



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 18, 2021 – 10:02 AM EST

PDB ID : 7LS0
Title : Structure of the Human ALK GRD bound to AUG
Authors : Stayrook, S.; Li, T.; Klein, D.E.
Deposited on : 2021-02-17
Resolution : 3.05 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.23.2
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) : 2.23.2

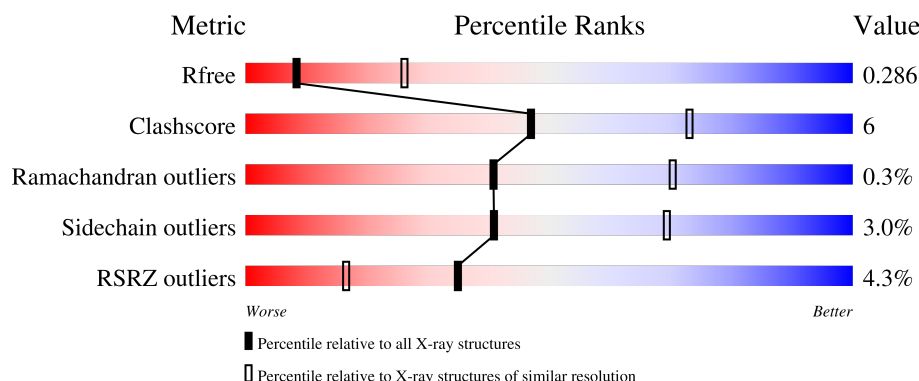
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.05 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	1754 (3.10-3.02)
Clashscore	141614	1864 (3.10-3.02)
Ramachandran outliers	138981	1794 (3.10-3.02)
Sidechain outliers	138945	1793 (3.10-3.02)
RSRZ outliers	127900	1713 (3.10-3.02)

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	451	<div> <div>70%</div> <div>10%</div> <div>20%</div> </div>
1	B	451	<div> <div>3%</div> <div>68%</div> <div>12%</div> <div>•</div> <div>20%</div> </div>
1	C	451	<div> <div>7%</div> <div>67%</div> <div>13%</div> <div>•</div> <div>20%</div> </div>
1	D	451	<div> <div>3%</div> <div>70%</div> <div>10%</div> <div>20%</div> </div>

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	CIT	A	1201	-	-	-	X
2	CIT	A	1202	-	-	-	X

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10672 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 ALK tyrosine kinase receptor fused with ALK and LTK ligand 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	361	Total	C	N	O	S	0	0	0
			2658	1656	479	505	18			
1	B	362	Total	C	N	O	S	0	0	0
			2658	1654	479	507	18			
1	C	362	Total	C	N	O	S	0	0	0
			2662	1658	478	508	18			
1	D	361	Total	C	N	O	S	0	0	0
			2654	1653	477	507	17			

There are 120 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	663	GLY	-	expression tag	UNP Q9UM73
A	664	HIS	-	expression tag	UNP Q9UM73
A	665	HIS	-	expression tag	UNP Q9UM73
A	666	HIS	-	expression tag	UNP Q9UM73
A	667	HIS	-	expression tag	UNP Q9UM73
A	668	HIS	-	expression tag	UNP Q9UM73
A	669	HIS	-	expression tag	UNP Q9UM73
A	670	HIS	-	expression tag	UNP Q9UM73
A	671	HIS	-	expression tag	UNP Q9UM73
A	672	ILE	-	expression tag	UNP Q9UM73
A	673	GLU	-	expression tag	UNP Q9UM73
A	674	GLY	-	expression tag	UNP Q9UM73
A	675	ARG	-	expression tag	UNP Q9UM73
A	676	SER	-	expression tag	UNP Q9UM73
A	677	LEU	-	expression tag	UNP Q9UM73
A	1031	GLY	-	linker	UNP Q9UM73
A	1032	GLY	-	linker	UNP Q9UM73
A	1033	SER	-	linker	UNP Q9UM73
A	1034	GLY	-	linker	UNP Q9UM73
A	1035	GLY	-	linker	UNP Q9UM73

