Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 01:59 AM GMT

PDB ID	:	2F4M
Title	:	The Mouse PNGase-HR23 Complex Reveals a Complete Remodulation of the Protein-Protein Interface Compared to its Yeast Orthologs
Authors	:	Zhao, G.; Zhou, X.; Wang, L.; Kisker, C.; Lennarz, W.J.; Schindelin, H.
Deposited on	:	2005-11-23
Resolution	:	1.85 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:

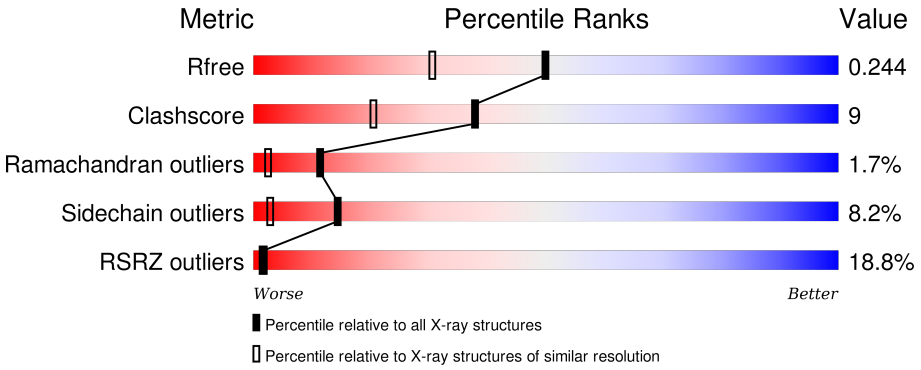
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 1.85 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	1745 (1.86-1.86)
Clashscore	102246	1898 (1.86-1.86)
Ramachandran outliers	100387	1875 (1.86-1.86)
Sidechain outliers	100360	1875 (1.86-1.86)
RSRZ outliers	91569	1747 (1.86-1.86)

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	295	21% 77% 18%
2	B	61	8% 74% 18% 7%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 3210 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 peptide N-glycanase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	295	Total	C	N	O	S	0	0	0
			2448	1544	441	448	15			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	451	LEU	-	CLONING ARTIFACT	UNP Q9JI78
A	452	GLU	-	CLONING ARTIFACT	UNP Q9JI78
A	453	HIS	-	EXPRESSION TAG	UNP Q9JI78
A	454	HIS	-	EXPRESSION TAG	UNP Q9JI78
A	455	HIS	-	EXPRESSION TAG	UNP Q9JI78
A	456	HIS	-	EXPRESSION TAG	UNP Q9JI78
A	457	HIS	-	EXPRESSION TAG	UNP Q9JI78
A	458	HIS	-	EXPRESSION TAG	UNP Q9JI78

- Molecule 2 is a protein called UV excision repair protein RAD23 homolog B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	61	Total	C	N	O	S	0	0	0
			505	318	95	90	2			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	333	GLY	-	CLONING ARTIFACT	UNP P54728

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Cl 1 1	0	0

- Molecule 5 is water.

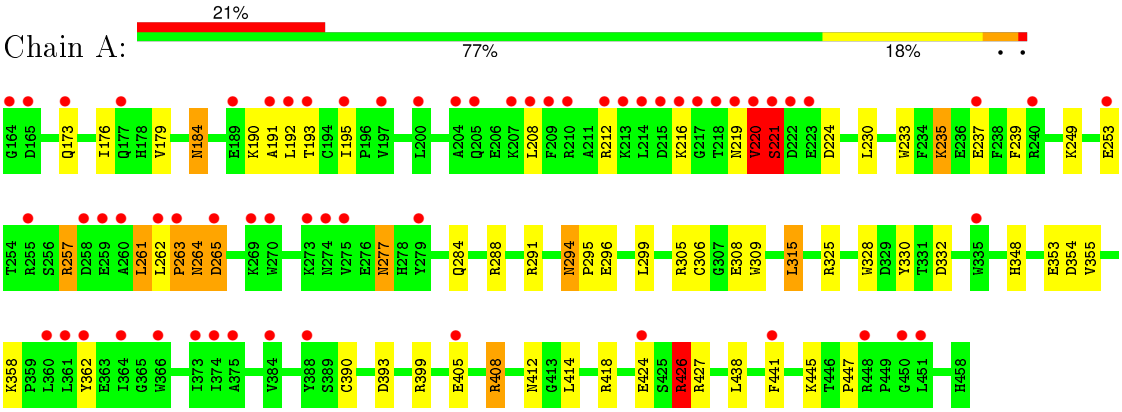
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	217	Total O 217 217	0	0
5	B	38	Total O 38 38	0	0

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 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.

