



wwPDB X-ray Structure Validation Summary Report i

Mar 1, 2014 – 01:07 AM GMT

PDB ID : 3PQ4
Title : Structure of I274C variant of E. coli KatE[] - Images 13-18
Authors : Loewen, P.C.; Jha, V.; Louis, S.; Chelikani, P.; Carpena, X.; Fita, I.
Deposited on : 2010-11-25
Resolution : 1.79 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

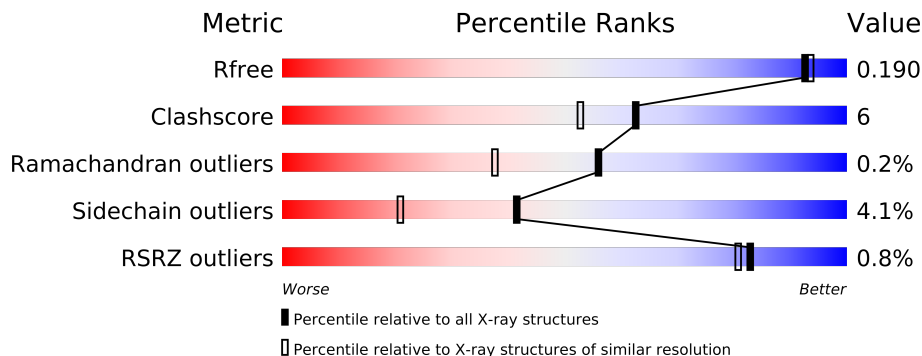
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 1.79 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}	66092	3513 (1.80-1.80)
Clashscore	79885	4461 (1.80-1.80)
Ramachandran outliers	78287	4404 (1.80-1.80)
Sidechain outliers	78261	4403 (1.80-1.80)
RSRZ outliers	66119	3515 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	753	
1	B	753	
1	C	753	
1	D	753	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	H2S	A	754	-	X
4	H2S	B	754	-	X
4	H2S	C	754	-	X
4	H2S	D	754	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 26821 atoms, of which 0 are hydrogen and 0 are deuterium.

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.

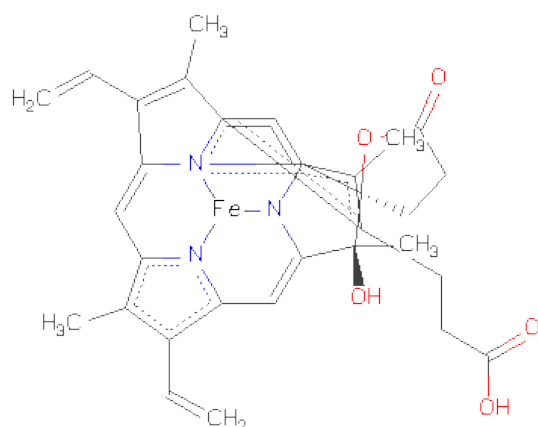
- Molecule 1 is a protein called Catalase HP11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	726	Total	C	N	O	S	0	5	0
			5755	3652	1009	1083	11			
1	B	726	Total	C	N	O	S	0	5	0
			5757	3654	1009	1083	11			
1	C	726	Total	C	N	O	S	0	2	0
			5747	3649	1007	1080	11			
1	D	726	Total	C	N	O	S	0	9	0
			5771	3662	1013	1085	11			

There are 12 discrepancies between the modelled and reference sequences:

