Full wwPDB X-ray Structure Validation Report (i)

Jun 23, 2024 – 09:49 AM EDT

PDB ID	:	6QZL
Title	:	Structure of the H1 domain of human KCTD12
Authors	:	Pinkas, D.M.; Bufton, J.C.; Fox, A.E.; Newman, J.A.; Kupinska, K.; Burgess-Brown, N.A.; von Delft, F.; Arrowsmith, C.H.; Edwards, A.M.; Bountra, C.; Bullock, A.N.; Structural Genomics Consortium (SGC)
Deposited on	:	2019-03-11
Resolution	:	1.98 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 (1)) were used in the production of this report:

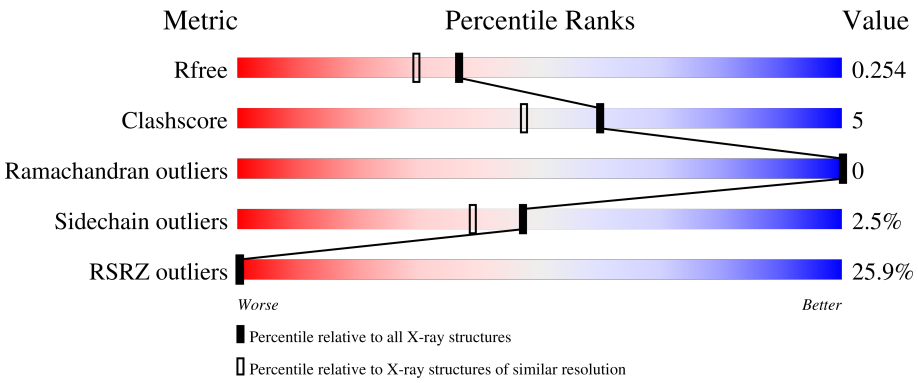
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 1.98 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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	126	25% 73% 8% • 18%
1	B	126	16% 72% 5% • 22%
1	C	126	20% 71% 7% • 21%
1	D	126	17% 65% 11% 24%
1	E	126	22% 65% 8% • 26%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4003 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 BTB/POZ domain-containing protein KCTD12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	103	796	502	131	158	5	0	0	0
1	B	98	775	489	132	150	4	0	0	0
1	C	100	795	500	139	152	4	0	0	0
1	D	96	747	471	123	149	4	0	0	0
1	E	93	709	446	117	142	4	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	200	SER	-	expression tag	UNP Q96CX2
A	201	MET	-	expression tag	UNP Q96CX2
B	200	SER	-	expression tag	UNP Q96CX2
B	201	MET	-	expression tag	UNP Q96CX2
C	200	SER	-	expression tag	UNP Q96CX2
C	201	MET	-	expression tag	UNP Q96CX2
D	200	SER	-	expression tag	UNP Q96CX2
D	201	MET	-	expression tag	UNP Q96CX2
E	200	SER	-	expression tag	UNP Q96CX2
E	201	MET	-	expression tag	UNP Q96CX2

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	31	31	31	0	0
2	B	47	47	47	0	0

Continued on next page...

Continued from previous page...

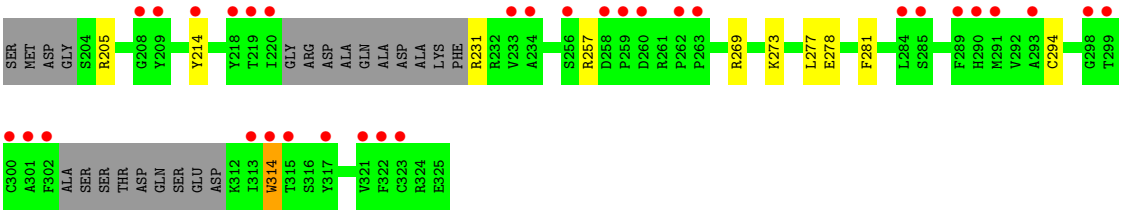
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	C	44	44	44	0	0
			Total	O		
2	D	29	29	29	0	0
			Total	O		
2	E	30	30	30	0	0

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.

