Therapeutic Opportunities and Approaches to Sequence Control for Nucleic Acids

Thesis by

Matthew Leroy Gethers III

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2018

(Defended April 2, 2018)

Matthew Leroy Gethers III

ORCID: 0000-0001-7455-4709

ACKNOWLEDGMENTS

I would like to thank my advisors, Bill Goddard and Paul Weiss, for allowing me the freedom to explore my own ideas, and my colleagues in the Materials and Process Simulation Center, the Weiss Group, and the Duan Lab for providing a supportive and fun environment in which to pursue them. I would especially like to thank Drs. Si-Ping Han, Tod Pascal, and John C. Thomas for being great mentors. I couldn't have done without the support I received from my family, friends, the Caltech Center for Diversity, and the broader Caltech community. I would also like to acknowledge support from the NSF ODISSEI program, Caltech's EAS Discovery Fund, and the US Department of Energy Grant # OE-SC-000525.

ABSTRACT

RNA interference (RNAi) is a powerful mechanism to regulate gene expression. A key feature of RNAi is its sequence specificity: a short interfering RNA (siRNA) assembles into the RNA induced silencing complex (RISC) and then targets cellular transcripts complementary to the siRNA for degradation. RNAi has been adapted for therapeutic applications, but is challenged by the need to identify unique target transcripts for each disease that are both effective and result in few off-target effects. This challenge could be eased if siRNAs could be activated only and specifically in diseased cells. If this were the case, rather than targeting a new transcript for each new disease, the same cellular housekeeping genes could be reused. Targeting housekeeping genes would result in greater potency, both effectively treating the disease and requiring less drug for treatment, alleviating problems associated with toxicity and delivery. A new class of nucleic acid therapeutics called conditional siRNAs (*Cond*-siRNA) is designed to act in this environment-specific manner. The first part of this thesis uses molecular dynamics simulations to understand the structure of *Cond*-siRNA and to suggest improvements in future designs.

Bioengineering like the work done in the development of *Cond*-siRNAs depends on the existence of tools that make work simple, fast, cheap, and reproducible. In the case of nucleic acids, *de novo* synthesis of custom constructs is a fundamental tool. While approaches to synthesis have improved immensely since their inception, increasing ambition demands increasingly powerful tools. As target constructs get longer, the synthesis can become intractably complicated, slowing the process, increasing costs, and making it less likely to be replicated by others. The source of complexity in nucleic acid synthesis is the inability to directly synthesize long fragments without errors. Finding a new means of sequence-controlled synthesis that results in fewer errors and perhaps allows for correction could address this challenge. The second part of this thesis looks at using graphene as a mask for patterning the deposition of molecules on a surface with an eye towards arranging and coupling reactants in a sequence-specific way.

PUBLISHED CONTENT AND CONTRIBUTIONS

Gethers, M., Thomas, J.C., Jiang, S., Weiss, N.O., Duan, X., Goddard III, W.A., Weiss, P.S. (2015). "Holey Graphene as a Weed Barrier for Molecules". In: *ACS Nano* 9, pp. 10909-10915. doi: 10.1021/acsnano.5b03936.

M.G. initiated the project, conducted the experiments with J.C.T., prepared the data, and participated in the writing of the manuscript.

Han, S., Scherer, L., Gethers, M., Mancusi, R., Kuo, Y., Marcucci, G., Rossi, J., Goddard III, W.A. (2018) "Conditional small interfering RNAs programmable for intracellular activation by specific RNA transcripts." (In prep)

M.G. performed molecular dynamics simulations to inform and explain experiments conducted by S.H., L.S., and R.M.

Gethers, M., Goddard W.A., Weiss, P.S., Randall, J., (2015) "Surface Mediated Synthesis of Polynucleotides, Polypeptides, and Polysaccharides and Related Materials, Methods, and Systems." (Patent Pending)

M.G. initiated the project and drafted the patent in collaboration with legal counsel.

TABLE OF CONTENTS

Acknowledgments..... iii
Abstract iv
Published Content and Contributions.....v
Table of Contents..... vi
List of Figures and Tables.....vii
Introduction 1
Chapter I: A Molecular Dynamics Investigation of *Cond*-siRNAs 2
    Introduction 3
    Background 4
    Materials and Methods 12
    Results 14
    Discussion 26
    References 29
Chapter II: Holey Graphene as a Weed Barrier for Molecules 31
    Introduction 32
    Background 33
    Materials and Methods 35
    Results 38
    Discussion 44
    References 46
Appendix A: Supplementary Material for *Cond*-siRNAs 54
Appendix B: Supplementary Material for Weed Barrier 95

LIST OF FIGURES AND TABLES

Number *Page*

1. Figure 1-1 Nucleosides 6
2. Figure 1-2 Modified Nucleosides 6
3. Figure 1-3 Nucleoside Torsions 7
4. Figure 1-4 Local Parameters 7
5. Figure 1-5 Step Parameters 8
6. Figure 1-6 *Cond*-siRNA Architecture 11
7. Figure 1-7 *Cond*-siRNA Activation 11
8. Figure 1-8 Initial and Final Structures of *Cond*-siRNAs 15
9. Figure 1-9 Averaged Structures of *Cond*-siRNAs 16
10. Figure 1-10 Chi Torsions 18
11. Figure 1-11 Pucker 19
12. Figure 1-12 Propeller Twist 20
13. Figure 1-13 Roll Angle 21
14. Table 1-1 Duplex Pitch 22
15. Figure 1-14 Direction of Linker Movement 23
16. Figure 1-15 Protection of Dicer Cleavage Site 25
17. Figure 2-1 Weed Barrier Procedure 37
18. Figure 2-2 TEM of Holey Graphene 38
19. Figure 2-3 STM of Graphene on Au {111} 39
20. Figure 2-4 STM of Molecules Deposited in Graphene Pores 41
21. Figure 2-5 STM of Graphene on Au {111} after Anneal 42

Introduction

This is an exciting era for the engineering of biology. In biological molecules and organisms, engineers have at their disposal a medium that can sense and respond to its environment, complete complex chemical reactions at room temperature and pressure, and evolve to become better at its work over time. The potential impact is particularly exciting in the field of medicine where the treatment of complex pathologies would benefit from more sophisticated therapeutics. The application of RNA interference (RNAi) to treat disease is an example of such a therapeutic. By introducing a small interfering RNA (siRNA) to a cell, a particular cellular mRNA transcript can be targeted for degradation and thereby regulate gene expression. The effect of this siRNA could be improved if it were activated only on the condition that it is in a diseased cell. If this were the case, an essential housekeeping gene common to all cells could be targeted. Targeting a housekeeping gene would have the effects of simplifying the design, increasing potency, and reducing the amount of drug required, lowering toxicity and easing the challenge of delivery. A new class of therapeutics named conditional siRNAs (*Cond*-siRNA) seeks to achieve this conditional activation.[1] Chapter I of this thesis presents the results of molecular dynamics simulations conducted to better understand the structure of *Cond*-siRNAs and to suggest improvements in future designs.

Successful engineering of biological systems is enabled by tools that make the work simple, fast, cheap, and reproducible. Nucleic acid products like *Cond*-siRNAs are made *via* solid-phase synthesis, a technique that sequentially adds monomers to a growing chain and is characterized by a limited product length due to failed coupling events. To achieve longer constructs, the products of solid-phase synthesis can be inputs to enzymatic assemblies, but these assemblies become more complicated and unsuccessful as the number of inputs increases. If the issue of error could be addressed at the stage of direct synthesis, the complications of the enzymatic assembly could be avoided. With this end in mind, the second part of this thesis explores a new graphene-based method of patterning molecules on a surface with the eventual goal of using the technique to arrange and couple monomers in a sequence-specific way.

C h a p t e r I

A Molecular Dynamics Investigation of *Cond*-siRNAs

Introduction

RNA interference (RNAi) is a powerful biological mechanism for gene regulation. Within a cell, double-stranded RNA is enzymatically cut down to ~20 base pair fragments that then join with protein to form the RNA-induced silencing complex (RISC). The RISC then knocks downs gene expression in a sequence-specific manner. Since its discovery in *C. elegans,*[2] we continue to learn about the prevalence of RNAi. It is used both for endogenous post-transcriptional gene regulation and for defense from exogenous sources.[3] The power of RNAi in regulating gene expression has made it an exciting candidate for therapeutic applications. Initial trials have shown that it can be effective against diseases that are intractable with typical small-molecule therapies, however, there are a number of limitations.[4] A unique mRNA target must be found for each disease. This target must be specific to the disease and not to healthy cells, and the disruption of this target must lead to remission of the disease. As always, delivery is a challenge. Each of these problems could be addressed if RNAi could be initiated only in cells where the disease is present. If this were the case, essential genes common to all cells could be the target of the siRNA. Targeting essential genes would make the drug more potent and broadly applicable. Further, a more potent drug could be delivered in smaller quantities, avoiding off-target toxicity and easing the burden of delivery. This environment-specific RNAi activity is the aim of conditional siRNA (*Cond*-siRNA).[1]

In the presence of a user-defined mRNA transcript, *Cond*-siRNA activates a siRNA that targets a separate user-defined transcript for knockdown. A *Cond*-siRNA comprises an siRNA and a sensor duplex covalently linked through a pair of carbon chains. The proximity of the two duplexes prevents the enzymatic processing of the siRNA and assembly into the RISC. The removal of the sensor duplex through toehold-mediated strand displacement leaves the siRNA accessible for processing and assembly into the RISC, leading to knockdown of the target. Given the role of the sensor duplex in protecting the siRNA, the structure of the *Cond*-siRNA is critical to its function. The initial *Cond*-siRNAs

were designed using only thermodynamic optimization to choose stable duplexes and to avoid problematic secondary structures. This thermodynamic optimization does not take into account the dynamic structure of the *Cond*-siRNA. The goal of this work is to use molecular dynamics simulations to give detailed structural information about the *Cond*-siRNA. Of special interest is how well the Dicer cleavage site of the *Cond*-siRNA is protected and the stability of base pairing in the face of extensive modifications. With these answers in hand, this work further seeks both to explain experimental results and suggest improvements for the design of the next generation of constructs.

Background

Structural Description of Nucleic Acids

Nucleic acids play essential roles in biology. They are the material in which genes are encoded and expressed. They also regulate gene expression and serve as a source of energy. The capacity for nucleic acids to play these roles is a function of their structure. The nucleoside, the basic unit, comprises a furanose ring and a nitrogenous base. Figure 1-1 shows both a purine and pyrimidine nucleoside, with each atom named according to convention. Modifications to naturally occurring nucleosides change both structure and chemical properties. This work on *Cond*-siRNA involves several modified nucleosides: the Locked Nucleic Acid (LNA), 2' O- methyl, and phosphorothioate. Figure 1-2 shows the structures of these modified nucleosides. The LNA involves a carbon that bridges the 2'O and the C4'. The result is a furanose ring locked in the 3' endo conformation, which stabilizes base pairing. The 2'O-methyl performs a similar role. The phosphorothioate prevents degradation *via* exonucleases.

Nucleosides are connected to other nucleosides *via* phosphodiester bonds to form single strands of nucleotides. When strands have complementary sequences, they can come together to form a duplex. The bases of each nucleoside interact with one another across strands through hydrogen bonding. Adenosine pairs with thymidine and uridine, and

guanosine pairs with cytidine. The specificity of base pairing is responsible for the fidelity of replication and transcription. Nucleosides also interact with one another within strands through π stacking between bases. These interactions influence the overall shape of the duplex.

Standard parameters have been defined to describe nucleic acid structures.[5] At the level of the single strand, seven torsions are used to describe the strand shape. Six of these torsions are along the backbone, and one describes the gycosidic linkage between the base and furanose ring. Figure 1-3 shows these torsions. Within each base, the pucker of the furanose ring is described as a phase angle in a pseudorotation cycle. Base pairs are described by six parameters: buckle, opening, propeller twist, shear, stagger, and stretch. These are shown in Figure 1-4. Steps between base pairs are described by six parameters: rise, roll, shift, slide, tilt, and twist. These are shown in Figure 1-5. The local and step parameters allow a complete description of helical structure. The duplex can also be described more coarsely in terms of pitch: base pairs per turn and rise per turn.

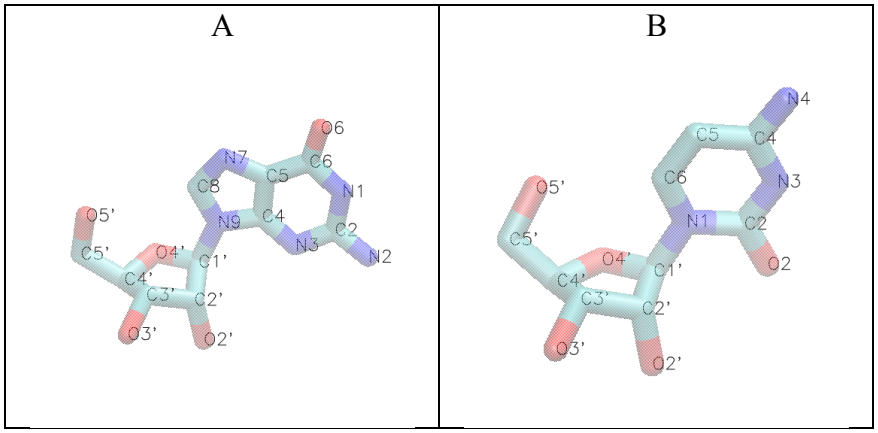


Figure 1-1 Purine (A) and Pyrimidine (B) nucleosides with atom names.

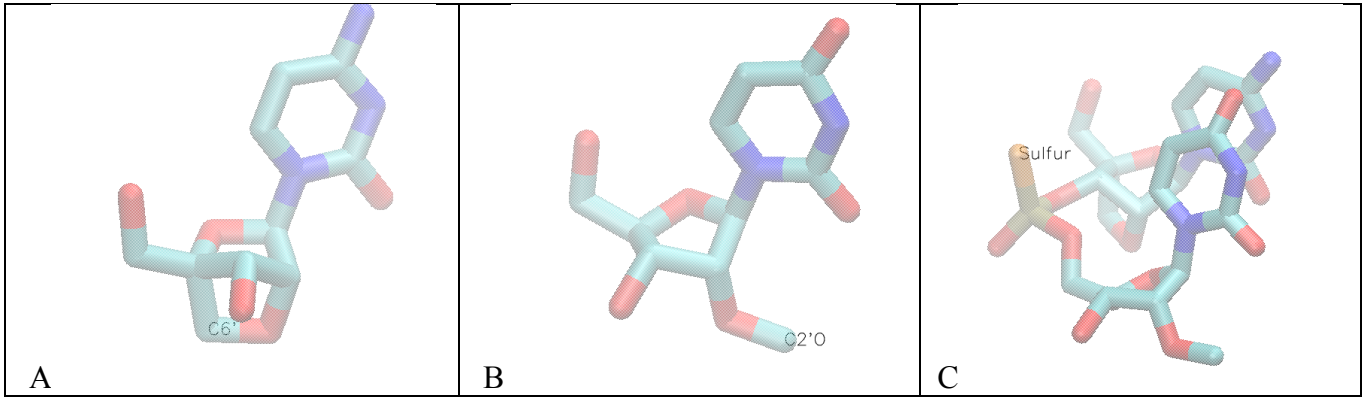


Figure 1-2 Modified Nucleosides. LNA (A), 2'-O methyl (B), and phosphorothioate (C).

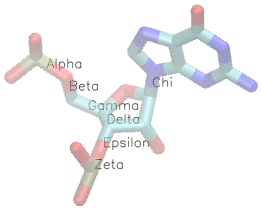


Figure **1-3** Nucleoside Torsions.

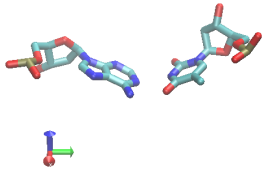
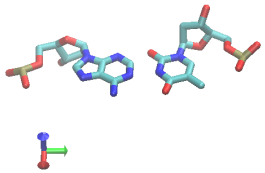
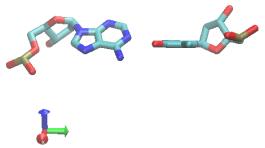
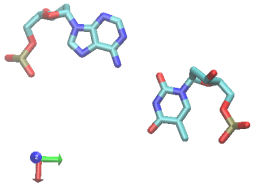
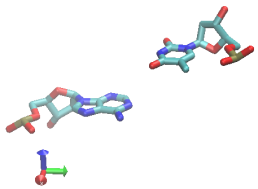
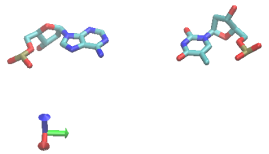
Buckle	Opening	Propeller
		
Shear	**Stagger**	**Stretch**
		

Figure **1-4** Nucleoside local parameters.

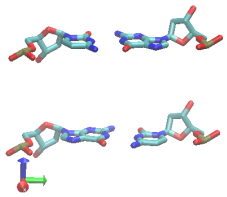
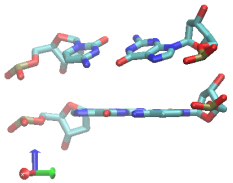
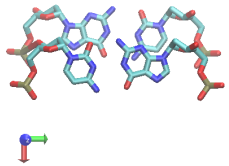
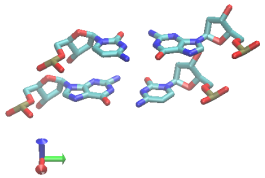
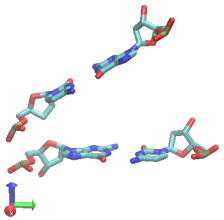
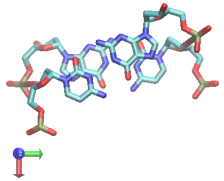
Rise	Roll	Shift
		
Slide	Tilt	Twist
		

Figure **1-5** Nucleoside step parameters.

RNA Interference

The first example of RNA interference (RNAi) was discovered in *C. elegans*[2] in which a micro RNA (miRNA) originating in the cell nucleus was found to be enzymatically processed and moved to the cytoplasm where it assembles into the RNA-induced silencing complex (RISC). The RISC then regulates function in the cell by sequestering or degrading mRNAs. This first example represented a new means of regulation of endogenous genes. Since this first discovery, different incarnations of the mechanism have been discovered in different organisms.[6] Another instance of RNA interference utilizes small interfering RNA (siRNA), which is much smaller than miRNA and may have both endogenous and exogenous sources.

The typical siRNA is 20-30 nt in length. Dicer, an enzyme, trims siRNA leaving a two-nucleotide overhang on each 3' terminus and phosphate at each 5' terminus. This processed RNA is then loaded into a ribonucleoprotein complex with a member of the Argonaut family. Once the double-stranded helix is in Argonaut, the phosphate of the guide strand is bound by PAZ domain and MID domains respectively, generating the RISC. At this point, the passenger strand is discarded, leaving a complex between a single strand of RNA (the guide strand) and the protein. The guide strand is then used to base pair with mRNAs in the cell. Depending on the level of sequence complementarity, the RISC can regulate the transcript in two ways. With a high degree of complementarity, the mRNA can be degraded. With low complementarity, the translation of the transcript can be inhibited.

This powerful means of regulating gene expression made RNAi an obvious target for drug development.[4] siRNAs targeting transcripts associated with disease could specifically and effectively treat that disease. Much work has been done along these lines. The use of siRNA for treating disease is met with several challenges, however. First, a suitable mRNA target must be chosen. This target must be specific to the disease and its down-regulation must lead to the disease going into remission. Further, the siRNA must have little toxicity and cause minimal off-target effects. These requirements mean that a new target must be found for each disease and toxicity must be determined each time. A powerful advance in the field of siRNA would be made if the siRNA could first assess the disease state of the cell, and act only if the cell is diseased. In this case, the siRNA could

target an endogenous housekeeping gene of the cell rather that something specific to the disease. This would be helpful because the same target could be used for myriad diseases, and because targeting housekeeping genes would likely result in greater potency. Greater potency, in turn, would require smaller doses of the drug. These aims could be achieved by *Cond*-siRNA.

Conditional RNA

The goal of the *Cond*-siRNA is to knockdown a user-defined target mRNA transcript in cells *via* the RNAi pathway in response to the presence of separate user-defined signal transcripts in the same cells.[1] Han *et al*. achieved this behavior through the architecture depicted in Figure 1-6. Three strands form two duplexes that are joined by a pair of three-carbon linkers. The guide strand and the inner part of the core strand form the siRNA duplex. The sensor strand and the termini of the core strand form the sensor duplex. The three-carbon linkers join the termini of the core strand to its middle. The key hypothesis is that the proximity of the two duplexes prevents Dicer from cleaving the siRNA duplex to render it viable for the RISC. In the event that RNA complementary to the sensor strand is present, it will remove the sensor strand through toehold-mediated strand displacement, leaving single stranded regions on the termini of the core strand. With the sensor duplex melted, Dicer can process the siRNA duplex, as shown in Figure 1-7.

In the experiments performed by Han *et al*., three different *Cond*-siRNA constructs were tested. In each case, the sensor sequence was selected by considering the sequence of the target mRNA, going over all frames 33 nucleotides long, and finding those sequences that did not have significant overlap with other endogenous mRNA sequences or significant secondary structures. These sequences were further modified with LNAs, 2'-O methyls, and phosporothioates to stabilize the desired base pairing over secondary structures and to prevent degradation by nucleases. It is important to note that the only design considerations were sequence overlap and thermodynamic stability of the intended duplex relative to potential secondary structures, and that each duplex was considered in isolation.

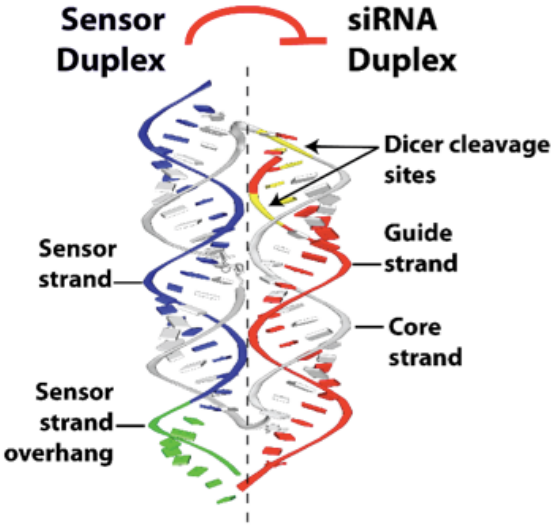


Figure **1-6** Architecture of *Cond*-siRNA. Reproduced with permission from Han *et al.*

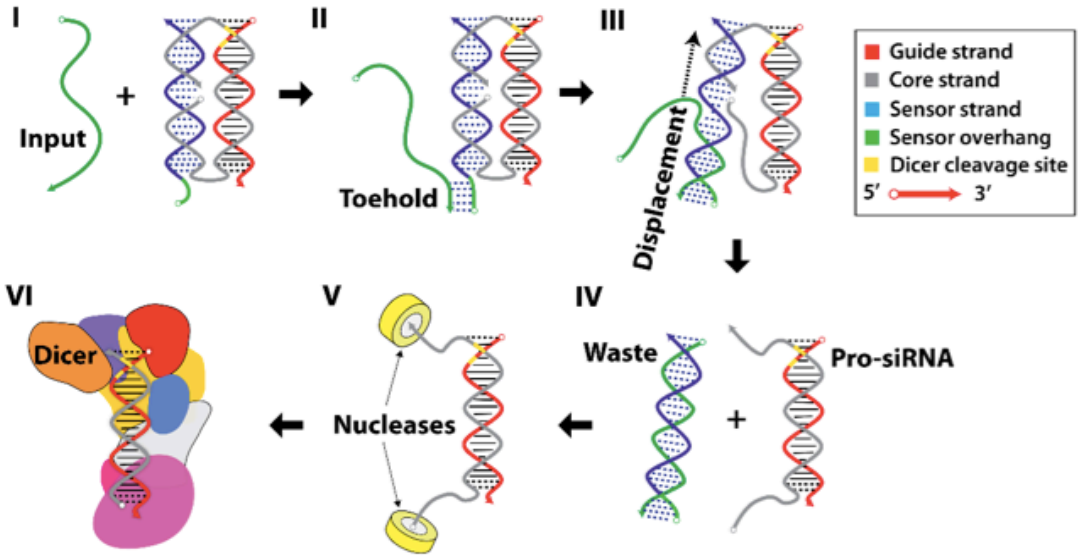


Figure **1-7** Activation of *Cond*-siRNA. Reproduced with permission from Han *et al.*

Molecular Dynamics

Molecular dynamics (MD) is a computational technique for studying the evolution of a molecular system through time.[7] Depending on software and hardware, MD can produce trajectories spanning from nanoseconds to milliseconds. Molecular dynamics is an important tool because it gives atomic resolution insights about systems that could

otherwise be difficult to study at that spatial and temporal resolution. Atoms are represented by their coordinates and masses. Bonds, angles, dihedrals, and two-body non-bond interactions are represented by potential functions that are parameterized in a force field. Building an atomic model of the initial state of the system can be done by using a crystal structure as template or from quantum mechanics calculations. Explicit water and ions can be added as well. Calculation of point charges associated with each atom is achieved through quantum mechanics techniques. The force field can be parameterized either through empirical data or through quantum mechanics calculations.

A typical simulation begins with a geometry optimization that moves the structure into a local energy minimum. This process improves the conformation of a solute molecule and reduces large forces due to clashes with the added solvent. The next step is to bring the system to the appropriate temperature and pressure or volume. This is achieved through integration while keeping certain variables constant. The NVT ensemble, for example, keeps the amount of mass in the system, the volume, and the temperature constant. Temperature is kept constant through the use of a thermostat. In the NPT ensemble, mass, pressure, and temperature are kept constant. Pressure is kept constant through the use of a barostat. Finally comes the production run, the step from which data will be collected for analysis. This step can utilize any of the aforementioned ensembles.

Materials and Methods

RNA structures were built using a combination of Nucleic Acid Builder,[8] Cerius[2], a GUI builder by Accelerys, and custom scripts written in both Python and Perl. These RNA structures were then solvated with a custom script using TIP3 explicit waters, creating a buffer of 15 Å on each face of the water box. Each structure was neutralized by both magnesium and sodium, with each ion equally represented by charge. The ions were placed using a custom script that calculates the electrostatic potential of the entire system and places the ions at the minima, removing waters that clash. To this neutralized structure, 150 mM NaCl was added using the same process. The forcefield is a combination of AMBER03[9] and General Amber Forcefield (GAFF).[10] Parameters and charges for the

LNA, phosphorthioate, and 2'O-Methyl came from Condon,[11] Venkateswarlu,[12] Lind,[13] respectively. Charges for the standard bases were assigned using LEAP[14] with the ff99SB parameter set. For the three-carbon linkers and amine and PEG termini, charges were calculated using the RESP ESP charge Derive (R.E.D.) server.[15]

Simulations were run using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS)[16] on Nvidia K80 GPUS. Structures were first minimized for 500 steps using the steepest descent algorithm, with a harmonic potential on all RNA atoms to keep them in place while the solvent moved into a better position. This minimization was then followed by a second 500-step minimization using the conjugate gradient algorithm. The restraint on the RNA atoms was then removed, and the entire system was minimized a third time for 500 steps using the conjugate gradient algorithm.

The minimization was followed by several equilibration steps. In each case, a 1 femtosecond timestep was used and the SHAKE algorithm was applied to the hydrogens in the system. The first equilibration brought the system to 310 K. Velocities were initialized randomly at 1 K. The NVT fix was used to bring temperature up from 1 K to 310 K over the course of 10 picoseconds. A relaxation time of 100 femtoseconds was used for the themostat. This equilibration was followed by a 10 picosecond NPT step to relax the size of the water box. The target temperature and relaxation time for the thermostat remained at 310 K and 100 femtoseconds, respectively. The target pressure and relaxation time were set at 350 bar and 1 picosecond. The high pressure was used to avoid negative pressure oscillations. The NPT equilibration was followed by the 20 ns NVT production run. A periodic box was used. A 12 Å cutoff was implemented, using the lj/charmm/coul/long/gpu pair style. The kspace style was pppm/gpu with an accuracy of 0.001. The amber special bonds style turns off 1-2 and 1-3 non-bond interactions, and scales Columbic interactions by 0.8333 and van der Waals interactions by 0.5. Off-diagonal van der Waals terms are handled by geometric combination rule. Snapshots were dumped every 5 picoseconds. Structures were visualized in VMD[17] and X3DNA[18] was used for calculating the helical parameters. MATLAB was used to process the data and generate plots.

Results

The two goals of the simulations are to determine if the extensive modifications have compromised the structure of the *Cond*-siRNA, and to test the hypothesis that the proximity of sensor duplex protects the Dicer cleavage site. NVT simulations were run for 20 ns on three *Cond*-siRNA constructs used in experiments. Simulations were also run for 20 ns on the isolated all-natural version of each duplex: the siRNA, and the sensor duplex for construct 1, 2, and 3. These all-natural simulations were used to obtain the mean and standard deviation of each parameter against which the modified constructs were plotted. The sequences for each construct can be found in appendix 1 Figures A1-1 through A1-3. Figure 1-8 shows the initial and energy-minimized final conformations for each construct. For analysis, the average structure of the final 5 ns of the trajectory was constructed using X3DNA. These structures are depicted in Figure 1-9. Considerable deformation of the siRNA and sensor is apparent for construct 1. Minor deformations are apparent for the other duplexes. To quantify these features, the helical parameters vs. time are plotted for each strand, base pair, and step of each duplex.

For each strand, the seven torsions and the pucker are examined. In all cases, the torsions are within the normal range. The chi torsion is presented as Figure 1-10 and the other torsions can be found in appendix 1 Figures A1-4 through A1-17. A chi torsion between 90° and 270° is said to be in the *anti* position and denotes a base pointed away from the furanose ring allowing for good Watson-Crick base pairing. It is, therefore, a good indicator of health of the duplex. Figure 1-10 shows that the glycosidic torsion is in the *anti* position for all bases. One feature of note in Figure 1-10 is that the plot has a high representation of points below the average. This is due to the isolated natural reference structure for the sensor of construct 1. Interactions between the toehold and the duplex during the trajectory caused the bottom of the duplex to deform and the affected base pairs adopted higher chi values. The reference structure is in effect bimodal, with a small

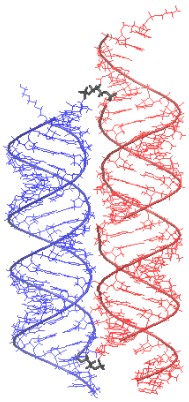
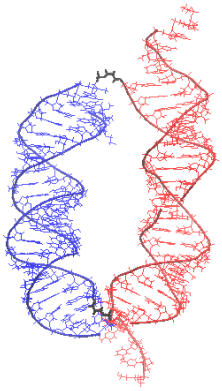
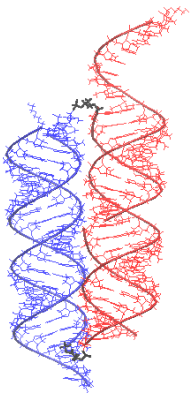
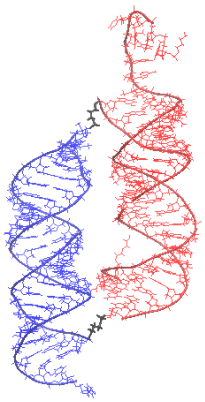
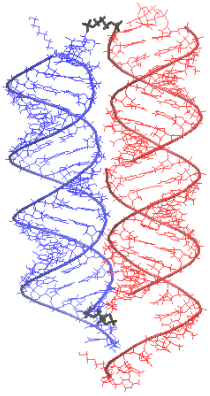
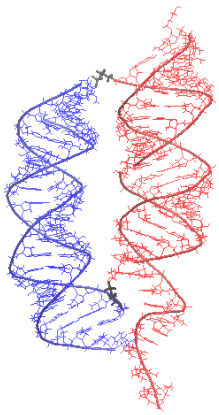
	Initial	20 ns
Construct 1		
Construct 2		
Construct 3		

Figure 1-8 Molecular Dynamics of three *Cond*-siRNA constructs. Images show the initial configuration and the energy-minimized final frame of a 20 ns simulation.

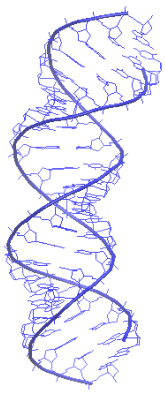
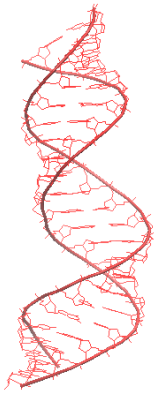
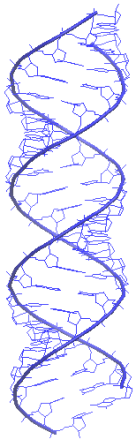
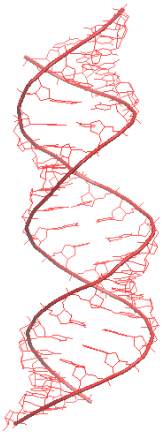
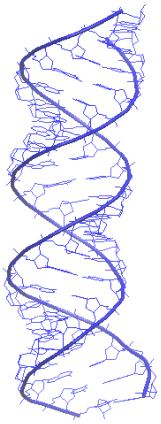
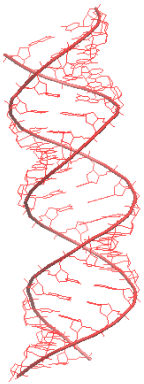
	siRNA	Sensor
Construct 1		
Construct 2		
Construct 3		

Figure 1-9 Average structures for each duplex of each construct. The last 5 ns of each 20 ns trajectory were used to calculate the average structure of each siRNA and sensor duplex.

population above the main population. This small population shifts the average upward, and thus makes most points appear below the mean. Figure 1-11 shows the pucker for each base of each strand represented as an angle along a pseudorotation cycle. RNA, and particularly LNA and RNA modified with 2' O Methyl groups, should be stabilized in the 3' endo confirmation, which corresponds to a phase angle of 18°. Figure 1-11 shows that puckers are stable around this value.

The six local parameters were calculated, and all values are within normal ranges. The propeller twist is shown in Figure 1-12 and the others can be found in appendix 1 Figures A1-19 through A1-29. Propeller twist is unique among the local parameters, as it is expected to have a non-zero average of about 12°. Base pairs are not perfectly planar because they twist so as to optimize base stacking with adjacent bases. The mean of -10° shown in Figure 1-12 is consistent with expectation, considering the convention for measuring the propeller twist is reversed by X3DNA.

The six step parameters were calculated, and all values are within normal ranges (see appendix 1 Figures A1-30 through A1-40). Of particular interest in describing the curvature of duplexes is the roll angle, which is presented in Figure 1-13. Of note are the steps 5, 16, and 20 in the siRNA, which are the only steps in the siRNA duplex that have extreme roll angles across constructs. Inspection of these steps shows that they occur where purine bases on opposite strands and adjacent base pairs can clash in the minor groove. Calladine predicted that due to non-zero propeller twist, purine bases in this situation would clash with one another and that this clash could be resolved by reducing the propeller twist, increasing the helical twist, and increasing the roll angle.[19] It seems that the increased roll angle in these three steps is the result of Calladine's purine-purine clashes. Examination of an isolated all-natural siRNA also shows extreme roll angles at these steps (see appendix 1 Figure A1-32). While this result appears to be as sequence-dependent feature, it should be noted that the degree of roll is not consistent across constructs at these steps. This variation could be a matter of sampling and the effects of the other duplex.

Steps that should be subject to steric clashes between purines are also present in the sensor duplexes. In construct sensor 1, these steps are 10, 16, 19, and 22. Examining the roll angles in Figure 1-13 shows extreme roll angles at 10 and 16, however, the roll angle in

step 10 is decreased relative to the average, opposite to what Calladine's rules predict. In the sensor of construct 2, steps 7, 9, and 14 may clash. There is some evidence of 7 and 9 increasing the roll. In construct 3, the steps are 3, 13, 19, and 22. These don't appear to have extreme roll angles. Proximity to the nick and presence of modifications likely have an effect on the behavior of the RNA in these examples.

