



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2017 – 08:37 AM EDT

PDB ID : 5GN0
Title : Structure of TAZ-TEAD complex
Authors : Kaan, H.Y.K.; Song, H.
Deposited on : 2016-07-18
Resolution : 2.90 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-20029077
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-20029077

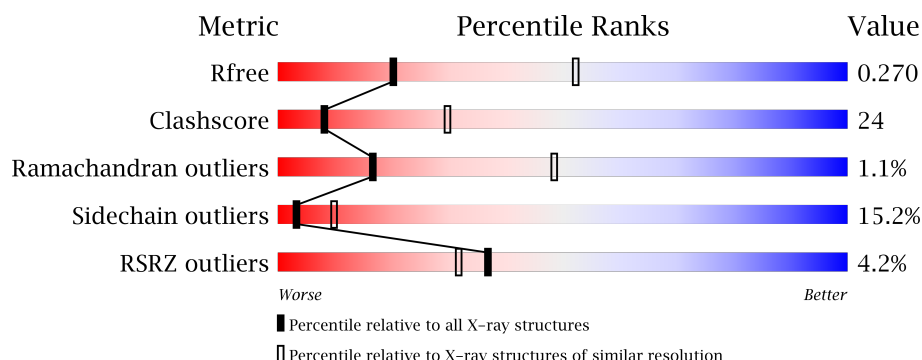
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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.



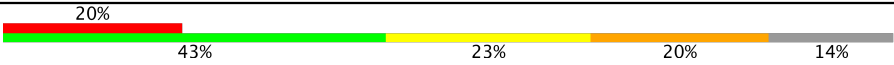


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	1586 (2.90-2.90)
Clashscore	112137	1807 (2.90-2.90)
Ramachandran outliers	110173	1768 (2.90-2.90)
Sidechain outliers	110143	1770 (2.90-2.90)
RSRZ outliers	101464	1596 (2.90-2.90)

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	225	<div> <div>2%</div> <div>52% 38% 8%</div> </div>
1	B	225	<div> <div>5%</div> <div>62% 31% 6%</div> </div>
1	C	225	<div> <div>%</div> <div>68% 24% 7%</div> </div>
1	D	225	<div> <div>4%</div> <div>61% 32% 5%</div> </div>
2	E	35	<div> <div>17%</div> <div>37% 43% 11% 9%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	35	
2	G	35	
2	H	35	

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
4	PLM	B	502	-	-	-	X
4	PLM	C	502	-	-	-	X
4	PLM	D	501	-	-	-	X

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 8228 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 Transcriptional enhancer factor TEF-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	222	Total	C	N	O	S	0	0	0
			1775	1144	297	324	10			
1	B	222	Total	C	N	O	S	0	0	0
			1780	1147	294	329	10			
1	C	222	Total	C	N	O	S	0	0	0
			1781	1147	295	329	10			
1	D	221	Total	C	N	O	S	0	0	0
			1797	1154	298	335	10			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	203	ASP	-	expression tag	UNP Q62296
A	204	LEU	-	expression tag	UNP Q62296
A	205	ASN	-	expression tag	UNP Q62296
A	206	TRP	-	expression tag	UNP Q62296
A	207	ILE	-	expression tag	UNP Q62296
A	208	SER	-	expression tag	UNP Q62296
A	209	MET	-	expression tag	UNP Q62296
B	203	ASP	-	expression tag	UNP Q62296
B	204	LEU	-	expression tag	UNP Q62296
B	205	ASN	-	expression tag	UNP Q62296
B	206	TRP	-	expression tag	UNP Q62296
B	207	ILE	-	expression tag	UNP Q62296
B	208	SER	-	expression tag	UNP Q62296
B	209	MET	-	expression tag	UNP Q62296
C	203	ASP	-	expression tag	UNP Q62296
C	204	LEU	-	expression tag	UNP Q62296
C	205	ASN	-	expression tag	UNP Q62296
C	206	TRP	-	expression tag	UNP Q62296
C	207	ILE	-	expression tag	UNP Q62296
C	208	SER	-	expression tag	UNP Q62296
C	209	MET	-	expression tag	UNP Q62296

