Journal of Medicinal Chemistry Article

J. Med. Chem. **2010,** *53,* 7979–7991 **7979**
DOI: 10.1021/jm100677s

Electron Density Guided Fragment-Based Lead Discovery of Ketohexokinase Inhibitors†

Alan C. Gibbs, Marta C. Abad, Xuqing Zhang, Brett A. Tounge, Francis A. Lewandowski, Geoffrey T. Struble, Weimei Sun, Zhihua Sui, and Lawrence C. Kuo*

Johnson & Johnson Pharmaceutical Research and Development, Welsh and McKean Roads, Spring House, Pennsylvania 19477, United States

Received June 7, 2010

A fragment-based drug design paradigm has been successfully applied in the discovery of lead series of ketohexokinase inhibitors. The paradigm consists of three iterations of design, synthesis, and X-ray crystallographic screening to progress low molecular weight fragments to leadlike compounds. Applying electron density of fragments within the protein binding site as defined by X-ray crystallography, one can generate target specific leads without the use of affinity data. Our approach contrasts with most fragment-based drug design methodology where solution activity is a main design guide. Herein we describe the discovery of submicromolar ketohexokinase inhibitors with promising druglike properties.

Introduction

Fragment-based drug design (FBDD[a]) has become an important tool for drug discovery over the past decade.[1−3] Typical FBDD campaigns employ screening of compounds with low molecular weight and low structural complexity against a therapeutic target of interest.[4] Fragments that bind or interact with the target are identified by way of a robust biophysical technique, including X-ray crystallography,[5−8] nuclear magnetic resonance spectroscopy,[9] mass spectrometry,[10] isothermal titration calorimetry,[11] and surface plasmon resonance.[12,13] These techniques, although disparate in atomic-level binding information, are ideal for FBDD screening because they are able to characterize fragments with weak affinity (millimolar to micromolar range) in a semiautomated to fully automated medium-throughput format. The screen is designed to identify fragments that interact with the target. These fragments are then evolved to leads using synthetic methods. Fragment elaboration, which is used throughout the FBDD campaign, routinely employs the growing, linking, or fusing of fragments alone or in combination.

Underlying FBDD are the following basic assumptions. First, the fragment as part of an elaborated molecule will bind in a similar orientation in the active site as an isolated molecule. This assumption, which holds in the majority of cases, forms the basis for rational structure based design of potent compounds. Second, the low molecular weight fragments (typically < 250 Da) used in fragment screening will have high binding efficiencies as optimized larger molecules.[14,15] Compounds with a high binding efficiency, defined as high free energy of binding per non-hydrogen atom, serve as optimal starting points for library design and compound elaboration. A successful FBDD program also requires the integrated collaboration between scientific disciplines including structural biology, medicinal chemistry, and computational chemistry.

Numerous FBDD success stories have been published, and more than a handful of FBDD-derived compounds have entered phase I clinical studies. A few of those have advanced to phase II/III clinical trials.[16] The successes originate from many different pharmaceutical companies covering a diverse range of target classes, including protein kinases, nuclear hormone receptors, and protein chaperones.[17] The FBDD methodologies leveraged apply varied detection methods, screening technologies, chemistry strategies, and protocols. Often more than one screening technology is used in concert with computational techniques such as virtual screening to further complement the discovery effort. Although successful FBDD campaigns differ in methodology and target class, they are often similar in that a major emphasis is placed on bioaffinity measurements to guide compound design.

The reported FBDD methodology herein is unique in that the rounds of elaboration from primary to secondary to tertiary library of compounds are conducted without any guidance from affinity data. Instead, they are guided solely with X-ray protein crystallographic data. Fragment electron density is utilized to reveal if, where, and how a molecular fragment binds. Comparison and superpositioning of two or more occurrences of fragment electron density, within a series of experiments, reveal the similarities and differences in binding interactions among the fragments and offer a "binding surface" map. This map is then used to direct the design of subsequent fragment libraries by either sprouting from a fragment of choice or fusion of non-hydrogen atoms from multiple fragments that occupy different areas in the binding map.

†Coordinates and structure factors for ketohexokinase complexes with compounds **1** and **2** have been deposited in the Protein Data Bank with accession code 3NBV, and those with compounds **3**, **4**, **5**, and **6** have been deposited as 3NBW, 3NC2, 3NCA, and 3NC9, respectively.

*To whom correspondence should be addressed. Phone: 215-628-7850. Fax: 215-540-4632. E-mail: LKUO@its.jnj.com.

[a] Abbreviations: AMP-PNP, adenylyl imidodiphosphate; ATP, adenosine triphosphate; CMC, Comprehensive Medicinal Chemistry database; FBDD, fragment-based drug design; HATU, *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate; rmsd, root-mean-square deviation.

Published on Web 10/29/2010

We believe that by not ranking compounds based on their solution activity, we enhance the design process with a "normalization" of weak and strong binders and encourage focus of the design to only the interactions between the fragment and target.

The FBDD protocol followed here is initiated by X-ray crystallography screening of a primary library of fragments followed by iterative design and synthesis of secondary and tertiary libraries. The primary library is a general purpose library; it is composed of approximately 900 fragments with no target class biases and no target-privileged functional groups or pharmacophores. Following screening of the primary library, a secondary library is designed based on the hits discovered. Similarly, a final tertiary library is designed following X-ray screening of the secondary library. In contrast to the primary library, the secondary and tertiary libraries are composed of 300−500 compounds each. Upon completion of tertiary library synthesis, bioaffinity data are collected. Tertiary library compounds meeting predefined thresholds in various bioactivity, pharmacokinetic, and selectivity assays will advance to "lead" status, and subsequent lead optimization will commence.

The value of FBDD as a lead discovery engine derives from its ability to act as an alternative source of leads for ongoing projects or as a primary source of leads for projects where traditional approaches have been unsuccessful. Using protein crystallography as the screening tool may only be realized if the protein target of interest possesses the following qualities: a robust supply of single crystals; crystals that are able to withstand fragment soaking at high concentration; a crystal form that allows fragments to soak into the active site of the protein; no occluded active sites when crystallized; better than 2.8 Å resolution X-ray diffraction. Our ketohexokinase crystals meeting all requisite crystallography qualities renders ketohexokinase an ideal candidate for FBDD.

Ketohexokinase catalyzes, with adenosine triphosphate (ATP) and potassium ion (K^+), the conversion of the furanose form of D-fructose to fructose 1-phosphate.[18] Ketohexokinase, aldolase B, and triokinase are the enzymes catalyzing the bulk of dietary fructose metabolism. A significant correlation has recently been observed between increased dietary sugar intake and the prevalence of non-insulin-dependent diabetes mellitus.[19] Non-insulin-dependent diabetes mellitus is often a consequence of the metabolic syndrome, a collection of disorders that include to varying degrees obesity, hypertriglyceridemia, insulin resistance, and hypertension.[20] It is because of this connection that regulation of dietary sugar metabolism may have therapeutic benefit for people suffering non-insulin-dependent diabetes mellitus, obesity, or hypertension. Ketohexokinase is directly implicated in regulation of dietary sugar metabolism; the product of its catalysis fructose-1-phosphate is shunted directly into the glycolytic pathway to provide a significant supply of carbon for the biosynthesis of fatty acids and very low density lipoproteins.[21] As dietary sugar intake increases, fatty acid and very low density lipoprotein syntheses are amplified. It is hypothesized that inhibition of ketohexokinase activity would suppress carbon supply for fatty acid and very low density lipoprotein syntheses. Small molecule inhibitors of ketohexokinase may be a practical means to achieve ketohexokinase modulation.

In this paper we describe the use of fragment electron density guided FBDD to design small molecule inhibitors of ketohexokinase. The eventual lead, compound **6**, exhibits selectivity against a panel of protein kinases as well as very favorable pharmacokinetic properties.

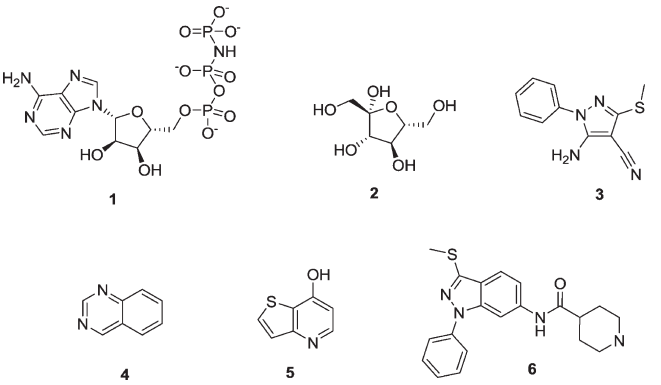


Figure 1. Two-dimensional representations of small molecules that bind ketohexokinase. X-ray crystallography derived structures of these molecules, together with ketohexokinse, have been determined during this study.

Results and Discussion

Ketohexokinase Structure. Ketohexokinase catalyzes the phosphorylation of a wide range of furanose substrates, most notably fructofuranose.[18] The ketohexokinase monomer is 298 amino acid residues in length and is a member of the ribokinase superfamily of kinases.[22] Enzymes within the ribokinase family show similar tertiary structures despite a low amino acid sequence homology. They catalyze the phosphorylation of a variety of substrates. Examples include D-fructose,[23] adenosine,[24] pyridoxal,[25] 4-amino-5-hydroxymethly-2-methylpyrimidine,[26] and 2-keto-3-deoxygluconate.[27]

An alternatively spliced gene gives rise to a hepatic active and a peripherally distributed, inactive isoform of KHK.[28] This study focuses specifically on the active ketohexokinase C isoform. The ketohexokinase protein construct used in this study contains 294 residues (amino acids 5−298). Three-dimensional structures of both isoforms have been determined via X-ray crystallography[29] with ketohexokinase A and ketohexokinase C in the apo form and ketohexokinase A in complex with AMP-PNP (an ATP surrogate). The two isoforms have high sequence homology, with differences confined to a region of 44 residues primarily localized in the β-sheet domain. Figure 1 illustrates two-dimensional representations of AMP-PNP (compound **1**) and pertinent small molecules whose ketohexokinase complex structures have been determined in this study.

Ketohexokinase tertiary structure is organized into two domains, a large $\alpha/\beta/\alpha$ fold and a small β-sheet region. The biologically relevant quaternary structure of ketohexokinase is a homodimer.[23] The dimer interface between the two monomers occurs through the β-sheet region. The two β-sheet regions interact with each other to form a twisted β-barrel, also known as the β-clasp[30] as shown in Figure 2A. The β-clasp (residues 13−40 and 98−114) acts as a lid for the active site and is thought to allow substrate, product, and cofactor to enter and exit. One active site per monomer is located between the $\alpha/\beta/\alpha$ fold and β-sheet domains. The active site contains residues responsible for positioning ATP, furanose, and a K^+ ion for chemistry. Figure 2B depicts an enlarged view of the ketohexokinase active site with bound AMP-PNP and D-fructose (compounds **1** and **2**).

