WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 20, 2024 – 12:18 pm GMT

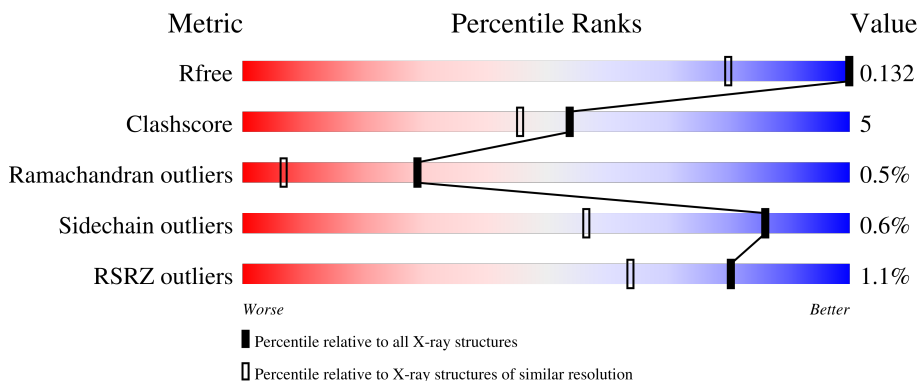|		
PDB ID	:	7AF2
Title	:	Salmonella typhimurium neuraminidase mutant (D62G)
Authors	:	Salinger, M.T.; Kuhn, P.; Laver, W.G.; Pape, T.; Schneider, T.R.; Sheldrick, G.M.; Vimr, E.R.; Garman, E.F.
Deposited on	:	2020-09-19
Resolution	:	0.79 Å(reported)

1   Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 0.79 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.132
Clashscore		5
Ramachandran outliers		0.5%
Sidechain outliers		0.6%
RSRZ outliers		1.1%

Worse   *Better*

▌ Percentile relative to all X-ray structures

❘ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1079 (1.04-0.56)
Clashscore	141614	1153 (1.04-0.56)
Ramachandran outliers	138981	1071 (1.04-0.60)
Sidechain outliers	138945	1072 (1.04-0.60)
RSRZ outliers	127900	1045 (1.04-0.56)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	AAA	379	87% 12% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GOL	AAA	503	-	-	X	-

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GOL	AAA	504	-	X	-	-
2	GOL	AAA	506	-	X	-	-

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 7168 atoms, of which 3300 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Sialidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	AAA	379	6514	2034	3260	566	641	13	206	49	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	62	GLY	ASP	engineered mutation	UNP P29768

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
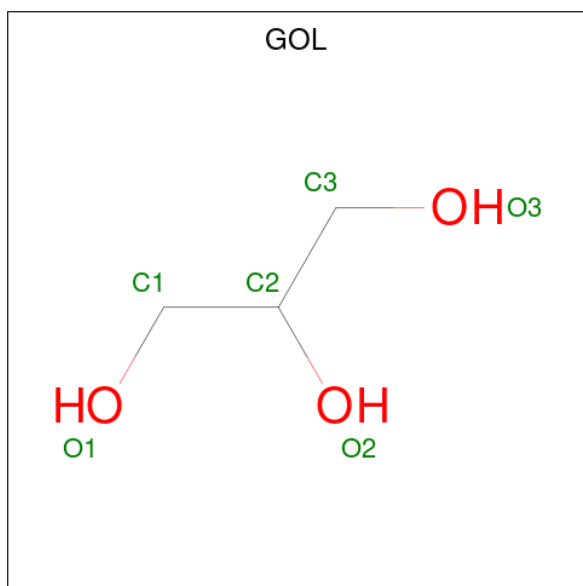
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	AAA	1	14	3	8	3	2	0
2	AAA	1	14	3	8	3	2	0
2	AAA	1	14	3	8	3	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	AAA	1	Total 14	C 3	H 8	O 3	2	0
2	AAA	1	Total 14	C 3	H 8	O 3	2	0

