wwPDB X-ray Structure Validation Summary Report (i)

Jan 28, 2024 – 12:55 AM EST

PDB ID	:	1GRL
Title	:	THE CRYSTAL STRUCTURE OF THE BACTERIAL CHAPERONIN GROEL AT 2.8 ANGSTROMS
Authors	:	Braig, K.; Otwinowski, Z.; Hegde, R.; Boisvert, D.C.; Joachimiak, A.; Horwich, A.L.; Sigler, P.B.
Deposited on	:	1995-03-07
Resolution	:	2.80 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*
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 (i)) were used in the production of this report:

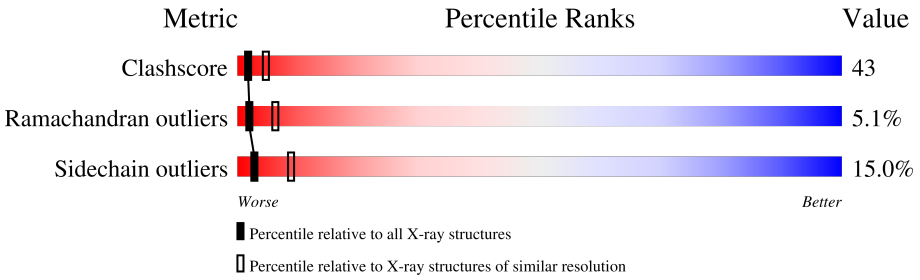
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.80 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	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	548	39% 46% 8% • 5%
1	B	548	39% 46% 8% • 5%
1	C	548	39% 46% 8% • 5%
1	D	548	38% 47% 7% • 5%
1	E	548	39% 47% 8% • 5%
1	F	548	39% 46% 8% • 5%
1	G	548	39% 47% 7% • 5%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 29274 atoms, of which 5278 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 GROEL (HSP60 CLASS).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	518	Total	C	H	N	O	S	0	0	59
			4182	2149	754	581	681	17			
1	B	518	Total	C	H	N	O	S	0	0	59
			4182	2149	754	581	681	17			
1	C	518	Total	C	H	N	O	S	0	0	59
			4182	2149	754	581	681	17			
1	D	518	Total	C	H	N	O	S	0	0	59
			4182	2149	754	581	681	17			
1	E	518	Total	C	H	N	O	S	0	0	59
			4182	2149	754	581	681	17			
1	F	518	Total	C	H	N	O	S	0	0	59
			4182	2149	754	581	681	17			
1	G	518	Total	C	H	N	O	S	0	0	59
			4182	2149	754	581	681	17			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	13	GLY	ARG	conflict	UNP P06139
A	126	VAL	ALA	conflict	UNP P06139
A	267	MET	ILE	conflict	UNP P06139
B	13	GLY	ARG	conflict	UNP P06139
B	126	VAL	ALA	conflict	UNP P06139
B	267	MET	ILE	conflict	UNP P06139
C	13	GLY	ARG	conflict	UNP P06139
C	126	VAL	ALA	conflict	UNP P06139
C	267	MET	ILE	conflict	UNP P06139
D	13	GLY	ARG	conflict	UNP P06139
D	126	VAL	ALA	conflict	UNP P06139
D	267	MET	ILE	conflict	UNP P06139
E	13	GLY	ARG	conflict	UNP P06139
E	126	VAL	ALA	conflict	UNP P06139
E	267	MET	ILE	conflict	UNP P06139

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	13	GLY	ARG	conflict	UNP P06139
F	126	VAL	ALA	conflict	UNP P06139
F	267	MET	ILE	conflict	UNP P06139
G	13	GLY	ARG	conflict	UNP P06139
G	126	VAL	ALA	conflict	UNP P06139
G	267	MET	ILE	conflict	UNP P06139

