

## Ligand–Protein Interactions

 $\pi$ -Hole Interactions Involving Nitro Aromatic Ligands in Protein StructuresAntonio Bauzá,<sup>[a]</sup> Antonio Frontera,<sup>\*[a]</sup> and Tidlo Jonathan Mooibroek<sup>\*[b]</sup>

**Abstract:** Studying noncanonical intermolecular interactions between a ligand and a protein constitutes an emerging research field. Identifying synthetically accessible molecular fragments that can engage in intermolecular interactions is a key objective in this area. Here, it is shown that so-called “ $\pi$ -hole interactions” are present between the nitro moiety in nitro aromatic ligands and lone pairs within protein structures (water and protein carbonyls and sulfurs). Ample structural evidence was found in a PDB analysis and computations reveal interaction energies of about  $-5 \text{ kcal mol}^{-1}$  for

ligand–protein  $\pi$ -hole interactions. Several examples are highlighted for which a  $\pi$ -hole interaction is implicated in the superior binding affinity or inhibition of a nitro aromatic ligand versus a similar non-nitro analogue. The discovery that  $\pi$ -hole interactions with nitro aromatics are significant within protein structures parallels the finding that halogen bonds are biologically relevant. This has implications for the interpretation of ligand–protein complexation phenomena, for example, involving the more than 50 approved drugs that contain a nitro aromatic moiety.

## Introduction

Noncovalent interactions largely determine how molecules interact with one another.<sup>[1]</sup>  $\sigma$ -Hole interactions,<sup>[2]</sup> such as hydrogen bonding, figure prominently among identifiable interactions that bear biological significance.<sup>[3]</sup>  $\sigma$ -Hole interaction can also occur with other nonmetals,<sup>[4]</sup> such as in the case of halogen bonding.<sup>[5]</sup> The biological significance of halogen bonding has become increasingly clear since about 2004<sup>[6]</sup> and systematic evaluations of the Protein Data Bank (PDB)<sup>[7]</sup> revealed that halogen bonding is orthogonal to hydrogen bonding from a structural perspective.<sup>[5a,6,8]</sup> Moreover, halogen bonding has been exploited in systematic drug discovery,<sup>[9]</sup> used to control the conformational isomer of a four-stranded DNA junction<sup>[10]</sup> and to control the conformation of peptides.<sup>[11]</sup>

The impact of a novel type of interaction on molecular biology begs the question what other yet underexplored noncovalent interactions await a similar exploitation. In this context,  $\pi$ -holes<sup>[2c,12]</sup> found on carbonyls<sup>[13]</sup> and  $\pi$ -acidic aromatics<sup>[14]</sup> must

be highlighted as known electron deficient moieties that interact favorably with electron rich partners.<sup>[15]</sup> It has been disclosed that  $\pi$ -holes are also found on nitro-derivatives (R-NO<sub>2</sub>).<sup>[16]</sup> Nitrate anions can even act as  $\pi$ -holes in crystal structures of proteins and small molecules when the negative charge is dampened by interactions such as hydrogen bonding<sup>[17]</sup> or coordination to a metal.<sup>[18]</sup> The gas-phase complex between trimethylamine and nitroethane has been recently elucidated with rotational spectroscopy, revealing that the  $\pi$ -hole bonding geometry prevails over competing arrangements involving C–H...O/N hydrogen bonding.<sup>[19]</sup>

The potency of a  $\sigma$ - or  $\pi$ -hole can easily be estimated by inspection of molecular electrostatic potential maps, according to which a more positive value is predictive of a stronger interaction. This potential is computed at  $+25 \text{ kcal mol}^{-1}$  for methylamine and  $+31 \text{ kcal mol}^{-1}$  for Br<sub>2</sub>, which are good hydrogen- and halogen-bond donors, respectively (see Figure 1). The moderate potential of the  $\pi$ -hole on nitrobenzene (**3**,  $+13 \text{ kcal mol}^{-1}$ ) is readily amplified by hydrogen bonding to water (**4**,  $+19 \text{ kcal mol}^{-1}$ ) or amides (**5**,  $+30 \text{ kcal mol}^{-1}$ ).

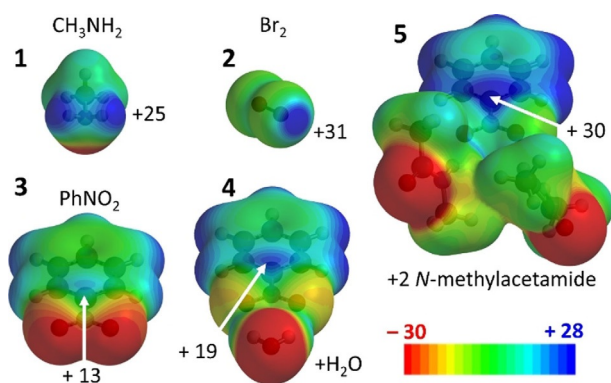
The  $\pi$ -hole bonding geometry of **3** or **4** with some O- or S-bearing molecules (e.g., water and dimethylthioether) already indicate an energy contribution up to about  $-5 \text{ kcal mol}^{-1}$  (see section 1.2. of the Supporting Information for details). Moreover, displacing the water molecule in a [3...OH<sub>2</sub>]  $\pi$ -hole bonding geometry away from N hardly impacts the interaction energy ( $<11\%$ ) below 0.25 Å displacement. Given these computational predictions, it was wondered if  $\pi$ -hole interactions with nitro aromatic ligands currently might have structural and/or functional significance in proteins. This is particularly interesting because nitro aromatics are synthetically accessible and incorporated within many approved drugs,<sup>[20]</sup> including the inexpensive and abundantly used antibiotic chlorampheni-

[a] Dr. A. Bauzá, Prof. A. Frontera  
Department of Chemistry, Universitat de les Illes Balears  
Ctra. de Valldemossa km 7.5, 07122 Palma (Balears) (Spain)  
E-mail: toni.frontera@uib.es

[b] Dr. T. J. Mooibroek  
van 't Hoff Institute for Molecular Sciences, Universiteit van Amsterdam  
Science Park 904, 1098 XH Amsterdam (The Netherlands)  
E-mail: t.j.mooibroek@uva.nl

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:  
<https://doi.org/10.1002/chem.201903404>.

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.



**Figure 1.** Electrostatic potential maps (MEPs) of model compounds. MEPS shown at isodensity surface of 0.002 a.u. of some  $\sigma$ -holes (1 and 2) and  $\pi$ -holes (3–5) with their electropositive potential maps in  $\text{kcal mol}^{-1}$  (MP2/aug-cc-pVDZ//B3LYP/6-31 + G\*, see also section 1.1. of the Supporting Information).

col,<sup>[21]</sup> psychoactive drugs like the valium derivative clonazepam,<sup>[22]</sup> the nonsteroidal antiandrogen flutamide<sup>[23]</sup> used to treat prostate cancer, venetoclax<sup>[24]</sup> used to treat chronic lymphocytic leukemia, dantrolene,<sup>[25]</sup> which is the only clinically available drug to treat malignant hypothermia, and nefidipine,<sup>[26]</sup> which is used to manage angina.

