



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 21, 2016 – 04:31 PM EDT

PDB ID : 4ZHX  
Title : Novel binding site for allosteric activation of AMPK  
Authors : Langendorf, C.G.; Ngoei, K.R.; Issa, S.M.A.; Ling, N.; Gorman, M.A.; Parker, M.W.; Sakamoto, K.; Scott, J.W.; Oakhill, J.S.; Kemp, B.E.  
Deposited on : 2015-04-27  
Resolution : 2.99 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027107  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0122  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027107

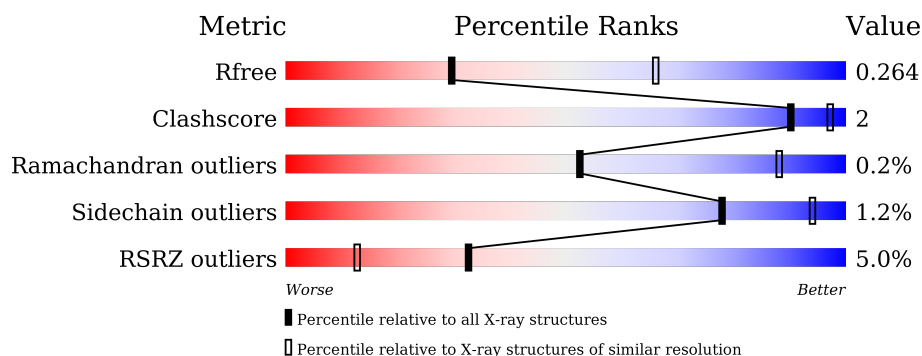
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.99 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))
$R_{free}$	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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	565	<div> <div>2%</div> <div> <div></div> <div>69%</div> <div>•</div> <div>28%</div> </div> </div>
1	C	565	<div> <div>3%</div> <div> <div></div> <div>74%</div> <div>6%</div> <div>20%</div> </div> </div>
2	B	270	<div> <div>7%</div> <div> <div></div> <div>56%</div> <div>8%</div> <div>36%</div> </div> </div>
2	D	270	<div> <div>4%</div> <div> <div></div> <div>57%</div> <div>8%</div> <div>35%</div> </div> </div>
3	E	336	<div> <div>5%</div> <div> <div></div> <div>79%</div> <div>10%</div> <div>11%</div> </div> </div>
3	F	336	<div> <div>5%</div> <div> <div></div> <div>83%</div> <div>6%</div> <div>11%</div> </div> </div>

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 14587 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 5'-AMP-activated protein kinase catalytic subunit alpha-2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	409	Total	C	N	O	P	S	0	0	0
			3255	2092	563	577	1	22			
1	C	453	Total	C	N	O	P	S	0	0	0
			3567	2283	617	640	1	26			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-12	MET	-	initiating methionine	UNP P54646
A	-11	GLY	-	expression tag	UNP P54646
A	-10	SER	-	expression tag	UNP P54646
A	-9	SER	-	expression tag	UNP P54646
A	-8	HIS	-	expression tag	UNP P54646
A	-7	HIS	-	expression tag	UNP P54646
A	-6	HIS	-	expression tag	UNP P54646
A	-5	HIS	-	expression tag	UNP P54646
A	-4	HIS	-	expression tag	UNP P54646
A	-3	HIS	-	expression tag	UNP P54646
A	-2	SER	-	expression tag	UNP P54646
A	-1	GLN	-	expression tag	UNP P54646
A	0	ASP	-	expression tag	UNP P54646
A	1	PRO	-	expression tag	UNP P54646
A	271	GLY	ASP	variant	UNP P54646
C	-12	MET	-	initiating methionine	UNP P54646
C	-11	GLY	-	expression tag	UNP P54646
C	-10	SER	-	expression tag	UNP P54646
C	-9	SER	-	expression tag	UNP P54646
C	-8	HIS	-	expression tag	UNP P54646
C	-7	HIS	-	expression tag	UNP P54646
C	-6	HIS	-	expression tag	UNP P54646
C	-5	HIS	-	expression tag	UNP P54646
C	-4	HIS	-	expression tag	UNP P54646
C	-3	HIS	-	expression tag	UNP P54646

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	SER	-	expression tag	UNP P54646
C	-1	GLN	-	expression tag	UNP P54646
C	0	ASP	-	expression tag	UNP P54646
C	1	PRO	-	expression tag	UNP P54646
C	271	GLY	ASP	variant	UNP P54646

- Molecule 2 is a protein called 5'-AMP-activated protein kinase subunit beta-1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	172	Total	C	N	O	P	S	0	0	0
			1310	845	221	240	1	3			
2	D	175	Total	C	N	O	P	S	0	0	0
			1373	884	230	252	1	6			

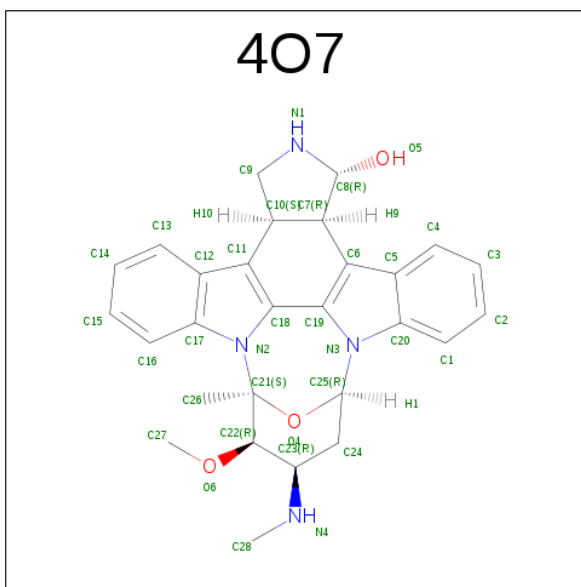
- Molecule 3 is a protein called 5'-AMP-activated protein kinase subunit gamma-1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	E	298	Total	C	N	O	S		0	2	0
			2381	1553	398	423	7				
3	F	300	Total	C	N	O	S		0	0	0
			2366	1535	396	428	7				

