### Citation

### As Published
http://dx.doi.org/10.1021/bi100947k

### Publisher
American Chemical Society (ACS)

### Version
Final published version

### Accessed
Mon Dec 10 18:58:45 EST 2018

### Citable Link
http://hdl.handle.net/1721.1/73054

### Terms of Use
Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.

### Detailed Terms

---

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Searching for the nik Operon: How a Ligand-Responsive Transcription Factor Hunts for Its DNA Binding Site†

Christine M. Phillips,‡,* Collin M. Stultz,‡,* and Catherine L. Drennan*‡,#

‡Department of Chemistry,§Research Laboratory of Electronics,‖Department of Biology,‡Department of Electrical Engineering and Computer Science, and ‡Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and ‡Howard Hughes Medical Institute

Received June 11, 2010; Revised Manuscript Received August 12, 2010

ABSTRACT: Transcription factors regulate a wide variety of genes in the cell and play a crucial role in maintaining cellular homeostasis. A major unresolved issue is how transcription factors find their specific DNA binding sequence in the vast expanse of the cell and how they do so at rates that appear faster than the diffusion limit. Here, we relate an atomic-detail model that has been developed to describe the transcription factor NikR’s mechanism of DNA binding to the broader theories of how transcription factors find their binding sites on DNA. NikR is the nickel regulatory transcription factor for many bacteria, and NikR from Escherichia coli is one of the best studied ligand-mediated transcription factors. For the E. coli NikR protein, there is a wide variety of structural, biochemical, and computational studies that provide significant insight into the NikR-DNA binding mechanism. We find that the two models, the atomic-level model for E. coli NikR and the cellular model for transcription factors in general, are in agreement, and the details laid out by the NikR system may lend additional credence to the current models for transcription factors searching for DNA.

Many cellular processes are regulated at the transcriptional level by proteins that can bind DNA and prevent or facilitate RNA polymerase binding to and transcribing DNA into mRNA. Transcription factors respond to a change in the cellular environment by binding or unbinding from DNA, thus altering the availability of a given gene for transcription. There are two main types of transcription factors: those regulated by a change in their own concentration that, in turn, changes the percentage of potential DNA sites having that transcription factor bound and those whose concentration remains more or less constant but whose affinity for DNA is regulated by a ligand binding mechanism. For ligand-binding transcription factors, the binding or dissociation of some ligand (small molecule, metal ion, etc.) alters the protein’s affinity for DNA. This review considers a ligand-binding transcription factor from Escherichia coli, NikR,† and how proposed mechanisms for NikR binding to DNA relate to current theories of how transcription factors find their specific DNA binding sites in the vast expanse of the genome.

The E. coli nickel regulatory protein, NikR, is a homotetrameric protein with a central metal binding domain (MBD) and two bordering dimeric ribbon–helix–helix (RHH) DNA binding domains (Figure 1a) (1). The MBD is a regulatory domain and the site of stoichiometric nickel binding (Figure 1b), while the dimeric RHH domains are responsible for making sequence-specific contacts with DNA (Figure 1c). Four nickel ions bind the NikR tetramer at strictly conserved square-planar sites in the MBD (Figure 1b), with a measured dissociation constant in the picomolar range (2, 3). Nickel binding to these sites induces high-affinity binding of NikR to the nik operon, with a nanomolar dissociation constant (2, 4), preventing the transcription of the nickel-specific uptake transporter NikABCDE and thereby limiting the amount of additional nickel entering the cell (5, 6). Interestingly, in the presence of an excess of stoichiometric nickel, the affinity of NikR for DNA increases, with the dissociation constant falling into the picomolar range (2, 4). Unlike stoichiometric nickel, however, excess nickel ions have a lower affinity for NikR, with dissociation constants in the micro- or millimolar range (2, 4). Excess or low-affinity nickel ions bind to nonconserved sites on the surface of the NikR protein (Figure 2) (7). Finally, NikR also has a binding site for a non-nickel metal. In addition to stoichiometric nickel, NikR requires potassium to bind to DNA with an apparent binding affinity of potassium for Ni2+-bound NikR of ≤1 mM (8). A potassium binding site is located at the interface of the MBD and the RHH domain (9, 10) (Figures 1c and 3). In addition to metal binding, the E. coli NikR protein has been extensively studied in terms of its DNA binding properties (2, 4, 6, 8, 11) and its structural states (1, 7, 10, 12). In fact, E. coli NikR is one of the few ligand-mediated transcription factors for which structures in all the various liganded states are known: apo NikR (Figure 1a), Ni2+-bound NikR (Figure 1b), and NikR-DNA complex (Figure 1c) (1, 10). Although the NikR from Helicobacter pylori has been the subject of several recent studies (13–19), the series of structures mentioned above is not available for this protein, making E. coli NikR our model system of choice. While some aspects of NikR function are still controversial (20, 21), here we will consider only the well-established features as we relate what we have learned about E. coli
NikR to the general question of how ligand-mediated transcriptional regulation is achieved at the molecular level.

