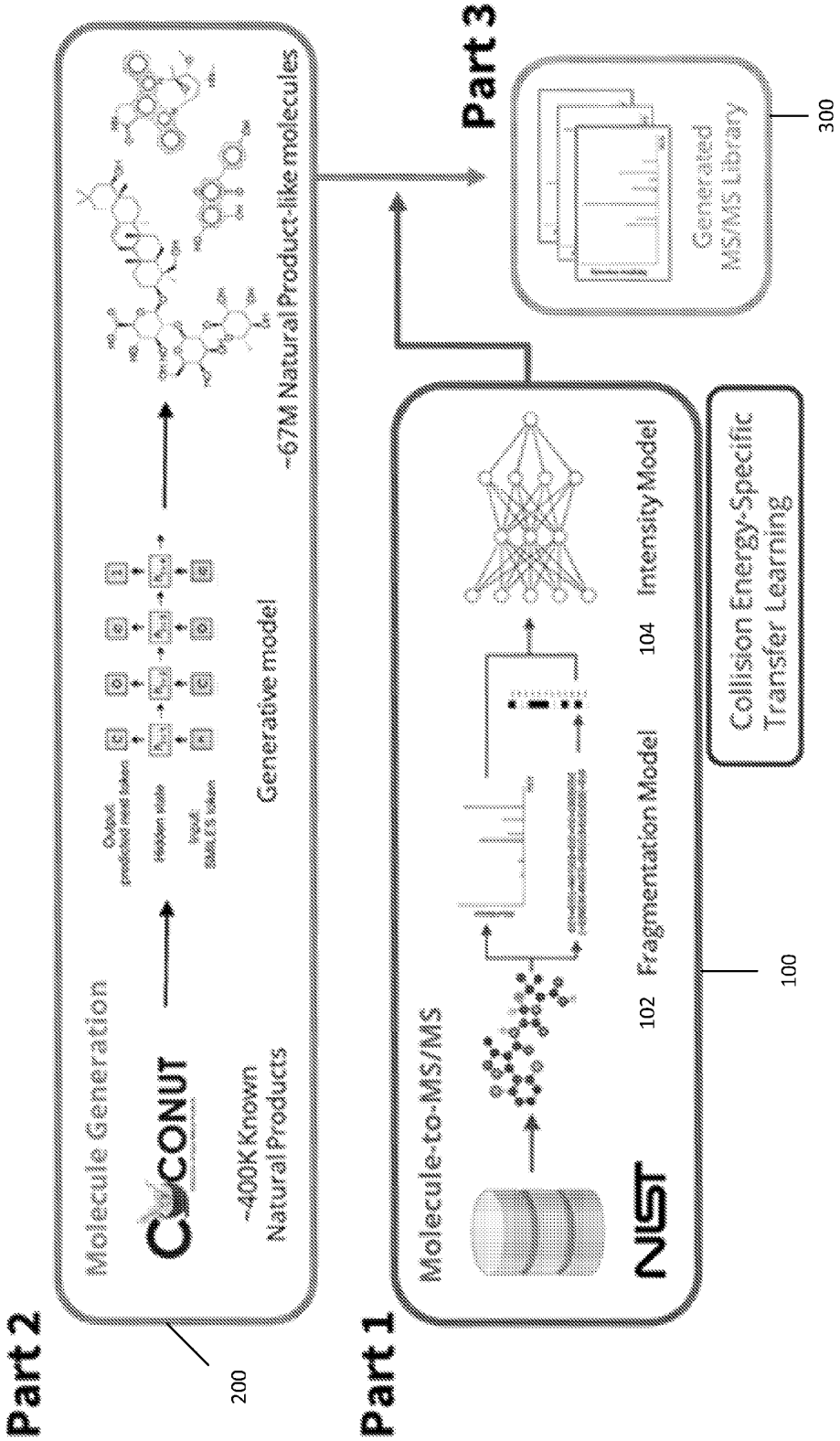
Part 2
Molecule Generation
COCONUT
~400K Known Natural Products
Generative model
~67M Natural Product-like molecules
200
Part 1
Molecule-to-MS/MS
NIST
102 Fragmentation Model
104 Intensity Model
100
Collision Energy-Specific Transfer Learning
Part 3
Generated MS/MS Library
300

