



Cite this: *Med. Chem. Commun.*,  
2019, **10**, 263

Received 2nd July 2018,  
Accepted 7th December 2018

DOI: 10.1039/c8md00327k

rsc.li/medchemcomm

## Design and characterization of a heterocyclic electrophilic fragment library for the discovery of cysteine-targeted covalent inhibitors†

A. Keeley, P. Ábrányi-Balogh and G. M. Keserű \*

A fragment library of electrophilic small heterocycles was characterized through cysteine-reactivity and aqueous stability tests that suggested their potential as covalent warheads. The analysis of theoretical and experimental descriptors revealed correlations between the electronic properties of the heterocyclic cores and their reactivity against GSH that are helpful in identifying suitable fragments for cysteines with specific nucleophilicity. The most important advantage of these fragments is that they show only minimal structural differences from non-electrophilic counterparts. Therefore, they could be used effectively in the design of targeted covalent inhibitors with minimal influence on key non-covalent interactions.

### Introduction

Heterocyclic rings and ring systems are among the key building blocks of drugs and drug-like compounds. It has been shown that 80 of the top 100 most frequently used ring systems among small-molecule drugs are nitrogen, oxygen and/or sulphur heterocycles.<sup>1</sup> The abundance of heterocycles can be traced back to HTS<sup>2</sup> and fragment-screening decks<sup>3</sup> that contribute significantly to lead discovery efforts. Most of the heterocyclic scaffolds form non-covalent interactions with the target; however, the emergence of targeted covalent inhibitors (TCIs) highlights their use as potential electrophilic warheads. The actual design paradigm of covalent inhibitors primarily involves targeting cysteines by attaching electrophilic functional groups to known non-covalent scaffolds.<sup>4</sup> These reactive warheads are typically acrylamides, but halo ketones, halo-acetyl amides, isothiocyanates, and maleimides are also prevalent.<sup>5,6</sup> As a consequence of this design strategy, the non-covalent ligand is equipped with at least three (isothiocyanate: -NCS), but most often more than four, extra atoms (e.g., the most common acrylamide: -CH=CH-C(O)NR<sub>2</sub>). This extension affects the molecular recognition and the binding mode of the non-covalent scaffold and requires the iterative optimization of electrophile geometry and linker length to engage the targeted cysteine successfully.

Starting from relevant heterocyclic cores, here we design and characterize an electrophilic fragment library having minimal influence on the potential non-covalent interactions. Based on the well-known electron-withdrawing character of heterocycles,<sup>7</sup> our intention was to turn them into fragment electrophiles with the introduction of the smallest available substituents (max. 1–2 atoms). We hypothesized that electron-withdrawing heterocycles activate the small electrophilic substituents and yield warheads suitable for cysteine-targeting covalent inhibitors. Testing this idea, we collected a wide range of five- and six-membered nitrogen-containing heterocycles combined with a selection of small electrophilic warheads. Library members were then subjected to detailed characterization that included the assessment of their cysteine reactivity, specificity and aqueous stability.

Screening electrophilic fragments is now an emerging strategy in both ligand discovery<sup>8</sup> and target identification,<sup>9</sup> and they can serve as starting points in TCI programs.<sup>10</sup> We believe that the library described here might facilitate the development of TCIs by replacing their heterocyclic scaffold with one of our electrophilic heterocycles. This approach would allow the precise positioning of the reactive group toward a catalytic/non-catalytic protein nucleophile in the proximity of the binding site<sup>11</sup> while maintaining the key non-covalent interactions.

### Results and discussion

The heterocyclic fragment library was composed of five- and six-membered heterocycles, particularly pyridine, pyrimidine, pyrazine, imidazole, pyrazole, oxazole, isoxazole and thiazole, while the corresponding six warheads were the halogens Cl, Br, and I reacting through aromatic nucleophilic substitution

Medicinal Chemistry Research Group, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Magyar tudósok krt 2, H-1117 Budapest, Hungary. E-mail: keseru.gyorgy@ttk.mta.hu

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8md00327k



( $S_NAr$ ), and the two-atom nitrile, vinyl and ethynyl groups reacting through nucleophilic addition ( $Ad_N$ ) to cysteines.

