



# Full wwPDB X-ray Structure Validation Report ⓘ

May 24, 2020 – 03:25 am BST

PDB ID : 4KHB  
Title : Structure of the Spt16D Pob3N heterodimer  
Authors : Stuwe, T.; Zhang, E.; Ladurner, A.G.  
Deposited on : 2013-04-30  
Resolution : 2.40 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

<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) : 1.13  
EDS : 2.11  
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) : 2.11

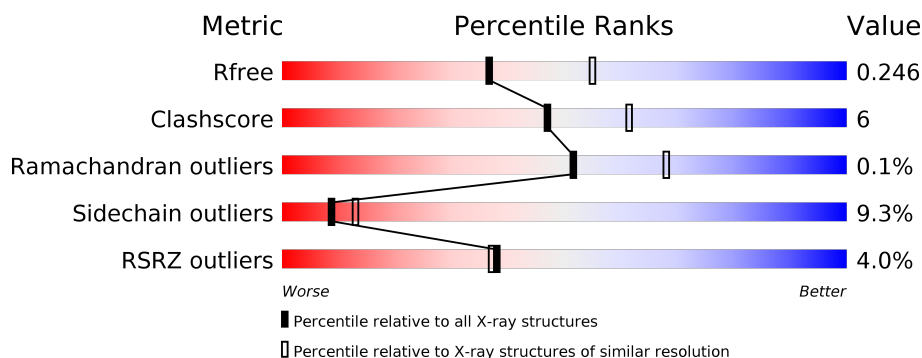
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	127	<div> <div>6%</div> <div> <div></div> <div>67%</div> <div>16%</div> <div>6%</div> <div>11%</div> </div> </div>
1	C	127	<div> <div>3%</div> <div> <div></div> <div>61%</div> <div>18%</div> <div>•</div> <div>17%</div> </div> </div>
1	E	127	<div> <div>3%</div> <div> <div></div> <div>69%</div> <div>9%</div> <div>5%</div> <div>17%</div> </div> </div>
1	G	127	<div> <div>4%</div> <div> <div></div> <div>63%</div> <div>17%</div> <div>•</div> <div>17%</div> </div> </div>
2	B	195	<div> <div>5%</div> <div> <div></div> <div>76%</div> <div>10%</div> <div>•</div> <div>10%</div> </div> </div>
2	D	195	<div> <div></div> <div> <div></div> <div>72%</div> <div>13%</div> <div>•</div> <div>13%</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	195	
2	H	195	

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	SO4	H	201	-	-	X	-

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 9291 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 Uncharacterized protein SPT16D.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	113	Total	C	N	O	S	0	0	0
			920	577	170	171	2			
1	G	106	Total	C	N	O	S	0	0	0
			859	541	157	159	2			
1	E	106	Total	C	N	O	S	0	0	0
			860	544	157	157	2			
1	C	106	Total	C	N	O	S	0	0	0
			865	547	157	159	2			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	521	MET	-	INITIATING METHIONINE	UNP G0SDN1
G	521	MET	-	INITIATING METHIONINE	UNP G0SDN1
E	521	MET	-	INITIATING METHIONINE	UNP G0SDN1
C	521	MET	-	INITIATING METHIONINE	UNP G0SDN1

- Molecule 2 is a protein called Uncharacterized protein POB3N.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	170	Total	C	N	O	S	0	0	0
			1357	865	235	255	2			
2	F	173	Total	C	N	O	S	0	0	0
			1378	876	239	261	2			
2	H	168	Total	C	N	O	S	0	0	0
			1345	858	233	252	2			
2	B	176	Total	C	N	O	S	0	0	0
			1397	888	243	264	2			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-2	GLY	-	EXPRESSION TAG	UNP G0SHK5
D	-1	SER	-	EXPRESSION TAG	UNP G0SHK5
D	0	HIS	-	EXPRESSION TAG	UNP G0SHK5
F	-2	GLY	-	EXPRESSION TAG	UNP G0SHK5
F	-1	SER	-	EXPRESSION TAG	UNP G0SHK5
F	0	HIS	-	EXPRESSION TAG	UNP G0SHK5
H	-2	GLY	-	EXPRESSION TAG	UNP G0SHK5
H	-1	SER	-	EXPRESSION TAG	UNP G0SHK5
H	0	HIS	-	EXPRESSION TAG	UNP G0SHK5
B	-2	GLY	-	EXPRESSION TAG	UNP G0SHK5
B	-1	SER	-	EXPRESSION TAG	UNP G0SHK5
B	0	HIS	-	EXPRESSION TAG	UNP G0SHK5