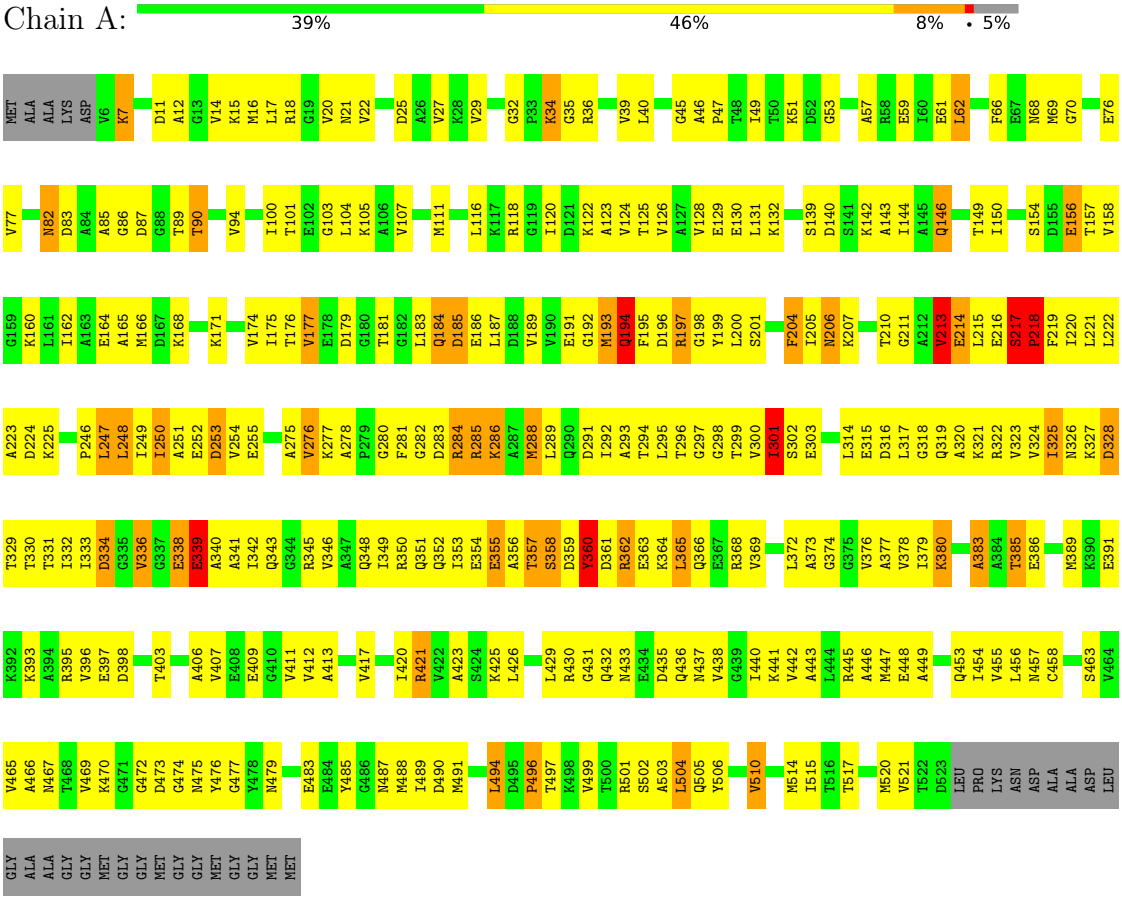
3 Residue-property plots (i)

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.

Note EDS was not executed.

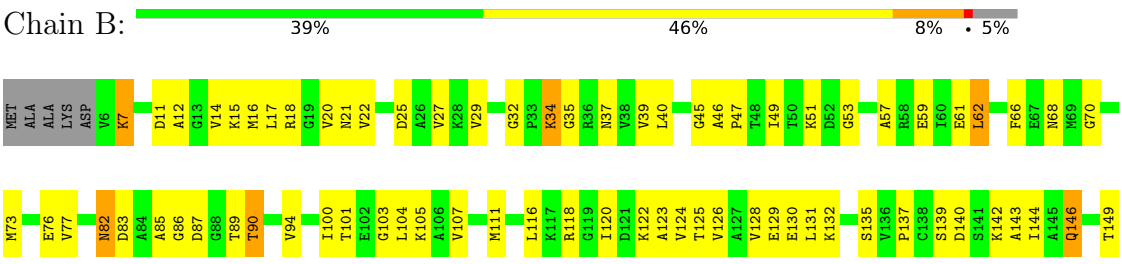
• Molecule 1: GROEL (HSP60 CLASS)