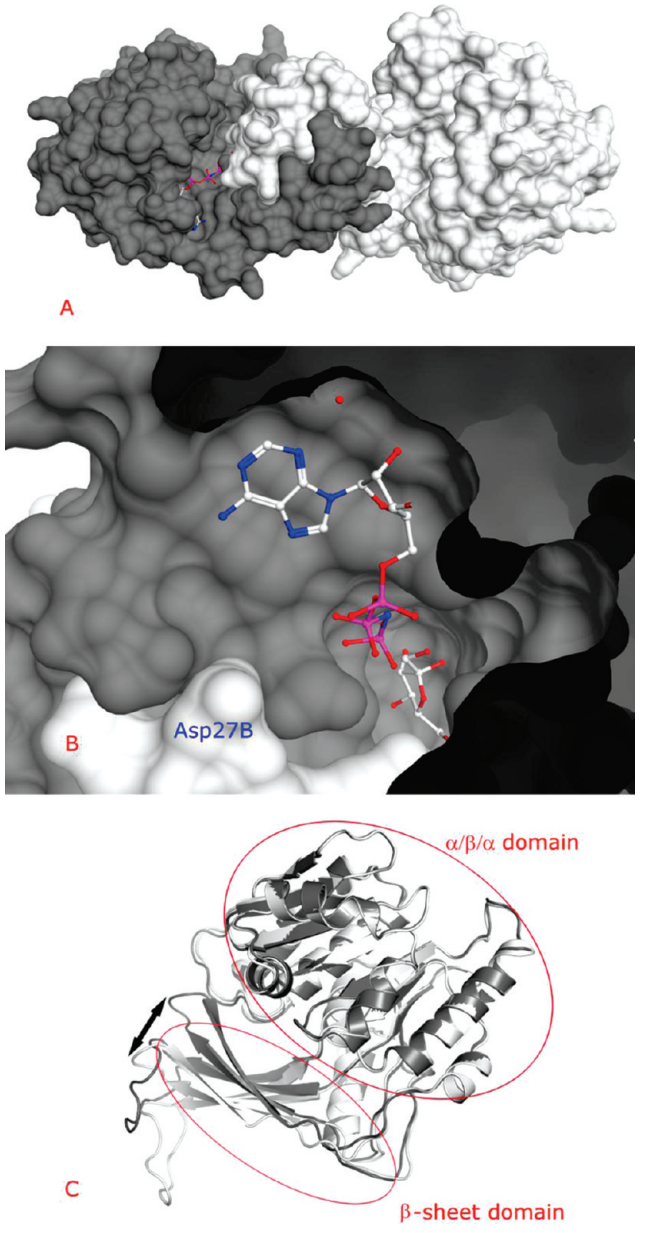


Figure 2. Structural features of ketohexokinase at 2.3 Å resolution. (A) Ketohexokinase dimer topology. The interlocking β-clasp is visible at the dimer interface. Monomer-A (gray) is in the closed conformation, and monomer-B (white) is in the open conformation. AMP-PNP is partially visible in monomer-A. (B) Catalytic site of ketohexokinase C with AMP-PNP (compound **1**) and D-fructose (compound **2**). AMP-PNP forms hydrogen bonds through N-3 of adenine to the conserved water (red sphere) and through HO-3′ of ribose to Gly229. D-Fructose is positioned below the terminal amidophosphate of AMP-PNP (γ-phosphate of ATP). Extension of Asp27B into the active site is clearly visible in the bottom of the figure. Although not visible in the density, Mg^{2+} or Mn^{2+} may also facilitate ATP and AMP-PNP binding to ketohexokinase. D-Fructose binds in a sterically constrained pocket and is held in place by a bidentate interaction between HO-3′ and HO-4′ with Asp15, a hydrogen bond between HO-1′ and Asn45, and a hydrogen bond between HO-6′ and Glu29B. (C) Ribbon representation of the backbone superimposition of monomer-A (gray) and monomer-B (white). The $\alpha/\beta/\alpha$ domains are very similar in conformation. In contrast, the β-sheet domain (comprising the β-clasp) conformations are very different and the difference between the open and closed states is dramatic. The arrow indicates a distance of 11 Å between the open and closed β-sheets. All three-dimensional structure figures were created by MOE (Chemical Computing Group, http://www.chemcomp.com).

The AMP-PNP bound structure of ketohexokinase A has previously been published,[29] the AMP-PNP bound structure of ketohexokinase C has not.

Inspection of the dimer backbone C_α trace reveals two distinct monomer conformations. One monomer possesses an open active site, and the other conformer possesses a closed active site. As shown in Figure 2C, the C_α traces of the two monomers are very similar throughout the $\alpha/\beta/\alpha$ domains but deviate extensively in the β-sheet domain, showing a root-mean-square deviation (rmsd) of 4.3 Å for the non-hydrogen atoms. A C_α backbone superimposition of ketohexokinase with two other known dimers (adenosine kinase[31] and ribokinase[32]) of the superfamily also reveals a low rmsd for the C_α backbone within the $\alpha/\beta/\alpha$ fold domain and a high backbone non-hydrogen atom rmsd for the β-sheet domain.

Our ketohexokinase crystals in space group $P2_12_12_1$ possess both open and closed conformations of ketohexokinase independent of ligand binding. On the other hand, ligand binding can be conformation dependent. Specifically, we have observed that some fragments bind only the closed-conformation monomer, some bind only the open-conformation monomer, and some bind both. Herein, we designate the closed conformation as monomer-A and the open conformation as monomer-B.

It is monomer-A that is catalytically competent. The closed conformation of ketohexokinase positions all catalytic residues, substrate, cofactor, and K^+ appropriately to allow phosphorylation to occur. Residues aligned for catalysis include the conjugate base, Asp258, and the β-clasp residue Arg108. Asp258 extracts the 1-OH proton from D-fructose, facilitating a 1-O^- nucleophilic substitution at the γ-phosphate of ATP which is anchored in position by Arg108.[30] Although its catalytic involvement is unclear, Asp27 in monomer-B (Asp27B) may be important for other processes. Asp27B extends from monomer-B into the active site of monomer-A to within approximately 6 Å of bound AMP-PNP, a feature not previously observed in any crystal structure from the ribokinase superfamily. This feature is exploited during fragment library design throughout our FBDD campaign.

Primary Library. In putting together the primary library, we applied a variety of filters to include the number of non-hydrogen atoms being ≥ 6 and ≤ 15, number of hydrogen bond acceptors being ≤ 3, number of hydrogen bond donors being ≤ 3, number of rings being ≤ 2, and no unspecified chiral centers. Compounds with moieties containing reactive, toxic, or nondruglike groups are removed. In total 45 specific functional groups and substructures were excluded (see Supporting Information for a complete list).

Next, the fragments were compared to the Comprehensive Medicinal Chemistry database (CMC),[33] a collection of over 8000 known pharmaceutical compounds, to determine if a fragment was a substructure in a CMC compound. If so, the compound was given a CMC subsimilarity (CMC_Subsim) score of 1. If not, it was assigned a score of 0. An overall FBDD score (FBDD_Score) was created by applying the equation

$$\text{FBDD\_Score} = \text{CMC\_Subsim} - [(\text{number of non-hydrogen atoms} - 6)/9] - (\text{percent heteroatom})$$

The result was a 900-member primary library of fragments.

Table 1. Crystallographic Data Collection and Refinement Statistics[a]

	compd, PDB accession no.				
parameter	**1** and **2**, 3NBV	**3**, 3NBW	**4**, 3NC2	**5**, 3NCA	**6**, 3NC9
space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
unit cell					
a, Å	82.8	82.7	83.4	82.9	82.7
b, Å	86.0	85.6	85.1	85.5	85.6
c, Å	136.9	136.6	137.2	137.5	136.8
resolution, Å	2.3	2.3	2.5	2.6	2.4
completeness, %	98.2 (98.3)	94.4 (78.8)	98.8 (99.1)	99.6 (99.5)	99.2 (99.9)
R_{merge},[b] %	7.1 (32.6)	5.3 (28.9)	4.5 (30.0)	6.1 (37.0)	5.9 (39.0)
$\langle I\rangle/\langle\sigma_I\rangle$	18.9 (3.6)	16.1 (3.1)	16.0 (3.4)	21.1 (3.9)	10.6 (2.7)
R_{factor},[c] %	20.7	22.4	23.3	19.2	23.6
R_{free},[d] %	23.6	27.4	27.1	23.9	28.6
rmsd bond, Å	0.008	0.008	0.008	0.009	0.009
rmsd angle, deg	1.12	1.17	1.10	1.17	1.16

[a] Values in parentheses refer to the highest resolution shell. [b] $R_{merge} = \sum_{hkl}\sum_I(|I_I - \langle I\rangle|/\langle I\rangle)$, where I_I is an individual intensity measurement and $\langle I\rangle$ is the average intensity for this reflection, with summation over all data. [c] $R_{factor} = \sum||F_o| - |F_c||/\sum|F_o|$. [d] 10% of the total reflections withheld.

Fragment Screening. Fragments or compounds from both the primary and secondary libraries were screened in groups of five by selecting five of the nearest-neighbor compounds in a structural-similarity space, based on substructure or two-dimensional topological descriptors. Although clustering by structural similarity would be counterintuitive if one were to deconvolute bound structures, grouping similar fragments reinforced identification of binding of a common chemotype, thus minimizing scaffold ambiguity when electron density was found in the active site. This was particularly advantageous in cases where electron density in the binding site lacked perfect definition. The combination of structurally similar fragments in a cocktail provided a "binary" answer of "yes" or "no" on binding. In our FBDD protocol, X-ray crystallographic screening was complete after the secondary library screen. Tertiary library compounds were not clustered into groups of five; selected compounds were soaked individually for nonscreening mode analysis.

ATP-Site Feature Map. Ketohexokinase crystals were soaked overnight with the primary library fragment cocktails and subjected to X-ray diffraction analysis. Protein structure determination was performed on all data sets. The resolved structures found to contain extra electron density within the protein active site due to fragment binding were further refined. During structure refinement, one of the five fragments in the cocktail was modeled into the observed extra electron density in the active site of the protein. The model was then optimized with electron density weighting giving the most reasonable orientations and binding modes as measured by typical crystallographic statistics, as shown in Table 1. Occasionally the electron density of the fragment was ambiguous allowing more than one orientation to fit or a second fragment from the cocktail to fit. In these cases, all possible structures were refined. In total approximately 60 data sets were refined, giving an unusually high hit rate of 30%, where a hit was defined as discernible active-site fragment density. The majority of fragment hits were found to bind monomer-A. For this reason, we used only the structure of this monomer in the design of compounds for the secondary and tertiary libraries. Unless stated otherwise, all discussions below refer to monomer-A.

