



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 02:18 am BST

PDB ID : 2HYB
Title : Crystal Structure of Hexameric DsrEFH
Authors : Shin, D.H.; Connie, H.; Schulte, A.; Dahl, C.; Kim, R.; Kim, S.H.; Berkeley
Structural Genomics Center (BSGC)
Deposited on : 2006-08-04
Resolution : 2.50 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtrriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

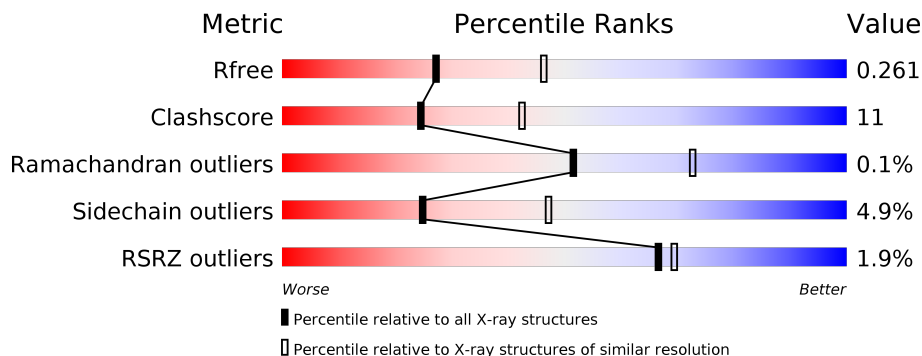
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria 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	130	
1	D	130	
1	G	130	
1	J	130	
1	M	130	
1	P	130	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	136	 6% 68% 28% ..
2	E	136	 3% 61% 35% ..
2	H	136	 3% 71% 24% ..
2	K	136	 3% 74% 22% ..
2	N	136	 5% 68% 26% ..
2	Q	136	 6% 71% 25% ..
3	C	102	 0% 81% 14% ..
3	F	102	 0% 75% 21% ..
3	I	102	 1% 80% 14% 5% ..
3	L	102	 0% 78% 19% ..
3	O	102	 2% 75% 20% ..
3	R	102	 1% 71% 25% ..

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 17611 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 Putative sulfurtransferase dsrE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	130	1029	644	188	194	3	0	0	0
1	D	130	1029	644	188	194	3	0	0	0
1	G	130	1029	644	188	194	3	0	0	0
1	J	130	1029	644	188	194	3	0	0	0
1	M	130	1029	644	188	194	3	0	0	0
1	P	130	1029	644	188	194	3	0	0	0

- Molecule 2 is a protein called Intracellular sulfur oxidation protein dsrF.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	132	1065	676	170	214	5	0	0	0
2	E	132	1065	676	170	214	5	0	0	0
2	H	132	1065	676	170	214	5	0	0	0
2	K	132	1065	676	170	214	5	0	0	0
2	N	132	1065	676	170	214	5	0	0	0
2	Q	132	1065	676	170	214	5	0	0	0

- Molecule 3 is a protein called DsrH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	101	Total	C	N	O	S	0	0	0
			771	493	125	151	2			
3	F	101	Total	C	N	O	S	0	0	0
			772	493	125	152	2			
3	I	101	Total	C	N	O	S	0	0	0
			771	493	125	151	2			
3	L	101	Total	C	N	O	S	0	0	0
			772	493	125	152	2			
3	O	101	Total	C	N	O	S	0	0	0
			771	493	125	151	2			
3	R	101	Total	C	N	O	S	0	0	0
			772	493	125	152	2			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	21	Total	O	0	0
			21	21		
4	B	17	Total	O	0	0
			17	17		
4	C	15	Total	O	0	0
			15	15		
4	D	23	Total	O	0	0
			23	23		
4	E	29	Total	O	0	0
			29	29		
4	F	20	Total	O	0	0
			20	20		
4	G	29	Total	O	0	0
			29	29		
4	H	11	Total	O	0	0
			11	11		
4	I	25	Total	O	0	0
			25	25		
4	J	28	Total	O	0	0
			28	28		
4	K	33	Total	O	0	0
			33	33		
4	L	22	Total	O	0	0
			22	22		
4	M	21	Total	O	0	0
			21	21		
4	N	18	Total	O	0	0
			18	18		

Continued on next page...


