wwPDB X-ray Structure Validation Summary Report (i)

May 21, 2020 – 11:59 pm BST

PDB ID	:	3TW7
Title	:	Structure of Rhizobium etli pyruvate carboxylase T882A crystallized without acetyl coenzyme-A
Authors	:	St Maurice, M.; Kumar, S.; Lietzan, A.D.
Deposited on	:	2011-09-21
Resolution	:	3.10 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 following versions of software and data (see references (1)) were used in the production of this report:

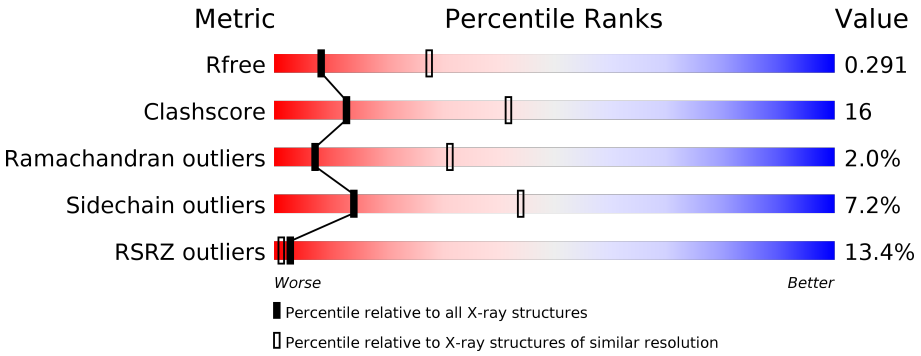
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 <=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	1165	13% 55% 28% . 14%
1	B	1165	10% 58% 25% . 14%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 15185 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 Pyruvate carboxylase protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1004	Total	C	N	O	S	0	15	0
			7544	4804	1275	1434	31			
1	B	1002	Total	C	N	O	S	0	16	0
			7623	4853	1290	1449	31			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-10	MET	-	EXPRESSION TAG	UNP Q2K340
A	-9	HIS	-	EXPRESSION TAG	UNP Q2K340
A	-8	HIS	-	EXPRESSION TAG	UNP Q2K340
A	-7	HIS	-	EXPRESSION TAG	UNP Q2K340
A	-6	HIS	-	EXPRESSION TAG	UNP Q2K340
A	-5	HIS	-	EXPRESSION TAG	UNP Q2K340
A	-4	HIS	-	EXPRESSION TAG	UNP Q2K340
A	-3	HIS	-	EXPRESSION TAG	UNP Q2K340
A	-2	HIS	-	EXPRESSION TAG	UNP Q2K340
A	-1	HIS	-	EXPRESSION TAG	UNP Q2K340
A	0	GLY	-	EXPRESSION TAG	UNP Q2K340
A	1	GLY	-	EXPRESSION TAG	UNP Q2K340
A	882	ALA	THR	ENGINEERED MUTATION	UNP Q2K340
B	-10	MET	-	EXPRESSION TAG	UNP Q2K340
B	-9	HIS	-	EXPRESSION TAG	UNP Q2K340
B	-8	HIS	-	EXPRESSION TAG	UNP Q2K340
B	-7	HIS	-	EXPRESSION TAG	UNP Q2K340
B	-6	HIS	-	EXPRESSION TAG	UNP Q2K340
B	-5	HIS	-	EXPRESSION TAG	UNP Q2K340
B	-4	HIS	-	EXPRESSION TAG	UNP Q2K340
B	-3	HIS	-	EXPRESSION TAG	UNP Q2K340
B	-2	HIS	-	EXPRESSION TAG	UNP Q2K340
B	-1	HIS	-	EXPRESSION TAG	UNP Q2K340
B	0	GLY	-	EXPRESSION TAG	UNP Q2K340
B	1	GLY	-	EXPRESSION TAG	UNP Q2K340

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	882	ALA	THR	ENGINEERED MUTATION	UNP Q2K340

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Mg 1 1	0	0
2	A	1	Total Mg 1 1	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Zn 1 1	0	0
3	A	1	Total Zn 1 1	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Cl 1 1	0	0
4	A	1	Total Cl 1 1	0	0

- Molecule 5 is water.