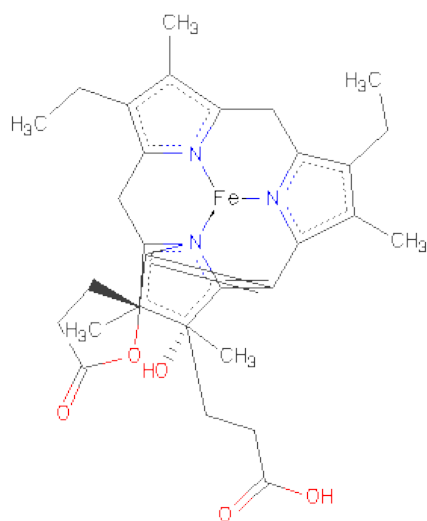
Chain	Residue	Modelled	Actual	Comment	Reference
A	274	CYS	ILE	ENGINEERED MUTATION	UNP P21179
A	438	ALA	CYS	ENGINEERED MUTATION	UNP P21179
A	669	ALA	CYS	ENGINEERED MUTATION	UNP P21179
B	274	CYS	ILE	ENGINEERED MUTATION	UNP P21179
B	438	ALA	CYS	ENGINEERED MUTATION	UNP P21179
B	669	ALA	CYS	ENGINEERED MUTATION	UNP P21179
C	274	CYS	ILE	ENGINEERED MUTATION	UNP P21179
C	438	ALA	CYS	ENGINEERED MUTATION	UNP P21179
C	669	ALA	CYS	ENGINEERED MUTATION	UNP P21179
D	274	CYS	ILE	ENGINEERED MUTATION	UNP P21179
D	438	ALA	CYS	ENGINEERED MUTATION	UNP P21179
D	669	ALA	CYS	ENGINEERED MUTATION	UNP P21179

- Molecule 2 is CIS-HEME D HYDROXYCHLORIN GAMMA-SPIROLACTONE (three-letter code: HDD) (formula: C₃₄H₃₂FeN₄O₅).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Fe	N	O	0	1
			44	34	1	4	5		
2	B	1	Total	C	Fe	N	O	0	1
			44	34	1	4	5		
2	C	1	Total	C	Fe	N	O	0	1
			44	34	1	4	5		
2	D	1	Total	C	Fe	N	O	0	1
			44	34	1	4	5		

- Molecule 3 is CIS-HEME D HYDROXYCHLORIN GAMMA-SPIROLACTONE 17R, 18S (three-letter code: HDE) (formula: $C_{34}H_{38}FeN_4O_5$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 44	C 34	Fe 1	N 4	O 5	0	1
3	B	1	Total 44	C 34	Fe 1	N 4	O 5	0	1
3	C	1	Total 44	C 34	Fe 1	N 4	O 5	0	1
3	D	1	Total 44	C 34	Fe 1	N 4	O 5	0	1

- Molecule 4 is HYDROSULFURIC ACID (three-letter code: H2S) (formula: H₂S).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	S 1	0	0
4	B	1	Total 1	S 1	0	0
4	C	1	Total 1	S 1	0	0
4	D	1	Total 1	S 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	896	Total 896	O 896	0	0
5	B	788	Total 788	O 788	0	0

Continued on next page...

Continued from previous page...

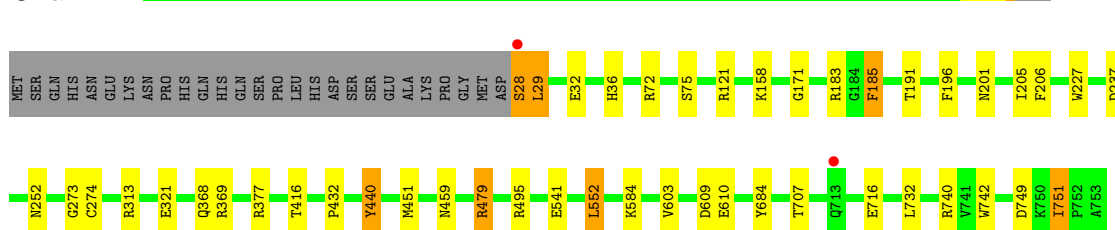
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	838	Total 838	O 838	0	0
5	D	913	Total 913	O 913	0	0

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

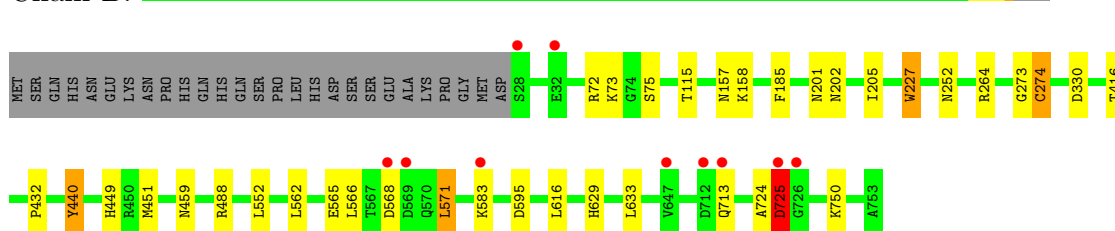
- Molecule 1: Catalase HP1I

Chain A:



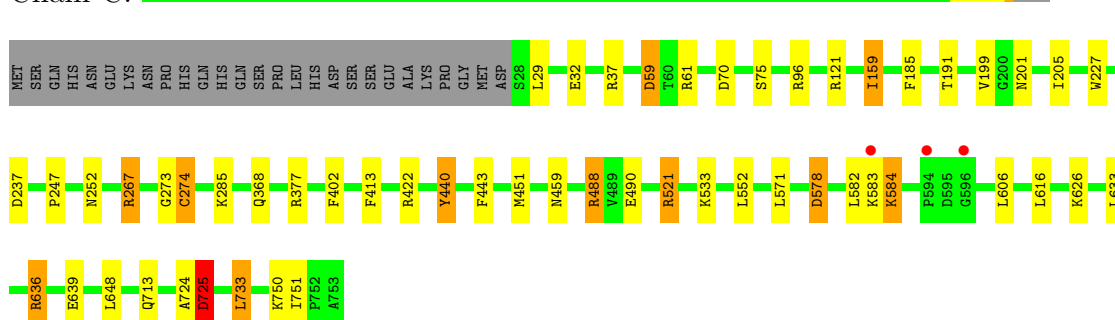
- Molecule 1: Catalase HP1I

Chain B:



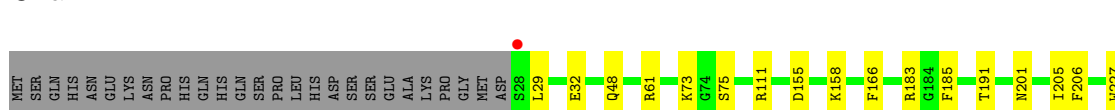
- Molecule 1: Catalase HP1I

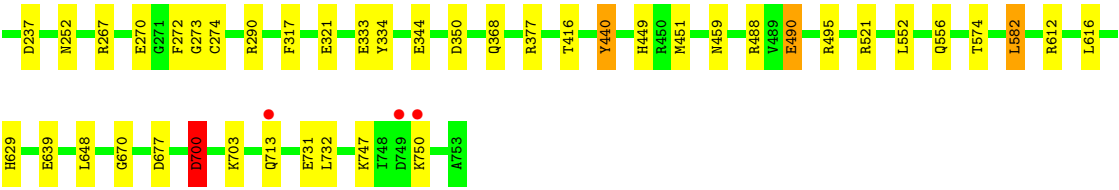
Chain C:



- Molecule 1: Catalase HP1I

Chain D:





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	93.51Å 133.03Å 122.65Å 90.00° 109.39° 90.00°	Depositor
Resolution (Å)	35.19 – 1.79 35.19 – 1.79	Depositor EDS
% Data completeness (in resolution range)	92.0 (35.19-1.79) 92.0 (35.19-1.79)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.16	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.48 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.141 , 0.189 0.142 , 0.190	Depositor DCC
R_{free} test set	12136 reflections (5.27%)	DCC
Wilson B-factor (Å ²)	11.7	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 33.8	EDS
Estimated twinning fraction	0.025 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 242617 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	26821	wwPDB-VP
Average B, all atoms (Å ²)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.21% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: HDE, HDD, H2S

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	1.18	9/5932 (0.2%)	0.96	10/8064 (0.1%)
1	B	1.11	2/5931 (0.0%)	0.94	5/8062 (0.1%)
1	C	1.14	3/5908 (0.1%)	0.97	12/8033 (0.1%)
1	D	1.19	7/5966 (0.1%)	0.97	14/8109 (0.2%)
All	All	1.15	21/23737 (0.1%)	0.96	41/32268 (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	C	0	2

