



HAL
open science

Sydnone-alkyne cycloaddition: Which factors are responsible for reaction rate ?

Timur Gimadiev, Olga Klimchuk, R.I. Nugmanov, Timur Madzhidov, Alexandre Varnek

► **To cite this version:**

Timur Gimadiev, Olga Klimchuk, R.I. Nugmanov, Timur Madzhidov, Alexandre Varnek. Sydnone-alkyne cycloaddition: Which factors are responsible for reaction rate ?. *Journal of Molecular Structure*, 2019, 1198, pp.126897. <10.1016/j.molstruc.2019.126897>. <hal-03133365>

HAL Id: hal-03133365

<https://hal.science/hal-03133365v1>

Submitted on 20 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

Sydnone-alkyne cycloaddition: which factors are responsible for reaction rate ?

T.R. Gimadiev ^{1, 2}, O. Klimchuk ¹, R.I. Nugmanov ², T.I. Madzhidov ², A. Varnek ^{1*}

¹ Laboratory of Chemoinformatics, Université de Strasbourg. 4 rue Blaise Pascal, 67000 Strasbourg, France

² Chemoinformatics and Molecular Modeling Laboratory, A.M. Butlerov Institute of Chemistry, Kazan Federal University. 18 Kremlyovskaya Str., 420008 Kazan, Russia

Abstract.

In this work we report extensive DFT study of sydnone-alkyne cyclization which included investigation of the reaction mechanism, analysis of different factors affecting sydnone and alkyne reactivity as well as attempt to reproduce quantitatively experimental activation free energy. The calculations were performed for a set of 18 sydnone-alkyne reactions with a help of a semi-automatized workflow involving reagent preparation and generation of starting structures for a plausible transition state. Reconstructed reaction path supported two-step mechanism: cycloaddition followed by retro-Diels-Alder reaction. Since the latter had a tiny barrier, the cycloaddition step was predicted to be the rate-limiting. For the ensemble of reactions, calculations reproduce activation free energies extracted from experimental reaction rates (k) with the accuracy of 2 kcal/mol. Accounting for solvation effects didn't change the overall trend of activation free energies as a function of substituents. A series of statistical model linking $\log k$ and sydnones structure was built using Support Vector Regression and Multiple Linear Regression machine-learning methods coupled with different types of molecular descriptors; none of them demonstrated a good performance at cross-validation stage. Detailed analysis of different factors affecting reaction rate variation as a function of substituents revealed particular role of the charge on C3 atom in the sydnone moiety as well as of the size of the substituent at C3.

Keywords: biorthogonal reactions, sydnones, DFT, reaction rate, structure-reactivity models

1. Introduction

The ability to selectively form and break chemical bonds in chemically complex and uncontrollable biological media is a long-standing goal of chemists interested in modifying biological materials. Biorthogonal chemical reactions [1], which don't interfere with biological processes, address precisely this challenge and therefore are of major importance in the fields of chemical biology and biochemistry. To fulfill the requirements of bioorthogonality, reaction partners must be stable and inert towards the plethora of chemical functionalities found in living systems while reacting selectively, efficiently and rapidly under physiological conditions with no or innocuous by-products [2,3]. The copper-catalyzed 1,3-dipolar cycloaddition reactions, e.g., azide-alkyne cycloaddition, are the most extensively studied biorthogonal transformations. However, the Cu salts used as catalyst are very toxic for living organisms [4]. This was a reason, to explore new reactions based on strain promoted cycloadditions [5] which proceed without any catalyst but, unfortunately, have rather low rate [6].

Recently sydnone were discovered to be a promising class of biorthogonal 1,3-diene which can participate in cyclization reaction with cyclic alkynes resulting in pyrazole formation [7]. This reaction proceeds in two step: cyclization with formation of reactive intermediate and retro-Diels-Alder reaction leading to CO₂ cleavage (Figure 1). Thermal cycloaddition of sydnones with alkynes require quite harsh conditions and proceeds with low regioselectivity [7] while copper catalysed reaction proceed selectivity in mild conditions with high yield [8–10]. Strained alkynes readily react with sydnones and iminosydnones without catalyst [11–13]. Substituents in sydnone ring [14,15] as well as alkyne structure [12] greatly impact the reaction rate switching it from moderate to ultra-fast. Recently, detailed stopped-flow kinetics study of reaction of some sydnones supported two-stage mechanism [15].

So far, theoretical studies were mostly focused on structure of alkynes or alkenes possessing the greatest reactivity in 1,3-cycloaddition. Thus, Narayanam et al [12] performed a DFT study of N-phenyl sydnone and diaryl-1,2,4,5-tetrazine

reaction with eight different alkenes and alkynes. It has been demonstrated that sydnone interactions with the most strained alkynes proceed with large reaction rate. The importance of steric strain of alkene or alkyne moieties to accelerate kinetics of 1,3-cycloaddition with organic azides or tetrazines was also demonstrated in [16,17]. It looks like strained alkenes and azides are well pre-organized toward the Diels–Alder transition structure and, therefore, interact easier.

Tao et al. [18] performed detailed analysis of cycloaddition rate enhancement for halogen-substituted sydnones. They revealed that distortion energies of sydnone moiety well correlated with activation energies. In such a way, high reactivity of halosydnones was explained by their low distortion energies. Analysis by Tao et al. [18] was limited to halogen-substituted sydnones and can't be directly used to predict reaction rate of cycloaddition involving sydnones with any substituents. Moreover, Tao et al. [18] considered only the first transition state of the reaction between sydnones and bicyclo[6.1.0]nonyne (BCN) but not retro-Diels-Alder reaction.

In this paper, we describe DFT-based full pathway of sydnone-alkyne reactions considering different sydnones and two alkynes - BCN and tetramethylthiacycloheptyne (TMTH). Our aim was to develop a fast semi-automatized screening workflow aiming to select reactants with desired reactivity. A linear correlation between predicted and experimental activation free energies for 18 sydnone-alkyne reactions was observed which shows that gas-phase DFT calculations can be used as a screening tool for reaction design. We also attempted to build structure-reactivity models linking reaction rate with sydnones structure which provides an insight of different factors responsible for sydnones reactivity.

2. Computational details

2.1 Experimental reaction rate data

The data on cycloaddition reaction rate were taken from publications of F. Taran's group [8,13]. The collected dataset contains rate constants at ambient temperatures of 15 reactions of sydnone with BCN and 3 reactions of sydnone with TMTH, see Table 1 and Figure 1. In the series of reactions involving BCN, the fastest is reaction 15 where sydnone bears F substituent at C₃ position. As it follows from Table 1, the same sydnone interacts with TMTH much faster than with BCN (see reactions **15** and **16**). Notice that second-order rate constants given in Table 1 correspond to cycloaddition step only because they were obtained by HPLC monitoring of sydnone conversion [14].