Of key interest is the mechanism by which ligand-mediated transcription factors respond to their ligands and bind DNA, preventing transcription. Outlined in Figure 4a is a general simplistic mechanism. In this simple scheme, adapted for the NikR case, the binding of the ligand (nickel) induces (1) a change in the electrostatic potential of the protein (i.e., making the overall net charge of the protein more positive) and/or (2) a conformational change in the protein that leads to the formation of a structure that is complementary to the nik operon. Both of these effects would, in principle, explain the increased affinity of NikR for DNA in the presence of nickel. However, the wealth of research on the NikR system suggests that the involvement of metal ions in the regulation of NikR binding to DNA is considerably more complicated than the mechanism outlined in Figure 4a.

The more detailed mechanism shown in Figure 4b arises from an analysis of several crystal structures of E. coli NikR that are now available (apo, Ni\(^{2+}\)-bound, DNA-bound, and with various transition metals bound) in conjunction with biochemical and computational work on the system (1–4, 6, 9–12, 22, 23). These data argue that the first step in Ni\(^{2+}\)-induced NikR–DNA binding is the crystallographically observed (Figure 1a,b) ordering of the α3 helices and their associated loops following stoichiometric binding of Ni\(^{2+}\) to the high-affinity sites (α3 helices and associated loops are represented by ovals in Figure 4b) (1, 12). When ordered, crystallographic data show that these α3-associated loops can make nonspecific hydrogen bonding contacts with the DNA phosphate backbone (Figures 1c and 3, black dashed lines), presumably helping to localize NikR on DNA (10). High-affinity binding of NikR to its operator DNA can be achieved when the RHH domains adopt a structurally observed down-cis conformation (Figure 1c) that enables these domains to make site-specific contacts with nucleotides in the major grooves at the DNA binding sites (Figure 3, red dashed lines) (9, 10). This down-cis conformation of NikR can be further stabilized by potassium binding to the interface between the MBD and RHH domains (Figure 3, cyan dashed lines, and Figure 4b, where the potassium is shown as a gray sphere).

**Figure 1:** Crystal structures of full-length E. coli NikR. (a) Apo-NikR structure with disordered α3 helices in an “open” conformation. (b) Ni\(^{2+}\)-bound NikR in an open conformation with ordered α3 helices and nickel ions represented as green spheres and an inset of the high-affinity nickel site (green spheres). (c) NikR–DNA complex with NikR in a “down-cis” conformation with potassium ions (purple spheres) bound between the MBD and RHH domains and an inset view of the potassium site. The NikR–DNA complex is shown in surface representation to illustrate contacts between NikR and DNA, and the DNA motif responsible for specific recognition by NikR is indicated.
The mechanism in panel a of Figure 4 differs from that in panel b in several significant ways. First, the ligand (nickel) exerts its effect on the regulatory domain (MBD), not the DNA binding domain (RHH domain). This variation in mechanism is supported by crystallographic data (1, 10), which show the only significant conformational change upon nickel binding is the ordering of the α3 helix and its accompanying loop of the MBD, while the RHH domains remain in open conformations (Figure 1b). In addition, both crystallographic and biochemical data show a correlation between the conformation of the α3 helix and DNA binding properties (1, 4, 10, 12). Second, Figure 4b depicts a mechanism that is a requisite multistep process as nickel binding alone does not directly invoke the down-cis NikR conformation. Third, the mechanism in Figure 4b invokes a role for an abundant cellular cation, potassium, in serving to stabilize the down-cis NikR conformation, consistent with the NikR–DNA crystal structure (Figure 1c) as well as biochemical data (8–10). Importantly, Figure 4b shows nickel binding to NikR before potassium. Recent data suggest that nickel binding to NikR increases its affinity for potassium (8). Preliminary metal analysis by ICP-AES shows that under the same conditions in which potassium is readily removed from apo-NikR, Ni²⁺-bound NikR retains a significant amount. This finding is also consistent with the X-ray structure, which reveals an extended hydrogen bonding network around the potassium site that is only fully formed in the presence of stoichiometric nickel when the α3 helix and its accompanying loop are ordered (Figure 3). Thus, it is easy to rationalize from the structure how nickel binding to NikR could promote the binding of potassium.

The scheme in Figure 4b can be modified to include an additional step in which the excess or low-affinity nickel binds to NikR (Figure 4c). In vitro, excess nickel increases the affinity of NikR for DNA (2, 4), presumably through an electrostatic mechanism. Crystallographic data show that excess nickel binds to similar sites on the surface of the NikR protein regardless of whether NikR is free or bound to a piece of its operator DNA (Figure 2a,b) (7, 10). Although it is not known if excess or low-affinity nickel plays a role in vivo, here we depict the possibility that excess nickel ions could bind to either free or bound NikR, serving to enhance NikR’s affinity for the nik operon.