The combined PDB study and computational analyses reported herein provide ample evidence to support the hypothesis that  $\pi$ -hole interactions with nitro aromatic ligands in protein structures have structural and functional significance.

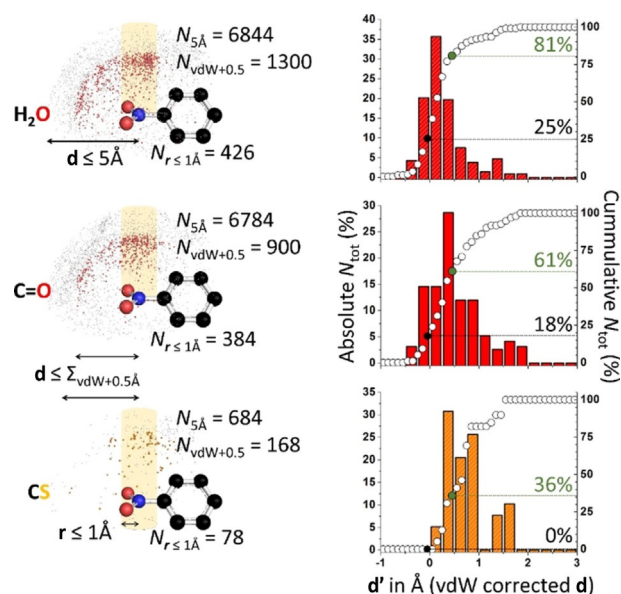
## Results

### Numerical PDB analysis

An initial dataset was retrieved from the PDB for which the distance ( $d$  in Figure 2, top left) between the N atom of any of the >1500 nitro aromatic ligands and some interacting atom was 5 Å or less (see section 2 of the Supporting Information for full technical details). The interacting atom could be the O atom of a water molecule (6844 hits, top left in Figure 2), a carbonyl moiety (middle left, 6784 hits), or any carbon-bound S atom (bottom left, 684 hits).

All these data are shown as small grey, red, or orange spheres in the left-hand side of Figure 2 and are confined within a hemisphere with a radius of 5 Å radius (hence the label " $N_{5\text{Å}}$ " in Figure 2). This hemisphere is centered on the N atom of a model nitro aromatic, which is shown in ball-and-stick representation. The model nitro aromatic was reconstructed from the actual data, as is detailed in section 2.2 of the Supporting Information. Data that enveloped the central N atom at a distance  $d \leq$  the sum of the van der Waals (vdW) radii of N+O(or S)+0.5 Å (i.e., " $N_{\text{vdW}+0.5}$ " in Figure 2) are colored in red (for O) or orange (for S). These subsets contain significant amounts of data and are distributed mainly above (/below) the  $-\text{NO}_2$  moiety.

To probe the distribution of data directly above (/below) the N atom, the data characterized by a parallel displacement from



**Figure 2.** Summary of the PDB data for nitro arenes interacting with a water O atom (top), a carbonyl O atom (middle) or a carbon-bound S atom (bottom). Left: distribution of data around the average model nitro arene (centered on the N atom); right: distribution of data with  $r \leq 1$  Å plotted as a function of  $d'$ .  $r$  = the parallel displacement parameter away from N in the  $\text{CNO}_2$  plane;  $d'$  = the van der Waals corrected  $\text{nitroN} \cdots \text{O/S}$  distance  $d$  [i.e.,  $d - \text{vdW}(\text{N}) - \text{vdW}(\text{O/S})$ ]. See the main text and section 2 of the Supporting Information for details.

the N atom in the nitro fragment of  $r \leq 1$  Å (see the bottom left) was further scrutinized. The graphs shown on the right-hand side of Figure 2 display the absolute (colored bars) and cumulative (white circles) percentages of data with  $r \leq 1$  Å as a function of  $d'$ , which is the van der Waals corrected N $\cdots$ O(or S) distance  $d$  [i.e.,  $d - \text{vdW}(\text{N}) - \text{vdW}(\text{O/S})$ ]. For the data involving water O atoms ( $N_{r \leq 1\text{Å}} = 426$ , top) a Gaussian-like distribution of data is observed and 25% of all hits are apparently involved in overlapping van der Waals shells (black circle and dotted line), which is a clear indicator for a bonding interaction. Loosening this rather strict criterion to data 0.5 Å longer than this benchmark (green circle and dotted line) even encompasses 81% of the data. Hits involving carbonyl O atoms resulted in a wider distribution of data and less van der Waals overlap (18%,  $N_{r \leq 1\text{Å}} = 384$ , middle right; these data might involve some  $\pi$ - $\pi$  stacking). Still, 61% of the data falls within the looser  $\Sigma_{\text{vdW}+0.5}$  Å criterion. For carbon bound sulfur atoms (bottom right) there were merely 78 data points. No van der Waals overlap was observed, yet 36% of the data falls within the  $\Sigma_{\text{vdW}+0.5}$  Å criterion. Further attempts to quantify the directionality of this interaction were inconclusive (see section 2.3 of the Supporting Information) due to the diverse steric effects of the nitro aromatic ligands and/or because there were insufficient data points.<sup>[15b, 16a,b, 17]</sup>

### Structural relevance

The above PDB analysis demonstrates that geometries consistent with  $\pi$ -hole interactions involving nitro aromatics are

abundant in protein structures. To assess whether existing literature contains evidence of the structural and/or functional significance of these interactions, all the data characterized by  $r \leq 1 \text{ \AA}$  and  $d \leq \sum_{\text{vdW}} + 0.5 \text{ \AA}$  were inspected manually. These data were found in 110 unique protein structures, as is detailed in section 3.1 of the Supporting Information. Manual inspection of these structures and the articles in which they were published yielded a total of 28 cases in which  $\pi$ -hole interactions appear to bear a structural and/or functional relevance. All these cases were analyzed as detailed in sections 3.2–3.9 of the Supporting Information and the most striking instances are highlighted below.

As is illustrated in Figure 3, three classes of structures were found in which the close proximity of a water ligand or protein amide O atom is highly conserved and indicative of a  $\pi$ -hole interaction. As is shown in the left-hand side of Figure 3, each of the various nitro aromatic ligands within 20 (*Homo sapiens*) cyclin-dependent kinase 2 structures are located in the same binding pocket comprising Gln-131, Ala-144 and Asp-145 (see also Figure S9, Supporting Information).<sup>[27]</sup> One nitrate O atom is H-bonded to H<sub>2</sub>O-743 (not shown) which likely enhances the  $\pi$ -hole potential (estimated at  $+14.3 \text{ kcal mol}^{-1}$  for a water complex). Of the four water molecules that fill the void in between the ligand and residues, water-375 is always in close proximity to the nitrate moiety. In structure PDB 3r28, the distance between this water molecule O atom (red sphere) and the nitrate N atom of ligand **6** (xa0-788, capped sticks) is  $2.98 \text{ \AA}$  (average =  $3.04 \text{ \AA}$ ), which is  $0.09 \text{ \AA}$  within the sum of the van der Waals radii of N and O ( $3.07 \text{ \AA}$ ). The four close contact distances found with water-375 suggest that water-375 acts with both electropositive (H) and electronegative sites (the O lone-pairs), where one interacting partner is the nitro group  $\pi$ -hole.

