Full wwPDB X-ray Structure Validation Report (i)

Sep 23, 2023 – 03:46 PM EDT

PDB ID	:	5SXF
Title	:	Crystal Structure of PI3Kalpha in complex with fragment 9
Authors	:	Gabelli, S.B.; Vogelstein, B.; Miller, M.S.; Amzel, L.M.
Deposited on	:	2016-08-09
Resolution	:	3.46 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 (i) 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 (i)) were used in the production of this report:

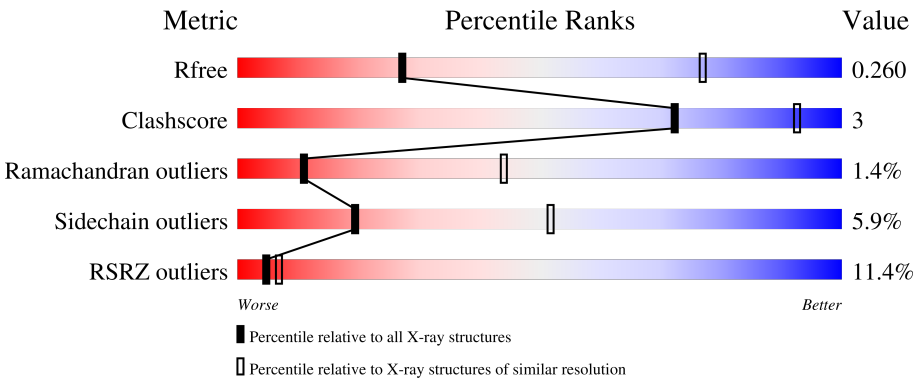
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.46 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}	130704	1291 (3.52-3.40)
Clashscore	141614	1372 (3.52-3.40)
Ramachandran outliers	138981	1337 (3.52-3.40)
Sidechain outliers	138945	1338 (3.52-3.40)
RSRZ outliers	127900	1205 (3.52-3.40)

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 <=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	1096	8% 78% 14% • 7%
2	B	279	18% 58% 8% 34%

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	HPP	A	1101	-	-	-	X

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9982 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 Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	1024	8389	5367	1436	1516	1	69	0	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-27	MET	-	initiating methionine	UNP P42336
A	-26	SER	-	expression tag	UNP P42336
A	-25	TYR	-	expression tag	UNP P42336
A	-24	TYR	-	expression tag	UNP P42336
A	-23	HIS	-	expression tag	UNP P42336
A	-22	HIS	-	expression tag	UNP P42336
A	-21	HIS	-	expression tag	UNP P42336
A	-20	HIS	-	expression tag	UNP P42336
A	-19	HIS	-	expression tag	UNP P42336
A	-18	HIS	-	expression tag	UNP P42336
A	-17	ASP	-	expression tag	UNP P42336
A	-16	TYR	-	expression tag	UNP P42336
A	-15	ASP	-	expression tag	UNP P42336
A	-14	ILE	-	expression tag	UNP P42336
A	-13	PRO	-	expression tag	UNP P42336
A	-12	THR	-	expression tag	UNP P42336
A	-11	THR	-	expression tag	UNP P42336
A	-10	GLU	-	expression tag	UNP P42336
A	-9	ASN	-	expression tag	UNP P42336
A	-8	LEU	-	expression tag	UNP P42336
A	-7	TYR	-	expression tag	UNP P42336
A	-6	PHE	-	expression tag	UNP P42336
A	-5	GLN	-	expression tag	UNP P42336
A	-4	GLY	-	expression tag	UNP P42336
A	-3	ALA	-	expression tag	UNP P42336
A	-2	MET	-	expression tag	UNP P42336

Continued on next page...

