

Natural Product-like Fragments Unlock Novel Chemotypes for a Kinase Target—Exploring Options beyond the Flatland

Anna Santura, Janis Müller, Madita Wolter, Ina-Charlotte Tutzschky, Moritz Ruf, Alexander Metz, Anna Sandner, Stefan Merkl, Gerhard Klebe, Serghei Glinca,* and Paul Czodrowski*



Cite This: *J. Chem. Inf. Model.* 2026, 66, 2249–2267



Read Online

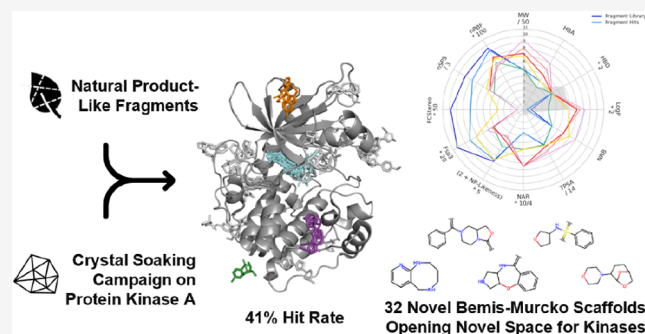
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: In this study, we utilized a high-performance soaking system of protein kinase A (PKA) to perform a crystallographic screening of a natural product-like fragment library. We resolved 36 fragment-bound structures, corresponding to a hit rate of 41%. Nine fragments bound within the ATP site, nine peripherally, and 18 interacted with both the ATP and peripheral sites. One fragment binds to the same site as the approved allosteric kinase inhibitor asciminib, while another induces an unexpected conformational change. Systematic database mining revealed that both the fragments and their natural product parents have not been previously associated with PKA or kinase activity. A scaffold/chemotype analysis further underscored their novelty. Cheminformatics analyses confirmed that these fragments occupy a distinct chemical space, enriched in saturation, spatial complexity and molecular three-dimensional character compared to kinase binders from reference data sets. These properties have previously been linked to increased selectivity, reduced CYP450 inhibition, and higher overall clinical success rates.



INTRODUCTION

Fragment Screening Libraries Benefit from Increased Molecular Three-Dimensionality. The success of fragment-based drug discovery (FBDD) heavily depends on the quality and structural variation, including the coverage of distinct chemotypes, of the fragment library.¹ Although chemical diversity in FBDD libraries has been achieved in many respects, considering a large variety of shapes—particularly in terms of fragments rich in three-dimensional (3D) character—remains a key challenge.^{2–4} In the past, many fragment libraries exhibited a preponderance of sp^2 -hybridized carbons, aromatic and heterocyclic rings, as a result of synthetic methods favoring sp^2 - sp^2 bond formation.⁵ However, the scope is changing and there is growing interest in, and a strong rationale for the inclusion of spatially more 3D building blocks in fragment libraries.^{2,4} Recent efforts have addressed synthetic challenges associated with the synthesis (and elaboration) of such sp^3 -rich fragments, and thus focused predominantly on tailored library design, ready for the screening against biological targets.^{4,6}

Compound shape has long been recognized as the primary factor for successful binding of a ligand and to its biological target(s). Incorporating out-of-plane substituents and starting with spatial core structures can optimize ligand-target complementarity. The spatial orientation of pharmacophoric features is essential for effective target binding and optimizing protein–ligand interactions, thus potentially improving

potency and selectivity while mitigating off-target effects. Indeed, systematic studies by Lovering et al.^{7,8} demonstrated that an increased fraction of sp^3 -hybridized/chiral carbons enhances selectivity and avoids cytochrome P450 inhibition, a key factor to control and design metabolism. Molecules incorporating significant 3D character often display improved aqueous solubility due to less efficient and therefore overall weaker solid-state crystal lattice packing.⁹ Taken together, an enhanced 3D character may contribute positively to both improved pharmacodynamics and pharmacokinetics, and has been correlated to overall success of drug candidates in clinical studies.⁷ At the same time, it can be argued that the increase in molecular complexity may reduce hit rates in screening campaigns.^{10–12} While the cited studies do not apply to FBDD approaches specifically, it is reasonable to propose that 3D fragments could be valuable starting points for drug development.⁴

While the 2009 publication by Lovering et al.⁷ was expected to be a turning point in medicinal chemistry, efforts in the last years mostly addressed the synthetic challenges associated with

While the 2009 publication by Lovering et al.⁷ was expected to be a turning point in medicinal chemistry, efforts in the last years mostly addressed the synthetic challenges associated with

Received: August 26, 2025
Revised: November 23, 2025
Accepted: December 9, 2025
Published: December 22, 2025



sp³-rich fragments. However, from a medicinal chemistry perspective, a reevaluation of the 15-year-old “*Escape from Flatland*” paper did not find a clear relationship between the highest (pre)clinical development phase reached by a compound and its fraction of sp³-hybridized carbons after 2009 anymore.¹³ This apparent disconnect in the last years was attributed to the changing trend in drug target selection, in particular an increase in the number of kinase inhibitors, as well as an increase in the use of metal-catalyzed cross coupling reactions simplifying the incorporation of sp²-hybridized systems into drugs.¹³ As described, however, there is still a strong rationale and promising perspective for the inclusion of sp³-rich fragments in screening libraries, also for kinase targets.

Given the inherent richness of natural products (NPs) in saturation, sp³-hybridized carbons, and stereo centers compared to synthetic molecules,¹⁴ NP-inspired molecules offer a promising strategy for creating shape-diverse fragment libraries.^{15–18} Additionally, NPs typically contain fewer aromatic rings which potentially enhances aqueous solubility, and suffer less from properties falling under the pan-assay interference compounds.^{14,19} In other words, NPs as well as NP-derived fragments occupy areas of the chemical space likely underexplored by molecules obtained by classical organic synthesis.²⁰ A proof-of-concept study by Over et al.¹⁵ demonstrated that NP-derived fragments can give access to novel chemotypes for established drug targets, such as p38 α kinase. Despite this success, NPs have so far received only little attention as a pool for fragment-based screening primarily due to their synthetic intractability and the challenging chemistry required for structural expansion and optimization.^{18,21} More recently, Huschmann et al.¹⁸ highlighted the utility of a commercially available library of NP-derived fragments for initial crystallographic screening, and subsequent hit validation using off-the-shelf follow-up compounds.

Kinases Are in the Focus of Drug Discovery. Protein kinases are among the largest protein families in the human genome, comprising 518 members.²² As approximately 30% of all human protein kinases are associated with disease, kinases are in the focus of the pharmaceutical industry and academic research.²³ Most protein kinases share a very similar 3D structure, the so-called eukaryotic protein kinase fold. Among them, PKA, also known as cAMP-dependent protein kinase, stands out as one of the best-studied and well-characterized protein kinase to date.^{24,25} As PKA is a well-established system for crystallization and qualifies for the “crystal structure first approach”,^{26,27} it can serve as a surrogate to study other members of the protein kinase family. Notably, the structure of the PKA catalytic subunit α was the first kinase structure to be determined.²⁸ As of today, more than 5000 crystal structures comprising the kinase catalytic domain have been deposited in the Protein Data Bank (PDB),²⁹ thereof about 350 structures of PKA, and many more are available in proprietary databases.

As of September 2024, 107 small molecule protein kinase inhibitors (PKIs) targeting around 20 different protein kinases have been approved, primarily for treating neoplasms, and many more are currently in clinical trials worldwide.^{30–32} Most PKIs currently on the market, target the highly conserved and ubiquitously present binding site of the ATP cofactor, making selectivity difficult to achieve as difference between kinases are found only in the back and front pocket or above and below the ATP binding site. Most orthosteric PKIs mimic the cofactor’s adenine ring by forming 1–3 hydrogen bonds with the kinase hinge motif. This conserved interaction pattern is

reflected in most chemotypes of the orthosteric inhibitors, which predominantly correspond to (flat, aromatic) N-heterocycles.³³ The nonselective natural product staurosporine and its derivatives also address the hinge motif and inhibit kinases which underscores that sterically more demanding compounds are amenable to modulate kinase activity.^{33,34} Notably, the introduction of sterically more complex metal organic centers provide the promiscuous staurosporine scaffold with remarkable selectivity features. Despite such systematic studies have shown the superiority of shapely sp³-rich compounds, even today kinase-targeted screening libraries are dominated by flat heteroaromatic compounds. As a consequence, presently the FDA-approved kinase inhibitors all possess a minimum of two aromatic rings and exhibit an assembly of rather planar building blocks.^{30,35}

PKIs binding to so-called allosteric binding sites are generally expected to offer a selectivity advantage compared to orthosteric PKIs, as they mostly bind to sites that are less conserved across the kinome. Additionally, such PKIs are thought to be less prone to the development of resistance. Recent systematic analyses of publicly available kinase ligand complex structures mapped in total 12 binding pockets in the kinase domain (designated A–L), including the ATP site (A) and 11 peripheral binding pockets (B–L), although not all are present in every kinase.^{36,37} A compelling example of a successful allosteric PKIs is asciminib, which specifically targets the E pocket of BCR-ABL1 kinase.³⁸ Approved in 2021 for the treatment of chronic myelogenous leukemia, asciminib is effective against both wild-type BCR-ABL1 and several mutated forms, thus meeting all the promises for allosteric PKIs.³⁸

As with drug candidates for other target classes, enhancing molecular three-dimensionality in general, and introducing carbon bond saturation and stereogenic centers in particular, may help to escape from flatland and to secure new intellectual property of future PKIs.^{7,31} Isolated examples of screening 3D-rich fragments against a kinase target have been published.^{39,40}

Objectives. This FBDD study aimed to address the challenge of enhancing molecular three-dimensionality in general, and introducing carbon bond saturation and stereogenic centers in specific, particularly for kinase targets. By utilizing a set of fragments from the “Fragments from Nature” library, we sought to leverage the inherent spatial complexity of NP-inspired fragments, to identify promising starting points for the development of new PKIs. In detail, we conducted a crystal soaking campaign on protein kinase A (PKA) using NP-like fragments, adhering to the “crystal structure first” paradigm. Our structure-based analyses focused on classifying the kinase conformation, including the conformation of the glycine-rich loop; on the fragment binding sites, with emphasis on the peripheral sites; and the discovery of potential novel kinase binding modalities. We further complemented these efforts with a systematic database mining to contextualize the identified fragment hits within the broader landscape of NPs and kinase ligands. Specifically, we assessed whether the bound fragments indeed represent novel scaffolds, and quantified the degree of spatial complexity shown by the fragments compared to reference molecules.

RESULTS AND DISCUSSION

Fragment Library and Crystal Soaking Campaign. A selection of 87 fragments from AnalytiCon Discovery’s FRGx (“Fragments from Nature”) library⁴¹ was utilized in a crystal

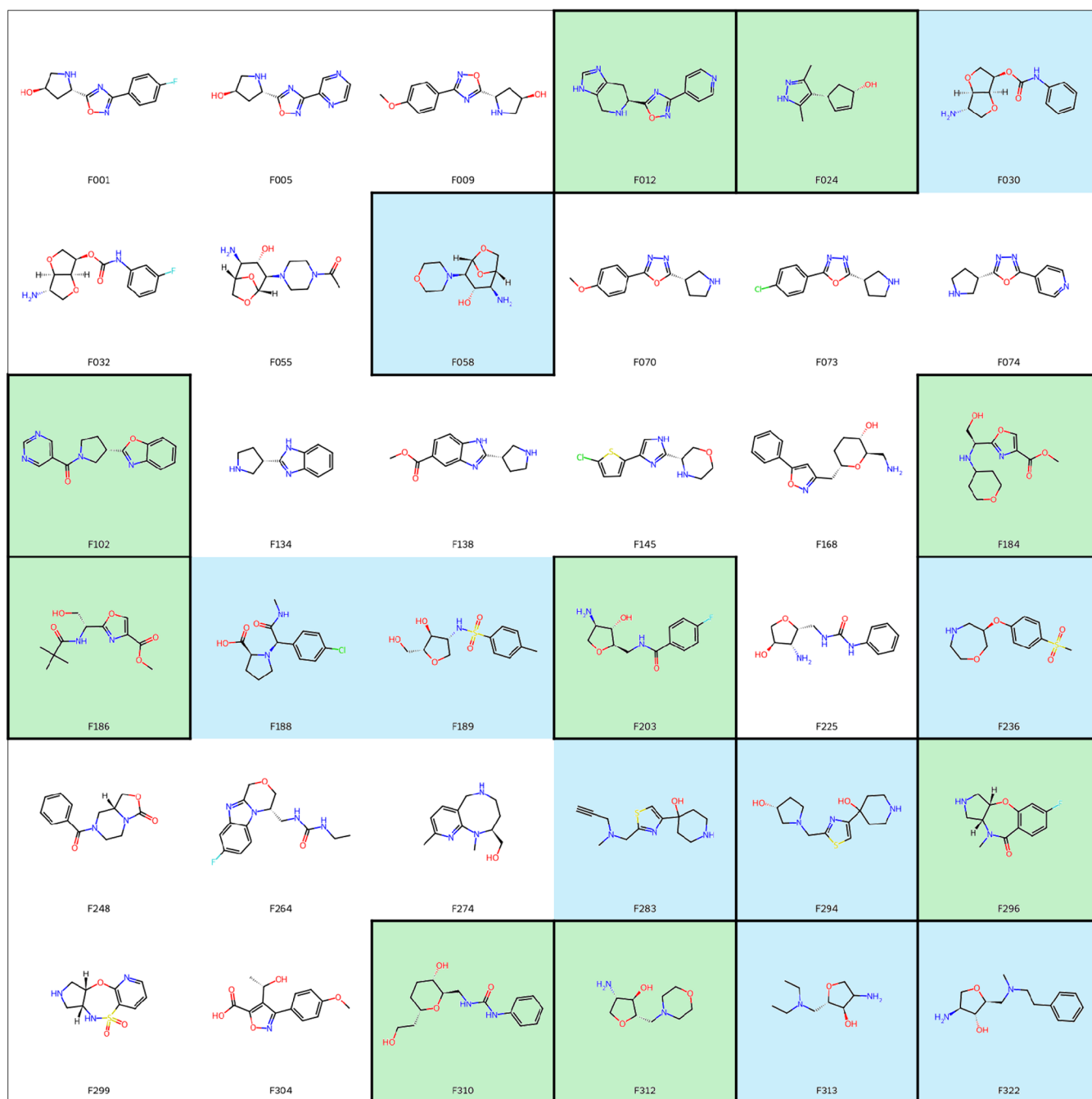


Figure 1. Chemical structures of the 36 fragment hits. Fragments that bind only in the ATP pocket are highlighted in green; fragments that bind only at one or more peripheral site(s) are highlighted in blue; fragments that bind to a single site are indicated by a black frame. All other fragments bind both in the ATP pocket and simultaneously to at least one additional peripheral site.

soaking campaign on PKA. The campaign delivered an exceptionally high hit rate of 41%, resulting in 36 high-resolution complex structures. All binding events were identified in the native electron density map and all structures could be fully refined. All complex structures were resolved at a resolution below 2 Å, with an average resolution of 1.5 Å. In 61% of the complex structures more than one fragment copy binds to the protein. Considering all of these copies, a total of 69 distinct binding poses were determined. The chemical structures of the 36 fragment hits are visualized in Figure 1.

