Structural conversion of the transformer protein RfaH: new insights derived from protein structure prediction and molecular dynamics simulations

Nicole Balasco, Daniela Barone and Luigi Vitagliano

Institute of Biostructures and Bioimaging, C.N.R., Via Mezzocannone 16, Naples I 80134, Italy; Dipartimento di Scienze e Tecnologie Ambientali Biologiche e Farmaceutiche, Seconda Università di Napoli, Caserta 81100, Italy

Communicated by Ramaswamy H. Sarma

(Received 27 October 2014; accepted 29 November 2014)

Recent structural investigations have shown that the C-terminal domain (CTD) of the transcription factor RfaH undergoes unique structural modifications that have a profound impact into its functional properties. These modifications cause a complete change in RfaH<sup>CTD</sup> topology that converts from an α-hairpin to a β-barrel fold. To gain insights into the determinants of this major structural conversion, we here performed computational studies (protein structure prediction and molecular dynamics simulations) on RfaH<sup>CTD</sup>. Although these analyses, in line with literature data, suggest that the isolated RfaH<sup>CTD</sup> has a strong preference for the β-barrel fold, they also highlight that a specific region of the protein is endowed with a chameleon conformational behavior. In particular, the Leu-rich region (residues 141–145) has a good propensity to adopt both α-helical and β-structured states. Intriguingly, in the RfaH homolog NusG, whose CTD uniquely adopts the β-barrel fold, the corresponding region is rich in residues as Val or Ile that present a strong preference for the β-structure. On this basis, we suggest that the presence of this Leu-rich element in RfaH<sup>CTD</sup> may be responsible for the peculiar structural behavior of the domain. The analysis of the sequences of RfaH family (PfamA code PF02357) unraveled that other members potentially share the structural properties of RfaH<sup>CTD</sup>. These observations suggest that the unusual conformational behavior of RfaH<sup>CTD</sup> may be rare but not unique.

Keywords: secondary structure prediction; conformational conversion; molecular dynamics; transformer proteins; amino acid propensities

Introduction

One of the basic paradigms of structural biology is that protein structure is dictated by the amino acid sequence (Anfinsen principle) (Anfinsen, 1973). The relation between sequence and structure is however not biunivocal. Indeed, similar sequences tend to assume rather similar structures, whereas the same structure is frequently assumed by multitudes of unrelated sequences. In this context, the analysis of similar/identical sequences that adopt significantly different structures has attracted the attention of the scientific community (see, e.g. the “Paracelsus Challenge”) (Dalal, Balasubramanian, & Regan, 1997). It is worth mentioning that short protein fragments in different contexts may easily assume different conformations (chameleon sequences) (Gendoo & Harrison, 2011). It has also been reported the case of specific protein regions that present significant differences between predicted and actual structures (discordant sequences) (Gendoo & Harrison, 2011). A special and rather widespread case of these structural ambiguities is represented by peptides and proteins that, in addition to their folded or unfolded physiological states, form in special conditions amyloid-like aggregates (Chiti & Dobson, 2006).

The most striking example among structural/functional conversions between completely different folds is represented by the carboxy-terminal domain of the transcription factor RfaH (Burmann et al., 2012). This protein belongs to the NusG family which includes important regulators of transcription in all living organisms (Tomar & Artsimovitch, 2013; Werner & Grohmann, 2011). Escherichia coli (E. coli) RfaH is characterized by a modular organization consisting of an N-terminal RNP-like domain (NTD – residues 1–100) and a C-terminal element with a KOW sequence motif (Belogurov et al., 2007; Kyripides, Woese, & Ouzounis, 1996). The X-ray structure of the full-length E. coli RfaH showed that the C-terminal domain (CTD – residues 101–162) folds as an α-hairpin that is tightly bound to the NTD (Belogurov et al., 2007). This interaction hides the RNAP-binding site of RfaH. Indeed, only when RNAP pauses at ops-containing operons, RfaH<sup>CTD</sup> loses its contacts with RfaH<sup>NTD</sup> allowing the interaction between RfaH and the polymerase (Belogurov, Mooney, Svetlov, Landick, & Artsimovitch, 2009). Very recently, Roesch and co-workers surprisingly observed that RfaH<sup>CTD</sup> undergoes a huge
conformational switch from the all-α state to an all-β fold when it is detached from the rest of the protein (Burmann et al., 2012; Svetlov & Nudler, 2012). By adopting this β-barrel fold, the domain acquires the ability to recruit ribosome, thereby activating translation (Tomar, Knauer, Nandymazumdar, Rosch, & Artsimovitch, 2013). Since RfaH simultaneously changes both topology and function, it has been considered a transformer protein (Knauer, Artsimovitch, & Rosch, 2012; Knauer, Rosch, & Artsimovitch, 2012).

