wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 10:50 PM EDT

PDB ID	:	3MKU
Title	:	Structure of a Cation-bound Multidrug and Toxin Compound Extrusion (MATE) transporter
Authors	:	He, X.; Szewczyk, P.; Karyakin, A.; Evin, M.; Hong, W.-X.; Zhang, Q.; Chang, G.
Deposited on	:	2010-04-15
Resolution	:	4.20 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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:

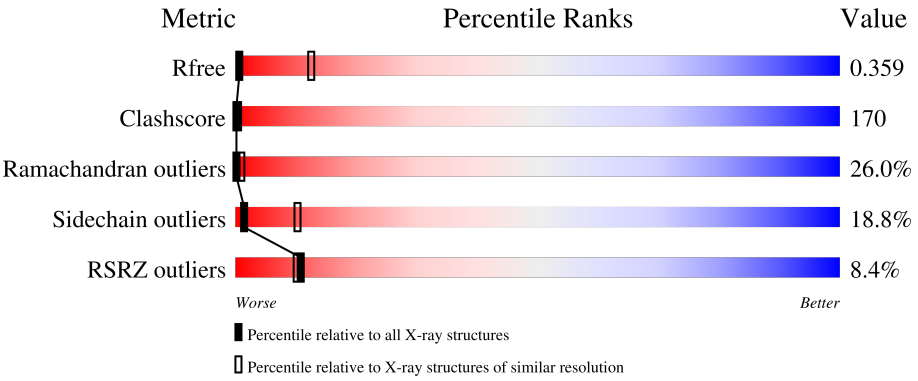
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 4.20 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	1005 (4.62-3.78)
Clashscore	141614	1044 (4.60-3.80)
Ramachandran outliers	138981	1000 (4.60-3.80)
Sidechain outliers	138945	1007 (4.62-3.78)
RSRZ outliers	127900	1063 (4.70-3.70)

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	460	11% 5% 58% 33% .
1	B	460	6% . 59% 33% .

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 7013 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 Multi antimicrobial extrusion protein (Na(+)/drug antiporter) MATE-like MDR efflux pump.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	460	Total	C	N	O	S	0	0	0
			3505	2326	569	589	21			
1	B	460	Total	C	N	O	S	0	0	0
			3506	2326	569	590	21			

- Molecule 2 is RUBIDIUM ION (three-letter code: RB) (formula: Rb).

