Full wwPDB X-ray Structure Validation Report (i)

Jan 11, 2022 – 09:17 pm GMT

PDB ID	:	7PEN
Title	:	Crystal Structure of Two-Domain Laccase mutant Y230A from Streptomyces griseoflavus
Authors	:	Gabdulkhakov, A.; Tishchenko, S.; Kolyadenko, I.
Deposited on	:	2021-08-11
Resolution	:	1.60 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

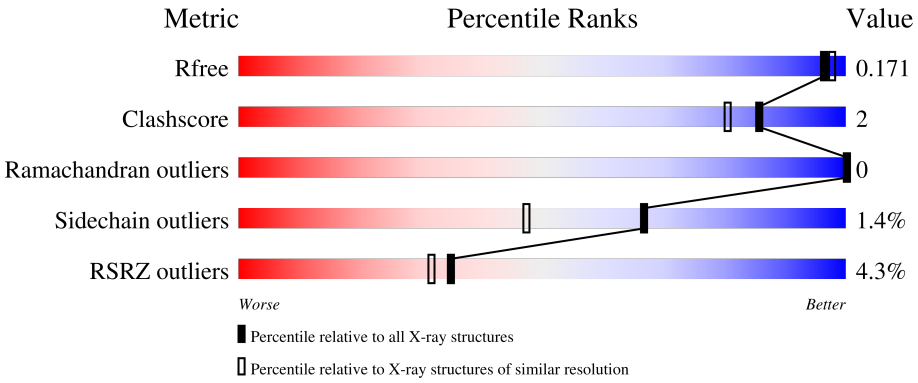
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.24
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.24

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 1.60 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.



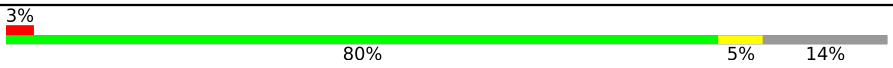
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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	322	3% 82% · 14%
1	B	322	3% 80% 7% 13%
1	C	322	7% 82% 5% 13%
1	D	322	4% 80% 6% 14%
1	E	322	2% 82% · 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	322	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 14384 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 Two-domain laccase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	278	Total	C	N	O	S	0	4	0
			2160	1343	397	407	13			
1	B	280	Total	C	N	O	S	0	5	0
			2186	1358	401	413	14			
1	C	279	Total	C	N	O	S	0	6	0
			2183	1358	400	410	15			
1	D	277	Total	C	N	O	S	0	1	0
			2127	1324	390	401	12			
1	E	277	Total	C	N	O	S	0	9	0
			2199	1367	406	412	14			
1	F	277	Total	C	N	O	S	8	5	0
			2164	1346	398	406	14			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	230	ALA	TYR	engineered mutation	UNP A0A0M4FJ81
B	230	ALA	TYR	engineered mutation	UNP A0A0M4FJ81
C	230	ALA	TYR	engineered mutation	UNP A0A0M4FJ81
D	230	ALA	TYR	engineered mutation	UNP A0A0M4FJ81
E	230	ALA	TYR	engineered mutation	UNP A0A0M4FJ81
F	230	ALA	TYR	engineered mutation	UNP A0A0M4FJ81

- Molecule 2 is COPPER (II) ION (three-letter code: CU) (formula: Cu) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	4	Total	Cu	0	0
			4	4		
2	B	4	Total	Cu	1	0
			4	4		
2	C	4	Total	Cu	0	0
			4	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	5	Total Cu 5 5	1	0
2	E	3	Total Cu 3 3	0	0
2	F	4	Total Cu 4 4	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Na 1 1	0	0
3	D	1	Total Na 1 1	0	0

- Molecule 4 is water.

