



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 12, 2021 – 08:27 AM EDT

PDB ID : 1YKM  
Title : Protocatechuate 3,4-Dioxygenase Y408E mutant  
Authors : Brown, C.K.; Ohlendorf, D.H.  
Deposited on : 2005-01-18  
Resolution : 2.22 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 : 2.23.2  
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) : 2.23.2

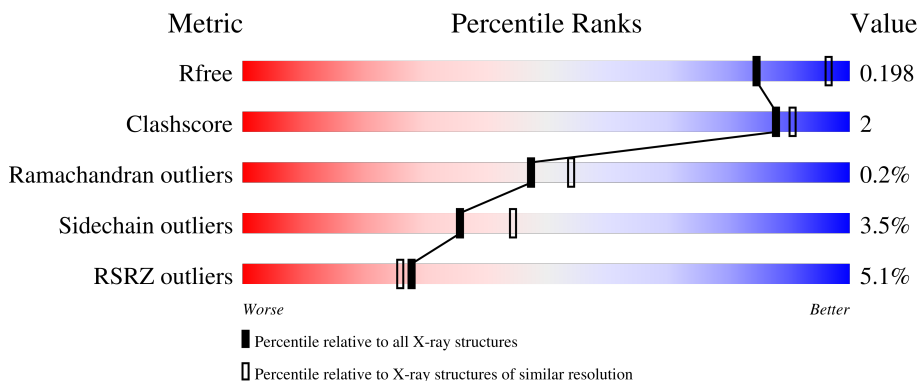
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.22 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.



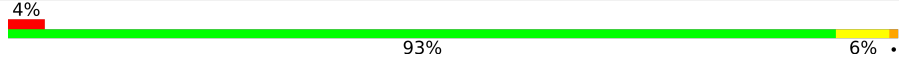
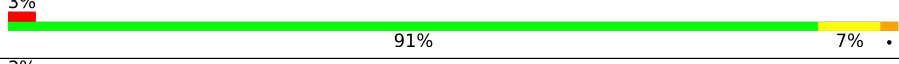
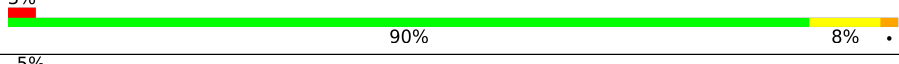
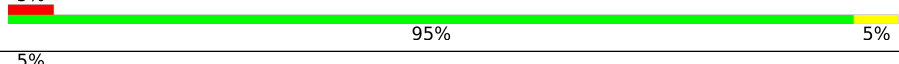
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5912 (2.24-2.20)
Clashscore	141614	6646 (2.24-2.20)
Ramachandran outliers	138981	6543 (2.24-2.20)
Sidechain outliers	138945	6544 (2.24-2.20)
RSRZ outliers	127900	5797 (2.24-2.20)

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  $\geq 3$ , 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	200	6% (poor fit), 95% (0-1 outliers), 3% (2 outliers), 2% (3+ outliers)
1	C	200	6% (poor fit), 95% (0-1 outliers), 3% (2 outliers), 2% (3+ outliers)
1	E	200	4% (poor fit), 94% (0-1 outliers), 6% (2 outliers), 2% (3+ outliers)
1	G	200	6% (poor fit), 95% (0-1 outliers), 3% (2 outliers), 2% (3+ outliers)
1	I	200	8% (poor fit), 94% (0-1 outliers), 5% (2 outliers), 2% (3+ outliers)

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	K	200	 9% 94% 5% .
2	B	238	 4% 93% 6% .
2	D	238	 4% 92% 7% .
2	F	238	 3% 91% 7% .
2	H	238	 3% 90% 8% .
2	J	238	 5% 95% 5% .
2	L	238	 5% 93% 5% .

