ORIGINAL PAPER

Role of methyl groups in dynamics and evolution of biomolecules

**Jonathan D. Nickels · Joseph E. Curtis · Hugh O'Neill ·
Alexei P. Sokolov**

Abstract Recent studies have discovered strong differences between the dynamics of nucleic acids (RNA and DNA) and proteins, especially at low hydration and low temperatures. This difference is caused primarily by dynamics of methyl groups that are abundant in proteins, but are absent or very rare in RNA and DNA. In this paper, we present a hypothesis regarding the role of methyl groups as intrinsic plasticizers in proteins and their evolutionary selection to facilitate protein dynamics and activity. We demonstrate the profound effect methyl groups have on protein dynamics relative to nucleic acid dynamics, and note the apparent correlation of methyl group content in protein classes and their need for molecular flexibility. Moreover, we note the fastest methyl groups of some enzymes appear around dynamical centers such as hinges or active sites. Methyl groups are also of tremendous importance from a hydrophobicity/folding/entropy perspective. These significant roles, however, complement our hypothesis rather than preclude the recognition of methyl groups in the dynamics and evolution of biomolecules.

J. D. Nickels · A. P. Sokolov
Joint Institute for Neutron Sciences, Oak Ridge National Laboratory, Oak Ridge,
TN 37831, USA

J. D. Nickels (✉) · A. P. Sokolov
Department of Chemistry, The University of Tennessee, Knoxville, TN 37996–1600, USA
e-mail: jnickels@alumni.nd.edu

A. P. Sokolov
e-mail: sokolov@utk.edu

J. E. Curtis
NIST Center for Neutron Research, 100 Bureau Drive, Gaithersburg, MD 20899, USA

H. O'Neill
Energy and the Environment Group, Biology and Soft Matter Division, Oak Ridge National
Laboratory, Oak Ridge, TN 37831, USA

1 Introduction

It is recognized that atomic and molecular motions (dynamics) are crucial to the function of biological macromolecules [1–5]. Yet our understanding of microscopic details of molecular motions, their relationship to chemical structure, and how they translate into the function of biomolecules remain very limited. Earlier analysis revealed that the dynamics of proteins at ambient conditions are faster than dynamics of RNA and DNA. Moreover, dynamics of RNA and DNA slow down with cooling much more rapidly than the dynamics of proteins [6]. This difference appears especially strong at lower hydration levels [6, 7]. Detailed experimental and computational studies ascribed this difference to dynamics of methyl groups, which are abundant in proteins and essentially are absent in tRNA [6, 8–11]. Methyl groups have relatively low energy barriers for rotation and, as a result, methyl group dynamics in proteins remain active, even at low hydration levels or low temperatures. This facilitates dynamics within the entire protein and allows the molecules to access a more diverse set of configurations and to sample various conformational states, thus increasing the conformational entropy.

The unexpected observation that methyl groups indeed facilitate the dynamics of proteins, calls our attention to the importance methyl groups might play in facilitating the associated biological functions. In this paper, we formulate a hypothesis (sketched in Fig. 1) based on the implications of these recent studies. We postulate that solvent-independent methyl dynamics were advantageous in the evolution of biomolecules from an RNA-dominated world, lacking methyl dynamics, to the protein/RNA/DNA life we see today. We present experimental evidence showing the difference in dynamics of RNA and
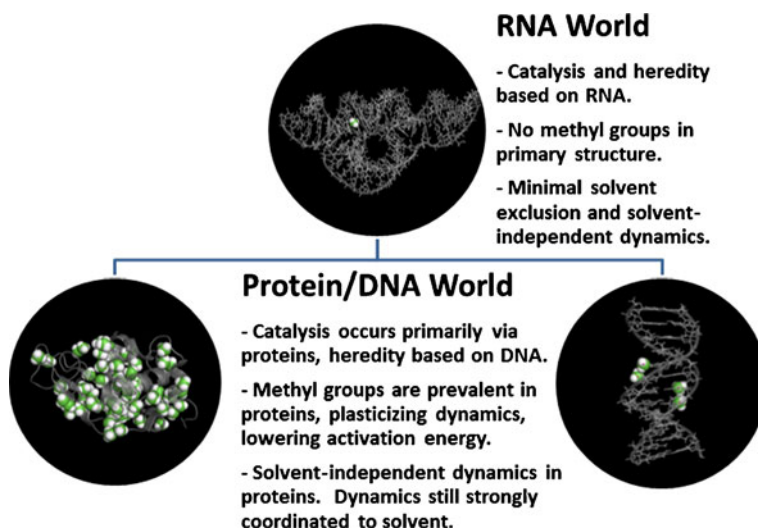