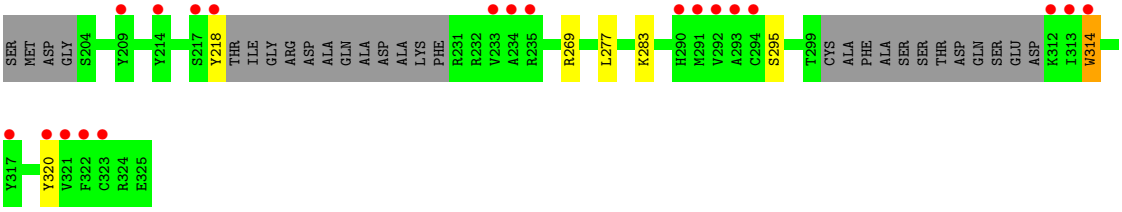
• Molecule 1: BTB/POZ domain-containing protein KCTD12

Chain A:



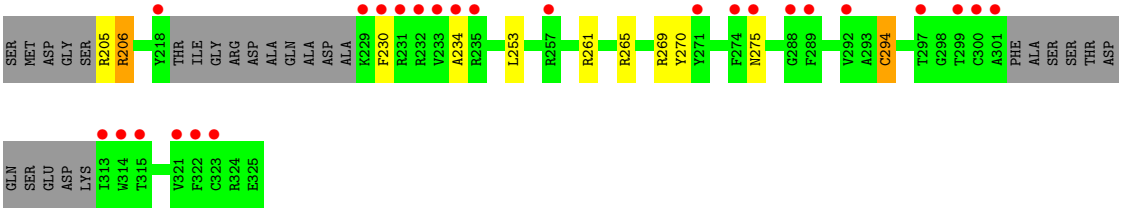
• Molecule 1: BTB/POZ domain-containing protein KCTD12

Chain B:



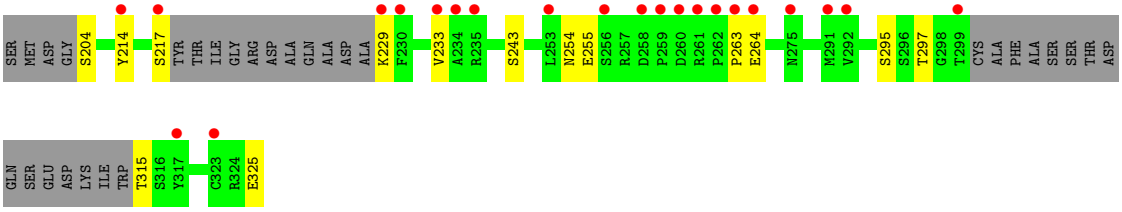
• Molecule 1: BTB/POZ domain-containing protein KCTD12

Chain C:



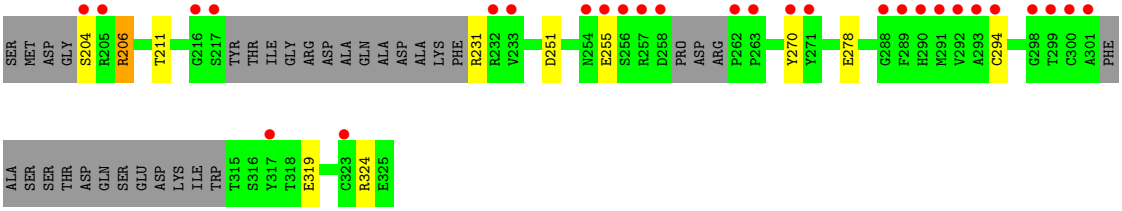
• Molecule 1: BTB/POZ domain-containing protein KCTD12

Chain D:



- Molecule 1: BTB/POZ domain-containing protein KCTD12

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	88.32Å 91.36Å 152.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.68 – 1.98 48.76 – 1.98	Depositor EDS
% Data completeness (in resolution range)	99.5 (45.68-1.98) 99.6 (48.76-1.98)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.67 (at 1.98Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.204 , 0.248 0.210 , 0.254	Depositor DCC
R_{free} test set	2181 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	48.4	Xtriage
Anisotropy	0.196	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 60.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.023 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4003	wwPDB-VP
Average B, all atoms (Å^2)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.73% 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)

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	0.40	0/813	0.55	0/1097
1	B	0.43	0/792	0.58	0/1066
1	C	0.39	0/811	0.55	0/1088
1	D	0.38	0/762	0.54	0/1025
1	E	0.40	0/721	0.59	0/967
All	All	0.40	0/3899	0.56	0/5243

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	796	0	722	9	0
1	B	775	0	717	5	0
1	C	795	0	741	10	0
1	D	747	0	687	9	0
1	E	709	0	648	6	0
2	A	31	0	0	2	0
2	B	47	0	0	1	0
2	C	44	0	0	1	0
2	D	29	0	0	1	0
2	E	30	0	0	0	0
All	All	4003	0	3515	34	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 5.

