The Structure of RNA-Free Rho Termination Factor Indicates a Dynamic Mechanism of Transcript Capture

Albert Canals¹,²*, Isabel Usón²,³ and Miquel Coll¹,²

¹Institute for Research in Biomedicine, Barcelona Science Park, Baldirí Reixac 10, 08028 Barcelona, Spain
²Institut de Biologia Molecular de Barcelona (CSIC), Barcelona Science Park, Baldirí Reixac 10, 08028 Barcelona, Spain
³Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Received 10 December 2009; received in revised form 15 April 2010; accepted 3 May 2010
Available online 7 May 2010

Edited by J. Doudna

© 2010 Elsevier Ltd. All rights reserved.

Keywords: helicase; Rho factor; RNA capture mechanism; transcription termination; X-ray crystal structure

The Rho factor is a ring-shaped ATP-dependent helicase that mediates transcription termination in most prokaryotic cells by disengaging the transcription elongation complex formed by the RNA polymerase, DNA, and the nascent RNA transcript. The crystal structures of key intermediates along the kinetic pathway of RNA binding to Rho unveiled an unprecedented mode of helicase loading and provided a model for the ATP turnover coupled to coordinated strand movement. Here we report the structure of the early RNA-free state of Rho, which had eluded crystallization for many years but now completes the series. The structure allows the characterization of the apo-form Rho from Thermotoga maritima to 2.3 Å resolution, reveals an RNA-recruiting site that becomes hidden after occupancy of the adjacent specific primary RNA-binding site, and suggests an enriched model for mRNA capture that is consistent with previous data.

The Rho factor is an essential ring-shaped ATP-dependent helicase that terminates transcription of about half of the open reading frames in most eubacteria.¹–³ The functional RecA-type enzyme is formed by six identical wedge-shaped Rho protomers that have an RNA-binding domain (RNA-BD) at the N-terminus linked by a connector loop to a C-terminal ATPase domain. The RNA-BD comprises a variable N-terminal helix bundle (NHB) subdomain and a conserved five-stranded β-barrel with an oligonucleotide/oligosaccharide binding (OB) fold that forms the primary RNA binding site, which has a preference for C-rich sequences. The C-terminal ATPase domain, highly homologous to the F₁-ATPase, comprises conserved Walker A (P loop) and Walker B motifs forming the ATP-binding site and two adjacent loops (known as Q and R loops) that form the secondary RNA binding site.

The conformational state of the Rho oligomer depends on the occupancy of the primary and secondary RNA binding sites.⁴,⁵ The crystal structures of Rho from Escherichia coli (ecRho from now on) in complex with RNA basically show two distinct nucleic-acid-dependent hexameric conformations, namely, the loading (RNA bound to the primary binding site) and the translocating (RNA bound to the secondary site) states. The first structure (3 Å resolution) captured the loading state of Rho, in which short RNA (or single-stranded DNA) chains are bound to both the primary and secondary RNA binding sites.⁶–⁷ The crystal structures of Rho from Thermotoga maritima (tmRho) to 2.3 Å resolution, reveals an RNA-recruiting site that becomes hidden after occupancy of the adjacent specific primary RNA-binding site, and suggests an enriched model for mRNA capture that is consistent with previous data.

Abbreviations used: RNA-BD, RNA-binding domain; NHB, N-terminal helix bundle; OB, oligonucleotide/oligosaccharide binding; tmRho, Thermotoga maritima Rho; ecRho, Escherichia coli Rho.
passing through the central channel—depicted the translocating state of the enzyme (Fig. 1b), which was further described in the recent report of a helicoidal (but still closed) hexameric, 2.8-Å-resolution structure of ecRho showing a remarkable asymmetry and with RNA bound only to the secondary site.