RNA World

- Catalysis and heredity based on RNA.

- No methyl groups in primary structure.

- Minimal solvent exclusion and solvent-independent dynamics.

Protein/DNA World

- Catalysis occurs primarily via proteins, heredity based on DNA.

- Methyl groups are prevalent in proteins, plasticizing dynamics, lowering activation energy.

- Solvent-independent dynamics in proteins. Dynamics still strongly coordinated to solvent.

Fig. 1 Summary of the hypothesis highlighting the evolutionary trend from RNA-based life, low in methyl content and lacking methyl plasticized dynamics, to protein and DNA-based life that has numerous methyl groups and solvent-independent methyl dynamics

proteins, an analysis of methyl group content in biological macromolecules from various classes of proteins and nucleic acids, and analyze positions of the fastest methyl groups in several proteins with respect to the active site. The results of the presented analysis are consistent with the formulated hypothesis. Thus, if true, the hypothesis creates a new way to look at the role of individual structural units (e.g., side groups) in dynamics, activity, and evolution of biological macromolecules.

2 Materials and methods

Neutron scattering data discussed for lysozyme and tRNA were measured with the High-Flux Back-Scattering instrument at the National Institute of Standards and Technology and previously published [7]. Similarly, for DNA measurements, we used the previously-published data obtained using the back-scattering spectrometer at the Research Center Julich in Germany [12]. Details of the measurement and sample preparation can be found in these references. Details of molecular dynamics simulations of lysozyme are presented in Roh et al. [13].

Analysis of methyl content in various classes of biomolecules is based on the number of methyl groups in each molecule, divided by the molecular weight. The content for RNA and DNA was calculated using the average value for each of the primary four nucleotides, whereas the content for protein was calculated as an average of the 78 proteins listed in the database contained in the Supplementary Information. The comparative analysis of methyl content in various classes of proteins was expressed as a ratio of methyl group content to total number of residues. Twenty-three ribosomal proteins, 43 enzymes, and 12 extracellular matrix proteins were analyzed. The exact proteins used are listed in the Supplementary Information.

3 Results and discussion

The mean-squared atomic displacements ($<r^2>$) on the $\sim$1 ns time scale, measured with neutron scattering, demonstrate that $<r^2>$ at biological temperatures ($\sim$300 K) are approximately three times larger in dehydrated proteins than in dehydrated tRNA (Fig. 2). The relaxation dynamics of these biomolecules, seen in Fig. 3, appear in scattering measurements as a quasielastic intensity, or a broadening of the central peak. The direct comparison of the neutron scattering spectra of proteins, RNA and DNA shows significant differences at low hydration levels and low temperatures. While protein spectra exhibit significant quasielastic intensity, the spectra of RNA and DNA under the same conditions present just the resolution function. The latter means that the dynamics of RNA and DNA on the ns time scale appear essentially frozen at these conditions [6, 7, 9]. These data show the presence of low activation-energy dynamic processes in the protein molecules, and not in the nucleic acids.