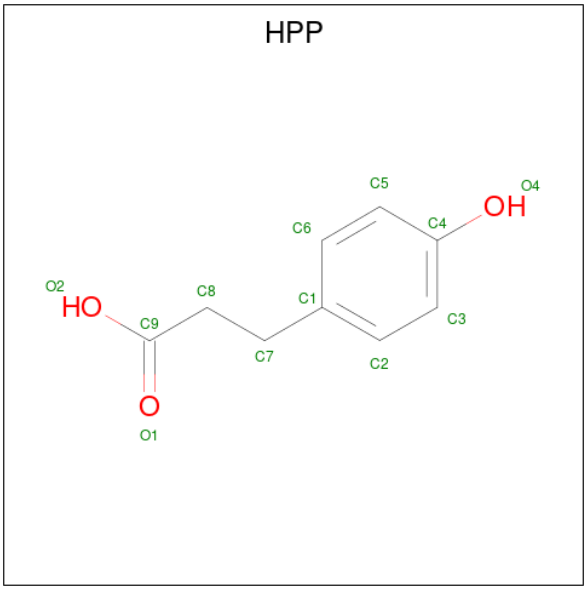
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP P42336
A	0	SER	-	expression tag	UNP P42336

- Molecule 2 is a protein called Phosphatidylinositol 3-kinase regulatory subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	183	1581	982	287	307	5	0	0	0

- Molecule 3 is HYDROXYPHENYL PROPIONIC ACID (three-letter code: HPP) (formula: $C_9H_{10}O_3$).



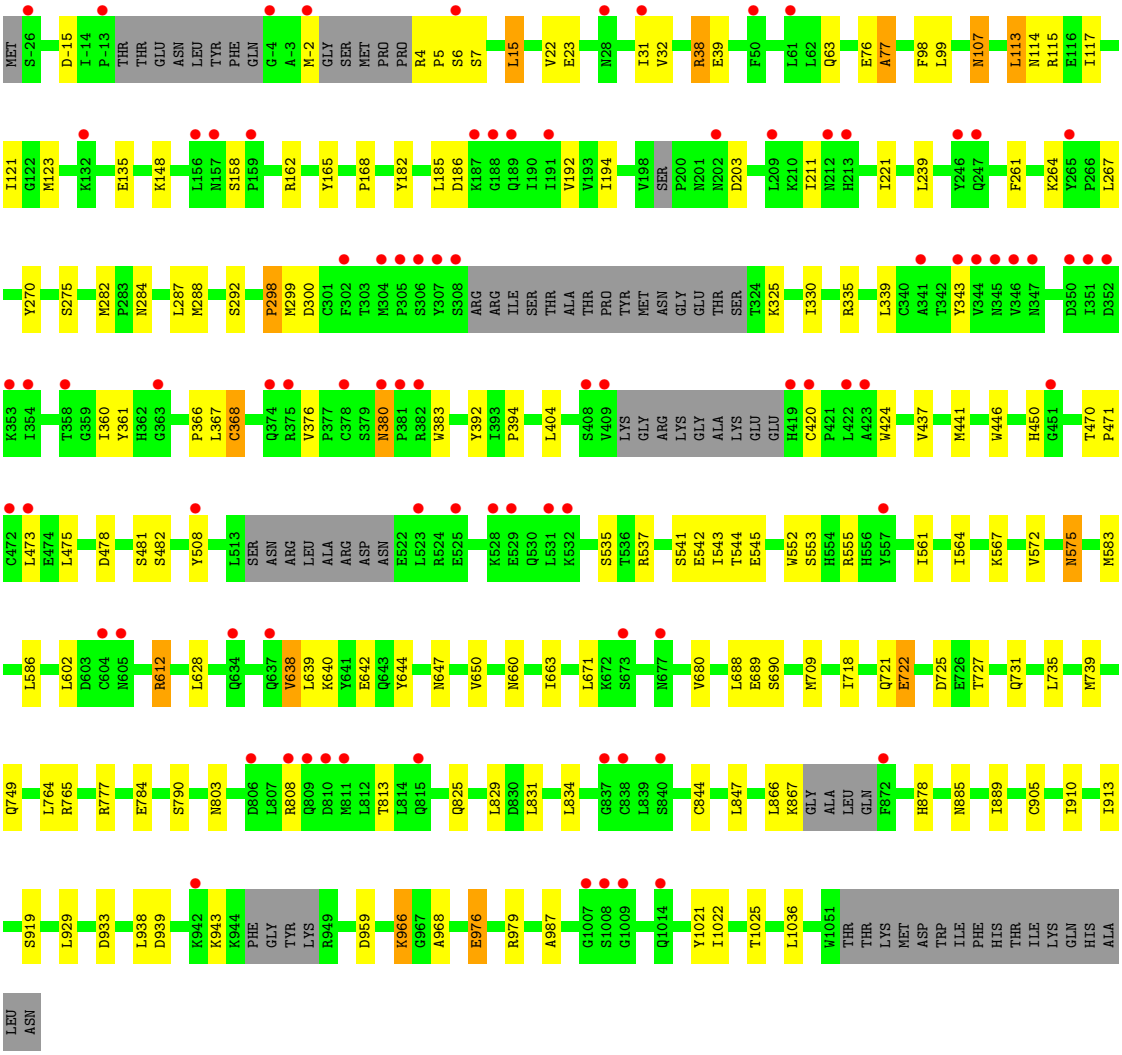
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	12	9	3	0	0