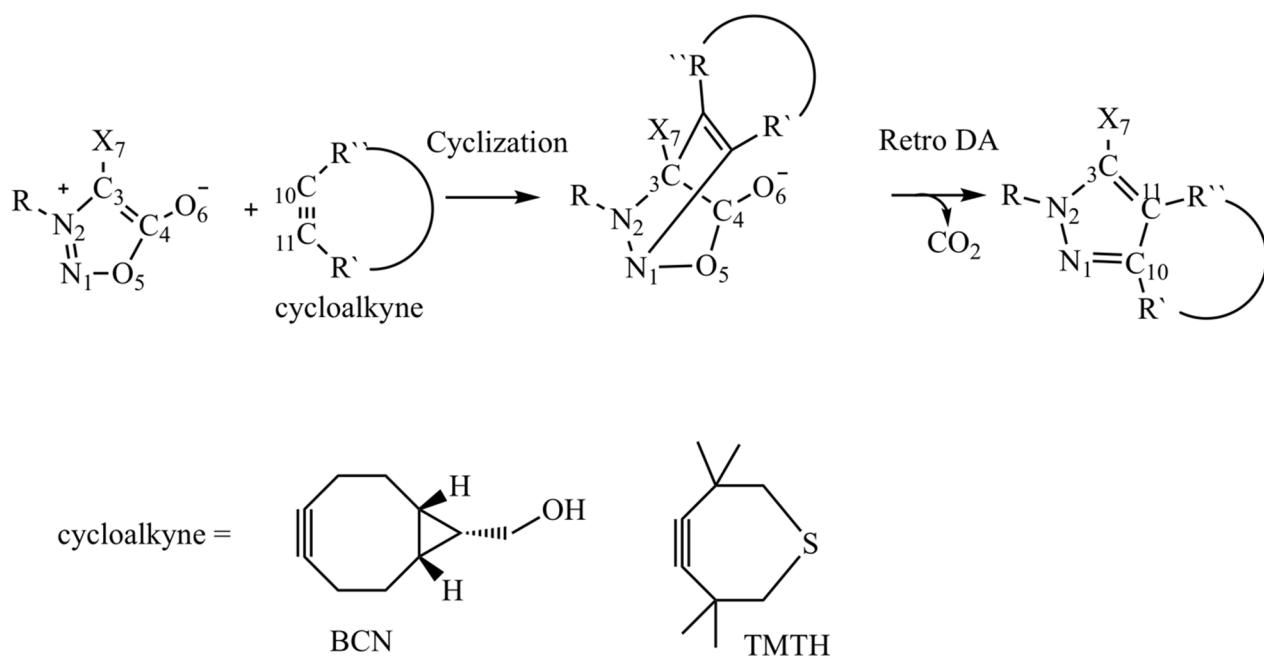


Figure 1. The two steps reactions considered in this study: (3+2) cyclisation followed by retro-Diels-Alder reaction.

Table 1. Rate constants of reaction shown on Figure 1. The measurements were performed in water - DMSO mixture.

No	R	X	Cycloalkyne	Rate constant (tolerance), M ⁻¹ s ⁻¹	Reference

1	p-MeO C ₆ H ₄	H	BCN	0.006 (0.001)	[14]
2	p-Me C ₆ H ₄	H	BCN	0.032 (0.001)	[15]
3	C ₆ H ₅	H	BCN	0.027 (0.002)	[14]
4	p-CO ₂ H C ₆ H ₄	H	BCN	0.059 (0.001)	[14]
5	p-CF ₃ C ₆ H ₄	H	BCN	0.199 (0.002)	[14]
6	p-NO ₂ C ₆ H ₄	H	BCN	0.289 (0.012)	[14]
7	C ₆ H ₅	CH ₃	BCN	0.018 (0.002)	[14]
8	C ₆ H ₅	C ₆ H ₅	BCN	0.027 (0.001)	[14]
9	C ₆ H ₅	CF ₃	BCN	0.008 (0.001)	[14]
10	p-Me C ₆ H ₄	Cl	BCN	0.872 (0.034)	[15]
11	p-Me C ₆ H ₄	Br	BCN	0.592 (0.021)	[15]
12	p-Me C ₆ H ₄	I	BCN	0.306 (0.008)	[15]
13	p-CO ₂ H C ₆ H ₄	Br	BCN	0.798 (0.065)	[14]
14	p-CO ₂ H C ₆ H ₄	Cl	BCN	1.593 (0.034)	[14]
15	p-Me C ₆ H ₄	F	BCN	42	[15]
16	p-Me C ₆ H ₄	F	TMTH	1500	[15]
17	p-F C ₆ H ₄	F	TMTH	3500	[15]
18	p-CF ₃ C ₆ H ₄	F	TMTH	12000	[15]

2.2 Quantum chemical calculations

Density Functional Theory (DFT) [19,20] calculations were performed with the Priroda 11 program [21,22] using PBE exchange and correlation functional [23] and built-in triple-zeta split valence basis set 3z which is equivalent to Schäfer's TZVP basis set [24]. Relativistic effects were neglected. Priroda11 is one of the fastest DFT software due to efficient evaluation of density functional exchange-correlation terms based on the electron density expansion [25].

Geometry optimization for reagents, intermediates and products as well as saddle point optimization were performed using built-in quasi-Newton-Raphson method and BFGS hessian update scheme. Scanning procedure was used to

localize reasonable structural guess for transition state. This was achieved in constrained local optimization along the N1...C_{alkyne} coordinate which varied from 2.1 to 3.1 Å with step 0.1 Å. In order to detect different kind of species (stable molecules, intermediates, transition state) geometry optimization was followed by the frequencies calculations. Thus, no imaginary frequencies must be found for stable species (reagents, products and intermediates) and only one large imaginary frequency must be detected for a transition state. The correctness of transition states was also checked using the intrinsic reaction coordinate (IRC) following approach.

Full reaction pathways were calculated for reaction **2** (also considered in early theoretical calculations [12]) as well as reactions **15** and **16** for which detailed kinetic experiments were performed by Liu et al [15].

Since continuum solvation is not available in PRIRODA, for some molecules and transition states, solvation free energy in water was calculated using the IEF-PCM model [26,27] with SMD parameters for non-electrostatic terms [28] implemented in the Gaussian 09 program [29]. In this case, geometry optimization and hessian were calculated with Gaussian 09 [29] using PBEPBE/6-311++G(d,p) method.

2.3 Workflow for transition state detection

All structures were drawn with the Marvin View program [30] following the same numeration of atoms in sydnone ring. ChemAxon Calculator Plugin *cxcalc* [31] was used to generate up to 500 conformers with diversity limit 0.1 Å followed by geometry optimization with the MMFF94 force field [32–34]. Then, geometry of each structure was optimized using the PM6 semi-empiric method [35] implemented in the MOPAC 11 program [36]. Duplicated conformers (with RMSD <0.1 Å) were removed using an *in-house* script which aligns molecules according to principal components of inertia. Remaining conformers were ranked according to their energies, and then 100 most stable conformers were retained for

further calculations. All structures issued from PM6 optimization were re-optimized with DFT PBE/3z. One lowest energy conformer was selected.

Four possible orientations of reagents were considered (Figure 2). They lead to formation of two pairs of enantiomeric 1st transition states (TS1). Further, at the retro Diels-Alder stage, a racemic mixture of two enantiomers was formed.

Four possible orientations of reagents were generated using in-house Python script in a way that distance between reaction center atoms (C3...C10 and N1...C11 or C3...C11 and N1...C10, depending on orientation) was 3.1 Å, i.e. much greater than those in TS1 (Figure 2). A reasonable guess for TS1 was found using scan procedure implemented in PRIRODA11 in which the distance between reaction center atoms varied from 3.1 Å to 2.1 Å with step size 0.1 Å. For a greatest energy structure, the hessian was calculated followed by saddle point optimization.

Theoretically, pair of enantiomeric structures of TS1, intermediate and TS2 should have the same energies. In practice, however, due to possibility to generate different conformations for the stereoisomers, the energies of TS1 resulted from 4 different orientations of reagents slightly differ. For further analysis the lowest energy transition state among the four structures was selected. Hessian for the selected structure was calculated and presence of one large imaginary frequency was checked. The correctness of transition state TS1 was checked using Intrinsic Reactional Coordinate following (IRC) procedure.

