



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 16, 2017 – 06:12 AM EDT

PDB ID : 3E2A  
Title : H. influenzae beta-carbonic anhydrase, variant Y181F with 100 mM bicarbonate  
Authors : Rowlett, R.S.; Lee, J.  
Deposited on : unknown  
Resolution : 2.30 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

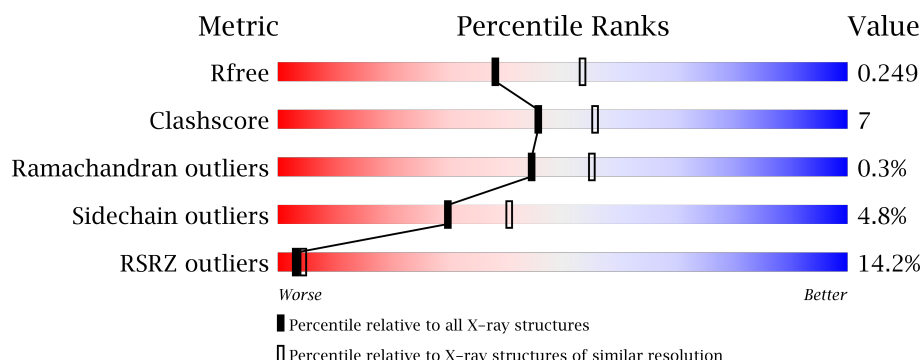
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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))
$R_{free}$	100719	4130 (2.30-2.30)
Clashscore	112137	4751 (2.30-2.30)
Ramachandran outliers	110173	4705 (2.30-2.30)
Sidechain outliers	110143	4704 (2.30-2.30)
RSRZ outliers	101464	4156 (2.30-2.30)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	229	<div> <div>11%</div> <div>78% 13% 9%</div> </div>
1	B	229	<div> <div>17%</div> <div>76% 12% 10%</div> </div>
1	C	229	<div> <div>15%</div> <div>73% 14% 10%</div> </div>
1	D	229	<div> <div>7%</div> <div>73% 14% 11%</div> </div>
1	E	229	<div> <div>15%</div> <div>72% 17% 10%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	229	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	SO4	A	232	-	-	X	X
3	SO4	B	232	-	-	X	-
3	SO4	C	232	-	-	X	-
3	SO4	D	232	-	-	X	-
3	SO4	E	233	-	-	X	-
3	SO4	F	231	-	-	X	X
4	BCT	F	232	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10134 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 Carbonic anhydrase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	209	Total	C	N	O	S	0	0	0
			1677	1067	301	300	9			
1	B	207	Total	C	N	O	S	0	0	0
			1659	1055	296	299	9			
1	C	205	Total	C	N	O	S	0	0	0
			1643	1044	294	296	9			
1	D	204	Total	C	N	O	S	0	0	0
			1632	1038	294	291	9			
1	E	207	Total	C	N	O	S	0	0	0
			1661	1055	298	299	9			
1	F	204	Total	C	N	O	S	0	0	0
			1634	1038	292	295	9			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	181	PHE	TYR	ENGINEERED	UNP P45148
B	181	PHE	TYR	ENGINEERED	UNP P45148
C	181	PHE	TYR	ENGINEERED	UNP P45148
D	181	PHE	TYR	ENGINEERED	UNP P45148
E	181	PHE	TYR	ENGINEERED	UNP P45148
F	181	PHE	TYR	ENGINEERED	UNP P45148

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	1	Total	Zn	0	0
			1	1		
2	E	1	Total	Zn	0	0
			1	1		
2	B	1	Total	Zn	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	1	Total	Zn	0	0
			1	1		
2	A	1	Total	Zn	0	0
			1	1		
2	F	1	Total	Zn	0	0
			1	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



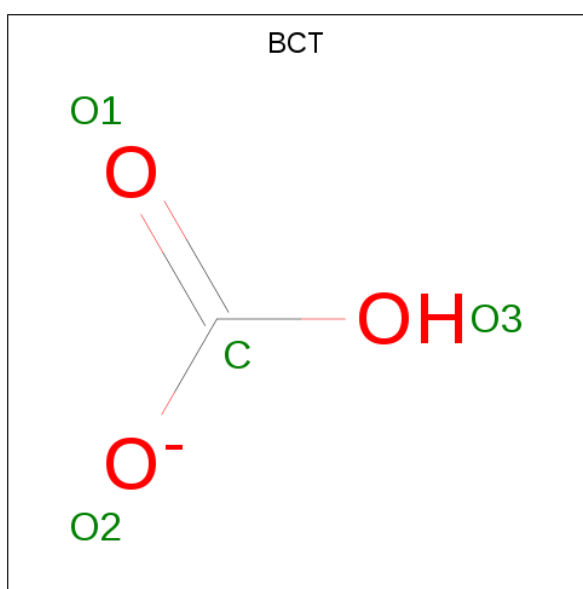
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	E	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is BICARBONATE ION (three-letter code: BCT) (formula:  $\text{CHO}_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	F	1	Total	C	O	0	0
			4	1	3		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	31	Total	O	0	0
			31	31		
5	B	18	Total	O	0	0
			18	18		
5	C	24	Total	O	0	0
			24	24		
5	D	27	Total	O	0	0
			27	27		

