
***ALGORITHM FOR COMPUTER-ASSISTED
MOLECULAR MODULARIZATION***

RELATED APPLICATIONS

This application claims the benefit of priority to U.S. Provisional Application No. 63/624,228, filed January 23, 2024; and U.S. Provisional Application No. 63/624,539, filed January 24, 2024.

STATEMENT OF GOVERNMENT SUPPORT

This invention was made with government support under 2019897 awarded by the National Science Foundation. The government has certain rights in the invention.

BACKGROUND

Establishing modular synthetic routes to target compounds is a process that, when performed traditionally, often requires extensive experimentation and laboratory resources. However, once a modular route to a drug is established, the drug may be freely derivatized, enabling the rapid discovery of structure-activity relationships and hastening the development of next-generation therapeutics. Computational approaches have the potential to greatly accelerate the modularization process through the rapid selection of optimized starting materials and reaction conditions. In view of the foregoing, there is an unmet need for the development of algorithmic molecular modularization methods.

SUMMARY OF THE INVENTION

In certain aspects, the present disclosure provides methods of forming a C-C, C-N, or C-O bond between two molecular building blocks, the method comprising:

providing a first molecular building block; wherein the first molecular building block comprises a first reactive moiety;

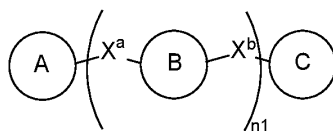
providing a second molecular building block; wherein the second molecular building block comprises a second reactive moiety and a third reactive moiety;

identifying suitable reaction conditions under which the first reactive moiety of the first molecular building block and the second reactive moiety of the second molecular building block will react to form the C-C, C-N, or C-O bond; and

combining the first molecular building block and the second molecular building block under the suitable reaction conditions;

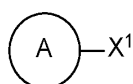
thereby forming a first coupled unit comprising the C-C, C-N, or C-O bond between the first molecular building block and the second molecular building block.

In further aspects, the present disclosure provides methods of making a compound of formula (I):



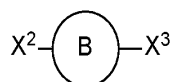
(I);

comprising combining a compound of formula (II):



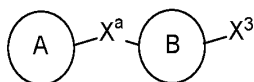
(II);

and a compound of formula (III):



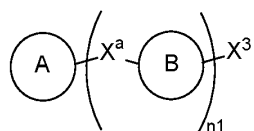
(III);

under a first set of suitable reaction conditions under which the compound of formula (II) and the compound of formula (III) will react to form a C-C, C-N, or C-O bond; thereby forming a compound of formula (IV):



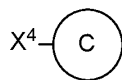
(IV);

optionally combining the compound of Formula (IV) with one or more additional compounds of Formula (III) to form a compound of Formula (IVa):



(IVa); and

combining the compound of formula (IV) or Formula (IVa) and a compound of formula (V):



(V);

under a second set of suitable reaction conditions under which the compound of formula (IV) or the compound of Formula (IVa) and the compound of formula (V) will react to form a C-C, C-N, or C-O bond; thereby forming the compound of formula (I);

wherein:

X¹ is selected from the group consisting of hydrogen, -B(OH)₂, -B(pin), -B(dan), -B(neop), -B(MIDA), -B(TIDA), -B(CbzTIDA), -NR¹H, -C(O)OR², -OH, -Cl, -Br, -I, -OMs, -OTs, and -OTf;

X², independently for each occurrence, is selected from the group consisting of -B(OH)₂, -B(pin), -B(dan), -B(neop), -B(MIDA), -B(TIDA), -B(CbzTIDA), -NR¹H, -C(O)OR², -OH, -Cl, -Br, -I, -OMs, -OTs, and -OTf;

X³, independently for each occurrence, is selected from the group consisting of -B(OH)₂, -B(pin), -B(dan), -B(neop), -B(MIDA), -B(TIDA), -B(CbzTIDA), -NR¹H, -C(O)OR², -OH, -Cl, -Br, -I, -OMs, -OTs, and -OTf;

X⁴ is selected from the group consisting of -N=C=O, -B(OH)₂, -B(pin), -B(dan), -B(neop), -B(MIDA), -B(TIDA), -B(CbzTIDA), -NR¹H, -C(O)OR², -OH, -Cl, -Br, -I, -OMs, -OTs, and -OTf;

R¹ is selected from the group consisting of hydrogen and a protecting group;

R² is selected from the group consisting of hydrogen and alkyl;

n1 is 1, 2, 3, 4, 5, 6, 7, 8, or 9;

A is selected from the group consisting of optionally substituted alkyl, optionally substituted alkenyl, optionally substituted aryl, optionally substituted heteroaryl, optionally substituted heterocyclyl, or optionally substituted cycloalkyl;

B, independently for each occurrence, is selected from the group consisting of optionally substituted aryl, optionally substituted heteroaryl, optionally substituted heterocyclyl, or optionally substituted cycloalkyl;

C is selected from the group consisting of optionally substituted alkyl, optionally substituted alkenyl, optionally substituted aryl, optionally substituted heteroaryl, optionally substituted heterocyclyl, or optionally substituted cycloalkyl;

X^a, independently for each occurrence, is selected from the group consisting of a bond, -C(O)NH-, -NHC(O)-, -O-, and -NR¹-; and

X^b, independently for each occurrence, is selected from the group consisting of a bond, -C(O)NH-, -NHC(O)-, -NH(CO)NH-, -O-, and -NR¹-.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows a partial dataset of FDA approved kinase inhibitors used to train the molecular modularization machine learning algorithm.

FIG. 2 is a schematic depicting the reaction toolkit used for molecular modularization, as well as describing the implementation of auxiliary reactions.

FIG. 3 shows an example of molecular modularization, highlighting the difference between modularization for traditional synthetic methods versus modularization for automated synthetic methods.

FIG. 4 shows a modular synthetic route that is compatible with automated synthetic methods.

FIG. 5 shows the various functional groups that, appended to a parent molecular moiety, define starting building blocks, propagation (also called extending, or linking) building blocks, or ending (also called terminal, or cap) building blocks.

FIG. 6A is a flowchart depicting the algorithmic logic for molecular modularization.

FIG. 6B shows a subset of the flowchart depicted in **FIG. 6A**, highlighting the first logical decision in the flowchart, as well as an example of the logical decision depicted in the subset.

FIG. 6C shows a subset of the flowchart depicted in **FIG. 6A**, highlighting the second logical decision in the flowchart, as well as an example of the logical decision depicted in the subset.

FIG. 6D shows a subset of the flowchart depicted in **FIG. 6A**, highlighting the third logical decision in the flowchart, as well as an example of the logical decision depicted in the subset.

FIG. 6E shows a subset of the flowchart depicted in **FIG. 6A**, highlighting the fourth logical decision in the flowchart, as well as an example of the logical decision depicted in the subset.

FIG. 6F shows a subset of the flowchart depicted in **FIG. 6A**, highlighting the fifth logical decision in the flowchart, as well as an example of the logical decision depicted in the subset.

FIG. 7 depicts a schematic demonstrating the transformation of molecular building blocks comprising competing and conflicting functional groups.

FIG. 8 shows examples of targets from the dataset that yield either no molecular building blocks, or large and/or complex molecular building blocks.

FIG. 9 depicts an additional set of rules for disconnecting large molecular building blocks, enabling access to the targets depicted in **FIG. 8**.

FIG. 10 shows the distribution of starting, extending, and cap building blocks based on their commercial availability.

FIG. 11 shows several exemplary structures of starting, extending, and cap building blocks, indicating their commercial availability.

FIG. 12 shows several reactions common to multiple target compounds, and depicts that 80 out of 153 training compounds are synthesizable in 3 generations, and that 93 out of 153 training compounds are synthesizable in 5 generations.

FIG. 13 is a scatter plot depicting the distribution and clustering of molecular building blocks grouped by reactive functional group using the ECFP6 algorithm.

FIG. 14 is a scatter plot depicting the distribution and clustering of molecular building blocks grouped by reactive functional group using the preferred ALLCH7 algorithm.

FIG. 15 shows representative structures of starting, extending, and terminal molecular building blocks, indicating their commercial availability.

FIG. 16 is a bar graph depicting, by number of carbon atoms in the target, the distribution of 1792 modularized FDA, Canada, and EU approved drugs. The leftmost and second leftmost bars depict drugs that have low molecular weights.

FIG. 17 depicts two bar graphs (top) and two scatter plots (bottom) collectively depicting the modularization of 1792 FDA, Canada, and EU approved drugs without incorporation of the sp^3 -Suzuki coupling reaction; the lefthand bar graph depicts the popularity of building blocks with no mass limit (number of drugs containing a given building block versus the number of building blocks); the righthand bar graph depicts the popularity of building blocks with fewer than 20 carbon atoms; the lefthand scatter plot depicts that 1193 out of 1792 drugs do not produce any building blocks; the righthand scatter plot depicts that 1198 out of 1792 drugs do not produce any building blocks.

FIG. 18 depicts two bar graphs collectively depicting the size distribution of the building blocks identified in **FIG. 17** as measured by the number of carbon atoms.

FIG. 19A shows several examples of building blocks as well as the parent drugs from which they are derived.

FIG. 19B shows several examples of building blocks as well as the parent drugs from which they are derived.

FIG. 19C shows several examples of building blocks as well as the parent drugs from which they are derived.

FIG. 19D shows several examples of building blocks as well as the parent drugs from which they are derived.

FIG. 20 depicts two bar graphs (top) and two scatter plots (bottom) collectively depicting the modularization of 1799 FDA, Canada, and EU approved drugs with incorporation of the sp^3 -Suzuki coupling reaction; the lefthand bar graph depicts the popularity of building blocks with no mass limit (number of drugs containing a given building block versus the number of building blocks); the righthand bar graph depicts the popularity of building blocks with fewer than 20 carbon atoms; the lefthand scatter plot depicts that 984 out of 1799 drugs do not produce any building blocks; the righthand scatter plot depicts that 946 out of 1799 drugs do not produce any building blocks. Incorporation of the sp^3 -Suzuki coupling reaction greatly improves modularization.

FIG. 21 depicts exemplary structures of drugs that are amenable to modularization; 348 FDA approved drugs can be modularized using approximately 750 unique building blocks.

FIG. 22 depicts exemplary modular synthetic routes for the drugs depicted in **FIG. 21**.

FIG. 23 depicts the modular synthetic route to the largest drug amenable to modularization: radotinib.

FIG. 24 depicts an exemplary computing node.

DETAILED DESCRIPTION OF THE INVENTION

This disclosure describes a computer algorithm that can perform *molecular modularization*. The input for the algorithm is a collection of organic small molecules that perform a useful function. The output is a collection of function-infused building blocks that can enable the discovery of new molecular functions via automated modular assembly under the guidance of artificial intelligence. A key aspect of the algorithm is the extraction of key building blocks that are both function-infused and compatible with modular automated

assembly. Another key aspect of the algorithm is real time analysis and optimization of all possible compatibility and incompatibility issues associated with the corresponding modular assembly process. In one example, the algorithm has been applied to the molecular modularization of all FDA approved kinase inhibitors. In a second example, the algorithm has been applied to all FDA approved drugs. In both cases, the algorithm yielded collections of building blocks that are well suited for the highly efficient discovery of new drug candidates via AI-guided automated modular synthesis.

The algorithm was designed to enable computer-driven disconnection modularization of small molecules into building blocks that are compatible with TIDA-based, automated, modular synthesis using a limited set of key reaction classes (typically Suzuki-Miyaura, Buchwald-Hartwig, and amide couplings). The block diagram of the algorithm is presented in **FIG. 6A**.

The modularization process begins by examining the molecule for competing reaction sites, such as peripheral bromides, amines, or carboxylic acids. This step is essential to ensure that the resulting building blocks promote structural growth in the intended direction. If such a site is detected, the functional group is either protected (e.g. Boc for amines, t-butyl esters for carboxylic acids etc.) or replaced with its bioisostere (e.g., bromides into methyl groups) that is compatible with the employed synthetic methodologies.

Once all potentially competing sites are either transformed or protected, the collection of modularizable reaction rules is applied to the molecules. This collection includes two types of chemical transformations: core methodologies and auxiliary reactions. The first type can be applied at any stage of the synthesis and typically includes, though is not limited to, Suzuki-Miyaura reaction, Buchwald-Hartwig amination and amide coupling. In contrast, auxiliary reactions are restricted to specific stages and can only be performed either at the beginning of the synthesis (e.g., Buchwald C-O coupling) or at the end (e.g., reaction of isocyanates with amines). The reactions are applied iteratively in the retrosynthetic direction, producing smaller molecules from the initial target until no further transformations are possible.

After generating the blocks, their compatibility with the modular chemistry approach is evaluated based on the number, type, and combinations of functional groups. The scope of functional groups is defined by the currently used set of modularizable reactions and typically includes aryl bromides, carboxylic acids, amines (or their CbzBTIDA protected equivalents) and boronic acid (or their BTIDA-protected forms). Only mono- and bifunctional building

blocks are accepted, with the proviso that one of two groups is either an amine or a boronic acid to ensure compatibility with automated synthesis. For example, an aryl bromide would be considered a valid building block, and similarly, a molecule with both a bromide and a boronic acid would also be accepted. However, a dibromide would be rejected. Depending on the type and number of functional groups, building blocks are also categorized into three classes based on their role in the synthesis: starters, extenders and caps.

Optionally, for building blocks that include bromides, boronic acids or both functional groups, an operation called “group swapping” is performed. In this step, bromides are converted into boronic acids, and boronic acids are replaced with bromides. This operation is optional and helps to increase diversity in how blocks can be connected.

Once the blocks are categorized, they are scrutinized for potential incompatibilities. Since one of the goals of modularization is to provide a set of building blocks compatible with any combination of methodologies applied in the forward direction this step is crucial. During this process, different categories of incompatible groups are assigned based on the type of building block. For starter molecules, incompatibilities include conflicting groups not only for the reactions in which a given block can participate but also for all potential future reactions, including those at the later stages of extension and capping. In contrast, caps, used only in the final iteration, have the least restrictive set of incompatibilities. If a conflicting group is detected for a particular building block, the algorithm checks whether this incompatibility can be transformed into a non-conflicting, bioisosteric equivalent.

Additionally, after performing all these operations the set of building blocks can be limited to those with a carbon atom count below a user-specified limit. This ensures that the final set of building blocks does not encompass overly large molecules.

As an optional extension to the described operations, an additional logic has been introduced. It is designed to handle overly large molecules for which application of modularizable reactions yield no results, but fragments derived from such compounds may still serve as valid building blocks for the modular approach.

In order to address this issue, we created a set of “virtual disconnections”. These transformations break down molecules that cannot undergo the initial modularizable reactions and generate additional building blocks. Examples of these additional rules are presented in **FIG. 9**, along with information which resulting blocks are potentially valid for the modular chemistry approach. Once these disconnections are applied, the same procedures – including categorization, inspection, and remedying potential conflicting groups – are

performed. Additionally, parent molecules that undergo these supplementary cuts can be retained and, if they meet all necessary criteria, they may also be accepted as valid building blocks.

In certain aspects, the present disclosure provides methods of forming a C-C, C-N, or C-O bond between two molecular building blocks, the method comprising:

providing a first molecular building block; wherein the first molecular building block comprises a first reactive moiety;

providing a second molecular building block; wherein the second molecular building block comprises a second reactive moiety and a third reactive moiety;

identifying suitable reaction conditions under which the first reactive moiety of the first molecular building block and the second reactive moiety of the second molecular building block will react to form the C-C, C-N, or C-O bond; and

combining the first molecular building block and the second molecular building block under the suitable reaction conditions;

thereby forming a first coupled unit comprising the C-C, C-N, or C-O bond between the first molecular building block and the second molecular building block.

In certain embodiments, the methods further comprise:

providing a third molecular building block; wherein the third molecular building block comprises a fourth reactive moiety;

identifying suitable reaction conditions under which the fourth reactive moiety of the third molecular building block and the first coupled unit will react to form a C-C, C-N, or C-O bond; and

combining the third molecular building block and the first coupled unit under the suitable reaction conditions;

thereby forming a second coupled unit comprising the C-C, C-N, or C-O bond between the third molecular building block and the first coupled unit.

In further embodiments, the first coupled unit comprises the third reactive moiety, and the third reactive moiety reacts with the fourth reactive moiety on the third molecular building block to form the second coupled unit. In yet further embodiments, the reaction between the first molecular building block and the second molecular building block is selected from the group consisting of Suzuki coupling, Buchwald C-N coupling, Buchwald C-O coupling, and amide coupling. In still further embodiments, the reaction between the first coupled unit and the third molecular building block is selected from the group consisting

of Suzuki coupling, Buchwald C-N coupling, Buchwald C-O coupling, amide coupling, and isocyanate-amine urea formation. In certain embodiments, the Suzuki coupling is selected from the group consisting of aryl-aryl Suzuki coupling, aryl-vinyl Suzuki coupling, and vinyl-vinyl Suzuki coupling. In further embodiments, the Buchwald C-N coupling is selected from the group consisting of aryl electrophile-aryl amine Buchwald C-N coupling, aryl electrophile-primary amine Buchwald C-N coupling, aryl electrophile-heterocyclic amine Buchwald C-N coupling, aryl electrophile-heteroaryl amine Buchwald C-N coupling, and aryl electrophile-secondary amine Buchwald C-N coupling. In yet further embodiments, the Buchwald C-O coupling is aryl electrophile-phenol Buchwald C-O coupling.

In still further embodiments, the methods further comprise:

identifying a first precursor molecule that does not comprise the first reactive moiety;
and

transforming the first precursor molecule to introduce the first reactive moiety;
thereby forming the first molecular building block.

In certain embodiments, the methods further comprise:

identifying a first precursor molecule comprising the first reactive moiety and one or more additional reactive moieties; and

protecting the one or more additional reactive moieties comprised by the first precursor molecule;

thereby forming the first molecular building block.

In further embodiments, the methods further comprise:

identifying a second precursor molecule that does not comprise the second reactive moiety and/or the third reactive moiety; and

transforming the second precursor molecule to introduce the second reactive moiety and/or the third reactive moiety;

thereby forming the second molecular building block.

In yet further embodiments, the methods further comprise:

identifying a second precursor molecule comprising the second reactive moiety, the third reactive moiety, and one or more additional reactive moieties; and

protecting the one or more additional reactive moieties comprised by the second precursor molecule;

thereby forming the second molecular building block.

In still further embodiments, the methods further comprise:

identifying a third precursor molecule that does not comprise the fourth reactive moiety; and

transforming the third precursor molecule to introduce the fourth reactive moiety;
thereby forming the third molecular building block.

In certain embodiments, the methods further comprise:

identifying a third precursor molecule comprising the fourth reactive moiety and one or more additional reactive moieties; and

protecting the one or more additional reactive moieties comprised by the third precursor molecule;

thereby forming the third molecular building block.

In further embodiments, the first reactive moiety, the second reactive moiety, the third reactive moiety, and the fourth reactive moiety are, independently at each occurrence, selected from the group consisting of an electrophilic moiety, a nucleophilic moiety, and a moiety which, upon reaction with a metal catalyst, becomes a nucleophilic moiety.

In yet further embodiments, the methods further comprise:

determining whether the first molecular building block and the second molecular building block are chemically compatible for the purpose of forming the C-C, C-N, or C-O bond; wherein the first molecular building block and the second molecular building block are chemically compatible if:

the first reactive moiety is a nucleophilic moiety or a moiety which, upon reaction with a metal catalyst, becomes a nucleophilic moiety, and the second reactive moiety is an electrophilic moiety; or

the first reactive moiety is an electrophilic moiety, and the second reactive moiety is a nucleophilic moiety or a moiety which, upon reaction with a metal catalyst, becomes a nucleophilic moiety.

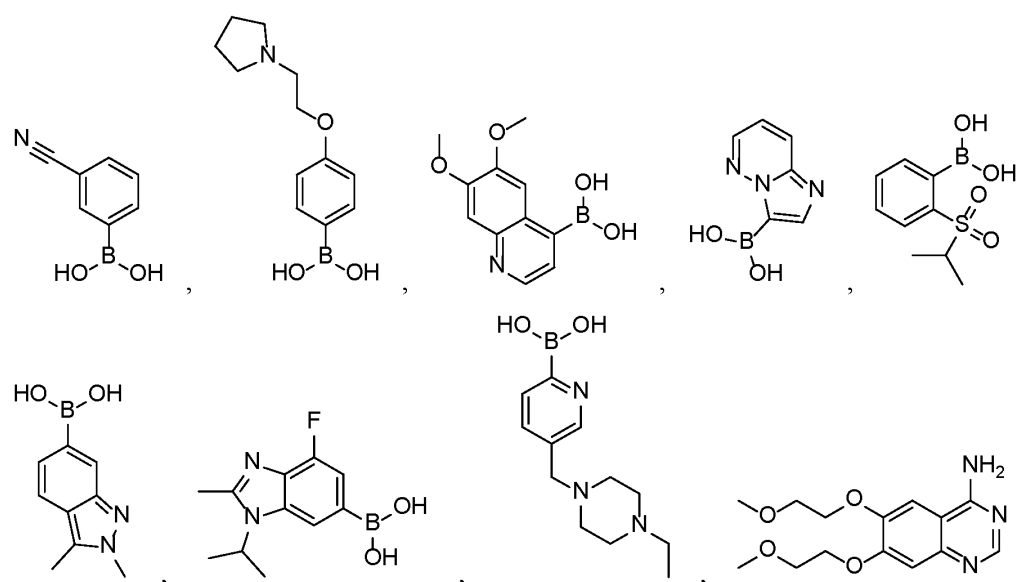
In still further embodiments, the methods further comprise:

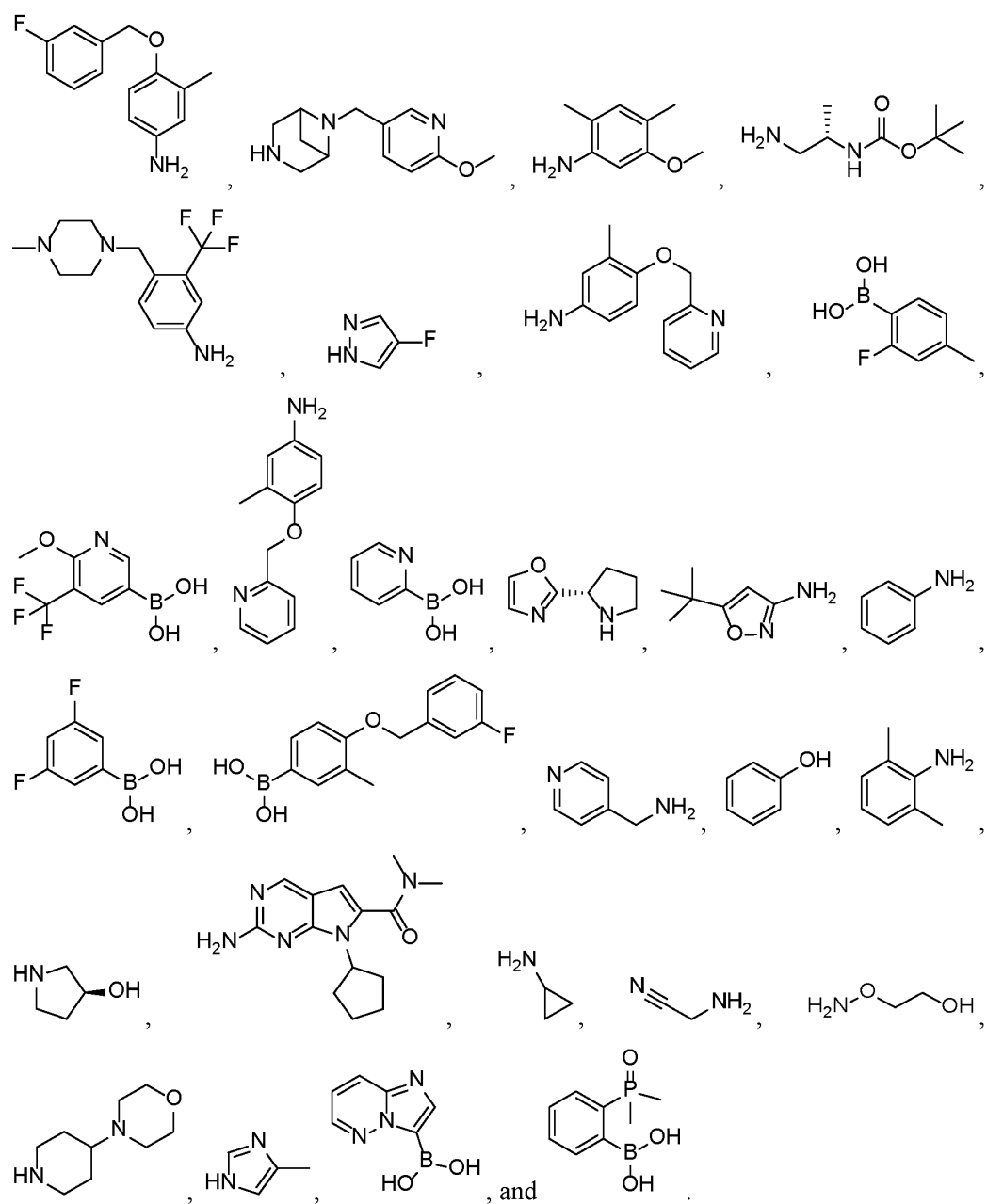
determining whether the third molecular building block and the first coupled unit are chemically compatible for the purpose of forming the C-C, C-N, or C-O bond; wherein the third molecular building block and the first coupled unit are chemically compatible if:

the third reactive moiety is a nucleophilic moiety or a moiety which, upon reaction with a metal catalyst, becomes a nucleophilic moiety, and the fourth reactive moiety is an electrophilic moiety; or