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	203	ASP	-	expression tag	UNP Q62296
D	204	LEU	-	expression tag	UNP Q62296
D	205	ASN	-	expression tag	UNP Q62296
D	206	TRP	-	expression tag	UNP Q62296
D	207	ILE	-	expression tag	UNP Q62296
D	208	SER	-	expression tag	UNP Q62296
D	209	MET	-	expression tag	UNP Q62296

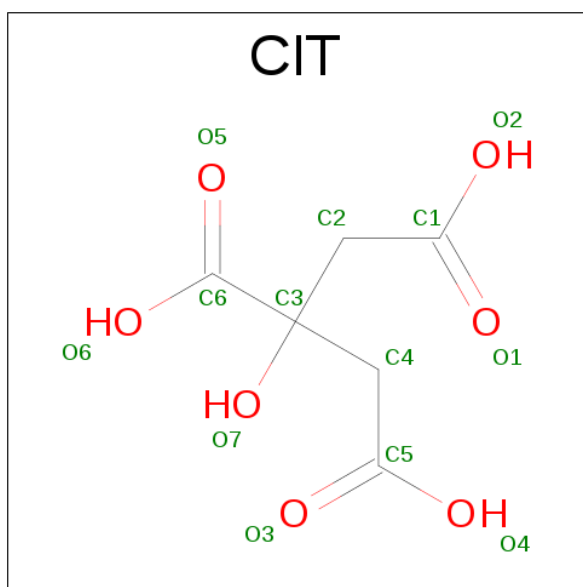
- Molecule 2 is a protein called WW domain-containing transcription regulator protein 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	32	Total	C	N	O	0	0	0
			250	164	40	46			
2	F	30	Total	C	N	O	0	0	0
			231	154	38	39			
2	G	33	Total	C	N	O	0	0	0
			264	171	42	50			
2	H	32	Total	C	N	O	0	0	0
			252	165	41	45			

There are 4 discrepancies between the modelled and reference sequences:

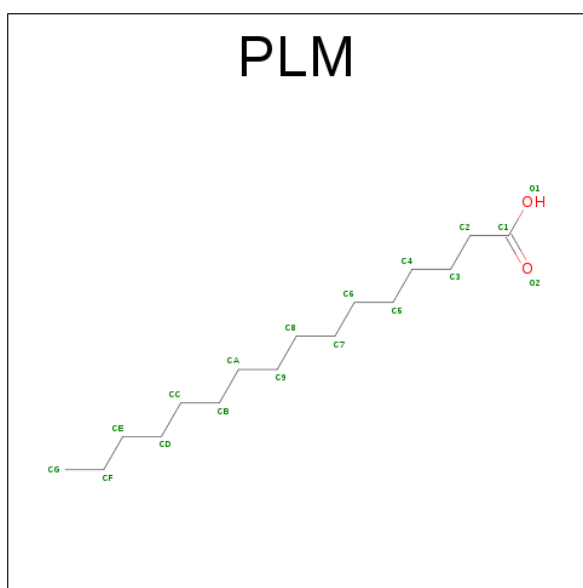
Chain	Residue	Modelled	Actual	Comment	Reference
E	23	PRO	-	expression tag	UNP Q9EPK5
F	23	PRO	-	expression tag	UNP Q9EPK5
G	23	PRO	-	expression tag	UNP Q9EPK5
H	23	PRO	-	expression tag	UNP Q9EPK5

- Molecule 3 is CITRIC ACID (three-letter code: CIT) (formula: C₆H₈O₇).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	0
			13	6	7		
3	C	1	Total	C	O	0	0
			13	6	7		

- Molecule 4 is PALMITIC ACID (three-letter code: PLM) (formula: $C_{16}H_{32}O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			18	16	2		
4	C	1	Total	C	O	0	0
			18	16	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	C	O	0	0
			17	16	1		