*Continued on next page...*

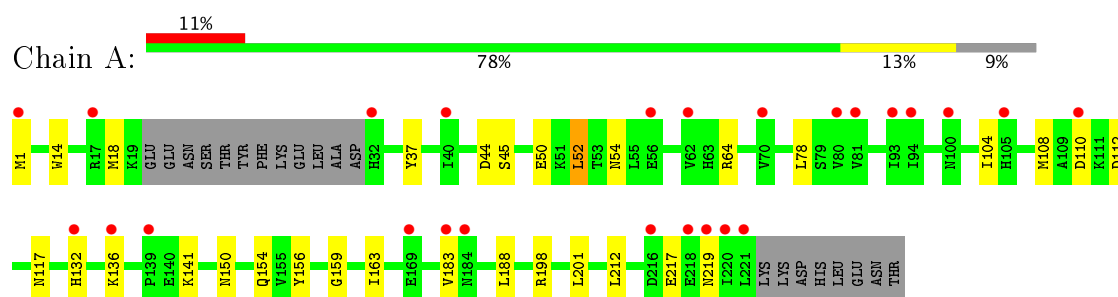
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	E	29	Total	O	0	0
			29	29		
5	F	29	Total	O	0	0
			29	29		

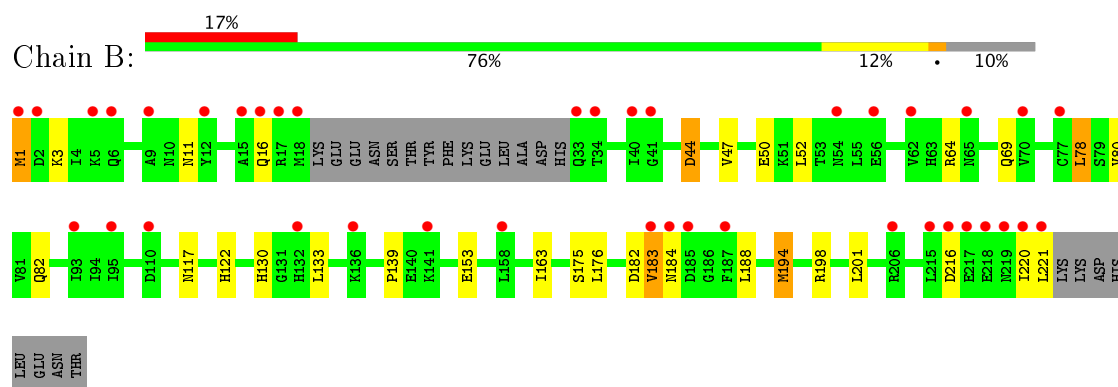
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

