wwPDB X-ray Structure Validation Summary Report (i)

Feb 3, 2024 – 11:01 PM EST

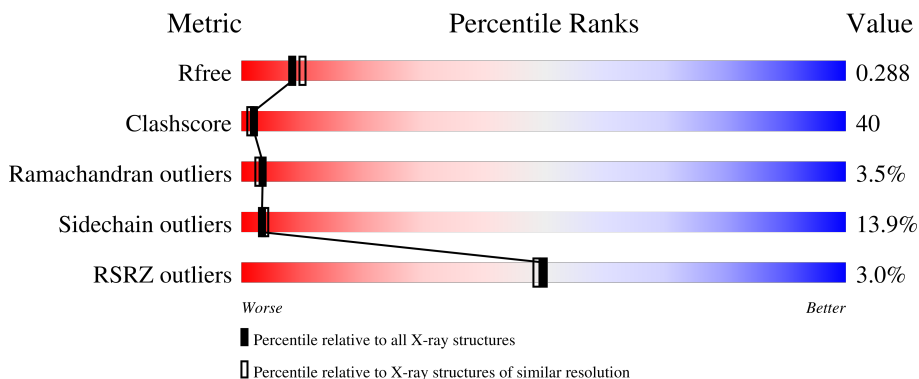|		
PDB ID	:	1NYE
Title	:	Crystal structure of OsmC from E. coli
Authors	:	Shin, D.H.; Choi, I.-G.; Busso, D.; Jancarik, J.; Yokota, H.; Kim, R.; Kim, S.-H.; Berkeley Structural Genomics Center (BSGC)
Deposited on	:	2003-02-12
Resolution	:	2.40 Å(reported)

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 2.40 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	Percentile Ranks	Value
Rfree		0.288
Clashscore		40
Ramachandran outliers		3.5%
Sidechain outliers		13.9%
RSRZ outliers		3.0%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	162	% 43% 36% 9% · 12%
1	B	162	51% 28% 9% · 12%
1	C	162	5% 52% 33% 7% · 6%
1	D	162	2% 52% 35% 11% ·
1	E	162	4% 47% 38% 9% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	162	

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 6900 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 Osmotically inducible protein C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	143	Total	C	N	O	S	0	0	0
			1059	666	177	211	5			
1	B	143	Total	C	N	O	S	0	0	0
			1059	666	177	211	5			
1	C	153	Total	C	N	O	S	0	0	0
			1144	720	191	228	5			
1	D	162	Total	C	N	O	S	0	0	0
			1227	772	208	242	5			
1	E	151	Total	C	N	O	S	0	0	0
			1130	712	189	224	5			
1	F	158	Total	C	N	O	S	0	0	0
			1194	751	204	234	5			

There are 114 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	HIS	-	expression tag	UNP P0C0L2
A	3	HIS	-	expression tag	UNP P0C0L2
A	4	HIS	-	expression tag	UNP P0C0L2
A	5	HIS	-	expression tag	UNP P0C0L2
A	6	ASP	-	expression tag	UNP P0C0L2
A	7	TYR	-	expression tag	UNP P0C0L2
A	8	ASP	-	expression tag	UNP P0C0L2
A	9	ILE	-	expression tag	UNP P0C0L2
A	10	PRO	-	expression tag	UNP P0C0L2
A	11	THR	-	expression tag	UNP P0C0L2
A	12	THR	-	expression tag	UNP P0C0L2
A	13	GLU	-	expression tag	UNP P0C0L2
A	14	ASN	-	expression tag	UNP P0C0L2
A	15	LEU	-	expression tag	UNP P0C0L2
A	16	TYR	-	expression tag	UNP P0C0L2
A	17	PHE	-	expression tag	UNP P0C0L2
A	18	GLN	-	expression tag	UNP P0C0L2

