W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 4, 2023 – 03:42 PM EDT

PDB ID	:	6OAO
Title	:	Structure of DBP in complex with human neutralizing antibody 092096
Authors	:	Urusova, D.; Tolia, N.H.
Deposited on	:	2019-03-18
Resolution	:	3.50 Å(reported)

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

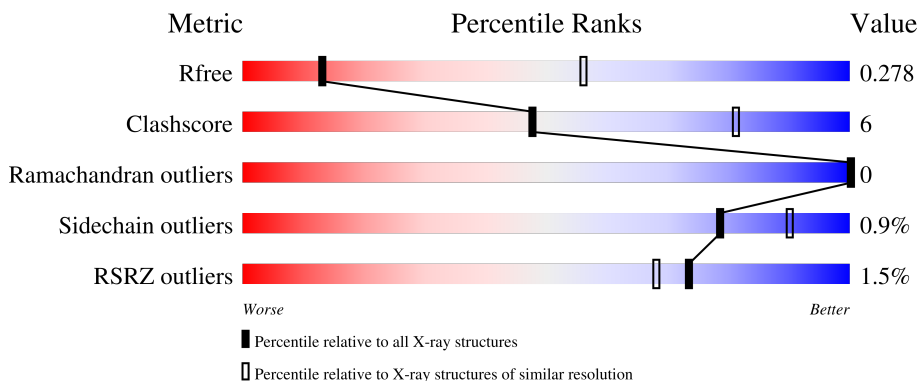|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.50 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	Percentile Ranks	Value
Rfree		0.278
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		0.9%
RSRZ outliers		1.5%

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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	315	% 79% · 10% · 10%
1	C	315	2% 78% · 13% · 9%
1	E	315	78% · 13% · 10%
1	G	315	% 80% · 10% · 10%
1	I	315	2% 85% · 6% · 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	315	.% 77% 14% 10%
2	B	251	3% 76% 14% • 10%
2	D	251	.% 77% 13% 10%
2	F	251	3% 80% 11% 9%
2	H	251	2% 78% 12% 9%
2	J	251	79% 12% 9%
2	L	251	.% 82% 8% 10%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 24648 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 Duffy binding surface protein region II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	284	Total 2374	C 1507	N 418	O 432	S 17	0	0	0
1	C	288	Total 2405	C 1523	N 424	O 441	S 17	0	0	0
1	E	285	Total 2383	C 1513	N 420	O 433	S 17	0	0	0
1	G	284	Total 2374	C 1507	N 418	O 432	S 17	0	0	0
1	I	287	Total 2397	C 1519	N 423	O 438	S 17	0	0	0
1	K	285	Total 2383	C 1513	N 420	O 433	S 17	0	0	0

- Molecule 2 is a protein called Antibody 092096 single chain variable fragment.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	227	Total 1712	C 1065	N 288	O 351	S 8	0	0	0
2	D	227	Total 1712	C 1065	N 288	O 351	S 8	0	0	0
2	F	229	Total 1725	C 1072	N 290	O 355	S 8	0	0	0
2	H	229	Total 1728	C 1074	N 292	O 354	S 8	0	0	0
2	J	228	Total 1718	C 1068	N 289	O 353	S 8	0	0	0
2	L	227	Total 1712	C 1065	N 288	O 351	S 8	0	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O3 O⁻ — S — O⁻ O4

S

O2

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	E	1	Total 5	O 4	S 1	0	0

- Molecule 4 is water.

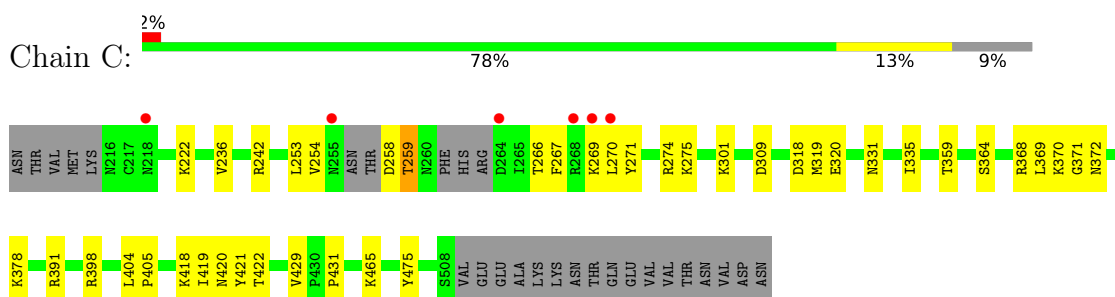| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	5	Total 5	O 5	0	0
4	B	1	Total 1	O 1	0	0
4	C	1	Total 1	O 1	0	0
4	E	3	Total 3	O 3	0	0
4	F	2	Total 2	O 2	0	0
4	G	1	Total 1	O 1	0	0
4	I	2	Total 2	O 2	0	0
4	K	4	Total 4	O 4	0	0
4	L	1	Total 1	O 1	0	0

