



# wwPDB X-ray Structure Validation Summary Report i

Jun 18, 2024 – 09:53 AM EDT

PDB ID : 4B2T  
Title : The crystal structures of the eukaryotic chaperonin CCT reveal its functional partitioning  
Authors : Kalisman, N.; Schroeder, G.F.; Levitt, M.  
Deposited on : 2012-07-17  
Resolution : 5.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>  
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 ①) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.20.1  
EDS : 2.37.1  
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) : 2.37.1

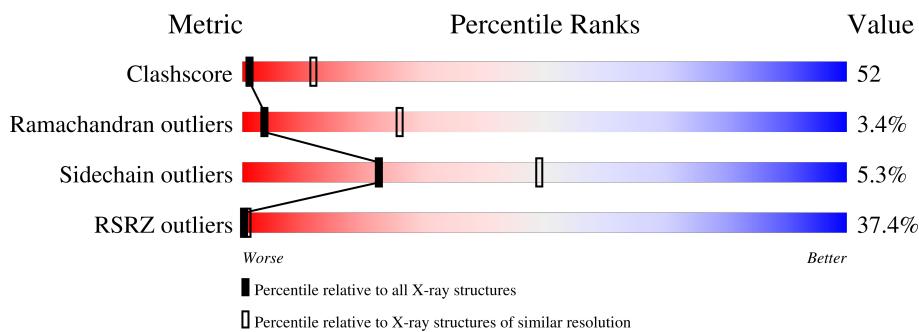
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## X-RAY DIFFRACTION

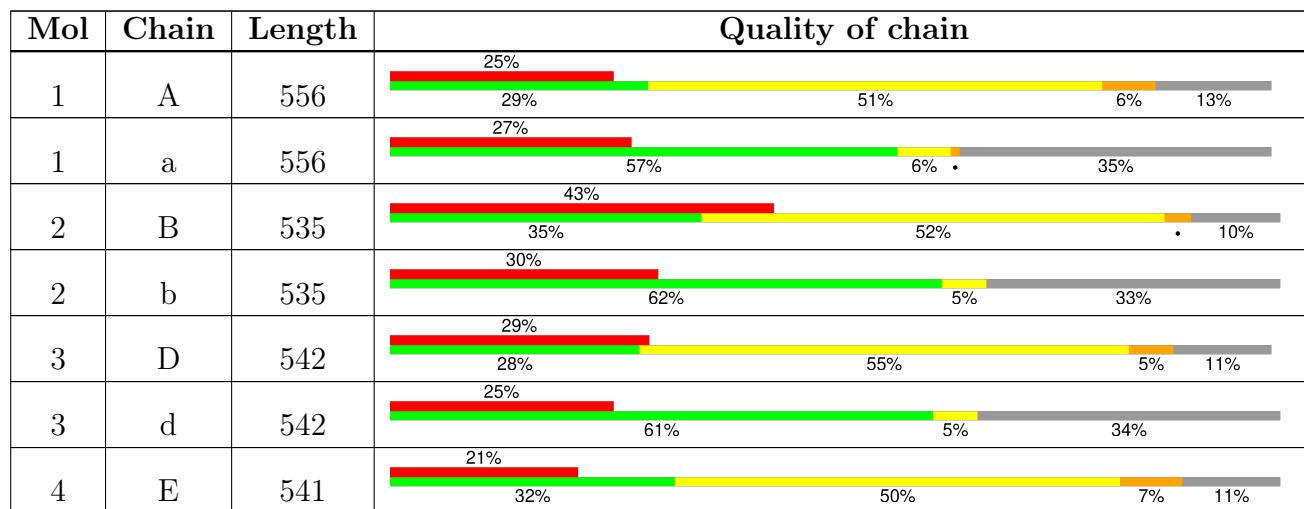
The reported resolution of this entry is 5.50 Å.

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1010 (7.10-3.90)
Ramachandran outliers	138981	1014 (7.12-3.82)
Sidechain outliers	138945	1191 (7.20-3.80)
RSRZ outliers	127900	1023 (7.08-3.76)

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.



Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain			
4	e	541	30%	61%	6%	34%
5	G	545	34%	31%	54%	12%
5	g	545	14%	61%	5%	34%
6	H	543	30%	34%	50%	11%
6	h	543	16%	61%	•	34%
7	Q	548	43%	28%	54%	5% 12%
7	q	548	28%	61%	• •	34%
8	Z	531	41%	33%	55%	• 9%
8	z	531	37%	84%	7%	9%

## 2 Entry composition [\(i\)](#)

There are 8 unique types of molecules in this entry. The entry contains 51877 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 T-COMPLEX PROTEIN 1 SUBUNIT ALPHA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	481	Total	C	N	O	S	0	0	0
			3625	2280	633	692	20			

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	a	359	Total	C	N	O	S	0	0	0
			2705	1703	469	520	13			

- Molecule 2 is a protein called T-COMPLEX PROTEIN 1 SUBUNIT BETA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	481	Total	C	N	O	S	0	0	0
			3602	2258	629	696	19			

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	b	359	Total	C	N	O	S	0	0	0
			2658	1652	469	524	13			

- Molecule 3 is a protein called T-COMPLEX PROTEIN 1 SUBUNIT DELTA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	481	Total	C	N	O	S	0	0	0
			3610	2259	627	703	21			

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	d	359	Total	C	N	O	S	0	0	0
			2690	1671	473	532	14			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	158	VAL	GLU	conflict	UNP Q2T9X2
D	510	LEU	GLN	conflict	UNP Q2T9X2
d	1158	VAL	GLU	conflict	UNP Q2T9X2
d	1510	LEU	GLN	conflict	UNP Q2T9X2

- Molecule 4 is a protein called T-COMPLEX PROTEIN 1 SUBUNIT EPSILON.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	481	Total C	N	O	S				
			3674	2299	644	703	28	0	0	0

4	e	359	Total C	N	O	S				
			2724	1688	486	528	22	0	0	0

- Molecule 5 is a protein called T-COMPLEX PROTEIN 1 SUBUNIT GAMMA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	G	481	Total C	N	O	S				
			3719	2326	661	705	27	0	0	0

5	g	359	Total C	N	O	S				
			2735	1711	480	523	21	0	0	0

- Molecule 6 is a protein called T-COMPLEX PROTEIN 1 SUBUNIT ETA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	H	481	Total C	N	O	S				
			3671	2320	633	693	25	0	0	0

6	h	359	Total C	N	O	S				
			2724	1719	472	517	16	0	0	0

- Molecule 7 is a protein called T-COMPLEX PROTEIN 1 SUBUNIT THETA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	Q	481	Total C	N	O	S				
			3673	2317	628	703	25	0	0	0

7	q	359	Total C	N	O	S				
			2739	1729	467	526	17	0	0	0

- Molecule 8 is a protein called T-COMPLEX PROTEIN 1 SUBUNIT ZETA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	Z	481	Total C	N	O	S				
			3664	2310	638	697	19	0	0	0

