WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 18, 2024 – 10:42 AM EST

PDB ID	:	4E5T
Title	:	Crystal structure of a putative Mandelate racemase/Muconate lactonizing enzyme (Target PSI-200750) from Labrenzia alexandrii DFL-11
Authors	:	Kumar, P.R.; Bonanno, J.; Chowdhury, S.; Foti, R.; Gizzi, A.; Glen, S.; Hammonds, J.; Hillerich, B.; Matikainen, B.; Seidel, R.; Toro, R.; Zencheck, W.; Almo, S.C.; New York Structural Genomics Research Consortium (NYSGRC)
Deposited on	:	2012-03-14
Resolution	:	2.90 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

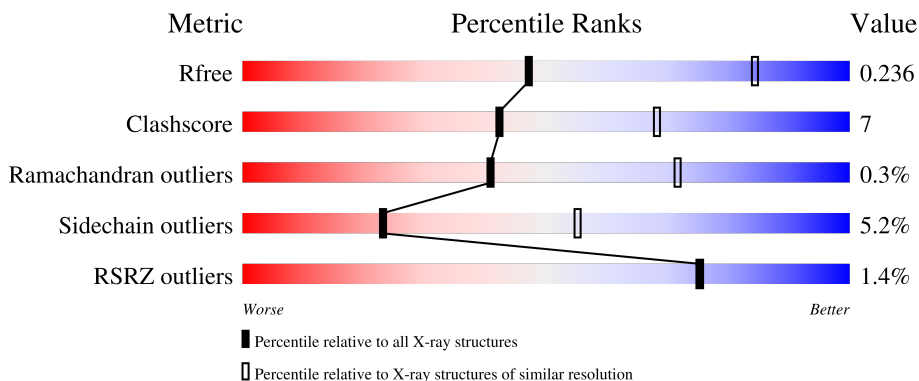|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.90 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.
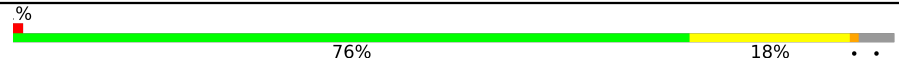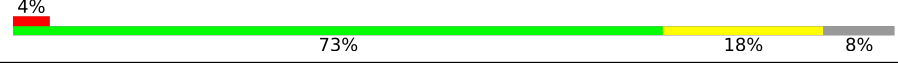
Metric	Percentile Ranks	Value
Rfree		0.236
Clashscore		7
Ramachandran outliers		0.3%
Sidechain outliers		5.2%
RSRZ outliers		1.4%

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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (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 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	404	76% 19% · ·
1	B	404	78% 18% ·
1	C	404	% 76% 18% · ·
1	D	404	75% 19% · ·
1	E	404	72% 18% · 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	404	% 76% 18%
1	G	404	3% 75% 16% 8%
1	H	404	4% 73% 18% 8%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 23681 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 Mandelate racemase / muconate lactonizing enzyme, C-terminal domain protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	392	Total	C	N	O	S	0	0	0
			3031	1939	513	564	15			
1	B	387	Total	C	N	O	S	0	1	0
			3006	1920	511	560	15			
1	C	386	Total	C	N	O	S	0	0	0
			2988	1910	506	557	15			
1	D	388	Total	C	N	O	S	0	0	0
			3004	1920	509	560	15			
1	E	371	Total	C	N	O	S	0	0	0
			2867	1834	485	533	15			
1	F	386	Total	C	N	O	S	0	0	0
			2991	1911	507	558	15			
1	G	370	Total	C	N	O	S	0	0	0
			2860	1830	484	531	15			
1	H	371	Total	C	N	O	S	0	0	0
			2867	1834	485	533	15			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP B9R417
B	0	SER	-	expression tag	UNP B9R417
C	0	SER	-	expression tag	UNP B9R417
D	0	SER	-	expression tag	UNP B9R417
E	0	SER	-	expression tag	UNP B9R417
F	0	SER	-	expression tag	UNP B9R417
G	0	SER	-	expression tag	UNP B9R417
H	0	SER	-	expression tag	UNP B9R417

