charge distribution on the discriminating protein group is more important than the side chain size. In the absence of the experimental crystallographic data on the RNAP-substrate complex structure, we propose a model that is certainly too tentative to predict the exact scheme of the hydrogen-bonding with the substrate but nonetheless allows to zero in on the three protein groups that likely interact with the substrate ribose: Asn<sup>458</sup> side chain, Asn<sup>458</sup> main chain carbonyl oxygen, and Arg<sup>425</sup> guanidinium group. In the high resolution crystal structure of the <i>T. thermophilus</i> RNAP holoenzyme (17) these three groups form an internal hydrogen-bonding network (Fig. 1B). Interestingly, the main chain conformation of Asn<sup>458</sup> does not fall in the most favorable region on the Ramachandran plot (ϕ = -68°; ψ = 77°). The stabilization of this unfavorable conformational conformation likely comes from the hydrogen-bonding of the Asn<sup>458</sup> main chain carbonyl with Arg<sup>425</sup> and with its own side chain (Fig. 1B). The latter interaction would prevent flipping of the Asn<sup>458</sup> side chain amido group that might be crucial for the proper sensing of the substrate ribose. This network of interaction would be enhanced upon binding of the rNTP substrate. In the model, Arg<sup>425</sup> as well as the main chain carbonyl oxygen and the side chain amide of Asn<sup>458</sup> make contacts with the 3'-OH group, whereas the Asn<sup>458</sup> side chain oxygen likely recognizes 2'-OH group of the substrate ribose. The modeling of N458S substitution showed that although the interactions with the sugar would be weakened due to the smaller size of the Ser side chain, its hydroxyl group may preserve the framework of protein-protein and protein-substrate contacts similar to that of Asn (data not shown). In contrast, the negatively charged Asp side chain would lack the interactions with the main chain carbonyl potentially affecting both the main chain and side chain conformations and thus distorting the optimal orientation of the discriminating residue. In addition, the acidic Asp<sup>458</sup> side chain might also form a salt bridge with the adjacent Arg<sup>425</sup> that would further perturb the Asp interactions with the substrate ribose.

Substitution of Asn<sup>458</sup> Leads to Profound Defects in Transcript Elongation—RNAP variants with altered substrate selection properties could be used to determine the contribution of individual transcript bases to recognition of regulatory signals. This approach is particularly important for the functional analysis of transient kinetic intermediates, such as those occurring during transcription termination and relies on the ability of nucleotide analogs to affect RNAP response to a particular signal (nucleotide analogs interference mapping (NAIM)). This analysis requires that RNAP is able to incorporate substrate analogs bearing modifications at various positions (22) but does not have defects in recognition of the transcription signals intended for study.

We studied the elongation properties of the N458D RNAP on pIA349 template that encodes several well characterized pause sites (Fig. 3). We found that N458D substitution confers a strong elongation defect; the rate of transcription elongation was dramatically reduced (more than 20-fold), and the enzyme paused strongly early in the transcribed sequence. We conclude that N458D RNAP, albeit able to efficiently incorporate dNTPs into the nascent RNA, is not suitable for NAIM due to its profound elongation defects. N458S RNAP also displayed a reduced elongation rate, but the defect was less pronounced (5-fold; Fig. 3 and data not shown). These results are also in a good agreement with our modeling. In the RNAP holoenzyme structure ␦␤␥Arg<sup>425</sup> makes strong hydrogen bonds with ␦␤Asp<sup>464</sup> from the catalytic triad. We presume that these interactions are crucial for the proper positioning of the Asp<sup>464</sup> side chain that is likely required to optimize the coordination of the
major catalytic Mg$^{2+}$ ion (cMg1, Fig. 1B). In N458D enzyme side chains of Arg$^{425}$ and Asp$^{465}$ may form a salt bridge (see above), thereby disrupting functionally important Arg$^{425}$/Asp$^{464}$ contacts and subsequently violating proper Mg$^{2+}$ coordination and compromising catalysis. Consistently, the smaller polar Ser residue would not affect strongly the orientation of the Arg$^{425}$ side chain and would not exhibit dramatic effect on catalysis.

