



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 30, 2017 – 08:45 AM EDT

PDB ID : 3PFQ  
Title : Crystal Structure and Allosteric Activation of Protein Kinase C beta II  
Authors : Leonard, T.A.; Rozycki, B.; Saidi, L.F.; Hummer, G.; Hurley, J.H.  
Deposited on : unknown  
Resolution : 4.00 Å(reported)

This is a wwPDB X-ray Structure 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/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 : rb-20030345  
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) : rb-20030345

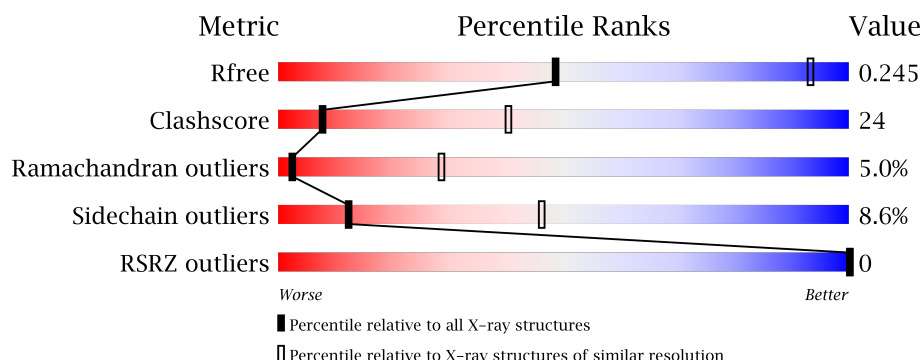
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 4.00 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	1088 (4.40-3.60)
Clashscore	112137	1187 (4.40-3.60)
Ramachandran outliers	110173	1139 (4.40-3.60)
Sidechain outliers	110143	1126 (4.40-3.60)
RSRZ outliers	101464	1099 (4.40-3.60)

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 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	674	<div> <div></div> <div>39%</div> <div>33%</div> <div>5%</div> <div>22%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4273 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 Protein kinase C beta type.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	523	Total	C	N	O	P	S	0	0	0
			4237	2711	708	785	3	30			

There are 31 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	EXPRESSION TAG	UNP P68403
A	71	SER	CYS	ENGINEERED MUTATION	UNP P68403
A	217	SER	CYS	ENGINEERED MUTATION	UNP P68403
A	?	-	ARG	ENGINEERED MUTATION	UNP P68403
A	622	SER	ASP	VARIANT	UNP P68403
A	623	GLY	LYS	VARIANT	UNP P68403
A	625	ASN	ASP	VARIANT	UNP P68403
A	626	ALA	THR	VARIANT	UNP P68403
A	627	GLU	SER	VARIANT	UNP P68403
A	631	ARG	LYS	VARIANT	UNP P68403
A	632	PHE	GLU	VARIANT	UNP P68403
A	636	HIS	GLN	VARIANT	UNP P68403
A	638	PRO	VAL	VARIANT	UNP P68403
A	639	VAL	GLU	VARIANT	UNP P68403
A	643	PRO	THR	VARIANT	UNP P68403
A	645	GLN	LYS	VARIANT	UNP P68403
A	646	GLU	LEU	VARIANT	UNP P68403
A	647	VAL	PHE	VARIANT	UNP P68403
A	649	ARG	MET	VARIANT	UNP P68403
A	651	ILE	LEU	VARIANT	UNP P68403
A	654	SER	ASN	VARIANT	UNP P68403
A	657	GLU	ALA	VARIANT	UNP P68403
A	661	PHE	TYR	VARIANT	UNP P68403
A	662	VAL	THR	VARIANT	UNP P68403
A	664	SER	PRO	VARIANT	UNP P68403
A	667	LEU	-	VARIANT	UNP P68403
A	668	LYS	-	VARIANT	UNP P68403

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	669	PRO	-	VARIANT	UNP P68403
A	670	GLU	-	VARIANT	UNP P68403
A	672	LYS	-	VARIANT	UNP P68403
A	673	SER	-	VARIANT	UNP P68403

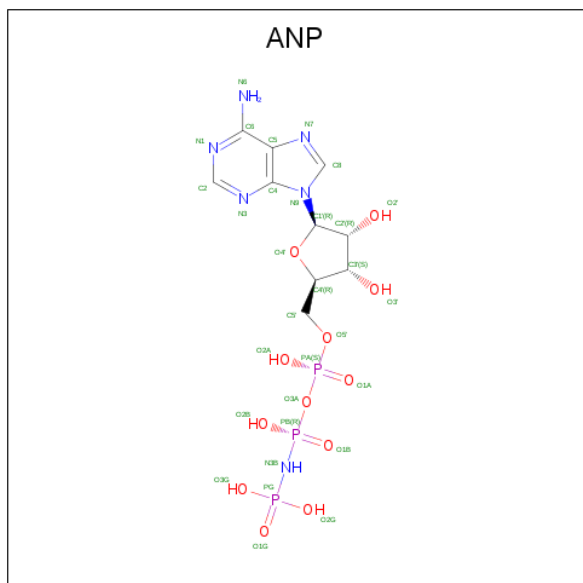
- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	3	Total Ca 3 3	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total Zn 2 2	0	0