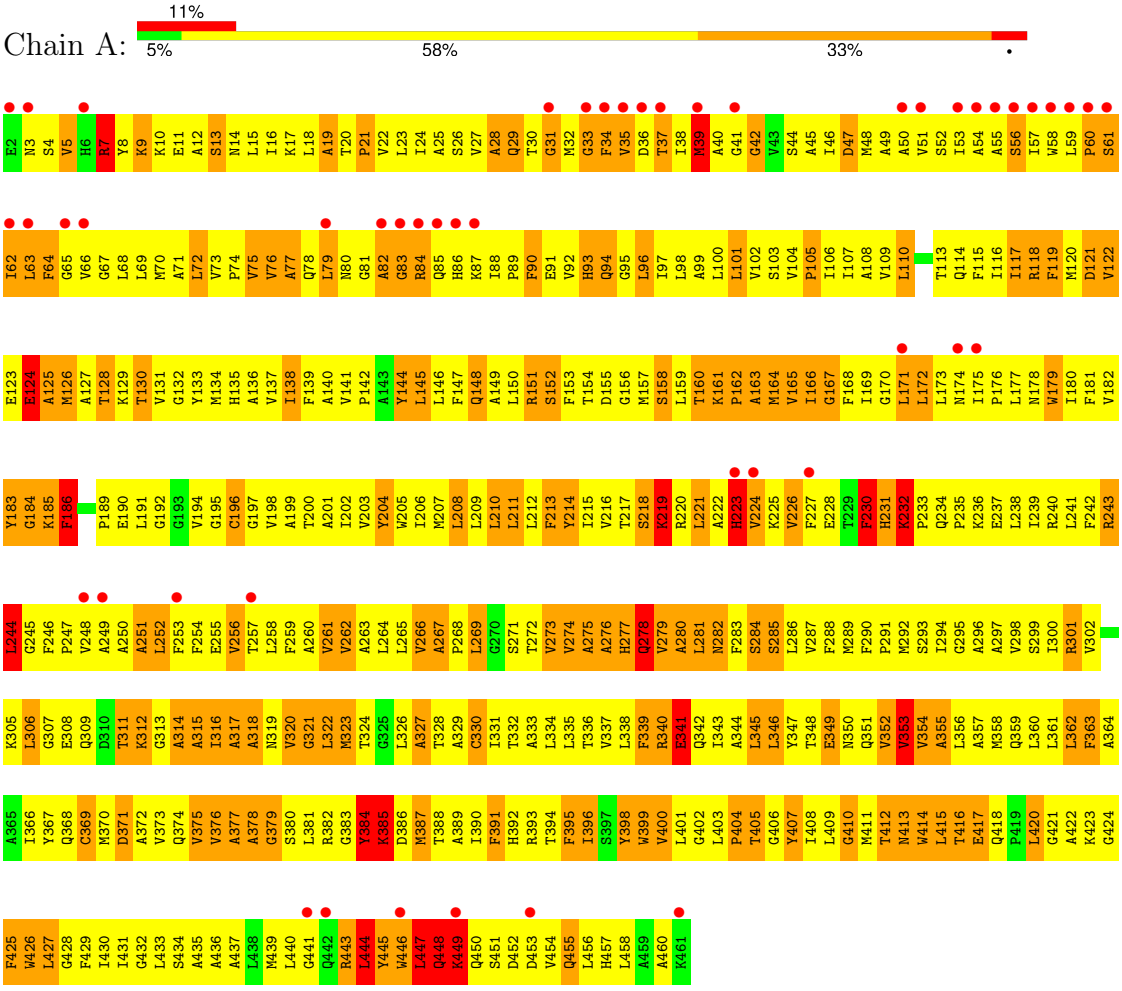
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Rb	0	0
			1	1		
2	B	1	Total	Rb	0	0
			1	1		

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.

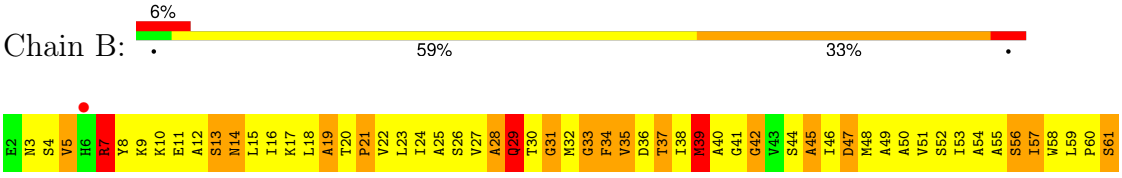
• Molecule 1: Multi antimicrobial extrusion protein (Na(+)/drug antiporter) MATE-like MDR efflux pump