Continued from previous page...

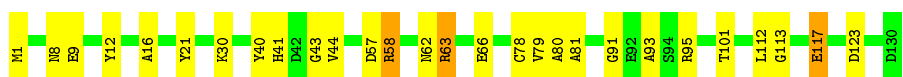
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	O	22	Total 22	O 22	0	0
4	P	32	Total 32	O 32	0	0
4	Q	30	Total 30	O 30	0	0
4	R	22	Total 22	O 22	0	0

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Putative sulfurtransferase dsrE

Chain A:  78% 19%




- Molecule 1: Putative sulfurtransferase dsrE

Chain D:  79% 18%



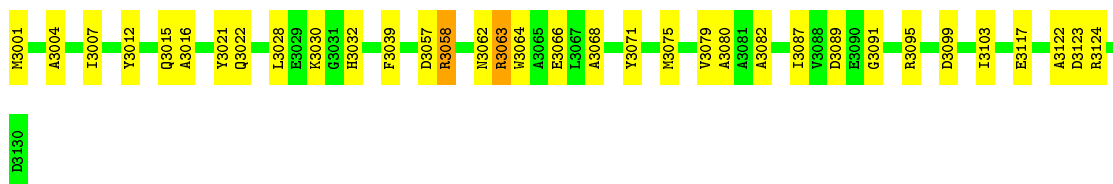
- Molecule 1: Putative sulfurtransferase dsrE

Chain G:  79% 18%




- Molecule 1: Putative sulfurtransferase dsrE

Chain J:  74% 25%



- Molecule 1: Putative sulfurtransferase dsrE

Chain M:  72% 26%





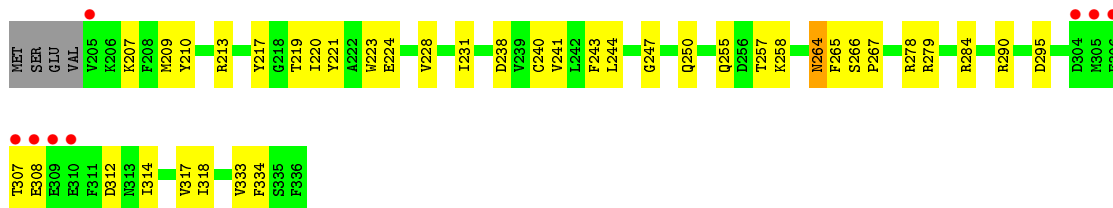
- Molecule 1: Putative sulfurtransferase dsrE

Chain P: 81% 17%



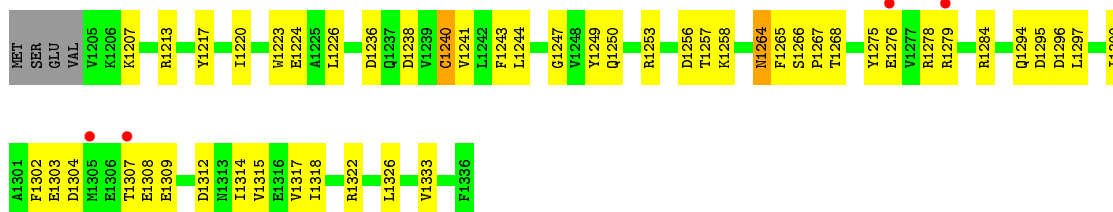
- Molecule 2: Intracellular sulfur oxidation protein dsrF

Chain B: 6% 68% 28%



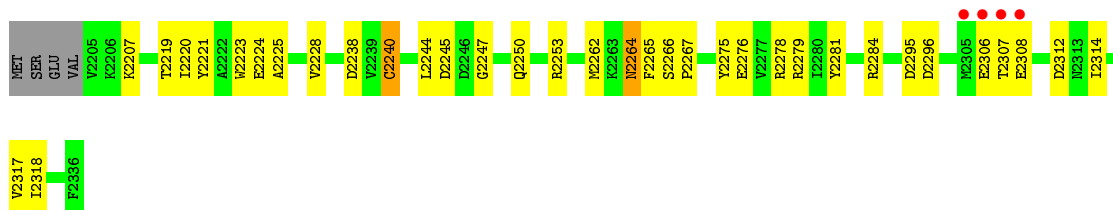
- Molecule 2: Intracellular sulfur oxidation protein dsrF