Figure 1A

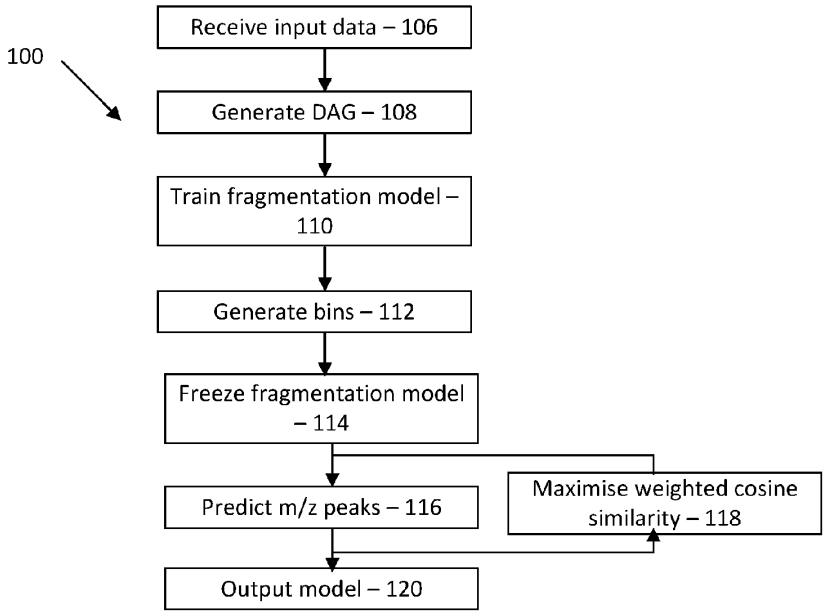
100
Receive input data – 106
Generate DAG – 108
Train fragmentation model – 110
Generate bins – 112
Freeze fragmentation model – 114
Predict m/z peaks – 116
Maximise weighted cosine similarity – 118
Output model – 120

