



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 13, 2020 – 10:26 PM BST

PDB ID : 5VYJ
Title : Crystal structure of the photosynthetic phosphoenolpyruvate carboxylase isoenzyme from maize in complex with Gly
Authors : Gonzalez-Segura, L.; Guemez-Toro, R.; Munoz-Clares, R.A.
Deposited on : 2017-05-25
Resolution : 3.30 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 : 2.14.4.dev1
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) : 2.14.4.dev1

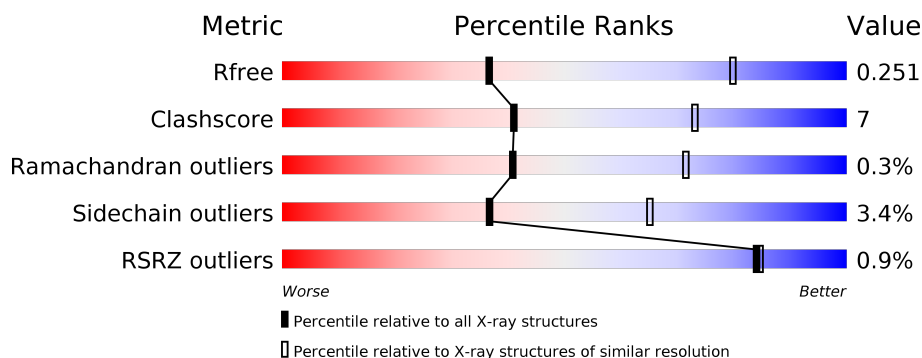
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 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	970	<div> <div>%</div> <div> <div></div> <div>73%</div> <div>20%</div> <div>• 5%</div> </div> </div>
1	B	970	<div> <div></div> <div>74%</div> <div>20%</div> <div>• 5%</div> </div>
1	C	970	<div> <div>%</div> <div> <div></div> <div>75%</div> <div>19%</div> <div>• 5%</div> </div> </div>
1	D	970	<div> <div>%</div> <div> <div></div> <div>76%</div> <div>18%</div> <div>• 5%</div> </div> </div>

2 Entry composition [i](#)

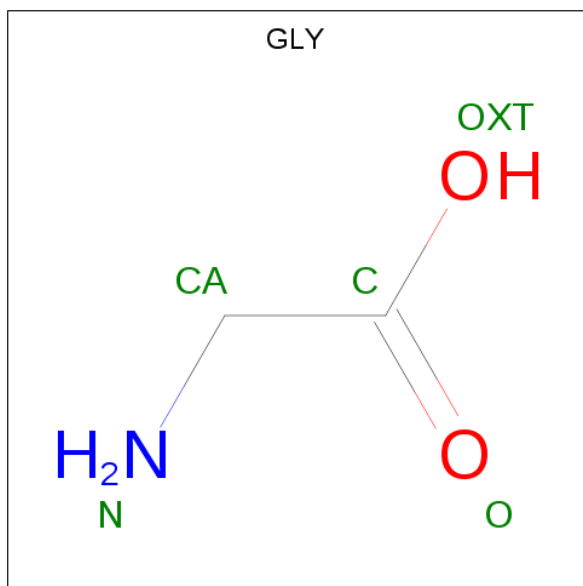
There are 3 unique types of molecules in this entry. The entry contains 29339 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Phosphoenolpyruvate carboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	918	Total	C	N	O	S	0	0	0
			7318	4646	1275	1367	30			
1	B	917	Total	C	N	O	S	0	0	0
			7309	4640	1274	1365	30			
1	C	918	Total	C	N	O	S	0	0	0
			7318	4645	1275	1368	30			
1	D	918	Total	C	N	O	S	0	0	0
			7317	4646	1275	1366	30			

- Molecule 2 is GLYCINE (three-letter code: GLY) (formula: $C_2H_5NO_2$).



