



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:47 AM GMT

PDB ID : 2YAJ
Title : CRYSTAL STRUCTURE OF GLYCYL RADICAL ENZYME WITH BOUND
SUBSTRATE
Authors : Martins, B.M.; Blaser, M.; Feliks, M.; Ullmann, G.M.; Selmer, T.
Deposited on : 2011-02-23
Resolution : 1.81 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
<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

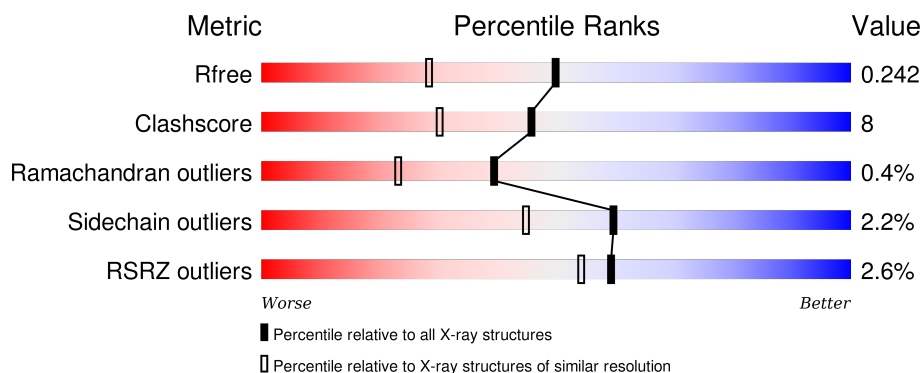
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.81 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	5422 (1.84-1.80)
Clashscore	102246	6347 (1.84-1.80)
Ramachandran outliers	100387	6276 (1.84-1.80)
Sidechain outliers	100360	6276 (1.84-1.80)
RSRZ outliers	91569	5439 (1.84-1.80)

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 ≥ 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	897	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, green 1%, green 81%, yellow 81%, yellow 96%, grey 96%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> 81% 15% </div> </div>
1	C	897	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 2%, green 2%, green 81%, yellow 81%, yellow 96%, grey 96%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> 81% 15% </div> </div>
2	B	86	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 9%, green 9%, green 79%, yellow 79%, yellow 98%, grey 98%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> 79% 19% </div> </div>
2	D	86	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 13%, green 13%, green 78%, yellow 78%, yellow 95%, grey 95%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> 78% 17% 5% </div> </div>

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 31755 atoms, of which 14606 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.

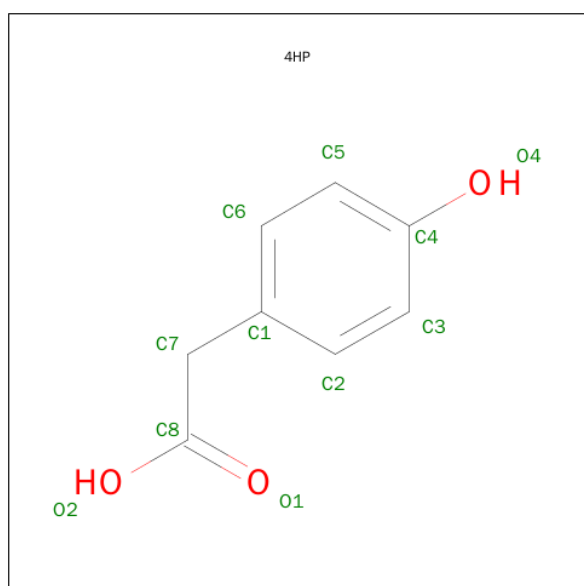
- Molecule 1 is a protein called 4-HYDROXYPHENYLACETATE DECARBOXYLASE LARGE SUBUNIT.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	869	Total	C	H	N	O	S	35	2	0
			13524	4347	6682	1126	1311	58			
1	C	871	Total	C	H	N	O	S	3	3	0
			13565	4360	6700	1129	1318	58			

- Molecule 2 is a protein called 4-HYDROXYPHENYLACETATE DECARBOXYLASE SMALL SUBUNIT.

