



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 3, 2023 – 01:26 am GMT

PDB ID : 2J5I
Title : Crystal Structure of Hydroxycinnamoyl-CoA Hydratase-Lyase
Authors : Leonard, P.M.; Brzozowski, A.M.; Lebedev, A.; Marshall, C.M.; Smith, D.J.;
Verma, C.S.; Walton, N.J.; Grogan, G.
Deposited on : 2006-09-18
Resolution : 1.80 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

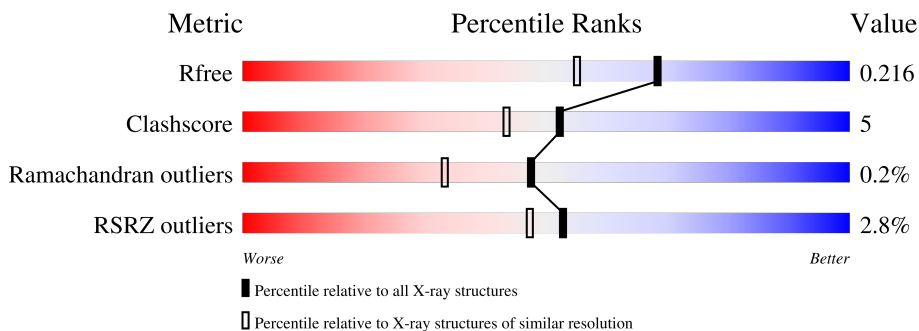
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 $\leq 5\%$. 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	276	 10% 80% 9% • 10%
2	B	276	 10% 79% 9% • 11%
2	C	276	 10% 82% 6% • 11%
2	D	276	 10% 4% 79% 10% • 10%
2	E	276	 10% 5% 82% 7% • 10%
2	F	276	 10% 3% 80% 9% • 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	276	 % 81% 7% • 11%
2	H	276	 2% 81% 7% • 11%
2	J	276	 4% 79% 10% 11%
2	K	276	 4% 79% 9% • 11%
2	L	276	 4% 78% 11% • 11%
3	I	276	 76% 13% 11%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 25264 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 P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	248	1958	1242	339	363	14	25	2	1

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	169	MET	TYR	conflict	UNP O69762

- Molecule 2 is a protein called P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	245	1934	1231	334	356	13	64	4	1
2	C	246	1959	1245	340	362	12	46	2	0
2	D	248	1968	1252	340	364	12	46	3	1
2	E	248	1959	1245	338	363	13	79	2	1
2	F	246	1946	1239	338	357	12	45	2	1
2	G	246	1960	1247	341	359	13	19	3	0
2	H	247	1946	1237	336	360	13	42	1	1
2	J	247	1941	1233	336	360	12	82	0	1
2	K	247	1952	1242	339	359	12	78	2	1
2	L	246	1936	1231	335	358	12	87	1	1

- Molecule 3 is a protein called P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	I	246	1944	1235	340	357	12	34	1	1

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	232	GLN	GLU	conflict	UNP O69762

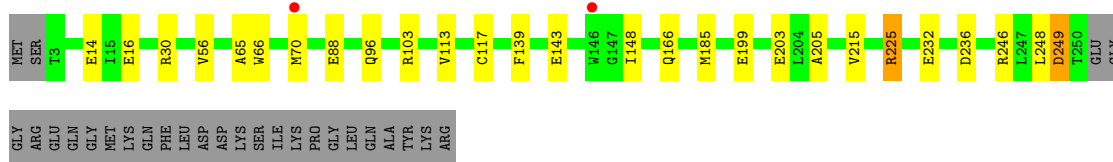
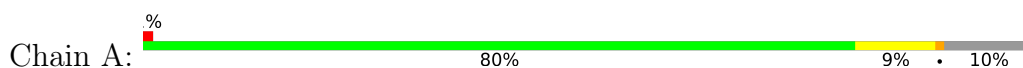
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	163	Total 163	O 163	0	0
4	B	115	Total 115	O 115	0	0
4	C	149	Total 149	O 149	0	0
4	D	149	Total 149	O 149	0	0
4	E	169	Total 169	O 169	0	0
4	F	175	Total 175	O 175	0	0
4	G	198	Total 198	O 198	0	0
4	H	152	Total 152	O 152	0	0
4	I	148	Total 148	O 148	0	0
4	J	146	Total 146	O 146	0	0
4	K	151	Total 151	O 151	0	0
4	L	146	Total 146	O 146	0	0