All (34) 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:C:205:ARG:NH2	2:C:401:HOH:O	2.14	0.77
1:D:325:GLU:OE2	2:D:401:HOH:O	2.06	0.72
1:C:234:ALA:O	1:C:275:ASN:ND2	2.23	0.71
1:A:257:ARG:O	1:A:269:ARG:NH1	2.23	0.71
1:B:283:LYS:NZ	2:B:401:HOH:O	2.26	0.69
1:E:231:ARG:NH1	1:E:278:GLU:OE1	2.32	0.61
1:C:206:ARG:NE	1:C:206:ARG:HA	2.18	0.59
1:D:217:SER:HA	1:D:315:THR:HA	1.84	0.59
1:A:273:LYS:NZ	2:A:403:HOH:O	2.36	0.58
1:C:206:ARG:HA	1:C:206:ARG:HE	1.73	0.54
1:D:214:TYR:CD2	1:D:233:VAL:HG22	2.43	0.53
1:C:261:ARG:HH11	1:C:265:ARG:HB3	1.73	0.53
1:A:205:ARG:O	2:A:401:HOH:O	2.19	0.52
1:D:243:SER:HB3	1:D:263:PRO:CB	2.40	0.52
1:E:206:ARG:NE	1:E:206:ARG:HA	2.25	0.51
1:E:255:GLU:HG2	1:E:270:TYR:CZ	2.45	0.51
1:C:269:ARG:HG2	1:D:229:LYS:HA	1.91	0.51
1:D:243:SER:HB3	1:D:263:PRO:HB2	1.93	0.51
1:E:231:ARG:HH12	1:E:278:GLU:CD	2.15	0.50
1:A:314:TRP:HZ2	1:D:297:THR:HG23	1.80	0.46
1:E:211:THR:CG2	1:E:319:GLU:HG3	2.45	0.46
1:B:295:SER:HB3	1:C:294:CYS:SG	2.56	0.45
1:B:277:LEU:HD21	1:B:320:TYR:CE1	2.52	0.45
1:C:261:ARG:NH1	1:C:265:ARG:HB3	2.31	0.45
1:A:294:CYS:SG	1:D:295:SER:HB3	2.59	0.43
1:B:218:TYR:HB2	1:B:314:TRP:CZ3	2.54	0.43
1:A:231:ARG:NH2	1:A:278:GLU:OE1	2.51	0.43
1:A:214:TYR:CD1	1:A:277:LEU:HD13	2.55	0.41
1:A:277:LEU:HG	1:A:281:PHE:CZ	2.55	0.41
1:E:204:SER:HA	1:E:324:ARG:HD2	2.02	0.41
1:C:253:LEU:HD21	1:C:270:TYR:CD2	2.55	0.41
1:D:204:SER:OG	1:D:325:GLU:O	2.32	0.41
1:A:231:ARG:HH22	1:A:278:GLU:CD	2.25	0.41
1:B:269:ARG:HG2	1:C:230:PHE:CD2	2.56	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	97/126 (77%)	95 (98%)	2 (2%)	0	100	100
1	B	92/126 (73%)	91 (99%)	1 (1%)	0	100	100
1	C	94/126 (75%)	93 (99%)	1 (1%)	0	100	100
1	D	90/126 (71%)	88 (98%)	2 (2%)	0	100	100
1	E	85/126 (68%)	83 (98%)	2 (2%)	0	100	100
All	All	458/630 (73%)	450 (98%)	8 (2%)	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	82/108 (76%)	81 (99%)	1 (1%)	71	67
1	B	81/108 (75%)	80 (99%)	1 (1%)	71	67
1	C	82/108 (76%)	80 (98%)	2 (2%)	49	41
1	D	79/108 (73%)	76 (96%)	3 (4%)	33	21
1	E	73/108 (68%)	70 (96%)	3 (4%)	30	18
All	All	397/540 (74%)	387 (98%)	10 (2%)	47	39

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	314	TRP
1	B	314	TRP
1	C	206	ARG
1	C	294	CYS
1	D	254	ASN
1	D	255	GLU
1	D	264	GLU
1	E	206	ARG
1	E	251	ASP
1	E	294	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	275	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)

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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	103/126 (81%)	1.70	32 (31%)	0	0	38, 56, 109, 125	0
1	B	98/126 (77%)	1.25	20 (20%)	1	0	35, 48, 97, 122	0
1	C	100/126 (79%)	1.46	25 (25%)	0	0	38, 52, 111, 115	0
1	D	96/126 (76%)	1.65	22 (22%)	0	0	39, 57, 97, 100	0
1	E	93/126 (73%)	1.57	28 (30%)	0	0	38, 56, 106, 130	0
All	All	490/630 (77%)	1.52	127 (25%)	0	0	35, 54, 105, 130	0