Figure 1B

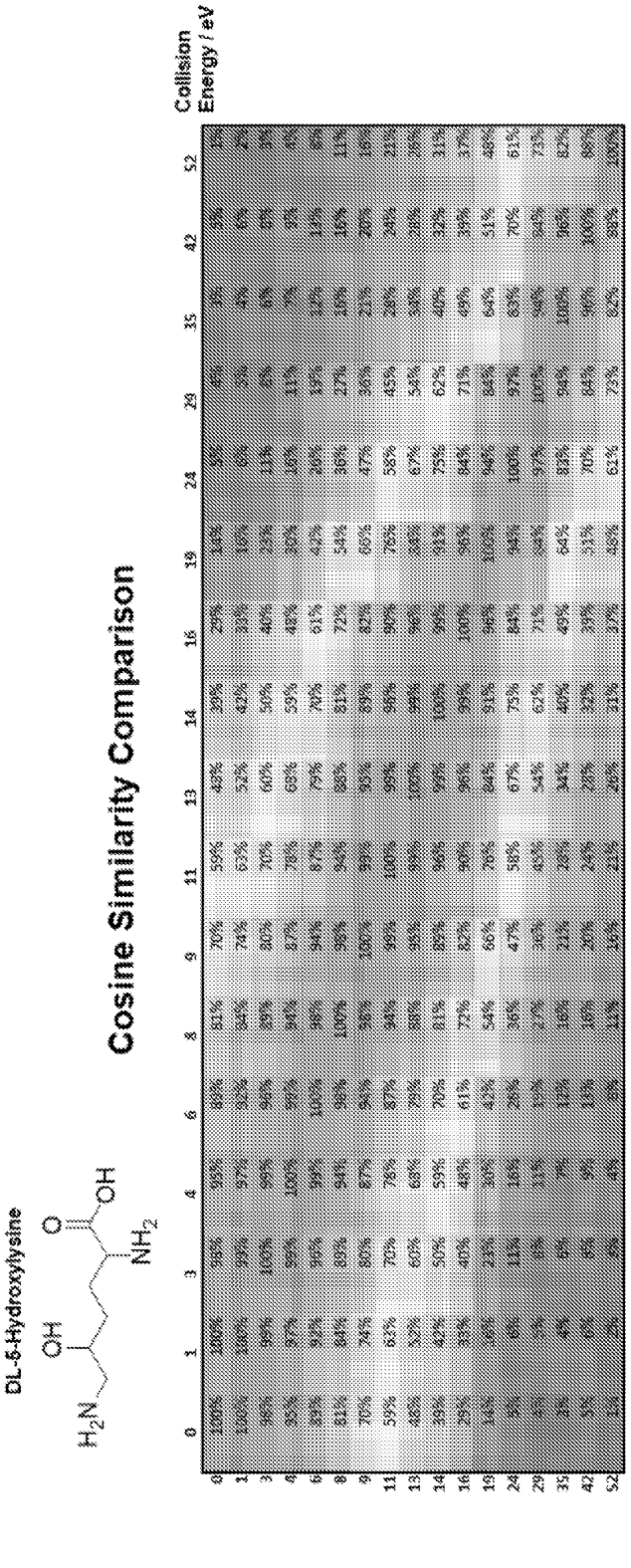
DL-5-Hydroxylysine
Cosine Similarity Comparison
Collision Energy / eV
Collision Energy / eV

