



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 05:33 AM BST

PDB ID : 5GPR  
Title : Crystal structure of chitinase-h from Ostrinia furnacalis  
Authors : Liu, T.; Zhou, Y.; Chen, L.; Yang, Q.  
Deposited on : 2016-08-04  
Resolution : 3.23 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](mailto: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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

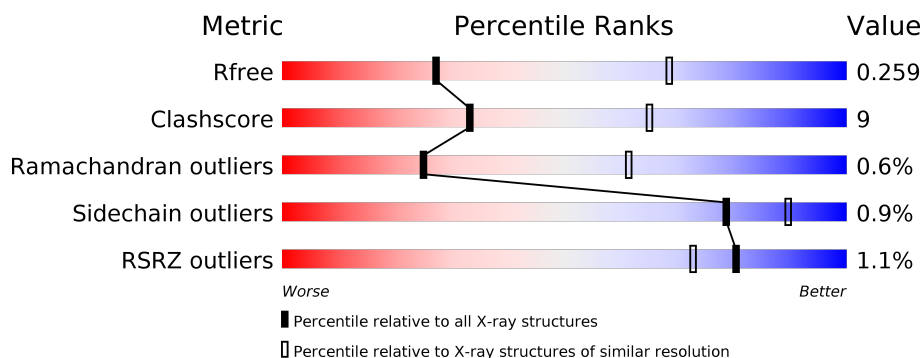
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.23 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	1619 (3.28-3.20)
Clashscore	141614	1755 (3.28-3.20)
Ramachandran outliers	138981	1728 (3.28-3.20)
Sidechain outliers	138945	1727 (3.28-3.20)
RSRZ outliers	127900	1567 (3.28-3.20)

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  $\geq 3$ , 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	553	<div> <div></div> <div>80%</div> <div>16%</div> <div>..</div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4226 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 Chitinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	536	Total	C	N	O	S	0	0	0
			4197	2685	705	792	15			

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is water.

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

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 ( $\text{RSRZ} > 2$ ). 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.

Chain A:

80% 16%

• •

Label	Color
Q295	Green
T296	Green
W297	Green
Y163	Green
P164	Green
D306	Green
F309	Green
G325	Green
V329	Green
L330	Green
E334	Green
L335	Green
R336	Green
T353	Green
S357	Green
D367	Green
Y368	Green
S369	Green
Y374	Green
M375	Green
D376	Green
H377	Green
D392	Green
Q397	Green
P403	Green
D404	Green
K429	Green
H450	Green
H451	Green
P452	Green
G463	Green
D467	Green
R473	Green
K490	Green
V491	Green
L505	Green
T506	Green
T507	Green
R512	Green
V521	Green
A155	Green
Y163	Green
P164	Green
R165	Green
K166	Green
F167	Green
P168	Green
R171	Green
L180	Green
G183	Green
F194	Green
I187	Green
G188	Green
G189	Green
I193	Green
I200	Green
E201	Green
R214	Green
S220	Green
I221	Green
H222	Green
D223	Green
P224	Green
W225	Green
Q229	Green
L236	Green
S236	Green
S237	Green
E240	Green
P241	Green
T242	Green
K243	Green
S264	Green
T269	Green
P273	Green
F277	Green
E280	Green
R283	Green
R284	Green
K291	Green
MET	Grey
GLY	Grey
ARG	Grey
LEU	Grey
ALA	Grey
ILE	Grey
VAL	Grey
VAL	Grey
ALA	Grey
THR	Grey
LEU	Grey
ALA	Grey
LEU	Grey
ALA	Grey
ALA	Grey
A118	Green
P19	Green
P20	Green
W27	Green
R30	Green
Q39	Green
H45	Green
V48	Green
D52	Green
W60	Green
R72	Green
V73	Green
L74	Green
L75	Green
H76	Green
D77	Green
K78	Green
E79	Green
F80	Green
T84	Green
G85	Green
G86	Green
R100	Green
E106	Green
S114	Green
I121	Green
I122	Green
E140	Green

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.91Å 114.42Å 122.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.97 – 3.23 45.43 – 3.12	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.97-3.23) 89.1 (45.43-3.12)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.99 (at 3.12Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.211 , 0.257 0.219 , 0.259	Depositor DCC
$R_{free}$ test set	1207 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	53.3	Xtriage
Anisotropy	0.219	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , -2.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	4226	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.38% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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: NAG

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	0.30	0/4313	0.44	0/5850

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	4197	0	4042	71	0
2	A	28	0	26	0	0
3	A	1	0	0	0	0
All	All	4226	0	4068	71	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 9.

