



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2026 – 08:24 PM UTC

PDB ID : 5RG5 / pdb_00005rg5
Title : Crystal Structure of Kemp Eliminase HG3.3b in unbound state, 277K
Authors : Broom, A.; Rakotoharisoa, R.V.; Thompson, M.C.; Fraser, J.S.; Chica, R.A.
Deposited on : 2020-03-19
Resolution : 1.62 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

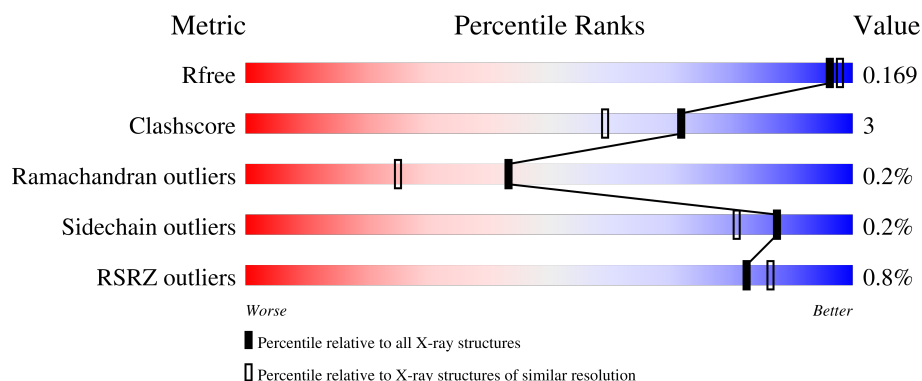
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

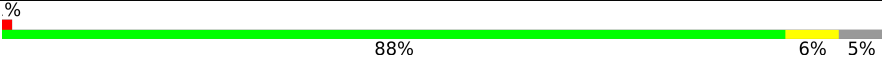
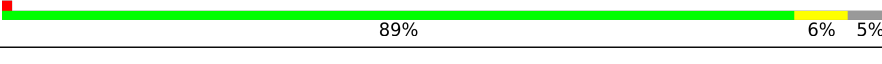
The reported resolution of this entry is 1.62 Å.

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}	180053	6728 (1.64-1.60)
Clashscore	190562	7023 (1.64-1.60)
Ramachandran outliers	187476	6898 (1.64-1.60)
Sidechain outliers	187428	6896 (1.64-1.60)
RSRZ outliers	180081	6727 (1.64-1.60)

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	318	 88% 6% 5%
1	B	318	 89% 6% 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	ACT	A	401	-	-	X	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10907 atoms, of which 5160 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 Kemp Eliminase HG3.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	301	Total	C	H	N	O	S	0	70	0
			5231	1661	2586	466	503	15			
1	B	301	Total	C	H	N	O	S	0	69	0
			5184	1642	2568	463	495	16			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	initiating methionine	UNP P23360
A	0	ALA	-	expression tag	UNP P23360
A	1	GLU	-	expression tag	UNP P23360
A	6	ILE	VAL	engineered mutation	UNP P23360
A	42	MET	GLN	engineered mutation	UNP P23360
A	44	TRP	THR	engineered mutation	UNP P23360
A	50	HIS	LYS	engineered mutation	UNP P23360
A	81	GLY	ARG	engineered mutation	UNP P23360
A	83	GLY	HIS	engineered mutation	UNP P23360
A	84	CYS	THR	engineered mutation	UNP P23360
A	89	ARG	SER	engineered mutation	UNP P23360
A	90	ASP	GLN	engineered mutation	UNP P23360
A	125	ASN	ALA	engineered mutation	UNP P23360
A	130	GLY	ASN	engineered mutation	UNP P23360
A	172	MET	ASN	engineered mutation	UNP P23360
A	234	SER	ALA	engineered mutation	UNP P23360
A	236	LEU	THR	engineered mutation	UNP P23360
A	237	MET	GLU	engineered mutation	UNP P23360
A	267	PHE	TRP	engineered mutation	UNP P23360
A	304	GLY	-	expression tag	UNP P23360
A	305	SER	-	expression tag	UNP P23360
A	306	ILE	-	expression tag	UNP P23360
A	307	GLU	-	expression tag	UNP P23360
A	308	GLY	-	expression tag	UNP P23360
A	309	ARG	-	expression tag	UNP P23360