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		
3	H	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

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

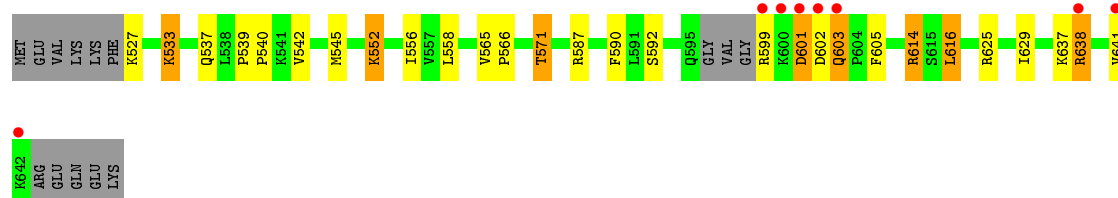
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	22	Total	O	0	0
			22	22		
4	D	51	Total	O	0	0
			51	51		
4	F	43	Total	O	0	0
			43	43		
4	H	50	Total	O	0	0
			50	50		
4	G	11	Total	O	0	0
			11	11		
4	E	16	Total	O	0	0
			16	16		
4	C	25	Total	O	0	0
			25	25		
4	B	47	Total	O	0	0
			47	47		

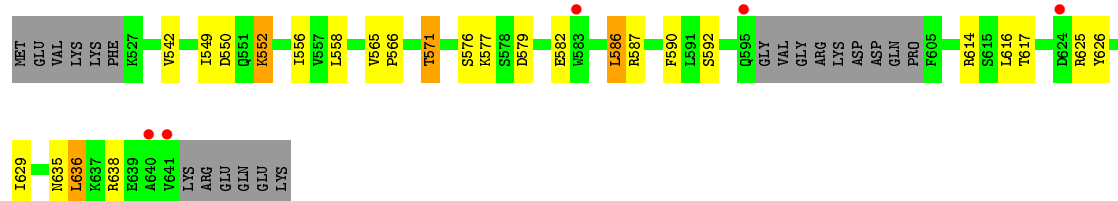
### 3 Residue-property plots [i](#)

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 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.

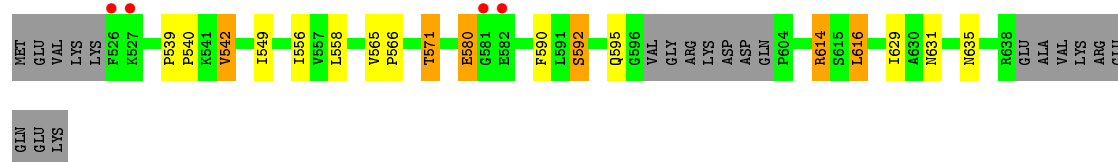
#### • Molecule 1: Uncharacterized protein SPT16D



#### • Molecule 1: Uncharacterized protein SPT16D

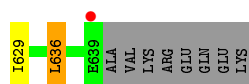


#### • Molecule 1: Uncharacterized protein SPT16D



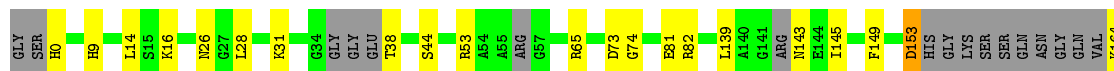
#### • Molecule 1: Uncharacterized protein SPT16D





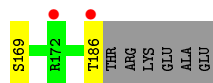
• Molecule 2: Uncharacterized protein POB3N

Chain D: 72% 13% 13%



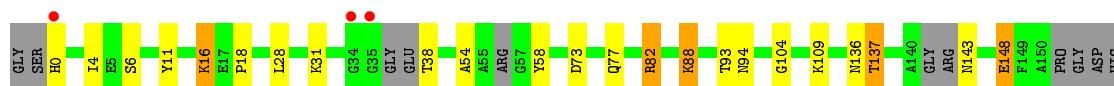
• Molecule 2: Uncharacterized protein POB3N