There are 12 discrepancies between the modelled and reference sequences:

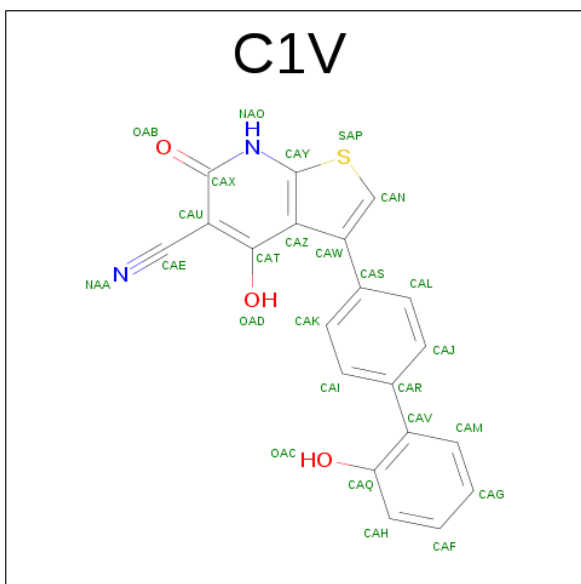
Chain	Residue	Modelled	Actual	Comment	Reference
E	-4	MET	-	initiating methionine	UNP P54619
E	-3	ALA	-	expression tag	UNP P54619
E	-2	ASP	-	expression tag	UNP P54619
E	-1	LEU	-	expression tag	UNP P54619
E	0	ASN	-	expression tag	UNP P54619
E	1	TRP	-	expression tag	UNP P54619
F	-4	MET	-	initiating methionine	UNP P54619
F	-3	ALA	-	expression tag	UNP P54619
F	-2	ASP	-	expression tag	UNP P54619
F	-1	LEU	-	expression tag	UNP P54619
F	0	ASN	-	expression tag	UNP P54619
F	1	TRP	-	expression tag	UNP P54619

- Molecule 4 is (5S,6R,7R,9R,13cR,14R,16aS)-6-methoxy-5-methyl-7-(methylamino)-6,7,8,9,14,15,16,16a-octahydro-5H,13cH-5,9-epoxy-4b,9a,15-triazadibenzo[b,h]cyclonona[1,2,3,4-jkl]cyclopenta[e]-as-indacen-14-ol (three-letter code: 4O7) (formula: C<sub>28</sub>H<sub>30</sub>N<sub>4</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			35	28	4	3		
4	C	1	Total	C	N	O	0	0
			35	28	4	3		

- Molecule 5 is 3-[4-(2-hydroxyphenyl)phenyl]-4-oxidanyl-6-oxidanylidene-7H-thieno[2,3-b]pyridine-5-carbonitrile (three-letter code: C1V) (formula:  $C_{20}H_{12}N_2O_3S$ ).



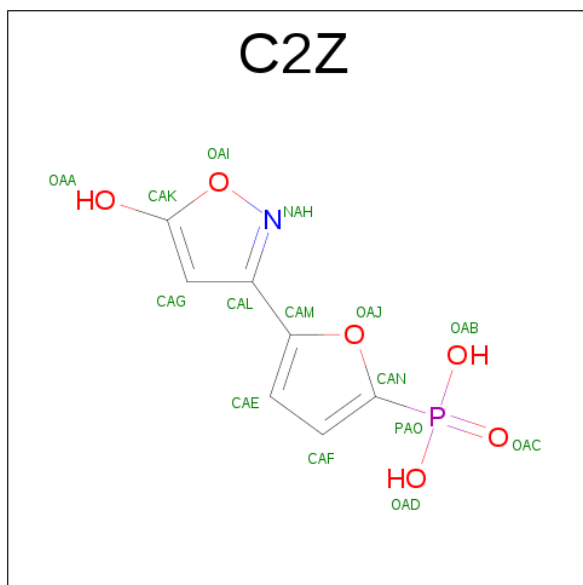
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total	C	N	O	S	0	0
			26	20	2	3	1		

*Continued on next page...*

Continued from previous page...

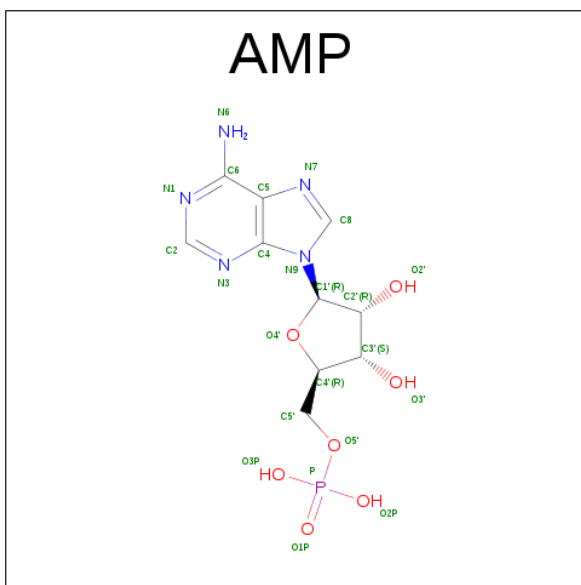
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	C	1	Total	C	N	O	S	0	0
			26	20	2	3	1		