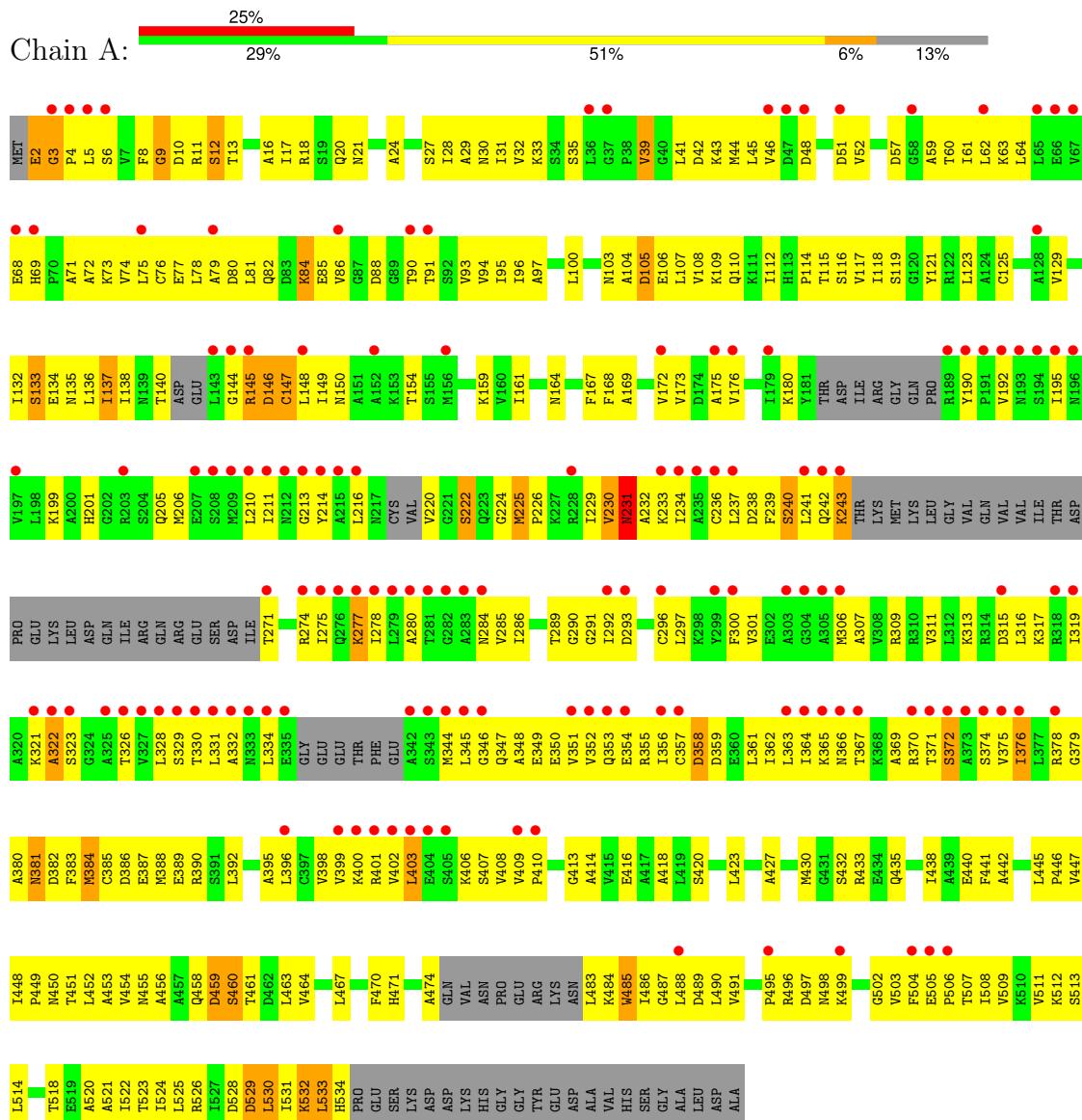
  

8	z	481	Total C	N	O	S				
			3664	2310	638	697	19	0	0	0

### 3 Residue-property plots [\(i\)](#)

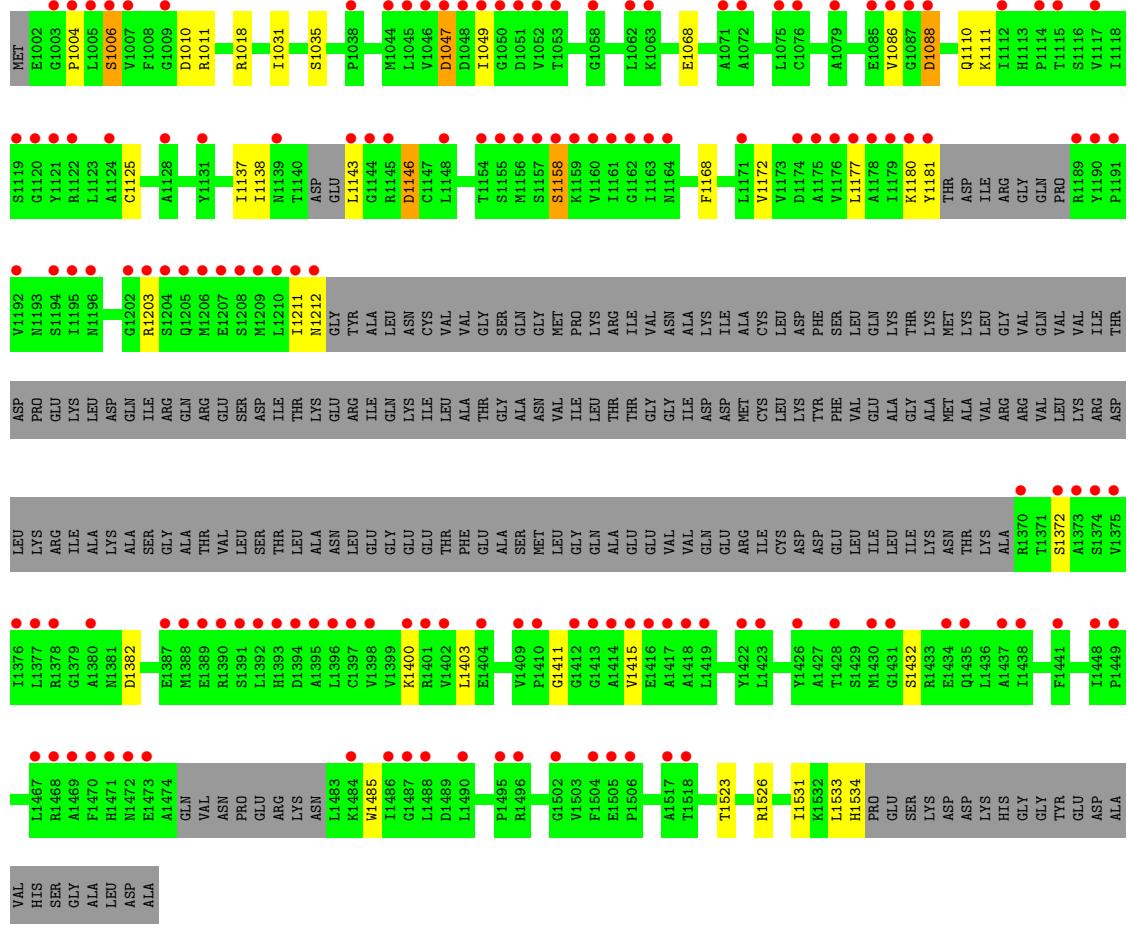
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: T-COMPLEX PROTEIN 1 SUBUNIT ALPHA



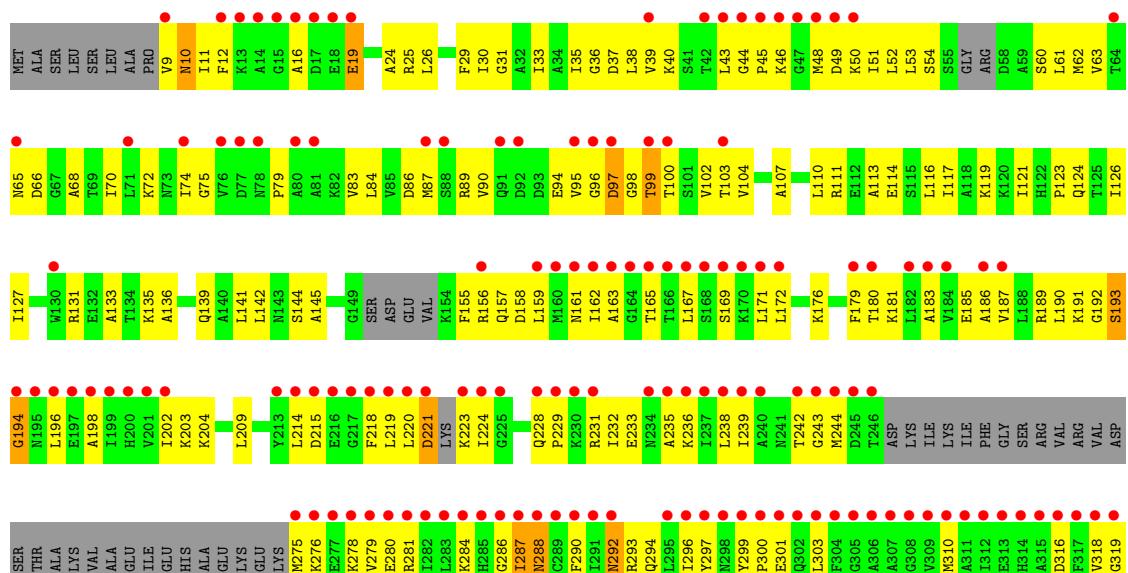
- Molecule 1: T-COMPLEX PROTEIN 1 SUBUNIT ALPHA

