



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 30, 2017 – 08:33 PM EDT

PDB ID : 4ZFL
Title : Ergothioneine-biosynthetic Ntn hydrolase variant EgtC_C2A with natural substrate
Authors : Vit, A.; Seebeck, F.P.; Blankenfeldt, W.
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-20029824
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-20029824

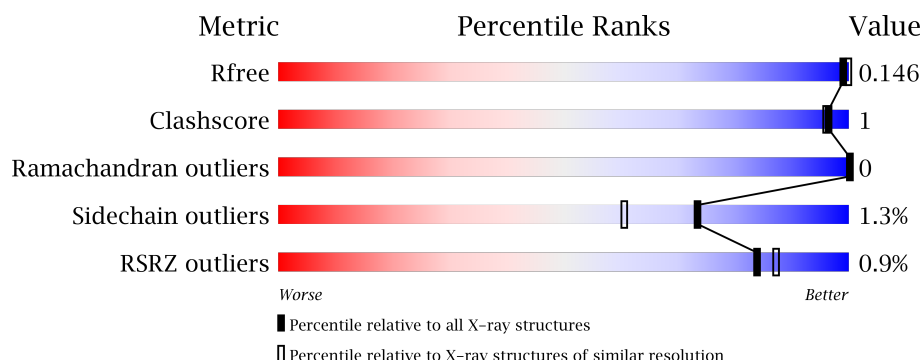
1 Overall quality at a glance

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	A	234	<div> <div>95%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> </div>
1	B	234	<div> <div>%</div> <div>95%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> </div>
1	C	234	<div> <div>%</div> <div>94%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> </div>
1	D	234	<div> <div>2%</div> <div>93%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> </div>
1	E	234	<div> <div>%</div> <div>96%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	234	
1	G	234	
1	H	234	
1	I	234	
1	J	234	
1	K	234	
1	L	234	

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	4NK	B	301	-	-	-	X
2	4NK	C	301	-	-	-	X
2	4NK	D	301	-	-	-	X
2	4NK	F	301	-	-	-	X
2	4NK	G	301	-	-	-	X
2	4NK	H	301	-	-	-	X
2	4NK	J	301	-	-	-	X
2	4NK	K	301	-	-	-	X
2	4NK	L	301	-	-	-	X
3	GOL	E	302	-	-	-	X
3	GOL	G	302	-	-	-	X
3	GOL	I	302	-	-	-	X
3	GOL	K	302	-	-	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 44594 atoms, of which 20537 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 Amidohydrolase EgtC.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	229	Total	C	H	N	O	S	0	7	0
			3434	1094	1701	313	322	4			
1	B	229	Total	C	H	N	O	S	0	4	0
			3393	1079	1679	311	321	3			
1	C	228	Total	C	H	N	O	S	0	7	0
			3396	1082	1680	307	323	4			
1	D	227	Total	C	H	N	O	S	0	2	0
			3343	1067	1656	302	314	4			
1	E	229	Total	C	H	N	O	S	0	4	0
			3380	1079	1669	306	322	4			
1	F	230	Total	C	H	N	O	S	0	4	0
			3436	1092	1703	316	321	4			
1	G	229	Total	C	H	N	O	S	0	7	0
			3387	1078	1677	307	321	4			
1	H	229	Total	C	H	N	O	S	0	3	0
			3419	1086	1695	312	322	4			
1	I	229	Total	C	H	N	O	S	0	4	0
			3420	1086	1696	312	323	3			
1	J	229	Total	C	H	N	O	S	0	4	0
			3419	1088	1698	312	317	4			
1	K	228	Total	C	H	N	O	S	0	3	0
			3375	1073	1671	310	318	3			
1	L	227	Total	C	H	N	O	S	0	3	0
			3369	1072	1668	305	320	4			

