



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:12 PM GMT

PDB ID : 1SJD  
Title : x-ray structure of o-succinylbenzoate synthase complexed with n-succinyl phenylglycine  
Authors : Thoden, J.B.; Taylor-Ringia, E.A.; Garrett, J.B.; Gerlt, J.A.; Holden, H.M.; Rayment, I.  
Deposited on : 2004-03-03  
Resolution : 1.87 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

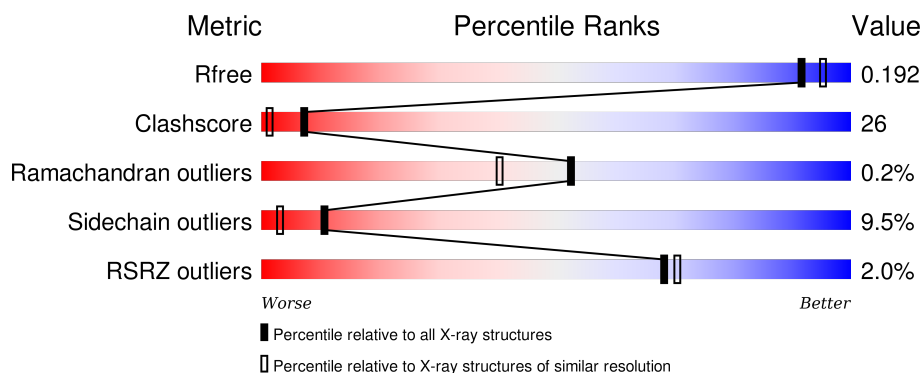
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.87 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	6965 (1.90-1.86)
Clashscore	102246	7778 (1.90-1.86)
Ramachandran outliers	100387	7691 (1.90-1.86)
Sidechain outliers	100360	7692 (1.90-1.86)
RSRZ outliers	91569	6979 (1.90-1.86)

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	368	<div> <div>57%</div> <div>32%</div> <div>10%</div> <div>.</div> </div>
1	B	368	<div> <div>67%</div> <div>23%</div> <div>9%</div> <div>.</div> </div>
1	C	368	<div> <div>61%</div> <div>32%</div> <div>6%</div> <div>.</div> </div>
1	D	368	<div> <div>7%</div> <div>47%</div> <div>38%</div> <div>12%</div> <div>.</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	NPG	B	1163	-	-	-	X
2	NPG	D	1500	-	-	X	X

## 2 Entry composition [i](#)

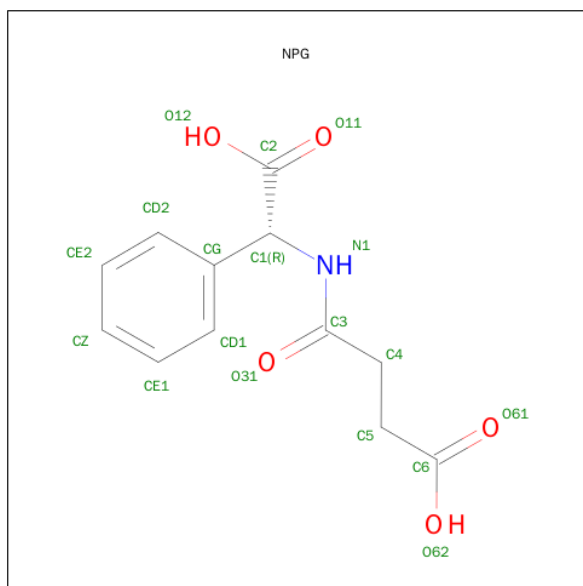
There are 3 unique types of molecules in this entry. The entry contains 12337 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 N-acylamino acid racemase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	367	Total	C	N	O	S	0	0	0
			2765	1758	480	515	12			
1	B	368	Total	C	N	O	S	0	0	0
			2772	1761	481	518	12			
1	C	367	Total	C	N	O	S	0	0	0
			2765	1758	480	515	12			
1	D	367	Total	C	N	O	S	0	0	0
			2765	1758	480	515	12			