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	19	GLY	-	expression tag	UNP P0C0L2
A	20	HIS	-	expression tag	UNP P0C0L2
B	202	HIS	-	expression tag	UNP P0C0L2
B	203	HIS	-	expression tag	UNP P0C0L2
B	204	HIS	-	expression tag	UNP P0C0L2
B	205	HIS	-	expression tag	UNP P0C0L2
B	206	ASP	-	expression tag	UNP P0C0L2
B	207	TYR	-	expression tag	UNP P0C0L2
B	208	ASP	-	expression tag	UNP P0C0L2
B	209	ILE	-	expression tag	UNP P0C0L2
B	210	PRO	-	expression tag	UNP P0C0L2
B	211	THR	-	expression tag	UNP P0C0L2
B	212	THR	-	expression tag	UNP P0C0L2
B	213	GLU	-	expression tag	UNP P0C0L2
B	214	ASN	-	expression tag	UNP P0C0L2
B	215	LEU	-	expression tag	UNP P0C0L2
B	216	TYR	-	expression tag	UNP P0C0L2
B	217	PHE	-	expression tag	UNP P0C0L2
B	218	GLN	-	expression tag	UNP P0C0L2
B	219	GLY	-	expression tag	UNP P0C0L2
B	220	HIS	-	expression tag	UNP P0C0L2
C	402	HIS	-	expression tag	UNP P0C0L2
C	403	HIS	-	expression tag	UNP P0C0L2
C	404	HIS	-	expression tag	UNP P0C0L2
C	405	HIS	-	expression tag	UNP P0C0L2
C	406	ASP	-	expression tag	UNP P0C0L2
C	407	TYR	-	expression tag	UNP P0C0L2
C	408	ASP	-	expression tag	UNP P0C0L2
C	409	ILE	-	expression tag	UNP P0C0L2
C	410	PRO	-	expression tag	UNP P0C0L2
C	411	THR	-	expression tag	UNP P0C0L2
C	412	THR	-	expression tag	UNP P0C0L2
C	413	GLU	-	expression tag	UNP P0C0L2
C	414	ASN	-	expression tag	UNP P0C0L2
C	415	LEU	-	expression tag	UNP P0C0L2
C	416	TYR	-	expression tag	UNP P0C0L2
C	417	PHE	-	expression tag	UNP P0C0L2
C	418	GLN	-	expression tag	UNP P0C0L2
C	419	GLY	-	expression tag	UNP P0C0L2
C	420	HIS	-	expression tag	UNP P0C0L2
D	602	HIS	-	expression tag	UNP P0C0L2
D	603	HIS	-	expression tag	UNP P0C0L2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	604	HIS	-	expression tag	UNP P0C0L2
D	605	HIS	-	expression tag	UNP P0C0L2
D	606	ASP	-	expression tag	UNP P0C0L2
D	607	TYR	-	expression tag	UNP P0C0L2
D	608	ASP	-	expression tag	UNP P0C0L2
D	609	ILE	-	expression tag	UNP P0C0L2
D	610	PRO	-	expression tag	UNP P0C0L2
D	611	THR	-	expression tag	UNP P0C0L2
D	612	THR	-	expression tag	UNP P0C0L2
D	613	GLU	-	expression tag	UNP P0C0L2
D	614	ASN	-	expression tag	UNP P0C0L2
D	615	LEU	-	expression tag	UNP P0C0L2
D	616	TYR	-	expression tag	UNP P0C0L2
D	617	PHE	-	expression tag	UNP P0C0L2
D	618	GLN	-	expression tag	UNP P0C0L2
D	619	GLY	-	expression tag	UNP P0C0L2
D	620	HIS	-	expression tag	UNP P0C0L2
E	802	HIS	-	expression tag	UNP P0C0L2
E	803	HIS	-	expression tag	UNP P0C0L2
E	804	HIS	-	expression tag	UNP P0C0L2
E	805	HIS	-	expression tag	UNP P0C0L2
E	806	ASP	-	expression tag	UNP P0C0L2
E	807	TYR	-	expression tag	UNP P0C0L2
E	808	ASP	-	expression tag	UNP P0C0L2
E	809	ILE	-	expression tag	UNP P0C0L2
E	810	PRO	-	expression tag	UNP P0C0L2
E	811	THR	-	expression tag	UNP P0C0L2
E	812	THR	-	expression tag	UNP P0C0L2
E	813	GLU	-	expression tag	UNP P0C0L2
E	814	ASN	-	expression tag	UNP P0C0L2
E	815	LEU	-	expression tag	UNP P0C0L2
E	816	TYR	-	expression tag	UNP P0C0L2
E	817	PHE	-	expression tag	UNP P0C0L2
E	818	GLN	-	expression tag	UNP P0C0L2
E	819	GLY	-	expression tag	UNP P0C0L2
E	820	HIS	-	expression tag	UNP P0C0L2
F	1002	HIS	-	expression tag	UNP P0C0L2
F	1003	HIS	-	expression tag	UNP P0C0L2
F	1004	HIS	-	expression tag	UNP P0C0L2
F	1005	HIS	-	expression tag	UNP P0C0L2
F	1006	ASP	-	expression tag	UNP P0C0L2
F	1007	TYR	-	expression tag	UNP P0C0L2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	1008	ASP	-	expression tag	UNP P0C0L2
F	1009	ILE	-	expression tag	UNP P0C0L2
F	1010	PRO	-	expression tag	UNP P0C0L2
F	1011	THR	-	expression tag	UNP P0C0L2
F	1012	THR	-	expression tag	UNP P0C0L2
F	1013	GLU	-	expression tag	UNP P0C0L2
F	1014	ASN	-	expression tag	UNP P0C0L2
F	1015	LEU	-	expression tag	UNP P0C0L2
F	1016	TYR	-	expression tag	UNP P0C0L2
F	1017	PHE	-	expression tag	UNP P0C0L2
F	1018	GLN	-	expression tag	UNP P0C0L2
F	1019	GLY	-	expression tag	UNP P0C0L2
F	1020	HIS	-	expression tag	UNP P0C0L2