There are 156 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	ALA	CYS	engineered mutation	UNP A0R5M9
A	53	ASP	GLU	engineered mutation	UNP A0R5M9
A	84	VAL	LEU	engineered mutation	UNP A0R5M9
A	137	LEU	VAL	engineered mutation	UNP A0R5M9
A	188	ARG	HIS	engineered mutation	UNP A0R5M9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	228	LEU	-	expression tag	UNP A0R5M9
A	229	GLU	-	expression tag	UNP A0R5M9
A	230	HIS	-	expression tag	UNP A0R5M9
A	231	HIS	-	expression tag	UNP A0R5M9
A	232	HIS	-	expression tag	UNP A0R5M9
A	233	HIS	-	expression tag	UNP A0R5M9
A	234	HIS	-	expression tag	UNP A0R5M9
A	235	HIS	-	expression tag	UNP A0R5M9
B	2	ALA	CYS	engineered mutation	UNP A0R5M9
B	53	ASP	GLU	engineered mutation	UNP A0R5M9
B	84	VAL	LEU	engineered mutation	UNP A0R5M9
B	137	LEU	VAL	engineered mutation	UNP A0R5M9
B	188	ARG	HIS	engineered mutation	UNP A0R5M9
B	228	LEU	-	expression tag	UNP A0R5M9
B	229	GLU	-	expression tag	UNP A0R5M9
B	230	HIS	-	expression tag	UNP A0R5M9
B	231	HIS	-	expression tag	UNP A0R5M9
B	232	HIS	-	expression tag	UNP A0R5M9
B	233	HIS	-	expression tag	UNP A0R5M9
B	234	HIS	-	expression tag	UNP A0R5M9
B	235	HIS	-	expression tag	UNP A0R5M9
C	2	ALA	CYS	engineered mutation	UNP A0R5M9
C	53	ASP	GLU	engineered mutation	UNP A0R5M9
C	84	VAL	LEU	engineered mutation	UNP A0R5M9
C	137	LEU	VAL	engineered mutation	UNP A0R5M9
C	188	ARG	HIS	engineered mutation	UNP A0R5M9
C	228	LEU	-	expression tag	UNP A0R5M9
C	229	GLU	-	expression tag	UNP A0R5M9
C	230	HIS	-	expression tag	UNP A0R5M9
C	231	HIS	-	expression tag	UNP A0R5M9
C	232	HIS	-	expression tag	UNP A0R5M9
C	233	HIS	-	expression tag	UNP A0R5M9
C	234	HIS	-	expression tag	UNP A0R5M9
C	235	HIS	-	expression tag	UNP A0R5M9
D	2	ALA	CYS	engineered mutation	UNP A0R5M9
D	53	ASP	GLU	engineered mutation	UNP A0R5M9
D	84	VAL	LEU	engineered mutation	UNP A0R5M9
D	137	LEU	VAL	engineered mutation	UNP A0R5M9
D	188	ARG	HIS	engineered mutation	UNP A0R5M9
D	228	LEU	-	expression tag	UNP A0R5M9
D	229	GLU	-	expression tag	UNP A0R5M9
D	230	HIS	-	expression tag	UNP A0R5M9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	231	HIS	-	expression tag	UNP A0R5M9
D	232	HIS	-	expression tag	UNP A0R5M9
D	233	HIS	-	expression tag	UNP A0R5M9
D	234	HIS	-	expression tag	UNP A0R5M9
D	235	HIS	-	expression tag	UNP A0R5M9
E	2	ALA	CYS	engineered mutation	UNP A0R5M9
E	53	ASP	GLU	engineered mutation	UNP A0R5M9
E	84	VAL	LEU	engineered mutation	UNP A0R5M9
E	137	LEU	VAL	engineered mutation	UNP A0R5M9
E	188	ARG	HIS	engineered mutation	UNP A0R5M9
E	228	LEU	-	expression tag	UNP A0R5M9
E	229	GLU	-	expression tag	UNP A0R5M9
E	230	HIS	-	expression tag	UNP A0R5M9
E	231	HIS	-	expression tag	UNP A0R5M9
E	232	HIS	-	expression tag	UNP A0R5M9
E	233	HIS	-	expression tag	UNP A0R5M9
E	234	HIS	-	expression tag	UNP A0R5M9
E	235	HIS	-	expression tag	UNP A0R5M9
F	2	ALA	CYS	engineered mutation	UNP A0R5M9
F	53	ASP	GLU	engineered mutation	UNP A0R5M9
F	84	VAL	LEU	engineered mutation	UNP A0R5M9
F	137	LEU	VAL	engineered mutation	UNP A0R5M9
F	188	ARG	HIS	engineered mutation	UNP A0R5M9
F	228	LEU	-	expression tag	UNP A0R5M9
F	229	GLU	-	expression tag	UNP A0R5M9
F	230	HIS	-	expression tag	UNP A0R5M9
F	231	HIS	-	expression tag	UNP A0R5M9
F	232	HIS	-	expression tag	UNP A0R5M9
F	233	HIS	-	expression tag	UNP A0R5M9
F	234	HIS	-	expression tag	UNP A0R5M9
F	235	HIS	-	expression tag	UNP A0R5M9
G	2	ALA	CYS	engineered mutation	UNP A0R5M9
G	53	ASP	GLU	engineered mutation	UNP A0R5M9
G	84	VAL	LEU	engineered mutation	UNP A0R5M9
G	137	LEU	VAL	engineered mutation	UNP A0R5M9
G	188	ARG	HIS	engineered mutation	UNP A0R5M9
G	228	LEU	-	expression tag	UNP A0R5M9
G	229	GLU	-	expression tag	UNP A0R5M9
G	230	HIS	-	expression tag	UNP A0R5M9
G	231	HIS	-	expression tag	UNP A0R5M9
G	232	HIS	-	expression tag	UNP A0R5M9
G	233	HIS	-	expression tag	UNP A0R5M9

