Supplementary Information (SI) for Physical Chemistry Chemical Physics.

Supplementary Information

Unveiling the Role of Non-Structural Supramolecular Water Assisted Coupling of Pyrroles and Isocyanates to Amidopyrroles

Sanat Kumar Mahapatra,[a] Bijoy Ghosh,[b] and Lisa Roy*[b]

---

[a] *Institute of Chemical Technology Mumbai – IOC Odisha Campus Bhubaneswar, Bhubaneswar 751013, India.*

[b] *Department of Education, Indian Institute of Technology Kharagpur, Kharagpur 721302, India. Email: L.Roy@edu.iitkgp.ac.in*

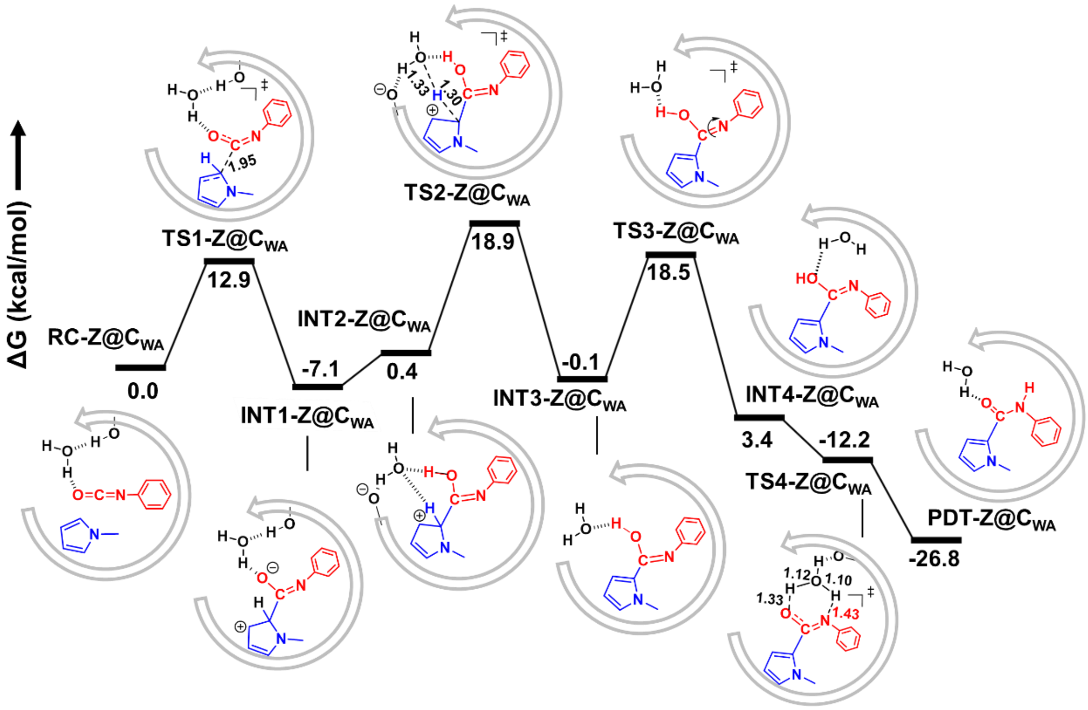


Figure S1. Gibbs free energy profiles in kcal/mol at B3LYP-D3BJ/Def2-TZVP/CPCM($CHCl_3$:H_2O=1:1) for the coupling of **2a** and **3a** to *Z*-selective **4aa** inside capsule (**C**) with water-assistance (**C**$_{\mathbf{WA}}$). Notably, the pyrrolic C-H activation via **TS2-Z@C**$_{\mathbf{WA}}$ featuring an activation free energy of ~ 26.0 kcal/mol is non-competitive with the other pathways with monomer and capsule assistance, in presence or absence of water, and has not been discussed in the main text.

Table S1. Binding energies of different reactant complexes calculated at B3LYP-D3BJ/Def2-TZVP/CPCM(Chloroform).