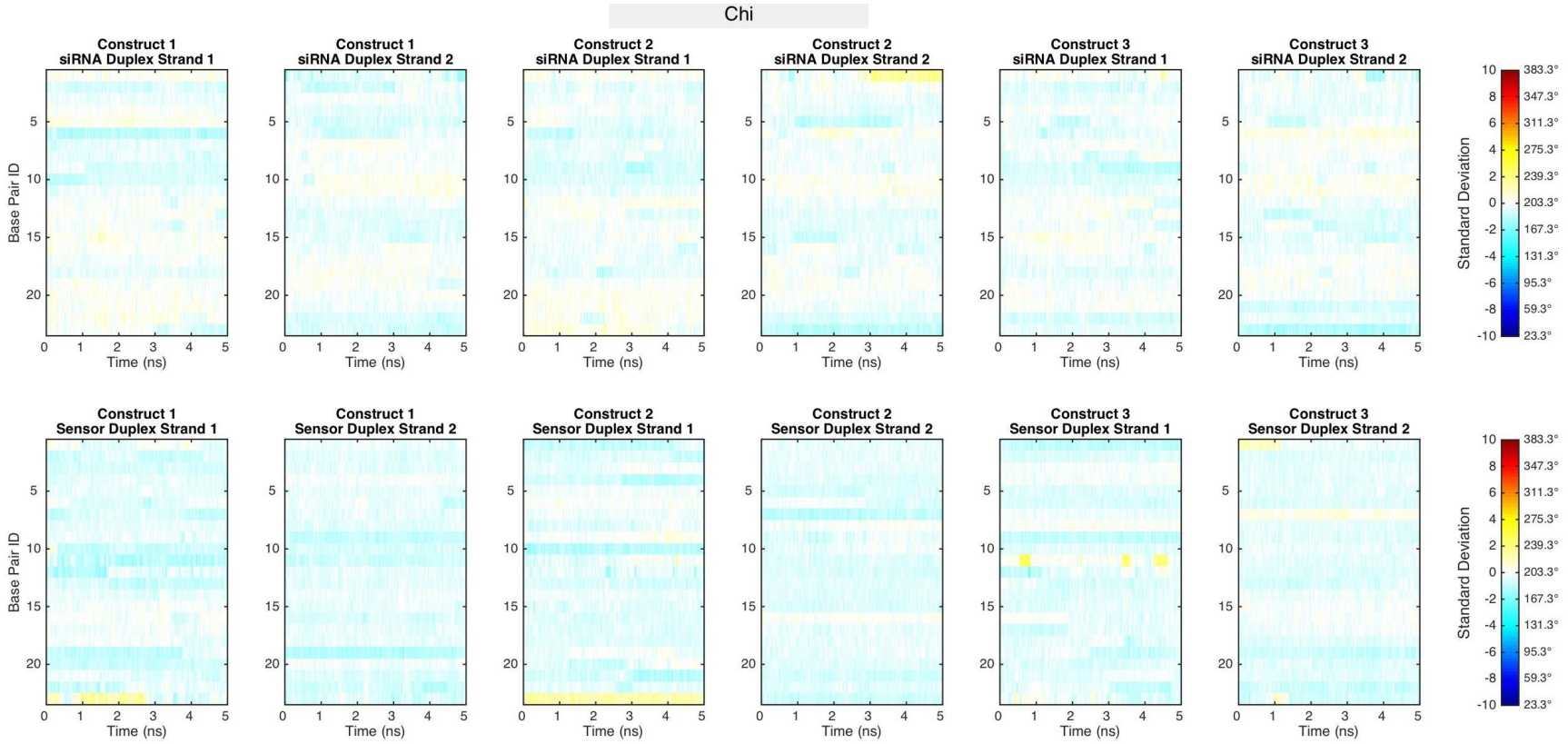


Figure 1-10 Each plot represents the chi torsion for each base in each strand across the last 5 ns of the trajectory. The high representation of points below the mean is due to spurious interactions in the reference structure increasing the mean.

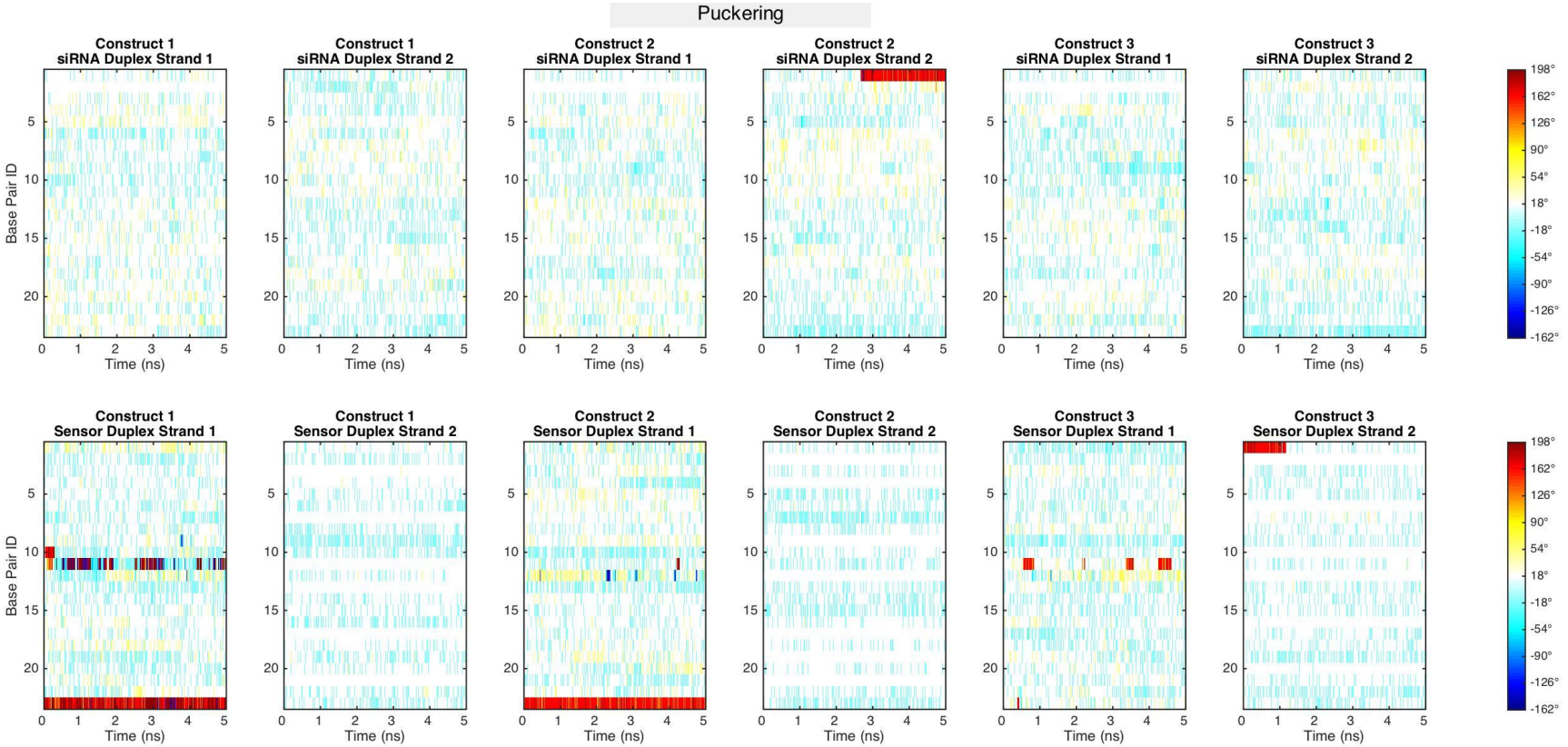


Figure 1-11 Pucker is represented as a phase angle in the pseudorotation cycle. An angle of 18° represents the 3' endo conformation, which is to be expected for RNA.

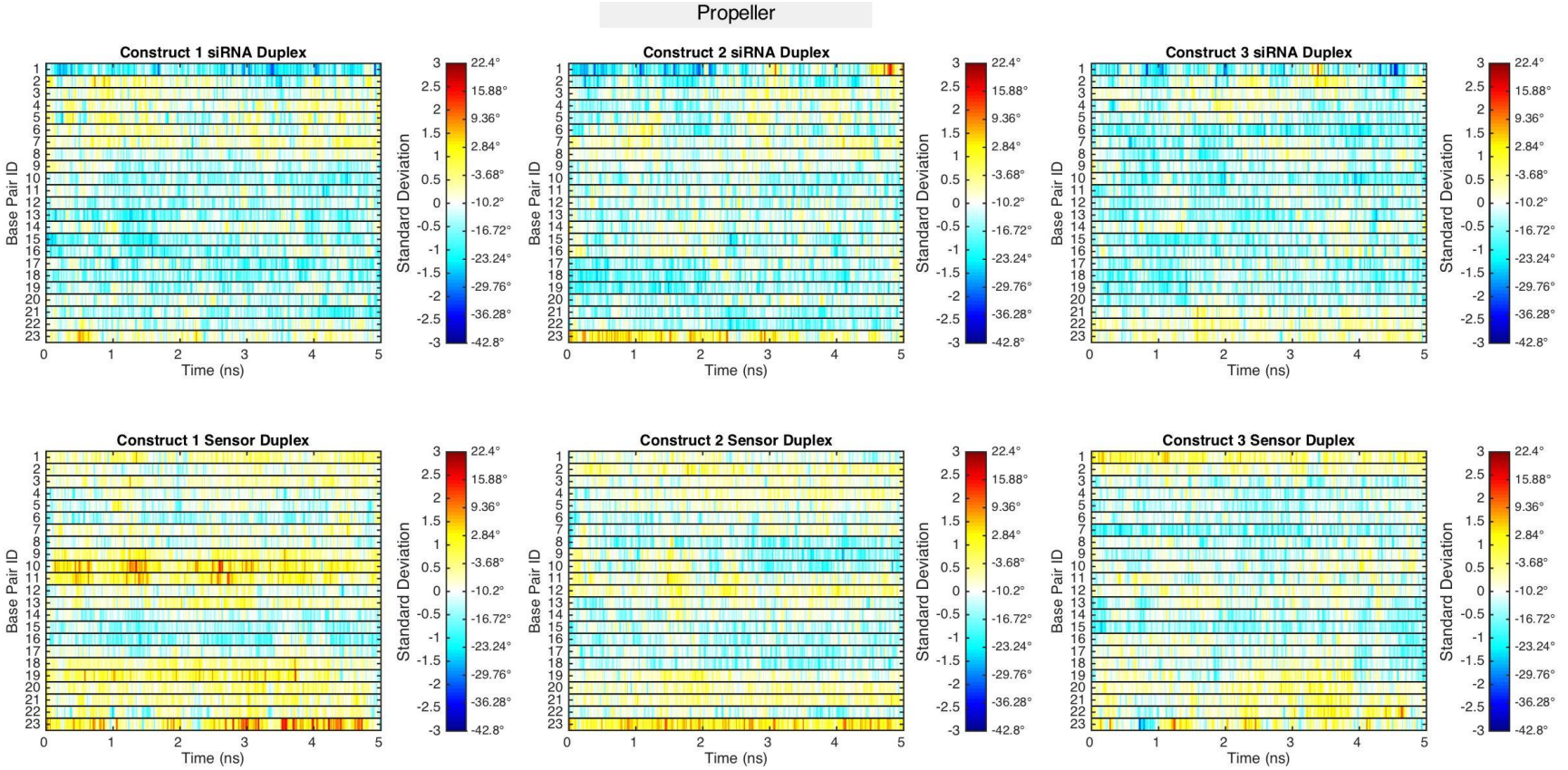


Figure 1-12 An average of -10° for propeller twist is consistent with expected value of 12°, considering the difference in convention adopted by X3DNA.

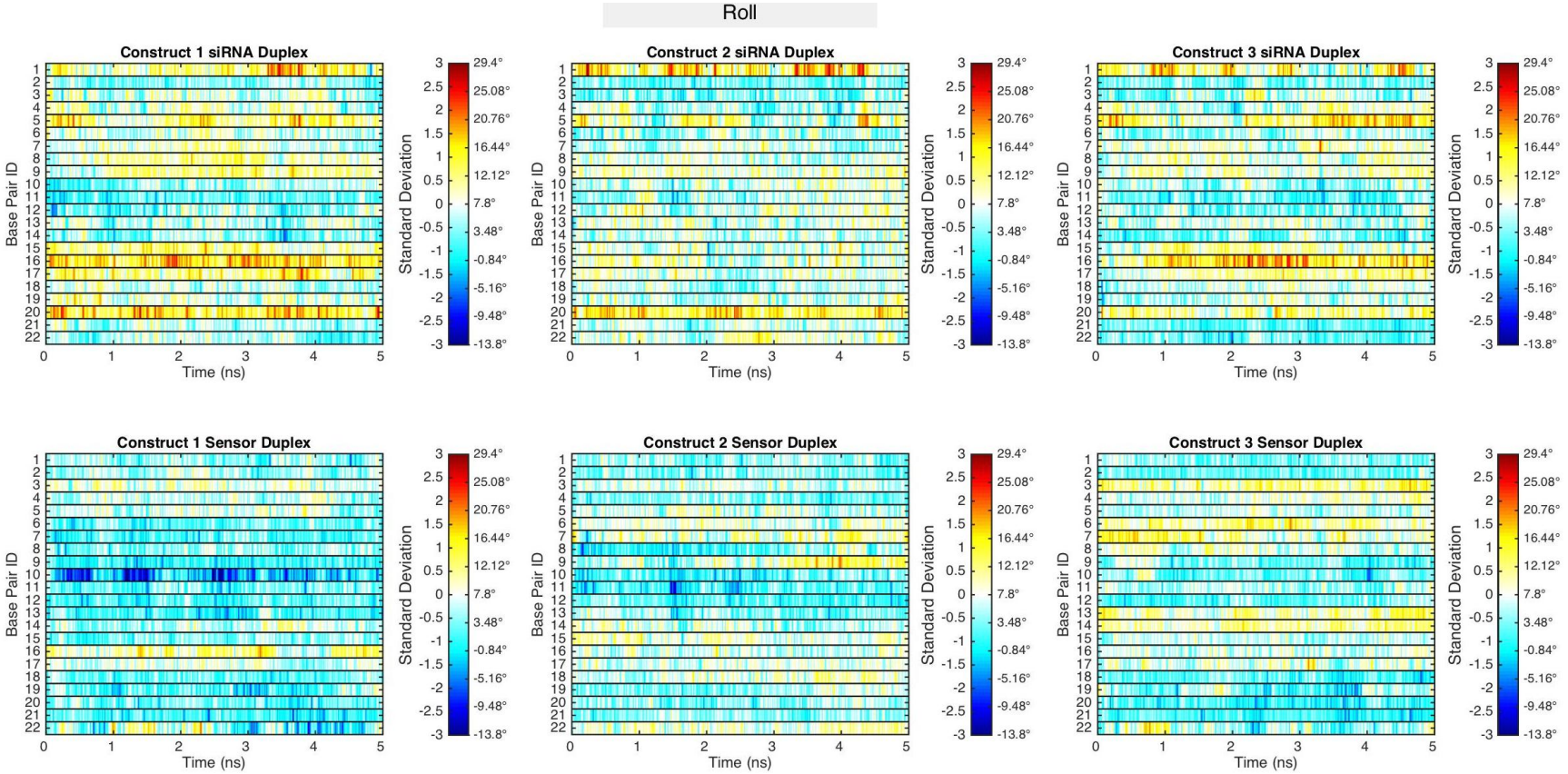


Figure **1-13** The roll angle averages at 8° and is consistent with expectation. Of note are steps 5, 16, and 20 in the siRNA. These are predicted to have large deviations by Calladine's rules. Other steps in the sensor duplexes are also predicted to have deviations, but do not. This could be due to the modifications or the nick.

The RNA duplex can also be described by its pitch both in terms of base pairs per turn and rise per turn. The values for ideal A-RNA are 11 base pairs per turn and 30 Å per turn, and A'-RNA is 12 base pairs per turn and a rise of 36 Å per turn. Table 1-1 shows the pitch of each duplex. Each construct is consistent either with A-RNA or A'-RNA. The sensor duplexes are consistently underwound and longer relative to the siRNAs. This could be due to both the modifications and the nick. The torsions, pucker, base pair and step parameters, and pitch all suggest that despite the extensive modifications and linkage, each RNA duplex behaves within the bounds of expected behavior.

Table 1-1 Pitch of each duplex

	Base Pairs per Turn	Rise per Turn (Å)
Construct 1 siRNA	11.38	30.30
Construct 1 Sensor	12.33	36.83
Construct 2 siRNA	11.37	31.05
Construct 2 Sensor	12.08	33.47
Construct 3 siRNA	11.30	31.03
Construct 3 Sensor	11.65	32.72

Even though they are within normal bounds, the overall shapes of the duplexes in construct 1 demonstrate significant curvature. These shapes can be explained by stress induced by the relaxation of the sensor duplex to its equilibrium pitch. The equilibrium pitch for the sensor is 12.33 base pairs per turn. Relative to the 11 base pairs per turn of ideal A-RNA in which it is initialized, the sensor duplex equilibrium is unwound. The motion of unwinding the sensor duplex is transferred to the siRNA through the three-carbon linkers and leads to the deformation of the siRNA. Looking down from the top and up from the bottom of the construct, in order to unwind, the sensor must rotate clockwise. At the top, the rotation brings the three-carbon linker across the junction between two duplexes. Given that the bottom of the siRNA duplex cannot move freely, the top of the duplex deforms in response to the movement of the sensor. At the bottom, the rotation of the sensor moves the three-carbon linker further away from the junction between the

duplexes. As the sensor duplex rotates, it makes the three-carbon linker taught, which brings the bottom of the two duplexes together. Figure 1-14 illustrates these motions.

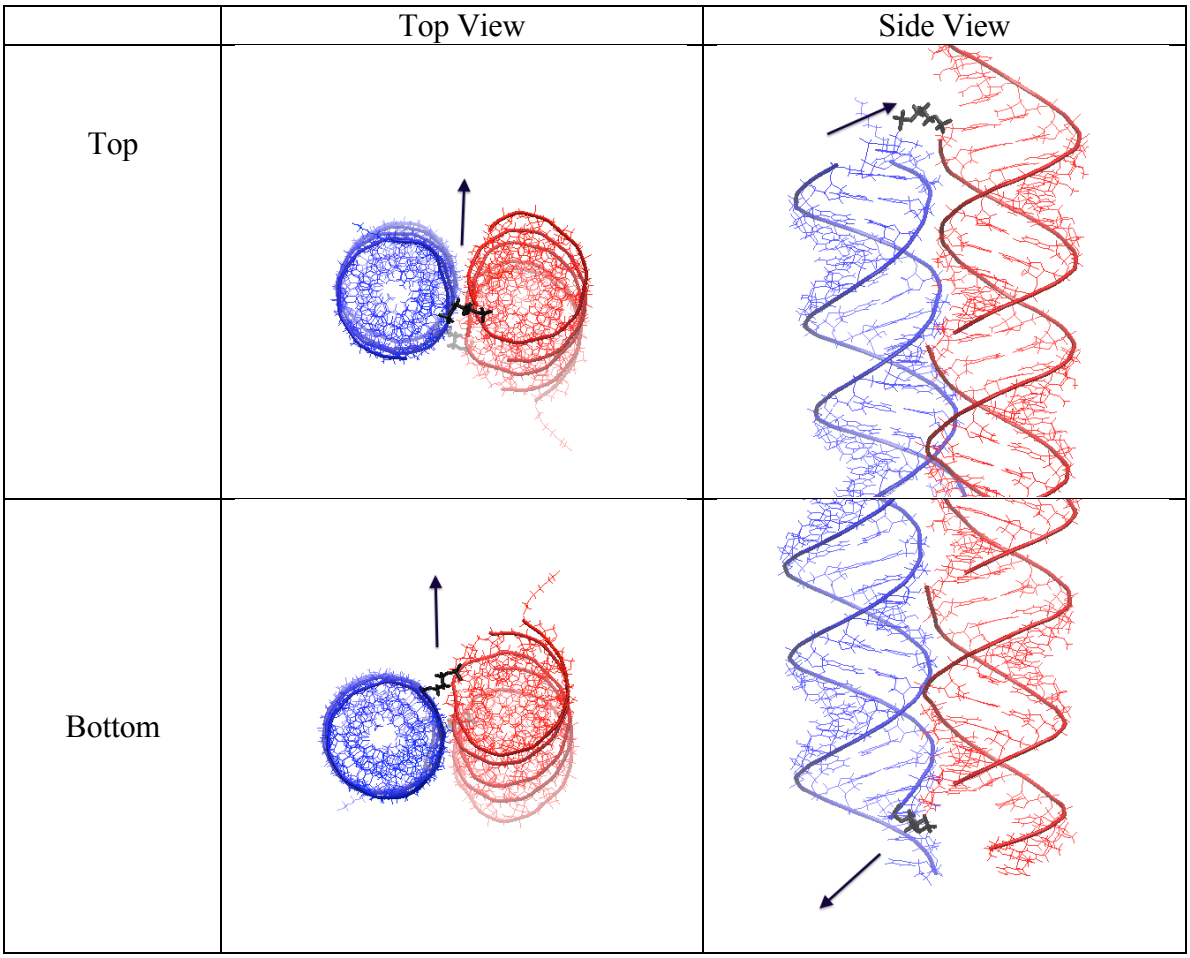


Figure 1-14 Direction of movement of the three-carbon linkers in response to unwinding of sensor duplex.

In the case of construct 2, the equilibrium pitch of the sensor is 12.08 base pairs per turn, so it is also driven to unwind relative to the ideal A-RNA. In the earlier portion of the trajectory, both siRNA and the sensor look similar to their counterparts in construct 1; the top of the siRNA is deformed, and the bottoms of the siRNA and sensor duplex are brought together. Call this conformation A. By the end of the trajectory, however, the sensor duplex has moved upward along the helical axis relative to the siRNA and neither duplex shows much deformation. Call this conformation B. Given the similarity in equilibrium

pitch, it is surprising that constructs 1 and 2 demonstrate such different final structures. One possible explanation is that construct 1 can achieve conformation B, but it simply didn't sample that conformation in the course of the trajectory. If true, the next questions to ask are whether there is a substantial energetic barrier to transition from A to B and what contributes to the barrier. Assuming there is a large barrier, one culprit could be the 5' toehold. Construct 1 has one, and construct 2 does not. It could be that the interaction of the toehold with the siRNA in construct 1 impedes the transition. Another possible explanation is that the conformation of the three-carbon linkers is important. In order for one duplex to translate along the helical axis relative to the other, each linker must adopt a particular conformation. If there are large barriers to the conformational sampling of the linkers, this could impede the transition. It may also be that conformation B is simply not as energetically favorable for construct 1, and so it does not move toward that conformation.

Regarding construct 3, both the siRNA and sensor duplex equilibrium pitches are close to the ideal A-RNA pitch of 11 base pairs per turn, so there is little unwinding. Even so, there is some minor deformation in both the siRNA and sensor duplexes. A unique feature of construct 3 is a long 5' toehold that is observed to interact with the siRNA during the trajectory. While the interactions are not strong, they may be sufficient to influence the overall structure of each duplex.

The second goal of the simulations is to determine how well the sensor duplex protects the Dicer cleavage site. Ultimately the best measure of protection would be to dock a molecular model of Dicer to the various conformations of the *Cond*-siRNA in the trajectory, but just observing the position of the Dicer cleavage site in a qualitative way still gives some insight. It is assumed that the initial structure is close to the ideal in terms of protection. Examining the trajectory shows that the two main modes of deviation from this ideal are the Dicer cleavage site translating away from the sensor duplex through bowing of one or both duplexes, and turning away from the sensor duplex through rotation of the siRNA along the helical axisfire. Figure 1-15 shows frames from the trajectory showing these deviations in the case of construct 1. Both the bowing and turning can be attributed to the deformations that occur as a result of the sensor duplex unwinding to its equilibrium pitch.

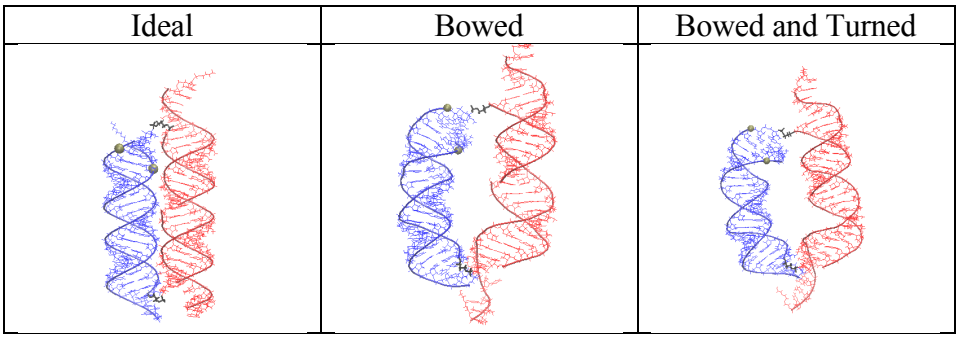


Figure 1-15 Variance in the position of the Dicer Cleavage Site (gold spheres) of the siRNA relative to the sensor duplex. The ideal is the initial position.

Discussion

Review of the helical parameters across all three constructs indicates that the extensive modifications and three-carbon linkage do not destabilize either duplex. However, a duplex with a pitch that deviates from the ideal for A-RNA will relax to the equilibrium pitch and may lead to deformation in a linked duplex. These deformations lead to some variance in the level of protection afforded by the sensor duplex, but on the whole, the Dicer cleavage site is well-protected. To know whether the variance in protection is problematic, it must be known where the sensor duplex must be in order to prevent Dicer from cleaving.

Assuming that the initial conformation is best for protection, and that Dicer activity is sensitive to the variation in protection observed in Figure 1-15, then avoiding deformations in each duplex should lead to better protection. The deformation seen among the constructs is largely attributable to the sensor duplex dragging the termini of the siRNA duplex with it as the sensor duplex moves towards its equilibrium pitch. One way to alleviate the problem is to increase the linker length, perhaps to a four- or five-carbon chain. The longer chain would allow the sensor duplex to relax while not forcing the siRNA to deform. The danger in this approach, of course, is that increasing the linker length would increase the region between the duplexes and allow a greater opportunity for

Dicer to cleave the siRNA. There may be an optimum linker length that prevents deformation even in the face of disparate pitches, but does not allow the duplexes to drift so far as to allow Dicer to approach the cleavage site. Another approach would be to design sensor duplexes that have natural pitches closer to that of the ideal A-RNA so that the unwinding and associated deformation does not occur. Construct 3 is an illustration of this approach.

Given a pitch and linker length, there are still some changes that could alleviate the deformations. One approach would be to modulate the mechanical properties of the duplex. The roll data for the siRNA suggests that bending tends to happen at steps where purine-purine clashes can take place. Perhaps eliminating these steps would avoid weak points for bending the duplex. Understanding the reason for the difference in final conformation of constructs 1 and 2 could yield another means of avoiding deformation. If a 5' toehold were primarily to blame for the barrier to transition between conformations A and B, then it would be best to use 3' toehold. If the limited conformations of the three-carbon linkers are to blame, then there is even more reason to use a longer linker that would allow for greater motion between the duplexes.

It is important to note that the protection of the dicer cleavage site only partly governs the behavior of the *Cond*-siRNA. The behavior is also governed by activation through toehold-mediated strand displacement. This process occurs on the order of milliseconds, so MD can't address it directly. However, some inferences may be made by considering how the structure of the *Cond*-siRNA might impact the process of strand displacement. For example, during strand displacement, a Holliday Junction will be present involving the sensor strand, the core strand, and the invading strand. Crowding around this branch point could inhibit the progression of the displacement. If that were true, then proximity between the terminus of the sensor duplex and the siRNA could inhibit activation. The final conformations of both construct 1 and 2 could lead to this sort of inhibition. In construct 1, the clash between the termini of each duplex could create an environment with little room for the junction. Even in construct 2 where there is no such clash, on the 5' side of the sensor strand, the initial invasion would take place on bases situated in the region between the siRNA and sensor duplexes where there would be greater

crowding. These hypotheses offer a plausible explanation for the experimental data in which Han *et al.* found that only the 3' toehold of construct 1 was viable for activation. Both of these hypotheses could be investigated by further simulation.

These results must be considered in the context of the limitations of the molecular dynamics simulation. With respect to the atomic model, a significant limitation is the use of fixed charges. Charges are environment-dependent. In places like the junction at the bottom of construct 1 where the siRNA and sensor duplex clash, the charges would certainly deviate from the fixed charges and perhaps lead to a different conformation. Using a technique like Charge Equilibration (Qeq) could help on this front.[20] Another concern is conformational sampling. The initial position for each construct was chosen under the assumption that it would afford the best protection for the Dicer cleavage site. The simulations can therefore be interpreted as asking how stable that initial orientation is. The simulation does not necessarily explore the entire conformational landscape of the conditional RNA. The global energetic minimum could look considerably different. Using techniques like umbrella sampling or metadynamics could explore this landscape.

Cond-siRNAs offer an exciting therapeutic opportunity. By causing the activation of RNAi to be context-dependent, more potent siRNAs can be used at lower dosages. Structure is critical to the proper function of the *Cond*-siRNA. These simulations show that the extensive modifications required to stabilize the desired structure over problematic secondary structures does not compromise either duplex. Further, while the equilibrium pitch of the sensor duplexes can deform the siRNA considerably, this deformation does not seem to significantly compromise the capacity of the sensor duplex to protect the Dicer cleavage site. These deformations may, however, impede the process of toehold-mediated strand displacement. In future designs of the *Cond*-siRNA, means of alleviating the stresses caused by non-ideal duplex pitch should be considered to make both cleavage site protection and activation more consistent.

References

1. Han, S.-P., Scherer, L., Gethers, M., Mancusi, R., Kuo, Y., Marcucci, G., Rossi, J., Goddard III, W.A., Conditional small interfering RNAs programmable for intracellular activation by specific RNA transcripts. In prep.

2. Lee, R. C.; Feinbaum, R. L.; Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **1993,** *75,* 843-854.

3. Carthew, R. W.; Sontheimer, E. J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **2009,** *136,* 642-55.

4. Wu, S. Y.; Lopez-Berestein, G.; Calin, G. A.; Sood, A. K. RNAi Therapies: Drugging the Undruggable. *Sci. Trans. Med.* **2014,** *6,* 240PS7.

5. Saenger, W., Principles of Nucleic Acid Structure. In Springer-Verlag 1984.

6. Wilson, R. C.; Doudna, J. A. Molecular Mechanisms of RNA Interference. *Annu. Rev. of Biophys.* **2013,** *42,* 217-239.

7. Jensen, F., Introduction to computational chemistry. In John wiley & sons: 2017.

8. Macke, T. J.; Case, D. A., Modeling Unusual Nucleic Acid Structures. In *Molecular Modeling of Nucleic Acids,* American Chemical Society: 1997; 682, 379-393.

9. Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000,** *21,* 1049-1074.

10. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004,** *25,* 1157-1174.

11. Condon, D. E.; Yildirim, I.; Kennedy, S. D.; Mort, B. C.; Kierzek, R.; Turner, D. H. Optimization of an AMBER Force Field for the Artificial Nucleic Acid, LNA, and Benchmarking with NMR of L(CAAU). *J. Phys. Chem. B* **2014,** *118,* 1216-1228.

12. Venkateswarlu, D.; Lind, K. E.; Mohan, V.; Manoharan, M.; Ferguson, D. M. Structural properties of DNA:RNA duplexes containing 2′-O-methyl and 2′-S-methyl substitutions: A molecular dynamics investigation. *Nucleic Acids Res.* **1999,** *27,* 2189-2195.

13. Lind, K. E.; Sherlin, L. D.; Mohan, V.; Griffey, R. H.; Ferguson, D. M. Parameterization and Simulation of the Physical Properties of Phosphorothioate Nucleic Acids. *ACS Symp. Ser.* **1997,** *682,* 41-54.

14. Case, D. A.; Darden, T. A.; Cheatham Iii, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M. AMBER 12; University of California: San Francisco, 2012. **2010,** 1-826.

15. Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. The R.E.D. Tools: Advances in RESP and ESP charge derivation and force field library building. *Phys. Chem. Chem. Phys. : PCCP* **2010,** *12,* 7821-7839.

16. Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comp. Phys.* **1995,** *117,* 1-19.

17. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996,** *14,* 33-38.

18. Lu, X. J.; Olson, W. K. 3DNA: a software package for the analysis, rebuilding and visualization of three‐dimensional nucleic acid structures. *Nucleic Acids Res.* **2003,** *31,* 5108-5121.

19. Calladine, C. R. Mechanics of sequence-dependent stacking of bases in B-DNA. *J. Mol. Biol.* **1982,** *161,* 343-352.

20. Rappe, A. K.; Goddard Iii, W. A. Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* **1991,** *95,* 3358-3363.

Chapter II

Holey Graphene as a Weed Barrier for Molecules

Adapted with permission from:
Gethers, M. *et al.* (2015). "Holey Graphene as a Weed Barrier for Molecules". In: *ACS Nano* 9, pp. 10909-10915. doi: 10.1021/acsnano.5b03936.

Introduction

DNA, RNA, proteins, and polysaccharides are polymers key to biology, biological engineering, and medicine. As these disciplines have advanced, they have increased the demand for longer and more complicated instances of these polymers at greater yields and lower costs and turn around times. Means of synthesis have improved, but their capabilities are being strained by increasingly ambitious goals. The workhorse for synthesis of these polymers is the solid-phase approach in which monomers are sequentially added to a growing chain until the full polymer is synthesized.[1] The problem with solid-phase synthesis is that monomer additions fail, leading to polymers with missing elements. Even with a coupling efficiency of 99.5%, 1000 coupling events leads to a yield of less than 1%. In the case of DNA, the route to longer products is to use the solid-phase approach to synthesize short oligos at high yield, and then enzymatically assemble these oligos into the longer construct.[2] While effective, these enzymatic assemblies become complicated with longer constructs, making success less certain.

A fast, cheap, and reliable way to synthesize long polymers would give a strong boost to research in a variety of disciplines. The sequence and dispersion control from a more powerful method would be useful for polymers generally, not just those important for biology. The key to an improved method would be to overcome the ruinous effect of failed coupling events between monomers. One approach to achieve this end is to arrange and fix the monomers in the desired sequence first, and then initiate coupling of all monomers simultaneously. If monomers could be arranged and fixed in this way, they could only interact with the correct partners. Further, by separating the arrangement and coupling steps, the approach would allow for correction in the case of a missing or incorrect monomer.

Using a surface as a scaffold for arrangement and coupling requires a means of precisely addressing monomers to the surface. The self-assembled monolayer (SAM) serves as a good starting point.[3-6] In the most popular example of a SAM, a Au{111}

surface is incubated with a solution of *n*-alkanethiols. The thiol group binds strongly to the gold and the alkyl "tail" group orients away from the surface and forms a semi-crystalline layer. Replacing the tail of the SAM with a monomer creates a means of attaching monomers to the surface in a high-throughput way. To meet the requirements for sequence control, self-assembly would further need to direct particular molecules to specified loci. To this end, a promising approach comes from Battaglini *et al.* who were able to use cobalt as a mask that could be used to pattern the deposition of the SAM.[7] The cobalt mask embodies a number of limitations, so for greater patterning control, we consider a different mask: graphene. While current techniques to pattern graphene place single molecule deposition out of reach, advances may one day make this possible, and the widespread use of both SAMs and graphene in other parts of materials science suggests that even achieving relatively large features with a graphene mask could be valuable.

Background

The extraordinary electronic, thermal, and mechanical properties of graphene have been elaborated and exploited.[8] Graphene's high carrier mobility and ambipolarity make it a potentially powerful component in electronic systems.[9] An equally intriguing aspect of graphene is its capacity to act as an impermeable or semipermeable membrane; Bunch *et al.* demonstrated the impermeability of graphene to helium through the inflation of a "nanoballoon".[10 11] Graphene can also act as an effective barrier to oxidation of metal surfaces under certain conditions.[12-14] The purposeful introduction of pores into graphene tunes this permeability by allowing certain molecules to pass through while others are inhibited. This use of graphene has led to proposals as varied as desalination and DNA sequencing.[15-21] Here, we demonstrate and explore the use of "holey" graphene as a molecular barrier by applying it to adsorption and self-assembly.