- Molecule 2 is N-SUCCINYL PHENYLGLYCINE (three-letter code: NPG) (formula: C<sub>12</sub>H<sub>13</sub>NO<sub>5</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			18	12	1	5		
2	B	1	Total	C	N	O	0	0
			18	12	1	5		

*Continued on next page...*

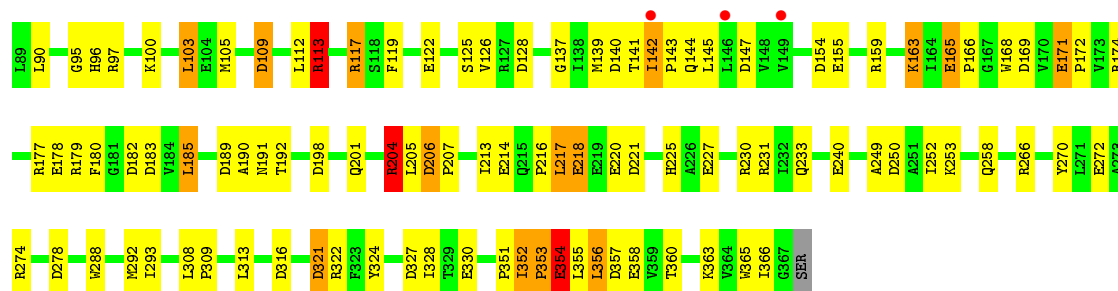
*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	C	1	Total	C	N	O	0	0
			18	12	1	5		
2	D	1	Total	C	N	O	0	0
			18	12	1	5		
2	B	1	Total	C	N	O	0	0
			18	12	1	5		
2	C	1	Total	C	N	O	0	0
			18	12	1	5		

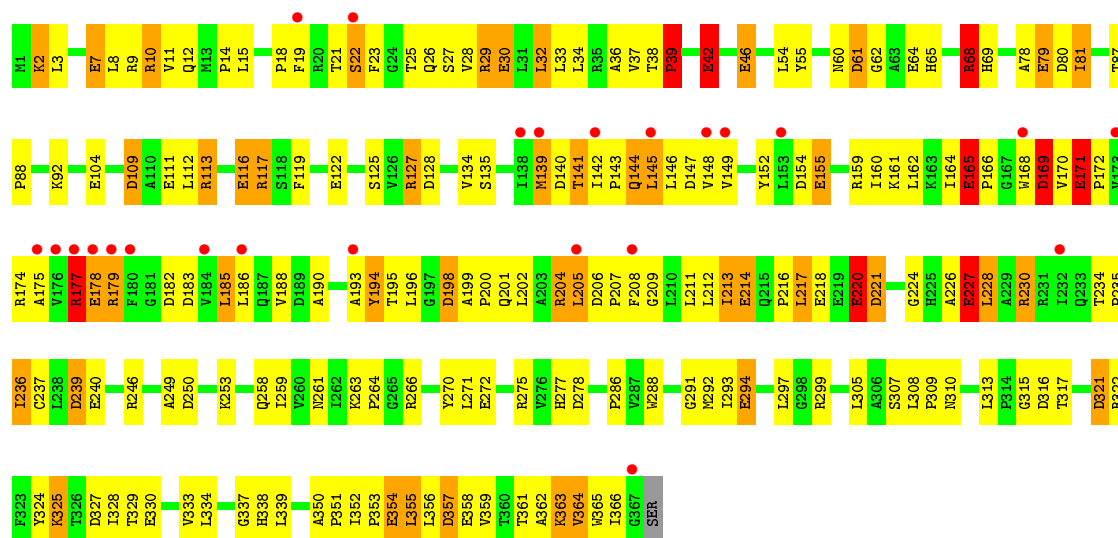
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	261	Total	O	0	0
			261	261		
3	B	375	Total	O	0	0
			375	375		
3	C	336	Total	O	0	0
			336	336		
3	D	190	Total	O	0	0
			190	190		





