



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 1, 2017 – 07:35 AM EST

PDB ID : 5KBQ
Title : Pak1 in complex with bis-anilino pyrimidine inhibitor
Authors : Ferguson, A.
Deposited on : 2016-06-03
Resolution : 2.58 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

<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 : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20028442
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) : rb-20028442

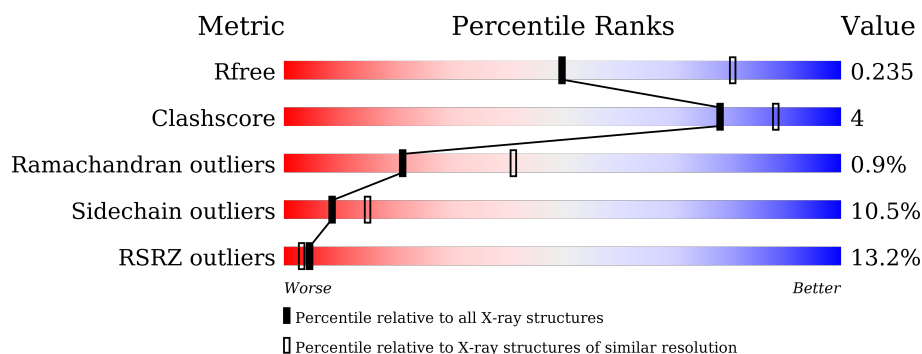
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.58 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	2636 (2.60-2.56)
Clashscore	102246	3003 (2.60-2.56)
Ramachandran outliers	100387	2956 (2.60-2.56)
Sidechain outliers	100360	2956 (2.60-2.56)
RSRZ outliers	91569	2642 (2.60-2.56)

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	289	<div> <div>18%</div> <div> <div></div> <div>78%</div> <div>19%</div> <div>.</div> </div> </div>
1	B	289	<div> <div>8%</div> <div> <div></div> <div>79%</div> <div>15%</div> <div>6%</div> </div> </div>

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4483 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.

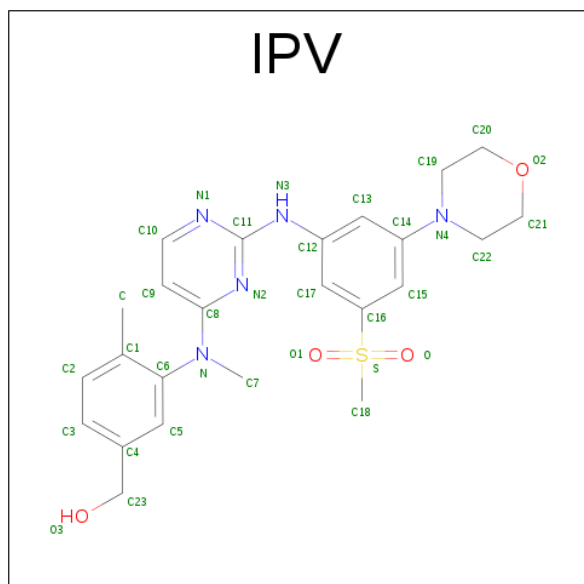
- Molecule 1 is a protein called Serine/threonine-protein kinase PAK 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	289	Total	C	N	O	S	0	0	0
			2257	1433	379	429	16			
1	B	273	Total	C	N	O	S	0	0	0
			2143	1360	361	407	15			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	389	ASN	ASP	engineered mutation	UNP Q13153
A	423	GLU	THR	engineered mutation	UNP Q13153
A	503	ASP	GLU	engineered mutation	UNP Q13153
B	389	ASN	ASP	engineered mutation	UNP Q13153
B	423	GLU	THR	engineered mutation	UNP Q13153
B	503	ASP	GLU	engineered mutation	UNP Q13153

- Molecule 2 is [4-methyl-3-[methyl-[2-[(3-methylsulfonyl-5-morpholin-4-yl-phenyl)amino]pyrimidin-4-yl]amino]phenyl]methanol (three-letter code: IPV) (formula: C₂₄H₂₉N₅O₄S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			34	24	5	4	1		

