



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 09:08 am BST

PDB ID : 4WJW  
Title : Crystal Structure of the Chs5-Chs6 Exomer Cargo Adaptor Complex Bound to portion of Chs3  
Authors : Weiskoff, A.M.; Fromme, J.C.  
Deposited on : 2014-10-01  
Resolution : 2.59 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](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
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

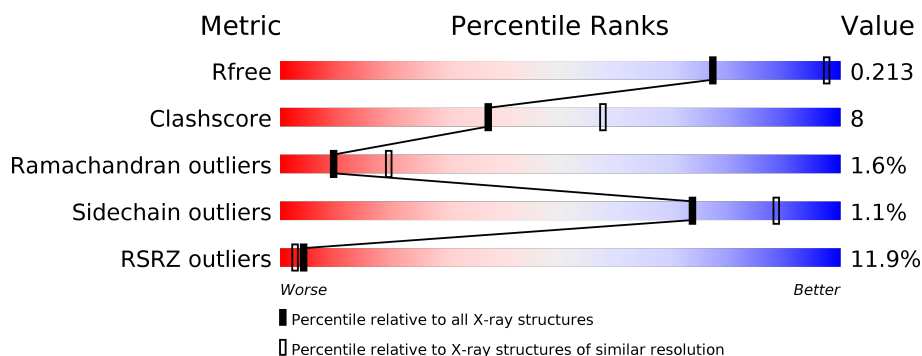
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.59 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (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 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	77	<div> <div>60%</div> <div> <div>68%</div> <div>19%</div> <div>•</div> <div>12%</div> </div> </div>
2	B	761	<div> <div>4%</div> <div> <div>74%</div> <div>13%</div> <div>•</div> <div>12%</div> </div> </div>
3	P	19	<div> <div>58%</div> <div> <div>42%</div> <div>21%</div> <div>11%</div> <div>26%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6098 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 Chitin biosynthesis protein CHS5.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	68	Total	C	N	O	0	0	0
			436	266	85	85			

- Molecule 2 is a protein called Chitin biosynthesis protein CHS6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	673	Total	C	N	O	S	0	0	0
			5446	3495	897	1015	39			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	747	GLY	-	expression tag	UNP P40955
B	748	THR	-	expression tag	UNP P40955
B	749	GLU	-	expression tag	UNP P40955
B	750	ASN	-	expression tag	UNP P40955
B	751	LEU	-	expression tag	UNP P40955
B	752	TYR	-	expression tag	UNP P40955
B	753	PHE	-	expression tag	UNP P40955
B	754	GLN	-	expression tag	UNP P40955
B	755	GLY	-	expression tag	UNP P40955
B	756	HIS	-	expression tag	UNP P40955
B	757	HIS	-	expression tag	UNP P40955
B	758	HIS	-	expression tag	UNP P40955
B	759	HIS	-	expression tag	UNP P40955
B	760	HIS	-	expression tag	UNP P40955
B	761	HIS	-	expression tag	UNP P40955

- Molecule 3 is a protein called CHITIN SYNTHASE 3.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	P	14	Total	C	N	O	0	0	0
			124	78	20	26			

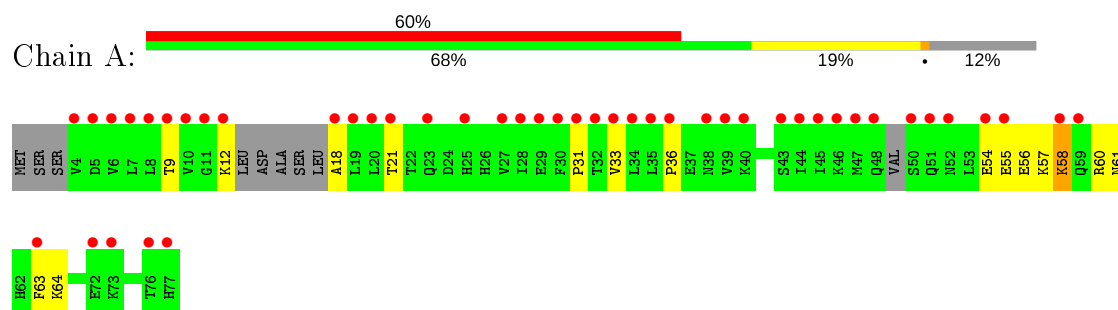
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	92	Total	O	0	0
			92	92		

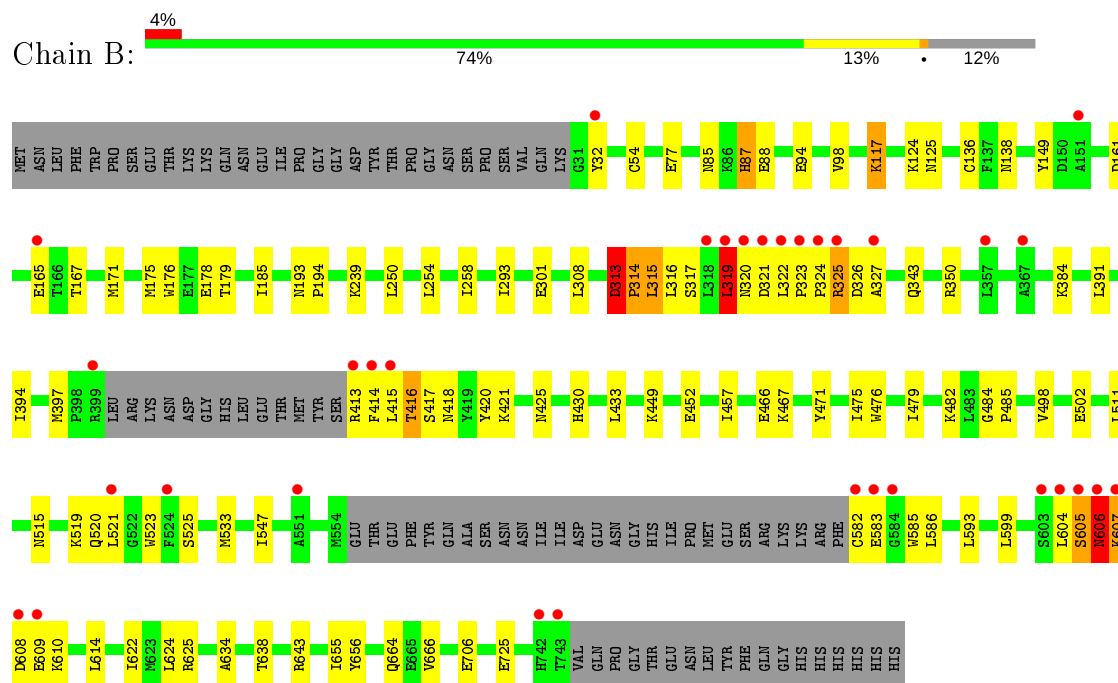
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