- Molecule 6 is 5-(5-hydroxyl-isoxazol-3-yl)-furan-2-phosphonic acid (three-letter code: C2Z) (formula:  $C_7H_6NO_6P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	E	1	Total	C	N	O	P	0	0
			15	7	1	6	1		
6	E	1	Total	C	N	O	P	0	0
			15	7	1	6	1		

- Molecule 7 is ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula:  $C_{10}H_{14}N_5O_7P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	F	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
7	F	1	Total	C	N	O	P	0	0
			23	10	5	7	1		

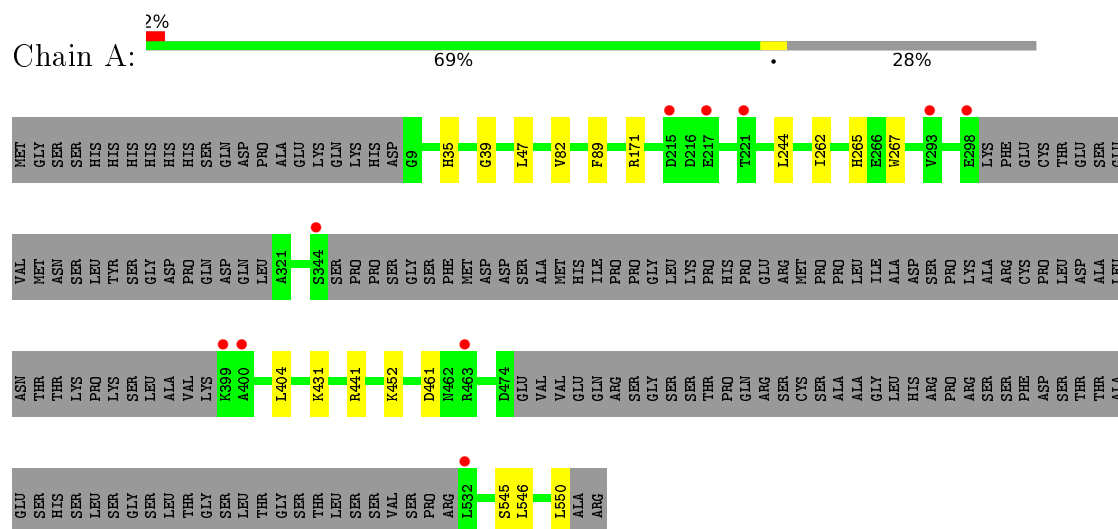
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	28	Total	O	0	0
			28	28		
8	B	13	Total	O	0	0
			13	13		
8	C	33	Total	O	0	0
			33	33		
8	D	16	Total	O	0	0
			16	16		
8	E	26	Total	O	0	0
			26	26		
8	F	21	Total	O	0	0
			21	21		

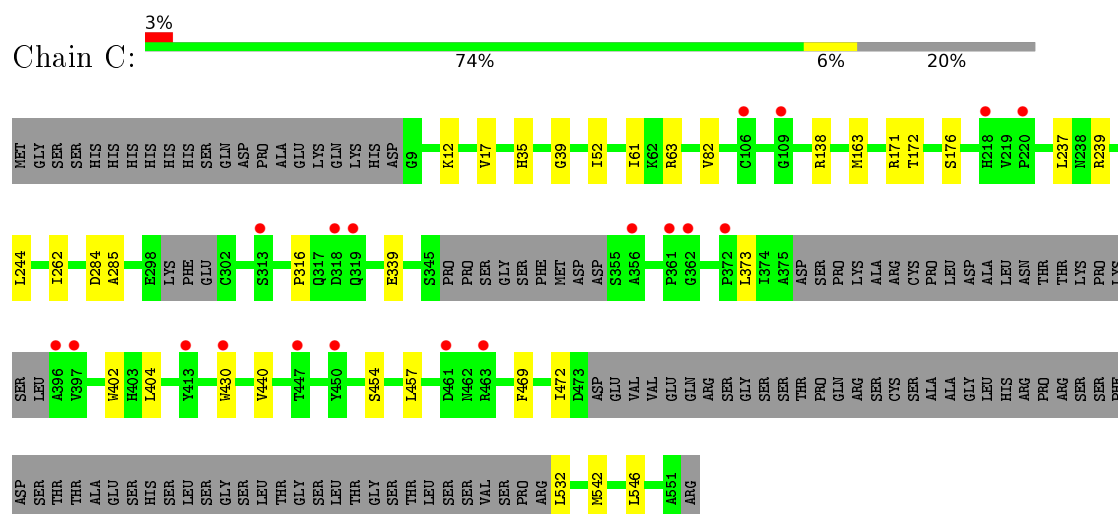
### 3 Residue-property plots

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: 5'-AMP-activated protein kinase catalytic subunit alpha-2



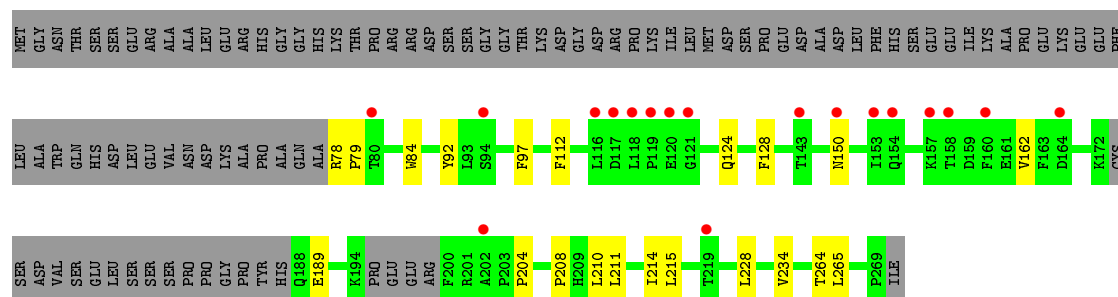
- Molecule 1: 5'-AMP-activated protein kinase catalytic subunit alpha-2



