



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 10, 2026 – 05:47 AM UTC

PDB ID : 4XIX / pdb_00004xix
Title : Carbonic anhydrase Cah3 from Chlamydomonas reinhardtii in complex with phosphate.
Authors : Hainzl, T.; Grundstrom, C.; Benlloch, R.; Shevela, D.; Shutova, T.; Messinger, J.; Samuelsson, G.; Sauer-Eriksson, A.E.
Deposited on : 2015-01-08
Resolution : 2.70 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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

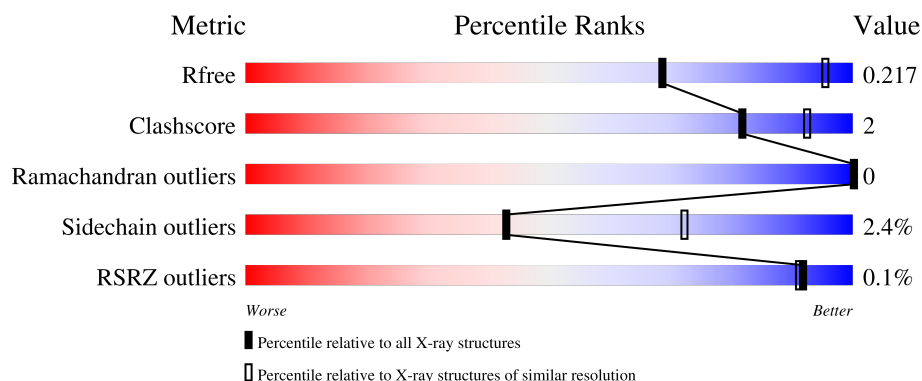
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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.



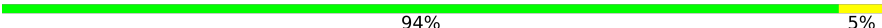
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	3538 (2.70-2.70)
Clashscore	190562	3843 (2.70-2.70)
Ramachandran outliers	187476	3778 (2.70-2.70)
Sidechain outliers	187428	3778 (2.70-2.70)
RSRZ outliers	180081	3538 (2.70-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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	239	 91% 8%
1	B	239	 90% 8% .
1	C	239	 89% 11%
1	D	239	 89% 10% .
1	E	239	 92% 7% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	239	 94%5%
1	G	239	 92%7%
1	H	239	 90%10%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 29590 atoms, of which 14584 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, alpha type.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	238	Total	C	H	N	O	S	0	0	0
			3654	1175	1822	311	334	12			
1	B	237	Total	C	H	N	O	S	0	0	0
			3644	1172	1817	310	333	12			
1	C	239	Total	C	H	N	O	S	0	0	0
			3671	1180	1831	312	335	13			
1	D	237	Total	C	H	N	O	S	0	0	0
			3644	1172	1817	310	333	12			
1	E	238	Total	C	H	N	O	S	0	0	0
			3654	1175	1822	311	334	12			
1	F	238	Total	C	H	N	O	S	0	0	0
			3654	1175	1822	311	334	12			
1	G	238	Total	C	H	N	O	S	0	0	0
			3654	1175	1822	311	334	12			
1	H	239	Total	C	H	N	O	S	0	0	0
			3671	1180	1831	312	335	13			

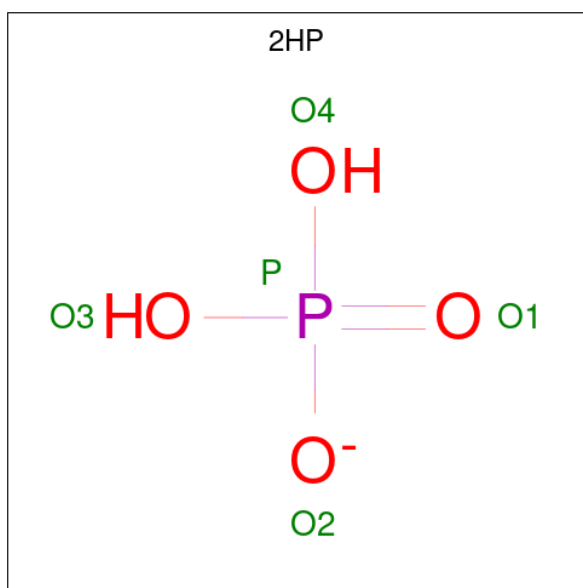
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	72	MET	-	initiating methionine	UNP Q39588
B	72	MET	-	initiating methionine	UNP Q39588
C	72	MET	-	initiating methionine	UNP Q39588
D	72	MET	-	initiating methionine	UNP Q39588
E	72	MET	-	initiating methionine	UNP Q39588
F	72	MET	-	initiating methionine	UNP Q39588
G	72	MET	-	initiating methionine	UNP Q39588
H	72	MET	-	initiating methionine	UNP Q39588

