



# Full wwPDB X-ray Structure Validation Report

Feb 26, 2014 – 02:29 PM GMT

PDB ID : 4C0Z  
Title : The N-terminal domain of the Streptococcus pyogenes pilus tip adhesin Cpa  
Authors : Linke-Winnebeck, C.; Paterson, N.; Baker, E.N.  
Deposited on : 2013-08-08  
Resolution : 2.00 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

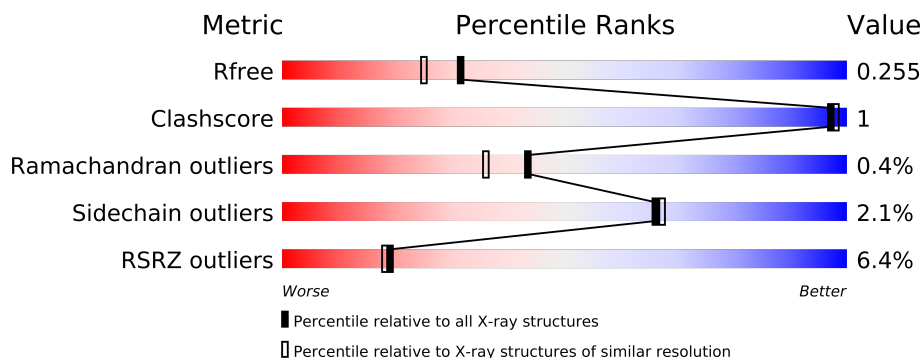
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.00 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))
$R_{free}$	66092	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	215	
1	B	215	
1	C	215	
1	D	215	
1	E	215	
1	F	215	
1	G	215	
1	H	215	
1	I	215	
1	J	215	
1	K	215	
1	L	215	

The following table lists non-polymeric compounds that are outliers for geometric or electron-

density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	CL	B	1221	-	X
2	CL	B	1222	-	X
2	CL	E	1220	-	X
2	CL	G	1221	-	X
3	SPD	A	1223	-	X
4	PO4	B	1220	-	X
4	PO4	F	1222	-	X
5	GOL	I	1220	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 21588 atoms, of which 0 are hydrogen and 0 are deuterium.

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 ANCILLARY PROTEIN 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	208	Total	C	N	O	S	0	2	0
			1718	1091	285	339	3			
1	B	208	Total	C	N	O	S	0	0	0
			1706	1084	284	335	3			
1	C	205	Total	C	N	O	S	0	3	0
			1710	1090	283	334	3			
1	D	208	Total	C	N	O	S	0	2	0
			1718	1091	285	339	3			
1	E	206	Total	C	N	O	S	0	1	0
			1700	1081	282	334	3			
1	F	209	Total	C	N	O	S	0	1	0
			1722	1094	286	339	3			
1	G	207	Total	C	N	O	S	0	1	0
			1701	1082	282	334	3			
1	H	208	Total	C	N	O	S	0	2	0
			1716	1091	284	338	3			
1	I	207	Total	C	N	O	S	0	1	0
			1703	1082	283	335	3			
1	J	207	Total	C	N	O	S	0	2	0
			1700	1082	280	335	3			
1	K	206	Total	C	N	O	S	0	0	0
			1680	1068	277	332	3			
1	L	209	Total	C	N	O	S	0	1	0
			1722	1094	286	339	3			

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

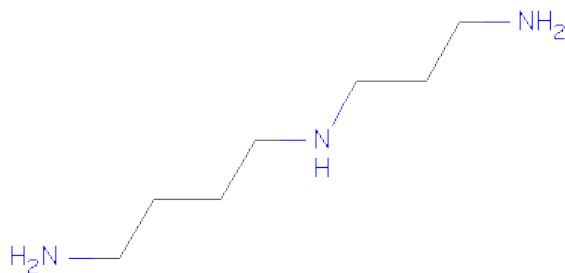
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	1	Total	Cl	0	0
			1	1		
2	J	1	Total	Cl	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	E	1	Total 1	Cl 1	0	0
2	H	3	Total 3	Cl 3	0	0
2	B	2	Total 2	Cl 2	0	0
2	C	1	Total 1	Cl 1	0	0
2	A	2	Total 2	Cl 2	0	0
2	L	2	Total 2	Cl 2	0	0
2	F	1	Total 1	Cl 1	0	0

- Molecule 3 is SPERMIDINE (three-letter code: SPD) (formula:  $C_7H_{19}N_3$ ).



