

# Full wwPDB X-ray Structure Validation Report (i)

#### Jun 23, 2024 – 11:14 AM EDT

PDB ID	:	5F7K
Title	:	Blood group antigen binding adhesin BabA of Helicobacter pylori strain 17875
		in complex with Nanobody Nb-ER19
Authors	:	Moonens, K.; Gideonsson, P.; Subedi, S.; Romao, E.; Oscarson, S.; Muylder-
		mans, S.; Boren, T.; Remaut, H.
Deposited on	:	2015-12-08
Resolution	:	2.17  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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
$\mathrm{EDS}$	:	2.37.1
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)	:	2.37.1

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

The reported resolution of this entry is 2.17 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	Similar resolution
	$(\# { m Entries})$	$(\# { m Entries},  { m resolution}  { m range}({ m \AA}))$
$R_{free}$	130704	6864 (2.20-2.16)
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)
RSRZ outliers	127900	6738 (2.20-2.16)

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	А	466	% • 82%	6% • 10%
1	В	466	<sup>2%</sup> 83%	6% • 10%
2	С	120	% 87%	7% • 6%
2	D	120	87%	8% 5%



# 2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 8518 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.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	Λ	420	Total	С	Ν	0	$\mathbf{S}$	0	2	0
	A	420	3145	1939	541	652	13	0	J	0
1	D	491	Total	С	Ν	0	S	0	1	0
1	Б	421	3146	1030	544	650	13	0	1	0

• Molecule 1 is a protein called Adhesin binding fucosylated histo-blood group antigen.

Chain	Residue	Modelled	Actual	Comment	Reference
А	3	ALA	-	expression tag	UNP O52269
А	4	SER	-	expression tag	UNP O52269
А	5	TRP	-	expression tag	UNP O52269
А	6	SER	-	expression tag	UNP O52269
А	7	HIS	-	expression tag	UNP O52269
А	8	PRO	-	expression tag	UNP O52269
А	9	GLN	-	expression tag	UNP O52269
А	10	PHE	-	expression tag	UNP O52269
А	11	GLU	-	expression tag	UNP O52269
А	12	LYS	-	expression tag	UNP O52269
А	13	SER	-	expression tag	UNP O52269
А	14	GLY	-	expression tag	UNP O52269
А	15	GLY	-	expression tag	UNP O52269
А	16	GLY	-	expression tag	UNP O52269
А	17	GLY	-	expression tag	UNP O52269
А	18	GLY	-	expression tag	UNP O52269
А	19	LEU	-	expression tag	UNP O52269
А	20	VAL	-	expression tag	UNP O52269
А	21	PRO	-	expression tag	UNP O52269
А	22	ARG	-	expression tag	UNP O52269
А	23	GLY	-	expression tag	UNP O52269
А	24	SER	-	expression tag	UNP O52269
А	461	GLY	-	expression tag	UNP O52269
А	462	SER	-	expression tag	UNP O52269
A	463	HIS	-	expression tag	UNP O52269

There are 60 discrepancies between the modelled and reference sequences:





Chain	Deciduo	Medelled	Actual	Commont	Defenence
Chain	Residue	Modelled	Actual	Comment	Reference
A	464	HIS	-	expression tag	UNP 052269
A	465	HIS	-	expression tag	UNP 052269
A	466	HIS	-	expression tag	UNP O52269
A	467	HIS	-	expression tag	UNP O52269
A	468	HIS	-	expression tag	UNP O52269
В	3	ALA	-	expression tag	UNP O52269
В	4	SER	-	expression tag	UNP O52269
В	5	TRP	-	expression tag	UNP O52269
В	6	SER	-	expression tag	UNP O52269
В	7	HIS	-	expression tag	UNP O52269
В	8	PRO	-	expression tag	UNP O52269
В	9	GLN	-	expression tag	UNP O52269
В	10	PHE	-	expression tag	UNP O52269
В	11	GLU	-	expression tag	UNP O52269
В	12	LYS	-	expression tag	UNP O52269
В	13	SER	-	expression tag	UNP O52269
В	14	GLY	-	expression tag	UNP O52269
В	15	GLY	-	expression tag	UNP O52269
В	16	GLY	-	expression tag	UNP O52269
В	17	GLY	-	expression tag	UNP O52269
В	18	GLY	-	expression tag	UNP O52269
В	19	LEU	-	expression tag	UNP O52269
В	20	VAL	-	expression tag	UNP O52269
В	21	PRO	-	expression tag	UNP O52269
В	22	ARG	-	expression tag	UNP O52269
В	23	GLY	-	expression tag	UNP O52269
В	24	SER	-	expression tag	UNP O52269
В	461	GLY	-	expression tag	UNP O52269
В	462	SER	_	expression tag	UNP O52269
В	463	HIS	-	expression tag	UNP O52269
В	464	HIS	_	expression tag	UNP 052269
В	465	HIS	-	expression tag	UNP O52269
В	466	HIS	-	expression tag	UNP 052269
В	467	HIS	-	expression tag	UNP 052269
В	468	HIS	-	expression tag	UNP O52269

