Full wwPDB X-ray Structure Validation Report (i)

Aug 16, 2016 – 08:43 AM EDT

PDB ID	:	5KX7
Title	:	Irak4-inhibitor co-structure
Authors	:	Fischmann, T.O.
Deposited on	:	2016-07-20
Resolution	:	2.80 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

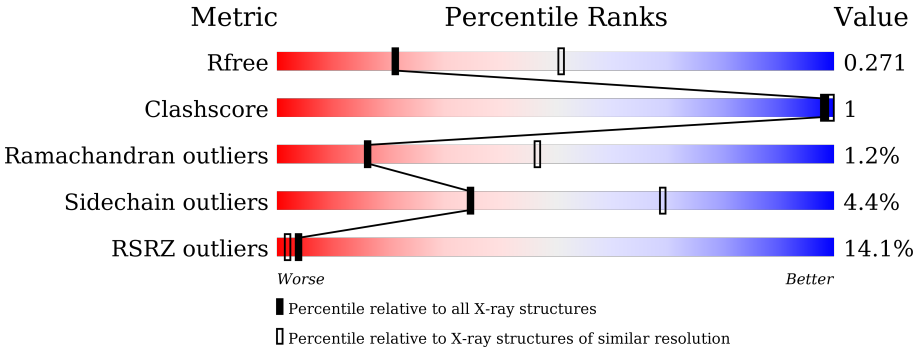
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

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 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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 <=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	301	15% 73% 7% 20%
1	B	301	9% 90% 6% •

2 Entry composition (i)

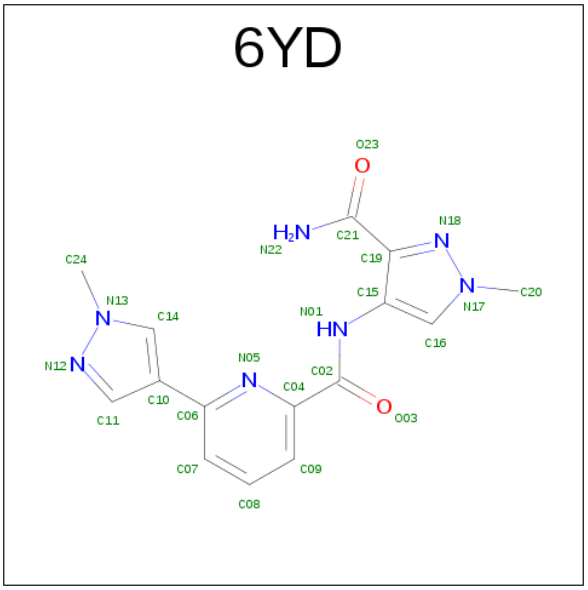
There are 2 unique types of molecules in this entry. The entry contains 4159 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Interleukin-1 receptor-associated kinase 4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	240	1868	1166	313	372	3	14	0	0	2
1	B	289	2267	1418	383	449	3	14	0	0	3

- Molecule 2 is {N}-(3-aminocarbonyl-1-methyl-pyrazol-4-yl)-6-(1-methylpyrazol-4-yl)pyridine-2-carboxamide (three-letter code: 6YD) (formula: $C_{15}H_{15}N_7O_2$).



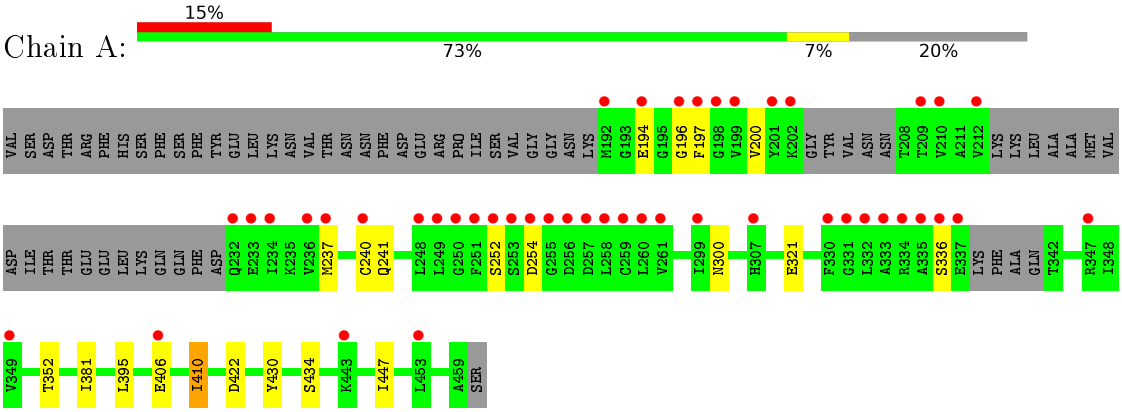
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	B	1	24	15	7	2	0	0

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.

