

## Linking ATP and allosteric sites to achieve superadditive binding with bivalent EGFR kinase inhibitors

Florian Wittlinger<sup>a</sup>, Blessing C. Ogboo<sup>b</sup>, Ekaterina Shevchenko<sup>a,c,d</sup>, Tahereh Damghani<sup>b</sup>, Calvin D. Pham<sup>b</sup>, Ilse K. Schaeffner<sup>e,f</sup>, Brandon T. Oligny<sup>b</sup>, Surbhi P. Chitnis<sup>b</sup>, Tyler S. Beyett<sup>e,f,g</sup>, Alexander Rasch<sup>a</sup>, Brian Buckley<sup>h</sup>, Daniel A. Urul<sup>i</sup>, Tatiana Shaurova<sup>j</sup>, Earl W. May<sup>i</sup>, Erik M. Schaefer<sup>i</sup>, Michael J. Eck<sup>e,f</sup>, Pamela A. Hershberger<sup>j</sup>, Antti Poso<sup>a,c,d,k</sup>, Stefan A. Laufer<sup>a,c,d\*</sup>, David E. Heppner<sup>b,j\*</sup>

a. Department of Pharmaceutical and Medicinal Chemistry, Institute of Pharmaceutical Sciences, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

b. Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, NY, 14260, USA

c. Cluster of Excellence iFIT (EXC 2180) “Image-Guided and Functionally Instructed Tumor Therapies” Eberhard Karls Universität Tübingen, 72076 Tübingen, Germany.

d. Tübingen Center for Academic Drug Discovery & Development (TüCAD2), 72076 Tübingen, Germany

e. Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, 02215 USA

f. Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, 02115 USA

g. Present Address: Department of Pharmacology and Chemical Biology, Emory University School of Medicine, 5119 Rollins Research Center, 1510 Clifton Rd, Atlanta, GA 30322, USA

h. Department of Cell Stress Biology, Roswell Park Comprehensive Cancer Center, Buffalo, NY, 14203, USA

i. AssayQuant Technologies, Inc., Marlboro, MA, 01752, USA

j. Department of Pharmacology and Therapeutics, Roswell Park Comprehensive Cancer Center, Buffalo, NY, 14203, USA

k. School of Pharmacy, University of Eastern Finland, 70210 Kuopio, Finland

To whom correspondence may be addressed:

David E. Heppner [davidhep@buffalo.edu](mailto:davidhep@buffalo.edu)

Stefan A. Laufer [stefan.laufer@uni-tuebingen.de](mailto:stefan.laufer@uni-tuebingen.de)

These authors contributed equally:

Florian Wittlinger, Blessing C. Ogboo

## Abstract

Bivalent molecules consisting of groups connected through bridging linkers often exhibit strong target binding and unique biological effects. However, developing bivalent inhibitors with the desired activity is challenging due to the dual motif architecture of these molecules and to the variability that can be introduced through differing linker structures and geometries. We report a set of alternatively linked bivalent EGFR inhibitors that simultaneously occupy the ATP substrate and allosteric pockets. Crystal structures show that initial and redesigned linkers between a trisubstituted imidazole ATP-site inhibitor and dibenzodiazepinone allosteric-site inhibitor proved successful in spanning these sites. However, the re-engineered linker yielded a compound with ~60 pM potency against the drug-resistant EGFR L858R/T790M and L858R/T790M/C797S, which was superadditive as compared with the parent molecules. The enhanced potency is attributed to factors stemming from the linker connection to the allosteric-site group and informs strategies to engineer linkers in bivalent agent design.

## Introduction

Molecules that simultaneously bind to distinct sites within biological targets are increasingly sought after in drug development. This binding strategy is often accomplished through bivalent (or heterobifunctional) compounds, which comprise dual functional motifs connected by a covalent linker.<sup>1,2</sup> Diverse and novel pharmacological strategies have emerged based on the design of such bivalent compounds that induce protein-protein neo-associations,<sup>2-9</sup> produce protein homodimers,<sup>10-14</sup> or associate within distinct sites of the same target.<sup>15-17</sup> A pivotal step in the development of these biological agents involves the optimization of the linker composition and structure, which relies on brute force exploration of synthetically-accessible structural and functional motifs motivating studies that enable more efficient design strategies.<sup>18-22</sup>

Optimization of bivalent small molecules is also a principle focus in drug development most commonly with respect to fragment-based drug discovery (FBDD) where low molecular-weight building blocks can be connected to generate high-affinity lead molecules spanning diverse binding sites.<sup>23-26</sup> The exploration of linked fragments is a highly attractive strategy in drug discovery due to the potential for superadditivity, such that the linked complex binds stronger than the sum of the parent fragments on their own.<sup>27-30</sup> Despite the simple premise, diverse efforts spanning several decades have shown that superadditivity is scarce and the majority of cases fall short of achieving the expected improvement in target affinity.<sup>20,21,28</sup> Additionally, case studies have offered suggestions for ideal binding properties of optimally linked compound,<sup>20,21,28,30</sup> but little is known regarding a general structural-based strategy for swiftly maximizing fragment linkers.

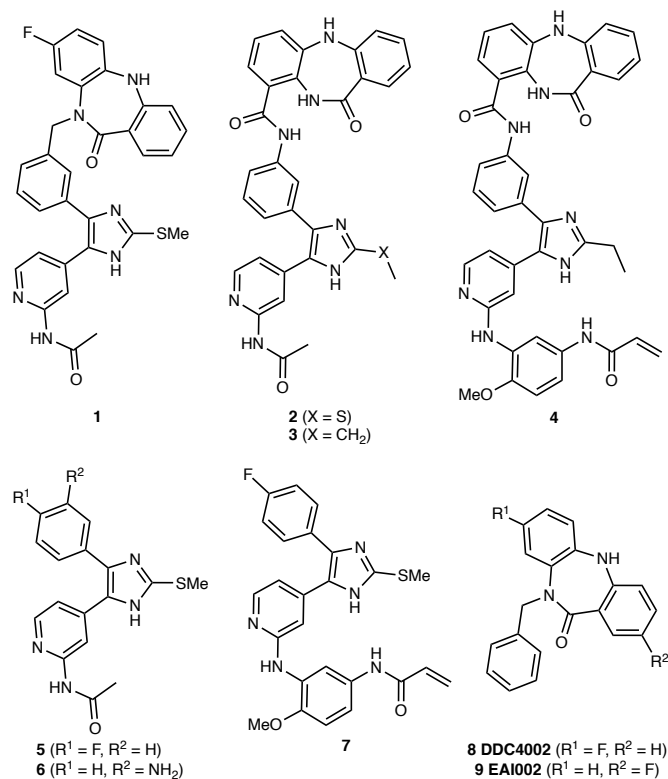
The design of effective bivalent inhibitors has been inspired, in part, by the search for more effective tyrosine kinase inhibitors (TKIs).<sup>31-33</sup> Kinase inhibitors most commonly bind a canonical orthosteric (ATP) binding site and more recently a set of distinct allosteric inhibitors have been reported.<sup>34-36</sup> The kinase domain of the epidermal growth factor receptor (EGFR) is an established drug target in non-small cell lung cancer (NSCLC) where oncogenic mutations often predict clinical responsiveness to treatment with certain TKIs.<sup>37</sup> Indeed, clinically-effective TKIs are often selective for EGFR activating mutations L858R (LR) and exon19del,<sup>38</sup> as well as drug resistant

T790M (TM) gatekeeper and C797S (CS) mutants. Promising pre-clinical results have been seen where combinations of ATP and allosteric inhibitors show synergistic tumor regression *in vivo* in addition to delayed acquired drug resistance.<sup>39,40</sup> Importantly, the EGFR allosteric inhibitor binding site is adjacent to the ATP pocket and cobinding of structurally-compatible inhibitors within both sites<sup>41,42</sup> has been shown to enable structural changes consistent with biochemical experiments where these two inhibitor types exhibit cooperative binding.<sup>42</sup> The unique effects enabled by combinations of ATP and allosteric inhibitors, as well as the structural proximity of their binding sites, have led to the recent development of ATP-allosteric bivalent molecules that simultaneously occupy these sites.<sup>43,44</sup>

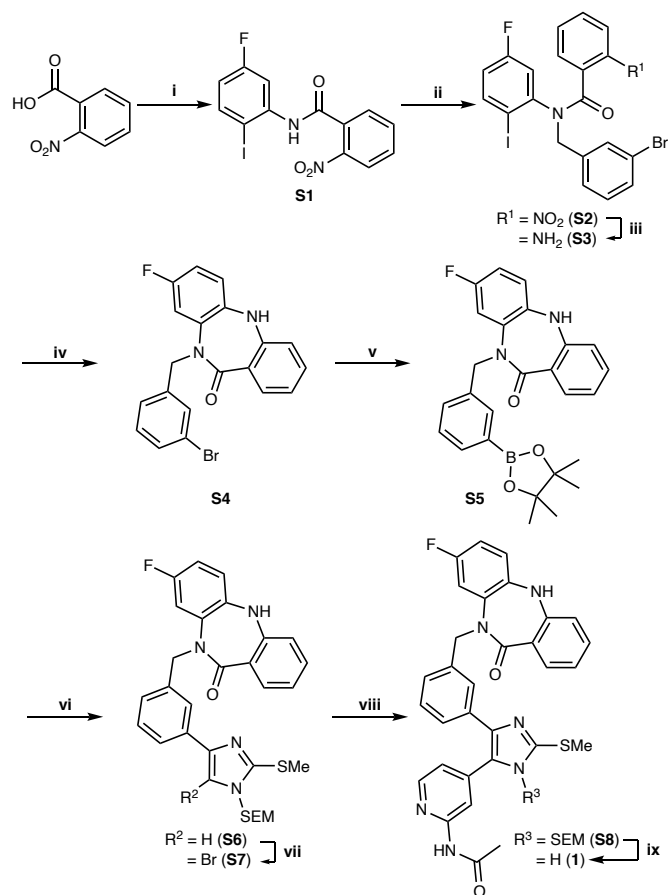
In this study, we synthesized a novel series of bivalent EGFR kinase inhibitors that simultaneously bind the ATP and allosteric sites and differ with respect to the structure of the site-bridging linker. Strikingly, we find that distinctly linked compounds exhibit considerable differences in potency where one linker exhibits superadditivity, and another is virtually inactive. Structural characterization indicates that the linker structure induces conformational differences and intermolecular interactions that provide a unique side-by-side comparison of functionally divergent linking strategies. Cocrystal structures and molecular dynamics simulations of these bivalent inhibitors enable the dissection of the specific properties of the linker that afford strong binding, which informs novel design strategies.

