

## Mini Review

# Crossroad Intermediates as 'Folding Pathway' Control Mechanism of Gene Expression

Sadaf F\*

Department of Biotechnology, Jamia Millia Islamia, India

**\*Corresponding author:** Sadaf F, Department of Biotechnology, Jamia Millia Islamia, India**Received:** March 15, 2017; **Accepted:** April 19, 2017;**Published:** April 26, 2017**Abstract**

Crossroad Intermediates (CRI) are those folding pathway intermediates that are situated at crossroads between protein folding and aggregation. CRI can either lead to folding of protein into native functional state (if one pathway is followed) or lead to formation of aggregated non-functional state (if another pathway is followed). We propose here for the first time, that formation of CRI may be the last control mechanism of gene expression chosen by Nature (after transcriptional, post-transcriptional, translational control mechanisms) - an event that occurs just before formation of native state in order to maintain concentration of particular protein. Since this step occurs during folding of protein, we name it as 'Folding Pathway' Control Mechanism of gene expression. We also propose that switch mechanism, for these CRI following a particular pathway lies in the concentration of that protein. Notion that Protein Folding pathway should also be subjected to control mechanism has long been ignored in the history of biological sciences. Information in this article will help to reshape and strengthen our understanding of Control of Gene Expression.

**Abbreviations**

DNA: Deoxyribonucleic Acid; RNA: Ribonucleic Acid; CRI: Cross Road Intermediate

**Introduction****Gene expression and its control**

Gene Expression is set of steps that starting from gene ultimately leads to production of proteins or functional RNA product. (In this communication, we are concerned with protein synthesis; hence in all further description gene expression would mean production of proteins from genes). It is the most important function of cell to maintain its structure and function. Gene Expression starts with Transcription (production of RNA from DNA) and then leads to Translation (production of proteins from RNA). When RNA is synthesized from DNA by process of transcription that occurs in nucleus of cell, it undergoes several modifications known as Post-Transcriptional Modifications and then transported out of nucleus. In cytoplasm, RNA is translated into protein sequence. Protein then undergoes Post-translational modifications (in some cases). Linear chain of amino-acids is then folded into 3-dimensional structure by process of Protein Folding before it becomes functionally active. Transcription, Post-Transcriptional Modifications, Translation, Post-translational modifications- all constitute the elaborate process of Gene Expression.

As much as Gene Expression is important, so is its control, in order to maintain timing and concentration of protein at a particular site in cell. Although same DNA sequence is present in all cells, it is control of gene expression that defines protein content (and subsequently content of other bio molecules) of a particular cell at a particular time and makes the cell specialized and distinct from other cells [1]. Control of gene expression is exercised at each step involved in this process. Transcriptional control forms the first step of control

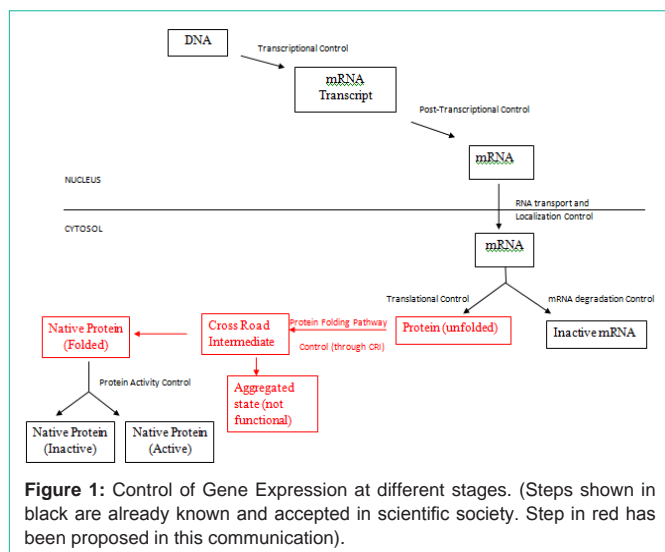
and is usually the most important step [2]. Post-transcriptional control [3] occurs after mRNA is synthesized. Once protein is being synthesized from mRNA, translational control mechanism takes over and tries to maintain concentration of protein being synthesized [4]. The protein can also be subjected to activity control mechanism [5] switching it between active or inactive states.

This basic information about steps involved in gene expression and its control mechanism have been discussed in all Molecular Biology books and are summarized here in Figure 1 (all steps shown in black).

**Protein folding**

After being synthesized by translation, native linear chain of polypeptide has to fold in order to achieve its functional state by the process known as Protein Folding. Protein folding is thus described as a process by which a disordered protein chain diffuses across a high-dimensional energy landscape and finally reaches the folded ensemble [6]. It has been inferred that proteins must fold to their unique native state through multiple unpredictable routes and intermediate conformations [7]. The search problem involved in folding however has been simplified through the evolution of folding energy landscapes that are funnelled [8]. Funnel-shaped energy landscape pictures that proteins must fold energetically downhill along with decrease in entropy in order to form native conformation [9]. The free energy landscape can provide a quantitative description of folding dynamics, if determined as a function of an optimally chosen reaction coordinate [10]. Single molecule studies have provided novel insights about how the dynamic sampling of the free energy landscape dictates all aspects of protein behaviour; from its folding to function [11].