Chain E: 3% 61% 35%



- Molecule 2: Intracellular sulfur oxidation protein dsrF

Chain H: 3% 71% 24%



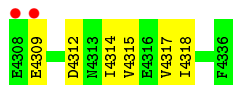
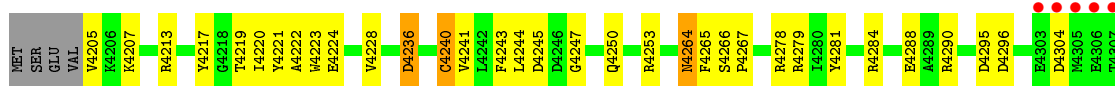
- Molecule 2: Intracellular sulfur oxidation protein dsrF

Chain K: 3% 74% 22%





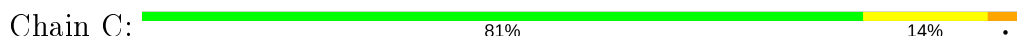
• Molecule 2: Intracellular sulfur oxidation protein dsrF



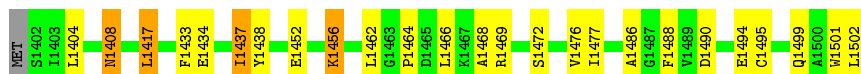
• Molecule 2: Intracellular sulfur oxidation protein dsrF



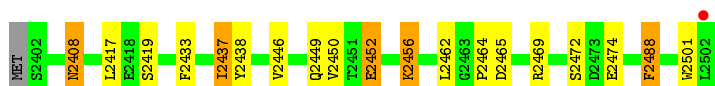
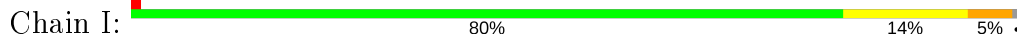
• Molecule 3: DsrH




• Molecule 3: DsrH



• Molecule 3: DsrH




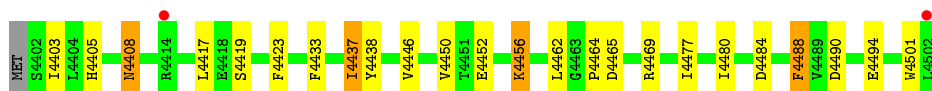
• Molecule 3: DsrH

Chain L:  78% 19%



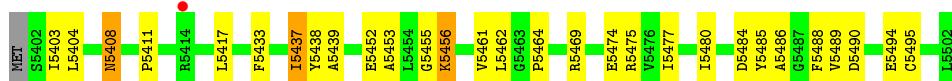
• Molecule 3: DsrH

Chain O:  2% 75% 20%



• Molecule 3: DsrH

Chain R:  71% 25%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	56.61Å 183.11Å 107.83Å 90.00° 99.57° 90.00°	Depositor
Resolution (Å)	19.98 – 2.50 47.66 – 2.50	Depositor EDS
% Data completeness (in resolution range)	92.7 (19.98-2.50) 92.8 (47.66-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.40 (at 2.51Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.206 , 0.262 0.205 , 0.261	Depositor DCC
R_{free} test set	7338 reflections (10.11%)	wwPDB-VP
Wilson B-factor (Å ²)	29.1	Xtrriage
Anisotropy	0.152	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 41.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	17611	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 23.61 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.5041e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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](#)

