



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 08:48 AM GMT

PDB ID : 4KP5  
Title : Crystal structure of catalytic domain of human carbonic anhydrase isozyme XII with 2-Chloro-4-[(pyrimidin-2-ylsulfanyl)acetyl]benzenesulfonamide  
Authors : Smirnov, A.; Manakova, E.; Grazulis, S.  
Deposited on : 2013-05-13  
Resolution : 1.45 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

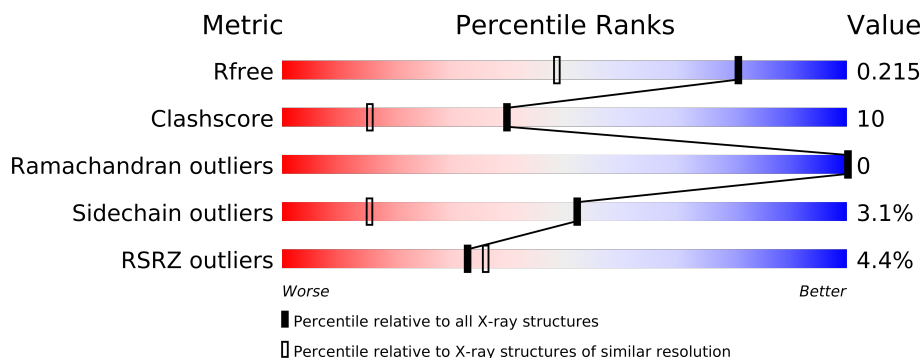
The following versions of software and data (see [references](#)) were used in the production of this report:

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 1.45 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	2476 (1.50-1.42)
Clashscore	79885	2834 (1.50-1.42)
Ramachandran outliers	78287	2769 (1.50-1.42)
Sidechain outliers	78261	2767 (1.50-1.42)
RSRZ outliers	66119	2477 (1.50-1.42)

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. 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	263	
1	B	263	
1	C	263	
1	D	263	

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

Mol	Type	Chain	Res	Geometry	Electron density
3	E1F	A	302	-	X
3	E1F	B	302	-	X
3	E1F	C	302	-	X
3	E1F	D	302	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
4	EDO	A	303	-	X
4	EDO	B	303	-	X
4	EDO	B	304	-	X
4	EDO	D	303	-	X
4	EDO	D	304	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10072 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 Carbonic anhydrase 12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	261	Total	C	N	O	S	0	10	0
			2187	1386	374	420	7			
1	B	261	Total	C	N	O	S	0	7	0
			2154	1365	368	413	8			
1	C	260	Total	C	N	O	S	0	16	0
			2223	1414	380	419	10			
1	D	260	Total	C	N	O	S	0	12	0
			2185	1383	370	424	8			

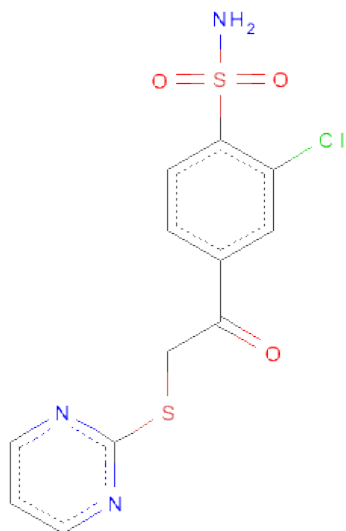
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	EXPRESSION TAG	UNP O43570
B	1	MET	-	EXPRESSION TAG	UNP O43570
C	1	MET	-	EXPRESSION TAG	UNP O43570
D	1	MET	-	EXPRESSION TAG	UNP O43570

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

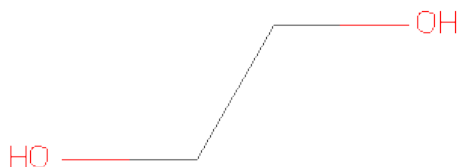
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Zn	0	0
			1	1		
2	A	1	Total	Zn	0	0
			1	1		
2	D	1	Total	Zn	0	0
			1	1		
2	C	1	Total	Zn	0	0
			1	1		

- Molecule 3 is 2-CHLORO-4-[(PYRIMIDIN-2-YLSULFANYL)ACETYL]BENZENESULFO  
NAMIDE (three-letter code: E1F) (formula: C<sub>12</sub>H<sub>10</sub>ClN<sub>3</sub>O<sub>3</sub>S<sub>2</sub>).



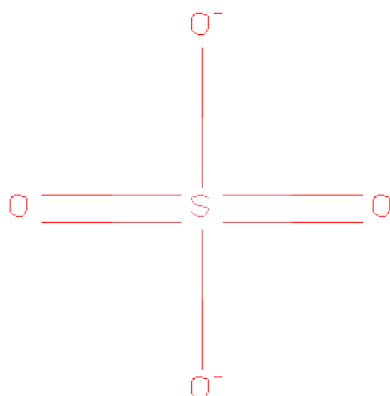
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	Cl	N	O	S	0	0
			21	12	1	3	3	2		
3	B	1	Total	C	Cl	N	O	S	0	0
			21	12	1	3	3	2		
3	C	1	Total	C	Cl	N	O	S	0	0
			21	12	1	3	3	2		
3	D	1	Total	C	Cl	N	O	S	0	0
			21	12	1	3	3	2		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0
4	D	1	Total C O 4 2 2	0	0
4	D	1	Total C O 4 2 2	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	1	Total O S 5 4 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	328	Total O 328 328	0	0
6	B	331	Total O 331 331	0	0
6	C	226	Total O 226 226	0	0