We planned to investigate the influence of the different heterocycles and the effect of warhead positions by evaluat-

ing the fragment's reactivity experimentally. Moreover, we aimed to identify theoretical descriptors supporting the design of new fragments with tailored reactivity. Furthermore, we intended to analyse the reactivity differences between the

**Table 1** GSH reactivity of heterocyclic electrophiles measured at concentrations of 250  $\mu$ M. Library members having  $t_{1/2(GSH)} > 72$  h were considered inactive and are shown in Table S1

ID	Structure	$k_{GSH}$	$t_{1/2(GSH)}$ [h]	ID	Structure	$k_{GSH}$	$t_{1/2(GSH)}$ [h]
A1		0.0100	69.4	K6		0.1412	4.9
A2		0.0100	69.2	L6		0.4044	1.7
A5		0.6887	1.0	N1		0.1533	4.5
B4		0.0176	41.0	N3		6.0573	0.1
C3		0.0100	69.0	N4		1.4615	0.5
C4		0.8837	0.8	N5		0.0153	45.2
C5		2.4150	0.3	N6		0.0127	54.6
C6		0.2900	2.4	O4		0.693	1.0
D4		0.3138	2.2	P6		0.1284	5.4
D6		0.0151	46.8	Q2		0.011	63.0
E4		0.0147	47.1	Q5		0.0156	44.5
G4		0.0314	22.5	Q6		0.0130	53.1
G5		0.0414	16.9	R3		0.7872	0.9
H3		0.116	6.0	R4		0.0868	8.0
H4		0.0132	52.3	R5		0.2583	2.7
I6		0.0144	48.0				



halogen atoms for the  $S_NAr$  and that of the nitrile, vinyl and ethynyl groups for the  $Ad_N$  reactions. We have compiled the library from 84 electrophilic heterocycles, out of which 27 were synthesized in our laboratory. The pyridine, pyrimidine, imidazole and pyrazole rings were substituted at three different positions; the oxazole and thiazole rings at two positions; and finally the pyrazine and isoxazole at one available position. Notably, in some cases, the 1-substituted compound was not available; therefore, 2-chlorobenzoxazole, 2-bromo- and 2-vinyl-5-phenyloxazole, 3,5-dimethylisoxazoles and 2-chlorobenzothiazole were considered (see Tables 1 and S1† for the chemical structures of the library).

First, the stability and reactivity of the library members were investigated in a GSH-based assay (Fig. 1) using HPLC-MS (Fig. 1 IIa) or NMR-based kinetic methods (Fig. 1 IIb).<sup>12</sup> We measured the decreasing amount of the electrophilic fragment up to 72 h in two parallel measurements. The aqueous stability of the compounds was characterized by the fragment half-life calculated from the equation  $t_{1/2(\text{deg})} = \ln 2/k_{\text{deg}}$ , where the degradation rate constant for auto-degradation ( $k_{\text{deg}}$ ) was calculated by linear regression of the measured datapoints in the absence of GSH. Thiol reactivity was assessed by measuring fragment depletion with a large excess of GSH<sup>12</sup> that provided the rate constant  $k_{\text{deg}+\text{GSH}}$  as the sum of the thiol-reactivity and the degradation. The GSH reactivity of the electrophile was then calculated from these two rate constants as  $k_{\text{GSH}} = k_{\text{deg}+\text{GSH}} - k_{\text{deg}}$ . The GSH half-life ( $t_{1/2(\text{GSH})}$ ) was determined from the  $k_{\text{GSH}}$  thiol-reactivity rate constant.

Stability data confirmed that all of the compounds showed the appropriate stability ( $>1$  h) required for biological testing (Table S1†).<sup>3</sup> Notably, the less stable species were found in isoxazoles ( $t_{1/2(\text{deg})} < 17$  h), while 2-chlorobenzoxazole and 2-chlorobenzothiazole also had  $t_{1/2(\text{deg})} < 24$  h.

