Supporting Information
Searching the chemical space of hetero-atom bridged NBDs

Nils Oberhof, Andreas Erbs Hillers-Bendtsen, Oscar Berlin Obel, Karoline Schjelde, Kurt V. Mikkelsen, and Andreas Dreuw

1 Molecule Generation

Looping over the 14 substituents as combinations with replacement (i.e. that every substituent can appear up to three times at up to three substituents), looped over the 6 substitution sites as combinations (meaning, each can only appear once per combination), gives $14 * 6 + 105 * 15 + 560 * 20 = 84 + 1575 + 11200 = 12859$ possible combinations over all. One then has to account for "double counting" of structures, since the positions B1&B5 as well as B2&B4 can be equivalent due to the symmetry of the phenyl ring. This leads to two sets of two times substituted combinations ($2 * 105 = 210$) to be excluded, where B1 is equivalent to B5 and where B2 is equivalent to B4. Additionally one set of thrice substituted combinations ($1 * 560 = 560$), where B1&B2 are equivalently substituted as B4&B5 in the other set, is excluded. Therefore, one arrives at $12859 - 210 - 560 = 12089$ molecules to be calculated. In the workflow this correction is automatically done by taking all generated molecules, translating them into canonical smiles strings (which are the same for rotamers) and collecting them in one set, which avoids doubly integrating the same smiles string in the final set of molecules. This way the overall amount of molecules in the large data set comes out to be 12089 per bridge head.

2 Screening Methodology

The screening of the 37,662 compounds within the selected chemical space was done with a previously developed screening procedure.[1–3] Within this workflow extended tight binding (xTB) methods[4,5] are employed to calculate storage energies (ΔE_S), thermal back-reaction barriers (ΔE_{tbr}) and the first vertical singlet excitation energy with oscillator strength above 0.01 for the reactant and product molecule. The procedure is depicted in Figure 1 and described in the following. At first, a SMILES string of the reactant system is constructed and the corresponding photoproduct SMILES string is generated using a reaction SMARTS string. Next a conformer search is performed for both using the ETKDG[6] method in RDkit[7] and afterwards a geometry optimization of the lowest energy conformers is done with GFN2-xTB[8]. Subsequently the transition state is approximated as the highest energy point in a 20 step scan of the symmetrical opening of the C-C

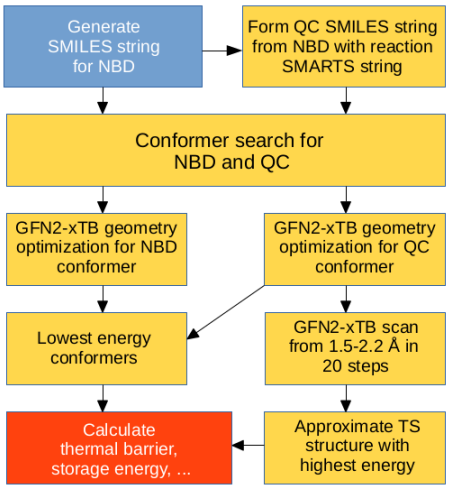

Figure 1: Flowchart of the screening procedure.

bonds broken in the thermal back-reaction, between 1.5 and 2.2 Å. This results in the thermal back-reaction barrier being an upper limit, since (e.g. for some quadricyclane systems) the actual transition state opens both bonds asymmetrically. This procedure is chosen to stay computationally efficient, since other methods, such as Nudged Elastic Band[9] or Reaction Path Methods also implemented in the xTB program[4], would be computationally more demanding. From this data the storage energy (ΔE_S) and thermal back-reaction barrier (ΔE_{tbr}) can be extracted. Finally, sTDA-xTB[5] is employed to calculate the optical properties of the reactant and product. The applicability of this screening procedure has been illustrated in great detail with benchmarking against density functional theory in previous publications.[1–3]

3 Solar Conversion Efficiency

To be able to score the individual compounds from the screening procedure consistently, the solar power conversion efficiency (SCE) is used.[10,11] Within the SCE the thermochemical as well as the photochemical properties are combined to a single value that then can be used to compare the individual compounds against each other and give insight into their applicability as MOST systems. The SCE can be given as

$$SCE(p) = \Delta E_S \frac{\int_{\Delta E_{cut}}^{\infty} \frac{P_{sun}(\omega)}{\hbar\omega}[1 - 10^{(\epsilon_R(\omega)c_R + \epsilon_P(\omega)c_P L)}]\frac{\epsilon_R(\omega)c_R}{\epsilon_R(\omega)c_R + \epsilon_P(\omega)c_P} d\omega}{\int_0^{\infty} P_{sun}(\omega) d\omega}. \quad (1)$$

