

Protein Structure-Based Pathogenicity Determination 100

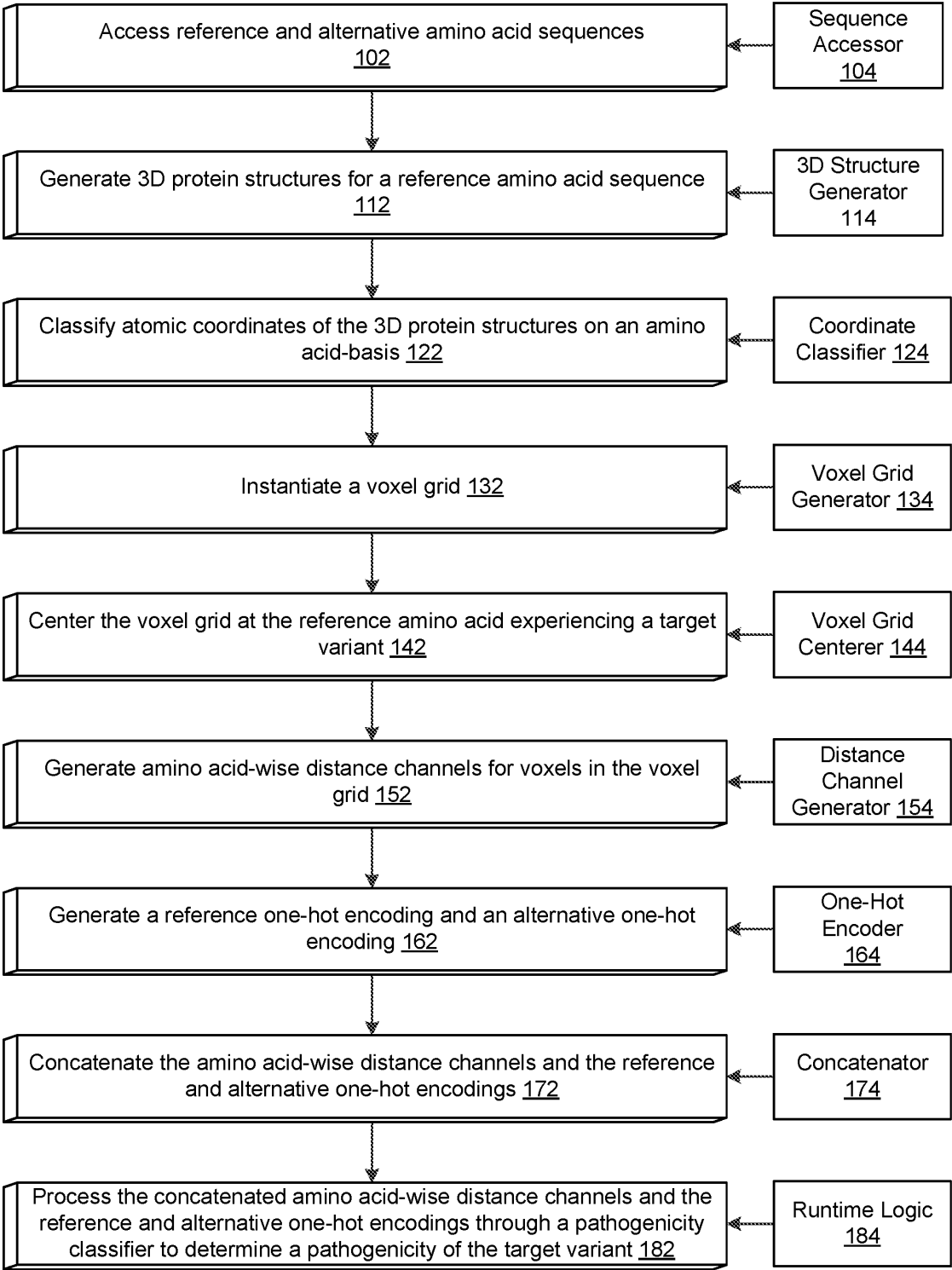


Figure 1

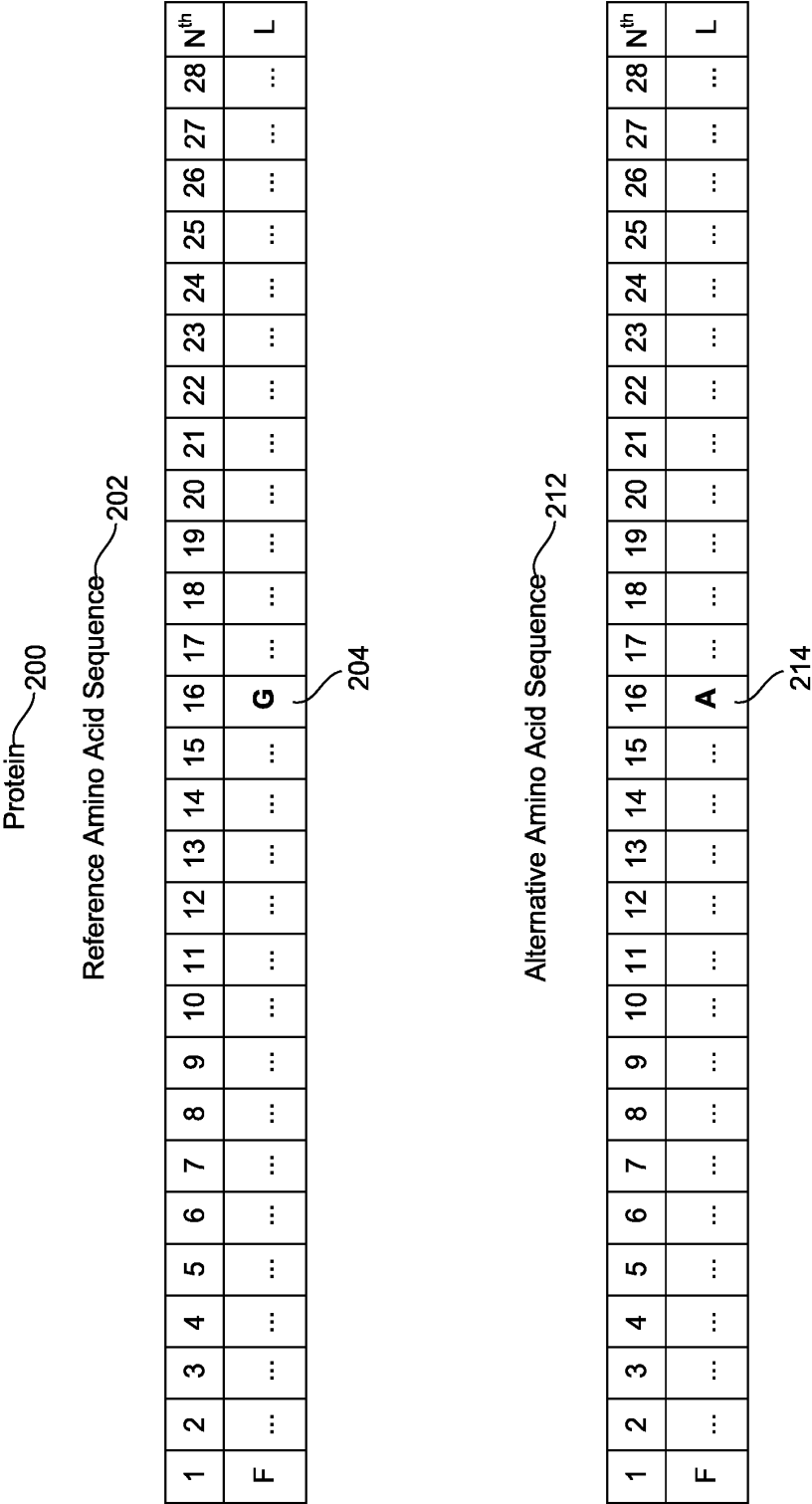


Figure 2

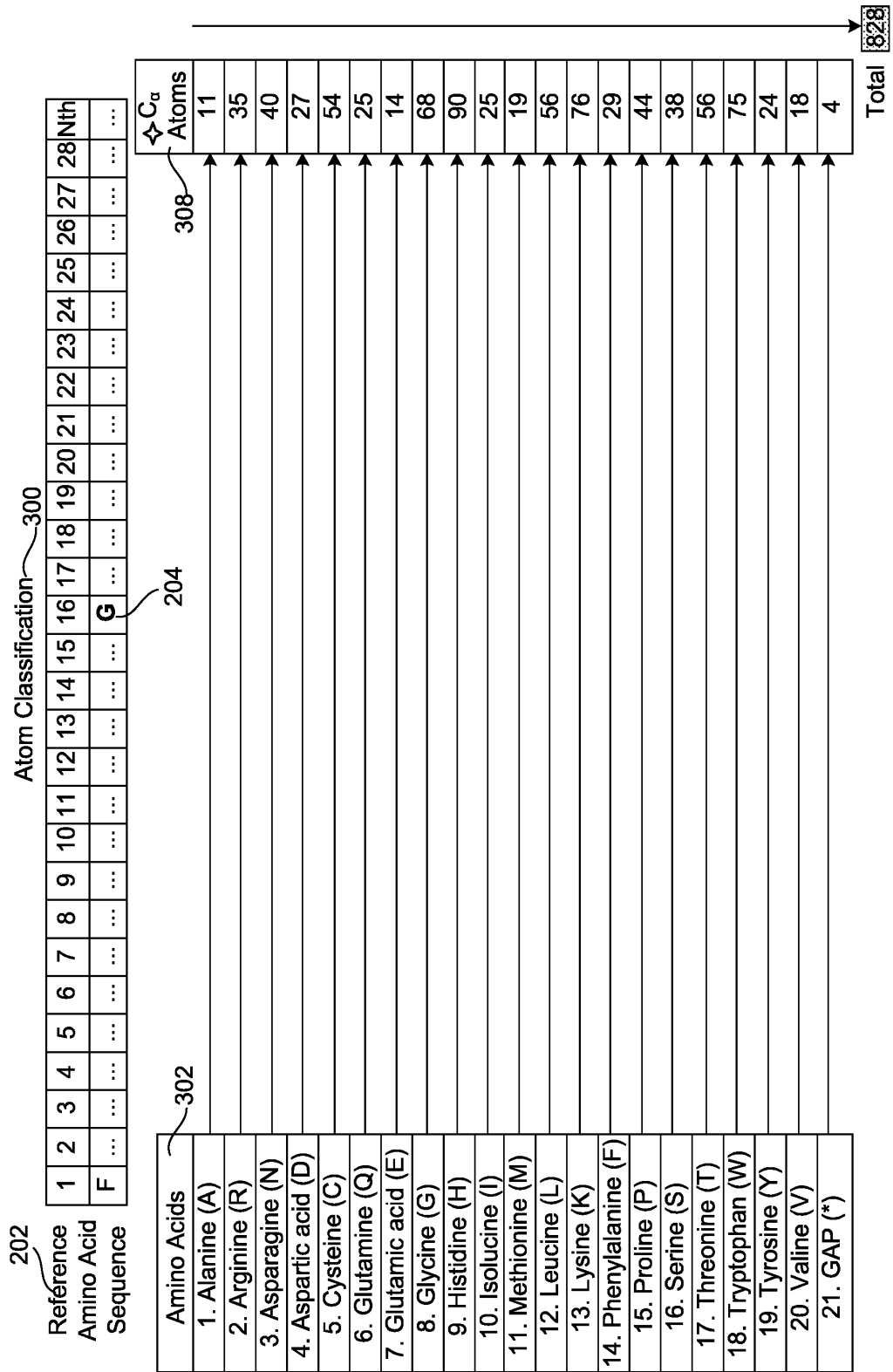


Figure 3

Atomic Coordinates Bucketing 400

1. Alanine (A) C_{α} Atom Coordinates <u>402</u> 1. x, y, z 2. x, y, z ... 11. x, y, z	2. Arginine (R) C_{α} Atom Coordinates <u>404</u> 1. x, y, z 2. x, y, z ... 35. x, y, z	3. Asparagine (N) C_{α} Atom Coordinates <u>406</u> 1. x, y, z 2. x, y, z ... 40. x, y, z	4. Aspartic acid (D) C_{α} Atom Coordinates <u>408</u> 1. x, y, z 2. x, y, z ... 27. x, y, z	5. Cysteine (C) C_{α} Atom Coordinates <u>410</u> 1. x, y, z 2. x, y, z ... 54. x, y, z
6. Glutamine (Q) C_{α} Atom Coordinates <u>412</u> 1. x, y, z 2. x, y, z ... 25. x, y, z	7. Glutamic acid (E) C_{α} Atom Coordinates <u>414</u> 1. x, y, z 2. x, y, z ... 14. x, y, z	8. Glycine (G) C_{α} Atom Coordinates <u>416</u> 1. x, y, z 2. x, y, z ... 68. x, y, z	9. Histidine (H) C_{α} Atom Coordinates <u>418</u> 1. x, y, z 2. x, y, z ... 90. x, y, z	10. Isoleucine (I) C_{α} Atom Coordinates <u>420</u> 1. x, y, z 2. x, y, z ... 25. x, y, z
11. Methionine (M) C_{α} Atom Coordinates <u>422</u> 1. x, y, z 2. x, y, z ... 19. x, y, z	12. Leucine (L) C_{α} Atom Coordinates <u>424</u> 1. x, y, z 2. x, y, z ... 56. x, y, z	13. Lysine (K) C_{α} Atom Coordinates <u>426</u> 1. x, y, z 2. x, y, z ... 76. x, y, z	14. Phenylalanine (F) C_{α} Atom Coordinates <u>428</u> 1. x, y, z 2. x, y, z ... 29. x, y, z	15. Proline (P) C_{α} Atom Coordinates <u>430</u> 1. x, y, z 2. x, y, z ... 44. x, y, z
16. Serine (S) C_{α} Atom Coordinates <u>432</u> 1. x, y, z 2. x, y, z ... 38. x, y, z	17. Threonine (T) C_{α} Atom Coordinates <u>434</u> 1. x, y, z 2. x, y, z ... 56. x, y, z	18. Tryptophan (W) C_{α} Atom Coordinates <u>436</u> 1. x, y, z 2. x, y, z ... 75. x, y, z	19. Tyrosine (Y) C_{α} Atom Coordinates <u>438</u> 1. x, y, z 2. x, y, z ... 24. x, y, z	20. Valine (V) C_{α} Atom Coordinates <u>440</u> 1. x, y, z 2. x, y, z ... 18. x, y, z
21. GAP (*) 1. x, y, z 2. x, y, z ... 4. x, y, z				

