Biophysical Chemistry, 37 (1990) 239–250
Elsevier

BIOCHE 01481

Thermodynamic analysis of the transcription cycle in *E. coli*

Stanley C. Gill [+], Thomas D. Yager [*] and Peter H. von Hippel

Institute of Molecular Biology and Department of Chemistry, University of Oregon, Eugene, OR 97403, U.S.A.

Received 12 February 1990
Accepted 28 February 1990

Transcription cycle; RNA polymerase interaction; Thermodynamics; Sigma factor; Protein, NuA; Transcription regulation

The *E. coli* RNA transcription cycle can be divided into three major phases, which are generally called initiation, elongation, and termination. In this paper, we review recent biophysical studies of the interactions of the transcriptional regulatory proteins, sigma^{70} and NusA, with themselves and with core RNA polymerase in solution, as well as with core polymerase within the transcription complex. The different affinities of sigma^{70} and NusA for core RNA polymerase at various stages in the transcription cycle, together with other quantitative data, are then used to construct a partial free energy diagram for the overall transcription process. This thermodynamic framework, which is interrupted by at least two irreversible steps, can be used to rationalize physiological aspects of the transcription cycle and its regulation, as well as to identify crucial points at which our knowledge is still incomplete.

1. Introduction

It is a pleasure to contribute a paper to this issue of Biophysical Chemistry, which honors the many contributions to chemistry and biology of Professor Jeffries Wyman. While this paper is not directly in Professor Wyman's present area of research, we feel that it may be appropriate, since the analysis we present here owes much to the ideas of linkage and allostery that he initially formulated and applied to various problems of molecular biology. It may also be fitting that aspects of this work were presented (by both S.C.G. and P.H.vH.) at a recent Symposium held at the University of Colorado in honor of Stanley J. Gill, who has been Jeffries Wyman's major collaborator in recent years.

Correspondence address: P.H. von Hippel, Institute of Molecular Biology and Department of Chemistry, University of Oregon, Eugene, OR 97403, U.S.A.

[+] Present address: Department of Chemistry, University of Colorado, Boulder, CO 80309, U.S.A.

[*] Present address: Department of Biophysics, King's College, University of London, 26-29 Drury Lane, London WC2B 5RL, U.K.

2. Analysis

2.1. The transcription cycle

The process of transcription is conveniently divided into three phases, called initiation, elongation and termination (for fairly recent mechanistic reviews, see refs 1–3). It is useful to think of transcription as a cycle of reactions through which the RNA polymerase passes in carrying out the various stages of transcript synthesis. A schematic representation of this cycle is shown as the inner circle in fig. 1. However, it is also important to realize that a polymerase molecule in its free solution conformation does not simply 'enter' the template DNA at a specific promoter and, after transcribing the gene, 'exit' unchanged at the terminator. Rather, the process involves changes in the polymerase itself as it moves through the cycle, and also, as the 'loops' drawn around the central

cycle in fig. 1 suggest, the participation of protein regulatory factors that bind to (and alter) the ternary complex of core polymerase, template DNA, and nascent RNA during specific phases of the transcription process.

In this paper, we shall consider only two of the many protein factors that can participate in the loops that regulate transcription of the *E. coli* genome in vivo. These proteins are the sigma specificity factor, which is an essential participant in initiation, and the NusA protein, which can serve both as a quantitative modulator of elongation and termination and as an 'anchor' protein on core RNA polymerase for the binding of a variety of additional regulatory factors. This overall sigma-NusA regulatory cycle, which we present here in quantitative terms, was initially formulated by Greenblatt and Li [4].

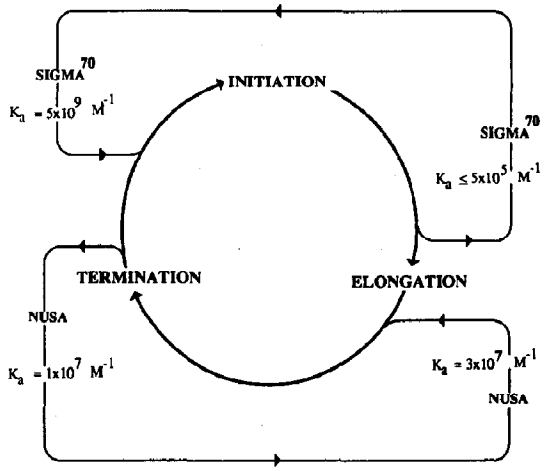

Fig. 1. Schematic representation of the overall transcription cycle, including the binding affinities of the sigma^{70} and NusA regulatory factors for core polymerase at various stages within the ternary complex.

2.2. Transcript initiation

Initiation of synthesis of a specific *E. coli* RNA transcript proceeds through a series of steps. First, the RNA polymerase holoenzyme is formed by association of the free core enzyme with a particular sigma factor. The holoenzyme is then able to recognize a given class of promoters and thus introduces specificity into the transcription process. Next, the holoenzyme must locate and bind to the sigma-specified promoter, forming a 'closed promoter complex'. Promoter location is a multistep process that may involve one-dimensional 'sliding' over regions of nonspecific DNA, as well as three dimensional diffusion (see ref. 5).

A local 'melting' of approx. 12 base-pairs of promoter DNA, coupled with a conformational change in the holoenzyme, then results in the formation of an 'open promoter complex' [6,7]. RNA synthesis is initiated at a specific template site (position +1) that marks the start of the coding region of the gene to be transcribed. Successive template-specified nucleotide residues are then added to the nascent RNA transcript by the holoenzyme. During these initial steps of synthesis the ternary initiation complex is quite unstable, and the nascent transcript may be released in a process called 'abortive initiation'. If this occurs the holoenzyme remains bound to the promoter and moves back to the +1 position on the template to initiate another RNA chain.

