



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:08 AM GMT

PDB ID : 2VZU  
Title : COMPLEX OF AMYCOLATOPSIS ORIENTALIS EXO-CHITOSANASE  
CSXA D469A WITH PNP-BETA-D-GLUCOSAMINE  
Authors : Lammerts Van Bueren, A.; Ghinet, M.G.; Gregg, K.; Fleury, A.; Brzezinski,  
R.; Boraston, A.B.  
Deposited on : 2008-08-05  
Resolution : 2.10 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

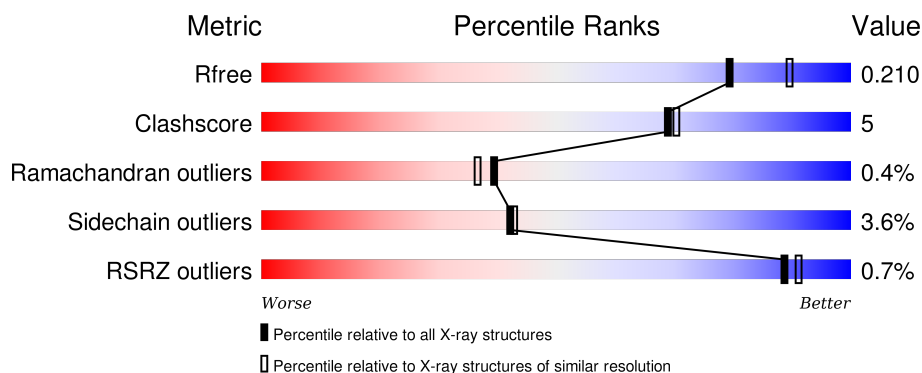
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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))
$R_{free}$	91344	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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. 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	1032	<div> <div></div> <div>74%</div> <div>7%</div> <div>•</div> <div>18%</div> </div>
1	B	1032	<div> <div>%</div> <div>73%</div> <div>8%</div> <div>•</div> <div>18%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	ACT	B	1903	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13943 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 EXO-BETA-D-GLUCOSAMINIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	851	Total	C	N	O	S	4	0	1
			6506	4088	1126	1275	17			
1	B	851	Total	C	N	O	S	0	0	1
			6506	4088	1126	1275	17			

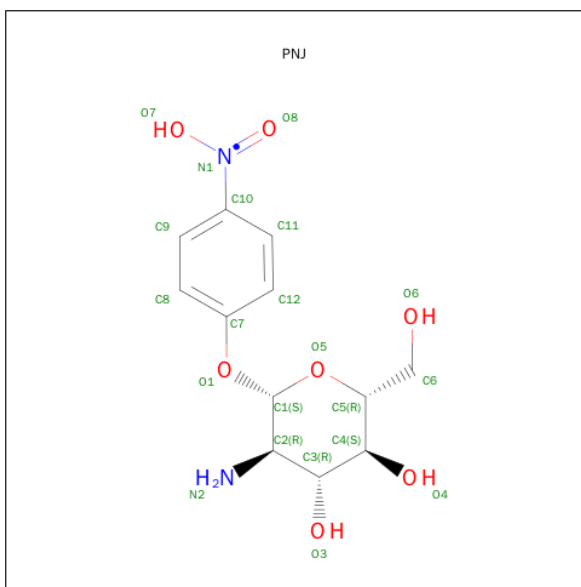
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	469	ALA	ASP	ENGINEERED MUTATION	UNP Q56F26
A	750	ASN	TRP	CONFLICT	UNP Q56F26
B	469	ALA	ASP	ENGINEERED MUTATION	UNP Q56F26
B	750	ASN	TRP	CONFLICT	UNP Q56F26

- Molecule 2 is CADMIUM ION (three-letter code: CD) (formula: Cd).