Continued on next page...

Continued from previous page...

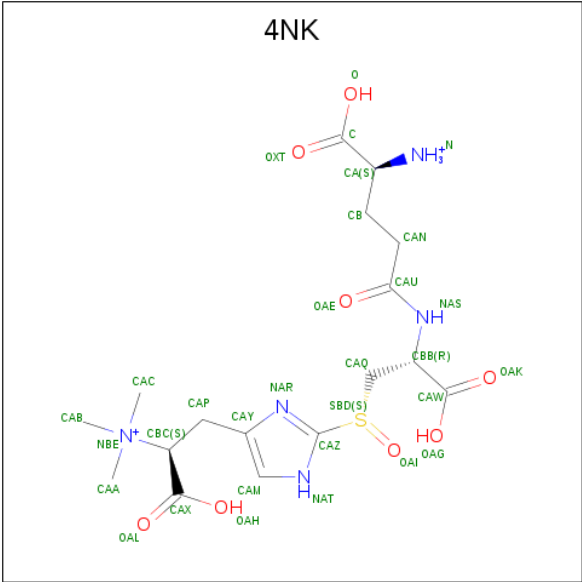
Chain	Residue	Modelled	Actual	Comment	Reference
G	234	HIS	-	expression tag	UNP A0R5M9
G	235	HIS	-	expression tag	UNP A0R5M9
H	2	ALA	CYS	engineered mutation	UNP A0R5M9
H	53	ASP	GLU	engineered mutation	UNP A0R5M9
H	84	VAL	LEU	engineered mutation	UNP A0R5M9
H	137	LEU	VAL	engineered mutation	UNP A0R5M9
H	188	ARG	HIS	engineered mutation	UNP A0R5M9
H	228	LEU	-	expression tag	UNP A0R5M9
H	229	GLU	-	expression tag	UNP A0R5M9
H	230	HIS	-	expression tag	UNP A0R5M9
H	231	HIS	-	expression tag	UNP A0R5M9
H	232	HIS	-	expression tag	UNP A0R5M9
H	233	HIS	-	expression tag	UNP A0R5M9
H	234	HIS	-	expression tag	UNP A0R5M9
H	235	HIS	-	expression tag	UNP A0R5M9
I	2	ALA	CYS	engineered mutation	UNP A0R5M9
I	53	ASP	GLU	engineered mutation	UNP A0R5M9
I	84	VAL	LEU	engineered mutation	UNP A0R5M9
I	137	LEU	VAL	engineered mutation	UNP A0R5M9
I	188	ARG	HIS	engineered mutation	UNP A0R5M9
I	228	LEU	-	expression tag	UNP A0R5M9
I	229	GLU	-	expression tag	UNP A0R5M9
I	230	HIS	-	expression tag	UNP A0R5M9
I	231	HIS	-	expression tag	UNP A0R5M9
I	232	HIS	-	expression tag	UNP A0R5M9
I	233	HIS	-	expression tag	UNP A0R5M9
I	234	HIS	-	expression tag	UNP A0R5M9
I	235	HIS	-	expression tag	UNP A0R5M9
J	2	ALA	CYS	engineered mutation	UNP A0R5M9
J	53	ASP	GLU	engineered mutation	UNP A0R5M9
J	84	VAL	LEU	engineered mutation	UNP A0R5M9
J	137	LEU	VAL	engineered mutation	UNP A0R5M9
J	188	ARG	HIS	engineered mutation	UNP A0R5M9
J	228	LEU	-	expression tag	UNP A0R5M9
J	229	GLU	-	expression tag	UNP A0R5M9
J	230	HIS	-	expression tag	UNP A0R5M9
J	231	HIS	-	expression tag	UNP A0R5M9
J	232	HIS	-	expression tag	UNP A0R5M9
J	233	HIS	-	expression tag	UNP A0R5M9
J	234	HIS	-	expression tag	UNP A0R5M9
J	235	HIS	-	expression tag	UNP A0R5M9
K	2	ALA	CYS	engineered mutation	UNP A0R5M9

