Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 06:45 AM GMT

PDB ID	:	2Y6E
Title	:	Structure of the D1D2 domain of USP4, the conserved catalytic domain
Authors	:	Luna-Vargas, M.P.A.; Faesen, A.C.; van Dijk, W.J.; Rape, M.; Fish, A.; Sixma, T.K.
Deposited on	:	2011-01-20
Resolution	:	2.40 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
http://wwpdb.org/validation/2016/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:

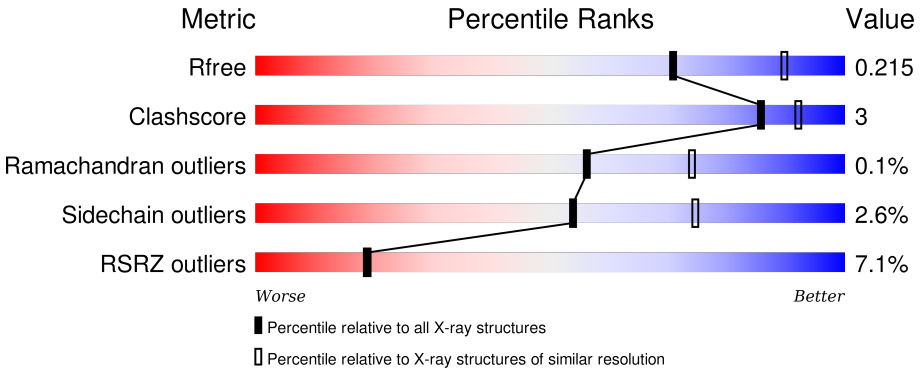
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.40 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.



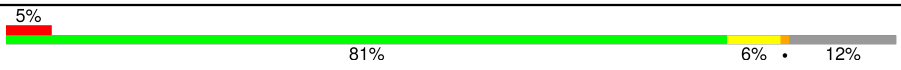
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2919 (2.40-2.40)
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (2.40-2.40)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	367	8% 85% 6% • 8%
1	B	367	5% 86% 7% 8%
1	C	367	4% 85% 5% 10%
1	D	367	10% 80% 9% • 11%
1	E	367	6% 81% 6% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	367	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
1	CME	D	311	-	-	X	-
1	CME	E	311	-	-	X	-
3	SO4	C	1925	-	-	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 17209 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 UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	337	Total	C	N	O	S	0	1	0
			2734	1750	464	502	18			
1	B	339	Total	C	N	O	S	0	0	0
			2738	1752	465	504	17			
1	C	331	Total	C	N	O	S	0	0	0
			2666	1709	453	487	17			
1	D	328	Total	C	N	O	S	0	0	0
			2646	1697	446	485	18			
1	E	319	Total	C	N	O	S	0	0	0
			2569	1649	432	473	15			
1	F	323	Total	C	N	O	S	0	0	0
			2591	1659	438	479	15			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	294	GLY	-	EXPRESSION TAG	UNP Q13107
A	295	MET	-	EXPRESSION TAG	UNP Q13107
A	491	GLY	-	LINKER	UNP Q13107
A	492	PRO	-	LINKER	UNP Q13107
B	294	GLY	-	EXPRESSION TAG	UNP Q13107
B	295	MET	-	EXPRESSION TAG	UNP Q13107
B	491	GLY	-	LINKER	UNP Q13107
B	492	PRO	-	LINKER	UNP Q13107
C	294	GLY	-	EXPRESSION TAG	UNP Q13107
C	295	MET	-	EXPRESSION TAG	UNP Q13107
C	491	GLY	-	LINKER	UNP Q13107
C	492	PRO	-	LINKER	UNP Q13107
D	294	GLY	-	EXPRESSION TAG	UNP Q13107
D	295	MET	-	EXPRESSION TAG	UNP Q13107
D	491	GLY	-	LINKER	UNP Q13107
D	492	PRO	-	LINKER	UNP Q13107
E	294	GLY	-	EXPRESSION TAG	UNP Q13107

Continued on next page...