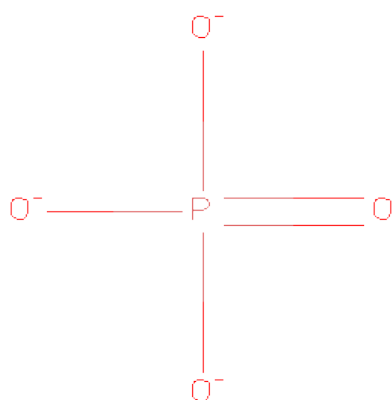
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 10	C 7	N 3	0	0
3	C	1	Total 10	C 7	N 3	0	0
3	F	1	Total 10	C 7	N 3	0	0
3	G	1	Total 10	C 7	N 3	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	I	1	Total	C	N	0	0
			10	7	3		
3	L	1	Total	C	N	0	0
			10	7	3		

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	O	P	0	0
			5	4	1		
4	C	1	Total	O	P	0	0
			5	4	1		
4	F	1	Total	O	P	0	0
			5	4	1		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	C	O	0	0
			6	3	3		
5	I	1	Total	C	O	0	0
			6	3	3		
5	L	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	108	Total	O	0	0
			108	108		
6	B	73	Total	O	0	0
			73	73		
6	C	117	Total	O	0	0
			117	117		
6	D	93	Total	O	0	0
			93	93		
6	E	79	Total	O	0	0
			79	79		
6	F	79	Total	O	0	0
			79	79		
6	G	93	Total	O	0	0
			93	93		
6	H	69	Total	O	0	0
			69	69		
6	I	66	Total	O	0	0
			66	66		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	J	83	Total 83	O 83	0	0
6	K	58	Total 58	O 58	0	0
6	L	67	Total 67	O 67	0	0

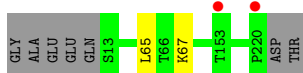


### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: ANCILLARY PROTEIN 2

Chain A: 



- Molecule 1: ANCILLARY PROTEIN 2

Chain B: 



- Molecule 1: ANCILLARY PROTEIN 2

Chain C: 



- Molecule 1: ANCILLARY PROTEIN 2

Chain D: 



- Molecule 1: ANCILLARY PROTEIN 2

Chain E: 



- Molecule 1: ANCILLARY PROTEIN 2

Chain F: 



- Molecule 1: ANCILLARY PROTEIN 2

Chain G:



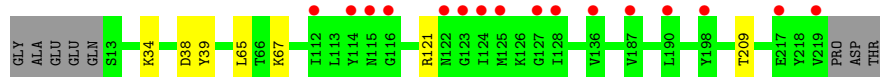
- Molecule 1: ANCILLARY PROTEIN 2

Chain H:



- Molecule 1: ANCILLARY PROTEIN 2

Chain I:



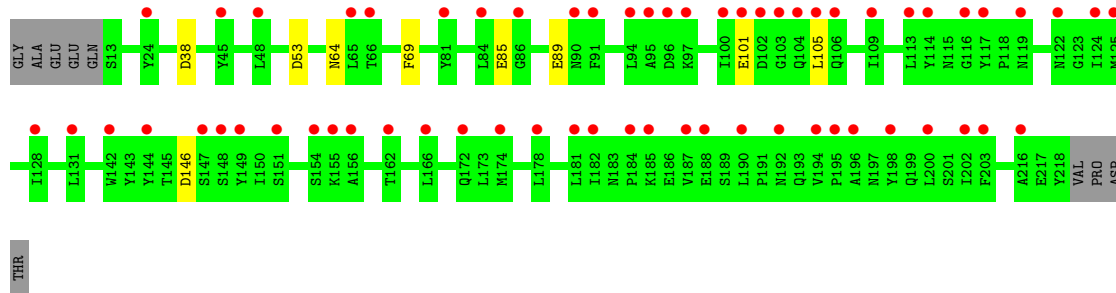
- Molecule 1: ANCILLARY PROTEIN 2

Chain J:



