wwPDB X-ray Structure Validation Summary Report (i)

Dec 12, 2023 – 04:33 pm GMT

PDB ID	:	2XSO
Title	:	CRYSTAL STRUCTURE OF P4 VARIANT OF BIPHENYL DIOXYGENASE FROM BURKHOLDERIA XENOVORANS LB400
Authors	:	Kumar, P.; Bolin, J.T.
Deposited on	:	2010-09-29
Resolution	:	2.20 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 (i) 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 (1)) were used in the production of this report:

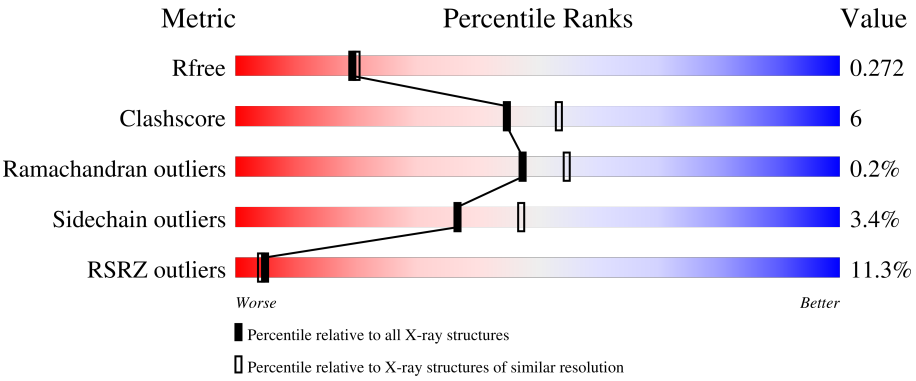
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 <=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	459	81% 13% • 6%
1	C	459	83% 10% • 6%
1	E	459	80% 14% • 6%
1	G	459	80% 13% • 6%
1	I	459	9% 76% 18% • 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	459	4% 78% 14% 6%
1	M	459	11% 80% 14% 6%
1	O	459	6% 79% 14% 6%
1	Q	459	14% 78% 15% 6%
1	S	459	10% 75% 18% 6%
1	U	459	42% 74% 19% 6%
1	W	459	61% 76% 17% 6%
2	B	188	% 87% 10%
2	D	188	% 80% 16%
2	F	188	% 83% 14%
2	H	188	2% 85% 11%
2	J	188	4% 83% 14%
2	L	188	3% 86% 11%
2	N	188	2% 79% 16%
2	P	188	5% 83% 13%
2	R	188	2% 83% 13%
2	T	188	% 82% 14%
2	V	188	5% 85% 11%
2	X	188	16% 81% 14%

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
3	FES	K	900	-	-	X	-
3	FES	O	900	-	-	X	-
3	FES	S	900	-	-	X	-
3	FES	W	900	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 61911 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 BIPHENYL DIOXYGENASE SUBUNIT ALPHA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	433	Total	C	N	O	S	0	0	0
			3427	2179	602	622	24			
1	C	433	Total	C	N	O	S	0	0	0
			3427	2179	602	622	24			
1	E	433	Total	C	N	O	S	0	0	0
			3427	2179	602	622	24			
1	G	433	Total	C	N	O	S	0	0	0
			3427	2179	602	622	24			
1	I	433	Total	C	N	O	S	0	0	0
			3427	2179	602	622	24			
1	K	433	Total	C	N	O	S	0	0	0
			3427	2179	602	622	24			
1	M	433	Total	C	N	O	S	0	0	0
			3427	2179	602	622	24			
1	O	433	Total	C	N	O	S	0	0	0
			3427	2179	602	622	24			
1	Q	433	Total	C	N	O	S	0	0	0
			3427	2179	602	622	24			
1	S	433	Total	C	N	O	S	0	0	0
			3427	2179	602	622	24			
1	U	433	Total	C	N	O	S	0	0	0
			3427	2179	602	622	24			
1	W	433	Total	C	N	O	S	0	0	0
			3427	2179	602	622	24			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	335	ALA	THR	engineered mutation	UNP P37333
A	336	MET	PHE	engineered mutation	UNP P37333
C	335	ALA	THR	engineered mutation	UNP P37333
C	336	MET	PHE	engineered mutation	UNP P37333
E	335	ALA	THR	engineered mutation	UNP P37333

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	336	MET	PHE	engineered mutation	UNP P37333
G	335	ALA	THR	engineered mutation	UNP P37333
G	336	MET	PHE	engineered mutation	UNP P37333
I	335	ALA	THR	engineered mutation	UNP P37333
I	336	MET	PHE	engineered mutation	UNP P37333
K	335	ALA	THR	engineered mutation	UNP P37333
K	336	MET	PHE	engineered mutation	UNP P37333
M	335	ALA	THR	engineered mutation	UNP P37333
M	336	MET	PHE	engineered mutation	UNP P37333
O	335	ALA	THR	engineered mutation	UNP P37333
O	336	MET	PHE	engineered mutation	UNP P37333
Q	335	ALA	THR	engineered mutation	UNP P37333
Q	336	MET	PHE	engineered mutation	UNP P37333
S	335	ALA	THR	engineered mutation	UNP P37333
S	336	MET	PHE	engineered mutation	UNP P37333
U	335	ALA	THR	engineered mutation	UNP P37333
U	336	MET	PHE	engineered mutation	UNP P37333
W	335	ALA	THR	engineered mutation	UNP P37333
W	336	MET	PHE	engineered mutation	UNP P37333

- Molecule 2 is a protein called BIPHENYL DIOXYGENASE SUBUNIT BETA.

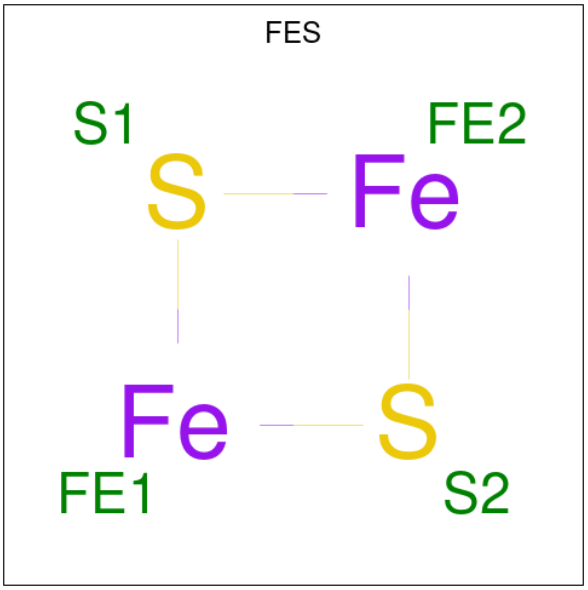
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	183	Total	C	N	O	S	0	0	0
			1524	968	270	282	4			
2	D	183	Total	C	N	O	S	0	0	0
			1524	968	270	282	4			
2	F	183	Total	C	N	O	S	0	0	0
			1524	968	270	282	4			
2	H	183	Total	C	N	O	S	0	0	0
			1524	968	270	282	4			
2	J	183	Total	C	N	O	S	0	0	0
			1524	968	270	282	4			
2	L	183	Total	C	N	O	S	0	0	0
			1524	968	270	282	4			
2	N	181	Total	C	N	O	S	0	0	0
			1507	957	266	280	4			
2	P	183	Total	C	N	O	S	0	0	0
			1524	968	270	282	4			
2	R	181	Total	C	N	O	S	0	0	0
			1507	957	266	280	4			
2	T	181	Total	C	N	O	S	0	0	0
			1507	957	266	280	4			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	V	181	Total	C	N	O	S	0	0	0
			1507	957	266	280	4			
2	X	181	Total	C	N	O	S	0	0	0
			1507	957	266	280	4			