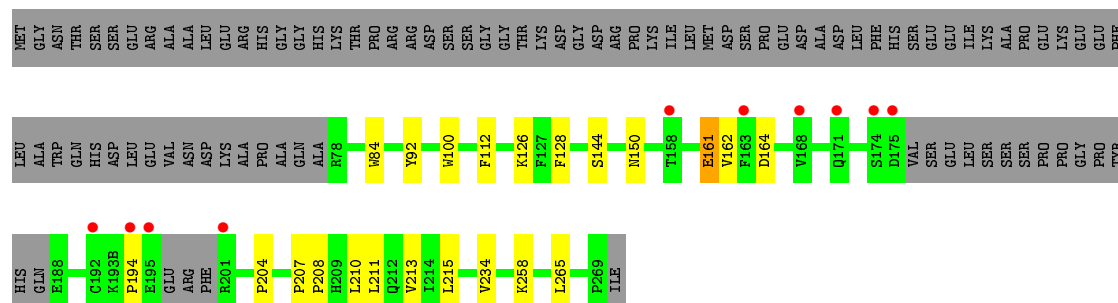
- Molecule 2: 5'-AMP-activated protein kinase subunit beta-1



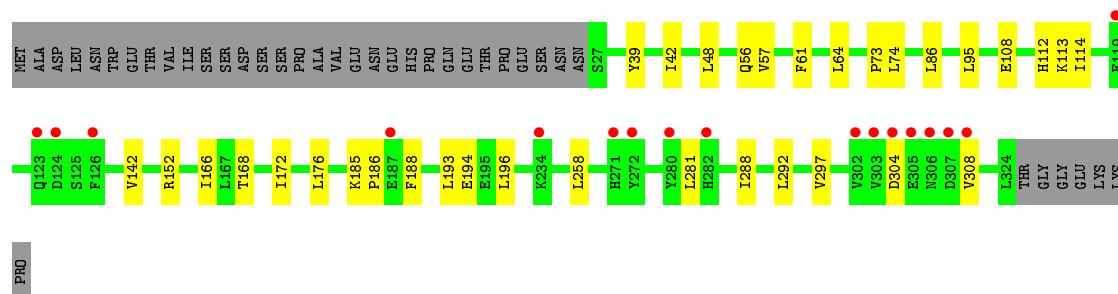
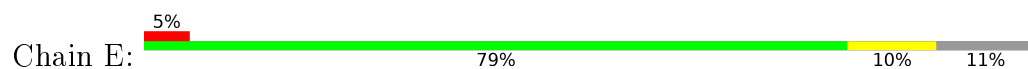




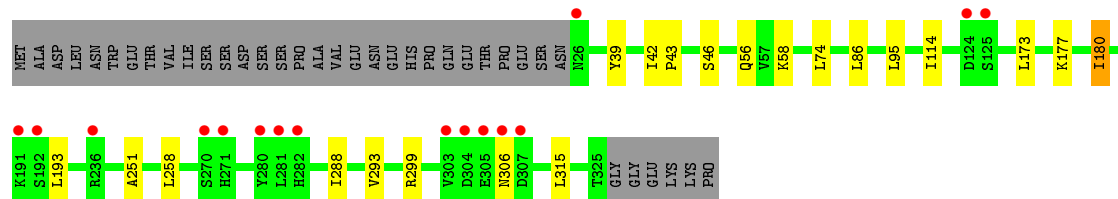
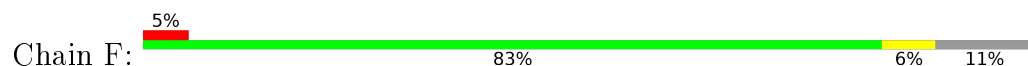
- Molecule 2: 5'-AMP-activated protein kinase subunit beta-1



- Molecule 3: 5'-AMP-activated protein kinase subunit gamma-1



- Molecule 3: 5'-AMP-activated protein kinase subunit gamma-1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.95Å 134.24Å 141.48Å 90.00° 93.02° 90.00°	Depositor
Resolution (Å)	43.78 – 2.99 48.66 – 2.99	Depositor EDS
% Data completeness (in resolution range)	98.6 (43.78-2.99) 98.6 (48.66-2.99)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.30 (at 3.01Å)	Xtriage
Refinement program	BUSTER-TNT	Depositor
R, $R_{free}$	0.225 , 0.243 0.245 , 0.264	Depositor DCC
$R_{free}$ test set	2867 reflections (5.34%)	DCC
Wilson B-factor (Å <sup>2</sup> )	62.6	Xtriage
Anisotropy	0.309	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 33.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	2 of 56573 reflections (0.004%)	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	14587	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: SEP, TPO, 4O7, C2Z, AMP, C1V

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	0.38	0/3316	0.55	0/4477
1	C	0.37	0/3637	0.56	0/4918
2	B	0.37	0/1335	0.55	0/1825
2	D	0.36	0/1400	0.55	0/1908
3	E	0.38	0/2437	0.56	0/3309
3	F	0.38	0/2414	0.56	0/3284
All	All	0.37	0/14539	0.55	0/19721

There are no bond length outliers.

There are no bond angle outliers.

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	3255	0	3224	9	0
1	C	3567	0	3505	16	1
2	B	1310	0	1239	12	0
2	D	1373	0	1342	12	1
3	E	2381	0	2443	17	0
3	F	2366	0	2400	11	0
4	A	35	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	35	0	0	0	0
5	B	26	0	12	0	0
5	C	26	0	12	0	0
6	E	30	0	0	0	0
7	F	46	0	24	0	0
8	A	28	0	0	0	0
8	B	13	0	0	0	0
8	C	33	0	0	1	0
8	D	16	0	0	0	0
8	E	26	0	0	0	0
8	F	21	0	0	0	0
All	All	14587	0	14201	67	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 2.

