



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 08:14 AM UTC

PDB ID : 5M21 / pdb_00005m21
Title : Crystal structure of hydroquinone 1,2-dioxygenase from *Sphingomonas* sp. TTNP3 with 4-hydroxybenzoate bound
Authors : Ferraroni, M.; Da Vela, S.; Scozzafava, A.; Kolvenbach, B.; Corvini, P.F.X.
Deposited on : 2016-10-11
Resolution : 1.99 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

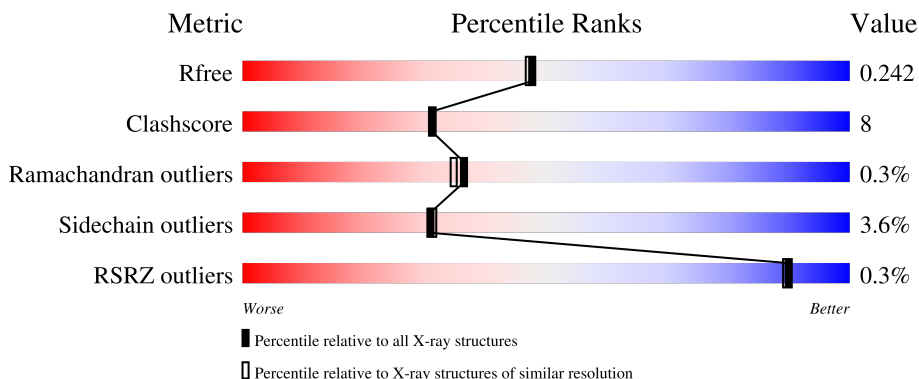
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.99 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}	180053	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)

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	170	 76% 20% ..
1	C	170	 81% 16% ..
1	E	170	 87% 11% ..
1	G	170	 73% 24% ..
2	B	341	 87% 7% 5% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	341	 76% 18% ..
2	F	341	 79% 14% ..
2	H	341	 83% 13% ..

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
4	PHB	D	402	-	-	X	-

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 16580 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 Hydroquinone dioxygenase small subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	166	1275	808	216	245	6	0	0	0
1	C	167	1284	813	217	248	6	0	0	0
1	E	167	1275	808	213	248	6	0	0	0
1	G	165	1251	796	209	240	6	0	0	0

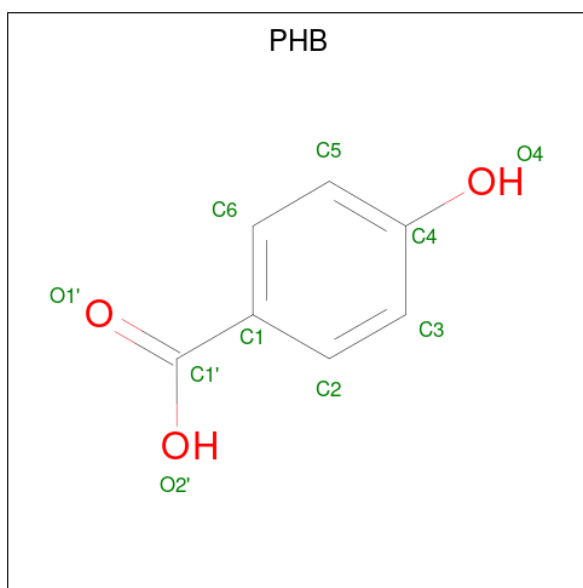
- Molecule 2 is a protein called Hydroquinone dioxygenase large subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	325	2562	1624	448	476	14	0	0	0
2	D	327	2572	1628	448	481	15	0	2	0
2	F	326	2572	1630	449	478	15	0	1	0
2	H	333	2604	1649	455	486	14	0	0	0

- Molecule 3 is FE (III) ION (CCD ID: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total 1	Fe 1	0	0
3	D	1	Total 1	Fe 1	0	0
3	F	1	Total 1	Fe 1	0	0
3	H	1	Total 1	Fe 1	0	0

- Molecule 4 is P-HYDROXYBENZOIC ACID (CCD ID: PHB) (formula: C₇H₆O₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C O 10 7 3	0	0
4	D	1	Total C O 10 7 3	0	0
4	F	1	Total C O 10 7 3	0	0
4	H	1	Total C O 10 7 3	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	97	Total O 97 97	0	0
5	B	213	Total O 213 213	0	0
5	C	117	Total O 117 117	0	0
5	D	209	Total O 209 209	0	0
5	E	89	Total O 89 89	0	0
5	F	177	Total O 177 177	0	0
5	G	50	Total O 50 50	0	0

Continued on next page...

