

Molecular Dynamics Simulation in RNA Interference

Xia Wang¹, Yonghua Wang², Lei Zheng^{1,*} and Jianxin Chen^{3,*}

¹School of Medical Engineering, Hefei University of Technology, Hefei, Anhui Province, 230009, China; ²College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, 712100, China; ³Beijing University of Chinese Medicine, ChaoYang District, Beijing, 100029, China

Abstract: RNA interference (RNAi) is a mechanism that utilizes small RNA molecules to silence gene expression after the gene has been transcribed. To understand the mechanisms of small RNA biogenesis, target nucleic acid recognition and cleavage, and how they are influenced by other regulators, one needs to know the structures and dynamics of the proteins or/and nucleic acids in these processes. Molecular dynamics (MD) simulation is a powerful tool for understanding motions and dynamics of macro-biomolecules at an atomic-scale via theoretical and empirical principles in physical chemistry. With its application to RNAi, an excellent overview of structural and dynamical mechanistic of RNAi processes has already emerged. In this review, we summarize the recent advances in MD simulations in the study of functional modules and their assemblies and target recognition and cleavage in RNAi processes. Additionally, we also present some perspectives on this technique.

Keywords: Molecular dynamics simulation, RNA interference, miRNA, siRNA, biogenesis.

1. INTRODUCTION

RNAi is a widespread eukaryotic regulatory process in which double-stranded RNA (dsRNA) triggers the silencing of cognate genes, and this process has been reviewed extensively [1-4]. In brief, this RNA-mediated genetic control mainly initiates from small interfering RNAs (siRNA) and microRNAs (miRNA), which are 20–30-nt in length [5,6] (Fig. 1). In the cell, dsRNA precursors are cleaved by RNase III enzyme Dicer into short dsRNA with 2-nt overhangs at the 3'-end. Then, one strand of the siRNA is incorporated into RNA-induced silencing complex (RISC) for selecting target mRNAs by base pairing. Subsequently, the mRNA is cleaved by RISC and thus silenced. As for endogenous miRNA pathway, it starts in nucleus, where long primary transcripts (pri-miRNAs) are cleaved by the RNase III enzyme Drosha and its cofactor DGCR8 (microprocessor complex) into precursor miRNA (pre-miRNA) hairpins of about 70-nt with a 2-nt 3' overhang. The pre-miRNAs are then transported to the cytoplasm by Exportin-5 and cleaved by Dicer to make miRNA-sized fragments. One strand of the fragment is then incorporated into the RISC complex. In contrast to siRNAs, miRNA shows imperfect sequence complementarity to a target mRNA and results in the inhibition of translation.

To deep understand functional modules and their assemblies in RNAi pathway, many structures of proteins and their nucleic acid complexes, such as Argonaute (Ago), PIWI, Tudor-SN Dicer, Drosha, and DGCR8, involved in RNA biogenesis have been determined, providing valuable

insights into the molecular mechanisms of nucleic acid recognition and cleavage [7,8]. Since functional modules in RNAi, such as recognition, traffic and catalysis mechanisms, are always kinetic and dynamic processes. Since crystallographic structures represent only the average conformation for a particular condition, thus important information regarding the mobility of protein and nucleic acid has received benefit from other experimental methods, such as nuclear magnetic resonance (NMR) spectroscopy [9], and small-angle scattering (SAS) with X-ray or neutrons [10]. However, the expense and extensive labor of them have urged many to seek computational approaches, such as molecular dynamics (MD) simulation [11,12], that can predict protein motions.

2. OUTLINE OF MD SIMULATIONS

MD simulation is a computational method based on empirically derived potential energy functions, which presents an understanding of the relationship between protein or nucleic acid structure and dynamics by numerically solving Newton's equations of motions [13]. This method reduces the complexity and computational resources of the calculations that are required to describe the absurd quantum-mechanical motions and chemical reactions of large molecular systems. The process of MD simulation consists of the following steps (Fig. 2): (1) The starting conditions are the positions of the atoms and their velocities. The initial positions of the atoms are prepared from a known crystal structure or previously undescribed three-dimensional (3D) structures constructed by homology modeling techniques (when a reported structure of a homolog is available); their velocities are generated from random numbers and scaled to the desired temperature. (2) Based on the initial positions and velocities, parameters of mathematical functions describing the

*Address correspondence to this author at the School of Medical Engineering, Hefei University of Technology, Hefei, Anhui Province, 230009, China; Tel: +86-551-2901285; E-mail: lzheng@hfut.edu.cn, cjx@bucm.edu.cn

potential energy of a system, termed the “force field”, are set to simulate the movements of atoms and molecules. Typically used force fields for proteins and nucleic acids, such as the “AMBER” [14,15] and “CHARMM” [16] force fields, have the formulae of covalent bonds, angles, dihedrals, van der Waals, and electrostatic potentials like that shown in Fig. (3). All atoms are assumed to be the charged spheres; chemical bonds and atomic angles are modeled as harmonic springs, and dihedral angles are defined by a sinusoidal function that approximates the energy differences between eclipsed and staggered conformations; non-bonded forces arise from van der Waals interactions are modeled using the Lennard-Jones potential [17]; the electrostatic interactions are computed based on Coulomb’s law. The system has to be calculated many times in the course of a full simulation, usually for many thousands of time steps. (3) The trajectory needs to be investigated by multiple analysis tools, such as principle component analysis (PCA) and dynamic cross correlation map (DCCM), to understand the structural dynamics of proteins or/and nucleic acids on timescales of nanoseconds to microseconds. (4) We should summarize the main findings for depicting the unique structural features and biological significance of proteins or/and nucleic acids [13,18,19].

