wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 02:01 PM BST

PDB ID	:	4UX8
EMDB ID:	:	EMD-2712
Title	:	RET recognition of GDNF-GFRalpha1 ligand by a composite binding site promotes membrane-proximal self-association
Authors	:	Goodman, K.; Kjaer, S.; Beuron, F.; Knowles, P.; Nawrotek, A.; Burns, E.; Purkiss, A.; George, R.; Santoro, M.; Morris, E.P.; McDonald, N.Q.
Deposited on	:	2014-08-19
Resolution	:	24.00 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

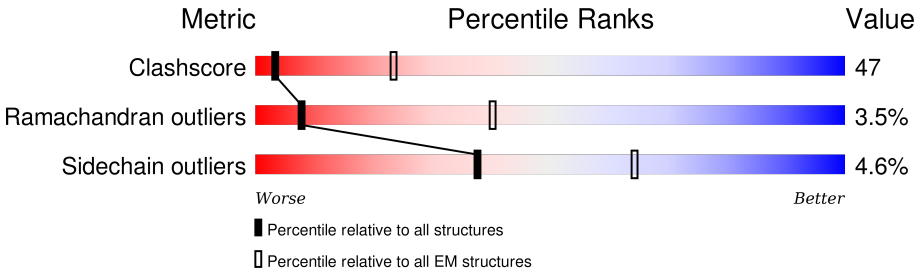
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 24.00 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	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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%

Mol	Chain	Length	Quality of chain
1	A	607	35% 27% 10% • 26%
1	B	607	36% 27% 10% • 26%
2	C	463	42% 18% • 39%
2	E	463	42% 18% • 39%
3	D	134	57% 8% • 31%
3	F	134	57% 8% • 31%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13046 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 PROTO-ONCOGENE TYROSINE-PROTEIN KINASE RECEPTOR RET.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	448	Total	C	N	O	S	0	0
			3569	2264	632	660	13		
1	B	448	Total	C	N	O	S	0	0
			3569	2264	632	660	13		

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	87	ARG	CYS	ENGINEERED MUTATION	UNP P07949
A	98	GLN	ASN	CONFLICT	UNP P07949
A	199	GLN	ASN	CONFLICT	UNP P07949
A	216	SER	CYS	ENGINEERED MUTATION	UNP P07949
B	87	ARG	CYS	ENGINEERED MUTATION	UNP P07949
B	98	GLN	ASN	CONFLICT	UNP P07949
B	199	GLN	ASN	CONFLICT	UNP P07949
B	216	SER	CYS	ENGINEERED MUTATION	UNP P07949

- Molecule 2 is a protein called GDNF FAMILY RECEPTOR ALPHA-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	284	Total	C	N	O	S	0	0
			2203	1349	398	422	34		
2	E	284	Total	C	N	O	S	0	0
			2203	1349	398	422	34		

- Molecule 3 is a protein called GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	93	Total	C	N	O	S	0	0
			720	450	126	138	6		
3	F	93	Total	C	N	O	S	0	0
			720	450	126	138	6		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	AltConf
4	B	3	Total Ca 3 3	0
4	A	3	Total Ca 3 3	0

- Molecule 5 is a polymer of unknown type called SUGAR (2-MER).