Figure 4

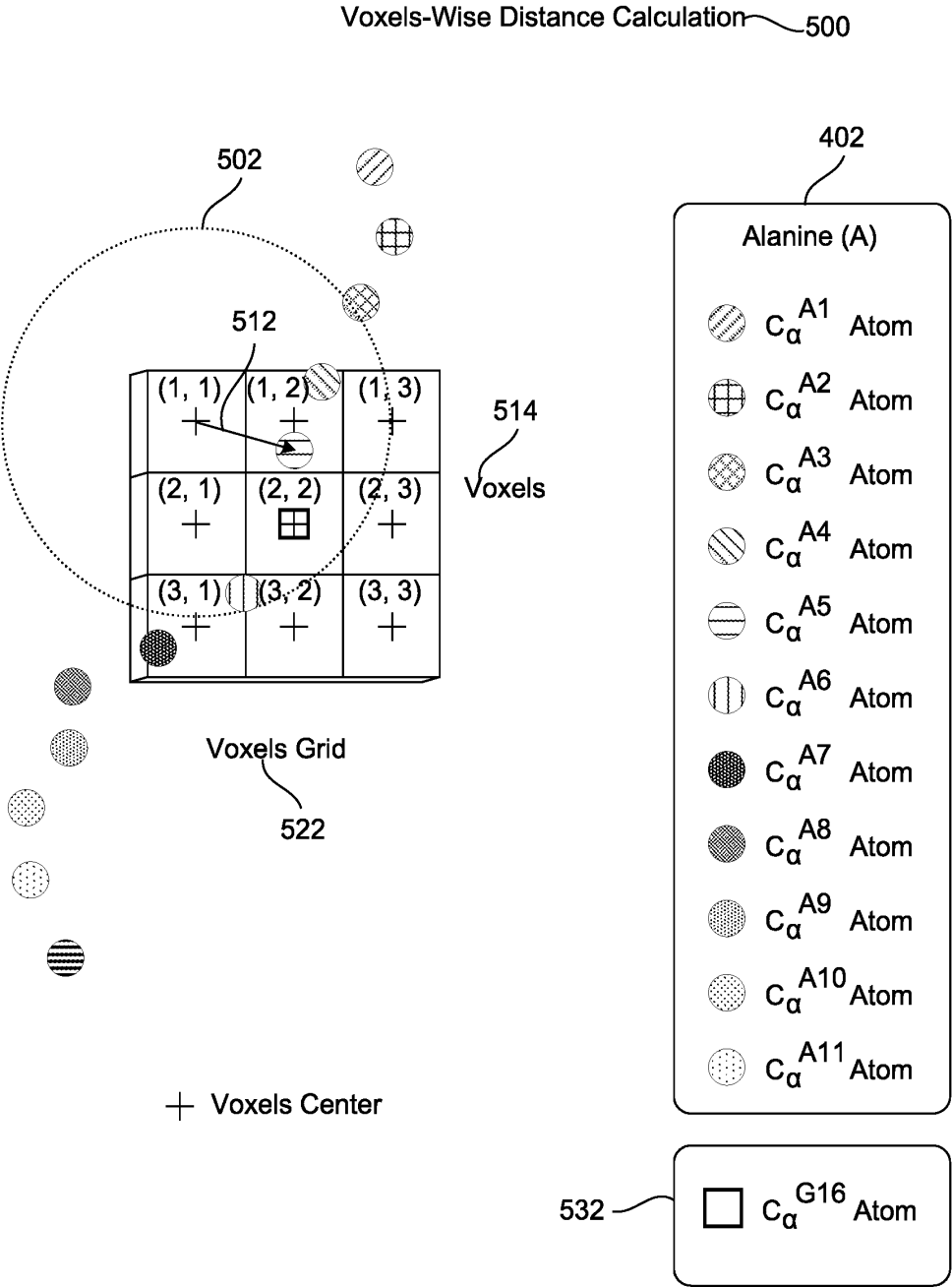


Figure 5

Amino Acid-Wise Distance Channels 600

514	602	604	606	608	610	612	614	616	618	620	622	624	626	628	630	632	634	636	638	640	642
Voxels	1. A	2. R	3. N	4. D	5. C	6. Q	7. E	8. G	9. H	10. I	11. M	12. L	13. K	14. F	15. P	16. S	17. T	18. W	19. Y	20. V	21. *
(1, 1)	1.55	0.34	3.49	0.71	2.8	4.9	0.25	0.15	0.59	1.8	2.5	3.2	4.7	0.61	2.72	5.4	1.5	2.4	0.36	2.9	4.2
(1, 2)	2.38	4.57	0.22	1.4	3.6	2.7	0.45	0.31	0.68	4.5	1.7	5.3	0.78	0.46	4.1	7.3	5.2	4.9	1.6	2.9	4.1
(1, 3)	5.16	1.77	0.38	0.84	1.22	2.3	0.56	0.28	0.41	3.7	1.85	4.52	0.92	3.12	1.98	5.21	3.53	1.34	0.47	8.9	3.3
(2, 1)	1.42	2.21	3.5	0.94	3.67	5.8	0.37	0.24	3.55	0.59	3.16	0.27	2.41	3.30	5.58	4.43	1.17	2.41	2.15	1.3	1.4
(2, 2)	4.27	0.36	1.38	4.54	3.15	4.92	1.25	2.33	0.60	1.82	2.66	3.25	2.72	5.99	4.21	1.26	1.35	3.15	3.36	2.7	1.1
(2, 3)	5.82	3.22	2.35	5.12	2.10	3.17	0.16	5.71	5.19	0.19	5.28	4.13	1.27	0.65	0.22	2.18	2.89	1.44	1.15	1.4	0.28
(3, 1)	2.19	4.34	1.19	3.15	1.56	2.16	1.25	3.25	4.22	1.86	5.52	3.14	0.42	4.29	1.34	3.36	3.81	0.44	0.28	4.9	1.22
(3, 2)	3.45	5.35	2.15	4.70	5.18	4.12	2.34	2.11	3.79	4.19	3.43	5.51	0.44	1.65	0.15	4.42	2.76	5.1	0.64	1.1	2.45
(3, 3)	1.35	2.98	1.17	0.70	3.2	3.16	3.27	6.56	1.16	5.25	2.75	2.12	1.24	0.27	1.72	5.41	1.53	1.15	0.71	0.77	3.90

Figure 6

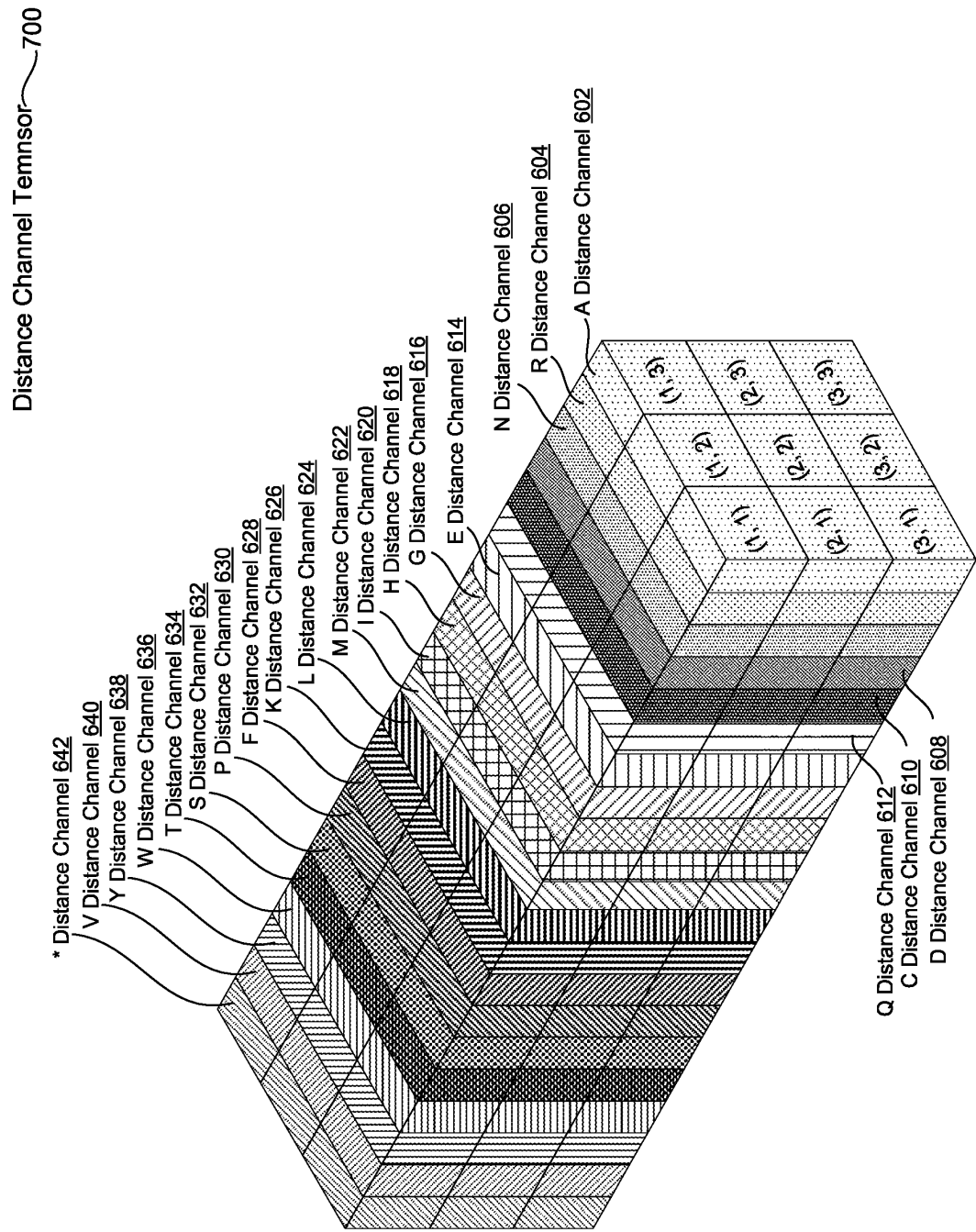
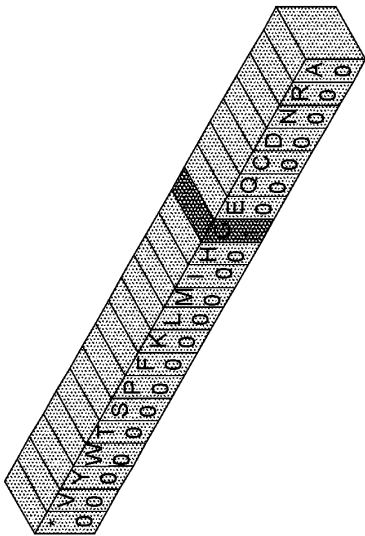


Figure 7

302	802	One-Hot Encodings		800
Amino Acids	One-Hot Encoded Variant- Experiencing Reference Amino Acid	G	One-Hot Encoded Variant/ Alternative Amino Acid	A
1. Alanine (A)		0		1
2. Arginine (R)		0		0
3. Asparagine (N)		0		0
4. Aspartic acid (D)		0		0
5. Cysteine (C)		0		0
6. Glutamine (Q)		0		0
7. Glutamic acid (E)		0		0
8. Glycine (G)		1		0
9. Histidine (H)		0		0
10. Isolucine (I)		0		0
11. Methionine (M)		0		0
12. Leucine (L)		0		0
13. Lysine (K)		0		0
14. Phenylalanine (F)		0		0
15. Proline (P)		0		0
16. Serine (S)		0		0
17. Threonine (T)		0		0
18. Tryptophan (W)		0		0
19. Tyrosine (Y)		0		0
20. Valine (V)		0		0
21. GAP (*)		0		0

Figure 8

Voxelized One-Hot Encoded Variant-Experiencing Reference Amino Acid 902



Voxelized One-Hot Encoded Variant/Alternative Amino Acid 912

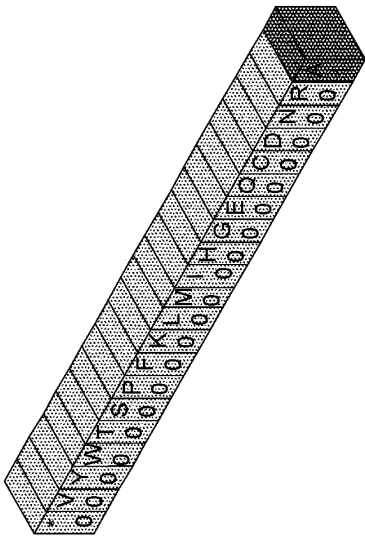
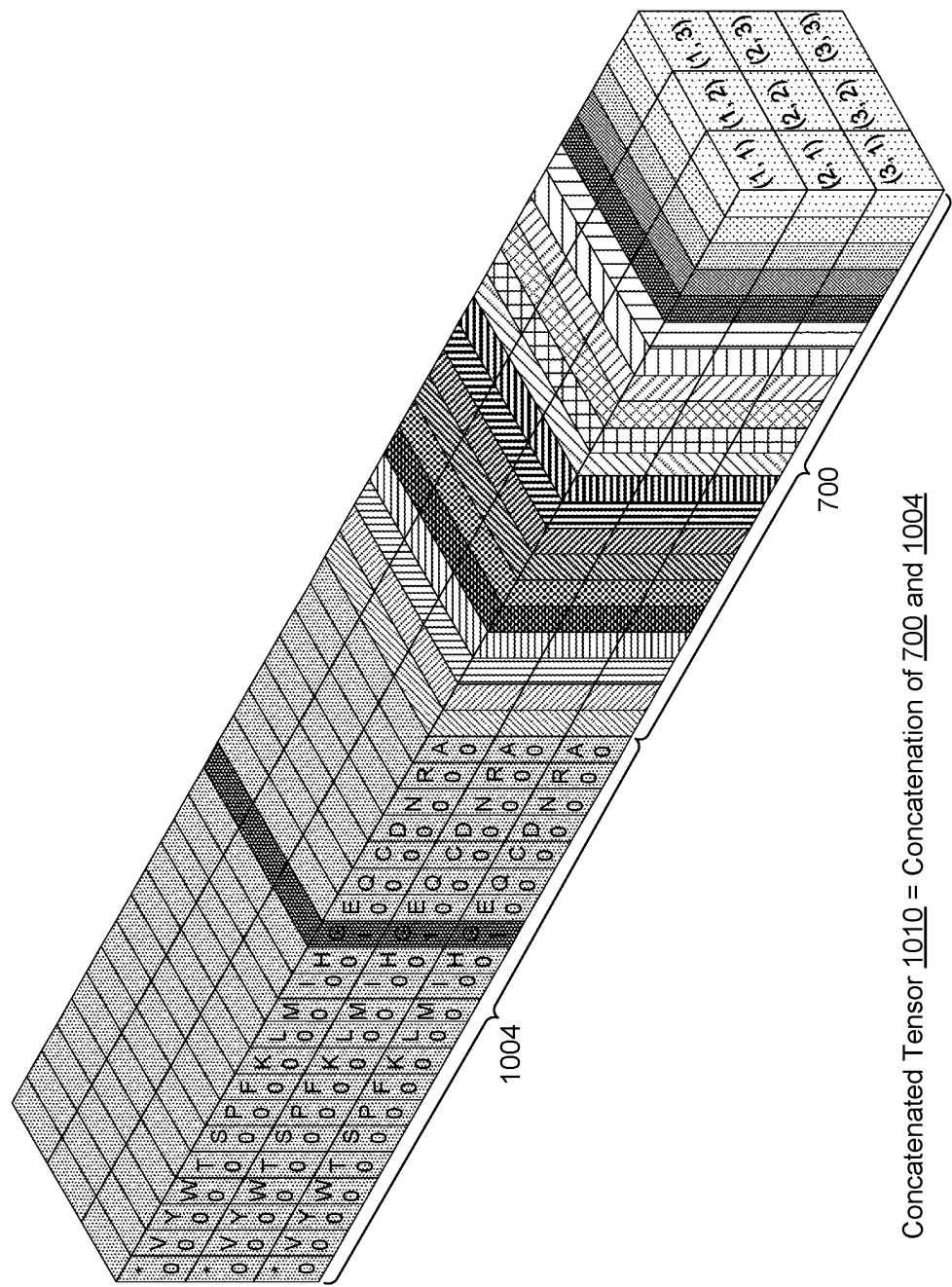