A wiggling version of the RNA-bound open ring

The full-length Rho transcription termination factor from Thermotoga maritima (tmRho from now on) was purified as a stable hexamer in 20 mM Tris–HCl (pH 7.6), 350 mM KCl, and 100 μM ethylenediaminetetraacetic acid, as described elsewhere. However, RNA-free tmRho crystallized only in a solution at a lower pH [0.1 M Na-acetate (pH 4.4) and 0.8 M MgSO4], and this condition shifted the equilibrium toward a lower-order dimeric association (Fig. 2a). Size-exclusion chromatography of tmRho confirmed that the hexamer dissociates at pH 4.5, and most of the protein is then eluted as a monomer regardless of the salt concentration (results not shown). In fact, early cryo-electron microscopy experiments had already shown that, under certain conditions, RNA-free ecRho can be a mixture of partial oligomerization states. However, that study also showed that the addition of (rC)23 oligomers under the same conditions results in a population of hexameric rings, either notched or closed. Binding of longer (rC)100 oligomers limits the observed conformations to closed hexamers. These observations suggest a stabilizing effect of the RNA cofactor on the Rho conformation and stress the difficulty of trapping an unbound hexamer in a crystal.

Our tmRho crystals show that the apo-tmRho protomers (Fig. 2b) have the same structural features that were reported in detail for the ecRho open ring (52.6% sequence identity and 72.8% similarity), with the exceptions of the NHB conformation and a subtle rearrangement at the adjacent OB fold. The structural elements within the ATPase domain—including loops P, Q, and R—are equivalent to those described by Skordalakes and Berger (r.m.s.d. of 1.3 Å for the Cα atoms) and show only minor differences as a result of their distinct residue sequences. Interestingly, electron density that can be attributed to a sulfate ion (0.8 M concentration in the crystallization solution) is found in the ATP-binding pocket, as was previously reported for one of the ecRho closed-ring structures. The NHB of protomer B is disordered and could not be traced, thereby suggesting that this is a mobile region in the apo form.