When the growing RNA chain reaches a (promoter-dependent) length of 8–12 nucleotide residues, the holoenzyme loses contact with the −10 and −35 regions of the promoter and releases the sigma subunit into solution [8–11]. This marks the end of the initiation phase of the transcription cycle, and the end of the sigma portion of the regulatory factor loop (see fig. 1).

2.3. Transcript elongation and termination

From this point on, the transcriptional complex exists in the 'elongation' mode, and continued RNA synthesis is catalyzed by the core enzyme. In this phase the ternary complex is very stable, and synthesis continues to the end of the gene in a highly processive manner.

There now exists a structural model of the elongation complex that includes interactions between its protein and nucleic acid components. This model is based on the studies of many workers [12–16], and has been formulated in detail in refs 3 and 17. The model contains two major features. These are an unpaired region (or 'bubble') in the DNA, consisting of approx. 18 'open' base-pairs, and an RNA-DNA hybrid, located within the

DNA bubble, that is approx. 12 base-pairs in length and consists of template DNA plus the 12 nucleotide residues at the 3′ end of the nascent RNA chain. This RNA-DNA hybrid is positioned so that the 3′-OH terminus of the nascent RNA is coincident with the first melted base-pair of the DNA bubble. The geometry of the elongation complex is thought to be approximately invariant as the polymerase moves along a DNA template during transcription. The precise definition of this model allows one to calculate the thermodynamic stability of the elongation complex, position-by-position, along the DNA template (see ref. 17).

Recent analysis [17,18] has suggested that the elongation complex is quite stable at most positions on the DNA template. We have proposed that the complex is 'trapped' within a deep potential energy well that (at non-termination positions in the DNA) is characterized by a free energy of activation barrier to RNA release of approx. 36 kcal/mol [18]. This large barrier accounts for the high degree of processivity observed for the transcript elongation reaction and indicates the magnitude of the free energy change needed to bring about termination.

Recent studies [19–22] show that, within the stable elongation mode, polymerization occurs in kinetic competition with an opposing pyrophosphorolysis. Thus, the ternary complex located at template position i can move either forward to position $i+1$ with the insertion of the next ribonucleotide specified by the template, or backward to position $i-1$ by a pyrophosphorolysis reaction that corresponds to the reverse of the chain extension process. However, as indicated above, the elongation complex is kinetically blocked from dissociating on this time scale. Within this stable context elongation is a thermodyamically and kinetically discontinuous synthesis process, characterized by sequence-defined 'pausing' events that may be important for regulation and may also be modulated by extrinsic protein factors.

Termination of transcription occurs at specific sites on the DNA template, either with or without the participation of additional protein factors. At these sites the standard free energy barrier to dissociation of the transcription complex appears to be significantly and abruptly lowered (often by 10–15 kcal/mol over one to two template positions; see ref. 17), and RNA release becomes thermodynamically (and kinetically) favored over further elongation. As a consequence, the nascent RNA chain is released from the polymerase and from the template DNA. This results in a free RNA transcript, a 'reclosed' DNA bubble, and a core RNA polymerase molecule that is either free in solution or nonspecifically bound to the DNA near the termination site. The core polymerase can then pick up a new sigma factor and the transcription cycle is ready to begin again.

2.4. *Physiological specificity*

A complete transcription cycle, including the sigma regulatory 'loop', is outlined in fig. 1. This system (core RNA polymerase, a sigma factor, and DNA containing the sigma-specified promoter, together with the necessary ribonucleotide triphosphates) forms the minimal transcription system that can operate with physiological specificity in vitro. This system can initiate properly, elongate properly, and terminate properly at factor-independent ('intrinsic') termination sites. With added rho protein (and *E. coli* RNA release factor) it can also terminate properly at rho-dependent termination sites.

2.5. *Regulatory factors*

In recent years a variety of protein factors have been shown to interact with the elongation complex and to influence elongation, pausing and termination (for a recent review, see ref. 3). Of these the NusA protein has been shown to play a central role, either alone or through interactions with other protein factors. A unifying interpretation is that NusA binds to the core polymerase, thus influencing elongation and termination as well as serving as an 'anchor' protein that allows other factors to bind and thus exert their specific effects.

As the lower loop of fig. 1 shows, NusA protein enters the transcription cycle and interacts with core RNA polymerase shortly after the point of sigma dissociation. NusA then remains bound un-

til termination, at which point it dissociates and sigma rebinds, thus completing the sigma-NusA regulatory cycle [4,23,24]. Unlike sigma, however, NusA protein is not necessary for the specific operation of the minimal transcription system in vitro, though it appears to be essential in vivo and may be needed to define physiological levels of termination efficiency in many operons [25–27].

3. Stoichiometry and thermodynamics of the interaction of sigma^{70} and NusA with RNA polymerase in the transcription cycle

3.1. Lack of self-association of sigma^{70} and NusA proteins

Prior to initiating a study of the interactions of sigma factor and NusA protein with core polymerase, either free in solution or as part of the ternary transcription complex, it was necessary to characterize the self-association states of both proteins. If oligomerization of sigma or NusA were to occur, these additional equilibria would alter the equilibrium state of the transcriptional system. * These studies have shown that both NusA and sigma^{70} exist as homogeneous monomers over a wide range of protein concentrations and salt conditions [28]. Earlier work [29] had shown that both core RNA polymerase and the sigma^{70}-containing holoenzyme also exist as monomers at salt concentrations in excess of 0.25 M KCl.