3 Residue-property plots ⓘ

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.
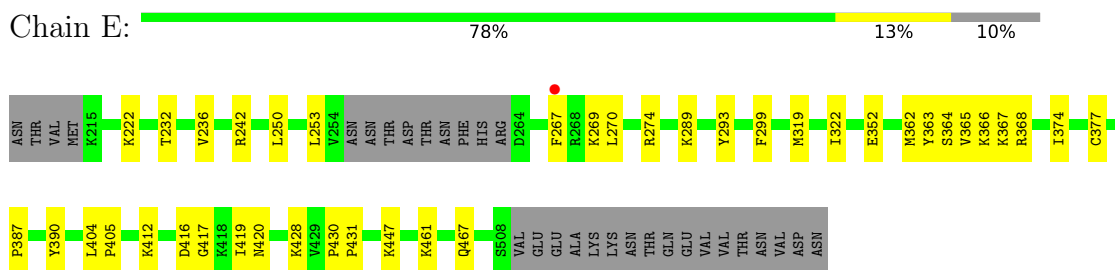
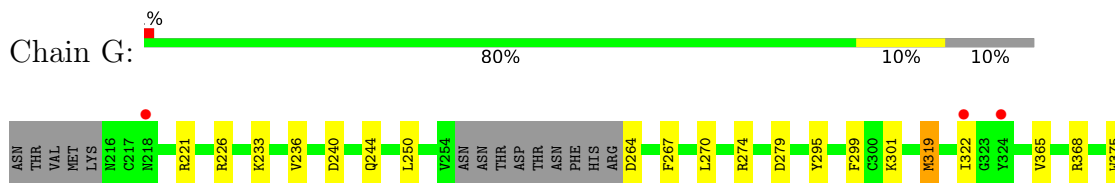• Molecule 1: Duffy binding surface protein region II

