



Full wwPDB X-ray Structure Validation Report ⓘ

Nov 7, 2023 – 05:53 PM JST

PDB ID : 8I16
Title : Crystal structure of the selenomethionine (SeMet)-derived Cas12g (D513A) mutant
Authors : Zhang, B.; Chen, J.; Ye, Y.M.; OuYang, S.Y.
Deposited on : 2023-01-12
Resolution : 2.24 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
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

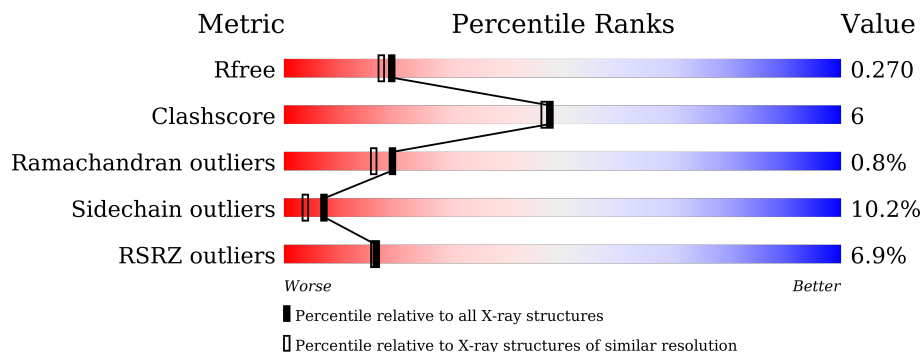
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.24 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	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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 $\leq 5\%$. 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	775	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5382 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 Cas12g.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	S				Se
1	A	639	5338	3389	1008	927	9	5	0	0	0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Zn	0	0
			2	2		

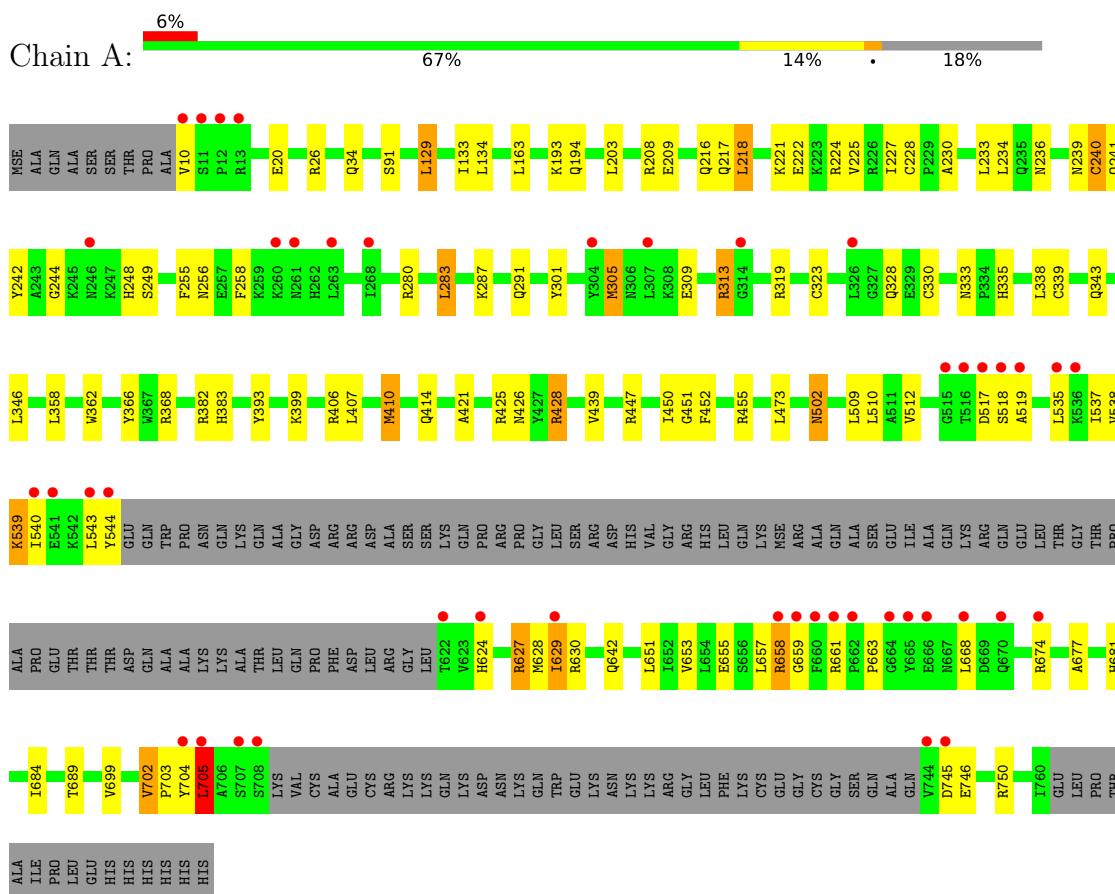
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	42	Total	O	0	0
			42	42		

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: Cas12g



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	64.07Å 95.12Å 167.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	62.86 – 2.24 62.86 – 2.24	Depositor EDS
% Data completeness (in resolution range)	99.3 (62.86-2.24) 99.3 (62.86-2.24)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.39 (at 2.25Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.224 , 0.268 0.231 , 0.270	Depositor DCC
R_{free} test set	2416 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å ²)	53.5	Xtrriage
Anisotropy	0.073	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 36.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	5382	wwPDB-VP
Average B, all atoms (Å ²)	67.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.09% 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:
ZN

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	0.71	0/5467	0.87	0/7353

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 [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	5338	0	5324	65	0
2	A	2	0	0	0	0
3	A	42	0	0	0	0
All	All	5382	0	5324	65	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 6.