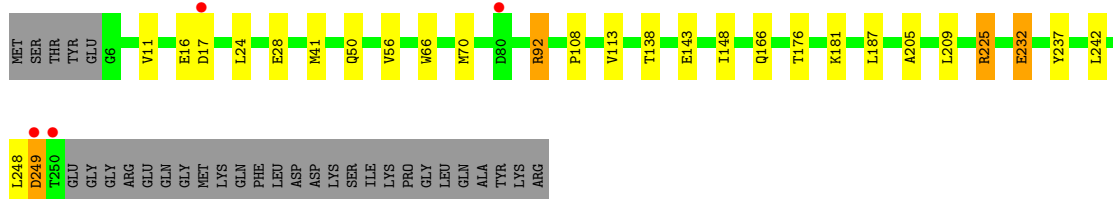
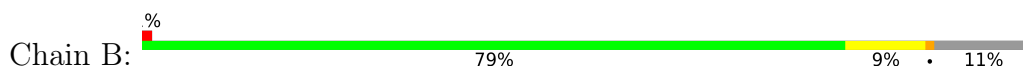
3 Residue-property plots [i](#)

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.

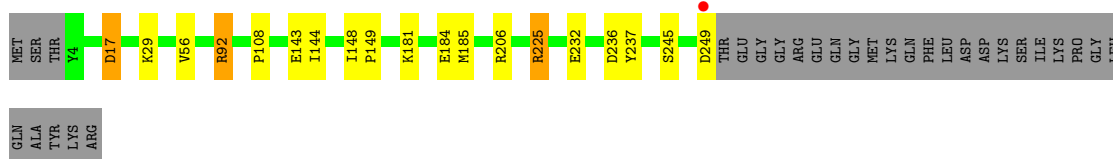
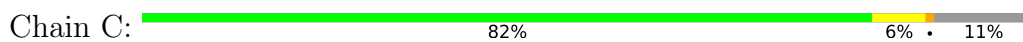
- Molecule 1: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE



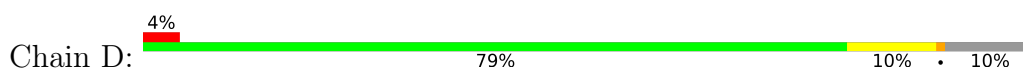
- Molecule 2: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE

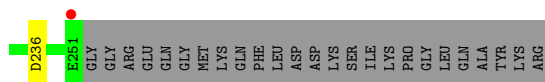


- Molecule 2: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE

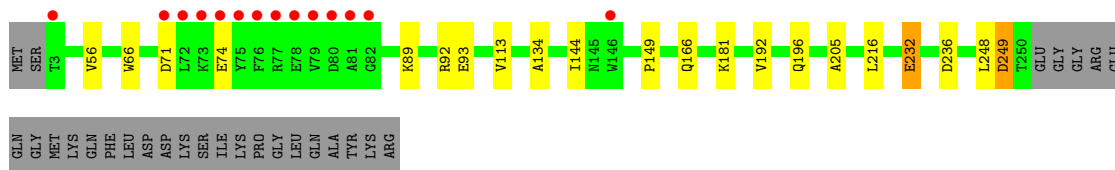
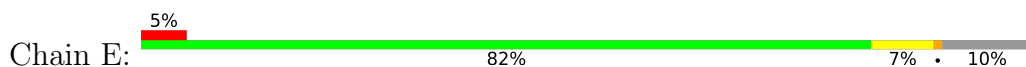


- Molecule 2: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE

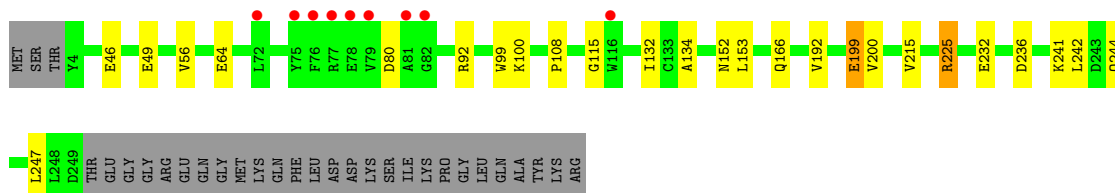
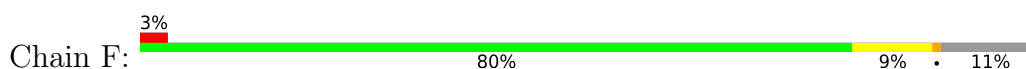




