Structural Perspective on Mutations Affecting the Function of Multisubunit RNA Polymerases†

Vincent Trinh,¹ Marie-France Langelier,¹ Jacques Archambault,² and Benoit Coulombe¹*  
Gene Transcription Laboratory¹ and Molecular Virology Laboratory,² Institut de Recherches Cliniques de Montréal, Montréal, Québec, Canada H2W 1R7

INTRODUCTION

In prokaryotes, a single DNA-dependent RNA polymerase (RNAP) synthesizes all classes of RNAs including mRNAs, rRNAs, and tRNAs. Prokaryotic core RNAPs are composed of five subunits (α2ββ′α′) which associate with a σ factor to form the holoenzyme required for promoter recognition (26). In eukaryotes, RNAP I synthesizes the tRNAs and RNAP II forms the mRNA and the small nuclear RNA, while RNAP III is responsible for the synthesis of the tRNA and the 5S rRNA.

RNAPs I, II, and III contain 14, 12, and 17 subunits, respectively. These three enzymes are functionally and structurally related; five subunits are common to all three enzymes, while another four are related (32, 198). The two largest subunits of prokaryotic (bacterial RNAP β′ and β) and eukaryotic RNAPs (RNAP I Rpa190 and Rpa135, RNAP II Rpb1 and Rpb2, and RNAP III Rpe160 and Rpe128) share a high degree of sequence similarity and form the catalytic center of the enzyme. The multisubunit archaeabacterial RNAP is composed of 13 polypeptides, the three largest subunits being the homologues of the two largest subunits of the eukaryotic and prokaryotic RNAPs (101). Six other archaeal subunits show sequence similarity with bacterial and eukaryotic RNAP polypeptides.

The resolution of the crystallographic structures of multisubunit RNAP has provided a framework for elucidating transcriptional mechanisms. To date, high-resolution structures are available for the bacterial RNAP and the eukaryotic RNAP II alone (7, 8, 28, 44, 116) as well as part of a complex with nucleic acids (57, 87, 118, 194, 195) or regulatory factors (11, 29, 86, 119, 185) (see Table 1 and Fig. 1A and B for models).

Both the prokaryotic and eukaryotic transcription reactions involve a number of steps, starting with promoter recognition. The bacterial σ factor (60) and the eukaryotic general transcription factors TFIIA, TFIIB, TFIID, TFIIE, TFII F, and TFIH (42, 62) are required for specific binding of the polymerase to promoters. In both systems, promoter binding is accomplished by bending and wrapping of the promoter DNA against the body of the polymerase (55, 100, 142–144). DNA wrapping is involved in promoter melting in the region of the transcriptional initiation site, probably through the induction of a torsional strain that generates unwinding of the DNA double helix (50, 54). On the basis of recent site-specific protein-DNA photo-cross-linking experiments (54), it was proposed that formation of a right-handed loop in the promoter DNA wrapped around a mobile portion of the enzyme, named the clamp domain, induces DNA unwinding, which can then be stabilized by some general transcription factors. The formation of this open complex allows pairing of incoming ribonucleoside triphosphates (NTPs) to the template DNA strand for phosphodiester bond formation (43, 69, 70).

It has been proposed that all polymerases catalyze phosphodiester bond formation by a two metal ion mechanism (166) (Fig. 1C). By this mechanism, a first Mg²⁺ ion, termed metal A, facilitates the nucleophilic attack of the 3′ oxygen on the 5′ α-phosphate. The second Mg²⁺ ion, metal B, facilitates the release of the pyrophosphate. In prokaryotic RNAP and eukaryotic RNAP II, metal A is coordinated by three strictly conserved aspartates of β′/Rpb1 contained in the NADFDGD motif (44, 57, 206). Metal B has a low apparent affinity for free RNAP II (44) and appears to enter the active site with the...
incoming NTP and is coordinated by three aspartates, two from β'/Rpb1 and one from β'Rpb2, located in a conserved ED motif (195). Formation of a phosphodiester bond is followed by translocation of the nucleic acids in order to present the next template register for a second nucleotide addition cycle.

Transcription initiation is characterized by a cycle of abortive initiation events, where RNAP synthesizes and releases small transcripts without disengaging from the DNA template (188). When the transcript reaches a length of approximately 10 to 12 nucleotides, stabilization of an early elongation complex occurs and RNAP breaks its contacts with the general transcription factors/ and clears the promoter for elongation (43, 69).

In the elongation phase, RNAPs require the help of protein factors to bypass impediments to elongation. Replacement of the general transcription factors by various elongation factors has been demonstrated for RNAP II (134) and numerous elongation factors have been identified to date (159). These factors affect transcription in various ways from rescue of elongation complexes stalled at pause and arrest sites to phosphor-ylation of the C-terminal domain of the Rpb1 subunit of RNAP II and chromatin remodeling and modification.

Because mRNA processing occurs cotranscriptionally in the RNAP II system (68), the complexity of the network of factors interacting with the elongating enzyme is much greater than in the RNAP I and III systems. Recruitment of both mRNA processing and elongation factors is mainly directed through specific interactions with the phosphory-lated RNAP II C-terminal domain. Nonetheless, evidence supports the existence of elongation factors during transcription by RNAP I (95) and RNAP III (113), which lack a C-terminal domain. In prokaryotes, the number of charac-terized elongation factors is smaller, the best-characterized being GreA, NusA, and Mfd (23).

The termination process is different for prokaryotic RNAPs than for eukaryotic RNAPs I, II, and III. Termination in pro-karyotes is the better understood. Two major mechanisms have been proposed. The first is independent of termination factors but rather requires formation of stem-loop structures in the transcribed RNA that interacts with the RNAP to induce paus-ing and which, combined with a weak RNA-DNA hybrid, stim-ulates transcript release from the template (140). The second mechanism involves the activity of the rho helicase to facilitate transcript release and is therefore referred to as rho-depen-dent termination (140).

Eukaryotic RNAP I termination, although poorly under-stood, is for some genes dependent on a short repeated sequence element (93) that is recognized by a nuclear ter-mination factor called Reb1 in Saccharomyces cerevisiae and TTF1 in the mouse (162). Termination depends on the in-teraction of this protein with RNAP I and could involve some protein-induced changes in DNA structure. Other types of signals have been identified but to date (85, 98, 187), no specific termination element has been unequivocally identified.
RNAP II termination is more complex because it is coupled to 3'-end processing of the transcript (39). It requires association of the 3'-end processing complexes CPSF and CstF (13, 47) with RNAP II via the phosphorylated C-terminal domain of Rpb1 as well as the polyadenylation signals at the 3' end of the pre-mRNA (41, 102, 109).

Many cleavage and polyadenylation factors have been identified and studies of thermosensitive alleles of cleavage factors...
show that some of them are required for RNAP II termination (21). RNAP III termination most resembles bacterial RNAP termination and involves recognition of short termination signals rich in T stretches (136). However, in contrast to prokaryotes, RNAP III termination does not involve the formation of stem-loop structures and no auxiliary factor has been identified to date.

Crystallographic studies have revealed that the structure of RNAP is highly conserved from prokaryotes to eukaryotes, with the regions of highest homology forming the enzyme’s active center (44, 206). The two largest RNAP subunits (β’ and β in bacterial RNAPs; Rpb1 and Rpb2 in RNAP II) form a positively charged cleft (also termed the main channel) that accommodates the nucleic acids during transcription (44, 206) (Fig. 1A). The two catalytic Mg$^{2+}$ ions, metals A and B, are buried deep in this cleft (Fig. 1B) (195). A domain of Rpb2/β called the wall closes the upstream extremity of this cleft and is a binding site for the upstream end of the RNA-DNA hybrid. The wall domain contains the flap feature, ordered in the prokaryotic structures and disordered in the eukaryotic structures, which serves as a binding site for transcription factors and would be implicated in the obstruction of the RNA exit channel by the σ factor region 4 (123).

