



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 08:26 am GMT

PDB ID : 4R5G  
Title : Crystal structure of the DnaK C-terminus with the inhibitor PET-16  
Authors : Leu, J.I.; Zhang, P.; Murphy, M.E.; Marmorstein, R.; George, D.L.  
Deposited on : 2014-08-21  
Resolution : 3.45 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](mailto: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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

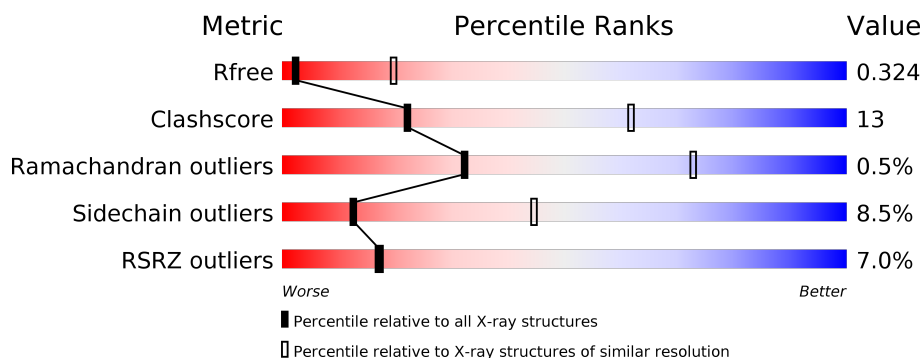
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.45 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))
$R_{free}$	100719	1135 (3.56-3.36)
Clashscore	112137	1040 (3.52-3.40)
Ramachandran outliers	110173	1009 (3.52-3.40)
Sidechain outliers	110143	1010 (3.52-3.40)
RSRZ outliers	101464	1017 (3.54-3.38)

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  $\geq 3$ , 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	230	<div> <div>9%</div> <div>72%</div> <div>20%</div> <div>• 5%</div> </div>
1	B	230	<div> <div>4%</div> <div>69%</div> <div>18%</div> <div>• 9%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	3JE	B	701	-	-	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2834 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.

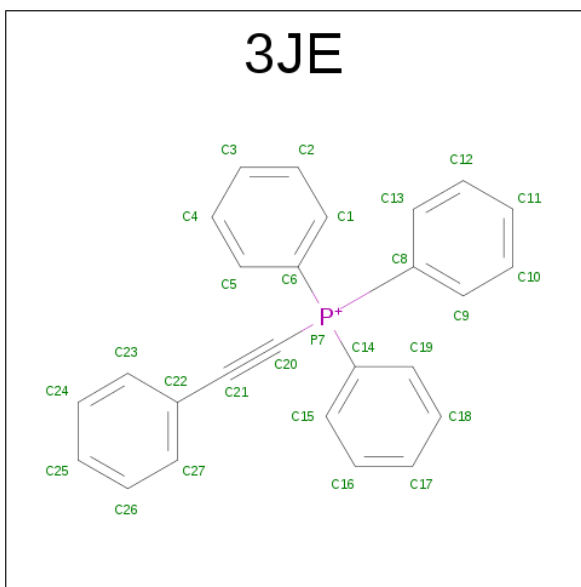
- Molecule 1 is a protein called Chaperone protein DnaK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	218	Total	C	N	O	S	0	0	0
			1413	867	250	292	4			
1	B	209	Total	C	N	O	S	3	0	0
			1393	859	247	284	3			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	378	MET	-	EXPRESSION TAG	UNP P0A6Y8
A	379	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	380	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	381	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	382	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	383	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	384	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	385	ILE	-	EXPRESSION TAG	UNP P0A6Y8
A	386	GLU	-	EXPRESSION TAG	UNP P0A6Y8
A	387	GLY	-	EXPRESSION TAG	UNP P0A6Y8
A	388	ARG	-	EXPRESSION TAG	UNP P0A6Y8
B	378	MET	-	EXPRESSION TAG	UNP P0A6Y8
B	379	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	380	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	381	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	382	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	383	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	384	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	385	ILE	-	EXPRESSION TAG	UNP P0A6Y8
B	386	GLU	-	EXPRESSION TAG	UNP P0A6Y8
B	387	GLY	-	EXPRESSION TAG	UNP P0A6Y8
B	388	ARG	-	EXPRESSION TAG	UNP P0A6Y8

- Molecule 2 is TRIPHENYL(PHENYLETHYNYL)PHOSPHONIUM (three-letter code: 3JE) (formula: C<sub>26</sub>H<sub>20</sub>P).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	C	P	0
			27	26	1	

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	O	0	0
			1	1		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.78Å 91.78Å 136.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.51 – 3.45 43.51 – 3.44	Depositor EDS
% Data completeness (in resolution range)	99.6 (43.51-3.45) 99.8 (43.51-3.44)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.38 (at 3.40Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.284 , 0.328 0.282 , 0.324	Depositor DCC
$R_{free}$ test set	394 reflections (4.84%)	DCC
Wilson B-factor (Å <sup>2</sup> )	100.3	Xtriage
Anisotropy	0.621	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 68.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	2834	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.19% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: 3JE