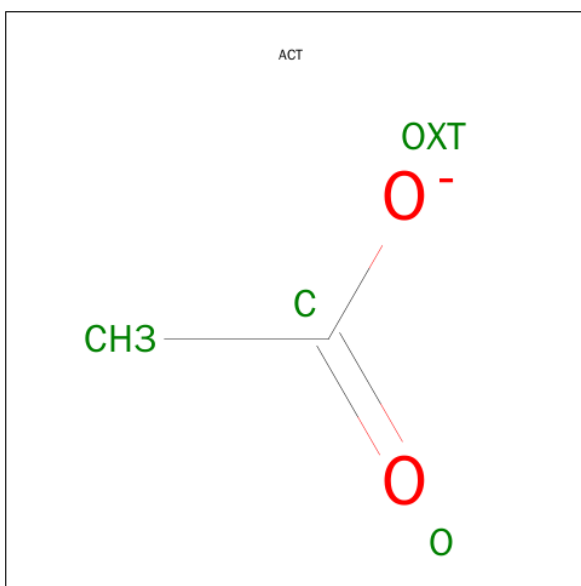
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	4	Total	Cd	0	0
			4	4		
2	A	2	Total	Cd	0	0
			2	2		

- Molecule 3 is PNP-BETA-D-GLUCOSAMINE (three-letter code: PNJ) (formula: C<sub>12</sub>H<sub>17</sub>N<sub>2</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			21	12	2	7		
3	B	1	Total	C	N	O	0	0
			21	12	2	7		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	451	Total 451	O 451	0	0
5	B	428	Total 428	O 428	0	0



[illegible]

SER	ALA	GLY	THR	TYR	ASP	VAL	VAL	ARG	TYR	ALA	ASN	GLY	THR	THR	SER	ARG	PRO	LEU	ASP	PHE	SER	VAL	GLY	GLY	SER	SER	ILE	SER	ALA	ALA	PHE	GLY	SER	THR	THR	THR	TRP	PRO	ALA	TRP	THR	THR	LYS	VAL	VAL	VAL	ARG	THR	THR	LEU	ALA	ALA	GLY	VAL	ASN	LYS
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

ILE  
LYS  
ALA  
VAL  
ALA  
THR  
THR  
ALA  
ASN  
GLY  
GLY  
PRO  
ASN  
VAL  
ASP  
LYS  
ILE  
THR  
LEU

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.60Å 121.96Å 91.81Å 90.00° 90.42° 90.00°	Depositor
Resolution (Å)	40.00 – 2.10 19.99 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (40.00-2.10) 99.8 (19.99-2.10)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.30 (at 2.09Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.172 , 0.207 0.176 , 0.210	Depositor DCC
$R_{free}$ test set	5546 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	13.2	Xtriage
Anisotropy	0.266	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 44.7	EDS
Estimated twinning fraction	0.027 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 110721 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13943	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: CD, PNJ, ACT

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	0.83	1/6669 (0.0%)	0.88	22/9100 (0.2%)
1	B	0.85	0/6669	0.92	27/9100 (0.3%)
All	All	0.84	1/13338 (0.0%)	0.90	49/18200 (0.3%)

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	3
1	B	0	3
All	All	0	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	430	GLU	CB-CG	5.41	1.62	1.52

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	335	ARG	NE-CZ-NH2	-17.55	111.52	120.30
1	B	454	ARG	NE-CZ-NH2	-17.24	111.68	120.30
1	B	454	ARG	NE-CZ-NH1	16.42	128.51	120.30
1	A	577	ARG	NE-CZ-NH1	12.97	126.78	120.30
1	A	335	ARG	NE-CZ-NH2	-12.68	113.96	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	468	SER	Peptide
1	A	596	ARG	Peptide
1	A	98	GLY	Peptide
1	B	468	SER	Peptide
1	B	98	GLY	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	6506	0	6287	57	0
1	B	6506	0	6287	62	0
2	A	2	0	0	0	0
2	B	4	0	0	0	0
3	A	21	0	17	1	0
3	B	21	0	17	1	0
4	B	4	0	3	0	0
5	A	451	0	0	18	0
5	B	428	0	0	15	0
All	All	13943	0	12611	119	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.

