



# wwPDB X-ray Structure Validation Summary Report

Mar 1, 2014 – 01:48 AM GMT

PDB ID : 2I3Z  
Title : rat DPP-IV with xanthine mimetic inhibitor #7  
Authors : Kurukulasuriya, R.; Rohde, J.J.; Szczepankiewicz, B.G; Basha, F.; Lai, C.; Winn, M.; Stewart, K.D.; Longenecker, K.L.; Lubben, T.W.; Ballaron, S.J.; Sham, H.L.; VonGeldern, T.W.  
Deposited on : 2006-08-21  
Resolution : 2.90 Å(reported)

This is a wwPDB validation summary 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/ValidationPDFNotes.html>

---

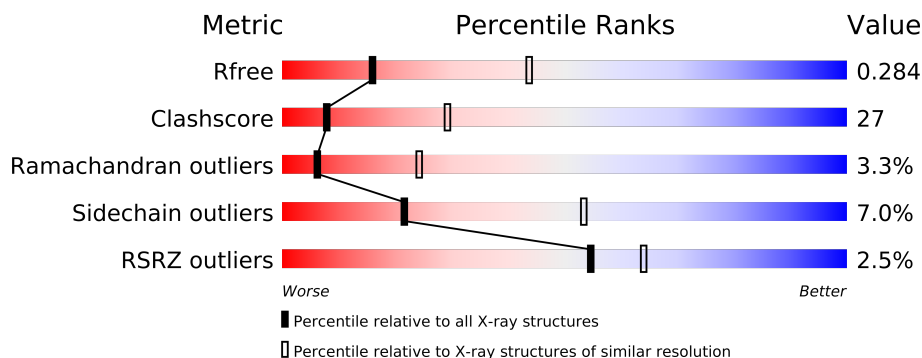
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.90 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	1053 (2.90-2.90)
Clashscore	79885	1326 (2.90-2.90)
Ramachandran outliers	78287	1290 (2.90-2.90)
Sidechain outliers	78261	1292 (2.90-2.90)
RSRZ outliers	66119	1054 (2.90-2.90)

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. 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	730	
1	B	730	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	LIR	A	900	-	X

## 2 Entry composition i

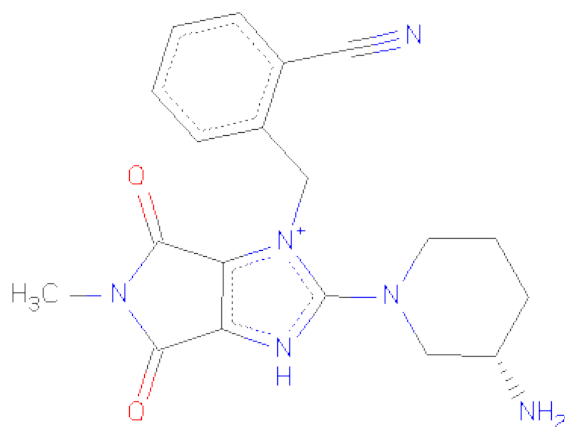
There are 2 unique types of molecules in this entry. The entry contains 11867 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Dipeptidyl peptidase 4 (Dipeptidyl peptidase IV) (DPP IV).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	730	Total	C	N	O	S	0	0	0
			5920	3789	981	1124	26			
1	B	730	Total	C	N	O	S	0	0	0
			5920	3789	981	1124	26			

- Molecule 2 is 2-[(3S)-3-AMINOPIPERIDIN-1-YL]-1-(2-CYANOBENZYL)-5-METHYL-4,6-DIOXO-3,4,5,6-TETRAHYDROPYRROLO[3,4-D]IMIDAZOL-1-IUM (three-letter code: LIR) (formula: C<sub>19</sub>H<sub>21</sub>N<sub>6</sub>O<sub>2</sub>).



