Full wwPDB X-ray Structure Validation Report (i)

Apr 22, 2025 – 03:46 AM EDT

PDB ID	:	5TCS / pdb_00005tcs
Title	:	Crystal structure of a Dwarf Ndc80 Tetramer
Authors	:	Valverde, R.; Harrison, S.C.
Deposited on	:	2016-09-15
Resolution	:	2.83 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:

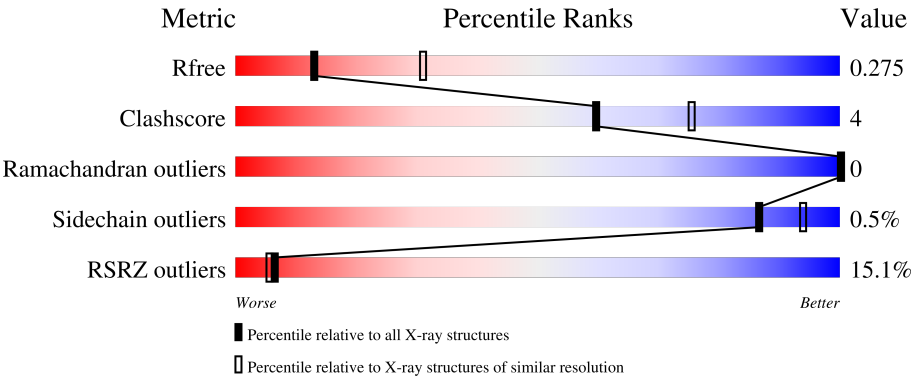
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

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 2.83 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}	164625	1367 (2.86-2.82)
Clashscore	180529	1455 (2.86-2.82)
Ramachandran outliers	177936	1422 (2.86-2.82)
Sidechain outliers	177891	1423 (2.86-2.82)
RSRZ outliers	164620	1368 (2.86-2.82)

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	277	6% 83% 12%
2	B	215	14% 90% 7%
3	C	100	26% 88% 9%
4	D	115	23% 85% 14%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 11358 atoms, of which 5650 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 Kinetochore protein NDC80.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S	Se			
1	A	266	4501	1437	2264	375	419	1	5	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	111	SER	-	expression tag	UNP P40460
A	112	ASN	-	expression tag	UNP P40460
A	113	ALA	-	expression tag	UNP P40460

- Molecule 2 is a protein called Kinetochore protein NUF2.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S	Se			
2	B	211	3420	1093	1691	280	344	4	8	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-17	SER	-	expression tag	UNP P33895
B	-16	ASN	-	expression tag	UNP P33895
B	-15	ALA	-	expression tag	UNP P33895
B	-14	SER	-	expression tag	UNP P33895
B	-13	ILE	-	expression tag	UNP P33895
B	-12	PHE	-	expression tag	UNP P33895
B	-11	LYS	-	expression tag	UNP P33895
B	-10	ASP	-	expression tag	UNP P33895
B	-9	LEU	-	expression tag	UNP P33895
B	-8	GLU	-	expression tag	UNP P33895
B	-7	ALA	-	expression tag	UNP P33895
B	-6	LEU	-	expression tag	UNP P33895
B	-5	SER	-	expression tag	UNP P33895

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-4	PHE	-	expression tag	UNP P33895
B	-3	GLN	-	expression tag	UNP P33895
B	-2	SER	-	expression tag	UNP P33895
B	-1	ASN	-	expression tag	UNP P33895
B	0	ALA	-	expression tag	UNP P33895

- Molecule 3 is a protein called Kinetochore protein SPC24.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	97	Total	C	H	N	O	Se	0	0	0
			1626	511	815	141	158	1			

- Molecule 4 is a protein called Kinetochore protein SPC25.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
4	D	114	Total	C	H	N	O	S	Se	0	0	0
			1771	559	880	164	164	1	3			

- Molecule 5 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Mg	0	0
			1	1		

- Molecule 6 is water.

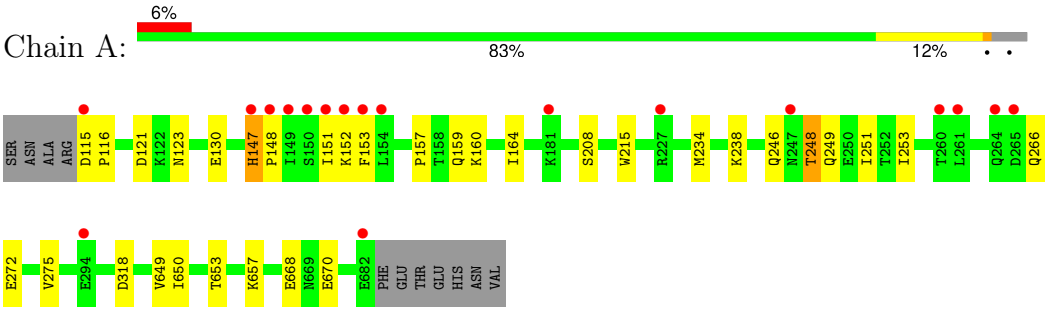
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	19	Total	O	0	0
			19	19		
6	B	15	Total	O	0	0
			15	15		
6	D	5	Total	O	0	0
			5	5		

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.

