wwPDB X-ray Structure Validation Summary Report (i)

May 25, 2020 – 07:14 am BST

PDB ID	:	3ZFX
Title	:	Crystal structure of EphB1
Authors	:	Debreczeni, J.E.; Overman, R.; Truman, C.; McAlister, M.; Attwood, T.K.
Deposited on	:	2012-12-12
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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

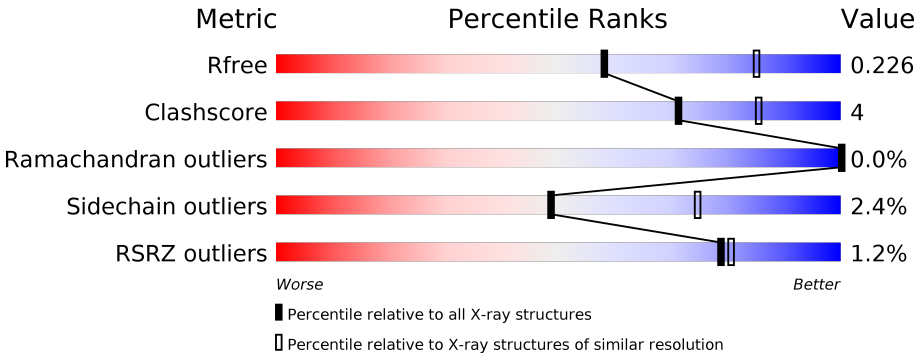
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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.



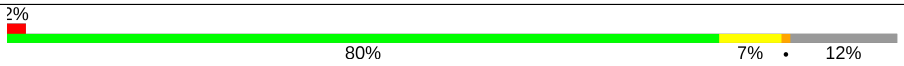
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 <=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	298	77% 12% 10%
1	B	298	79% 6% 15%
1	C	298	2% 76% 7% 17%
1	D	298	78% 10% 11%
1	E	298	78% 9% 12%
1	F	298	3% 79% 6% 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	298	
1	H	298	
1	I	298	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 18158 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 EPHRIN TYPE-B RECEPTOR 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	267	Total	C	N	O	S	0	0	0
			2075	1330	350	380	15			
1	B	254	Total	C	N	O	S	0	0	0
			1961	1257	330	359	15			
1	C	248	Total	C	N	O	S	0	0	0
			1909	1223	318	353	15			
1	D	265	Total	C	N	O	S	0	0	0
			2052	1314	345	378	15			
1	E	262	Total	C	N	O	S	0	0	0
			2021	1294	339	373	15			
1	F	254	Total	C	N	O	S	0	0	0
			1946	1248	317	366	15			
1	G	253	Total	C	N	O	S	0	0	0
			1974	1261	332	366	15			
1	H	266	Total	C	N	O	S	0	0	0
			2051	1316	344	376	15			
1	I	263	Total	C	N	O	S	0	0	0
			2030	1300	343	372	15			

There are 27 discrepancies between the modelled and reference sequences:

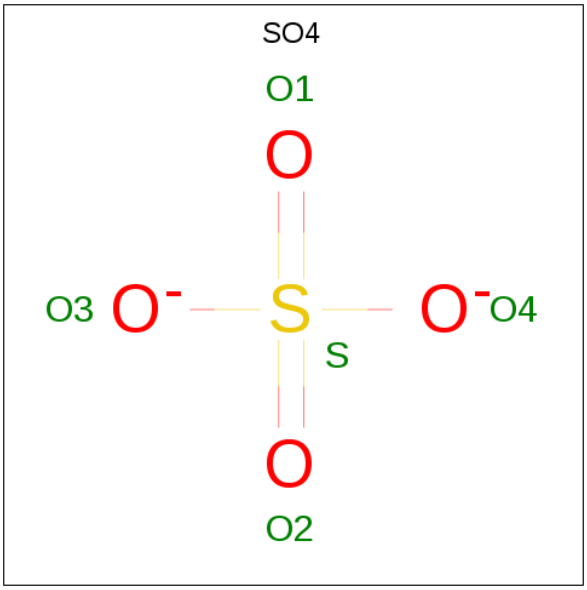
Chain	Residue	Modelled	Actual	Comment	Reference
A	599	GLY	-	expression tag	UNP P54762
A	600	SER	-	expression tag	UNP P54762
A	601	SER	-	expression tag	UNP P54762
B	599	GLY	-	expression tag	UNP P54762
B	600	SER	-	expression tag	UNP P54762
B	601	SER	-	expression tag	UNP P54762
C	599	GLY	-	expression tag	UNP P54762
C	600	SER	-	expression tag	UNP P54762
C	601	SER	-	expression tag	UNP P54762
D	599	GLY	-	expression tag	UNP P54762
D	600	SER	-	expression tag	UNP P54762

