

The image displays a large table of interatomic distances and deviation matrices for 22 specific main chain oxygen atoms in selected neuraminidase proteins. The table is organized into five sections, each with a right-pointing arrow. Each section contains a list of atom IDs (e.g., O196, O197, O198) and a corresponding matrix of numerical values representing distances and deviations.

deviation: distance matrices → deviation matrix

DIST AND DEVIATION TUTORIAL: MAKING A POPULATION STANDARD DEVIATION MATRIX FOR THE INTERATOMIC DISTANCE OF 22 SPECIFIC MAIN CHAIN OXYGEN ATOMS IN SELECTED NEURAMINIDASE PROTEINS

Overview

Instructions

Summary

Tutorial Overview

This tutorial shows how to easily create a population standard deviation matrix for the interatomic distances within separate molecules. This tutorial will first show how to use **dist** to create distance matrices from selected lists of atoms (in PDB format). This tutorial will then show how to use **deviation** to create a population standard deviation matrix from these distance matrices.

- **dist** creates a distance matrix from a points list. There are options for handling input and output labels and specifying output format. **dist** can optionally parse a PDB file, and has options for specifying user defined output atom labels.
- **deviation** reads in multiple files and writes out the population standard deviation of numerical values that match string token positions across all input files.

The **dist** and **deviation** Overview Pages describes **dist** and **deviation** command line syntax.

Specific neuraminidases (N6 (1W1X), N10 (4FVK), N11 (4K3Y), influenza B (1A4G), and streptococcus (3H72)) contain distributed atoms that have common spatial occupancy, i.e. atoms that occupy similar relative positions in space.

Weininger, A.; Weininger, S. (2015) Using Common Spatial Distributions of Atoms to Relate Functionally Divergent Influenza Virus N10 and N11 Protein Structures to Functionally Characterized Neuraminidase Structures, Toxin Cell Entry Domains, and Non-Influenza Virus Cell Entry Domains.

PloS One 10(2):e0117499.

[dx.doi.org/10.1371/journal.pone.0117499](https://doi.org/10.1371/journal.pone.0117499)

(16 MB PDF – LOCAL COPY) (3 MB PDF – LOCAL COPY)

This tutorial will create a population standard deviation matrix for the interatomic distances of 22 selected main chain oxygen atoms from these neuraminidase proteins.

Instructions

1. Request and install **dist** and **deviation** (if you have not already done so).

Download, (from weiningerworks.com by using the links below) five files that are subsets of PDB structure files from rcsb.org:

1W1Xatoms.pdb

4FVKatoms.pdb

4K3Yatoms.pdb

1A4Gatoms.pdb

3H72atoms.pdb

2. These files should have a similar format to the following (and 1W1Xatoms.pdb should have identical content):

ATOM	293	O	ARG	A	124	32.542	4.910	59.543	1.00	14.80	O
ATOM	770	O	TRP	A	185	39.672	4.897	56.458	1.00	13.17	O
ATOM	784	O	SER	A	186	40.878	0.283	56.594	1.00	15.41	O
ATOM	825	O	ASP	A	192	34.555	-19.986	56.865	1.00	18.02	O
ATOM	904	O	GLY	A	203	48.222	7.443	56.825	1.00	18.27	O
ATOM	983	O	TYR	A	214	43.328	-13.981	64.025	1.00	16.85	O
ATOM	1105	O	LEU	A	230	44.646	2.786	51.864	1.00	18.46	O
ATOM	1113	O	ARG	A	231	41.929	-0.134	49.208	1.00	16.63	O
ATOM	1149	O	SER	A	235	33.584	-4.759	53.629	1.00	14.85	O
ATOM	1202	O	GLY	A	242	40.087	-18.835	53.462	1.00	18.72	O
ATOM	1255	O	ASP	A	250	46.557	2.834	43.371	1.00	19.94	O
ATOM	1263	O	GLY	A	251	44.703	6.760	44.397	1.00	20.15	O
ATOM	1389	O	GLY	A	267	46.987	-13.160	56.350	1.00	17.71	O
ATOM	1511	O	GLU	A	283	34.396	1.036	45.104	1.00	15.09	O
ATOM	1877	O	ASP	A	331	24.511	8.068	41.321	1.00	16.01	O
ATOM	1898	O	ARG	A	334	22.914	13.063	37.830	1.00	16.82	O
ATOM	2137	O	TRP	A	368	19.911	-4.598	43.297	1.00	19.40	O
ATOM	2218	O	ARG	A	378	22.228	12.223	51.392	1.00	17.66	O
ATOM	2235	O	GLY	A	380	18.788	8.473	49.691	1.00	17.40	O
ATOM	2492	O	SER	A	413	26.956	0.641	50.022	1.00	16.30	O
ATOM	2659	O	GLU	A	433	23.619	6.960	55.288	1.00	16.40	O
ATOM	2914	O	TRP	A	466	16.074	5.722	58.706	1.00	14.95	O

