SPECIAL
ISSUE

Discovery of a Potent BTK Inhibitor with a Novel Binding Mode by Using Parallel Selections with a DNA-Encoded Chemical Library

John W. Cuzzo,* Paolo A. Centrella, Diana Gikunju, Sevan Habeshian, Christopher D. Hupp, Anthony D. Keefe, Eric A. Sigel, Holly H. Soutter, Heather A. Thomson, Ying Zhang, and Matthew A. Clark^[a]

We have identified and characterized novel potent inhibitors of Bruton's tyrosine kinase (BTK) from a single DNA-encoded library of over 110 million compounds by using multiple parallel selection conditions, including variation in target concentration and addition of known binders to provide competition information. Distinct binding profiles were observed by comparing enrichments of library building block combinations under these conditions; one enriched only at high concentrations of BTK and was competitive with ATP, and another enriched at both high and low concentrations of BTK and was not compet-

itive with ATP. A compound representing the latter profile showed low nanomolar potency in biochemical and cellular BTK assays. Results from kinetic mechanism of action studies were consistent with the selection profiles. Analysis of the co-crystal structure of the most potent compound demonstrated a novel binding mode that revealed a new pocket in BTK. Our results demonstrate that profile-based selection strategies using DNA-encoded libraries form the basis of a new methodology to rapidly identify small molecule inhibitors with novel binding modes to clinically relevant targets.

Introduction

DNA-encoded chemical libraries have begun to emerge as valuable sources of new small molecule leads^[1] for a wide variety of biologically relevant disease targets, such as kinases,^[2] a phosphatase,^[3] metalloenzymes,^[4] hydrolases,^[5] and protein–protein interactions.^[6] In the most basic application of this technology, a soluble protein target with an affinity tag is mixed with a large library of small molecules, each covalently tagged with a unique DNA sequence. These encoded small molecule libraries can be made by split-and-pool synthesis with DNA-recorded synthetic histories.^[1d,2a,5,7] Hits are identified by affinity-mediated selection, followed by sequencing. Affinity-mediated selection is carried out by first forming complexes of the target and bound small molecules in solution.^[2a,8] After sufficient incubation time to allow components to reach equilibrium, the complexes are captured on an affinity resin by an affinity tag on the target, such as His₆, GST, or biotin. The complexes are washed under conditions that stringently remove unbound library members while retaining complexes. The bound molecules are then eluted by heat denaturation of the protein or by other means. High concentrations of carrier

DNA are often added to block interactions between the target and the encoding DNA in the library.^[2a] This selection process can then be repeated multiple times with fresh addition of protein target to further refine the population of small molecule binders. The final eluted fraction is then amplified, clustered, and sequenced by using high-throughput sequencing^[9] to deeply sample the identities of the small molecules comprising the output population. Highly enriched compounds are resynthesized without the encoding DNA and then tested for their activity in the appropriate target assay.

The ability to rapidly query very large compound libraries, containing 10⁸–10¹¹ compounds, by using miniaturized affinity-based selection methods requiring only microgram quantities of protein, allows multiple selection experiments with different targets or varying conditions to be explored in parallel. These conditions can be varied to focus the selection campaign on desired characteristics of the small molecule output. Altering target concentration, inclusion of selectivity targets, and addition of known binders to compete with library compounds are all conditions that can be run in parallel in a single selection campaign. As target concentration is the primary driver of complex formation at equilibrium in DNA-encoded library selections, altering the target concentration should significantly affect the population of selected molecules. High target concentration should retain low- and high-affinity library members, whereas low target concentration will only retain high-affinity binders. Inclusion of selectivity targets in the campaign can aid in the identification of selective compounds. Addition of saturating concentrations of known small molecule binders to the

[a] J. W. Cuzzo, P. A. Centrella, D. Gikunju, S. Habeshian, C. D. Hupp, A. D. Keefe, E. A. Sigel, H. H. Soutter, H. A. Thomson, Y. Zhang, M. A. Clark
X-Chem Pharmaceuticals
100 Beaver Street, Waltham, MA 02453 (USA)
E-mail: jcuzzo@x-chemrx.com

Supporting information and the ORCID identification numbers for the authors of this article can be found under <http://dx.doi.org/10.1002/cbic.201600573>.

This manuscript is part of a Special Issue on DNA-Encoded Chemistry.

target will compete away library molecules that bind to the same site. The outputs of these parallel selection conditions are evaluated so as to identify compounds with particular potency ranges, selectivity profiles, and binding modes directly from the selection output sequence data.

We have tested the concept of multiple parallel selection conditions to identify novel inhibitors of a clinically validated target, Bruton's tyrosine kinase (BTK). BTK is primarily expressed in B-cells and is activated by signaling pathways downstream of the B-cell receptor that recruit BTK to the plasma membrane and result in phosphorylation on Tyr551.^[10] Ibrutinib, a potent and selective BTK inhibitor that covalently binds to Cys481 in the active site of BTK, is an effective monotherapy for CLL, but recent reports indicate that mutations at Cys481 can lead to ibrutinib resistance in some CLL patients.^[11] BTK is also a target for inflammatory diseases, such as rheumatoid arthritis.^[12] Our selection conditions allowed us to identify a potent BTK inhibitor with a novel binding mode, as demonstrated by co-crystallography. This inhibitor might provide a useful template for new clinical candidates to treat both B-cell malignancies and autoimmune diseases.