- Molecule 2 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

- Molecule 3 is DIHYDROGENPHOSPHATE ION (CCD ID: 2HP) (formula: $\text{H}_2\text{O}_4\text{P}$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	P	0	0
			5	4	1		
3	A	1	Total	O	P	0	0
			5	4	1		
3	A	1	Total	O	P	0	0
			5	4	1		
3	A	1	Total	O	P	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		
3	C	1	Total	O	P	0	0
			5	4	1		
3	C	1	Total	O	P	0	0
			5	4	1		
3	C	1	Total	O	P	0	0
			5	4	1		
3	C	1	Total	O	P	0	0
			5	4	1		
3	D	1	Total	O	P	0	0
			5	4	1		
3	D	1	Total	O	P	0	0
			5	4	1		
3	E	1	Total	O	P	0	0
			5	4	1		
3	E	1	Total	O	P	0	0
			5	4	1		
3	E	1	Total	O	P	0	0
			5	4	1		
3	F	1	Total	O	P	0	0
			5	4	1		
3	F	1	Total	O	P	0	0
			5	4	1		
3	G	1	Total	O	P	0	0
			5	4	1		
3	G	1	Total	O	P	0	0
			5	4	1		
3	H	1	Total	O	P	0	0
			5	4	1		
3	H	1	Total	O	P	0	0
			5	4	1		
3	H	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	46	Total 46	O 46	0	0
4	B	34	Total 34	O 34	0	0
4	C	36	Total 36	O 36	0	0
4	D	16	Total 16	O 16	0	0
4	E	21	Total 21	O 21	0	0
4	F	15	Total 15	O 15	0	0
4	G	20	Total 20	O 20	0	0
4	H	28	Total 28	O 28	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: Carbonic anhydrase, alpha type

Chain A:  91% 8%



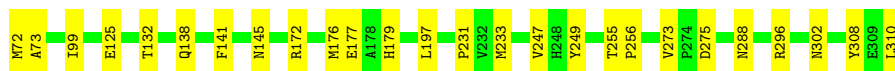
- Molecule 1: Carbonic anhydrase, alpha type

Chain B:  90% 8%



- Molecule 1: Carbonic anhydrase, alpha type

Chain C:  89% 11%

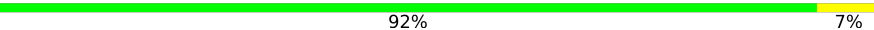


- Molecule 1: Carbonic anhydrase, alpha type

Chain D:  89% 10%

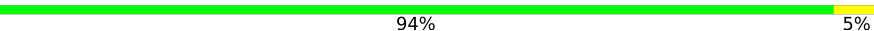


- Molecule 1: Carbonic anhydrase, alpha type

Chain E:  92% 7%



- Molecule 1: Carbonic anhydrase, alpha type

Chain F:  94% 5%



- Molecule 1: Carbonic anhydrase, alpha type

Chain G:  92% 7%



- Molecule 1: Carbonic anhydrase, alpha type

Chain H:  90% 10%



4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	138.96Å 138.96Å 202.81Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.45 – 2.70 48.45 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.45-2.70) 99.9 (48.45-2.70)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.65 (at 2.69Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.156 , 0.230 (Not available) , 0.217	Depositor DCC
R_{free} test set	3178 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	48.5	Xtriage
Anisotropy	0.719	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 37.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.025 for -h,-k,l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	29590	wwPDB-VP
Average B, all atoms (Å ²)	33.0	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 22.88 % 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 5.2126e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, 2HP