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Mg 1	0	0
2	B	1	Total 1	Mg 1	0	0
2	C	1	Total 1	Mg 1	0	0
2	D	1	Total 1	Mg 1	0	0
2	E	1	Total 1	Mg 1	0	0
2	F	1	Total 1	Mg 1	0	0
2	G	1	Total 1	Mg 1	0	0
2	H	1	Total 1	Mg 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	8	Total 8	O 8	0	0
3	B	9	Total 9	O 9	0	0
3	C	8	Total 8	O 8	0	0
3	D	10	Total 10	O 10	0	0
3	E	9	Total 9	O 9	0	0
3	F	5	Total 5	O 5	0	0
3	G	6	Total 6	O 6	0	0
3	H	4	Total 4	O 4	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 t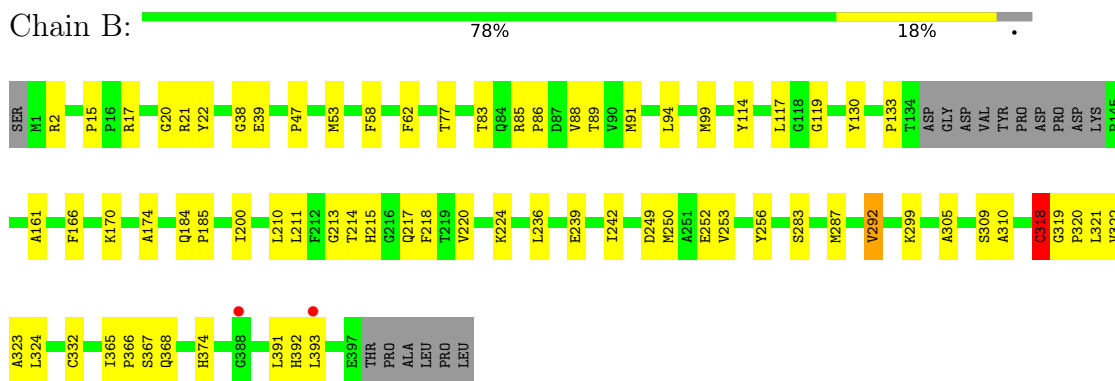he various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 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.
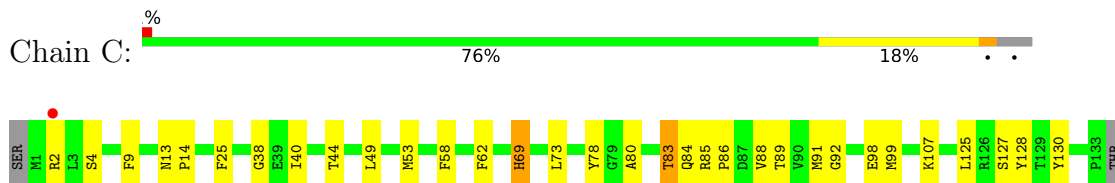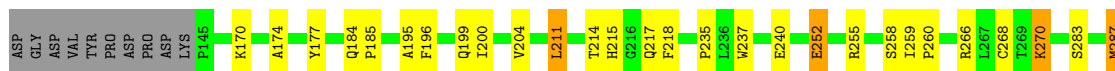
- Molecule 1: Mandelate racemase / muconate lactonizing enzyme, C-terminal domain protein

- Molecule 1: Mandelate racemase / muconate lactonizing enzyme, C-terminal domain protein

- Molecule 1: Mandelate racemase / muconate lactonizing enzyme, C-terminal domain protein

ASP GLY ASP VAL TYR PRO ASP PRO ASP LYS P145 K170 A174 Y177 Q184 P185 A195 F196 Q199 I200 V204 L211 T214 H215 G216 Q217 F218 P235 L236 W237 E240 E252 R255 S258 I259 P260 R266 L267 C268 T269 K270 S283 M287

N288 L289 A298 I301 M304 S309 L321 L329 F348 E351 T354 T355 R358 Y363 I364 I365 D375 V380 T387 G388 S389 D390 L393 G394 F395 Q396 E397 THR PRO ALA LEU PRO LEU

- Molecule 1: Mandelate racemase / muconate lactonizing enzyme, C-terminal domain protein

Chain D: 75% 19% · ·

SER M1 D5 T8 P14 G20 R21 Y22 V26 G38 E39 F45 G46 P47 K52 W74 Y82 T83 Q84 R85 T89 V90 M91 L94 M99 W102 E115 L116 L117 G118 G119 K120 V121 R126 S127 Y130 P133 T134 ASP GLY ASP VAL

TYR PRO ASP ASP K144 D152 K170 F171 D172 P173 A174 D181 G182 H183 Q184 Q199 I200 L210 L211 T214 F218 G222 R225 L236 W237 F238 E252 Y256 T257 S258 R266 T269 E272 V276 L277 E278 A281 L285 Q286 M287 N288

