



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2018 – 11:42 am GMT

PDB ID : 5J1V
Title : Crystal structure of human CLK1 in complex with pyrido[3,4-g]quinazoline derivative ZW29 (compound 13)
Authors : Chaikuad, A.; Esvan, Y.J.; Zeinyeh, W.; Boibessot, T.; Nauton, L.; They, V.; Loaec, N.; Meijer, L.; Giraud, F.; Moreau, P.; Anizon, F.; von Delft, F.; Bountra, C.; Arrowsmith, C.H.; Edwards, A.M.; Knapp, S.; Structural Genomics Consortium (SGC)
Deposited on : 2016-03-29
Resolution : 2.52 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

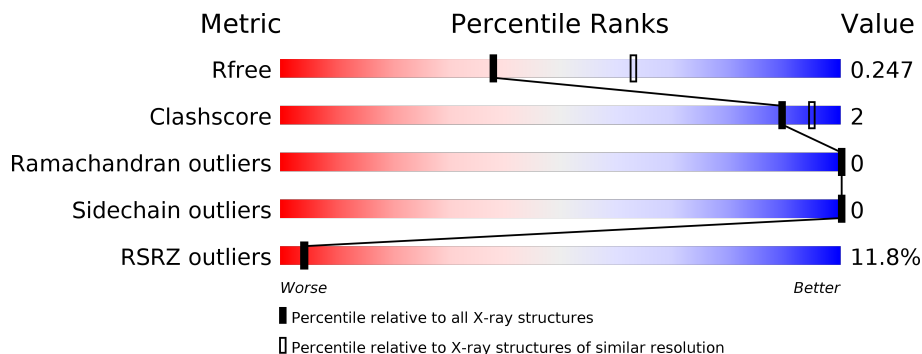
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.52 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}	111664	5045 (2.54-2.50)
Clashscore	122126	5751 (2.54-2.50)
Ramachandran outliers	120053	5650 (2.54-2.50)
Sidechain outliers	120020	5652 (2.54-2.50)
RSRZ outliers	108989	4938 (2.54-2.50)

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 ≥ 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	339	
1	B	339	
1	C	339	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8082 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.

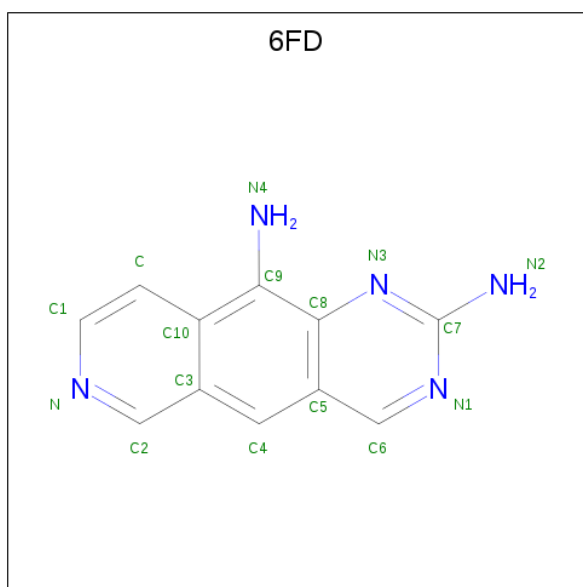
- Molecule 1 is a protein called Dual specificity protein kinase CLK1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	338	2721	1740	469	495	17	0	0	0
1	B	337	2725	1743	472	493	17	0	0	0
1	C	314	2481	1592	418	454	17	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	expression tag	UNP P49759
A	0	MET	-	expression tag	UNP P49759
A	432	ALA	ARG	conflict	UNP P49759
B	-1	SER	-	expression tag	UNP P49759
B	0	MET	-	expression tag	UNP P49759
B	432	ALA	ARG	conflict	UNP P49759
C	-1	SER	-	expression tag	UNP P49759
C	0	MET	-	expression tag	UNP P49759
C	432	ALA	ARG	conflict	UNP P49759

- Molecule 2 is pyrido[3,4-g]quinazoline-2,10-diamine (three-letter code: 6FD) (formula: C₁₁H₉N₅).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	N	0	0
			16	11	5		
2	B	1	Total	C	N	0	0
			16	11	5		
2	C	1	Total	C	N	0	0
			16	11	5		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		

Continued on next page...

Continued from previous page...