3.2. Binding affinities of sigma^{70} and NusA proteins to core polymerase in solution

Next the binding interactions of sigma^{70} and NusA with core polymerase were studied in the absence of DNA, under conditions where all three proteins exist separately as monomers in solution. Binding experiments were performed by monitoring changes in the fluorescence intensity or anisotropy of fluorescently labelled NusA or sigma^{70} upon titration with core polymerase. These titrations were carried out directly, and also as competitive binding reactions in which the labelled sigma^{70} or NusA were displaced by unlabelled NusA or sigma^{70}, respectively. Additional binding studies were performed by analyzing mixtures of NusA, sigma^{70}, and core polymerase, using HPLC gel filtration procedures.

These studies [23] have shown that NusA and sigma^{70} bind competitively to core polymerase in 0.25 M KCl, to form either R : σ^{70} or R : NusA complexes *. Binding stoichiometries (one-to-one) and association constants are presented in eqs 1 and 2:

$$\mathrm{R} + \sigma^{70} \rightleftharpoons \mathrm{R} : \sigma^{70} \qquad K_{\mathrm{a,R}:\sigma^{70}} = 5 \times 10^{9}\ \mathrm{M}^{-1} \tag{1}$$

$$\mathrm{R} + \mathrm{NusA} \rightleftharpoons \mathrm{R} : \mathrm{NusA} \qquad K_{\mathrm{a,R}:\mathrm{NusA}} = 1 \times 10^{7}\ \mathrm{M}^{-1} \tag{2}$$

We conclude that sigma^{70} binds to free core polymerase with a 500-fold higher association constant than does NusA.

3.3. The affinities of sigma^{70} and NusA protein for core polymerase within the initiation complex

We next determined that a 300-fold molar excess of NusA over sigma^{70} has no effect on the binding of holoenzyme to a sigma^{70}-specific promoter. This suggests that sigma^{70} binds to the open promoter complex with an affinity that is at least 10^4 greater than that of NusA for this complex. ** It is clear from these results that sigma^{70} can easily displace NusA from the core polymerase during the initiation phase of transcription. Thus a high concentration of NusA protein

* All of our thermodynamic studies have been carried out using the major sigma factor of *E. coli*, sigma^{70}. (The superscript indicates the approximate molecular mass of the sigma protomer in kDa.)

* Here R and σ^{70} correspond to core RNA polymerase and sigma^{70}, respectively. The same abbreviations apply in fig. 2, where we also use $\mathrm{DNA}_{\mathrm{nsp}}$ for nonspecific DNA, $\mathrm{P_c}$ for the closed promoter, $\mathrm{P_o}$ for the open promoter, RNP′ for a second core polymerase conformation, rNTPs for ribonucleotide triphosphates, E for DNA in the elongation complex mode, and T for DNA in the termination complex mode.

** See footnote on p. 243.

should present no impediment to specific transcript initiation.

3.4. The affinities of sigma^{70} and NusA proteins for core polymerase within the elongation complex

We have utilized several NusA-specific activity assays to measure an apparent K_a for the binding interaction between core polymerase and NusA within the elongation complex [24]. This has allowed us to determine how the binding affinities of NusA and sigma^{70} for core polymerase change as a consequence of forming the elongation complex. Results from four independent assays, performed in our laboratory and in others [30–34] are in good agreement and yield an apparent binding constant of approx. 3×10^7 M^{-1}, assuming a 1 : 1 binding stoichiometry for NusA to core polymerase in this phase of transcription (this stoichiometry is consistent with our results in free solution, and the data of ref. 35).

We note that the apparent NusA binding constant is an average value that has been determined at different pause sites, and with both intrinsic and factor-dependent terminators located in natural DNA templates [24]. The differences in the values of this parameter measured at these various sites are quite small, and thus we conclude that the apparent binding affinity of NusA for core polymerase within the ternary complex is likely to be constant over the various phases of transcript elongation (and termination).

** It is important to note that this difference in apparent binding affinities could, in part, have a kinetic rather than an equilibrium basis, since the rate of dissociation of holoenzyme from the open promoter complex is quite slow. However this apparent affinity difference is also perfectly compatible with an equilibrium effect, since the binding constant of the holoenzyme to the open promoter complex is of the order of $1–5 \times 10^{11}$ M^{-1} [40], which requires that sigma^{70} exhibit a similar or greater binding affinity for the core polymerase-DNA complex in the open promoter than it does for free core polymerase. An additional assumption that underlies this statement is that the binding constant of NusA to core polymerase in the open promoter complex (or to core polymerase nonspecifically bound to DNA) is the same as that of NusA to core polymerase free in solution, since these equilibria involving DNA cannot be measured directly.

The same approach (based on NusA activity assays) was used to measure the binding affinity for sigma^{70} to the elongation complex. Here we have looked for an inhibition of NusA activity in the presence of sigma^{70}. We find that sigma^{70} is incapable of competing effectively with NusA at as much as a 50-fold molar excess of sigma^{70} over NusA. Assuming that NusA activity would be inhibited if sigma^{70} were to displace NusA from the elongation complex, we estimate the binding affinity between sigma^{70} and the elongation complex to be less than or equal to 5×10^5 M^{-1}. This result is in good agreement with other studies that have shown that sigma^{70} does not affect elongation or termination events or copurify with isolated elongation complexes [35–37].

3.5. Two polymerase conformations?

The equilibrium constants we have determined for the binding of sigma^{70} and NusA to core polymerase during the transcription cycle are summarized in fig. 1. These values also reflect the fact that there may be as few as two discrete binding conformations of RNA polymerase over the entire transcription cycle ($R : \sigma^{70}$ during initiation, and $R' :$ NusA in the elongation mode), at least to the extent that these conformations are reflected in different affinities of the polymerase within the ternary complex for sigma^{70} and NusA.