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.

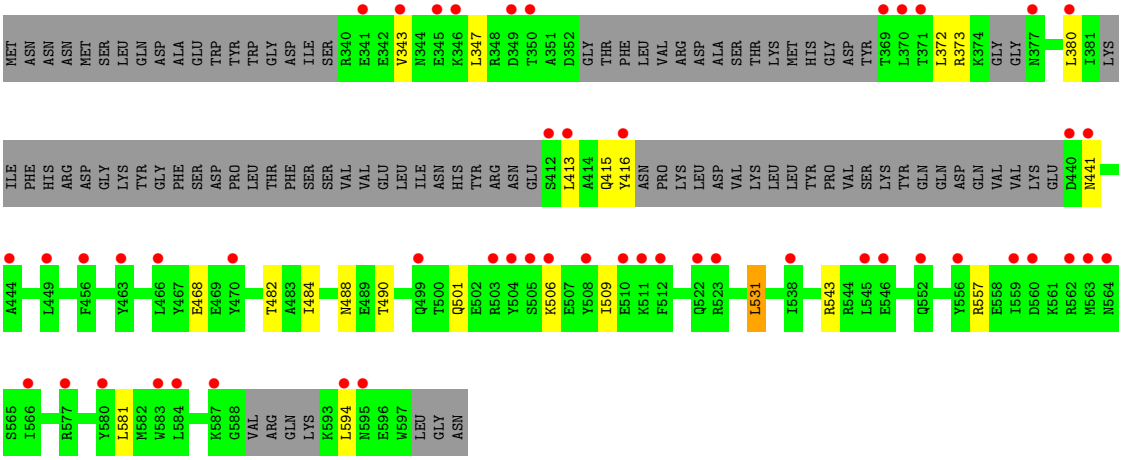
• Molecule 1: Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform

Chain A: 8% 78% 14% • 7%



• Molecule 2: Phosphatidylinositol 3-kinase regulatory subunit alpha

Chain B: 18% 58% 8% 34%


MET ASN ASN ASN MET SER LEU GLN ASP ALA GLU TRP TYR TRP GLY ASP ILE SER R340 E341 E342 V343 N344 E345 K346 L347 R348 D349 T350 A351 D352 GLY THR PHE LEU VAL ARG ASP ALA SER THR LYS MET HIS GLY ASP TYR T369 L370 T371 L372 R373 K374 GLY GLY N377 L380 I381 LYS
ILE PHE HIS ARG ASP GLY LYS TYR GLY PHE SER ASP PRO LEU THR PHE SER SER VAL VAL GLU LEU ILE ASN HIS TYR ARG ASN GLU S412 L413 A414 Q415 Y416 ASN PRO LYS LEU ASP VAL LYS LEU LEU TYR PRO VAL SER LYS TYR GLN GLN ASP GLN VAL VAL LYS GLU D440 N441
A444 L449 F456 Y463 L466 Y467 E468 E469 Y470 T482 A483 I484 N488 E489 T490 Q499 T500 Q501 E502 R503 Y504 S505 K506 E507 Y508 I509 E510 K511 F512 Q522 R523 L531 I538 R543 R544 L545 E546 Q552 Y556 R557 E558 I559 D560 K561 R562 M563 N564
S565 I566 R577 Y580 L581 M582 W583 L584 K587 G588 VAL ARG GLN LYS K593 L594 N595 E596 W597 LEU GLY ASN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	112.35Å 115.92Å 147.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	91.07 – 3.46 48.62 – 3.46	Depositor EDS
% Data completeness (in resolution range)	98.8 (91.07-3.46) 98.9 (48.62-3.46)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.38 (at 3.48Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.195 , 0.267 0.197 , 0.260	Depositor DCC
R_{free} test set	1314 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å^2)	109.5	Xtriage
Anisotropy	0.374	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 92.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.022 for k,h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9982	wwPDB-VP
Average B, all atoms (Å^2)	129.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.33% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: HPP, SEP