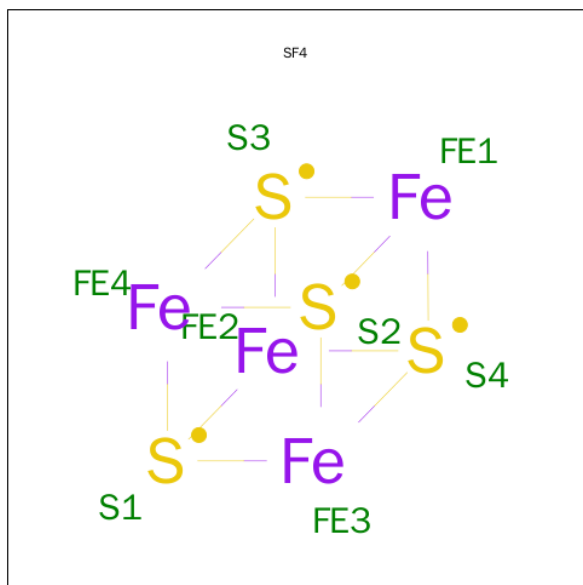
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	86	Total	C	H	N	O	S	0	0	0
			1249	398	604	108	127	12			
2	D	86	Total	C	H	N	O	S	0	0	0
			1249	398	604	108	127	12			

- Molecule 3 is 4-HYDROXYPHENYLACETATE (three-letter code: 4HP) (formula: $C_8H_8O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			19	8	8	3		
3	C	1	Total	C	H	O	0	0
			19	8	8	3		

- Molecule 4 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	Fe	S	0	0
			8	4	4		
4	B	1	Total	Fe	S	0	0
			8	4	4		
4	D	1	Total	Fe	S	0	0
			8	4	4		
4	D	1	Total	Fe	S	0	0
			8	4	4		

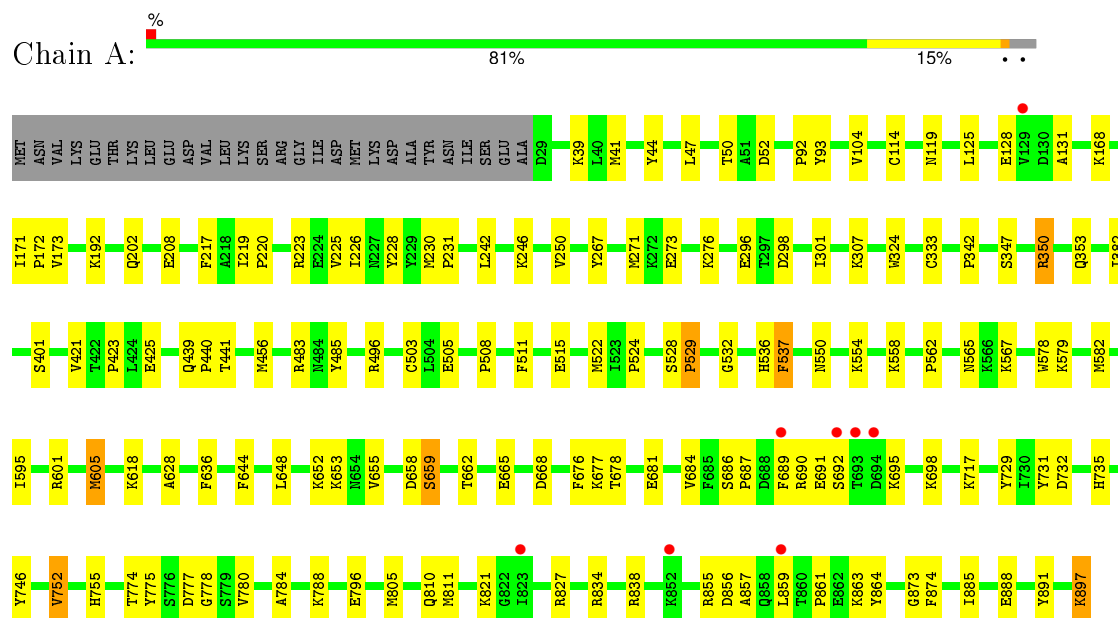
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1015	Total	O	0	0
			1015	1015		
5	B	66	Total	O	0	0
			66	66		
5	C	950	Total	O	0	0
			950	950		
5	D	67	Total	O	0	0
			67	67		

