



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 23, 2021 – 12:39 PM EDT

PDB ID : 3PCD
Title : PROTOCATECHUATE 3,4-DIOXYGENASE Y447H MUTANT
Authors : Orville, A.M.; Lipscomb, J.D.; Ohlendorf, D.H.
Deposited on : 1997-11-24
Resolution : 2.10 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 ⓘ 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) : NOT EXECUTED
EDS : NOT EXECUTED
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.23.2

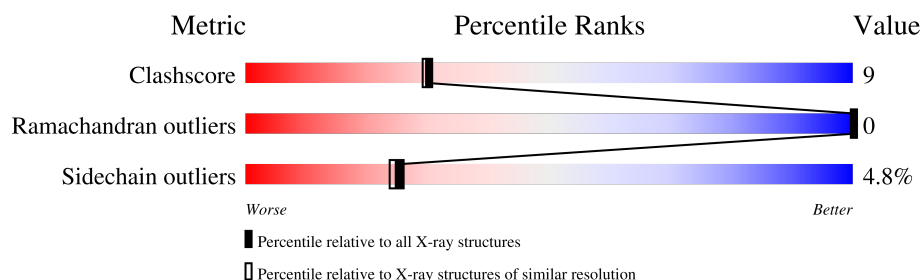
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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))
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)

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 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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	200	80% 16% .
1	B	200	78% 18% .
1	C	200	76% 20% .
1	D	200	78% 18% .
1	E	200	72% 22% 6%
1	F	200	71% 22% 6%
2	M	238	73% 20% . .
2	N	238	78% 16% . .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	O	238	<div><div></div><div>75%</div><div>19%</div><div></div><div>...</div></div>
2	P	238	<div><div></div><div>74%</div><div>19%</div><div></div><div>...</div></div>
2	Q	238	<div><div></div><div>76%</div><div>17%</div><div></div><div>...</div></div>
2	R	238	<div><div></div><div>71%</div><div>23%</div><div></div><div>..</div></div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 21930 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 PROTOCATECHUATE 3,4-DIOXYGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	200	Total	C	N	O	S	0	0	0
			1571	993	276	299	3			
1	B	200	Total	C	N	O	S	0	0	0
			1571	993	276	299	3			
1	C	200	Total	C	N	O	S	0	0	0
			1571	993	276	299	3			
1	D	200	Total	C	N	O	S	0	0	0
			1571	993	276	299	3			
1	E	200	Total	C	N	O	S	0	0	0
			1571	993	276	299	3			
1	F	200	Total	C	N	O	S	0	0	0
			1571	993	276	299	3			

- Molecule 2 is a protein called PROTOCATECHUATE 3,4-DIOXYGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	M	233	Total	C	N	O	S	0	0	0
			1838	1168	336	327	7			
2	N	233	Total	C	N	O	S	0	0	0
			1838	1168	336	327	7			
2	O	233	Total	C	N	O	S	0	0	0
			1838	1168	336	327	7			
2	P	233	Total	C	N	O	S	0	0	0
			1838	1168	336	327	7			
2	Q	233	Total	C	N	O	S	0	0	0
			1838	1168	336	327	7			
2	R	233	Total	C	N	O	S	0	0	0
			1838	1168	336	327	7			

There are 6 discrepancies between the modelled and reference sequences:

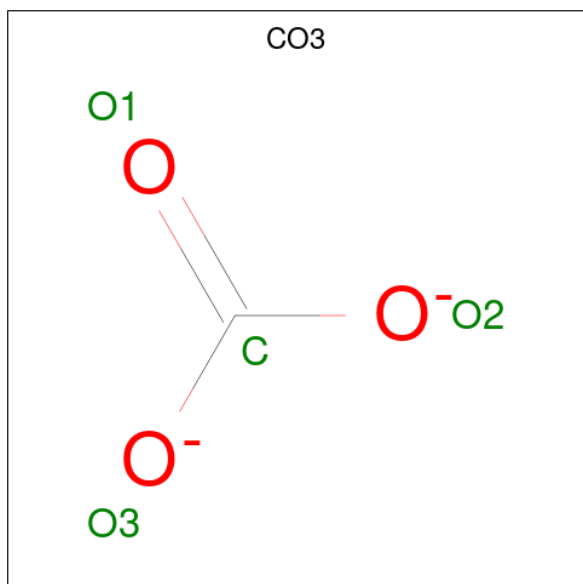
Chain	Residue	Modelled	Actual	Comment	Reference
M	447	HIS	TYR	engineered mutation	UNP P00437

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
N	447	HIS	TYR	engineered mutation	UNP P00437
O	447	HIS	TYR	engineered mutation	UNP P00437
P	447	HIS	TYR	engineered mutation	UNP P00437
Q	447	HIS	TYR	engineered mutation	UNP P00437
R	447	HIS	TYR	engineered mutation	UNP P00437

- Molecule 3 is CARBONATE ION (three-letter code: CO3) (formula: CO₃).