Continued from previous page...

• Molecule 2 is a protein called Nanobody Nb-ER19.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
9	С	112	Total	С	Ν	0	S	0	0	0
	U	115	863	540	155	163	5		0	
0	р	114	Total	С	Ν	0	S	0	0	0
	D	114	873	546	158	164	5	0	0	0



• Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula:  $O_4P$ ).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	
3	А	1	Total 5	0 4	Р 1	0	0



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	
4	А	1	Total 4	C 2	O 2	0	0

• Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	С	1	Total Na 1 1	0	0

• Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	А	176	Total O 176 176	0	0
6	С	54	$\begin{array}{cc} \text{Total} & \text{O} \\ 54 & 54 \end{array}$	0	0
6	В	158	Total O 158 158	0	0
6	D	93	Total O 93 93	0	0



# 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: Adhesin binding fucosylated histo-blood group antigen



## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	50.99Å 131.66Å 123.46Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $94.77^{\circ}$ $90.00^{\circ}$	Depositor
Bosolution (Å)	58.04 - 2.17	Depositor
Resolution (A)	58.04 - 2.17	EDS
% Data completeness	99.4 (58.04-2.17)	Depositor
(in resolution range)	99.4(58.04-2.17)	EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	$2.17 (at 2.16 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
B B.	0.171 , $0.207$	Depositor
$n, n_{free}$	0.178 , $0.211$	DCC
$R_{free}$ test set	4261 reflections $(5.00%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	46.1	Xtriage
Anisotropy	0.109	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.34,48.1	EDS
L-test for twinning <sup>2</sup>	$ < L >=0.48, < L^2>=0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8518	wwPDB-VP
Average B, all atoms $(Å^2)$	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.53% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical 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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 5 Model quality (i)

## 5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, NA, EDO

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).

Mal	Mol Chain RMS		nd lengths	Bond angles	
			# Z  > 5	RMSZ	# Z  > 5
1	А	0.94	3/3197~(0.1%)	0.88	5/4344~(0.1%)
1	В	0.88	0/3193	0.86	2/4338~(0.0%)
2	С	0.94	1/882~(0.1%)	1.01	4/1195~(0.3%)
2	D	1.09	2/893~(0.2%)	1.00	2/1210~(0.2%)
All	All	0.94	6/8165~(0.1%)	0.90	13/11087~(0.1%)

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	В	0	1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	18	SER	CB-OG	-5.79	1.34	1.42
1	А	105	GLN	CG-CD	5.74	1.64	1.51
2	С	18	SER	CB-OG	-5.34	1.35	1.42
2	D	98	ARG	CZ-NH2	-5.27	1.26	1.33
1	А	317[A]	GLU	CD-OE2	5.25	1.31	1.25
1	А	317[B]	GLU	CD-OE2	5.25	1.31	1.25

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	А	318	LEU	CA-CB-CG	7.89	133.45	115.30
1	В	318	LEU	CA-CB-CG	7.88	133.42	115.30
2	С	99	LEU	CA-CB-CG	6.82	130.98	115.30

Continued on next page...



Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	D	44	LYS	CB-CA-C	-5.78	98.84	110.40
2	С	99	LEU	CB-CG-CD2	5.66	120.63	111.00
2	С	20	ARG	NE-CZ-NH2	-5.64	117.48	120.30
1	А	317[A]	GLU	CA-CB-CG	5.54	125.60	113.40
1	А	317[B]	GLU	CA-CB-CG	5.54	125.60	113.40
1	В	295	ARG	CA-CB-CG	5.37	125.20	113.40
2	С	98	ARG	NE-CZ-NH1	5.27	122.94	120.30
1	А	139	ARG	NE-CZ-NH2	-5.24	117.68	120.30
2	D	72	ARG	NE-CZ-NH1	5.20	122.90	120.30
1	А	257	ASP	CB-CG-OD2	-5.16	113.66	118.30

Continued from previous page...