The results of the GSH reactivity assay revealed that heterocyclic electrophiles cover a wide range of thiol reactivity (Tables 1 and S1†). The library contained fragments reacting under 1 h (C4 (0.8 h), C5 (0.3 h), N3 (0.1 h), N4 (0.5 h), Table 1) to compounds considered practically non-

reactive, with  $t_{1/2(\text{GSH})} > 72$  h (shown in Table S1†). Since the library is intended for use in labelling cysteine nucleophiles, we considered two major subsets of the compounds. The first subset involves compounds with nitrile, vinyl and ethynyl substituents that react through nucleophilic addition ( $Ad_N$ ). The second set consists of halogenated derivatives that label cysteine in nucleophilic substitution reactions ( $S_NAr$ ). Comparing the six-membered fragments, we found that halogenated compounds showed weak reactivity (for A1, A2, A3, B2, C1, C2, C3, D2, E1, F1, F2, G1,  $t_{1/2(\text{GSH})} > 69$  h, and B1, B3, D1, D3, F3, G2, G3 were essentially non-reactive). In contrast, the cyano-, vinyl- and ethynylpyridines at position 4 (C4 (0.8 h), C5 (0.3 h), C6 (2.4 h), respectively) reacted quickly (Table 1). In the pyrimidine subset equipped with CN or ethynyl groups, position 2 between the two nitrogens gave the shortest half-lives (D4 (2.2 h), D6 (46.8 h) (Table 1)). Among vinylated six-membered heterocycles, pyridines were the most potent electrophiles (A5 (1.0 h), C5 (0.3 h), Table 1).

Taking a closer look at the cyano derivatives, position 4 of the pyridine (C4, 0.8 h) and position 2 of the pyrimidine (D4, 2.2 h) and the pyrazine (G4, 22.5 h) rings were most reactive (Table 1). Focusing on the five-membered heterocycles, among the imidazole derivatives, only 2-iodoimidazole (H3, 6.0 h) showed considerable reactivity, and among pyrazoles, only 3-ethynyl- (K6, 4.9 h) and 4-ethynylpyrazole (L6, 1.7 h) were reactive (Table 1). In the case of the oxazole core, the 2-iodo- (N3, 0.1 h), 2-cyano- (N4, 0.5 h) and 4-cyanooxazole (O4, 1.0) showed remarkable reactivity (Table 1). From the 3,5-dimethylisoxazoles, only the 4-ethynyl derivative (P6, 5.4 h) was reactive (Table 1). Thiazoles were, in particular, the most reactive heterocycles in the five-membered group. Their nitrile and vinyl derivatives were most active when located between the heteroatoms at position 2 (R4 (8.0 h), R5 (2.7 h), respectively, Table 1). In contrast, bromine and ethynyl derivatives (Q2 (63.0 h) and Q6 (53.1 h), respectively) performed best at position 5 (Table 1).

Next, we analysed the reactivity trends quantitatively using computed descriptors and experimental  $(\log)t_{1/2(\text{GSH})}$  values

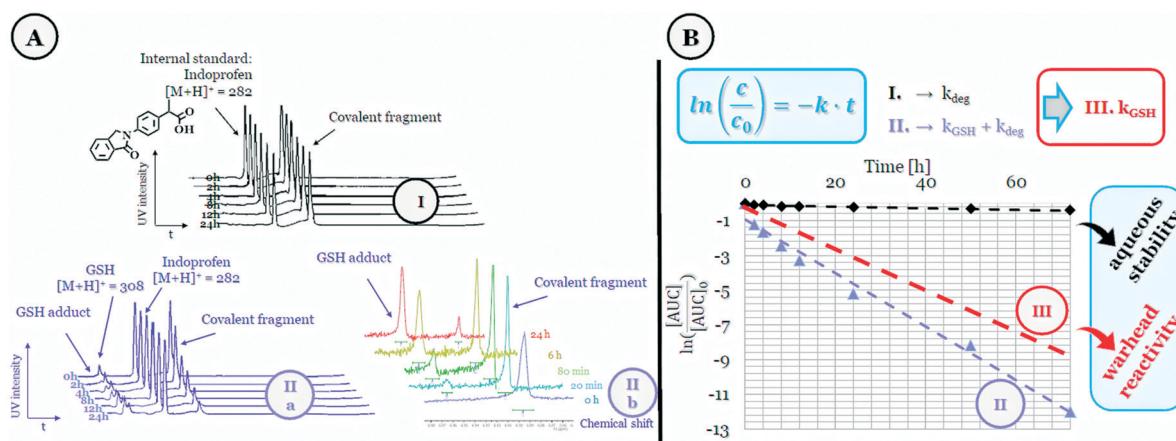


Fig. 1 Representation of the (A) HPLC- (IIa) and NMR-based (IIb) thiol-reactivity studies with the (B) corresponding calculations.

