What is claimed is:

1. A method of predicting viscosity of potential protein formulations, the method comprising:

classifying, by one or more processors of a computing system, a set of formulation descriptors as belonging to a specific group of a plurality of predetermined groups that each correspond to a different value range for viscosity, wherein the set of formulation descriptors includes (i) one or more molecule descriptors, (ii) one or more excipient descriptors, and (iii) one or more solution descriptors, and wherein classifying the set of formulation descriptors includes applying at least a first portion of the set of formulation descriptors as inputs to a first machine learning model;

selecting, by the one or more processors and based on the specific group, a second machine learning model from among a plurality of machine learning models that each correspond to a different one of the plurality of predetermined groups;

predicting, by the one or more processors, a value of viscosity by applying at least a second portion of the set of formulation descriptors as inputs to the second machine learning model; and

causing, by the one or more processors, the value of viscosity to be one or both of (i) displayed to a user, and (ii) stored in a memory.

2. The method of claim 1, wherein the set of formulation descriptors includes one or more numerical values.

3. The method of claim 1 or 2, wherein the set of formulation descriptors includes one or more categories.

4. The method of any one of claims 1 through 3, wherein the set of formulation descriptors includes one or more of (i) a molecule identifier, (ii) an antibody type, (iii) a target pH level, and (iv) a target protein concentration.

5. The method of any one of claims 1 through 4, wherein the set of formulation descriptors includes one or more of (i) molecular weight of an excipient, (ii) charge of the excipient, (iii) solubility of the excipient, (iv) bonds of the excipient, and (v) a type of the excipient.

6. The method of any one of claims 1 through 5, wherein the plurality of predetermined groups consists of three groups.

7. The method of any one of claims 1 through 6, wherein each of the plurality of machine learning models includes a regression and/or classification model.

8. The method of claim 7, wherein at least one of the plurality of machine learning models is a random forest model.

9. The method of any one of claims 1 through 8, wherein:

classifying the set of formulation descriptors includes applying the set of formulation descriptors as inputs to the first machine learning model; and

predicting the value of viscosity includes applying the set of formulation descriptors as inputs to the second machine learning model.

10. The method of any one of claims 1 through 9, further comprising:

receiving, by the one or more processors and via a first graphical user interface, the set of formulation descriptors from a user.

11. The method of claim 10, further comprising:

determining, by the one or more processors and based in part on the predicted value of viscosity, a rank-ordering of a plurality of candidate protein formulations; and

causing, by the one or more processors, an indication of the rank-ordering to be displayed via the first graphical user interface or a second graphical user interface.

12. A computing system comprising:

one or more processors; and

one or more memories storing instructions that, when executed by the one or more processors, cause the computing system to perform the method of any one of claims 1 through 11.

13. A non-transitory computer-readable medium storing instructions that, when executed by one or more processors of a computing system, cause the computing system to perform the method of any one of claims 1 through 11.