



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 03:36 pm BST

PDB ID : 1EKJ
Title : THE X-RAY CRYSTALLOGRAPHIC STRUCTURE OF BETA CARBONIC ANHYDRASE FROM THE C3 DICOT PISUM SATIVUM
Authors : Kimber, M.S.; Pai, E.F.
Deposited on : 2000-03-08
Resolution : 1.93 Å(reported)

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

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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

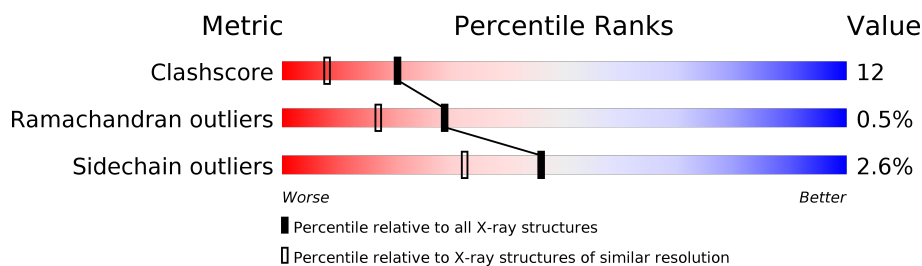
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.93 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	221	
1	B	221	
1	C	221	
1	D	221	
1	E	221	
1	F	221	
1	G	221	
1	H	221	

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
2	ACT	A	3001	-	-	X	-
2	ACT	C	3007	-	-	X	-
2	ACT	E	3005	-	-	X	-
2	ACT	F	3008	-	-	X	-

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 14070 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 BETA-CARBONIC ANHYDRASE.

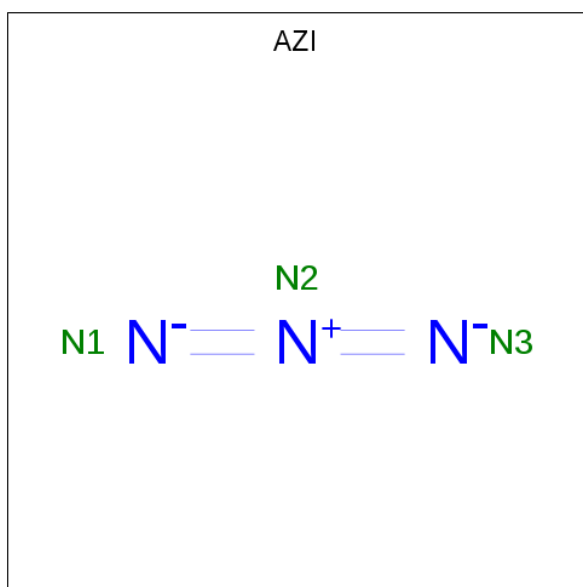
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	210	Total	C	N	O	S	0	0	0
			1618	1051	264	297	6			
1	B	213	Total	C	N	O	S	0	0	0
			1640	1065	268	301	6			
1	C	217	Total	C	N	O	S	0	1	0
			1672	1083	273	309	7			
1	D	211	Total	C	N	O	S	0	1	0
			1630	1057	266	300	7			
1	E	212	Total	C	N	O	S	0	0	0
			1633	1060	267	300	6			
1	F	214	Total	C	N	O	S	0	0	0
			1648	1071	269	302	6			
1	G	220	Total	C	N	O	S	0	1	0
			1691	1093	276	315	7			
1	H	211	Total	C	N	O	S	0	1	0
			1630	1057	266	300	7			

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula: C₂H₃O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		
2	E	1	Total	C	O	0	0
			4	2	2		
2	F	1	Total	C	O	0	0
			4	2	2		
2	G	1	Total	C	O	0	0
			4	2	2		
2	G	1	Total	C	O	0	0
			4	2	2		
2	H	1	Total	C	O	0	0
			4	2	2		

- Molecule 3 is AZIDE ION (three-letter code: AZI) (formula: N₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total N 3 3	0	0
3	D	1	Total N 3 3	0	0
3	E	1	Total N 3 3	0	0
3	H	1	Total N 3 3	0	0