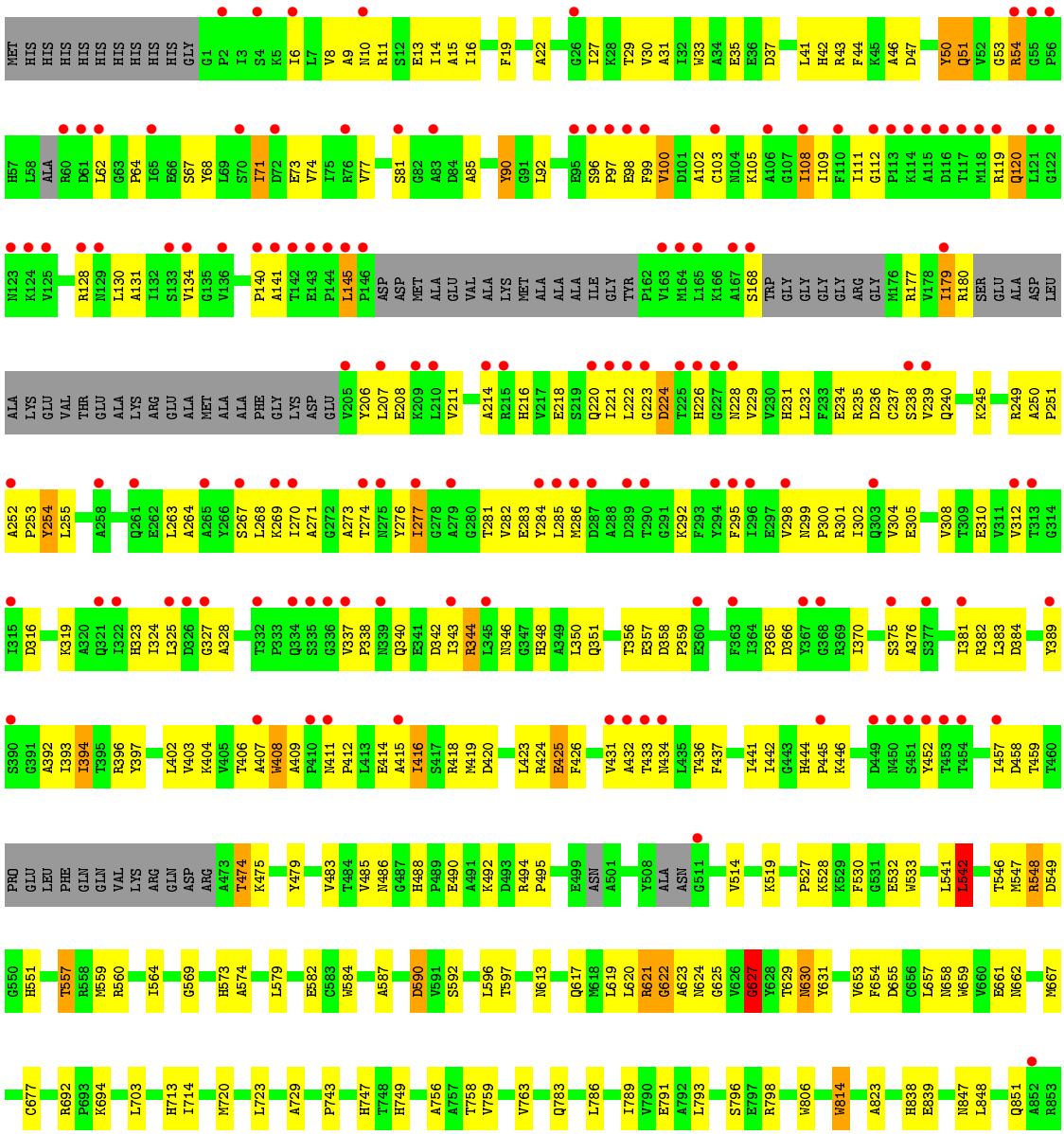
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	7	Total O 7 7	0	0
5	B	5	Total O 5 5	0	0