Reactant Complex	Binding Energy in kcal/mol
RC@S$_{\mathbf{U}}$	-3.4
RC@S$_{\mathbf{WA}}$	-9.3
RC@M$_{\mathbf{U}}$	-21.0
RC@M$_{\mathbf{WA}}$	-21.2
RC@C$_{\mathbf{WA}}$	-12.2
RC@C$_{\mathbf{U}}$	-19.4

Reactant Complex[a]	Binding Energy in kcal/mol
RC@C$_{\mathbf{WA\_1}}$	-46.2
RC@C$_{\mathbf{2WA\_1}}$	-58.0
RC@C$_{\mathbf{3WA\_1}}$	-77.1
RC@C$_{\mathbf{U\_1}}$	-42.0

[a] Computed at B3LYP-D3BJ/Def2-TZVP/ALPB($CHCl_3$) on ONIOM[B3LYP-D3(BJ)/ALPB($CHCl_3$)/def2-SVP: XTB] optimized geometries with ORCA 5.0.3 software.[1]

Table S2. Calculated relative free energy (ΔG) for all species with different functionals in mixed solvent.

Entry	**Stationary State**	**M062X/Def2-TZVP/CPCM ($CHCl_3$:H_2O=1:1)**	**B3LYP-D3BJ/Def2-TZVP/CPCM ($CHCl_3$:H_2O=1:1)**	**B3LYP/Def2-TZVP/CPCM ($CHCl_3$:H_2O=1:1)**
1	**Separated Reactants**	0.0	0.0	0.0
	RC@S_U	3.1	1.1	3.1
	TS1@S_U	27.1	22.5	26.7
	INT1@S_U	26.1	22.1	25.9
	TS2@S_U	47.8	41.5	27.3
	PDT@S_U	-10.5	-12.3	-8.9
2	**Separated Reactants**	0.0	0.0	0.0
	RC@S_{WA}	9.9	1.0	9.9
	TS1@S_{WA}	32.2	21.4	34.2
	INT1@S_{WA}	27.5	17.6	30.1
	TS2@S_{WA}	38.4	26.0	36.5
	PDT@S_{WA}	-4.0	-13.1	-5.2
3	**Separated Reactants**	0.0	0.0	0.0
	RC@M_U	-5.1	-6.7	32.7
	TS1@M_U	14.7	13.2	38.2
	INT1@M_U	5.5	-4.1	28.5
	TS2@M_U	11.1	1.2	29.9
	PDT@M_U	-18.7	-20.7	11.0
4	**Separated Reactants**	0.0	0.0	0.0
	RC@M_{WA}	-5.4	-9.0	31.0
	TS1@M_{WA}	8.8	5.2	45.7
	INT1@M_{WA}	-0.6	-7.9	37.7
	TS2@M_{WA}	7.9	-0.4	37.1
	PDT@M_{WA}	-25.5	-29.5	12.0
5	**RC@C_U**	0.0	0.0	0.0
	TS1@C_U	20.0	20.1	19.6
	INT1@C_U	3.7	-1.9	-4.3

	TS2@C$_{U}$	12.2	5.0	3.9
	PDT@C$_{U}$	-16.6	-14.3	-18.4
6	**RC@C$_{WA}$**	0.0	0.0	0.0
	TS1@C$_{WA}$	18.3	15.2	16.9
	INT1@C$_{WA}$	0.7	-6.1	-6.0
	TS2@C$_{WA}$	3.3	-3.3	-3.4
	PDT@C$_{WA}$	-33.6	-32.9	-29.0
7	**RC-Z@C$_{WA}$**	0.0	0.0	0.0
	TS1-Z@C$_{WA}$	16.2	13.3	12.3
	INT1-Z@C$_{WA}$	-2.9	-4.8	-6.8
	INT2-Z@C$_{WA}$	2.3	2.0	3.9
	TS2-Z@C$_{WA}$	21.9	21.2	17.8
	INT3-Z@C$_{WA}$	-1.7	0.3	-17.0
	TS3-Z@C$_{WA}$	17.2	18.4	0.1
	INT4-Z@C$_{WA}$	-7.6	-3.2	-14.6
	TS4-Z@C$_{WA}$	-14.0	-11.7	-20.8
	PDT-Z@C$_{WA}$	-26.2	-26.9	-36.7

Table S3. Calculated relative free energy (ΔG) for all species with different functionals in chloroform