The relative orientation of the two tmRho protomers is that of any two contiguous subunits in the E. coli open-ring hexameric structure (Fig. 2c), with an r.m.s.d. of 1.4 Å for the positions of the Cα atoms when the conserved and ordered C-terminal ATPase domains of both dimers are superimposed. Conversely, our dimeric structure does not match any contiguous pairs in the available ecRho closed-ring structures (r.m.s.d. of 3.5 Å when superimposing the dimers forming the flat Rho structure and an average of 2.6 Å when superimposing contiguous pairs forming the helicoidal asymmetric Rho structure). The hexamer that was partially broken by the severe crystallization conditions can be pieced back together by joining six tmRho protomers, using the relative orientation of the two molecules that form the surviving dimers as a guide for the assembly. The operation that relates protomers A and B of our dimer is defined by spherical polar angles ω = 47.0°, ϕ = 100.3°, ψ = 55.6°, and a translation of (x, y, z) = (28.1, −2.7, −3.9) Å. Starting from protomer A, the most complete subunit in our crystal, the repetition of this operation five times with identical copies generates an open-ring hexamer (Fig. 2d) that is essentially like the one reported for ecRho. The superposition of the two homolog rings (r.m.s.d. of 2.8 Å for the equivalent Cα atoms at the six ATPase domains) reveals that the apo-tmRho hexamer generated is only slightly wider at its opening and has the NHB subdomains spread out (Fig. 2e). The apo-Rho assembly is possibly much more mobile in solution than the RNA-Rho complex because of thermally driven fluctuations that RNA binding may suppress in the bound form, as discussed later.
Fig. 2. Overall structure of the RNA-free Rho form. (a) Stereoview of the apo-tmRho dimer forming the crystals (one dimer per asymmetric unit), as seen from the side corresponding to the central channel in the hexameric form. (b) Structure of the apo-Rho protomer A showing the upper N-terminal (light blue) and lower C-terminal (blue) domains with their secondary-structure elements labeled. The N-terminal domain has two subdomains: a three-helix bundle (NHB) and an OB fold that forms the primary RNA-binding site. Loops P (yellow), Q (orange), and R (red) of the C-terminal ATPase domain are highlighted. (c) Comparison of the apo-tmRho dimer forming the crystal (blue) and two adjacent protomers of the E. coli open-ring structure (light orange). Cα atoms of the ATPase domain from apo-tmRho protomer A (tmA) are superimposed on those of protomer C from E. coli (ecC) and their respective couples tmB and ecD become superposed as a result. (d) Perspective and top–down diagrams of a hexameric apo-tmRho ring generated by propagation of the rotation and translation operation that relates the two protomers forming the crystals. (e) Top–down view of the superposition of the generated apo-tmRho open ring (blue) and the crystallographic open ring from E. coli (light orange). Only protomers A of the two rings have been explicitly superimposed. The preparation and crystallization of the recombinant Rho from T. maritima have been described in detail elsewhere.9 Crystals were grown by vapor diffusion at room temperature. One volume of the protein solution [3.3 mg/ml of Rho in 20 mM Tris–HCl (pH 7.6), 350 mM KCl, and 0.1 mM ethylenediaminetetraacetic acid] was mixed with an equal volume of the reservoir solution containing 0.1 M sodium acetate (pH 4.4) and 0.8 M magnesium sulfate. Crystals were flash-frozen in 0.1 M sodium acetate (pH 4.4), 20% ethylene glycol, and 1.0 M magnesium sulfate and belong to the space group P4322 with unit cell dimensions of a = b = 139.3 Å and c = 150.5 Å. Diffraction data were collected at the European Synchrotron Radiation Facility (Grenoble) beamlines ID14-3 and ID29 at 100 K and processed and scaled with XDS17 and XPREP.18 The structure was solved by the SIRAS method using two native 2.5-Å-resolution data sets and a uranium heavy-atom derivative data set at 2.3 Å resolution. The positions of the heavy atoms were determined with SHELXD, and the phasing and density modification were accomplished with SHELXE.19 The refinement was carried out using REFMAC520 with ARP/wARP,21 interspersed with inspection of electron-density maps, water positioning, and manual model building with Coot.22 Non-crystallographic symmetry restraints were not used during refinement.
The NHB as an RNA-recruiting subdomain

The analysis of the electrostatic surfaces of our apo-tmRho dimer reveals that, whereas the inner surface of the ring is, in general, positively charged, the outer surface is clearly negative except from the region corresponding to the NHB (Fig. 3a). The folding of the NHB from the open peripheral

Fig. 3. RNA-recruiting and cohesive roles of the NHB subdomain. (a) Inner (left) and outer (right) electrostatic surfaces of the apo-tmRho dimer of our crystals showing that the only electropositive patch on the outer surface of the ring is that corresponding to the NHB. (b) Sequence-based alignment of the N-terminal regions of tmRho and ecRho, with their secondary-structure elements. (c) Top–down stereoview of the NHB and OB-fold regions after the superposition of protomers tmA from the apo-tmRho (blue) and ecC from the loaded E. coli open ring (light orange). The red rod indicates the RNA chain bound to the OB fold of ecRho. Secondary-structure elements of tmRho are labeled in white or gray, and those from ecRho are in black. (d) Front stereoview of the same elements depicted in (d).
conformation found in the apo-Rho to the closed conformation of the loading ecRho (see Fig. 1a for reference) removes a significant electropositive patch from the outer surface of the ring and transfers it to the inner part of the molecule. As a result, this patch is buried in all the subunits of the loading notched hexamer except one, that of chain A, which flanks the ring opening and hence cannot contact any neighboring subunit. The electropositive patch corresponds to residues Lys19, Lys26, Arg32, Arg37, Lys38, Arg39, and Lys47 of the NHB. The resulting distribution of charges points to a likely scenario whereby the flexible NHB flops and dangles around the RNA-binding crown of the ring, weakly sticking to the electronegative RNA backbone, and thereby assists the contiguous primary site in the recognition of the cytosine-rich rut target sequence.

