wwPDB X-ray Structure Validation Summary Report (i)

Oct 3, 2026 – 10:22 AM EDT

PDB ID	:	9ZFX / pdb_00009zfx
Title	:	Mitochondrial malate dehydrogenase from Acanthamoeba castellanii
Authors	:	Fry, M.Y.; Luce, B.E.; Mootha, V.K.; Chao, L.H.
Deposited on	:	2025-12-02
Resolution	:	2.33 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 (1)) were used in the production of this report:

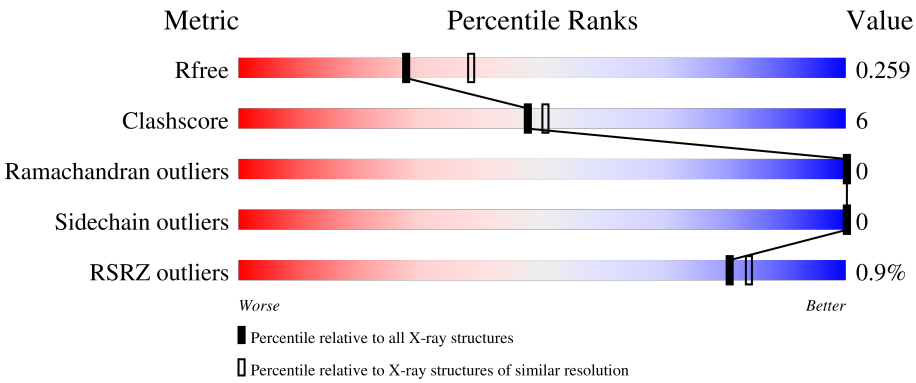
MolProbity	:	4-5-2 with Phenix2.2
Xtriage (Phenix)	:	2.2
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.51

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.33 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}	180053	3031 (2.36-2.32)
Clashscore	190562	3127 (2.36-2.32)
Ramachandran outliers	187476	3095 (2.36-2.32)
Sidechain outliers	187428	3095 (2.36-2.32)
RSRZ outliers	180081	3033 (2.36-2.32)

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	454	66% 7% 27%
1	B	454	% 62% 11% 27%
1	C	454	% 65% 9% 26%
1	D	454	63% 11% 26%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 20552 atoms, of which 10117 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 malate dehydrogenase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	D	335	5079	1595	2536	455	479	14	0	0	0
1	A	332	5043	1582	2522	452	473	14	0	0	0
1	B	332	5041	1583	2518	452	474	14	0	0	0
1	C	336	5089	1598	2541	456	480	14	0	0	0