The nucleic acids are held in the cleft by a number of protein domains, including the upper and lower jaws, which grab the downstream DNA, and by the clamp domain, which locks the nucleic acids in the cleft (57). Comparison of the structures of the free core enzyme (44) with that of the elongating core enzyme (57) from yeast led to the hypothesis that the mobile clamp domain folds over DNA to lock it during elongation. However, the backbone models (7, 28) and refined atomic model (8) of complete RNAP containing Rpb4 and Rpb7 show the clamp domain in the same position as in the elongating core structure, suggesting that promoter DNA is first loaded on the top of the clamp as seen for bacterial RNAP (7, 28, 118). The template strand would then reach the active site only after promoter melting had occurred. A pore structure (also termed secondary channel), which has the shape of an inverting funnel that opens in the cleft near the active site, was proposed as an entry route for the incoming NTP and an exit for the released pyrophosphate (44, 57). The mRNA exit channel lies between the wall and the clamp domain (Fig. 1A).

In addition to their polymerization activity, DNA-dependent RNAPs can also catalyze the 3’ endonucleolytic cleavage of transcripts under certain circumstances (40, 158). First discovered in Escherichia coli (92, 172), this cleavage activity takes place when RNAP has backtracked at pause and arrest sites (53, 196). Weakly associated DNA-RNA hybrids can induce backtracking of the enzyme to a more stable register (125). To resume elongation from this more stable position, RNAP needs to cleave the 3’ end of the transcript that now protrudes from the catalytic site through the pore.

The 3’ cleavage activity is enhanced by the cleavage factors GreA and GreB in bacteria (24, 25) and TFIIIS in eukaryotes (53, 196). The backbone structure of TFIIIS in a complex with complete RNAP II (86) provided insight into the mechanism by which this factor enhances transcript cleavage. More specifically, it revealed that two conserved essential acidic residues of TFIIIS (79) are located in the vicinity of the polymerase metal B and could participate in its coordination. Therefore it was suggested that the two metal ion mechanism invoked for NTP polymerization is also involved in the 3’ endonucleolytic activity of RNAP II. Similar studies on GreB, the prokaryotic homologue of TFIIIS, have also reached similar conclusions for bacterial RNAP (128, 164).

Structures of yeast elongating core RNAP have revealed the presence of an 8- to 9-base-pair DNA-RNA hybrid in the cleft (57, 194). The crystallographic data have also shown that a number of loops and helices of Rpb1/β’ and Rpb2/β are located in this region close to the catalytic site (Fig. 1B). These structural motifs have been named either according to their location, aspect or presumed role in the transcription reaction. For example, the Rpb1/β’ bridge helix separates the main channel and is located near the template DNA at the +1 site.

Because crystallographic data revealed two different conformations for this structure, bent (or distorted) (206) or straight (or continuous) (44, 57, 194), it was proposed that the bridge helix might be involved in translocation of the nucleic acids during transcription. The location of the rudder, the lid and the fork loop 1 suggests that these loops are involved in DNA-RNA strand separation, in order to maintain an 8- to 9-base-pair hybrid (194). The fork loop 2 and the zipper could be involved in delineating the downstream and upstream boundaries of the transcription bubble respectively (44, 57). Five loops of Rpb1 and Rpb2 have been termed the switches and could participate in controlling the position of the clamp or, for switches 1 to 3, in forming a binding site for the DNA-RNA hybrid (7, 44, 57).

Crystallographic studies of yeast core RNAP in complex with the general transcription factor TFIIIB have shown contacts of this factor with the dock domain (29) of the enzyme. This domain of Rpb1 is located between the wall and the clamp and lies on the surface of the structure (Fig. 1A). The finger domain of TFIIIB enters deep into the active site after passing across the saddle, between the wall and clamp domains. This feature is reminiscent of the 3.2 linker loop of the σ factor in prokaryotic RNAP, which follows a similar path through the RNAP flap and clamp domains (185).

Superposition of the crystallographic structures of the core RNAP/TFIIIB complex and the core elongation complex (57) shows that the RNA would clash with the TFIIIB finger domain in the active site beyond synthesis of the fifth residue and that TFIIIB, if it were not displaced, would compete with RNA for binding to the saddle after synthesis of the 10th nucleotide. These observations provide a molecular basis for abortive initiation before synthesis of the 10th nucleotide and explain why TFIIIB is released from the complex beyond synthesis of this register.

Crystallographic structures of RNAPs in complexes with known catalytic inhibitors have also been resolved and have brought new insight into the catalytic mechanisms of RNAPs. Observation of these inhibitors at their binding sites allows a better understanding of their mode of action and the role of the structural features they interact with. The structures of coenzymes of α-amanitin (27), rifampin (30), sorangicin (31), and streptolydigin (176, 182) with core RNAP have all been resolved. α-Amanitin, an inhibitor of the translocation step, is seen binding to a region located between the funnel and the bridge helix (27) of yeast RNAP II (Fig. 1B).
Many observations support the hypothesis that restriction in the movement of the bridge helix is required for translocation of nucleic acids in the active site (61, 184). In prokaryotes, rifampin and sorangicin both bind in a pocket juxtaposed to fork loop 2 (30, 31). Modeling of an RNA-DNA hybrid in a bacterial RNAP structure (90) combined with positioning of rifampin to its binding site reveals that a steri.c clash would occur after synthesis of the second or third nucleotide (30), which is in good agreement with observations that rifampin and sorangicin specifically inhibit RNAP during initiation at a step before addition of the second or third nucleotide (90, 115).

An additional hypothesis about the mode of action of rifampin has been suggested and involves an allosteric signal transmitted through the binding site of rifampin to the catalytic aspartates in order to lower the affinity for the Mg2+ ion (12). The location of streptolydigin in the RNAP structure and its binding to the fork 2, bridge helix, and trigger loop features indicate that this drug could interfere with the translocation of nucleic acids in the active site (182). However, an alternative hypothesis suggests an allosteric mechanism trapping RNAP in a conformation inappropriate for transcript elongation (176).

Recent reports have enlightened the evolutionary connection between prokaryotic and eukaryotic RNAPs (75, 76). The presence of the catalytic double-psi β-barrel domain, which contains a signature metal-coordinating motif, in both eukaryotic RNA-dependent RNA polymerases and the universally conserved β’ subunit of DNA-dependent RNA polymerases, coupled to the absence of other common domains, suggests that they have evolved through the early divergence of a common ancestor. Interestingly, the presence of another, although highly diverged, double-psi β-barrel domain in the β subunit of DNA-dependent RNA polymerases supports this idea of a common ancestor that diverged through lineage-specific insertions of domains and motifs.

During the past two decades many mutational studies of RNAP have been performed with the aim of understanding the function and regulation of this enzyme. A great number of these studies were performed before the resolution of RNAP crystal structures. We surmised that reexamining the effect of amino acid alterations in RNAP in light of their available three-dimensional structures might provide new insight into the structure-activity relationship of these enzymes.

We therefore created an extensive catalogue of published mutations affecting the function of multisubunit RNAP and mapped them on the structures of the prokaryotic and eukaryotic enzymes. This catalogue is presented in the form of a table (see Table S2 in the supplemental material for mutations discussed in this article and additional mutations; the complete catalogue is also available in a Web format at http://www.ircm.qc.ca/microsites/mutationsaffecting/fr/index.html) indicating (i) the reference in which the mutation is described, (ii) the name of the mutant allele when available, (iii) the organism and subunit, (iv) the position of the altered amino acids on the linear amino acid sequence of the altered subunit, (v) the equivalent position on either S. cerevisiae RNAP II or Thermus aquaticus RNAP (the amino acid numbering in the text and figures refers to these specific positions), (vi) the name of the structural domain where the altered amino acid is located, (vii) the homology region, and (viii) the function affected by the altered amino acids or phenotype for 6-azauracil sensitivity or drug resistance (blanks stand for mutants with phenotypes for which no affected function could be attributed). This work, as presented in this review, has allowed us to make a series of observations concerning the functions of domains and structural elements of RNAP.