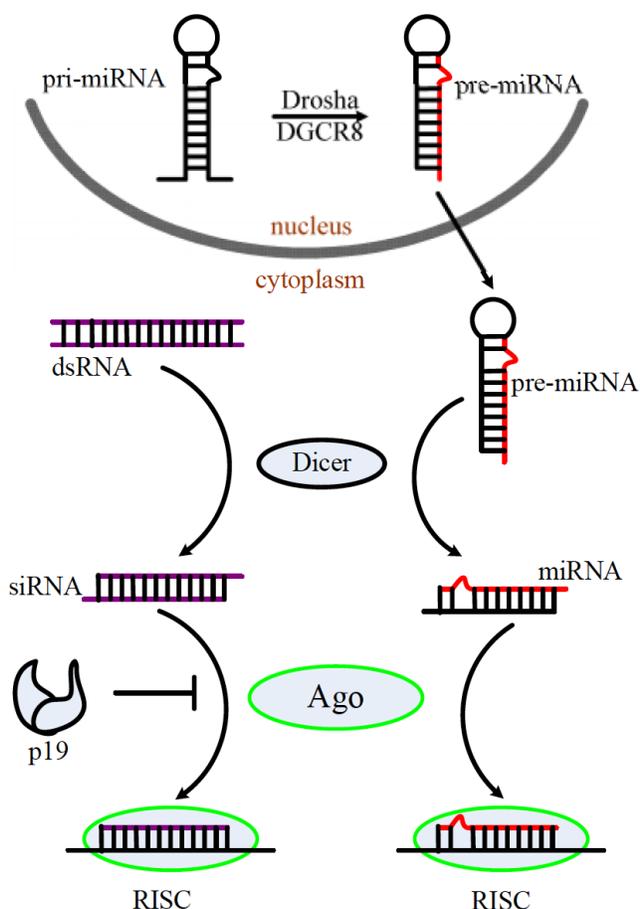


Fig. (1). Major RNA interference pathways mediated by small RNAs. pri-miRNA are processed by Drosha-DGCR8 in the nucleus to generate pre-miRNA, then this product is transported into cytoplasm by Exportin-5 and processed by Dicer. siRNAs start as dsRNA and are processed by Dicer. In both cases the end product 20–30-nt RNA is incorporated into the RISC complex.

3. MD SIMULATION IN RNAI

In this section, we will summarize applications of MD simulations in a series of RNAi processes.

3.1. Flexibility of dsRNA-Binding Domain (dsRBD)

The ~70-residue dsRBD is present in many proteins involved in RNAi pathway, such as Drosha, DGCR8, and Dicer. Besides, dsRBDs usually coexist with the RNase III domains in RNase III proteins [20,21], suggesting a functional role in binding and/or cleavage for the dsRBD. Thus understanding the flexibility of dsRBD and how it facilitates dsRBD to recognize dsRNA is undoubtedly germane to RNAi. Structures of several dsRBDs have been determined. They all comprise an α helix, followed by three β strands forming an antiparallel β sheet, and terminating with a second α helix ($\alpha\beta\beta\alpha$ fold) [22] (Fig.4). Showalter and colleagues applied three 250 ns MD simulations to explore the correlated motions of two dsRBDs from DGCR8 [23]. By using isotropically distributed ensemble (IDE) analysis [24], they found that the two domains reorient mutually, with the well-ordered linker and C-terminus serving as a pivot. Their results indicate that the conformation of the RNA-binding surface is dynamically influenced by motions at the domain interface, this probably provides an adaptive separation distance to accommodate various pri-miRNAs with heterogeneous structures.

Interestingly, they performed electrophoretic mobility shift assays (EMSAs) of pri-miR-16-1 binding by Drosha-dsRBD (1259-1337) and DGCR8-dsRBD1 (505-583) and found that Drosha-dsRBD does not bind this representative pri-miRNA but isolated DGCR8-dsRBD1 does [25]. Motivated by this finding, they combined NMR spin relaxation and MD simulations to exploit dynamic origins of differential RNA binding function in two dsRBDs from two proteins. The results displayed that, although loop 2 in both dsRBDs is extremely dynamic (generalized order parameter S^2 of loop 2 is about 0.6–0.65, while the rest is about 0.9), the pattern of the correlations for the two proteins is different. Furthermore, the extended loop 1 of Drosha-dsRBD is more flexible than that of DGCR8-dsRBD1 but has no correlation with loop 2. This finding potentially interprets the lack of dsRNA binding by Drosha-dsRBD in the absence of the RNase III domains. Subsequently, with same approaches, they found that the isolated Dicer-dsRBD can bind dsRNA and present similar backbone dynamics to DGCR8-dsRBD1, which are different from those in Drosha-dsRBD [26]. This binding is only minimally influenced by the length of RNA or by a single-stranded–double-stranded junction caused by large terminal loops observed in miRNA precursors. Therefore, the Dicer-dsRBD contributes directly to substrate binding but not to selection between pre-miRNA and pre-siRNA. Combined, these results highlight key structural and dynamic features of dsRBDs that contribute to the binding mechanism of these domains to dsRNA.