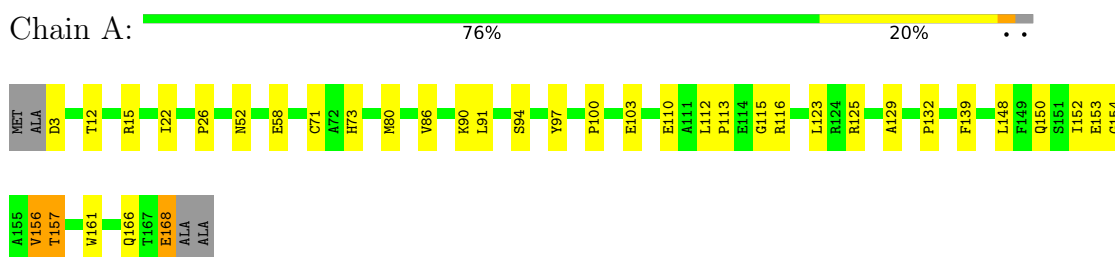
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	H	189	Total 189	O 189	0	0

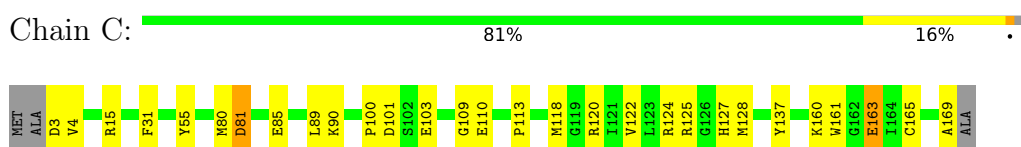
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

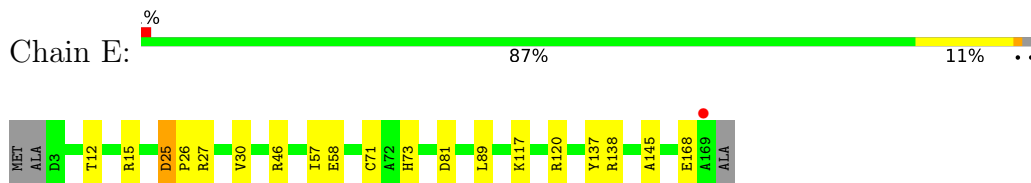
- Molecule 1: Hydroquinone dioxygenase small subunit



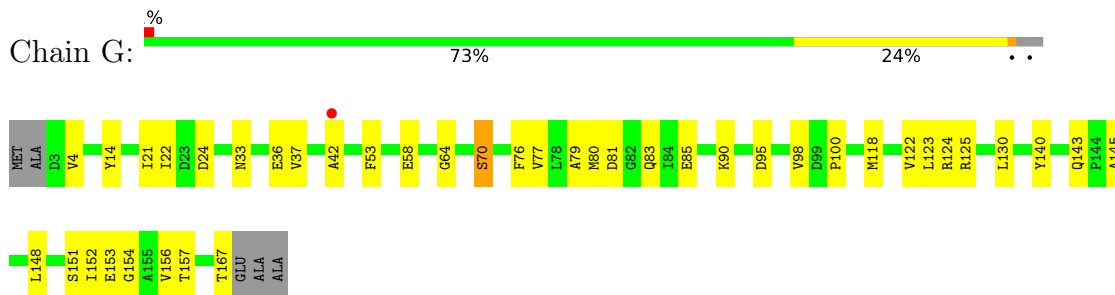
- Molecule 1: Hydroquinone dioxygenase small subunit