In order to gain further insights into this unique structural transformation, we here preliminarily performed some sequence analyses aimed at unraveling the intrinsic preferences of RfaH<sup>CTD</sup>. We then evaluated the stability of the isolated α-hairpin (hereafter referred as αRfaH<sup>CTD</sup>) and β-barrel (βRfaH<sup>CTD</sup>) folds of the domain by performing molecular dynamics simulations in explicit solvent. Since these analyses provided some clues on the sequence determinants of RfaH<sup>CTD</sup> structural transition, we used our results to search KOW sequence motifs in order to identify other proteins that potentially share this peculiar conformational behavior with RfaH.

**Materials and methods**

**Sequence analyses**

RfaH<sup>CTD</sup> secondary structure predictions were performed using the servers PsiPred (http://bioinf.cs.ucl.ac.uk/psipred/) (McGuffin, Bryson, & Jones, 2000) and Net-CSSP which computes the contact-dependent secondary structure propensities (http://cssp2.sookmyung.ac.kr/) (Kim, Choi, Lee, Welsh, & Yoon, 2009).

**Molecular dynamics: models and protocols**

Molecular dynamics simulations were performed on both αRfaH<sup>CTD</sup> and βRfaH<sup>CTD</sup>. αRfaH<sup>CTD</sup> consists of an extended helix-turn-helix motif in which the two α-helices span residues 118–130 (helix 1) and 134–156 (helix 2). On the other hand, βRfaH<sup>CTD</sup> is a compact five-stranded β-sheet. The starting model of the helical RfaH<sup>CTD</sup> was extracted from the crystal structure of the entire protein solved at 2.1 Å resolution (PDB: 2OUG) (Belogurov et al., 2007). Some regions (residues 109–114 and 158–162), that were missing in the crystal structure, were manually modeled as coil. In the simulation of βRfaH<sup>CTD</sup>, the solution structure of the β-barrel CTD (PDB: 2LCL) (Burmann et al., 2012) was used as starting model.

Molecular dynamics simulations were performed using the GROMACS software package 4.5.5 (Van Der Spoel et al., 2005) with Amber99sb as force field and TIP4P-EW as water model. The models were immersed in cubic boxes filled with water molecules. The dimensions of the boxes were 7.59 × 7.19 × 7.25 nm<sup>3</sup> (number of water molecules 12,626) and 7.02 × 6.86 × 6.49 nm<sup>3</sup> (number of water molecules 10,067) for αRfaH<sup>CTD</sup> and βRfaH<sup>CTD</sup>, respectively. The simulations were run with periodic boundary conditions. Equilibration was conducted in two phases. In the first phase, conducted under an NVT ensemble, the temperature of the systems (300 K) was stabilized. Equilibration of pressure (1 atm) was then carried out under an NPT ensemble. Before starting the MD simulations, energies were minimized by fixing the protein atoms and then without restraints. The system temperature was brought to 300 K in a stepwise manner. In particular, 100 ps MD runs were carried out at 50, 100, 150, 200, 250, and 300 K. The time step was 0.002 ps for both systems. The timescales were 250 and 160 ns for αRfaH<sup>CTD</sup> and βRfaH<sup>CTD</sup>, respectively. Bond lengths were constrained using the LINCS procedure. Lennard-Jones interactions were calculated with a cutoff of 10 Å. Electrostatic interactions were treated using the particle mesh Ewald method with a grid spacing of 0.12 nm.

In order to examine the collective motions of the proteins in the MD simulations, the essential degrees of freedom were extracted according to the ED method (Amadei, Linssen, & Berendsen, 1993). This method is based on the construction of the covariance matrix of the coordinate fluctuations. This matrix was diagonalized to obtain eigenvectors and eigenvalues. The first 10–20 eigenvectors usually represent most (80–90%) of the motion in the protein.