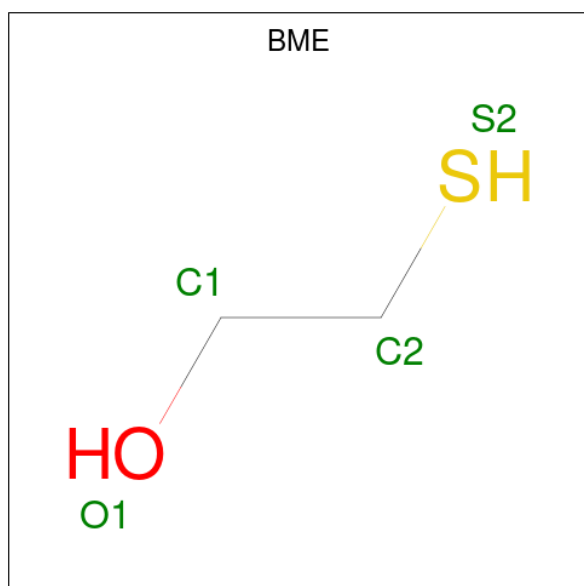


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	M	1	Total	C	O	0	0
			4	1	3		
3	N	1	Total	C	O	0	0
			4	1	3		
3	O	1	Total	C	O	0	0
			4	1	3		
3	P	1	Total	C	O	0	0
			4	1	3		
3	Q	1	Total	C	O	0	0
			4	1	3		
3	R	1	Total	C	O	0	0
			4	1	3		

- Molecule 4 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	M	1	Total Fe 1 1	0	0
4	N	1	Total Fe 1 1	0	0
4	O	1	Total Fe 1 1	0	0
4	P	1	Total Fe 1 1	0	0
4	Q	1	Total Fe 1 1	0	0
4	R	1	Total Fe 1 1	0	0

- Molecule 5 is BETA-MERCAPTOETHANOL (three-letter code: BME) (formula: C_2H_6OS).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	M	1	Total C O S 4 2 1 1	0	0
5	N	1	Total C O S 4 2 1 1	0	0
5	O	1	Total C O S 4 2 1 1	0	0
5	P	1	Total C O S 4 2 1 1	0	0
5	Q	1	Total C O S 4 2 1 1	0	0
5	R	1	Total C O S 4 2 1 1	0	0

- Molecule 6 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	73	Total 73	O 73	0	0
6	M	161	Total 161	O 161	0	0
6	B	79	Total 79	O 79	0	0
6	N	163	Total 163	O 163	0	0
6	C	78	Total 78	O 78	0	0
6	O	156	Total 156	O 156	0	0
6	D	81	Total 81	O 81	0	0
6	P	151	Total 151	O 151	0	0
6	E	82	Total 82	O 82	0	0
6	Q	159	Total 159	O 159	0	0
6	F	80	Total 80	O 80	0	0
6	R	159	Total 159	O 159	0	0

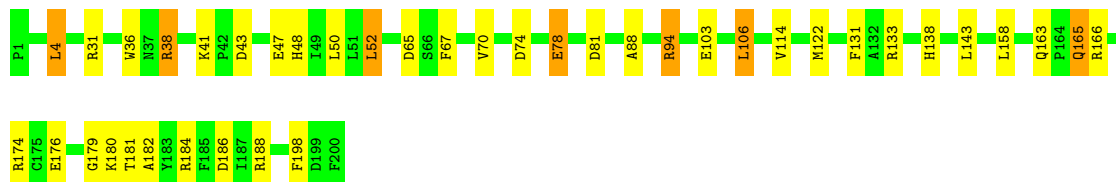
3 Residue-property plots [i](#)

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. 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. 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.


Note EDS was not executed.