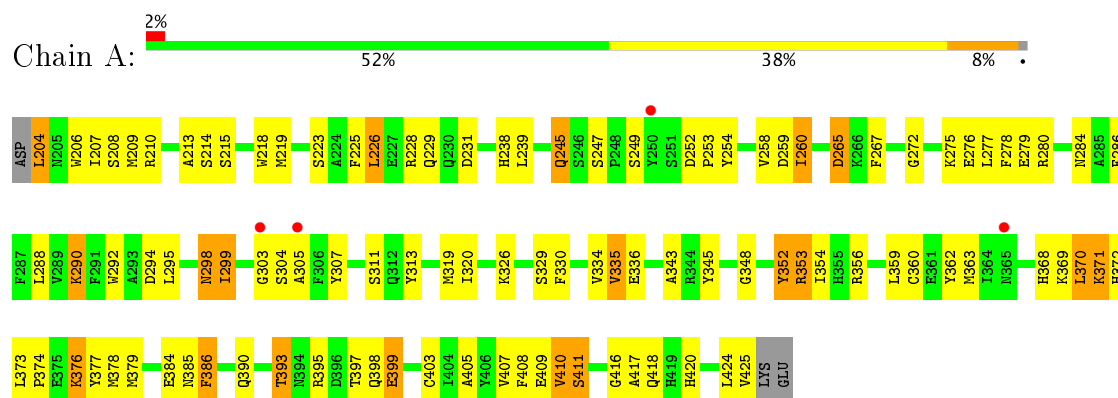
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	5	Total	O	0	0
			5	5		
5	B	7	Total	O	0	0
			7	7		
5	C	4	Total	O	0	0
			4	4		
5	D	2	Total	O	0	0
			2	2		
5	F	1	Total	O	0	0
			1	1		

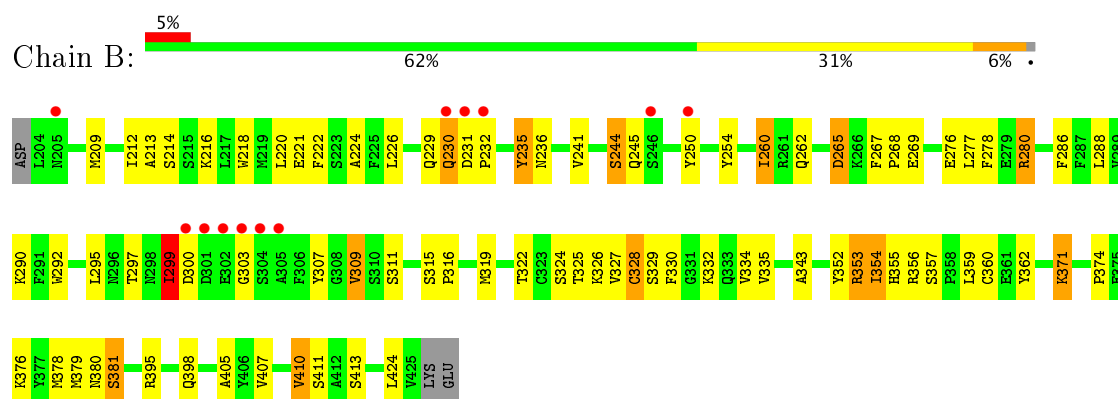
3 Residue-property plots [i](#)

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.

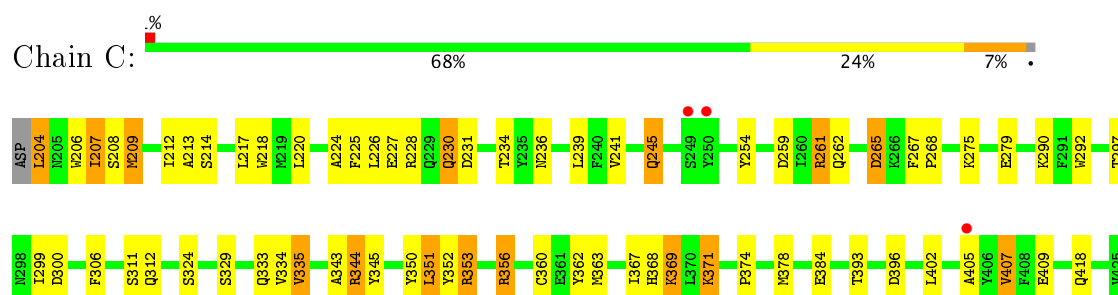
• Molecule 1: Transcriptional enhancer factor TEF-3