Figure 3 illustrates the structural diversity and scope of binding within the active site of ketohexokinase of hits from the primary library fragments. The hit fragments are

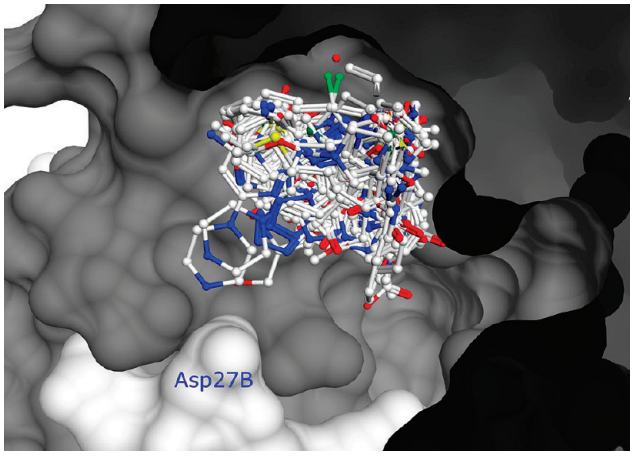


Figure 3. X-ray crystallographic determination of all primary library "hits" in the active site of ketohexokinase. Sixty fragments are shown. The fragments are accommodated in a wide range of binding orientations. The resolution of the bound structures ranges from 2.3 to 3.0 Å.

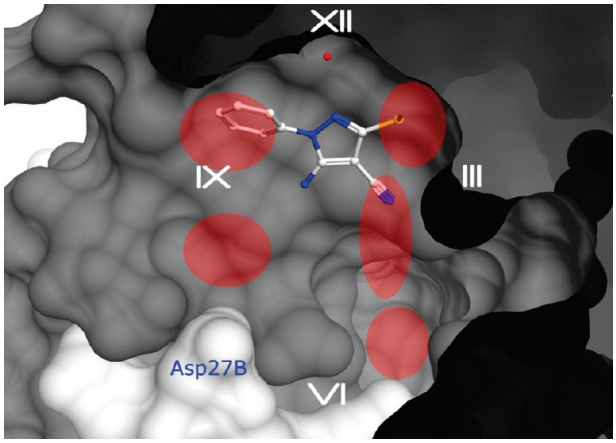


Figure 4. Ketohexokinase ATP-site feature map at 2.3 Å resolution. Regions of interest, largely defined by primary library fragment binding, are highlighted red; see text for details. A primary library pyrazole (compound **3**) is shown hydrogen-bonding to the conserved water, red sphere. The water molecule, in the reference frame, is at 12 o'clock.

accommodated in a moderately sized volume roughly characterized by a 14 Å × 11 Å × 3 Å box, or ~460 Å^3, to yield an "ATP-site feature map".

Figure 4 details the ATP-site feature map as defined by hits from the primary library screen. At the 12 o'clock position of the figure, an explicit density from a conserved water molecule, donating a hydrogen bond to the main-chain Cys282 carbonyl, is often well resolved in all our ketohexokinase structures. The conserved water molecule is observed in many related kinases of the ribokinase superfamily although the residue it interacts with is not conserved. In the majority of cases, we find that the conserved water also donates a hydrogen bond to the bound ligand. This protein−water−ligand interaction plays a key role in mediating binding of ATP, AMP-PNP (through N-3 of adenine; see also Figure 2B), and the majority of fragments to ketohexokinase. Interestingly, among all fragment-bound ketohexokinase structures solved in this study, only one fragment displaced the conserved water.

At the 2 o'clock position, the side chains of Cys282 and Phe260 define a small pocket that accommodates the 2′ and 3′ ribose hydroxyl groups of the endogenous ligand ATP (see also Figure 2B). Fragments that bind near this pocket reveal the inherent specificity of the pocket for small alkyl (with apparent preference for S-methyl) and aromatic groups of non-ATP molecules. Often observed are $\pi-\pi$ interactions between the aromatic side chain of Phe260 and (hetero)aromatic fragments. T-shaped (edge-to-face) interactions are usually observed, although face-to-face interactions are also represented.

At the 4 o'clock position is the "phosphate channel", where the endogenous ATP phosphate groups reside. This is a polar region with few primary library fragment hits. A common observation among our structures is the presence of a sulfate ion, a constituent of the crystallization buffer, occupying a similar position as the γ-phosphate of ATP and AMP-PNP. Residues Gly255, Gly257, and Arg108 coordinate the sulfate ion. Coincidentally, a sulfate ion is also observed in the apo crystal structure of monomer-A of ketohexokinase.[29] Because of its polar nature, this region provides sought-after protein hotspots for lead design.

At the 5 o'clock position is a small, extremely hydrophilic pocket where the furanose form of D-fructose binds. In the absence of sugar, the pocket recruits four water molecules. The oxygen atoms of these water molecules overlay with the four D-fructose hydroxyl oxygen atoms when superimposed, with a low rmsd. Several fragments are found to bind in this pocket, but they are not part of the design of compounds in this study.

At the 6 o'clock position, or the central region of the active site, is where the adenine heterocycle of AMP-PNP, or ATP, binds. This region is surrounded by nonaromatic, hydrophobic residues from both the $\alpha/\beta/\alpha$ and β-sheet domains and binds almost exclusively flat (hetero)aromatic fragments. The pocket is very narrow in the closed conformation of the enzyme, limiting accommodation and accessibility primarily to flat fragments, but displays little preference as to the composition of allowed aromatic rings. Five-member, six-member, fused bicyclic, fused tricyclic-aromatic, and heteroaromatic rings have all been found to bind.

At the 7 o'clock position, the carboxylate side chain of β-clasp residue Asp27B enters the catalytic site. Although it is possible that other residues from the monomer partner are involved in binding and/or catalysis, our observations indicate that Asp27B is the only residue involved in ATP-site fragment binding. The negatively charged carboxylate of Asp27B attracts fragments containing charged amines and basic groups, revealing another potentially favorable spot for ligand interaction. This interaction is successfully exploited in the design of a number of tertiary library compounds (described below).

Between the 8 and 10 o'clock positions is the egress of the active site to the bulk solvent, possibly the entrance for ATP and D-fructose as well as exit for ADP and fructose 1-phosphate. At the 10 o'clock position, three sequential proline residues (Pro246, Pro247, and Pro248) define the perimeter of a small hydrophobic binding pocket. Fragments and functional groups located in this region are composed of single ring aromatics or small (cyclo)alkyl moieties.

Secondary and Tertiary Libraries. The aim of the secondary library screen is to explore sprouting or merging of fragment hits found from the primary library. The aim of the tertiary library screen is to focus on a small number of chemotypes and to apply a greater emphasis on the synthesis of compounds with leadlike properties. The secondary library contains greater compound diversity for a larger number of chemotypes, whereas the tertiary library contains a concentration of leadlike compounds centered on a fewer number of chemotypes. How strict one selects primary library hits for the design of a secondary library is largely dictated by the hit rate. In the case of ketohexokinase, the primary library hit rate is high allowing a more stringent selection process to include novelty, diversity, ease of synthesis, and quality of fragment electron density.

Increasing bias is given to structural novelty as the chemotype progresses throughout secondary and tertiary library design, synthesis, and screening. The designed libraries must also be synthetically accessible, contain features that help elucidate the nature of specific protein−chemotype interactions, and possess leadlike physicochemical properties. Often various synthetic routes are required in order to accomplish these goals.

Our FBDD design method for ketohexokinase focuses on fragment electron density found within a radius of 5 Å from the location of the two bound ligands defined in the X-ray structure of ketohexokinase crystallized with AMP-PNP and D-fructose (shown in Figure 2B). The electron density features tracked during the course of this study are size, complementarity (goodness of fit), and location (see Figure 5).

Screening the Secondary Library. The secondary library was designed around hits found from the primary library to produce 6 scaffolds: II_A (pyrazoles), II_B (arylamides), II_C (benzotriazoles), II_D (pyridines), II_E (sulfones), and II_F (triazolones). They are shown in Figure 6 together with the substituents proposed for elaboration. Substituents are chosen on the basis of the electron density of the hits and the ATP-site feature map derived thereof. Approximately 350 compounds were synthesized. These compounds were screened in a manner identical to that performed for the primary library fragments. Hits were only discovered for compounds belonging to the sublibraries II_A, II_B, II_C, and II_D.

Screening the Tertiary Library. In contrast to the secondary library where fragment elaboration consisted primarily of growing, fragment fusion was heavily used in the preparation of fragments in the tertiary library. This is illustrated in Figure 7 for the tertiary sublibrary designated as III_A. Electron density relationships between pyrazole-containing fragments and various six-member aromatic rings found

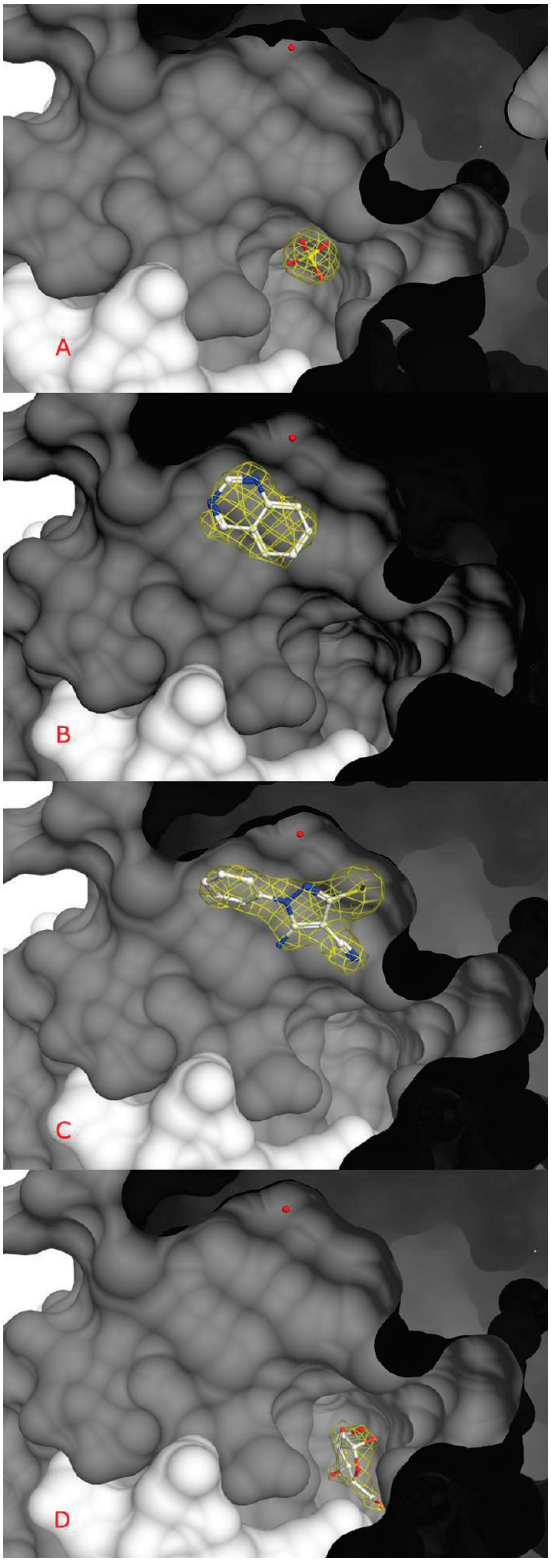


Figure 5. Electron density characteristics used for defining fragment constitution and relationships. A small sulfate ion is illustrated in panel A at 2.3 Å resolution. The density (at 2.5 Å resolution) of compound **4** in panel B allows more than one orientation of the symmetric fragment to fit. The density (at 2.3 Å resolution) of compound **3** in panel C allows an unambiguous fit. Compound **2** in panel D resides in the fructose binding site (2.3 Å resolution). For comparison the O-1′ of compound **2** is 9 Å from the nitrile N of compound **3** (C) when superimposed.