The worst 5 of 119 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:855:ALA:HB1	5:A:2430:HOH:O	1.19	1.24
1:B:855:ALA:HB1	5:B:2406:HOH:O	1.46	1.14
1:A:821:VAL:HB	5:A:2413:HOH:O	1.46	1.12
1:A:201:TRP:HE1	1:A:212:ASN:HD21	1.10	0.97
1:B:201:TRP:HE1	1:B:212:ASN:HD21	1.05	0.93

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	849/1032 (82%)	833 (98%)	13 (2%)	3 (0%)	39	37
1	B	849/1032 (82%)	825 (97%)	20 (2%)	4 (0%)	34	30
All	All	1698/2064 (82%)	1658 (98%)	33 (2%)	7 (0%)	39	37

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	469	ALA
1	B	597	TYR
1	A	202	ILE
1	A	541	GLU
1	B	202	ILE

### 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	697/832 (84%)	675 (97%)	22 (3%)	46	48
1	B	697/832 (84%)	669 (96%)	28 (4%)	38	38
All	All	1394/1664 (84%)	1344 (96%)	50 (4%)	42	43

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	99	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	211	GLN
1	B	808	ASN
1	B	131	ASP
1	B	144	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	116	GLN
1	B	176	GLN
1	B	682	HIS
1	A	796	ASN
1	A	808	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](#)

Of 9 ligands modelled in this entry, 6 are monoatomic - leaving 3 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	PNJ	A	1901	-	21,22,22	2.62	2 (9%)	28,31,31	1.38	4 (14%)
4	ACT	B	1903	-	1,3,3	4.76	1 (100%)	0,3,3	0.00	-
3	PNJ	B	1904	-	21,22,22	2.53	4 (19%)	28,31,31	1.26	4 (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	PNJ	A	1901	-	-	0/8/30/30	0/2/2/2
4	ACT	B	1903	-	-	0/0/0/0	0/0/0/0
3	PNJ	B	1904	-	-	0/8/30/30	0/2/2/2

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1904	PNJ	C10-N1	-5.48	1.34	1.45
3	A	1901	PNJ	C10-N1	-4.92	1.35	1.45
3	B	1904	PNJ	O5-C1	2.01	1.47	1.41
3	B	1904	PNJ	C1-C2	2.69	1.57	1.52
4	B	1903	ACT	CH3-C	4.76	1.55	1.48

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1901	PNJ	O1-C1-C2	-3.35	102.75	106.95
3	A	1901	PNJ	O5-C5-C4	-3.19	103.69	109.68
3	B	1904	PNJ	O1-C1-C2	-2.09	104.33	106.95
3	B	1904	PNJ	C9-C10-N1	-2.05	117.91	119.52
3	B	1904	PNJ	C11-C10-N1	2.26	121.30	119.52

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1901	PNJ	1	0
3	B	1904	PNJ	1	0

## 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	851/1032 (82%)	-0.53	3 (0%) 93 94	4, 11, 22, 32	1 (0%)
1	B	851/1032 (82%)	-0.50	9 (1%) 82 86	4, 11, 22, 34	0
All	All	1702/2064 (82%)	-0.51	12 (0%) 89 91	4, 11, 22, 34	1 (0%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	130	ASP	2.9
1	B	50	ASN	2.7
1	B	828	GLY	2.7
1	A	49	GLY	2.6
1	B	430	GLU	2.5

### 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ACT	B	1903	4/4	0.56	0.25	15.85	17,19,20,20	0
3	PNJ	B	1904	21/21	0.96	0.09	0.70	5,9,23,25	0
3	PNJ	A	1901	21/21	0.96	0.09	0.64	5,10,26,28	0
2	CD	B	1902	1/1	0.93	0.12	-0.53	97,97,97,97	0
2	CD	A	1900	1/1	0.85	0.09	-0.84	117,117,117,117	0
2	CD	B	1899	1/1	1.00	0.03	-1.73	10,10,10,10	0
2	CD	B	1900	1/1	1.00	0.03	-2.33	10,10,10,10	0
2	CD	B	1901	1/1	0.82	0.23	-	203,203,203,203	0
2	CD	A	1899	1/1	0.87	0.27	-	99,99,99,99	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.