*Continued on next page...*

*Continued from previous page...*

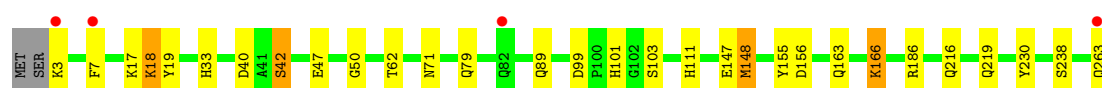
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	321	Total 321	O 321	0	0

### 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.

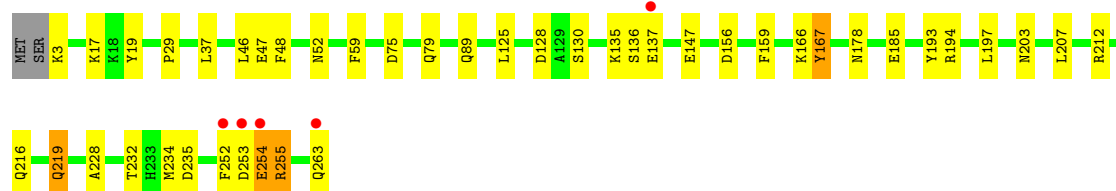
- Molecule 1: Carbonic anhydrase 12

Chain A: 



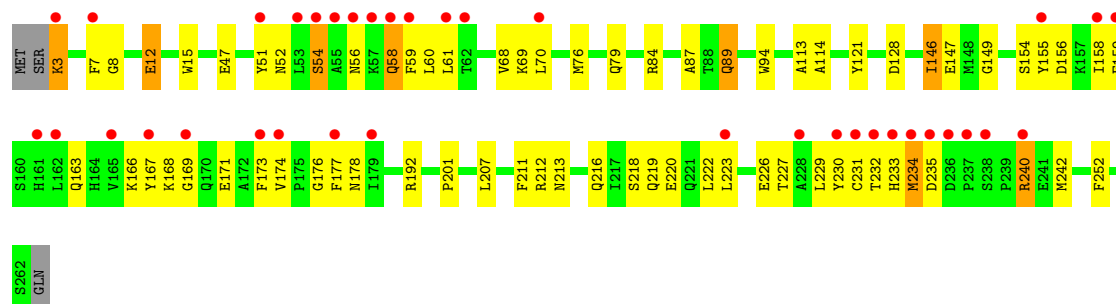
- Molecule 1: Carbonic anhydrase 12

Chain B: 



- Molecule 1: Carbonic anhydrase 12

Chain C: 



- Molecule 1: Carbonic anhydrase 12

Chain D: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.30Å 74.16Å 91.39Å 90.00° 108.81° 90.00°	Depositor
Resolution (Å)	73.13 – 1.45 73.17 – 1.45	Depositor EDS
% Data completeness (in resolution range)	97.3 (73.13-1.45) 97.3 (73.17-1.45)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.02 (at 1.45Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.177 , 0.216 0.176 , 0.215	Depositor DCC
$R_{free}$ test set	16774 reflections (11.08%)	DCC
Wilson B-factor (Å <sup>2</sup> )	12.1	Xtriage
Anisotropy	0.303	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 34.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	6 of 168151 reflections (0.004%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10072	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 52.62 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 4.7385e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities 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: ZN, EDO, SO4, E1F

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	1.42	2/2250 (0.1%)	1.22	2/3057 (0.1%)
1	B	1.45	8/2218 (0.4%)	1.32	6/3019 (0.2%)
1	C	1.44	6/2287 (0.3%)	1.26	4/3107 (0.1%)
1	D	1.47	10/2248 (0.4%)	1.31	8/3058 (0.3%)
All	All	1.44	26/9003 (0.3%)	1.28	20/12241 (0.2%)

All (26) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	59	PHE	CG-CD2	5.79	1.47	1.38
1	A	230	TYR	CD2-CE2	5.76	1.48	1.39
1	B	167	TYR	CG-CD1	-5.73	1.31	1.39
1	C	146	ILE	N-CA	5.73	1.57	1.46
1	A	50	GLY	N-CA	-5.73	1.37	1.46
1	D	252	PHE	CD2-CE2	5.66	1.50	1.39
1	D	39	TYR	CE2-CZ	-5.64	1.31	1.38
1	D	166[A]	LYS	CD-CE	5.63	1.65	1.51
1	D	166[B]	LYS	CD-CE	5.63	1.65	1.51
1	D	39	TYR	CE1-CZ	5.58	1.45	1.38
1	D	47	GLU	CB-CG	5.43	1.62	1.52
1	D	258	TYR	CD2-CE2	-5.39	1.31	1.39
1	C	121	TYR	CD1-CE1	5.38	1.47	1.39
1	B	37	LEU	N-CA	5.37	1.57	1.46
1	C	59	PHE	CE1-CZ	5.31	1.47	1.37
1	C	12	GLU	CD-OE2	5.28	1.31	1.25
1	D	155	TYR	CG-CD1	5.26	1.46	1.39
1	C	211	PHE	CD2-CE2	5.26	1.49	1.39
1	B	228	ALA	CA-CB	-5.22	1.41	1.52
1	D	94	TRP	CD2-CE2	5.16	1.47	1.41
1	B	159	PHE	CG-CD2	5.10	1.46	1.38
1	B	255	ARG	CA-C	5.09	1.66	1.52