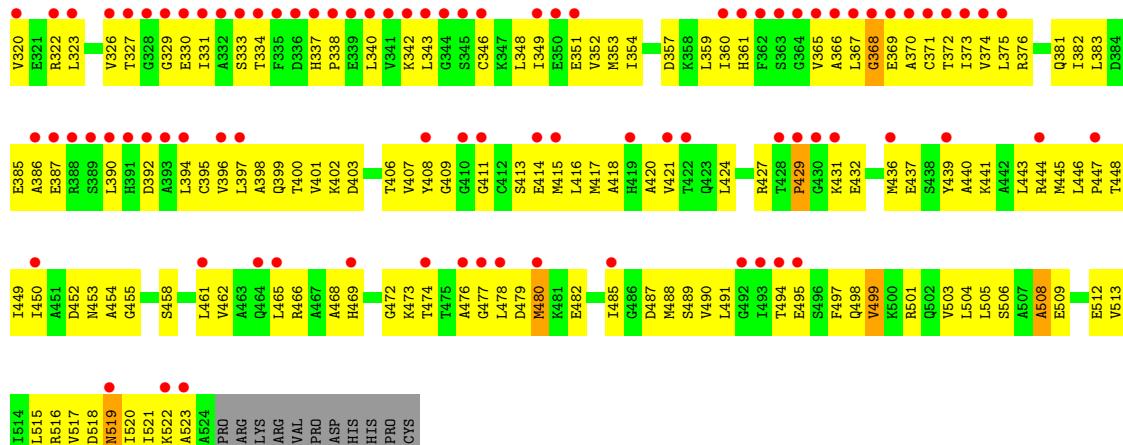




- Molecule 2: T-COMPLEX PROTEIN 1 SUBUNIT BETA

Chain B: 43% (Red), 35% (Green), 52% (Yellow), 10% (Grey)

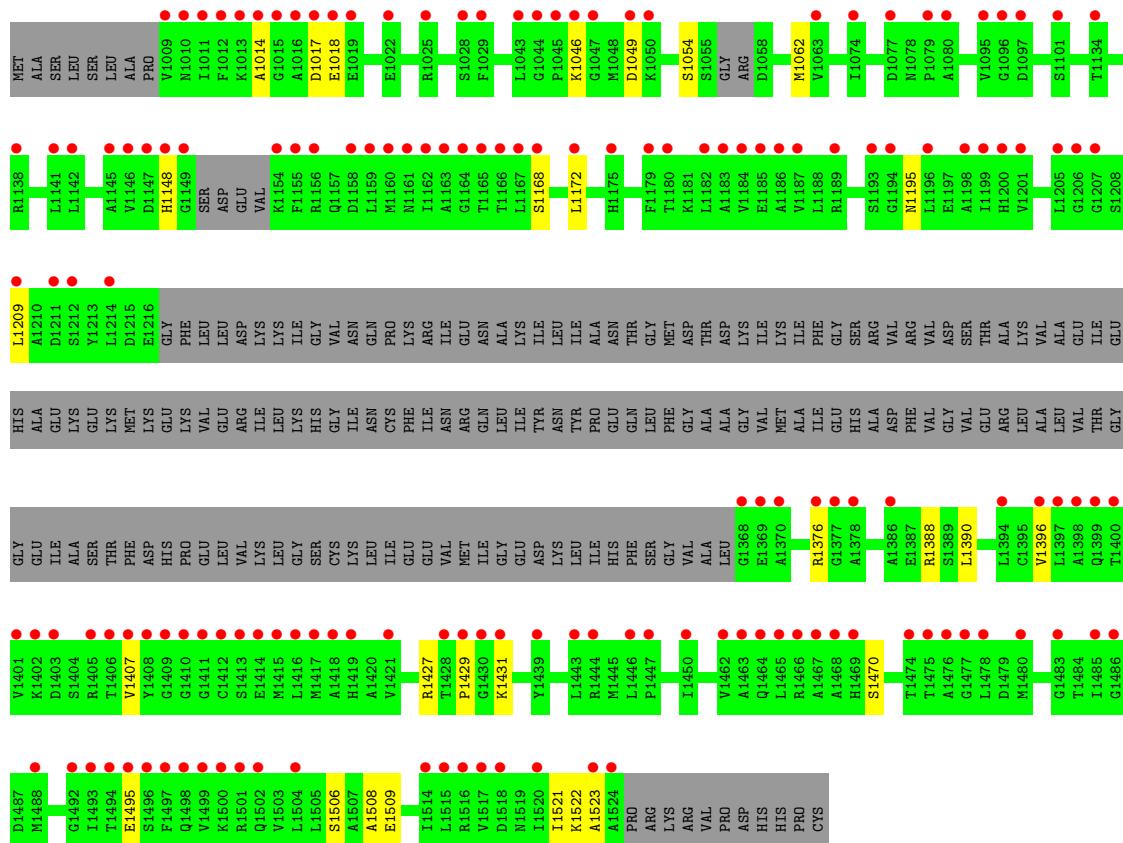




- Molecule 2: T-COMPLEX PROTEIN 1 SUBUNIT BETA

A horizontal bar chart illustrating the distribution of Chain b across four categories. The categories are represented by colored bars: Red, Green, Grey, and White. The percentages for each category are: Red (30%), Green (62%), Grey (5%), and White (33%).