Prior Knowledge on the Fragments and Their (Natural Product) Parents from Database Mining. No

Reported Kinase Activity for the Fragment Hits in Public Databases. First, we analyzed what was already known about the fragments detected as hits in the public domain, querying PubChem and ChEMBL.^{42,43} Although 35 fragments possess a PubChem entry, none of them is associated with any target or bioactivity annotation in this database. In ChEMBL four fragments are annotated, namely F012 (ChEMBL3437495), F074 (ChEMBL3437482), F186 (ChEMBL4937704), and F236 (ChEMBL4916673), whereby bioactivity data and target annotations are only available for the first two. For both, bioactivity data are reported for bacterial and protozoan

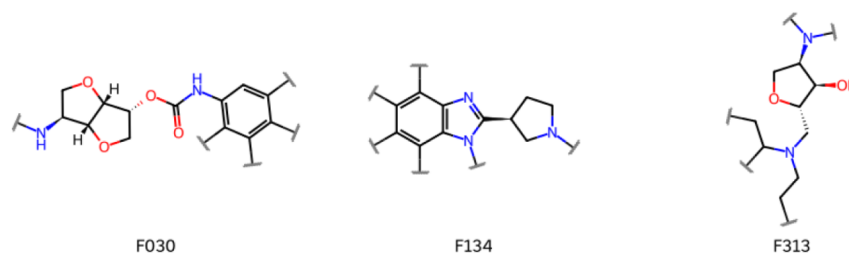


Figure 2. Growth vectors identified in the fragments with more than 100 associated NP parents (F030, F134, and F313, see previous section), based on derivatization patterns of the latter. Growth vectors are indicated by squiggled lines perpendicular to a bond. Note that not all of them may be compatible with the binding mode evident from the complex crystal structure.

protein targets only. Notably, no activity toward any kinase target has yet been reported.

Natural Product Parent Molecules of the Fragment Hits.

As the fragments were deduced from NPs, we investigated what is already known about NPs comprising the fragments as substructures, particularly with respect to their biological activities. A substructure search in the COCONUT database^{44,45} yielded 1675 different NPs. Depending on the individual fragment, the number of associated NPs largely varied from 0 to 554 (Figure S2). More than a hundred NPs are associated with F030, F134, and F313. In a later section, we will pinpoint potential growth vectors in these fragments that may be of possible interest for the follow-up development of the identified fragment hits.

No Reported Protein Kinase Activity for the NP Parents in Public Databases. Of the 1675 NPs found in COCONUT comprising the fragments as substructures, 54 NPs were found to possess target annotations in PubChem. For one of them, activity toward inositol hexakisphosphate kinase 1 (IP6K1) is reported. Although some structural homologies have been reported between kinases from IPK family (PFAM-ID: PF03770⁴⁶) and the eukaryotic protein kinase (ePK) family, which includes PKA, they are only very distantly related with respect to their enzymatic properties.^{47,48} Hence, no assumptions on the NP's bioactivity toward ePK or PKA can be drawn solely based on the reported IP6K1 bioactivity. ChEMBL provides target annotations including bioactivity data with an assay confidence score of 9—indicating direct testing of a single protein target—for 23 parent NPs. Two NPs have been tested for their inhibitory activity against human glycogen synthase kinase-3 beta (GSK3 β) and one NP against P-glycoprotein. However, all three NPs are inactive, with IC₅₀ values above 50 μ M.

Two Parent Molecules from ChEMBL Are Protein Kinase Inhibitors. An analogous substructure search across all ChEMBL molecules, not limited to NPs, identified 175 compounds with bioactivity data, covering 48 distinct human target proteins, including two kinases, GSK3 β and mTOR. Seven compounds, all parent molecules of fragment F030, were tested against GSK3 β (Figure S4): Six of them were inactive (IC₅₀ > 50 μ M), including those two identified in the COCONUT-based approach described above, whereas CHEMBL4751957 displayed an IC₅₀ value of 1.08 μ M.⁴⁹ For mTOR, CHEMBL1514008, a parent molecule of F134, exhibited a potency of 26 μ M. Both molecules, active toward a kinase, are about twice the size of the corresponding fragment hit.

Overall, little is known in the public domain about the bioactivity of the NPs or synthetic molecules that comprise the

discovered fragments as substructures, or the fragments themselves.

Growth Vector Analysis and Fragment Sociability.

Many FBDD projects have been discontinued in the past due to synthetic intractability or the effort of chemistry required for structural expansion and optimization of the fragment hits.²¹ This is also why NPs in particular have received little attention as a pool for fragment-based screening earlier.¹⁸ However, this is expected not to be an issue in our case, as the concept of “fragment sociability”²¹ has already been integrated into the initial fragment design/selection. A fragment is considered as “social” if a significant number of close analogs is commercially available and if robust synthetic methodology allows the elaboration of growth vectors.²¹

In this context, we analyzed the derivatization points/patterns of the NP parents to reveal possible growth vectors in the fragments (Figure 2 and S3). In addition, we will next address the aspect of the commercial availability of possible follow-up compounds.

To quantify “fragment sociability” we mined the NATx library: On average, each fragment has 405 follow-up compounds sharing the same chemotype, 172 with the same scaffold, and 28 containing it as a substructure available from the NATx library. However, when selecting candidate molecules it must be considered that not all modifications may be compatible with the crystallographically determined binding mode in general, or may be favorably contributing to kinase binding/inhibition. Overall, the NATx library alone already provides multiple follow-up compounds with exemplified growth vectors off-the-shelf, hence enabling efficient validation, examination of structure–activity relationships (SAR), and ultimately rapid optimization of the fragment hits (“growing by catalog”).^{18,41,50}

Structure-Based Analyses. Structural Classification of Protein Kinases and Their Ligands. When describing a kinase and its role in ligand binding, two aspects are to be considered, the conformational state of the protein and the type of ligand in the complex.

All 36 PKA fragment complexes exhibit the DFG-in, α -helix-in conformation, which is also the most frequently observed conformation for ligand-bound PKA structures in the PDB (99% of the 239 PKA:ligand complex structures used as reference herein). While the DFG-in (and DFG-out) label only provides a broad description of a more complex kinase conformational landscape, the more sophisticated Modi and Dunbrack scheme distinguishes 8 kinase conformations based on the dihedral angles of the xDFG motif residues. They are named after the region occupied in the Ramachandran plot (A = alpha; B = beta; L = left) and the phenylalanine rotamer (minus, plus, trans).^{51,52} According to this scheme, 6

subclasses of the DFG-in conformation were identified: BLAminus, BLBplus, ABAmminus, BLBminus, BLBtrans, and BLAplus. 35 of the complex structures are in the so-called BLAminus conformation, corresponding to the kinase active form, while one occupies the active-like ABAmminus conformation (F274). ABAmminus structures strongly resemble the active BLAminus state, involving only a peptide flip of the x and D residues of the xDFG motif. In the remaining part, the positions of the activation loop and the α C-helix are very similar.^{51,52} Nevertheless, the ABAmminus conformation corresponds to an inactive kinase state, as the aspartate does not facilitate the positioning of the cofactors Mg²⁺ and ATP.⁵¹ These two conformational classes also correspond to the most common ones found for ligand-bound PKA structures in the PDB, with abundances of 80% and 18%, respectively.

Following the nomenclature introduced by Dar and Shokat⁵³ and extended by Gavrin and Saiah,⁵⁴ all 36 fragments correspond to type I and/or type IV ligands. Most of the fragments bind to multiple binding sites in the protein, most frequently in the ATP binding pocket and one of the peripheral sites. The fact that fragments tend to bind at several individual binding sites, has been frequently observed and is due to their smaller size and fewer options to establish interactions, as well as to the fact that fragments are usually added in higher concentrations to the crystallization conditions (vide infra).²⁷ However, 9 fragments bind exclusively to peripheral sites and 9 are pure ATP site binders.

The fragment binding modes will be discussed in later sections.

Detailed Pocket Analysis. Next, we analyze in greater detail, which binding sites are occupied by our fragments (Figure 3)

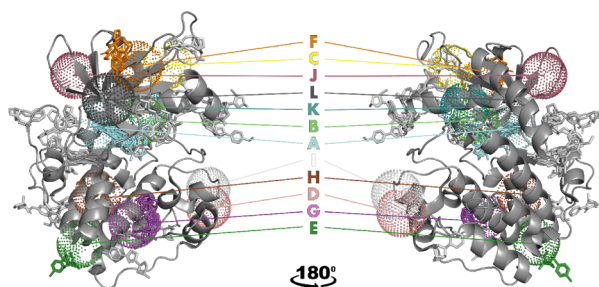


Figure 3. The kinase catalytic domain comprises 12 binding pockets in total (designated A–L), including the ATP site (A), and 11 additional binding pockets (B–L) as described by Laufkötter et al.³⁶ Pockets (in dot representation), defined by a 5 Å radius around the center of mass of all reference ligands (see Materials and Methods), are mapped on a PKA structure (PDB-ID: 3FJQ, cartoon representation). Note that according to Laufkötter et al.³⁶ no individual kinase contains all the pockets. Fragments are colored the same way as the pockets they occupy (no element-specific coloration). Fragments that could not be assigned to any of the known kinase pockets (52.2% of all fragment copies) are displayed in light gray. For 83.3% of the unique fragment hits, at least one copy has been assigned.

and other known PKA ligands, respectively (Figure S5). PDB ligands were found to occupy mostly the ATP pocket A, but also the peripheral pockets B, F, G and K, defined by Laufkötter et al.³⁶ Those bound fragments, which were successfully assigned to the Laufkötter pockets, bind to almost the same set of sites as the PDB ligands, namely A, E, F and G.

F189 Binds to Peripheral Site E. Site E is located in the C-terminal lobe, distal to the active site. To date, in no PDB structure of PKA a ligand is found to bind to this site, while it is occupied by our F189 (Figure S6). Site E has been reported to accommodate type IV inhibitors as well as activators, with the most prominent example being the approved PKI asciminib, targeting the BCR-ABL1 tyrosine kinase domain (Figure S6).^{38,55} The C-terminal α I-helix of tyrosine kinases such as ABL1 exhibits significant conformational flexibility and borders the allosteric pocket E in its kinked conformation,⁵⁶ which is not the case for PKA. Ligands binding to allosteric site E in kinases with a straight α I-helix, which does not enclose the pocket, have also been identified. However, only those that induce α I-helix bending function as allosteric inhibitors (in ABL1).⁵⁶ Asciminib reinforces the pharmacological relevance of allosteric kinase ligands and is an example of the general promise that allosteric inhibitors are likely to be more selective and resistant to mutation than ATP site binders as they target sites with low conservation throughout the kinome. Whether our fragment F189 can bind to the equivalent site in a pharmacologically relevant target such as ABL1 kinase and trigger the conformational change required for inhibition will be addressed in a subsequent study.

Other Peripheral Sites Occupied by Our Fragment Hits. The other fragments assigned to a Laufkötter peripheral site are F134 and F264 (both binding to site F, the ATP site and at least one additional peripheral site), as well as F030, F138 and F283 (all binding to peripheral site G and at least one other site).

Crystal Contact Analysis for Peripheral Binders. For 83.3% of the fragment hits, at least one copy has been assigned to at least one of the Laufkötter kinase pockets. No assignment to any of the Laufkötter pockets could be found for F058, F188, F236, F294, F313, and F322, all of them binding peripherally. Since the bioactivity of the fragments was not the scope of our study, it is possible that some of them may represent (surface-bound) fragments without functional relevance, rather than being true modulators of kinase activity. Moreover, the assignment may be hindered by imperfect superimposition of the PKA and the reference structures that originate from various kinases and different kinase families. Also, a general caveat of the crystallographic fragment screening approach should be borne in mind: Contact sites in the crystal packing are often detected as pockets or at spots and accommodate fragments. Such hits are typically specific to the crystal environment and will likely not bind to the protein in solution. In our study, this accounts for only a minor portion of the obtained hits which means 14 out of 69 (20.3%) of all fragment copies bind in proximity to more than one PKA molecule in the crystal lattice. These include the individual copies of F058 and F294. In contrast, the copies of F188, F236, F313, and F322 each bind in proximity to a single PKA molecule only. Even for F058 and F294, however, binding is unlikely to be dominated by crystal-packing contacts since >60% of their surface area is in close contact to the asymmetric unit under consideration than to that of the symmetry related crystal mate, and they form more polar interactions with the asymmetric unit than with the symmetry mate. We are therefore confident that none of our fragment hits are detected as hits only due to the crystal packing environment, while we admit that this may be the case for some of the additional copies of e.g. ATP site binders. Even if we exclude the fragments for which no copy could be assigned to one of the

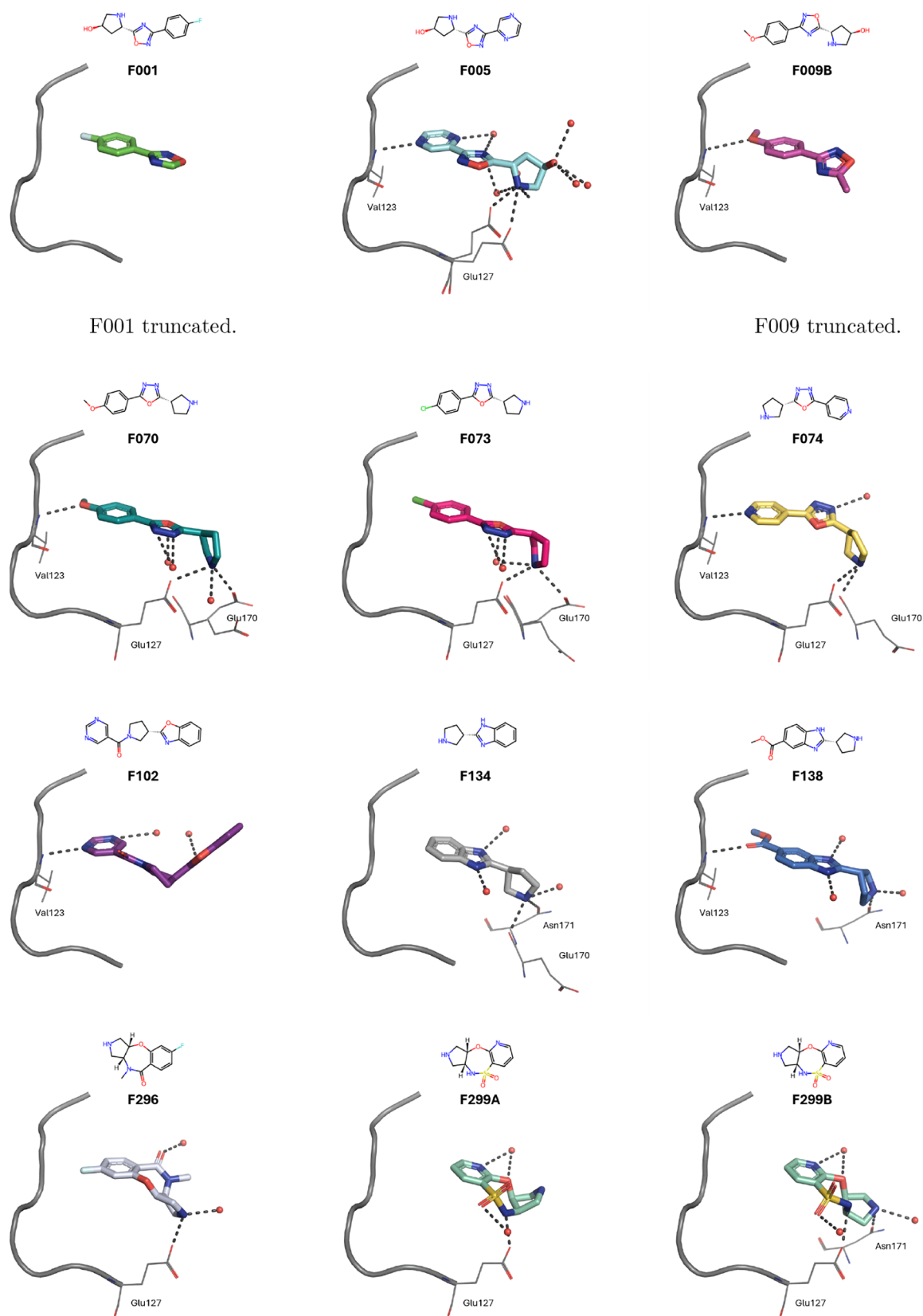


Figure 4. Binding modes of selected fragments within PKA's ATP pocket. If multiple conformations of the fragments were resolved, we show each individually (A/B). Fragments are shown as sticks, amino acid residues involved in polar interactions as lines, waters involved in polar interactions as red spheres, and polar interactions as dashed lines. Additionally, the kinase hinge and linker (residues 120–127) are shown in cartoon representation. For clarity, only amino acid residues involved in polar interactions are labeled. The assigned molecules were truncated to the regions with well-defined electron density at $\sigma = 1$ to minimize model bias. More fragment binding mode depictions can be found in Figure S1.

known Laufkötter pockets hereinafter, the conclusions of the descriptor analysis (*vide infra*) remain unchanged.