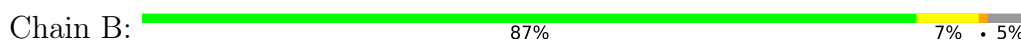
- Molecule 1: Hydroquinone dioxygenase small subunit

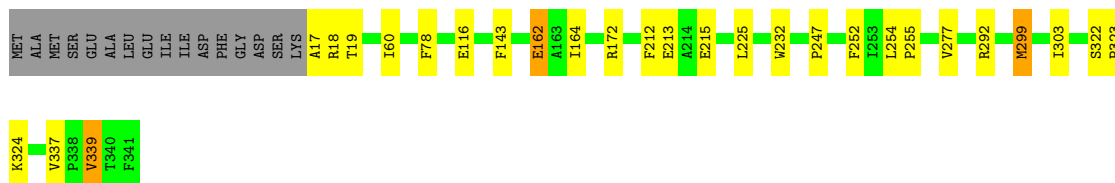


- Molecule 1: Hydroquinone dioxygenase small subunit



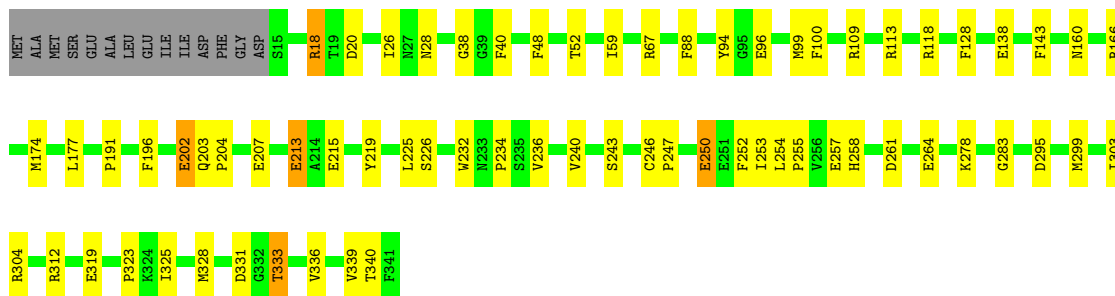
- Molecule 2: Hydroquinone dioxygenase large subunit





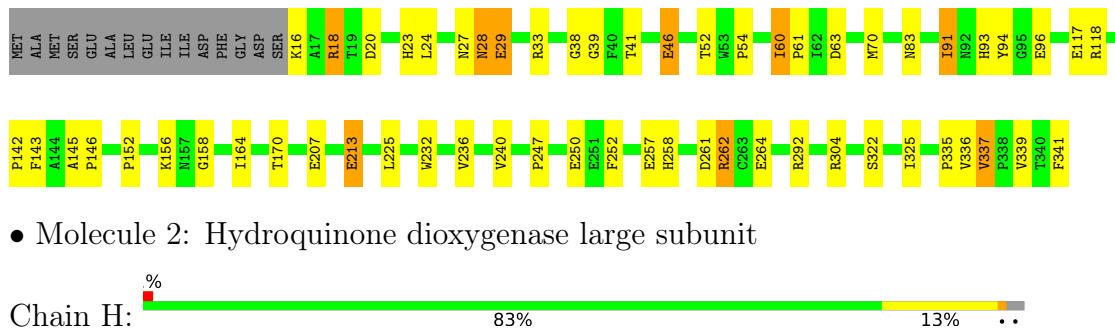
- Molecule 2: Hydroquinone dioxygenase large subunit

Chain D: 76% 18%



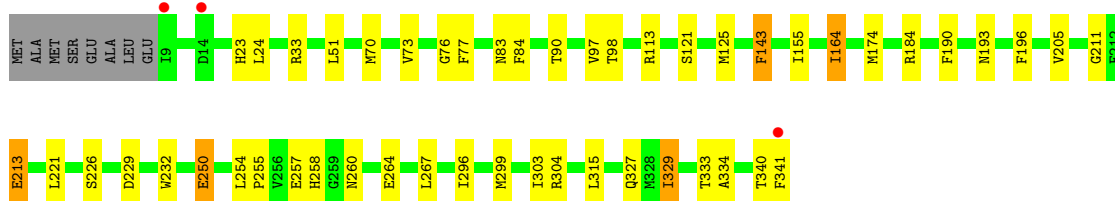
- Molecule 2: Hydroquinone dioxygenase large subunit