The second transition state TS2 was found after IRC downhill shift towards the products. Resulting intermediate structure was used as starting structure for scanning reaction coordinate (N1-O5 bond length in this case). We found that small step of 0.01 Å was required because TS2 looks like a small “hubble” on the potential energy surface followed by steep energy drop.

Detailed description of the workflow and corresponding Python scripts are given in Appendix 1 in Supporting Information.

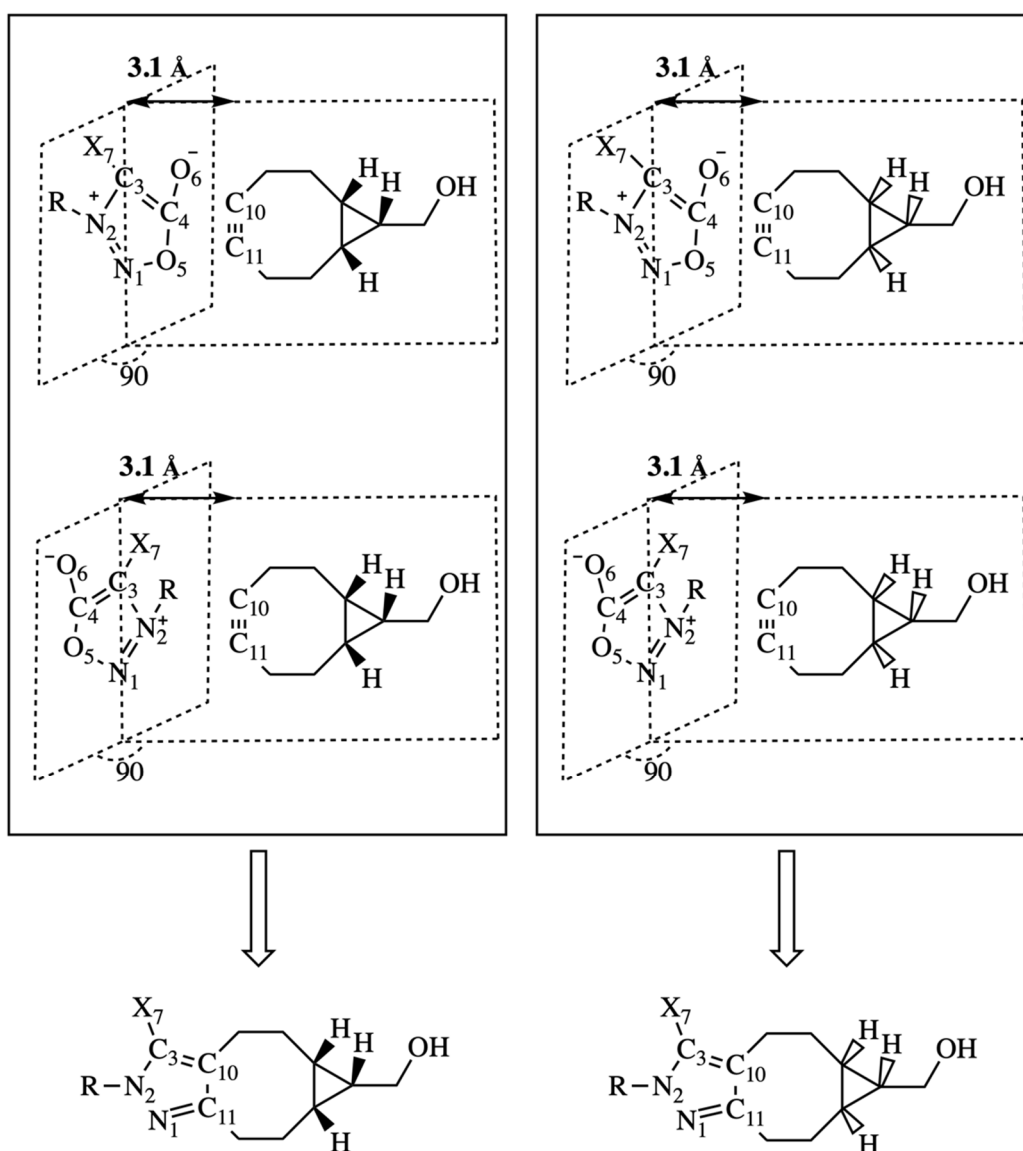


Figure 2. Four possible pre-reaction complexes sydnone and alkyne (*top*) and related products (*bottom*)

2.4 Activation free energy calculation

Free energy of molecule and transition state formation was calculated using built-in thermochemical calculation in rigid ideal rotator and harmonic oscillator approximation at 298 K in Priroda11 program. Energy includes zero-point vibrational energy corrections.

Formula (1) was used for calculation of activation free energy on the basis of quantum chemical calculations:

$$\Delta G^\ddagger = \Delta G_{298}^{TS} - (\Delta G_{298}^{sydnone} + \Delta G_{298}^{alkyne}) \quad (1)$$

Calculation of activation free energy based on experimentally measured reaction rate was performed using formula (3) derived from Eyring equation of the Transition state theory (2) :

$$k = \kappa \frac{K_B T}{h} e^{\frac{-\Delta G^\ddagger}{RT}} \quad (2)$$

$$\Delta G^\ddagger = -RT \ln\left(\frac{k \cdot h}{\kappa K_B T}\right) \quad (3)$$

where k - rate constant; κ – transmission coefficient (here, $\kappa=1$); T - Temperature in Kelvin; K_B -Boltzmann constant; ΔG^\ddagger - Difference of free energies of reagents and TS at 298.15 K; h - Planck constant; R - gas constant.

2.5 Conceptual DFT indices calculation

Several Conceptual DFT [37] indices were used in order to reveal structural factors responsible for reaction rate variation. Electrophilicity index, ω , measures the stabilization energy when system acquires an additional electronic charge [38]:

$$\omega = \frac{\mu^2}{2\eta}$$

where μ is electronic chemical potential, and η – chemical hardness. The latter could be expressed in terms of HOMO (ϵ_{HOMO}) and LUMO (ϵ_{LUMO}) energies as:

$$\mu = (\epsilon_{HOMO} + \epsilon_{LUMO})/2$$

$$\eta = \epsilon_{LUMO} - \epsilon_{HOMO}$$

The HOMO and LUMO energies were obtained within DFT scheme [20] for sydnone molecules.

Fukui indexes are used to characterize atom ability to share/withdraw electronic charge. For calculation of Fukui nucleophilicity F_A^- , electrophilicity F_A^+ and radical attack susceptibility F_A^0 indexes of atom A single point calculations of molecules with added and removed electron were done:

$$F_A^- = P_A(N) - P_A(N - 1)$$

$$F_A^+ = P_A(N + 1) - P_A(N)$$

$$F_A^0 = 0.5 * (P_A(N + 1) - P_A(N - 1))$$

where $P_A(M)$ – Hirshfield charge on the atom A in molecule with M electrons, N - number of electrons in neutral molecule. Geometry of cation-radical and anion-radical molecules corresponded to lowest energy structure of neutral molecule.

Charge transfer from alkyne to sydnone in the transition state was calculated as absolute value of the entire electric charge of one of the species.

2.6 QSPR modeling

Quantitative Structure-Property Relationships (QSPR) for $\log k$ were established using Support Vector Regression (SVR) and Multiple Linear Regression (MLR) machine-learning methods as well as ISIDA fragment descriptors [39] or quantum mechanics descriptors generated for sydnones.