All (65) 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:255:PHE:HA	1:A:305:MSE:HE1	1.31	1.12

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:702:VAL:HG12	1:A:703:PRO:HD2	1.22	1.11
1:A:702:VAL:HG12	1:A:703:PRO:CD	1.96	0.94
1:A:627:ARG:HH21	1:A:627:ARG:HG3	1.39	0.86
1:A:627:ARG:O	1:A:627:ARG:NH2	2.09	0.86
1:A:255:PHE:HA	1:A:305:MSE:CE	2.04	0.85
1:A:658:ARG:HG3	1:A:704:TYR:CE2	2.14	0.82
1:A:255:PHE:CA	1:A:305:MSE:HE1	2.09	0.81
1:A:26:ARG:HD3	1:A:421:ALA:O	1.81	0.81
1:A:627:ARG:HG3	1:A:627:ARG:NH2	1.97	0.75
1:A:240:CYS:SG	1:A:242:TYR:N	2.59	0.70
1:A:228:CYS:SG	1:A:240:CYS:HA	2.31	0.69
1:A:681:HIS:O	1:A:684:ILE:HG22	1.92	0.69
1:A:702:VAL:CG1	1:A:703:PRO:CD	2.72	0.67
1:A:702:VAL:CG1	1:A:703:PRO:HD2	2.13	0.66
1:A:629:ILE:HD13	1:A:629:ILE:N	2.12	0.63
1:A:407:LEU:HB2	1:A:410:MSE:HE3	1.81	0.61
1:A:657:LEU:HD21	1:A:677:ALA:HB1	1.82	0.61
1:A:233:LEU:O	1:A:343:GLN:NE2	2.34	0.61
1:A:248:HIS:ND1	1:A:330:CYS:HB2	2.17	0.59
1:A:323:CYS:SG	1:A:330:CYS:HA	2.42	0.59
1:A:20:GLU:O	1:A:455:ARG:HG3	2.04	0.58
1:A:283:LEU:HD22	1:A:291:GLN:HG2	1.86	0.57
1:A:248:HIS:CE1	1:A:328:GLN:O	2.59	0.55
1:A:208:ARG:HD3	1:A:362:TRP:CE2	2.40	0.55
1:A:224:ARG:HB3	1:A:663:PRO:HG3	1.89	0.55
1:A:217:GLN:C	1:A:218:LEU:O	2.42	0.55
1:A:519:ALA:HB1	1:A:535:LEU:O	2.07	0.55
1:A:428:ARG:NE	1:A:428:ARG:HA	2.22	0.54
1:A:653:VAL:HG13	1:A:702:VAL:CG2	2.39	0.53
1:A:653:VAL:CG1	1:A:702:VAL:CG2	2.89	0.51
1:A:658:ARG:HA	1:A:704:TYR:CG	2.46	0.50
1:A:309:GLU:OE2	1:A:313:ARG:NH1	2.45	0.49
1:A:216:GLN:O	1:A:218:LEU:O	2.31	0.49
1:A:627:ARG:HH21	1:A:627:ARG:CG	2.14	0.49
1:A:225:VAL:HG12	1:A:227:ILE:HG13	1.94	0.49
1:A:335:HIS:HB3	1:A:339:CYS:HB2	1.96	0.48
1:A:502:ASN:C	1:A:502:ASN:HD22	2.17	0.48
1:A:539:LYS:HA	1:A:539:LYS:HE2	1.96	0.47
1:A:240:CYS:SG	1:A:242:TYR:HB2	2.55	0.46
1:A:629:ILE:N	1:A:629:ILE:CD1	2.79	0.45
1:A:330:CYS:SG	1:A:333:ASN:HB2	2.56	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:705:LEU:HD12	1:A:705:LEU:HA	1.82	0.45
1:A:657:LEU:HD12	1:A:657:LEU:HA	1.85	0.45
1:A:406:ARG:HG3	1:A:410:MSE:HG2	1.98	0.44
1:A:439:VAL:HA	1:A:451:GLY:O	2.18	0.44
1:A:258:PHE:HB2	1:A:305:MSE:HE2	2.00	0.43
1:A:230:ALA:HB1	1:A:346:LEU:HD13	2.01	0.43
1:A:509:LEU:HA	1:A:651:LEU:O	2.18	0.43
1:A:630:ARG:HE	1:A:630:ARG:HB2	1.68	0.43
1:A:240:CYS:SG	1:A:241:GLN:N	2.92	0.43
1:A:382:ARG:O	1:A:383:HIS:CG	2.72	0.43
1:A:627:ARG:HG3	1:A:627:ARG:O	2.18	0.42
1:A:129:LEU:HD22	1:A:133:ILE:CD1	2.49	0.42
1:A:450:ILE:HD12	1:A:452:PHE:CZ	2.55	0.42
1:A:255:PHE:N	1:A:305:MSE:HE1	2.35	0.42
1:A:393:TYR:HA	1:A:406:ARG:HB3	2.02	0.42
1:A:228:CYS:HB2	1:A:233:LEU:HD13	2.01	0.41
1:A:193:LYS:HG3	1:A:366:TYR:HB3	2.02	0.41
1:A:509:LEU:C	1:A:509:LEU:HD23	2.40	0.41
1:A:689:THR:HA	1:A:699:VAL:HG21	2.03	0.41
1:A:655:GLU:OE2	1:A:705:LEU:HB2	2.20	0.40
1:A:244:GLY:HA3	1:A:248:HIS:O	2.21	0.40
1:A:258:PHE:CE1	1:A:301:TYR:HA	2.57	0.40
1:A:230:ALA:HB1	1:A:346:LEU:CD1	2.51	0.40

