



ChemComm

**Sequence-directed Dynamic Covalent Assembly of Base-4-  
encoded Oligomers**

Journal:	<i>ChemComm</i>
Manuscript ID	CC-COM-02-2020-001083.R1
Article Type:	Communication

SCHOLARONE™  
Manuscripts

## Sequence-directed Dynamic Covalent Assembly of Base-4-encoded Oligomers

Received 00<sup>th</sup> January 20xx,  
Accepted 00<sup>th</sup> January 20xx

S. C. Leguizamon,<sup>a†</sup> M. F. Dunn,<sup>a†</sup> and T. F. Scott<sup>\*a,b,c</sup>

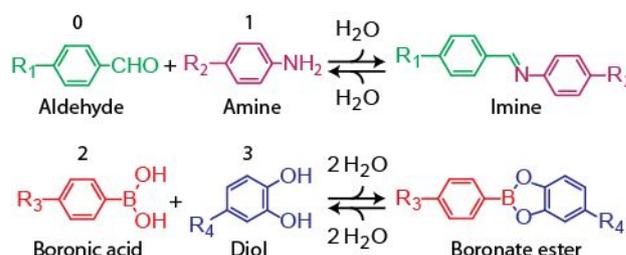
DOI: 10.1039/x0xx00000x

[www.rsc.org/](http://www.rsc.org/)

**As an information-bearing biomacromolecule, DNA is encoded in base-4, where each residue site can be occupied by any one of four nucleobases. Mimicking the information dense, sequence-selective hybridization of DNA, we demonstrate two orthogonal dynamic covalent interactions to effect the selective assembly of molecular ladders and grids from base-4-encoded oligo(peptoid)s.**

Information in DNA is encoded in the nucleobase sequence affixed as pendant groups on a deoxyribose phosphate polymer backbone. The double and triple hydrogen bonding between paired nucleobase residues of adenine–thymine and guanine–cytosine, respectively, enable complementary DNA sequences to selectively self-assemble, yielding molecular ladders encoded in base-4.<sup>1</sup> Owing to this sequence-selective self-assembly, DNA has emerged as a versatile nanoconstruction medium where, upon thermal melting and annealing, mixtures of designed DNA strands can hybridize to afford complex, multi-dimensional structures.<sup>2–4</sup> Nevertheless, despite the success of DNA-based nanotechnology, reliance on interstrand hydrogen bonds and a sugar-phosphate backbone can compromise the mechanical, thermal, and chemical stability of the resultant structures.<sup>5,6</sup> Mimicking the sequence-selective hybridization of DNA in robust, abiotic systems, the self-assembly of molecular ladders incorporating covalent bond rungs from complementary oligomeric precursor strands bearing covalently coreactive pendant groups has been previously described.<sup>7–9</sup> Employing dynamic covalent interactions to mediate these self-assembly processes provides a mechanism for interstrand connectivity

rearrangement, necessary for error correction and enable the annealing of intermediate, out-of-registry species to yield the thermodynamically favored product while maintaining the strength of covalent bonds. The systems examined to date have exclusively employed a single reactant pair such that residue sequences can only be encoded in binary; in contrast, mimicking the base-4 encoding of DNA requires the use of dual concurrent and orthogonal dynamic covalent reactant pairs. Several orthogonal dynamic covalent interactions have been described.<sup>10–12</sup> In particular, imine- and boronate ester-yielding condensation reactions have been employed concurrently to synthesize macrocycles, molecular cages, and dendritic nanostructures.<sup>13–15</sup> Here, we demonstrate an abiotic, DNA-mimetic dynamic covalent system encoded in base-4 by demonstrating the sequence-directed self-assembly of molecular ladders and grids from oligomers bearing two covalently coreactive pendant group pairs whose condensation reactions are orthogonal but proceed under the same reaction conditions (Scheme S1).



**Scheme 1.** pH-mediated dynamic covalent condensation reactions. The addition of amines (0) to aldehydes (1) yields imines, while the reaction of boronic acids (2) with diols (3) affords boronate esters.

To date, studies examining the dynamic covalent assembly of molecular ladders from single-stranded oligomeric precursors have exclusively employed one pair of pendant groups whose co-reaction proceeds under dissimilar reaction conditions.<sup>8,9,16</sup> To ensure that coreaction of dual reactant pairs affixed to base-4-encoded oligomers would proceed concurrently, we first determined reaction conditions under which both boronic

<sup>a</sup> Department of Chemical Engineering, University of Michigan, Ann Arbor, MI 48109, USA.