These processes are the rotation and fluctuations of methyl groups, the lowest activation-energy dynamic processes in biological macromolecules. Our observation on methyl group dynamics suggests another evolutionary advantage of proteins—their greater diversity of dynamical processes and the lower energies required for conformational changes. Our initial investigation of this hypothesis is based on the direct comparison of the number of methyl groups in different classes of biomolecules. We will not consider here post-transcriptional
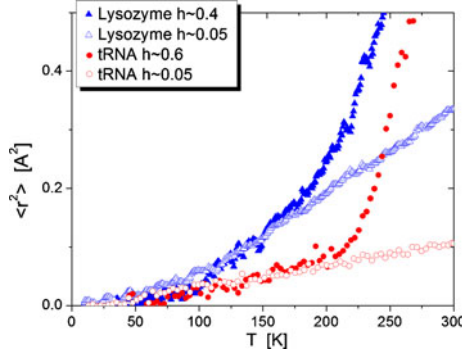
Fig. 2 Differences in the dynamics of RNA and proteins. Mean squared atomic displacements on a ~1 ns time scale. Comparing hydrated and dry lysozyme (*filled and open triangles*, respectively) to hydrated and dry tRNA (*filled and open circles*). Hydration levels, h, are in g of water per 1 g of biomolecules. The solvent-independent, methyl-group dynamic onset in protein appears at ~100 K and the subsequent hydration-dependent onset at ~200 K in both RNA and proteins (data from [7])

Fig. 3 Quasielastic neutron scattering spectra of lysozyme, tRNA and DNA (*symbols*); *lines* show the resolution function of spectrometers. Data for lysozyme and tRNA are from [7], for DNA from [20]. Lysozyme spectra exhibit relaxation contribution even in the dry state and at T = 200 K, while no quasielastic scattering is observed for dry tRNA and DNA, or for hydrated tRNA and DNA below 210 K
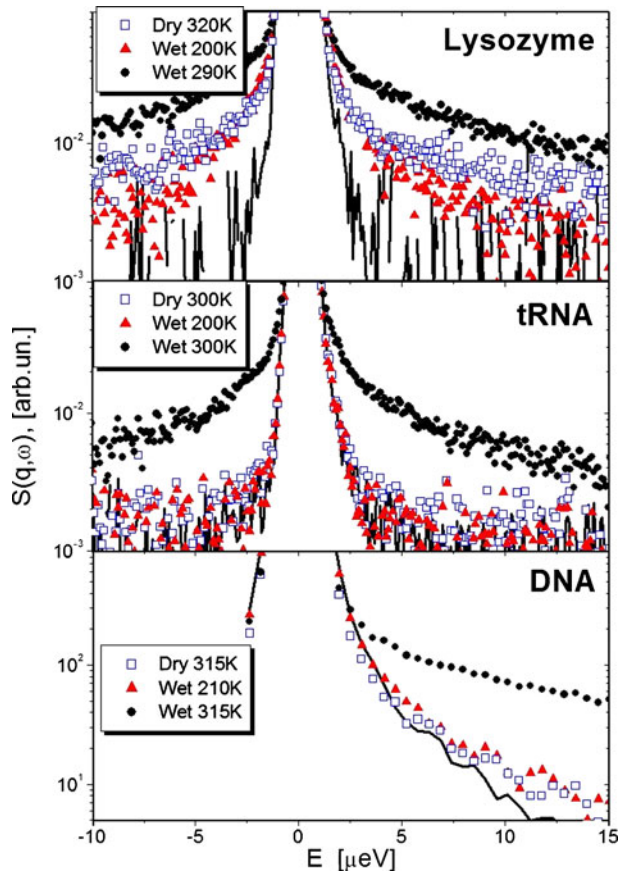
methylations of RNA and DNA, although it is conceivable that methylations of nucleic acids such as the cap of mRNA will also have important functions in solvent-excluded regions (such as within the ribosome). There are no methyl groups in the primary structure of RNA while one of DNA's four nucleic acids contains a single methyl group. Six of the 20 standard amino-acids found in proteins, however, contain methyl groups. Clearly, the number of methyl groups in proteins is significantly higher than in DNA or RNA (Fig. 4).