3. Run **dist** to create distance matrices for each of these lists of atoms. **dist** will output a single distance matrix file (1W1Xatoms.dist, 4FVKatoms.dist, 4K3Yatoms.dist, 1A4Gatoms.dist, 3H72atoms.dist in the following) for each **dist** execution. (Option '-i' specifies the input list filename, option '-o' specifies the output distance matrix filename, and option '-a' specifies that the input list is in PDB format.)

```
dist -i 1W1Xatoms.pdb -o 1W1Xatoms.dist -a
dist -i 4FVKatoms.pdb -o 4FVKatoms.dist -a
dist -i 4K3Yatoms.pdb -o 4K3Yatoms.dist -a
dist -i 1A4Gatoms.pdb -o 1A4Gatoms.dist -a
dist -i 3H72atoms.pdb -o 3H72atoms.dist -a
```

These files should have a similar format to the following (and 1W1Xatoms.dist should have identical content):

	O:293:R	O:770:W	O:784:S	O:825:D	O:904:G	O:983:Y	O:1105:L	O:1113:R	O:1149:S	O:1202:G	O:1255:D	O:1263:G	O:1389:G	O:1511:E	O:1877:D	O:1898:R	O:2137:W	O:2218:R	O:2235:G	O:2492:S	O:2659:E	O:2914:W	
O:293:R	0.000	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----
O:770:W	7.769	0.000	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----
O:784:S	9.980	4.771	0.000	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----
O:825:D	25.120	25.407	21.234	0.000	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----
O:904:G	16.114	8.929	10.259	30.645	0.000	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----
O:983:Y	22.210	20.664	16.269	12.818	23.125	0.000	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----
O:1105:L	14.491	7.092	6.545	25.405	7.687	20.755	0.000	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----
O:1113:R	14.845	9.109	7.472	22.519	12.451	20.328	4.792	0.000	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----
O:1149:S	11.382	11.760	9.350	15.597	19.323	16.973	13.506	10.515	0.000	----	----	----	----	----	----	----	----	----	----	----	----	----	----
O:1202:G	25.646	23.924	19.389	6.596	27.713	12.068	22.154	19.267	15.506	0.000	----	----	----	----	----	----	----	----	----	----	----	----	----
O:1255:D	21.500	14.931	14.615	29.101	14.319	26.828	8.705	8.019	18.198	24.764	0.000	----	----	----	----	----	----	----	----	----	----	----	----