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	155	TYR	CB-CG	-5.04	1.44	1.51
1	B	3	LYS	N-CA	5.04	1.56	1.46
1	B	48	PHE	CD1-CE1	5.03	1.49	1.39
1	B	159	PHE	CD1-CE1	5.01	1.49	1.39

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	166[A]	LYS	CD-CE-NZ	5.89	125.25	111.70
1	D	166[B]	LYS	CD-CE-NZ	5.89	125.25	111.70
1	D	192	ARG	CD-NE-CZ	5.86	131.80	123.60
1	D	84	ARG	CB-CG-CD	5.76	126.58	111.60
1	D	84	ARG	NE-CZ-NH1	5.67	123.14	120.30
1	C	12	GLU	OE1-CD-OE2	-5.64	116.53	123.30
1	D	84	ARG	NE-CZ-NH2	-5.60	117.50	120.30
1	C	84	ARG	NE-CZ-NH1	-5.54	117.53	120.30
1	B	19	TYR	CB-CG-CD1	-5.47	117.72	121.00
1	B	235	ASP	CB-CG-OD1	5.37	123.14	118.30
1	B	193	TYR	CB-CG-CD1	-5.36	117.78	121.00
1	B	46	LEU	CB-CG-CD1	-5.29	102.02	111.00
1	A	147	GLU	OE1-CD-OE2	-5.28	116.97	123.30
1	C	252	PHE	CB-CG-CD2	-5.23	117.14	120.80
1	B	59	PHE	CB-CG-CD2	-5.23	117.14	120.80
1	B	212	ARG	NE-CZ-NH2	-5.14	117.73	120.30
1	D	39	TYR	CZ-CE2-CD2	5.13	124.42	119.80
1	C	113	ALA	N-CA-C	5.08	124.73	111.00
1	A	155	TYR	CD1-CE1-CZ	5.05	124.35	119.80
1	D	206	VAL	O-C-N	5.01	130.72	122.70

There are no chirality outliers.

There are no planarity outliers.

## 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	2187	0	2073	40	0
1	B	2154	0	2033	38	0
1	C	2223	0	2118	83	0
1	D	2185	0	2055	12	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
3	A	21	0	9	0	0
3	B	21	0	9	0	0
3	C	21	0	9	2	0
3	D	21	0	9	0	0
4	A	8	0	12	0	0
4	B	8	0	12	0	0
4	D	8	0	12	3	0
5	C	5	0	0	2	0
6	A	328	0	0	12	0
6	B	331	0	0	4	0
6	C	226	0	0	8	0
6	D	321	0	0	4	0
All	All	10072	0	8351	165	0

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 10.

All (165) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:166[B]:LYS:HE2	6:D:700:HOH:O	1.30	1.30
1:D:75:ASP:HB3	6:D:669:HOH:O	1.43	1.17
1:A:3[B]:LYS:HE2	1:A:18[B]:LYS:HZ3	1.09	1.13
1:A:71[B]:ASN:OD1	6:A:711:HOH:O	1.63	1.11
1:A:3[B]:LYS:HE2	1:A:18[B]:LYS:NZ	1.65	1.10
1:B:253:ASP:HA	6:B:583:HOH:O	1.51	1.07
1:B:137:GLU:HB3	1:B:194:ARG:HH12	0.94	1.07
1:C:168[A]:LYS:HB2	1:C:231[A]:CYS:O	1.56	1.04
1:B:136:SER:O	1:B:137:GLU:HG2	1.60	1.00
1:A:7:PHE:CE1	1:B:253:ASP:HB3	1.97	0.99
1:B:185:GLU:HG2	6:B:721:HOH:O	1.61	0.98
1:B:137:GLU:HB3	1:B:194:ARG:NH1	1.78	0.98
1:C:94:TRP:CH2	1:C:242[B]:MET:HE3	2.01	0.95
1:B:252:PHE:CZ	1:B:255:ARG:HB2	2.02	0.95
1:A:3[A]:LYS:HD2	6:A:694:HOH:O	1.67	0.93
1:B:137:GLU:CB	1:B:194:ARG:HH12	1.80	0.93