In this p gives the conversion rate between the reactant (R) and the product (P) giving a concentration of each as $c_R = (1-p)c$ and $c_P = pc$ with c being the total concentration. $\epsilon_R(\omega)$ and $\epsilon_P(\omega)$ are the extinction coefficients of reactant and product at the respective frequency ω. ΔE_S is the storage energy of the system, which also contributes to the cut-off energy ΔE_{cut}, which is given by $\Delta E_{cut} = \Delta E_S + \Delta E_{tbr}$. $P_{sun}(\omega)$ is the reference spectrum according to the AM1.5G standard spectrum over which is integrated, where the denominator serves as normalization. The factor $[1-10^{(\epsilon_R(\omega)c_R+\epsilon_P(\omega)c_P L)}]$ gives the attenuation of sunlight at frequency ω, while $\frac{\epsilon_R(\omega)c_R}{\epsilon_R(\omega)c_R+\epsilon_P(\omega)c_P}$ is a scaling factor originating from the interconversion of reactant and product and their competing absorption. E_{cut} makes the system fulfill the requirement of an absorbed photon having enough energy to uphold a full conversion cycle. The extinction coefficients $\epsilon(\omega)$ are calculated from the vertical excitation energies ω_k and the corresponding oscillator strengths f_k from the computations via convolution with a Gaussian function $g(\omega_k, \sigma)$ as

$$\epsilon(\omega) = \frac{N_A}{ln(10)} \frac{\pi e^2}{2\epsilon_0 m_e c} \sum_k f_k g(\omega_k, \sigma) \quad (2)$$

where σ is the full width half maximum (FWHM) of the absorption. Here we employ a total concentration of $1M$, a conversion ratio p of 50 %, and a FWHM of 0.25 eV to evaluate the SCE of each screened system.

4 Contributing metrics per hetero-atom bridge

For a comprehensive understanding of the correlation between the contributing properties and their correlation with the SCE, scatter plots of each parameter against the resulting SCE have been generated for each bridge system. They are shown in the following. The general trends between the bridges are discussed in the main manuscript.

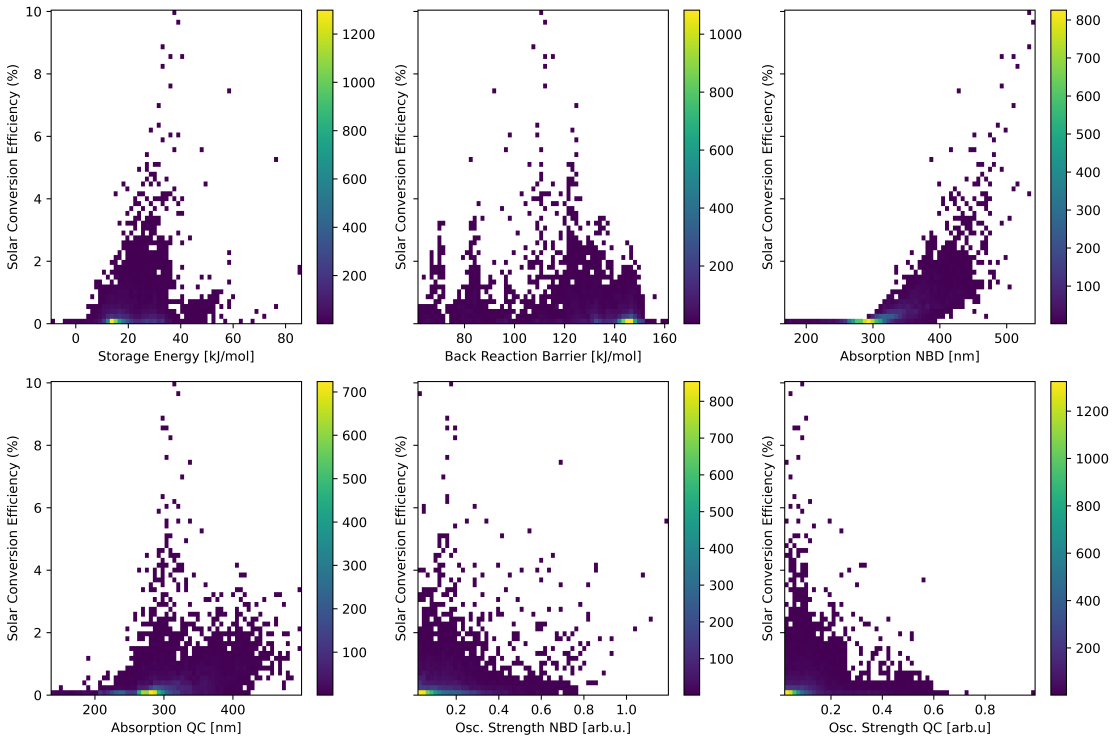

Figure 2: Scatter plots of the SCE calculated according to eq.1 for the -CH_2- bridged systems. The contributing metrics storage energy (ΔE_S, top left), thermal back-reaction barrier (ΔE_{tbr}, top right), the absorption wavelengths and the oscillator strengths of the lowest-lying singlet excited state with an oscillator strength above 0.01 of the NBD (middle) and QC (bottom) isomers are explored.