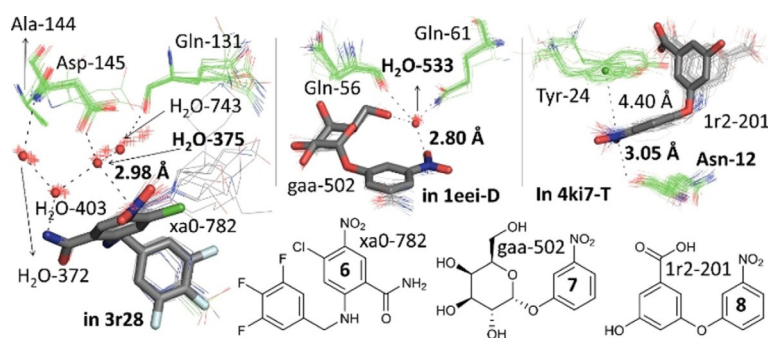
Shown in the middle of Figure 3 is the alignment of 35 nitro aromatic ligands found in nine pentamers of (*Vibrio cholerae*) cholera toxin (CT) B<sup>[28]</sup> that have two close contacts to water H<sub>2</sub>O-553 (see also Figure S10, Supporting Information). One close contact is to the nitro group N atom with a N...O distance of  $2.80 \text{ \AA}$  involving ligand **7** (gaa-502) in chain D of structure 1eei. This is  $0.27 \text{ \AA}$  below the van der Waals benchmark (N+O =  $3.07 \text{ \AA}$ ). This water molecule is also bound by residues Gln-56 and Gln-61, thus displaying a total of four close contact distances. One of the nitro O atoms is H-bonded to the amide NH of Gln-56 (not shown), which will enhance the magnitude of the  $\pi$ -hole (estimated at  $+26.3 \text{ kcal mol}^{-1}$  for the water complex). Interestingly, it is known that CT is 100 times better inhibited by **7** ( $IC_{50} = 1700 \text{ \mu M}$ ) than regular D-galactose.<sup>[28e]</sup> This nitro aromatic ligand has consequently been exploited as an anchoring platform to develop nonspanning bivalent CT inhibitors with  $IC_{50}$ -values up to  $9 \text{ \mu M}$ .<sup>[28e]</sup> In the best inhibitor the binding pocket shown in Figure 3 is preserved (shortest N...O distance =  $3.17 \text{ \AA}$  with ligand BV3 in chain X of structure 1rdp).<sup>[28e]</sup>

The 38 nitro aromatic ligands found within two (*Mycobacterium tuberculosis*) type II dehydroquinases are highlighted in the right-hand side of Figure 3 (see also Figure S11, Supporting Information).<sup>[29]</sup> In these examples, the nitro aromatic group is not H-bonded to any residue, yet consistently sandwiched in between Tyr-24 and the amidic O atom of Asn-12. The shown N...O distance ( $3.05 \text{ \AA}$ ) involving ligand **8** (1r2-201) and Asn-12 is found in chain T of structure 4ki7 ( $\pi$ -hole estimated at  $+22.0 \text{ kcal mol}^{-1}$ ). Such nitro aromatic ligands are potent inhibitors for type II dehydroquinase with up to 92% inhibition and  $K_i = 20 \text{ \mu M}$ .<sup>[29]</sup>

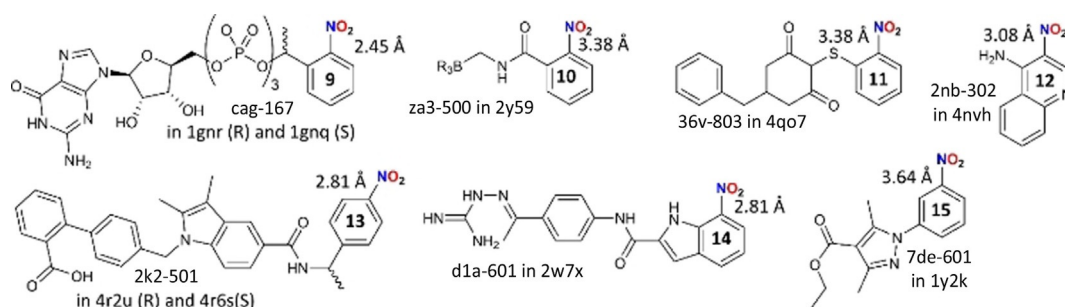
### Functional relevance

Besides these three classes of structures with highly preserved  $\pi$ -hole interaction geometries, various examples were encountered for which the  $\pi$ -hole interaction can be significant in the function of the protein. Shown in Figure 4 is an overview of the ligands involved in these cases and detailed structural information is given in section 3 of the Supporting Information.

Structures 1gnr and 1gnq are (*Homo sapiens*) oncogene product p21<sup>H-RAS</sup> guanine nucleotide binding proteins that have their catalytic domains occupied by enantiomerically pure R/S **9** (ligands cag-167, see Figure 4).<sup>[30]</sup> The binding pockets of both these protein structures are very similar as is detailed in Figure S12 (Supporting Information). However, in structure 1gnr (the R isomer) a water molecule (H<sub>2</sub>O-523) is found in close contact to the nitro group (N...O =  $2.45 \text{ \AA}$ ). This interaction may well contribute to the slower phosphate hydrolysis observed (half-life of 1000 days versus 70 days for the S isomer in 1gnq) by stabilizing the water molecule involved in the hydrolysis reaction.<sup>[30]</sup>



**Figure 3.**  $\pi$ -Hole interactions with nitro aromatic ligands found within protein–ligand complexes. Left: Perspective view of 20 aligned nitro arene ligands (grey wires) and relevant residues (green wires) as found within 20 cyclin-dependent kinases 2 structures. The ligand shown in capped grey sticks is ligand **6** found in PDB 3r28. The red spheres represent the center of mass of the waters 372, 375, 403 and 743. Middle: Perspective view of 35 aligned nitro arene ligands (grey wires) and relevant residues (green wires) as found within 9 cholera toxin B pentameric structures. The ligand shown in capped grey sticks is ligand **7** found in chain D of 1eei. The red sphere represents the center of mass of all water-553 O-atoms. Right: Perspective view of 38 aligned nitro arene ligands (grey wires) and relevant residues (green wires) as found within 2 type II dehydroquinase structures. The ligands shown in capped grey sticks is ligand **8** found in chain T of 4ki7. The green sphere represents the center of mass of all aromatic rings of the Tyr-24 residues.



**Figure 4.** Overview of nitro aromatic ligands 9–15 found within the PDB that have their nitro group N atom in close proximity (as indicated) to a water O atom (9), a carbonyl O atom (10–14) or a residue S atom (15). These short distances could explain the functional significance of these  $\pi$ -hole interactions, as detailed in the main text.

