Full wwPDB X-ray Structure Validation Report (i)

Mar 9, 2026 – 06:59 AM UTC

PDB ID	:	5OXD / pdb_00005oxd
Title	:	Complex of a C. perfringens O-GlcNAcase with a fragment hit
Authors	:	Rafie, K.; van Aalten, D.M.F.
Deposited on	:	2017-09-06
Resolution	:	2.60 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 (i) 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 (1)) were used in the production of this report:

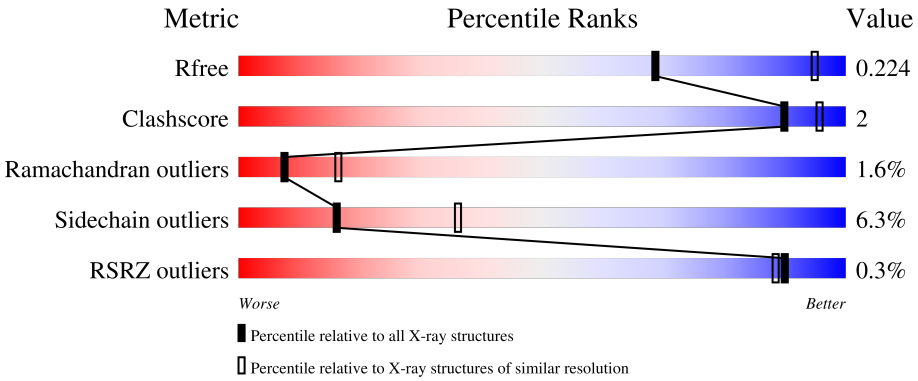
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.60 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))
R_{free}	180053	4008 (2.60-2.60)
Clashscore	190562	4347 (2.60-2.60)
Ramachandran outliers	187476	4277 (2.60-2.60)
Sidechain outliers	187428	4277 (2.60-2.60)
RSRZ outliers	180081	4008 (2.60-2.60)

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	A	590	86% 10%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4690 atoms, of which 0 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 O-GlcNAcase NagJ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	580	4586	2882	753	934	17	0	0	0

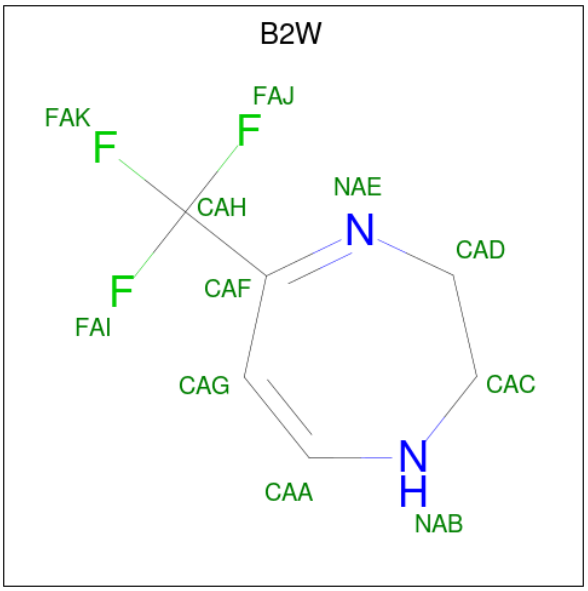
There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	29	GLY	-	expression tag	UNP Q0TR53
A	30	SER	-	expression tag	UNP Q0TR53
A	388	ASP	ASN	conflict	UNP Q0TR53

- Molecule 2 is CADMIUM ION (CCD ID: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cd		
2	A	22	22	22	0	0

- Molecule 3 is 5-(trifluoromethyl)-2,3-dihydro-1 {H}-1,4-diazepine (CCD ID: B2W) (formula: $C_6H_7F_3N_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C F N 11 6 3 2	0	0

- Molecule 4 is water.