Whether our fragments or derived lead compounds also bind and modulate a kinase in solution will be examined in a subsequent study. To assess the biological relevance of fragments that bind only at one or more peripheral site(s), in solution binding should be confirmed by an orthogonal biophysical assay, accompanied by functional biochemical inhibition assays and ATP-competition experiments to check for orthosteric versus allosteric behavior. Despite this, these ligands may still be important even if they do not exhibit an allosteric inhibition mechanism, particularly with regard to the development of degraders, e.g., proteolysis-targeting chimeras (PROTACs)—bifunctional molecules that induce selective proteasomal degradation of disease-relevant target proteins by recruiting them to E3 ubiquitin ligases.

Crystal Contact Analysis for F189 Binding to Peripheral Site E. Below, we will focus on describing the interaction of F189 with the two neighboring PKA molecules in the crystal lattice (Figure S7) in order to verify its specific binding. The interactions of F189 within the asymmetric unit involve seven hydrogen bonds—to the residues His260, Glu140, and Arg144, as well as to three water molecules. In contrast, the interactions of F189 with the protein of a crystal mate include only a hydrogen bond to a water molecule in the first hydration shell and a hydrophobic interaction of the phenyl ring with the hydrophobic pocket. Concerning the fragment's surface area, however, both PKA macromolecules contribute almost equally to the binding site, with 54.1% and 45.4%. This suggests that crystal packing effects may contribute to the observed binding mode, but the hydrogen bonding network indicates that the fragment is recognizing complementary structural features rather than just binding to the crystal-packing interface. Nevertheless, unequivocal proof remains pending, which is why we would like to present this binding mode primarily as interesting and worthy of further experimental investigation.

Fragment Binding Modes. Figures 4 and S1 display the binding modes of all ATP site binders. Nearly half of them are hinge binders, forming one or two hydrogen bonds to the protein backbone of the hinge residues (Glu121, Tyr122, Val123), and/or featuring an ionic interaction with the side chain of Glu127, which is part of the linker region following the hinge. Furthermore, many of the ATP site binding fragments form a hydrogen bond and/or ionic interaction with the Thr183 and the Asp184 residues of the conserved xDFG motif; and/or form one or two hydrogen bonds/salt bridges with residues from the catalytic loop (Glu170, Asn171); and/or the glycine-rich loop (Leu49, Thr51, Phe54).

Some of the fragment hits possess the same scaffold/chemotype. Their binding modes will be compared in more detail in the following. Binding modes of fragments sharing their chemotypes with PKA ligands found in the PDB are described in the Supporting Information.

F001, F005 and F009 belong to the same chemotype, so do F070, F073, and F074. All of which bind within the ATP binding site, with their 1,2,4-oxadiazole or 1,3,4-oxadiazole ring in roughly the same position, and a (substituted) 6-membered aromatic ring pointing toward the hinge region. The latter is involved in hydrogen/halogen bonding with the hinge residues in all cases, except for F001, which orients its polar fluorine atom toward the hinge residues. On the contrary, the oxadiazole does not form a hydrogen bond to any amino acid residue, which presumably explains why its orientation is

not spatially conserved. F134 and F138 share the same scaffold and bind in the same orientation within the ATP binding site. However, their exact binding location is slightly shifted, as the additional methyl ester group of F138 occupies the space in the proximity of the hinge residues. Its carbonyl oxygen is involved in a hydrogen bond to Val123. The positively charged pyrrolidine nitrogen of F134 forms a salt bridge to Glu170 as well as a hydrogen bond to Asn171, while only the latter is found for F138. Although F102 exhibits the same chemotype as F134 and F138, its orientation in the ATP binding site is rotated, so that the nitrogen of the additional pyrimidine ring builds a hydrogen bond with hinge residue Val123. F296 and the structurally similar F299 are also laterally displaced within the ATP binding pocket and form salt bridges with the aforementioned acidic residues. F203, F312, F313 and F322 all represent a 2-substituted (2*S*,3*R*,4*S*)-4-amino-3-hydroxy-tetrahydrofuran. While F203 and F312 bind within the ATP pocket in proximity to the hinge residues (Figure S1), and in the case of F312 forming a hydrogen bond with Val123, the other two fragments bind at the other end of the pocket, in the phosphate subpocket and beyond. Beyond the ATP site binders, F030 and F032 feature a conserved binding mode within peripheral pocket G (*cf.* Detailed Pocket Analysis). On the contrary, the binding sites/modes of three structurally similar fragment sets (F055 + F058, F184 + F186, F283 + F294) are not conserved. While the binding modes are preserved for most of the structurally similar fragments, this is not the case for all of them. This observation exemplifies why it is generally advantageous to incorporate also structurally similar molecules into a screening library.

Unexpected Binding Modality of F274. As already discussed above, the complex structure with F274 occupies the active-like AB_Aminus conformation characterized by a flip of the peptide bond connecting the x (Thr183) and D (Asp184) residues of the xDFG motif relative to the active BL_Aminus kinase conformation. Consequently, the side chains of both residues also occupy alternative conformations, enabling the formation of a hydrogen bond between the Asp184 side chain and the ligand's hydroxyl group (Figure 5). With an occupancy of 0.4, the typical xDFG conformation of the kinase in the active state can also be observed in the electron density. As the structural adaptation is minor, it is

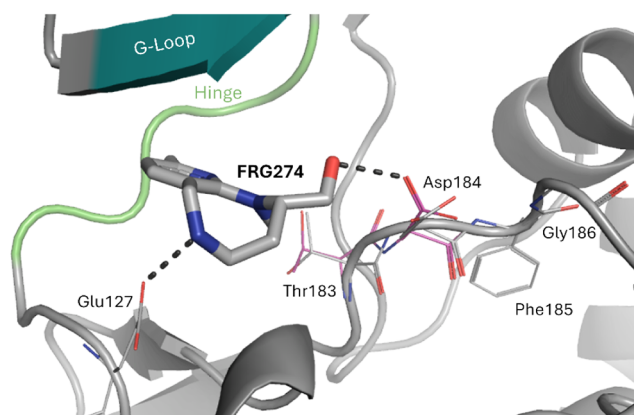


Figure 5. Unexpected binding modality of F274. The xDFG residues in the AB_Aminus conformation are displayed in color, while their alternative conformation is shown in gray. Hydrogen bonds are visualized as black dashed lines. For better clarity, water molecules are not displayed and the view differs from the common one.

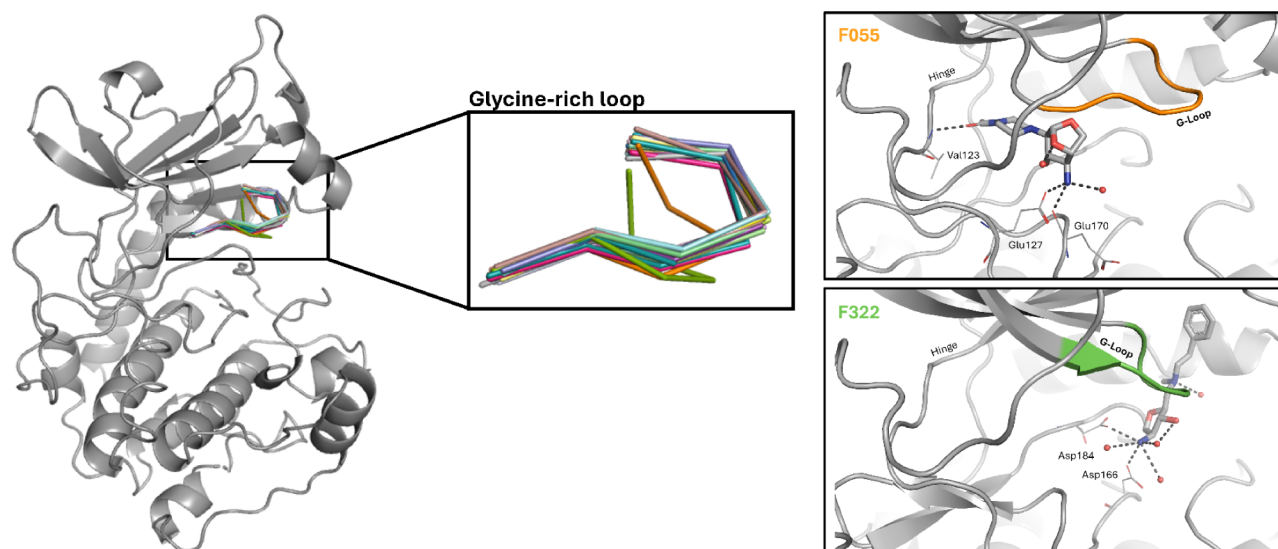


Figure 6. Superimposition of the glycine-rich loop (amino acid residues ⁵⁰GTGSFG⁵⁵, ribbon representation) of all our crystal structures. Note that in case of the complexes with F134, F138 and F274 not all G-loop residues have been modeled due to missing electron density support. The boxes on the right show the complex structures for which the G-loop conformation stood out, as it is stacked down onto the ligand. Therein, the G-loop residues are highlighted in the same color as in the left part of the figure. The protein is shown in cartoon representation and fragments are shown as sticks. Additionally, we show amino acid residues involved in polar interactions as lines, water molecules involved in polar interactions as red spheres, and polar interactions as black dashed lines.

plausible that this change could even be induced in a soaking experiment, where ligands are added to preformed apo protein crystals, which typically occupy the BLAminus conformation, as seen in PDB structures 4WIH and 6EH0.⁵¹

It should be noted that Modi and Dunbrack⁵¹ have demonstrated with the help of the electron density score for individual atoms (EDIA)⁵⁷ that in many cases ABAMinus structures are in fact mismodelled BLAminus conformations. An analogous examination proves that this is not the case for our F274 complex structure (xDFG motif residues' EDIA_m scores ≥ 0.78 , i.e. well-covered by the electron density).

Conformational Analysis: Glycine-Rich Loop (G-Loop). Although known to undergo significant movement, the kinase G-loop conformation and its implication for ligand binding are only scarcely explored in kinase research.^{58,59} The G-loop is located above the ATP binding site and contains the highly conserved GxGxxG motif (⁵⁰GTGSFG⁵⁵ in PKA α). The latter is also known as the nucleotide-binding motif in accordance with its native function. The flexibility of the G-loop is an essential but seemingly an unpredictable determinant of kinase inhibitor binding strengths.⁶⁰ For instance, as described by Möbitz,⁵⁹ a stacking down of the loop onto the ligand creates a more buried cavity that may contribute to tighter binding and thus higher ligand efficiency.

A qualitative analysis of the G-loop conformations occupied in our PKA fragment complexes shows that two conformations stand out in particular, the one occupied with F055 and that with F322 (Figure 6), in which the G-loop is stacked down onto the ligand, although neither interacts directly with the G-loop residues. On the contrary, F236 forms hydrogen bonds to the backbone of the G-loop residues Thr51, Ser53, and Phe54, while F186 interacts with the preceding and following residues Leu49 and Val57 (Figure S1). In both complex structures, the common G-loop conformation is occupied. Owing to its high flexibility, the G-loop may not be fully resolved in X-ray crystallographic structures, as it is the case for the complexes with F134, F138, and F274. In other words, binding of these

three fragments does not appear to stabilize a distinct G-loop conformation, as it appears for the other complex structures, possibly due to lower affinity or shorter residence time. Overall, the crystallographic data shown alone cannot provide any information about a putative coupling between the G-loop conformation and ligand binding. We plan to investigate this further in follow-up studies, using advanced methodologies such as molecular dynamics simulations.

Chemical Space Analysis. Identification of Novel Scaffolds/Chemotypes. To assess the novelty of our fragment hits beyond the systematic database mining and the NP-centric analysis we compared their scaffolds across five reference data sets. In brief, these reference data sets correspond to two collections of compounds that bind PKA, extracted from the PDB and the ChEMBL database, respectively. Furthermore, we analyzed the data with respect to: A collection of compounds bioactive toward any typical kinase, as found in the ChEMBL database. A data set encompassing the 107 PKIs currently on the market and a data set of all approved oral drugs regardless of their biological target.

The Fragment Hits Incorporate Novel Scaffolds for Kinase Ligands. The 36 fragments represent 32 different Bemis–Murcko scaffolds, none of which have been previously reported in the reference data sets described above (Figure S8). Among the 36 fragments, 21 different cyclic skeletons are included, of which two are not found in any of the reference data sets (Figure 7), and 13 are absent from the PKA-specific reference data sets (Figure S9).