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	0.53	0/8570	0.76	2/11575 (0.0%)
2	B	0.47	0/1596	0.68	0/2121
All	All	0.52	0/10166	0.75	2/13696 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	113	LEU	CB-CG-CD1	5.19	119.82	111.00
1	A	612	ARG	NE-CZ-NH1	5.07	122.84	120.30

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	8389	0	8368	61	0
2	B	1581	0	1578	13	0
3	A	12	0	8	0	0
All	All	9982	0	9954	68	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 3.

All (68) 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:A:572:VAL:HG21	1:A:583:MET:HG2	1.76	0.68
1:A:98:PHE:CE1	2:B:490:THR:HG23	2.32	0.65
1:A:749:GLN:HE21	1:A:764:LEU:H	1.44	0.65
1:A:910:ILE:O	1:A:1025:THR:HG21	1.98	0.63
1:A:31:ILE:HD11	2:B:531:LEU:HD13	1.81	0.62
1:A:878:HIS:NE2	1:A:966:LYS:O	2.34	0.60
1:A:640:LYS:HE2	1:A:680:VAL:HG11	1.85	0.59
1:A:544:THR:HG22	2:B:380:LEU:HB3	1.85	0.58
1:A:709:MET:HE1	1:A:847:LEU:HD21	1.86	0.55
1:A:543:ILE:HD11	1:A:567:LYS:HD3	1.90	0.54
1:A:182:TYR:O	1:A:185:LEU:HB2	2.08	0.54
1:A:803:ASN:O	1:A:844:CYS:O	2.25	0.54
1:A:15:LEU:HB2	1:A:718:ILE:HD11	1.89	0.53
1:A:162:ARG:NH2	1:A:298:PRO:O	2.41	0.53
1:A:361:TYR:HA	1:A:366:PRO:HD3	1.91	0.52
1:A:537:ARG:HB3	1:A:541:SER:OG	2.09	0.52
2:B:372:LEU:HD23	2:B:373:ARG:N	2.24	0.52
1:A:561:ILE:O	1:A:564:ILE:HG22	2.11	0.51
2:B:506:LYS:HA	2:B:509:ILE:HG22	1.91	0.51
1:A:671:LEU:HB2	1:A:688:LEU:HD21	1.92	0.51
1:A:660:ASN:OD1	1:A:660:ASN:C	2.49	0.51
1:A:192:VAL:HG21	1:A:211:ILE:HD11	1.93	0.50
1:A:602:LEU:O	1:A:612:ARG:NH2	2.45	0.49
1:A:360:ILE:N	1:A:360:ILE:HD12	2.28	0.49
1:A:831:LEU:HD11	1:A:987:ALA:HB2	1.94	0.49
1:A:31:ILE:CD1	2:B:531:LEU:HD13	2.43	0.47
1:A:552:TRP:CZ3	1:A:555:ARG:HD3	2.49	0.47
2:B:581:LEU:HD11	2:B:594:LEU:HD23	1.98	0.46
1:A:114:ASN:HA	1:A:117:ILE:HD12	1.98	0.46
1:A:424:TRP:CE2	1:A:446:TRP:HB2	2.51	0.46
1:A:1022:ILE:HA	1:A:1025:THR:HG22	1.98	0.46
1:A:575:ASN:C	1:A:575:ASN:HD22	2.19	0.46
1:A:639:LEU:HD22	1:A:650:VAL:HG22	1.98	0.46
1:A:367:LEU:HD12	1:A:368:CYS:O	2.17	0.45
1:A:642:GLU:HG2	1:A:647:ASN:CG	2.37	0.45
1:A:555:ARG:HD2	1:A:586:LEU:HD11	1.99	0.45