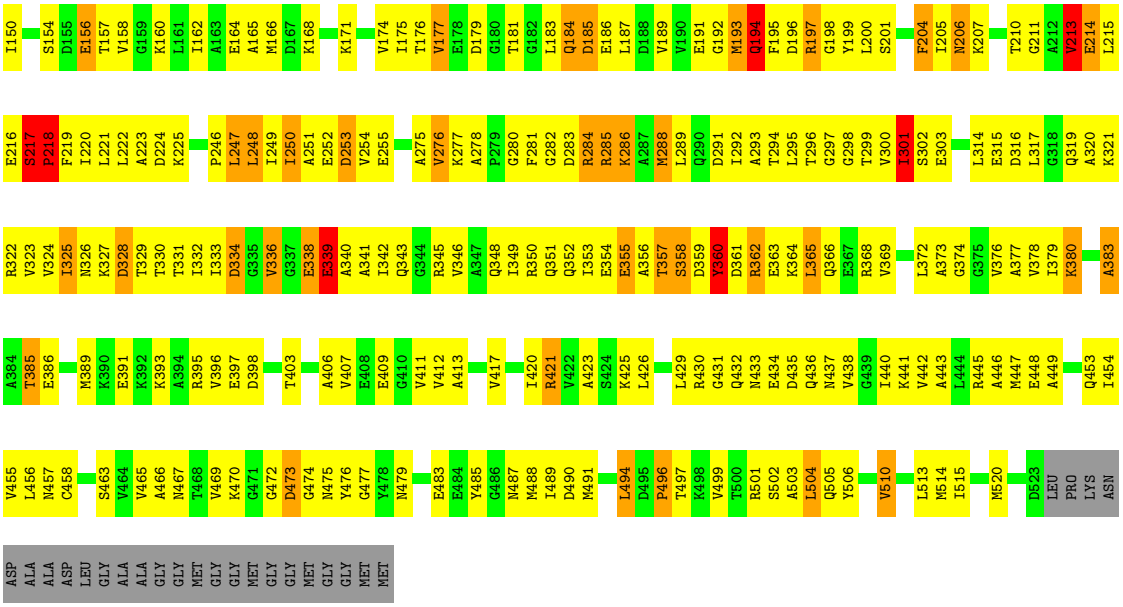
Chain A:



• Molecule 1: GROEL (HSP60 CLASS)

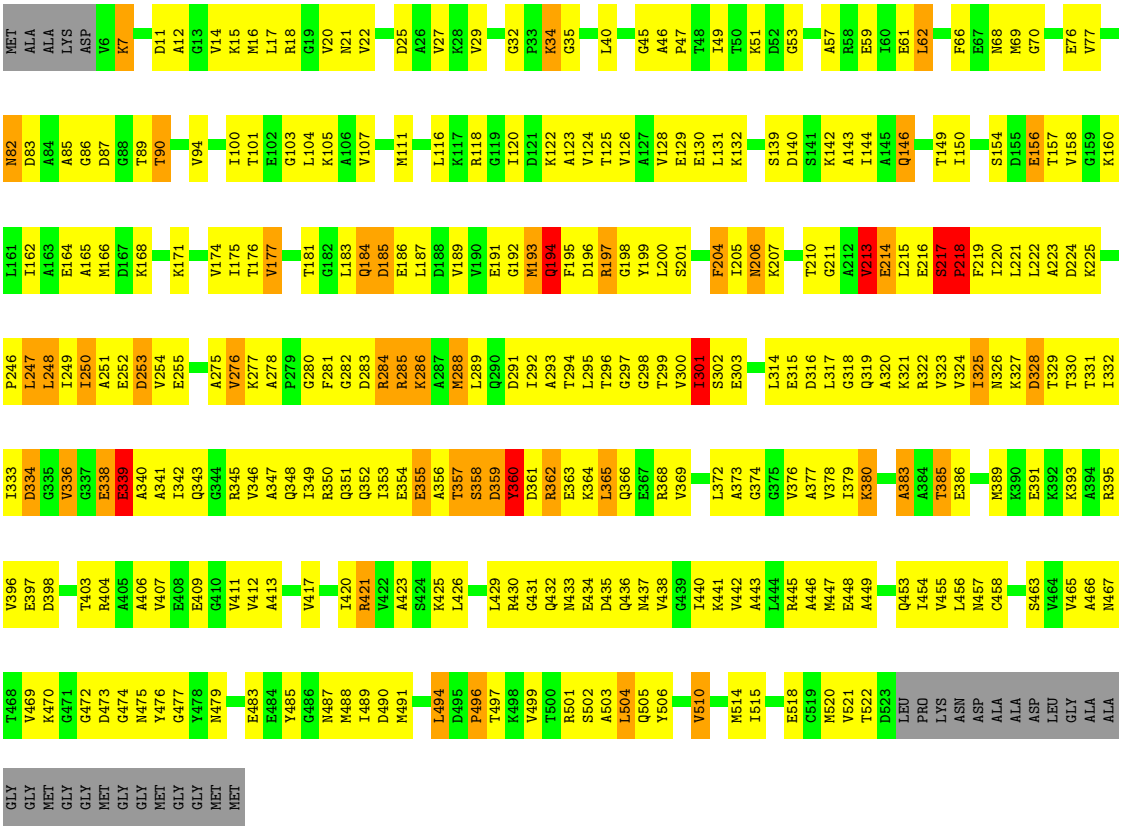
Chain B:





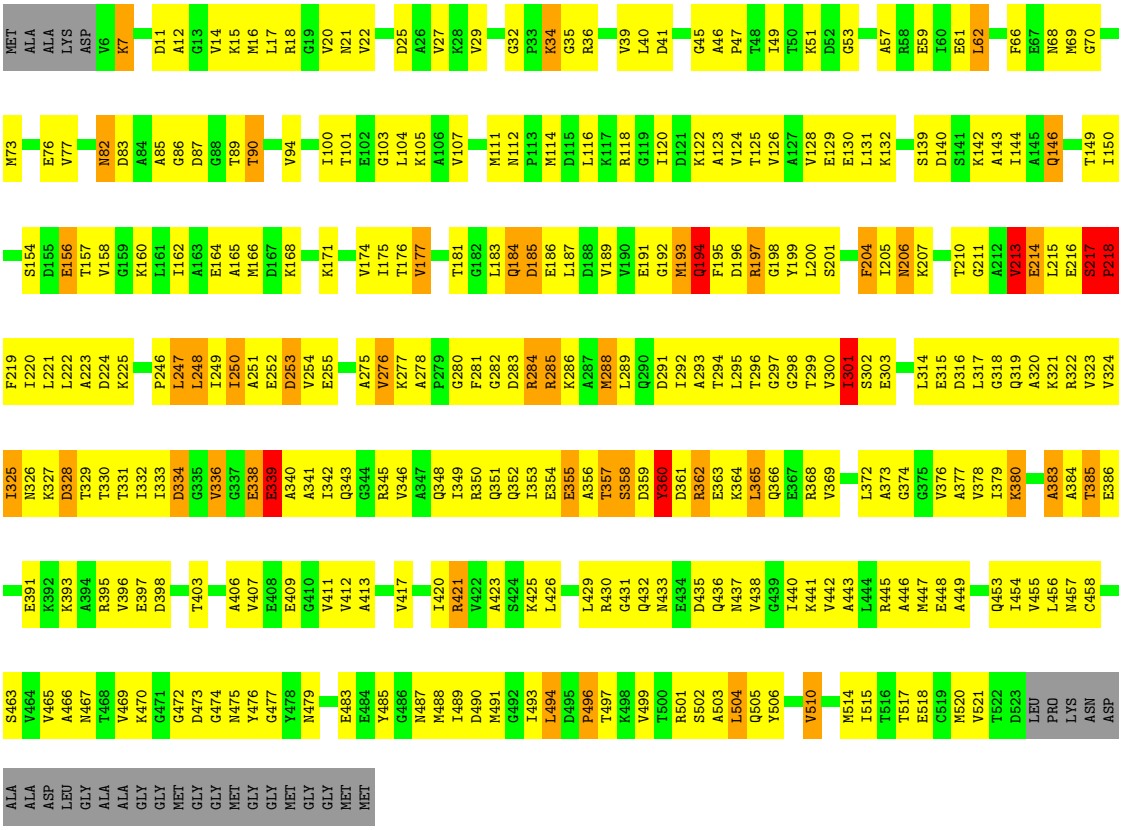
• Molecule 1: GROEL (HSP60 CLASS)

Chain C:



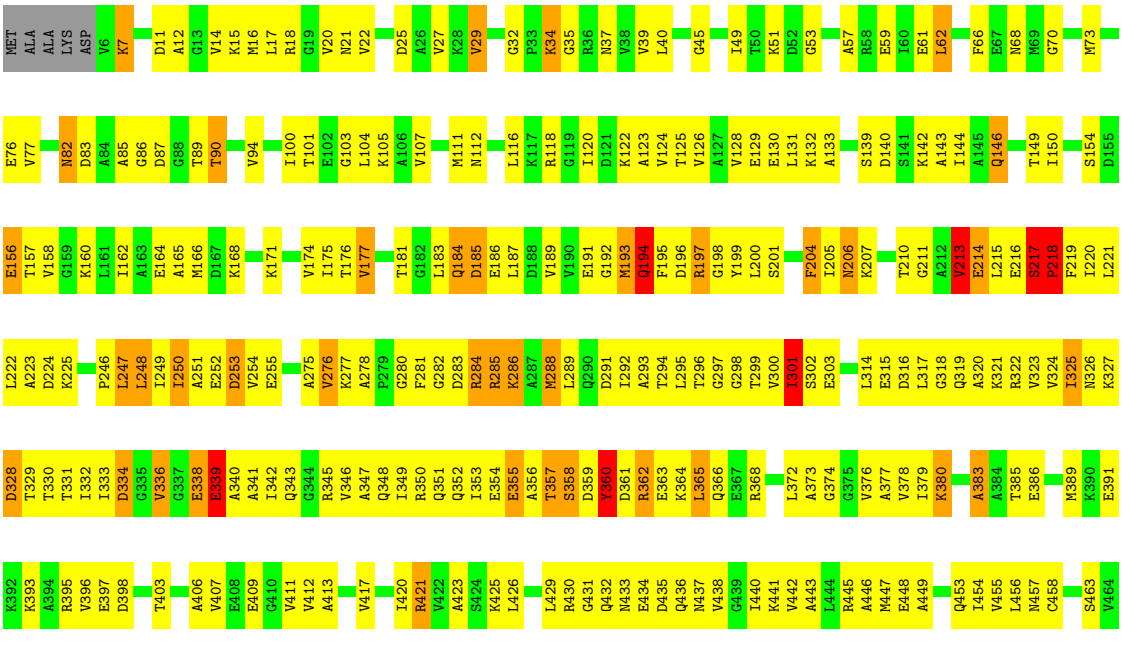
• Molecule 1: GROEL (HSP60 CLASS)

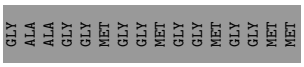
Chain D:



• Molecule 1: GROEL (HSP60 CLASS)

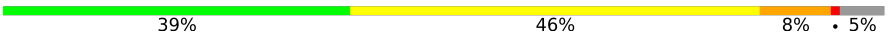
Chain E:

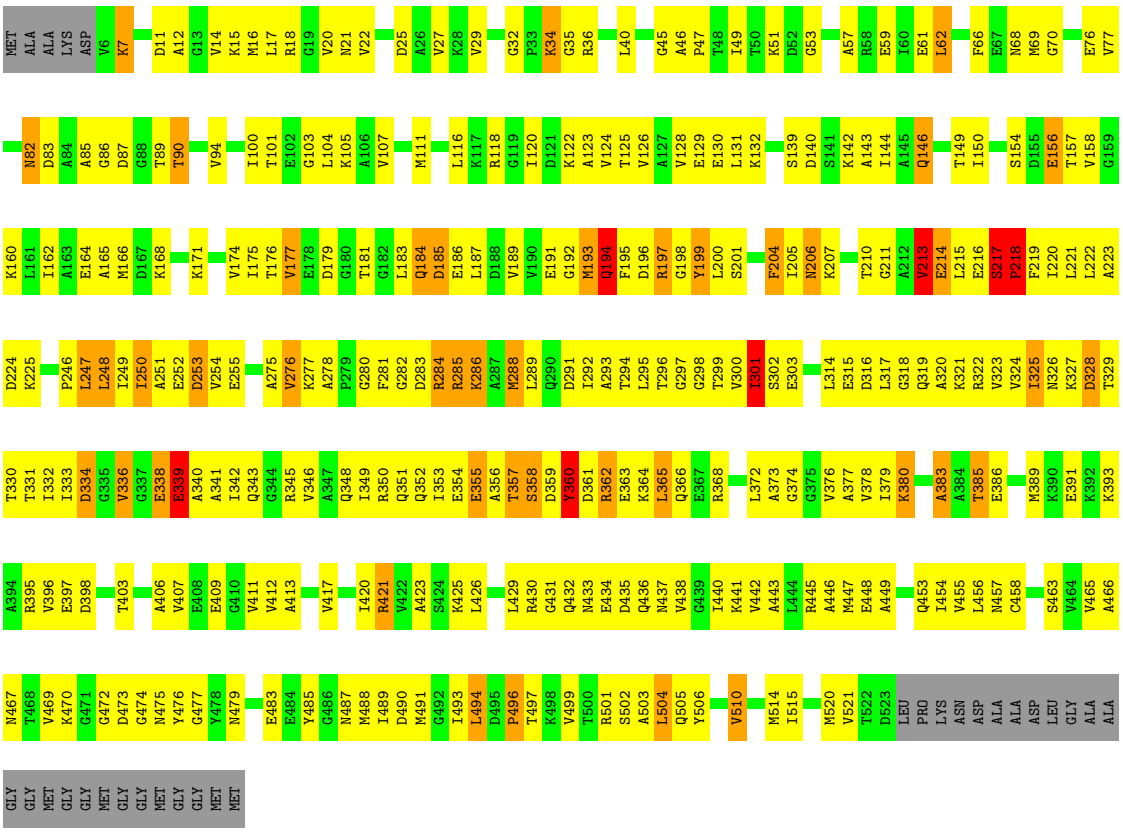




• Molecule 1: GROEL (HSP60 CLASS)

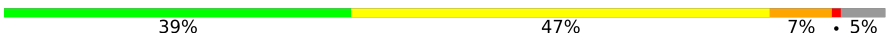
Chain F:

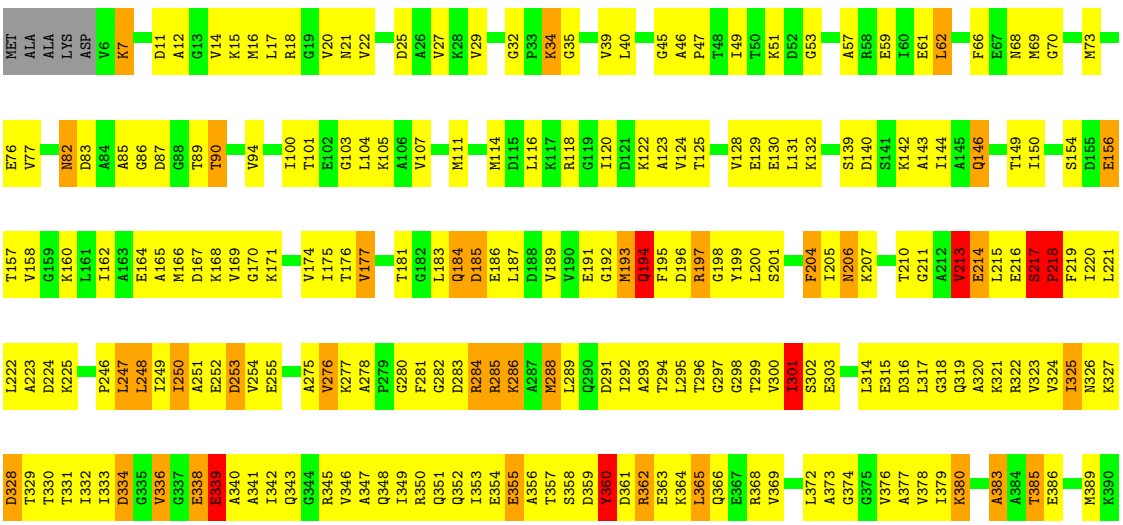




• Molecule 1: GROEL (HSP60 CLASS)

Chain G:






E391 K392 K393 A394 R395 V396 E397 D398 T403 A406 V407 E408 E409 G410 V411 V412 A413 V417 I420 R421 V422 A423 S424 K425 L429 R430 G431 Q432 N433 E434 D435 Q436 M437 V438 G439 I440 K441 V442 A443 L444 R445 A446 M447 E448 A449 Q453 I454 V455 L456 N457 C458 S463 V464
V465 A466 N467 T468 V469 K470 G471 G472 D473 G474 N475 Y476 G477 Y478 N479 E483 E484 Y485 G486 N487 M488 I489 D490 M491 G492 I493 L494 D495 P496 T497 K498 V499 T500 R501 S502 A503 L504 Q505 Y506 V510 M514 I515 E518 C519 M520 V521 T522 D523 LEU PRO LYS ASN ASP ALA ALA ASP
LEU GLY ALA ALA GLY GLY MET GLY GLY MET GLY GLY MET GLY GLY MET MET