Inspection of the Most Similar Reference Molecules per Fragment. The Bemis–Murcko scaffolds are by definition sensitive to minor structural modifications in the core of the molecule. Hence, we additionally inspected the three in terms of the Morgan fingerprint Tanimoto coefficient top scored most similar reference molecules per fragment, irrespective from which reference data set they originated (Figures 8 and S10). The majority of the most similar reference molecules are contained in the ChEMBL kinases data set, which is by far the

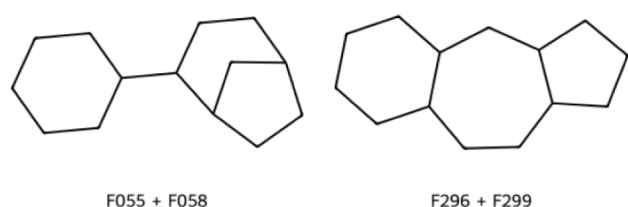


Figure 7. Cyclic skeletons of our fragments not found in any of the reference data sets.

largest of our reference data sets. A few are PDB ligands, approved oral drugs or approved PKIs. The overall highest Tanimoto similarity of 0.50 was determined for F304 and mofezolac, an approved oral nonsteroidal anti-inflammatory drug that selectively inhibits cyclooxygenase (COX1). F304 and putative follow-up compounds may also inhibit COX1 and thus may cause undesirable side effects. F012 and the PDB ligand T7W (PDB-ID: 7BB0) exhibit the second highest Tanimoto similarity of 0.46. F012 is a substructure of T7W, but shows (*S*)-configuration, while T7W is the (*R*)-enantiomer. Three more PDB ligands appear in the list of most similar reference molecules (Figure 8), namely 8KB (PDB-ID: 5N3G), 8PQ (PDB-ID: 5N7U), and 7QI (PDB-ID: 7PID). 8KB represents a substructure of F236 and is thus even smaller than our fragment. 8PQ is an analog of F024 with a bromine substituent in the 4-position of the pyrazole ring, compared to the larger cyclopent-2-en-1-ol moiety in F024. 7QI and F145 share only a small proportion of their structure, namely the morpholine ring (linked to a biaryl). The binding modes of the fragment-PDB ligand pairs of the same chemotype are compared in the Supporting Information (Figure S11). For approved PKIs, the highest Tanimoto similarity to a fragment amounts to 0.29, albeit largely differing in the size of the central ring, and in the type of N-heterocycle incorporated.

Although we have seen a few similar molecules in the reference data sets, this was only the case for a minority of the binders. Overall, the Tanimoto similarities are low, with a mean value of 0.31 for the most similar reference molecule per fragment. This supports our argument that (most of) our fragment hits represent novel chemotypes for kinase targets. Altogether, the scaffold/chemotype analyses again highlight

the novelty of our fragments, demonstrating their potential for further exploration, while generally avoiding conflicts in the already tightly occupied intellectual property space of kinase modulators, and enabling new insights into structural biology not observed before.

Delineating the Fragment's Chemical Space through Molecular Descriptors. To quantify the NP similarity of the fragments, their 3D character and physicochemical properties relevant to medicinal chemistry, and thus to delineate the occupied chemical space in comparison to the reference molecules, we analyzed various molecular descriptors (Figure 9).

Fundamental Physicochemical Properties. Unsurprisingly, our fragments possess lower values in the lower molecular weight (MW), number of rotatable bonds (NRB), number of aromatic rings (NAR), lipophilicity (logP), and topological polar surface area (TPSA) compared to the molecules from the reference data sets described earlier. The average number of hydrogen bond donor (HBD) atoms is identical with a value of 2 across all data sets, including our fragments, despite their smaller molecular size. The average number of hydrogen bond acceptor (HBA) atoms spreads out over a larger range, whereby our fragments fall into the lower range compared to the reference data sets. Approved PKIs stand out, as they possess the highest median value of all seven aforementioned physicochemical descriptors, higher than that of the oral drugs. This is in concordance with the previous assessments in the literature^{31,35} and is presented here for bookkeeping purposes only.

Natural Product Likeness. For our fragment hits, the NP likeness⁶² ranges from -1.82 to $+0.88$, with a median of -0.46 . Lower median values are observed for the kinase-targeted reference data sets indicating that our fragments exhibit greater similarity to NPs than the reference molecules (all *p*-values ≤ 0.03 , all Cohen's d^{63} effect sizes ≥ 0.4 , Figure S12). In contrast, the NP likeness of the fragments is lower than that of approved oral drugs (d -value = -0.25), but higher compared to the NPs comprising the fragments as substructures (median = -0.97) (Figure S12). While the larger, more complex NPs include additional structural elements that may contribute to an increase in the NP likeness score, this is offset by the larger size of the NPs. In other words, the fragments exhibit a

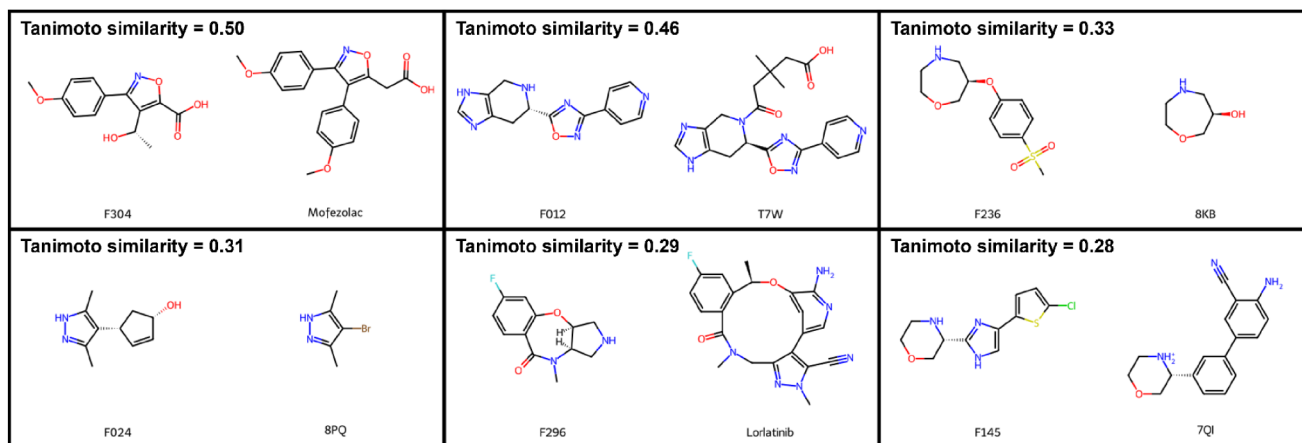


Figure 8. 2D Structures of the in terms of the Tanimoto coefficient most similar reference ligands for the reference data set of oral drugs, PDB ligands and approved PKIs, along with the related fragments. These also include the pairs with the overall highest Tanimoto similarity. Tanimoto similarity was calculated from Morgan fingerprints (radius = 2, 4096 bits).

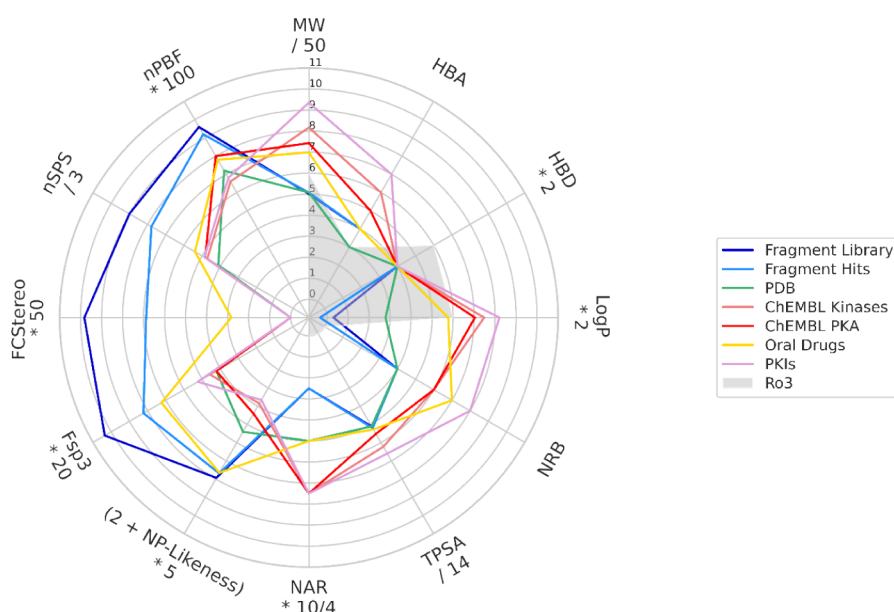


Figure 9. Median values of molecular descriptors. The descriptor space of the Rule of Three (Ro3)⁶¹ is indicated as gray shade. MW = Molecular weight. HBA and HBD = Number of hydrogen acceptor and donor atoms, respectively. LogP = calculated logarithm of the n-octanol/water partition coefficient. NRB = Number of rotatable bonds. TPSA = Topological polar surface area. NAR = Number of aromatic rings. NP likeness = Natural product likeness. Fsp³ = Fraction of sp³-hybridized carbons. FC_{Stereo} = Fraction of stereogenic carbons. nSPS = Normalized spatial score. nPBF = Normalized deviation from the plane of best fit, calculated from in silico-generated ligand conformations. For detailed descriptions of the molecular descriptors refer to the eponymous Methods section. Exact median values (Table S1) descriptor value distributions, and a tabular overview of the compounds/fragments associated with the highest/lowest descriptor value per data set in the Supporting Information (Figure S12 and Tables S1 and S2).

high(er) NP likeness score, as their structures are (more) densely packed with typical NP substructures than the parent NPs.

Saturation and Spatial Complexity. Looking at the fraction of sp³-hybridized carbons (Fsp³), the fraction of stereogenic carbons (FC_{Stereo}), and the normalized spatial score (nSPS) descriptors, our fragments possess a higher median value compared to the reference molecules. Dunn's statistical test verified that the descriptor value distributions for the fragments differ significantly from those in the reference data sets with p-values ≤ 0.03 and Cohen's d-values ≥ 0.3 . For the kinase-targeted reference data sets, all p-values are $\leq 10^{-7}$ and d-values are ≥ 1.25 , indicating a large effect size (Figure S12). In line with the findings by Lovering et al.⁷ who reported that the Fsp³ measure correlates with overall clinical success, the data sets of approved PKIs and oral drugs feature higher Fsp³ median values compared to the data sets from ChEMBL and the PDB, which are dominated by early stage discovery compounds (all p-values ≤ 0.025 , d-values ≥ 0.3 , data not shown).

An outstanding proportion of 55.6% of our fragments surpasses the threshold value of Fsp³ ≥ 0.42 reported in the literature,⁶⁴ while the percentages of compliant molecules in the kinase-targeted reference data sets amount to 16.2% at most (Figure S13). The percentage of molecules surpassing the Fsp³ threshold of the data set of oral drugs on the market is quite high (44.0%), again reflecting the finding by Lovering et al.⁷ On the contrary, it amounts to only 15.9% for approved PKIs, mirroring the fact that most of them contain multiple aromatic rings which do not contribute to Fsp³, also indicated by the higher median number of aromatic rings in PKIs versus all approved oral drugs.

3D Descriptors Depend on the Underlying Ligand Conformation. While all previously mentioned molecular descriptors were calculated from the molecule's 2D structure, the deviation from the plane of best fit (PBF) descriptor was computed from the 3D structure, which was either generated in silico using RDKit's ETKDG method^{65,66} or obtained from crystallographic data representing the protein-bound state. This two-legged approach was chosen to facilitate comparisons with reference data sets lacking 3D structural information, such as those derived from ChEMBL, while also leveraging available experimental structural data where possible. As the molecules differ largely in size across and within the data sets, the deviation from the plane of best fit descriptor has been normalized by the number of non-hydrogen atoms (nPBF).

Comparing Ligand Conformations and Derived Descriptor Values. One frequently criticized aspect in the field is the lack of attention to how molecular conformers are generated for descriptor calculation. To validate our approach, we compared (i) the in silico-generated and crystallographically determined conformations by means of the non-hydrogen atom root-mean-square deviation (RMSD) value, and (ii) the difference in the nPBF value calculated from both conformations. The median RMSD value amounted to 1.0 Å for the fragments and 1.5 Å for the PDB ligands. 95% of all fragment copies and 78% of all PDB ligand copies showed an RMSD ≤ 2.0 Å. This demonstrates the convincing relevance of the RDKit's ETKDG-generated conformers,^{65–67} in particular for fragment-sized molecules. The largest difference in the nPBF value [Δ nPBF = nPBF(protein-bound) – nPBF(in silico)] was observed for PDB ligand GGB (5N3K), with Δ nPBF = -0.051 and RMSD = 0.4 Å. GGB is one of eight ligands for which no fully resolved copy was present in any crystal structure (see Materials and Methods). The second-

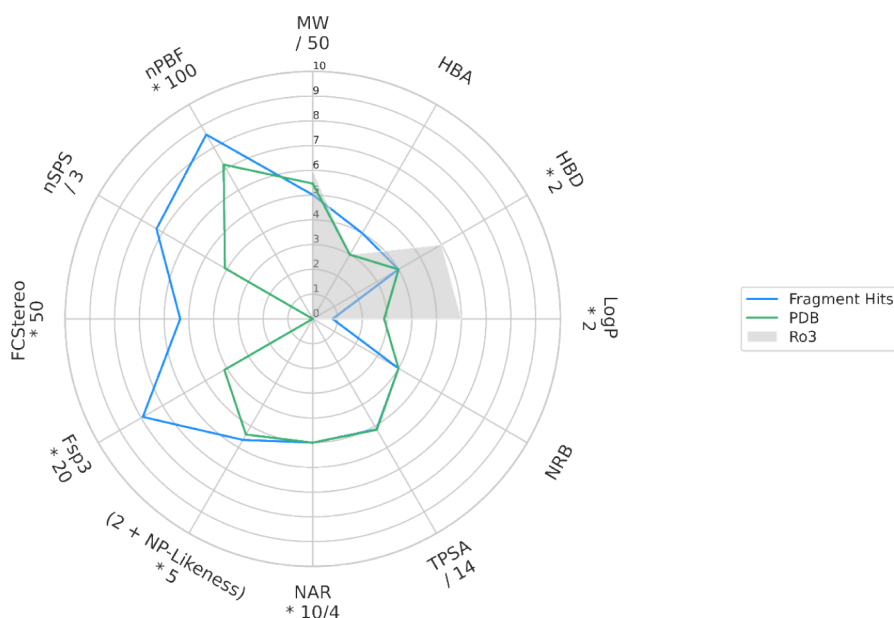


Figure 10. Median values of molecular descriptors for the subset of ATP site binders. The descriptor space of the Rule of Three (Ro3)⁶¹ is indicated as gray shade. MW = Molecular weight. HBA and HBD = Number of hydrogen acceptor and donor atoms, respectively. LogP = calculated logarithm of the n-octanol/water partition coefficient. NRB = Number of rotatable bonds. TPSA = Topological polar surface area. NAR = Number of aromatic rings. NP likeness = Natural product likeness. Fsp³ = Fraction of sp³-hybridized carbons. FC_{Stereo} = Fraction of stereogenic carbons. nSPS = Normalized spatial score. nPBF = Normalized deviation from the plane of best fit, calculated from in silico-generated ligand conformations. For detailed descriptions of the molecular descriptors refer to the eponymous Methods section. Exact median values can be found in Table S1. In addition to the median values, we present the descriptor value distributions in Figure S12.

largest difference, however, was found for the fully resolved PDB ligand 46L (5N3E), with Δ nPBF = -0.039 and RMSD = 0.6 \AA . Since only the largest Δ nPBF involves a truncated ligand and such cases make up a minuscule fraction of the data set, their inclusion is not expected to bias the overall outcome, especially given that our analysis is mainly based on medians. We therefore retained these ligands to keep the data sets invariant and reflect real-world conditions in structure-based drug discovery, where partially resolved ligands are frequently encountered. Among the fragments, the largest difference was obtained for fully resolved F012 (Δ nPBF = -0.028 , RMSD = 0.7 \AA). Importantly, the nPBF values were not systematically higher or lower for the protein-bound versus the in silico-generated ligand conformations. Furthermore, no correlation was observed between RMSD and Δ nPBF values, as confirmed by Pearson and Spearman correlation analyses performed across both data sets (data not shown).