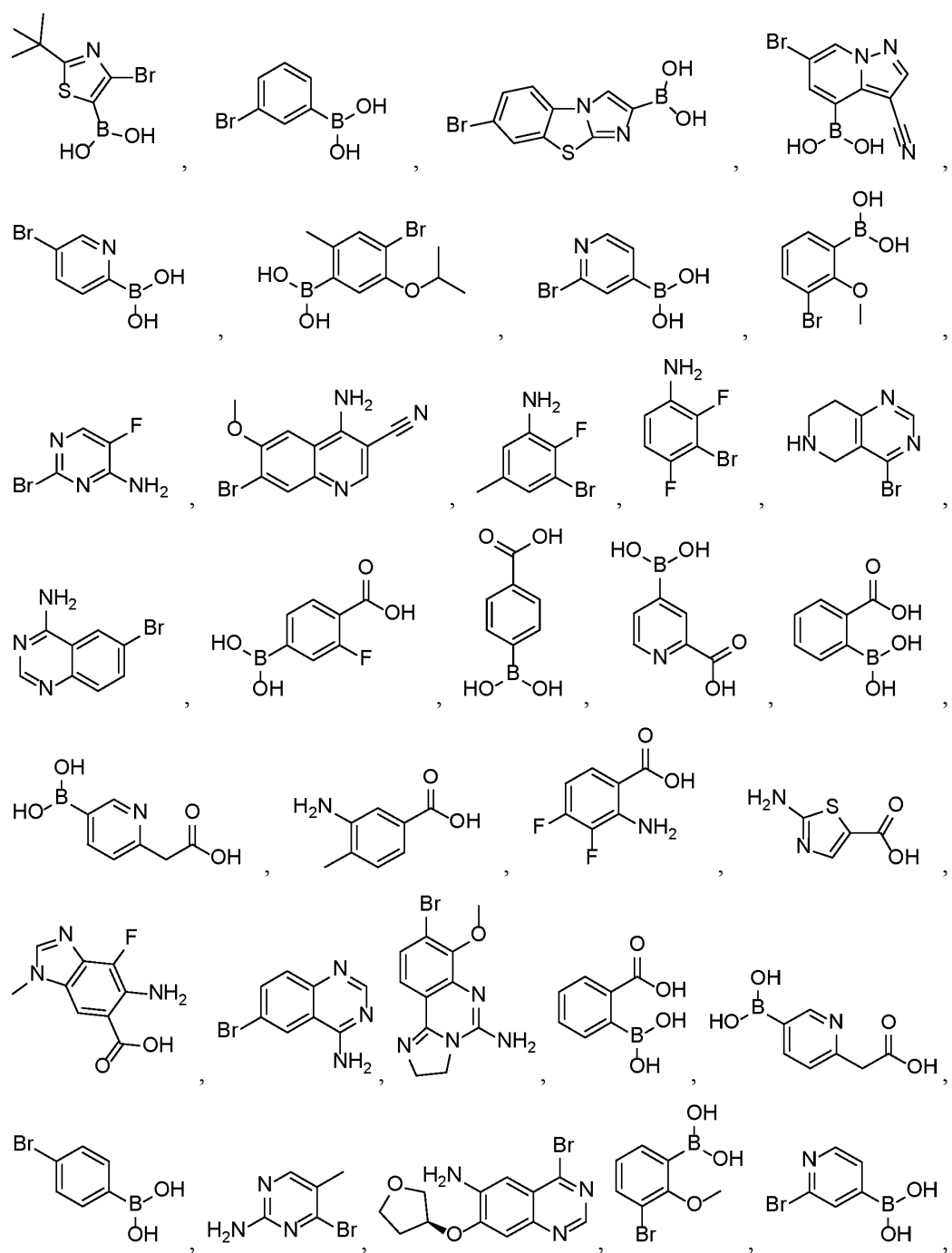
the third reactive moiety is an electrophilic moiety, and the fourth reactive moiety is a nucleophilic moiety or a moiety which, upon reaction with a metal catalyst, becomes a nucleophilic moiety.

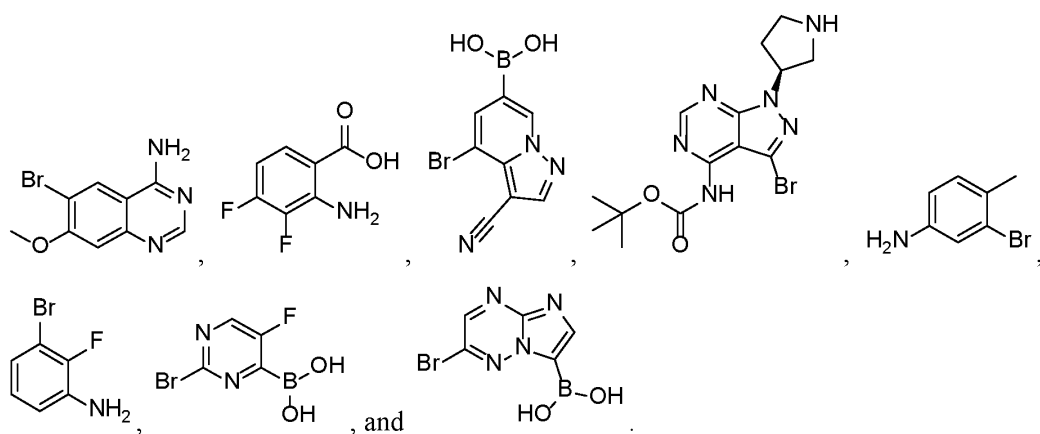
In certain embodiments, the first reactive moiety, the second reactive moiety, the third reactive moiety, or the fourth reactive moiety is, independently at each occurrence, a nucleophilic moiety or a moiety which, upon reaction with a metal catalyst, becomes a nucleophilic moiety. In further embodiments, the first reactive moiety, the second reactive moiety, the third reactive moiety, or the fourth reactive moiety is, independently at each occurrence, selected from the group consisting of: $-B(OH)_2$, $-B(pin)$, $-B(dan)$, $-B(neop)$, $-B(MIDA)$, $-B(TIDA)$, $-B(CbzTIDA)$, $-NR^1H$, and $-OH$; and R^1 is selected from the group consisting of hydrogen and a protecting group. In yet further embodiments, the first reactive moiety, the second reactive moiety, or the third reactive moiety is, independently at each occurrence, an electrophilic moiety. In still further embodiments, the first reactive moiety, the second reactive moiety, or the third reactive moiety is, independently at each occurrence, selected from the group consisting of: $-C(O)OR^2$, $-Cl$, $-Br$, $-I$, $-OMs$, $-OTs$, and $-OTf$; and R^2 is selected from the group consisting of hydrogen and alkyl. In certain embodiments, the fourth reactive moiety is an electrophilic moiety. In further embodiments, the fourth reactive moiety is selected from the group consisting of: $-C(O)OR^2$, $-Cl$, $-Br$, $-I$, $-OMs$, $-OTs$, $-OTf$, and $-N=C=O$; and R^2 is selected from the group consisting of hydrogen and alkyl. In yet further embodiments, the first molecular building block is selected from the group consisting of:



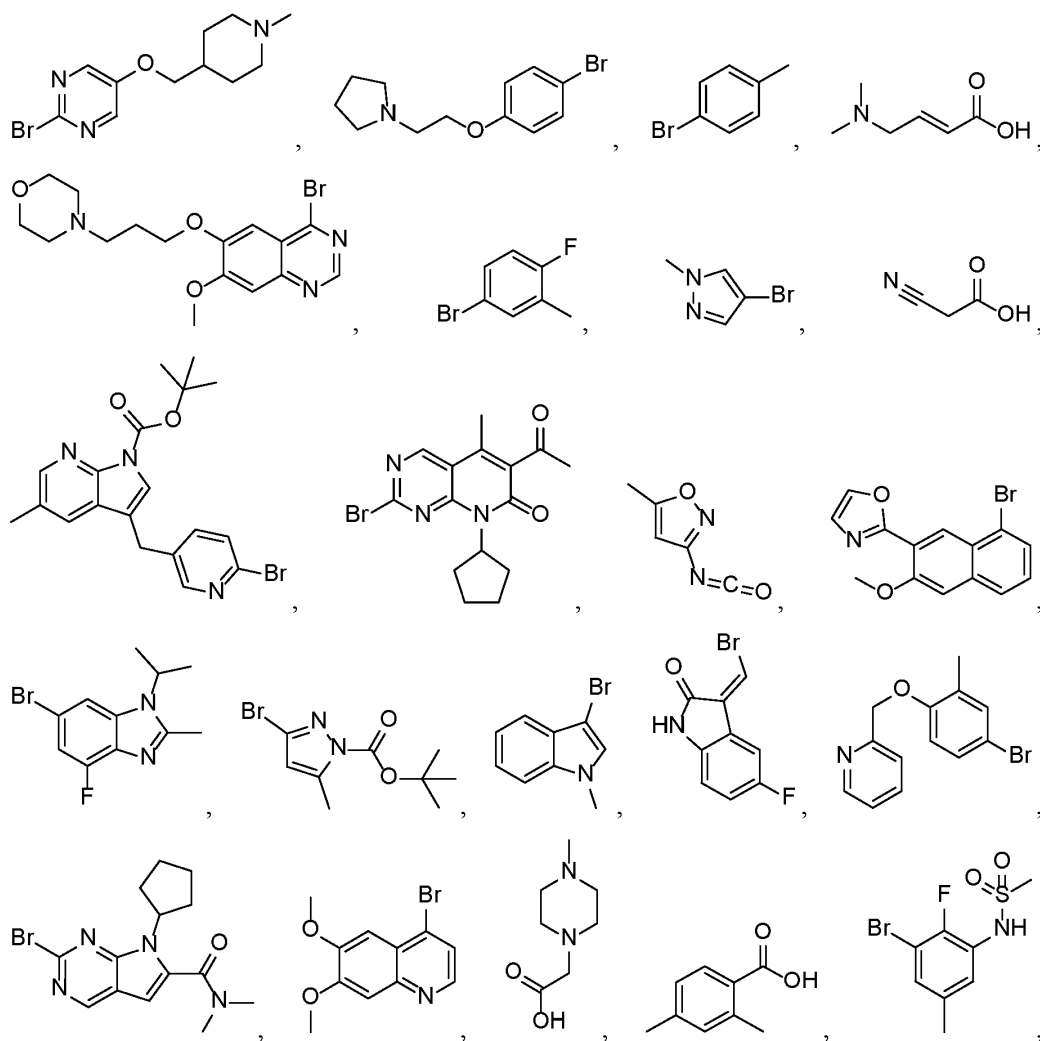


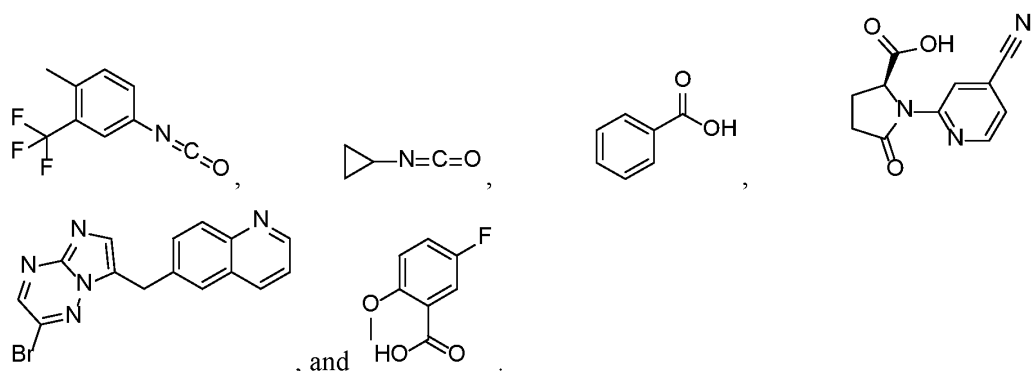
In still further embodiments, the second molecular building block is selected from the group consisting of:



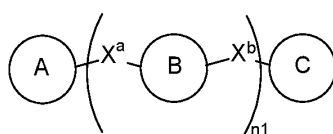


In certain embodiments, the third molecular building block is selected from the group consisting of:



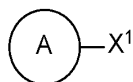


In further aspects, the present disclosure provides methods of making a compound of formula (I):



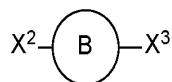
(I);

comprising combining a compound of formula (II):



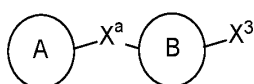
(II);

and a compound of formula (III):



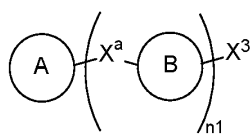
(III);

under a first set of suitable reaction conditions under which the compound of formula (II) and the compound of formula (III) will react to form a C-C, C-N, or C-O bond; thereby forming a compound of formula (IV):



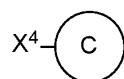
(IV);

optionally combining the compound of Formula (IV) with one or more additional compounds of Formula (III) to form a compound of Formula (IVa):



(IVa); and

combining the compound of formula (IV) or Formula (IVa) and a compound of formula (V):



(V);

under a second set of suitable reaction conditions under which the compound of formula (IV) or the compound of Formula (IVa) and the compound of formula (V) will react to form a C-C, C-N, or C-O bond; thereby forming the compound of formula (I);

wherein:

X^1 is selected from the group consisting of hydrogen, $-\text{B}(\text{OH})_2$, $-\text{B}(\text{pin})$, $-\text{B}(\text{dan})$, $-\text{B}(\text{neop})$, $-\text{B}(\text{MIDA})$, $-\text{B}(\text{TIDA})$, $-\text{B}(\text{CbzTIDA})$, $-\text{NR}^1\text{H}$, $-\text{C}(\text{O})\text{OR}^2$, $-\text{OH}$, $-\text{Cl}$, $-\text{Br}$, $-\text{I}$, $-\text{OMs}$, $-\text{OTs}$, and $-\text{OTf}$;

X^2 , independently for each occurrence, is selected from the group consisting of $-\text{B}(\text{OH})_2$, $-\text{B}(\text{pin})$, $-\text{B}(\text{dan})$, $-\text{B}(\text{neop})$, $-\text{B}(\text{MIDA})$, $-\text{B}(\text{TIDA})$, $-\text{B}(\text{CbzTIDA})$, $-\text{NR}^1\text{H}$, $-\text{C}(\text{O})\text{OR}^2$, $-\text{OH}$, $-\text{Cl}$, $-\text{Br}$, $-\text{I}$, $-\text{OMs}$, $-\text{OTs}$, and $-\text{OTf}$;

X^3 , independently for each occurrence, is selected from the group consisting of $-\text{B}(\text{OH})_2$, $-\text{B}(\text{pin})$, $-\text{B}(\text{dan})$, $-\text{B}(\text{neop})$, $-\text{B}(\text{MIDA})$, $-\text{B}(\text{TIDA})$, $-\text{B}(\text{CbzTIDA})$, $-\text{NR}^1\text{H}$, $-\text{C}(\text{O})\text{OR}^2$, $-\text{OH}$, $-\text{Cl}$, $-\text{Br}$, $-\text{I}$, $-\text{OMs}$, $-\text{OTs}$, and $-\text{OTf}$;

X^4 is selected from the group consisting of $-\text{N}=\text{C}=\text{O}$, $-\text{B}(\text{OH})_2$, $-\text{B}(\text{pin})$, $-\text{B}(\text{dan})$, $-\text{B}(\text{neop})$, $-\text{B}(\text{MIDA})$, $-\text{B}(\text{TIDA})$, $-\text{B}(\text{CbzTIDA})$, $-\text{NR}^1\text{H}$, $-\text{C}(\text{O})\text{OR}^2$, $-\text{OH}$, $-\text{Cl}$, $-\text{Br}$, $-\text{I}$, $-\text{OMs}$, $-\text{OTs}$, and $-\text{OTf}$;

R^1 is selected from the group consisting of hydrogen and a protecting group;

R^2 is selected from the group consisting of hydrogen and alkyl;

$n1$ is 1, 2, 3, 4, 5, 6, 7, 8, or 9;

A is selected from the group consisting of optionally substituted alkyl, optionally substituted alkenyl, optionally substituted aryl, optionally substituted heteroaryl, optionally substituted heterocyclyl, or optionally substituted cycloalkyl;

B, independently for each occurrence, is selected from the group consisting of optionally substituted aryl, optionally substituted heteroaryl, optionally substituted heterocyclyl, or optionally substituted cycloalkyl;

C is selected from the group consisting of optionally substituted alkyl, optionally substituted alkenyl, optionally substituted aryl, optionally substituted heteroaryl, optionally substituted heterocyclyl, or optionally substituted cycloalkyl;

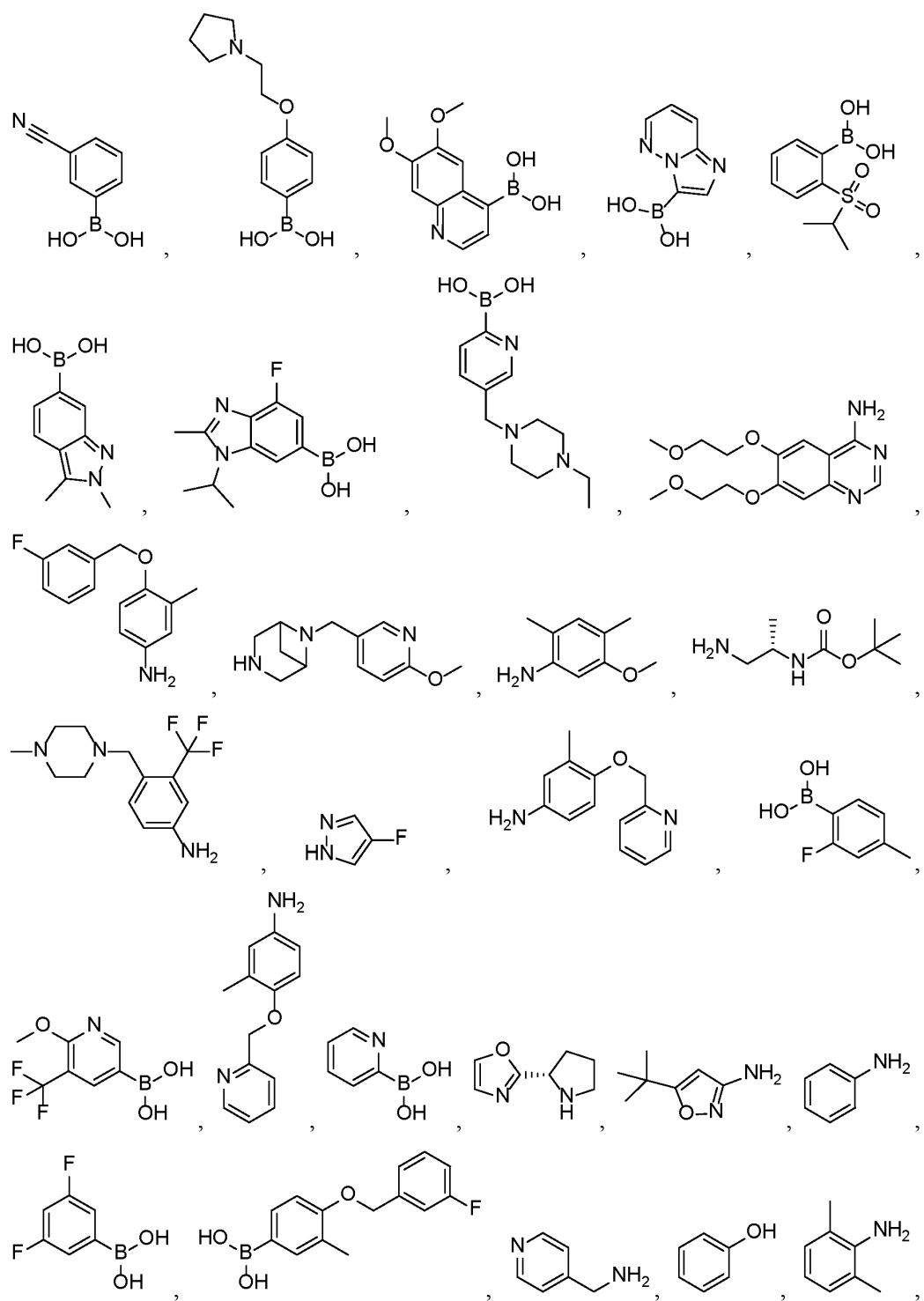
X^a, independently for each occurrence, is selected from the group consisting of a bond,

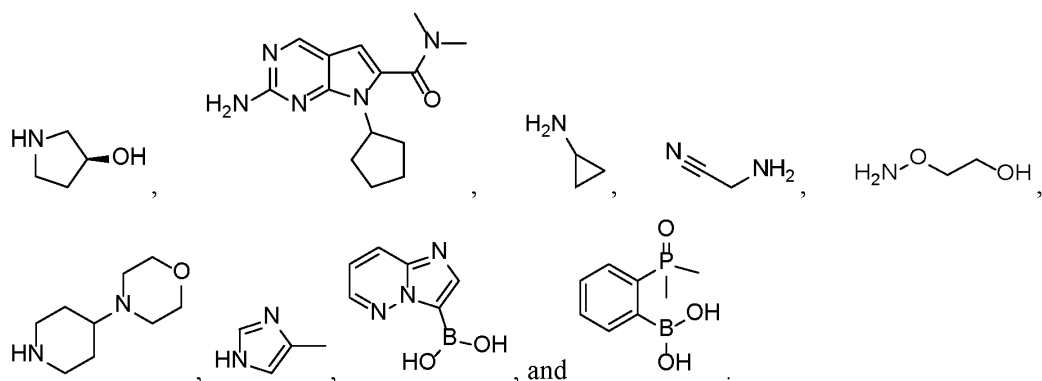
-C(O)NH-, -NHC(O)-, -O-, and -NR¹-; and

X^b, independently for each occurrence, is selected from the group consisting of a bond,

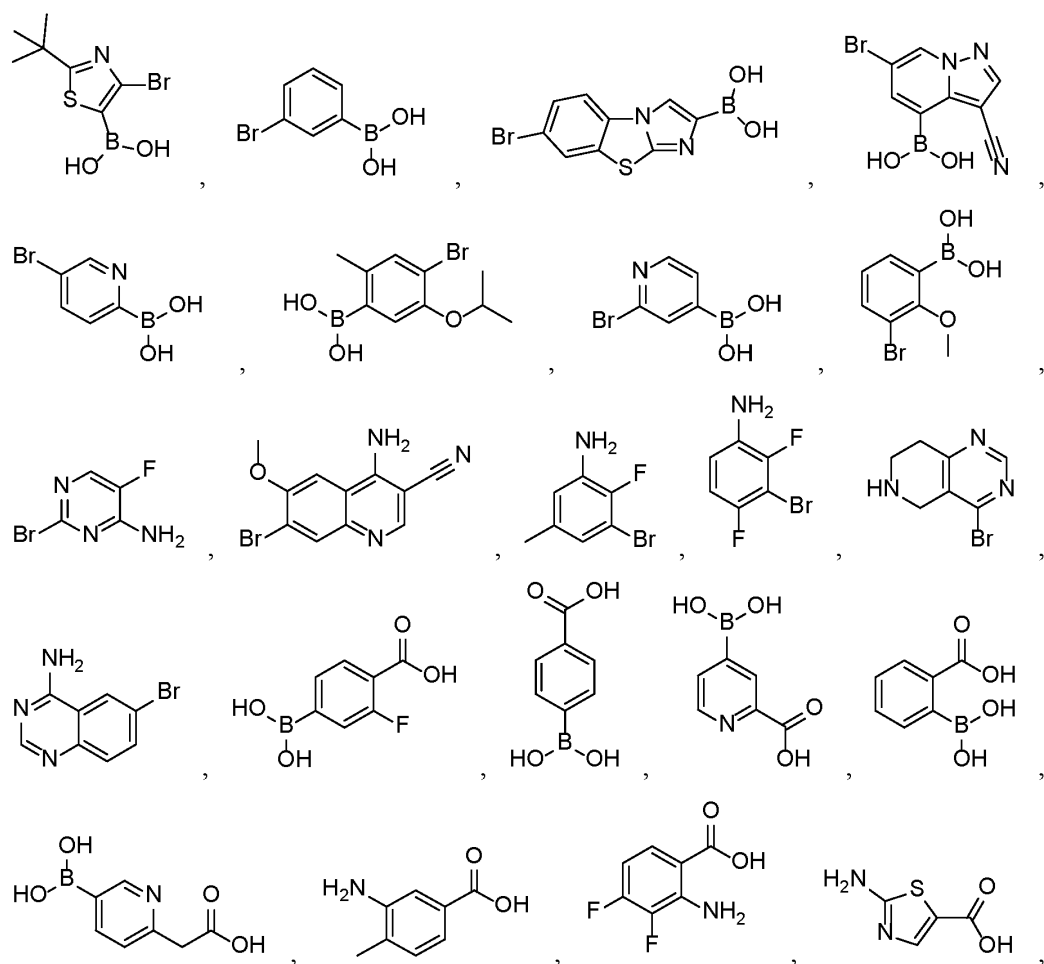
-C(O)NH-, -NHC(O)-, -NH(CO)NH-, -O-, and -NR¹-.