Figure 9

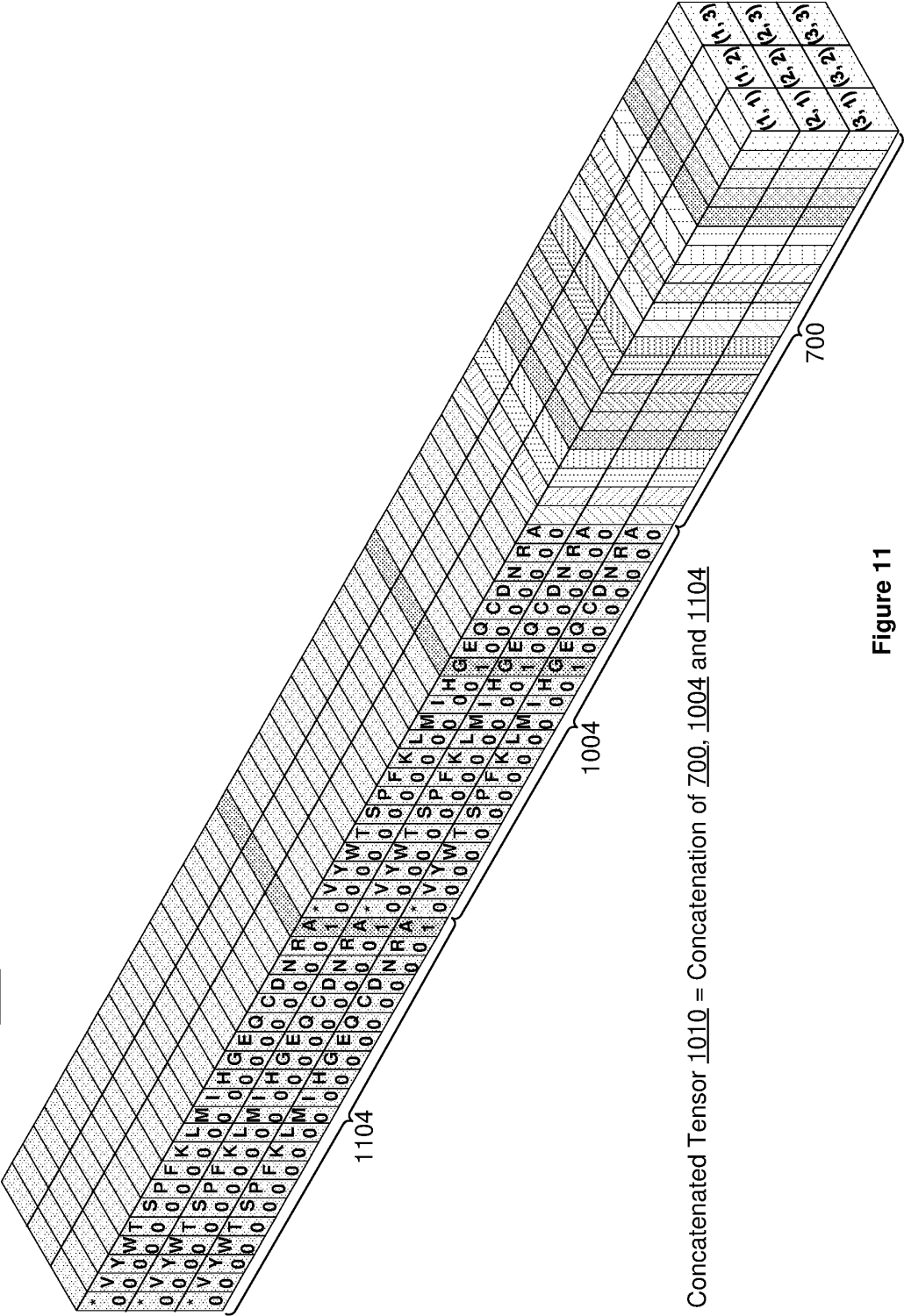
Distance Channel Tensor + Reference Allele Tensor 1000



Concatenated Tensor 1010 = Concatenation of 700 and 1004

Figure 10

Distance Channel Tensor + Reference Allele Tensor + Alternative Allel Tensor
1100



Concatenated Tensor 1010 = Concatenation of 700, 1004 and 1104

Figure 11

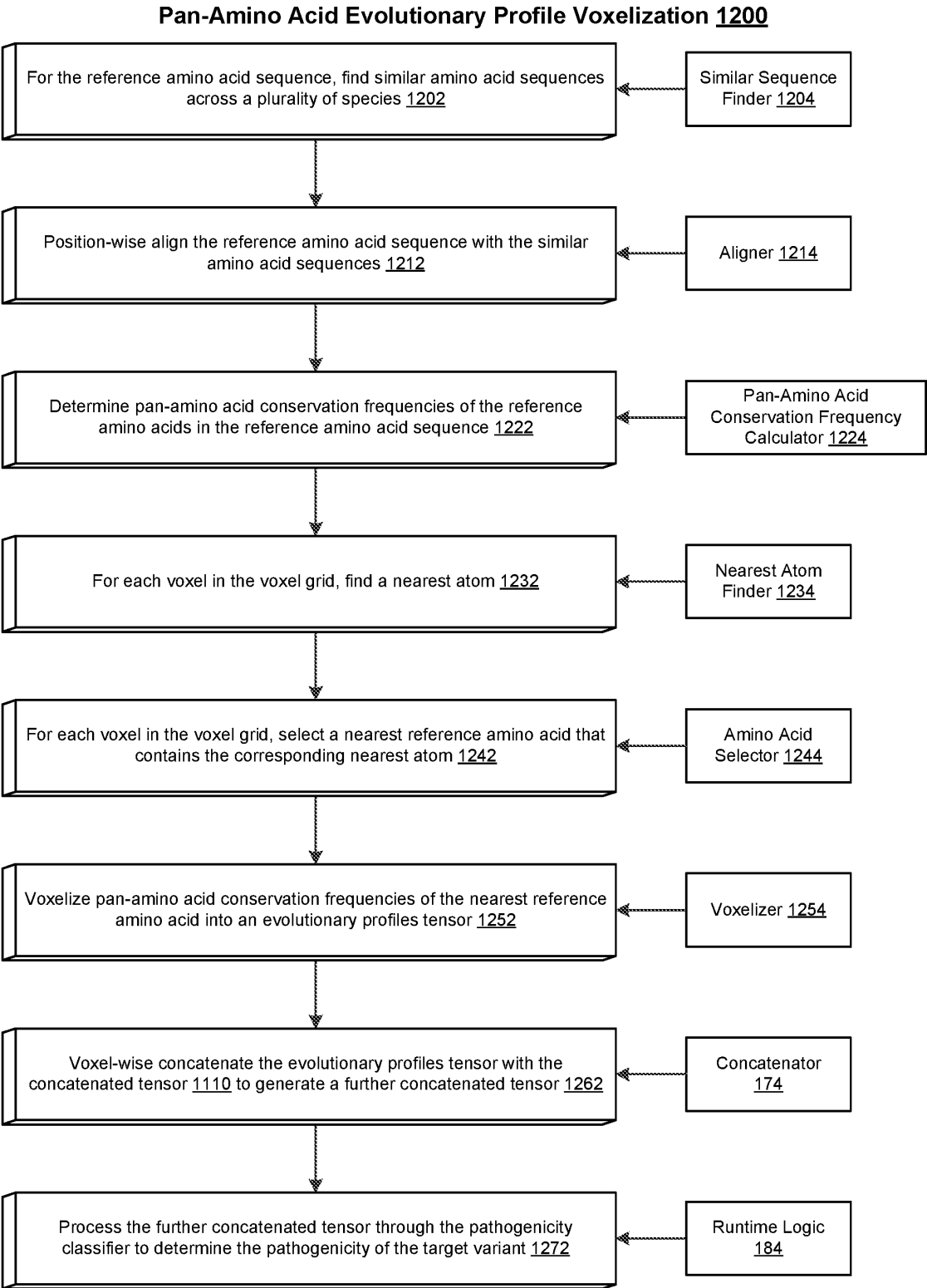


Figure 12

Voxels-to-Nearest Amino Acids Mapping

514	1302	1312	1300
Voxels	Nearest Atoms	Nearest Reference Amino Acids	
(1, 1)	Carbon ¹⁵ _α	Aspartic acid (D) ¹⁵	
(1, 2)	Oxygen ²⁷	Arginine (R) ²⁷	
(1, 3)	Hydrogen ¹⁴	Isolucine (I) ¹⁴	
(2, 1)	Carbon ⁴⁵ _β	Threonine (T) ⁴⁵	
(2, 2)	Oxygen ²²	Tyrosine (Y) ²²	
(2, 3)	Nitrogen ⁶³	Glycine (G) ⁶³	
(3, 1)	Carbon ³⁸ _α	Tryptophan (W) ³⁸	
(3, 2)	Carbon ⁵ _β	Valine (V) ⁵	
(3, 3)	Carbon ²³ _α	Proline (P) ²³	