## Results

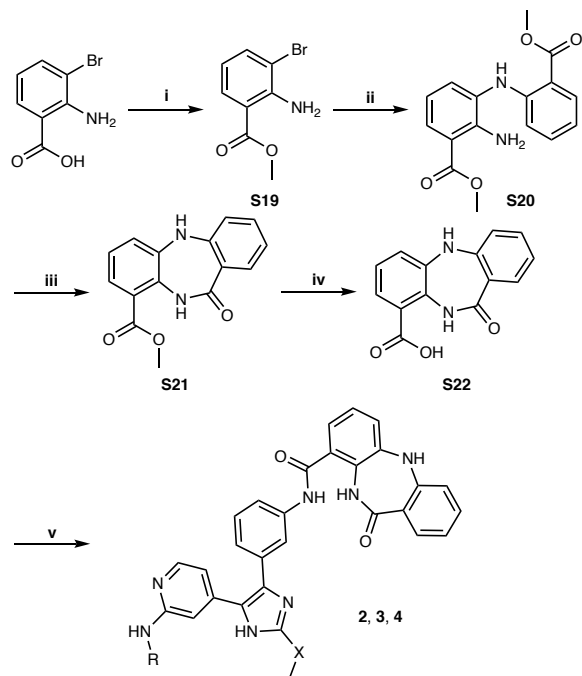
Due to the structural proximity of the allosteric and ATP (orthosteric) sites within the EGFR kinase, we sought to explore alternatively linked bivalent compounds that span these pockets. Starting motifs were selected and derived from established ATP-competitive inhibitors based on trisubstituted imidazole molecules (**5-7**),<sup>45-49</sup> and the mutant-selective allosteric 5,10-dihydro-11*H*-dibenzo[*b,e*][1,4]diazepin-11-one inhibitors **8** and **9** (herein denoted as “benzo” for simplicity).<sup>50</sup> We synthesized a set of bivalent ATP-allosteric inhibitors bridged by an *N*-linked methylene (**1**) and *C*-linked amide (**2-4**) (Scheme 1). To combine fragments for the *N*-linked derivative **1** we utilized a cross-coupling focused reaction route. The motif of the allosteric site was thereby assembled by slight adjustments of previously described conditions for derivatives of the allosteric inhibitor **8**.<sup>50</sup> The subsequent Miyaura borylation afforded the corresponding boronic acid pinacol ester, which was Suzuki coupled with the imidazole core of the orthosteric scaffold. Bromination and Suzuki coupling with the hinge-binding motif yielded **1** after acidic deprotection (Scheme 2). For the synthesis of the allosteric motif of the *C*-linked series we applied a Buchwald-Hartwig amination for coupling of methyl anthranilate with 3-bromo methyl anthranilate. The product was refluxed in acetic acid to obtain the methyl-dibenzodiazepine-9-carboxylate by means of an intramolecular aminolysis. Saponification of the remaining ester and amide coupling of thereof resulting carboxylic acid with amines of corresponding orthosteric motifs yielded *C*-linked derivatives **2**, **3** and **4** after deprotection (Scheme 3, Supplementary Scheme S2).



Scheme 1. Bivalent inhibitors of EGFR developed in this study and relevant ATP and allosteric analogues for functional comparisons. Chemical structures of bivalent ATP-allosteric inhibitors consisting of *N*-linked reversible (**1**) as well as *C*-linked reversible (**2**, **3**) and covalent (**4**) scaffolds. Parent ATP site imidazole reversible (**5**, **6**) and covalent (**7**) inhibitors as well as dibenzodiazepinone allosteric inhibitors (**8**, **9**).



**Scheme 2. Synthesis of the *N*-linked bivalent inhibitor (1) inspired by the parent allosteric inhibitors 8 and 9.** - Reagents and conditions are as follows: i) a) (COCl)<sub>2</sub>, DMF cat., DCM, rt; b) 5-Fluoro-2-iodoaniline, Et<sub>3</sub>N, DCM, rt, 38% over two steps; ii) 3-Bromobenzyl bromide, NaH (60% dispersion in mineral oil), THF, rt, 91%; iii) Fe, NH<sub>4</sub>Cl, THF/MeOH/H<sub>2</sub>O, 50 °C, quant.; iv) CuI, K<sub>2</sub>CO<sub>3</sub>, DMSO, 135 °C, 76%; v) Bis(pinacolato)diboron, KOAc, Pd(dppf)Cl<sub>2</sub>, 1,4-dioxane, 90 °C, quant.; vi) 4-Bromo-2-(methylthio)-1-((2-(trimethylsilyl)ethoxy)methyl)-1*H*-imidazole, K<sub>3</sub>PO<sub>4</sub> trihydrate, P(*t*-Bu)<sub>3</sub> Pd G3, 1,4-dioxane/H<sub>2</sub>O, 50 °C, 70%; vii) NBS, ACN, -30 °C, 70%; viii) *N*-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)acetamide, K<sub>3</sub>PO<sub>4</sub> trihydrate, P(*t*-Bu)<sub>3</sub> Pd G3, 1,4-dioxane/H<sub>2</sub>O, 50 °C, 64%; ix) 33% TFA in DCM, rt, 62%. With adaptations from 44,50



**Scheme 3. Synthesis of the C-linked amide bridged bivalent inhibitors (2-4).** - Reagents and conditions are as follows: i) MeOH, H<sub>2</sub>SO<sub>4</sub>, rf, 56%; ii) Methyl anthranilate, Cs<sub>2</sub>CO<sub>3</sub>, BrettPhos Pd G3, 1,4-dioxane, rf, 73%; iii) AcOH, rf, 63%; iv) 2 N NaOH (aq), MeOH, rt, 95%; v) a) appropriate aniline (see supporting information), HATU, TEA, DMF, rt; b) 33% TFA in DCM or MSA in DCM, rt, 30-65% over two steps. For the substitutions at position X and R see Scheme 1.

We first sought to understand the degree to which these alternatively linked motifs influence the ability to inhibit recombinant EGFR kinase activity. We carried out biochemical IC<sub>50</sub> value determination using HTRF-based activity assays with purified EGFR kinase domains (Table 1, Figure 1). Strikingly, the *N*-linked **1** was observed to be limitedly potent against WT and mutant EGFR with IC<sub>50</sub> values  $\geq 1 \mu\text{M}$  while the *C*-linked inhibitors **2-3** show substantially lower IC<sub>50</sub> values of 1.2-1.5 nM for LR and 51-64 pM for LRTM and LRTMCS (Table 1, Figure 1). The C797-targeting irreversible *C*-linked analogue **4** was slightly less potent as a reversible inhibitor of LRTMCS, and additional time-dependent activity measurements showed that this molecule was most effective against LR (Table 2). To put these biochemical IC<sub>50</sub> values into proper context, we next compared them to structurally related ATP- and allosteric-site analogues **5-9**. The ATP-site imidazole motifs **5,6** and the original allosteric inhibitors **8,9** inhibit LRTM and LRTMCS at IC<sub>50</sub> values  $\geq 6-10 \mu\text{M}$  and  $\sim 39-59 \text{ nM}$ , respectively, indicating that the *C*-linked bivalent molecules are 10<sup>3</sup>-to-10<sup>6</sup>-fold more potent over the parent motifs. The matched covalent analogue **4** inhibits LRTMCS reversibly with an IC<sub>50</sub> value 100-fold better than the orthosteric-only **7**, showcasing the additional reversible binding gained by interactions in the allosteric pocket in this covalent scaffold. Seeing as how the allosteric motif and linker in **2-4** is different from the *N*-linked **8,9**, we synthesized matching *C*-linked compounds of the benzo scaffolds and assayed them against LRTM (Supplementary Schemes S1 & S3 and Supplementary Figure S1). To our surprise, **10**, which is the closest structural analogue to **2-4**, is virtually inactive and related aminothiazole-containing analogues **11** and **12** are slightly more active but with IC<sub>50</sub> values  $\geq 10 \mu\text{M}$  (Supplementary Figure S1). The relative inactivity of the matched benzo analogues **10-12** ( $\geq 10 \mu\text{M}$ ) and ATP site

analogues **5** and **6** ( $\geq 6 \mu\text{M}$ ) demonstrates that the linker in the C-linked bivalent inhibitors **2** and **3** (51-59 pM) enables the markedly improvement in potency. Analysis of these  $\text{IC}_{50}$  values allows for the estimation of the higher-limit of the linking coefficients, which are consistent with superadditivity ( $E < 1$ ), as done previously (Supplementary Table S1).<sup>28,30</sup> While the overall values are estimated to be  $0.5\text{-}1.0 \text{ M}^{-1}$ , they represent upper limits due to the C-linked allosteric **10** exhibiting virtually no activity against LRTM. The estimate of a lower-limit linking coefficient less than 1 is in line with **2** and **3** exhibiting superadditivity and confirms the C-linked amide as one of only a few examples of linked bivalent compounds exhibiting this behavior.<sup>20,21,28</sup> Corresponding calculations for lower limits of N-linked bivalent **1** indicate that this compound has distinctively higher linking coefficients ( $>1.9 \times 10^7 \text{ M}^{-1}$ ) in line with the  $\text{IC}_{50}$  activity measurements and show  $10^7$ -fold differences compared to C-linked amides.

**Table 1.** Biochemical EGFR  $\text{IC}_{50}$  values (nanomolar) against WT and mutant EGFR kinase domains.

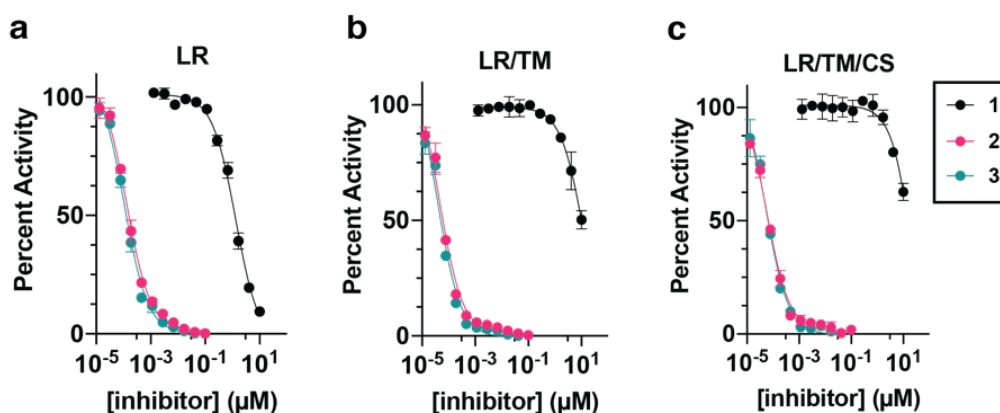
Compound	WT	LR	LRTM	LRTMCS
<b>1</b>	> 10000	1300 $\pm$ 100	> 10000	> 10000
<b>2</b>	< 10	1.5 $\pm$ 0.1	0.059 $\pm$ 0.005	0.064 $\pm$ 0.004
<b>3</b>	< 10	1.2 $\pm$ 0.09	0.051 $\pm$ 0.005	0.063 $\pm$ 0.005
<b>4</b>	- <sup>a</sup>	- <sup>a</sup>	- <sup>a</sup>	1.8 $\pm$ 0.3
<b>5</b> <sup>45</sup>	n.d.	n.d.	5800 $\pm$ 300	6000 $\pm$ 500
<b>6</b>	n.d.	n.d.	> 10000	> 10000
<b>7</b> <sup>44</sup> (LN2057)	- <sup>a</sup>	- <sup>a</sup>	- <sup>a</sup>	130 $\pm$ 40
<b>8</b> (DDC4002) <sup>50</sup>	> 1000	690 $\pm$ 120	39 $\pm$ 4	59 $\pm$ 8
<b>9</b> (EAI002) <sup>9</sup>	> 1000	n.d.	52	n.d.

Total enzyme concentrations WT EGFR 10 nM, LR 0.1 nM, LRTM at 0.02 nM and LRTMCS at 0.02 nM. <sup>a</sup>An  $\text{IC}_{50}$  value is not adequate to describe the potency of a covalent inhibitor. n.d. – “not determined.”

**Table 2.** Time-dependent kinetic parameters for **4** targeting WT, LR, and LRTM EGFR recombinant proteins.

<b>4</b>	WT	LR	LRTM
$k_{\text{inact}}/K_{\text{I}}$ ( $\text{M}^{-1}\text{s}^{-1}$ )	2500±30	14100±300	1070±40
$k_{\text{inact}}$ ( $\text{min}^{-1}$ )	0.18±0.004	0.52±0.02	0.10±0.004
$K_{\text{I}}$ ( $\mu\text{M}$ )	1.20±0.04	0.61±0.03	1.6±0.1

Values obtained from global fits of progress curves ± standard errors. Total enzyme concentrations were WT EGFR at 2.0 nM, LR at 1.0 nM, and LRTM at 2.0 nM.