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1036	SER	-	linker	UNP Q9UM73
A	1037	GLY	-	linker	UNP Q9UM73
A	1038	GLY	-	linker	UNP Q9UM73
A	1039	SER	-	linker	UNP Q9UM73
A	1040	GLY	-	linker	UNP Q9UM73
A	1041	GLY	-	linker	UNP Q9UM73
A	1042	SER	-	linker	UNP Q9UM73
A	1043	GLY	-	linker	UNP Q9UM73
A	1044	GLY	-	linker	UNP Q9UM73
A	1045	ALA	-	linker	UNP Q9UM73
B	663	GLY	-	expression tag	UNP Q9UM73
B	664	HIS	-	expression tag	UNP Q9UM73
B	665	HIS	-	expression tag	UNP Q9UM73
B	666	HIS	-	expression tag	UNP Q9UM73
B	667	HIS	-	expression tag	UNP Q9UM73
B	668	HIS	-	expression tag	UNP Q9UM73
B	669	HIS	-	expression tag	UNP Q9UM73
B	670	HIS	-	expression tag	UNP Q9UM73
B	671	HIS	-	expression tag	UNP Q9UM73
B	672	ILE	-	expression tag	UNP Q9UM73
B	673	GLU	-	expression tag	UNP Q9UM73
B	674	GLY	-	expression tag	UNP Q9UM73
B	675	ARG	-	expression tag	UNP Q9UM73
B	676	SER	-	expression tag	UNP Q9UM73
B	677	LEU	-	expression tag	UNP Q9UM73
B	1031	GLY	-	linker	UNP Q9UM73
B	1032	GLY	-	linker	UNP Q9UM73
B	1033	SER	-	linker	UNP Q9UM73
B	1034	GLY	-	linker	UNP Q9UM73
B	1035	GLY	-	linker	UNP Q9UM73
B	1036	SER	-	linker	UNP Q9UM73
B	1037	GLY	-	linker	UNP Q9UM73
B	1038	GLY	-	linker	UNP Q9UM73
B	1039	SER	-	linker	UNP Q9UM73
B	1040	GLY	-	linker	UNP Q9UM73
B	1041	GLY	-	linker	UNP Q9UM73
B	1042	SER	-	linker	UNP Q9UM73
B	1043	GLY	-	linker	UNP Q9UM73
B	1044	GLY	-	linker	UNP Q9UM73
B	1045	ALA	-	linker	UNP Q9UM73
C	663	GLY	-	expression tag	UNP Q9UM73
C	664	HIS	-	expression tag	UNP Q9UM73

Continued on next page...

Continued from previous page...