3.2. Pre-miRNA Recognition by lin28

After processed by microprocessor, pri-miRNA becomes a pre-miRNA which can be regulated by some regulators. For instance, Lin28 (or homolog Lin28B) blocks the processing of precursor of let-7 family miRNAs thus inhibiting let-7

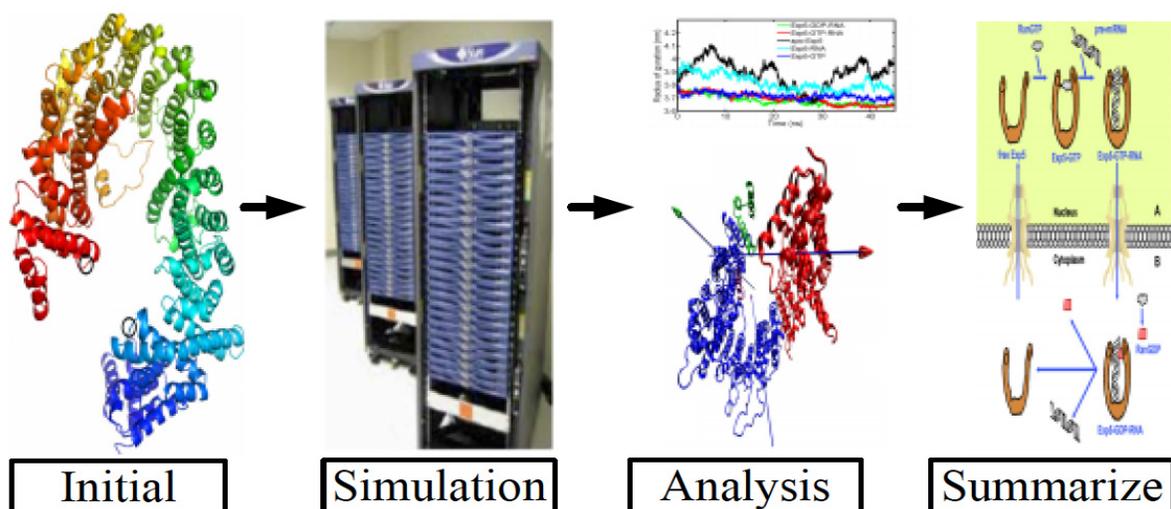


Fig. (2). The molecular dynamics simulation process. Firstly, an initial 3D structure of the system is prepared. Secondly, a potential energy function is used to estimate the forces acting on each of the system atoms. The positions of the atoms in the system are moved according to Newton's equations of motion. The simulation time is advanced, and the process is repeated many times. Thirdly, the trajectory needs to be investigated by many analysis tools to probe the dynamics of the system. Finally, the main findings should be summarized for depicting the unique structural features and biological significance of the system.

$$U(\vec{R}) = \underbrace{\sum_{bonds} k_i^{bond} (r_i - r_0)^2}_{U_{bond}} + \underbrace{\sum_{angles} k_i^{angle} (\theta_i - \theta_0)^2}_{U_{angle}} + \underbrace{\sum_{dihedrals} k_i^{dihedral} [1 + \cos(n_i \phi_i + \delta_i)]}_{U_{dihedral}} + \underbrace{\sum_i \sum_{j \neq i} 4 \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}}_{U_{nonbond}}$$

Fig. (3). Potential energy function of biopolymer. The first three terms are the harmonic bond, angle and dihedral potentials, respectively. The fourth term is the nonbond potential. r governs bond stretching; θ represents the bond angle; ϕ gives the dihedral angle.

maturation [27-29]. However, how pre-miRNA is recognized by Lin28B dynamically and which residues of Lin28B are involved in binding to pre-miRNA remain elusive. Lin28 proteins contain a cold shock domain (CSD) followed by two consecutive CCHC type zinc finger domains forming a zinc knuckle domain (ZKD) [30]. By homology modeling and MD simulations, Syed Ali and colleagues [31] found that ZKD is a disordered region and it can bind to the let-7g GGAG motif with the help of zinc, which promotes the CCHC type zinc finger formation. The flexibility of ZKD facilitates the binding of Lin28B to the let-7 family of miRNAs, with a shorter or longer loop preceding the GGAG motif. Moreover, the pi stacking and H-bonding are central to interaction between ZKD of Lin28B and pre-let-7g GGAG motif that is required for recognition by the ZKD. In addition, interactions of G2, G3 and G4 to the CSD of Lin28B could be pivotal in recognizing the pre-let-7g stem loop.

3.3. Pre-miRNA Binding and Export by Exportin-5

Pre-miRNAs are generated in the cell nucleus and then needed to be exported to cytoplasm for further maturation. This translocation process is mediated by Exportin-5 which is a RanGTP-dependent dsRNA-binding receptor [32,33].

However, the thermodynamic and kinetic mechanisms of this pathway are still needed to be further investigated. MD simulations have been used to study the basis of the structural dynamics that allow pre-miRNA to be recognized, protected and transported by Exportin-5 [12]. By the root mean square fluctuations (RMSFs), end-to-end distance and radius of gyration (R_g) analyses, the results indicated that free Exportin-5 structure intrinsically undergoes large conformational changes (as evidenced by the end-to-end distance increased from 5.6 to 6.9 nm, and R_g ranged from 3.66 to 4.3 nm) and fluctuations (~ 0.5 nm for the N and C termini of apo-Exportin-5). Indicating by PCA, Exportin-5 is divided into two segments which rotate and bend mutually, presenting an open conformation to accommodate RanGTP. RanGTP binds to Exportin-5 and connects its flexible N and C termini to form an overall shape complementarity, paving the way for pre-miRNA binding. Intriguingly, *in silico* mutation and H-bonding analyses indicate that pre-miRNA binds to Exportin-5 in a sequence-independent manner through flipping motion of the 2-nt 3' overhang and the widening of the minor groove of the stem. This explains that different pre-miRNAs even with no consensus sequences can all be recognized and exported by Exportin-5. By steered MD simulation and Molecular Mechanics/Generalized Born/Solvent Accessible Area (MM/GBSA) free energy

analysis, we observed that the existence of **Ran ensure Exportin-5** an appropriate structural geometry for pre-miRNA dismounting from the complex. Thus we speculate that RanGTP probably dissociates at the same time or after the release of pre-miRNA from receptor. In summary, this study offers a deep understanding of the nucleocytoplasmic translocation process of pre-miRNAs, especially on the coordination of interactions and the basis for molecular recognition and dissociation that underlies the process.