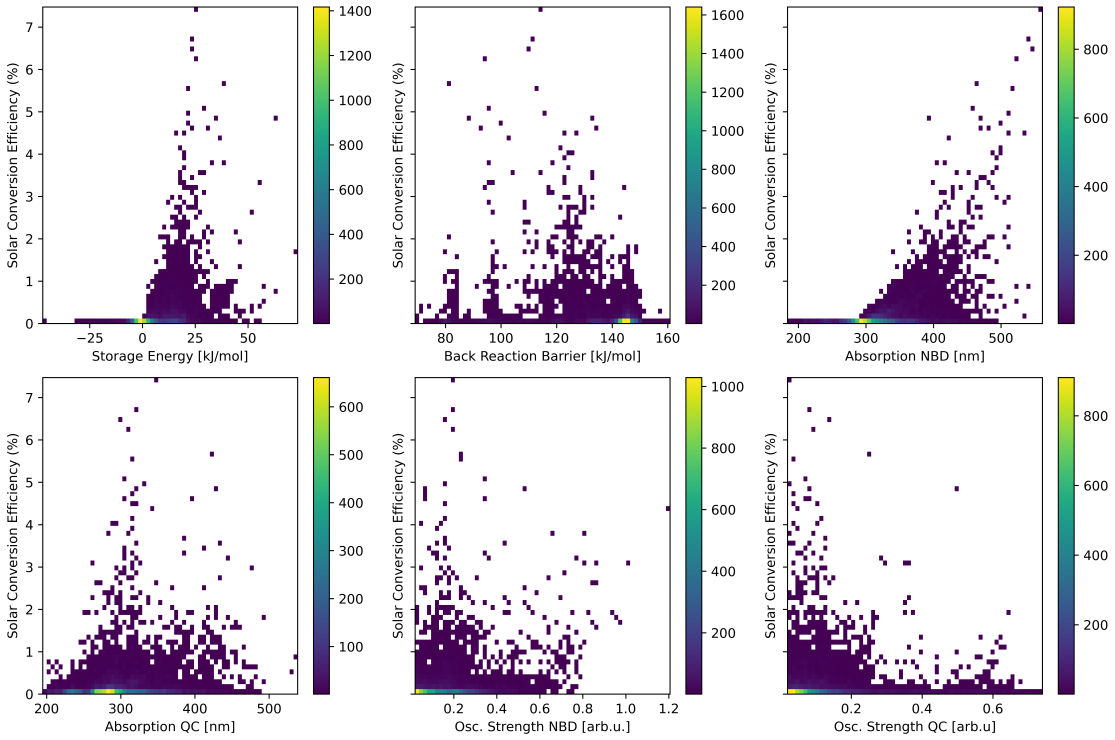

Figure 3: Scatter plots of the SCE calculated according to eq.1 for the -$NC(O)OMe$- bridged systems. The contributing metrics storage energy (ΔE_S, top left), thermal back-reaction barrier (ΔE_{tbr}, top right), the absorption wavelengths and the oscillator strengths of the lowest-lying singlet excited state with an oscillator strength above 0.01 of the NBD (middle) and QC (bottom) isomers are explored.

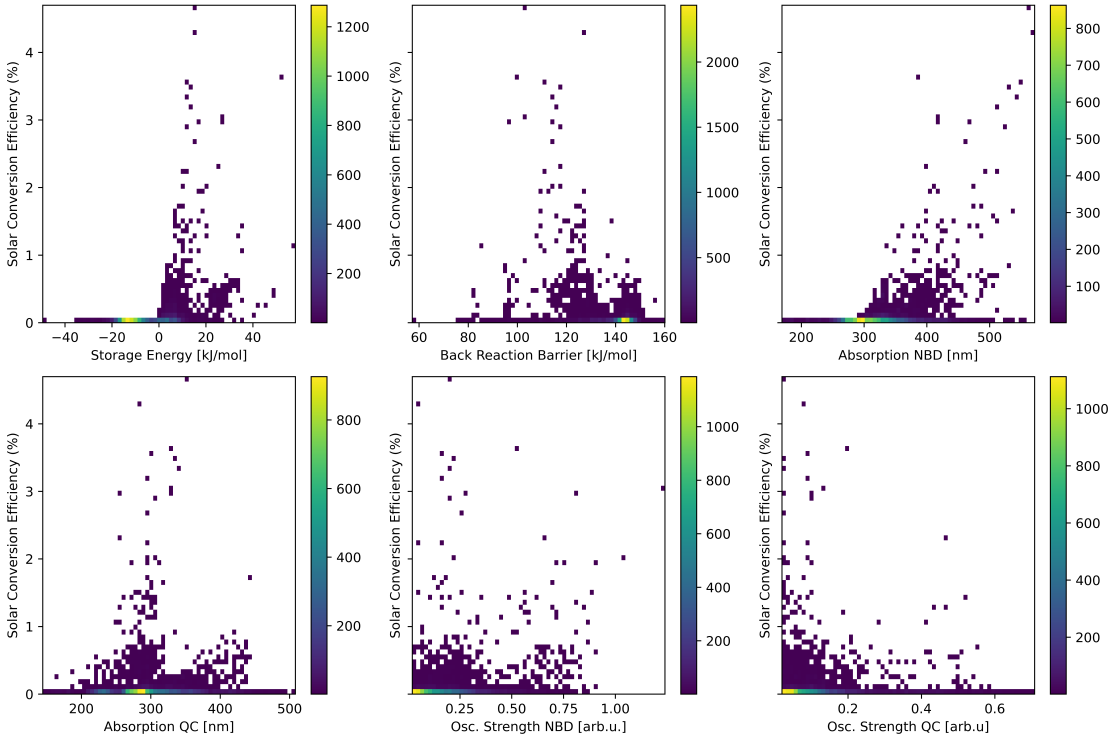

Figure 4: Scatter plots of the SCE calculated according to eq.1 for the -*O*- bridged systems. The contributing metrics storage energy (ΔE_S, top left), thermal back-reaction barrier (ΔE_{tbr}, top right), the absorption wavelengths and the oscillator strengths of the lowest-lying singlet excited state with an oscillator strength above 0.01 of the NBD (middle) and QC (bottom) isomers are explored.