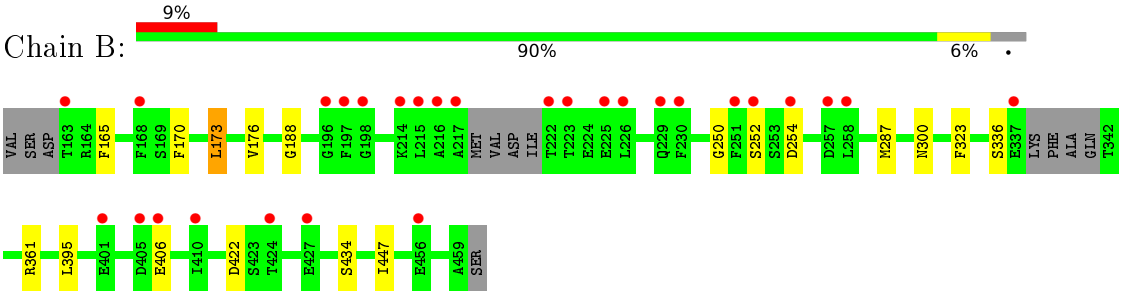
- Molecule 1: Interleukin-1 receptor-associated kinase 4

Chain A:



- Molecule 1: Interleukin-1 receptor-associated kinase 4

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	87.80Å 114.28Å 139.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.32 – 2.80 39.32 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.7 (39.32-2.80) 99.7 (39.32-2.80)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.48 (at 2.81Å)	Xtriage
Refinement program	BUSTER 2.11.6	Depositor
R, R_{free}	0.211 , 0.250 0.222 , 0.271	Depositor DCC
R_{free} test set	871 reflections (5.21%)	DCC
Wilson B-factor (Å^2)	69.4	Xtriage
Anisotropy	0.839	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 81.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4159	wwPDB-VP
Average B, all atoms (Å^2)	93.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 33.62 % of the origin peak, indicating pseudo translational symmetry.The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.7717e-04.The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: TPO, 6YD, SEP

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	0.53	0/1862	0.64	0/2508
1	B	0.53	0/2271	0.64	0/3060
All	All	0.53	0/4133	0.64	0/5568

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 (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 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	1868	0	1843	4	0
1	B	2267	0	2225	5	0
2	B	24	0	0	0	0
All	All	4159	0	4068	8	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 1.

All (8) 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:410:ILE:HD12	1:A:430:TYR:HB2	1.90	0.53
1:A:352:THR:CG2	1:B:361:ARG:HA	2.42	0.49
1:B:165:PHE:HB3	1:B:250:GLY:HA2	1.99	0.44
1:B:173:LEU:HA	1:B:176:VAL:HG22	1.99	0.44
1:A:381:ILE:HG21	1:A:410:ILE:HD11	2.02	0.41
1:A:300:ASN:HA	1:A:447:ILE:HG21	2.01	0.41
1:B:300:ASN:HA	1:B:447:ILE:HG21	2.01	0.41
1:B:287:MET:HE1	1:B:323:PHE: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	230/301 (76%)	212 (92%)	15 (6%)	3 (1%)	15 44
1	B	281/301 (93%)	264 (94%)	14 (5%)	3 (1%)	17 50
All	All	511/602 (85%)	476 (93%)	29 (6%)	6 (1%)	16 47

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	406	GLU
1	B	406	GLU
1	B	170	PHE
1	B	188	GLY
1	A	321	GLU
1	A	196	GLY

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	203/259 (78%)	190 (94%)	13 (6%)	22	52
1	B	247/259 (95%)	240 (97%)	7 (3%)	51	84
All	All	450/518 (87%)	430 (96%)	20 (4%)	35	69

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	194	GLU
1	A	197	PHE
1	A	200	VAL
1	A	237	MET
1	A	240	CYS
1	A	241	GLN
1	A	252	SER
1	A	254	ASP
1	A	336	SER
1	A	395	LEU
1	A	410	ILE
1	A	422	ASP
1	A	434	SER
1	B	173	LEU
1	B	252	SER
1	B	254	ASP
1	B	336	SER
1	B	395	LEU
1	B	422	ASP
1	B	434	SER

Some 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)

6 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	$\#\|Z\| > 2$
1	TPO	A	342	1	7,10,11	0.86	0	10,14,16	1.39	2 (20%)
1	TPO	A	345	1	7,10,11	1.48	2 (28%)	10,14,16	1.44	1 (10%)
1	SEP	A	346	1	7,9,10	0.92	0	8,12,14	1.53	2 (25%)
1	TPO	B	342	1	7,10,11	0.92	0	10,14,16	1.58	2 (20%)
1	TPO	B	345	1	7,10,11	1.49	1 (14%)	10,14,16	1.38	1 (10%)
1	SEP	B	346	1	7,9,10	0.92	0	8,12,14	1.56	3 (37%)

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
1	TPO	A	342	1	-	0/8/11/13	0/0/0/0
1	TPO	A	345	1	-	1/8/11/13	0/0/0/0
1	SEP	A	346	1	-	0/5/8/10	0/0/0/0
1	TPO	B	342	1	-	0/8/11/13	0/0/0/0
1	TPO	B	345	1	-	1/8/11/13	0/0/0/0
1	SEP	B	346	1	-	0/5/8/10	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	345	TPO	CG2-CB	2.01	1.56	1.51
1	A	345	TPO	CB-CA	3.02	1.59	1.54
1	B	345	TPO	CB-CA	3.10	1.59	1.54