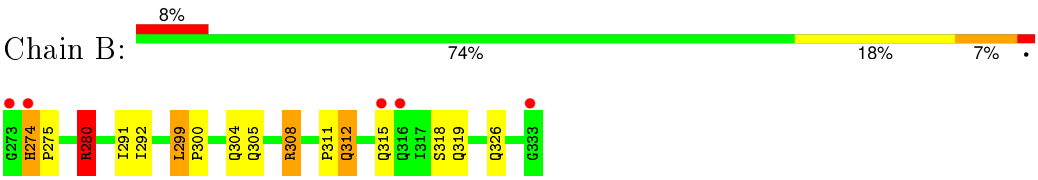
- Molecule 1: peptide N-glycanase

Chain A:



- Molecule 2: UV excision repair protein RAD23 homolog B

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	98.22Å 52.30Å 82.55Å 90.00° 115.33° 90.00°	Depositor
Resolution (Å)	20.00 – 1.85 37.55 – 1.85	Depositor EDS
% Data completeness (in resolution range)	96.0 (20.00-1.85) 96.0 (37.55-1.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	2.47 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.186 , 0.239 0.194 , 0.244	Depositor DCC
R_{free} test set	1578 reflections (5.33%)	DCC
Wilson B-factor (Å^2)	31.6	Xtriage
Anisotropy	0.508	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 56.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 31216 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3210	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.47% 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: ZN, 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	0.80	0/2506	0.82	4/3386 (0.1%)
2	B	0.74	0/516	0.80	3/698 (0.4%)
All	All	0.79	0/3022	0.82	7/4084 (0.2%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	399	ARG	NE-CZ-NH1	5.99	123.30	120.30
2	B	280	ARG	NE-CZ-NH1	5.75	123.18	120.30
2	B	308	ARG	NE-CZ-NH2	-5.72	117.44	120.30
1	A	441	PHE	CB-CG-CD1	5.21	124.45	120.80
1	A	325	ARG	NE-CZ-NH2	-5.20	117.70	120.30
1	A	426	ARG	NE-CZ-NH1	5.15	122.87	120.30
2	B	280	ARG	NE-CZ-NH2	-5.02	117.79	120.30

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	2448	0	2395	41	1
2	B	505	0	501	15	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	1	0	0	0	0
4	A	1	0	0	0	0
5	A	217	0	0	12	3
5	B	38	0	0	1	0
All	All	3210	0	2896	54	3

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 9.