Table 1: Properties of the 40 best systems for the -CH_2- bridge, xTB and DFT data sorted by their respective SCE.

xTB	ΔE_S (kJ/mol)	ΔE_{tbr} (kJ/mol)	λ_R (nm)	f_R	λ_P (nm)	f_R	η (%)
C1	37.8	110.1	535.2	0.18	313.3	0.09	10.04
C2	38.6	112.0	542.6	0.03	319.1	0.04	9.61
C3	33.1	107.1	534.9	0.15	300.7	0.10	8.94
C4	41.2	115.3	490.6	0.15	299.4	0.06	8.50
C5	35.7	111.6	509.7	0.20	304.5	0.07	8.49
C6	33.6	112.0	515.3	0.19	310.6	0.08	8.23
C7	35.9	111.8	494.1	0.15	305.1	0.07	7.60
C8	58.0	91.4	429.5	0.69	339.4	0.03	7.45
C9	30.9	124.1	508.4	0.15	328.3	0.04	7.04
C10	31.5	109.8	486.5	0.16	300.7	0.06	6.33
C11	28.0	122.8	499.6	0.16	312.8	0.12	6.15
C12	35.7	98.0	475.1	0.05	318.1	0.04	6.03
C13	39.7	109.0	449.0	0.28	305.6	0.07	5.97
C14	33.3	125.1	479.8	0.04	293.9	0.10	5.83
C15	48.4	96.6	418.1	1.20	340.4	0.19	5.61
C16	37.5	110.8	449.1	0.25	305.6	0.07	5.60
C17	38.3	110.7	444.3	0.34	321.9	0.02	5.53
C18	27.9	125.4	482.0	0.20	301.0	0.13	5.49
C19	76.1	82.0	386.0	0.55	357.6	0.24	5.19
C20	30.8	110.3	466.3	0.12	304.5	0.06	5.16
C21	30.9	106.4	456.2	0.42	306.8	0.07	5.12
C22	26.5	123.6	481.4	0.16	311.7	0.10	5.10
C23	27.2	121.1	478.7	0.14	311.1	0.08	5.09
C24	27.4	123.7	474.3	0.17	302.2	0.13	5.01
C25	26.5	123.4	476.8	0.16	303.3	0.13	4.94
C26	23.6	123.2	494.7	0.13	331.1	0.04	4.90
C27	28.9	125.2	463.3	0.17	306.4	0.03	4.80
C28	29.9	110.5	460.8	0.11	304.0	0.06	4.75
C29	35.3	121.7	439.2	0.20	300.3	0.11	4.69
C30	27.3	123.4	467.4	0.14	294.5	0.09	4.68
C31	30.2	110.4	456.4	0.14	312.0	0.04	4.65
C32	29.4	121.7	456.9	0.15	309.6	0.07	4.56
C33	31.4	109.0	450.9	0.11	308.8	0.07	4.51
C34	49.3	115.0	400.3	0.65	272.9	0.03	4.50
C35	25.0	123.0	475.3	0.11	316.9	0.02	4.46
C36	28.7	124.7	456.2	0.18	318.0	0.05	4.46
C37	18.4	122.2	516.8	0.12	291.3	0.03	4.43
C38	37.2	107.5	422.1	0.60	316.2	0.05	4.39
C39	19.2	121.9	508.7	0.11	300.7	0.04	4.37
C40	35.1	118.1	427.3	0.29	293.9	0.05	4.23
DFT	ΔE_S (kJ/mol)	ΔE_{tbr} (kJ/mol)	λ_R (nm)	f_R	λ_P (nm)	f_R	η (%)
C15	73.7	129.4	392.3	0.67	320.2	0.14	5.85
C9	56.5	159.4	387.2	0.15	275.7	0.03	3.78
C23	63.7	171.0	377.8	0.10	280.0	0.02	3.47
C11	45.0	53.8	395.0	0.14	277.7	0.11	3.39
C6	49.3	182.9	398.1	0.14	272.9	0.05	3.37
C2	62.0	148.7	383.0	0.02	246.5	0.03	3.24
C1	48.9	185.7	407.0	0.14	273.1	0.09	3.15
C5	48.8	186.7	387.6	0.16	261.0	0.09	3.07
C14	62.4	183.8	377.5	0.02	251.4	0.05	2.87
C22	43.6	194.6	382.3	0.11	276.9	0.07	2.50
C4	55.0	48.0	367.4	0.12	251.1	0.07	2.49
C7	49.6	183.2	371.3	0.13	248.0	0.07	2.46
C17	50.8	182.7	362.6	0.41	256.9	0.07	2.32
C25	44.0	55.5	373.1	0.13	266.3	0.11	2.25
C3	46.3	169.9	370.6	0.10	252.9	0.06	2.20
C18	41.6	59.0	370.7	0.17	266.3	0.17	2.09

C19	108.3	91.3	330.2	0.47	289.9	0.35	2.02
C27	58.6	166.8	355.7	0.14	255.2	0.02	2.02
C13	53.1	177.6	355.3	0.31	246.1	0.06	1.98
C10	40.6	181.5	370.4	0.14	279.0	0.02	1.98
C12	54.7	96.6	361.3	0.04	253.8	0.04	1.91
C8	72.8	158.0	339.2	0.47	236.0	0.08	1.79
C30	44.6	37.2	361.7	0.14	253.0	0.04	1.79
C21	48.9	176.3	354.0	0.32	253.3	0.05	1.77
C36	55.4	160.6	350.8	0.15	268.2	0.02	1.69
C38	61.4	162.8	342.6	0.50	252.4	0.08	1.68
C16	50.8	181.6	350.2	0.27	245.6	0.05	1.63
C40	48.5	189.9	350.3	0.38	256.9	0.08	1.61
C26	36.7	187.3	366.7	0.11	298.4	0.03	1.60
C34	66.0	175.4	337.5	0.44	237.2	0.07	1.54
C35	42.3	187.6	361.2	0.06	261.0	0.09	1.53
C29	55.6	175.1	342.9	0.35	261.9	0.05	1.48
C37	33.4	192.8	361.8	0.06	256.0	0.06	1.24
C33	56.3	175.8	337.3	0.20	256.1	0.08	1.20
C31	51.8	172.3	341.5	0.12	246.9	0.04	1.18
C32	49.3	184.5	340.8	0.13	246.9	0.02	1.11
C20	55.5	172.5	333.5	0.24	251.4	0.09	1.08
C39	32.8	193.3	355.3	0.05	278.2	0.02	0.99
C28	55.3	172.8	331.2	0.22	252.9	0.09	0.98
C24	45.3	236.6	370.8	0.14	265.7	0.12	0.86