for hits from the primary and secondary (II_A) libraries revealed non-hydrogen atom overlaps (see Figure 8) and were the basis for the incorporation of 6:5 fused heterocycles into tertiary sublibrary III_A. Additional support for the incorporation of 6:5 analogues came from pyrazole hits containing 4-carboxamide and 5-amine functional groups that formed pseudo-six-member rings via intramolecular hydrogen bonding, as revealed by molecular mechanics and virtual docking exercises. The 6:5 fused systems of III_A (indazoles, imidazopyridines, pyrazolopyrimidines, and pyrazolopyrimidinones) provided the following compound properties: (1) As basic amines were commonly located around the 7 o'clock region of the active site (see Figure 4) making favorable interactions with Asp27B, extension of the scaffold relative to the parent pyrazole provided alternative, well positioned "handles" for attachment of amines. (2) The 6:5 analogues maintained hydrogen bonding with the conserved water via an appropriately positioned atom within the five-member heterocycle (equivalent to N-2 of the pyrazole). (3) Both aryl and small alkyl substituents on either side of the five-member ring, as dictated by the ATP-site feature map, could be maintained by the 6:5 systems.

Figure 9 outlines the design scheme for the tertiary sublibrary III_B. This series of fragments required a carboxamide for accepting a hydrogen bond from the conserved water and tolerated little diversity in the amide substituent relative to the parent cyclopropylethyl. Similar to sublibrary III_A, the core aromatic scaffold was extended on the basis of electron density comparison and virtual docking experiments, in attempts to make the available Asp27B interaction with a pendent basic amine.

In contrast to sublibraries III_A and III_B, compounds belonging to the III_C sublibrary primarily consisted of modification to the core scaffold shown in Figure 10 with maintenance of a 6:5 heterocycle. This core scaffold did not interact directly with the conserved water. Interactions with the water molecule were observed for parent analogues in secondary library II_C that contained hydrogen bonding substituents extending off the scaffold, e.g., pendent imidazole rings. Analogues were designed into this tertiary library to follow this trend.

Fragments of secondary sublibrary II_D were positioned in the active site by assuming interaction of the pyridine nitrogen with the conserved active site water. This nitrogen in the context of a six-member ring was held constant mostly in the design of the final tertiary sublibrary, III_D. Figure 11 illustrates the tertiary sublibrary III_D design scheme. In some cases, the six-member ring was extended to isoquinoline or semisaturated heterocycles. It was hoped that extension of the pyridine scaffold would provide a reasonable handle for not only basic amine attachment to interact with Asp27B but also phosphate-channel attractive substituents (see the 4 o'clock position in Figure 4). Unfortunately, compounds in any series with polar functional groups that were predicted to produce favorable interactions in the phosphate channel region did not show any discernible electron density in the bound structure of ketohexokinase.

Compound 6. All tertiary library compounds were tested for solution activity in an enzyme assay. This step was taken to transition the FBDD effort into a lead optimization effort. We found many tertiary library scaffolds offered at least one representative displaying an IC_{50} of $< 5\ \mu M$ in the ketohexokinase enzyme activity assay. These scaffolds are highlighted by the boxed structures in Figures 7, 9, and 10. The most promising ketohexokinase inhibitors came from sublibrary III_A, the indazoles, which led to the discovery of compound **6**.

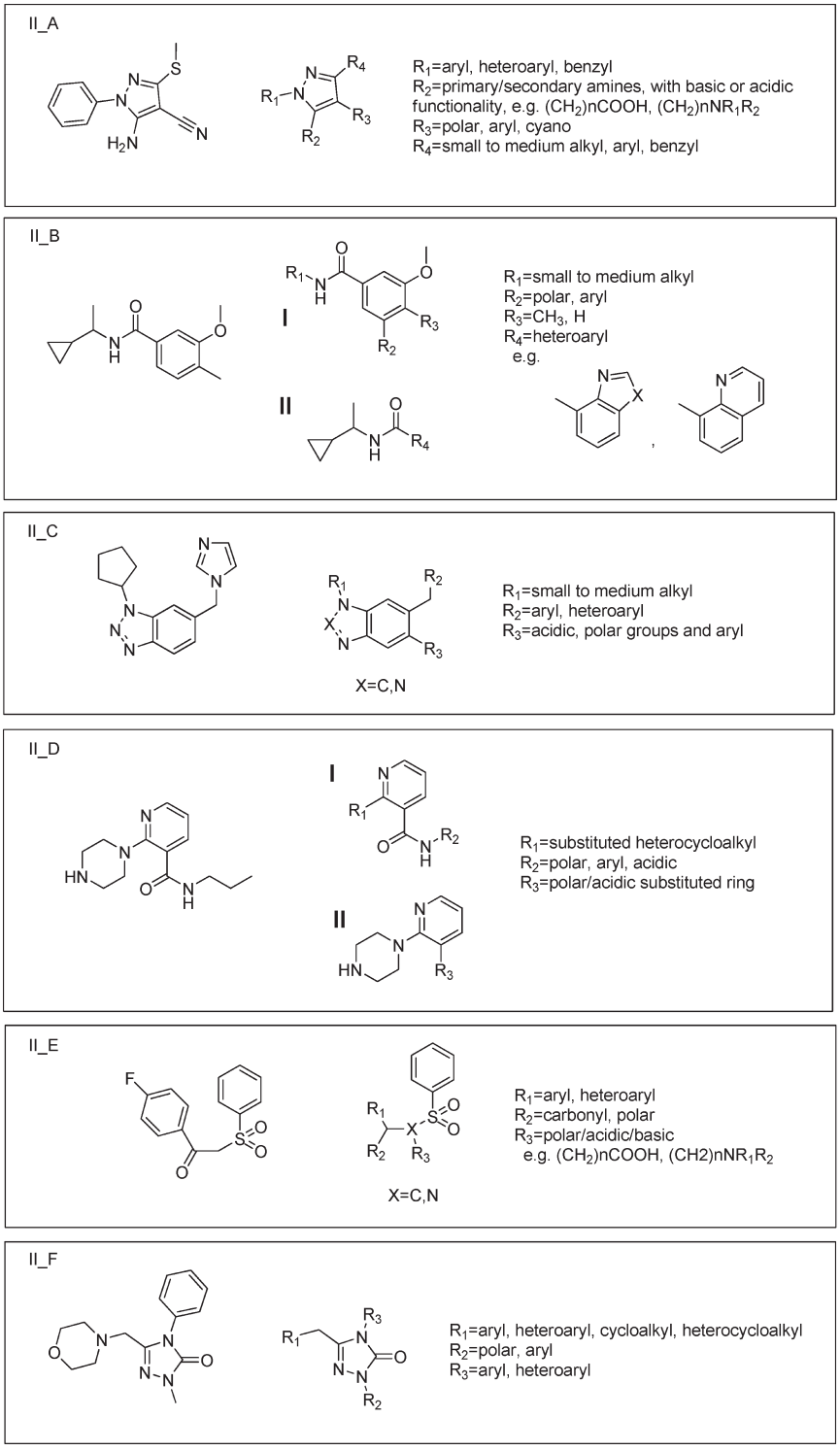


Figure 6. Secondary library design scheme. The six primary library fragments selected for secondary library design are shown at left. Corresponding constitution of the six sublibraries are illustrated at right: II_A (pyrazoles); II_B (aryl amides); II_C (benzotriazoles); II_D (pyridines); II_E (sulfones); II_F (triazolones).

The structural evolution of lead compound **6** from a representative pyrazole hit (compound **3**) is depicted in Figure 12. Our data indicate that large modifications are tolerated around certain regions of the parent pyrazole scaffold, whereas in other regions little change is tolerated. Although the search for optimal substituents was not exhaustive, minimal tolerance to modification of the *S*-methyl (at 1 o'clock as shown in Figure 4) and *N*-phenyl (at 11 o'clock as shown in Figure 4) is observed. In contrast, large structural modifications to the pyrazole ring extending the scaffold into the 6 and 7 o'clock regions of the active site (as shown in Figure 4), via a 6:5 fusion, are well tolerated and very advantageous for inhibitor design. Although many tertiary library indazoles displayed submicromolar affinity, all tertiary library pyrazoles or precursors of the indazoles showed IC_{50} values greater than 5 μM in the enzyme assay.

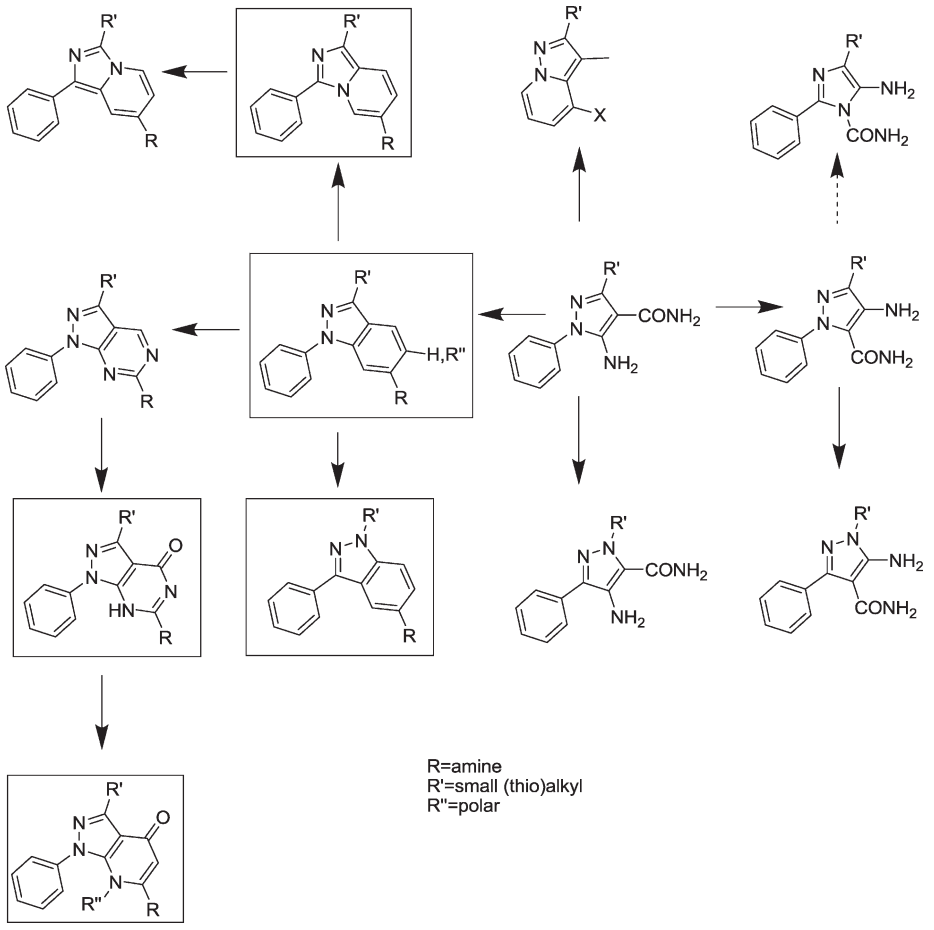


Figure 7. Tertiary sublibrary III_A design scheme. Various single ring and fused ring heterocycles make up the sublibrary. Boxed structures indicate scaffolds with at least one tertiary library representative testing below 5 μM in the ketohexokinase enzyme assay.