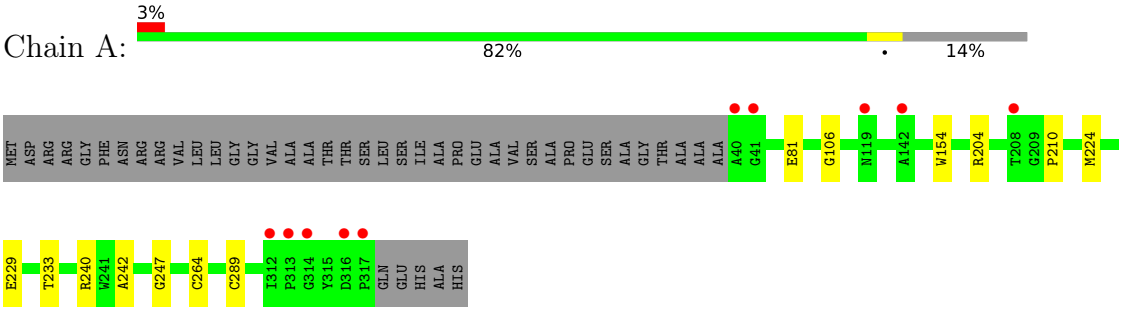
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	241	Total O 241 241	0	0
4	B	216	Total O 216 216	0	0
4	C	198	Total O 198 198	0	0
4	D	223	Total O 223 223	0	0
4	E	241	Total O 241 241	0	0
4	F	220	Total O 220 220	0	0

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.

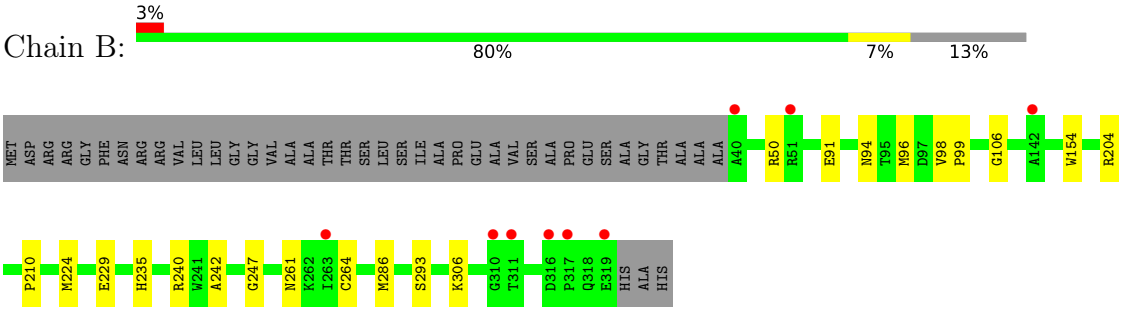
• Molecule 1: Two-domain laccase

Chain A:



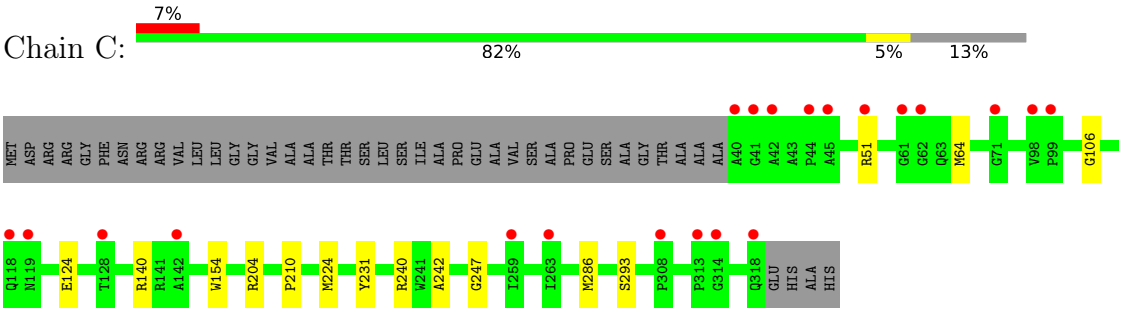
• Molecule 1: Two-domain laccase

Chain B:



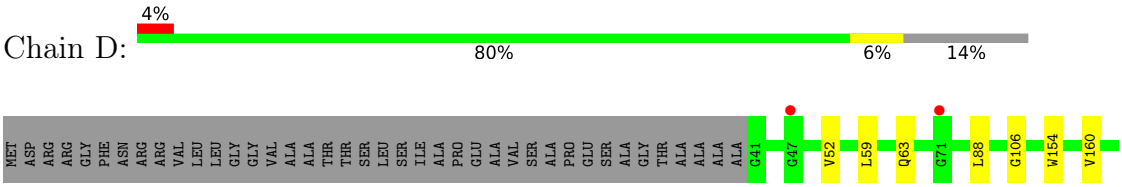
• Molecule 1: Two-domain laccase

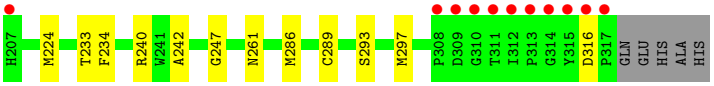
Chain C:



• Molecule 1: Two-domain laccase

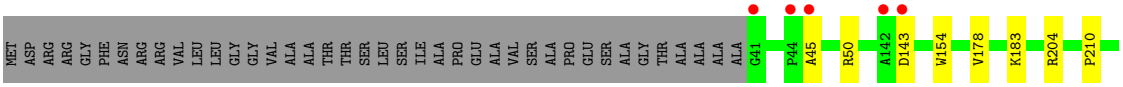
Chain D:

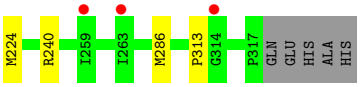




• Molecule 1: Two-domain laccase

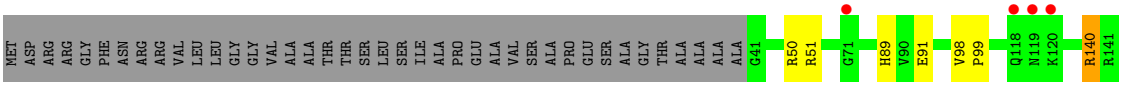
Chain E:

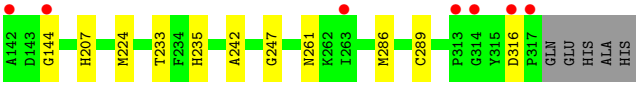




• Molecule 1: Two-domain laccase

Chain F:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	77.26Å 94.98Å 116.27Å 90.00° 91.67° 90.00°	Depositor
Resolution (Å)	47.10 – 1.60 47.10 – 1.60	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.10-1.60) 99.9 (47.10-1.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.63 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.8.0230, PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.135 , 0.165 0.147 , 0.171	Depositor DCC
R_{free} test set	8823 reflections (4.00%)	wwPDB-VP
Wilson B-factor (Å^2)	20.1	Xtriage
Anisotropy	0.334	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.016 for h,-k,-l	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	14384	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.85% 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

5.1 Standard geometry

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

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	0.36	0/2221	0.63	0/3016
1	B	0.35	0/2247	0.64	0/3049
1	C	0.33	0/2244	0.62	0/3047
1	D	0.36	0/2188	0.64	0/2972
1	E	0.36	0/2260	0.62	0/3068
1	F	0.37	0/2226	0.63	0/3021
All	All	0.35	0/13386	0.63	0/18173

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	2160	0	2030	6	0
1	B	2186	0	2051	15	0
1	C	2183	0	2056	9	0
1	D	2127	0	1999	11	0
1	E	2199	0	2074	9	0
1	F	2164	0	2031	11	0
2	A	4	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	4	0	0	0	0
2	C	4	0	0	0	0
2	D	5	0	0	0	0
2	E	3	0	0	0	0
2	F	4	0	0	0	0
3	A	1	0	0	0	0
3	D	1	0	0	0	0
4	A	241	0	0	2	0
4	B	216	0	0	3	0
4	C	198	0	0	0	0
4	D	223	0	0	4	0
4	E	241	0	0	4	0
4	F	220	0	0	3	0
All	All	14384	0	12241	56	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 2.