To evaluate whether the simulation reached an adequate convergence in the essential space (Amadei et al., 1993), the root mean square inner product (RMSIP) between two halves of the equilibrated trajectory was calculated as described by Di Nola and coworkers (Amadei, Ceruso, & Di Nola, 1999; Merlino, Vitagliano, Ceruso, Di Nola, & Mazzarella, 2002; Merlino, Vitagliano, Ceruso, & Mazzarella, 2003). The RMSIP between the first 10 eigenvectors is defined as

$$
\sqrt{\frac{1}{10} \sum_{i=1}^{10} \sum_{j=1}^{10} (\eta^a_i \eta^b_j)^2}
$$

where η<sup>a</sup><sub>i</sub> and η<sup>b</sup><sub>j</sub> are the <i>i</i><sup>th</sup> and <i>j</i><sup>th</sup> eigenvectors from the first and second half of the equilibrated trajectory, respectively.

Trajectories were checked to assess the quality of the simulation using GROMACS routines and the program VMD (Humphrey, Dalke, & Schulten, 1996). In order to identify H-bond interactions, both donor–acceptor distances (<3.5 Å) and hydrogen-hydrogen donor-hydrogen acceptor angles (<30°) were checked using GROMACS utilities (Van Der Spoel et al., 2005).
Results

Intrinsic structural preferences of RfaH$^{CTD}$ and NusG$^{CTD}$: secondary structure prediction

The servers Psipred and Net-CSSP were used to gain insights into the intrinsic structural preferences of RfaH$^{CTD}$. For comparative purposes, these analyses were also conducted on the C-terminal region of the homolog *E. coli* NusG (NusG$^{CTD}$), which uniquely folds as a β-barrel and does not display any structural conversion. With the exception of the fragment 137–146 of RfaH$^{CTD}$, Psipred predicts an overall preference for the β-structure for CTDs of both proteins (Figure S1(A–B)). PsiPred predicts a β-structure preference for the RfaH C-terminal region also when the sequence of the full-length protein was used (Figure S1(C)). Taking into account that this domain is helical in the experimental structure of the full-length protein, RfaH$^{CTD}$ represents a major discordant region. Therefore, the dual conformational behavior of this portion could have been highlighted prior to the NMR characterization of the isolated RfaH$^{CTD}$ by simple comparisons between experimental structure and prediction.

Net-CSSP provides similar predictive results showing that both RfaH$^{CTD}$ and NusG$^{CTD}$ have a significant preference for the β-structure. However, a deeper inspection of Figure 1 indicates that the RfaH$^{CTD}$ region 137–146, which is a β-strand in the NMR structure of the isolated domain has a significant preference for the helical state. This feature is not exhibited by the corresponding region of NusG$^{CTD}$.

Molecular dynamics simulation of RfaH$^{CTD}$

In order to gain further insights into the huge structural transformation of RfaH$^{CTD}$, we performed all-atom molecular dynamics studies on both conformers of the domain, βRfaH$^{CTD}$, and αRfaH$^{CTD}$.

The β-barrel structure of βRfaH$^{CTD}$ is stable and rigid

The solution structure of βRfaH$^{CTD}$ (PDB: 2LCL) was employed as starting model (Figure 2(A)) in the MD simulation which was performed using GROMACS software (see Materials and methods). Indicators commonly adopted to check system stability in MD analyses reveal that the β-motif is extremely stable within time. Furthermore, the very high value of the RMSIP, 0.92, calculated between the two halves of the equilibrated trajectory (20–160 ns) confirms that the simulation has reached an appropriate convergence in the simulation timescale. βRfaH$^{CTD}$ does not exhibit significant deviations from the starting model with average RMSD values, computed on the C$\alpha$ atoms, of ~1.5 Å (Figure S2(A)). βRfaH$^{CTD}$ gyration radius (Rg) is stable (variations of about 1.5 Å) along the time (Figure S2(C)). The time evolution of the protein secondary structure elements (Figure 3(A)) is in good agreement with the very high stability observed. βRfaH$^{CTD}$ totally retains its secondary structure content throughout the simulation. The root mean square fluctuation (RMSF) values, calculated on C$\alpha$ atoms in the equilibrated region of the trajectory, are rather low for the structured elements (below 0.5 Å). Higher fluctuations (~1–2 Å) are exhibited by residues of loops (Figure S3).

In order to gain further insights into the dynamics of this highly stable folding, we analyzed the collective motions of βRfaH$^{CTD}$ by essential dynamics. This analysis showed that most of the atomic displacements were contained in the first few eigenvectors. In particular, the essential subspace spanned by the first 10 eigenvectors covers about 72% of the total fluctuations. The regions involved in the motions described by the first eigenvector (26% of the total fluctuations) can be visualized plotting the protein along this principal component in a film-like fashion (Figure 2(B)). In agreement with the RMSF profile, significant atomic displacements observed for the first eigenvector are mainly concentrated in loop regions.

