



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 11:56 am BST

PDB ID : 3HD8
Title : Crystal structure of the Triticum aestivum xylanase inhibitor-IIA in complex with bacillus subtilis xylanase
Authors : Sansen, S.; Pollet, A.; Raedschelders, G.; Gebruers, K.; Rabijns, A.; Courtin, C.M.
Deposited on : 2009-05-07
Resolution : 2.39 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

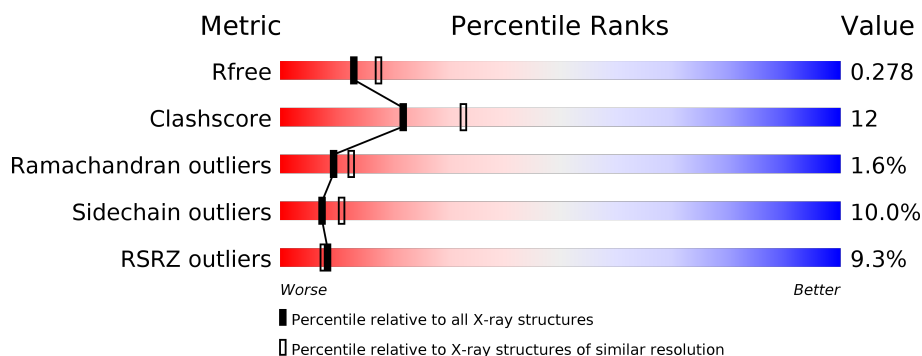
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.39 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}	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 on 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 $\leq 5\%$. 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	389	<div> <div>13%</div> <div> <div></div> <div>67%</div> <div>23%</div> <div>• • 5%</div> </div> </div>
1	C	389	<div> <div>11%</div> <div> <div></div> <div>74%</div> <div>17%</div> <div>• • 5%</div> </div> </div>
2	B	185	<div> <div>2%</div> <div> <div></div> <div>75%</div> <div>19%</div> <div>• • •</div> </div> </div>
2	D	185	<div> <div>3%</div> <div> <div></div> <div>72%</div> <div>23%</div> <div>• •</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8282 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 Xylanase inhibitor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	370	Total	C	N	O	S	0	0	0
			2682	1711	455	498	18			
1	C	368	Total	C	N	O	S	0	0	0
			2672	1705	453	496	18			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLU	LYS	SEE REMARK 999	UNP Q53IQ4
A	163	PRO	ALA	SEE REMARK 999	UNP Q53IQ4
A	314	MET	LEU	SEE REMARK 999	UNP Q53IQ4
C	1	GLU	LYS	SEE REMARK 999	UNP Q53IQ4
C	163	PRO	ALA	SEE REMARK 999	UNP Q53IQ4
C	314	MET	LEU	SEE REMARK 999	UNP Q53IQ4

- Molecule 2 is a protein called Endo-1,4-beta-xylanase A.

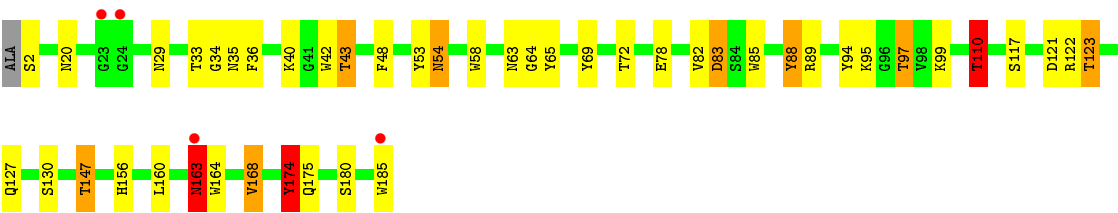
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	184	Total	C	N	O	S	0	0	0
			1443	909	248	284	2			
2	D	184	Total	C	N	O	S	0	0	0
			1443	909	248	284	2			

There are 2 discrepancies between the modelled and reference sequences:

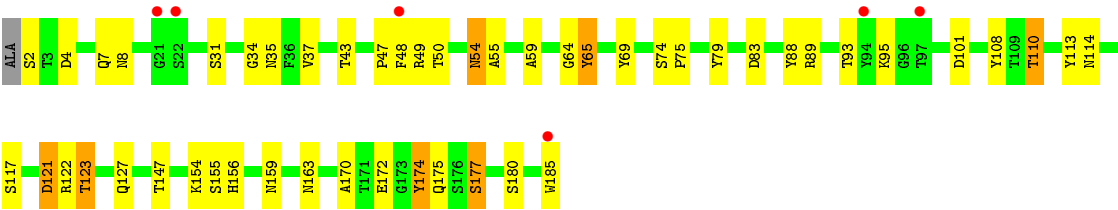
Chain	Residue	Modelled	Actual	Comment	Reference
B	147	THR	SER	SEE REMARK 999	UNP P18429
D	147	THR	SER	SEE REMARK 999	UNP P18429