Self-assembly provides a convenient route towards the bottom-up placement of single molecules with applications ranging from nanotechnology to biology.[4 22-24] Molecules for self-assembly typically comprise an attaching head group, an interacting

backbone, and a functional tail group. The head group binds the molecule to a substrate, backbone intermolecular interactions lead to crystalline packing (through design), and the exposed terminal functional group can tune interfacial chemical properties between the substrate and its environment.[25] Molecular monolayers enable controllable surface functionalization and can be used to isolate and to study individual molecules.[26-28] Self-assembly is made even more powerful when combined with patterning. Currently, patterning of SAMs is achieved through conventional, soft, or hybrid lithographies that are limited by the conflicting requirements of feature resolution and large-scale fabrication, where manufacturing cost and assembly time each play key roles.[29-33] Inspired by the approach of Battaglini *et al.*, we pattern SAMs by masking the surface with an inert material.[7] We find that graphene can function as such a mask, as it is a material with relatively inert chemistry[34] and functions as an impermeable barrier against other molecules. Our choice was also influenced by the number of techniques that enable the introduction of nanoscale pores of arbitrary size and location to graphene, including both electron-beam bombardment and chemical approaches.[35-44] These techniques will ultimately provide flexibility and precision in pattern shape and scale, including precise hole size and pitch. Used in this way, graphene acts as a "weed barrier", where it restricts covered areas from forming Au-S attachments (molecules "taking root") and thus enables only the open regions exposed by the pores to form substrates covered by self-assembled monolayers (akin to gardening where, ideally, plants only grow in the holes deliberately placed in the weed barrier).

At the same time, the graphene-Au interaction is relatively weak, as observed in a number of quantitative and spectroscopic studies of aromatics on Au{111}.[45-48] For example, aromatics sit much farther from and interact much less strongly with the Au{111} surface than from the close-packed Ag{111} and Cu{111} surfaces.[48]

Materials and Methods

Holey Graphene Sample Preparation

Graphene was synthesized on a 25-μm-thick copper foil (99.8%, Alfa Aesar, Ward Hill, MA) that was treated with hydrochloric acid/deionized water (1:10) (36.5-38.0%, Sigma-Aldrich, St. Louis, MO) for 30 min and rinsed by isopropyl alcohol (99.8%, Sigma-Aldrich, St. Louis, MO) for 10 min. After drying under an N_2 stream, the copper foil was loaded into the chemical vapor deposition (CVD) furnace (1-inch tube diameter; Lindbergh/Blue M, Thermo Scientific, Waltham, MA). The system was pumped down to a vacuum of 10 mTorr in 30 min and refilled with 300 sccm H_2/Ar flow (25 sccm/475 sccm) and heated to 1040 ˚C within 25 min. Next, diluted methane and Ar were introduced into the CVD system for graphene growth at 1040 ˚C for 90 min (500 ppm methane in Ar, 35 sccm) with H_2/Ar (25 sccm/440 sccm). All process gases were supplied by Airgas, Inc. (Burbank, CA).

Graphene was grown on both sides of the copper foil, and one side of the graphene/copper surface was spin-coated with poly(methyl methacrylate) (PMMA; 495 PMMA C_2, MicroChem, Newton, MA) and baked at 140 ˚C for 5 min. The other side of the copper foil was exposed to O_2 plasma for 1 min to remove the graphene. After that, the Cu foil was etched away using copper etchant (ferric chloride, Transene), resulting in a free-standing PMMA/graphene membrane floating on the surface of the etchant bath. The PMMA/graphene film was washed with HCl/deionized H_2O (1:10) and deionized water several times, and then transferred onto a 300-nm-thick SiO_2 substrate. After air-drying, the PMMA was dissolved by acetone and the substrate was rinsed with isopropyl alcohol to yield a graphene film on the substrate.

A 2-nm-thick gold film was deposited using thermal evaporation onto the graphene/SiO_2 substrate. After annealing at 350 ˚C for 15 min, gold nanoparticles were found on the substrate. The holey graphene is oxidized by exposure to oxygen in the ambient air, with gold nanoparticles acting as the catalyst. The gold nanoparticles were removed by gold etchant (KI/I_2 solution, Sigma-Aldrich, St. Louis, MO) and washed with isopropyl alcohol and deionized water. The graphene/SiO_2 was again spin-coated with

PMMA, and the SiO_2 substrate was etched away using a buffered oxide etch. The PMMA-coated holey graphene was washed in deionized water and transferred to a deionized water bath. A H_2 flame-annealed (at a rate of 1 Hz, 10 passes) Au{111}/mica substrate (Agilent, Santa Clara, CA) was then used to scoop the PMMA-coated graphene from the water bath. The PMMA/graphene/Au substrate was allowed to air dry overnight, and then the PMMA was dissolved in acetone and the graphene/Au substrate was washed with isopropyl alcohol.

Transmission Electron Microscopy

The morphology and structure of the graphene were characterized with field emission high-resolution transmission electron microscopy (TEM; FEI Titan S/TEM), typically with an accelerating voltage of 300 kV. The diffraction patterns were collected with accelerating voltages of 300 kV to assess whether the beam energy played a role in graphene surface changes. Specimens for TEM analysis were prepared by the same as the process of graphene transfer onto 200 mesh formvar/copper grids purchased from Ted Pella, Inc. (Redding, CA).

Scanning Tunneling Microscope Sample Preparation

Holey graphene was deposited onto flame annealed, commercially available Au{111} on mica substrates. Samples were imaged and then subsequently annealed at 100 ˚C for a period of 24 h in a gasketed glass v-vial (Wheaton, Millville, NJ). Samples were heated in a chamber of a Barstead Thermolyne 1400 furnace (ThermoFischer Scientific, Waltham, MA). Samples were taken out and imaged with STM, then placed back into a clean v-vial above a solution of 1 mM commercially available 1-adamantanethiol (Sigma-Aldrich, St. Louis, MO) in ethanol for vapor deposition. Vials were placed back into a preheated furnace at 78 ˚C for a period of 24 h. Inserted 1-adamantanethiolate holey graphene samples were taken out for STM imaging. After sufficient experiments were performed, samples were placed back into a preheated furnace

at 250 ˚C for a period of 24 h for molecular desorption. Samples were then taken out for subsequent imaging and desorption confirmation.

Imaging

All STM measurements were performed in air using a custom beetle-style scanning tunneling microscope and a platinum/iridium tip (80:20).[49] The known lattice of 1-dodecanethiolate SAMs on Au{111} was used to calibrate the piezoelectric scanners. The sample was held between -1 V to -0.1 V bias range, and 256 × 256 pixel images were collected, at varying size, in constant-current mode with a tunneling current ranging from 2 to 80 pA. There is a strong tip dependence for imaging cage molecules, as reported previously.[50]

Image Analyses

All STM images were initially processed with automated routines developed in MATLAB (Mathworks, Natick, MA) to remove any high-frequency noise and intensity spikes that may otherwise impair reliable segmentation.[51] Images used to obtain nearest-neighbor spacings were resized to account for drift at room temperature. Transmission electron microscopy images were thresholded to segment both graphene holes and the graphene layer that was used to create a binary mask, where the average diameter of the holes was computed. Orientations in diffraction images obtained by TEM were highlighted by computing local maxima within a defined pixel radius (10 pixels), referenced, and plotted for clarity. The nearest-neighbor spacing of graphene was computed in Fourier space for the pre-1-adamantanethiol deposition and post-annealing experiments. The spacings of assembled 1-adamantanethiol and the surrounding graphene were determined by fitting the centers of the molecules using a binary mask generated through thresholding and the Regionprops function in Matlab. Values obtained by Fourier analysis and Regionprops fitting on the same image were compared to ensure that results were similar. Apparent height was also used for image segmentation and to determine separation distances in the z-direction.

Results

The process developed for producing spatially patterned monolayers on Au{111} using a graphene mesh is shown schematically in Figure 2-1. “Holey” graphene is fabricated by depositing graphene on a SiO_2 substrate,[52-54] then evaporating a thin layer of Au (2 nm) onto the exposed graphene layer. Subsequent annealing forms surface-bound Au nanoparticles. The Au nanoparticles catalyze oxidation of the graphene by oxygen in the air, thereby forming pores. The Au nanoparticles are then etched *via* brief immersion into an etchant solution (see Materials and Methods for details). A thin protecting layer of poly(methyl methacrylate) (PMMA) is added to facilitate transfer, and the “holey” graphene is transferred onto a Au{111} substrate. The protecting layer is removed and samples are ready for characterization. Further annealing at 100 ˚C removes any excess solvent, and the covered Au{111} substrate is thereby primed for molecular deposition.

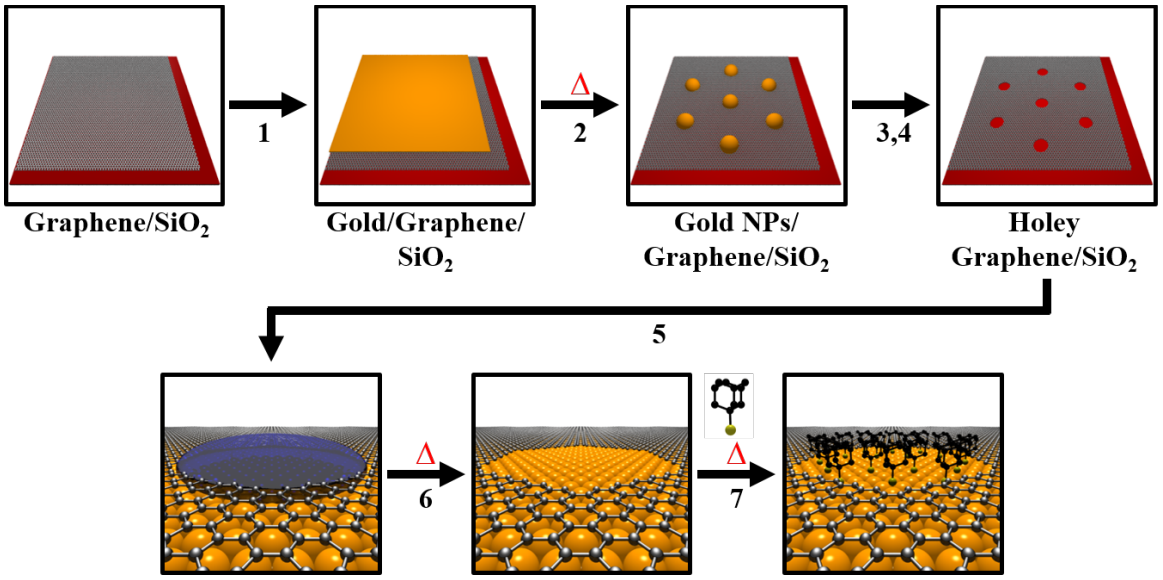


Figure 2-1 Process for producing spatially patterned monolayers on Au{111} using a graphene mesh. From a monolayer sheet of graphene on a SiO_2 substrate, (1) 2 nm of Au is deposited and (2) then annealed for 15 min at 350 ˚C. (3) The Au is etched (KI/I_2, solution) for 30 sec and (4) washed in DI water for 30 sec. (5) "Holey" graphene is then transferred to a Au{111}/mica substrate and (6) annealed at 100 ˚C for 24 h. (7) The same substrate is then exposed to the vapor of a 1 mM ethanolic solution of 1-adamantanethiol (**1AD**) at 78 ˚C for 24 h for deposition.

The fabrication of porous graphene is confirmed by transmission electron microscopy (TEM), where TEM images show a graphene mesh with randomly distributed holes; measured holes have an average diameter of 37 ± 8 Å (Figure 2-2 and appendix Figure A2-1). Images also depict cracks in the graphene induced by the transfer and annealing processes. Graphene is known to retain the surface morphology of the substrate on which it was synthesized even when attached to the PMMA overlayer.[55] When transferred to the final substrate, this morphology results in gaps between the graphene and the substrate that can cause folding and cracking when the PMMA is removed. Further, water caught between the graphene and the substrate can leave gaps between the graphene and substrate upon drying that lead to folds, thereby appearing like graphite in images.[55] After the mesh is successfully transferred to Au{111} and annealed, scanning tunneling microscopy (STM) is used to probe the local environment.

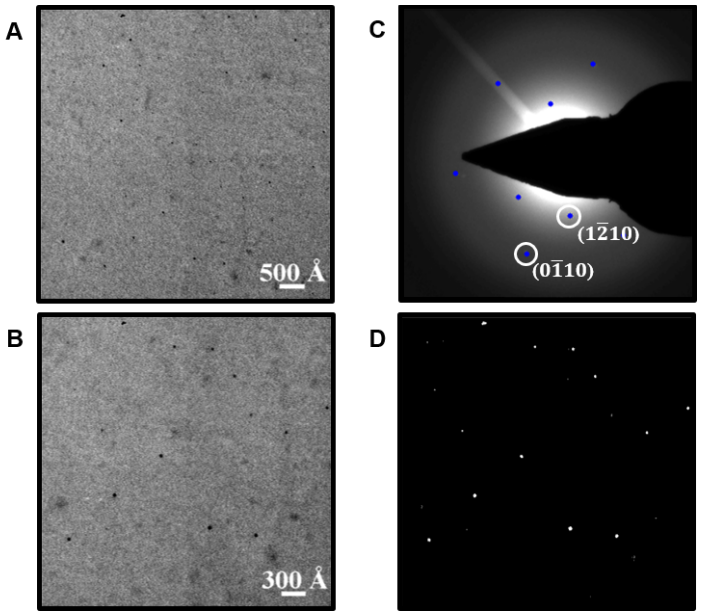


Figure 2-2 (A, B) "Holey" graphene measured with transmission electron microscopy (TEM) supported on a 200 mesh formvar/copper grid. Each image was acquired at an accelerating voltage of 300 kV using a FEI Titan microscope. Holes measured with TEM are 37 ± 8 Å in diameter and are randomly distributed across the graphene layer. (C) A diffraction image of B is shown, where the hexagonal pattern of graphene is observed. Orientation points are highlighted, for clarity, in blue, where local image maxima were computed within a defined pixel window. (D) A thresholded binary of B that highlights randomly distributed holes within the graphene mesh.

The scanning tunneling microscope provides a window into the nanoscopic world, where constant-current imaging measures a convolution of electronic and topographic structure as a function of position across surfaces.[51 56 57] Measurements are recorded on a custom-built, ultrastable microscope held at ambient temperature and pressure.[49] Scanning tunneling micrographs before annealing are shown in appendix Figure A2-2, where a large depression (pore) in the center of the image is noted that is surrounded by other pores filled with residual solvent from the transfer step. Annealing removes the solvent within the pores. The annealed graphene-gold surface is shown in Figure 2-3, where images depict porous graphene with hole diameters that match TEM measurements. The surrounding

graphene Moiré pattern shows a six-fold symmetry with a nearest-neighbor distance of 5.0 ± 0.5 Å, which is in good agreement with the predicted and energetically favorable (2 × 2) superstructure for graphene on a Au{111} substrate.[58]

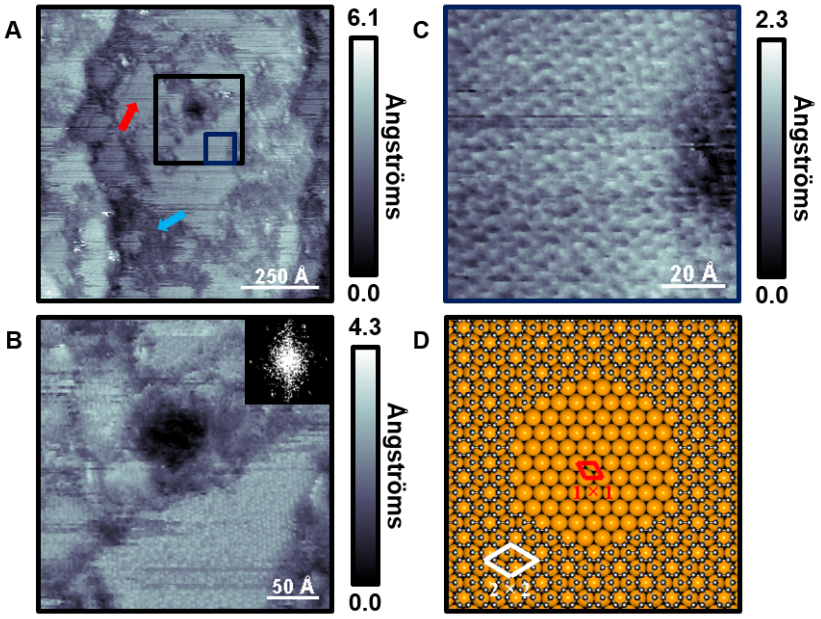


Figure 2-3 (A) Scanning tunneling micrograph ($I_{tunneling}$ = 3 pA, V_{sample} = -1.0 V) of "holey" graphene on Au{111}/mica along two monoatomic step edges after annealing at 100 ˚C for 24 h. We highlight an ordered region (red arrow, higher in conductance), and an area of disorder (blue arrow, lower in conductance) that we attribute to possible cracks, folds, and varied underlying graphene-Au attachments. (B) Higher resolution of the larger box in A. (C) Higher resolution image of the smaller box in (A). Inset in B shows a fast Fourier transform, where graphene displays hexagonal nearest-neighbor spacings of 5.0 ± 0.5 Å. (D) A schematic showing a pore in graphene exposing the underlying Au{111} substrate that further depicts the measured (2 × 2) Moiré superstructure of graphene on Au.

The structure of graphene on Au{111} is difficult to predict and likely to be locally varied, where measured superlattices are highly influenced by both the underlying Au substrate and the detailed structure of the STM tip.[52 59] With this caveat in mind, acquired STM images confirm a single transferred layer of holey graphene with exposed Au regions, where image differences were quantified in real and Fourier space. Thresholding and

masking techniques, performed in MATLAB, enable gold and graphene regions to be segmented and compared (appendix Figure A2-3). In STM images, under the conditions used, graphene layers are 2.1 ± 1.1 Å more protruding in apparent height compared to exposed Au regions. The same sample is then exposed to the vapor of a 1 mM ethanolic solution of the self-assembling cage molecule 1-adamantanethiol (**1AD**) and subsequently imaged.[60]

The diamondoid **1AD** is ideal for an initial patterning test, in that it is commercially available, forms well-ordered monolayers with few defects (due to limited degrees of freedom), and has a well-defined structure.[61-64] Scanning tunneling micrographs recorded after deposition show islands of molecular protrusions consistent with the diameters of the pores (Figure 2-4). Nearest-neighbor distances within measured molecular protrusions (7.2 ± 1.1 Å) are near previously recorded distances of **1AD** on Au{111} (6.9 ± 0.4 Å) (appendix Figure A2-4).[61, 65] The areas surrounding the islands proved difficult to resolve; however, there were small, well-resolved areas bearing molecular features that were analyzed (appendix Figure A2-5). Several of these areas topographically resemble the graphene overlayer imaged prior to **1AD** deposition. The average nearest-neighbor spacings of these areas (5.0 ± 1.1 Å) is consistent with the graphene overlayer. The difference in nearest-neighbor spacings provides evidence of separate molecular domains brought about due to the graphene masking effect against the deposition of **1AD**. Graphene regions show a different apparent height than **1AD** islands and can thereby be segmented (appendix Figure A2-6), where **1AD** patches differ by 1.1 ± 0.5 Å in the z-direction under the STM imaging conditions used. Measured spacings, both in the lateral and surface normal (apparent height) directions, and consistent hole diameters confirm the blocking effect of the graphene layer. The same samples are then annealed again to test if molecular desorption can be achieved, and thus if the bare surface in the pores of the graphene mask can be regenerated.

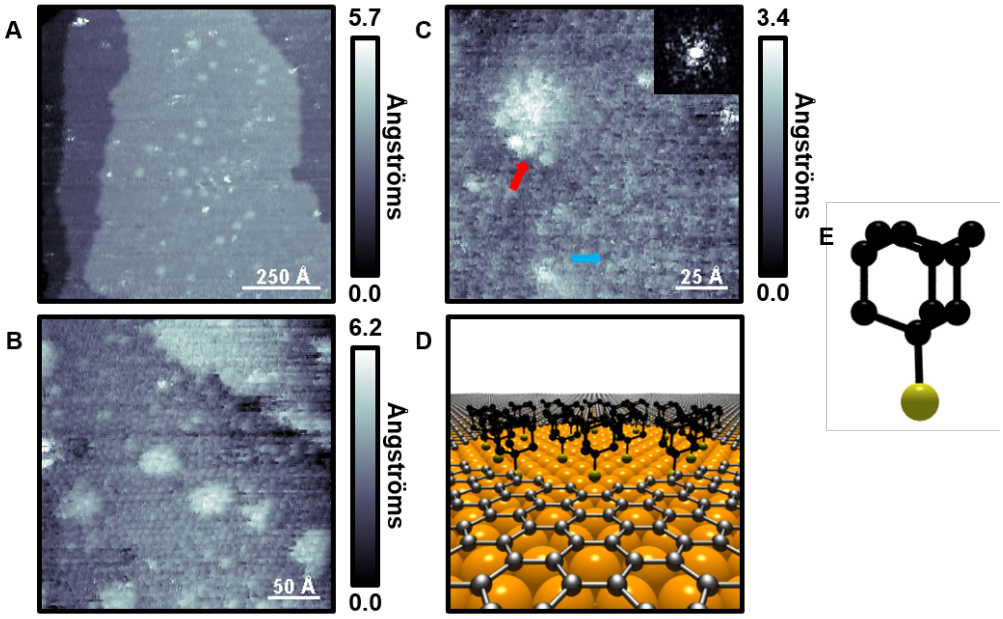


Figure 2-4. (A) Scanning tunneling micrograph ($I_{tunneling}$ = 3 pA, V_{sample} = -1.0 V) of "holey" graphene on Au{111}/mica after exposure to the vapor of a 1 mM ethanolic solution of 1-adamantanethiol (**1AD**). (B, C) Two regions where **1AD** has assembled on Au{111} within the confines of the pores of the holey graphene. Here, we highlight protrusions due to **1AD** (red arrow) and areas of graphene (blue arrow) for clarity. (C, inset) A fast Fourier transform shows local order of the graphene overlayer with nearest-neighbor spacings of 5.0 ± 1.1 Å. Molecular distances for **1AD** were computed in real space, with nearest-neighbor spacings of 7.2 ± 1.1 Å (see Figure S4 for more detail). (D) A schematic of the arrangement in (C) where the graphene pore is filled with assembled **1AD**. (E) Ball-and-stick model of the **1AD** molecule with hydrogens not shown.

Scanning tunneling topographs before and after this second anneal, to 250 ˚C, are shown in Figure 2-5, where evidence of molecular desorption is obtained.[24 66 67] Once-filled holes are now empty and the hexagonal spacing of 5.0 ± 0.5 Å is recovered outside graphene pores. We confirm desorption by topographic imaging, where the exposed areas within the mask are destructively regenerated and thus prepared for further molecular deposition steps (appendix Figure A2-7). We note that non-destructive methods such as displacement techniques can also be applied, since **1AD** has been shown to be labile upon exposure *via* solution, vapor, or contact to more strongly bound self-assembling molecules.[63 64]

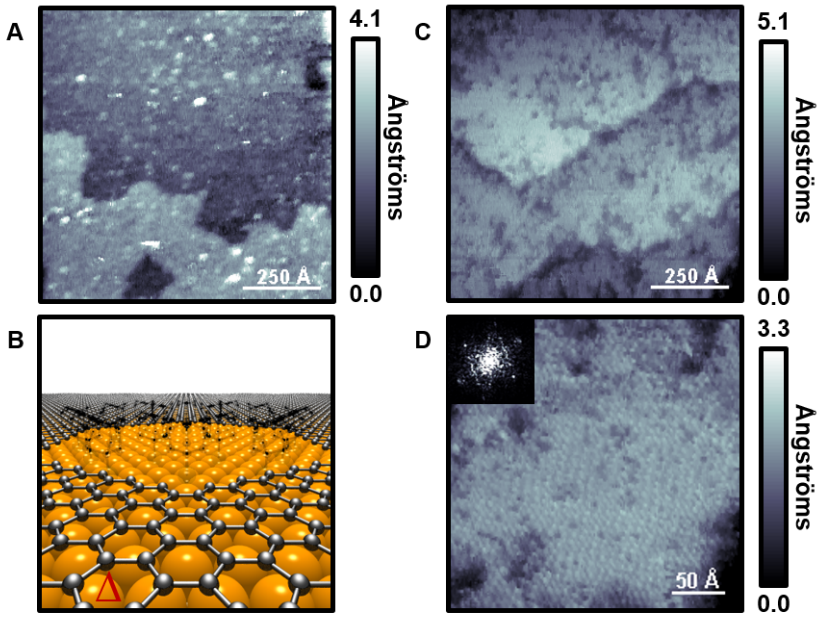


Figure 2-5 (A) Scanning tunneling micrograph ($I_{tunneling}$ = 3 pA, V_{sample} = -1.0 V) of "holey" graphene with the 2D pores filled with assembled 1-adamantanethiol on Au{111}/mica. (B) Annealing at 250 ˚C for 24 h removes adsorbates from the pores, as shown schematically. (C, D) Scanning tunneling micrographs ($I_{tunneling}$ = 3 pA, V_{sample} = -1.0 V) of the same sample after complete molecular desorption, recorded at two different resolutions, as indicated. (D, inset) A fast Fourier transform shows the recovered hexagonal spacing (5.0 ± 0.5 Å) measured previously.

Discussion

The effectiveness of graphene as a mask against adsorption depends on the integrity of the graphene. Defects, folds, wrinkles, and strain can all compromise the impermeability of the mask. Small defects and cracking permit penetration of adsorbates through to the underlying substrate. Folding leads to multilayered regions where a pore in one layer can be occluded by another layer that is not porous. A strained lattice could open gaps in the mesh and induce tearing at pores. Second to the synthesis of the graphene itself, the transfer of the graphene to the target substrate is the most important process in minimizing these

undesirable features. Improving the transfer method can alleviate some of these issues. Liang *et al.* point out several ways in which the transfer can be improved, including increasing the hydrophobicity of the target substrate, annealing the graphene/substrate complex before dissolving the PMMA, and using the modified "RCA Clean" method to get rid of residual agents.[55, 68-71] While our method patterns graphene destructively, another approach is to employ bottom-up methods to graphene synthesis that enable the placement and design of desired structures with predetermined pore size, shape, and pitch.[72-75]

The graphene mask approach embodies a wholly different set of strengths and weaknesses than photolithographic and soft-lithography approaches.[30, 32] Patterning is achieved with electron-beam and chemical methods, surpassing the diffraction limit of photolithographic methods and placing it in league with scanning probe lithography in feature size. The graphene mask, akin to the PDMS stamp and unlike the scanning probe lithography techniques, is reusable, and can cover large areas. Our approach permits control over the pattern, overcoming a major limitation of Battaglini's approach, though it is not removable and thus cannot be backfilled. The value of graphene as a mask is further enhanced by its thermal stability, making it resilient to high-temperature fabrication processes.[76] Importantly for SAMs, thermal stability implies that the integrity of the graphene should not be compromised when annealing the masked substrate to remove adsorbates and thereby to regenerate the mask.

Our results show that graphene can serve as a barrier to adsorption and open up a plethora of future patterning experiments. Since graphene pores can be readily manufactured, masks can be used with a wide variety adsorbates with implications ranging from constructing well-defined nanoarchitectures to patterning biosensors.

References

1. Kosuri, S.; Church, G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Meth.* **2014,** *11,* 499-507.

2. Stemmer, W. P. C.; Crameri, A.; Ha, K. D.; Brennan, T. M.; Heyneker, H. L. Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* **1995,** *164,* 49-53.

3. Ulman, A. Formation and Structure of Self-Assembled Monolayers. *Chem. Rev.* **1996,** *96,* 1533-1554.

4. Poirier, G. E. Characterization of Organosulfur Molecular Monolayers on Au (111) Using Scanning Tunneling Microscopy. *Chem. Rev.* **1997,** *97,* 1117-1128.

5. Smith, R. K.; Lewis, P. A.; Weiss, P. S. Patterning Self-Assembled Monolayers. *Prog. Surf. Sci.* **2004,** *75,* 1-68.

6. Love, J. C.; Estroff, L. A.; Kriebel, J. K.; Nuzzo, R. G.; Whitesides, G. M. Self-Assembled Monolayers of Thiolates on Metals as a Form of Nanotechnology. *Chem. Rev.* **2005,** *105,* 1103-1170.

7. Battaglini, N.; Qin, Z.; Campiglio, P.; Repain, V.; Chacon, C.; Rousset, S.; Lang, P. Directed Growth of Mixed Self-Assembled Monolayers on a Nanostructured Template: A Step Toward the Patterning of Functional Molecular Domains. *Langmuir* **2012,** *28,* 15095-15105.

8. Novoselov, K. S.; Geim, A. K.; Morozov, S. V.; Jiang, D.; Zhang, Y.; Dubonos, S. V.; Grigorieva, I. V.; Firsov, A. A. Electric Field Effect in Atomically Thin Carbon Films. *Science* **2004,** *306,* 666-669.

9. Allen, M. J.; Tung, V. C.; Kaner, R. B. Honeycomb Carbon: A Review of Graphene. *Chem. Rev.* **2010,** *110,* 132-145.

10. Bunch, J. S.; Verbridge, S. S.; Alden, J. S.; van der Zande, A. M.; Parpia, J. M.; Craighead, H. G.; McEuen, P. L. Impermeable Atomic Membranes from Graphene Sheets. *Nano. Lett.* **2008,** *8,* 2458-2462.

11. Berry, V. Impermeability of Graphene and Its Applications. *Carbon* **2013,** *62,* 1-10.

12. Prasai, D.; Tuberquia, J. C.; Harl, R. R.; Jennings, G. K.; Bolotin, K. I. Graphene: Corrosion-Inhibiting Coating. *ACS Nano* **2012,** *6,* 1102-1108.

13. Nilsson, L.; Andersen, M.; Balog, R.; Lægsgaard, E.; Hofmann, P.; Besenbacher, F.; Hammer, B.; Stensgaard, I.; Hornekær, L. Graphene Coatings: Probing the Limits of the One Atom Thick Protection Layer. *ACS Nano* **2012,** *6,* 10258-10266.

14. Schriver, M.; Regan, W.; Gannett, W. J.; Zaniewski, A. M.; Crommie, M. F.; Zettl, A. Graphene as a Long-Term Metal Oxidation Barrier: Worse Than Nothing. *ACS Nano* **2013,** *7,* 5763-5768.

15. Sint, K.; Wang, B.; Král, P. Selective Ion Passage through Functionalized Graphene Nanopores. *J. Am. Chem. Soc.* **2008,** *130,* 16448-16449.

16. Schneider, G. F.; Kowalczyk, S. W.; Calado, V. E.; Pandraud, G.; Zandbergen, H. W.; Vandersypen, L. M. K.; Dekker, C. DNA Translocation through Graphene Nanopores. *Nano Lett.* **2010,** *10,* 3163–3167.

17. Merchant, C. A.; Healy, K.; Wanunu, M.; Ray, V.; Peterman, N.; Bartel, J.; Fischbein, M. D.; Venta, K.; Luo, Z.; Johnson, A. T. C. DNA Translocation through Graphene Nanopores. *Nano Lett.* **2010,** *10,* 2915–2921.

18. Garaj, S.; Hubbard, W.; Reina, A.; Kong, J.; Branton, D.; Golovchenko, J. A. Graphene as a Subnanometre Trans-Electrode Membrane. *Nature* **2010,** *467,* 190-193.

19. Siwy, Z. S.; Davenport, M. Nanopores: Graphene Opens up to DNA. *Nat. Nanotechnol.* **2010,** *5,* 697-698.

20. Cohen-Tanugi, D.; Grossman, J. C. Water Desalination Across Nanoporous Graphene. *Nano Lett.* **2012,** *12,* 3602–3608.

21. Russo, P.; Hu, A.; Compagnini, G. Synthesis, Properties and Potential Applications of Porous Graphene: A Review. *Nano-Micro Lett.* **2013,** *5,* 260–273.

22. Ulman, A. Formation and Structure of Self-Assembled Monolayers. *Chem. Rev.* **1996,** *96,* 1533–1554.

23. Smith, R. K.; Lewis, P. A.; Weiss, P. S. Patterning Self-Assembled Monolayers. *Prog. Surf. Sci.* **2004,** *75,* 1–68.

24. Love, J. C.; Estroff, L. A.; Kriebel, J. K.; Nuzzo, R. G.; Whitesides, G. M. Self-Assembled Monolayers of Thiolates on Metals as a Form of Nanotechnology. *Chem. Rev.* **2005,** *105,* 1103–1170.

25. Claridge, S. A.; Liao, W.-S.; Thomas, J. C.; Zhao, Y.; Cao, H. H.; Cheunkar, S.; Serino, A. C.; Andrews, A. M.; Weiss, P. S. From the Bottom Up: Dimensional Control and Characterization in Molecular Monolayers. *Chem. Soc. Rev.* **2013,** *42,* 2725–2745.

26. Donhauser, Z. J.; Mantooth, B. A.; Kelly, K. F.; Bumm, L. A.; Monnell, J. D.; Stapleton, J. J.; Price, D. W.; Rawlett, A. M.; Allara, D. L.; Tour, J. M.*, et al.* Conductance Switching in Single Molecules through Conformational Changes. *Science* **2001,** *292,* 2303-2307.

27. Shuster, M. J.; Vaish, A.; Szapacs, M. E.; Anderson, M. E.; Weiss, P. S.; Andrews, A. M. Biospecific Recognition of Tethered Small Molecules Diluted in Self-Assembled Monolayers. *Adv. Mater.* **2008,** *20,* 164–167.

28. Weiss, P. S. Functional Molecules and Assemblies in Controlled Environments: Formation and Measurements. *Acc. Chem. Res.* **2008,** *41,* 1772-1781.

29. Gates, B. D.; Xu, Q.; Stewart, M.; Ryan, D.; Willson, C. G.; Whitesides, G. M. New Approaches to Nanofabrication: Molding, Printing, and Other Techniques. *Chem. Rev.* **2005,** *105,* 1171–1196.

30. Srinivasan, C.; Mullen, T. J.; Hohman, J. N.; Anderson, M. E.; Dameron, A. A.; Andrews, A. M.; Dickey, E. C.; Horn, M. W.; Weiss, P. S. Scanning Electron Microscopy of Nanoscale Chemical Patterns. *ACS Nano* **2007,** *1,* 191-201.

31. Willson, C. G.; Roman, B. J. The Future of Lithography: SEMATECH Litho Forum 2008. *ACS Nano* **2008,** *2,* 1323-1328.

32. Saavedra, H. M.; Mullen, T. J.; Zhang, P.; Dewey, D. C.; Claridge, S. A.; Weiss, P. S. Hybrid Strategies in Nanolithography. *Rep. Prog. Phys.* **2010,** *73,* 036501.

33. Liao, W.-S.; Cheunkar, S.; Cao, H. H.; Bednar, H. R.; Weiss, P. S.; Andrews, A. M. Subtractive Patterning via Chemical Lift-Off Lithography. *Science* **2012,** *337,* 1517-1521.

34. Yan, L.; Zheng, Y. B.; Zhao, F.; Li, S.; Gao, X.; Xu, B.; Weiss, P. S.; Zhao, Y. Chemistry and Physics of a Single Atomic Layer: Strategies and Challenges for

Functionalization of Graphene and Graphene-Based Materials. *Chem. Soc. Rev.* **2012,** *41,* 97–114.

35. Fischbein, M. D.; Drndić, M. Electron Beam Nanosculpting of Suspended Graphene Sheets. *Appl. Phys. Lett.* **2008,** *93,* 113107.

36. Jin, Z.; Sun, W.; Yin, P.; Strano, M. S., Nanolithography Based on Metalized DNA Templates for Graphene Patterning. In *Current Protocols in Chemical Biology,* John Wiley & Sons, Inc.: Hoboken, NJ, 2009; 6, 53-64.