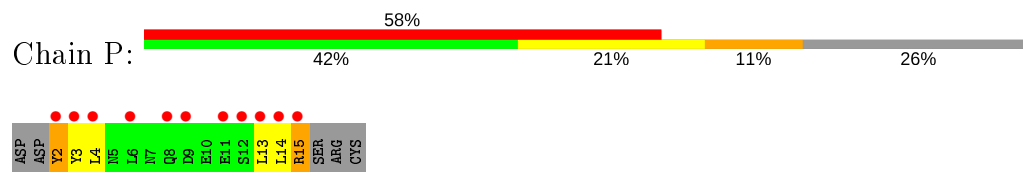
#### • Molecule 1: Chitin biosynthesis protein CHS5



#### • Molecule 2: Chitin biosynthesis protein CHS6



#### • Molecule 3: CHITIN SYNTHASE 3



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	218.61Å 218.61Å 137.85Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.65 – 2.59 49.64 – 2.59	Depositor EDS
% Data completeness (in resolution range)	98.9 (44.65-2.59) 96.3 (49.64-2.59)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.09 (at 2.58Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.8.4_1496)	Depositor
R, $R_{free}$	0.180 , 0.211 0.188 , 0.213	Depositor DCC
$R_{free}$ test set	2000 reflections (3.34%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	48.7	Xtriage
Anisotropy	0.108	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 63.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6098	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.76% 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

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	0.39	0/437	0.68	2/588 (0.3%)
2	B	0.31	0/5555	0.55	5/7516 (0.1%)
3	P	0.37	0/125	0.69	1/168 (0.6%)
All	All	0.32	0/6117	0.56	8/8272 (0.1%)

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
2	B	0	5

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	319	LEU	CA-CB-CG	14.56	148.80	115.30
2	B	319	LEU	CB-CA-C	-6.83	97.23	110.20
1	A	58	LYS	CD-CE-NZ	-6.82	96.02	111.70
2	B	319	LEU	N-CA-C	6.35	128.15	111.00
2	B	607	LYS	N-CA-C	-5.89	95.09	111.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	313	ASP	Peptide
2	B	319	LEU	Peptide
2	B	416	THR	Peptide
2	B	605	SER	Peptide
2	B	606	ASN	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	436	0	322	12	0
2	B	5446	0	5494	81	0
3	P	124	0	115	5	0
4	B	92	0	0	5	0
All	All	6098	0	5931	93	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 8.

The worst 5 of 93 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:B:316:LEU:O	2:B:319:LEU:HG	1.58	1.02
2:B:317:SER:HA	2:B:319:LEU:HD12	1.56	0.86
1:A:55:GLU:HA	1:A:58:LYS:HE2	1.55	0.85
2:B:413:ARG:NH2	4:B:874:HOH:O	2.10	0.85
2:B:414:PHE:O	2:B:416:THR:N	2.09	0.84

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	62/77 (80%)	50 (81%)	10 (16%)	2 (3%)	<b>4</b>	<b>6</b>
2	B	667/761 (88%)	632 (95%)	25 (4%)	10 (2%)	<b>10</b>	<b>21</b>

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	P	12/19 (63%)	12 (100%)	0	0	100	100
All	All	741/857 (86%)	694 (94%)	35 (5%)	12 (2%)	9	19

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	31	PRO
2	B	315	LEU
2	B	319	LEU
2	B	325	ARG
2	B	415	LEU

### 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	26/70 (37%)	26 (100%)	0	100	100
2	B	616/696 (88%)	610 (99%)	6 (1%)	76	90
3	P	14/19 (74%)	13 (93%)	1 (7%)	14	29
All	All	656/785 (84%)	649 (99%)	7 (1%)	73	88

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	582	CYS
3	P	2	TYR
2	B	606	ASN
2	B	117	LYS
2	B	609	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	61	ASN
2	B	163	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 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, 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	68/77 (88%)	3.43	46 (67%) 0 0	65, 128, 149, 160	0
2	B	673/761 (88%)	0.41	33 (4%) 29 23	33, 56, 114, 157	0
3	P	14/19 (73%)	3.84	11 (78%) 0 0	76, 108, 129, 153	0
All	All	755/857 (88%)	0.74	90 (11%) 4 3	33, 59, 134, 160	0

The worst 5 of 90 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	582	CYS	12.8
1	A	77	HIS	9.2
1	A	30	PHE	9.1
2	B	322	LEU	8.8
1	A	12	LYS	8.2

### 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.