Annual Report for Emory Project G-33-C23

PI: Nicholas V. Hud, School of Chemistry and Biochemistry, Georgia Institute of Technology.

Contract Information: This project represents a subcontract of "Towards synthetic biology: the replication of synthetic polymers", an NSF Collaborative Research Center awarded to David Lynn of Emory University (NSF grant number 0404677).

Period of Report: June 1, 2006-May 31, 2007

Personnel Support: The past year of funding has been used to support the research of two graduate students in the Hud laboratory, Eric Englehart and Eric Horowitz.

Summary of Progress: *Aim II Activities and Findings:* Identify and characterize molecular midwifes for template-substrate assembly.

We are pleased to report that our progress on Specific Aim II has resulted in the acceptance of three peer-review publications this past year (attached). The paper by Horowitz and Hud, which appeared as a communication in *J. Am. Chem. Soc.*, describes our discovery that the small molecule proflavine binds to 2',5'-linked RNA with an association constant that is 25-fold greater than that measured for the natural 3',5'-linked RNA. We have been using proflavine as a "molecular midwife" to promote the assembly of nucleic acids for template-directed synthesis. Thus, the greater affinity of proflavine for 2',5'-linked RNA indicates that nucleic acids with this alternative backbone linkage will be even more amendable to intercalation-mediated assembly.

The experimental results presented in our *J. Am. Chem. Soc.* publication are fully consistent with proflavine binding to 2',5'-linked RNA by base pair intercalation. Nevertheless, we believed that is was necessary to obtain a high resolution structure of 2',5'-linked RNA with bound proflavine, as this is the only unequivocal means by which the mode of binding of a small molecule to a nucleic acid can be confirmed. We are pleased to report that we have recently obtained the solution-state structure of proflavine bound to the 2',5'-linked RNA duplex $[GCCGCGGC]_2$. At a ratio of two proflavine molecules per duplex, we observe proflavine bound between the two CpG steps of the duplex (Figure X). This structure is quite unique in that it represents the first solution-state structure of a simple intercalator (as opposed to a bisintercalator) bound to a nucleic acid. The greater association constant, and slower off-rate, allowed structure determination, whereas past attempts to solve similar structures with natural RNA or DNA have been stymied by the fast exchange of the intercalator between the nucleic acid and solution. We have also measured the relatively binding constants for a number of other known intercalators for 2',5'-linked RNA. This data is expected to be valuable for our development of an optimally matched pair of molecular midwife and nucleic acid backbone structures.

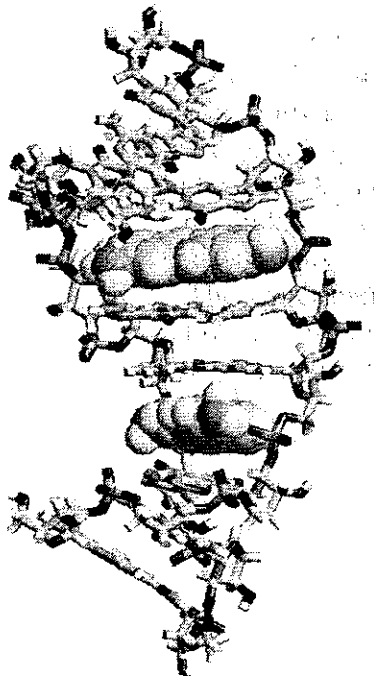

Figure X. A model of the solution-state structure of the 2',5'-linked RNA duplex $[GCCGCGGC]_2$ with bound proflavine. The model is based upon distance constraints obtained by 2D NMR spectroscopy and a refinement using the Amber suite of molecular dynamics programs.