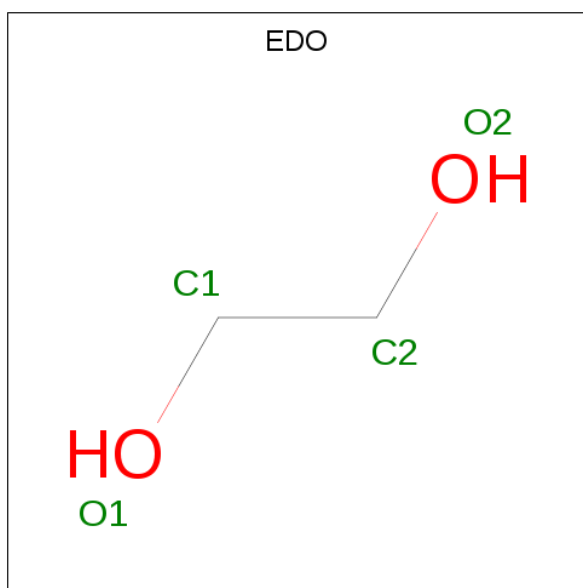
- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	G	1	Total Zn 1 1	0	0
4	D	1	Total Zn 1 1	0	0
4	E	1	Total Zn 1 1	0	0
4	H	1	Total Zn 1 1	0	0
4	B	1	Total Zn 1 1	0	0
4	C	1	Total Zn 1 1	0	0
4	A	1	Total Zn 1 1	0	0
4	F	1	Total Zn 1 1	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	G	1	Total Cl 1 1	0	0
5	D	1	Total Cl 1 1	0	0
5	E	1	Total Cl 1 1	0	0
5	H	1	Total Cl 1 1	0	0
5	B	1	Total Cl 1 1	0	0
5	C	1	Total Cl 1 1	0	0
5	A	1	Total Cl 1 1	0	0
5	F	1	Total Cl 1 1	0	0

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



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

Continued on next page...

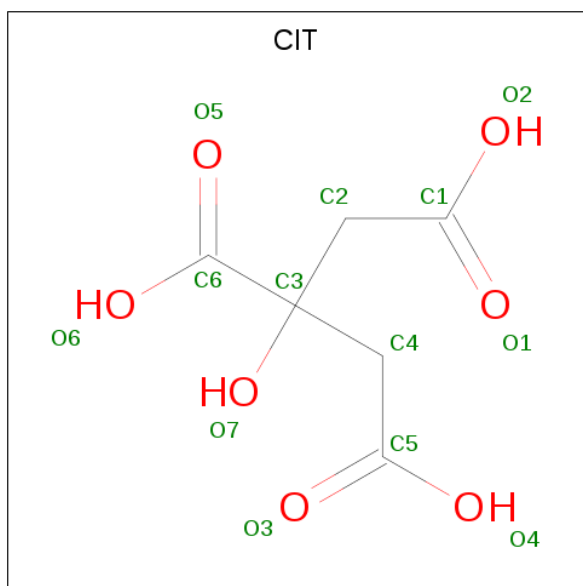
Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	H	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	G	1	Total	Cu	0	0
			1	1		
7	B	1	Total	Cu	0	0
			1	1		
7	C	1	Total	Cu	0	0
			1	1		
7	E	1	Total	Cu	0	0
			1	1		

- Molecule 8 is CITRIC ACID (three-letter code: CIT) (formula: C₆H₈O₇).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	D	1	Total	C	O	0	0
			13	6	7		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	50	Total	O	0	0
			50	50		

Continued on next page...

Continued from previous page...

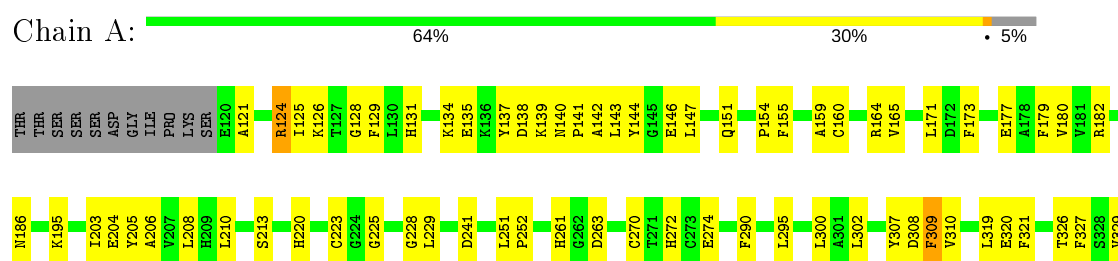
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	B	43	Total 43	O 43	0	0
9	C	138	Total 138	O 138	0	0
9	D	131	Total 131	O 131	0	0
9	E	71	Total 71	O 71	0	0
9	F	84	Total 84	O 84	0	0
9	G	150	Total 150	O 150	0	0
9	H	144	Total 144	O 144	0	0