Category	Percentage
Red	30%
Green	62%
Grey	5%
White	33%

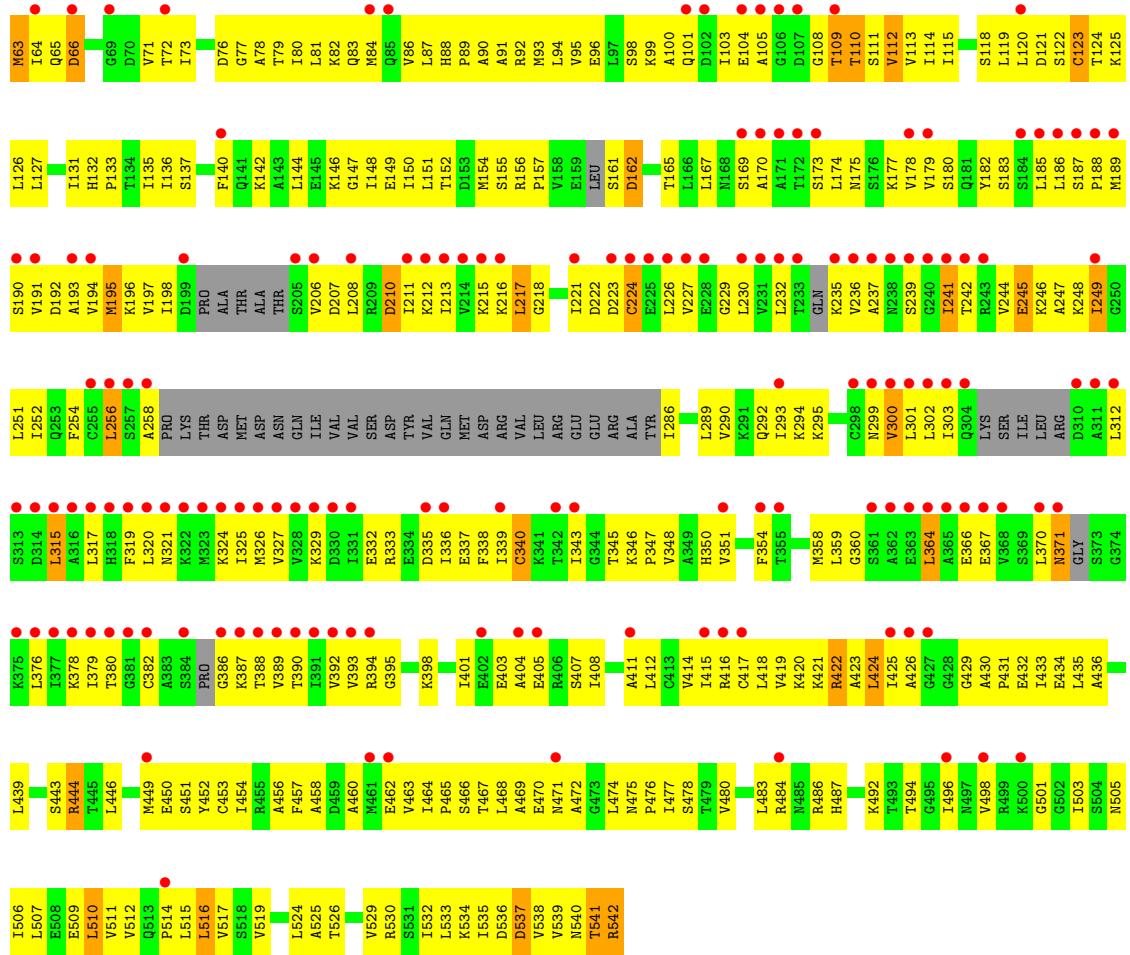


- Molecule 3: T-COMPLEX PROTEIN 1 SUBUNIT DELTA

A horizontal bar chart titled "Chain D" showing its distribution across four categories. The categories are represented by colored bars: Red (29%), Green (28%), Yellow (55%), and Orange (5%).

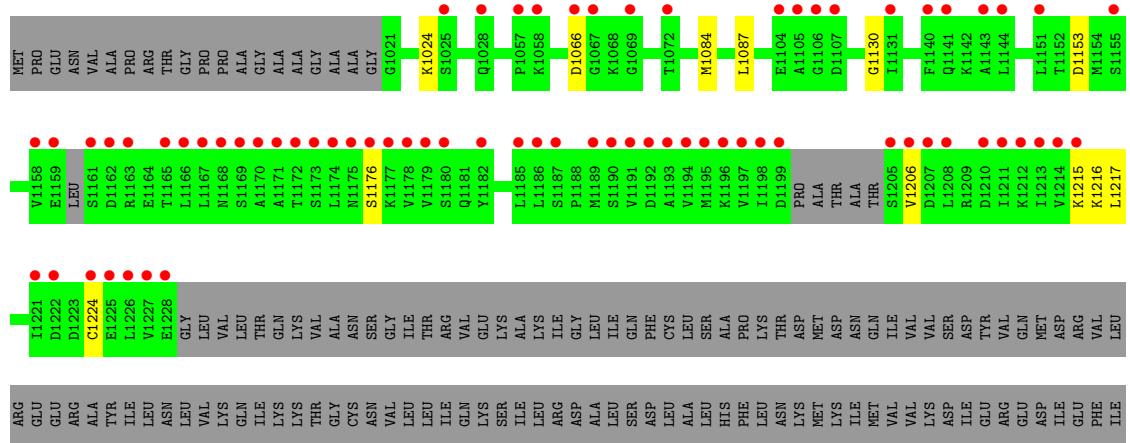
Category	Percentage
Red	29%
Green	28%
Yellow	55%
Orange	5%

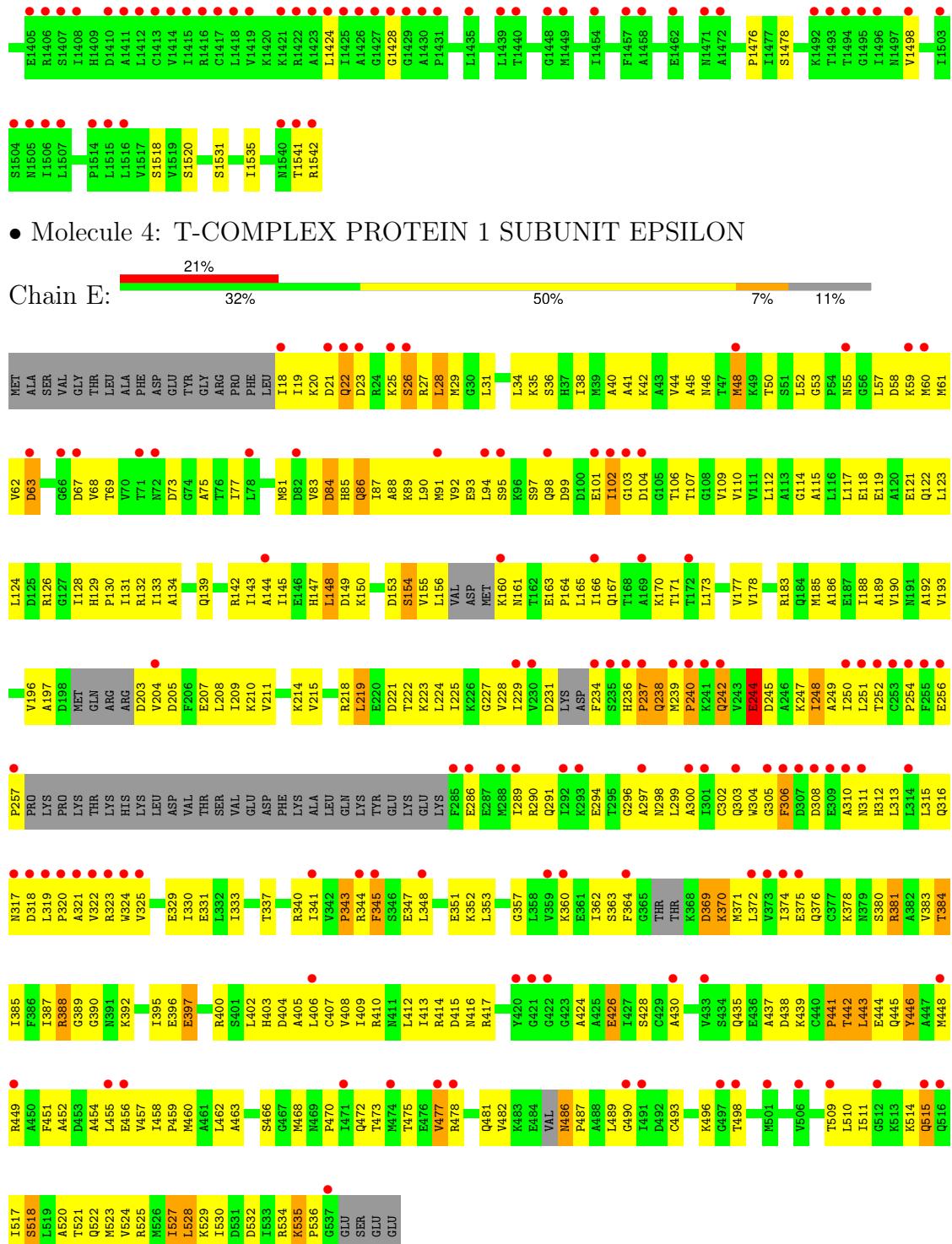




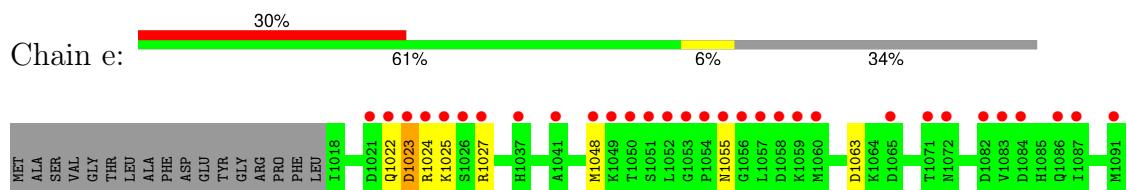
- Molecule 3: T-COMPLEX PROTEIN 1 SUBUNIT DELTA