Entry	**Stationary State**	**M062X/Def2-TZVP/CPCM (CHCl$_3$)**	**B3LYP-D3BJ/Def2-TZVP/CPCM (CHCl$_3$)**	**B3LYP/Def2-TZVP/CPCM (CHCl$_3$)**
1	**RC@S$_{U}$**	0.0	0.0	0.0
	TS1@S$_{U}$	23.3	22.9	23.0
	INT1@S$_{U}$	23.2	22.4	22.8
	TS2@S$_{U}$	43.9	41.3	40.3
	PDT@S$_{U}$	-15.3	-13.2	-13.7
2	**RC@S$_{WA}$**	0.0	0.0	0.0
	TS1@S$_{WA}$	16.1	15.8	18.1
	INT1@S$_{WA}$	13.1	13.6	15.4
	TS2@S$_{WA}$	23.1	21.4	21.1
	PDT@S$_{WA}$	-18.9	-18.0	-20.1
3	**RC@M$_{U}$**	0.0	0.0	0.0

	TS1@M$_{U}$	15.9	15.4	16.6
	INT1@M$_{U}$	8.7	2.0	11.5
	TS2@M$_{U}$	13.4	6.4	-0.6
	PDT@M$_{U}$	-17.3	-16.3	-20.3
4	**RC@M$_{WA}$**	0.0	0.0	0.0
	TS1@M$_{WA}$	10.7	9.3	11.1
	INT1@M$_{WA}$	3.5	-0.3	5.3
	TS2@M$_{WA}$	10.7	5.9	3.3
	PDT@M$_{WA}$	-23.9	-24.3	-11.3
5	**RC@C$_{U}$**	0.0	0.0	0.0
	TS1@C$_{U}$	21.4	19.0	21.6
	INT1@C$_{U}$	6.3	0.5	-1.8
	TS2@C$_{U}$	13.6	6.5	5.3
	PDT@C$_{U}$	-15.1	-12.8	-16.8
6	**RC@C$_{WA}$**	0.0	0.0	0.0
	TS1@C$_{WA}$	18.9	15.7	17.5
	INT1@C$_{WA}$	0.7	-5.5	-5.4
	TS2@C$_{WA}$	3.9	-2.6	-2.7
	PDT@C$_{WA}$	-32.9	-32.2	-28.3
7	**RC-Z@C$_{WA}$**	0.0	0.0	0.0
	TS1-Z@C$_{WA}$	16.8	13.7	15.9
	INT1-Z@C$_{WA}$	-0.1	-2.2	-4.1
	INT2-Z@C$_{WA}$	4.4	4.0	6.0
	TS2-Z@C$_{WA}$	23.7	22.9	23.3
	INT3-Z@C$_{WA}$	-0.5	1.3	-16.0
	TS3-Z@C$_{WA}$	18.2	19.2	1.4
	INT4-Z@C$_{WA}$	-7.4	-3.2	-14.6
	TS4-Z@C$_{WA}$	-13.5	-11.4	-20.4
	PDT-Z@C$_{WA}$	-25.2	-26.1	-36.9