There are 164 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	217	VAL	SER	conflict	UNP L8HFT9
D	218	GLN	ARG	conflict	UNP L8HFT9
D	219	GLY	ALA	conflict	UNP L8HFT9
D	220	LYS	-	insertion	UNP L8HFT9
D	221	ALA	-	insertion	UNP L8HFT9
D	222	LEU	-	insertion	UNP L8HFT9
D	223	ASN	-	insertion	UNP L8HFT9
D	224	LYS	-	insertion	UNP L8HFT9
D	225	VAL	-	insertion	UNP L8HFT9
D	226	ALA	-	insertion	UNP L8HFT9
D	227	ASN	-	insertion	UNP L8HFT9
D	228	ARG	-	insertion	UNP L8HFT9
D	229	ASP	-	insertion	UNP L8HFT9
D	230	THR	-	insertion	UNP L8HFT9
D	231	LEU	-	insertion	UNP L8HFT9
D	232	ARG	-	insertion	UNP L8HFT9
D	233	VAL	-	insertion	UNP L8HFT9
D	234	CYS	-	insertion	UNP L8HFT9
D	235	VAL	-	insertion	UNP L8HFT9
D	236	VAL	-	insertion	UNP L8HFT9
D	237	GLY	-	insertion	UNP L8HFT9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	238	ASN	-	insertion	UNP L8HFT9
D	239	PRO	-	insertion	UNP L8HFT9
D	240	ALA	-	insertion	UNP L8HFT9
D	241	ASN	-	insertion	UNP L8HFT9
D	242	THR	-	insertion	UNP L8HFT9
D	243	ASN	-	insertion	UNP L8HFT9
D	244	ALA	-	insertion	UNP L8HFT9
D	245	LEU	-	insertion	UNP L8HFT9
D	246	ILE	-	insertion	UNP L8HFT9
D	247	ALA	-	insertion	UNP L8HFT9
D	248	SER	-	insertion	UNP L8HFT9
D	249	ALA	-	insertion	UNP L8HFT9
D	250	ASN	-	insertion	UNP L8HFT9
D	251	ALA	-	insertion	UNP L8HFT9
D	252	PRO	-	insertion	UNP L8HFT9
D	253	ASN	-	insertion	UNP L8HFT9
D	254	ILE	-	insertion	UNP L8HFT9
D	255	HIS	-	insertion	UNP L8HFT9
D	256	PRO	-	insertion	UNP L8HFT9
D	257	ARG	-	insertion	UNP L8HFT9
A	217	VAL	SER	conflict	UNP L8HFT9
A	218	GLN	ARG	conflict	UNP L8HFT9
A	219	GLY	ALA	conflict	UNP L8HFT9
A	220	LYS	-	insertion	UNP L8HFT9
A	221	ALA	-	insertion	UNP L8HFT9
A	222	LEU	-	insertion	UNP L8HFT9
A	223	ASN	-	insertion	UNP L8HFT9
A	224	LYS	-	insertion	UNP L8HFT9
A	225	VAL	-	insertion	UNP L8HFT9
A	226	ALA	-	insertion	UNP L8HFT9
A	227	ASN	-	insertion	UNP L8HFT9
A	228	ARG	-	insertion	UNP L8HFT9
A	229	ASP	-	insertion	UNP L8HFT9
A	230	THR	-	insertion	UNP L8HFT9
A	231	LEU	-	insertion	UNP L8HFT9
A	232	ARG	-	insertion	UNP L8HFT9
A	233	VAL	-	insertion	UNP L8HFT9
A	234	CYS	-	insertion	UNP L8HFT9
A	235	VAL	-	insertion	UNP L8HFT9
A	236	VAL	-	insertion	UNP L8HFT9
A	237	GLY	-	insertion	UNP L8HFT9
A	238	ASN	-	insertion	UNP L8HFT9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	239	PRO	-	insertion	UNP L8HFT9
A	240	ALA	-	insertion	UNP L8HFT9
A	241	ASN	-	insertion	UNP L8HFT9
A	242	THR	-	insertion	UNP L8HFT9
A	243	ASN	-	insertion	UNP L8HFT9
A	244	ALA	-	insertion	UNP L8HFT9
A	245	LEU	-	insertion	UNP L8HFT9
A	246	ILE	-	insertion	UNP L8HFT9
A	247	ALA	-	insertion	UNP L8HFT9
A	248	SER	-	insertion	UNP L8HFT9
A	249	ALA	-	insertion	UNP L8HFT9
A	250	ASN	-	insertion	UNP L8HFT9
A	251	ALA	-	insertion	UNP L8HFT9
A	252	PRO	-	insertion	UNP L8HFT9
A	253	ASN	-	insertion	UNP L8HFT9
A	254	ILE	-	insertion	UNP L8HFT9
A	255	HIS	-	insertion	UNP L8HFT9
A	256	PRO	-	insertion	UNP L8HFT9
A	257	ARG	-	insertion	UNP L8HFT9
B	217	VAL	SER	conflict	UNP L8HFT9
B	218	GLN	ARG	conflict	UNP L8HFT9
B	219	GLY	ALA	conflict	UNP L8HFT9
B	220	LYS	-	insertion	UNP L8HFT9
B	221	ALA	-	insertion	UNP L8HFT9
B	222	LEU	-	insertion	UNP L8HFT9
B	223	ASN	-	insertion	UNP L8HFT9
B	224	LYS	-	insertion	UNP L8HFT9
B	225	VAL	-	insertion	UNP L8HFT9
B	226	ALA	-	insertion	UNP L8HFT9
B	227	ASN	-	insertion	UNP L8HFT9
B	228	ARG	-	insertion	UNP L8HFT9
B	229	ASP	-	insertion	UNP L8HFT9
B	230	THR	-	insertion	UNP L8HFT9
B	231	LEU	-	insertion	UNP L8HFT9
B	232	ARG	-	insertion	UNP L8HFT9
B	233	VAL	-	insertion	UNP L8HFT9
B	234	CYS	-	insertion	UNP L8HFT9
B	235	VAL	-	insertion	UNP L8HFT9
B	236	VAL	-	insertion	UNP L8HFT9
B	237	GLY	-	insertion	UNP L8HFT9
B	238	ASN	-	insertion	UNP L8HFT9
B	239	PRO	-	insertion	UNP L8HFT9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	240	ALA	-	insertion	UNP L8HFT9
B	241	ASN	-	insertion	UNP L8HFT9
B	242	THR	-	insertion	UNP L8HFT9
B	243	ASN	-	insertion	UNP L8HFT9
B	244	ALA	-	insertion	UNP L8HFT9
B	245	LEU	-	insertion	UNP L8HFT9
B	246	ILE	-	insertion	UNP L8HFT9
B	247	ALA	-	insertion	UNP L8HFT9
B	248	SER	-	insertion	UNP L8HFT9
B	249	ALA	-	insertion	UNP L8HFT9
B	250	ASN	-	insertion	UNP L8HFT9
B	251	ALA	-	insertion	UNP L8HFT9
B	252	PRO	-	insertion	UNP L8HFT9
B	253	ASN	-	insertion	UNP L8HFT9
B	254	ILE	-	insertion	UNP L8HFT9
B	255	HIS	-	insertion	UNP L8HFT9
B	256	PRO	-	insertion	UNP L8HFT9
B	257	ARG	-	insertion	UNP L8HFT9
C	217	VAL	SER	conflict	UNP L8HFT9
C	218	GLN	ARG	conflict	UNP L8HFT9
C	219	GLY	ALA	conflict	UNP L8HFT9
C	220	LYS	-	insertion	UNP L8HFT9
C	221	ALA	-	insertion	UNP L8HFT9
C	222	LEU	-	insertion	UNP L8HFT9
C	223	ASN	-	insertion	UNP L8HFT9
C	224	LYS	-	insertion	UNP L8HFT9
C	225	VAL	-	insertion	UNP L8HFT9
C	226	ALA	-	insertion	UNP L8HFT9
C	227	ASN	-	insertion	UNP L8HFT9
C	228	ARG	-	insertion	UNP L8HFT9
C	229	ASP	-	insertion	UNP L8HFT9
C	230	THR	-	insertion	UNP L8HFT9
C	231	LEU	-	insertion	UNP L8HFT9
C	232	ARG	-	insertion	UNP L8HFT9
C	233	VAL	-	insertion	UNP L8HFT9
C	234	CYS	-	insertion	UNP L8HFT9
C	235	VAL	-	insertion	UNP L8HFT9
C	236	VAL	-	insertion	UNP L8HFT9
C	237	GLY	-	insertion	UNP L8HFT9
C	238	ASN	-	insertion	UNP L8HFT9
C	239	PRO	-	insertion	UNP L8HFT9
C	240	ALA	-	insertion	UNP L8HFT9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	241	ASN	-	insertion	UNP L8HFT9
C	242	THR	-	insertion	UNP L8HFT9
C	243	ASN	-	insertion	UNP L8HFT9
C	244	ALA	-	insertion	UNP L8HFT9
C	245	LEU	-	insertion	UNP L8HFT9
C	246	ILE	-	insertion	UNP L8HFT9
C	247	ALA	-	insertion	UNP L8HFT9
C	248	SER	-	insertion	UNP L8HFT9
C	249	ALA	-	insertion	UNP L8HFT9
C	250	ASN	-	insertion	UNP L8HFT9
C	251	ALA	-	insertion	UNP L8HFT9
C	252	PRO	-	insertion	UNP L8HFT9
C	253	ASN	-	insertion	UNP L8HFT9
C	254	ILE	-	insertion	UNP L8HFT9
C	255	HIS	-	insertion	UNP L8HFT9
C	256	PRO	-	insertion	UNP L8HFT9
C	257	ARG	-	insertion	UNP L8HFT9