Kinetic studies with Rho from E. coli are consistent with a four-step sequential mechanism of RNA binding and suggest a first, fast binding event followed by RNA wrapping around the perimeter of the six-membered ring. The presence of a peripheral, nonspecific RNA-recruiting site is in keeping with the rapid formation of a first protein–RNA intermediate structure in the RNA binding pathway. A sophisticated evolution of the NHB as an RNA-recruitment element had already been reported for Rho from Micrococcus luteus, where the locally chaotic arginine-rich NHB subdomain was assigned the role of facilitating the binding of RNA transcripts with base-paired secondary structures at the rut site. Finally, the observation that the mutation of residue 3 in ecRho from Leu to Phe increases the affinity for polynucleotides at the primary binding site, probably by means of stacking interactions that are relatively base-independent, is again consistent with the notion that the NHB is involved in RNA recruitment.

A cohesive association toward the RNA-bound Rho

Although tmRho has seven extra residues at the N-terminus compared to its E. coli homolog (ecRho), the secondary structure elements of their NHB subdomains are equivalent—three α-helices, the last two connected by an ordered solvent-exposed loop in ecRho, which is extended to a 3_10 helix in tmRho—and essentially share the same boundaries (Fig. 3b). However, the relative orientation of these elements is strikingly different as revealed by superimposition of chain A from our apo-tmRho and any chain from the open-ring ecRho structure (Fig. 3c and d). While the ecRho NHB is folded inward, right beside the recognition loop β3_4β5 of the primary binding site, the β3_4β5 loop folds in to wrap the nucleotide. As a result, strands β4 and β5 are twisted and pulled inward, thereby favoring new hydrophobic interactions between helix α3 and both strand β5 and loop β2_3. Helix α2 joins helix α3, and this tight packing displaces the α_2_α3 loop toward the linker that joins the RNA-BD and the ATPase domain of a neighboring molecule. Thus, new subunit–subunit interactions can occur. After RNA binding to an OB fold in ecRho, several intermolecular hydrogen bonds and a stacking interaction of Arg30 with Phe133 of the neighboring molecule become possible. Consequently, the side chain of Arg28 is buried in the adjacent protomer and may form an additional salt bridge with Asp95 of the neighboring chain, thereby further reinforcing the association.

As the loading process may start at any of the OB folds, a certain degree of flexibility within the ring can be advantageous to recruit the nascent RNA and allows for the best possible positioning of the C-rich transcript sequence (the so-called rut site) around the primary RNA-binding crown. In contrast, the subsequent RNA internalization is localized between the two subunits separated by the notch, which has to be brought to a standstill for proper passage of the transcript into the central channel. Each binding event at the OB folds contributes to fix the whole ring in a synergic way through the intersubunit interactions described above, and hence, the notch location is much more stable after several of these events. The same stability of the interfaces between adjacent subunits—where the ATP-binding site resides—may be responsible for the improved ATP uptake registered in kinetic assays with RNA-bound Rho compared to that of the RNA-free enzyme. Moreover, this dependence on subunit–subunit interactions provides an explanation as to why truncations of 22 and 28 residues at the N-terminal region of ecRho result in impaired and depleted termination activity, respectively.