Results and Discussion

DNA-encoded library construction

A DNA-encoded library was constructed by using standard split-and-pool methodology in three cycles of synthesis with three different sets of chemical building blocks, or synthons, to create a 110-million-compound DNA-encoded trisynthon library (Scheme 1). In the first cycle of library synthesis, cycle B,

Scheme 1. Synthetic scheme used to generate the 110-million-compound DNA-encoded chemical library used for BTK selections.

the bifunctional starting material was split into 300 wells and ligated to a unique oligonucleotide tag to encode each of the 300 Fmoc amino acids. The amino acids were installed by acylation reaction onto the primary amine on the oligonucleotide headpiece. After pooling, deprotection of N-Fmoc, and purification, the cycle B product was split into 157 wells, and encoded with 157 DNA tags prior to installing 157 formyl acids by acylation reaction (cycle C). The resulting cycle C product was pooled and purified prior to being split again into 2341 wells. The final diversification step, cycle D, was achieved by in-

stallation of 2341 amines by reductive amination reaction after encoding with 2341 unique oligonucleotide tags. Upon completion of the final step, the wells were pooled, and the combined material was purified by RP-HPLC. The reactivity of building blocks for each cycle was confirmed by reactions in a model system during library development prior to the synthesis of this library. We applied criteria, including molecular weight (MW), functional group compatibility, and percent conversion in a model reaction system (as assessed by LC-MS) for the inclusion of building blocks in library synthesis. Prior to use in selection, the library was separated into multiple aliquots, each of which was ligated to a closing tailpiece containing an encoding tag sequence that identified both library identity as well as experiment identity. This additional encoding permitted the pooling of multiple selection output samples for the purposes of sequencing.

BTK selection experiments

Four individual selection conditions were run side-by-side to distinguish library compounds with different binding modes and potentially different mechanisms of action. The BTK construct used in selection was unphosphorylated and tagged on its N terminus with 6xHis. Ni-NTA agarose was used for all selections. The selection conditions were as follows: no protein, 1 μ M BTK only, and 1 μ M BTK plus saturating amounts of either ATP or dasatinib, a known tight-binding small-molecule BTK inhibitor. Prior to selection, the protocol was evaluated by using an on-DNA form of dasatinib as a positive control. In this experiment, the positive control was enriched 3000-fold over one cycle of selection, as assessed by qPCR, confirming the suitability of the selection conditions for the discovery of BTK active site binders.

After the selection outputs were amplified, clustered, and sequenced by using Illumina/Solexa technology, analysis of the building-block combinations that were enriched under the various selection conditions produced two dramatically different profiles. The selection output for the 1 μ M BTK selection was shown in a pair of cube plots (Figures 1 A, B). Each cube plots the identities of each of a pair of contiguous building blocks on the horizontal axes and an enrichment metric as the vertical axis. Each point represents a fully elaborated instance (trisynthon); each enriched cluster represents a contiguous pair of building blocks (disynthon). Each point is sized by normalized count and randomly displaced by 0.5% to permit the viewer to see the extent to which the unrepresented building block in each enriched cluster is tolerated. These plots were used to identify compounds **1**, **2**, and **3** (Scheme 2), prior to their re-synthesis off-DNA. Profiles were then determined by measuring the extent of enrichment of each enriched building block combination in each of the parallel selections (Figures 1 C, D). The first profile, exemplified by **1**, showed enrichment against BTK alone but none in the absence of target, indicating that this enrichment was dependent on the presence of BTK and not the Ni-NTA agarose. The first profile was further characterized by an approximately 20-fold reduction in enrichment of **1** by the presence of either ATP or dasatinib. The second profile, as

Figure 1. A) The 110-million-compound three-cycle encoded small molecule library was selected against 1 μ M BTK, and the selection output was plotted on a three-dimensional scatter plot, with cycle D (amine) and cycle C (formyl acid) building block identities represented on the x- and y-axes, respectively. The enrichment of each CD disynthon pair is represented on the z-axis, and the point size represents counts. Trisynton instances are randomly displaced by 0.5% to indicate the extent to which variation at cycle B (amino acid) is tolerated. The enriched CD disynthon clusters that contain compounds 2 and 3 are indicated. B) As in (A) except that the cycle B (amino acid) and cycle C (formyl acid) building block identities are represented on the x- and y-axes, respectively, the enrichment of each BC disynthon pair is represented on the z-axis, the trisynton instances are randomly displaced by 0.5% to indicate the extent to which variation at cycle D (amine) is tolerated, and the enriched BC disynthon clusters that contain compounds 1 and 3 are indicated. C) Normalized frequencies of occurrence of building block combinations representing compounds 1 (■), 2 (■), and 3 (■) across the indicated BTK selection outputs. Histogram heights are normalized counts (total number of counts of each compound DNA sequence/total number of sequences $\times 10^6$; counts of < 2 for individual instances were presumed to be noise and disregarded) of the BC disynthon for 1, the CD disynthon for 2, and the BCD trisynton for 3 under each individual selection condition: BTK selected alone at a concentration of 1 μ M, 1 μ M BTK in the presence of 5 mM ATP, 1 μ M BTK in the presence of 5 μ M dasatinib. D) As in (C) except that two different BTK concentrations were used.

