Full wwPDB X-ray Structure Validation Report ⓘ

Jun 1, 2021 – 06:01 pm BST

PDB ID	:	6Z6Z
Title	:	Crystal structure of an Anticalin directed towards colchicine without ligand
Authors	:	Skerra, A.; Eichinger, A.
Deposited on	:	2020-05-29
Resolution	:	1.78 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see references ⓘ) were used in the production of this report:

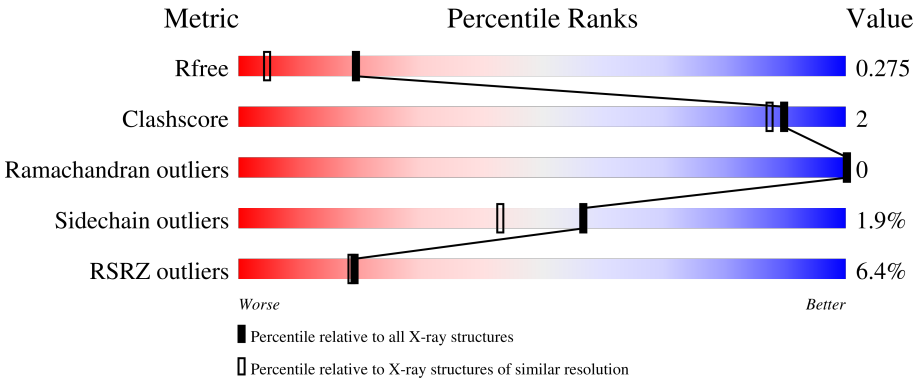
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.19
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.19

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.78 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	9185 (1.80-1.76)
Clashscore	141614	10184 (1.80-1.76)
Ramachandran outliers	138981	10051 (1.80-1.76)
Sidechain outliers	138945	10050 (1.80-1.76)
RSRZ outliers	127900	9032 (1.80-1.76)

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	187	6% 87% 5% · 8%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 1466 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 Neutrophil gelatinase-associated lipocalin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	172	Total	C	N	O	S	0	2	0
			1404	910	230	257	7			

There are 33 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P80188
A	28	HIS	GLN	engineered mutation	UNP P80188
A	36	VAL	LEU	engineered mutation	UNP P80188
A	40	GLY	ALA	engineered mutation	UNP P80188
A	41	PHE	ILE	engineered mutation	UNP P80188
A	49	ILE	GLN	engineered mutation	UNP P80188
A	52	ALA	TYR	engineered mutation	UNP P80188
A	68	PHE	SER	engineered mutation	UNP P80188
A	69	MET	VAL	engineered mutation	UNP P80188
A	70	LYS	LEU	engineered mutation	UNP P80188
A	72	PRO	ARG	engineered mutation	UNP P80188
A	73	MET	LYS	engineered mutation	UNP P80188
A	77	GLN	ASP	engineered mutation	UNP P80188
A	79	MET	TRP	engineered mutation	UNP P80188
A	80	THR	ILE	engineered mutation	UNP P80188
A	81	ASP	ARG	engineered mutation	UNP P80188
A	83	LEU	PHE	engineered mutation	UNP P80188
A	87	SER	CYS	engineered mutation	UNP P80188
A	100	GLU	TYR	engineered mutation	UNP P80188
A	103	TYR	LEU	engineered mutation	UNP P80188
A	106	TRP	TYR	engineered mutation	UNP P80188
A	125	ALA	LYS	engineered mutation	UNP P80188
A	127	GLN	SER	engineered mutation	UNP P80188
A	132	ASP	TYR	engineered mutation	UNP P80188
A	134	PHE	LYS	engineered mutation	UNP P80188
A	179	SER	-	expression tag	UNP P80188
A	180	ALA	-	expression tag	UNP P80188

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	181	HIS	-	expression tag	UNP P80188
A	182	HIS	-	expression tag	UNP P80188
A	183	HIS	-	expression tag	UNP P80188
A	184	HIS	-	expression tag	UNP P80188
A	185	HIS	-	expression tag	UNP P80188
A	186	HIS	-	expression tag	UNP P80188

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

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

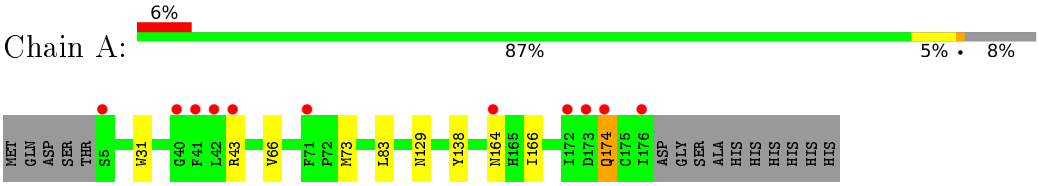
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	61	Total O 61 61	0	0

3 Residue-property plots (i)

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: Neutrophil gelatinase-associated lipocalin



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	38.56Å 75.58Å 116.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	63.34 – 1.78 63.34 – 1.78	Depositor EDS
% Data completeness (in resolution range)	99.3 (63.34-1.78) 99.3 (63.34-1.78)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	4.09 (at 1.78Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.201 , 0.259 0.210 , 0.275	Depositor DCC
R_{free} test set	811 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å^2)	26.6	Xtriage
Anisotropy	0.899	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 41.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	1466	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 9.53% 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

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	0.63	0/1445	0.77	0/1958

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1404	0	1390	5	0
2	A	1	0	0	0	0
3	A	61	0	0	0	1
All	All	1466	0	1390	5	1

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.