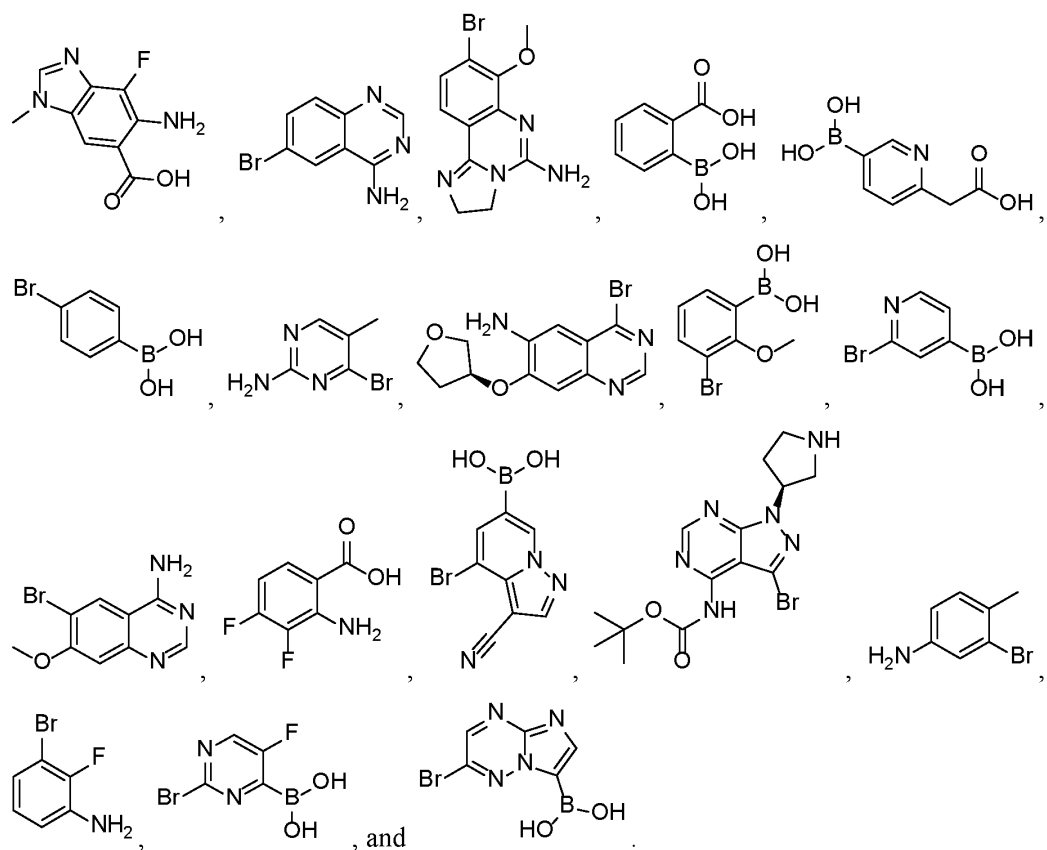
In certain embodiments, X¹ is hydrogen. In further embodiments, X¹ is -B(OH)₂. In yet further embodiments, X¹ is -B(pin). In still further embodiments, X¹ is -B(TIDA). In certain embodiments, X¹ is -B(CbzTIDA). In further embodiments, X¹ is -NR¹H. In yet further embodiments, X¹ is -OH. In still further embodiments, X² is -Br. In certain embodiments, X² is -C(O)OR². In further embodiments, X³ is -B(OH)₂. In yet further embodiments, X³ is -B(pin). In still further embodiments, X³ is -B(TIDA). In certain embodiments, X³ is -B(CbzTIDA). In further embodiments, X³ is -NR¹H. In yet further embodiments, X³ is -OH. In still further embodiments, X⁴ is -Br. In certain embodiments, X⁴ is -C(O)OR². In further embodiments, X⁴ is -N=C=O. In yet further embodiments, R¹ is hydrogen. In still further embodiments, R² is hydrogen. In certain embodiments, R² is alkyl. In further embodiments, n₁ is 1. In yet further embodiments, n₁ is 2. In still further embodiments, n₁ is 3. In certain embodiments, n₁ is 4. In further embodiments, n₁ is 5. In yet further embodiments, n₁ is 6. In still further embodiments, A is optionally substituted alkyl. In certain embodiments, A is optionally substituted alkenyl. In further embodiments, A is optionally substituted aryl. In yet further embodiments, A is optionally substituted heteroaryl. In still further embodiments, A is optionally substituted heterocyclyl. In certain embodiments, A is optionally substituted cycloalkyl. In further embodiments, A is selected from the group consisting of:



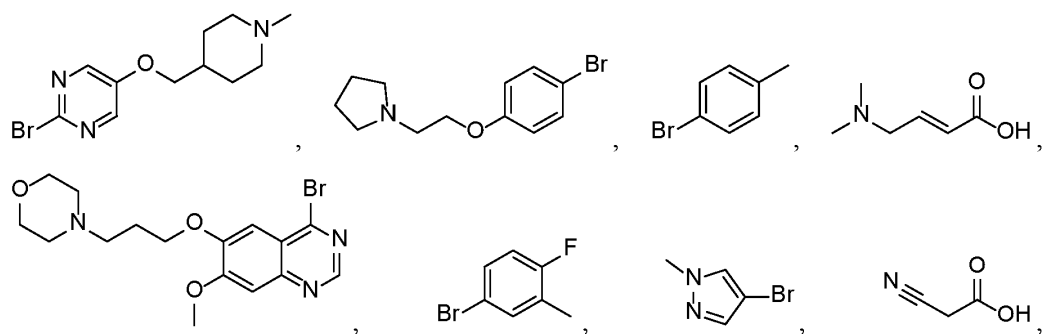


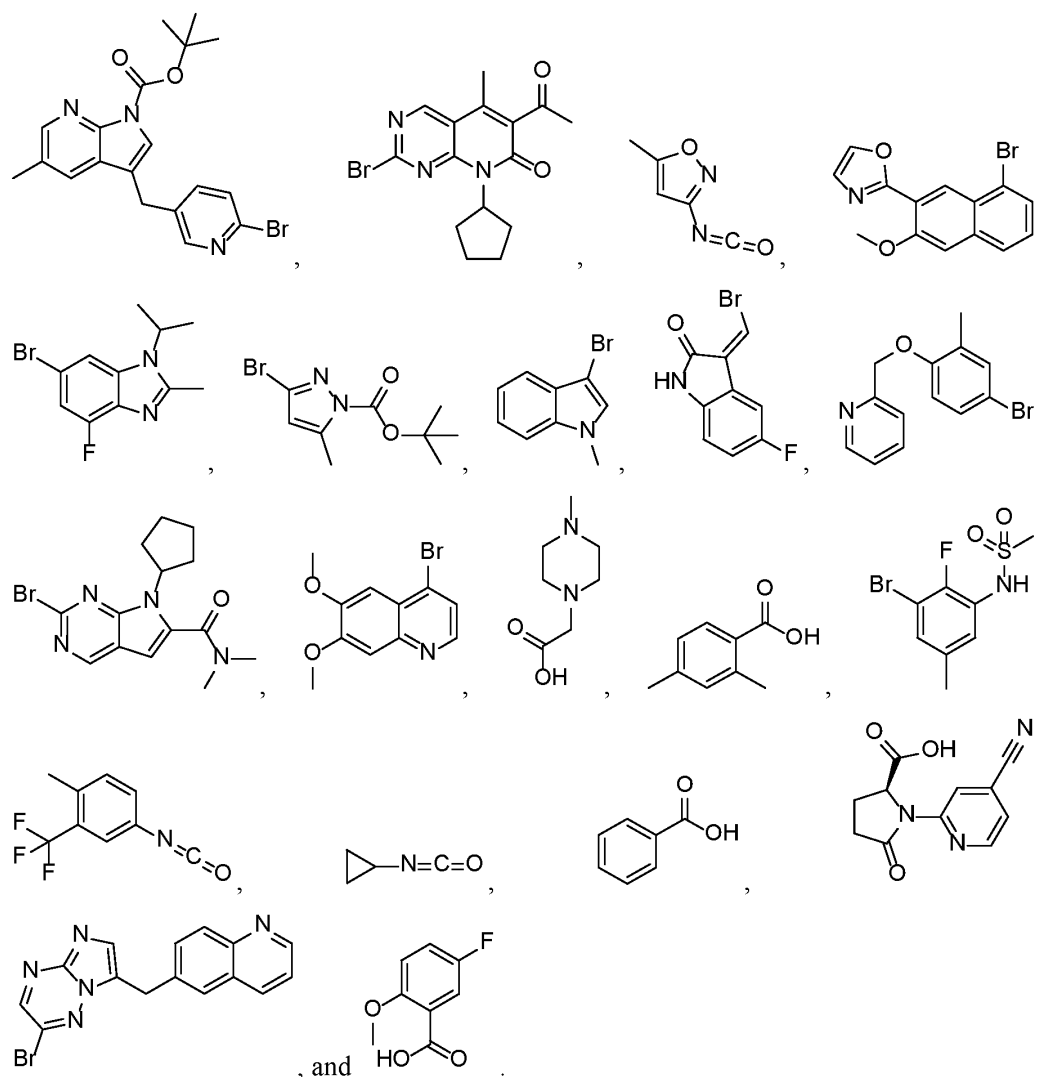
In yet further embodiments, B is optionally substituted aryl. In still further embodiments, B is optionally substituted heteroaryl. In certain embodiments, B is optionally substituted heterocyclyl. In further embodiments, B is optionally substituted cycloalkyl. In yet further embodiments, B is selected from the group consisting of:





In still further embodiments, C is optionally substituted alkyl. In certain embodiments, C is optionally substituted alkenyl. In further embodiments, C is optionally substituted aryl. In yet further embodiments, C is optionally substituted heteroaryl. In still further embodiments, C is optionally substituted heterocyclyl. In certain embodiments, C is optionally substituted cycloalkyl. In further embodiments, C is selected from the group consisting of:





In yet further embodiments, X^a is a bond. In still further embodiments, X^a is $-C(O)NH-$. In certain embodiments, X^a is $-NHC(O)-$. In further embodiments, X^a is $-NR^1-$. In yet further embodiments, X^a is $-O-$. In still further embodiments, X^b is a bond. In certain embodiments, X^b is $-C(O)NH-$. In further embodiments, X^b is $-NHC(O)-$. In yet further embodiments, X^b is $-NHC(O)NH-$. In still further embodiments, X^b is $-NR^1-$. In certain embodiments, X^b is $-O-$.

In certain embodiments, catalysts used in the methods of the present disclosure are so-called “Buchwald precatalyst.” Buchwald precatalysts are transition metal complexes used in, for example, cross-coupling reactions, and are of the general form $(L(M) G_1-G_4)$, wherein L is a ligand, M is a transition metal, such as palladium, and G indicates the

generational series, for example, (XPhos)Pd G3. A list of transition metal catalysts, including Buchwald precatalysts, used in this disclosure and their standard nomenclature follows below. In still further embodiments, the palladium catalyst is selected from the group consisting of (t-Bu-XPhos)Pd G3 [(2-dicyclohexylphosphino-2',4',6'-triisopropyl-1,1'-biphenyl)[2-(2'-amino-1,1'-biphenyl)] palladium(II) methanesulfonate], (t-Bu-XPhos)Pd G4 [(SP-4-3)-[2-di-tert-butylphosphino-2',4',6'-triisopropylbiphenyl[1,1'-biphenyl]-2-yl]phosphine](methanesulfonato-κO)[2'-(methylamino-κN)[1,1'-biphenyl]-2-yl-κC]palladium], (XPhos)Pd G3 [(2-dicyclohexylphosphino-2',4',6'-triisopropyl-1,1'-biphenyl)[2-(2'-amino-1,1'-biphenyl)] palladium(II) methanesulfonate], (XPhos)Pd G4 [(SP-4-3)-[dicyclohexyl[2',4',6'-tris(1-methylethyl)[1,1'-biphenyl]-2-yl]phosphine](methanesulfonato-κO)[2'-(methylamino-κN)[1,1'-biphenyl]-2-yl-κC]palladium], Pd(PPh₃)₄ [tetrakis(triphenylphosphine)palladium(0)], (RuPhos)Pd G3 [(2-dicyclohexylphosphino-2',6'-diisopropoxy-1,1'-biphenyl)[2-(2'-amino-1,1'-biphenyl)]palladium(II) methanesulfonate], (RuPhos)Pd G4 [[dicyclohexyl(2',6'-diisopropoxy-2-biphenyl)phosphine-κP](methanesulfonatato-κO)[2'-(methylamino-κN)-2-biphenyl-κC₂]palladium], (XantPhos)Pd G3 [[[(4,5-bis(diphenylphosphino)-9,9-dimethylxanthene)-2-(2'-amino-1,1'-biphenyl)]palladium(II) methanesulfonate], (XantPhos)Pd G4 [(SP-4-3)-[[5-(diphenylphosphino)-9,9-dimethyl-9H-xanthen-4-yl]diphenylphosphine-κP](methanesulfonato-κO)[2'-(methylamino-κN)[1,1'-biphenyl]-2-yl-κC]- palladium], (t-Bu-P)₃Pd G3 [(tris(2-methyl-2-propanyl)phosphino-2',6'-diisopropoxy-1,1'-biphenyl)[2-(2'-amino-1,1'-biphenyl)]palladium(II) methanesulfonate], (t-Bu-P)₃Pd G4 [(methanesulfonatato-κO)[2'-(methylamino-κN)-2-biphenyl-κC₂][tris(2-methyl-2-propanyl)phosphine]palladium], (Sphos)Pd G3 [(2-dicyclohexylphosphino-2',6'-dimethoxybiphenyl) [2-(2'-amino-1,1'-biphenyl)]palladium(II) methanesulfonate], (Sphos)Pd G4 [(methanesulfonato-κO)[2'-(methylamino)-2-biphenyl]palladium – dicyclohexyl(2',6'-dimethoxy-2-biphenyl)phosphine], Pd(OAc)₂ [palladium(II) acetate], (Pd)₂(DBA)₃ [tris(dibenzylideneacetone)dipalladium(0)], and Pd(PPh₃)₂(Cl)₂ [bis(triphenylphosphine)palladium(II) dichloride].

Computing Nodes

Referring now to **FIG. 24**, a schematic of an example of a computing node is shown. Computing node **10** is only one example of a suitable computing node and is not intended to suggest any limitation as to the scope of use or functionality of embodiments described

herein. Regardless, computing node **10** is capable of being implemented and/or performing any of the functionality set forth hereinabove.

In computing node **10** there is a computer system/server **12**, which is operational with numerous other general purpose or special purpose computing system environments or configurations. Examples of well-known computing systems, environments, and/or configurations that may be suitable for use with computer system/server **12** include, but are not limited to, personal computer systems, server computer systems, thin clients, thick clients, handheld or laptop devices, multiprocessor systems, microprocessor-based systems, set top boxes, programmable consumer electronics, network PCs, minicomputer systems, mainframe computer systems, and distributed cloud computing environments that include any of the above systems or devices, and the like.

Computer system/server **12** may be described in the general context of computer system-executable instructions, such as program modules, being executed by a computer system. Generally, program modules may include routines, programs, objects, components, logic, data structures, and so on that perform particular tasks or implement particular abstract data types. Computer system/server **12** may be practiced in distributed cloud computing environments where tasks are performed by remote processing devices that are linked through a communications network. In a distributed cloud computing environment, program modules may be located in both local and remote computer system storage media including memory storage devices.

As shown in **FIG. 24**, computer system/server **12** in computing node **10** is shown in the form of a general-purpose computing device. The components of computer system/server **12** may include, but are not limited to, one or more processors or processing units **16**, a system memory **28**, and a bus **18** that couples various system components including system memory **28** to processor **16**.

Bus **18** represents one or more of any of several types of bus structures, including a memory bus or memory controller, a peripheral bus, an accelerated graphics port, and a processor or local bus using any of a variety of bus architectures. By way of example, and not limitation, such architectures include Industry Standard Architecture (ISA) bus, Micro Channel Architecture (MCA) bus, Enhanced ISA (EISA) bus, Video Electronics Standards Association (VESA) local bus, Peripheral Component Interconnect (PCI) bus, Peripheral Component Interconnect Express (PCIe), and Advanced Microcontroller Bus Architecture (AMBA).

Computer system/server **12** typically includes a variety of computer system readable media. Such media may be any available media that is accessible by computer system/server **12**, and it includes both volatile and non-volatile media, removable and non-removable media.

System memory **28** can include computer system readable media in the form of volatile memory, such as random-access memory (RAM) **30** and/or cache memory **32**. Computer system/server **12** may further include other removable/non-removable, volatile/non-volatile computer system storage media. By way of example only, storage system **34** can be provided for reading from and writing to a non-removable, non-volatile magnetic media (not shown and typically called a "hard drive"). Although not shown, a magnetic disk drive for reading from and writing to a removable, non-volatile magnetic disk (*e.g.*, a "floppy disk"), and an optical disk drive for reading from or writing to a removable, non-volatile optical disk such as a CD-ROM, DVD-ROM or other optical media can be provided. In such instances, each can be connected to bus **18** by one or more data media interfaces. As will be further depicted and described below, memory **28** may include at least one program product having a set (*e.g.*, at least one) of program modules that are configured to carry out the functions of embodiments of the disclosure.

Program/utility **40**, having a set (at least one) of program modules **42**, may be stored in memory **28** by way of example, and not limitation, as well as an operating system, one or more application programs, other program modules, and program data. Each of the operating system, one or more application programs, other program modules, and program data or some combination thereof, may include an implementation of a networking environment. Program modules **42** generally carry out the functions and/or methodologies of embodiments as described herein.

Computer system/server **12** may also communicate with one or more external devices **14** such as a keyboard, a pointing device, a display **24**, etc.; one or more devices that enable a user to interact with computer system/server **12**; and/or any devices (*e.g.*, network card, modem, etc.) that enable computer system/server **12** to communicate with one or more other computing devices. Such communication can occur via Input/Output (I/O) interfaces **22**. Still yet, computer system/server **12** can communicate with one or more networks such as a local area network (LAN), a general wide area network (WAN), and/or a public network (*e.g.*, the Internet) via network adapter **20**. As depicted, network adapter **20** communicates with the other components of computer system/server **12** via bus **18**. Although not shown,

other hardware and/or software components could be used in conjunction with computer system/server 12. Examples, include, but are not limited to: microcode, device drivers, redundant processing units, external disk drive arrays, RAID systems, tape drives, and data archival storage systems, etc.

The present disclosure may be embodied as a system, a method, and/or a computer program product. The computer program product may include a computer readable storage medium (or media) having computer readable program instructions thereon for causing a processor to carry out aspects of the present disclosure.

The computer readable storage medium can be a tangible device that can retain and store instructions for use by an instruction execution device. The computer readable storage medium may be, for example, but is not limited to, an electronic storage device, a magnetic storage device, an optical storage device, an electromagnetic storage device, a semiconductor storage device, or any suitable combination of the foregoing. A non-exhaustive list of more specific examples of the computer readable storage medium includes the following: a portable computer diskette, a hard disk, a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM or Flash memory), a static random access memory (SRAM), a portable compact disc read-only memory (CD-ROM), a digital versatile disk (DVD), a memory stick, a floppy disk, a mechanically encoded device such as punch-cards or raised structures in a groove having instructions recorded thereon, and any suitable combination of the foregoing. A computer readable storage medium, as used herein, is not to be construed as being transitory signals per se, such as radio waves or other freely propagating electromagnetic waves, electromagnetic waves propagating through a waveguide or other transmission media (*e.g.*, light pulses passing through a fiber-optic cable), or electrical signals transmitted through a wire.

Computer readable program instructions described herein can be downloaded to respective computing/processing devices from a computer readable storage medium or to an external computer or external storage device via a network, for example, the Internet, a local area network, a wide area network and/or a wireless network. The network may comprise copper transmission cables, optical transmission fibers, wireless transmission, routers, firewalls, switches, gateway computers and/or edge servers. A network adapter card or network interface in each computing/processing device receives computer readable program instructions from the network and forwards the computer readable program instructions for

storage in a computer readable storage medium within the respective computing/processing device.

Computer readable program instructions for carrying out operations of the present disclosure may be assembler instructions, instruction-set-architecture (ISA) instructions, machine instructions, machine dependent instructions, microcode, firmware instructions, state-setting data, or either source code or object code written in any combination of one or more programming languages, including an object oriented programming language such as Smalltalk, C++ or the like, and conventional procedural programming languages, such as the “C” programming language or similar programming languages. The computer readable program instructions may execute entirely on the user’s computer, partly on the user’s computer, as a stand-alone software package, partly on the user’s computer and partly on a remote computer or entirely on the remote computer or server. In the latter scenario, the remote computer may be connected to the user’s computer through any type of network, including a local area network (LAN) or a wide area network (WAN), or the connection may be made to an external computer (for example, through the Internet using an Internet Service Provider). In some embodiments, electronic circuitry including, for example, programmable logic circuitry, field-programmable gate arrays (FPGA), or programmable logic arrays (PLA) may execute the computer readable program instructions by utilizing state information of the computer readable program instructions to personalize the electronic circuitry, in order to perform aspects of the present disclosure.

Aspects of the present disclosure are described herein with reference to flowchart illustrations and/or block diagrams of methods, apparatus (systems), and computer program products according to embodiments of the disclosure. It will be understood that each block of the flowchart illustrations and/or block diagrams, and combinations of blocks in the flowchart illustrations and/or block diagrams, can be implemented by computer readable program instructions.

These computer readable program instructions may be provided to a processor of a general purpose computer, special purpose computer, or other programmable data processing apparatus to produce a machine, such that the instructions, which execute via the processor of the computer or other programmable data processing apparatus, create means for implementing the functions/acts specified in the flowchart and/or block diagram block or blocks. These computer readable program instructions may also be stored in a computer readable storage medium that can direct a computer, a programmable data processing

apparatus, and/or other devices to function in a particular manner, such that the computer readable storage medium having instructions stored therein comprises an article of manufacture including instructions which implement aspects of the function/act specified in the flowchart and/or block diagram block or blocks.

The computer readable program instructions may also be loaded onto a computer, other programmable data processing apparatus, or other device to cause a series of operational steps to be performed on the computer, other programmable apparatus or other device to produce a computer implemented process, such that the instructions which execute on the computer, other programmable apparatus, or other device implement the functions/acts specified in the flowchart and/or block diagram block or blocks.

The flowchart and block diagrams in the Figures illustrate the architecture, functionality, and operation of possible implementations of systems, methods, and computer program products according to various embodiments of the present disclosure. In this regard, each block in the flowchart or block diagrams may represent a module, segment, or portion of instructions, which comprises one or more executable instructions for implementing the specified logical function(s). In some alternative implementations, the functions noted in the block may occur out of the order noted in the figures. For example, two blocks shown in succession may, in fact, be executed substantially concurrently, or the blocks may sometimes be executed in the reverse order, depending upon the functionality involved. It will also be noted that each block of the block diagrams and/or flowchart illustration, and combinations of blocks in the block diagrams and/or flowchart illustration, can be implemented by special purpose hardware-based systems that perform the specified functions or acts or carry out combinations of special purpose hardware and computer instructions.

The descriptions of the various embodiments of the present disclosure have been presented for purposes of illustration, but are not intended to be exhaustive or limited to the embodiments disclosed. Many modifications and variations will be apparent to those of ordinary skill in the art without departing from the scope and spirit of the described embodiments. The terminology used herein was chosen to best explain the principles of the embodiments, the practical application or technical improvement over technologies found in the marketplace, or to enable others of ordinary skill in the art to understand the embodiments disclosed herein.

Definitions

The articles “a” and “an” are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element.