<sup>b</sup> Department of Chemical Engineering, Monash University, Clayton, VIC 3800, Australia. E-mail: [timothy.scott@monash.edu](mailto:timothy.scott@monash.edu)

<sup>c</sup> Department of Materials Science and Engineering, Monash University, Clayton, VIC 3800, Australia.

<sup>†</sup> These authors contributed equally to this work

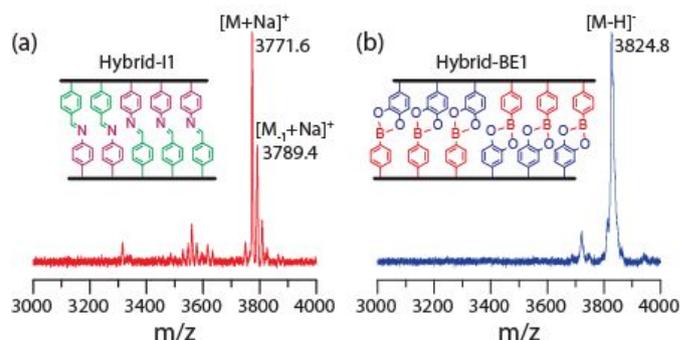
Electronic Supplementary Information (ESI) available: Experimental procedures for monomer and oligomer synthesis, purification, and self-assembly. Small molecule cross-reaction, oligomer, and assembled hybrid characterization. See DOI: 10.1039/x0xx00000x

acid/catechol and amine/aldehyde condensation and rearrangement reactions would mediate the assembly of molecular ladders from binary-encoded strands. Sequences bearing multiple co-reactive functional group types necessitate *in situ* functional group deprotection to preclude premature reaction of the precursors, anticipated to yield cross-linked, intractable material. We recently described an approach for the information-directed assembly of peptoid (i.e., oligo(*N*-substituted glycine)) strands bearing amine and acetal-protected aldehyde pendant groups whereby high concentrations of a Lewis acidic rare earth metal triflate facilitated the simultaneous *in situ* aldehyde deprotection and, by shifting the amine/aldehyde reaction equilibrium, dissociation of strands; subsequent extraction of the Lewis acid enabled the imine-yielding condensation reaction and consequent hybridization of information-bearing ladders to proceed.<sup>17</sup> Other work has described acetonide-protected diol deprotection upon Lewis acid treatment,<sup>18,19</sup> suggesting that the reaction conditions previously employed to deprotect acetal-protected aldehydes could similarly effect *in situ* deprotection of acetonide-protected catechol groups. Nevertheless, attempts to deprotect oligomers bearing acetonide-protected diols using scandium triflate (a strong Lewis acid)<sup>20</sup> proved unsuccessful, where no deprotection was observed for a peptoid bearing three acetonide-protected catechol pendant groups even at high Lewis acid concentrations (up to 0.33 equivalents of scandium triflate per functional group, an amount exceeding that required for quantitative acetal deprotection,<sup>18</sup> see Fig. S2).

This unsuccessful attempt at Lewis acid-mediated *in situ* deprotection prompted the examination of an alternative approach to affect the reaction equilibria of the boronic acid/catechol and amine/aldehyde condensation reactions by acidifying the reaction mixture with trifluoroacetic acid (TFA) to dissociate the oligomeric sequences, then raising the mixture pH by adding a base to effect ladder assembly. Here, peptoid oligomers were synthesized bearing acetal-protected aldehyde and amine pendant groups (denoted as '0' and '1', respectively). Synthesized peptoids bearing 11000 and 00111 sequences were dissolved in acetonitrile and mixed with TFA to simultaneously deprotect the acetal-protected aldehyde and shift the amine/aldehyde condensation reaction equilibrium to favor dissociated strands. Chloroform was then added, the mixture was adjusted to a pH of 14 by addition of aqueous sodium hydroxide and, upon thorough mixing, washing, and phase separation, the organic layer revealed the formation of in-registry 00111×11000 molecular ladder species (Hybrid-I1, Fig. 1a). This approach was similarly employed to afford 10101×01010 molecular ladders from 10101 and 01010 precursor strands (Fig. S3a).