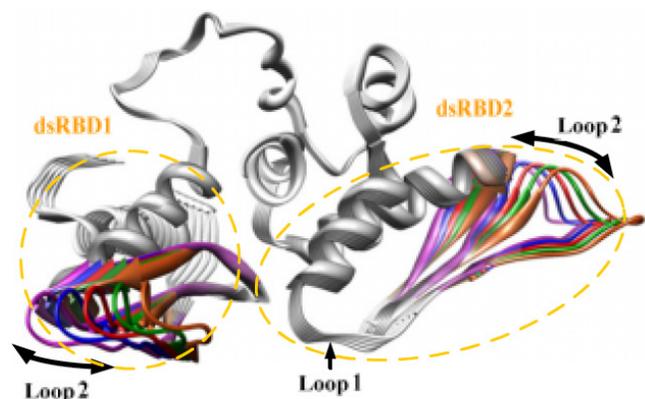


Fig. (4). Ribbon bundle from the Core simulation in which dsRBD2 and dsRBD1 are superimposed. Loop 2 displays maximal displacement. (This figure was adapted from a version originally created by Christopher Wostenberg.)

3.4. Flexibility of Dicer

Structure of eukaryotic Dicer contains a PAZ domain and two RNase III domains (RNase IIIa and RNase IIIb), a dsRBD, a domain of unknown function (DUF283) and a helicase domain [21]. The PAZ domain contains the RNA binding pocket that is recognized by 2-nt 3' overhang. It is connected with the RNase IIIa by a long α helix which is supported by a Platform domain formed by N-terminal residues. Thus the Platform domain may help in the spatial arrangement of PAZ and the RNase III domains (catalytic regions) which is critical for generating small RNAs of defined length. In a recent study [34], Sarzyńska et al examined the whole Dicer (from *Giardia intestinalis*) and the Dicer with a deleted Platform domain (Δ Plf Dicer) by MD simulation in an implicit solvent model with Generalized Born (GB) solvation model. By domain motion analysis through the DYNAMO program, the authors found that the Platform in Dicer clearly helps to maintain the spatial arrangement of the PAZ mainly by restricting its rotation about the axis passing the hinge region, while the removal of the Platform substantially changed the flexibility and arrangement of those domains. This result suggests that the implicit solvent methods can be applied to study of biomolecules, including large scale motion of protein and protein-nucleic acids complexes.

3.5. siRNA Recognition by Viral p19

The p19 protein encoded from tombusviruses suppresses RNAi in their infected hosts by sequestering siRNAs and preventing their incorporation into the RISC (Fig. 1) [35]. The p19 homodimer binds to the siRNA with very high affinity ($K_d = 0.17$ nM) [36]. Several experiments suggest that the two hydrophobic residues Trp39 and Trp42 are key

amino acids for p19 homodimer to contact and recognize the two siRNA termini, thus mutations on these two tryptophan residues of p19 will strongly suppress the RNA silencing [36,37]. To address the molecular details on how exactly the Trp39 and Trp42 residues interact and recognize the 21-nt siRNA dynamically, and how much binding affinity each Trp residue contributes, Xia et al performed standard MD and steered MD simulations of the wild-type and mutant p19 protein (Trp39 and Trp42Gly) in complex with a 21-nt siRNA duplex [38]. The authors found that the double mutation breaks the stacking interactions between Trp39-Trp42 of p19 and 5'- and 3'-terminal bases of siRNA in the wild-type, thus causing the double mutant structure much less stable than the wild-type. This result is further supported by steered MD simulations which indicate that the mutant p19-siRNA complex is "decompounded" very fast under a constant external force, while the wild-type remains largely intact under the same steering force. By free energy perturbation, they calculated a binding affinity of about 6.98 kcal/mol loss for the single mutation Trp39Gly, and about 12.8 kcal/mol loss for the double mutation Trp39/42Gly, with the van der Waals interactions dominating the contribution (~90%). These findings probably provide new insights to the interactions and co-evolution relationship between RNA virus proteins and their hosts.

Besides strong stacking interaction between 5'- and 3'-terminal bases of siRNA with end-capping tryptophans, p19 binding stability is also demonstrated to be sensitive to changes in pH. However, which specific amino acid residues involved remains unclear. Constant pH MD simulations have been applied to address this issue [39]. By predicting the variation of the free energy contribution from each titratable residue at various pH ranges, several key pH-dependent residues that affect *Carnation Italian* ringspot virus p19-siRNA binding stability have been identified. For examples, at high pH, the deprotonation of basic residues (Lys60, Lys67, Lys71, and Cys134) has the largest effect on the binding stability; while at low pH, deprotonation of acidic residues (Asp9, Glu12, Asp20, Glu35, and/or Glu41) decreases the binding stability. Moreover, the optimal pH range for siRNA binding is between 7.0 and 10.0.

3.6. Target Recognition and Cleavage

Following Dicer cleavage, mature miRNA or siRNA must be loaded into RISC, and then Ago, the core of RISC, utilizes siRNA or miRNA to target and cleave mRNA. Structurally, Ago is comprised of four domains: N-terminal, PAZ, MID, and PIWI domains. Except for the N domain, each has been ascribed with a specific function. Generally, the PAZ domain recognizes 3'-end of a guide RNA; the RNase H-like PIWI domain houses the slicing active site; MID domain forms a binding pocket for the 5'-phosphate of the guide strand [40,41].