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).
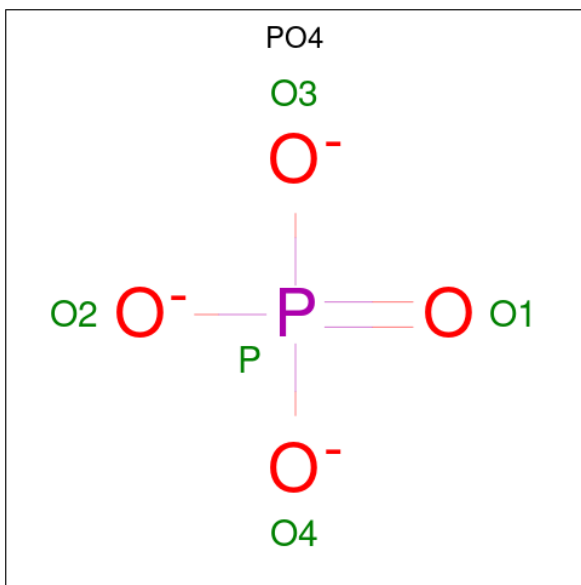
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	AAA	1	Total 10	O 8	P 2	0	1

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	502	Total 574	O 574	0	71

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: Sialidase
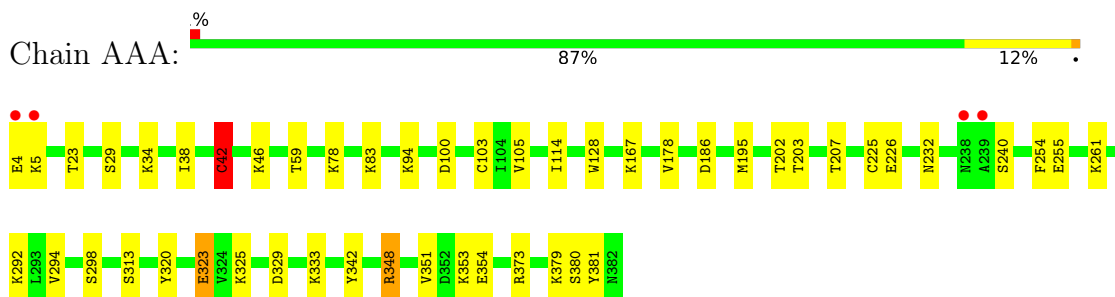
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	46.91Å 81.36Å 91.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	18.10 – 0.79 18.10 – 0.79	Depositor EDS
% Data completeness (in resolution range)	92.9 (18.10-0.79) 92.9 (18.10-0.79)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	0.94 (at 0.79Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.121 , 0.132 0.121 , 0.132	Depositor DCC
R_{free} test set	17531 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	12.3	Xtriage
Anisotropy	0.070	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.55 , 87.3	EDS
L-test for twinning[2]	$< |L| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.99	EDS
Total number of atoms	7168	wwPDB-VP
Average B, all atoms (Å^2)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.26% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, PO4

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$ >5	RMSZ	#$|Z|$ >5
1	AAA	0.84	10/3387 (0.3%)	1.15	19/4565 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	AAA	0	4

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	29[A]	SER	C-O	-8.10	1.07	1.23
1	AAA	29[B]	SER	C-O	-8.10	1.07	1.23
1	AAA	323[A]	GLU	CD-OE1	6.70	1.33	1.25
1	AAA	323[B]	GLU	CD-OE1	6.70	1.33	1.25
1	AAA	83[A]	LYS	C-O	6.24	1.35	1.23

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	AAA	195[A]	MET	CG-SD-CE	-22.85	63.64	100.20
1	AAA	195[B]	MET	CG-SD-CE	-22.85	63.64	100.20
1	AAA	42[A]	CYS	O-C-N	-9.24	107.91	122.70
1	AAA	42[B]	CYS	O-C-N	-9.24	107.91	122.70
1	AAA	373[A]	ARG	NE-CZ-NH1	7.01	123.80	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	379[A]	LYS	Mainchain
1	AAA	379[B]	LYS	Mainchain
1	AAA	42[A]	CYS	Mainchain
1	AAA	42[B]	CYS	Mainchain