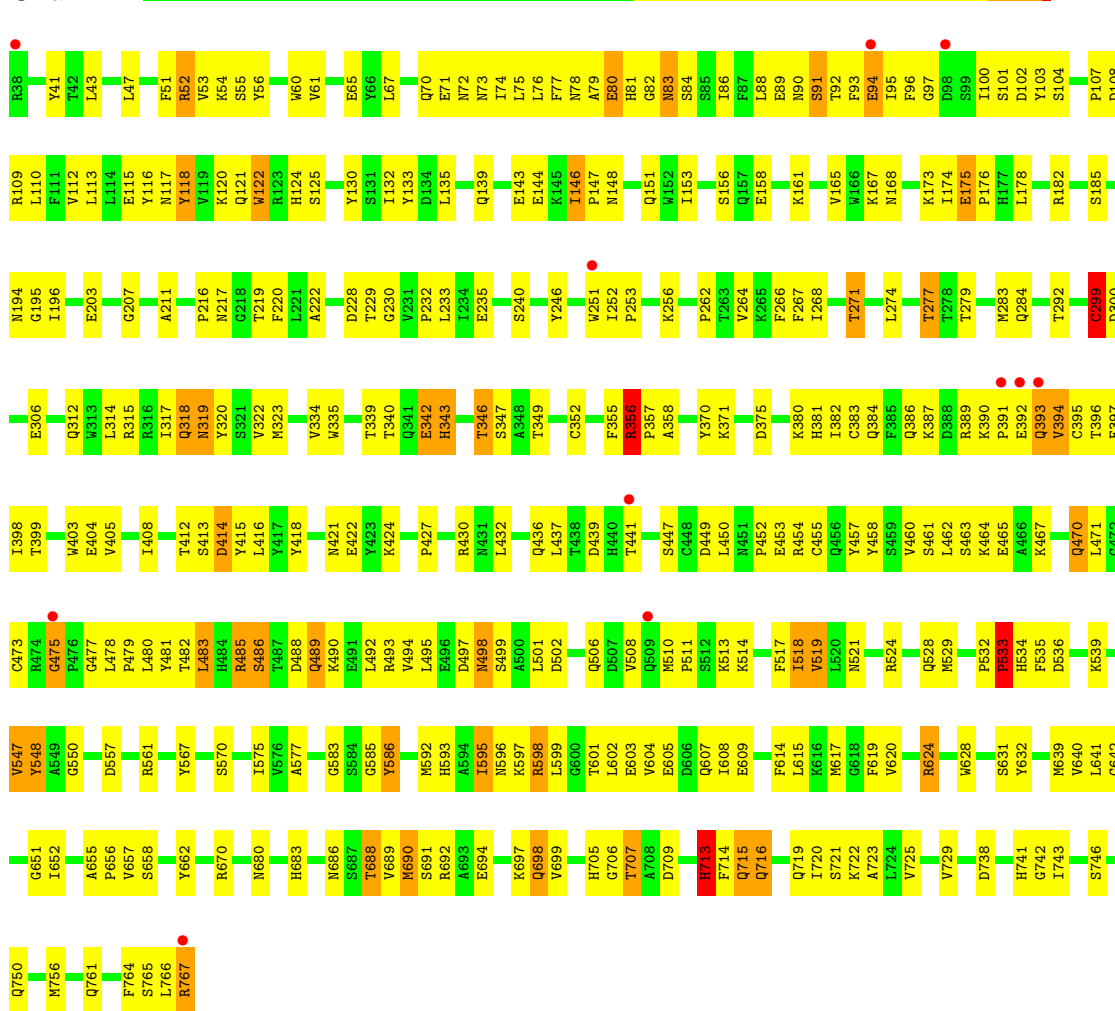
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			27	19	6	2		

### 3 Residue-property plots

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.

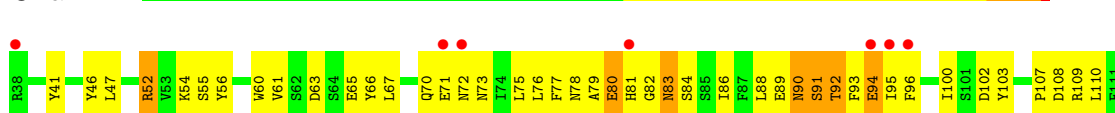
- Molecule 1: Dipeptidyl peptidase 4 (Dipeptidyl peptidase IV) (DPP IV)

Chain A:



- Molecule 1: Dipeptidyl peptidase 4 (Dipeptidyl peptidase IV) (DPP IV)

Chain B:



I743	G651	Y541	Y469	K409	Q312	G195	V112
S746	I652	F542	Q470	G401	R316	I196	L113
Q750	L543	L543	L471	A402	I317	G207	Y116
M756	P656	V547	C472	W403	Q318	S210	N117
S757	S658	Y548	C473	E404	I319	A211	Y118
Q761	R659	A549	G475	V405	Y320		V119
F764	Y662	P551	P476	I408	S321	P216	K120
S765	C552	S553	L480	E409	M322	N217	W122
L766	N563	N563	Y481	L411	G218	G218	R123
R767	T566		T482	S413	T219	H124	S125
	A569		L483	D414	D329	S125	Y126
	S570		H484	Y415	K350	A222	
	S570		R485	L416	V334	Q225	I132
	I576		S486	I419	W335	F226	L135
	G585		T487	S420	N227	N227	N136
	Y586		D488	N421	H343	P232	K137
	M592		Q489	E422	T339	L233	R138
	H593		K490	P427	E342	D228	Q139
	A594		L492	B430	H343	E234	E143
	N596		R493	I435	T346	E235	E144
	K597		V494	Q436	S347		K145
	R598		L495	L437	A348	Y246	I146
	L599		E496	T438	T349	W251	N148
	G600		D502	T439	R354	I252	Q151
	T601		K503	D439	F355	P253	W152
	L602			H440	R356	K256	I153
	E603		Q506	C444	P357		S156
	E609		D507	L446	A358	P262	Q157
	F614		V508	S447	Y370	T263	E158
	L615		Q509	C448	K371	V264	K161
	K616		M510	D449	D375	K285	
	M617		P511	L449	Y379	F266	Y164
	G618		S512	N451	K380	F267	V165
	F619		K513	P452	H381	T271	V166
	V620		I518	E453	I382	D272	K167
	R624		V519	R454	C383	S273	N168
	W628		B524	C455	Q384	L274	
	Y632		M529	Q456	F385	S276	K173
	Y635		P532	Y457	K387	T277	I174
	V636		P533	S458	K390	T278	E175
	T637		H534	V460	P391	T279	P176
	S638		P535	L462	E392	Q284	H177
	S639		D536	S463	Q393	P179	L178
	V640		K537	K464	V394	T292	
	L641		S538	E465	C395	Y297	R182
			K539	A466	T396	L298	S185
				K467	F397	C299	F193
				Y468	I398	E306	N194

