



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2017 – 12:06 PM EST

PDB ID : 5B2V  
Title : Crystal Structure of P450BM3 with N-perfluorohexanoyl-L-tryptophan  
Authors : Cong, Z.; Shoji, O.; Kasai, C.; Sugimoto, H.; Shiro, Y.; Watanabe, Y.  
Deposited on : 2016-02-03  
Resolution : 2.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20028442  
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) : rb-20028442

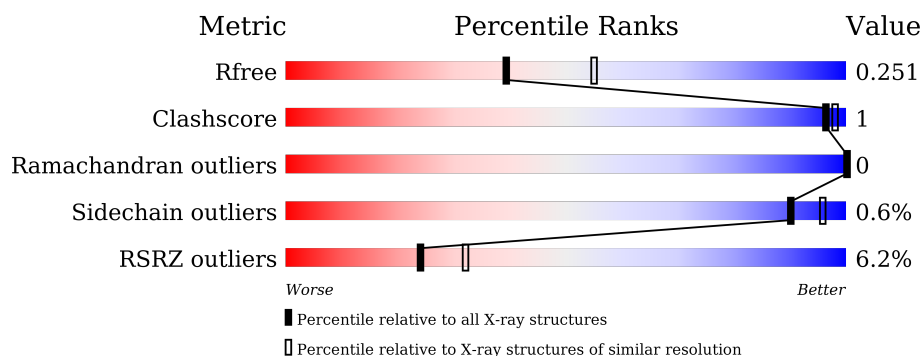
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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	456	<div> <div>6%</div> <div> <div></div> <div>96%</div> <div>..</div> </div> </div>
1	B	456	<div> <div>6%</div> <div> <div></div> <div>97%</div> <div>..</div> </div> </div>

## 2 Entry composition [i](#)

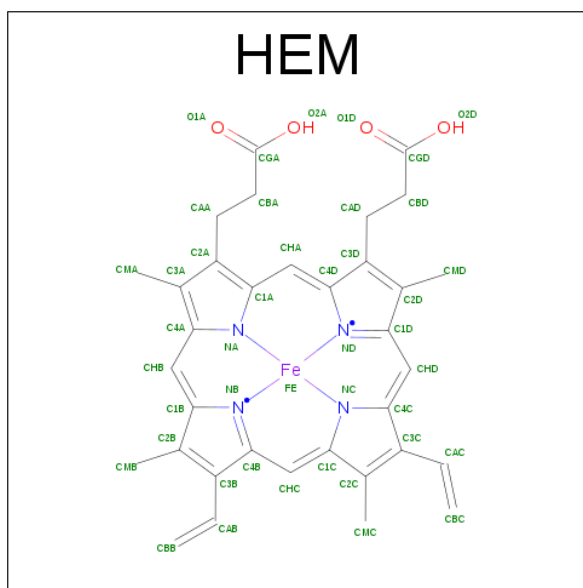
There are 4 unique types of molecules in this entry. The entry contains 7600 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Bifunctional cytochrome P450/NADPH-P450 reductase.

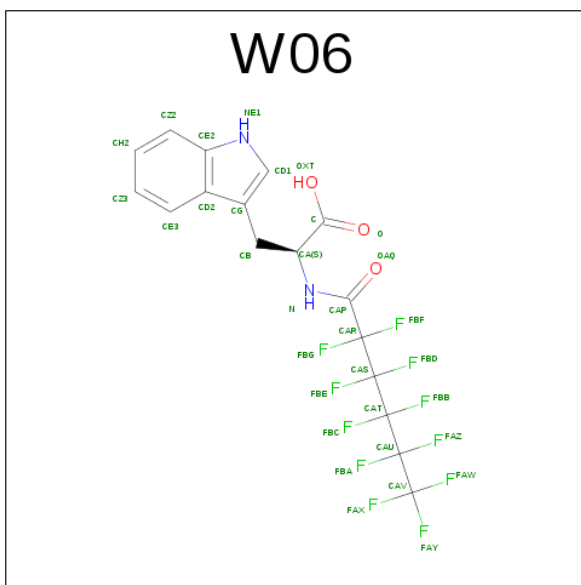
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	453	Total	C	N	O	S	0	1	0
			3664	2340	622	685	17			
1	B	453	Total	C	N	O	S	0	0	0
			3655	2335	621	682	17			