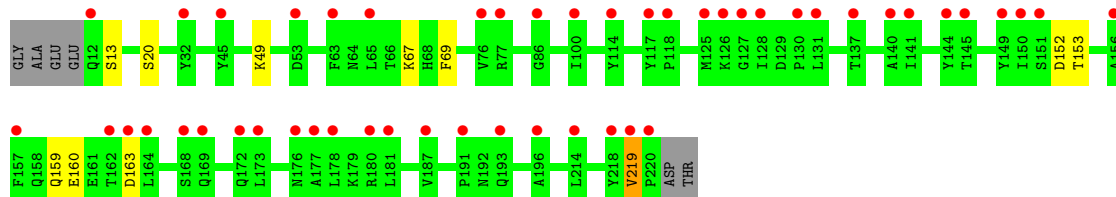
- Molecule 1: ANCILLARY PROTEIN 2

Chain K:



- Molecule 1: ANCILLARY PROTEIN 2

Chain L:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	132.22Å 132.22Å 136.58Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.50 – 2.00 47.50 – 2.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.50-2.00) 99.6 (47.50-2.00)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.42 (at 2.00Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, $R_{free}$	0.201 , 0.232 0.218 , 0.255	Depositor DCC
$R_{free}$ test set	9251 reflections (5.40%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.5	Xtriage
Anisotropy	0.206	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 46.6	EDS
Estimated twinning fraction	0.024 for -h,-k,l 0.031 for h,-h-k,-l 0.021 for -k,-h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 180592 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	21588	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	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 58.09 % 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 2.1526e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, PO4, SPD, CL

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	0.48	0/1760	0.61	0/2387
1	B	0.48	0/1747	0.62	0/2368
1	C	0.51	0/1753	0.64	0/2373
1	D	0.49	0/1760	0.62	0/2387
1	E	0.48	0/1741	0.64	0/2360
1	F	0.48	0/1764	0.62	0/2392
1	G	0.46	0/1745	0.62	0/2366
1	H	0.47	0/1761	0.63	0/2388
1	I	0.46	0/1744	0.62	0/2364
1	J	0.47	0/1747	0.64	0/2369
1	K	0.54	0/1721	0.70	0/2335
1	L	0.59	0/1764	0.74	0/2392
All	All	0.49	0/21007	0.64	0/28481

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1718	0	0	3	0
1	B	1706	0	0	0	0
1	C	1710	0	0	2	0
1	D	1718	0	0	0	0
1	E	1700	0	0	1	0
1	F	1722	0	0	1	0
1	G	1701	0	0	2	0
1	H	1716	0	0	0	0
1	I	1703	0	0	4	0
1	J	1700	0	0	1	0
1	K	1680	0	0	1	0
1	L	1722	0	0	4	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	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	3	0	0	0	0
2	J	1	0	0	0	0
2	L	2	0	0	0	0
3	A	10	0	16	3	0
3	C	10	0	15	1	0
3	F	10	0	15	2	0
3	G	10	0	15	2	0
3	I	10	0	15	2	0
3	L	10	0	15	4	0
4	B	5	0	0	0	0
4	C	5	0	0	0	0
4	F	5	0	0	0	0
5	C	6	0	8	1	0
5	I	6	0	8	2	0
5	L	6	0	8	0	0
6	A	108	0	0	0	0
6	B	73	0	0	0	0
6	C	117	0	0	0	0
6	D	93	0	0	0	0
6	E	79	0	0	0	0
6	F	79	0	0	0	0
6	G	93	0	0	0	0
6	H	69	0	0	0	0
6	I	66	0	0	0	0
6	J	83	0	0	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	K	58	0	0	0	0
6	L	67	0	0	0	0
All	All	21588	0	115	19	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 1.