Ligand **10** (za3-500) is found within penicillin-binding protein complex 2y59.<sup>[31]</sup> In chain B, the nitro moiety is nearby the O atom of Thr-413 ( $N\cdots O=3.38$  Å, see Figure S14, Supporting Information). Nitro aromatic **10** is by far the strongest binder of the ligands tested with  $K_i=0.36$   $\mu\text{M}$ , 117 times more potent than the isostructural chloride analogue ( $K_i=42$   $\mu\text{M}$ ).<sup>[31]</sup>

Nitro aromatic **11** (36v-803) is found in protein complex 4qo7, a (*Homo sapiens*) lactate dehydrogenase A.<sup>[32]</sup> As is detailed in Figure S15 (Supporting Information), the carboxylate of Asp-165 is located directly above the nitro group ( $N\cdots O=3.38$  Å in chain A). In a study testing 60 ligands as inhibitors, **11** was found to be 31 times more potent than the -Br analogue ( $LC_{50}=1.7$  vs. 53  $\mu\text{M}$ ).<sup>[32]</sup>

Compound **12** is found as ligand 2nb-302 in (*Saccharomyces cerevisiae*) cytochrome c peroxidase structure 4nvh.<sup>[33]</sup> Within the binding pocket (detailed in Figure S16, Supporting Information) the nitro N atom is merely 3.08 Å away from the O atom of Met-228. This methionine residue is preserved in the nearly identical binding pockets of 4nvi (containing a -Br analogue) and 4nvg (containing an ethyl ester analogue). The affinity ( $K_d$ ) of **12** for the protein (7  $\mu\text{M}$ ) is 2.5 and 10 times higher than the isostructural -Br (18  $\mu\text{M}$ ) and ethyl ester (71  $\mu\text{M}$ ) analogues.<sup>[33]</sup>

The *S* (3k2-501) and *R* (3k2-501) enantiomers of **13** were found in (*Homo sapiens*) nuclear receptor peroxisome proliferator-activated receptor gamma (PPAR $\gamma$ ) structures 4r6s and 4r2u, respectively.<sup>[34]</sup> Interestingly, the *S* enantiomer binds 40 times stronger to PPAR $\gamma$  ( $K_d=2$  nM) than the *R* enantiomer ( $K_d=80$  nM). A major difference in the binding pockets of these protein complexes is the orientation of an O-atom of side-chain amide of Gln-286 relative to the nitro aromatic. As is detailed in Figure S17 (Supporting Information), this O atom is pointing towards the  $\pi$ -hole for the stronger binding *S* isomer with a  $N\cdots O$  distance of 2.81 Å, which is 0.26 Å below the van der Waals benchmark. This O atom is pointing away from the nitro aromatic in the weaker bound *R* isomer.

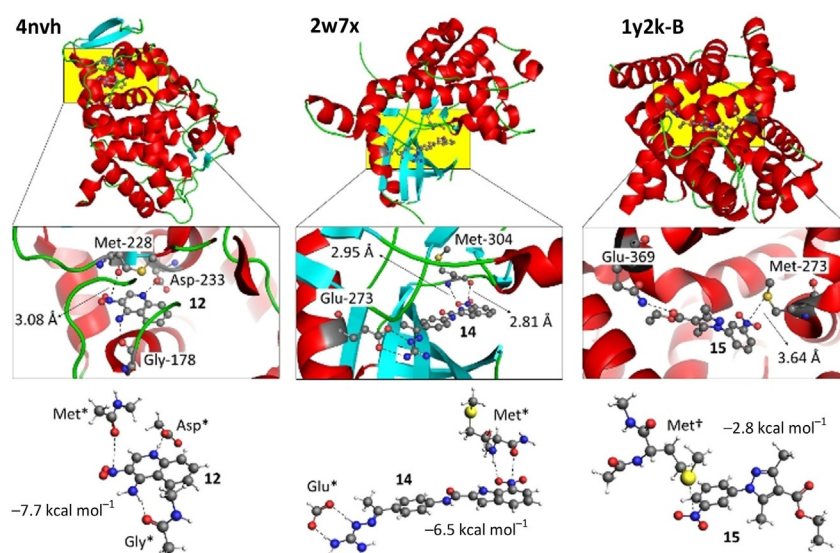
7-Nitro-1*H*-indole derivative **14** (d1a-601) was found in (*Homo sapiens*) checkpoint kinase 2 structure 2w7x (see also Figure 5), in which it displays a very short nitro  $N\cdots O$  (Met-304) distance of 2.81 Å (i.e., 0.26 Å below the van der Waals benchmark).<sup>[35]</sup> As is detailed in Figure S18 (Supporting Information), this complex is part of a series of seven structures containing

very similar ligands, four of which contain the 7-nitro-1*H*-indole fragment.<sup>[35,36]</sup> The residues surrounding the ligands as well as their relative orientation is highly preserved throughout these seven structures. The short  $N\cdots O$  distance is also preserved in the four structures involving a 7-nitro-1*H*-indole ligand (average distance = 2.87 Å) and this interaction actually is one of only two polar contacts with the ligands. Interestingly, **14** displays the shortest  $N\cdots O$  contact distance and is also the best inhibitor tested with  $IC_{50}=0.03$  nM, which is at least 45 times more potent than ligands that lack the nitro group (e.g.,  $IC_{50}=1.36$  nM for nxp-600 in 2yca, see Figure S18, Supporting Information).<sup>[35,36]</sup>

Nitro aromatic ligand **15** (7de-601) co-crystallized with (*Homo sapiens*) phosphodiesterase (PDE) 4D (1y2k) and 4B (1y2j) where it is situated in nearly identical binding pockets (see Figure S22 in the Supporting Information for details).<sup>[37]</sup> The nitro moiety is in close contact with the S atom of Met-273 with  $N\cdots S=3.64$  Å (in chain B of 1y2k, see also Figure 5). Interestingly, **15** was found to be the most potent inhibitor for both PDE-4D ( $IC_{50}=21$  nM) and PDE-4B ( $IC_{50}=33$  nM). The  $IC_{50}$  values for the other 20 inhibitors tested are typically in the (higher) mM range.<sup>[37]</sup>

## Computational studies

To probe the energetic contribution of the likely  $\pi$ -hole interactions involving **6–15**, model binding pockets were evaluated computationally (see sections 3.1 and 3.9 of the Supporting Information). Atomic coordinates of relevant C, N and O atoms were retrieved from the PDB, H atoms were added and their position optimized at the BP86<sup>[38]</sup>-D3<sup>[39]</sup>/def2-TZVP<sup>[40]</sup> level of theory after which an energy calculation was conducted at the much higher SCS-MP2<sup>[41]</sup>/def2-TZVP<sup>[40]</sup> level of theory ( $\Delta E^1$ ). The def2-TZVP basis set was chosen due to the very low basis set superposition error.<sup>[40]</sup> For comparison, interaction energies were also computed of isostructural nitroso arenes (i.e., in silico -NO<sub>2</sub> to -NO mutation;  $\Delta E^2$ ), which are sterically similar but have a negligible  $\pi$ -hole (+3.1 kcal mol<sup>-1</sup>, see Figure S8, Supporting Information). The differences in interacting energies ( $\Delta\Delta E=\Delta E^1-\Delta E^2$ ) were always negative and as large as -4.4 kcal mol<sup>-1</sup> for **13** in 4r6s, which is indicative of the energetic contribution of the nitro moiety  $\pi$ -hole interaction.



