Full wwPDB X-ray Structure Validation Report (i)

Jan 21, 2024 – 01:09 am GMT

PDB ID	:	7AR4
Title	:	Crystal structure of beta-catenin in complex with cyclic peptide inhibitor
Authors	:	Wendt, M.; Pearce, N.M.; Grossmann, T.N.; Hennig, S.
Deposited on	:	2020-10-23
Resolution	:	2.60 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:

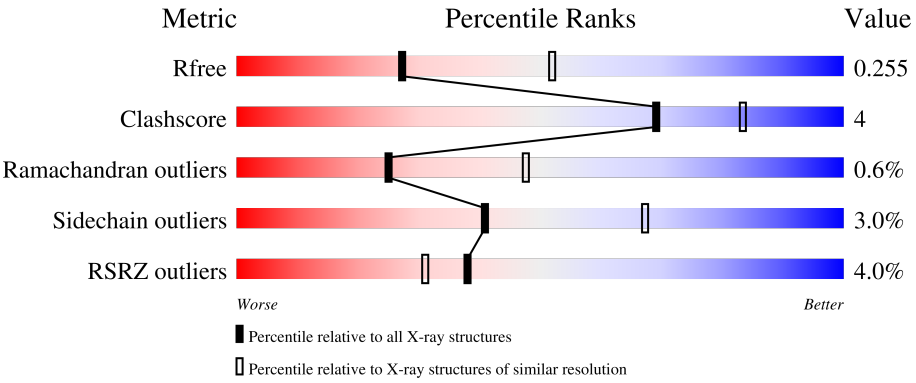
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (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 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	AAA	536	3% 85% 10%
2	PaP	16	19% 88% 12%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8287 atoms, of which 4203 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 Catenin beta-1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	AAA	512	7967	2451	4065	707	717	27	208	1	1

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	130	GLY	-	expression tag	UNP P35222
AAA	131	ALA	-	expression tag	UNP P35222
AAA	132	MET	-	expression tag	UNP P35222
AAA	133	ALA	-	expression tag	UNP P35222

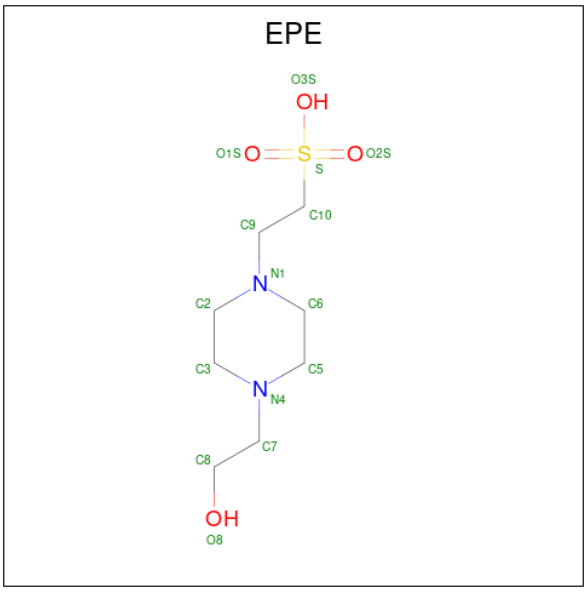
- Molecule 2 is a protein called Cadherin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
2	PaP	16	241	77	121	20	23	4	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
PaP	1	BAL	GLU	engineered mutation	UNP P09803
PaP	8	DPR	ALA	engineered mutation	UNP P09803
PaP	10	ASP	THR	engineered mutation	UNP P09803
PaP	11	SER	LEU	engineered mutation	UNP P09803
PaP	12	LEU	MET	engineered mutation	UNP P09803
PaP	13	LEU	SER	engineered mutation	UNP P09803
PaP	15	PHE	PRO	engineered mutation	UNP P09803
PaP	16	BAL	GLN	engineered mutation	UNP P09803

- Molecule 3 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	AAA	1	Total	C	H	N	O	S	1	0
			32	8	17	2	4	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	1	Total	Cl	0	0
			1	1		

- Molecule 5 is water.