In thermodynamic terms the conformational difference between these states manifests itself as a large change in the binding affinity of sigma^{70} for core polymerase, with little or no change (over the entire cycle) in the affinity of core polymerase for NusA. This difference permits core polymerase, in the presence of molar excesses of both proteins, to bind sigma^{70} in the initiation phase and NusA in the elongation and termination phases. This result is also compatible with both specific (sigma-directed) transcript initiation and the subsequent modulation of elongation and termination by NusA.

4. A free energy map of the transcription cycle

We have used quantitative information from the literature to construct a preliminary thermody-

namic model of the transcription cycle. This model certainly is not complete, and does not adequately address important kinetic features of the cycle. However, it does provide information on the thermodynamic limits within which transcription must occur, and may help us to determine which steps are governed by simple thermodynamic equilibria and which by irreversible kinetic processes.

This thermodynamic model presupposes a more detailed 'molecular' view of the transcription cycle than is given in fig. 1. Such a view is presented in fig. 2. To simplify fig. 2 we have omitted the NusA loop.

4.1. Two irreversible steps

On the time scale of a normal transcription reaction, there appear to be two irreversible steps in the cycle represented in fig. 2. The first occurs as polymerase moves out of the abortive initiation phase and into the elongation phase. However, the irreversibility of this step is difficult to demonstrate experimentally, since (in the presence of rNTPs) additional RNA synthesis occurs immediately after sigma release. This effectively removes the RNA polymerase from the promoter and prevents a possible back reaction from taking place.

The second irreversible step occurs during termination, which has, of course, been shown experimentally to be irreversible overall. Here, we imagine (see ref. 17) that after the nucleic acid and protein rearrangements characterizing termination have occurred, but before the components have diffused apart, the ternary complex may exist as a transient species containing the core RNA polymerase in its solution conformation, together with the 'reclosed' DNA and the nascent RNA transcript. To re-enter the elongation mode at this point would require reopening the DNA bubble, re-forming the correctly aligned DNA-RNA hybrid, and re-establishing the termination confor-

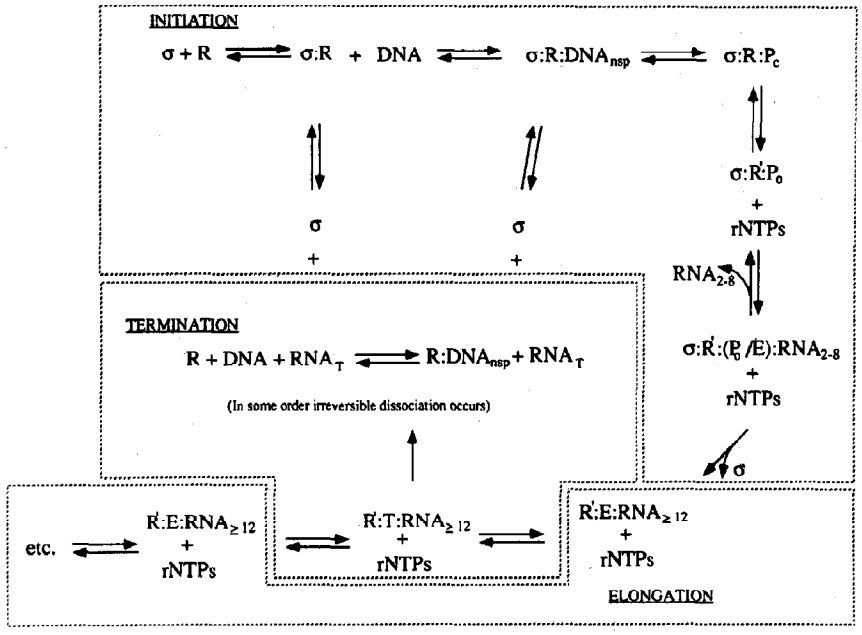

Fig. 2. Molecular representation of the overall transcription cycle. R, core RNA polymerase; σ^{70}, sigma^{70}; DNA$_{nsp}$, nonspecific DNA; P$_c$, closed promoter; P$_o$, open promoter; R′, a second core polymerase conformation; rNTPs, ribonucleotide triphosphates; E, DNA in the elongation complex mode; T, DNA in the termination complex mode. The cycle begins in the upper left-hand corner and proceeds clockwise, and the boundaries of the three phases of the overall cycle are indicated with dashed boxes. See also text and footnote in section 3.2.

mation of the core polymerase. These are all very improbable events.

4.2. Core polymerase undergoes at least two (possibly three) conformational changes in the transcription cycle

Many conformational rearrangements of the RNA polymerase molecule have been proposed to explain various functional and regulator aspects of transcription. However, experimental evidence has been obtained for only three specific conformational changes.

(i) The first documented conformational change occurs during the transition from the closed to open promoter complex [6,7]. This conformational change entails a shift in protease sensitivity, as well as a change in the pattern of nuclease protection that the polymerase confers on the DNA.

(ii) The second documented change may occur upon entry into the elongation mode as the holoenzyme loses contact with the open promoter and releases the sigma factor. During this process several physical and structural properties of the polymerase-DNA complex are altered (reviewed in ref. 3). These changes have been attributed to a conformational change in polymerase. An alternative view (cited above) is that the observed changes may reflect the growth of the DNA bubble and RNA-DNA hybrid to their mature sizes, and that the release of sigma^{70} is due to the movement of polymerase away from the promoter region.

