Biochimica et Biophysica Acta xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Review

Single-molecule studies of riboswitch folding ☆

Andrew Savinov [a], Christian F. Perez [b], Steven M. Block [c,d,*]

[a] Biophysics Program, Stanford University, Stanford, CA 94305, USA
[b] Department of Physics, Stanford University, Stanford, CA 94305, USA
[c] Department of Applied Physics, Stanford University, Stanford, CA 94305, USA
[d] Department of Biology, Stanford University, Stanford, CA 94305, USA

ARTICLE INFO

Article history:
Received 20 December 2013
Received in revised form 27 March 2014
Accepted 3 April 2014
Available online xxxx

Keywords:
Single molecule
Optical trap
Optical tweezers
Gene regulation
Regulatory mechanism

ABSTRACT

The folding dynamics of riboswitches are central to their ability to modulate gene expression in response to environmental cues. In most cases, a structural competition between the formation of a ligand-binding aptamer and an expression platform (or some other competing off-state) determines the regulatory outcome. Here, we review single-molecule studies of riboswitch folding and function, predominantly carried out using single-molecule FRET or optical trapping approaches. Recent results have supplied new insights into riboswitch folding energy landscapes, the mechanisms of ligand binding, the roles played by divalent ions, the applicability of hierarchical folding models, and kinetic vs. thermodynamic control schemes. We anticipate that future work, based on improved data sets and potentially combining multiple experimental techniques, will enable the development of more complete models for complex RNA folding processes. This article is part of a Special Issue entitled: Riboswitches.

1. Introduction

A foundational principle of molecular biology is that the explanations of living processes are bound up in the relationship between structure and function in biomolecules. A grand challenge of biology has been to build predictive models of organisms by leveraging our understanding of the structure–function relationship. One milestone in this pursuit would be the ability to predict the three-dimensional structures of proteins and nucleic acids directly from their genetic sequences [1–6]. But static structural information alone is insufficient: biomolecules are also highly dynamic. The interplay between structural dynamics and function is therefore a key to understanding the mechanisms underlying biological activity at the molecular level. Because RNA molecules can form comparatively simple structures that are capable of gene regulation and catalysis, in addition to storing and conveying genetic information, they have become model systems for studying the relationship between molecular structure and function in cellular processes. The riboswitch, a regulatory element that acts *in cis* on one or more genes encoded by its mRNA [7,8], is an exquisite example. The dynamics of riboswitch folding determine the regulatory fate of the gene(s) under its control. Any quantitative model of riboswitch function must therefore, of necessity, involve a detailed description of its folding dynamics. Folding behavior may be modeled as a diffusive process over a multi-dimensional energy landscape (Section 2). The physical mechanisms underpinning riboswitch function can then be understood in terms of this folding landscape and its response to a target ligand in the context of the regulatory mechanism employed by the switch (for example, the regulation of transcription, translation, or alternative splicing).

Accurately reconstructing folding energy landscapes requires the compilation of molecular trajectories in sufficient numbers, and with sufficient observation time, to ensure that all relevant subpopulations have been sampled and identified. Single-molecule approaches afford a means to observe directly individual folding trajectories in a manner free of issues that arise from population averaging and lack of synchronization, which otherwise hamper the interpretations of ensemble-based measurements. In particular, single-molecule Förster resonance energy transfer (smFRET) [9] or fluorescence quenching [10] can measure conformational changes subtending distances of 2–10 nm on as many as hundreds of molecules simultaneously, with a typical time resolution ranging from ~1 to 100 ms (Section 3.2). Optical trapping (OT) [11], atomic force microscopy (AFM) [12], and magnetic tweezers (MT) [13,14] facilitate observations of folding trajectories via the measurement of end-to-end changes in molecular extension. These techniques have achieved spatial resolution down to ~0.3 nm [15] and temporal resolution below 100 μs [16] and allow for applied forces ranging from ~0.1 to ~100 pN (Section 3.3). The various available single-molecule methods are versatile and, in many respects, complementary.

☆ This article is part of a Special Issue entitled: Riboswitches.

* Corresponding author at: Department of Biology, Stanford University, Stanford, CA 94305, USA. Tel.: +1 650 724 4046.
E-mail address: sblock@stanford.edu (S.M. Block).

http://dx.doi.org/10.1016/j.bbagrm.2014.04.005
1874-9399/

2. Riboswitch folding: models and mechanisms

Folding studies of riboswitches are organized around one or more key concepts: these include free energy landscapes, the effects of ligand binding, aptamer-expression platform competition, kinetic versus thermodynamic control, the role of counterions, the nature of intermediates and misfolded states, and the interplay between secondary and tertiary structures. Riboswitches generally employ a modular architecture consisting of an aptamer domain and expression platform. To date, most folding investigations have been confined to studies of the isolated riboswitch aptamer. The ligand-bound form of the aptamer will either interfere with, or promote, the formation of the expression platform and thereby, depending upon the context, up- or down-regulate gene expression. Folding experiments typically measure the lifetimes of various conformational states and the responses of those lifetimes to ligands, cations, forces, or chemical denaturants, in order to ascertain the physical mechanisms responsible for riboswitch behavior.

