W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 12:36 AM EDT

PDB ID	:	2NLZ
Title	:	Crystal structure of cephalosporin acylase from Bacillus halodurans
Authors	:	Patskovsky, Y.; Ramagopal, U.; Sauder, J.M.; Dickey, M.; Adams, J.M.; Ozyurt, S.; Wasserman, S.R.; Burley, S.K.; Almo, S.C.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2006-10-20
Resolution	:	2.70 Å(reported)

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

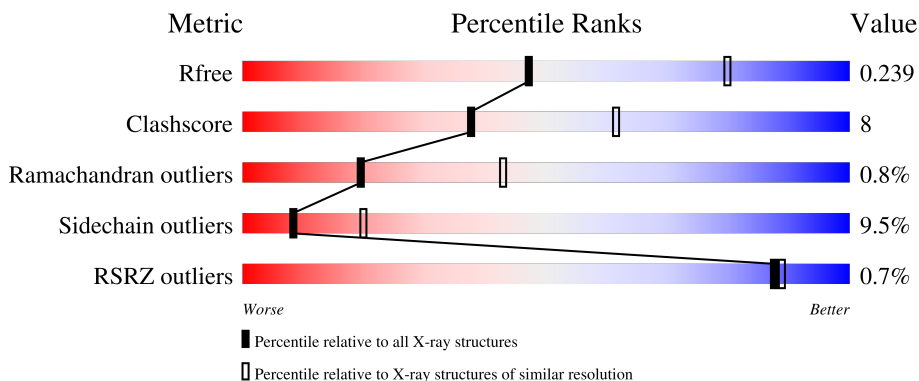|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.70 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	Percentile Ranks	Value
Rfree		0.239
Clashscore		8
Ramachandran outliers		0.8%
Sidechain outliers		9.5%
RSRZ outliers		0.7%

Worse *Better*

▉ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 <=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	547	77% 18%
1	B	547	77% 18%
1	C	547	77% 18%
1	D	547	78% 18%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 16871 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
1	A	536	Total	C	N	O	S	0	3	0
			4153	2645	713	778	17			
1	B	535	Total	C	N	O	S	0	1	0
			4137	2634	711	775	17			
1	C	535	Total	C	N	O	S	0	1	0
			4137	2634	711	775	17			
1	D	537	Total	C	N	O	S	0	2	0
			4163	2651	718	777	17			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	170	VAL	PHE	engineered mutation	UNP Q9KEI5
A	540	GLU	-	cloning artifact	UNP Q9KEI5
A	541	GLY	-	cloning artifact	UNP Q9KEI5
A	542	HIS	-	expression tag	UNP Q9KEI5
A	543	HIS	-	expression tag	UNP Q9KEI5
A	544	HIS	-	expression tag	UNP Q9KEI5
A	545	HIS	-	expression tag	UNP Q9KEI5
A	546	HIS	-	expression tag	UNP Q9KEI5
A	547	HIS	-	expression tag	UNP Q9KEI5
B	170	VAL	PHE	engineered mutation	UNP Q9KEI5
B	540	GLU	-	cloning artifact	UNP Q9KEI5
B	541	GLY	-	cloning artifact	UNP Q9KEI5
B	542	HIS	-	expression tag	UNP Q9KEI5
B	543	HIS	-	expression tag	UNP Q9KEI5
B	544	HIS	-	expression tag	UNP Q9KEI5
B	545	HIS	-	expression tag	UNP Q9KEI5
B	546	HIS	-	expression tag	UNP Q9KEI5
B	547	HIS	-	expression tag	UNP Q9KEI5
C	170	VAL	PHE	engineered mutation	UNP Q9KEI5
C	540	GLU	-	cloning artifact	UNP Q9KEI5
C	541	GLY	-	cloning artifact	UNP Q9KEI5

