



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 04:53 am BST

PDB ID : 4E57  
Title : Crystal Structure of spacer 6aa-shortened cephalosporin acylase mutant  
Authors : Yin, J.; Zhang, Z.; Wu, G.; Huang, X.  
Deposited on : 2012-03-14  
Resolution : 2.00 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

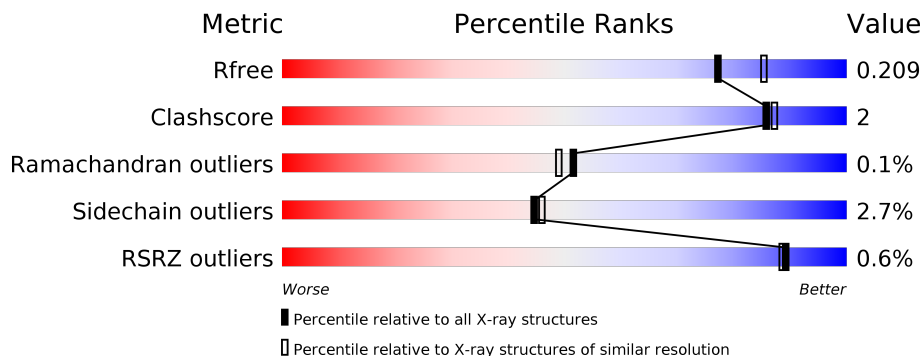
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	698	 % <span style="margin-left: 100px;">90%</span> <span style="margin-left: 100px;">6%</span> . .
1	B	698	 % <span style="margin-left: 100px;">91%</span> <span style="margin-left: 100px;">5%</span> . .

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 12015 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 Cephalosporin acylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	677	5329	3373	940	1004	12	0	0	0
1	B	677	5329	3373	940	1004	12	0	0	0

There are 46 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	?	-	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	166	ALA	SER	ENGINEERED MUTATION	UNP O86089
A	687	LEU	-	EXPRESSION TAG	UNP O86089
A	688	ALA	-	EXPRESSION TAG	UNP O86089
A	689	ALA	-	EXPRESSION TAG	UNP O86089
A	690	ALA	-	EXPRESSION TAG	UNP O86089
A	691	LEU	-	EXPRESSION TAG	UNP O86089
A	692	GLU	-	EXPRESSION TAG	UNP O86089
A	693	HIS	-	EXPRESSION TAG	UNP O86089
A	694	HIS	-	EXPRESSION TAG	UNP O86089
A	695	HIS	-	EXPRESSION TAG	UNP O86089
A	696	HIS	-	EXPRESSION TAG	UNP O86089
A	697	HIS	-	EXPRESSION TAG	UNP O86089
A	698	HIS	-	EXPRESSION TAG	UNP O86089
B	1	MET	-	EXPRESSION TAG	UNP O86089
B	2	GLY	-	EXPRESSION TAG	UNP O86089

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	3	ILE	-	EXPRESSION TAG	UNP O86089
B	127	GLU	ASP	SEE REMARK 999	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	166	ALA	SER	ENGINEERED MUTATION	UNP O86089
B	687	LEU	-	EXPRESSION TAG	UNP O86089
B	688	ALA	-	EXPRESSION TAG	UNP O86089
B	689	ALA	-	EXPRESSION TAG	UNP O86089
B	690	ALA	-	EXPRESSION TAG	UNP O86089
B	691	LEU	-	EXPRESSION TAG	UNP O86089
B	692	GLU	-	EXPRESSION TAG	UNP O86089
B	693	HIS	-	EXPRESSION TAG	UNP O86089
B	694	HIS	-	EXPRESSION TAG	UNP O86089
B	695	HIS	-	EXPRESSION TAG	UNP O86089
B	696	HIS	-	EXPRESSION TAG	UNP O86089
B	697	HIS	-	EXPRESSION TAG	UNP O86089
B	698	HIS	-	EXPRESSION TAG	UNP O86089

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	14	Total Cd 14 14	0	0
2	A	19	Total Cd 19 19	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	681	Total O 681 681	0	0
3	B	643	Total O 643 643	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.60Å 73.60Å 383.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.28 – 2.00 48.28 – 2.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.28-2.00) 100.0 (48.28-2.00)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.65 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.182 , 0.207 0.185 , 0.209	Depositor DCC
$R_{free}$ test set	6866 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.2	Xtrriage
Anisotropy	0.046	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 27.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.470 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	12015	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.81% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.50	6/5484 (0.1%)	0.54	2/7493 (0.0%)
1	B	0.50	5/5484 (0.1%)	0.54	1/7493 (0.0%)
All	All	0.50	11/10968 (0.1%)	0.54	3/14986 (0.0%)

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	71	TRP	CD2-CE2	5.21	1.47	1.41
1	A	191	TRP	CD2-CE2	5.17	1.47	1.41
1	B	417	TRP	CD2-CE2	5.13	1.47	1.41
1	A	605	TRP	CD2-CE2	5.12	1.47	1.41
1	B	191	TRP	CD2-CE2	5.10	1.47	1.41

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	668	LEU	CA-CB-CG	5.96	129.00	115.30
1	B	668	LEU	CA-CB-CG	5.79	128.60	115.30
1	A	163	ASP	CB-CG-OD2	5.21	122.99	118.30

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	5329	0	5070	28	0
1	B	5329	0	5071	17	0
2	A	19	0	0	0	0
2	B	14	0	0	0	0
3	A	681	0	0	2	0
3	B	643	0	0	1	0
All	All	12015	0	10141	45	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 2.