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	0.26	0/1426	0.52	0/1955
1	B	0.28	0/1406	0.56	2/1924 (0.1%)
All	All	0.27	0/2832	0.54	2/3879 (0.1%)

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	A	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	404	MET	CB-CG-SD	-5.64	95.47	112.40
1	B	484	LEU	CA-CB-CG	5.08	126.99	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	469	MET	Peptide

### 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	1413	0	1193	36	0
1	B	1393	0	1213	37	0
2	B	27	0	20	1	0
3	A	1	0	0	0	0
All	All	2834	0	2426	69	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 13.

All (69) 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:403:THR:O	1:B:404:MET:O	1.52	1.24
1:B:404:MET:O	1:B:406:GLY:N	1.91	1.03
1:A:483:ILE:HD11	1:A:500:THR:HG22	1.65	0.77
1:B:404:MET:C	1:B:406:GLY:H	1.87	0.77
1:B:445:ARG:HE	1:B:519:ALA:HA	1.50	0.77
1:B:393:ASP:HB2	1:B:418:ILE:HD11	1.71	0.72
1:A:502:LYS:NZ	1:B:500:THR:OG1	2.26	0.69
1:A:398:SER:HB2	1:A:444:GLU:HB3	1.74	0.68
1:A:392:LEU:HD21	1:A:480:ALA:HA	1.75	0.68
1:B:481:ASP:HB2	1:B:483:ILE:HG22	1.77	0.67
1:B:505:SER:N	1:B:506:GLY:HA2	2.11	0.65
1:A:426:PHE:HB2	1:A:472:ILE:HD11	1.77	0.65
1:B:378:MET:HB3	1:B:382:HIS:HB3	1.79	0.64
1:A:427:SER:OG	1:A:428:THR:N	2.29	0.63
1:B:403:THR:O	1:B:404:MET:C	2.29	0.63
1:B:417:THR:O	1:B:420:THR:OG1	2.16	0.63
1:A:405:GLY:HA3	1:A:533:VAL:HG13	1.82	0.62
1:B:475:THR:HB	1:B:487:SER:HB3	1.82	0.62
1:A:591:LEU:O	1:A:595:SER:OG	2.19	0.60
1:A:397:LEU:HD13	1:A:443:GLY:HA2	1.85	0.59
1:B:403:THR:HG1	1:B:409:THR:HG1	1.48	0.58
1:A:442:GLN:HB2	1:A:454:LEU:HD11	1.86	0.57
1:B:408:MET:SD	1:B:451:ASN:ND2	2.76	0.57
1:A:479:ASP:N	1:A:479:ASP:OD1	2.25	0.57
1:B:547:ARG:HH21	1:B:569:LEU:HG	1.69	0.57
1:B:402:GLU:HB2	1:B:441:LEU:HD11	1.87	0.56
1:B:399:LEU:HD11	2:B:701:3JE:H3	1.86	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:442:GLN:O	1:A:451:ASN:HB3	2.08	0.54
1:A:443:GLY:O	1:A:451:ASN:ND2	2.41	0.53
1:B:382:HIS:CG	1:B:383:HIS:H	2.26	0.53
1:B:547:ARG:NH2	1:B:569:LEU:HG	2.23	0.53
1:B:382:HIS:CD2	1:B:383:HIS:H	2.27	0.53
1:A:561:ASP:O	1:A:565:ILE:HG13	2.10	0.52
1:B:532:LEU:O	1:B:535:THR:OG1	2.20	0.51
1:A:484:LEU:HD23	1:A:485:HIS:N	2.26	0.50
1:B:445:ARG:NE	1:B:519:ALA:HA	2.24	0.50
1:A:499:ILE:HG22	1:B:501:ILE:HG22	1.95	0.49
1:B:445:ARG:HH21	1:B:518:ASP:C	2.16	0.49
1:A:500:THR:HB	1:B:500:THR:HB	1.95	0.48
1:A:581:LYS:O	1:A:585:GLU:HG2	2.14	0.47
1:A:411:LEU:HD11	1:A:426:PHE:HZ	1.80	0.47
1:A:586:ALA:O	1:A:590:GLU:HG2	2.15	0.46
1:A:384:HIS:ND1	1:A:384:HIS:N	2.62	0.46
1:B:560:ASP:N	1:B:561:ASP:HA	2.30	0.46
1:B:403:THR:C	1:B:404:MET:O	2.39	0.46
1:A:442:GLN:HG2	1:A:443:GLY:H	1.80	0.46
1:A:465:ALA:HB3	1:A:470:PRO:HG3	1.97	0.46
1:A:441:LEU:HD23	1:A:452:LYS:O	2.16	0.45
1:A:426:PHE:O	1:A:472:ILE:HG13	2.16	0.45
1:B:384:HIS:N	1:B:384:HIS:ND1	2.54	0.44
1:A:425:VAL:HG23	1:A:472:ILE:O	2.18	0.44
1:A:479:ASP:OD1	1:A:483:ILE:HG23	2.17	0.43
1:A:485:HIS:ND1	1:A:485:HIS:N	2.66	0.43
1:A:433:GLN:NE2	1:A:435:ALA:O	2.42	0.43
1:B:399:LEU:HD13	1:B:484:LEU:HD12	2.01	0.43
1:A:473:GLU:O	1:A:488:ALA:HA	2.18	0.43
1:A:411:LEU:HD11	1:A:426:PHE:CZ	2.54	0.42
1:A:456:GLN:HB3	1:B:456:GLN:HG2	2.01	0.42
1:B:591:LEU:HA	1:B:591:LEU:HD23	1.78	0.42
1:B:454:LEU:HD11	1:B:503:ALA:O	2.20	0.41
1:B:393:ASP:N	1:B:393:ASP:OD1	2.53	0.41
1:B:435:ALA:HA	1:B:460:ASP:HA	2.01	0.41
1:A:484:LEU:CD2	1:A:486:VAL:HG23	2.50	0.41
1:B:440:VAL:HG21	1:B:486:VAL:HG21	2.02	0.41
1:A:490:ASP:HB3	1:A:492:ASN:O	2.21	0.41
1:B:477:ASP:HB3	1:B:485:HIS:HB2	2.02	0.41
1:B:445:ARG:HE	1:B:519:ALA:CA	2.28	0.41
1:A:511:GLU:O	1:A:515:MET:HG3	2.22	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:442:GLN:HB2	1:A:454:LEU:CD1	2.49	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	214/230 (93%)	186 (87%)	28 (13%)	0	100	100
1	B	203/230 (88%)	178 (88%)	23 (11%)	2 (1%)	18	60
All	All	417/460 (91%)	364 (87%)	51 (12%)	2 (0%)	32	73

