



Full wwPDB X-ray Structure Validation Report ⓘ

Dec 12, 2016 – 06:00 PM EST

PDB ID : 4YVX
Title : Crystal structure of AKR1C3 complexed with glimepiride
Authors : Zhao, Y.; Zheng, X.; Zhang, H.; Hu, X.
Deposited on : 2015-03-20
Resolution : 2.30 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20028442
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) : rb-20028442

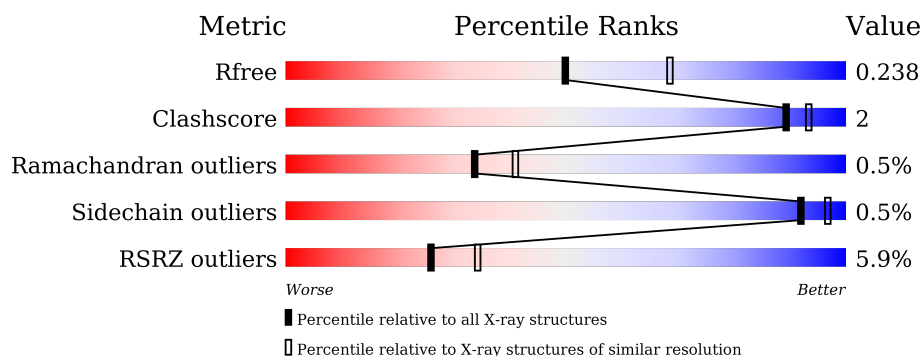
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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	323	 4% 92% 5% ..
1	B	323	 8% 91% 7% •

2 Entry composition [i](#)

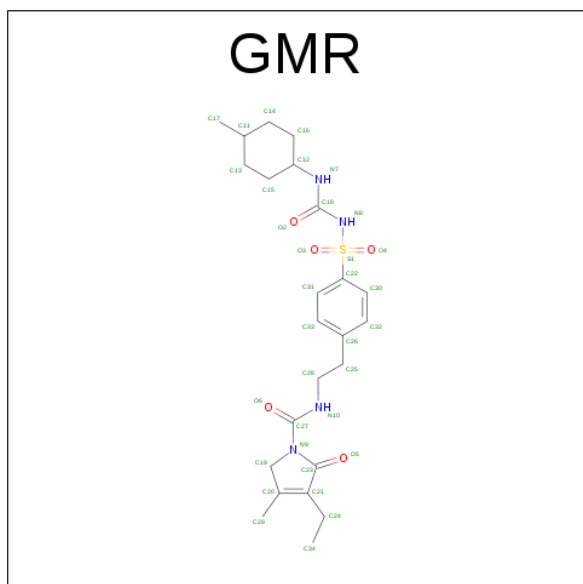
There are 4 unique types of molecules in this entry. The entry contains 5460 atoms, of which 0 are hydrogens and 0 are deuteriums.

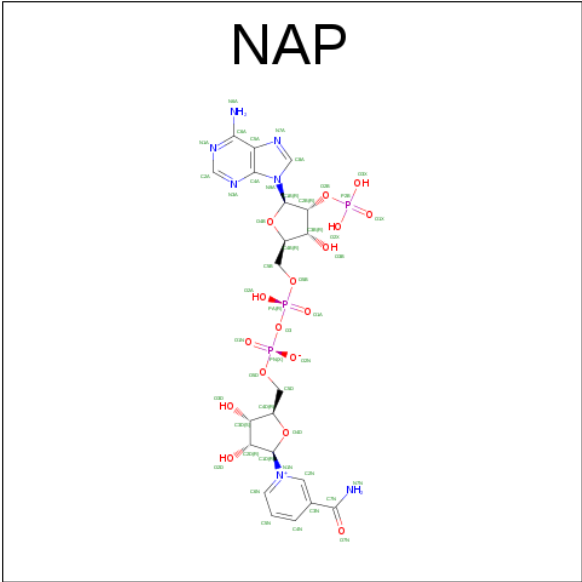
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Aldo-keto reductase family 1 member C3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	315	Total	C	N	O	S	0	0	0
			2524	1609	440	463	12			
1	B	315	Total	C	N	O	S	0	0	0
			2524	1609	440	463	12			