(see Table S2†). The Gaussian09 program package with the B3LYP/6-311++(2d,2p) method and basis set was used to calculate the HOMO ( $\varepsilon_H$ ) and LUMO energies ( $\varepsilon_L$ ) and electron distribution on the reacting carbon atom ( $\rho_c^0$  and  $\rho_c^{-1}$ ) at the neutral and also at the -1 charged state, based on the atomic charge distributions in terms of natural population analysis (NPA). From these values, we have calculated the chemical hardness ( $\eta = \varepsilon_L - \varepsilon_H$ ), electronic chemical potential ( $\mu = (\varepsilon_L + \varepsilon_H)/2$ ), Parr-index (global electrophilicity,  $\omega = \mu^2/2\eta$ ), Fukui-function (frontier function,  $f^+(\vec{r}) = \rho_c^{-1} - \rho_c^0$ ) and local electrophilicity index ( $\omega(\vec{r}) = \omega \cdot f^+(\vec{r})$ ).<sup>13</sup> Furthermore, we have computed the transition state enthalpies and Gibbs free energies for the reaction of the chloro- and vinyl-derivatives (modelling S<sub>N</sub>Ar and Ad<sub>N</sub> reactions, respectively) with the MeS<sup>-</sup> anion as a cysteine surrogate. Inductive sigma constants for the heterocyclic rings were calculated by ACD/Percepta (see Table S2†).

Calculation of the Pearson and Spearman correlation coefficients between each of the descriptors and experimental GSH half-lives revealed no significant correlation when considering the whole library. Taking into account, however, the reaction types between the electrophilic fragments and the targeted nucleophile, we identified some interesting trends. In the case of halogen derivatives reacting in nucleophilic substitutions, we found that the GSH half-life decreases with increasing electronic chemical potential for the halogen set (Fig. 2). Changes in the electronic chemical potential are opposite that of the electron affinity<sup>14</sup> and related to the corresponding HOMO energies. This explains why a higher potential is associated with more reactive species with higher  $\varepsilon_H$  and with low  $(\log)t_{1/2}(\text{GSH})$ . Notably, the more electrophilic fragments with higher HOMO energies are more reactive towards the nucleophilic cysteine.

In the case of ethynyl-substituted fragments reacting in nucleophilic additions, we found that the GSH half-life decreases with increasing atomic charge, which indicates decreasing electron density ( $\rho_c^0$ ) for the ethynyl set (Fig. 3). This trend showed that the lower electron density is advantageous for thiol reactivity, since readily ionisable electrophiles and the resulting electron-poor carbon atoms can be considered more reactive towards the thiol nucleophile. This analysis revealed that the prediction of the GSH reactivity requires different descriptors depending on the reaction mechanism. In the case of Ad<sub>N</sub> reactions, the warhead itself does not influence the reactivity, but the local electron distribution on the reacting carbon caused by the electron-withdrawing effect of the aromatic substituent is determining. Moreover, the negative charge is localised on the warhead in the transition state (TS), and the intermediate in Ad<sub>N</sub> reactions implicates the use of a local electronic descriptor. Taking a closer look at S<sub>N</sub>Ar reactions, the aromatic core and the halogen atom should be treated together due to the characteristic leav-

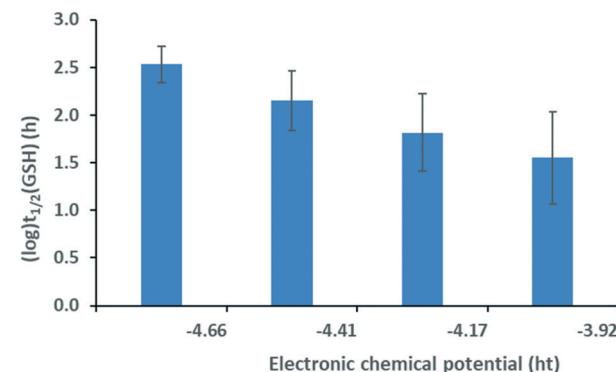


Fig. 2 Logarithm of the GSH half-life vs. electronic potential for the halogen set. Standard errors are also shown.