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 22002 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 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	200	1571	993	276	299	3	0	0	0
1	C	200	1571	993	276	299	3	0	0	0
1	E	200	1571	993	276	299	3	0	0	0
1	G	200	1571	993	276	299	3	0	0	0
1	I	200	1571	993	276	299	3	0	0	0
1	K	200	1571	993	276	299	3	0	0	0

- Molecule 2 is a protein called Protocatechuate 3,4-dioxygenase beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	238	1876	1186	342	339	9	0	0	0
2	D	238	1876	1186	342	339	9	0	0	0
2	F	238	1876	1186	342	339	9	0	0	0
2	H	238	1876	1186	342	339	9	0	0	0
2	J	238	1876	1186	342	339	9	0	0	0
2	L	238	1876	1186	342	339	9	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	408	GLU	TYR	engineered mutation	UNP P00437

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	429	CME	CYS	modified residue	UNP P00437
D	408	GLU	TYR	engineered mutation	UNP P00437
D	429	CME	CYS	modified residue	UNP P00437
F	408	GLU	TYR	engineered mutation	UNP P00437
F	429	CME	CYS	modified residue	UNP P00437
H	408	GLU	TYR	engineered mutation	UNP P00437
H	429	CME	CYS	modified residue	UNP P00437
J	408	GLU	TYR	engineered mutation	UNP P00437
J	429	CME	CYS	modified residue	UNP P00437
L	408	GLU	TYR	engineered mutation	UNP P00437
L	429	CME	CYS	modified residue	UNP P00437

- Molecule 3 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Fe 1 1	0	0
3	D	1	Total Fe 1 1	0	0
3	F	1	Total Fe 1 1	0	0
3	H	1	Total Fe 1 1	0	0
3	J	1	Total Fe 1 1	0	0
3	L	1	Total Fe 1 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	78	Total O 78 78	0	0
4	B	140	Total O 140 140	0	0
4	C	78	Total O 78 78	0	0
4	D	146	Total O 146 146	0	0
4	E	76	Total O 76 76	0	0
4	F	141	Total O 141 141	0	0

*Continued on next page...*

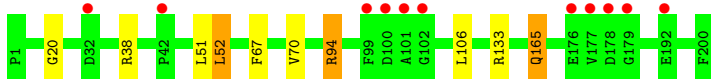
*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	G	77	Total O 77 77	0	0
4	H	137	Total O 137 137	0	0
4	I	75	Total O 75 75	0	0
4	J	147	Total O 147 147	0	0
4	K	78	Total O 78 78	0	0
4	L	141	Total O 141 141	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.

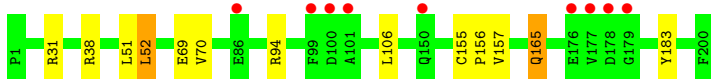
- Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain



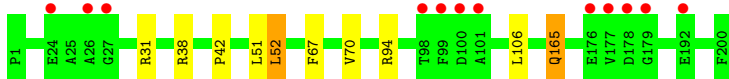
- Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain



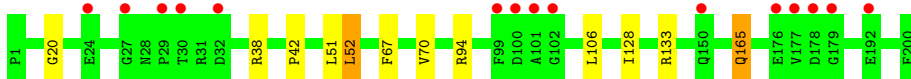
- Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain



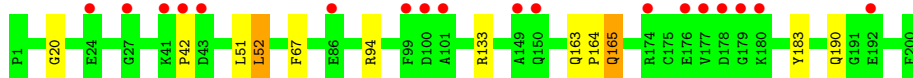
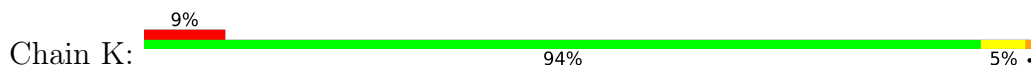
- Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain



- Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain



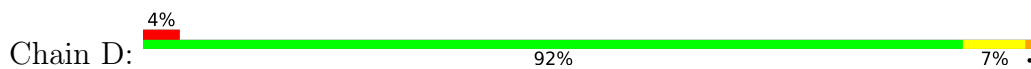
- Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain



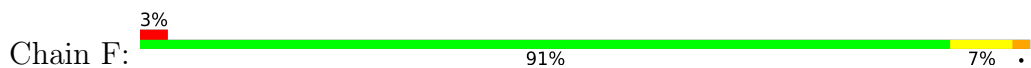
• Molecule 2: Protocatechuate 3,4-dioxygenase beta chain



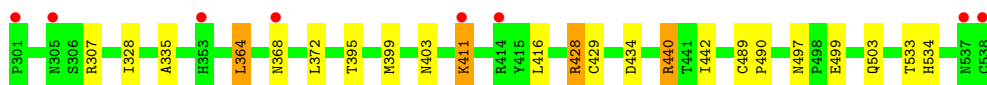
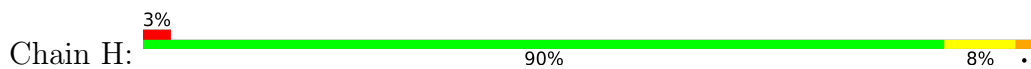
• Molecule 2: Protocatechuate 3,4-dioxygenase beta chain