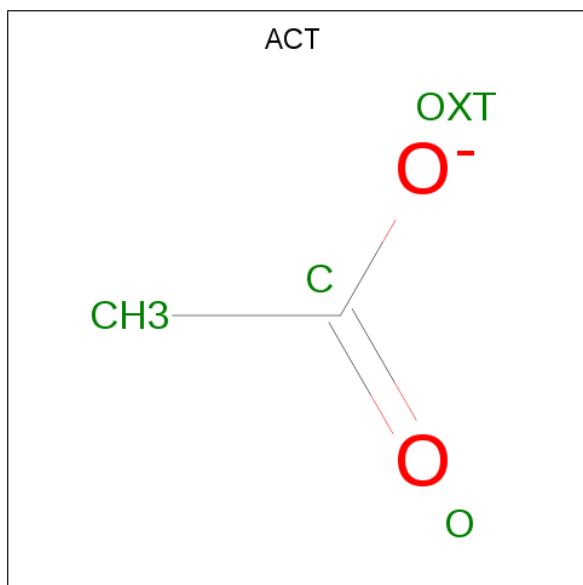
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			5	2	1	2		
2	B	1	Total	C	N	O	0	0
			5	2	1	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	N	O	0	0
			5	2	1	2		
2	C	1	Total	C	N	O	0	0
			5	2	1	2		
2	D	1	Total	C	N	O	0	0
			5	2	1	2		

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		
3	C	1	Total	C	O	0	0
			4	2	2		
3	C	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

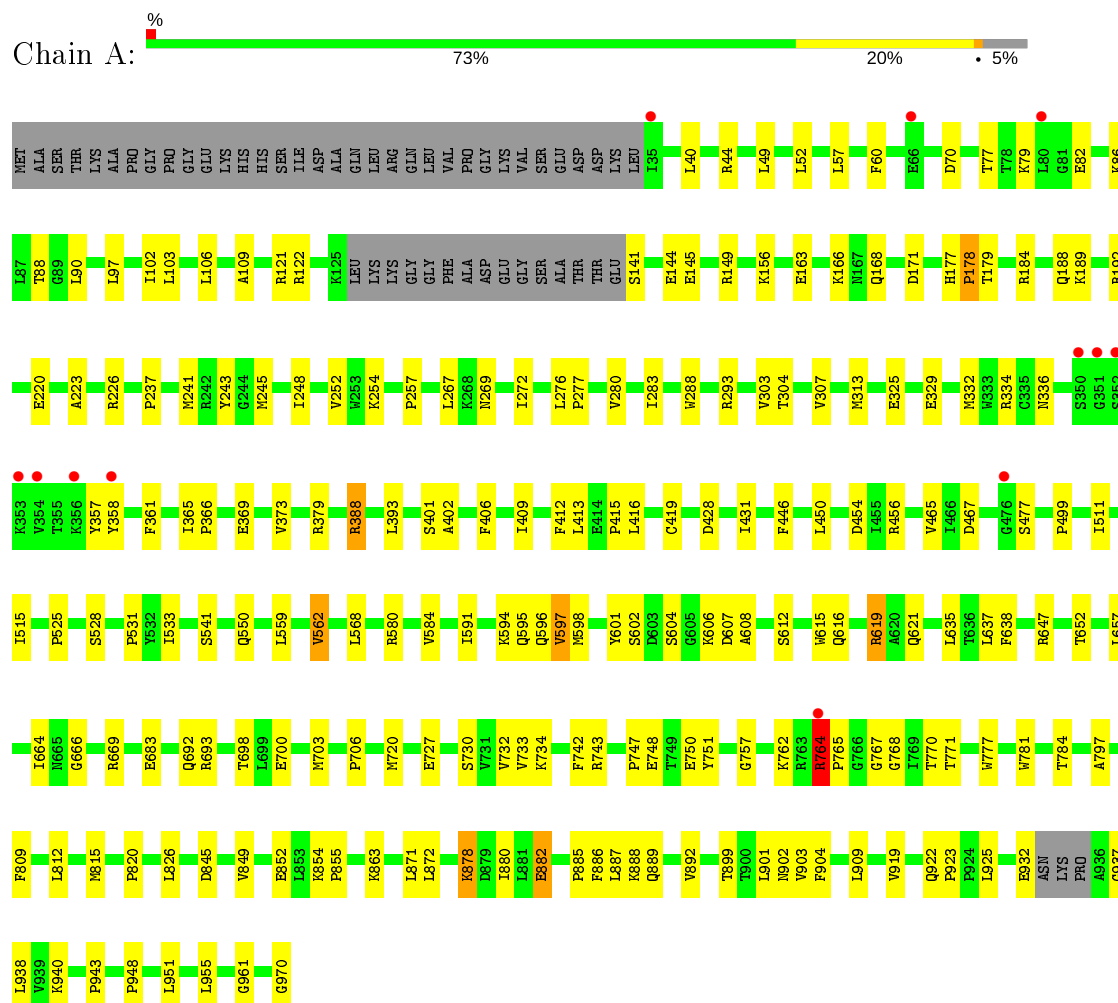
Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total 4	C 2	O 2	0	0
3	C	1	Total 4	C 2	O 2	0	0
3	D	1	Total 4	C 2	O 2	0	0
3	D	1	Total 4	C 2	O 2	0	0
3	D	1	Total 4	C 2	O 2	0	0