**AMINO ACID SUBSTITUTIONS IN RNAP THAT LOWER THE POLYMERIZATION RATE ARE LOCATED CLOSE TO THE ACTIVE SITE**

An important, albeit predictable, observation is that a great number of substitutions inducing a low polymerization rate map in proximity to the active site. RNAP mutants that have been included in this category are defective in transcription assays in vitro. In most cases, these mutants have not been characterized sufficiently to define which step of the transcription process is impaired.

In Fig. 2, we have modeled in the structure of S. cerevisiae RNAP II (87) (Fig. 2A and C) and T. aquaticus RNAP (30) (Fig. 2B and D) a series of substitutions in both prokaryotic and eukaryotic enzymes that induce a low polymerization phenotype. Site-directed (191, 205) and random mutagenesis (48, 117) of the three strictly conserved aspartates of β’/Rpb1 binding metal A, which are shown in Fig. 1C and 2 (for β’, D739, 741, and 743; and for Rpb1, D481, 483, and 485), generates mutants that are completely inactive in transcription (191, 205) with the surprising exception of the random mutation D743G conferring microcin J25 resistance in E. coli. (The amino acid numbering in the text and figures refers to amino acid positions on either S. cerevisiae RNAP II or T. aquaticus RNAP.)

Mutants arising from site-directed (99, 163, 191) and random mutagenesis (103) of residues ED of β’/Rpb2 (β, E685K/A and D686A; Rpb2, E836A and D837A) present a reduced polymerization activity, especially at low NTP and Mg2+ concentrations (99, 103, 163, 191) (Fig. 2A). Interestingly, the D837A substitution in yeast Rpb2 gives rise to a dominant negative phenotype, indicating that the mutated subunit interferes with the function and/or assembly of the wild-type enzyme (99).

The substitutions that cause a low polymerization rate and map near the catalytic acidic residues are indicated in Fig. 2. Bacterial residue N737 and its eukaryotic counterpart N479 are juxtaposed to the three aspartates of β’/Rpb1. In yeast, the N479Y substitution is lethal (4, 48), but when combined with an S476I substitution results in a low-catalytic-rate phenotype in vitro (48, 66) as seen in the site-directed E. coli mutant N737A (163).

Based on the structure of RNAP, it was proposed that residue N737/N479 binds the ribose 2’-OH group and is involved in discrimination of NTP versus dNTP (87, 173). Replacement of this amino acid might allow inclusion of dNTPs in the RNA chain, a situation that may eventually lead to a block in polymerization. This block could be a consequence of the fact that the nucleic acid binding site is highly specific for a particular conformation of the DNA-RNA hybrid that is modified upon addition of a dNTP (57). Alternatively the N737/N479 replacement could impair binding of the incoming NTP to the active site. In addition, because this amino acid is located in close proximity to the catalytic aspartates of the largest subunit, it is
possible that its replacement causes a distortion of the active center, leading to a defect in transcription.

In prokaryotes, the engineered β R879A replacement (Fig. 2B) lowers the polymerization rate of the enzyme (163). In the crystallographic structure of yeast RNAP II (194), the equivalent amino acid, R1020, is involved in a salt bridge with the catalytic residue D837 of Rpb2. Disruption of this bridge could cause changes in the conformation of the active site, having a
negative effect on the catalytic activity of the enzyme. Two thermostensitive and 6-azauracil-sensitive replacements in the Rpb2 subunit of yeast RNAP II, A1016T (135, 150) and P1018S (105, 135, 150), are located near R1020, and could have a similar effect by modifying the position of R1020 and its interaction with D837.

The bacterial replacement S1091F (154) which confers resistance to streptolydigin is located in the bridge helix, an important helix of β/Rpb1 proposed to be involved in nucleic acid translocation (44, 57) (Fig. 2B). In the structure of T. aquaticus RNAP (30) with a DNA-RNA hybrid modeled in the active site, S1091 could contact the template DNA strand near +2. The replacement of a small serine for a large phenylalanine probably can result in a steric clash with the template DNA and disturb its presence in the active site during polymerization.

The I1252V replacement in E. coli also alters the bridge helix function, but indirectly through an effect on the juxtaposed trigger loop. This mutant is also overreactive to regulatory signals such as terminators and pause signals (16). The site-directed K838A and K838R replacements, identified by affinity cross-linking studies as residues being close to the active site, confer a defect at the promoter clearance step of the transcription cycle (83, 120, 148). These two K838 mutants synthesize abortive transcripts but are unable to escape the promoter and proceed into elongation.

Interestingly, the positively charged side chain of K838 is directed toward the RNA around position +2. One can propose that K838 contacts the RNA at this position, leading to its stabilization in the active site. Replacement of this important lysine with a neutral alanine residue or arginine, a residue with a less-flexible and bulkier side chain, possibly results in the inability to retain the transcript in the catalytic center, leading to abortive initiation and impairment of the transition to elongation.

The following mutants were engineered in order to investigate the properties of the conserved residues surrounding K838. K846A (148) (Fig. 2B) is similarly oriented toward the DNA-RNA hybrid. It is possible that the positively charged K846 amino acid is also involved in the interaction with the nucleic acids and their stabilization in the active site. The G844A replacement could similarly affect the polymerization rate by changing the conformation of this important hybrid binding region. Two replacements in the switch 2 region of Rpb1, K332A and R344A, confer an increase in abortive initiation (110). Both replacements are located in close proximity to the template DNA strand around positions +1 to −3. Although the R344A mutant is able to bind DNA as well as the wild type, this mutant is unable to form stable RNAP II-DNA-RNA complexes containing a 5-nucleotide RNA. This observation suggests that the observed defects are linked to an inability to stabilize the hybrid in the active site.

Among the amino acid replacements near the catalytic center that lower the polymerization rate, many are located in proximity to fork loop 2 (Fig. 2C and D). This β/Rpb2 structural element was proposed to be involved in maintaining the downstream end of the transcription bubble (44, 57, 86). In prokaryotes, the replacement F425S (64) conferring streptolydigin resistance is located directly in fork loop 2. Considering its location, the neutral phenylalanine F425 could be implicated in an interaction with the DNA helix during strand separation (Fig. 2D). We speculate that the replacement for a serine could lead to a promoter opening defect and therefore to a low polymerization rate. Similarly, replacements Q393P (81, 82, 202), conferring rifampin resistance, and G446D (174, 175), isolated based on its increased termination phenotype, both surround fork loop 2 and possibly destabilize the structure of this element resulting in the inability to maintain the transcription bubble open.

Because R1078, a residue located in the E. coli β′ bridge helix, forms a hydrogen bond with the backbone of fork loop 2 at position H431, the disruption of this bond by the R1078C (17) replacement, isolated from its ability to mimic the ppGpp effect (see below), could induce a conformational change that alters an important interaction of the bridge helix with fork loop 2. Figure 2C shows that the eukaryotic replacements S480I, G503K, G503K, L508D, and K510N (22), isolated from their altered termination properties, are localized in fork loop 2 and induce a low-polymerization phenotype.

