



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 15, 2024 – 08:54 AM JST

PDB ID : 6IV8  
Title : the selenomethionine(SeMet)-derived Cas13d binary complex  
Authors : Zhang, B.; Ye, Y.M.; Ye, W.W.; OuYang, S.Y.  
Deposited on : 2018-12-02  
Resolution : 2.15 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 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 : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

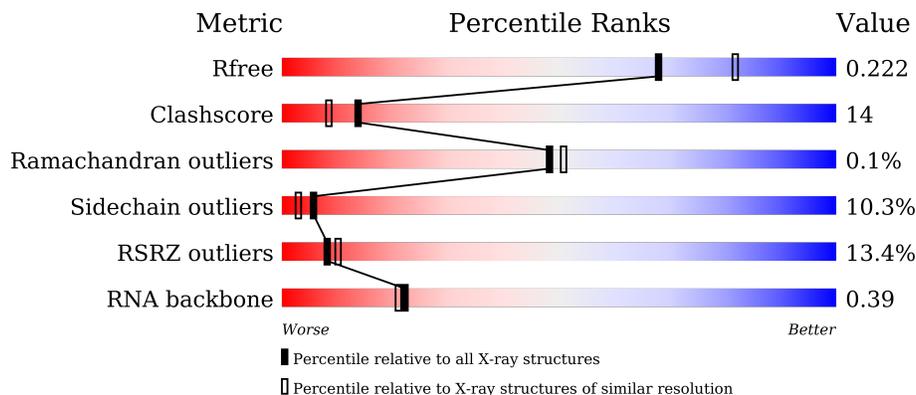
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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))
$R_{free}$	164625	1881 (2.16-2.16)
Clashscore	180529	2047 (2.16-2.16)
Ramachandran outliers	177936	2027 (2.16-2.16)
Sidechain outliers	177891	2026 (2.16-2.16)
RSRZ outliers	164620	1882 (2.16-2.16)
RNA backbone	3690	1055 (2.50-1.82)

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  $\geq 3$ , 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	930	
1	C	930	
2	B	51	
3	D	53	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 16939 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 The selenomethionine (SeMet)-labeled Cas13d.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	869	7065	4494	1195	1341	10	25	0	0	0
1	C	862	7011	4465	1188	1323	10	25	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	288	ALA	ARG	engineered mutation	UNP A0A1C5SD84
A	823	ALA	ARG	engineered mutation	UNP A0A1C5SD84
A	923	LEU	-	expression tag	UNP A0A1C5SD84
A	924	GLU	-	expression tag	UNP A0A1C5SD84
A	925	HIS	-	expression tag	UNP A0A1C5SD84
A	926	HIS	-	expression tag	UNP A0A1C5SD84
A	927	HIS	-	expression tag	UNP A0A1C5SD84
A	928	HIS	-	expression tag	UNP A0A1C5SD84
A	929	HIS	-	expression tag	UNP A0A1C5SD84
A	930	HIS	-	expression tag	UNP A0A1C5SD84
C	288	ALA	ARG	engineered mutation	UNP A0A1C5SD84
C	823	ALA	ARG	engineered mutation	UNP A0A1C5SD84
C	923	LEU	-	expression tag	UNP A0A1C5SD84
C	924	GLU	-	expression tag	UNP A0A1C5SD84
C	925	HIS	-	expression tag	UNP A0A1C5SD84
C	926	HIS	-	expression tag	UNP A0A1C5SD84
C	927	HIS	-	expression tag	UNP A0A1C5SD84
C	928	HIS	-	expression tag	UNP A0A1C5SD84
C	929	HIS	-	expression tag	UNP A0A1C5SD84
C	930	HIS	-	expression tag	UNP A0A1C5SD84

- Molecule 2 is a RNA chain called RNA (51-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	51	1077	484	190	352	51	0	0	0

- Molecule 3 is a RNA chain called RNA (53-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	D	53	1119	503	198	365	53	0	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Mg	0	0
			1	1		
4	D	1	Total	Mg	0	0
			1	1		

