



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:21 AM GMT

PDB ID : 2QE2
Title : Structure of HCV NS5B Bound to an Anthranilic Acid Inhibitor
Authors : Chopra, R.; Svenson, K.; Bard, J.
Deposited on : 2007-06-22
Resolution : 2.90 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
<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 (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

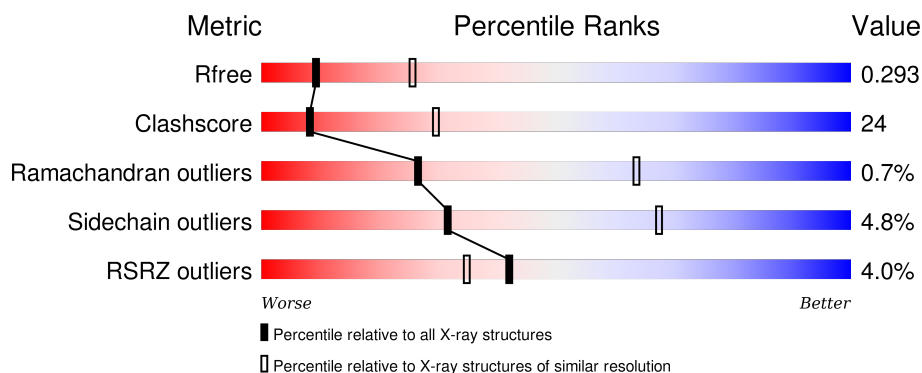
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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	578	<div> <div>3%</div> <div> <div></div> <div>52%</div> <div>35%</div> <div>•</div> <div>10%</div> </div> </div>
1	B	578	<div> <div>4%</div> <div> <div></div> <div>52%</div> <div>34%</div> <div>•</div> <div>10%</div> </div> </div>

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 8193 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.

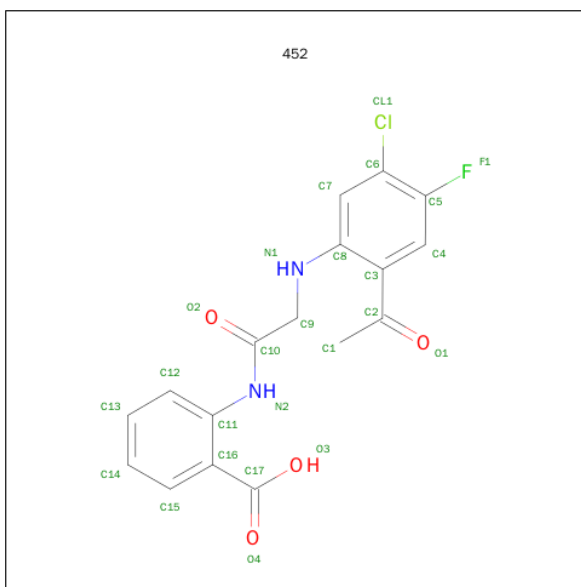
- Molecule 1 is a protein called RNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	521	Total	C	N	O	S	0	0	0
			4065	2567	721	748	29			
1	B	523	Total	C	N	O	S	0	0	0
			4078	2573	725	751	29			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	571	GLY	-	EXPRESSION TAG	UNP Q99AU2
A	572	SER	-	EXPRESSION TAG	UNP Q99AU2
A	573	HIS	-	EXPRESSION TAG	UNP Q99AU2
A	574	HIS	-	EXPRESSION TAG	UNP Q99AU2
A	575	HIS	-	EXPRESSION TAG	UNP Q99AU2
A	576	HIS	-	EXPRESSION TAG	UNP Q99AU2
A	577	HIS	-	EXPRESSION TAG	UNP Q99AU2
A	578	HIS	-	EXPRESSION TAG	UNP Q99AU2
B	571	GLY	-	EXPRESSION TAG	UNP Q99AU2
B	572	SER	-	EXPRESSION TAG	UNP Q99AU2
B	573	HIS	-	EXPRESSION TAG	UNP Q99AU2
B	574	HIS	-	EXPRESSION TAG	UNP Q99AU2
B	575	HIS	-	EXPRESSION TAG	UNP Q99AU2
B	576	HIS	-	EXPRESSION TAG	UNP Q99AU2
B	577	HIS	-	EXPRESSION TAG	UNP Q99AU2
B	578	HIS	-	EXPRESSION TAG	UNP Q99AU2