Other replacements from these studies, including V536E, L541S, K537N, and N538Y (22), are positioned in a region that interacts with both the N-terminal and C-terminal ends of fork loop 2; they possibly induce misfolding of this element. Finally, the R512C replacement in yeast Rpb2 is also located in the fork 2 element. Isolated as a suppressor of the ssu72-2 defect (131), this mutation affects transcription. It would be interesting to assess whether all of these fork 2 mutations alter the architecture of the transcription bubble in vitro.

AMINO ACID REPLACEMENTS IN RNAP II AFFECTING SELECTION OF THE TRANSCRIPTION INITIATION SITE

In eukaryotes, selection of the transcription initiation site is dictated by different promoter elements, such as the TATA box and the initiator (161). The general transcription factor TFII B was also shown to play a role in initiation site selection. Indeed, many mutants of TFII B that alter initiation site selection have been isolated (20, 132, 133). Consistently, photo-cross-linking studies have shown that a domain of TFII B approaches the DNA template near +1 in the initiation complex (35). The Rpb9 subunit of RNAP II also plays a role in initiation site selection and many replacements affecting this process affect residues of this RNAP II subunit (71, 171). Rpb9 forms the jaw-lobe module, in combination with the jaw domain of Rpb1 and the lobe domain of Rpb2 (see Fig. 1A). This module may grip the downstream DNA during initiation, suggesting that these contacts are important for selection of the initiation site.

Figure 3 shows the locations of amino acid changes affecting initiation site selection by RNAP II. Replacements C317Y (65) and E368K (65) were isolated for their ability to restore the expression of genes rendered inactive by insertion of a Ty retrotransposon in their promoter regions. Both of these replacements map to the Rpb2 lobe domain which interacts with Rpb9. We can hypothesize that these replacements alter the lobe in a way that affects its interaction with Rpb9, thereby affecting initiation site selection. Alternatively, these replacements may alter the ability of the jaw-lobe module to firmly contact downstream DNA, impairing the accurate positioning in the +1 register of the active site.
The G369D (65) and G369S (34) replacements also affect the Rpb2 lobe domain. The G369S mutation was isolated as a suppressor of a TFIIB mutation affecting initiation site selection (34). The Rpb2 replacement R512C (131) located in the fork 2 domain also induces a downstream shift in transcription start site. This mutation has been isolated as a suppressor of the ssu72-2 temperature-sensitive defect, an essential uncharacterized protein known to interact with TFIIB. The location of this replacement within the fork 2 element suggests that this residue participates in stabilizing contacts with nucleic acids or maintaining the transcription bubble architecture. The proximity of the fork 2 element to the finger domain of TFIIB raises the possibility of an altered interaction between these two features, leading to a defect in start site selection.

The trigger loop is a structural element that was recently shown to play a role in elongation by modulating the oscillating properties of the adjacent bridge helix (16) (Fig. 3). Two initiation site selection mutants carry alterations in residues of the trigger loop (T1080I and S1096F) (65). In the available structures, a large portion of the trigger loop is missing, consistent with this element being mobile and capable of adopting different conformations relative to the +1 site. Therefore, replacing T1080 and S1096 could affect the selection of the initiation site by directly altering the structure and/or function of the trigger loop. Alternatively, replacement of T1080 and S1096 could affect the interaction of the trigger loop with the bridge helix, affecting the lateral mobility of RNAP (16) and consequently altering start site selection.

Many amino acid replacements affecting start site selection could exert their effect by altering the interaction between RNAP II and TFIIB. A402R (20) is located in the dock domain of Rpb1 (Fig. 3) and has been isolated as a suppressor of a TFIIB mutation affecting start site selection. Crystallographic structures of TFIIB bound to RNAP II revealed that TFIIB contacts the dock domain (29). Replacement N445T (4) displays a sit phenotype (see below) and this replacement affects start site selection. The N445S (20) replacement also affects start site selection and both N445T and N445S locate close to the saddle and thus could destabilize TFIIB function by altering its interaction with RNAP II.

Finally, replacements K332A and R344A in the switch 2 element of Rpb1 have been isolated from their reduced growth phenotype and exhibit an increase in abortive initiation but also a downstream shift in start site utilization (110). These two residues are close to the template DNA at positions +1 to −3 and the authors suggest that the mutant RNAPs are unable to form a stable complex with the DNA-RNA hybrid. It should be noted that K332A is within 5 angstroms of the finger domain of TFIIB, providing another basis to the observed start site selection defect.

**AMINO ACID DELETIONS AFFECTING FORMATION OF THE OPEN COMPLEX**

Figure 4 shows the locations of three prokaryotic RNAP mutations that affect open complex formation prior to initiation. The engineered mutant Δ174-311 (121, 152) carries an almost complete deletion of the β dispensable region which is located in the lobe domain. The transcription bubble formed by this mutant is shorter at its downstream end. A shorter
deletion engineered in the lobe (Δ314-325) also leads to the formation of an open complex that is shortened in the downstream region but shows a slightly more severe phenotype in that it fails to fully protect the promoter DNA in a DNase I footprinting assay (121). Even though the deleted region is longer for Δ174-311 than Δ314-325, the N- and C-terminal ends of the deleted regions in Δ314-325 are spatially more distant than they are in Δ174-311, where they appear almost juxtaposed (Fig. 4).

It was proposed that the movement of the lobe domain toward DNA is necessary to maintain a grip on the DNA (90). A suggested mechanism for open complex formation involves downstream promoter contacts with the lobe regions of RNAP to allow for the downstream propagation of the transcription bubble (121). We can hypothesize that the formation of the transcription bubble requires a certain degree of restriction in the axial rotation of DNA in order to generate an unwound helix at the active site.

Mutant Δ586-601 is also defective in downstream promoter opening (94). Indeed, Δ586-601 is unable to initiate transcription unless the template is premelted in the −7 to +1 region. This deletion mutant was constructed to assess the role of the β’ rudder which is an element that does not contact the template DNA in the RNAP structure. The position of the non-template DNA is not revealed in the different crystallographic structures of RNAP; it therefore remains possible that the rudder interacts with the non-template DNA strand and thereby contributes to establishing or maintaining an open promoter.

RNAP II MUTATIONS ASSOCIATED WITH 6- AZAURACIL SENSITIVITY

TFIIS was first identified as an elongation factor that assists RNAP II to elongate through arrest sites in vitro (137, 139). TFIIS, like its bacterial counterparts GreA and GreB (24, 25), greatly stimulates the 3’ transcript cleavage activity of RNAP (77, 138, 189). This intrinsic endonucleolytic cleavage activity takes place when RNAP has backtracked at pause and arrest sites (53, 196). To resume elongation, RNAP needs to cleave the 3’ end of the transcript that now extrudes from the catalytic site through the pore.

The crystal structure of TFIIS in a complex with RNAP II (87) revealed that TFIIS binds RNAP II along the floor of the cleft and repositions the RNA in the active site. An acidic hairpin of TFIIS reaches all the way into the pore in such a way that it positions two highly conserved residues, D290 and E291, essential for stimulating cleavage activity (79), near the metal B ion of the enzyme’s active site. These acidic amino acids have been proposed to participate in the coordination of metal B and to directly position and activate a nucleophilic water mol-
ecule (87), thereby stimulating cleavage activity and/or elongation by RNAP.

Inhibition of the enzyme IMP dehydrogenase by 6-azauracil results in depletion of the intracellular pools of GTP and to a lesser extent that of UTP in eukaryotes (52). In yeast, mutants with a defect in transcriptional elongation often show a growth defect on 6-azauracil-containing medium (141). Not surprisingly, a number of RNAP II in-frame linker insertion mutants with 6-azauracil sensitivity have been isolated (Fig. 5). For some of these mutants, overexpression of TFIIS (by increased dosage of the \textit{PPR2} gene) was shown to suppress the 6-azauracil phenotype.