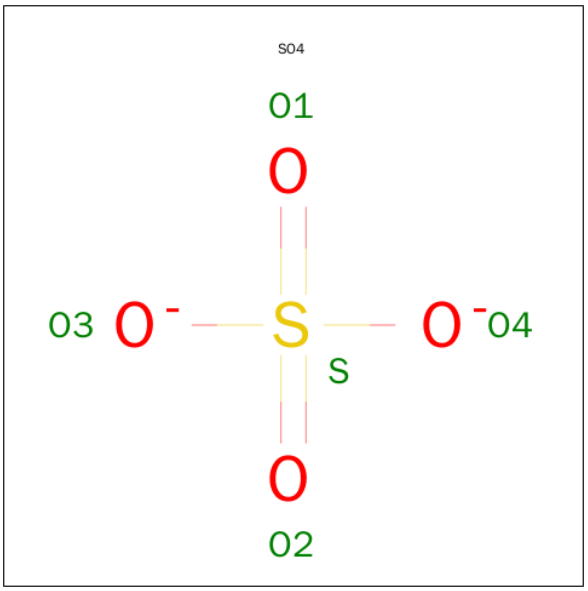
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	295	MET	-	EXPRESSION TAG	UNP Q13107
E	491	GLY	-	LINKER	UNP Q13107
E	492	PRO	-	LINKER	UNP Q13107
F	294	GLY	-	EXPRESSION TAG	UNP Q13107
F	295	MET	-	EXPRESSION TAG	UNP Q13107
F	491	GLY	-	LINKER	UNP Q13107
F	492	PRO	-	LINKER	UNP Q13107

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	1	Total Zn 1 1	0	0
2	E	1	Total Zn 1 1	0	0
2	B	1	Total Zn 1 1	0	0
2	C	1	Total Zn 1 1	0	0
2	A	1	Total Zn 1 1	0	0
2	F	1	Total Zn 1 1	0	0

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



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

- Molecule 4 is water.

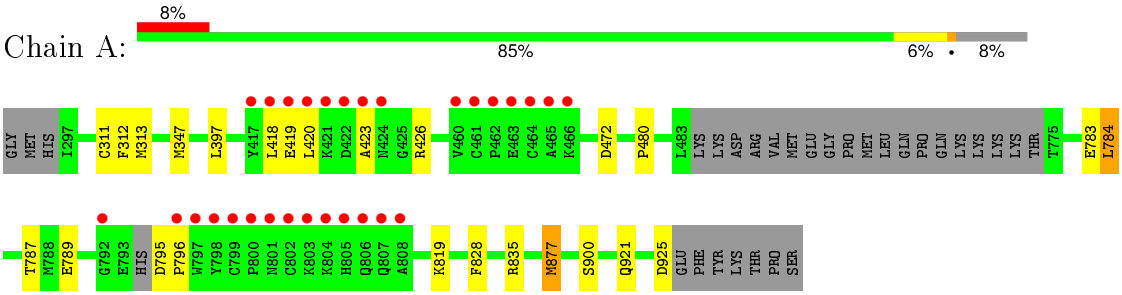
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	250	Total O 250 250	0	0
4	B	232	Total O 232 232	0	0
4	C	204	Total O 204 204	0	0
4	D	214	Total O 214 214	0	0
4	E	166	Total O 166 166	0	0
4	F	183	Total O 183 183	0	0

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 errors 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.

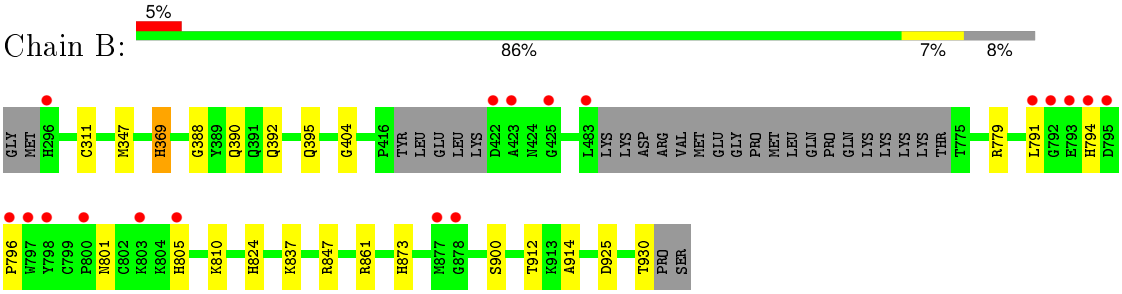
- Molecule 1: UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 4

Chain A:



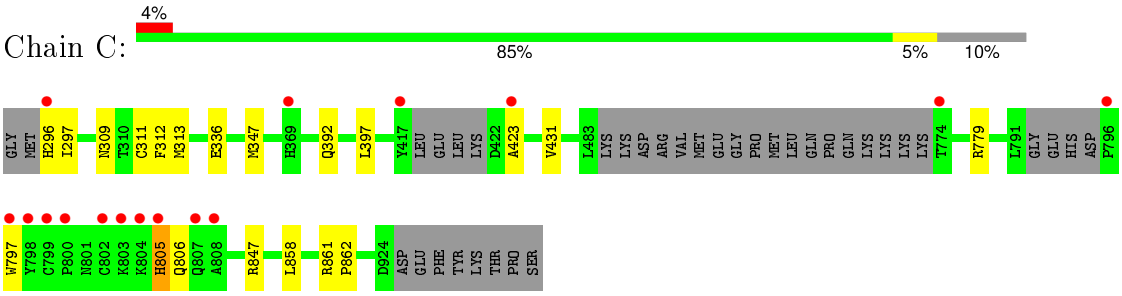
- Molecule 1: UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 4

Chain B:



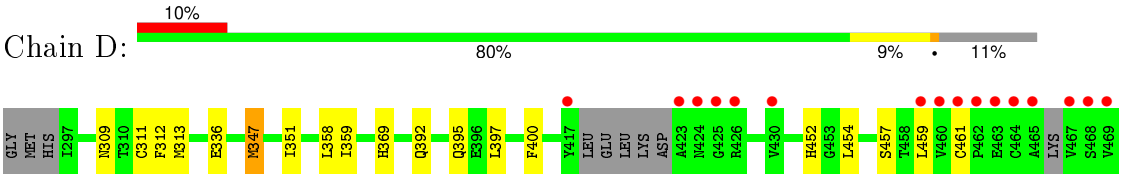
- Molecule 1: UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 4

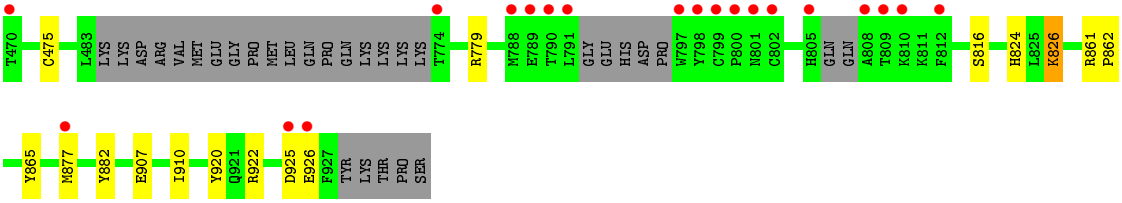
Chain C:



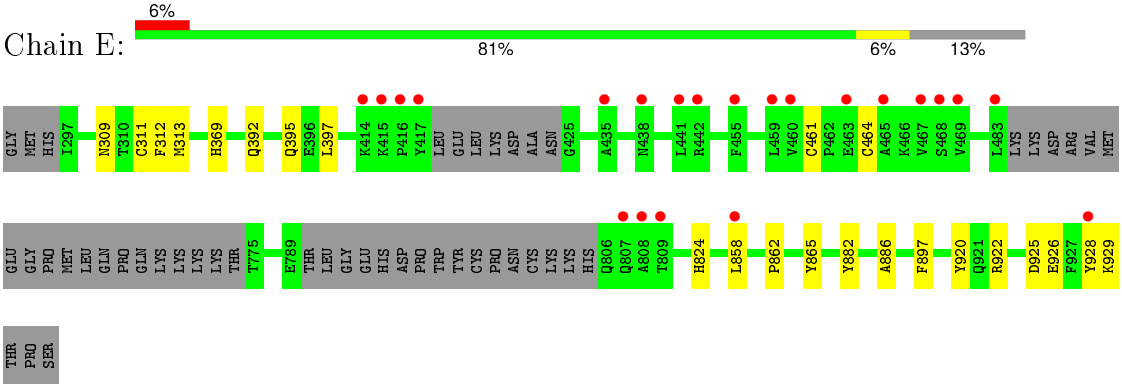
- Molecule 1: UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 4

Chain D:

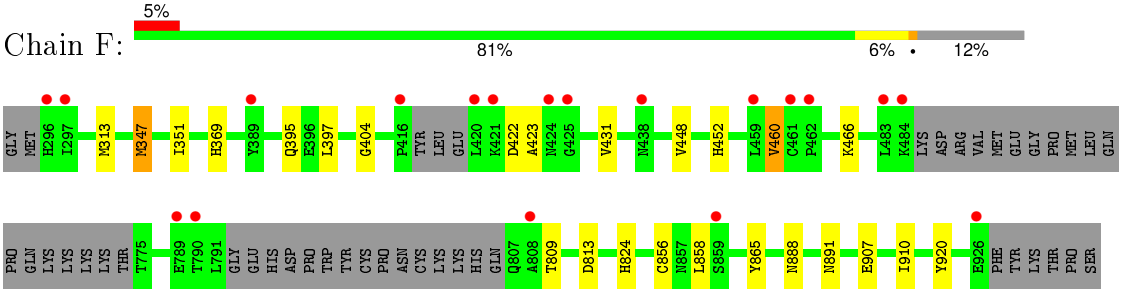




• Molecule 1: UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 4



• Molecule 1: UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 4



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	110.50Å 151.03Å 178.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.59 – 2.40 44.59 – 2.40	Depositor EDS
% Data completeness (in resolution range)	(Not available) (44.59-2.40) 94.8 (44.59-2.40)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.69 (at 2.39Å)	Xtriage
Refinement program	BUSTER 2.8.0	Depositor
R, R_{free}	0.178 , 0.210 0.183 , 0.215	Depositor DCC
R_{free} test set	5583 reflections (5.29%)	DCC
Wilson B-factor (Å^2)	40.2	Xtriage
Anisotropy	0.108	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 64.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 111095 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	17209	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.23% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: ZN, CME, 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	0.50	0/2791	0.64	0/3780
1	B	0.50	0/2798	0.63	0/3793
1	C	0.48	0/2723	0.62	0/3691
1	D	0.50	0/2700	0.65	0/3656
1	E	0.47	0/2620	0.63	0/3547
1	F	0.49	0/2641	0.64	0/3577
All	All	0.49	0/16273	0.63	0/22044

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	2734	0	2668	13	0
1	B	2738	0	2646	15	0
1	C	2666	0	2588	12	0
1	D	2646	0	2565	21	0
1	E	2569	0	2501	17	0
1	F	2591	0	2525	13	0
2	A	1	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	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	5	0	0	0	0
3	C	5	0	0	0	0
4	A	250	0	0	0	0
4	B	232	0	0	1	0
4	C	204	0	0	1	0
4	D	214	0	0	1	0
4	E	166	0	0	1	0
4	F	183	0	0	0	0
All	All	17209	0	15493	87	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 3.