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	0.38	0/1050	0.61	0/1419
1	D	0.39	0/1050	0.63	0/1419
1	G	0.38	0/1050	0.62	0/1419
1	J	0.38	0/1050	0.63	0/1419
1	M	0.38	0/1050	0.62	0/1419
1	P	0.39	0/1050	0.61	0/1419
2	B	0.39	0/1083	0.63	0/1464
2	E	0.40	0/1083	0.62	0/1464
2	H	0.37	0/1083	0.61	0/1464
2	K	0.41	0/1083	0.63	0/1464
2	N	0.38	0/1083	0.62	0/1464
2	Q	0.40	0/1083	0.62	0/1464
3	C	0.39	0/784	0.67	0/1064
3	F	0.37	0/785	0.63	0/1064
3	I	0.41	0/784	0.68	0/1064
3	L	0.38	0/785	0.65	0/1064
3	O	0.40	0/784	0.67	0/1064
3	R	0.40	0/785	0.66	0/1064
All	All	0.39	0/17505	0.63	0/23682

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	E	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	E	1249	TYR	Sidechain

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	1029	0	992	21	0
1	D	1029	0	989	20	0
1	G	1029	0	989	19	0
1	J	1029	0	989	25	0
1	M	1029	0	989	30	0
1	P	1029	0	989	20	0
2	B	1065	0	1021	30	0
2	E	1065	0	1021	33	0
2	H	1065	0	1021	24	0
2	K	1065	0	1021	21	0
2	N	1065	0	1021	36	0
2	Q	1065	0	1021	25	0
3	C	771	0	771	14	0
3	F	772	0	771	20	0
3	I	771	0	771	12	0
3	L	772	0	771	13	0
3	O	771	0	771	21	0
3	R	772	0	771	17	0
4	A	21	0	0	2	0
4	B	17	0	0	1	0
4	C	15	0	0	0	0
4	D	23	0	0	3	0
4	E	29	0	0	1	0
4	F	20	0	0	2	0
4	G	29	0	0	1	0
4	H	11	0	0	1	0
4	I	25	0	0	2	0
4	J	28	0	0	1	0
4	K	33	0	0	2	0
4	L	22	0	0	0	0
4	M	21	0	0	3	0
4	N	18	0	0	5	0
4	O	22	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	P	32	0	0	1	0
4	Q	30	0	0	1	0
4	R	22	0	0	1	0
All	All	17611	0	16689	369	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 11.

The worst 5 of 369 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:M:4091:GLY:O	1:M:4095:ARG:HG3	1.74	0.87
1:J:3091:GLY:O	1:J:3095:ARG:HG3	1.76	0.85
1:G:2062:ASN:O	1:G:2066:GLU:HG3	1.80	0.82
2:H:2253:ARG:HG2	2:H:2296:ASP:O	1.79	0.82
3:C:408:ASN:HD22	3:C:408:ASN:H	1.30	0.80

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	128/130 (98%)	120 (94%)	8 (6%)	0	100	100
1	D	128/130 (98%)	121 (94%)	7 (6%)	0	100	100
1	G	128/130 (98%)	123 (96%)	5 (4%)	0	100	100
1	J	128/130 (98%)	122 (95%)	6 (5%)	0	100	100
1	M	128/130 (98%)	122 (95%)	6 (5%)	0	100	100
1	P	128/130 (98%)	123 (96%)	5 (4%)	0	100	100
2	B	130/136 (96%)	120 (92%)	10 (8%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	E	130/136 (96%)	119 (92%)	11 (8%)	0	100	100
2	H	130/136 (96%)	120 (92%)	10 (8%)	0	100	100
2	K	130/136 (96%)	125 (96%)	5 (4%)	0	100	100
2	N	130/136 (96%)	123 (95%)	5 (4%)	2 (2%)	10	18
2	Q	130/136 (96%)	123 (95%)	7 (5%)	0	100	100
3	C	99/102 (97%)	96 (97%)	3 (3%)	0	100	100
3	F	99/102 (97%)	95 (96%)	4 (4%)	0	100	100
3	I	99/102 (97%)	95 (96%)	4 (4%)	0	100	100
3	L	99/102 (97%)	95 (96%)	3 (3%)	1 (1%)	15	28
3	O	99/102 (97%)	97 (98%)	2 (2%)	0	100	100
3	R	99/102 (97%)	94 (95%)	5 (5%)	0	100	100
All	All	2142/2208 (97%)	2033 (95%)	106 (5%)	3 (0%)	51	73