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	12	Total 12	O 12	0	0
2	B	16	Total 16	O 16	0	0
2	C	8	Total 8	O 8	0	0
2	D	16	Total 16	O 16	0	0
2	E	20	Total 20	O 20	0	0
2	F	15	Total 15	O 15	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.
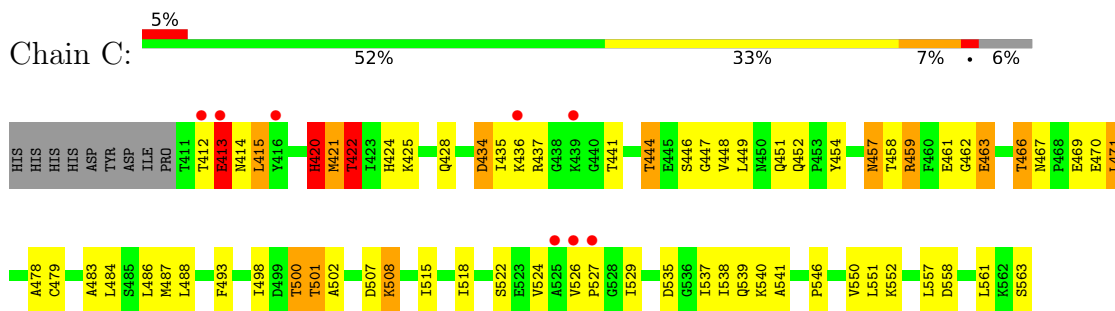
• Molecule 1: Osmotically inducible protein C

Chain A:

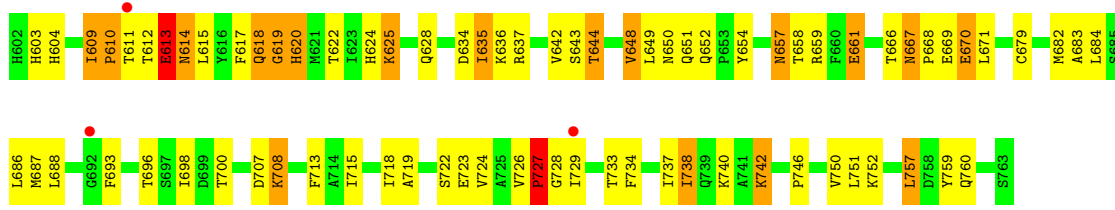• Molecule 1: Osmotically inducible protein C