• Molecule 2: Protocatechuate 3,4-dioxygenase beta chain



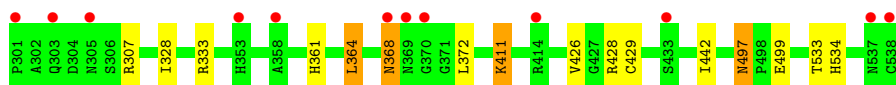
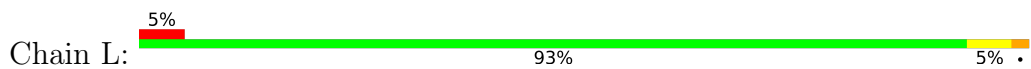
• Molecule 2: Protocatechuate 3,4-dioxygenase beta chain



• Molecule 2: Protocatechuate 3,4-dioxygenase beta chain



• Molecule 2: Protocatechuate 3,4-dioxygenase beta chain





## 4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	196.61Å 127.73Å 134.32Å 90.00° 97.66° 90.00°	Depositor
Resolution (Å)	7.81 – 2.22 19.99 – 1.95	Depositor EDS
% Data completeness (in resolution range)	69.4 (7.81-2.22) 68.0 (19.99-1.95)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.00 (at 1.94Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.169 , 0.199 0.170 , 0.198	Depositor DCC
$R_{free}$ test set	1950 reflections (1.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.9	Xtrriage
Anisotropy	0.260	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.44 , 52.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	22002	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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: FE, CME

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	0.62	0/1611	0.75	1/2195 (0.0%)
1	C	0.61	0/1611	0.74	1/2195 (0.0%)
1	E	0.63	0/1611	0.74	1/2195 (0.0%)
1	G	0.63	0/1611	0.75	1/2195 (0.0%)
1	I	0.64	0/1611	0.74	1/2195 (0.0%)
1	K	0.63	0/1611	0.73	1/2195 (0.0%)
2	B	0.61	0/1920	0.78	0/2612
2	D	0.61	0/1920	0.78	0/2612
2	F	0.61	0/1920	0.77	0/2612
2	H	0.62	0/1920	0.76	1/2612 (0.0%)
2	J	0.63	0/1920	0.76	0/2612
2	L	0.62	0/1920	0.77	0/2612
All	All	0.62	0/21186	0.76	7/28842 (0.0%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	K	52	LEU	CA-CB-CG	7.61	132.80	115.30
1	A	52	LEU	CA-CB-CG	7.48	132.50	115.30
1	E	52	LEU	CA-CB-CG	7.44	132.40	115.30
1	G	52	LEU	CA-CB-CG	7.19	131.84	115.30
1	I	52	LEU	CA-CB-CG	6.94	131.27	115.30
1	C	52	LEU	CA-CB-CG	6.19	129.53	115.30
2	H	416	LEU	CA-CB-CG	5.27	127.43	115.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-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 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	6	0
1	C	1571	0	1499	7	0
1	E	1571	0	1499	8	0
1	G	1571	0	1499	7	0
1	I	1571	0	1499	9	0
1	K	1571	0	1499	8	0
2	B	1876	0	1820	6	0
2	D	1876	0	1820	9	0
2	F	1876	0	1820	11	0
2	H	1876	0	1820	10	0
2	J	1876	0	1820	4	0
2	L	1876	0	1820	8	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
3	F	1	0	0	0	0
3	H	1	0	0	0	0
3	J	1	0	0	0	0
3	L	1	0	0	0	0
4	A	78	0	0	0	0
4	B	140	0	0	1	0
4	C	78	0	0	1	0
4	D	146	0	0	1	0
4	E	76	0	0	0	0
4	F	141	0	0	0	0
4	G	77	0	0	0	0
4	H	137	0	0	1	0
4	I	75	0	0	0	0
4	J	147	0	0	1	0
4	K	78	0	0	0	0
4	L	141	0	0	0	0
All	All	22002	0	19914	79	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 2.