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	207.29Å 207.29Å 207.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.90 19.95 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.6 (20.00-2.90) 99.7 (19.95-2.90)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.96 (at 2.88Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.253 , 0.290 0.248 , 0.284	Depositor DCC
$R_{free}$ test set	3294 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	54.7	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 19.1	EDS
Estimated twinning fraction	0.004 for l,-k,h	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.36$	Xtriage
Outliers	2 of 65257 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	11867	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.68% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>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: LIR

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	0.51	2/6088 (0.0%)	0.73	2/8278 (0.0%)
1	B	0.54	1/6088 (0.0%)	0.73	5/8278 (0.1%)
All	All	0.52	3/12176 (0.0%)	0.73	7/16556 (0.0%)

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 (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	299	CYS	CB-SG	-6.05	1.72	1.82
1	A	352	CYS	CB-SG	-5.39	1.73	1.81
1	A	299	CYS	CB-SG	-5.27	1.73	1.81

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	713	HIS	N-CA-C	7.13	130.24	111.00
1	A	713	HIS	N-CA-C	6.90	129.63	111.00
1	B	548	TYR	N-CA-C	-6.11	94.50	111.00
1	A	548	TYR	N-CA-C	-6.07	94.61	111.00
1	B	585	GLY	N-CA-C	6.04	128.19	113.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	736	TYR	Sidechain

## 5.2 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 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	5920	0	5632	306	0
1	B	5920	0	5632	323	0
2	A	27	0	21	2	0
All	All	11867	0	11285	626	0

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 27.

The worst 5 of 626 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:595:ILE:HD11	1:A:603:GLU:HB3	1.33	1.10
1:A:52:ARG:HH11	1:A:52:ARG:HB2	1.21	1.05
1:B:595:ILE:HD11	1:B:603:GLU:HB3	1.37	1.02
1:A:349:THR:HB	1:A:593:HIS:HD2	1.22	1.02
1:B:235:GLU:HG2	1:B:251:TRP:HB3	1.41	1.01

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	728/730 (100%)	638 (88%)	67 (9%)	23 (3%)	6 25

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	728/730 (100%)	634 (87%)	69 (10%)	25 (3%)	6	23
All	All	1456/1460 (100%)	1272 (87%)	136 (9%)	48 (3%)	6	24

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	80	GLU
1	A	143	GLU
1	A	450	LEU
1	A	475	GLY
1	A	533	PRO

### 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	643/651 (99%)	595 (92%)	48 (8%)	19	49
1	B	643/651 (99%)	601 (94%)	42 (6%)	24	58
All	All	1286/1302 (99%)	1196 (93%)	90 (7%)	21	53

5 of 90 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	686	ASN
1	B	94	GLU
1	B	686	ASN
1	A	688	THR
1	A	715	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	49	ASN
1	B	284	GLN
1	B	716	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	148	ASN
1	B	319	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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	LIR	A	900	-	30,30,30	2.00	6 (20%)	35,44,44	2.33	11 (31%)

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	LIR	A	900	-	-	0/6/20/20	0/2/4/4

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	900	LIR	C2-N6	5.60	1.46	1.39
2	A	900	LIR	C3-C2	-4.69	1.30	1.44
2	A	900	LIR	C7-N12	4.45	1.43	1.35
2	A	900	LIR	C1-C2	3.21	1.47	1.40
2	A	900	LIR	C4-C3	3.17	1.46	1.40

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	900	LIR	C7-N6-C2	-6.08	106.11	111.74
2	A	900	LIR	C13-N12-C7	-5.51	110.48	122.05
2	A	900	LIR	C2-C3-N8	5.13	110.23	106.07
2	A	900	LIR	C17-N12-C7	-4.23	112.91	122.01
2	A	900	LIR	C20-C25-C26	3.09	122.57	119.99

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers ⓘ

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<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	730/730 (100%)	-0.18	11 (1%) 70 79	21, 41, 69, 122	0
1	B	730/730 (100%)	-0.05	25 (3%) 43 51	22, 48, 81, 124	0
All	All	1460/1460 (100%)	-0.11	36 (2%) 54 64	21, 44, 75, 124	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	767	ARG	7.8
1	B	392	GLU	3.9
1	A	393	GLN	3.8
1	A	251	TRP	3.7
1	B	393	GLN	3.5

### 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<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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	LIR	A	900	27/27	0.25	2.83	45,45,45,45	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.