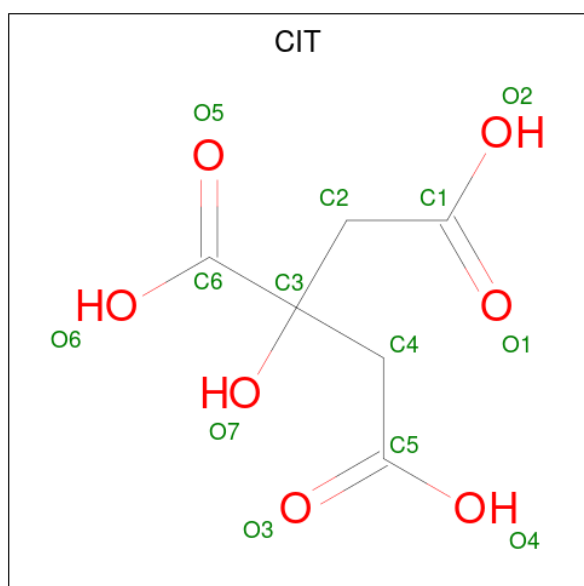
Chain	Residue	Modelled	Actual	Comment	Reference
C	665	HIS	-	expression tag	UNP Q9UM73
C	666	HIS	-	expression tag	UNP Q9UM73
C	667	HIS	-	expression tag	UNP Q9UM73
C	668	HIS	-	expression tag	UNP Q9UM73
C	669	HIS	-	expression tag	UNP Q9UM73
C	670	HIS	-	expression tag	UNP Q9UM73
C	671	HIS	-	expression tag	UNP Q9UM73
C	672	ILE	-	expression tag	UNP Q9UM73
C	673	GLU	-	expression tag	UNP Q9UM73
C	674	GLY	-	expression tag	UNP Q9UM73
C	675	ARG	-	expression tag	UNP Q9UM73
C	676	SER	-	expression tag	UNP Q9UM73
C	677	LEU	-	expression tag	UNP Q9UM73
C	1031	GLY	-	linker	UNP Q9UM73
C	1032	GLY	-	linker	UNP Q9UM73
C	1033	SER	-	linker	UNP Q9UM73
C	1034	GLY	-	linker	UNP Q9UM73
C	1035	GLY	-	linker	UNP Q9UM73
C	1036	SER	-	linker	UNP Q9UM73
C	1037	GLY	-	linker	UNP Q9UM73
C	1038	GLY	-	linker	UNP Q9UM73
C	1039	SER	-	linker	UNP Q9UM73
C	1040	GLY	-	linker	UNP Q9UM73
C	1041	GLY	-	linker	UNP Q9UM73
C	1042	SER	-	linker	UNP Q9UM73
C	1043	GLY	-	linker	UNP Q9UM73
C	1044	GLY	-	linker	UNP Q9UM73
C	1045	ALA	-	linker	UNP Q9UM73
D	663	GLY	-	expression tag	UNP Q9UM73
D	664	HIS	-	expression tag	UNP Q9UM73
D	665	HIS	-	expression tag	UNP Q9UM73
D	666	HIS	-	expression tag	UNP Q9UM73
D	667	HIS	-	expression tag	UNP Q9UM73
D	668	HIS	-	expression tag	UNP Q9UM73
D	669	HIS	-	expression tag	UNP Q9UM73
D	670	HIS	-	expression tag	UNP Q9UM73
D	671	HIS	-	expression tag	UNP Q9UM73
D	672	ILE	-	expression tag	UNP Q9UM73
D	673	GLU	-	expression tag	UNP Q9UM73
D	674	GLY	-	expression tag	UNP Q9UM73
D	675	ARG	-	expression tag	UNP Q9UM73
D	676	SER	-	expression tag	UNP Q9UM73

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	677	LEU	-	expression tag	UNP Q9UM73
D	1031	GLY	-	linker	UNP Q9UM73
D	1032	GLY	-	linker	UNP Q9UM73
D	1033	SER	-	linker	UNP Q9UM73
D	1034	GLY	-	linker	UNP Q9UM73
D	1035	GLY	-	linker	UNP Q9UM73
D	1036	SER	-	linker	UNP Q9UM73
D	1037	GLY	-	linker	UNP Q9UM73
D	1038	GLY	-	linker	UNP Q9UM73
D	1039	SER	-	linker	UNP Q9UM73
D	1040	GLY	-	linker	UNP Q9UM73
D	1041	GLY	-	linker	UNP Q9UM73
D	1042	SER	-	linker	UNP Q9UM73
D	1043	GLY	-	linker	UNP Q9UM73
D	1044	GLY	-	linker	UNP Q9UM73
D	1045	ALA	-	linker	UNP Q9UM73

- Molecule 2 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			13	6	7		
2	A	1	Total	C	O	0	0
			13	6	7		