• Molecule 1: Duffy binding surface protein region II

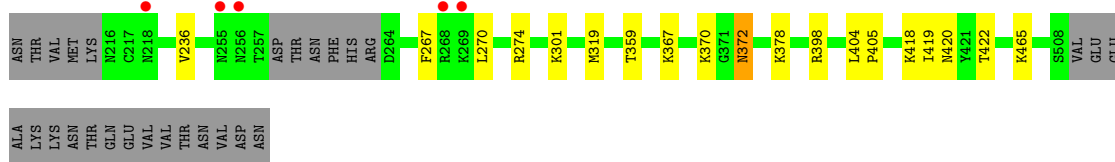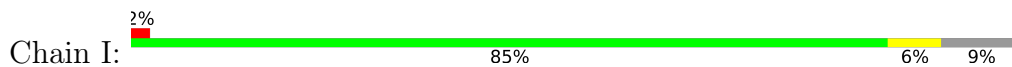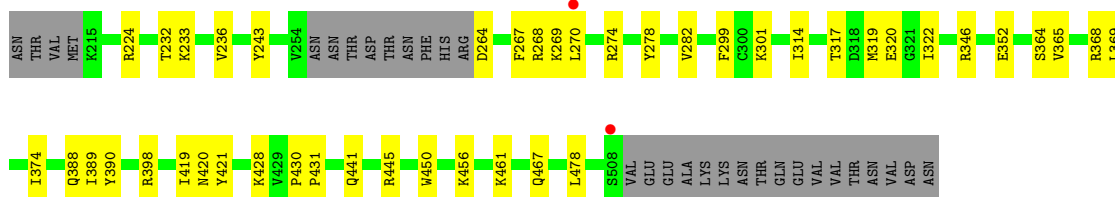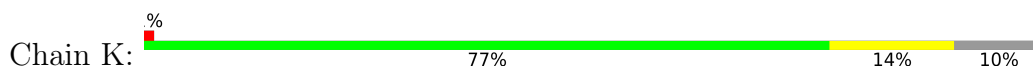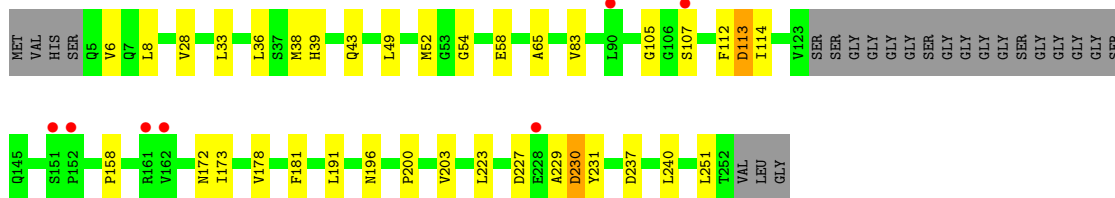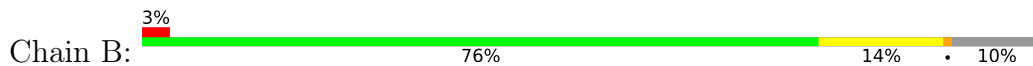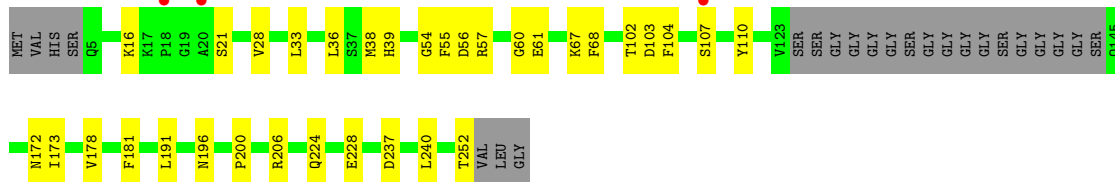• Molecule 1: Duffy binding surface protein region II

• Molecule 1: Duffy binding surface protein region II

I376 R398 L404 P405 I419 N420 K428 V429 P430 P431 N462 T468 S508 VAL GLU GLU ALA LYS LYS ASN THR GLN GLU VAL VAL THR ASN VAL ASP ASN
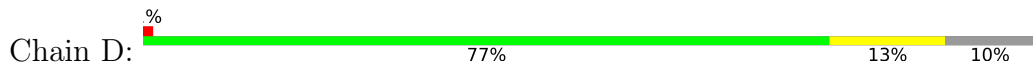
• Molecule 1: Duffy binding surface protein region II

Chain I:

2%

85% 6% 9%

ASN THR VAL MET LYS N216 C217 N218 V236 N255 N256 T257 ASP THR ASN PHE HIS ARG D264 F267 R268 K269 L270 R274 K301 M319 T359 K367 K370 G371 N372 K378 R398 L404 P405 K418 I419 N420 Y421 T422 K465 S508 VAL GLU GLU

ALA LYS LYS ASN THR GLN GLU VAL VAL THR ASN VAL ASP ASN

• Molecule 1: Duffy binding surface protein region II

Chain K:

%

77% 14% 10%

ASN THR VAL MET K215 R224 T232 K233 V236 Y243 V254 ASN ASN THR ASP THR ASN PHE HIS ARG D264 F267 R268 K269 L270 R274 Y278 V282 F299 C300 K301 I314 T317 D318 M319 E320 G321 I322 R346 E352 S364 V365 R368 L369

I374 Q388 I389 Y390 R398 I419 N420 Y421 K428 V429 P430 P431 Q441 R445 W450 K456 K461 Q467 L478 S508 VAL GLU GLU ALA LYS LYS ASN THR GLN GLU VAL VAL THR ASN VAL ASP ASN

• Molecule 2: Antibody 092096 single chain variable fragment

Chain B:

3%

76% 14% • 10%

MET VAL HIS SER Q5 V6 Q7 L8 V28 L33 L36 S37 M38 H39 Q43 L49 M52 G53 G54 E58 A65 V83 L90 G105 G106 S107 F112 D113 I114 V123 SER SER GLY GLY GLY GLY SER GLY GLY GLY SER GLY GLY GLY SER

Q145 S151 P152 P158 R161 V162 N172 I173 V178 F181 L191 N196 P200 V203 L223 D227 E228 A229 D230 Y231 D237 L240 L251 T252 VAL LEU GLY

• Molecule 2: Antibody 092096 single chain variable fragment

Chain D:

%

77% 13% 10%

MET VAL HIS SER Q5 K16 K17 P18 G19 A20 S21 V28 L33 L36 S37 M38 H39 G54 F55 D56 R57 G60 E61 K67 F68 T102 D103 F104 S107 Y110 V123 SER SER GLY GLY GLY GLY SER GLY GLY GLY SER GLY GLY GLY SER GLY GLY GLY SER Q145

N172 I173 V178 F181 L191 N196 P200 R206 Q224 E228 D237 L240 T252 VAL LEU GLY

• Molecule 2: Antibody 092096 single chain variable fragment
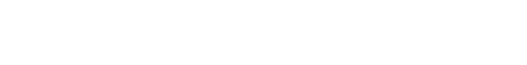
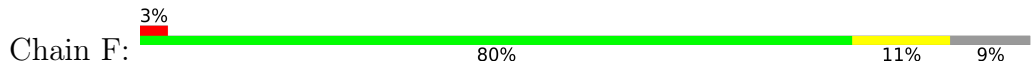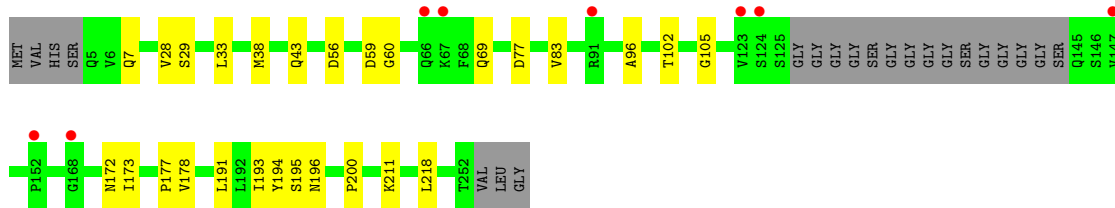Chain F:

3%

80% 11% 9%

• Molecule 2: Antibody 092096 single chain variable fragment

Chain H:
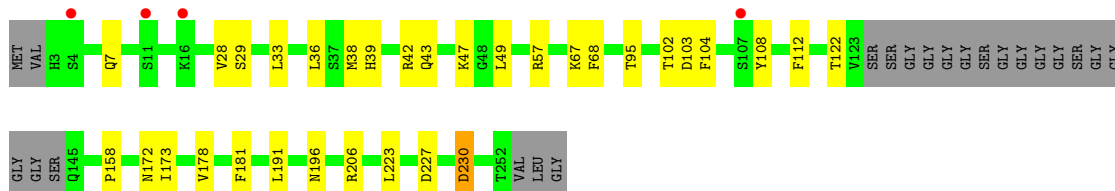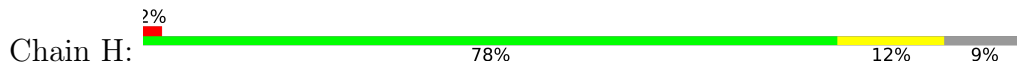