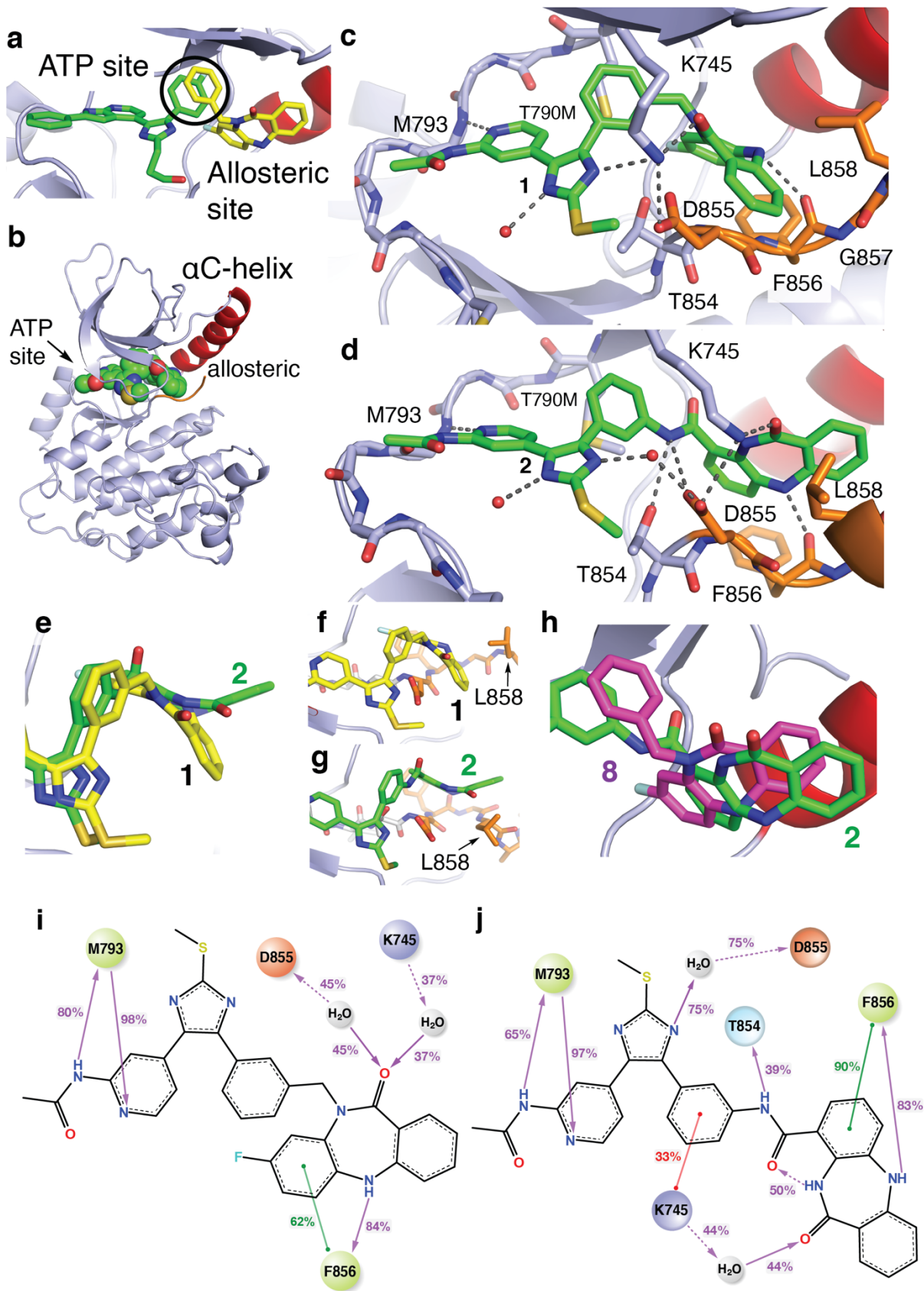


**Figure 1.** The linker structure bridging the ATP and allosteric site enables significantly altered biochemical activities of EGFR bivalent inhibitors. HTRF-based biochemical activity dose-response curves for the reversible binding bivalent EGFR inhibitors **1-3** against A) LR, B) LRTM, and C) LRTMCS mutant EGFR recombinant kinase domains. Total enzyme concentrations: LR 0.1 nM, LRTM and LRTMCS at 0.02 nM.

To characterize the binding modes of these bivalent inhibitors inspired by overlapping features in cocrystal structures (Figure 2A), we determined X-ray cocrystal structures through soaking EGFR(T790M/V948R) crystals with the compounds, which reliably crystallize EGFR in the inactive ( $\alpha\text{C}$ -helix “out”) conformation (Figure 2B, , Supplementary Table S2). A 2.1 Å-resolution cocrystal structure of **1** shows the imidazole and benzo groups bound within the ATP and allosteric sites, respectively, with the benzo moiety adopting an “outward” conformation (Figure 2C, Supplementary Figure S2A-C). Analogously, a 2.2 Å-resolution cocrystal structure of **2** indicates that this compound is bound identically at the ATP site as **1**, but with an opposite “inward” conformation within the allosteric pocket (Figure 2D-E, Supplementary Figure S2D-F). Additional intermolecular interactions are observed for **2** such as H-bonding with T854 and D855 enabled by the *C*-linked amide, which are not possible in the *N*-linked methylene **1** (Figure 2C-D). The side chain of K745, the catalytic lysine, exhibits a “swing” toward the benzo ketone in the case of **2** binding opening a position on the imidazole, which now binds a solvent water (Figure 2D). The



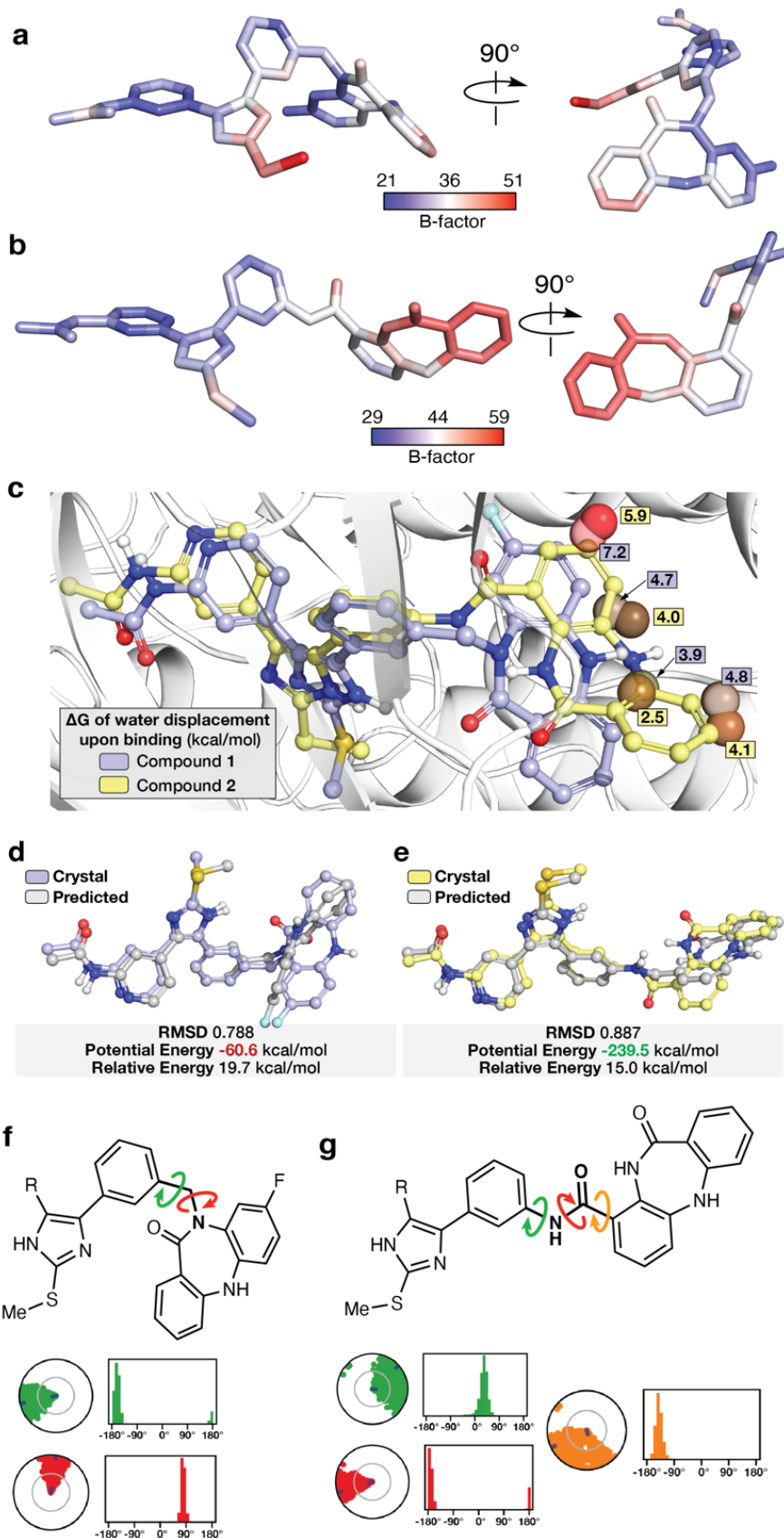
conformation of the allosteric benzo moiety influences the position of the A-loop for the cocrystal structure of **1** and **2** (Figure 2F-G, Supplementary Figure S3). Despite the parent allosteric inhibitors **8** and **9** being best matched in terms of the *N*-linked **1**, the binding conformation of the *C*-linked **2** corresponds most closely to the allosteric inhibitor, **8** (Figure 2H). Despite this difference in binding mode, the length of the linker is comparable between **1** and **2** (Supplementary Figure S4). To gain a more complete understanding of the inhibitors binding and provide the insights in the activity differences, we have performed 20  $\mu$ s long molecular dynamic (MD) simulations (10 replicas x 1  $\mu$ s per compound) based on our cocrystal structures of **1** and **2** (Supplementary Table S3, Supplementary Figure S5-S6). We find excellent correspondence of the ligand interaction patterns between the simulations and experimental structures including some minor variations with respect to water-mediated H-bonds with ligands not evident from the cocrystal structures (Figure 2I-J). Additionally, computer-aided docking has provided a pose for the covalent **4** similar to what is observed for **2** with the expected orientation for covalent bond formation with C797 and consistent with our earlier structural and functional studies (Supplementary Figure S7).<sup>44,49</sup> These cocrystal structures and MD simulations indicate these structurally related compounds with different linker structures exhibit alternative inhibitor binding modes within the allosteric site, side chain orientations, and intermolecular interactions.