All (79) 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:I:165:GLN:H	1:I:165:GLN:HE21	1.07	1.00
1:C:165:GLN:HE21	1:C:165:GLN:H	1.16	0.92
1:G:165:GLN:HE21	1:G:165:GLN:H	1.36	0.71
1:I:165:GLN:HE21	1:I:165:GLN:N	1.90	0.59
1:C:67:PHE:HZ	1:C:94:ARG:HD2	1.68	0.59
1:I:51:LEU:HD12	1:I:106:LEU:HD23	1.85	0.59
2:F:411:LYS:H	2:F:411:LYS:HZ2	1.49	0.58
1:G:51:LEU:HD12	1:G:106:LEU:HD23	1.85	0.58
1:I:165:GLN:H	1:I:165:GLN:NE2	1.90	0.57
2:H:335:ALA:HB2	2:L:328:ILE:HD12	1.88	0.56
2:L:307:ARG:HG2	2:L:533:THR:HG22	1.90	0.54
2:J:497:ASN:HD22	2:J:499:GLU:H	1.56	0.53
2:D:307:ARG:HG2	2:D:533:THR:HG22	1.90	0.53
1:E:165:GLN:NE2	1:E:165:GLN:H	2.07	0.53
1:K:67:PHE:HZ	1:K:94:ARG:HD2	1.73	0.53
2:H:364:LEU:HD22	2:H:440:ARG:HD3	1.91	0.53
1:C:20:GLY:HA2	2:D:426:VAL:HG13	1.89	0.52
1:A:67:PHE:HZ	1:A:94:ARG:HD2	1.73	0.52
2:D:497:ASN:HD22	2:D:499:GLU:H	1.58	0.52
1:G:67:PHE:HZ	1:G:94:ARG:HD2	1.73	0.52
1:E:51:LEU:HD12	1:E:106:LEU:HD23	1.91	0.52
2:F:497:ASN:HD22	2:F:499:GLU:H	1.58	0.51
1:G:165:GLN:H	1:G:165:GLN:NE2	2.05	0.51
2:B:335:ALA:HB2	2:F:328:ILE:HD12	1.93	0.50
1:I:20:GLY:HA2	2:J:426:VAL:HG13	1.93	0.49
1:A:51:LEU:HD12	1:A:106:LEU:HD23	1.94	0.49
1:A:20:GLY:HA2	2:B:426:VAL:HG13	1.95	0.48
1:K:165:GLN:NE2	1:K:165:GLN:H	2.11	0.48
1:A:165:GLN:NE2	1:A:165:GLN:H	2.10	0.48
2:D:538:CYS:HA	2:H:440:ARG:HH22	1.78	0.48
1:I:67:PHE:HZ	1:I:94:ARG:HD2	1.78	0.48
2:B:307:ARG:HG2	2:B:533:THR:HG22	1.95	0.48
2:H:328:ILE:HD12	2:J:335:ALA:HB2	1.96	0.48
1:C:165:GLN:HE21	1:C:165:GLN:N	1.98	0.48
2:D:328:ILE:HD12	2:F:335:ALA:HB2	1.94	0.47
2:H:411:LYS:H	2:H:411:LYS:NZ	2.13	0.47
2:F:411:LYS:H	2:F:411:LYS:NZ	2.13	0.46
2:H:307:ARG:HG2	2:H:533:THR:HG22	1.97	0.46
1:A:70:VAL:HG21	1:A:106:LEU:HD21	1.98	0.46