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 3 is (2 {S})-3-(1 {H}-indol-3-yl)-2-[2,2,3,3,4,4,5,5,6,6,6-undecakis(fluoranyl)hexanoylamino]propanoic acid (three-letter code: W06) (formula:  $C_{17}H_{11}F_{11}N_2O_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 33	C 17	F 11	N 2	O 3	0	0
3	B	1	Total 33	C 17	F 11	N 2	O 3	0	0

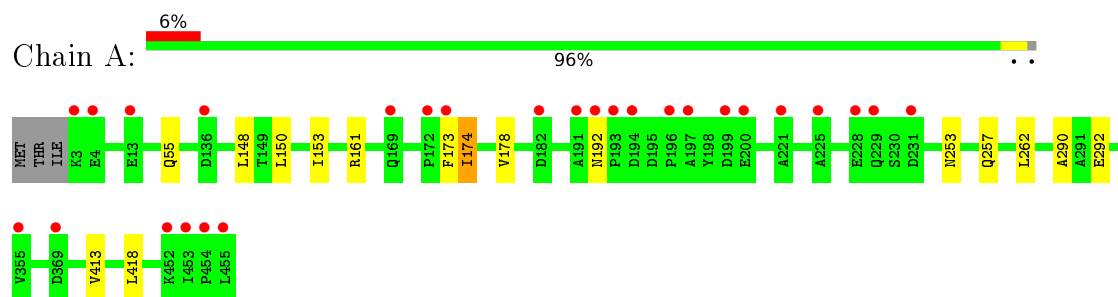
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	70	Total O 70 70	0	0
4	B	59	Total O 59 59	0	0

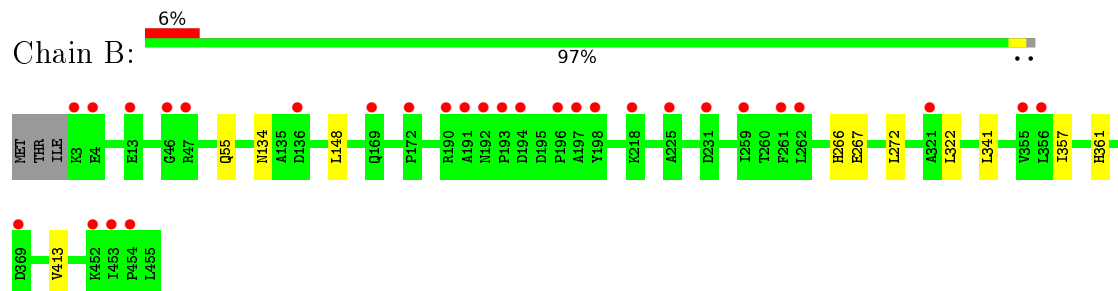
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 ( $\text{RSRZ} > 2$ ). 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: Bifunctional cytochrome P450/NADPH-P450 reductase



- Molecule 1: Bifunctional cytochrome P450/NADPH-P450 reductase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.88 Å 147.49 Å 64.25 Å 90.00° 98.23° 90.00°	Depositor
Resolution (Å)	19.91 – 2.30 19.91 – 2.30	Depositor EDS
% Data completeness (in resolution range)	88.9 (19.91-2.30) 89.0 (19.91-2.30)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.80 (at 2.30 Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.215 , 0.255 0.216 , 0.251	Depositor DCC
$R_{free}$ test set	2176 reflections (5.36%)	DCC
Wilson B-factor (Å <sup>2</sup> )	42.4	Xtriage
Anisotropy	0.082	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 29.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7600	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.35% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: HEM, W06

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	0.27	0/3749	0.45	0/5068
1	B	0.27	0/3740	0.46	0/5056
All	All	0.27	0/7489	0.45	0/10124

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

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	3664	0	3629	9	0
1	B	3655	0	3624	5	0
2	A	43	0	30	3	0
2	B	43	0	30	4	0
3	A	33	0	0	1	0
3	B	33	0	0	0	0
4	A	70	0	0	0	0
4	B	59	0	0	0	0
All	All	7600	0	7313	21	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 1.