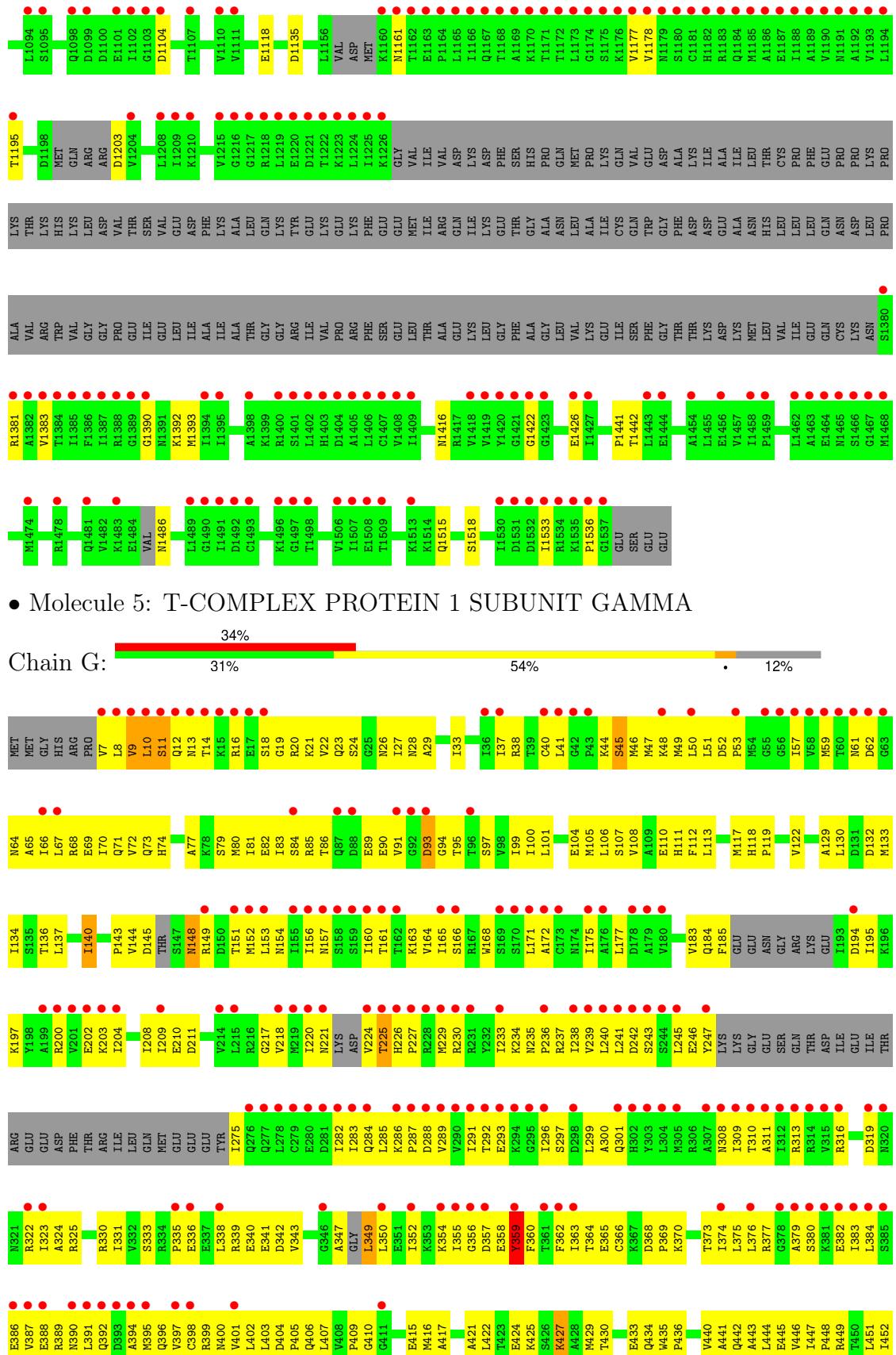
Chain d: 25% 61% 5% 34%

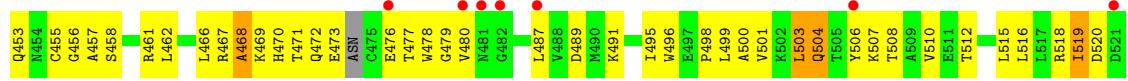




- Molecule 4: T-COMPLEX PROTEIN 1 SUBUNIT EPSILON

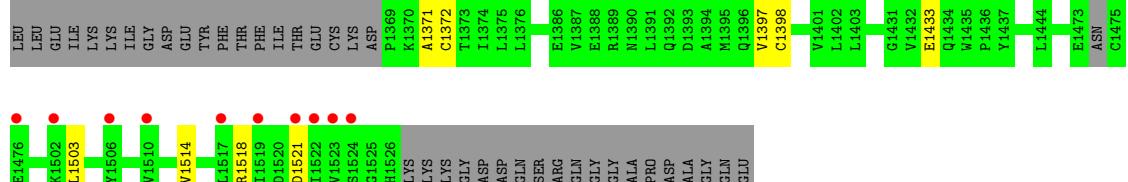
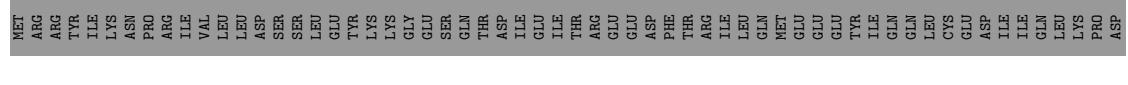
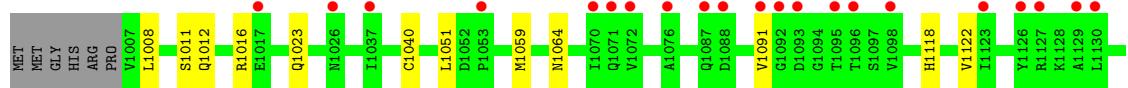






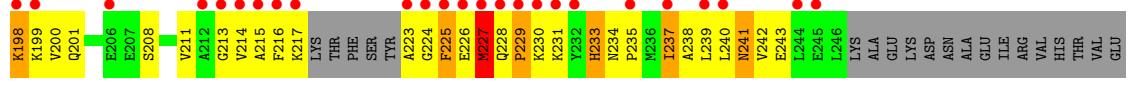
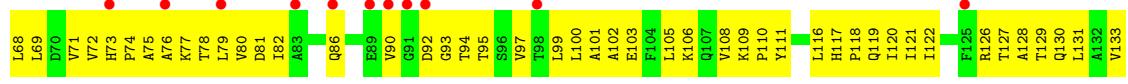
- Molecule 5: T-COMPLEX PROTEIN 1 SUBUNIT GAMMA

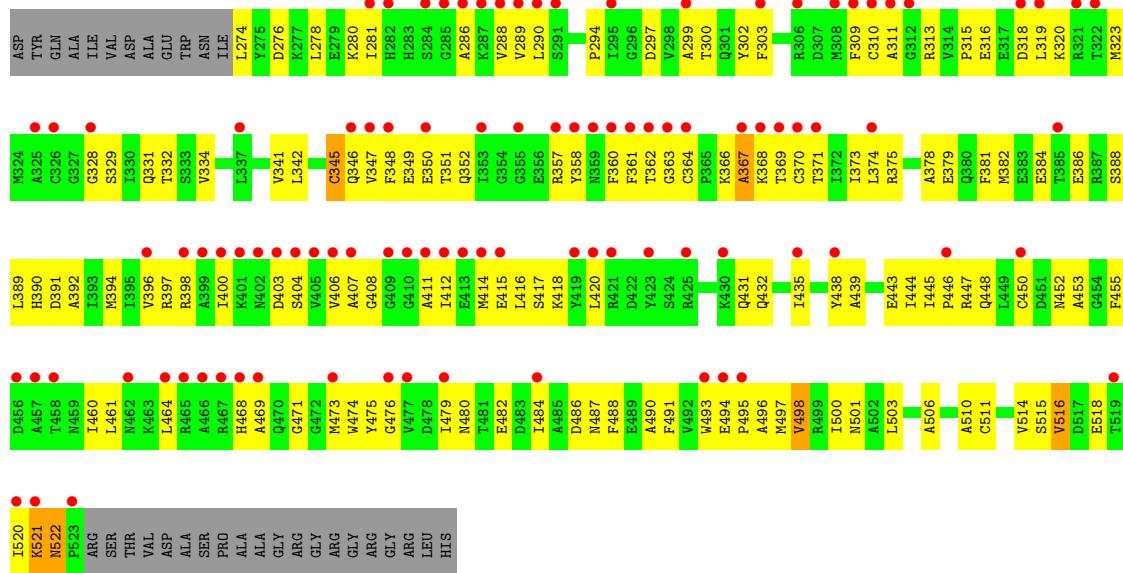
A horizontal bar chart titled "Chain g:" at the top left. The bar is divided into four segments: a red segment at the beginning labeled "14%", a long green segment labeled "61%", a short yellow segment labeled "5%", and a grey segment at the end labeled "34%".