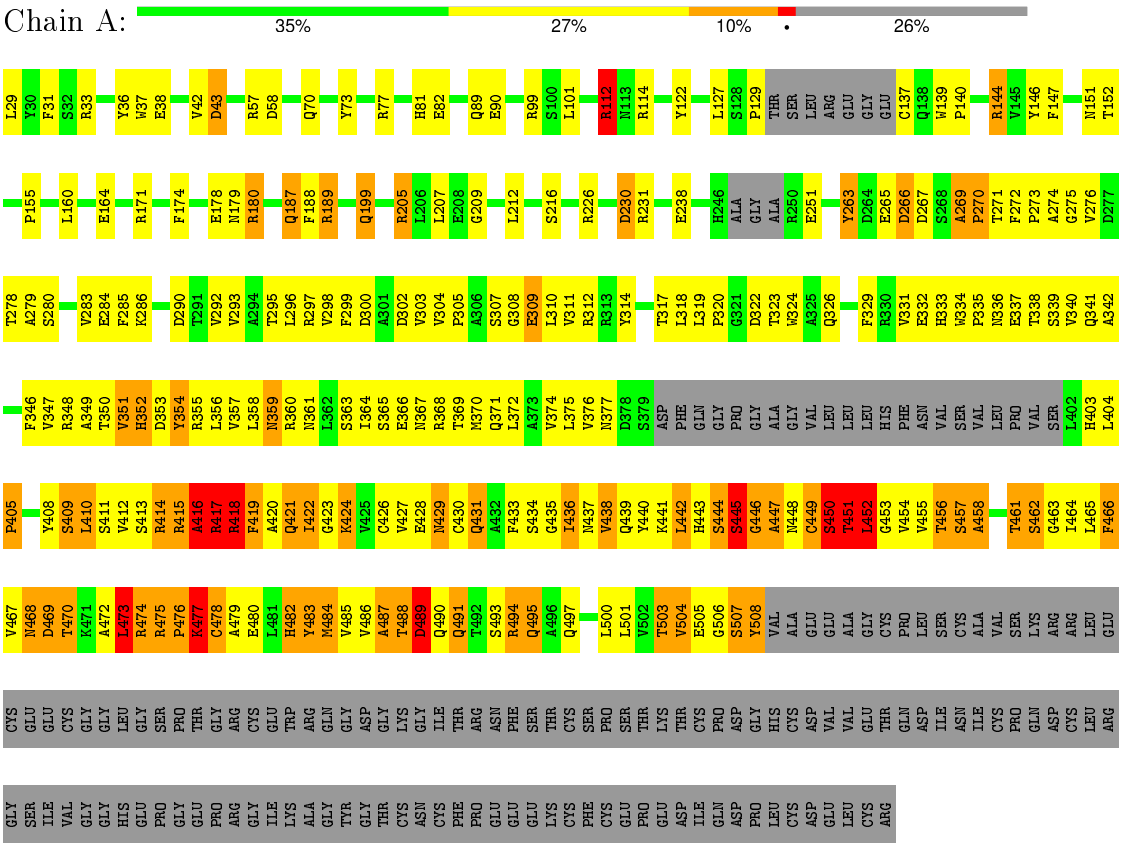
Mol	Chain	Residues	Atoms	AltConf
5	D	2	Total C N O 28 16 2 10	0
5	F	2	Total C N O 28 16 2 10	0

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. 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.

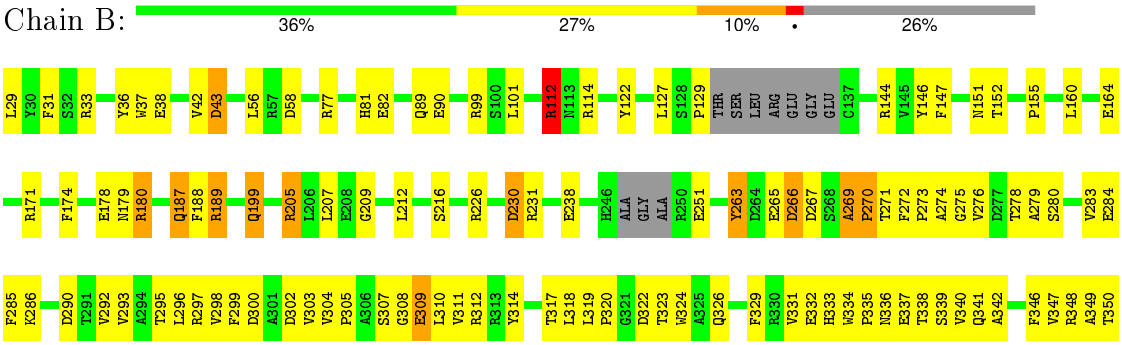
- Molecule 1: PROTO-ONCOGENE TYROSINE-PROTEIN KINASE RECEPTOR RET