Continued on next page...

Continued from previous page...

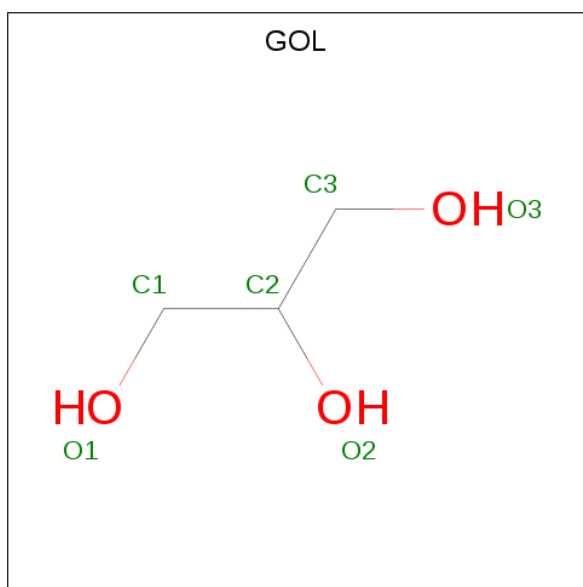
Chain	Residue	Modelled	Actual	Comment	Reference
K	53	ASP	GLU	engineered mutation	UNP A0R5M9
K	84	VAL	LEU	engineered mutation	UNP A0R5M9
K	137	LEU	VAL	engineered mutation	UNP A0R5M9
K	188	ARG	HIS	engineered mutation	UNP A0R5M9
K	228	LEU	-	expression tag	UNP A0R5M9
K	229	GLU	-	expression tag	UNP A0R5M9
K	230	HIS	-	expression tag	UNP A0R5M9
K	231	HIS	-	expression tag	UNP A0R5M9
K	232	HIS	-	expression tag	UNP A0R5M9
K	233	HIS	-	expression tag	UNP A0R5M9
K	234	HIS	-	expression tag	UNP A0R5M9
K	235	HIS	-	expression tag	UNP A0R5M9
L	2	ALA	CYS	engineered mutation	UNP A0R5M9
L	53	ASP	GLU	engineered mutation	UNP A0R5M9
L	84	VAL	LEU	engineered mutation	UNP A0R5M9
L	137	LEU	VAL	engineered mutation	UNP A0R5M9
L	188	ARG	HIS	engineered mutation	UNP A0R5M9
L	228	LEU	-	expression tag	UNP A0R5M9
L	229	GLU	-	expression tag	UNP A0R5M9
L	230	HIS	-	expression tag	UNP A0R5M9
L	231	HIS	-	expression tag	UNP A0R5M9
L	232	HIS	-	expression tag	UNP A0R5M9
L	233	HIS	-	expression tag	UNP A0R5M9
L	234	HIS	-	expression tag	UNP A0R5M9
L	235	HIS	-	expression tag	UNP A0R5M9