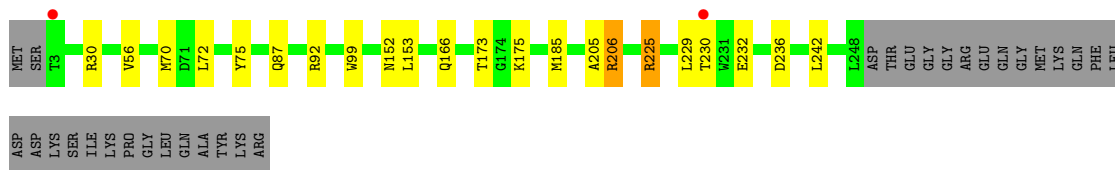
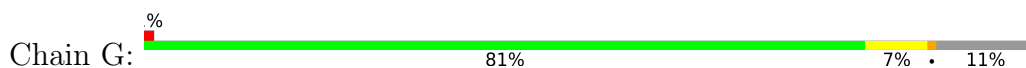
- Molecule 2: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE



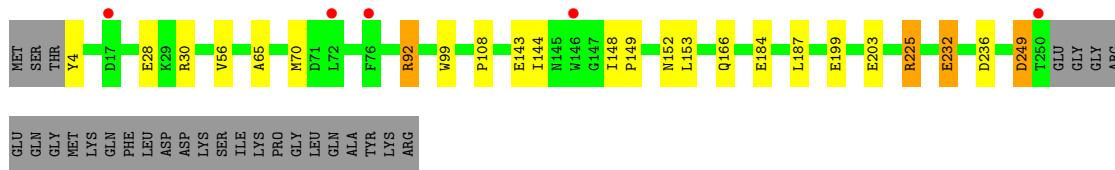
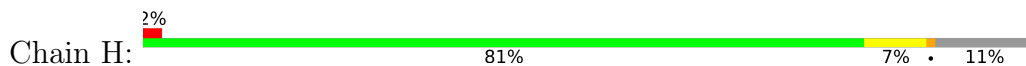
- Molecule 2: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE



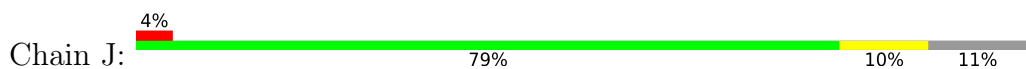
- Molecule 2: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE

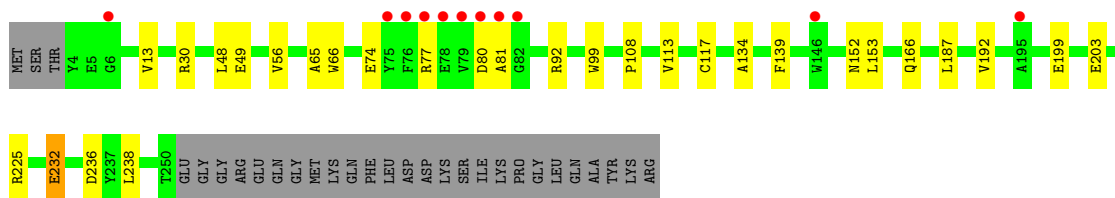


- Molecule 2: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE

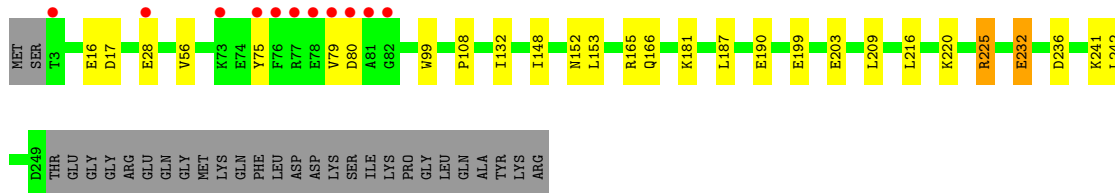
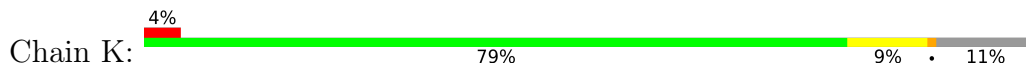