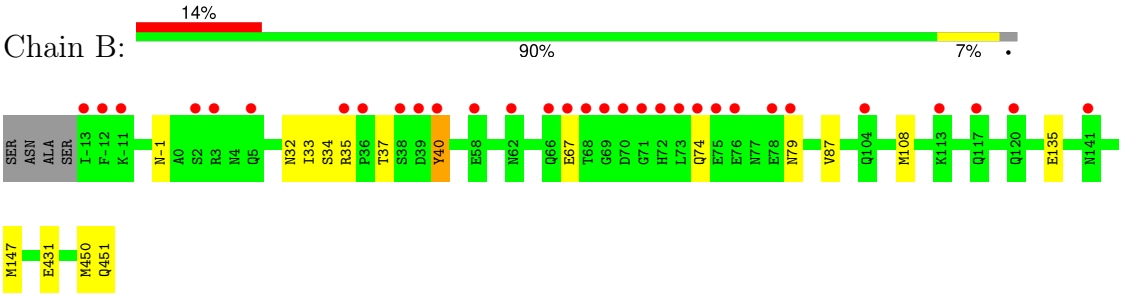
• Molecule 1: Kinetochore protein NDC80

Chain A:



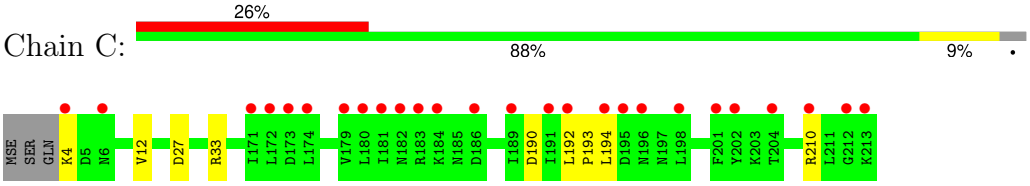
• Molecule 2: Kinetochore protein NUF2

Chain B:



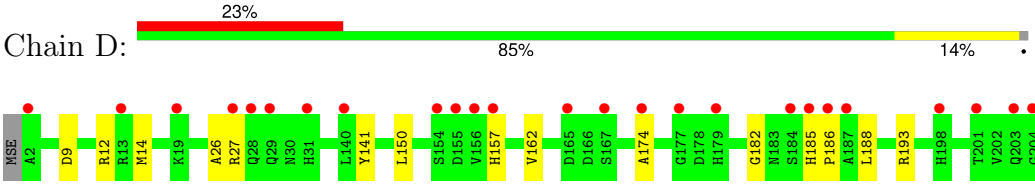
• Molecule 3: Kinetochore protein SPC24

Chain C:



• Molecule 4: Kinetochore protein SPC25

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	169.38Å 186.59Å 122.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.75 – 2.83 41.75 – 2.83	Depositor EDS
% Data completeness (in resolution range)	99.3 (41.75-2.83) 99.9 (41.75-2.83)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.39 (at 2.69Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, R_{free}	0.239 , 0.266 0.249 , 0.275	Depositor DCC
R_{free} test set	2334 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	80.2	Xtriage
Anisotropy	0.272	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 67.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	11358	wwPDB-VP
Average B, all atoms (Å^2)	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.13% 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

5.1 Standard geometry

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

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	0.27	0/2275	0.41	0/3055
2	B	0.27	0/1746	0.41	0/2338
3	C	0.24	0/820	0.46	0/1101
4	D	0.25	0/904	0.44	0/1220
All	All	0.26	0/5745	0.42	0/7714

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
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	147	HIS	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	2237	2264	2264	28	1
2	B	1729	1691	1693	15	1
3	C	811	815	815	6	1
4	D	891	880	880	9	0
5	B	1	0	0	0	0
6	A	19	0	0	0	0
6	B	15	0	0	0	0
6	D	5	0	0	1	0
All	All	5708	5650	5652	48	2

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 4.