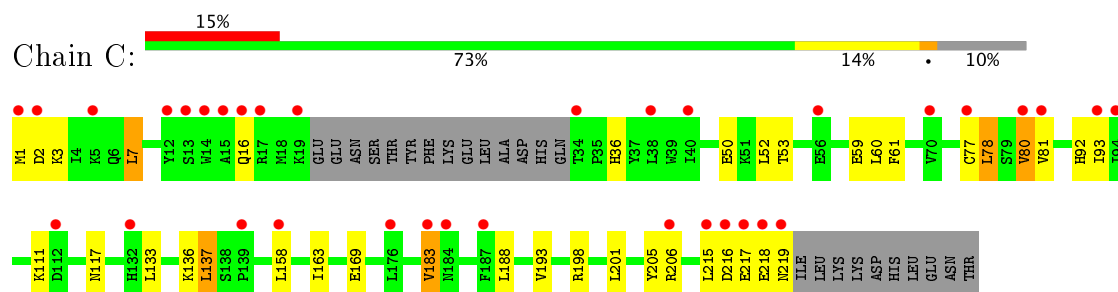
#### • Molecule 1: Carbonic anhydrase 2



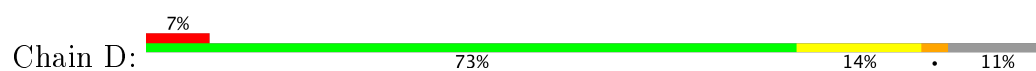
#### • Molecule 1: Carbonic anhydrase 2



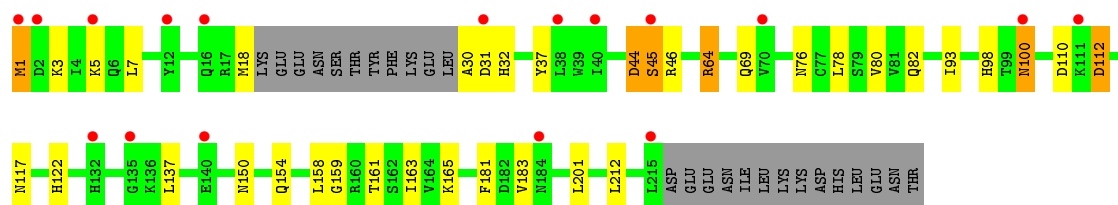
#### • Molecule 1: Carbonic anhydrase 2



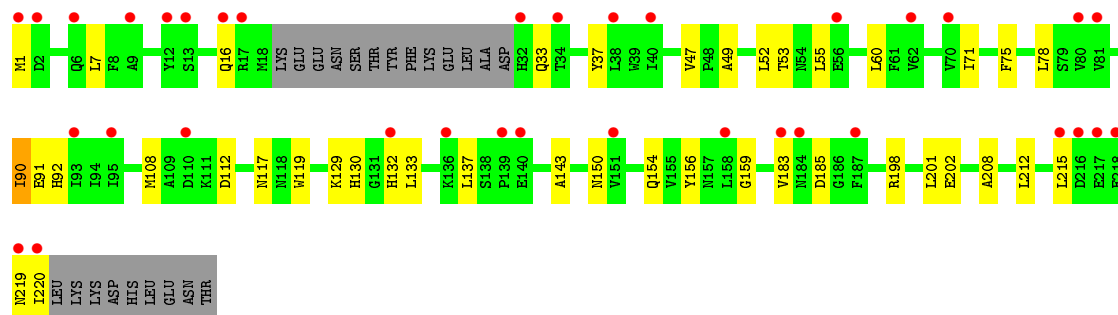
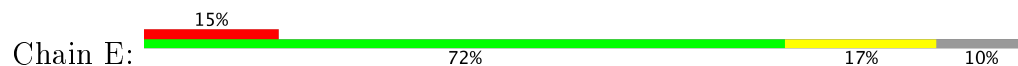
#### • Molecule 1: Carbonic anhydrase 2



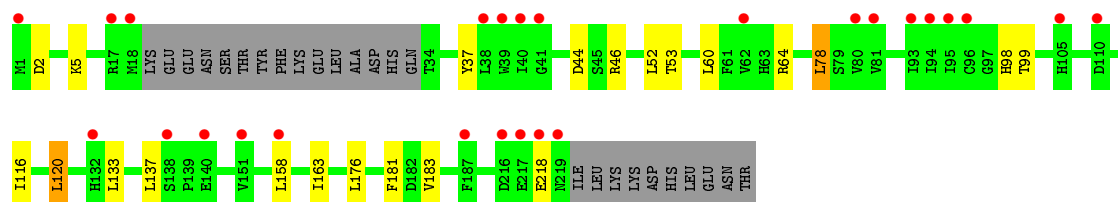
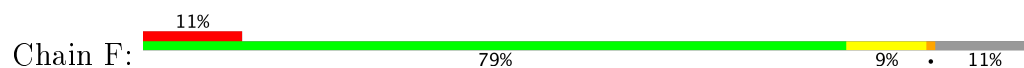




• Molecule 1: Carbonic anhydrase 2



• Molecule 1: Carbonic anhydrase 2



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	230.61Å 145.51Å 53.23Å 90.00° 93.78° 90.00°	Depositor
Resolution (Å)	24.78 – 2.30 24.77 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (24.78-2.30) 99.9 (24.77-2.30)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.65 (at 2.31Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.201 , 0.244 0.207 , 0.249	Depositor DCC
$R_{free}$ test set	7808 reflections (11.22%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.0	Xtriage
Anisotropy	0.236	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 50.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	10134	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.54% 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: BCT, ZN, SO4

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	0.54	0/1713	0.69	3/2318 (0.1%)
1	B	0.55	1/1694 (0.1%)	0.62	1/2294 (0.0%)
1	C	0.51	0/1678	0.61	0/2271
1	D	0.58	0/1668	0.68	2/2259 (0.1%)
1	E	0.53	0/1697	0.63	0/2298
1	F	0.58	0/1669	0.67	3/2260 (0.1%)
All	All	0.55	1/10119 (0.0%)	0.65	9/13700 (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	D	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	11	ASN	CG-ND2	5.55	1.46	1.32

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	44	ASP	CB-CA-C	-6.88	96.64	110.40
1	F	64	ARG	NE-CZ-NH2	-6.83	116.89	120.30
1	A	64	ARG	NE-CZ-NH1	5.89	123.25	120.30
1	A	52	LEU	CA-CB-CG	5.70	128.41	115.30
1	F	64	ARG	NE-CZ-NH1	5.62	123.11	120.30
1	A	64	ARG	NE-CZ-NH2	-5.50	117.55	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	32	HIS	N-CA-C	5.28	125.25	111.00
1	D	64	ARG	NE-CZ-NH2	-5.22	117.69	120.30
1	B	44	ASP	CB-CA-C	-5.11	100.18	110.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	31	ASP	Peptide

## 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	1677	0	1671	20	1
1	B	1659	0	1651	23	0
1	C	1643	0	1634	29	0
1	D	1632	0	1623	30	0
1	E	1661	0	1647	37	0
1	F	1634	0	1621	15	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	10	0	0	3	0
3	B	10	0	0	2	0
3	C	10	0	0	3	0
3	D	10	0	0	3	0
3	E	15	0	0	3	0
3	F	5	0	0	2	0
4	F	4	0	1	2	0
5	A	31	0	0	0	1
5	B	18	0	0	0	0
5	C	24	0	0	2	0
5	D	27	0	0	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	E	29	0	0	0	0
5	F	29	0	0	0	0
All	All	10134	0	9848	145	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 7.