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	15	Total 15	O 15	0	0
3	B	7	Total 7	O 7	0	0
3	C	13	Total 13	O 13	0	0
3	D	7	Total 7	O 7	0	0



● Molecule 2: Endo-1,4-beta-xylanase A



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	77.35Å 60.30Å 134.19Å 90.00° 101.49° 90.00°	Depositor
Resolution (Å)	38.52 – 2.39 38.52 – 2.39	Depositor EDS
% Data completeness (in resolution range)	98.8 (38.52-2.39) 98.7 (38.52-2.39)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.13 (at 2.39Å)	Xtriage
Refinement program	CNS, REFMAC 5.1.19	Depositor
R, R_{free}	0.223 , 0.275 0.236 , 0.278	Depositor DCC
R_{free} test set	2427 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	32.1	Xtriage
Anisotropy	0.267	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 37.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	8282	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.12% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

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	1.23	10/2746 (0.4%)	1.17	16/3747 (0.4%)
1	C	1.28	9/2736 (0.3%)	1.16	15/3733 (0.4%)
2	B	1.39	11/1491 (0.7%)	1.12	6/2043 (0.3%)
2	D	1.30	8/1491 (0.5%)	1.17	9/2043 (0.4%)
All	All	1.29	38/8464 (0.4%)	1.16	46/11566 (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	A	0	5
1	C	0	5
2	B	0	1
All	All	0	11

The worst 5 of 38 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	164	TRP	CB-CG	-8.75	1.34	1.50
1	C	195	ASP	CB-CG	8.64	1.69	1.51
1	A	385	PHE	CD1-CE1	7.37	1.53	1.39
2	B	163	ASN	CB-CG	-7.36	1.34	1.51
1	C	148	SER	CA-CB	7.33	1.64	1.52

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	213	ARG	CA-CB-CG	14.79	145.94	113.40
1	A	213	ARG	CA-CB-CG	11.86	139.49	113.40
1	C	307	ASP	CB-CG-OD2	10.07	127.36	118.30
1	A	384	ASN	N-CA-C	8.71	134.53	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	163	ASN	CB-CA-C	8.60	127.60	110.40

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1	GLU	Peptide
1	A	146	ALA	Peptide
1	A	225	SER	Peptide
1	A	384	ASN	Peptide
1	A	48	ALA	Peptide

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	2682	0	2679	71	0
1	C	2672	0	2669	69	0
2	B	1443	0	1314	33	0
2	D	1443	0	1314	31	0
3	A	15	0	0	2	0
3	B	7	0	0	2	0
3	C	13	0	0	3	0
3	D	7	0	0	0	0
All	All	8282	0	7976	191	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 12.

The worst 5 of 191 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)
2:D:117:SER:OG	2:D:123:THR:HG21	1.53	1.09
1:C:35:LEU:HD21	1:C:206:PRO:CB	1.81	1.09
1:A:385:PHE:HB3	3:A:404:HOH:O	1.54	1.07

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:35:LEU:HD21	1:A:206:PRO:CB	1.85	1.05
2:B:43:THR:HG22	3:B:188:HOH:O	1.59	1.01

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	362/389 (93%)	334 (92%)	19 (5%)	9 (2%)	5	6
1	C	360/389 (92%)	335 (93%)	17 (5%)	8 (2%)	6	7
2	B	182/185 (98%)	177 (97%)	5 (3%)	0	100	100
2	D	182/185 (98%)	177 (97%)	5 (3%)	0	100	100
All	All	1086/1148 (95%)	1023 (94%)	46 (4%)	17 (2%)	9	13

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	69	PRO
1	A	230	ALA
1	A	283	LEU
1	A	384	ASN
1	C	69	PRO

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	283/301 (94%)	252 (89%)	31 (11%)	6	8
1	C	283/301 (94%)	254 (90%)	29 (10%)	7	10
2	B	151/151 (100%)	138 (91%)	13 (9%)	10	16
2	D	151/151 (100%)	137 (91%)	14 (9%)	9	13
All	All	868/904 (96%)	781 (90%)	87 (10%)	7	11

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	123	THR
1	C	116	SER
2	D	110	THR
2	B	163	ASN
1	C	41	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	163	ASN
1	C	107	ASN
2	D	163	ASN
2	B	175	GLN
1	C	61	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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(Å ²)	Q<0.9
1	A	370/389 (95%)	0.79	50 (13%) 3 2	24, 43, 69, 85	0
1	C	368/389 (94%)	0.89	43 (11%) 4 4	24, 43, 68, 84	0
2	B	184/185 (99%)	0.20	4 (2%) 62 60	23, 32, 40, 52	0
2	D	184/185 (99%)	0.28	6 (3%) 46 45	24, 32, 40, 52	0
All	All	1106/1148 (96%)	0.64	103 (9%) 8 8	23, 37, 63, 85	0

The worst 5 of 103 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	339	ASP	10.7
1	A	340	ALA	9.6
1	C	338	VAL	8.2
1	C	342	ASP	7.3
1	C	227	ARG	7.1

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.