All (54) 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:A:408:ARG:NE	5:A:688:HOH:O	2.05	0.90
1:A:176:ILE:HD11	1:A:414:LEU:HD13	1.58	0.85
2:B:312:GLN:HA	2:B:315:GLN:HE21	1.49	0.78
2:B:315:GLN:O	2:B:319:GLN:NE2	2.16	0.76
1:A:195:ILE:O	5:A:665:HOH:O	2.03	0.74
1:A:328:TRP:CH2	1:A:330:TYR:HD1	2.04	0.74
1:A:192:LEU:O	5:A:665:HOH:O	2.07	0.72
1:A:328:TRP:HH2	1:A:330:TYR:HD1	1.38	0.69
1:A:328:TRP:CH2	1:A:330:TYR:CD1	2.87	0.63
1:A:348:HIS:HD2	1:A:358:LYS:H	1.49	0.61
1:A:291:ARG:NH1	1:A:308:GLU:OE1	2.34	0.60
1:A:332:ASP:HA	5:A:613:HOH:O	2.02	0.59
2:B:304:GLN:NE2	5:B:180:HOH:O	2.37	0.58
1:A:412:ASN:OD1	5:A:683:HOH:O	2.17	0.58
1:A:328:TRP:HH2	1:A:330:TYR:CD1	2.19	0.57
2:B:274:HIS:HB3	2:B:275:PRO:HD3	1.86	0.57
2:B:305:GLN:NE2	2:B:308:ARG:HH11	2.03	0.56
1:A:390:CYS:SG	1:A:447:PRO:HG3	2.46	0.56
1:A:393:ASP:HB3	5:A:700:HOH:O	2.06	0.54
2:B:274:HIS:HB3	2:B:275:PRO:CD	2.38	0.53
2:B:274:HIS:CB	2:B:275:PRO:CD	2.87	0.53
1:A:261:LEU:HD13	1:A:277:ASN:HB2	1.91	0.53
1:A:184:ASN:ND2	1:A:184:ASN:N	2.56	0.53
1:A:184:ASN:N	1:A:184:ASN:HD22	2.07	0.52
1:A:445:LYS:HD3	5:A:529:HOH:O	2.10	0.52
1:A:208:LEU:HD13	1:A:221:SER:O	2.09	0.51
2:B:280:ARG:HH22	2:B:326:GLN:HE22	1.59	0.51
1:A:263:PRO:O	1:A:264:ASN:C	2.50	0.50
1:A:306:CYS:HB2	5:A:669:HOH:O	2.12	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:233:TRP:CD1	1:A:237:GLU:HB3	2.49	0.47
1:A:296:GLU:HA	1:A:299:LEU:HD12	1.96	0.47
1:A:438:LEU:HD11	2:B:291:ILE:HD12	1.95	0.47
2:B:318:SER:HB2	2:B:319:GLN:NE2	2.29	0.47
2:B:299:LEU:HB3	2:B:300:PRO:HD3	1.97	0.47
1:A:405:GLU:HG3	5:A:662:HOH:O	2.14	0.46
1:A:294:ASN:HD22	1:A:295:PRO:HD2	1.79	0.46
1:A:191:ALA:HB2	1:A:315:LEU:HD13	1.98	0.46
1:A:355:VAL:HG13	1:A:358:LYS:HD2	1.97	0.46
1:A:348:HIS:HE1	1:A:362:TYR:OH	1.99	0.45
1:A:220:VAL:O	1:A:221:SER:CB	2.64	0.44
2:B:311:PRO:O	2:B:315:GLN:HG3	2.17	0.44
1:A:257:ARG:HB2	1:A:277:ASN:HB3	1.99	0.44
1:A:249:LYS:NZ	5:A:695:HOH:O	2.50	0.44
1:A:191:ALA:HB2	1:A:315:LEU:CD1	2.49	0.43
2:B:280:ARG:HH22	2:B:326:GLN:NE2	2.17	0.43
1:A:426:ARG:HD3	5:A:567:HOH:O	2.19	0.43
1:A:305:ARG:O	1:A:309:TRP:HD1	2.01	0.42
1:A:262:LEU:HA	1:A:263:PRO:HD3	1.92	0.41
1:A:190:LYS:O	1:A:193:THR:HB	2.19	0.41
1:A:221:SER:HA	1:A:224:ASP:HB2	2.04	0.40
2:B:292:ILE:HD13	2:B:292:ILE:HA	1.93	0.40
1:A:418:ARG:HD3	5:A:612:HOH:O	2.21	0.40
1:A:438:LEU:HD11	2:B:291:ILE:CD1	2.51	0.40
1:A:235:LYS:HG3	1:A:354:ASP:HB2	2.02	0.40

All (3) 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)
5:A:573:HOH:O	5:A:696:HOH:O[4_555]	2.01	0.19
5:A:708:HOH:O	5:A:715:HOH:O[4_546]	2.06	0.14
1:A:353:GLU:OE1	5:A:530:HOH:O[2_656]	2.12	0.08

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	293/295 (99%)	275 (94%)	13 (4%)	5 (2%)	11 2
2	B	59/61 (97%)	57 (97%)	1 (2%)	1 (2%)	11 2
All	All	352/356 (99%)	332 (94%)	14 (4%)	6 (2%)	11 2

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	274	HIS
1	A	220	VAL
1	A	265	ASP
1	A	264	ASN
1	A	221	SER
1	A	263	PRO