- Molecule 4 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: C<sub>10</sub>H<sub>17</sub>N<sub>6</sub>O<sub>12</sub>P<sub>3</sub>).

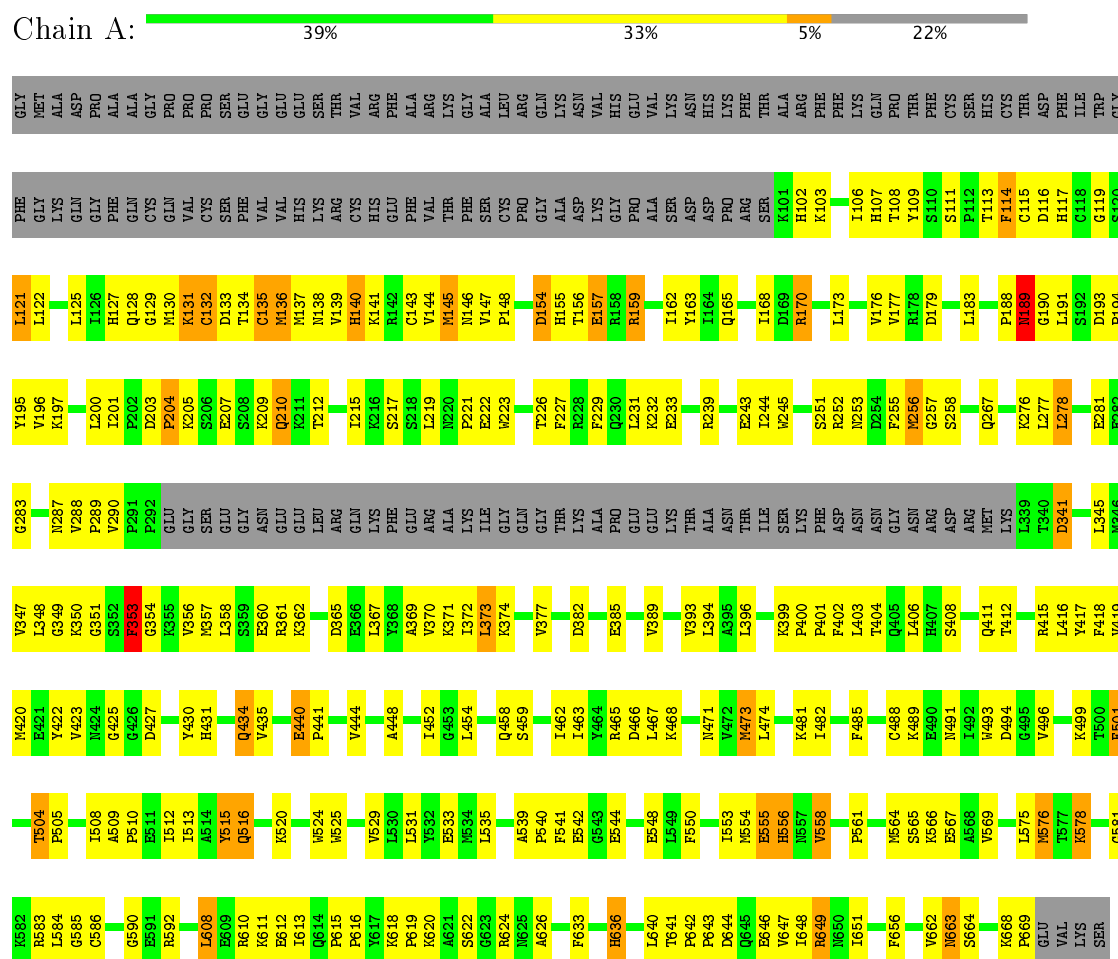


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O P 31 10 6 12 3	0	0

### 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 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: Protein kinase C beta type



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	114.27Å 114.27Å 170.84Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	57.10 – 4.00 57.14 – 4.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (57.10-4.00) 93.9 (57.14-4.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.05 (at 4.01Å)	Xtriage
Refinement program	CNS 1.3	Depositor
R, $R_{free}$	0.193 , 0.244 0.198 , 0.245	Depositor DCC
$R_{free}$ test set	508 reflections (4.76%)	DCC
Wilson B-factor (Å <sup>2</sup> )	88.8	Xtriage
Anisotropy	0.017	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 102.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.049 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	4273	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	110.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.22% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: TPO, ZN, ANP, CA, 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	0.26	0/4311	0.44	0/5811

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	4237	0	4156	204	0
2	A	3	0	0	0	0
3	A	2	0	0	0	0
4	A	31	0	13	0	0
All	All	4273	0	4169	204	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 24.