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	272	PHE	CE2-CZ	7.73	1.52	1.37
1	D	333	GLU	CD-OE2	-7.34	1.17	1.25
1	A	684	TYR	CD2-CE2	6.67	1.49	1.39
1	A	603	VAL	CB-CG1	-6.55	1.39	1.52
1	D	334	TYR	CD2-CE2	5.89	1.48	1.39

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	636	ARG	NE-CZ-NH1	13.18	126.89	120.30
1	C	636	ARG	NE-CZ-NH2	-12.14	114.23	120.30
1	A	740	ARG	NE-CZ-NH1	7.97	124.29	120.30
1	B	72	ARG	NE-CZ-NH1	7.85	124.22	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	59	ASP	CB-CG-OD1	7.66	125.19	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	724	ALA	Peptide
1	C	725	ASP	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5755	0	0	20	0
1	B	5757	0	0	18	0
1	C	5747	0	0	20	1
1	D	5771	0	0	24	0
2	A	44	0	0	6	0
2	B	44	0	0	4	0
2	C	44	0	0	6	0
2	D	44	0	0	6	0
3	A	44	0	36	9	0
3	B	44	0	36	13	0
3	C	44	0	36	12	0
3	D	44	0	36	10	0
4	A	1	0	0	6	0
4	B	1	0	0	3	0
4	C	1	0	0	3	0
4	D	1	0	0	5	0
5	A	896	0	0	9	2
5	B	788	0	0	7	0
5	C	838	0	0	11	0
5	D	913	0	0	19	2
All	All	26821	0	144	131	3

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

The worst 5 of 131 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:C:760[A]:HDD:CBB	4:C:754:H2S:S	2.06	1.43
1:A:274:CYS:CB	4:A:754:H2S:S	2.19	1.31
2:A:760[A]:HDD:CBB	4:A:754:H2S:S	2.19	1.30
1:B:449[A]:HIS:CD2	5:B:3614:HOH:O	1.67	1.26
1:D:449[B]:HIS:CE1	5:D:3616:HOH:O	1.65	1.25

All (3) 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	Distance(Å)	Clash(Å)
5:D:2178:HOH:O	5:D:2976:HOH:O[1_655]	2.00	0.20
1:C:584:LYS:NZ	5:A:1624:HOH:O[1_554]	2.06	0.14
5:A:3241:HOH:O	5:D:2457:HOH:O[2_646]	2.15	0.05

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	729/753 (97%)	707 (97%)	21 (3%)	1 (0%)	59	41
1	B	729/753 (97%)	709 (97%)	18 (2%)	2 (0%)	50	31
1	C	726/753 (96%)	702 (97%)	22 (3%)	2 (0%)	50	31
1	D	733/753 (97%)	713 (97%)	19 (3%)	1 (0%)	59	41
All	All	2917/3012 (97%)	2831 (97%)	80 (3%)	6 (0%)	56	38

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	725	ASP
1	A	75	SER
1	C	75	SER
1	C	725	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	75	SER

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	614/634 (97%)	595 (97%)	19 (3%)	52	33
1	B	614/634 (97%)	592 (96%)	22 (4%)	47	27
1	C	611/634 (96%)	577 (94%)	34 (6%)	30	11
1	D	618/634 (98%)	592 (96%)	26 (4%)	40	20
All	All	2457/2536 (97%)	2356 (96%)	101 (4%)	41	21

5 of 101 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	191	THR
1	C	488	ARG
1	D	582	LEU
1	C	227	TRP
1	C	274	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 4 are modelled with single atom - leaving 8 for Mogul analysis.

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	HDD	A	760[A]	1,5	52,52,52	2.63	19 (36%)	70,89,89	2.09	23 (32%)
3	HDE	A	761[B]	1,5	52,52,52	3.13	21 (40%)	75,89,89	2.69	35 (46%)
2	HDD	B	760[A]	1,5	52,52,52	2.33	19 (36%)	70,89,89	2.44	22 (31%)
3	HDE	B	761[B]	1,5	52,52,52	3.06	22 (42%)	75,89,89	2.70	33 (44%)
2	HDD	C	760[A]	1,5	52,52,52	2.33	24 (46%)	70,89,89	2.58	27 (38%)
3	HDE	C	761[B]	1,5	52,52,52	3.03	24 (46%)	75,89,89	2.69	33 (44%)
2	HDD	D	760[A]	1,5	52,52,52	2.56	20 (38%)	70,89,89	2.38	27 (38%)
3	HDE	D	761[B]	1,5	52,52,52	2.75	19 (36%)	75,89,89	2.75	36 (48%)