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:51:TYR:HA	1:C:76[B]:MET:HE3	1.48	0.93
1:D:163:GLN:OE1	1:D:166[A]:LYS:HD2	1.70	0.92
1:C:94:TRP:CZ2	1:C:242[B]:MET:HE3	2.04	0.92
1:B:252:PHE:CE1	1:B:255:ARG:HB2	2.05	0.91
1:C:58:GLN:OE1	1:C:58:GLN:N	2.07	0.88
1:B:252:PHE:CZ	1:B:255:ARG:CB	2.59	0.85
1:A:7:PHE:HE1	1:B:253:ASP:HB3	1.33	0.85
1:D:47:GLU:HB2	1:D:79[B]:GLN:HG3	1.60	0.83
1:C:177:PHE:HA	6:C:599:HOH:O	1.79	0.83
1:A:163:GLN:HG3	6:A:675:HOH:O	1.79	0.82
1:C:7:PHE:HD1	1:C:8:GLY:N	1.78	0.81
1:C:94:TRP:CZ2	1:C:242[B]:MET:CE	2.64	0.81
1:C:229:LEU:HD12	1:C:242[B]:MET:CE	2.11	0.80
1:C:51:TYR:HA	1:C:76[B]:MET:CE	2.11	0.80
1:A:40:ASP:OD1	1:A:42:SER:HB3	1.83	0.79
1:A:3[B]:LYS:O	1:A:3[B]:LYS:HD3	1.83	0.78
1:A:148:MET:SD	1:A:219:GLN:CA	2.71	0.78
1:C:7:PHE:CD1	1:C:8:GLY:N	2.53	0.77
1:C:231[B]:CYS:SG	1:C:240:ARG:HD2	2.24	0.76
1:B:166[A]:LYS:HE2	6:D:700:HOH:O	1.85	0.76
1:C:94:TRP:CH2	1:C:242[B]:MET:CE	2.70	0.74
1:A:148:MET:SD	1:A:219:GLN:N	2.61	0.73
1:D:47:GLU:HB3	1:D:79[B]:GLN:HE21	1.54	0.72
1:C:229:LEU:HD12	1:C:242[B]:MET:HE1	1.70	0.72
1:C:218:SER:HB2	5:C:303:SO4:O1	1.90	0.71
1:C:192:ARG:HD2	1:C:207[B]:LEU:HD21	1.74	0.70
1:A:148:MET:SD	1:A:219:GLN:HA	2.32	0.69
1:C:219:GLN:HE21	1:C:223:LEU:HD22	1.56	0.69
1:C:232:THR:HG23	1:C:240:ARG:HG3	1.73	0.69
1:A:3[B]:LYS:C	1:A:3[B]:LYS:HD3	2.14	0.68
1:A:163:GLN:CG	6:A:675:HOH:O	2.39	0.67
1:A:7:PHE:CE1	1:B:253:ASP:CB	2.76	0.66
1:B:252:PHE:CE1	1:B:255:ARG:CB	2.76	0.65
1:A:18[B]:LYS:HD2	1:A:19:TYR:CZ	2.32	0.65
1:C:167:TYR:CD2	1:C:234:MET:HG2	2.32	0.65
1:C:229:LEU:HD12	1:C:242[B]:MET:HE2	1.79	0.64
1:C:219:GLN:NE2	1:C:223:LEU:HD22	2.12	0.64
1:B:252:PHE:CZ	1:B:255:ARG:HB3	2.33	0.64
1:C:192:ARG:CD	1:C:207[B]:LEU:HD21	2.29	0.63
1:D:163:GLN:HB2	6:D:688:HOH:O	1.99	0.63
1:A:166[B]:LYS:HE2	6:A:555:HOH:O	2.00	0.61
1:A:163:GLN:O	1:A:166[A]:LYS:HG3	2.01	0.61

