



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 16, 2023 – 08:42 pm GMT

PDB ID : 2X05
Title : Inhibition of the exo-beta-D-glucosaminidase CsxA by a glucosamine- configured castanospermine and an amino-australine analogue
Authors : Pluvinae, B.; Ghinet, M.G.; Brzezinski, R.; Boraston, A.B.; Stubbs, K.A.
Deposited on : 2009-12-05
Resolution : 2.30 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
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) : 2.36

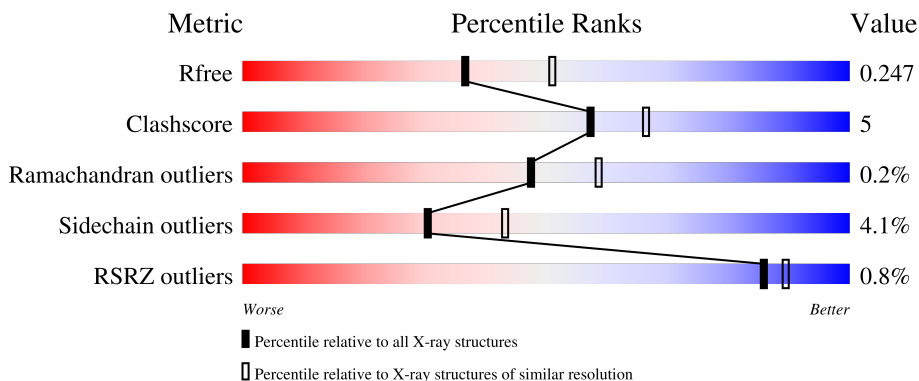
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 $\leq 5\%$. 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	
1	B	1032	

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
3	CD	A	1902	-	-	X	-

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 14302 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.

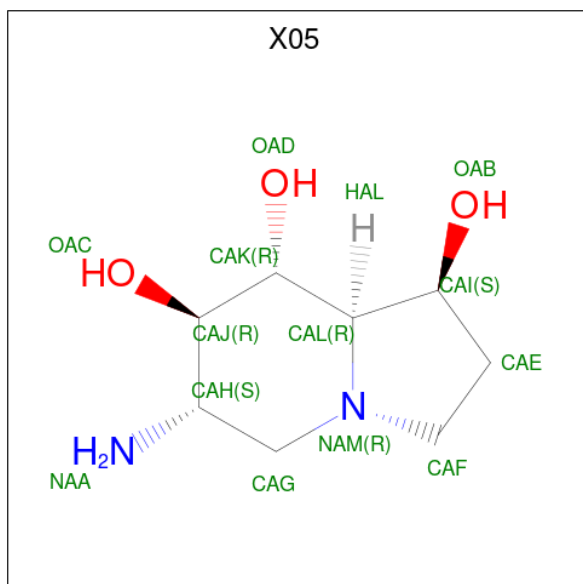
- Molecule 1 is a protein called EXO-BETA-D-GLUCOSAMINIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	852	6520	4095	1127	1281	17	4	1	1
1	B	852	6532	4104	1128	1282	18	0	3	1

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	750	ASN	TRP	conflict	UNP Q56F26
B	750	ASN	TRP	conflict	UNP Q56F26

- Molecule 2 is AMINO-CASTANOSPERMINE (three-letter code: X05) (formula: C₈H₁₆N₂O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	13	8	2	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	N	O	0	0
			13	8	2	3		

- Molecule 3 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	9	Total	Cd	0	0
			9	9		
3	B	10	Total	Cd	0	0
			10	10		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	604	Total	O	0	0
			604	604		
4	B	601	Total	O	0	0
			601	601		

M634	K839	SER	VAL
L639	K846	VAL	THR
W642	P847	TRP	LEU
M643	V848	ALA	ALA
W648	L877	THR	GLY
H652	S880	VAL	VAL
F656	W889	PRO	ASN
M660	T895	ILE	LYS
D661	D899	ALA	ALA
L676	GLY	THR	THR
S681	SER	ALA	THR
H682	GLY	VAL	ALA
T701	GLY	VAL	ASN
A702	GLY	VAL	ASN
T703	ASP	ARG	GLY
T704	PRO	TYR	GLY
L709	PRO	ALA	PRO
K714	ALA	ARG	LEU
V724	GLU	LEU	ASP
T744	ASP	PHE	PHE
T745	THR	SER	SER
A748	ILE	VAL	VAL
Y765	ASN	ASN	ASN
W766	GLN	GLY	GLY
L767	ILE	GLY	ILE
S768	VAL	SER	SER
T769	VAL	ILE	ALA
K770	GLN	ALA	GLY
W781	ALA	ALA	VAL
A790	PHE	ALA	ALA
G804	GLY	PHE	THR
M808	TYR	GLY	TRP
S809	THR	THR	PRO
L823	GLY	SER	ALA
K824	THR	THR	TRP
R830	THR	THR	THR
F834	LYS	THR	LYS
	THR	THR	THR
	VAL	VAL	VAL
	ALA	ALA	ALA
	GLY	GLY	GLY
	SER	SER	SER
	ARG	ARG	ARG