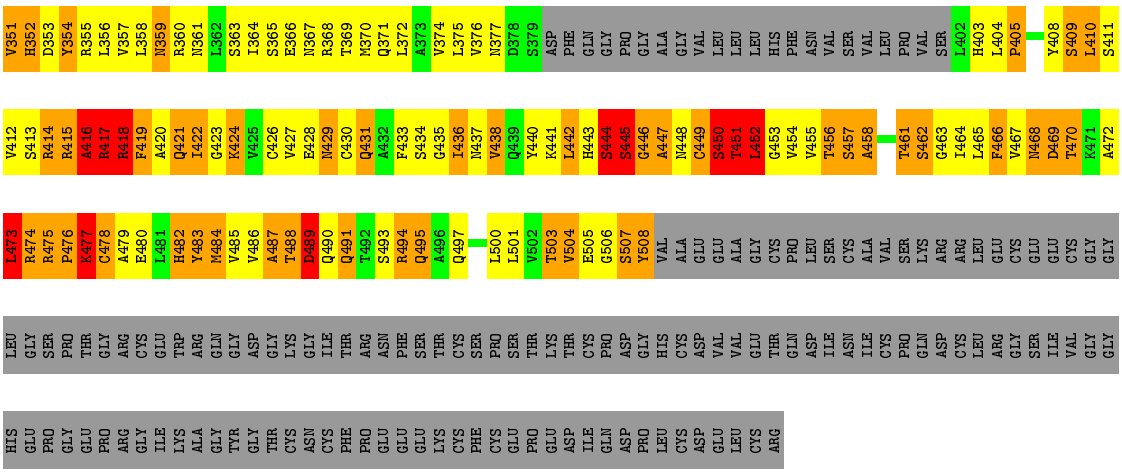
Chain A:



- Molecule 1: PROTO-ONCOGENE TYROSINE-PROTEIN KINASE RECEPTOR RET

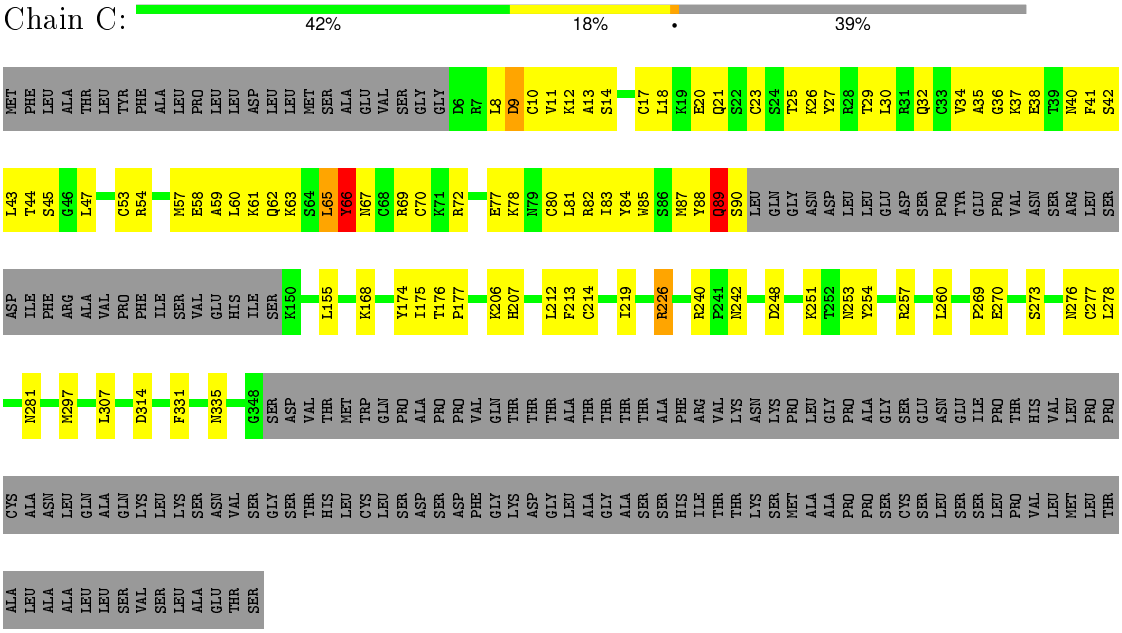
Chain B:





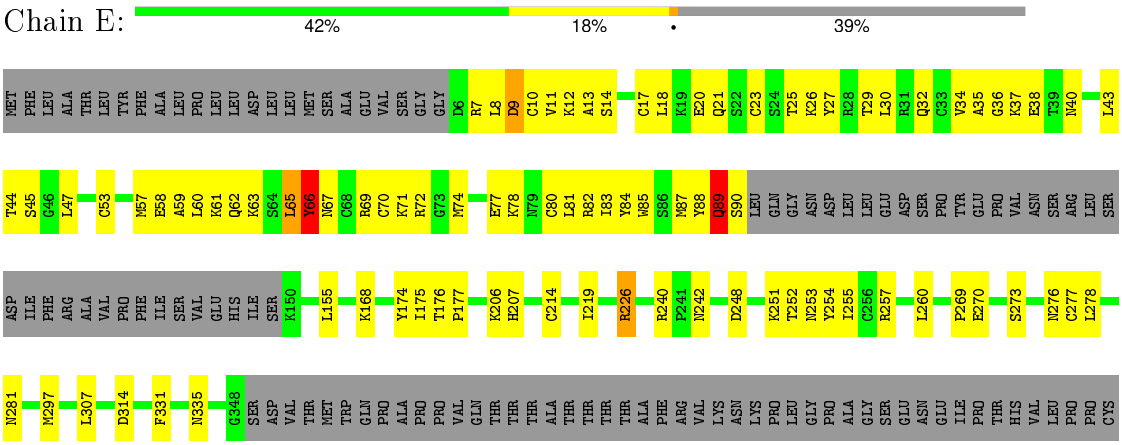
• Molecule 2: GDNF FAMILY RECEPTOR ALPHA-1

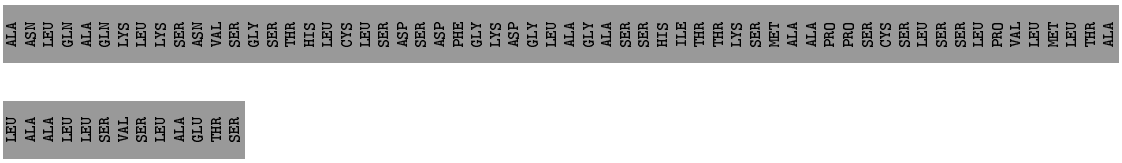
Chain C:



• Molecule 2: GDNF FAMILY RECEPTOR ALPHA-1

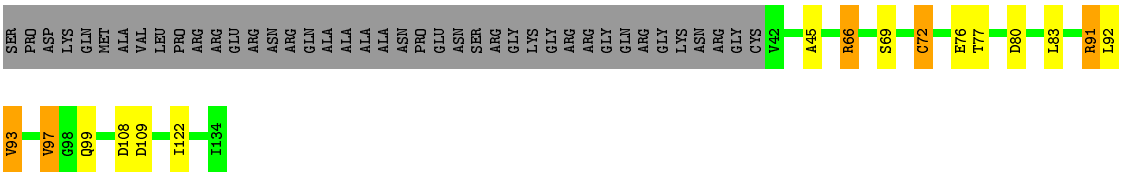
Chain E:





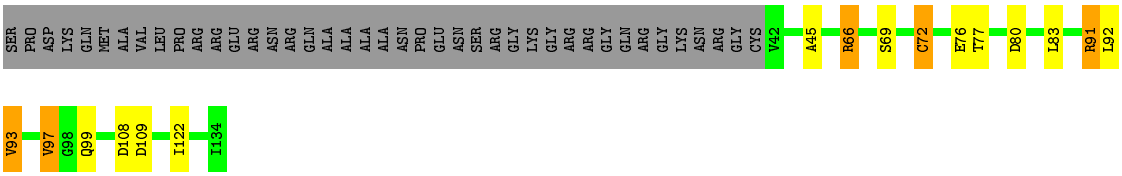
• Molecule 3: GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR

Chain D:



• Molecule 3: GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR

Chain F:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	100	Depositor
Minimum defocus (nm)	900	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	80000	Depositor
Image detector	TVIPS TEMCAM-F415 (4K 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: CA, NAG

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	#$\|Z\|$ >2	RMSZ	#$\|Z\|$ >2
1	A	1.59	43/3655 (1.2%)	1.94	130/4977 (2.6%)
1	B	1.59	45/3655 (1.2%)	1.94	128/4977 (2.6%)
2	C	0.62	0/2236	0.80	3/3005 (0.1%)
2	E	0.62	0/2236	0.80	3/3005 (0.1%)
3	D	0.37	0/730	0.56	0/985
3	F	0.36	0/730	0.56	0/985
All	All	1.24	88/13242 (0.7%)	1.53	264/17934 (1.5%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	4	2
1	B	4	2
All	All	8	4

The worst 5 of 88 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	269	ALA	C-N	16.06	1.64	1.34
1	B	269	ALA	C-N	16.06	1.64	1.34
1	B	405	PRO	N-CD	15.82	1.70	1.47
1	A	405	PRO	N-CD	15.81	1.70	1.47
1	A	410	LEU	N-CA	15.03	1.76	1.46