Self-complementary peptoid oligomers bearing boronic acid and catechol pendant groups (denoted as '2' and '3', respectively) were synthesized and also subjected to sequential acidic and alkaline reaction conditions to dissociate and hybridize these sequences. As with the amine/aldehyde system, the boronic acid/catechol condensation reaction equilibrium is sensitive to pH, where the binding constant between

phenylboronic acid and catechol increases from 150 M<sup>-1</sup> at a pH of 6.5 to 3300 M<sup>-1</sup> at pH 8.5.<sup>21</sup> For these oligomers, TFA was employed to effect oligopeptoid cleavage and deprotection, ensuring acidic conditions to prevent premature hybridization. Solutions of peptoid sequences were added to an aqueous alkaline solution with a pH of 9 to yield the desired, in-registry molecular ladders 222333×222333 (Hybrid-BE1, Fig. 1b) and 223233×223233 (Fig. S3b). Out-of-registry species were not identified in the mass spectra of the reaction mixtures examined.



**Figure 1. Dynamic covalent assembly of information-bearing molecular ladders.** Idealized schematics and MALDI-TOF mass spectra of molecular ladders bearing (a) imine and (b) boronate ester rungs. Expected exact masses:  $[M_{\text{Hybrid-I1}} + \text{Na}]^+ = 3768.0$ ;  $[M_{\text{Hybrid-BE1}} - \text{H}]^- = 3816.9$ . For imine-bearing ladders, peaks at multiples of +18 m/z values (e.g., M<sub>n</sub>) are attributable to ladders species with progressively fewer rungs.

Having successfully demonstrated hybridization of binary oligomeric sequences to afford molecular ladders bearing imine- and boronate ester-based rungs by adjusting the reaction mixture pH, the orthogonality of boronic acid/catechol and amine/aldehyde condensation reactions was examined. Here, pairs of model compounds bearing the functional groups employed as dynamic covalent reactant pairs, including catechol, phenylboronic acid, aniline, and benzaldehyde, were mixed in deuterated acetonitrile and the resultant solutions were characterized using <sup>11</sup>B NMR (Fig. 2a) and <sup>1</sup>H NMR (Fig. S4). Of the reaction mixtures examined, <sup>11</sup>B NMR spectrum indicated that phenylboronic acid exclusively reacted with catechol owing to the emergence of a new peak only for the phenylboronic acid/catechol reaction mixture. Similarly, <sup>1</sup>H NMR revealed co-reaction between benzaldehyde and aniline, evidenced by the emergence of a peak at 8.58 ppm attributable to imine generation, while neither compound reacted with phenylboronic acid or catechol (Fig. S4), confirming the orthogonality of these condensation reactions.

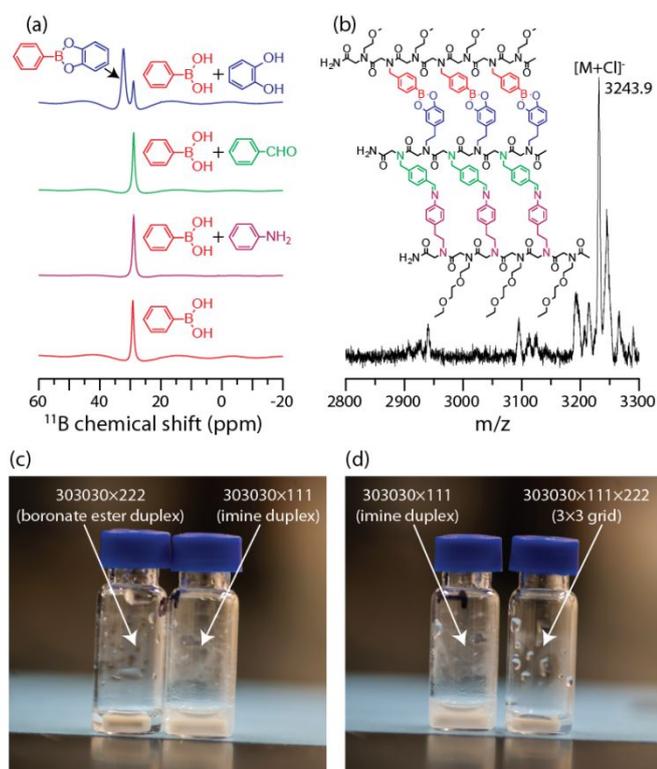
The use of both boronic acid/catechol and amine/aldehyde reactions to mediate the dynamic covalent assembly of base-4-encoded peptoid oligomers was initially explored via the assembly of a molecular grid bearing both imine- and boronate ester-based rungs. A 3×3 molecular grid (i.e., three oligomeric strands linked by three dynamic covalent interactions per oligomer pair) was assembled from a core peptoid oligomer bearing six covalently reactive pendant groups and flanked by two peptoids each bearing three reactive pendant groups that complement the core. By exploiting the 'Z-strand' conformation adopted by peptoids, where the pendant groups of adjacent