All (71) 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:100:ARG:NE	1:A:122:ILE:HD11	1.72	1.03

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:100:ARG:CD	1:A:122:ILE:HD11	1.93	0.99
1:A:100:ARG:CD	1:A:122:ILE:CD1	2.44	0.96
1:A:100:ARG:HG3	1:A:122:ILE:CD1	2.05	0.87
1:A:546:ASN:HB3	1:A:551:ASN:HD22	1.40	0.85
1:A:100:ARG:HE	1:A:122:ILE:HD11	1.38	0.85
1:A:163:TYR:HB2	1:A:164:PRO:HD3	1.60	0.83
1:A:100:ARG:HD2	1:A:122:ILE:CD1	2.09	0.82
1:A:100:ARG:HG3	1:A:122:ILE:HD13	1.62	0.80
1:A:100:ARG:HD2	1:A:122:ILE:HD12	1.65	0.79
1:A:100:ARG:CG	1:A:122:ILE:CD1	2.63	0.77
1:A:397:GLN:H	1:A:507:THR:HG21	1.49	0.77
1:A:353:THR:HG22	1:A:377:HIS:HB2	1.69	0.74
1:A:167:PHE:CD1	1:A:171:ARG:NH2	2.56	0.73
1:A:100:ARG:CG	1:A:122:ILE:HD11	2.18	0.72
1:A:473:ARG:NH2	1:A:535:ASP:O	2.23	0.72
1:A:167:PHE:CE1	1:A:171:ARG:NH2	2.59	0.70
1:A:100:ARG:HG3	1:A:122:ILE:HD11	1.77	0.67
1:A:542:LEU:O	1:A:546:ASN:ND2	2.29	0.66
1:A:237:SER:HB2	1:A:240:GLU:HG2	1.80	0.63
1:A:546:ASN:H	1:A:546:ASN:HD22	1.49	0.61
1:A:84:THR:O	1:A:86:GLY:N	2.34	0.60
1:A:100:ARG:NE	1:A:122:ILE:CD1	2.54	0.59
1:A:140:GLU:OE2	1:A:473:ARG:NH1	2.36	0.58
1:A:45:ASN:OD1	1:A:214:ARG:NH1	2.35	0.58
1:A:163:TYR:HB2	1:A:164:PRO:CD	2.31	0.58
1:A:235:LEU:HB3	1:A:243:LYS:HG2	1.85	0.56
1:A:163:TYR:OH	1:A:223:ASP:OD2	2.20	0.56
1:A:166:LYS:O	1:A:166:LYS:HG3	2.05	0.55
1:A:463:GLY:HA3	1:A:467:ASP:HA	1.90	0.54
1:A:375:MET:O	1:A:429:LYS:NZ	2.41	0.54
1:A:27:TRP:HE3	1:A:30:ARG:HH22	1.57	0.52
1:A:30:ARG:HD3	1:A:121:ILE:HG22	1.91	0.52
1:A:264:SER:OG	1:A:306:ASP:OD2	2.26	0.52
1:A:277:PHE:O	1:A:283:ARG:NH1	2.42	0.52
1:A:52:ASP:N	1:A:52:ASP:OD1	2.33	0.50
1:A:546:ASN:N	1:A:546:ASN:HD22	2.09	0.50
1:A:72:ARG:NH1	1:A:106:GLU:OE2	2.42	0.50
1:A:189:GLY:HA3	1:A:193:ILE:HG21	1.95	0.48
1:A:353:THR:CG2	1:A:377:HIS:HB2	2.41	0.48
1:A:291:LYS:O	1:A:295:GLN:HG3	2.13	0.48
1:A:367:ASP:OD1	1:A:369:SER:OG	2.22	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:220:SER:OG	1:A:221:ILE:N	2.45	0.47
1:A:60:TRP:HZ2	1:A:85:GLY:HA3	1.79	0.47
1:A:309:PHE:CE1	1:A:357:SER:HB2	2.51	0.46
1:A:330:LEU:O	1:A:334:GLU:HG3	2.16	0.46
1:A:225:TRP:HA	1:A:229:GLN:HB2	1.98	0.45
1:A:546:ASN:ND2	1:A:546:ASN:N	2.65	0.45
1:A:546:ASN:HB3	1:A:551:ASN:HB2	1.98	0.45
1:A:452:PRO:HB2	1:A:505:LEU:HD22	1.99	0.44
1:A:74:LEU:HD23	1:A:79:GLU:HA	1.99	0.44
1:A:168:PRO:HD2	1:A:171:ARG:HD2	2.00	0.43
1:A:183:GLY:HA2	1:A:184:PHE:HA	1.58	0.43
1:A:325:GLY:O	1:A:329:VAL:HG23	2.19	0.43
1:A:521:VAL:HG13	1:A:526:LEU:HB2	2.01	0.43
1:A:77:ASP:O	1:A:78:LYS:HD3	2.19	0.42
1:A:376:ASP:C	1:A:377:HIS:ND1	2.73	0.42
1:A:201:GLU:H	1:A:201:GLU:CD	2.23	0.42
1:A:240:GLU:HA	1:A:241:PRO:HD3	1.92	0.42
1:A:336:ARG:HD2	1:A:374:TYR:O	2.20	0.42
1:A:490:LYS:HE2	1:A:512:ARG:HH21	1.85	0.42
1:A:167:PHE:HD1	1:A:171:ARG:NH2	2.13	0.42
1:A:100:ARG:CG	1:A:122:ILE:HD13	2.38	0.41
1:A:280:GLU:O	1:A:284:ARG:HG3	2.21	0.41
1:A:48:VAL:HG21	1:A:297:TRP:HZ2	1.85	0.41
1:A:155:ALA:O	1:A:180:LEU:HD12	2.20	0.41
1:A:75:LEU:HD12	1:A:80:PHE:CD2	2.56	0.41
1:A:200:ILE:HD11	1:A:269:THR:HG21	2.03	0.40
1:A:532:TRP:HA	1:A:533:GLU:HA	1.67	0.40
1:A:187:ILE:HB	1:A:273:PRO:HG2	2.04	0.40
1:A:20:PRO:O	1:A:114:SER:OG	2.25	0.40