L289 A298 H308 V322 A323 L324 L329 S333 P334 V338 T344 F348 T354 T355 W359 S367 H374 E378 R382 T387 H392 E397 THR PRO ALA LEU PRO LEU

- Molecule 1: Mandelate racemase / muconate lactonizing enzyme, C-terminal domain protein

Chain E: 72% 18% · 8%

SER M1 R2 Y22 C32 G38 E39 I40 Y41 T44 F45 G46 P47 D48 L49 G64 H68 H69 T83 Q84 R85 P86 D87 V88 T89 V90 M91 G92 V93 E98 M99 W102 I105 Y114 E115 L116 L117 G118 G119 K120 Y130 L131 Y132 PRO THR ASP GLY ASP

VAL TYR PRO ASP PRO ASP LYS P145 K170 T178 Q184 P185 E191 L211 F212 G213 T214 P235 L236 W237 F238 E239 E248 V253 T257 S258 I259 T263 R266 K270 Y271 E272 F273 S274 R275 V276 L277 E278 T279 G280 A281 M287 K299 K300 I301 A302

A303 M304 A305 E306 S309 C332 S333 P334 L339 E340 S341 I342 ARG THR PHE ASP G347 T354 T355 I364 S367 Q368 E378 R382 T387 GLY SER ASP LEU HIS LEU GLY PHE GLN GLU THR PRO ALA LEU PRO LEU
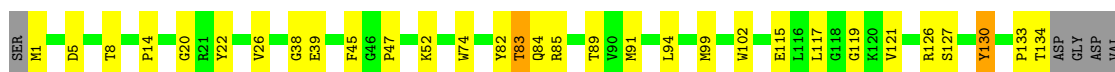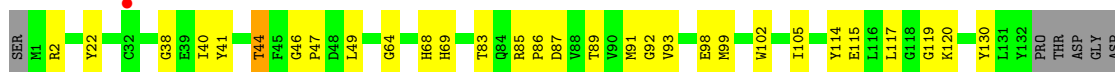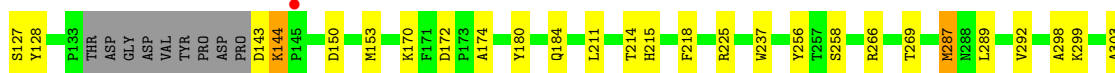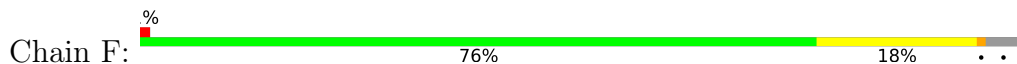
- Molecule 1: Mandelate racemase / muconate lactonizing enzyme, C-terminal domain protein

Chain F: % 76% 18% · ·

SER M1 R2 E7 T8 F9 G20 F25 V26 K27 L28 V29 G33 I34 I35 G36 I40 T44 L49 M53 F58 F62 L73 W74 Y78 T83 Q84 R85 P86 M91 G92 V93 L94 S95 E98 M99 I104 E115 L116 L117 G118 G119

S127 Y128 P133 THR ASP GLY VAL TYR PRO ASP PRO D143 K144 P145 D150 M153 K170 F171 D172 P173 A174 Y180 Q184 L211 T214 H215 F218 R225 W237 Y256 T257 S258 R266 T269 M287 N288 L289 V292 A298 K299 A303

M304 A305 E306 C307 H308 S309 I312 H315 C318 G319 P320 L321 L324 L329 T344 E351 L352 L353 T354 THR PRO I357 E378 D379 V380 A381 R382 A383 N384 H392 E397 THR PRO ALA LEU PRO LEU

