wwPDB X-ray Structure Validation Summary Report (i)

Aug 7, 2020 – 04:01 AM BST

PDB ID	:	1DE4
Title	:	HEMOCHROMATOSIS PROTEIN HFE COMPLEXED WITH TRANSFERRIN RECEPTOR
Authors	:	Bennett, M.J.; Lebron, J.A.; Bjorkman, P.J.
Deposited on	:	1999-11-12
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 following versions of software and data (see references (i)) were used in the production of this report:

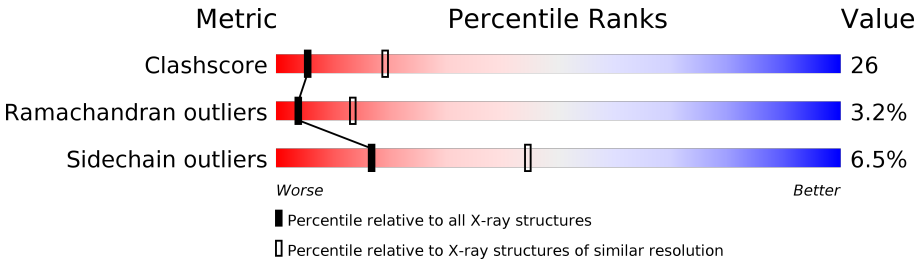
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 on 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	275	50% 41% 8% •
1	D	275	51% 40% 7% •
1	G	275	51% 41% 7% ••
2	B	99	52% 42% 6%
2	E	99	49% 43% 7%
2	H	99	52% 41% 7%
3	C	640	55% 39% 5% •
3	F	640	56% 39% 5% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	I	640	

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 24315 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 HEMOCHROMATOSIS PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	272	Total	C	N	O	S	0	0	0
			2242	1424	390	416	12			
1	D	272	Total	C	N	O	S	0	0	0
			2242	1424	390	416	12			
1	G	272	Total	C	N	O	S	0	0	0
			2242	1424	390	416	12			

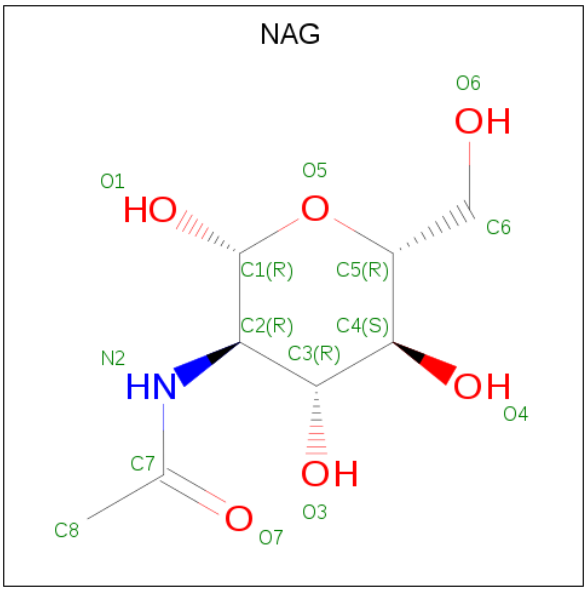
- Molecule 2 is a protein called BETA-2-MICROGLOBULIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			821	522	138	158	3			
2	E	99	Total	C	N	O	S	0	0	0
			821	522	138	158	3			
2	H	99	Total	C	N	O	S	0	0	0
			821	522	138	158	3			

- Molecule 3 is a protein called TRANSFERRIN RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	635	Total	C	N	O	S	0	0	0
			5022	3223	845	940	14			
3	F	635	Total	C	N	O	S	0	0	0
			5022	3223	845	940	14			
3	I	635	Total	C	N	O	S	0	0	0
			5022	3223	845	940	14			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