• Molecule 2: Antibody 092096 single chain variable fragment

Chain J:
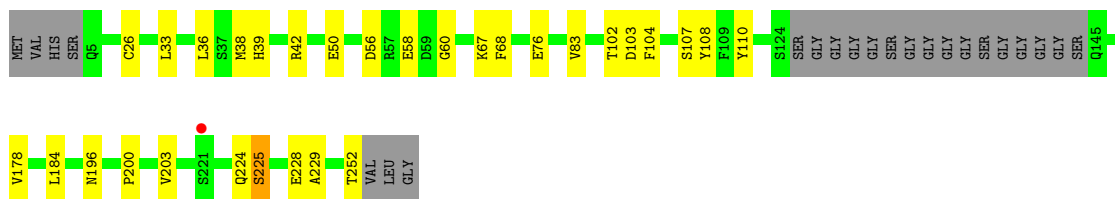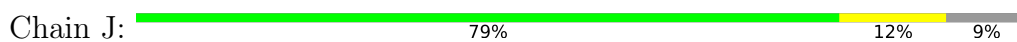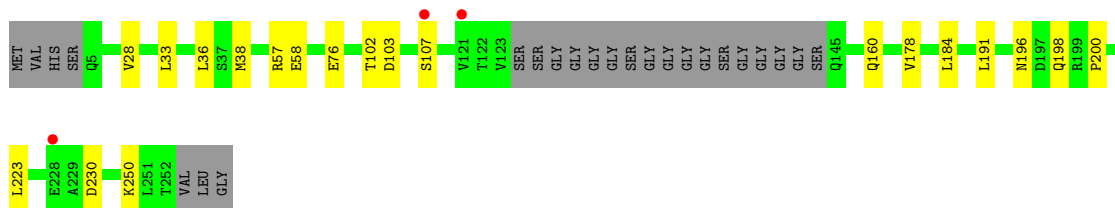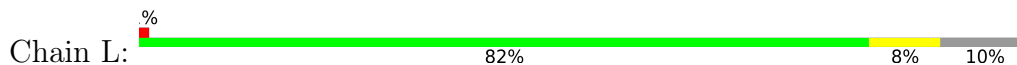

• Molecule 2: Antibody 092096 single chain variable fragment

Chain L:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	130.66Å 89.56Å 193.71Å 90.00° 108.05° 90.00°	Depositor
Resolution (Å)	19.97 – 3.50 19.97 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.6 (19.97-3.50) 99.7 (19.97-3.50)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.27 (at 3.52Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.237 , 0.279 0.240 , 0.278	Depositor DCC
R_{free} test set	2787 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å^2)	78.3	Xtriage
Anisotropy	0.308	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 17.7	EDS
L-test for twinning [2]	$<|L|>$ = 0.47, $<L^2>$ = 0.30	Xtriage
Estimated twinning fraction	0.073 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	24648	wwPDB-VP
Average B, all atoms (Å^2)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.67% 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: 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	0.34	0/2420	0.46	0/3254
1	C	0.35	0/2450	0.47	0/3294
1	E	0.37	0/2429	0.48	0/3265
1	G	0.33	0/2420	0.47	0/3254
1	I	0.32	0/2443	0.47	0/3286
1	K	0.38	0/2429	0.49	0/3265
2	B	0.37	0/1749	0.57	0/2371
2	D	0.41	0/1749	0.59	0/2371
2	F	0.49	1/1762 (0.1%)	0.61	0/2389
2	H	0.36	0/1766	0.56	0/2394
2	J	0.39	0/1755	0.54	0/2379
2	L	0.40	0/1749	0.56	0/2371
All	All	0.37	1/25121 (0.0%)	0.52	0/33893

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	69	GLN	C-N	-5.06	1.24	1.33

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	2374	0	2371	32	0
1	C	2405	0	2393	37	0
1	E	2383	0	2384	35	0
1	G	2374	0	2371	21	0
1	I	2397	0	2390	13	0
1	K	2383	0	2384	31	0
2	B	1712	0	1608	32	0
2	D	1712	0	1608	26	0
2	F	1725	0	1622	23	0
2	H	1728	0	1620	23	0
2	J	1718	0	1613	24	0
2	L	1712	0	1608	21	0
3	E	5	0	0	1	0
4	A	5	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	E	3	0	0	1	0
4	F	2	0	0	0	0
4	G	1	0	0	0	0
4	I	2	0	0	0	0
4	K	4	0	0	0	0
4	L	1	0	0	0	0
All	All	24648	0	23972	294	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 294 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:E:362:MET:HE3	1:E:377:CYS:HB2	1.39	1.02
1:E:364:SER:O	1:E:368:ARG:HG3	1.70	0.91
2:L:178:VAL:HB	2:L:196:ASN:ND2	1.88	0.89
1:E:362:MET:HE3	1:E:377:CYS:CB	2.07	0.84
1:E:363:TYR:HD1	2:F:105:GLY:O	1.59	0.83

There are no symmetry-related clashes.