Figure 13

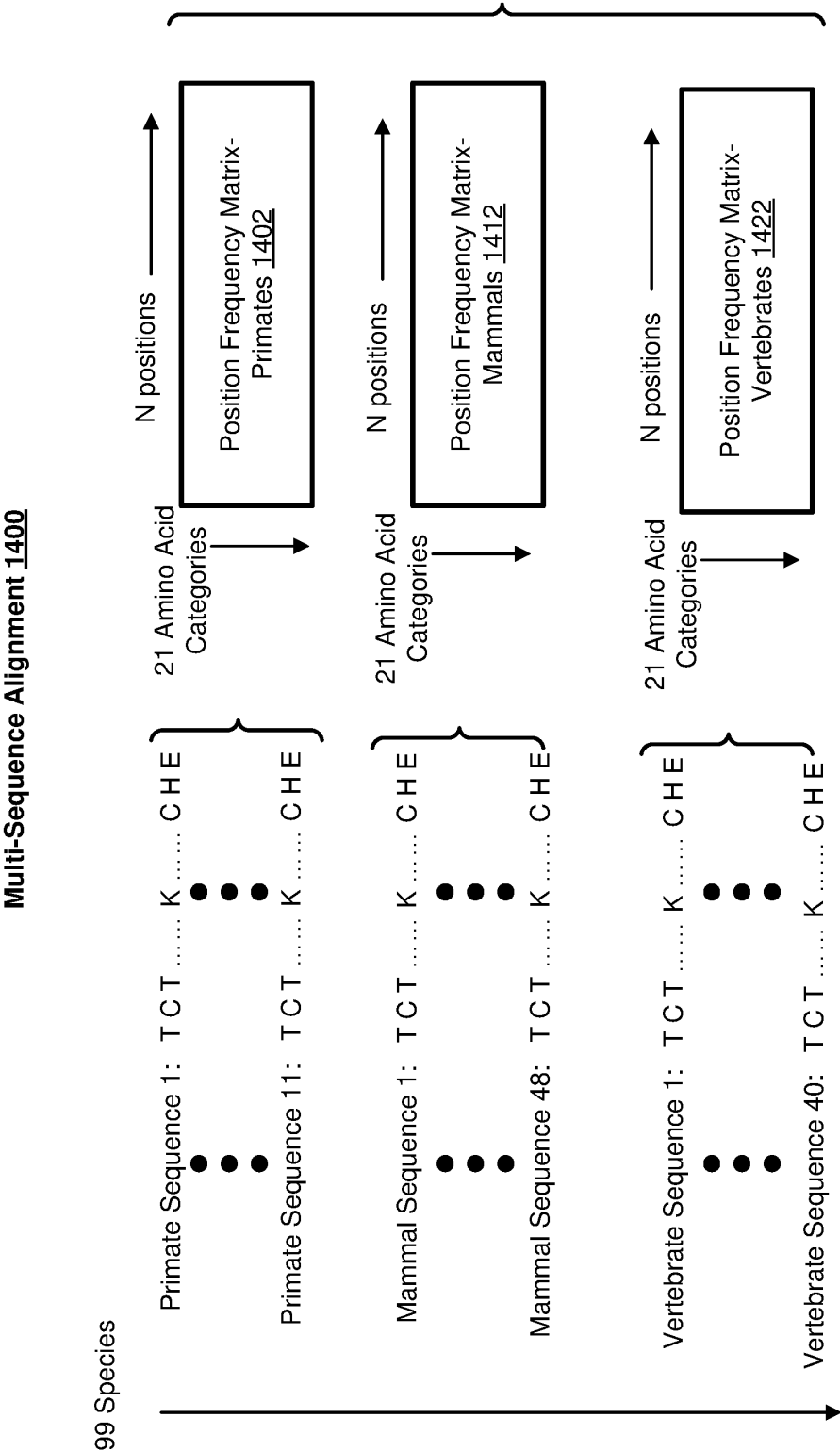


Figure 14

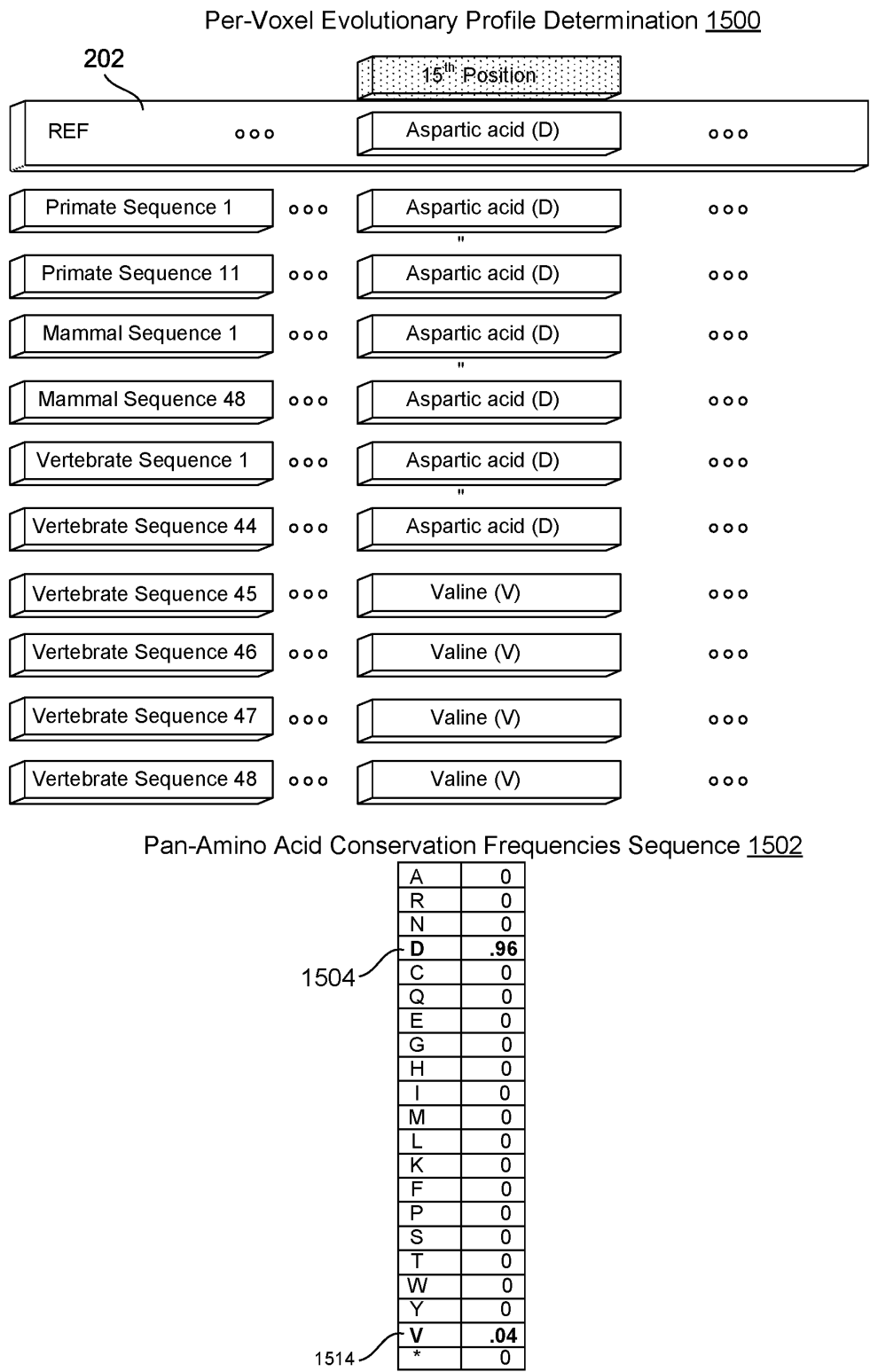


Figure 15

Voxels-to-Evolutionary Profiles Mapping

1600

514

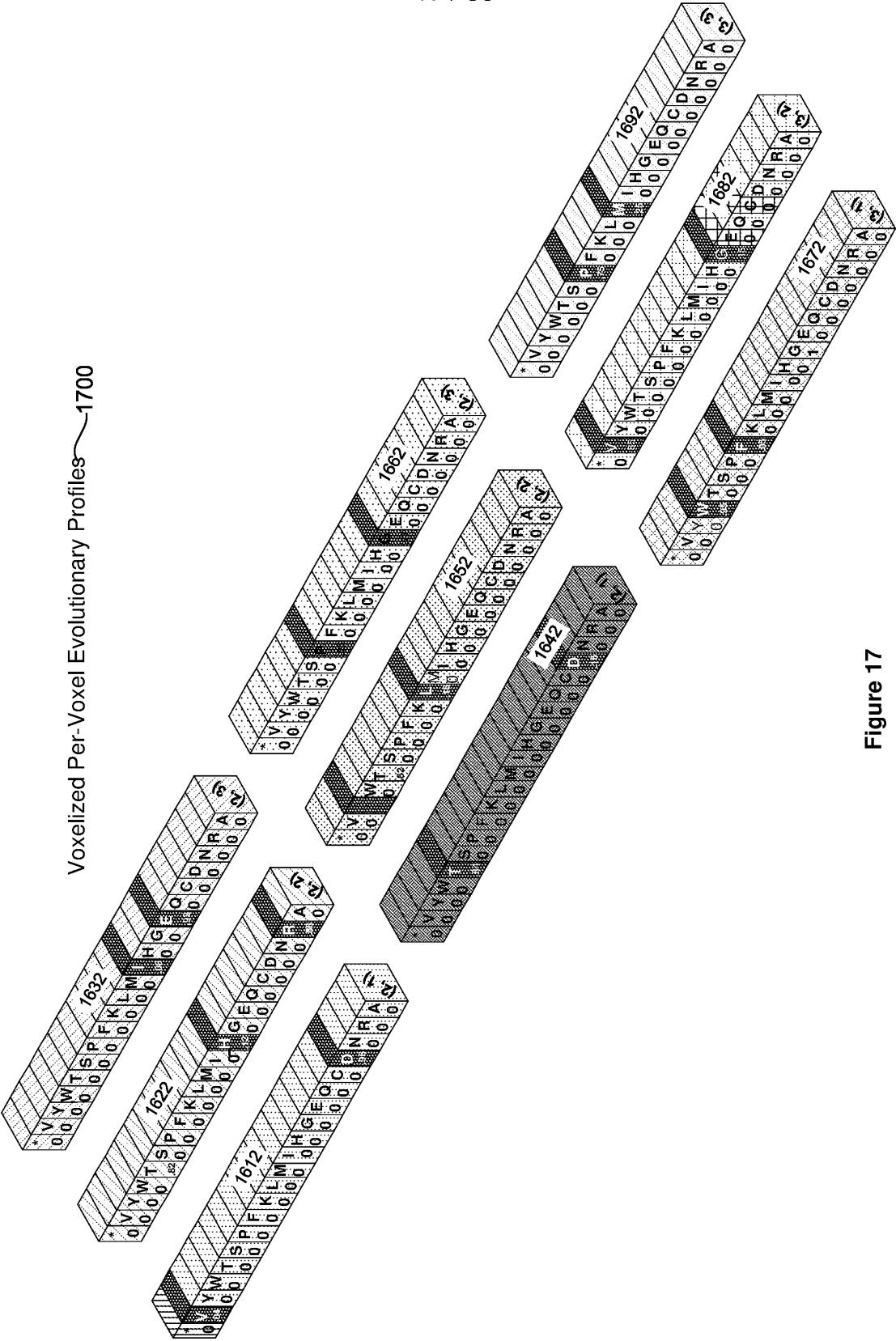
1302

1312

1602

Voxels	Nearest Atoms	Nearest Reference Amino Acids	Per-Voxel Evolutionary Profiles																					
			A	R	N	D	C	Q	E	G	H	I	M	L	K	F	P	S	T	W	Y	V	*	
(1, 1)	Carbon ¹⁵ _α	Aspartic acid (D) ¹⁵	0	0	0	.96	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.04	0	1612
(1, 2)	Oxygen ²⁷	Arginine (R) ²⁷	0	.88	0	0	0	0	0	0	.12	0	0	0	0	0	0	0	0	0	0	0	0	1622
(1, 3)	Hydrogen ¹⁴	Isolucine (I) ¹⁴	0	0	0	0	0	0	.08	0	0	.92	0	0	0	0	0	0	0	0	0	0	0	1632
(2, 1)	Carbon ⁴⁵ _β	Threonine (T) ⁴⁵	0	0	.16	0	0	0	0	0	0	0	0	0	0	0	0	0	.84	0	0	0	0	1642
(2, 2)	Oxygen ²²	Tyrosine (Y) ²²	0	0	0	0	0	0	0	0	0	0	0	.18	0	0	0	0	0	0	0	.82	0	1652
(2, 3)	Nitrogen ⁶³	Glycine (G) ⁶³	0	0	0	0	0	0	0	.77	0	0	0	0	0	0	.23	0	0	0	0	0	0	1662
(3, 1)	Carbon ³⁸ _α	Tryptophan (W) ³⁸	0	0	0	0	0	0	0	0	0	0	0	0	0	.05	0	0	0	.95	0	0	0	1672
(3, 2)	Carbon ⁵ _β	Valine (V) ⁵	0	0	0	0	0	0	0	.15	0	0	0	0	0	0	0	0	0	0	0	.85	0	1682
(3, 3)	Carbon ²³ _α	Proline (P) ²³	0	0	0	0	0	0	0	0	0	0	.25	0	0	0	.75	0	0	0	0	0	0	1692