**Figure 5.** Examples of high level computational analysis (SCS-MP2/def2-TZVP) to estimate the energetic contribution of the  $\pi$ -hole interactions found in structures 4nhv (left) 2w7x (middle) and 1y2k (right). The top picture shows the protein structure with helices in red, sheets in cyan and loops in green. The yellow area highlights the binding pockets as magnified in the middle. The bottom is a model of the binding pocket with optimized energy.

Three examples were scrutinized further; **12** in structure 4nhv and **14** in structure 2w7x highlighted on the left and in the middle of Figure 5 serve as an example for a  $\pi$ -hole interaction with an amidic carbonyl O-atom. Compound **15** in 1y2k highlighted in the right-hand side of Figure 5 involves a carbon-bound S atom.

Inhibitor **12** is shown in Figure 5 (left), bound within the active site of cytochrome c peroxidase structure 4nhv by hydrogen bonding to Gly-178 and Asp-233 (see also Figure S16, Supporting Information). Additionally, the carbonyl O atom of Met-228 is in close proximity of the ligand nitro moiety with  $\text{nitroN}\cdots\text{O}^{\text{amide}} = 3.08 \text{ \AA}$ . The interaction energy of a model for the nitroarene–amide interaction (bottom left in Figure 5) is  $-7.7 \text{ kcal mol}^{-1}$ . An energy decomposition analysis (see section 1.3. of the Supporting Information) indicates that electrostatics (53%), orbital interactions (22%) and dispersion (25%) all contribute significantly to the binding energy. Mutating the amide C=O to a CH<sub>2</sub> group resulted in an interaction energy of merely  $-2.1 \text{ kcal mol}^{-1}$ , which means that the  $\pi$ -hole interaction can account for about  $-5.6 \text{ kcal mol}^{-1}$ . Interestingly, the bromide analogue of ligand **12** is found within the structure 4nvi. Both binding pockets are nearly identical to one another, as detailed in Figure S16-2 (middle; Supporting Information). The binding energies for model binding pockets are  $-119.0 \text{ kcal mol}^{-1}$  for the -NO<sub>2</sub> ligand and  $-109.3 \text{ kcal mol}^{-1}$  for the -Br ligand (SCS-MP2-def2-SVP, see section 1.4. of the Supporting Information). Mutating the NO<sub>2</sub> ligand in silico to the -Br analogue while retaining the model binding pocket of 4nhv gave an energy penalty of  $+9.0 \text{ kcal mol}^{-1}$ . Conversely, in silico -Br  $\rightarrow$  -NO<sub>2</sub> mutation of the ligand in 4nvi gave an energy gain of  $-5.5 \text{ kcal mol}^{-1}$ . These calculations are consistent with the observations that nitro aromatic **12** binds stronger to the protein ( $K_d = 7 \text{ }\mu\text{M}$ ) than its -Br analogue ( $K_d = 18 \text{ }\mu\text{M}$ ).<sup>[33]</sup>

Inhibitor **14** is anchored within the active site of checkpoint kinase 2 structure 2w7x by a strong salt bridge to Glu-273 (Figure 5, middle). An additional polar interaction involving Met-304 is a combined H-bonding ( $\text{N}\cdots\text{O} = 2.95 \text{ \AA}$ ) and  $\pi$ -hole ( $\text{N}\cdots\text{O} = 2.81 \text{ \AA}$ ) interaction. Curiously, no other H-bonding interaction is established between the inhibitor and the active site, whereas both NHs of the inhibitor interact with water molecules (not shown). The interaction energy of the complex  $[\text{14H}^+\cdots\text{O}_2\text{CH}\cdots\text{Met}^*]$  (bottom center in Figure 5) is  $-6.5 \text{ kcal mol}^{-1}$ , which accounts for both the N–H $\cdots$ O H-bond and the  $\pi$ -hole interaction. An energy decomposition analysis (see section 1.3. of the Supporting Information) revealed that 51% of the attractive forces are electrostatic in nature, with orbital interactions and dispersion accounting for 24 and 25% respectively. Mutating the amide C=O involved in the  $\pi$ -hole interaction, or the amide N–H involved in the hydrogen bond to a H atom (i.e. CH<sub>2</sub> or CH) resulted in interaction energies of  $\Delta E^{\text{H-bond}} = -2.9 \text{ kcal mol}^{-1}$  (i.e., only the H-bonding interaction) and  $\Delta E^{\pi\text{-hole}} = -4.4 \text{ kcal mol}^{-1}$  (i.e., only the  $\pi$ -hole interaction). This implies that the contribution of the  $\pi$ -hole interaction is of similar magnitude as the H-bond, and actually most significant at about  $-4.0 \text{ kcal mol}^{-1}$ . An energy decomposition analysis (see section 1.3. of the Supporting Information) of the two mutants revealed that dispersion contributes somewhat more in the  $\pi$ -hole interaction (38% $\pi$ -hole vs. 24%<sup>H-bond</sup>), whereas the orbital contribution is larger in the H-bonding interaction (13% $\pi$ -hole vs. 30%<sup>H-bond</sup>). The electrostatic component is about the same on both cases (49% $\pi$ -hole vs. 46%<sup>H-bond</sup>).

Inhibitor **15** is found within the active site of phosphodiesterase structure 1y2k (Figure 5, right), in which it is H-bonded to only one residue (Glu-369). The nitro-group is H-bonded to a water molecule (not shown) and is in close proximity to the S atom of Met-273. The interaction energy of the complex  $\text{15}\cdots\text{Met}^{\dagger}$  (bottom right in Figure 5) is  $-2.8 \text{ kcal mol}^{-1}$ . Addition

of a water molecule H-bonded to the nitro group increased the interaction energy to  $-4.0 \text{ kcal mol}^{-1}$ . An energy decomposition analysis (see section 1.3. of the Supporting Information) revealed that about 25% of the attractive forces are electrostatic in nature, with orbital interactions and dispersion each accounting for about 15 and 60%, respectively.

### Approved nitro aromatic drugs within the PDB

Considering that the evidences from structure data, literature studies, and our model computations all point towards the relevance of  $\pi$ -hole interactions involving nitro aromatic ligands in protein structures, we assessed which approved drugs contain a nitro aromatic and looked for their occurrence in the PDB (see section 4 of the Supporting Information for details). In total, 54 such drugs were identified (see Table S5, Supporting Information) and nine of these drugs (or very similar molecules) were present as a ligand: chloramphenicol (and related ligands),<sup>[21]</sup> nimodipine,<sup>[42]</sup> hydroxy flutamide (the active metabolite of flutamide),<sup>[23]</sup> nifursol-like ligands,<sup>[43]</sup> a venetoclax-like ligand,<sup>[24]</sup> tolcapone,<sup>[44]</sup> nimesulide,<sup>[45]</sup> nitroxoline,<sup>[46]</sup> and oxaminiquine.<sup>[47]</sup> Manual inspection of the 5 Å envelope surrounding the  $\text{NO}_2$  moiety revealed that in 32 of the 57 structures (56%) there is at least one electron-rich moiety near the nitro group (i.e., atoms with lone-pair electrons and/or aromatic rings). These 32 structures included all of the nine abovementioned drugs—except for a finurisol-like ligand—and are summarized in Figures S37–S41 (Supporting Information).