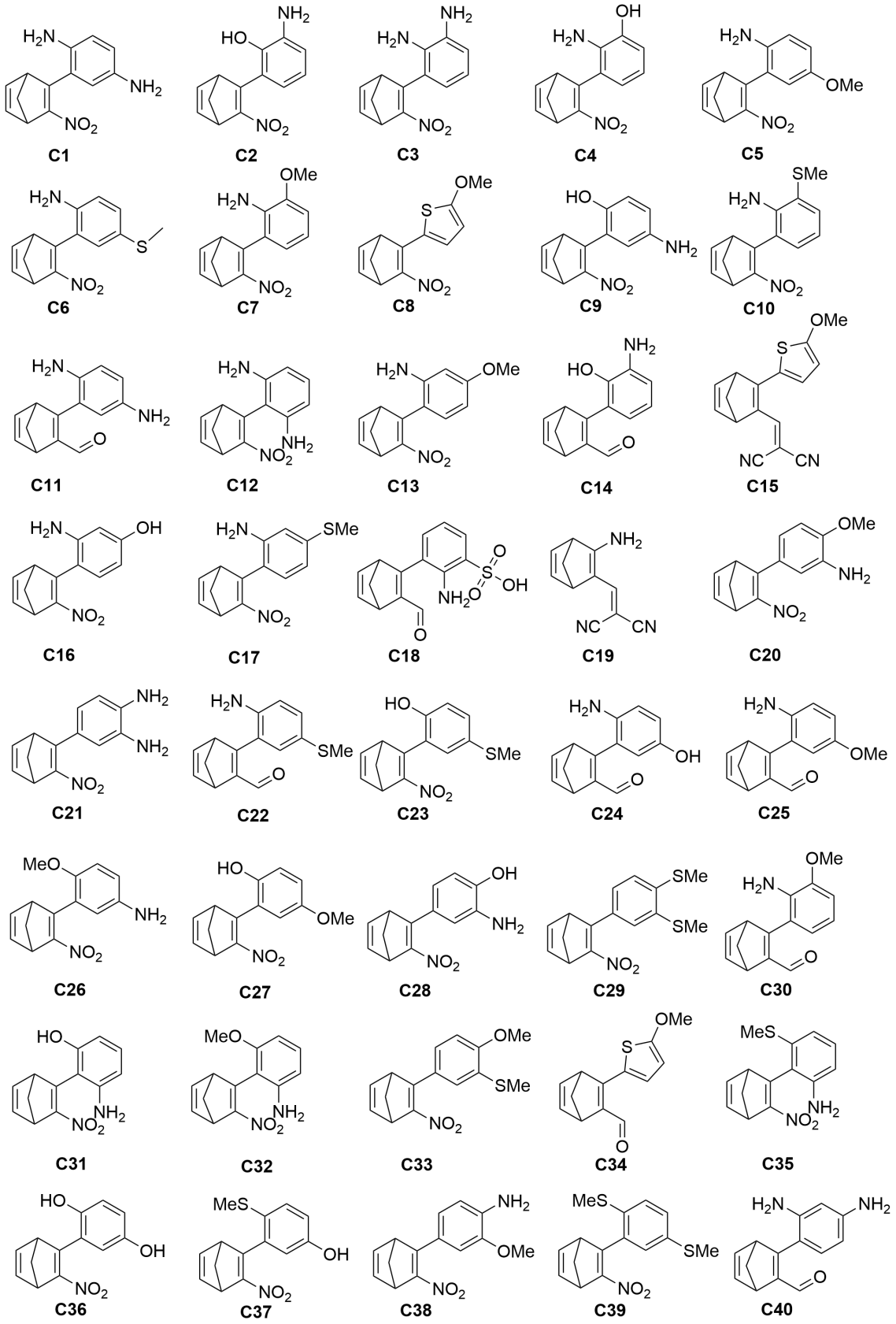

Figure 5: Molecular structures of the predicted 40 best systems after xTB scoring for the -CH_2- bridge-head.