Figure 1. Contact-dependent secondary structure propensities (CSSP) of RfaH$^{CTD}$ (A) and NusG$^{CTD}$ (B) computed using the Net-CSSP server. Helix and beta propensities are represented in red and blue, respectively.
The αRfaHCTD fold is highly unstable

The structure of αRfaHCTD, used as starting model in the MD simulation, corresponds to the residues 109–162 of the crystal structure of the full-length protein (see Materials and methods) (Figure 2(C)). Since the early stages of the simulation, αRfaHCTD displayed a completely different dynamic behavior compared to βRfaHCTD. The trend of the RMSD values of the trajectory structures shows that αRfaHCTD undergoes significant variations in the interval 0–160 ns. Indeed, RMSD values, computed on Cα atoms, are in the range 3–13 Å (Figure S2(B)). In the 190–250 ns interval, RMSD values become less variable with an average of ~10 Å. An analogous trend can be observed in the evolution of αRfaHCTD gyration radius (Figure S2(D)). After significant fluctuations, Rg assumes a lower value (about 11 Å) beyond 180 ns indicating a more compact structure.

The instability of αRfaHCTD fold is confirmed by the time evolution of the secondary structure elements (Figure 3(B)) which clearly shows the unfolding of most of the helical regions of the protein. Interestingly, a short segment belonging to helix 2 (residues 135–145) retains the α-helical content in the simulation timescale.

Discussion

In line with the experimental evidence, prediction programs and molecular dynamics simulations indicate that the isolated C-terminal region of RfaH has a strong preference for the β-structure. The MD simulation clearly indicates that αRfaHCTD α-hairpin is highly unstable. Nevertheless, this analysis shows that the two helices of the α-hairpin motif display a rather different behavior. Indeed, helix 1 rapidly unfolds, whereas a significant portion of helix 2 remains structured in the simulation. This finding is in line with previous MD studies carried out on this system by using rather different approaches (replica exchange in implicit solvent, targeted and high-temperature MD in implicit solvent) (Gc, Bhandari, Gerstman, & Chapagain, 2014; Li et al., 2014). In both of these reports, which were published while the present work was in progress, a higher stability of helix 2 was observed.

According to these authors, the rapid unfolding of helix 1 in the isolated RfaHCTD was ascribed to the loss of favorable interdomain contacts with RfaHNTD that stabilize this helix in the full-length protein. However, a detailed analysis of the interfaces between RfaHCTD and the two helices clearly indicates that the interactions
formed by helix 2 with the N-terminal domain are comparable or even larger than those involving helix 1. Indeed, the overall surface, computed using the server PDBePISA (http://www.ebi.ac.uk/pdbe/pisa/), of RfaHNTD/helix 2 and RfaHNTD/helix 1 interfaces is 590 Å² and 320 Å², respectively. RfaHNTD/helix 2 interface is characterized by stabilizing interactions such as a salt bridge (Glu 48 of the NTD and Arg138 of helix 2) and several hydrogen bonds. On the contrary, no electrostatic interaction is formed by RfaHNTD with helix 1. These observations indicate that helix 2 is able to maintain the helical structure despite the loss of important interactions. These considerations suggest that residues of helix 2 are intrinsically endowed with a higher propensity for helical states compared to residues of helix 1. The analysis of helix 2 sequence shows the presence of a Leu-rich segment (LLLNL, residues 141–145) which has a high intrinsic propensity for the helical conformation as shown by Net-CSSP (Figure 1). This reflects the good intrinsic propensity of Leu for α-helix (Fersht, 1999; Myers, Pace, & Scholtz, 1997). It is interesting to note that the corresponding region in NusG CTD, the RfaH homolog whose CTD forms a stable non-transient β-structure, shows an abundance of valine/isoleucine (VSVSI, residues 160–164). These residues present a marked preference for the β-structure and a very poor propensity for the α-helix (Fersht, 1999; Myers et al., 1997). This is consistent with the analysis of the PDB structures containing the LLLNL sequence motif present in RfaHCTD. Indeed, in addition to RfaH, we identified this motif in other eleven proteins. Notably, these fragments adopt a helical structure in nine cases (PDB codes: 1C8R, 1H8L, 2BL0, 2WZF, 3F7F, 3VE9, 4AN8, 4H3T, and 4IMP). In the other two cases, the motif is either disordered (PDB code: 2E0Z) or it adopts a mixed β/irregular structure (PDB code: 4O5A).