All (145) 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:C:78:LEU:HD22	1:C:163:ILE:CD1	1.80	1.11
1:C:78:LEU:HD22	1:C:163:ILE:HD12	1.31	1.07
1:A:183:VAL:HG11	3:A:232:SO4:O1	1.53	1.07
1:D:78:LEU:HD22	1:D:163:ILE:HD12	1.44	0.98
1:E:220:ILE:HD12	1:E:220:ILE:O	1.64	0.96
1:D:183:VAL:HG21	3:D:232:SO4:O1	1.68	0.93
1:E:47:VAL:HG12	1:E:52:LEU:HD12	1.52	0.92
1:A:78:LEU:HD22	1:A:163:ILE:HD12	1.58	0.86
1:C:78:LEU:CD2	1:C:163:ILE:HD12	2.05	0.86
1:C:80:VAL:HG21	5:C:240:HOH:O	1.78	0.81
1:F:52:LEU:HD22	1:F:181:PHE:CE1	2.15	0.81
1:D:78:LEU:HD22	1:D:163:ILE:CD1	2.17	0.74
1:E:90:ILE:HD11	1:E:92:HIS:O	1.87	0.74
1:D:1:MET:O	1:D:5:LYS:NZ	2.22	0.72
1:C:183:VAL:CG1	3:C:232:SO4:O3	2.39	0.71
1:A:112:ASP:HA	1:A:117:ASN:ND2	2.06	0.69
1:A:183:VAL:CG1	3:A:232:SO4:O1	2.35	0.69
1:B:1:MET:HE2	1:B:1:MET:H1	1.58	0.69
1:B:52:LEU:HD21	1:B:188:LEU:HD21	1.75	0.69
1:E:90:ILE:HD12	1:E:92:HIS:H	1.56	0.68
1:D:45:SER:OG	5:D:249:HOH:O	2.10	0.68
1:A:78:LEU:HD22	1:A:163:ILE:CD1	2.25	0.67
1:D:44:ASP:O	1:D:45:SER:C	2.34	0.66
1:C:53:THR:HG22	1:E:7:LEU:HD21	1.78	0.65
1:E:90:ILE:CD1	1:E:92:HIS:O	2.45	0.65
1:F:52:LEU:CD2	1:F:181:PHE:CE1	2.79	0.64
1:B:1:MET:HE2	1:B:1:MET:N	2.11	0.63
1:C:7:LEU:CD2	1:E:55:LEU:HD11	2.29	0.62
1:E:208:ALA:O	1:E:212:LEU:HD23	1.99	0.62
1:E:108:MET:HE1	1:E:143:ALA:HB2	1.82	0.62
1:F:116:ILE:HG13	1:F:120:LEU:HD22	1.81	0.61