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

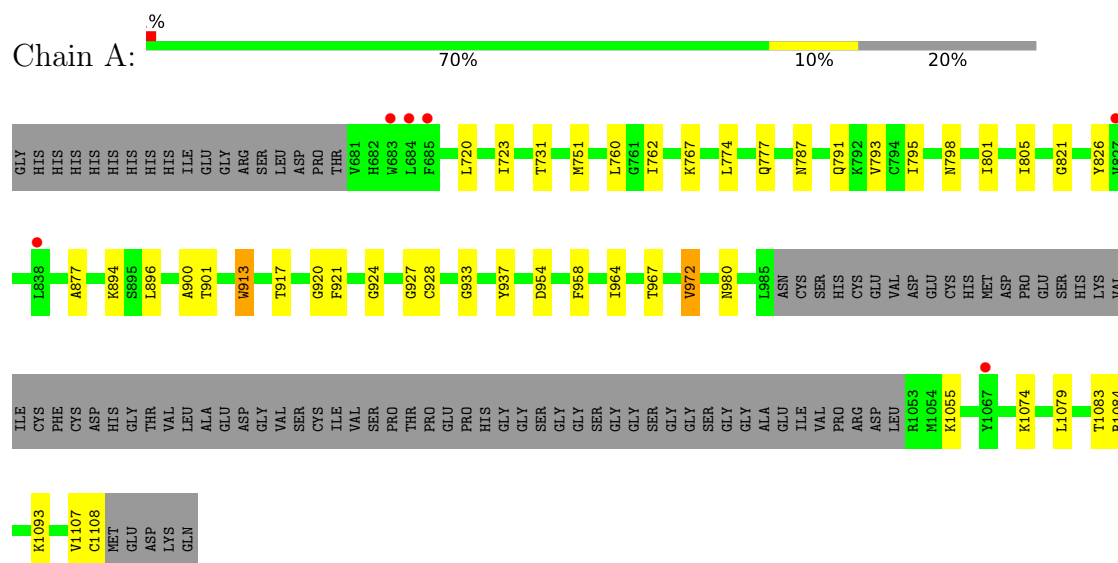


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	D	1	14	8	1	5	0	0

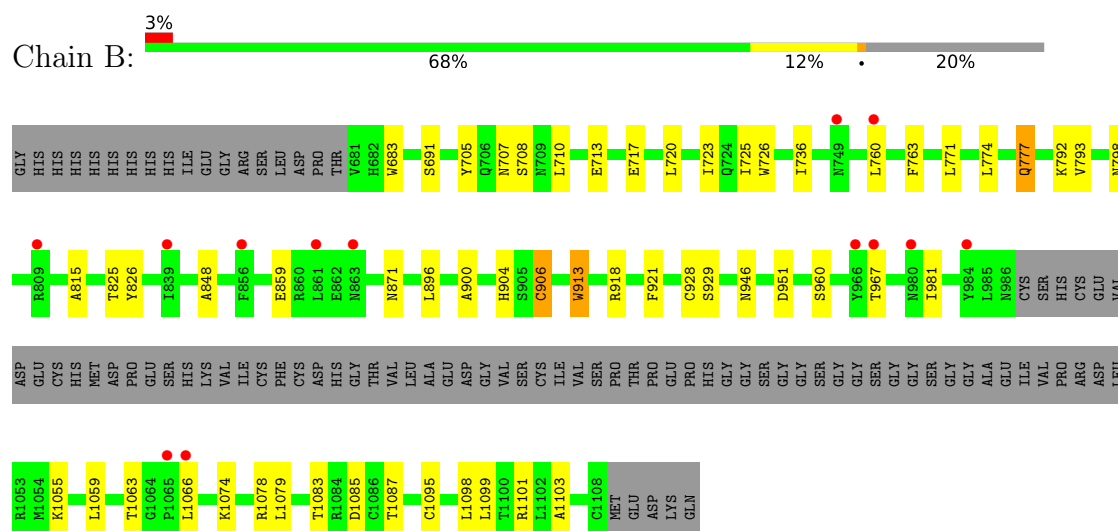
3 Residue-property plots [i](#)

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: ALK tyrosine kinase receptor fused with ALK and LTK ligand 2