- Molecule 2 is water.

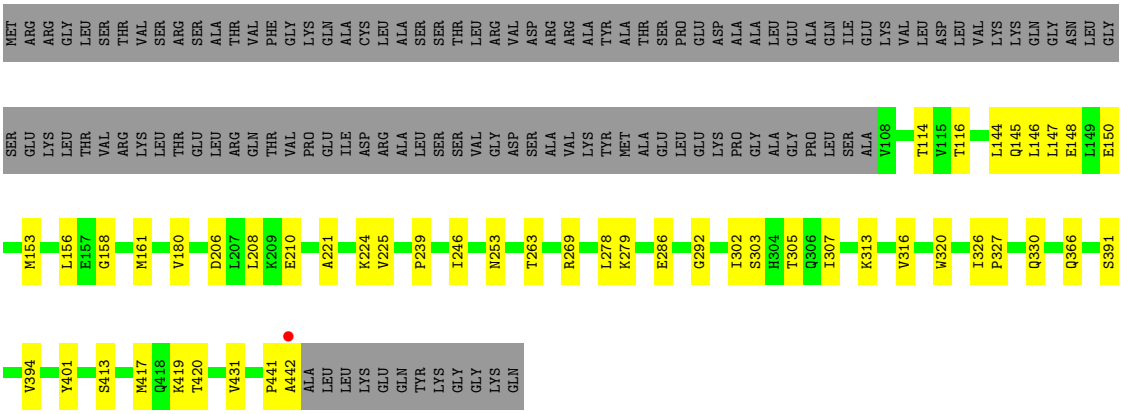
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	D	76	76	76	0	0
			Total	O		
2	A	74	74	74	0	0
			Total	O		
2	B	70	70	70	0	0
			Total	O		
2	C	80	80	80	0	0