Although the structures of Ago provide preliminary insights into how multiple domains within a single protein may cooperatively bind and cleave an RNA target, several questions remain. MD simulations of Ago have been used to address the dynamic behavior of the PAZ, MID and PIWI domains. The results indicate that PIWI and MID domains are relative rigid and static, whereas PAZ domain is intrinsically flexible [42]. Upon binding of siRNA or

miRNA, PAZ becomes more rigid and stable [42,43]. The previous NMR experiment [44] suggests RNA binding can induce significant conformational change in the PAZ domain. However, the mechanism by which folding is coupled to binding is poorly understood. By kinetic analysis of high-temperature MD simulations, the author found that, albeit both bound PAZ and apo-PAZ unfold via a two-state process, their unfolding pathways are distinct [43]. Notably, dynamics of Ago with and without RNA clearly show that each domain moves relative to the others, thereby opens and closes the boat-like shape of the nucleic acid binding channel. These inter-domain correlated movements not only facilitate the relaxation of the interactions between residues surrounding the RNA binding channel but also expand the cavity of the active site and open putative pathways for both the substrate uptake and product release [42].

Following RISC loading, Ago-guide strand complex binds its target mRNA through base pairing of the guide strand with target mRNA sequence. Full guide-target strand complementarity is needed for target mRNA cleavage through the endonucleolytic 'slicer' activity of Ago [45]. In the MD simulations of lin-4:lin-14 miRNA-mRNA complex (partially complementary) and lin-4:lin-14 siRNA-mRNA complex (fully complementary) [46], a kinked structure of miRNA complex induced by mismatch base pairs and bulge makes the scissile phosphate of the mRNA strand less accessible to the catalytic site as compared to the canonical helical structure of siRNA-mRNA complex. This finding suggests that mismatches between guide strand and the target mRNA decrease cleavage efficiency. Recently, MD simulations of Ago-miRNA:mRNA ternary complexes with introduction of G:U wobbles and mismatches at various positions of the miRNA 'seed' region have been reported [47]. The results, coupled with HITS-CLIP data, reveal that many modifications in the seed region are admissible and do not affect overall complex stability. However, with introduction of "damaging" C•C mismatches at different positions in the seed region of the DNA-RNA duplex (in Ago-DNA:mRNA ternary complex) [48], they found the complex can survive less than three such mismatches. Moreover, four-site mismatch mutant that intrigues severe distortions of the guide-target heteroduplex, can result in a bending motion of the PAZ domain along the pivotal of L1/L2 segment, leading to the opening of the nucleic-acid-binding channel. These findings indicate the central role of the seed region in the guide-target strands recognition. To further explore molecular basis for the target sequence selectivity in RNAi, thermodynamic integration (TI) simulations of all possible single-base mismatches in seed region [49] display that individual Ago residues located near the mismatched base pair contribute significantly to binding affinity changes due to single-base mismatches. Furthermore, several mismatched sequences of the miRNA have increased affinity to Ago, suggesting that some mismatches may decrease the probability for the escape of miRNAs from the silencing pathway. These results give a thermodynamic basis for RNAi target sequence selectivity. In another MD simulation [50], they found nucleotides in the 'seed' region have a stronger relationship with the protein than other nucleotides, and any of several seed region guide-target mismatches causes certain Ago residues to have modified correlations with the rest of the protein. These

results imply that the mismatch site communicates with effector sites (i.e., the active site, binding groove, and surface) of Ago through the cumulative effects of a large number of internal allosteric pathways.

After mRNA recruited by Ago protein, Ago, which contains a RNase H-like PIWI domain, slices mRNA strands in a Mg^{2+} -dependent behavior [51]. Structurally, three structural Mg^{2+} cations coexist in the complex: one (Mg1) bound to Gln414 embedded in the MID domain, one (Mg2) bound to Asp459, Ala460, and Asp635, and the third one (Mg3) coordinated by the catalytic triad (Asp459, Glu490 and Asp524) of PIWI domain. MD simulations [52] showed that Mg1 and Mg2 significantly enhance a locally cooperated movement of the PAZ, PIWI and MID domains, producing an "open-closed" motion between the PAZ and PIWI domains. Meanwhile, Mg3 globally stabilizes the whole Ago protein. Particularly, after mRNA taken up by Ago, MD simulations [42] show that the Mg3 interface in PIWI domain switches from closed to open conformations. This conformation switch makes the mRNA backbone approach to the catalytic center and form a stable octahedral geometry at the RNA-binding site. Furthermore, structural water molecules surrounding the Mg^{2+} -binding regions stabilize these ions, facilitating the target binding. For instance, inside the binding pocket of PIWI domain, one water molecule is critical to the conformational changes of Mg3 which contributes >10% of total ΔG for the binding of mRNA to Ago complex [42,52].

CONCLUSIONS AND PROSPECTS

Our current view of the molecular mechanisms of RNA silencing owes much to structural insights of RNAi-related proteins. However, siRNA/miRNA biogenesis, assembly with RISC and target recognition are very dynamic processes. The challenge is to understand how the different components work together harmoniously in RNAi pathway. For example, although FXR1 is recruited to assemble miRNA:mRNA and Ago2 complex under the condition of serum starvation [53,54], its exact role in this process is unknown. The MD simulation results reveal that the two strands of dsRNA are separated upon binding of Ago2 and FXR1, thus facilitating the disassemble of mRNA from the miRNA:mRNA in the polysomes which may improve the translation efficiency of mRNA [55]. Although the examples highlighted in this review demonstrate the power of modeling and MD simulation to decipher the dynamics as well as thermodynamics features of RNAi pathway, quite a few questions in RNAi pathway are still open: What is the mechanism of RanBP1 (or RanBP2) and RanGAP to release pre-miRNA-cargos efficiently from Exportin-5 and RanGTP? What are the functional and structural roles of additional RISC components that confer specificity and selectivity in different RISCs? What are the roles of water molecules in target recognition and cleavage? To address these problems by using MD simulation, it should be emphasized that adequate sampling needs to be performed to explore the conformational space of structures and assure convergence of the free energies of the systems.