WORLDWIDE
PDB
PROTEIN DATA BANK

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	542	HIS	-	expression tag	UNP Q9KEI5
C	543	HIS	-	expression tag	UNP Q9KEI5
C	544	HIS	-	expression tag	UNP Q9KEI5
C	545	HIS	-	expression tag	UNP Q9KEI5
C	546	HIS	-	expression tag	UNP Q9KEI5
C	547	HIS	-	expression tag	UNP Q9KEI5
D	170	VAL	PHE	engineered mutation	UNP Q9KEI5
D	540	GLU	-	cloning artifact	UNP Q9KEI5
D	541	GLY	-	cloning artifact	UNP Q9KEI5
D	542	HIS	-	expression tag	UNP Q9KEI5
D	543	HIS	-	expression tag	UNP Q9KEI5
D	544	HIS	-	expression tag	UNP Q9KEI5
D	545	HIS	-	expression tag	UNP Q9KEI5
D	546	HIS	-	expression tag	UNP Q9KEI5
D	547	HIS	-	expression tag	UNP Q9KEI5

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	79	Total 79	O 79	0	0
2	B	73	Total 73	O 73	0	0
2	C	57	Total 57	O 57	0	0
2	D	72	Total 72	O 72	0	0

3   Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Cephalosporin acylase

Chain A:

• Molecule 1: Cephalosporin acylase

Chain B:

• Molecule 1: Cephalosporin acylase

Chain C:

MET SER VAL M4 F5 D6 P7 Q8 Y12 P13 S14 R15 V26 A27 T28 S29 L38 L41 A50 A51 I52 T59 N66 G67 I68 D71 V76 L85 N86 G87 S88 G89 R90 A91 P92 L95 E98 K101 Y105 L109 P110 V119 P120 G121

A122 P123 G124 A125 W126 A127 E128 L129 M132 E150 K164 A165 A166 R169 V170 K171 W174 D177 V178 R191 A192 G203 H204 A205 D206 T207 I211 E223 L239 R248 P249 S256 I257 W265 P269 N270 L277 L280 V283 F288 Y289

H290 K291 Q300 F308 V309 S319 D320 V325 L329 P348 T352 P353 PRO ARG GLY GLY T358 V359 Y360 L361 A362 T363 M370 V371 S372 F373 I374 Q375 S376 N377 Y378 V387 T390 R397 N400 H407 D408 N409 A410 L411 R416 I421 I422 L426

T427 P432 I433 M439 Q444 M453 F458 G459 L460 N461 P462 L466 A490 L504 D505 T520 T521 G526 G527 T528 V536 H542 HIS HIS HIS HIS

• Molecule 1: Cephalosporin acylase

Chain D:

%

78%   18%   • • •

MET SER VAL M4 P7 Q8 R15 V26 A27 L41 A50 A51 I52 T59 N66 D71 V76 G87 S88 G89 R90 A91 P92 L95 E98 Y105 L109 P110 V119 P120 G121 A122 P123 W126 A127 E128 L129 M132 A139 E150

A166 R169 V170 K171 W174 D177 V178 Y179 K189 G190 R191 G203 H204 A205 D206 T207 I211 A212 S218 G222 E223 L239 R248 P249 S256 I257 W265 P269 N270 Q271 Q272 L277 L280 V283 F288 Q300 F308 V309 M312

D320 V325 L329 P348 T352 P353 PRO ARG GLY GLY T358 V359 Y360 L361 A362 T363 M370 V371 S372 F373 I374 Q375 S376 V387 T390 H407 D408 N409 A410 L411 K412 P413 R416 T417 Y418 I421 I422 L426 T427 P432 I433 M439 Q444 M453

L460 N461 P462 L466 V479 I500 L504 G507 G510 R517 T520 T521 G526 G527 T528 V536 H544 HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	105.75Å 105.75Å 385.06Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.70 19.98 – 2.70	Depositor EDS
% Data completeness (in resolution range)	95.2 (20.00-2.70) 94.9 (19.98-2.70)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	1.84 (at 2.71Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.210 , 0.245 0.209 , 0.239	Depositor DCC
R_{free} test set	2015 reflections (3.03%)	wwPDB-VP
Wilson B-factor (Å^2)	37.5	Xtriage
Anisotropy	0.198	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 1.0	EDS
L-test for twinning[2]	$< |L| > = 0.44, < L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.055 for -h,-k,l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	16871	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.40% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0.42	0/4277	0.61	2/5824 (0.0%)
1	B	0.40	0/4255	0.62	2/5794 (0.0%)
1	C	0.42	0/4255	0.61	1/5794 (0.0%)
1	D	0.41	0/4286	0.63	6/5835 (0.1%)
All	All	0.41	0/17073	0.62	11/23247 (0.0%)