All (87) 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:D:311:CME:SG	1:D:311:CME:SD	2.32	1.28
1:B:311:CME:SG	1:B:311:CME:SD	2.42	1.16
1:C:311:CME:HZ3	1:C:392:GLN:HB2	1.30	1.11
1:C:309:ASN:OD1	1:C:311:CME:HE2	1.75	0.85
1:C:311:CME:CZ	1:C:392:GLN:HB2	2.09	0.82
1:B:912:THR:HG22	1:B:914:ALA:H	1.49	0.76
1:C:311:CME:HZ3	1:C:392:GLN:CB	2.15	0.73
1:E:309:ASN:OD1	1:E:311:CME:HE2	1.91	0.70
1:C:313:MET:HA	1:C:397:LEU:HD22	1.73	0.70
1:A:877:MET:CE	1:B:369:HIS:HB2	2.21	0.70
1:E:313:MET:HA	1:E:397:LEU:HD22	1.74	0.69
1:F:313:MET:HA	1:F:397:LEU:HD22	1.74	0.68
1:A:819:LYS:HE3	1:A:921:GLN:HE21	1.60	0.65
1:A:877:MET:HE1	1:B:369:HIS:HB2	1.79	0.65
1:D:313:MET:HA	1:D:397:LEU:HD22	1.79	0.64
1:B:311:CME:HE2	1:B:392:GLN:HB2	1.79	0.64
1:F:813:ASP:HB3	1:F:856:CYS:SG	2.39	0.63
1:E:862:PRO:HD2	1:E:928:TYR:CD1	2.33	0.63
1:A:313:MET:HA	1:A:397:LEU:HD22	1.82	0.62
1:A:311:CME:HE2	1:A:312:PHE:HB3	1.81	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:395:GLN:HG3	1:E:824:HIS:CD2	2.37	0.59
1:C:313:MET:HA	1:C:397:LEU:CD2	2.32	0.59
1:B:912:THR:HG21	4:B:2209:HOH:O	2.02	0.59
1:A:877:MET:HE2	1:B:369:HIS:HB2	1.87	0.56
1:F:888:ASN:HD22	1:F:891:ASN:H	1.53	0.56
1:D:395:GLN:HG3	1:D:824:HIS:CD2	2.40	0.56
1:F:423:ALA:CB	1:F:431:VAL:HG22	2.37	0.55
1:D:347:MET:HG3	1:D:351:ILE:CG1	2.37	0.54
1:E:395:GLN:HG3	1:E:824:HIS:CG	2.42	0.54
1:D:907:GLU:HA	1:D:910:ILE:HD12	1.90	0.54
1:E:392:GLN:HG3	4:E:2048:HOH:O	2.07	0.54
1:D:311:CME:HZ3	1:D:392:GLN:HB2	1.91	0.53
1:B:837:LYS:NZ	1:B:873:HIS:HD2	2.07	0.53
1:F:448:VAL:HA	1:F:452:HIS:ND1	2.24	0.53
1:F:347:MET:HG3	1:F:351:ILE:CG1	2.39	0.53
1:E:311:CME:HZ3	1:E:392:GLN:CB	2.39	0.52
1:D:395:GLN:HG3	1:D:824:HIS:CG	2.44	0.51
1:C:805:HIS:CD2	1:C:805:HIS:H	2.29	0.49
1:E:886:ALA:HB3	1:E:897:PHE:CE1	2.48	0.49
1:A:828:PHE:CD1	1:A:835[B]:ARG:HD3	2.47	0.49
1:A:795:ASP:N	1:A:796:PRO:HD3	2.28	0.49
1:B:347:MET:HE1	1:B:404:GLY:HA2	1.96	0.48
1:C:423:ALA:HB2	1:C:431:VAL:HG11	1.96	0.48
1:B:861:ARG:HB3	1:B:930:THR:HG22	1.96	0.48
1:E:461:CYS:SG	1:E:464:CYS:HB3	2.52	0.48
1:C:336:GLU:HB2	4:C:2020:HOH:O	2.13	0.48
1:A:877:MET:HE3	1:B:900:SER:HB2	1.95	0.47
1:A:423:ALA:HA	1:A:426:ARG:HH11	1.79	0.47
1:E:311:CME:HZ2	1:E:882:TYR:CE1	2.49	0.47
1:B:388:GLY:HA3	1:B:390:GLN:OE1	2.14	0.47
1:B:395:GLN:HG2	1:B:824:HIS:CE1	2.50	0.47
1:D:309:ASN:OD1	1:D:311:CME:HE2	2.15	0.47
1:A:783:GLU:O	1:A:787:THR:HG23	2.15	0.46
1:C:296:HIS:CD2	1:C:297:ILE:HD12	2.51	0.46
1:F:907:GLU:HA	1:F:910:ILE:HD12	1.96	0.46
1:D:311:CME:SD	1:D:312:PHE:N	2.85	0.45
1:D:311:CME:HZ2	1:D:882:TYR:HE1	1.81	0.45
1:F:460:VAL:HG13	1:F:809:THR:OG1	2.16	0.45
1:D:865:TYR:HB3	1:D:920:TYR:HB3	1.99	0.44
1:F:865:TYR:HB3	1:F:920:TYR:HB3	1.98	0.44
1:F:347:MET:HG3	1:F:351:ILE:HG12	1.99	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:454:LEU:HB2	1:D:816:SER:OG	2.17	0.43
1:D:347:MET:HE2	1:D:400:PHE:CZ	2.54	0.43
1:E:311:CME:HZ2	1:E:882:TYR:HE1	1.82	0.43
1:E:865:TYR:CE2	1:E:922:ARG:HB2	2.54	0.43
1:E:865:TYR:HB3	1:E:920:TYR:HB3	2.00	0.43
1:B:791:LEU:HD21	1:B:810:LYS:HD2	2.01	0.43
1:C:311:CME:SD	1:C:312:PHE:N	2.89	0.42
1:D:336:GLU:HB2	4:D:2018:HOH:O	2.19	0.42
1:C:861:ARG:NH1	1:C:862:PRO:O	2.53	0.42
1:F:395:GLN:HG2	1:F:824:HIS:CE1	2.54	0.42
1:A:423:ALA:HA	1:A:426:ARG:HD3	2.02	0.42
1:F:347:MET:HE1	1:F:404:GLY:HA2	2.02	0.42
1:D:347:MET:HG3	1:D:351:ILE:HG12	2.01	0.42
1:E:311:CME:HZ3	1:E:392:GLN:HB2	2.02	0.42
1:D:861:ARG:NH1	1:D:862:PRO:O	2.53	0.42
1:A:480:PRO:HG2	1:A:784:LEU:HD13	2.02	0.41
1:D:311:CME:HZ3	1:D:392:GLN:H	1.84	0.41
1:E:311:CME:HZ3	1:E:392:GLN:HB3	2.02	0.41
1:E:312:PHE:CD2	1:E:392:GLN:HG2	2.55	0.41
1:D:452:HIS:CD2	1:D:475:CYS:HB3	2.56	0.41
1:D:359:ILE:HD13	1:D:359:ILE:HA	1.80	0.41
1:D:311:CME:HZ2	1:D:882:TYR:CE1	2.57	0.40
1:F:423:ALA:CB	1:F:431:VAL:CG2	3.00	0.40
1:D:861:ARG:NH1	1:D:922:ARG:HD2	2.37	0.40
1:B:796:PRO:HG2	1:B:805:HIS:HB3	2.03	0.40
1:E:312:PHE:CG	1:E:392:GLN:HG2	2.57	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	331/367 (90%)	324 (98%)	7 (2%)	0	100	100
1	B	332/367 (90%)	322 (97%)	10 (3%)	0	100	100
1	C	322/367 (88%)	316 (98%)	6 (2%)	0	100	100
1	D	315/367 (86%)	307 (98%)	7 (2%)	1 (0%)	46	63
1	E	310/367 (84%)	303 (98%)	6 (2%)	1 (0%)	46	63
1	F	314/367 (86%)	309 (98%)	5 (2%)	0	100	100
All	All	1924/2202 (87%)	1881 (98%)	41 (2%)	2 (0%)	56	74