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:7:PHE:CD1	1:B:253:ASP:HB3	2.36	0.60
1:A:216:GLN:HG3	6:A:593:HOH:O	2.02	0.59
1:C:69[A]:LYS:HD2	6:C:473:HOH:O	2.02	0.59
1:A:7:PHE:HE1	1:B:253:ASP:CB	2.09	0.59
1:B:136:SER:C	1:B:137:GLU:HG2	2.22	0.59
1:A:47:GLU:CD	1:A:79:GLN:HE21	2.06	0.59
1:A:148:MET:SD	1:A:219:GLN:HB3	2.43	0.59
1:C:69[B]:LYS:HE2	6:C:473:HOH:O	2.03	0.58
1:C:212[A]:ARG:HH11	1:C:213:ASN:HD21	1.50	0.58
1:C:61[B]:LEU:CD2	1:C:70:LEU:HD13	2.34	0.58
1:B:252:PHE:CE2	1:B:255:ARG:HB2	2.37	0.57
1:C:229:LEU:HB2	1:C:242[A]:MET:HB3	1.87	0.57
1:C:61[B]:LEU:HG	1:C:174:VAL:HG23	1.86	0.57
1:A:3[B]:LYS:HE2	1:A:18[B]:LYS:HZ1	1.62	0.57
1:D:47:GLU:CB	1:D:79[B]:GLN:HE21	2.17	0.57
1:C:219:GLN:HB3	5:C:303:SO4:O2	2.06	0.55
1:C:69[B]:LYS:NZ	6:C:582:HOH:O	2.39	0.55
1:C:54:SER:HB2	1:C:56:ASN:HB2	1.89	0.55
1:C:229:LEU:HB2	1:C:242[A]:MET:CB	2.37	0.55
1:C:8:GLY:HA2	1:D:25:LEU:HD21	1.90	0.54
1:B:128:ASP:OD2	1:B:130[A]:SER:HB2	2.07	0.54
1:C:7:PHE:CD1	1:C:7:PHE:C	2.81	0.54
1:C:167:TYR:CG	1:C:234:MET:HA	2.43	0.54
1:C:94:TRP:HH2	1:C:242[B]:MET:CE	2.21	0.54
1:B:135:LYS:NZ	6:B:698:HOH:O	2.34	0.54
1:C:7:PHE:HE2	1:C:240:ARG:HH11	1.57	0.53
1:B:252:PHE:CE1	1:B:255:ARG:NE	2.78	0.52
1:A:18[A]:LYS:HG2	6:A:642:HOH:O	2.10	0.52
1:C:51:TYR:CD2	1:C:76[B]:MET:HE3	2.44	0.52
1:B:167:TYR:CE2	1:B:234[B]:MET:HG3	2.44	0.52
1:C:232:THR:CG2	1:C:240:ARG:HG3	2.40	0.52
1:C:94:TRP:HZ2	1:C:242[B]:MET:CE	2.18	0.51
1:C:230:TYR:HA	1:C:240:ARG:O	2.11	0.51
1:B:47:GLU:HB2	1:B:79:GLN:HB3	1.93	0.51
1:C:230:TYR:CD2	1:C:230:TYR:N	2.78	0.51
1:C:47:GLU:HB2	1:C:79:GLN:HB3	1.92	0.51
1:A:148:MET:SD	1:A:219:GLN:CB	2.99	0.51
1:C:168[B]:LYS:HD3	1:C:169:GLY:N	2.26	0.51
1:C:158[B]:ILE:N	1:C:158[B]:ILE:HD12	2.26	0.51
1:C:94:TRP:HZ2	1:C:242[B]:MET:HE2	1.74	0.50
1:C:94:TRP:CZ2	1:C:242[A]:MET:HG2	2.46	0.50
1:A:18[B]:LYS:HD2	1:A:19:TYR:CE2	2.47	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:252:PHE:CE2	1:B:255:ARG:C	2.85	0.50
1:C:154:SER:N	6:C:572:HOH:O	2.28	0.50
1:C:207[B]:LEU:C	1:C:207[B]:LEU:HD23	2.32	0.49
1:C:212[A]:ARG:NH1	1:C:213:ASN:HD21	2.09	0.49
1:C:76[A]:MET:HB3	1:C:87:ALA:HB3	1.95	0.49
1:B:252:PHE:CD1	1:B:255:ARG:HB2	2.45	0.49
1:A:7:PHE:CD1	1:B:253:ASP:CB	2.96	0.48
1:A:79:GLN:OE1	6:A:707:HOH:O	2.20	0.48
1:C:54:SER:O	1:C:176:GLY:HA3	2.13	0.48
1:C:3[B]:LYS:O	1:C:3[B]:LYS:CG	2.61	0.48
1:C:58:GLN:HG2	1:C:173:PHE:CD1	2.48	0.48
1:C:154:SER:OG	6:C:593:HOH:O	2.20	0.48
1:B:252:PHE:CE2	1:B:255:ARG:O	2.66	0.47
1:C:114:ALA:HB3	1:C:146:ILE:HB	1.95	0.47
1:C:149:GLY:HA3	1:C:216:GLN:OE1	2.14	0.47
1:C:229:LEU:C	1:C:230:TYR:CD2	2.88	0.47
1:C:233[A]:HIS:CG	1:C:234:MET:N	2.78	0.47
1:C:3[B]:LYS:HG3	1:C:3[B]:LYS:O	2.14	0.47
1:A:18[B]:LYS:O	1:A:18[B]:LYS:HD3	2.16	0.46
1:C:61[B]:LEU:HD23	1:C:70:LEU:HD13	1.97	0.46
1:A:3[B]:LYS:HD2	1:A:19:TYR:OH	2.16	0.46
1:D:43:LEU:HA	4:D:304:EDO:H12	1.97	0.46
1:C:220:GLU:HG3	1:C:220:GLU:O	2.15	0.46
1:A:17[B]:LYS:HG2	1:B:17:LYS:HG2	1.96	0.45
1:A:186:ARG:HH11	1:A:186:ARG:HG3	1.80	0.45
1:C:94:TRP:CZ2	1:C:242[B]:MET:HE2	2.47	0.45
1:A:99:ASP:OD1	1:A:101:HIS:HD2	1.99	0.45
1:D:192:ARG:HD2	1:D:207:LEU:HD11	1.99	0.45
1:C:61[A]:LEU:HD11	1:C:68:VAL:HG22	1.99	0.44
1:C:201:PRO:CD	3:C:302:E1F:H1	2.47	0.44
1:D:253:ASP:O	1:D:254:GLU:HB2	2.17	0.44
1:C:60:LEU:HD11	1:C:171:GLU:HB3	2.00	0.44
1:C:219:GLN:HG2	6:C:589:HOH:O	2.16	0.43
1:C:52:ASN:HA	1:C:178:ASN:HA	2.00	0.43
1:C:155:TYR:HB3	1:C:159:PHE:CE2	2.52	0.43
1:A:103:SER:O	1:A:111:HIS:HD2	2.02	0.43
1:C:147:GLU:O	1:C:216:GLN:HA	2.18	0.43
1:A:163:GLN:HG3	6:A:608:HOH:O	2.18	0.43
1:A:163:GLN:OE1	6:A:582:HOH:O	2.20	0.43
1:C:163:GLN:OE1	1:C:166:LYS:HD2	2.17	0.43
1:B:219:GLN:HB3	1:B:219:GLN:HE21	1.68	0.43
1:B:147:GLU:HG3	1:B:216:GLN:HG2	2.01	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:252:PHE:O	1:B:254:GLU:N	2.51	0.42
1:C:192:ARG:CD	1:C:207[B]:LEU:CD2	2.97	0.42
1:B:52:ASN:HA	1:B:178:ASN:HA	2.00	0.42
1:C:7:PHE:HD1	1:C:8:GLY:CA	2.32	0.42
1:A:148:MET:HE3	1:A:148:MET:C	2.40	0.42
1:B:207:LEU:HD12	1:B:207:LEU:HA	1.93	0.42
1:D:80:GLY:O	4:D:304:EDO:H21	2.19	0.42
1:C:69[A]:LYS:HG3	1:C:89:GLN:HB2	2.02	0.42
1:B:253:ASP:O	1:B:254:GLU:HB2	2.18	0.41
1:C:12:GLU:HA	1:C:15:TRP:CE2	2.55	0.41
1:C:149:GLY:O	1:C:218:SER:HA	2.20	0.41
1:C:61[B]:LEU:HD23	1:C:70:LEU:CD1	2.51	0.41
1:B:197:LEU:HD13	1:B:203[A]:ASN:HB2	2.02	0.41
1:C:201:PRO:HD2	3:C:302:E1F:H1	2.02	0.40
1:A:62[A]:THR:HG21	6:A:600:HOH:O	2.20	0.40
1:C:3[A]:LYS:O	6:C:551:HOH:O	2.22	0.40
1:B:166[B]:LYS:CB	6:B:487:HOH:O	2.69	0.40
1:C:51:TYR:CG	1:C:76[B]:MET:CE	3.04	0.40
1:C:223:LEU:O	1:C:227:THR:OG1	2.32	0.40
1:C:222:LEU:O	1:C:226:GLU:HG3	2.21	0.40
1:A:33:HIS:HE1	6:A:607:HOH:O	2.04	0.40
1:D:44:THR:O	4:D:304:EDO:H22	2.21	0.40

