



wwPDB X-ray Structure Validation Summary Report i

Feb 26, 2014 – 05:44 PM GMT

PDB ID : 3I3T
Title : Crystal structure of covalent ubiquitin-USP21 complex
Authors : Neculai, D.; Avvakumov, G.V.; Walker, J.R.; Xue, S.; Butler-Cole, C.; Weigelt, J.; Bountra, C.; Edwards, A.M.; Arrowsmith, C.H.; Bochkarev, A.; Dhe-Paganon, S.; Structural Genomics Consortium (SGC)
Deposited on : 2009-06-30
Resolution : 2.59 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

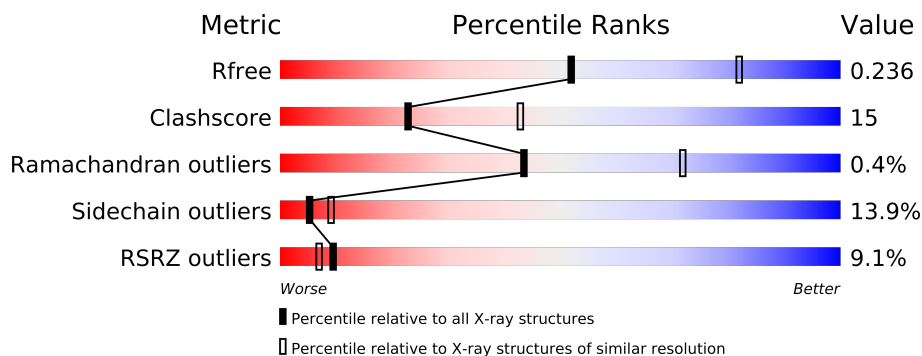
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.59 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}	66092	1718 (2.60-2.60)
Clashscore	79885	2154 (2.60-2.60)
Ramachandran outliers	78287	2113 (2.60-2.60)
Sidechain outliers	78261	2113 (2.60-2.60)
RSRZ outliers	66119	1718 (2.60-2.60)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	355	
1	C	355	
1	E	355	
1	G	355	
2	B	75	
2	D	75	
2	F	75	
2	H	75	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12456 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Ubiquitin carboxyl-terminal hydrolase 21.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	0	1	0
			2460	1549	442	452	17			
1	C	310	Total	C	N	O	S	0	1	0
			2506	1579	452	458	17			
1	E	310	Total	C	N	O	S	0	0	0
			2497	1574	450	456	17			
1	G	304	Total	C	N	O	S	0	1	0
			2468	1555	446	450	17			

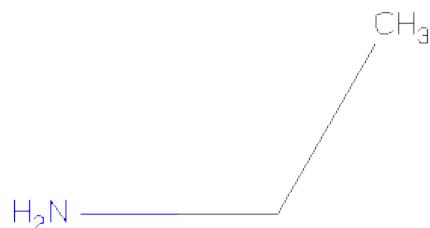
- Molecule 2 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	75	Total	C	N	O	S	0	0	0
			597	376	104	116	1			
2	D	75	Total	C	N	O	S	0	0	0
			597	376	104	116	1			
2	F	75	Total	C	N	O	S	0	0	0
			597	376	104	116	1			
2	H	75	Total	C	N	O	S	0	0	0
			597	376	104	116	1			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	G	1	Total	Zn	0	0
			1	1		
3	A	1	Total	Zn	0	0
			1	1		
3	C	1	Total	Zn	0	0
			1	1		
3	E	1	Total	Zn	0	0
			1	1		

- Molecule 4 is ETHANAMINE (three-letter code: NEH) (formula: C₂H₇N).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	N	0	0
			3	2	1		
4	D	1	Total	C	N	0	0
			3	2	1		
4	F	1	Total	C	N	0	0
			3	2	1		
4	H	1	Total	C	N	0	0
			3	2	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	20	Total	O	0	0
			20	20		
5	C	19	Total	O	0	0
			19	19		
5	E	31	Total	O	0	0
			31	31		
5	G	33	Total	O	0	0
			33	33		
5	B	2	Total	O	0	0
			2	2		
5	D	7	Total	O	0	0
			7	7		
5	F	6	Total	O	0	0
			6	6		

