wwPDB EM Validation Summary Report

Jun 9, 2024 – 12:06 AM JST

PDB ID	:	8WYI
EMDB ID	:	EMD-37929
Title	:	T cell receptor delta 2 gamma 9 with TCRD TM domain chimera of TRAC
Authors	:	Xin, W.; Huang, B.; Chi, X.; Liu, Y.; Xu, M.; Zhang, Y.; Li, X.; Su, Q.; Zhou, Q.
Deposited on	:	2023-10-31
Resolution	:	3.90 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 ⓘ 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 ⓘ) were used in the production of this report:

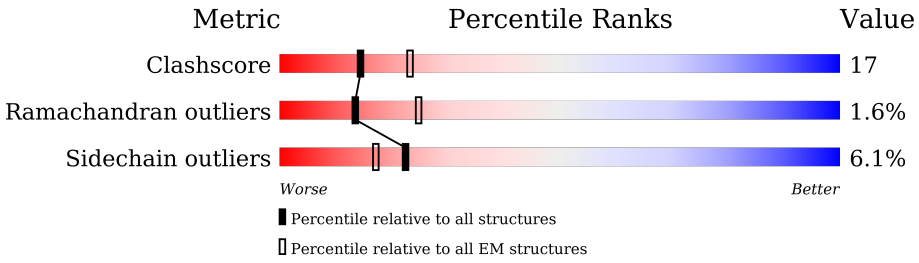
EMDB validation analysis	:	FAILED
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.90 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	a	195	15% · 85%
1	b	195	13% · 86%
2	d	171	59% •• 38%
3	e	207	57% · 42%
3	f	207	52% 5% 43%
4	g	182	55% 8% 37%
5	m	310	10% · 88%
6	n	332	10% · 89%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 4659 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 T-cell surface glycoprotein CD3 zeta chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	a	30	Total	C	N	O	S	0	0
			245	172	35	37	1		
1	b	28	Total	C	N	O	S	0	0
			226	157	33	35	1		

There are 62 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
a	165	ALA	-	expression tag	UNP P20963
a	166	ALA	-	expression tag	UNP P20963
a	167	ALA	-	expression tag	UNP P20963
a	168	TRP	-	expression tag	UNP P20963
a	169	SER	-	expression tag	UNP P20963
a	170	HIS	-	expression tag	UNP P20963
a	171	PRO	-	expression tag	UNP P20963
a	172	GLN	-	expression tag	UNP P20963
a	173	PHE	-	expression tag	UNP P20963
a	174	GLU	-	expression tag	UNP P20963
a	175	LYS	-	expression tag	UNP P20963
a	176	GLY	-	expression tag	UNP P20963
a	177	GLY	-	expression tag	UNP P20963
a	178	GLY	-	expression tag	UNP P20963
a	179	SER	-	expression tag	UNP P20963
a	180	GLY	-	expression tag	UNP P20963
a	181	GLY	-	expression tag	UNP P20963
a	182	GLY	-	expression tag	UNP P20963
a	183	SER	-	expression tag	UNP P20963
a	184	GLY	-	expression tag	UNP P20963
a	185	GLY	-	expression tag	UNP P20963
a	186	SER	-	expression tag	UNP P20963
a	187	ALA	-	expression tag	UNP P20963
a	188	TRP	-	expression tag	UNP P20963
a	189	SER	-	expression tag	UNP P20963
a	190	HIS	-	expression tag	UNP P20963