All (127) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	233	VAL	13.3
1	D	230	PHE	11.5
1	A	300	CYS	11.2
1	C	230	PHE	9.7
1	B	314	TRP	7.9
1	B	313	ILE	7.8
1	A	233	VAL	7.5
1	D	229	LYS	7.5
1	B	218	TYR	7.2
1	E	233	VAL	7.0
1	D	260	ASP	6.8
1	A	218	TYR	6.7
1	E	262	PRO	6.5
1	D	256	SER	6.5
1	E	217	SER	6.3
1	B	233	VAL	6.1
1	D	259	PRO	5.9
1	C	313	ILE	5.8
1	C	299	THR	5.7
1	C	300	CYS	5.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	301	ALA	5.2
1	D	262	PRO	5.1
1	D	263	PRO	5.0
1	D	214	TYR	4.9
1	C	314	TRP	4.9
1	D	217	SER	4.7
1	E	256	SER	4.7
1	E	300	CYS	4.7
1	A	219	THR	4.6
1	E	292	VAL	4.6
1	A	313	ILE	4.5
1	D	234	ALA	4.4
1	E	216	GLY	4.2
1	B	214	TYR	4.1
1	D	233	VAL	4.0
1	E	298	GLY	3.9
1	C	218	TYR	3.9
1	A	262	PRO	3.9
1	E	263	PRO	3.9
1	A	234	ALA	3.8
1	D	258	ASP	3.8
1	A	259	PRO	3.8
1	E	301	ALA	3.7
1	E	258	ASP	3.7
1	A	315	THR	3.7
1	A	263	PRO	3.7
1	A	299	THR	3.7
1	A	317	TYR	3.6
1	A	260	ASP	3.4
1	D	323	CYS	3.4
1	B	234	ALA	3.4
1	C	229	LYS	3.4
1	D	275	ASN	3.3
1	A	220	ILE	3.3
1	B	323	CYS	3.3
1	E	255	GLU	3.3
1	C	322	PHE	3.3
1	D	292	VAL	3.3
1	C	275	ASN	3.2
1	A	323	CYS	3.1
1	C	323	CYS	3.1
1	A	256	SER	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	214	TYR	3.0
1	A	314	TRP	2.9
1	E	257	ARG	2.9
1	C	235	ARG	2.8
1	E	254	ASN	2.8
1	A	298	GLY	2.8
1	B	291	MET	2.8
1	C	271	TYR	2.8
1	A	322	PHE	2.8
1	B	292	VAL	2.8
1	A	289	PHE	2.7
1	E	323	CYS	2.7
1	B	321	VAL	2.7
1	A	302	PHE	2.7
1	E	288	GLY	2.7
1	E	232	ARG	2.7
1	A	321	VAL	2.6
1	C	292	VAL	2.6
1	E	270	TYR	2.6
1	A	208	GLY	2.6
1	C	231	ARG	2.6
1	D	261	ARG	2.6
1	E	299	THR	2.6
1	C	301	ALA	2.5
1	A	209	TYR	2.5
1	E	204	SER	2.5
1	E	317	TYR	2.5
1	C	297	THR	2.5
1	D	291	MET	2.5
1	E	289	PHE	2.5
1	A	258	ASP	2.5
1	C	289	PHE	2.5
1	B	293	ALA	2.4
1	B	217	SER	2.4
1	A	284	LEU	2.4
1	D	235	ARG	2.4
1	B	294	CYS	2.4
1	E	293	ALA	2.4
1	B	322	PHE	2.4
1	E	205	ARG	2.4
1	A	291	MET	2.4
1	E	291	MET	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	257	ARG	2.3
1	C	321	VAL	2.3
1	B	209	TYR	2.2
1	E	271	TYR	2.2
1	A	285	SER	2.2
1	B	235	ARG	2.2
1	D	317	TYR	2.2
1	A	290	HIS	2.2
1	C	234	ALA	2.2
1	E	290	HIS	2.2
1	B	317	TYR	2.2
1	C	288	GLY	2.2
1	C	232	ARG	2.2
1	C	315	THR	2.1
1	A	293	ALA	2.1
1	B	320	TYR	2.1
1	E	294	CYS	2.1
1	D	299	THR	2.1
1	B	312	LYS	2.1
1	D	253	LEU	2.1
1	C	274	PHE	2.0
1	D	264	GLU	2.0
1	B	290	HIS	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 monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.