- Molecule 2 is 3-ethyl-4-methyl-N-[2-(4-[(cis-4-methylcyclohexyl)carbamoyl]sulfamoyl]phenyl)ethyl]-2-oxo-2,5-dihydro-1H-pyrrole-1-carboxamide (three-letter code: GMR) (formula: C₂₄H₃₄N₄O₅S).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	143	Total	O	0	0
			143	143		
4	B	105	Total	O	0	0
			105	105		

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: Aldo-keto reductase family 1 member C3

Chain A: 



- Molecule 1: Aldo-keto reductase family 1 member C3

Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	47.00Å 49.27Å 83.72Å 74.65° 86.04° 69.41°	Depositor
Resolution (Å)	80.71 – 2.30 25.50 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (80.71-2.30) 90.9 (25.50-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.15	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.35 (at 2.31Å)	Xtriage
Refinement program	REFMAC 5.8.0123	Depositor
R, R_{free}	0.184 , 0.232 0.192 , 0.238	Depositor DCC
R_{free} test set	1514 reflections (5.30%)	DCC
Wilson B-factor (Å ²)	22.1	Xtriage
Anisotropy	0.049	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 49.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.014 for -h,-k,-k+l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5460	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.35% 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

5.1 Standard geometry

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

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	0.83	0/2582	0.90	6/3494 (0.2%)
1	B	0.76	0/2582	0.88	5/3494 (0.1%)
All	All	0.80	0/5164	0.89	11/6988 (0.2%)

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	301	ARG	NE-CZ-NH1	9.18	124.89	120.30
1	A	301	ARG	NE-CZ-NH2	-7.13	116.73	120.30
1	A	47	ARG	NE-CZ-NH1	6.87	123.73	120.30
1	A	132	ASP	CB-CG-OD1	6.39	124.05	118.30
1	B	250	ARG	NE-CZ-NH1	5.70	123.15	120.30
1	B	96	ARG	NE-CZ-NH1	5.55	123.08	120.30
1	B	220	GLY	N-CA-C	-5.51	99.31	113.10
1	B	263	ARG	NE-CZ-NH2	5.24	122.92	120.30
1	A	223	ARG	NE-CZ-NH1	5.22	122.91	120.30
1	B	76	ARG	NE-CZ-NH1	5.05	122.82	120.30
1	A	226	ARG	NE-CZ-NH2	5.02	122.81	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	2524	0	2531	12	0
1	B	2524	0	2531	10	0
2	A	34	0	34	2	0
2	B	34	0	34	2	0
3	A	48	0	25	2	0
3	B	48	0	25	1	0
4	A	143	0	0	1	0
4	B	105	0	0	0	0
All	All	5460	0	5180	25	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 (25) 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:311:PHE:CD1	1:B:317:TYR:CZ	2.83	0.67
1:A:308:SER:HB3	1:A:311:PHE:HB3	1.77	0.66
1:B:311:PHE:CE1	1:B:312:ALA:HB2	2.32	0.64
1:A:226:ARG:NE	4:A:501:HOH:O	2.35	0.59
1:A:311:PHE:HB2	2:A:401:GMR:H32	1.87	0.56
1:B:311:PHE:CD1	1:B:312:ALA:HB2	2.41	0.55
1:B:291:GLU:N	1:B:291:GLU:OE1	2.40	0.54
1:A:310:SER:O	1:A:313:SER:HB3	2.10	0.52
1:B:226:ARG:HB2	1:B:227:TRP:CE3	2.47	0.49
1:B:311:PHE:CD1	1:B:312:ALA:N	2.81	0.48
1:B:217:SER:HA	3:B:402:NAP:O2A	2.14	0.48
1:A:314:HIS:CD2	1:A:316:ASN:H	2.34	0.45
1:A:314:HIS:CD2	1:A:316:ASN:HB2	2.52	0.45
1:B:145:CYS:HB3	1:B:179:LYS:HD2	1.99	0.44
1:B:196:TYR:OH	1:B:238:ASP:OD1	2.26	0.43
2:B:401:GMR:H5	2:B:401:GMR:O3	2.18	0.43
1:A:270:LYS:O	3:A:402:NAP:H8A	2.19	0.42
1:A:31:ARG:O	1:A:60:GLN:HG2	2.19	0.42
1:A:145:CYS:HB3	1:A:179:LYS:HD2	2.02	0.41
2:A:401:GMR:C18	3:A:402:NAP:C4N	2.98	0.41
2:B:401:GMR:H5	2:B:401:GMR:S1	2.59	0.41
1:A:120:MET:HG2	1:A:314:HIS:CD2	2.55	0.41
1:B:163:ILE:O	1:B:187:VAL:HG22	2.21	0.41
1:A:120:MET:HE1	1:A:311:PHE:HA	2.03	0.40
1:A:120:MET:HG2	1:A:314:HIS:HD2	1.86	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	313/323 (97%)	303 (97%)	9 (3%)	1 (0%)	46	57
1	B	313/323 (97%)	298 (95%)	13 (4%)	2 (1%)	30	36
All	All	626/646 (97%)	601 (96%)	22 (4%)	3 (0%)	34	41