• Molecule 1: Transcriptional enhancer factor TEF-3

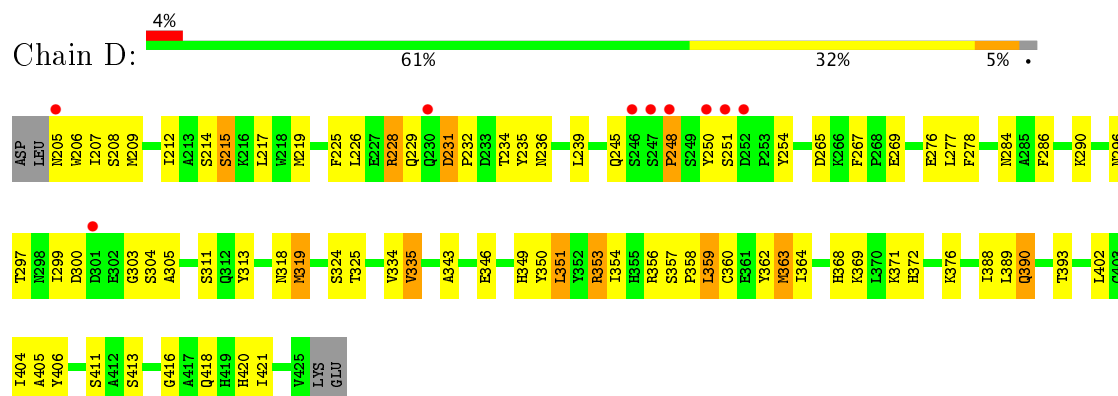


• Molecule 1: Transcriptional enhancer factor TEF-3

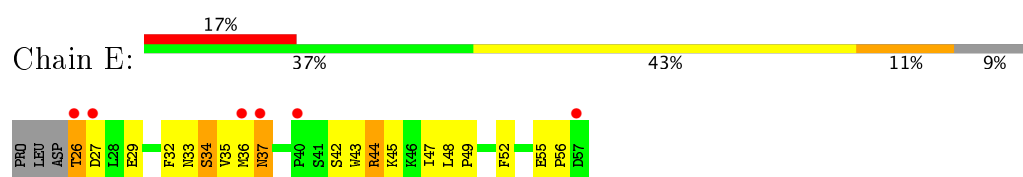


LYS
GLU

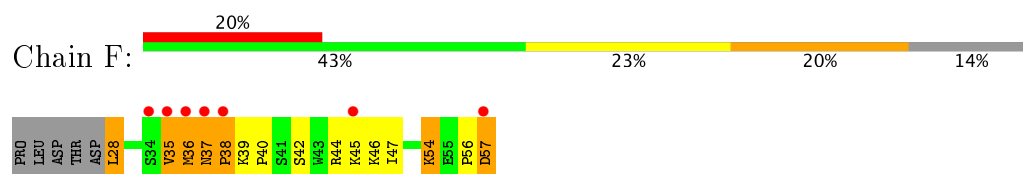
- Molecule 1: Transcriptional enhancer factor TEF-3



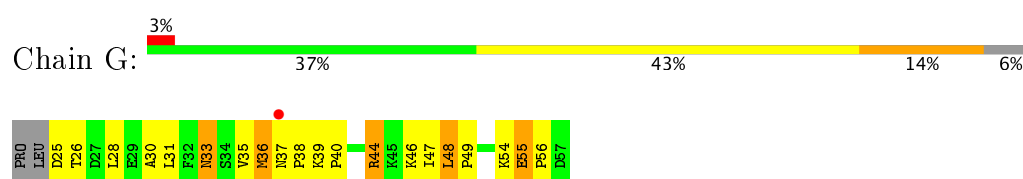
- Molecule 2: WW domain-containing transcription regulator protein 1