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	925	ASP
1	D	826	LYS

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	294/321 (92%)	284 (97%)	10 (3%)	44	65
1	B	293/321 (91%)	287 (98%)	6 (2%)	63	81
1	C	285/321 (89%)	278 (98%)	7 (2%)	55	76
1	D	283/321 (88%)	272 (96%)	11 (4%)	39	59
1	E	274/321 (85%)	270 (98%)	4 (2%)	72	87
1	F	277/321 (86%)	271 (98%)	6 (2%)	60	79
All	All	1706/1926 (89%)	1662 (97%)	44 (3%)	54	74

All (44) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	347	MET
1	A	418	LEU
1	A	419	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	420	LEU
1	A	472	ASP
1	A	784	LEU
1	A	789	GLU
1	A	877	MET
1	A	900	SER
1	A	925	ASP
1	B	369	HIS
1	B	779	ARG
1	B	794	HIS
1	B	801	ASN
1	B	847	ARG
1	B	925	ASP
1	C	347	MET
1	C	779	ARG
1	C	797	TRP
1	C	805	HIS
1	C	806	GLN
1	C	847	ARG
1	C	858	LEU
1	D	347	MET
1	D	358	LEU
1	D	369	HIS
1	D	457	SER
1	D	459	LEU
1	D	461	CYS
1	D	779	ARG
1	D	826	LYS
1	D	877	MET
1	D	925	ASP
1	D	926	GLU
1	E	369	HIS
1	E	858	LEU
1	E	926	GLU
1	E	929	LYS
1	F	347	MET
1	F	369	HIS
1	F	422	ASP
1	F	460	VAL
1	F	466	LYS
1	F	858	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	857	ASN
1	A	921	GLN
1	B	873	HIS
1	C	296	HIS
1	C	805	HIS
1	D	391	GLN
1	F	296	HIS
1	F	391	GLN
1	F	807	GLN
1	F	888	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)

6 non-standard protein/DNA/RNA residues are 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	$\#\|Z\| > 2$
1	CME	A	311	1	8,9,10	1.11	1 (12%)	6,9,11	2.58	4 (66%)
1	CME	B	311	1	8,9,10	1.67	1 (12%)	6,9,11	3.83	4 (66%)
1	CME	C	311	1	8,9,10	1.10	1 (12%)	6,9,11	2.10	3 (50%)
1	CME	D	311	1	8,9,10	1.33	1 (12%)	6,9,11	2.64	3 (50%)
1	CME	E	311	1	8,9,10	1.08	1 (12%)	6,9,11	2.81	2 (33%)
1	CME	F	311	1	8,9,10	0.80	0	6,9,11	2.21	2 (33%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CME	A	311	1	-	0/5/8/10	0/0/0/0
1	CME	B	311	1	-	0/5/8/10	0/0/0/0
1	CME	C	311	1	-	0/5/8/10	0/0/0/0
1	CME	D	311	1	-	0/5/8/10	0/0/0/0
1	CME	E	311	1	-	0/5/8/10	0/0/0/0
1	CME	F	311	1	-	0/5/8/10	0/0/0/0