37. Bai, J.; Zhong, X.; Jiang, S.; Huang, Y.; Duan, X. Graphene Nanomesh. *Nat. Nanotechnol.* **2010,** *5,* 190-194.

38. Song, B.; Schneider, G. F.; Xu, Q.; Pandraud, G.; Dekker, C.; Zandbergen, H. Atomic-Scale Electron-Beam Sculpting of Near-Defect-Free Graphene Nanostructures. *Nano Lett.* **2011,** *11,* 2247–2250.

39. Liu, J.; Cai, H.; Yu, X.; Zhang, K.; Li, X.; Li, J.; Pan, N.; Shi, Q.; Luo, Y.; Wang, X. Fabrication of Graphene Nanomesh and Improved Chemical Enhancement for Raman Spectroscopy. *J. Phys. Chem. C* **2012,** *116,* 15741–15746.

40. Radich, J. G.; Kamat, P. V. Making Graphene Holey. Gold-Nanoparticle-Mediated Hydroxyl Radical Attack on Reduced Graphene Oxide. *ACS Nano* **2013,** *7,* 5546-5557.

41. Xu, Q.; Wu, M.-Y.; Schneider, G. F.; Houben, L.; Malladi, S. K.; Dekker, C.; Yucelen, E.; Dunin-Borkowski, R. E.; Zandbergen, H. W. Controllable Atomic Scale Patterning of Freestanding Monolayer Graphene at Elevated Temperature. *ACS Nano* **2013,** *7,* 1566-1572.

42. Han, X.; Funk, M. R.; Shen, F.; Chen, Y.-C.; Li, Y.; Campbell, C. J.; Dai, J.; Yang, X.; Kim, J.-W.; Liao, Y. Scalable Holey Graphene Synthesis and Dense Electrode Fabrication toward High-Performance Ultracapacitors. *ACS Nano* **2014,** *8,* 8255-8265.

43. Nam, S.; Choi, I.; Fu, C.-c.; Kim, K.; Hong, S.; Choi, Y.; Zettl, A.; Lee, L. P. Graphene Nanopore with a Self-Integrated Optical Antenna. *Nano Lett.* **2014,** *14,* 5584–5589.

44. Kuan, A. T.; Lu, B.; Xie, P.; Szalay, T.; Golovchenko, J. A. Electrical Pulse Fabrication of Graphene Nanopores in Electrolyte Solution. *Appl. Phys. Lett.* **2015,** *106,* 203109.

45. Han, P.; Mantooth, B. A.; Sykes, E. C. H.; Donhauser, Z. J.; Weiss, P. S. Benzene on Au{111} at 4 K: Monolayer Growth and Tip-Induced Molecular Cascades. *J. Am. Chem. Soc.* **2004,** *126,* 10787-10793.

46. Sykes, E. C. H.; Mantooth, B. A.; Han, P.; Donhauser, Z. J.; Weiss, P. S. Substrate-Mediated Intermolecular Interactions: A Quantitative Single Molecule Analysis. *J. Am. Chem. Soc.* **2005,** *127,* 7255-7260.

47. Mantooth, B. A.; Sykes, E. C. H.; Han, P.; Moore, A. M.; Donhauser, Z. J.; Crespi, V. H.; Weiss, P. S. Analyzing the Motion of Benzene on Au{111}: Single Molecule Statistics from Scanning Probe Images. *J. Phys. Chem. C* **2007,** *111,* 6167-6182.

48. Heimel, G.; Duhm, S.; Salzmann, I.; Gerlach, A.; Strozecka, A.; Niederhausen, J.; Bürker, C.; Hosokai, T.; Fernandez-Torrente, I.; Schulze, G.*, et al.* Charged and metallic molecular monolayers through surface-induced aromatic stabilization. *Nat. Chem.* **2013,** *5,* 187-194.

49. Bumm, L. A.; Arnold, J. J.; Charles, L. F.; Dunbar, T. D.; Allara, D. L.; Weiss, P. S. Directed Self-Assembly to Create Molecular Terraces with Molecularly Sharp Boundaries in Organic Monolayers. *J. Am. Chem. Soc.* **1999,** *121,* 8017-8021.

50. Hohman, J. N.; Zhang, P.; Morin, E. I.; Han, P.; Kim, M.; Kurland, A. R.; McClanahan, P. D.; Balema, V. P.; Weiss, P. S. Self-Assembly of Carboranethiol Isomers on Au{111}: Intermolecular Interactions Determined by Molecular Dipole Orientations. *ACS Nano* **2009,** *3,* 527–536.

51. Han, P.; Kurland, A. R.; Giordano, A. N.; Nanayakkara, S. U.; Blake, M. M.; Pochas, C. M.; Weiss, P. S. Heads and Tails: Simultaneous Exposed and Buried Interface Imaging of Monolayers. *ACS Nano* **2009,** *3,* 3115-3121.

52. Ishigami, M.; Chen, J. H.; Cullen, W. G.; Fuhrer, M. S.; Williams, E. D. Atomic Structure of Graphene on SiO_2. *Nano Lett.* **2007,** *7,* 1643–1648.

53. Liao, L.; Bai, J.; Cheng, R.; Lin, Y.-C.; Jiang, S.; Huang, Y.; Duan, X. Top-Gated Graphene Nanoribbon Transistors with Ultrathin High-k Dielectrics. *Nano Lett.* **2010,** *10,* 1917–1921.

54. Liao, L.; Bai, J.; Qu, Y.; Huang, Y.; Duan, X. Single-Layer Graphene on Al_2O_3/Si Substrate: Better Contrast and Higher Performance of Graphene Transistors. *Nanotechnology* **2010,** *21,* 015705.

55. Liang, X.; Sperling, B. A.; Calizo, I.; Cheng, G.; Hacker, C. A.; Zhang, Q.; Obeng, Y.; Yan, K.; Peng, H.; Li, Q. Toward Clean and Crackless Transfer of Graphene. *ACS Nano* **2011,** *5,* 9144-9153.

56. Monnell, J. D.; Stapleton, J. J.; Dirk, S. M.; Reinerth, W. A.; Tour, J. M.; Allara, D. L.; Weiss, P. S. Relative Conductances of Alkaneselenolate and Alkanethiolate Monolayers on Au{111}. *J. Phys. Chem. B* **2005,** *109,* 20343–20349.

57. Claridge, S. A.; Schwartz, J. J.; Weiss, P. S. Electrons, Photons, and Force: Quantitative Single-Molecule Measurements from Physics to Biology. *ACS Nano* **2011,** *5,* 693-729.

58. Khomyakov, P. A.; Giovannetti, G.; Rusu, P. C.; Brocks, G.; van den Brink, J.; Kelly, P. J. First-Principles Study of the Interaction and Charge Transfer between Graphene and Metals. *Phys. Rev. B* **2009,** *79,* 195425.

59. Wofford, J. M.; Starodub, E.; Walter, A. L.; Nie, S.; Bostwick, A.; Bartelt, N. C.; Thürmer, K.; Rotenberg, E.; McCarty, K. F.; Dubon, O. D. Extraordinary Epitaxial Alignment of Graphene Islands on Au(111). *New J. Phys.* **2012,** *14,* 053008.

60. Donhauser, Z. J.; Price, D. W.; Tour, J. M.; Weiss, P. S. Control of Alkanethiolate Monolayer Structure Using Vapor-Phase Annealing. *J. Am. Chem. Soc.* **2003,** *125,* 11462-11463.

61. Dameron, A. A.; Charles, L. F.; Weiss, P. S. Structures and Displacement of 1-Adamantanethiol Self-Assembled Monolayers on Au{111}. *J. Am. Chem. Soc.* **2005,** *127,* 8697-8704.

62. Mullen, T. J.; Dameron, A. A.; Saavedra, H. M.; Williams, M. E.; Weiss, P. S. Dynamics of Solution Displacement in 1-Adamantanethiolate Self-Assembled Monolayers. *J. Phys. Chem. C* **2007,** *111,* 6740–6746.

63. Dameron, A. A.; Mullen, T. J.; Hengstebeck, R. W.; Saavedra, H. M.; Weiss, P. S. Origins of Displacement in 1-Adamantanethiolate Self-Assembled Monolayers. *J. Phys. Chem. C* **2007,** *111,* 6747–6752.

64. Saavedra, H. M.; Barbu, C. M.; Dameron, A. A.; Mullen, T. J.; Crespi, V. H.; Weiss, P. S. 1-Adamantanethiolate Monolayer Displacement Kinetics Follow a Universal Form. *J. Am. Chem. Soc.* **2007,** *129,* 10741–10746.

65. Fujii, S.; Akiba, U.; Fujihira, M. Geometry for Self-Assembling of Spherical Hydrocarbon Cages with Methane Thiolates on Au(111). *J. Am. Chem. Soc.* **2002,** *124,* 13629-13635.

66. Kondoh, H.; Kodama, C.; Nozoye, H. Structure-Dependent Change of Desorption Species from *n*-Alkanethiol Monolayers Adsorbed on Au(111): Desorption of Thiolate Radicals from Low-Density Structures. *J. Phys. Chem. B* **1998,** *102,* 2310–2312.

67. Kondoh, H.; Kodama, C.; Sumida, H.; Nozoye, H. Molecular Processes of Adsorption and Desorption of Alkanethiol Monolayers on Au(111). *J. Chem. Phys.* **1999,** *111,* 1175–1184.

68. Gao, L.; Ren, W.; Xu, H.; Jin, L.; Wang, Z.; Ma, T.; Ma, L.-P.; Zhang, Z.; Fu, Q.; Peng, L.-M. Repeated Growth and Bubbling Transfer of Fraphene with Millimetre-Size Single-Crystal Grains Using Platinum. *Nat. Commun.* **2012,** *3,* 699.

69. Suk, J. W.; Kitt, A.; Magnuson, C. W.; Hao, Y.; Ahmed, S.; An, J.; Swan, A. K.; Goldberg, B. B.; Ruoff, R. S. Transfer of CVD-Grown Monolayer Graphene onto Arbitrary Substrates. *ACS Nano* **2011,** *5,* 6916-6924.

70. Kang, J.; Hwang, S.; Kim, J. H.; Kim, M. H.; Ryu, J.; Seo, S. J.; Hong, B. H.; Kim, M. K.; Choi, J.-B. Efficient Transfer of Large-Area Graphene Films onto Rigid Substrates by Hot Pressing. *ACS Nano* **2012,** *6,* 5360-5365.

71. Chen, X.-D.; Liu, Z.-B.; Zheng, C.-Y.; Xing, F.; Yan, X.-Q.; Chen, Y.; Tian, J.-G. High-Quality and Efficient Transfer of Large-Area Graphene Films onto Different Substrates. *Carbon* **2013,** *56,* 271-278.

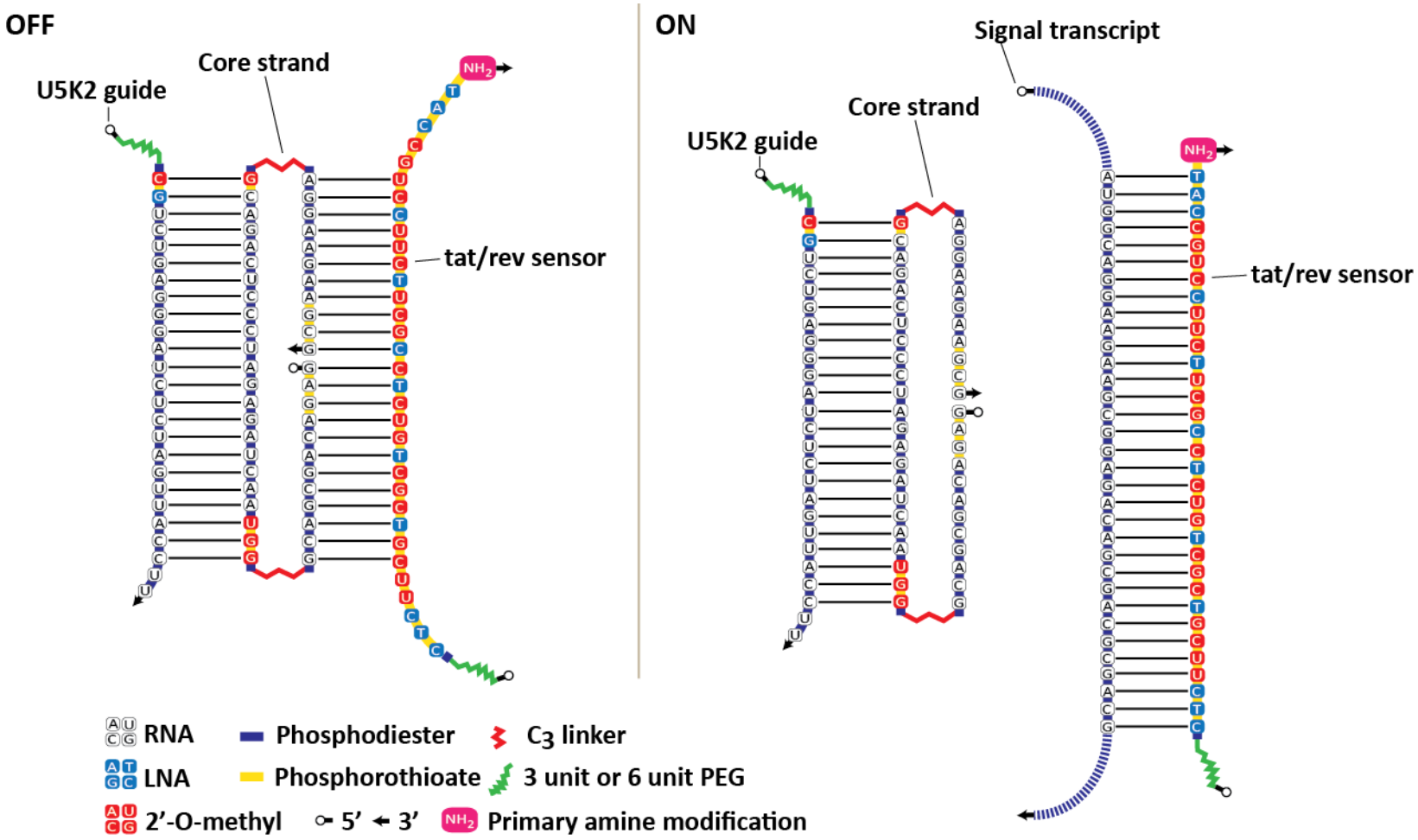
72. Müllen, K. Evolution of Graphene Molecules: Structural and Functional Complexity as Driving Forces behind Nanoscience. *ACS Nano* **2014,** *8,* 6531-6541.

73. Han, P.; Akagi, K.; Canova, F. F.; Mutoh, H.; Shiraki, S.; Iwaya, K.; Weiss, P. S.; Asao, N.; Hitosugi, T. Bottom-Up Graphene-Nanoribbon Fabrication Reveals Chiral Edges and Enantioselectivity. *ACS Nano* **2014,** *8,* 9181-9187.

74. Narita, A.; Verzhbitskiy, I. A.; Frederickx, W.; Mali, K. S.; Jensen, S. A.; Hansen, M. R.; Bonn, M.; De Feyter, S.; Casiraghi, C.; Feng, X. Bottom-Up Synthesis of Liquid-Phase-Processable Graphene Nanoribbons with Near-Infrared Absorption. *ACS Nano* **2014,** *8,* 11622-11630.
75. Zhang, R.; Lyu, G.; Chen, C.; Lin, T.; Liu, J.; Liu, P. N.; Lin, N. Two-Dimensional Superlattices of Bi Nanoclusters Formed on a Au(111) Surface Using Porous Supramolecular Templates. *ACS Nano* **2015,** *9,* 8547-8553.
76. Leong, W. S.; Nai, C. T.; Thong, J. T. L. What Does Annealing Do to Metal–Graphene Contacts? *Nano Lett.* **2014,** *14,* 3840–3847.

Appendix A

Supplementary Material for A Molecular Dynamics Investigation of *Cond*-siRNA.


OFF
U5K2 guide
Core strand
tat/rev sensor
ON
Signal transcript
Core strand
U5K2 guide
tat/rev sensor
RNA
LNA
2'-O-methyl
Phosphodiester
Phosphorothioate
5' 3'
C3 linker
3 unit or 6 unit PEG
Primary amine modification

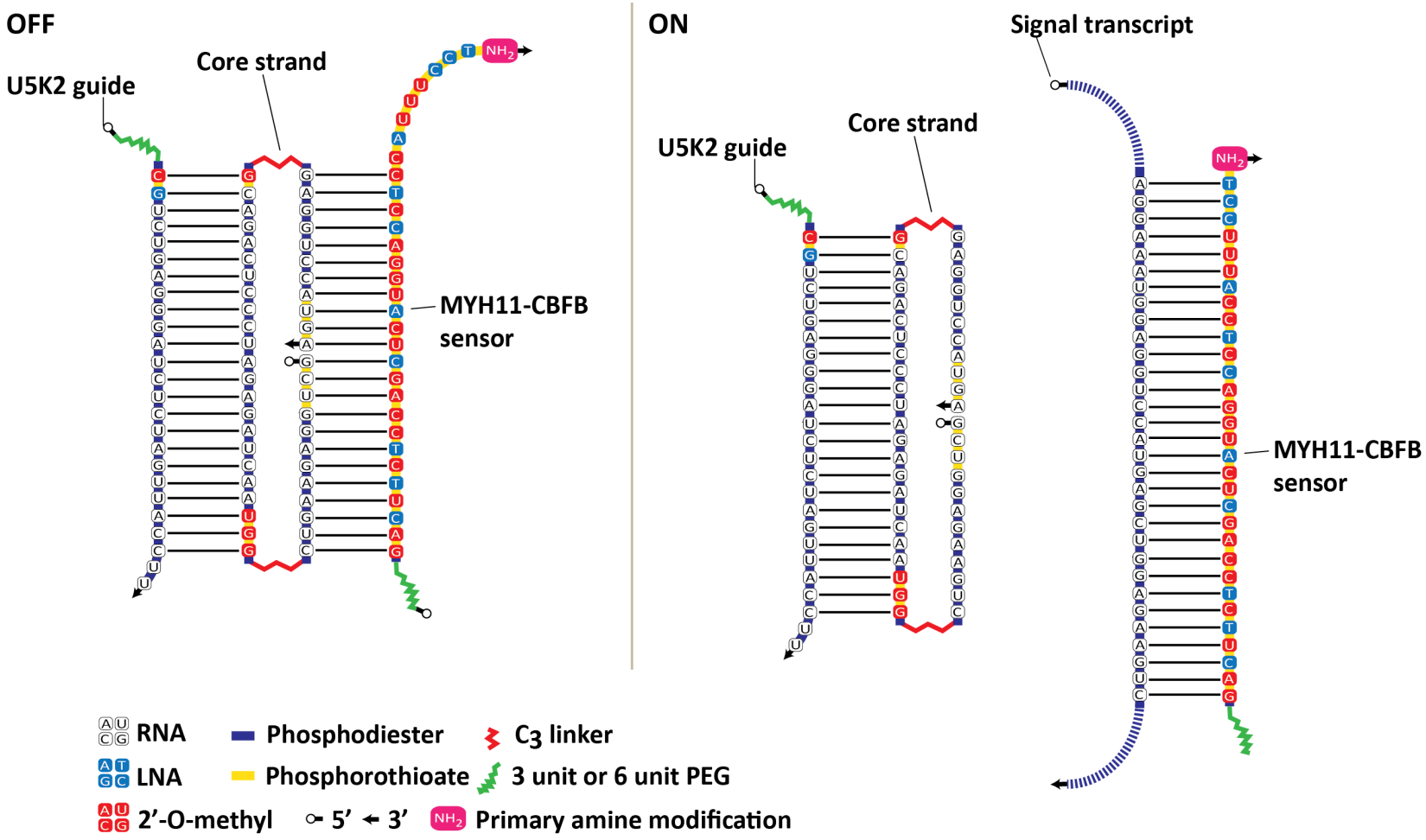


Figure A1-2 Sequence of Cond-RNA Construct 2. Reproduced with permission from Han *et al.*

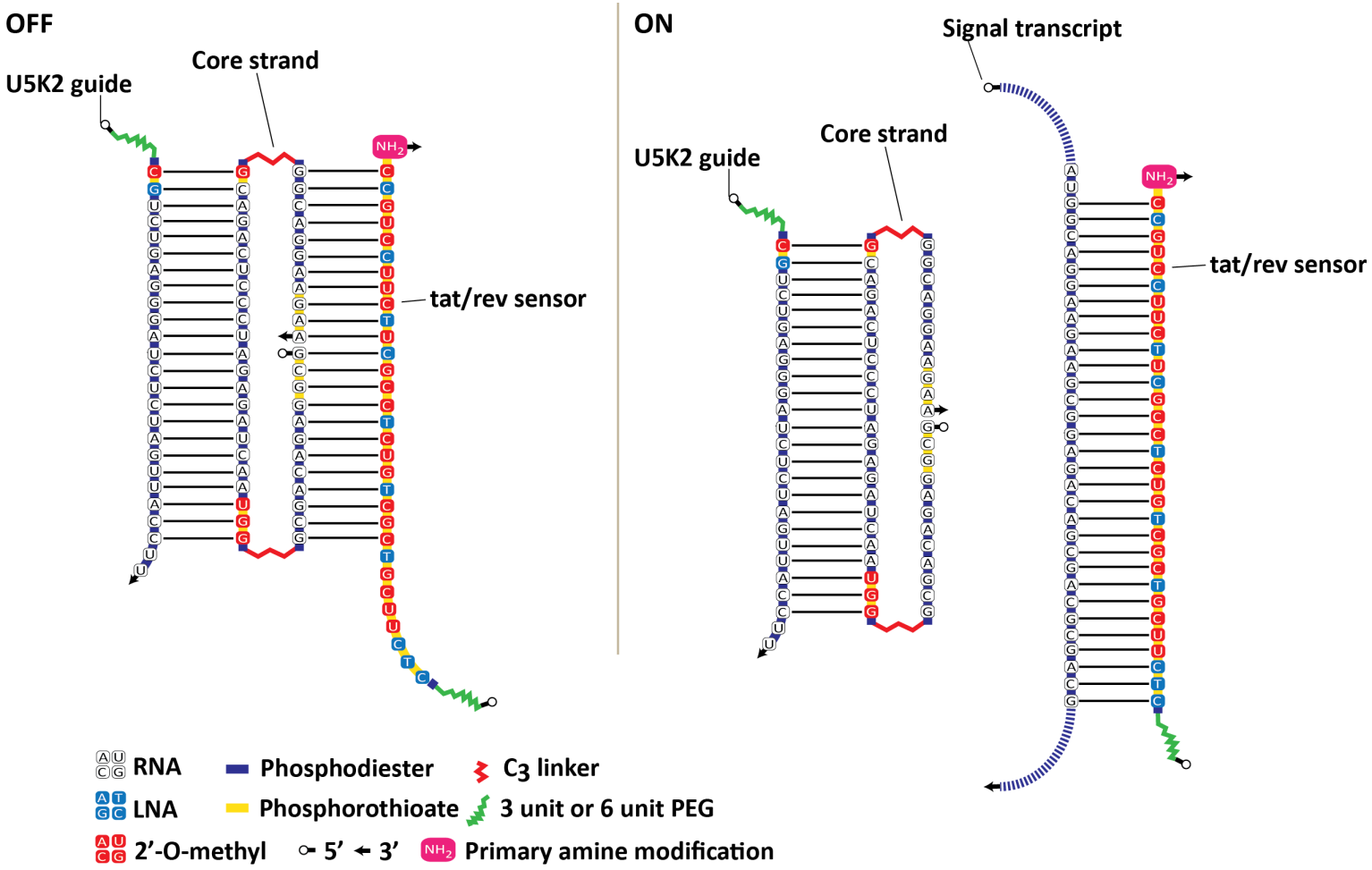


Figure A1-3 Sequence of Cond-RNA Construct 3. Reproduced with permission from Han *et al.*

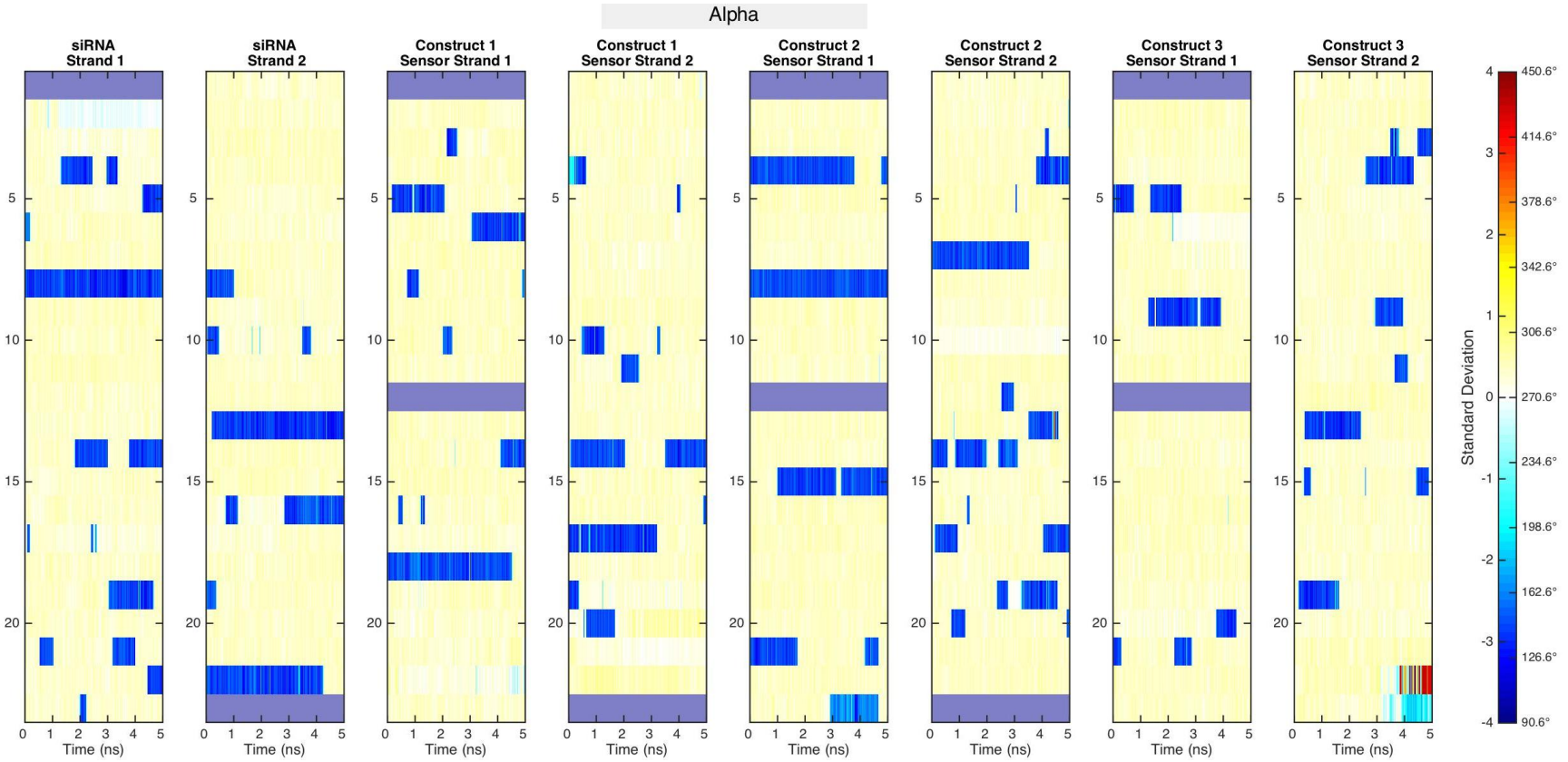


Figure A1-4 Alpha torsions of isolated natural duplexes.

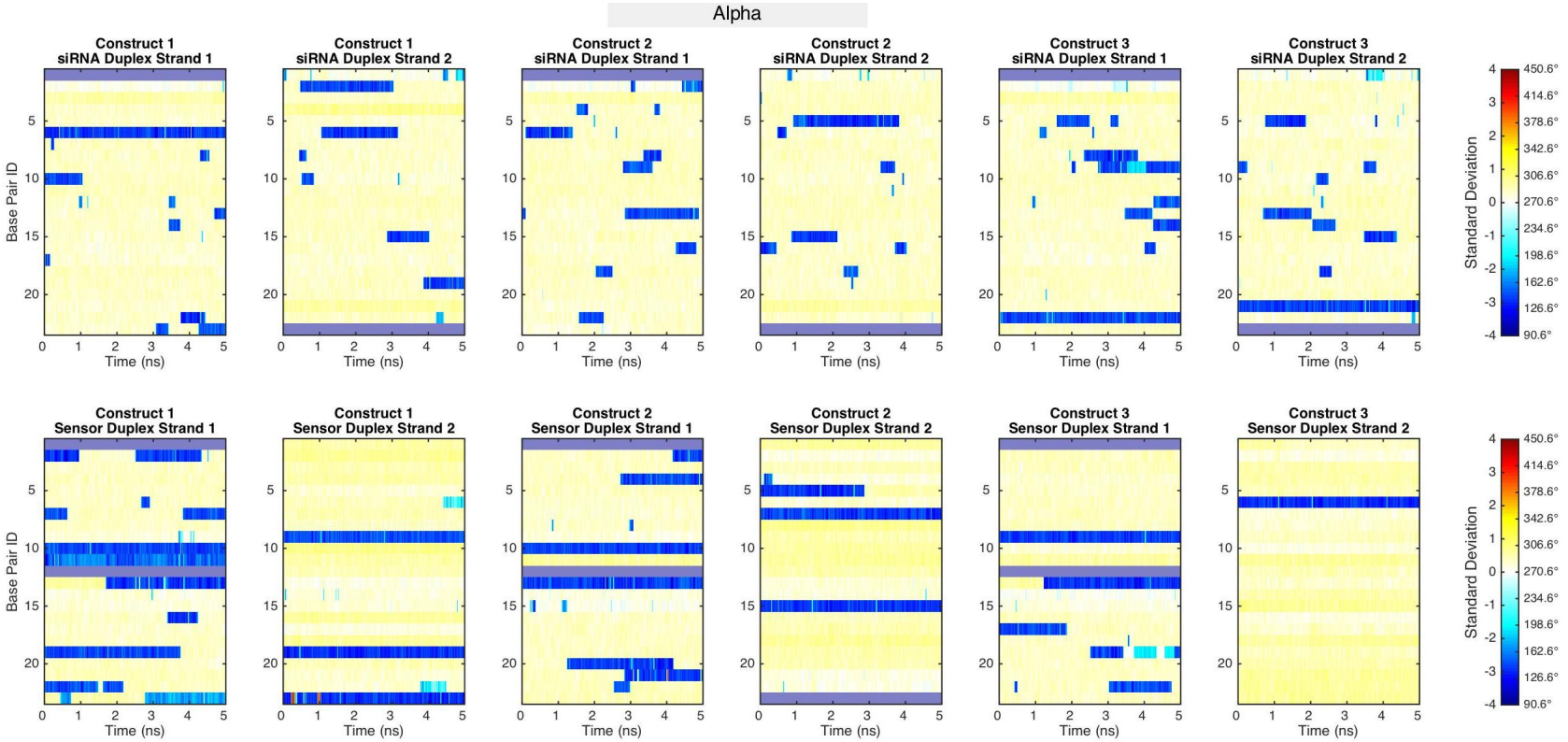


Figure A1-5 Alpha torsions of *Cond*-siRNAs.

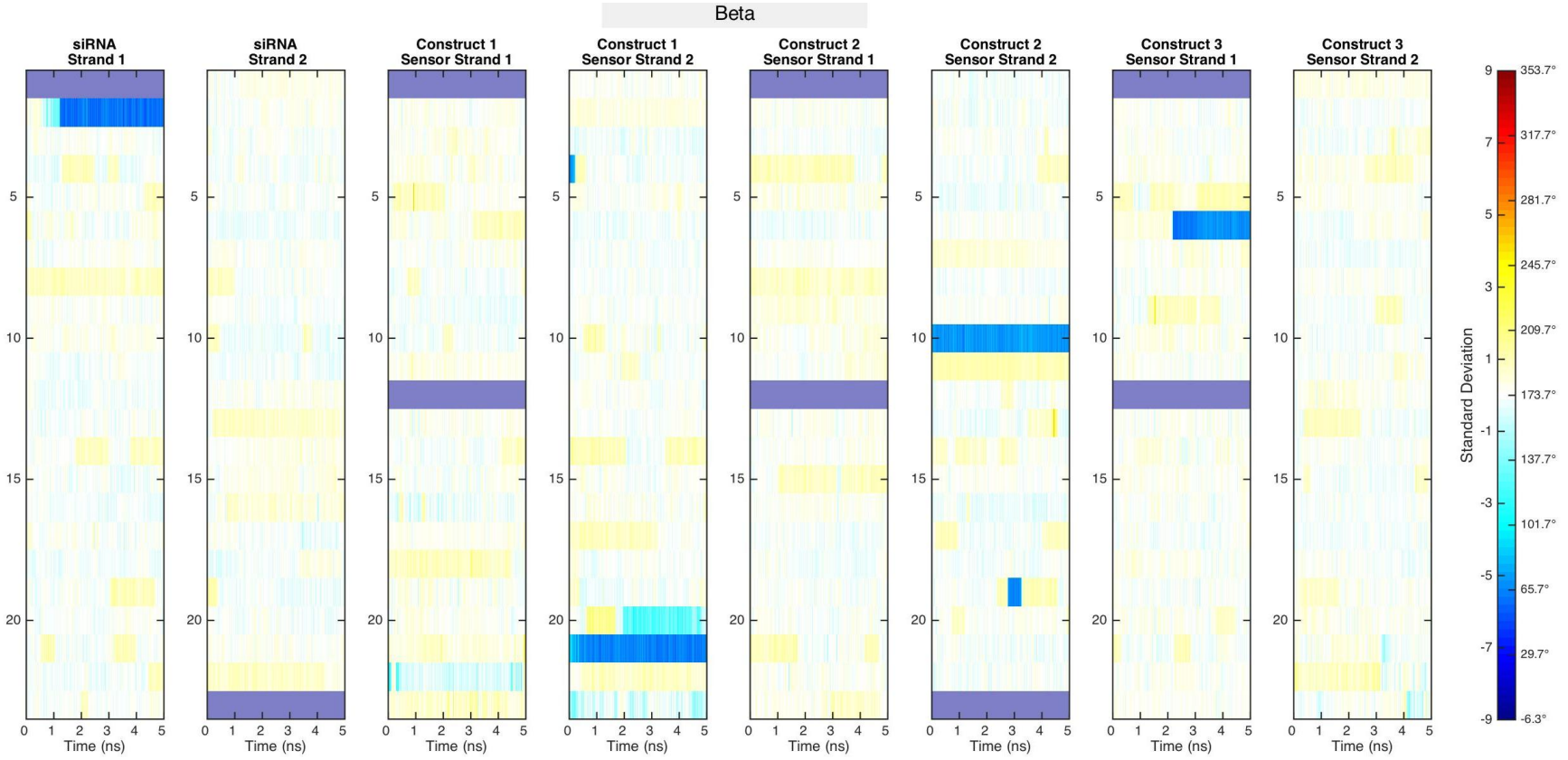


Figure A1-6 Beta torsions of isolated natural duplexes.

