



wwPDB X-ray Structure Validation Summary Report i

Feb 26, 2014 – 07:42 PM GMT

PDB ID : 4E55
Title : Crystal Structure of spacer removed cephalosporin acylase mutant
Authors : Yin, J.; Zhang, Z.; Wu, G.; Huang, X.
Deposited on : 2012-03-14
Resolution : 2.30 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

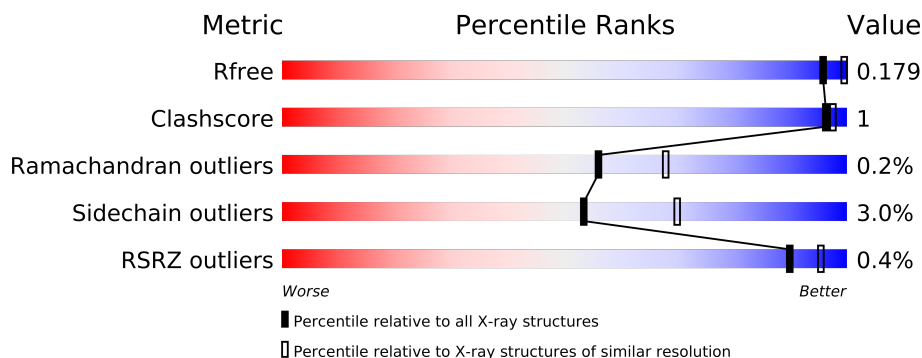
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance



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}	66092	2929 (2.30-2.30)
Clashscore	79885	3679 (2.30-2.30)
Ramachandran outliers	78287	3642 (2.30-2.30)
Sidechain outliers	78261	3641 (2.30-2.30)
RSRZ outliers	66119	2930 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	694	
1	B	694	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11411 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Cephalosporin acylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	665	Total	C	N	O	S	0	0	0
			5247	3323	925	987	12			
1	B	665	Total	C	N	O	S	0	0	0
			5247	3323	925	987	12			

There are 54 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	EXPRESSION TAG	UNP O86089
A	2	GLY	-	EXPRESSION TAG	UNP O86089
A	3	ILE	-	EXPRESSION TAG	UNP O86089
A	127	GLU	ASP	SEE REMARK 999	UNP O86089
A	?	-	GLY	DELETION	UNP O86089
A	?	-	ASP	DELETION	UNP O86089
A	?	-	PRO	DELETION	UNP O86089
A	?	-	PRO	DELETION	UNP O86089
A	?	-	ASP	DELETION	UNP O86089
A	?	-	LEU	DELETION	UNP O86089
A	?	-	ALA	DELETION	UNP O86089
A	?	-	ASP	DELETION	UNP O86089
A	?	-	GLN	DELETION	UNP O86089
A	?	-	GLY	DELETION	UNP O86089
A	162	ALA	SER	ENGINEERED MUTATION	UNP O86089
A	683	LEU	-	EXPRESSION TAG	UNP O86089
A	684	ALA	-	EXPRESSION TAG	UNP O86089
A	685	ALA	-	EXPRESSION TAG	UNP O86089
A	686	ALA	-	EXPRESSION TAG	UNP O86089
A	687	LEU	-	EXPRESSION TAG	UNP O86089
A	688	GLU	-	EXPRESSION TAG	UNP O86089
A	689	HIS	-	EXPRESSION TAG	UNP O86089
A	690	HIS	-	EXPRESSION TAG	UNP O86089
A	691	HIS	-	EXPRESSION TAG	UNP O86089
A	692	HIS	-	EXPRESSION TAG	UNP O86089

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	693	HIS	-	EXPRESSION TAG	UNP O86089
A	694	HIS	-	EXPRESSION TAG	UNP O86089
B	1	MET	-	EXPRESSION TAG	UNP O86089
B	2	GLY	-	EXPRESSION TAG	UNP O86089
B	3	ILE	-	EXPRESSION TAG	UNP O86089
B	127	GLU	ASP	SEE REMARK 999	UNP O86089
B	?	-	GLY	DELETION	UNP O86089
B	?	-	ASP	DELETION	UNP O86089
B	?	-	PRO	DELETION	UNP O86089
B	?	-	PRO	DELETION	UNP O86089
B	?	-	ASP	DELETION	UNP O86089
B	?	-	LEU	DELETION	UNP O86089
B	?	-	ALA	DELETION	UNP O86089
B	?	-	ASP	DELETION	UNP O86089
B	?	-	GLN	DELETION	UNP O86089
B	?	-	GLY	DELETION	UNP O86089
B	162	ALA	SER	ENGINEERED MUTATION	UNP O86089
B	683	LEU	-	EXPRESSION TAG	UNP O86089
B	684	ALA	-	EXPRESSION TAG	UNP O86089
B	685	ALA	-	EXPRESSION TAG	UNP O86089
B	686	ALA	-	EXPRESSION TAG	UNP O86089
B	687	LEU	-	EXPRESSION TAG	UNP O86089
B	688	GLU	-	EXPRESSION TAG	UNP O86089
B	689	HIS	-	EXPRESSION TAG	UNP O86089
B	690	HIS	-	EXPRESSION TAG	UNP O86089
B	691	HIS	-	EXPRESSION TAG	UNP O86089
B	692	HIS	-	EXPRESSION TAG	UNP O86089
B	693	HIS	-	EXPRESSION TAG	UNP O86089
B	694	HIS	-	EXPRESSION TAG	UNP O86089

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Ca 1 1	0	0
2	A	1	Total Ca 1 1	0	0

- Molecule 3 is water.