seen with both 2 and 3, showed enrichment in the presence of BTK and in the presence of saturating amounts of ATP. No enrichment was seen in the presence of dasatinib. These two strikingly different selection profiles strongly suggested different binding modes for these two sets of compounds.

In a separate selection experiment, the concentration of BTK was varied (Figure 1D). In one condition, the concentration of BTK was the same as the selection described above (1 μ M), and in a second condition, the BTK concentration was lowered to 50 nM. All of the compounds described above were again observed. Compound 3 was enriched at both 1 μ M and 50 nM BTK concentrations, whereas 1 and 2 were enriched only when

1 μ M BTK was used. As compound enrichment is dependent on both target concentration and the affinity of the selected molecule, these results predicted that these compounds would have significantly different affinities for BTK, with 3 expected to be more potent than 1 and 2. The similarity between the profiles of 1 and 2 in this experiment contrasted with the result from the previous selection experiment, thus indicating that these compounds might potentially represent three different combinations of both binding mode and potency range. The individual enrichment factors for each of the compounds as trisyntons were calculated for the selection experiment by using 1 μ M BTK and were 300-, 700-, and 2200-fold for 1, 2,

Scheme 2. Chemical structures of compounds 1–3 identified in the BTK selection outputs.

and 3, respectively. These enrichment factors were lower than what was seen for the on-DNA dasatinib enrichment experiment, presumably because the synthetic yields of these compounds in the library were less than 100%, but the observed enrichment factors were sufficient to identify these compounds as binders in the context of the screen.

The predicted structures of compounds 1, 2, and 3 fall into three structurally distinct categories (Scheme 2). Compound 1 contains a 7-azatryptophan building block close to the former DNA attachment point and an indazole that was introduced in the second cycle of the library synthesis. Compound 2 consists of an aminobenzhydrol group attached to a pyrazole containing formyl acid that allowed for linkage to DNA by aminophenyl acetic acid in the first cycle of the library synthesis. Compound 3 contains a tetrahydro-β-carboline building block, which allowed for linkage to DNA as well as a potential hinge-binding quinoxaline motif distal from the DNA attachment point. Despite the divergence of building blocks present in compounds 1 and 2, we speculated that the indazole and pyrazole motifs, the cycle C building blocks in each compound, formed hydrogen bond interactions with the hinge region of BTK, presumably similar to the hinge-binding quinoxaline in compound 3, which was later confirmed in an X-ray crystal structure of the BTK–compound 3 complex (Figure 3).

BTK assays and mechanism of action studies

Compounds 1–3 were synthesized off-DNA (for analytical data, see the Supporting Information), and were tested for BTK activity in a competition binding assay that measures the ability of

a compound to displace a known fluorescently labeled BTK small molecule binder. This assay was chosen, as BTK used in the selection experiments was in the unphosphorylated kinase-inactive state. All of the compounds were active in the competition binding assay (Table 1). Compound 1 showed

Table 1. Activity of compounds 1–3 in biochemical, cell-based, cytotoxicity, and Cyp assays.

Biochemical BTK binding	IC ₅₀ [μM]				Cyp inhibition 3A4, 2C9, 2D6
	Ramos	Jurkat	HWB	Huh-7	
1	2.5	n.d.	n.d.	n.d.	n.d.
2	0.21	0.13	2.8	7.6	> 50
3	0.00055	0.002	0.03	0.83	15.8
n.d.: not determined.					

