Diffusion of nucleoside triphosphates and role of the entry site to the RNA polymerase II active center

Nizar N. Batada, Kenneth D. Westover, David A. Bushnell, Michael Levitt, and Roger D. Kornberg*

Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305-5126

Contributed by Roger D. Kornberg, November 4, 2004

Nucleoside triphosphates (NTPs) diffuse to the active center of RNA polymerase II through a funnel-shaped opening that narrows to a negatively charged pore. Computer simulation shows that the funnel and pore reduce the rate of diffusion by a factor of $\approx 2 \times 10^{-7}$. The resulting limitation on the rate of RNA synthesis under conditions of low NTP concentration may be overcome by NTP binding to an entry site adjacent to the active center. Binding to the entry site greatly enhances the lifetime of an NTP in the active center region, and it prevents “backtracking” and the consequent occlusion of the active site.

computation | electrostatics | transcription

The RNA polymerases responsible for cellular gene transcription are large multiprotein complexes. Structure determination has revealed a core devoted to nucleic acid polymerization and an elaborate superstructure for the recognition of diverse promoter sequences and also for regulation (1–3). The proportion of the structure devoted to these different tasks may be gauged by comparison with the single-subunit viral RNA polymerases (4, 5), which are approximately one-fifth of the mass of the multisubunit enzymes and transcribe only one or a small number of promoters, with rudimentary requirements for control.

Among the many structural accoutrements of the multisubunit enzymes, one that has attracted recent attention is an undercarriage that serves an important purpose in transcript elongation but that also imposes an apparent limitation on the enzyme mechanism. This undercarriage harbors a funnel-shaped opening, narrowing to a pore beneath the active center (Fig. 1) through which the 3' end of the RNA is extruded during “backtracking” (retrograde motion of the enzyme on the DNA template) and through which substrate nucleoside triphosphates (NTPs) must enter for addition to the 3' end during RNA synthesis. Backtracking is crucial for recovery from pausing, for exposing DNA damage for repair, and possibly for proofreading the nascent transcript (reviewed in ref. 6). Pausing is of particular interest for its regulatory role. Recovery from pausing generally requires transcript cleavage in the polymerase active center, induced by the general transcription factors SII (TFIIS) in eukaryotes and GreA/B in bacteria. Defects in pausing and recovery underlie human diseases such as Cockayne syndrome, Wolf–Hirschhorn syndrome, and von Hippel–Lindau cancer-predisposition syndrome.

The polymerase funnel and pore facilitate pausing and recovery. A backtracked transcript is sequestered, possibly with its 3' end bound to a site of nucleotide interaction in the funnel at approximately nine residues from the active center (7). The transcript is protected from nucleases, accessible only to TFIIS or GreA/B, which bind to the surface of the funnel and possess a long, slender protrusion that is capable of reaching up through the pore to the active center (8, 9). The transcript and protrusion can occupy the pore and funnel simultaneously. The protrusion is thought to supply a Mg ion that is needed for transcript cleavage at the polymerase active center (10).

While protecting the nascent transcript and controlling its cleavage, the funnel and pore unavoidably restrict diffusion of

NTPs to the active center. In this study, we have investigated the nature of this restriction by computer simulation, taking into account all relevant factors of funnel and pore topography, electrostatics, and NTP interaction. We have explored the effect of the recently described entry site (E site) (7, 11) from which NTPs are delivered to the nucleotide addition site (A site) at the 3' end of the transcript (Fig. 1). All NTPs bind to the E site, whereas only an NTP able to base pair with the DNA template can bind the A site. We have asked whether binding at the E site, which has no base specificity and a much lower affinity for a matched NTP than the A site, can nevertheless enhance the rate of transcription.

Materials and Methods

Computation of Electrostatic Potential. The electrostatic potential was computed for the vicinity of RNA polymerase II (pol II) with a 1-Å grid spacing by using APBS software (12), which solves the Poisson–Boltzmann equation by the multilevel finite-element method. For this calculation, an ionic strength of 150 mM, a solvent dielectric constant of 78.54, and a protein dielectric constant of 1 were used.