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	601	SER	-	expression tag	UNP P54762
E	599	GLY	-	expression tag	UNP P54762
E	600	SER	-	expression tag	UNP P54762
E	601	SER	-	expression tag	UNP P54762
F	599	GLY	-	expression tag	UNP P54762
F	600	SER	-	expression tag	UNP P54762
F	601	SER	-	expression tag	UNP P54762
G	599	GLY	-	expression tag	UNP P54762
G	600	SER	-	expression tag	UNP P54762
G	601	SER	-	expression tag	UNP P54762
H	599	GLY	-	expression tag	UNP P54762
H	600	SER	-	expression tag	UNP P54762
H	601	SER	-	expression tag	UNP P54762
I	599	GLY	-	expression tag	UNP P54762
I	600	SER	-	expression tag	UNP P54762
I	601	SER	-	expression tag	UNP P54762

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0

- Molecule 3 is water.

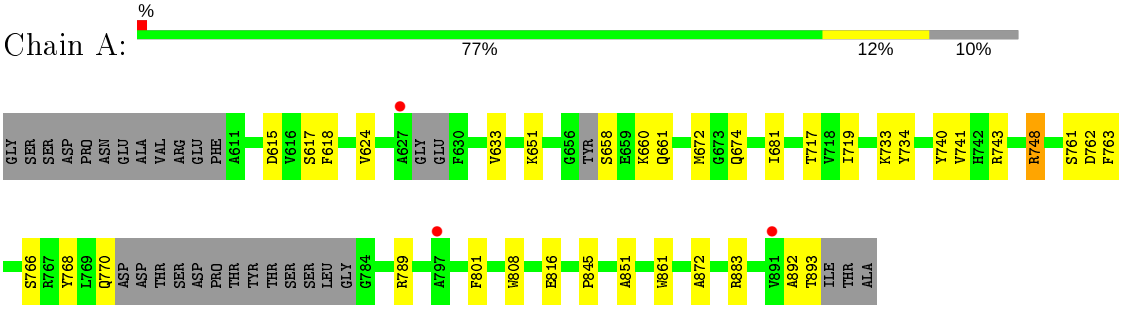
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	52	Total O 52 52	0	0
3	B	7	Total O 7 7	0	0
3	C	1	Total O 1 1	0	0
3	D	39	Total O 39 39	0	0
3	E	3	Total O 3 3	0	0
3	F	4	Total O 4 4	0	0
3	G	5	Total O 5 5	0	0
3	H	17	Total O 17 17	0	0
3	I	6	Total O 6 6	0	0

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.

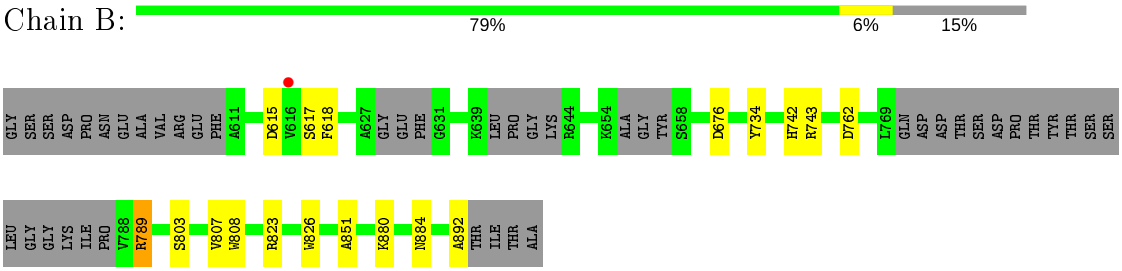
- Molecule 1: EPHRIN TYPE-B RECEPTOR 1

Chain A:



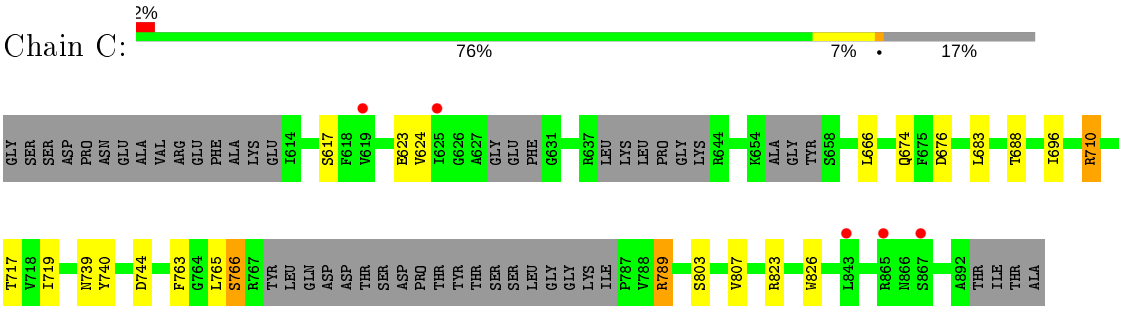
- Molecule 1: EPHRIN TYPE-B RECEPTOR 1

Chain B:



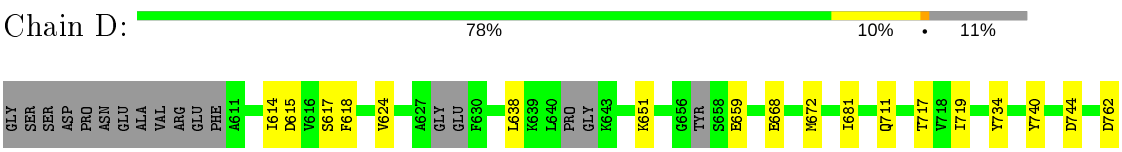
- Molecule 1: EPHRIN TYPE-B RECEPTOR 1

Chain C:



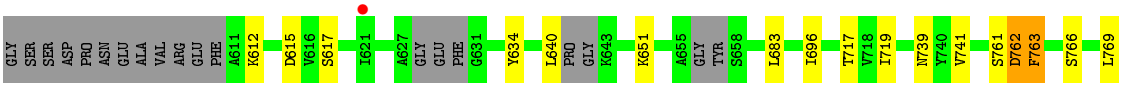
- Molecule 1: EPHRIN TYPE-B RECEPTOR 1

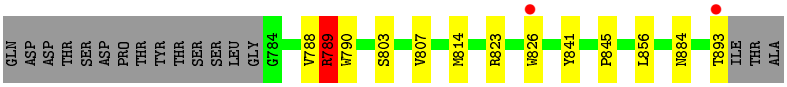
Chain D:



• Molecule 1: EPHRIN TYPE-B RECEPTOR 1

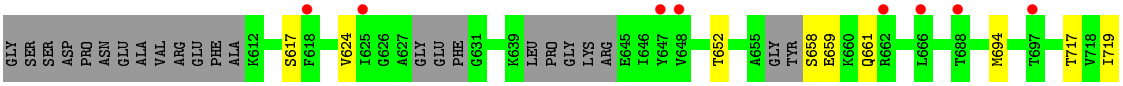
Chain E:





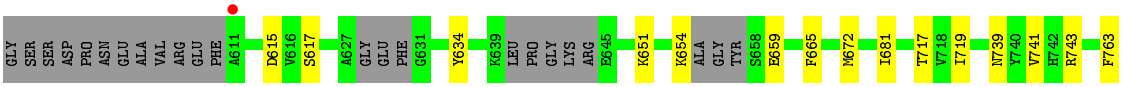
• Molecule 1: EPHRIN TYPE-B RECEPTOR 1

Chain F:



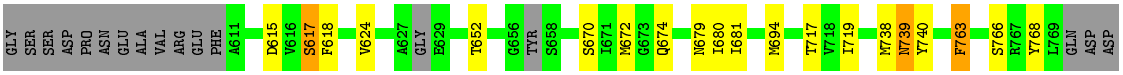
• Molecule 1: EPHRIN TYPE-B RECEPTOR 1

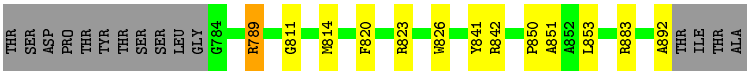
Chain G:



• Molecule 1: EPHRIN TYPE-B RECEPTOR 1

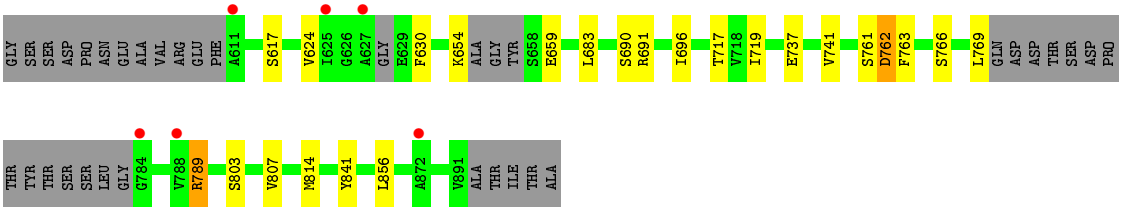
Chain H:





• Molecule 1: EPHRIN TYPE-B RECEPTOR 1

Chain I:


GLY SER SER ASP PRO ASN GLU ALA VAL ARG GLU PHE A611 S617 V624 I625 G626 A627 GLY E629 F630 K654 ALA GLY TYR S658 E659 L683 S690 R691 I696 T717 V718 I719 E737 V741 S761 D762 F763 S766 L769 GLN ASP ASP THR SER ASP PRO

THR TYR THR SER SER LEU GLY G784 V788 R789 S803 V807 M814 Y841 L856 A872 V891 ALA THR ILE THR ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	195.98Å 195.98Å 60.24Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	15.00 – 2.50 84.86 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.8 (15.00-2.50) 100.0 (84.86-2.50)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.93 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.188 , 0.217 0.203 , 0.226	Depositor DCC
R_{free} test set	4548 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	32.7	Xtriage
Anisotropy	0.149	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 43.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.000 for -h,-k,l 0.000 for h,-h-k,-l 0.097 for -k,-h,-l	Xtriage
Reported twinning fraction	0.031 for H, K, L 0.030 for -K, -H, -L 0.432 for -h,-k,l 0.507 for K, H, -L	Depositor
Outliers	0 of 89479 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	18158	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 36.38 % 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 5.0458e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: SO4

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	0.70	2/2116 (0.1%)	0.95	6/2863 (0.2%)
1	B	0.77	4/1998 (0.2%)	0.74	2/2705 (0.1%)
1	C	0.50	0/1946	0.93	5/2638 (0.2%)
1	D	0.65	0/2090	0.86	4/2828 (0.1%)
1	E	0.75	8/2059 (0.4%)	0.74	3/2790 (0.1%)
1	F	0.49	0/1983	0.69	2/2692 (0.1%)
1	G	0.86	12/2011 (0.6%)	0.81	3/2723 (0.1%)
1	H	0.81	8/2092 (0.4%)	0.84	3/2833 (0.1%)
1	I	0.65	4/2070 (0.2%)	0.75	3/2804 (0.1%)
All	All	0.70	38/18365 (0.2%)	0.82	31/24876 (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	C	0	1

The worst 5 of 38 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	734	TYR	CE1-CZ	-14.83	1.19	1.38
1	E	634	TYR	CE1-CZ	-12.59	1.22	1.38
1	I	841	TYR	CE1-CZ	-12.56	1.22	1.38
1	B	734	TYR	CG-CD2	-11.71	1.24	1.39
1	G	841	TYR	CE1-CZ	-11.59	1.23	1.38

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	710	ARG	NE-CZ-NH2	-24.68	107.96	120.30
1	A	748	ARG	NE-CZ-NH1	19.30	129.95	120.30
1	C	710	ARG	NE-CZ-NH1	17.56	129.08	120.30
1	D	789	ARG	NE-CZ-NH2	-15.96	112.32	120.30
1	I	789	ARG	NE-CZ-NH2	-15.64	112.48	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	710	ARG	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	2075	0	2028	21	1
1	B	1961	0	1896	7	1
1	C	1909	0	1831	11	0
1	D	2052	0	1999	15	2
1	E	2021	0	1953	15	1
1	F	1946	0	1850	9	1
1	G	1974	0	1910	13	0
1	H	2051	0	1993	19	0
1	I	2030	0	1969	18	0
2	A	5	0	0	0	0
3	A	52	0	0	4	0
3	B	7	0	0	0	0
3	C	1	0	0	0	0
3	D	39	0	0	0	0
3	E	3	0	0	0	0
3	F	4	0	0	0	0
3	G	5	0	0	0	0
3	H	17	0	0	0	0
3	I	6	0	0	0	0
All	All	18158	0	17429	124	3

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 4.