- Molecule 2 is 2-{[N-(2-ACETYL-5-CHLORO-4-FLUOROPHENYL)GLYCYL]AMINO}BE NZOIC ACID (three-letter code: 452) (formula: C₁₇H₁₄ClFN₂O₄).

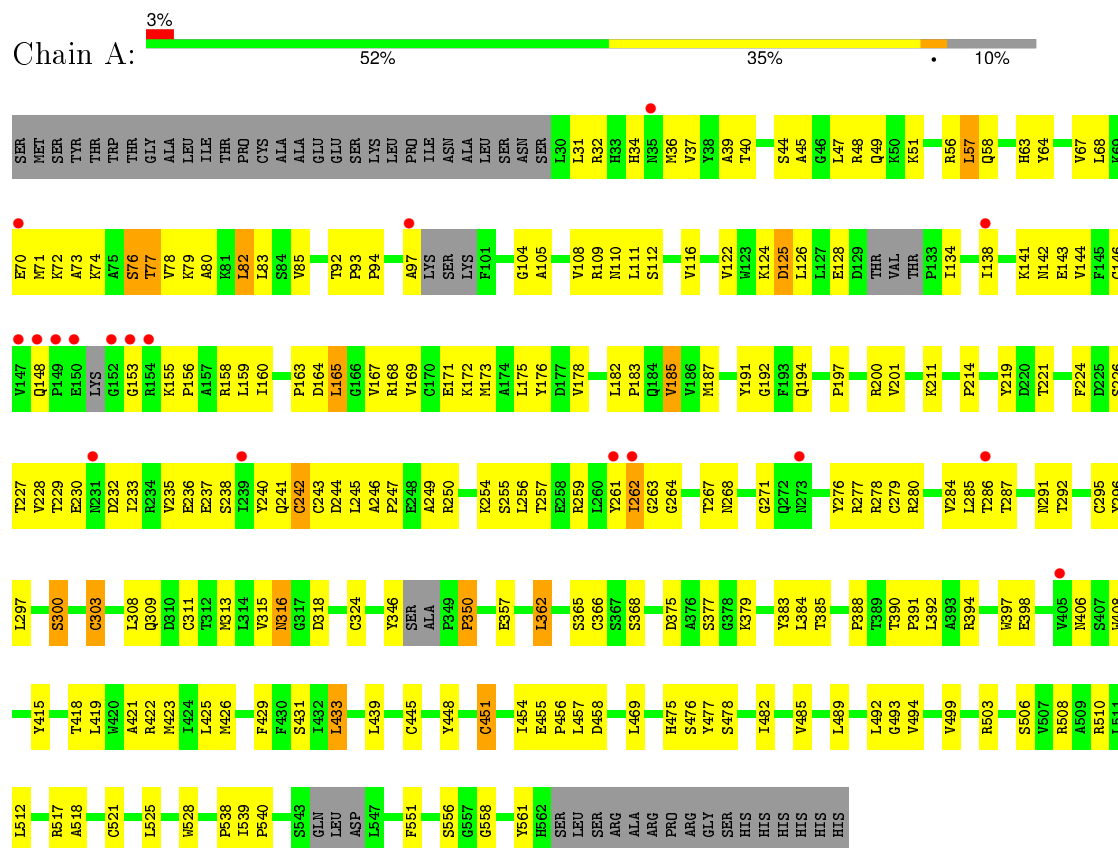


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	Cl	F	N	O		
2	B	1	25	17	1	1	2	4	0	0
2	A	1	25	17	1	1	2	4	0	0