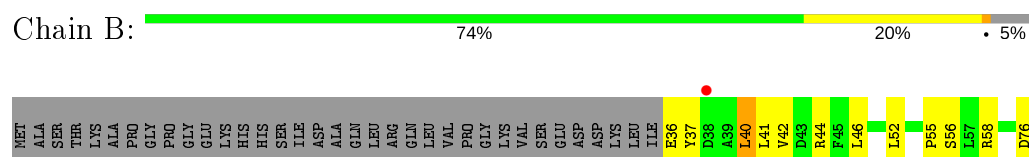
3 Residue-property plots [i](#)

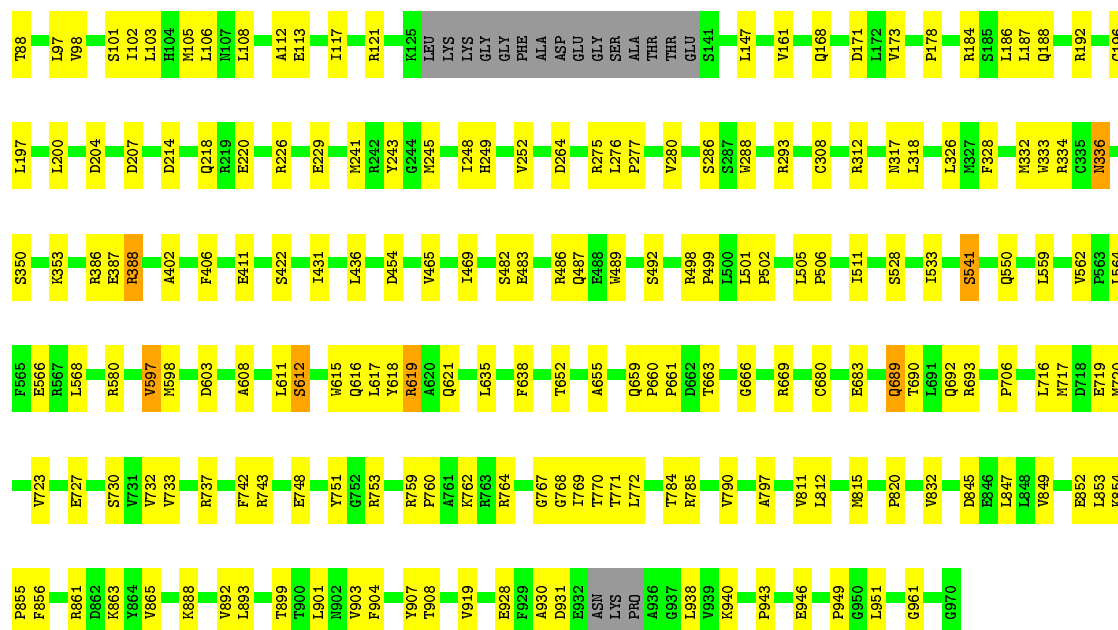
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Phosphoenolpyruvate carboxylase