Quantifying Molecular 3D Character. With a median of 0.046 , the nPBF value computed from the in silico-generated conformations is higher for the fragments compared to all reference data sets (Figures 9 and S12). Dunn's statistical test confirmed that the descriptor value distribution for the fragments differs significantly from those of the reference data sets, with p-values ≤ 0.02 . Cohen's d amounted to ≥ 0.25 in general and to 0.58 for the fragment hits versus the PDB data set. Comparison with the kinase-targeted reference data sets only yields p-values $\leq 8 \times 10^{-4}$ and d-values ≥ 0.58 (Figure S12). Calculation of the nPBF descriptor from the protein-bound ligand conformations was only applicable to the fragments and the PDB reference data set. Like the nPBF values computed from the in silico-generated conformers, the distribution is shifted toward higher values for the fragments compared to PDB ligands (medians = 0.037 and 0.032 ; p-value = 5×10^{-4} , d-value = 0.636). Altogether, leveraging both

crystallographically determined and in silico-generated ligand conformations enhances the reliability of our nPBF calculations, offering a more comprehensive view of molecular shape across data sets. We observe consistent trends for both approaches, with the fragments exhibiting higher nPBF values than the reference data set(s) on average.

Comparing the Fragment Hits versus the Screening Library. Comparison of the Fsp³, FC_{Stereo}, nSPS, and nPBF descriptor distributions between the 36 fragment hits and the 87-membered fragment library screened reveals that the hits span almost the entire range of descriptor values as found for all library members, but are concentrated toward the lower end of the distributions.

The Same Trends Are Found for the Subset of ATP Site Binders. So far, the descriptor analyses have been binding site agnostic. However, for the fragments as well as for the PDB reference data set, the binding sites have been classified earlier (section Detailed Pocket Analysis). While allosteric inhibitors such as asciminib are gaining attention in kinase research, most PKIs continue to target the conserved ATP site. Ligands binding to this site typically comprise rather planar (sub)-structures, including at least two aromatic rings.^{30,35} To explore whether the trends observed across the full data set also apply to this key subset, we carried out a focused analysis of the ATP site ligands, encompassing 27 unique fragments and 151 distinct PDB ligands.

Indeed, we found that the ATP site-occupying fragments show the same statistically significant, favorable trends in their property distributions (Figures 10, S12 and S13, Table S1). Also, the percentage of molecules exceeding the threshold of Fsp³ ≥ 0.42 is much higher for the ATP site fragments (40.7%) compared to the PDB ligands occupying the same site (13.9%).

Populating Less Frequently Sampled Regions of Chemical Space. Taken together, the descriptor analysis revealed that our fragment hits exhibit above-average molecular three-dimensionality and spatial complexity, reduced aromaticity, as well as a higher degree of saturation compared to known kinase binders. This also holds true when looking at the subset of ATP site binders alone. Although the descriptor distributions for our fragments are statistically distinct from those of the reference sets, in some cases even with a large effect size (Cohen's $d \geq 0.8$),⁶³ there is considerable overlap. In other words, even within the reference sets, some molecules also exhibit high values of these descriptors. Thus, our fragment hits rather expand than redefine the accessible chemical space for kinase binders by populating less frequently sampled regions. Nevertheless, the different distribution patterns highlight that our library of natural product-like fragments is tailored toward improved three-dimensional properties, while preserving a high hit rate.

Looking at Individual Descriptor Values Obtained for Our Fragment Hits. Finally, we take a closer look at the individual descriptor values obtained for our fragment hits (Table 1). In

Table 1. Exemplary selection of fragments, along with their descriptor values and Ro3 compliance.

	F024	F055	F264	F312	F313
ATP Site	✓	✓	✓	✓	✗
Ro3	✓	✓	✓	✓	✓
Fsp ³	0.50	0.92	0.43	1.00	1.00
FC _{Stereo}	0.20	0.42	0.07	0.33	0.33
nSPS	27	44	18	40	34
nPBF ^a	0.060	0.049	0.052	0.073	0.067
nPBF ^b	0.052	0.048	0.055	0.061	0.057
RMSD ^c [Å]	0.6	1.7	2.1	1.4	1.3

^a Computed from the *in silico*-generated ligand conformation.

^b Mean value across all crystallographically determined conformations.

^c Mean structural deviation between the fragment's *in silico*-generated conformer and all its observed conformations in the protein-bound state.

the case of Fsp³, FC_{Stereo}, and nSPS the five highest descriptor values are associated with the same set of fragments (F055, F058, F312, F313), albeit in diverging order (Table S2). An exception to this is the nPBF descriptor, for which F312 and F313 score the highest. Additionally, F024 and F189 or F058 and F264 are among the top-scoring fragments depending on whether the nPBF descriptor has been computed from the *in silico*-generated conformation or the protein-bound one. Interestingly, F313, F058, F312, F322, and F055 also exhibit the highest NP likeness values of all fragments (in decreasing order). This reflects the general trend that natural products, unlike many synthetic compounds, typically possess a higher degree of saturation and more stereocenters¹⁴—a characteristic also captured by our NP-like fragment hits.

Most Recently Approved PKIs Reflect the Trend. Recently, Bournez et al.³⁵ forecasted changes in the PKI landscape/chemical space in general, and new trends in terms of structures and molecular shapes in particular, based on their analysis of PKIs in clinical trials. Indeed, our analysis confirms this trend, as all the approved PKI for which the highest molecular descriptors for saturation, spatial complexity, and molecular three-dimensionality were reported, namely gilteritinib (2018), peficitinib (2019) and repotrectinib (2023), have all been approved by the regulatory authorities in the past few years. Similarly, every PKI with an overall Fsp³ ≥ 0.5 ,

namely gilteritinib, abrocitinib (2022), peficitinib, maribavir (2021), and leniolisib (2023), has entered the market within the last seven years. Since fully optimized inhibitors typically display saturation primarily in the peripheral decorations, we further calculated the Fsp³ for their Bemis-Murcko scaffolds. This again highlighted recently approved PKIs, peficitinib, gilteritinib, and delgocitinib (2024), as the ones with the highest carbon saturation in their central scaffold.

Despite these examples, a broader temporal analysis of approved PKIs reveals only a weak correlation between approval year and Fsp³, FC_{Stereo}, nSPS, or nPBF descriptor values (Figure S14, Spearman's rank correlation coefficient $|\rho| \leq 0.25$). This suggests that, although individual success stories exist, the approach of increasing saturation and 3D character so far had only little impact on approved PKIs. A recently published reevaluation of the 15-year-old "Escape from Flatland" paper⁷ by Churcher et al.¹³ likewise did not find a clear relationship between the highest phase reached and Fsp³ after 2009 anymore. This apparent disconnect in the last years was attributed to the focus on kinases and the widespread usage of metal-catalyzed cross-coupling reactions, which made it easier to introduce sp²-hybridized systems into drugs.¹³ Nonetheless, the success of more saturated and 3D inhibitors such as gilteritinib and repotrectinib, together with our own results, suggests that exploring this chemical subspace holds promise for diversifying and advancing next-generation PKIs.

Strategies for Fragment Evolution and Translation to Other Kinases. Our study provides a proof-of-concept that incorporating NP-like fragments enriched in sp³-hybridized carbons, stereo centers, and molecular 3D character into screening libraries can yield competitive hit rates, even against kinase targets. By accessing novel scaffolds, such an approach provides diverse derivatization vectors and geometries, increasing the chances of generating selective and potent molecules for targets that have so far proven difficult to modulate. This concept is broadly applicable beyond PKA, and holds potential to find starting points across diverse kinases and other protein families. For instance, Huschmann et al.¹⁸ recently highlighted the utility of the FRGx library⁴¹ for crystallographic hit identification against a protease, followed by fragment growth and rapid (off-the-shelf) SAR exploration using readily accessible follow-up compounds. Fragment merging or linking strategies may similarly be applied to NP-like fragment hits: Our ATP site kinase binders could potentially be combined with fragments from the kinase-focused fragmentation library (KinFragLib),^{68,69} in a subpocket-informed manner. As the fragments themselves, we expect the potential follow-up molecules to possess the same favorable properties with all the mentioned benefits, including a higher likelihood of progressing to the next phase in the drug development process.

Beyond the development of traditional small-molecule inhibitors, novel kinase binders offer promising avenues for PROTAC design. Both the ATP pocket and allosteric sites can in principle serve as anchor points for kinase-targeting degraders, as illustrated by BCR-ABL1-targeting PROTACs derived from type I (e.g., dasatinib), type II (e.g., ponatinib), and type IV (e.g., asciminib) kinase inhibitors.^{70,71} In particular, Yang et al.⁷¹ showed that engaging different sites on a protein of interest can tune degradation efficiency and mutant coverage. Another example where PROTACs have shown promise in kinase research is Bruton's tyrosine kinase (BTK), a critical target in B-cell malignancies such as mantle

cell lymphoma and chronic lymphocytic leukemia.⁷² Several BTK-directed PROTACs, such as NX-2127 and BGB-16673, are already in clinical trials.⁷² Given the high hit rate and flexible derivatization potential of NP-derived fragments, fragment-based approaches like ours could accelerate PROTAC development: Their defined 3D shape may confer scaffold rigidity and provide alternative exit vectors, allowing optimal spatial orientation of both ligase- and target-binding moieties and thereby facilitating the formation of the productive ternary complex.

In summary, the integration of natural product-inspired, three-dimensional fragments holds significant potential not only for the development of novel kinase inhibitors but also for advancing the design of modern drug modalities like PROTACs. This approach may be particularly beneficial in targeting kinases that are resistant to traditional inhibition strategies.

CONCLUSIONS

This study presented a crystal soaking campaign using 87 NP-like fragments against PKA, achieving an exceptionally high hit rate of 41%. These fragments are expected to be enriched in sp³-hybridized/chiral carbons compared to conventional fragment libraries which usually exhibit a preponderance of sp²-hybridized carbons, aromatic and heterocyclic rings. This disproves the initial concerns that enhanced molecular 3D character would have a negative impact on the hit identification success.^{2,10,11} None of the fragments detected as hits, nor the natural products comprising them as substructures, have previously been reported as active against a typical protein kinase with the conserved catalytic domain shared by PKA in publicly available data. The 36 determined fragment-bound crystal structures, have been fully refined and represent a high-quality data set for computational analyses. 35 structures displayed the kinase in the active (BLAminus) conformation, while the kinase in complex with F274 occupied the active-like (ABAminus) conformation. The latter was characterized by an unexpected fragment-induced flip of the peptide bond connecting the x (Thr183) and D (Asp184) residues of the xDFG motif. Out of the fragment hits, 9 corresponded to purely peripheral binders, 9 exclusively bound within the ATP site, and 18 fragments occupy simultaneously both orthosteric and peripheral sites. Notably, F189 was found to bind specifically at the underexplored allosteric site E, which is yet not occupied by any known PKA ligand stored in the PDB, but whose pharmacological relevance has been exemplified by asciminib (in BCR-ABL1 tyrosine kinase). A comparison of the fragments' Bemis-Murcko scaffolds to the ones from five reference data sets, as well as examination of the most similar reference molecules further highlights the novelty of the fragment binders demonstrating their potential for further exploration, while avoiding conflicts in the already crowded intellectual property space of kinase binders, and enabling insights into structural biology not observed before. Further cheminformatics analysis revealed that these novel binders indeed populate less frequently sampled regions of chemical space characterized by enriched saturation, (spatial) complexity, and molecular 3D character. Specifically, we demonstrated that the fragments bound to PKA exhibit statistically significant shifts toward higher values in the Fsp³, FC_{Stereo}, nSPS, and nPBF descriptors compared to established kinase binders from reference data sets. The latter was computed from in silico-generated as well as from the

crystallographically determined protein-bound ligand conformations, thereby enhancing the robustness of the shape analysis and ensuring comparability across data sets with and without experimentally resolved 3D structures. Notably, these favorable trends persist even within the subset of ATP site binders. These findings collectively emphasize how the resolved fragment hits challenge the planarity typical for kinase binders and venture beyond the flatland. While the spatial complexity is already inherent to the fragment's scaffold, it can possibly be further enhanced through peripheral decorations in the hit or lead optimization phase. Rather than relying on tedious incremental growth and saturation of sp²-carbon-rich hits into 3D-shaped molecules by sp³-carbon chemistry, we already start with 3D-rich kinase-binding fragments, which can be expanded by established chemistry. In conclusion, the fragment hits identified using a high-performance soaking system for PKA present novel perspective on hit identification and lead optimization in kinase drug discovery.

MATERIALS AND METHODS

Fragment Selection. A selection of 87 fragments from AnalytiCon Discovery's FRGx ('Fragments from Nature') library⁴¹ was subjected to the soaking campaign. The selection was based on the availability and diversity of the fragments for the soaking experiments first, and second, a random component. The number 87 corresponds to the number of crystals that could be measured within the allocated synchrotron beam time. As a result of the included randomness, some similar compounds were included in the selection, which we consider as advantageous in general, as it allows to assess how subtle structural variations influence binding modes. In order to be consistent with the molecule identifiers used in the FRGx library, we herein stick with a discontinuous numbering of the fragment hits.

Crystallography. The full-length catalytic subunit α of PKA from Chinese hamster (*Cricetus griseus*; UniProt-ID P25321, residues 1–350) was expressed, purified and crystallized as described by C. Siefker.⁷³ Compared to the human enzyme, 6 amino acid replacements are given at positions 26, 35, 40, 43, 45 and 349, all of which distant (>10 Å) from the ATP binding site, corresponding to a sequence identity of 98.3%. Fragments were dissolved in 100% DMSO to concentrations of 1 M or 500 mM. When fragments were not soluble at 500 mM, they were used as a saturated slurry. Soaking was performed by transferring one crystal each into a 2 μ L drop of SmartSoak⁷⁴-derived soaking condition which contained 10% of the fragment ligand stock. The exact soaking-buffer composition is proprietary, but can be provided upon reasonable request or within a collaborative framework. After 24 h of soaking, the crystals were harvested and flash-frozen. Data sets were collected at BESSY BL14.2 in Berlin.⁷⁵ The required number of images, degree of rotation per image, and exposition time were calculated by taking test images before and after a 90° rotation and following the proposed strategy as computed by Mosflm.⁷⁶ Data indexing, scaling, and integration were performed with XDS.⁷⁷ Afterward, a molecular replacement using an in-house structure of the PKA catalytic subunit was performed using PHASER.⁷⁸ The structures were refined by iterative cycles with Phenix,⁷⁹ while model building was performed with Coot.⁸⁰ The assigned molecules were truncated to the regions with well-defined electron density at σ

= 1 to minimize model bias. Crystallographic tables can be found in the [Supporting Information](#).