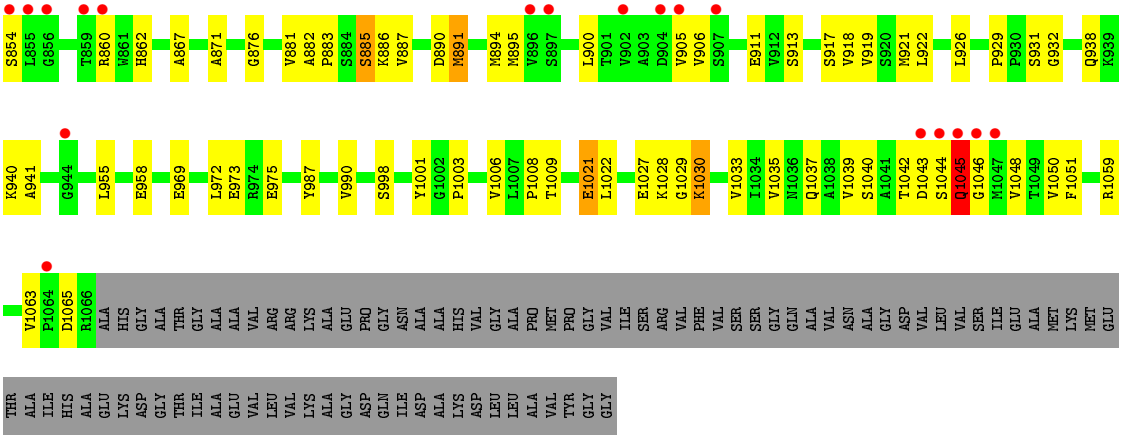
3 Residue-property plots (i)

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: Pyruvate carboxylase protein

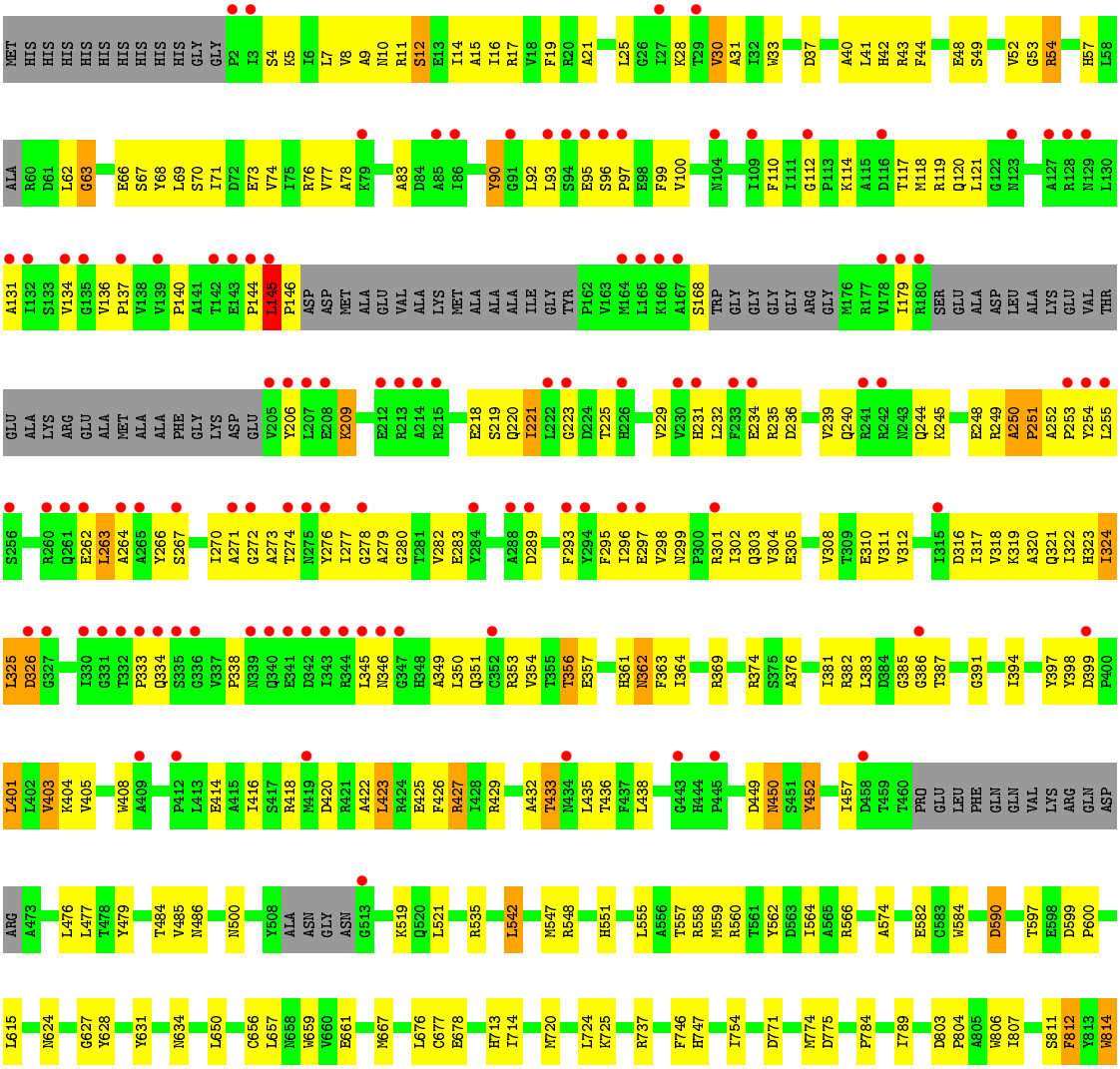
Chain A:

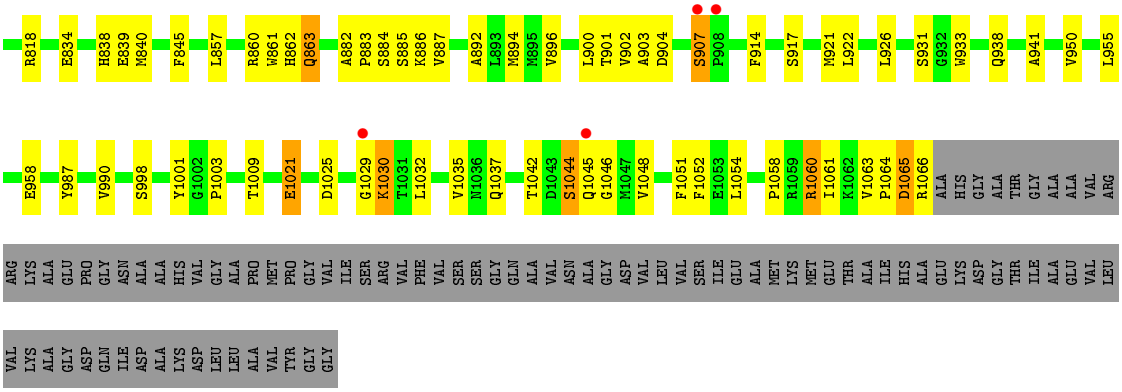




• Molecule 1: Pyruvate carboxylase protein

Chain B:




R818 E834 H838 E839 M840 F845 L857 R860 W861 H862 Q863 A882 P883 S884 S885 K886 V887 A892 L893 M894 M895 V896 L900 T901 V902 A903 D904 S907 P908 F914 S917 M921 L922 L926 S931 G932 W933 Q938 A941 V950 L955
E958 Y987 V990 S998 Y1001 G1002 P1003 T1009 E1021 D1025 G1029 K1030 T1031 L1032 V1035 N1036 Q1037 T1042 D1043 S1044 Q1045 G1046 M1047 V1048 F1051 F1052 E1053 L1054 P1058 R1059 R1060 I1061 K1062 V1063 P1064 D1065 R1066 ALA HIS GLY ALA THR GLY ALA ALA VAL ARG
ARG LYS ALA GLU PRO GLY ASN ALA ALA HIS VAL GLY ALA PRO MET PRO GLY VAL ILE SER ARG VAL PHE VAL SER SER GLY GLN ALA VAL ASN ALA ALA GLY ASP VAL LEU VAL SER ILE GLU ALA MET LYS MET GLU THR ALA ILE HIS ALA GLU LYS ASP GLY THR ILE ALA GLU VAL LEU
VAL LYS ALA GLY ASP GLN ILE ASP ALA LYS ASP LEU LEU ALA VAL TYR GLY GLY

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, α, β, γ	264.16Å 264.16Å 91.79Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.10 45.30 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (50.00-3.10) 100.0 (45.30-3.10)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	8.13 (at 3.12Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.250 , 0.292 0.248 , 0.291	Depositor DCC
R_{free} test set	2964 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å^2)	54.7	Xtriage
Anisotropy	0.039	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 66.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.023 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	15185	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.61% 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, ZN, KCX, CL

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	0.69	8/7719 (0.1%)	0.79	7/10500 (0.1%)
1	B	0.70	2/7799 (0.0%)	0.84	6/10603 (0.1%)
All	All	0.70	10/15518 (0.1%)	0.82	13/21103 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
All	All	0	2