A dynamic mechanism for RNA capture

Considering our findings, and in accordance with previous studies, we propose a mechanism for RNA capture by Rho (Fig. 4). To grasp the whole process, one must imagine a highly dynamic scenario. Before contacting the RNA transcript, Rho is a hexameric open ring with the notch rapidly shifting between its six possible locations. The motion is even greater around the crown of RNA-BDs, where the six NHBs randomly sweep the periphery of the enzyme in search of negatively charged molecules (Fig. 4a).
The loading process involves a first unspecific binding event when the nascent RNA is transiently bound to the positively charged patch of any of the six NHBs of Rho (Fig. 4a-II). One of the moving positively charged NHBs recruits the ribosome-free RNA transcript (red) by sticking to its negatively charged phosphate backbone (II). This transient binding assists the adjacent OB fold in the specific C-rich RNA recognition, which triggers a conformational change of the NHB that leads it to interact with the neighboring ring protomer (III). Additional binding events at the rest of the OB folds contribute to a steadier conformation, and the ring opening becomes fixed in a single location (IV). As a side effect, the RNA-recruiting subdomain stops wobbling and remains folded and fixed in a region where it will no longer interact with RNA, which might be detrimental for the subsequent loading of neighboring subunits. After the first specific binding to an OB fold, there is a greater probability of occupancy of adjacent subunits and each additional binding event increases the firmness of the entire hexamer until reaching a ring with a fixed opening. As an alternative for fixing the open-ring conformation, RNA may also help to lock ring neighbors by sticking to unloaded subunits through weaker interactions at their peripheral, and still unfolded, positively charged NHBs (Fig. 4a-IV). This auxiliary binding discharges the RNA chain from the need to zigzag between each and every neighboring primary site, as would be mandatory if all the OB folds had to be occupied. Furthermore, the presence of a peripheral binding site—even if “tertiary”—explains how Rho also terminates transcription of RNA containing hairpins downstream from the rut site.

Surprisingly, what we have deduced from our structure suggests that the RNA chain might enter the ring more easily from the most uplifted RNA-BD (that of subunit A in the available structures of the loading open ring) than from the nethermost one (subunit F) as proposed from the earlier ecRho structures (see Fig. 1 for reference). In fact, the folding of the NHB of subunit A upon RNA binding can prevent the RNA chain from sliding out of the ring through the cleft as it leaves the OB fold of chain A and favor the interaction of the 3′ end of the transcript with the adjacent Q loop from the same subunit (Fig. 4b). In the reverse polarity (the one depicted in Fig. 1), subunit F lacks any element capable of playing an equivalent role.

At a later stage, and according to the structure of the mature asymmetric closed-ring Rho, the NHBs have the same conformation and relative position (with respect to the rest of the protomer) that they...
have at the earlier RNA-bound open-ring state. However, the overall reorganization of the hexamer leaves them solvent-exposed again and this may explain the fact that most of the NHs of the closed-ring structures are slightly disordered and some residues could not be traced. The enzyme is from then on free to rearrange with the necessary asymmetry to work properly.

The structure reported here, together with previous studies, inspires an innovative and simple model for transcript internalization and sheds some light on hitherto unexplained observations of Rho behavior. The model assumes that RNA binding plays a major stabilizing role, but it has to be noted that the oligomeric state of Rho and, hence, the association between its subunits are also dependent on other factors such as the nature of the ionic environment, the presence of nucleotides, or the pH as we have discussed earlier. Nevertheless, our structural data provide a framework for designing new biochemical and biophysical experiments that should be able to accurately determine how the highly unsteady association of the oligomer components is overcome during the initial stages of Rho function.

### Protein Data Bank accession number

Coordinates and structure factors have been deposited in the Protein Data Bank under accession number 3L0O.

### Acknowledgements

This work was supported by the “Ministerio de Educación y Ciencia de España” (Grants BFU2005-24123-E and BFU2008-02372/BMC to M.C. and Grants BIO2006-14139 and BIO2009-10576 to I.U.), by the “Generalitat de Catalunya” (Grant 2009SGR-1309), by the EU (Spine2-Complexes LSHG-2006-031220 and 3DReperoire LSHG-CT-2005-512028 projects), and by the “Fundació La Marató de TV3” (Grant 052810). Synchrotron data collection was supported by the European Synchrotron Radiation Facility. A.C. acknowledges a postdoctoral fellowship from the “Ministerio de Educación y Ciencia de España”.

### Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2010.05.004

### References