Chain B:
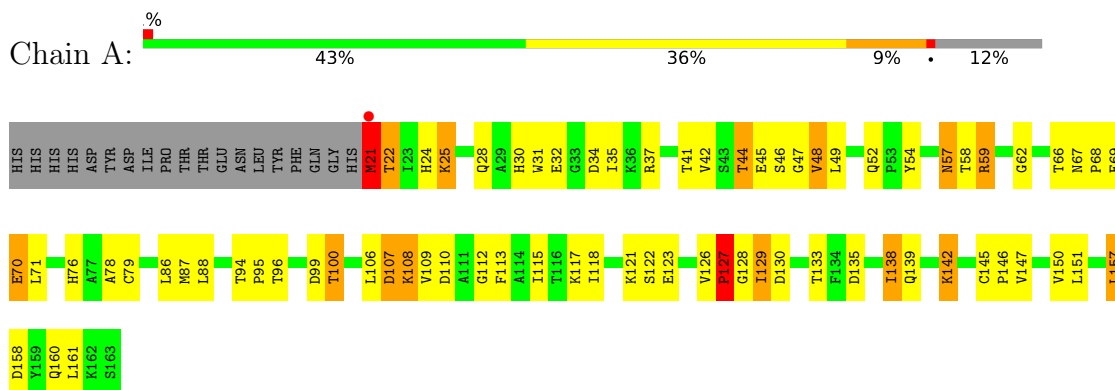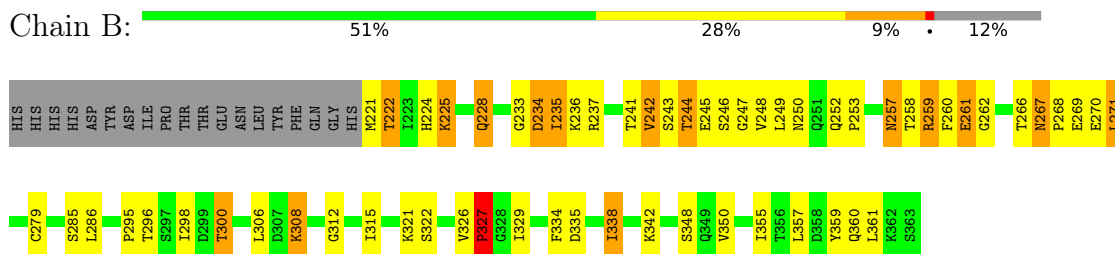
• Molecule 1: Osmotically inducible protein C
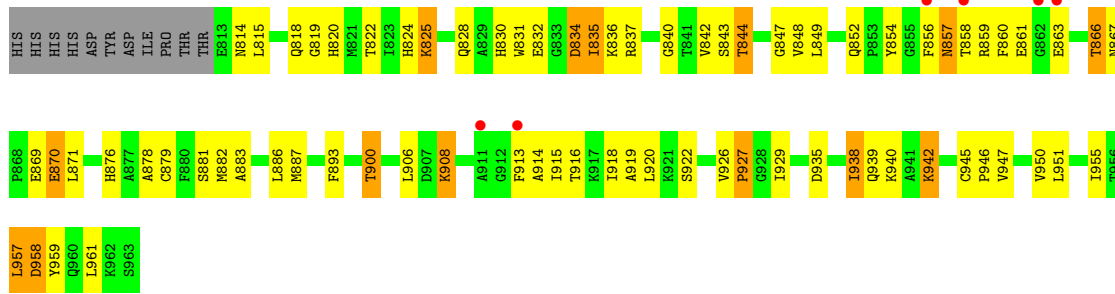
Chain C:

• Molecule 1: Osmotically inducible protein C

Chain D:

• Molecule 1: Osmotically inducible protein C

Chain E:

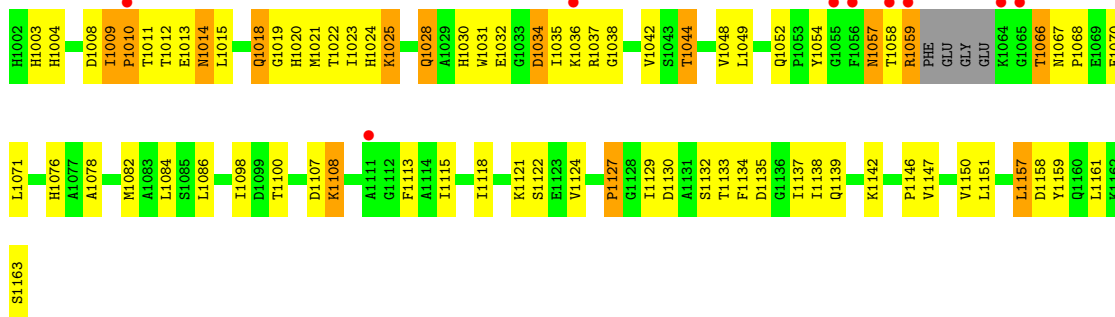• Molecule 1: Osmotically inducible protein C

Chain F:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	49.53Å 90.29Å 112.68Å 90.00° 93.90° 90.00°	Depositor
Resolution (Å)	19.96 – 2.40 46.41 – 2.40	Depositor EDS
% Data completeness (in resolution range)	81.6 (19.96-2.40) 92.7 (46.41-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	3.57 (at 2.39Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.218 , 0.277 0.234 , 0.288	Depositor DCC
R_{free} test set	3838 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	44.5	Xtriage
Anisotropy	0.617	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 43.8	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6900	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.23% 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 ⓘ