All (21) 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)
2:A:501:HEM:HMC2	2:A:501:HEM:HBC2	1.79	0.65
2:A:501:HEM:HMB2	2:A:501:HEM:HBB2	1.78	0.63
2:B:501:HEM:HBC2	2:B:501:HEM:HMC2	1.81	0.62
2:B:501:HEM:HBB2	2:B:501:HEM:HMB2	1.82	0.62
1:A:253:ASN:HD21	1:A:257:GLN:HE21	1.53	0.55
1:A:161:ARG:H	1:B:134:ASN:HD21	1.57	0.52
1:A:148:LEU:HD21	1:A:413:VAL:HG21	1.93	0.49
1:A:174:ILE:HD12	1:A:178:VAL:HG23	1.94	0.49
1:A:150:LEU:HD11	1:A:174:ILE:HD13	1.97	0.47
1:A:290:ALA:HB1	1:A:418:LEU:HD13	1.98	0.45
1:B:272:LEU:HD13	1:B:322:LEU:HG	1.99	0.45
2:B:501:HEM:HBC2	2:B:501:HEM:CMC	2.47	0.45
1:B:357:ILE:HG22	1:B:361:HIS:CE1	2.52	0.44
1:A:173:PHE:CE1	1:A:262:LEU:HD13	2.53	0.44
1:A:292[A]:GLU:C	1:A:292[A]:GLU:CD	2.77	0.43
2:A:501:HEM:HBC2	2:A:501:HEM:CMC	2.46	0.43
1:B:266:HIS:CG	1:B:267:GLU:N	2.88	0.42
3:A:502:W06:CD1	3:A:502:W06:C	2.99	0.41
2:B:501:HEM:HBB2	2:B:501:HEM:CMB	2.50	0.41
1:B:148:LEU:HD21	1:B:413:VAL:HG21	2.01	0.41
1:A:153:ILE:HG23	1:A:262:LEU:HD23	2.02	0.40

There are no symmetry-related clashes.

## 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	452/456 (99%)	439 (97%)	13 (3%)	0	100	100
1	B	451/456 (99%)	437 (97%)	14 (3%)	0	100	100
All	All	903/912 (99%)	876 (97%)	27 (3%)	0	100	100



There are no Ramachandran outliers to report.

### 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	398/400 (100%)	395 (99%)	3 (1%)	86	94
1	B	397/400 (99%)	395 (100%)	2 (0%)	92	97
All	All	795/800 (99%)	790 (99%)	5 (1%)	90	96

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	55	GLN
1	A	174	ILE
1	A	192	ASN
1	B	55	GLN
1	B	341	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	7	GLN
1	A	55	GLN
1	A	95	ASN
1	A	128	GLN
1	A	134	ASN
1	A	186	ASN
1	A	201	ASN
1	A	253	ASN
1	B	21	ASN
1	B	55	GLN
1	B	95	ASN
1	B	134	ASN
1	B	186	ASN
1	B	189	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	201	ASN
1	B	387	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

4 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	HEM	A	501	1	24,50,50	0.79	1 (4%)	16,82,82	1.16	1 (6%)
3	W06	A	502	-	29,34,34	1.42	3 (10%)	40,56,56	0.90	2 (5%)
2	HEM	B	501	1	24,50,50	0.75	1 (4%)	16,82,82	1.10	1 (6%)
3	W06	B	502	-	29,34,34	2.17	4 (13%)	40,56,56	1.36	7 (17%)

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	HEM	A	501	1	-	0/6/54/54	0/0/8/8
3	W06	A	502	-	-	0/49/54/54	0/2/2/2
2	HEM	B	501	1	-	0/6/54/54	0/0/8/8
3	W06	B	502	-	-	0/49/54/54	0/2/2/2