residues are presented on opposite sides of the backbone,<sup>22</sup> the precursor oligomers were designed such that the imine- and boronate ester-based rungs of the assembled grid were organized on opposing flanks of the core peptoid sequence. This core strand incorporated alternating catechol and aldehyde pendant groups, whereas the flanking strands were composed of alternating reactive pendant groups, either amine- or boronic acid-based, respectively, and spacer residues to match the reactive pendant group spacing between the flanking and core strands. To begin, stock solutions of the core sequence (303030) and the amine- (111) and boronic acid-bearing (222) flanking strands in acetonitrile/water were each added at a 2:1:1 stoichiometric ratio to a pH 9 aqueous solution containing a catalytic amount of scandium triflate and gently stirred overnight. The resultant, clear hybridization solution was characterized and the target 3×3 molecular grid was identified by mass spectrometry (Fig. 2b), confirming successful dynamic covalent self-assembly using the two orthogonal condensation reactions. Interestingly, whereas the addition of the sequence 303030 to the alkaline aqueous solution yielded a cloudy mixture which persisted upon addition of the sequence 111, the reaction mixture rapidly became clear upon addition of sequence 222 prior to 111, (Fig. 2c). To examine the influence of the sequence addition order on the 3×3 molecular grid fabrication, the process for grid assembly was performed *via* two alternative, sequential assembly routes (see Scheme 2).

of exterior strands and (d) the final 303030×111×222 molecular grid reaction solution by way of a 303030×111 intermediate.

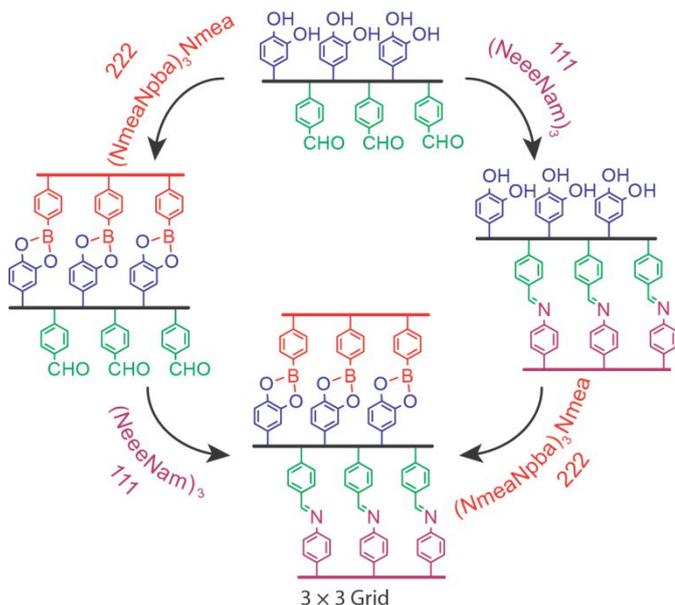
Two alkaline aqueous solutions containing sequence 303030 and either sequence 111 or 222 were allowed to react overnight with a catalytic amount of scandium triflate to form intermediate ladder species. As before, the 303030×111 and 303030×222 reaction mixtures were cloudy and clear solutions, respectively. MALDI-TOF mass spectrometry confirmed the formation of the target 303030×111 and 303030×222 molecular ladders (Fig. S5); however, misaligned or out-of-registry ladder species were also identified in the cloudy 303030×111 mixture, whereas the clear 303030×222 mixture contained exclusively in-registry duplexes. The final molecular grid was generated by addition of the remaining complementary sequences and again mixing overnight. Although no visible change was observed upon addition of the sequence 111 to the clear 303030×222 mixture, addition of sequence 222 to the cloudy 303030×111 reaction mixture yielded a transparent solution (Fig. 2d), suggesting that the boronate ester functionality strongly influences aqueous solubility of the generated structures.