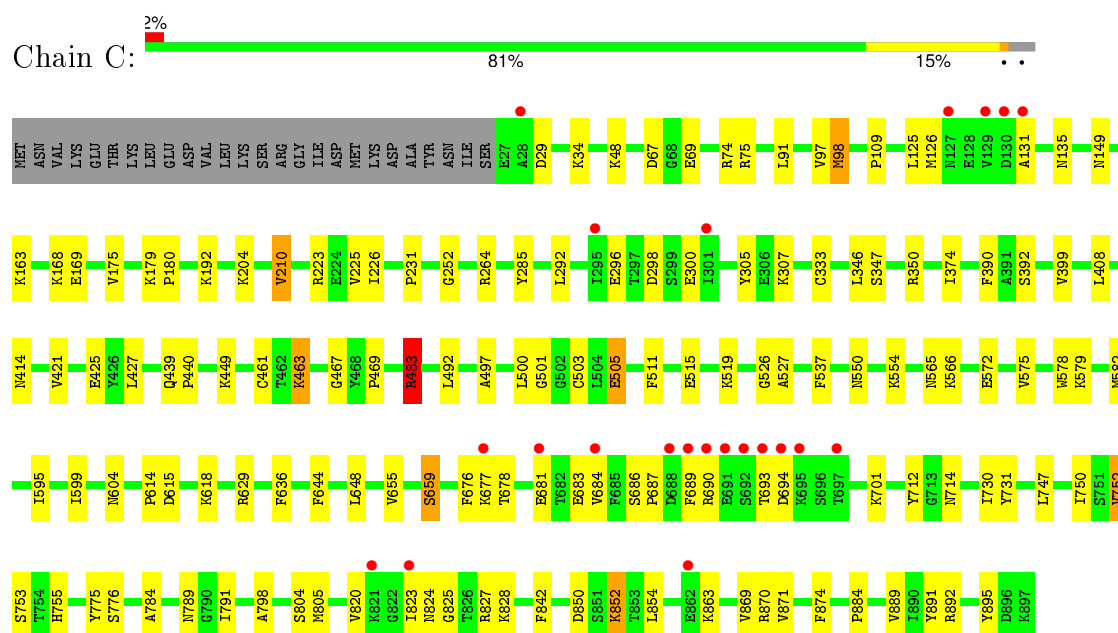
3 Residue-property plots

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: 4-HYDROXYPHENYLACETATE DECARBOXYLASE LARGE SUBUNIT

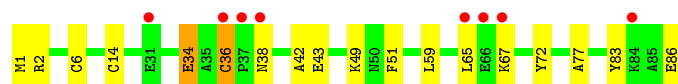


• Molecule 1: 4-HYDROXYPHENYLACETATE DECARBOXYLASE LARGE SUBUNIT




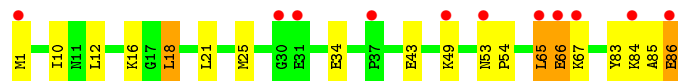
● Molecule 2: 4-HYDROXYPHENYLACETATE DECARBOXYLASE SMALL SUBUNIT

Chain B:  9% 79% 19%



● Molecule 2: 4-HYDROXYPHENYLACETATE DECARBOXYLASE SMALL SUBUNIT

Chain D:  13% 78% 17% 5%



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	133.05Å 228.59Å 148.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.16 – 1.81 29.16 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.8 (29.16-1.81) 99.3 (29.16-1.75)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.55 (at 1.75Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.188 , 0.244 0.185 , 0.242	Depositor DCC
R_{free} test set	10134 reflections (5.26%)	DCC
Wilson B-factor (Å ²)	17.0	Xtriage
Anisotropy	0.120	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.42 , 86.4	EDS
Estimated twinning fraction	0.055 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.075 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.38$, $\langle L^2 \rangle = 0.21$	Xtriage
Outliers	0 of 223564 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	31755	wwPDB-VP
Average B, all atoms (Å ²)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: SF4, 4HP

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	0.80	1/7008 (0.0%)	0.77	2/9477 (0.0%)
1	C	0.82	1/7031 (0.0%)	0.80	6/9508 (0.1%)
2	B	0.68	0/659	0.74	0/887
2	D	0.69	0/659	0.71	0/887
All	All	0.80	2/15357 (0.0%)	0.78	8/20759 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	527	ALA	CA-CB	6.95	1.67	1.52
1	A	114	CYS	CB-SG	-5.69	1.72	1.81

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	483	ARG	NE-CZ-NH2	-7.58	116.51	120.30
1	C	483	ARG	NE-CZ-NH1	7.47	124.04	120.30
1	A	350	ARG	NE-CZ-NH2	-6.29	117.16	120.30
1	C	629	ARG	NE-CZ-NH2	6.05	123.32	120.30
1	C	75	ARG	NE-CZ-NH1	5.79	123.19	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	6842	6682	6656	105	0
1	C	6865	6700	6672	101	0
2	B	645	604	604	12	1
2	D	645	604	604	14	0
3	A	11	8	7	0	0
3	C	11	8	7	0	0
4	B	16	0	0	1	0
4	D	16	0	0	1	0
5	A	1015	0	0	38	0
5	B	66	0	0	2	0
5	C	950	0	0	29	0
5	D	67	0	0	1	0
All	All	17149	14606	14550	228	1

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 8.