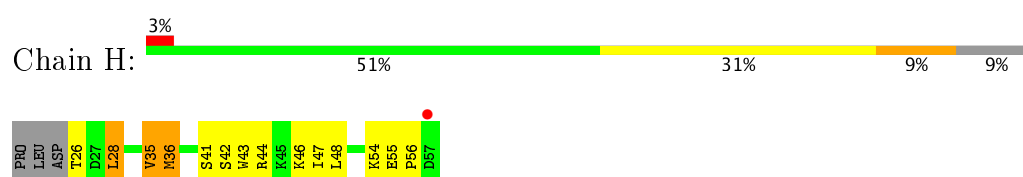
- Molecule 2: WW domain-containing transcription regulator protein 1



- Molecule 2: WW domain-containing transcription regulator protein 1



- Molecule 2: WW domain-containing transcription regulator protein 1



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	79.66Å 120.90Å 196.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.90 29.98 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.5 (30.00-2.90) 98.6 (29.98-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.63 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, R_{free}	0.200 , 0.260 0.217 , 0.270	Depositor DCC
R_{free} test set	2131 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	46.9	Xtriage
Anisotropy	1.361	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 52.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	8228	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.74% 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: PLM, 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	1.19	0/1824	0.87	1/2480 (0.0%)
1	B	1.29	0/1829	0.90	2/2486 (0.1%)
1	C	1.35	0/1829	0.88	1/2485 (0.0%)
1	D	1.38	0/1846	0.89	4/2505 (0.2%)
2	E	0.91	0/258	0.83	0/351
2	F	1.09	0/239	0.81	0/325
2	G	1.13	0/272	1.00	1/369 (0.3%)
2	H	1.16	0/260	0.87	0/353
All	All	1.28	0/8357	0.89	9/11354 (0.1%)

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	299	ILE	N-CA-C	7.26	130.59	111.00
1	D	357	SER	C-N-CD	-6.33	106.67	120.60
1	A	303	GLY	N-CA-C	-5.95	98.24	113.10
2	G	55	GLU	C-N-CD	-5.90	107.62	120.60
1	C	351	LEU	CA-CB-CG	5.76	128.55	115.30

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	1775	0	1671	106	0
1	B	1780	0	1675	75	0
1	C	1781	0	1681	65	0
1	D	1797	0	1701	60	0
2	E	250	0	235	24	0
2	F	231	0	220	22	0
2	G	264	0	250	27	0
2	H	252	0	242	15	0
3	B	13	0	5	4	0
3	C	13	0	5	2	0
4	B	18	0	31	2	0
4	C	18	0	31	2	0
4	D	17	0	31	5	0
5	A	5	0	0	0	0
5	B	7	0	0	0	0
5	C	4	0	0	0	0
5	D	2	0	0	0	0
5	F	1	0	0	0	0
All	All	8228	0	7778	376	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 24.

The worst 5 of 376 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)
2:E:47:ILE:C	2:E:48:LEU:HD23	1.51	1.31
2:E:47:ILE:O	2:E:48:LEU:HD23	1.24	1.26
1:A:370:LEU:CD1	1:A:379:MET:HG2	1.73	1.17
1:A:379:MET:HE3	1:A:410:VAL:HG21	1.28	1.15
2:F:37:ASN:CB	2:F:38:PRO:HD3	1.74	1.14

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	220/225 (98%)	214 (97%)	5 (2%)	1 (0%)	32	68
1	B	220/225 (98%)	215 (98%)	5 (2%)	0	100	100
1	C	220/225 (98%)	213 (97%)	7 (3%)	0	100	100
1	D	219/225 (97%)	209 (95%)	9 (4%)	1 (0%)	32	68
2	E	30/35 (86%)	26 (87%)	1 (3%)	3 (10%)	1	1
2	F	28/35 (80%)	23 (82%)	2 (7%)	3 (11%)	0	1
2	G	31/35 (89%)	27 (87%)	3 (10%)	1 (3%)	5	19
2	H	30/35 (86%)	25 (83%)	3 (10%)	2 (7%)	1	4
All	All	998/1040 (96%)	952 (95%)	35 (4%)	11 (1%)	17	48