- Molecule 2 is (1S)-1-carboxy-4-({(1R)-1-carboxy-2-[(S)-{4-[(2S)-2-carboxy-2-(trimethylammonio)ethyl]-1H-imidazol-2-yl}sulfinyl]ethyl}amino)-4-oxobutan-1-aminium (three-letter code: 4NK) (formula: C₁₇H₂₉N₅O₈S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	S	0	0
			57	17	26	5	8	1		
2	B	1	Total	C	H	N	O	S	0	0
			57	17	26	5	8	1		
2	C	1	Total	C	H	N	O	S	0	0
			57	17	26	5	8	1		
2	D	1	Total	C	H	N	O	S	0	0
			57	17	26	5	8	1		
2	E	1	Total	C	H	N	O	S	0	0
			57	17	26	5	8	1		
2	F	1	Total	C	H	N	O	S	0	0
			57	17	26	5	8	1		
2	G	1	Total	C	H	N	O	S	0	0
			57	17	26	5	8	1		
2	H	1	Total	C	H	N	O	S	0	0
			57	17	26	5	8	1		
2	I	1	Total	C	H	N	O	S	0	0
			57	17	26	5	8	1		
2	J	1	Total	C	H	N	O	S	0	0
			57	17	26	5	8	1		
2	K	1	Total	C	H	N	O	S	0	0
			57	17	26	5	8	1		
2	L	1	Total	C	H	N	O	S	0	0
			57	17	26	5	8	1		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	E	1	Total	C	H	O	0	0
			14	3	8	3		
3	G	1	Total	C	H	O	0	0
			14	3	8	3		
3	I	1	Total	C	H	O	0	0
			14	3	8	3		
3	K	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	298	Total	O	0	0
			298	298		
4	B	280	Total	O	0	0
			280	280		
4	C	245	Total	O	0	0
			245	245		
4	D	232	Total	O	0	0
			232	232		
4	E	250	Total	O	0	0
			250	250		
4	F	266	Total	O	0	0
			266	266		
4	G	230	Total	O	0	0
			230	230		
4	H	266	Total	O	0	0
			266	266		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	I	268	Total 268	O 268	0	0
4	J	258	Total 258	O 258	0	0
4	K	246	Total 246	O 246	0	0
4	L	244	Total 244	O 244	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.

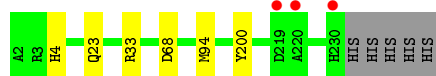
- Molecule 1: Amidohydrolase EgtC

Chain A:  95%



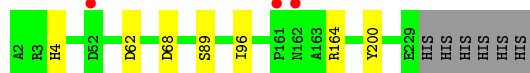
- Molecule 1: Amidohydrolase EgtC

Chain B:  95%



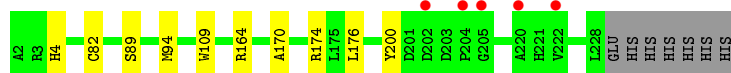
- Molecule 1: Amidohydrolase EgtC

Chain C:  94%



- Molecule 1: Amidohydrolase EgtC

Chain D:  93%



- Molecule 1: Amidohydrolase EgtC

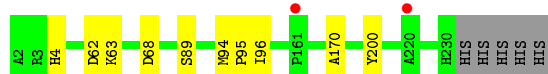
Chain E:  96%



- Molecule 1: Amidohydrolase EgtC



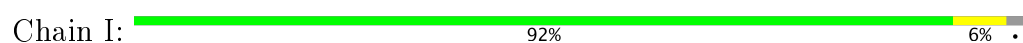
- Molecule 1: Amidohydrolase EgtC



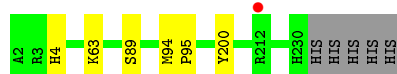
- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



- Molecule 1: Amidohydrolase EgtC



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	128.47Å 69.36Å 159.41Å 90.00° 94.67° 90.00°	Depositor
Resolution (Å)	49.31 – 1.70 49.31 – 1.70	Depositor EDS
% Data completeness (in resolution range)	97.6 (49.31-1.70) 97.6 (49.31-1.70)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.37 (at 1.70Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.144 , 0.171 0.144 , 0.146	Depositor DCC
R_{free} test set	14992 reflections (5.00%)	DCC
Wilson B-factor (Å ²)	19.8	Xtriage
Anisotropy	0.130	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 45.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	44594	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 45.57 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.2958e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² 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: GOL, 4NK