activity in the low micromolar range, whereas the two compounds that were enriched in the presence of ATP during selection showed sub-micromolar activity. Compound 3, which was enriched at both high and low BTK concentrations during selection, showed the lowest IC₅₀ value (0.55 nM) in the binding assay. The two most potent compounds were also tested for inhibition of calcium release resulting from B-cell receptor or T-cell receptor stimulation in Ramos and Jurkat cells, respectively, and B-cell signaling in human whole blood assays. In the Ramos assay, which is dependent on cellular BTK activity, both compounds showed sub-micromolar activity. Compound 3 also showed sub-micromolar activity in the human whole blood B-cell signaling assay. Compound 3 was 15-fold more selective in the Ramos cell assay as compared to the Jurkat cell assay. The comparison between activity in Ramos (B-cell) and Jurkat (T-cell) assays was used as an initial assessment of selectivity for the BTK compounds, as T-cell activation is not dependent on BTK but is dependent on other tyrosine kinases, such as Lck^[13] and Zap-70.^[14] The 15-fold selectivity achieved by 3 is acceptable for a BTK hit compound but would need to be further improved to progress towards a clinical candidate. Testing for general cell-based cytotoxicity of 3 by using a HuH-7 cell line, an assay typically run for many small-molecule drug-discovery projects as a predictor of toxic potential in animal and human studies,^[15] showed toxicity only at a concentration over 7000-fold higher than its IC₅₀ value in the Ramos assay, which is an excellent place to start with an initial hit compound. Compound 3 was also tested for cytochrome P450 (Cyp) inhibition. Cyp enzymes such as Cyp3A4, Cyp2C9, and Cyp2D6 are largely responsible for drug metabolism in the liver, and lack of Cyp inhibition greatly reduces the likelihood of drug–drug interactions that could lead to potential adverse events in patients by inhibiting metabolism of other drugs.^[16] Compound 3 showed no inhibition of these three major cytochromes, further strengthening the case that this molecule, identified directly from an on-DNA library, is a good starting point for a clinical candidate. Taken together, the cellular potency and selectivity of this novel inhibitor suggest a unique binding mode that

could be relatively insensitive to high intracellular ATP concentrations, as suggested by the selection profile.

In order to learn more about the binding mode of these compounds, we performed mechanism of action studies by using a BTK kinase assay that allowed us to monitor the rate of substrate phosphorylation. First, the IC_{50} values for **1**, **2** and **3** were determined by using this assay and were found to be 3.6 μM , 64 nM, and 6 nM respectively. We also ran an initial ATP titration to measure the K_m value for ATP and V_{max} value for BTK in this assay. The ATP K_m value was found to be 52.7 ± 4 nM, and the V_{max} for the concentration of BTK used in this assay was 9088 ± 463 fluorescence units min^{-1} . ATP titrations were then conducted in the presence of various concentrations of compound. Increasing concentrations of **1** produced little change in the V_{max} of the enzyme but slightly more than a twofold shift in the K_m value of ATP from 50 to 114 nM (Figure 2A). These results are consistent with an ATP-competitive mode of action for **1**. Compound **2** showed both a slight increase in K_m value and a 1.6-fold reduction in V_{max} , typical of a mixed mode of inhibition (Figure 2B). Interestingly, increasing concentrations of **3** caused a nearly threefold reduction in the V_{max} of the kinase and showed very little change in the K_m value for ATP (Figure 2C). These changes indicate either a non-competitive mode of action or a very tight binding mode for **3** that prevented it from being displaced by ATP.^[17] Overall, the mechanism of action studies strongly suggested that **3** has a different binding mode than those exhibited by **1** and **2**.

Crystal structure of BTK kinase domain in complex with compound **3**

To learn more about the binding mode of the most potent compound identified by BTK selection, we determined the X-ray protein crystal structures of unphosphorylated BTK kinase domain in complex with **3** (1.33 Å resolution, PDB ID: 5U9D). Compound **3** formed interactions with key residues in the ATP binding pocket, glycine-rich loop, and DFG motif (Figure 3). The ATP binding pocket was occupied by the quinoxaline moiety of **3**, which formed several interactions with the hinge region. The backbone amide nitrogen of Met477 formed a hydrogen bond with the quinoxaline nitrogen in the *para* position relative to the linker, and the backbone carbonyl of Met477 formed a favorable interaction with the CH group adjacent to the quinoxaline nitrogen. The tetrahydro- β -carboline of **3** bound a hydrophobic pocket defined by the side chains of Phe413 (glycine-rich loop), Ile432 and Met437 (N-terminal to helix C), Ile472 (N-terminal to the hinge region), and Leu542 (activation loop). The indole NH of **3** formed a hydrogen bond with the side chain of Asp539 in the DFG motif. The carbonyl oxygen adjacent to the methylamide formed a direct hydrogen bond to the backbone amide nitrogen of Phe413 in the glycine-rich loop, and the side chain of Cys481 formed a hydrophobic interaction with the furan ring.

The overall structure of BTK in complex with **3** is very similar to previously reported structures—such as BTK in complex with a phenoxy-pyridone inhibitor (PDB ID: 3PJ2^[18])—in which helix C adopts the “out” conformation, thereby preventing the

Figure 2. Michaelis–Menten plots of the kinetics for compounds A) **1**, B) **2**, and C) **3**.

side chains of Lys430 and Glu445 from forming a salt bridge. When comparing the BTK–compound **3** co-crystal structure to a structure of BTK in complex with an amino-pyrazole inhibitor (PDB ID: 3PIZ^[18]), in which helix C adopts the “in” conformation, binding of compound **3** was observed to cause helix C to

Figure 3. X-ray crystal structure of BTK–3 complex in ball-and-stick representation (PDB ID: 5U9D). Compound **3** is displayed in green, adjacent amino acids in yellow, and selected water molecules in red. Black dashes indicate hydrogen bonds, and gray dashes indicate CH–carbonyl or van der Waals interactions.