- Molecule 2: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE

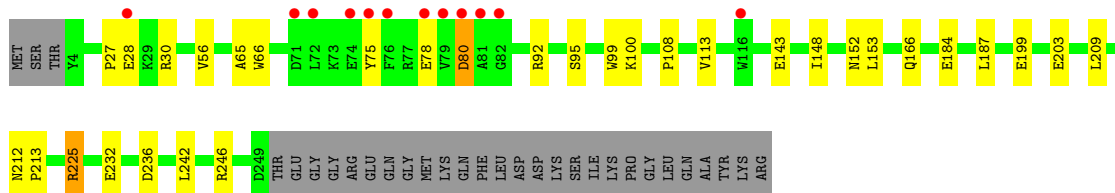
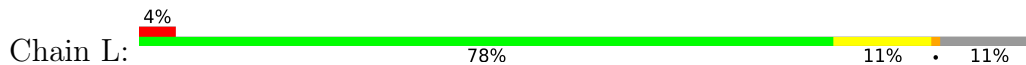




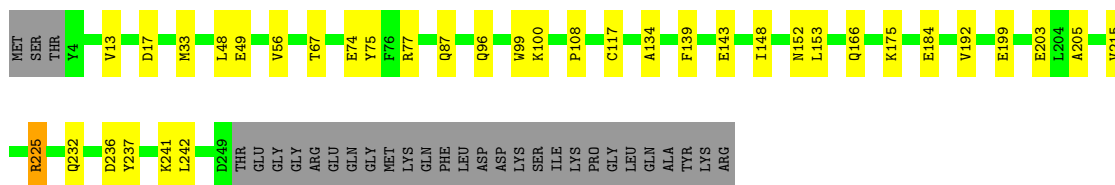
- Molecule 2: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE



- Molecule 2: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE



- Molecule 3: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	154.24Å 167.49Å 130.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	111.80 – 1.80 30.02 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.5 (111.80-1.80) 99.5 (30.02-1.80)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.09 (at 1.80Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.179 , 0.215 0.180 , 0.216	Depositor DCC
R_{free} test set	15557 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	20.1	Xtrriage
Anisotropy	0.621	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 53.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	25264	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 24.10 % 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 4.0906e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical 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

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	0.85	0/1999	0.82	4/2704 (0.1%)
2	B	0.84	2/1984 (0.1%)	0.88	5/2685 (0.2%)
2	C	0.81	0/1998	0.87	5/2703 (0.2%)
2	D	0.87	0/2017	0.85	4/2730 (0.1%)
2	E	0.87	1/2004 (0.0%)	0.80	1/2712 (0.0%)
2	F	0.94	1/1991 (0.1%)	0.84	3/2694 (0.1%)
2	G	0.90	1/2008 (0.0%)	0.88	4/2715 (0.1%)
2	H	0.86	2/1988 (0.1%)	0.88	5/2690 (0.2%)
2	J	0.82	1/1980 (0.1%)	0.80	2/2680 (0.1%)
2	K	0.94	4/1997 (0.2%)	0.90	7/2703 (0.3%)
2	L	0.89	3/1978 (0.2%)	0.84	4/2677 (0.1%)
3	I	0.92	1/1983 (0.1%)	0.84	3/2683 (0.1%)
All	All	0.88	16/23927 (0.1%)	0.85	47/32376 (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
2	E	0	1
2	G	1	0
All	All	1	1

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	K	79	VAL	CB-CG1	-11.09	1.29	1.52
2	K	80	ASP	CA-CB	-9.70	1.32	1.53
2	K	28	GLU	CB-CG	6.98	1.65	1.52
2	L	78	GLU	CB-CG	-6.63	1.39	1.52
2	H	232	GLU	CD-OE1	6.61	1.32	1.25

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	K	80	ASP	N-CA-CB	11.46	131.24	110.60
2	K	225	ARG	NE-CZ-NH2	11.30	125.95	120.30
2	H	92	ARG	NE-CZ-NH2	-11.27	114.66	120.30
2	K	225	ARG	NE-CZ-NH1	-11.21	114.69	120.30
2	D	225	ARG	NE-CZ-NH2	10.26	125.43	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	G	230	THR	CB