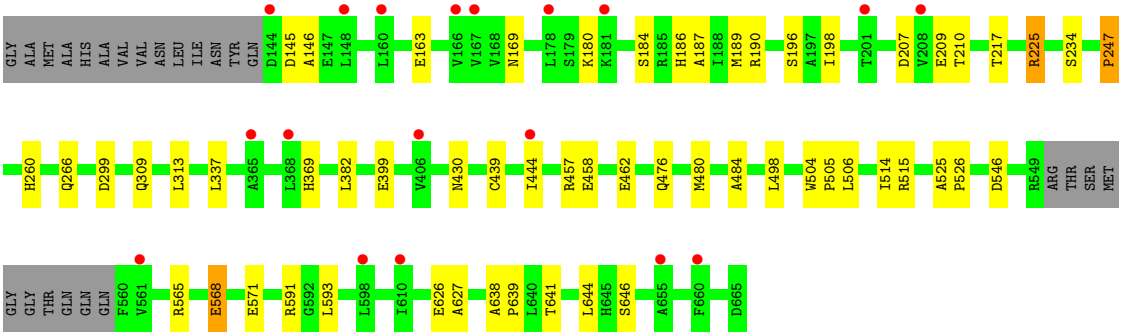
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	45	Total	O	0	0
			45	45		
5	PaP	1	Total	O	0	0
			1	1		

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.

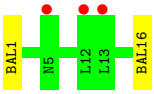
• Molecule 1: Catenin beta-1

Chain AAA:



• Molecule 2: Cadherin-1

Chain PaP:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	71.38Å 85.62Å 142.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	85.62 – 2.60 85.62 – 2.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (85.62-2.60) 100.0 (85.62-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.38 (at 2.62Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.203 , 0.242 0.212 , 0.255	Depositor DCC
R_{free} test set	1380 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	68.9	Xtriage
Anisotropy	0.481	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.43 , 57.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8287	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.82% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: DPR, EPE, BAL, CL

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	AAA	0.67	0/3959	0.78	1/5372 (0.0%)
2	PaP	0.66	0/103	0.89	0/138
All	All	0.67	0/4062	0.78	1/5510 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	AAA	247	PRO	O-C-N	5.90	132.14	122.70

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	AAA	3902	4065	4045	30	1
2	PaP	120	121	120	0	0
3	AAA	15	17	18	0	0
4	AAA	1	0	0	0	0
5	AAA	45	0	0	0	0
5	PaP	1	0	0	0	0
All	All	4084	4203	4183	30	1

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 4.

All (30) 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:AAA:638:ALA:HB3	1:AAA:639:PRO:CD	2.34	0.58
1:AAA:169:ASN:HD22	1:AAA:210:THR:CG2	2.21	0.54
1:AAA:444:ILE:HG21	1:AAA:484:ALA:HB3	1.90	0.54
1:AAA:515:ARG:C	1:AAA:515:ARG:HD3	2.28	0.54
1:AAA:591:ARG:HD3	1:AAA:627:ALA:HB2	1.91	0.53
1:AAA:591:ARG:NH1	1:AAA:626:GLU:OE1	2.42	0.53
1:AAA:525:ALA:HB3	1:AAA:526:PRO:HD3	1.91	0.52
1:AAA:641:THR:O	1:AAA:644:LEU:HB2	2.10	0.51
1:AAA:169:ASN:HD21	1:AAA:209:GLU:HG2	1.75	0.51
1:AAA:260:HIS:HE1	1:AAA:299:ASP:OD2	1.93	0.51
1:AAA:207:ASP:OD2	1:AAA:210:THR:HG23	2.12	0.50
1:AAA:169:ASN:ND2	1:AAA:207:ASP:OD2	2.45	0.50
1:AAA:638:ALA:HB3	1:AAA:639:PRO:HD3	1.94	0.50
1:AAA:498:LEU:HD11	1:AAA:514:ILE:CD1	2.45	0.47
1:AAA:163:GLU:HA	1:AAA:163:GLU:OE1	2.15	0.46
1:AAA:591:ARG:NH1	1:AAA:626:GLU:HB3	2.30	0.46
1:AAA:169:ASN:HA	1:AAA:210:THR:CG2	2.47	0.45
1:AAA:337:LEU:HD13	1:AAA:369:HIS:CE1	2.51	0.45
1:AAA:565:ARG:O	1:AAA:568:GLU:HB2	2.17	0.45
1:AAA:184:SER:O	1:AAA:187:ALA:HB3	2.17	0.45
1:AAA:439:CYS:HB3	1:AAA:480:MET:HE1	1.99	0.44
1:AAA:458:GLU:HG2	1:AAA:506:LEU:HD22	2.00	0.43
1:AAA:198:ILE:HD13	1:AAA:217:THR:HG21	1.99	0.43
1:AAA:439:CYS:HB3	1:AAA:480:MET:CE	2.49	0.43
1:AAA:169:ASN:HD22	1:AAA:210:THR:HG23	1.82	0.43
1:AAA:382:LEU:HD23	1:AAA:382:LEU:HA	1.91	0.43
1:AAA:504:TRP:N	1:AAA:505:PRO:CD	2.82	0.42
1:AAA:309:GLN:O	1:AAA:313:LEU:HG	2.20	0.42
1:AAA:225:ARG:NH2	1:AAA:266:GLN:OE1	2.52	0.41
1:AAA:186:HIS:HA	1:AAA:189:MET:HE2	2.03	0.41