- Molecule 3 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe_2S_2).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	Fe	S		0	0
			4	2	2			
3	C	1	Total	Fe	S		0	0
			4	2	2			
3	E	1	Total	Fe	S		0	0
			4	2	2			
3	G	1	Total	Fe	S		0	0
			4	2	2			
3	I	1	Total	Fe	S		0	0
			4	2	2			
3	K	1	Total	Fe	S		0	0
			4	2	2			
3	M	1	Total	Fe	S		0	0
			4	2	2			
3	O	1	Total	Fe	S		0	0
			4	2	2			
3	Q	1	Total	Fe	S		0	0
			4	2	2			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	S	1	Total Fe S 4 2 2	0	0
3	U	1	Total Fe S 4 2 2	0	0
3	W	1	Total Fe S 4 2 2	0	0

- Molecule 4 is FE (II) ION (three-letter code: FE2) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Fe 1 1	0	0
4	C	1	Total Fe 1 1	0	0
4	E	1	Total Fe 1 1	0	0
4	G	1	Total Fe 1 1	0	0
4	I	1	Total Fe 1 1	0	0
4	K	1	Total Fe 1 1	0	0
4	M	1	Total Fe 1 1	0	0
4	O	1	Total Fe 1 1	0	0
4	Q	1	Total Fe 1 1	0	0
4	S	1	Total Fe 1 1	0	0
4	U	1	Total Fe 1 1	0	0
4	W	1	Total Fe 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	185	Total O 185 185	0	0
5	B	122	Total O 122 122	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	147	Total O 147 147	0	0
5	D	113	Total O 113 113	0	0
5	E	210	Total O 210 210	0	0
5	F	131	Total O 131 131	0	0
5	G	181	Total O 181 181	0	0
5	H	122	Total O 122 122	0	0
5	I	88	Total O 88 88	0	0
5	J	48	Total O 48 48	0	0
5	K	123	Total O 123 123	0	0
5	L	88	Total O 88 88	0	0
5	M	96	Total O 96 96	0	0
5	N	76	Total O 76 76	0	0
5	O	106	Total O 106 106	0	0
5	P	38	Total O 38 38	0	0
5	Q	90	Total O 90 90	0	0
5	R	89	Total O 89 89	0	0
5	S	116	Total O 116 116	0	0
5	T	69	Total O 69 69	0	0
5	U	115	Total O 115 115	0	0
5	V	45	Total O 45 45	0	0
5	W	89	Total O 89 89	0	0

Continued on next page...

Continued from previous page...

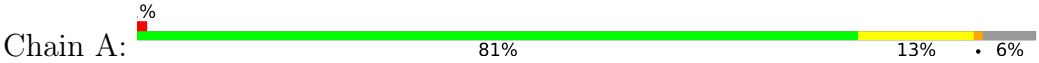
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	X	37	37	37	0	0

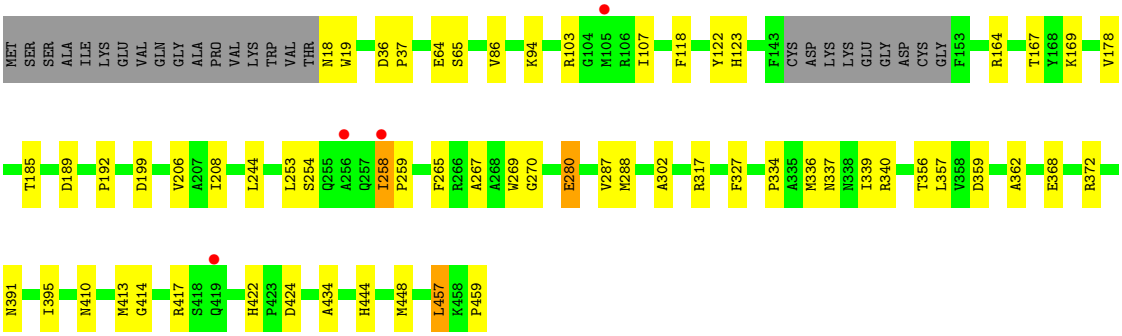
3 Residue-property plots (i)

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: BIPHENYL DIOXYGENASE SUBUNIT ALPHA

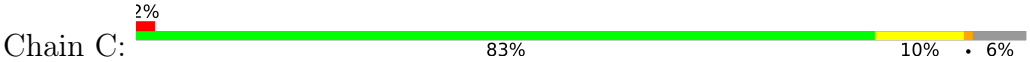
Chain A:

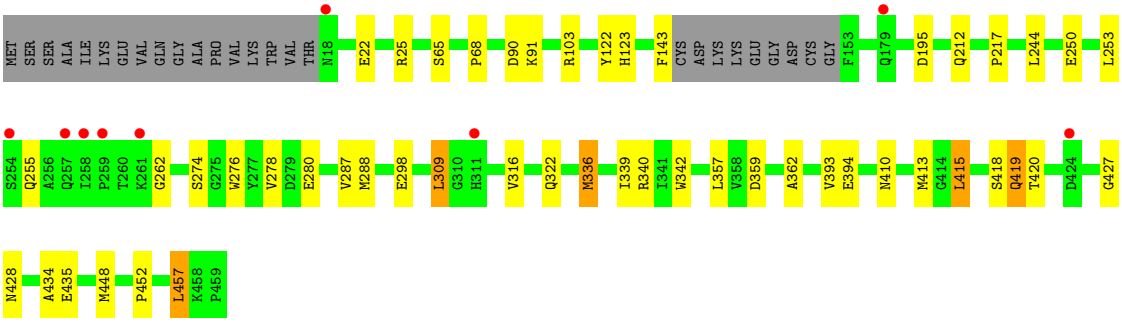




• Molecule 1: BIPHENYL DIOXYGENASE SUBUNIT ALPHA

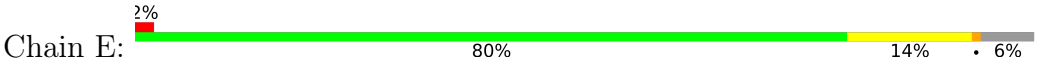
Chain C:

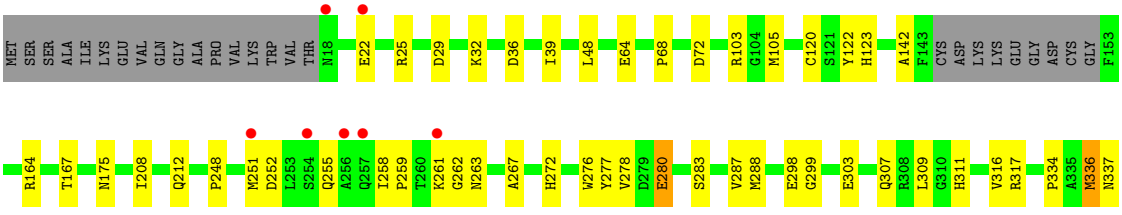


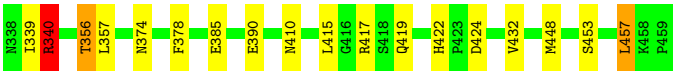


• Molecule 1: BIPHENYL DIOXYGENASE SUBUNIT ALPHA

Chain E:

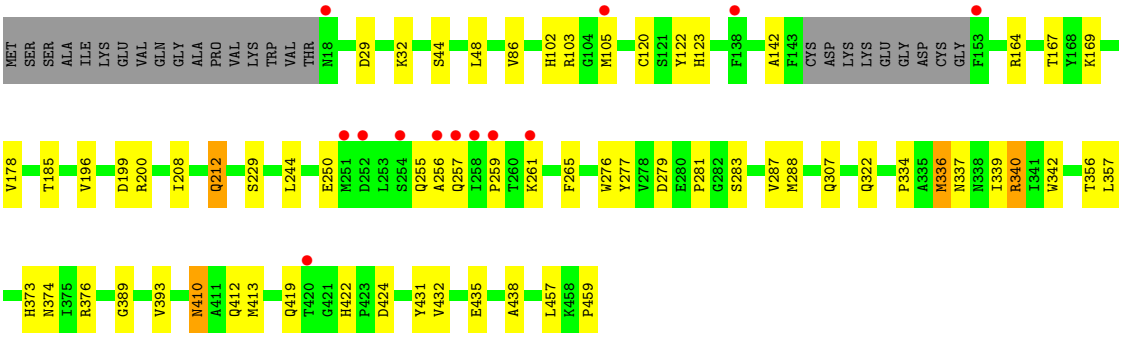






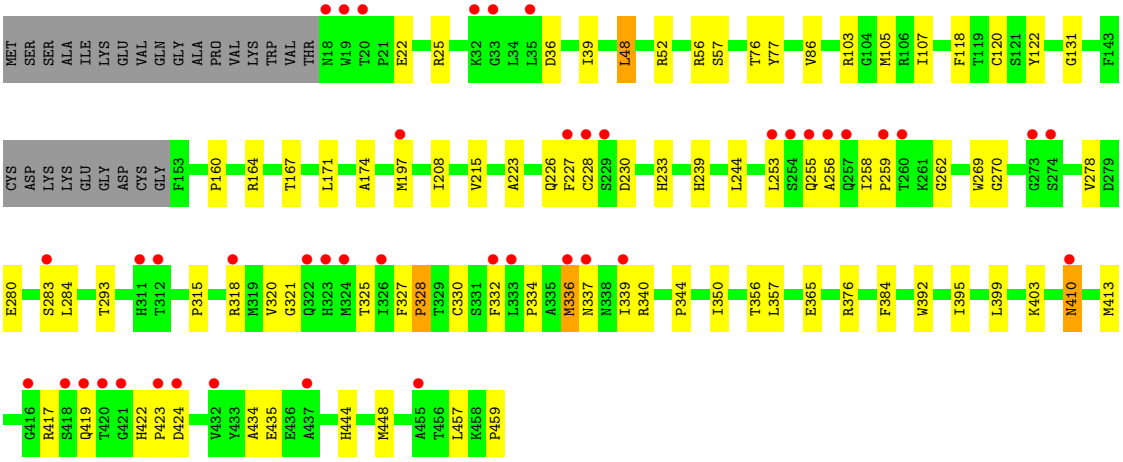
• Molecule 1: BIPHENYL DIOXYGENASE SUBUNIT ALPHA

Chain G:



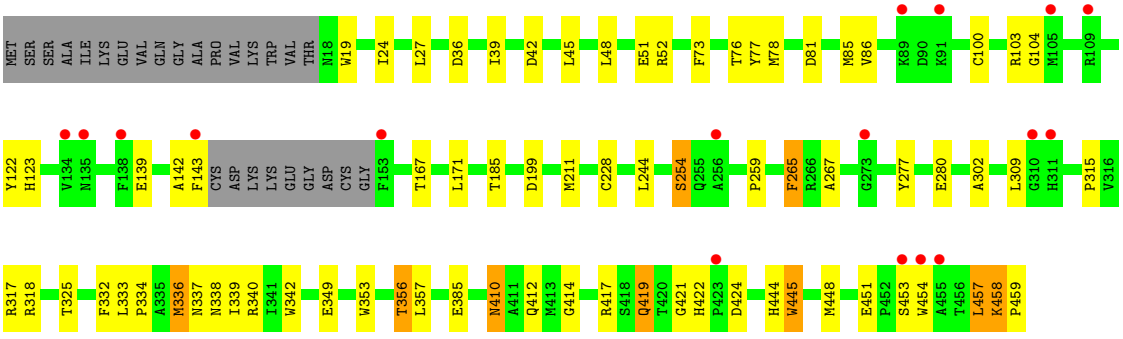
• Molecule 1: BIPHENYL DIOXYGENASE SUBUNIT ALPHA

Chain I:

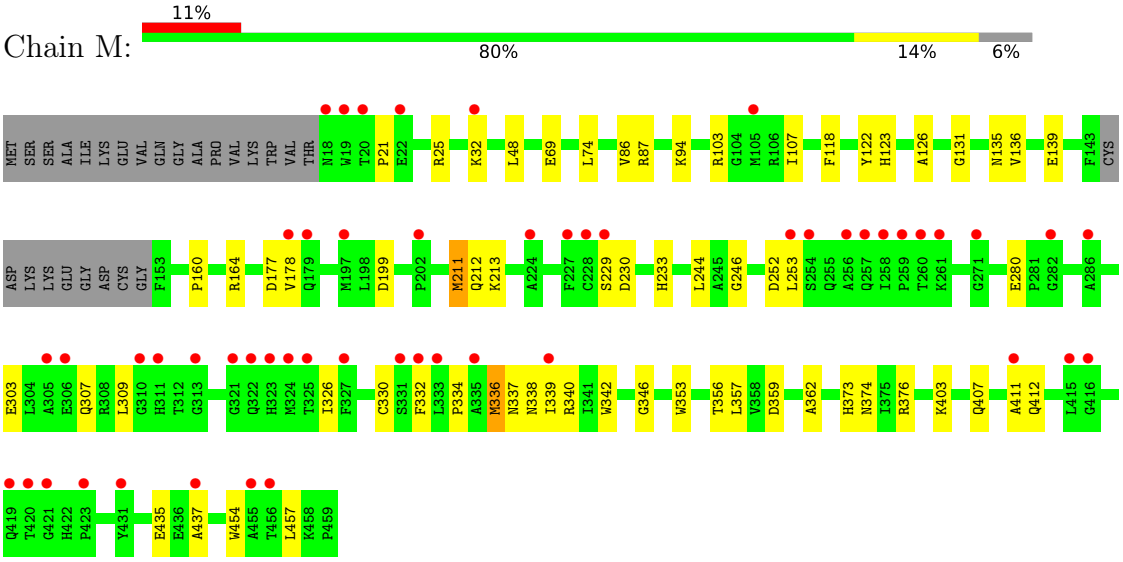


• Molecule 1: BIPHENYL DIOXYGENASE SUBUNIT ALPHA

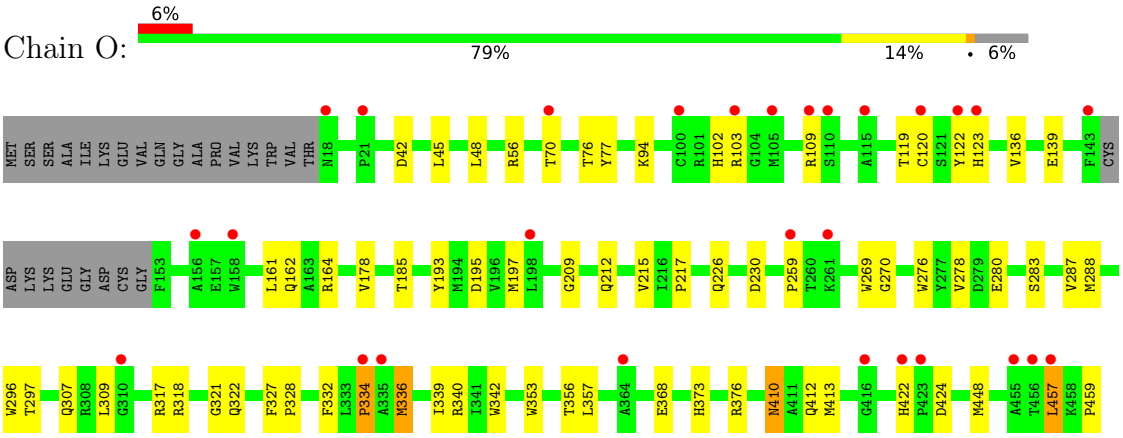
Chain K:



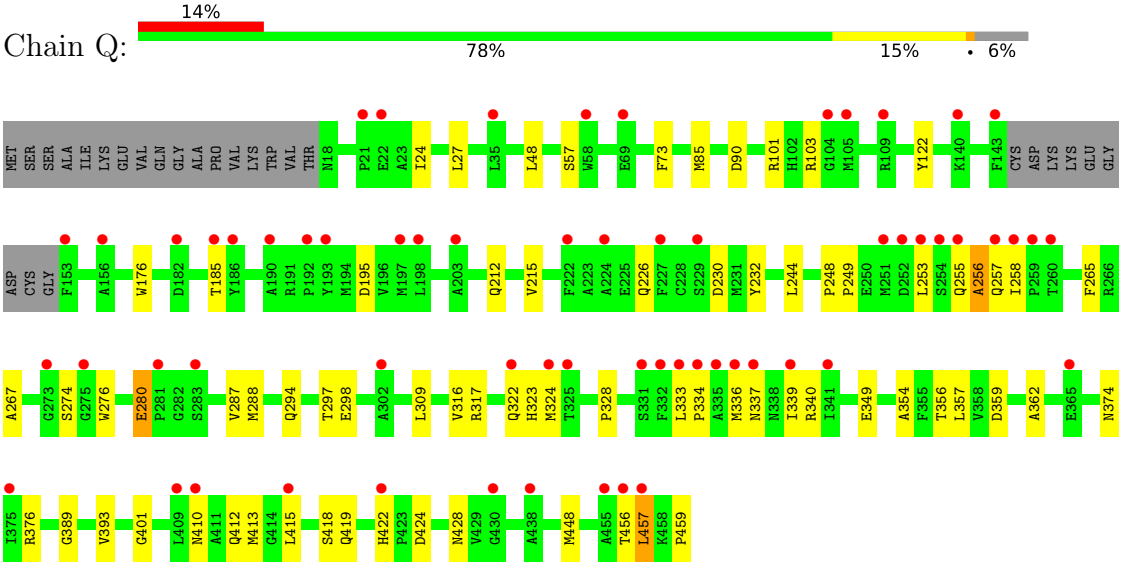
• Molecule 1: BIPHENYL DIOXYGENASE SUBUNIT ALPHA



• Molecule 1: BIPHENYL DIOXYGENASE SUBUNIT ALPHA

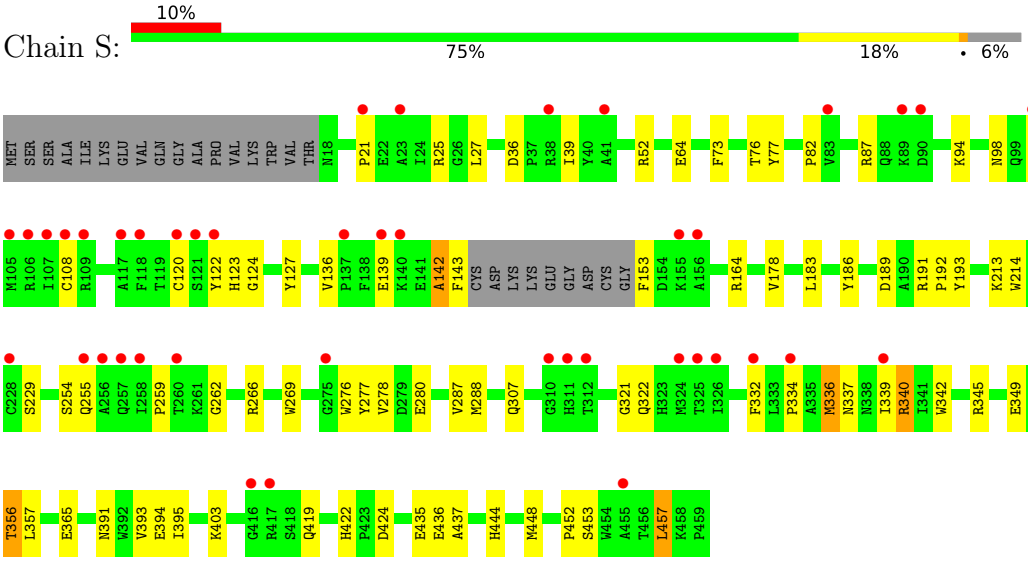


• Molecule 1: BIPHENYL DIOXYGENASE SUBUNIT ALPHA



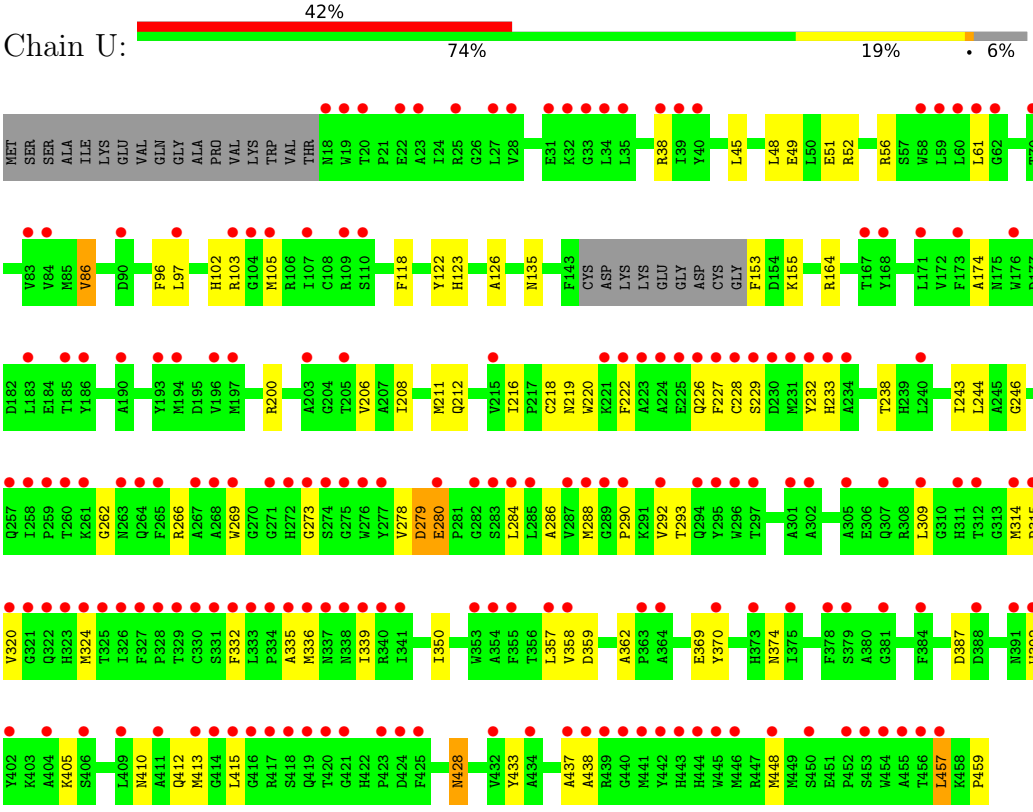
• Molecule 1: BIPHENYL DIOXYGENASE SUBUNIT ALPHA

Chain S:



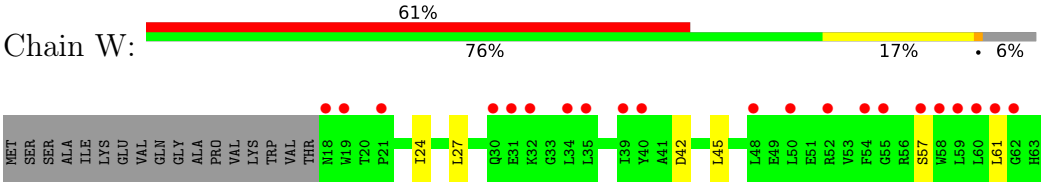
• Molecule 1: BIPHENYL DIOXYGENASE SUBUNIT ALPHA

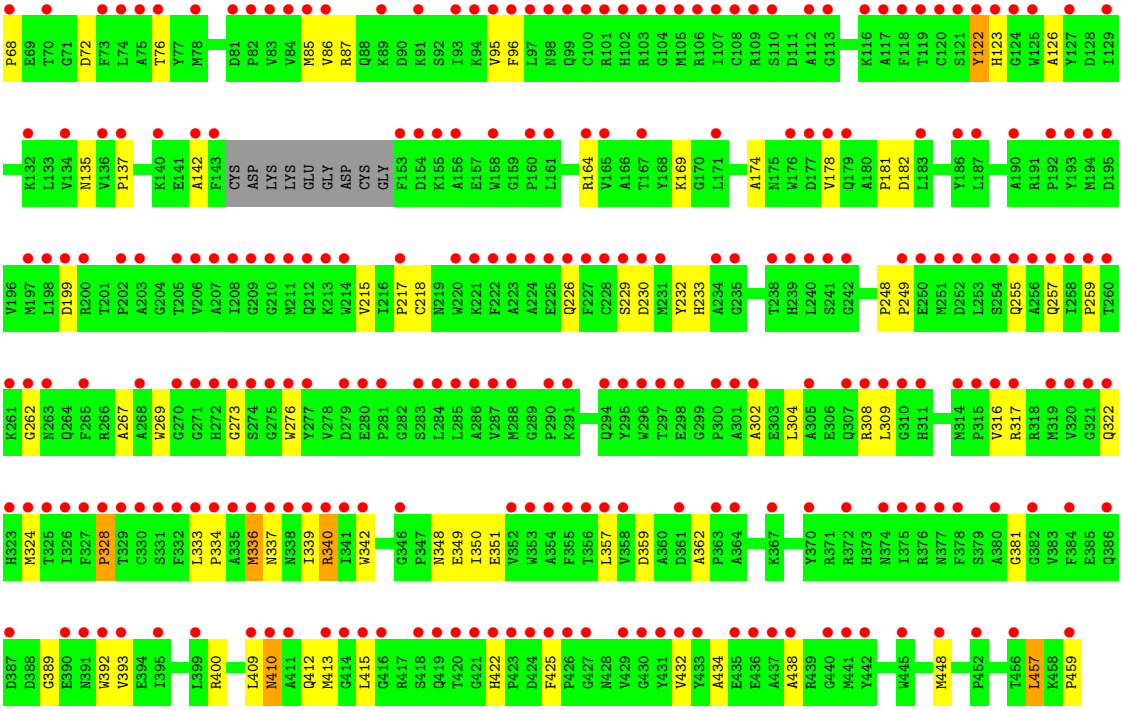
Chain U:



• Molecule 1: BIPHENYL DIOXYGENASE SUBUNIT ALPHA

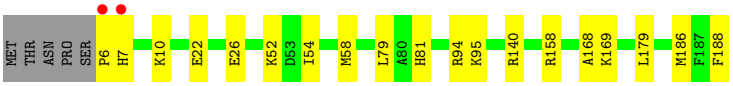
Chain W:





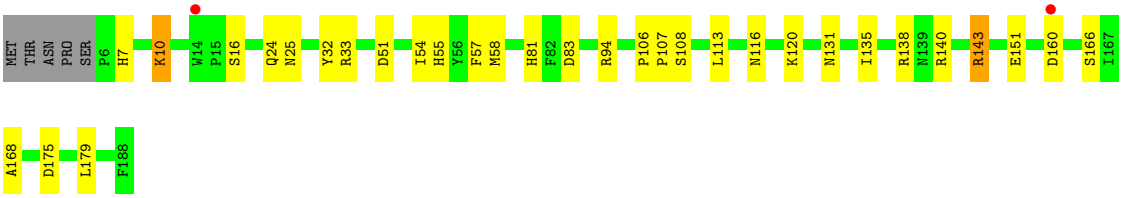
• Molecule 2: BIPHENYL DIOXYGENASE SUBUNIT BETA

Chain B:



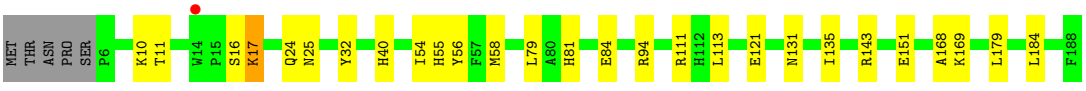
• Molecule 2: BIPHENYL DIOXYGENASE SUBUNIT BETA

Chain D:



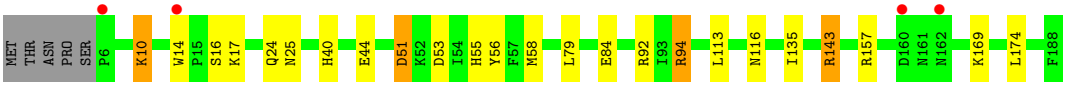
• Molecule 2: BIPHENYL DIOXYGENASE SUBUNIT BETA

Chain F:

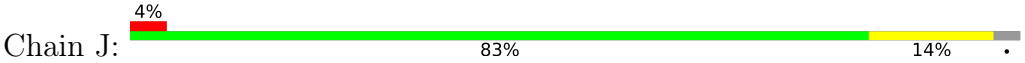


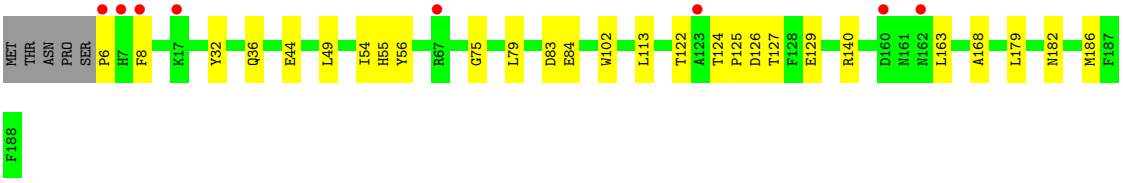
• Molecule 2: BIPHENYL DIOXYGENASE SUBUNIT BETA

Chain H:

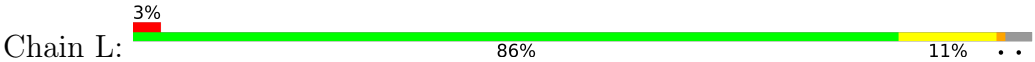


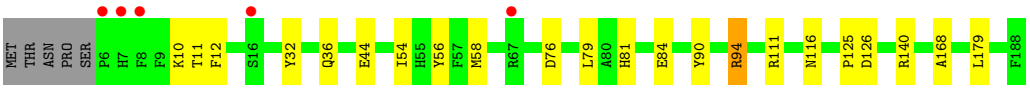
• Molecule 2: BIPHENYL DIOXYGENASE SUBUNIT BETA