1:K:67:PHE:CZ	1:K:94:ARG:HD2	2.50	0.46
1:G:67:PHE:CZ	1:G:94:ARG:HD2	2.51	0.46
1:I:70:VAL:HG21	1:I:106:LEU:HD21	1.98	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:497:ASN:HD22	2:B:499:GLU:H	1.64	0.45
2:L:497:ASN:HD22	2:L:499:GLU:H	1.64	0.45
1:C:67:PHE:CZ	1:C:94:ARG:HD2	2.50	0.44
2:F:537:ASN:O	2:F:538:CYS:HB2	2.17	0.44
2:F:364:LEU:HD22	2:F:440:ARG:HD3	2.00	0.44
1:G:31:ARG:NH1	2:H:428:ARG:HG2	2.32	0.44
1:K:163:GLN:HA	1:K:164:PRO:HD2	1.90	0.44
1:E:31:ARG:NH1	2:F:428:ARG:HG2	2.32	0.44
1:C:132:ALA:HB3	1:C:135:ILE:HD12	2.00	0.44
2:D:368:ASN:ND2	2:D:371:GLY:H	2.16	0.43
2:D:403:ASN:HB2	4:D:1620:HOH:O	2.18	0.43
2:L:411:LYS:HZ3	2:L:411:LYS:H	1.66	0.43
2:J:403:ASN:HB2	4:J:4620:HOH:O	2.17	0.43
2:L:361:HIS:H	2:L:361:HIS:CD2	2.36	0.43
1:E:69:GLU:HG2	1:E:94:ARG:HG2	2.00	0.43
2:H:403:ASN:HB2	4:H:3620:HOH:O	2.18	0.43
2:L:364:LEU:HD11	2:L:442:ILE:HG23	2.01	0.43
2:B:328:ILE:HD12	2:D:335:ALA:HB2	1.99	0.42
2:D:361:HIS:CD2	2:D:361:HIS:H	2.37	0.42
1:A:67:PHE:CZ	1:A:94:ARG:HD2	2.52	0.42
2:F:359:HIS:O	2:F:366:ASN:HB3	2.19	0.42
1:I:67:PHE:CZ	1:I:94:ARG:HD2	2.54	0.42
1:K:165:GLN:H	1:K:165:GLN:HE21	1.67	0.42
1:K:51:LEU:HD23	1:K:183:TYR:HB2	2.01	0.42
2:H:489:CYS:HA	2:H:490:PRO:HD3	1.92	0.41
2:H:364:LEU:HD11	2:H:442:ILE:HG23	2.02	0.41
1:G:70:VAL:HG21	1:G:106:LEU:HD21	2.03	0.41
1:E:51:LEU:HD23	1:E:183:TYR:HB2	2.03	0.41
1:E:155:CYS:HA	1:E:156:PRO:HD3	1.95	0.41
1:E:157:VAL:HG13	2:F:339:ILE:HG21	2.03	0.41
2:F:361:HIS:H	2:F:361:HIS:CD2	2.38	0.41
1:K:20:GLY:HA2	2:L:426:VAL:HG13	2.03	0.41
1:E:70:VAL:HG21	1:E:106:LEU:HD21	2.02	0.41
1:C:74:ASP:HB2	4:C:1690:HOH:O	2.21	0.40
1:I:70:VAL:HG12	1:I:128:ILE:HG12	2.02	0.40
2:B:403:ASN:HB2	4:B:620:HOH:O	2.21	0.40
1:K:190:GLN:HG3	2:L:333:ARG:HG2	2.04	0.40