Figure 16



Evolutionary Profiles Tensor—1800

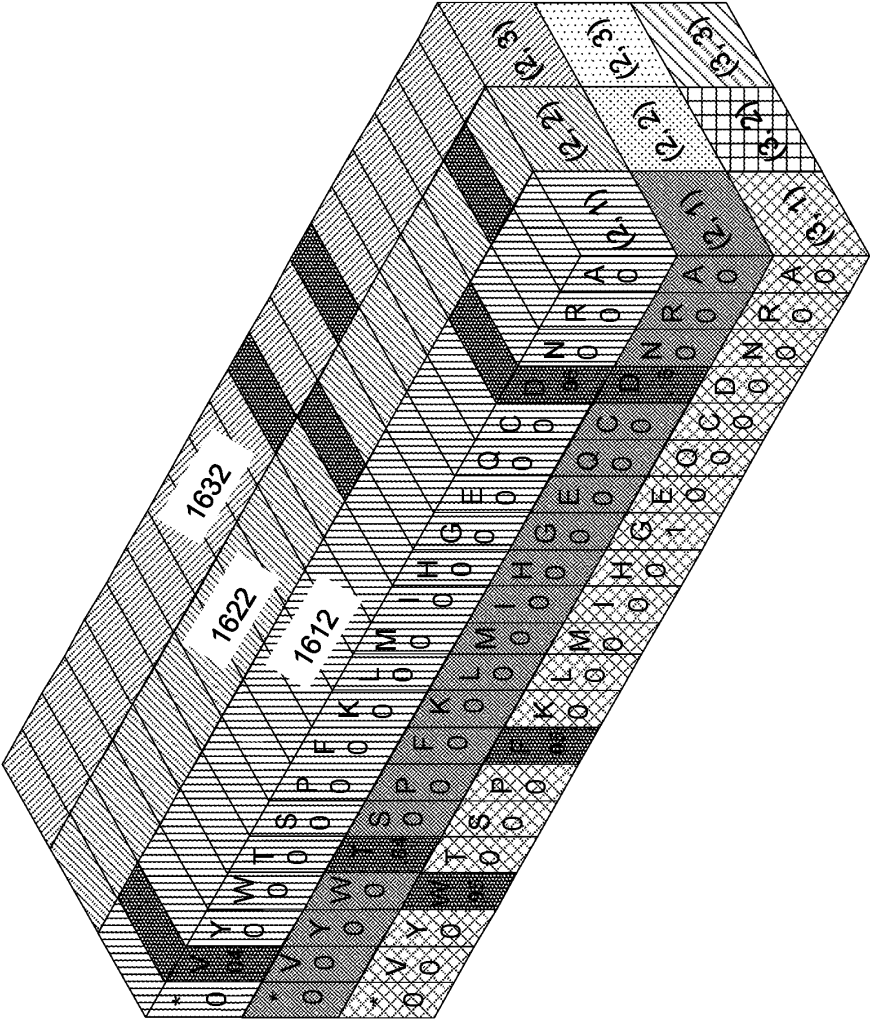


Figure 18

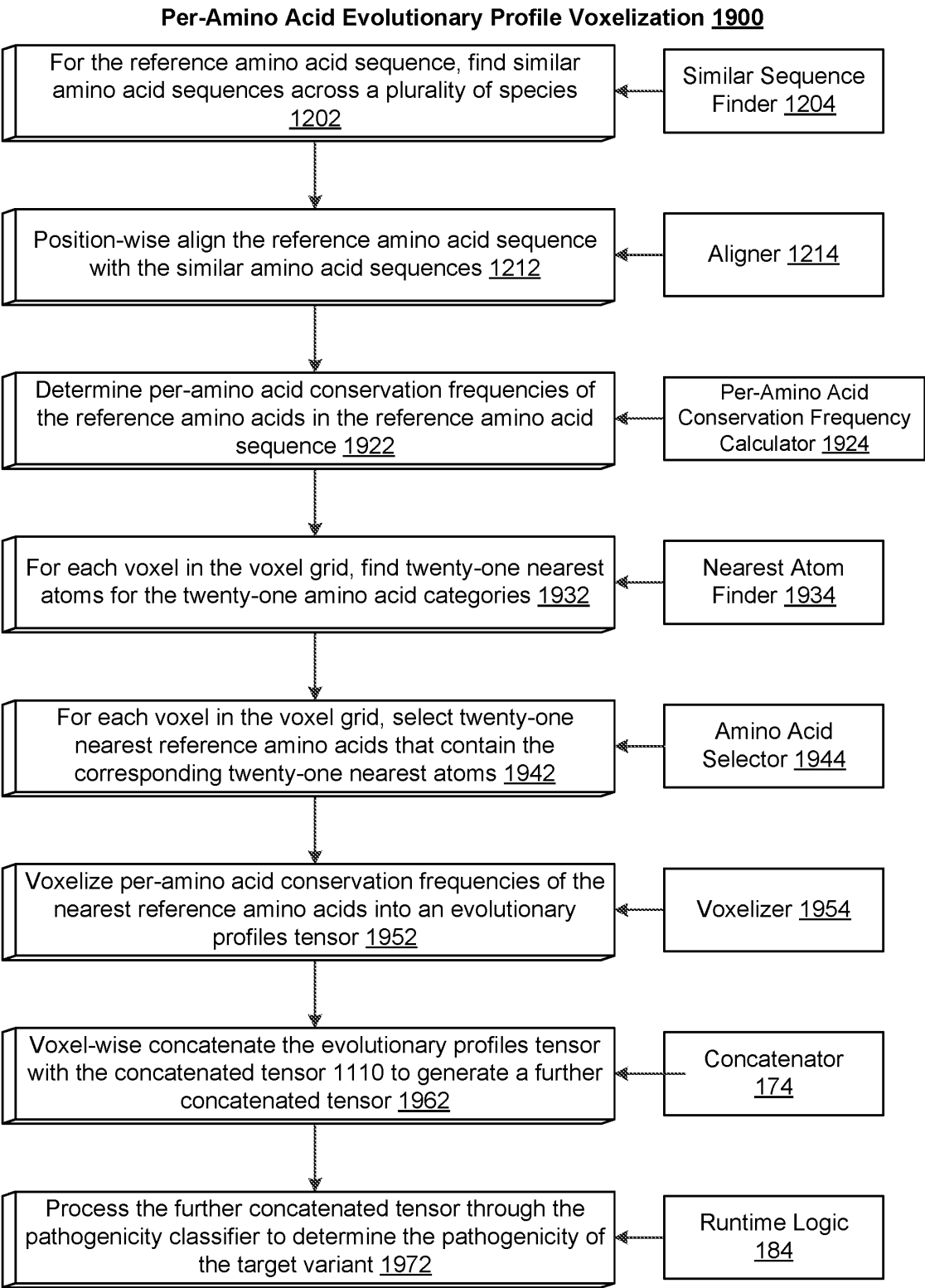


Figure 19

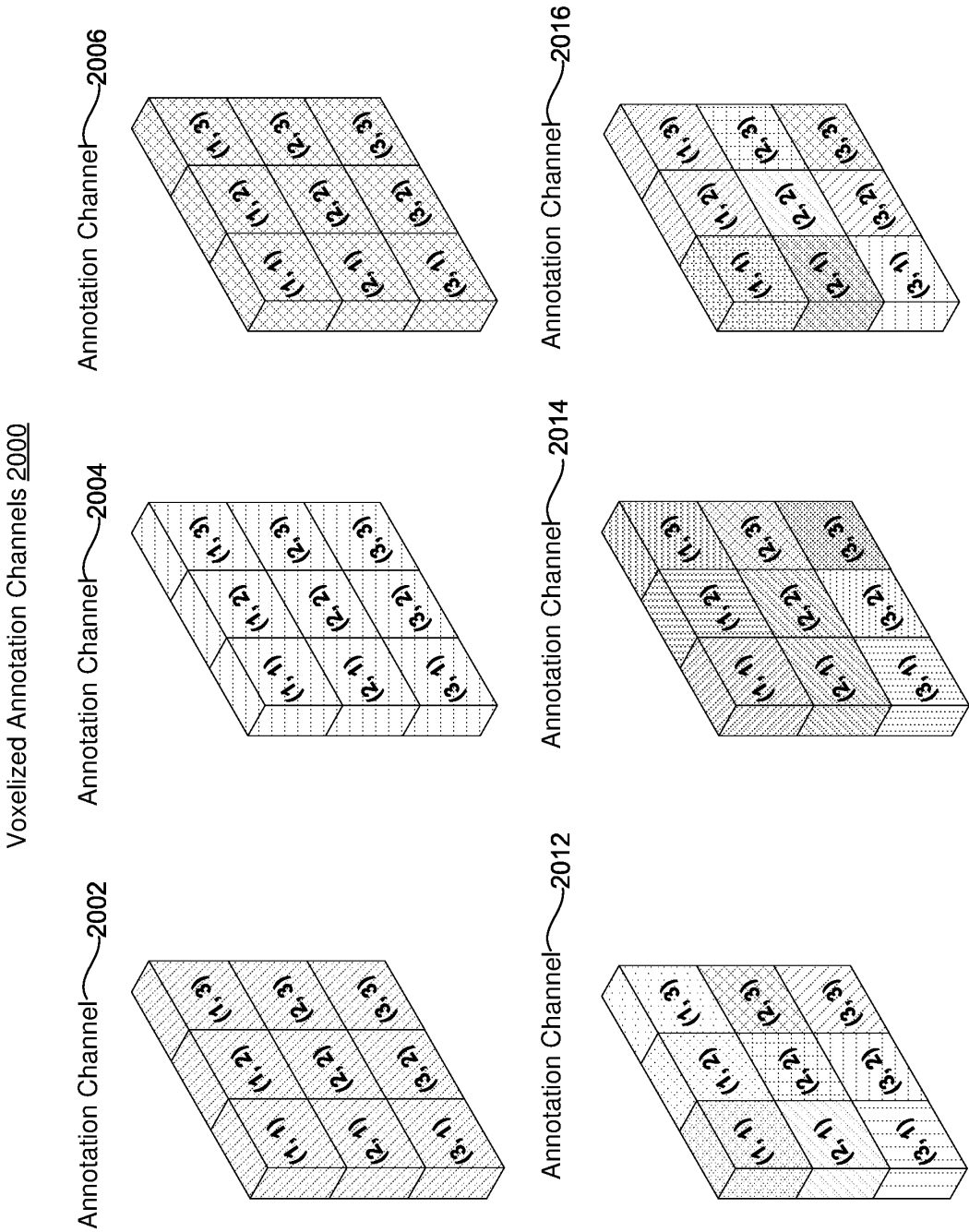


Figure 20

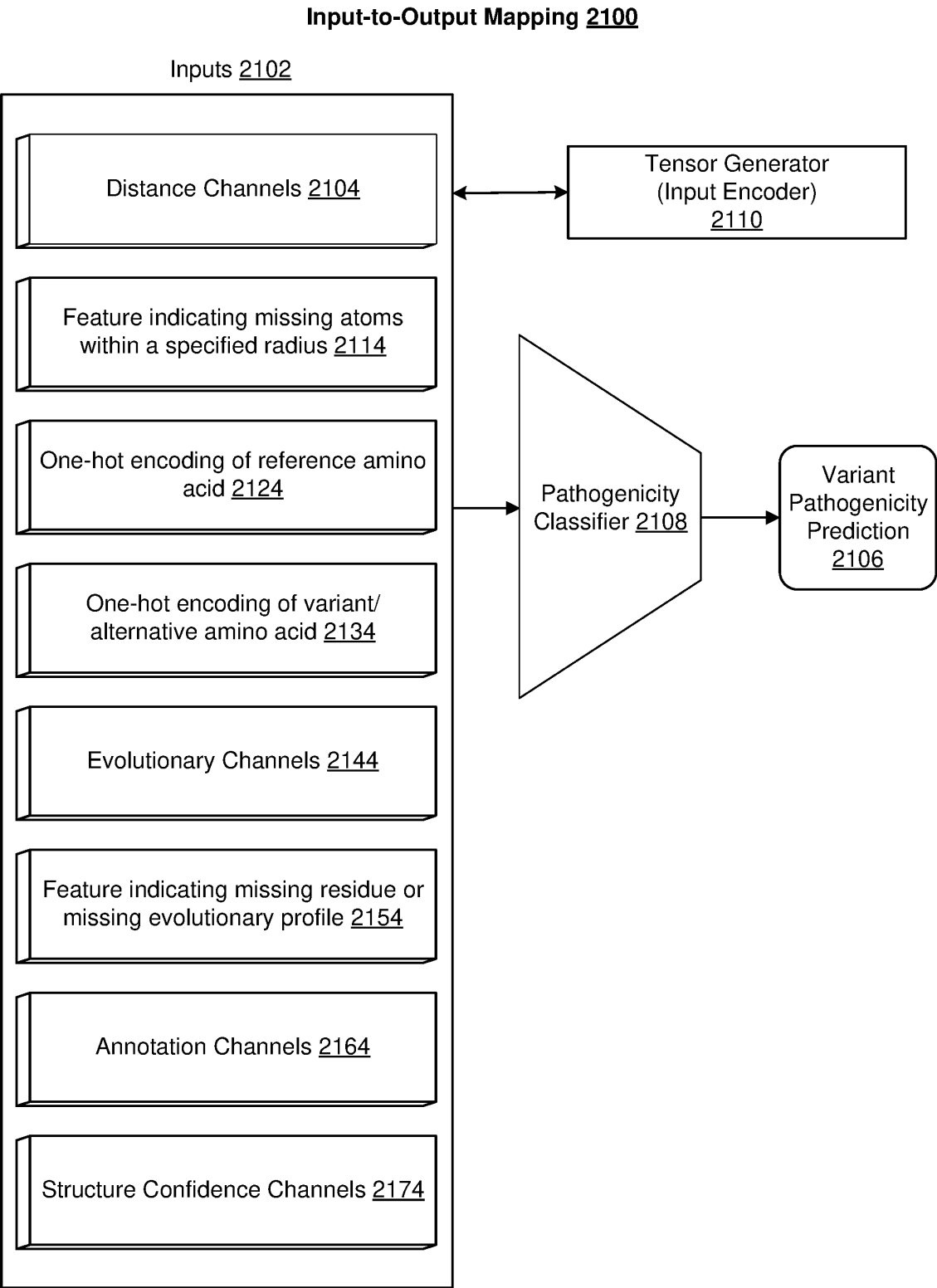


Figure 21

22 / 36

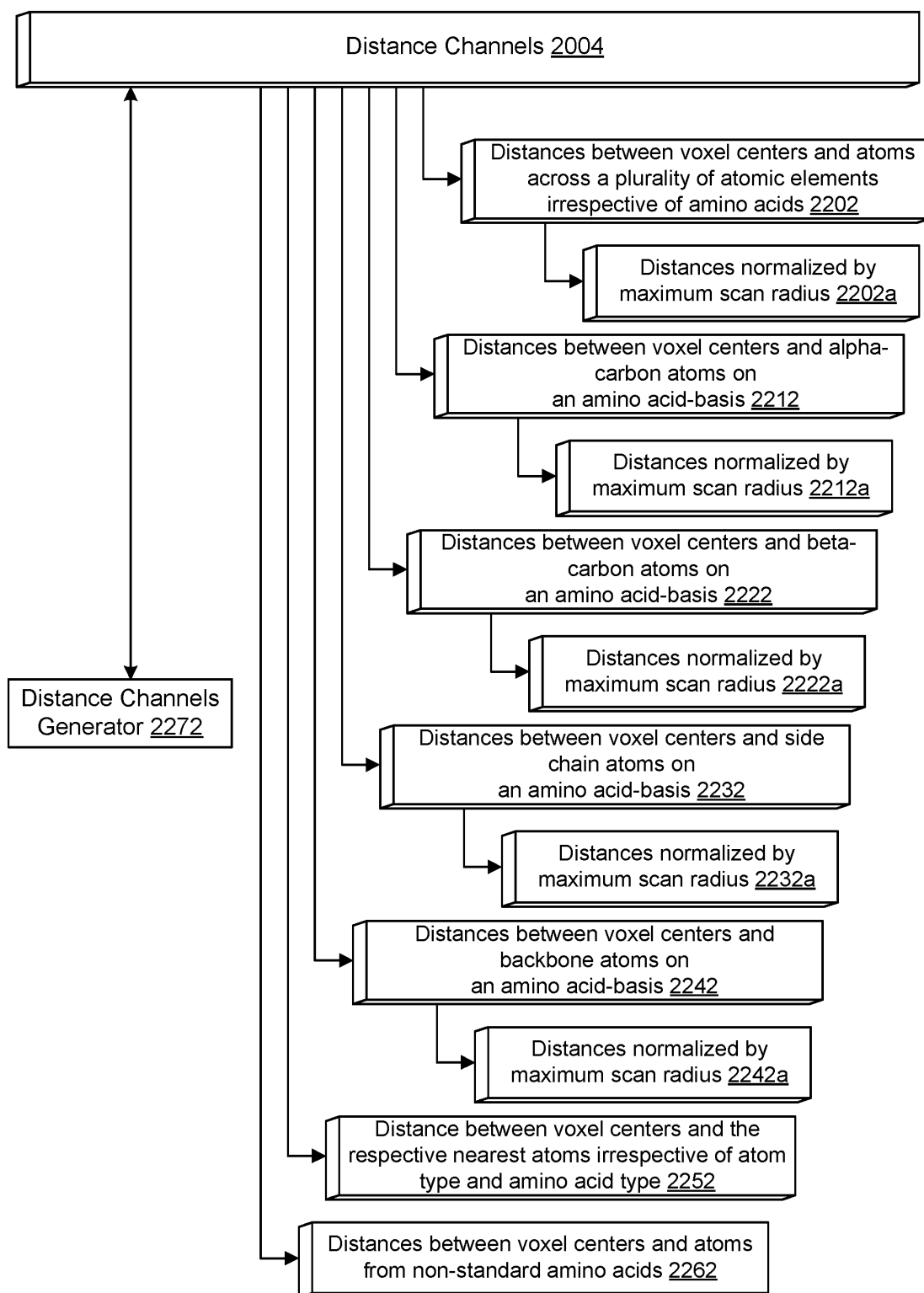