Proteins fold through many independent pathways forming multiple partially folded intermediate states. Characterization of partially folded intermediate states has long been done by scientists

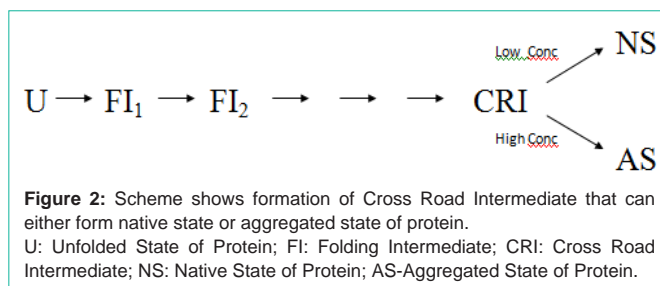


in order to understand folding mechanism. By subjecting the protein to denaturing conditions in order to induce and capture partially folded structures, it is characterized with respect to its secondary and tertiary (or quaternary) structure. While unfolding reaction is studied in-vitro, it is imagined that similar intermediates would have been stabilized while actual folding reaction proceeds in-vivo. Scientists also perform continuous-flow micro fluidic mixing and spectroscopic study to observe the first several milliseconds of folding, to capture and study early intermediate states [12]. Our group has also been involved in characterization of folding intermediates for a number of proteins [13-19].

### Crossroad Intermediates (CRI)

While studying intermediate states, scientists have come across certain states which can follow two pathways- one of which leads to formation of native functional fully folded form of protein, while the other leads to aggregation of protein forming misfolded non-functional form. Such intermediate states have been termed as Crossroad Intermediates (CRI). Aggregation prone intermediate states have been frequently reported in different proteins [20,21] indicating that a lot of data has accumulated over a number of years showing the presence of CRI, an intermediate that can either form native state or aggregated state. Formation of CRI from unfolded protein on the pathway crossroad between folding and aggregation has been depicted in Figure 2. Such a conformation state at the crossroad between folding and aggregation has been reported [22] at pH 4.8 and 10% TFE for human stefin B variant (bearing Y at site 31). A Native-like Intermediate has been reported as a branching point between the folding and aggregation pathways of the mouse prion protein [23]. Folding intermediate of a small protein module--the FF domain- acts as a central player in both folding and misfolding pathways and illustrates how incomplete folding can lead to the formation of higher-order structures [24].

Recently we have also characterized CRI of Lima Bean Trypsin Inhibitor at pH 4.0 (unpublished result). The CRI state stabilized at pH 4.0 lies between MG state (at pH 2.0) and native state (at pH 7.0). Our results show that when present at low concentration (0.5mg.ml<sup>-1</sup>) upon increasing from pH 4.0 to 7.0, native state can be



achieved. However, if present at high concentration (5mg.ml<sup>-1</sup>), this state is converted into aggregated form incapable of performing any function or retaining structure. LBTI is not prone to aggregation in the entire pH range except at pH 4.0, showing that regions of protein responsible for aggregation are exposed at mild pH while they are buried safely in native state [25]. CRI at pH 4.0 showed concentration-dependant increase in turbidity measurement experiment. This result was further confirmed by Thioflavin T and Congo-Red binding, dyes used as probe for fibril formation [26]. TEM images also showed formation of fibrils in concentration dependant manner.

### Outlook

Formation of CRI decides whether final product will be a native functional state of protein or non-native aggregated state. We propose that formation of CRI states is the final step chosen by Nature for control of gene expression that lies on the protein folding pathway. The steps shown in red in Figure 2 have been proposed in this article. Switch mechanism for these CRI following a particular pathway lies in concentration of protein. At low concentration of protein, CRI follows the path leading to formation of native protein that is required to carry out necessary function in the cell. If linear chain of protein is synthesized in more than sufficient quantity, CRI is formed during protein folding procedure; it leads to formation of aggregated state. Aggregates would then be removed by cellular machinery. In this way, formation of CRI takes care of maintaining proper required concentration of protein in the cell- requisite function of gene expression control mechanisms.

It is new proposal for existence of CRI as gene expression mechanism. This proposal also opens up new questions such as need to characterize CRI for lots of proteins and also some counterparts to CRI need to be addressed in case of natively denatured proteins. However an important question need to be addressed: When all the steps involved in synthesis of proteins have been regulated, it stands to reason out why elaborate protein folding mechanism has not been subjected to any control?