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

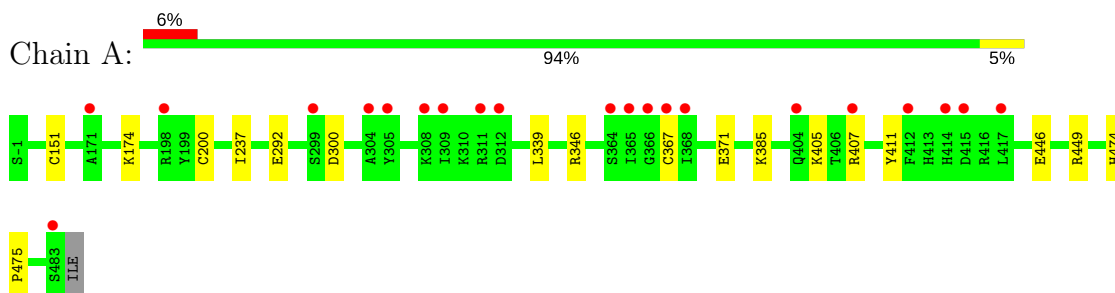
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	30	Total	O	0	0
			30	30		
4	B	26	Total	O	0	0
			26	26		
4	C	39	Total	O	0	0
			39	39		

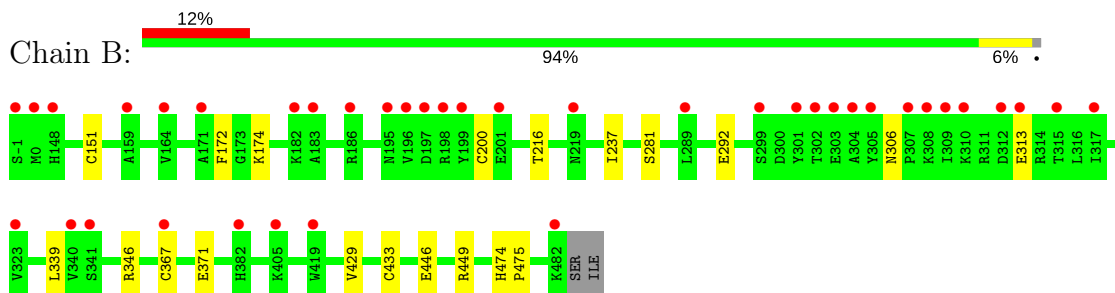
3 Residue-property plots [i](#)

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.

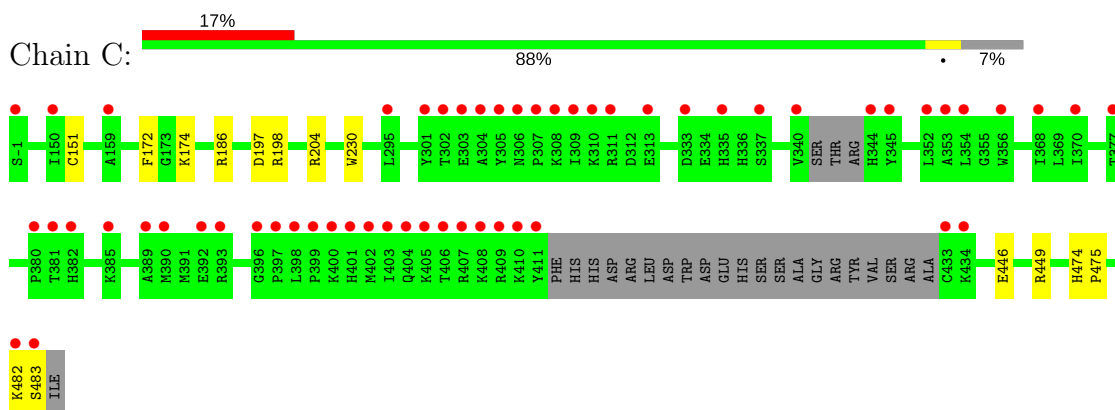
- Molecule 1: Dual specificity protein kinase CLK1



- Molecule 1: Dual specificity protein kinase CLK1



- Molecule 1: Dual specificity protein kinase CLK1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	56.47Å 116.02Å 90.11Å 90.00° 99.28° 90.00°	Depositor
Resolution (Å)	35.47 – 2.52 35.29 – 2.52	Depositor EDS
% Data completeness (in resolution range)	99.7 (35.47-2.52) 99.8 (35.29-2.52)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.16 (at 2.51Å)	Xtrriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.199 , 0.244 0.206 , 0.247	Depositor DCC
R_{free} test set	1897 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	56.3	Xtrriage
Anisotropy	0.242	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 50.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	8082	wwPDB-VP
Average B, all atoms (Å ²)	71.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.32% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, 6FD

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	0.63	1/2791 (0.0%)	0.62	0/3780
1	B	0.58	1/2795 (0.0%)	0.61	0/3783
1	C	0.60	0/2542	0.61	0/3448
All	All	0.60	2/8128 (0.0%)	0.61	0/11011

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	292	GLU	CD-OE1	6.36	1.32	1.25
1	B	292	GLU	CD-OE1	6.17	1.32	1.25