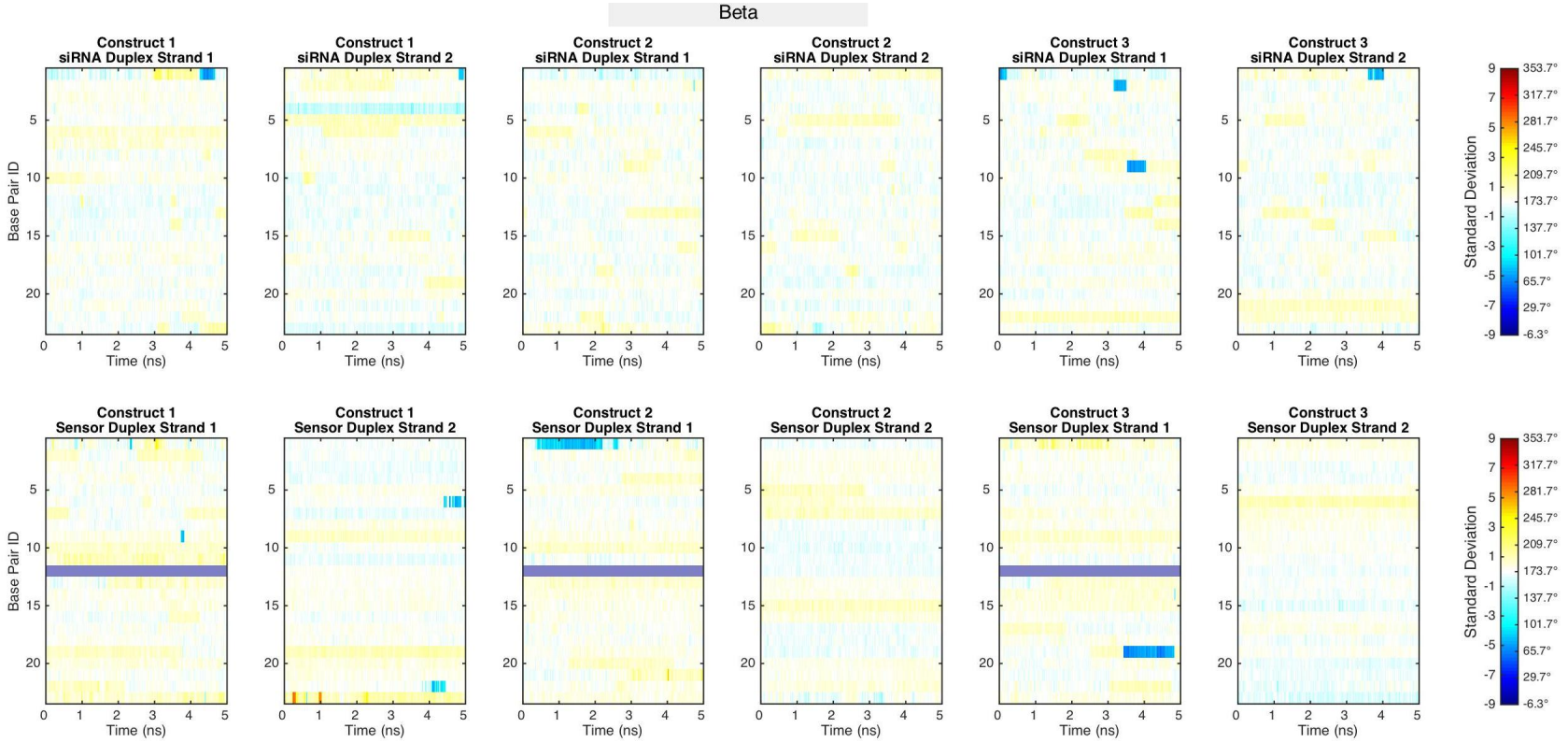


Figure A1-7 Beta torsions of *Cond*-siRNAs.

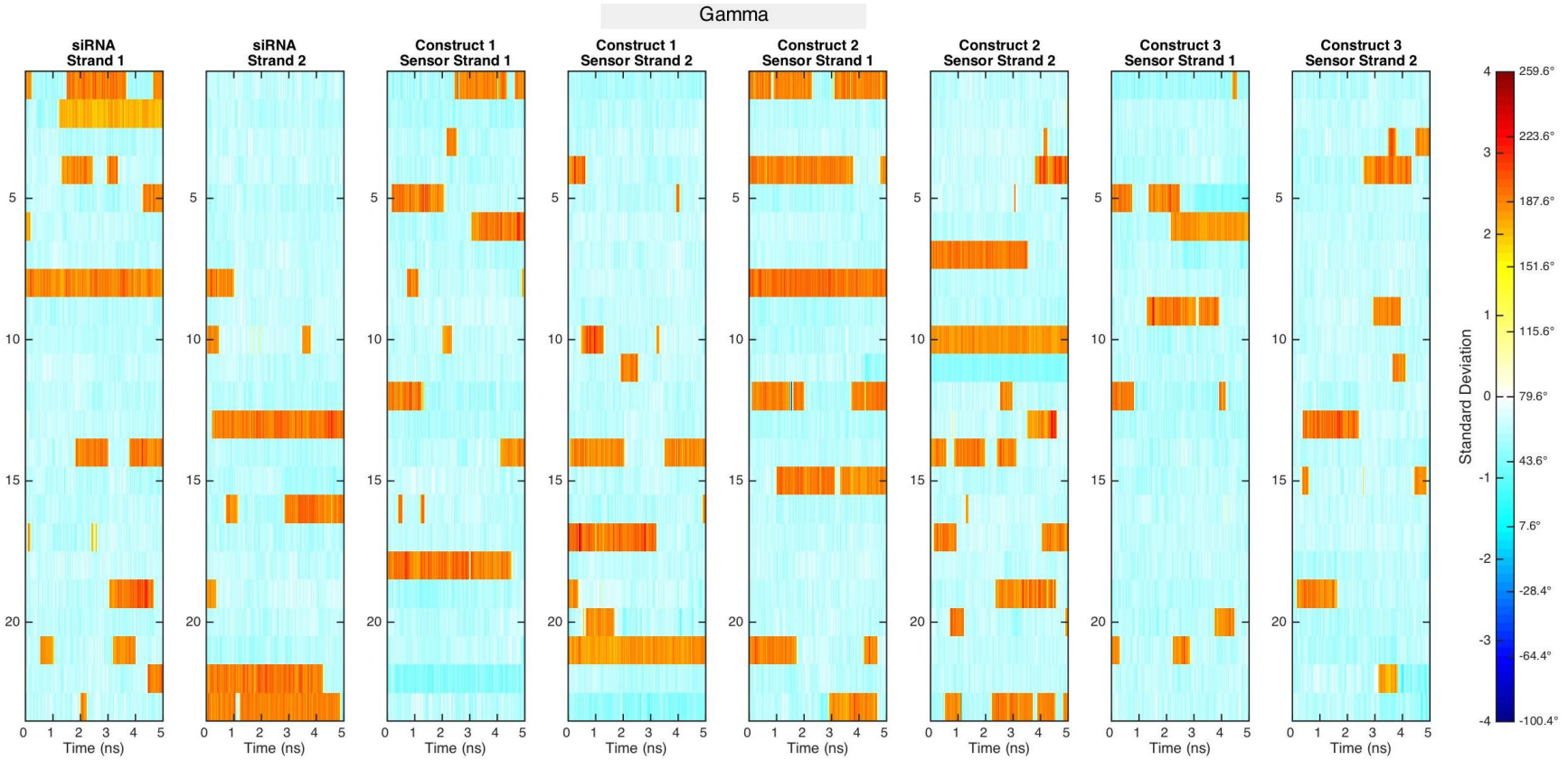


Figure A1-8 Gamma torsions of isolated natural duplexes.

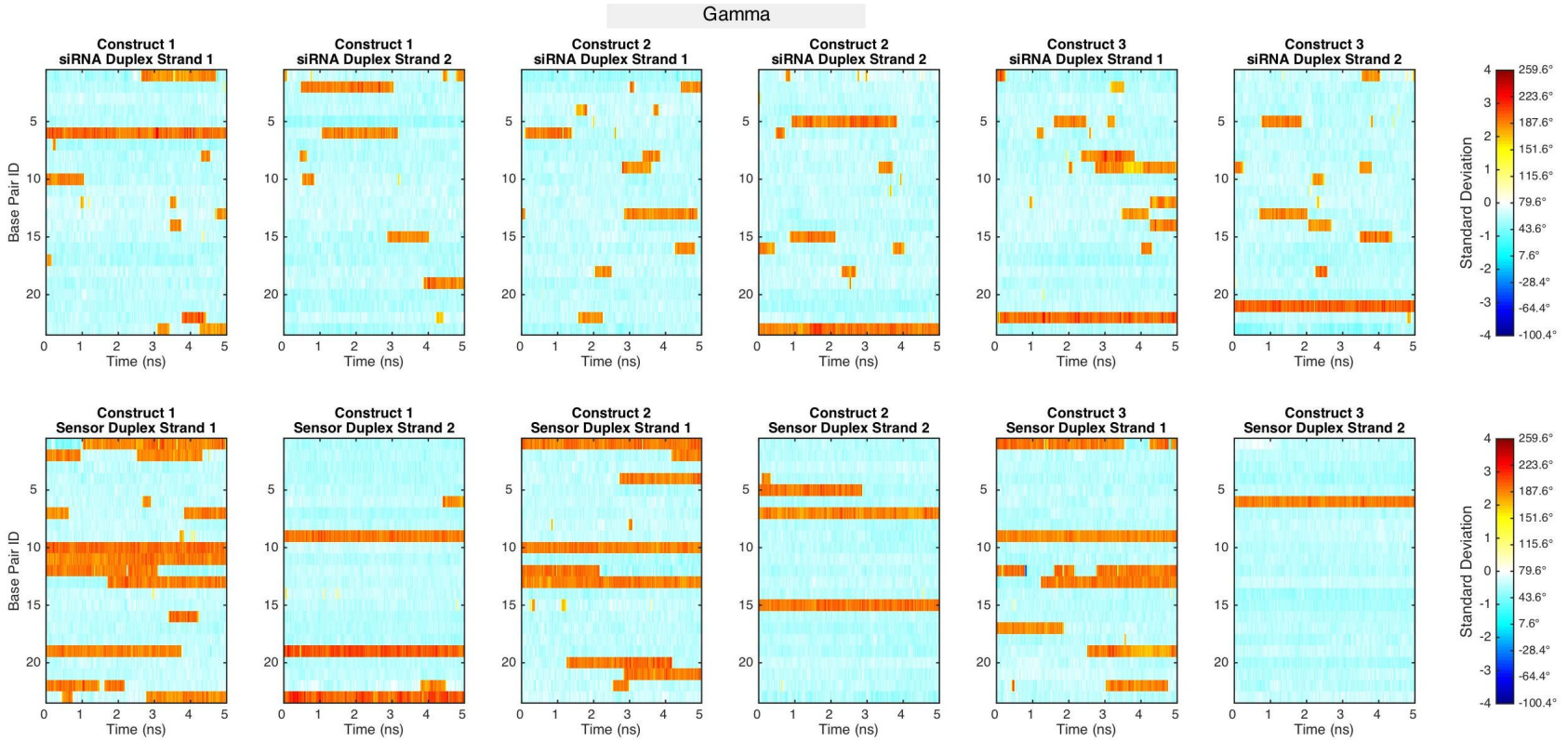


Figure A1-9 Gamma torsions of *Cond*-siRNAs.

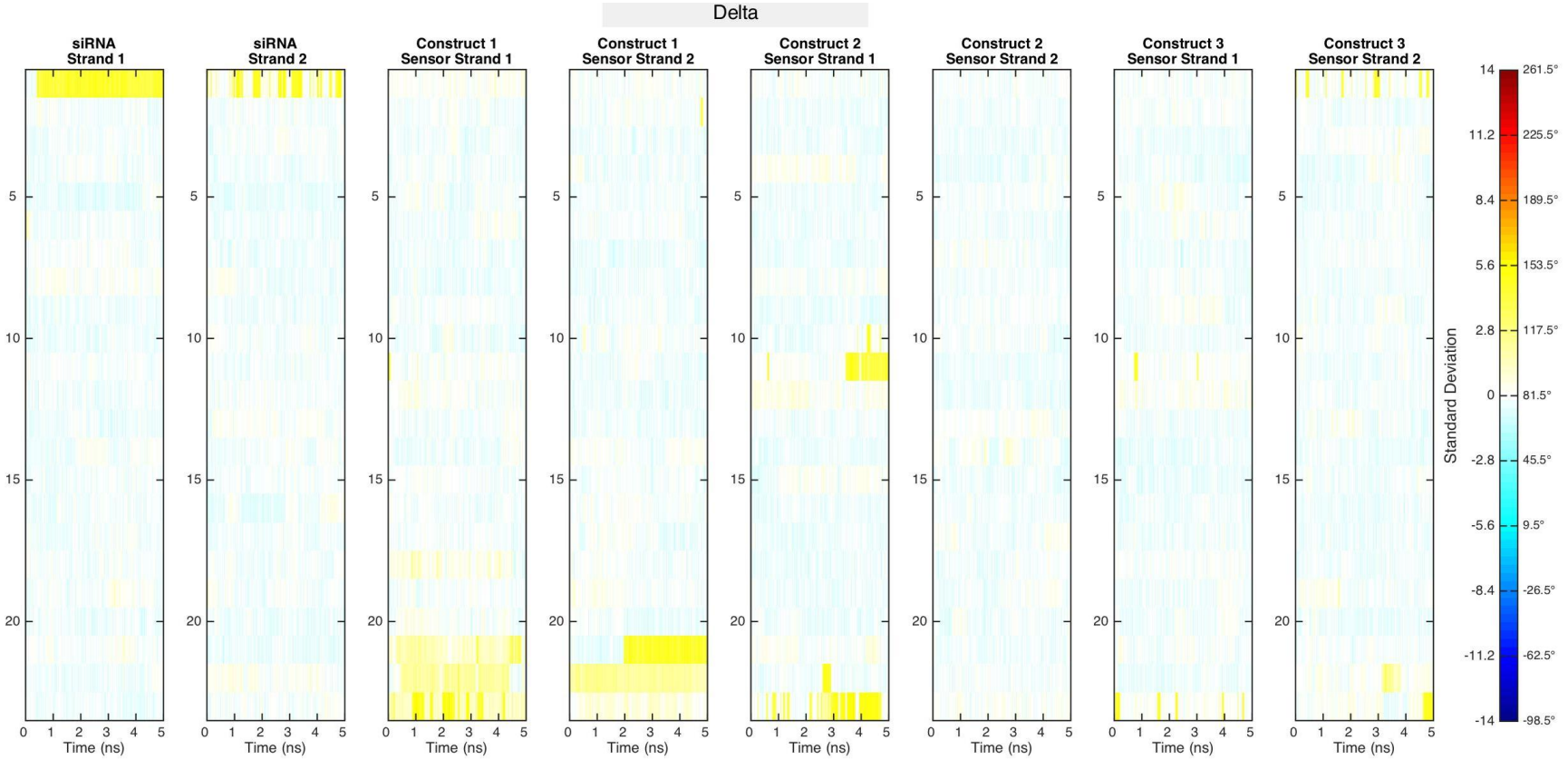


Figure A1-10 Delta torsions of isolated natural duplexes.

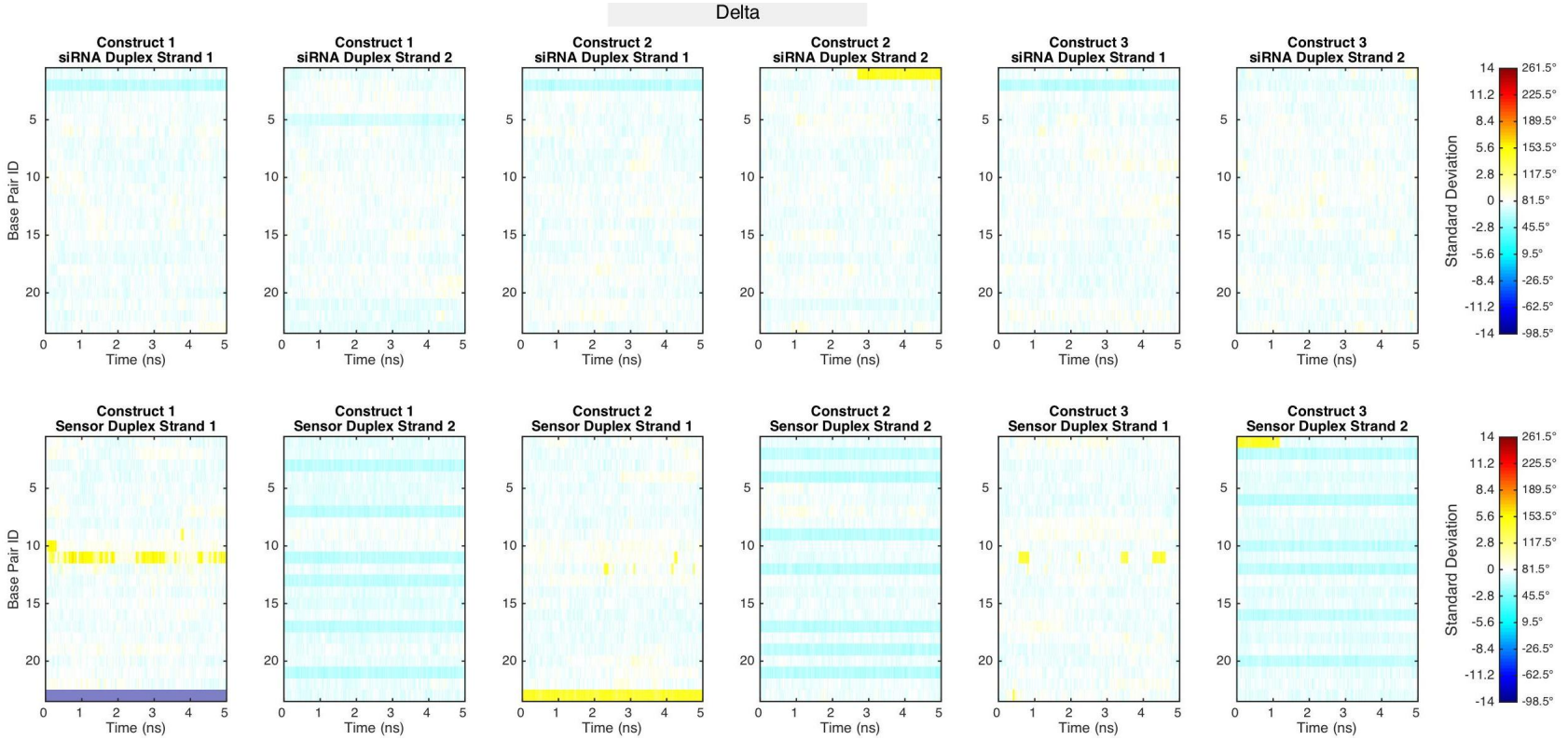


Figure A1-11 Delta torsions of *Cond*-siRNAs.

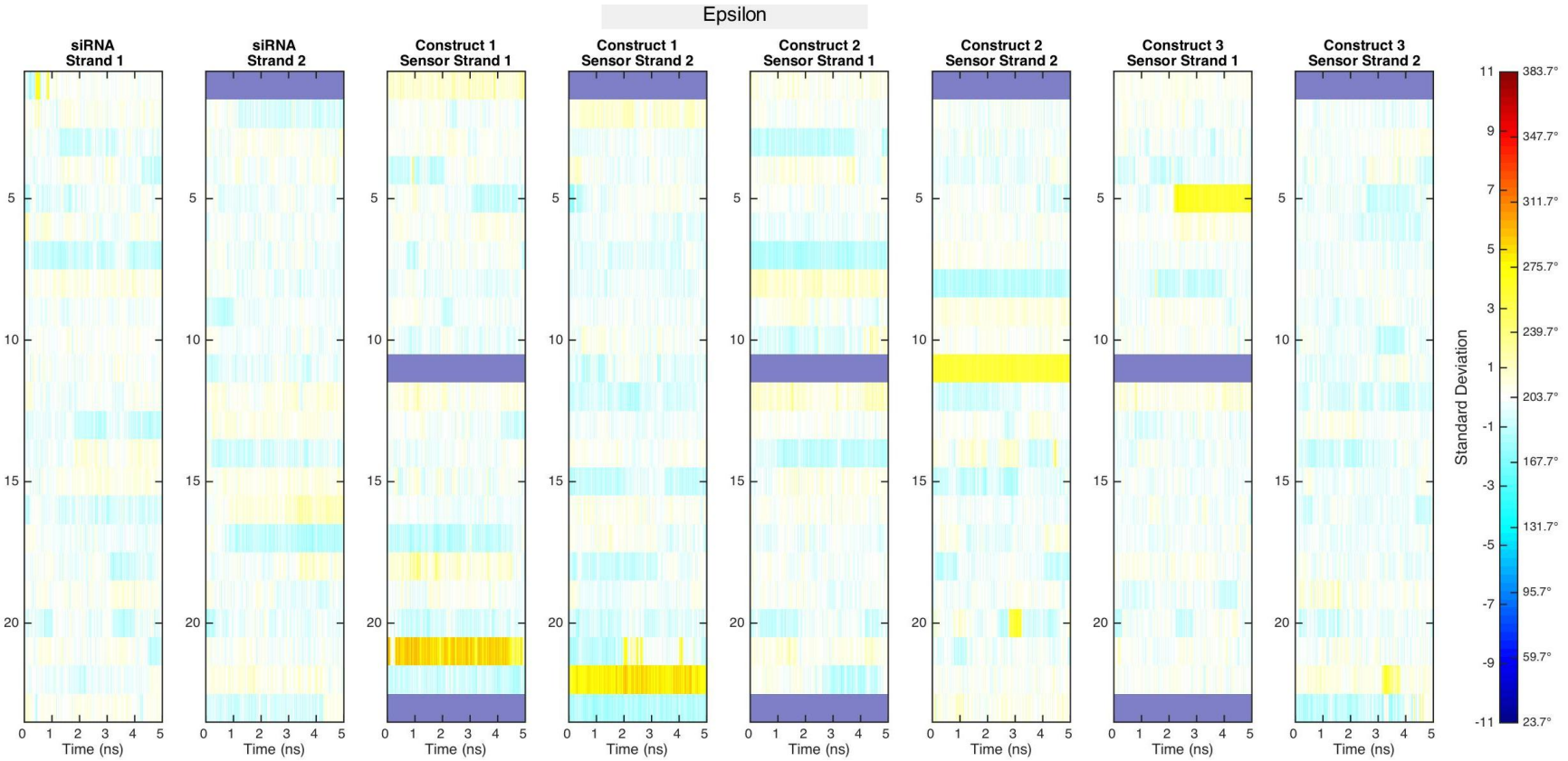


Figure A1-12 Epsilon torsions of isolated natural duplexes.

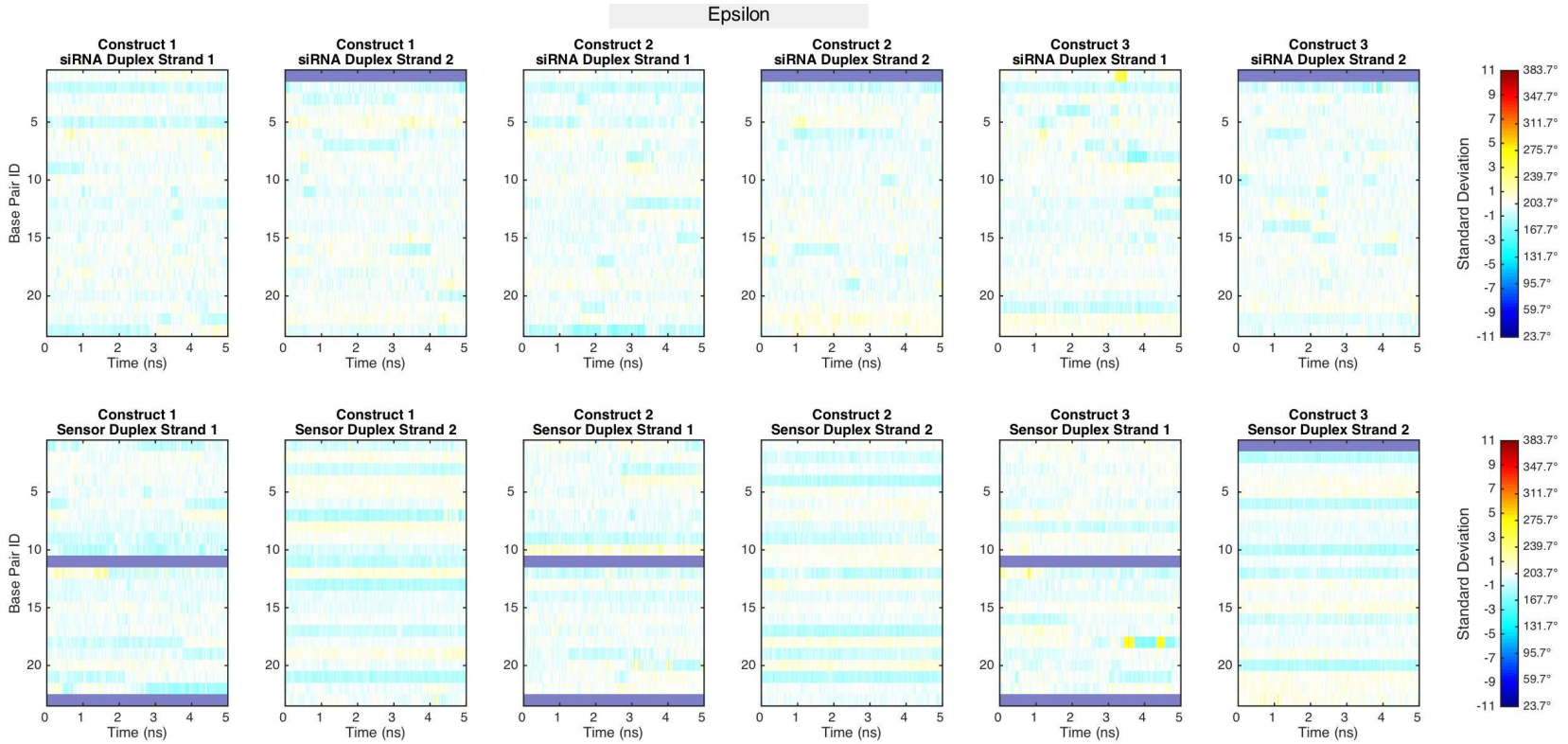


Figure A1-13 Epsilon torsions of *Cond*-siRNAs.

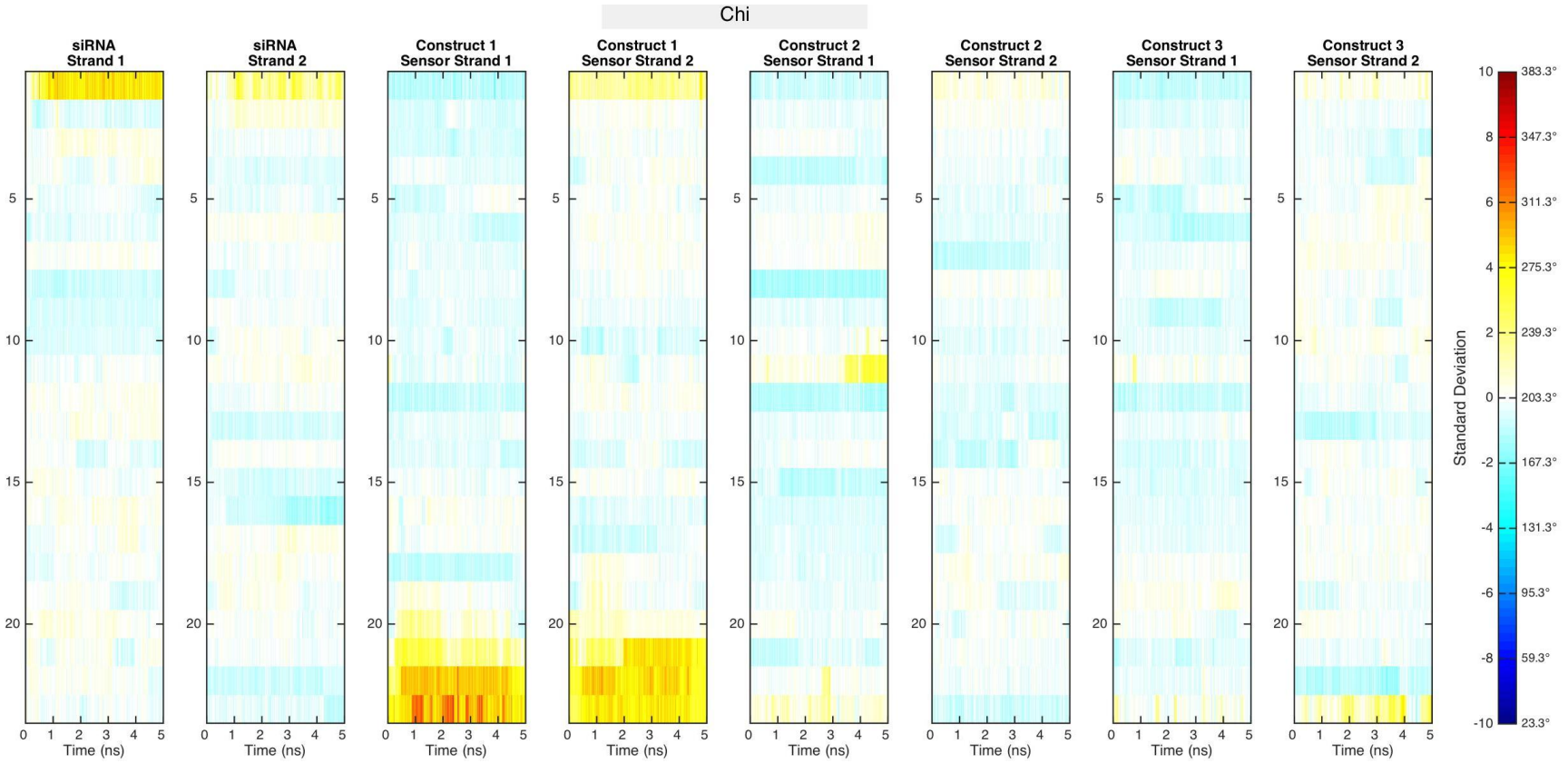


Figure A1-14 Chi torsions of isolated natural duplexes.

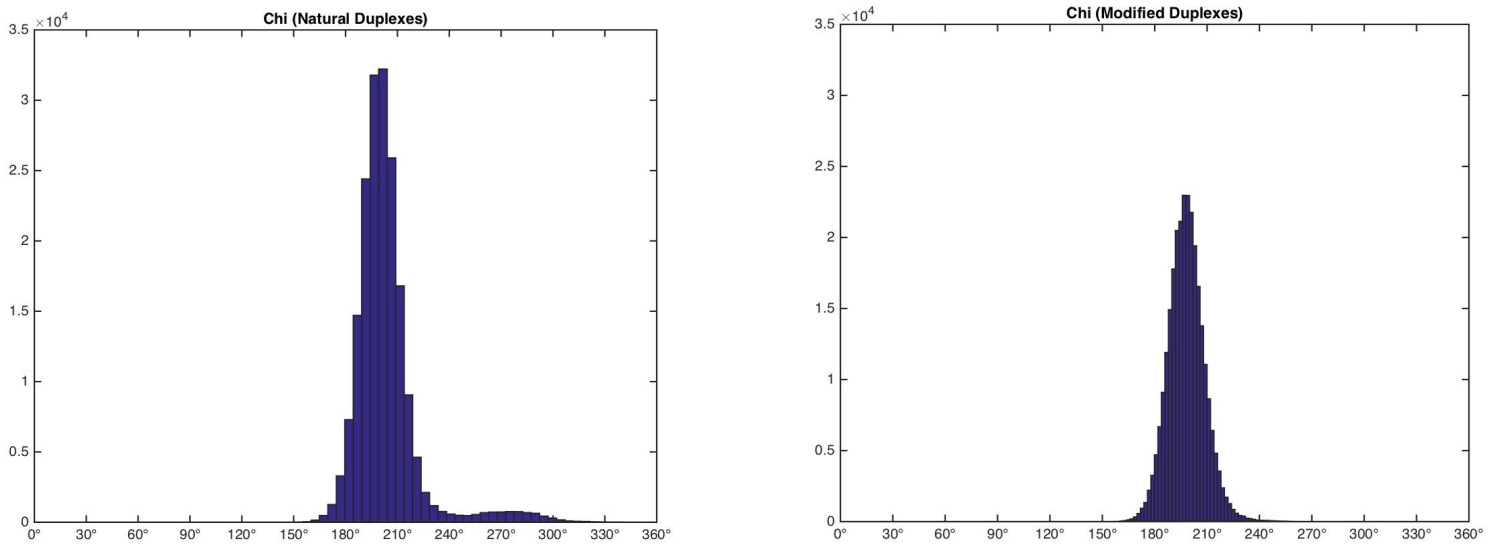


Figure A1-15 Histogram of chi torsions of isolated natural duplexes (left) and *Cond*-siRNA (right).

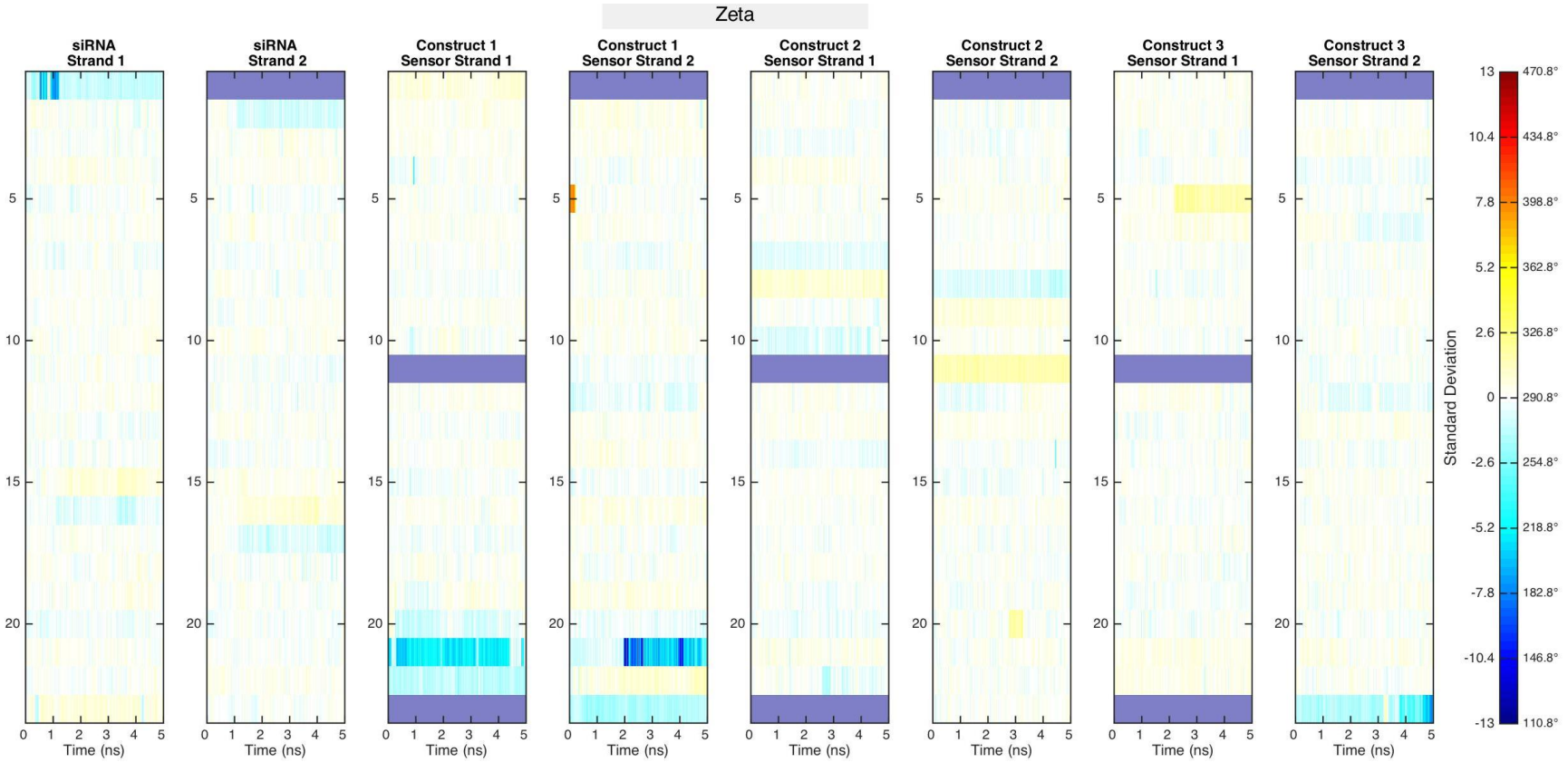


Figure A1-16 Zeta torsions of isolated natural duplexes.