Having realized dynamic covalent assembly employing dual concurrent, orthogonal dynamic covalent interactions, we evaluated the ability of the orthogonal reaction strategy to direct the selective hybridization of information-encoded species. To this end, the oligomeric sequences 213112 and 300203 were synthesized as complementary sequences designed to afford a six-rung molecular ladder. To enable facile identification of assembled species by mass spectrometry, sequence 213112 was synthesized with an additional inert spacer residue prior to acetylation. Overnight incubation of sequences 213112 and 300203 at a 1:1 stoichiometric ratio in alkaline aqueous solution afforded the target molecular ladder (Hybrid-O1), as identified by MALDI-TOF mass spectrometry, with no evidence of non-complementary dimeric (Fig. 3a) or multimeric species (Fig. S6). Notably, parallel orientation of the 213112 and 300203 strands would preclude the generation of an imine and a boronate ester owing to misalignment of the reactive pendant groups; thus, the absence of a peak at +54 m/z (i.e., the sum of +18 and +36 to account for the unreacted imine and boronate ester precursors, respectively) suggests that the assembled molecular ladder consisted of antiparallel-orientated strands. In contrast to the 1:1 stoichiometric system, incubation of sequence 213112 with a 50% excess of 300203 yielded both dimeric and trimeric (consisting of one 213112 strand and two 300203 strands) species as identified by MALDI-TOF mass spectrometry (Fig. S6). To further explore sequence-selective hybridization using concurrent and orthogonal dynamic covalent interactions, a 3×3 molecular grid was designed in which the hybridization of the flanking strands is directed by the information encoded in the core strand. Here, again owing to the  $\Sigma$ -strand conformation, the spacer-free oligo(peptoid) sequence 123321 displays the sequence 132 on opposite sides of the backbone. Hybridization of this core strand with two stoichiometric equivalents of the sequence 320 afforded the base-4, 3×3 molecular grid, Hybrid-G1. Characterization of the reaction solution by MALDI-TOF mass



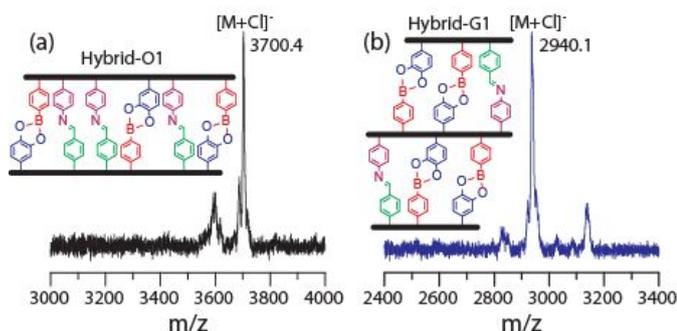
**Figure 2: Orthogonality between imine and boronate ester systems.** (a) <sup>11</sup>B-NMR spectra highlighting the specificity of boronic acids to diols. (b) Structure and MALDI-TOF spectra of a 3×3 molecular grid assembled from the concurrent addition of a core peptoid strand with aldehyde and catechol functionality and complementary, exterior sequences bearing amine- and boronic acid-based pendant groups. Expected exact mass: [M+Cl]<sup>-</sup> = 3235.5. Digital image capturing the alternative pathways for a 3×3 molecular grid *via* (c) the sequential addition

spectrometry confirmed the formation of the molecular grid (Fig.3b). Unfortunately, a non-negligible amount of the cross-product  $123321 \times 123321 \times 320$  ( $[M]=3160.4$  m/z) was also observed, potentially attributable to sequence brevity and consequent lack of mismatched dynamic covalent reactant pairs.



**Scheme 2.** Alternate pathways for the assembly of a base-4 molecular grid via the sequential addition of oligomeric sequences bearing boronic acid or amine pendant groups.