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:50:GLU:OE2	1:B:50:GLU:N	2.31	0.61
1:A:141:LYS:NZ	1:A:219:ASN:ND2	2.49	0.61
1:A:141:LYS:HZ2	1:A:219:ASN:HD22	1.48	0.60
1:B:78:LEU:HD13	1:B:163:ILE:HD12	1.84	0.59
1:A:44:ASP:O	1:A:45:SER:C	2.41	0.59
1:E:47:VAL:CG1	1:E:52:LEU:HD12	2.32	0.58
1:B:183:VAL:CG1	3:B:232:SO4:O4	2.51	0.58
1:D:1:MET:HE3	1:D:3:LYS:HD2	1.85	0.58
1:A:112:ASP:HA	1:A:117:ASN:HD21	1.69	0.57
1:B:176:LEU:O	1:B:194:MET:HE2	2.04	0.57
1:F:37:TYR:HB2	1:F:60:LEU:HD23	1.84	0.57
1:E:130:HIS:HA	1:E:132:HIS:CE1	2.40	0.57
1:D:112:ASP:HA	1:D:117:ASN:HD21	1.70	0.56
1:B:198:ARG:NH1	1:B:201:LEU:HD23	2.20	0.56
1:C:217:GLU:O	1:C:219:ASN:N	2.39	0.56
1:E:183:VAL:HG13	3:E:233:SO4:O2	2.06	0.56
1:C:137:LEU:HD13	1:C:215:LEU:HD13	1.88	0.55
1:C:61:PHE:CE2	1:C:80:VAL:HG12	2.41	0.55
1:D:76:ASN:O	1:D:80:VAL:HG23	2.06	0.55
1:E:183:VAL:HG12	3:E:233:SO4:O3	2.07	0.55
1:A:141:LYS:NZ	1:A:219:ASN:HD22	2.04	0.55
1:B:175:SER:HB2	1:B:194:MET:HE1	1.88	0.55
1:D:64:ARG:HH21	4:F:232:BCT:C	2.19	0.55
1:A:52:LEU:HD21	1:A:188:LEU:HD21	1.89	0.54
1:C:183:VAL:HG12	3:C:232:SO4:O3	2.07	0.54
1:E:49:ALA:HA	1:E:52:LEU:HD13	1.89	0.54
1:C:183:VAL:HG13	3:C:232:SO4:O4	2.08	0.53
1:C:7:LEU:HD21	1:E:55:LEU:HD11	1.91	0.53
1:D:112:ASP:HA	1:D:117:ASN:ND2	2.24	0.53
1:C:80:VAL:CG2	5:C:240:HOH:O	2.44	0.53
1:D:7:LEU:HD21	1:F:53:THR:HG22	1.90	0.53
1:C:137:LEU:CD1	1:C:215:LEU:HD13	2.39	0.52
1:D:1:MET:C	1:D:5:LYS:NZ	2.63	0.52
1:C:50:GLU:HG3	1:C:60:LEU:HD12	1.90	0.52
1:D:100:ASN:HD22	1:D:100:ASN:C	2.13	0.52
1:D:69:GLN:HE22	1:D:122:HIS:HB2	1.73	0.52
1:B:1:MET:CE	1:B:1:MET:N	2.73	0.51
1:C:7:LEU:HD11	1:E:53:THR:HG22	1.91	0.51
1:B:130:HIS:NE2	1:B:153:GLU:OE2	2.34	0.50
1:D:161:THR:O	1:D:165:LYS:HG3	2.10	0.50
1:F:52:LEU:HD21	1:F:181:PHE:CZ	2.46	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1:MET:CE	1:B:3:LYS:HB2	2.41	0.50
1:F:133:LEU:O	1:F:137:LEU:HD13	2.12	0.50
1:B:1:MET:CE	1:B:1:MET:H1	2.24	0.49
1:D:181:PHE:CE2	1:D:183:VAL:HG22	2.48	0.49
1:E:183:VAL:CG1	3:E:233:SO4:O3	2.61	0.49
1:A:141:LYS:HZ3	1:A:219:ASN:ND2	2.10	0.49
1:F:99:THR:O	3:F:231:SO4:O3	2.32	0.48
1:B:69:GLN:HE22	1:B:122:HIS:HB2	1.78	0.48
1:C:136:LYS:O	1:C:219:ASN:ND2	2.42	0.48
1:D:44:ASP:O	1:D:46:ARG:N	2.46	0.48
1:A:150:ASN:O	1:A:154:GLN:HG2	2.13	0.48
1:D:78:LEU:HD12	1:E:75:PHE:CE1	2.49	0.48
1:E:37:TYR:HB2	1:E:60:LEU:HD23	1.94	0.47
1:C:93:ILE:HG21	1:C:158:LEU:HD21	1.97	0.47
1:A:132:HIS:O	1:A:136:LYS:HG3	2.13	0.47
1:E:47:VAL:O	1:E:52:LEU:HD11	2.14	0.47
1:E:220:ILE:CD1	1:E:220:ILE:O	2.52	0.47
1:A:14:TRP:O	1:A:18:MET:HG2	2.14	0.47
1:C:78:LEU:HD22	1:C:163:ILE:HD11	1.87	0.47
1:E:71:ILE:HD12	1:E:119:TRP:CH2	2.50	0.47
1:C:52:LEU:HD21	1:C:188:LEU:HD11	1.95	0.46
1:B:183:VAL:HG13	3:B:232:SO4:O4	2.15	0.46
1:E:159:GLY:HA3	1:E:201:LEU:HD22	1.98	0.46
1:B:1:MET:HE1	1:B:3:LYS:HB2	1.98	0.46
1:D:183:VAL:CG2	3:D:232:SO4:O1	2.51	0.46
1:C:198:ARG:NH1	1:C:201:LEU:HD23	2.31	0.46
1:E:7:LEU:HD23	1:E:7:LEU:O	2.17	0.45
1:B:139:PRO:HD3	1:B:221:LEU:HD13	1.98	0.45
1:F:52:LEU:HD22	1:F:181:PHE:CD1	2.50	0.45
1:E:198:ARG:O	1:E:198:ARG:HD2	2.17	0.45
1:F:52:LEU:CD2	1:F:181:PHE:CZ	3.00	0.45
1:E:137:LEU:HD21	1:E:215:LEU:HD22	1.97	0.45
1:D:82:GLN:HA	1:D:163:ILE:HG21	1.99	0.45
1:D:30:ALA:HB1	1:F:183:VAL:CG2	2.46	0.45
1:B:82:GLN:HA	1:B:163:ILE:HG21	1.99	0.44
1:D:80:VAL:HG22	5:D:233:HOH:O	2.17	0.44
1:E:129:LYS:O	1:E:132:HIS:HE1	2.00	0.44
1:D:93:ILE:HG21	1:D:158:LEU:HD21	2.00	0.44
1:B:52:LEU:CD2	1:B:188:LEU:HD11	2.48	0.44
1:F:158:LEU:CD2	1:F:176:LEU:HD22	2.48	0.44
1:B:133:LEU:HD23	1:B:133:LEU:C	2.38	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:98:HIS:ND1	3:F:231:SO4:O4	2.51	0.44
1:C:193:VAL:HG22	1:C:205:TYR:HA	2.00	0.44
1:A:156:TYR:HA	1:A:201:LEU:HD11	2.00	0.43
1:C:133:LEU:HD23	1:C:133:LEU:C	2.38	0.43
1:C:77:CYS:O	1:C:81:VAL:HG23	2.19	0.43
1:E:156:TYR:OH	1:E:202:GLU:OE2	2.32	0.43
1:F:78:LEU:HD13	1:F:163:ILE:HD12	2.00	0.43
1:A:183:VAL:CG1	3:A:232:SO4:S	3.06	0.43
1:E:132:HIS:CD2	1:E:133:LEU:H	2.36	0.43
1:B:52:LEU:HD21	1:B:188:LEU:HD11	2.00	0.43
1:E:90:ILE:HD12	1:E:92:HIS:N	2.29	0.43
1:A:159:GLY:HA3	1:A:201:LEU:HD22	2.00	0.43
1:E:47:VAL:HG12	1:E:52:LEU:CD1	2.35	0.43
1:B:175:SER:HB2	1:B:194:MET:CE	2.47	0.43
1:D:150:ASN:O	1:D:154:GLN:HG2	2.19	0.43
1:B:182:ASP:OD1	1:B:184:ASN:ND2	2.48	0.43
1:E:90:ILE:HD12	1:E:91:GLU:N	2.34	0.42
1:E:150:ASN:O	1:E:154:GLN:HG2	2.19	0.42
1:F:183:VAL:HG22	1:F:183:VAL:O	2.20	0.42
1:E:37:TYR:HA	1:E:90:ILE:HD13	2.01	0.41
1:D:159:GLY:HA3	1:D:201:LEU:HD22	2.00	0.41
1:E:108:MET:CE	1:E:143:ALA:HB2	2.49	0.41
1:C:92:HIS:NE2	1:E:1:MET:CE	2.83	0.41
1:D:183:VAL:HG21	3:D:232:SO4:S	2.58	0.41
1:C:1:MET:HG3	1:C:2:ASP:N	2.35	0.41
1:D:78:LEU:HD12	1:E:75:PHE:HE1	1.84	0.41
1:A:104:ILE:O	1:A:108:MET:HG3	2.20	0.41
1:D:44:ASP:HB2	1:D:98:HIS:CE1	2.56	0.41
1:C:1:MET:HG2	1:C:3:LYS:HB2	2.03	0.40
1:D:64:ARG:NH2	4:F:232:BCT:O3	2.55	0.40
1:A:78:LEU:HD23	1:A:78:LEU:HA	1.96	0.40
1:C:36:HIS:HD2	1:C:59:GLU:OE1	2.04	0.40