1:A:866:LEU:HD23	1:A:867:LYS:N	2.31	0.44
1:A:392:TYR:HB3	1:A:394:PRO:HD2	1.99	0.44
1:A:121:ILE:HD11	1:A:689:GLU:HA	2.00	0.44
1:A:194:ILE:N	1:A:194:ILE:HD12	2.33	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:325:LYS:HE2	1:A:330:ILE:HD11	2.01	0.43
1:A:165:TYR:O	1:A:168:PRO:HD3	2.18	0.43
1:A:261:PHE:HA	1:A:270:TYR:CE2	2.54	0.43
1:A:938:LEU:HD22	1:A:1021:TYR:HB3	2.00	0.43
1:A:404:LEU:CD1	1:A:473:LEU:HD23	2.48	0.43
1:A:721:GLN:O	1:A:722:GLU:HB2	2.18	0.43
1:A:765:ARG:HD2	1:A:784:GLU:HG3	2.00	0.43
1:A:5:PRO:HB3	1:A:721:GLN:HE22	1.84	0.43
1:A:343:TYR:OH	2:B:557:ARG:NH1	2.48	0.43
1:A:628:LEU:HD23	1:A:663:ILE:HD13	2.00	0.42
1:A:535:SER:HA	1:A:564:ILE:HD12	2.00	0.42
1:A:612:ARG:NH1	1:A:642:GLU:OE1	2.52	0.42
1:A:135:GLU:OE2	1:A:644:TYR:HB3	2.20	0.42
1:A:76:GLU:O	1:A:77:ALA:HB3	2.20	0.42
1:A:721:GLN:O	1:A:722:GLU:CB	2.68	0.42
2:B:531:LEU:HD12	2:B:531:LEU:O	2.20	0.41
2:B:343:VAL:HG12	2:B:347:LEU:HD12	2.01	0.41
1:A:885:ASN:HB3	1:A:889:ILE:HG22	2.03	0.41
2:B:413:LEU:HD22	2:B:416:TYR:CD2	2.56	0.41
1:A:282:MET:O	1:A:284:ASN:ND2	2.52	0.41
1:A:376:VAL:HG11	1:A:380:ASN:HD22	1.83	0.41
1:A:339:LEU:O	1:A:383:TRP:O	2.39	0.41
1:A:602:LEU:HB3	1:A:638:VAL:HG11	2.02	0.41
1:A:731:GLN:OE1	1:A:777:ARG:NH2	2.51	0.41
1:A:976:GLU:N	1:A:976:GLU:OE2	2.54	0.41
1:A:542:GLU:HG2	2:B:380:LEU:HD13	2.01	0.40
1:A:221:ILE:HG23	1:A:287:LEU:CD2	2.52	0.40
2:B:484:ILE:HG22	2:B:488:ASN:ND2	2.36	0.40

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	1005/1096 (92%)	916 (91%)	72 (7%)	17 (2%)	9 40
2	B	171/279 (61%)	160 (94%)	11 (6%)	0	100 100
All	All	1176/1375 (86%)	1076 (92%)	83 (7%)	17 (1%)	11 44

All (17) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	380	ASN
1	A	722	GLU
1	A	77	ALA
1	A	107	ASN
1	A	264	LYS
1	A	939	ASP
1	A	968	ALA
1	A	38	ARG
1	A	471	PRO
1	A	508	TYR
1	A	186	ASP
1	A	298	PRO
1	A	450	HIS
1	A	933	ASP
1	A	368	CYS
1	A	725	ASP
1	A	913	ILE

5.3.2 Protein sidechains (i)

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	937/998 (94%)	879 (94%)	58 (6%)	18 50
2	B	173/259 (67%)	166 (96%)	7 (4%)	31 63
All	All	1110/1257 (88%)	1045 (94%)	65 (6%)	19 52