Chain F: 79% 14%



- Molecule 2: Hydroquinone dioxygenase large subunit

Chain H: 83% 13%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	88.78Å 124.87Å 92.37Å 90.00° 105.15° 90.00°	Depositor
Resolution (Å)	29.80 – 1.99 29.80 – 1.99	Depositor EDS
% Data completeness (in resolution range)	98.6 (29.80-1.99) 98.5 (29.80-1.99)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.52 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.179 , 0.242 0.180 , 0.242	Depositor DCC
R_{free} test set	6615 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	24.4	Xtrriage
Anisotropy	0.582	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 40.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	0.036 for l,-k,h	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	16580	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.91% 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: FE, PHB

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	1.05	1/1302 (0.1%)	1.15	6/1766 (0.3%)
1	C	1.03	0/1311	1.09	2/1778 (0.1%)
1	E	1.01	0/1302	1.09	3/1768 (0.2%)
1	G	0.91	0/1278	1.12	5/1735 (0.3%)
2	B	1.09	0/2635	1.15	8/3583 (0.2%)
2	D	1.09	1/2651 (0.0%)	1.09	9/3607 (0.2%)
2	F	1.02	0/2648	1.11	9/3599 (0.3%)
2	H	1.05	0/2677	1.11	4/3641 (0.1%)
All	All	1.05	2/15804 (0.0%)	1.11	46/21477 (0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	52	ASN	N-CA	-5.35	1.40	1.46
2	D	236	VAL	C-O	5.27	1.29	1.24

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	60	ILE	CA-C-N	-7.92	111.60	119.76
2	F	60	ILE	C-N-CA	-7.92	111.60	119.76
1	G	53	PHE	CB-CA-C	-7.00	108.48	116.54
1	A	22	ILE	N-CA-C	-6.91	105.61	111.56
1	G	83	GLN	N-CA-C	6.34	118.91	109.59

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	1275	0	1220	30	0
1	C	1284	0	1229	28	0
1	E	1275	0	1209	16	0
1	G	1251	0	1188	35	0
2	B	2562	0	2426	14	0
2	D	2572	0	2424	67	0
2	F	2572	0	2430	44	0
2	H	2604	0	2443	49	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
3	F	1	0	0	0	0
3	H	1	0	0	0	0
4	B	10	0	4	1	0
4	D	10	0	5	5	0
4	F	10	0	4	1	0
4	H	10	0	4	1	0
5	A	97	0	0	5	0
5	B	213	0	0	1	0
5	C	117	0	0	6	0
5	D	209	0	0	15	0
5	E	89	0	0	3	0
5	F	177	0	0	15	0
5	G	50	0	0	5	0
5	H	189	0	0	14	0
All	All	16580	0	14586	247	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 8.

The worst 5 of 247 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:G:76:PHE:CD1	2:H:174:MET:CE	2.11	1.34
1:C:128:MET:HG2	2:D:174:MET:CE	1.60	1.30
2:F:240:VAL:HB	5:F:658:HOH:O	1.13	1.30

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:76:PHE:CE1	2:H:174:MET:HE3	1.68	1.26
2:H:97:VAL:HB	5:H:648:HOH:O	1.07	1.25

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	164/170 (96%)	158 (96%)	6 (4%)	0	100	100
1	C	165/170 (97%)	154 (93%)	10 (6%)	1 (1%)	21	17
1	E	165/170 (97%)	156 (94%)	8 (5%)	1 (1%)	21	17
1	G	163/170 (96%)	149 (91%)	14 (9%)	0	100	100
2	B	323/341 (95%)	307 (95%)	15 (5%)	1 (0%)	36	35
2	D	327/341 (96%)	310 (95%)	16 (5%)	1 (0%)	36	35
2	F	325/341 (95%)	309 (95%)	15 (5%)	1 (0%)	36	35
2	H	331/341 (97%)	312 (94%)	19 (6%)	0	100	100
All	All	1963/2044 (96%)	1855 (94%)	103 (5%)	5 (0%)	36	35

