wwPDB EM Validation Summary Report (i)

Jun 25, 2023 – 12:08 AM JST

PDB ID	:	8HPO
EMDB ID	:	EMD-34935
Title	:	Cryo-EM structure of a SIN3/HDAC complex from budding yeast
Authors	:	Guo, Z.; Zhan, X.; Wang, C.
Deposited on	:	2022-12-12
Resolution	:	2.60 Å(reported)

This is a wwPDB EM 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/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

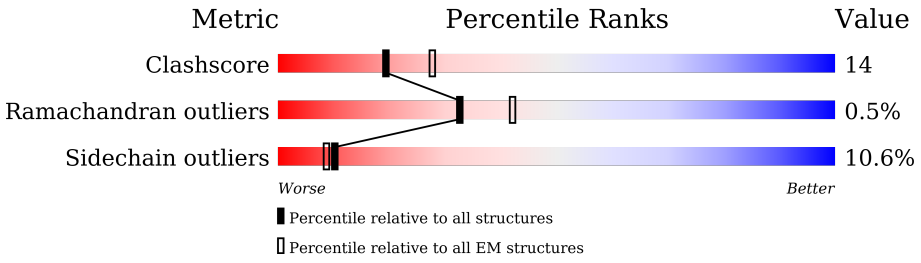
EMDB validation analysis	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 2.60 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.



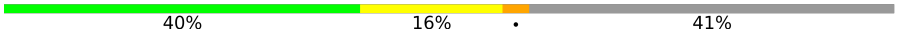
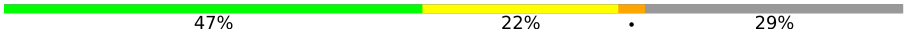
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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%

Mol	Chain	Length	Quality of chain
1	K	460	86% · 10%
2	D	327	28% 16% 6% 50%
3	F	433	66% 21% · · 9%
4	G	433	63% 22% · 11%
5	I	201	38% 23% · · 35%
6	A	1536	26% 14% · 59%
6	B	1536	22% 16% · 60%
7	H	405	36% 18% · 42%
8	C	330	27% 13% · 58%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	E	430	
10	J	294	

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
3	SEP	F	265	-	-	X	-
3	TPO	F	365	-	-	X	-
9	TPO	E	167	-	-	X	-

2 Entry composition (i)

There are 13 unique types of molecules in this entry. The entry contains 27951 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Transcriptional regulatory protein UME1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
1	K	413	2038	1212	413	413	0	0

- Molecule 2 is a protein called Transcriptional regulatory protein SDS3.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	N	O	P	S		
2	D	165	1385	863	248	271	1	2	0	0

- Molecule 3 is a protein called Histone deacetylase RPD3.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	N	O	P	S		
3	F	392	3132	1982	524	596	5	25	0	0

- Molecule 4 is a protein called Histone deacetylase RPD3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	G	385	3057	1948	513	571	25	0	0

- Molecule 5 is a protein called Transcriptional regulatory protein SAP30.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	N	O	P	S		
5	I	130	1104	696	203	201	1	3	0	0

- Molecule 6 is a protein called Transcriptional regulatory protein SIN3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	A	637	5283	3392	887	987	17	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
6	B	613	Total	C	N	O	S	0	0
			5129	3295	862	955	17		

- Molecule 7 is a protein called Transcriptional regulatory protein DEP1.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	H	233	Total	C	N	O	S	0	0
			1936	1218	340	368	10		

- Molecule 8 is a protein called Transcriptional regulatory protein PHO23.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	C	138	Total	C	N	O	S	0	0
			1116	703	195	211	7		

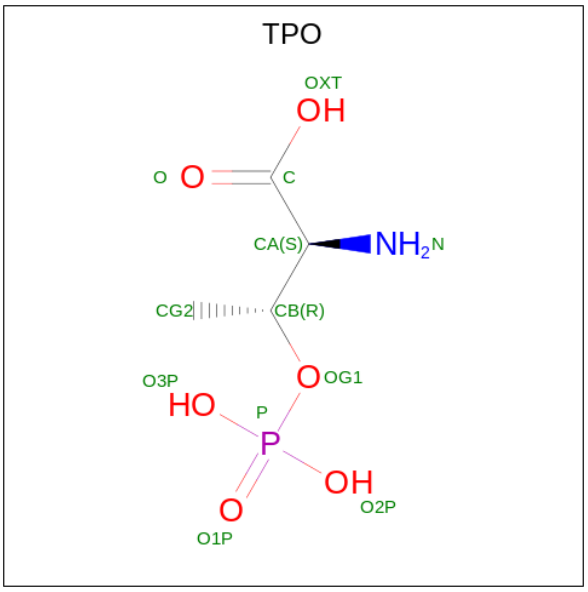
- Molecule 9 is a protein called Transcriptional regulatory protein RXT2.

Mol	Chain	Residues	Atoms						AltConf	Trace
9	E	255	Total	C	N	O	P	S	0	0
			2077	1306	375	391	2	3		

- Molecule 10 is a protein called Transcriptional regulatory protein RXT3.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	209	Total	C	N	O	S	0	0
			1676	1069	284	321	2		

- Molecule 11 is PHOSPHOTHREONINE (three-letter code: TPO) (formula: $C_4H_{10}NO_6P$).