Compound **6** and its synthesis are shown in Scheme 1. Cyclization of 2-amino-4-nitrobenzoic acid **7** with *N*,*N*′-carbonyldiimidazole yielded the 7-nitro-1*H*-benzo[*d*][1,3]-oxazine-2,4-dione intermediate, which was then reacted with phenylhydrazine to generate compound **8** (~34% yield in two steps). Compound **8** was treated with $NaNO_2$ and HCl to give the indazolone intermediate which was then reacted with P_2S_5 to yield the corresponding indazolethione **9** (~11% yield in two steps). Methylation of compound **9**, with MeI in the presence of Cs_2CO_3 followed by reduction of the nitro group with hydrogenation at normal atmospheric pressure via the catalysis of Pd on carbon, gave compound **10** (~77% yield in two steps). Compound **10** was coupled with piperidine-1,4-dicarboxylic acid mono-*tert*-butyl ester in the presence of *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HATU). The adduct was then deprotected using HCl to yield the final product, compound **6**, in two steps with a 40% yield.

Compound **6** displays an IC_{50} value of 330 nM in the ketohexokinase enzyme assay. In inhibition assays with recombinant CYP450 enzymes, compound **6** shows no significant activity against any of the five major human CYP450 enzymes (CYP1A2, CYP3A4, CYP2C9, CYP2C19, CYP2D6, >10 μM). The compound is stable in human and rat liver microsomes (96% remaining at 10 min in HLM and 95% in RLM). Compound **6** possesses high aqueous solubility at both pH 7.4 and pH 2.0 (45.2 μM in PBS, pH 7.4; 35.3 μM in PBS at pH 2.0; 40.7 μM in water; 40 μM in ACN). The compound is found to be permeable through a monolayer of Caco-2 cells in vitro, with efflux ratio of 5.74 (P_{app}(A→B) = 1.65×10^{-6} cm/s; P_{app} (B→A) = 9.47×10^{-6} cm/s). Compound **6** is highly orally bioavailable, exhibiting an oral bioavailability of 75.4% in rats when dosed at 10 mg/kg using 20% hydroxypropyl-β-cyclodextrin (HPBCD) as the vehicle (10 mg/kg po, C_{max} = 529 ng/mL, AUC_{last} = 9031 h·ng/mL). Compound **6** shows moderate in vivo systemic clearance (11.6 (mL/min)/kg) and a high volume of distribution (8.53 L/kg).

The potential for cardiovascular side effects with compound **6** was investigated. The compound shows an IC_{50} of 8.13 μM in hERG channel astemizole binding assays. In single-cell patch clamp studies, inhibitions of the membrane K^+ current Ikr are observed at the three concentrations tested (18% at 0.1 μM, 27% at 0.3 μM, and 57% at 3 μM, compared to 10%, 16%, and 19% with solvent as the control). Selectivity of compound **6** against protein kinases (Invitrogen), G-protein-coupled receptors (Cerep), and ion channels (Cerep) were also evaluated. The Invitrogen protein kinase panel contains 200 kinases; compound **6** inhibits protein kinase A (99% inhibition) at 10 μM compound and 10 μM ATP. It does not inhibit any of the other protein kinases greater than 37% at 10 μM. As indicated by the Cerep panel screening, compound **6** inhibits three receptors greater than 70% at 10 μM; they are 5-HT1A (83% inhibition), 5-HT2B (91% inhibition), and 5-HT7 (95% inhibition). No ion channels are inhibited by compound **6**. As these apparent inhibition values are based on a single concentration of compound **6**, true inhibitions would need to be determined by dose response. Selectivity within the ribokinase super family was not tested.

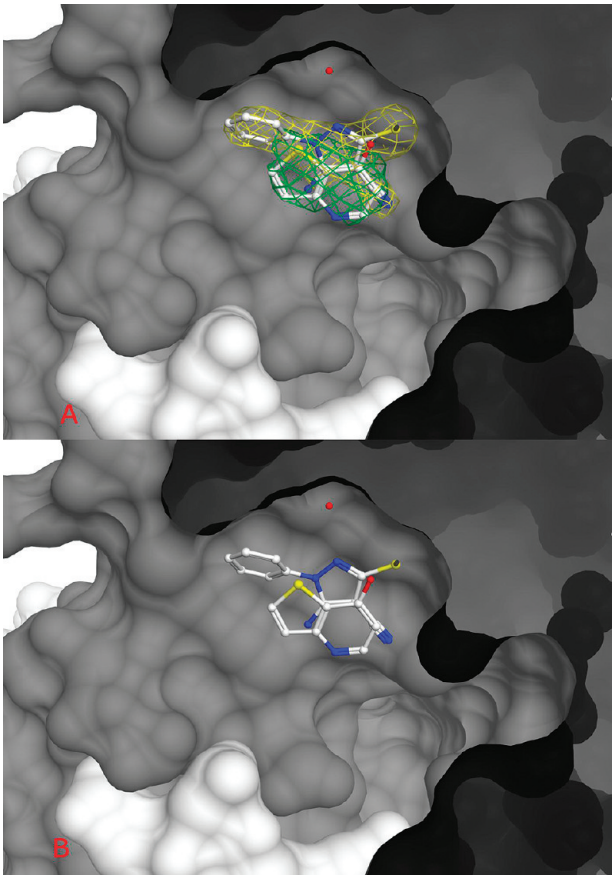


Figure 8. Example of electron density comparison used to fuse two fragment substructures. Panel A illustrates the density overlap between two experiments as shown by yellow and green density mesh (yellow density = compound **3**, green density = compound **5**). Panel B shows the non-hydrogen atom overlap between the fragments, the genesis of 6:5 fused rings from compound **3** as shown in sublibrary III_A.

Conclusion

The methodology described here provides a unique route to the discovery of lead compounds by FBDD-based approaches. The discovery of the lead series of ketohexokinase inhibitors shown here is driven with fragments from a primary library, followed by two rounds of compound elaboration, and directed by X-ray electron density maps without guidance from affinity data. The outcomes are compounds that are structurally unique, are leadlike in pharmacokinetic properties, are minimally restrained by restricted intellectual property space, and possess significant binding affinity toward the target of interest. All of these properties are embodied in compound **6**.

Experimental Section

Chemistry. ^{1}H nuclear magnetic resonance spectra were determined with a Bruker Biospin International AG-400 spectrometer at 400 MHz. Chemical shifts (δ) are reported in parts per million (ppm) relative to residual chloroform (7.26 ppm), TMS (0 ppm), or CD_3OD (4.87 ppm) as an internal reference with coupling constants (J) reported in hertz (Hz). The peak shapes are denoted as follows: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br, broad. Electrospray mass spectra were recorded in positive or negative mode on a Micromass Platform spectrometer. The purity of all compounds described here was determined by analytical LC/MS using a Shimadzu LCMS-2110EV instrument with analytical LC using Shiseido Capcell Pak C18 MGS3 column (3.0 mm × 50 mm, 3.5 μm), 0.05% TFA in water as mobile phase A, and 0.05% TFA in acetonitrile as mobile phase B at 1 mL/min flow, gradient from 10% to 100% B in 1.7 min, 100% B for 1.5 min, 100% to 10% B in 0.13 min, monitored by UV absorption at 215 nm using PDA and ELSD. The purity of all compounds was found to be >95%. Thin-layer chromatography was performed on Merck PLC prescored plates $60F_{254}$. Unless otherwise noted, reagents were obtained from commercial sources and were used without further purification.

Preparation of 6: *N*-(3-(Methylthio)-1-phenyl-1*H*-indazol-6-yl) piperidine-4-carboxamide Hydrochloride Salt. A solution of 2-amino-4-nitrobenzoic acid (54.6 g, 300.00 mmol, 1.00 equiv) in tetrahydrofuran (500 mL) and di(1*H*-imidazol-1-yl)methanone (58.32 g, 360.00 mmol) was added to a 1000 mL four-necked round-bottom flask, then purged and maintained with an inert atmosphere of nitrogen. The resulting solution was stirred for 2 h at 10 °C. The solids were collected by filtration. This resulted in 56 g (90%) of 7-nitro-1*H*-benzo[*d*][1,3]oxazine-2,4-dione as a yellow solid.

A solution of 7-nitro-1*H*-benzo[*d*][1,3]oxazine-2,4-dione (56 g, 269.23 mmol, 1.00 equiv) in ethanol (500 mL) and 1-phenylhydrazine (31.98 g, 296.11 mmol, 1.10 equiv) was added to a 1000 mL three-necked round-bottom flask. The resulting solution was heated to reflux for 5 min. The resulting mixture was concentrated under vacuum. The residue was applied onto a silica gel column with dichloromethane/petroleum ether (1:2 to 1:0) and ethyl acetate/petroleum ether (1:1). This resulted in 28 g (38%) of 2-amino-4-nitro-*N*′-phenylbenzohydrazide as a yellow solid. MS (*m*/*z*): 273 [MH^+].

A solution of 2-amino-4-nitro-*N*′-phenylbenzohydrazide (23 g, 84.56 mmol, 1.00 equiv) in ethanol (250 mL), 1 N hydrogen chloride (250 mL), and tetrahydrofuran (100 mL) was added to a 1000 mL four-necked round-bottom flask. This was followed by the addition of a solution of $NaNO_2$ (17.5 g, 253.62 mmol, 3.00 equiv) in water (45 mL) dropwise with stirring at 72 °C in 40 min. The resulting solution was heated to reflux for 1 h. The reaction mixture was cooled to 25 °C. The solids were collected by filtration. This resulted in 11.5 g (53%) of 6-nitro-1-phenyl-1,2-dihydroindazol-3-one as a brown solid. MS (*m*/*z*): 256 [MH^+].