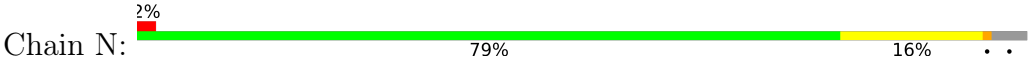


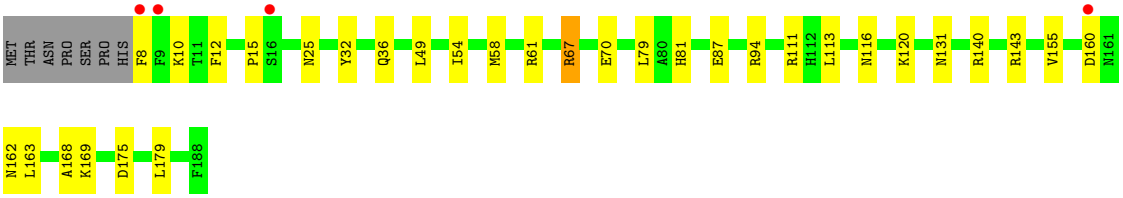
• Molecule 2: BIPHENYL DIOXYGENASE SUBUNIT BETA



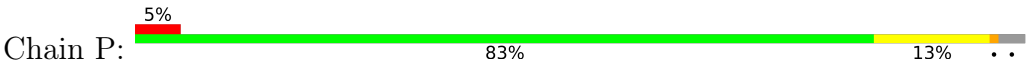


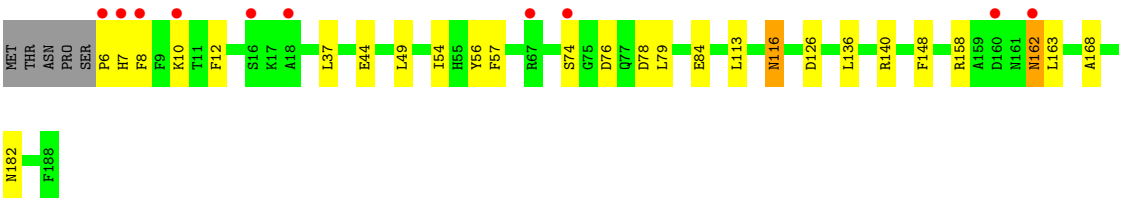
• Molecule 2: BIPHENYL DIOXYGENASE SUBUNIT BETA



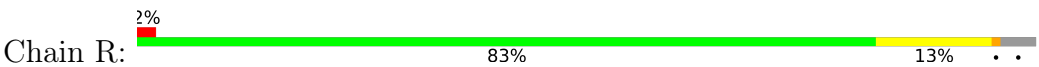


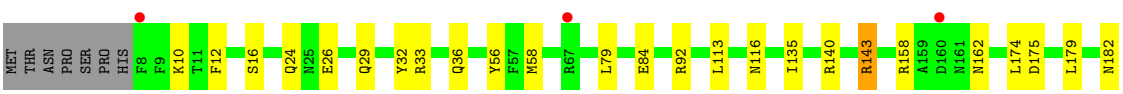
• Molecule 2: BIPHENYL DIOXYGENASE SUBUNIT BETA





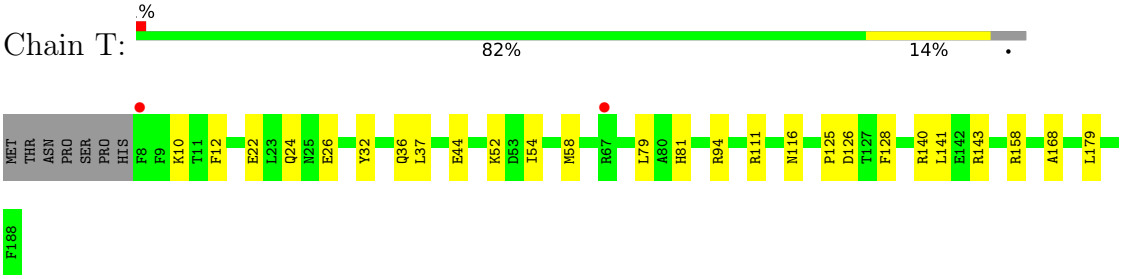
• Molecule 2: BIPHENYL DIOXYGENASE SUBUNIT BETA





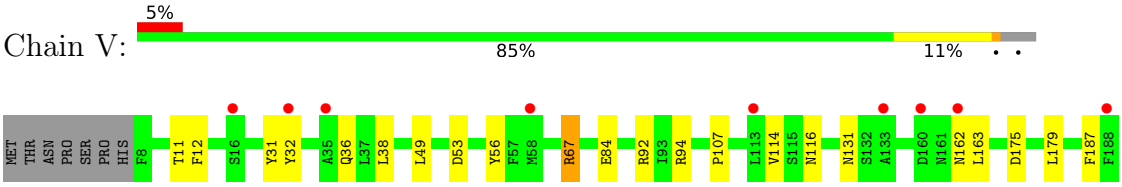
• Molecule 2: BIPHENYL DIOXYGENASE SUBUNIT BETA

Chain T:



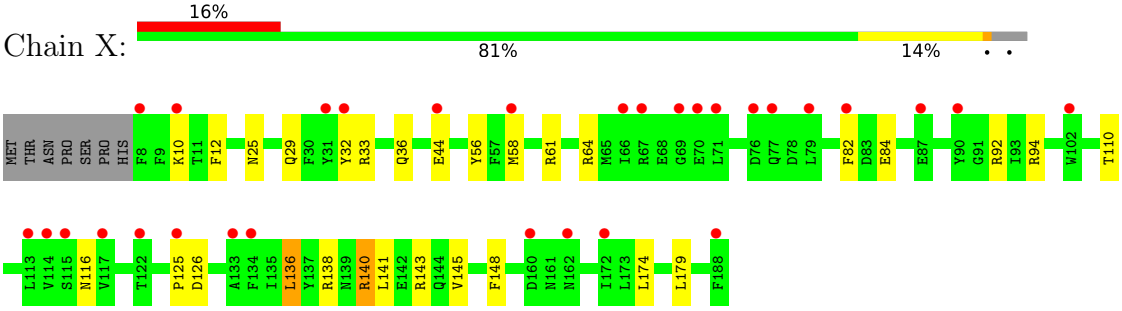
• Molecule 2: BIPHENYL DIOXYGENASE SUBUNIT BETA

Chain V:



• Molecule 2: BIPHENYL DIOXYGENASE SUBUNIT BETA

Chain X:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	133.47Å 133.59Å 133.23Å 102.51° 104.99° 102.75°	Depositor
Resolution (Å)	125.00 – 2.20 28.40 – 2.20	Depositor EDS
% Data completeness (in resolution range)	90.0 (125.00-2.20) 81.0 (28.40-2.20)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.15 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.215 , 0.266 0.234 , 0.272	Depositor DCC
R_{free} test set	17060 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	31.8	Xtriage
Anisotropy	0.250	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 37.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.080 for k,l,h 0.080 for l,h,k 0.018 for -l,-k,-h 0.022 for -h,-l,-k 0.023 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	61911	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.41% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: FE2, FES