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	252	PHE
2	F	252	PHE
1	C	81	ASP
2	B	252	PHE
1	E	30	VAL

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	128/131 (98%)	125 (98%)	3 (2%)	44	49
1	C	129/131 (98%)	123 (95%)	6 (5%)	23	22
1	E	127/131 (97%)	124 (98%)	3 (2%)	43	47
1	G	124/131 (95%)	120 (97%)	4 (3%)	34	35
2	B	268/284 (94%)	260 (97%)	8 (3%)	36	38
2	D	270/284 (95%)	258 (96%)	12 (4%)	25	24
2	F	268/284 (94%)	256 (96%)	12 (4%)	24	23
2	H	269/284 (95%)	260 (97%)	9 (3%)	33	34
All	All	1583/1660 (95%)	1526 (96%)	57 (4%)	31	31

5 of 57 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	340	THR
2	H	329	ILE
2	F	29	GLU
2	H	327	GLN
2	H	143	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	40	ASN
1	C	93	ASN
2	H	306	GLN
2	D	160	ASN
2	B	83	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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	PHB	H	402	3	10,10,10	0.62	0	13,13,13	0.56	0
4	PHB	F	402	3	10,10,10	0.70	1 (10%)	13,13,13	0.66	0
4	PHB	D	402	3	10,10,10	0.67	0	13,13,13	0.53	0
4	PHB	B	402	3	10,10,10	0.69	0	13,13,13	0.55	0

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	PHB	H	402	3	-	0/4/4/4	0/1/1/1
4	PHB	F	402	3	-	1/4/4/4	0/1/1/1
4	PHB	D	402	3	-	0/4/4/4	0/1/1/1
4	PHB	B	402	3	-	0/4/4/4	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	402	PHB	O2'-C1'	-2.05	1.24	1.30

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	F	402	PHB	C6-C1-C1'-O1'

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	402	PHB	1	0
4	F	402	PHB	1	0
4	D	402	PHB	5	0
4	B	402	PHB	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	166/170 (97%)	-0.36	0 100 100	14, 25, 40, 52	0
1	C	167/170 (98%)	-0.36	0 100 100	16, 26, 39, 55	0
1	E	167/170 (98%)	-0.26	1 (0%) 85 85	17, 27, 44, 61	0
1	G	165/170 (97%)	0.25	1 (0%) 85 85	24, 37, 49, 59	0
2	B	325/341 (95%)	-0.50	0 100 100	15, 21, 34, 54	0
2	D	327/341 (95%)	-0.44	0 100 100	12, 22, 41, 69	2 (0%)
2	F	326/341 (95%)	-0.18	0 100 100	15, 26, 44, 67	1 (0%)
2	H	333/341 (97%)	-0.23	3 (0%) 81 80	15, 26, 50, 70	0
All	All	1976/2044 (96%)	-0.29	5 (0%) 90 89	12, 25, 45, 70	3 (0%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	9	ILE	4.0
2	H	14	ASP	3.0
1	G	42	ALA	2.7
1	E	169	ALA	2.6
2	H	341	PHE	2.2

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 oligosaccharides 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
4	PHB	D	402	10/10	0.87	0.14	23,28,44,46	0
4	PHB	H	402	10/10	0.92	0.10	27,32,36,40	0
4	PHB	F	402	10/10	0.94	0.08	29,33,55,58	0
4	PHB	B	402	10/10	0.94	0.08	24,26,28,36	0
3	FE	H	401	1/1	0.99	0.03	35,35,35,35	0
3	FE	F	401	1/1	1.00	0.03	35,35,35,35	0
3	FE	B	401	1/1	1.00	0.04	28,28,28,28	0
3	FE	D	401	1/1	1.00	0.03	32,32,32,32	0

6.5 Other polymers [i](#)

There are no such residues in this entry.