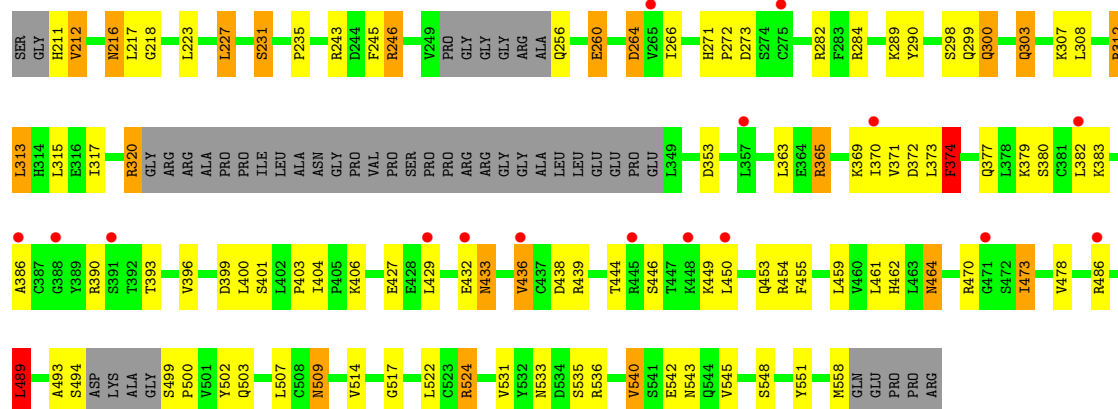
Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	H	3	Total	O	0	0
			3	3		

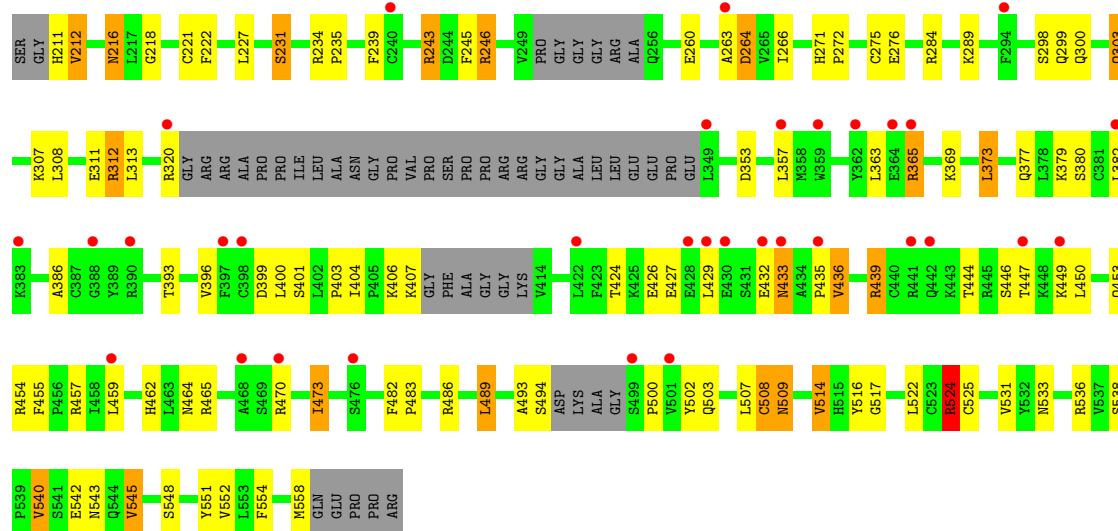
- Molecule 1: Ubiquitin carboxyl-terminal hydrolase 21

Chain E:



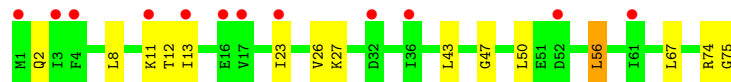
- Molecule 1: Ubiquitin carboxyl-terminal hydrolase 21

Chain G:



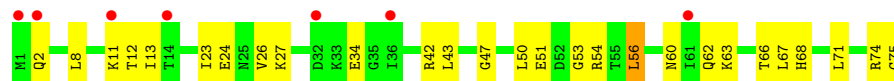
- Molecule 2: Ubiquitin

Chain B:

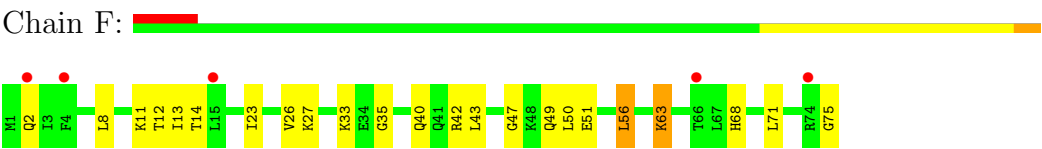


- Molecule 2: Ubiquitin

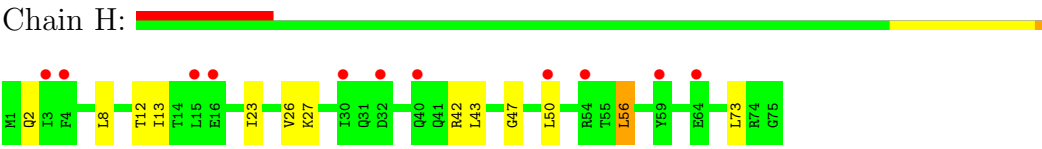
Chain D:



- Molecule 2: Ubiquitin



● Molecule 2: Ubiquitin



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	58.43Å 83.66Å 118.79Å 88.71° 75.73° 85.11°	Depositor
Resolution (Å)	83.37 – 2.59 83.36 – 2.59	Depositor EDS
% Data completeness (in resolution range)	95.8 (83.37-2.59) 99.3 (83.36-2.59)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.14 (at 2.58Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, R_{free}	0.186 , 0.218 0.209 , 0.236	Depositor DCC
R_{free} test set	3401 reflections (5.05%)	DCC
Wilson B-factor (Å ²)	46.9	Xtriage
Anisotropy	0.094	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 35.3	EDS
Estimated twinning fraction	0.540 for H, K, L 0.460 for -H, -K, -H+L 0.059 for -h,-k,-h+l	Xtriage
Reported twinning fraction	0.540 for H, K, L 0.460 for -H, -K, -H+L	Depositor
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 67375 reflections	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	12456	wwPDB-VP
Average B, all atoms (Å ²)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.91% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	1.09	5/2509 (0.2%)	1.17	17/3382 (0.5%)
1	C	1.01	3/2557 (0.1%)	1.02	8/3445 (0.2%)
1	E	1.07	5/2548 (0.2%)	1.05	10/3433 (0.3%)
1	G	1.01	4/2520 (0.2%)	1.05	8/3395 (0.2%)
2	B	0.79	0/603	0.84	0/811
2	D	0.85	0/603	0.89	1/811 (0.1%)
2	F	0.79	0/603	0.87	0/811
2	H	0.81	0/603	0.89	1/811 (0.1%)
All	All	1.00	17/12546 (0.1%)	1.04	45/16899 (0.3%)

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	545	VAL	CB-CG2	-13.44	1.24	1.52
1	A	545	VAL	CB-CG1	-13.13	1.25	1.52
1	E	374	PHE	CE1-CZ	-10.01	1.18	1.37
1	E	374	PHE	CG-CD1	-9.85	1.24	1.38
1	E	374	PHE	CE2-CZ	-9.35	1.19	1.37

The worst 5 of 45 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	365	ARG	NE-CZ-NH1	-15.99	112.31	120.30
1	A	365	ARG	NE-CZ-NH2	15.76	128.18	120.30
1	A	545	VAL	CG1-CB-CG2	-15.26	86.48	110.90
1	A	246	ARG	NE-CZ-NH1	11.43	126.01	120.30
1	E	264	ASP	CB-CG-OD1	-11.36	108.08	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2460	0	2412	77	0
1	C	2506	0	2464	96	0
1	E	2497	0	2457	103	0
1	G	2468	0	2433	67	0
2	B	597	0	626	13	0
2	D	597	0	626	26	0
2	F	597	0	626	23	0
2	H	597	0	626	13	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
3	E	1	0	0	0	0
3	G	1	0	0	0	0
4	B	3	0	5	1	0
4	D	3	0	4	0	0
4	F	3	0	4	0	0
4	H	3	0	5	1	0
5	A	20	0	0	3	0
5	B	2	0	0	0	0
5	C	19	0	0	3	0
5	D	7	0	0	4	0
5	E	31	0	0	3	0
5	F	6	0	0	1	0
5	G	33	0	0	0	0
5	H	3	0	0	1	0
All	All	12456	0	12288	363	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 15.