In this alternative model we assume that the conformational change that occurs upon the closed-to-open promoter transition is responsible for the decreased binding affinity observed between polymerase and sigma^{70} in the elongation phase. We must then also assume that sigma^{70} does not dissociate from the open promoter complex because of the combined binding interactions of sigma^{70} to polymerase and to the DNA promoter. At some point, as polymerase proceeds to initiate RNA synthesis, sigma^{70} loses contact with either the promoter or the polymerase (or both?). With the loss of at least one of these sets of contacts the net stability of the remaining complex becomes sufficiently low to cause sigma^{70} to dissociate. Currently, there is no evidence to permit us to choose between these interpretations.

$\sigma + R$	$\rightleftharpoons$	$\sigma{:}R$	$K = 5 \times 10^{9}\ M^{-1}$
$\sigma{:}R + DNA$	$\rightleftharpoons$	$\sigma{:}R{:}DNA_{nsp}$	$K = ?$
$\sigma{:}R{:}DNA_{nsp}$	$\rightleftharpoons$	$\sigma{:}R{:}P_c$	$K = ?$
$\sigma{:}R{:}P_c$	$\rightleftharpoons$	$\sigma{:}R{:}P_o$	$K = 1\text{-}5 \times 10^{4}$
$R + DNA$	$\rightleftharpoons$	$R{:}DNA_{nsp}$	$K = ?$
$\sigma + R{:}DNA_{nsp}$	$\rightleftharpoons$	$\sigma{:}R{:}DNA_{nsp}$	$K = ?$
$RNA_i + NTP$	$\rightleftharpoons$	$RNA_{i+1} + PPi$	$K = 1 \times 10^{2}$
$\sigma{:}R + DNA$	$\rightleftharpoons$	$\sigma{:}R{:}P_c$	$K = 1 \times 10^{7}\ M^{-1}$
$\sigma{:}R + DNA$	$\rightleftharpoons$	$\sigma{:}R{:}P_o$	$K = 1\text{-}5 \times 10^{11}\ M^{-1}$

Fig. 3. List of the equilibria involved in the initiation phase of transcription.

(iii) In any event polymerase must eventually re-adopt its initiation conformation in order for the transcription cycle to start again. This (third) conformational change is thought to occur at some point during termination, and results in restoring the core polymerase to the free solution form characterized by a high affinity for sigma^{70}. Since sigma^{70} has no observable effect on transcription termination [24,38] this conformational change probably occurs after the polymerase has been irreversibly committed to termination.

4.3. A free energy diagram for open promoter formation

The equilibrium constants that have been measured for various steps relative to initiation are summarized in fig. 3. From these data we can begin to establish relative free energy values for the different states along a transcriptional reaction coordinate.

The left-hand side of fig. 4 shows a free energy diagram, calculated using the equilibrium constants of fig. 3, for the formation of the open promoter complex during the initiation phase of transcription (we have arbitrarily used holoenzyme plus free DNA as a reference state to set the zero point of the relative free energy scale). Clearly, with respect to this reference state, the formation

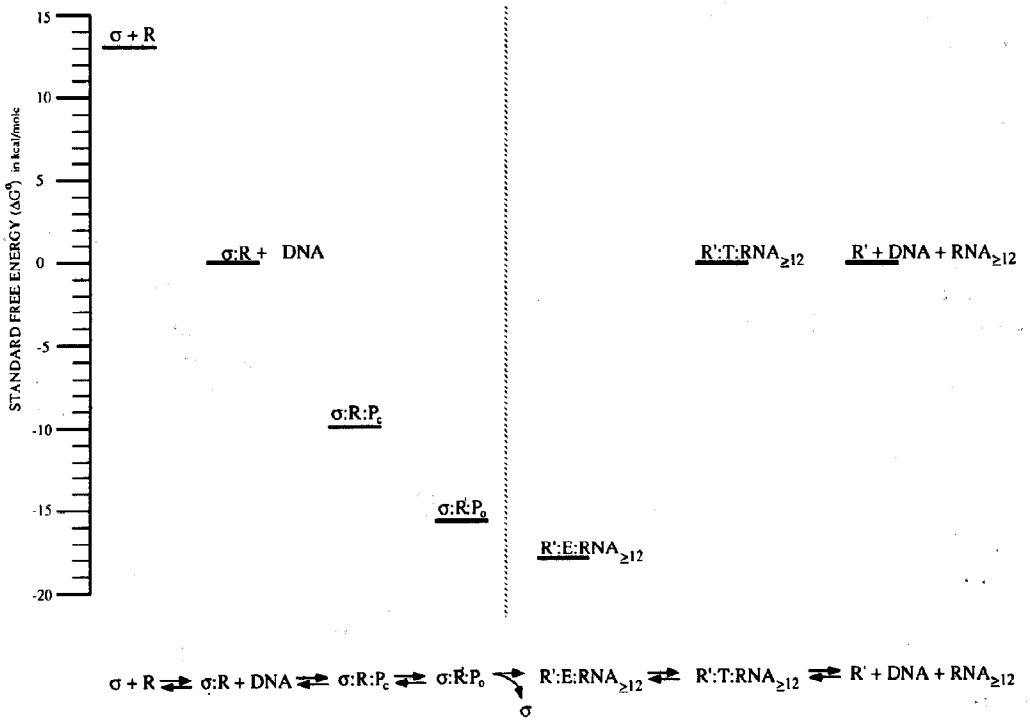

Fig. 4. Free energy diagram for some of the steps of the transcription cycle. The left-hand side of the diagram represents the formation of the open promoter complex, based on the equilibria of fig. 3. The right-hand side represents the relative free energies of the elongation and termination complexes (note that these are standard free energies (1 M standard states), so the ternary and the dissociated termination complexes have the same free energies). The break between the two sides of the diagram represents the irreversible step that separates the initiation and elongation phases. The overall reaction is shown at the bottom of the figure. For further details, see text.

of the open promoter complex is energetically favored, with $\Delta G^0 \sim -16$ kcal/mol. *

We have previously cited the experimental evidence for a conformational change as the holoenzyme undergoes the closed to open promoter transition. By comparing the free energies of these two states we can begin to determine how this conformational change influences holoenzyme binding and the closed-to-open transition.