Figure 22

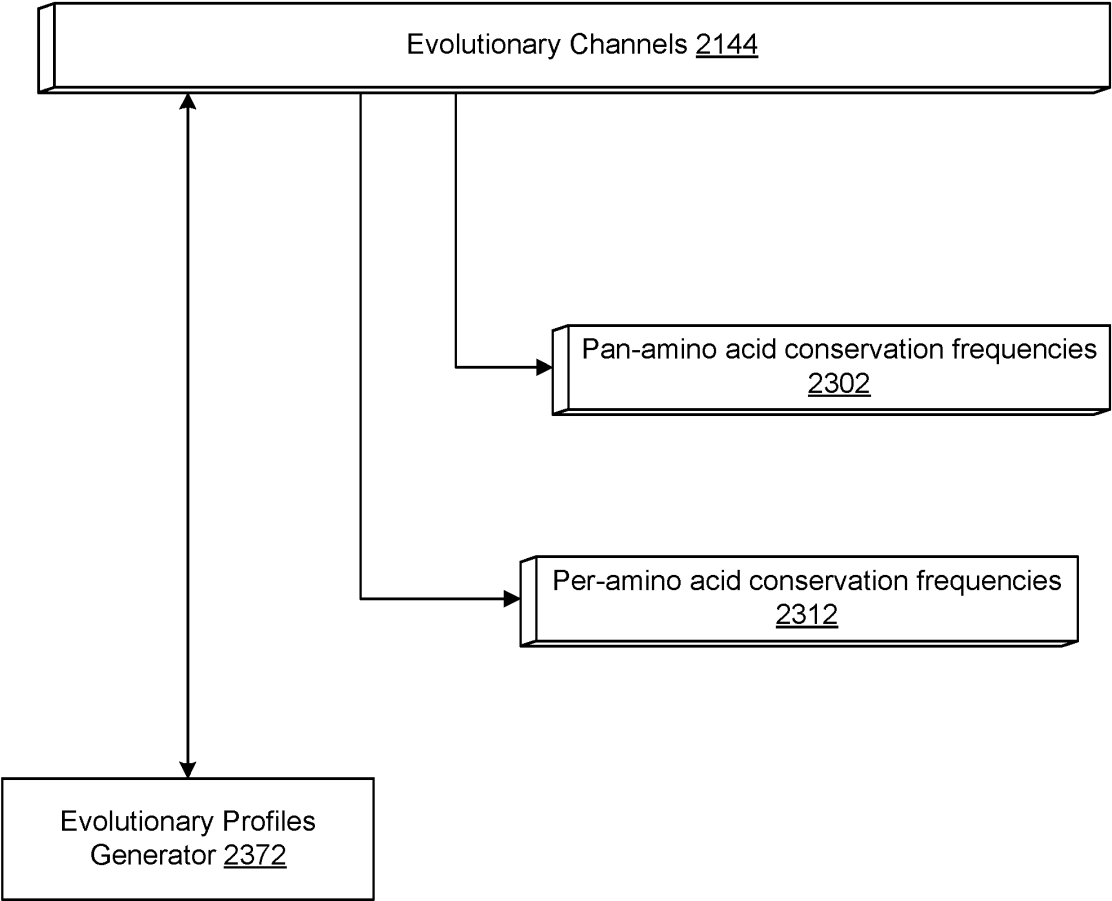


Figure 23

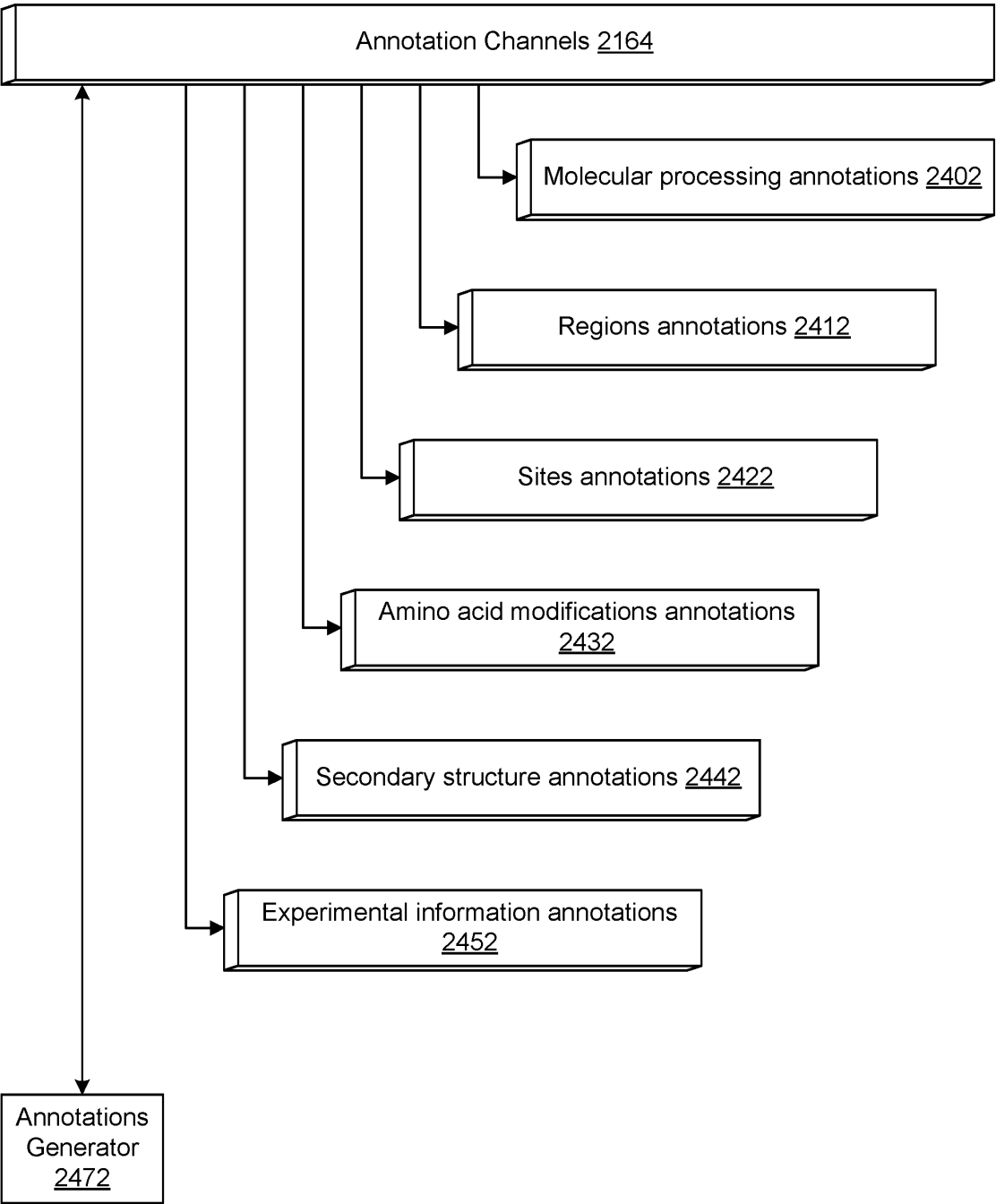


Figure 24

25 / 36

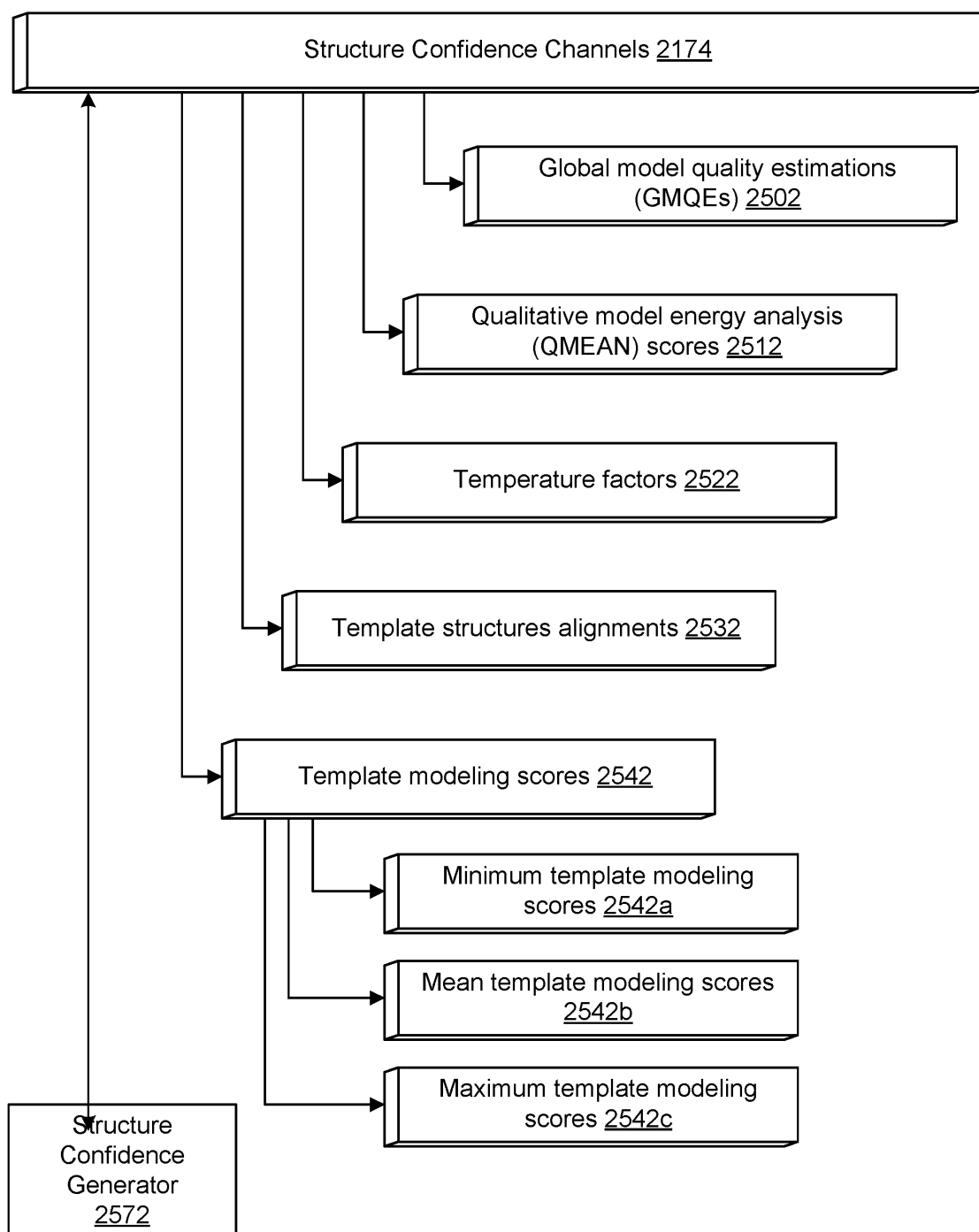


Figure 25

26 / 36

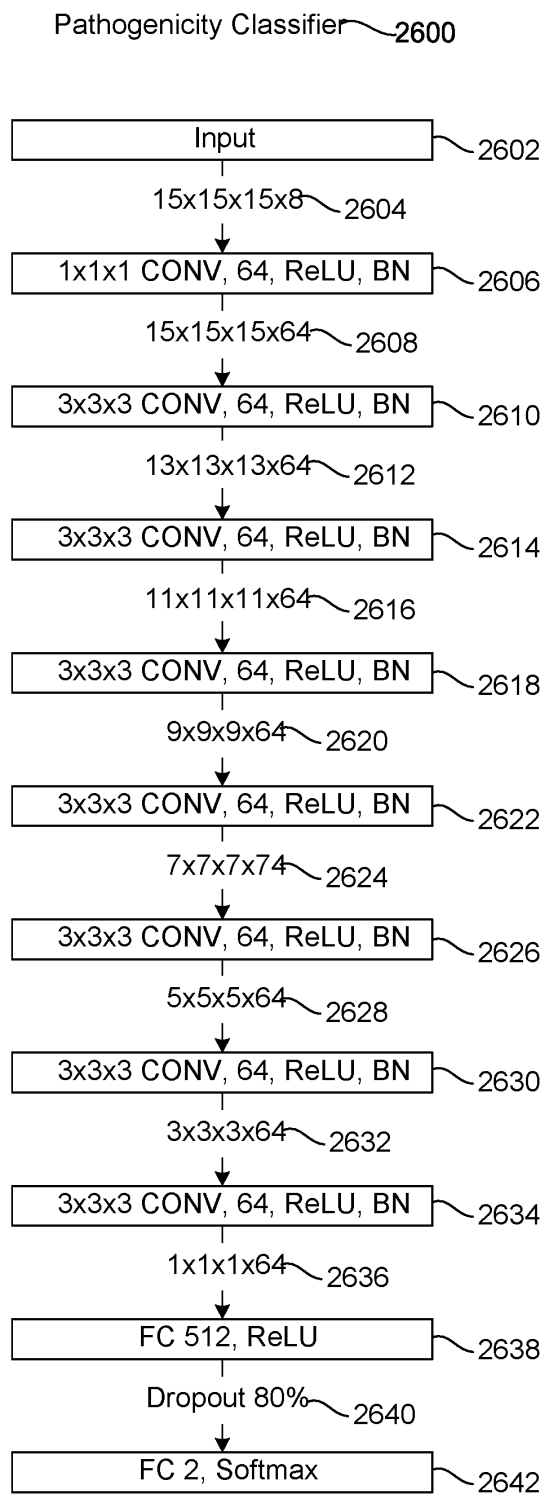


Figure 26

27 / 36

Pathogenicity Classifier 2700