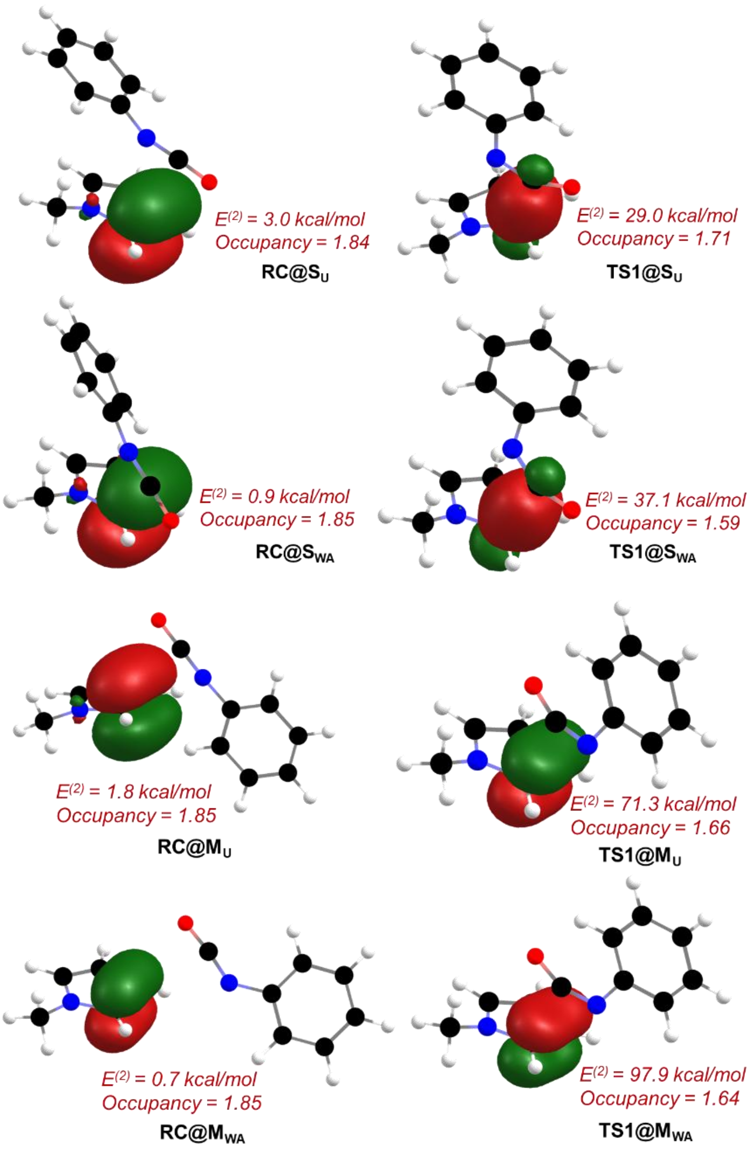
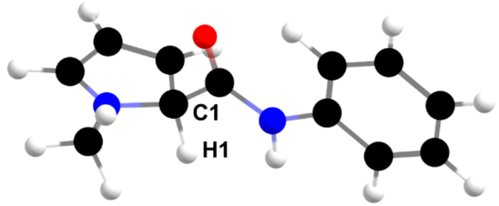
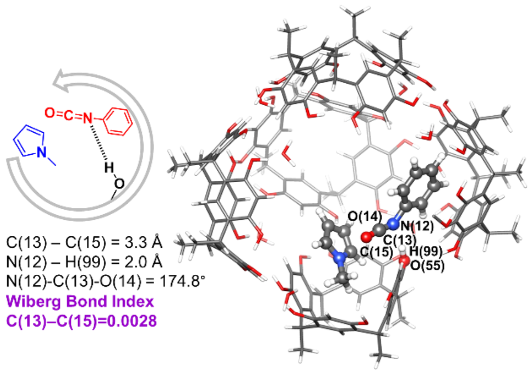


Figure S2: Plots of the donor natural bond orbitals in fragments of the reactant complexes and transition states for the coupling of **2a** and **3a** to **4aa** in implicit solvent (S) and assisted by a resorcin-arene monomer (M). Second order perturbation theory energies and orbital-occupancy at B3LYP/D3(BJ)/6-31G(d,p) level of theory are shown.

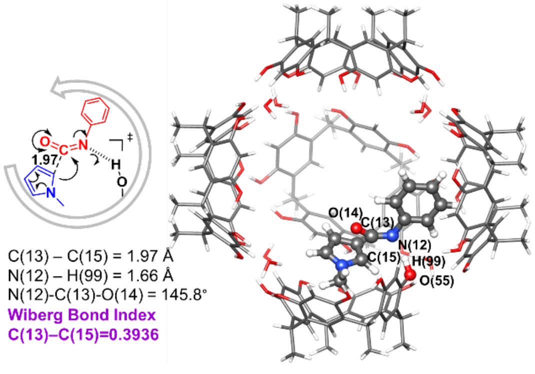
Table S4: Calculated electrophilicity (ω^{+}) in fragments of the intermediate in different chemical environment.

<table>
<tr><td colspan="2">
</td></tr>
<tr><td>Stationary State</td><td>Electrophilicity (ω^{+}) on C1-H1 unit in eV</td></tr>
<tr><td>INT1@M$_{U}$</td><td>14.8</td></tr>
<tr><td>INT1@M$_{WA}$</td><td>15.0</td></tr>
<tr><td>INT1@C$_{U}$</td><td>13.7</td></tr>
<tr><td>INT1@C$_{WA}$</td><td>15.1</td></tr>
</table>