All (19) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:39:TYR:OH	5:C:1221:GOL:H2	1.81	0.80
1:I:65:LEU:O	3:I:1221:SPD:H51	2.02	0.60
1:L:67:LYS:O	3:L:1223:SPD:H52	2.04	0.58
1:C:67:LYS:O	3:C:1223:SPD:H42	2.06	0.56
1:G:67:LYS:O	3:G:1220:SPD:H51	2.08	0.53
1:G:67:LYS:O	3:G:1220:SPD:C5	2.58	0.52
1:L:67:LYS:O	3:L:1223:SPD:C5	2.60	0.49
1:I:67:LYS:O	3:I:1221:SPD:H52	2.13	0.49
1:E:67:LYS:O	3:F:1223:SPD:H71	2.17	0.45
1:L:69:PHE:CE2	3:L:1223:SPD:H41	2.52	0.45
1:A:65:LEU:O	3:A:1223:SPD:H51	2.17	0.44
1:J:29:TYR:CD1	1:J:29:TYR:N	2.86	0.44
1:F:67:LYS:O	3:F:1223:SPD:H42	2.18	0.43
1:K:69:PHE:CE1	3:L:1223:SPD:H91	2.55	0.41
1:I:39:TYR:OH	5:I:1220:GOL:H32	2.20	0.41
1:I:34:LYS:N	5:I:1220:GOL:HO3	2.19	0.41
1:L:219:VAL:CG1	1:L:219:VAL:O	2.69	0.41
1:A:67:LYS:O	3:A:1223:SPD:H71	2.21	0.41
1:A:67:LYS:O	3:A:1223:SPD:H52	2.21	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	208/215 (97%)	203 (98%)	5 (2%)	0	100	100
1	B	206/215 (96%)	201 (98%)	4 (2%)	1 (0%)	38	29
1	C	202/215 (94%)	198 (98%)	3 (2%)	1 (0%)	38	29
1	D	208/215 (97%)	203 (98%)	4 (2%)	1 (0%)	38	29
1	E	205/215 (95%)	202 (98%)	3 (2%)	0	100	100
1	F	208/215 (97%)	204 (98%)	3 (1%)	1 (0%)	38	29
1	G	206/215 (96%)	199 (97%)	6 (3%)	1 (0%)	38	29
1	H	208/215 (97%)	202 (97%)	5 (2%)	1 (0%)	38	29
1	I	206/215 (96%)	199 (97%)	7 (3%)	0	100	100
1	J	207/215 (96%)	201 (97%)	5 (2%)	1 (0%)	38	29
1	K	204/215 (95%)	189 (93%)	13 (6%)	2 (1%)	22	12
1	L	208/215 (97%)	195 (94%)	12 (6%)	1 (0%)	38	29
All	All	2476/2580 (96%)	2396 (97%)	70 (3%)	10 (0%)	43	36

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	K	53	ASP
1	K	64	ASN
1	B	64	ASN
1	D	64	ASN
1	F	64	ASN
1	G	64	ASN
1	H	64	ASN
1	J	64	ASN
1	C	64	ASN
1	L	219	VAL

### 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	194/197 (98%)	194 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	192/197 (98%)	188 (98%)	4 (2%)	66	67
1	C	192/197 (98%)	184 (96%)	8 (4%)	40	34
1	D	194/197 (98%)	191 (98%)	3 (2%)	76	79
1	E	191/197 (97%)	187 (98%)	4 (2%)	66	67
1	F	194/197 (98%)	192 (99%)	2 (1%)	85	88
1	G	192/197 (98%)	187 (97%)	5 (3%)	59	58
1	H	194/197 (98%)	189 (97%)	5 (3%)	59	58
1	I	192/197 (98%)	189 (98%)	3 (2%)	75	77
1	J	192/197 (98%)	188 (98%)	4 (2%)	66	67
1	K	188/197 (95%)	182 (97%)	6 (3%)	51	47
1	L	194/197 (98%)	186 (96%)	8 (4%)	41	35
All	All	2309/2364 (98%)	2257 (98%)	52 (2%)	66	63

All (52) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	49	LYS
1	B	91	PHE
1	B	100	ILE
1	B	102	ASP
1	C	38	ASP
1	C	91[A]	PHE
1	C	91[B]	PHE
1	C	102	ASP
1	C	153	THR
1	C	172[A]	GLN
1	C	172[B]	GLN
1	C	209	THR
1	D	102	ASP
1	D	119	ASN
1	D	209	THR
1	E	38	ASP
1	E	52	LEU
1	E	96	ASP
1	E	209	THR
1	F	38	ASP
1	F	209	THR
1	G	38	ASP