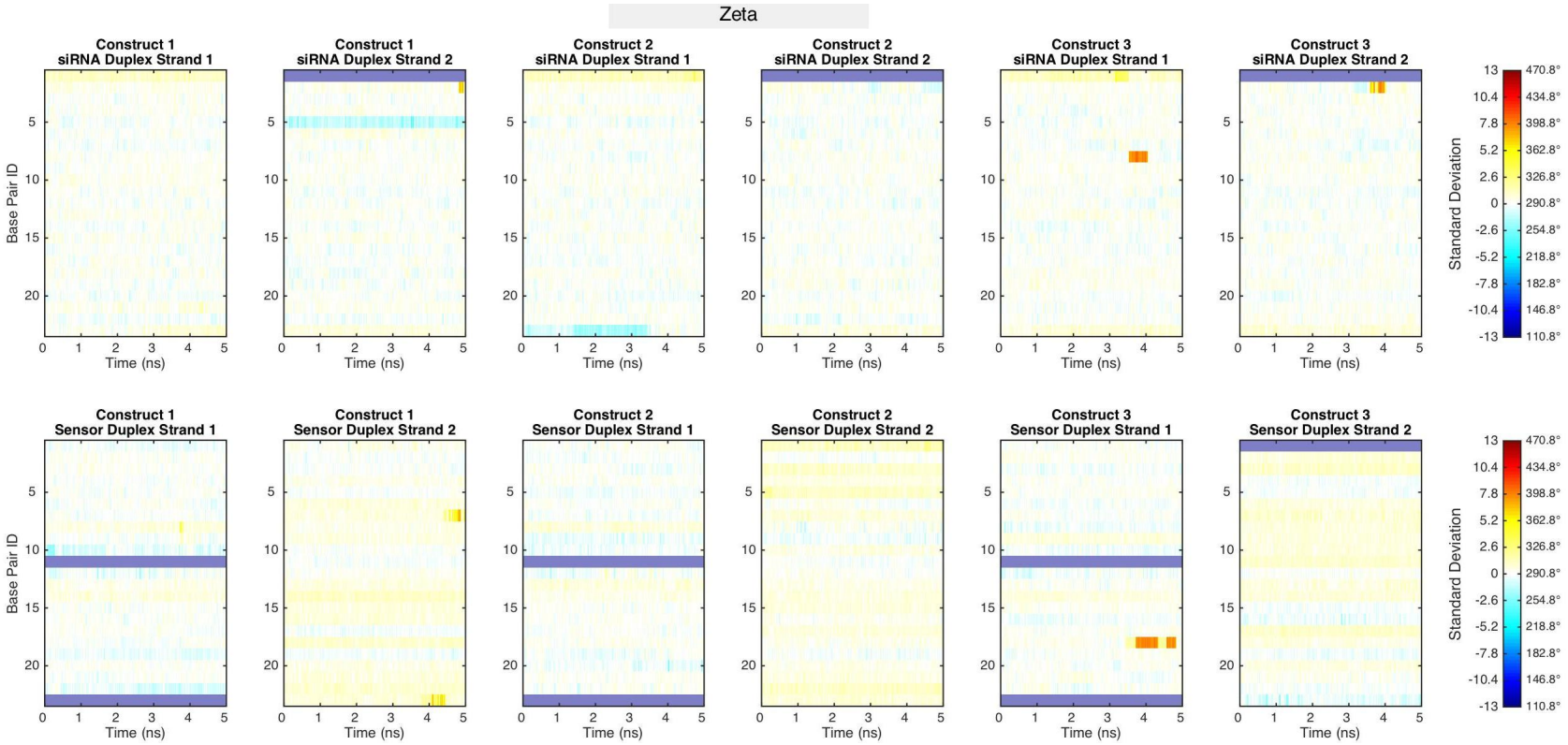


Figure A1-17 Zeta torsions of *Cond*-siRNAs.

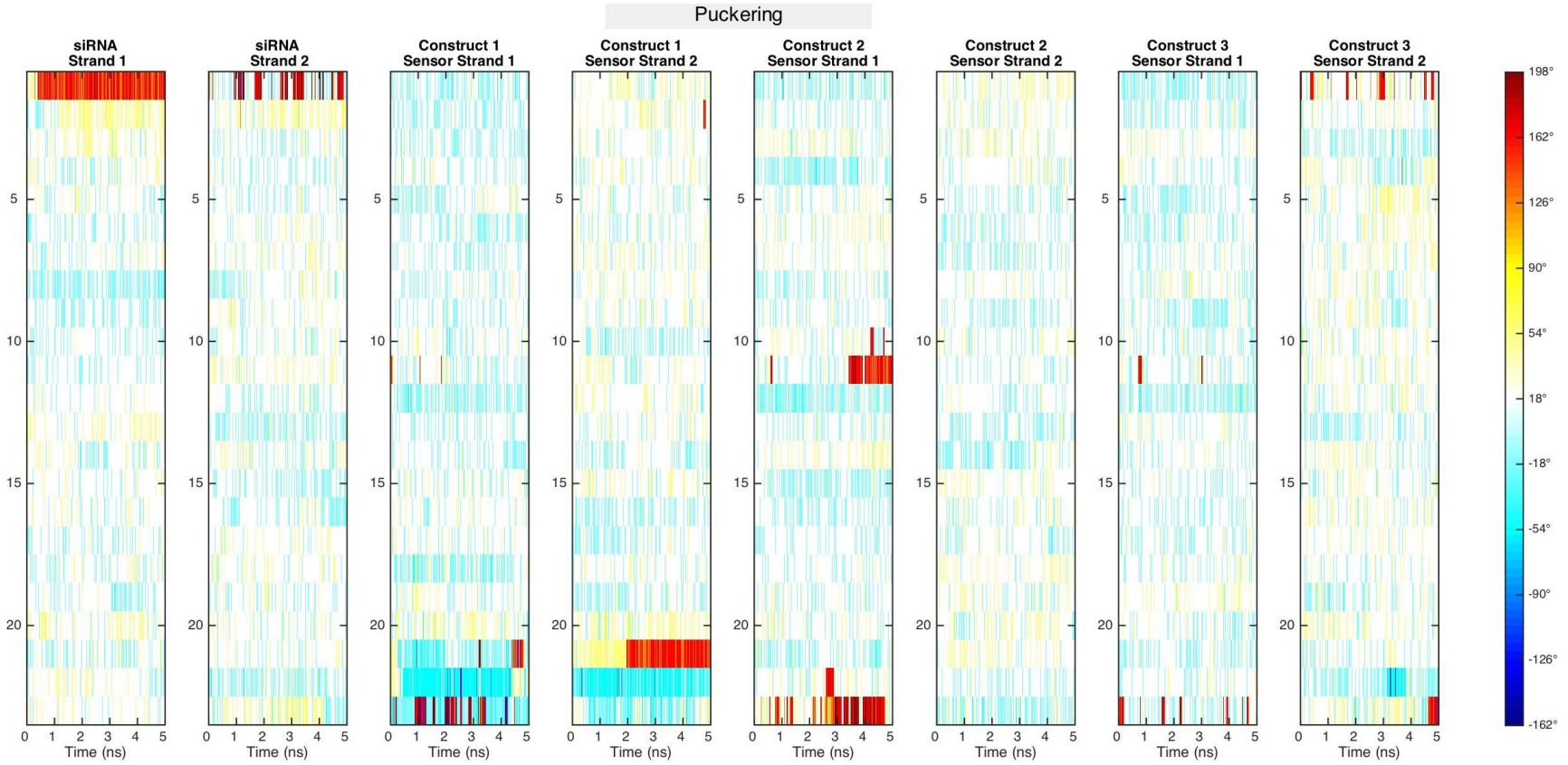


Figure A1-18 Puckering of isolated natural duplexes.

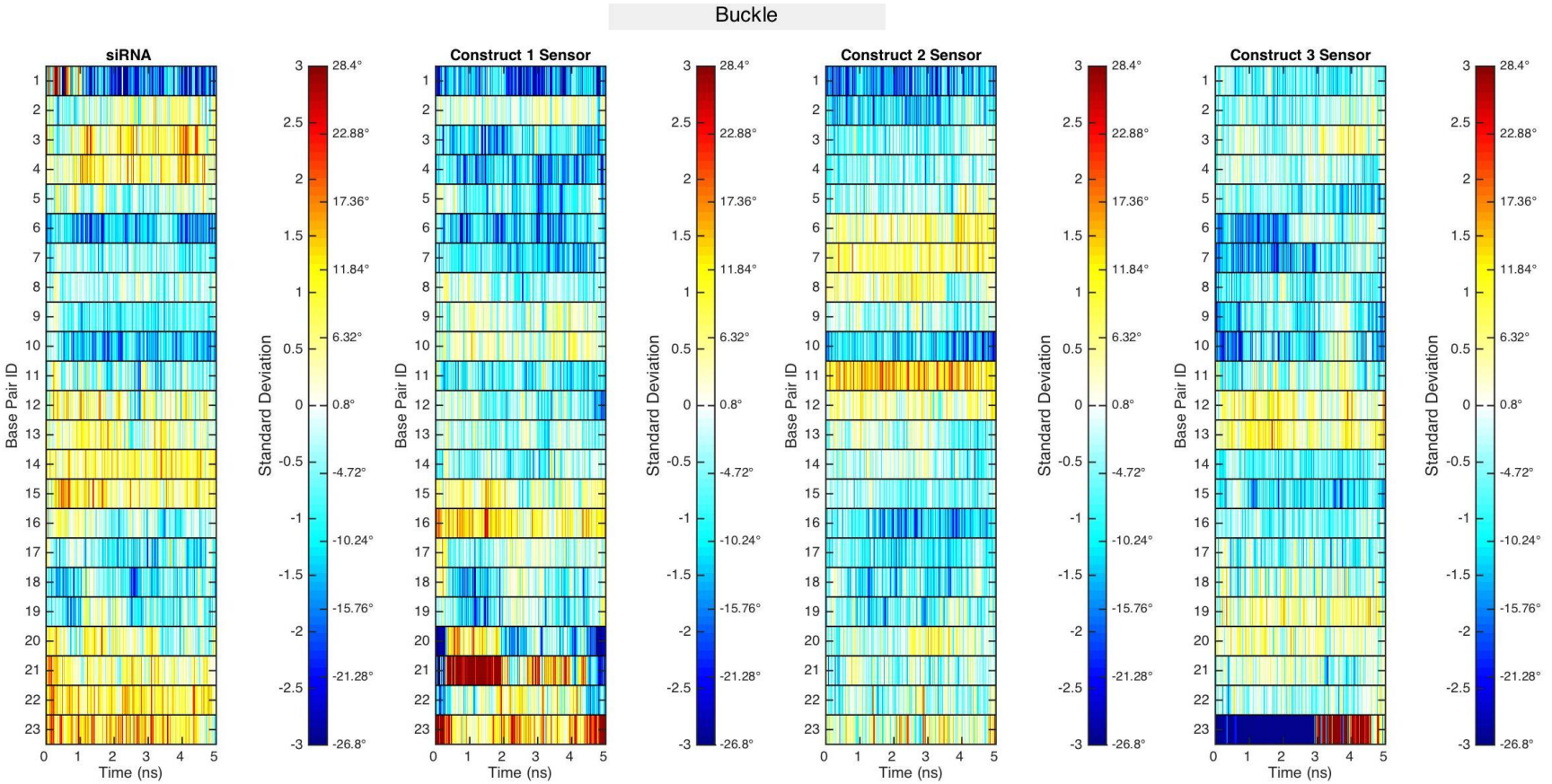


Figure A1-19 Buckle of isolated natural duplexes.

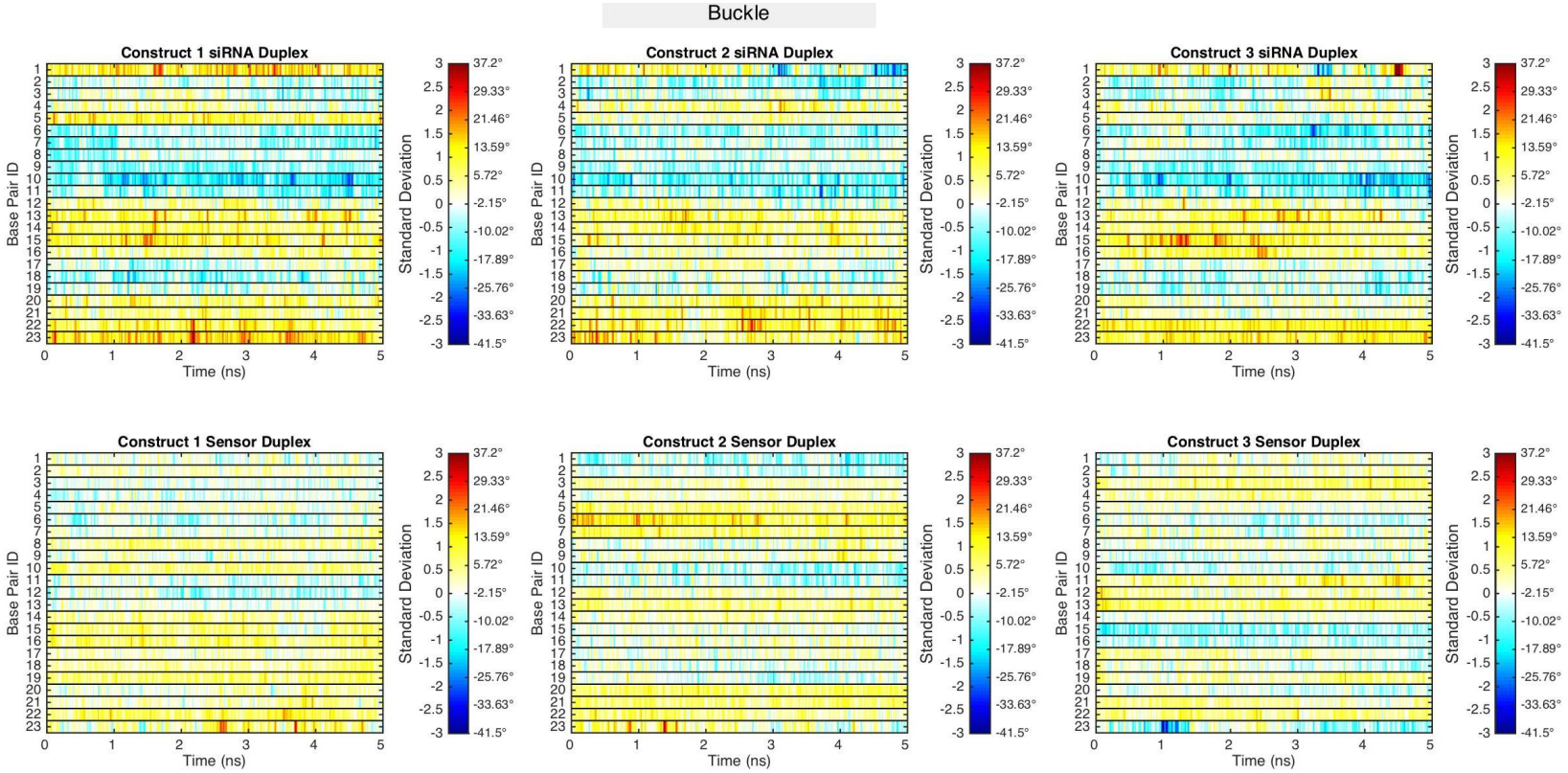


Figure A1-20 Buckle of *Cond*-siRNAs.

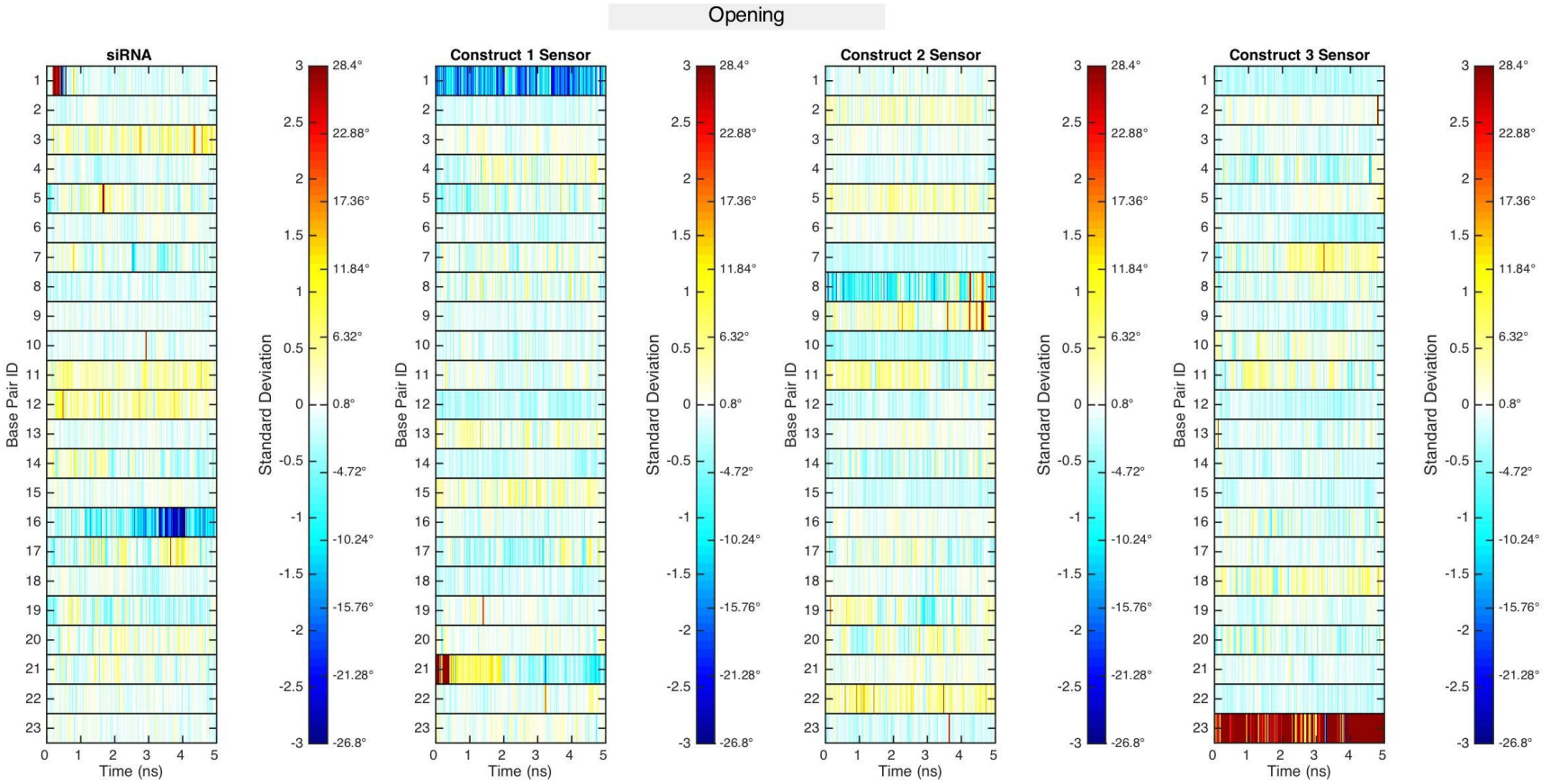


Figure A1-21 Opening of isolated natural duplexes.

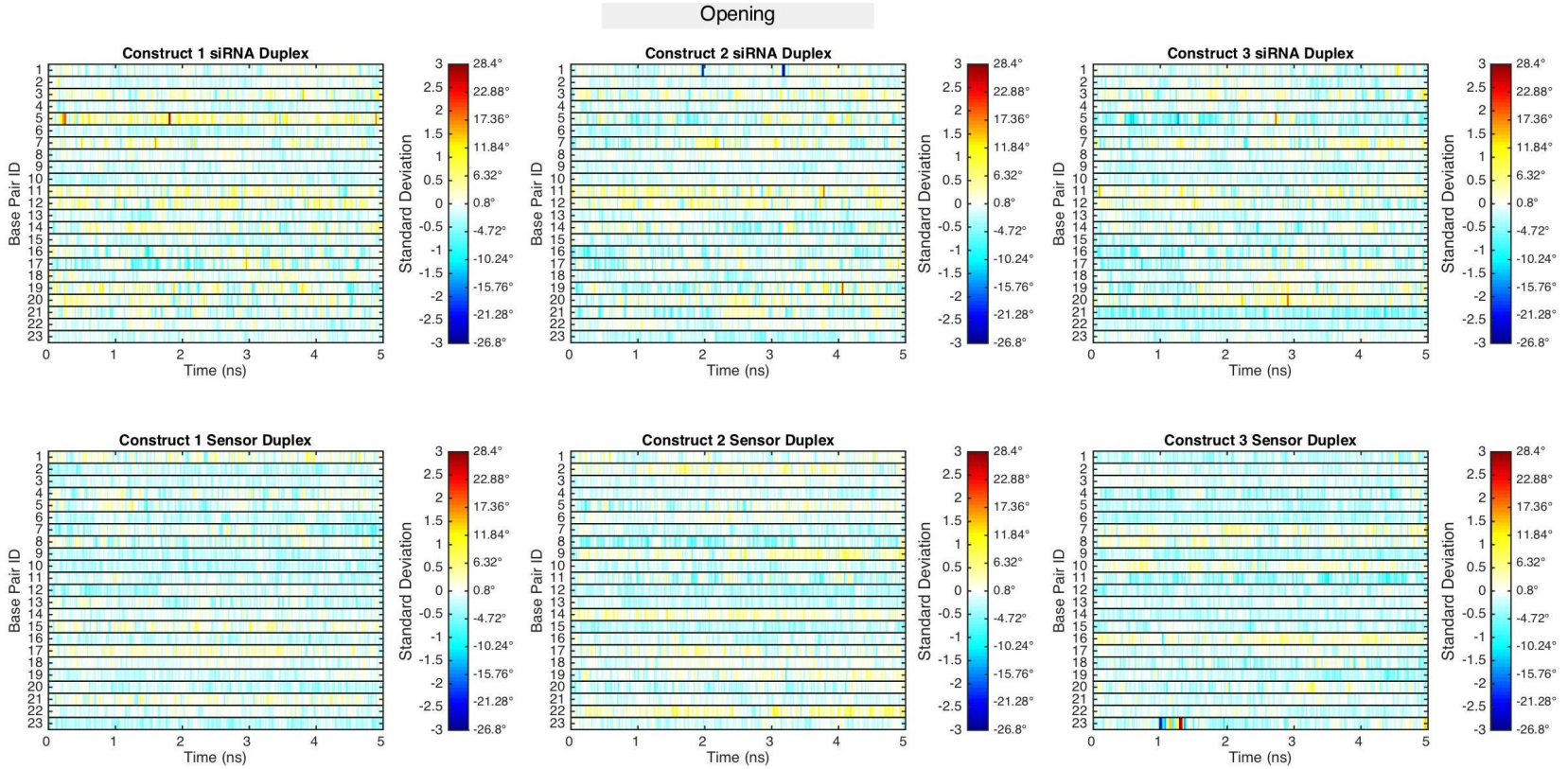


Figure A1-22 Opening of *Cond*-siRNAs.

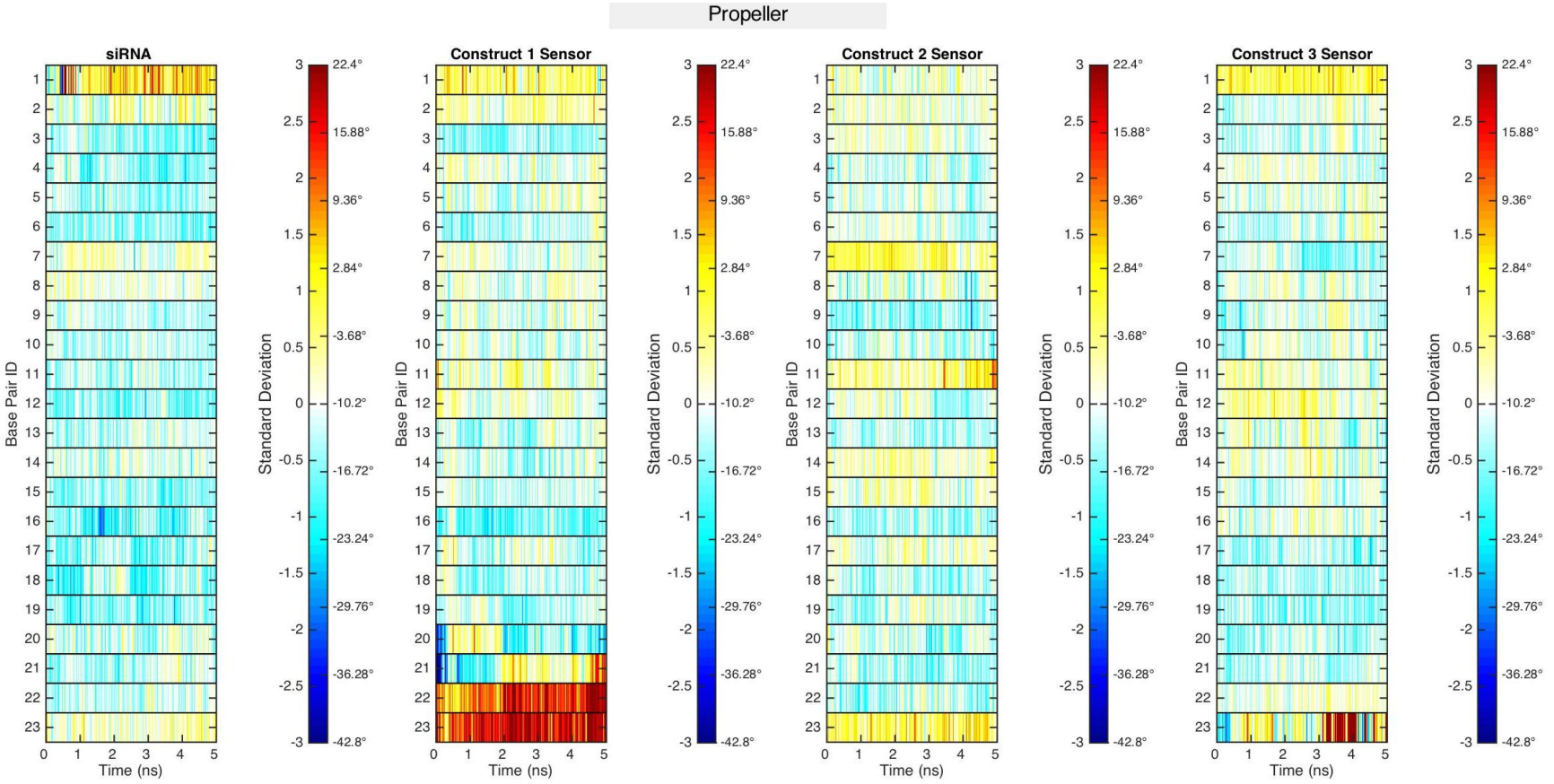


Figure A1-23 Propeller of isolated natural duplexes.

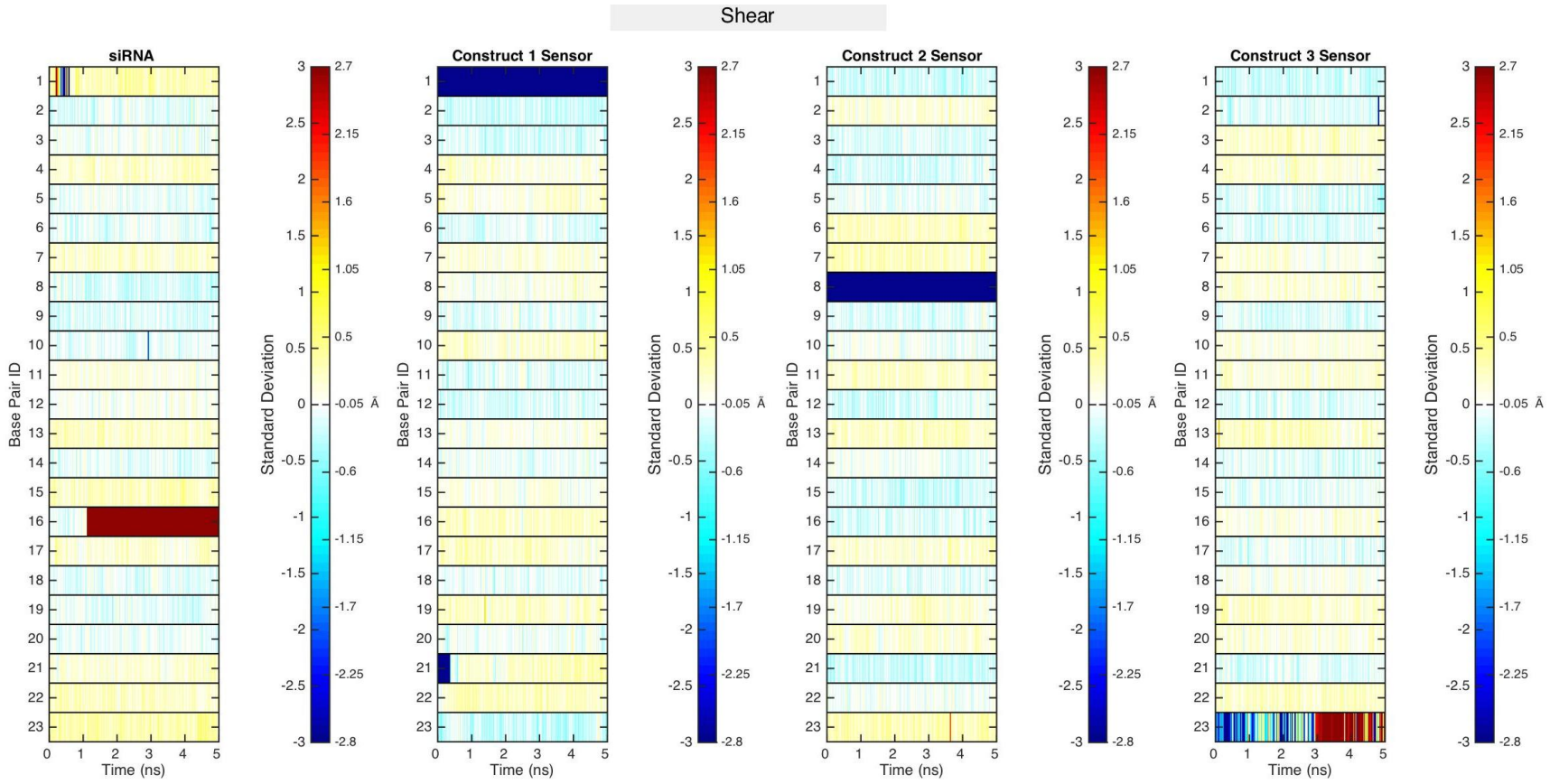


Figure A1-24 Shear of isolated natural duplexes.

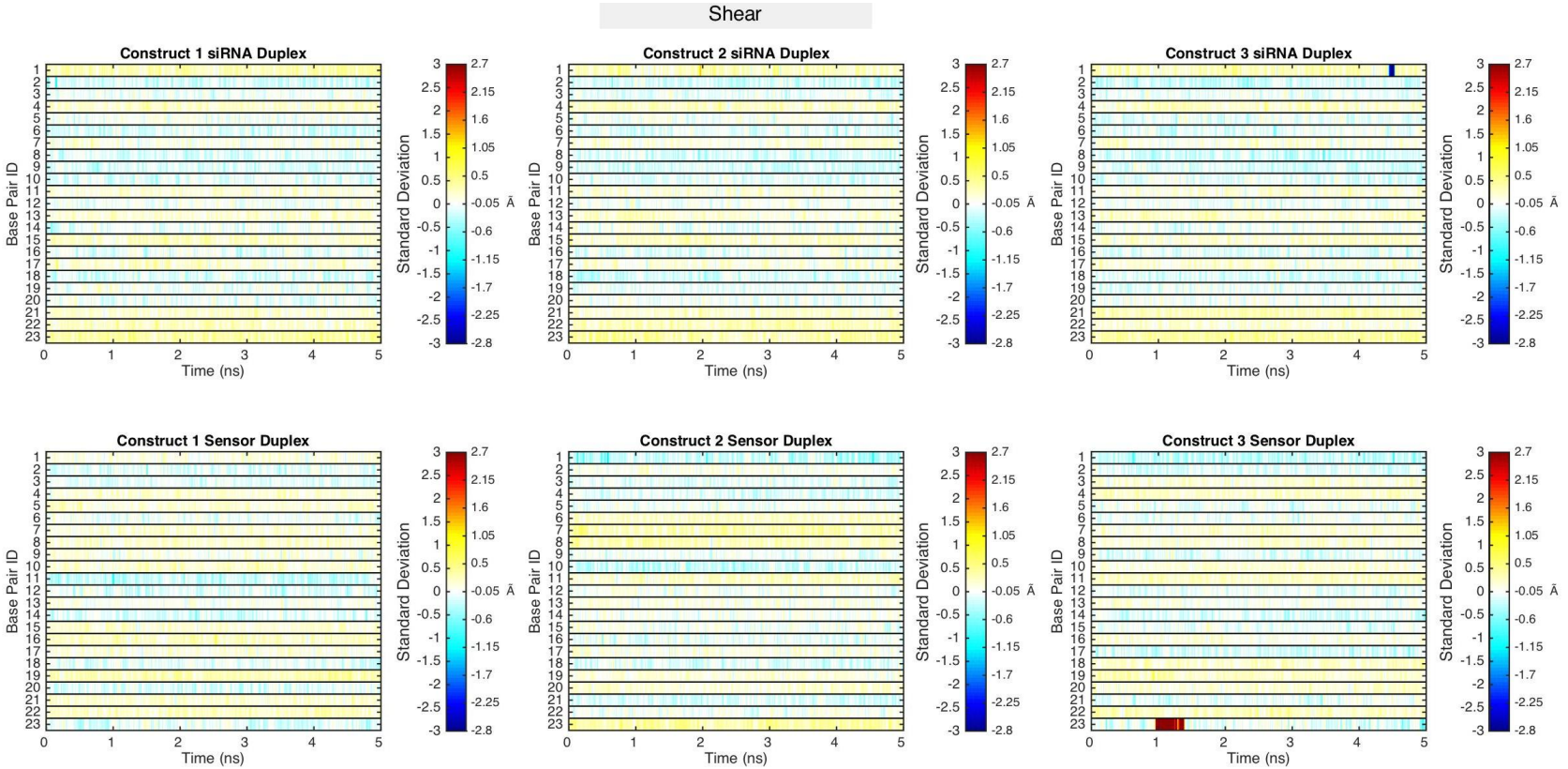


Figure A1-25 Shear of *Cond*-siRNAs.