The worst 5 of 204 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:145:MET:HE2	1:A:146:ASN:H	1.32	0.94
1:A:130:MET:HG2	1:A:141:LYS:HA	1.55	0.87
1:A:162:ILE:HD12	1:A:244:ILE:HD11	1.58	0.83
1:A:372:ILE:HD13	1:A:640:LEU:HD11	1.59	0.83
1:A:189:ASN:HD22	1:A:191:LEU:H	1.27	0.81

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	516/674 (77%)	412 (80%)	78 (15%)	26 (5%)	<b>2</b> 28

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	131	LYS
1	A	135	CYS
1	A	136	MET
1	A	217	SER
1	A	353	PHE

### 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.

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
-----	-------	----------	-----------	----------	-------------

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	463/589 (79%)	423 (91%)	40 (9%)	12 46

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	341	ASP
1	A	382	ASP
1	A	608	LEU
1	A	365	ASP
1	A	434	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	491	ASN
1	A	516	GLN
1	A	614	GLN
1	A	471	ASN
1	A	557	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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	$\# Z  > 2$
1	TPO	A	500	1	9,10,11	1.27	1 (11%)	10,14,16	1.02	1 (10%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	TPO	A	641	1	9,10,11	1.56	2 (22%)	10,14,16	0.89	0
1	SEP	A	660	1	9,9,10	1.12	1 (11%)	9,12,14	1.66	1 (11%)

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	500	1	-	0/8/11/13	0/0/0/0
1	TPO	A	641	1	-	0/8/11/13	0/0/0/0
1	SEP	A	660	1	-	0/5/8/10	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	660	SEP	CA-C	2.08	1.53	1.50
1	A	500	TPO	CA-C	2.17	1.53	1.50
1	A	641	TPO	CA-C	2.20	1.53	1.50
1	A	641	TPO	P-OG1	3.08	1.65	1.59

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	660	SEP	P-OG-CB	-3.96	107.38	118.30
1	A	500	TPO	CG2-CB-CA	-2.32	108.90	113.22

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
1	A	641	TPO	1	0

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 6 ligands modelled in this entry, 5 are 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
4	ANP	A	800	-	29,33,33	3.46	9 (31%)	28,52,52	2.26	9 (32%)

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
4	ANP	A	800	-	-	0/13/38/38	0/3/3/3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	800	ANP	C2'-C1'	-2.77	1.49	1.53
4	A	800	ANP	PA-O1A	2.05	1.58	1.50
4	A	800	ANP	C5-C4	2.26	1.45	1.40
4	A	800	ANP	PG-N3B	4.24	1.74	1.63
4	A	800	ANP	PB-N3B	4.51	1.75	1.63

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	800	ANP	N3-C2-N1	-7.83	122.04	128.86
4	A	800	ANP	C4-C5-N7	-3.87	105.67	109.41
4	A	800	ANP	O2'-C2'-C1'	-2.66	103.28	111.61
4	A	800	ANP	O1B-PB-N3B	-2.54	107.99	111.79
4	A	800	ANP	O2'-C2'-C3'	-2.19	104.80	111.83

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

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	520/674 (77%)	-0.29	0 100 100	35, 105, 167, 234	0

There are no RSRZ outliers to report.

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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(Å <sup>2</sup> )	Q<0.9
1	SEP	A	660	10/11	0.91	0.12	-	128,143,143,143	0
1	TPO	A	641	11/12	0.91	0.18	-	171,171,177,177	0
1	TPO	A	500	11/12	0.87	0.24	-	93,112,112,112	0

### 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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ANP	A	800	31/31	0.78	0.31	-0.06	78,133,182,200	0
2	CA	A	674	1/1	0.88	0.20	-1.09	76,76,76,76	0
3	ZN	A	751	1/1	0.98	0.13	-1.13	96,96,96,96	0
3	ZN	A	750	1/1	0.99	0.06	-1.90	113,113,113,113	0
2	CA	A	675	1/1	0.95	0.15	-	61,61,61,61	0
2	CA	A	676	1/1	0.94	0.16	-	75,75,75,75	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.