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	G	106	GLN
1	G	172	GLN
1	G	209[A]	THR
1	G	209[B]	THR
1	H	38	ASP
1	H	96	ASP
1	H	101	GLU
1	H	159	GLN
1	H	209	THR
1	I	38	ASP
1	I	121	ARG
1	I	209	THR
1	J	52	LEU
1	J	101	GLU
1	J	168	SER
1	J	209	THR
1	K	38	ASP
1	K	85	GLU
1	K	89	GLU
1	K	101	GLU
1	K	105	LEU
1	K	146	ASP
1	L	13	SER
1	L	20	SER
1	L	49	LYS
1	L	152	ASP
1	L	153	THR
1	L	159	GLN
1	L	160	GLU
1	L	163	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 26 ligands modelled in this entry, 14 are monoatomic - leaving 12 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	SPD	A	1223	1	9,9,9	0.69	0	8,8,8	0.75	0
4	PO4	B	1220	-	4,4,4	1.71	0	6,6,6	0.30	0
5	GOL	C	1221	-	5,5,5	0.21	0	5,5,5	0.48	0
4	PO4	C	1222	-	4,4,4	1.64	0	6,6,6	0.31	0
3	SPD	C	1223	1	9,9,9	0.63	0	8,8,8	0.98	1 (12%)
4	PO4	F	1222	-	4,4,4	1.71	0	6,6,6	0.30	0
3	SPD	F	1223	1	9,9,9	0.70	0	8,8,8	1.14	1 (12%)
3	SPD	G	1220	1	9,9,9	0.86	0	8,8,8	0.62	0
5	GOL	I	1220	-	5,5,5	0.17	0	5,5,5	0.47	0
3	SPD	I	1221	1	9,9,9	0.65	0	8,8,8	0.71	0
5	GOL	L	1222	-	5,5,5	0.13	0	5,5,5	0.46	0
3	SPD	L	1223	1	9,9,9	0.69	0	8,8,8	0.71	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	SPD	A	1223	1	-	0/7/7/7	0/0/0/0
4	PO4	B	1220	-	-	0/0/0/0	0/0/0/0
5	GOL	C	1221	-	-	0/4/4/4	0/0/0/0
4	PO4	C	1222	-	-	0/0/0/0	0/0/0/0
3	SPD	C	1223	1	-	0/7/7/7	0/0/0/0
4	PO4	F	1222	-	-	0/0/0/0	0/0/0/0
3	SPD	F	1223	1	-	0/7/7/7	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SPD	G	1220	1	-	0/7/7/7	0/0/0/0
5	GOL	I	1220	-	-	0/4/4/4	0/0/0/0
3	SPD	I	1221	1	-	0/7/7/7	0/0/0/0
5	GOL	L	1222	-	-	0/4/4/4	0/0/0/0
3	SPD	L	1223	1	-	0/7/7/7	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	1223	SPD	C8-C7-N6	-3.00	104.12	112.09
3	C	1223	SPD	C8-C7-N6	-2.48	105.49	112.09

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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	208/215 (96%)	0.04	2 (0%) 79 80	22, 37, 65, 88	0
1	B	208/215 (96%)	0.31	9 (4%) 34 32	21, 47, 87, 100	0
1	C	205/215 (95%)	-0.01	2 (0%) 79 80	20, 37, 67, 88	0
1	D	208/215 (96%)	0.14	3 (1%) 72 72	23, 45, 76, 102	0
1	E	206/215 (95%)	0.21	2 (0%) 79 80	24, 46, 75, 83	0
1	F	209/215 (97%)	0.20	2 (0%) 79 80	27, 49, 81, 116	0
1	G	207/215 (96%)	0.14	4 (1%) 64 64	24, 43, 71, 87	0
1	H	208/215 (96%)	0.19	5 (2%) 56 56	24, 43, 77, 98	0
1	I	207/215 (96%)	0.60	16 (7%) 13 12	28, 58, 104, 121	0
1	J	207/215 (96%)	0.19	5 (2%) 56 56	22, 43, 76, 93	0
1	K	206/215 (95%)	1.60	62 (30%) 1 1	38, 62, 91, 112	0
1	L	209/215 (97%)	1.44	49 (23%) 1 1	31, 55, 87, 104	0
All	All	2488/2580 (96%)	0.42	161 (6%) 19 18	20, 47, 84, 121	0