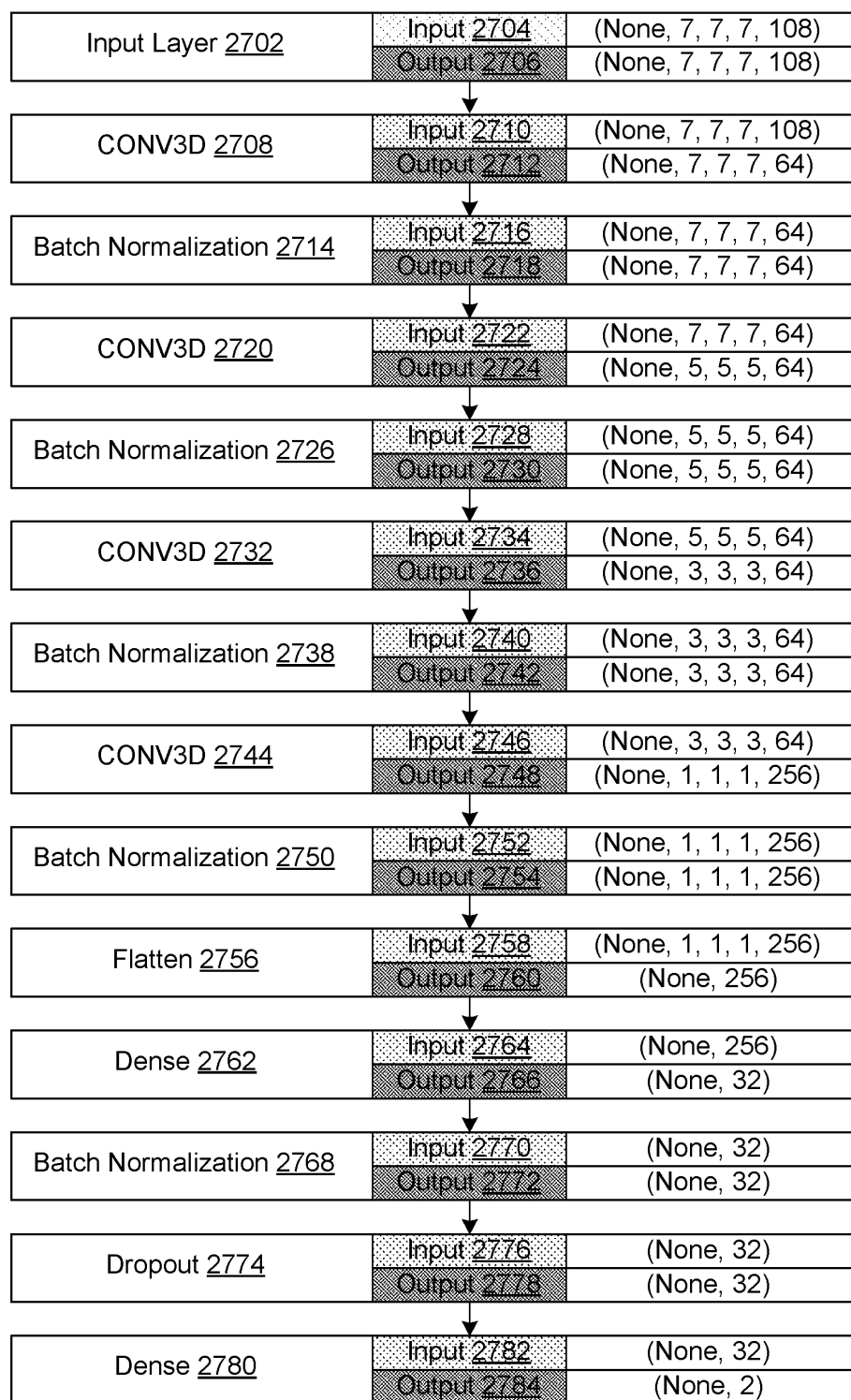


Figure 27

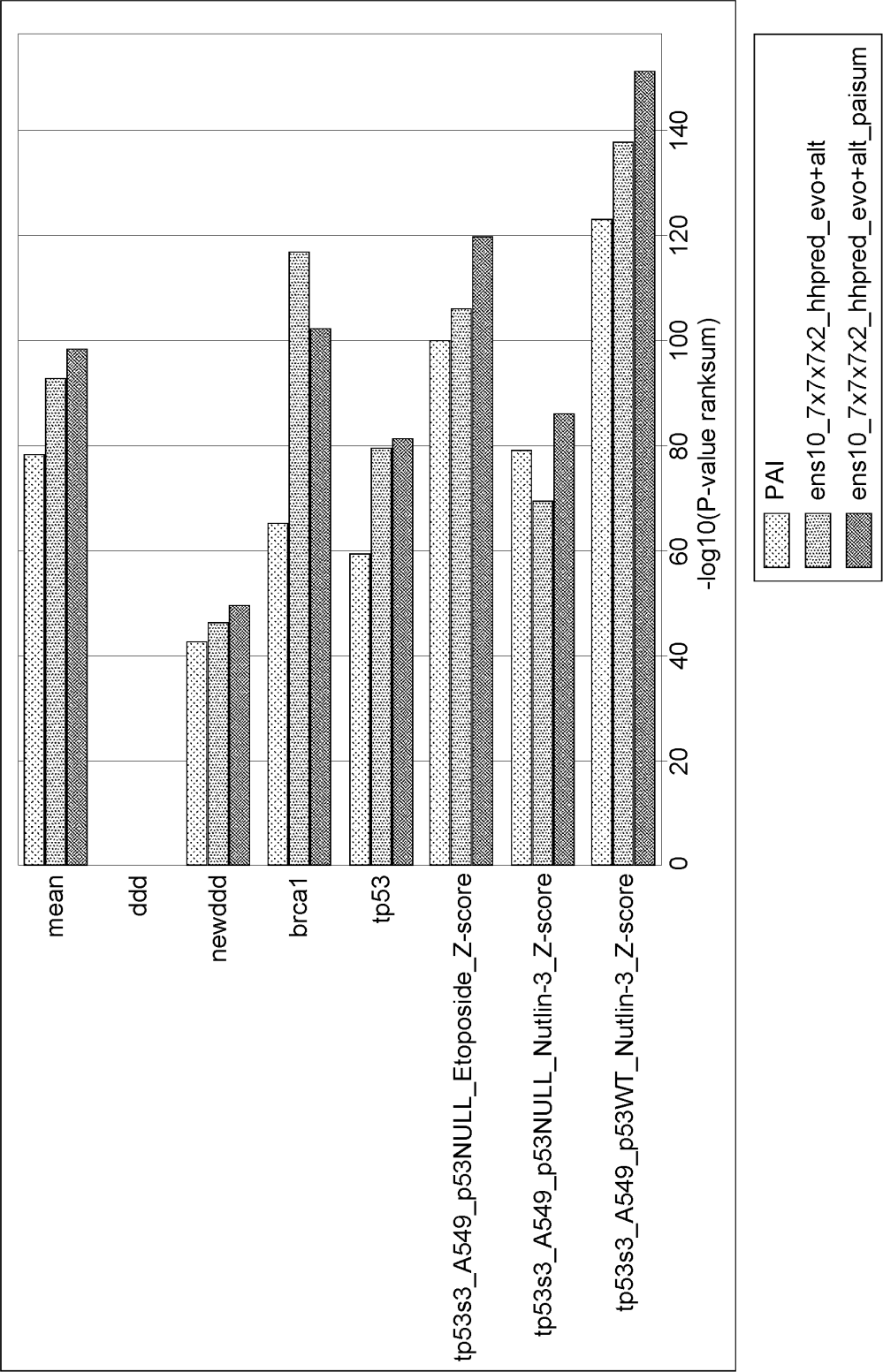


Figure 28

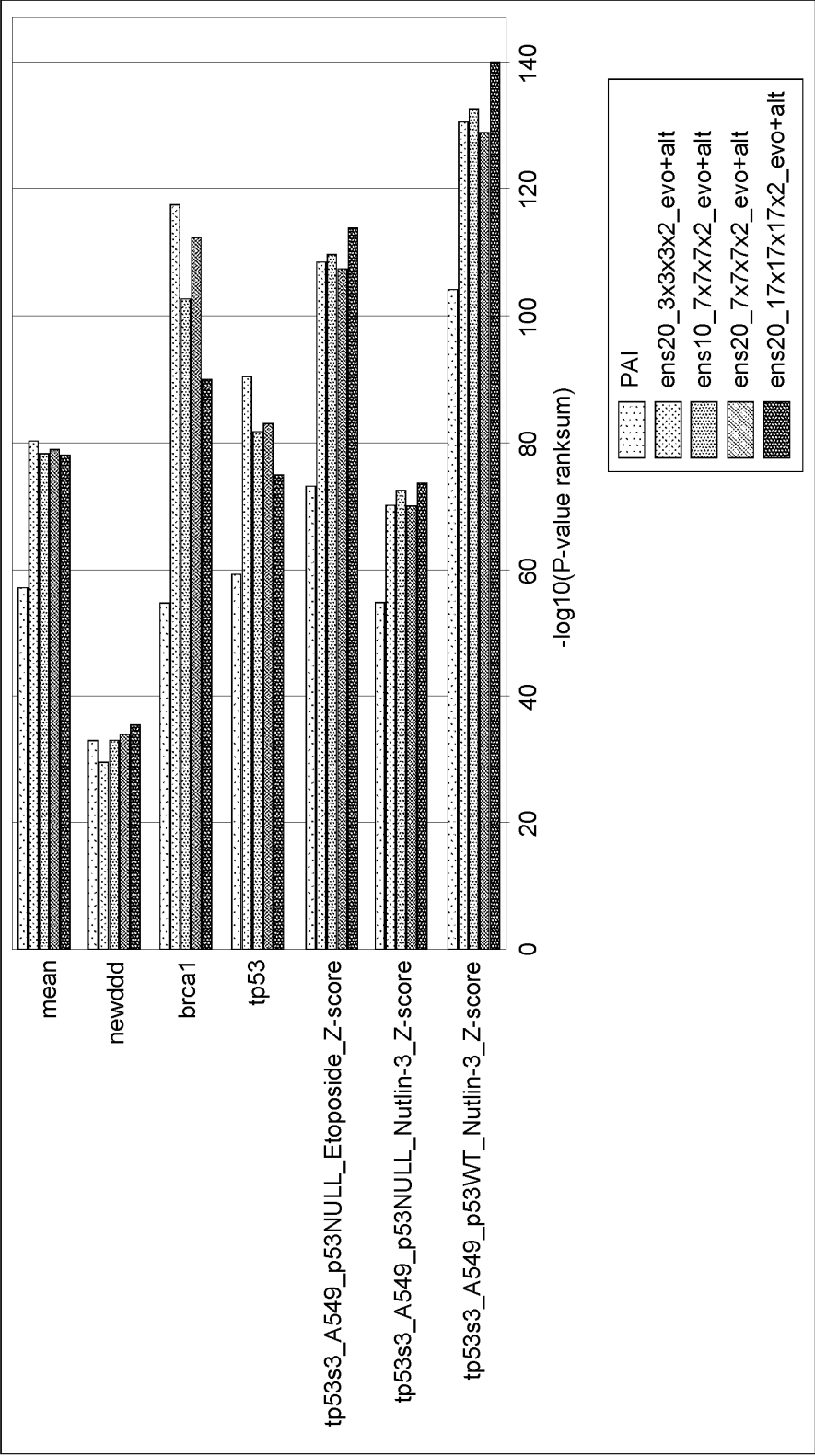


Figure 29

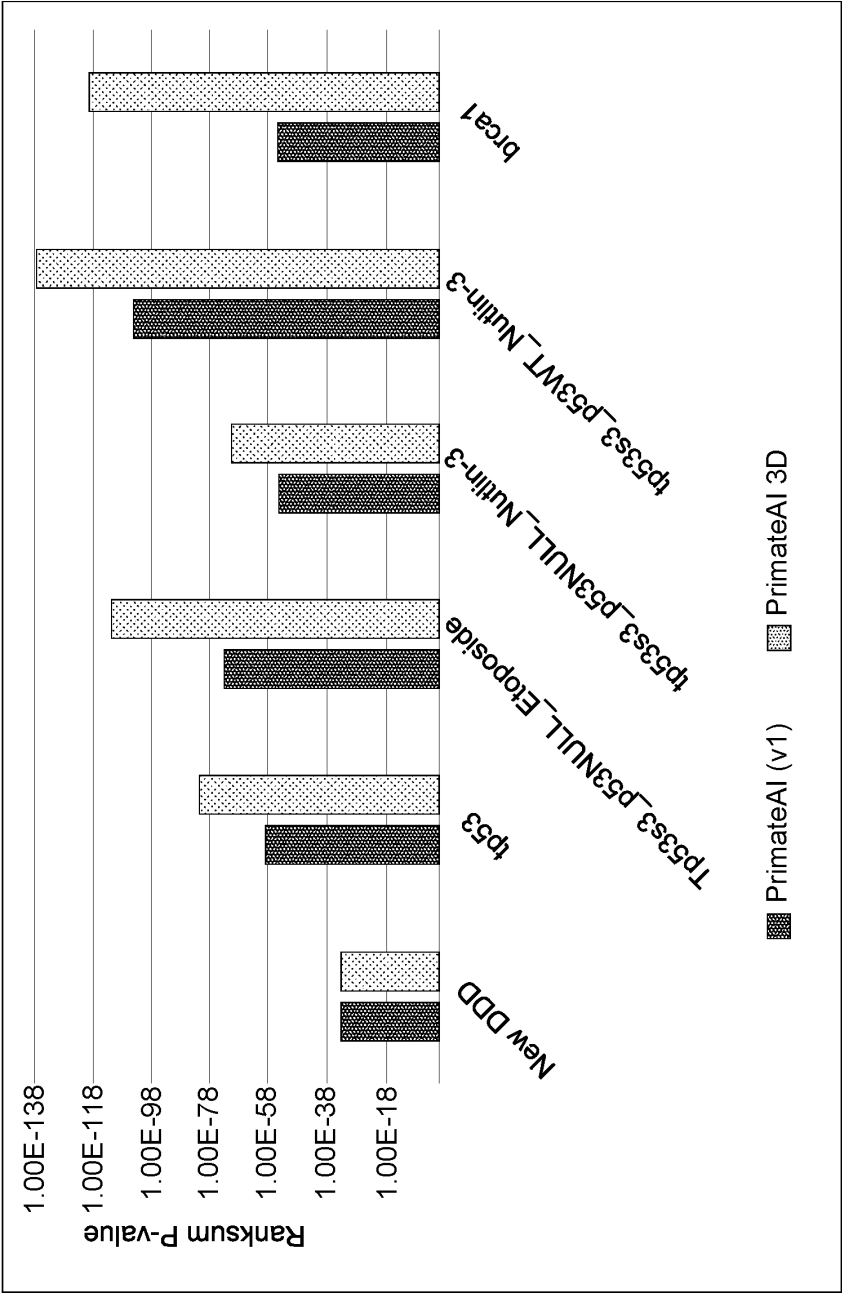


Figure 30

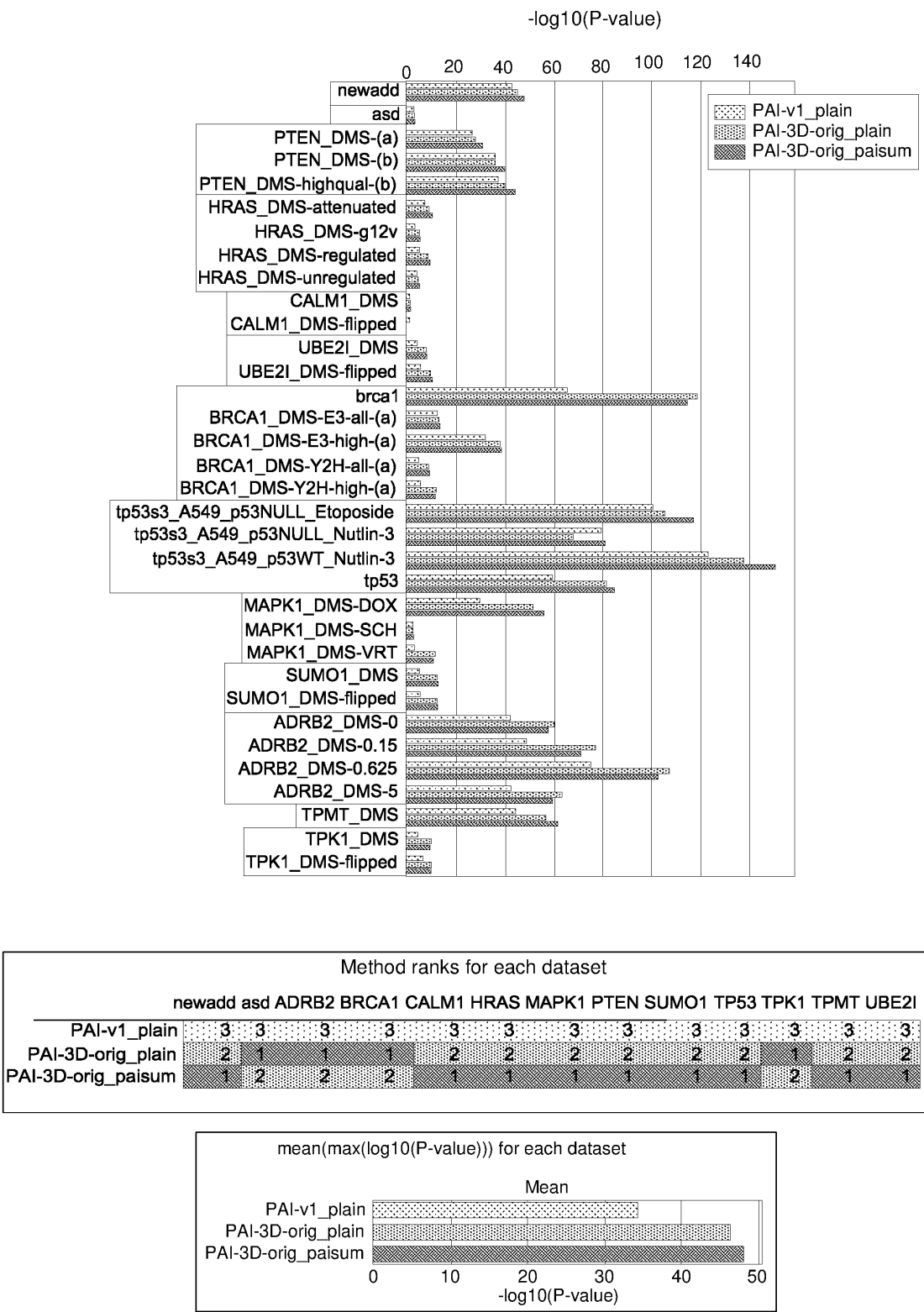


Figure 31

SUBSTITUTE SHEET (RULE 26)