It is interesting therefore to consider that the amount of methyl groups in biological macromolecules might reflect the evolutionary process. The RNA world hypothesis [14] proposes that RNA preceded proteins and DNA as the molecule of life. This was spawned by the discovery of enzymatic RNAs in the 1980s [15], leading to Gilbert's theory [14]. RNA-based organisms would have fulfilled the fundamental requirements for life by being both self-replicating and metabolically/catalytically active. The prevailing thought is that the hereditary fidelity of DNA was superior to RNA and proteins are better catalytic agents than ribozymes, and recent work supports this assertion [16]. Why then did RNA also lose its role as primary catalytic agent? The diversity of chemical groups is one of the obvious benefits of proteins over RNA as catalysts. Chemically and structurally, the 'repertoire' of biomolecules was expanded from four bases, and a backbone, with functional groups highly interactive with the solvent and each other, to 20 distinct side chains which contain a far wider array of functionality and interaction with each other, potential ligands, and the aqueous environment. However, this also introduced large regions of solvent exclusion and the molecular motions contained within these regions should not be ignored. Such solvent-independent dynamics are typified by methyl group dynamics, which we again point out, are absent in RNA. One can envision how such flexibility (or lack thereof) could translate into the energy barrier involved in enzymatic action. In material terms, we suggest that methyl groups are plasticizing agents in proteins, facilitating dynamics and, subsequently, activity.
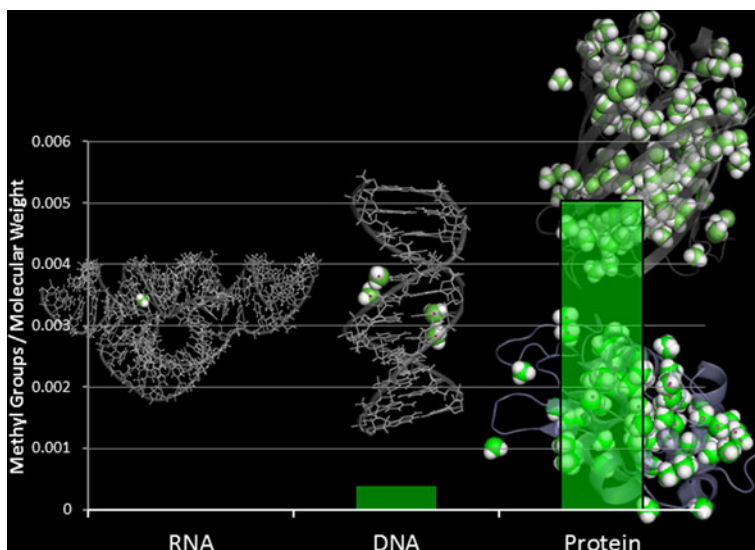

Fig. 4 Methyl content in different classes of biomolecules, presented in number of methyl groups per molecular weight. Post-translational and post-transcriptional modifications have been disregarded

Consideration of this hypothesis suggests that we should look at the role of methyl dynamics in various proteins. For example, traditional model enzymes such as lysozyme and ribonuclease A, and transport proteins such as myoglobin and hemoglobin, have significant numbers of methyl groups in their primary structure, ~0.6 methyl groups per residue, while the structural proteins such as collagen have as few as ~0.3 methyl groups per residue (Fig. 5). This suggests a connection between the required functional flexibility or dynamic activity and the amount of methyl groups in the protein. To investigate this hypothesis further, we have averaged the quantity of methyl groups in a sample of ribosomal, enzymatic, and extracellular proteins (Fig. 5; a detailed list of proteins is presented in the Supplementary Information). Indeed, proteins involved in extracellular matrix roles (collagen, laminin, etc.) have as many as 25% fewer methyl groups per residue. This result is potentially significant: If methyl groups are acting to plasticize proteins and to lower energy barriers for catalytic activity, then it is logical to expect that proteins which function enzymatically would be enriched in plasticizers (methyl groups), and proteins functioning structurally would be depleted in methyl groups.