All (56) 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)
4:D:600:HOH:O	1:E:286[B]:MET:HE2	1.57	1.04
1:F:207[B]:HIS:HD2	4:F:511:HOH:O	1.61	0.83
1:F:50:ARG:NH1	4:F:501:HOH:O	2.16	0.78
1:C:124[B]:GLU:H	1:C:124[B]:GLU:CD	1.94	0.71
1:B:50[A]:ARG:NH1	4:B:502:HOH:O	2.24	0.70
4:D:600:HOH:O	1:E:286[A]:MET:SD	2.52	0.66
4:D:600:HOH:O	1:E:286[B]:MET:CE	2.30	0.66
1:E:45:ALA:HB2	1:E:183:LYS:HE3	1.82	0.62
1:A:229:GLU:OE2	1:B:293:SER:OG	2.19	0.60
1:D:293:SER:HB2	1:D:297:MET:HE1	1.87	0.56
1:E:204[A]:ARG:HD2	1:E:210:PRO:HD3	1.87	0.56
1:D:52:VAL:HG13	1:D:88:LEU:HD11	1.88	0.55
1:A:81:GLU:OE2	4:A:501:HOH:O	2.18	0.55
4:B:507:HOH:O	1:C:286[A]:MET:HG3	2.06	0.55
1:B:229[B]:GLU:OE2	1:C:293:SER:N	2.35	0.54
1:C:51:ARG:HH11	1:C:51:ARG:HG2	1.74	0.53
1:B:235:HIS:HB2	1:B:261:ASN:ND2	2.25	0.51
1:A:204:ARG:HD2	1:A:210:PRO:HD3	1.93	0.50
1:D:293:SER:HB2	1:D:297:MET:CE	2.42	0.50
4:A:512:HOH:O	1:B:286[A]:MET:HG3	2.12	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:229[B]:GLU:O	1:C:231:TYR:OH	2.29	0.47
1:C:106:GLY:HA3	1:C:154:TRP:CD2	2.51	0.46
1:D:286:MET:HG3	4:F:506:HOH:O	2.15	0.46
4:E:509:HOH:O	1:F:286:MET:HG3	2.16	0.45
1:F:51[B]:ARG:NH1	1:F:91:GLU:OE2	2.49	0.45
1:B:106:GLY:HA3	1:B:154:TRP:CD2	2.52	0.45
1:D:106:GLY:HA3	1:D:154:TRP:CD2	2.52	0.44
1:B:242:ALA:O	1:B:247:GLY:HA2	2.18	0.44
1:E:154:TRP:HB2	1:E:178[B]:VAL:HG13	1.99	0.44
1:B:91:GLU:OE1	4:B:501:HOH:O	2.20	0.44
1:A:106:GLY:HA3	1:A:154:TRP:CD2	2.54	0.43
1:C:204[B]:ARG:HD3	1:C:210:PRO:HD3	2.00	0.43
1:F:51[B]:ARG:HG2	1:F:89:HIS:HB2	2.01	0.43
1:D:293:SER:O	1:D:297:MET:HE3	2.18	0.43
1:B:204:ARG:HD3	1:B:210:PRO:HD3	2.00	0.43
1:D:160:VAL:HG12	4:D:531:HOH:O	2.19	0.43
1:E:143:ASP:OD2	4:E:501:HOH:O	2.21	0.43
1:B:99:PRO:HD2	1:F:98:VAL:HG12	2.00	0.42
1:F:235:HIS:HB2	1:F:261:ASN:OD1	2.19	0.42
1:F:242:ALA:O	1:F:247:GLY:HA2	2.20	0.42
1:B:235:HIS:HB2	1:B:261:ASN:HD22	1.83	0.42
1:B:306:LYS:HE2	1:B:306:LYS:HB2	1.69	0.42
1:B:98:VAL:HG12	1:F:99:PRO:HD2	2.00	0.42
1:A:242:ALA:O	1:A:247:GLY:HA2	2.20	0.42
1:A:233:THR:O	1:A:289:CYS:HA	2.20	0.41
1:D:242:ALA:O	1:D:247:GLY:HA2	2.20	0.41
1:F:233:THR:O	1:F:289:CYS:HA	2.20	0.41
1:B:94:ASN:OD1	1:B:96:MET:HB2	2.21	0.41
1:D:234:PHE:O	1:D:261:ASN:HA	2.21	0.41
1:E:313:PRO:O	4:E:502:HOH:O	2.21	0.41
1:D:233:THR:O	1:D:289:CYS:HA	2.21	0.41
1:C:51:ARG:HG2	1:C:51:ARG:NH1	2.35	0.41
1:D:59:LEU:HB2	1:D:63[A]:GLN:HB2	2.02	0.41
1:C:242:ALA:O	1:C:247:GLY:HA2	2.20	0.40
1:E:50[B]:ARG:HD2	4:E:507:HOH:O	2.21	0.40
1:F:140:ARG:HD3	1:F:144:GLY:O	2.22	0.40

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	280/322 (87%)	277 (99%)	3 (1%)	0	100	100
1	B	283/322 (88%)	279 (99%)	4 (1%)	0	100	100
1	C	283/322 (88%)	279 (99%)	4 (1%)	0	100	100
1	D	276/322 (86%)	271 (98%)	5 (2%)	0	100	100
1	E	284/322 (88%)	281 (99%)	3 (1%)	0	100	100
1	F	280/322 (87%)	275 (98%)	5 (2%)	0	100	100
All	All	1686/1932 (87%)	1662 (99%)	24 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	222/248 (90%)	219 (99%)	3 (1%)	67	47
1	B	225/248 (91%)	222 (99%)	3 (1%)	69	50
1	C	225/248 (91%)	220 (98%)	5 (2%)	52	27
1	D	219/248 (88%)	216 (99%)	3 (1%)	67	47
1	E	227/248 (92%)	225 (99%)	2 (1%)	78	65
1	F	223/248 (90%)	219 (98%)	4 (2%)	59	36
All	All	1341/1488 (90%)	1321 (98%)	20 (2%)	67	44

