



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 23, 2023 – 03:22 AM EDT

PDB ID : 2ZF4
Title : Crystal Structure of VioE complexed with phenylpyruvic acid
Authors : Hirano, S.; Shiro, Y.; Nagano, S.
Deposited on : 2007-12-20
Resolution : 2.18 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

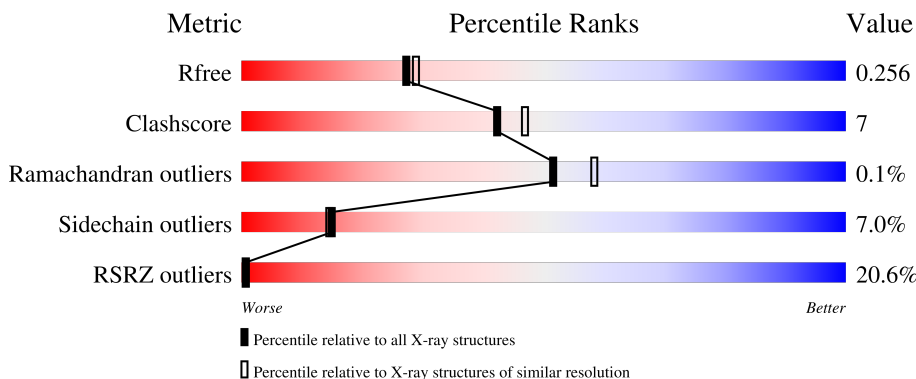
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.18 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}	130704	6864 (2.20-2.16)
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)
RSRZ outliers	127900	6738 (2.20-2.16)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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	194	
1	B	194	
1	C	194	
1	D	194	
1	E	194	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	194	<p>20% 79% 14% • 6%</p>

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	PPY	A	193	-	-	-	X
2	PPY	B	193	-	-	-	X

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9238 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 Hypothetical protein VioE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	187	1508	950	273	278	7	0	2	0
1	B	180	1448	912	262	268	6	0	0	0
1	C	181	1461	922	264	269	6	0	1	0
1	D	179	1442	909	260	266	7	0	1	0
1	E	177	1423	897	256	264	6	0	0	0
1	F	183	1492	937	273	275	7	0	3	0

There are 18 discrepancies between the modelled and reference sequences:

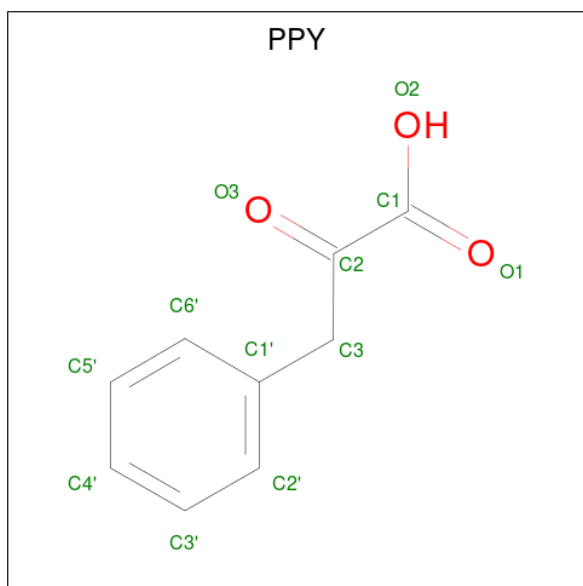
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP Q7NSZ5
A	-1	SER	-	expression tag	UNP Q7NSZ5
A	0	HIS	-	expression tag	UNP Q7NSZ5
B	-2	GLY	-	expression tag	UNP Q7NSZ5
B	-1	SER	-	expression tag	UNP Q7NSZ5
B	0	HIS	-	expression tag	UNP Q7NSZ5
C	-2	GLY	-	expression tag	UNP Q7NSZ5
C	-1	SER	-	expression tag	UNP Q7NSZ5
C	0	HIS	-	expression tag	UNP Q7NSZ5
D	-2	GLY	-	expression tag	UNP Q7NSZ5
D	-1	SER	-	expression tag	UNP Q7NSZ5
D	0	HIS	-	expression tag	UNP Q7NSZ5
E	-2	GLY	-	expression tag	UNP Q7NSZ5
E	-1	SER	-	expression tag	UNP Q7NSZ5
E	0	HIS	-	expression tag	UNP Q7NSZ5
F	-2	GLY	-	expression tag	UNP Q7NSZ5
F	-1	SER	-	expression tag	UNP Q7NSZ5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	HIS	-	expression tag	UNP Q7NSZ5