* Ideally one would also like to know the binding affinity between holoenzyme and nonspecific DNA, to define the selectivity of the holoenzyme for a given promoter. However, this comparison is difficult to make, due in part to the competitive formation of other types of complexes between holoenzyme and DNA (e.g., end-binding complexes; tight-binding sites, etc.; see ref. 39).

We may write the following equation to define the standard free energy change involved in the closed-to-open promoter transition:

$$\Delta G^0_{\mathrm{H:Pc} \rightarrow \mathrm{H:Po}} = \left(\Delta G^0_{\mathrm{f,H:Po}} - G^0_{\mathrm{f,H:Pc}}\right) + \Delta G^0_{\mathrm{f,DNA\ bubble}} \quad (3)$$

where $\Delta G^0_{\mathrm{H:Pc} \rightarrow \mathrm{H:Po}}$ is the free energy change for the closed-to-open promoter transition (with holoenzyme bound), $\Delta G^0_{\mathrm{f,H:Po}}$ and $\Delta G^0_{\mathrm{f,H:Pc}}$ are the free energies of formation of the holoenzyme-P_0 and holoenzyme-P_c binary complexes, respectively, and $\Delta G^0_{\mathrm{f,DNA\ bubble}}$ is the free energy of opening the 12 base-pair DNA bubble at the promoter site.

Available data [6,40] show that the transition from the closed to open promoter complex corresponds to a standard free energy difference ($\Delta G^0_{\mathrm{H:Pc} \rightarrow \mathrm{H:Po}}$) of ~ -6 kcal/mol (see fig. 4).

The free energy change for forming the 12 base-pair DNA bubble ($\Delta G^0_{f,DNA\ bubble}$) may be calculated as the sum of disrupting 13 nearest-neighbor stacking interactions, plus the (entropically unfavorable) 'bubble initiation' event (see ref. 17). For an average promoter, we thus estimate that $\Delta G^0_{f,DNA\ bubble} = 13 \times (1.8) + 5.5$ kcal/mol = 29 kcal/mol. Using these free energy terms and rearranging eq. 3, we obtain $(\Delta G^0_{f,H:Po} - \Delta G^0_{f,H:Pc}) = -35$ kcal/mol, corresponding to the difference (between closed and open promoter complexes) of holoenzyme binding affinity for DNA.

The term $(\Delta G^0_{f,H:Po} - \Delta G^0_{f,H:Pc})$ in eq. 3 represents the standard free energy difference between closed and open promoter complex formation with holoenzyme. This free energy difference may be resolved into two components, one of which is due to the change in binding contacts between the holoenzyme and the DNA ($\Delta\Delta G^0_{f,binding}$), and the other to a change in holoenzyme conformation (ΔG^0_{conf}). At this time, we have no way to determine either of these two free energies independently. If ΔG^0_{conf} is positive then $\Delta\Delta G^0_{f,binding}$ will be less than or equal to -35 kcal/mol, while if ΔG^0_{conf} is negative then $\Delta\Delta G^0_{f,binding}$ will be greater than or equal to -35 kcal/mol. Qualitatively, we do know that the sum of these two components must increase the binding affinity between the holoenzyme and the promoter (in going from the closed to the open promoter complex) sufficiently to offset the cost of melting the DNA in open promoter formation.

4.4. Free energy change for the transition from the initiation mode to the elongation mode

The step that follows initiation and initial RNA chain extension in the detailed transcription cycle outlined in fig. 2 involves the detachment of the holoenzyme from the promoter, ejection of the sigma factor, and switching of the resulting ternary complex into the fully processive elongation mode. During the initial phases of transcript synthesis, short RNA fragments (2–8 nucleotide residues in length) may be released in the abortive initiation process [41,42]. This release is accompanied by movement of the bound holoenzyme to re-form the open promoter complex and reinitiate RNA synthesis. When a critical RNA length has been reached, sigma dissociates and core polymerase continues RNA synthesis within the highly stable and processive elongation mode.

The abortive initiation process should, in theory, be reversible. Such reversibility has been demonstrated experimentally using di- and trinucleotide fragments as RNA chain initiators [43]. One might expect the release of sigma to be reversible as well, but definitive experiments on this point are lacking.

The transition of RNA polymerase from the open promoter complex to the elongation complex can be thought of as a translocation of polymerase along the template DNA between two highly stable states that are in deep potential energy wells. This movement (or translocation) of polymerase along the DNA is stimulated by the corresponding synthesis of RNA. Very little, if any, polymerase dissociates from the DNA during this process, although more than 50% of the nascent RNA transcript may be 'aborted' here.

In order for polymerase to pass from the open promoter complex to the elongation complex mode, it must traverse an activation energy barrier. The free energy maximum of this transition is characterized by a complex set of determinants, because at this point: (i) the DNA bubble and RNA-DNA hybrid are growing to their mature lengths, (ii) the sigma factor is being released, (iii) a polymerase conformational change may be occurring, and (iv) a change in polymerase-nucleic acid contacts may also be taking place. Because of this complexity, we cannot yet formulate a convincing quantitative model of the initiation-to-elongation step. This is a topic for future research.