Mol	Chain	Residues	Atoms					AltConf
11	F	1	Total	C	N	O	P	0
			12	4	1	6	1	

- Molecule 12 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
12	F	1	Total	Zn	0
			1	1	
12	G	1	Total	Zn	0
			1	1	

- Molecule 13 is POTASSIUM ION (three-letter code: K) (formula: K).

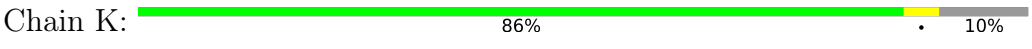
Mol	Chain	Residues	Atoms		AltConf
13	F	2	Total	K	0
			2	2	
13	G	2	Total	K	0
			2	2	

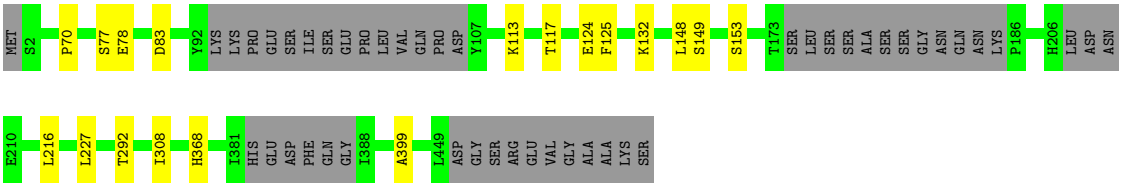
3 Residue-property plots (i)

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. 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. 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: Transcriptional regulatory protein UME1

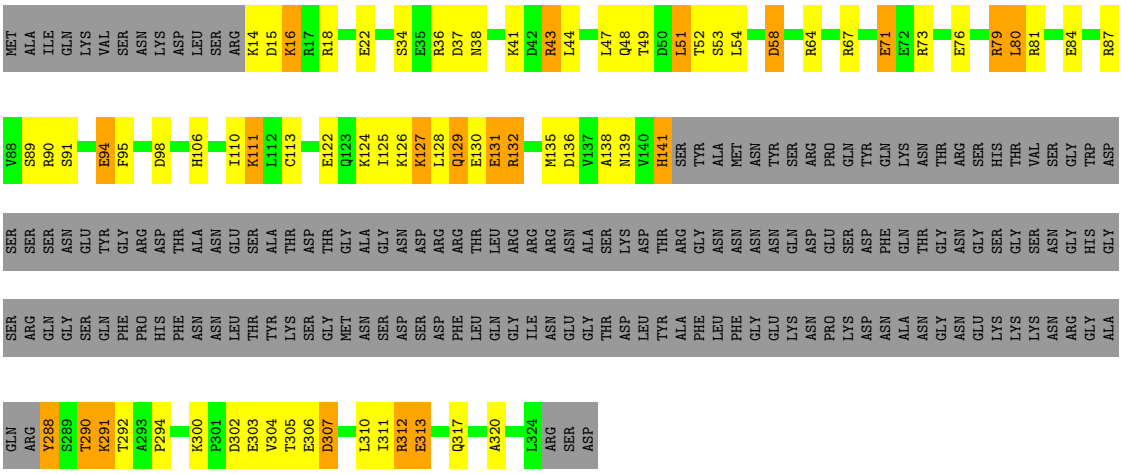
Chain K:





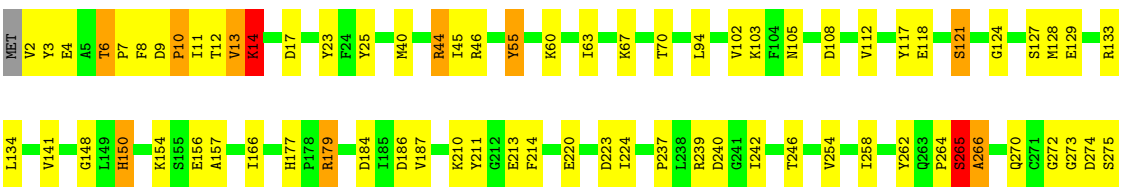
- Molecule 2: Transcriptional regulatory protein SDS3

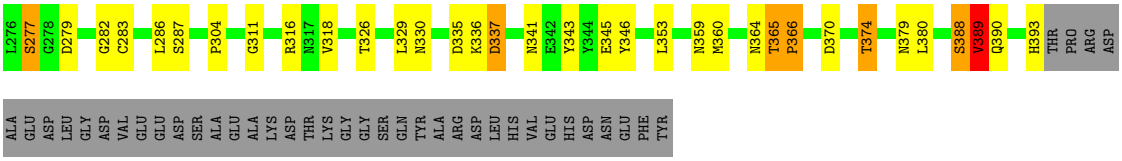
Chain D:



- Molecule 3: Histone deacetylase RPD3

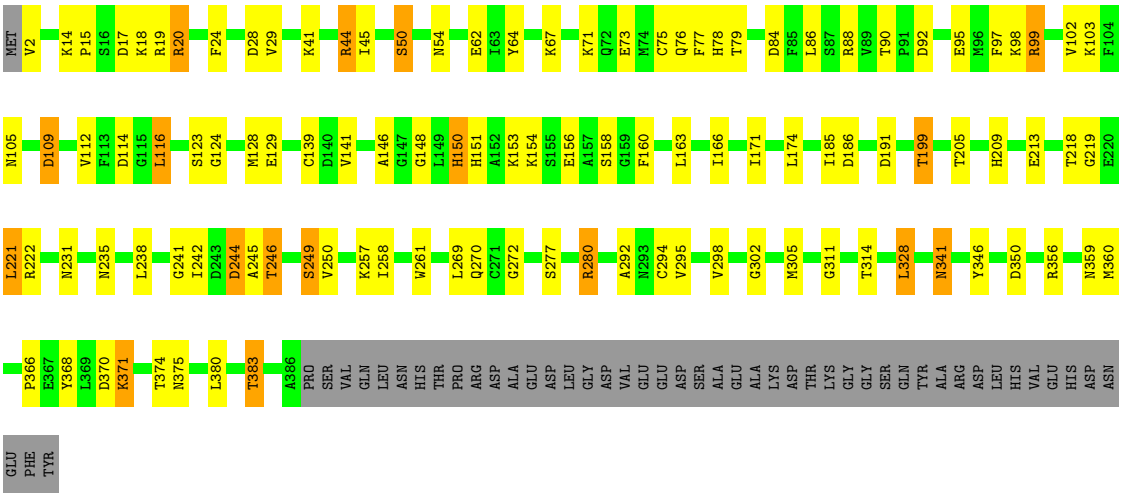
Chain F:





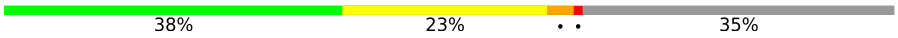
• Molecule 4: Histone deacetylase RPD3

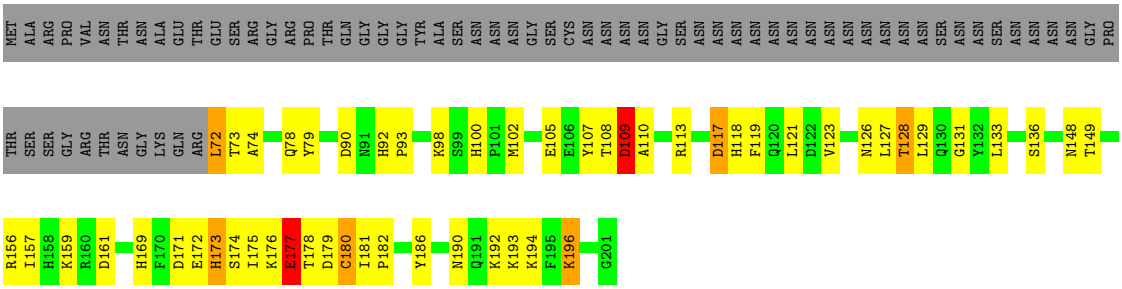
Chain G:



• Molecule 5: Transcriptional regulatory protein SAP30

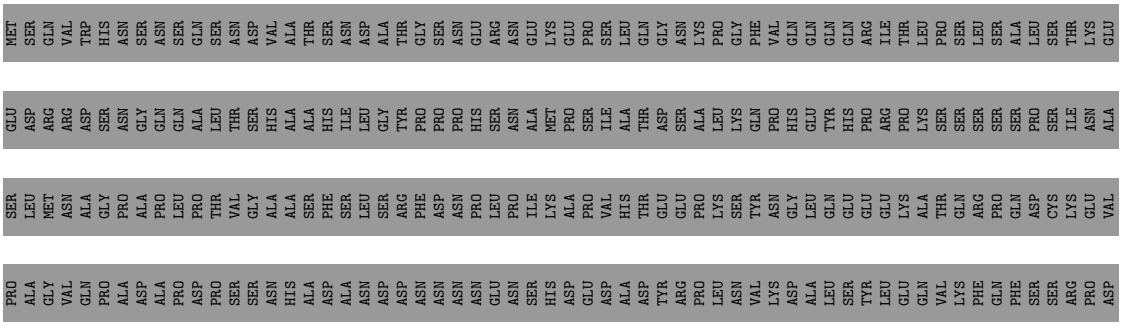
Chain I:

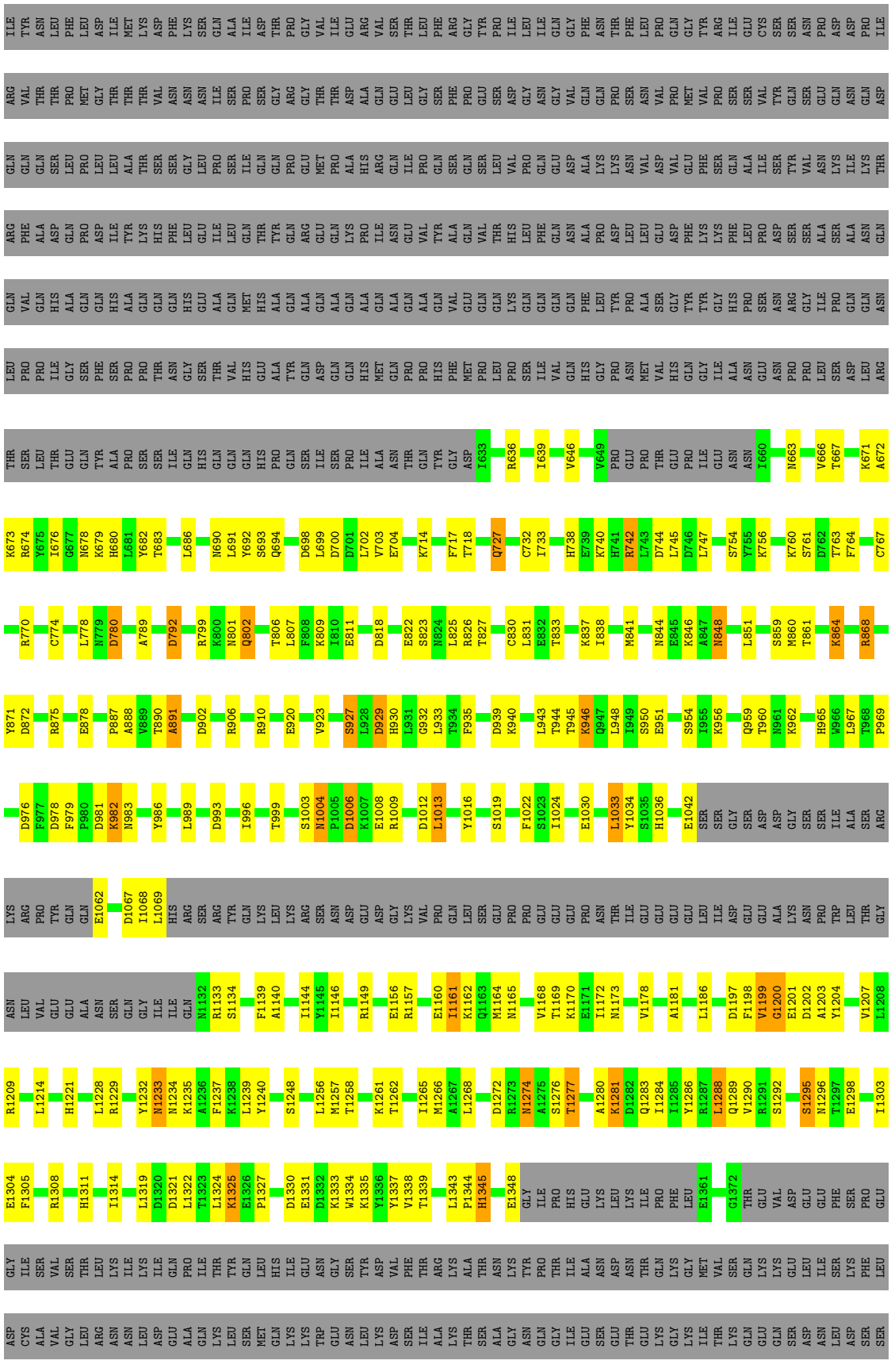




• Molecule 6: Transcriptional