All (48) 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)
4:D:157:HIS:HA	4:D:174:ALA:HB2	1.79	0.63
3:C:194:LEU:O	3:C:194:LEU:HG	2.00	0.61
1:A:668:GLU:OE1	3:C:33:ARG:NH1	2.38	0.56
3:C:190:ASP:OD2	3:C:210:ARG:NH2	2.40	0.54
1:A:657:LYS:NZ	3:C:27:ASP:OD2	2.35	0.54
4:D:26:ALA:O	6:D:301:HOH:O	2.18	0.54
2:B:33:ILE:O	2:B:33:ILE:HG22	2.09	0.52
1:A:253:ILE:HD13	1:A:270:ARG:HH12	1.75	0.52
1:A:267:ARG:NH2	1:A:318:ASP:OD2	2.41	0.52
1:A:115:ASP:HB3	1:A:116:PRO:HD3	1.92	0.50
1:A:159:GLN:NE2	1:A:208:SER:O	2.44	0.50
1:A:147:HIS:HB2	1:A:148:PRO:CD	2.41	0.50
1:A:272:GLU:HA	2:B:147:MSE:HE1	1.95	0.49
4:D:182:GLY:O	4:D:193:ARG:NH2	2.42	0.48
1:A:670:GLU:HG2	4:D:27:ARG:CZ	2.44	0.48
2:B:32:ASN:HB3	2:B:40:TYR:CZ	2.48	0.48
1:A:275:VAL:HB	2:B:147:MSE:HE1	1.96	0.47
1:A:160:LYS:O	1:A:164:ILE:HD13	2.15	0.46
1:A:238:LYS:CE	2:B:67:GLU:HG2	2.45	0.46
2:B:35:ARG:O	2:B:37:THR:N	2.44	0.45
2:B:33:ILE:HG12	2:B:40:TYR:OH	2.17	0.45
2:B:34:SER:O	2:B:37:THR:OG1	2.26	0.45
1:A:148:PRO:C	1:A:151:ILE:HG22	2.37	0.44
3:C:192:LEU:HD12	3:C:193:PRO:HD2	2.00	0.44
2:B:79:ASN:O	2:B:79:ASN:OD1	2.36	0.44
1:A:157:PRO:HG3	1:A:215:TRP:CE2	2.53	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:649:VAL:CG2	1:A:650:ILE:N	2.80	0.43
2:B:431:GLU:HG3	3:C:12:VAL:HG13	2.00	0.43
1:A:248:THR:O	1:A:251:ILE:HG22	2.19	0.43
4:D:185:HIS:HB3	4:D:186:PRO:HD3	2.01	0.43
4:D:188:LEU:HD21	4:D:217:LEU:HD23	2.01	0.42
1:A:121:ASP:OD1	1:A:123:ASN:N	2.53	0.42
1:A:246:GLN:O	1:A:249:GLN:HG2	2.18	0.42
2:B:450:MSE:O	2:B:451:GLN:HB2	2.19	0.42
1:A:115:ASP:CB	1:A:116:PRO:HD3	2.50	0.42
4:D:9:ASP:O	4:D:12:ARG:HB2	2.20	0.42
1:A:653:THR:HA	4:D:14:MSE:HE1	2.02	0.42
2:B:87:VAL:HG11	2:B:108:MSE:HE1	2.02	0.42
1:A:152:LYS:HG2	1:A:153:PHE:HD2	1.84	0.41
1:A:275:VAL:HB	2:B:147:MSE:CE	2.50	0.41
1:A:266:GLN:O	1:A:270:ARG:HG2	2.20	0.41
2:B:74:GLN:O	2:B:74:GLN:HG2	2.20	0.41
1:A:147:HIS:HB3	1:A:164:ILE:HG12	2.02	0.41
1:A:148:PRO:C	1:A:151:ILE:CG2	2.89	0.41
1:A:148:PRO:HB2	1:A:151:ILE:CG2	2.50	0.41
1:A:238:LYS:HE2	2:B:67:GLU:HG2	2.02	0.41
1:A:234:MSE:HG3	1:A:238:LYS:HZ1	1.86	0.40
4:D:150:LEU:HD12	4:D:162:VAL:HG21	2.03	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:-1:ASN:H	2:B:135:GLU:OE2[4_545]	1.52	0.08
1:A:130:GLU:OE2	3:C:4:LYS:HZ3[3_655]	1.56	0.04

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	264/277 (95%)	234 (89%)	30 (11%)	0	100	100
2	B	209/215 (97%)	198 (95%)	11 (5%)	0	100	100
3	C	95/100 (95%)	92 (97%)	3 (3%)	0	100	100
4	D	112/115 (97%)	104 (93%)	8 (7%)	0	100	100
All	All	680/707 (96%)	628 (92%)	52 (8%)	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	256/261 (98%)	255 (100%)	1 (0%)	89	95
2	B	198/193 (103%)	197 (100%)	1 (0%)	86	94
3	C	92/93 (99%)	92 (100%)	0	100	100
4	D	95/92 (103%)	94 (99%)	1 (1%)	70	86
All	All	641/639 (100%)	638 (100%)	3 (0%)	86	94

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	248	THR
2	B	40	TYR
4	D	141	TYR

Sometimes 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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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.