Chain F: 5% 78% 8% 11%



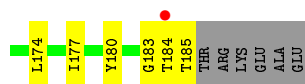
• Molecule 2: Uncharacterized protein POB3N

Chain H: 3% 69% 14% 14%



• Molecule 2: Uncharacterized protein POB3N

Chain B: 5% 76% 10% 10%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.39 Å 128.08 Å 132.33 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.78 – 2.40 47.78 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.78-2.40) 95.4 (47.78-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.27 (at 2.39 Å)	Xtriage
Refinement program	PHENIX 1.8_1069	Depositor
R, $R_{free}$	0.202 , 0.243 0.208 , 0.246	Depositor DCC
$R_{free}$ test set	2873 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.6	Xtriage
Anisotropy	0.090	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 45.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.015 for -h,l,k	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9291	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.92% 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, 0.333 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: SO4

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	0.35	0/938	0.50	0/1261
1	C	0.27	0/884	0.50	0/1190
1	E	0.32	0/879	0.48	0/1182
1	G	0.26	0/876	0.50	0/1179
2	B	0.43	2/1428 (0.1%)	0.55	3/1925 (0.2%)
2	D	0.41	1/1387 (0.1%)	0.51	1/1867 (0.1%)
2	F	0.52	1/1408 (0.1%)	0.63	4/1896 (0.2%)
2	H	0.33	0/1374	0.51	2/1849 (0.1%)
All	All	0.38	4/9174 (0.0%)	0.53	10/12349 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	F	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	58	TYR	CD1-CE1	-6.65	1.29	1.39
2	B	58	TYR	CD2-CE2	-6.03	1.30	1.39
2	F	74	GLY	C-O	-5.78	1.14	1.23
2	D	183	PRO	N-CD	5.26	1.55	1.47

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	73	ASP	CB-CA-C	12.29	134.99	110.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	54	ALA	CB-CA-C	-5.92	101.23	110.10
2	B	143	ASN	N-CA-C	-5.74	95.52	111.00
2	B	28	LEU	CA-CB-CG	5.64	128.28	115.30
2	B	183	GLY	N-CA-C	5.63	127.18	113.10
2	H	165	SER	N-CA-C	-5.52	96.08	111.00
2	F	57	GLY	N-CA-C	-5.43	99.52	113.10
2	F	73	ASP	CB-CG-OD1	5.38	123.14	118.30
2	F	28	LEU	CA-CB-CG	5.32	127.53	115.30
2	D	182	ILE	C-N-CD	5.26	139.45	128.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	F	73	ASP	Peptide

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of 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	920	0	914	18	0
1	C	865	0	854	15	0
1	E	860	0	852	17	0
1	G	859	0	852	14	0
2	B	1397	0	1354	16	0
2	D	1357	0	1307	17	0
2	F	1378	0	1324	5	0
2	H	1345	0	1300	24	0
3	A	5	0	0	0	0
3	B	10	0	0	1	0
3	C	5	0	0	0	0
3	D	10	0	0	0	0
3	E	5	0	0	0	0
3	F	5	0	0	0	0
3	H	5	0	0	3	0
4	A	22	0	0	0	0
4	B	47	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	25	0	0	0	0
4	D	51	0	0	1	0
4	E	16	0	0	0	0
4	F	43	0	0	0	0
4	G	11	0	0	0	0
4	H	50	0	0	2	0
All	All	9291	0	8757	110	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 6.