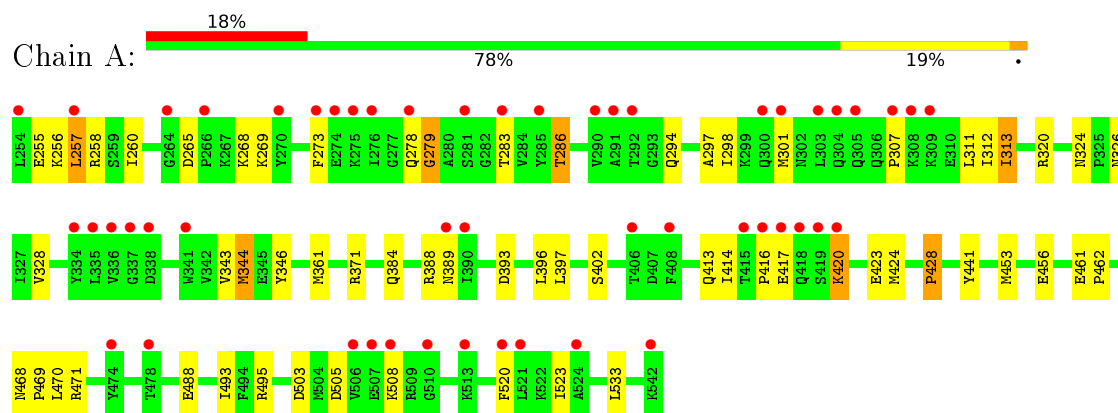
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	17	Total	O	0	0
			17	17		
3	B	32	Total	O	0	0
			32	32		

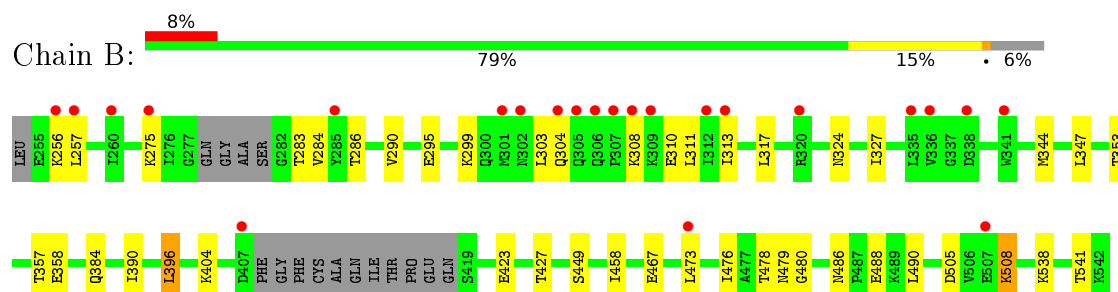
3 Residue-property plots [i](#)

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: Serine/threonine-protein kinase PAK 1



- Molecule 1: Serine/threonine-protein kinase PAK 1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	61.91Å 80.94Å 66.02Å 90.00° 107.07° 90.00°	Depositor
Resolution (Å)	34.37 – 2.58 34.07 – 2.58	Depositor EDS
% Data completeness (in resolution range)	99.2 (34.37-2.58) 99.5 (34.07-2.58)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.78 (at 2.57Å)	Xtriage
Refinement program	BUSTER 2.11.7	Depositor
R, R_{free}	0.175 , 0.230 0.182 , 0.235	Depositor DCC
R_{free} test set	1004 reflections (5.41%)	DCC
Wilson B-factor (Å ²)	65.3	Xtriage
Anisotropy	0.255	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 75.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	4483	wwPDB-VP
Average B, all atoms (Å ²)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.64% 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: IPV

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	0.52	0/2295	0.72	0/3104
1	B	0.51	0/2176	0.71	0/2939
All	All	0.51	0/4471	0.72	0/6043

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	2257	0	2298	21	0
1	B	2143	0	2192	11	0
2	A	34	0	0	0	0
3	A	17	0	0	0	0
3	B	32	0	0	0	0
All	All	4483	0	4490	32	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 4.