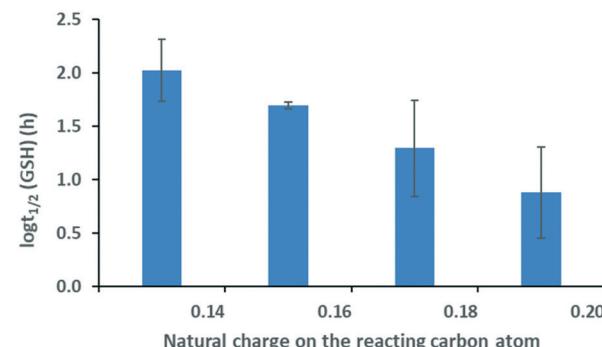


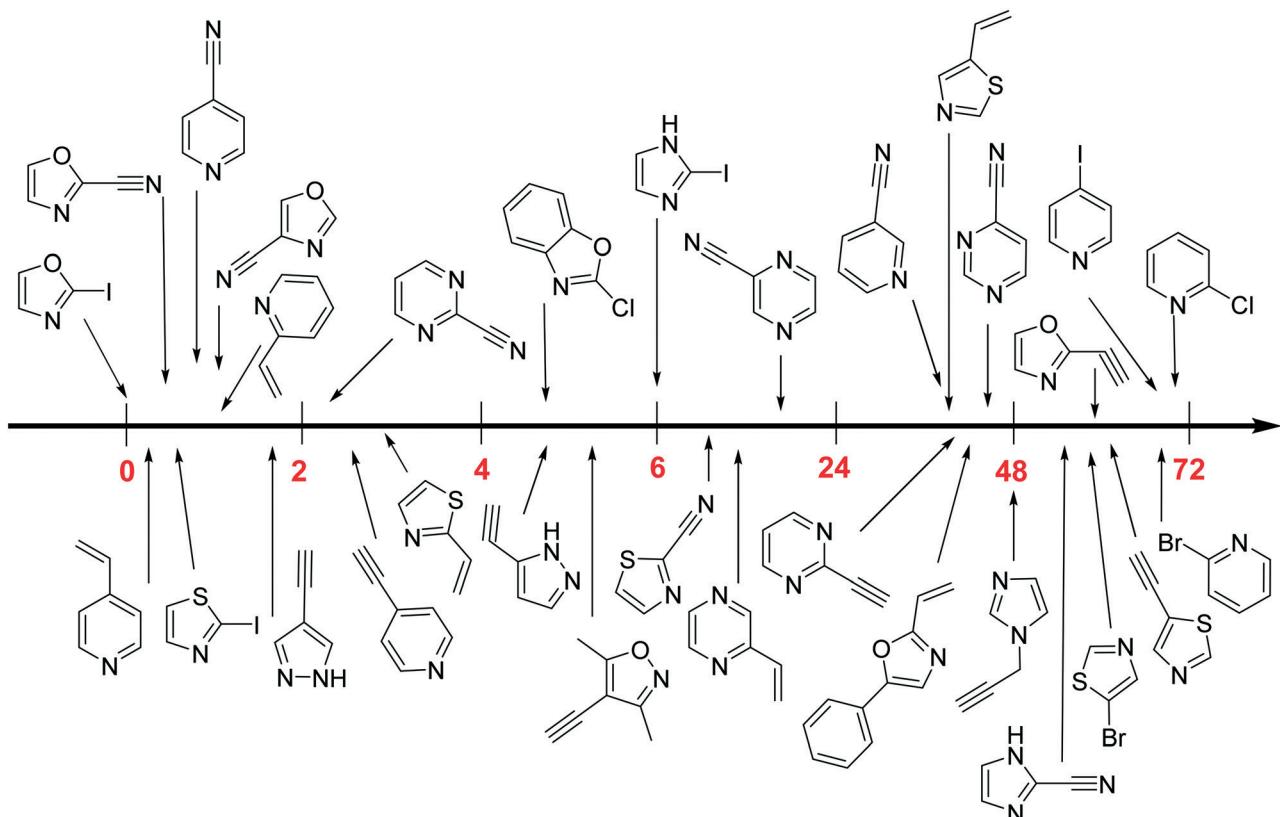
Fig. 3 Logarithm of the GSH half-life vs. atomic charge distribution for the ethynyl set. Standard errors are also shown.

ing group effect. Furthermore, the negative charge is dissipated in the TS. These two factors indicate that a global descriptor is more accurate for the prediction of the reactivity.