All (5) 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:174:GLN:OE1	1:A:174:GLN:HA	1.97	0.65

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:31:TRP:CE3	1:A:138:TYR:HB3	2.47	0.49
1:A:66[A]:VAL:HG21	1:A:83:LEU:HG	1.98	0.45
1:A:43:ARG:NH1	1:A:166:ILE:O	2.52	0.43
1:A:129:ASN:OD1	1:A:129:ASN:N	2.50	0.42

All (1) 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)
3:A:311:HOH:O	3:A:330:HOH:O[3_555]	2.06	0.14

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	172/187 (92%)	163 (95%)	9 (5%)	0	100 100

There are no Ramachandran outliers to report.

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	158/169 (94%)	155 (98%)	3 (2%)	57 43

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	73	MET
1	A	164	ASN
1	A	174	GLN

Sometimes 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 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

Of 1 ligands modelled in this entry, 1 is 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

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, 95^{th} 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	172/187 (91%)	0.27	11 (6%)	19	18	22, 41, 85, 95	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	174	GLN	4.1
1	A	176	ILE	4.0
1	A	40	GLY	3.6
1	A	71	PHE	3.3
1	A	164	ASN	3.3
1	A	43	ARG	2.8
1	A	42	LEU	2.7
1	A	173	ASP	2.6
1	A	5	SER	2.5
1	A	172	ILE	2.4
1	A	41	PHE	2.0

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

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,

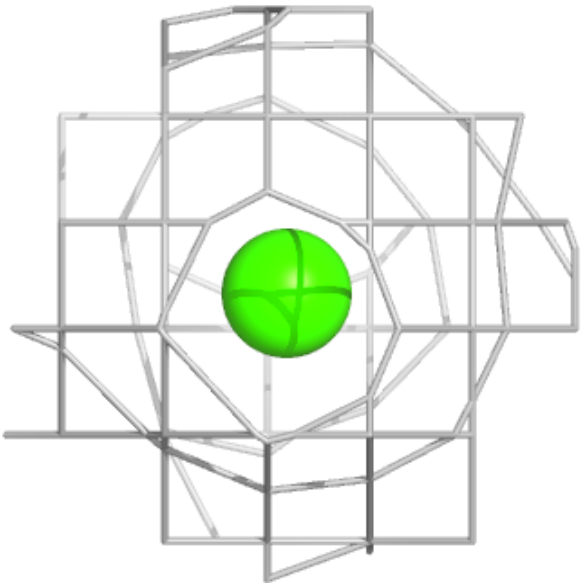
median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

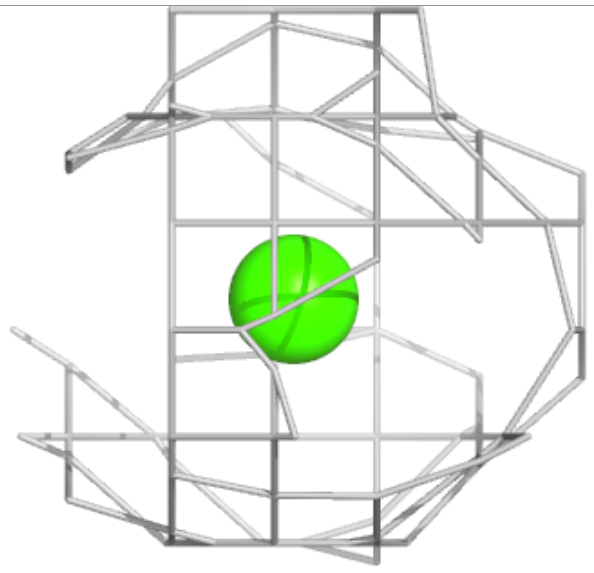
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CA	A	201	1/1	0.99	0.06	25,25,25,25	0

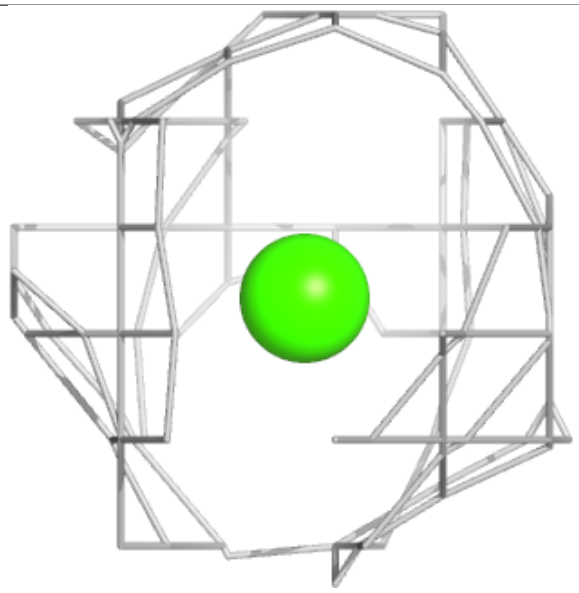
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around CA A 201:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)







6.5 Other polymers (i)

There are no such residues in this entry.