6-Nitro-1-phenyl-1,2-dihydroindazol-3-one (15 g, 58.82 mmol, 1.00 equiv), xylene (150 mL), and P_2S_5 (52.2 g, 235.14 mmol) were added to a 250 mL three-necked round-bottom flask. The resulting solution was heated to reflux for 40 min in an oil bath. The reaction mixture was cooled to 25 °C. The resulting solution was diluted with 50 mL of ethyl acetate. The solids were filtered out. This resulted in 3.1 g (19%) of 6-nitro-1-phenyl-1,2-dihydroindazole-3-thione as a yellow solid. MS (*m*/*z*): 272 [MH^+].

A solution of 6-nitro-1-phenyl-1,2-dihydroindazole-3-thione (3.1 g, 11.44 mmol, 1.00 equiv) in CH_3CN (20 mL) and Cs_2CO_3 (21 g, 64.42 mmol, 5.63 equiv) was added to a 100 mL three-necked round-bottom flask. This was followed by the addition of a solution of iodomethane (11.3 g, 79.58 mmol, 6.96 equiv) in CH_3CN (15 mL) dropwise with stirring at 10 °C in 30 min. The resulting solution was stirred for 30 min at 10 °C. The solids were filtered out. The resulting mixture was concentrated under vacuum. The residue was applied onto a silica gel column with ethyl acetate/petroleum ether (1:100). This resulted in 2.8 g (86%) of 3-(methylthio)-6-nitro-1-phenyl-1*H*-indazole as a yellow solid. MS (*m*/*z*): 285 [MH^+].

A 100 mL round-bottom flask was purged, flushed, and maintained with a hydrogen atmosphere. A solution of 3-(methylthio)-6-nitro-1-phenyl-1*H*-indazole (3.5 g, 12.28 mmol, 1.00 equiv) in ethyl acetate (30 mL) was added to the flask followed by 5% palladium on carbon (2.2 g), then bubbled with H_2 (gas). The resulting solution was stirred for 48 min at room temperature. The solids were filtered out. The resulting mixture was concentrated under vacuum. This resulted in 2.8 g (89%) of 3-(methylthio)-1-phenyl-1*H*-indazol-6-amine as brown oil. MS (*m*/*z*): 256 [MH^+].

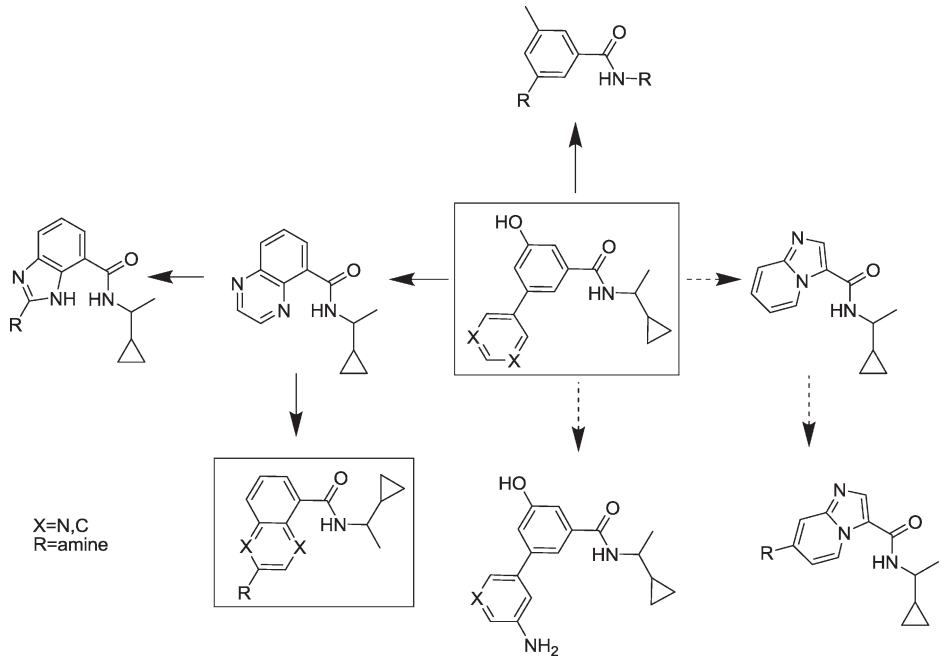


Figure 9. Tertiary (III_B) sublibrary design scheme. The carboxamide, important for hydrogen bonding to the conserved water, is maintained in all compounds. Boxed structures indicate scaffolds with at least one tertiary library representative testing below 5 μM in the ketohexokinase enzyme assay.

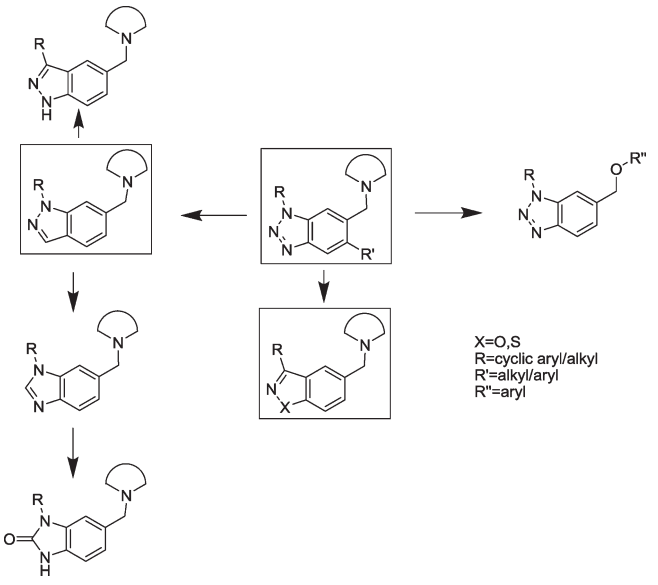


Figure 10. Tertiary (III_C) sublibrary design scheme. No fragment fusion was involved in the design of sublibrary III_C. Boxed structures indicate scaffolds with at least one tertiary library representative testing below 5 μM in the ketohexokinase enzyme assay.

A 50 mL three-necked round-bottom flask was purged and maintained with an inert atmosphere of nitrogen. A solution of 3-(methylthio)-1-phenyl-1*H*-indazol-6-amine (500 mg, 1.96 mmol, 1.00 equiv) in *N*,*N*-dimethylformamide (5 mL), triethylamine (0.5 mL), piperidine-1,4-dicarboxylic acid mono-*tert*-butyl ester (536 mg, 2.34 mmol, 1.19 equiv), and HATU (2.235 g, 5.88 mmol, 3.00 equiv) was added to the flask. The resulting solution was stirred for 2 h at 15 °C. The solids were filtered out. The resulting mixture was concentrated under vacuum. The residue was applied onto a silica gel column with ethyl acetate/petroleum ether (1:5 to 1:3). This step resulted in 705 mg (77%) of 4-(3-methylsulfanyl-1-phenyl-1*H*-indazol-6-ylcarbamoyl)piperidine-1-carboxylic acid *tert*-butyl ester as a white solid.

A solution of 4-(3-methylsulfanyl-1-phenyl-1*H*-indazol-6-ylcarbamoyl)piperidine-1-carboxylic acid *tert*-butyl ester (112 mg, 0.24 mmol, 1.00 equiv) in dichloromethane (4 mL) was added to a 50 mL three-necked round-bottom flask. The solution was bubbled with hydrogen chloride (gas) and stirred for 30 min at 0 °C. The solids were collected by filtration. This resulted in 46 mg (52%) of *N*-(3-(methylthio)-1-phenyl-1*H*-indazol-6-yl)piperidine-4-carboxamide hydrochloride as a yellow solid. MS (*m*/*z*): 367 $[M - HCl + H]^+$. ^{1}H NMR (400 MHz, MeOD, ppm): δ 10.082 (1H, s), 8.303 (1H, s), 7.628−7.609 (2H, d), 7.568−7.546 (1H, d), 7.498−7.461 (2H, t), 7.316−7.279 (1H, t), 7.151−7.130 (1H, d), 3.421−3.390 (2H, d), 3.044−3.983 (2H, t), 2.703−2.648 (1H, d), 2.597 (3H, s), 2.054−2.020 (2H, d), 1.947−1.857 (2H, m).

Enzyme Assay. An enzymatic assay was developed to measure ketohexokinase-catalyzed conversion of D-fructose to fructose 1-phosphate, using a LCMS instrument to quantify the product. This assay was adapted to a high throughput format with use of the BioTrove RapidFire.

Inhibition of ketohexokinase was analyzed with determination of IC_{50} values. Twelve-point titration curves were generated with concentration of a compound ranging from 511 to 0.5 μM or from 50 μM to 50 nM, depending on the tightness of binding. All kinetic reactions were conducted under steady-state conditions with use of 200 μM D-fructose, 100 μM ATP, and 2 nM ketohexokinase at 25 °C. The reaction time was 60 min, and initial velocity was measured. The assay was carried out in 384-well plate format and referenced to an internal standard compound included in every plate as a control.

Rat Pharmacokinetics. Male SPF Sprague−Dawley rats, each weighing ~250 g, were used in pharmacokinetics (PK) studies. Compound **6** as HCl salt was dissolved in 20% HPBCD for total dissolution, and the solution was adjusted to pH 4.5 with NaOH. Before dosing, the formulations were stored at room temperature, protected from light, and analyzed quantitatively; the stability of the formulations was checked at the day of dosing. Oral dose (10 mg/kg) administration was by gastric intubation, and blood samples were collected at 0.5, 1, 2, 4, 8, and 24 h after administration. Intravenous dose administration (2.0 mg/kg) was in the jugular vein, and blood samples were collected at 7 and 20 min and at 1, 2, 4, 8, and 24 h after administration.

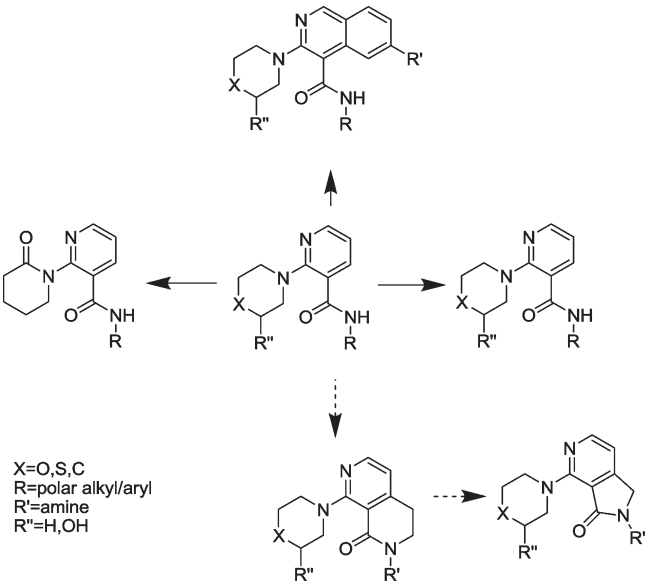


Figure 11. Tertiary (III_D) sublibrary design scheme. Although numerous fused ring analogues were designed for sublibrary III_D, the internal pyridine ring was held constant in order to maintain the hydrogen bond to conserved water.

