WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 30, 2023 – 07:51 PM JST

PDB ID	:	4YRA
Title	:	mouse TDH in the apo form
Authors	:	He, C.; Li, F.
Deposited on	:	2015-03-14
Resolution	:	2.65 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 (i) 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 (i)) were used in the production of this report:

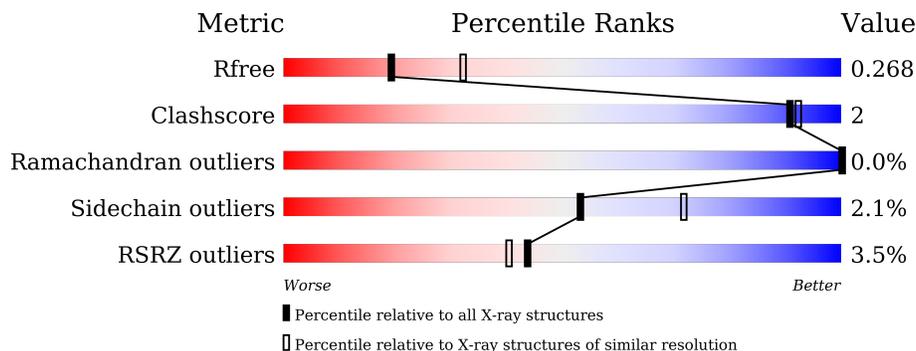|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.65 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	Percentile Ranks	Value
Rfree		0.268
Clashscore		2
Ramachandran outliers		0.0%
Sidechain outliers		2.1%
RSRZ outliers		3.5%

Worse　　　　　　　　　　　　　　　　　　　　Better

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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 <=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	329	5%　82%　6%　11%
1	B	329	5%　83%　6%　11%
1	C	329	4%　82%　6% ·　11%
1	D	329	2%　86%　· 11%
1	E	329	2%　84%　· ·　12%
1	F	329	2%　85%　· 13%

Continued on next page...

wwPDB X-ray Structure Validation Summary Report

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	329	2% 81% 8% 11%
1	H	329	3% 84% 5% 10%
1	I	329	2% 84% 12%
1	J	329	3% 81% 8% 11%
1	K	329	5% 83% 13%
1	L	329	2% 86% 9%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 27791 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 L-threonine 3-dehydrogenase, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	292	Total	C	N	O	S	0	0	0
			2295	1465	396	421	13			
1	B	294	Total	C	N	O	S	0	0	0
			2324	1483	404	424	13			
1	C	293	Total	C	N	O	S	0	0	0
			2321	1485	403	420	13			
1	D	292	Total	C	N	O	S	0	0	0
			2307	1484	398	412	13			
1	E	291	Total	C	N	O	S	0	0	0
			2313	1483	401	416	13			
1	F	287	Total	C	N	O	S	0	0	0
			2296	1475	399	409	13			
1	G	294	Total	C	N	O	S	0	0	0
			2310	1482	399	416	13			
1	H	295	Total	C	N	O	S	0	0	0
			2334	1494	406	421	13			
1	I	291	Total	C	N	O	S	0	0	0
			2308	1480	398	417	13			
1	J	292	Total	C	N	O	S	0	0	0
			2294	1466	398	417	13			
1	K	286	Total	C	N	O	S	0	0	0
			2269	1451	393	412	13			
1	L	299	Total	C	N	O	S	0	0	0
			2350	1503	407	427	13			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	45	HIS	-	expression tag	UNP Q8K3F7
A	46	MET	-	expression tag	UNP Q8K3F7
B	45	HIS	-	expression tag	UNP Q8K3F7
B	46	MET	-	expression tag	UNP Q8K3F7
C	45	HIS	-	expression tag	UNP Q8K3F7

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	46	MET	-	expression tag	UNP Q8K3F7
D	45	HIS	-	expression tag	UNP Q8K3F7
D	46	MET	-	expression tag	UNP Q8K3F7
E	45	HIS	-	expression tag	UNP Q8K3F7
E	46	MET	-	expression tag	UNP Q8K3F7
F	45	HIS	-	expression tag	UNP Q8K3F7
F	46	MET	-	expression tag	UNP Q8K3F7
G	45	HIS	-	expression tag	UNP Q8K3F7
G	46	MET	-	expression tag	UNP Q8K3F7
H	45	HIS	-	expression tag	UNP Q8K3F7
H	46	MET	-	expression tag	UNP Q8K3F7
I	45	HIS	-	expression tag	UNP Q8K3F7
I	46	MET	-	expression tag	UNP Q8K3F7
J	45	HIS	-	expression tag	UNP Q8K3F7
J	46	MET	-	expression tag	UNP Q8K3F7
K	45	HIS	-	expression tag	UNP Q8K3F7
K	46	MET	-	expression tag	UNP Q8K3F7
L	45	HIS	-	expression tag	UNP Q8K3F7
L	46	MET	-	expression tag	UNP Q8K3F7