All (161) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	187	VAL	11.6
1	K	187	VAL	7.7
1	K	105	LEU	7.2
1	L	150	ILE	6.7
1	K	100	ILE	6.7
1	B	187	VAL	6.1
1	L	162	THR	5.9
1	L	219	VAL	5.6
1	K	182	ILE	5.4
1	H	187	VAL	5.4
1	K	106	GLN	5.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	L	140	ALA	5.1
1	K	124	ILE	5.1
1	K	198	TYR	5.1
1	I	114	TYR	5.0
1	K	196	ALA	5.0
1	L	177	ALA	4.9
1	L	187	VAL	4.9
1	K	166	LEU	4.9
1	L	181	LEU	4.7
1	L	157	PHE	4.6
1	B	154	SER	4.6
1	I	187	VAL	4.5
1	L	178	LEU	4.5
1	K	109	ILE	4.4
1	K	144	TYR	4.4
1	L	176	ASN	4.4
1	K	113	LEU	4.3
1	G	187	VAL	4.1
1	K	194	VAL	4.1
1	K	96	ASP	4.1
1	I	125	MET	4.1
1	K	200	LEU	4.0
1	K	174	MET	4.0
1	J	189	SER	4.0
1	L	191	PRO	3.9
1	K	116	GLY	3.8
1	K	184	PRO	3.8
1	L	180	ARG	3.7
1	L	131	LEU	3.7
1	K	188	GLU	3.7
1	K	86	GLY	3.6
1	K	178	LEU	3.6
1	L	193	GLN	3.6
1	K	142	TRP	3.6
1	C	187	VAL	3.6
1	K	117	TYR	3.6
1	K	155	LYS	3.5
1	L	144	TYR	3.5
1	L	141	ILE	3.5
1	K	202	ILE	3.4
1	L	220	PRO	3.4
1	K	103	GLY	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	K	147	SER	3.4
1	J	188	GLU	3.4
1	L	156	ALA	3.3
1	L	127	GLY	3.3
1	K	90	ASN	3.3
1	L	114	TYR	3.2
1	L	149	TYR	3.2
1	K	149	TYR	3.1
1	I	127	GLY	3.1
1	K	94	LEU	3.1
1	K	101	GLU	3.1
1	I	219	VAL	3.0
1	L	118	PRO	3.0
1	L	77	ARG	3.0
1	L	169	GLN	3.0
1	L	164	LEU	3.0
1	I	136	VAL	3.0
1	K	125	MET	3.0
1	B	102	ASP	2.9
1	L	125	MET	2.9
1	G	151	SER	2.9
1	G	219	VAL	2.9
1	L	130	PRO	2.9
1	I	124	ILE	2.9
1	B	100	ILE	2.9
1	L	100	ILE	2.9
1	K	65	LEU	2.8
1	A	220	PRO	2.8
1	K	181	LEU	2.8
1	B	153	THR	2.8
1	K	203	PHE	2.8
1	K	66	THR	2.8
1	K	162	THR	2.8
1	J	190	LEU	2.8
1	K	190	LEU	2.8
1	K	148	SER	2.8
1	L	126	LYS	2.8
1	D	196	ALA	2.7
1	K	216	ALA	2.7
1	K	84	LEU	2.7
1	K	156	ALA	2.7
1	K	104	GLN	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	H	190	LEU	2.6
1	L	12	GLN	2.6
1	C	91[A]	PHE	2.6
1	I	128	ILE	2.6
1	K	91	PHE	2.6
1	L	168	SER	2.5
1	L	128	ILE	2.5
1	J	155	LYS	2.5
1	K	119	ASN	2.5
1	L	196	ALA	2.5
1	I	190	LEU	2.5
1	L	145	THR	2.5
1	K	102	ASP	2.5
1	L	137	THR	2.4
1	K	97	LYS	2.4
1	K	48	LEU	2.4
1	L	173	LEU	2.4
1	K	95	ALA	2.4
1	K	114	TYR	2.4
1	F	219	VAL	2.4
1	L	45	TYR	2.4
1	K	154	SER	2.3
1	H	188	GLU	2.3
1	L	32	TYR	2.3
1	G	190	LEU	2.3
1	K	131	LEU	2.3
1	K	24	TYR	2.3
1	H	91	PHE	2.3
1	I	112	ILE	2.3
1	L	86	GLY	2.2
1	E	22	GLN	2.2
1	I	115	ASN	2.2
1	L	53	ASP	2.2
1	L	151	SER	2.2
1	I	122	ASN	2.2
1	K	172	GLN	2.2
1	L	63	PHE	2.2
1	D	220	PRO	2.2
1	K	195	PRO	2.2
1	B	77	ARG	2.2
1	L	117	TYR	2.2
1	K	122	ASN	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	I	123	GLY	2.2
1	L	163	ASP	2.2
1	F	77	ARG	2.1
1	B	151	SER	2.1
1	E	219	VAL	2.1
1	I	198	TYR	2.1
1	L	218	TYR	2.1
1	I	217	GLU	2.1
1	K	128	ILE	2.1
1	K	151	SER	2.1
1	L	76	VAL	2.1
1	B	107	GLN	2.1
1	L	172	GLN	2.1
1	I	116	GLY	2.1
1	K	45	TYR	2.1
1	K	185	LYS	2.1
1	B	91	PHE	2.1
1	H	189	SER	2.1
1	K	81	TYR	2.1
1	L	65	LEU	2.0
1	L	214	LEU	2.0
1	K	192	ASN	2.0
1	D	77	ARG	2.0
1	A	153	THR	2.0