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	280/315 (89%)	269 (96%)	11 (4%)	0	100	100
1	C	282/315 (90%)	268 (95%)	14 (5%)	0	100	100
1	E	281/315 (89%)	267 (95%)	14 (5%)	0	100	100
1	G	280/315 (89%)	267 (95%)	13 (5%)	0	100	100
1	I	283/315 (90%)	269 (95%)	14 (5%)	0	100	100
1	K	281/315 (89%)	267 (95%)	14 (5%)	0	100	100
2	B	223/251 (89%)	206 (92%)	17 (8%)	0	100	100
2	D	223/251 (89%)	205 (92%)	18 (8%)	0	100	100
2	F	225/251 (90%)	208 (92%)	17 (8%)	0	100	100
2	H	225/251 (90%)	210 (93%)	15 (7%)	0	100	100
2	J	224/251 (89%)	205 (92%)	19 (8%)	0	100	100
2	L	223/251 (89%)	205 (92%)	18 (8%)	0	100	100
All	All	3030/3396 (89%)	2846 (94%)	184 (6%)	0	100	100

There are no Ramachandran outliers to report.

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	261/291 (90%)	256 (98%)	5 (2%)	57	80
1	C	265/291 (91%)	261 (98%)	4 (2%)	65	84

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	262/291 (90%)	261 (100%)	1 (0%)	91	96
1	G	261/291 (90%)	257 (98%)	4 (2%)	65	84
1	I	264/291 (91%)	262 (99%)	2 (1%)	81	91
1	K	262/291 (90%)	260 (99%)	2 (1%)	81	91
2	B	188/199 (94%)	186 (99%)	2 (1%)	73	88
2	D	188/199 (94%)	188 (100%)	0	100	100
2	F	190/199 (96%)	190 (100%)	0	100	100
2	H	190/199 (96%)	188 (99%)	2 (1%)	73	88
2	J	189/199 (95%)	188 (100%)	1 (0%)	88	94
2	L	188/199 (94%)	187 (100%)	1 (0%)	88	94
All	All	2708/2940 (92%)	2684 (99%)	24 (1%)	78	90

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	319	MET
2	H	230	ASP
2	H	28	VAL
1	I	267	PHE
2	B	230	ASP

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

Mol	Chain	Res	Type
2	B	39	HIS
2	L	196	ASN

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 (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

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	#$|Z| > 2$
3	SO4	E	601	-	4,4,4	1.24	0	6,6,6	0.45	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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	601	SO4	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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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	284/315 (90%)	0.06	4 (1%)	75	69	67, 84, 113, 122	0
1	C	288/315 (91%)	-0.02	6 (2%)	63	58	56, 77, 107, 134	0
1	E	285/315 (90%)	-0.07	1 (0%)	92	90	53, 71, 101, 126	0
1	G	284/315 (90%)	0.04	3 (1%)	80	75	63, 82, 108, 116	0
1	I	287/315 (91%)	-0.01	5 (1%)	70	64	56, 75, 108, 127	0
1	K	285/315 (90%)	-0.04	2 (0%)	87	83	49, 70, 103, 119	0
2	B	227/251 (90%)	0.22	7 (3%)	49	43	82, 109, 130, 138	0
2	D	227/251 (90%)	0.10	3 (1%)	77	71	64, 85, 102, 113	0
2	F	229/251 (91%)	0.26	8 (3%)	44	39	75, 99, 129, 155	0
2	H	229/251 (91%)	0.20	4 (1%)	70	64	79, 109, 131, 138	0
2	J	228/251 (90%)	0.14	1 (0%)	92	90	67, 92, 117, 137	0
2	L	227/251 (90%)	0.10	3 (1%)	77	71	71, 94, 127, 147	0
All	All	3080/3396 (90%)	0.07	47 (1%)	73	68	49, 86, 122, 155	0

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	255	ASN	4.7
2	B	228	GLU	4.6
1	C	270	LEU	3.7
1	C	268	ARG	3.7
2	D	107	SER	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 monosaccharides 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, 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
3	SO4	E	601	5/5	0.89	0.31	30,30,30,30	0

6.5 Other polymers ⓘ

There are no such residues in this entry.