More detailed analysis focuses on the fact that not all methyl groups are equal in terms of dynamics. Molecular dynamics simulations and NMR data reveal [11, 17–19] that among the methyl groups the methionine methyl is the most dynamic (i.e., has the lowest energy barrier for rotation), followed by leucine and then the delta methyl group on isoleucine. Of course, their individual energy barriers depend on the position of the residue in the protein 3-D structure. Nevertheless, on average, they remain the fastest, and we will refer to them as 'highly mobile methyl' groups. Analysis of the content of highly mobile methyl groups (Fig. 5) reveals that the extracellular matrix proteins have on average ~32% less highly mobile methyl groups than enzymes.
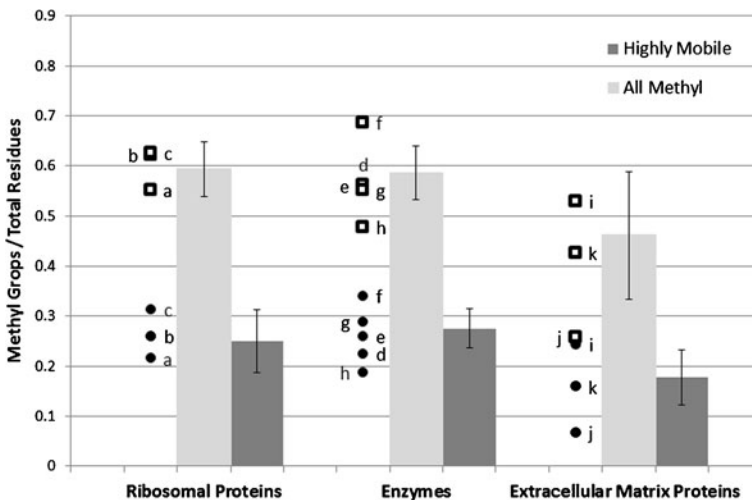

Fig. 5 Number of methyl groups and highly mobile (Ile δ, Leu, Met) methyl groups per residue for selected protein classes. *Error bars* depict the standard deviation. Representative data points have been included; (**a**) 40S ribosomal protein S11, (**b**) 40S ribosomal protein S17, (**c**) 60S ribosomal protein L30, (**d**) glucose oxidase, (**e**) calmodulin, (**f**) adenylate cyclase, (**g**) dihydrofolate reductase, (**h**) lysozyme, (**i**) laminin α-chain, (**j**) collagen, and (**k**) plasminogen

Detailed analysis of distribution of methyl groups in a 3-D structure of biomolecules has only been done for a few proteins. It has been noted that the highly mobile methyl groups in lysozyme are buried near the active site of the enzyme, thus facilitating dynamics in the most functionally important part of the molecule [20]. We illustrate this point in Fig. 6, which shows the spatial distribution of methyl groups in lysozyme. We have superimposed the methyl carbons (red spheres) over the translations associated with binding the substrate (red lines). We can see the abundance of methyl groups in the area centered on this hinge motion (lower right quadrant). Krishnan et al. also analyzed the dynamics of myoglobin using computer simulations [8] and found that the fastest methyl groups are placed in the vicinity of the xenon cavities in the protein. Xenon cavities in myoglobin are crucial for substrate migration to and from the active site. These correlations suggest that the fastest methyl groups might be directly involved in, if not critical to, the function of this protein. Experimental studies of the influence of methyl group deleting mutations on the activity of calmodulin [13] present even more direct evidence of the role of methyl-containing residues in the protein's function. It was observed that removal of the methyl groups found in the latch region of calmodulin resulted in a significant reduction of activity [20], independent of substrate binding. This observation too is consistent with our hypothesis that methyl groups, and their associated dynamics are critical to protein activity.

