Electronic Supplementary Information (ESI)

A Combined BET and IQA-REG Study of the Activation Energy of non-polar *zw-type* [3+2] Cycloaddition Reactions

Mar Ríos-Gutiérrez,[a]* Fabio Falcioni,[b] Luis R. Domingo,[a] and Paul L. A. Popelier[b]*

[a] Department of Organic Chemistry, University of Valencia, Dr. Moliner 50, 46100 Burjassot, Valencia, Spain. E-mail: m.mar.rios@uv.es

[b] Department of Chemistry, University of Manchester, Oxford Road, Manchester, M13 9PL, (Great) Britain

Index

S2 Analysis of the potential energy surface associated with the 32CA reaction between nitrone **3** and ethylene **5**.

S4 **Table S2** with the ELF valence basin populations, relative energies, IRC values, and C1–C5 and O3–C4 distances, for the last structures of *Phases I – III* along the activation IRC path of the *zw-type* 32CA reaction between nitrone **3** and ethylene **5**.

S5 **Table S3** includes relative ratios and Pearson correlation coefficients *R* of the twelve fully-partitioned IQA terms with largest REG_i values along the activation IRC path associated with the *zw-type* 32CA reaction between nitrone **3** and ethylene **5**.

S5 **Table S4** lists the twelve largest REG_i values and Pearson correlation coefficients *R* of the full IQA partitioning obtained with D3(BJ) dispersion corrections along the activation IRC path associated with the *zw-type* 32CA reaction between nitrone **3** and ethylene **5**.

S6 **Table S5** includes relative ratios and Pearson correlation coefficients *R* for the five most relevant V_{inter}^{AB} energy terms coming from IQA partitioning along the activation IRC path associated with the *zw-type* 32CA reaction between nitrone **3** and ethylene **5**.

S6 **Table S6** gathers the relative ratios and Pearson correlation coefficients *R* of the twelve fully-partitioned IQA terms with largest REG_i values along *Phases I – III* of the activation IRC path associated with the *zw-type* 32CA reaction between nitrone **3** and ethylene **5**.

S7 Summary description of the Ramer-Douglas-Peucker algorithm used for point selection of the IQA-REG analyses.

S8 **References**

Analysis of the potential energy surface associated with the 32CA reaction between nitrone 3 and ethylene 5

Due to the symmetry of both reagents, the *zw-type*[1] 32CA reaction between nitrone **3** and ethylene **5** can only take place through a single pathway. A molecular complex, **MC**, a transition state **TS**, and the final isozazolidine **6** were found along that pathway, indicating that this 32CA reaction takes place through a one-step mechanism. The relative energies of the stationary points are given in Table S1.

Table S1. B3LYP/6-311G(d,p) gas phase absolute and relative energies of the stationary points involved in the 32CA reaction of nitrone **3** with ethylene **5**.

	E	DE
nitrone **3**	-169.847794	
ethylene **5**	-78.613979	
MC	-248.465461	-2.3
TS	-248.440555	13.3
isoxazolidine **6**	-248.506267	-27.9

The **MC** formed at the earliest stage of the reaction is 2.3 kcal·mol^{-1} below the reagents. This low negative relative energy of **MC** suggests a non-polar character for the reaction, as strongly stabilised Electron Density Transfer Complexes (EDTCs) are associated with polar processes[2]. Note that although nitrone **3** is classified as a moderate nucleophile[3] and a moderate electrophile[3] with a nucleophilicity *N* index of 2.93 eV and an electrophilicity ω index of 1.06 eV, ethylene **5** is one of the most unreactive species, *N*=1.87 eV and ω=0.73 eV, being classified as a marginal nucleophile and a marginal electrophile. The activation energy associated with **TS** is 13.3 kcal·mol^{-1}, the reaction being exothermic by about 28 kcal·mol^{-1}.

The geometry of **TS** is represented in Figure S1. The distances between the C1/C5 and O3/C4 atoms, 2.163 and 2.153 Å, respectively, shows a rather symmetric structure in which formation of the new single bonds has not started yet. Note that the formation of C–C and O–C single bonds generally begins at distance ranges of approximately 2.0-1.9 [4] and 1.8-1.7 Å [5]. Nonetheless, these data suggest that, despite the shorter O3–C4 distance at TS, formation of the C1–C5 single bond is slightly more advanced than that of the O3–C4 bond.