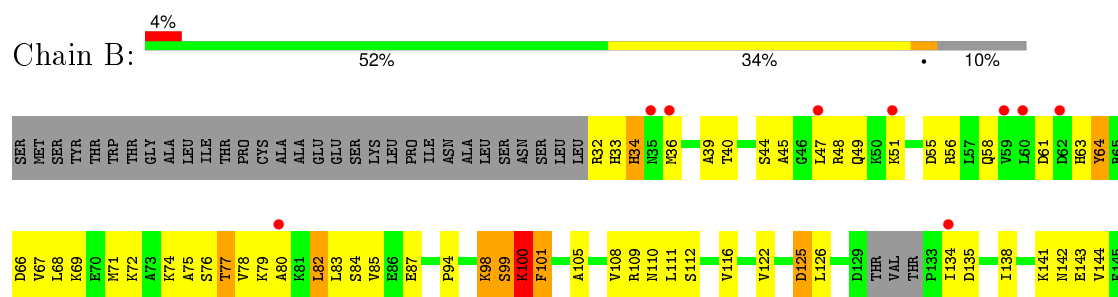
3 Residue-property plots

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: RNA-directed RNA polymerase



• Molecule 1: RNA-directed RNA polymerase



L384	I482	L394	G290	T221	C146
T385	V485	T386	T291	Y147	Y147
P388		P389	T292	F224	Q148
T389	P495	T390	L293		P149
P391	P496	T391	Y296	V228	E150
L392	V499	T392	L297	T229	K151
A393		L393	S300	E230	G152
R394	R503	A394	A301	T231	G153
K397	W528	E398	C302	D232	R154
E398	P538		R304	T233	K155
				R234	P156
				V235	A157
				E236	R158
				E237	L159
					I160
					V161
A542	A542	P403	L308	Y240	F162
S543	S543	P404	Q309	Q241	P163
GLN	GLN	V405	D310	C242	D164
LEU	LEU	S407	T312	C243	G165
ASP	ASP	W408	K313	D244	G166
L547	L547		L314	L245	V167
		Y415	V315	A246	
F551	F551		N316	P247	R168
S556	S556	T418	G317	E248	V169
G558	G558	L419	D318	A249	C170
D559	D559	R422	D319	R250	E171
I560	I560	N423	L320		K172
Y561	Y561	L425	V321	K254	M173
H562	H562	I424	V322	S255	A174
SER	SER	F429	I323	L256	L175
LEU	LEU	F430	C324	T257	Y176
SER	SER	S431	G324	E258	D177
ARG	ARG	I432	E325	R259	V178
ALA	ALA	L433	F339	L260	V179
ARG	ARG	L439	N343	Y261	
PRO	PRO	C445	Y346	I262	L182
ARG	ARG		SER	G263	P183
GLY	GLY		ALA	G264	Q184
SER	SER	Y448	P349	P265	V185
HIS	HIS		P350	N268	V186
HIS	HIS		D359		M187
HIS	HIS	C451	I362	G271	
HIS	HIS	I454	S365	Q272	S190
HIS	HIS	E455	C366	W273	Y191
HIS	HIS	P456	R278	Y276	G192
		L457	S367	R277	F193
		D458	S368	C279	Q194
				R280	Y195
		R465	H374	A281	S196
				S282	P197
		L469	S377	G283	
		H475	G378	V284	R200
		S476	K379	L285	V201
		Y477	R380	T286	K211
		S478	Y383	T287	P214
				S288	Y219
				C289	D220

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	70.83Å 70.88Å 251.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.77 – 2.90 49.14 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.9 (19.77-2.90) 99.0 (49.14-2.40)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.91 (at 2.39Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.279 , 0.296 0.276 , 0.293	Depositor DCC
R_{free} test set	1470 reflections (5.36%)	DCC
Wilson B-factor (Å ²)	42.3	Xtriage
Anisotropy	0.707	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 21.8	EDS
Estimated twinning fraction	0.488 for k,h,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 50241 reflections	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	8193	wwPDB-VP
Average B, all atoms (Å ²)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.73% 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, 0.375 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: 452