The term “heteroatom” is art-recognized and refers to an atom of any element other than carbon or hydrogen. Illustrative heteroatoms include boron, nitrogen, oxygen, phosphorus, sulfur and selenium, and alternatively oxygen, nitrogen or sulfur.

The term “alkyl” as used herein is a term of art and refers to saturated aliphatic groups, including straight-chain alkyl groups, branched-chain alkyl groups, cycloalkyl (alicyclic) groups, alkyl substituted cycloalkyl groups, and cycloalkyl substituted alkyl groups. In certain embodiments, a straight-chain or branched-chain alkyl has about 30 or fewer carbon atoms in its backbone (e.g., C₁-C₃₀ for straight chain, C₃-C₃₀ for branched chain), and alternatively, about 20 or fewer, or 10 or fewer. In certain embodiments, the term “alkyl” refers to a C₁-C₁₀ alkyl group. In certain embodiments, the term “alkyl” refers to a C₁-C₆ alkyl group, for example a C₁-C₆ straight-chain alkyl group. In certain embodiments, the term “alkyl” refers to a C₃-C₁₂ branched-chain alkyl group. In certain embodiments, the term “alkyl” refers to a C₃-C₈ branched-chain alkyl group. Representative examples of alkyl include, but are not limited to, methyl, ethyl, n-propyl, iso-propyl, n-butyl, sec-butyl, iso-butyl, tert-butyl, n-pentyl, isopentyl, neopentyl, and n-hexyl.

The term “cycloalkyl” means mono- or bicyclic or bridged saturated carbocyclic rings, each having from 3 to 12 carbon atoms. Certain cycloalkyls have from 5-12 carbon atoms in their ring structure, and may have 6-10 carbons in the ring structure. Preferably, cycloalkyl is (C₃-C₇)cycloalkyl, which represents a monocyclic saturated carbocyclic ring, having from 3 to 7 carbon atoms. Examples of monocyclic cycloalkyls include cyclopropyl, cyclobutyl, cyclopentyl, cyclopentenyl, cyclohexyl, cyclohexenyl, cycloheptyl, and cyclooctyl. Bicyclic cycloalkyl ring systems include bridged monocyclic rings and fused bicyclic rings. Bridged monocyclic rings contain a monocyclic cycloalkyl ring where two non-adjacent carbon atoms of the monocyclic ring are linked by an alkylene bridge of between one and three additional carbon atoms (*i.e.*, a bridging group of the form -(CH₂)_w-, where *w* is 1, 2, or 3). Representative examples of bicyclic ring systems include, but are not limited to, bicyclo[3.1.1]heptane, bicyclo[2.2.1]heptane, bicyclo[2.2.2]octane, bicyclo[3.2.2]nonane, bicyclo[3.3.1]nonane, and bicyclo[4.2.1]nonane. Fused bicyclic cycloalkyl ring systems contain a monocyclic cycloalkyl ring fused to either a phenyl, a

monocyclic cycloalkyl, a monocyclic cycloalkenyl, a monocyclic heterocyclyl, or a monocyclic heteroaryl. The bridged or fused bicyclic cycloalkyl is attached to the parent molecular moiety through any carbon atom contained within the monocyclic cycloalkyl ring. Cycloalkyl groups are optionally substituted. In certain embodiments, the fused bicyclic cycloalkyl is a 5 or 6 membered monocyclic cycloalkyl ring fused to either a phenyl ring, a 5 or 6 membered monocyclic cycloalkyl, a 5 or 6 membered monocyclic cycloalkenyl, a 5 or 6 membered monocyclic heterocyclyl, or a 5 or 6 membered monocyclic heteroaryl, wherein the fused bicyclic cycloalkyl is optionally substituted.

The term “(cycloalkyl)alkyl” as used herein refers to an alkyl group substituted with one or more cycloalkyl groups. An example of (cycloalkyl)alkyl is cyclohexylmethyl group.

The term “heterocycloalkyl” as used herein refers to a radical of a non-aromatic ring system, including, but not limited to, monocyclic, bicyclic, and tricyclic rings, which can be completely saturated or which can contain one or more units of unsaturation, for the avoidance of doubt, the degree of unsaturation does not result in an aromatic ring system, and having 3 to 12 atoms including at least one heteroatom, such as nitrogen, oxygen, or sulfur. For purposes of exemplification, which should not be construed as limiting the scope of this invention, the following are examples of heterocyclic rings: aziridinyl, azirinyl, oxiranyl, thiiranyl, thiirenyl, dioxiranyl, diazirinyl, diazepanyl, 1,3-dioxanyl, 1,3-dioxolanyl, 1,3-dithiolanyl, 1,3-dithianyl, imidazolidinyl, isothiazolinyl, isothiazolidinyl, isoxazolinyl, isoxazolidinyl, azetyl, oxetanyl, oxetyl, thietanyl, thietyl, diazetidinyl, dioxetanyl, dioxetenyl, dithietanyl, dithietyl, dioxalanyl, oxazolyl, thiazolyl, triazinyl, isothiazolyl, isoxazolyl, azepines, azetidiny, morpholinyl, oxadiazolinyl, oxadiazolidinyl, oxazolinyl, oxazolidinyl, oxopiperidinyl, oxopyrrolidinyl, piperazinyl, piperidinyl, pyranyl, pyrazolinyl, pyrazolidinyl, pyrrolinyl, pyrrolidinyl, quinuclidinyl, thiomorpholinyl, tetrahydropyranyl, tetrahydrofuranyl, tetrahydrothienyl, thiadiazolinyl, thiadiazolidinyl, thiazolinyl, thiazolidinyl, thiomorpholinyl, 1,1-dioxidothiomorpholinyl (thiomorpholine sulfone), thiopyranyl, and trithianyl. A heterocycloalkyl group is optionally substituted by one or more substituents as described below.

The term “(heterocycloalkyl)alkyl” as used herein refers to an alkyl group substituted with one or more heterocycloalkyl (i.e., heterocyclyl) groups.

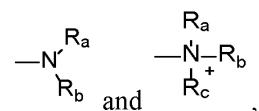
The term “alkenyl” as used herein means a straight or branched chain hydrocarbon radical containing from 2 to 10 carbons and containing at least one carbon-carbon double bond formed by the removal of two hydrogens. Representative examples of alkenyl include,

but are not limited to, ethenyl, 2-propenyl, 2-methyl-2-propenyl, 3-butenyl, 4-pentenyl, 5-hexenyl, 2-heptenyl, 2-methyl-1-heptenyl, and 3-decenyl. The unsaturated bond(s) of the alkenyl group can be located anywhere in the moiety and can have either the (Z) or the (E) configuration about the double bond(s).

The term “alkynyl” as used herein means a straight or branched chain hydrocarbon radical containing from 2 to 10 carbon atoms and containing at least one carbon-carbon triple bond. Representative examples of alkynyl include, but are not limited, to acetylenyl, 1-propynyl, 2-propynyl, 3-butylnyl, 2-pentylnyl, and 1-butylnyl.

The term “alkylene” is art-recognized, and as used herein pertains to a diradical obtained by removing two hydrogen atoms of an alkyl group, as defined above. In one embodiment an alkylene refers to a disubstituted alkane, i.e., an alkane substituted at two positions with substituents such as halogen, azide, alkyl, aralkyl, alkenyl, alkynyl, cycloalkyl, hydroxyl, alkoxy, amino, nitro, sulfhydryl, imino, amido, phosphonate, phosphinate, carbonyl, carboxyl, silyl, ether, alkylthio, sulfonyl, sulfonamido, ketone, aldehyde, ester, heterocyclyl, aromatic or heteroaromatic moieties, fluoroalkyl (such as trifluoromethyl), cyano, or the like. That is, in one embodiment, a “substituted alkyl” is an “alkylene”.

The term “amino” is a term of art and as used herein refers to both unsubstituted and substituted amines, e.g., a moiety that may be represented by the general formulas:



wherein R_a , R_b , and R_c each independently represent a hydrogen, an alkyl, an alkenyl, $-(\text{CH}_2)_x\text{—R}_d$, or R_a and R_b , taken together with the N atom to which they are attached complete a heterocycle having from 4 to 8 atoms in the ring structure; R_d represents an aryl, a cycloalkyl, a cycloalkenyl, a heterocyclyl or a polycyclyl; and x is zero or an integer in the range of 1 to 8. In certain embodiments, only one of R_a or R_b may be a carbonyl, e.g., R_a , R_b , and the nitrogen together do not form an imide. In other embodiments, R_a and R_b (and optionally R_c) each independently represent a hydrogen, an alkyl, an alkenyl, or $-(\text{CH}_2)_x\text{—R}_d$. In certain embodiments, the term “amino” refers to $-\text{NH}_2$.

In certain embodiments, the term “alkylamino” refers to $-\text{NH}(\text{alkyl})$.

In certain embodiments, the term “dialkylamino” refers to $-\text{N}(\text{alkyl})_2$.

The term “amido”, as used herein, means $-\text{NHC}(=\text{O})-$, wherein the amido group is bound to the parent molecular moiety through the nitrogen. Examples of amido include alkylamido such as $\text{CH}_3\text{C}(=\text{O})\text{N}(\text{H})-$ and $\text{CH}_3\text{CH}_2\text{C}(=\text{O})\text{N}(\text{H})-$.

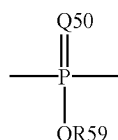
The term “acyl” is a term of art and as used herein refers to any group or radical of the form RCO- where R is any organic group, e.g., alkyl, aryl, heteroaryl, aralkyl, and heteroaralkyl. Representative acyl groups include acetyl, benzoyl, and malonyl.

The term “aminoalkyl” as used herein refers to an alkyl group substituted with one or more amino groups. In one embodiment, the term “aminoalkyl” refers to an aminomethyl group.

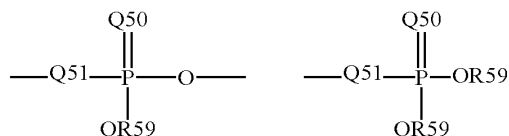
The term “aminoacyl” is a term of art and as used herein refers to an acyl group substituted with one or more amino groups.

The term “aminothionyl” as used herein refers to an analog of an aminoacyl in which the O of RC(O)- has been replaced by sulfur, hence is of the form RC(S)-.

The term “phosphoryl” is a term of art and as used herein may in general be represented by the formula:



wherein Q50 represents S or O, and R59 represents hydrogen, a lower alkyl or an aryl; for example, -P(O)(OMe)- or -P(O)(OH)₂. When used to substitute, e.g., an alkyl, the phosphoryl group of the phosphorylalkyl may be represented by the general formulas:



wherein Q50 and R59, each independently, are defined above, and Q51 represents O, S or N; for example, -O-P(O)(OH)OMe or -NH-P(O)(OH)₂. When Q50 is S, the phosphoryl moiety is a “phosphorothioate.”

The term “aminophosphoryl” as used herein refers to a phosphoryl group substituted with at least one amino group, as defined herein; for example, -P(O)(OH)NMe₂.

The term “azide” or “azido”, as used herein, means an -N₃ group.

The term “carbonyl” as used herein refers to -C(=O)-.

The term “thiocarbonyl” as used herein refers to -C(=S)-.

The term “alkylphosphoryl” as used herein refers to a phosphoryl group substituted with at least one alkyl group, as defined herein; for example, -P(O)(OH)Me.

The term “alkylthio” as used herein refers to alkyl-S-. The term “(alkylthio)alkyl” refers to an alkyl group substituted by an alkylthio group.

The term “carboxy”, as used herein, means a $\text{-CO}_2\text{H}$ group.

The term “aryl” is a term of art and as used herein refers to includes monocyclic, bicyclic and polycyclic aromatic hydrocarbon groups, for example, benzene, naphthalene, anthracene, and pyrene. Typically, an aryl group contains from 6-10 carbon ring atoms (i.e., $(\text{C}_6\text{-C}_{10})\text{aryl}$). The aromatic ring may be substituted at one or more ring positions with one or more substituents, such as halogen, azide, alkyl, aralkyl, alkenyl, alkynyl, cycloalkyl, hydroxyl, alkoxy, amino, nitro, sulfhydryl, imino, amido, phosphonate, phosphinate, carbonyl, carboxyl, silyl, ether, alkylthio, sulfonyl, sulfonamido, ketone, aldehyde, ester, heterocyclyl, aromatic or heteroaromatic moieties, fluoroalkyl (such as trifluoromethyl), cyano, or the like. The term “aryl” also includes polycyclic ring systems having two or more cyclic rings in which two or more carbons are common to two adjoining rings (the rings are “fused rings”) wherein at least one of the rings is an aromatic hydrocarbon, e.g., the other cyclic rings may be cycloalkyls, cycloalkenyls, cycloalkynyls, aryls, heteroaryls, and/or heterocyclyls. In certain embodiments, the term “aryl” refers to a phenyl group.

The term “arylene” means a diradical obtained by removing two hydrogen atoms of an aryl group, as defined above. In certain embodiments an arylene refers to a disubstituted arene, i.e., an arene substituted at two positions with substituents such as halogen, azide, alkyl, aralkyl, alkenyl, alkynyl, cycloalkyl, hydroxyl, alkoxy, amino, nitro, sulfhydryl, imino, amido, phosphonate, phosphinate, carbonyl, carboxyl, silyl, ether, alkylthio, sulfonyl, sulfonamido, ketone, aldehyde, ester, heterocyclyl, aromatic or heteroaromatic moieties, fluoroalkyl (such as trifluoromethyl), cyano, or the like. That is, in certain embodiments, a “substituted aryl” is an “arylene”.

The term “heteroaryl” is a term of art and as used herein refers to a monocyclic, bicyclic, and polycyclic aromatic group having 3 to 12 total atoms including one or more heteroatoms such as nitrogen, oxygen, or sulfur in the ring structure. Exemplary heteroaryl groups include azaindolyl, benzo(b)thienyl, benzimidazolyl, benzofuranyl, benzoxazolyl, benzothiazolyl, benzothiadiazolyl, benzotriazolyl, benzoxadiazolyl, furanyl, imidazolyl, imidazopyridinyl, indolyl, indolynyl, indazolyl, isoindolynyl, isoxazolyl, isothiazolyl, isoquinolynyl, oxadiazolyl, oxazolyl, purinyl, pyranyl, pyrazinyl, pyrazolyl, pyridinyl, pyrimidinyl, pyrrolyl, pyrrolo[2,3-d]pyrimidinyl, pyrazolo[3,4-d]pyrimidinyl, quinolynyl, quinazolinyl, triazolyl, thiazolyl, thiophenyl, tetrahydroindolyl, tetrazolyl, thiadiazolyl, thienyl, thiomorpholynyl, triazolyl or tropanyl, and the like. The “heteroaryl” may be substituted at one or more ring positions with one or more substituents such as halogen, azide,

alkyl, aralkyl, alkenyl, alkynyl, cycloalkyl, hydroxyl, alkoxy, amino, nitro, sulfhydryl, imino, amido, phosphonate, phosphinate, carbonyl, carboxyl, silyl, ether, alkylthio, sulfonyl, sulfonamido, ketone, aldehyde, ester, heterocyclyl, aromatic or heteroaromatic moieties, fluoroalkyl (such as trifluoromethyl), cyano, or the like. The term “heteroaryl” also includes polycyclic ring systems having two or more cyclic rings in which two or more carbons are common to two adjoining rings (the rings are “fused rings”) wherein at least one of the rings is an aromatic group having one or more heteroatoms in the ring structure, e.g., the other cyclic rings may be cycloalkyls, cycloalkenyls, cycloalkynyls, aryls, heteroaryls, and/or heterocyclyls.

The term “heteroarylene” means a diradical obtained by removing two hydrogen atoms of a heteroaryl group, as defined above. In certain embodiments an heteroarylene refers to a disubstituted heteroarene, i.e., a heteroarene substituted at two positions with substituents such as halogen, azide, alkyl, aralkyl, alkenyl, alkynyl, cycloalkyl, hydroxyl, alkoxy, amino, nitro, sulfhydryl, imino, amido, phosphonate, phosphinate, carbonyl, carboxyl, silyl, ether, alkylthio, sulfonyl, sulfonamido, ketone, aldehyde, ester, heterocyclyl, aromatic or heteroaromatic moieties, fluoroalkyl (such as trifluoromethyl), cyano, or the like. That is, in certain embodiments, a “substituted heteroaryl” is an “heteroarylene”.

The term “aralkyl” or “arylalkyl” is a term of art and as used herein refers to an alkyl group substituted with an aryl group, wherein the moiety is appended to the parent molecule through the alkyl group.

The term “heteroaralkyl” or “heteroarylalkyl” is a term of art and as used herein refers to an alkyl group substituted with a heteroaryl group, appended to the parent molecular moiety through the alkyl group.

The term “alkoxy” as used herein means an alkyl group, as defined herein, appended to the parent molecular moiety through an oxygen atom. Representative examples of alkoxy include, but are not limited to, methoxy, ethoxy, propoxy, 2-propoxy, butoxy, *tert*-butoxy, pentyloxy, and hexyloxy.

The term “alkoxyalkyl” refers to an alkyl group substituted by an alkoxy group.

The term “alkoxycarbonyl” means an alkoxy group, as defined herein, appended to the parent molecular moiety through a carbonyl group, represented by $-C(=O)-$, as defined herein. Representative examples of alkoxycarbonyl include, but are not limited to, methoxycarbonyl, ethoxycarbonyl, and *tert*-butoxycarbonyl.

The term “alkylcarbonyl”, as used herein, means an alkyl group, as defined herein, appended to the parent molecular moiety through a carbonyl group, as defined herein. Representative examples of alkylcarbonyl include, but are not limited to, acetyl, 1-oxopropyl, 2,2-dimethyl-1-oxopropyl, 1-oxobutyl, and 1-oxopentyl.

The term “arylcabonyl”, as used herein, means an aryl group, as defined herein, appended to the parent molecular moiety through a carbonyl group, as defined herein. Representative examples of arylcarbonyl include, but are not limited to, benzoyl and (2-pyridinyl)carbonyl.

The term “alkylcarbonyloxy” and “arylcabonyloxy”, as used herein, means an alkylcarbonyl or arylcarbonyl group, as defined herein, appended to the parent molecular moiety through an oxygen atom. Representative examples of alkylcarbonyloxy include, but are not limited to, acetyloxy, ethylcarbonyloxy, and *tert*-butylcarbonyloxy. Representative examples of arylcarbonyloxy include, but are not limited to phenylcarbonyloxy.

The term “alkenoxy” or “alkenoxyl” means an alkenyl group, as defined herein, appended to the parent molecular moiety through an oxygen atom. Representative examples of alkenoxyl include, but are not limited to, 2-propen-1-oxyl (i.e., $\text{CH}_2=\text{CH}-\text{CH}_2-\text{O}-$) and vinyloxy (i.e., $\text{CH}_2=\text{CH}-\text{O}-$).

The term “aryloxy” as used herein means an aryl group, as defined herein, appended to the parent molecular moiety through an oxygen atom.

The term “heteroaryloxy” as used herein means a heteroaryl group, as defined herein, appended to the parent molecular moiety through an oxygen atom.

The term “carbocyclyl” as used herein means a monocyclic or multicyclic (e.g., bicyclic, tricyclic, etc.) hydrocarbon radical containing from 3 to 12 carbon atoms that is completely saturated or has one or more unsaturated bonds, and for the avoidance of doubt, the degree of unsaturation does not result in an aromatic ring system (e.g., phenyl). Examples of carbocyclyl groups include 1-cyclopropyl, 1-cyclobutyl, 2-cyclopentyl, 1-cyclopentenyl, 3-cyclohexyl, 1-cyclohexenyl and 2-cyclopentenylmethyl.

The term “cyano” is a term of art and as used herein refers to $-\text{CN}$.

The term “halo” is a term of art and as used herein refers to $-\text{F}$, $-\text{Cl}$, $-\text{Br}$, or $-\text{I}$.

The term “haloalkyl” as used herein refers to an alkyl group, as defined herein, wherein some or all of the hydrogens are replaced with halogen atoms.

The term “hydroxy” is a term of art and as used herein refers to $-\text{OH}$.

The term “hydroxyalkyl”, as used herein, means at least one hydroxy group, as defined herein, is appended to the parent molecular moiety through an alkyl group, as defined herein. Representative examples of hydroxyalkyl include, but are not limited to, hydroxymethyl, 2-hydroxyethyl, 3-hydroxypropyl, 2,3-dihydroxypentyl, and 2-ethyl-4-hydroxyheptyl.

The term “silyl”, as used herein, includes hydrocarbyl derivatives of the silyl ($\text{H}_3\text{Si}-$) group (i.e., $(\text{hydrocarbyl})_3\text{Si}-$), wherein a hydrocarbyl groups are univalent groups formed by removing a hydrogen atom from a hydrocarbon, e.g., ethyl, phenyl. The hydrocarbyl groups can be combinations of differing groups which can be varied in order to provide a number of silyl groups, such as trimethylsilyl (TMS), tert-butyldiphenylsilyl (TBDPS), tert-butyldimethylsilyl (TBS/TBDMS), triisopropylsilyl (TIPS), and [2-(trimethylsilyl)ethoxy]methyl (SEM).

The term “silyloxy”, as used herein, means a silyl group, as defined herein, is appended to the parent molecule through an oxygen atom.

It will be understood that “substitution” or “substituted with” includes the implicit proviso that such substitution is in accordance with permitted valence of the substituted atom and the substituent, and that the substitution results in a stable compound, e.g., which does not spontaneously undergo transformation such as by rearrangement, fragmentation, decomposition, cyclization, elimination, or other reaction.

The term “substituted” is also contemplated to include all permissible substituents of organic compounds. In a broad aspect, the permissible substituents include acyclic and cyclic, branched and unbranched, carbocyclic and heterocyclic, aromatic and nonaromatic substituents of organic compounds. Illustrative substituents include, for example, those described herein above. The permissible substituents may be one or more and the same or different for appropriate organic compounds. For purposes of this invention, the heteroatoms such as nitrogen may have hydrogen substituents and/or any permissible substituents of organic compounds described herein which satisfy the valences of the heteroatoms. This invention is not intended to be limited in any manner by the permissible substituents of organic compounds.

As used herein, the term “optionally substituted” or “substituted or unsubstituted” when it precedes a list of chemical moieties means that the list of chemical moieties that follow are each substituted or unsubstituted. For example, “substituted or unsubstituted aryl, heteroaryl, and cycloalkyl” or “optionally substituted aryl, heteroaryl, and cycloalkyl” means

substituted or unsubstituted aryl, substituted or unsubstituted heteroaryl, and substituted or unsubstituted cycloalkyl.

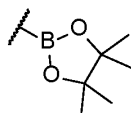
In certain embodiments, the optional substituents can include, for example, halogen, haloalkyl, hydroxyl, carbonyl (such as carboxyl (-COOH), alkoxycarbonyl, formyl, or acyl), thiocarbonyl (such as thioester, thioacetate, or thioformate), alkoxyl, alkenyloxy, alkynyloxy, phosphoryl, phosphate, phosphonate, phosphinate, amino (including alkyl- and dialkylamino), amido, amidine, imine, cyano, nitro, oxo (=O), azido, sulfhydryl, alkylthio, sulfate, sulfonate, sulfamoyl, sulfonamido, sulfonyl, silyl, silyloxy, heterocycloalkyl, cycloalkyl, alkyl, alkenyl, alkynyl, aryl, heteroaryl, arylalkyl, or heteroarylalkyl group.

For purposes of the invention, the chemical elements are identified in accordance with the Periodic Table of the Elements, CAS version, Handbook of Chemistry and Physics, 67th Ed., 1986-87, inside cover.

Other chemistry terms herein are used according to conventional usage in the art, as exemplified by The McGraw-Hill Dictionary of Chemical Terms (ed. Parker, S., 1985), McGraw-Hill, San Francisco, incorporated herein by reference). Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains.

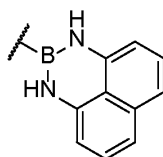
The phrase “protecting group”, as used herein, means temporary substituents which protect a potentially reactive functional group from undesired chemical transformations. Examples of such protecting groups include esters of carboxylic acids, silyl ethers of alcohols, and acetals and ketals of aldehydes and ketones, respectively. The field of protecting group chemistry has been reviewed (Greene, T.W.; Wuts, P.G.M. *Protective Groups in Organic Synthesis*, 2nd ed.; Wiley: New York, 1991). Protected forms of the inventive compounds are included within the scope of this invention.

As used herein, the term “B(pin)” is art recognized and refers to a pinacol boronate moiety. The pinacol boronate moiety has the following structure:



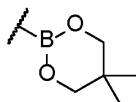
wherein the wavy bond represents the point of attachment to the parent molecular moiety.