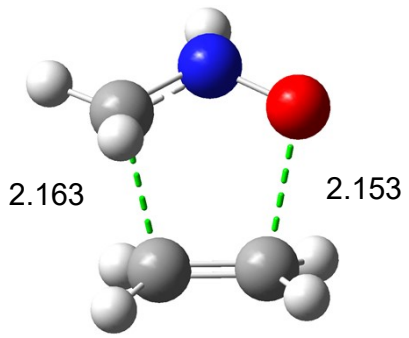

Figure S1. The B3LYP/6-311G(d,p) gas phase optimised geometry of the TS associated with the 32CA reaction between nitrone **3** and ethylene **5**.

Finally, the global electron density transfer (GEDT)[4] at the TS was computed to quantify the polar character of the reaction. Reactions with GEDT values lower than 0.10 e have been classified as non-polar process, while GEDT values higher than 0.20 e characterise polar processes[4]. The GEDT at **TS**, -0.01 e, is negligible, which confirms the non-polar character of this *zw-type* 32CA reaction and allows classifying it as of null electron density flux (NEDF)[6]. This non-polar character accounts for the relatively high activation energy of the *zw-type* 32CA reaction between nitrone **3** and ethylene **5** [1].

Table S2. ELF valence basin populations (in average number of electron, e), relative energies with respect to the first structure of the IRC (ΔE, in kcal·mol^{-1}), IRC values (in Å·amu$^{1/2}$), and C1–C5 and O3–C4 distances (in Ångströms, Å), for the last structures of *Phases I – III* along the activation IRC path of the *zw-type* 32CA reaction between nitrone **3** and ethylene **5**. These structures are labelled as "**Sx(-1)**", where the "(-1)" index refers to the structure immediately before the parent **Sx** structures, whose ELF analyses are given in Table 1 of the main text.

Structures	**S1**	**S1'(-1)**	**S2(-1)**	**S3(-1)**	**TS**
d(C1-C5)	3.413	3.206	2.342	2.282	2.163
d(O3-C4)	2.984	2.871	2.273	2.233	2.153
ΔE	0.0	0.5	11.8	12.9	13.9
IRC	-7.84	-6.51	-1.05	-0.70	0.00
V(C1,N2)	3.69	3.69	3.89	2.64	2.36
V(N2)				1.01	1.35
V(N2,O3)	1.46	1.46	1.32	1.29	1.24
V(C4,C5)	1.70	1.69	1.66	1.50	3.14
V'(C4,C5)	1.66	1.66	1.55	1.69	
V(O3)	3.01	2.99	2.97	2.98	2.95
V'(O3)	2.84	2.83	2.85	2.85	2.86
V(C1)				0.35	0.47

Table S3. Relative ratios (defined in the main text and drawn from Table 2) and Pearson correlation coefficients *R* of the twelve IQA terms with largest REG_i values along the activation IRC path associated with the *zw-type* 32CA reaction between nitrone **3** and ethylene **5**.

Energy term	Ratio	R
V_{xc}(C4,C5)	1.00	0.97
V_{xc}(C1,N2)	0.51	0.96
V_{xc}(N2,O3)	0.50	0.97
V_{cl}(C1,N2)	0.43	0.85
E_{intra}(O3)	0.32	0.99
E_{intra}(C4)	0.31	0.99
V_{xc}(N2,C8)	-0.12	-0.99
V_{cl}(N2,O3)	-0.12	-0.95
V_{cl}(C4,C5)	-0.19	-0.97
E_{intra}(N2)	-0.47	-0.89
V_{xc}(O3,C4)	-0.78	-0.98
V_{xc}(C1,C5)	-0.92	-0.97

Table S4. Twelve largest REG_i values and Pearson correlation coefficients *R* corresponding to the full IQA partitioning with D3(BJ) dispersion corrections along the activation IRC path associated with the *zw-type* 32CA reaction between nitrone **3** and ethylene **5**.