3 Residue-property plots

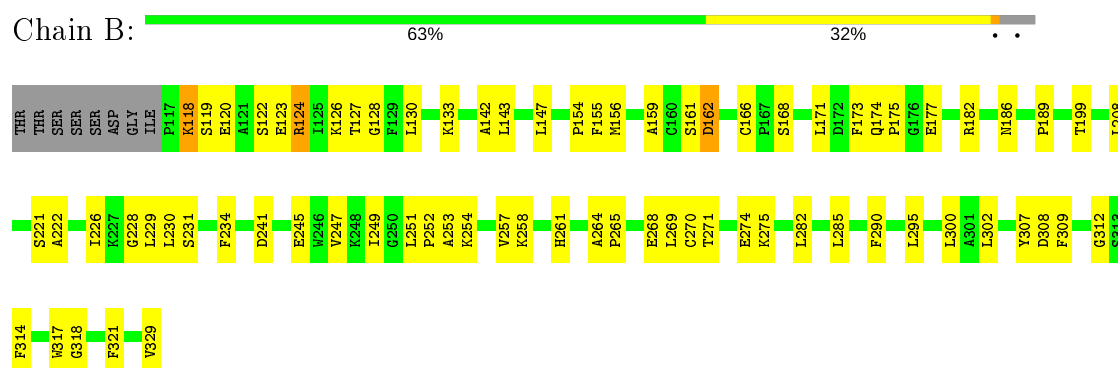
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according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

Note EDS was not executed.

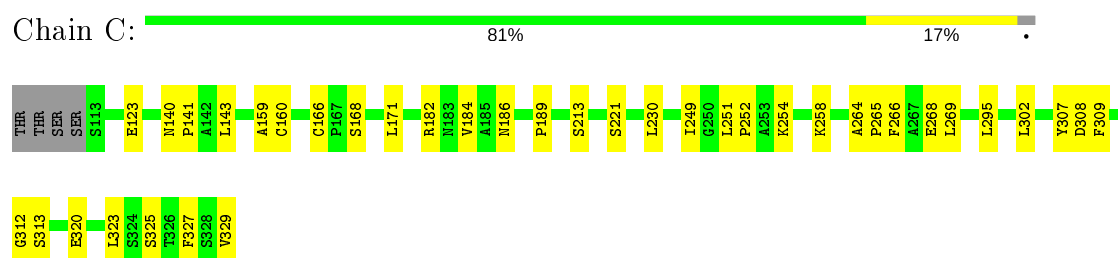
• Molecule 1: BETA-CARBONIC ANHYDRASE



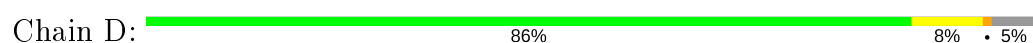
• Molecule 1: BETA-CARBONIC ANHYDRASE



• Molecule 1: BETA-CARBONIC ANHYDRASE



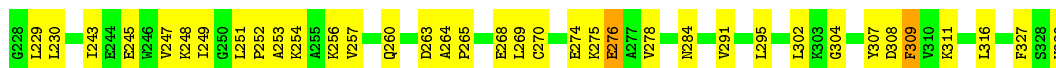
• Molecule 1: BETA-CARBONIC ANHYDRASE





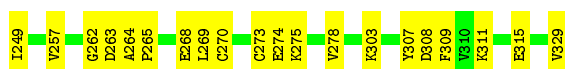
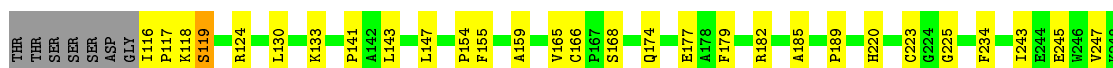
• Molecule 1: BETA-CARBONIC ANHYDRASE