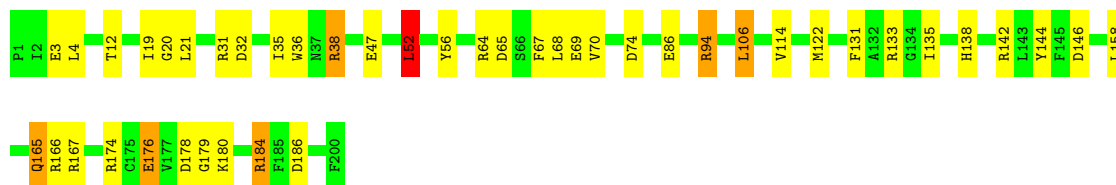
• Molecule 1: PROTOCATECHUATE 3,4-DIOXYGENASE

Chain A: 




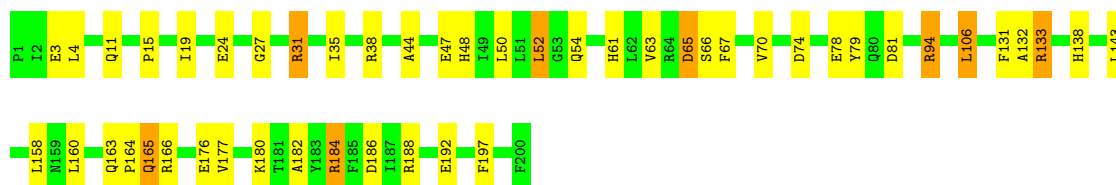
• Molecule 1: PROTOCATECHUATE 3,4-DIOXYGENASE

Chain B: 




• Molecule 1: PROTOCATECHUATE 3,4-DIOXYGENASE

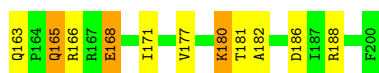
Chain C: 



• Molecule 1: PROTOCATECHUATE 3,4-DIOXYGENASE

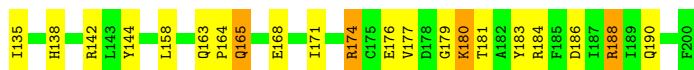
Chain D: 





• Molecule 1: PROTOCATECHUATE 3,4-DIOXYGENASE

Chain E: 72% 22% 6%



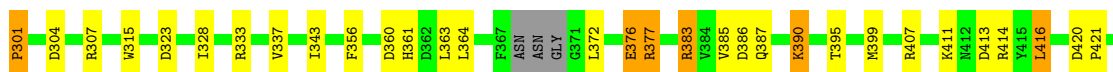
• Molecule 1: PROTOCATECHUATE 3,4-DIOXYGENASE

Chain F: 71% 22% 6%



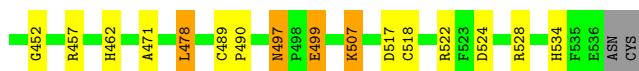
• Molecule 2: PROTOCATECHUATE 3,4-DIOXYGENASE

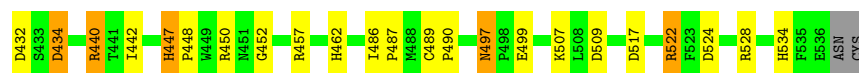
Chain M: 73% 20% 7%



• Molecule 2: PROTOCATECHUATE 3,4-DIOXYGENASE

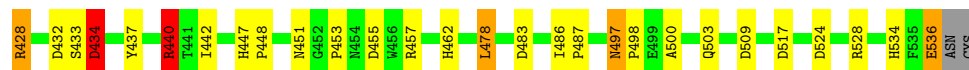
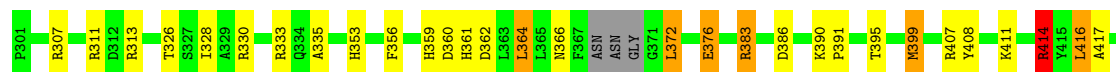
Chain N: 78% 16% 6%





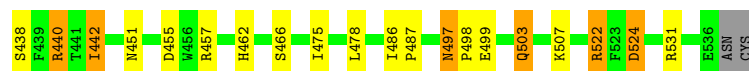
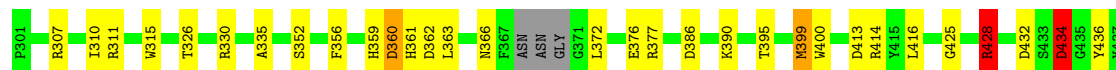
• Molecule 2: PROTOCATECHUATE 3,4-DIOXYGENASE

Chain P: 74% 19% . . .