### Discussion

The numerical analysis of the data compiled in the PDB (Figure 2) shows clustering of data near the  $\pi$ -hole region. These data and the examples highlighted in Figures 3–5 show that  $\pi$ -hole interactions with nitro aromatic ligands are a common structural phenomenon within the PDB and might thus be functionally relevant. The lack of apparent van der Waals overlap involving carbon bound S vis-à-vis water and carbonyl O atoms is in line with the model computations involving dimethyl thioether (see section 1 of the Supporting Information). Moreover, it has been reported that close proximity of the S atom of methionine and the  $\pi$ -system of histidine is conserved within at least 80 protein structures, without displaying van der Waals overlap.<sup>[48]</sup> This interaction was furthermore shown to stabilize transition state 2 of phosphite dehydrogenase with  $7.9 \text{ kcal mol}^{-1}$ .<sup>[48]</sup> This suggests that for S atoms, overlap of van der Waals shells is not a strict prerequisite for the relevance of an interaction with that S atom. In light of this, it is even more remarkable that so many data were found for which water and carbonyl O atoms are involved in apparent van der Waals overlap with the nitro group N atom and presents experimental evidence for the significance of such  $\pi$ -hole interactions within proteins.

Several examples were encountered in the literature for which a nitro aromatic ligand displayed a better inhibition or binding affinity compared to a very similar non-nitro analogue. In all cases highlighted, the protein–ligand interaction includes

a  $\pi$ -hole interaction with the nitro aromatic ligand, as is evident from the crystal structures. This clearly suggests a functional relevance of these  $\pi$ -hole interactions.

The computational data involving simple models of nitro arene ligands (3 and 4, Table S1, Supporting Information) predict that  $\pi$ -hole interactions are about  $-5 \text{ kcal mol}^{-1}$ . These interaction energies are likely a conservative estimate, as nitro aromatics bound within a protein can bear a deeper  $\pi$ -hole due to (charge assisted) H-bonding, like in the bis-amide model complex 5 (Figure 1). The interaction energies estimated with inhibitors 12, 14 and 15 highlighted in Figure 5 indeed show that  $\pi$ -hole interactions can contribute about  $5 \text{ kcal mol}^{-1}$  to the overall binding energy of a protein–ligand complex. This is twice as strong as the H-bonding interaction between the  $\alpha$ -H of side chain amino acids with water ( $1.9$ – $2.5 \text{ kcal mol}^{-1}$ )<sup>[49]</sup> and of the same magnitude as a water dimer.<sup>[50]</sup> Thus, the strength of these  $\pi$ -hole interactions is in general in the order of a moderately strong hydrogen bond.<sup>[51]</sup> It is worth pointing out that interactions of that magnitude are well known to play an important role in biological systems. For example, carbonyl–carbonyl  $\pi$ -hole interactions can be a structural determinant in protein folding.<sup>[13b]</sup> C–H $\cdots\pi$  interactions are a well-preserved structural feature in protein carbohydrate complexes<sup>[52]</sup> and have even been successfully exploited as important design elements for making artificial carbohydrate receptors.<sup>[53]</sup>

The search for approved nitro aromatic drugs within the PDB revealed nine drugs within 57 PDB structures. In 32 (56%) of these structures the nitro moiety is in closest proximity to an aromatic and/or a lone pair bearing atom of a residue. This, again, suggests that  $\pi$ -hole interactions with nitro aromatics is a common structural feature. That these structural features were encountered with eight out of the nine approved drugs found as ligand in the PDB again hints to the functional significance of these  $\pi$ -hole interactions.

### Conclusions

It is thus concluded that  $\pi$ -hole interactions of nitro aromatic ligands within proteins are a force to be reckoned with. Explicit examples are provided for  $\pi$ -hole interactions in biological systems. This discovery likens to the finding that halogen bonds are biologically significant. We thus expect that taking  $\pi$ -hole interactions into account will aid future interpretations of the structure/function relationship involving nitro aromatics, including the more than 50 currently approved drugs that contain the nitro aromatic moiety. Finally, our finding may serve as an aid to discover/design new ligands as potent inhibitors of a target protein.

### Acknowledgements

A.F. and A.B. thank the MINECO of Spain (projects CTQ2014-57393-C2-1-P and CTQ2017-85821-R, FEDER funds) for funding. We also thank the “Centre de Tecnologies de la Informació” (CTI) at the UIB for computational facilities. T.J.M. partially con-

ducted the work with funds from the research program "VIDI" with project number 723.015.006, which is financed by the Netherlands Organization for Scientific Research (NWO).

## Conflict of interest

The authors declare no conflict of interest.

**Keywords:** density functional calculations · ligand–protein interactions · nitro compounds · PDB analysis · pi–hole interactions