Chain E: 66% 29%



• Molecule 1: BETA-CARBONIC ANHYDRASE

Chain F: 75% 22%



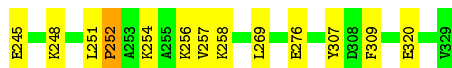
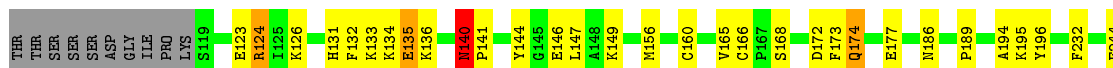
• Molecule 1: BETA-CARBONIC ANHYDRASE

Chain G: 73% 24%



• Molecule 1: BETA-CARBONIC ANHYDRASE

Chain H: 76% 18% 5%



4 Data and refinement statistics

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

Property	Value	Source
Space group	C 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	136.91Å 143.32Å 202.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 1.93	Depositor
% Data completeness (in resolution range)	87.5 (40.00-1.93)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 0.5	Depositor
R, R_{free}	0.229 , 0.250	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	14070	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

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: AZI, ZN, CL, CIT, EDO, ACT, CU

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	0.46	0/1662	0.64	0/2250
1	B	0.43	0/1685	0.63	1/2280 (0.0%)
1	C	0.62	0/1717	0.72	0/2325
1	D	0.75	0/1674	0.80	0/2267
1	E	0.48	0/1677	0.66	0/2269
1	F	0.51	0/1693	0.66	1/2292 (0.0%)
1	G	0.67	0/1736	0.78	2/2351 (0.1%)
1	H	0.75	0/1674	0.77	1/2267 (0.0%)
All	All	0.60	0/13518	0.71	5/18301 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	156	MET	N-CA-C	-5.12	97.16	111.00
1	F	179	PHE	N-CA-C	-5.07	97.32	111.00
1	G	112	SER	N-CA-C	5.06	124.67	111.00
1	B	156	MET	N-CA-C	-5.03	97.42	111.00
1	H	156	MET	N-CA-C	-5.01	97.47	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1618	0	1593	57	0
1	B	1640	0	1619	59	0
1	C	1672	0	1646	31	0
1	D	1630	0	1603	16	0
1	E	1633	0	1611	53	0
1	F	1648	0	1629	37	0
1	G	1691	0	1664	47	0
1	H	1630	0	1604	44	0
2	A	8	0	6	3	0
2	C	8	0	6	2	0
2	E	4	0	3	2	0
2	F	4	0	3	2	0
2	G	8	0	6	2	0
2	H	4	0	3	0	0
3	A	3	0	0	0	0
3	D	3	0	0	0	0
3	E	3	0	0	0	0
3	H	3	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
4	G	1	0	0	0	0
4	H	1	0	0	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0
5	F	1	0	0	0	0
5	G	1	0	0	0	0
5	H	1	0	0	0	0
6	A	4	0	6	0	0
6	D	4	0	6	0	0
6	E	4	0	6	1	0
6	H	4	0	6	0	0
7	B	1	0	0	0	0
7	C	1	0	0	0	0
7	E	1	0	0	0	0
7	G	1	0	0	0	0
8	D	13	0	4	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	A	50	0	0	0	0
9	B	43	0	0	4	0
9	C	138	0	0	3	0
9	D	131	0	0	2	0
9	E	71	0	0	2	0
9	F	84	0	0	4	0
9	G	150	0	0	12	0
9	H	144	0	0	4	0
All	All	14070	0	13024	325	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 12.

The worst 5 of 325 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:G:111:SER:HA	1:G:117:PRO:HA	1.29	1.10
1:H:134:LYS:HG3	1:H:135:GLU:H	1.23	1.03
1:F:166:CYS:HB2	9:F:1130:HOH:O	1.63	0.99
1:G:110:THR:O	1:G:111:SER:HB2	1.60	0.99
1:H:146:GLU:O	1:H:149:LYS:HG2	1.66	0.94

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	208/221 (94%)	190 (91%)	17 (8%)	1 (0%)	29	17
1	B	211/221 (96%)	192 (91%)	17 (8%)	2 (1%)	17	7
1	C	216/221 (98%)	209 (97%)	7 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	210/221 (95%)	205 (98%)	5 (2%)	0	100	100
1	E	210/221 (95%)	194 (92%)	15 (7%)	1 (0%)	29	17
1	F	212/221 (96%)	200 (94%)	10 (5%)	2 (1%)	17	7
1	G	219/221 (99%)	206 (94%)	12 (6%)	1 (0%)	29	17
1	H	210/221 (95%)	199 (95%)	9 (4%)	2 (1%)	15	6
All	All	1696/1768 (96%)	1595 (94%)	92 (5%)	9 (0%)	29	17