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	313	SER
1	B	310	SER
1	B	316	ASN

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	279/287 (97%)	277 (99%)	2 (1%)	88	95
1	B	279/287 (97%)	278 (100%)	1 (0%)	93	97
All	All	558/574 (97%)	555 (100%)	3 (0%)	92	97

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

Mol	Chain	Res	Type
1	A	304	HIS
1	A	309	ASP
1	B	171	ARG

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

Mol	Chain	Res	Type
1	B	314	HIS

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](#)

4 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	GMR	A	401	-	35,36,36	0.57	0	43,51,51	2.24	9 (20%)
3	NAP	A	402	-	45,52,52	1.17	5 (11%)	55,80,80	1.89	9 (16%)
2	GMR	B	401	-	35,36,36	0.42	0	43,51,51	1.88	4 (9%)
3	NAP	B	402	-	45,52,52	1.24	6 (13%)	55,80,80	1.83	8 (14%)

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	GMR	A	401	-	-	0/27/53/53	0/3/3/3
3	NAP	A	402	-	-	0/27/67/67	0/5/5/5
2	GMR	B	401	-	-	0/27/53/53	0/3/3/3
3	NAP	B	402	-	-	0/27/67/67	0/5/5/5

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	402	NAP	C2D-C1D	-2.45	1.49	1.53
3	B	402	NAP	O5D-C5D	-2.18	1.36	1.44
3	B	402	NAP	PA-O2A	-2.12	1.46	1.55
3	A	402	NAP	O4D-C1D	2.16	1.44	1.41
3	A	402	NAP	O4B-C1B	2.36	1.44	1.41
3	A	402	NAP	C2A-N3A	2.47	1.36	1.32
3	A	402	NAP	C2A-N1A	2.52	1.38	1.33
3	B	402	NAP	C2A-N3A	2.56	1.36	1.32
3	B	402	NAP	O4B-C1B	2.70	1.45	1.41
3	A	402	NAP	C5A-C4A	3.61	1.48	1.40
3	B	402	NAP	C5A-C4A	3.65	1.48	1.40

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	402	NAP	N3A-C2A-N1A	-8.47	122.22	128.87
3	A	402	NAP	N3A-C2A-N1A	-7.36	123.09	128.87
2	B	401	GMR	O2-C18-N8	-5.87	110.62	121.64
3	A	402	NAP	C4B-O4B-C1B	-5.24	104.09	109.64
2	A	401	GMR	O2-C18-N8	-5.18	111.92	121.64
3	A	402	NAP	O2B-P2B-O1X	-4.76	96.13	107.48
3	B	402	NAP	O4D-C4D-C5D	-4.03	94.86	109.29
3	B	402	NAP	C4B-O4B-C1B	-3.41	106.03	109.64
3	A	402	NAP	C1B-N9A-C4A	-2.75	123.74	126.81
3	B	402	NAP	C1B-N9A-C4A	-2.47	124.05	126.81
2	A	401	GMR	O2-C18-N7	-2.46	117.00	122.77
2	A	401	GMR	O3-S1-N8	-2.22	100.40	106.73
3	A	402	NAP	C4N-C3N-C7N	-2.07	115.62	121.11
2	A	401	GMR	C28-C25-C26	2.12	117.54	112.83
2	B	401	GMR	C12-N7-C18	2.14	128.40	123.38