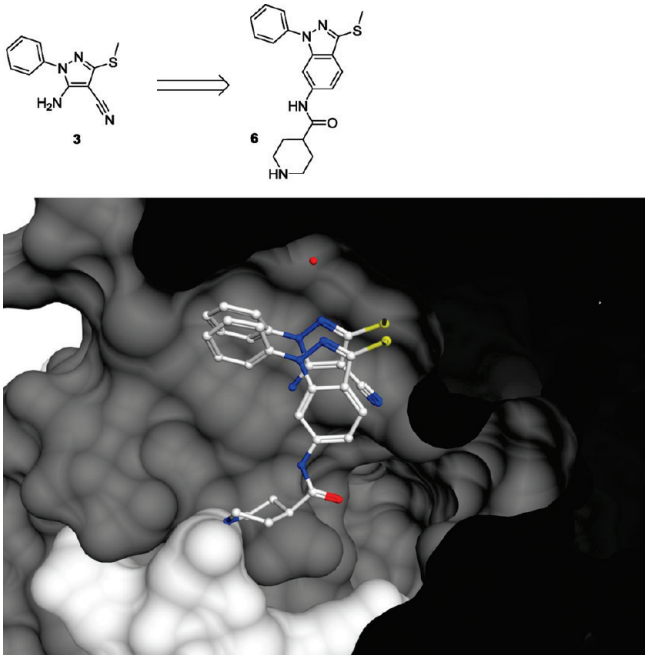


Figure 12. Evolution of hit compound **6**. Compared here are the primary library pyrazole (compound **3**) at 2.3 Å resolution and the tertiary library indazole lead compound (compound **6**) at 2.4 Å resolution.

Complete plasma profiles were taken from four different animals. A limited PK analysis was performed using WinNonlintm Professional (version 3.3, Pharsight, Mountain View, CA). Analysis was carried out on the plasma concentration time profiles obtained from each animal ($n=4$). Oral PK parameters observed were maximum plasma concentration (C_{max}), time to reach the maximum plasma concentration (T_{max}), plasma half-life ($t_{1/2}$), and exposure of the compound calculated by the area under the curve (AUC_{last} and AUC_{inf}). PK parameters after iv administration were volume of distribution, total plasma clearance, and area under the curve (AUC_{last} and AUC_{inf}).

In Vitro Kinetic Solubility Assay.[34,35] In a typical kinetic solubility assay, solid material was separated from the liquid phase after an incubation period by filtration or centrifugation. The supernatant was analyzed for compound in solution by a UV plate reader and/or HPLC analysis. Aliquots of DMSO stock solution (8 μL of 5 mM DMSO solution) were diluted into buffer solutions (0.01 N HCl; buffer at pH 4.0; buffer at pH 7.4) to a final DMSO concentration of 2% (representing a maximum possible solubility of 100 μM) and shaken for 4 h at room temperature. Insoluble material was removed by centrifugation at 3700 rpm for 10 min, and the concentration of compound in the supernatant was measured by UPLC/UV using a three-point reference calibration prepared by diluting the original DMSO stock solution.

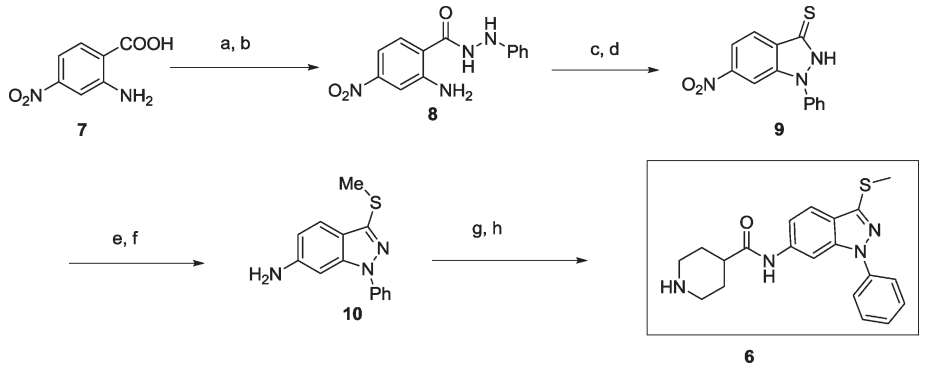
Caco-2 Permeability Assay.[36,37] These studies employed Caco-2 cells after a 21 day cell culture in 24-well Transwell plates. Test and reference compounds (propranolol and vinblastine) were prepared in Hanks' balanced salt solution containing 25 mM HEPES (pH 7.4) and added to either the apical or basolateral chambers of the Transwell plate assembly at 10 μM. Lucifer Yellow (LY) was added to the donor buffer in all wells to assess integrity of the cell layers by monitoring LY permeation. As LY could not freely permeate lipophilic barriers, a high degree of LY transport indicates poor integrity of the cell layer. After 1 h of incubation at 37 °C, aliquots were taken from both apical (A) and basal (B) chambers and added to acetonitrile containing analytical internal standard (carbamazepine) in a 96-well plate. Concentrations of compound in the samples were measured by high performance liquid chromatography/mass spectrometry (LC−MS/MS). Apparent permeability (P_{app}) values are calculated as given in references 36 and 37. The efflux ratios, as an indication of active efflux from the apical cell surface, were calculated using the ratio of $P_{app}(B\rightarrow A)/P_{app}(A\rightarrow B)$.

P450 Inhibition Assays.[38,39] Compounds were incubated in human liver microsomes over a concentration range appropriate for characterizing the IC_{50} for inhibition of the probe substrate by the test inhibitor. Positive control incubations in which a reference inhibitor was used were also included for each individual assay. The protein concentration, incubation time, and probe substrate concentration for each isoform/substrate pairing could vary. After a defined incubation time, the reaction was stopped and centrifuged and the supernatant prepared for analysis typically by LC/MS or fluorescence measurements. An appropriate internal standard could also be added. The measured concentration of the metabolite formed by the probe substrate was plotted against inhibitor concentration and an IC_{50} calculated by curve fitting.

Cloning and Expression.[40] Human ketohexokinase (EC 2.7.1.3) construct 6H.Tb.ketohexokinase 5-298 was cloned into pET28s and transformed into BL21(DE3) cells. Cells were grown in LB supplemented with 0.5% glucose and kanamycin (50 μg/mL) until the OD (A600) reached 0.8. The cultures were chilled to 15 °C on ice and induced with 1 mM IPTG and grown for an additional 16 h at 15 °C.

Purification[41] **and Crystallization.** All manipulations were performed at 4 °C. Bacterial cell pellet was thawed, suspended in lysis buffer [25 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10 mM imidazole, 10 mM β-mercaptoethanol, 1× complete protease inhibitor cocktail (Roche)], and lysed by sonication. Crude cell lysate was cleared by centrifugation at 1000*g* for 20 min. Cell lysate was further clarified by centrifugation at 100000*g* for 60 min. Supernatant containing 6His-tagged ketohexokinase was applied on 10 mL Ni-NTA agarose column (Qiagen) equilibrated in lysis buffer. The column was washed with 10 column volumes of lysis buffer and developed by 10 column volumes of lysis buffer supplemented with 250 mM imidazole. Protein-containing fractions were pooled and concentrated to 8 mg/mL using a centrifugal filtering device (Millipore, 10 kDa molecular-weight cutoffs). Protein homogeneity and identity

Scheme 1. Synthesis of Compound **6**[a]



[a] (a) *N*,*N*′-Carbonyldiimidazole, 90%; (b) $PhNHNH_2$, 38%; (C) $NaNO_2$, HCl, 53%; (d) P_2S_5, 19%; (e) Cs_2CO_3, MeI, 86%; (f) H_2, Pd/C, 89%; (g) piperidine-1,4-dicarboxylic acid mono-*tert*-butyl ester, HATU, 77%; (h) HCl, 52%.

were confirmed by SDS−PAGE (purity >98%) and mass spectrometry (34 078 Da). For crystallography, purified ketohexokinase was dialyzed overnight against 25 mM Tris, pH 8.0, 250 mM NaCl using Pierce Slide-A-Lyzer dialysis cassette with 10 kDa molecular-weight cutoff. Ketohexokinase crystallized at 295 K from a solution containing 17% PEG 8K, 100 mM Na citrate, pH 4.5, and 0.2 M ammonium sulfate.[29]

Crystal Soaking, Data Collection, and Structure Determination. Fragment plating was performed by first solubilizing each fragment cocktail in methanol to 20 mM/fragment. Once solubilized 6.5 μL of each fragment cocktail was pipetted into the well of an Imp@ct plate (Hampton Research) and allowed to evaporate at room temperature. Fragment cocktails contained a maximum of five fragments in each well. The fragment cocktails in each well were resuspended in 6.5 μL of 24% PEG 8K, 100 mM Na citrate, pH 4.5, and 0.2 M ammonium sulfate. Crystals were soaked in the fragment cocktails overnight. The next day the crystals were transferred to a cryoprotectant solution containing 24% PEG 8K, 100 mM Na citrate, pH 4.5, 0.2 M ammonium sulfate, and 15% glycerol, mounted, and quickly frozen by immersion in liquid nitrogen.

X-ray diffraction data for the various fragment structures were collected with resolutions between 2.3 and 3.0 Å at the IMCA-CAT beamlines ID-17 and BM-17 at the Argonne National Laboratory. Diffraction data were indexed, integrated, and scaled using d*trek.[42] The crystals belong to the $P2_12_12_1$ space group, with two ketohexokinase molecules in the asymmetric unit. The unit cell parameters for selected structure representatives are listed in Table 1. The structure was determined by molecular replacement using Auto-MR from the PHENIX suite.[43] One ketohexokinase molecule from the tetrameric structure of ketohexokinase-A (PDB code 2HLZ) was used as the search model.[44] Model building was performed using Coot.[45] Ligand building was performed using eLBOW.[43] Refinement and map calculations were carried out using PHENIX.[43] Refinement statistics for structure representatives are listed in Table 1. The atomic coordinates and structure factors for the ketohexokinase complexes with compounds **1**−**6** have been deposited in the Protein Data Bank under accession codes 3NBV, 3NBW, 3NC2, 3NCA, and 3NC9, respectively.[46]

Computational Fragment Docking. Various in-house ketohexokinase structures were used as protein coordinates in Glide docking studies. All fragment 3D coordinates were generated with an in-house implementation of a stochastic proximity embedding (SPE) algorithm,[47] followed by minimization using the MMFF94s force field.[48] Explicit hydrogens were added to the proteins using the "all-atom with no lone pair treatment" followed by restrained minimization, as implemented in PPREP utility of Glide.[49,50] Various Coulombic and van der Waals grids were generated with and without hydrogen bond constraints to the conserved active site water. Fragments considered for synthesis were allowed to flexibly dock into the receptors using standard precision (SP) or extra precision (XP) Glide. GlideScore, the default internal docking scoring function in Glide, evaluated the interaction energies of fragments with the Coulombic and van der Waals grids. The best scoring poses were minimized using the OPLS-AA force field[51] on precomputed van der Waals and electrostatic grids.