Simulation Parameters. The position of the diffusing NTP evolves according to the following Langevin stochastic differential equation (13):

$$dX(t) = -\frac{1}{\gamma} \frac{\partial}{\partial x} \phi(x, y, z) dt + \sqrt{2D} dW_x(t) \quad [1]$$

$$dY(t) = -\frac{1}{\gamma} \frac{\partial}{\partial y} \phi(x, y, z) dt + \sqrt{2D} dW_y(t) \quad [2]$$

$$dZ(t) = -\frac{1}{\gamma} \frac{\partial}{\partial z} \phi(x, y, z) dt + \sqrt{2D} dW_z(t), \quad [3]$$

where $W_x(t)$ represents a one-dimensional standard Brownian motion, $\phi(x, y, z)$ is the electrostatic potential, D is the diffusion coefficient, and $\gamma = kT/D$ (where k is the Boltzmann constant and T is the temperature). For computation of trajectories, a time step of 50 ps and a diffusion coefficient of $75 \mu\text{m}^2/\text{s}$ were used, and simulations were performed for 0.1 ms. Because the potential was computed on a grid size of 1 Å, polynomial interpolation was used to determine the potential at nonlattice points.

Results

The computer simulation of NTP diffusion performed in this study was realistic, taking into account the protein structure and aqueous solution conditions. The surface topography of the funnel and pore were derived from the 2.3-Å structure of yeast pol II (7). The electrostatic potential within the funnel and pore

Abbreviations: A site, addition site; E site, entry site; NTP, nucleoside triphosphate; pol II, RNA polymerase II.

*To whom correspondence should be addressed. E-mail: kornberg@stanford.edu.

© 2004 by The National Academy of Sciences of the USA

base pairing also come into play. NTP binding could be an order of magnitude more rapid at the E site than at the A site, for this reason.

The second advantage of NTP binding at the E site relates to backtracking. Pol II is believed to oscillate between forward translocation and backtracking (reverse translocation) at every step of transcription. Forward translocation clears the A site for entry of the next NTP. Backtracking returns the nucleotide that was just added to the RNA back to the A site. One or more additional steps of backtracking result in a stalled transcription complex, which may persist for many minutes or even be irreversible. As mentioned above, regulatory factors may intervene to restart transcription, or the polymerase may be ubiquitinated and destroyed, to avoid gene inactivation and cell death.

Therefore, maintenance of the transcription rate requires the entry of an NTP in the A site immediately after forward translocation and before backtracking can occur and lead to stalling. Entry of an NTP in the E site will suffice for this purpose because the E site overlaps the A site (7). NTP binding at the E site increases the barrier height for backtracking. The strong negative electrostatic potential in the vicinity of the E site may increase the barrier height as well. Thus, the E site and its associated features may be crucial for the overall purpose of the polymerase undercarriage in the pausing and restoration of transcription.

K.D.W. was supported by the Medical Scientist Training Program. This work was supported by National Institutes of Health Grants GM49985 (to R.D.K.) and GM63817 (to M.L.).

1. Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K. & Darst, S. A. (1999) *Cell* **98**, 811–824.
2. Cramer, P., Bushnell, D. A., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., Burgess, R. R., Edwards, A. M., David, P. R. & Kornberg, R. D. (2000) *Science* **288**, 640–649.
3. Bushnell, D. A., Westover, K. D., Davis, R. E. & Kornberg, R. D. (2004) *Science* **303**, 983–988.
4. Temiakov, D., Patlan, V., Anikin, M., McAllister, W. T., Yokoyama, S. & Vassylyev, D. G. (2004) *Cell* **116**, 381–391.
5. Yin, Y. W. & Steitz, T. A. (2004) *Cell* **116**, 393–404.
6. Shilatifard, A., Conaway, R. C. & Conaway, J. W. (2003) *Annu. Rev. Biochem.* **72**, 693–715.
7. Westover, K. D., Bushnell, D. A. & Kornberg, R. D. (2004) *Cell* **120**, 481–489.
8. Opalka, N., Chlenov, M., Chacon, P., Rice, W. J., Wriggers, W. & Darst, S. A. (2003) *Cell* **114**, 335–345.
9. Kettenberger, H., Armache, K. J. & Cramer, P. (2003) *Cell* **114**, 347–357.
10. Sosunova, E., Sosunov, V., Kozlov, M., Nikiforov, V., Goldfarb, A. & Mustaev, A. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 15469–15474.
11. Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V. & Goldfarb, A. (2003) *EMBO J.* **22**, 2234–2244.
12. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 10037–10041.
13. Gardiner, C. W. (1997) *Handbook of Stochastic Methods: For Physics, Chemistry, and the Natural Sciences* (Springer, Berlin).
14. Bochner, B. R. & Ames, B. N. (1982) *J. Biol. Chem.* **257**, 9759–9769.
15. Wang, M. D., Schnitzer, M. J., Yin, H., Landick, R., Gelles, J. & Block, S. M. (1998) *Science* **282**, 902–907.
16. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. (2001) *Science* **292**, 1876–1882.