All (65) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	-15	ASP
1	A	-2	MET
1	A	4	ARG
1	A	6	SER
1	A	7	SER
1	A	15	LEU
1	A	22	VAL
1	A	23	GLU
1	A	32	VAL
1	A	38	ARG
1	A	39	GLU
1	A	63	GLN
1	A	99	LEU
1	A	107	ASN
1	A	113	LEU
1	A	115	ARG
1	A	123	MET
1	A	148	LYS
1	A	158	SER
1	A	203	ASP
1	A	239	LEU
1	A	267	LEU
1	A	275	SER
1	A	288	MET
1	A	292	SER
1	A	299	MET
1	A	300	ASP
1	A	335	ARG
1	A	420	CYS
1	A	437	VAL
1	A	441	MET
1	A	470	THR
1	A	475	LEU
1	A	478	ASP
1	A	481	SER
1	A	482	SER
1	A	545	GLU
1	A	553	SER
1	A	575	ASN
1	A	638	VAL
1	A	690	SER
1	A	727	THR
1	A	735	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	739	MET
1	A	808	ARG
1	A	813	THR
1	A	825	GLN
1	A	829	LEU
1	A	834	LEU
1	A	905	CYS
1	A	919	SER
1	A	929	LEU
1	A	943	LYS
1	A	959	ASP
1	A	966	LYS
1	A	976	GLU
1	A	979	ARG
1	A	1036	LEU
2	B	415	GLN
2	B	441	ASN
2	B	468	GLU
2	B	482	THR
2	B	501	GLN
2	B	531	LEU
2	B	543	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	14	HIS
1	A	107	ASN
1	A	247	GLN
1	A	331	ASN
1	A	374	GLN
1	A	380	ASN
1	A	384	ASN
1	A	575	ASN
1	A	721	GLN
1	A	749	GLN
1	A	763	ASN
1	A	861	GLN
1	A	931	HIS
1	A	1042	GLN
2	B	453	ASN
2	B	475	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	564	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)

1 non-standard protein/DNA/RNA residue is 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	SEP	A	790	1	8,9,10	0.60	0	8,12,14	1.73	3 (37%)

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	SEP	A	790	1	-	4/5/8/10	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	790	SEP	OG-CB-CA	2.40	110.48	108.14
1	A	790	SEP	P-OG-CB	2.34	124.74	118.30
1	A	790	SEP	OG-P-O1P	-2.10	100.58	106.47

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	790	SEP	N-CA-CB-OG
1	A	790	SEP	CB-OG-P-O1P
1	A	790	SEP	CB-OG-P-O2P
1	A	790	SEP	CB-OG-P-O3P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is 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	$\#\|Z\| > 2$
3	HPP	A	1101	-	12,12,12	0.80	0	15,15,15	0.62	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	HPP	A	1101	-	-	2/5/5/5	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1101	HPP	C7-C8-C9-O1
3	A	1101	HPP	C7-C8-C9-O2

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	1023/1096 (93%)	0.53	87 (8%) 10 13	68, 114, 176, 235	0
2	B	183/279 (65%)	1.37	51 (27%) 0 0	105, 177, 216, 242	0
All	All	1206/1375 (87%)	0.65	138 (11%) 5 7	68, 122, 196, 242	0