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	633/775 (82%)	603 (95%)	25 (4%)	5 (1%)	19 16

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	218	LEU
1	A	518	SER
1	A	668	LEU
1	A	705	LEU
1	A	659	GLY

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	559/665 (84%)	502 (90%)	57 (10%)	7 4

All (57) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	VAL
1	A	34	GLN
1	A	91	SER
1	A	129	LEU
1	A	134	LEU
1	A	163	LEU
1	A	194	GLN
1	A	203	LEU
1	A	209	GLU
1	A	221	LYS
1	A	222	GLU
1	A	234	LEU
1	A	236	ASN
1	A	239	ASN
1	A	240	CYS
1	A	249	SER
1	A	256	ASN
1	A	280	ARG
1	A	283	LEU
1	A	287	LYS
1	A	305	MSE
1	A	313	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	319	ARG
1	A	338	LEU
1	A	358	LEU
1	A	368	ARG
1	A	399	LYS
1	A	410	MSE
1	A	414	GLN
1	A	425	ARG
1	A	426	ASN
1	A	428	ARG
1	A	447	ARG
1	A	473	LEU
1	A	502	ASN
1	A	510	LEU
1	A	512	VAL
1	A	517	ASP
1	A	537	ILE
1	A	538	VAL
1	A	539	LYS
1	A	540	ILE
1	A	543	LEU
1	A	544	TYR
1	A	624	HIS
1	A	627	ARG
1	A	628	MSE
1	A	629	ILE
1	A	642	GLN
1	A	658	ARG
1	A	661	ARG
1	A	674	ARG
1	A	702	VAL
1	A	705	LEU
1	A	745	ASP
1	A	746	GLU
1	A	750	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	239	ASN
1	A	241	GLN
1	A	264	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	343	GLN
1	A	430	GLN
1	A	502	ASN
1	A	639	GLN
1	A	647	ASN
1	A	670	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

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.

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

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	634/775 (81%)	0.42	44 (6%) 16 16	35, 60, 129, 172	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	516	THR	11.9
1	A	665	TYR	9.2
1	A	10	VAL	9.0
1	A	515	GLY	7.8
1	A	660	PHE	6.1
1	A	708	SER	6.0
1	A	670	GLN	5.8
1	A	11	SER	5.5
1	A	744	VAL	5.1
1	A	544	TYR	5.1
1	A	659	GLY	4.8
1	A	707	SER	4.4
1	A	666	GLU	4.3
1	A	314	GLY	4.3
1	A	260	LYS	3.9
1	A	745	ASP	3.8
1	A	517	ASP	3.6
1	A	307	LEU	3.6
1	A	662	PRO	3.5
1	A	518	SER	3.4
1	A	622	THR	3.2
1	A	704	TYR	3.1
1	A	705	LEU	3.1
1	A	543	LEU	3.0
1	A	540	ILE	2.9
1	A	541	GLU	2.9
1	A	658	ARG	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	661	ARG	2.7
1	A	304	TYR	2.6
1	A	268	ILE	2.6
1	A	536	LYS	2.5
1	A	535	LEU	2.4
1	A	261	ASN	2.4
1	A	12	PRO	2.3
1	A	246	ASN	2.2
1	A	629	ILE	2.2
1	A	519	ALA	2.1
1	A	13	ARG	2.1
1	A	624	HIS	2.1
1	A	674	ARG	2.1
1	A	664	GLY	2.1
1	A	263	LEU	2.1
1	A	668	LEU	2.0
1	A	326	LEU	2.0

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 monosaccharides 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, 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	B-factors(Å ²)	Q<0.9
2	ZN	A	802	1/1	0.95	0.15	80,80,80,80	0
2	ZN	A	801	1/1	0.96	0.15	63,63,63,63	0

6.5 Other polymers [i](#)

There are no such residues in this entry.