There are no symmetry-related clashes.

## 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	269/263 (102%)	266 (99%)	3 (1%)	0	100	100
1	B	266/263 (101%)	258 (97%)	8 (3%)	0	100	100
1	C	273/263 (104%)	256 (94%)	17 (6%)	0	100	100
1	D	270/263 (103%)	264 (98%)	6 (2%)	0	100	100
All	All	1078/1052 (102%)	1044 (97%)	34 (3%)	0	100	100



There are no Ramachandran outliers to report.

### 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	244/235 (104%)	234 (96%)	10 (4%)	41 8
1	B	240/235 (102%)	231 (96%)	9 (4%)	44 10
1	C	248/235 (106%)	238 (96%)	10 (4%)	42 9
1	D	244/235 (104%)	241 (99%)	3 (1%)	82 56
All	All	976/940 (104%)	944 (97%)	32 (3%)	52 14

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18[A]	LYS
1	A	18[B]	LYS
1	A	42	SER
1	A	89	GLN
1	A	148	MET
1	A	156	ASP
1	A	166[A]	LYS
1	A	166[B]	LYS
1	A	238	SER
1	A	263	GLN
1	B	29	PRO
1	B	75	ASP
1	B	89	GLN
1	B	125	LEU
1	B	156	ASP
1	B	219	GLN
1	B	232	THR
1	B	254	GLU
1	B	263	GLN
1	C	3[A]	LYS
1	C	3[B]	LYS
1	C	54	SER
1	C	58	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	89	GLN
1	C	128	ASP
1	C	156	ASP
1	C	234	MET
1	C	235	ASP
1	C	240	ARG
1	D	75	ASP
1	D	89	GLN
1	D	156	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	33	HIS
1	A	58	GLN
1	A	79	GLN
1	A	110	GLN
1	A	163	GLN
1	A	263	GLN
1	B	219	GLN
1	C	49	GLN
1	C	213	ASN
1	C	219	GLN

### 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 15 ligands modelled in this entry, 4 are monoatomic - leaving 11 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	E1F	A	302	2	22,22,22	2.55	8 (36%)	31,31,31	2.63	14 (45%)
4	EDO	A	303	-	3,3,3	1.02	0	2,2,2	0.32	0
4	EDO	A	304	-	3,3,3	0.93	0	2,2,2	0.75	0
3	E1F	B	302	2	22,22,22	2.79	10 (45%)	31,31,31	2.42	12 (38%)
4	EDO	B	303	-	3,3,3	1.46	0	2,2,2	0.88	0
4	EDO	B	304	-	3,3,3	1.18	0	2,2,2	0.37	0
3	E1F	C	302	2	22,22,22	2.11	5 (22%)	31,31,31	2.18	11 (35%)
5	SO4	C	303	-	4,4,4	0.11	0	6,6,6	0.35	0
3	E1F	D	302	2	22,22,22	2.01	4 (18%)	31,31,31	2.44	18 (58%)
4	EDO	D	303	-	3,3,3	0.95	0	2,2,2	1.04	0
4	EDO	D	304	-	3,3,3	0.75	0	2,2,2	0.55	0