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	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:190:ARG:HE	1:AAA:399:GLU:OE2[4_445]	1.52	0.08

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	AAA	509/536 (95%)	491 (96%)	15 (3%)	3 (1%)	25 47
2	PaP	13/16 (81%)	12 (92%)	1 (8%)	0	100 100
All	All	522/552 (95%)	503 (96%)	16 (3%)	3 (1%)	25 47

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	146	ALA
1	AAA	457	ARG
1	AAA	430	ASN

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	AAA	424/442 (96%)	411 (97%)	13 (3%)	40 66
2	PaP	13/13 (100%)	13 (100%)	0	100 100
All	All	437/455 (96%)	424 (97%)	13 (3%)	41 67

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	145	ASP
1	AAA	180	LYS
1	AAA	196	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	AAA	225	ARG
1	AAA	234	SER
1	AAA	247	PRO
1	AAA	462	GLU
1	AAA	476	GLN
1	AAA	546	ASP
1	AAA	568	GLU
1	AAA	571	GLU
1	AAA	593	LEU
1	AAA	646	SER

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)

3 non-standard protein/DNA/RNA residues 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	BAL	PaP	1	2	4,4,5	0.45	0	3,3,5	1.44	1 (33%)
2	BAL	PaP	16	2	4,4,5	0.45	0	3,3,5	2.46	1 (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
2	BAL	PaP	1	2	-	0/1/2/3	-
2	BAL	PaP	16	2	-	1/1/2/3	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	PaP	16	BAL	CB-CA-C	4.09	117.50	111.42
2	PaP	1	BAL	CB-CA-C	2.30	114.84	111.42

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	PaP	16	BAL	C-CA-CB-N

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	$\#\|Z\| > 2$
3	EPE	AAA	701	-	15,15,15	1.91	1 (6%)	18,20,20	1.31	3 (16%)

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	EPE	AAA	701	-	-	3/9/19/19	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	AAA	701	EPE	C10-S	-7.10	1.67	1.77

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	AAA	701	EPE	O2S-S-C10	2.79	110.28	106.92
3	AAA	701	EPE	O1S-S-C10	2.38	109.78	106.92
3	AAA	701	EPE	O3S-S-C10	2.38	109.61	105.77

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	AAA	701	EPE	C8-C7-N4-C5
3	AAA	701	EPE	C10-C9-N1-C6
3	AAA	701	EPE	C10-C9-N1-C2

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

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	AAA	512/536 (95%)	0.85	18 (3%)	44	36	60, 75, 102, 145	0
2	PaP	13/16 (81%)	1.20	3 (23%)	0	0	87, 94, 101, 106	0
All	All	525/552 (95%)	0.86	21 (4%)	38	31	60, 75, 102, 145	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	561	VAL	6.4
1	AAA	144	ASP	3.0
1	AAA	148	LEU	2.8
1	AAA	201	THR	2.7
1	AAA	160	LEU	2.6
2	PaP	13	LEU	2.4
2	PaP	12	LEU	2.4
1	AAA	598	LEU	2.3
2	PaP	5	ASN	2.3
1	AAA	181	LYS	2.3
1	AAA	208	VAL	2.3
1	AAA	178	LEU	2.3
1	AAA	444	ILE	2.3
1	AAA	610	ILE	2.3
1	AAA	166	VAL	2.3
1	AAA	368	LEU	2.3
1	AAA	655	ALA	2.2
1	AAA	365	ALA	2.1
1	AAA	406	VAL	2.1
1	AAA	167	VAL	2.1
1	AAA	660	PHE	2.1

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, 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^2)	Q<0.9
2	BAL	PaP	16	5/6	0.86	0.55	83,90,97,104	0
2	BAL	PaP	1	5/6	0.94	0.31	90,94,96,97	0
2	DPR	PaP	8	7/8	0.96	0.21	97,100,105,105	0

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^2)	Q<0.9
3	EPE	AAA	701	15/15	0.81	0.51	126,136,138,139	1
4	CL	AAA	702	1/1	0.87	0.31	81,81,81,81	0

6.5 Other polymers (i)

There are no such residues in this entry.