### References

- Control of Gene Expression - Garland Science.
- Satou Y, Imai KS. Gene regulatory systems that control gene expression in the Ciona embryo. Proc Jpn Acad Ser B Phys Biol Sci. 2015; 91: 33-51.
- Weil TT. Post-transcriptional regulation of early embryogenesis. F1000 Prime Rep. 2010; 7: 31.
- Istomine R, Pavey N, Piccirillo CA. Posttranscriptional and Translational Control of Gene Regulation in CD4+ T cell Subsets. J Immunol. 2016; 196: 533-540.
- Weber S, Meyer-Roxlau S, Wagner M, Dobrev D, El-Armouche A. Counteracting Protein Kinase Activity in the Heart: The Multiple Roles of

- Protein Phosphatases. *Front Pharmacol.* 2015; 6: 270.
6. Whitford PC, Onuchic JN. What protein folding teaches us about biological function and molecular machines. *Curr Opin Struct Biol.* 2015; 30: 57-62.
  7. Englander SW, Mayne L. The nature of protein folding pathways. *Proc Natl Acad Sci USA.* 2014; 111, 15873-15880.
  8. Wolynes PG. Evolution, energy landscapes and the paradoxes of protein folding. *Biochimie.* 2015; 119: 218-230.
  9. Plotkin SS, Onuchic JN. Understanding protein folding with energy landscape theory. Part I: Basic concepts. *Q Rev Biophys.* 2002; 35: 111-167.
  10. Banushkina PV, Krivov SV. High-resolution free energy landscape analysis of protein folding. *Biochem Soc Trans.* 2015; 43: 157-161.
  11. Bavishi K, Hatzakis NS. Shedding light on protein folding, structural and functional dynamics by single molecule studies. *Molecules.* 2014; 19: 19407-19434.
  12. Rosen LE, Kathuria SV, Matthews CR, Bilsel O, Marqusee S. Non-native structure appears in microseconds during the folding of *E. coli* RNase H. *J Mol Biol.* 2015; 427: 443-453.
  13. Rabbani G, Ahmad E, Zaidi N, Fatima S, Khan RH. pH-Induced molten globule state of *Rhizopus niveus* lipase is more resistant against thermal and chemical denaturation than its native state. *Cell Biochem Biophys.* 2012; 62: 487-499.
  14. Ahmad E, Fatima S, Khan MM, Khan RH. More stable structure of wheat germ lipase at low pH than its native state. *Biochimie.* 2010; 92: 885-893.
  15. Fatima S, Khan RH. Stability check of succinylated concanavalin A: presence of functionally active conformational state. *Protein Pept Lett.* 2009; 16: 423-429.
  16. Fatima S, Mishra A, Sen P, Khan RH. Characterization of fluoroalcohols-induced intermediates of *Mucor miehei* lipase at low pH. *Protein Pept Lett.* 2008; 15: 346-352.
  17. Fatima S, Ahmad B, Khan RH. Native-like tertiary structure in the *Mucor miehei* lipase molten globule state obtained at low pH. *IUBMB Life.* 2007; 59: 179-186.
  18. Fatima S, Ahmad B, Khan RH. Fluoroalcohols induced unfolding of Succinylated Con A: native like beta-structure in partially folded intermediate and alpha-helix in molten globule like state. *Arch Biochem Biophys.* 2006; 454: 170-180.
  19. Naeem A, Fatima S, Khan RH. Characterization of partially folded intermediates of papain in presence of cationic, anionic, and nonionic detergents at low pH. *Biopolymers.* 2006; 83: 1-10.
  20. Das NK, Ghosh N, Kale AP, Mondal R, Anand U, Ghosh S, et al. Temperature induced morphological transitions from native to unfolded aggregated States of human serum albumin. *J Phys Chem B.* 2014; 118: 7267-7276.
  21. Neudecker P, Robustelli P, Cavalli A, Walsh P, Lundström P, Zarrine-Afsar A, et al. Structure of an intermediate state in protein folding and aggregation. *Science.* 2012; 336: 362-366.
  22. Smajlović A, Berbić S, Žerovnik E. The cross-road between the mechanisms of protein folding and aggregation; study of human stefin B and its H75W mutant. *Biochem Biophys Res Commun.* 2011; 415: 337-341.
  23. Honda RP, Xu M, Yamaguchi K, Roder H, Kuwata K. A Native-like Intermediate Serves as a Branching Point between the Folding and Aggregation Pathways of the Mouse Prion Protein. *Structure.* 2015; 23: 1735-1742.
  24. Korzhnev DM, Religa TL, Kay LE. Transiently populated intermediate functions as a branching point of the FF domain folding pathway. *Proc Natl Acad Sci USA.* 2012; 109: 17777-17782.
  25. Gianni S, Camilloni C, Giri R, Toto A, Bonetti D, Morrone A, et al. Understanding the frustration arising from the competition between function, misfolding, and aggregation in a globular protein. *Proc Natl Acad Sci USA.* 2014; 111: 14141-14146.
  26. Reinke AA, Gestwicki JE. Insight into amyloid structure using chemical probes. *Chem Biol Drug Des.* 2011; 77: 399-411.