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	311	CME	SD-SG	2.55	2.25	2.03
1	C	311	CME	SD-SG	2.63	2.26	2.03
1	A	311	CME	SD-SG	2.73	2.27	2.03
1	D	311	CME	SD-SG	3.34	2.32	2.03
1	B	311	CME	SD-SG	4.49	2.42	2.03

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	311	CME	O-C-CA	-2.57	118.78	125.49
1	F	311	CME	O-C-CA	-2.39	119.27	125.49
1	A	311	CME	O-C-CA	-2.34	119.38	125.49
1	B	311	CME	OH-CZ-CE	-2.22	101.10	110.83
1	B	311	CME	O-C-CA	-2.16	119.85	125.49
1	C	311	CME	O-C-CA	-2.11	119.98	125.49
1	D	311	CME	O-C-CA	-2.09	120.04	125.49
1	A	311	CME	CZ-CE-SD	2.18	118.50	113.16
1	D	311	CME	CZ-CE-SD	2.38	118.99	113.16
1	C	311	CME	OH-CZ-CE	2.39	121.29	110.83
1	A	311	CME	CB-SG-SD	3.03	109.86	103.95
1	C	311	CME	CB-SG-SD	3.54	110.84	103.95
1	A	311	CME	CE-SD-SG	4.31	125.91	103.56
1	F	311	CME	CB-SG-SD	4.70	113.12	103.95
1	D	311	CME	CB-SG-SD	5.16	114.00	103.95
1	E	311	CME	CB-SG-SD	6.11	115.85	103.95
1	B	311	CME	CE-SD-SG	6.16	135.46	103.56
1	B	311	CME	CB-SG-SD	6.34	116.30	103.95

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	311	CME	1	0
1	B	311	CME	2	0
1	C	311	CME	5	0
1	D	311	CME	7	0
1	E	311	CME	6	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 6 are monoatomic - leaving 2 for Mogul analysis.

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	A	1926	-	4,4,4	0.52	0	6,6,6	0.24	0
3	SO4	C	1925	-	4,4,4	0.43	0	6,6,6	0.15	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SO4	A	1926	-	-	0/0/0/0	0/0/0/0
3	SO4	C	1925	-	-	0/0/0/0	0/0/0/0

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	336/367 (91%)	-0.03	29 (8%)	13	13	23, 39, 116, 140	0
1	B	338/367 (92%)	0.07	18 (5%)	30	30	23, 40, 103, 125	0
1	C	330/367 (89%)	-0.16	16 (4%)	34	35	24, 45, 103, 152	0
1	D	327/367 (89%)	0.15	36 (11%)	7	7	26, 45, 119, 265	0
1	E	318/367 (86%)	-0.11	22 (6%)	20	19	29, 46, 96, 139	0
1	F	322/367 (87%)	-0.17	19 (5%)	26	26	26, 49, 94, 145	0
All	All	1971/2202 (89%)	-0.04	140 (7%)	19	19	23, 45, 103, 265	0