SVR models were built and validated using the ε -SVR algorithm implemented in the libSVM package [40] using ISIDA fragment descriptors. The modeling was performed using the evolutionary SVR optimizer [41], which can be used to perform both descriptor space selection and optimization of the operational parameters (epsilon, kernel type, cost, gamma) of the SVR method.

The MLR models were built using CODESSA software [42]. Three different sets of descriptors were explored: (i) quantum-chemical CODESSA-PRO descriptors [42], (ii) descriptors based on Bader's quantum theory of Atoms in Molecules [43] calculated with CODESSA-3 [44] program, and, (iii) conceptual DFT [37] indices calculated for key atoms of sydnone with Priroda 11 [21,22].

Predictive performance of the models has been estimated by root mean squared error and squared determination coefficient (R^2) calculated in 10-fold cross-validation. Notice that predictive performance is acceptable only for the models with $R^2 > 0.5$.

3 Results and Discussion

3.1 Reaction pathway investigation

The full reaction path was explored for the reaction **2** and **15** (see Table 1) sharing the same alkyne (BCN) and sydrones differed by the substituent X (X = H and F, respectively). The procedure included scanning along the reaction coordinate, optimization of the transition state, motion along the intrinsic reaction coordinate which links transition state with reagents and products. At the initial stage, the first transition state (TS1) was localized. An intermediate was found in course of intrinsic reaction coordinate downhill movement with following local optimization. The intermediate's geometry was used to identify the second transition state (TS2) corresponding to retro-Diels-Alder reaction. In this case, an elongation of the N1-O5 bond was used as reaction coordinate (Figure 2). Since TS2 structural was close to intermediate, a small step size of 0.01 Å was used in the scan procedure. Geometry of TS2 was found by saddle point optimization of the highest energy structure found in the scan procedure. IRC procedure supported that TS2 belong to reaction minimal energy pathway. Descent from TS2 resulted in either the intermediate or the product molecules. Energetic profile of reaction (Figure 3) shows that the N1-O5 bond length smoothly elongates along the reaction path. Thus, for reaction **15** this bond length is equal to 1.379 Å in reagents, 1.414 Å in TS1, 1.546 Å in the intermediate and 1.721 Å in TS2.

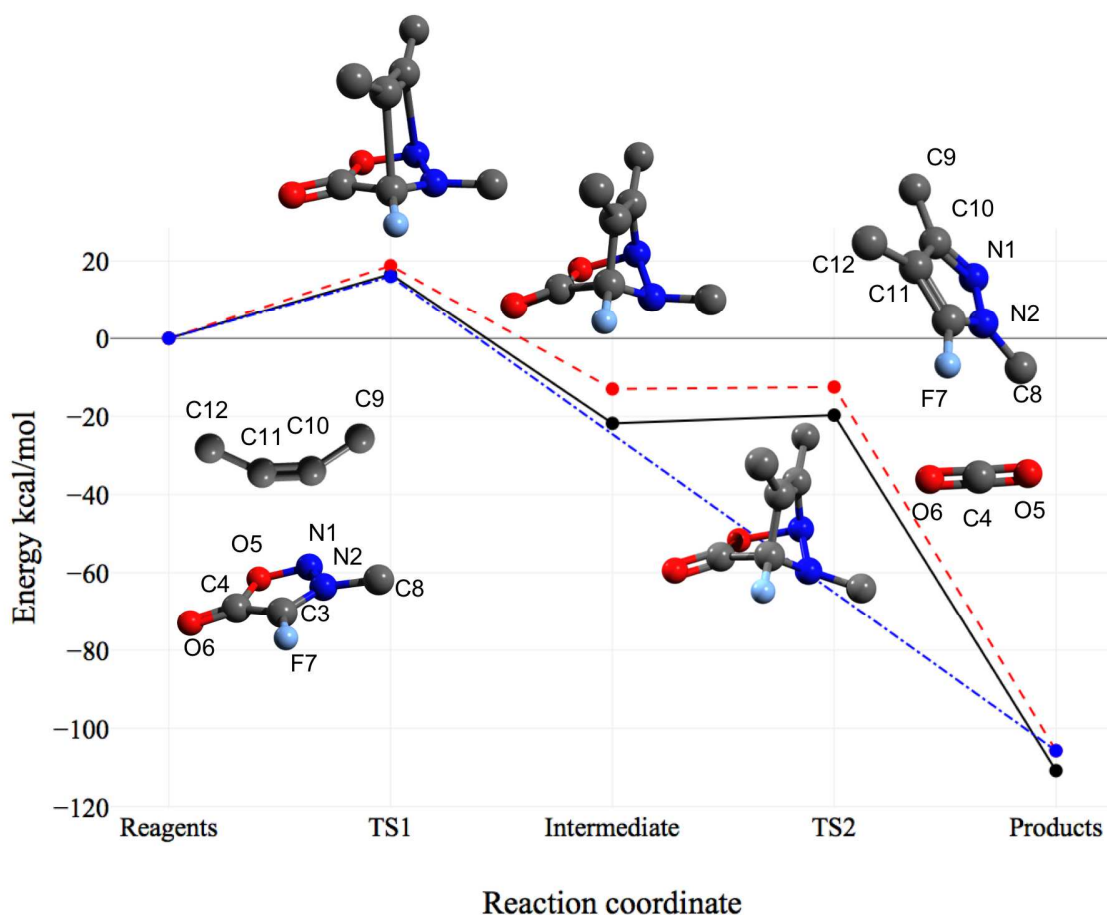


Figure 3. Reaction pathways for reactions **2** (dashed red line), **15** (solid black line) and **16** (dashed blue line). Relative free energies at 298K of molecules with respect to reagents are shown. Structure of reagents, transition states and intermediate for reaction **15** is shown. Substituent R of sydnone and some atoms of alkyne are omitted for the sake of clarity. Bond orders correspond to molecule representation on Figure 1.

Figure 3 shows that the limiting step of reaction **15** is the first transition state (TS1), free energy of the intermediate is 21 kcal/mol lower than that of reagents. The intermediate looks rather unstable since it is separated from products by transition state (TS2) with a tiny barrier (some 1 kcal/mol). Decomposition of the intermediate is exergonic by some 85 kcal/mol. The pathway of reaction **2** looks very similar to that of **15**. Much smaller, compared to **15**, rate constant of **2** (Table 1) can be explained by ~2 kcal/mol higher activation free energy which is in line with the results by Tao et al [18]. Quite small TS2 activation free energies decrease from reaction **2** to **15** together with TS1 energies. Unlike reactions **2** and **15**, reaction **16** between fluorosydnone and TMTH proceeds as synchronous one-step

addition /elimination leading immediately to the products. Similar observations concerning very small second barrier in reactions between sydnone and BCN were also reported by Narayanam et al [12]. These observations, however, contradict with experimental results by Liu et al [15] who demonstrated that the second step was rate limiting and quite slow compared to the first one. A contradiction between experiment and computations could be related to cage effect of solvent or some specific interactions TS2 with solvent. On the other hand, solvation free energies of reactants and TS1 estimated with the IEF-PCM model [26,27] were found very similar for all studied reactions. Thus, the first-step reaction barrier was found insensitive to solvent effects.