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	0.85	2/1884 (0.1%)	0.95	4/2561 (0.2%)
1	B	0.88	0/1879	0.96	3/2554 (0.1%)
1	C	0.87	1/1892 (0.1%)	1.00	3/2571 (0.1%)
1	D	0.85	0/1879	0.94	3/2554 (0.1%)
1	E	0.86	2/1884 (0.1%)	0.86	2/2561 (0.1%)
1	F	0.76	0/1884	0.90	1/2561 (0.0%)
1	G	0.84	1/1884 (0.1%)	0.94	1/2561 (0.0%)
1	H	0.82	1/1892 (0.1%)	0.93	2/2571 (0.1%)
All	All	0.84	7/15078 (0.0%)	0.94	19/20494 (0.1%)

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	248	HIS	C-O	-7.22	1.15	1.23
1	C	256	PRO	CA-C	6.23	1.55	1.51
1	A	178	ALA	CA-CB	-6.09	1.44	1.53
1	H	74	ALA	CA-CB	-5.95	1.43	1.53
1	A	265	PHE	CA-C	-5.71	1.45	1.52
1	G	79	GLU	CA-C	-5.62	1.45	1.53
1	E	247	VAL	C-O	-5.04	1.17	1.24

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	82	GLY	CA-C-N	6.65	127.23	120.38
1	B	82	GLY	C-N-CA	6.65	127.23	120.38
1	D	82	GLY	CA-C-N	6.13	126.69	120.38
1	D	82	GLY	C-N-CA	6.13	126.69	120.38
1	D	83	PRO	N-CA-C	6.06	118.09	110.70
1	C	302	ASN	CB-CA-C	-5.71	109.98	116.54

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	83	PRO	CA-C-N	-5.66	114.28	119.82
1	E	83	PRO	C-N-CA	-5.66	114.28	119.82
1	A	269	GLN	CA-C-N	-5.62	114.96	120.52
1	A	269	GLN	C-N-CA	-5.62	114.96	120.52
1	G	79	GLU	N-CA-C	-5.52	101.73	109.69
1	C	255	THR	CA-C-N	-5.41	115.71	119.66
1	C	255	THR	C-N-CA	-5.41	115.71	119.66
1	A	83	PRO	CA-C-N	-5.39	114.26	119.87
1	A	83	PRO	C-N-CA	-5.39	114.26	119.87
1	F	137	MET	N-CA-C	-5.26	102.26	110.42
1	H	83	PRO	N-CA-C	5.17	117.01	110.70
1	B	256	PRO	O-C-N	5.11	123.66	121.31
1	H	229	ILE	N-CA-C	5.03	114.51	107.37