All (67) 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)
3:F:43:PRO:HG2	3:F:46:SER:HB3	1.76	0.66
1:A:35:HIS:HB3	1:A:39:GLY:H	1.70	0.56
1:C:138:ARG:HH12	1:C:172:TPO:HB	1.73	0.54
3:F:177:LYS:HA	3:F:180:ILE:HD12	1.90	0.54
3:E:61:PHE:HA	3:E:64:LEU:HD12	1.90	0.53
2:B:78:ARG:CB	2:B:79:PRO:HD3	2.39	0.52
1:A:441:ARG:HG3	1:A:452:LYS:HB3	1.91	0.52
1:C:171:ARG:HH22	2:D:204:PRO:HB3	1.74	0.52
3:E:108:GLU:O	3:E:112:HIS:HB2	2.09	0.52
1:A:82:VAL:HG13	2:B:162:VAL:HG21	1.92	0.51
2:B:92:TYR:HB2	2:B:128:PHE:HB3	1.91	0.51
2:D:207:PRO:HD2	2:D:210:LEU:HD12	1.91	0.51
1:C:402:TRP:HB2	2:D:213:VAL:HG11	1.95	0.49
3:F:56:GLN:HE21	3:F:58:LYS:HB3	1.78	0.48
1:C:454:SER:HB2	1:C:472:ILE:HD11	1.96	0.48
1:C:82:VAL:HG13	2:D:162:VAL:HG21	1.96	0.48
1:C:284:ASP:HB2	8:C:713:HOH:O	2.14	0.47
3:E:39:TYR:HA	3:E:42:ILE:HD12	1.95	0.47
2:B:264:THR:HG22	3:E:48:LEU:HD23	1.96	0.47
1:A:404:LEU:HD22	2:B:210:LEU:HB3	1.95	0.47
2:D:208:PRO:HA	2:D:211:LEU:HD12	1.96	0.47
1:C:52:ILE:HD12	1:C:61:ILE:HG13	1.96	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:173:LEU:HD22	3:F:315:LEU:HD22	1.96	0.47
1:C:244:LEU:HD11	1:C:262:ILE:HG23	1.96	0.47
3:E:142:VAL:HG22	3:E:172:ILE:HD13	1.98	0.46
3:F:39:TYR:HA	3:F:42:ILE:HD12	1.96	0.46
3:E:56:GLN:HA	3:E:113:LYS:HA	1.97	0.46
3:E:188:PHE:HB2	3:E:196:LEU:HD21	1.96	0.46
1:C:469:PHE:HE2	1:C:546:LEU:HD23	1.81	0.46
2:B:208:PRO:HA	2:B:211:LEU:HD12	1.98	0.45
3:E:74:LEU:HD22	3:E:114:ILE:HG21	1.97	0.45
2:D:84:TRP:HB3	2:D:112:PHE:HB2	1.98	0.45
1:A:244:LEU:HD11	1:A:262:ILE:HG23	1.98	0.44
1:A:171:ARG:HH22	2:B:204:PRO:HB3	1.82	0.44
3:F:177:LYS:HG2	3:F:293:VAL:HG21	1.99	0.44
1:C:404:LEU:HD22	2:D:210:LEU:HB3	2.00	0.44
3:E:152:ARG:HG2	3:E:168:THR:HG22	2.00	0.44
3:E:292:LEU:HD23	3:E:297:VAL:HG23	2.00	0.44
2:D:161:GLU:HB3	2:D:164:ASP:HB2	1.99	0.43
2:D:100:TRP:HZ3	2:D:126:LYS:HB2	1.84	0.43
1:C:12:LYS:HG2	1:C:17:VAL:HG22	1.99	0.43
2:D:215:LEU:HD21	2:D:265:LEU:HD11	2.00	0.43
1:C:138:ARG:NH1	1:C:172:TPO:HB	2.34	0.43
3:E:193:LEU:HD13	3:E:288:ILE:HD13	2.01	0.43
3:F:74:LEU:HD22	3:F:114:ILE:HG21	2.01	0.43
1:C:35:HIS:HB3	1:C:39:GLY:H	1.84	0.42
1:A:265:HIS:HD2	1:A:267:TRP:H	1.66	0.42
3:E:95:LEU:HD22	3:E:258:LEU:HD11	2.01	0.42
2:B:214:ILE:HG12	2:B:228:LEU:HD22	2.02	0.42
2:B:124:GLN:HB3	2:B:150:ASN:HD22	1.85	0.42
2:B:215:LEU:HD21	2:B:265:LEU:HD11	2.01	0.42
3:E:194:GLU:HG3	3:E:281:LEU:HD22	2.02	0.42
3:F:193:LEU:HD13	3:F:288:ILE:HD13	2.02	0.41
3:F:95:LEU:HD22	3:F:258:LEU:HD11	2.02	0.41
3:F:74:LEU:HD21	3:F:86:LEU:HB2	2.02	0.41
1:C:373:LEU:HD12	3:F:251:ALA:HA	2.01	0.41
1:A:47:LEU:HB2	1:A:89:PHE:HB2	2.03	0.41
3:E:304:ASP:HB3	3:E:308:VAL:HB	2.03	0.41
1:A:431:LYS:HG3	2:B:189:GLU:HG2	2.02	0.41
1:C:430:TRP:HB3	1:C:440:VAL:HG12	2.03	0.41
2:D:126:LYS:HE2	2:D:150:ASN:HB3	2.02	0.41
3:E:73:PRO:HD3	3:E:166:ILE:HD11	2.03	0.41
2:B:84:TRP:HB3	2:B:112:PHE:HB2	2.03	0.40