## Conclusions

Recent strategies of covalent drug design have been focusing on the optimization of the non-covalent scaffold followed by the attachment of extended warheads, commonly Michael acceptors. In this work, we have developed a heterocyclic electrophilic fragment library from the combination of eight five- or six-membered heterocycles and six small one- or two-atom warheads acting through aromatic nucleophilic substitution or nucleophilic addition. The systematic characterization of this library resulted in compounds with appropriate stability and a wide range of thiol reactivity (Fig. 4) that allow the selection of the most appropriate heterocyclic warhead depending on the nucleophilicity of the targeted cysteine. These heterocyclic electrophiles could be used to replace aromatic moieties in known non-covalent ligands and could present a useful alternative design strategy aiming at targeted covalent inhibitors. The characterized library is available for screening against relevant targets upon request from the authors.





**Fig. 4** Measured GSH reactivity of diverse heterocycle moieties expressed in  $t_{1/2}$  (h). A high value indicates poor electrophilicity, whereas small values are attributed to more reactive electrophiles. Not represented are compounds whose  $t_{1/2}$  (h) is  $> 72$  h.

## Conflicts of interest

The authors declare no competing financial interests.

## Acknowledgements

This study was supported by the MSCA ITN FRAGNET (project 675899) grant to G. M. Keserű and A. Keeley and by the Hungarian Science Foundation OTKA (project K116904) to G. M. Keserű. P. Ábrányi-Balogh was supported by the postdoctoral fellowship of the Hungarian Academy of Sciences and Hungarian Science Foundation OTKA (PD124598) grant. The authors are grateful to ACD/Labs and Chemicro LLC making the Percepta platform available for descriptor calculations.

## Notes and references

- 1 R. D. Taylor, M. MacCoss and A. D. G. Lawson, *J. Med. Chem.*, 2014, **57**, 5845.
- 2 G. M. Keserű and G. M. Makara, *Nat. Rev. Drug Discovery*, 2009, **8**, 203.
- 3 G. M. Keserű, D. A. Erlanson, G. G. Ferenczy, M. M. Hann, C. W. Murray and S. D. Pickett, *J. Med. Chem.*, 2016, **59**, 8189.
- 4 J. Singh, R. C. Petter, T. A. Baillie and A. Whitty, *Nat. Rev. Drug Discovery*, 2011, **10**, 307.
- 5 T. A. Baillie, *Angew. Chem., Int. Ed.*, 2016, **55**, 13408.
- 6 D. A. Shannon and E. Weerapana, *Curr. Opin. Chem. Biol.*, 2015, **24**, 18.
- 7 J. A. Joule and K. Mills, *Heterocyclic Chemistry*, Blackwell Publishing, Chichester, 5th edn, 2010.
- 8 K. M. Backus, B. E. Correia, K. M. Lum, S. Forli, B. D. Horning, G. E. González-Páez, S. Chatterjee, B. R. Lanning, J. R. Teijaro, A. J. Olson, D. W. Wolan and B. Cravatt, *Nature*, 2016, **534**, 570.
- 9 G. G. Parker, A. Galmozzi, Y. Wang, B. E. Correira, K. Sasaki, C. M. Joslyn, A. S. Kim, C. L. Cavallaro, R. M. Lawrence, S. R. Johnson, I. Narvaiza, E. Saez and B. F. Cravatt, *Cell*, 2017, **168**, 527.
- 10 R. Lonsdale and R. A. Ward, *Chem. Soc. Rev.*, 2018, **47**, 3816.
- 11 S. G. Kathman and A. V. Statsyuk, *Med. Chem. Commun.*, 2016, **7**, 576.
- 12 M. E. Flanagan, J. A. Abramite, D. P. Anderson, A. Aulabaugh, U. P. Dahal, A. M. Gilbert, C. Li, J. Montgomery, S. R. Oppenheimer, T. Ryder, B. P. Schuff, D. P. Uccello, G. S. Walker, Y. Wu, M. F. Brown, J. M. Chen, M. M. Hayward, M. C. Noe, R. S. Obach, L. Philippe, V. Shanmugashundaram, M. J. Shapiro, J. Starr, J. Stroh and Y. Che, *J. Med. Chem.*, 2014, **57**, 10072.
- 13 L. R. Domingo, M. Ríos-Gutiérrez and P. Pérez, *Molecules*, 2016, **21**, 748.
- 14 M. Franco-Pérez, J. L. Gázquez, P. W. Ayers and A. Vela, *J. Chem. Phys.*, 2015, **143**, 154103.