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%)	193 (98%)	5 (2%)	0	100	100
1	C	198/200 (99%)	194 (98%)	4 (2%)	0	100	100
1	E	198/200 (99%)	193 (98%)	5 (2%)	0	100	100
1	G	198/200 (99%)	192 (97%)	6 (3%)	0	100	100
1	I	198/200 (99%)	194 (98%)	4 (2%)	0	100	100
1	K	198/200 (99%)	193 (98%)	5 (2%)	0	100	100
2	B	235/238 (99%)	226 (96%)	8 (3%)	1 (0%)	34	37
2	D	235/238 (99%)	229 (97%)	5 (2%)	1 (0%)	34	37
2	F	235/238 (99%)	226 (96%)	8 (3%)	1 (0%)	34	37
2	H	235/238 (99%)	228 (97%)	7 (3%)	0	100	100
2	J	235/238 (99%)	226 (96%)	8 (3%)	1 (0%)	34	37
2	L	235/238 (99%)	225 (96%)	9 (4%)	1 (0%)	34	37
All	All	2598/2628 (99%)	2519 (97%)	74 (3%)	5 (0%)	47	54

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	368	ASN
2	D	368	ASN
2	F	368	ASN
2	J	368	ASN
2	L	368	ASN

### 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%)	157 (97%)	5 (3%)	40	50
1	C	162/163 (99%)	158 (98%)	4 (2%)	47	58
1	E	162/163 (99%)	159 (98%)	3 (2%)	57	69
1	G	162/163 (99%)	158 (98%)	4 (2%)	47	58
1	I	162/163 (99%)	157 (97%)	5 (3%)	40	50
1	K	162/163 (99%)	158 (98%)	4 (2%)	47	58
2	B	199/201 (99%)	191 (96%)	8 (4%)	31	38
2	D	199/201 (99%)	192 (96%)	7 (4%)	36	44
2	F	199/201 (99%)	189 (95%)	10 (5%)	24	28
2	H	199/201 (99%)	186 (94%)	13 (6%)	17	18
2	J	199/201 (99%)	193 (97%)	6 (3%)	41	51
2	L	199/201 (99%)	192 (96%)	7 (4%)	36	44
All	All	2166/2184 (99%)	2090 (96%)	76 (4%)	36	44

All (76) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	ARG
1	A	52	LEU
1	A	94	ARG
1	A	133	ARG
1	A	165	GLN
2	B	364	LEU
2	B	368	ASN
2	B	372	LEU
2	B	395	THR
2	B	411	LYS
2	B	428	ARG
2	B	497	ASN
2	B	534	HIS
1	C	38	ARG
1	C	42	PRO
1	C	52	LEU
1	C	165	GLN
2	D	364	LEU
2	D	395	THR

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
2	D	399	MET
2	D	411	LYS
2	D	428	ARG
2	D	497	ASN
2	D	534	HIS
1	E	38	ARG
1	E	52	LEU
1	E	165	GLN
2	F	368	ASN
2	F	395	THR
2	F	399	MET
2	F	408	GLU
2	F	411	LYS
2	F	428	ARG
2	F	478	LEU
2	F	497	ASN
2	F	534	HIS
2	F	538	CYS
1	G	38	ARG
1	G	42	PRO
1	G	52	LEU
1	G	165	GLN
2	H	364	LEU
2	H	368	ASN
2	H	372	LEU
2	H	395	THR
2	H	399	MET
2	H	411	LYS
2	H	428	ARG
2	H	434	ASP
2	H	440	ARG
2	H	497	ASN
2	H	499	GLU
2	H	503	GLN
2	H	534	HIS
1	I	38	ARG
1	I	42	PRO
1	I	52	LEU
1	I	133	ARG
1	I	165	GLN
2	J	395	THR
2	J	399	MET

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	J	411	LYS
2	J	428	ARG
2	J	497	ASN
2	J	534	HIS
1	K	42	PRO
1	K	52	LEU
1	K	133	ARG
1	K	165	GLN
2	L	364	LEU
2	L	368	ASN
2	L	372	LEU
2	L	411	LYS
2	L	428	ARG
2	L	497	ASN
2	L	534	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (44) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	163	GLN
1	A	165	GLN
2	B	361	HIS
2	B	368	ASN
2	B	412	ASN
2	B	497	ASN
2	B	503	GLN
1	C	163	GLN
1	C	165	GLN
2	D	361	HIS
2	D	368	ASN
2	D	412	ASN
2	D	422	ASN
2	D	497	ASN
2	D	503	GLN
1	E	163	GLN
1	E	165	GLN
2	F	361	HIS
2	F	368	ASN
2	F	422	ASN
2	F	497	ASN
1	G	59	ASN
1	G	165	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	H	361	HIS
2	H	368	ASN
2	H	412	ASN
2	H	422	ASN
2	H	497	ASN
2	H	503	GLN
1	I	165	GLN
2	J	334	GLN
2	J	361	HIS
2	J	368	ASN
2	J	412	ASN
2	J	422	ASN
2	J	497	ASN
2	J	503	GLN
1	K	165	GLN
2	L	361	HIS
2	L	368	ASN
2	L	412	ASN
2	L	422	ASN
2	L	497	ASN
2	L	503	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](#)

6 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	$\# Z  > 2$
2	CME	D	429	2	8,9,10	0.86	1 (12%)	5,9,11	1.17	1 (20%)
2	CME	H	429	2	8,9,10	0.89	1 (12%)	5,9,11	1.17	1 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	CME	B	429	2	8,9,10	0.86	1 (12%)	5,9,11	1.15	1 (20%)
2	CME	L	429	2	8,9,10	0.91	1 (12%)	5,9,11	1.16	1 (20%)
2	CME	F	429	2	8,9,10	0.91	1 (12%)	5,9,11	1.11	1 (20%)
2	CME	J	429	2	8,9,10	0.88	1 (12%)	5,9,11	1.16	1 (20%)