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	L	3455	GLY
2	N	4236	ASP
2	N	4217	TYR

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	105/105 (100%)	102 (97%)	3 (3%)	42	69
1	D	105/105 (100%)	101 (96%)	4 (4%)	33	58
1	G	105/105 (100%)	100 (95%)	5 (5%)	25	48
1	J	105/105 (100%)	102 (97%)	3 (3%)	42	69
1	M	105/105 (100%)	101 (96%)	4 (4%)	33	58
1	P	105/105 (100%)	102 (97%)	3 (3%)	42	69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	115/119 (97%)	111 (96%)	4 (4%)	36	62
2	E	115/119 (97%)	110 (96%)	5 (4%)	29	53
2	H	115/119 (97%)	110 (96%)	5 (4%)	29	53
2	K	115/119 (97%)	110 (96%)	5 (4%)	29	53
2	N	115/119 (97%)	109 (95%)	6 (5%)	23	44
2	Q	115/119 (97%)	110 (96%)	5 (4%)	29	53
3	C	84/85 (99%)	78 (93%)	6 (7%)	14	28
3	F	84/85 (99%)	77 (92%)	7 (8%)	11	22
3	I	84/85 (99%)	78 (93%)	6 (7%)	14	28
3	L	84/85 (99%)	79 (94%)	5 (6%)	19	37
3	O	84/85 (99%)	78 (93%)	6 (7%)	14	28
3	R	84/85 (99%)	76 (90%)	8 (10%)	8	17
All	All	1824/1854 (98%)	1734 (95%)	90 (5%)	25	47

5 of 90 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	I	2417	LEU
2	K	3264	ASN
3	R	5417	LEU
3	I	2437	ILE
1	J	3058	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	2096	ASN
1	J	3045	ASN
1	P	5083	GLN
2	H	2264	ASN
3	I	2408	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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, 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	130/130 (100%)	-0.37	0 100 100	9, 30, 60, 90	0
1	D	130/130 (100%)	-0.56	0 100 100	15, 27, 47, 78	0
1	G	130/130 (100%)	-0.46	1 (0%) 86 87	15, 28, 46, 84	0
1	J	130/130 (100%)	-0.51	0 100 100	16, 28, 51, 75	0
1	M	130/130 (100%)	-0.36	1 (0%) 86 87	14, 29, 56, 83	0
1	P	130/130 (100%)	-0.50	0 100 100	14, 26, 55, 76	0
2	B	132/136 (97%)	0.25	8 (6%) 21 22	19, 40, 105, 165	0
2	E	132/136 (97%)	-0.09	4 (3%) 50 53	16, 31, 103, 154	0
2	H	132/136 (97%)	0.24	4 (3%) 50 53	23, 40, 101, 184	0
2	K	132/136 (97%)	-0.24	4 (3%) 50 53	15, 30, 85, 188	0
2	N	132/136 (97%)	0.27	7 (5%) 26 28	16, 39, 114, 167	0
2	Q	132/136 (97%)	-0.03	8 (6%) 21 22	14, 30, 96, 184	0
3	C	101/102 (99%)	-0.22	0 100 100	16, 29, 58, 78	0
3	F	101/102 (99%)	-0.20	0 100 100	19, 33, 62, 84	0
3	I	101/102 (99%)	-0.37	1 (0%) 82 84	10, 30, 62, 99	0
3	L	101/102 (99%)	-0.18	0 100 100	21, 34, 61, 104	0
3	O	101/102 (99%)	-0.26	2 (1%) 65 68	17, 29, 51, 78	0
3	R	101/102 (99%)	-0.28	1 (0%) 82 84	15, 30, 61, 82	0
All	All	2178/2208 (98%)	-0.21	41 (1%) 66 69	9, 31, 72, 188	0

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	Q	5305	MET	10.5
2	H	2305	MET	8.8
2	K	3305	MET	7.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	E	1305	MET	7.1
2	H	2308	GLU	6.7

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.