- Molecule 6: T-COMPLEX PROTEIN 1 SUBUNIT ETA

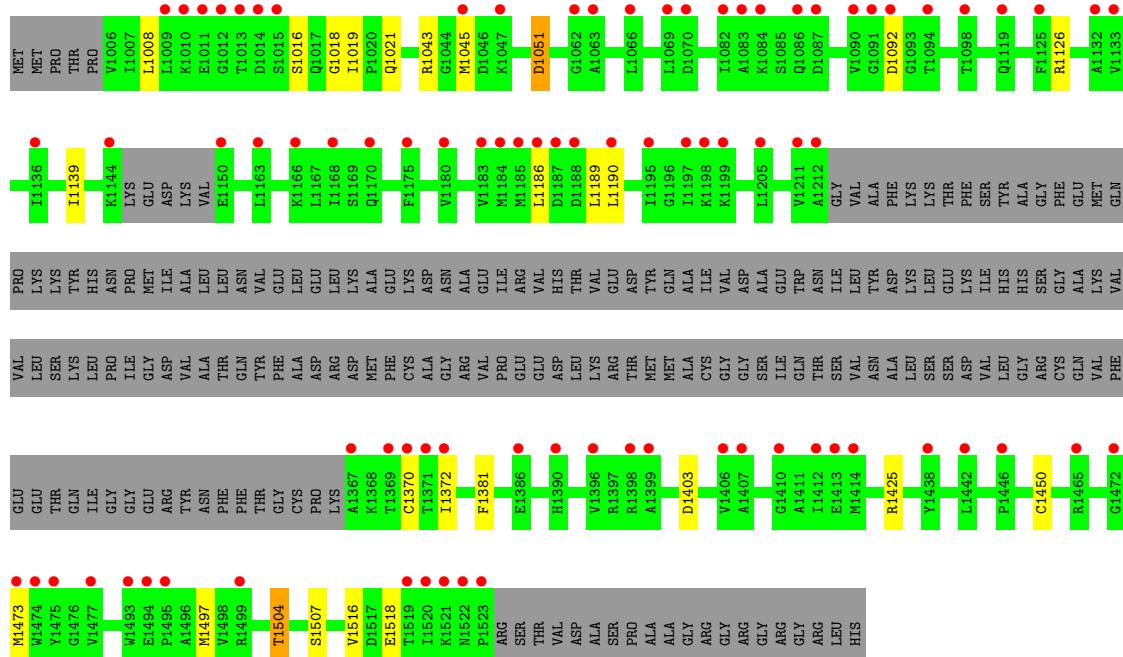
Category	Percentage
Red	30%
Green	34%
Yellow	50%
Grey	11%





- Molecule 6: T-COMPLEX PROTEIN 1 SUBUNIT ETA

Chain h: 16% • 61% • 34%

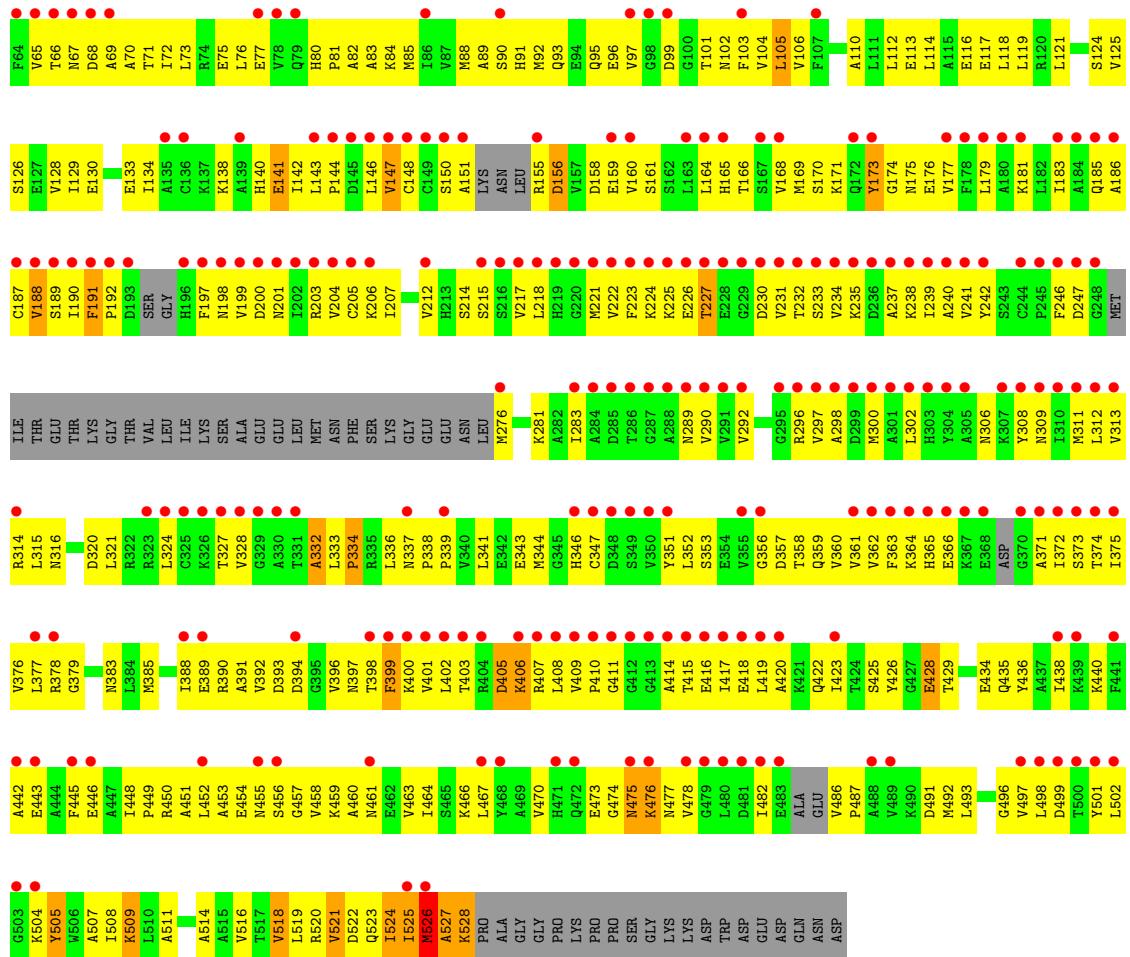


• Molecule 7: T-COMPLEX PROTEIN 1 SUBUNIT THETA

A horizontal bar chart illustrating the distribution of Chain Q across various categories. The total length of the bar is 100%.

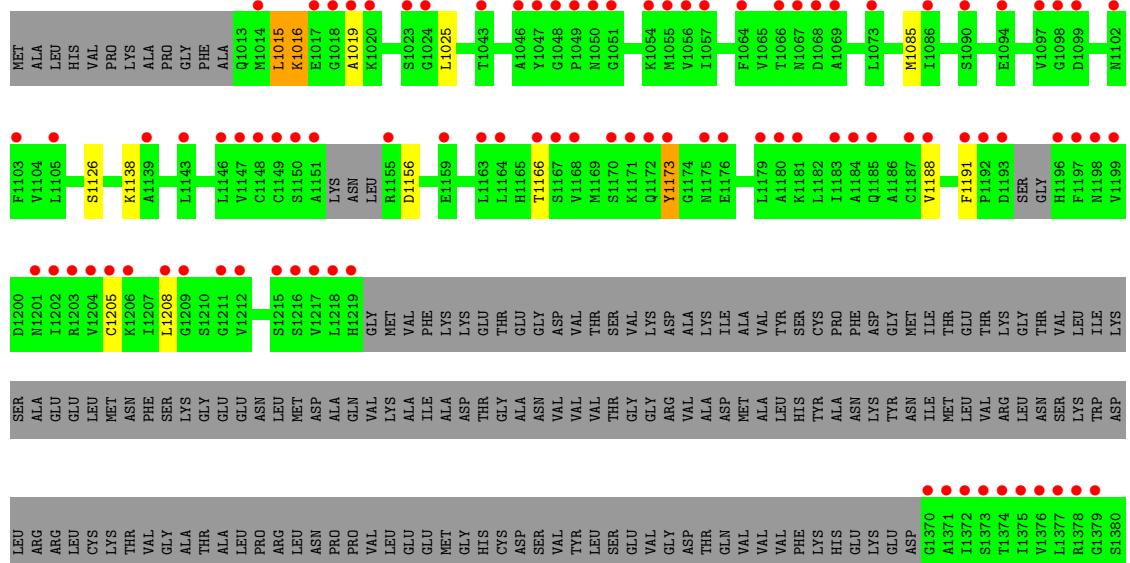
Category	Percentage
Green	28%
Yellow	43%
Orange	54%
Grey	5%
Red	12%