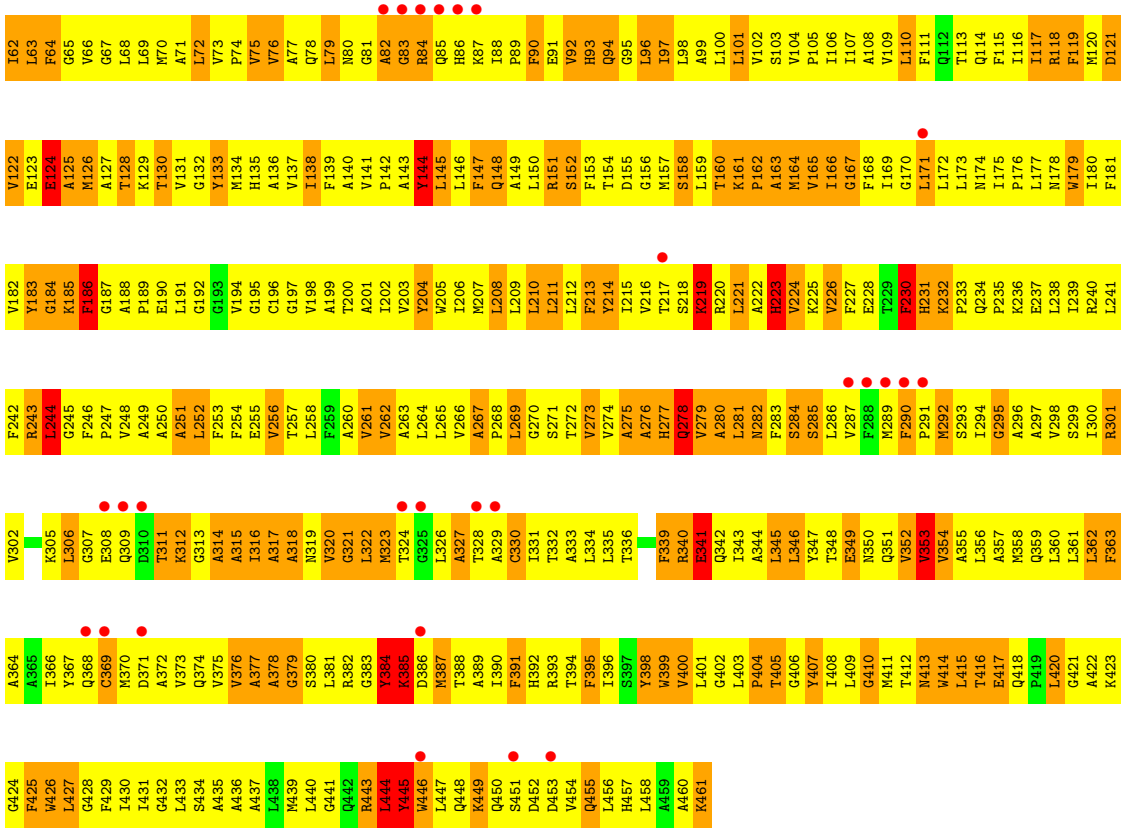
Chain A:



• Molecule 1: Multi antimicrobial extrusion protein (Na(+)/drug antiporter) MATE-like MDR efflux pump

Chain B:




I62 L63 F64 G65 V66 G67 L68 L69 M70 A71 L72 V73 P74 V75 V76 A77 Q78 L79 N80 G81 A82 G83 R84 Q85 H86 K87 I88 P89 F90 E91 V92 H93 Q94 G95 L96 I97 L98 A99 L100 L101 V102 S103 V104 P105 I106 I107 A108 V109 L110 F111 Q112 T113 Q114 F115 I116 I117 R118 F119 M120 D121
V122 E123 E124 A125 M126 A127 T128 K129 T130 V131 G132 Y133 M134 H135 A136 V137 I138 F139 A140 V141 P142 A143 Y144 L145 L146 F147 Q148 A149 L150 R151 S152 F153 T154 D155 G156 M157 S158 L159 T160 K161 P162 A163 M164 V165 I166 G167 F168 I169 G170 L171 L172 L173 N174 I175 P176 L177 N178 W179 I180 F181
V182 Y183 G184 K185 F186 G187 A188 P189 E190 L191 G192 G193 V194 G195 C196 G197 V198 A199 T200 A201 I202 V203 Y204 W205 I206 M207 L208 L209 L210 L211 L212 F213 Y214 I215 V216 T217 S218 K219 R220 L221 A222 H223 V224 K225 V226 F227 E228 T229 P230 H231 K232 P233 Q234 P235 K236 E237 L238 I239 R240 L241
F242 R243 L244 G245 F246 P247 V248 A249 A250 A251 L252 F253 F254 E255 V256 T257 L258 F259 A260 V261 V262 A263 L264 L265 V266 A267 P268 L269 G270 S271 T272 V273 V274 A275 A276 H277 Q278 V279 A280 L281 N282 F283 S284 S285 L286 V287 F288 M289 F290 P291 M292 S293 I294 G295 A296 A297 V298 S299 I300 R301
V302 K305 L306 G307 E308 Q309 D310 T311 K312 G313 A314 A315 I316 A317 A318 N319 V320 G321 L322 M323 T324 G325 L326 A327 T328 A329 C330 I331 T332 A333 L334 L335 T336 F339 R340 E341 Q342 I343 A344 L345 L346 Y347 T348 E349 N350 Q351 V352 V353 V354 A355 L356 A357 M358 Q359 L360 I361 L362 F363
A364 A365 I366 Y367 Q368 C369 M370 D371 A372 V373 Q374 V375 V376 A377 A378 G379 S380 L381 R382 G383 Y384 K385 D386 M387 T388 A389 I390 F391 H392 R393 T394 F395 I396 S397 Y398 W399 V400 L401 G402 L403 P404 T405 G406 Y407 I408 L409 G410 M411 T412 N413 W414 L415 T416 E417 Q418 P419 L420 G421 A422 K423
G424 F425 W426 L427 G428 F429 I430 I431 G432 L433 S434 A435 A436 A437 L438 M439 L440 G441 Q442 R443 L444 Y445 W446 L447 Q448 K449 Q450 S451 D452 D453 V454 Q455 L456 H457 L458 A459 A460 K461