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	346	SEP	O-C-CA	-2.38	119.33	125.72
1	B	346	SEP	O-C-CA	-2.36	119.40	125.72
1	B	342	TPO	O-C-CA	-2.25	119.54	125.69
1	A	342	TPO	O-C-CA	-2.13	119.86	125.69
1	B	346	SEP	OG-P-O1P	2.26	112.77	107.08
1	A	346	SEP	OG-CB-CA	2.51	110.45	108.26
1	B	346	SEP	OG-CB-CA	2.66	110.58	108.26
1	A	342	TPO	O2P-P-OG1	3.36	116.68	106.62
1	B	345	TPO	O2P-P-OG1	3.60	117.38	106.62
1	A	345	TPO	O2P-P-OG1	3.63	117.47	106.62
1	B	342	TPO	OG1-P-O1P	4.09	117.24	107.48

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	345	TPO	OG1-CB-CA-N
1	B	345	TPO	OG1-CB-CA-N

There are no ring outliers.

No monomer is involved in short contacts.

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	6YD	B	9901	-	23,26,26	0.95	1 (4%)	24,37,37	0.80	1 (4%)

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	6YD	B	9901	-	-	0/6/16/16	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	9901	6YD	C16-C15	2.78	1.42	1.37

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	9901	6YD	C14-C10-C06	-2.21	125.55	128.11

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 (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	237/301 (78%)	1.18	46 (19%)	1	1	54, 91, 169, 182	0
1	B	286/301 (95%)	0.58	28 (9%)	10	5	50, 83, 123, 139	0
All	All	523/602 (86%)	0.86	74 (14%)	4	2	50, 86, 158, 182	0

All (74) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	259	CYS	10.3
1	A	234	ILE	9.7
1	A	252	SER	8.9
1	A	253	SER	8.8
1	A	197	PHE	8.2
1	A	333	ALA	7.9
1	A	251	PHE	7.6
1	A	233	GLU	7.4
1	A	258	LEU	7.0
1	A	236	VAL	6.8
1	B	163	THR	6.5
1	A	332	LEU	6.5
1	A	334	ARG	6.1
1	A	198	GLY	6.0
1	A	196	GLY	6.0
1	A	336	SER	5.5
1	A	249	LEU	5.5
1	A	256	ASP	5.3
1	A	260	LEU	5.2
1	A	212	VAL	4.9
1	A	248	LEU	4.9
1	A	237	MET	4.7
1	A	201	TYR	4.6
1	A	255	GLY	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	226	LEU	4.4
1	A	261	VAL	4.4
1	B	216	ALA	4.4
1	A	250	GLY	4.3
1	B	223	THR	4.3
1	A	331	GLY	4.1
1	A	257	ASP	4.1
1	B	225	GLU	4.0
1	A	335	ALA	3.9
1	A	349	VAL	3.9
1	A	330	PHE	3.9
1	A	254	ASP	3.8
1	B	215	LEU	3.8
1	A	347	ARG	3.7
1	A	202	LYS	3.7
1	B	230	PHE	3.7
1	A	337	GLU	3.7
1	B	197	PHE	3.4
1	B	254	ASP	3.4
1	A	210	VAL	3.3
1	B	217	ALA	3.3
1	B	258	LEU	3.3
1	A	194	GLU	3.2
1	A	232	GLN	3.1
1	B	257	ASP	3.0
1	A	240	CYS	3.0
1	B	222	THR	2.8
1	A	209	THR	2.7
1	B	196	GLY	2.7
1	A	199	VAL	2.6
1	B	406	GLU	2.6
1	B	424	THR	2.5
1	B	214	LYS	2.5
1	A	406	GLU	2.4
1	B	198	GLY	2.4
1	B	168	PHE	2.4
1	A	443	LYS	2.3
1	B	229	GLN	2.3
1	A	299	ILE	2.2
1	B	401	GLU	2.2
1	B	252	SER	2.2
1	B	405	ASP	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	427	GLU	2.1
1	B	456	GLU	2.1
1	B	251	PHE	2.1
1	A	307	HIS	2.0
1	A	453	LEU	2.0
1	B	410	ILE	2.0
1	B	337	GLU	2.0
1	A	192	MET	2.0

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. 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
1	TPO	A	345	11/12	0.94	0.20	-	140,143,145,146	0
1	SEP	B	346	10/11	0.75	0.17	-	115,120,128,128	0
1	SEP	A	346	10/11	0.66	0.26	-	144,146,152,152	0
1	TPO	B	342	11/12	0.82	0.22	-	125,131,138,139	0
1	TPO	B	345	11/12	0.94	0.12	-	110,115,117,117	0
1	TPO	A	342	11/12	0.71	0.35	-	145,147,151,152	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. 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	6YD	B	9901	24/24	0.95	0.20	-0.38	47,64,79,80	0

6.5 Other polymers (i)

There are no such residues in this entry.