Notably, among different approaches, atomistic classical MD simulation is mostly used in studying the mechanism of RNAi process. Though this method can principally sampled

many biological processes in which motions span from microseconds to seconds and tens of nanometers or more given sufficient computer resources (both software and hardware) [56,57], the utility of classical MD simulation in exploring the landscape of large scale conformational changes at atomistic detail poses a considerable computational challenge. Recently, to overcome this challenge, various enhanced sampling methods and simplified models have been used in combination with MD simulation to access larger scale motions with less expensive computation resource. For examples, a powerful method to overcome the inherent time limitation of MD simulations is utilizing a potential of mean force (PMF) calculation [58]. In this method an external or biasing potential, usually referred to as an umbrella potential, is employed to overcome large energy barriers along a reaction coordinate between conformational states. Another useful method is accelerated molecular dynamics (aMD) where large energy barriers are artificially reduced [59]. Although some artifacts are inevitably introduced into this process, it allows the system to shift between conformations that would not be accessible given the time scales of classic MD. Besides, coarse-grained (CG) models have also been introduced to overcome the time-scale limitations of conventional MD simulations. In all the CG models, polypeptide chains and nucleic acids are represented using a reduced description, for instances, three interaction site (TIS) model for a RNA hairpin obtained by representing each nucleotide by three sites one each for phosphate, ribose and base; C α -SCM model for polymers obtained by replacing each amino acid by two sites one centered on the α -carbon and the other at the center of the side chain [60]. It needs to be pointed out that significant benefit of CG models is that their conformational space can be exhaustively sampled. However, even with simplification, accurate results for thermodynamics can only be obtained using enhanced sampling methods. Thus multiple models and sampling methods can be used together to explore the dynamic features of bio-macromolecules.

Since the possible applications and shortcomings of MD simulations have been extensively reviewed elsewhere [13,60-62] and a detailed discussion is beyond the scope of this review. Here, we only surveyed recent progress in MD simulation, focusing on the mechanistic insights into siRNA/miRNA biogenesis, assembly with RISC and target recognition and cleavage. In the future, we expect that, with more high resolution structures determined, MD simulations (especially CG MD) are likely to play an increasingly important role in clarifying the mechanisms of RNAi.

CONFLICT OF INTEREST

The author(s) confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

This work was financially supported by the specialized Research Fund for the Doctoral Program of Higher Education (20120111110024), the Fundamental Research Funds for the Central Universities (2012HGCX0003), (2012HGZY0021), the National Key Technologies R&D Program (2012BAD07B01), the Funds for Huangshan Professorship

of Hefei University of Technology, National Science Foundation of China (Grant No. 81173463), New Century Excellent Talent Support Plan of the Ministry of Education (NCET-12-0804), and Beijing Nova Program (Grant No. xx2013032).

ABBREVIATIONS

MD	=	molecular dynamics
RNAi	=	RNA interference
dsRNA	=	double-stranded RNA
siRNA	=	small interfering RNAs
miRNA	=	microRNA
RISC	=	RNA-induced silencing complex
pre-miRNA	=	precursor miRNA
Ago	=	Argonaute
NMR	=	nuclear magnetic resonance
SAS	=	small-angle scattering
3D	=	three-dimensional
PCA	=	principle component analysis
DCCM	=	dynamic cross correlation map
dsRBD	=	dsRNA-binding domain
IDE	=	isotropically distributed ensemble
EMSAs	=	electrophoretic mobility shift assays
CSD	=	cold shock domain
ZKD	=	zinc knuckle domain
RMSFs	=	root mean square fluctuations
Rg	=	radius of gyration
MM/GBSA	=	Mechanics/Generalized Born/Solvent Accessible Area
PMF	=	potential of mean force
CG	=	coarse-grained
TIS	=	three interaction site

REFERENCES

- [1] Bushati, N.; Cohen, S.M. MicroRNA functions. *Annu. Rev. Cell. Dev. Biol.*, **2007**, *23*, 175-205.
- [2] Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **2004**, *116*, 281-297.
- [3] Tomari, Y.; Zamore, P.D. Perspective: machines for RNAi. *Genes. Dev.*, **2005**, *19*, 517-529.
- [4] Rana, T.M. Illuminating the silence: understanding the structure and function of small RNAs. *Nat. Rev. Mol. Cell. Biol.*, **2007**, *8*, 23-36.
- [5] Chapman, E.J.; Carrington, J.C. Specialization and evolution of endogenous small RNA pathways. *Nat. Rev. Genet.*, **2007**, *8*, 884-896.
- [6] Zamore, P.D.; Haley, B. Ribo-gnome: the big world of small RNAs. *Science*, **2005**, *309*, 1519-1524.
- [7] Sashital, D.G.; Doudnam, J.A. Structural insights into RNA interference. *Curr. Opin. Struct. Biol.*, **2010**, *20*, 90-97.
- [8] Nowotny, M.; Yang, W. Structural and functional modules in RNA interference. *Curr. Opin. Struct. Biol.*, **2009**, *19*, 286-293.
- [9] Dunn, W.B.; Broadhurst, D.I.; Atherton, H.J.; Goodacre, R.; Griffin, J.L. Systems level studies of mammalian metabolomes: the

- roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chem. Soc. Rev.*, **2011**, *40*, 387-426.
- [10] Hennig, J.; Wang, I.; Sonntag, M.; Gabel, F.; Sattler, M. Combining NMR and small angle X-ray and neutron scattering in the structural analysis of a ternary protein-RNA complex. *J. Biomol. NMR.*, **2013**, *56*, 17-30.
- [11] Xu, X.; Yang, W.; Wang, X.; Li, Y.; Wang, Y.; Ai, C. Dynamic communication between androgen and coactivator: Mutually induced conformational perturbations in androgen receptor ligand-binding domain. *Proteins.*, **2011**, *79*, 1154-1171.
- [12] Wang, X.; Xu, X.; Ma, Z.; Huo, Y.; Xiao, Z.; Li, Y.; Wang, Y. Dynamic mechanisms for pre-miRNA binding and export by Exportin-5. *RNA*, **2011**, *17*, 1511-1528.
- [13] Karplus, M.; McCammon, J.A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.*, **2002**, *9*, 646-652.
- [14] Pearlman, D.A.; Case, D.A.; Caldwell, J.W.; Ross, W.S.; Cheatham, III, T.E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.*, **1995**, *91*, 1-41.
- [15] Case, D.A.; Cheatham, T.E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.*, **2005**, *26*, 668-1688.
- [16] Brooks, B.R.; Brooks, C.L.; 3rd, Mackerell, A.D. Jr.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A.R.; Feig, M.; Fischer, S.; Gao, J.; Hodoseck, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R.W.; Post, C.B.; Pu, J.Z.; Schaefer, M.; Tidor, B.; Venable, R.M.; Woodcock, H.L.; Wu, X.; Yang, W.; York, D.M.; Karplus, M. CHARMM: the biomolecular simulation program. *J. Comput. Chem.*, **2009**, *30*, 1545-1614.
- [17] Jones, J. On the determination of molecular fields. II. From the equation of state of a gas. *Proc. R. Soc. A.*, **1924**, *106*, 463-477.
- [18] Karplus, M.; Kuriyan, J. Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. USA*, **2005**, *102*, 6679-6685.
- [19] Mackerell, A.D. Jr.; Nilsson, L. Molecular dynamics simulations of nucleic acid-protein complexes. *Curr. Opin. Struct. Biol.*, **2008**, *18*, 194-199.
- [20] Tian, B.; Bevilacqua, P.C.; Diegelman-Parente, A.; Mathews, M.B. The double-stranded-RNA-binding motif: interference and much more. *Nat. Rev. Mol. Cell. Biol.*, **2004**, *5*, 1013-1023.
- [21] Jinek, M.; Doudna, J.A. A three-dimensional view of the molecular machinery of RNA interference. *Nature*, **2009**, *457*, 405-412.
- [22] Doyle, M.; Jantsch, M.F. New and old roles of the double-stranded RNA-binding domain. *J. Struct. Biol.*, **2002**, *140*, 147-153.
- [23] Wostenberg, C.; Noid, W.G.; Showalter, S.A. MD simulations of the dsRBP DGCR8 reveal correlated motions that may aid pri-miRNA binding. *Biophys. J.*, **2010**, *99*, 248-256.
- [24] Prompers, J.J.; Brusweiler, R. Dynamic and structural analysis of isotropically distributed molecular ensembles. *Proteins*, **2002**, *46*, 177-189.
- [25] Wostenberg, C.; Quarles, K.A.; Showalter, S.A. Dynamic origins of differential RNA binding function in two dsRBDs from the miRNA "microprocessor" complex. *Biochemistry*, **2010**, *49*, 10728-10736.
- [26] Wostenberg, C.; Lary, J.W.; Sahu, D.; Acevedo, R.; Quarles, K.A.; Cole, J.L.; Showalter, S.A. The role of human Dicer-dsRBD in processing small regulatory RNAs. *PLoS One*, **2012**, *7*, e51829.
- [27] Rybak, A.; Fuchs, H.; Smirnova, L.; Brandt, C.; Pohl, E.E.; Nitsch, R.; Wulczyn, F.G. A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nat. Cell Biol.*, **2008**, *10*, 987-993.
- [28] Hagan, J.P.; Piskounova, E.; Gregory, R.I. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.*, **2009**, *16*, 1021-1025.
- [29] Newman, M.A.; Thomson, J.M. Hammond SM: Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA*, **2008**, *14*, 1539-1549.
- [30] Nam, Y.; Chen, C.; Gregory, R.I.; Chou, J.J.; Sliz, P. Molecular basis for interaction of let-7 microRNAs with Lin28. *Cell*, **2011**, *147*, 1080-1091.
- [31] Ali, P.S.; Ghoshdastider, U.; Hoffmann, J.; Brutschy, B.; Filipek, S. Recognition of the let-7g miRNA precursor by human Lin28B. *FEBS Lett.*, **2012**, *586*, 3986-3990.
- [32] Lund, E.; Güttinger, S.; Calado, A.; Dahlberg, J.E.; Kutay, U. Nuclear Export of MicroRNA Precursors. *Science*, **2004**, *303*, 95-98.
- [33] Yi, R.; Qin, Y.; Macara, I.G.; Cullen, B.R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.*, **2003**, *17*, 3011-3016.
- [34] Sarzyńska, J.; Mickiewicz, A.; Miłostan, M.; Łukasiak, P.; Błażewicz, J.; Figlerowicz, M.; Kuliński, T. Flexibility of Dicer Studied by Implicit Solvent Molecular Dynamics Simulations. *Comput. Methods Sci. Tech.*, **2010**, *16*, 97-104.
- [35] Li, W.X.; Ding, S.W. Viral suppressors of RNA silencing. *Curr. Opin. Biotechnol.*, **2001**, *12*, 150-154.
- [36] Vargason, J.M.; Szitty, G.; Burgyan, J.; Hall, T.M. Size selective recognition of siRNA by an RNA silencing suppressor. *Cell*, **2003**, *115*, 799-811.
- [37] Chu, M.; Desvoyes, B.; Turina, M.; Noad, R.; Scholthof, H.B. Genetic dissection of tomato bushy stunt virus p19-protein-mediated host-dependent symptom induction and systemic invasion. *Virology*, **2000**, *266*, 79-87.
- [38] Xia, Z.; Zhu, Z.; Zhu, J.; Zhou, R. Recognition mechanism of siRNA by viral p19 suppressor of RNA silencing: a molecular dynamics study. *Biophys. J.*, **2009**, *96*, 1761-1769.
- [39] Law, S.M.; Zhang, B.W.; Brooks, C.L.; 3rd. pH-sensitive residues in the p19 RNA silencing suppressor protein from carnation Italian ringspot virus affect siRNA binding stability. *Protein Sci.*, **2013**, *22*, 595-604.
- [40] Wang, Y.; Juranek, S.; Li, H.; Sheng, G.; Tuschl, T.; Patel, D.J. Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature*, **2008**, *456*, 921-926.
- [41] Wang, Y.; Sheng, G.; Juranek, S.; Tuschl, T.; Patel, D.J. Structure of the guide-strand-containing argonaute silencing complex. *Nature*, **2008**, *456*, 209-213.
- [42] Wang, Y.; Li, Y.; Ma, Z.; Yang, W.; Ai, C. Mechanism of MicroRNA-Target Interaction: Molecular Dynamics Simulations and Thermodynamics Analysis. *PLoS comput. Biol.*, **2010**, *6*, e1000866.
- [43] Chen, H-F. Mechanism of coupled folding and binding in the siRNA-PAZ complex. *J. Chem. Theory Comput.*, **2008**, *4*, 1360-1368.
- [44] Yan, K.S.; Yan, S.; Farooq, A.; Han, A.; Zeng, L.; Zhou, M.M. Structure and conserved RNA binding of the PAZ domain. *Nature*, **2003**, *426*, 468-474.
- [45] Tolia, N.H.; Joshua-Tor, L. Slicer and the argonautes. *Nat. Chem. Biol.*, **2007**, *3*, 36-43.
- [46] Balasubramanian, C.; Ojha, R.P.; Maiti, S.; Desideri, A. Sampling the structure of the noncanonical lin-4:lin-14 microRNA:mRNA complex by molecular dynamics simulations. *J. Phys. Chem. B.*, **2010**, *114*, 16443-16449.
- [47] Xia, Z.; Clark, P.; Huynh, T.; Loher, P.; Zhao, Y.; Chen, H.W.; Ren, P.; Rigoutsos, I.; Zhou, R. Molecular dynamics simulations of Ago silencing complexes reveal a large repertoire of admissible 'seed-less' targets. *Sci. Rep.*, **2012**, *2*, 569.
- [48] Xia, Z.; Huynh, T.; Ren, P.; Zhou, R. Large domain motions in Ago protein controlled by the guide DNA-strand seed region determine the Ago-DNA-mRNA complex recognition process. *PLoS One*, **2013**, *8*, e54620.
- [49] Joseph, T.T.; Osman, R. Thermodynamic basis of selectivity in guide-target-mismatched RNA interference. *Proteins*, **2012**, *80*, 1283-1298.
- [50] Joseph, T.T.; Osman, R. Convergent transmission of RNAi guide-target mismatch information across Argonaute internal allosteric network. *PLoS Comput. Biol.*, **2012**, *8*, e1002693.
- [51] Schwarz, D.S.; Tomari, Y.; Zamore, P.D. The RNA-induced silencing complex is a Mg²⁺-dependent endonuclease. *Curr. Biol.*, **2004**, *14*, 787-791.
- [52] Ma, Z.; Xue, Z.; Zhang, H.; Li, Y.; Wang, Y. Local and global effects of Mg²⁺ on Ago and miRNA-target interactions. *J. Mol. Model.*, **2012**, *18*, 3769-3781.
- [53] Plante, I.; Davidovic, L.; Ouellet, D.L.; Gobeil, L.A.; Tremblay, S.; Khandjian, E.W.; Provost, P. Dicer-derived microRNAs are utilized by the fragile X mental retardation protein for assembly on target RNAs. *J. Biomed. Biotechnol.*, **2006**, *2006*, 64347.