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	E	248	LEU	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	1958	0	1962	19	0
2	B	1934	0	1952	43	0
2	C	1959	0	1961	17	1
2	D	1968	0	1973	27	0
2	E	1959	0	1963	17	0
2	F	1946	0	1959	21	0
2	G	1960	0	1977	23	0
2	H	1946	0	1948	24	0
2	J	1941	0	1939	21	0
2	K	1952	0	1964	19	1
2	L	1936	0	1940	27	0
3	I	1944	0	1949	26	0
4	A	163	0	0	4	0
4	B	115	0	0	8	0
4	C	149	0	0	2	0
4	D	149	0	0	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	169	0	0	3	0
4	F	175	0	0	5	0
4	G	198	0	0	5	0
4	H	152	0	0	6	0
4	I	148	0	0	2	0
4	J	146	0	0	2	0
4	K	151	0	0	5	0
4	L	146	0	0	5	0
All	All	25264	0	23487	249	1

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 5.

The worst 5 of 249 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)
2:E:71:ASP:CG	2:E:74:GLU:HG3	1.71	1.11
2:B:56[B]:VAL:HG12	2:B:108:PRO:HG2	1.31	1.10
2:B:11:VAL:CG2	2:B:41:MET:HE1	1.85	1.05
2:E:71:ASP:OD2	2:E:74:GLU:HG3	1.62	0.99
2:K:56[B]:VAL:HG12	2:K:108:PRO:HG2	1.42	0.98

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:29:LYS:NZ	2:K:17:ASP:OD1[2_664]	1.89	0.31

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	248/276 (90%)	238 (96%)	9 (4%)	1 (0%)	34	21
2	B	247/276 (90%)	237 (96%)	9 (4%)	1 (0%)	34	21
2	C	246/276 (89%)	238 (97%)	8 (3%)	0	100	100
2	D	249/276 (90%)	240 (96%)	9 (4%)	0	100	100
2	E	248/276 (90%)	239 (96%)	8 (3%)	1 (0%)	34	21
2	F	246/276 (89%)	240 (98%)	6 (2%)	0	100	100
2	G	247/276 (90%)	241 (98%)	6 (2%)	0	100	100
2	H	246/276 (89%)	236 (96%)	9 (4%)	1 (0%)	34	21
2	J	245/276 (89%)	236 (96%)	9 (4%)	0	100	100
2	K	247/276 (90%)	237 (96%)	9 (4%)	1 (0%)	34	21
2	L	245/276 (89%)	236 (96%)	9 (4%)	0	100	100
3	I	245/276 (89%)	237 (97%)	8 (3%)	0	100	100
All	All	2959/3312 (89%)	2855 (96%)	99 (3%)	5 (0%)	47	33

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	249	ASP
2	E	249	ASP
2	H	249	ASP
1	A	249	ASP
2	K	75	TYR

5.3.2 Protein sidechains [i](#)

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

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 monosaccharides 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	247/276 (89%)	-0.38	2 (0%) 86 84	12, 19, 32, 40	8 (3%)
2	B	245/276 (88%)	-0.18	4 (1%) 72 68	14, 23, 36, 44	16 (6%)
2	C	246/276 (89%)	-0.35	1 (0%) 92 90	13, 21, 30, 38	12 (4%)
2	D	248/276 (89%)	-0.10	11 (4%) 34 28	12, 20, 35, 45	12 (4%)
2	E	247/276 (89%)	-0.07	14 (5%) 23 19	12, 19, 34, 44	19 (7%)
2	F	246/276 (89%)	-0.16	9 (3%) 41 36	12, 19, 34, 41	12 (4%)
2	G	246/276 (89%)	-0.42	2 (0%) 86 84	11, 18, 30, 38	4 (1%)
2	H	247/276 (89%)	-0.19	5 (2%) 65 61	13, 22, 32, 44	12 (4%)
2	J	246/276 (89%)	-0.16	11 (4%) 33 27	14, 21, 35, 40	17 (6%)
2	K	247/276 (89%)	-0.16	11 (4%) 33 27	13, 20, 36, 41	19 (7%)
2	L	246/276 (89%)	-0.17	12 (4%) 29 24	13, 21, 37, 44	21 (8%)
3	I	246/276 (89%)	-0.34	0 100 100	13, 21, 31, 39	8 (3%)
All	All	2957/3312 (89%)	-0.22	82 (2%) 53 47	11, 20, 33, 45	160 (5%)

The worst 5 of 82 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	J	81	ALA	8.0
2	E	72	LEU	7.9
2	B	250	THR	7.7
2	E	79	VAL	7.2
2	L	81	ALA	6.7

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.