On this basis, we can reasonably propose that the specific secondary structure propensity of the Leu-rich segment may be one of the determining factors for the peculiar and hitherto unique conformational properties of RfaHCTD. In order to identify new proteins that potentially share this behavior with RfaHCTD, we analyzed the sequence of this segment in proteins belonging to NusG family. The family (PfamA code PF02357) includes 3544 sequences with the NusG-KOW architecture. The vast majority of these proteins presents valine/isoleucine-rich sequences in the considered region. For example, as many as 1254 sequences exhibit the motif VSV. On the other hand, very few members present residues with good helix propensity. Indeed, only 4 sequences exhibit...
the Leu-rich segment LLL of RfaH\textsuperscript{CTD}. This observation is confirmed by the analysis of the motifs XLL, LXL, and LLX, where \(X\) is an \(\alpha\)-helix preferer (Ala, Arg, or Met) (Fersht, 1999; Myers et al., 1997). In particular, the motifs ALL, LML, and LRL are observed only in 65, 29, and 1 sequences, respectively. By considering in the final ensemble only sequences with identities lower than 80\%, we obtained 2, 17, 1, and 1 sequences containing LLL, ALL, LML, and LRL motifs, respectively (Table S1). The analysis of the structural propensity of the selected sequences, carried out by using the server Net-CSSP, shows a close similarity with the trends observed for RfaH\textsuperscript{CTD}. As shown in Figure 4, in which representative examples of the Net-CSSP plots are reported, the chameleon behavior of these sequences is evident. It is important to note that the propensity of the Leu-rich region is higher for helical states compared to \(\beta\)-structure. This finding suggests that the identified sequences encode potential transformer proteins, which may undergo the structural transition observed in RfaH.

In conclusion, by combining prediction tools and MD simulation, we highlighted the role of specific sequence features. With a single exception (uniprot code C2NS75), in all sequences the charged residues involved in the interdomain interaction in RfaH (Glu 48 and Arg 138) are functionally conserved.

Currently, this sequence ensemble is essentially made by predicted or inferred from homology proteins. Only in one case, the encoded protein has been experimentally detected (uniprot code Q5XE43).

In conclusion, by combining prediction tools and MD simulation, we highlighted the role of specific sequence regions that may play a role in the peculiar, but functional, conversion of RfaH\textsuperscript{CTD}. Our data suggest that the conformational behavior of RfaH\textsuperscript{CTD} is rare but not unique. Future characterizations of the proteins encoded by the sequences here identified may lead to the discovery of new transformer proteins.

Present analyses also suggest that other examples of transformer proteins, unrelated to RfaH, may be unveiled by the analyses of discrepancies between secondary structure predictions and “actual” structures reported in the PDB. Finally, the identification of the structural determinants of this structural switch opens the possibility for the design and the development of new active biomolecules endowed with this intriguing property.

**Abbreviations**

| MD         | RMSF         | RMSD         | RMSIP       |
|------------|--------------|--------------|-------------|
| molecular dynamics | root mean square fluctuation | root mean square deviation | root mean square inner product |

**Supplementary material**

The supplementary material for this paper is available online at http://dx.doi.org/10.1080/07391102.2014.994188.

**Acknowledgments**

The authors thank Luca De Luca for skillful technical assistance.

**Funding**

This work was financially supported by MIUR (FIRB RBRN07BMCT).

**References**

Amadei, A., Cenuso, M. A., & Di Nola, A. (1999). On the convergence of the conformational coordinates basis set obtained by the essential dynamics analysis of proteins’ molecular dynamics simulations. *Proteins: Structure, Function, and Genetics*, 36, 419–424.

Amadei, A., Linssen, A. B., & Berendsen, H. J. (1993). Essential dynamics of proteins. *Proteins: Structure, Function, and Genetics*, 17, 412–425. doi:10.1002/pro.340170408

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*, 181, 223–230.

Belogurov, G. A., Mooney, R. A., Svetlov, V., Landick, R., & Artsimovitch, I. (2007). Structural basis for converting a general transcription elongation factor into a translational regulator. *The EMBO Journal*, 28, 112–122. doi:10.1038/emboj.2008.268

Belogurov, G. A., Vassylyeva, M. N., Svetlov, V., Klyuyev, S., Grishin, N. V., Vassylyev, D. G., & Artsimovitch, I. (2007). Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Molecular Cell*, 26, 117–129. doi:10.1016/j.molcel.2007.02.021

Burmann, B. M., Knauer, S. H., Sevostyanova, A., Schweimer, K., Mooney, R. A., Landick, R., … & Artsimovitch, I. (2012). An \(\alpha\) helix to \(\beta\) barrel domain switch transforms the transcription factor RfaH into a translation factor. *Cell*, 150, 291–303. doi:10.1016/j.cell.2012.05.042

Chiti, F., & Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry*, 75, 333–366. doi:10.1146/annurev.biochem.75.101304.123901

Dalal, S., Balasubramanian, S., & Regan, L. (1997). Protein alchemy: Changing \(\beta\)-sheet into \(\alpha\)-helix. *Nature Structural Biology*, 4, 548–552.