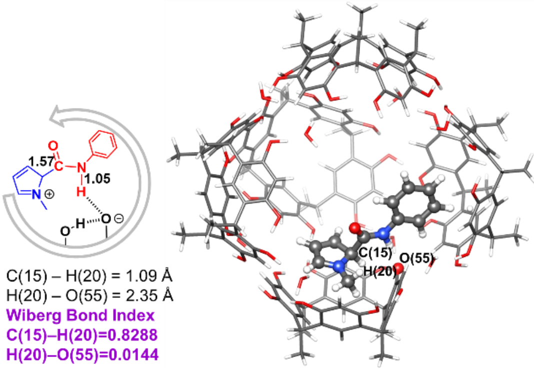
C(13) – C(15) = 3.3 Å
N(12) – H(99) = 2.0 Å
N(12)-C(13)-O(14) = 174.8°
Wiberg Bond Index
C(13)–C(15)=0.0028

RC@C$_\mathrm{U}$



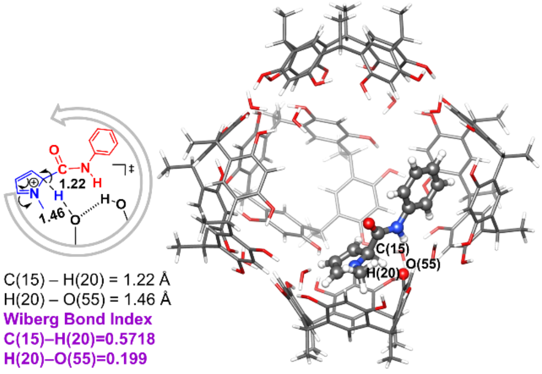
C(13) – C(15) = 1.97 Å
N(12) – H(99) = 1.66 Å
N(12)-C(13)-O(14) = 145.8°
Wiberg Bond Index
C(13)–C(15)=0.3936

TS1@C$_\mathrm{U}$



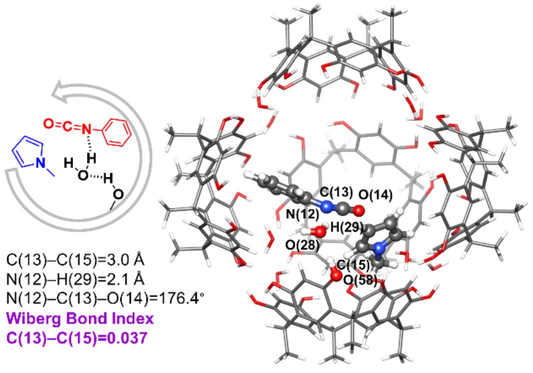
C(15) – H(20) = 1.09 Å
H(20) – O(55) = 2.35 Å
Wiberg Bond Index
C(15)–H(20)=0.8288
H(20)–O(55)=0.0144

INT1@C$_\mathrm{U}$



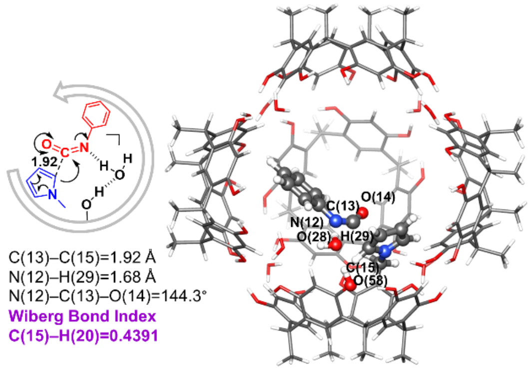
C(15) – H(20) = 1.22 Å
H(20) – O(55) = 1.46 Å
Wiberg Bond Index
C(15)–H(20)=0.5718
H(20)–O(55)=0.199

TS2@C$_\mathrm{U}$



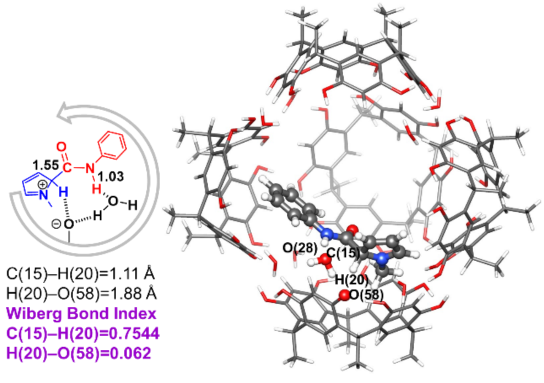
C(13)–C(15)=3.0 Å
N(12)–H(29)=2.1 Å
N(12)–C(13)–O(14)=176.4°
Wiberg Bond Index
C(13)–C(15)=0.037