As used herein, the term “B(dan)” is art recognized and refers to a 2,3-dihydro-1H-naphtho[1,8-de][1,3,2]diazaborinine moiety. The 2,3-dihydro-1H-naphtho[1,8-de][1,3,2]diazaborinine moiety has the following structure:



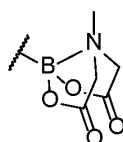
wherein the wavy bond represents the point of attachment to the parent molecular moiety.

As used herein, the term “B(neop)” is art recognized and refers to a neopentyl glycol boronate moiety. The neopentyl glycol boronate moiety has the following structure:



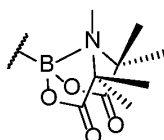
wherein the wavy bond represents the point of attachment to the parent molecular moiety.

As used herein, the term “B(MIDA)” is art recognized and refers to an *N*-methyliminodiacetic acid boronate moiety. The *N*-methyliminodiacetic acid boronate moiety has the following structure:



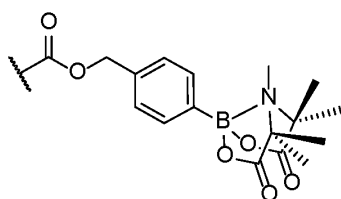
wherein the wavy bond represents the point of attachment to the parent molecular moiety.

As used herein, the term “B(TIDA)” is art recognized and refers to a tetramethyl *N*-methyliminodiacetic acid boronate moiety. The *N*-methyliminodiacetic acid boronate moiety has the following structure:



wherein the wavy bond represents the point of attachment to the parent molecular moiety.

As used herein, the terms “Cbz(BTIDA)” or “NCbz(BTIDA)” are art recognized and refer to a benzyloxycarbonyl tetramethyl *N*-methyliminodiacetic acid boronate moiety. The benzyloxycarbonyl *N*-methyliminodiacetic acid boronate moiety has the following structure:



wherein the wavy bond represents the point of attachment to the parent molecular moiety.

EXAMPLES

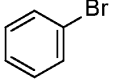
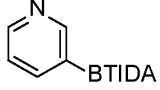
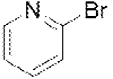
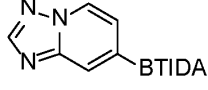
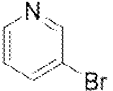
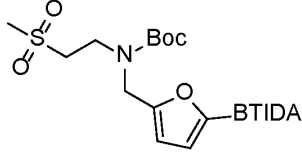
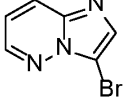
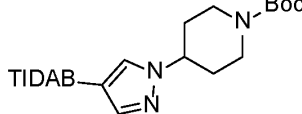
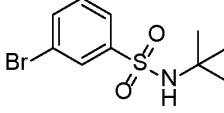
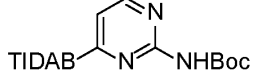
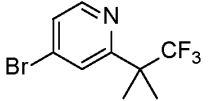
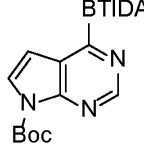
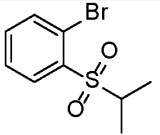
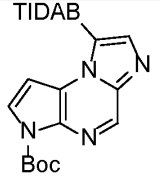
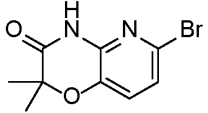
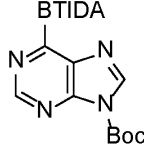
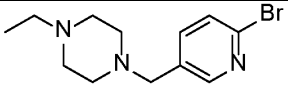
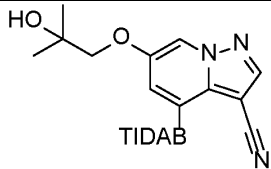
The invention now being generally described, it will be more readily understood by reference to the following examples which are included merely for purposes of illustration of certain aspects and embodiments of the present invention, and they are not intended to limit the scope of the invention.

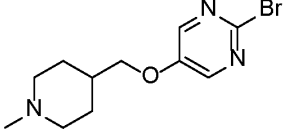
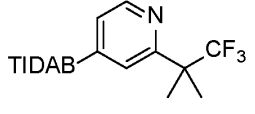
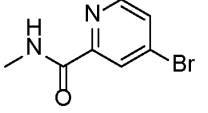
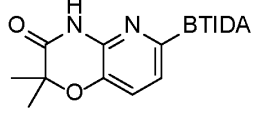
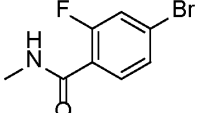
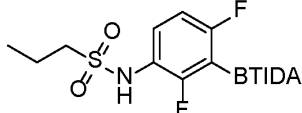
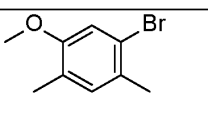
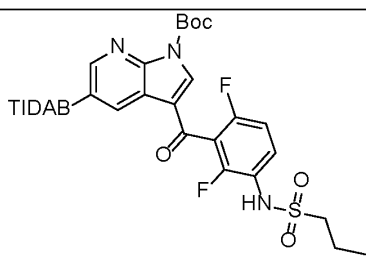
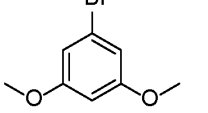
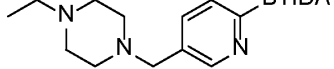
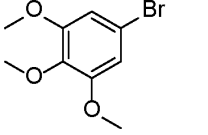
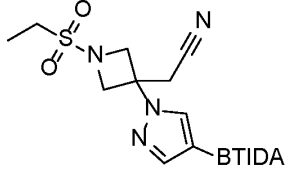
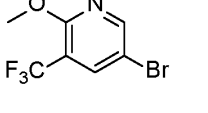
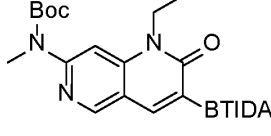
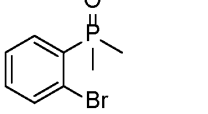
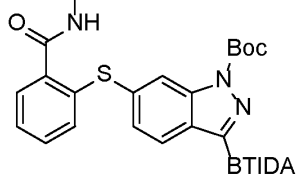
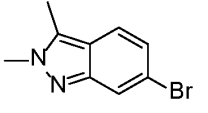
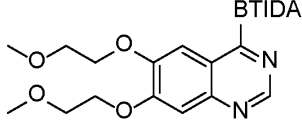
Example 1 – Exemplary Materials and Methods

General automated and manual synthetic procedures, general methods for the synthesis of molecular building blocks, and general automated and manual cross-coupling/deprotection strategies incorporating protected boronates (e.g., Cbz(BTIDA) and NCbz(BTIDA)) are disclosed in U.S. Provisional Patent Application No. 63/569,496, filed March 25, 2024; the contents of which are incorporated herein by reference in their entirety.

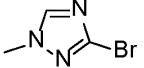
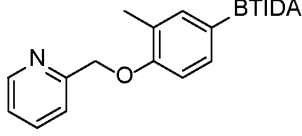
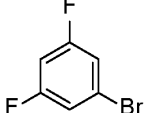
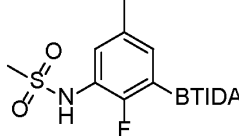
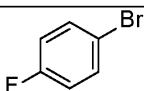
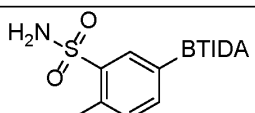
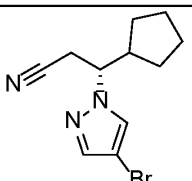
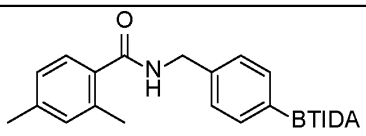
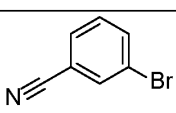
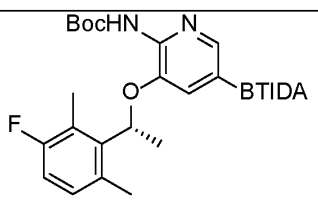
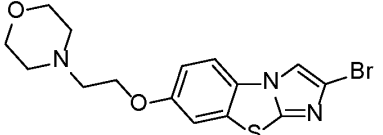
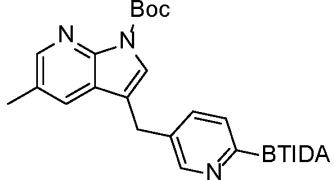
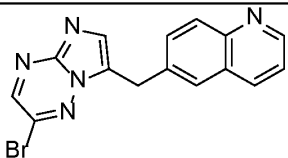
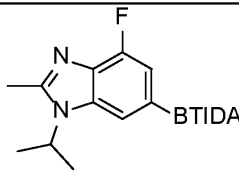
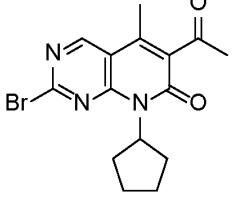
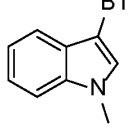
Table 1. Exemplary Molecular Building Blocks of This Disclosure

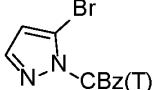
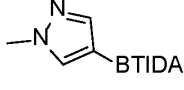
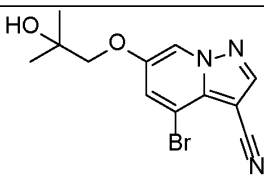
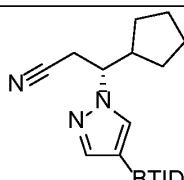
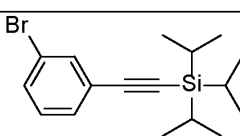
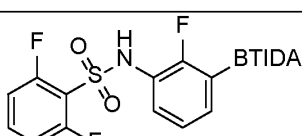
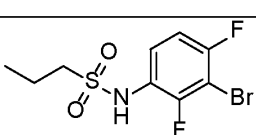
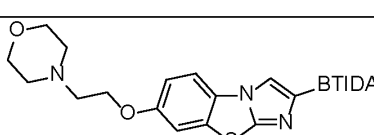
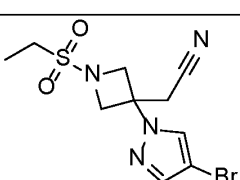
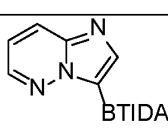
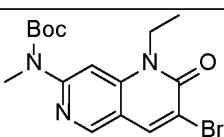
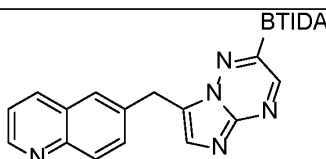
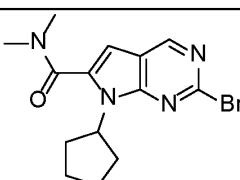
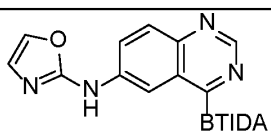
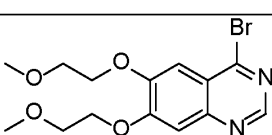
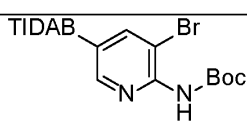
Identifier	Structure	Identifier	Structure
Commercial Halide Building Blocks		TIDA Boronate Building Blocks (from Commercial Boronic Acids) (cont.)	
A0		A316	
A1		A319	
A3		A320	

A4		A321	
A5		A322	
Custom TIDA Boronate Building Blocks			
A6		A23	
A8		A25	
A21		A33	
A66		A43	
A72		A45	
A79		A47	
A89		A63	

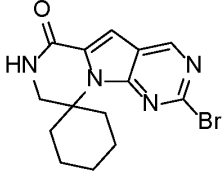
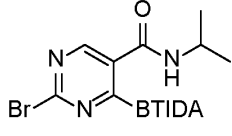
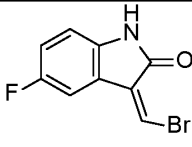
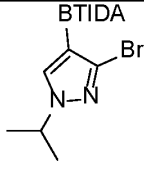
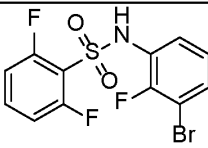
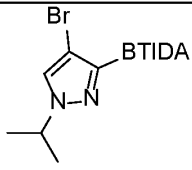
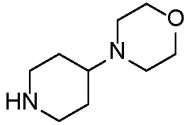
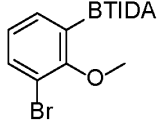
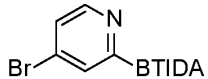
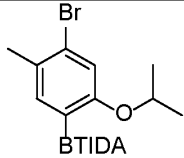
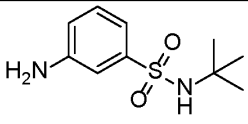
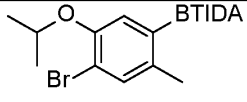
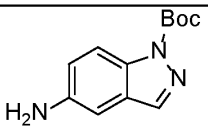
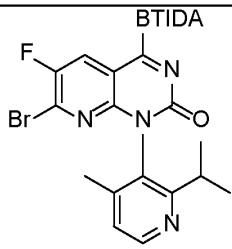
A102		A65	
A109		A78	
A111		A83	
A131		A85	
A132		A88	
A133		A92	
A159		A97	
A161		A114	
A171		A120	

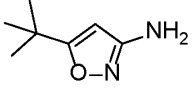
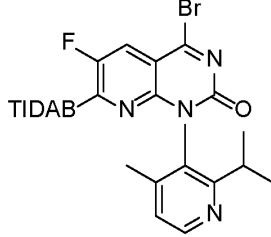
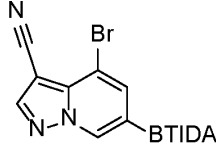
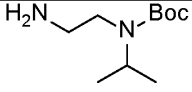
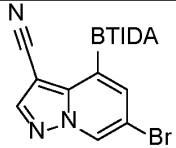
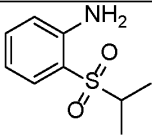
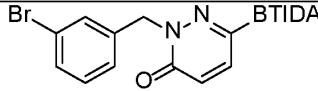
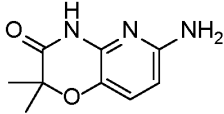
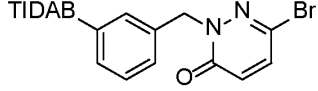
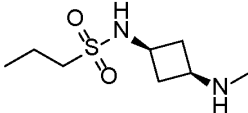
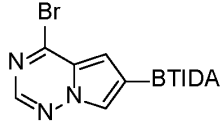
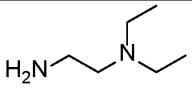
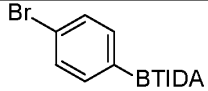
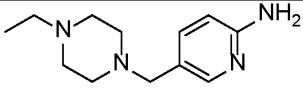
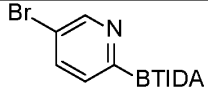
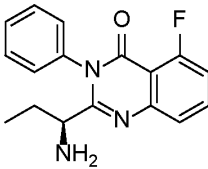
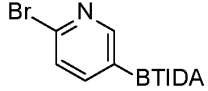
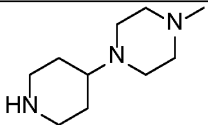
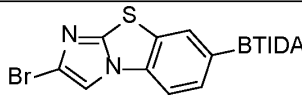
A184		A139	
A206		A140	
A207		A147	
A208		A148	
A209		A151	
A214		A160	
A236		A166	
A239		A176	
A240		A178	

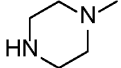
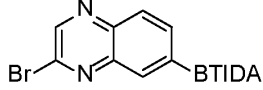
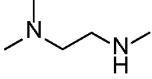
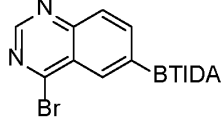
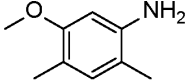
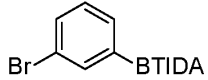
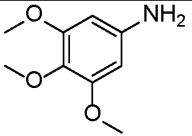
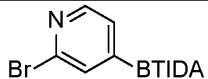
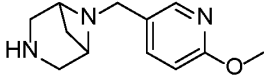
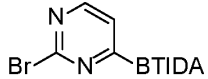
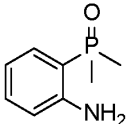
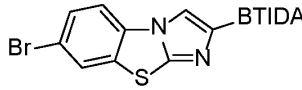
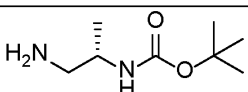
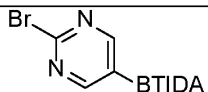
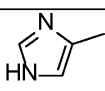
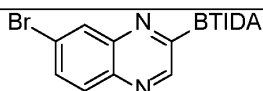
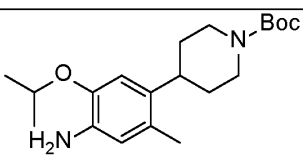
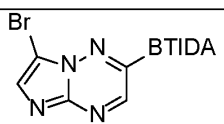
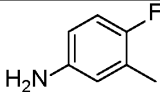
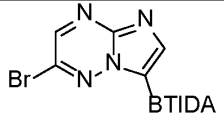
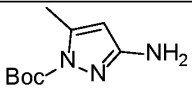
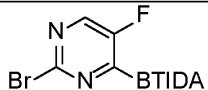
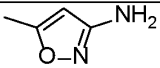
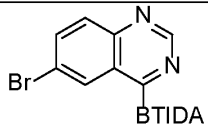
A241		A179	
A243		A201	
A244		A205	
A251		A210	
A253		A216	
Custom Halide Building Blocks			
A7		A232	
A9		A235	
A17		A237	

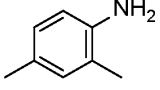
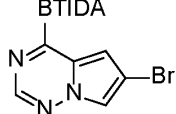
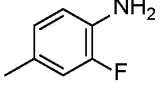
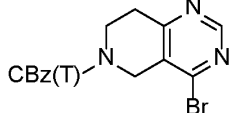
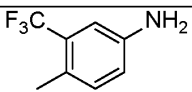
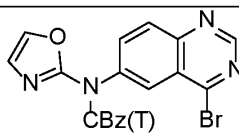
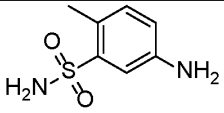
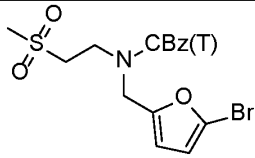
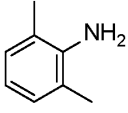
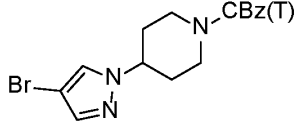
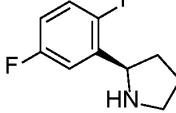
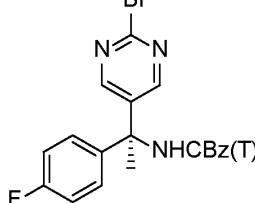
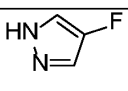
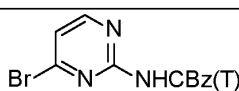
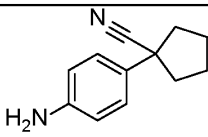
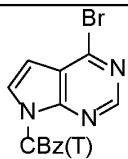
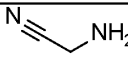
A58		A238	
A64		A250	
A74		A301	
A84		A326	
A93		A329	
A98		A332	
A99		A335	
		Halide-TIDA Boronate Bifunctional Building Blocks	
A121		A30	

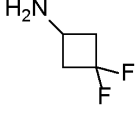
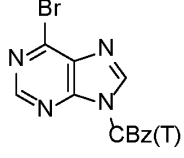
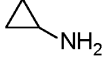
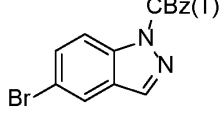
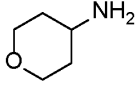
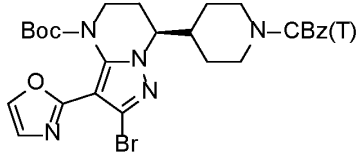
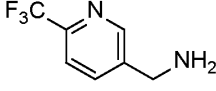
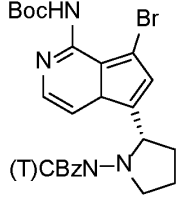
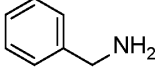
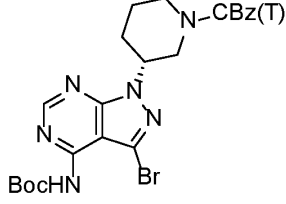
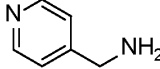
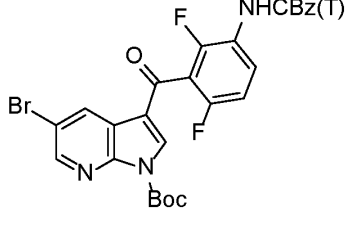
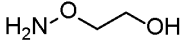
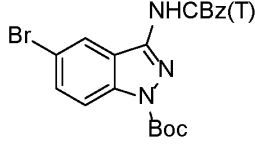
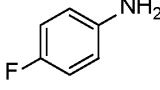
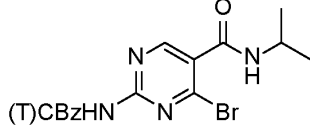
A141		A31	
A142		A38	
A149		A39	
A150		A40	
A153		A41	
A172		A49	
A181		A51	
A186		A61	
A187		A62	

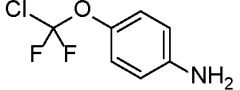
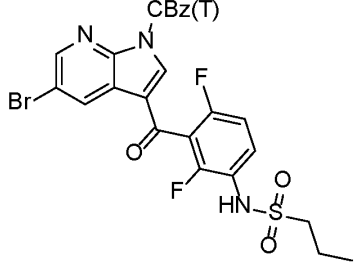
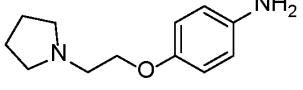
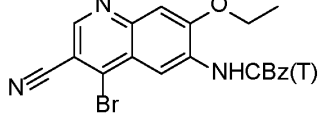
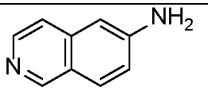
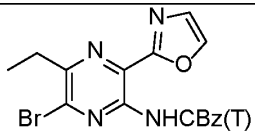
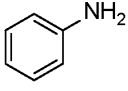
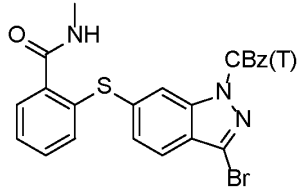
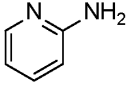
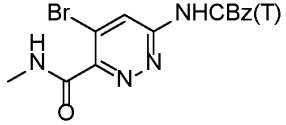
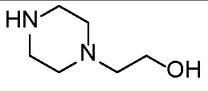
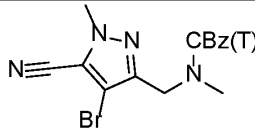
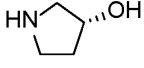
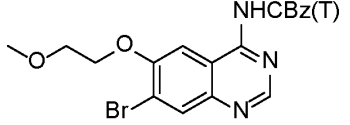
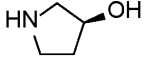
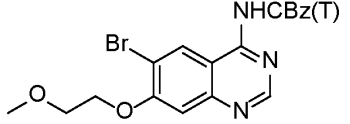
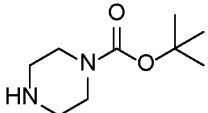
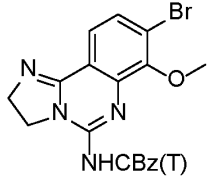
A296		A69	
A297		A75	
A302		A76	
Commercial Amine Building Blocks			
A12		A126	
A13		A305	
A19		A175	
A22		A182	
A56		A226	

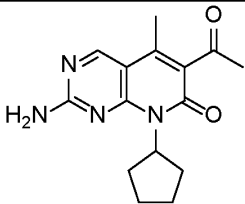
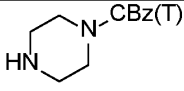
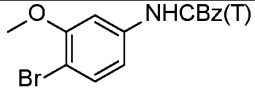
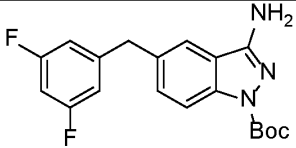
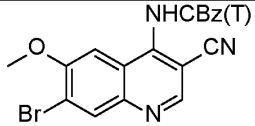
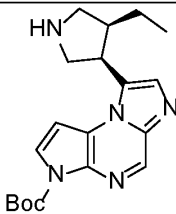
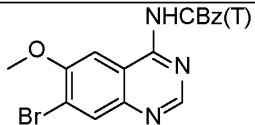
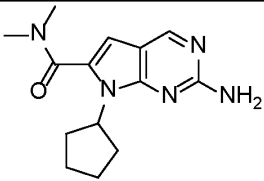
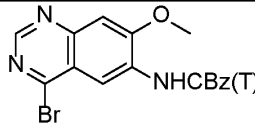
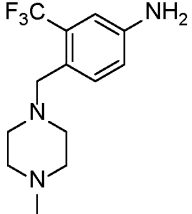
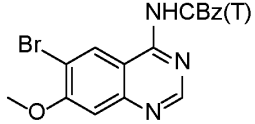
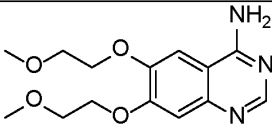
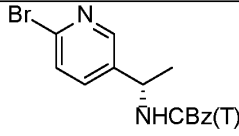
A59		A227	
A67		A255	
A68		A256	
A73		A303	
A80		A304	
A82		A307	
A87		A308	
A90		A309	
A95		A310	
A103		A314	