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	2721	0	2642	12	0
1	B	2725	0	2659	12	0
1	C	2481	0	2369	13	0
2	A	16	0	0	0	0
2	B	16	0	0	0	0
2	C	16	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	6	0	8	0	0
3	C	6	0	8	2	0
4	A	30	0	0	0	0
4	B	26	0	0	0	0
4	C	39	0	0	0	0
All	All	8082	0	7686	31	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (31) 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:B:281:SER:O	1:C:230:TRP:HZ2	1.54	0.91
1:B:281:SER:O	1:C:230:TRP:CZ2	2.42	0.72
1:A:407:ARG:HA	1:C:198:ARG:HG3	1.71	0.70
1:C:186:ARG:HD2	3:C:502:GOL:H11	1.74	0.67
1:B:151:CYS:SG	1:B:174:LYS:NZ	2.66	0.65
1:B:216:THR:HG23	1:C:204:ARG:NH1	2.14	0.63
1:A:151:CYS:SG	1:A:174:LYS:NZ	2.67	0.62
1:C:151:CYS:SG	1:C:174:LYS:NZ	2.67	0.61
1:A:405:LYS:HG2	1:C:197:ASP:OD2	2.02	0.59
1:A:446:GLU:OE2	1:A:449:ARG:NH1	2.37	0.56
1:B:446:GLU:OE2	1:B:449:ARG:NH1	2.38	0.56
1:C:446:GLU:OE2	1:C:449:ARG:NH1	2.38	0.56
1:A:407:ARG:HA	1:C:198:ARG:CG	2.38	0.54
1:A:385:LYS:HG2	1:A:411:TYR:CD1	2.42	0.53
1:B:200:CYS:SG	1:B:237:ILE:HD12	2.53	0.48
1:A:200:CYS:SG	1:A:237:ILE:HD12	2.53	0.48
1:A:474:HIS:CG	1:A:475:PRO:HD2	2.49	0.48
1:C:474:HIS:CG	1:C:475:PRO:HD2	2.50	0.47
1:B:474:HIS:CG	1:B:475:PRO:HD2	2.50	0.47
1:C:482:LYS:O	1:C:483:SER:C	2.53	0.47
1:A:385:LYS:HG2	1:A:411:TYR:CE1	2.51	0.45
1:B:306:ASN:CB	1:B:313:GLU:OE1	2.64	0.45
1:B:429:VAL:O	1:B:433:CYS:HB2	2.16	0.45
1:C:186:ARG:HD2	3:C:502:GOL:C1	2.46	0.44
1:A:339:LEU:HD11	1:A:346:ARG:NH2	2.34	0.43
1:B:339:LEU:HD11	1:B:346:ARG:NH2	2.34	0.42
1:A:300:ASP:OD1	1:A:300:ASP:N	2.53	0.41
1:B:367:CYS:O	1:B:371:GLU:HG3	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:367:CYS:O	1:A:371:GLU:HG3	2.21	0.41
1:C:172:PHE:CD1	1:C:172:PHE:C	2.95	0.40
1:B:172:PHE:CD1	1:B:172:PHE:C	2.95	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	336/339 (99%)	320 (95%)	16 (5%)	0	100	100
1	B	335/339 (99%)	323 (96%)	12 (4%)	0	100	100
1	C	308/339 (91%)	296 (96%)	12 (4%)	0	100	100
All	All	979/1017 (96%)	939 (96%)	40 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	296/308 (96%)	296 (100%)	0	100	100
1	B	297/308 (96%)	297 (100%)	0	100	100
1	C	264/308 (86%)	264 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	857/924 (93%)	857 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

5 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$
2	6FD	A	501	-	17,18,18	0.67	0	24,26,26	0.92	1 (4%)
3	GOL	A	502	-	5,5,5	0.44	0	5,5,5	0.23	0
2	6FD	B	501	-	17,18,18	0.68	0	24,26,26	0.97	2 (8%)
2	6FD	C	501	-	17,18,18	0.67	1 (5%)	24,26,26	0.94	1 (4%)
3	GOL	C	502	-	5,5,5	0.54	0	5,5,5	0.52	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
2	6FD	A	501	-	-	0/0/0/0	0/3/3/3
3	GOL	A	502	-	-	0/4/4/4	0/0/0/0
2	6FD	B	501	-	-	0/0/0/0	0/3/3/3
2	6FD	C	501	-	-	0/0/0/0	0/3/3/3
3	GOL	C	502	-	-	0/4/4/4	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	501	6FD	C9-C10	-2.29	1.39	1.43

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	501	6FD	C6-C5-C8	-2.24	112.48	117.57
2	A	501	6FD	C4-C5-C6	2.90	128.98	122.08
2	C	501	6FD	C4-C5-C6	2.91	129.00	122.08
2	B	501	6FD	C4-C5-C6	3.20	129.68	122.08