RC@C$_\mathrm{WA}$



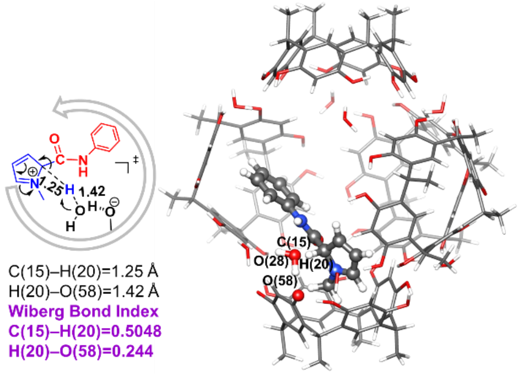
C(13)–C(15)=1.92 Å
N(12)–H(29)=1.68 Å
N(12)–C(13)–O(14)=144.3°
Wiberg Bond Index
C(15)–H(20)=0.4391

TS1@C$_\mathrm{WA}$



C(15)–H(20)=1.11 Å
H(20)–O(58)=1.88 Å
Wiberg Bond Index
C(15)–H(20)=0.7544
H(20)–O(58)=0.062

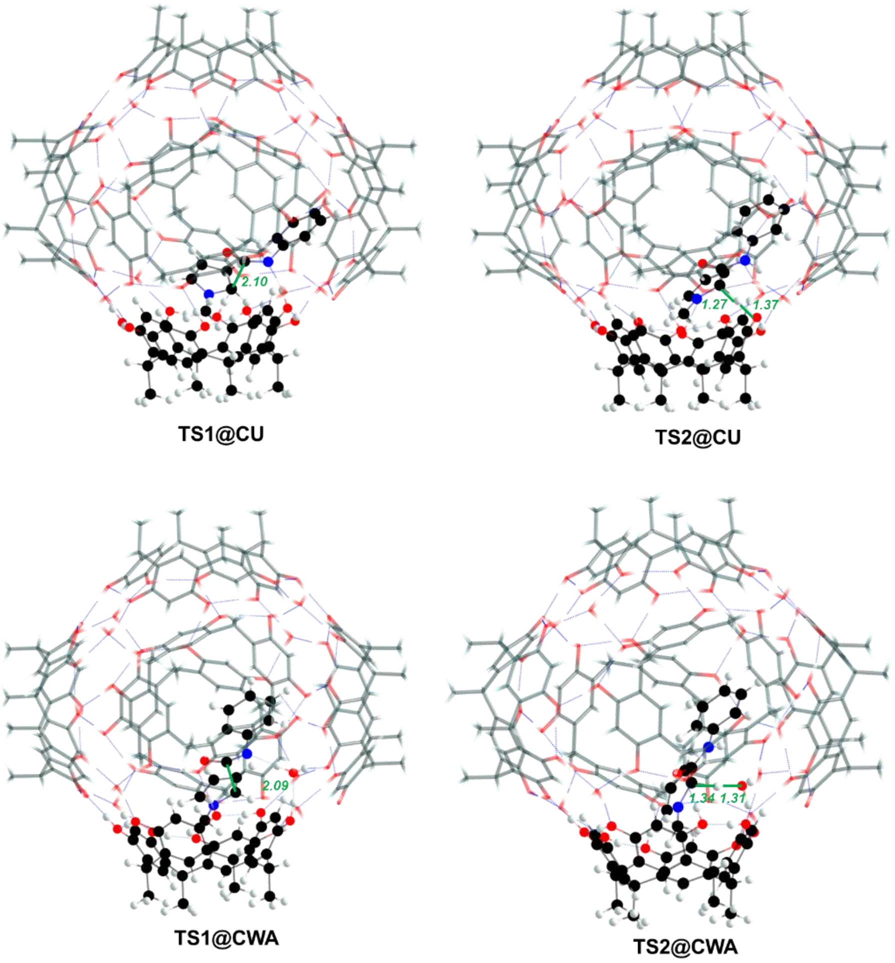
INT1@C$_\mathrm{WA}$



C(15)–H(20)=1.25 Å
H(20)–O(58)=1.42 Å
Wiberg Bond Index
C(15)–H(20)=0.5048
H(20)–O(58)=0.244

TS2@C$_\mathrm{WA}$

Figure S3: Optimized geometries and calculated Wiberg Bond Indices of stationary states, in presence or absence of additional water. One of the resorcin-arene monomers are not shown for clarity.