Folding free energy landscapes [17,18] provide a unifying framework for interpreting results from different experimental techniques and disparate riboswitches. A ribonucleotide polymer chain has a great many conformational degrees of freedom, and the folding energy associated with any given molecular configuration can therefore be represented on a hyper-dimensional "surface," or folding energy landscape. For simplicity, this hyper-surface is generally flattened (projected) onto just a one- or two-dimensional surface, where the multiple conformational degrees of freedom become represented by one or two well-chosen *reaction coordinates* [19] (Fig. 1). As RNA folds, it samples, in principle, from a vast conformational space – the ensemble of all possible structures – to settle into its functional, native state (or, in some cases, one of several native states [20–22]), which is often the minimum free-energy state. The fact that biological polymers manage to fold into native conformations on biological timescales, despite the vastness of the conformational space, has been dubbed the "Levinthal Paradox" [23]. The development of a "folding funnel" picture of the folding free energy landscape offered one resolution to the paradox. In this picture, the conformational search takes place on a funnel-like energy landscape leading to the native state(s), thereby biasing the entire folding process and constraining the space that gets explored [24]. There are often a series of locally accessible, non-native intermediate states found on the funnel landscape between the unfolded ensemble and the native state (Fig. 1; Section 5.1). Prior work has suggested that RNA folding landscapes tend to be fairly rugged, with substantial activation energy barriers separating multiple intermediate [25,26] and native [20–22] states. Occasionally, the intermediate states compete with longer-lived, misfolded states [27], which – in the case of riboswitches – can act to delay or prevent aptamer formation (Section 4.2).

Generally speaking, RNA folding tends to be hierarchical, with secondary structures (which tend to have lower energy than tertiary structures) typically folding first, followed by tertiary contact formation [28–31]. A *hierarchical folding* model has been invoked to describe the folding of many RNAs fairly well (Section 4.4). However, numerous counterexamples have been identified that violate a strict hierarchy (for example, ribonucleoprotein assembly [32]), demonstrating a more complex interplay between secondary and tertiary structure formation (Sections 4.1 and 4.2). In recognition of the limitations and approximations implicit in the hierarchical model, an alternative "quasi-hierarchical" view of RNA folding has been proposed [33]. The hierarchical framework may nevertheless serve as a starting point for RNA structure prediction, without necessarily capturing the detailed RNA folding kinetics [29].

RNA is a polyanion at physiological pH, and the formation of compact structures requires cations to screen its negatively-charged phosphate backbone [30,34–38]. Divalent magnesium ions, in particular, facilitate tertiary interactions through both non-specific "screening" and specific "binding" interactions, and are sometimes necessary for riboswitch ligand binding (Sections 4.4 and 4.5). The significant role played by counterions in structured RNAs compels most folding studies to explore the effect of monovalent and divalent salts, even when no specific interactions are anticipated.

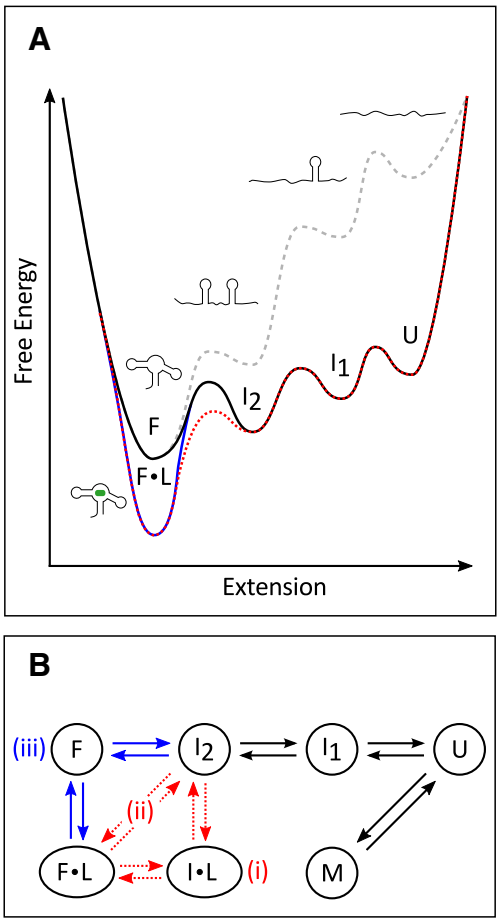

Fig. 1. Riboswitch aptamer folding. (A) Notional free energy landscape for the folding and ligand binding of a generic riboswitch aptamer domain. Energy is plotted vs. end-to-end extension, the reaction coordinate used in optical trapping experiments. Example conformational states (**U**: unfolded, $\mathbf{I_{1,2}}$: intermediates; **F**: folded, **F·L**: folded and bound to ligand) are shown for each energy well (bound ligand in green). Ligand-free energy landscapes at zero applied load (gray dashed curve) and under external tension (black curve) are plotted against extension. Also shown are folding landscapes in the presence of ligand, following either a conformational selection (blue curve) or an induced fit (red dotted curve) model for binding, both under applied tension. (B) Kinetic scheme matching the energy landscape in (A), with additional potential states represented (**M**: a misfolded state, **I·L**: an induced-fit intermediate). Roman numerals indicate two possible induced-fit pathways (**i**, **ii**) and the conformational selection pathway (**iii**).

Single-molecule experiments can probe both the effects of ligand binding on folding dynamics and the mechanisms that dictate how such effects propagate through the riboswitch to produce genetic regulation. Riboswitch aptamers are able to bind their cognate ligands with high specificity and affinity, and studies of mutant aptamer folding dynamics can reveal the origin of these characteristics (Sections 4.2, 4.5, and 4.8). Ligand binding perturbs the free energy landscape – at a minimum, decreasing the energy of the fully folded state – and can potentially influence folding *rates*, as well. The effect of ligand binding on RNA folding has been conceptualized in terms of two models: (1) induced fit and (2) conformational selection (also called conformational capture).

To define these models in the context of riboswitch folding, we consider the case of an aptamer folding from an unfolded state (**U**) via a series of intermediates to its fully folded RNA state (**F**, or **F·L** with ligand bound), with a notional landscape shown in Fig. 1A. We then focus on the transition from the final intermediate, $\mathbf{I_2}$, to the fully folded aptamer, **F**, where **F** is the "ON" switch state promoted by ligand. [In principle, an