An additional objective of Aim II was to explore the possibility of alternative base pairs (i.e. non-Watson-Crick pairs) that could be used for information transfer. Within this objective, we have been interested in the possibility that low barrier hydrogen bonds (LBHB) could be used as a means for increasing the selectivity of hydrogen bonded structures in aqueous solution. Additionally, we have been interested in the possibility of purine-purine informational systems. We had previously discovered that the small molecular coralyne can be used to propose adenine-adenine base pairs, and the greater for propensity purines to stack (as compared to pyrimidines) suggested to us that artificial purine-purine informational systems my be more amenable to self-assembly than Watson-Crick pairing systems. Combining these two ideas, we performed a computational study of adenine and related purine bases to determine if purine-purine base pairs could be formed that contain LBHBs. This work, performed in collaboration with Tom Morton of UC Riverside, is now in press at *J. Phys Chem.* Briefly, we have identified an adenine-adenine base pair that is held together by a LBHB, based upon energy levels determined by B3LYP/6-31G** level calculations. The structure of this base pair is extremely close to that of the purine-purine G·A N7-N1 amino-carbonyl base pair that is observed in naturally occurring RNA structures (Figure Y). Thus, we believe that it will be possible to use this base pair in non-natural informational systems that have helical structures similar to that of RNA. Interestingly, the asymmetry of this base pair requires that one adenine be the natural N9-glycoside, whereas the other is the N7-glycoside (Figure Y). It may therefore be possible to create a non-natural informational system with high specificity from two isomers of adenosine. We are beginning to explore this possibility in the laboratory.

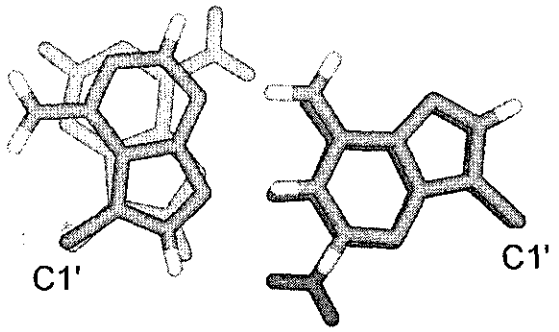

Figure Y. The structure of a (proton bound) base pair between adenosine and the N7-nucleoside of adenine that is predicted to contain a low barrier hydrogen bond between the N1 of adenine and the N9 of the N7-nucleoside of adenine. This proposed base pair is superimposed on top the G·A N7-N1 amino-carbonyl base pair that is known to occur in ribosomal RNA.

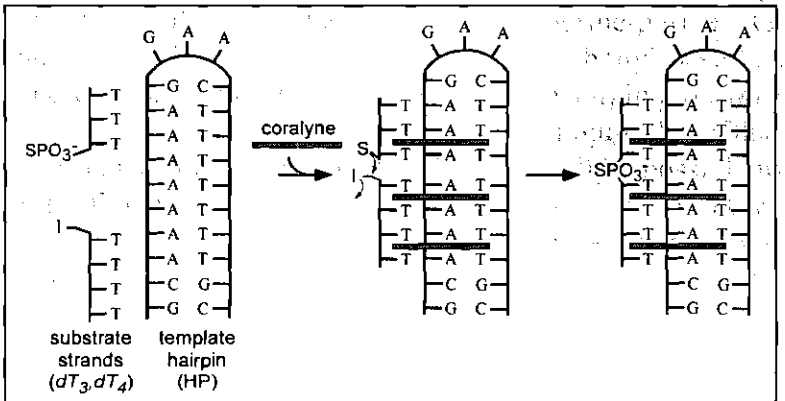

Fig. 4. An experimental test system used to explore the ability for coralyne to act as a midwife in a template-directed ligation reaction that includes A·T·T base *pairs triplets*. (Left to Right) 5'-phosphorothioate-$(dT)_3$ and 3'-iodo-$(dT)_4$ substrate strands that are too short to form a stable triplex with the indicated DNA hairpin. The addition of coralyne promotes the assembly of a ligation-active complex, which leads to the formation of the $(dT)_7$ product within an internal phosphorothioate linkage. Experimental data is shown in Hud *et al.* (Chemistry & Biodiversity, in press).

In a third publication, resulting from collaborations between the Hud and Lynn laboratories (*Chemistry & Biodiversity*, in press), we have demonstrated our ability to use intercalation to drive template-directed synthesis in a non-duplex system. Specifically, we have shown that is possible to use coralyne, a crescent-shaped molecule, to transfer sequence information via a triple helix structure (Figure Z). In contrast proflavine, which works as a midwife for

information transfer in a duplex system, is ineffective in the triplex system. Likewise, coralyne does not promote information transfer in a duplex system. These results provide substantial support for our proposal that information transfer between natural (and non-natural) pairing motifs can be selectively promoted by intercalation, if the structure of the intercalator matches that of the desired base assembly.