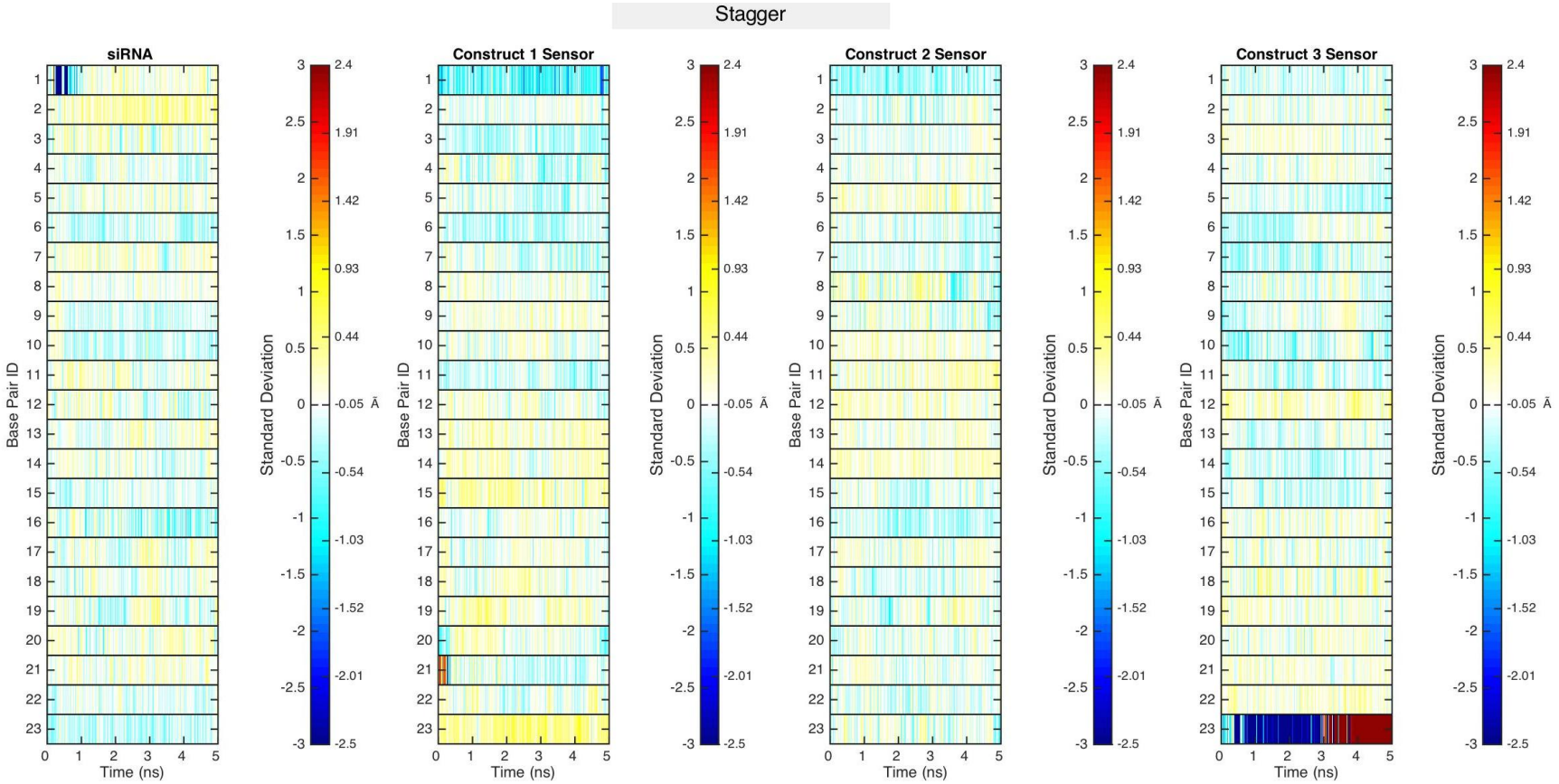


Figure A1-26 Stagger of isolated natural duplexes.

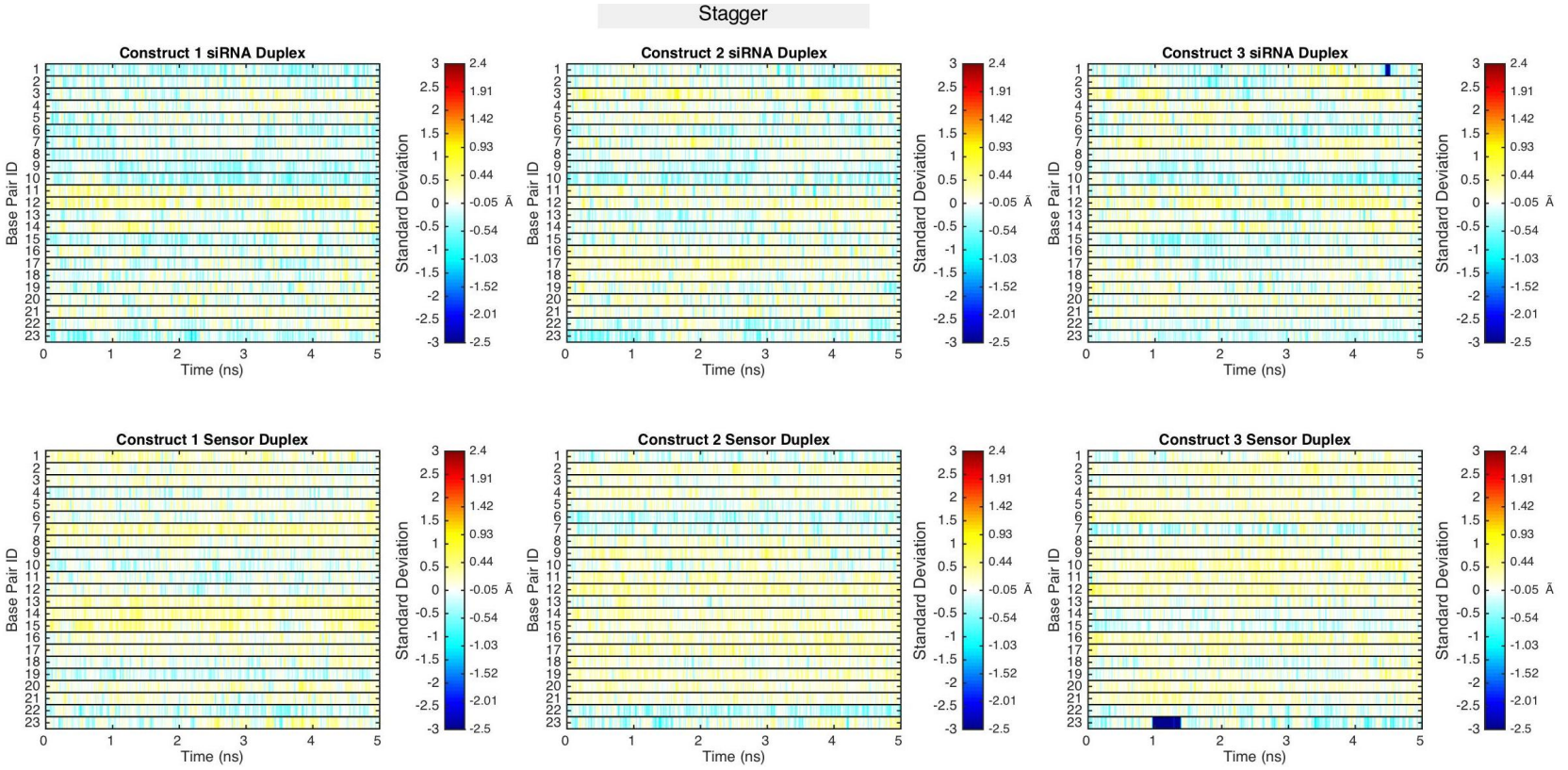


Figure A1-27 Stagger of *Cond*-siRNAs.

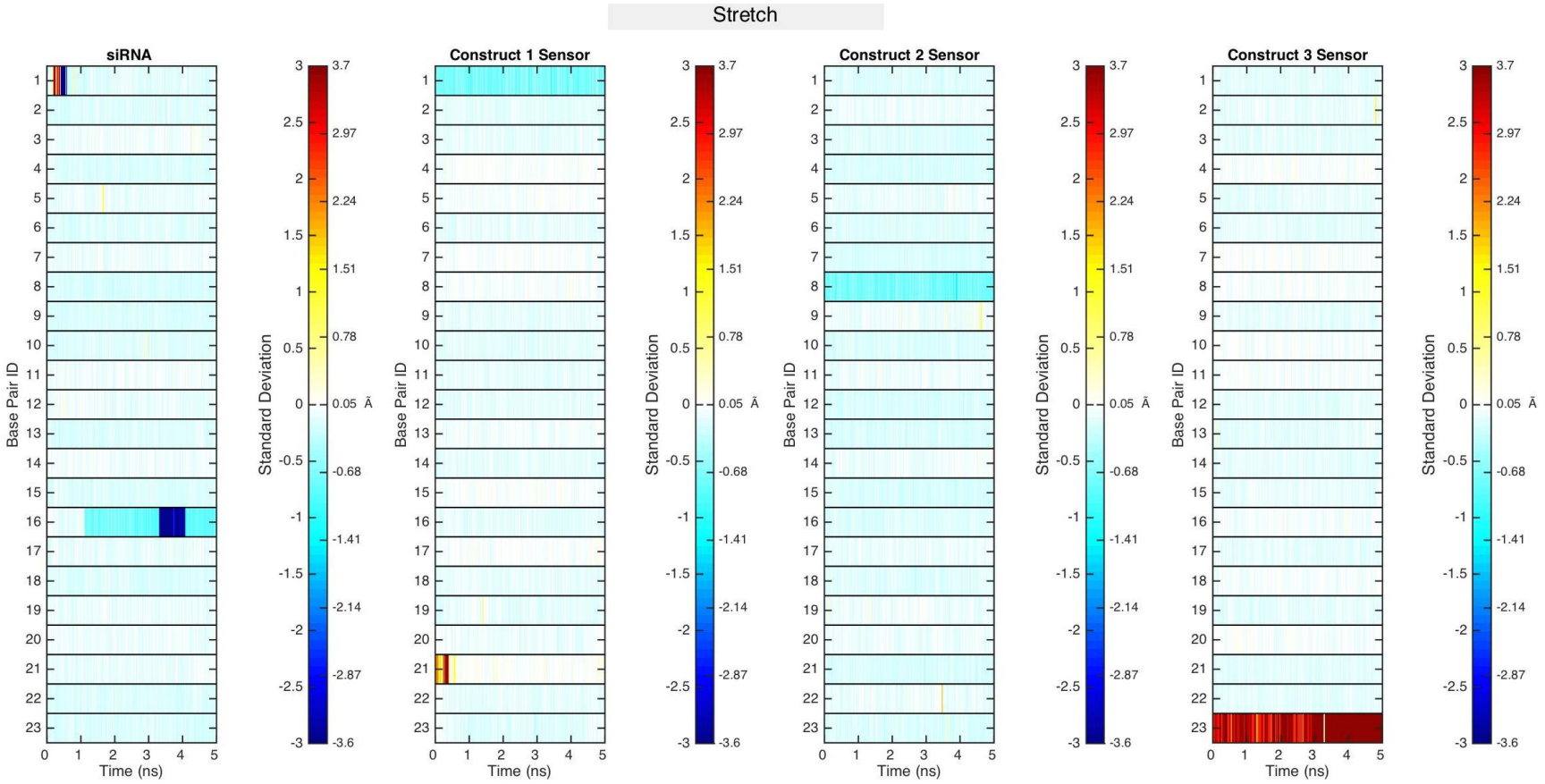


Figure A1-28 Stretch of isolated natural duplexes.

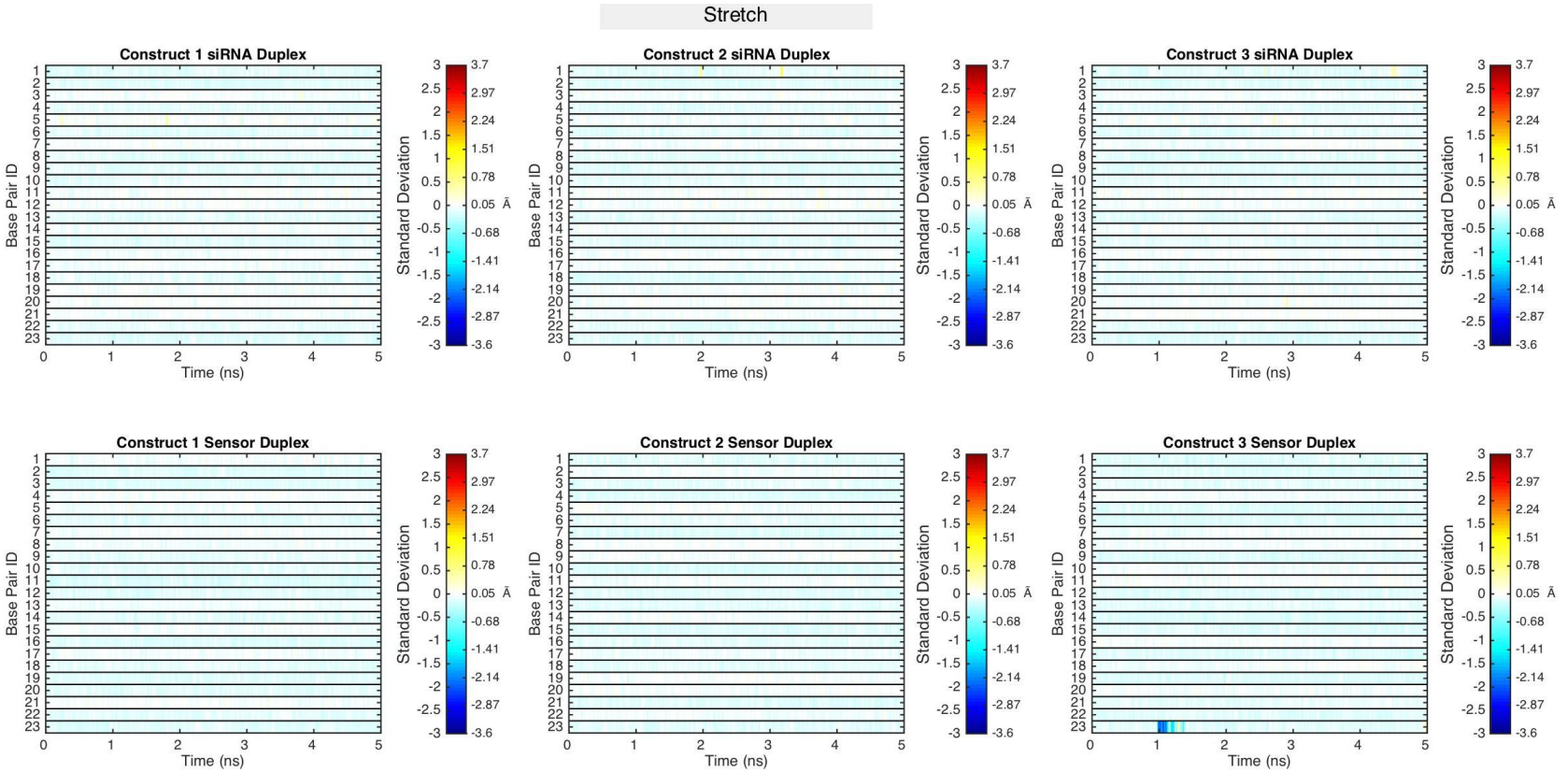


Figure A1-29 Stretch of *Cond*-siRNAs.

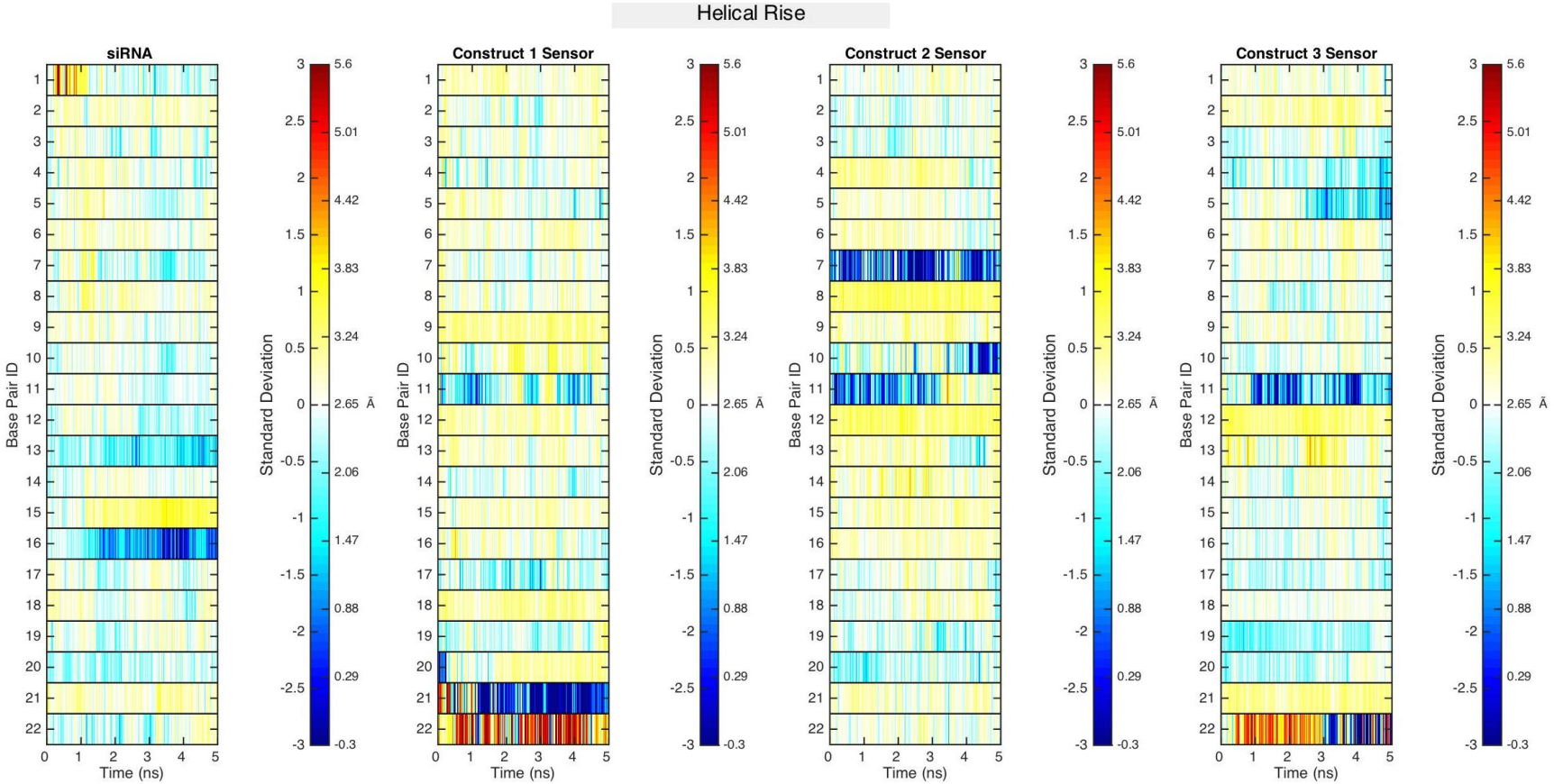


Figure A1-30 Helical Rise of isolated natural duplexes.

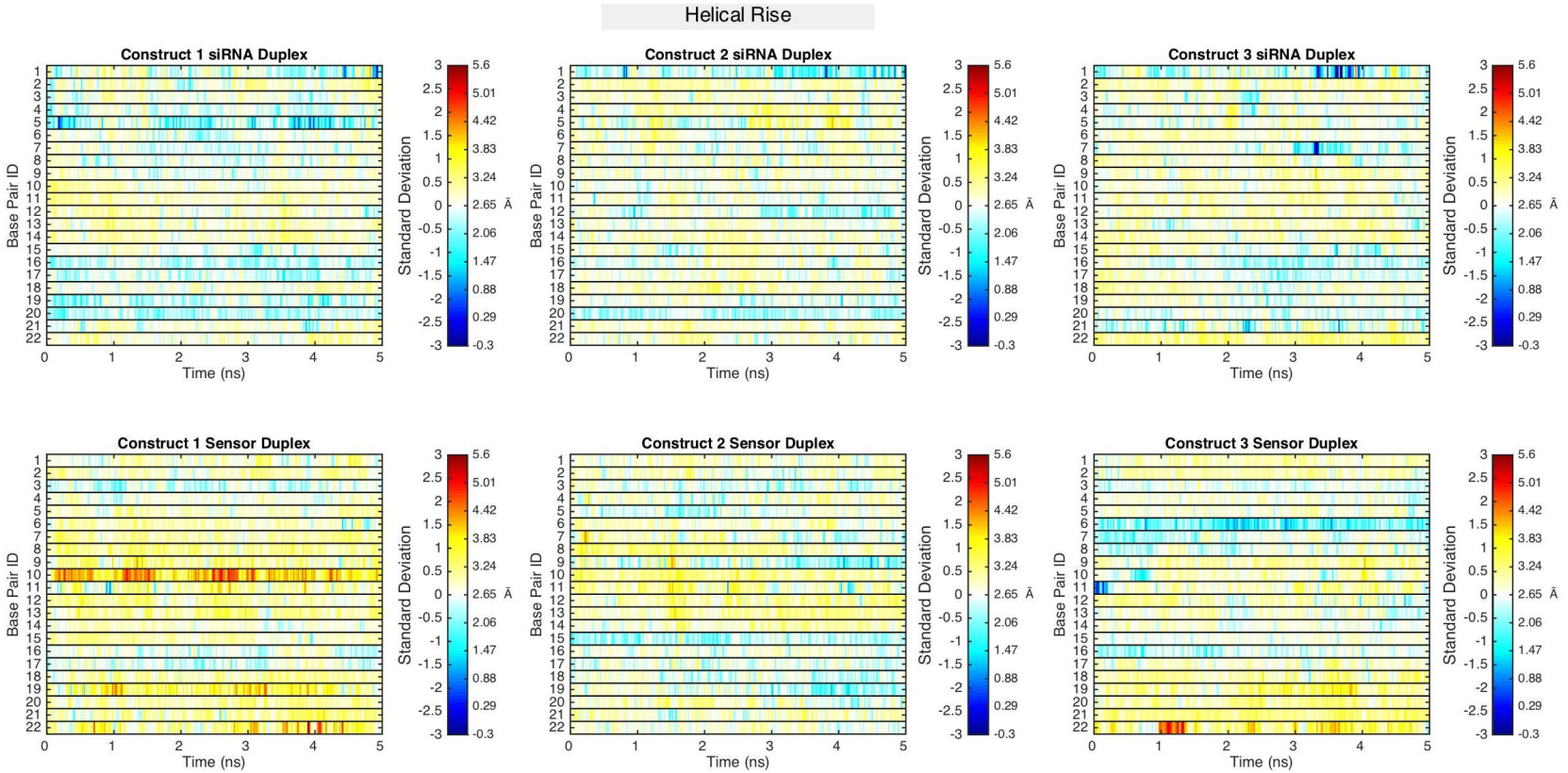


Figure A1-31 Helical Rise of *Cond*-siRNAs.

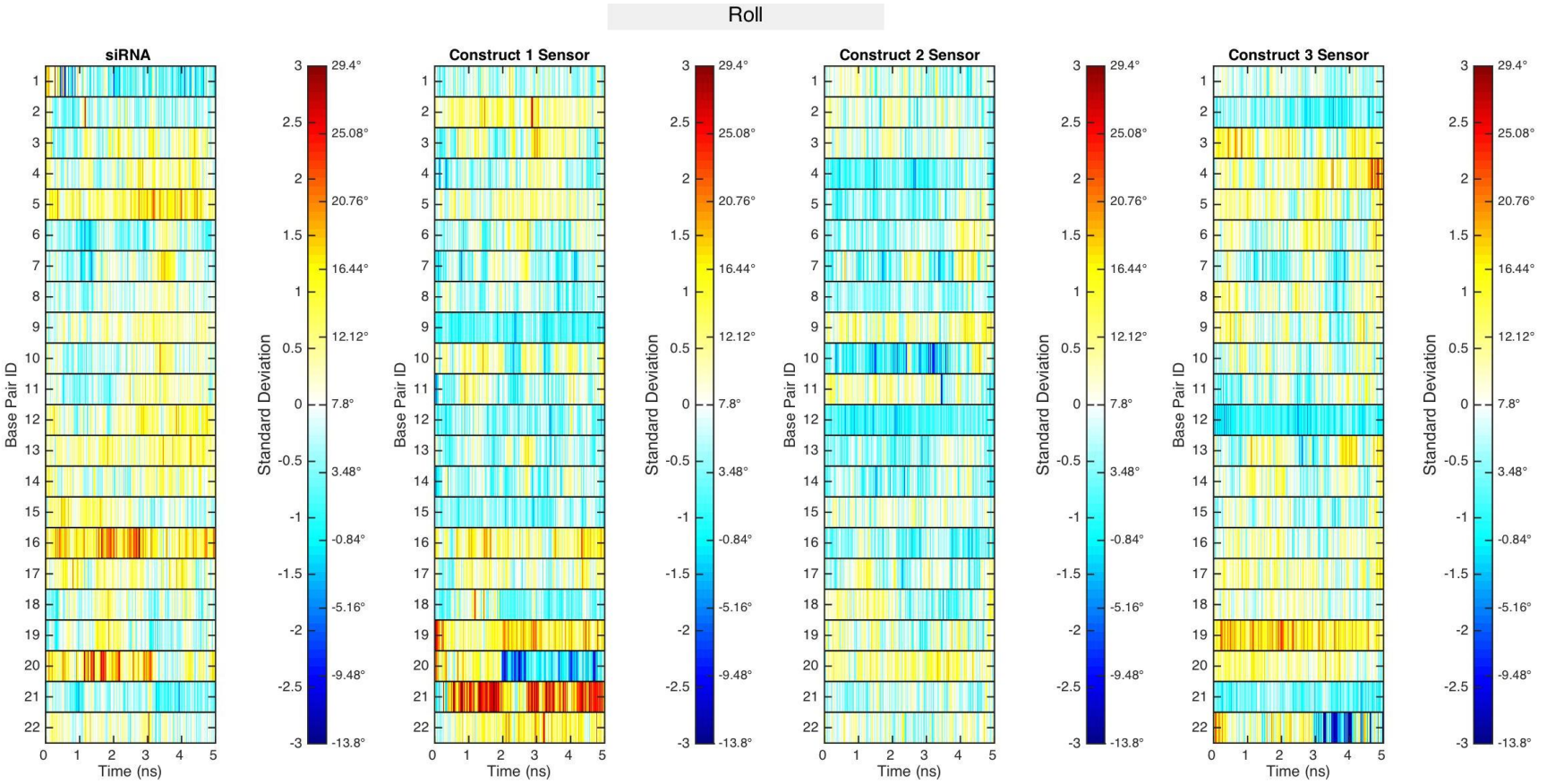


Figure A1-32 Roll of isolated natural duplexes.

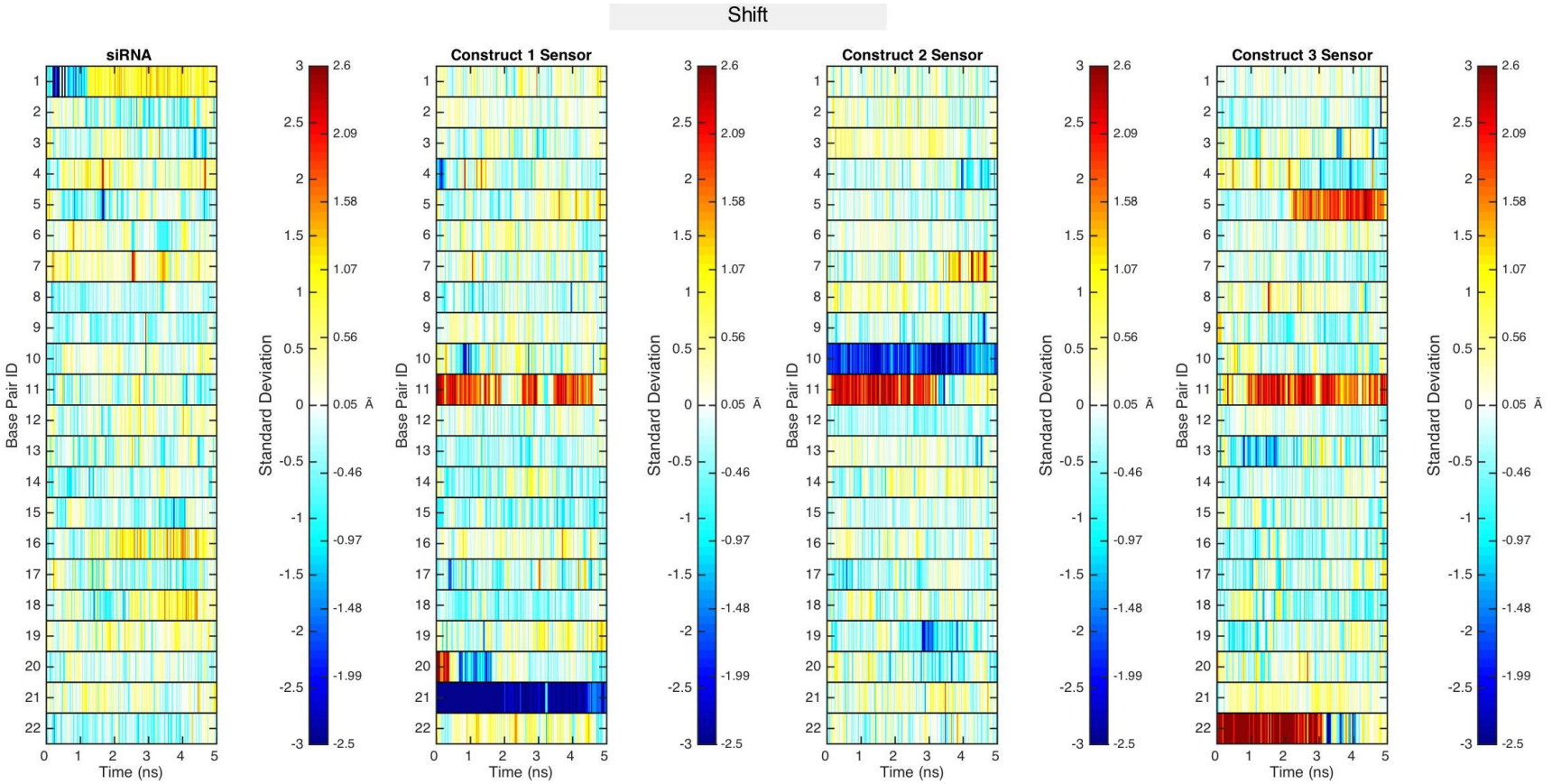


Figure A1-33 Shift of isolated natural duplexes.

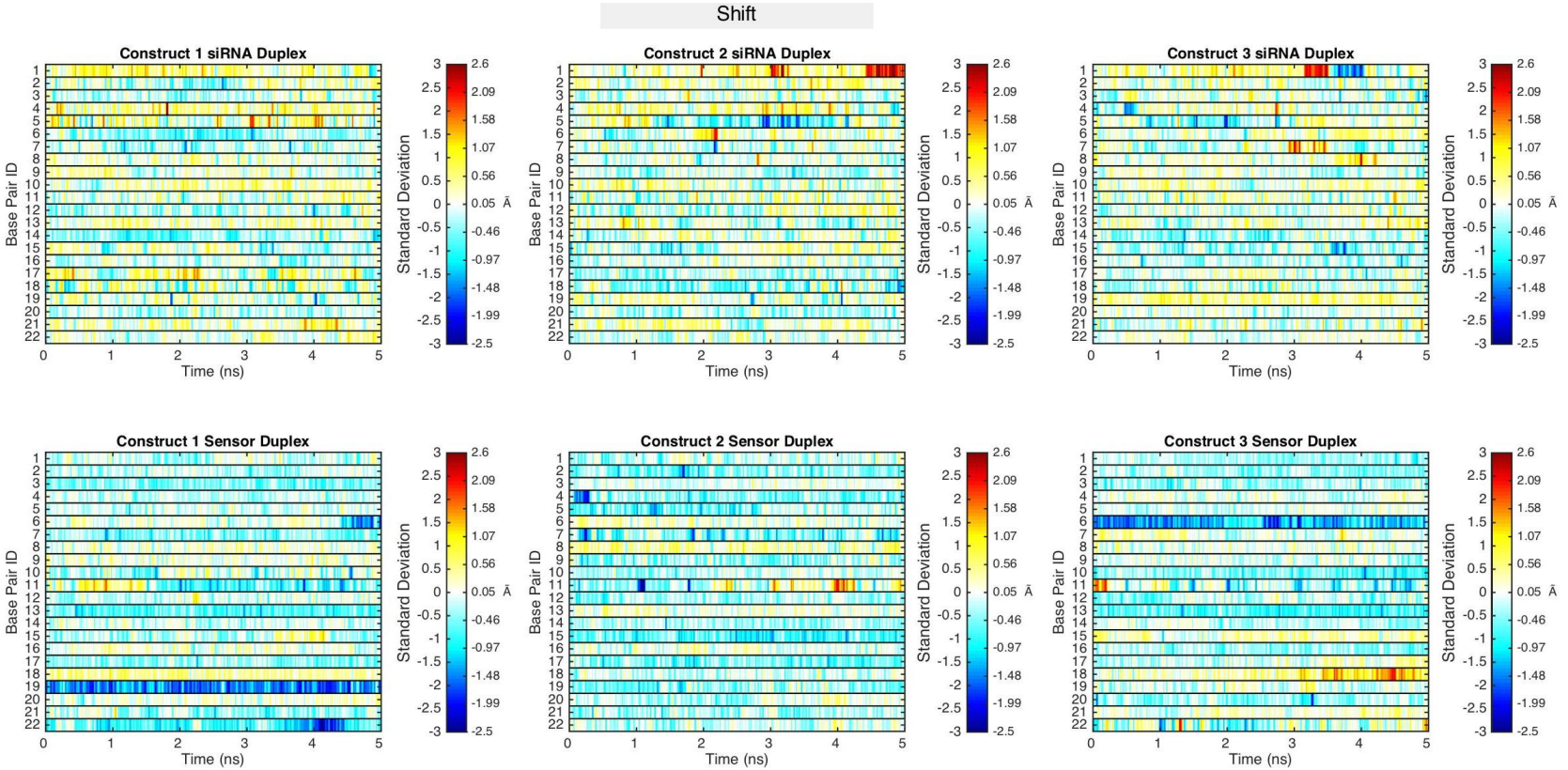


Figure A1-34 Shift of *Cond*-siRNAs.

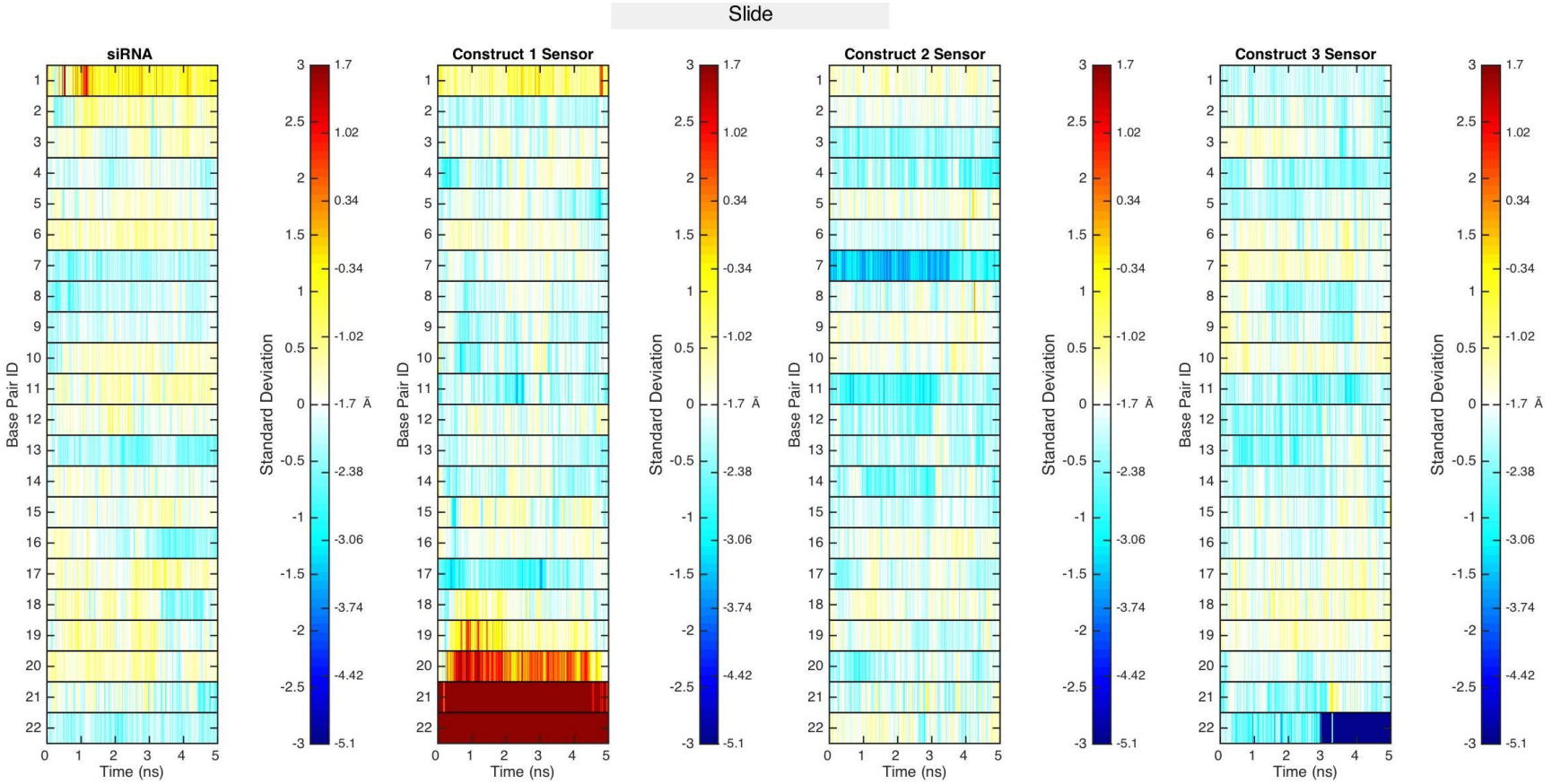


Figure A1-35 Slide of isolated natural duplexes.