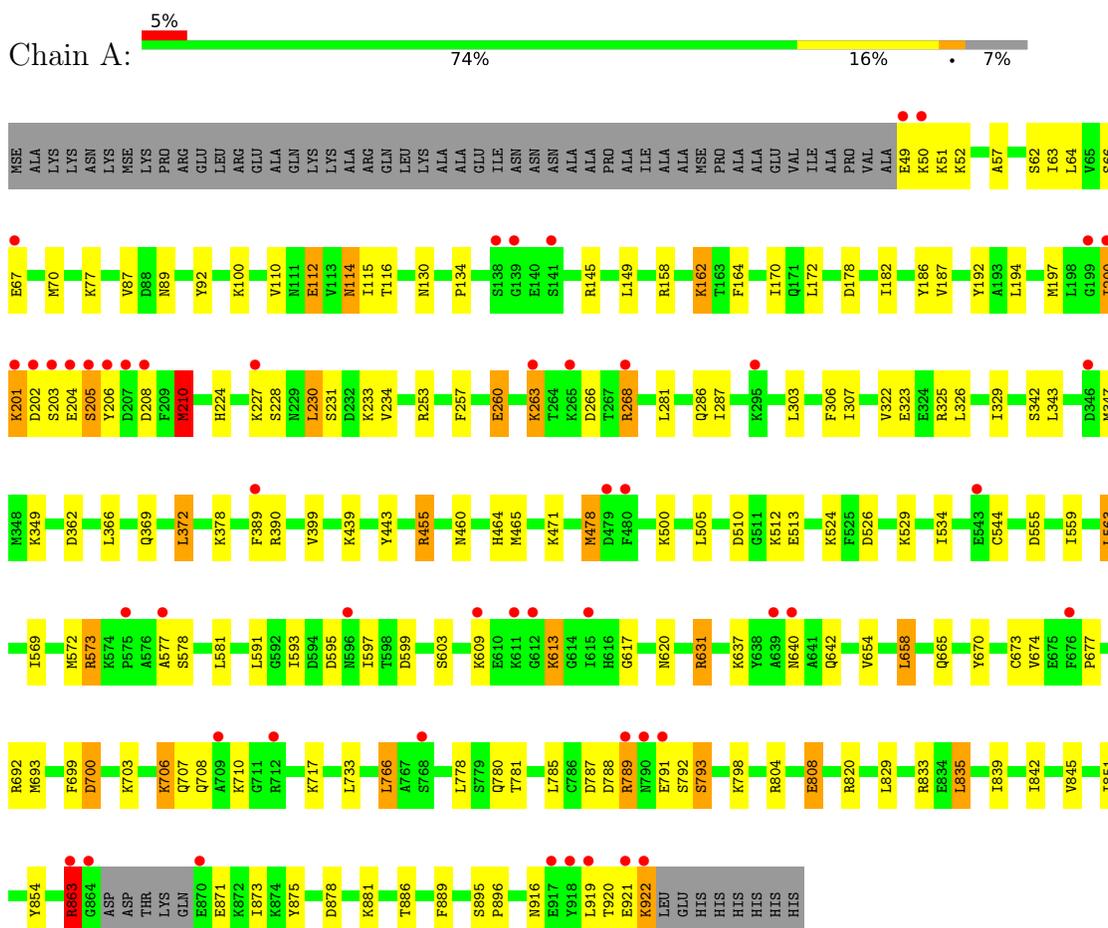
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	444	Total	O	0	0
			444	444		
5	B	106	Total	O	0	0
			106	106		
5	D	46	Total	O	0	0
			46	46		
5	C	69	Total	O	0	0
			69	69		

### 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: The selenomethionine (SeMet)-labeled Cas13d



- Molecule 1: The selenomethionine (SeMet)-labeled Cas13d





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.16Å 145.73Å 249.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	62.92 – 2.15 62.92 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.6 (62.92-2.15) 99.6 (62.92-2.15)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.53 (at 2.14Å)	Xtrriage
Refinement program	REFMAC 5.8.0222	Depositor
R, $R_{free}$	0.189 , 0.210 0.194 , 0.222	Depositor DCC
$R_{free}$ test set	6627 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	30.2	Xtrriage
Anisotropy	0.014	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 41.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	16939	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% 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: MG