TS1@CU

TS2@CU

TS1@CWA

TS2@CWA

Figure S4. 2-layer ONIOM method used for optimizations at M06-2X/6-31G(d,p)//PM6 level of theory. Ball and stick represent High-Level and Thin sticks represents Low-Level.

Table S5. S-value Test (Energies are in units of Hartree).

S. No.		E^{PM6}	$E^{\mathrm{M06\text{-}2X/6\text{-}31(d,p)}}$	$S_{\mathrm{low}} = E_{\mathrm{low,real}} - E_{\mathrm{low,model}}$	$S_{\mathrm{high}} = E_{\mathrm{high,real}} - E_{\mathrm{high,model}}$	Err = S_{low}-S_{high}
1	RC@C_U^{real}	-3.8598136	-12299.20774	0.010508326	0.00735259	0.003155736
	RC@C_U^{model}	-0.4200407	-2488.77442			
	INT1@C_U^{real}	-3.8345726	-12299.18988			
	INT1@C_U^{model}	-0.405308	-2488.763914			
2	RC@C_{WA}^{real}	-3.9651628	-12375.60123	-0.004254494	-0.01244567	0.008191176
	RC@C_{WA}^{model}	-0.5177151	-2565.169098			
	INT1@C_{WA}^{real}	-3.9616831	-12375.60378			
	INT1@C_U^{model}	-0.5099808	-2565.159201			

The S-value test is a crucial diagnostic feature to demonstrate the accuracy of the chosen methodology in a two-level ONIOM2 calculation based on integrated molecular orbital + molecular orbital hypothesis, i.e. whether the two levels correctly describe the electronic effect or the “substituent effect” of the large system. It describes the amount of error (Err) incurred in avoiding the high-level calculation for the real system.[2-3] Here the following partitioning is utilized:

$\mathrm{Err} = E_{\mathrm{ONIOM2(high:low)}} - E_{\mathrm{high,real}}$ (1)

$\mathrm{Err} = (E_{\mathrm{high,model}} + E_{\mathrm{low,real}} - E_{\mathrm{low,model}}) - E_{\mathrm{high,real}}$

$= (E_{\mathrm{low,real}} - E_{\mathrm{low,model}}) - (E_{\mathrm{high,real}} - E_{\mathrm{high,model}})$

The S-values (substituent-values) for high- and low-level methodology are then defined as follows

$S_{\mathrm{low}} = E_{\mathrm{low,real}} - E_{\mathrm{low,model}}$

$S_{\mathrm{high}} = E_{\mathrm{high,real}} - E_{\mathrm{high,model}}$

Thus, if S_{low} is equal to S_{high}, the ONIOM2 approximation with the QM2 (low-level method) has no error. ΔS-value or Err-value has been calculated for two cases: for the C-C coupling within capsules with and without assistance of non-structural supramolecular water molecule. As observed from Table S5, the Err-value is ~0 for both the cases suggesting that the low-level method PM6 can be appropriately used in combination with the present high-level method, M06-2X/6-31(d,p) for geometry optimizations. This strongly confirms the reliability of the chosen high- and low-level methodology in our study.

Table S6. Reaction energies for optimization in bulk solvent. Energies reported are in kcal/mol.

	Δ*G* (B3LYP-D3BJ/def2-TZVP/CPCM(Chloroform:H_2O=1:1))	Δ*G* (B3LYP-D3BJ/Def2-TZVP/CPCM($CHCl_3$:H_2O=1:1)//M06-2X/CPCM($CHCl_3$)/6-31G(d,p))
RC@S_U	1.1	1.1
TS1@S_U	22.2	22.5
INT1@S_U	22.1	22.1
TS2@S_U	41.6	41.5
PDT@S_U	-12.9	-12.3
RC@S_{WA}	-1.4	1.0
TS1@S_{WA}	19.6	21.4
INT1@S_{WA}	15.9	17.6
TS2@S_{WA}	24.8	26.0
PDT@S_{WA}	-14.5	-13.1