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	310	GLY	-	expression tag	UNP P23360
A	311	HIS	-	expression tag	UNP P23360
A	312	HIS	-	expression tag	UNP P23360
A	313	HIS	-	expression tag	UNP P23360
A	314	HIS	-	expression tag	UNP P23360
A	315	HIS	-	expression tag	UNP P23360
A	316	HIS	-	expression tag	UNP P23360
B	-1	MET	-	initiating methionine	UNP P23360
B	0	ALA	-	expression tag	UNP P23360
B	1	GLU	-	expression tag	UNP P23360
B	6	ILE	VAL	engineered mutation	UNP P23360
B	42	MET	GLN	engineered mutation	UNP P23360
B	44	TRP	THR	engineered mutation	UNP P23360
B	50	HIS	LYS	engineered mutation	UNP P23360
B	81	GLY	ARG	engineered mutation	UNP P23360
B	83	GLY	HIS	engineered mutation	UNP P23360
B	84	CYS	THR	engineered mutation	UNP P23360
B	89	ARG	SER	engineered mutation	UNP P23360
B	90	ASP	GLN	engineered mutation	UNP P23360
B	125	ASN	ALA	engineered mutation	UNP P23360
B	130	GLY	ASN	engineered mutation	UNP P23360
B	172	MET	ASN	engineered mutation	UNP P23360
B	234	SER	ALA	engineered mutation	UNP P23360
B	236	LEU	THR	engineered mutation	UNP P23360
B	237	MET	GLU	engineered mutation	UNP P23360
B	267	PHE	TRP	engineered mutation	UNP P23360
B	304	GLY	-	expression tag	UNP P23360
B	305	SER	-	expression tag	UNP P23360
B	306	ILE	-	expression tag	UNP P23360
B	307	GLU	-	expression tag	UNP P23360
B	308	GLY	-	expression tag	UNP P23360
B	309	ARG	-	expression tag	UNP P23360
B	310	GLY	-	expression tag	UNP P23360
B	311	HIS	-	expression tag	UNP P23360
B	312	HIS	-	expression tag	UNP P23360
B	313	HIS	-	expression tag	UNP P23360
B	314	HIS	-	expression tag	UNP P23360
B	315	HIS	-	expression tag	UNP P23360
B	316	HIS	-	expression tag	UNP P23360

- Molecule 2 is ACETATE ION (CCD ID: ACT) (formula: C₂H₃O₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			7	2	3	2		
2	B	1	Total	C	H	O	0	0
			7	2	3	2		

- Molecule 3 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	229	Total	O	0	9
			236	236		
4	B	220	Total	O	0	7
			227	227		

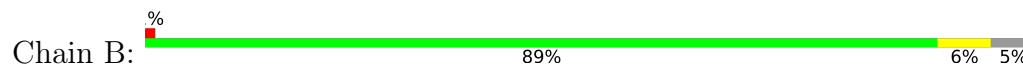
3 Residue-property plots [i](#)

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: Kemp Eliminate HG3



• Molecule 1: Kemp Eliminate HG3



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	76.23Å 80.03Å 98.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.09 – 1.62 37.09 – 1.62	Depositor EDS
% Data completeness (in resolution range)	99.8 (37.09-1.62) 99.8 (37.09-1.62)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.27 (at 1.62Å)	Xtriage
Refinement program	PHENIX 1.14 _3260	Depositor
R, R_{free}	0.142 , 0.167 0.144 , 0.169	Depositor DCC
R_{free} test set	1998 reflections (2.58%)	wwPDB-VP
Wilson B-factor (Å ²)	19.5	Xtriage
Anisotropy	0.404	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 42.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.005 for k,h,-l	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	10907	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.02% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ACT, SO4

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	0.18	0/2952	0.45	0/4032
1	B	0.18	0/2953	0.43	0/4033
All	All	0.18	0/5905	0.44	0/8065

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	2645	2586	2326	19	0
1	B	2616	2568	2294	16	0
2	A	4	3	3	2	0
2	B	4	3	3	1	0
3	A	5	0	0	0	0
3	B	10	0	0	0	0
4	A	236	0	0	5	0
4	B	227	0	0	8	0
All	All	5747	5160	4626	34	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 3.