All (140) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	417	TYR	11.6
1	C	798	TYR	10.6
1	A	422	ASP	9.1
1	A	800	PRO	8.5
1	D	800	PRO	8.4
1	A	802	CYS	8.1
1	D	461	CYS	8.0
1	A	798	TYR	8.0
1	B	794	HIS	7.8
1	D	798	TYR	7.7
1	D	802	CYS	7.7
1	A	804	LYS	7.1
1	A	805	HIS	6.7
1	D	467	VAL	6.7
1	D	468	SER	6.7
1	A	806	GLN	6.5
1	C	797	TRP	6.4
1	D	791	LEU	6.4
1	D	462	PRO	6.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	797	TRP	5.8
1	A	417	TYR	5.8
1	B	296	HIS	5.5
1	D	423	ALA	5.5
1	B	798	TYR	5.5
1	B	878	GLY	5.4
1	A	420	LEU	5.3
1	D	463	GLU	5.3
1	F	808	ALA	5.2
1	A	797	TRP	5.2
1	C	805	HIS	5.2
1	D	417	TYR	5.0
1	A	803	LYS	4.9
1	A	792	GLY	4.7
1	E	928	TYR	4.7
1	A	796	PRO	4.7
1	D	809	THR	4.7
1	E	483	LEU	4.7
1	D	460	VAL	4.7
1	A	799	CYS	4.7
1	B	792	GLY	4.6
1	A	461	CYS	4.6
1	D	926	GLU	4.6
1	C	804	LYS	4.6
1	A	419	GLU	4.6
1	D	805	HIS	4.5
1	C	807	GLN	4.5
1	D	801	ASN	4.4
1	D	424	ASN	4.4
1	E	467	VAL	4.3
1	D	465	ALA	4.3
1	A	421	LYS	4.2
1	A	418	LEU	4.2
1	F	420	LEU	4.2
1	D	808	ALA	4.2
1	A	462	PRO	4.2
1	A	463	GLU	4.1
1	D	464	CYS	4.0
1	D	425	GLY	4.0
1	F	790	THR	4.0
1	D	799	CYS	3.9
1	E	468	SER	3.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	460	VAL	3.9
1	C	296	HIS	3.8
1	A	465	ALA	3.7
1	B	797	TRP	3.7
1	D	774	THR	3.7
1	D	925	ASP	3.6
1	D	469	VAL	3.6
1	B	483	LEU	3.4
1	A	808	ALA	3.4
1	E	417	TYR	3.4
1	D	788	MET	3.3
1	C	799	CYS	3.3
1	A	807	GLN	3.3
1	F	296	HIS	3.3
1	C	803	LYS	3.3
1	E	808	ALA	3.3
1	C	796	PRO	3.3
1	C	774	THR	3.2
1	D	459	LEU	3.2
1	E	858	LEU	3.2
1	B	791	LEU	3.2
1	F	424	ASN	3.2
1	E	809	THR	3.1
1	C	808	ALA	3.1
1	B	803	LYS	3.1
1	B	796	PRO	3.1
1	B	423	ALA	3.1
1	E	416	PRO	3.0
1	E	807	GLN	3.0
1	E	435	ALA	3.0
1	C	800	PRO	2.9
1	B	795	ASP	2.9
1	A	801	ASN	2.8
1	A	466	LYS	2.8
1	D	790	THR	2.8
1	E	441	LEU	2.8
1	D	789	GLU	2.7
1	F	484	LYS	2.7
1	E	455	PHE	2.7
1	B	425	GLY	2.6
1	B	877	MET	2.6
1	E	415	LYS	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	430	VAL	2.5
1	F	483	LEU	2.5
1	E	469	VAL	2.5
1	D	470	THR	2.5
1	D	426	ARG	2.5
1	F	461	CYS	2.5
1	B	422	ASP	2.5
1	A	423	ALA	2.5
1	C	423	ALA	2.5
1	F	416	PRO	2.5
1	F	425	GLY	2.4
1	F	421	LYS	2.4
1	B	800	PRO	2.4
1	F	462	PRO	2.4
1	F	859	SER	2.4
1	F	789	GLU	2.3
1	F	389	TYR	2.3
1	A	424	ASN	2.3
1	B	793	GLU	2.3
1	C	369	HIS	2.3
1	F	297	ILE	2.3
1	E	414	LYS	2.3
1	C	802	CYS	2.3
1	E	442	ARG	2.3
1	A	460	VAL	2.2
1	E	463	GLU	2.2
1	F	438	ASN	2.2
1	F	459	LEU	2.1
1	D	812	PHE	2.1
1	E	459	LEU	2.1
1	D	877	MET	2.1
1	F	926	GLU	2.1
1	A	464	CYS	2.1
1	D	810	LYS	2.1
1	E	465	ALA	2.1
1	B	805	HIS	2.1
1	E	438	ASN	2.0

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

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. LLDF column lists the quality of electron

density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
1	CME	E	311	10/11	0.92	0.18	-	32,37,56,58	0
1	CME	C	311	10/11	0.88	0.18	-	38,41,58,59	0
1	CME	A	311	10/11	0.90	0.19	-	27,32,45,47	0
1	CME	B	311	10/11	0.90	0.21	-	23,28,44,44	0
1	CME	D	311	10/11	0.88	0.19	-	31,37,53,53	0
1	CME	F	311	10/11	0.90	0.17	-	35,43,66,68	0

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
3	SO4	C	1925	5/5	0.78	0.23	6.63	118,121,122,125	0
3	SO4	A	1926	5/5	0.91	0.17	1.04	100,104,105,105	0
2	ZN	D	1000	1/1	0.87	0.24	-1.85	162,162,162,162	0
2	ZN	A	1000	1/1	0.90	0.15	-2.23	159,159,159,159	0
2	ZN	B	1000	1/1	0.99	0.04	-2.47	88,88,88,88	0
2	ZN	C	1000	1/1	0.89	0.04	-2.51	103,103,103,103	0
2	ZN	F	1000	1/1	0.85	0.05	-	99,99,99,99	0
2	ZN	E	1000	1/1	0.92	0.08	-	123,123,123,123	0

6.5 Other polymers (i)

There are no such residues in this entry.