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	6	Total 6	O 6	0	0
2	B	11	Total 11	O 11	0	0
2	C	6	Total 6	O 6	0	0
2	D	6	Total 6	O 6	0	0
2	E	6	Total 6	O 6	0	0
2	F	9	Total 9	O 9	0	0
2	G	8	Total 8	O 8	0	0
2	H	2	Total 2	O 2	0	0
2	I	3	Total 3	O 3	0	0
2	J	4	Total 4	O 4	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	K	2	Total 2	O 2	0	0
2	L	7	Total 7	O 7	0	0

3 Residue-property plots ⓘ

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.

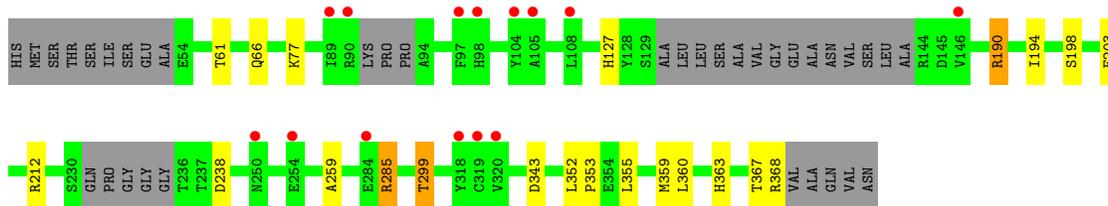- Molecule 1: L-threonine 3-dehydrogenase, mitochondrial

- Molecule 1: L-threonine 3-dehydrogenase, mitochondrial
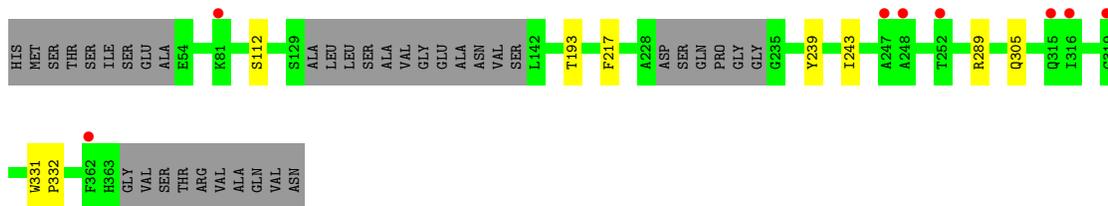

- Molecule 1: L-threonine 3-dehydrogenase, mitochondrial
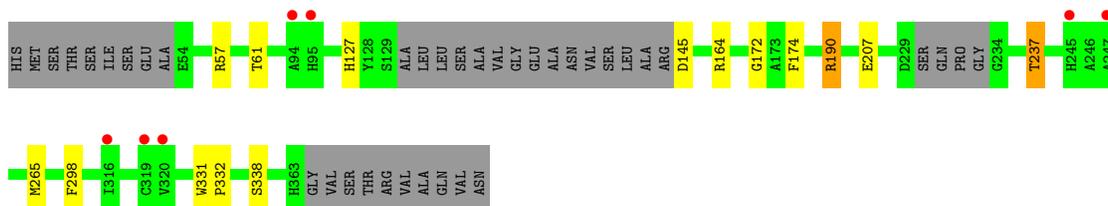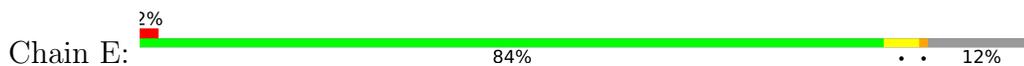