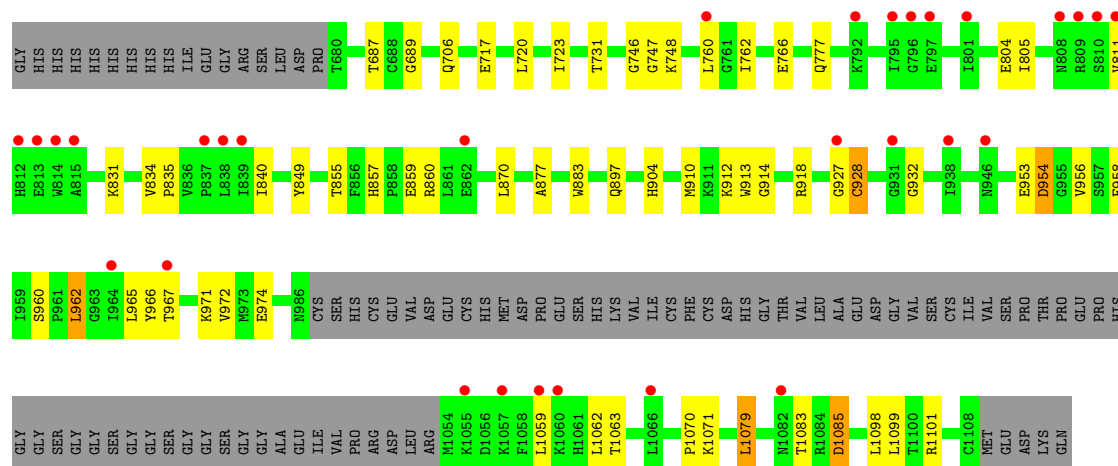


- Molecule 1: ALK tyrosine kinase receptor fused with ALK and LTK ligand 2

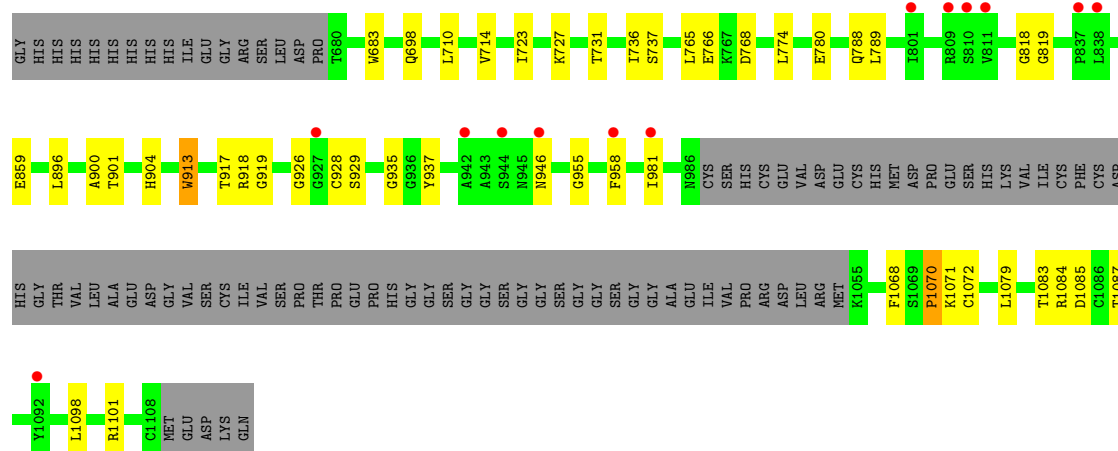


- Molecule 1: ALK tyrosine kinase receptor fused with ALK and LTK ligand 2





- Molecule 1: ALK tyrosine kinase receptor fused with ALK and LTK ligand 2



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	66.26Å 177.75Å 94.86Å 90.00° 104.80° 90.00°	Depositor
Resolution (Å)	49.77 – 3.05 60.27 – 3.05	Depositor EDS
% Data completeness (in resolution range)	99.5 (49.77-3.05) 99.5 (60.27-3.05)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.41 (at 3.07Å)	Xtriage
Refinement program	PHENIX dev_3915	Depositor
R, R_{free}	0.238 , 0.288 0.236 , 0.286	Depositor DCC
R_{free} test set	1946 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å ²)	70.9	Xtriage
Anisotropy	0.481	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , 38.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	10672	wwPDB-VP
Average B, all atoms (Å ²)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.70% 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: NAG, CIT

