wwPDB X-ray Structure Validation Summary Report (i)

Feb 28, 2014 – 08:48 AM GMT

PDB ID	:	1UNG
Title	:	STRUCTURAL MECHANISM FOR THE INHIBITION OF CDK5-P25 BY ROSCOVITINE, ALOISINE AND INDIRUBIN.
Authors	:	Mapelli, M.; Crovace, C.; Massimiliano, L.; Musacchio, A.
Deposited on	:	2003-09-10
Resolution	:	2.30 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

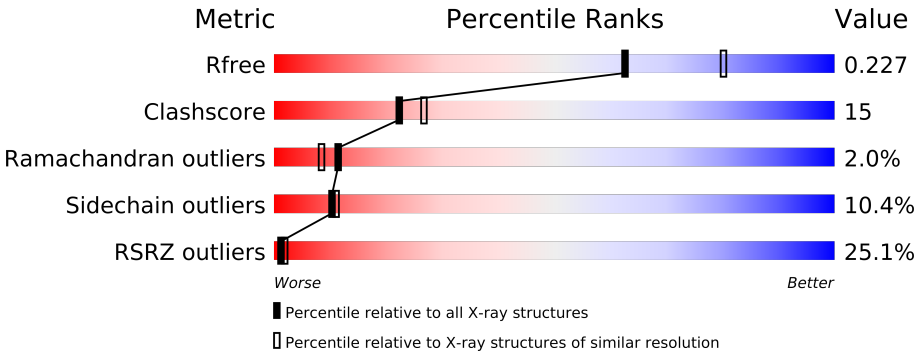
The following versions of software and data (see references) were used in the production of this report:

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 2.30 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	2929 (2.30-2.30)
Clashscore	79885	3679 (2.30-2.30)
Ramachandran outliers	78287	3642 (2.30-2.30)
Sidechain outliers	78261	3641 (2.30-2.30)
RSRZ outliers	66119	2930 (2.30-2.30)

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	292	
1	B	292	
2	D	208	
2	E	208	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7204 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 CELL DIVISION PROTEIN KINASE 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	288	Total	C	N	O	S	0	0	0
			2322	1492	399	419	12			
1	B	269	Total	C	N	O	S	0	0	1
			2155	1391	373	381	10			

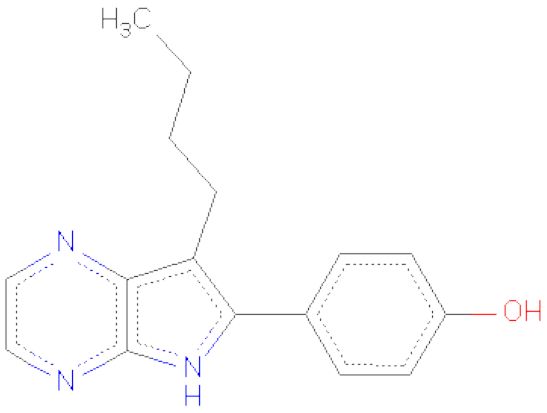
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	144	ASN	ASP	ENGINEERED MUTATION	UNP Q00535
B	144	ASN	ASP	ENGINEERED MUTATION	UNP Q00535

- Molecule 2 is a protein called CYCLIN-DEPENDENT KINASE 5 ACTIVATOR 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	150	Total	C	N	O	S	0	0	1
			1202	771	199	221	11			
2	E	148	Total	C	N	O	S	0	0	1
			1191	765	197	218	11			

- Molecule 3 is 6-PHENYL[5H]PYRROLO[2,3-B]PYRAZINE (three-letter code: ALH) (formula: $C_{16}H_{17}N_3O$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 20 16 3 1	0	0
3	B	1	Total C N O 20 16 3 1	0	0

- Molecule 4 is water.