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:E:185:LYS:HA	3:E:186:PRO:HD3	1.98	0.40
1:C:63:ARG:HH21	1:C:163:MET:HB2	1.87	0.40
2:D:92:TYR:HB2	2:D:128:PHE:HB3	2.04	0.40
3:E:74:LEU:HD21	3:E:86:LEU:HB2	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:C:285:ALA:O	2:D:144:SER:OG[2_946]	2.17	0.03

## 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	400/565 (71%)	388 (97%)	11 (3%)	1 (0%)	46	84
1	C	442/565 (78%)	423 (96%)	18 (4%)	1 (0%)	52	88
2	B	165/270 (61%)	156 (94%)	9 (6%)	0	100	100
2	D	168/270 (62%)	161 (96%)	5 (3%)	2 (1%)	16	56
3	E	298/336 (89%)	291 (98%)	7 (2%)	0	100	100
3	F	298/336 (89%)	290 (97%)	8 (3%)	0	100	100
All	All	1771/2342 (76%)	1709 (96%)	58 (3%)	4 (0%)	52	88

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	461	ASP
1	C	316	PRO
2	D	258	LYS
2	D	194	PRO

### 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	344/497 (69%)	341 (99%)	3 (1%)	84	95
1	C	378/497 (76%)	371 (98%)	7 (2%)	65	90
2	B	135/239 (56%)	133 (98%)	2 (2%)	72	92
2	D	152/239 (64%)	150 (99%)	2 (1%)	76	93
3	E	265/308 (86%)	263 (99%)	2 (1%)	86	96
3	F	262/308 (85%)	259 (99%)	3 (1%)	80	94
All	All	1536/2088 (74%)	1517 (99%)	19 (1%)	78	94

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	545	SER
1	A	546	LEU
1	A	550	LEU
2	B	97	PHE
2	B	234	VAL
1	C	176	SER
1	C	237	LEU
1	C	239	ARG
1	C	339	GLU
1	C	457	LEU
1	C	532	LEU
1	C	542	MET
2	D	161	GLU
2	D	234	VAL
3	E	57	VAL
3	E	176	LEU
3	F	180	ILE
3	F	299	ARG
3	F	306	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
2	D	132	GLN
2	D	237	ASN
3	E	93	ASN
3	F	56	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

4 non-standard protein/DNA/RNA residues are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	TPO	A	172	1	7,10,11	0.70	0	10,14,16	1.58	3 (30%)
2	SEP	B	108	2	7,9,10	1.06	1 (14%)	8,12,14	2.37	2 (25%)
1	TPO	C	172	1	7,10,11	0.75	0	10,14,16	1.66	3 (30%)
2	SEP	D	108	2	7,9,10	1.01	0	8,12,14	1.80	2 (25%)

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
1	TPO	A	172	1	-	0/8/11/13	0/0/0/0
2	SEP	B	108	2	-	0/5/8/10	0/0/0/0
1	TPO	C	172	1	-	0/8/11/13	0/0/0/0
2	SEP	D	108	2	-	0/5/8/10	0/0/0/0

All (1) bond length outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	108	SEP	P-OG	-2.35	1.53	1.59

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	172	TPO	P-OG1-CB	-2.65	109.84	121.42
1	C	172	TPO	P-OG1-CB	-2.17	111.91	121.42
2	B	108	SEP	O-C-CA	-2.13	120.01	125.72
1	A	172	TPO	O-C-CA	-2.12	119.91	125.69
2	D	108	SEP	O-C-CA	-2.08	120.13	125.72
1	C	172	TPO	O-C-CA	-2.03	120.14	125.69
1	A	172	TPO	C-CA-N	2.30	115.02	109.95
1	C	172	TPO	C-CA-N	3.38	117.42	109.95
2	D	108	SEP	OG-CB-CA	3.81	111.58	108.26
2	B	108	SEP	OG-CB-CA	5.89	113.39	108.26

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	172	TPO	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

8 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	4O7	A	601	-	30,42,42	2.81	9 (30%)	22,68,68	1.95	8 (36%)
5	C1V	B	301	-	28,29,29	3.41	9 (32%)	29,42,42	7.48	6 (20%)
4	4O7	C	601	-	30,42,42	2.80	9 (30%)	22,68,68	1.94	8 (36%)
5	C1V	C	602	-	28,29,29	3.37	9 (32%)	29,42,42	7.64	6 (20%)
6	C2Z	E	401	-	8,16,16	3.73	5 (62%)	10,24,24	2.26	5 (50%)
6	C2Z	E	402	-	8,16,16	3.70	5 (62%)	10,24,24	2.05	4 (40%)
7	AMP	F	401	-	22,25,25	0.64	0	22,38,38	0.55	0
7	AMP	F	402	-	22,25,25	0.51	0	22,38,38	0.43	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
4	4O7	A	601	-	-	0/4/58/58	0/0/8/8
5	C1V	B	301	-	-	0/9/10/10	0/4/4/4
4	4O7	C	601	-	-	0/4/58/58	0/0/8/8
5	C1V	C	602	-	-	0/9/10/10	0/4/4/4
6	C2Z	E	401	-	-	0/0/10/10	0/0/2/2
6	C2Z	E	402	-	-	0/0/10/10	0/0/2/2
7	AMP	F	401	-	-	0/6/26/26	0/3/3/3
7	AMP	F	402	-	-	0/6/26/26	0/3/3/3