3.3 Activation free energies: prediction vs experiment

The question arises: whether rate constants estimated from activation free energy correlate with the experimentally measured values? Activation free energies, $\Delta G^{\ddagger}_{\text{calc}}$, were assessed for all 18 studied reactions according to eq. (1). For the comparison purposes, “experimental” free energies, $\Delta G^{\ddagger}_{\text{exp}}$, were calculated from the rate constants according to eq. (3). Figure 4 demonstrates that calculated free energies well correlated with the “experimental” ones. Thus, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) resulted from deviations between predicted and experimental ΔG^{\ddagger} are 1.97 kcal/mol and 1.69 kcal/mol, respectively. With respect to the correlation line, the errors are even smaller (RMSE = 1.23 kcal/mol, MAE = 0.98 kcal/mol, p-value is $8.6 \cdot 10^{-6}$) whereas the correlation coefficient $r=0.85$, is fairly high. The errors of calculations are still larger than “chemical accuracy” estimated by 1 kcal/mol, but the latter can hardly be reached by common DFT functionals [45]. Thus, developed workflow provides relatively fast, accurate, and inexpensive (in comparison with experimental measurement) computational screening aiming to identify fast reaction.

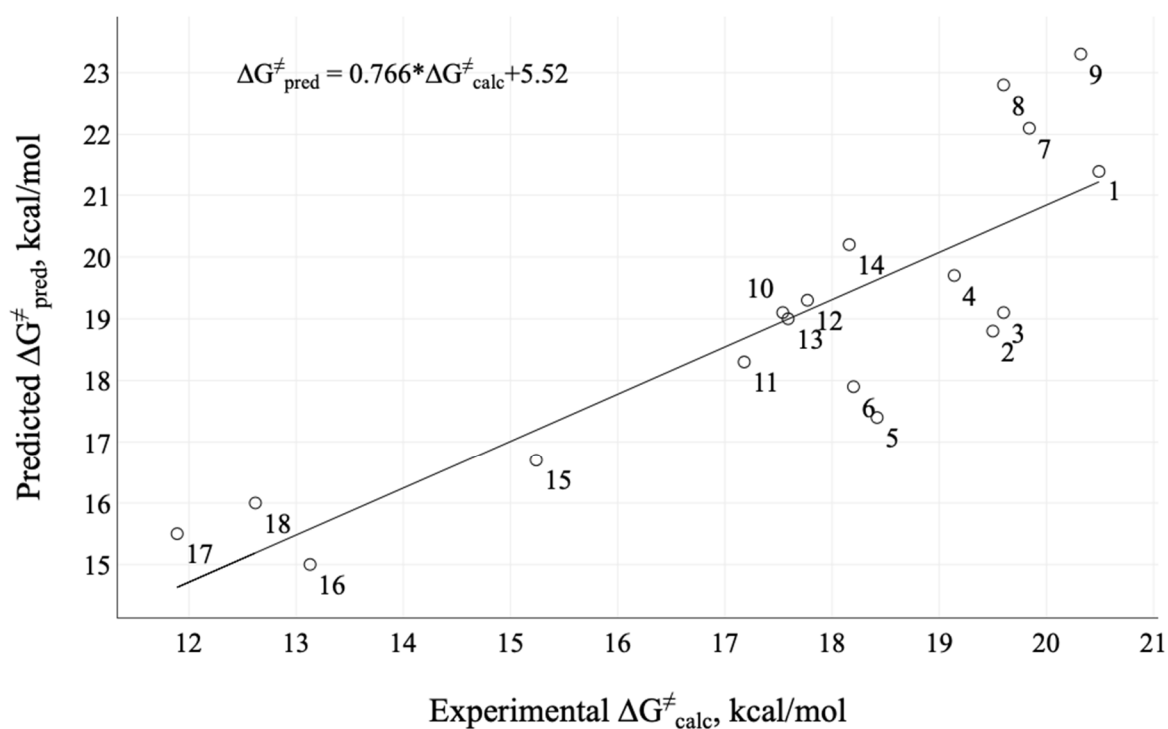


Figure 4. Free energies of activations predicted by quantum chemistry calculations and those calculated with formula (3). Reactions numbers corresponds to those in Table 1. Solid line shows best linear correlation $\Delta G^{\ddagger}_{\text{pred}}$ VS $\Delta G^{\ddagger}_{\text{exp}}$

3.4 QSPR models

Quantum chemical assessment of reaction rate is rather reliable but, computational resources consuming task. In this case, QSPR models for $\log k$ could become an efficient solution allowing to accelerate the screening of potentially interesting reactants. A large panel of models was built using Multiple Linear Regression and Support Vector Regression [46] machine learning methods and different types of molecular descriptors for sydnones including conceptual DFT indices (see Section 2.6 for details). Despite of our efforts, no models with acceptable predictive performance in cross-validation were obtained. This can be explained by small number of considered cycloaddition reactions and relatively large structural diversity of sydnones listed in Table 1.

3.5 Analysis of structural and electronic factors affecting reaction rate

Conceptual Density Functional Theory (DFT) indices [37] are often used to analyze cycloaddition reactions [47,48]. To reveal structural factors affecting the rate constant we looked for linear correlations within restricted series of reactions involving the same alkyne (BCN). However, neither sydnone electrophilicity index [38] nor chemical potential (calculated as half-sum of HOMO and LUMO energies) nor chemical hardness (calculated as HOMO – LUMO gap) correlate with $\log k$.

Analysis of orbital symmetries and energies shows that electrons are transferred from HOMO of alkyne to LUMO of sydnone (Figure 5) in agreement with earlier publication by Liang et al [49] studied azide and tetrazine cycloaddition to alkynes. No valid correlation of LUMO of sydnones with reaction rate was found. In contrast to observations of Domingo et al [48] who studied Diels-Alder reactions between cyclopentadiene and some ethylene derivatives, no correlation of $\log k$ with the charge transfer in the transition state was found. So even within this congeneric series of reactions there is no evidence that orbital interactions represent the major factor affecting the reaction rate variations.

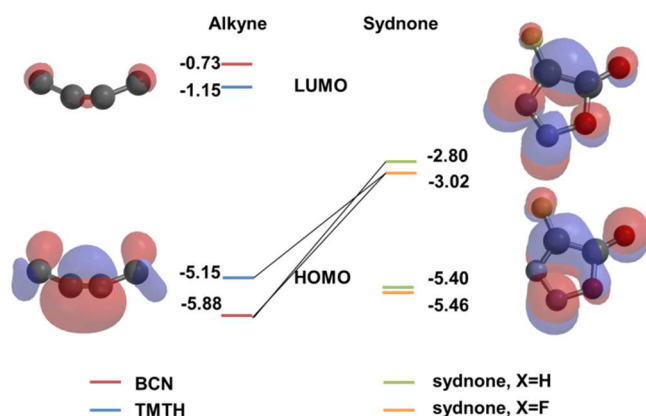


Figure 5. HOMO and LUMO energies obtained in DFT PBE/3z (TZVP) calculations for reactions **2**, **15**, **16**. The energies are given in eV.

Different descriptors characterizing electrostatic interactions (Fukui nucleophilicity, electrophilicity and radical attack susceptibility) using Conceptual DFT [37] were calculated. However, none of them correlated with $\log k$ values for all 15 considered sydnones – BCN reactions. One may suggest that there exists a

complex interplay of structural and electronic factors which cannot be explained by simple linear correlations.

In order to reveal factors responsible for reaction rate, we considered two small congeneric series of reactions. The first one assembled reactions of BCN with sydnone having different substituents X and the same R = p-Me C₆H₄ (**2**, **10-12**, **15**), whereas the second one included three reactions of TMTH with fluorosydnone (X=F) having different substituents R (**16-18**). Calculations for the first series show significant variation of charges on C3 atom, which change from negative (X = H) to positive value (X = F), whereas the charge on N2 atom slightly varies as a function of substituents. In this series, log*k* well correlate with the C3 charges (Figure 6a), $r=0.98$ and RMSE = 0.13 log*k*, p value $4 \cdot 10^{-4}$. In line with observations reported by Tao et al. [18], no correlation with LUMO energy was observed (Figure 6b).