Acknowledgment. The authors thank Keli Dzordzorme, Shariff Bayoumy, Robyn William, Alexander Barnakov, Ludmila Barnakova, and Cynthia Milligan for ketohexokinase cloning, expression, purification, and crystallization. The authors also thank the ADME, Analytical Research, and Lead Generation Biology teams for bioanalysis support. The use of the IMCA-CAT beamlines 17-ID and 17-BM at the Advanced Photon Source is supported by the companies of the Industrial Macromolecular Crystallography Association through a contract with Hauptman-Woodward Medical Research Institute. Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

Supporting Information Available: List of structures and substructures used to filter compounds to produce candidates for the primary library. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discovery* **2007**, *6*, 211–219.
(2) Warr, W. A. Fragment-based drug discovery. *J. Comput.-Aided Mol. Des.* **2009**, *23*, 453–458.
(3) Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J. Recent developments in fragment-based drug discovery. *J. Med. Chem.* **2008**, *51*, 3661–3680.
(4) Hajduk, P. J. Puzzling through fragment-based drug design. *Nat. Chem. Biol.* **2006**, *2*, 658–659.
(5) Murray, C. W.; Callaghan, O.; Chessari, G.; Cleasby, A.; Congreve, M.; Frederickson, M.; Hartshorn, M. J.; McMenamin, R.; Patel, S.; Wallis, N. Application of fragment screening by X-ray crystallography to β-secretase. *J. Med. Chem.* **2007**, *50*, 1116–1123.
(6) Congreve, M.; Aharony, D.; Albert, J.; Callaghan, O.; Campbell, J.; Carr, R. A. E.; Chessari, G.; Cowan, S.; Edwards, P. D.; Frederickson, M.; McMenamin, R.; Murray, C. W.; Patel, S.; Wallis, N. Application of fragment screening by X-ray crystallography to the discovery of aminopyridines as inhibitors of β-secretase. *J. Med. Chem.* **2007**, *50*, 1124–1132.
(7) Murray, C. W.; Callaghan, O.; Chessari, G.; Cleasby, A.; Congreve, M.; Frederickson, M.; Hartshorn, M. J.; McMenamin, R.; Patel, S.; Wallis, N. Application of fragment screening by X-ray

crystallography to beta-secretase. *J. Med. Chem.* **2007**, *50*, 1116–1123.
(8) Nienaber, V. L.; Richardson, P. L.; Klighofer, V.; Bouska, J. J.; Giranda, V. L.; Greer, J. Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nat. Biotechnol.* **2000**, *18*, 1105–1108.
(9) Jhoti, H.; Cleasby, A.; Verdonk, M.; Williams, G. Fragment-based screening using X-ray crystallography and NMR spectroscopy. *Curr. Opin. Chem. Biol.* **2007**, *11*, 485–493.
(10) Poulsen, S.-A.; Kruppa, G. H. In situ fragment-based medicinal chemistry: screening by mass spectrometry. *Fragment-Based Drug Discovery* **2008**, 159–198.
(11) Ciulli, A.; Williams, G.; Smith, A. G.; Blundell, T. L.; Abell, C. Probing hot spots at protein−ligand binding sites: a fragment-based approach using biophysical methods. *J. Med. Chem.* **2006**, *49*, 4992–5000.
(12) Neumann, T.; Junker, H.; Schmidt, K.; Sekul, R. SPR-based fragment screening: advantages and applications. *Curr. Top. Med. Chem.* **2007**, *7*, 1630–1642.
(13) Miura, T. Fragment screening using surface plasmon resonance optical biosensor technology. *Yakugaku Zasshi* **2010**, *130*, 341–348.
(14) Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. The maximal affinity of ligands. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9997.
(15) Bembenek, S. D.; Tounge, B. A.; Reynolds, C. H. Ligand efficiency and fragment-based drug discovery. *Drug Discovery Today* **2009**, *14*, 278–283.
(16) Chessari, G.; Woodhead, A. J. From fragment to clinical candidate—a historical perspective. *Drug Discovery Today* **2009**, *14*, 668–675.
(17) de Kloe, G. E.; Bailey, D.; Leurs, R.; de Esch, I. J. Transforming fragments into candidates: small becomes big in medicinal chemistry. *Drug Discovery Today* **2009**, *14*, 630–646.
(18) Raushel, F. M.; Cleland, W. W. Bovine liver fructokinase: purification and kinetic properties. *Biochemistry* **1977**, *16*, 2169–2175.
(19) Gross, L. S.; Li, L.; Ford, E. S.; Liu, S. Increased consumption of refined carbohydrates and the epidemic of type 2 diabetes in the United States: an ecologic assessment. *Am. J. Clin. Nutr.* **2004**, *79*, 774–779.
(20) Basciano, H.; Federico, L.; Adeli, K. Fructose, insulin resistance, and metabolic dyslipidemia. *Nutr. Metab.* **2005**, *2*, 5.
(21) Elliott, S. S.; Keim, N. L.; Stern, J. S.; Teff, K.; Havel, P. J. Fructose, weight gain, and the insulin resistance syndrome. *Am. J. Clin. Nutr.* **2002**, *76*, 911–922.
(22) Bork, P.; Sander, C.; Valencia, A. Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. *Protein Sci.* **1993**, *2*, 31–40.
(23) Renze, B.; James, H. M.; Allan, M.; Conyers, R. A. The purification and properties of human liver ketohexokinase. *Biochem. J.* **1985**, *230*, 53–60.
(24) Muchmore, S. W.; Smith, R. A.; Stewart, A. O.; Cowart, M. D.; Gomtsyan, A.; Matulenko, M. A.; Yu, H.; Severin, J. M.; Bhagwat, S. S.; Lee, C.; Kowaluk, E. A.; Jarvis, M. F.; Jakob, C. L. Crystal structures of human adenosine kinase inhibitor complexes reveal two distinct binding modes. *J. Med. Chem.* **2006**, *49*, 6726–6731.
(25) Tang, L.; Li, M.; Cao, P.; Wang, F.; Chang, W.; Bach, S.; Reinhardt, J.; Ferandin, Y.; Galons, H.; Wan, Y.; Gray, N.; Meijer, L.; Jiang, T.; Liang, D. Crystal structure of pyridoxal kinase in complex with roscovitine and derivatives. *J. Biol. Chem.* **2005**, *280*, 31220–31229.
(26) Cheng, G.; Bennett, E. M.; Begley, T. P.; Ealick, S. E. Crystal structure of 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate kinase from *Salmonella typhimurium* at 2.3 Å resolution. *Structure* **2002**, *10*, 225–235.
(27) Ohshima, N.; Inagaki, E.; Yasuike, K.; Takio, K.; Tahirov, T. H. Structure of *Thermus thermophilus* 2-keto-3-deoxygluconate kinase: evidence for recognition of an open chain substrate. *J. Mol. Biol.* **2004**, *340*, 477–489.
(28) Hayward, B. E.; Bonthron, D. T. Structure and alternative splicing of the ketohexokinase gene. *Eur. J. Biochem.* **1998**, *257*, 85–91.
(29) Trinh, C. H.; Asipu, A.; Bonthron, D. T.; Phillips, S. E. V. Structures of alternatively spliced isoforms of human ketohexokinase. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2009**, *65*, 201–211.
(30) Sigrell, J. A.; Cameron, A. D.; Jones, T. A.; Mowbray, S. L. Structure of *Escherichia coli* ribokinase in complex with ribose and dinucleotide determined to 1.8 Å resolution: insights into a new family of kinase structures. *Structure* **1998**, *6*, 183–193.
(31) Mathews, I. I.; Erion, M. D.; Ealick, S. E. Structure of human adenosine kinase at 1.5 Å resolution. *Biochemistry* **1998**, *37*, 15607–15620.
(32) Sigrell, J. A.; Cameron, A. D.; Jones, T. A.; Mowbray, S. L. Purification, characterization, and crystallization of *Escherichia coli* ribokinase. *Protein Sci.* **1997**, *6*, 2474–2476.
(33) Symyx. Comprehensive Medicinal Chemistry Database, CMC. http://www.symyx.com/products/databases/bioactivity/cmc/index.jsp.
(34) Alsenz, J.; Kansy, M. High throughput solubility measurement in drug discovery and development. *Adv. Drug Delivery Rev.* **2007**, *59*, 546–567.
(35) Kerns, E. D.; Di, L.; Carter, G. T. In vitro solubility assays in drug discovery. *Curr. Drug Metab.* **2008**, *9*, 879−885.
(36) Irvine, J.; Takahashi, L.; Lockhart, K.; Cheong, J.; Tolan, J.; Selick, H.; Grove, J. MDCK cells: a tool for membrane permeability screening. *J. Pharm. Sci.* **1999**, *88*, 28–33.
(37) Artursson, P. Epithelial transport of drugs in cell culture. I: A model for studying the passive diffusion of drugs over intestinal absorbtive (Caco-2) cells. *J. Pharm. Sci.* **1990**, *79*, 476–482.
(38) Walsky, R.; Obach, R. Validated Assays for Human Cytochrome P450 Activities. *Drug Metab. Dispos.* **2004**, *32*, 647–660.
(39) Tucker, G.; Houston, J.; Huang, S. Optimizing drug development: strategies to assess metabolism/transporter interaction potential—towards a consensus. *Br. J. Clin. Pharmacol.* **2001**, *52*, 107–117.
(40) Bayoumy, S. S.; Dzordzorme, K. C. Unpublished results.
(41) Barnakov, A. N.; Barnakova, L. A. Unpublished results.
(42) Pflugrath, J. W. The finer things in X-ray diffraction data collection. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 1718–1725.
(43) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.; 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.
(44) Rabeh, W.; Tempel, W.; Nedyalkova, L.; Landry, R.; Arrowsmith, C.; Edwards, A.; Sundstrom, M.; Weigelt, J.; Bochkarev, A.; Park, H. Deposited in the Protein Data Bank Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org), PDB entry 2HLZ. Unpublished results.
(45) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr* **2004**, *60*, 2126–2132.
(46) Rutgers University, New Brunswick, NJ. Protein Data Bank, Research Collaboratory for Structural Bioinformatics. http://www.rcsb.org.
(47) Xu, H.; Izrailev, S.; Agrafiotis, D. K. Conformational sampling by self-organization. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1186–1191.
(48) Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490–519.
(49) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
(50) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.
(51) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.