**Figure 2. Binding modes of bivalent EGFR inhibitors exhibit linker-dependent conformations within the allosteric pocket.** A) The binding modes of ATP-site trisubstituted imidazole inhibitor (green, PDB ID 6V5N) and the allosteric inhibitor DDC4002 (yellow, PDB ID 6P1D). The overlapping phenyl rings of these two compounds (bold circle) is the group at which the bivalent inhibitors are merged. B) Overall protein-ligand structure of the EGFR(T790M/V948R) kinase domain in the  $\alpha$ C-helix ( $\alpha$ C) outward conformation in complex with **1** (green spheres, PDB ID 8FV3). C) Active site view and binding mode of **1** showing ATP-site binding and “outward” benzo conformation at the allosteric site (PDB ID 8FV3). D) Active site view and binding mode of **2** showing ATP-site binding and “inward” benzo conformation at the allosteric site **2** (PDB ID 8FV4) in complex with EGFR(T790M/V948R). E) Top view overlay of **1** and **2** cocrystal structures within the allosteric pocket demonstrating the full conformational change of the benzo moiety. The conformation of the allosteric moiety influences the positioning of L858 in F) **1** and G) **2**. H) View of allosteric pocket featuring an overlay of the C-linked bivalent **2** (PDB ID 8FV4) and allosteric inhibitor **8** (PDB ID 6P1D). 2D-representation of compound **1** I) and compound **2** J) interaction frequencies with EGFR(T790M/V948R) based on 10 $\mu$ s/compound MD simulations. Only interactions occurring in more than 20% of the simulation time are shown (full data is available in Supplementary Table S3, Supplementary Figure S5). The residue and interaction color schemes are consistent for **I**) and **J**). Polar residues are blue, hydrophobic residues are green, negative charged residues are orange, and positive charged are purple. A green line represents  $\pi$ - $\pi$  stacking, a red line represents the  $\pi$ -cations, and a purple line represents the H-bonds. A dashed line is used as an indication of side-chain interaction and a straight line of the backbone one. The interaction strength along the simulation time is shown by the percentage on the line.

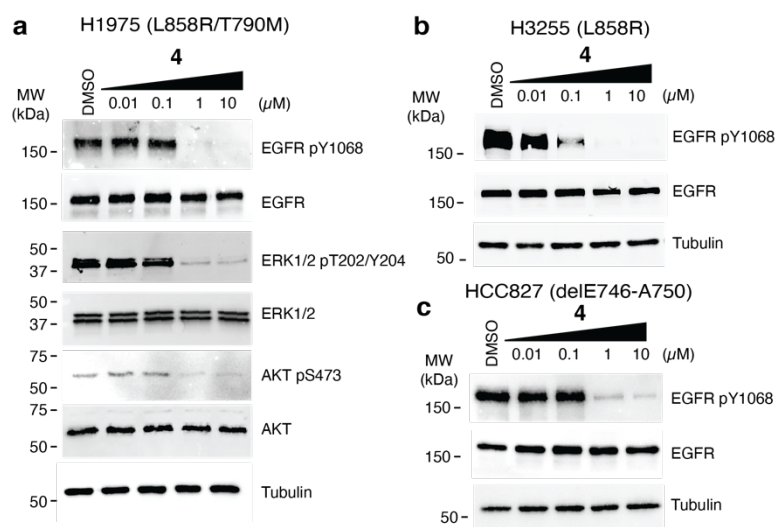
The wide range of potency observed for the C-linked **2** and the N-linked **1** motivated us to more completely understand the structural basis that enables this difference in activity. Appreciating that mobility can contribute to binding, we assessed compound rigidity from crystallographic B-factors of the bound ligands **1** and **2** (Figure 3A-B, Supplementary Figure S8).<sup>51</sup> This is made possible due to several commonalities shared between these cocrystal structures, including as they originate from the same protein, unit cells, and atomic resolutions (2.1 Å for **1** and 2.2 Å for **2**). Generally, the ATP-binding imidazole in both compounds are comparably rigid with B-factors below the structure average while a notable increase is observed for the allosteric moiety in **2** and to a much lesser extent in **1** (Figure 3A-B). To gain deeper insight, we performed generalized Born and surface area solvation (MM-GBSA) calculations on MD trajectories using our cocrystal structures (Figure 1 C-D). These calculations provide Free energies of binding where **2** exhibits greater affinity than **1** ( $\Delta\Delta G = 9.5$  kcal/mol), consistent with the difference in IC<sub>50</sub> values (Table 1), which is enabled by enhanced van der Waals and H-bonding interactions (Supplementary Table S4-5, Supplementary Figure S9). Additionally, MM-GBSA ligand energy calculations indicate a ~3.4-fold lower energy for **2** compared to **1**, implying that **2** possesses a greater degree of structural complementarity within the kinase binding sites (Supplementary Table S4, Supplementary Figure S10). Superior binding of **2** is also aided by the “inward” benzo binding mode as this conformation is capable of full displacement of energetically unfavorable water molecules and the “outward” conformation of **1** does not allow complete displacement (Figure 3C, Supplementary Figure S11). Further energetic analysis indicates that the potential energy of **2** pertaining to conformation within the binding mode of **2** is ~4-fold more favorable than the corresponding conformation of **1** (Figure 3D-E). The overall pictures obtained by the

crystallographic B-factors and MD simulations indicate that the superior potency of the *C*-linked compounds is due to a variety of factors that all stem from the structure of the linker allowing for improved mobility and pocket complementarity within the allosteric site. How the structure of the linker impacts the binding mode of these compounds is best visualized in terms of torsion angles observed in the MD simulations (Figure 3F-G, Supplementary Figure S12). The most unique rotatable bond in **2** is the C-C bond that connects the linker amide to the benzo via the back pocket phenyl ring and allows for enhanced mobility of the group within the allosteric pocket (orange arrow Figure 3G). This rotatable bond is the key structural element that allows for tight binding of this compound to EGFR since the other rotatable bonds in the linkers of **1** and **2** are comparably rigid and anchored to the relatively static ATP-site imidazole. These experimental and theoretical studies reveal the molecular factors that enable effective bivalent inhibitor binding, which can all be attributed to the “inward” conformation within the allosteric pocket enabled by the enhanced linker-enabled mobility of the *C*-linked scaffold.



**Figure 3. Structural basis for linker-dependent strong binding of the C-linked bivalent inhibitor 2.** Variations in crystallographic B-factors of the bound ligands A) **1** and B) **2** in complex with EGFR. The blue-to-red gradient is averaged to the overall average B-factors are 36 for **1** (PDB ID 8FV3) and 44 for **2** (PDB ID 8FV4). C) WaterMap simulation of the unbound EGFR(T790M/V948R) binding pocket for compounds **1** and **2**, highlighting unfavorable (brown spheres) to highly unfavorable hydration sites (red spheres) and their relative location to bound **1** and **2**. Color-coded  $\Delta G$  values reflect the compound associated with each water molecule (See Supplementary Figure S11 for complete analysis). Conformational analysis of compound D) **1** and E) **2** (MacroModel). The potential energy values highlight that **2** adopts a conformation that is over 3-fold more energetically favorable in the binding site compared to **1** in the binding site. Relative potential energy indicates the energy difference to the lowest energy conformation in the predicted set and highlights a closer alignment to the ideal conformation for **2** compared to **1**. The root-mean-square deviation (RMSD) was used to justify the ligand conformation closest to the crystal structure. Ligands are displayed with ball and stick representations using violet and yellow colors for carbon-atoms of **1** and **2**, respectively. Chemical structures of F) **1** and G) **2** denoting linker rotatable bonds and their torsion angles from MD simulations (see Supplementary Figure S12 for complete analysis).

We next sought to gauge the biological activity of our C-linked scaffolds as novel ATP-allosteric bivalent inhibitors. The Michael acceptor-containing C-linked analogue **4**, designed to target C797 as done previously,<sup>44,46-48</sup> was found most effective at suppressing LRTM phosphorylation (pY1068) as well as downstream pERK and pAKT in the human NSCLC cell line H1975 below 1  $\mu\text{M}$  concentration dosed for 6 hours (Figure 4A). Additional studies with H3255 (LR), H3255GR (LRTM) and HCC827 (exon19 delE746-A750) cells exhibit analogous suppression of EGFR pY1068 phosphorylation, slightly better potency in H3255 and H3255GR, cells indicating that **4** broadly targets diverse EGFR mutations (Figure 4B-C, Supplementary Figure S13). The reversible binding **2-3** are also effective in H1975 cells, however to a lesser extent than **4**, while the N-linked **1** exhibits limited ability to suppress phosphorylation (Supplementary Figure S13). We have also determined antiproliferation effects in human cancer cell lines H1975 and HCC827 treated with **4**, which show that this novel EGFR inhibitor is active at concentrations  $\sim 100\text{-}500$  nM overall  $\sim 60$ -fold less potent than AZD9291 in both cell lines (Supplementary Table S6, Supplementary Figure S14). While activity in H1975 and H3255 cells is expected based on biochemical data (Figure 1), **4** is unexpectedly effective against HCC827 cells that harbor the prominent EGFR exon19 delE746-A750 mutation as allosteric pocket binding compounds are typically ineffective against this mutation (Figure 4C).<sup>52</sup> Additional antiproliferative activity experiments in Ba/F3 cells are consistent with above observations, and the  $\sim 220$  nM  $\text{EC}_{50}$  value for LR is markedly improved compared to our earlier covalent bivalent ATP-allosteric inhibitors indicating that the benzo-derived scaffold exhibits improved cellular activity (Supplementary Table S7).<sup>44</sup> Furthermore, we confirm that **4** is selective across the kinome exhibiting a selectivity score of  $S(35) = 0.084$  (Supplementary Table S8) and metabolically stable in liver microsome assays (Supplementary Figure S15). These experiments indicate that the C-linked bivalent inhibitor **4** is capable of targeting cellular EGFR in biological context of several prevalent oncogenic activating mutants and motivates further efforts to optimize the potency and medicinal chemistry properties for translational studies.



**Figure 4. The bivalent inhibitor 4 suppresses EGFR phosphorylation across diverse human NSCLC cell lines.** Impact of **4** on EGFR phosphorylation (pY1068) in A) H1975, B) H3255, and C) HCC827 cell line models. A) Dosing of **4** in H1975 cells also diminishes phosphorylation of ERK1/2 and AKT downstream signaling kinases. Western blot experiments performed after 6 hour treatments. Western blots are representative of at least three independent experiments.

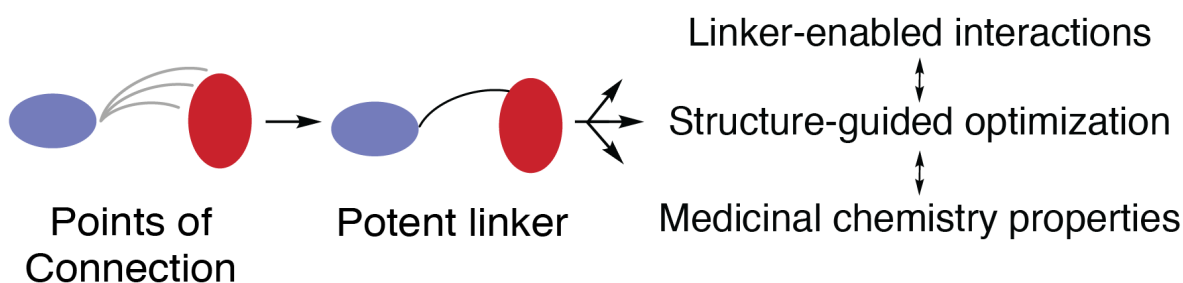
## Discussion

Bivalent (heterobifunctional) compounds are attractive molecules in chemical biology and drug discovery for their unique properties and accessibility toward classically “undruggable” biological targets.<sup>1,2,53-55</sup> However, their complex dual-motif chemical structure is challenging to optimize, especially with respect to the structure of the linking group as subtle changes to structure can have considerable effects on biological function.<sup>22,56,57</sup> Furthermore, the process by which compounds are connected in FBDD, although theoretically guaranteed to yield superiorly active compounds, is rarely achieved despite decades-long efforts.<sup>20-22,30</sup> For these reasons we rationalize that studies of alternatively linked bivalent molecules, which exhibit a broad range of potency, would reveal structural insights that can be used to improve the processes in optimizing bivalent compounds. The bivalent ATP-allosteric EGFR inhibitors reported here provide a structural basis for linker potency and offer previously unconsidered strategies for linker design.