- Molecule 2 is 3-PHENYLPYRUVIC ACID (three-letter code: PPY) (formula: C₉H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	12	9	3	0	0
2	A	1	12	9	3	0	0
2	B	1	12	9	3	0	0
2	B	1	12	9	3	0	0
2	C	1	12	9	3	0	0
2	C	1	12	9	3	0	0
2	D	1	12	9	3	0	0
2	D	1	12	9	3	0	0
2	E	1	12	9	3	0	0
2	E	1	12	9	3	0	0
2	F	1	12	9	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	F	1	Total	C	O	0	0
			12	9	3		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	68	Total	O	0	0
			68	68		
3	B	36	Total	O	0	0
			36	36		
3	C	67	Total	O	0	0
			67	67		
3	D	41	Total	O	0	0
			41	41		
3	E	49	Total	O	0	0
			49	49		
3	F	59	Total	O	0	0
			59	59		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	83.42Å 91.03Å 159.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.93 – 2.18 29.81 – 2.18	Depositor EDS
% Data completeness (in resolution range)	97.0 (29.93-2.18) 97.0 (29.81-2.18)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.11 (at 2.18Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.206 , 0.254 0.208 , 0.256	Depositor DCC
R_{free} test set	3115 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	36.6	Xtrriage
Anisotropy	0.444	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 44.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	9238	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.58% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: PPY

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	0/1556	0.71	1/2112 (0.0%)
1	B	0.72	6/1489 (0.4%)	0.69	2/2022 (0.1%)
1	C	0.52	0/1505	0.70	1/2042 (0.0%)
1	D	0.49	0/1484	0.66	1/2014 (0.0%)
1	E	0.49	0/1461	0.69	2/1981 (0.1%)
1	F	0.53	0/1536	0.70	3/2086 (0.1%)
All	All	0.55	6/9031 (0.1%)	0.69	10/12257 (0.1%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	154	SER	CB-OG	12.52	1.58	1.42
1	B	131	ARG	CZ-NH1	8.29	1.43	1.33
1	B	153	ALA	C-O	6.60	1.35	1.23
1	B	153	ALA	C-N	5.96	1.47	1.34
1	B	154	SER	CA-CB	5.18	1.60	1.52

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	45	ARG	NE-CZ-NH1	7.74	124.17	120.30
1	B	131	ARG	NE-CZ-NH2	-7.27	116.67	120.30
1	B	131	ARG	NE-CZ-NH1	7.03	123.81	120.30
1	C	129	LEU	CA-CB-CG	6.14	129.43	115.30
1	F	45	ARG	NE-CZ-NH2	-5.66	117.47	120.30

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	1508	0	1457	25	0
1	B	1448	0	1385	15	0
1	C	1461	0	1406	10	0
1	D	1442	0	1380	32	0
1	E	1423	0	1355	21	0
1	F	1492	0	1429	18	0
2	A	24	0	14	3	0
2	B	24	0	14	3	0
2	C	24	0	14	3	0
2	D	24	0	14	5	0
2	E	24	0	14	0	0
2	F	24	0	14	5	0
3	A	68	0	0	2	0
3	B	36	0	0	3	0
3	C	67	0	0	1	0
3	D	41	0	0	1	0
3	E	49	0	0	2	0
3	F	59	0	0	0	0
All	All	9238	0	8496	120	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 7.

The worst 5 of 120 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:109:PHE:HB3	2:D:192:PPY:H32	1.50	0.92
1:F:109:PHE:HB3	2:F:193:PPY:H32	1.51	0.90
1:A:36:TRP:CZ3	1:A:45:ARG:HD2	2.09	0.88
1:D:111:PRO:HD2	1:D:114:VAL:CG1	2.06	0.86
1:A:87:THR:HG22	1:A:89:LEU:H	1.38	0.86

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	187/194 (96%)	180 (96%)	7 (4%)	0	100	100
1	B	176/194 (91%)	169 (96%)	7 (4%)	0	100	100
1	C	178/194 (92%)	172 (97%)	6 (3%)	0	100	100
1	D	174/194 (90%)	169 (97%)	5 (3%)	0	100	100
1	E	169/194 (87%)	164 (97%)	4 (2%)	1 (1%)	25	24
1	F	182/194 (94%)	175 (96%)	7 (4%)	0	100	100
All	All	1066/1164 (92%)	1029 (96%)	36 (3%)	1 (0%)	51	58

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	119	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	156/160 (98%)	142 (91%)	14 (9%)	9	8
1	B	150/160 (94%)	142 (95%)	8 (5%)	22	25
1	C	151/160 (94%)	138 (91%)	13 (9%)	10	9
1	D	149/160 (93%)	139 (93%)	10 (7%)	16	16
1	E	147/160 (92%)	138 (94%)	9 (6%)	18	19
1	F	154/160 (96%)	145 (94%)	9 (6%)	20	21

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	907/960 (94%)	844 (93%)	63 (7%)	15	15

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	115	LEU
1	F	29	GLN
1	D	45	ARG
1	F	8	LEU
1	F	108	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	29	GLN
1	D	92	GLN
1	F	29	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