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	0.47	0/1808	0.67	1/2472 (0.0%)
1	B	0.45	0/1772	0.62	0/2427
1	C	0.42	0/1788	0.59	0/2449
1	D	0.40	0/1734	0.58	0/2377
1	E	0.41	0/1769	0.62	0/2424
1	F	0.42	0/1787	0.59	0/2447
1	G	0.41	0/1784	0.59	0/2443
1	H	0.43	0/1775	0.61	0/2431
1	I	0.43	0/1782	0.63	1/2440 (0.0%)
1	J	0.41	0/1775	0.58	0/2431
1	K	0.42	0/1751	0.60	0/2398
1	L	0.42	0/1751	0.62	0/2399
All	All	0.42	0/21276	0.61	2/29138 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	133	ASP	CB-CG-OD1	6.05	123.75	118.30
1	I	133	ASP	CB-CG-OD1	5.05	122.84	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	A	1733	1701	1687	3	0
1	B	1714	1679	1667	3	0
1	C	1716	1680	1657	4	0
1	D	1687	1656	1657	7	0
1	E	1711	1669	1658	2	0
1	F	1733	1703	1709	4	0
1	G	1710	1677	1645	7	0
1	H	1724	1695	1698	4	0
1	I	1724	1696	1683	6	0
1	J	1721	1698	1703	3	0
1	K	1704	1671	1667	2	0
1	L	1701	1668	1671	3	0
2	A	31	26	26	0	0
2	B	31	26	26	1	0
2	C	31	26	26	2	0
2	D	31	26	26	2	0
2	E	31	26	26	0	0
2	F	31	26	26	2	0
2	G	31	26	26	2	0
2	H	31	26	26	1	0
2	I	31	26	26	0	0
2	J	31	26	26	2	0
2	K	31	26	26	0	0
2	L	31	26	26	2	0
3	E	6	8	8	0	0
3	G	6	8	8	0	0
3	I	6	8	8	0	0
3	K	6	8	8	0	0
4	A	298	0	0	2	0
4	B	280	0	0	1	0
4	C	245	0	0	1	0
4	D	232	0	0	1	0
4	E	250	0	0	1	0
4	F	266	0	0	2	0
4	G	230	0	0	1	0
4	H	266	0	0	2	1
4	I	268	0	0	1	1
4	J	258	0	0	2	0
4	K	246	0	0	1	0
4	L	244	0	0	1	0
All	All	24057	20537	20446	49	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 1.

The worst 5 of 49 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:H:188:ARG:NH2	1:H:220:ALA:O	2.12	0.83
1:J:63[A]:LYS:NZ	4:J:401:HOH:O	2.12	0.82
1:F:89[B]:SER:OG	2:F:301:4NK:OAI	1.96	0.81
1:G:89[B]:SER:OG	2:G:301:4NK:OAI	2.01	0.78
1:L:164:ARG:NH2	2:L:301:4NK:OAK	2.20	0.73

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)
4:H:538:HOH:O	4:I:581:HOH:O[2_545]	1.89	0.31