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	0.34	0/2720	0.52	0/3672
1	B	0.33	0/2720	0.52	0/3675
1	C	0.32	0/2724	0.50	0/3679
1	D	0.32	0/2716	0.50	0/3669
All	All	0.33	0/10880	0.51	0/14695

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	2658	0	2552	29	0
1	B	2658	0	2536	37	0
1	C	2662	0	2552	37	0
1	D	2654	0	2542	31	0
2	A	26	0	10	0	0
3	D	14	0	13	0	0
All	All	10672	0	10205	129	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.

The worst 5 of 129 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:962:LEU:H	1:C:962:LEU:HD23	1.18	1.02
1:A:980:ASN:HD22	1:A:1084:ARG:HD2	1.26	1.00
1:A:980:ASN:ND2	1:A:1084:ARG:HD2	1.88	0.87
1:A:760:LEU:HB2	1:A:967:THR:HG22	1.63	0.80
1:C:962:LEU:H	1:C:962:LEU:CD2	1.96	0.79

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	357/451 (79%)	334 (94%)	21 (6%)	2 (1%)	25	55
1	B	358/451 (79%)	340 (95%)	18 (5%)	0	100	100
1	C	358/451 (79%)	335 (94%)	22 (6%)	1 (0%)	41	70
1	D	357/451 (79%)	333 (93%)	23 (6%)	1 (0%)	41	70
All	All	1430/1804 (79%)	1342 (94%)	84 (6%)	4 (0%)	41	70

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	1070	PRO
1	D	1070	PRO
1	A	1107	VAL
1	A	972	VAL

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	271/345 (79%)	265 (98%)	6 (2%)	52 76
1	B	270/345 (78%)	262 (97%)	8 (3%)	41 69
1	C	272/345 (79%)	257 (94%)	15 (6%)	21 50
1	D	271/345 (79%)	267 (98%)	4 (2%)	65 83
All	All	1084/1380 (79%)	1051 (97%)	33 (3%)	41 69

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	1085	ASP
1	D	737	SER
1	D	913	TRP
1	B	1055	LYS
1	B	913	TRP

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	980	ASN
1	D	698	GLN
1	D	788	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](#)

3 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	CIT	A	1202	-	3,12,12	1.41	0	3,17,17	2.78	2 (66%)
2	CIT	A	1201	-	3,12,12	1.32	0	3,17,17	1.67	1 (33%)
3	NAG	D	1201	1	14,14,15	0.25	0	17,19,21	0.63	0

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	CIT	A	1202	-	-	0/6/16/16	-
2	CIT	A	1201	-	-	0/6/16/16	-
3	NAG	D	1201	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1202	CIT	C3-C2-C1	-3.54	109.32	114.98
2	A	1202	CIT	C3-C4-C5	-3.25	109.78	114.98
2	A	1201	CIT	C3-C2-C1	-2.64	110.76	114.98

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	A	361/451 (80%)	0.15	6 (1%) 70 46	32, 67, 106, 131	0
1	B	362/451 (80%)	0.20	13 (3%) 42 21	33, 64, 109, 137	0
1	C	362/451 (80%)	0.35	30 (8%) 11 4	38, 85, 152, 192	0
1	D	361/451 (80%)	0.18	13 (3%) 42 21	30, 61, 103, 144	0
All	All	1446/1804 (80%)	0.22	62 (4%) 35 16	30, 69, 123, 192	0

The worst 5 of 62 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	810	SER	5.4
1	A	1067	TYR	5.0
1	C	809	ARG	5.0
1	A	684	LEU	4.6
1	C	812	HIS	4.3

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(\AA^2)	Q<0.9
3	NAG	D	1201	14/15	0.64	0.29	111,143,152,152	0
2	CIT	A	1201	13/13	0.78	0.51	81,98,103,103	0
2	CIT	A	1202	13/13	0.79	0.41	115,117,127,145	0

6.5 Other polymers ⓘ

There are no such residues in this entry.