Table 2: Properties of the 20 best systems for the -$NC(O)OMe$- bridge, xTB and DFT data sorted by their respective SCE.

xTB	ΔE_S (kJ/mol)	ΔE_{tbr} (kJ/mol)	λ_R (nm)	f_R	λ_P (nm)	f_R	η (%)
N1	25.4	113.9	561.3	0.19	349.3	0.03	7.47
N2	23.7	111.2	540.3	0.19	319.4	0.08	6.70
N3	22.5	110.0	546.2	0.15	300.1	0.14	6.49
N4	24.5	94.3	521.1	0.20	308.0	0.09	6.25
N5	38.2	81.5	465.0	0.23	420.6	0.25	5.63
N6	22.0	113.3	516.7	0.22	313.7	0.04	5.50
N7	30.0	95.1	459.1	0.34	305.4	0.07	5.03
N8	20.9	115.6	510.8	0.16	334.1	0.03	4.91
N9	34.4	133.1	472.8	0.07	430.5	0.08	4.86
N10	63.4	88.7	391.4	0.53	323.0	0.50	4.83
N11	22.5	100.4	501.6	0.06	315.4	0.03	4.74
N12	31.0	133.7	470.6	0.07	397.2	0.08	4.67
N13	22.9	93.0	479.1	0.35	306.3	0.07	4.56
N14	19.1	123.5	512.6	0.15	315.1	0.11	4.54
N15	16.4	127.3	536.6	0.15	322.6	0.04	4.44
N16	36.0	102.6	420.9	1.21	340.4	0.05	4.34
N17	18.7	121.8	501.8	0.15	316.9	0.12	4.13
N18	20.0	124.1	487.0	0.14	287.2	0.04	4.01
N19	19.8	125.0	498.5	0.05	296.2	0.10	4.01
N20	19.2	123.2	492.5	0.12	316.5	0.05	3.96
DFT	ΔE_S (kJ/mol)	ΔE_{tbr} (kJ/mol)	λ_R (nm)	f_R	λ_P (nm)	f_R	η (%)
N16	66.7	144.1	391.7	0.68	312.8	0.08	5.28
N18	48.0	57.7	401.9	0.08	256.6	0.06	3.85
N15	45.1	175.9	404.8	0.14	276.0	0.03	3.75
N2	40.2	178.9	406.2	0.15	277.4	0.06	3.43
N19	52.8	178.9	395.4	0.02	254.8	0.05	3.37
N4	41.4	28.7	395.4	0.15	300.7	0.02	3.12
N13	46.3	168.9	383.2	0.26	301.5	0.02	3.01
N3	37.8	174.1	389.1	0.09	253.9	0.06	2.52
N10	99.7	106.6	340.3	0.44	286.7	0.44	2.51
N7	43.5	177.3	373.6	0.44	294.3	0.02	2.49
N14	40.9	198.2	427.7	0.11	274.8	0.06	2.37
N6	33.3	173.5	389.1	0.20	258.7	0.02	2.36
N8	39.5	198.2	383.0	0.13	248.5	0.06	2.33
N17	52.5	198.7	399.6	0.11	276.0	0.09	2.32
N11	42.9	4.4	377.2	0.06	256.1	0.04	2.21
N1	35.3	202.5	428.2	0.17	313.3	0.02	2.05
N20	49.1	188.1	348.1	0.23	252.9	0.11	1.46
N5	25.7	43.4	353.0	0.23	328.0	0.19	0.82
N12	59.9	92.1	321.9	0.11	307.2	0.02	0.68
N9	53.3	242.7	322.8	0.13	320.5	0.08	0.43