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	234/234 (100%)	228 (97%)	6 (3%)	0	100	100
1	B	231/234 (99%)	224 (97%)	7 (3%)	0	100	100
1	C	233/234 (100%)	227 (97%)	6 (3%)	0	100	100
1	D	227/234 (97%)	221 (97%)	6 (3%)	0	100	100
1	E	231/234 (99%)	225 (97%)	6 (3%)	0	100	100
1	F	232/234 (99%)	226 (97%)	6 (3%)	0	100	100
1	G	234/234 (100%)	228 (97%)	6 (3%)	0	100	100
1	H	230/234 (98%)	223 (97%)	7 (3%)	0	100	100
1	I	231/234 (99%)	224 (97%)	7 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	J	231/234 (99%)	224 (97%)	7 (3%)	0	100	100
1	K	228/234 (97%)	222 (97%)	6 (3%)	0	100	100
1	L	228/234 (97%)	222 (97%)	6 (3%)	0	100	100
All	All	2770/2808 (99%)	2694 (97%)	76 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	180/183 (98%)	178 (99%)	2 (1%)	78	68
1	B	177/183 (97%)	175 (99%)	2 (1%)	78	68
1	C	179/183 (98%)	177 (99%)	2 (1%)	78	68
1	D	173/183 (94%)	171 (99%)	2 (1%)	75	64
1	E	176/183 (96%)	174 (99%)	2 (1%)	78	68
1	F	179/183 (98%)	177 (99%)	2 (1%)	78	68
1	G	178/183 (97%)	176 (99%)	2 (1%)	78	68
1	H	179/183 (98%)	177 (99%)	2 (1%)	78	68
1	I	179/183 (98%)	176 (98%)	3 (2%)	66	50
1	J	177/183 (97%)	175 (99%)	2 (1%)	78	68
1	K	174/183 (95%)	171 (98%)	3 (2%)	66	50
1	L	176/183 (96%)	174 (99%)	2 (1%)	78	68
All	All	2127/2196 (97%)	2101 (99%)	26 (1%)	73	64

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	200	TYR
1	H	4	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	L	4	HIS
1	G	4	HIS
1	G	200	TYR

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](#)

16 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	4NK	A	301	-	18,31,31	1.42	2 (11%)	17,44,44	2.23	5 (29%)
2	4NK	B	301	-	18,31,31	1.37	2 (11%)	17,44,44	2.83	5 (29%)
2	4NK	C	301	-	18,31,31	1.24	0	17,44,44	2.83	5 (29%)
2	4NK	D	301	-	18,31,31	1.28	1 (5%)	17,44,44	2.92	5 (29%)
2	4NK	E	301	-	18,31,31	1.37	3 (16%)	17,44,44	2.34	5 (29%)
3	GOL	E	302	-	5,5,5	0.29	0	5,5,5	0.47	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	4NK	F	301	-	18,31,31	1.34	2 (11%)	17,44,44	3.12	6 (35%)
2	4NK	G	301	-	18,31,31	1.26	1 (5%)	17,44,44	2.60	5 (29%)
3	GOL	G	302	-	5,5,5	0.34	0	5,5,5	0.30	0
2	4NK	H	301	-	18,31,31	1.27	1 (5%)	17,44,44	2.37	5 (29%)
2	4NK	I	301	-	18,31,31	1.34	3 (16%)	17,44,44	2.14	3 (17%)
3	GOL	I	302	-	5,5,5	0.60	0	5,5,5	0.78	0
2	4NK	J	301	-	18,31,31	1.36	3 (16%)	17,44,44	2.46	3 (17%)
2	4NK	K	301	-	18,31,31	1.21	2 (11%)	17,44,44	2.45	3 (17%)
3	GOL	K	302	-	5,5,5	0.46	0	5,5,5	0.24	0
2	4NK	L	301	-	18,31,31	1.24	2 (11%)	17,44,44	3.05	3 (17%)

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	4NK	A	301	-	-	0/23/39/39	0/1/1/1
2	4NK	B	301	-	-	0/23/39/39	0/1/1/1
2	4NK	C	301	-	-	0/23/39/39	0/1/1/1
2	4NK	D	301	-	-	0/23/39/39	0/1/1/1
2	4NK	E	301	-	-	0/23/39/39	0/1/1/1
3	GOL	E	302	-	-	0/4/4/4	0/0/0/0
2	4NK	F	301	-	-	0/23/39/39	0/1/1/1
2	4NK	G	301	-	-	0/23/39/39	0/1/1/1
3	GOL	G	302	-	-	0/4/4/4	0/0/0/0
2	4NK	H	301	-	-	0/23/39/39	0/1/1/1
2	4NK	I	301	-	-	0/23/39/39	0/1/1/1
3	GOL	I	302	-	-	0/4/4/4	0/0/0/0
2	4NK	J	301	-	-	0/23/39/39	0/1/1/1
2	4NK	K	301	-	-	0/23/39/39	0/1/1/1
3	GOL	K	302	-	-	0/4/4/4	0/0/0/0
2	4NK	L	301	-	-	0/23/39/39	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	301	4NK	CBC-NBE	-2.66	1.49	1.54
2	L	301	4NK	CBC-NBE	-2.65	1.49	1.54
2	E	301	4NK	CBC-NBE	-2.53	1.49	1.54
2	F	301	4NK	CBC-NBE	-2.41	1.49	1.54