4.5. Thermodynamic stability of the elongation mode

Because of the complex nature of the initiation to elongation transition, we also cannot connect the elongation complex to the free energy diagram for open promoter complex formation that is presented on the left-hand side of fig. 4. We can, however, analyze the relative stability of the elongation complex as a function of position along the DNA template (right-hand side of fig. 4).

In the mature elongation complex, the DNA bubble has reached a size of 18 ± 1 open base-pairs, and the RNA-DNA hybrid has reached a length of 12 base-pairs. Elsewhere [17], we have calculated the free energy of formation of the elongation complex as a function of template position, relative to a zero stability state defined at the precise position at which an intrinsic terminator of transcription exerts its effect. At nonterminating positions, the average free energy of formation (disregarding 'initiation' terms for the DNA bubble and RNA-DNA hybrid) is approx. -18 kcal/mol. We also have estimated [18] that a further approx. 18 kcal/mol of activation free energy is required to overcome the barrier to RNA release at such template positions.

These calculations predict that: (i) transcript elongation should proceed in a deep potential energy well, (ii) 'stalled' elongation complexes should demonstrate long-term stability, and (iii) elongation should be highly processive. These predictions are consistent with experiment [31,37,41].

4.6. Energetics of RNA synthesis

The net free energy change that drives transcription in the elongation mode is due to a phosphoryl group transfer reaction in which a ribonucleotide monophosphate is added to the 3′ end of the nascent RNA chain, with the concomitant release of pyrophosphate. This reaction is characterized by a net standard free energy change of approx. -1 to -2 kcal/mol [22]. Elsewhere (see also ref. 22), we have analyzed the free energy change for RNA synthesis in terms of its elementary components.

It should be recalled that the free energy change for RNA synthesis is almost completely decoupled from the free energy terms that describe the average stability of the elongation complex. Thus, the conversion of rNTP units ('fuel') during the synthesis of the RNA transcript does not affect the stability of the elongation complex itself.

We also note that the efficiency of converting the initiation complex to the elongation complex may be governed by an irreversible and rate-limiting kinetic step (see above). Transcript extension past the critical 12-nucleotide level may occur more efficiently when a large phosphoryl-group transfer potential (driving force) is present.

4.7. Thermodynamics and kinetics of termination

Elsewhere, we have considered the thermodynamics and kinetics of intrinsic transcript termination [17,18]. In essence, we argue that the thermodynamic stability of the elongation complex approaches zero at many intrinsic termination sites, and that the kinetic barrier to termination becomes comparable to that for elongation at these positions. Thus, in terms of a free energy diagram such as that of fig. 4, the standard state thermodynamic stability of the complex will decrease by about 18 kcal/mol as we enter the termination phase (see right-hand side of fig. 4).

Finally, as indicated above, an irreversible (and incompletely defined) step occurs at this point, which prevents the elongation complex from being reassembled. The only way to reinitiate the cycle is for core polymerase to rebind sigma. This reforms holoenzyme, which can then rebind to a promoter and begin the transcription process again.

5. Summary and perspectives

Our understanding of the function of macromolecular complexes has traditionally developed through five levels of increasing structural and energetic detail (although these levels are not always passed through in the same order). These stages of representation of our knowledge are: (i) the chemical reaction or equation representation (though the stoichiometry and/or subunit composition of some or all of macromolecular components is often not known at this stage); (ii) the 'cartoon' representation, which often incorporates some functional information as well as facts about active sites, interactions, and shapes of some of the macromolecular components; (iii) a thermodynamic description of the complex and its interactions, including the affinities of substrates, ligands, and small molecule effectors for the macromolecular components of the complex and the stoichiometry and affinities of the macromolecular compo-

nents; (iv) a kinetic description of the same interactions, so that pathways, rates, and intermediate states can start to be elucidated; and (v) a detailed (atomic level) structural description of the complex and its associated ligands which is usually provided by X-ray crystallographers or two-dimensional NMR spectroscopists). Sometimes, if crystallization is rapid and physical biochemists are slow, stage (v) may be reached before the details of stages (iii) and (iv) are fully in hand.

It seems to us that our understanding of the mechanisms and regulation of transcription has passed through stages (i) and (ii) (see any recent literature or textbook review of the subject), and that we are now significantly into stages (iii) and (iv). Stage (v) is still a long way off!

In this paper we have attempted to systematize and to summarize what is known about the thermodynamics of the transcription cycle. To this end we have developed a partial free energy diagram for the transcription cycle (fig. 4).

We note that the cycle is interrupted at two points by reactions that appear to be irreversible. These points come at the transition from the initiation to the elongation phase, and during the last part of the termination phase. Both of these irreversible reactions require further definition.

In addition to describing the basic cycle, we have also described loops involving the regulatory proteins sigma^{70} and NusA. We have specifically included the sigma loop in the free energy analysis of the initiation phase of the cycle, since specific initiation is totally dependent on the sigma protein. We have not conducted a similar analysis of the NusA regulatory loop, although the results of Gill et al. [24] might allow this. Such an analysis might be premature, in part because NusA usually (though not always) functions in conjunction with other regulatory factors [23]. Also it appears that NusA probably does not influence the overall stability of the elongation complex, since it has been shown that this protein does not alter the rates of decay of stalled elongation complexes [37].

Although the overall analysis we have presented in this paper is incomplete, it may be useful in clarifying our ideas of what is and what is not known about *E. coli* transcription as a thermodynamic and kinetic process. We hope that this partial energetic and molecular framework will be helpful to others as well in designing and conducting experiments to extend our knowledge of this biologically central set of reactions.