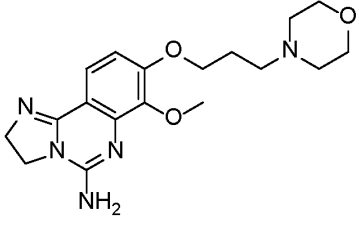
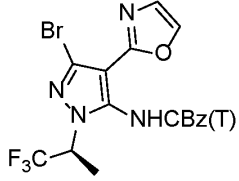
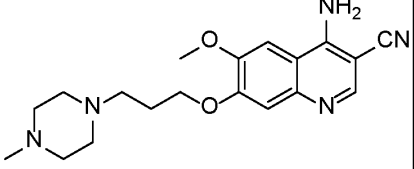
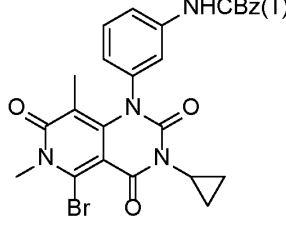
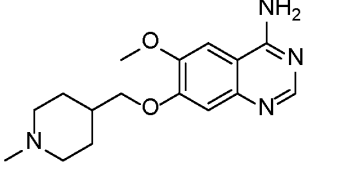
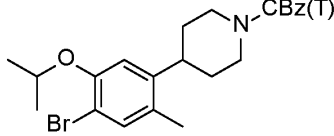
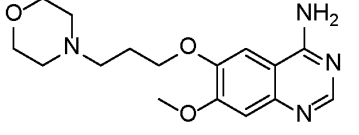
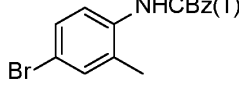
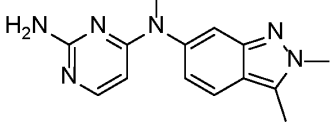
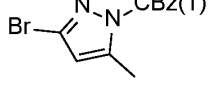
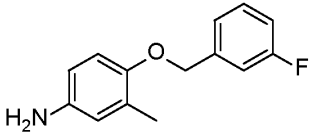
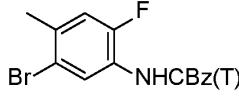
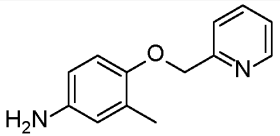
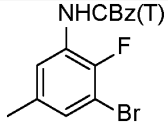
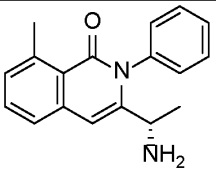
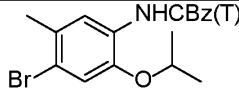
A107		A315	
A117		A317	
A135		A318	
A136		A323	
A156		A324	
A162		A325	
A163		A327	
A174		A328	
A194		A330	
A196		A331	
A200		A333	
A201		A334	

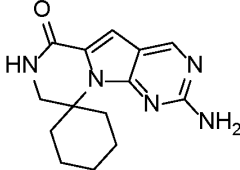
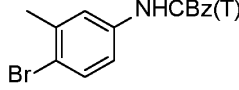
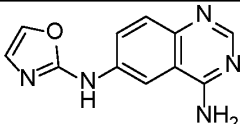
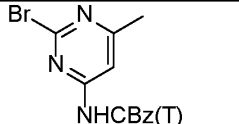
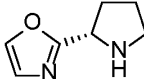
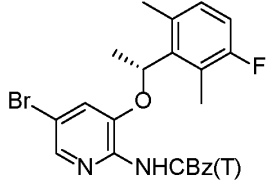
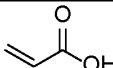
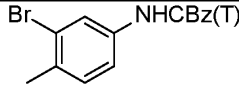
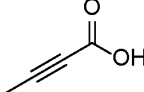
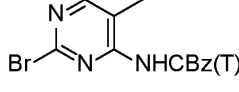
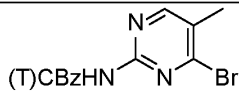
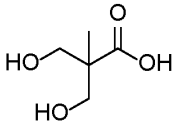
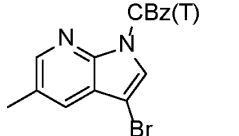
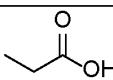
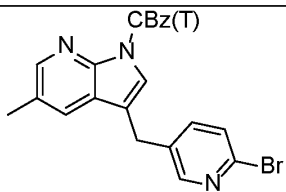
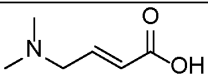
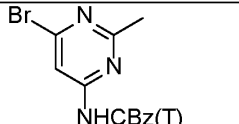
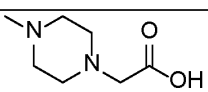
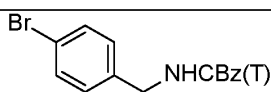
A218		A336	BTIDA 
		Halide-NCBz(TIDA) Bifunctional Building Blocks ((T) = TIDA Boronate)	
A221		A10	
A223		A11	
A224		A24	
A226		A26	
A245		A29	
A246		A34	
A247		A44	
A249		A46	

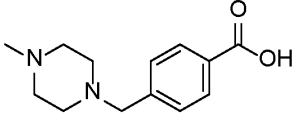
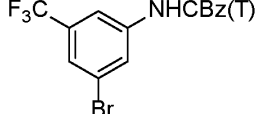
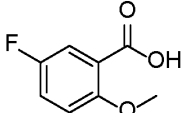
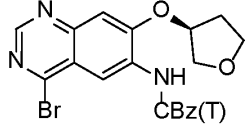
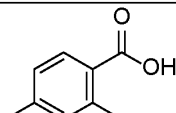
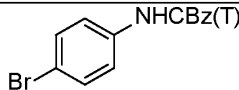
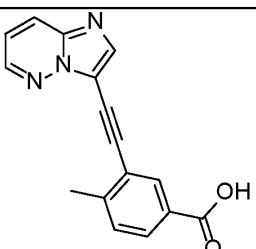
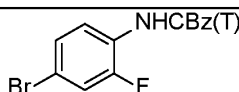
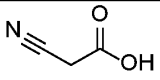
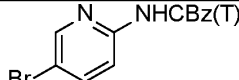
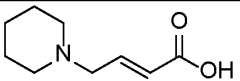
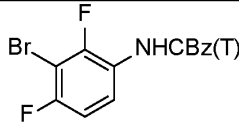
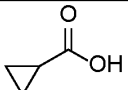
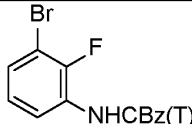
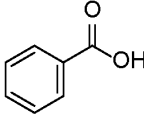
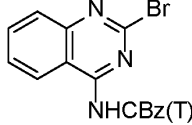
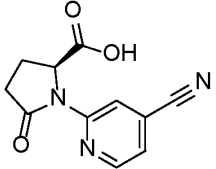
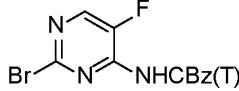
A257		A48	
A260		A55	
A261		A28	
A263		A35	
A264		A36	
A265		A42	
A266		A52	
A274		A70	

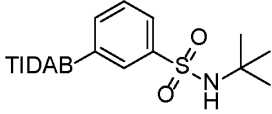
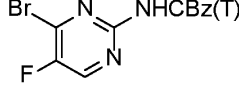
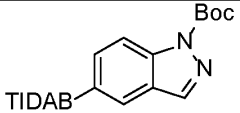
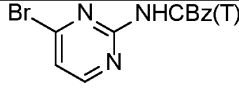
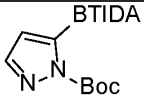
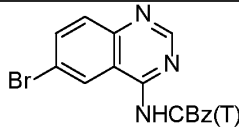
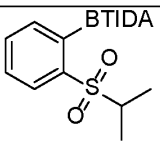
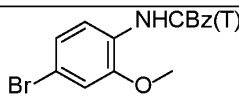
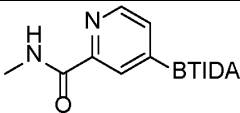
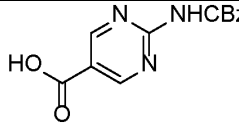
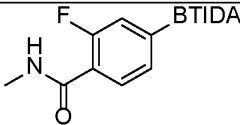
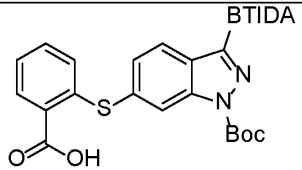
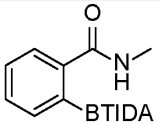
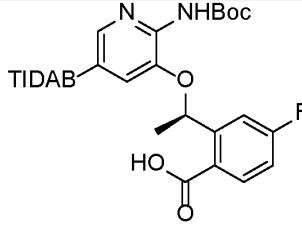
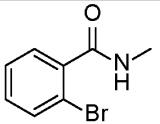
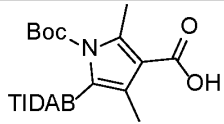
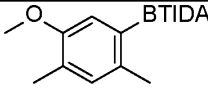
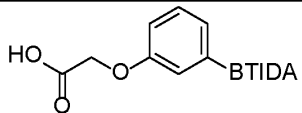
A275		A86	
A276		A91	
A277		A96	
A278		A115	
A280		A116	
A339		A118	
A340		A119	
A341		A122	
A27		A125	

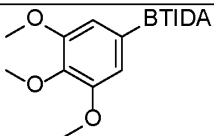
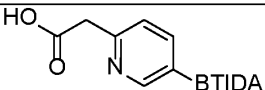
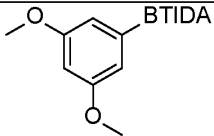
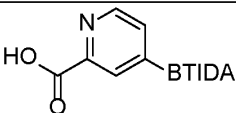
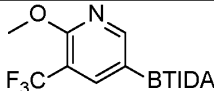
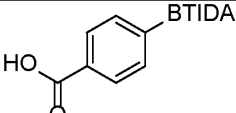
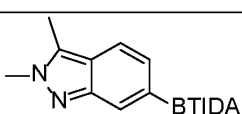
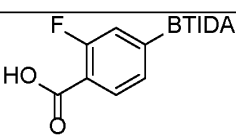
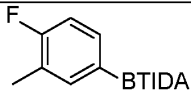
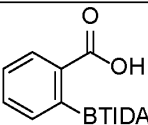
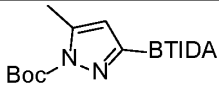
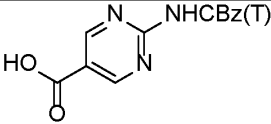
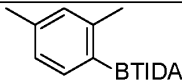
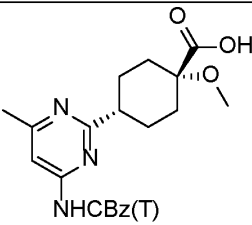
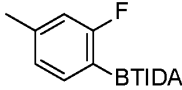
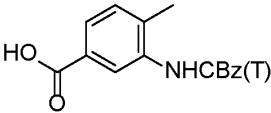
Custom Amine Building Blocks			
A18		A127	
A27		A137	
A53		A143	
A94		A145	
A100		A152	
A106		A154	
A123		A165	

A128		A167	
A144		A168	
A146		A183	
A155		A185	
A173		A188	
A195		A189	
A196		A190	
A225		A191	

A283		A193	
A286		A197	
A341		A217	
Commercial Carboxylic Acid Building Blocks			
A16		A220	
A17		A228	
A18		A229	
A79		A231	
A83		A233	
A103		A234	
A106		A260	

A107		A266	
A159		A267	
A213		A268	
A214		A269	
A250		A270	
A289		A271	
A290		A276	
A296		A279	
Custom Carboxylic Acid Building Blocks			
A254		A280	
TIDA Boronate Building Blocks (from Commercial Boronic Acids)			

A20		A282	
A54		A284	
A57		A285	
A71		A134	
Carboxylic Acid-TIDA Boronate Bifunctional Building Blocks			
A108		A32	
A110		A50	
A112		A164	
A113		A169	
A129		A289	

A130		A290	
A138		A291	
A158		A292	
A170		A293	
A177		A295	
		Carboxylic Acid-NCBz(T) Bifunctional Building Blocks ((T) = TIDA Boronate)	
A180		A32	
A202		A124	
A203		A213	

A204		A215	
A230		A242	
A252		A265	
A306		A281	
		Excluded Building Blocks	
A311		A37	
A312		A60	
A313		A200	
		A223	
		A298	
		A299	

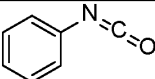
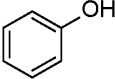
		A300	
		A340	

Table 2. Categories of Exemplary Molecular Building Blocks Shown in Table 1

Type of Block	Quantity
Commercial Halides	35
Commercial Amines	52
Commercial Carboxylic Acids	15
TIDA Boronates (Commercial Boronic Acids)	29
Custom Halides	23
Custom Amines	18
Custom Carboxylic Acid	1
Custom TIDA-Boronates	40
Halide-TIDA Boronates Bifunctional	42
Halide-NCBz(T) Bifunctional	62
Carboxylic Acid – TIDA Boronates Bifunctional	10
Carboxylic Acid-NCBz(T) Bifunctional	7
Excluded (isocyanates and phenols)	7

Example 2 – Exemplary Modular Synthesis of Trimers*Trimer Synthesis*

Trimers were synthesized through N–C bond formation using the Buchwald–Hartwig coupling reaction, facilitated by a palladium oxidative addition complex (OAC). Three distinct trimers were synthesized manually, all sharing a commercially available dimeric building block (**compound 4**, CAS Number 66521-66-2), which serves as a core component of imatinib. Structural diversity was introduced by varying the third building block in each

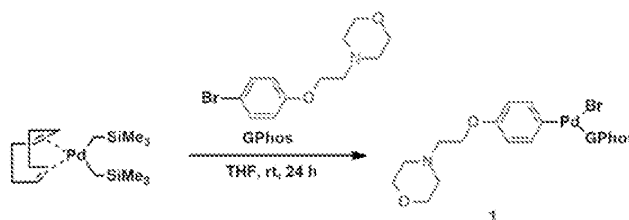
synthesis. The building blocks used in this procedure were identified using the molecular modularization algorithm. The detailed synthetic procedure is described below.

Synthesis of Palladium Oxidative Addition Complex (OAC)

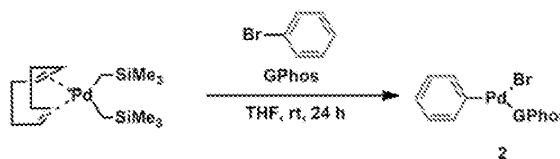
General Procedure

A reaction tube equipped with a magnetic stir bar was charged with aryl bromide (1.0 equiv.) and transferred into a glovebox. Inside the glovebox, GPhos (Compounds **1–3**) (1.0 equiv.) and (COD)Pd(CH₂TMS)₂ (1.0 equiv., thermally sensitive) were added, followed by the addition of anhydrous THF as the solvent. The tube was sealed with a PTFE septum cap and removed from the glovebox. The reaction mixture was stirred at room temperature (500 rpm) for 24 hours. After 24 hours, the tube was exposed to air, and the solvent was evaporated under reduced pressure. The resulting solid was subjected to high vacuum for 2 hours to remove any residual volatiles. The obtained solid was used directly in the subsequent step of the synthesis.

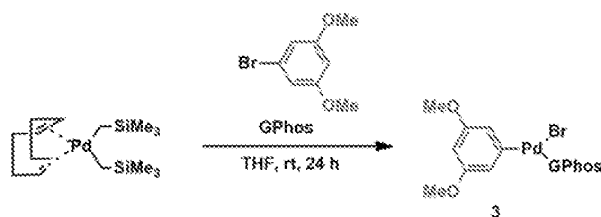
Synthesis of Oxidative Addition Complex (OAC) 1



Synthesis of Oxidative Addition Complex (OAC) 2

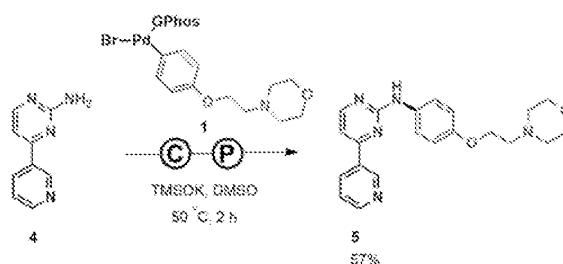


Synthesis of Oxidative Addition Complex (OAC) 3



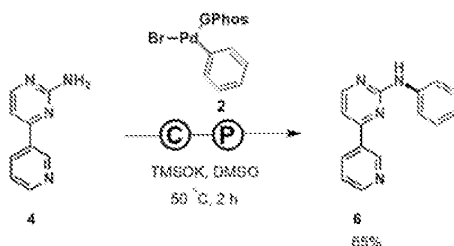
Manual synthesis of Trimers 5, 6, and 7

Synthesis of N-(4-(2-morpholinoethoxy)phenyl)-4-(pyridin-3-yl)pyrimidin-2-amine (5)



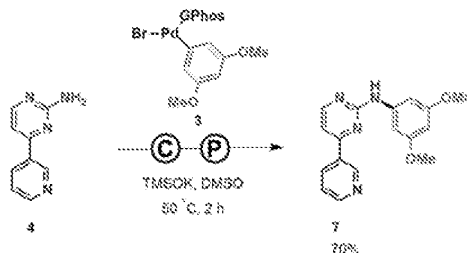
In an argon-filled glovebox, an oven-dried 20 mL vial equipped with a magnetic stir bar was charged with compound **1** (232 mg, 0.25 mmol, 1.0 equiv), compound **4** (43 mg, 0.25 mmol, 1.0 equiv), and TMSOK (38 mg, 0.3 mmol, 1.2 equiv). The vial was sealed with a PTFE septum cap. The sealed vial was removed from the glovebox and placed under a positive pressure of argon via a needle connected to a Schlenk line. Anhydrous, nitrogen-sparged DMSO (3 mL) was added to the vial, and the reaction mixture was stirred at 50 °C for 2 hours. After completion, the reaction mixture was diluted with ethyl acetate, and a brine solution was added. The aqueous layer was extracted 2–3 times with ethyl acetate. The combined organic layers were then washed with chilled water (3 times) followed by brine. The organic phase was dried over anhydrous sodium sulfate, filtered, and concentrated under reduced pressure using a rotary evaporator to afford the crude product as a brown solid. Purification of the crude product was performed using medium-pressure liquid chromatography (MPLC) with an ethyl acetate/methanol solvent system. (53 mg, 57% yield).

Synthesis of N-phenyl-4-(pyridin-3-yl)pyrimidin-2-amine (6)



In an argon-filled glovebox, an oven-dried 20 mL vial equipped with a magnetic stir bar was charged with compound **2** (200 mg, 0.25 mmol, 1.0 equiv), compound **4** (43 mg, 0.25 mmol, 1.0 equiv), and TMSOK (38 mg, 0.3 mmol, 1.2 equiv). The vial was sealed with a PTFE septum cap. The sealed vial was removed from the glovebox and placed under a positive pressure of argon via a needle connected to a Schlenk line. Anhydrous, nitrogen-sparged DMSO (3 mL) was added to the vial, and the reaction mixture was stirred at 50 °C for 2 hours. After completion, the reaction mixture was diluted with ethyl acetate, and a brine solution was added. The aqueous layer was extracted 2–3 times with ethyl acetate. The combined organic layers were then washed with chilled water (3 times) followed by brine. The organic phase was dried over anhydrous sodium sulfate, filtered, and concentrated under reduced pressure using a rotary evaporator to afford the crude product as a brown solid. Purification of the crude product was performed using medium-pressure liquid chromatography (MPLC) with hexane/ethyl acetate solvent system. (40 mg, 65% yield).

Synthesis of N-(3,5-dimethoxyphenyl)-4-(pyridin-3-yl)pyrimidin-2-amine (7)



In an argon-filled glovebox, an oven-dried 20 mL vial equipped with a magnetic stir bar was charged with compound **3** (215 mg, 0.25 mmol, 1.0 equiv), compound **4** (43 mg, 0.25 mmol, 1.0 equiv), and TMSOK (38 mg, 0.3 mmol, 1.2 equiv). The vial was sealed with a PTFE septum cap. The sealed vial was removed from the glovebox and placed under a positive pressure of argon via a needle connected to a Schlenk line. Anhydrous, nitrogen-sparged DMSO (3 mL) was added to the vial, and the reaction mixture was stirred at 50 °C for 2 hours. After completion, the reaction mixture was diluted with ethyl acetate, and a brine

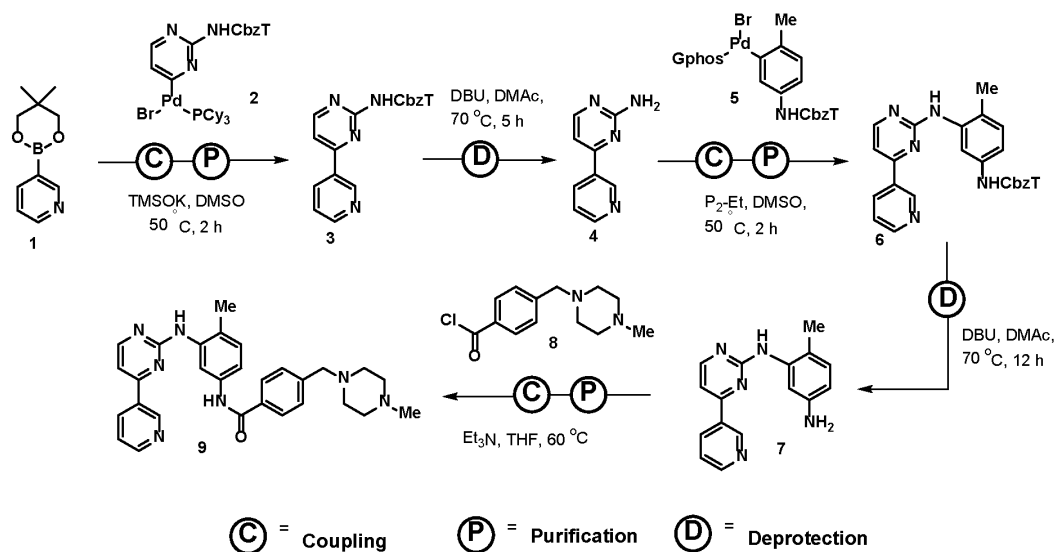
solution was added. The aqueous layer was extracted 2–3 times with ethyl acetate. The combined organic layers were then washed with chilled water (3 times) followed by brine. The organic phase was dried over anhydrous sodium sulfate, filtered, and concentrated under reduced pressure using a rotary evaporator to afford the crude product as a brown solid. Purification of the crude product was performed using medium-pressure liquid chromatography (MPLC) with hexane/ethyl acetate solvent system. (53 mg, 70% yield).

Example 3 – Exemplary Modular Synthesis of Imatinib

Synthesis of Imatinib through Oxidative Addition Complex Enabled Iterative Cross-Coupling

Synthesis of small molecule kinase inhibitor Imatinib of which the representative building blocks are connected through C-C or C-N bond has been accomplished by iterative cross coupling promoted by pre-synthesized oxidative addition complexes. The building blocks used in this procedure were identified by the molecular modularization algorithm. In the first step, Suzuki-Mayura coupling performed between pyridyl boronate ester **1** with prior synthesized oxidative addition complex **2** from the appropriate bromide and tricyclohexylphosphine ligand generated the coupled product **3** with very high yield of 84%. Unique polarity of the product **3** due presence of CbzTIDA functional group enabled a simple catch and release purification of the product by opting specific solvent sequence in silica. In the following step, the deprotection of the CbzTIDA functional group was achieved by simple addition of DBU in DMAc at 70 °C. The biproduct benzyl TIDA was isolated through a similar catch and release purification by choosing appropriate solvent and the product **3** was isolated in very good yield of 80%. In the next step the C-N bond was formed by Buchwald-Hartwig coupling with dimeric amine **4** and separately prepared oxidative addition complex **5** from the specific bromide and Gphos phosphine ligand. The reaction performed in DMSO and P2-Et base produced the trimeric scaffold **6** in 78% yield. A similar catch and release protocol was pursued to purify the product. The deprotection of CbzTIDA from trimeric building block **6** was again achieved by DBU in DMAc at 70 °C and the purification was accomplished in a similar fashion to isolate trimeric block **7** having free amine. In the final step, the amide bond formation was performed with the 4-((4-methylpiperazin-1-yl)methyl)benzoyl chloride **8** in the presence of triethyl amine in THF at 60 °C. A basic work up gave us the final Imatinib in almost pure form. Nevertheless, the crude Imatinib was finally purified by HPLC to obtain the compound in ultrapure form with 86% yield. This synthetic

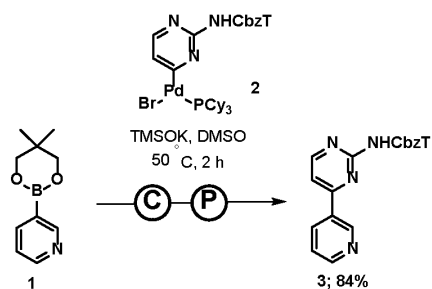
protocol using separately synthesized oxidative addition complexes enhanced the overall yield (25%) as compared to the yield through catalytic method (yield 9.9%).