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	В	286	SER	Peptide

### 5.2 Too-close contacts (i)

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	А	3145	0	3067	24	0
1	В	3146	0	3063	15	0
2	С	863	0	845	3	0
2	D	873	0	852	4	0
3	А	5	0	0	0	0
4	А	4	0	6	0	0
5	С	1	0	0	0	0
6	А	176	0	0	1	0
6	В	158	0	0	5	0
6	С	54	0	0	1	0
6	D	93	0	0	2	0
All	All	8518	0	7833	46	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 3.

All (46) close contacts within the same asymmetric unit are listed below, sorted by their clash



magnitude.

Atom-1 Atom-2 Interator		Interatomic	Clash
Atom-1	Atom-2	distance $(\text{\AA})$	overlap (Å)
1:B:317:GLU:HG2	6:B:644:HOH:O	1.47	1.12
1:A:202:THR:HG22	1:A:204:VAL:H	1.43	0.82
1:A:129:ARG:NH1	6:A:602:HOH:O	2.15	0.80
1:A:198:SER:O	1:A:202:THR:HB	1.87	0.75
1:A:243:VAL:HG12	1:A:244:SER:H	1.53	0.72
1:A:234:SER:CB	1:A:243:VAL:HA	2.25	0.67
2:D:20:ARG:NH1	6:D:201:HOH:O	2.21	0.65
1:B:234:SER:OG	6:B:501:HOH:O	2.14	0.64
1:B:352[A]:GLN:HG3	6:B:548:HOH:O	1.99	0.63
2:C:103:ARG:HG2	6:C:353:HOH:O	1.98	0.62
2:C:3:VAL:N	2:C:26:SER:HG	1.98	0.60
2:D:14:GLN:OE1	2:D:116:HIS:HB3	2.01	0.60
1:A:243:VAL:CG1	1:A:244:SER:H	2.16	0.58
1:B:204:VAL:HG12	1:B:206:ASP:O	2.05	0.57
1:A:235:GLN:O	1:A:236:ALA:O	2.22	0.57
1:A:39:LEU:C	1:A:39:LEU:HD13	2.25	0.56
1:B:273:ILE:HG23	1:B:298:LEU:HD23	1.88	0.55
1:A:42:ARG:NH1	1:A:43:TYR:OH	2.34	0.55
1:A:105:GLN:OE1	1:A:279:TYR:CD2	2.60	0.54
1:B:275:THR:OG1	6:B:502:HOH:O	2.19	0.53
1:A:234:SER:HB2	1:A:243:VAL:HA	1.89	0.53
1:A:235:GLN:O	1:A:236:ALA:C	2.49	0.51
1:A:287:GLY:O	1:A:288:GLY:O	2.30	0.51
1:B:287:GLY:O	1:B:288:GLY:O	2.29	0.50
1:B:307:ALA:C	1:B:311:MET:HE2	2.31	0.50
2:D:87:LYS:HE3	6:D:253:HOH:O	2.10	0.50
1:A:202:THR:HG23	1:A:204:VAL:HG23	1.94	0.49
1:A:111:ASN:OD1	1:A:113:GLN:HB2	2.12	0.48
1:B:307:ALA:O	1:B:311:MET:HE2	2.12	0.48
2:D:3:VAL:HG13	2:D:29:ILE:HG13	1.95	0.48
1:B:156:LEU:C	1:B:156:LEU:HD23	2.34	0.48
1:A:307:ALA:O	1:A:311:MET:HE3	2.14	0.47
1:B:376:ASN:HB3	6:B:551:HOH:O	2.15	0.47
1:A:279:TYR:CE1	1:A:295:ARG:HG3	2.50	0.46
1:B:111:ASN:OD1	1:B:111:ASN:N	2.46	0.46
1:A:307:ALA:C	1:A:311:MET:HE3	2.36	0.45
1:A:156:LEU:C	1:A:156:LEU:HD23	2.39	0.43
1:B:240:THR:O	1:B:242:ARG:N	2.52	0.43
1:A:282:VAL:HG22	1:A:294:THR:HA	2.01	0.43
1:A:337:ASN:O	1:A:340:LYS:CG	2.67	0.43
1:A:237:LYS:O	1:A:239:ASN:N	2.53	0.41

Continued on next page...