The worst 5 of 124 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:I:741:VAL:HG23	1:I:769:LEU:HD11	1.49	0.94
1:D:651:LYS:NZ	1:D:668:GLU:OE1	2.09	0.84
1:G:741:VAL:HG12	1:G:743:ARG:HG2	1.60	0.83
1:A:816:GLU:OE2	3:A:2026:HOH:O	2.03	0.75
1:I:741:VAL:CG2	1:I:769:LEU:HD11	2.17	0.75

All (3) 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:A:734:TYR:OH	1:B:884:ASN:ND2[1_556]	1.93	0.27
1:D:734:TYR:OH	1:E:884:ASN:OD1[3_554]	2.09	0.11
1:D:865:ARG:NH1	1:F:768:TYR:OH[1_554]	2.18	0.02

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	259/298 (87%)	248 (96%)	11 (4%)	0	100	100
1	B	244/298 (82%)	237 (97%)	7 (3%)	0	100	100
1	C	238/298 (80%)	230 (97%)	8 (3%)	0	100	100
1	D	255/298 (86%)	245 (96%)	10 (4%)	0	100	100
1	E	252/298 (85%)	240 (95%)	11 (4%)	1 (0%)	34	54
1	F	244/298 (82%)	237 (97%)	7 (3%)	0	100	100
1	G	243/298 (82%)	237 (98%)	6 (2%)	0	100	100
1	H	258/298 (87%)	251 (97%)	7 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	I	255/298 (86%)	246 (96%)	9 (4%)	0	100	100
All	All	2248/2682 (84%)	2171 (97%)	76 (3%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	788	VAL

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	214/256 (84%)	211 (99%)	3 (1%)	67	86
1	B	200/256 (78%)	196 (98%)	4 (2%)	55	79
1	C	195/256 (76%)	190 (97%)	5 (3%)	46	72
1	D	211/256 (82%)	204 (97%)	7 (3%)	38	64
1	E	207/256 (81%)	199 (96%)	8 (4%)	32	57
1	F	198/256 (77%)	194 (98%)	4 (2%)	55	79
1	G	205/256 (80%)	200 (98%)	5 (2%)	49	74
1	H	210/256 (82%)	206 (98%)	4 (2%)	57	80
1	I	209/256 (82%)	204 (98%)	5 (2%)	49	74
All	All	1849/2304 (80%)	1804 (98%)	45 (2%)	49	74

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	615	ASP
1	E	762	ASP
1	I	659	GLU
1	E	640	LEU
1	E	789	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	862	GLN
1	F	862	GLN
1	G	862	GLN
1	H	862	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	SO4	A	1890	-	4,4,4	0.72	0	6,6,6	0.28	0

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	267/298 (89%)	0.02	3 (1%) 80 82	17, 26, 40, 44	0
1	B	254/298 (85%)	-0.00	1 (0%) 92 93	28, 35, 46, 49	0
1	C	248/298 (83%)	0.10	5 (2%) 65 68	35, 41, 48, 52	0
1	D	265/298 (88%)	-0.13	1 (0%) 92 93	14, 23, 33, 42	0
1	E	262/298 (87%)	0.02	3 (1%) 80 82	28, 36, 47, 53	0
1	F	254/298 (85%)	0.21	8 (3%) 49 52	34, 40, 52, 58	0
1	G	253/298 (84%)	-0.01	1 (0%) 92 93	26, 34, 47, 52	0
1	H	266/298 (89%)	-0.18	0 100 100	14, 24, 34, 42	0
1	I	263/298 (88%)	0.35	6 (2%) 60 63	37, 43, 51, 56	0
All	All	2332/2682 (86%)	0.04	28 (1%) 79 80	14, 35, 48, 58	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	611	ALA	3.8
1	F	647	TYR	3.6
1	E	893	THR	3.3
1	I	784	GLY	3.2
1	I	625	ILE	3.2

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 carbohydrates in this entry.

6.4 Ligands

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, 95^{th} 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
2	SO4	A	1890	5/5	0.96	0.10	25,28,32,38	0

6.5 Other polymers

There are no such residues in this entry.