Scheme 1. Schematic diagram for the synthesis of Imatinib

Synthetic Procedures:

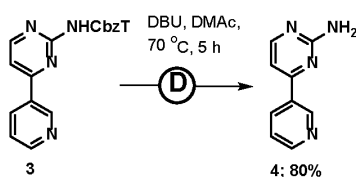
Synthesis of 4-(3,3,4,5,5-pentamethyl-2,6-dioxotetrahydro-2H-4,8[4,1,3,2]oxazaborolo[2,3-b][1,3,2]oxazaborol-8-yl)benzyl (4-(pyridin-3-yl)pyrimidin-2-yl)carbamate (1)



An oven dried 8 mL vial equipped with a stir bar was charged with 3-(5,5-dimethyl-1,3,2-dioxaborinan-2-yl) pyridine **1** (73 mg, 0.38 mmol, 1.5 equiv), oxidative addition complex **2** (226 mg, 0.25 mmol, 1.0 equiv.), TMSOK (38.4 mg, 0.3 mmol, 1.2 equiv). The vial was evacuated and backfilled with nitrogen using Schenk line 5 times followed by anhydrous DMSO (2.5 mL) was added via syringe. The vial was placed in a pre-equilibrated

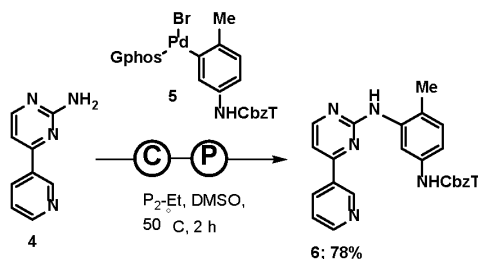
aluminum heating block set to 50 °C and stirred at 500 rpm for 2 h. The vial was removed from the heating block and allowed to cool to room temperature. The reaction mixture was diluted with methyl *tert*-butyl ether (10 mL) and filtered through a silica plug. The plug was washed with hexanes (100 mL), *tert*-butyl ether (100 mL) and EtOAc (100 mL), and then the eluents were discarded. The silica plug was then washed with acetone (100 mL) to elute the product. The acetone eluent was concentrated via rotary evaporation to yield **3** as a white solid (108.6 mg, 84% yield).

Deprotection of 3 to prepare 4-(pyridin-3-yl)pyrimidin-2-amine (4)



To an 8 mL vial with a stir bar was added **3** (108 mg, 0.21 mmol, 1.0 equiv) and dissolved in dimethyl acetamide (0.8 mL) then DBU (63.5 mg, 0.42 mmol, 2.0 equiv) was added by syringe. The reaction mixture was stirred at 70 °C for 5 hours. The reaction mixture was diluted with ether (10 mL) and filtered through a silica plug. The plug was washed with hexanes (50 mL), Et₂O (50 mL) and EtOAc (50 mL) and then the eluent was discarded. The silica plug was washed with acetone (100 mL) to release the product. The acetone eluent was then concentrated via rotary evaporation to yield **4** as a white solid (28.9 mg, 80% yield).

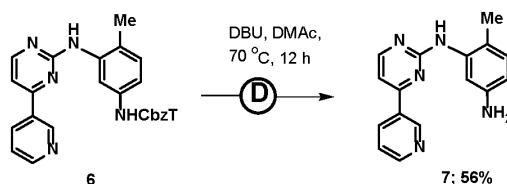
Synthesis of 4-(3,3,4,5,5-pentamethyl-2,6-dioxotetrahydro-2H-4[1,3,2]oxazaborolo[2,3-b][1,3,2]oxazaborol-8-yl)benzyl (4-methyl-3-((4-(pyridin-3-yl)pyrimidin-2-yl)amino)phenyl)carbamate (6)



An oven dried 8 mL vial equipped with a stir bar was charged with 4-(pyridin-3-yl)pyrimidin-2-amine **4** (28.9 mg, 0.17 mmol, 1.0 equiv), oxidative addition complex **5** (202 mg, 0.17 mmol, 1.0 equiv.), P₂-Et (86.4 mg, 0.26 mmol, 1.5 equiv). The vial was evacuated and backfilled with nitrogen using Schenk line 5 times followed by anhydrous DMSO (1.7

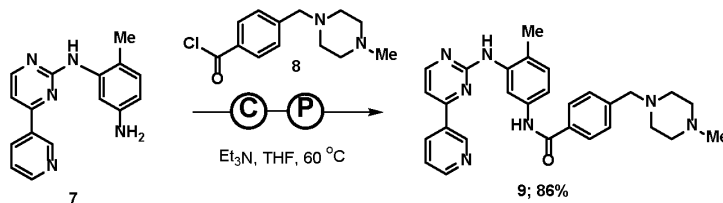
mL) was added via syringe. The vial was placed in a pre-equilibrated aluminum heating block set to 50 °C and stirred at 500 rpm for 2 h. The vial was removed from the heating block and allowed to cool to room temperature. The reaction mixture was diluted with methyl *tert*-butyl ether (10 mL) and filtered through a silica plug. The plug was washed with hexanes (50 mL), *tert*-butyl ether (50 mL) and EtOAc (100 mL), and then the eluents were discarded. The silica plug was then washed with acetone (50 mL) to elute the product. The acetone eluent was concentrated via rotary evaporation to yield **6** as a white solid (82.5 mg, 78% yield).

Deprotection of 6 to prepare 6-methyl-N¹-(4-(pyridin-3-yl)pyrimidin-2-yl)benzene 1,3-diamine (7)



To an 8 mL vial with a stir bar was added **6** (82.5 mg, 0.13 mmol, 1.0 equiv) and dissolved in dimethyl acetamide (0.5 mL) then DBU (39.5 mg, 0.26 mmol, 2.0 equiv) was added by syringe. The reaction mixture was stirred at 70 °C for 12 hours. The reaction mixture was diluted with ether (10 mL) and filtered through a silica plug. The plug was washed with hexanes (50 mL), Et₂O (50 mL) and EtOAc (50 mL) and then the eluent was discarded. The silica plug was washed with acetone (50 mL) to release the product. The acetone eluent was then concentrated via rotary evaporation to yield **7** as a white solid (20.2 mg, 56% yield).

Preparation of N-(4-methyl-3-((4-(pyridin-3-yl)pyrimidin-2-yl)amino)phenyl)-4-((4-methylpiperazin-1-yl)methyl)benzamide 9 (Imatinib)



Amine **7** (20.2 mg, 0.072 mmol, 1 equiv) and 4-(4-methyl-piperazinomethyl)-benzoyl chloride **8** (22 mg, 0.086 mmol, 1.2 equiv) were added to an oven-dried 8 mL vial with a stir bar. The vial was sealed with a PTFE septum cap, and then evacuated and backfilled with nitrogen (3x) via a needle attached to a Schlenk line. Anhydrous THF (1 mL) was added by syringe followed by addition of Et₃N (30.6 mg, 0.3 mmol, 4.0 equiv). The mixture was then

placed in a pre-equilibrated 60° C heat block and stirred for 4 hours. The solvent was concentrated on a rotary evaporator. Water (6 ml) was added, and then the solution was basified to pH 10 with 6 N NaOH. The aqueous solution was extracted with dichloromethane (3×10 mL), and the collected organic phase dried over anhydrous sodium sulfate. The crude reaction mixture was resuspended in 1:1 H₂O:ACN 0.1% Formic Acid for preparative HPLC purification with an Agilent 10 Prep-C18 250 x 50.0mm column. The final product imatinib was isolated as a white solid (27.2 mg, 86% yield).

INCORPORATION BY REFERENCE

All US patents and US and PCT patent application publications mentioned herein are hereby incorporated by reference in their entirety as if each individual publication or patent was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control.

EQUIVALENTS

While specific embodiments of the subject invention have been discussed, the above specification is illustrative and not restrictive. Many variations of the invention will become apparent to those skilled in the art upon review of this specification and the claims below. The full scope of the invention should be determined by reference to the claims, along with their full scope of equivalents, and the specification, along with such variations.

CLAIMS

What is claimed is:

1. A method of forming a C-C, C-N, or C-O bond between two molecular building blocks, the method comprising:
 - providing a first molecular building block; wherein the first molecular building block comprises a first reactive moiety;
 - providing a second molecular building block; wherein the second molecular building block comprises a second reactive moiety and a third reactive moiety;
 - identifying suitable reaction conditions under which the first reactive moiety of the first molecular building block and the second reactive moiety of the second molecular building block will react to form the C-C, C-N, or C-O bond; and
 - combining the first molecular building block and the second molecular building block under the suitable reaction conditions;
 - thereby forming a first coupled unit comprising the C-C, C-N, or C-O bond between the first molecular building block and the second molecular building block.
2. The method of claim 1, further comprising:
 - providing a third molecular building block; wherein the third molecular building block comprises a fourth reactive moiety;
 - identifying suitable reaction conditions under which the fourth reactive moiety of the third molecular building block and the first coupled unit will react to form a C-C, C-N, or C-O bond; and
 - combining the third molecular building block and the first coupled unit under the suitable reaction conditions;
 - thereby forming a second coupled unit comprising the C-C, C-N, or C-O bond between the third molecular building block and the first coupled unit.
3. The method of claim 2, wherein the first coupled unit comprises the third reactive moiety, and the third reactive moiety reacts with the fourth reactive moiety on the third molecular building block to form the second coupled unit.

4. The method of any one of claims 1-3, wherein the reaction between the first molecular building block and the second molecular building block is selected from the group consisting of Suzuki coupling, Buchwald C-N coupling, Buchwald C-O coupling, and amide coupling.
5. The method of any one of claims 2-4, wherein the reaction between the first coupled unit and the third molecular building block is selected from the group consisting of Suzuki coupling, Buchwald C-N coupling, Buchwald C-O coupling, amide coupling, and isocyanate-amine urea formation.
6. The method of claim 4 or 5, wherein the Suzuki coupling is selected from the group consisting of aryl-aryl Suzuki coupling, aryl-vinyl Suzuki coupling, and vinyl-vinyl Suzuki coupling.
7. The method of claim 4 or 5, wherein the Buchwald C-N coupling is selected from the group consisting of aryl electrophile-aryl amine Buchwald C-N coupling, aryl electrophile-primary amine Buchwald C-N coupling, aryl electrophile-heterocyclic amine Buchwald C-N coupling, aryl electrophile-heteroaryl amine Buchwald C-N coupling, and aryl electrophile-secondary amine Buchwald C-N coupling.
8. The method of claim 7, wherein the Buchwald C-O coupling is aryl electrophile-phenol Buchwald C-O coupling.
9. The method of any one of claims 1-8, further comprising:
 - identifying a first precursor molecule that does not comprise the first reactive moiety; and
 - transforming the first precursor molecule to introduce the first reactive moiety; thereby forming the first molecular building block.
10. The method of any one of claims 1-8, further comprising:
 - identifying a first precursor molecule comprising the first reactive moiety and one or more additional reactive moieties; and

protecting the one or more additional reactive moieties comprised by the first precursor molecule;

thereby forming the first molecular building block.

11. The method of any one of claims 1-10, further comprising:

identifying a second precursor molecule that does not comprise the second reactive moiety and/or the third reactive moiety; and

transforming the second precursor molecule to introduce the second reactive moiety and/or the third reactive moiety;

thereby forming the second molecular building block.

12. The method of any one of claims 1-10, further comprising:

identifying a second precursor molecule comprising the second reactive moiety, the third reactive moiety, and one or more additional reactive moieties; and

protecting the one or more additional reactive moieties comprised by the second precursor molecule;

thereby forming the second molecular building block.

13. The method of any one of claims 2-12, further comprising:

identifying a third precursor molecule that does not comprise the fourth reactive moiety; and

transforming the third precursor molecule to introduce the fourth reactive moiety;

thereby forming the third molecular building block.

14. The method of any one of claims 2-12, further comprising:

identifying a third precursor molecule comprising the fourth reactive moiety and one or more additional reactive moieties; and

protecting the one or more additional reactive moieties comprised by the third precursor molecule;

thereby forming the third molecular building block.

15. The method of any one of claims 2-14, wherein the first reactive moiety, the second reactive moiety, the third reactive moiety, and the fourth reactive moiety are, independently

at each occurrence, selected from the group consisting of an electrophilic moiety, a nucleophilic moiety, and a moiety which, upon reaction with a metal catalyst, becomes a nucleophilic moiety.

16. The method of any one of claims 1-15, further comprising:

determining whether the first molecular building block and the second molecular building block are chemically compatible for the purpose of forming the C-C, C-N, or C-O bond; wherein the first molecular building block and the second molecular building block are chemically compatible if:

the first reactive moiety is a nucleophilic moiety or a moiety which, upon reaction with a metal catalyst, becomes a nucleophilic moiety, and the second reactive moiety is an electrophilic moiety; or

the first reactive moiety is an electrophilic moiety, and the second reactive moiety is a nucleophilic moiety or a moiety which, upon reaction with a metal catalyst, becomes a nucleophilic moiety.

17. The method of any one of claims 2-16, further comprising:

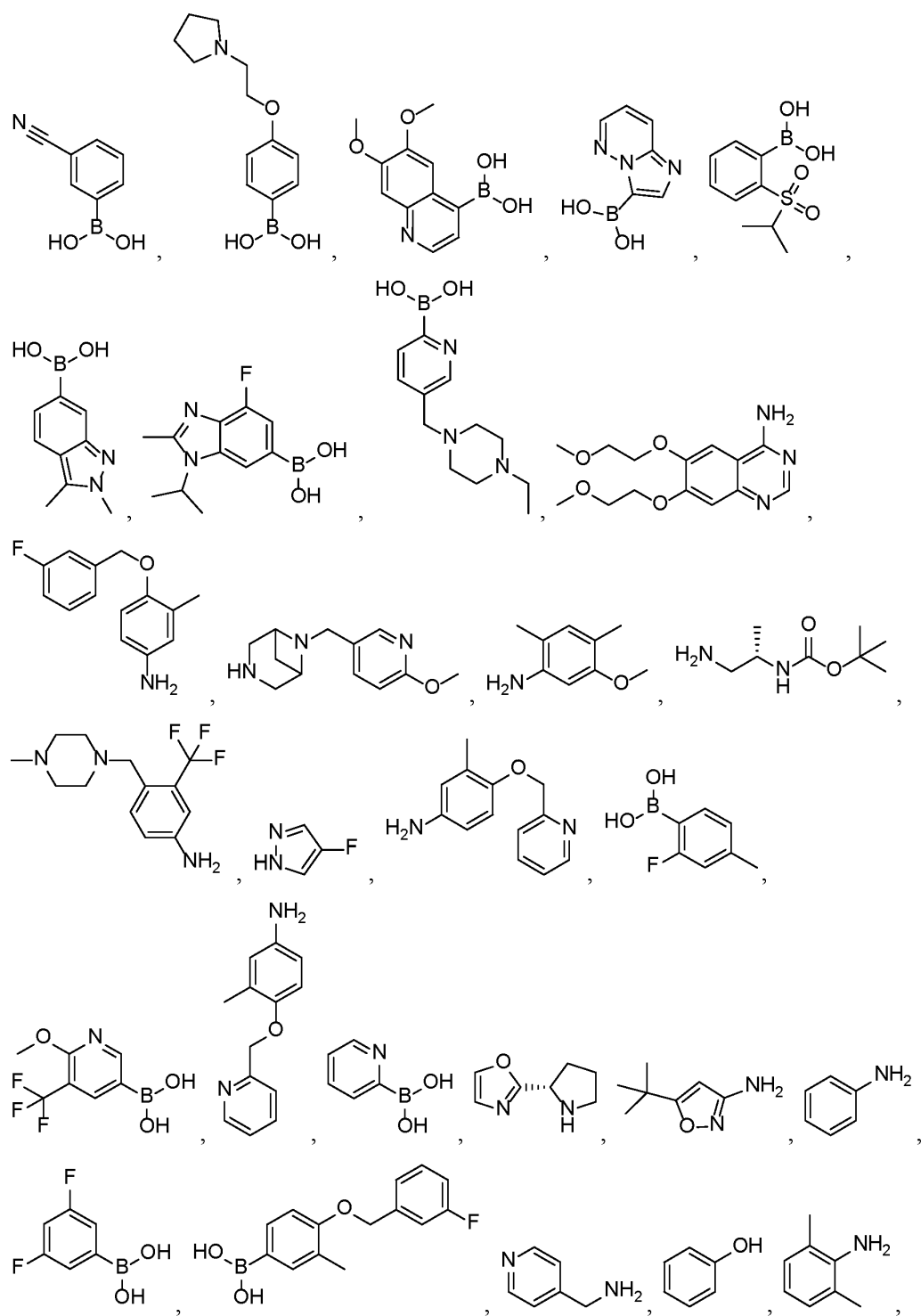
determining whether the third molecular building block and the first coupled unit are chemically compatible for the purpose of forming the C-C, C-N, or C-O bond; wherein the third molecular building block and the first coupled unit are chemically compatible if:

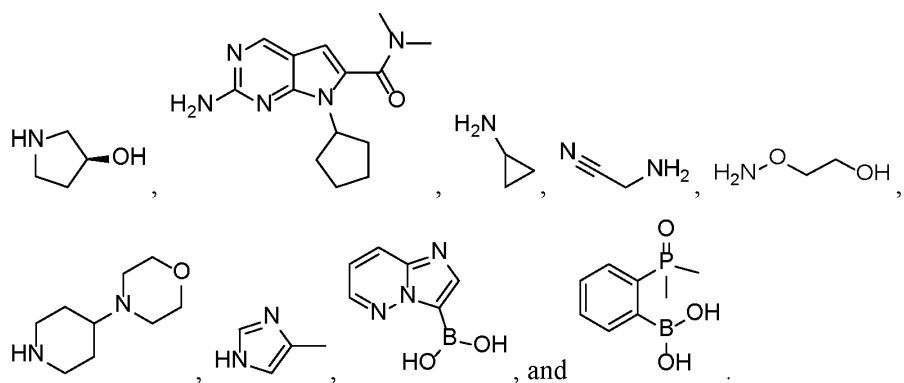
the third reactive moiety is a nucleophilic moiety or a moiety which, upon reaction with a metal catalyst, becomes a nucleophilic moiety, and the fourth reactive moiety is an electrophilic moiety; or

the third reactive moiety is an electrophilic moiety, and the fourth reactive moiety is a nucleophilic moiety or a moiety which, upon reaction with a metal catalyst, becomes a nucleophilic moiety.

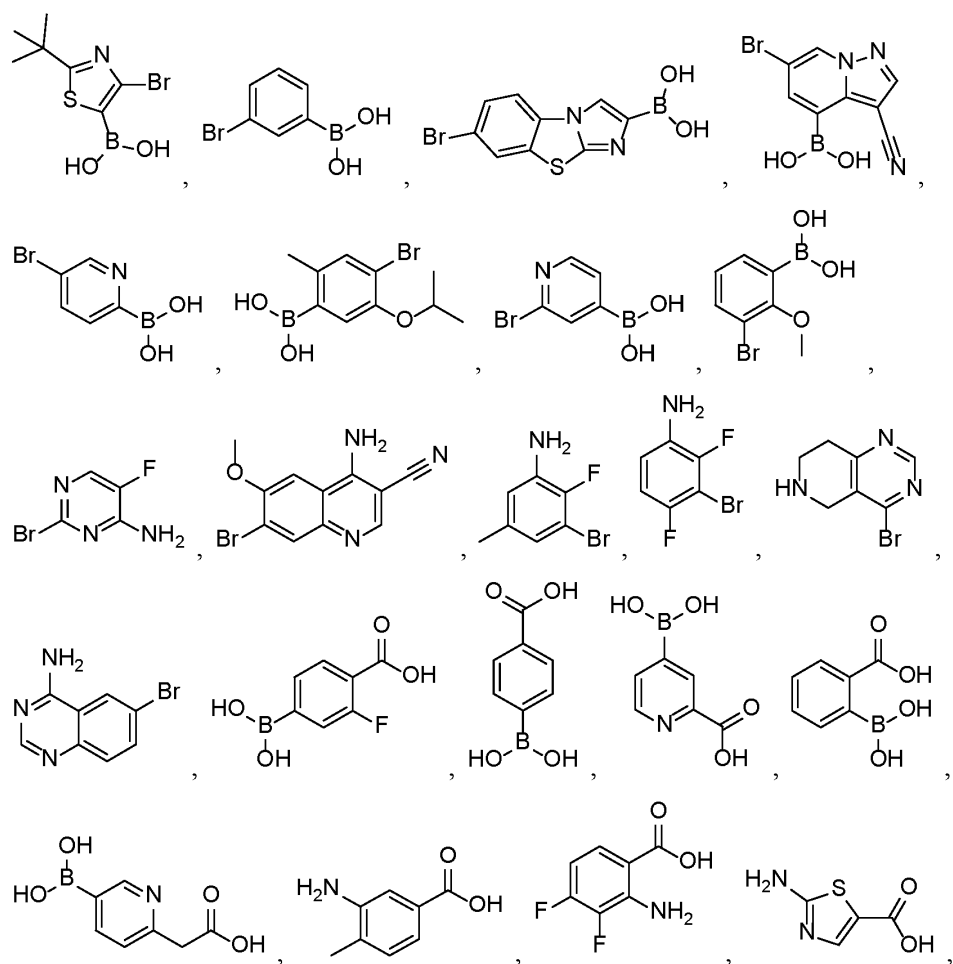
18. The method of any one of claims 2-17, wherein the first reactive moiety, the second reactive moiety, the third reactive moiety, or the fourth reactive moiety is, independently at each occurrence, a nucleophilic moiety or a moiety which, upon reaction with a metal catalyst, becomes a nucleophilic moiety.