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:3:VAL:HB	2:C:26:SER:HB3	2.01	0.41
1:B:337:ASN:O	1:B:340:LYS:HG2	2.21	0.41
1:B:282:VAL:HG22	1:B:294:THR:HA	2.01	0.41
1:A:236:ALA:O	1:A:237:LYS:O	2.39	0.40
1:A:111:ASN:CG	1:A:113:GLN:HG2	2.42	0.40

Continued from previous page...

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	Perce	entiles
1	А	417/466~(90%)	399~(96%)	14 (3%)	4 (1%)	15	12
1	В	416/466~(89%)	396~(95%)	16 (4%)	4 (1%)	15	12
2	С	111/120~(92%)	107~(96%)	4 (4%)	0	100	100
2	D	112/120~(93%)	110 (98%)	2(2%)	0	100	100
All	All	1056/1172~(90%)	1012 (96%)	36 (3%)	8 (1%)	19	17

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	237	LYS
1	В	288	GLY
1	А	236	ALA
1	А	288	GLY
1	В	241	THR
1	В	237	LYS
1	А	238	GLY
1	В	205	SER



#### 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 side chain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Perce	ntiles
1	А	349/381~(92%)	339~(97%)	10 (3%)	42	51
1	В	348/381~(91%)	336~(97%)	12 (3%)	37	44
2	$\mathbf{C}$	92/99~(93%)	89~(97%)	3(3%)	38	46
2	D	93/99~(94%)	93 (100%)	0	100	100
All	All	882/960~(92%)	857 (97%)	25 (3%)	43	53

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	$\mathbf{Res}$	Type
1	А	42	ARG
1	А	44	SER
1	А	145	TYR
1	А	202	THR
1	А	294	THR
1	А	295	ARG
1	А	318	LEU
1	А	320	ASN
1	А	329	GLU
1	А	340	LYS
2	С	14	GLN
2	С	99	LEU
2	С	115	SER
1	В	44	SER
1	В	111	ASN
1	В	145	TYR
1	В	206	ASP
1	В	214	LYS
1	В	235	GLN
1	В	260	GLN
1	В	294	THR
1	В	298	LEU
1	В	318	LEU
1	В	320	ASN
1	В	329	GLU



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 monosaccharides in this entry.

### 5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 for Mogul analysis.

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).

Mal	Type Chain			T inl.	Bond lengths			Bond angles		
	туре	Chain	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
4	EDO	А	502	-	3,3,3	0.58	0	2,2,2	0.60	0
3	PO4	А	501	-	4,4,4	0.30	0	6,6,6	1.71	2 (33%)

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
4	EDO	А	502	-	-	0/1/1/1	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
3	А	501	PO4	04-P-03	2.94	117.39	107.97
3	А	501	PO4	03-P-01	-2.53	101.65	110.89

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 (i)

## 6.1 Protein, DNA and RNA chains (i)

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	< <b>RSRZ</b> >	#RSRZ>2	$OWAB(Å^2)$	Q<0.9
1	А	420/466~(90%)	0.14	6 (1%) 75 75	36, 48, 95, 128	0
1	В	421/466~(90%)	0.21	9 (2%) 63 64	35, 52, 98, 128	0
2	С	113/120~(94%)	-0.11	1 (0%) 84 84	39, 49, 66, 85	0
2	D	114/120~(95%)	0.05	0 100 100	32, 42, 57, 80	0
All	All	1068/1172~(91%)	0.13	16 (1%) 73 74	32, 49, 95, 128	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	241	THR	4.2
1	В	232	VAL	3.7
1	А	41	THR	3.6
1	В	204	VAL	3.3
1	В	207	GLN	3.3
2	С	3	VAL	2.7
1	А	32	ASN	2.6
1	В	41	THR	2.5
1	В	205	SER	2.5
1	А	33	TYR	2.4
1	А	338	ASN	2.4
1	А	295	ARG	2.3
1	А	205	SER	2.3
1	В	240	THR	2.2
1	В	237	LYS	2.0
1	В	206	ASP	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 <sup>2</sup> )	Q < 0.9
4	EDO	А	502	4/4	0.90	0.11	33,39,46,46	0
3	PO4	А	501	5/5	0.92	0.18	73,79,97,108	0
5	NA	С	201	1/1	0.92	0.19	48,48,48,48	0

### 6.5 Other polymers (i)

There are no such residues in this entry.