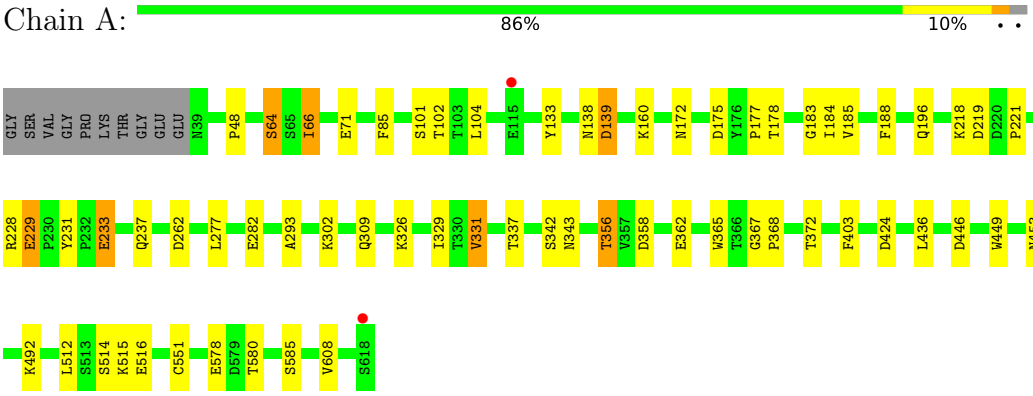
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	71	Total O 71 71	0	0

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 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: O-GlcNAcase NagJ

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	117.79Å 117.79Å 147.23Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.01 – 2.60 50.01 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.1 (50.01-2.60) 99.1 (50.01-2.60)	Depositor EDS
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.51 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.172 , 0.221 0.180 , 0.224	Depositor DCC
R_{free} test set	1880 reflections (5.28%)	wwPDB-VP
Wilson B-factor (Å^2)	58.1	Xtriage
Anisotropy	0.008	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 33.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.043 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4690	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% 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: B2W, CD

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	0.99	1/4681 (0.0%)	1.12	12/6356 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	178	THR	C-O	-5.32	1.18	1.24

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	225	GLU	N-CA-C	-7.17	101.04	110.43
1	A	367	GLY	CA-C-N	6.82	127.38	119.47
1	A	367	GLY	C-N-CA	6.82	127.38	119.47
1	A	226	LYS	N-CA-C	5.40	119.03	111.90
1	A	231	TYR	CA-C-N	5.36	126.54	119.84
1	A	231	TYR	C-N-CA	5.36	126.54	119.84
1	A	356	THR	N-CA-C	5.32	117.77	111.33
1	A	467	LEU	N-CA-C	5.18	119.31	112.89
1	A	331	VAL	CB-CA-C	5.14	116.23	110.52
1	A	358	ASP	CA-C-N	-5.10	114.79	120.45
1	A	358	ASP	C-N-CA	-5.10	114.79	120.45
1	A	188	PHE	N-CA-C	5.01	117.00	110.43

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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	4586	0	4371	16	0
2	A	22	0	0	0	1
3	A	11	0	0	0	0
4	A	71	0	0	0	0
All	All	4690	0	4371	16	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 2.

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:221:PRO:O	1:A:225:GLU:O	2.19	0.60
1:A:66:ILE:HG22	1:A:102:THR:HB	1.94	0.49
1:A:85:PHE:CD1	1:A:160:LYS:HG2	2.48	0.48
1:A:436:LEU:HD11	1:A:551:CYS:HB3	1.94	0.48
1:A:224:ARG:C	1:A:225:GLU:O	2.55	0.46
1:A:262:ASP:C	1:A:262:ASP:OD1	2.59	0.46
1:A:227:TRP:C	1:A:228:ARG:O	2.57	0.46
1:A:228:ARG:O	1:A:229:GLU:CB	2.64	0.45
1:A:48:PRO:HB3	1:A:446:ASP:HA	2.01	0.42
1:A:177:PRO:HB3	1:A:449:TRP:CE3	2.54	0.42
1:A:133:TYR:CE1	1:A:175:ASP:HB3	2.54	0.41
1:A:337:THR:HB	1:A:368:PRO:HA	2.01	0.41
1:A:293:ALA:HA	1:A:329:ILE:O	2.21	0.40
1:A:138:ASN:O	1:A:139:ASP:C	2.64	0.40
1:A:183:GLY:C	1:A:184:ILE:HD12	2.47	0.40
1:A:514:SER:O	1:A:516:GLU:N	2.55	0.40

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)
2:A:705:CD:CD	2:A:722:CD:CD[2_565]	1.96	0.24

5.3 Torsion angles

5.3.1 Protein backbone

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	578/590 (98%)	544 (94%)	25 (4%)	9 (2%)	7 16