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	224	MET
1	A	240	ARG
1	A	264	CYS
1	B	224	MET
1	B	240	ARG
1	B	264	CYS
1	C	64	MET
1	C	140	ARG
1	C	224[A]	MET
1	C	224[B]	MET
1	C	240	ARG
1	D	224	MET
1	D	240	ARG
1	D	316	ASP
1	E	224	MET
1	E	240	ARG
1	F	140	ARG
1	F	224[A]	MET
1	F	224[B]	MET
1	F	316	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	261	ASN
1	C	63	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 26 ligands modelled in this entry, 26 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

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^{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	278/322 (86%)	0.23	10 (3%)	42	40	14, 20, 36, 63	9 (3%)
1	B	280/322 (86%)	0.37	9 (3%)	47	44	15, 21, 40, 79	3 (1%)
1	C	279/322 (86%)	0.42	21 (7%)	14	12	15, 23, 40, 61	5 (1%)
1	D	277/322 (86%)	0.26	13 (4%)	31	28	14, 21, 36, 72	5 (1%)
1	E	277/322 (86%)	0.20	8 (2%)	51	49	14, 20, 35, 49	4 (1%)
1	F	277/322 (86%)	0.25	11 (3%)	38	35	15, 21, 36, 58	7 (2%)
All	All	1668/1932 (86%)	0.29	72 (4%)	35	32	14, 21, 38, 79	33 (1%)

All (72) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	40	ALA	9.7
1	B	40	ALA	9.1
1	A	314	GLY	8.6
1	B	316	ASP	8.2
1	D	316	ASP	7.9
1	C	318	GLN	7.0
1	C	142	ALA	6.7
1	E	142	ALA	6.5
1	B	319	GLU	6.0
1	A	316	ASP	5.6
1	C	40	ALA	5.4
1	F	316	ASP	5.0
1	F	142	ALA	4.6
1	B	311	THR	4.6
1	D	314	GLY	4.3
1	D	317	PRO	3.8
1	C	314	GLY	3.8
1	E	314	GLY	3.7
1	E	45	ALA	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	F	120	LYS	3.4
1	F	119	ASN	3.3
1	F	314	GLY	3.2
1	C	42	ALA	3.1
1	F	71	GLY	3.1
1	F	118	GLN	3.1
1	C	41	GLY	3.0
1	A	41	GLY	3.0
1	A	142	ALA	3.0
1	A	313	PRO	3.0
1	D	309	ASP	2.9
1	B	310	GLY	2.9
1	C	61	GLY	2.9
1	B	142	ALA	2.9
1	A	312	ILE	2.9
1	D	207	HIS	2.8
1	D	313	PRO	2.8
1	D	315	TYR	2.8
1	C	99	PRO	2.7
1	B	263	ILE	2.6
1	C	51	ARG	2.6
1	C	263	ILE	2.6
1	C	71	GLY	2.5
1	D	311	THR	2.5
1	C	44	PRO	2.5
1	C	98	VAL	2.5
1	C	128	THR	2.4
1	C	62	GLY	2.4
1	C	259	ILE	2.3
1	C	313	PRO	2.3
1	D	47	GLY	2.3
1	C	308	PRO	2.2
1	F	313	PRO	2.2
1	D	308	PRO	2.2
1	A	208	THR	2.2
1	C	119	ASN	2.2
1	F	263	ILE	2.2
1	E	41	GLY	2.2
1	F	317	PRO	2.2
1	E	143	ASP	2.2
1	A	119	ASN	2.1
1	D	310	GLY	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	317	PRO	2.1
1	D	71	GLY	2.1
1	E	259	ILE	2.1
1	B	317	PRO	2.1
1	B	51	ARG	2.1
1	C	45	ALA	2.1
1	D	312	ILE	2.1
1	E	263	ILE	2.1
1	C	118	GLN	2.1
1	E	44	PRO	2.1
1	F	144	GLY	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 monosaccharides in this entry.

6.4 Ligands (i)

LIGAND-RSR INFOmissingINFO

6.5 Other polymers (i)

There are no such residues in this entry.