The molecules in this study are informative since they differ only in terms of linker structure and exhibit significantly different biochemical potencies ( $> 10^6$ -fold in LRTM and LRTMCS, Figure 1). The origins of this sizable potency range can be understood on the basis of how these compounds bind to the EGFR kinase domain (Figure 2). The “inward” allosteric benzo conformation of the C-linked scaffold **2** is found to best match the parent allosteric fragment **8** (Figure 2G) and MD simulations indicate that stronger binding of **2** is due, in part, to structural complementarity with the allosteric site. Matching the binding modes of parent fragments is a well-appreciated objective in FBDD linking,<sup>28</sup> which is consistent with the structures and biochemical potencies of **1** versus **2**. Interestingly, the binding mode “inward” conformation of **2** most closely resembles the allosteric-only **8** despite possessing the N-linked structure of **1**, and demonstrates

how minute alterations in linker structure can manifest as major conformational differences in target binding sites.<sup>50</sup> We further characterize the linker structure to be the key factor in enabling benzo mobility provided by variations crystallographic B-factors of **1** and **2** in cocrystal structures (Figures 3A-D, Supplementary Figure S8), MM-GBSA ligand energy calculations (Supplementary Table S4, Supplementary Figure S10) and rotatable bond torsion profiles (Supplementary Figure S12), which attributes the enhanced binding of **2** to a single unique C-C bond within the linker (orange in Figure 3G, Supplementary Figure S12). This potency-enabling aspect of the structure of **2** is only possible by shifting the linker connection (*N*- to *C*-linked) on the benzo motif (Table 1, Figure 1 & 4). Several previous studies have highlighted the importance of enabling proper flexibility for effective linkers,<sup>30,58</sup> but to our knowledge our compounds are distinctive with respect to the range of potency (>10<sup>6</sup>-fold) and binding energy despite the relatively subtle differences in linker structure (Figure 1,  $\Delta G$  Supplementary Table S4).

Given the notable potency enhancement between the *N*- and *C*-linked scaffolds, we further considered how these compounds may inform strategies for more efficient linker optimization. Several recent reviews on FBDD linking offer perspectives from dozens of examples, but very little is understood regarding efficient design strategies beyond iterative trial and error.<sup>20-22,30,59</sup> If we consider the evolution of the *N*- to *C*-linked scaffold, the structural basis for the large potency enhancement is related to how the linker is connected with the benzo allosteric motif. The shift in the linker “point of connection” from **1** to **2** introduces several attributes that are not necessarily possible in varying what is considered to be more conventional linker structural properties, such as length or alternative functional groups. To our knowledge, linker optimization of this sort has not been discussed previously, however, it is worth noting the parallels with pioneering work by Fesik and co-workers and their work pertaining to Bxl-x<sub>L</sub> inhibitors.<sup>60,61</sup> Specifically their linked lead compound, connected through a central *trans*-olefin, was shown to exhibit improved binding by a shift to a linear acylsulfonamide linker structure that varies with respect to different point of connection, and ultimately lead to the FDA-approved drug venetoclax.<sup>61,62</sup> In terms of general optimization strategies, we propose that once linked, early-phase exploration of points of connection to the linked motifs are likely to result in large changes to drug potency (Figure 5). While understandably cumbersome in terms of synthesis, we rationalize that arriving at the optimal linker structure at the onset of a project is the most efficient scenario and as such would streamline lead optimization strategy. Additionally, this strategy may be especially helpful in cases where structural information regarding the linked molecule and target is challenging to obtain.



**Figure 5. Points of connection serve as an effective strategy for linker optimization in early-phase drug discovery efforts.** To best optimize for compound binding, flexibility and fragment binding mode can be altered through alternative points of connection between the linker and



fragment would be surveyed. Subsequent optimization of linker-enabled interactions as well as structure-guided optimization for potency and medicinal chemistry properties would follow using more conventional lead optimization strategies.

The iterative nature of drug optimization makes developing new medicines reliant on brute force structural explorations and serendipity,<sup>20,59,63</sup> but results from our work offer new strategies to more effectively optimize highly sensitive regions of bivalent inhibitors. The dissection of the structural basis of our alternatively linked molecules informs a new perspective on linker design, and also demonstrates that insights can be afforded through studies of linker-dependent effects on three-dimensional binding and how they translate to function. We rationalize that multi-site model systems, such as EGFR and others, represent constructive frameworks for evaluating drug design strategies amenable to structure determination with the ultimate goal to streamline structure-based drug optimization. Additionally, this work presents new examples of bivalent scaffolds unique within the diverse repertoire of EGFR tyrosine kinase inhibitors, which we show here are active in human NSCLC cell lines across the most prevalent oncogenic activating mutations (L858R, T790M, and exon19del). Future work will be directed toward expanding evaluations of bivalent inhibitor applicability broadly across the human kinome and improve our understanding of structure-based drug design.

## Methods.

### Protein expression and purification

The EGFR kinase domain (residues 696-1022) was cloned into pTriEx with an N-terminal 6xHis-glutathione S-transferase (GST) fusion tag followed by a TEV protease cleavage site. EGFR WT, L858R, L858R/T790M, L858R/T790M/C797S was expressed after baculoviral infection in SF9 cells and EGFR(T790M/V948R) was expressed in SF21 cells. Briefly, cells were pelleted and resuspended in lysis buffer composed of 50 mM Tris pH 8.0, 500 mM NaCl, 1 mM tris(2-carboxyethyl) phosphine (TCEP), and 5% glycerol. Cells were lysed via sonication prior to ultracentrifugation at >200,000 g for 1 h. Imidazole pH 8.0 was added to the supernatant for a final concentration of 40 mM and flowed through a column containing Ni-NTA agarose beads. The resin was washed with lysis buffer supplemented with 40 mM imidazole and eluted with lysis buffer containing 200 mM imidazole. Eluted EGFR kinase domain was dialyzed overnight in the presence of 5% (w/w) TEV protease against dialysis buffer containing 50 mM Tris pH 8.0, 500 mM NaCl, 1 mM TCEP, and 5% glycerol. The cleaved protein was passed through Ni-NTA resin to remove the 6xHis-GST fusion protein and TEV prior to size exclusion chromatography on a prep-grade Superdex S200 column in 50 mM Tris pH 8.0, 500 mM NaCl, 1 mM TCEP, and 5% glycerol. Fractions containing EGFR kinase of  $\geq 95\%$  purity as assessed by Coomassie-stained SDS-PAGE were concentrated to approximately 4 mg/mL as determined by Bradford assay or absorbance.

### Crystallization and structure determination

EGFR(T790M/V948R) pre-incubated with 1 mM AMP-PNP and 10 mM MgCl<sub>2</sub> on ice was prepared by hanging-drop vapor diffusion over a reservoir solution containing 0.1 M Bis-Tris (pH = 5.5), 25% PEG-3350, and 5 mM TCEP (buffer A for crystals soaked with compound 1) or 0.1 M Bis-Tris (pH = 5.7), 30% PEG-3350 TCEP (buffer B for crystals soaked with compound 2). Drops containing crystals in buffer A and B were exchanged with solutions of both buffers

containing ~1.0 mM **1** or **2** were exchanged three times for an hour and then left to soak overnight. Crystals were flash frozen after rapid immersion in a cryoprotectant solution with buffer A or BA containing 25% ethylene glycol. X-ray diffraction data of T790M/V948R-compound **2** crystals was collected at 100K at the Advanced Light Source a part of the Northeastern Collaborative Access Team (NE-CAT) on Beamline 24-ID-C. While data on T790M/V948R-compound **1** crystals were collected at 100K at the National Synchrotron Light Source II 17-ID-2<sup>64</sup>. Diffraction data was processed and merged in Xia2 using aimless and dials. The structure was determined by molecular replacement with the program PHASER using the inactive kinase domain EGFR(T790M/V948R) kinase from our previous work excluding the LN3844 ligand (PDB 6WXN). Repeated rounds of manual refitting and crystallographic refinement were performed using COOT and Phenix. The inhibitor was modeled into the closely fitting positive  $F_o - F_c$  electron density and then included in following refinement cycles. Statistics for diffraction data processing and structure refinement are shown in Table S1. Due to a mixture of difference map density with contributions from both AMP-PNP and **2** in in Chain C we elected to leave this chain without bound ligands.

### Time-dependent Kinase Inhibition Assays

Biochemical assays were performed with commercially available EGFR WT, cytoplasmic domain (669-1210), GST-tagged, Carna (Cat#/Lot#: 08-115/21CBS-0127H), EGFR [L858R], cytoplasmic domain (668-end), GST-tagged, SignalChem (Cat#/Lot#: E10-122BG/D2411-4), EGFR [T790M/L858R] cytoplasmic domain (669-1210), GST-tagged, Carna (Cat#/Lot#: 08-510/12CBS-0765M). Reactions were performed with kinase domain enzyme concentrations of WT EGFR at 2.0 nM, LR at 1.0 nM, and LRTM at 2.0 nM in final solutions of 52 mM HEPES pH 7.5, 1 mM ATP, 0.5 mM TCEP, 0.011% Brij-35, 0.25% glycerol, 0.1 mg/ml BSA, 0.52 mM EGTA, 10 mM MgCl<sub>2</sub>, 15 μM Sox-based substrate (AQT0734). BSA was not included in this experiment to prevent interference with irreversible inhibitor characterization via off-target binding. All reactions were run for 240 minutes at 30 °C. Time-dependent fluorescence from the Sox-based substrate was monitored in PerkinElmer ProxiPlate-384 Plus, white shallow well microplates (Cat. #6008280) Biotek Synergy Neo 2 microplate reader with excitation (360 nm) and emission (485 nm) wavelengths. **4** was dosed between 0 and 10 μM in 24-point curves with 1.5-fold dilutions. Fluorescence, determined with identical reactions but lacking purified enzyme or crude cell lysate was subtracted from the total fluorescence signal for each time point, with both determined in duplicate, to obtain corrected relative fluorescence units (RFU). Corrected RFU values then were plotted vs. time and the reaction velocity for the first ~40 min (initial reaction rates) was determined from the slope using GraphPad Prism (La Jolla, CA) with units of RFU/min.

### HTRF Assays

Biochemical assays for EGFR domains were carried out using a homogeneous time-resolved fluorescence (HTRF) KinEASE-TK (Cisbio) assay, as described previously.<sup>65</sup> Assays were optimized for ATP concentration of 100 μM with enzyme concentrations WT EGFR 10 nM, L858R 0.1 nM, L858R/T790M at 0.02 nM and L858R/T790M/C797S at 0.02 nM. Inhibitor compounds in DMSO were dispensed directly in 384-well plates with the D300 digital dispenser (Hewlett Packard) followed immediately by the addition of aqueous buffered solutions using the Multidrop Combi Reagent Dispenser (Thermo Fischer). Compound IC<sub>50</sub> values were determined by 11-point inhibition curves (from 1.0 to 0.00130 μM) in triplicate. The data was

graphically displayed using GraphPad Prism version 7.0, (GraphPad software). The curves were fitted using a non-linear regression model with a sigmoidal dose response.