All (110) 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:614:ARG:HG3	1:E:614:ARG:HH11	0.97	1.10
1:E:614:ARG:HH11	1:E:614:ARG:CG	1.66	1.06
1:E:614:ARG:NH1	1:E:614:ARG:HG3	1.61	1.04
1:A:601:ASP:HB3	1:A:602:ASP:HB3	1.65	0.77
2:H:58:TYR:CE1	2:H:77:GLN:HB2	2.24	0.73
1:A:614:ARG:HD2	2:B:180:TYR:HB2	1.71	0.71
1:C:577:LYS:NZ	1:C:623:GLY:O	2.23	0.71
2:H:136:ASN:ND2	2:H:148:GLU:OE2	2.24	0.70
1:E:558:LEU:HD13	1:E:616:LEU:HD11	1.73	0.68
2:H:0:HIS:N	3:H:201:SO4:O4	2.27	0.67
1:C:593:PRO:O	1:C:614:ARG:NH2	2.27	0.67
2:F:34:GLY:O	2:F:36:GLY:N	2.28	0.67
2:D:153:ASP:N	2:D:153:ASP:OD1	2.27	0.67
2:H:58:TYR:HE1	2:H:77:GLN:HB2	1.60	0.64
2:F:88:LYS:NZ	2:F:94:ASN:OD1	2.30	0.64
2:D:0:HIS:O	2:B:142:ARG:NH2	2.32	0.63
2:F:31:LYS:HB2	2:F:38:THR:HG22	1.81	0.62
2:H:88:LYS:NZ	2:H:94:ASN:OD1	2.30	0.62
1:E:614:ARG:NH1	1:E:614:ARG:CG	2.36	0.61
2:B:22:ARG:NH2	4:B:341:HOH:O	2.29	0.60
2:H:0:HIS:N	3:H:201:SO4:S	2.76	0.59
1:E:590:PHE:O	1:E:595:GLN:NE2	2.35	0.58
2:H:0:HIS:N	3:H:201:SO4:O2	2.32	0.56
2:B:174:LEU:HD21	2:B:177:ILE:HD11	1.87	0.56
2:D:44:SER:O	2:D:65:ARG:NH1	2.39	0.56
2:D:28:LEU:O	2:D:28:LEU:HD12	2.06	0.55
1:E:592:SER:HB2	1:E:614:ARG:HH12	1.71	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:549:ILE:HD11	1:E:629:ILE:HG23	1.89	0.54
2:D:171:SER:OG	1:C:545:MET:O	2.20	0.53
2:D:28:LEU:C	2:D:28:LEU:HD12	2.29	0.52
1:G:550:ASP:OD1	1:G:552:LYS:HG2	2.09	0.52
1:A:552:LYS:CE	1:A:552:LYS:HA	2.40	0.52
1:A:552:LYS:HA	1:A:552:LYS:HE2	1.91	0.52
1:E:571:THR:HG23	1:E:590:PHE:HD1	1.74	0.52
2:D:9:HIS:ND1	4:D:336:HOH:O	2.34	0.52
2:H:143:ASN:HB3	2:H:182:ILE:O	2.10	0.52
2:H:58:TYR:CE1	2:H:77:GLN:CB	2.93	0.51
1:A:571:THR:HG23	1:A:590:PHE:HD1	1.75	0.51
2:B:5:GLU:OE1	2:B:82:ARG:NE	2.38	0.50
1:C:556:ILE:HD12	1:C:558:LEU:HD21	1.94	0.50
1:G:571:THR:HG23	1:G:590:PHE:HD1	1.77	0.50
2:H:11:TYR:HB2	2:H:73:ASP:OD1	2.12	0.49
1:G:577:LYS:HE2	1:G:579:ASP:OD2	2.13	0.49
2:D:143:ASN:HB3	2:D:182:ILE:O	2.12	0.49
2:B:103:ARG:HD3	3:B:201:SO4:O4	2.13	0.49
1:C:603:GLN:HA	1:C:603:GLN:OE1	2.11	0.48
2:H:181:TYR:HB2	1:G:614:ARG:HH11	1.78	0.48
1:C:571:THR:HG23	1:C:590:PHE:HD1	1.77	0.48
1:G:625:ARG:O	1:G:629:ILE:HG13	2.14	0.48
1:A:558:LEU:HD13	1:A:616:LEU:HD11	1.94	0.48
1:A:637:LYS:HE3	1:A:641:VAL:HG13	1.96	0.48
1:G:552:LYS:HD3	1:G:552:LYS:N	2.28	0.48
2:H:18:PRO:HG3	1:E:580:GLU:OE1	2.14	0.47
2:B:142:ARG:HB3	2:B:185:THR:HG21	1.96	0.47
1:A:625:ARG:O	1:A:629:ILE:HG13	2.14	0.47
2:D:26:ASN:HB2	2:H:4:ILE:HD11	1.97	0.47
1:A:533:LYS:N	1:A:537:GLN:OE1	2.41	0.47
1:C:574:ASN:OD1	1:C:589:ASN:ND2	2.47	0.47
1:A:603:GLN:HB2	1:A:605:PHE:HD1	1.80	0.47
1:C:527:LYS:HD3	1:C:528:ARG:N	2.30	0.47
2:H:137:THR:HG23	2:B:44:SER:OG	2.16	0.46