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)
1:A:50:GLU:OE1	5:A:246:HOH:O[2_556]	1.92	0.28



## 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	205/229 (90%)	199 (97%)	6 (3%)	0	100	100
1	B	203/229 (89%)	197 (97%)	6 (3%)	0	100	100
1	C	201/229 (88%)	193 (96%)	7 (4%)	1 (0%)	32	39
1	D	200/229 (87%)	195 (98%)	4 (2%)	1 (0%)	32	39
1	E	203/229 (89%)	197 (97%)	6 (3%)	0	100	100
1	F	200/229 (87%)	193 (96%)	5 (2%)	2 (1%)	18	20
All	All	1212/1374 (88%)	1174 (97%)	34 (3%)	4 (0%)	44	55

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	45	SER
1	F	2	ASP
1	C	218	GLU
1	F	218	GLU

### 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	182/201 (90%)	175 (96%)	7 (4%)	38	52
1	B	180/201 (90%)	168 (93%)	12 (7%)	19	24
1	C	178/201 (89%)	167 (94%)	11 (6%)	21	28

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	176/201 (88%)	167 (95%)	9 (5%)	28	37
1	E	180/201 (90%)	172 (96%)	8 (4%)	33	45
1	F	177/201 (88%)	173 (98%)	4 (2%)	56	73
All	All	1073/1206 (89%)	1022 (95%)	51 (5%)	30	40

All (51) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	37	TYR
1	A	54	ASN
1	A	110	ASP
1	A	198	ARG
1	A	212	LEU
1	A	217	GLU
1	B	1	MET
1	B	16	GLN
1	B	44	ASP
1	B	47	VAL
1	B	64	ARG
1	B	78	LEU
1	B	80	VAL
1	B	117	ASN
1	B	183	VAL
1	B	194	MET
1	B	216	ASP
1	B	220	ILE
1	C	7	LEU
1	C	16	GLN
1	C	78	LEU
1	C	80	VAL
1	C	111	LYS
1	C	117	ASN
1	C	137	LEU
1	C	169	GLU
1	C	183	VAL
1	C	206	ARG
1	C	216	ASP
1	D	1	MET
1	D	18	MET
1	D	37	TYR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	44	ASP
1	D	100	ASN
1	D	110	ASP
1	D	112	ASP
1	D	137	LEU
1	D	212	LEU
1	E	16	GLN
1	E	33	GLN
1	E	78	LEU
1	E	90	ILE
1	E	112	ASP
1	E	117	ASN
1	E	185	ASP
1	E	219	ASN
1	F	5	LYS
1	F	46	ARG
1	F	78	LEU
1	F	120	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (21) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	54	ASN
1	A	117	ASN
1	A	219	ASN
1	B	16	GLN
1	B	54	ASN
1	B	69	GLN
1	B	117	ASN
1	B	132	HIS
1	C	36	HIS
1	C	54	ASN
1	D	36	HIS
1	D	54	ASN
1	D	69	GLN
1	D	100	ASN
1	D	117	ASN
1	E	6	GLN
1	E	36	HIS
1	E	92	HIS
1	E	117	ASN
1	E	132	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	54	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 19 ligands modelled in this entry, 6 are monoatomic - leaving 13 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	SO4	A	231	-	4,4,4	0.20	0	6,6,6	0.27	0
3	SO4	A	232	-	4,4,4	0.17	0	6,6,6	0.43	0
3	SO4	B	231	-	4,4,4	0.14	0	6,6,6	0.26	0
3	SO4	B	232	-	4,4,4	0.17	0	6,6,6	0.26	0
3	SO4	C	231	-	4,4,4	0.24	0	6,6,6	0.39	0
3	SO4	C	232	-	4,4,4	0.15	0	6,6,6	0.31	0
3	SO4	D	231	-	4,4,4	0.15	0	6,6,6	0.34	0
3	SO4	D	232	-	4,4,4	0.20	0	6,6,6	0.38	0
3	SO4	E	231	-	4,4,4	0.11	0	6,6,6	0.52	0
3	SO4	E	232	-	4,4,4	0.15	0	6,6,6	0.47	0
3	SO4	E	233	-	4,4,4	0.18	0	6,6,6	0.48	0
3	SO4	F	231	-	4,4,4	0.35	0	6,6,6	0.38	0
4	BCT	F	232	-	0,3,3	0.00	-	0,3,3	0.00	-