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	301	4NK	CBC-NBE	-2.38	1.49	1.54

The worst 5 of 53 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	301	4NK	CB-CA-N	-3.49	105.69	110.78
2	G	301	4NK	CB-CAN-CAU	-3.44	105.40	113.18
2	F	301	4NK	CB-CAN-CAU	-3.40	105.49	113.18
2	E	301	4NK	CB-CAN-CAU	-3.37	105.56	113.18
2	E	301	4NK	CB-CA-N	-3.21	106.10	110.78

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	301	4NK	1	0
2	C	301	4NK	2	0
2	D	301	4NK	2	0
2	F	301	4NK	2	0
2	G	301	4NK	2	0
2	H	301	4NK	1	0
2	J	301	4NK	2	0
2	L	301	4NK	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 ⓘ

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	229/234 (97%)	-0.35	1 (0%) 92 93	14, 18, 29, 48	0
1	B	229/234 (97%)	-0.34	3 (1%) 77 82	15, 20, 36, 54	0
1	C	228/234 (97%)	-0.35	3 (1%) 77 82	15, 22, 34, 48	0
1	D	227/234 (97%)	-0.10	5 (2%) 62 68	15, 24, 40, 61	0
1	E	229/234 (97%)	-0.27	3 (1%) 77 82	15, 22, 39, 59	0
1	F	230/234 (98%)	-0.28	2 (0%) 84 87	13, 21, 35, 70	0
1	G	229/234 (97%)	-0.28	2 (0%) 84 87	15, 22, 41, 62	0
1	H	229/234 (97%)	-0.41	2 (0%) 84 87	14, 20, 36, 52	0
1	I	229/234 (97%)	-0.33	1 (0%) 92 93	15, 21, 36, 53	0
1	J	229/234 (97%)	-0.30	1 (0%) 92 93	15, 23, 44, 54	0
1	K	228/234 (97%)	-0.25	1 (0%) 92 93	16, 22, 39, 57	0
1	L	227/234 (97%)	-0.34	1 (0%) 92 93	15, 22, 35, 56	0
All	All	2743/2808 (97%)	-0.30	25 (0%) 84 87	13, 21, 38, 70	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	204	PRO	5.1
1	L	228	LEU	4.8
1	E	220	ALA	4.2
1	D	220	ALA	3.9
1	J	212	ARG	3.5

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 carbohydrates 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. 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	4NK	D	301	31/31	0.86	0.21	7.21	26,40,78,78	38
2	4NK	F	301	31/31	0.89	0.19	4.41	19,36,70,74	33
3	GOL	E	302	6/6	0.84	0.14	3.92	23,38,45,46	0
2	4NK	L	301	31/31	0.83	0.17	3.80	26,46,71,75	0
3	GOL	G	302	6/6	0.88	0.14	3.61	33,47,57,58	0
2	4NK	J	301	31/31	0.89	0.15	3.54	25,41,68,78	0
2	4NK	H	301	31/31	0.92	0.12	3.33	22,33,57,60	16
2	4NK	K	301	31/31	0.87	0.13	3.20	27,42,58,58	0
2	4NK	G	301	31/31	0.92	0.12	2.84	24,37,55,58	24
3	GOL	I	302	6/6	0.92	0.16	2.63	25,33,40,40	0
2	4NK	C	301	31/31	0.90	0.15	2.56	21,32,71,75	36
2	4NK	B	301	31/31	0.91	0.14	2.49	22,34,60,66	24
3	GOL	K	302	6/6	0.84	0.16	2.17	31,47,57,57	0
2	4NK	E	301	31/31	0.94	0.10	1.44	25,33,49,51	21
2	4NK	A	301	31/31	0.93	0.10	1.15	24,32,51,56	10
2	4NK	I	301	31/31	0.92	0.10	0.67	21,30,38,40	0

6.5 Other polymers ⓘ

There are no such residues in this entry.