2:D:149:PHE:HE2	2:D:178:ILE:HG13	1.81	0.46
2:H:38:THR:OG1	4:H:350:HOH:O	2.20	0.46
1:E:592:SER:OG	1:E:614:ARG:NH1	2.48	0.46
2:F:17:GLU:OE2	2:F:33:VAL:O	2.34	0.46
2:B:143:ASN:OD1	2:B:185:THR:HB	2.16	0.46
2:D:175:LEU:HD21	2:D:178:ILE:HD11	1.99	0.45
1:G:552:LYS:H	1:G:552:LYS:HD3	1.81	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:175:LEU:HD21	2:H:178:ILE:HD11	1.99	0.45
2:D:181:TYR:HB2	1:C:614:ARG:HD3	1.99	0.44
1:E:556:ILE:HD12	1:E:558:LEU:HD21	1.98	0.44
1:A:539:PRO:HA	1:A:540:PRO:HD3	1.89	0.44
1:E:565:VAL:HA	1:E:566:PRO:HD3	1.85	0.44
1:E:631:ASN:OD1	1:E:635:ASN:ND2	2.51	0.44
2:D:181:TYR:HB3	1:C:612:PHE:HB2	2.00	0.43
1:A:545:MET:O	2:B:170:SER:OG	2.21	0.43
2:F:23:PHE:CE1	2:F:28:LEU:HD12	2.54	0.43
2:H:6:SER:O	2:H:82:ARG:NH2	2.45	0.43
1:C:636:LEU:HD12	1:C:636:LEU:HA	1.84	0.43
1:E:539:PRO:O	1:E:542:VAL:HB	2.18	0.43
2:H:16:LYS:H	2:H:16:LYS:HD3	1.83	0.43
1:A:565:VAL:HA	1:A:566:PRO:HD3	1.85	0.43
1:A:638:ARG:HB3	1:A:638:ARG:HE	1.63	0.42
1:C:539:PRO:HA	1:C:540:PRO:HD3	1.83	0.42
2:B:63:LEU:HD23	2:B:69:VAL:HG12	2.01	0.42
1:A:533:LYS:HA	1:A:533:LYS:HD3	1.89	0.42
2:H:137:THR:HG21	2:B:45:ASN:HB2	2.02	0.42
2:D:73:ASP:OD1	2:D:74:GLY:N	2.53	0.41
1:E:595:GLN:OE1	1:E:614:ARG:HD3	2.20	0.41
2:H:137:THR:CG2	2:B:45:ASN:HB2	2.50	0.41
1:A:556:ILE:HD12	1:A:558:LEU:HD21	2.02	0.41
1:G:565:VAL:HA	1:G:566:PRO:HD3	1.88	0.41
1:G:556:ILE:HD12	1:G:558:LEU:HD21	2.02	0.41
2:D:14:LEU:HD21	1:C:562:GLY:HA3	2.03	0.41
2:H:173:ASP:OD1	1:G:625:ARG:NH2	2.51	0.41
2:B:16:LYS:H	2:B:16:LYS:CD	2.32	0.41
1:C:625:ARG:O	1:C:629:ILE:HG13	2.21	0.41
2:D:181:TYR:CE2	2:D:183:PRO:HG3	2.55	0.41
2:B:4:ILE:HD12	2:B:4:ILE:HA	1.92	0.41
1:C:586:LEU:HB2	1:C:626:TYR:CD2	2.56	0.41
2:D:165:SER:O	2:D:169:SER:OG	2.38	0.40
2:H:73:ASP:OD2	4:H:335:HOH:O	2.22	0.40
1:G:586:LEU:HB2	1:G:626:TYR:CD2	2.56	0.40
1:A:527:LYS:HB2	1:A:527:LYS:HE2	1.82	0.40
1:A:565:VAL:HG21	2:B:127:PHE:HB3	2.03	0.40
1:G:635:ASN:OD1	1:G:636:LEU:N	2.55	0.40
2:H:88:LYS:HD3	2:H:88:LYS:HA	1.83	0.40
2:H:104:GLY:HA2	1:G:566:PRO:O	2.21	0.40
1:E:539:PRO:HA	1:E:540:PRO:HD3	1.85	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:636:LEU:HD23	1:G:636:LEU:HA	1.92	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	109/127 (86%)	103 (94%)	6 (6%)	0	100	100
1	C	102/127 (80%)	96 (94%)	6 (6%)	0	100	100
1	E	102/127 (80%)	98 (96%)	4 (4%)	0	100	100
1	G	102/127 (80%)	98 (96%)	4 (4%)	0	100	100
2	B	170/195 (87%)	166 (98%)	3 (2%)	1 (1%)	25	36
2	D	160/195 (82%)	159 (99%)	1 (1%)	0	100	100
2	F	163/195 (84%)	159 (98%)	4 (2%)	0	100	100
2	H	158/195 (81%)	156 (99%)	2 (1%)	0	100	100
All	All	1066/1288 (83%)	1035 (97%)	30 (3%)	1 (0%)	51	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	34	GLY