- [54] Mortensen, R.D.; Serra, M.; Steitz, J.A.; Vasudevan, S. Posttranscriptional activation of gene expression in *Xenopus laevis* oocytes by microRNA-protein complexes (microRNPs). *Proc. Natl. Acad. Sci. USA*, **2011**, *108*, 8281-8286.
- [55] Ye, W.; Qin, F.; Zhang, J.; Luo, R.; Chen, H.F. Atomistic mechanism of microRNA translation upregulation via molecular dynamics simulations. *PLoS One*, **2012**, *7*, e43788.
- [56] Grant, B.J.; Gorfe, A.A.; McCammon, J.A. Large conformational changes in proteins: signaling and other functions. *Curr. Opin. Struct. Biol.*, **2010**, *20*, 142-147.
- [57] Klepeis, J.L.; Lindorff-Larsen, K.; Dror, R.O.; Shaw, D.E. Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struct. Biol.*, **2009**, *19*, 120-127.
- [58] Allen, T.W.; Andersen, O.S.; Roux, B. Molecular dynamics - potential of mean force calculations as a tool for understanding ion permeation and selectivity in narrow channels. *Biophys. Chem.*, **2006**, *124*, 251-267.
- [59] Hamelberg, D.; Mongan, J.; McCammon, J.A. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J. Chem. Phys.*, **2004**, *120*, 11919-11929.
- [60] Hyeon, C.; Thirumalai, D. Capturing the essence of folding and functions of biomolecules using coarse-grained models. *Nat. Commun.*, **2011**, *2*, 487.
- [61] Durrant, J.D.; McCammon, J.A. Molecular dynamics simulations and drug discovery. *BMC Biol.*, **2011**, *9*, 71.
- [62] Adcock, S.A.; McCammon, J.A. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem. Rev.*, **2006**, *106*, 1589-1615.