Fersht, A. (1999). *Structure and mechanism in protein science*. New York, NY: W. H. Freeman.

Ge, J. B., Bhandary, Y. R., Gerstman, B. S., & Chapagain, P. P. (2014). Molecular dynamics investigations of the \(\alpha\)-helix to \(\beta\)-barrel conformational transformation in the RfaH transcription factor. *The Journal of Physical Chemistry B*, 118, 5101–5108. doi:10.1021/jp502193v
Gendoo, D. M., & Harrison, P. M. (2011). Discordant and chameleon sequences: Their distribution and implications for amyloidogenicity. *Protein Science, 20*, 567–579. doi:10.1002/pro.590

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics, 14*, 27–38.

Kim, C., Choi, J., Lee, S. J., Welsh, W. J., & Yoon, S. (2009). NetCSSP: Web application for predicting chameleon sequences and amyloid fibril formation. *Nucleic Acids Research, 37*, W469–W473. doi:10.1093/nar/gkp351

Knauer, S. H., Artsimovitch, I., & Rösch, P. (2012). Transformer proteins. *Cell Cycle, 11*, 4289–4290. doi:10.4161/cc.22468

Knauer, S. H., Rösch, P., & Artsimovitch, I. (2012). Transformation. *RNA Biology, 9*, 1418–1423. doi:10.4161/rta.22724

Kyrpides, N. C., Woese, C. R., & Ouzounis, C. A. (1996). KOW: A novel motif linking a bacterial transcription factor with ribosomal proteins. *Trends in Biochemical Sciences, 21*, 425–426.

Li, S. S., Xiong, B., Xu, Y., Lu, T., Luo, X. M., Luo, C., … & Shen, J. (2014). Mechanism of the all-α to all-β conformational transition of RfaH-CTD: Molecular dynamics simulation and markov state model. *Journal of Chemical Theory and Computation, 10*, 2255–2264. doi:10.1021/ct5002279

McGuffin, L. J., Bryson, K., & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics, 16*, 404–405.

Merlino, A., Vitagliano, L., Antoine Ceruso, M. A., Di Nola, A., & Mazzarella, L. (2002). Global and local motions in ribonuclease A: A molecular dynamics study. *Biopolymers, 65*, 274–283. doi:10.1002/bip.10225

Merlino, A., Vitagliano, L., Ceruso, M. A., & Mazzarella, L. (2003). Subtle functional collective motions in pancreatic-like ribonucleases: From ribonuclease A to angiogenin. *Proteins: Structure, Function, and Bioinformatics, 53*, 101–110. doi:10.1002/prot.10466

Myers, J. K., Pace, C. N., & Scholtz, J. M. (1997). Helix propensities are identical in proteins and peptides. *Biochemistry, 36*, 10923–10929. doi:10.1021/bi9707180

Svetlov, V., & Nucler, E. (2012). Unfolding the bridge between transcription and translation. *Cell, 150*, 243–245. doi:10.1016/j.cell.2012.06.025

Tomar, S. K., & Artsimovitch, I. (2013). NusG-Spt5 proteins – Universal tools for transcription modification and communication. *Chemical Reviews, 113*, 8604–8619. doi:10.1021/cr400064k

Tomar, S. K., Knauer, S. H., NandyMazumdar, M., Rosch, P., & Artsimovitch, I. (2013). Interdomain contacts control folding of transcription factor RfaH. *Nucleic Acids Research, 41*, 10077–10085. doi:10.1093/nar/gkt779

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. (2005). GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry, 26*, 1701–1718. doi:10.1002/jcc.20291

Werner, F., & Grohmann, D. (2011). Evolution of multsubunit RNA polymerases in the three domains of life. *Nature Reviews Microbiology, 9*, 85–98. doi:10.1038/nrmicro2507

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. (2005). GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry, 26*, 1701–1718. doi:10.1002/jcc.20291