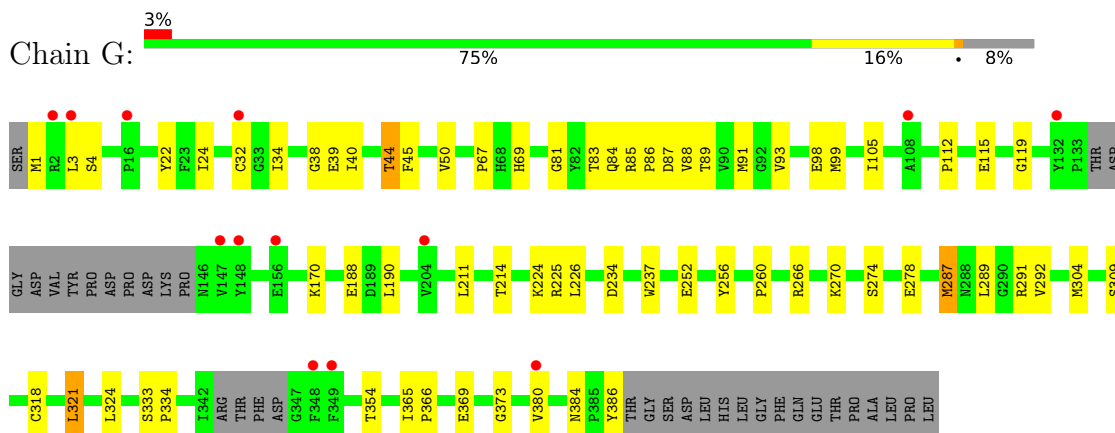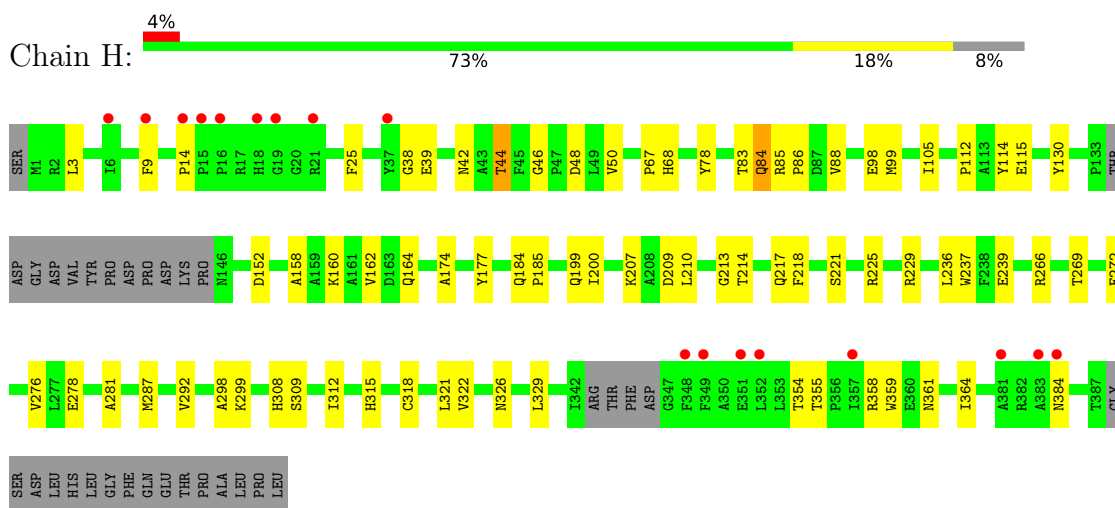W O R L D W I D E
wwPDB
PROTEIN DATA BANK

• Molecule 1: Mandelate racemase / muconate lactonizing enzyme, C-terminal domain protein

Chain G:

• Molecule 1: Mandelate racemase / muconate lactonizing enzyme, C-terminal domain protein

Chain H:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	98.90Å 156.88Å 218.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.45 – 2.90 49.45 – 2.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.45-2.90) 100.0 (49.45-2.90)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.70 (at 2.91Å)	Xtriage
Refinement program	PHENIX 1.7.3_928	Depositor
R, R_{free}	0.164 , 0.242 0.157 , 0.236	Depositor DCC
R_{free} test set	3817 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	55.7	Xtriage
Anisotropy	0.354	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 50.7	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	0.95	EDS
Total number of atoms	23681	wwPDB-VP
Average B, all atoms (Å^2)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.31% 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 ⓘ

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

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	0.45	0/3110	0.62	0/4227
1	B	0.42	0/3083	0.61	0/4187
1	C	0.41	0/3065	0.58	0/4163
1	D	0.41	0/3081	0.59	0/4185
1	E	0.44	0/2939	0.61	0/3991
1	F	0.37	0/3066	0.54	0/4161
1	G	0.38	0/2932	0.55	0/3982
1	H	0.36	0/2939	0.53	0/3992
All	All	0.41	0/24215	0.58	0/32888

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	3031	0	2966	50	0
1	B	3006	0	2936	49	0
1	C	2988	0	2917	48	0
1	D	3004	0	2936	47	0
1	E	2867	0	2809	48	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	2991	0	2918	48	0
1	G	2860	0	2801	41	0
1	H	2867	0	2808	43	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	A	8	0	0	0	0
3	B	9	0	0	0	0
3	C	8	0	0	0	0
3	D	10	0	0	0	0
3	E	9	0	0	0	0
3	F	5	0	0	0	0
3	G	6	0	0	0	0
3	H	4	0	0	0	0
All	All	23681	0	23091	321	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 7.