While the mechanisms shown in Figure 4 address the molecular changes that must occur for NikR to bind DNA, we have not yet considered how these regulatory mechanisms relate to the “search” NikR would have to undertake to find the nik operon. It is interesting to consider how our mechanisms shown in panels b and c of Figure 4 relate to current theories of how transcription factors find their binding sites on DNA. It was noted as early as 1970 by Riggs and co-workers that transcription factors can find their binding sites with surprising efficiency in vivo (24). Specifically, it was documented that the LacI repressor could find its
The binding site on DNA at a rate of \( \sim 10^{10} \text{ M}^{-1} \text{s}^{-1} \), which is faster than the diffusion limit in three dimensions, \( 10^8 \text{ M}^{-1} \text{s}^{-1} \) (24). To explain this phenomenon, the dimensionality of the search was proposed to vary with time (25, 26). Such variation in search dimensionality could be achieved if the protein spent some amount of time performing a one-dimensional search by sliding along the DNA in addition to some amount of time diffusing around the cell in a three-dimensional search (24–27). This sliding–diffusion model has been supported by a variety of biochemical and single-molecule experiments (28–38).

Recently, it has been suggested that DNA-bound sliding must incur little energetic penalty for this sliding–diffusion mechanism to account for the biologically observed rates at which transcription factors can find their appropriate binding sites on DNA (39). It follows that sliding must be associated with a weak affinity for DNA, while recognition of a specific operon requires the transcription factor to bind its complementary nucleic acid sequence with high affinity. A possible mechanism to account for the paradox of both weak and strong DNA binding is that the protein must have at least two distinct conformational states that
can bind DNA with different affinities (39). The weakly bound or "nonspecific binding" conformation binds different DNA sequences all with low affinity and is therefore impartial with respect to the precise DNA sequence. The second, "specific binding" conformation binds only the complementary DNA sequence with high affinity. In this model, the protein is able to sample both conformations as it continues its search until it finds its DNA binding site.

The wealth of data arising from crystallographic, biochemical, and computational studies on *E. coli* NikR supports a sliding—diffusion mechanism that employs at least three distinct NikR conformers, with each conformer having a different affinity for the *nik* operon (Figure 5). When no metal is bound to NikR, the protein can freely diffuse with no or little affinity [less than micromolar affinity (2)] for DNA. Once nickel ions bind to NikR, the α3 helices and loops preceding them become ordered, and this allows the protein to interact with DNA through nonspecific hydrogen bonds to backbone phosphate groups (Figure 3, black dashed lines). In this sense, the one-dimensional search along the DNA sequence does not begin until nickel ions bind at high-affinity nickel sites on NikR. However, because nickel binding is not sufficient to induce the RHH domains to adopt a down-cis conformation, the presence of nickel does not ensure that the protein will recognize the specific nucleic acid sequence in the major groove of DNA. The nickel-bound structure of NikR with the RHH domains in an open conformation therefore represents a nonspecific binding conformation of NikR, whose localization on DNA is sequence-independent (Figure 5). While we do not have a good experimental number for the affinity of such a nonspecific interaction between Ni\(^{2+}\)-bound NikR and DNA, because it is below the detection level of the standard gel shift assay (4, 8), an estimate for the nonspecific DNA affinity can be obtained from base substitution experiments in which bases in the *nik* operon binding motif are individually substituted. Individual mutations can lead to significant reductions in affinity (>1000-fold) (6).

Crystallographic studies of Ni\(^{2+}\)-bound NikR suggest that the RHH domains remain flexible when stoichiometric Ni\(^{2+}\) is bound at the high-affinity sites (10). Consequently, it is likely that thermal fluctuations cause these domains to sample a variety of orientations relative to the MBD in the nonspecific binding conformation. Fluctuations that cause the RHH domains to adopt a down-cis conformation enable the protein to "see" the precise nucleic acid sequence in the major groove. Therefore, as the protein slides along DNA, it can in principle sample different nucleic acid sequences along the DNA chain (Figure 5). Sampling of down-cis conformations is enhanced during sliding because

potassium, which is present at a millimolar concentration in the cell (40), binds at a site between the RHH domain and MBD and establishes a network of hydrogen bonds that involves additional interactions between the RHH domain and the phosphate backbone (Figure 3, cyan dashed lines). The hydrogen bonding network organized by potassium ion binding contains a number of fully conserved residues, including E30, D34, and R33, any of whose mutations abrogates DNA binding (Figure 3) (8, 10). Aspartic acid 34 coordinates the potassium ion in a bidentate fashion; E30 has a monodentate coordination to potassium, and the second oxygen participates in a hydrogen bonding network with conserved R33 (Figure 3). Arginine 33 then makes hydrogen bonds to the DNA phosphate backbone and participates in the ordering of the hydrogen bonding networks contacting DNA in both the MBD and RHH regions (Figure 3). Thus, this potassium site is well positioned to play a key role in organizing or maintaining a complex hydrogen bonding network between both domains of NikR and DNA that could stabilize the down-cis NikR conformation.

