



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 24, 2017 – 11:39 PM EDT

PDB ID : 5TZ9
Title : Structure of plasma kallikrein protease domain with inhibitor
Authors : Partridge, J.R.; Choy, R.
Deposited on : unknown
Resolution : 1.70 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

<http://wwpdb.org/validation/2016/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
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030345
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030345

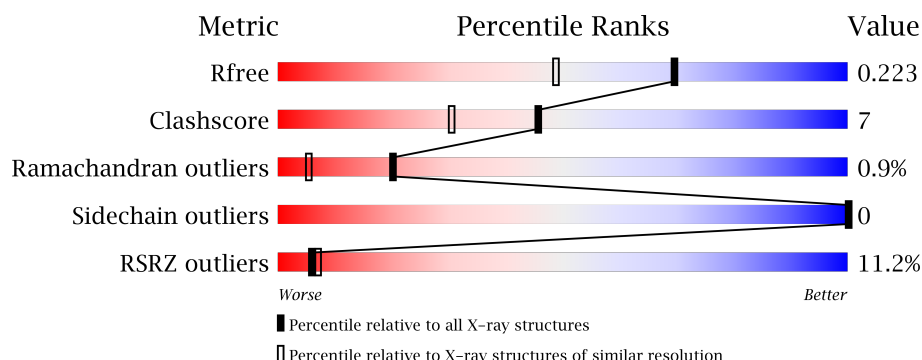
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION


The reported resolution of this entry is 1.70 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	3453 (1.70-1.70)
Clashscore	112137	3876 (1.70-1.70)
Ramachandran outliers	110173	3815 (1.70-1.70)
Sidechain outliers	110143	3815 (1.70-1.70)
RSRZ outliers	101464	3491 (1.70-1.70)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	E	263	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 2103 atoms, of which 25 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.

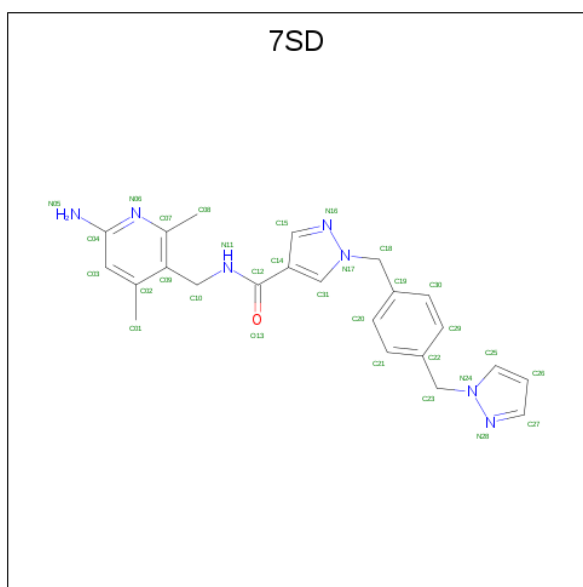
- Molecule 1 is a protein called Plasma kallikrein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	E	232	1861	1185	317	347	12	0	3	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	396	GLU	ASN	engineered mutation	UNP P03952
E	453	GLU	ASN	engineered mutation	UNP P03952
E	494	GLU	ASN	engineered mutation	UNP P03952
E	503	SER	CYS	engineered mutation	UNP P03952

- Molecule 2 is N-[(6-amino-2,4-dimethylpyridin-3-yl)methyl]-1-({4-[(1H-pyrazol-1-yl)methyl]phenyl}methyl)-1H-pyrazole-4-carboxamide (three-letter code: 7SD) (formula: C₂₃H₂₅N₇O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
2	E	1	56	23	25	7	1	0	0

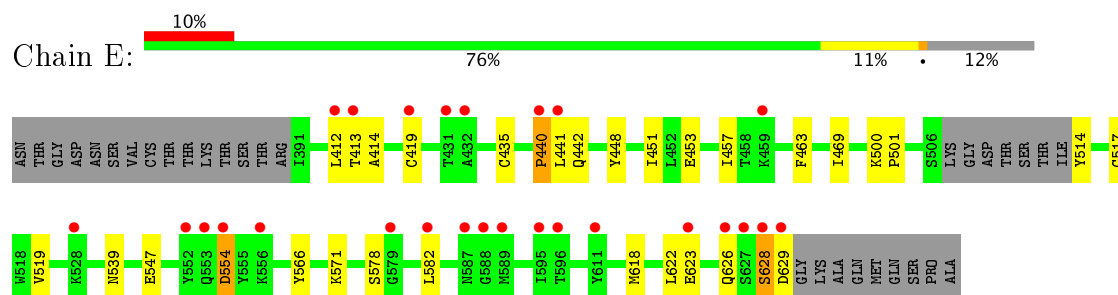
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	186	Total 186	O 186	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 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: Plasma kallikrein



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	55.69Å 57.45Å 86.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.84 – 1.70 27.84 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.9 (27.84-1.70) 97.1 (27.84-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.49 (at 1.70Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.182 , 0.219 0.182 , 0.223	Depositor DCC
R_{free} test set	1941 reflections (6.43%)	DCC
Wilson B-factor (Å ²)	21.4	Xtriage
Anisotropy	0.540	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 46.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.024 for k,h,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	2103	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 5.83% 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: 7SD

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	E	0.34	0/1910	0.51	1/2588 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	E	554	ASP	CB-CG-OD2	5.24	123.02	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	E	1861	0	1816	27	0
2	E	31	25	0	0	0
3	E	186	0	0	7	0
All	All	2078	25	1816	27	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 7.

The worst 5 of 27 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:E:554:ASP:OD1	3:E:801:HOH:O	1.79	0.99
1:E:628:SER:HB3	3:E:813:HOH:O	1.92	0.69
1:E:442:GLN:NE2	3:E:805:HOH:O	2.23	0.62
1:E:517:CYS:O	1:E:539:ASN:ND2	2.31	0.60
1:E:453:GLU:OE1	3:E:802:HOH:O	2.18	0.53

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	E	231/263 (88%)	220 (95%)	9 (4%)	2 (1%)	20 5

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	440	PRO
1	E	628	SER

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	E	203/226 (90%)	203 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

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

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](#)

1 ligand is modelled in this entry.

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	7SD	E	701	-	33,34,34	2.37	12 (36%)	39,47,47	1.76	6 (15%)

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	7SD	E	701	-	-	0/17/17/17	0/4/4/4

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	701	7SD	C31-N17	-5.74	1.29	1.35
2	E	701	7SD	C04-N06	-4.90	1.26	1.35
2	E	701	7SD	C02-C09	-4.48	1.33	1.40
2	E	701	7SD	N16-N17	-3.82	1.30	1.35
2	E	701	7SD	C25-N24	-3.69	1.29	1.35

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	701	7SD	C31-N17-N16	-4.14	108.03	111.57
2	E	701	7SD	C19-C18-N17	-3.58	106.90	112.17
2	E	701	7SD	C22-C23-N24	-2.92	107.87	112.17
2	E	701	7SD	C25-N24-N28	-2.56	109.38	111.57
2	E	701	7SD	C27-N28-N24	5.34	109.69	104.22

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, 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	E	232/263 (88%)	0.62	26 (11%) 6 7	13, 27, 59, 115	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	628	SER	12.0
1	E	627	SER	9.6
1	E	553	GLN	7.0
1	E	629	ASP	7.0
1	E	554	ASP	6.5

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. 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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	7SD	E	701	31/31	0.93	0.09	-0.47	17,26,35,42	0

6.5 Other polymers [i](#)

There are no such residues in this entry.