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	159.80Å 241.85Å 46.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 4.20 133.33 – 4.00	Depositor EDS
% Data completeness (in resolution range)	99.0 (20.00-4.20) 99.0 (133.33-4.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.87 (at 4.01Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.309 , 0.342 0.324 , 0.359	Depositor DCC
R_{free} test set	791 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	105.5	Xtriage
Anisotropy	1.074	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.17 , 145.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.36$, $< L^2 > = 0.19$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.73	EDS
Total number of atoms	7013	wwPDB-VP
Average B, all atoms (Å^2)	118.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.00% 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: RB

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	0.65	0/3585	0.97	13/4879 (0.3%)
1	B	0.64	0/3586	0.93	9/4879 (0.2%)
All	All	0.64	0/7171	0.95	22/9758 (0.2%)

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	B	0	2

There are no bond length outliers.

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	353	VAL	CB-CA-C	-8.84	94.61	111.40
1	A	410	GLY	N-CA-C	-8.24	92.50	113.10
1	A	384	TYR	CB-CA-C	7.94	126.28	110.40
1	B	410	GLY	N-CA-C	-7.78	93.65	113.10
1	A	447	LEU	N-CA-C	-7.57	90.56	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	133	TYR	Sidechain
1	B	144	TYR	Sidechain

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	3505	0	3676	1224	0
1	B	3506	0	3676	1223	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
All	All	7013	0	7352	2440	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 170.

The worst 5 of 2440 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:447:LEU:CD1	1:A:450:GLN:HE21	1.27	1.48
1:A:447:LEU:CD1	1:A:450:GLN:NE2	1.96	1.27
1:A:447:LEU:HD13	1:A:450:GLN:NE2	1.50	1.24
1:B:162:PRO:O	1:B:165:VAL:HG12	1.38	1.17
1:A:162:PRO:O	1:A:165:VAL:HG12	1.43	1.17

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	458/460 (100%)	218 (48%)	118 (26%)	122 (27%)	0 1
1	B	458/460 (100%)	221 (48%)	121 (26%)	116 (25%)	0 1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	916/920 (100%)	439 (48%)	239 (26%)	238 (26%)	0 1

5 of 238 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	19	ALA
1	A	35	VAL
1	A	56	SER
1	A	76	VAL
1	A	121	ASP

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	364/364 (100%)	296 (81%)	68 (19%)	1 10
1	B	364/364 (100%)	295 (81%)	69 (19%)	1 10
All	All	728/728 (100%)	591 (81%)	137 (19%)	1 10

5 of 137 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	371	ASP
1	B	391	PHE
1	B	426	TRP
1	A	385	LYS
1	A	384	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	14	ASN
1	B	114	GLN
1	B	455	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	85	GLN
1	B	148	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 monosaccharides 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	460/460 (100%)	0.12	49 (10%)	6	6	34, 111, 174, 267	0
1	B	460/460 (100%)	-0.03	28 (6%)	21	17	28, 117, 180, 253	0
All	All	920/920 (100%)	0.05	77 (8%)	11	10	28, 114, 180, 267	0

The worst 5 of 77 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	309	GLN	7.6
1	A	36	ASP	6.7
1	B	83	GLY	5.8
1	A	84	ARG	5.3
1	A	442	GLN	5.0

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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
2	RB	B	5002	1/1	0.66	0.20	186,186,186,186	0
2	RB	A	5001	1/1	0.80	0.11	182,182,182,182	0

6.5 Other polymers (i)

There are no such residues in this entry.