



Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 10:26 am BST

PDB ID : 2H5K
Title : Crystal Structure of Complex Between the Domain-Swapped Dimeric Grb2 SH2 Domain and Shc-Derived Ligand, Ac-NH-pTyr-Val-Asn-NH2
Authors : Benfield, A.P.; Whiddon, B.B.; Martin, S.F.
Deposited on : 2006-05-26
Resolution : 3.25 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

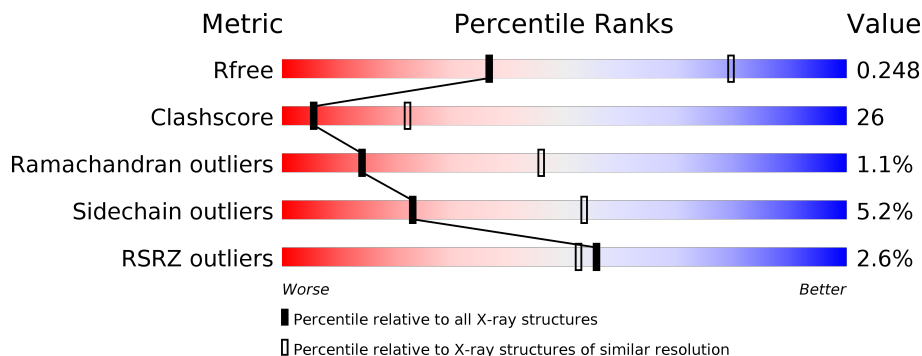
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.25 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	1191 (3.30-3.22)
Clashscore	141614	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)
RSRZ outliers	127900	1154 (3.30-3.22)

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 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 $\leq 5\%$. 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	116	 45% 37% 17%
1	B	116	 43% 38% 17%
2	C	5	 80% 20%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	CAC	B	10	-	-	-	X

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 1631 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 Growth factor receptor-bound protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	96	792	509	142	140	1	0	0	0
1	B	96	790	508	142	139	1	0	0	0

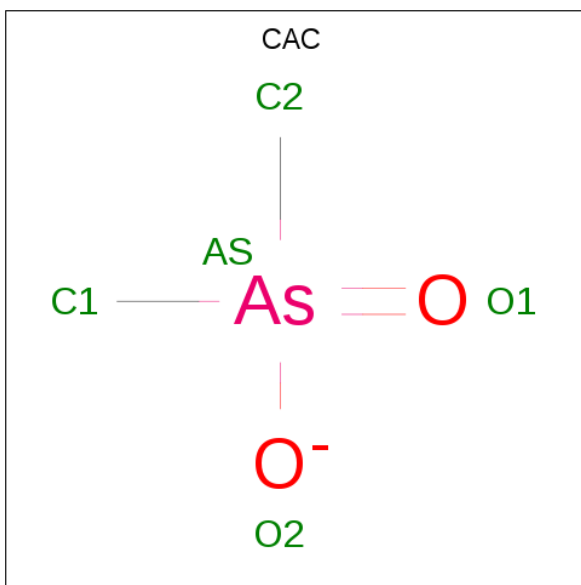
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	163	HIS	-	EXPRESSION TAG	UNP P62993
A	164	HIS	-	EXPRESSION TAG	UNP P62993
A	165	HIS	-	EXPRESSION TAG	UNP P62993
A	166	HIS	-	EXPRESSION TAG	UNP P62993
A	167	HIS	-	EXPRESSION TAG	UNP P62993
A	168	HIS	-	EXPRESSION TAG	UNP P62993
B	163	HIS	-	EXPRESSION TAG	UNP P62993
B	164	HIS	-	EXPRESSION TAG	UNP P62993
B	165	HIS	-	EXPRESSION TAG	UNP P62993
B	166	HIS	-	EXPRESSION TAG	UNP P62993
B	167	HIS	-	EXPRESSION TAG	UNP P62993
B	168	HIS	-	EXPRESSION TAG	UNP P62993

- Molecule 2 is a protein called Shc-Derived Ligand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	5	35	20	5	9	1	0	0	1

- Molecule 3 is CACODYLATE ION (three-letter code: CAC) (formula: C₂H₆AsO₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	As	C	O		
3	B	1	5	1	2	2	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	6	Total	O	0	0
			6	6		
4	B	3	Total	O	0	0
			3	3		

4 Data and refinement statistics

Property	Value	Source
Space group	P 62 2 2	Depositor
Cell constants a, b, c, α , β , γ	94.86Å 94.86Å 139.24Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.25 19.70 – 3.25	Depositor EDS
% Data completeness (in resolution range)	98.8 (20.00-3.25) 99.1 (19.70-3.25)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.56 (at 3.22Å)	Xtrriage
Refinement program	CNS	Depositor
R, R_{free}	0.247 , 0.300 0.241 , 0.248	Depositor DCC
R_{free} test set	284 reflections (4.59%)	wwPDB-VP
Wilson B-factor (Å ²)	57.0	Xtrriage
Anisotropy	0.045	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 47.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	1631	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.99% 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: CAC, PTR, ACE, NH2

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	0.57	0/814	0.69	0/1094
1	B	0.49	0/812	0.69	0/1091
2	C	0.45	0/14	0.62	0/18
All	All	0.53	0/1640	0.69	0/2203

There are no bond length outliers.