All (46) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	E	401	C2Z	CAM-CAL	-6.56	1.34	1.49
4	A	601	4O7	C8-N1	-6.53	1.35	1.45
4	C	601	4O7	C8-N1	-6.52	1.35	1.45
6	E	402	C2Z	CAM-CAL	-6.52	1.34	1.49
4	C	601	4O7	C10-C7	-5.20	1.48	1.55
4	A	601	4O7	C10-C7	-5.20	1.48	1.55
6	E	401	C2Z	CAG-CAL	-4.45	1.33	1.40
6	E	402	C2Z	CAG-CAL	-4.33	1.33	1.40
4	C	601	4O7	C9-C10	-2.14	1.45	1.53
4	A	601	4O7	C9-C10	-2.13	1.45	1.53
5	B	301	C1V	OAC-CAQ	2.24	1.41	1.36
5	C	602	C1V	OAC-CAQ	2.45	1.41	1.36
5	B	301	C1V	CAV-CAR	2.52	1.53	1.49
5	C	602	C1V	CAV-CAR	2.52	1.53	1.49

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	601	4O7	C19-C6	2.82	1.48	1.41
4	A	601	4O7	C19-C6	2.84	1.48	1.41
6	E	402	C2Z	PAO-OAD	3.02	1.60	1.54
4	C	601	4O7	C18-C11	3.24	1.49	1.41
4	A	601	4O7	C18-C11	3.26	1.49	1.41
6	E	401	C2Z	PAO-OAB	3.26	1.61	1.54
6	E	402	C2Z	PAO-OAB	3.29	1.61	1.54
6	E	401	C2Z	PAO-OAD	3.38	1.61	1.54
5	C	602	C1V	CAW-CAS	3.45	1.56	1.49
5	B	301	C1V	CAW-CAS	3.65	1.56	1.49
6	E	401	C2Z	PAO-OAC	4.82	1.61	1.50
5	C	602	C1V	CAW-CAZ	4.97	1.49	1.41
5	B	301	C1V	CAW-CAZ	4.97	1.49	1.41
6	E	402	C2Z	PAO-OAC	5.04	1.61	1.50
4	C	601	4O7	C5-C20	5.10	1.48	1.41
4	A	601	4O7	C5-C20	5.13	1.48	1.41
4	C	601	4O7	C6-C5	5.19	1.46	1.40
4	A	601	4O7	C6-C5	5.22	1.46	1.40
4	A	601	4O7	C12-C17	5.22	1.49	1.41
4	C	601	4O7	C12-C17	5.23	1.49	1.41
4	C	601	4O7	C11-C12	5.67	1.46	1.40
4	A	601	4O7	C11-C12	5.73	1.46	1.40
5	C	602	C1V	CAU-CAT	5.89	1.43	1.38
5	B	301	C1V	CAU-CAT	5.99	1.43	1.38
5	B	301	C1V	OAB-CAX	6.92	1.42	1.24
5	C	602	C1V	OAB-CAX	6.93	1.42	1.24
5	B	301	C1V	CAE-NAA	7.49	1.32	1.14
5	C	602	C1V	CAE-NAA	7.51	1.32	1.14
5	B	301	C1V	CAU-CAE	7.51	1.53	1.44
5	C	602	C1V	CAU-CAE	7.56	1.53	1.44
5	C	602	C1V	CAN-CAW	7.70	1.41	1.37
5	B	301	C1V	CAN-CAW	8.08	1.41	1.37

All (37) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	301	C1V	CAU-CAE-NAA	-34.85	120.02	177.37
5	C	602	C1V	CAU-CAE-NAA	-34.72	120.24	177.37
5	C	602	C1V	CAW-CAN-SAP	-19.52	107.08	112.53
5	B	301	C1V	CAW-CAN-SAP	-17.44	107.66	112.53
5	B	301	C1V	CAU-CAX-NAO	-4.57	120.64	124.15
5	C	602	C1V	CAU-CAX-NAO	-4.42	120.75	124.15

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	301	C1V	CAN-CAW-CAS	-4.39	118.87	125.69
5	C	602	C1V	CAN-CAW-CAS	-3.95	119.56	125.69
6	E	401	C2Z	CAG-CAL-CAM	-3.32	124.17	129.24
4	A	601	4O7	C16-C17-C12	-3.03	116.78	120.58
4	C	601	4O7	C16-C17-C12	-3.02	116.79	120.58
5	C	602	C1V	CAU-CAT-CAZ	-2.93	118.76	121.10
5	B	301	C1V	CAU-CAT-CAZ	-2.92	118.78	121.10
4	C	601	4O7	C1-C20-C5	-2.77	117.11	120.58
4	A	601	4O7	C1-C20-C5	-2.73	117.16	120.58
6	E	401	C2Z	OAD-PAO-OAC	-2.51	105.38	112.09
6	E	401	C2Z	OAB-PAO-OAC	-2.47	105.49	112.09
4	A	601	4O7	C3-C4-C5	-2.44	117.50	120.88
4	C	601	4O7	C3-C4-C5	-2.43	117.51	120.88
6	E	402	C2Z	OAD-PAO-OAC	-2.40	105.65	112.09
6	E	402	C2Z	OAB-PAO-OAC	-2.20	106.20	112.09
4	A	601	4O7	C14-C13-C12	-2.07	118.01	120.88
4	C	601	4O7	C14-C13-C12	-2.01	118.09	120.88
4	C	601	4O7	C4-C5-C20	2.28	122.13	119.83
4	A	601	4O7	C4-C5-C20	2.29	122.13	119.83
4	C	601	4O7	C1-C20-N3	2.51	135.21	132.18
4	A	601	4O7	C1-C20-N3	2.54	135.25	132.18
4	C	601	4O7	C9-C10-C11	2.86	126.15	116.59
4	A	601	4O7	C9-C10-C11	2.96	126.50	116.59
6	E	401	C2Z	OAD-PAO-CAN	3.11	112.48	106.73
6	E	402	C2Z	OAB-PAO-CAN	3.31	112.84	106.73
6	E	401	C2Z	OAB-PAO-CAN	3.47	113.12	106.73
6	E	402	C2Z	OAD-PAO-CAN	3.64	113.45	106.73
4	C	601	4O7	C16-C17-N2	4.95	138.28	132.22
4	A	601	4O7	C16-C17-N2	5.03	138.36	132.22
5	B	301	C1V	CAN-SAP-CAY	6.67	99.06	91.02
5	C	602	C1V	CAN-SAP-CAY	7.06	99.53	91.02