All (32) 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:B:347:LEU:HD13	1:B:396:LEU:HB3	1.78	0.66
1:A:324:ASN:HD22	1:A:326:ASN:H	1.46	0.62
1:B:324:ASN:HB3	1:B:327:ILE:HG12	1.81	0.62
1:A:298:ILE:HG12	1:A:343:VAL:HG22	1.81	0.60
1:A:328:VAL:HG12	1:A:344:MET:HG2	1.88	0.55
1:B:476:ILE:O	1:B:480:GLY:HA2	2.08	0.54
1:A:388:ARG:HH12	1:A:423:GLU:HG2	1.73	0.54
1:A:328:VAL:HG21	1:A:396:LEU:HD12	1.92	0.52
1:A:265:ASP:HB3	1:A:268:LYS:HB2	1.91	0.51
1:A:453:MET:O	1:A:456:GLU:HB2	2.11	0.51
1:B:347:LEU:HD11	1:B:404:LYS:HD2	1.92	0.50
1:A:420:LYS:HB3	1:A:441:TYR:O	2.10	0.50
1:A:256:LYS:O	1:A:260:ILE:HG12	2.12	0.50
1:B:284:VAL:HG22	1:B:299:LYS:HG3	1.95	0.49
1:B:353:THR:O	1:B:357:THR:HG23	2.14	0.48
1:A:493:ILE:HD12	1:A:520:PHE:HA	1.96	0.47
1:B:358:GLU:HG3	1:B:541:THR:HG23	1.96	0.46
1:B:458:ILE:HG12	1:B:490:LEU:HD22	1.97	0.46
1:B:257:LEU:HB3	1:B:313:ILE:HD12	1.97	0.46
1:A:388:ARG:HH12	1:A:423:GLU:CG	2.29	0.46
1:A:397:LEU:HA	1:A:402:SER:O	2.16	0.45
1:A:505:ASP:HB3	1:A:508:LYS:HB2	1.98	0.44
1:B:505:ASP:HB3	1:B:508:LYS:HB2	1.99	0.44
1:A:273:PHE:HA	1:A:286:THR:O	2.18	0.43
1:A:428:PRO:HG2	1:A:469:PRO:HB3	1.99	0.43
1:A:297:ALA:HB2	1:A:346:TYR:HD1	1.83	0.43
1:A:461:GLU:HB2	1:A:462:PRO:HD2	2.02	0.41
1:A:384:GLN:HE22	1:A:416:PRO:HD3	1.86	0.41
1:A:278:GLN:HB2	1:A:279:GLY:HA3	2.04	0.40
1:A:361:MET:HG2	1:A:533:LEU:HD13	2.03	0.40
1:A:257:LEU:HG	1:A:313:ILE:HG22	2.03	0.40
1:B:390:ILE:HB	1:B:449:SER:HB2	2.04	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	287/289 (99%)	270 (94%)	14 (5%)	3 (1%)	19	38
1	B	267/289 (92%)	253 (95%)	12 (4%)	2 (1%)	26	49
All	All	554/578 (96%)	523 (94%)	26 (5%)	5 (1%)	21	41

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	279	GLY
1	A	307	PRO
1	A	424	MET
1	B	303	LEU
1	B	283	THR

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	249/250 (100%)	221 (89%)	28 (11%)	7	12
1	B	238/250 (95%)	215 (90%)	23 (10%)	10	18
All	All	487/500 (97%)	436 (90%)	51 (10%)	8	15

All (51) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	255	GLU
1	A	257	LEU
1	A	258	ARG
1	A	269	LYS
1	A	283	THR
1	A	286	THR
1	A	294	GLN
1	A	301	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	311	LEU
1	A	312	ILE
1	A	313	ILE
1	A	320	ARG
1	A	344	MET
1	A	371	ARG
1	A	389	ASN
1	A	393	ASP
1	A	413	GLN
1	A	414	ILE
1	A	417	GLU
1	A	420	LYS
1	A	428	PRO
1	A	468	ASN
1	A	470	LEU
1	A	471	ARG
1	A	488	GLU
1	A	495	ARG
1	A	503	ASP
1	A	523	ILE
1	B	256	LYS
1	B	275	LYS
1	B	286	THR
1	B	290	VAL
1	B	295	GLU
1	B	304	GLN
1	B	308	LYS
1	B	310	GLU
1	B	311	LEU
1	B	317	LEU
1	B	344	MET
1	B	384	GLN
1	B	396	LEU
1	B	423	GLU
1	B	427	THR
1	B	467	GLU
1	B	473	LEU
1	B	478	THR
1	B	479	ASN
1	B	486	ASN
1	B	488	GLU
1	B	508	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	538	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	324	ASN
1	A	384	GLN
1	A	389	ASN
1	A	413	GLN
1	B	375	GLN