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	1832	1822	1818	8	0
1	B	1827	1817	1813	12	0
1	C	1840	1831	1827	13	0
1	D	1827	1817	1813	10	0
1	E	1832	1822	1818	10	0
1	F	1832	1822	1818	7	0
1	G	1832	1822	1818	9	0
1	H	1840	1831	1827	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
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	20	0	0	0	0
3	B	15	0	0	0	0
3	C	20	0	0	0	0
3	D	10	0	0	0	0
3	E	15	0	0	0	0
3	F	10	0	0	0	0
3	G	10	0	0	0	0
3	H	20	0	0	0	0
4	A	46	0	0	0	0
4	B	34	0	0	0	0
4	C	36	0	0	0	0
4	D	16	0	0	0	0
4	E	21	0	0	0	0
4	F	15	0	0	0	0
4	G	20	0	0	0	0
4	H	28	0	0	0	0
All	All	15006	14584	14552	73	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 (73) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:177:GLU:OE2	1:E:179:HIS:NE2	2.28	0.61
1:E:217:GLU:HG2	1:H:72:MET:HE2	1.88	0.56
1:H:278:ILE:HG22	1:H:282:MET:CE	2.37	0.55
1:A:268:MET:HA	1:A:268:MET:HE2	1.90	0.54
1:D:247:VAL:HG13	1:D:268:MET:HE3	1.90	0.54
1:B:203:ILE:HB	1:F:203:ILE:HD11	1.90	0.53
1:A:177:GLU:OE2	1:A:179:HIS:NE2	2.39	0.53
1:F:103:LEU:HD21	1:F:268:MET:SD	2.48	0.53
1:B:141:PHE:CE2	1:B:231:PRO:HG3	2.45	0.52
1:E:112:ALA:HB1	1:E:114:MET:HE3	1.91	0.52
1:C:176:MET:HB2	1:C:197:LEU:HB2	1.93	0.51
1:H:103:LEU:O	1:H:105:THR:N	2.42	0.51
1:E:275:ASP:N	1:E:275:ASP:OD1	2.44	0.50
1:C:72:MET:O	1:C:73:ALA:HB3	2.12	0.49
1:H:176:MET:HB2	1:H:197:LEU:HB2	1.94	0.49
1:D:206:PRO:HG2	1:D:233:MET:HG3	1.95	0.48
1:C:177:GLU:OE2	1:C:179:HIS:NE2	2.45	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:121:TYR:OH	1:A:182:HIS:NE2	2.40	0.48
1:D:247:VAL:HG13	1:D:268:MET:CE	2.44	0.48
1:G:141:PHE:CZ	1:G:231:PRO:HG3	2.48	0.48
1:D:169:MET:HE2	1:D:196:MET:HE2	1.96	0.48
1:H:130:LEU:HB3	1:H:222:LYS:HG2	1.96	0.48
1:B:167:HIS:NE2	1:B:249:TYR:OH	2.42	0.47
1:G:148:PHE:CE2	1:G:153:GLU:HG3	2.50	0.47
1:H:275:ASP:OD1	1:H:275:ASP:N	2.47	0.47
1:E:167:HIS:HE1	1:E:264:TRP:CZ3	2.33	0.47
1:H:268:MET:HA	1:H:268:MET:HE2	1.97	0.47
1:H:136:THR:HG21	1:H:160:HIS:HB2	1.96	0.47
1:A:83:PRO:HA	1:A:86:TRP:CD2	2.50	0.46
1:A:275:ASP:OD1	1:A:275:ASP:N	2.48	0.46
1:C:141:PHE:CE1	1:C:231:PRO:HG3	2.50	0.46
1:H:83:PRO:HA	1:H:86:TRP:CD2	2.50	0.46
1:F:203:ILE:HG22	1:F:204:LYS:O	2.15	0.46
1:F:233:MET:HE2	1:F:233:MET:HA	1.98	0.46
1:E:233:MET:HA	1:E:233:MET:HE2	1.98	0.46
1:D:124:PHE:HB2	1:D:229:ILE:O	2.16	0.45
1:B:148:PHE:CE1	1:B:153:GLU:HB2	2.52	0.45
1:F:141:PHE:CZ	1:F:231:PRO:HG3	2.52	0.45
1:D:193:LEU:HD22	1:D:265:PHE:HB2	1.99	0.45