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	272/272 (100%)	248 (91%)	24 (9%)	12 2
2	B	57/57 (100%)	54 (95%)	3 (5%)	28 10
All	All	329/329 (100%)	302 (92%)	27 (8%)	14 3

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	173	GLN
1	A	179	VAL
1	A	184	ASN
1	A	212	ARG
1	A	216	LYS
1	A	219	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	220	VAL
1	A	221	SER
1	A	230	LEU
1	A	235	LYS
1	A	239	PHE
1	A	253	GLU
1	A	257	ARG
1	A	261	LEU
1	A	265	ASP
1	A	277	ASN
1	A	284	GLN
1	A	288	ARG
1	A	294	ASN
1	A	315	LEU
1	A	408	ARG
1	A	424	GLU
1	A	426	ARG
1	A	427	ARG
2	B	280	ARG
2	B	299	LEU
2	B	312	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	184	ASN
1	A	219	ASN
1	A	277	ASN
1	A	294	ASN
1	A	344	GLN
1	A	348	HIS
1	A	456	HIS
2	B	295	ASN
2	B	304	GLN
2	B	305	GLN
2	B	315	GLN
2	B	323	HIS
2	B	326	GLN

5.3.3 RNA

There are no RNA molecules 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 2 ligands modelled in this entry, 2 are 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

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	295/295 (100%)	1.35	62 (21%) 1 1	39, 46, 58, 74	0
2	B	61/61 (100%)	0.82	5 (8%) 14 14	42, 47, 53, 58	0
All	All	356/356 (100%)	1.26	67 (18%) 2 1	39, 46, 58, 74	0

All (67) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	217	GLY	11.7
1	A	218	THR	10.8
1	A	219	ASN	9.2
1	A	216	LYS	8.7
1	A	215	ASP	7.4
1	A	164	GLY	7.1
2	B	274	HIS	7.0
1	A	262	LEU	7.0
1	A	220	VAL	6.8
1	A	209	PHE	6.4
1	A	214	LEU	5.4
1	A	221	SER	5.4
1	A	450	GLY	5.0
1	A	212	ARG	4.7
2	B	273	GLY	4.6
1	A	265	ASP	4.4
1	A	213	LYS	4.0
1	A	205	GLN	3.6
1	A	259	GLU	3.6
1	A	195	ILE	3.6
1	A	210	ARG	3.5
1	A	270	TRP	3.3
1	A	335	TRP	3.3
1	A	258	ASP	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	375	ALA	3.1
1	A	384	VAL	3.1
1	A	197	VAL	3.0
1	A	373	ILE	3.0
1	A	273	LYS	3.0
1	A	424	GLU	3.0
1	A	207	LYS	2.8
1	A	208	LEU	2.7
1	A	189	GLU	2.7
1	A	364	ILE	2.7
1	A	361	LEU	2.6
1	A	165	ASP	2.6
1	A	173	GLN	2.6
1	A	222	ASP	2.6
1	A	279	TYR	2.6
1	A	263	PRO	2.5
1	A	360	LEU	2.5
1	A	366	TRP	2.5
1	A	255	ARG	2.5
1	A	204	ALA	2.4
1	A	274	ASN	2.4
1	A	388	TYR	2.4
1	A	374	ILE	2.3
2	B	316	GLN	2.3
1	A	362	TYR	2.3
1	A	223	GLU	2.3
1	A	260	ALA	2.3
1	A	253	GLU	2.3
1	A	191	ALA	2.3
2	B	315	GLN	2.3
1	A	448	ARG	2.2
1	A	275	VAL	2.1
1	A	240	ARG	2.1
1	A	269	LYS	2.1
1	A	237	GLU	2.1
1	A	405	GLU	2.1
1	A	192	LEU	2.1
2	B	333	GLY	2.1
1	A	193	THR	2.1
1	A	451	LEU	2.1
1	A	441	PHE	2.0
1	A	200	LEU	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	177	GLN	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
4	CL	A	1	1/1	0.96	0.07	-2.51	45,45,45,45	0
3	ZN	A	501	1/1	1.00	0.07	-3.13	47,47,47,47	0

6.5 Other polymers (i)

There are no such residues in this entry.