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	0/3529	0.63	0/4791
1	C	0.47	0/3529	0.59	1/4791 (0.0%)
1	E	0.52	0/3529	0.64	1/4791 (0.0%)
1	G	0.52	0/3529	0.64	0/4791
1	I	0.43	0/3529	0.57	0/4791
1	K	0.46	0/3529	0.59	0/4791
1	M	0.44	0/3529	0.57	0/4791
1	O	0.43	0/3529	0.58	0/4791
1	Q	0.43	0/3529	0.56	0/4791
1	S	0.43	0/3529	0.56	0/4791
1	U	0.41	0/3529	0.54	0/4791
1	W	0.39	0/3529	0.52	0/4791
2	B	0.60	0/1561	0.68	0/2110
2	D	0.57	0/1561	0.69	1/2110 (0.0%)
2	F	0.59	0/1561	0.69	0/2110
2	H	0.60	0/1561	0.68	1/2110 (0.0%)
2	J	0.42	0/1561	0.54	0/2110
2	L	0.50	0/1561	0.63	0/2110
2	N	0.49	0/1542	0.60	0/2084
2	P	0.44	0/1561	0.59	0/2110
2	R	0.48	0/1542	0.61	0/2084
2	T	0.49	0/1542	0.60	0/2084
2	V	0.43	0/1542	0.55	0/2084
2	X	0.41	0/1542	0.54	0/2084
All	All	0.47	0/60985	0.59	4/82682 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	415	LEU	CA-CB-CG	5.98	129.06	115.30
1	E	340	ARG	NE-CZ-NH1	-5.56	117.52	120.30
2	D	143	ARG	NE-CZ-NH2	-5.53	117.54	120.30
2	H	143	ARG	NE-CZ-NH2	-5.14	117.73	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	3427	0	3276	41	0
1	C	3427	0	3276	30	0
1	E	3427	0	3276	45	0
1	G	3427	0	3276	42	0
1	I	3427	0	3276	52	0
1	K	3427	0	3276	51	0
1	M	3427	0	3276	36	0
1	O	3427	0	3276	52	0
1	Q	3427	0	3276	51	0
1	S	3427	0	3276	61	0
1	U	3427	0	3276	61	0
1	W	3427	0	3276	58	0
2	B	1524	0	1471	15	0
2	D	1524	0	1471	25	0
2	F	1524	0	1471	33	0
2	H	1524	0	1471	29	0
2	J	1524	0	1471	17	0
2	L	1524	0	1471	20	0
2	N	1507	0	1456	34	0
2	P	1524	0	1471	16	0
2	R	1507	0	1456	22	0
2	T	1507	0	1456	24	0
2	V	1507	0	1456	18	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	X	1507	0	1456	21	0
3	A	4	0	0	1	0
3	C	4	0	0	1	0
3	E	4	0	0	1	0
3	G	4	0	0	1	0
3	I	4	0	0	0	0
3	K	4	0	0	2	0
3	M	4	0	0	1	0
3	O	4	0	0	3	0
3	Q	4	0	0	0	0
3	S	4	0	0	4	0
3	U	4	0	0	1	0
3	W	4	0	0	2	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
4	E	1	0	0	0	0
4	G	1	0	0	0	0
4	I	1	0	0	0	0
4	K	1	0	0	0	0
4	M	1	0	0	0	0
4	O	1	0	0	0	0
4	Q	1	0	0	0	0
4	S	1	0	0	0	0
4	U	1	0	0	0	0
4	W	1	0	0	0	0
5	A	185	0	0	5	0
5	B	122	0	0	4	0
5	C	147	0	0	1	0
5	D	113	0	0	4	0
5	E	210	0	0	6	0
5	F	131	0	0	2	0
5	G	181	0	0	1	0
5	H	122	0	0	10	0
5	I	88	0	0	4	0
5	J	48	0	0	2	0
5	K	123	0	0	3	0
5	L	88	0	0	3	0
5	M	96	0	0	2	0
5	N	76	0	0	5	0
5	O	106	0	0	7	0
5	P	38	0	0	0	0
5	Q	90	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	R	89	0	0	1	0
5	S	116	0	0	13	0
5	T	69	0	0	2	0
5	U	115	0	0	10	0
5	V	45	0	0	1	0
5	W	89	0	0	9	0
5	X	37	0	0	2	0
All	All	61911	0	56889	753	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 6.

The worst 5 of 753 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:H:58:MET:HE3	2:H:174:LEU:HD22	1.15	1.14
1:G:287:VAL:HG12	1:G:288:MET:CE	1.87	1.03
2:H:58:MET:CE	2:H:174:LEU:HD22	1.91	1.00
2:B:188:PHE:C	5:B:2106:HOH:O	2.03	0.97
1:A:339:ILE:HD11	1:A:357:LEU:HG	1.49	0.93

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	429/459 (94%)	417 (97%)	12 (3%)	0	100 100
1	C	429/459 (94%)	410 (96%)	19 (4%)	0	100 100
1	E	429/459 (94%)	413 (96%)	15 (4%)	1 (0%)	47 55
1	G	429/459 (94%)	407 (95%)	22 (5%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	I	429/459 (94%)	414 (96%)	14 (3%)	1 (0%)	47 55
1	K	429/459 (94%)	407 (95%)	21 (5%)	1 (0%)	47 55
1	M	429/459 (94%)	402 (94%)	26 (6%)	1 (0%)	47 55
1	O	429/459 (94%)	407 (95%)	20 (5%)	2 (0%)	29 31
1	Q	429/459 (94%)	405 (94%)	23 (5%)	1 (0%)	47 55
1	S	429/459 (94%)	412 (96%)	15 (4%)	2 (0%)	29 31
1	U	429/459 (94%)	401 (94%)	24 (6%)	4 (1%)	17 16
1	W	429/459 (94%)	396 (92%)	33 (8%)	0	100 100
2	B	181/188 (96%)	172 (95%)	9 (5%)	0	100 100
2	D	181/188 (96%)	174 (96%)	7 (4%)	0	100 100
2	F	181/188 (96%)	174 (96%)	7 (4%)	0	100 100
2	H	181/188 (96%)	174 (96%)	7 (4%)	0	100 100
2	J	181/188 (96%)	172 (95%)	8 (4%)	1 (1%)	25 26
2	L	181/188 (96%)	171 (94%)	10 (6%)	0	100 100
2	N	179/188 (95%)	173 (97%)	6 (3%)	0	100 100
2	P	181/188 (96%)	170 (94%)	9 (5%)	2 (1%)	14 12
2	R	179/188 (95%)	174 (97%)	5 (3%)	0	100 100
2	T	179/188 (95%)	172 (96%)	6 (3%)	1 (1%)	25 26
2	V	179/188 (95%)	174 (97%)	5 (3%)	0	100 100
2	X	179/188 (95%)	168 (94%)	11 (6%)	0	100 100
All	All	7310/7764 (94%)	6959 (95%)	334 (5%)	17 (0%)	47 55

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	J	8	PHE
2	P	8	PHE
1	Q	256	ALA
1	M	253	LEU
1	K	445	TRP

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	350/372 (94%)	340 (97%)	10 (3%)	42	54
1	C	350/372 (94%)	338 (97%)	12 (3%)	37	47
1	E	350/372 (94%)	335 (96%)	15 (4%)	29	36
1	G	350/372 (94%)	336 (96%)	14 (4%)	31	40
1	I	350/372 (94%)	339 (97%)	11 (3%)	40	51
1	K	350/372 (94%)	336 (96%)	14 (4%)	31	40
1	M	350/372 (94%)	340 (97%)	10 (3%)	42	54
1	O	350/372 (94%)	340 (97%)	10 (3%)	42	54
1	Q	350/372 (94%)	340 (97%)	10 (3%)	42	54
1	S	350/372 (94%)	340 (97%)	10 (3%)	42	54
1	U	350/372 (94%)	336 (96%)	14 (4%)	31	40
1	W	350/372 (94%)	342 (98%)	8 (2%)	50	63
2	B	162/167 (97%)	157 (97%)	5 (3%)	40	51
2	D	162/167 (97%)	154 (95%)	8 (5%)	25	31
2	F	162/167 (97%)	155 (96%)	7 (4%)	29	36
2	H	162/167 (97%)	155 (96%)	7 (4%)	29	36
2	J	162/167 (97%)	159 (98%)	3 (2%)	57	71
2	L	162/167 (97%)	157 (97%)	5 (3%)	40	51
2	N	160/167 (96%)	154 (96%)	6 (4%)	33	42
2	P	162/167 (97%)	156 (96%)	6 (4%)	34	43
2	R	160/167 (96%)	155 (97%)	5 (3%)	40	51
2	T	160/167 (96%)	155 (97%)	5 (3%)	40	51
2	V	160/167 (96%)	156 (98%)	4 (2%)	47	60
2	X	160/167 (96%)	153 (96%)	7 (4%)	28	35
All	All	6134/6468 (95%)	5928 (97%)	206 (3%)	37	47

5 of 206 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	122	TYR
1	Q	90	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	X	10	LYS
1	M	336	MET
1	O	212	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 126 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	307	GLN
1	U	428	ASN
1	M	422	HIS
1	U	419	GLN
1	W	307	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 monosaccharides in this entry.