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

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^{th} 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	261/277 (94%)	0.34	18 (6%) 24 20	39, 76, 129, 268	0
2	B	203/215 (94%)	0.61	31 (15%) 6 6	30, 67, 175, 220	0
3	C	96/100 (96%)	1.24	26 (27%) 2 1	58, 134, 207, 228	0
4	D	111/115 (96%)	1.28	26 (23%) 2 2	71, 109, 178, 196	0
All	All	671/707 (94%)	0.71	101 (15%) 6 5	30, 82, 189, 268	0

All (101) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	156	VAL	12.4
1	A	149	ILE	11.4
2	B	-13	ILE	9.5
1	A	147	HIS	8.2
4	D	221	LEU	7.9
2	B	35	ARG	7.9
4	D	157	HIS	7.2
1	A	150	SER	7.1
2	B	-12	PHE	6.8
4	D	155	ASP	6.6
2	B	-11	LYS	6.6
4	D	185	HIS	6.3
2	B	68	THR	5.6
1	A	154	LEU	5.3
2	B	75	GLU	5.3
4	D	2	ALA	5.2
3	C	172	LEU	5.2
2	B	5	GLN	5.1
2	B	39	ASP	5.1
2	B	72	HIS	5.0
4	D	186	PRO	5.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	C	201	PHE	4.9
3	C	174	LEU	4.8
2	B	73	LEU	4.8
2	B	67	GLU	4.7
2	B	69	GLY	4.5
4	D	28	GLN	4.3
3	C	202	TYR	4.2
3	C	194	LEU	4.2
3	C	180	LEU	4.1
2	B	141	ASN	4.0
1	A	261	LEU	4.0
4	D	187	ALA	4.0
2	B	40	TYR	3.9
1	A	151	ILE	3.7
4	D	31	HIS	3.7
3	C	4	LYS	3.7
3	C	213	LYS	3.6
2	B	36	PRO	3.6
2	B	104	GLN	3.6
2	B	74	GLN	3.6
1	A	115	ASP	3.5
4	D	184	SER	3.5
1	A	153	PHE	3.5
3	C	6	ASN	3.5
3	C	182	ASN	3.4
3	C	198	LEU	3.4
2	B	66	GLN	3.3
2	B	120	GLN	3.3
1	A	148	PRO	3.3
3	C	183	ARG	3.2
3	C	173	ASP	3.1
2	B	71	GLY	3.1
4	D	198	HIS	3.1
2	B	79	ASN	3.1
1	A	260	THR	3.1
4	D	179	HIS	3.0
3	C	179	VAL	3.0
2	B	78	GLU	3.0
3	C	192	LEU	2.9
1	A	247	ASN	2.8
2	B	2	SER	2.8
4	D	204	GLY	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
4	D	27	ARG	2.7
1	A	152	LYS	2.7
1	A	682	GLU	2.7
2	B	113	LYS	2.6
1	A	264	GLN	2.6
4	D	177	GLY	2.6
4	D	174	ALA	2.6
2	B	3	ARG	2.6
2	B	117	GLN	2.5
3	C	204	THR	2.5
4	D	201	THR	2.5
2	B	62	ASN	2.4
3	C	171	ILE	2.4
1	A	294	GLU	2.4
3	C	212	GLY	2.4
4	D	154	SER	2.4
4	D	29	GLN	2.3
3	C	210	ARG	2.3
4	D	13	ARG	2.3
3	C	189	ILE	2.3
4	D	165	ASP	2.3
3	C	181	ILE	2.3
2	B	38	SER	2.3
1	A	227	ARG	2.2
3	C	196	ASN	2.2
3	C	191	ILE	2.2
4	D	203	GLN	2.2
4	D	140	LEU	2.2
2	B	58	GLU	2.1
1	A	265	ASP	2.1
3	C	186	ASP	2.1
3	C	184	LYS	2.1
3	C	195	ASP	2.1
2	B	70	ASP	2.0
4	D	19	LYS	2.0
4	D	167	SER	2.0
2	B	76	GLU	2.0
1	A	181	LYS	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)

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^{th} 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
5	MG	B	501	1/1	0.96	0.41	27,27,27,27	0

6.5 Other polymers (i)

There are no such residues in this entry.