regulatory protein SIN3

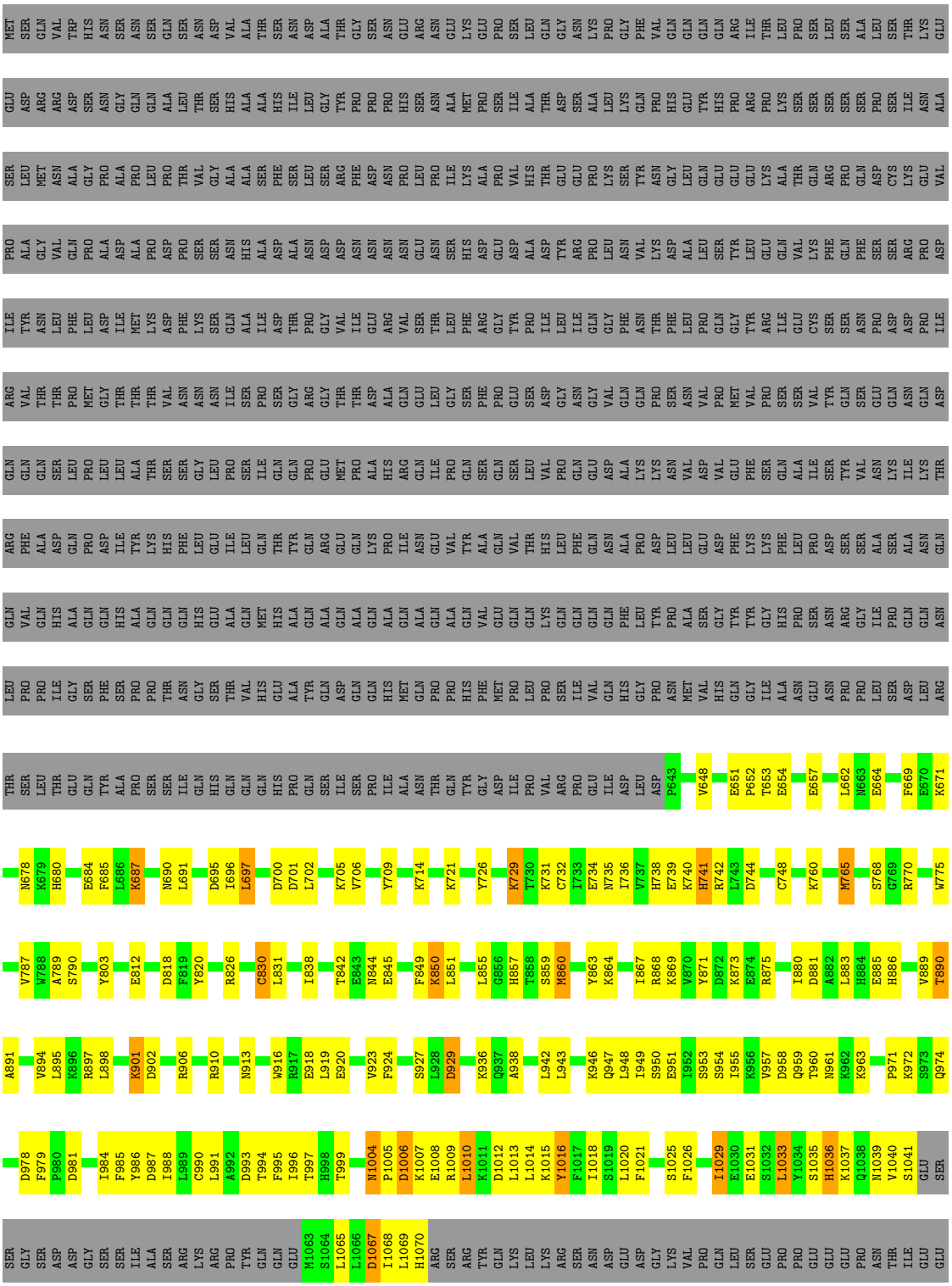
Chain A:

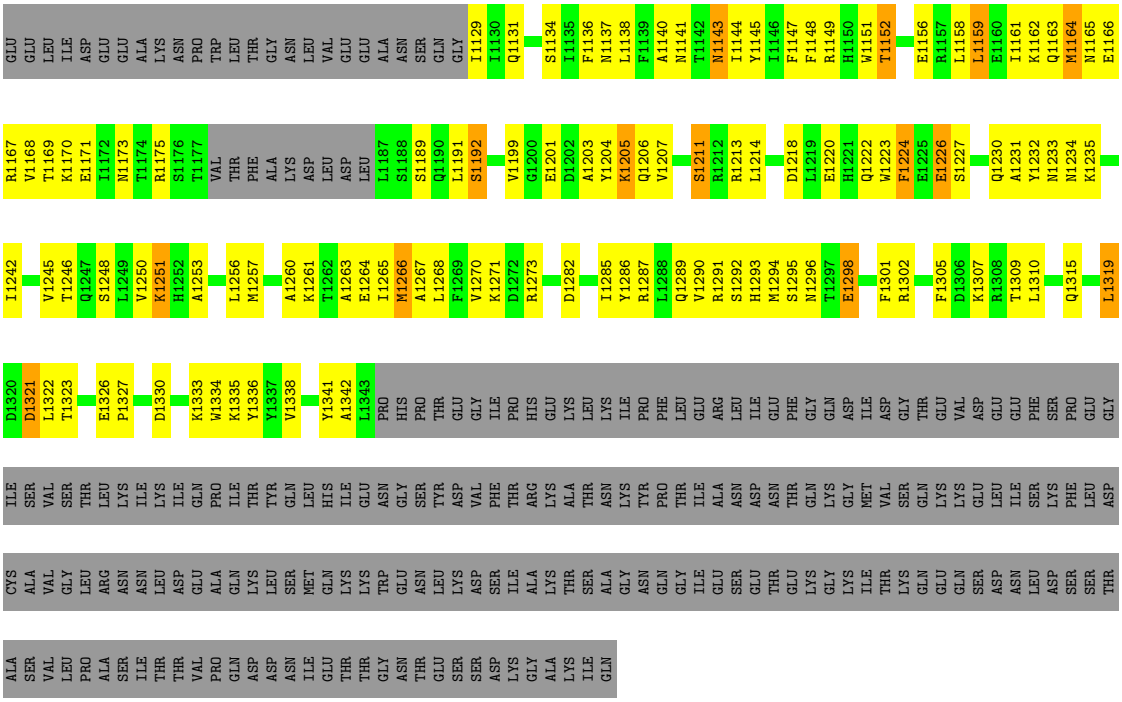




• Molecule 6: Transcriptional regulatory protein SIN3

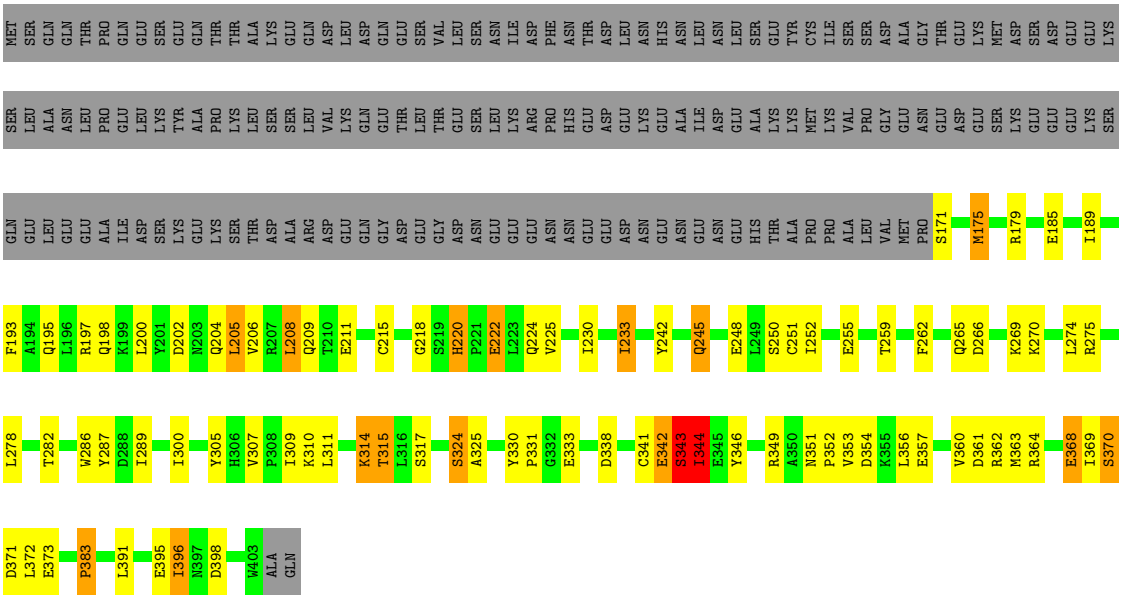
Chain B:





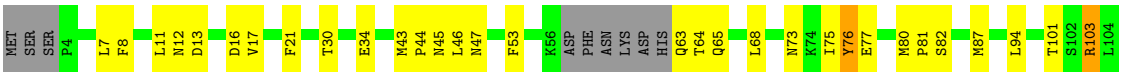
• Molecule 7: Transcriptional regulatory protein DEP1

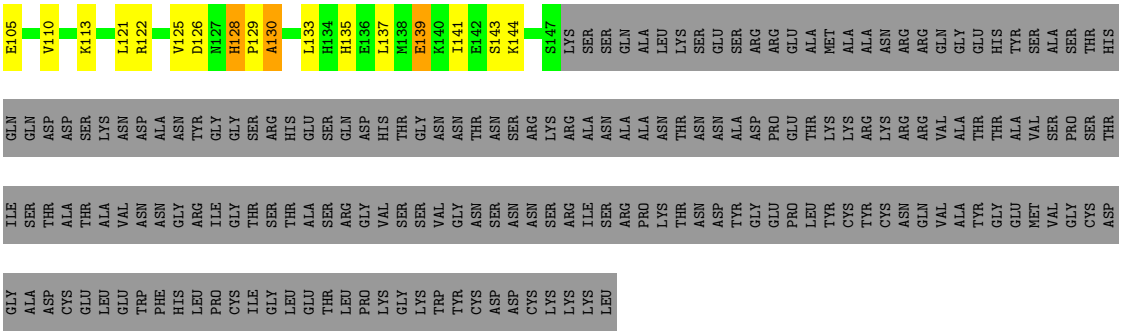
Chain H:



• Molecule 8: Transcriptional regulatory protein PHO23

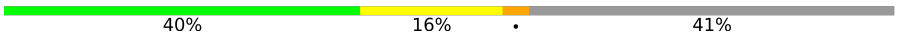
Chain C:

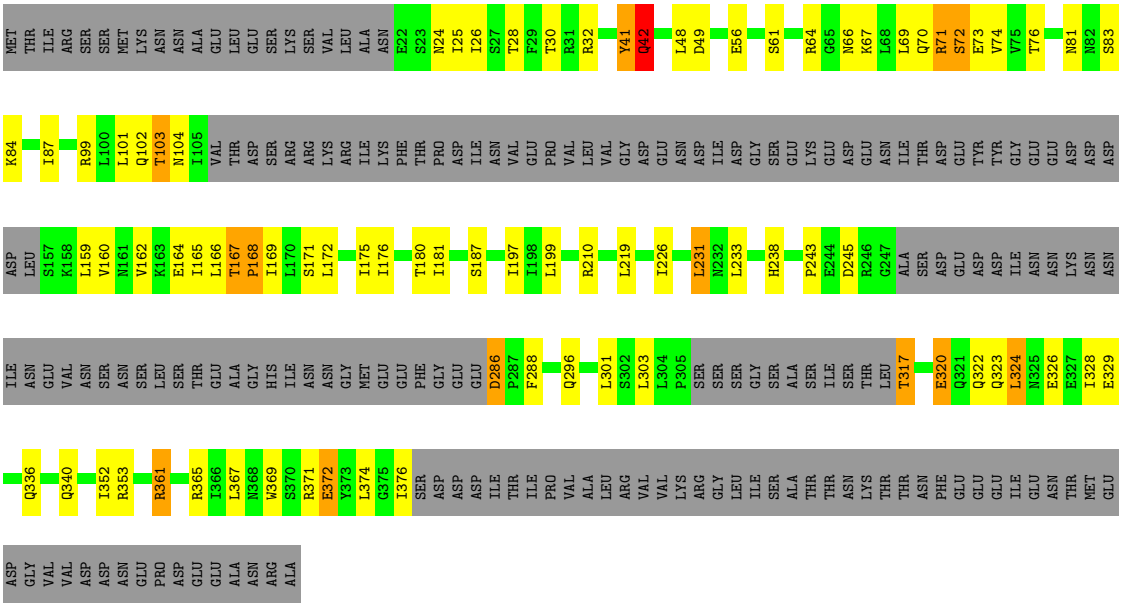




• Molecule 9: Transcriptional regulatory protein RXT2

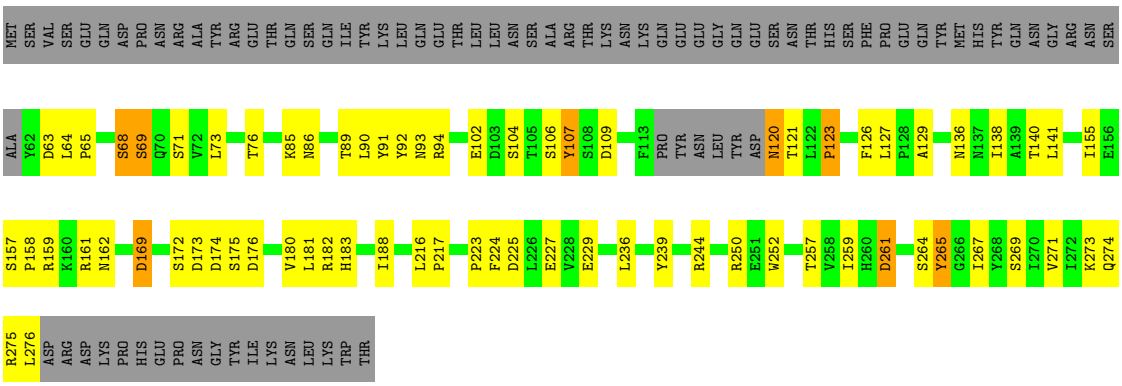
Chain E:





• Molecule 10: Transcriptional regulatory protein RXT3

Chain J:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	665105	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	1800	Depositor
Maximum defocus (nm)	2300	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SEP, K, ZN, TPO

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	K	0.24	0/2033	0.44	0/2822
2	D	0.50	1/1393 (0.1%)	0.69	1/1866 (0.1%)
3	F	0.55	3/3156 (0.1%)	0.63	6/4262 (0.1%)
4	G	0.35	0/3137	0.48	0/4246
5	I	0.38	0/1119	0.61	1/1500 (0.1%)
6	A	0.32	0/5400	0.48	0/7289
6	B	0.30	0/5245	0.51	1/7077 (0.0%)
7	H	0.34	0/1976	0.50	0/2674
8	C	0.30	0/1134	0.50	0/1529
9	E	0.42	0/2082	0.58	1/2799 (0.0%)
10	J	0.38	1/1718 (0.1%)	0.54	0/2335
All	All	0.37	5/28393 (0.0%)	0.53	10/38399 (0.0%)

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
3	F	0	1
5	I	0	1
6	A	0	1
6	B	0	2
7	H	0	4
9	E	0	2
All	All	0	11

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	10	PRO	N-CA	13.06	1.69	1.47

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	294	PRO	C-N	8.63	1.50	1.34
3	F	14	LYS	C-N	8.37	1.50	1.34
10	J	123	PRO	C-N	7.87	1.49	1.34
3	F	9	ASP	C-N	6.27	1.46	1.34

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	307	ASP	CB-CA-C	-11.25	87.91	110.40
3	F	9	ASP	CB-CA-C	-7.49	95.42	110.40
3	F	266	ALA	N-CA-CB	7.36	120.40	110.10
3	F	10	PRO	CA-N-CD	-6.63	102.21	111.50
6	B	1033	LEU	CA-CB-CG	6.06	129.25	115.30

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
6	A	891	ALA	Peptide
3	F	265	SEP	Mainchain
7	H	314	LYS	Peptide
7	H	342	GLU	Peptide
5	I	174	SEP	Mainchain

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	K	2038	0	890	9	0
2	D	1385	0	1388	84	0
3	F	3132	0	2978	85	0
4	G	3057	0	2932	69	0
5	I	1104	0	1087	46	0
6	A	5283	0	5216	166	0
6	B	5129	0	5096	193	0
7	H	1936	0	1896	65	0
8	C	1116	0	1133	38	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	E	2077	0	2136	75	0
10	J	1676	0	1637	44	0
11	F	12	0	5	3	0
12	F	1	0	0	0	0
12	G	1	0	0	0	0
13	F	2	0	0	0	0
13	G	2	0	0	0	0
All	All	27951	0	26394	767	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 14.

The worst 5 of 767 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)
3:F:10:PRO:N	3:F:10:PRO:CA	1.69	1.36
3:F:6:TPO:CG2	3:F:7:PRO:HD2	1.68	1.23
9:E:176:ILE:CD1	9:E:374:LEU:HD12	1.77	1.15
9:E:167:TPO:CG2	9:E:168:PRO:HD2	1.81	1.10
5:I:177:GLU:O	5:I:179:ASP:N	1.82	1.09

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 PDB entries followed by that with respect to all EM entries.