Acknowledgements

These studies have been supported, in part, by USPHS Research Grants GM-15792 and GM-29158 (to P.H.vH.), by USPHS Individual Postdoctoral Fellowship GM-10227 (to T.D.Y.), and by a grant to the Institute of Molecular Biology from the Lucille P. Markey Charitable Trust. S.C.G. was a pre-doctoral trainee on USPHS Institutional Training Grant GM-07759. P.H.vH. is an American Cancer Society Research Professor of Chemistry. We are grateful to many colleagues for several discussions that have helped us to clarify the ideas presented in this overview.

References

1 P.H. von Hippel, D.G. Bear, W.D. Morgan and J.A. McSwiggen, Annu. Rev. Biochem. 53 (1984) 389.
2 W.R. McClure, Annu. Rev. Biochem. 54 (1985) 171.
3 T.D. Yager and P.H. von Hippel, in: *Escherichia coli* and *Salmonella typhimurium*, ed. F.C. Neidhardt (Am. Soc. Microbiol., Washington, DC, 1987) p. 1241.
4 J. Greenblatt and J. Li, Cell 24 (1981) 421.
5 P.H. von Hippel and O.G. Berg, J. Biol. Chem. 264 (1989) 675.
6 J.H. Roe, R.R. Burgess and M.T. Record Jr, J. Mol. Biol. 184 (1985) 441.
7 H. Buc and W.R. McClure, Biochemistry 24 (1985) 2712.
8 A.A. Travers and R.R. Burgess, Nature 222 (1969) 537.
9 U.M. Hansen and W.R. McClure, J. Biol. Chem. 255 (1980) 9564.
10 D.C. Straney and D.M. Crothers, Cell 43 (1985) 449.
11 B. Krummel and M.J. Chamberlin, Biochemistry 28 (1989) 7829.
12 H. Jehle, Proc. Natl. Acad. Sci. U.S.A. 53 (1965) 1451.
13 J.P. Richardson, J. Mol. Biol. 98 (1975) 565.
14 H.B. Gamper and J.E. Hearst, Cell 29 (1982) 81.
15 M.M. Hanna and C.F. Meares, Proc. Natl. Acad. Sci. U.S.A. 80 (1983) 4238.
16 Y. Shi, H. Gamper, B. Van Houten and J.E. Hearst, J. Mol. Biol. 199 (1988) 277.
17 T.D. Yager and P.H. von Hippel, Biochemistry 29 (1990) in the press.

18 P.H. von Hippel and T.D. Yager, Proc. Natl. Acad. Sci. U.S.A. 86 (1990) in the press.
19 T.A. Rozovskaya, A.A. Chenchik, N.B. Tarusova, R.S. Beabealashvilli and R.M. Khomutov, Mol. Biol. 15 (Moscow) (1982) 931. (Engl. Transl.)
20 G.A. Kassavetis, P.G. Zentner and E.P. Geiduschek, J. Biol. Chem. 261 (1986) 14256.
21 J.D. Kahn and J.E. Hearst, J. Mol. Biol. 205 (1989) 291.
22 T.D. Yager, J. Korte and P.H. von Hippel, manuscript in preparation.
23 S.C. Gill, S. Weitzel and P.H. von Hippel, J. Mol. Biol. (1990) submitted.
24 S.C. Gill, S. Weitzel and P.H. von Hippel, J. Mol. Biol. (1990) submitted.
25 Y. Nakamura and H. Uchida, Mol. Gen. Genet. 190 (1983) 196.
26 Y. Nakamura and A. Tsugawa, in: Sequence specificity in transcription and translation. UCLA Symp. Mol. Cell. Biol., vol. 30, eds R. Calendar and L. Gold (A.R. Liss, New York, 1985) p. 196.
27 A.T. Schauer, L.D. Carver, B. Bigelow, L.S. Baron and D.I. Friedman, J. Mol. Biol. 194 (1987) 679.
28 S.C. Gill, T.D. Yager and P.H. von Hippel, J. Mol. Biol. (1990) submitted.
29 S.L. Shaner, D.M. Piatt, G.C. Wensley, H. Yu, R.R. Burgess and M.T. Record Jr, Biochemistry 21 (1982) 5539.
30 M. Schmidt and M.J. Chamberlin, Biochemistry 23 (1984) 197.
31 R. Landick and C. Yanofsky, J. Mol. Biol. 196 (1987) 363.
32 W. Whalen, B. Ghosh and A. Das, Proc. Natl. Acad. Sci. U.S.A. 85 (1988) 2494.
33 C.D. Sigmund and E.A. Morgan, Biochemistry 27 (1988) 5622.
34 R. Reynolds, Thesis, University of California, Berkeley, CA (1988).
35 R.J. Horowitz, J. Li and J. Greenblatt, Cell 51 (1987) 631.
36 R.R. Burgess, Annu. Rev. Biochem. 40 (1971) 711.
37 K.M. Arndt and M.J. Chamberlin, J. Mol. Biol. 213 (1990) 79.
38 K.M. Arndt and M.J. Chamberlin, J. Mol. Biol. 202 (1988) 271.
39 P. Melancon, R.R. Burgess and M.T. Record Jr, Biochemistry 22 (1983) 5169.
40 C.J. Dayton, D.E. Prosen, K.L. Parker and C.L. Cech, J. Biol. Chem. 259 (1984) 1616.
41 A.J. Carpousis and J.D. Gralla, Biochemistry 19 (1980) 3245.
42 L.M. Munson and W.S. Reznikoff, Biochemistry 20 (1981) 2081.
43 N. Ruetsch and D. Dennis, J. Biol. Chem. 262 (1987) 1674.
44 G. Rhodes and M.J. Chamberlin, J. Biol. Chem. 249 (1974) 667.