O:1263:G	19.512	13.200	14.330	31.205	12.935	28.589	8.459	8.853	18.481	27.542	4.461	0.000	----	----	----	----	----	----	----	----	----	----	----
O:1389:G	23.353	19.483	14.768	14.192	20.645	8.542	16.730	15.693	16.051	9.389	20.602	23.343	0.000	----	----	----	----	----	----	----	----	----	----
O:1511:E	15.064	13.102	13.214	24.089	19.225	25.755	12.403	8.658	10.340	22.296	12.415	11.811	22.057	0.000	----	----	----	----	----	----	----	----	----
O:1877:D	20.162	21.657	23.701	33.608	28.337	36.820	23.334	20.805	19.958	33.373	22.751	20.467	34.375	12.707	0.000	----	----	----	----	----	----	----	----
O:1898:R	25.112	26.354	28.950	39.876	32.139	42.829	27.836	25.791	26.098	39.456	26.350	23.614	40.127	18.149	6.300	0.000	----	----	----	----	----	----	----
O:2137:W	22.669	25.571	25.303	25.206	33.608	32.650	27.198	23.231	17.138	26.704	27.663	27.292	31.254	15.647	13.620	18.730	0.000	----	----	----	----	----	----
O:2218:R	15.043	19.586	22.747	34.919	26.982	35.937	24.328	23.358	20.551	35.886	27.284	24.164	35.803	17.685	11.131	13.605	18.811	0.000	----	----	----	----	----
O:2235:G	17.290	22.242	24.550	33.316	30.304	36.220	26.565	24.695	20.236	34.837	29.032	26.506	36.180	17.887	10.148	13.371	14.594	5.366	0.000	----	----	----	----
O:2492:S	11.835	14.874	15.399	23.023	23.341	26.037	17.915	15.015	9.279	23.740	20.815	19.597	25.135	8.927	11.698	17.869	11.059	12.585	11.321	0.000	----	----	----
O:2659:E	10.096	16.227	18.552	29.123	24.656	30.055	21.709	20.556	15.472	30.658	26.176	23.732	30.855	15.967	14.039	18.507	17.062	6.694	7.547	8.877	0.000	----	
O:2914:W	16.509	23.719	25.481	31.715	32.249	34.048	29.526	28.160	21.029	34.744	34.245	32.023	36.300	23.295	19.466	23.162	18.938	11.560	9.808	14.820	8.375	0.000	

4. Run **deviation** to create a single population standard deviation matrix, named 'neuraminidases.deviation', from the newly created five distance matrices.

(Option '-o' specifies the output population standard deviation matrix filename, the input files are listed at the end of the **deviation** command.)

```
deviation -o neuraminidases.deviation 1W1Xatoms.dist
4FVKatoms.dist 4K3Yatoms.dist 1A4Gatoms.dist 3H72atoms.dist
```

The newly created file 'neuraminidases.deviation' should have identical content to the following:

1W1X->	O:293:R	O:770:W	O:784:S	O:825:D	O:904:G	O:983:Y	O:1105:L	O:1113:R	O:1149:S	O:1202:G	O:1255:D	O:1263:G	O:1389:G	O:1511:E	O:1877:D	O:1898:R	O:2137:W	O:2218:R	O:2235:G	O:2492:S	O:2659:E	O:2914:W	
4FVK->	O:283:R	O:557:W	O:756:S	O:796:D	O:881:G	O:973:Y	O:1090:L	O:1098:R	O:1132:S	O:1184:G	O:1239:D	O:1247:G	O:1378:G	O:1509:E	O:1887:D	O:1913:R	O:2129:W	O:2454:R	O:2224:G	O:2465:S	O:2633:E	O:2851:W	
4K3Y->	O:273:R	O:856:W	O:670:S	O:709:D	O:795:G	O:851:Y	O:964:L	O:972:R	O:1004:S	O:1052:G	O:1109:D	O:1117:G	O:1246:G	O:1364:E	O:1748:D	O:1770:R	O:1990:W	O:2331:R	O:2086:G	O:2342:S	O:2510:E	O:2708:W	
1A4G->	O:312:R	O:796:W	O:810:S	O:847:D	O:939:G	O:1019:Y	O:1152:L	O:1160:R	O:1196:S	O:1239:G	O:1300:D	O:1308:G	O:1429:G	O:1559:E	O:1943:D	O:1965:R	O:2232:W	O:2321:R	O:2340:G	O:2610:S	O:2748:E	O:2956:W	
3H72->	O:207:R	O:405:W	O:744:D	O:791:D	O:943:G	O:1737:Y	O:1949:L	O:1957:G	O:1983:T	O:2066:G	O:2119:Y	O:2131:T	O:2302:G	O:2569:E	O:2848:D	O:2873:R	O:3064:Y	O:3143:R	O:3171:G	O:3408:N	O:3525:E	O:3678:W	
O:293:R	0.000																						
O:293:R		2.581																					
O:273:R																							
O:312:R																							
O:207:R																							
O:770:W																							
O:557:W																							
O:656:W																							
O:796:W																							
O:405:W																							
O:784:S																							
O:756:S																							
O:670:S																							
O:810:S																							
O:744:D	0.410	3.789	0.000																				
O:825:D																							
O:796:D																							
O:709:D																							
O:847:D																							
O:791:D	0.485	3.293	0.530	0.000																			
O:904:G																							
O:881:G																							
O:795:G																							
O:939:G																							
O:943:G	0.601	0.686	1.357	1.027	0.000																		
O:983:Y																							
O:973:Y																							
O:851:Y																							
O:1019:Y																							
O:1737:Y	1.174	1.800	0.525	0.276	1.410	0.000																	
O:1105:L																							
O:1090:L																							
O:964:L																							
O:1152:L																							
O:1949:L	0.908	3.665	0.729	0.600	1.639	0.473	0.000																
O:1113:R																							
O:1098:R																							
O:972:R																							
O:1160:R																							
O:1957:G	0.287	3.954	0.112	0.189	1.234	0.140	0.069	0.000															
O:1149:S																							
O:1132:S																							
O:1004:S																							
O:1196:S																							
O:1983:T	1.331	5.232	1.125	0.730	2.121	0.230	1.348	1.091	0.000														
O:1202:G																							
O:1184:G																							
O:1052:G																							
O:1239:G																							
O:2066:G	0.198	3.937	0.407	0.950	1.559	2.002	0.580	0.304	0.916	0.000													
O:1255:D																							
O:1239:D																							
O:1109:D																							
O:1300:D																							
O:2119:Y	1.369	2.445	1.559	1.629	1.488	1.598	1.389	1.787	0.867	1.615	0.000												
O:1263:G																							
O:1247:G																							
O:1117:G																							
O:1308:G																							
O:2131:T	0.434	2.990	0.148	0.892	1.341	0.450	0.328	0.558	0.219	1.294	0.452	0.000											
O:1389:G																							
O:1378:G																							
O:1246:G																							
O:1429:G																							
O:2302:G	0.648	3.367	0.676	0.940	2.324	0.988	0.128	0.484	0.587	0.367	0.865	0.366	0.000										
O:1511:E																							
O:1509:E																							
O:1364:E																							
O:1559:E																							
O:2569:E	0.253	4.215	0.377	0.296	1.260	0.398	0.506	0.212	0.331	0.730	0.856	0.878	0.369	0.000									
O:1877:D																							
O:1887:D																							
O:1748:D																							
O:1943:D																							
O:2848:D	3.899	5.631	1.155	1.464	1.529																		

Tutorial Summary

dist created distance matrices from selected lists of atoms (in PDB format).

deviation created a population standard deviation matrix from these distance matrices.

If you are interested in identifying atoms in molecules that have common spatial occupancy (i.e. atoms that occupy similar relative positions), check out the Weinger Works™ program [wwavePDB](#).

Arthur Weinger and Susan Weinger do not accept confidential information under any circumstance. Weinger Works™ is a trademark of Arthur Weinger and Susan Weinger. All images constitute content of this website and are copyrighted material of this website.

Copyright © 2013 – 2023 by Arthur Weinger and Susan Weinger. All Rights Reserved.