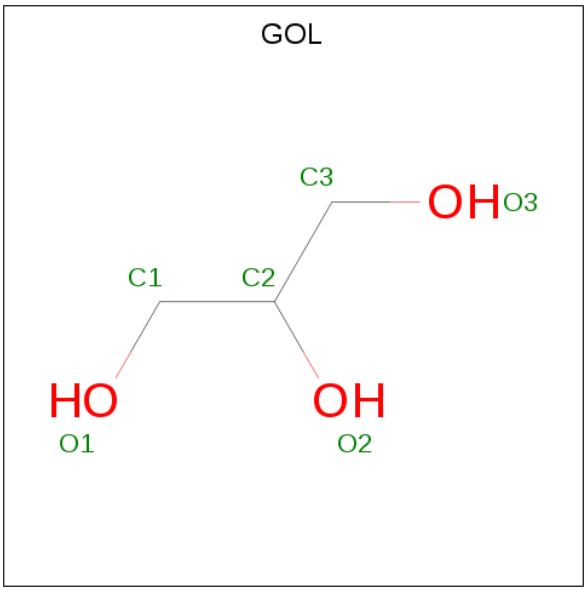


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	1	Total C N O 14 8 1 5	0	0
4	F	1	Total C N O 14 8 1 5	0	0
4	I	1	Total C N O 14 8 1 5	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	I	1	Total Ca 1 1	0	0
5	C	1	Total Ca 1 1	0	0
5	F	1	Total Ca 1 1	0	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	G	1	Total C O 6 3 3	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	C	4	Total O 4 4	0	0
7	F	2	Total O 2 2	0	0
7	I	3	Total O 3 3	0	0

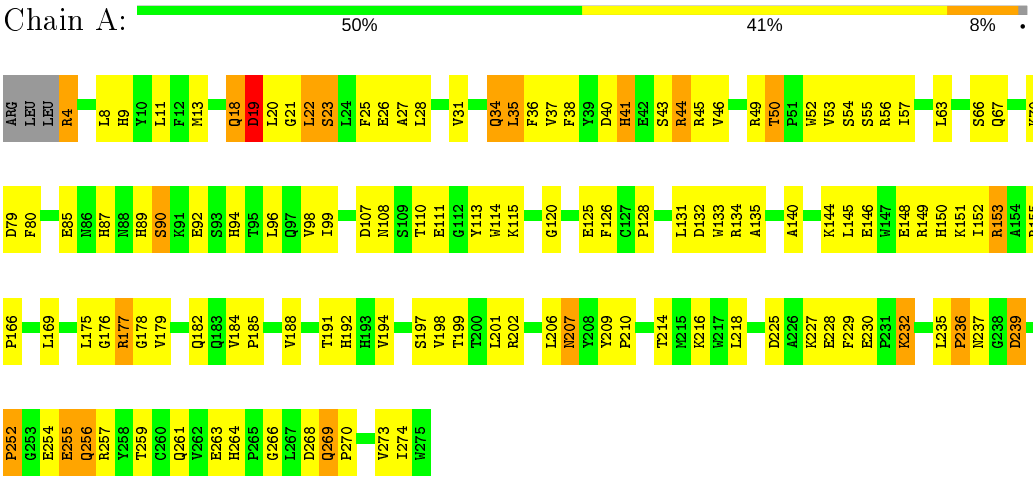
3 Residue-property plots (i)

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. 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.

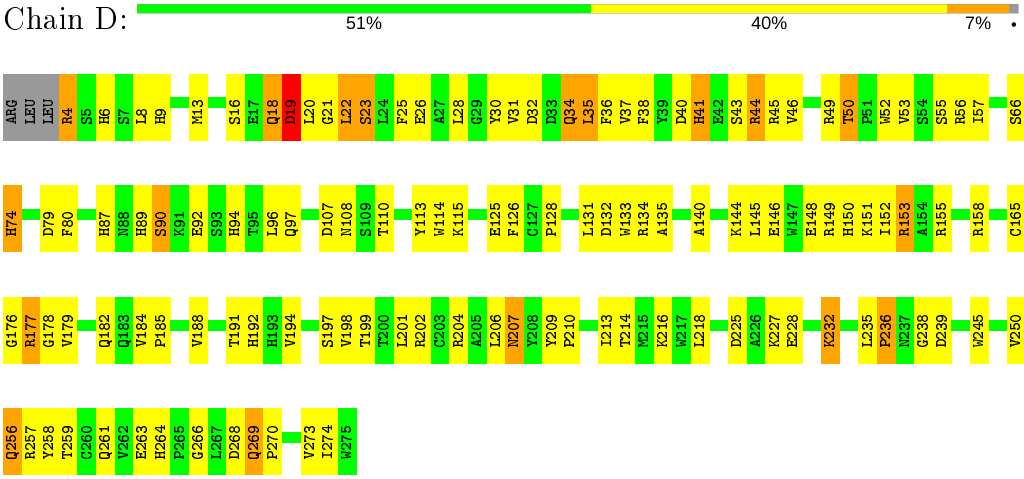
- Molecule 1: HEMOCHROMATOSIS PROTEIN

Chain A:



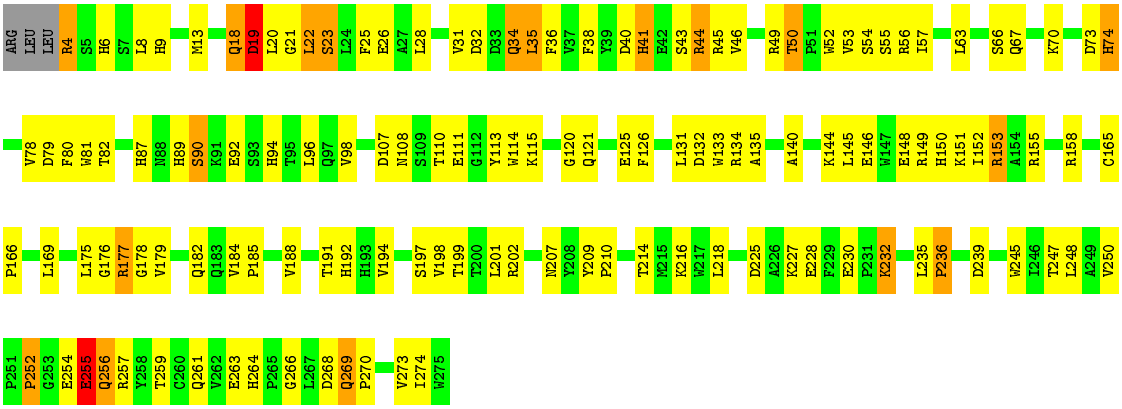
- Molecule 1: HEMOCHROMATOSIS PROTEIN

Chain D:



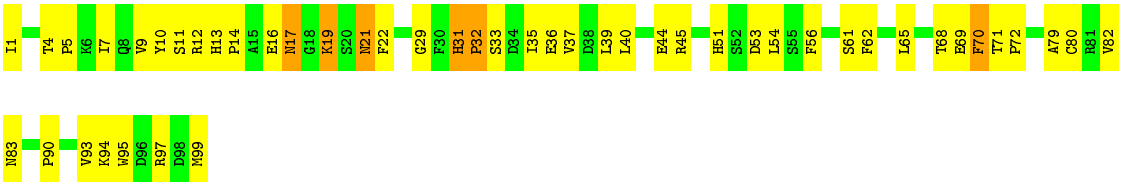
- Molecule 1: HEMOCHROMATOSIS PROTEIN

Chain G:



- Molecule 2: BETA-2-MICROGLOBULIN

Chain B:

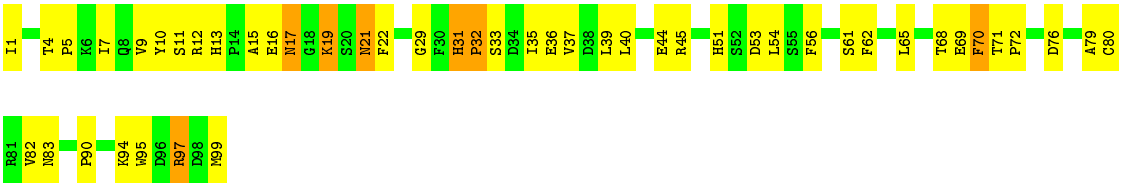


- Molecule 2: BETA-2-MICROGLOBULIN

Chain E:

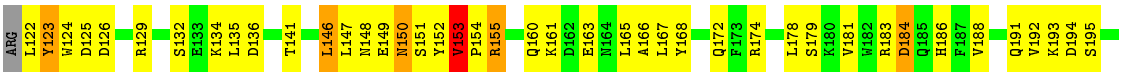
- Molecule 2: BETA-2-MICROGLOBULIN

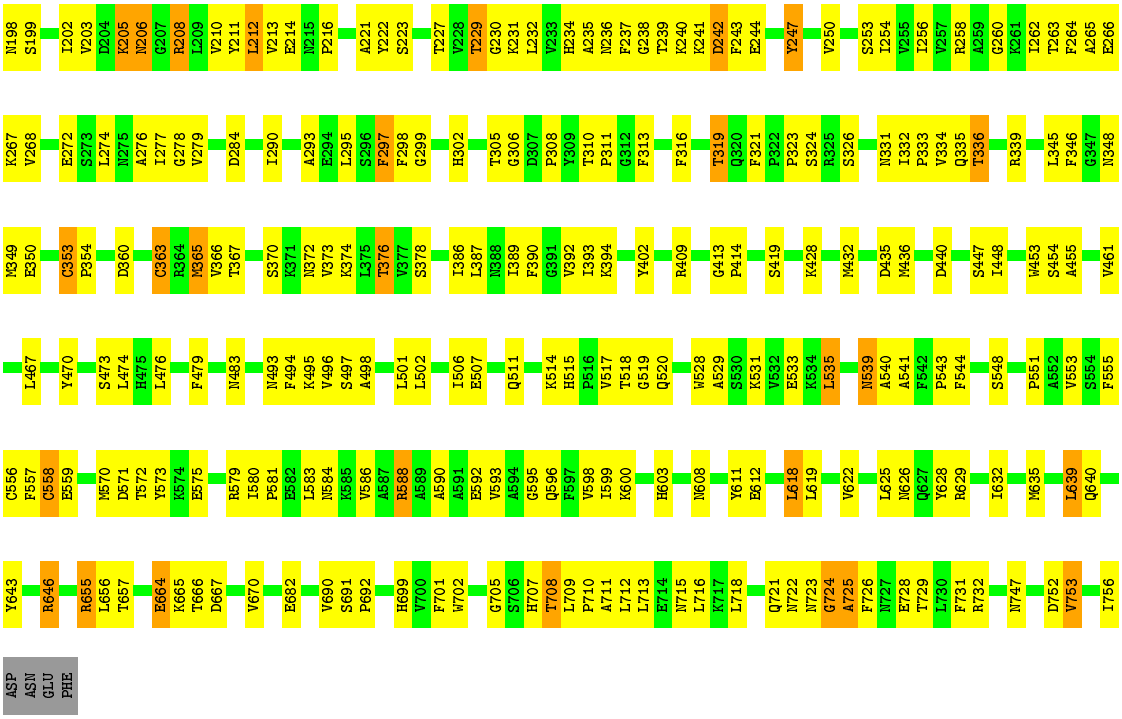
Chain H:



- Molecule 3: TRANSFERRIN RECEPTOR

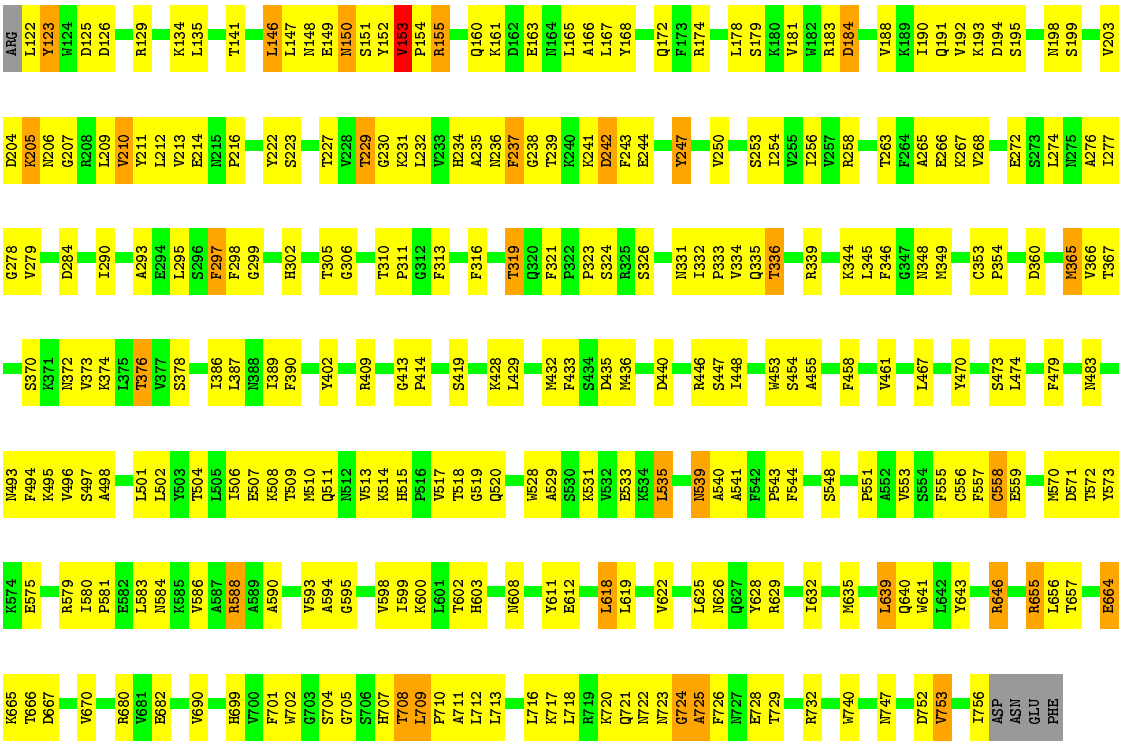
Chain C:





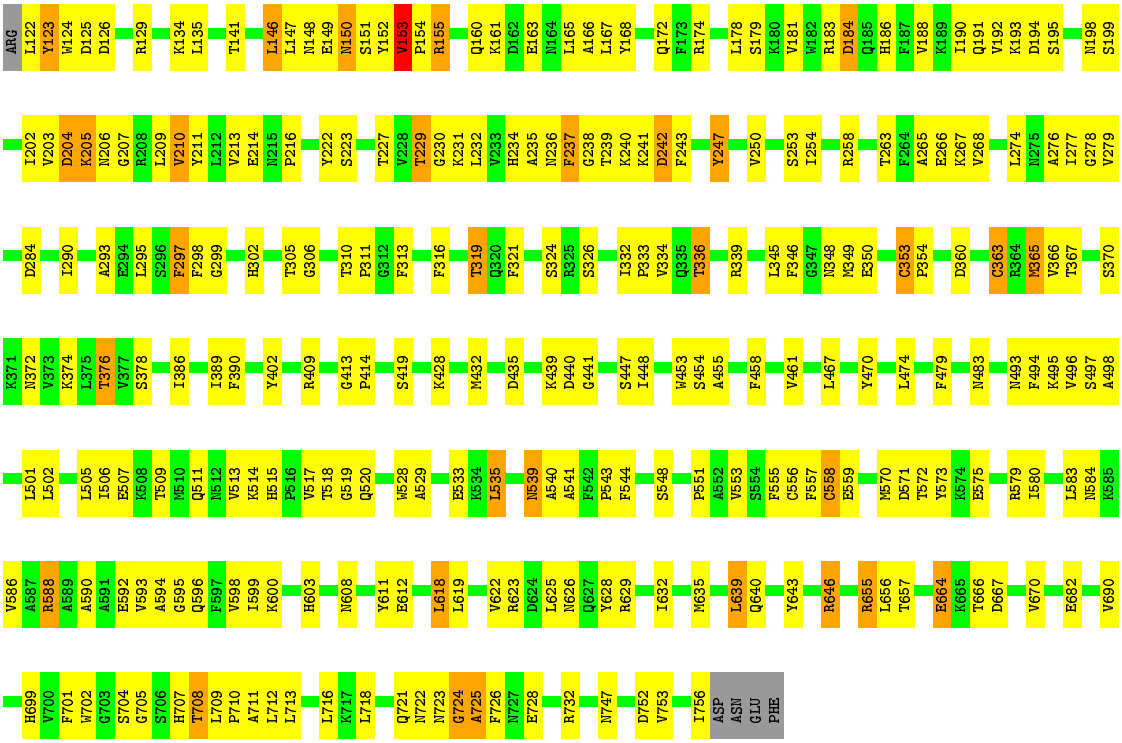
- Molecule 3: TRANSFERRIN RECEPTOR

Chain F:



- Molecule 3: TRANSFERRIN RECEPTOR

Chain I:


ARG L122 Y123 W124 D125 D126 R129 K134 L135 T141 L146 L147 N148 E149 N150 S151 Y152 V153 P154 R155 Q160 K161 D162 E163 N164 L165 A166 L167 Y168 Q172 F173 R174 L178 S179 K180 V181 W182 R183 D184 Q185 H186 F187 V188 K189 I190 Q191 V192 K193 D194 S195 N198 S199
I202 V203 D204 K205 N206 G207 R208 L209 V210 Y211 L212 V213 E214 N215 P216 Y222 S223 T227 V228 T229 G230 K231 L232 V233 H234 A235 N236 F237 G238 T239 K240 K241 D242 F243 Y247 V250 S253 I254 R258 T263 F264 A265 E266 K267 V268 L274 N275 A276 I277 G278 V279
D284 I290 A293 E294 L295 S296 F297 F298 G299 H302 T305 G306 T310 P311 G312 F313 F316 T319 Q320 F321 S324 R325 S326 I332 P333 V334 Q335 T336 R339 L345 F346 G347 N348 M349 E350 C353 P354 D360 C363 R364 M365 V366 T367 S370
K371 N372 V373 K374 L375 T376 V377 S378 I386 I389 F390 Y402 R409 G413 P414 S419 K428 M432 D435 K439 D440 G441 S447 I448 W453 S454 A455 F458 V461 L467 Y470 L474 F479 N483 N493 F494 K495 V496 S497 A498
L501 L502 L505 I506 E507 K508 T509 M510 Q511 M512 V513 K514 H515 P516 V517 T518 G519 Q520 W528 A529 E533 K534 L535 N539 A540 A541 F542 P543 F544 S548 P551 A552 V553 S554 F555 C556 F557 C558 E559 M570 D571 T572 Y573 K574 E575 R579 I580 L583 N584 K585
V586 A587 R588 A589 A590 A591 E592 V593 A594 G595 Q596 F597 V598 I599 K600 H603 N608 Y611 E612 L618 L619 V622 R623 D624 L625 N626 Q627 Y628 R629 I632 M635 L639 Q640 Y643 R646 R655 L656 T657 E664 K665 T666 D667 V670 E682 V690
H699 V700 F701 W702 G703 S704 G705 S706 H707 T708 L709 P710 A711 L712 L713 L716 K717 L718 Q721 N722 N723 G724 A725 F726 N727 E728 R732 N747 D752 V753 I756 ASP ASN GLU PHE

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	110.40Å 144.40Å 327.10Å 90.00° 93.60° 90.00°	Depositor
Resolution (Å)	30.00 – 2.80	Depositor
% Data completeness (in resolution range)	94.3 (30.00-2.80)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 0.5	Depositor
R, R_{free}	0.231 , 0.265	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	24315	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, 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	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.34	0/2310	0.60	0/3143
1	D	0.34	0/2310	0.60	0/3143
1	G	0.35	0/2310	0.60	0/3143
2	B	0.33	0/844	0.56	0/1144
2	E	0.34	0/844	0.56	0/1144
2	H	0.33	0/844	0.56	0/1144
3	C	0.40	0/5142	0.65	2/6973 (0.0%)
3	F	0.40	0/5142	0.65	1/6973 (0.0%)
3	I	0.43	0/5142	0.66	2/6973 (0.0%)
All	All	0.38	0/24888	0.63	5/33780 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	557	PHE	N-CA-C	-6.66	93.01	111.00
3	F	557	PHE	N-CA-C	-6.58	93.23	111.00
3	I	557	PHE	N-CA-C	-6.44	93.61	111.00
3	C	363	CYS	CA-CB-SG	5.07	123.13	114.00
3	I	363	CYS	CA-CB-SG	5.03	123.05	114.00

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	2242	0	2126	136	0
1	D	2242	0	2126	139	0
1	G	2242	0	2126	133	0
2	B	821	0	772	51	0
2	E	821	0	772	55	0
2	H	821	0	772	50	0
3	C	5022	0	4965	253	0
3	F	5022	0	4965	250	0
3	I	5022	0	4965	238	0
4	C	14	0	13	0	0
4	F	14	0	13	0	0
4	I	14	0	13	0	0
5	C	1	0	0	0	0
5	F	1	0	0	0	0
5	I	1	0	0	0	0
6	G	6	0	8	3	0
7	C	4	0	0	0	0
7	F	2	0	0	0	0
7	I	3	0	0	0	0
All	All	24315	0	23636	1231	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 26.