Variation of aryl substituents at N3 atom in reaction of fluorosydnone with TMTH (**16-18**), doesn't lead to significant variations of charge at C3 atom (Figure 6a). The LUMO energy in this series decreases from -3.57 to -3.05 eV which could partially explain considerable raise of the rate constant (Figure 6b).

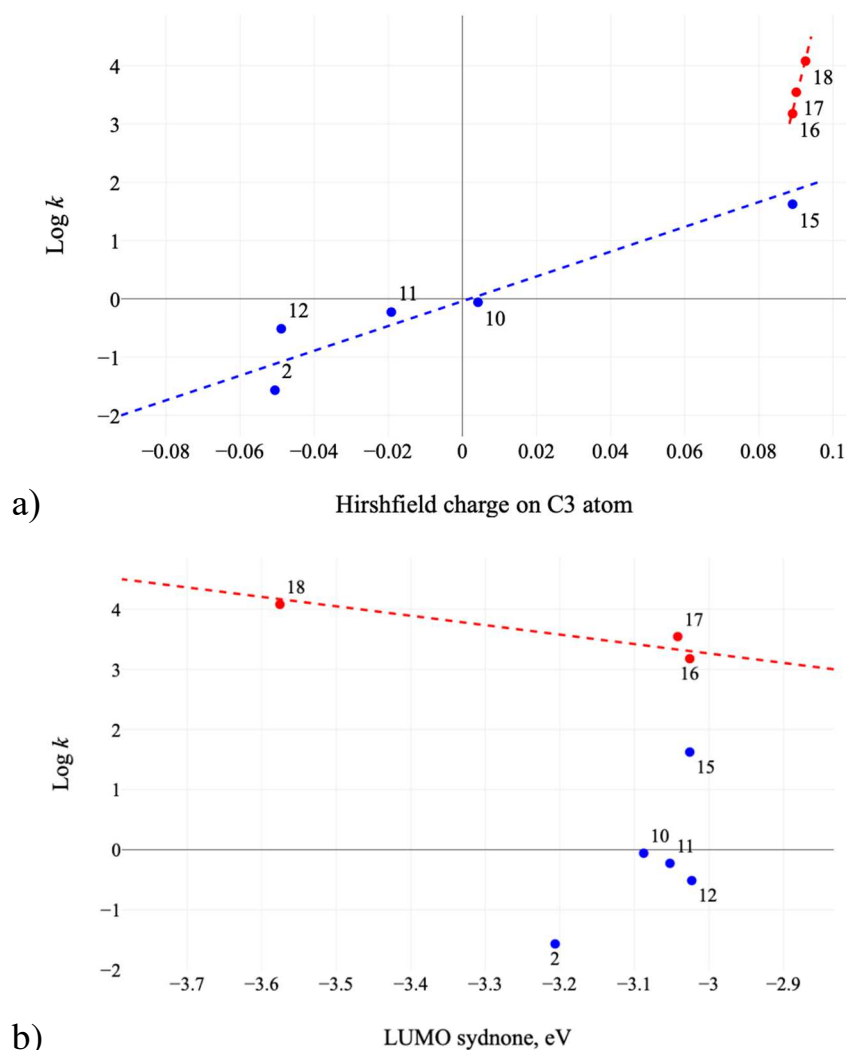


Figure 6. Variation of $\log k$ in congeneric series of halo-sydnone as a function of (a) Hirshfield charge of C3 atom of sydnone and (b) LUMO of sydnone. The numbers correspond to reactions in Table 1. Reactions with TMTH and BCN are shown in red and blue, respectively. The reaction numbers correspond to those in Table 1.

Notice that larger steric strain of alkyne in TMTH compared to BCN affects both charges on triple bond carbon atoms (-0.663 and -0.652 in BCN and -0.636 for both atoms in TMTH) and HOMO energies (-5.88 eV in BCN and -5.15 eV in TMTH, Figure 6). The latter favors larger reactivity of TMTH. This effect was mentioned in previous works on optimization of alkyne reactivity [50].

DFT calculations of the transition state of the cycloaddition of sydnone with BCN reveal an increase of the distance between alkyne C atom and C3 atom ($C_{\text{alkyne}}\text{-C3}$) with the size of the substituent at C4 atom: 2.283 Å, 2.345 Å and 2.348 Å for X = H, F and I, respectively (Figure 7). This could be explained by steric

repulsion between substituent **X** and aliphatic CH₂ group next to triple bond in the alkyne moiety in the transition state. The shortest contact H...I (3.486 Å) in transition state TS1 is almost equal to the sum of the Pauling vdW radius of H and I (3.4 Å). While fluorine atom is almost located in the plane of sydnone ring in TS1, iodine atom deviates from it avoiding steric clashes, Figure 7. It can also be noticed that distances C3-C10 and N1-C11 in case of fluoro- and unsubstituted sydnones are almost equal (see Supplementary Information) which supports synchronous cyclization mechanism. At the same time, the N1-C11 bond in sydnones 10-12 (X = Cl, Br, I) is shorter than the C3-C10 bond. This supports our hypothesis that steric clashes can lower reaction rate in case of bulky substituent **X**. Notice that steric and electronic effects for halogens vary in the same direction.

The sum of valence angles around C3 atoms can be considered as an implicit measure of its planarity: more its value differs from 360°, less planar (and more strained) the sydnone is. The calculation show that this value gradually decreases in TS1 of reactions **15**, **2**, **10**, **11** and **12**: 353.6°, 353.2°, 351.9°, 351.5° and 351.2° respectively. One can see that in TS1, the sydnone with X=F (reaction **15**) is as planar as the sydnone with X=H (reaction **2**). One may suggest that attractive interactions between negative charge on the F atom and positive charge on hydrogen atoms of alkyne stabilize TS1 of reaction **15**.

Notice that these observations disagree with conclusions drawn by Liang et al [49] that electronically more reactive electrophiles should be sterically more encumbered. Indeed, iodo-sydnone in **12** is sterically more encumbered than fluoro-sydnone in **15**. However, the former is less reactive due to weaker electron acceptor ability, on one hand, and steric repulsion with alkyne hydrogens, on the other hand. Replacement of F with more bulky electron-withdrawing substituent leads to reduction of the reaction rate. Thus reaction **9** involving sydnone with X=CF₃ is drastically slower than reaction **10** with X=Cl, despite the fact that atomic charges on C3 in these molecules are similar (0.009 and 0.0042 a.u., respectively).

In the context of rational design of highly reactive substrates, our

calculations lead to the following observations:

- Strong electron acceptor X at C3 atom increases its positive charge and, thus, improves its affinity to alkyne in cycloaddition. However, this substituent should not be bulky because it may cause steric hindrance in the transition state. In this context, fluorine substituent at C4 represents an optimal choice.
- Strong electron acceptor at N3 atom weakly affects charge distribution in sydnone moiety but decreases its LUMO energy. This, in turn, favors cycloaddition, especially in reactions with TMTH.
- Compared to BCN, more sterically strained TMTH has higher HOMO energy which explains its better reactivity.