Analysis of Ligand-Bound Protein Kinase a Crystal Structures. Crystal structures were analyzed using Schrödinger's PyMOL (version 2.6.0)⁸¹ together with home-written Python scripts. All structures were aligned to an ATP-bound PKA structure, 3FJQ. To put the results of the analysis of the novel 36 crystal structures into context, we further analyzed all PKA complex structures in the PDB (vide infra).

Pocket occupation of a ligand/fragment was defined as "being located in a 5 Å radius of an added and well-considered pseudo atom." In the case of the orthosteric ATP-pocket, the pseudo atom was placed at the center of mass of the adenine substructure of ATP, to capture all hinge binders. In the case of the allosteric kinase pockets, the reference structures as listed by Xerxa et al.³⁷ were exploited, and the pseudo atom was placed in the center of mass of all ligands described to occupy the respective pocket.

To analyze if the binding of (peripheral) ligands may be influenced by the crystal environment we generated the symmetry mates within a cutoff distance of 6 Å of each ligand copy. Next, we determined if any ligand atom is closer to a neighboring symmetry mate in the crystal lattice than to the asymmetric unit itself. If so, we visualized the interactions of the ligand with each mate and computed the percentage of the ligand's surface area closer to the respective mate.

For the structural classification of PKA and the fragment hits we utilized the KINCORE⁵² stand alone program embedded in a bash script for batch processing of .pdb files.

Additionally, we evaluated the electron density support for the individual atoms in the FRG274:PKA complex structure, using the EDIASCORER 1.1.0 command line version. As we were primarily interested in the AB₂ conformation (occupancy = 0.6), we only kept this conformation of the residues Thr183 and Asp184 and manually removed the other from the .pdb file using a text editor. The EDIA⁵⁷ quantifies the electron density support of each atom in a crystallographically resolved structure. An EDIA value ≥ 0.8 indicates that the atom is well-covered with electron density. Values between 0.4 and 0.8 or values < 0.4 point out minor or substantial inconsistencies with the electron density fit, respectively. Multiple EDIA values can be combined with the help of the power mean to compute EDIA_m, the electron density score for multiple atoms to score a set of atoms such as a residue or a ligand.

Cheminformatic Analyses. The cheminformatics data analysis was performed using Python 3.10.13, together with JupyterLab 8.6.0. Python script preparation and data analysis were done in Visual Studio Code 1.99.3 using a virtual environment managed in Mamba 1.5.8, with code version control via Git 2.34.1. Data retrieval through APIs was conducted using the requests 2.31.0 library for general HTTP requests, e.g., for querying COCONUT, KLIFS and PubChem, the BIOTITE.DATABASE.RCSB module from biotite 0.39.0 for accessing PDB data, and the CHEMBL_WEBRESOURCE_CLIENT 0.10.9 for retrieving information from the ChEMBL database. Data processing and analysis was performed with the help of Pandas 2.1.4, and NumPy 1.26.4 packages. Chemical data processing, including molecular descriptor calculation, substructure search, and scaffold generation, was conducted with RDKit 2024.03.5.⁸² Presented plots were prepared with matplotlib 3.8.3, seaborn 0.11.0 or ptitprince 0.2.7. All JupyterNotebooks and the.yml environment file are available

on GitHub (<https://github.com/czodrowskilab/NPFrag2Kinase>).

Reference Data Sets. Natural Products. The COLleCtion of Open NatUral producTs (COCONUT) is one of the biggest and best annotated open-source databases for NPs.^{44,45} Canonical SMILES of the 695 133 NPs were downloaded from the COCONUT, version August 2024 (<https://coconut.naturalproducts.net/download/>). A stereospecific substructure search was conducted to identify all those NPs, comprising one of the fragments confirmed crystallographically as bound hits in this study as a substructure. Using the NPs' SMILES as input, we queried two additional databases, PubChem⁴² and ChEMBL,^{43,83} via their APIs, for additional information on the molecules, including biological target annotations and bioactivity values.

Kinase-Targeted Reference Data Sets. For the kinase-specific reference data sets, data mining was primarily based on the name or ID of the protein family (Pfam)⁴⁶ of typical kinases that share the same fold as PKA, i.e., PROTEIN KINASE DOMAIN or PF00069. For PKA-specific data sets, filtering was refined using PKA's enzyme commission (EC) number, being 2.7.11.11, along with limitation to the α -isoform of PKA's catalytic subunit (PKAC α).

For the ChEMBL35⁸³-based reference data sets, we first retrieved the target IDs of single proteins belonging to the protein kinase family, with protein classification "enzyme" and "kinase". In a subsequent step, the associated small molecules and bioactivity data were retrieved. After filtering for compounds that exhibit a pChEMBL value of ≥ 5.0 and/or a ligand efficiency (LE)⁸⁴ of ≥ 0.3 toward any kinase in an assay with a confidence score of 9, and molecule standardization (vide infra), a total of 38721 compounds was obtained.

From the data set described before, we created a subset of molecules bioactive toward PKAC α from different species. Applied filters were PKA's enzyme commission number (EC = 2.7.11.11), as it is nonspecies-specific, as well as the subunit and isoform indicators "catalytic" and "alpha" in the target name. In total, 71 molecules were identified to possess a pChEMBL value ≥ 5.0 and/or a LE ≥ 0.3 , toward PKAC α , originating from three different species, namely Human (*Homo sapiens*), European Rabbit (*Oryctolagus cuniculus*) and Brown Rat (*Rattus norvegicus*), all of which possessing >96% sequence identity.

For the PDB²⁹-based reference data set, first, the IDs of all structures tagged with the kinase domain Pfam name and the EC number of PKA were retrieved. Next, we used a GraphQL-based approach to retrieve the metadata associated with the structures. Structures corresponding to or comprising the regulatory subunit as well as complexes with protein substrates (of up to similar size as the PKA catalytic subunit), were excluded based on gene names. Subsequently, information on the nonpolymer entities, including their molecular representation as SMILES, were extracted and uncomplexed structures were eliminated. No more than one ligand per PDB structure is considered in the analysis, yielding a total of 160 unique ligands, that are next subjected to molecular standardization (vide infra). It must be noted that some of these ligands have been cocrystallized multiple times, corresponding to a total of 239 PKA:ligand complex structures in the PDB, all of them determined by X-ray crystallography, with a median resolution of 1.8 Å. The protein component originates from five different species, namely Human (*Homo sapiens*), Chinese hamster

(*Cricetulus griseus*), Cattle (*Bos taurus*), House mouse (*Mus musculus*), and Wild boar (*Sus scrofa*), all of which possessing >97% sequence identity. The most recently deposited complex structure included in our analysis is 8SF8 (February 2024).

A data set of approved drugs and clinical drugs targeting kinases is publicly available at <https://klifs.net/drugs.php>. This compilation integrates data from the Database of Protein Kinase Inhibitors in Clinical Trials (PKIDB)^{31,35} and the Kinase-Ligand Fingerprints and Structures (KLIFS) database.^{32,85} Therein included are 107 approved (clinical phase 4) PKIs. The most recently added PKI included in our analysis is Deuruxolitinib (July 2024).

Oral Drugs. In addition, we exploited ChEMBL as described earlier to compile a data set of organic small molecules in clinical phase 4, i.e., approved drugs that are administered orally. With the exclusion of prodrugs and compounds withdrawn from the market, and after molecule standardization (vide infra) and deduplication, the total number of drugs in this data set amounts to 1064.

Data Preparation. While each of the mined databases have already implemented molecular standardization procedures, molecules for which no structural information was available were prepared using RDKit's MolStandardize module to ensure consistency. This includes the disconnection of covalent bonds between metals and organic atoms (the disconnected metal is not preserved), desalting (keeping the largest fragment), neutralizing charges, SMILES canonicalization, and finally, deduplication. Alternatively, for those data sets, for which structural information was available, we used the protonation state(s) and tautomeric form(s) of the protein-bound ligand as determined by PROTOSS (vide infra).

Although few molecules were present in more than one of the data sets, none of the fragment hits has been found in the public data sets.

Scaffold/Chemotype Analysis. We compute the Bemis-Murcko scaffolds⁸⁶ and cyclic skeletons of the fragments and all reference molecules using RDKit,⁸² and analyze, in case which of our fragments represent scaffolds absent from the reference data sets. In contrast to the stringent scaffold definition, the term “chemotype” is only loosely defined.⁸⁷ We will use this term to describe molecules sharing similar structural features. Molecules of the same chemotype may, but not necessarily share the same scaffold. We additionally inspect the three top-scored and thus most similar reference molecules (irrespective from which of the reference data sets) per fragment. To this end, we computed the pairwise Tanimoto similarity for each fragment and reference molecule based on their Morgan fingerprints (radius = 2, 4096 bits) using RDKit.^{82,88}

Molecular Descriptors. All but one of the descriptors were calculated from the molecules' 2D structure. One additional descriptor was computed from the molecules' 3D structure, which was either generated in silico or obtained from crystallographic data representing the protein-bound state. This two-legged approach was chosen to facilitate comparisons with reference data sets lacking 3D structural information, such as those derived from ChEMBL, while also leveraging available experimental 3D data where possible.

For our fragment hits as well as for the PDB reference data set for which 3D structural information was available, we used the protonation state(s) and tautomeric form(s) of the protein-bound ligand as assigned by PROTOSS (vide infra).

In cases where multiple ligand copies have been resolved (in any crystal structure), possibly even representing different protomers/tautomers, the descriptor values were averaged per ligand to yield a total of 160 and 36 descriptor values for the PDB reference data set and our fragments, respectively.

Crystallographic Protein-Bound Conformation. We first identified the missing hydrogen atoms in the protein–ligand complex using PROTOSS,⁸⁹ filtered the ligands based on our definition, as well as for ligand copies for which all atoms have been resolved only. For eight ligands, no fully resolved copy was available in any complex structure, namely fragments F009 and F012, and PDB ligands 46P, 7W8, 9NT, AO8, BVZ and GGB. Instead, we selected the copy with the highest number of resolved non-hydrogen atoms for those. After filtering, a total of 270 ligand instances remained in the PDB data set and 57 fragment copies, which were subsequently subjected to descriptor calculation. The percentage of ligands for which more than one copy has been considered amounted to 39% for the fragments and 27% for the PDB ligands, respectively.

Computational Conformer Generation. Computational conformer generation relied on RDKit's EMBEDMOLECULE functionality starting from random coordinates⁹⁰ with a fixed seed to ensure reproducibility. All other parameter settings were in concordance with the preconfigured parameter object `RDKit.ALLCHEM.ETKDGv3()`. The RDKit's ETKDG approach combines experimental torsional-angle and ring geometry preferences with additional “basic knowledge” terms and the stochastic distance geometry-based approach.^{65,66} It has been shown to outperform other free and commercial conformer generators in producing a conformation closely resembling the bioactive conformation, typically with a root-mean-square deviation (RMSD) of less than 2 Å.^{67,91} This way, even if the generated ligand conformation may not equal the lowest energy conformation or the most stable in solution, it will closely match the conformation found in protein–ligand complexes.⁶⁷

Comparison of Conformations. To compare the in silico-generated conformers and the crystallographic ligands, the non-hydrogen atom RMSD was computed. For this purpose, an abstracted representation of the 3D molecules was used in which all non-hydrogen atoms were converted to carbon and all bond types to single bonds. This approach focuses on molecular shape rather than chemical detail, thereby avoiding mismatches.

Introduction to and Calculation of Descriptors. The Natural Product Likeness (NP likeness) score quantifies the overall similarity of a molecule to the structural space covered by NPs.⁶² Broadly speaking, it is calculated by comparing the structural features of a compound against the ones of a database of known NPs and normalizing for molecular size.

The fraction of sp³-hybridized carbons (F_{sp³}) and the fraction of stereogenic carbons (FC_{Stereo}) are widely employed, easily interpretable descriptors for characterizing saturation and spatial complexity of molecules. These metrics are defined as the number of sp³-hybridized carbons or stereogenic/chiral carbons, respectively, divided by the total number of all carbons in the molecule.⁷ When evaluating molecular 3D character using F_{sp³}, Kombo et al.⁶⁴ suggested a cutoff value of ≥ 0.42. In other words, at least 42% of the carbon atoms should be sp³-hybridized. The main limitation of both, F_{sp³} and FC_{Stereo} is that they disregard noncarbon atoms, which may constitute a considerable portion of the molecule and contribute to its 3D structure. Consequently, these descriptors

do not fully capture molecular topology and do not align with the chemical intuition of complexity. For instance, they may yield the lowest possible score of 0 even for molecules with considerable topological complexity, such as triphenylphosphine, or the highest score of 1 for relatively simple structures such as an aliphatic hydrocarbon chain.

The spatial score (SPS) is based on four parameters, calculated for each non-hydrogen atom in a molecule and are then summed across the whole structure: an atom hybridization term, a stereoisomeric term (strongly related to FC_{Stereo}), a non-aromatic ring term and the number of directly bonded non-hydrogen atoms.⁹² As the SPS depends on molecular size, the authors suggested normalizing the score by dividing by the total number of non-hydrogen atoms. The resulting nSPS was reported to correlate with biologically relevant properties, such as selectivity and potency.⁹²

The deviation from the plane of best fit (PBF) descriptor was the only descriptor computed from the molecules' 3D structures. Notably, as the conformers generated may differ among different software solutions, values cannot be directly compared between publications. The PBF descriptor measures the sum of the distances of all non-hydrogen atoms from the plane of best fit.⁹³ In general, a higher PBF value corresponds to a higher degree of three-dimensionality. While its theoretical range extends from zero to infinity, in practice, the PBF value tends to be < 2 for small drug-like molecules even if they are rich in 3D-character. Two examples of particularly globular or planar molecules are the NPs and approved drugs adamantine and salicylic acid, for which PBF values of 1.02 and 0.13 are obtained. However, to account for its size-dependency, the PBF value should be either only used to compare similarly sized compounds or should be normalized,⁹⁴ e.g. divided by the number of non-hydrogen atoms, which was also applied in this study. The two example molecules feature nPBF values of 0.093 and 0.013, respectively.

While Fsp^3 , FC_{Stereo} and nSPS descriptors have originally been introduced as measures of saturation and the molecular (spatial) complexity,^{7,92} nPBF expresses the molecular shape/three-dimensionality.^{93,95} All the aforementioned descriptors were calculated using the respective RDKit functions.