The plasticization of proteins by the presence of methyl groups may also affect the existence of stable tertiary structures in proteins compared to those observed in RNA. Dynamically, methyl protons are the largest contributor to the observed amplitude of motion in proteins. One can speculate that methyl groups influence both the folding process through solvent exclusion and the attainment of well-defined, thermodynamically stable structures. The motion of methyl groups in proteins contributes a significant amount of entropy that stabilizes the structure of proteins beyond that which is achievable in RNA. Thus methyl groups provide both the well-recognized hydrophobic contribution to protein folding and an additional entropic contribution to the Gibbs free energy. From an evolutionary perspective, the inclusion of methyl groups may give proteins the dual advantage over RNA of both
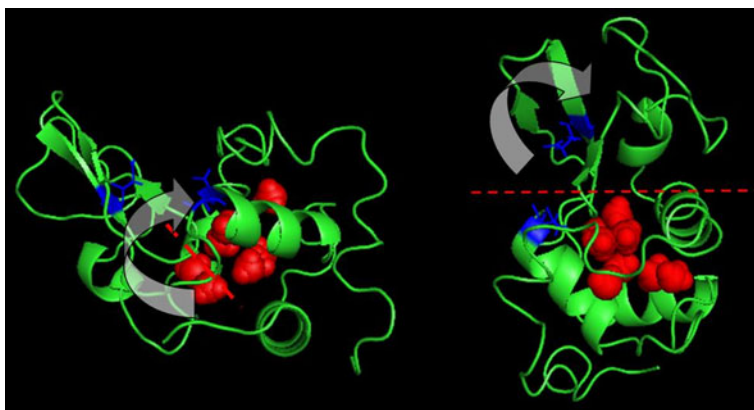

Fig. 6 Distribution of the most dynamically active methyl groups in lysozyme (*red spheres*), based on simulation and defined as the methyl groups with the lowest activation barrier to rotation. The protein is presented along the hinge axis and rotated 90 degrees. The *blue residues* are the critical active residues (E35 and D52). We direct the reader's attention to the proximity of these most dynamic methyl groups and the axis of the hinge bending

facilitating dynamically required motions and providing the thermodynamically favorable species to expedite protein folding and/or stabilize the uniquely defined equilibrium structures. We note that stabilization of RNA structure is quite different from that of proteins due to the significant role of charge interactions in RNA.

It is also known that modern nucleic acid biomolecules are often post-transcriptionally methylated. We have no way of knowing if ancient RNA-based life was capable of utilizing methylation, but it does exist in modern organisms. DNA methylation is a common mechanism of gene deactivation in higher animals, and is one of the most common post-transcriptional modifications of RNA products. Methylations have been reported in rRNA, tRNA, and in the $3'$ cap of mRNA. The dynamics of these molecules have not been investigated as closely as proteins to this point but we hypothesize that methyl groups act as intrinsic plasticizers in these systems as well. The role of post-transcriptional methylation is an interesting topic; however, it is beyond the scope of the current paper, and is a subject for future work.

4 Conclusions

We have presented a hypothesis that early life based on RNA for both catalysis and gene storage lacked methyl groups, and subsequently, lacked access to the low-energy dynamic processes provided by methyl groups. These processes plasticize the dynamics of proteins due to their low activation energies and remain active even at low hydrations and temperatures, particularly important for extreme conditions. This is a new perspective with which to consider the progression of biomolecule evolution: A lowering of activation energy required to initiate dynamic processes associated with catalysis provided an advantage to organisms functioning with early proteins. We suggest that this be added to the list of the driving forces of evolution from RNA, for both storage of genetic material and metabolic catalysis, to DNA for storage and proteins for the vast majority of catalysis. Finally, we suggest a comparative study of ribozyme and enzyme dynamics. We hypothesize that the dynamical differences will be strongly connected with the energetic efficiency of their catalytic activity, illustrating what we interpret as an evolutionary advantage.

References

1. Beece, D., Eisenstein, L., Frauenfelder, H., Good, D., Marden, M.C., Reinisch, L., Reynolds, A.H., Sorensen, L.B., Yue, K.T.: Solvent viscosity and protein dynamics. Biochemistry **19**, 5147–5157 (1980)
2. Eisenmesser, E.Z., Bosco, D.A., Akke, M., Kern, D.: Enzyme dynamics during catalysis. Science **295**, 1520–1523 (2002)
3. Eisenmesser, E.Z., Millet, O., Labeikovsky, W., Korzhnev, D.M., Wolf-Watz, M., Bosco, D.A., Skalicky, J.J., Kay, L.E., Kern, D.: Intrinsic dynamics of an enzyme underlies catalysis. Nature **438**, 117–121 (2005)
4. Henzler-Wildman, K., Kern, D.: Dynamic personalities of proteins. Nature **450**, 964–972 (2007)
5. Tolman, J.R.: Structural biology: protein dynamics from disorder. Nature **459**, 1063–1064 (2009)