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	E1F	A	302	2	-	0/15/15/15	0/2/2/2
4	EDO	A	303	-	-	0/1/1/1	0/0/0/0
4	EDO	A	304	-	-	0/1/1/1	0/0/0/0
3	E1F	B	302	2	-	0/15/15/15	0/2/2/2
4	EDO	B	303	-	-	0/1/1/1	0/0/0/0
4	EDO	B	304	-	-	0/1/1/1	0/0/0/0
3	E1F	C	302	2	-	0/15/15/15	0/2/2/2
5	SO4	C	303	-	-	0/0/0/0	0/0/0/0
3	E1F	D	302	2	-	0/15/15/15	0/2/2/2
4	EDO	D	303	-	-	0/1/1/1	0/0/0/0
4	EDO	D	304	-	-	0/1/1/1	0/0/0/0

All (27) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	302	E1F	C3-S16	7.51	1.88	1.77

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	302	E1F	O17-S16	6.97	1.55	1.43
3	D	302	E1F	C3-S16	6.73	1.87	1.77
3	A	302	E1F	O17-S16	6.08	1.54	1.43
3	A	302	E1F	O18-S16	5.27	1.52	1.43
3	B	302	E1F	C9-S8	5.02	1.82	1.75
3	C	302	E1F	C3-S16	4.51	1.84	1.77
3	B	302	E1F	C4-C3	3.99	1.43	1.39
3	B	302	E1F	C7-S8	3.99	1.89	1.80
3	B	302	E1F	C9-N10	3.69	1.41	1.33
3	D	302	E1F	O18-S16	3.61	1.49	1.43
3	A	302	E1F	C21-C2	3.48	1.44	1.38
3	A	302	E1F	C7-S8	3.47	1.87	1.80
3	A	302	E1F	C4-C5	3.42	1.45	1.38
3	B	302	E1F	C21-C1	-3.42	1.34	1.39
3	A	302	E1F	C3-S16	3.25	1.82	1.77
3	A	302	E1F	C2-CL1	-3.20	1.65	1.73
3	C	302	E1F	C7-S8	2.76	1.86	1.80
3	D	302	E1F	O17-S16	2.70	1.48	1.43
3	B	302	E1F	C21-C2	2.39	1.42	1.38
3	D	302	E1F	C7-C6	-2.37	1.49	1.51
3	B	302	E1F	C7-C6	-2.20	1.49	1.51
3	B	302	E1F	O18-S16	2.10	1.47	1.43
3	C	302	E1F	C9-N10	2.09	1.37	1.33
3	A	302	E1F	C9-N11	2.05	1.37	1.33
3	B	302	E1F	C9-N11	2.01	1.37	1.33
3	C	302	E1F	C9-N11	2.00	1.37	1.33

All (55) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	302	E1F	C4-C3-C2	6.70	123.79	118.97
3	C	302	E1F	C7-S8-C9	-6.03	94.76	101.53
3	C	302	E1F	O18-S16-N19	5.20	115.34	107.28
3	A	302	E1F	O12-C6-C7	4.84	128.26	120.59
3	D	302	E1F	C13-N10-C9	4.83	122.90	114.92
3	B	302	E1F	C13-N10-C9	4.80	122.85	114.92
3	D	302	E1F	C4-C3-C2	4.77	122.40	118.97
3	A	302	E1F	O12-C6-C1	-4.73	112.96	120.68
3	B	302	E1F	C4-C3-C2	4.72	122.36	118.97
3	B	302	E1F	C5-C1-C21	4.62	125.05	119.24
3	B	302	E1F	N10-C9-N11	-4.16	119.68	126.97
3	D	302	E1F	C15-C13-N10	-4.01	116.63	123.42