tilt away from the active site and rotate. The tip of the glycine ring loop was disordered in the previously reported phenoxy-pyridone and amino-pyrazole co-crystal structures, but ordered in the BTK–3 complex. Comparison of the conformation of the glycine-rich loop in the BTK–3 co-crystal structure to that of the BTK complex with RN486, another sub-nanomolar BTK inhibitor (BTK $K_d = 0.3$ nM),^[19] which contains an ordered glycine-rich loop (PDB ID: 4RFZ^[20]), the methylamide moiety of compound **3** pushes residues Gly411, Gln412, and Phe413 away from the binding site by up to 3 Å (Figure 4). The main-chain shift of the glycine-rich loop results in the side chain of Gln412 being rotated out of the binding pocket and toward solvent and a 4 Å shift of the side chain of Phe413 away from helix C, thereby blocking access to the selectivity pocket around Tyr551.^[21]

Conclusion

Using a profile-based approach to DNA-encoded library screening with multiple parallel selections under different conditions, we identified three structurally distinct compounds with different potencies and binding modes that confirmed predictions made from the selection profile analysis of the original on-DNA library compounds. Compound **1** was enriched by BTK, but the enrichment was reduced to background levels by either ATP or dasatinib, consistent with the ATP-competitive mode of action, as demonstrated by its effects on the kinetics of BTK kinase activity. When the concentration of BTK was lowered to 50 nM in the selection, the enrichment of **1** decreased to background levels, suggesting that its affinity was in the micromolar range, as also confirmed by both competitive binding and kinase assays. The selection profile of **2** showed non-competitive behavior with ATP, but it was not enriched at the low concentration of BTK. These selection results also match what was observed in kinetic studies, where **2** showed a mixed mode of inhibition and was less potent in both the BTK competition binding assay and the kinase assay than **3**. The behavior of **3** in se-

lections was strikingly different than **1**. Compound **3** did not compete with ATP in the selection; this is consistent with the observation that increasing amounts of ATP in the presence of **3** did not restore BTK to its full activity, as seen in the mechanism of action experiments. Compound **3** was the only one of these three compounds that enriched at the lowest concentration of BTK used in selection, which predicted the nanomolar IC_{50} values observed in both the competitive binding and kinase assays. Interestingly, **3** was competed away by another tight binding inhibitor of BTK, dasatinib, which bound in the ATP binding pocket, suggesting that, although the compound appeared to be non-competitive with ATP, it still bound in or near the ATP pocket, and its apparent non-competitive behavior was more likely attributable to its tight binding to the target.

The BTK–compound **3** co-crystal structure revealed a novel binding pocket near the ATP binding site occupied by tetrahydro- β -carboline. This portion of the compound occupies a sub-pocket in which the side chain of Lys430 is typically observed (Figure 3). Despite forming only minimal interactions with the hinge region, compound **3** is a potent, ATP-competitive inhibitor. The unoccupied space near the gatekeeper residue, Thr474, allows the side chain of Lys430 to shift toward the ATP binding pocket, opening up the novel hydrophobic pocket into which **3** extends. Formation of this pocket is also aided by shifts in the glycine-rich loop induced by the methylamide moiety; these result in large rearrangements of the side chains of Gln412 and Phe413, opening up space for rearrangements of the side chains of Met437 and Leu542, which form part of the hydrophobic pocket (Figure 4). The novel hydrophobic pocket accessed by **3** might explain the tight binding of this compound. Overall, the discovery of the unique binding mode of **3** will be very useful for designing future compounds with increased potency and selectivity for BTK.

The discovery of three novel BTK inhibitor series that have significantly different potencies and binding modes from a single DNA-encoded small molecule library demonstrates the predictive power of this technology when it is utilized with multiple parallel selection conditions. This, in turn, opens up the potential for the rapid identification of small molecules with novel binding modes to potentially any therapeutic protein target. Careful design of selection conditions using prior knowledge about the enzymology of the target and any small or large molecule binders as competitors in selection can yield predictive information for how the compounds will behave once synthesized without the attached encoding DNA. Moving forward, it will be of interest to see how multiple selection conditions can be applied to rapidly identify novel small molecule binders to both well-established and new drug targets.

Experimental Section

Affinity-mediated selection: N-terminally His₆-tagged full-length human BTK was produced in baculovirus cells and purified based on a procedure similar to the one described below for protein crystallization, with the exception that the tag was not cleaved. This was combined with the DNA-encoded chemical library in solution,

Figure 4. X-ray crystal structures of BTK-inhibitor complexes in ribbon representation: A) BTK-inhibitor complex containing an ordered glycine-rich loop (PDB ID: 4RFZ); B) BTK-compound **3** complex. Inhibitors (green) and selected amino acid side chains are shown in ball-and-stick representation.