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	CME	D	429	2	-	2/5/8/10	-
2	CME	H	429	2	-	1/5/8/10	-
2	CME	B	429	2	-	2/5/8/10	-
2	CME	L	429	2	-	1/5/8/10	-
2	CME	F	429	2	-	2/5/8/10	-
2	CME	J	429	2	-	1/5/8/10	-

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	L	429	CME	OH-CZ	-2.16	1.31	1.42
2	F	429	CME	OH-CZ	-2.15	1.31	1.42
2	H	429	CME	OH-CZ	-2.14	1.31	1.42
2	B	429	CME	OH-CZ	-2.12	1.31	1.42
2	D	429	CME	OH-CZ	-2.12	1.31	1.42
2	J	429	CME	OH-CZ	-2.11	1.31	1.42

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	429	CME	OH-CZ-CE	2.50	120.69	110.83
2	B	429	CME	OH-CZ-CE	2.49	120.64	110.83
2	L	429	CME	OH-CZ-CE	2.47	120.59	110.83
2	H	429	CME	OH-CZ-CE	2.46	120.55	110.83
2	J	429	CME	OH-CZ-CE	2.45	120.51	110.83
2	F	429	CME	OH-CZ-CE	2.37	120.17	110.83

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	429	CME	SD-CE-CZ-OH
2	F	429	CME	SD-CE-CZ-OH
2	L	429	CME	SD-CE-CZ-OH
2	B	429	CME	SD-CE-CZ-OH
2	B	429	CME	CZ-CE-SD-SG
2	F	429	CME	CZ-CE-SD-SG
2	J	429	CME	CZ-CE-SD-SG
2	D	429	CME	CZ-CE-SD-SG
2	H	429	CME	CZ-CE-SD-SG

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	200/200 (100%)	-0.14	11 (5%) 25 23	8, 24, 51, 64	0
1	C	200/200 (100%)	-0.09	13 (6%) 18 17	8, 24, 52, 63	0
1	E	200/200 (100%)	-0.09	9 (4%) 33 31	8, 24, 51, 64	0
1	G	200/200 (100%)	-0.09	12 (6%) 21 20	9, 24, 52, 64	0
1	I	200/200 (100%)	0.02	15 (7%) 14 12	10, 26, 52, 64	0
1	K	200/200 (100%)	0.12	18 (9%) 9 8	11, 27, 52, 64	0
2	B	237/238 (99%)	-0.40	9 (3%) 40 38	9, 16, 42, 65	0
2	D	237/238 (99%)	-0.41	9 (3%) 40 38	9, 16, 42, 64	0
2	F	237/238 (99%)	-0.48	7 (2%) 50 48	9, 16, 42, 64	0
2	H	237/238 (99%)	-0.45	8 (3%) 45 43	9, 16, 42, 65	0
2	J	237/238 (99%)	-0.43	11 (4%) 32 30	11, 19, 43, 65	0
2	L	237/238 (99%)	-0.44	12 (5%) 28 26	10, 18, 44, 64	0
All	All	2622/2628 (99%)	-0.26	134 (5%) 28 26	8, 21, 48, 65	0

All (134) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	99	PHE	7.6
2	B	301	PRO	7.3
2	L	370	GLY	7.0
1	I	101	ALA	7.0
1	K	177	VAL	6.9
2	D	369	ASN	6.3
1	K	99	PHE	6.3
2	B	538	CYS	6.2
1	A	99	PHE	6.1
2	J	370	GLY	6.1
1	C	100	ASP	6.0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
2	L	301	PRO	5.8
1	K	178	ASP	5.5
2	D	368	ASN	5.5
2	L	368	ASN	5.5
2	J	368	ASN	5.4
2	B	370	GLY	5.3
1	C	99	PHE	5.3
2	D	301	PRO	5.3
2	F	368	ASN	5.2
2	H	538	CYS	5.2
2	F	301	PRO	5.2
2	J	538	CYS	5.0
2	B	368	ASN	5.0
1	I	99	PHE	5.0
2	L	369	ASN	4.9
1	I	177	VAL	4.9
2	J	369	ASN	4.8
1	A	178	ASP	4.7
1	K	179	GLY	4.7
1	I	100	ASP	4.7
2	F	370	GLY	4.6
2	H	301	PRO	4.6
1	G	178	ASP	4.5
1	A	101	ALA	4.5
1	G	99	PHE	4.4
1	I	178	ASP	4.4
1	K	100	ASP	4.2
1	G	176	GLU	4.2
2	B	369	ASN	4.2
1	G	177	VAL	4.2
2	D	370	GLY	4.1
2	D	538	CYS	4.0
2	L	538	CYS	4.0
1	C	98	THR	3.9
2	F	369	ASN	3.9
2	H	414	ARG	3.9
1	G	100	ASP	3.8
2	B	414	ARG	3.7
1	E	177	VAL	3.7
1	I	176	GLU	3.7
2	J	537	ASN	3.7
2	B	537	ASN	3.7