- Molecule 1: L-threonine 3-dehydrogenase, mitochondrial

HIS MET SER THR ILE SER GLU ALA **E54 K81 S112 S129** ALA LEU LEU SER ALA VAL GLY GLU ALA ASN VAL SER **L142 T193 F217 A228** ASP GLN PRO GLY GLY **G235 Y239 I243 A247 A248 T252 R289 Q305 Q315 I316 C319**
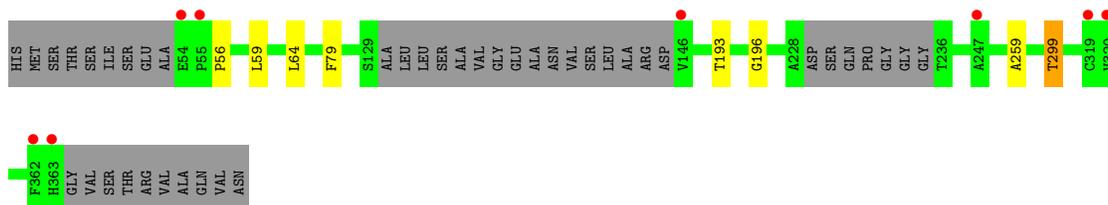
W331 P332 F362 H363 GLY VAL SER THR ARG VAL ALA GLN VAL ASN

• Molecule 1: L-threonine 3-dehydrogenase, mitochondrial

Chain E:

2%
84% 12%

HIS MET SER THR ILE SER GLU ALA **E54 R57 T61 A94 H95 H127 Y128 S129** ALA LEU LEU SER ALA VAL GLY GLU ALA ASN VAL SER LEU ARG **D145 R164 G172 A173 F174 R190 E207 D229** SER GLN PRO GLY **G234 T237 H245 A246 A247**

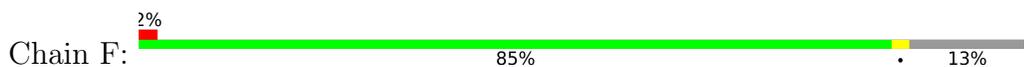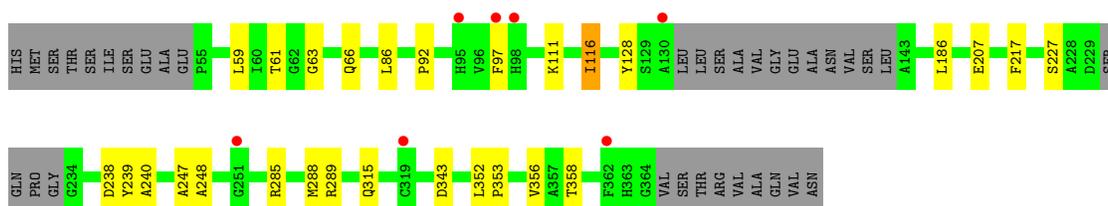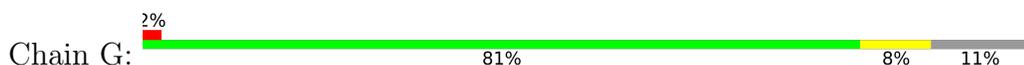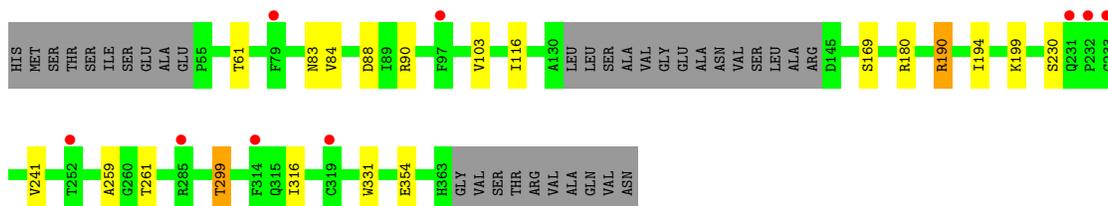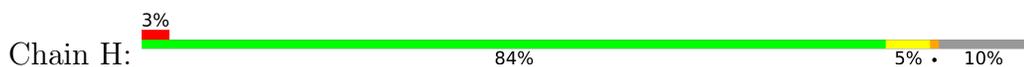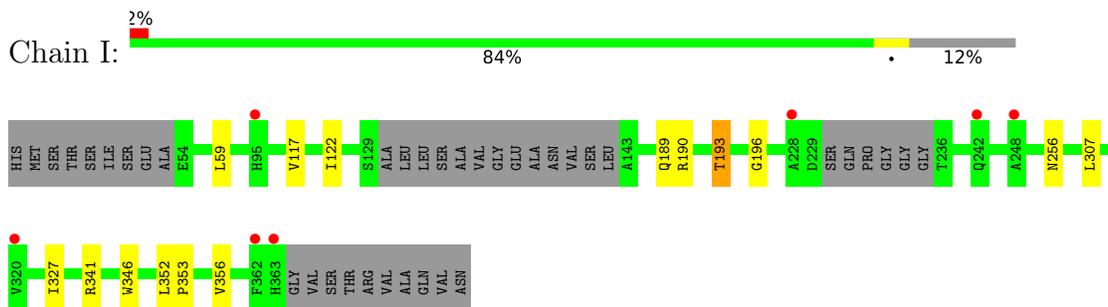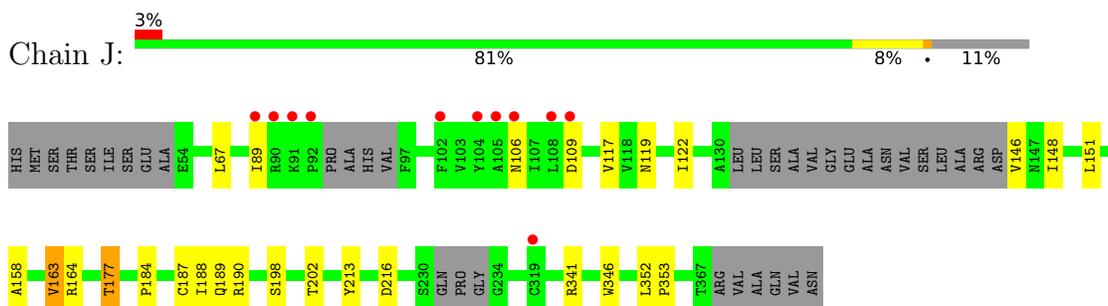**M265 F298 I316 C319 V320 W331 P332 S338 H363** GLY VAL SER THR ARG VAL ALA GLN VAL ASN