Descriptor Analysis. On the one hand, we analyzed the descriptor values of all ligands, on the other hand, we made use of the binding site annotation determined earlier and examined the subdata set of ATP site ligands. To assess whether the descriptor values differed significantly across data sets, first, the nonparametric Kruskal–Wallis test was employed.⁹⁶ As it indicated significant differences, Dunn's posthoc test⁹⁷ with Bonferroni correction was used to perform pairwise comparisons between the fragments and the reference data sets for each descriptor. Descriptor differences were considered statistically significant if the p-value was below 0.05. The tests were implemented using `SCIPY.STATS.KRUSKAL` and `SCIKIT_POSTHOC.POSTHOC_DUNN` functions. Additionally, we report the effect size, in the form of Cohen's d .⁶³

■ ASSOCIATED CONTENT

Data Availability Statement

All data and code underlying this study are openly available from [GitHub](https://github.com/czodrowskilab/NPFrag2Kinase) (<https://github.com/czodrowskilab/NPFrag2Kinase>) and the PDB. The reference data sets were compiled from sources in the public domain, namely COCONUT, PDB, ChEMBL and KLIFS databases, as

described in the Methods section, and are stored in the same [GitHub](#) repository.

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.5c01952>.

Fragment copy information, including binding site annotations, and computed descriptor values (XLSX)
X-ray and crystallographic data (XLSX)
SMILES of the fragment hits and the 87-membered fragment library (CSV)
Additional figures and tables, e.g., descriptor value distributions, etc. (PDF)

Accession Codes

The authors will release the atomic coordinates of all 36 disclosed PDB structures upon article publication: F001 (9RDU), F005 (9RDV), F009 (9RDW), F012 (9RDX), F024 (9RDY), F030 (9RDZ), F032 (9RE0), F055 (9RE1), F058 (9RE2), F070 (9RE3), F073 (9RE4), F074 (9RE5), F102 (9RE6), F134 (9RE7), F138 (9RE8), F145 (9RE9), F168 (9REA), F184 (9REB), F186 (9REC), F188 (9RED), F189 (9REE), F203 (9REF), F225 (9REG), F236 (9REH), F248 (9REI), F264 (9REJ), F274 (9REK), F283 (9REL), F294 (9REM), F296 (9REN), F299 (9REO), F304 (9REP), F310 (9REQ), F312 (9RER), F313 (9RES), F322 (9RET).

■ AUTHOR INFORMATION

Corresponding Authors

Sergei Glinca – *CrystalsFirst GmbH, Marburg 35037, Germany*; orcid.org/0000-0002-0394-8468;
Email: sergei.glinca@crystalsfirst.com

Paul Czodrowski – *Johannes Gutenberg University Mainz, Department of Chemistry, Mainz 55128, Germany*;
orcid.org/0000-0002-7390-8795; Email: czodpaul@uni-mainz.de

Authors

Anna Santura – *Johannes Gutenberg University Mainz, Department of Chemistry, Mainz 55128, Germany*

Janis Müller – *CrystalsFirst GmbH, Marburg 35037, Germany*

Madita Wolter – *CrystalsFirst GmbH, Marburg 35037, Germany*

Ina-Charlotte Tutzschky – *Johannes Gutenberg University Mainz, Department of Chemistry, Mainz 55128, Germany*

Moritz Ruf – *CrystalsFirst GmbH, Marburg 35037, Germany*

Alexander Metz – *Philipps-University Marburg, Department of Pharmaceutical Chemistry, Marburg 35037, Germany*;
Present Address: Current address: UCB Biopharma SRL, Braine-l'Alleud, Belgium

Anna Sandner – *Philipps-University Marburg, Department of Pharmaceutical Chemistry, Marburg 35037, Germany*

Stefan Merkl – *Philipps-University Marburg, Department of Pharmaceutical Chemistry, Marburg 35037, Germany*

Gerhard Klebe – *Philipps-University Marburg, Department of Pharmaceutical Chemistry, Marburg 35037, Germany*

Complete contact information is available at:
<https://pubs.acs.org/doi/10.1021/acs.jcim.5c01952>

Author Contributions

A.Santura. is the main contributor of the manuscript, involved in conceptualization, coding, analysis, visualization, and writing of the original draft. J.M., M.W., M.R., A.M., A.S., S.M., and

S.G. conducted the wet lab experiments and/or the determination, refinement, and final validation of the protein–ligand crystal structures. I.-C.T. aided in parts of the descriptor analysis. J.M., M.W., G.K., S. G., and P.C. were involved in the internal review process of the manuscript. S.G. and P.C. supervised the project and provided funding and resources.

Notes

The work described herein was produced by the joint collaboration between employees of the universities of Mainz and Marburg and CrystalsFirst. As a commercial organization, CrystalsFirst provides drug discovery and structural biology services.

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank the Helmholtz-Zentrum Berlin für Materialien und Energie for the allocation of synchrotron radiation beamtime. In particular, they thank Manfred Weiss. The authors acknowledge the use of artificial intelligence tools, namely ChatGPT, BlackboxAI, and DeepL, which supported language refinement and coding assistance. These tools were not used for the de novo generation of text or content conceptualization.

REFERENCES

- (1) Bon, M.; Bilslund, A.; Bower, J.; McAulay, K. Fragment-Based Drug Discovery - The Importance of High-Quality Molecule Libraries. *Mol. Oncol.* **2022**, *16*, 3761–3777.
- (2) Morley, A. D.; Pugliese, A.; Birchall, K.; Bower, J.; Brennan, P.; Brown, N.; Chapman, T.; Drysdale, M.; Gilbert, I. H.; Hoelder, S.; Jordan, A.; Ley, S. V.; Merritt, A.; Miller, D.; Swarbrick, M. E.; Wyatt, P. G. Fragment-Based Hit Identification: Thinking in 3D. *Drug Discovery Today* **2013**, *18*, 1221–1227.
- (3) Meyers, J.; Carter, M.; Mok, N. Y.; Brown, N. On The Origins of Three-Dimensionality in Drug-Like Molecules. *Future Med. Chem.* **2016**, *8*, 1753.
- (4) Hamilton, D. J.; Dekker, T.; Klein, H. F.; Janssen, G. V.; Wijnmans, M.; O'Brien, P.; de Esch, I. J. Escape from Planarity in Fragment-Based Drug Discovery: A Physicochemical and 3D Property Analysis of synthetic 3D fragment libraries. *Drug Discovery Today: Technol.* **2020**, *38*, 77–90.
- (5) Walters, W. P.; Green, J.; Weiss, J. R.; Murcko, M. A. What Do Medicinal Chemists Actually Make? A 50-Year Retrospective. *J. Med. Chem.* **2011**, *54*, 6405–6416.
- (6) Klein, H. F.; Hamilton, D. J.; de Esch, I. J.; Wijnmans, M.; O'Brien, P. Escape from Planarity in Fragment-Based Drug Discovery: A Synthetic Strategy Analysis of Synthetic 3D Fragment Libraries. *Drug Discovery Today* **2022**, *27*, 2484–2496.
- (7) Lovering, F.; Bikker, J.; Humblet, C. Escape From Flatland: Increasing Saturation as an Approach to Improving Clinical Success. *J. Med. Chem.* **2009**, *52*, 6752–6756.
- (8) Lovering, F. Escape from Flatland 2: Complexity and Promiscuity. *Med. Chem. Commun.* **2013**, *4*, 515–519.
- (9) Ishikawa, M.; Hashimoto, Y. Improvement in Aqueous Solubility in Small Molecule Drug Discovery Programs by Disruption of Molecular Planarity and Symmetry. *J. Med. Chem.* **2011**, *54*, 1539–1554.
- (10) Hann, M. M.; Leach, A. R.; Harper, G. Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856–864.
- (11) Schuffenhauer, A.; Brown, N.; Selzer, P.; Ertl, P.; Jacoby, E. Relationships between Molecular Complexity, Biological Activity, and Structural Diversity. *J. Chem. Inf. Model.* **2006**, *46*, 525–535.
- (12) Johnson, J. A.; Nicolaou, C. A.; Kirberger, S. E.; Pandey, A. K.; Hu, H.; Pomerantz, W. C. K. Evaluating the Advantages of Using 3D-Enriched Fragments for Targeting BET Bromodomains. *ACS Med. Chem. Lett.* **2019**, *10*, 1648–1654.
- (13) Churcher, I.; Newbold, S.; Murray, C. W. Return to Flatland. *Nat. Rev. Chem.* **2025**, *9*, 140–141.
- (14) Feher, M.; Schmidt, J. M. Property Distributions: Differences between Drugs, Natural Products, and Molecules from Combinatorial Chemistry. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 218–227.
- (15) Over, B.; Wetzel, S.; Grütter, C.; Nakai, Y.; Renner, S.; Rauh, D.; Waldmann, H. Natural-Product-Derived Fragments for Fragment-Based Ligand Discovery. *Nat. Chem.* **2013**, *5*, 21–28.
- (16) Prescher, H.; Koch, G.; Schuhmann, T.; Ertl, P.; Bussenault, A.; Glick, M.; Dix, I.; Petersen, F.; Lizos, D. E. Construction of a 3D-shaped, Natural Product Like Fragment Library by Fragmentation and Diversification of Natural Products. *Bioorg. Med. Chem.* **2017**, *25*, 921–925.
- (17) Hanby, A. R.; Troelsen, N. S.; Osberger, T. J.; Kidd, S. L.; Mortensen, K. T.; Spring, D. R. Fsp³-Rich and Diverse Fragments Inspired by Natural Products as a Collection to Enhance Fragment-Based Drug Discovery. *Chem. Commun.* **2020**, *56*, 2280–2283.
- (18) Huschmann, F. U.; Mueller, J.; Metz, A.; Ruf, M.; Senst, J.; Glinca, S.; Schiebel, J.; Heine, A.; Klebe, G. Screening, Growing, and Validation by Catalog: Using Synthetic Intermediates from Natural Product Libraries to Discover Fragments for an Aspartic Protease Through Crystallography. *Crystals* **2024**, *14*, 755.
- (19) Baell, J. B. Feeling Nature's PAINS: Natural Products, Natural Product Drugs, and Pan Assay Interference Compounds (PAINS). *J. Nat. Prod.* **2016**, *79*, 616–628.
- (20) Wetzel, S.; Schuffenhauer, A.; Roggo, S.; Ertl, P.; Waldmann, H. Cheminformatic Analysis of Natural Products and their Chemical Space. *Chimia* **2007**, *61*, 355.
- (21) Denis, J. D. S.; Hall, R. J.; Murray, C. W.; Heightman, T. D.; Rees, D. C. Fragment-Based Drug Discovery: Opportunities for Organic Synthesis. *RSC Med. Chem.* **2021**, *12*, 321–329.
- (22) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The Protein Kinase Complement of the Human Genome. *Science* **2002**, *298*, 1912–1934.
- (23) Kooistra, A. J.; Volkamer, A. Kinase-Centric Computational Drug Development. *Annu. Rep. Med. Chem.* **2017**, *50*, 197–236.
- (24) Roskoski, R. A Historical Overview of Protein Kinases and Their Targeted Small Molecule Inhibitors. *Pharmacol. Res.* **2015**, *100*, 1–23.
- (25) Kudlinzki, D.; Linhard, V. L.; Saxena, K.; Sreeramulu, S.; Gande, S.; Schieborr, U.; Dreyer, M.; Schwalbe, H. High-resolution Crystal Structure of cAMP-Dependent Protein Kinase from *Cricetulus griseus*. *Struct. Biol. Cryst. Commun.* **2015**, *71*, 1088–1093.
- (26) Schiebel, J.; Radeva, N.; Krimmer, S. G.; Wang, X.; Stieler, M.; Ehrmann, F. R.; Fu, K.; Metz, A.; Huschmann, F. U.; Weiss, M. S.; Mueller, U.; Heine, A.; Klebe, G. Six Biophysical Screening Methods Miss a Large Proportion of Crystallographically Discovered Fragment Hits: A Case Study. *ACS Chem. Biol.* **2016**, *11*, 1693–1701.
- (27) Müller, J.; Klein, R.; Tarkhanova, O.; Gryniukova, A.; Borysko, P.; Merkl, S.; Ruf, M.; Neumann, A.; Gastreich, M.; Moroz, Y. S.; Klebe, G.; Glinca, S. Magnet for the Needle in Haystack: “Crystal Structure First” Fragment Hits Unlock Active Chemical Matter Using Targeted Exploration of Vast Chemical Spaces. *J. Med. Chem.* **2022**, *65*, 15663–15678.
- (28) Knighton, D. R.; Zheng, J.; Eyck, L. F. T.; Huu Xuong, N.; Taylor, S. S.; Sowadski, J. M. Structure of a Peptide Inhibitor Bound to the Catalytic Subunit of Cyclic Adenosine Monophosphate-Dependent Protein Kinase. *Science* **1991**, *253*, 414–420.
- (29) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.; Weissig, H.; Shindyalov, I.; Bourne, P. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (30) Roskoski, R. Properties of FDA-Approved Small Molecule Protein Kinase Inhibitors: A 2024 Update. *Pharmacol. Res.* **2024**, *200*, 107059.
- (31) Carles, F.; Bourg, S.; Meyer, C.; Bonnet, P. PKIDB: A Curated, Annotated and Updated Database of Protein Kinase Inhibitors in Clinical Trials. *Molecules* **2018**, *23*, 908.