and affinity-mediated selection for BTK binders was initiated by multiple incubations in 60 μL of a model cytosolic incubation buffer containing HEPES (20 mM), potassium acetate (134 mM), sodium acetate (8 mM), sodium chloride (4 mM), magnesium acetate (0.8 mM), sheared salmon sperm DNA (1 mg mL⁻¹, Invitrogen), and Tween 20 (0.02%) at pH 7.2. Different incubation samples contained different combinations of BTK (1 μM), dasatinib (5 μM), or ATP (5 mM). The library concentration was 1 μM . After 1 h incubation, the mixture was applied to a bed of nickel affinity matrix (5 μL ; His-Select High-Flow nickel affinity gel; Phynexus, CA) with 20 passages, followed by washes with incubation buffer (8 $\times$ 200 μL). Retained library members were eluted by incubation with incubation buffer (60 μL) at 72 $^{\circ}\text{C}$ for 5 min, followed by further incubation with a second resin bed of His-Select High-Flow nickel affinity gel (5 μL) to remove any eluted protein. This entire selection protocol was repeated with the addition of fresh BTK and competitor where appropriate to half of the first-round eluate to regenerate a 1 μM BTK concentration. Encoding oligonucleotides present in the output of the second selection round were amplified by using Platinum PCR Supermix (Invitrogen) with denaturation at 94 $^{\circ}\text{C}$, annealing at 55 $^{\circ}\text{C}$, and extension at 72 $^{\circ}\text{C}$ for 24 cycles by using 5'- and 3'-primer oligonucleotides (each at 0.5 μM) that each incorporated sequences complementary to the tailpiece or headpiece along with Illumina READ1 or READ2 sequences required to support clustering and subsequent single-read 100-base pair sequencing on an Illumina HiSeq 2500. Sequencing was also performed for PCR-amplified samples of the naïve (unselected) library and the output of a no-target selection performed in the absence of BTK. A total of 17.9 million sequence reads were determined

from the four combined samples. Sequence data were converted back into encoded chemical information computationally, and demographic and statistical information were calculated for individual building block combinations. In a parallel positive-control experiment, an on-DNA form of dasatinib was enriched by a factor of 3000-fold over one cycle of selection, as assessed by qPCR and the selection protocol described above.

Assays: The BTK time-resolved FRET-based competitive binding assay and cell-based BTK assays have been previously described.^[19] For the BTK Omnia kinetic assay used in the mechanism of action experiments, assay reagents were prepared from the Omnia Y peptide 5 kit as follows: kinase reaction buffer was diluted to 1X in dH₂O. A 10X substrate (100 μM) was prepared in dH₂O. ATP was prepared at 1 mM (10 $\times$) in dH₂O, and 2 mM (10 $\times$) DTT was prepared in dH₂O.

A master mix of kinase buffer (10 $\times$), peptide (10 $\times$), and DTT (10 $\times$), all at 2 μL , and dH₂O (4 μL) was prepared per well. Untagged BTK enzyme was run in a final concentration of 20 nM. Compounds at varying concentrations and ATP dilutions were prepared in 1 $\times$ kinase buffer. Final compound DMSO concentration was 1.25%.

In a black non-binding plate (Corning 3676), enzyme (5 μL) was added to compound/ATP mix (5 μL). The plate was incubated for 20 min at 30 $^{\circ}\text{C}$, then master mix (10 μL) was added into each well and mixed. The plate was incubated at 30 $^{\circ}\text{C}$, and fluorescence intensity readings at $\lambda_{\text{ex}} = 360$ nm, $\lambda_{\text{em}} = 485$ nm was collected at pre-determined intervals on the Tecan M1000. The data were fitted to the Michaelis–Menten equation in GraphPad Prism.

Crystallization, data collection, structure determination, and refinement: Wild-type human BTK(aa 389–659) was expressed with a TEV-cleavable GST tag in insect cells utilizing established protocols of the Bac-to-Bac expression system (Invitrogen). The purification procedure included affinity chromatography, tag cleavage, negative affinity chromatography, and size-exclusion chromatography steps. A ligand in the same chemical series as compounds **3** and **4** was used in co-crystallization experiments and was added to a final concentration of 2 mM from a 100 mM stock solution in DMSO. A grid screen around established conditions was set up, and crystallization hits were refined by systematically varying pH or precipitant concentrations, in addition to screening for additional compounds enhancing crystallization. Crystallization conditions contained poly(ethylene glycol)s (PEG3350/ PEG5000MME), (NH₄)₂SO₄, and MES buffer at pH 6.0–6.7. A BTK structure in complex with compound **3** was solved from crystals prepared by soaking. Diffraction data were measured at a temperature of 100 K. The crystallographic data were collected at the Swiss Light Source (SLS, Villigen, Switzerland) under cryogenic conditions. The structure was solved by molecular replacement with a previously solved structure of BTK as a search model. Subsequent model building and refinement were performed according to standard protocols with the software packages CCP4 and COOT.^[22]

Acknowledgement

The X-ray structure of BTK in complex with compound **3** was generated in the laboratory of Dr. Stephan Krapp at Proteros Biostructures GmbH, Martinsried, Germany.

