Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 11:27 AM GMT

PDB ID	:	3P1J
Title	:	Crystal structure of human GTPase IMAP family member 2 in the nucleotide-free state
Authors	:	Shen, L.; Tempel, W.; Tong, Y.; Guan, X.; Nedyalkova, L.; Wernimont, A.K.; Mackenzie, F.; Arrowsmith, C.H.; Edwards, A.M.; Bountra, C.; Weigelt, J.; Bochkarev, A.; Andrews, D.W.; Park, H.; Structural Genomics Consortium (SGC)
Deposited on	:	2010-09-30
Resolution	:	2.58 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

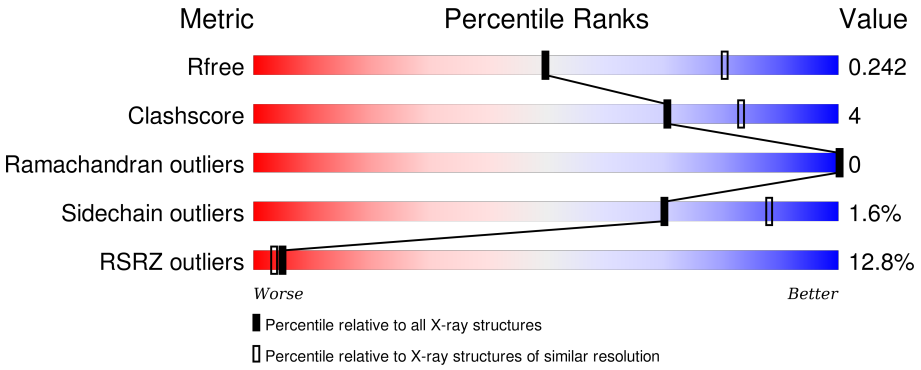
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.58 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}	91344	2636 (2.60-2.56)
Clashscore	102246	3003 (2.60-2.56)
Ramachandran outliers	100387	2956 (2.60-2.56)
Sidechain outliers	100360	2956 (2.60-2.56)
RSRZ outliers	91569	2642 (2.60-2.56)

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 <=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	209	9% 82% 8% 9%
1	B	209	19% 73% 13% 14%
1	C	209	10% 79% 10% 11%
1	D	209	9% 84% 8% 8%

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	UNX	A	1	-	-	-	X
2	UNX	A	12	-	-	-	X
2	UNX	A	9	-	-	-	X
2	UNX	B	15	-	-	-	X
2	UNX	B	2	-	-	-	X
2	UNX	B	3	-	-	-	X
2	UNX	D	17	-	-	-	X
2	UNX	D	7	-	-	-	X
2	UNX	D	8	-	-	-	X

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5399 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 GTPase IMAP family member 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	190	Total	C	N	O	S	0	0	0
			1392	880	238	262	12			
1	B	180	Total	C	N	O	S	0	0	0
			1286	809	224	242	11			
1	C	186	Total	C	N	O	S	0	0	0
			1314	828	227	248	11			
1	D	193	Total	C	N	O	S	0	0	0
			1388	878	242	258	10			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	18	GLY	-	EXPRESSION TAG	UNP Q9UG22
B	18	GLY	-	EXPRESSION TAG	UNP Q9UG22
C	18	GLY	-	EXPRESSION TAG	UNP Q9UG22
D	18	GLY	-	EXPRESSION TAG	UNP Q9UG22

- Molecule 2 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

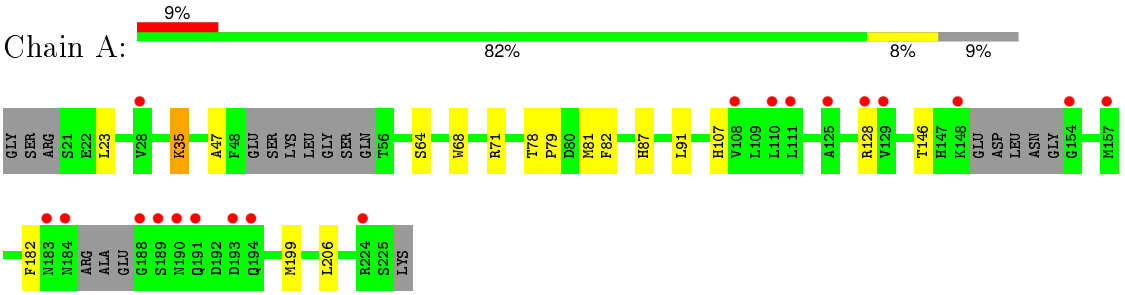
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	4	Total	X	0	0
			4	4		
2	A	6	Total	X	0	0
			6	6		
2	D	8	Total	X	0	0
			8	8		
2	C	1	Total	X	0	0
			1	1		

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.