1:B:159:PHE:C	1:B:159:PHE:CD1	2.96	0.44
1:B:180:LEU:HD12	1:B:180:LEU:N	2.31	0.44
1:D:83:PRO:HA	1:D:86:TRP:CE2	2.52	0.44
1:B:114:MET:HE1	1:B:248:HIS:CE1	2.53	0.44
1:C:233:MET:HA	1:C:233:MET:HE2	2.00	0.44
1:C:275:ASP:N	1:C:275:ASP:OD1	2.51	0.43
1:B:275:ASP:OD1	1:B:275:ASP:N	2.51	0.43
1:D:83:PRO:HA	1:D:86:TRP:CD2	2.53	0.43
1:A:77:TYR:CE2	1:G:294:ASN:HB2	2.52	0.43
1:G:203:ILE:HG22	1:G:204:LYS:O	2.19	0.43
1:C:172:ARG:HH22	1:H:72:MET:HG2	1.82	0.43
1:G:176:MET:HB2	1:G:197:LEU:HB2	1.99	0.42
1:B:268:MET:HE2	1:B:268:MET:HA	2.00	0.42
1:A:268:MET:HG3	1:A:310:LEU:CD2	2.49	0.42
1:C:72:MET:HE1	1:E:219:PRO:HB2	2.01	0.42
1:C:99:ILE:HG12	1:C:249:TYR:CZ	2.55	0.42
1:C:125:GLU:OE2	1:C:145:ASN:ND2	2.44	0.42
1:H:121:TYR:OH	1:H:182:HIS:NE2	2.46	0.42
1:G:230:ASN:HA	1:G:231:PRO:HD3	1.96	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:285:VAL:HG12	1:A:292:ALA:HB3	2.02	0.42
1:G:275:ASP:OD1	1:G:275:ASP:N	2.53	0.41
1:G:268:MET:HA	1:G:268:MET:HE2	2.02	0.41
1:B:114:MET:HE1	1:B:248:HIS:ND1	2.35	0.41
1:B:141:PHE:O	1:C:72:MET:O	2.39	0.41
1:F:211:ALA:O	1:F:215:ALA:HB2	2.21	0.41
1:D:285:VAL:HG12	1:D:292:ALA:HB3	2.03	0.41
1:C:308:TYR:HB3	1:C:310:LEU:HG	2.03	0.41
1:F:159:PHE:C	1:F:159:PHE:CD1	2.99	0.41
1:G:103:LEU:O	1:G:109:LYS:NZ	2.46	0.41
1:E:220:LEU:HD12	1:E:220:LEU:N	2.36	0.40
1:C:275:ASP:OD2	1:E:276:SER:HA	2.21	0.40
1:B:77:TYR:O	1:H:80:VAL:CG1	2.70	0.40
1:E:83:PRO:HA	1:E:86:TRP:CD2	2.56	0.40
1:D:94:LYS:HB3	1:D:302:ASN:HD22	1.87	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	236/239 (99%)	220 (93%)	16 (7%)	0	100	100
1	B	235/239 (98%)	222 (94%)	13 (6%)	0	100	100
1	C	237/239 (99%)	219 (92%)	18 (8%)	0	100	100
1	D	235/239 (98%)	225 (96%)	10 (4%)	0	100	100
1	E	236/239 (99%)	220 (93%)	16 (7%)	0	100	100
1	F	236/239 (99%)	218 (92%)	18 (8%)	0	100	100
1	G	236/239 (99%)	220 (93%)	16 (7%)	0	100	100
1	H	237/239 (99%)	223 (94%)	14 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1888/1912 (99%)	1767 (94%)	121 (6%)	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	198/199 (100%)	193 (98%)	5 (2%)	42	71
1	B	198/199 (100%)	192 (97%)	6 (3%)	36	66
1	C	199/199 (100%)	193 (97%)	6 (3%)	36	66
1	D	198/199 (100%)	191 (96%)	7 (4%)	32	61
1	E	198/199 (100%)	193 (98%)	5 (2%)	42	71
1	F	198/199 (100%)	196 (99%)	2 (1%)	68	86
1	G	198/199 (100%)	195 (98%)	3 (2%)	57	81
1	H	199/199 (100%)	195 (98%)	4 (2%)	48	76
All	All	1586/1592 (100%)	1548 (98%)	38 (2%)	43	72