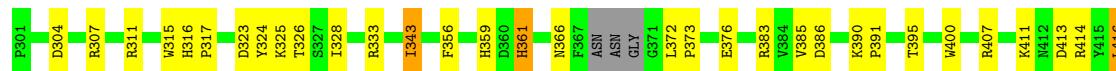
• Molecule 2: PROTOCATECHUATE 3,4-DIOXYGENASE

Chain Q: 76% 17% . . .



• Molecule 2: PROTOCATECHUATE 3,4-DIOXYGENASE

Chain R: 71% 23% . .



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	197.40 Å 127.20 Å 134.60 Å 90.00° 97.70° 90.00°	Depositor
Resolution (Å)	6.00 – 2.10	Depositor
% Data completeness (in resolution range)	78.0 (6.00-2.10)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	0.180 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	21930	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CO3, BME, FE

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	1.03	1/1611 (0.1%)	1.70	29/2195 (1.3%)
1	B	1.04	2/1611 (0.1%)	1.79	31/2195 (1.4%)
1	C	1.03	0/1611	1.71	26/2195 (1.2%)
1	D	1.05	2/1611 (0.1%)	1.79	28/2195 (1.3%)
1	E	1.00	1/1611 (0.1%)	1.60	28/2195 (1.3%)
1	F	1.02	0/1611	1.65	25/2195 (1.1%)
2	M	1.10	1/1893 (0.1%)	1.67	35/2577 (1.4%)
2	N	1.07	2/1893 (0.1%)	1.63	36/2577 (1.4%)
2	O	1.09	2/1893 (0.1%)	1.56	32/2577 (1.2%)
2	P	1.10	1/1893 (0.1%)	1.75	35/2577 (1.4%)
2	Q	1.13	3/1893 (0.2%)	1.70	27/2577 (1.0%)
2	R	1.13	2/1893 (0.1%)	1.62	30/2577 (1.2%)
All	All	1.07	17/21024 (0.1%)	1.68	362/28632 (1.3%)

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	B	0	1
2	P	0	1
All	All	0	2

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	94	ARG	CD-NE	-7.13	1.34	1.46
1	A	94	ARG	CD-NE	-6.99	1.34	1.46
2	R	452	GLY	N-CA	6.47	1.55	1.46
2	Q	440	ARG	CD-NE	-6.29	1.35	1.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	428	ARG	CD-NE	-6.24	1.35	1.46

The worst 5 of 362 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	P	440	ARG	NE-CZ-NH2	-31.71	104.45	120.30
1	D	133	ARG	CD-NE-CZ	30.50	166.31	123.60
1	B	133	ARG	CD-NE-CZ	29.98	165.58	123.60
2	N	440	ARG	NE-CZ-NH2	-23.03	108.78	120.30
2	P	440	ARG	NE-CZ-NH1	22.20	131.40	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	184	ARG	Sidechain
2	P	440	ARG	Sidechain

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	1571	0	1499	27	0
1	B	1571	0	1499	25	0
1	C	1571	0	1499	33	0
1	D	1571	0	1499	24	0
1	E	1571	0	1499	42	0
1	F	1571	0	1499	47	0
2	M	1838	0	1791	46	0
2	N	1838	0	1791	19	0
2	O	1838	0	1791	30	0
2	P	1838	0	1791	39	0
2	Q	1838	0	1791	33	0
2	R	1838	0	1791	39	0
3	M	4	0	0	1	0
3	N	4	0	0	0	0
3	O	4	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	P	4	0	0	1	0
3	Q	4	0	0	0	0
3	R	4	0	0	1	0
4	M	1	0	0	0	0
4	N	1	0	0	0	0
4	O	1	0	0	0	0
4	P	1	0	0	0	0
4	Q	1	0	0	0	0
4	R	1	0	0	0	0
5	M	4	0	5	0	0
5	N	4	0	5	0	0
5	O	4	0	5	1	0
5	P	4	0	5	0	0
5	Q	4	0	5	2	0
5	R	4	0	5	0	0
6	A	73	0	0	0	0
6	B	79	0	0	0	0
6	C	78	0	0	0	0
6	D	81	0	0	0	0
6	E	82	0	0	1	0
6	F	80	0	0	0	0
6	M	161	0	0	4	0
6	N	163	0	0	2	0
6	O	156	0	0	3	0
6	P	151	0	0	3	0
6	Q	159	0	0	2	0
6	R	159	0	0	2	0
All	All	21930	0	19770	376	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 9.