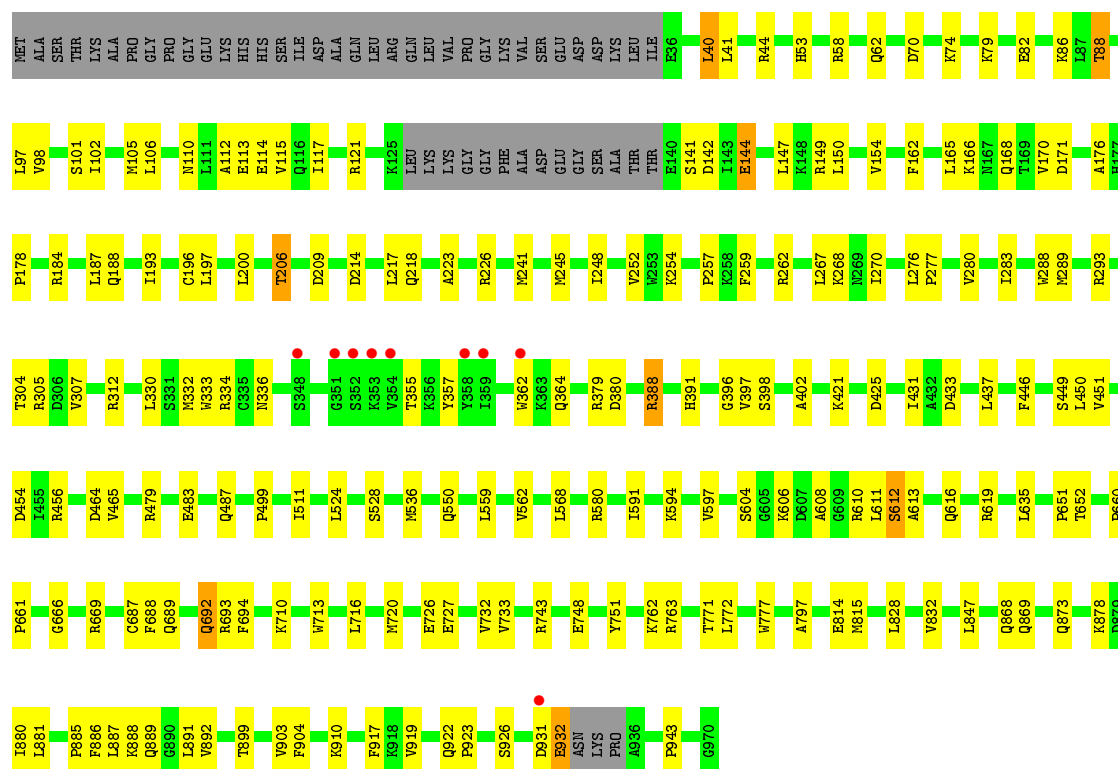


• Molecule 1: Phosphoenolpyruvate carboxylase





• Molecule 1: Phosphoenolpyruvate carboxylase



• Molecule 1: Phosphoenolpyruvate carboxylase



I898	Y751	P540	Y357	L213	V98	MET
L901	A761	S541	Y358	D214	I102	ALA
F904	K762	Q550	F361	Q218	THR	SER
T908	R763	L559	K362	R219	LYS	LYS
L909	R764	L564	K363	E220	N110	ALA
K910	P765	L568	Y371	A223	E113	PRO
F917	G766	Q572	R386	R226	E114	GLY
K918	T770	L581	E387	R241	I117	PRO
V919	T771	Q595	R388	K242	R121	GLY
T920	P776	L581	H591	Y243	K125	GLY
E928	W777	Q596	A394	G244	L126	ASP
E932	S780	Q597	S398	M245	LYS	ILE
ASN	W781	V597	S401	I248	LYS	ALA
LYS	T784	M598	A402	V252	GLY	ALA
PRO	V790	K606	S422	F259	PHE	LEU
A936	G793	R610	L423	I270	ALA	ARG
G937	W819	L611	D428	L276	ASP	GLN
L938	P820	S612	I431	F277	GLY	LEU
L941	V824	Q616	F446	V280	GLY	LEU
I956	T825	R619	L450	I283	VAL	VAL
M959	L828	L657	D454	S286	SER	PRO
G970	V832	Q659	L473	W288	GLU	GLY
	D845	R669	D484	R293	S141	PRO
	E852	L866	W489	T304	D142	GLY
	K854	Q689	S492	R312	I143	ASP
	P855	Q692	P499	A315	E144	ASP
	F856	P706	L500	F328	L147	LYS
	R861	T725	P502	K332	L150	LEU
	D862	E726	P503	R336	V151	ILE
	K863	E727	L505	D337	V154	E36
	K878	Y728	P506	V341	D171	Q50
	L881	R729	I511	S350	L172	D51
	E882	V732	I515	G351	V173	L52
	L887	V733	S528	K352	P178	P55
	K888	F742	R743	V354	R184	R53
	V892	R894	E748	K356	Q188	Q62
	L893				R192	E63
	R894				C196	V67
					L200	R74
					T206	L83
					T355	K86
					E212	L87
						T88
						I96
						L97