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	502	W06	FBD-CAS	-8.83	0.94	1.35
3	B	502	W06	CZ2-CE2	-4.37	1.33	1.41
3	A	502	W06	CZ2-CE2	-4.31	1.34	1.41
3	B	502	W06	CE3-CD2	-3.92	1.34	1.42
3	A	502	W06	CE3-CD2	-3.78	1.34	1.42
3	B	502	W06	CD2-CE2	-2.98	1.34	1.42
3	A	502	W06	CD2-CE2	-2.95	1.34	1.42
2	A	501	HEM	C1B-NB	-2.45	1.33	1.36
2	B	501	HEM	C1B-NB	-2.21	1.33	1.36

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	502	W06	FBE-CAS-CAR	-3.19	100.78	108.48
3	B	502	W06	FBD-CAS-CAR	-2.94	101.36	108.48
3	B	502	W06	CB-CG-CD1	-2.92	124.36	127.97
3	A	502	W06	CB-CG-CD1	-2.83	124.47	127.97
3	B	502	W06	CB-CA-N	-2.66	103.23	108.66
3	B	502	W06	CAR-CAS-CAT	-2.61	107.20	116.87
2	A	501	HEM	CBA-CAA-C2A	-2.47	108.15	112.49
3	B	502	W06	FBG-CAR-CAP	-2.46	107.01	109.95
3	A	502	W06	CB-CA-N	-2.24	104.09	108.66
2	B	501	HEM	CBA-CAA-C2A	-2.09	108.81	112.49
3	B	502	W06	FBD-CAS-FBE	2.17	116.70	107.15

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	HEM	3	0
3	A	502	W06	1	0
2	B	501	HEM	4	0

## 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	453/456 (99%)	0.27	27 (5%) 25 33	31, 43, 76, 110	0
1	B	453/456 (99%)	0.24	29 (6%) 23 31	30, 43, 69, 113	0
All	All	906/912 (99%)	0.25	56 (6%) 24 32	30, 43, 75, 113	0

All (56) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	191	ALA	7.0
1	B	193	PRO	6.9
1	A	3	LYS	5.9
1	B	196	PRO	5.6
1	A	194	ASP	4.8
1	A	4	GLU	4.7
1	A	454	PRO	4.7
1	B	3	LYS	4.6
1	A	225	ALA	4.5
1	A	196	PRO	4.4
1	A	193	PRO	4.3
1	A	191	ALA	4.1
1	B	453	ILE	4.1
1	A	192	ASN	4.0
1	A	453	ILE	3.9
1	B	194	ASP	3.7
1	A	169	GLN	3.5
1	B	197	ALA	3.5
1	B	225	ALA	3.2
1	A	199	ASP	3.2
1	B	46	GLY	3.2
1	A	136	ASP	3.1
1	A	173	PHE	3.0
1	B	192	ASN	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	4	GLU	2.9
1	A	229	GLN	2.8
1	A	355	VAL	2.7
1	A	231	ASP	2.6
1	A	200	GLU	2.6
1	B	190	ARG	2.6
1	B	452	LYS	2.6
1	B	355	VAL	2.6
1	B	454	PRO	2.5
1	B	169	GLN	2.5
1	B	172	PRO	2.5
1	A	182	ASP	2.5
1	A	13	GLU	2.4
1	A	197	ALA	2.4
1	B	198	TYR	2.4
1	B	262	LEU	2.3
1	B	356	LEU	2.3
1	B	261	PHE	2.3
1	B	136	ASP	2.3
1	A	452	LYS	2.3
1	A	221	ALA	2.3
1	B	321	ALA	2.2
1	B	259	ILE	2.2
1	A	455	LEU	2.2
1	A	228	GLU	2.2
1	B	47	ARG	2.2
1	B	369	ASP	2.2
1	B	13	GLU	2.1
1	B	218	LYS	2.1
1	A	369	ASP	2.1
1	A	172	PRO	2.1
1	B	231	ASP	2.0

## 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(Å <sup>2</sup> )	Q<0.9
3	W06	A	502	33/33	0.86	0.17	0.18	60,68,71,72	0
3	W06	B	502	33/33	0.89	0.16	0.09	55,59,63,63	0
2	HEM	B	501	43/43	0.97	0.10	-0.63	28,30,32,33	0
2	HEM	A	501	43/43	0.96	0.11	-0.81	33,35,36,38	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.