The worst 5 of 264 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	263	TYR	O-C-N	-27.83	78.17	122.70
1	B	263	TYR	O-C-N	-27.79	78.23	122.70
1	A	112	ARG	NE-CZ-NH1	-27.21	106.70	120.30
1	B	112	ARG	NE-CZ-NH1	-27.05	106.78	120.30
1	A	450	SER	CB-CA-C	-22.51	67.33	110.10

5 of 8 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	445	SER	CA
1	A	452	LEU	CA
1	A	491	GLN	CA
1	A	507	SER	CA
1	B	445	SER	CA

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	112	ARG	Sidechain
1	A	263	TYR	Peptide
1	B	112	ARG	Sidechain
1	B	263	TYR	Peptide

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	3569	0	3460	442	0
1	B	3569	0	3462	437	0
2	C	2203	0	2105	143	0
2	E	2203	0	2105	148	0
3	D	720	0	697	19	0
3	F	720	0	697	17	0
4	A	3	0	0	0	0
4	B	3	0	0	0	0
5	D	28	0	25	0	0
5	F	28	0	25	0	0
All	All	13046	0	12576	1195	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 47.

The worst 5 of 1195 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:A:486:VAL:CA	1:A:497:GLN:HE22	1.06	1.65
1:B:486:VAL:CA	1:B:497:GLN:HE22	1.06	1.62
1:A:482:HIS:CD2	1:A:501:LEU:HD21	1.37	1.60
2:C:78:LYS:CE	2:C:254:TYR:CA	1.75	1.59
1:B:482:HIS:CD2	1:B:501:LEU:HD21	1.37	1.59

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	A	440/607 (72%)	385 (88%)	33 (8%)	22 (5%)	3 31
1	B	440/607 (72%)	385 (88%)	33 (8%)	22 (5%)	3 31
2	C	280/463 (60%)	260 (93%)	15 (5%)	5 (2%)	11 53
2	E	280/463 (60%)	260 (93%)	15 (5%)	5 (2%)	11 53
3	D	91/134 (68%)	84 (92%)	6 (7%)	1 (1%)	17 63
3	F	91/134 (68%)	84 (92%)	6 (7%)	1 (1%)	17 63
All	All	1622/2408 (67%)	1458 (90%)	108 (7%)	56 (4%)	8 39

5 of 56 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	349	ALA
1	A	351	VAL
1	A	414	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	416	ALA
1	A	417	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	
1	A	390/528 (74%)	372 (95%)	18 (5%)	33	68
1	B	390/528 (74%)	372 (95%)	18 (5%)	33	68
2	C	247/408 (60%)	236 (96%)	11 (4%)	34	69
2	E	247/408 (60%)	236 (96%)	11 (4%)	34	69
3	D	78/113 (69%)	74 (95%)	4 (5%)	29	66
3	F	78/113 (69%)	74 (95%)	4 (5%)	29	66
All	All	1430/2098 (68%)	1364 (95%)	66 (5%)	38	68

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	238	GLU
2	C	66	TYR
2	E	277	CYS
1	B	267	ASP
1	B	473	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	187	GLN
1	B	359	ASN
2	E	89	GLN
1	B	199	GLN
1	B	361	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

4 carbohydrates 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	$\#\|Z\| > 2$
5	NAG	D	1135	3,5	14,14,15	0.51	0	15,19,21	1.06	1 (6%)
5	NAG	D	1136	5	14,14,15	0.74	1 (7%)	15,19,21	1.29	1 (6%)
5	NAG	F	1135	3,5	14,14,15	0.53	0	15,19,21	1.06	1 (6%)
5	NAG	F	1136	5	14,14,15	0.73	0	15,19,21	1.30	1 (6%)

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
5	NAG	D	1135	3,5	-	0/6/23/26	0/1/1/1
5	NAG	D	1136	5	-	0/6/23/26	0/1/1/1
5	NAG	F	1135	3,5	-	0/6/23/26	0/1/1/1
5	NAG	F	1136	5	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	1136	NAG	C1-C2	2.01	1.55	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	D	1136	NAG	C4-C3-C2	3.01	116.00	111.34
5	F	1136	NAG	C4-C3-C2	3.02	116.03	111.34
5	D	1135	NAG	C1-O5-C5	3.11	116.71	112.14
5	F	1135	NAG	C1-O5-C5	3.17	116.81	112.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

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 (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.