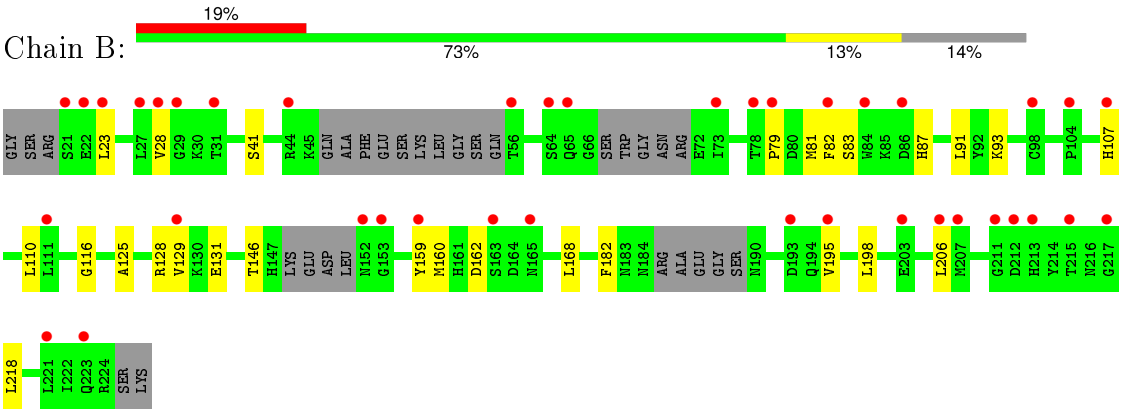
- Molecule 1: GTPase IMAP family member 2

Chain A:



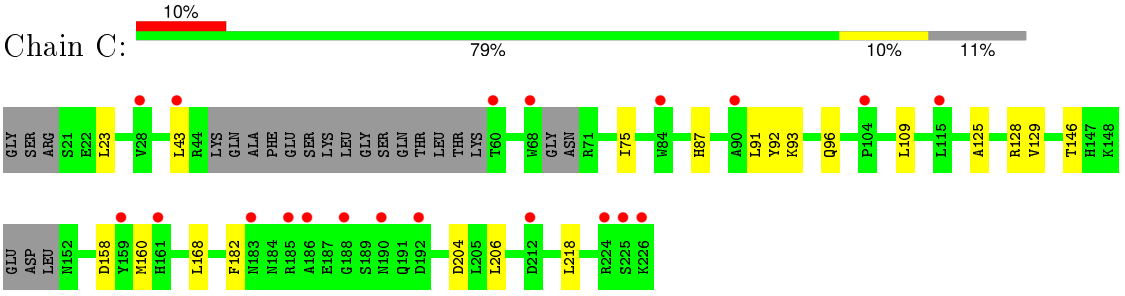
- Molecule 1: GTPase IMAP family member 2

Chain B:



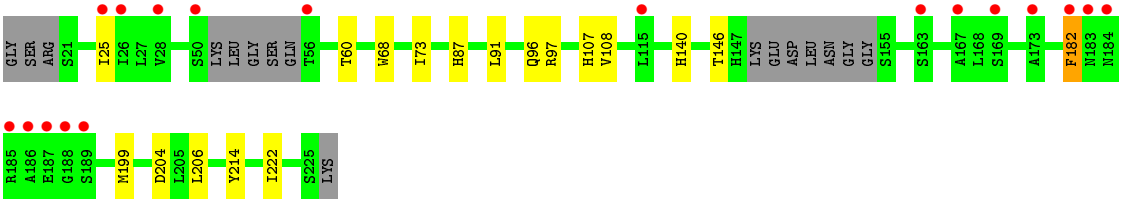
- Molecule 1: GTPase IMAP family member 2