### 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	102/114 (90%)	90 (88%)	12 (12%)	5	7
1	C	96/114 (84%)	86 (90%)	10 (10%)	7	10
1	E	95/114 (83%)	89 (94%)	6 (6%)	18	28
1	G	95/114 (83%)	82 (86%)	13 (14%)	3	4
2	B	143/158 (90%)	132 (92%)	11 (8%)	13	20
2	D	139/158 (88%)	128 (92%)	11 (8%)	12	19
2	F	141/158 (89%)	129 (92%)	12 (8%)	10	16
2	H	138/158 (87%)	125 (91%)	13 (9%)	8	13
All	All	949/1088 (87%)	861 (91%)	88 (9%)	9	13

All (88) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	533	LYS
1	A	542	VAL
1	A	552	LYS
1	A	571	THR
1	A	587	ARG
1	A	592	SER
1	A	599	ARG
1	A	601	ASP
1	A	603	GLN
1	A	614	ARG
1	A	616	LEU
1	A	638	ARG
2	D	16	LYS
2	D	31	LYS
2	D	38	THR
2	D	53	ARG
2	D	81	GLU
2	D	82	ARG
2	D	139	LEU
2	D	145	ILE
2	D	153	ASP
2	D	164	LYS
2	D	169	SER
2	F	4	ILE
2	F	8	ASP

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	F	17	GLU
2	F	31	LYS
2	F	33	VAL
2	F	82	ARG
2	F	88	LYS
2	F	96	GLU
2	F	109	LYS
2	F	142	ARG
2	F	169	SER
2	F	186	THR
2	H	16	LYS
2	H	28	LEU
2	H	31	LYS
2	H	82	ARG
2	H	88	LYS
2	H	93	THR
2	H	109	LYS
2	H	137	THR
2	H	148	GLU
2	H	166	LYS
2	H	169	SER
2	H	172	ARG
2	H	183	PRO
1	G	542	VAL
1	G	549	ILE
1	G	552	LYS
1	G	571	THR
1	G	576	SER
1	G	582	GLU
1	G	586	LEU
1	G	587	ARG
1	G	592	SER
1	G	616	LEU
1	G	617	THR
1	G	636	LEU
1	G	638	ARG
1	E	542	VAL
1	E	571	THR
1	E	580	GLU
1	E	592	SER
1	E	614	ARG
1	E	616	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	527	LYS
1	C	542	VAL
1	C	571	THR
1	C	576	SER
1	C	580	GLU
1	C	586	LEU
1	C	587	ARG
1	C	592	SER
1	C	603	GLN
1	C	636	LEU
2	B	4	ILE
2	B	16	LYS
2	B	28	LEU
2	B	31	LYS
2	B	78	GLU
2	B	82	ARG
2	B	103	ARG
2	B	136	ASN
2	B	142	ARG
2	B	161	GLN
2	B	184	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	F	138	ASN
1	G	595	GLN
2	B	77	GLN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