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	169	MET
1	A	198	GLU
1	A	273	VAL
1	A	295	THR
1	A	303	SER
1	B	130	LEU
1	B	153	GLU
1	B	158	GLN
1	B	186	SER
1	B	217	GLU
1	B	273	VAL
1	C	132	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	138	GLN
1	C	247	VAL
1	C	273	VAL
1	C	288	ASN
1	C	296	ARG
1	D	128	ASP
1	D	172	ARG
1	D	183	LYS
1	D	213	GLU
1	D	273	VAL
1	D	288	ASN
1	D	303	SER
1	E	185	LYS
1	E	193	LEU
1	E	233	MET
1	E	275	ASP
1	E	295	THR
1	F	127	CYS
1	F	288	ASN
1	G	153	GLU
1	G	166	GLU
1	G	169	MET
1	H	85	THR
1	H	104	ASN
1	H	295	THR
1	H	303	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	288	ASN
1	B	299	GLN
1	C	134	HIS
1	C	299	GLN
1	D	145	ASN
1	D	189	ASN
1	E	205	ASN
1	E	248	HIS
1	F	138	GLN
1	F	140	ASN
1	F	299	GLN
1	G	76	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	G	145	ASN
1	H	248	HIS

5.3.3 RNA ⓘ

There are no RNA molecules 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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 32 ligands modelled in this entry, 8 are monoatomic - leaving 24 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	2HP	E	404	-	4,4,4	1.23	1 (25%)	6,6,6	0.48	0
3	2HP	C	402	2	4,4,4	1.22	1 (25%)	6,6,6	0.48	0
3	2HP	E	403	-	4,4,4	1.22	1 (25%)	6,6,6	0.47	0
3	2HP	D	403	-	4,4,4	1.23	1 (25%)	6,6,6	0.47	0
3	2HP	A	405	-	4,4,4	1.23	1 (25%)	6,6,6	0.47	0
3	2HP	B	402	2	4,4,4	1.25	1 (25%)	6,6,6	0.49	0
3	2HP	F	403	-	4,4,4	1.24	1 (25%)	6,6,6	0.48	0
3	2HP	C	405	-	4,4,4	1.23	1 (25%)	6,6,6	0.48	0
3	2HP	H	405	-	4,4,4	1.22	1 (25%)	6,6,6	0.47	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	2HP	A	403	-	4,4,4	1.23	1 (25%)	6,6,6	0.48	0
3	2HP	B	404	-	4,4,4	1.24	1 (25%)	6,6,6	0.48	0
3	2HP	H	404	-	4,4,4	1.24	1 (25%)	6,6,6	0.49	0
3	2HP	F	402	2	4,4,4	1.22	1 (25%)	6,6,6	0.47	0
3	2HP	A	404	-	4,4,4	1.23	1 (25%)	6,6,6	0.47	0
3	2HP	C	403	-	4,4,4	1.16	1 (25%)	6,6,6	0.42	0
3	2HP	H	403	-	4,4,4	1.25	1 (25%)	6,6,6	0.48	0
3	2HP	C	404	-	4,4,4	1.24	1 (25%)	6,6,6	0.48	0
3	2HP	G	402	2	4,4,4	1.24	1 (25%)	6,6,6	0.47	0
3	2HP	D	402	-	4,4,4	1.24	1 (25%)	6,6,6	0.49	0
3	2HP	B	403	-	4,4,4	1.23	1 (25%)	6,6,6	0.47	0
3	2HP	G	403	-	4,4,4	1.16	1 (25%)	6,6,6	0.41	0
3	2HP	E	402	2	4,4,4	1.23	1 (25%)	6,6,6	0.48	0
3	2HP	H	402	2	4,4,4	1.16	1 (25%)	6,6,6	0.41	0
3	2HP	A	402	2	4,4,4	1.23	1 (25%)	6,6,6	0.48	0

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	402	2HP	P-O1	2.37	1.56	1.50
3	H	403	2HP	P-O1	2.36	1.56	1.50
3	D	402	2HP	P-O1	2.36	1.56	1.50
3	H	404	2HP	P-O1	2.36	1.56	1.50
3	F	403	2HP	P-O1	2.36	1.56	1.50
3	C	404	2HP	P-O1	2.35	1.56	1.50
3	B	404	2HP	P-O1	2.34	1.56	1.50
3	G	402	2HP	P-O1	2.34	1.56	1.50
3	A	403	2HP	P-O1	2.33	1.56	1.50
3	E	404	2HP	P-O1	2.33	1.56	1.50
3	A	405	2HP	P-O1	2.33	1.56	1.50
3	E	402	2HP	P-O1	2.33	1.56	1.50
3	B	403	2HP	P-O1	2.33	1.56	1.50
3	D	403	2HP	P-O1	2.33	1.56	1.50
3	C	405	2HP	P-O1	2.33	1.56	1.50
3	E	403	2HP	P-O1	2.32	1.56	1.50
3	A	402	2HP	P-O1	2.32	1.56	1.50
3	H	405	2HP	P-O1	2.32	1.56	1.50
3	A	404	2HP	P-O1	2.31	1.56	1.50
3	F	402	2HP	P-O1	2.30	1.56	1.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	402	2HP	P-O1	2.29	1.56	1.50
3	C	403	2HP	P-O1	2.16	1.55	1.50
3	G	403	2HP	P-O1	2.14	1.55	1.50
3	H	402	2HP	P-O1	2.13	1.55	1.50