Energy term	REG	R
V_{xc}(C4,C5)	4.01	0.97
V_{xc}(C1,N2)	2.05	0.97
V_{xc}(N2,O3)	1.99	0.97
V_{cl}(C1,N2)	1.76	0.87
E_{intra}(O3)	1.27	0.99
E_{intra}(C4)	1.24	0.99
V_{xc}(N2,C4)	-0.47	-0.99
V_{cl}(N2,O3)	-0.48	-0.96
V_{cl}(C4,C5)	-0.78	0.98
E_{intra}(N2)	-1.92	-0.91
V_{xc}(O3,C4)	-3.11	-0.99
V_{xc}(C1,C5)	-3.69	-0.98

Table S5. Relative ratios (drawn from Table 3) and Pearson correlation coefficients *R* for the five most relevant V_{inter}^{AB} energy terms calculated by IQA partitioning along the activation IRC path associated with the *zw-type* 32CA reaction between nitrone **3** and ethylene **5**.

Energy term	Ratio	R
V_{inter}(C1,N2)	1.00	0.92
V_{inter}(C4,C5)	0.86	0.97
V_{inter}(N2,O3)	0.40	0.97
V_{inter}(O3,C4)	-0.93	-0.98
V_{inter}(C1,C5)	-1.01	-0.97

Table S6. Relative ratios (drawn from Table 4 where available) and Pearson correlation coefficients *R* of the twelve fully-partitioned IQA terms with largest REG_i values along *Phases I – III* of the activation IRC path associated with the *zw-type* 32CA reaction between nitrone **3** and ethylene **5**.

Phase I-1			*Phase I-2*			*Phase II*			*Phase III*		
Energy term	Ratio	R	Energy term	Ratio	R	Energy term	Ratio	R	Energy term	Ratio	R
V_{xc}(C4,C5)	1.00	0.99	V_{xc}(C4,C5)	1.00	0.98	V_{xc}(C4,C5)	1.00	1.00	V_{xc}(C4,C5)	1.00	0.96
E_{intra}(C4)	0.99	1.00	V_{xc}(N2,O3)	0.50	0.98	V_{cl}(C1,N2)	0.71	1.00	V_{cl}(C1,N2)	0.83	0.95
E_{intra}(C5)	0.76	1.00	V_{xc}(C1,N2)	0.50	0.97	V_{xc}(C1,N2)	0.53	1.00	V_{xc}(C1,N2)	0.53	0.96
E_{intra}(C1)	0.76	0.99	E_{intra}(C4)	0.40	0.99	V_{xc}(N2,O3)	0.49	1.00	V_{xc}(N2,O3)	0.50	0.96
E_{intra}(O3)	0.71	1.00	E_{intra}(O3)	0.39	1.00	E_{intra}(O3)	0.21	1.00	E_{intra}(O3)	0.19	0.97
E_{intra}(N2)	0.48	0.97	E_{intra}(C5)	0.25	0.98	E_{intra}(C4)	0.18	1.00	E_{intra}(C4)	0.16	0.97
V_{xc}(O3,C5)	-0.27	-1.00	V_{xc}(N2,C4)	-0.14	-1.00	V_{cl}(N2,O3)	-0.13	-1.00	V_{cl}(N2,O3)	-0.15	-0.96
V_{cl}(N2,C4)	-0.31	-1.00	V_{xc}(N2,C5)	-0.15	-0.99	V_{cl}(C4,C5)	-0.19	-1.00	V_{cl}(C4,C5)	-0.19	-0.96
V_{xc}(N2,C5)	-0.49	-1.00	V_{cl}(C4,C5)	-0.20	-0.98	E_{intra}(C1)	-0.24	-0.98	E_{intra}(C1)	-0.30	-0.94
V_{cl}(C1,N2)	-0.54	-0.95	E_{intra}(N2)	-0.33	-0.84	V_{xc}(O3,C4)	-0.65	-1.00	V_{xc}(O3,C4)	-0.65	-0.96
V_{xc}(C1,C5)	-1.09	-1.00	V_{xc}(O3,C4)	-0.86	-1.00	E_{intra}(N2)	-0.69	-1.00	E_{intra}(N2)	-0.76	-0.96
V_{xc}(O3,C4)	-1.37	-1.00	V_{xc}(C1,C5)	-0.93	-0.99	V_{xc}(C1,C5)	-0.91	-1.00	V_{xc}(C1,C5)	-0.94	-0.96

Description of the Ramer-Douglas-Peucker algorithm.