There are no bond angle outliers.

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	792	0	772	46	0
1	B	790	0	771	45	0
2	C	35	0	25	3	0
3	B	5	0	0	1	0
4	A	6	0	0	0	0
4	B	3	0	0	0	0
All	All	1631	0	1568	82	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 26.

All (82) 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:151:ILE:H	1:A:151:ILE:HD12	1.34	0.92
1:A:151:ILE:N	1:A:151:ILE:HD12	2.01	0.73
1:B:90:SER:O	1:B:92:PRO:HD3	1.90	0.72
1:B:127:SER:OG	1:B:130:GLU:HG3	1.92	0.69
1:B:83:PHE:HB3	1:B:99:VAL:HG22	1.75	0.68
1:A:128:LEU:HD21	1:B:110:VAL:HG11	1.76	0.68
1:A:100:LYS:HD3	1:A:101:PHE:N	2.09	0.68
1:B:149:ARG:HH11	1:B:149:ARG:HB2	1.59	0.67
1:B:83:PHE:CB	1:B:99:VAL:HG22	2.26	0.66
1:A:86:ARG:HH11	1:A:86:ARG:HG3	1.62	0.65
1:B:86:ARG:HH22	3:B:10:CAC:C2	2.09	0.64
1:A:100:LYS:C	1:A:100:LYS:HD3	2.18	0.64
1:A:151:ILE:CD1	1:A:151:ILE:H	2.11	0.62
1:B:149:ARG:CB	1:B:149:ARG:HH11	2.12	0.62
1:B:84:LEU:C	1:B:84:LEU:HD12	2.20	0.62
1:B:139:SER:HA	1:B:146:ILE:O	2.00	0.61
1:B:131:LEU:HD23	1:B:132:VAL:N	2.19	0.57
1:B:85:ILE:CD1	1:B:97:LEU:HD13	2.34	0.57
1:A:128:LEU:O	1:A:132:VAL:HG23	2.04	0.57
1:B:66:PRO:HA	1:B:89:GLU:OE2	2.05	0.56
1:A:127:SER:OG	1:A:130:GLU:HG3	2.05	0.56
1:B:97:LEU:O	1:B:107:HIS:HA	2.05	0.56
1:A:139:SER:HA	1:A:146:ILE:O	2.07	0.55
1:A:132:VAL:HG13	1:A:148:LEU:HD13	1.88	0.55
1:A:112:ARG:HA	1:A:117:LYS:O	2.07	0.55
1:A:61:PHE:CZ	1:A:63:GLY:HA2	2.43	0.54
1:A:83:PHE:CG	1:B:148:LEU:HD22	2.43	0.54
1:A:58:HIS:CE1	1:B:129:ASN:HD22	2.25	0.54
1:A:136:ARG:O	1:A:147:PHE:HB3	2.08	0.54
1:B:85:ILE:HD11	1:B:97:LEU:HD13	1.89	0.54
1:B:131:LEU:HD23	1:B:131:LEU:C	2.29	0.53
1:A:151:ILE:HG23	1:B:77:GLN:HG3	1.90	0.52
1:A:111:LEU:HB2	1:A:119:PHE:CE1	2.44	0.52
1:A:58:HIS:NE2	1:B:129:ASN:ND2	2.59	0.50
1:B:61:PHE:CE1	1:B:63:GLY:HA2	2.47	0.50
1:A:109:LYS:HG3	2:C:1:PTR:CG	2.42	0.49
1:A:86:ARG:HG3	1:A:86:ARG:NH1	2.27	0.49
1:A:83:PHE:O	1:B:151:ILE:HD11	2.14	0.48
1:B:61:PHE:HD1	1:B:85:ILE:O	1.98	0.47
1:B:125:PHE:HZ	1:B:134:TYR:CD2	2.33	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:84:LEU:HD12	1:A:84:LEU:C	2.35	0.47
1:B:82:ALA:O	1:B:99:VAL:HA	2.16	0.46
1:B:133:ASP:O	1:B:136:ARG:HG3	2.16	0.46
1:A:136:ARG:HG2	1:A:136:ARG:HH11	1.81	0.46
1:A:125:PHE:HB3	1:A:130:GLU:HB2	1.97	0.45
1:B:69:LYS:O	1:B:72:GLU:HB3	2.17	0.45
1:A:62:PHE:O	1:A:65:ILE:HG23	2.17	0.45
1:A:61:PHE:CE2	1:A:63:GLY:HA2	2.52	0.45
1:A:85:ILE:HD13	1:A:97:LEU:HD13	1.99	0.45
1:A:148:LEU:HD22	1:B:83:PHE:CG	2.51	0.45
1:B:143:ASN:O	1:B:144:GLN:HG3	2.16	0.45
1:A:117:LYS:HE2	1:B:126:ASN:OD1	2.17	0.44
1:B:118:TYR:O	1:B:119:PHE:HB3	2.17	0.44
1:A:58:HIS:HA	1:A:59:PRO:HD2	1.89	0.44
1:A:125:PHE:CD2	1:A:131:LEU:HA	2.52	0.44
1:A:85:ILE:CD1	1:A:97:LEU:HD13	2.48	0.44
2:C:1:PTR:CE1	2:C:1:PTR:O2P	2.66	0.44
1:A:137:SER:O	1:A:147:PHE:CE1	2.71	0.44
1:B:98:SER:OG	1:B:107:HIS:CE1	2.71	0.44
1:B:98:SER:HG	1:B:107:HIS:CE1	2.33	0.43
1:A:151:ILE:HG22	1:A:152:GLU:H	1.84	0.43
1:B:67:ARG:NH1	1:B:89:GLU:OE1	2.51	0.43
1:A:146:ILE:HG23	1:B:101:PHE:CD1	2.54	0.43
1:A:100:LYS:C	1:A:100:LYS:CD	2.86	0.42
1:B:98:SER:OG	1:B:107:HIS:ND1	2.44	0.42
1:A:139:SER:OG	1:A:144:GLN:O	2.28	0.42
1:A:63:GLY:O	1:A:87:GLU:HG3	2.20	0.42
1:B:85:ILE:HD13	1:B:97:LEU:HD13	2.01	0.42
1:B:83:PHE:HB2	1:B:97:LEU:HD11	2.02	0.42
1:A:67:ARG:HD2	2:C:1:PTR:O2P	2.19	0.41
1:B:83:PHE:HB2	1:B:99:VAL:HG22	1.99	0.41
1:B:133:ASP:O	1:B:134:TYR:C	2.59	0.41
1:B:142:ARG:O	1:B:142:ARG:HG2	2.19	0.41
1:A:65:ILE:HG13	1:A:65:ILE:O	2.20	0.41
1:A:136:ARG:HG2	1:A:136:ARG:NH1	2.35	0.41
1:A:64:LYS:HB2	1:A:87:GLU:OE2	2.21	0.41
1:A:131:LEU:HD21	1:B:120:LEU:HG	2.03	0.41
1:A:145:GLN:HA	1:A:145:GLN:OE1	2.21	0.41
1:B:100:LYS:HG2	1:B:101:PHE:N	2.36	0.41
1:A:135:HIS:CG	1:A:140:VAL:HG12	2.57	0.40
1:B:84:LEU:O	1:B:84:LEU:HD12	2.20	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:83:PHE:HA	1:B:98:SER:O	2.21	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	94/116 (81%)	83 (88%)	10 (11%)	1 (1%)	14	46
1	B	94/116 (81%)	74 (79%)	19 (20%)	1 (1%)	14	46
2	C	2/5 (40%)	1 (50%)	1 (50%)	0	100	100
All	All	190/237 (80%)	158 (83%)	30 (16%)	2 (1%)	14	46