4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	86.12Å 121.80Å 91.56Å 90.00° 90.51° 90.00°	Depositor
Resolution (Å)	43.79 – 2.30 43.79 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.0 (43.79-2.30) 99.0 (43.79-2.30)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.37 (at 2.29Å)	Xtrriage
Refinement program	REFMAC 5.5.0072	Depositor
R, R_{free}	0.184 , 0.248 0.185 , 0.247	Depositor DCC
R_{free} test set	4151 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	20.3	Xtrriage
Anisotropy	0.201	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 34.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.025 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	14302	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: X05, 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	0.57	1/6686 (0.0%)	0.68	3/9123 (0.0%)
1	B	0.56	0/6701	0.68	4/9143 (0.0%)
All	All	0.57	1/13387 (0.0%)	0.68	7/18266 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	430	GLU	CB-CG	13.47	1.77	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	430	GLU	CA-CB-CG	-7.71	96.45	113.40
1	B	230	ARG	NE-CZ-NH2	-6.84	116.88	120.30
1	B	144	LEU	CA-CB-CG	6.45	130.13	115.30
1	A	335	ARG	NE-CZ-NH2	-5.62	117.49	120.30
1	A	144	LEU	CA-CB-CG	5.15	127.16	115.30

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	6520	0	6295	65	0
1	B	6532	0	6312	60	3
2	A	13	0	16	4	0
2	B	13	0	16	5	0
3	A	9	0	0	0	3
3	B	10	0	0	0	0
4	A	604	0	0	7	0
4	B	601	0	0	9	0
All	All	14302	0	12639	125	3

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 125 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:B:201:TRP:HE1	1:B:212:ASN:HD21	1.03	1.01
1:A:246:ASP:HB3	1:A:292:LEU:HD11	1.53	0.91
1:A:608:ARG:HG3	1:A:889:TRP:CZ3	2.07	0.90
1:A:201:TRP:HE1	1:A:212:ASN:ND2	1.69	0.90
1:A:541:GLU:OE1	2:A:1900:X05:HAG2	1.76	0.84

All (3) 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)
1:B:246:ASP:OD1	3:A:1902:CD:CD[2_557]	1.99	0.21
1:B:244:HIS:NE2	3:A:1902:CD:CD[2_557]	2.09	0.11
1:B:296:ASP:OD2	3:A:1905:CD:CD[2_557]	2.12	0.08

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	851/1032 (82%)	824 (97%)	26 (3%)	1 (0%)	51 64
1	B	853/1032 (83%)	817 (96%)	33 (4%)	3 (0%)	34 42
All	All	1704/2064 (83%)	1641 (96%)	59 (4%)	4 (0%)	47 58

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	202	ILE
1	B	202	ILE
1	B	541	GLU
1	B	205	ALA

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	700/834 (84%)	667 (95%)	33 (5%)	26 37
1	B	702/834 (84%)	678 (97%)	24 (3%)	37 51
All	All	1402/1668 (84%)	1345 (96%)	57 (4%)	30 43

5 of 57 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	830	ARG
1	B	877	LEU
1	B	143	VAL
1	B	846	LYS
1	B	701	THR

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

Mol	Chain	Res	Type
1	B	176	GLN
1	B	299	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	796	ASN
1	B	212	ASN
1	B	348	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 21 ligands modelled in this entry, 19 are monoatomic - leaving 2 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
2	X05	B	1900	-	14,14,14	0.94	0	14,21,21	0.89	0
2	X05	A	1900	-	14,14,14	1.18	1 (7%)	14,21,21	1.06	1 (7%)

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	X05	B	1900	-	-	-	0/2/2/2

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	X05	A	1900	-	-	-	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1900	X05	CAI-CAL	-2.27	1.52	1.54

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1900	X05	CAE-CAI-CAL	-2.28	99.95	102.80

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1900	X05	5	0
2	A	1900	X05	4	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, 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	852/1032 (82%)	-0.19	9 (1%) 80 85	8, 18, 28, 47	1 (0%)
1	B	852/1032 (82%)	-0.23	5 (0%) 89 92	8, 18, 27, 42	0
All	All	1704/2064 (82%)	-0.21	14 (0%) 86 89	8, 18, 28, 47	1 (0%)

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	48	ALA	3.5
1	A	48	ALA	3.1
1	A	181	VAL	3.0
1	A	49	GLY	2.9
1	A	128	ASN	2.9

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, 95th 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(\AA^2)	Q<0.9
3	CD	A	1907	1/1	0.76	0.18	118,118,118,118	0
3	CD	B	1910	1/1	0.81	0.17	154,154,154,154	0
3	CD	A	1908	1/1	0.86	0.26	169,169,169,169	0
3	CD	B	1908	1/1	0.88	0.15	146,146,146,146	0
2	X05	A	1900	13/13	0.89	0.17	14,17,20,21	0
2	X05	B	1900	13/13	0.91	0.16	15,16,18,19	0
3	CD	A	1904	1/1	0.93	0.12	76,76,76,76	0
3	CD	B	1907	1/1	0.94	0.18	139,139,139,139	0
3	CD	B	1903	1/1	0.94	0.17	93,93,93,93	0
3	CD	B	1909	1/1	0.94	0.16	112,112,112,112	0
3	CD	B	1906	1/1	0.94	0.21	132,132,132,132	0
3	CD	B	1904	1/1	0.95	0.24	97,97,97,97	0
3	CD	B	1901	1/1	0.95	0.08	69,69,69,69	0
3	CD	A	1906	1/1	0.96	0.22	116,116,116,116	0
3	CD	B	1902	1/1	0.97	0.12	60,60,60,60	0
3	CD	B	1905	1/1	0.98	0.26	93,93,93,93	0
3	CD	A	1903	1/1	0.98	0.09	60,60,60,60	0
3	CD	A	1902	1/1	0.99	0.02	11,11,11,11	0
3	CD	A	1901	1/1	0.99	0.03	29,29,29,29	0
3	CD	A	1905	1/1	1.00	0.02	12,12,12,12	0
3	CD	A	1909	1/1	1.00	0.03	14,14,14,14	0

6.5 Other polymers [i](#)

There are no such residues in this entry.