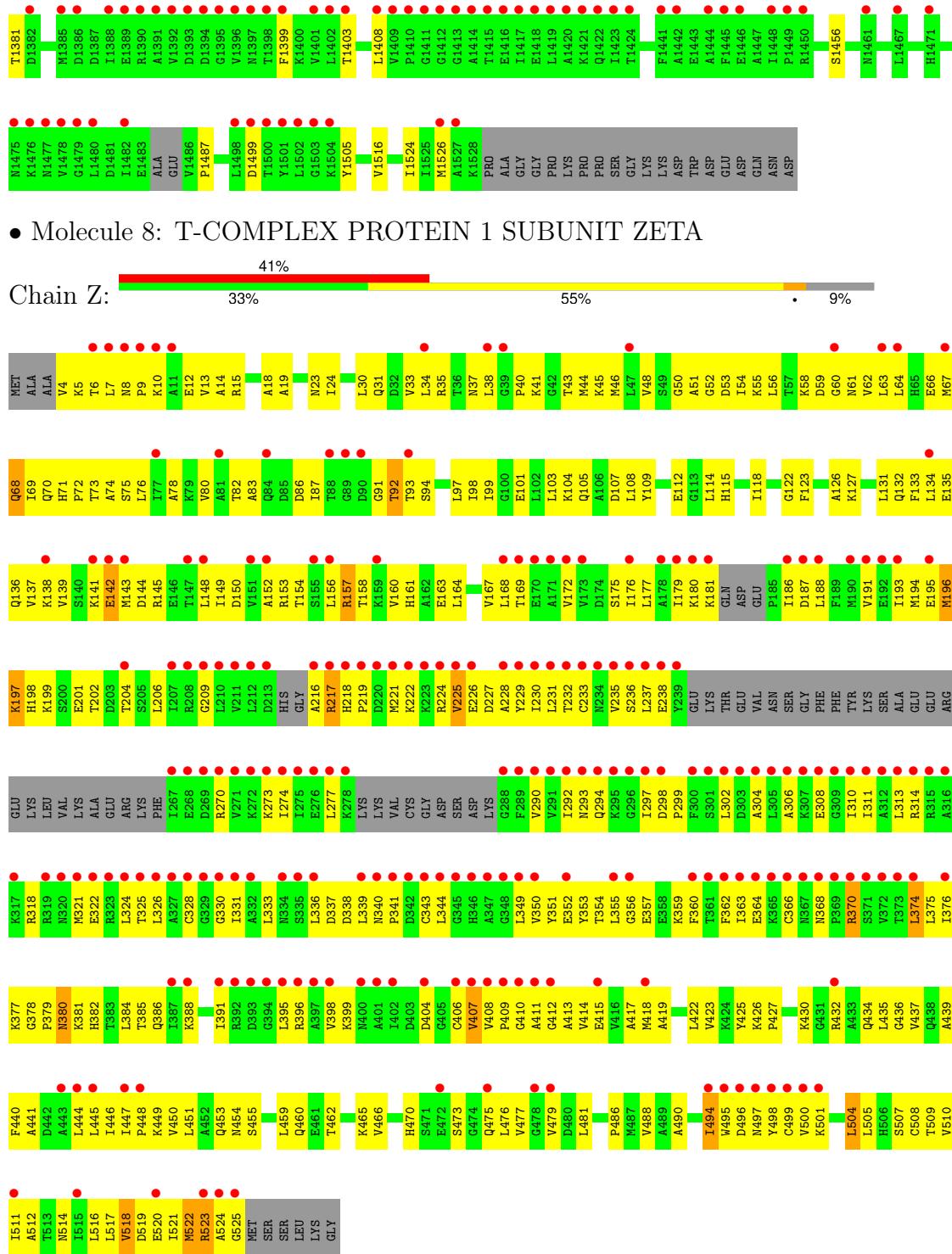




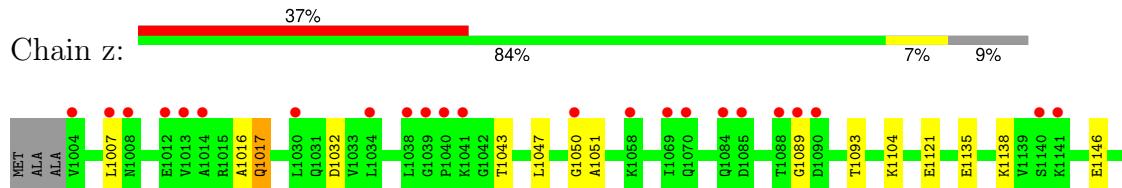
- Molecule 7: T-COMPLEX PROTEIN 1 SUBUNIT THETA

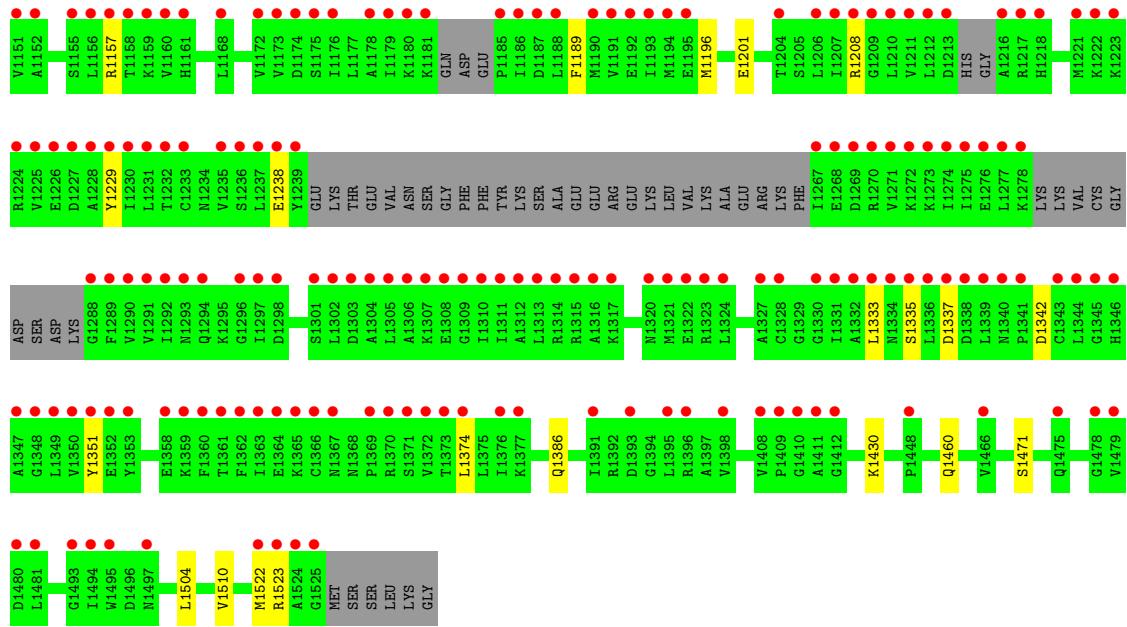
Chain q: 61%





- Molecule 8: T-COMPLEX PROTEIN 1 SUBUNIT ZETA





## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	272.70Å 313.50Å 158.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	200.00 – 5.50 97.58 – 5.44	Depositor EDS
% Data completeness (in resolution range)	99.1 (200.00-5.50) 98.4 (97.58-5.44)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) >$ <sup>1</sup>	1.37 (at 5.41Å)	Xtriage
Refinement program	CNS 1.3	Depositor
$R$ , $R_{free}$	0.340 , 0.399 0.342 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	257.8	Xtriage
Anisotropy	0.400	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 425.9	EDS
L-test for twinning <sup>2</sup>	$<  L  > = 0.32$ , $< L^2 > = 0.16$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.75	EDS
Total number of atoms	51877	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	277.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.63% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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 [\(i\)](#)