• Molecule 1: L-threonine 3-dehydrogenase, mitochondrial

Chain F:

2%
85% 13%

HIS MET SER THR ILE SER GLU ALA **E54 P55 P56 L59 L64 F79 S129** ALA LEU LEU SER ALA VAL GLY GLU ALA ASN VAL SER LEU ALA ARG ASP **V146 T193 G196 A228** ASP SER GLN PRO GLY GLY GLY **T236 A247 A259 T299 C319 V320**

F362 H363 GLY VAL SER THR ARG VAL ALA GLN VAL ASN

• Molecule 1: L-threonine 3-dehydrogenase, mitochondrial

Chain G:

2%
81% 8% 11%

HIS MET SER THR ILE SER GLU ALA GLU **P55 L59 I60 T61 G62 G63 Q66 L86 P92 H95 V96 F97 H98 K111 I116 Y128 S129 A130** LEU LEU SER ALA VAL GLY GLU ALA ASN VAL SER LEU **A143 L186 E207 F217 S227 A228 D229** SER

GLN PRO GLY **G234 D238 Y239 A240 A247 A248 G251 R285 M288 R289 Q315 C319 D343 L352 P353 V356 A357 T358 F362 H363 G364** VAL SER THR ARG VAL ALA GLN VAL ASN

• Molecule 1: L-threonine 3-dehydrogenase, mitochondrial

Chain H:

3%
84% 5% 10%

HIS MET SER THR ILE SER GLU ALA GLU **P55 T61 F79 N83 V84 D88 I89 R90 F97 V103 I116 A130** LEU LEU SER ALA VAL GLY GLU ALA ASN VAL SER LEU ALA ARG **D145 S169 R180 R190 I194 K199 S230 Q231 P232 G233**