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	HDD	A	760[A]	1,5	-	0/5/89/89	0/1/9/9
3	HDE	A	761[B]	1,5	-	1/9/89/89	0/1/9/9
2	HDD	B	760[A]	1,5	-	0/5/89/89	0/1/9/9
3	HDE	B	761[B]	1,5	-	2/9/89/89	0/1/9/9
2	HDD	C	760[A]	1,5	-	0/5/89/89	0/1/9/9
3	HDE	C	761[B]	1,5	-	1/9/89/89	0/1/9/9
2	HDD	D	760[A]	1,5	-	0/5/89/89	0/1/9/9
3	HDE	D	761[B]	1,5	-	1/9/89/89	0/1/9/9

The worst 5 of 168 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	761[B]	HDE	CHC-C4B	-8.96	1.43	1.51
3	A	761[B]	HDE	C4C-C3C	8.50	1.49	1.37
3	C	761[B]	HDE	CHC-C4B	-8.29	1.43	1.51
3	C	761[B]	HDE	C4C-C3C	8.10	1.48	1.37
2	B	760[A]	HDD	C3C-CAC	7.04	1.55	1.49

The worst 5 of 236 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	760[A]	HDD	O1D-CGD-O2D	7.71	128.42	120.81
2	C	760[A]	HDD	O1D-CGD-O2D	7.50	128.20	120.81
2	C	760[A]	HDD	C3B-CAB-CBB	-6.91	111.64	125.95
3	B	761[B]	HDE	C2B-C1B-NB	6.88	114.60	109.41
3	C	761[B]	HDE	C2C-C1C-NC	6.64	117.35	110.11

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	761[B]	HDE	C2B-C3B-CAB-CBB
3	A	761[B]	HDE	C2B-C3B-CAB-CBB
3	C	761[B]	HDE	C2B-C3B-CAB-CBB
3	B	761[B]	HDE	C2B-C3B-CAB-CBB
3	B	761[B]	HDE	C2C-C3C-CAC-CBC

There are no ring outliers.

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	726/753 (96%)	-0.64	2 (0%) 91 90	3, 10, 27, 45	2 (0%)
1	B	726/753 (96%)	-0.50	10 (1%) 72 68	4, 12, 32, 48	2 (0%)
1	C	726/753 (96%)	-0.55	3 (0%) 90 89	4, 12, 32, 47	2 (0%)
1	D	726/753 (96%)	-0.64	4 (0%) 86 85	3, 10, 28, 45	1 (0%)
All	All	2904/3012 (96%)	-0.58	19 (0%) 83 83	3, 11, 30, 48	7 (0%)

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	28	SER	4.1
1	B	726	GLY	3.8
1	A	28	SER	3.3
1	D	749	ASP	3.2
1	B	28	SER	3.2

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 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	H2S	C	754	1/1	0.20	8.86	24,24,24,24	1
4	H2S	D	754	1/1	0.21	8.03	19,19,19,19	1
4	H2S	B	754	1/1	0.21	6.80	25,25,25,25	1
4	H2S	A	754	1/1	0.19	5.78	22,22,22,22	1
2	HDD	C	760[A]	44/44	0.07	0.93	2,4,8,9	44
2	HDD	B	760[A]	44/44	0.07	0.87	2,4,8,8	44
3	HDE	C	761[B]	44/44	0.07	0.85	6,10,13,15	44
3	HDE	B	761[B]	44/44	0.07	0.84	3,11,14,15	44
2	HDD	D	760[A]	44/44	0.06	0.20	2,3,6,13	44
2	HDD	A	760[A]	44/44	0.07	0.16	2,4,8,12	44
3	HDE	D	761[B]	44/44	0.06	0.16	2,9,13,15	44
3	HDE	A	761[B]	44/44	0.07	0.11	2,8,11,15	44

6.5 Other polymers ⓘ

There are no such residues in this entry.