### **Cellular Antiproliferative Activity Assays**

H1975 and HCC827 cells were obtained from the lab of Dr. Pasi Jänne (Dana-Farber Cancer Institute, 2022) cultured at 37°C in RPMI 1640 media (Corning, 1004-CV) supplemented with 10% fetal bovine serum (Tissue Culture Biologicals, 35-010-CV) and 1% penicillin and streptomycin (P/S, Corning, 30-002-CI) and seeded overnight in a 96-well plate at a density of 60000 cells/mL. Cells were dosed with inhibitors to a final DMSO 1% in triplicate for 37°C for 72 hours. Cellular inhibition of growth was assessed by MTT viability assay according to the manufactures protocol (OZ Biosciences). Parental Ba/F3 cells was a generous gift from the laboratory of Dr. David Weinstock (in 2014), Dr. Pasi Jänne (2020) both of the Dana-Farber Cancer Institute and was used to generate the wildtype EGFR, L858R, and L858R/T790M EGFR mutant Ba/F3 cells as performed previously.<sup>52 66</sup> Briefly, all Ba/F3 cells were cultured in RPMI1640 media with 10% fetal bovine serum and 1% penicillin and streptomycin. All cell lines were tested negative for *Mycoplasma* using Mycoplasma Plus PCR Primer Set (Agilent) and were passaged and/or used for no longer than 4 weeks for all experiments. Assay reagents were purchased from MilliporeSigma (Cat# R7017-5G). Ba/F3 cells were plated and treated with increasing concentrations of inhibitors in triplicate for 72 hours. Compounds were dispensed using the Tecan D300e Digital Dispenser. Cellular growth or the inhibition of growth was assessed by resazurin viability assay to a final 1% DMSO. All experiments were repeated at least 3 times and values were reported as an average with standard deviation.

### **Western blotting**

H1975, H3255, H3255GR and HCC827 lung adenocarcinoma cells were cultured in RPMI 1640 media (Corning, 1004-CV) supplemented with 10% FBS (Tissue Culture Biologicals, 35-010-CV) and 1% penicillin-streptomycin (P/S, Corning, 30-002-CI). H1975 and HCC-827 cells were seeded in 6-well plates with 400,000 cells per well and incubated overnight. H3255 and H3255GR cells were seeded in 12-well plates with 200,000 cells per well and were grown to confluency after 48 hours. Cells were treated the next day after replacing fresh media for 6 hours as indicated in the figure legends. Culture medium was removed, cells washed with PBS, and lysed with lysis buffer containing 5M NaCl, 1M TRIS pH 8.0, 10% SDS, 10% Triton X-100 and a tablet of protease and phosphatase inhibitor. Protein lysate concentration was analyzed using Pierce BCA kit (ThermoFisher, 23225). Protein samples (10 µg) were resolved on hand-cast Criterion Tris-HCl protein gels. Primary antibodies used; phospho-EGFR (Tyr1068; #2234, 1:1,000), EGFR (#4267; 1:1,000), phospho-AKT (Ser473; #4060, 1:1,000), AKT (#9272, 1:1,000), phospho-ERK1/2 (Thr202/Tyr204; #4370, 1:1,000), and ERK1/2 (#4695, 1:1,000) antibodies; were purchased from Cell Signaling Technology. Secondary Goat anti-rabbit IgG starbright blue 700 (Biorad, 64484700) and Anti-tubulin hFAB Rhodamine Tubulin (Bio-Rad, 64512248). Western blots were visualized with the ChemiDoc MP imager (Bio-Rad) utilizing the Image Lab Touch Software (version 2.4.0.03). All Western blot images are representatives of at least 3 independent replicate experiments.

### **Metabolic stability in Human Liver Microsomes (HLM)**

Pooled liver microsomes from humans (male) were purchased from *Sekisui XenoTech, LLC*, Kansas City, KS, USA. Metabolic stability assays were performed in the presence of an NADPH-

regenerating system consisting of 5 mM glucose-6-phosphate, 5 U/mL glucose-6-phosphate dehydrogenase, and 1 mM NADP<sup>+</sup>. Liver microsomes (20 mg/mL), NADPH-regenerating system, and 4 mM MgCl<sub>2</sub>·6 H<sub>2</sub>O in 0.1 M TRIS-HCl-buffer (pH 7.4) were preincubated for 5 min at 37 °C and 750 rpm on a shaker. The reaction was started by adding the preheated compound at 10 mM resulting in a final concentration of 0.1 mM. The reaction was quenched at selected time points (0, 10, 20, 30, 60, and 120 min) by pipetting 100 µL of internal standard (ketoprofen) at a concentration of 150 µM in acetonitrile. The samples were vortexed for 30 s and centrifuged (21910 relative centrifugal force, 4 °C, 20 min). The supernatant was used directly for LC-MS analysis. All compound incubations were conducted at least in triplicates. Additionally, a negative control containing BSA (20 mg/mL) instead of liver microsomes and a positive control using verapamil instead of compound were performed. A limit of 1% organic solvent during incubation was not exceeded. Sample separation and detection were performed on an *Alliance 2695 Separations Module* HPLC system (*Waters Corporation*, Milford, MA, USA) equipped with a *Phenomenex Kinetex 2.6 µm XB-C18 100 Å 50 x 3 mm* column (*Phenomenex Inc.*, Torrance, CA, USA) coupled to an *Alliance 2996 Photodiode Array Detector* and a *MICROMASS QUATTRO micro API* mass spectrometer (both *Waters Corporation*, Milford, MA, USA) using electrospray ionization in positive mode. Mobile phase A: 90% water, 10% acetonitrile and additionally 0.1% formic acid (v/v), mobile phase B: 100% acetonitrile with additionally 0.1% formic acid (v/v). The gradient was set to: 0-2.5 min 0% B, 2.5-10 min from 0 to 40% B, 10-12 min 40% B, 12.01-15 min from 40 to 0% B at a flow rate of 0.7 mL/min. Samples were maintained at 10 °C, the column temperature was set to 20 °C with an injection volume of 5 µL. Spray, cone, extractor, and RF lens voltages were at 4 kV, 30 V, 8 V and 2 V, respectively. The source and desolvation temperatures were set to 120 °C and 350 °C, respectively, and the desolvation gas flow was set to 750 L/h. Data analysis was conducted using *MassLynx 4.1* software (*Waters Corporation*, Milford, MA, USA).

### Compound Docking

Computer-aided compound docking was performed with Glide (Schrödinger, LLC, New York, NY, 2021, re. 2021-2) and Maestro 12.8.117. The receptor grid was generated from the EGFR(T790M/V948R) kinase domain from Chain D of PDB ID 8FV4 (compound **2**), and ligands were prepared with LigPrep. The best binding poses were ranked on the basis of the lowest docking and glide score values.<sup>67,68</sup>

### Modeling and structure preparation for MD simulations

Molecular modeling was conducted using Maestro (Schrödinger Release 2023-1, Schrödinger LLC, New York, NY, 2021) and the OPLS4 force field.<sup>69</sup> The crystal structures **1** (PDB ID 8FV3) and **2** (PDB ID 8FV4) were utilized for modeling the complexes of compound **1** and compound **2**, respectively. Prior to modeling, the protein structures were prepared using the Protein Preparation Wizard<sup>70</sup> (Maestro 2021.4, Schrödinger LLC, New York, NY, USA) with default settings, which involved adding hydrogen atoms and correcting any missing side chains. In the compound **1** complex, the disordered residues in the activation loop required were rebuilt as following: The A859-A871 region was rebuilt based on the chain B of the 8FV4 crystal structure using the chimera homology modeling approach, followed by further minimization of the region using the OPLS4 force field within a selected interval and a 3 Å region around the selected residue interval. Additionally, the E872-E874 residues were added using Maestro's cross-link proteins tool, followed by region minimization in the OPLS4 force field. For the compound **2** complex, the L862-A871 region of the activation loop was rebuilt based on the chain B of the 8FV4 crystal structure using chimera homology modeling. Similarly, the E872-K875 residues and S784 residue

were added using Maestro's cross-link proteins tool, followed by region minimization in the OPLS4 force field within a selected interval and a 3 Å region around the selected residue interval. The initial validation of the individual models was assessed using the Ramachandran plot.

### **Molecular Dynamics Simulations**

Desmond MD engine<sup>71</sup> was used for the MD simulations with OPLS4<sup>69</sup> force field. The system was solvated in an orthorhombic box (minimum distance of 10 Å to the edges from the protein). A temperature of 300 K was used for membrane patch pre-equilibration. The water was described with the TIP3P<sup>72</sup> model. The final systems comprised ~44 k atoms. All simulations were run in the NpT ensemble (T = 310 K, Nosé-Hoover method; p = 1.01325 bar, Martyna-Tobias-Klein method) with default Desmond settings. Reversible reference system propagator algorithms (RESPA) integrator with 2 fs, 2 fs, and 6 fs timesteps were used for bonded, near and far, respectively. Short-range coulombic interactions were calculated using 1 fs time steps and 9.0 Å cut-off value, whereas long-range coulombic interactions were estimated using the Smooth Particle Mesh Ewald method, which is a sufficiently good approximation to treat long-range interactions on large timescales.<sup>73</sup> The system was relaxed using the default Desmond protocol before the production simulation.

A total of 20 simulation replicas were prepared individually, each simulated with a random seed for a duration of 1 μs for compound 1 and compound 2. This resulted in a cumulative simulation time of 20 μs (10 replicas x 1 μs x 2 compounds). All production simulations were conducted using consistent settings as previously described. The simulation interaction diagram provided by the Maestro package (Schrödinger, LLC, New York, NY) served as the foundation for the analysis of the simulations.

### **Interaction analyses.**

Protein–ligand interactions, as well as hydrophobic interaction frequency, RMSD, and torsional conformations of rotatable bonds were analyzed by the Simulation Interaction Analysis tool (scripts: event\_analysis.py; analyze\_simulation.py) (Schrödinger LLC). The default settings were used in the definition of the interactions, where the following parameters were applied: H-bonds, a distance of 2.5 Å between the donor and acceptor with  $\geq 120^\circ$  and  $\geq 90^\circ$  for donor and acceptor angles, respectively;  $\pi$ -cation interactions, a 4.5 Å distance between the positively charged and aromatic group;  $\pi$ - $\pi$  interactions, stacking of two aromatic groups face-to-face or face-to-edge; water bridges, a distance of 2.8 Å between the donor and acceptor with  $\geq 110^\circ$  and  $\geq 90^\circ$  for donor and acceptor angles, respectively.

### **MM-GBSA**

The molecular mechanics energies with generalized Born and surface area continuum solvation (MM-GBSA) were calculated with Prime Thermal MM/GBSA.<sup>74,75</sup> Each 2nd frame of MD was used for MM-GBSA calculations (5010 complexes proceeded for an individual complex x 2 ligands). MM-GBSA calculations report energies for the ligand, receptor, and complex structures as well as energy differences relating to strain and binding and are broken down into contributions from various terms in the energy expression.

### **WaterMap**

WaterMap<sup>76,77</sup> simulations were using Maestro, and the system was solvated in TIP3P water box extending at least 10 Å beyond the truncated protein in all directions. 5 ns MD simulation was performed, following a standard relaxation protocol, and the water molecule trajectories were then clustered into distinct hydration sites. Entropy and enthalpy values for each hydration site were calculated using inhomogeneous solvation theory.

### **Conformational analysis**

Conformational analysis of compound 1 and 2 was assessed with Conformational Search tool from MacroModel module (Schrödinger LLC). Analysis was conducted with OPLS4<sup>69</sup> force field force field in a water solvent was employed for the analysis. A mixed torsional/Low mode sampling approach was chosen with the default settings. The Polak-Ribier Conjugate Gradient<sup>78</sup> (PRCG) method with restarts every 3N iterations (maximum of 2500 iterations) was utilized for energy minimization, with a convergence threshold of 0.05. For compound 1, a total of 139 conformers were generated, while for compound 2, 278 conformers were generated. The ligand conformation that best matched the crystal structure was determined by SMARTS superimposition of the structure scaffolds with the reference ligand conformation. The selection of the final conformation was justified based on the lowest superimposition root mean square deviation (RMSD) values obtained from both conformational datasets.

### **Data Visualization**

Results were plotted with Seaborn library for Python 3.8.8.<sup>79</sup> Protein structures were visualized with PyMOL (The PyMOL Molecular Graphics System, Version 2.5.2 Schrödinger, LLC.) Graphical representations of figures were arranged using Adobe Illustrator©.