V241 T252 A259 G260 T261 R285 T299 F314 Q315 I316 C319 W331 E354 **H363** GLY VAL SER THR ARG VAL ALA GLN VAL ASN
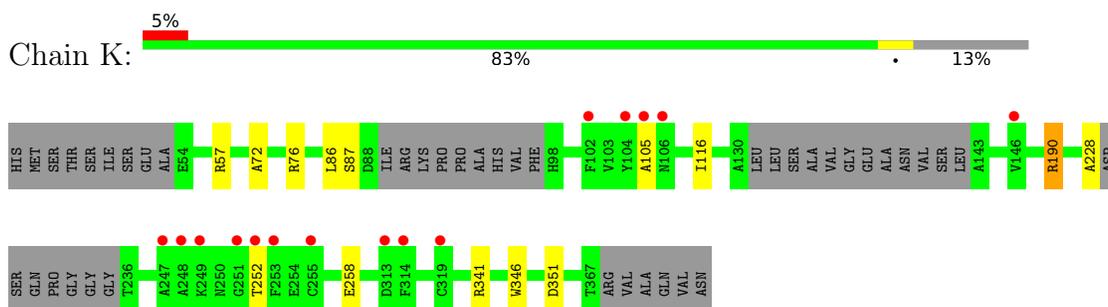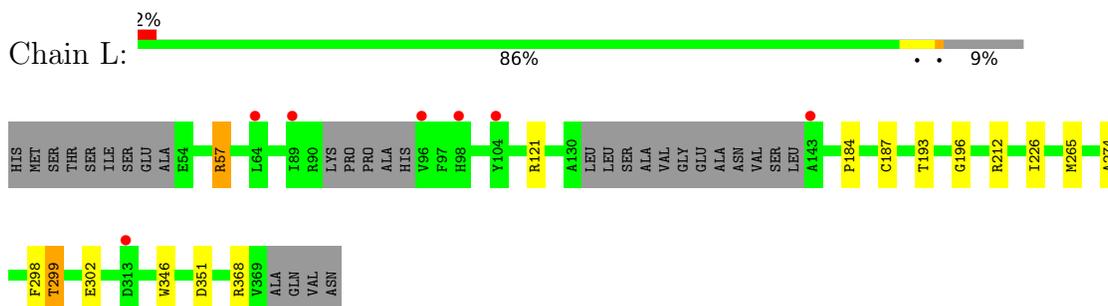
• Molecule 1: L-threonine 3-dehydrogenase, mitochondrial

Chain I:

• Molecule 1: L-threonine 3-dehydrogenase, mitochondrial

Chain J:

• Molecule 1: L-threonine 3-dehydrogenase, mitochondrial

Chain K:

• Molecule 1: L-threonine 3-dehydrogenase, mitochondrial

Chain L:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	94.44Å 157.31Å 162.34Å 90.00° 97.00° 90.00°	Depositor
Resolution (Å)	44.35 – 2.65 44.35 – 2.65	Depositor EDS
% Data completeness (in resolution range)	91.5 (44.35-2.65) 91.5 (44.35-2.65)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.95 (at 2.65Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.217 , 0.271 0.217 , 0.268	Depositor DCC
R_{free} test set	6300 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	58.5	Xtriage
Anisotropy	0.096	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 31.7	EDS
L-test for twinning[2]	$<|L|> = 0.48, <L^2> = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	27791	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.75% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<|L|>$, $<L^2>$ 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 ⓘ

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	0.29	0/2353	0.52	0/3200
1	B	0.30	0/2383	0.53	0/3239
1	C	0.29	0/2380	0.50	0/3236
1	D	0.28	0/2369	0.49	0/3224
1	E	0.28	0/2375	0.48	0/3230
1	F	0.29	0/2358	0.50	0/3206
1	G	0.29	0/2372	0.52	0/3228
1	H	0.29	0/2398	0.49	0/3264
1	I	0.30	0/2370	0.53	0/3227
1	J	0.31	0/2351	0.53	0/3196
1	K	0.29	0/2325	0.50	0/3159
1	L	0.31	0/2410	0.54	0/3279
All	All	0.29	0/28444	0.51	0/38688

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	2295	0	2237	11	0
1	B	2324	0	2277	8	0
1	C	2321	0	2264	11	0
1	D	2307	0	2263	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	2313	0	2265	8	0
1	F	2296	0	2266	3	0
1	G	2310	0	2253	14	0
1	H	2334	0	2282	8	0
1	I	2308	0	2248	8	0
1	J	2294	0	2242	16	0
1	K	2269	0	2230	8	0
1	L	2350	0	2291	9	0
2	A	6	0	0	0	0
2	B	11	0	0	0	0
2	C	6	0	0	0	0
2	D	6	0	0	0	0
2	E	6	0	0	0	0
2	F	9	0	0	0	0
2	G	8	0	0	0	0
2	H	2	0	0	0	0
2	I	3	0	0	0	0
2	J	4	0	0	1	0
2	K	2	0	0	0	0
2	L	7	0	0	0	0
All	All	27791	0	27118	95	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 2.