The worst 5 of 321 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:B:287:MET:HG2	1:B:292:VAL:HG21	1.56	0.86
1:B:85:ARG:NH1	1:B:86:PRO:O	2.12	0.83
1:E:85:ARG:NH1	1:E:86:PRO:O	2.14	0.80
1:B:38:GLY:HA2	1:B:99:MET:HE2	1.65	0.78
1:H:213:GLY:HA2	1:H:239:GLU:HB3	1.71	0.71

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	388/404 (96%)	372 (96%)	16 (4%)	0	100	100
1	B	384/404 (95%)	360 (94%)	22 (6%)	2 (0%)	29	61
1	C	382/404 (95%)	359 (94%)	22 (6%)	1 (0%)	41	71
1	D	384/404 (95%)	355 (92%)	26 (7%)	3 (1%)	19	51
1	E	365/404 (90%)	339 (93%)	26 (7%)	0	100	100
1	F	380/404 (94%)	359 (94%)	19 (5%)	2 (0%)	29	61
1	G	364/404 (90%)	341 (94%)	22 (6%)	1 (0%)	41	71
1	H	365/404 (90%)	345 (94%)	19 (5%)	1 (0%)	41	71
All	All	3012/3232 (93%)	2830 (94%)	172 (6%)	10 (0%)	41	71

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	269	THR
1	C	69	HIS
1	B	133	PRO
1	D	392	HIS
1	F	144	LYS

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	311/322 (97%)	294 (94%)	17 (6%)	21	53

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	308/322 (96%)	301 (98%)	7 (2%)	50	80
1	C	306/322 (95%)	287 (94%)	19 (6%)	18	47
1	D	308/322 (96%)	288 (94%)	20 (6%)	17	45
1	E	293/322 (91%)	273 (93%)	20 (7%)	16	42
1	F	306/322 (95%)	291 (95%)	15 (5%)	25	57
1	G	292/322 (91%)	280 (96%)	12 (4%)	30	64
1	H	293/322 (91%)	277 (94%)	16 (6%)	21	53
All	All	2417/2576 (94%)	2291 (95%)	126 (5%)	23	55

5 of 126 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	287	MET
1	H	42	ASN
1	E	238	PHE
1	G	321	LEU
1	H	237	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	368	GLN
1	D	311	GLN
1	F	184	GLN
1	H	384	ASN

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 ⓘ

Of 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	392/404 (97%)	-0.53	1 (0%)	94	94	23, 42, 64, 97	0
1	B	387/404 (95%)	-0.46	2 (0%)	91	91	24, 45, 80, 113	0
1	C	386/404 (95%)	-0.51	4 (1%)	82	82	30, 49, 80, 127	0
1	D	388/404 (96%)	-0.50	0	100	100	25, 52, 74, 101	0
1	E	371/404 (91%)	-0.50	1 (0%)	94	94	24, 44, 78, 117	0
1	F	386/404 (95%)	-0.21	4 (1%)	82	82	34, 64, 94, 116	0
1	G	370/404 (91%)	-0.10	13 (3%)	44	38	35, 66, 97, 119	0
1	H	371/404 (91%)	-0.07	17 (4%)	32	29	37, 68, 104, 119	0
All	All	3051/3232 (94%)	-0.36	42 (1%)	75	75	23, 53, 91, 127	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	390	ASP	4.1
1	H	16	PRO	4.0
1	H	349	PHE	3.9
1	H	15	PRO	3.7
1	G	16	PRO	3.3

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 (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(Å^2)	Q<0.9
2	MG	B	501	1/1	0.92	0.26	69,69,69,69	0
2	MG	G	501	1/1	0.92	0.23	93,93,93,93	0
2	MG	D	501	1/1	0.93	0.34	68,68,68,68	0
2	MG	E	501	1/1	0.95	0.34	90,90,90,90	0
2	MG	A	501	1/1	0.96	0.32	73,73,73,73	0
2	MG	H	501	1/1	0.96	0.17	75,75,75,75	0
2	MG	C	501	1/1	0.97	0.35	67,67,67,67	0
2	MG	F	501	1/1	0.97	0.24	83,83,83,83	0

6.5 Other polymers (i)

There are no such residues in this entry.