The Ramer-Douglas-Peucker (RDP) algorithm[7,8] is a method that is mainly used in image processing and cartography to cut down the number of points of polylines (i.e. lines defined by multiple segments) while maintaining a reasonable graphic resolution of the image. In the context of potential energy surfaces, which is at the base of this combined BET and IQA-REG study, this algorithm can be used to consider the minimal number of points on which to run the IQA-REG analysis, thereby reducing its computational cost, while obtaining chemically meaningful results.

The RDP algorithm is heuristic because it defines the number of points through a cut-off parameter usually named ε, which is dimensionless. When applied to potential energy surfaces it is then useful to translate the parameter to a more meaningful metric such as an RMSE, which is expressed in units of energy. This is carried out by a mixture of recursive cross-validation and linear interpolation. Thus, the algorithm outputs the number of most suitable points to consider along the activation energy path given a specific RMSE of tolerance.

For instance, applied to the activation IRC path of the 32CA reaction between nitrone **3** and ethylene **5**, the RDP algorithm yielded only 11 geometries out of a total of 136 with an RMSE of confidence value of 0.05 kcal·mol^{-1} (see Figure S2). This represents a substantial reduction in the number of geometries, which saves much computation time given the high cost of IQA calculations.

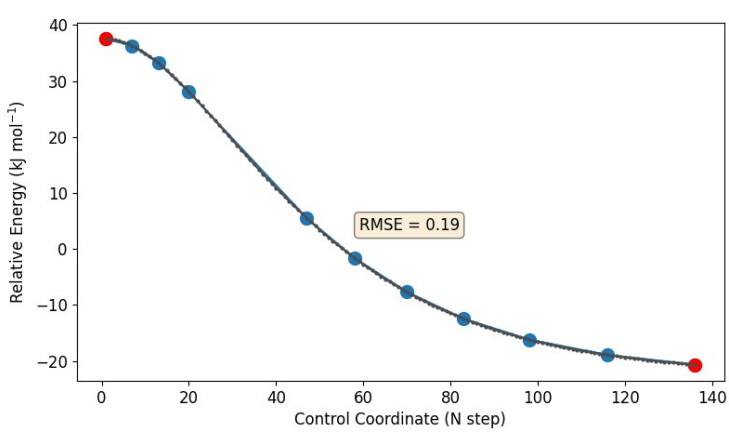

Figure S2. Eleven most suitable geometries to represent the full activation IRC path of the 32CA reaction between nitrone **3** and ethylene **5** with an RMSE value of 0.05 kcal·mol^{-1}.

References

1. M. Ríos-Gutiérrez and L. R. Domingo, Unravelling the Mysteries of the [3+2] Cycloaddition Reactions, *European J Organic Chem*, 2019, **2019**, 267–282.
2. L. R. Domingo and M. Ríos-Gutiérrez, On the nature of organic electron density transfer complexes within molecular electron density theory, *Org. Biomol. Chem.*, 2019, **17**, 6478–6488.
3. L. R. Domingo, M. Ríos-Gutiérrez and P. Pérez, Applications of the Conceptual Density Functional Theory Indices to Organic Chemistry Reactivity, *Molecules*, 2016, **21**, 748.
4. L. R. Domingo, A new C–C bond formation model based on the quantum chemical topology of electron density, *RSC Adv.*, 2014, **4**, 32415–32428.
5. L. R. Domingo, M. Ríos-Gutiérrez and P. Pérez, A Molecular Electron Density Theory Study of the Reactivity and Selectivities in [3 + 2] Cycloaddition Reactions of *C* , *N* -Dialkyl Nitrones with Ethylene Derivatives, *J. Org. Chem.*, 2018, **83**, 2182–2197.
6. L. R. Domingo, K. Kula and M. Ríos-Gutiérrez, Unveiling the Reactivity of Cyclic Azomethine Ylides in [3+2] Cycloaddition Reactions within the Molecular Electron Density Theory, *Eur. J. Org. Chem.*, 2020, **2020**, 5938–5948.
7. D. H. Douglas and T. K. Peucker, ALGORITHMS FOR THE REDUCTION OF THE NUMBER OF POINTS REQUIRED TO REPRESENT A DIGITIZED LINE OR ITS CARICATURE, *Cartographica*, 1973, **10**, 112–122.
8. U. Ramer, An iterative procedure for the polygonal approximation of plane curves, *Computer Graphics and Image Processing*, 1972, **1**, 244–256.