6. Khodadadi, S., Roh, J.H., Kisliuk, A., Mamontov, E., Tyagi, M., Woodson, S.A., Briber, R.M., Sokolov, A.P.: Dynamics of biological macromolecules: not a simple slaving by hydration water. Biophys. J. **98**, 1321–1326 (2010)

7. Caliskan, G., Briber, R.M., Thirumalai, D., Garcia-Sakai, V., Woodson, S.A., Sokolov, A.P.: Dynamic transition in tRNA is solvent induced. J. Am. Chem. Soc. **128**, 32–33 (2006)

8. Krishnan, M., Kurkal-Siebert, V., Smith, J.C.: Methyl group dynamics and the onset of anharmonicity in myoglobin. J. Phys. Chem. B **112**, 5522–5533 (2008)

9. Roh, J.H., Briber, R.M., Damjanovic, A., Thirumalai, D., Woodson, S.A., Sokolov, A.P.: Dynamics of tRNA at different levels of hydration. Biophys. J. **96**, 2755–2762 (2009)

10. Roh, J.H., Novikov, V.N., Gregory, R.B., Curtis, J.E., Chowdhuri, Z., Sokolov, A.P.: Onsets of anharmonicity in protein dynamics. Phys. Rev. Lett. **95**, 038101 (2005)

11. Wood, K., Tobias, D.J., Kessler, B., Gabel, F., Oesterhelt, D., Mulder, F.A.A., Zaccai, G., Weik, M.: The low-temperature inflection observed in neutron scattering measurements of proteins is due to methyl rotation: direct evidence using isotope labeling and molecular dynamics simulations. J. Am. Chem. Soc. **132**, 4990–4991 (2010)

12. Sokolov, A.P., Grimm, H., Kisliuk, A.: Slow relaxation process in DNA. J. Biol. Phys. **27**, 313–327 (2001)

13. Vugmeyster, L., Ostrovsky, D., Ford, J.J., Lipton, A.S.: Freezing of dynamics of a methyl group in a protein hydrophobic core at cryogenic temperatures by deuteron NMR spectroscopy. J. Am. Chem. Soc. **132**, 4038–4039 (2010)

14. Gilbert, W.: Origin of life—the RNA world. Nature **319**, 618–618 (1986)

15. Guerrier-Takada, C.G.K., Marsh, T., Pace, N., Altman, S.: The RNA moiety of ribonuclease-P is the catalytic subunit of the enzyme. Cell **35**, 849–857 (1983)

16. Leu, K., Obermayer, B., Rajamani, S., Gerland, U., Chen, I.A.: The prebiotic evolutionary advantage of transferring genetic information from RNA to DNA. Nucleic Acids Res. **39**, 8135–8147 (2011)

17. Wong, K.B., Daggett, V.: Barstar has a highly dynamic hydrophobic core: evidence from molecular dynamics simulations and nuclear magnetic resonance relaxation data. Biochemistry **37**, 11182–11192 (1998)

18. Bajaj, V.S., van der Wel, P.C.A., Griffin, R.G.: Observation of a low-temperature, dynamically driven structural transition in a polypeptide by solid-state NMR spectroscopy. J. Am. Chem. Soc. **131**, 118–128 (2009)

19. Roh, J.H., Curtis, J.E., Azzam, S., Novikov, V.N., Peral, I., Chowdhuri, Z., Gregory, R.B., Sokolov, A.P.: Influence of hydration on the dynamics of lysozyme. Biophys. J. **91**, 2573–2588 (2006)

20. Su, Z.H., Fan, D.J., George, S.E.: Role of domain-3 of calmodulin in activation of calmodulin-stimulated phosphodiesterase and smooth-muscle myosin light-chain kinase. J. Biol. Chem. **269**, 16761–16765 (1994)