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
3	SO4	A	231	-	-	0/0/0/0	0/0/0/0
3	SO4	A	232	-	-	0/0/0/0	0/0/0/0
3	SO4	B	231	-	-	0/0/0/0	0/0/0/0
3	SO4	B	232	-	-	0/0/0/0	0/0/0/0
3	SO4	C	231	-	-	0/0/0/0	0/0/0/0
3	SO4	C	232	-	-	0/0/0/0	0/0/0/0
3	SO4	D	231	-	-	0/0/0/0	0/0/0/0
3	SO4	D	232	-	-	0/0/0/0	0/0/0/0
3	SO4	E	231	-	-	0/0/0/0	0/0/0/0
3	SO4	E	232	-	-	0/0/0/0	0/0/0/0
3	SO4	E	233	-	-	0/0/0/0	0/0/0/0
3	SO4	F	231	-	-	0/0/0/0	0/0/0/0
4	BCT	F	232	-	-	0/0/0/0	0/0/0/0

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.

7 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	232	SO4	3	0
3	B	232	SO4	2	0
3	C	232	SO4	3	0
3	D	232	SO4	3	0
3	E	233	SO4	3	0
3	F	231	SO4	2	0
4	F	232	BCT	2	0

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<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	209/229 (91%)	0.70	25 (11%) 5 7	23, 42, 54, 63	0
1	B	207/229 (90%)	1.11	39 (18%) 1 2	38, 45, 60, 63	0
1	C	205/229 (89%)	0.80	34 (16%) 2 3	36, 42, 55, 65	0
1	D	204/229 (89%)	0.61	17 (8%) 12 16	33, 40, 51, 59	0
1	E	207/229 (90%)	0.90	35 (16%) 2 2	37, 43, 55, 62	0
1	F	204/229 (89%)	0.64	26 (12%) 4 6	33, 41, 52, 65	0
All	All	1236/1374 (89%)	0.79	176 (14%) 3 4	23, 42, 55, 65	0

All (176) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	16	GLN	6.6
1	E	220	ILE	6.4
1	D	215	LEU	6.0
1	B	12	TYR	5.6
1	A	220	ILE	5.4
1	A	32	HIS	5.4
1	C	183	VAL	5.1
1	B	184	ASN	5.1
1	B	217	GLU	5.0
1	E	183	VAL	4.9
1	B	54	ASN	4.8
1	C	217	GLU	4.8
1	B	183	VAL	4.7
1	C	184	ASN	4.7
1	B	187	PHE	4.7
1	B	221	LEU	4.6
1	B	2	ASP	4.6
1	B	1	MET	4.5
1	B	219	ASN	4.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	40	ILE	4.4
1	E	219	ASN	4.4
1	B	220	ILE	4.3
1	C	1	MET	4.3
1	C	5	LYS	4.2
1	E	16	GLN	4.2
1	E	13	SER	4.2
1	F	40	ILE	4.2
1	A	1	MET	4.0
1	B	218	GLU	4.0
1	D	2	ASP	4.0
1	F	95	ILE	4.0
1	A	221	LEU	3.9
1	C	17	ARG	3.9
1	C	187	PHE	3.9
1	A	132	HIS	3.8
1	E	218	GLU	3.8
1	D	1	MET	3.7
1	E	1	MET	3.7
1	A	110	ASP	3.7
1	E	187	PHE	3.6
1	D	16	GLN	3.6
1	E	2	ASP	3.6
1	F	218	GLU	3.6
1	B	136	LYS	3.6
1	C	2	ASP	3.5
1	D	132	HIS	3.5
1	C	12	TYR	3.5
1	E	217	GLU	3.5
1	E	40	ILE	3.5
1	B	132	HIS	3.5
1	E	93	ILE	3.4
1	B	56	GLU	3.4
1	F	93	ILE	3.4
1	E	132	HIS	3.4
1	D	31	ASP	3.4
1	E	32	HIS	3.4
1	F	1	MET	3.4
1	A	40	ILE	3.4
1	C	16	GLN	3.3
1	B	9	ALA	3.3
1	B	17	ARG	3.3