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	37	ASN
2	F	36	MET
2	F	38	PRO
2	H	35	VAL
2	E	34	SER

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	189/208 (91%)	158 (84%)	31 (16%)	2	8
1	B	191/208 (92%)	162 (85%)	29 (15%)	3	10
1	C	191/208 (92%)	167 (87%)	24 (13%)	5	15
1	D	196/208 (94%)	167 (85%)	29 (15%)	3	11
2	E	27/34 (79%)	23 (85%)	4 (15%)	3	11
2	F	24/34 (71%)	19 (79%)	5 (21%)	1	4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	G	30/34 (88%)	24 (80%)	6 (20%)	1	4
2	H	28/34 (82%)	23 (82%)	5 (18%)	2	6
All	All	876/968 (90%)	743 (85%)	133 (15%)	3	10

5 of 133 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	204	LEU
1	C	335	VAL
2	G	25	ASP
1	C	207	ILE
1	C	261	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	390	GLN
1	B	415	HIS
1	D	230	GLN
1	B	296	ASN
1	B	312	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 carbohydrates 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	$\# Z > 2$
3	CIT	B	501	-	3,12,12	1.50	1 (33%)	3,17,17	2.08	1 (33%)
4	PLM	B	502	-	14,17,17	0.17	0	13,17,17	0.98	0
3	CIT	C	501	-	3,12,12	0.79	0	3,17,17	1.90	2 (66%)
4	PLM	C	502	-	14,17,17	0.07	0	13,17,17	0.88	0
4	PLM	D	501	1	16,16,17	0.25	0	15,15,17	0.79	1 (6%)

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
3	CIT	B	501	-	-	0/6/16/16	0/0/0/0
4	PLM	B	502	-	-	0/13/15/15	0/0/0/0
3	CIT	C	501	-	-	0/6/16/16	0/0/0/0
4	PLM	C	502	-	-	0/13/15/15	0/0/0/0
4	PLM	D	501	1	-	0/13/14/15	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	501	CIT	C2-C3	2.58	1.58	1.54

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	501	PLM	O2-C1-C2	-2.14	115.17	125.72
3	C	501	CIT	C4-C3-C2	2.16	115.12	109.75
3	C	501	CIT	C3-C2-C1	2.38	118.66	114.95
3	B	501	CIT	C3-C2-C1	3.24	120.02	114.95

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	501	CIT	4	0
4	B	502	PLM	2	0
3	C	501	CIT	2	0
4	C	502	PLM	2	0
4	D	501	PLM	5	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	222/225 (98%)	0.17	4 (1%) 69 66	25, 61, 98, 126	0
1	B	222/225 (98%)	0.07	12 (5%) 26 22	19, 44, 83, 111	0
1	C	222/225 (98%)	-0.11	3 (1%) 75 74	27, 45, 70, 92	0
1	D	221/225 (98%)	0.06	9 (4%) 38 32	23, 41, 86, 116	0
2	E	32/35 (91%)	1.02	6 (18%) 1 1	30, 81, 110, 121	0
2	F	30/35 (85%)	0.86	7 (23%) 1 0	34, 80, 139, 148	0
2	G	33/35 (94%)	0.08	1 (3%) 51 44	37, 58, 77, 101	0
2	H	32/35 (91%)	0.12	1 (3%) 49 43	40, 51, 104, 121	0
All	All	1014/1040 (97%)	0.11	43 (4%) 37 32	19, 48, 97, 148	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	38	PRO	6.3
1	D	205	ASN	3.9
1	B	302	GLU	3.9
2	E	57	ASP	3.9
1	B	246	SER	3.8

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 [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. 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
4	PLM	B	502	18/18	0.84	0.48	7.33	56,62,89,94	0
4	PLM	D	501	17/18	0.82	0.41	4.43	34,52,75,77	0
4	PLM	C	502	18/18	0.73	0.39	4.28	59,70,90,94	0
3	CIT	C	501	13/13	0.78	0.20	-0.12	34,44,60,62	0
3	CIT	B	501	13/13	0.82	0.19	-0.26	32,43,50,51	0

6.5 Other polymers [i](#)

There are no such residues in this entry.