All (138) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	377	ASN	6.1
1	A	-26	SER	5.2
1	A	378	CYS	4.9
1	A	409	VAL	4.8
2	B	504	TYR	4.8
1	A	307	TYR	4.7
1	A	347	ASN	4.7
1	A	353	LYS	4.7
1	A	308	SER	4.6
1	A	346	VAL	4.5
2	B	594	LEU	4.4
2	B	456	PHE	4.3
2	B	510	GLU	4.3
2	B	346	LYS	4.0
1	A	31	ILE	3.8
2	B	349	ASP	3.8
2	B	369	THR	3.8
2	B	563	MET	3.7
2	B	559	ILE	3.7
2	B	350	THR	3.7
1	A	246	TYR	3.7
2	B	371	THR	3.6
1	A	408	SER	3.5
2	B	583	TRP	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	557	TYR	3.5
1	A	344	VAL	3.4
1	A	351	ILE	3.4
1	A	422	LEU	3.4
1	A	-13	PRO	3.4
1	A	341	ALA	3.4
1	A	532	LYS	3.4
1	A	673	SER	3.4
2	B	564	ASN	3.3
1	A	806	ASP	3.3
1	A	1009	GLY	3.3
2	B	556	TYR	3.3
1	A	531	LEU	3.3
1	A	382	ARG	3.2
1	A	375	ARG	3.2
1	A	156	LEU	3.2
1	A	381	PRO	3.2
2	B	341	GLU	3.1
1	A	345	ASN	3.1
1	A	343	TYR	3.1
1	A	872	PHE	3.1
1	A	350	ASP	3.1
1	A	1008	SER	3.1
2	B	440	ASP	3.1
1	A	305	PRO	3.0
2	B	538	ILE	3.0
2	B	580	TYR	3.0
2	B	416	TYR	3.0
1	A	473	LEU	3.0
1	A	354	ILE	3.0
1	A	187	LYS	3.0
2	B	560	ASP	3.0
1	A	809	GLN	2.9
2	B	584	LEU	2.9
1	A	423	ALA	2.9
1	A	380	ASN	2.9
2	B	511	LYS	2.8
1	A	352	ASP	2.8
2	B	441	ASN	2.8
1	A	189	GLN	2.8
2	B	463	TYR	2.8
1	A	159	PRO	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1007	GLY	2.8
1	A	529	GLU	2.7
1	A	363	GLY	2.7
1	A	838	CYS	2.7
2	B	470	TYR	2.7
2	B	508	TYR	2.7
1	A	811	MET	2.7
2	B	562	ARG	2.7
1	A	419	HIS	2.7
2	B	587	LYS	2.7
1	A	815	GLN	2.7
2	B	343	VAL	2.6
2	B	499	GLN	2.6
2	B	449	LEU	2.6
1	A	508	TYR	2.6
2	B	370	LEU	2.6
1	A	132	LYS	2.5
1	A	528	LYS	2.5
1	A	810	ASP	2.5
2	B	523	ARG	2.5
1	A	304	MET	2.5
2	B	444	ALA	2.5
1	A	604	CYS	2.5
1	A	213	HIS	2.4
1	A	942	LYS	2.4
2	B	566	ILE	2.4
1	A	-2	MET	2.4
1	A	840	SER	2.4
2	B	345	GLU	2.4
1	A	212	ASN	2.4
1	A	306	SER	2.4
1	A	28	ASN	2.4
1	A	157	ASN	2.4
1	A	265	TYR	2.4
2	B	503	ARG	2.4
1	A	209	LEU	2.4
2	B	413	LEU	2.4
1	A	451	GLY	2.3
1	A	202	ASN	2.3
1	A	302	PHE	2.3
1	A	523	LEU	2.3
1	A	191	ILE	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	420	CYS	2.3
1	A	61	LEU	2.3
1	A	374	GLN	2.3
2	B	505	SER	2.3
2	B	522	GLN	2.3
1	A	-4	GLY	2.3
2	B	412	SER	2.3
2	B	545	LEU	2.2
1	A	525	GLU	2.2
2	B	380	LEU	2.2
1	A	472	CYS	2.2
1	A	247	GLN	2.2
1	A	634	GLN	2.2
1	A	358	THR	2.2
1	A	605	ASN	2.2
1	A	837	GLY	2.2
2	B	512	PHE	2.1
1	A	1014	GLN	2.1
1	A	677	ASN	2.1
1	A	188	GLY	2.1
1	A	808	ARG	2.1
2	B	466	LEU	2.1
2	B	506	LYS	2.1
2	B	577	ARG	2.1
1	A	6	SER	2.1
2	B	546	GLU	2.0
2	B	552	GLN	2.0
2	B	595	ASN	2.0
1	A	50	PHE	2.0
1	A	637	GLN	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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. 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
1	SEP	A	790	10/11	0.95	0.14	92,110,175,176	0

6.3 Carbohydrates (i)

There are no monosaccharides 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. 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	HPP	A	1101	12/12	0.51	0.44	152,176,190,191	0

6.5 Other polymers (i)

There are no such residues in this entry.