- Molecule 1: N-acylamino acid racemase



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	215.20 Å   215.20 Å   257.60 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	50.00 – 1.87 49.66 – 1.87	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.00-1.87) 99.7 (49.66-1.87)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.87 (at 1.87 Å)	Xtriage
Refinement program	TNT	Depositor
R, $R_{free}$	0.196 , 0.241 0.193 , 0.192	Depositor DCC
$R_{free}$ test set	18681 reflections (11.11%)	DCC
Wilson B-factor (Å <sup>2</sup> )	26.0	Xtriage
Anisotropy	0.037	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 115.9	EDS
Estimated twinning fraction	0.017 for -2/3*h-1/3*k+2/3*l,-1/3*h-2/3*k-2/3*l,2/3*h-2/3*k+1/3*l 0.014 for -h,1/3*h-1/3*k+2/3*l,2/3*h+4/3*k+1/3*l 0.011 for -1/3*h+1/3*k-2/3*l,-k,-4/3*h-2/3*k+1/3*l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 186821 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	12337	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.48% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: NPG

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	0.98	20/2821 (0.7%)	1.36	43/3842 (1.1%)
1	B	0.96	18/2828 (0.6%)	1.41	51/3850 (1.3%)
1	C	0.98	18/2821 (0.6%)	1.46	51/3842 (1.3%)
1	D	1.04	23/2821 (0.8%)	1.45	46/3842 (1.2%)
All	All	0.99	79/11291 (0.7%)	1.42	191/15376 (1.2%)

The worst 5 of 79 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	220	GLU	CD-OE2	9.11	1.35	1.25
1	D	358	GLU	CD-OE2	7.63	1.34	1.25
1	D	178	GLU	CD-OE2	7.59	1.33	1.25
1	A	178	GLU	CD-OE2	7.56	1.33	1.25
1	A	358	GLU	CD-OE2	7.54	1.33	1.25

The worst 5 of 191 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	230	ARG	NE-CZ-NH1	13.27	126.94	120.30
1	C	140	ASP	CB-CG-OD2	-10.82	108.56	118.30
1	B	159	ARG	NE-CZ-NH1	10.75	125.67	120.30
1	D	299	ARG	NE-CZ-NH1	9.66	125.13	120.30
1	B	154	ASP	CB-CG-OD1	9.42	126.77	118.30

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	2765	0	2805	130	1
1	B	2772	0	2810	85	0
1	C	2765	0	2805	111	0
1	D	2765	0	2805	260	1
2	A	18	0	11	1	0
2	B	36	0	21	1	0
2	C	36	0	21	7	0
2	D	18	0	11	12	0
3	A	261	0	0	10	1
3	B	375	0	0	8	0
3	C	336	0	0	11	0
3	D	190	0	0	7	1
All	All	12337	0	11289	585	2

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 26.

The worst 5 of 585 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:D:38:THR:HB	1:D:39:PRO:HD2	1.20	1.17
1:C:10:ARG:HD2	1:C:32:LEU:CD1	1.74	1.14
1:D:141:THR:HG23	1:D:143:PRO:HD2	1.16	1.12
1:A:1:MET:CE	1:A:38:THR:HG21	1.83	1.09
1:A:1:MET:HE3	1:A:38:THR:HG21	1.06	1.03

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:111:GLU:OE2	1:D:117:ARG:NH2[17_555]	1.85	0.35
3:A:1310:HOH:O	3:D:1604:HOH:O[17_555]	1.97	0.23

## 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	365/368 (99%)	350 (96%)	15 (4%)	0	100	100
1	B	366/368 (100%)	358 (98%)	8 (2%)	0	100	100
1	C	365/368 (99%)	349 (96%)	15 (4%)	1 (0%)	46	33
1	D	365/368 (99%)	334 (92%)	29 (8%)	2 (0%)	34	20
All	All	1461/1472 (99%)	1391 (95%)	67 (5%)	3 (0%)	52	40