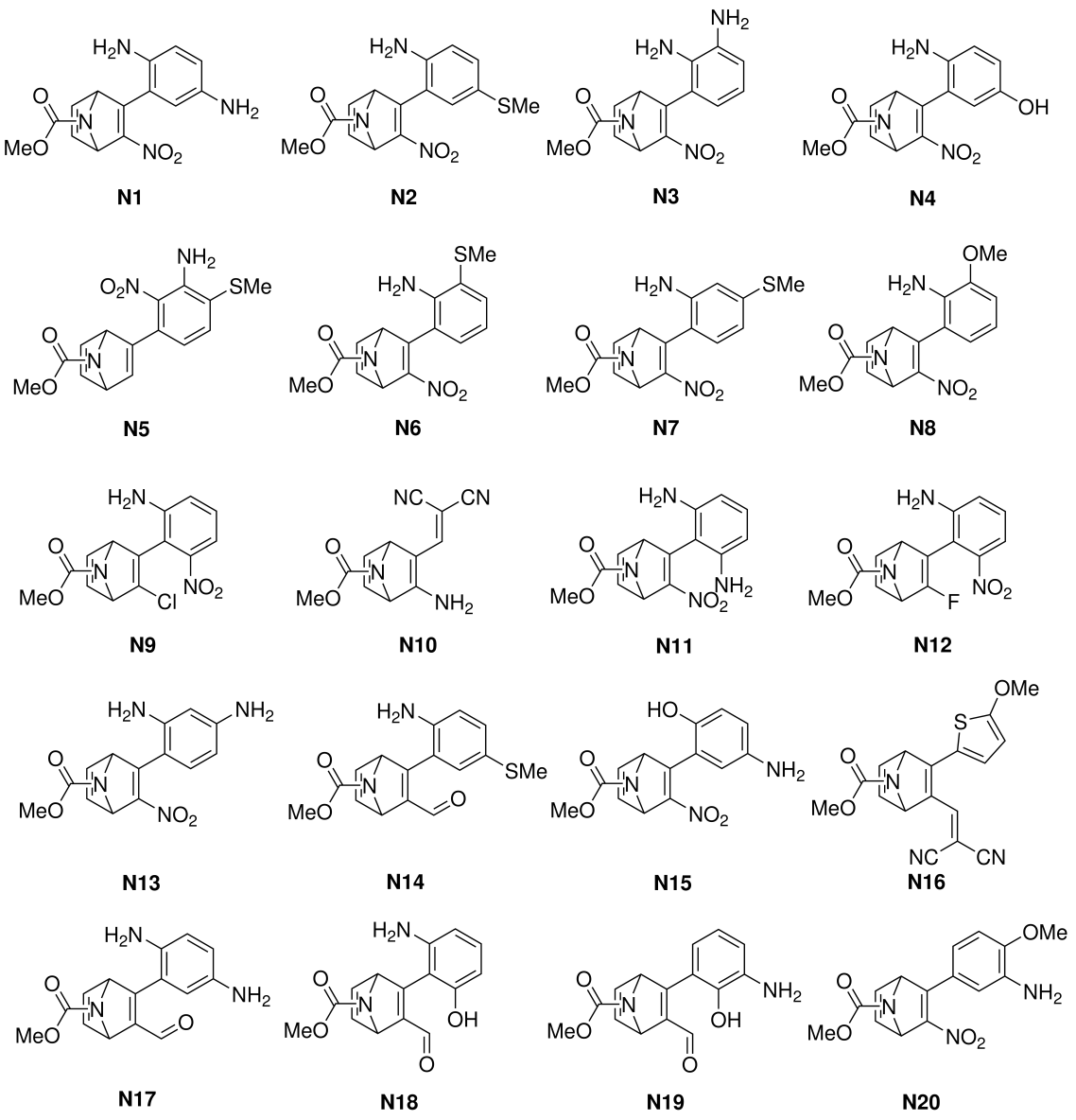

Figure 6: Molecular structures of the predicted 20 best systems after xTB scoring for the -$NC(O)OMe$- bridge-head.

Table 3: Properties of the 20 best systems for the -*O*- bridge, xTB and DFT data sorted by their respective SCE.

xTB	ΔE_S (kJ/mol)	ΔE_{tbr} (kJ/mol)	λ_R (nm)	f_R	λ_P (nm)	f_R	η (%)
O1	15.2	103.4	560.7	0.20	350.4	0.02	4.70
O2	15.3	126.5	571.7	0.04	285.7	0.07	4.33
O3	51.6	99.2	386.9	0.52	328.9	0.20	3.66
O4	12.2	110.4	546.6	0.16	301.0	0.09	3.54
O5	13.0	116.8	529.6	0.23	336.6	0.02	3.47
O6	11.7	114.4	542.3	0.19	338.6	0.02	3.32
O7	13.1	116.2	514.1	0.17	294.3	0.09	3.17
O8	26.7	102.8	416.4	1.24	330.1	0.14	3.05
O9	17.2	114.9	465.7	0.28	326.5	0.02	3.00
O10	26.7	96.0	414.5	0.82	257.9	0.10	2.99
O11	11.1	116.8	524.5	0.21	305.0	0.10	2.88
O12	15.4	114.8	464.0	0.27	294.9	0.02	2.67
O13	25.3	118.1	399.6	0.66	257.6	0.47	2.30
O14	9.4	111.7	513.7	0.17	292.4	0.09	2.27
O15	11.6	111.3	491.4	0.05	297.4	0.03	2.23
O16	10.8	116.8	475.9	0.23	295.1	0.03	2.05
O17	19.5	115.1	408.7	1.03	307.0	0.09	2.03
O18	17.2	113.4	418.0	0.72	312.9	0.03	1.96
O19	18.9	127.0	407.1	0.90	274.7	0.03	1.94
O20	17.3	109.9	415.7	0.75	296.2	0.10	1.93
DFT	ΔE_S (kJ/mol)	ΔE_{tbr} (kJ/mol)	λ_R (nm)	f_R	λ_P (nm)	f_R	η (%)
O8	74.0	139.8	393.7	0.74	308.3	0.15	6.07
O2	61.2	158.4	406.2	0.03	246.5	0.03	4.67
O17	67.5	150.1	379.0	0.65	280.1	0.06	4.41
O19	60.4	15.4	366.3	0.56	252.2	0.09	3.07
O4	46.3	165.6	382.6	0.11	253.8	0.06	2.80
O18	66.5	16.5	356.0	0.61	272.7	0.02	2.69
O7	47.1	194.0	386.1	0.16	296.2	0.02	2.63
O15	45.0	179.4	387.4	0.03	254.9	0.05	2.60
O9	51.8	187.9	365.9	0.34	295.5	0.02	2.48
O5	46.7	196.3	400.9	0.19	303.0	0.02	2.46
O11	45.0	196.2	398.2	0.17	303.8	0.02	2.46
O6	44.5	195.4	417.1	0.14	313.5	0.02	2.46
O3	103.9	106.7	337.4	0.45	282.5	0.37	2.41
O1	46.4	195.3	425.8	0.16	324.1	0.02	2.40
O14	35.0	165.7	386.5	0.15	258.4	0.05	2.33
O16	45.5	195.3	369.8	0.21	280.6	0.02	2.25
O12	48.7	192.5	362.3	0.31	281.2	0.02	2.14
O13	64.9	181.6	343.3	0.42	246.4	0.10	1.79
O10	66.9	164.2	339.7	0.55	235.8	0.07	1.70
O20	60.7	9.5	341.7	0.57	254.0	0.10	1.64