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	404	MET
1	B	405	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	114/191 (60%)	105 (92%)	9 (8%)	14	48
1	B	120/191 (63%)	109 (91%)	11 (9%)	11	40
All	All	234/382 (61%)	214 (92%)	20 (8%)	12	44

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

Mol	Chain	Res	Type
1	A	384	HIS
1	A	393	ASP
1	A	436	VAL
1	A	462	ILE
1	A	479	ASP
1	A	483	ILE
1	A	485	HIS
1	A	490	ASP
1	A	573	GLU
1	B	384	HIS
1	B	394	VAL
1	B	401	ILE
1	B	403	THR
1	B	404	MET
1	B	411	LEU
1	B	417	THR
1	B	437	THR
1	B	485	HIS
1	B	535	THR
1	B	569	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules 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 ⓘ

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	3JE	B	701	-	29,30,30	1.72	4 (13%)	39,40,40	0.84	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	3JE	B	701	-	-	0/20/24/24	0/4/4/4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	701	3JE	C22-C21	-6.62	1.28	1.44
2	B	701	3JE	P7-C6	-3.84	1.74	1.80
2	B	701	3JE	P7-C8	-3.63	1.74	1.80
2	B	701	3JE	P7-C14	-3.15	1.75	1.80

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	701	3JE	C14-P7-C6	2.71	112.76	108.15

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	701	3JE	1	0

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	218/230 (94%)	0.39	20 (9%) 10 11	8, 41, 106, 120	0
1	B	209/230 (90%)	0.32	10 (4%) 31 27	8, 46, 92, 107	1 (0%)
All	All	427/460 (92%)	0.36	30 (7%) 17 17	8, 43, 98, 120	1 (0%)

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	443	GLY	6.2
1	A	549	GLN	5.1
1	A	552	GLU	4.8
1	A	551	GLU	4.8
1	B	546	THR	4.2
1	A	555	ASP	4.1
1	B	444	GLU	3.9
1	B	574	THR	3.8
1	A	550	VAL	3.7
1	A	546	THR	3.6
1	B	576	LEU	3.5
1	B	379	HIS	3.4
1	B	577	LYS	3.3
1	A	431	ASP	3.1
1	A	563	THR	3.1
1	A	560	ASP	3.0
1	A	463	ASN	3.0
1	B	450	ASP	2.9
1	A	564	ALA	2.8
1	A	559	ALA	2.7
1	A	553	ALA	2.5
1	A	588	MET	2.5
1	A	388	ARG	2.4
1	A	378	MET	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	575	ALA	2.3
1	A	545	SER	2.2
1	A	529	PHE	2.2
1	B	567	SER	2.1
1	A	379	HIS	2.1
1	A	558	PRO	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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	3JE	B	701	27/27	0.73	0.40	2.39	30,54,70,85	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.