The worst 5 of 45 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:317:GLN:HE22	1:B:338:GLN:H	1.25	0.84
1:A:317:GLN:HE22	1:A:338:GLN:H	1.27	0.82
1:B:166:ALA:H	1:B:409:ASN:HD21	1.26	0.81
1:A:166:ALA:H	1:A:409:ASN:HD21	1.25	0.79
1:A:409:ASN:HD22	1:A:439:ARG:HH22	1.38	0.71

There are no symmetry-related clashes.

## 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	675/698 (97%)	658 (98%)	16 (2%)	1 (0%)	51	49
1	B	675/698 (97%)	660 (98%)	14 (2%)	1 (0%)	51	49
All	All	1350/1396 (97%)	1318 (98%)	30 (2%)	2 (0%)	51	49

All (2) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	B	342	PHE
1	A	342	PHE

### 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	547/564 (97%)	531 (97%)	16 (3%)	42 43
1	B	547/564 (97%)	534 (98%)	13 (2%)	49 51
All	All	1094/1128 (97%)	1065 (97%)	29 (3%)	44 46

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	617	VAL
1	B	9	GLN
1	B	617	VAL
1	A	640	TYR
1	B	58	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	442	GLN
1	B	148	ASN
1	B	409	ASN
1	B	9	GLN
1	B	188	HIS

### 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 33 ligands modelled in this entry, 33 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.

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	677/698 (96%)	-0.30	4 (0%) 89 88	18, 26, 44, 55	0
1	B	677/698 (96%)	-0.28	4 (0%) 89 88	18, 26, 45, 56	1 (0%)
All	All	1354/1396 (96%)	-0.29	8 (0%) 89 88	18, 26, 44, 56	1 (0%)

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	684	ASN	2.7
1	A	685	PHE	2.2
1	A	684	ASN	2.2
1	A	683	PHE	2.2
1	B	685	PHE	2.2

### 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](#)

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( $\text{\AA}^2$ )	Q<0.9
2	CD	A	718	1/1	0.69	0.09	81,81,81,81	1
2	CD	B	713	1/1	0.82	0.11	128,128,128,128	0
2	CD	A	719	1/1	0.83	0.10	129,129,129,129	0
2	CD	B	714	1/1	0.84	0.09	80,80,80,80	1
2	CD	B	705	1/1	0.85	0.09	45,45,45,45	0
2	CD	A	705	1/1	0.85	0.21	123,123,123,123	0
2	CD	B	703	1/1	0.85	0.12	117,117,117,117	0
2	CD	A	716	1/1	0.86	0.12	129,129,129,129	0
2	CD	A	702	1/1	0.88	0.04	109,109,109,109	0
2	CD	A	703	1/1	0.89	0.07	98,98,98,98	0
2	CD	A	701	1/1	0.90	0.05	83,83,83,83	0
2	CD	B	702	1/1	0.90	0.05	75,75,75,75	0
2	CD	A	704	1/1	0.90	0.04	105,105,105,105	0
2	CD	A	708	1/1	0.91	0.08	45,45,45,45	0
2	CD	A	717	1/1	0.91	0.04	73,73,73,73	0
2	CD	B	710	1/1	0.91	0.11	63,63,63,63	1
2	CD	B	708	1/1	0.92	0.07	52,52,52,52	1
2	CD	A	711	1/1	0.92	0.07	52,52,52,52	1
2	CD	A	714	1/1	0.93	0.10	63,63,63,63	1
2	CD	A	713	1/1	0.94	0.05	55,55,55,55	1
2	CD	A	709	1/1	0.94	0.08	52,52,52,52	1
2	CD	B	709	1/1	0.94	0.06	53,53,53,53	1
2	CD	A	715	1/1	0.95	0.05	61,61,61,61	0
2	CD	B	712	1/1	0.95	0.04	62,62,62,62	0
2	CD	B	711	1/1	0.95	0.05	57,57,57,57	1
2	CD	A	710	1/1	0.96	0.04	51,51,51,51	1
2	CD	B	706	1/1	0.96	0.06	54,54,54,54	1
2	CD	B	707	1/1	0.97	0.05	50,50,50,50	1
2	CD	A	706	1/1	0.98	0.16	30,30,30,30	0
2	CD	B	701	1/1	0.98	0.17	14,14,14,14	1
2	CD	B	704	1/1	0.98	0.08	36,36,36,36	0
2	CD	A	707	1/1	0.98	0.08	36,36,36,36	0
2	CD	A	712	1/1	0.99	0.18	52,52,52,52	1

## 6.5 Other polymers i

There are no such residues in this entry.