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
a	191	PRO	-	expression tag	UNP P20963
a	192	GLN	-	expression tag	UNP P20963
a	193	PHE	-	expression tag	UNP P20963
a	194	GLU	-	expression tag	UNP P20963
a	195	LYS	-	expression tag	UNP P20963
b	165	ALA	-	expression tag	UNP P20963
b	166	ALA	-	expression tag	UNP P20963
b	167	ALA	-	expression tag	UNP P20963
b	168	TRP	-	expression tag	UNP P20963
b	169	SER	-	expression tag	UNP P20963
b	170	HIS	-	expression tag	UNP P20963
b	171	PRO	-	expression tag	UNP P20963
b	172	GLN	-	expression tag	UNP P20963
b	173	PHE	-	expression tag	UNP P20963
b	174	GLU	-	expression tag	UNP P20963
b	175	LYS	-	expression tag	UNP P20963
b	176	GLY	-	expression tag	UNP P20963
b	177	GLY	-	expression tag	UNP P20963
b	178	GLY	-	expression tag	UNP P20963
b	179	SER	-	expression tag	UNP P20963
b	180	GLY	-	expression tag	UNP P20963
b	181	GLY	-	expression tag	UNP P20963
b	182	GLY	-	expression tag	UNP P20963
b	183	SER	-	expression tag	UNP P20963
b	184	GLY	-	expression tag	UNP P20963
b	185	GLY	-	expression tag	UNP P20963
b	186	SER	-	expression tag	UNP P20963
b	187	ALA	-	expression tag	UNP P20963
b	188	TRP	-	expression tag	UNP P20963
b	189	SER	-	expression tag	UNP P20963
b	190	HIS	-	expression tag	UNP P20963
b	191	PRO	-	expression tag	UNP P20963
b	192	GLN	-	expression tag	UNP P20963
b	193	PHE	-	expression tag	UNP P20963
b	194	GLU	-	expression tag	UNP P20963
b	195	LYS	-	expression tag	UNP P20963

- Molecule 2 is a protein called T-cell surface glycoprotein CD3 delta chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	d	106	Total	C	N	O	S	0	0
			821	523	136	156	6		

- Molecule 3 is a protein called T-cell surface glycoprotein CD3 epsilon chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	e	121	Total	C	N	O	S	0	0
			959	609	151	191	8		
3	f	118	Total	C	N	O	S	0	0
			933	594	148	183	8		

- Molecule 4 is a protein called T-cell surface glycoprotein CD3 gamma chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	g	114	Total	C	N	O	S	0	0
			899	578	149	165	7		

- Molecule 5 is a protein called Signal peptide,flag tag,T cell receptor delta variable 2,T cell receptor delta constant,T cell receptor alpha chain constant,T cell receptor delta variable 2,T cell receptor delta constant,T cell receptor alpha chain constant.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	m	37	Total	C	N	O	S	0	0
			292	192	50	47	3		

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
m	116	LEU	-	linker	UNP A0JD36
m	117	GLY	-	linker	UNP A0JD36
m	118	MET	-	linker	UNP A0JD36
m	119	GLY	-	linker	UNP A0JD36
m	120	GLY	-	linker	UNP A0JD36
m	121	GLU	-	linker	UNP A0JD36
m	122	TYR	-	linker	UNP A0JD36
m	123	THR	-	linker	UNP A0JD36
m	124	ASP	-	linker	UNP A0JD36
m	125	LYS	-	linker	UNP A0JD36
m	126	LEU	-	linker	UNP A0JD36
m	127	ILE	-	linker	UNP A0JD36
m	128	PHE	-	linker	UNP A0JD36
m	129	GLY	-	linker	UNP A0JD36
m	130	LYS	-	linker	UNP A0JD36
m	131	GLY	-	linker	UNP A0JD36
m	132	THR	-	linker	UNP A0JD36
m	133	ARG	-	linker	UNP A0JD36
m	134	VAL	-	linker	UNP A0JD36

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
m	135	THR	-	linker	UNP A0JD36
m	136	VAL	-	linker	UNP A0JD36
m	137	GLU	-	linker	UNP A0JD36
m	138	PRO	-	linker	UNP A0JD36
m	139	ARG	-	linker	UNP A0JD36

- Molecule 6 is a protein called Signal peptide,flag tag,T cell receptor gamma variable 9,T cell receptor gamma constant 1,T cell receptor gamma variable 9,T cell receptor gamma constant 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	n	35	Total	C	N	O	S	0	0
			284	193	41	47	3		

There are 21 discrepancies between the modelled and reference sequences:

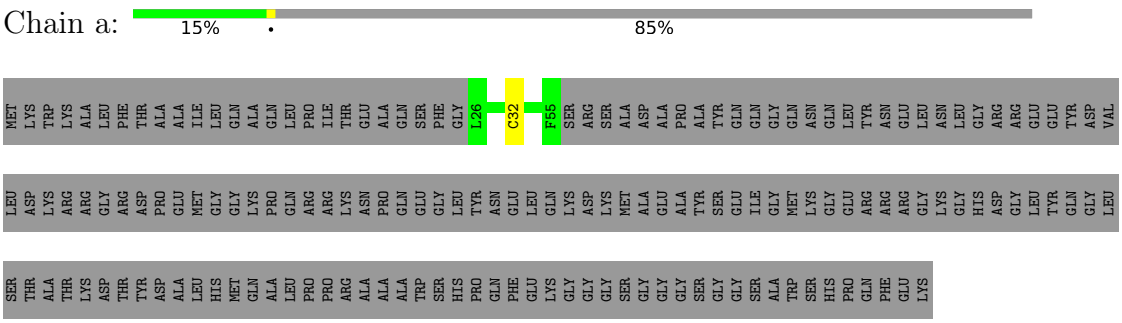
Chain	Residue	Modelled	Actual	Comment	Reference
n	123	ALA	-	linker	UNP Q99603
n	124	GLN	-	linker	UNP Q99603
n	125	GLN	-	linker	UNP Q99603
n	126	GLU	-	linker	UNP Q99603
n	127	LEU	-	linker	UNP Q99603
n	128	GLY	-	linker	UNP Q99603
n	129	LYS	-	linker	UNP Q99603
n	130	LYS	-	linker	UNP Q99603
n	131	ILE	-	linker	UNP Q99603
n	132	LYS	-	linker	UNP Q99603
n	133	VAL	-	linker	UNP Q99603
n	134	PHE	-	linker	UNP Q99603
n	135	GLY	-	linker	UNP Q99603
n	136	PRO	-	linker	UNP Q99603
n	137	GLY	-	linker	UNP Q99603
n	138	THR	-	linker	UNP Q99603
n	139	LYS	-	linker	UNP Q99603
n	140	LEU	-	linker	UNP Q99603
n	141	ILE	-	linker	UNP Q99603
n	142	ILE	-	linker	UNP Q99603
n	143	THR	-	linker	UNP Q99603

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.

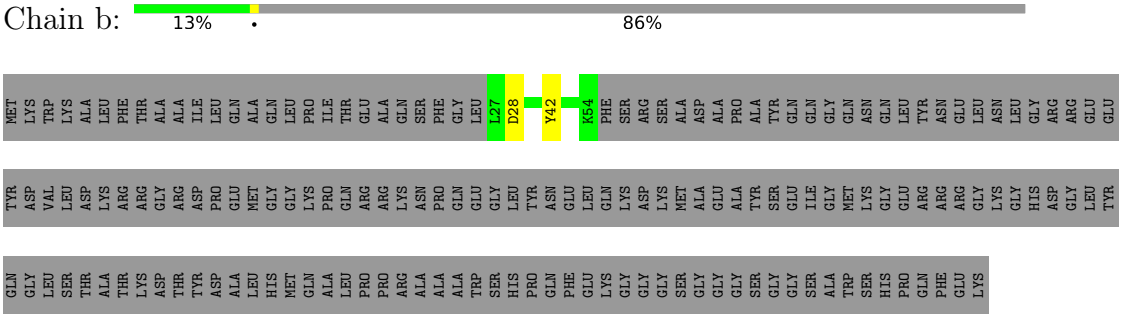
• Molecule 1: T-cell surface glycoprotein CD3 zeta chain

Chain a:



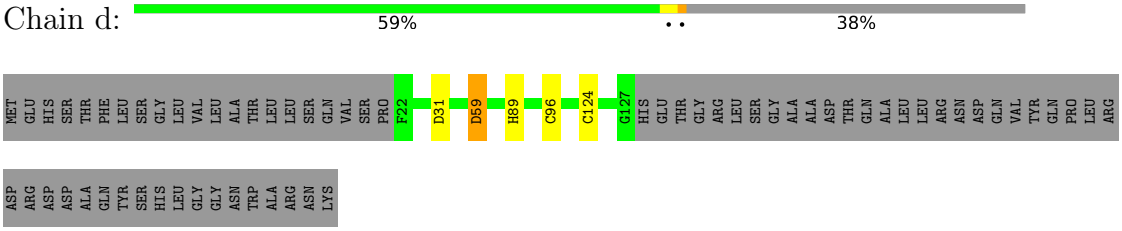
• Molecule 1: T-cell surface glycoprotein CD3 zeta chain

Chain b:



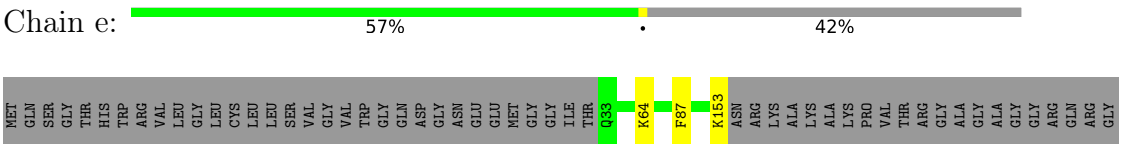
• Molecule 2: T-cell surface glycoprotein CD3 delta chain

Chain d:



• Molecule 3: T-cell surface glycoprotein CD3 epsilon chain

Chain e:

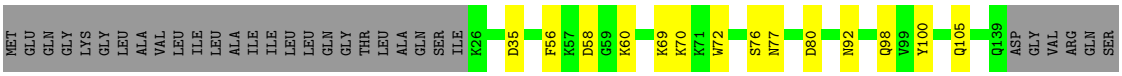


• Molecule 3: T-cell surface glycoprotein CD3 epsilon chain

Chain f:

• Molecule 4: T-cell surface glycoprotein CD3 gamma chain

Chain g:

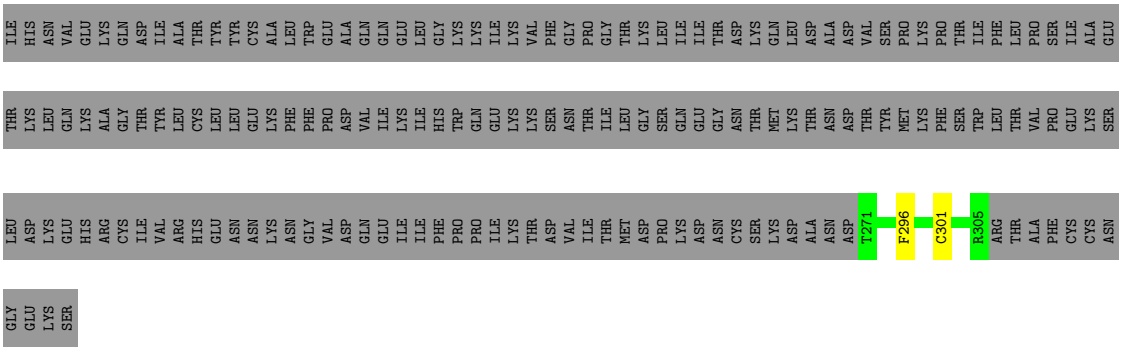


• Molecule 5: Signal peptide,flag tag,T cell receptor delta variable 2,T cell receptor delta constant,T cell receptor alpha chain constant,T cell receptor delta variable 2,T cell receptor delta constant,T cell receptor alpha chain constant

Chain m:

• Molecule 6: Signal peptide,flag tag,T cell receptor gamma variable 9,T cell receptor gamma constant 1,T cell receptor gamma variable 9,T cell receptor gamma constant 1

Chain n:


ILE HIS ASN VAL GLU LYS GLN ASP ILE ALA THR TYR TYR CYS ALA LEU TRP GLU ALA GLN GLN GLU LEU GLY LYS LYS ILE LYS VAL PHE GLY PRO GLY THR LYS LEU ILE ILE THR ASP LYS GLN LEU ASP ALA ASP VAL SER PRO LYS PRO THR ILE PHE LEU PRO SER ILE ALA GLU
THR LYS LEU GLN LYS ALA GLY THR TYR LEU CYS LEU LEU GLU LYS PHE PHE PRO ASP VAL ILE LYS ILE HIS TRP GLN GLU LYS LYS SER ASN THR ILE LEU GLY SER GLN GLU GLY ASN THR MET LYS THR ASN ASP THR TYR MET LYS PHE SER TRP LEU THR VAL PRO GLU LYS SER
LEU ASP LYS GLU HIS ARG CYS ILE VAL ARG HIS GLU ASN ASN LYS ASN GLY VAL ASP GLN GLU ILE ILE PHE PRO PRO ILE LYS THR ASP VAL ILE THR MET ASP PRO LYS ASP ASN CYS SER LYS ASP ALA ASN ASP T271 F296 C301 R305 ARG THR ALA PHE CYS CYS ASN
GLY GLU LYS SER

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	275838	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

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	0.26	0/250	0.48	0/338
1	b	0.28	0/230	0.49	0/311
2	d	0.28	0/833	0.70	2/1132 (0.2%)
3	e	0.31	0/980	0.57	0/1331
3	f	0.29	0/953	0.56	0/1294
4	g	0.29	0/917	0.65	0/1233
5	m	0.26	0/295	0.66	0/397
6	n	0.27	0/288	0.52	0/392
All	All	0.29	0/4746	0.60	2/6428 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	d	96	CYS	CA-CB-SG	5.17	123.30	114.00
2	d	59	ASP	CB-CG-OD1	5.04	122.83	118.30

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	245	0	270	0	0
1	b	226	0	250	0	0
2	d	821	0	831	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	e	959	0	919	0	0
3	f	933	0	899	0	0
4	g	899	0	893	0	0
5	m	292	0	326	0	0
6	n	284	0	312	0	0
All	All	4659	0	4700	0	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 17.

There are no clashes within the asymmetric unit.

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	28/195 (14%)	28 (100%)	0	0	100 100
1	b	26/195 (13%)	25 (96%)	0	1 (4%)	3 28
2	d	104/171 (61%)	87 (84%)	16 (15%)	1 (1%)	15 52
3	e	119/207 (58%)	107 (90%)	12 (10%)	0	100 100
3	f	114/207 (55%)	99 (87%)	12 (10%)	3 (3%)	5 35
4	g	112/182 (62%)	87 (78%)	23 (20%)	2 (2%)	8 42
5	m	35/310 (11%)	32 (91%)	1 (3%)	2 (6%)	1 21
6	n	33/332 (10%)	33 (100%)	0	0	100 100
All	All	571/1799 (32%)	498 (87%)	64 (11%)	9 (2%)	13 44

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	d	59	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	g	105	GLN
5	m	289	LEU
3	f	57	ILE
3	f	122	CYS

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	27/154 (18%)	26 (96%)	1 (4%)	34	60
1	b	25/154 (16%)	24 (96%)	1 (4%)	31	58
2	d	93/147 (63%)	90 (97%)	3 (3%)	39	63
3	e	109/177 (62%)	106 (97%)	3 (3%)	43	66
3	f	106/177 (60%)	99 (93%)	7 (7%)	16	46
4	g	96/155 (62%)	84 (88%)	12 (12%)	4	23
5	m	34/273 (12%)	31 (91%)	3 (9%)	10	37
6	n	33/298 (11%)	31 (94%)	2 (6%)	18	48
All	All	523/1535 (34%)	491 (94%)	32 (6%)	22	48

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	m	288	ARG
5	m	290	TRP
3	f	112	LEU
3	f	107	ASP
6	n	296	PHE

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

Mol	Chain	Res	Type
2	d	38	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
6	n	278	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

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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.