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	156.95Å 167.24Å 242.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	57.08 – 3.30 57.08 – 3.30	Depositor EDS
% Data completeness (in resolution range)	93.2 (57.08-3.30) 93.3 (57.08-3.30)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.87 (at 3.33Å)	Xtriage
Refinement program	PHENIX 1.8.1_1168	Depositor
R, R_{free}	0.210 , 0.253 0.212 , 0.251	Depositor DCC
R_{free} test set	4518 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	45.0	Xtriage
Anisotropy	0.079	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 29.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	29339	wwPDB-VP
Average B, all atoms (Å ²)	27.0	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 38.98 % 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 3.4006e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² Theoretical 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: ACT

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	0/7473	0.51	0/10119
1	B	0.35	0/7464	0.53	0/10108
1	C	0.33	0/7473	0.50	0/10120
1	D	0.33	0/7472	0.50	0/10119
All	All	0.34	0/29882	0.51	0/40466

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	7318	0	7314	114	0
1	B	7309	0	7303	114	0
1	C	7318	0	7309	100	0
1	D	7317	0	7314	98	0
2	A	5	0	2	1	0
2	B	10	0	4	1	0
2	C	5	0	2	1	0
2	D	5	0	2	0	0
3	A	12	0	9	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	12	0	9	0	0
3	C	16	0	12	0	0
3	D	12	0	9	1	0
All	All	29339	0	29289	414	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 414 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:388:ARG:HH12	1:B:402:ALA:HA	1.39	0.88
1:D:743:ARG:HH22	1:D:762:LYS:HE3	1.43	0.82
1:B:386:ARG:NH1	1:B:387:GLU:OE1	2.17	0.77
1:C:362:TRP:NE1	1:C:380:ASP:OD2	2.19	0.75
1:B:568:LEU:HD22	1:B:608:ALA:HB2	1.67	0.74

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	912/970 (94%)	855 (94%)	53 (6%)	4 (0%)	34	66
1	B	911/970 (94%)	863 (95%)	44 (5%)	4 (0%)	34	66
1	C	912/970 (94%)	867 (95%)	42 (5%)	3 (0%)	41	71
1	D	912/970 (94%)	857 (94%)	54 (6%)	1 (0%)	51	81
All	All	3647/3880 (94%)	3442 (94%)	193 (5%)	12 (0%)	41	71

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	943	PRO
1	B	943	PRO
1	C	330	LEU
1	A	764	ARG
1	A	943	PRO

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	791/832 (95%)	768 (97%)	23 (3%)	42	69
1	B	790/832 (95%)	762 (96%)	28 (4%)	36	64
1	C	791/832 (95%)	761 (96%)	30 (4%)	33	62
1	D	791/832 (95%)	763 (96%)	28 (4%)	36	64
All	All	3163/3328 (95%)	3054 (97%)	109 (3%)	37	65

5 of 109 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	904	PHE
1	C	336	ASN
1	D	619	ARG
1	B	940	LYS
1	C	88	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	211	GLN
1	D	692	GLN
1	D	336	ASN
1	C	180	GLN
1	D	53	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 ⓘ