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	402	NAP	O4B-C1B-C2B	2.19	110.53	106.60
3	B	402	NAP	C2A-N1A-C6A	2.19	122.68	118.77
3	B	402	NAP	O3X-P2B-O2X	2.19	115.50	107.44
3	A	402	NAP	O3D-C3D-C4D	2.36	118.05	111.01
2	A	401	GMR	C12-N7-C18	2.53	129.32	123.38
3	A	402	NAP	N6A-C6A-N1A	2.62	122.91	118.52
3	A	402	NAP	C2N-C3N-C4N	2.79	121.44	118.27
2	A	401	GMR	C14-C16-C12	2.80	115.28	111.51
2	A	401	GMR	O4-S1-N8	2.83	114.79	106.73
3	B	402	NAP	O4D-C1D-N1N	3.23	111.60	108.10
3	B	402	NAP	C4D-O4D-C1D	3.49	113.34	109.64
2	B	401	GMR	C18-N8-S1	4.65	133.77	123.66
2	A	401	GMR	C18-N8-S1	7.13	139.15	123.66
2	B	401	GMR	N8-C18-N7	8.20	129.67	114.48
2	A	401	GMR	N8-C18-N7	8.94	131.03	114.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	GMR	2	0
3	A	402	NAP	2	0
2	B	401	GMR	2	0
3	B	402	NAP	1	0

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, 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	315/323 (97%)	0.20	12 (3%) 44 53	11, 20, 47, 118	0
1	B	315/323 (97%)	0.35	25 (7%) 15 22	11, 25, 66, 127	0
All	All	630/646 (97%)	0.27	37 (5%) 26 34	11, 22, 57, 127	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	311	PHE	10.6
1	A	312	ALA	10.2
1	A	313	SER	8.5
1	B	317	TYR	6.5
1	A	311	PHE	6.5
1	A	309	ASP	6.5
1	A	320	SER	6.4
1	B	320	SER	6.3
1	B	315	PRO	6.1
1	B	312	ALA	5.5
1	B	309	ASP	5.2
1	A	314	HIS	5.1
1	B	313	SER	5.0
1	B	308	SER	4.7
1	B	319	TYR	4.4
1	B	135	GLY	4.4
1	B	310	SER	4.3
1	B	126	GLU	4.2
1	B	133	GLU	4.1
1	B	134	ASN	4.0
1	A	315	PRO	3.8
1	A	317	TYR	3.7
1	B	314	HIS	3.6
1	B	131	THR	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	319	TYR	3.1
1	B	307	ASN	3.1
1	B	132	ASP	3.0
1	B	302	ASN	3.0
1	A	310	SER	2.9
1	A	126	GLU	2.8
1	B	231	ASN	2.7
1	A	318	PRO	2.6
1	B	230	PRO	2.5
1	B	318	PRO	2.4
1	B	127	GLU	2.2
1	B	137	VAL	2.2
1	B	316	ASN	2.1

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, 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	LLDF	B-factors(Å ²)	Q<0.9
2	GMR	B	401	34/34	0.87	0.21	0.67	38,60,73,75	0
2	GMR	A	401	34/34	0.93	0.15	-0.30	18,25,29,31	0
3	NAP	B	402	48/48	0.95	0.14	-0.61	16,19,23,24	0
3	NAP	A	402	48/48	0.97	0.12	-1.07	10,12,14,16	0

6.5 Other polymers [i](#)

There are no such residues in this entry.