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.

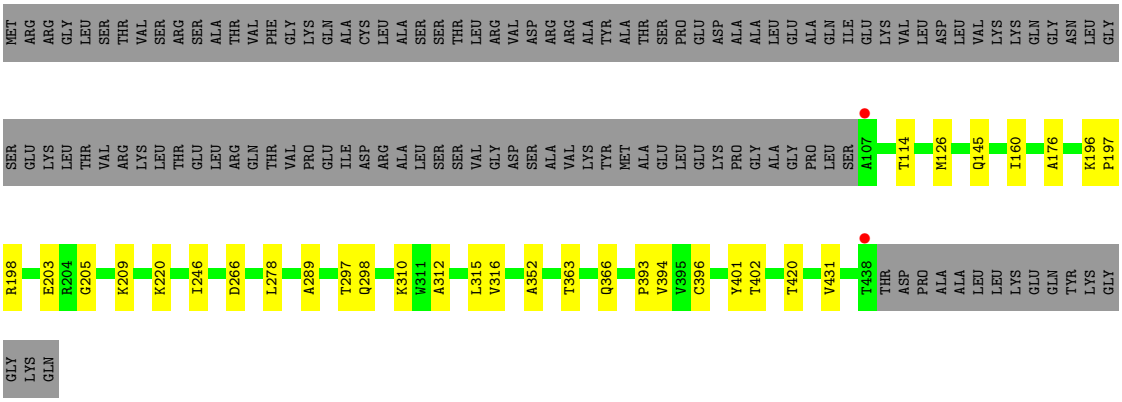
• Molecule 1: malate dehydrogenase

Chain D:



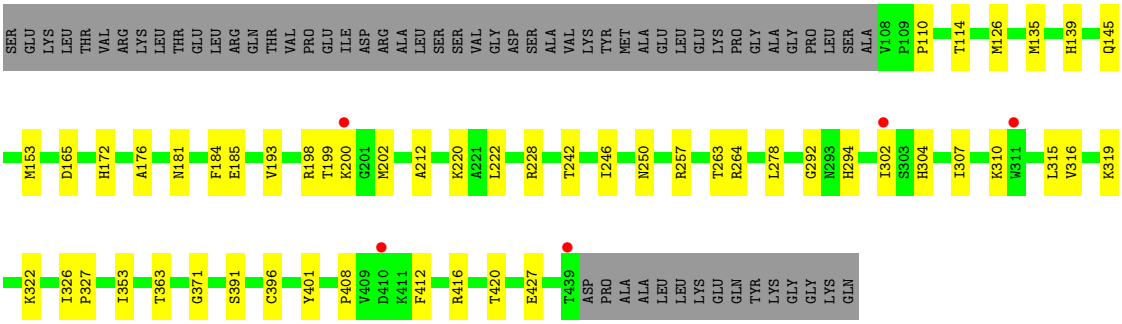
• Molecule 1: malate dehydrogenase

Chain A:



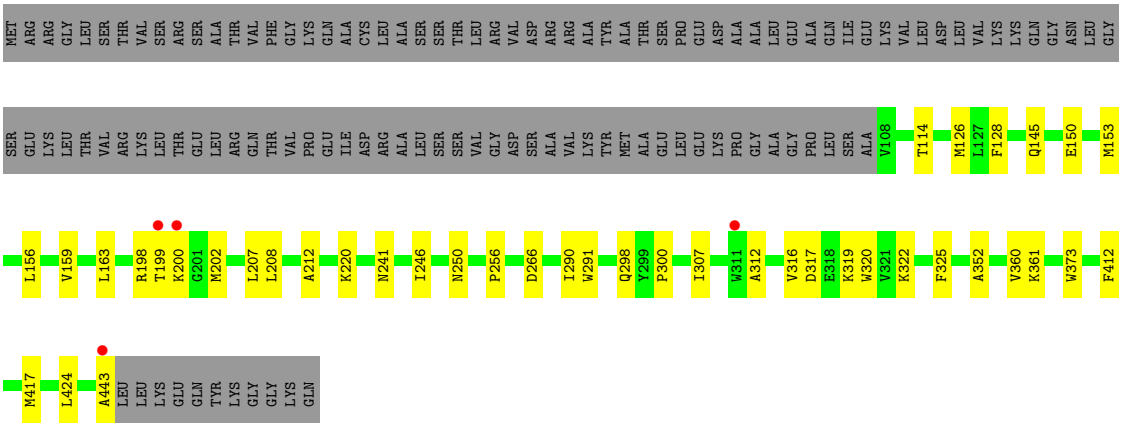
• Molecule 1: malate dehydrogenase

Chain B:



- Molecule 1: malate dehydrogenase

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	50.41Å 136.50Å 129.55Å 90.00° 100.85° 90.00°	Depositor
Resolution (Å)	34.54 – 2.33 34.54 – 2.33	Depositor EDS
% Data completeness (in resolution range)	96.2 (34.54-2.33) 96.2 (34.54-2.33)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.56 (at 2.10Å)	Xtriage
Refinement program	PHENIX (1.21.2_5419: ???)	Depositor
R, R_{free}	0.204 , 0.259 0.210 , 0.259	Depositor DCC
R_{free} test set	4765 reflections (4.73%)	wwPDB-VP
Wilson B-factor (Å^2)	27.9	Xtriage
Anisotropy	0.404	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 33.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	20552	wwPDB-VP
Average B, all atoms (Å^2)	35.0	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 42.65 % 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 1.9769e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

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	0.58	0/2569	0.73	0/3481
1	B	0.54	0/2571	0.68	0/3484
1	C	0.57	0/2597	0.73	0/3521
1	D	0.57	0/2592	0.71	0/3514
All	All	0.57	0/10329	0.71	0/14000

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 (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	2521	2522	2522	22	0
1	B	2523	2518	2524	37	0
1	C	2548	2541	2545	26	0
1	D	2543	2536	2540	31	0
2	A	74	0	0	0	0
2	B	70	0	0	7	0
2	C	80	0	0	0	0
2	D	76	0	0	1	0
All	All	10435	10117	10131	113	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 113 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:A:394:VAL:HG11	1:A:401:TYR:HB2	1.60	0.83
1:A:205:GLY:O	1:A:209:LYS:HG2	1.83	0.77
1:C:316:VAL:HG11	1:C:320:TRP:CE3	2.23	0.74
1:A:363:THR:CG2	1:A:396:CYS:H	2.01	0.74
1:C:199:THR:H	1:C:202:MET:HE3	1.56	0.71

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	330/454 (73%)	322 (98%)	8 (2%)	0	100	100
1	B	330/454 (73%)	324 (98%)	6 (2%)	0	100	100
1	C	334/454 (74%)	330 (99%)	4 (1%)	0	100	100
1	D	333/454 (73%)	325 (98%)	8 (2%)	0	100	100
All	All	1327/1816 (73%)	1301 (98%)	26 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	263/362 (73%)	263 (100%)	0	100	100
1	B	264/362 (73%)	264 (100%)	0	100	100
1	C	266/362 (74%)	266 (100%)	0	100	100
1	D	266/362 (74%)	266 (100%)	0	100	100
All	All	1059/1448 (73%)	1059 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

Mol	Chain	Res	Type
1	D	330	GLN
1	A	273	GLN
1	A	308	ASN
1	A	330	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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	332/454 (73%)	-0.23	2 (0%)	85	88	19, 32, 52, 64	0
1	B	332/454 (73%)	-0.13	5 (1%)	72	76	22, 36, 55, 77	0
1	C	336/454 (74%)	-0.23	4 (1%)	76	80	20, 32, 51, 86	0
1	D	335/454 (73%)	-0.22	1 (0%)	90	91	22, 32, 52, 68	0
All	All	1335/1816 (73%)	-0.20	12 (0%)	81	84	19, 33, 53, 86	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	443	ALA	4.5
1	B	311	TRP	3.6
1	D	442	ALA	3.6
1	C	311	TRP	3.2
1	A	107	ALA	3.1

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 oligosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.