There are no symmetry-related clashes.

## 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	534/553 (97%)	506 (95%)	25 (5%)	3 (1%)	25 61

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	163	TYR
1	A	85	GLY
1	A	39	GLN

### 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	436/446 (98%)	432 (99%)	4 (1%)	78 89

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	100	ARG
1	A	171	ARG
1	A	280	GLU
1	A	546	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	546	ASN

### 5.3.3 RNA ⓘ

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](#)

2 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	NAG	A	602	1	14,14,15	0.33	0	17,19,21	0.51	0
2	NAG	A	601	1	14,14,15	0.48	0	17,19,21	0.72	1 (5%)

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	NAG	A	602	1	-	1/6/23/26	0/1/1/1
2	NAG	A	601	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	601	NAG	C1-O5-C5	2.43	115.49	112.19

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	601	NAG	O5-C5-C6-O6
2	A	601	NAG	C4-C5-C6-O6
2	A	602	NAG	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	536/553 (96%)	0.06	6 (1%) 80 73	29, 34, 45, 54	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	27	TRP	3.9
1	A	392	ASP	2.6
1	A	404	ASP	2.5
1	A	450	ASN	2.3
1	A	491	VAL	2.2
1	A	403	PRO	2.1

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	NAG	A	602	14/15	0.81	0.53	36,37,38,38	0
2	NAG	A	601	14/15	0.90	0.30	34,36,38,38	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.