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	0.42	0/1076	0.80	2/1451 (0.1%)
1	B	0.42	0/1076	0.73	0/1451
1	C	0.45	2/1164 (0.2%)	0.74	2/1571 (0.1%)
1	D	0.40	0/1253	0.69	1/1694 (0.1%)
1	E	0.41	0/1150	0.69	0/1551
1	F	0.41	0/1218	0.68	0/1646
All	All	0.42	2/6937 (0.0%)	0.72	5/9364 (0.1%)

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
1	A	1	0
1	C	0	1
All	All	1	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	422	THR	N-CA	5.99	1.58	1.46
1	C	413	GLU	N-CA	5.10	1.56	1.46

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	21	MET	CB-CA-C	11.93	134.26	110.40
1	C	420	HIS	C-N-CA	8.09	141.94	121.70
1	C	422	THR	N-CA-CB	6.16	122.00	110.30
1	A	22	THR	N-CA-C	-5.26	96.80	111.00
1	D	613	GLU	C-N-CA	5.10	134.46	121.70

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	21	MET	CA

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	421	MET	Peptide

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	1059	0	1059	118	0
1	B	1059	0	1059	99	0
1	C	1144	0	1132	115	0
1	D	1227	0	1195	124	0
1	E	1130	0	1118	104	1
1	F	1194	0	1170	124	0
2	A	12	0	0	0	0
2	B	16	0	0	3	0
2	C	8	0	0	3	0
2	D	16	0	0	2	1
2	E	20	0	0	3	0
2	F	15	0	0	1	0
All	All	6900	0	6733	548	1

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 40.

The worst 5 of 548 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:C:421:MET:O	1:D:708:LYS:HE3	1.24	1.24
1:C:420:HIS:O	1:D:708:LYS:HE2	1.48	1.13
1:C:508:LYS:H	1:C:508:LYS:HD3	1.14	1.13
1:E:908:LYS:H	1:E:908:LYS:HD3	1.02	1.12

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:1108:LYS:HD3	1:F:1108:LYS:H	1.11	1.12

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:958:ASP:OD2	2:D:1223:HOH:O[1_455]	2.17	0.03

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	141/162 (87%)	130 (92%)	8 (6%)	3 (2%)	7	8
1	B	141/162 (87%)	129 (92%)	8 (6%)	4 (3%)	5	4
1	C	151/162 (93%)	132 (87%)	12 (8%)	7 (5%)	2	1
1	D	160/162 (99%)	141 (88%)	12 (8%)	7 (4%)	2	2
1	E	149/162 (92%)	133 (89%)	13 (9%)	3 (2%)	7	9
1	F	154/162 (95%)	139 (90%)	8 (5%)	7 (4%)	2	2
All	All	896/972 (92%)	804 (90%)	61 (7%)	31 (4%)	3	3

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	48	VAL
1	A	127	PRO
1	B	327	PRO
1	C	413	GLU
1	C	415	LEU

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	112/130 (86%)	93 (83%)	19 (17%)	2	2
1	B	112/130 (86%)	92 (82%)	20 (18%)	2	2
1	C	121/130 (93%)	106 (88%)	15 (12%)	4	5
1	D	130/130 (100%)	112 (86%)	18 (14%)	3	4
1	E	119/130 (92%)	103 (87%)	16 (13%)	4	4
1	F	127/130 (98%)	115 (91%)	12 (9%)	8	13
All	All	721/780 (92%)	621 (86%)	100 (14%)	3	4

5 of 100 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	628	GLN
1	D	757	LEU
1	F	1157	LEU
1	D	644	THR
1	D	700	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	857	ASN
1	F	1018	GLN
1	F	1139	GLN
1	F	1057	ASN
1	E	939	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 ⓘ

There are no ligands in this entry.

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, 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	143/162 (88%)	-0.13	1 (0%)	87	86	24, 46, 87, 139	0
1	B	143/162 (88%)	-0.14	0	100	100	26, 47, 92, 130	0
1	C	153/162 (94%)	0.15	8 (5%)	27	26	39, 67, 124, 152	0
1	D	162/162 (100%)	0.07	3 (1%)	66	64	35, 57, 104, 164	0
1	E	151/162 (93%)	0.01	6 (3%)	38	37	26, 53, 128, 148	0
1	F	158/162 (97%)	0.18	9 (5%)	23	22	27, 56, 141, 172	0
All	All	910/972 (93%)	0.03	27 (2%)	50	49	24, 54, 121, 172	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	1056	PHE	7.2
1	F	1058	THR	6.8
1	F	1064	LYS	5.5
1	F	1010	PRO	5.5
1	F	1111	ALA	5.5

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.