### **Data availability**

The datasets generated during the in-silico study are available in the Zenodo repository (DOI 10.5281/zenodo.8020238). The available data includes the molecular dynamic trajectories for **1** and **2**, along with the raw data on interaction patterns, separated for interaction components (i.e. Hydrophobic interactions/Hydrogen Bonding), full-component MM/GBSA tables, WaterMap Pymol sessions, and raw output of ligand conformational analysis. (This DOI is reserved, and files will be released for the public access once the manuscript is accepted).

## Chemistry

All starting materials, reagents, and (anhydrous) solvents were commercially available and were used as received without any further purification or drying procedures unless otherwise noted. All NMR spectra were obtained with Bruker Avance 200 MHz and Bruker Avance 400 MHz spectrometers or with a Bruker Avance 600 MHz spectrometer (NMR Department, Institute of Organic Chemistry, Eberhard-Karls-Universität Tübingen) or Bruker Ascend 400 MHz and Bruker Ascend 500 MHz (Magnetic Resonance Center, Department of Chemistry, SUNY at Buffalo). Solvents for NMR are noted in the experimental procedures for each compound. Residual solvent peaks were used to calibrate the chemical shifts. Chemical shifts ( $\delta$ ) are reported in parts per million. Mass spectra were obtained by Advion TLC-MS (ESI) and from the MASS Spectrometry Department (ESI-HRMS), Institute of Organic Chemistry, Eberhard-Karls-Universität Tübingen or by Thermo Scientific LTQ XL Linear Ion Trap Mass Spectrometer (Small Instrument Center, Department of Chemistry, SUNY at Buffalo). The purity of the tested compounds was determined via HPLC analysis on an Agilent 1100 Series LC with a Phenomenex Luna C8 column (150  $\times$  4.6 mm, 5  $\mu$ m), and detection was performed with a UV diode array detector (DAD) at 254 and 230 nm wavelengths via elution condition (A) or on an Agilent 1200 series LC with an Agilent Eclipse XDB-C18 column (4.6x150 mm, 5  $\mu$ m) and detection was performed using an Agilent 1200 series Multiple Wavelength Detector (MWD) at 254, 280, and 305 nm wavelengths via elution condition (B) and was >95%. Elution was carried out with the following conditions: (A) 0.01 M KH<sub>2</sub>PO<sub>4</sub>, pH 2.30 (solvent A), and MeOH (solvent B), 40% B to 85% B in 8 min, 85% B for 5 min, 85% to 40% B in 1 min, 40% B for 2 min, stop time of 16 min, 5  $\mu$ L injection volume, flow rate of 1.5 mL/min, and 25 °C oven temperature or (B) 60% MeOH, 40% 0.1% formic acid in H<sub>2</sub>O, stop time of 16 min, 5  $\mu$ L injection volume, flow rate of 1.5 mL/min, 25°C oven temperature. Thin-layer chromatography (TLC) analyses were performed on fluorescent silica gel 60 F254 plates (Merck) and visualized via UV illumination at 254 and 366 nm. Column chromatography was performed on Davisil LC60A 20–45  $\mu$ m silica from Grace Davison as the stationary phase and Geduran Si60 63–200  $\mu$ m silica from Merck for the precolumn using an Interchim PuriFlash XS 420 automated flash chromatography system.

## Contributions

F.W., B.C.O., E.S., T.D., C.D.P., I.K.S., B.T.O, S.P.C., T.S.B., A.R., B.B., D.U., T.S., E.M., E.S., M.J.E., P.A.H., A.P., S.A.L., D.E.H. conceived and designed the experiments; A.P. and E.S. conducted MD simulations with energy, torsional, and water profiling; F.W. and B.T.O. performed synthesis; B.C.O., T.D., C.D.P., S.P.C., I.K.S., T.S.B., T.S. and B.B. performed biological assays; D.U. and E.M. determined time-dependent kinetic parameters; B.C.O., C.D.P. T.S.B. and D.E.H. performed X-ray crystallography; S.P.C. performed molecular docking; A.R. performed metabolic stability experiments; F.W., B.C.O., S.A.L. and D.E.H. interpreted the data; F.W., B.C.O., S.A.L. and D.E.H. wrote the paper. The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript.

**Acknowledgements** We acknowledge support from startup funds from The State University of New York (DEH) and support by the National Center for Advancing Translational Sciences of the National Institutes of Health under award Number UL1TR001412-08 (BTC K Scholar Award to DEH). S.A.L. and iFIT are funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC 2180-390900677). TüCAD2 is funded by the Federal Ministry of Education and Research (BMBF) and the Baden-Württemberg Ministry of Science as part of the Excellence Strategy of the German Federal and State

Governments. The authors acknowledge CSC-IT Center for Science, Finland for computational resources. National Institutes of Health Grants R01CA201049, R01CA116020, and R35CA242461 (to M.J.E), Roswell Park Alliance Foundation (PAH). Research reported in this publication was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under award Number UL1TR001412-07. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. T.S.B is supported by a Ruth L. Kirschstein National Research Service Award (5F32CA247198-02). This work was partially done in the Drug Discovery Core Facility of Roswell Park Comprehensive Cancer Center supported by National Cancer Institute (R01CA197967 to K.V. Gurova and P30CA016056 to Roswell Park Cancer Center). This work is based on research utilizing resources of the Frontier Microfocusing Macromolecular Crystallography (17-ID-2, FMX) beamline at National Synchrotron Light Source II at Brookhaven National Laboratory to Block Allocation Group 308246. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines (P30 GM124165, P41 GM103403) utilizing resources of the Advanced Photon Source at the Argonne National Laboratory (DE-AC02-06CH11357). We thank Mr. Bryan Renzoni for contributions to synthesis. We thank Dr. David Hangauer, Dr. Andrew Gulick, and Dr. Steven Diver for insightful comments and discussions. We thank Dr. Michael Malkowski and Dr. Liang Dong for SF21 cells and access to their tissue culture lab. We also acknowledge Dr. Diana Monteiro and Dr. Edward Snell for access to laboratory space and equipment for protein purification and crystallization resources.

#### Literature Cited

- 1 Corson, T. W., Aberle, N. & Crews, C. M. Design and Applications of Bifunctional Small Molecules: Why Two Heads Are Better Than One. *ACS Chemical Biology* **3**, 677-692, doi:10.1021/cb8001792 (2008).
- 2 Schreiber, S. L. The Rise of Molecular Glues. *Cell* **184**, 3-9, doi:<https://doi.org/10.1016/j.cell.2020.12.020> (2021).
- 3 Henning, N. J. *et al.* Deubiquitinase-targeting chimeras for targeted protein stabilization. *Nature chemical biology* **18**, 412-421 (2022).
- 4 Siriwardena, S. U. *et al.* Phosphorylation-inducing chimeric small molecules. *J. Am. Chem. Soc.* **142**, 14052-14057 (2020).
- 5 Wang, W. W. *et al.* Targeted Protein Acetylation in Cells Using Heterobifunctional Molecules. *J. Am. Chem. Soc.* **143**, 16700-16708, doi:10.1021/jacs.1c07850 (2021).
- 6 Schapira, M., Calabrese, M. F., Bullock, A. N. & Crews, C. M. Targeted protein degradation: expanding the toolbox. *Nature reviews Drug discovery* **18**, 949-963 (2019).
- 7 Chen, P.-H. *et al.* Modulation of phosphoprotein activity by phosphorylation targeting chimeras (PhosTACs). *ACS Chemical Biology* **16**, 2808-2815 (2021).
- 8 Gibson, W. J. *et al.* Bifunctional small molecules that induce nuclear localization and targeted transcriptional regulation. *bioRxiv*, 2023.2007.2007.548101, doi:10.1101/2023.07.07.548101 (2023).
- 9 Gourisankar, S. *et al.* Rewiring cancer drivers to activate apoptosis. *Nature*, doi:10.1038/s41586-023-06348-2 (2023).
- 10 Burgess, L. E. *et al.* Potent selective nonpeptidic inhibitors of human lung tryptase. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8348-8352 (1999).



- 11 Waring, M. J. *et al.* Potent and selective bivalent inhibitors of BET bromodomains. *Nature chemical biology* **12**, 1097-1104 (2016).
- 12 Tanaka, M. *et al.* Design and characterization of bivalent BET inhibitors. *Nature chemical biology* **12**, 1089-1096 (2016).
- 13 Matsui, Y. *et al.* Discovery and structure-guided fragment-linking of 4-(2,3-dichlorobenzoyl)-1-methyl-pyrrole-2-carboxamide as a pyruvate kinase M2 activator. *Bioorganic Med. Chem. Lett.* **25**, 3540-3546, doi:<https://doi.org/10.1016/j.bmc.2017.05.004> (2017).
- 14 Bamborough, P. *et al.* Design and Characterization of 1,3-Dihydro-2H-benzo[d]azepin-2-ones as Rule-of-5 Compliant Bivalent BET Inhibitors. *ACS Med. Chem. Lett.*, doi:10.1021/acsmchemlett.3c00242 (2023).
- 15 Hill, Z. B., Perera, B. G. K. & Maly, D. J. A Chemical Genetic Method for Generating Bivalent Inhibitors of Protein Kinases. *J. Am. Chem. Soc.* **131**, 6686-6688, doi:10.1021/ja900871y (2009).
- 16 Hill, Z. B., Perera, B. G. K., Andrews, S. S. & Maly, D. J. Targeting Diverse Signaling Interaction Sites Allows the Rapid Generation of Bivalent Kinase Inhibitors. *ACS Chemical Biology* **7**, 487-495, doi:10.1021/cb200387g (2012).
- 17 Meyer, S. C., Shomin, C. D., Gaj, T. & Ghosh, I. Tethering Small Molecules to a Phage Display Library: Discovery of a Selective Bivalent Inhibitor of Protein Kinase A. *J. Am. Chem. Soc.* **129**, 13812-13813, doi:10.1021/ja076197d (2007).
- 18 Nowak, R. P. *et al.* Plasticity in binding confers selectivity in ligand-induced protein degradation. *Nature Chemical Biology* **14**, 706-714, doi:10.1038/s41589-018-0055-y (2018).
- 19 Tran, N. L., Leconte, G. A. & Ferguson, F. M. Targeted Protein Degradation: Design Considerations for PROTAC Development. *Current Protocols* **2**, e611, doi:<https://doi.org/10.1002/cpz1.611> (2022).
- 20 Bancet, A. *et al.* Fragment Linking Strategies for Structure-Based Drug Design. *J. Med. Chem.* **63**, 11420-11435, doi:10.1021/acs.jmedchem.0c00242 (2020).
- 21 Bedwell, E. V., McCarthy, W. J., Coyne, A. G. & Abell, C. Development of potent inhibitors by fragment-linking strategies. *Chemical Biology & Drug Design* **100**, 469-486, doi:<https://doi.org/10.1111/cbdd.14120> (2022).
- 22 Wu, X. *et al.* Applications of “linkers” in fragment-based drug design. *Bioorg. Chem.* **127**, 105921, doi:<https://doi.org/10.1016/j.bioorg.2022.105921> (2022).
- 23 Rees, D. C., Congreve, M., Murray, C. W. & Carr, R. Fragment-based lead discovery. *Nature Reviews Drug Discovery* **3**, 660-672, doi:10.1038/nrd1467 (2004).
- 24 Ciulli, A. & Abell, C. Fragment-based approaches to enzyme inhibition. *Current opinion in biotechnology* **18**, 489-496 (2007).
- 25 Murray, C. W. & Blundell, T. L. Structural biology in fragment-based drug design. *Current Opinion in Structural Biology* **20**, 497-507, doi:<https://doi.org/10.1016/j.sbi.2010.04.003> (2010).
- 26 Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W. & Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nature reviews Drug discovery* **15**, 605 (2016).
- 27 Murray, C. W. & Verdonk, M. L. The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *Journal of Computer-Aided Molecular Design* **16**, 741-753, doi:10.1023/A:1022446720849 (2002).