Figure 2

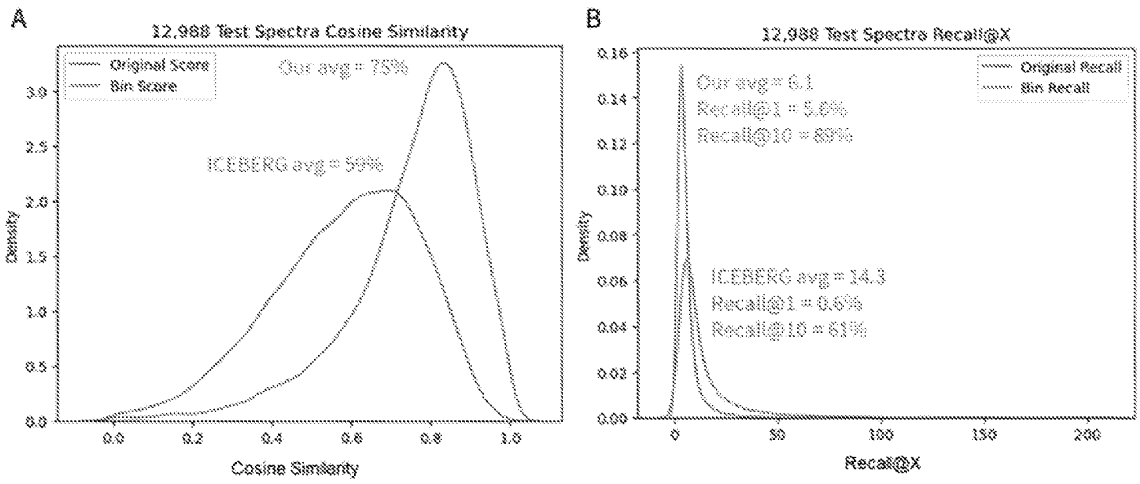

Figure 3

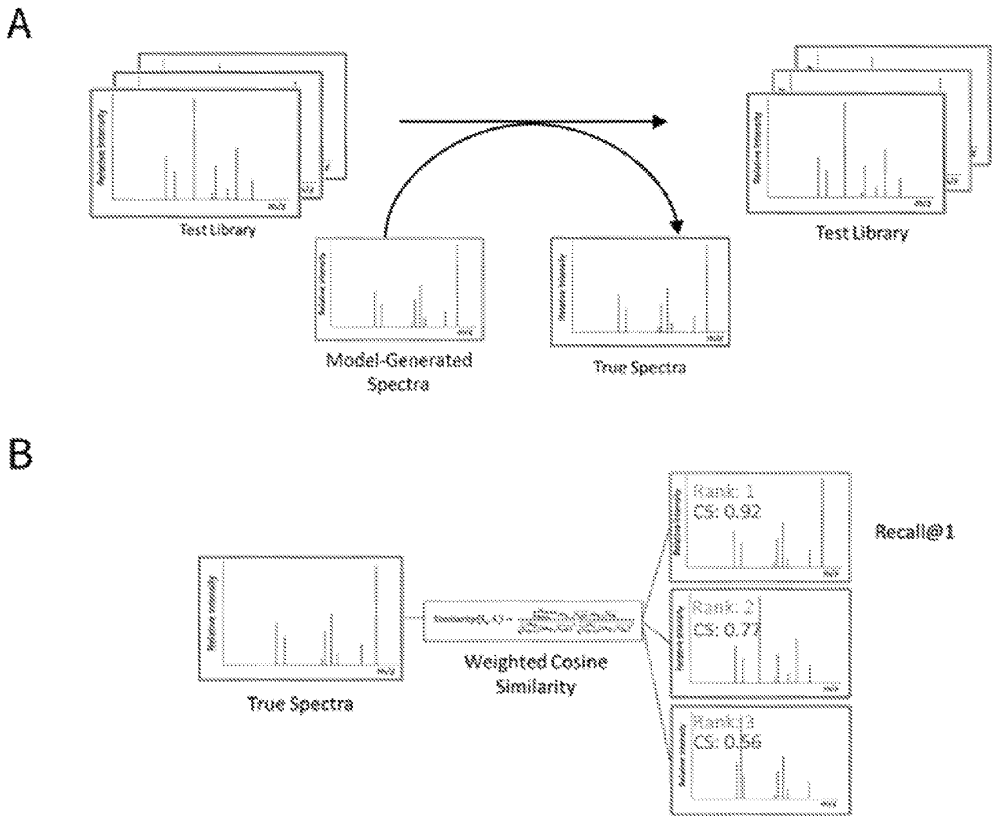

Figure 4

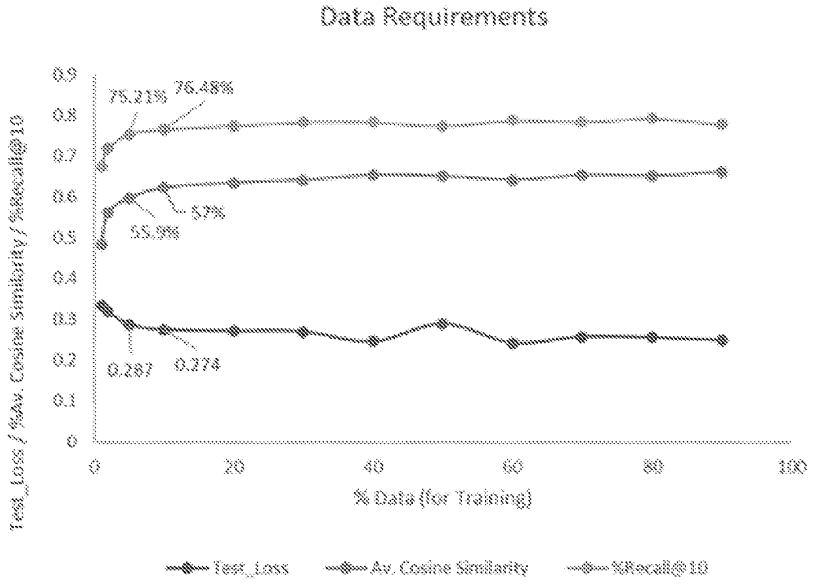
Data Requirements
Test_Loss / %Av. Cosine Similarity / %Recall@10
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
75.21%
76.48%
57%
55.9%
0.287
0.274
0
20
40
60
80
100
% Data (for Training)
Test_Loss
Av. Cosine Similarity
%Recall@10