### 5.1 Standard geometry [\(i\)](#)

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	0.22	0/3657	0.44	0/4934
1	a	0.22	0/2733	0.46	0/3695
2	B	0.22	0/3638	0.43	0/4903
2	b	0.21	0/2680	0.44	0/3615
3	D	0.21	0/3632	0.46	0/4891
3	d	0.21	0/2707	0.45	0/3650
4	E	0.22	0/3712	0.44	0/4997
4	e	0.21	0/2743	0.44	0/3687
5	G	0.21	0/3758	0.45	0/5073
5	g	0.21	0/2763	0.46	0/3733
6	H	0.23	0/3716	0.43	0/5008
6	h	0.22	0/2751	0.45	0/3711
7	Q	0.23	0/3724	0.44	0/5032
7	q	0.23	0/2774	0.43	0/3746
8	Z	0.22	0/3702	0.44	0/4995
8	z	0.21	0/3702	0.45	0/4995
All	All	0.22	0/52392	0.44	0/70665

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 [\(i\)](#)

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	3625	0	3783	380	0
1	a	2705	0	2799	0	0
2	B	3602	0	3705	353	0
2	b	2658	0	2733	0	0
3	D	3610	0	3810	456	0
3	d	2690	0	2818	0	0
4	E	3674	0	3781	382	0
4	e	2724	0	2822	0	0
5	G	3719	0	3870	398	0
5	g	2735	0	2851	0	0
6	H	3671	0	3783	366	0
6	h	2724	0	2842	0	0
7	Q	3673	0	3719	383	0
7	q	2739	0	2777	0	0
8	Z	3664	0	3820	413	0
8	z	3664	0	3820	0	0
All	All	51877	0	53733	2912	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 52.

The worst 5 of 2912 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)
5:G:456:GLY:HA3	8:Z:118:ILE:HD11	1.37	1.06
1:A:211:ILE:HG22	1:A:213:GLY:H	1.21	1.04
3:D:540:ASN:HB3	3:D:542:ARG:HD3	1.36	1.03
3:D:31:ASP:HB3	3:D:36:ILE:HB	1.39	1.03
4:E:94:LEU:HD21	4:E:523:MET:HB2	1.39	1.03

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	471/556 (85%)	390 (83%)	63 (13%)	18 (4%)	3 24
1	a	351/556 (63%)	283 (81%)	50 (14%)	18 (5%)	2 19
2	B	473/535 (88%)	398 (84%)	62 (13%)	13 (3%)	5 31
2	b	353/535 (66%)	301 (85%)	42 (12%)	10 (3%)	5 30
3	D	469/542 (86%)	385 (82%)	70 (15%)	14 (3%)	4 28
3	d	351/542 (65%)	300 (86%)	42 (12%)	9 (3%)	5 31
4	E	471/541 (87%)	392 (83%)	60 (13%)	19 (4%)	3 23
4	e	351/541 (65%)	295 (84%)	39 (11%)	17 (5%)	2 21
5	G	471/545 (86%)	396 (84%)	62 (13%)	13 (3%)	5 30
5	g	351/545 (64%)	291 (83%)	51 (14%)	9 (3%)	5 31
6	H	473/543 (87%)	384 (81%)	73 (15%)	16 (3%)	3 26
6	h	353/543 (65%)	289 (82%)	54 (15%)	10 (3%)	5 30
7	Q	473/548 (86%)	406 (86%)	45 (10%)	22 (5%)	2 21
7	q	351/548 (64%)	292 (83%)	48 (14%)	11 (3%)	4 27
8	Z	473/531 (89%)	400 (85%)	57 (12%)	16 (3%)	3 26
8	z	473/531 (89%)	403 (85%)	57 (12%)	13 (3%)	5 31
All	All	6708/8682 (77%)	5605 (84%)	875 (13%)	228 (3%)	3 26

5 of 228 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	GLY
1	A	145	ARG
1	A	146	ASP
1	A	230	VAL
2	B	54	SER

### 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	395/461 (86%)	368 (93%)	27 (7%)	16	42
1	a	297/461 (64%)	269 (91%)	28 (9%)	8	29
2	B	382/429 (89%)	372 (97%)	10 (3%)	46	67
2	b	281/429 (66%)	263 (94%)	18 (6%)	17	44
3	D	407/454 (90%)	388 (95%)	19 (5%)	26	52
3	d	303/454 (67%)	285 (94%)	18 (6%)	19	46
4	E	400/455 (88%)	371 (93%)	29 (7%)	14	40
4	e	298/455 (66%)	283 (95%)	15 (5%)	24	50
5	G	416/470 (88%)	402 (97%)	14 (3%)	37	60
5	g	307/470 (65%)	288 (94%)	19 (6%)	18	45
6	H	394/445 (88%)	381 (97%)	13 (3%)	38	61
6	h	292/445 (66%)	274 (94%)	18 (6%)	18	45
7	Q	398/452 (88%)	380 (96%)	18 (4%)	27	53
7	q	296/452 (66%)	279 (94%)	17 (6%)	20	47
8	Z	398/440 (90%)	386 (97%)	12 (3%)	41	63
8	z	398/440 (90%)	374 (94%)	24 (6%)	19	46
All	All	5662/7212 (78%)	5363 (95%)	299 (5%)	22	49

5 of 299 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	g	1162	THR
8	z	1196	MET
5	g	1503	LEU
7	q	1085	MET
5	G	503	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 144 such sidechains are listed below:

Mol	Chain	Res	Type
6	h	1191	GLN
8	z	1503	GLN
6	h	1431	GLN
7	q	1523	GLN
6	H	25	ASN

### 5.3.3 RNA [\(i\)](#)

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains [\(i\)](#)

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 [\(i\)](#)

There are no ligands in this entry.

### 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.

## 6 Fit of model and data [\(i\)](#)

### 6.1 Protein, DNA and RNA chains [\(i\)](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	481/556 (86%)	1.59	140 (29%) 0   1	260, 260, 260, 260	0
1	a	359/556 (64%)	2.09	151 (42%) 0   0	282, 282, 282, 282	0
2	B	481/535 (89%)	2.59	229 (47%) 0   0	281, 281, 281, 281	0
2	b	359/535 (67%)	2.16	159 (44%) 0   0	272, 272, 272, 272	0
3	D	481/542 (88%)	1.82	158 (32%) 0   1	289, 289, 289, 289	0
3	d	359/542 (66%)	1.78	134 (37%) 0   1	266, 266, 266, 266	0
4	E	481/541 (88%)	1.26	112 (23%) 0   2	262, 262, 262, 262	0
4	e	359/541 (66%)	2.37	164 (45%) 0   0	283, 283, 283, 283	0
5	G	481/545 (88%)	1.97	185 (38%) 0   1	283, 283, 283, 283	0
5	g	359/545 (65%)	1.18	76 (21%) 0   2	275, 275, 275, 275	0
6	H	481/543 (88%)	1.70	163 (33%) 0   1	272, 272, 272, 272	0
6	h	359/543 (66%)	1.18	85 (23%) 0   2	258, 258, 258, 258	0
7	Q	481/548 (87%)	2.56	233 (48%) 0   0	308, 308, 308, 308	0
7	q	359/548 (65%)	2.01	156 (43%) 0   0	261, 261, 261, 261	0
8	Z	481/531 (90%)	2.52	219 (45%) 0   0	297, 297, 297, 297	0
8	z	481/531 (90%)	2.37	194 (40%) 0   1	276, 276, 276, 276	0
All	All	6842/8682 (78%)	1.96	2558 (37%) 0   1	258, 276, 308, 308	0

The worst 5 of 2558 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
8	Z	288	GLY	25.1
8	z	1209	GLY	21.8
6	H	224	GLY	18.3
8	z	1288	GLY	16.3
4	e	1226	LYS	15.8

## 6.2 Non-standard residues in protein, DNA, RNA chains [\(i\)](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [\(i\)](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [\(i\)](#)

There are no ligands in this entry.

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.