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
3	C	502	GOL	2	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, 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	338/339 (99%)	0.37	21 (6%) 20 22	36, 60, 112, 155	0
1	B	337/339 (99%)	0.56	39 (11%) 4 4	42, 70, 119, 164	0
1	C	314/339 (92%)	0.98	57 (18%) 1 1	36, 65, 142, 177	0
All	All	989/1017 (97%)	0.63	117 (11%) 4 4	36, 66, 128, 177	0

All (117) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	304	ALA	11.3
1	C	433	CYS	9.7
1	C	406	THR	8.8
1	C	305	TYR	8.6
1	C	434	LYS	8.0
1	C	309	ILE	7.0
1	C	404	GLN	6.9
1	C	303	GLU	6.2
1	A	304	ALA	6.2
1	C	307	PRO	6.0
1	C	308	LYS	6.0
1	B	304	ALA	5.8
1	B	482	LYS	5.8
1	C	411	TYR	5.7
1	B	307	PRO	5.5
1	C	403	ILE	5.2
1	C	483	SER	5.1
1	C	352	LEU	4.6
1	C	306	ASN	4.5
1	C	385	LYS	4.4
1	B	340	VAL	4.4
1	C	402	MET	4.4
1	B	313	GLU	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	482	LYS	4.3
1	A	483	SER	4.3
1	A	404	GLN	4.2
1	B	186	ARG	4.2
1	C	409	ARG	4.2
1	B	-1	SER	4.0
1	C	382	HIS	4.0
1	B	301	TYR	4.0
1	A	311	ARG	4.0
1	C	407	ARG	4.0
1	C	354	LEU	3.9
1	C	399	PRO	3.9
1	C	310	LYS	3.9
1	C	344	HIS	3.8
1	B	310	LYS	3.8
1	B	219	ASN	3.8
1	A	415	ASP	3.8
1	C	397	PRO	3.8
1	C	398	LEU	3.7
1	B	299	SER	3.6
1	C	-1	SER	3.6
1	B	303	GLU	3.6
1	A	412	PHE	3.5
1	B	309	ILE	3.4
1	B	182	LYS	3.4
1	B	305	TYR	3.4
1	C	302	THR	3.4
1	A	309	ILE	3.3
1	A	364	SER	3.3
1	C	396	GLY	3.2
1	B	317	ILE	3.2
1	C	337	SER	3.2
1	C	311	ARG	3.2
1	A	312	ASP	3.1
1	B	312	ASP	3.1
1	C	335	HIS	3.1
1	C	333	ASP	3.1
1	B	164	VAL	3.1
1	C	408	LYS	3.0
1	C	400	LYS	3.0
1	B	195	ASN	3.0
1	C	301	TYR	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	302	THR	2.9
1	A	407	ARG	2.8
1	B	148	HIS	2.8
1	B	199	TYR	2.8
1	B	198	ARG	2.7
1	C	389	ALA	2.7
1	C	390	MET	2.6
1	C	392	GLU	2.6
1	C	150	ILE	2.6
1	C	401	HIS	2.6
1	C	410	LYS	2.5
1	A	171	ALA	2.5
1	B	197	ASP	2.5
1	B	405	LYS	2.5
1	B	159	ALA	2.5
1	B	196	VAL	2.4
1	A	365	ILE	2.4
1	B	367	CYS	2.4
1	C	377	THR	2.4
1	A	305	TYR	2.4
1	A	198	ARG	2.4
1	C	345	TYR	2.4
1	B	183	ALA	2.4
1	C	380	PRO	2.3
1	B	171	ALA	2.3
1	A	308	LYS	2.3
1	A	368	ILE	2.3
1	A	367	CYS	2.3
1	C	393	ARG	2.3
1	A	299	SER	2.3
1	B	382	HIS	2.3
1	C	405	LYS	2.2
1	B	315	THR	2.2
1	C	340	VAL	2.2
1	C	381	THR	2.2
1	A	366	GLY	2.2
1	B	419	TRP	2.2
1	C	313	GLU	2.2
1	C	368	ILE	2.2
1	C	356	TRP	2.1
1	C	370	ILE	2.1
1	B	308	LYS	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	323	VAL	2.1
1	B	0	MET	2.1
1	B	289	LEU	2.1
1	C	295	LEU	2.1
1	B	201	GLU	2.1
1	B	341	SER	2.0
1	C	353	ALA	2.0
1	C	159	ALA	2.0
1	A	414	HIS	2.0
1	A	417	LEU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no 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, 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	GOL	A	502	6/6	0.72	0.18	86,95,101,102	0
3	GOL	C	502	6/6	0.94	0.20	57,59,64,70	0
2	6FD	A	501	16/16	0.97	0.19	31,35,41,42	0
2	6FD	B	501	16/16	0.98	0.14	35,39,44,46	0
2	6FD	C	501	16/16	0.98	0.18	28,31,35,36	0

6.5 Other polymers [i](#)

There are no such residues in this entry.