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	0.93	2/7153 (0.0%)	0.82	4/9550 (0.0%)
1	C	0.71	0/7098	0.81	8/9475 (0.1%)
2	B	1.19	12/1203 (1.0%)	1.48	30/1869 (1.6%)
3	D	1.14	8/1250 (0.6%)	1.01	5/1942 (0.3%)
All	All	0.88	22/16704 (0.1%)	0.90	47/22836 (0.2%)

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	15
1	C	0	28
All	All	0	43

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	-9	G	O3'-P	-10.32	1.48	1.61
2	B	-17	U	C2'-O2'	8.80	1.53	1.41
2	B	10	A	O3'-P	-7.76	1.51	1.61
3	D	-7	C	O3'-P	-7.76	1.51	1.61
3	D	-8	U	O3'-P	-7.70	1.51	1.61

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	-17	U	O5'-P-OP1	-16.73	90.62	110.70
2	B	-22	C	O5'-P-OP1	12.10	125.22	110.70

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	18	A	O5'-P-OP2	-10.55	96.21	105.70
2	B	-24	U	O5'-P-OP2	-9.58	97.08	105.70
2	B	-17	U	O5'-P-OP2	9.31	121.87	110.70

There are no chirality outliers.

5 of 43 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	145	ARG	Sidechain
1	A	158	ARG	Sidechain
1	A	253	ARG	Sidechain
1	A	268	ARG	Sidechain
1	A	325	ARG	Sidechain

## 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	7065	0	7131	113	0
1	C	7011	0	7094	290	0
2	B	1077	0	549	34	0
3	D	1119	0	571	39	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
5	A	444	0	0	5	0
5	B	106	0	0	3	0
5	C	69	0	0	0	0
5	D	46	0	0	1	0
All	All	16939	0	15345	452	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 14.

The worst 5 of 452 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:C:876:GLU:HG3	1:C:880:LEU:HD11	1.20	1.14
1:C:137:ARG:NH2	1:C:140:GLU:HG2	1.65	1.10
1:C:342:SER:HB3	1:C:478:MSE:HE3	1.31	1.08
2:B:-30:C:H4'	2:B:-29:A:H5'	1.33	1.07
1:C:263:LYS:HE3	1:C:263:LYS:H	1.15	1.07

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	865/930 (93%)	855 (99%)	10 (1%)	0	100	100
1	C	856/930 (92%)	840 (98%)	15 (2%)	1 (0%)	48	51
All	All	1721/1860 (92%)	1695 (98%)	25 (2%)	1 (0%)	48	51

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	790	ASN

### 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	783/802 (98%)	728 (93%)	55 (7%)	12	8

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	C	777/802 (97%)	671 (86%)	106 (14%)	3 1
All	All	1560/1604 (97%)	1399 (90%)	161 (10%)	6 3

5 of 161 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	555	ASP
1	C	815	ASP
1	C	594	ASP
1	C	710	LYS
1	C	846	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	650	ASN
1	C	882	ASN
1	C	773	ASN
1	C	102	ASN
1	C	466	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	51/51 (100%)	14 (27%)	6 (11%)
3	D	52/53 (98%)	15 (28%)	4 (7%)
All	All	103/104 (99%)	29 (28%)	10 (9%)

5 of 29 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	-29	A
2	B	-22	C
2	B	-18	U
2	B	-17	U
2	B	-16	U

5 of 10 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	D	-17	U
3	D	-4	A
3	D	9	U
2	B	-10	A
2	B	-4	A

#### 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 oligosaccharides 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 [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	844/930 (90%)	0.18	50 (5%) 29 35	13, 29, 71, 179	0
1	C	837/930 (90%)	1.25	170 (20%) 3 5	26, 57, 104, 168	0
2	B	51/51 (100%)	0.22	13 (25%) 2 3	18, 38, 134, 160	0
3	D	53/53 (100%)	0.52	7 (13%) 8 10	30, 51, 161, 193	0
All	All	1785/1964 (90%)	0.69	240 (13%) 8 10	13, 44, 99, 193	0

The worst 5 of 240 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	803	LEU	8.7
1	C	200	ILE	7.6
1	A	922	LYS	6.3
1	C	201	LYS	5.7
1	A	206	TYR	5.4

### 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, 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
4	MG	D	101	1/1	0.98	0.04	36,36,36,36	0
4	MG	B	101	1/1	0.99	0.02	16,16,16,16	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.