## 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 ⓘ

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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CL	E	1220	1/1	0.19	5.05	69,69,69,69	0
2	CL	B	1221	1/1	0.24	4.82	78,78,78,78	0
5	GOL	I	1220	6/6	0.20	4.25	59,71,73,74	0
3	SPD	A	1223	10/10	0.15	3.22	22,25,28,33	0
2	CL	G	1221	1/1	0.20	3.05	56,56,56,56	0
4	PO4	F	1222	5/5	0.17	3.01	80,80,81,82	0
2	CL	B	1222	1/1	0.22	2.76	58,58,58,58	0
4	PO4	B	1220	5/5	0.24	2.61	86,88,90,90	0
3	SPD	G	1220	10/10	0.16	1.94	29,32,33,33	0
3	SPD	C	1223	10/10	0.14	1.62	23,28,31,33	0
3	SPD	F	1223	10/10	0.15	1.34	23,30,34,36	0
3	SPD	I	1221	10/10	0.14	1.07	24,28,36,40	0
4	PO4	C	1222	5/5	0.13	0.65	52,54,57,58	0
3	SPD	L	1223	10/10	0.24	0.55	40,45,48,51	0
2	CL	H	1223	1/1	0.14	0.45	79,79,79,79	0
5	GOL	C	1221	6/6	0.16	0.21	36,51,55,56	0
2	CL	A	1222	1/1	0.15	0.07	72,72,72,72	0
2	CL	H	1222	1/1	0.11	-0.70	82,82,82,82	0
2	CL	F	1221	1/1	0.12	-0.74	37,37,37,37	0
2	CL	J	1220	1/1	0.12	-0.81	31,31,31,31	0
2	CL	C	1220	1/1	0.10	-1.16	29,29,29,29	0
5	GOL	L	1222	6/6	0.18	-1.27	66,69,72,72	0
2	CL	H	1221	1/1	0.09	-2.11	33,33,33,33	0
2	CL	L	1221	1/1	0.11	-2.30	40,40,40,40	0
2	CL	A	1221	1/1	0.09	-3.07	27,27,27,27	0
2	CL	L	1224	1/1	0.07	-5.13	79,79,79,79	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.