12 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	PPY	A	192	-	12,12,12	1.73	1 (8%)	15,15,15	1.16	2 (13%)
2	PPY	F	192	-	12,12,12	1.64	2 (16%)	15,15,15	0.85	0
2	PPY	D	193	-	12,12,12	1.71	2 (16%)	15,15,15	1.23	1 (6%)
2	PPY	C	192	-	12,12,12	1.81	1 (8%)	15,15,15	1.15	1 (6%)
2	PPY	D	192	-	12,12,12	1.33	1 (8%)	15,15,15	1.09	1 (6%)
2	PPY	E	193	-	12,12,12	1.51	1 (8%)	15,15,15	0.96	0
2	PPY	B	192	-	12,12,12	1.61	1 (8%)	15,15,15	1.57	1 (6%)
2	PPY	A	193	-	12,12,12	1.42	1 (8%)	15,15,15	1.12	1 (6%)
2	PPY	F	193	-	12,12,12	1.40	1 (8%)	15,15,15	1.10	1 (6%)
2	PPY	E	192	-	12,12,12	1.74	2 (16%)	15,15,15	1.07	0
2	PPY	C	193	-	12,12,12	1.45	1 (8%)	15,15,15	1.07	1 (6%)
2	PPY	B	193	-	12,12,12	1.39	1 (8%)	15,15,15	1.16	1 (6%)

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	PPY	A	192	-	-	5/8/8/8	0/1/1/1
2	PPY	F	192	-	-	3/8/8/8	0/1/1/1
2	PPY	D	193	-	-	7/8/8/8	0/1/1/1
2	PPY	C	192	-	-	3/8/8/8	0/1/1/1
2	PPY	D	192	-	-	3/8/8/8	0/1/1/1
2	PPY	E	193	-	-	3/8/8/8	0/1/1/1
2	PPY	B	192	-	-	5/8/8/8	0/1/1/1
2	PPY	A	193	-	-	2/8/8/8	0/1/1/1
2	PPY	F	193	-	-	2/8/8/8	0/1/1/1
2	PPY	E	192	-	-	3/8/8/8	0/1/1/1
2	PPY	C	193	-	-	2/8/8/8	0/1/1/1
2	PPY	B	193	-	-	2/8/8/8	0/1/1/1

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	192	PPY	C2-C1	-5.42	1.46	1.53
2	E	192	PPY	C2-C1	-5.13	1.46	1.53
2	D	193	PPY	C2-C1	-4.89	1.47	1.53
2	F	192	PPY	C2-C1	-4.85	1.47	1.53
2	A	192	PPY	C2-C1	-4.51	1.47	1.53

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	192	PPY	C1'-C3-C2	4.89	119.98	113.64
2	D	193	PPY	C1'-C3-C2	3.24	117.84	113.64
2	F	193	PPY	C3-C2-C1	2.64	122.32	117.85
2	C	192	PPY	C1'-C3-C2	2.64	117.06	113.64
2	A	192	PPY	C3-C2-C1	2.55	122.17	117.85

There are no chirality outliers.

5 of 40 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	192	PPY	O1-C1-C2-O3
2	A	192	PPY	O1-C1-C2-C3
2	A	192	PPY	O2-C1-C2-C3
2	B	192	PPY	O1-C1-C2-O3
2	B	192	PPY	O1-C1-C2-C3

There are no ring outliers.

6 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	193	PPY	1	0
2	D	192	PPY	4	0
2	A	193	PPY	3	0
2	F	193	PPY	5	0
2	C	193	PPY	3	0
2	B	193	PPY	3	0

5.7 Other polymers [i](#)

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 [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	187/194 (96%)	1.05	34 (18%) 1 1	15, 25, 33, 45	0
1	B	180/194 (92%)	1.45	51 (28%) 0 0	17, 28, 37, 41	0
1	C	181/194 (93%)	0.86	27 (14%) 2 2	17, 23, 34, 38	0
1	D	179/194 (92%)	1.06	32 (17%) 1 1	16, 25, 33, 39	0
1	E	177/194 (91%)	1.27	41 (23%) 0 0	13, 20, 31, 35	0
1	F	183/194 (94%)	1.14	39 (21%) 0 0	12, 21, 34, 40	0
All	All	1087/1164 (93%)	1.14	224 (20%) 1 1	12, 24, 35, 45	0

The worst 5 of 224 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	33	GLY	5.3
1	F	125	ARG	5.0
1	D	17	TYR	5.0
1	F	152	ALA	4.9
1	E	85	GLU	4.9

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 monosaccharides 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. 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	B-factors(Å ²)	Q<0.9
2	PPY	B	193	12/12	0.47	0.41	55,57,60,60	0
2	PPY	A	193	12/12	0.61	0.41	49,50,55,56	0
2	PPY	C	193	12/12	0.62	0.34	31,34,46,47	0
2	PPY	D	192	12/12	0.68	0.32	42,44,49,49	0
2	PPY	B	192	12/12	0.75	0.25	34,37,39,40	0
2	PPY	E	193	12/12	0.76	0.29	46,47,49,50	0
2	PPY	F	193	12/12	0.78	0.28	28,30,43,43	0
2	PPY	D	193	12/12	0.84	0.21	47,50,51,51	0
2	PPY	E	192	12/12	0.87	0.19	32,35,38,38	0
2	PPY	C	192	12/12	0.87	0.18	34,36,41,42	0
2	PPY	F	192	12/12	0.87	0.22	26,29,31,32	0
2	PPY	A	192	12/12	0.87	0.20	31,33,35,35	0

6.5 Other polymers [i](#)

There are no such residues in this entry.