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	0.48	0/4151	0.77	2/5622 (0.0%)
1	B	0.51	1/4166 (0.0%)	0.80	9/5643 (0.2%)
All	All	0.49	1/8317 (0.0%)	0.78	11/11265 (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	B	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	100	LYS	C-O	-6.71	1.10	1.23

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	151	LYS	N-CA-C	6.45	128.42	111.00
1	A	558	GLY	N-CA-C	-5.78	98.66	113.10
1	B	558	GLY	N-CA-C	-5.67	98.92	113.10
1	B	99	SER	N-CA-C	-5.62	95.83	111.00
1	B	101	PHE	N-CA-C	5.62	126.17	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	100	LYS	Mainchain

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	A	4065	0	4071	204	0
1	B	4078	0	4084	198	0
2	A	25	0	13	0	0
2	B	25	0	13	0	0
All	All	8193	0	8181	400	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 24.

The worst 5 of 400 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:261:TYR:HE1	1:A:284:VAL:HG11	1.17	1.06
1:B:230:GLU:HB3	1:B:262:ILE:HD11	1.39	1.03
1:A:254:LYS:HE3	1:B:254:LYS:HE3	1.42	0.97
1:B:377:SER:HB2	1:B:379:LYS:HG3	1.45	0.96
1:A:230:GLU:HB3	1:A:262:ILE:HD11	1.48	0.94

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	509/578 (88%)	466 (92%)	41 (8%)	2 (0%)	39	74
1	B	515/578 (89%)	464 (90%)	46 (9%)	5 (1%)	19	54
All	All	1024/1156 (89%)	930 (91%)	87 (8%)	7 (1%)	26	63

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	98	LYS
1	A	165	LEU
1	B	34	HIS
1	B	33	HIS
1	A	271	GLY

5.3.2 Protein sidechains ⓘ

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	443/492 (90%)	420 (95%)	23 (5%)	29	64
1	B	444/492 (90%)	424 (96%)	20 (4%)	34	70
All	All	887/984 (90%)	844 (95%)	43 (5%)	31	67

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	433	LEU
1	B	77	THR
1	B	425	LEU
1	A	451	CYS
1	A	456	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	213	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	241	GLN
1	B	194	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 ⓘ

2 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	452	A	579	-	23,26,26	2.40	12 (52%)	32,36,36	1.34	5 (15%)
2	452	B	579	-	23,26,26	2.51	12 (52%)	32,36,36	1.35	6 (18%)

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	452	A	579	-	-	0/13/17/17	0/2/2/2
2	452	B	579	-	-	0/13/17/17	0/2/2/2

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	579	452	C10-N2	2.05	1.40	1.35
2	A	579	452	C11-N2	2.08	1.45	1.41
2	A	579	452	C8-N1	2.08	1.43	1.37
2	A	579	452	C3-C2	2.20	1.53	1.48
2	B	579	452	C9-N1	2.30	1.49	1.45

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	579	452	C8-C3-C2	-2.86	120.89	122.39
2	B	579	452	C15-C16-C17	-2.16	116.91	120.23
2	A	579	452	C5-C6-CL1	2.06	121.95	120.06
2	B	579	452	C5-C6-CL1	2.07	121.96	120.06
2	B	579	452	C8-C7-C6	2.15	123.42	119.78

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, 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	521/578 (90%)	0.01	18 (3%) 48 40	18, 43, 77, 119	0
1	B	523/578 (90%)	0.04	24 (4%) 36 30	19, 43, 79, 115	0
All	All	1044/1156 (90%)	0.03	42 (4%) 42 35	18, 43, 78, 119	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	152	GLY	9.8
1	A	150	GLU	7.4
1	B	150	GLU	6.5
1	B	149	PRO	5.5
1	B	152	GLY	5.0

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, 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(\AA^2)	Q<0.9
2	452	A	579	25/25	0.93	0.16	0.50	22,32,37,42	0
2	452	B	579	25/25	0.93	0.15	0.05	22,31,37,39	0

6.5 Other polymers [i](#)

There are no such residues in this entry.