- [1] J. M. Lehn, *Supramolecular Chemistry: Concepts and Perspectives*, 1st ed., VCH, Weinheim, **1995**.
- [2] a) T. Clark, *WIREs Comput. Mol. Sci.* **2013**, *3*, 13–20; b) P. Politzer, J. S. Murray, T. Clark, *Phys. Chem. Chem. Phys.* **2013**, *15*, 11178–11189; c) A. Bauzá, T. J. Mooibroek, A. Frontera, *ChemPhysChem* **2015**, *16*, 2496–2517; d) R. H. Crabtree, *Chem. Soc. Rev.* **2017**, *46*, 1720–1729.
- [3] S. J. Grabowski, *Hydrogen Bonding: New Insights*, Springer, Heidelberg, **2006**.
- [4] a) A. Bauzá, T. J. Mooibroek, A. Frontera, *Angew. Chem. Int. Ed.* **2013**, *52*, 12317–12321; *Angew. Chem.* **2013**, *125*, 12543–12547; b) A. Bauzá, T. J. Mooibroek, A. Frontera, *Chem. Rec.* **2016**, *16*, 473–487; c) A. Bauzá, A. Frontera, T. J. Mooibroek, *Phys. Chem. Chem. Phys.* **2016**, *18*, 1693–1698; d) A. Bauzá, T. J. Mooibroek, A. Frontera, *Phys. Chem. Chem. Phys.* **2014**, *16*, 19192–19197; e) A. Bauzá, T. J. Mooibroek, A. Frontera, *Chem. Commun.* **2014**, *50*, 12626–12629.
- [5] a) G. Cavallo, P. Metrangolo, R. Milani, T. Pilati, A. Priimagi, G. Resnati, G. Terraneo, *Chem. Rev.* **2016**, *116*, 2478–2601.
- [6] P. Auffinger, F. A. Hays, E. Westhof, P. Shing Ho, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16789–16794.
- [7] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235–242.
- [8] a) A. R. Voth, P. Khuu, K. Oishi, P. S. Ho, *Nat. Chem.* **2009**, *1*, 74–79; b) L. A. Hardegger, B. Kuhn, B. Spinnler, L. Anselm, R. Ecabert, M. Stihle, B. Gsell, R. Thoma, J. Diez, J. Benz, J.-M. Plancher, G. Hartmann, D. W. Banner, W. Haap, F. Diederich, *Angew. Chem. Int. Ed.* **2011**, *50*, 314–318; *Angew. Chem.* **2011**, *123*, 329–334; c) Q. Zhang, Z. J. Xu, J. Y. Shi, W. Zhu, *J. Chem. Inf. Model.* **2017**, *57*, 1529–1534.
- [9] a) P. S. Ho, *Future Med. Chem.* **2017**, *9*, 637–640; b) L. Mendez, G. Henriquez, S. Sirimulla, M. Narayan, *Molecules* **2017**, *22*, 1397.
- [10] A. R. Voth, F. A. Hays, P. S. Ho, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 6188–6193.
- [11] E. Danelius, H. Andersson, P. Jarvoll, K. Lood, J. Gräfenstein, M. Erdélyi, *Biochemistry* **2017**, *56*, 3265–3272.
- [12] J. S. Murray, P. Lane, T. Clark, K. E. Riley, P. Politzer, *J. Mol. Model.* **2012**, *18*, 541–548.
- [13] a) H. B. Bürgi, *Angew. Chem. Int. Ed. Engl.* **1975**, *14*, 460–473; *Angew. Chem.* **1975**, *87*, 461–475; b) G. J. Bartlett, A. Choudhary, R. T. Raines, D. N. Woolfson, *Nat. Chem. Biol.* **2010**, *6*, 615–620; c) M. T. Doppert, H. van Overeem, T. J. Mooibroek, *Chem. Commun.* **2018**, *54*, 12049–12052; d) A. Ruigrok van der Werve, Y. R. van Dijk, T. J. Mooibroek, *Chem. Commun.* **2018**, *54*, 10742–10745; e) J. Lewiński, W. Bury, W. Justyniak, *Eur. J. Inorg. Chem.* **2005**, 4490–4492; f) J. Echeverría, *Chem. Commun.* **2018**, *54*, 3061–3064.
- [14] a) P. Gamez, T. J. Mooibroek, S. J. Teat, J. Reedijk, *Acc. Chem. Res.* **2007**, *40*, 435–444; b) P. Gamez, *Inorg. Chem. Front.* **2014**, *1*, 35–43.
- [15] a) B. L. Schottel, H. T. Chifotides, K. R. Dunbar, *Chem. Soc. Rev.* **2008**, *37*, 68–83; b) T. J. Mooibroek, P. Gamez, *CrystEngComm* **2012**, *14*, 1027–1030; c) A. Frontera, P. Gamez, M. Mascal, T. J. Mooibroek, J. Reedijk, *Angew. Chem. Int. Ed.* **2011**, *50*, 9564–9583; *Angew. Chem.* **2011**, *123*, 9736–9756; d) T. J. Mooibroek, P. Gamez, J. Reedijk, *CrystEngComm* **2008**, *10*, 1501–1515.
- [16] a) A. Bauzá, T. J. Mooibroek, A. Frontera, *Chem. Commun.* **2015**, *51*, 1491–1493; b) A. Báuza, A. Frontera, T. J. Mooibroek, *Cryst. Growth Des.* **2016**, *16*, 5520–5524; c) A. Bauzá, A. V. Sharko, G. A. Senchyk, E. B. Rusanov, A. Frontera, K. V. Domasevitch, *CrystEngComm* **2017**, *19*, 1933–1937.
- [17] A. Bauzá, A. Frontera, T. J. Mooibroek, *Nat. Commun.* **2017**, *8*, 14522.
- [18] T. J. Mooibroek, *CrystEngComm* **2017**, *19*, 4485–4488.
- [19] W. Li, L. Spada, N. Tasinato, S. Rampino, L. Evangelisti, A. Gualandi, P. G. Cozzi, S. Melandri, V. Barone, C. Puzzarini, *Angew. Chem. Int. Ed.* **2018**, *57*, 13853–13857; *Angew. Chem.* **2018**, *130*, 14049–14053.
- [20] D. Olender, J. Zwawiak, L. Zaprutko, *Pharmaceuticals* **2018**, *11*, 54.
- [21] H. C. Neu, *Science* **1992**, *257*, 1064–1073.
- [22] C. P. O'Brien, *J. Clin. Psychiatry* **2005**, *66*, 28–33.
- [23] Prostate Cancer Trialists Collaborative Group, *Lancet* **2000**, *355*, P1491–P1498.
- [24] A. W. Roberts, M. S. Davids, J. M. Pagel, B. S. Kahl, S. D. Puvvada, J. F. Gerecitano, T. J. Kipps, M. A. Anderson, J. R. Brown, L. Gressick, S. Wong, M. Dunbar, M. Zhu, M. B. Desai, E. Cerri, S. Heitner Enschede, R. A. Humerickhouse, W. G. Wierda, J. F. Seymour, *N. Engl. J. Med.* **2016**, *374*, 311–322.
- [25] T. Krause, M. U. Gerbershagen, M. Fiege, R. Weisshorn, F. Wappler, *Anaesthesia* **2004**, *59*, 364–373.
- [26] P. A. Poole-Wilson, J. Lubsen, B. A. Kirwan, F. J. van Dalen, G. Wagener, N. Danchin, H. Just, K. A. Fox, S. J. Pocock, T. C. Clayton, M. Motro, J. D. Parker, M. G. Bourassa, A. M. Dart, P. Hildebrandt, A. Hjalmarson, J. A. Kragten, G. P. Molhoek, J. E. Otterstad, R. Seabra-Gomes, J. Soler-Soler, S. Weber, *Lancet* **2004**, *364*, 849–857.
- [27] E. Schonbrunn, A. Becker, S. Betzi, et al., unpublished in journal format but structures were released in the PDB in **2012** (3r28, 3qx4, 3r1q, 3qzi, 3qwk, 3qwj, 3r7v, 3r7e, 3r7i, 3r7u, 3qx2, 3r7y, 3r83, 3rm6, 3qzh, 3r6x, 3r71, 3rpo, 3rai and 3r73).
- [28] a) E. Fan, E. A. Merritt, Z. Zhang, et al., *Acta Crystallogr. D Biol. Crystallogr.* **2001**, *57*, 201–212; b) E. A. Merritt, S. Sarfaty, I. K. Feil, W. G. Hol, *Structure* **1997**, *5*, 1485–1499; c) E. A. Merritt, Z. Zhang, J. C. Pickens, M. Ahn, W. G. J. Hol, E. Fan, *J. Am. Chem. Soc.* **2002**, *124*, 8818–8824; d) J. C. Pickens, D. D. Mitchell, J. Liu, X. Tan, Z. Zhang, C. L. M. J. Verlinde, W. G. J. Hol, E. Fan, *Chem. Biol.* **2004**, *11*, 1205–1215; e) J. C. Pickens, E. A. Merritt, M. Ahn, C. L. Verlinde, W. G. Hol, E. Fan, *Chem. Biol.* **2002**, *9*, 215–224.
- [29] N. I. Howard, M. V. Dias, F. Peyrot, L. Chen, M. F. Schmidt, T. L. Blundell, C. Abell, *ChemMedChem* **2015**, *10*, 116–133.
- [30] A. J. Scheidig, S. M. Franken, J. E. Corrie, G. P. Reid, A. Wittinghofer, E. F. Pai, R. S. Goody, *J. Mol. Biol.* **1995**, *253*, 132–150.
- [31] A. Zervosen, R. Herman, F. Kerff, A. Herman, A. Bouillez, F. Prati, R. F. Pratt, J.-M. Frère, B. Joris, A. Luxen, P. Charlier, E. Sauvage, *J. Am. Chem. Soc.* **2011**, *133*, 10839–10848.
- [32] P. S. Dragovich, B. P. Fauber, J. Boggs, J. Chen, L. B. Corson, C. Z. Ding, C. Eigenbrot, H. Ge, A. M. Giannetti, T. Hunsaker, S. Labadie, C. Li, Y. Liu, Y. Liu, S. Ma, S. Malek, D. Peterson, K. E. Pitts, H. E. Purkey, K. Robarge, L. Salphati, S. Sideris, M. Ultsch, E. VanderPorten, J. Wang, B. Wei, Q. Xu, I. Yen, Q. Yue, H. Zhang, X. Zhang, A. Zhou, *Bioorg. Med. Chem. Lett.* **2014**, *24*, 3764–3771.
- [33] M. Fischer, R. G. Coleman, J. S. Fraser, B. K. Shoichet, *Nat. Chem.* **2014**, *6*, 575–583.
- [34] D. P. Marciano, D. S. Kuruvilla, S. V. Boregowda, A. Asteian, T. S. Hughes, R. Garcia-Ordóñez, C. A. Corzo, T. M. Khan, S. J. Novick, H. Park, D. J. Kojetkin, D. G. Phinney, J. B. Bruning, T. M. Kamenecka, P. R. Griffin, *Nat. Commun.* **2015**, *6*, 7443.
- [35] A. G. Jobson, G. T. Lountos, P. L. Lorenzi, J. Llamas, J. Connelly, D. Cerna, J. E. Tropea, A. Onda, G. Zoppoli, S. Kondapaka, G. Zhang, N. J. Caplen, J. H. Cardellina 2nd, S. S. Yoo, A. Monks, C. Self, D. S. Waugh, R. H. Shoemaker, Y. Pommier, *J. Pharmacol. Exp. Ther.* **2009**, *331*, 816–826.
- [36] a) G. T. Lountos, A. G. Jobson, J. E. Tropea, C. R. Self, G. Zhang, Y. Pommier, R. H. Shoemaker, D. S. Waugh, *FEBS Lett.* **2011**, *585*, 3245–3249; b) G. T. Lountos, A. G. Jobson, J. E. Tropea, C. R. Self, G. Zhang, Y. Pommier, R. H. Shoemaker, D. S. Waugh, *J. Struct. Biol.* **2011**, *176*, 292–301.
- [37] G. L. Card, L. Blasdel, B. P. England, C. Zhang, Y. Suzuki, S. Gillette, D. Fong, P. N. Ibrahim, D. R. Artis, G. Bollag, M. V. Milburn, S. H. Kim, J. Schlessinger, K. Y. Zhang, *Nat. Biotechnol.* **2005**, *23*, 201–207.
- [38] a) C. T. Lee, W. T. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789; b) A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.
- [39] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.