An interesting observation is that almost all the 6-azauracil mutations suppressed by overexpression of TFIIS, R1135(LELE)S (3, 5), T1141(SSSS)T (3, 5), E1230K (5), E1230K E1264K (5), L1236(TRARV)I (5), and D1257(SSSS)H (5), affect residues in Rpb1 located along the TFIIS-RNAP II binding interface (86). Alterations of these RNAP II residues probably affect the interaction with TFIIS, conferring an elongation defect and the 6-azauracil sensitivity phenotype. In fact, it was confirmed by biochemical means for two of these RNAP II mutants, E1230K and L1236(TRARV)I, that they indeed have a lower affinity for TFIIS (199).

A series of additional replacements conferring 6-azauracil sensitivity are not suppressed by TFIIS overexpression and are scattered throughout the structure of the enzyme (Fig. 5). The V32(ARAR)A insertion (3, 5) in Rpb1 is located on the outside of the clamp domain. It is possible that this insertion alters the interaction of RNAP II with TFIIS.

Indeed, TFIIF was found to localize to the clamp domain of RNAP II (38, 54). The Q124(ARAR)A insertion (3, 5) is located on the clamp domain at the opening of the cleft, where downstream DNA is located. Therefore, this amino acid is possibly important for the interaction of RNAP II with DNA. The G369S (34) replacement is located in the lobe region. Notably, this replacement was identified in a screen for suppressors of TFIIB mutations affecting initiation site selection. The G369S replacement may well alter the conformation of the lobe region, resulting in the correction of the TFIIB defect, but also conferring an elongation defect leading to 6-azauracil sensitivity.

Finally, three additional RNAP II mutants, the A1016T (135, 150), P1018S (105, 135, 150), and S476I\_N479Y (147) mutants, present a 6-azauracil sensitivity phenotype because of amino acid alterations in Rpb2. Two of these replacements, A1016T and P1018S, are located in the hybrid binding region, while S476I and N479Y is located near the aspartates coordinating metal A. It is therefore likely that the 6-azauracil sensitivity of these mutants results from a transcriptional elongation defect brought about by distortion of the enzyme’s active site, rather than a defect in its interaction with elongation factors.

**RNA POLYMERASE MUTANTS PRESENTING HIGHER POLYMERIZATION RATES OR SUPPRESSING POLYMERIZATION DEFECTS**

Yeast \textit{sit} mutants (suppressor of initiation of transcription defect) were isolated in a genetic screen for suppression of a...
transcriptional defect at the \textit{HIS4} gene (10) brought about by removal of three transcription factors required for \textit{HIS4} expression. In a strain carrying deletions of the \textit{BAS1}, \textit{BAS2}, and \textit{GCN4} genes, which encode transcription factors that act at the \textit{HIS4} promoter, the \textit{HIS4} gene is poorly transcribed, leading to auxotrophy for histidine (9, 10). Interestingly, the low level of \textit{HIS4} transcription observed in such a yeast strain can be increased by mutations in \textit{RPB1} or \textit{RPB2}. It was proposed that these mutations increase the processivity of the polymerase and/or the number of initiation events at the \textit{HIS4} promoter. Unfortunately, only the \textit{sit} mutations mapping in RPB1, not those in RPB2, have been sequenced. In either case the mutant RNAP II enzymes remain to be characterized in vitro in order to pinpoint the mechanistic basis for suppression.

RNAP mutants having a higher polymerization rate or suppressing polymerization defects appear to affect different regions of the enzyme. A first group of amino acid replacements are located near the catalytic aspartates of Rpb1/β′ or near the active center (Fig. 6A). N445T (4) and M487L (4) are located in this region and were isolated as \textit{sit} mutations. It is possible that these replacements affect the position of the catalytic aspartates, thereby affecting the polymerization rate.

Because N445S (20) was found to affect start site selection, and because position N445 falls near the saddle which interacts with TFIIIB, it is conceivable that N445T also affects the number of initiation events. The Rpb2 S771C (36) replacement, identified as an intergenic suppressor of the synthetic lethality of the \textit{Drosophila} Rpl215 K1 mutation, is located in the hybrid binding region, near the RNA at position +9 and the fork loop 2 segment. Replacement with cysteine could cause an increased interaction between the hybrid region and fork loop 2. Therefore, and as replacements in the fork loop 2 region can induce lower polymerization rates (see above), it is possible that S771C has the opposite effect.

A second group of replacements are localized in the Rpb1/β′ funnel domain (Fig. 6). The role of the funnel in transcription has not yet been defined. However, it was shown that the funnel and the bridge helix form the binding site for the inhibitor α-amanitin (27). Many prokaryotic mutations, including A1028T (190) and M1045I (154, 190) isolated from altered termination and streptolydigin resistance, and eukaryotic ones, including L702S (147), a suppressor of elongation defects of RNAP III mutant rpc160-112; G730D (4), a \textit{sit} mutation, K744S (147); A759P (4); and C764F D668N (4), that cause either a higher polymerization rate, a suppressed polymerization defect, or a \textit{sit} phenotype are found in this region, suggesting a role for the funnel in transcript elongation. Because the funnel is juxtaposed to the bridge helix, we propose that it regulates its ability to translocate the nucleic acids during polymerization. The funnel is also exposed at the surface of the enzyme and could theoretically be contacted by certain transcriptional regulators.

The bridge helix is the site of many replacements that confer a higher polymerization activity or can suppress mutations inducing polymerization defects (Fig. 6). In eukaryotes, replacement of glycine 820 with an aspartate (G820D) (4) probably alters the conformation of the bridge helix in a way that facilitates translocation (Fig. 6A). The A832V (4, 190) replace-
ment is located near the base moiety of the +1 nucleotide in the template DNA. It was proposed that the invariant alanine 832 helps position the +1 nucleotide through van der Waals interactions (57).

The replacement of the alanine for a valine, a bulkier hydrophobic residue, might allow a better interaction with the base of this nucleotide. This could stabilize the nucleic acids in the cleft, resulting in a more efficient polymerization. The L841I replacement (147) has its side chain oriented toward the two helices that form the N- and C-terminal ends of the trigger loop. The L841 residue could be involved in a number of interactions with neutral amino acids in these structural elements. Replacement of a leucine with an isoleucine is not a drastic modification. It may be that this region of the bridge helix is critical for the processivity of the enzyme. In prokaryotes, the R1078H (153, 190) replacement is part of the bridge helix, but is oriented toward fork loop 2 (Fig. 6B). The R1078 residue is involved in an interaction with the backbone of fork loop 2. Replacement of an arginine for a histidine could potentially stabilize an important interaction of the bridge helix with fork loop 2, leading to a higher rate of polymerization. Consistently, replacement of arginine 1078 with a cysteine (14, 17), which is too small to allow any interaction with fork loop 2, produces a low-polymerization phenotype as discussed above.