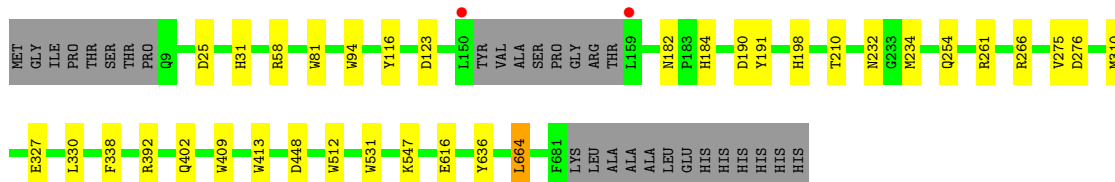
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	472	Total 472	O 472	0	0
3	B	443	Total 443	O 443	0	0

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 errors 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.

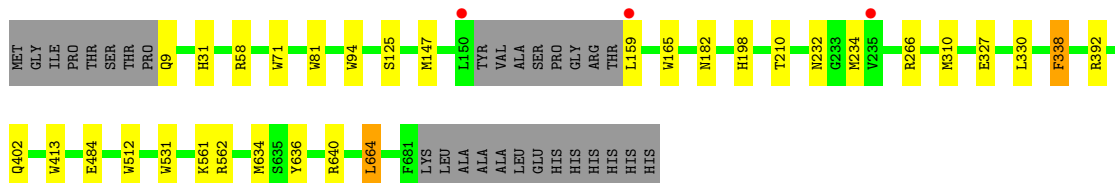
- Molecule 1: Cephalosporin acylase

Chain A: 



- Molecule 1: Cephalosporin acylase

Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α , β , γ	74.24Å 74.24Å 383.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.30 48.46 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.3 (50.00-2.30) 99.4 (48.46-2.30)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.85 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.171 , 0.199 0.167 , 0.179	Depositor DCC
R_{free} test set	4551 reflections (5.27%)	DCC
Wilson B-factor (Å ²)	31.2	Xtriage
Anisotropy	0.019	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 16.0	EDS
Estimated twinning fraction	0.467 for h,-k,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 90915 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	11411	wwPDB-VP
Average B, all atoms (Å ²)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.13% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	0.56	6/5398 (0.1%)	0.60	1/7373 (0.0%)
1	B	0.56	7/5398 (0.1%)	0.61	2/7373 (0.0%)
All	All	0.56	13/10796 (0.1%)	0.61	3/14746 (0.0%)

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	531	TRP	CD2-CE2	5.71	1.48	1.41
1	A	531	TRP	CD2-CE2	5.53	1.48	1.41
1	A	81	TRP	CD2-CE2	5.46	1.47	1.41
1	B	81	TRP	CD2-CE2	5.45	1.47	1.41
1	B	512	TRP	CD2-CE2	5.42	1.47	1.41

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	664	LEU	CA-CB-CG	5.67	128.33	115.30
1	B	634	MET	CG-SD-CE	-5.54	91.33	100.20
1	B	664	LEU	CA-CB-CG	5.37	127.64	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the

chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5247	0	0	6	0
1	B	5247	0	0	5	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	472	0	0	0	0
3	B	443	0	0	1	0
All	All	11411	0	0	11	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 1.

The worst 5 of 11 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:327:GLU:OE2	1:B:392:ARG:NH2	2.40	0.54
1:B:31:HIS:ND1	1:B:198:HIS:CD2	2.79	0.51
1:B:562:ARG:NH1	3:B:1085:HOH:O	2.47	0.48
1:B:147:MET:CE	1:B:338:PHE:CE1	2.98	0.47
1:A:327:GLU:OE2	1:A:392:ARG:NH2	2.50	0.45

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	661/694 (95%)	644 (97%)	16 (2%)	1 (0%)	56	68
1	B	661/694 (95%)	645 (98%)	15 (2%)	1 (0%)	56	68
All	All	1322/1388 (95%)	1289 (98%)	31 (2%)	2 (0%)	56	68

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	338	PHE
1	B	338	PHE

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	539/562 (96%)	522 (97%)	17 (3%)	51	67
1	B	539/562 (96%)	524 (97%)	15 (3%)	56	73
All	All	1078/1124 (96%)	1046 (97%)	32 (3%)	53	70

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	547	LYS
1	B	9	GLN
1	B	636	TYR
1	A	664	LEU
1	B	58	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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	665/694 (95%)	-0.37	2 (0%) 91 96	22, 30, 49, 67	0
1	B	665/694 (95%)	-0.36	3 (0%) 88 94	22, 30, 50, 68	1 (0%)
All	All	1330/1388 (95%)	-0.36	5 (0%) 90 95	22, 30, 50, 68	1 (0%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	150	LEU	3.4
1	B	150	LEU	2.8
1	B	235	VAL	2.7
1	A	159	LEU	2.1
1	B	159	LEU	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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, 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	CA	B	701	1/1	0.10	-0.40	30,30,30,30	0
2	CA	A	701	1/1	0.10	-0.74	32,32,32,32	0

6.5 Other polymers ⓘ

There are no such residues in this entry.