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	E	100	ASP	3.6
1	K	176	GLU	3.6
1	C	101	ALA	3.6
1	I	179	GLY	3.6
2	J	414	ARG	3.6
1	A	100	ASP	3.5
2	F	414	ARG	3.5
2	F	538	CYS	3.5
1	G	101	ALA	3.5
1	E	178	ASP	3.5
1	K	180	LYS	3.4
1	I	32	ASP	3.3
1	E	176	GLU	3.2
1	K	86	GLU	3.2
1	K	43	ASP	3.2
1	K	101	ALA	3.2
2	J	301	PRO	3.2
1	A	192	GLU	3.2
2	L	305	ASN	3.2
1	E	101	ALA	3.1
1	G	192	GLU	3.0
2	H	368	ASN	3.0
1	G	179	GLY	3.0
1	C	192	GLU	3.0
2	L	537	ASN	3.0
2	B	305	ASN	3.0
2	L	414	ARG	3.0
1	A	179	GLY	2.9
1	G	27	GLY	2.9
1	I	30	THR	2.9
2	H	411	LYS	2.9
2	D	411	LYS	2.9
1	G	24	GLU	2.9
2	J	433	SER	2.8
1	A	176	GLU	2.8
1	C	29	PRO	2.8
1	I	24	GLU	2.8
1	K	42	PRO	2.8
2	B	303	GLN	2.8
1	E	179	GLY	2.7
1	C	150	GLN	2.7
1	K	24	GLU	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	I	150	GLN	2.7
2	D	414	ARG	2.7
1	K	41	LYS	2.6
1	A	177	VAL	2.6
2	J	411	LYS	2.5
2	J	371	GLY	2.5
1	K	150	GLN	2.5
2	D	537	ASN	2.4
2	J	412	ASN	2.4
1	A	102	GLY	2.4
2	H	305	ASN	2.3
1	E	86	GLU	2.3
1	C	86	GLU	2.3
1	K	149	ALA	2.3
1	E	150	GLN	2.3
2	L	353	HIS	2.3
1	I	29	PRO	2.3
2	L	303	GLN	2.3
1	C	27	GLY	2.2
2	H	537	ASN	2.2
1	I	27	GLY	2.2
1	C	24	GLU	2.2
1	K	192	GLU	2.2
2	L	433	SER	2.2
1	K	27	GLY	2.2
1	C	32	ASP	2.1
2	L	358	ALA	2.1
1	I	192	GLU	2.1
1	K	174	ARG	2.1
1	A	42	PRO	2.1
2	H	353	HIS	2.1
1	I	102	GLY	2.1
1	A	32	ASP	2.1
2	D	305	ASN	2.1
1	G	26	ALA	2.1
1	C	180	LYS	2.0
1	G	98	THR	2.0
1	C	179	GLY	2.0
2	F	537	ASN	2.0



## 6.2 Non-standard residues in protein, DNA, RNA chains [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	CME	J	429	10/11	0.94	0.10	19,21,35,38	0
2	CME	D	429	10/11	0.95	0.11	16,18,35,36	0
2	CME	H	429	10/11	0.96	0.11	16,18,33,37	0
2	CME	F	429	10/11	0.96	0.11	17,18,35,38	0
2	CME	L	429	10/11	0.96	0.09	16,21,35,37	0
2	CME	B	429	10/11	0.97	0.11	15,18,33,36	0

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	FE	B	600	1/1	0.95	0.04	41,41,41,41	0
3	FE	F	600	1/1	0.98	0.03	39,39,39,39	0
3	FE	H	600	1/1	0.98	0.06	40,40,40,40	0
3	FE	J	600	1/1	0.98	0.03	46,46,46,46	0
3	FE	L	600	1/1	0.98	0.04	38,38,38,38	0
3	FE	D	600	1/1	0.99	0.02	40,40,40,40	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.