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	1
1	B	0	1
1	C	0	1
1	D	0	1
All	All	0	4

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	361	LEU	CA-CB-CG	6.76	130.84	115.30
1	C	361	LEU	CA-CB-CG	6.52	130.30	115.30
1	D	517	ARG	NE-CZ-NH1	-6.39	117.10	120.30
1	A	361	LEU	CA-CB-CG	5.75	128.51	115.30
1	B	90	ARG	NE-CZ-NH2	-5.66	117.47	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	269	PRO	Peptide
1	B	269	PRO	Peptide
1	C	269	PRO	Peptide
1	D	269	PRO	Peptide

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	4153	0	4035	68	0
1	B	4137	0	4013	62	0
1	C	4137	0	4013	65	0
1	D	4163	0	4040	63	0
2	A	79	0	0	3	0
2	B	73	0	0	2	0
2	C	57	0	0	2	0
2	D	72	0	0	1	0
All	All	16871	0	16101	251	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 251 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:376:SER:HB3	1:A:421:ILE:HD11	1.50	0.91
1:A:426:LEU:HD23	1:A:433:ILE:HD11	1.56	0.88
1:D:376:SER:HB3	1:D:421:ILE:HD11	1.59	0.84
1:B:376:SER:HB3	1:B:421:ILE:HD11	1.62	0.82
1:C:248:ARG:HH11	1:C:248:ARG:CG	1.93	0.82

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	535/547 (98%)	509 (95%)	22 (4%)	4 (1%)	22	46
1	B	532/547 (97%)	511 (96%)	16 (3%)	5 (1%)	17	40
1	C	532/547 (97%)	506 (95%)	21 (4%)	5 (1%)	17	40
1	D	535/547 (98%)	509 (95%)	22 (4%)	4 (1%)	22	46
All	All	2134/2188 (98%)	2035 (95%)	81 (4%)	18 (1%)	19	43

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	439	MET
1	C	439	MET
1	D	439	MET
1	A	7	PRO
1	B	7	PRO

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	428/434 (99%)	385 (90%)	43 (10%)	7	18
1	B	425/434 (98%)	387 (91%)	38 (9%)	9	22
1	C	425/434 (98%)	381 (90%)	44 (10%)	7	16
1	D	428/434 (99%)	386 (90%)	42 (10%)	8	18
All	All	1706/1736 (98%)	1539 (90%)	167 (10%)	8	18

5 of 167 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	352	THR
1	D	191	ARG
1	C	433	ILE
1	D	76	VAL
1	D	270	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	407	HIS
1	D	66	ASN
1	C	444	GLN
1	C	513	GLN
1	D	290	HIS

5.3.3　RNA ⓘ

There are no RNA molecules 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 monosaccharides in this entry.

5.6　Ligand geometry ⓘ

There are no ligands in this entry.

5.7　Other polymers ⓘ

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 ⓘ

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	536/547 (97%)	-0.51	5 (0%)	84	85	11, 39, 93, 162	0
1	B	535/547 (97%)	-0.51	2 (0%)	92	93	11, 38, 93, 162	0
1	C	535/547 (97%)	-0.51	4 (0%)	87	89	11, 39, 93, 162	0
1	D	537/547 (98%)	-0.52	3 (0%)	89	91	12, 39, 94, 161	0
All	All	2143/2188 (97%)	-0.51	14 (0%)	87	89	11, 39, 93, 162	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	4	MET	4.2
1	A	289	TYR	4.0
1	C	289	TYR	3.0
1	D	4	MET	2.9
1	D	105	TYR	2.9

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 monosaccharides in this entry.

6.4   Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.