Figure 5

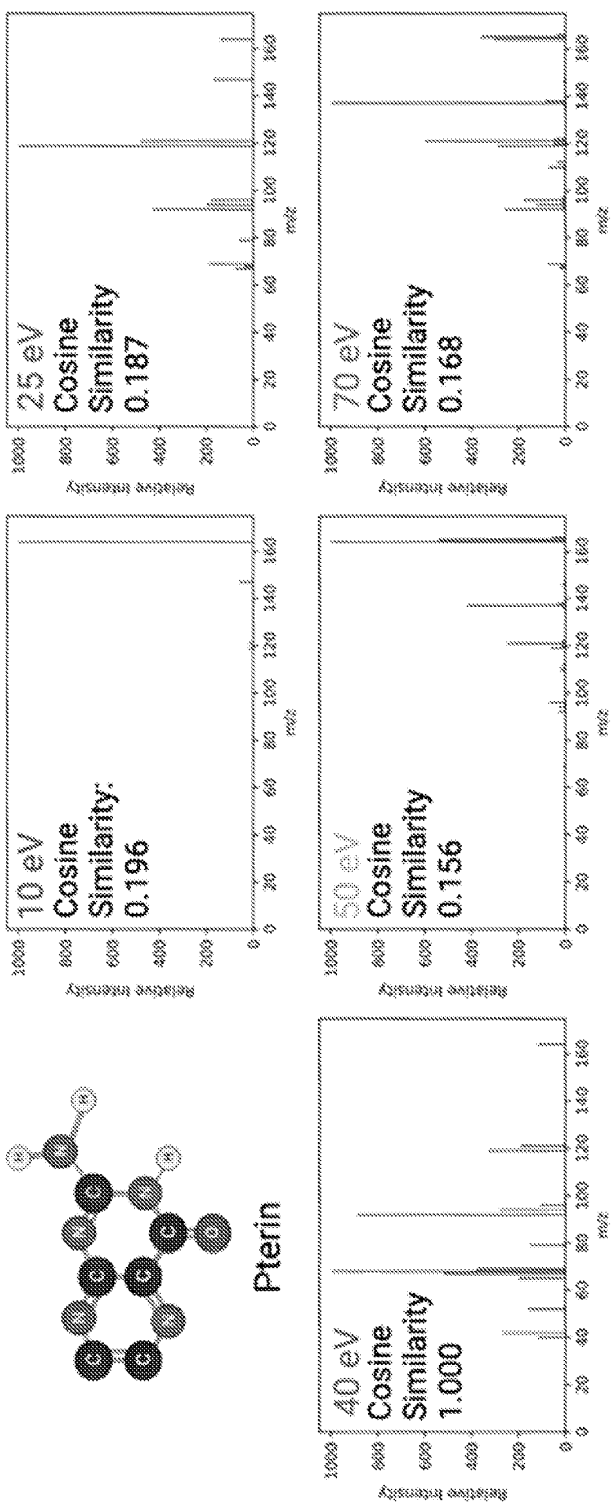
Pterin
10 eV
Cosine
Similarity:
0.196
25 eV
Cosine
Similarity
0.187
40 eV
Cosine
Similarity
1.000
50 eV
Cosine
Similarity
0.156
70 eV
Cosine
Similarity
0.168
Relative Intensity
m/z

Figure 6