Other prokaryotic and eukaryotic RNAP mutants are affected in residues close to the bridge helix. The replacements S1436F (190) and A1437V (190) in E. coli RNAP are found in a β-sheet surrounded by two β-strands. One of these β-strands is located in the region adjacent to the trigger loop and is proximal to the bridge helix. Therefore it is possible that this replacement also increases the enzyme’s polymerization rate by affecting the functions of the trigger loop or the bridge helix. Finally, a replacement in E. coli, Q1254S (16), which is located in the trigger loop, produces a fast-elongating RNAP in vitro and is thought to affect the oscillating properties of the bridge helix (16).
FIG. 8. (A) Distribution of RNAP II substitutions conferring α-amanitin resistance and mapped onto the *S. cerevisiae* RNAP II structure 1K83. Residues are colored differently for clarity. (B) Distribution of RNAP substitutions conferring rifampin resistance and mapped onto the *T. aquaticus* RNAP structure 1I6V. Substitutions inducing low levels of rifampin resistance are shown in pink. (C) Distribution of RNAP substitutions and insertions conferring streptolydigin resistance and mapped onto the *T. aquaticus* RNAP structure 1ZYR. Red, pink, and magenta residues are located in the bridge helix, fork domain, and trigger loop, respectively. (D) Distribution of RNAP substitutions conferring sorangicin resistance and mapped onto the *T. aquaticus* RNAP structure 1YNJ. (E) Distribution of RNAP substitutions conferring microcin J25 resistance and mapped onto the *T. aquaticus* RNAP structure 1I6V.
AMINO ACID REPLACEMENTS AFFECTING TERMINATION

A great number of replacements in E. coli and eukaryotic RNAPs were reported to affect the termination step of the transcription reaction (Fig. 7) (16, 22, 78, 82, 97, 147, 154, 157, 165, 174, 190). These mutants were isolated from their altered ability to read through terminators located in specific genes. Many terminators have been used, including SUP4, T1, ρ-dependent and ρ-independent terminators, trp attenuator, thr attenuator, rntT1, T7, his, Tr2, P14, T3, lambda Tr2, and ops, in various contexts.
In some cases the ability to read through a terminator leads to the normal expression of a gene required for survival, allowing isolation of RNAP mutants with decreased termination ability. In contrast, expression of toxic genes from readthrough of terminators by wild-type RNAP allows for screening of mutants with increased termination properties. Most of these replacements affect the lobe, including G338E (22, 157), T339T (22, 157), T339T K353K (22, 157), A340K (22, 78, 157), L341I (22), D354K (22, 157), D399E (22), D399A (22), R405A (22), and T240C (97); the cleft, including I1252V (16), S1091F (154), Q1254S (16), R1078H (154, 190), and H1367Y (190); the fork, including S480I G503K (22, 157), V536E L541S (22, 157), K537N (22, 157), N538Y (22, 157), K510N (22, 157), G503K (22, 157), L508D (22, 157), S480I (22, 157), G446D (22, 157), Q393P (81, 82), P440S T443I (97), P440L (175), and S454F S591F (175); and the funnel, including L702S (147), M1023I (190), S1026F (154, 190), A1028T (190), and M1045I (190), regions of the enzyme or are into or close to switches, including R1129K E1149D (22, 157), I1139L R1129S (22, 157), M1133L (22, 157), S1436F (190), A1437V (190), G609D (190), S626F (190), S1439L (190), G1011S (175), P1020L (175), and G1028D (175).

These domains also regroup replacements carrying particular phenotypes such as altered polymerization rates and unstable open promoter complexes. For example, the funnel and the fork domains are altered in mutants showing various polymerization defects. It is tempting to speculate that the termination phenotypes observed in this case derive from an altered elongation phenotype. The lobe seems important for opening and maintaining the transcription bubble. We can hypothesize that a transcription complex in which the transcription bubble has a tendency to collapse is more prone to arrest at a termination site.

Many termination mutants also show altered switch domains. The switches were proposed to form a portion of the binding site for the DNA-RNA hybrid (7). Therefore, it is possible that replacements in these structural elements affect the stability of the RNA and its retention in the active site, leading to increased or reduced ability to terminate transcription.

**DRUG-RESISTANT RNAP MUTANTS**

Over the past years, numerous drug-resistant RNAP mutants have been isolated and characterized. α-Amanitin, rifampin, streptolydigin, sorangicin, and microcin J25 resistances have all been studied thoroughly allowing new insights regarding the transcriptional mechanisms of RNAPs to be gained. Moreover, understanding the mode of action of these and other inhibitors is pivotal for the design of novel and potent antipathogen drugs.

α-Amanitin is a bicyclic octapeptide used for a long time as a specific RNAP II inhibitor (106). The crystallographic structures of RNAP II in a complex with α-amanitin were resolved...
by Bushnell and colleagues in 2002 (27) and show binding of this drug to the funnel and the bridge helix features. Known replacements conferring α-amanitin resistance, including L722F (18), R726H (37)/P (19), I756F (19), C764Y (19), S769D (18, 19), and G772E (19), have been mapped onto the crystal structure of RNAP II in Fig. 8A to show that they all cluster around this binding site.

Because the bridge helix feature has been proposed to be involved in nucleic acid translocation, α-amanitin most probably impairs the translocation of nucleic acids by interfering with the function of the bridge (58). This is further supported by studies showing that after addition of α-amanitin to a transcribing RNAP II complex, synthesis of a single phosphodiester bond can still occur (61, 184).

Rifampin, a member of the rifamycin family, is an efficient antibiotic against bacterial pathogens and binds with high affinity to bacterial RNAPs (63). Crystals of RNAP in a complex with rifampin were obtained by Campbell and colleagues in 2001 (30) and permitted the identification of the binding pocket and the description of a mechanism of action for this drug. Direct interactions of rifampin with RNAP are mostly made with a region of the β subunit forming the fork domain.

In accordance with these findings, most of the rifampin-resistant mutations isolated so far map directly in or adjacent to the fork domain (Fig. 8B). More specifically, they are found in the protrusion, fork and external 1 and 2 domains of β (14, 31, 81, 82, 97, 108, 120, 127, 129, 130, 156, 160, 174, 178, 186, 201, 202, 207). Thus, these replacements most probably affect the conformation of the binding pocket and lower its affinity for rifampin. Other mutations conferring lower levels of rifampin resistance are mostly located away from this binding pocket. Many of them are found in the lobe region and suggest that binding of rifampin to the fork region is somehow linked to the lobe’s conformation when holding the downstream DNA. Thus, it appears that binding of the lobe domain to DNA is necessary, either directly or indirectly, for the formation of the rifampin-binding pocket.

Numerous rifampin-resistant mutations have been isolated through the years and their locations overlap a small region where some streptolydigin-resistant mutations (74) cluster. Nonetheless, the biochemical effects of these two drugs on transcription are different. Rifampin inhibits initiating RNAPs at the promoter and has no effects once RNAP has elongated past the promoter (183). Consistently, rifampin’s mechanism of action based on the crystallographic studies implies a steric clash between the drug and the RNA hybrid when it reaches 2 or 3 nucleotides in length (30). In contrast, streptolydigin can block transcription of initiating as well as elongating RNAPs (114).

Nonetheless, this steric mechanism alone cannot explain why some rifampin-resistant mutations can only confer resistance to certain members of the rifamycin family, without impeding the antibiotic binding to RNAP. Why, for example, can an elevated Mg²⁺ concentration confer a certain Rif resistance phenotype and why is rifabutin, a member of the Rif family, able to inhibit formation of the first phosphodiester bond? From the study of cocrystals of RNAP structures with different members of the Rif family (rifampin and rifabutin), the authors proposed two allosteric mechanisms, the σ pathway, which implicates an interaction between the drug and the σ factor, and the β pathway, which implicates specific residues for transmission of a signal from the Rif binding site to the catalytic aspartates, altering binding of Mg²⁺ in the active site and leading to inhibition of transcription. The σ pathway was evidenced by the finding that rifabutin, but not rifampin, can bind a portion of σ⁵₂, an interaction that could account for its specific effect on the formation of the first phosphodiester bond. The β pathway was defined by the engineering of specific RNAP mutations lying in the signal transmission path, far from the Rif binding pocket, which induce resistance without altering Rif binding (12).