Chain C:



- Molecule 1: GTPase IMAP family member 2

Chain D:


GLY
SER
ARG
S21
I25
I26
L27
V28
S50
LYS
LEU
GLY
SER
GLN
T56
T60
W68
I73
H87
L91
Q96
R97
H107
V108
L115
H140
T146
H147
LYS
GLU
ASP
LEU
ASN
GLY
GLY
S155
S163
A167
L168
S169
A173
F182
N183
N184
R185
A186
E187
G188
S189
M199
D204
L205
L206
Y214
I222
S225
LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	132.47Å 132.47Å 84.16Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.58 29.76 – 2.58	Depositor EDS
% Data completeness (in resolution range)	100.0 (30.00-2.58) 100.0 (29.76-2.58)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.01 (at 2.57Å)	Xtriage
Refinement program	BUSTER 2.8.0	Depositor
R, R_{free}	0.205 , 0.230 0.216 , 0.242	Depositor DCC
R_{free} test set	1005 reflections (3.93%)	DCC
Wilson B-factor (Å^2)	59.9	Xtriage
Anisotropy	0.323	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 66.7	EDS
Estimated twinning fraction	0.039 for h,-h-k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 26556 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5399	wwPDB-VP
Average B, all atoms (Å^2)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: UNX

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	0.51	0/1411	0.64	0/1907
1	B	0.48	0/1302	0.62	0/1762
1	C	0.49	0/1333	0.60	0/1810
1	D	0.49	0/1407	0.63	0/1907
All	All	0.49	0/5453	0.62	0/7386

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

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	1392	0	1298	10	0
1	B	1286	0	1157	17	0
1	C	1314	0	1162	12	0
1	D	1388	0	1281	10	0
2	A	6	0	0	0	0
2	B	4	0	0	0	0
2	C	1	0	0	0	0
2	D	8	0	0	0	0
All	All	5399	0	4898	45	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 4.