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	178.00Å 203.00Å 278.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.80	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-2.80)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.326 , 0.368	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	29274	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

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	0.62	0/3389	0.84	2/4571 (0.0%)
1	B	0.62	0/3389	0.84	2/4571 (0.0%)
1	C	0.62	0/3389	0.84	2/4571 (0.0%)
1	D	0.62	0/3389	0.84	2/4571 (0.0%)
1	E	0.62	0/3389	0.84	2/4571 (0.0%)
1	F	0.62	0/3389	0.84	2/4571 (0.0%)
1	G	0.62	0/3389	0.84	2/4571 (0.0%)
All	All	0.62	0/23723	0.84	14/31997 (0.0%)

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	F	360	TYR	N-CA-C	6.51	128.57	111.00
1	B	360	TYR	N-CA-C	6.50	128.55	111.00
1	G	360	TYR	N-CA-C	6.50	128.55	111.00
1	D	360	TYR	N-CA-C	6.49	128.53	111.00
1	E	360	TYR	N-CA-C	6.49	128.52	111.00

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	3428	754	3441	314	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	3428	754	3441	324	9
1	C	3428	754	3440	334	54
1	D	3428	754	3441	323	0
1	E	3428	754	3441	323	18
1	F	3428	754	3440	324	1
1	G	3428	754	3441	315	55
All	All	23996	5278	24085	2090	75

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 43.

The worst 5 of 2090 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:181:THR:O	1:G:282:GLY:CA	1.65	1.39
1:D:282:GLY:CA	1:E:181:THR:O	1.69	1.37
1:F:282:GLY:CA	1:G:181:THR:O	1.72	1.36
1:A:282:GLY:HA3	1:B:181:THR:O	1.28	1.30
1:A:386:GLU:OE2	1:G:285:ARG:NH2	1.71	1.23

The worst 5 of 75 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:358:SER:OG	1:G:169:VAL:CB[5_455]	0.65	1.55
1:C:358:SER:CB	1:G:169:VAL:CA[5_455]	0.74	1.46
1:C:358:SER:CA	1:G:169:VAL:CA[5_455]	0.89	1.31
1:C:359:ASP:N	1:G:170:GLY:H[5_455]	0.31	1.29
1:C:359:ASP:CB	1:G:167:ASP:CA[5_455]	0.97	1.23

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	449/548 (82%)	358 (80%)	68 (15%)	23 (5%)	2	6
1	B	449/548 (82%)	359 (80%)	67 (15%)	23 (5%)	2	6
1	C	449/548 (82%)	358 (80%)	68 (15%)	23 (5%)	2	6
1	D	449/548 (82%)	358 (80%)	68 (15%)	23 (5%)	2	6
1	E	449/548 (82%)	358 (80%)	68 (15%)	23 (5%)	2	6
1	F	449/548 (82%)	358 (80%)	68 (15%)	23 (5%)	2	6
1	G	449/548 (82%)	358 (80%)	68 (15%)	23 (5%)	2	6
All	All	3143/3836 (82%)	2507 (80%)	475 (15%)	161 (5%)	2	6

5 of 161 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	184	GLN
1	A	185	ASP
1	A	214	GLU
1	A	253	ASP
1	A	301	ILE

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	353/414 (85%)	300 (85%)	53 (15%)	3	9
1	B	353/414 (85%)	300 (85%)	53 (15%)	3	9
1	C	353/414 (85%)	300 (85%)	53 (15%)	3	9
1	D	353/414 (85%)	300 (85%)	53 (15%)	3	9
1	E	353/414 (85%)	300 (85%)	53 (15%)	3	9
1	F	353/414 (85%)	300 (85%)	53 (15%)	3	9
1	G	353/414 (85%)	300 (85%)	53 (15%)	3	9
All	All	2471/2898 (85%)	2100 (85%)	371 (15%)	3	9

5 of 371 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	199	TYR
1	F	206	ASN
1	E	248	LEU
1	E	421	ARG
1	F	329	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 72 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	194	GLN
1	G	453	GLN
1	F	436	GLN
1	G	82	ASN
1	C	82	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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.