Keywords: Bruton's tyrosine kinase • DNA-encoded library • kinase inhibitors • small molecules • X-ray crystallography

- [1] a) O. Eidam, A. L. Satz, *MedChemComm* **2016**, *7*, 1323–1331; b) R. M. Franzini, C. Randolph, *J. Med. Chem.* **2016**, *59*, 6629–6644; c) R. A. Goodnow, Jr., *A Handbook for DNA-Encoded Chemistry: Theory and Applications for Exploring Chemical Space and Drug Discovery*, John Wiley & Sons, **2014**; d) R. A. Goodnow, Jr., C. E. Dumelin, A. D. Keefe, *Nat. Rev. Drug Discovery* **2016**; DOI: 10.1038/nrd.2016.213; e) L. Mannocci, M. Leimbacher, M. Wichert, J. Scheuermann, D. Neri, *Chem. Commun. (Camb.)* **2011**, *47*, 12747–12753.
- [2] a) M. A. Clark, R. A. Acharya, C. C. Arico-Muendel, S. L. Belyanskaya, D. R. Benjamin, N. R. Carlson, P. A. Centrella, C. H. Chiu, S. P. Creaser, J. W. Cuozzo, C. P. Davie, Y. Ding, G. J. Franklin, K. D. Franzen, M. L. Geftter, S. P. Hale, N. J. Hansen, D. I. Israel, J. Jiang, M. J. Kavarana, M. S. Kelley, C. S. Kollmann, F. Li, K. Lind, S. Mataruse, P. F. Medeiros, J. A. Messer, P. Myers, H. O'Keefe, M. C. Oliff, C. E. Rise, A. L. Satz, S. R. Skinner, J. L. Svendsen, L. Tang, K. van Vloten, R. W. Wagner, G. Yao, B. Zhao, B. A. Morgan, *Nat. Chem. Biol.* **2009**, *5*, 647–654; b) H. Yang, P. F. Medeiros, K. Raha, P. Elkins, K. E. Lind, R. Lehr, N. D. Adams, J. L. Burgess, S. J. Schmidt, S. D. Knight, K. R. Auger, M. D. Schaber, G. J. Franklin, Y. Ding, J. L. DeLorey, P. A. Centrella, S. Mataruse, S. R. Skinner, M. A. Clark, J. W. Cuozzo, G. Evindar, *ACS Med. Chem. Lett.* **2015**, *6*, 531–536.
- [3] A. G. Gilmartin, T. H. Faltg, M. Richter, A. Groy, M. A. Seefeld, M. G. Darcy, X. Peng, K. Federowicz, J. Yang, S. Y. Zhang, E. Minthorn, J. P. Jaworski, M. Schaber, S. Martens, D. E. McNulty, R. H. Sinnamon, H. Zhang, R. B. Kirkpatrick, N. Nevins, G. Cui, B. Pietrak, E. Diaz, A. Jones, M. Brandt, B. Schwartz, D. A. Heerding, R. Kumar, *Nat. Chem. Biol.* **2014**, *10*, 181–187.
- [4] a) F. Buller, M. Steiner, K. Frey, D. Mirsof, J. Scheuermann, M. Kalisch, P. Buhlmann, C. T. Supuran, D. Neri, *ACS Chem. Biol.* **2011**, *6*, 336–344; b) H. Deng, H. O'Keefe, C. P. Davie, K. E. Lind, R. A. Acharya, G. J. Franklin, J. Larkin, R. Matico, M. Neeb, M. M. Thompson, T. Lohr, J. W. Gross, P. A. Centrella, G. K. O'Donovan, K. L. Bedard, K. van Vloten, S. Mataruse, S. R. Skinner, S. L. Belyanskaya, T. Y. Carpenter, T. W. Shearer, M. A. Clark, J. W. Cuozzo, C. C. Arico-Muendel, B. A. Morgan, *J. Med. Chem.* **2012**, *55*, 7061–7079; c) Y. Ding, H. O'Keefe, J. L. DeLorey, D. I. Israel, J. A. Messer, C. H. Chiu, S. R. Skinner, R. E. Matico, M. F. Murray-Thompson, F. Li, M. A. Clark, J. W. Cuozzo, C. C. Arico-Muendel, B. A. Morgan, *ACS Med. Chem. Lett.* **2015**, *6*, 888–893; d) F. Samain, T. Ekblad, G. Mikutis, N. Zhong, M. Zimmermann, A. Nauer, D. Bajic, W. Decurtins, J. Scheuermann, P. J. Brown, J. Hall, S. Graslund, H. Schuler, D. Neri, R. M. Franzini, *J. Med. Chem.* **2015**, *58*, 5143–5149.
- [5] A. Litovchick, C. E. Dumelin, S. Habeshian, D. Gikunju, M. A. Guie, P. Centrella, Y. Zhang, E. A. Sigel, J. W. Cuozzo, A. D. Keefe, M. A. Clark, *Sci. Rep.* **2015**, *5*, 10916.