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	98[A]	GLU	CD-OE1	7.78	1.34	1.25
1	A	98[B]	GLU	CD-OE1	7.78	1.34	1.25
1	A	269	LYS	CE-NZ	7.16	1.67	1.49
1	A	1045	GLN	CD-NE2	6.86	1.50	1.32
1	A	1045	GLN	CG-CD	6.82	1.66	1.51

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	145	LEU	CA-CB-CG	7.13	131.69	115.30
1	A	548	ARG	NE-CZ-NH2	6.72	123.66	120.30
1	A	542	LEU	CA-CB-CG	-6.47	100.42	115.30
1	A	798	ARG	NE-CZ-NH2	-6.16	117.22	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	542	LEU	CA-CB-CG	-6.16	101.13	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1046	GLY	Peptide
1	B	1065	ASP	Peptide

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	7544	0	7260	247	0
1	B	7623	0	7392	249	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	7	0	0	0	0
5	B	5	0	0	0	0
All	All	15185	0	14652	492	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 16.

The worst 5 of 492 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:4:SER:O	1:B:28:LYS:HG3	1.53	1.08
1:A:359:PRO:HD3	1:A:433:THR:O	1.55	1.06
1:A:90:TYR:HB2	1:A:301:ARG:HH11	1.21	1.05
1:A:1029:GLY:HA3	1:A:1030:LYS:HB2	1.39	1.04
1:B:1029:GLY:CA	1:B:1030:LYS:HB2	1.88	1.04

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	1001/1165 (86%)	875 (87%)	107 (11%)	19 (2%)	8 33
1	B	1002/1165 (86%)	899 (90%)	82 (8%)	21 (2%)	7 30
All	All	2003/2330 (86%)	1774 (89%)	189 (9%)	40 (2%)	7 31

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	92	LEU
1	A	622	GLY
1	A	623	ALA
1	A	1030	LYS
1	B	92	LEU

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	749/933 (80%)	699 (93%)	50 (7%)	16 46
1	B	769/933 (82%)	709 (92%)	60 (8%)	12 40
All	All	1518/1866 (81%)	1408 (93%)	110 (7%)	14 44

5 of 110 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1045	GLN
1	B	219	SER
1	B	907[B]	SER
1	B	30	VAL
1	B	120	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	10	ASN
1	B	220	GLN
1	B	873	GLN
1	B	57	HIS
1	B	120	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)

2 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
1	KCX	A	718[A]	1,3	7,11,12	0.70	0	4,12,14	1.28	1 (25%)
1	KCX	B	718[A]	1,3	7,11,12	0.93	0	4,12,14	0.71	0

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
1	KCX	A	718[A]	1,3	-	1/7/10/12	-
1	KCX	B	718[A]	1,3	-	3/7/10/12	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	718[A]	KCX	CD-CG-CB	-2.37	105.24	113.62

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	718[A]	KCX	O-C-CA-CB
1	B	718[A]	KCX	O-C-CA-CB
1	B	718[A]	KCX	CE-CD-CG-CB
1	B	718[A]	KCX	CG-CD-CE-NZ

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 6 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, 95^{th} 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	1003/1165 (86%)	0.52	156 (15%)	2	1	19, 72, 163, 191	8 (0%)
1	B	1001/1165 (85%)	0.30	113 (11%)	5	2	19, 56, 139, 190	7 (0%)
All	All	2004/2330 (86%)	0.41	269 (13%)	3	1	19, 65, 156, 191	15 (0%)

The worst 5 of 269 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	335	SER	11.1
1	A	142	THR	8.1
1	B	288	ALA	7.7
1	A	285	LEU	7.5
1	B	346	ASN	7.1

6.2 Non-standard residues in protein, DNA, RNA chains

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, 95^{th} 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
1	KCX	A	718[A]	12/13	0.98	0.13	20,21,22,22	0
1	KCX	B	718[A]	12/13	0.98	0.12	19,21,22,22	0

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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, 95^{th} 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	A	1200	1/1	0.80	0.19	25,25,25,25	0
2	MG	B	1200	1/1	0.88	0.24	40,40,40,40	0
3	ZN	A	1201	1/1	0.97	0.09	56,56,56,56	0
4	CL	B	1202	1/1	0.97	0.11	30,30,30,30	0
3	ZN	B	1201	1/1	0.98	0.05	59,59,59,59	0
4	CL	A	1202	1/1	0.98	0.17	24,24,24,24	0

6.5 Other polymers

There are no such residues in this entry.