The worst 5 of 376 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:E:165:GLN:H	1:E:165:GLN:NE2	1.64	0.95
1:E:165:GLN:H	1:E:165:GLN:HE21	1.05	0.93
1:F:165:GLN:H	1:F:165:GLN:NE2	1.66	0.93
1:F:165:GLN:H	1:F:165:GLN:HE21	0.97	0.92
2:M:497:ASN:HD22	2:M:499:GLU:H	1.22	0.87

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	198/200 (99%)	190 (96%)	8 (4%)	0	100	100
1	B	198/200 (99%)	193 (98%)	5 (2%)	0	100	100
1	C	198/200 (99%)	193 (98%)	5 (2%)	0	100	100
1	D	198/200 (99%)	191 (96%)	7 (4%)	0	100	100
1	E	198/200 (99%)	189 (96%)	9 (4%)	0	100	100
1	F	198/200 (99%)	191 (96%)	7 (4%)	0	100	100
2	M	229/238 (96%)	223 (97%)	6 (3%)	0	100	100
2	N	229/238 (96%)	223 (97%)	6 (3%)	0	100	100
2	O	229/238 (96%)	221 (96%)	8 (4%)	0	100	100
2	P	229/238 (96%)	224 (98%)	5 (2%)	0	100	100
2	Q	229/238 (96%)	221 (96%)	8 (4%)	0	100	100
2	R	229/238 (96%)	224 (98%)	5 (2%)	0	100	100
All	All	2562/2628 (98%)	2483 (97%)	79 (3%)	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	162/163 (99%)	156 (96%)	6 (4%)	34	35
1	B	162/163 (99%)	157 (97%)	5 (3%)	40	43
1	C	162/163 (99%)	156 (96%)	6 (4%)	34	35
1	D	162/163 (99%)	155 (96%)	7 (4%)	29	29
1	E	162/163 (99%)	154 (95%)	8 (5%)	25	23
1	F	162/163 (99%)	155 (96%)	7 (4%)	29	29
2	M	196/202 (97%)	185 (94%)	11 (6%)	21	18
2	N	196/202 (97%)	187 (95%)	9 (5%)	27	26
2	O	196/202 (97%)	188 (96%)	8 (4%)	30	31
2	P	196/202 (97%)	183 (93%)	13 (7%)	16	14
2	Q	196/202 (97%)	185 (94%)	11 (6%)	21	18
2	R	196/202 (97%)	184 (94%)	12 (6%)	18	16
All	All	2148/2190 (98%)	2045 (95%)	103 (5%)	25	24

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	P	416	LEU
1	E	165	GLN
2	R	478	LEU
2	P	434	ASP
1	E	4	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
2	O	497	ASN
1	F	165	GLN
1	D	165	GLN
2	R	497	ASN
2	Q	497	ASN

5.3.3 RNA ⓘ

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](#)

Of 18 ligands modelled in this entry, 6 are monoatomic - leaving 12 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	$\# Z > 2$
3	CO3	N	550	4	0,3,3	-	-	0,3,3	-	-
5	BME	N	601	2	3,3,3	0.23	0	1,2,2	0.52	0
3	CO3	M	550	4	0,3,3	-	-	0,3,3	-	-
5	BME	P	601	2	3,3,3	0.46	0	1,2,2	1.15	0
3	CO3	Q	550	4	0,3,3	-	-	0,3,3	-	-
3	CO3	O	550	4	0,3,3	-	-	0,3,3	-	-
5	BME	R	601	2	3,3,3	0.29	0	1,2,2	0.07	0
5	BME	Q	601	2	3,3,3	0.55	0	1,2,2	1.14	0
5	BME	M	601	2	3,3,3	0.37	0	1,2,2	0.46	0
5	BME	O	601	2	3,3,3	0.34	0	1,2,2	1.01	0
3	CO3	P	550	4	0,3,3	-	-	0,3,3	-	-
3	CO3	R	550	4	0,3,3	-	-	0,3,3	-	-

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
5	BME	N	601	2	-	0/1/1/1	-
5	BME	P	601	2	-	0/1/1/1	-
5	BME	R	601	2	-	0/1/1/1	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	BME	Q	601	2	-	0/1/1/1	-
5	BME	M	601	2	-	1/1/1/1	-
5	BME	O	601	2	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	M	601	BME	O1-C1-C2-S2

There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	M	550	CO3	1	0
5	Q	601	BME	2	0
5	O	601	BME	1	0
3	P	550	CO3	1	0
3	R	550	CO3	1	0

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](#)

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.