The worst 5 of 228 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:483:ARG:HD2	5:A:2946:HOH:O	1.76	0.86
1:A:735:HIS:HE1	5:A:2670:HOH:O	1.65	0.79
2:B:49:LYS:HG3	2:B:86:GLU:HA	1.65	0.79
1:C:684:VAL:HG11	1:C:689:PHE:HB3	1.70	0.74
1:C:820:VAL:O	1:C:825:GLY:HA3	1.87	0.73

All (1) 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)
2:B:34:GLU:OE2	2:B:34:GLU:OE2[3_655]	2.16	0.04

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	869/897 (97%)	829 (95%)	37 (4%)	3 (0%)	46	29
1	C	872/897 (97%)	833 (96%)	34 (4%)	5 (1%)	30	14
2	B	84/86 (98%)	83 (99%)	1 (1%)	0	100	100
2	D	84/86 (98%)	79 (94%)	5 (6%)	0	100	100
All	All	1909/1966 (97%)	1824 (96%)	77 (4%)	8 (0%)	39	23

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	392	SER
1	C	537	PHE
1	A	752	VAL
1	C	752	VAL
1	A	537	PHE

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	738/761 (97%)	725 (98%)	13 (2%)	66	54
1	C	740/761 (97%)	725 (98%)	15 (2%)	63	49
2	B	69/69 (100%)	65 (94%)	4 (6%)	25	9
2	D	69/69 (100%)	63 (91%)	6 (9%)	13	3
All	All	1616/1660 (97%)	1578 (98%)	38 (2%)	60	40

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	29	ASP
1	C	483	ARG
2	D	65	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	300	GLU
1	C	503[A]	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	812	ASN
2	B	53	ASN
1	C	789	ASN
1	A	789	ASN
1	C	565	ASN

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	$\# Z > 2$
3	4HP	A	1898	-	11,11,11	1.15	0	14,14,14	1.37	3 (21%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	SF4	B	87	2	0,12,12	0.00	-	0,24,24	0.00	-
4	SF4	B	88	2	0,12,12	0.00	-	0,24,24	0.00	-
3	4HP	C	1898	-	11,11,11	1.54	2 (18%)	14,14,14	1.48	3 (21%)
4	SF4	D	87	2	0,12,12	0.00	-	0,24,24	0.00	-
4	SF4	D	88	2	0,12,12	0.00	-	0,24,24	0.00	-

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
3	4HP	A	1898	-	-	0/4/4/4	0/1/1/1
4	SF4	B	87	2	-	0/0/48/48	0/6/5/5
4	SF4	B	88	2	-	0/0/48/48	0/6/5/5
3	4HP	C	1898	-	-	0/4/4/4	0/1/1/1
4	SF4	D	87	2	-	0/0/48/48	0/6/5/5
4	SF4	D	88	2	-	0/0/48/48	0/6/5/5

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1898	4HP	C3-C2	2.13	1.42	1.38
3	C	1898	4HP	C6-C5	2.24	1.42	1.38

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1898	4HP	C7-C1-C6	-2.76	116.72	120.86
3	A	1898	4HP	C7-C1-C6	-2.70	116.81	120.86
3	C	1898	4HP	C5-C6-C1	-2.68	117.36	121.04
3	A	1898	4HP	O1-C8-C7	-2.37	117.15	123.19
3	A	1898	4HP	O2-C8-C7	2.01	119.54	113.96

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	88	SF4	1	0
4	D	87	SF4	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 [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	869/897 (96%)	-0.36	8 (0%) 85 84	13, 28, 51, 131	3 (0%)
1	C	871/897 (97%)	-0.30	22 (2%) 61 56	13, 27, 58, 134	1 (0%)
2	B	86/86 (100%)	0.33	8 (9%) 11 8	29, 39, 62, 85	0
2	D	86/86 (100%)	0.60	11 (12%) 5 3	28, 45, 68, 86	0
All	All	1912/1966 (97%)	-0.26	49 (2%) 59 54	13, 29, 58, 134	4 (0%)

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	692	SER	7.3
1	A	689	PHE	6.6
1	C	693	THR	6.1
1	C	689	PHE	5.8
1	C	129	VAL	5.3

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, 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	LLDF	B-factors(\AA^2)	Q<0.9
3	4HP	A	1898	11/11	0.95	0.08	0.30	15,25,36,36	0
3	4HP	C	1898	11/11	0.96	0.07	-0.18	17,26,36,42	0
4	SF4	B	88	8/8	0.99	0.04	-1.39	28,31,33,34	0
4	SF4	B	87	8/8	0.99	0.06	-1.42	42,46,50,54	0
4	SF4	D	88	8/8	0.99	0.03	-1.60	31,34,37,39	0
4	SF4	D	87	8/8	0.98	0.04	-2.01	38,43,50,53	0

6.5 Other polymers [i](#)

There are no such residues in this entry.