Streptolydigin is an antibiotic that also specifically inhibits bacterial RNAPs (114). Mapping of amino acid replacements conferring streptolydigin resistance onto the RNAP structure, as shown in Fig. 8C, shows clustering in the fork loop 2 region, including A423V/P/T (64, 182), G424R (64, 182), G424D-F425S (107), F425S/C/I (64, 182), D426V (64), and 6His insertion at 420 to 421 (156); the fork domain, including G450C (182), and L451R (182); the bridge helix, including R1078H (156), L1086M/V (182), A1089G (182), S1091F (156, 182), and R1096I; the trigger loop, including P1232L, L1236M, T1237I/N, V1246V, Q1254R, Q1254S P1257A, G1255D, L1256V, P1257T, and P1257R/L/S (182); the funnel domain (M1045I) (156); and a region adjacent to the trigger loop (V1361G) (182).

This location for the streptolydigin binding pocket strongly suggests that the mechanism of action of this drug involves an impaired fork loop 2 function and a restriction of the bridge helix and trigger loop movement, these being required for translocation of nucleic acids in the active site, reminiscent of the α-amanitin mode of action (176). However, some recent results show that streptolydigin blocks neither translocation nor phosphodiester bond formation (182). These findings raise important concerns about the respective implication of the bridge helix/trigger loop movement, these being required for transmission of a signal from the Rif binding site to the catalytic aspartates and forming the first phosphodiester bond. The β pathway was defined by the engineering of specific RNAP mutations lying in the signal transmission path, far from the Rif binding pocket, which induce resistance without altering Rif binding (12).

Sorangicin is a macrolide polyether antibiotic that inhibits transcription initiation in a way very similar to rifampin (72). It has also been observed that some rifampin-resistant mutations can confer sorangicin resistance, suggesting largely overlapping binding sites for these two drugs (31, 127, 145). However, the structures of these drugs are different and observation that some rifampin-resistant mutants are sorangicin sensitive, while most sorangicin-resistant mutations confer rifampin resistance, suggests that subtle differences exist with regard to their modes of action (145).

In order to better understand the basis of similarities and differences between these drugs, crystallographic studies of sorangicin in a complex with T. aquaticus RNAP have been performed which demonstrated that both drugs bind to the same RNAP region (31). The mapping of mutations conferring sorangicin resistance is shown in Fig. 8D. All (31, 127) cluster to the protrusion and fork domains, suggesting that the modes of action of these two drugs are the same, and that the higher conformational flexibility of sorangicin renders it more potent to adapt to structural alterations of its binding pocket.

Microcin J25 is an antibacterial peptide inhibiting transcription by RNAP (149). Recent studies by Adelman and colleagues and Mukhopadhyay and colleagues (1, 117), including
saturation mutagenesis and mapping of mutations conferring microcin J25 resistance, have led to the identification of the microcin J25 binding site made of more than 50 amino acids. These replacements and those isolated in previous reports have been mapped in Fig. 5E. The defined region surrounds the secondary channel which is proposed to be the entry site of NTPs. Thus, based on these studies, the suggested mode of action of microcin J25 consists in the occlusion of the secondary channel, interfering with NTP entry to the active site and leading to inhibition of transcription.

**RNAP MUTANTS MIMICKING THE EFFECTS OF ALARMONE ppGpp**

In nutrient-limiting conditions, such as amino acid starvation, bacterial cells rapidly accumulate a guanine derivative, ppGpp. The production of this metabolite is dependent on the proteins RelA and SpoT. The presence of ppGpp leads to the inhibition of the expression of the translational machinery and to an increase in the synthesis of many enzymes involved in the biosynthesis of amino acids and many other products (33) in a process called the stringent response. RelA and SpoT strains are auxotrophic for many amino acids (200) and this phenotype, due to the absence of ppGpp, can be suppressed by changes in the β, β', and σ subunits of RNAP.

Although recent crystallographic studies of *T. thermophilus* RNAP in a complex with alarmone ppGpp have been resolved (11), the precise mechanism of inhibition and activation of respective promoters by this metabolite is still not completely understood. Studies of promoters like *rrnB P1, rrnD P1,* and *tyrT* have showed that open promoter complexes at these growth rate-regulated promoters are unstable (56, 59, 104). In addition to the short-lived open complexes, other determinants for regulation by ppGpp have been identified in promoters, including a C-rich sequence called discriminator located between the −10 box and +1 site on the nontemplate DNA strand (179), a shortened 16-bp spacer between the −35 and −10 elements, and a noncanonical sequence in the −35 region (14, 15, 33).

Based in part on the crystallographic structures showing the ppGpp molecule bound to RNAP in the pore near the active site, it has been proposed that alarmone could make sequence-specific contacts with the nontemplate strand (11), this interaction being possibly responsible for the reduced stability of open complexes at these promoters.

The mapping of the RNAP replacements mimicking ppGpp effect on the *T. aquaticus* RNAP structure is shown in Fig. 9. A great number of β replacements cluster in the fork domain (14, 178, 201, 202, 207) and some in the protrusion domain, including Q139P (178), P144L (178), R142S (178), and R334H (14, 17). Interestingly, many of them also confer rifampin resistance (178, 201, 207) and some in the protrusion domain, including A490E (178) replacement and the Δ586-601 deletion (178), affects the rudder element. This element is thought to help separate the RNA transcript from DNA by preventing reannealing of the two strands (94). Moreover, the complete deletion of this element (Δ586-601) leads to an altered transcription bubble in the downstream part (94), and thus the Δ587-590 replacement is consistent with the mode of action of ppGpp.

The A490E (178) replacement and the Δ490-495 (17) deletion are located in the clamp head near the downstream part of the DNA in the cleft. It is possible that this region of the clamp helps hold DNA in a fashion similar to the lobe and jaw regions. Thus, effects on the stability of the open complex arise from an altered capacity of RNAP to grip the downstream DNA and maintain a stable transcription bubble. Two other replacements, E1264D (178) and R1266C (14), are located in the cleft region and may act similarly to replacements in the lobe and clamp head regions because this part of the cleft also participates in holding the downstream DNA.

Finally, two replacements are located in the bridge helix. This element has been proposed to be involved in translocation of nucleic acids and to maintain a specific conformation in the template strand in the active site. Because this element interacts with nucleic acids, replacements could easily destabilize the open complex. The mutated residue R1078C (14) is oriented toward the fork loop 2 element and may affect open promoter stability by disrupting an important interaction between this element and the bridge helix. The other replacement in the bridge helix, R1087Q (178), is oriented toward the nucleic acids in the prokaryotic structure, but in the opposite direction, toward the trigger loop, in eukaryotic structures. This could be attributable to conformational changes between the bent and straight conformations observed in these respective structures. Because this replacement is near the point of passage of nucleic acids over the bridge helix, replacement to glutamine could interfere with the movement of nucleic acids in the cleft. All the replacements discussed in this section support a destabilizing effect of ppGpp on open promoter complexes.
This section focuses on mutations affecting the correct assembly of RNAPs. Many of the assembly mutations most probably alter the correct folding of the subunit either locally or globally, rather than abolish specific key interactions that disrupt the contacts between subunits.

Many zinc atoms are known to bind and stabilize the RNAP structures (44). The eukaryotic mutation T69N (192), isolated in the largest subunit of RNAP II by random mutagenesis, and temperature-sensitive mutations G166S (89, 150) and P172S (89, 150), isolated by random hydroxylamine mutagenesis, are located near or in Zn2+/H11001 binding regions and show altered RNAP composition after chromatography (Fig. 10).

The RNAP I temperature-sensitive mutations H80Y (197, 203) and G82D (197, 203), isolated by hydroxylamine mutagenesis, are located near or in Zn2+/H11001 binding regions and show altered RNAP composition after chromatography (Fig. 10).

Finally, the temperature-sensitive prokaryotic β' subunit mutations R1189H (122) and G1206D (122) also map to a Zn2+ region and these subunits fail to reassemble after denaturation and renaturation. This occurrence of assembly defective mutations in the Zn2+ binding regions suggests an important contribution of these regions and of zinc to the correct folding of subunits and the assembly of RNAP.