5.6 Ligand geometry

Of 24 ligands modelled in this entry, 12 are monoatomic - leaving 12 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	$\#\|Z\| > 2$
3	FES	W	900	1,5	0,4,4	-	-	-		

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	FES	O	900	1,5	0,4,4	-	-	-		
3	FES	G	900	1	0,4,4	-	-	-		
3	FES	C	900	1	0,4,4	-	-	-		
3	FES	M	900	1	0,4,4	-	-	-		
3	FES	S	900	1,5	0,4,4	-	-	-		
3	FES	I	900	1	0,4,4	-	-	-		
3	FES	U	900	1	0,4,4	-	-	-		
3	FES	K	900	1,5	0,4,4	-	-	-		
3	FES	Q	900	1	0,4,4	-	-	-		
3	FES	E	900	1	0,4,4	-	-	-		
3	FES	A	900	1	0,4,4	-	-	-		

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	FES	W	900	1,5	-	-	0/1/1/1
3	FES	O	900	1,5	-	-	0/1/1/1
3	FES	G	900	1	-	-	0/1/1/1
3	FES	C	900	1	-	-	0/1/1/1
3	FES	M	900	1	-	-	0/1/1/1
3	FES	S	900	1,5	-	-	0/1/1/1
3	FES	I	900	1	-	-	0/1/1/1
3	FES	U	900	1	-	-	0/1/1/1
3	FES	K	900	1,5	-	-	0/1/1/1
3	FES	Q	900	1	-	-	0/1/1/1
3	FES	E	900	1	-	-	0/1/1/1
3	FES	A	900	1	-	-	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	W	900	FES	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	O	900	FES	3	0
3	G	900	FES	1	0
3	C	900	FES	1	0
3	M	900	FES	1	0
3	S	900	FES	4	0
3	U	900	FES	1	0
3	K	900	FES	2	0
3	E	900	FES	1	0
3	A	900	FES	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

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^{th} 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	433/459 (94%)	-0.46	4 (0%) 84 83	-7, 10, 35, 51	18 (4%)
1	C	433/459 (94%)	-0.01	9 (2%) 63 61	3, 29, 56, 66	18 (4%)
1	E	433/459 (94%)	-0.46	7 (1%) 72 70	-11, 7, 33, 47	18 (4%)
1	G	433/459 (94%)	-0.17	13 (3%) 50 48	-3, 17, 42, 60	18 (4%)
1	I	433/459 (94%)	0.52	43 (9%) 7 6	7, 50, 89, 101	18 (4%)
1	K	433/459 (94%)	0.03	17 (3%) 39 37	12, 29, 54, 99	18 (4%)
1	M	433/459 (94%)	0.63	52 (12%) 4 3	17, 47, 83, 92	18 (4%)
1	O	433/459 (94%)	0.48	28 (6%) 18 17	20, 49, 71, 100	18 (4%)
1	Q	433/459 (94%)	1.00	62 (14%) 2 2	37, 64, 92, 108	18 (4%)
1	S	433/459 (94%)	0.64	46 (10%) 6 5	15, 48, 80, 113	18 (4%)
1	U	433/459 (94%)	2.11	192 (44%) 0 0	43, 88, 126, 135	18 (4%)
1	W	433/459 (94%)	2.83	280 (64%) 0 0	71, 111, 142, 155	18 (4%)
2	B	183/188 (97%)	-0.62	2 (1%) 80 79	-12, -2, 16, 28	4 (2%)
2	D	183/188 (97%)	-0.61	2 (1%) 80 79	-12, 2, 28, 40	4 (2%)
2	F	183/188 (97%)	-0.61	1 (0%) 91 90	-16, -5, 19, 32	4 (2%)
2	H	183/188 (97%)	-0.53	4 (2%) 62 59	-15, -4, 33, 48	4 (2%)
2	J	183/188 (97%)	-0.02	8 (4%) 34 32	1, 22, 47, 58	4 (2%)
2	L	183/188 (97%)	-0.46	5 (2%) 54 52	3, 12, 33, 52	4 (2%)
2	N	181/188 (96%)	-0.16	4 (2%) 62 59	8, 25, 40, 46	4 (2%)
2	P	183/188 (97%)	-0.16	10 (5%) 25 24	-11, 12, 45, 67	4 (2%)
2	R	181/188 (96%)	-0.41	3 (1%) 70 68	7, 22, 41, 53	4 (2%)
2	T	181/188 (96%)	-0.34	2 (1%) 80 79	5, 20, 49, 75	4 (2%)
2	V	181/188 (96%)	0.51	9 (4%) 28 27	25, 50, 81, 99	4 (2%)
2	X	181/188 (96%)	1.04	30 (16%) 1 1	42, 69, 103, 117	4 (2%)

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
All	All	7382/7764 (95%)	0.36	833 (11%)	5	4	-16, 33, 105, 155	264 (3%)

The worst 5 of 833 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	256	ALA	8.0
1	U	421	GLY	7.9
1	W	429	VAL	7.7
1	U	321	GLY	7.6
1	W	333	LEU	7.6

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
4	FE2	U	901	1/1	0.84	0.27	67,67,67,67	0
3	FES	W	900	4/4	0.91	0.10	65,67,68,69	0
3	FES	S	900	4/4	0.93	0.09	48,48,50,50	0
4	FE2	Q	901	1/1	0.94	0.08	45,45,45,45	0
3	FES	O	900	4/4	0.94	0.06	41,42,43,43	0
4	FE2	I	901	1/1	0.95	0.13	48,48,48,48	0
3	FES	K	900	4/4	0.96	0.06	32,33,34,35	0
3	FES	U	900	4/4	0.96	0.07	21,23,23,27	0
3	FES	A	900	4/4	0.97	0.09	19,20,22,22	0
4	FE2	O	901	1/1	0.97	0.15	32,32,32,32	0
3	FES	G	900	4/4	0.97	0.10	25,26,26,27	0
4	FE2	S	901	1/1	0.97	0.10	26,26,26,26	0
3	FES	Q	900	4/4	0.97	0.07	31,32,37,37	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	FE2	W	901	1/1	0.97	0.08	57,57,57,57	0
3	FES	E	900	4/4	0.98	0.10	13,13,14,15	0
3	FES	M	900	4/4	0.98	0.09	21,21,21,21	0
4	FE2	M	901	1/1	0.98	0.08	37,37,37,37	0
3	FES	I	900	4/4	0.98	0.10	14,15,15,17	0
4	FE2	A	901	1/1	0.99	0.14	13,13,13,13	0
4	FE2	C	901	1/1	0.99	0.14	29,29,29,29	0
4	FE2	E	901	1/1	0.99	0.12	20,20,20,20	0
4	FE2	G	901	1/1	0.99	0.14	14,14,14,14	0
3	FES	C	900	4/4	0.99	0.10	17,18,19,19	0
4	FE2	K	901	1/1	0.99	0.14	19,19,19,19	0

6.5 Other polymers (i)

There are no such residues in this entry.