9 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	$\# Z  > 2$
3	SO4	D	201	-	4,4,4	0.18	0	6,6,6	0.05	0
3	SO4	B	202	-	4,4,4	0.17	0	6,6,6	0.06	0
3	SO4	B	201	-	4,4,4	0.18	0	6,6,6	0.05	0
3	SO4	D	202	-	4,4,4	0.17	0	6,6,6	0.05	0
3	SO4	C	701	-	4,4,4	0.15	0	6,6,6	0.07	0
3	SO4	E	701	-	4,4,4	0.14	0	6,6,6	0.06	0
3	SO4	A	701	-	4,4,4	0.15	0	6,6,6	0.06	0
3	SO4	H	201	-	4,4,4	0.14	0	6,6,6	0.08	0
3	SO4	F	201	-	4,4,4	0.15	0	6,6,6	0.05	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	201	SO4	1	0
3	H	201	SO4	3	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 ⓘ

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	113/127 (88%)	0.20	8 (7%) 16 14	9, 25, 54, 101	0
1	C	106/127 (83%)	0.04	4 (3%) 40 39	9, 25, 47, 67	0
1	E	106/127 (83%)	0.27	4 (3%) 40 39	8, 27, 46, 69	0
1	G	106/127 (83%)	0.11	5 (4%) 31 30	12, 32, 53, 61	0
2	B	176/195 (90%)	0.09	10 (5%) 23 22	6, 21, 51, 81	0
2	D	170/195 (87%)	-0.15	0 100 100	6, 17, 40, 54	0
2	F	173/195 (88%)	0.09	9 (5%) 27 26	6, 17, 52, 75	0
2	H	168/195 (86%)	-0.07	5 (2%) 50 49	6, 20, 44, 75	0
All	All	1118/1288 (86%)	0.06	45 (4%) 38 37	6, 21, 51, 101	0

All (45) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1	MET	8.2
2	B	162	VAL	8.0
2	F	186	THR	7.8
1	A	600	LYS	6.4
1	A	601	ASP	6.3
2	B	161	GLN	6.0
2	F	1	MET	5.6
1	A	602	ASP	5.1
1	G	641	VAL	4.9
1	C	603	GLN	4.7
2	F	153	ASP	4.7
2	H	35	GLY	4.3
2	B	2	ALA	4.2
1	A	599	ARG	4.0
1	E	527	LYS	4.0
1	C	604	PRO	4.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	F	152	GLY	3.8
1	A	641	VAL	3.8
2	H	34	GLY	3.6
1	A	603	GLN	3.6
2	H	184	GLY	3.5
1	E	526	PHE	3.4
1	A	642	LYS	3.4
2	F	0	HIS	3.4
1	G	640	ALA	3.3
1	G	624	ASP	3.2
2	H	0	HIS	3.1
2	F	37	GLU	3.0
2	H	172	ARG	3.0
1	G	583	TRP	2.8
2	F	142	ARG	2.6
2	B	142	ARG	2.6
1	A	638	ARG	2.5
2	B	3	ALA	2.5
2	B	4	ILE	2.4
2	F	17	GLU	2.3
1	C	526	PHE	2.3
1	E	582	GLU	2.3
2	F	172	ARG	2.2
1	G	595	GLN	2.2
1	C	639	GLU	2.2
2	B	37	GLU	2.2
1	E	581	GLY	2.2
2	B	34	GLY	2.1
2	B	184	THR	2.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands [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, 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	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SO4	H	201	5/5	0.85	0.28	51,57,67,75	0
3	SO4	A	701	5/5	0.90	0.12	43,58,62,77	0
3	SO4	F	201	5/5	0.91	0.14	53,54,70,86	0
3	SO4	C	701	5/5	0.92	0.16	48,54,68,70	0
3	SO4	E	701	5/5	0.97	0.08	35,42,51,55	0
3	SO4	D	202	5/5	0.99	0.15	8,10,11,11	0
3	SO4	B	202	5/5	0.99	0.17	8,8,10,10	0
3	SO4	B	201	5/5	0.99	0.18	7,10,12,12	0
3	SO4	D	201	5/5	1.00	0.15	9,9,12,14	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.