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	515	LYS
1	A	64	SER
1	A	233	GLU
1	A	403	PHE
1	A	578	GLU
1	A	139	ASP
1	A	229	GLU
1	A	489	THR
1	A	608	VAL

5.3.2 Protein sidechains

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	495/502 (99%)	464 (94%)	31 (6%)	16 36

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	64	SER
1	A	66	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	71	GLU
1	A	101	SER
1	A	104	LEU
1	A	172	ASN
1	A	185	VAL
1	A	196	GLN
1	A	218	LYS
1	A	219	ASP
1	A	233	GLU
1	A	237	GLN
1	A	277	LEU
1	A	282	GLU
1	A	302	LYS
1	A	309	GLN
1	A	326	LYS
1	A	331	VAL
1	A	342	SER
1	A	343	ASN
1	A	356	THR
1	A	362	GLU
1	A	365	TRP
1	A	372	THR
1	A	424	ASP
1	A	453	ASN
1	A	489	THR
1	A	492	LYS
1	A	512	LEU
1	A	580	THR
1	A	585	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	138	ASN
1	A	201	GLN
1	A	413	HIS
1	A	433	HIS
1	A	442	HIS
1	A	453	ASN
1	A	538	ASN
1	A	592	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 23 ligands modelled in this entry, 22 are monoatomic - leaving 1 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	$\#\|Z\| > 2$
3	B2W	A	723	-	7,11,11	3.55	4 (57%)	7,15,15	2.29	1 (14%)

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
3	B2W	A	723	-	-	3/6/15/15	0/0/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	723	B2W	CAA-CAG	5.65	1.40	1.33
3	A	723	B2W	FAK-CAH	5.44	1.54	1.32
3	A	723	B2W	FAI-CAH	3.84	1.48	1.32

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	723	B2W	CAC-NAB	2.93	1.52	1.46

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	723	B2W	CAC-NAB-CAA	-5.26	118.54	125.33

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	723	B2W	NAE-CAF-CAH-FAI
3	A	723	B2W	NAE-CAF-CAH-FAJ
3	A	723	B2W	NAE-CAF-CAH-FAK

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

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

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, 95^{th} 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(Å^2)	Q<0.9
1	A	580/590 (98%)	-0.29	2 (0%)	90	88	43, 61, 91, 112	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	618	SER	2.9
1	A	115	GLU	2.8

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 oligosaccharides 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, 95^{th} 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(Å^2)	Q<0.9
3	B2W	A	723	11/11	0.74	0.25	97,117,141,145	0
2	CD	A	716	1/1	0.85	0.18	206,206,206,206	0
2	CD	A	713	1/1	0.86	0.16	186,186,186,186	0
2	CD	A	708	1/1	0.88	0.16	181,181,181,181	0
2	CD	A	717	1/1	0.92	0.17	186,186,186,186	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CD	A	720	1/1	0.92	0.13	187,187,187,187	0
2	CD	A	715	1/1	0.92	0.16	233,233,233,233	0
2	CD	A	721	1/1	0.94	0.11	190,190,190,190	0
2	CD	A	714	1/1	0.94	0.14	158,158,158,158	0
2	CD	A	719	1/1	0.96	0.21	185,185,185,185	0
2	CD	A	718	1/1	0.96	0.12	157,157,157,157	0
2	CD	A	710	1/1	0.98	0.09	187,187,187,187	0
2	CD	A	712	1/1	0.98	0.04	126,126,126,126	0
2	CD	A	703	1/1	0.98	0.11	134,134,134,134	0
2	CD	A	702	1/1	0.98	0.04	93,93,93,93	0
2	CD	A	707	1/1	0.99	0.07	104,104,104,104	0
2	CD	A	701	1/1	0.99	0.03	104,104,104,104	0
2	CD	A	705	1/1	1.00	0.10	206,206,206,206	0
2	CD	A	709	1/1	1.00	0.01	74,74,74,74	0
2	CD	A	706	1/1	1.00	0.02	63,63,63,63	0
2	CD	A	711	1/1	1.00	0.01	65,65,65,65	0
2	CD	A	722	1/1	1.00	0.03	62,62,62,62	0
2	CD	A	704	1/1	1.00	0.02	75,75,75,75	0

6.5 Other polymers (i)

There are no such residues in this entry.