5.3.3 RNA [i](#)

There are no RNA molecules 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](#)

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	$\# Z > 2$
2	IPV	A	601	-	37,37,37	0.45	1 (2%)	49,53,53	0.69	1 (2%)

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	IPV	A	601	-	-	0/24/32/32	0/4/4/4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	IPV	C6-N	2.28	1.47	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	601	IPV	C12-N3-C11	2.69	136.12	129.34

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 ⓘ

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	289/289 (100%)	0.94	51 (17%) 2 1	47, 84, 144, 166	0
1	B	273/289 (94%)	0.48	23 (8%) 14 10	49, 76, 130, 160	0
All	All	562/578 (97%)	0.71	74 (13%) 4 3	47, 79, 142, 166	0

All (74) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	291	ALA	9.6
1	A	336	VAL	7.3
1	A	419	SER	6.3
1	B	338	ASP	6.1
1	A	276	ILE	5.8
1	B	305	GLN	5.6
1	B	257	LEU	5.0
1	A	474	TYR	4.8
1	A	335	LEU	4.8
1	A	266	PRO	4.8
1	B	309	LYS	4.6
1	A	420	LYS	4.5
1	B	313	ILE	4.5
1	B	306	GLN	4.4
1	A	281	SER	4.3
1	A	416	PRO	4.2
1	A	417	GLU	4.0
1	A	305	GLN	4.0
1	B	304	GLN	3.9
1	A	264	GLY	3.8
1	A	341	TRP	3.8
1	B	312	ILE	3.8
1	B	336	VAL	3.8
1	B	307	PRO	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	418	GLN	3.6
1	A	292	THR	3.6
1	B	308	LYS	3.5
1	A	415	THR	3.4
1	B	407	ASP	3.3
1	A	309	LYS	3.3
1	A	254	LEU	3.3
1	A	303	LEU	3.2
1	A	283	THR	3.2
1	B	256	LYS	3.2
1	A	338	ASP	3.2
1	A	275	LYS	3.2
1	A	390	ILE	3.1
1	B	260	ILE	3.0
1	A	304	GLN	3.0
1	A	278	GLN	2.9
1	B	285	TYR	2.9
1	B	320	ARG	2.9
1	A	308	LYS	2.8
1	A	520	PHE	2.8
1	A	274	GLU	2.8
1	A	257	LEU	2.7
1	B	473	LEU	2.7
1	A	389	ASN	2.7
1	A	478	THR	2.6
1	A	542	LYS	2.5
1	B	341	TRP	2.5
1	A	285	TYR	2.5
1	A	301	MET	2.5
1	A	506	VAL	2.4
1	A	270	TYR	2.4
1	A	521	LEU	2.4
1	B	302	ASN	2.3
1	B	507	GLU	2.3
1	B	275	LYS	2.3
1	A	300	GLN	2.3
1	A	513	LYS	2.3
1	A	273	PHE	2.2
1	A	307	PRO	2.2
1	B	335	LEU	2.2
1	A	334	TYR	2.2
1	A	337	GLY	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	290	VAL	2.1
1	A	408	PHE	2.1
1	B	301	MET	2.1
1	A	406	THR	2.1
1	A	507	GLU	2.1
1	A	508	LYS	2.1
1	A	524	ALA	2.1
1	A	510	GLY	2.0

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(Å ²)	Q<0.9
2	IPV	A	601	34/34	0.88	0.22	-0.03	68,77,105,107	0

6.5 Other polymers [i](#)

There are no such residues in this entry.