All (45) 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:91:LEU:HD21	1:C:128:ARG:HH12	1.48	0.77
1:A:23:LEU:HD13	1:A:206:LEU:HD11	1.75	0.69
1:B:23:LEU:HD13	1:B:206:LEU:HD11	1.76	0.67
1:B:91:LEU:HD21	1:B:128:ARG:HH22	1.69	0.57
1:D:25:ILE:HG12	1:D:108:VAL:HG13	1.87	0.57
1:B:218:LEU:HD13	1:C:93:LYS:HA	1.88	0.56
1:C:87:HIS:CD2	1:C:91:LEU:HD23	2.41	0.56
1:D:96:GLN:OE1	1:D:222:ILE:HD11	2.08	0.54
1:B:82:PHE:O	1:B:128:ARG:HG3	2.09	0.53
1:B:93:LYS:HA	1:C:218:LEU:HD13	1.90	0.53
1:D:87:HIS:CD2	1:D:91:LEU:HD23	2.44	0.52
1:A:35:LYS:NZ	1:A:78:THR:O	2.42	0.52
1:A:87:HIS:CD2	1:A:91:LEU:HD23	2.44	0.52
1:C:43:LEU:HG	1:C:75:ILE:HD12	1.91	0.52
1:A:146:THR:HA	1:A:182:PHE:O	2.10	0.51
1:B:87:HIS:CD2	1:B:91:LEU:HD23	2.45	0.51
1:A:91:LEU:HD21	1:A:128:ARG:HH22	1.76	0.50
1:D:60:THR:HB	1:D:97:ARG:NH1	2.27	0.50
1:B:28:VAL:HG21	1:B:129:VAL:HG21	1.94	0.50
1:A:82:PHE:O	1:A:128:ARG:HG3	2.13	0.49
1:B:218:LEU:HD11	1:C:96:GLN:CB	2.42	0.49
1:C:23:LEU:HD13	1:C:206:LEU:HD11	1.94	0.48
1:C:125:ALA:O	1:C:129:VAL:HG23	2.14	0.47
1:A:107:HIS:HB3	1:A:206:LEU:HD21	1.96	0.47
1:B:107:HIS:HB3	1:B:206:LEU:HD21	1.96	0.47
1:C:146:THR:HA	1:C:182:PHE:O	2.15	0.47
1:D:68:TRP:HB3	1:D:73:ILE:HD11	1.97	0.46
1:C:91:LEU:HD21	1:C:128:ARG:NH1	2.23	0.46
1:D:68:TRP:CD1	1:D:199:MET:HG3	2.50	0.46
1:D:60:THR:HB	1:D:97:ARG:HH12	1.80	0.45
1:B:160:MET:HE1	1:B:168:LEU:HD21	1.97	0.45
1:B:146:THR:HA	1:B:182:PHE:O	2.17	0.45
1:B:131:GLU:HB3	1:C:92:TYR:HE2	1.81	0.44
1:A:68:TRP:CD1	1:A:199:MET:HG3	2.53	0.44
1:B:79:PRO:HB2	1:B:81:MET:HG2	1.99	0.44
1:B:110:LEU:HD21	1:B:198:LEU:HD11	1.99	0.44
1:B:41:SER:HB3	1:B:195:VAL:HG22	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:160:MET:HE1	1:C:168:LEU:HD21	2.01	0.42
1:D:140:HIS:CG	1:D:214:TYR:HB2	2.56	0.41
1:A:47:ALA:HB1	1:A:64:SER:CB	2.51	0.41
1:A:79:PRO:HB2	1:A:81:MET:HG2	2.03	0.41
1:B:83:SER:HB2	1:B:125:ALA:HB2	2.02	0.41
1:B:116:GLY:HA2	1:B:159:TYR:CE2	2.56	0.40
1:D:146:THR:HA	1:D:182:PHE:O	2.21	0.40
1:D:107:HIS:HB3	1:D:206:LEU:HD21	2.04	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	182/209 (87%)	181 (100%)	1 (0%)	0	100	100
1	B	170/209 (81%)	168 (99%)	2 (1%)	0	100	100
1	C	178/209 (85%)	176 (99%)	2 (1%)	0	100	100
1	D	187/209 (90%)	185 (99%)	2 (1%)	0	100	100
All	All	717/836 (86%)	710 (99%)	7 (1%)	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	133/177 (75%)	131 (98%)	2 (2%)	72 89
1	B	116/177 (66%)	115 (99%)	1 (1%)	84 94
1	C	116/177 (66%)	113 (97%)	3 (3%)	54 78
1	D	126/177 (71%)	124 (98%)	2 (2%)	70 88
All	All	491/708 (69%)	483 (98%)	8 (2%)	70 88

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	35	LYS
1	A	71	ARG
1	B	162	ASP
1	C	109	LEU
1	C	158	ASP
1	C	204	ASP
1	D	182	PHE
1	D	204	ASP

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)

Of 19 ligands modelled in this entry, 19 are unknown - 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

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^{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	190/209 (90%)	0.36	19 (10%)	9	7	41, 59, 88, 109	0
1	B	180/209 (86%)	0.88	39 (21%)	1	0	47, 71, 99, 134	0
1	C	186/209 (88%)	0.53	20 (10%)	8	5	49, 77, 109, 132	0
1	D	193/209 (92%)	0.45	18 (9%)	11	8	41, 63, 101, 114	0
All	All	749/836 (89%)	0.55	96 (12%)	5	3	41, 68, 102, 134	0