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	38	LEU	3.2
1	E	136	LYS	3.2
1	D	12	TYR	3.2
1	F	94	ILE	3.1
1	C	40	ILE	3.1
1	F	110	ASP	3.1
1	E	34	THR	3.1
1	B	95	ILE	3.1
1	E	12	TYR	3.0
1	F	62	VAL	3.0
1	F	105	HIS	3.0
1	E	56	GLU	3.0
1	E	17	ARG	3.0
1	A	216	ASP	3.0
1	A	105	HIS	3.0
1	B	158	LEU	2.9
1	B	216	ASP	2.9
1	A	80	VAL	2.9
1	E	184	ASN	2.9
1	A	17	ARG	2.9
1	C	139	PRO	2.8
1	F	39	TRP	2.8
1	A	56	GLU	2.8
1	F	140	GLU	2.8
1	A	62	VAL	2.8
1	E	9	ALA	2.8
1	B	5	LYS	2.8
1	C	219	ASN	2.8
1	D	184	ASN	2.8
1	E	140	GLU	2.8
1	E	6	GLN	2.8
1	C	132	HIS	2.8
1	E	215	LEU	2.7
1	A	184	ASN	2.7
1	A	218	GLU	2.7
1	F	158	LEU	2.7
1	E	110	ASP	2.7
1	B	206	ARG	2.7
1	F	219	ASN	2.7
1	C	38	LEU	2.7
1	D	45	SER	2.6
1	F	216	ASP	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	93	ILE	2.6
1	E	95	ILE	2.6
1	B	110	ASP	2.6
1	C	216	ASP	2.6
1	E	216	ASP	2.6
1	F	80	VAL	2.6
1	B	62	VAL	2.6
1	B	33	GLN	2.6
1	C	93	ILE	2.6
1	C	15	ALA	2.6
1	C	13	SER	2.5
1	E	62	VAL	2.5
1	B	15	ALA	2.5
1	D	135	GLY	2.5
1	C	77	CYS	2.5
1	C	218	GLU	2.5
1	D	40	ILE	2.5
1	F	96	CYS	2.5
1	C	81	VAL	2.5
1	B	93	ILE	2.5
1	C	19	LYS	2.4
1	F	17	ARG	2.4
1	C	158	LEU	2.4
1	F	217	GLU	2.4
1	B	34	THR	2.4
1	A	94	ILE	2.4
1	E	81	VAL	2.4
1	B	141	LYS	2.4
1	B	6	GLN	2.3
1	E	70	VAL	2.3
1	C	176	LEU	2.3
1	C	215	LEU	2.3
1	A	183	VAL	2.3
1	A	139	PRO	2.3
1	D	111	LYS	2.3
1	C	56	GLU	2.3
1	A	219	ASN	2.3
1	A	81	VAL	2.2
1	F	41	GLY	2.2
1	B	70	VAL	2.2
1	D	140	GLU	2.2
1	E	139	PRO	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	94	ILE	2.2
1	E	38	LEU	2.2
1	B	41	GLY	2.2
1	B	77	CYS	2.2
1	A	100	ASN	2.1
1	F	81	VAL	2.1
1	F	187	PHE	2.1
1	D	38	LEU	2.1
1	C	206	ARG	2.1
1	B	18	MET	2.1
1	A	70	VAL	2.1
1	E	151	VAL	2.1
1	D	100	ASN	2.1
1	F	132	HIS	2.1
1	C	34	THR	2.1
1	E	80	VAL	2.1
1	C	70	VAL	2.1
1	D	70	VAL	2.1
1	F	151	VAL	2.1
1	A	169	GLU	2.0
1	B	185	ASP	2.0
1	C	112	ASP	2.0
1	C	14	TRP	2.0
1	D	5	LYS	2.0
1	B	215	LEU	2.0
1	B	65	ASN	2.0
1	F	138	SER	2.0
1	C	80	VAL	2.0
1	E	158	LEU	2.0
1	F	18	MET	2.0
1	A	136	LYS	2.0

## 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 carbohydrates 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. 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, 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	SO4	F	231	5/5	0.83	0.49	7.15	103,103,104,104	0
3	SO4	A	232	5/5	0.95	0.30	2.64	54,56,58,58	0
3	SO4	C	232	5/5	0.93	0.35	1.86	83,83,83,84	0
3	SO4	B	232	5/5	0.80	0.29	1.24	81,81,82,82	0
3	SO4	E	233	5/5	0.93	0.24	0.81	76,76,78,78	0
4	BCT	F	232	4/4	0.95	0.20	-0.13	42,42,43,43	0
3	SO4	E	232	5/5	0.95	0.14	-0.77	60,62,63,63	0
3	SO4	C	231	5/5	0.94	0.13	-0.93	64,64,66,66	0
3	SO4	D	231	5/5	0.98	0.14	-0.94	56,58,59,59	0
2	ZN	C	230	1/1	1.00	0.11	-1.20	42,42,42,42	0
2	ZN	B	230	1/1	0.99	0.10	-1.30	42,42,42,42	0
3	SO4	D	232	5/5	0.97	0.10	-1.38	45,45,48,49	0
2	ZN	E	230	1/1	0.99	0.07	-1.40	41,41,41,41	0
2	ZN	F	230	1/1	0.99	0.10	-1.48	44,44,44,44	0
2	ZN	D	230	1/1	0.98	0.08	-1.73	41,41,41,41	0
3	SO4	B	231	5/5	0.97	0.11	-1.77	53,55,57,58	0
2	ZN	A	230	1/1	0.95	0.07	-2.23	48,48,48,48	0
3	SO4	A	231	5/5	0.96	0.11	-3.20	60,63,63,63	0
3	SO4	E	231	5/5	0.98	0.10	-4.10	56,58,58,59	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.