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, 95th 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 ²)	Q<0.9
1	A	238/239 (99%)	-0.73	1 (0%) 88 87	13, 29, 52, 68	0
1	B	237/239 (99%)	-0.74	0 100 100	10, 27, 50, 73	0
1	C	239/239 (100%)	-0.79	0 100 100	12, 26, 47, 64	0
1	D	237/239 (99%)	-0.68	0 100 100	14, 31, 56, 68	0
1	E	238/239 (99%)	-0.73	0 100 100	13, 30, 59, 90	0
1	F	238/239 (99%)	-0.56	0 100 100	12, 40, 69, 94	0
1	G	238/239 (99%)	-0.73	1 (0%) 88 87	10, 31, 62, 84	0
1	H	239/239 (100%)	-0.75	0 100 100	14, 30, 60, 80	0
All	All	1904/1912 (99%)	-0.71	2 (0%) 92 91	10, 30, 58, 94	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	73	ALA	2.2
1	G	73	ALA	2.1

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 oligosaccharides 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. The B-factors column lists the minimum, median, 95th 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 ²)	Q<0.9
3	2HP	B	404	5/5	0.73	0.29	21,32,36,38	5
3	2HP	E	404	5/5	0.88	0.21	13,22,28,28	5
3	2HP	D	403	5/5	0.89	0.10	35,47,78,89	0
3	2HP	A	403	5/5	0.89	0.14	38,46,52,63	0
3	2HP	B	403	5/5	0.90	0.21	15,21,28,44	5
3	2HP	H	404	5/5	0.90	0.11	25,54,57,58	0
3	2HP	C	405	5/5	0.92	0.10	37,44,58,72	0
3	2HP	H	405	5/5	0.93	0.06	41,41,47,66	0
3	2HP	A	405	5/5	0.94	0.08	29,32,36,41	0
3	2HP	B	402	5/5	0.95	0.10	21,25,40,41	0
3	2HP	F	403	5/5	0.95	0.06	41,44,52,66	0
3	2HP	E	402	5/5	0.95	0.09	10,23,37,41	0
3	2HP	E	403	5/5	0.95	0.11	44,55,63,68	0
3	2HP	H	402	5/5	0.96	0.06	25,27,35,39	0
3	2HP	C	403	5/5	0.96	0.06	20,23,25,30	0
3	2HP	G	403	5/5	0.96	0.06	35,44,51,64	0
3	2HP	A	402	5/5	0.97	0.07	18,21,31,36	0
3	2HP	A	404	5/5	0.98	0.04	28,29,34,38	0
3	2HP	G	402	5/5	0.98	0.06	20,22,26,32	0
3	2HP	C	404	5/5	0.98	0.05	15,16,20,24	0
3	2HP	C	402	5/5	0.98	0.08	25,26,37,43	0
3	2HP	H	403	5/5	0.98	0.07	29,29,30,33	0
3	2HP	D	402	5/5	0.98	0.07	21,25,32,35	0
3	2HP	F	402	5/5	0.98	0.07	22,23,36,36	0
2	ZN	A	401	1/1	1.00	0.01	18,18,18,18	0
2	ZN	B	401	1/1	1.00	0.02	20,20,20,20	0
2	ZN	C	401	1/1	1.00	0.02	24,24,24,24	0
2	ZN	D	401	1/1	1.00	0.02	26,26,26,26	0
2	ZN	E	401	1/1	1.00	0.02	19,19,19,19	0
2	ZN	F	401	1/1	1.00	0.03	26,26,26,26	0
2	ZN	G	401	1/1	1.00	0.04	16,16,16,16	0
2	ZN	H	401	1/1	1.00	0.04	19,19,19,19	0

6.5 Other polymers

There are no such residues in this entry.