Asn$^{458}$ is located on the catalytic loop next to the active site. Thus the alternative possible explanation of the observed alterations in the substrate selection and catalysis exhibited by the mutant enzymes would be allosteric effects on the active site configuration. However, substitutions of an adjacent Tyr$^{457}$ residue for His or Val did not confer either altered substrate selection or defects in transcription elongation (Fig. 2C and data not shown).

DISCUSSION

The major conclusion of this work is that the β’ Asn$^{458}$ residue in the E. coli RNAP provides the recognition of sugar moiety on the incoming NTP substrate. No significant changes in misincorporation of non-templated NTP substrates arose from Asn$^{458}$ substitutions (see supplemental data), indicating that the determinants for the sugar and base selection are non-overlapping, as was observed for the single-subunit T7 RNAP (11). The proposed ribose selection mechanism would also be likely valid for eukaryotic RNAPs that contain a highly conserved Asn$^{458}$ residue, indeed this role for the corresponding Asn$^{759}$ residue in Rpb1 had been proposed earlier (14).

For both T7 and E. coli RNAPs the selectivity is far below the values reported for the DNAPs that exclude the rNTP substrates with several thousand-to-a million-fold efficiency (1, 2). These differences could be explained by the fact that the levels of rNTPs are at least 10-fold higher in the cell than the levels of the corresponding dNTPs (25). Thus, to prevent incorporation of rNTPs DNAPs must impose the strict discrimination mechanism, which is achieved by the steric exclusion of the ribose 2’-hydroxyl. On the other hand, RNAPs might face the opposite problem: the relatively inefficient (e.g., via a single H bond) discrimination between r- and dNTPs could result not only in the synthesis of compromised messages but also in “draining” of the dNTPs pool. Cellular RNAPs would be expected to evolve tighter control mechanisms as compared with the phage ones, since at the time phage RNAP becomes engaged in transcription the host cell is usually moribund.

In DNAPs, the substrate recognition is thought to occur exclusively in the insertion site (23). In contrast, in the T7 TECs substrate can be bound in either the pre-insertion or in the insertion sites located 10 Å apart, suggesting that RNAPs may select substrates in two rather than in a single site (9, 10). The presence of a pre-insertion site in which the substrate can be “sampled” prior to catalysis was also proposed recently for multisubunit RNAPs (10, 24). Our present data on Asn$^{458}$, which is adjacent to the active site, suggest that in bacterial RNAPs substrate selection, at least in part the rNTP/dNTP selection, occurs in the insertion site. Thus the sugar selection may principally occur in the pre-insertion site in T7 but in the insertion site in bacterial RNAPs. If discrimination in favor of rNTP binding were to occur predominantly in the insertion site, persistence of these interactions after catalysis might hinder the movement of the incorporated nucleotide from the n to the n − 1 site, thereby slowing translocation and the rate of polymerization. In T7 RNAP, where the pre-insertion site binding seems to be preferable, the Tyr$^{385}$-2’-OH contact might be compromised during the transition to the closed form. This might explain a faster rate of T7 RNAP transcription as compared with the multisubunit cellular enzymes (25).

One cannot, however, rule out the possibility that rNTP/dNTP discrimination occurs both in pre-insertion and insertion sites by the different sets of residues. Indeed, substitution of the E. coli β’ Thr$^{790}$ that is located in the bridge helix and likely belongs to the putative pre-insertion site for Val not only did not decrease the enzyme preference for the ribose but instead increased discrimination up to ~5-fold (Fig. 2C), implying the direct interactions of β’ Thr$^{790}$ (or an adjacent) residue with the substrate. Since β’ Thr$^{790}$ is located 18 Å away from the active (insertion) site and ~8 Å from the modeled substrate ribose, this result provides strong support to the hypothesis of existence of the pre-insertion site in multisubunit RNAPs. It is therefore possible that some other residue from the pre-insertion site may play a ribose discriminating role similar to that of Asn$^{458}$ in the insertion site. Alternatively, upon the substrate binding to the insertion site the structural elements that we are now assigning to the pre-insertion site might move toward the RNAP active center to constitute a single closed insertion site. We are currently dissecting the roles of residues in the pre-insertion site of the bacterial RNAP in selection of sugar and the base moieties of the incoming substrate.

REFERENCES