- [6] a) C. S. Kollmann, X. Bai, C. H. Tsai, H. Yang, K. E. Lind, S. R. Skinner, Z. Zhu, D. I. Israel, J. W. Cuozzo, B. A. Morgan, K. Yuki, C. Xie, T. A. Springer, M. Shimaoka, G. Evindar, *Bioorg. Med. Chem.* **2014**, *22*, 2353–2365; b) M. Leimbacher, Y. Zhang, L. Mannocci, M. Stravs, T. Geppert, J. Scheuermann, G. Schneider, D. Neri, *Chem. Eur. J.* **2012**, *18*, 7729–7737; c) S. Melkko, L. Mannocci, C. E. Dumelin, A. Villa, R. Somavilla, Y. Zhang, M. G. Grutter, N. Keller, L. Jermutus, R. H. Jackson, J. Scheuermann, D. Neri, *ChemMedChem* **2010**, *5*, 584–590.
- [7] F. Buller, L. Mannocci, Y. Zhang, C. E. Dumelin, J. Scheuermann, D. Neri, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5926–5931.
- [8] S. Melkko, J. Scheuermann, C. E. Dumelin, D. Neri, *Nat. Biotechnol.* **2004**, *22*, 568–574.
- [9] L. Mannocci, Y. Zhang, J. Scheuermann, M. Leimbacher, G. De Bellis, E. Rizzi, C. Dumelin, S. Melkko, D. Neri, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17670–17675.
- [10] M. Spaargaren, M. F. de Rooij, A. P. Kater, E. Eldering, *Oncogene* **2015**, *34*, 2426–2436.
- [11] J. A. Woyach, R. R. Furman, T. M. Liu, H. G. Ozer, M. Zapatka, A. S. Ruppert, L. Xue, D. H. Li, S. M. Steggerda, M. Versele, S. S. Dave, J. Zhang, A. S. Yilmaz, S. M. Jaglowski, K. A. Blum, A. Lozanski, G. Lozanski, D. F. James, J. C. Barrientos, P. Lichter, S. Stilgenbauer, J. J. Buggy, B. Y. Chang, A. J. Johnson, J. C. Byrd, *N Engl. J. Med.* **2014**, *370*, 2286–2294.
- [12] J. A. Di Paolo, T. Huang, M. Balazs, J. Barbosa, K. H. Barck, B. J. Bravo, R. A. Carano, J. Darrow, D. R. Davies, L. E. DeForge, L. Diehl, R. Ferrando, S. L. Gallion, A. M. Giannetti, P. Gribbling, V. Hurez, S. G. Hymowitz, R. Jones, J. E. Kropf, W. P. Lee, P. M. Maciejewski, S. A. Mitchell, H. Rong, B. L. Staker, J. A. Whitney, S. Yeh, W. B. Young, C. Yu, J. Zhang, K. Reif, K. S. Currie, *Nat. Chem. Biol.* **2011**, *7*, 41–50.
- [13] S. Danielian, R. Fagard, A. Alcover, O. Acuto, S. Fischer, *Eur. J. Immunol.* **1989**, *19*, 2183–2189.
- [14] Y. J. Kim, F. Sekiya, B. Poulin, Y. S. Bae, S. G. Rhee, *Mol. Cell. Biol.* **2004**, *24*, 9986–9999.
- [15] L. Pohjala, P. Tammela, S. K. Samanta, J. Yli-Kauhaluoma, P. Vuorela, *Anal. Biochem.* **2007**, *362*, 221–228.
- [16] S. Kumar, R. Sharma, A. Roychowdhury, *Curr. Med. Chem.* **2012**, *19*, 3605–3621.
- [17] R. A. Copeland, *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*, 2nd ed., Wiley, New York, **2000**.
- [18] A. Kuglstatter, A. Wong, S. Tsing, S. W. Lee, Y. Lou, A. G. Villaseñor, J. M. Bradshaw, D. Shaw, J. W. Barnett, M. F. Browner, *Protein Sci.* **2011**, *20*, 428–436.
- [19] D. Xu, Y. Kim, J. Postelnek, M. D. Vu, D. Q. Hu, C. Liao, M. Bradshaw, J. Hsu, J. Zhang, A. Pashine, D. Srinivasan, J. Woods, A. Levin, A. O'Mahony, T. D. Owens, Y. Lou, R. J. Hill, S. Narula, J. DeMartino, J. S. Fine, *J. Pharmacol. Exp. Ther.* **2012**, *341*, 90–103.
- [20] Y. Lou, Z. K. Sweeney, A. Kuglstatter, D. Davis, D. M. Goldstein, X. Han, J. Hong, B. Kocer, R. K. Kondru, R. Litman, J. McIntosh, K. Sarma, J. Suh, J. Taygerly, T. D. Owens, *Bioorg. Med. Chem. Lett.* **2015**, *25*, 367–371.
- [21] Y. Lou, T. D. Owens, A. Kuglstatter, R. K. Kondru, D. M. Goldstein, *J. Med. Chem.* **2012**, *55*, 4539–4550.
- [22] a) N. Collaborative Computational Project, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1994**, *50*, 760–763; b) P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.

Manuscript received: October 28, 2016

Accepted Article published: January 5, 2017

Final Article published: February 8, 2017