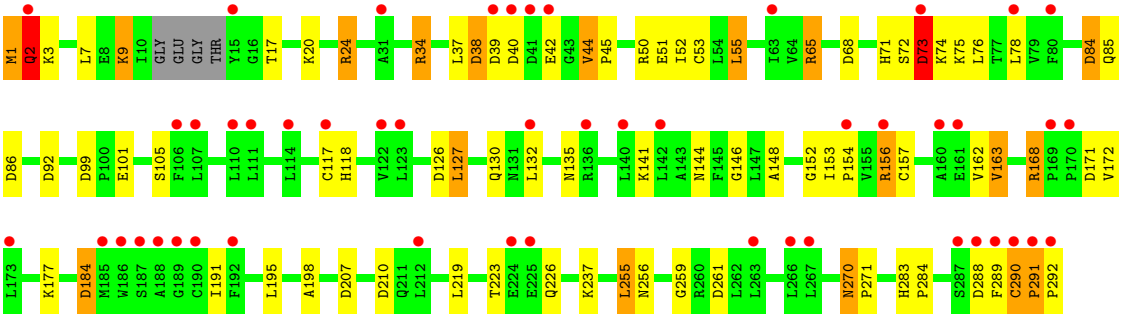
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	146	Total O 146 146	0	0
4	B	76	Total O 76 76	0	0
4	D	61	Total O 61 61	0	0
4	E	11	Total O 11 11	0	0

3 Residue-property plots (i)

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.

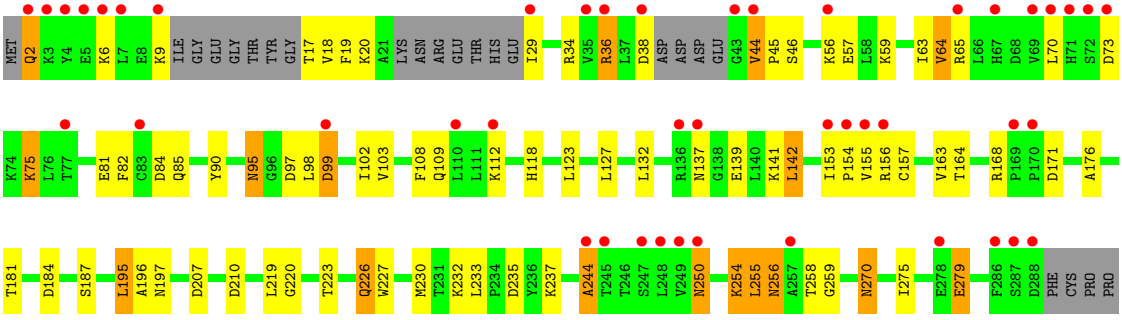
• Molecule 1: CELL DIVISION PROTEIN KINASE 5

Chain A:



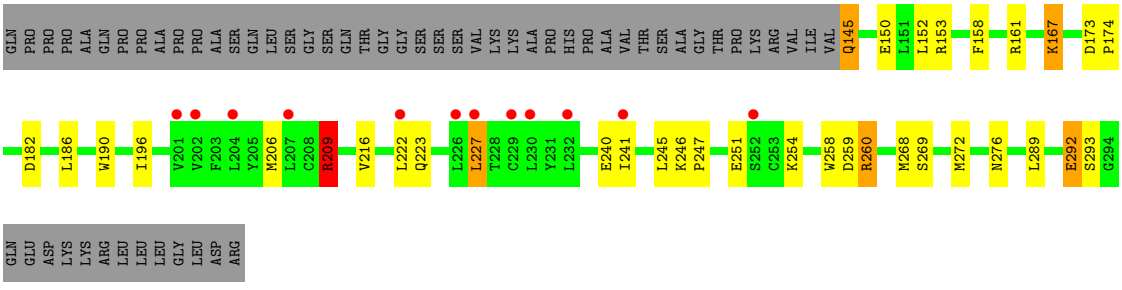
• Molecule 1: CELL DIVISION PROTEIN KINASE 5

Chain B:



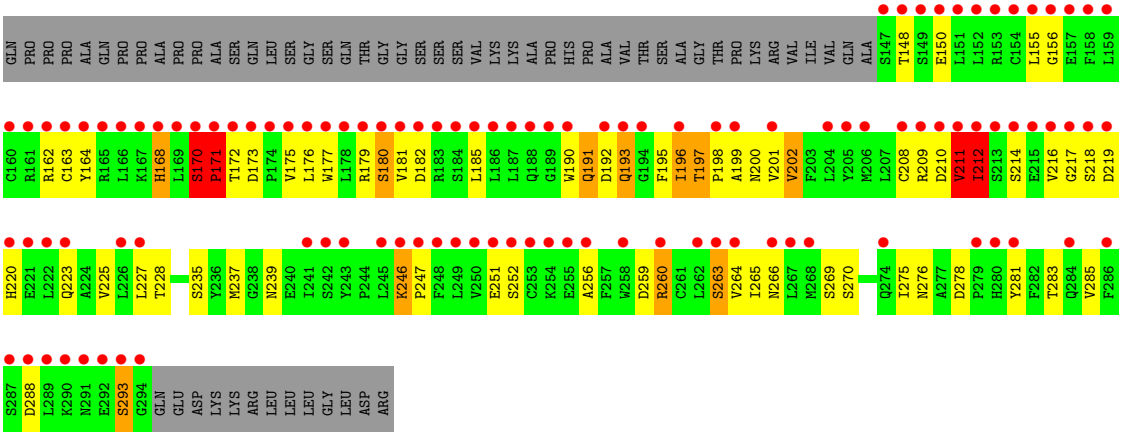
• Molecule 2: CYCLIN-DEPENDENT KINASE 5 ACTIVATOR 1

Chain D:



• Molecule 2: CYCLIN-DEPENDENT KINASE 5 ACTIVATOR 1

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	117.72Å 117.72Å 156.79Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.76 – 2.30 19.73 – 2.30	Depositor EDS
% Data completeness (in resolution range)	95.5 (19.76-2.30) 95.5 (19.73-2.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.34 (at 2.30Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.216 , 0.225 0.215 , 0.227	Depositor DCC
R_{free} test set	2713 reflections (5.33%)	DCC
Wilson B-factor (Å^2)	30.1	Xtriage
Anisotropy	0.078	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.44 , 48.1	EDS
Estimated twinning fraction	0.025 for -h,-k,l	Xtriage
L-test for twinning	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 53636 reflections	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	7204	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.02% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities 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: ALH

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	1.01	2/2377 (0.1%)	1.08	19/3216 (0.6%)
1	B	0.80	0/2203	0.95	12/2979 (0.4%)
2	D	1.03	1/1230 (0.1%)	1.01	5/1669 (0.3%)
2	E	0.92	2/1219 (0.2%)	0.97	11/1654 (0.7%)
All	All	0.94	5/7029 (0.1%)	1.01	47/9518 (0.5%)

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 (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	171	PRO	N-CD	20.82	1.76	1.47
1	A	117	CYS	CB-SG	-5.98	1.72	1.81
1	A	135	ASN	CB-CG	-5.75	1.37	1.51
2	D	145	GLN	C-O	-5.37	1.13	1.23
2	E	210	ASP	N-CA	-5.16	1.36	1.46

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	65	ARG	NE-CZ-NH2	-11.86	114.37	120.30
2	D	209	ARG	NE-CZ-NH2	-11.57	114.52	120.30
2	E	171	PRO	CA-N-CD	-10.23	97.17	111.50
1	A	288	ASP	CB-CG-OD2	8.45	125.90	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	97	ASP	CB-CG-OD2	7.80	125.32	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	244	ALA	Peptide

5.2 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 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	2322	0	2345	69	0
1	B	2155	0	2200	68	0
2	D	1202	0	1187	32	0
2	E	1191	0	1177	46	0
3	A	20	0	16	9	0
3	B	20	0	16	7	0
4	A	146	0	0	5	0
4	B	76	0	0	3	1
4	D	61	0	0	8	0
4	E	11	0	0	2	0
All	All	7204	0	6941	206	1

