WHAT IS CLAIMED IS:

1. A computer-implemented method for determining one or more potential drug candidates during a drug discovery process, the method comprising:

applying one or more edit heuristics included in a plurality of edit heuristics to a query molecule specification to generate a plurality of derived molecule specifications;

performing one or more mapping operations on the plurality of derived molecule specifications to generate a plurality of mapped molecule specifications that are compressed digital representations of the plurality of derived molecule specifications;

performing one or more hash-based search operations on a mapped catalog of available molecules that can be acquired based on the plurality of mapped molecule specifications to determine the one or more potential drug candidates; and

causing one or more search results that include the one or more potential drug candidates to be outputted.

2. The computer-implemented method of claim 1, further comprising performing one or more mapping operations on a plurality of molecule specifications associated with a catalog of molecules to generate the mapped catalog of molecules.

3. The computer-implemented method of claim 1, wherein generating the plurality of derived molecule specifications comprises recursively applying the plurality of edit heuristics to the query molecule specification to generate a derivation tree that includes the plurality of derived molecule specifications.

4. The computer-implemented method of claim 1, wherein generating the plurality of derived molecule specifications comprises:

applying a first edit heuristic included in the plurality of edit heuristics to the query molecule specification to generate a first derived molecule specification; and

applying a second edit heuristic included in the plurality of edit heuristics to the first derived molecule specification to generate a second derived molecule specification.

5. The computer-implemented method of claim 1, wherein the plurality of edit heuristics includes at least one edit heuristic that, when applied to the query molecule specification, adds a Nitrogen or a substituent to the query molecule specification, removes a Nitrogen or a substituent from the query molecule specification, or repositions a substituent included in the query molecule specification to generate a derived molecule specification.

6. The computer-implemented method of claim 1, wherein the plurality of edit heuristics includes at least one edit heuristic that, when applied to the query molecule specification, repositions a Nitrogen included in the query molecule specification to generate a derived molecule specification that represents an isomer of a query molecule corresponding to the query molecule specification.

7. The computer-implemented method of claim 1, wherein the one or more mapping operations comprise applying a hash function to the plurality of derived molecule specifications to generate a plurality of array indices included in the plurality of mapped molecule specifications.

8. The computer-implemented method of claim 1, wherein performing the one or more hash-based search operations comprises performing a hash-based search on the mapped catalog of molecules based on a first mapped molecule specification included in the plurality of mapped molecule specifications to determine that a first derived molecule specification included in the plurality of derived molecule specifications matches a first molecule specification included in the mapped catalog of molecules.

9. The computer-implemented method of claim 8, further comprising:

performing another hash-based search on another mapped catalog of molecules to determine that the first derived molecule specification matches a second molecule specification included in the another mapped catalog of molecules;

determining that a first catalog of molecules corresponding to the mapped catalog of molecules has a first preference ranking that is lower than a

second catalog of molecules corresponding to the another mapped catalog of molecules; and

displaying on a computing device, via a graphical user interface, that a first derived molecule corresponding to the first derived molecule specification is a first potential drug candidate and is located in the second catalog of molecules.

10. The computer-implemented method of claim 1, wherein the query molecule specification represents a drug development candidate of interest associated with the drug discovery process.

11. One or more non-transitory computer readable media including instructions that, when executed by one or more processors, cause the one or more processors to determine one or more potential drug candidates during a drug discovery process by performing the steps of:

applying one or more edit heuristics included in a plurality of edit heuristics to a query molecule specification to generate a plurality of derived molecule specifications;

performing one or more mapping operations on the plurality of derived molecule specifications to generate a plurality of mapped molecule specifications that are compressed digital representations of the plurality of derived molecule specifications;

performing one or more hash-based search operations on a mapped catalog of available molecules that can be acquired based on the plurality of mapped molecule specifications to determine the one or more potential drug candidates; and

causing one or more search results that include the one or more potential drug candidates to be outputted.

12. The one or more non-transitory computer readable media of claim 11, further comprising performing one or more mapping operations on a plurality of molecule specifications associated with a catalog of molecules to generate the mapped catalog of molecules.

13. The one or more non-transitory computer readable media of claim 11, wherein

generating the plurality of derived molecule specifications comprises recursively applying the plurality of edit heuristics to the query molecule specification to generate a derivation tree that includes the plurality of derived molecule specifications.

14. The one or more non-transitory computer readable media of claim 11, wherein generating the plurality of derived molecule specifications comprises:

applying a first edit heuristic included in the plurality of edit heuristics to the query molecule specification to generate a first derived molecule specification; and

applying a second edit heuristic included in the plurality of edit heuristics to the query molecule specification to generate a second derived molecule specification.

15. The one or more non-transitory computer readable media of claim 11, wherein the plurality of edit heuristics includes at least one edit heuristic that, when applied to the query molecule specification, adds a Nitrogen or a substituent to the query molecule specification, removes a Nitrogen or a substituent from the query molecule specification, or repositions a substituent included in the query molecule specification to generate a derived molecule specification.

16. The one or more non-transitory computer readable media of claim 11, wherein the plurality of edit heuristics includes at least one edit heuristic that, when applied to the query molecule specification, replaces an Oxygen included in the query molecule specification with a Sulfur or replaces a Sulfur included in the query molecule specification with an Oxygen.

17. The one or more non-transitory computer readable media of claim 11, wherein the one or more mapping operations map the plurality of derived molecule specifications to a vector space to generate the plurality of mapped molecule specifications.

18. The one or more non-transitory computer readable media of claim 11, wherein performing the one or more hash-based search operations comprises performing a hash-based search on the mapped catalog of molecules based on a first mapped

molecule specification included in the plurality of mapped molecule specifications to determine that a first derived molecule specification include in the plurality of derived molecule specifications matches a first molecule specification included in the mapped catalog of molecules.

19. The one or more non-transitory computer readable media of claim 11, wherein the query molecule specification represents a drug development candidate of interest associated with the drug discovery process.

20. A system comprising:

one or more memories storing instructions; and

one or more processors coupled to the one or more memories that, when executing the instructions, perform the steps of:

applying one or more edit heuristics included in a plurality of edit heuristics to a query molecule specification to generate a plurality of derived molecule specifications;

performing one or more mapping operations on the plurality of derived molecule specifications to generate a plurality of mapped molecule specifications that are compressed digital representations of the plurality of derived molecule specifications;

performing one or more hash-based search operations on a mapped catalog of available molecules that can be acquired based on the plurality of mapped molecule specifications to determine one or more potential drug candidates; and

causing one or more search results that include the one or more potential drug candidates to be outputted.