- 28 Ichihara, O., Barker, J., Law, R. J. & Whittaker, M. Compound Design by Fragment-Linking. *Molecular Informatics* **30**, 298-306, doi:<https://doi.org/10.1002/minf.201000174> (2011).
- 29 Jencks, W. P. On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 4046-4050, doi:10.1073/pnas.78.7.4046 (1981).
- 30 Yu, H. S. *et al.* General Theory of Fragment Linking in Molecular Design: Why Fragment Linking Rarely Succeeds and How to Improve Outcomes. *Journal of Chemical Theory and Computation* **17**, 450-462, doi:10.1021/acs.jctc.0c01004 (2021).
- 31 Zhang, J., Yang, P. L. & Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nature Reviews Cancer* **9**, 28-39, doi:10.1038/nrc2559 (2009).
- 32 Attwood, M. M., Fabbro, D., Sokolov, A. V., Knapp, S. & Schiöth, H. B. Trends in kinase drug discovery: targets, indications and inhibitor design. *Nature Reviews Drug Discovery* **20**, 839-861, doi:10.1038/s41573-021-00252-y (2021).
- 33 Müller, S., Chaikuad, A., Gray, N. S. & Knapp, S. The ins and outs of selective kinase inhibitor development. *Nature Chemical Biology* **11**, 818-821, doi:10.1038/nchembio.1938 (2015).
- 34 Wu, P., Clausen, M. H. & Nielsen, T. E. Allosteric small-molecule kinase inhibitors. *Pharmacology & Therapeutics* **156**, 59-68, doi:<https://doi.org/10.1016/j.pharmthera.2015.10.002> (2015).
- 35 Lu, X., Smaill, J. B. & Ding, K. New promise and opportunities for allosteric kinase inhibitors. *Angewandte Chemie International Edition* **59**, 13764-13776 (2020).
- 36 Heppner, D. E. & Eck, M. J. A structural perspective on targeting the RTK/Ras/MAP kinase pathway in cancer. *Protein Sci.* **30**, 1535-1553, doi:<https://doi.org/10.1002/pro.4125> (2021).
- 37 Paez, J. G. *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304**, 1497-1500 (2004).
- 38 van Alderwerelt van Rosenburgh, I. K. *et al.* Biochemical and structural basis for differential inhibitor sensitivity of EGFR with distinct exon 19 mutations. *Nat. Commun.* **13**, 6791, doi:10.1038/s41467-022-34398-z (2022).
- 39 To, C. *et al.* Single and dual targeting of mutant EGFR with an allosteric inhibitor. *Cancer Discov.* **9**, 926-943 (2019).
- 40 To, C. *et al.* An allosteric inhibitor against the therapy-resistant mutant forms of EGFR in non-small cell lung cancer. *Nat. Cancer* **3**, 402-417, doi:10.1038/s43018-022-00351-8 (2022).
- 41 Niggenaber, J. *et al.* Complex Crystal Structures of EGFR with Third-Generation Kinase Inhibitors and Simultaneously Bound Allosteric Ligands. *ACS Med. Chem. Lett.* **11**, 2484-2490 (2020).
- 42 Beyett, T. S. *et al.* Molecular basis for cooperative binding and synergy of ATP-site and allosteric EGFR inhibitors. *Nat. Commun.* **13**, 2530, doi:10.1038/s41467-022-30258-y (2022).
- 43 Li, Q. *et al.* Discovery of Potent and Noncovalent Reversible EGFR Kinase Inhibitors of EGFR L858R/T790M/C797S. *ACS Med. Chem. Lett.* **10**, 869-873, doi:10.1021/acsmchemlett.8b00564 (2019).
- 44 Wittlinger, F. *et al.* Design of a “Two-in-One” Mutant-Selective Epidermal Growth Factor Receptor Inhibitor That Spans the Orthosteric and Allosteric Sites. *J. Med. Chem.* **65**, 1370-1383, doi:10.1021/acs.jmedchem.1c00848 (2022).

- 45 Bühler, S., Goettert, M., Schollmeyer, D., Albrecht, W. & Laufer, S. A. Chiral Sulfoxides as Metabolites of 2-Thioimidazole-Based p38 $\alpha$  Mitogen-Activated Protein Kinase Inhibitors: Enantioselective Synthesis and Biological Evaluation. *J. Med. Chem.* **54**, 3283-3297, doi:10.1021/jm101623p (2011).
- 46 Günther, M., Juchum, M., Kelter, G., Fiebig, H. & Laufer, S. Lung Cancer: EGFR Inhibitors with Low Nanomolar Activity against a Therapy-Resistant L858R/T790M/C797S Mutant. *Angewandte Chemie International Edition* **55**, 10890-10894, doi:<https://doi.org/10.1002/anie.201603736> (2016).
- 47 Günther, M. *et al.* Trisubstituted pyridinylimidazoles as potent inhibitors of the clinically resistant L858R/T790M/C797S EGFR mutant: targeting of both hydrophobic regions and the phosphate binding site. *J. Med. Chem.* **60**, 5613-5637 (2017).
- 48 Juchum, M. *et al.* Trisubstituted imidazoles with a rigidized hinge binding motif act as single digit nM inhibitors of clinically relevant EGFR L858R/T790M and L858R/T790M/C797S mutants: an example of target hopping. *J. Med. Chem.* **60**, 4636-4656 (2017).
- 49 Heppner, D. E., Günther, M., Wittlinger, F., Laufer, S. A. & Eck, M. J. Structural Basis for EGFR Mutant Inhibition by Trisubstituted Imidazole Inhibitors. *J. Med. Chem.* **63**, 4293-4305, doi:10.1021/acs.jmedchem.0c00200 (2020).
- 50 De Clercq, D. J. H. *et al.* Discovery and Optimization of Dibenzodiazepinones as Allosteric Mutant-Selective EGFR Inhibitors. *ACS Med. Chem. Lett.* **10**, 1549-1553, doi:10.1021/acsmchemlett.9b00381 (2019).
- 51 Sun, Z., Liu, Q., Qu, G., Feng, Y. & Reetz, M. T. Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. *Chemical Reviews* **119**, 1626-1665, doi:10.1021/acs.chemrev.8b00290 (2019).
- 52 Jia, Y. *et al.* Overcoming EGFR(T790M) and EGFR(C797S) resistance with mutant-selective allosteric inhibitors. *Nature* **534**, 129-132, doi:10.1038/nature17960 (2016).
- 53 Hua, L. *et al.* Beyond Proteolysis-Targeting Chimeric Molecules: Designing Heterobifunctional Molecules Based on Functional Effectors. *J. Med. Chem.* **65**, 8091-8112, doi:10.1021/acs.jmedchem.2c00316 (2022).
- 54 Dong, G., Ding, Y., He, S. & Sheng, C. Molecular Glues for Targeted Protein Degradation: From Serendipity to Rational Discovery. *J. Med. Chem.* **64**, 10606-10620, doi:10.1021/acs.jmedchem.1c00895 (2021).
- 55 Liu, X. & Ciulli, A. Proximity-Based Modalities for Biology and Medicine. *ACS Central Science*, doi:10.1021/acscentsci.3c00395 (2023).
- 56 Troup, R. I., Fallan, C. & Baud, M. G. Current strategies for the design of PROTAC linkers: a critical review. *Exploration of Targeted Anti-tumor Therapy* **1**, 273 (2020).
- 57 Cyrus, K. *et al.* Impact of linker length on the activity of PROTACs. *Molecular BioSystems* **7**, 359-364 (2011).
- 58 Chung, S., Parker, J. B., Bianchet, M., Amzel, L. M. & Stivers, J. T. Impact of linker strain and flexibility in the design of a fragment-based inhibitor. *Nature Chemical Biology* **5**, 407-413, doi:10.1038/nchembio.163 (2009).
- 59 de Souza Neto, L. R. *et al.* In silico Strategies to Support Fragment-to-Lead Optimization in Drug Discovery. *Frontiers in Chemistry* **8**, doi:10.3389/fchem.2020.00093 (2020).
- 60 Oltersdorf, T. *et al.* An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **435**, 677-681, doi:10.1038/nature03579 (2005).

- 61 Petros, A. M. *et al.* Discovery of a Potent Inhibitor of the Antiapoptotic Protein Bcl-xL from NMR and Parallel Synthesis. *J. Med. Chem.* **49**, 656-663, doi:10.1021/jm0507532 (2006).
- 62 Deeks, E. D. Venetoclax: First Global Approval. *Drugs* **76**, 979-987, doi:10.1007/s40265-016-0596-x (2016).
- 63 Jorgensen, W. L. Efficient Drug Lead Discovery and Optimization. *Accounts of Chemical Research* **42**, 724-733, doi:10.1021/ar800236t (2009).
- 64 Schneider, D. K. *et al.* FMX - the Frontier Microfocusing Macromolecular Crystallography Beamline at the National Synchrotron Light Source II. *Journal of Synchrotron Radiation* **28**, 650-665, doi:10.1107/S1600577520016173 (2021).
- 65 Hong, L., Quinn, C. M. & Jia, Y. Evaluating the utility of the HTRF® Transcreeper™ ADP assay technology: A comparison with the standard HTRF assay technology. *Analytical biochemistry* **391**, 31-38 (2009).
- 66 Zhou, W. *et al.* Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature* **462**, 1070-1074, doi:10.1038/nature08622 (2009).
- 67 Friesner, R. A. *et al.* Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. *Journal of medicinal chemistry* **49**, 6177-6196 (2006).
- 68 Banks, J. L. *et al.* Integrated modeling program, applied chemical theory (IMPACT). *Journal of computational chemistry* **26**, 1752-1780 (2005).
- 69 Lu, C. *et al.* OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. *Journal of Chemical Theory and Computation* **17**, 4291-4300, doi:10.1021/acs.jctc.1c00302 (2021).
- 70 Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design* **27**, 221-234, doi:10.1007/s10822-013-9644-8 (2013).
- 71 Bowers, K. J. *et al.* in *SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*. 43-43.
- 72 Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926-935, doi:10.1063/1.445869 (1983).
- 73 Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **98**, 10089-10092, doi:10.1063/1.464397 (1993).
- 74 Jacobson, M. P. *et al.* A hierarchical approach to all-atom protein loop prediction. *Proteins: Structure, Function, and Bioinformatics* **55**, 351-367, doi:<https://doi.org/10.1002/prot.10613> (2004).
- 75 Jacobson, M. P., Friesner, R. A., Xiang, Z. & Honig, B. On the Role of the Crystal Environment in Determining Protein Side-chain Conformations. *Journal of Molecular Biology* **320**, 597-608, doi:[https://doi.org/10.1016/S0022-2836\(02\)00470-9](https://doi.org/10.1016/S0022-2836(02)00470-9) (2002).
- 76 Abel, R., Young, T., Farid, R., Berne, B. J. & Friesner, R. A. Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **130**, 2817-2831, doi:10.1021/ja0771033 (2008).

- 77 Young, T., Abel, R., Kim, B., Berne, B. J. & Friesner, R. A. Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 808-813, doi:doi:10.1073/pnas.0610202104 (2007).
- 78 Polak, E. & Ribiere, G. Note sur la convergence de méthodes de directions conjuguées. *Revue française d'informatique et de recherche opérationnelle. Série rouge* **3**, 35-43 (1969).
- 79 Waskom, M. L. Seaborn: statistical data visualization. *Journal of Open Source Software* **6**, 3021 (2021).