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 206 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:E:171:PRO:N	2:E:171:PRO:CD	1.77	1.27
1:B:56:LYS:NZ	2:E:265:ILE:HG22	1.52	1.24
1:A:53:CYS:SG	4:D:2004:HOH:O	1.95	1.20
1:A:53:CYS:HB3	4:D:2004:HOH:O	1.51	1.10
2:D:158:PHE:HB2	2:D:292:GLU:HG3	1.37	1.07

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
4:B:2019:HOH:O	4:B:2019:HOH:O[5_555]	1.91	0.29

5.3 Torsion angles

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	284/292 (97%)	261 (92%)	17 (6%)	6 (2%)	11 8
1	B	261/292 (89%)	240 (92%)	17 (6%)	4 (2%)	15 13
2	D	148/208 (71%)	146 (99%)	2 (1%)	0	100 100
2	E	146/208 (70%)	113 (77%)	26 (18%)	7 (5%)	4 1
All	All	839/1000 (84%)	760 (91%)	62 (7%)	17 (2%)	11 8

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	GLN
1	A	44	VAL
1	B	44	VAL
2	E	170	SER
1	A	42	GLU

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	258/260 (99%)	239 (93%)	19 (7%)	20 24
1	B	239/260 (92%)	216 (90%)	23 (10%)	12 13

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	D	138/186 (74%)	132 (96%)	6 (4%)	40 52
2	E	137/186 (74%)	105 (77%)	32 (23%)	1 1
All	All	772/892 (86%)	692 (90%)	80 (10%)	10 11

5 of 80 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	254	LYS
2	D	209	ARG
2	E	235	SER
1	B	255	LEU
1	B	270	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	95	ASN
1	B	197	ASN
2	E	239	ASN
1	B	118	HIS
1	B	131	ASN

5.3.3 RNA (i)

There are no RNA chains 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 carbohydrates in this entry.

5.6 Ligand geometry (i)

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	$\#\|Z\| > 2$
3	ALH	A	1293	-	22,22,22	2.42	4 (18%)	28,30,30	8.39	12 (42%)
3	ALH	B	1288	-	22,22,22	2.38	3 (13%)	28,30,30	5.59	11 (39%)

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
3	ALH	A	1293	-	-	0/8/8/8	0/1/3/3
3	ALH	B	1288	-	-	0/8/8/8	0/1/3/3

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1293	ALH	CAR-NAQ	7.45	1.44	1.34
3	B	1288	ALH	CAR-NAQ	7.06	1.43	1.34
3	A	1293	ALH	CAM-CAL	-6.08	1.39	1.47
3	B	1288	ALH	CAM-CAL	-5.98	1.39	1.47
3	B	1288	ALH	CAJ-NAI	5.07	1.44	1.37

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1293	ALH	CAR-CAJ-NAI	-40.42	119.64	124.05
3	B	1288	ALH	CAR-CAJ-NAI	-24.66	121.36	124.05
3	A	1293	ALH	CAE-CAM-CAL	-8.73	114.17	120.65
3	A	1293	ALH	CAL-CAK-CAJ	-7.61	99.60	106.58
3	B	1288	ALH	CAJ-CAR-NAS	7.52	111.84	106.09

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers (i)

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	288/292 (98%)	1.00	49 (17%) 2 4	14, 21, 38, 60	0
1	B	269/292 (92%)	0.96	45 (16%) 2 4	14, 24, 34, 41	0
2	D	150/208 (72%)	0.73	12 (8%) 12 19	17, 21, 30, 33	0
2	E	148/208 (71%)	4.20	109 (73%) 0 0	16, 26, 30, 31	0
All	All	855/1000 (85%)	1.49	215 (25%) 1 2	14, 23, 34, 60	0

The worst 5 of 215 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	187	LEU	15.2
1	A	292	PRO	12.4
2	E	212	ILE	10.8
2	E	170	SER	10.6
2	E	171	PRO	10.6

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, 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	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	ALH	B	1288	20/20	0.18	0.22	22,26,31,32	0
3	ALH	A	1293	20/20	0.18	-0.18	18,23,29,34	0

6.5 Other polymers (i)

There are no such residues in this entry.