The worst 5 of 363 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:E:503:GLN:NE2	1:E:524:ARG:HH12	1.12	1.42
1:G:503:GLN:NE2	1:G:524:ARG:HH12	1.07	1.42
1:G:503:GLN:HE21	1:G:524:ARG:NH1	1.15	1.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:289:LYS:HE2	1:E:315:LEU:CD2	1.48	1.40
1:C:503:GLN:NE2	1:C:524:ARG:HH12	1.26	1.29

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	294/355 (83%)	272 (92%)	21 (7%)	1 (0%)	50	77
1	C	303/355 (85%)	287 (95%)	14 (5%)	2 (1%)	30	58
1	E	302/355 (85%)	279 (92%)	22 (7%)	1 (0%)	50	77
1	G	295/355 (83%)	272 (92%)	21 (7%)	2 (1%)	30	58
2	B	73/75 (97%)	68 (93%)	5 (7%)	0	100	100
2	D	73/75 (97%)	71 (97%)	2 (3%)	0	100	100
2	F	73/75 (97%)	69 (94%)	4 (6%)	0	100	100
2	H	73/75 (97%)	69 (94%)	4 (6%)	0	100	100
All	All	1486/1720 (86%)	1387 (93%)	93 (6%)	6 (0%)	43	72

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	478	VAL
1	E	478	VAL
1	G	263	ALA
1	G	465	ARG
1	C	478	VAL

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	276/309 (89%)	232 (84%)	44 (16%)	4	6
1	C	279/309 (90%)	239 (86%)	40 (14%)	5	8
1	E	278/309 (90%)	237 (85%)	41 (15%)	4	8
1	G	277/309 (90%)	229 (83%)	48 (17%)	3	4
2	B	68/68 (100%)	63 (93%)	5 (7%)	20	38
2	D	68/68 (100%)	61 (90%)	7 (10%)	10	19
2	F	68/68 (100%)	63 (93%)	5 (7%)	20	38
2	H	68/68 (100%)	65 (96%)	3 (4%)	39	68
All	All	1382/1508 (92%)	1189 (86%)	193 (14%)	5	9

5 of 193 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	227	LEU
1	E	433	ASN
2	D	8	LEU
1	E	271	HIS
1	E	313	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 45 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	219	ASN
1	E	464	ASN
2	F	40	GLN
1	E	299	GLN
1	E	503	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
4	NEH	B	76	2	2,2,2	0.45	0	1,1,1	0.14	0
4	NEH	D	76	2	2,2,2	0.62	0	1,1,1	0.23	0
4	NEH	F	76	2	2,2,2	0.31	0	1,1,1	0.16	0
4	NEH	H	76	2	2,2,2	0.62	0	1,1,1	0.06	0

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	NEH	B	76	2	-	0/0/0/0	0/0/0/0
4	NEH	D	76	2	-	0/0/0/0	0/0/0/0
4	NEH	F	76	2	-	0/0/0/0	0/0/0/0
4	NEH	H	76	2	-	0/0/0/0	0/0/0/0

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.

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	303/355 (85%)	0.82	37 (12%) 5 4	22, 39, 70, 83	0
1	C	310/355 (87%)	0.77	19 (6%) 21 17	19, 39, 70, 83	0
1	E	310/355 (87%)	0.80	16 (5%) 26 22	22, 39, 70, 83	0
1	G	304/355 (85%)	0.80	33 (10%) 6 5	22, 39, 70, 83	0
2	B	75/75 (100%)	1.15	12 (16%) 3 1	27, 49, 61, 62	0
2	D	75/75 (100%)	0.79	7 (9%) 9 6	26, 49, 61, 61	0
2	F	75/75 (100%)	0.83	5 (6%) 17 15	27, 49, 61, 62	1 (1%)
2	H	75/75 (100%)	0.97	11 (14%) 3 2	27, 49, 61, 62	1 (1%)
All	All	1527/1720 (88%)	0.82	140 (9%) 9 6	19, 42, 67, 83	2 (0%)

The worst 5 of 140 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	436	VAL	8.1
1	A	440	CYS	6.1
1	A	451	THR	6.0
2	B	32	ASP	6.0
2	B	1	MET	5.3

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 carbohydrates 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. 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
4	NEH	D	76	3/3	0.20	0.64	22,22,22,23	0
4	NEH	F	76	3/3	0.19	0.07	21,21,22,24	0
4	NEH	B	76	3/3	0.16	-0.08	23,23,24,24	0
3	ZN	C	700	1/1	0.07	-1.68	71,71,71,71	0
4	NEH	H	76	3/3	0.12	-1.68	29,29,30,31	0
3	ZN	E	700	1/1	0.09	-1.75	68,68,68,68	0
3	ZN	A	700	1/1	0.05	-1.99	74,74,74,74	0
3	ZN	G	700	1/1	0.03	-3.03	74,74,74,74	0

6.5 Other polymers

There are no such residues in this entry.