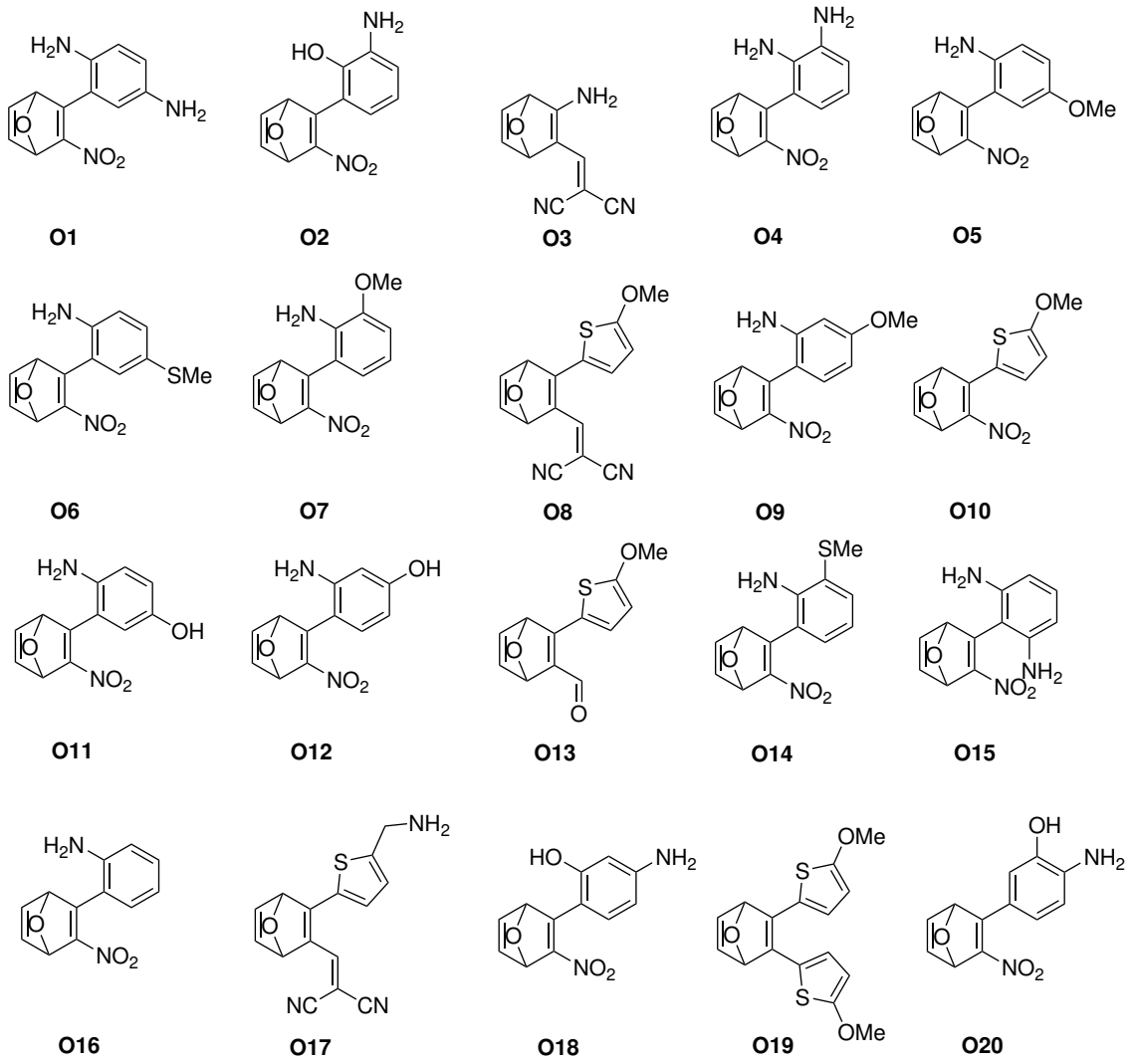

Figure 7: Molecular structures of the predicted 20 best systems after xTB scoring for the -*O*-bridge-head.

References

[1] N. Ree, M. Koerstz, K. V. Mikkelsen and J. H. Jensen, *The Journal of Chemical Physics*, 2021, **155**, 184105.

[2] J. L. Elholm, A. E. Hillers-Bendtsen, H. Hölzel, K. Moth-Poulsen and K. V. Mikkelsen, *Physical Chemistry Chemical Physics*, 2022, **24**, 28956–28964.

[3] A. E. Hillers-Bendtsen, J. L. Elholm, O. B. Obel, H. Hölzel, K. Moth-Poulsen and K. V. Mikkelsen, *Angewandte Chemie International Edition*, 2023, **62**, e202309543.

[4] C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher and S. Grimme, *WIREs Computational Molecular Science*, 2021, **11**, e1493.

[5] S. Grimme and C. Bannwarth, *The Journal of Chemical Physics*, 2016, **145**, 054103.

[6] S. Riniker and G. A. Landrum, *Journal of Chemical Information and Modeling*, 2015, **55**, 2562–2574.

[7] RDKit: Open-source cheminformatics. https://www.rdkit.org.

[8] C. Bannwarth, S. Ehlert and S. Grimme, *Journal of Chemical Theory and Computation*, 2019, **15**, 1652–1671.

[9] H. Jónsson, G. Mills and K. W. Jacobsen, *Classical and Quantum Dynamics in Condensed Phase Simulations*, 1998, 385–404.

[10] K. Börjesson, A. Lennartson and K. Moth-Poulsen, *ACS Sustainable Chemistry and Engineering*, 2013, **1**, 585–590.

[11] M. Kuisma, A. Lundin, K. Moth-Poulsen, P. Hyldgaard and P. Erhart, *ChemSusChem*, 2016, **9**, 1786–1794.