- [40] a) F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305; b) F. Weigend, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
- [41] S. Grimme, *J. Chem. Phys.* **2003**, *118*, 9095–9102.
- [42] a) N. G. Wahlgren, D. G. Macmahon, J. Dekeyser, B. Indredavik, T. Ryman, *Cerebrovasc. Dis.* **1994**, *4*, 204–210; b) F. G. Barker, C. S. Ogilvy, *J. Neurosurg.* **1996**, *84*, 405–414.
- [43] L. R. McDougald, *Avian Dis.* **2005**, *49*, 462–476.
- [44] M. C. Kurth, C. H. Adler, M. StHilaire, C. Singer, C. Waters, P. LeWitt, D. A. Chernik, E. E. Dorflinger, K. Yoo, *Neurology* **1997**, *48*, 81–87.
- [45] R. Davis, R. N. Brogden, *Drugs* **1994**, *48*, 431–454.
- [46] a) J. S. Shim, Y. Matsui, S. Bhat, B. A. Nacev, J. Xu, H. E. Bhang, S. Dhara, K. C. Han, C. R. Chong, M. G. Pomper, A. So, J. O. Liu, *JNCI-J. Natl. Cancer Inst.* **2010**, *102*, 1855–1873; b) H. C. Jiang, J. E. Taggart, X. X. Zhang, D. M. Benbrook, S. E. Lind, W. Q. Ding, *Cancer Lett.* **2011**, *312*, 11–17.
- [47] J. Utzinger, J. Keiser, *Expert Opin. Pharmacother.* **2004**, *5*, 263–285.
- [48] K. E. Ranaghan, J. E. Hung, G. J. Bartlett, T. J. Mooibroek, J. N. Harvey, D. N. Woolfson, W. A. van der Donk, A. J. Mulholland, *Chem. Sci.* **2014**, *5*, 2191–2199.
- [49] S. Scheiner, T. Kar, Y. L. Gu, *J. Biol. Chem.* **2001**, *276*, 9832–9837.
- [50] W. Klopper, J. van Duijneveldt-van de Rijdt, F. B. van Duijneveldt, *Phys. Chem. Chem. Phys.* **2000**, *2*, 2227–2234.
- [51] a) G. Nemethy, M. S. Pottle, H. A. Scheraga, *J. Phys. Chem.* **1983**, *87*, 1883–1887; b) F. Biedermann, H.-J. Schneider, *Chem. Rev.* **2016**, *116*, 5216–5300.
- [52] a) N. K. Vyas, M. N. Vyas, F. A. Quioco, *Science* **1988**, *242*, 1290–1295; b) K. L. Hudson, G. J. Bartlett, R. C. Diehl, J. Agirre, T. Gallagher, L. L. Kiesling, D. N. Woolfson, *J. Am. Chem. Soc.* **2015**, *137*, 15152–15160.
- [53] a) T. J. Mooibroek, M. P. Crump, A. P. Davis, *Org. Biomol. Chem.* **2016**, *14*, 1930–1933; b) P. Ríos, T. J. Mooibroek, T. S. Carter, C. Williams, M. R. Wilson, M. P. Crump, A. P. Davis, *Chem. Sci.* **2017**, *8*, 4056–4061; c) T. S. Carter, T. J. Mooibroek, P. F. N. Stewart, M. P. Crump, M. C. Galan, A. P. Davis, *Angew. Chem. Int. Ed.* **2016**, *55*, 9311–9315; *Angew. Chem.* **2016**, *128*, 9457–9461.

---

Manuscript received: July 25, 2019

Accepted manuscript online: August 27, 2019

Version of record online: September 17, 2019