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 ⓘ

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	408/565 (72%)	0.44	10 (2%) 61 30	42, 69, 110, 163	0
1	C	452/565 (80%)	0.44	19 (4%) 40 16	33, 65, 108, 167	0
2	B	171/270 (63%)	0.74	18 (10%) 8 3	41, 85, 113, 142	0
2	D	174/270 (64%)	0.44	10 (5%) 27 10	43, 62, 103, 156	0
3	E	298/336 (88%)	0.43	17 (5%) 27 10	37, 62, 94, 113	0
3	F	300/336 (89%)	0.41	16 (5%) 30 12	36, 61, 93, 108	0
All	All	1803/2342 (76%)	0.46	90 (4%) 32 13	33, 66, 106, 167	0

All (90) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	319	GLN	6.3
3	E	306	ASN	5.9
3	F	304	ASP	5.6
3	E	272	TYR	4.4
2	D	194	PRO	4.0
3	E	123	GLN	4.0
1	C	413	TYR	3.8
3	F	303	VAL	3.7
2	D	175	ASP	3.7
2	B	94	SER	3.7
3	E	307	ASP	3.7
3	E	124	ASP	3.7
3	F	124	ASP	3.6
1	A	463	ARG	3.6
3	F	281	LEU	3.5
3	F	306	ASN	3.5
1	C	109	GLY	3.4
3	F	125	SER	3.3
2	D	174	SER	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	160	PHE	3.0
1	C	397	VAL	3.0
2	D	158	THR	3.0
3	E	303	VAL	2.9
2	D	195	GLU	2.9
1	A	293	VAL	2.8
2	B	116	LEU	2.8
1	A	298	GLU	2.8
3	E	126[A]	PHE	2.8
2	B	120	GLU	2.7
1	C	220	PRO	2.7
1	C	430	TRP	2.7
2	B	157	LYS	2.7
3	E	304	ASP	2.7
2	B	219	THR	2.6
3	E	302	VAL	2.6
1	A	399	LYS	2.6
2	B	164	ASP	2.6
2	B	154	GLN	2.6
3	E	282	HIS	2.6
3	F	307	ASP	2.6
1	C	218	HIS	2.6
3	E	308	VAL	2.6
1	C	313	SER	2.6
2	D	168	VAL	2.5
3	E	234	LYS	2.5
1	A	215	ASP	2.5
2	B	121	GLY	2.5
2	B	202	ALA	2.5
1	C	372	PRO	2.5
2	B	119	PRO	2.5
3	E	271	HIS	2.4
1	C	447	THR	2.4
1	C	461	ASP	2.4
2	B	143	THR	2.4
3	F	280	TYR	2.4
3	F	236	ARG	2.4
2	D	201	ARG	2.4
3	F	305	GLU	2.4
1	C	450	TYR	2.4
3	F	26	ASN	2.4
2	B	80	THR	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	356	ALA	2.3
1	C	463	ARG	2.3
2	D	192	CYS	2.3
3	F	282	HIS	2.3
3	E	280	TYR	2.3
3	F	270	SER	2.3
3	F	191	LYS	2.3
1	C	396	ALA	2.3
2	B	158	THR	2.2
2	B	153	ILE	2.2
3	E	187	GLU	2.2
1	C	362	GLY	2.2
2	B	118	LEU	2.2
1	C	361	PRO	2.2
2	D	163	PHE	2.2
1	A	532	LEU	2.2
1	A	400	ALA	2.2
1	C	106	CYS	2.2
2	D	171	GLN	2.2
3	F	271	HIS	2.1
3	E	119	GLU	2.1
3	F	192	SER	2.1
1	A	217	GLU	2.1
1	A	344	SER	2.0
2	B	117	ASP	2.0
1	A	221	THR	2.0
3	E	305	GLU	2.0
2	B	150	ASN	2.0
1	C	318	ASP	2.0

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

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
2	SEP	D	108	10/11	0.93	0.15	-	60,63,67,67	0
2	SEP	B	108	10/11	0.91	0.19	-	76,78,85,86	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	TPO	C	172	11/12	0.89	0.20	-	70,71,74,76	4
1	TPO	A	172	11/12	0.86	0.20	-	60,62,65,66	4

### 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( $\text{\AA}^2$ )	Q<0.9
6	C2Z	E	402	15/15	0.86	0.26	0.89	95,95,95,95	0
6	C2Z	E	401	15/15	0.88	0.23	0.24	88,88,89,89	0
5	C1V	B	301	26/26	0.89	0.23	0.14	58,61,67,69	0
4	4O7	C	601	35/35	0.96	0.18	-0.47	33,35,36,37	0
4	4O7	A	601	35/35	0.96	0.18	-0.52	37,39,41,41	0
5	C1V	C	602	26/26	0.97	0.20	-0.56	39,42,43,45	0
7	AMP	F	401	23/23	0.95	0.17	-0.88	63,68,71,72	0
7	AMP	F	402	23/23	0.94	0.19	-0.91	60,64,68,69	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.