In conclusion, we have demonstrated the information-directed assembly of molecular ladders and grids bearing covalent bond-based rungs from encoded precursor strands using dual concurrent, orthogonal dynamic covalent interactions. We anticipate that the information density offered by this system to have broad utility in the development of complex and robust self-assembled nanostructures.



**Figure 3.** Base-4 molecular assembly. Idealized schematics and MALDI-TOF spectrum confirming the formation of (a) the molecular ladder, Hybrid-O1, and (b) the  $3 \times 3$  molecular grid, Hybrid-G1. Expected exact masses:  $[M_{\text{Hybrid-O1}} + \text{Cl}]^- = 3692.7$ ;  $[M_{\text{Hybrid-G1}} + \text{Cl}]^- = 2961.4$ .

This work was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under Award # DE-SC0012479. SCL and MFD would like to acknowledge the National Science Foundation Graduate Research Fellowship

Program. MFD would also like to acknowledge financial support from the Ford Foundation, through its fellowship program, and the National Academies of Sciences, Engineering, and Medicine, the administrative agency for the fellowship program.

## References

- 1 F. Crick and J. Watson, *Nature*, 1953, **171**, 737.
- 2 N. C. Seeman, *J. Theor. Biol.*, 1982, **99**, 237–247.
- 3 P. W. K. Rothemund, *Nature*, 2006, **440**, 297–302.
- 4 B. Wei, M. Dai and P. Yin, *Nature*, 2012, **485**, 623–626.
- 5 M. Santosh and P. K. Maiti, *J. Phys. Condens. Matter.*, 2009, **21**, 101–139.
- 6 U. Bockelmann, B. Essevaz-Roulet and F. Heslot, *Phys. Rev. Lett.*, 1997, **79**, 4489–4492.
- 7 C. S. Hartley, E. L. Elliott and J. S. Moore, *J. Am. Chem. Soc.*, 2007, **129**, 4512–4513.
- 8 T. Wei, J. C. Furgal, J. H. Jung and T. F. Scott, *Polym. Chem.*, 2017, **8**, 520–527.
- 9 J. F. Reuther, J. L. Dees, I. V. Kolesnichenko, E. T. Hernandez, D. V. Ukraintsev, R. Guduru, M. Whiteley and E. V. Anslyn, *Nat. Chem.*, 2018, **10**, 45–50.
- 10 H. M. Seifert, K. Ramirez Trejo and E. V. Anslyn, *J. Am. Chem. Soc.*, 2016, **138**, 10916–10924.
- 11 M. E. Bracchi and D. A. Fulton, *Chem. Commun.*, 2015, **51**, 11052–11055.
- 12 Z. Rodriguez-Docampo and S. Otto, *Chem. Commun.*, 2008, **42**, 5301–5303.
- 13 M. Hutin, G. Bernardinelli and J. R. Nitschke, *Chem. Eur. J.*, 2008, **14**, 4585–4593.
- 14 N. Christinat, R. Scopelliti and K. Severin, *J. Org. Chem.*, 2007, **72**, 2192–2200.
- 15 N. Christinat, R. Scopelliti and K. Severin, *Angew. Chem. Int. Ed.*, 2008, **47**, 1848–1852.
- 16 M. F. Dunn, T. Wei, R. N. Zuckermann and T. F. Scott, *Polym. Chem.*, 2019, **10**, 2337–2343.
- 17 S. C. Leguizamon and T. F. Scott, *Nat. Commun.*, 2020, **11**, 784.
- 18 T. Wei, J. C. Furgal and T. F. Scott, *Chem. Commun.*, 2017, **53**, 3874–3877.
- 19 J. S. Yadav, B. V. S. Reddy and K. S. Reddy, *Chem. Lett.*, **30** 2001, 430–431.
- 20 S. Kobayashi, M. Sugiura, H. Kitagawa and W. W. L. Lam, *Chem. Rev.*, 2002, **102**, 2227–2302.
- 21 J. Yan, G. Springsteen, S. Deeter and B. Wang, *Tetrahedron*, 2004, **60**, 11205–11209.
- 22 J. R. Edison, R. K. Spencer, G. L. Butterfoss, B. C. Hudson, A. I. Hochbaum, A. K. Paravastu, R. N. Zuckermann and S. Whitelam, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, 5647–5651.

Molecular ladders and molecular grids were self-assembled employing two orthogonal dynamic covalent reactant pairs in an aqueous environment.