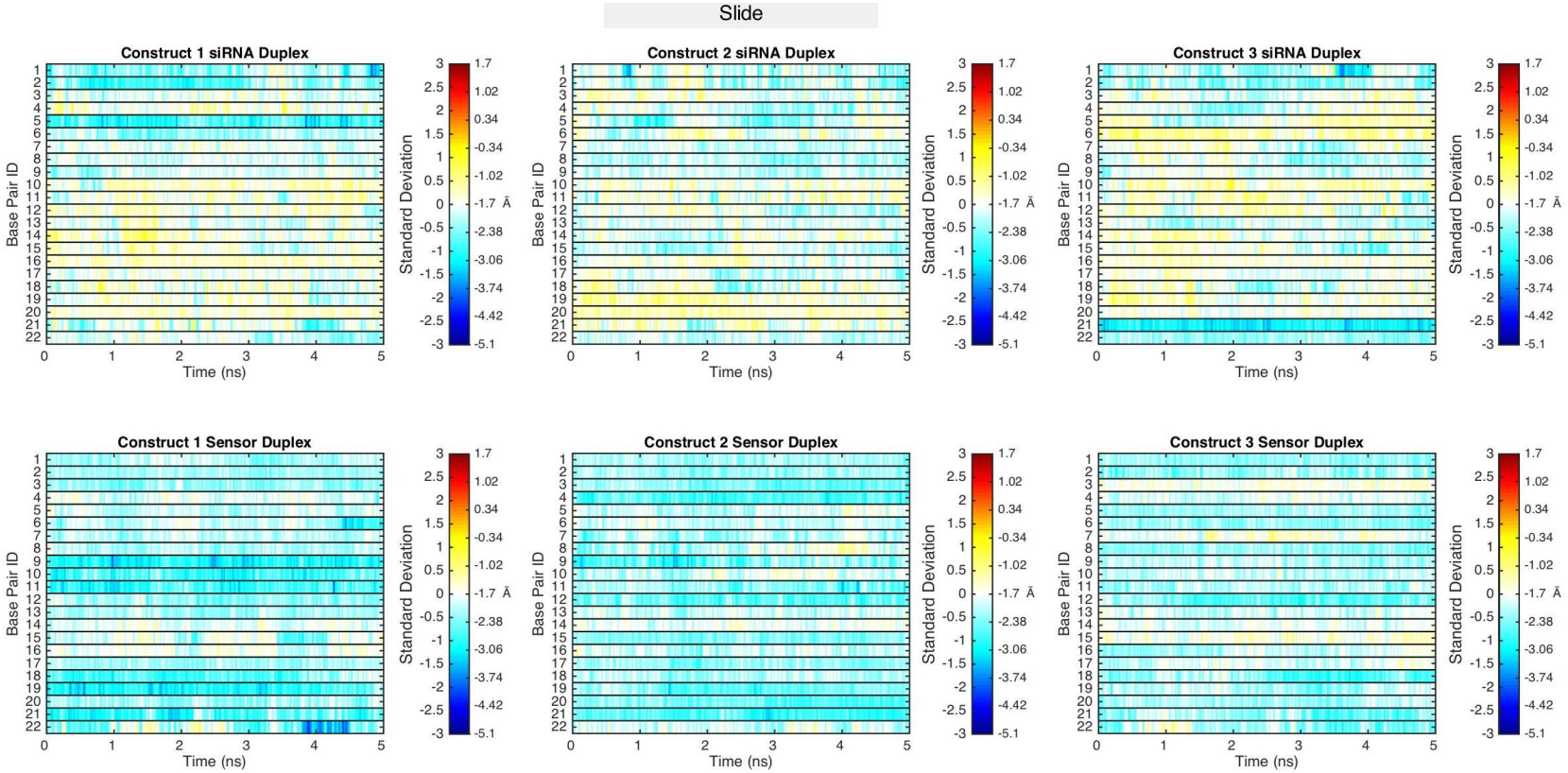


Figure A1-36 Slide of *Cond*-siRNAs.

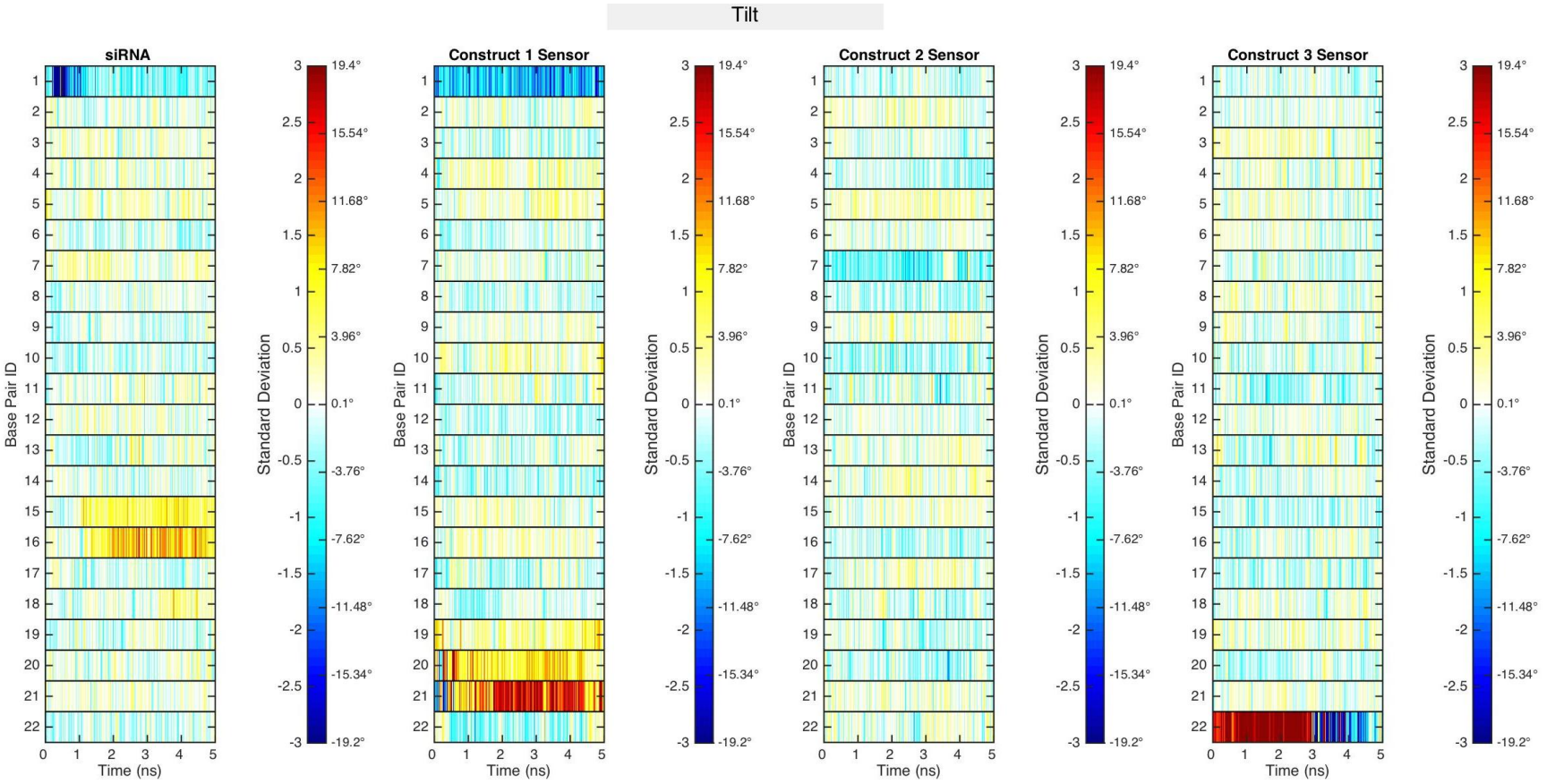


Figure A1-37 Tilt of isolated natural duplexes.

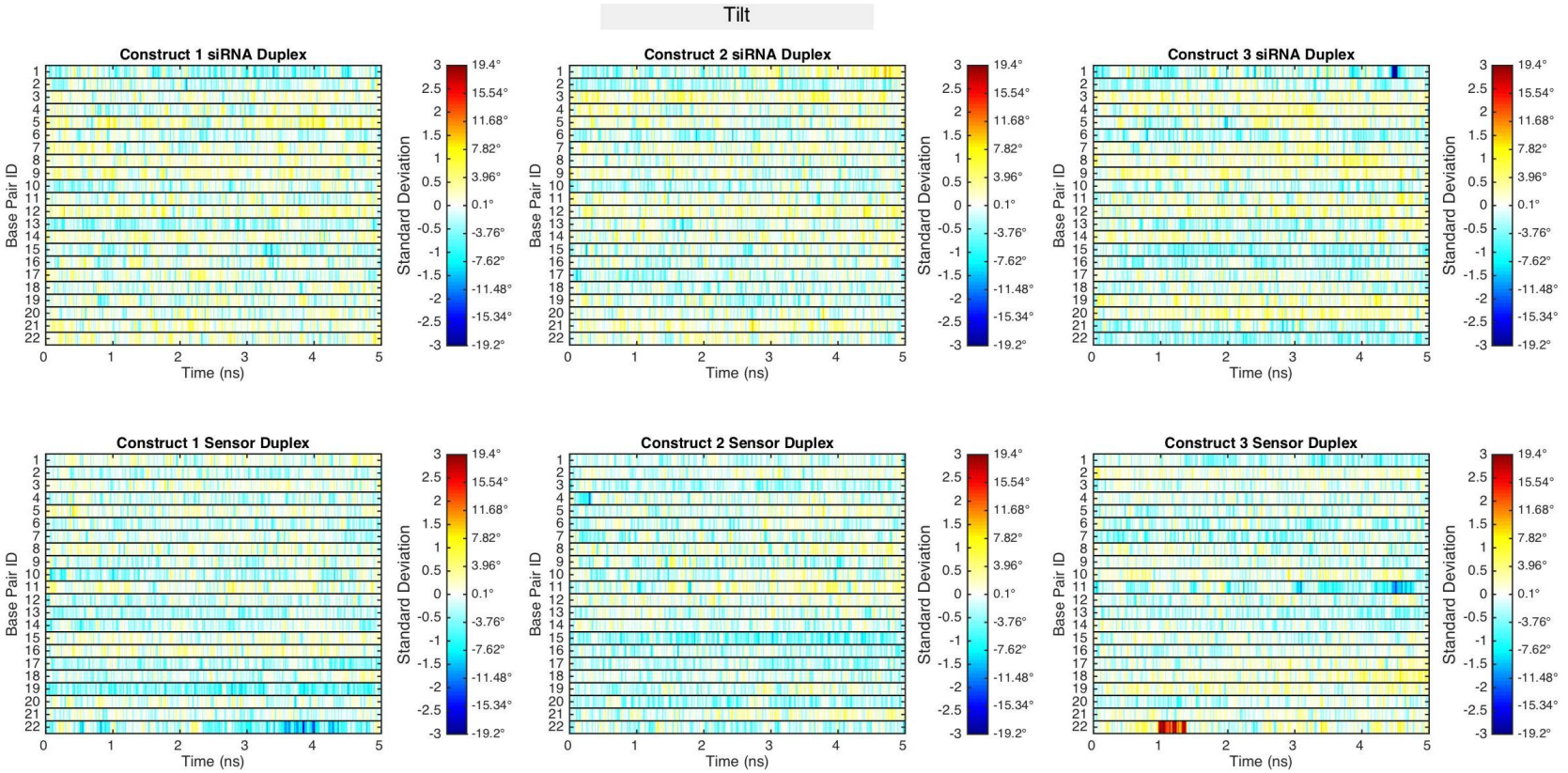


Figure A1-38 Tilt of *Cond*-siRNAs.

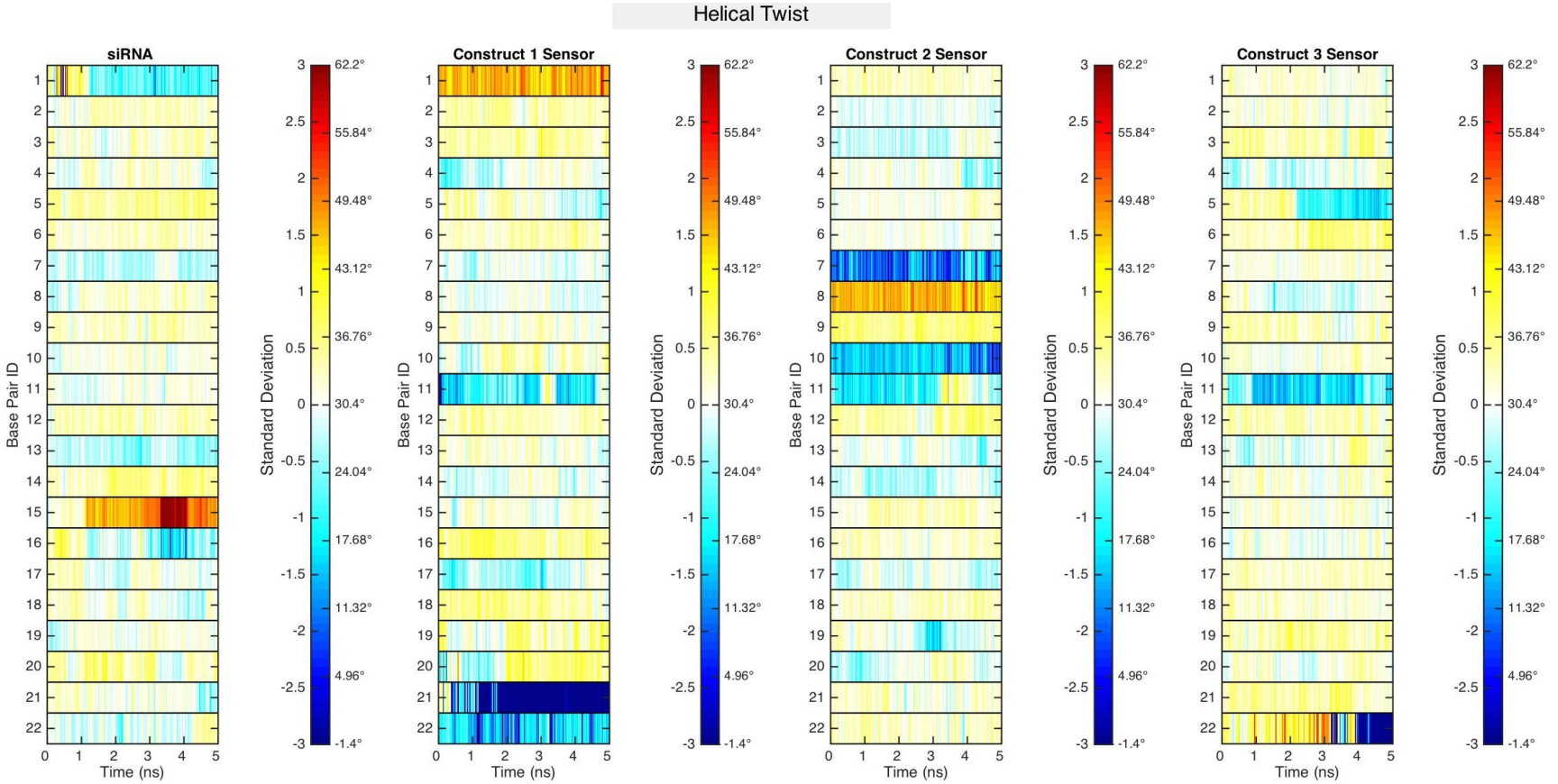


Figure A1-39 Helical Twist of isolated natural duplexes.

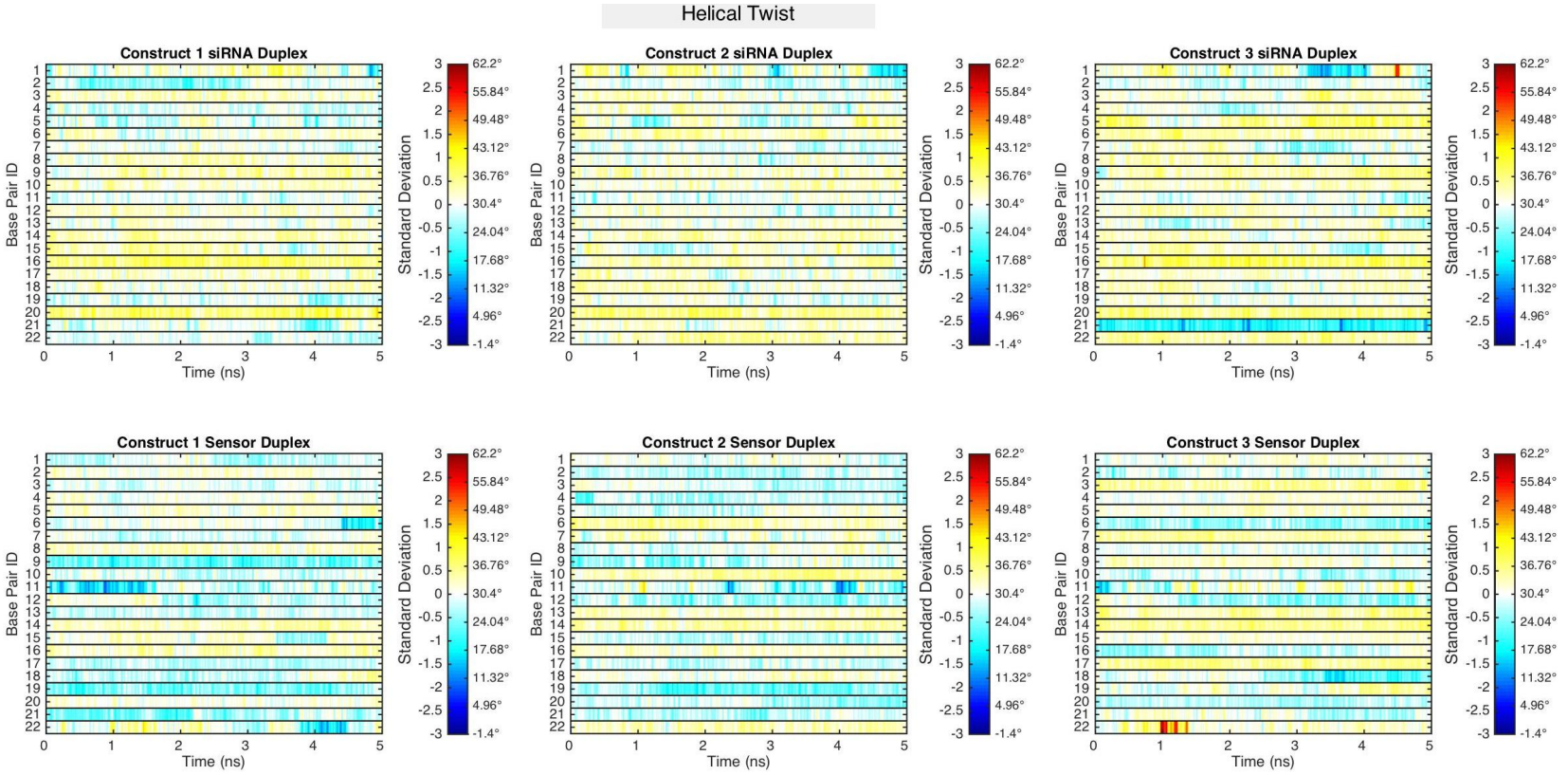


Figure A1-40 Helical Twist of *Cond*-siRNAs.

Appendix B

Supplementary Material for Holey Graphene as a Weed Barrier for Molecules.

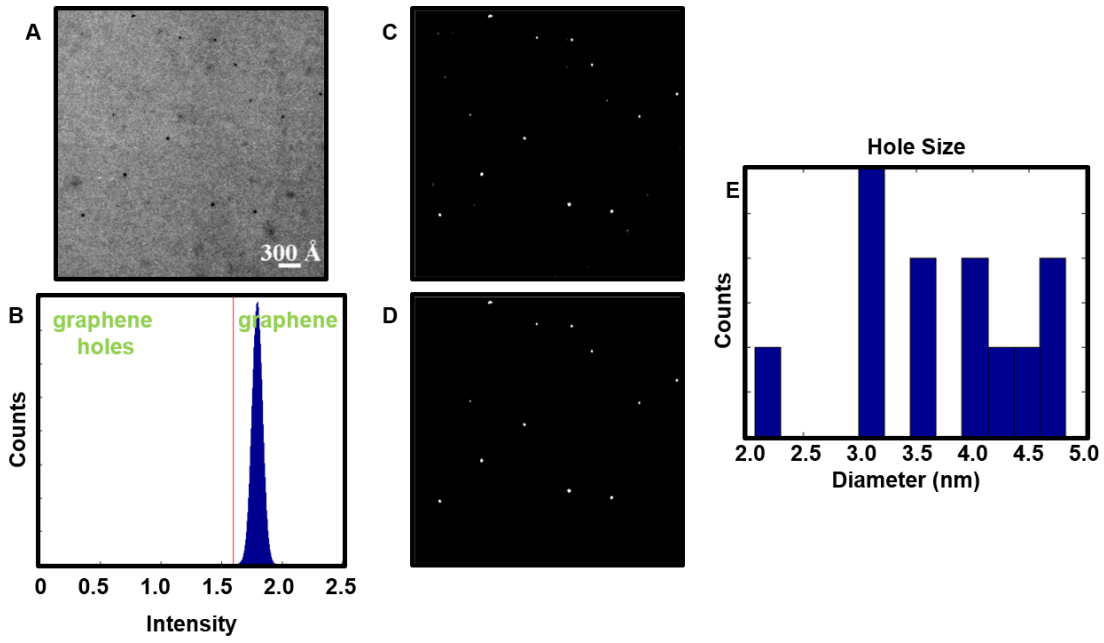


Figure A2-1 (A) Original transmission electron microscopy image from Figure 2 before segmentation. (B) An image histogram of the data in A showing the intensity threshold cut off used to create an image binary. (C) Resulting binary mask, where graphene holes are separated from the graphene layer. (D) Small outlier artifacts in the image binary are removed. (E) The diameters of the remaining holes are displayed in a bar graph, binned by diameter (10 Å bin width); we measure an average 37 ± 8 Å hole size.

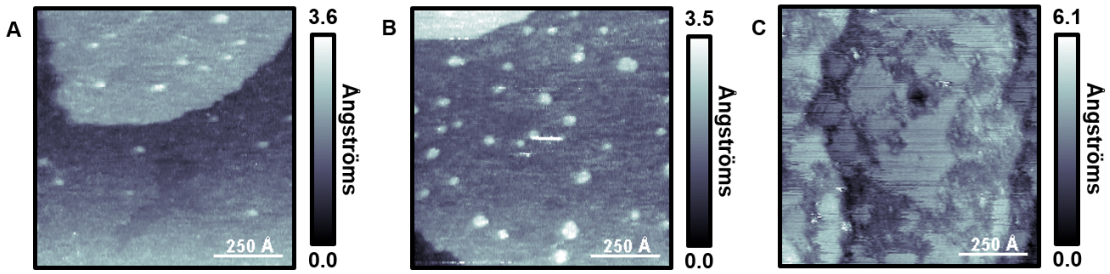


Figure A2-2 (A,B) Scanning tunneling micrographs ($I_{tunneling}$ = 3 pA, V_{sample} = -1.0 V) of "holey" graphene on Au{111}/mica directly after deposition from solution of water and acetone. Images show protrusions and depressions, displayed as brighter and dimmer, respectively. We attribute the higher protrusions as solvent that has not desorbed from the holes, and depressions as holes (without solvent) within the graphene overlayer. (C) After annealing at 100 ˚C for 24 h, all solvent is evaporated and only the depressions (holes) remain.

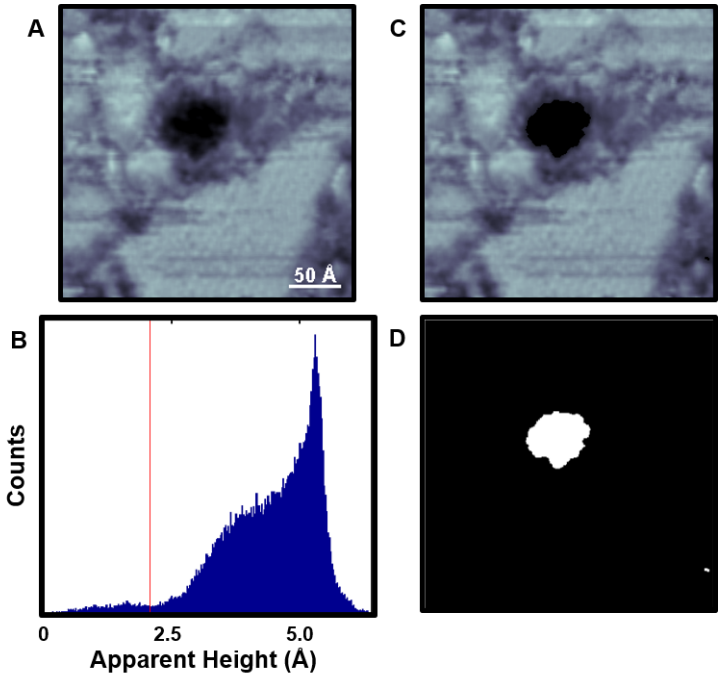
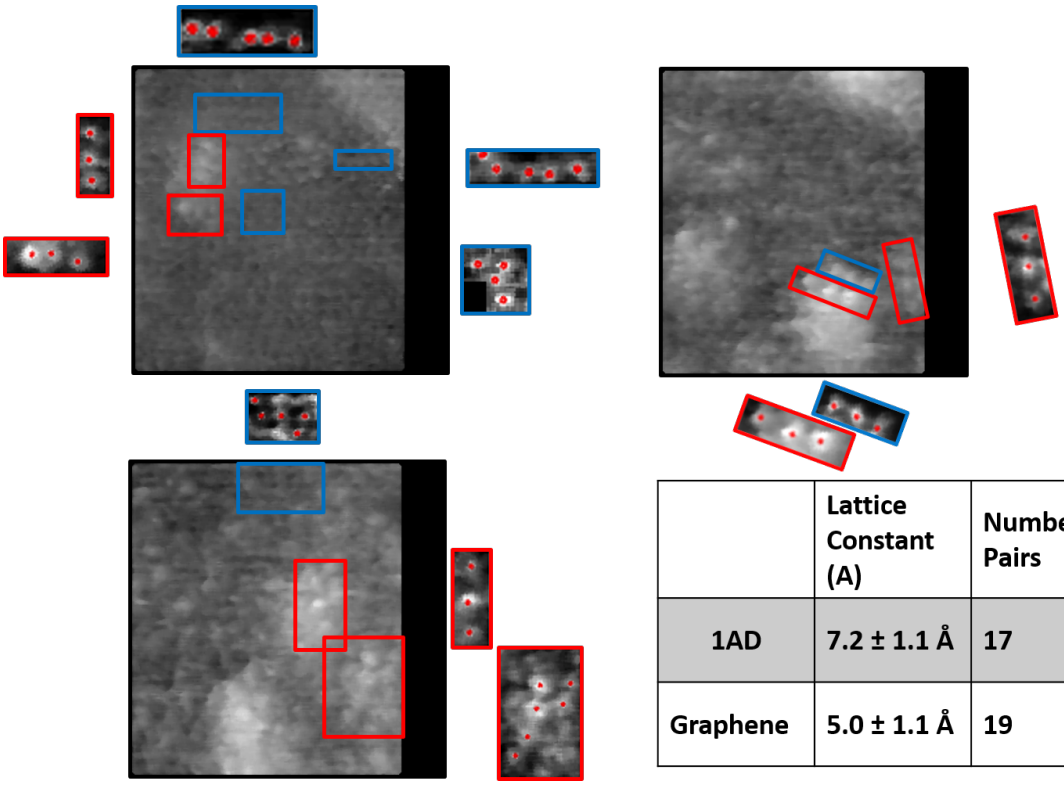


Figure A2-3 (A) Scanning tunneling micrograph ($I_{tunneling}$ = 3 pA, V_{sample} = -1.0 V) of "holey" graphene on Au{111}/mica with (B) a corresponding apparent height histogram. Masking techniques, performed in MATLAB, enable "holey" regions and graphene regions to be isolated and analyzed independently. (C) The image in A is segmented by apparent height. The graphene layer is 2.1 ± 1.2 Å higher in average apparent height compared to (D) the exposed Au region.



	Lattice Constant (A)	Number of Pairs
1AD	7.2 ± 1.1 Å	17
Graphene	5.0 ± 1.1 Å	19

Figure A2-4 Scanning tunneling micrographs ($I_{tunneling}$ = 3 pA, V_{sample} = -1.0 V) of "holey" graphene filled with 1-adamantanethiolate (1AD) on Au{111}/mica, where the spacing between adjacent 1AD molecules and graphene atoms is recorded. Images of the molecules were first smoothed and then analyzed using the Regionprops function in Matlab in the molecular regions highlighted. The inserted molecular layer shows an average spacing (across multiple images) of 7.2 ± 1.1 Å, while the graphene mask shows an average spacing of 5.0 ± 1.1 Å.

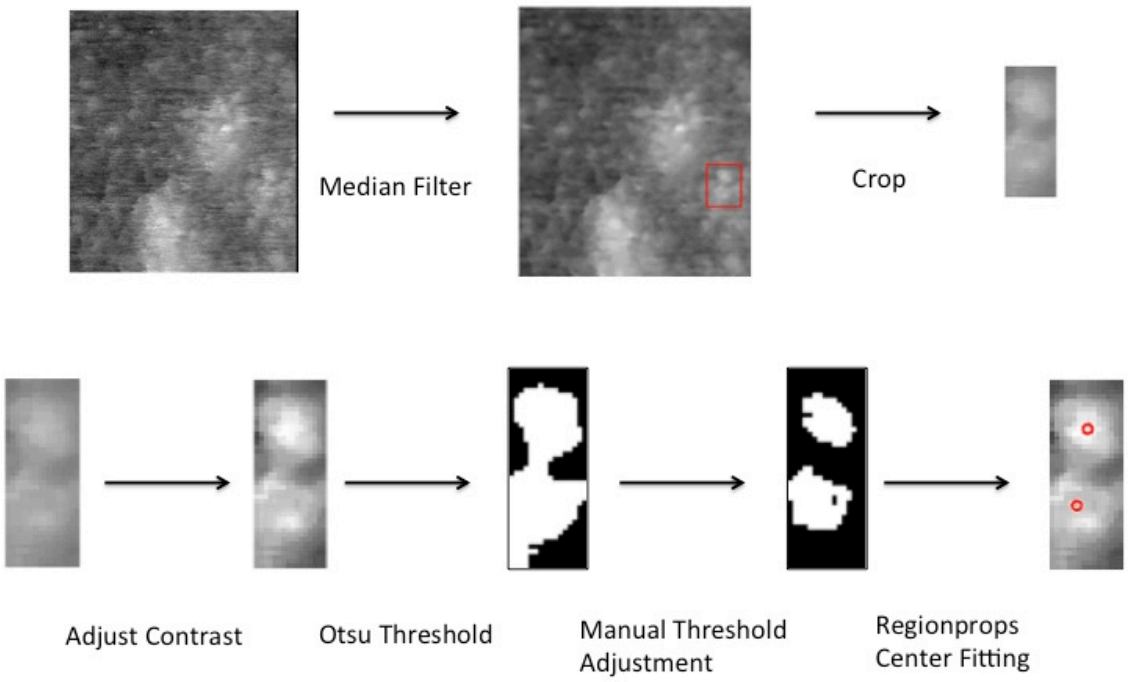


Figure A2-5 Molecule Fitting Methodology. To determine nearest-neighbor spacings between molecules post-1AD deposition, molecules were fit using the Regionprops function in Matlab. A median filter is applied to remove intensity spikes, and then the region of interest is cropped for analysis. The contrast of the cropped image is enhanced, and then the image is thresholded using the Otsu cutoff. The cutoff was increased until sufficient segmentation was achieved. The average adjustment was 0.16 where images were set to a grayscale. Finally, the center of each segmented molecule was determined. The locations of these centers were used to calculate nearest-neighbor distances. Fittings were also performed on regions that were analyzed in Fourier space to crosscheck results.

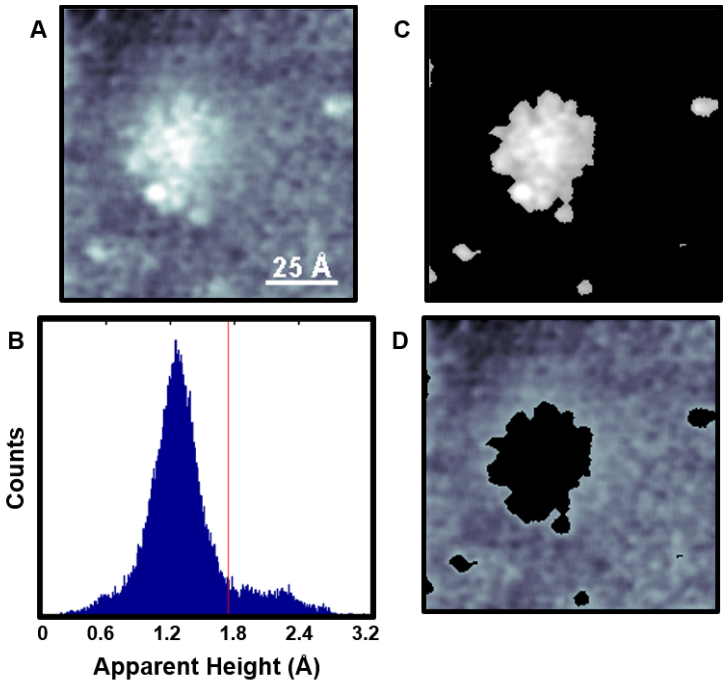


Figure A2-6 (A) Scanning tunneling micrograph ($I_{tunneling}$ = 3 pA, V_{sample} = -1.0 V) of "holey" graphene filled with 1-adamantanethiolate on Au{111}/mica with (B) a corresponding apparent height histogram. Masking techniques, performed in MATLAB, enable filled regions and bare graphene regions to be isolated and analyzed independently. (C,D) The image in A is segmented by apparent height and displayed. A 1-admantanethiolate patch appears on average 1.1 ± 0.5 Å than the graphene layer.

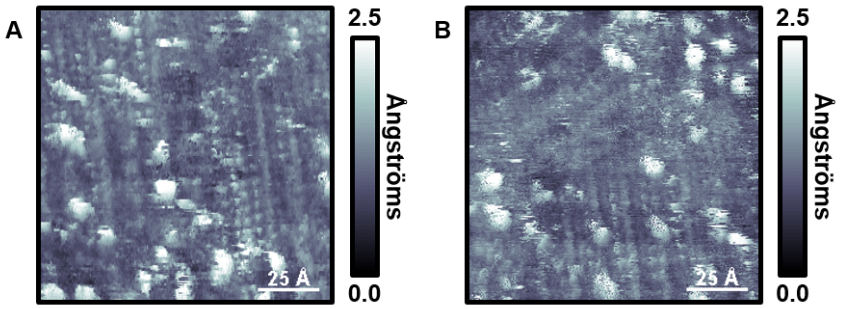


Figure A2-7. (A, B) Scanning tunneling micrographs ($I_{tunneling}$ = 3 pA, V_{sample} = -1.0 V) of "holey" graphene on Au{111}/mica after a second 1-adamantanethiolate vapor deposition for 24 h. Each sample was regenerated, prior to the second deposition step, by annealing at 250 ˚C. Images depict 1AD molecules within a "holey" graphene framework.