The worst 5 of 1231 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:C:708:THR:HG22	3:C:711:ALA:H	1.18	1.09
3:C:708:THR:HG23	3:C:710:PRO:HD2	1.29	1.08
3:I:708:THR:HG23	3:I:710:PRO:HD2	1.29	1.07
3:F:708:THR:HG23	3:F:710:PRO:HD2	1.31	1.04
3:I:708:THR:HG22	3:I:711:ALA:H	1.18	1.03

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	270/275 (98%)	233 (86%)	24 (9%)	13 (5%)	2	7
1	D	270/275 (98%)	230 (85%)	26 (10%)	14 (5%)	2	6
1	G	270/275 (98%)	230 (85%)	27 (10%)	13 (5%)	2	7
2	B	97/99 (98%)	83 (86%)	9 (9%)	5 (5%)	2	6
2	E	97/99 (98%)	84 (87%)	8 (8%)	5 (5%)	2	6
2	H	97/99 (98%)	84 (87%)	8 (8%)	5 (5%)	2	6
3	C	633/640 (99%)	575 (91%)	45 (7%)	13 (2%)	7	23
3	F	633/640 (99%)	572 (90%)	47 (7%)	14 (2%)	6	22
3	I	633/640 (99%)	573 (90%)	47 (7%)	13 (2%)	7	23
All	All	3000/3042 (99%)	2664 (89%)	241 (8%)	95 (3%)	4	13

5 of 95 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	19	ASP
1	A	22	LEU
1	A	44	ARG
2	B	31	HIS
3	C	153	VAL

5.3.2 Protein sidechains

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	244/249 (98%)	230 (94%)	14 (6%)	20	50
1	D	244/249 (98%)	230 (94%)	14 (6%)	20	50
1	G	244/249 (98%)	230 (94%)	14 (6%)	20	50
2	B	92/94 (98%)	88 (96%)	4 (4%)	29	62
2	E	92/94 (98%)	88 (96%)	4 (4%)	29	62
2	H	92/94 (98%)	88 (96%)	4 (4%)	29	62
3	C	544/549 (99%)	504 (93%)	40 (7%)	13	37
3	F	544/549 (99%)	506 (93%)	38 (7%)	15	40
3	I	544/549 (99%)	505 (93%)	39 (7%)	14	38
All	All	2640/2676 (99%)	2469 (94%)	171 (6%)	17	44

5 of 171 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	152	TYR
3	F	376	THR
3	I	559	GLU
3	F	155	ARG
3	F	229	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 113 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	222	GLN
3	F	302	HIS
3	I	493	ASN
1	D	256	GLN
2	E	83	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)

Of 7 ligands modelled in this entry, 3 are monoatomic - leaving 4 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	$\#\|Z\| > 2$
4	NAG	F	901	3	14,14,15	0.58	0	17,19,21	0.83	1 (5%)
6	GOL	G	309	-	5,5,5	1.75	2 (40%)	5,5,5	0.91	0
4	NAG	C	900	3	14,14,15	0.64	0	17,19,21	0.79	1 (5%)
4	NAG	I	902	3	14,14,15	0.46	0	17,19,21	1.00	2 (11%)

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
4	NAG	F	901	3	-	2/6/23/26	0/1/1/1
6	GOL	G	309	-	-	2/4/4/4	-
4	NAG	C	900	3	-	2/6/23/26	0/1/1/1
4	NAG	I	902	3	-	2/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	G	309	GOL	O2-C2	2.89	1.52	1.43
6	G	309	GOL	C3-C2	2.05	1.60	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	F	901	NAG	C2-N2-C7	-2.63	119.16	122.90
4	I	902	NAG	C2-N2-C7	-2.49	119.36	122.90
4	C	900	NAG	C2-N2-C7	-2.35	119.56	122.90
4	I	902	NAG	C4-C3-C2	-2.04	108.03	111.02

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	G	309	GOL	C1-C2-C3-O3
6	G	309	GOL	O2-C2-C3-O3
4	F	901	NAG	O5-C5-C6-O6
4	I	902	NAG	C4-C5-C6-O6
4	F	901	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	G	309	GOL	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.

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.