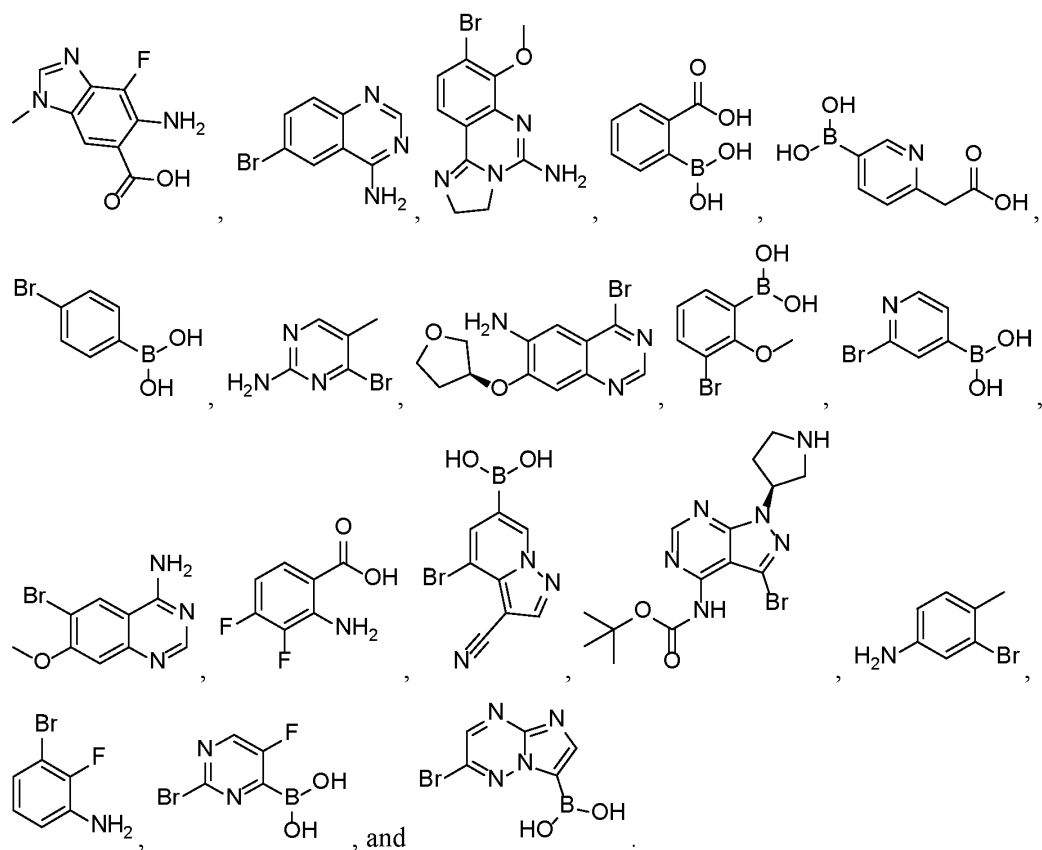
19. The method of claim 18, wherein the first reactive moiety, the second reactive moiety, the third reactive moiety, or the fourth reactive moiety is, independently at each occurrence, selected from the group consisting of: -B(OH)₂, -B(pin), -B(dan), -B(neop), -B(MIDA), -B(TIDA), -B(CbzTIDA), -NR¹H, and -OH; and R¹ is selected from the group consisting of hydrogen and a protecting group.
20. The method of any one of claims 2-17, wherein the first reactive moiety, the second reactive moiety, or the third reactive moiety is, independently at each occurrence, an electrophilic moiety.
21. The method of claim 20, wherein the first reactive moiety, the second reactive moiety, or the third reactive moiety is, independently at each occurrence, selected from the group consisting of: -C(O)OR², -Cl, -Br, -I, -OMs, -OTs, and -OTf; and R² is selected from the group consisting of hydrogen and alkyl.
22. The method of any one of claims 2-17, wherein the fourth reactive moiety is an electrophilic moiety.
23. The method of claim 22, wherein the fourth reactive moiety is selected from the group consisting of: -C(O)OR², -Cl, -Br, -I, -OMs, -OTs, -OTf, and -N=C=O; and R² is selected from the group consisting of hydrogen and alkyl.
24. The method of any one of claims 1-23, wherein the first molecular building block is selected from the group consisting of:



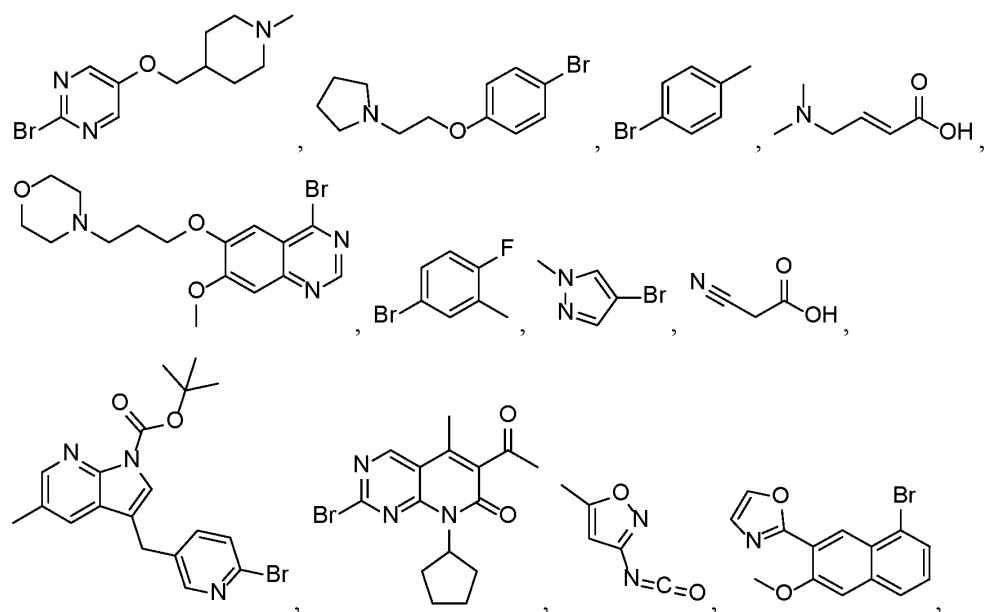


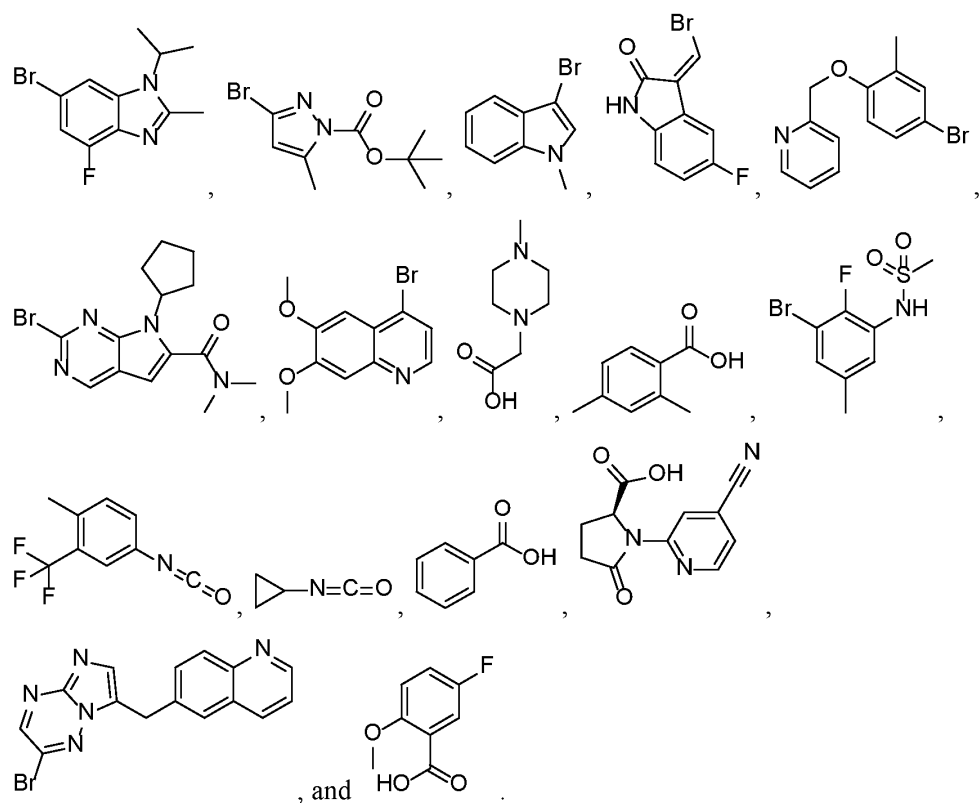
25. The method of any one of claims 1-24, wherein the second molecular building block is selected from the group consisting of:



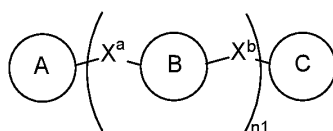


26. The method of any one of claims 2-25, wherein the third molecular building block is selected from the group consisting of:



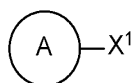


27. A method of making a compound of formula (I):

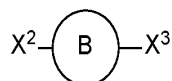


(I);

comprising combining a compound of formula (II):

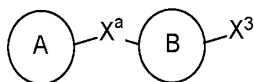

$$(II);$$

and a compound of formula (III):



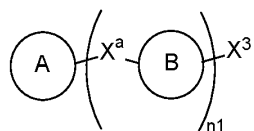
(III);

under a first set of suitable reaction conditions under which the compound of formula (II) and the compound of formula (III) will react to form a C-C, C-N, or C-O bond; thereby forming a compound of formula (IV):



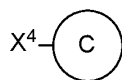
(IV);

optionally combining the compound of Formula (IV) with one or more additional compounds of Formula (III) to form a compound of Formula (IVa):



(IVa); and

combining the compound of formula (IV) or Formula (IVa) and a compound of formula (V):



(V);

under a second set of suitable reaction conditions under which the compound of formula (IV) or the compound of Formula (IVa) and the compound of formula (V) will react to form a C-C, C-N, or C-O bond; thereby forming the compound of formula (I);

wherein:

X^1 is selected from the group consisting of hydrogen, -B(OH)₂, -B(pin), -B(dan), -B(neop), -B(MIDA), -B(TIDA), -B(CbzTIDA), -NR¹H, -C(O)OR², -OH, -Cl, -Br, -I, -OMs, -OTs, and -OTf;

X^2 , independently for each occurrence, is selected from the group consisting of -B(OH)₂, -B(pin), -B(dan), -B(neop), -B(MIDA), -B(TIDA), -B(CbzTIDA), -NR¹H, -C(O)OR², -OH, -Cl, -Br, -I, -OMs, -OTs, and -OTf;

X^3 , independently for each occurrence, is selected from the group consisting of -B(OH)₂, -B(pin), -B(dan), -B(neop), -B(MIDA), -B(TIDA), -B(CbzTIDA), -NR¹H, -C(O)OR², -OH, -Cl, -Br, -I, -OMs, -OTs, and -OTf;

X⁴ is selected from the group consisting of -N=C=O, -B(OH)₂, -B(pin), -B(dan), -B(neop), -B(MIDA), -B(TIDA), -B(CbzTIDA), -NR¹H, -C(O)OR², -OH, -Cl, -Br, -I, -OMs, -OTs, and -OTf;

R¹ is selected from the group consisting of hydrogen and a protecting group;

R² is selected from the group consisting of hydrogen and alkyl;

n1 is 1, 2, 3, 4, 5, 6, 7, 8, or 9;

A is selected from the group consisting of optionally substituted alkyl, optionally substituted alkenyl, optionally substituted aryl, optionally substituted heteroaryl, optionally substituted heterocyclyl, or optionally substituted cycloalkyl;

B, independently for each occurrence, is selected from the group consisting of optionally substituted aryl, optionally substituted heteroaryl, optionally substituted heterocyclyl, or optionally substituted cycloalkyl;

C is selected from the group consisting of optionally substituted alkyl, optionally substituted alkenyl, optionally substituted aryl, optionally substituted heteroaryl, optionally substituted heterocyclyl, or optionally substituted cycloalkyl;

X^a, independently for each occurrence, is selected from the group consisting of a bond,

-C(O)NH-, -NHC(O)-, -O-, and -NR¹-; and

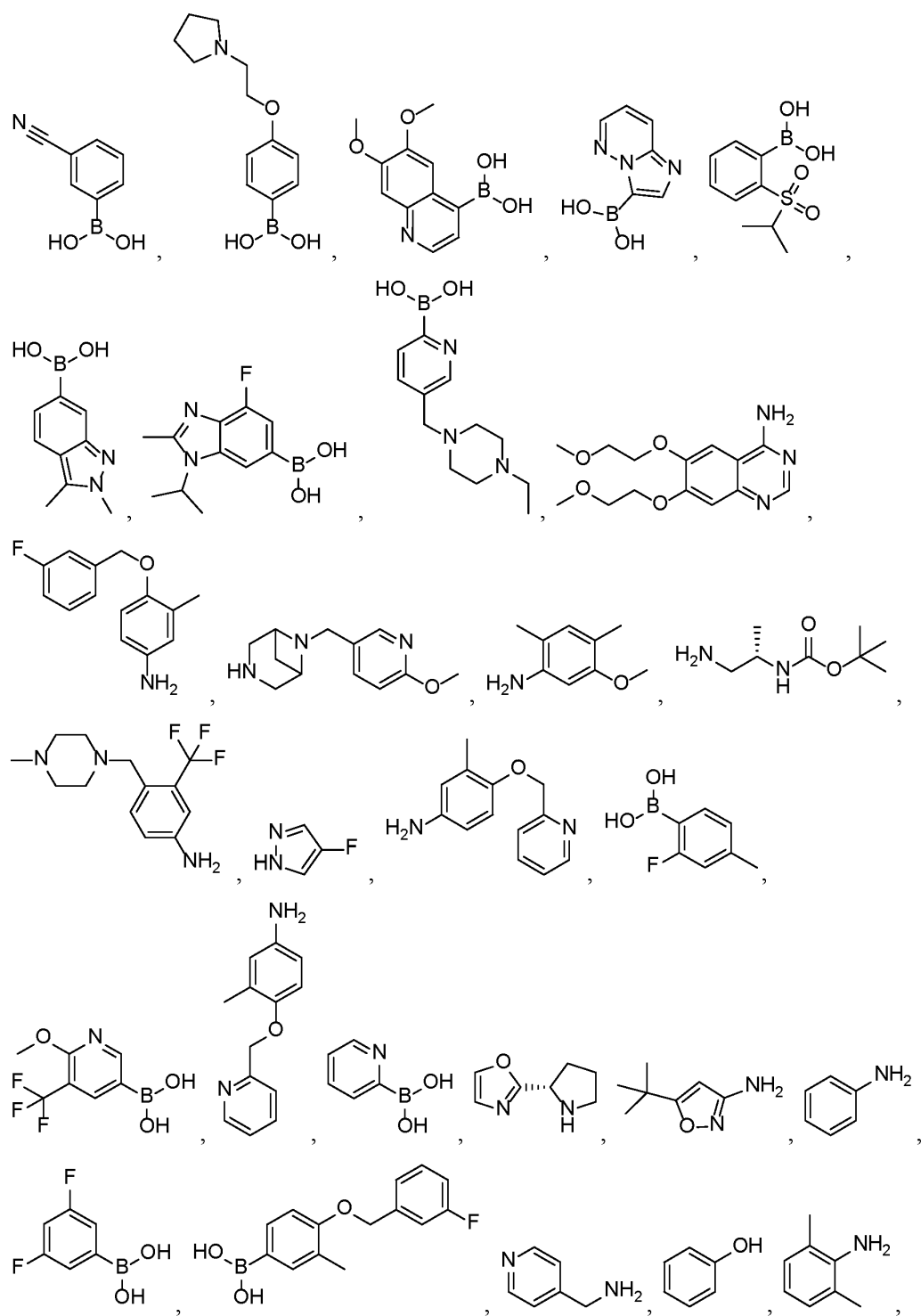
X^b, independently for each occurrence, is selected from the group consisting of a bond,

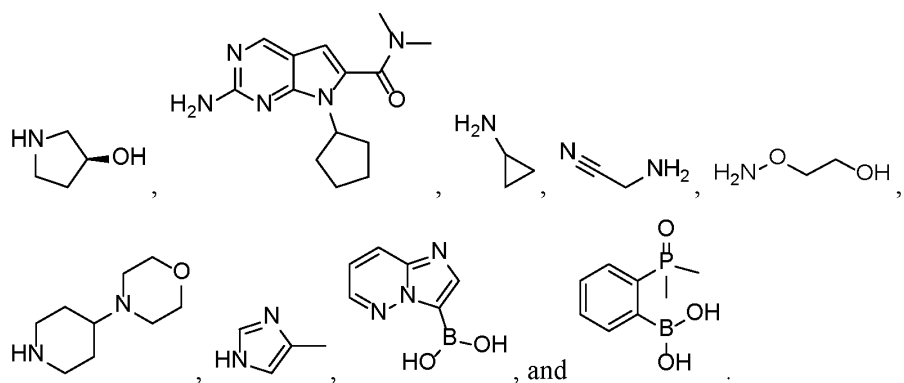
-C(O)NH-, -NHC(O)-, -NH(CO)NH-, -O-, and -NR¹-.

28. The method of claim 27, wherein X¹ is hydrogen.
29. The method of claim 27, wherein X¹ is -B(OH)₂.
30. The method of claim 27, wherein X¹ is -B(pin).
31. The method of claim 27, wherein X¹ is -B(TIDA).
32. The method of claim 27, wherein X¹ is -B(CbzTIDA).

33. The method of claim 27, wherein X^1 is $-NR^1H$.
34. The method of claim 27, wherein X^1 is $-OH$.
35. The method of any one of claims 27-34, wherein X^2 is $-Br$.
36. The method of any one of claims 27-34, wherein X^2 is $-C(O)OR^2$.
37. The method of any one of claims 27-36, wherein X^3 is $-B(OH)_2$.
38. The method of any one of claims 27-36, wherein X^3 is $-B(pin)$.
39. The method of any one of claims 27-36, wherein X^3 is $-B(TIDA)$.
40. The method of any one of claims 27-36, wherein X^3 is $-B(CbzTIDA)$.
41. The method of any one of claims 27-36, wherein X^3 is $-NR^1H$.
42. The method of any one of claims 27-36, wherein X^3 is $-OH$.
43. The method of any one of claims 27-42, wherein X^4 is $-Br$.
44. The method of any one of claims 27-42, wherein X^4 is $-C(O)OR^2$.
45. The method of any one of claims 27-42, wherein X^4 is $-N=C=O$.
46. The method of any one of claims 27-45, wherein R^1 is hydrogen.
47. The method of any one of claims 27-46, wherein R^2 is hydrogen.
48. The method of any one of claims 27-46, wherein R^2 is alkyl.
49. The method of any one of claims 27-48, wherein $n1$ is 1.

50. The method of any one of claims 27-48, wherein n1 is 2.
51. The method of any one of claims 27-48, wherein n1 is 3.
52. The method of any one of claims 27-48, wherein n1 is 4.
53. The method of any one of claims 27-48, wherein n1 is 5.
54. The method of any one of claims 27-48, wherein n1 is 6.
55. The method of any one of claims 27-54, wherein A is optionally substituted alkyl.
56. The method of any one of claims 27-54, wherein A is optionally substituted alkenyl.
57. The method of any one of claims 27-54, wherein A is optionally substituted aryl.
58. The method of any one of claims 27-54, wherein A is optionally substituted heteroaryl.
59. The method of any one of claims 27-54, wherein A is optionally substituted heterocyclyl.
60. The method of any one of claims 27-54, wherein A is optionally substituted cycloalkyl.
61. The method of any one of claims 27-54, wherein A is selected from the group consisting of:





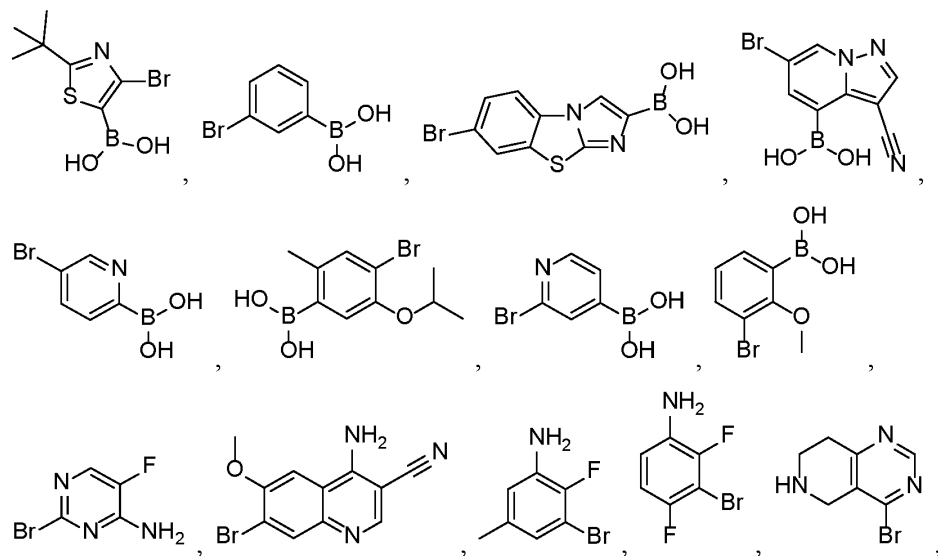
62. The method of any one of claims 27-61, wherein B is optionally substituted aryl.

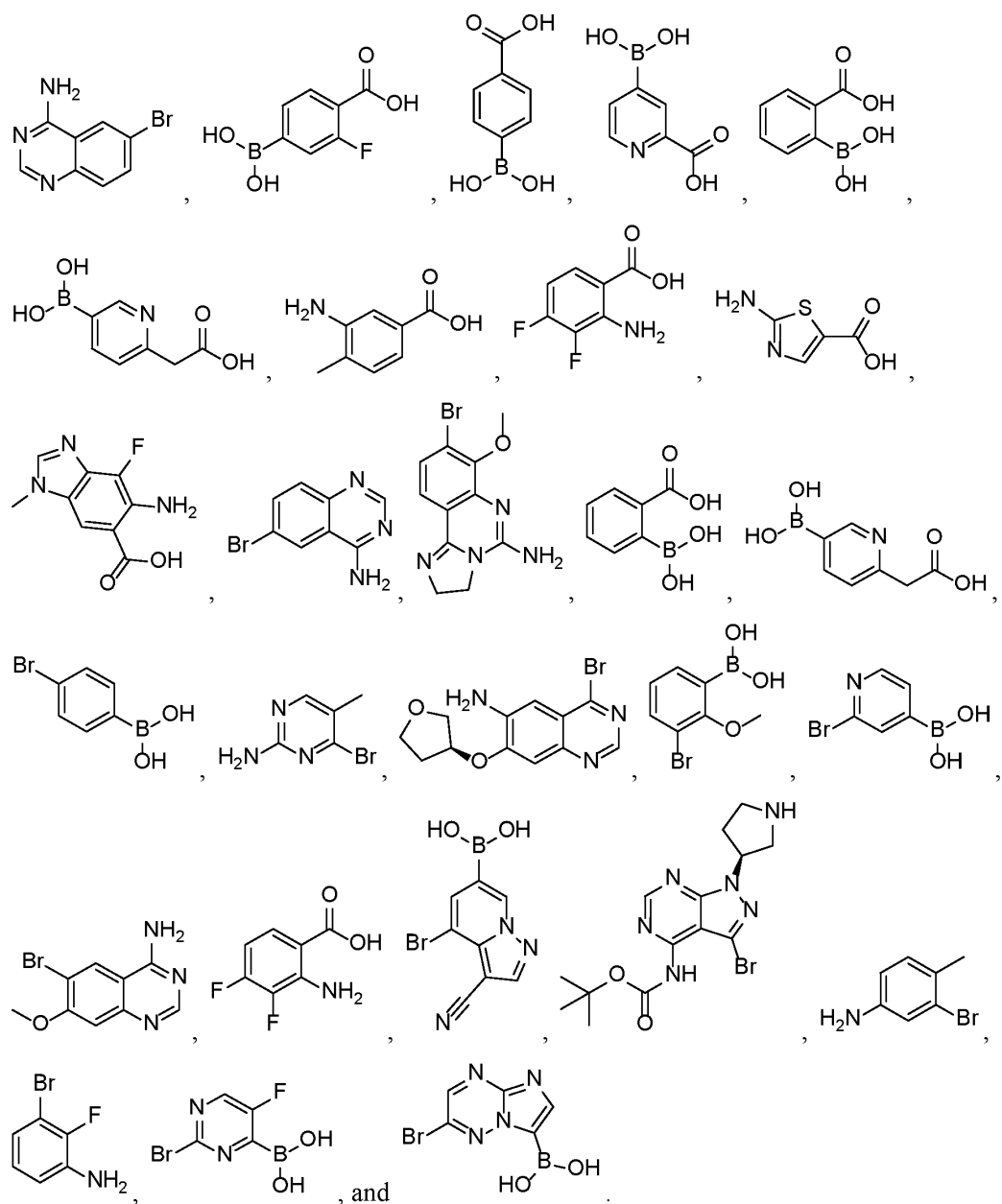
63. The method of any one of claims 27-61, wherein B is optionally substituted heteroaryl.

64. The method of any one of claims 27-61, wherein B is optionally substituted heterocyclyl.

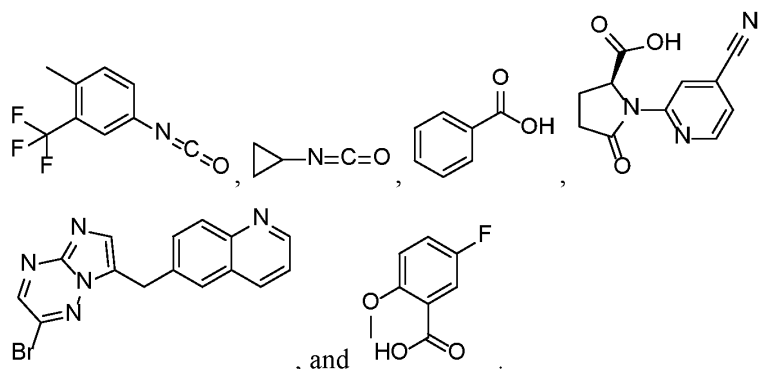
65. The method of any one of claims 27-61, wherein B is optionally substituted cycloalkyl.

66. The method of any one of claims 27-61, wherein B is selected from the group consisting of:





67. The method of any one of claims 27-66, wherein C is optionally substituted alkyl.
68. The method of any one of claims 27-66, wherein C is optionally substituted alkenyl.
69. The method of any one of claims 27-66, wherein C is optionally substituted aryl.



74. The method of any one of claims 27-73, wherein X^a is a bond.
75. The method of any one of claims 27-73, wherein X^a is $-C(O)NH-$.
76. The method of any one of claims 27-73, wherein X^a is $-NHC(O)-$.
77. The method of any one of claims 27-73, wherein X^a is $-NR^1-$.
78. The method of any one of claims 27-73, wherein X^a is $-O-$.
79. The method of any one of claims 27-78, wherein X^b is a bond.
80. The method of any one of claims 27-78, wherein X^b is $-C(O)NH-$.
81. The method of any one of claims 27-78, wherein X^b is $-NHC(O)-$.
82. The method of any one of claims 27-78, wherein X^b is $-NHC(O)NH-$.
83. The method of any one of claims 27-78, wherein X^b is $-NR^1-$.
84. The method of any one of claims 27-78, wherein X^b is $-O-$.