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	K	403/460 (88%)	396 (98%)	7 (2%)	0	100 100
2	D	160/327 (49%)	157 (98%)	3 (2%)	0	100 100
3	F	385/433 (89%)	356 (92%)	26 (7%)	3 (1%)	19 39
4	G	383/433 (88%)	360 (94%)	23 (6%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
5	I	127/201 (63%)	110 (87%)	14 (11%)	3 (2%)	6 10
6	A	627/1536 (41%)	593 (95%)	32 (5%)	2 (0%)	41 64
6	B	605/1536 (39%)	551 (91%)	54 (9%)	0	100 100
7	H	231/405 (57%)	208 (90%)	20 (9%)	3 (1%)	12 24
8	C	134/330 (41%)	128 (96%)	3 (2%)	3 (2%)	6 12
9	E	245/430 (57%)	217 (89%)	24 (10%)	4 (2%)	9 19
10	J	205/294 (70%)	188 (92%)	17 (8%)	0	100 100
All	All	3505/6385 (55%)	3264 (93%)	223 (6%)	18 (0%)	32 52

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	I	177	GLU
5	I	178	THR
7	H	344	ILE
9	E	42	GLN
9	E	71	ARG

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 PDB entries followed by that with respect to all EM entries.

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
2	D	153/289 (53%)	124 (81%)	29 (19%)	1 2
3	F	328/362 (91%)	301 (92%)	27 (8%)	11 22
4	G	326/367 (89%)	292 (90%)	34 (10%)	7 13
5	I	119/178 (67%)	105 (88%)	14 (12%)	5 9
6	A	585/1391 (42%)	524 (90%)	61 (10%)	7 13
6	B	574/1391 (41%)	513 (89%)	61 (11%)	6 12
7	H	213/371 (57%)	190 (89%)	23 (11%)	6 12
8	C	129/290 (44%)	123 (95%)	6 (5%)	26 50
9	E	234/391 (60%)	210 (90%)	24 (10%)	7 13

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
10	J	190/269 (71%)	167 (88%)	23 (12%)	5 9
All	All	2851/5299 (54%)	2549 (89%)	302 (11%)	10 12

5 of 302 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	B	765	MET
10	J	109	ASP
6	B	881	ASP
6	B	1152	THR
10	J	273	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 64 such sidechains are listed below:

Mol	Chain	Res	Type
6	B	1233	ASN
6	B	1296	ASN
6	A	848	ASN
6	A	836	ASN
10	J	120	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)

9 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	$\#\|Z\| > 2$
9	SEP	E	171	9	8,9,10	0.69	0	8,12,14	1.52	1 (12%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	SEP	I	174	5	8,9,10	0.96	0	8,12,14	0.82	0
3	SEP	F	388	3	8,9,10	0.73	0	8,12,14	1.08	1 (12%)
3	TPO	F	365	3	8,10,11	0.86	0	10,14,16	1.19	2 (20%)
2	SEP	D	309	2	8,9,10	0.94	0	8,12,14	0.94	0
9	TPO	E	167	9	8,10,11	2.98	1 (12%)	10,14,16	1.56	2 (20%)
3	SEP	F	265	3	8,9,10	1.97	1 (12%)	8,12,14	2.22	3 (37%)
3	TPO	F	6	3	8,10,11	0.99	0	10,14,16	1.33	1 (10%)
3	TPO	F	12	3	8,10,11	0.81	0	10,14,16	1.12	1 (10%)

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
9	SEP	E	171	9	-	2/5/8/10	-
5	SEP	I	174	5	-	1/5/8/10	-
3	SEP	F	388	3	-	3/5/8/10	-
3	TPO	F	365	3	-	7/9/11/13	-
2	SEP	D	309	2	-	2/5/8/10	-
9	TPO	E	167	9	-	6/9/11/13	-
3	SEP	F	265	3	-	4/5/8/10	-
3	TPO	F	6	3	-	0/9/11/13	-
3	TPO	F	12	3	-	1/9/11/13	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	E	167	TPO	P-OG1	-7.95	1.44	1.59
3	F	265	SEP	O-C	-4.57	1.01	1.19

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	265	SEP	O2P-P-OG	-4.51	94.74	106.73
9	E	167	TPO	P-OG1-CB	-3.55	112.48	123.21
9	E	171	SEP	O3P-P-OG	-2.75	99.41	106.73
3	F	265	SEP	OG-P-O1P	2.72	114.11	106.47
3	F	388	SEP	OG-CB-CA	2.70	110.78	108.14

There are no chirality outliers.

5 of 26 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	F	12	TPO	O-C-CA-CB
3	F	265	SEP	N-CA-CB-OG
3	F	265	SEP	CB-OG-P-O2P
3	F	265	SEP	CB-OG-P-O3P
3	F	365	TPO	N-CA-CB-CG2

There are no ring outliers.

5 monomers are involved in 28 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	F	388	SEP	1	0
3	F	365	TPO	8	0
9	E	167	TPO	8	0
3	F	265	SEP	6	0
3	F	6	TPO	5	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 6 are monoatomic - leaving 1 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	$\#\|Z\| > 2$
11	TPO	F	501	3	9,11,11	0.83	0	13,16,16	1.79	1 (7%)

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
11	TPO	F	501	3	-	5/13/13/13	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
11	F	501	TPO	CB-CA-C	5.25	122.46	110.32

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
11	F	501	TPO	O-C-CA-N
11	F	501	TPO	OXT-C-CA-N
11	F	501	TPO	O-C-CA-CB
11	F	501	TPO	OXT-C-CA-CB
11	F	501	TPO	CG2-CB-OG1-P

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	F	501	TPO	3	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.