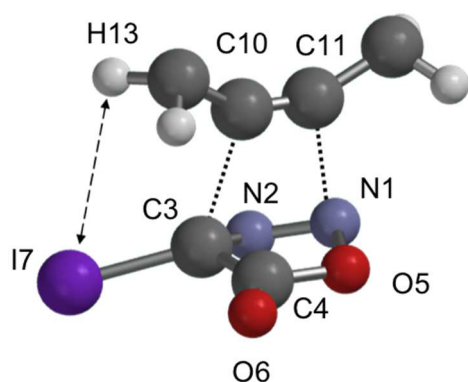


Figure 7. Transition state of cycloaddition reaction **12**, between sydnone (X=I) and B□N. Only few atoms of both reactants are shown for clarity. Dashed line shows steric contacts H...I between halogen atom of sydnone and H atoms attached to one of carbons at the triple bond of alkyne, dotted lines indicate new single bonds forming in this reaction.

4 Conclusions

Extensive DFT calculations on a set of sydnone-alkyne reactions confirmed previously suggested two-step mechanism: cycloaddition followed by retro-Diels-Alder reaction. Since the latter had a tiny barrier, the cycloaddition step was

predicted to be the rate-limiting. For the ensemble of 18 reactions, calculations reproduce activation free energies extracted from experimental reaction rates (k) with the accuracy of 2 kcal/mol. Accounting for solvation effects didn't change the overall trend of activation free energies as a function of substituents. As an alternative, a series of statistical model linking $\log k$ and sydnones structure was built using Support Vector Regression and Multiple Linear Regression machine-learning methods coupled with different types of molecular descriptors; none of them demonstrated a good performance at cross-validation stage.

Detailed analysis of different factors affecting reaction rate variation as a function of substituents revealed particular role of the charge on C3 atom in the sydnone moiety (Figure 1) as well as of the size of the substituent at C3. Large positive charge on C3 enhances reaction rate due to transition state energy lowering, whereas bulky electron withdrawing substituent X at C3 atom leads to steric hindrance with alkyne hydrogens upon the transition state formation. Thus, only sterically unencumbered substituents X allow gaining reaction speed.

Analysis of orbital symmetries and energies shows that in the course of reaction, electrons are transferred from HOMO of alkyne to LUMO of sydnone. The lower LUMO energy, the larger is reaction rate. Since LUMO energy is affected by π -electron withdrawing ability of substituent R, the usage of stronger electron acceptors will favor reaction.

Acknowledgement

The authors thank the Agence Nationale de la Recherche (ANR ClickReal, 2014-2019) for support.

References:

- [1] E.M. Sletten, C.R. Bertozzi, From Mechanism to Mouse: A Tale of Two Bioorthogonal Reactions, *Acc. Chem. Res.* 44 (2011) 666–676.
doi:10.1021/ar200148z.
- [2] R.K. V Lim, Q. Lin, Bioorthogonal chemistry: recent progress and future

- directions, *Chem. Commun.* 46 (2010) 1589–1600. doi:10.1039/B925931G.
- [3] Y. Gong, L. Pan, Recent advances in bioorthogonal reactions for site-specific protein labeling and engineering, *Tetrahedron Lett.* 56 (2015) 2123–2132. doi:10.1016/j.tetlet.2015.03.065.
- [4] D.C. Kennedy, C.S. McKay, M.C.B. Legault, D.C. Danielson, J.A. Blake, A.F. Pegoraro, A. Stolow, Z. Mester, J.P. Pezacki, Cellular Consequences of Copper Complexes Used To Catalyze Bioorthogonal Click Reactions, *J. Am. Chem. Soc.* 133 (2011) 17993–18001. doi:10.1021/ja2083027.
- [5] N.E. Mbuja, J. Guo, M.A. Wolfert, R. Steet, G.-J. Boons, Strain-Promoted Alkyne–Azide Cycloadditions (SPAAC) Reveal New Features of Glycoconjugate Biosynthesis, *ChemBioChem.* 12 (2011) 1912–1921. doi:10.1002/cbic.201100117.
- [6] E.M. Sletten, C.R. Bertozzi, Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality, *Angew. Chemie Int. Ed.* 48 (2009) 6974–6998. doi:10.1002/anie.200900942.
- [7] D.L. Browne, M.D. Helm, A. Plant, J.P.A. Harrity, A Sydnone Cycloaddition Route to Pyrazole Boronic Esters, *Angew. Chemie Int. Ed.* 46 (2007) 8656–8658. doi:10.1002/anie.200703767.
- [8] S. Kolodych, E. Rasolofonjatovo, M. Chaumontet, M. Nevers, C. Créminon, F. Taran, Discovery of Chemoselective and Biocompatible Reactions Using a High-Throughput Immunoassay Screening, *Angew. Chemie Int. Ed.* 52 (2013) 12056–12060. doi:10.1002/anie.201305645.
- [9] S. Specklin, E. Decuypere, L. Plougastel, S. Aliani, F. Taran, One-Pot Synthesis of 1,4-Disubstituted Pyrazoles from Arylglycines via Copper-Catalyzed Sydnone–Alkyne Cycloaddition Reaction, *J. Org. Chem.* 79 (2014) 7772–7777. doi:10.1021/jo501420r.
- [10] E. Decuypere, S. Specklin, S. Gabillet, D. Audisio, H. Liu, L. Plougastel, S. Kolodych, F. Taran, Copper(I)-Catalyzed Cycloaddition of 4-Bromosydnone and Alkynes for the Regioselective Synthesis of 1,4,5-Trisubstituted Pyrazoles, *Org. Lett.* 17 (2015) 362–365.

- doi:10.1021/ol503482a.
- [11] S. Wallace, J.W. Chin, Strain-promoted sydnone bicyclo-[6.1.0]-nonyne cycloaddition, *Chem. Sci.* 5 (2014) 1742–1744. doi:10.1039/C3SC53332H.
- [12] M.K. Narayanam, Y. Liang, K.N. Houk, J.M. Murphy, Discovery of new mutually orthogonal bioorthogonal cycloaddition pairs through computational screening, *Chem. Sci.* 7 (2016) 1257–1261. doi:10.1039/C5SC03259H.
- [13] S. Bernard, D. Audisio, M. Riomet, S. Bregant, A. Sallustrau, L. Plougastel, E. Decuypere, S. Gabillet, R.A. Kumar, J. Elyian, M.N. Trinh, O. Koniev, A. Wagner, S. Kolodych, F. Taran, Bioorthogonal Click and Release Reaction of Iminosydnes with Cycloalkynes, *Angew. Chemie Int. Ed.* 56 (2017) 15612–15616. doi:10.1002/anie.201708790.
- [14] L. Plougastel, O. Koniev, S. Specklin, E. Decuypere, C. Creminon, D.-A. Buisson, A. Wagner, S. Kolodych, F. Taran, 4-Halogeno-sydnes for fast strain promoted cycloaddition with bicyclo-[6.1.0]-nonyne, *Chem. Commun.* 50 (2014) 9376–9378. doi:10.1039/C4CC03816A.
- [15] H. Liu, D. Audisio, L. Plougastel, E. Decuypere, D.-A. Buisson, O. Koniev, S. Kolodych, A. Wagner, M. Elhabiri, A. Krzyczmonik, S. Forsback, O. Solin, V. Gouverneur, F. Taran, Ultrafast Click Chemistry with Fluorosydnes, *Angew. Chemie Int. Ed.* 55 (2016) 12073–12077. doi:10.1002/anie.201606495.
- [16] S.A. Lopez, K.N. Houk, Alkene Distortion Energies and Torsional Effects Control Reactivities, and Stereoselectivities of Azide Cycloadditions to Norbornene and Substituted Norbornenes, *J. Org. Chem.* 78 (2013) 1778–1783. doi:10.1021/jo301267b.
- [17] F. Liu, Y. Liang, K.N. Houk, Theoretical Elucidation of the Origins of Substituent and Strain Effects on the Rates of Diels–Alder Reactions of 1,2,4,5-Tetrazines, *J. Am. Chem. Soc.* 136 (2014) 11483–11493. doi:10.1021/ja505569a.
- [18] H. Tao, F. Liu, R. Zeng, Z. Shao, L. Zou, Y. Cao, J.M. Murphy, K.N. Houk,