- (32) Kanev, G. K.; de Graaf, C.; Westerman, B. A.; de Esch, I. J.; Kooistra, A. J. KLIFS: An Overhaul After The First 5 Years of Supporting Kinase Research. *Nucleic Acids Res.* **2021**, *49*, D562–D569.
- (33) Xing, L.; Klug-Mcleod, J.; Rai, B.; Lunney, E. A. Kinase Hinge Binding Scaffolds and Their Hydrogen Bond Patterns. *Bioorg. Med. Chem.* **2015**, *23*, 6520–6527.
- (34) Baier, A.; Szyszka, R. Compounds from Natural Sources as Protein Kinase Inhibitors. *Biomolecules* **2020**, *10*, 1546.
- (35) Bournez, C.; Carles, F.; Peyrat, G.; Aci-Sèche, S.; Bourg, S.; Meyer, C.; Bonnet, P. Comparative Assessment of Protein Kinase Inhibitors in Public Databases and in PKIDB. *Molecules* **2020**, *25*, 3226.
- (36) Laufkötter, O.; Hu, H.; Miljković, F.; Bajorath, J. Structure- and Similarity-Based Survey of Allosteric Kinase Inhibitors, Activators, and Closely Related Compounds. *J. Med. Chem.* **2022**, *65*, 922–934.
- (37) Xerxa, E.; Laufkötter, O.; Bajorath, J. Systematic Analysis of Covalent and Allosteric Protein Kinase Inhibitors. *Molecules* **2023**, *28*, 5805.
- (38) Schuld, P.; Grzesiek, S.; Schlotte, J.; Habazettl, J. M.; Jahnke, W.; Barys, L.; Cowan-Jacob, S. W.; Loo, A.; Wiget, A.; Manley, P. W. Structural and Biochemical Studies Confirming the Mechanism of Action of Asciminib, an Agent Specifically Targeting The ABL Myristoyl Pocket (STAMP). *Blood* **2020**, *136*, 34–35.
- (39) Wang, Y.; Wach, J.-Y.; Sheehan, P.; Zhong, C.; Zhan, C.; Harris, R.; Almo, S. C.; Bishop, J.; Haggarty, S. J.; Ramek, A.; Berry, K. N.; O’Herin, C.; Koehler, A. N.; Hung, A. W.; Young, D. W. Diversity-Oriented Synthesis as a Strategy for Fragment Evolution against GSK3 β . *ACS Med. Chem. Lett.* **2016**, *7*, 852–856.
- (40) Zhang, R.; McIntyre, P. J.; Collins, P. M.; Foley, D. J.; Arter, C.; von Delft, F.; Bayliss, R.; Warriner, S.; Nelson, A. Construction of a Shape-Diverse Fragment Set: Design, Synthesis and Screen against Aurora-A Kinase. *Chem.–Eur. J.* **2019**, *25*, 6831–6839.
- (41) AnalytiConDiscovery, *FRGx - Fragments from Nature*. <https://ac-discovery.com/fragments-nature/>.
- (42) Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.; Thiessen, P.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E. E. PubChem 2025 Update. *Nucleic Acids Res.* **2025**, *53*, D1516–D1525.
- (43) Davies, M.; Nowotka, M.; Papadatos, G.; Dedman, N.; Gaulton, A.; Atkinson, F.; Bellis, L.; Overington, J. P. ChEMBL Web Services: Streamlining Access to Drug Discovery Data and Utilities. *Nucleic Acids Res.* **2015**, *43*, W612–W620.
- (44) Sorokina, M.; Merseburger, P.; Rajan, K.; Yirik, M. A.; Steinbeck, C. COCONUT online: Collection of Open Natural Products Database. *J. Cheminf.* **2021**, *13*, 2.
- (45) Chandrasekhar, V.; Rajan, K.; Kanakam, S.; Sharma, N.; Weissenborn, V.; Schaub, J.; Steinbeck, C. COCONUT 2.0: A Comprehensive Overhaul and Curation of the COllection of Open NatUral productTs database. *Nucleic Acids Res.* **2025**, *53*, D634–D643.
- (46) Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G.; Sonnhammer, E. L. L.; Tosatto, S. C. E.; Paladin, L.; Raj, S.; Richardson, L. J.; Finn, R. D.; Bateman, A. Pfam: The Protein Families Database in 2021. *Nucleic Acids Res.* **2021**, *49*, D412–D419.
- (47) Wang, H.; De Rose, E. F.; London, R. E.; Shears, S. B. IP6K Structure and the Molecular Determinants of Catalytic Specificity in an Inositol Phosphate Kinase Family. *Nat. Commun.* **2014**, *5*, 4178.
- (48) Shears, S. B.; Wang, H. Inositol Phosphate Kinases: Expanding the Biological Significance of the Universal Core of the Protein Kinase Fold. *Adv. Biol. Regul.* **2019**, *71*, 118–127.
- (49) Davies, M. P.; Benitez, R.; Perez, C.; Jakupovic, S.; Welsby, P.; Rzepecka, K.; Alder, J.; Davidson, C.; Martinez, A.; Hayes, J. M. Structure-Based Design of Potent Selective Nanomolar Type-II Inhibitors of Glycogen Synthase Kinase-3 β . *J. Med. Chem.* **2021**, *64*, 1497–1509.
- (50) AnalytiConDiscovery, *Synthetic Screening Compounds NATx*. <https://ac-discovery.com/synthetic-screening-compounds/>.
- (51) Modi, V.; Dunbrack, R. L. Defining A New Nomenclature for The Structures of Active And Inactive Kinases. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 6818–6827.
- (52) Modi, V.; Dunbrack, R. L. Kincore: A Web Resource for Structural Classification of Protein Kinases and their Inhibitors. *Nucleic Acids Res.* **2022**, *50*, D654–D664.
- (53) Dar, A. C.; Shokat, K. M. The Evolution of Protein Kinase Inhibitors from Antagonists to Agonists of Cellular Signaling. *Annu. Rev. Biochem.* **2011**, *80*, 769–795.
- (54) Gavrin, L. K.; Saiah, E. Approaches to Discover Non-ATP Site Kinase Inhibitors. *MedChemcomm* **2013**, *4*, 41–51.
- (55) Wylie, A. A.; Schoepfer, J.; Jahnke, W.; Cowan-Jacob, S. W.; Loo, A.; Furet, P.; Marzinzik, A. L.; Pelle, X.; Donovan, J.; Zhu, W.; Buonamici, S.; Hassan, A. Q.; Lombardo, F.; Iyer, V.; Palmer, M.; Berellini, G.; Dodd, S.; Thohan, S.; Bitter, H.; Branford, S.; Ross, D. M.; Hughes, T. P.; Petruzzelli, L.; Vanasse, K. G.; Warmuth, M.; Hofmann, F.; Keen, N. J.; Sellers, W. R. The Allosteric Inhibitor ABL001 Enables Dual Targeting of BCR-ABL1. *Nature* **2017**, *543*, 733–737.
- (56) Paladini, J.; Maier, A.; Habazettl, J. M.; Hertel, I.; Sonti, R.; Grzesiek, S. The Molecular Basis of Abelson Kinase Regulation by its α I-Helix. *eLife* **2024**, *12*, RP92324.
- (57) Meyder, A.; Nittinger, E.; Lange, G.; Klein, R.; Rarey, M. Estimating Electron Density Support for Individual Atoms and Molecular Fragments in X-ray Structures. *J. Chem. Inf. Model.* **2017**, *57*, 2437–2447.
- (58) Patel, R. Y.; Doerksen, R. J. Protein Kinase Inhibitor Database: Structural Variability of and Inhibitor Interactions with the Protein Kinase P-Loop. *J. Proteome Res.* **2010**, *9*, 4433–4442.
- (59) Möbitz, H. The ABC of Protein Kinase Conformations. *Biochim. Biophys. Acta, Proteins Proteomics.* **2015**, *1854*, 1555–1566.
- (60) Alam, K. A.; Gani, O. A. S. B. M.; Engh, R. A. Inhibitor Binding to Mutants of Protein Kinase A with GGGxxG and GxGxxA Glycine-Rich Loop Motifs. *J. Mol. Recognit.* **2021**, *34*, No. e2882.
- (61) Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A ‘Rule of Three’ for Fragment-Based Lead Discovery? *Drug Discovery Today* **2003**, *8*, 876–877.
- (62) Ertl, P.; Roggo, S.; Schuffenhauer, A. Natural Product-Likeness Score and its Application for Prioritization of Compound Libraries. *J. Chem. Inf. Model.* **2008**, *48*, 68–74.
- (63) Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Routledge, 1988; p 567.
- (64) Kombo, D. C.; Tallapragada, K.; Jain, R.; Chewing, J.; Mazurov, A. A.; Speake, J. D.; Hauser, T. A.; Toler, S. 3D Molecular Descriptors Important for Clinical Success. *J. Chem. Inf. Model.* **2013**, *53*, 327–342.
- (65) Riniker, S.; Landrum, G. A. Better Informed Distance Geometry: Using What We Know To Improve Conformation Generation. *J. Chem. Inf. Model.* **2015**, *55*, 2562–2574.
- (66) Wang, S.; Witek, J.; Landrum, G. A.; Riniker, S. Improving Conformer Generation for Small Rings and Macrocycles Based on Distance Geometry and Experimental Torsional-Angle Preferences. *J. Chem. Inf. Model.* **2020**, *60*, 2044–2058.
- (67) McNutt, A. T.; Bisiriyu, F.; Song, S.; Vyas, A.; Hutchison, G. R.; Koes, D. R. Conformer Generation for Structure-Based Drug Design: How Many and How Good? *J. Chem. Inf. Model.* **2023**, *63*, 6598–6607.
- (68) Sydow, D.; Schmiel, P.; Mortier, J.; Volkamer, A. KinFragLib: Exploring The Kinase Inhibitor Space Using Subpocket-Focused Fragmentation And Recombination. *J. Chem. Inf. Model.* **2020**, *60*, 6081–6094.
- (69) Kramer, P. L.; Buchthal, K.; Sydow, D.; Leo, K. S.; Volkamer, A. CustomKinFragLib: Filtering The Kinase-Focused Fragmentation Library. *chemRxiv* **2025**.
- (70) Burslem, G. M.; Schultz, A. R.; Bondeson, D. P.; Eide, C. A.; Stevens, S. L. S.; Druker, B. J.; Crews, C. M. Targeting BCR-ABL1 in Chronic Myeloid Leukemia by PROTAC-Mediated Targeted Protein Degradation. *Cancer Res.* **2019**, *79*, 4744–4753.

- (71) Yang, Y.; Gao, H.; Sun, X.; Sun, Y.; Qiu, Y.; Weng, Q.; Rao, Y. Global PROTAC Toolbox for Degrading BCR-ABL Overcomes Drug-Resistant Mutants and Adverse Effects. *J. Med. Chem.* **2020**, *63*, 8567–8583.
- (72) Danilov, A.; Tees, M. T.; Patel, K.; Wierda, W. G.; Patel, M.; Flinn, I. W.; Latif, T.; Ai, W.; Thompson, M. C.; Wang, M. L.; Sun, C.; Stephens, D. M.; Thirman, M. J.; Gessner, M.; Wolff, J.; Schwab, A.; Tan, M.; Chan, D.; Meredith, E.; Wiestner, A. A First-in-Human Phase 1 Trial of NX-2127, a First-in-Class Bruton's Tyrosine Kinase (BTK) Dual-Targeted Protein Degradator with Immunomodulatory Activity, in Patients with Relapsed/Refractory B Cell Malignancies. *Blood* **2023**, *142*, 4463–4463.
- (73) Siefker, C. *Characterisation and Differentiation of Kinase Binding Pockets in PKA and PIM1 by Small Molecule Fragments Using Protein Crystallography*, Dissertation. University of Marburg, 2018.
- (74) Patent Cooperation Treaty (PCT) applications: WO 2,022,053,530 A1, WO 2,022,053,525 A1, WO 2,022,053,528 A1.
- (75) Mueller, U.; Förster, R.; Hellmig, M.; Huschmann, F. U.; Kastner, A.; Malecki, P.; Pühringer, S.; Röwer, M.; Sparta, K.; Steffien, M.; Uhlein, M.; Wilk, P.; Weiss, M. S. The Macromolecular Crystallography Beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current Status and Perspectives. *Eur. Phys. J. Plus.* **2015**, *130*, 141.
- (76) Dauter, Z. Data-Collection Strategies. *Acta Crystallogr., Sect. d: Biol. Crystallogr.* **1999**, *55*, 1703–1717.
- (77) Kabsch, W. XDS. *Acta Crystallogr., Sect. d: Biol. Crystallogr.* **2010**, *66*, 125–132.
- (78) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
- (79) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr., Sect. d: Biol. Crystallogr.* **2010**, *66*, 213–221.
- (80) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr., Sect. d: Biol. Crystallogr.* **2010**, *66*, 486–501.
- (81) Schrödinger LLC., *The PyMOL Molecular Graphics System, Version 1.3r1*. Schrödinger, 2010.
- (82) Landrum, G.; *RDKit: Open-source Cheminformatics Software*. <https://www.rdkit.org>. 2024.
- (83) Zdrzil, B.; Felix, E.; Hunter, F.; Manners, E. J.; Blackshaw, J.; Corbett, S.; de Veij, M.; Ioannidis, H.; Lopez, D. M.; Mosquera, J. F.; Magarinos, M.; Bosc, N.; Arcila, R.; Kizilören, T.; Gaulton, A.; Bento, A. P.; Adasme, M. F.; Monecke, P.; Landrum, G. A.; Leach, A. R. The ChEMBL Database in 2023: A Drug Discovery Platform Spanning Multiple Bioactivity Data Types and Time Periods. *Nucleic Acids Res.* **2024**, *52*, D1180–D1192.
- (84) Schultes, S.; de Graaf, C.; Haaksma, E.; de Esch, I.; Leurs, R.; Krämer, O. Ligand Efficiency as a Guide in Fragment Hit Selection and Optimization. *Drug Discovery Today: Technol.* **2010**, *7*, No. e157–e162.
- (85) Kooistra, A. J.; Kanev, G. K.; Linden, O. P. V.; Leurs, R.; Esch, I. J. D.; Graaf, C. D. KLIFS: A Structural Kinase-Ligand Interaction Database. *Nucleic Acids Res.* **2016**, *44*, D365–D371.
- (86) Bemis, G. W.; Murcko, M. A. The Properties of Known Drugs: Molecular Frameworks. *J. Med. Chem.* **1996**, *39*, 2887–2893.
- (87) Bajorath, J. Improving the Utility of Molecular Scaffolds for Medicinal and Computational Chemistry. *Future Med. Chem.* **2018**, *10*, 1645–1648.
- (88) Maggiora, G.; Vogt, M.; Stumpfe, D.; Bajorath, J. Molecular Similarity in Medicinal Chemistry. *J. Med. Chem.* **2014**, *57*, 3186–3204.
- (89) Bietz, S.; Urbaczek, S.; Schulz, B.; Rarey, M. Protoss: A Holistic Approach to Predict Tautomers and Protonation States in Protein-Ligand Complexes. *J. Cheminf.* **2014**, *6*, 12.
- (90) Spellmeyer, D. C.; Wong, A. K.; Bower, M. J.; Blaney, J. M. Conformational Analysis using Distance Geometry Methods. *J. Mol. Graphics Modell.* **1997**, *15*, 18–36.
- (91) Friedrich, N.-O.; Meyder, A.; de Bruyn Kops, C.; Sommer, K.; Flachsenberg, F.; Rarey, M.; Kirchmair, J. High-Quality Dataset of Protein-Bound Ligand Conformations and Its Application to Benchmarking Conformer Ensemble Generators. *J. Chem. Inf. Model.* **2017**, *57*, 529–539.
- (92) Krzyzanowski, A.; Pahl, A.; Grigalunas, M.; Waldmann, H. Spacial Score – A Comprehensive Topological Indicator for Small-Molecule Complexity. *J. Med. Chem.* **2023**, *66*, 12739–12750.
- (93) Firth, N. C.; Brown, N.; Blagg, J. Plane of Best fit: A Novel Method to Characterize the Three-Dimensionality of Molecules. *J. Chem. Inf. Model.* **2012**, *52*, 2516–2525.
- (94) Dekker, T.; Wijnmans, M.; de Esch, I. J. Comparative Analysis of Molecular 3D Metrics in Fragment Space and Beyond. *Drug Discovery Today* **2025**, *30*, 104443.
- (95) Sauer, W. H.; Schwarz, M. K. Molecular Shape Diversity of Combinatorial Libraries: A Prerequisite For Broad Bioactivity. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 987–1003.
- (96) Kruskal, W. H.; Wallis, W. A. Use of Ranks in One-Criterion Variance Analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583–621.
- (97) Dunn, O. J. Multiple Comparisons among Means. *J. Am. Stat. Assoc.* **1961**, *56*, 52–64.



CAS BIOFINDER DISCOVERY PLATFORM™

**PRECISION DATA
FOR FASTER
DRUG
DISCOVERY**

CAS BioFinder helps you identify targets, biomarkers, and pathways

Unlock insights

CAS
A Division of the
American Chemical Society