Efficient Voxelization 3200

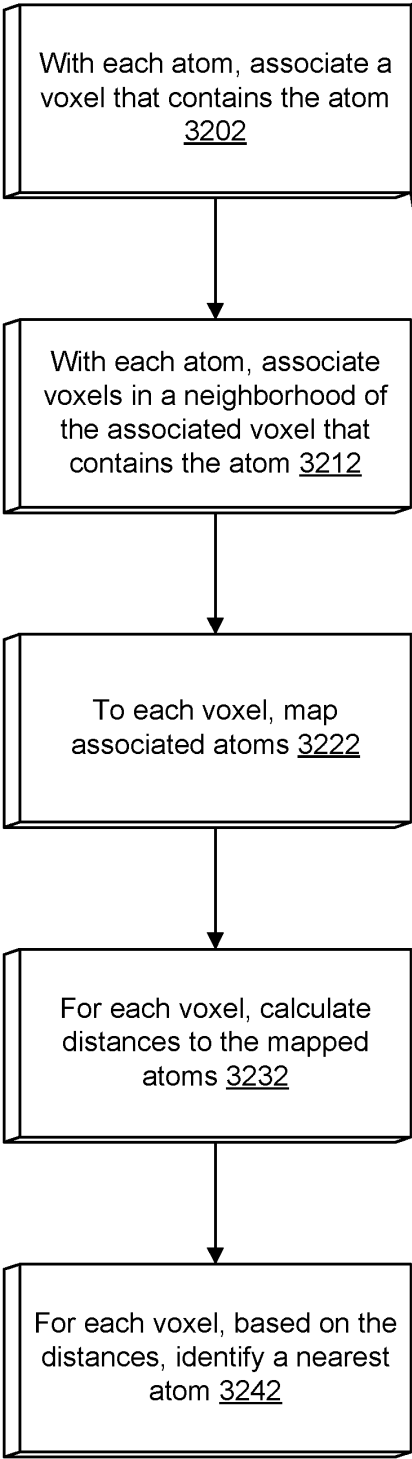


Figure 32A

Particular Atom

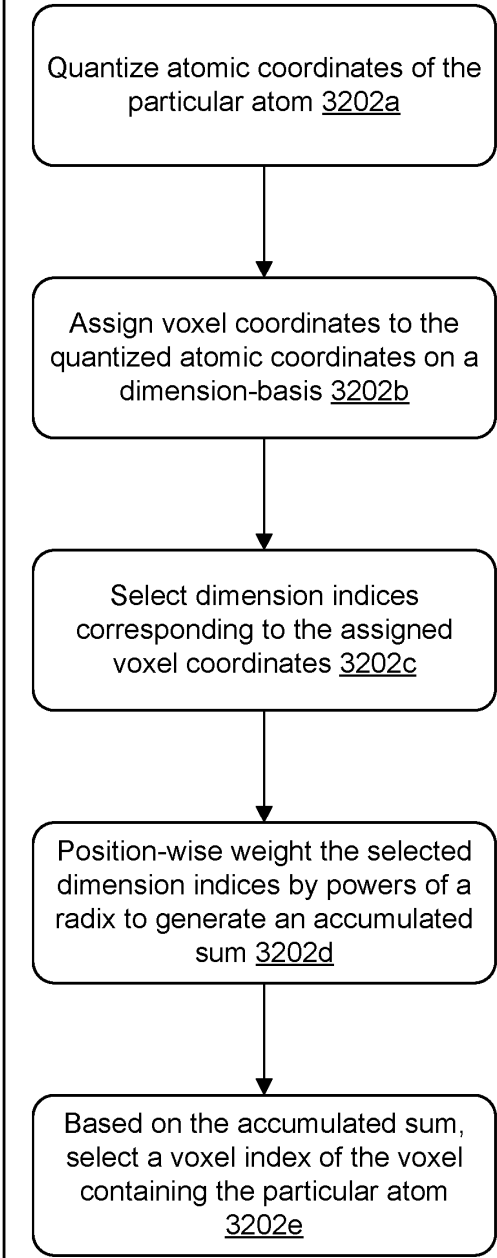
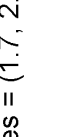
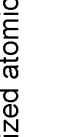
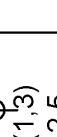
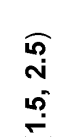
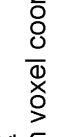


Figure 32B

$\underline{0}$	$\underline{1}$	$\underline{2}$
Voxel 0 	Voxel 1 	Voxel 2 
Voxel 3 	Voxel 4 	Voxel 5 
Voxel 6 	Voxel 7 	Voxel 8 

Voxel Grid 522

First Dimension Indices

First Dimension Voxel Coordinates Second Dimension Voxel Coordinates

Voxel Indices

Figure 33

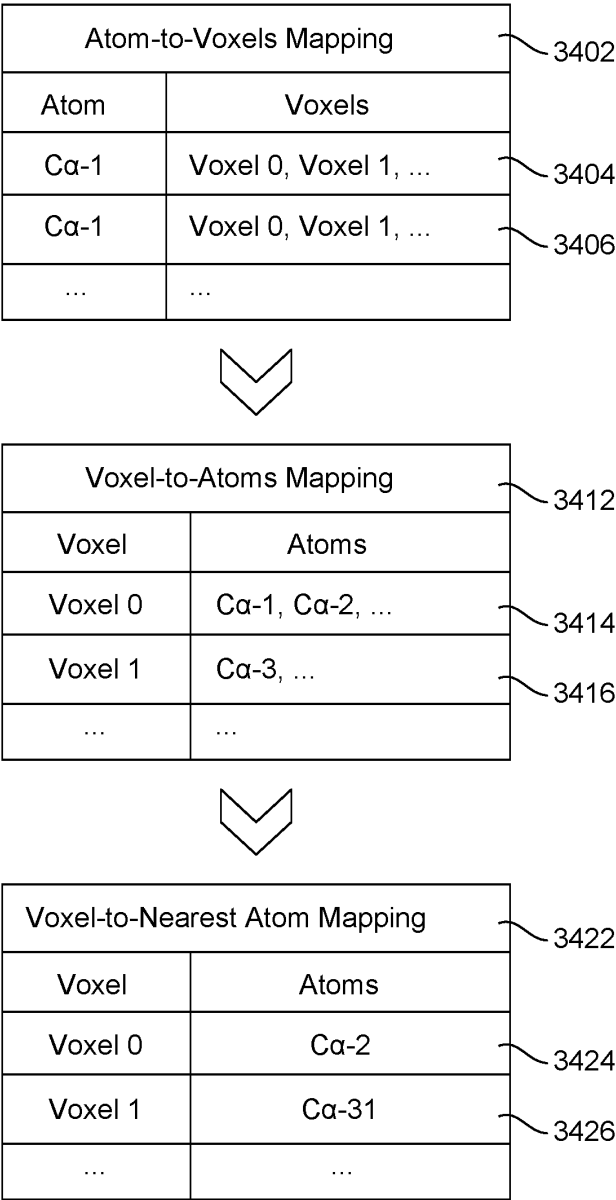


Figure 34

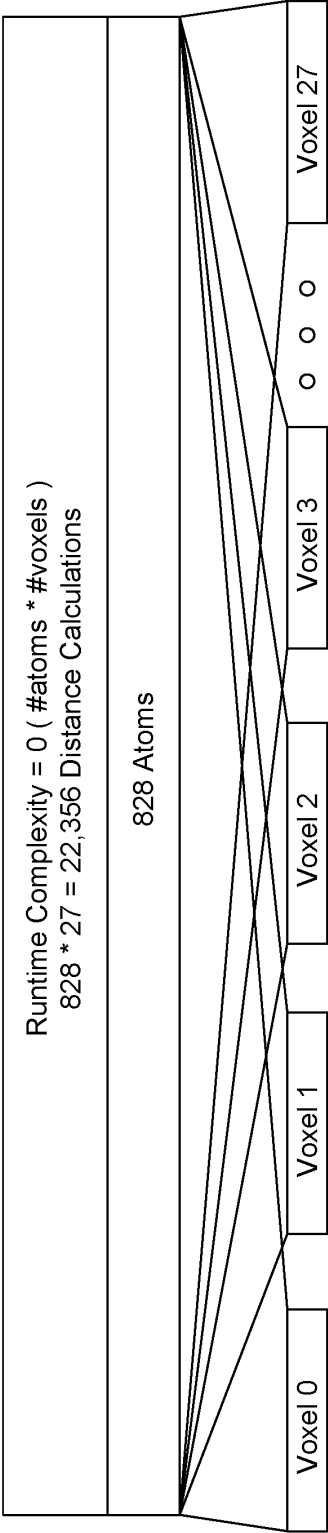


Figure 35A

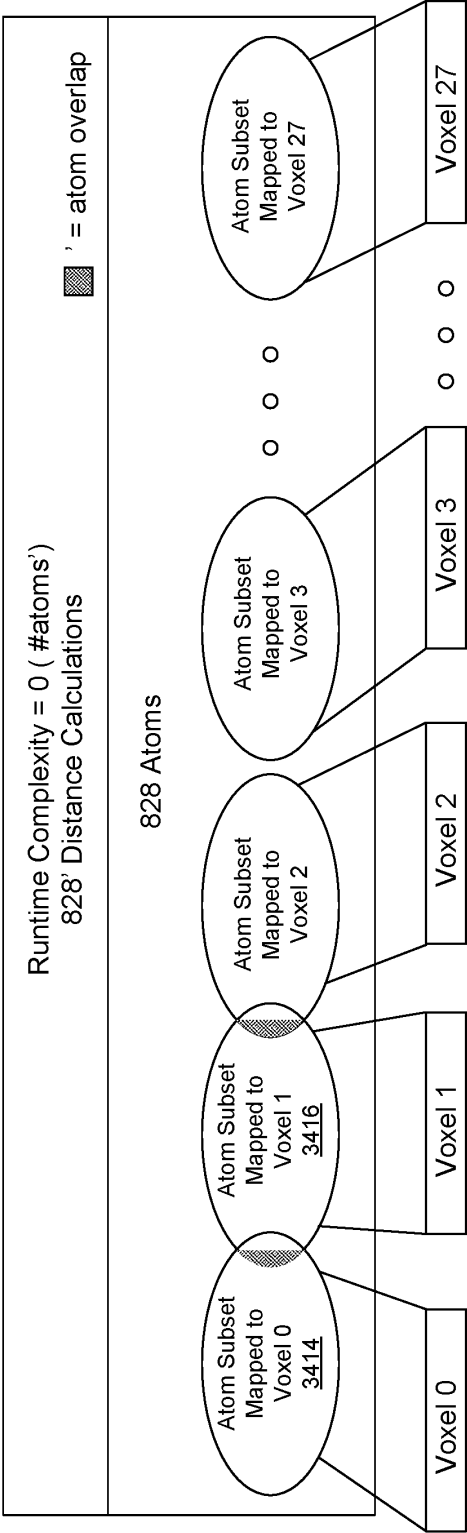


Figure 35B

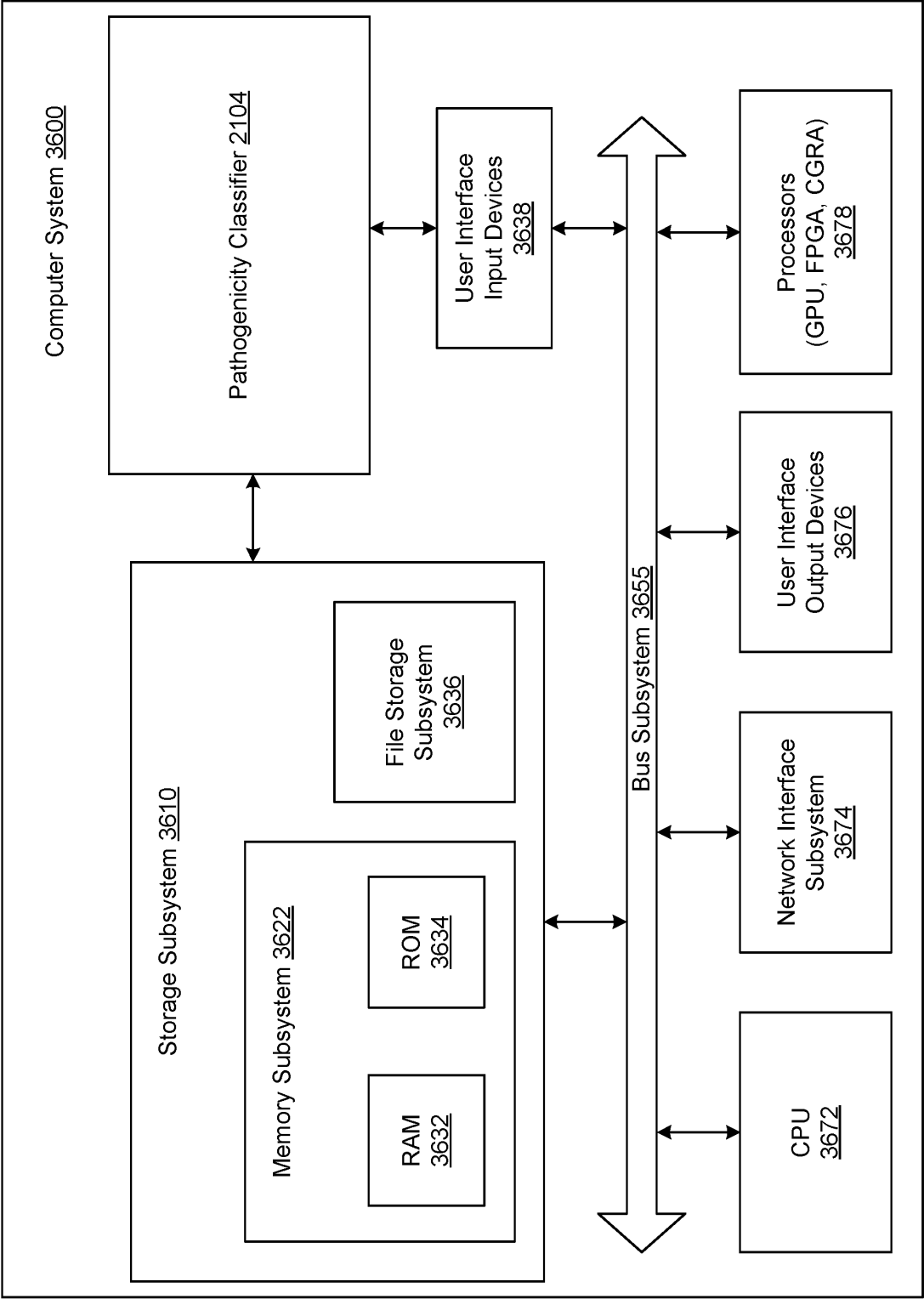


Figure 36