All (96) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	188	GLY	6.3
1	D	186	ALA	6.1
1	B	206	LEU	5.3
1	D	188	GLY	4.7
1	B	29	GLY	4.6
1	A	189	SER	4.6
1	C	188	GLY	4.4
1	A	190	ASN	4.4
1	D	185	ARG	4.3
1	B	215	THR	4.2
1	B	84	TRP	4.2
1	D	183	ASN	4.2
1	B	65	GLN	4.1
1	C	90	ALA	4.1
1	B	28	VAL	4.0
1	B	73	ILE	3.9
1	B	211	GLY	3.8
1	B	21	SER	3.8
1	D	173	ALA	3.7
1	A	193	ASP	3.6
1	B	223	GLN	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	82	PHE	3.4
1	B	78	THR	3.2
1	C	104	PRO	3.2
1	B	212	ASP	3.2
1	D	56	THR	3.2
1	A	125	ALA	3.1
1	C	192	ASP	3.1
1	C	68	TRP	3.1
1	A	191	GLN	3.1
1	B	163	SER	3.1
1	D	184	ASN	3.1
1	B	56	THR	3.0
1	C	225	SER	2.9
1	C	84	TRP	2.9
1	C	183	ASN	2.9
1	B	44	ARG	2.9
1	A	129	VAL	2.9
1	B	129	VAL	2.9
1	D	169	SER	2.9
1	B	165	ASN	2.9
1	B	22	GLU	2.9
1	A	154	GLY	2.7
1	A	194	GLN	2.7
1	D	187	GLU	2.7
1	C	190	ASN	2.7
1	B	153	GLY	2.7
1	B	27	LEU	2.7
1	C	43	LEU	2.6
1	C	186	ALA	2.6
1	A	128	ARG	2.6
1	A	110	LEU	2.6
1	B	86	ASP	2.6
1	B	64	SER	2.6
1	B	159	TYR	2.6
1	B	111	LEU	2.6
1	B	152	ASN	2.6
1	B	213	HIS	2.6
1	C	224	ARG	2.5
1	D	50	SER	2.5
1	B	79	PRO	2.5
1	C	28	VAL	2.4
1	B	23	LEU	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	107	HIS	2.4
1	A	111	LEU	2.4
1	A	148	LYS	2.4
1	A	183	ASN	2.4
1	D	28	VAL	2.3
1	A	108	VAL	2.3
1	B	31	THR	2.3
1	B	104	PRO	2.3
1	D	189	SER	2.3
1	B	221	LEU	2.3
1	C	60	THR	2.2
1	C	226	LYS	2.2
1	D	167	ALA	2.2
1	D	25	ILE	2.2
1	C	212	ASP	2.2
1	B	203	GLU	2.2
1	B	98	CYS	2.2
1	A	184	ASN	2.2
1	C	161	HIS	2.2
1	C	159	TYR	2.1
1	B	217	GLY	2.1
1	D	182	PHE	2.1
1	B	193	ASP	2.1
1	C	185	ARG	2.1
1	C	115	LEU	2.1
1	A	157	MET	2.1
1	B	207	MET	2.1
1	D	115	LEU	2.1
1	D	163	SER	2.0
1	A	224	ARG	2.0
1	D	26	ILE	2.0
1	A	28	VAL	2.0
1	B	195	VAL	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. 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^{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	LLDF	B-factors(Å^2)	Q<0.9
2	UNX	A	1	1/1	0.86	0.66	14.45	50,50,50,50	0
2	UNX	D	17	1/1	0.66	0.54	9.97	30,30,30,30	0
2	UNX	A	12	1/1	0.89	0.58	8.78	30,30,30,30	0
2	UNX	B	3	1/1	0.98	0.86	7.14	51,51,51,51	0
2	UNX	D	8	1/1	0.94	0.44	6.55	47,47,47,47	0
2	UNX	A	9	1/1	0.94	0.34	4.98	30,30,30,30	0
2	UNX	D	7	1/1	0.78	0.34	4.91	53,53,53,53	0
2	UNX	B	2	1/1	0.83	0.71	4.52	51,51,51,51	0
2	UNX	B	15	1/1	0.85	0.48	3.79	30,30,30,30	0
2	UNX	A	10	1/1	0.59	0.17	0.81	30,30,30,30	0
2	UNX	C	16	1/1	0.98	0.17	-0.23	30,30,30,30	0
2	UNX	B	14	1/1	0.66	0.20	-0.32	30,30,30,30	0
2	UNX	A	13	1/1	0.90	0.19	-	30,30,30,30	0
2	UNX	A	11	1/1	0.82	0.28	-	30,30,30,30	0
2	UNX	D	5	1/1	0.78	0.28	-	46,46,46,46	0
2	UNX	D	6	1/1	0.71	0.41	-	61,61,61,61	0
2	UNX	D	227	1/1	0.97	0.53	-	30,30,30,30	0
2	UNX	D	228	1/1	0.86	0.37	-	30,30,30,30	0
2	UNX	D	229	1/1	0.94	0.14	-	30,30,30,30	0

6.5 Other polymers (i)

There are no such residues in this entry.