The worst 5 of 95 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:K:87:SER:OG	1:K:105:ALA:HB3	1.92	0.69
1:L:226:ILE:HD11	1:L:265:MET:CE	2.24	0.68
1:I:190:ARG:NH2	1:J:189:GLN:OE1	2.29	0.65
1:J:341:ARG:HG2	1:J:346:TRP:O	1.98	0.64
1:A:61:THR:O	1:A:128:TYR:N	2.32	0.63

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	286/329 (87%)	275 (96%)	11 (4%)	0	100	100
1	B	288/329 (88%)	279 (97%)	9 (3%)	0	100	100
1	C	285/329 (87%)	276 (97%)	9 (3%)	0	100	100
1	D	286/329 (87%)	275 (96%)	11 (4%)	0	100	100
1	E	285/329 (87%)	280 (98%)	5 (2%)	0	100	100
1	F	281/329 (85%)	271 (96%)	10 (4%)	0	100	100
1	G	288/329 (88%)	275 (96%)	13 (4%)	0	100	100
1	H	291/329 (88%)	282 (97%)	9 (3%)	0	100	100
1	I	285/329 (87%)	275 (96%)	10 (4%)	0	100	100
1	J	284/329 (86%)	278 (98%)	5 (2%)	1 (0%)	34	48
1	K	278/329 (84%)	268 (96%)	10 (4%)	0	100	100
1	L	293/329 (89%)	283 (97%)	10 (3%)	0	100	100
All	All	3430/3948 (87%)	3317 (97%)	112 (3%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	J	89	ILE

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	244/277 (88%)	240 (98%)	4 (2%)	62	78
1	B	248/277 (90%)	243 (98%)	5 (2%)	55	73
1	C	247/277 (89%)	239 (97%)	8 (3%)	39	56
1	D	244/277 (88%)	239 (98%)	5 (2%)	55	73
1	E	246/277 (89%)	240 (98%)	6 (2%)	49	67
1	F	246/277 (89%)	243 (99%)	3 (1%)	71	84
1	G	243/277 (88%)	238 (98%)	5 (2%)	53	72
1	H	248/277 (90%)	240 (97%)	8 (3%)	39	56
1	I	245/277 (88%)	242 (99%)	3 (1%)	71	84
1	J	243/277 (88%)	237 (98%)	6 (2%)	47	66
1	K	242/277 (87%)	238 (98%)	4 (2%)	60	77
1	L	248/277 (90%)	244 (98%)	4 (2%)	62	78
All	All	2944/3324 (89%)	2883 (98%)	61 (2%)	53	72

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	59	LEU
1	K	252	THR
1	G	358	THR
1	K	190	ARG
1	L	299	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	66	GLN
1	H	66	GLN
1	L	361	ASN
1	J	152	HIS
1	L	152	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4   Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6   Ligand geometry ⓘ

There are no ligands in this entry.

5.7   Other polymers ⓘ

There are no such residues in this entry.

5.8   Polymer linkage issues ⓘ

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	292/329 (88%)	0.36	16 (5%)	25	21	48, 62, 94, 128	0
1	B	294/329 (89%)	0.23	15 (5%)	28	25	41, 58, 91, 122	0
1	C	293/329 (89%)	0.37	14 (4%)	30	27	52, 68, 103, 140	0
1	D	292/329 (88%)	0.18	8 (2%)	54	50	48, 67, 100, 118	0
1	E	291/329 (88%)	0.15	7 (2%)	59	54	53, 69, 101, 118	0
1	F	287/329 (87%)	0.11	8 (2%)	53	49	42, 63, 97, 116	0
1	G	294/329 (89%)	-0.01	7 (2%)	59	54	42, 61, 95, 113	0
1	H	295/329 (89%)	0.20	9 (3%)	49	45	52, 73, 101, 131	0
1	I	291/329 (88%)	0.01	7 (2%)	59	54	41, 60, 92, 107	0
1	J	292/329 (88%)	0.20	11 (3%)	40	36	37, 57, 94, 126	0
1	K	286/329 (86%)	0.17	15 (5%)	27	24	41, 62, 96, 114	0
1	L	299/329 (90%)	-0.00	7 (2%)	60	56	35, 50, 98, 128	0
All	All	3506/3948 (88%)	0.16	124 (3%)	44	40	35, 64, 98, 140	0

The worst 5 of 124 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	319	CYS	5.6
1	L	98	HIS	4.9
1	K	252	THR	4.7
1	K	248	ALA	4.6
1	C	146	VAL	4.5

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 monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.