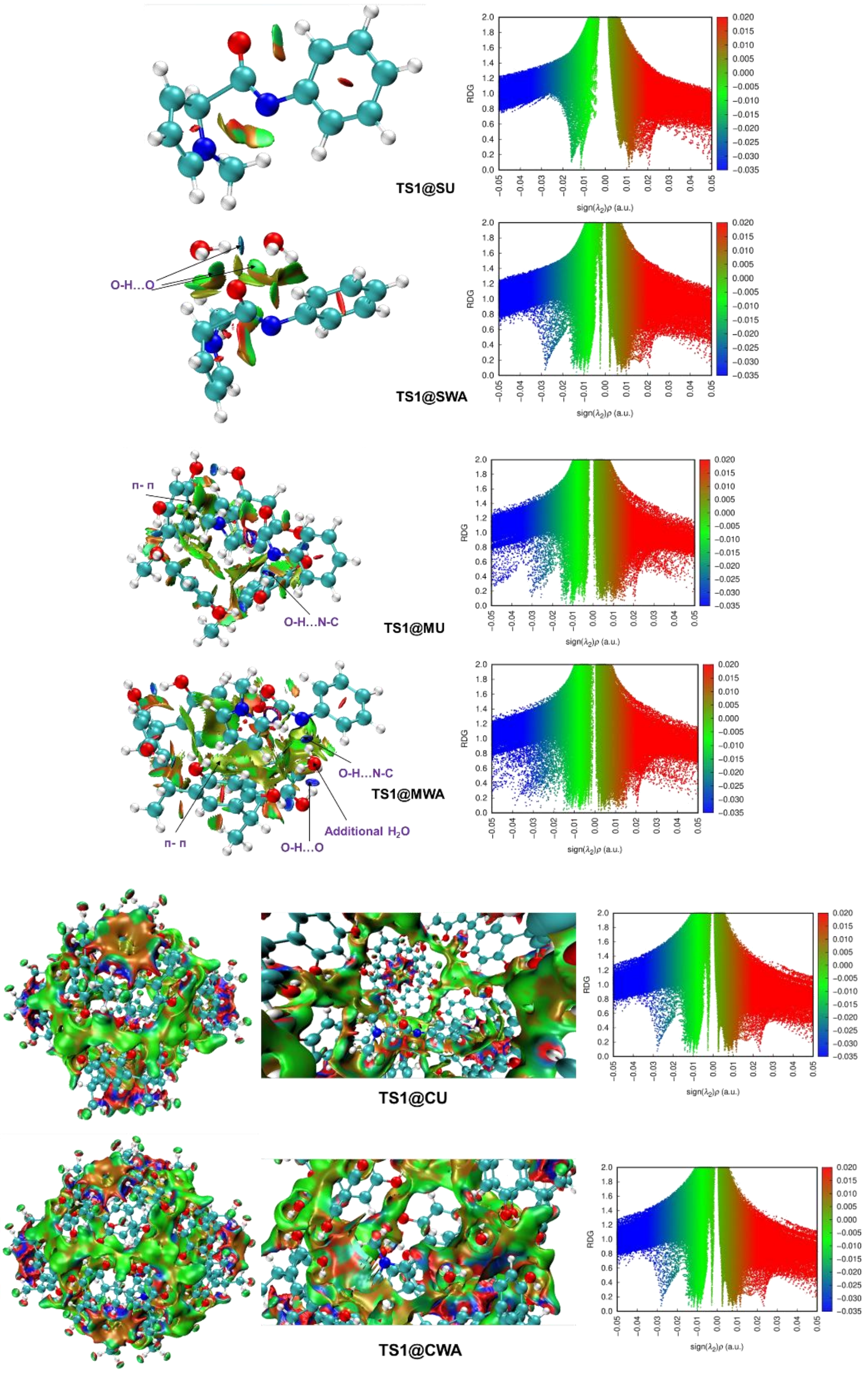


Figure S5. NCIs and RDG plots in transition states for C-C coupling.

References

1. F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2022, **12**, e1606.

2. L. W. Chung, W. M. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. Ke, F. Liu, H.-B. Li, L. Ding and K. Morokuma, *Chem. Rev.* 2015, **115**, 5678.

3. T. Vreven, K. Morokuma, *J. Comput. Chem.* 2000, **21**, 1419