The next group of replacements localize in regions more or less distant from the interface of subunit-subunit interactions. These replacements most probably affect RNAP assembly by altering either globally or locally the folding of the given subunit.

![Fig. 9. Distribution of substitutions and deletions mimicking the alarmone ppGpp effect on transcription mapped onto the T. aquaticus RNAP structure 1H6V. ppGpp has been placed by overlaying the T. thermophilus RNAP structure 1SMY. Substitutions located in the lobe and protrusion domain are shown in orange and pink, respectively. The many substitutions in the fork domain are shown in red and are not labeled. Substitutions in the cleft, switches, and jaw domains are shown in magenta, green, and blue, respectively. Cyan-colored altered amino acids are located in the clamp core and clamp head region.](http://mmbr.asm.org/)
temperature-sensitive eukaryotic Rpb1 mutation A301D (169) is located in a helix of the clamp domain. An assembly defect has been proposed based on the altered cellular distribution of this Rpb1 mutant, which localizes in the cytoplasm instead of the nucleus at the nonpermissive temperature. The replacement could disturb the correct folding of the clamp domain known to interact with the C-terminal part of Rpb2.

In the second largest subunit of eukaryotic RNAP, the R857K (89, 150) replacement maps to the wall domain. The replacement of R857 with a lysine may also disturb the folding of this domain bound by Rpb3, Rpb10, and Rpb12. Finally, replacements in the largest Rpb1 subunit, I1327N (89, 150) and C1240Y (89, 150), affect residues located in regions interacting with other subunits, Rpb5 and Rpb9, respectively. This observation suggests an alteration of the interaction domains of Rpb1 with these subunits. However, these mutants were isolated based upon their ability to assemble stably with Rpb3, which is quite distant from these Rpb1 regions interacting with

**FIG. 10.** (A) Distribution of substitutions and insertions affecting assembly of eukaryotic RNAP mapped onto the *S. cerevisiae* RNAP II structure. (B) Distribution of substitutions and insertions affecting assembly of prokaryotic RNAP mapped onto the *T. aquaticus* structure.
Rpb5 and Rpb9. In the crystal structure, Rpb3 interacts mostly with Rpb2 and Rpb11. It is then more probable that these replacements alter the stability and/or folding of Rpb1 which no longer associates stably with Rpb3, consequently altering interaction with either Rpb2 or Rpb11.

The W954(LELE)P insertion (3, 5, 6, 51) is positioned in the foot domain of Rpb1 between the interaction domains of TFIIS and Rpb6. The slow-growth, temperature-sensitive, and inositol auxotrophy phenotypes exhibited by this mutant can be suppressed by overexpression of Rpb6, suggesting that the observed defects arise from an altered ability of W954(LELE)P to bind Rpb6. Observed in the crystal structure, the W954(LELE)P insertion in the foot domain of Rpb1 is located too far away from Rpb6 to interfere directly with its association. Thus, this insertion may lead to instability of this Rpb1 region, which can be overcome by formation of the complete RNAP complex, this later step being favored by overexpression of Rpb6.

Finally, the prokaryotic β’ replacement G457E (122) is located in the N-terminal part of a helix of the clamp and replacement for a glutamate may alter the correct folding in this region and, consequently, its association with the β subunit.

The final set of mutants discussed here come from deletion or insertion mutagenesis. Even though large deletions of dispensable regions in the β subunit have been reported (155), these mutations are generally more prone to severely affect the folding of the subunits, given the complexity of the RNAP structure. A number of XhoI linker insertions have been identified (96) in the β subunit (Xho-42, 25-RDS-25; Xho-5, 178-SEV-184; Xho-6, 38-LEA-47; Xho-16, 214-LEE-220; Xho-29, 786-LEK-786; Xho-31, 83-HSSSSSVV-101; Xho-1, 309-TREK-317; Xho-4, 166-SEP-178; Xho-19, 756-LEK-762; Xho-36, 15-LEF-48; Xho-54, 906-LEK-910; and Xho-73, 163-HSS-RVM-242) (Fig. 10). Assembly defects have been confirmed by measuring the amount of rifampin-sensitive (Rif') RNAP activity arising from association of these Rif' mutant β subunits into RNAP when overexpressed in a strain carrying a rifampin-resistant chromosomal copy of β. A number of these mutants show detectable (Xho-42, Xho-5, Xho-6, Xho-29, and Xho-31), inhibitory (Xho-1, Xho-4, and Xho-19) or no catalytic activity (Xho-16, Xho-36, Xho-54, and Xho-73). These insertions most probably induce more or less severe misfolding of the respective subunits or domains.

CONCLUSIONS AND PERSPECTIVES

Mapping the amino acid changes reported to affect the function of multisubunit RNAP on the structure of the prokaryotic and eukaryotic enzymes revealed interesting features on the role of many structural elements during the transcription reaction. Many of these changes have not been fully characterized and knowing their position within the structure now allows us to make novel predictions that can be tested experimentally. For example, fork loop 2, which, based on its location, was proposed to be involved in setting the downstream boundary of the transcription bubble, is the site of many replacements that negatively regulate the polymerization rate. It would be interesting to further characterize these mutant RNAPs in vitro in order to determine whether or not they are altered in their ability to form and maintain a transcription bubble; these experiments would either confirm the implication of fork loop 2 in promoter opening or reveal another role for this element in the transcription reaction.

The trigger loop is targeted in a number of mutants affected in initiation site selection; these mutations could impair either direct contact of nucleic acids around +1 or the movement of the bridge helix. A functional analysis of mutants with an altered trigger loop would be required to discriminate between these two possibilities. Yet another example is the funnel domain, which is the site of many replacements that increase the polymerization rate or suppress polymerization defects. In vitro characterization of amino acid changes in this domain should allow us to clarify the role of the funnel in transcription.

It is important to note that a significant number of mutations listed in the catalogue (see Table S2 in the supplemental material) have not been described in this article (2, 45, 46, 49, 66, 67, 73, 78, 80, 84, 88, 91, 97, 108, 111, 112, 124, 126, 129, 130, 146, 151, 157, 160, 165, 167, 168, 170, 177, 180, 181, 186, 193, 204). For most of them, only in vivo phenotypes have been described and no in vitro characterization has yet been performed. Other mutations listed in Table S2 in the supplemental material that have not been discussed in this article are those whose effect on the enzyme has been characterized in vitro but whose location in the RNAP structure does not suggest a mechanistic basis for their effect. Some of these replacements may act indirectly by affecting known functional elements of the enzyme or may define new functional regions. As a whole, all these mutations are of great interest and have been included in our table in the supplemental material. Their positions within the structure and the domain in which they lie are indicated.

We hope that the structural perspective on mutations affecting the function of multisubunit RNA polymerases provided herein will stimulate the design of new mutations to be characterized and provide novel information on transcriptional mechanisms.

ACKNOWLEDGMENTS

We apologize to our colleagues whose work on RNAP could not be incorporated in this review due to space limitation. We are grateful to the members of our laboratory for helpful discussions and comments on this paper. We thank Diane Bourque for artwork and Julie Edwards for critical reading of the manuscript.

This work is supported by grants from the Canadian Institutes for Health Research (to B.C. and J.A.), Genome Canada (to B.C.), and Genome Québec (to B.C.). M.F.L. holds a studentship from the Fonds de la Recherche en Santé du Québec. J.A. is a senior scholar from the Fonds de la Recherche en Santé du Québec.

REFERENCES


25. Fish, R. N., and C. M. Kane. 2002. Promoting elongation with transcript cayacation

26. Fricke, S., and I. Fricke. 1997. Photo-cross-linking of a purified preinitiation complex reveals central roles...