Crystallographic and biochemical data taken together suggest that a number of factors are required for the transcription factor to recognize its operon with high affinity: (1) the formation of nonspecific hydrogen bonds between the DNA phosphate backbone and both the MBD and RHH domain [note that RHH domains have no measurable affinity for DNA in absence of the MBD (6)], (2) binding of potassium to the secondary site in an octahedral geometry (8–10), and (3) hydrogen bonds between the RHH domain and specific nucleic acid bases in the operon (6, 10). When the RHH domains adopt a down-cis conformation, residues of these domains are poised to make hydrogen bonds with the nucleic acids in the DNA major groove that are known to be required for high-affinity binding (6, 10). Therefore, this down-cis conformation of NikR corresponds to a specific binding conformation that can recognize and bind the *nik* operon with the measured "high affinity" or nanomolar dissociation constant (2, 4) (Figure 5). However, a down-cis conformation of the protein at sites other than the *nik* operon corresponds to metastable states that would be expected to have short lifespans relative to that of the state where the protein is bound at its operon, with affinities decreased by an estimated 1000-fold for even one mistake in the operon binding motif (6). When the protein finds the correct nucleic acid sequence, the required specific interactions are formed and high-affinity binding is achieved (Figures 1 and 3).

To further illustrate these concepts, we can consider a putative energy landscape that describes the energy of different bound
conformations of the NikR nickel ions being able to stabilize both the nonspecific and specific can bind to both free and bound NikR, one can imagine excess depth of each of the wells on the landscape. Because excess nickel in vitro observation that the affinity of NikR for DNA is increased the DNA sequence (Figure 6c).

The down-cis conformation of NikR is a "specific" DNA binding conformation in which the RHH domains are sampling DNA. The down-cis conformation of NikR corresponds to NikR adopting a down-cis conformation at the correct DNA binding site (Figure 6a, blue well). By contrast, when NikR adopts an "open" or "up-cis" conformation, it can form hydrogen bonds only with the phosphate backbone of the DNA chain and therefore cannot distinguish between sequence differences in the DNA itself. This corresponds to a relatively rugged region of the energy landscape populated by many local energy minima that are separated by small energy barriers (Figure 6b). Each energy minimum corresponds to a different bound conformation of the protein, with NikR only hydrogen bonding to the phosphate backbone. Barriers on the energy surface arise because NikR needs to break hydrogen bonds with the phosphate backbone to "slide" to a new position on the DNA. The down-cis conformation of NikR is a "specific" DNA binding conformation in which the RHH domains are sampling the DNA sequence (Figure 6c). We could adapt the energy landscape in Figure 6a to consider the in vitro observation that the affinity of NikR for DNA is increased in the presence of excess nickel binding by simply increasing the depth of each of the wells on the landscape. Because excess nickel can bind to both free and bound NikR, one can imagine excess nickel ions being able to stabilize both the nonspecific and specific conformations of the NikR–DNA complex via favorable electrostatic interactions between the more positively charged NikR molecule and the negatively charged DNA. While stabilization of the nonspecific conformation of NikR bound to DNA may worsen the efficiency of the search, it would extend the lifetime of the bound NikR conformation until the nickel stress is over.

The ideas presented here represent a synthesis of experimental, theoretical, and computational data on the NikR system, which are consistent with NikR exploiting a sliding—diffusion mechanism to find its operon. An efficient search for the correct DNA sequence is achieved because NikR uses both a nonspecific, low-affinity, conformation to move along the DNA chain and a specific, high-affinity, conformation to recognize and bind to the operon. Thus, while one can propose a simplistic model for how NikR binding to DNA could be regulated by nickel (Figure 4a), this model fails when one considers the features needed for efficient binding site searching, perhaps explaining the need for a more complex mechanism of action (Figure 4b,c). Although the model we propose for NikR binding to DNA in panels b and c of Figure 4 and Figure 5 may not include all caveats of the NikR system, it does illustrate that ligand-bonding transcription factors can have very complicated mechanisms of regulation. In addition, this review opens the door for further studies to be conducted on the NikR system and other systems, testing the theories and comparisons we suggest here and extrapolating these ideas to a broader array of DNA binding proteins.

ACKNOWLEDGMENT

The authors thank Charles Fischer for his help with Figure 6.

REFERENCES