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	90	SER
1	A	92	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	85/105 (81%)	80 (94%)	5 (6%)	19	50
1	B	85/105 (81%)	81 (95%)	4 (5%)	26	57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	C	2/2 (100%)	2 (100%)	0	100	100
All	All	172/212 (81%)	163 (95%)	9 (5%)	23	53

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	104	ASP
1	A	129	ASN
1	A	138	THR
1	A	142	ARG
1	A	151	ILE
1	B	65	ILE
1	B	96	SER
1	B	104	ASP
1	B	149	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	129	ASN
1	B	129	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

1 non-standard protein/DNA/RNA residue is 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	PTR	C	1	2	15,16,17	1.29	1 (6%)	19,22,24	1.13	2 (10%)

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	PTR	C	1	2	-	0/10/11/13	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1	PTR	P-O1P	3.34	1.61	1.50

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	C	1	PTR	O3P-P-OH	2.32	112.49	105.24
2	C	1	PTR	P-OH-CZ	-2.05	117.19	123.75

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1	PTR	3	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand is 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
3	CAC	B	10	-	0,4,4	0.00	-	0,6,6	0.00	-

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	10	CAC	1	0

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	96/116 (82%)	-0.24	1 (1%) 82 82	5, 28, 52, 72	3 (3%)
1	B	96/116 (82%)	0.05	4 (4%) 36 33	7, 41, 76, 84	3 (3%)
2	C	2/5 (40%)	0.14	0 100 100	20, 20, 20, 29	0
All	All	194/237 (81%)	-0.09	5 (2%) 56 52	5, 30, 73, 84	6 (3%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	57	PRO	3.6
1	B	103	ASN	2.6
1	A	57	PRO	2.5
1	B	78	ARG	2.5
1	B	76	LYS	2.5

6.2 Non-standard residues in protein, DNA, RNA chains [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, 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	B-factors(Å ²)	Q<0.9
2	PTR	C	1	16/17	0.94	0.18	13,15,17,18	0

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. 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	B-factors(Å ²)	Q<0.9
3	CAC	B	10	5/5	0.80	0.45	59,63,64,64	0

6.5 Other polymers [i](#)

There are no such residues in this entry.