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	144	GLN
1	C	353	PRO
1	D	39	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	290/291 (100%)	260 (90%)	30 (10%)	9	2
1	B	291/291 (100%)	268 (92%)	23 (8%)	15	5
1	C	290/291 (100%)	269 (93%)	21 (7%)	18	7
1	D	290/291 (100%)	254 (88%)	36 (12%)	6	1
All	All	1161/1164 (100%)	1051 (90%)	110 (10%)	11	3

5 of 110 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	339	LEU
1	C	142	ILE
1	D	220	GLU
1	B	358	GLU
1	C	8	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	65	HIS
1	B	258	GLN
1	D	65	HIS
1	B	12	GLN
1	C	258	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](#)

6 ligands 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
2	NPG	A	1200	-	12,18,18	1.42	2 (16%)	15,23,23	1.45	3 (20%)
2	NPG	B	1163	-	12,18,18	1.48	2 (16%)	15,23,23	3.04	7 (46%)
2	NPG	B	1300	-	12,18,18	1.42	2 (16%)	15,23,23	2.28	3 (20%)
2	NPG	C	1164	-	12,18,18	1.45	2 (16%)	15,23,23	3.16	6 (40%)
2	NPG	C	1400	-	12,18,18	1.39	2 (16%)	15,23,23	0.66	0
2	NPG	D	1500	-	12,18,18	1.33	2 (16%)	15,23,23	1.26	2 (13%)

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	NPG	A	1200	-	-	0/11/17/17	0/1/1/1
2	NPG	B	1163	-	-	0/11/17/17	0/1/1/1
2	NPG	B	1300	-	-	0/11/17/17	0/1/1/1
2	NPG	C	1164	-	-	0/11/17/17	0/1/1/1
2	NPG	C	1400	-	-	0/11/17/17	0/1/1/1
2	NPG	D	1500	-	-	0/11/17/17	0/1/1/1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1163	NPG	CG-C1	-3.74	1.47	1.52
2	C	1164	NPG	CG-C1	-3.58	1.47	1.52
2	A	1200	NPG	CG-C1	-2.97	1.48	1.52
2	B	1300	NPG	CG-C1	-2.82	1.48	1.52
2	C	1400	NPG	CG-C1	-2.69	1.48	1.52

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1300	NPG	C4-C3-N1	-5.89	106.24	115.83
2	B	1300	NPG	C4-C5-C6	-4.37	104.74	112.75
2	B	1163	NPG	CD1-CG-C1	-4.28	113.49	120.77
2	B	1163	NPG	C5-C4-C3	-4.01	103.18	113.01
2	B	1163	NPG	CG-C1-N1	-3.15	105.30	111.18

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1200	NPG	1	0
2	B	1163	NPG	1	0
2	C	1164	NPG	3	0
2	C	1400	NPG	4	0
2	D	1500	NPG	12	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 [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, 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	367/368 (99%)	-0.40	1 (0%) 94 94	17, 37, 69, 96	0
1	B	368/368 (100%)	-0.47	2 (0%) 91 92	17, 27, 55, 91	0
1	C	367/368 (99%)	-0.50	3 (0%) 87 88	19, 30, 61, 97	0
1	D	367/368 (99%)	0.06	24 (6%) 22 24	15, 46, 89, 100	0
All	All	1469/1472 (99%)	-0.33	30 (2%) 68 70	15, 33, 78, 100	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	145	LEU	6.2
1	D	184	VAL	5.9
1	D	208	PHE	4.2
1	D	168	TRP	4.2
1	D	149	VAL	4.2

### 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	NPG	B	1163	18/18	0.90	0.12	8.78	24,54,100,100	0
2	NPG	D	1500	18/18	0.82	0.25	3.07	50,84,100,100	0
2	NPG	A	1200	18/18	0.94	0.12	1.65	30,53,98,100	0
2	NPG	B	1300	18/18	0.94	0.11	1.56	29,44,65,67	0
2	NPG	C	1164	18/18	0.87	0.14	1.52	29,79,100,100	0
2	NPG	C	1400	18/18	0.97	0.07	-0.10	23,33,47,48	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.