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	111	SER
1	A	139	LYS
1	F	119	SER
1	B	222	ALA
1	H	135	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	170/180 (94%)	166 (98%)	4 (2%)	49	36
1	B	173/180 (96%)	169 (98%)	4 (2%)	50	38
1	C	177/180 (98%)	174 (98%)	3 (2%)	60	49
1	D	172/180 (96%)	168 (98%)	4 (2%)	50	38
1	E	172/180 (96%)	166 (96%)	6 (4%)	36	21
1	F	174/180 (97%)	170 (98%)	4 (2%)	50	38
1	G	180/180 (100%)	174 (97%)	6 (3%)	38	24
1	H	172/180 (96%)	167 (97%)	5 (3%)	42	28
All	All	1390/1440 (96%)	1354 (97%)	36 (3%)	46	32

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	124	ARG
1	E	309	PHE
1	H	140	ASN
1	E	263	ASP
1	F	118	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	169	HIS
1	C	186	ASN
1	F	209	HIS
1	B	260	GLN
1	D	186	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 38 ligands modelled in this entry, 20 are monoatomic - leaving 18 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	ACT	G	3009	-	1,3,3	2.54	1 (100%)	0,3,3	0.00	-
2	ACT	G	3002	4	1,3,3	3.36	1 (100%)	0,3,3	0.00	-
6	EDO	A	3303	-	3,3,3	0.79	0	2,2,2	0.40	0
2	ACT	E	3005	4	1,3,3	2.10	1 (100%)	0,3,3	0.00	-
2	ACT	C	3007	-	1,3,3	3.47	1 (100%)	0,3,3	0.00	-
2	ACT	F	3008	-	1,3,3	3.14	1 (100%)	0,3,3	0.00	-
3	AZI	H	3201	-	0,2,2	0.00	-	0,1,1	0.00	-
6	EDO	D	3302	-	3,3,3	0.84	0	2,2,2	0.99	0
3	AZI	D	3202	-	0,2,2	0.00	-	0,1,1	0.00	-
2	ACT	C	3004	-	1,3,3	3.23	1 (100%)	0,3,3	0.00	-
6	EDO	H	3301	-	3,3,3	1.25	0	2,2,2	0.52	0
2	ACT	A	3003	4	1,3,3	2.94	1 (100%)	0,3,3	0.00	-
3	AZI	E	3204	-	0,2,2	0.00	-	0,1,1	0.00	-
6	EDO	E	3304	-	3,3,3	0.75	0	2,2,2	0.42	0
2	ACT	H	3006	4	1,3,3	2.28	1 (100%)	0,3,3	0.00	-
8	CIT	D	3101	-	3,12,12	4.01	1 (33%)	3,17,17	0.41	0
3	AZI	A	3203	-	0,2,2	0.00	-	0,1,1	0.00	-
2	ACT	A	3001	4	1,3,3	2.94	1 (100%)	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
6	EDO	H	3301	-	-	0/1/1/1	-
6	EDO	A	3303	-	-	0/1/1/1	-
6	EDO	E	3304	-	-	0/1/1/1	-
8	CIT	D	3101	-	-	6/6/16/16	-
6	EDO	D	3302	-	-	0/1/1/1	-

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	D	3101	CIT	O7-C3	-6.94	1.32	1.43
2	C	3007	ACT	CH3-C	3.47	1.53	1.48
2	G	3002	ACT	CH3-C	3.36	1.53	1.48
2	C	3004	ACT	CH3-C	3.23	1.52	1.48
2	F	3008	ACT	CH3-C	3.14	1.52	1.48

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	D	3101	CIT	C1-C2-C3-O7
8	D	3101	CIT	C1-C2-C3-C4
8	D	3101	CIT	C1-C2-C3-C6
8	D	3101	CIT	C2-C3-C4-C5
8	D	3101	CIT	O7-C3-C4-C5

There are no ring outliers.

9 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	3009	ACT	1	0
2	G	3002	ACT	1	0
2	E	3005	ACT	2	0
2	C	3007	ACT	2	0
2	F	3008	ACT	2	0
2	A	3003	ACT	1	0
6	E	3304	EDO	1	0
8	D	3101	CIT	1	0
2	A	3001	ACT	2	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.