- Y. Liang, Origins of halogen effects in bioorthogonal sydnone cycloadditions, *Chem. Commun.* 54 (2018) 5082–5085.
doi:10.1039/C8CC02128G.
- [19] P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas, *Phys. Rev.* 136 (1964) B864--B871. doi:10.1103/PhysRev.136.B864.
- [20] W. Kohn, L.J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, *Phys. Rev.* 140 (1965) A1133–A1138.
doi:10.1103/PhysRev.140.A1133.
- [21] D.N. Laikov, Y.A. Ustynyuk, PRIRODA-04: a quantum-chemical program suite. New possibilities in the study of molecular systems with the application of parallel computing, *Russ. Chem. Bull.* 54 (2005) 820–826.
doi:10.1007/s11172-005-0329-x.
- [22] D.N. Laikov, *Priroda 2011*, (2011). <http://rad.chem.msu.ru/~laikov/>.
- [23] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* 77 (1996) 3865–3868.
doi:10.1103/PhysRevLett.77.3865.
- [24] A. Schäfer, C. Huber, R. Ahlrichs, Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr, *J. Chem. Phys.* 100 (1994) 5829–5835. doi:10.1063/1.467146.
- [25] D.N. Laikov, Fast evaluation of density functional exchange-correlation terms using the expansion of the electron density in auxiliary basis sets, *Chem. Phys. Lett.* 281 (1997) 151–156. doi:10.1016/S0009-2614(97)01206-2.
- [26] J. Tomasi, B. Mennucci, E. Cancès, The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM ab initio level, *J. Mol. Struct. THEOCHEM.* 464 (1999) 211–226.
doi:10.1016/S0166-1280(98)00553-3.
- [27] J. Tomasi, B. Mennucci, R. Cammi, Quantum Mechanical Continuum Solvation Models, *Chem. Rev.* 105 (2005) 2999–3094.
doi:10.1021/cr9904009.

- [28] C.J. Cramer, D.G. Truhlar, An SCF Solvation Model for the Hydrophobic Effect and Absolute Free Energies of Aqueous Solvation, *Science* (80-.). 256 (1992) 213–217. doi:10.1126/science.256.5054.213.
- [29] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J. V Ortiz, J. Cioslowski, D.J. Fox, Gaussian 09 Revision D.01, (2009).
- [30] ChemAxon, Marvin 17.5.2, (2018). <http://www.chemaxon.com>.
- [31] ChemAxon, JChem Calculator Plugins 15.8.3, (n.d.). <http://www.chemaxon.com>.
- [32] T.A. Halgren, Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94, *J. Comput. Chem.* 17 (1996) 490–519. doi:10.1002/(sici)1096-987x(199604)17:5/6<490::aid-jcc1>3.0.co;2-p.
- [33] T.A. Halgren, Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions, *J. Comput. Chem.* 17 (1996) 520–552. doi:10.1002/(sici)1096-987x(199604)17:5/6<520::aid-jcc2>3.0.co;2-w.
- [34] T.A. Halgren, Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94, *J. Comput. Chem.* 17 (1996) 553–586.

- doi:10.1002/(SICI)1096-987X(199604)17:5/6<553::AID-JCC3>3.0.CO;2-T.
- [35] J. Stewart, Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements, *J. Mol. Model.* 13 (2007) 1173–1213. doi:10.1007/s00894-007-0233-4.
- [36] J.J.P. Stewart, MOPAC2011, (2011).
- [37] P. Geerlings, F. De Proft, W. Langenaeker, Conceptual Density Functional Theory, *Chem. Rev.* 103 (2003) 1793–1874. doi:10.1021/cr990029p.
- [38] R.G. Parr, L. v. Szentpály, S. Liu, Electrophilicity Index, *J. Am. Chem. Soc.* 121 (1999) 1922–1924. doi:10.1021/ja983494x.
- [39] A. Varnek, D. Fourches, F. Hoonakker, V.P. Solov'ev, Substructural fragments: an universal language to encode reactions, molecular and supramolecular structures, *J. Comput. Aided. Mol. Des.* 19 (2005) 693–703. doi:10.1007/s10822-005-9008-0.
- [40] C.-C. Chang, C.-J. Lin, LIBSVM: A Library for Support Vector Machines, *ACM Trans. Intell. Syst. Technol.* 2 (2011) 27:1--27:27. doi:10.1145/1961189.1961199.
- [41] D. Horvath, J. Brown, G. Marcou, A. Varnek, An Evolutionary Optimizer of libsvm Models, *Challenges.* 5 (2014) 450–472. doi:10.3390/challe5020450.
- [42] M. Karelson, U. Maran, Y. Wang, A.R. Katritzky, QSPR and QSAR Models Derived Using Large Molecular Descriptor Spaces. A Review of CODESSA Applications, *Collect. Czechoslov. Chem. Commun.* 64 (1999) 1551–1571.
- [43] R.F.W. Bader, *Atoms in Molecules: A Quantum Theory*, Clarendon Press, Oxford, 1994. <http://books.google.co.in/books?id=tyVpQgAACAAJ>.
- [44] A.R. Katritzky, M. Karelson, R. Petrukhin, CODESSA 3 program, (n.d.). <http://www.semichem.com/default.php>.
- [45] R. Peverati, D.G. Truhlar, Quest for a universal density functional: the accuracy of density functionals across a broad spectrum of databases in chemistry and physics, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 372 (2014).
- [46] H. Drucker, C.J.C. Burges, L. Kaufman, A. Smola, V. Vapnik, Support

- vector regression machines, in: M.C. Mozer, J.I. Jordan, J.I. Petsche (Eds.), *Adv. Neural Inf. Process. Syst.*, MIT Press, 1997: pp. 155–161.
doi:10.1.1.10.4845.
- [47] L.R. Domingo, M.J. Aurell, P. Pérez, R. Contreras, Quantitative characterization of the global electrophilicity power of common diene/dienophile pairs in Diels–Alder reactions, *Tetrahedron*. 58 (2002) 4417–4423. doi:[https://doi.org/10.1016/S0040-4020\(02\)00410-6](https://doi.org/10.1016/S0040-4020(02)00410-6).
- [48] L.R. Domingo, J.A. Saez, Understanding the mechanism of polar Diels–Alder reactions, *Org. Biomol. Chem.* 7 (2009) 3576–3583. doi:10.1039/B909611F.
- [49] Y. Liang, J.L. Mackey, S.A. Lopez, F. Liu, K.N. Houk, Control and Design of Mutual Orthogonality in Bioorthogonal Cycloadditions, *J. Am. Chem. Soc.* 134 (2012) 17904–17907. doi:10.1021/ja309241e.
- [50] F. Liu, Y. Liang, K.N. Houk, Bioorthogonal Cycloadditions: Computational Analysis with the Distortion/Interaction Model and Predictions of Reactivities, *Acc. Chem. Res.* 50 (2017) 2297–2308.
doi:10.1021/acs.accounts.7b00265.