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	AAA	3254	3260	3192	32	1
2	AAA	30	40	39	6	0
3	AAA	10	0	0	1	0
4	AAA	574	0	0	17	1
All	All	3868	3300	3231	34	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 34 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:AAA:505[A]:PO4:O4	4:AAA:601:HOH:O	1.71	1.08
1:AAA:46:LYS:HE2	4:AAA:997[A]:HOH:O	1.60	0.99
1:AAA:261[A]:LYS:HE2	4:AAA:957:HOH:O	1.77	0.84
1:AAA:323[B]:GLU:OE2	4:AAA:603:HOH:O	1.97	0.82
1:AAA:226[B]:GLU:OE2	4:AAA:602:HOH:O	1.96	0.82

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:78[B]:LYS:HZ3	4:AAA:1096:HOH:O[1_455]	1.41	0.19

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	AAA	424/379 (112%)	407 (96%)	15 (4%)	2 (0%)	29	7

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	178	VAL
1	AAA	351	VAL

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	AAA	372/324 (115%)	369 (99%)	3 (1%)	81	50

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	128	TRP
1	AAA	167[A]	LYS
1	AAA	167[B]	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

7 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	GOL	AAA	506	-	5,5,5	0.93	0	5,5,5	2.12	2 (40%)
2	GOL	AAA	504	-	5,5,5	1.02	0	5,5,5	2.16	3 (60%)
2	GOL	AAA	501	-	5,5,5	0.31	0	5,5,5	1.16	1 (20%)
2	GOL	AAA	502	-	5,5,5	0.17	0	5,5,5	0.42	0
2	GOL	AAA	503	-	5,5,5	0.88	0	5,5,5	1.93	1 (20%)
3	PO4	AAA	505[A]	-	4,4,4	1.12	0	6,6,6	0.91	0
3	PO4	AAA	505[C]	-	4,4,4	4.63	2 (50%)	6,6,6	0.60	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	AAA	506	-	-	4/4/4/4	-
2	GOL	AAA	504	-	-	4/4/4/4	-

Continued on next page...

W O R L D W I D E
PDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	AAA	501	-	-	0/4/4/4	-
2	GOL	AAA	502	-	-	0/4/4/4	-
2	GOL	AAA	503	-	-	1/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	AAA	505[C]	PO4	P-O1	-8.38	1.31	1.50
3	AAA	505[C]	PO4	P-O2	-3.52	1.44	1.54

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	AAA	503	GOL	O1-C1-C2	-3.87	91.67	110.20
2	AAA	506	GOL	O2-C2-C1	3.57	124.87	109.12
2	AAA	504	GOL	O3-C3-C2	3.03	124.72	110.20
2	AAA	504	GOL	O1-C1-C2	2.56	122.47	110.20
2	AAA	506	GOL	O2-C2-C3	2.34	119.45	109.12

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	AAA	504	GOL	O1-C1-C2-C3
2	AAA	504	GOL	C1-C2-C3-O3
2	AAA	506	GOL	O1-C1-C2-C3
2	AAA	506	GOL	C1-C2-C3-O3
2	AAA	504	GOL	O2-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	AAA	504	GOL	1	0
2	AAA	503	GOL	5	0
3	AAA	505[A]	PO4	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	AAA	379/379 (100%)	-0.56	4 (1%)	80	64	10, 13, 20, 118	1 (0%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	239	ALA	8.6
1	AAA	5[A]	LYS	6.8
1	AAA	4[B]	GLU	3.2
1	AAA	238	ASN	3.1

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
2	GOL	AAA	502	6/6	0.79	0.12	15,16,19,19	14
2	GOL	AAA	503	6/6	0.82	0.20	15,27,58,58	14
2	GOL	AAA	506	6/6	0.89	0.10	17,17,29,29	14

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	PO4	AAA	505[A]	5/5	0.90	0.15	12,13,16,17	5
3	PO4	AAA	505[C]	5/5	0.90	0.15	12,22,25,28	5
2	GOL	AAA	504	6/6	0.92	0.14	14,15,29,29	14
2	GOL	AAA	501	6/6	0.96	0.13	17,25,46,46	14

6.5 Other polymers ⓘ

There are no such residues in this entry.