All (34) 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:50[A]:HIS:ND1	4:A:503[A]:HOH:O	2.13	0.81
1:B:125[B]:ASN:ND2	4:B:510[B]:HOH:O	2.02	0.81
1:A:90[B]:ASP:O	4:A:502:HOH:O	2.09	0.70
1:A:236[B]:LEU:HA	1:A:265[B]:THR:HG22	1.81	0.62
1:A:89[B]:ARG:NE	1:A:131:GLU:OE1	2.35	0.59
1:B:8[B]:GLN:OE1	4:B:501:HOH:O	2.16	0.58
1:A:75[B]:GLN:NE2	4:A:507:HOH:O	2.37	0.57
1:B:276[A]:ARG:NH1	4:B:506[A]:HOH:O	2.36	0.56
1:B:265[B]:THR:HG23	4:B:511:HOH:O	2.06	0.56
1:B:117[B]:THR:HG22	4:B:529:HOH:O	2.06	0.56
1:B:8[B]:GLN:NE2	4:B:508:HOH:O	2.39	0.55
1:A:117[B]:THR:HG22	4:A:514:HOH:O	2.05	0.54
1:B:110:ASN:ND2	4:B:505:HOH:O	2.36	0.51
1:B:236[A]:LEU:HD21	4:B:510[A]:HOH:O	2.11	0.51
1:A:237[B]:MET:H	1:A:265[B]:THR:HG23	1.76	0.50
1:B:236[B]:LEU:HD21	2:B:401:ACT:H2	1.93	0.50
1:A:236[B]:LEU:HD12	1:A:265[B]:THR:HG21	1.93	0.49
1:A:235[A]:ILE:HD13	1:A:251:VAL:HG13	1.96	0.48
1:A:21:ALA:HB3	1:A:267[B]:PHE:O	2.15	0.47
1:A:237[A]:MET:CE	2:A:401:ACT:H3	2.47	0.45
1:B:203:GLY:HA2	1:B:232:GLU:O	2.18	0.44
1:A:235[A]:ILE:HD13	1:A:251:VAL:CG1	2.48	0.43
1:A:303[B]:GLN:NE2	4:A:512:HOH:O	2.52	0.43
1:B:54:THR:O	1:B:61:PHE:HA	2.19	0.42
1:A:203:GLY:HA2	1:A:232:GLU:O	2.19	0.42
1:A:237[A]:MET:HE2	2:A:401:ACT:H3	2.00	0.42
1:B:84[B]:CYS:SG	1:B:87[B]:TRP:HB2	2.60	0.41
1:B:209:HIS:NE2	1:B:237[B]:MET:HE2	2.36	0.41
1:A:273:ASP:OD1	1:B:273:ASP:OD1	2.40	0.40
1:A:21:ALA:HA	1:A:44[B]:TRP:O	2.22	0.40
1:B:21:ALA:HA	1:B:44[B]:TRP:O	2.21	0.40

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	368/318 (116%)	357 (97%)	9 (2%)	2 (0%)	24	10
1	B	367/318 (115%)	357 (97%)	10 (3%)	0	100	100
All	All	735/636 (116%)	714 (97%)	19 (3%)	2 (0%)	43	20

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	91[A]	LEU
1	A	91[B]	LEU

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	307/256 (120%)	307 (100%)	0	100	100
1	B	306/256 (120%)	304 (99%)	2 (1%)	76	62
All	All	613/512 (120%)	611 (100%)	2 (0%)	87	79

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	125[A]	ASN
1	B	125[B]	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	185	GLN
1	B	62	ASN
1	B	74	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

5 ligands are 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	ACT	B	401	-	3,3,3	1.20	0	3,3,3	1.54	0
3	SO4	B	403	-	4,4,4	0.24	0	6,6,6	0.07	0
2	ACT	A	401	-	3,3,3	1.29	0	3,3,3	1.40	0
3	SO4	A	402	-	4,4,4	0.24	0	6,6,6	0.07	0
3	SO4	B	402	-	4,4,4	0.23	0	6,6,6	0.11	0

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.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	401	ACT	1	0
2	A	401	ACT	2	0

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	A	301/318 (94%)	-0.60	3 (0%)	79 83	7, 20, 34, 53	44 (14%)
1	B	301/318 (94%)	-0.54	2 (0%)	84 88	9, 22, 36, 90	41 (13%)
All	All	602/636 (94%)	-0.57	5 (0%)	82 86	7, 21, 35, 90	85 (14%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	3[A]	ALA	6.1
1	A	3	ALA	3.9
1	A	303[A]	GLN	3.1
1	A	90[A]	ASP	2.5
1	B	303	GLN	2.4

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
3	SO4	A	402	5/5	0.69	0.13	97,99,100,100	0
3	SO4	B	403	5/5	0.71	0.15	95,96,99,102	5
3	SO4	B	402	5/5	0.75	0.14	56,58,61,68	5
2	ACT	A	401	4/4	0.80	0.21	41,56,78,78	0
2	ACT	B	401	4/4	0.85	0.20	34,52,63,63	7

6.5 Other polymers [i](#)

There are no such residues in this entry.