18 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	GLY	D	1001	-	1,4,4	0.04	0	0,4,4	0.00	-
2	GLY	B	1001	-	1,4,4	0.05	0	0,4,4	0.00	-
3	ACT	C	1005	-	1,3,3	1.57	0	0,3,3	0.00	-
3	ACT	D	1003	-	1,3,3	1.35	0	0,3,3	0.00	-
3	ACT	D	1004	-	1,3,3	0.88	0	0,3,3	0.00	-
3	ACT	C	1002	-	1,3,3	1.61	0	0,3,3	0.00	-
3	ACT	A	1002	-	1,3,3	1.55	0	0,3,3	0.00	-
3	ACT	D	1002	-	1,3,3	2.08	1 (100%)	0,3,3	0.00	-
3	ACT	A	1003	-	1,3,3	1.55	0	0,3,3	0.00	-
3	ACT	B	1005	-	1,3,3	1.96	0	0,3,3	0.00	-
2	GLY	C	1001	-	1,4,4	0.05	0	0,4,4	0.00	-
2	GLY	A	1001	-	1,4,4	0.05	0	0,4,4	0.00	-
3	ACT	A	1004	-	1,3,3	1.97	0	0,3,3	0.00	-
3	ACT	B	1004	-	1,3,3	1.30	0	0,3,3	0.00	-
3	ACT	C	1004	-	1,3,3	1.32	0	0,3,3	0.00	-
3	ACT	C	1003	-	1,3,3	1.62	0	0,3,3	0.00	-
3	ACT	B	1003	-	1,3,3	1.79	0	0,3,3	0.00	-
2	GLY	B	1002	-	1,4,4	0.07	0	0,4,4	0.00	-

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	GLY	D	1001	-	-	0/0/2/2	-
2	GLY	B	1001	-	-	0/0/2/2	-
2	GLY	A	1001	-	-	0/0/2/2	-
2	GLY	C	1001	-	-	0/0/2/2	-
2	GLY	B	1002	-	-	0/0/2/2	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	1002	ACT	CH3-C	2.08	1.51	1.48

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	1004	ACT	1	0
3	A	1002	ACT	1	0
2	C	1001	GLY	1	0
2	A	1001	GLY	1	0
3	A	1004	ACT	1	0
2	B	1002	GLY	1	0

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 [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	918/970 (94%)	-0.10	12 (1%) 77 77	3, 24, 57, 85	0
1	B	917/970 (94%)	-0.23	1 (0%) 95 97	0, 17, 50, 88	0
1	C	918/970 (94%)	-0.13	9 (0%) 82 82	8, 24, 60, 101	0
1	D	918/970 (94%)	-0.04	12 (1%) 77 77	2, 30, 74, 97	0
All	All	3671/3880 (94%)	-0.12	34 (0%) 84 84	0, 23, 63, 101	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	351	GLY	4.5
1	D	352	SER	4.3
1	C	352	SER	3.9
1	D	351	GLY	3.7
1	D	361	PHE	3.5

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	ACT	C	1003	4/4	0.83	0.22	30,30,30,30	0
3	ACT	C	1005	4/4	0.84	0.27	37,37,37,37	0
3	ACT	D	1002	4/4	0.86	0.21	21,21,21,21	0
3	ACT	A	1003	4/4	0.89	0.24	40,40,40,40	0
2	GLY	B	1001	5/5	0.89	0.23	22,22,22,22	0
3	ACT	D	1003	4/4	0.90	0.28	39,39,39,39	0
3	ACT	D	1004	4/4	0.91	0.24	22,22,22,22	0
2	GLY	C	1001	5/5	0.92	0.22	18,18,18,18	0
3	ACT	B	1005	4/4	0.92	0.21	11,11,11,11	0
3	ACT	B	1003	4/4	0.92	0.23	11,11,11,11	0
2	GLY	B	1002	5/5	0.92	0.19	8,8,8,8	0
2	GLY	D	1001	5/5	0.93	0.18	8,8,8,8	0
3	ACT	C	1002	4/4	0.94	0.28	17,17,17,17	0
3	ACT	A	1002	4/4	0.94	0.18	16,16,16,16	0
2	GLY	A	1001	5/5	0.94	0.20	9,9,9,9	0
3	ACT	C	1004	4/4	0.95	0.22	17,17,17,17	0
3	ACT	A	1004	4/4	0.95	0.14	20,20,20,20	0
3	ACT	B	1004	4/4	0.97	0.16	10,10,10,10	0

6.5 Other polymers ⓘ

There are no such residues in this entry.