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	302	E1F	O12-C6-C1	-3.96	114.22	120.68
3	A	302	E1F	N10-C9-N11	-3.95	120.05	126.97
3	B	302	E1F	C21-C2-C3	-3.83	116.31	121.25
3	B	302	E1F	C15-C13-N10	-3.77	117.04	123.42
3	D	302	E1F	C7-S8-C9	3.75	105.74	101.53
3	A	302	E1F	C14-N11-C9	3.63	120.93	114.92
3	A	302	E1F	C7-S8-C9	-3.62	97.47	101.53
3	A	302	E1F	C21-C2-C3	-3.42	116.84	121.25
3	C	302	E1F	O12-C6-C7	3.40	125.97	120.59
3	B	302	E1F	O18-S16-N19	3.34	112.45	107.28
3	D	302	E1F	N10-C9-N11	-3.18	121.39	126.97
3	D	302	E1F	O17-S16-N19	3.18	112.21	107.28
3	A	302	E1F	C4-C3-S16	-3.12	113.49	117.68
3	B	302	E1F	O18-S16-C3	-3.07	102.94	107.32
3	C	302	E1F	C13-N10-C9	3.06	119.99	114.92
3	B	302	E1F	C4-C5-C1	-3.03	117.14	120.76
3	D	302	E1F	C15-C14-N11	-3.00	118.33	123.42
3	B	302	E1F	C21-C1-C6	-2.81	114.28	120.35
3	D	302	E1F	C2-C3-S16	-2.74	119.08	123.21
3	D	302	E1F	O18-S16-N19	2.72	111.49	107.28
3	A	302	E1F	O18-S16-N19	2.69	111.44	107.28
3	D	302	E1F	C13-C15-C14	2.65	121.51	116.77
3	A	302	E1F	C21-C1-C6	-2.62	114.68	120.35
3	D	302	E1F	C5-C4-C3	-2.58	115.79	120.24
3	D	302	E1F	C14-N11-C9	2.57	119.17	114.92
3	D	302	E1F	O12-C6-C1	-2.45	116.68	120.68
3	B	302	E1F	C2-C3-S16	-2.43	119.54	123.21
3	A	302	E1F	C5-C4-C3	-2.40	116.09	120.24
3	D	302	E1F	O17-S16-C3	-2.38	103.92	107.32
3	A	302	E1F	C3-C2-CL1	2.36	123.47	121.62
3	C	302	E1F	C15-C13-N10	-2.30	119.52	123.42
3	A	302	E1F	C13-N10-C9	2.29	118.71	114.92
3	D	302	E1F	C3-C2-CL1	2.29	123.41	121.62
3	B	302	E1F	C14-N11-C9	2.24	118.62	114.92
3	C	302	E1F	O17-S16-C3	-2.19	104.20	107.32
3	C	302	E1F	O17-S16-N19	-2.15	103.96	107.28
3	D	302	E1F	O12-C6-C7	2.13	123.96	120.59
3	C	302	E1F	C21-C2-CL1	-2.12	115.22	118.55
3	C	302	E1F	C4-C3-S16	2.12	120.52	117.68
3	A	302	E1F	O17-S16-N19	2.08	110.51	107.28
3	D	302	E1F	C5-C1-C21	2.06	121.84	119.24
3	C	302	E1F	N10-C9-N11	-2.05	123.38	126.97

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	302	E1F	C7-C6-C1	2.04	119.35	117.81

There are no chirality outliers.

There are no torsion outliers.

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, 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	261/263 (99%)	-0.17	4 (1%) 70 76	6, 13, 25, 39	0
1	B	261/263 (99%)	-0.14	5 (1%) 64 70	5, 11, 24, 35	0
1	C	260/263 (98%)	0.66	37 (14%) 3 3	8, 20, 36, 47	0
1	D	260/263 (98%)	-0.26	0 100 100	7, 13, 22, 31	0
All	All	1042/1052 (99%)	0.02	46 (4%) 33 36	5, 14, 29, 47	0

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	55	ALA	7.9
1	C	158[A]	ILE	6.3
1	B	253	ASP	5.6
1	C	56	ASN	4.6
1	C	3[A]	LYS	4.6
1	C	174	VAL	4.0
1	B	254	GLU	3.9
1	B	252	PHE	3.8
1	C	58	GLN	3.7
1	C	235	ASP	3.7
1	C	159	PHE	3.6
1	C	234	MET	3.6
1	C	167	TYR	3.6
1	C	155	TYR	3.5
1	C	7	PHE	3.5
1	C	165	VAL	3.3
1	A	7	PHE	3.2
1	C	230	TYR	2.9
1	C	228	ALA	2.9
1	C	162	LEU	2.9
1	C	53	LEU	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	233[A]	HIS	2.8
1	C	177	PHE	2.8
1	C	231[A]	CYS	2.7
1	C	237	PRO	2.7
1	C	238	SER	2.7
1	C	61[A]	LEU	2.6
1	B	263	GLN	2.5
1	C	179	ILE	2.5
1	C	240	ARG	2.4
1	A	263	GLN	2.4
1	C	51	TYR	2.4
1	C	54	SER	2.3
1	A	3[A]	LYS	2.3
1	C	169	GLY	2.3
1	C	62	THR	2.3
1	B	137	GLU	2.2
1	C	236	ASP	2.2
1	A	82	GLN	2.2
1	C	223	LEU	2.2
1	C	59	PHE	2.1
1	C	173	PHE	2.1
1	C	57	LYS	2.1
1	C	232	THR	2.1
1	C	161	HIS	2.0
1	C	70	LEU	2.0

## 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, 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	EDO	D	303	4/4	0.15	5.46	21,21,23,25	0
3	E1F	A	302	21/21	0.12	4.28	7,14,31,33	0
3	E1F	C	302	21/21	0.15	3.85	9,19,32,33	0
3	E1F	D	302	21/21	0.12	3.72	7,13,22,23	0
3	E1F	B	302	21/21	0.13	3.57	5,13,32,34	0
4	EDO	B	304	4/4	0.10	3.51	15,19,24,24	0
4	EDO	D	304	4/4	0.15	3.31	25,26,30,32	0
4	EDO	B	303	4/4	0.12	3.11	14,18,24,26	0
4	EDO	A	303	4/4	0.12	2.43	19,23,26,26	0
5	SO4	C	303	5/5	0.22	1.60	64,64,67,67	0
4	EDO	A	304	4/4	0.08	1.10	22,22,23,23	0
2	ZN	B	301	1/1	0.07	-0.22	5,5,5,5	0
2	ZN	C	301	1/1	0.07	-0.66	12,12,12,12	0
2	ZN	A	301	1/1	0.05	-1.12	7,7,7,7	0
2	ZN	D	301	1/1	0.05	-1.40	7,7,7,7	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.