



Full wwPDB X-ray Structure Validation Report ⓘ

May 20, 2024 – 02:17 PM EDT

PDB ID : 8SX0
Title : Bordetella filamentous hemagglutinin (FhaB) C-terminal domain
Authors : Cuthbert, B.J.; Goulding, C.W.; Hayes, C.S.; Costello, M.S.
Deposited on : 2023-05-19
Resolution : 1.65 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : **FAILED**
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36.2

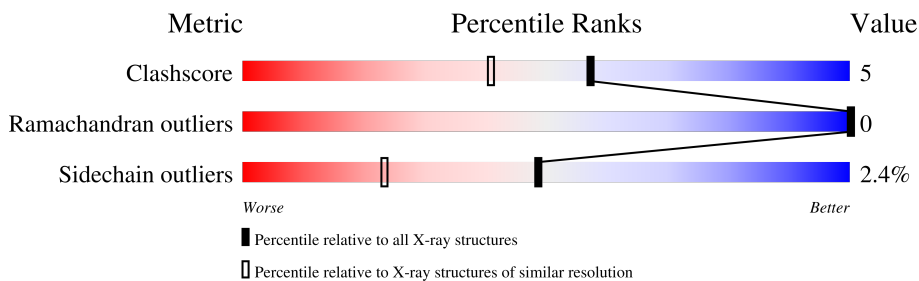
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.65 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)

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 $\leq 5\%$.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	112	 81% . . 13%
1	B	112	 75% 11% . 13%

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 3358 atoms, of which 1623 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 Filamentous hemagglutinin.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S	Se			
1	A	96	1634	503	827	152	150	1	1	0	6	0
1	B	96	1568	489	788	142	147	1	1	0	4	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3600	MSE	-	initiating methionine	UNP A0A381A265
A	3601	ALA	-	expression tag	UNP A0A381A265
A	3602	LYS	-	expression tag	UNP A0A381A265
A	3603	SER	-	expression tag	UNP A0A381A265
A	3604	HIS	-	expression tag	UNP A0A381A265
A	3605	HIS	-	expression tag	UNP A0A381A265
A	3606	HIS	-	expression tag	UNP A0A381A265
A	3607	HIS	-	expression tag	UNP A0A381A265
A	3608	HIS	-	expression tag	UNP A0A381A265
A	3609	HIS	-	expression tag	UNP A0A381A265
A	3610	THR	-	expression tag	UNP A0A381A265
A	3611	SER	-	expression tag	UNP A0A381A265
B	3600	MSE	-	initiating methionine	UNP A0A381A265
B	3601	ALA	-	expression tag	UNP A0A381A265
B	3602	LYS	-	expression tag	UNP A0A381A265
B	3603	SER	-	expression tag	UNP A0A381A265
B	3604	HIS	-	expression tag	UNP A0A381A265
B	3605	HIS	-	expression tag	UNP A0A381A265
B	3606	HIS	-	expression tag	UNP A0A381A265
B	3607	HIS	-	expression tag	UNP A0A381A265
B	3608	HIS	-	expression tag	UNP A0A381A265
B	3609	HIS	-	expression tag	UNP A0A381A265
B	3610	THR	-	expression tag	UNP A0A381A265
B	3611	SER	-	expression tag	UNP A0A381A265

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O P 5 4 1	0	0
2	A	1	Total O P 5 4 1	0	0
2	B	1	Total O P 5 4 1	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Ca 1 1	0	0
3	B	1	Total Ca 1 1	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 5 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	61	Total	O	0	0
			61	61		
5	B	64	Total	O	0	0
			64	64		

3 Residue-property plots [i](#)

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. 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. 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.

Note EDS failed to run properly.

- Molecule 1: Filamentous hemagglutinin

Chain A:  81% 13%



- Molecule 1: Filamentous hemagglutinin

Chain B:  75% 11% 13%



4 Data and refinement statistics i

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	23.06Å 46.58Å 75.99Å 90.00° 91.69° 90.00°	Depositor
Resolution (Å)	39.71 – 1.65	Depositor
% Data completeness (in resolution range)	98.7 (39.71-1.65)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.19 (at 1.65Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.179 , 0.204	Depositor
Wilson B-factor (Å ²)	22.6	Xtrriage
Anisotropy	0.107	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.044 for h,-k,-l	Xtrriage
Total number of atoms	3358	wwPDB-VP
Average B, all atoms (Å ²)	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.70% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: PO4, CA, GOL

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	A	0.42	0/824	0.71	0/1106
1	B	0.48	0/794	0.69	0/1065
All	All	0.45	0/1618	0.70	0/2171

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	A	807	827	827	7	0
1	B	780	788	787	10	0
2	A	10	0	0	0	0
2	B	5	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	B	6	8	8	2	0
5	A	61	0	0	1	0
5	B	64	0	0	0	0
All	All	1735	1623	1622	16	0

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.

All (16) 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)
1:A:3693[A]:ARG:NH1	5:A:302:HOH:O	2.29	0.65
1:B:3634:ASN:HD21	1:B:3702:TYR:H	1.47	0.60
1:A:3618:VAL:HG22	1:A:3706[A]:TYR:CD2	2.37	0.60
1:B:3682:THR:OG1	1:B:3684:LEU:HD23	2.09	0.52
1:B:3693:ARG:HH12	1:B:3695:GLU:HG2	1.76	0.50
1:B:3634:ASN:OD1	4:B:202:GOL:H12	2.11	0.50
1:B:3651:LYS:O	1:B:3652:LEU:HD23	2.11	0.49
1:A:3703[C]:ARG:HH11	1:A:3703[C]:ARG:CG	2.26	0.49
1:A:3703[C]:ARG:HH11	1:A:3703[C]:ARG:HG3	1.78	0.48
1:A:3703[A]:ARG:NH1	1:B:3698:GLY:N	2.65	0.43
1:B:3634:ASN:OD1	4:B:202:GOL:C1	2.67	0.43
1:B:3709:ASN:OD1	1:B:3709:ASN:N	2.54	0.41
1:B:3690:PHE:CD2	1:B:3706[B]:TYR:HD2	2.39	0.41
1:B:3693:ARG:NH1	1:B:3695:GLU:HG2	2.35	0.41
1:A:3637:LYS:HB3	1:A:3637:LYS:NZ	2.37	0.40
1:A:3706[A]:TYR:CD1	1:A:3706[A]:TYR:C	2.94	0.40

There are no symmetry-related clashes.

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	A	98/112 (88%)	97 (99%)	1 (1%)	0	100	100
1	B	96/112 (86%)	95 (99%)	1 (1%)	0	100	100
All	All	194/224 (87%)	192 (99%)	2 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	A	90/100 (90%)	88 (98%)	2 (2%)	52	27
1	B	85/100 (85%)	82 (96%)	3 (4%)	36	11
All	All	175/200 (88%)	170 (97%)	5 (3%)	49	16

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3621	GLN
1	A	3637	LYS
1	B	3693	ARG
1	B	3705[A]	PHE
1	B	3705[B]	PHE

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

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 for Mogul analysis.

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	PO4	A	201	-	4,4,4	0.69	0	6,6,6	0.90	0
4	GOL	B	202	-	5,5,5	0.71	0	5,5,5	0.75	0
2	PO4	B	201	-	4,4,4	1.01	0	6,6,6	0.47	0
2	PO4	A	202	-	4,4,4	0.81	0	6,6,6	0.52	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	B	202	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	202	GOL	C1-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	202	GOL	2	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

6.4 Ligands

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers

EDS failed to run properly - this section is therefore empty.