wwPDB X-ray Structure Validation Summary Report (i)

Jun 17, 2024 – 12:26 PM EDT

PDB ID	:	3EWE
Title	:	Crystal Structure of the Nup85/Seh1 Complex
Authors	:	Brohawn, S.G.; Leksa, N.C.; Rajashankar, K.R.; Schwartz, T.U.
Deposited on	:	2008-10-14
Resolution	:	3.50 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*
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:

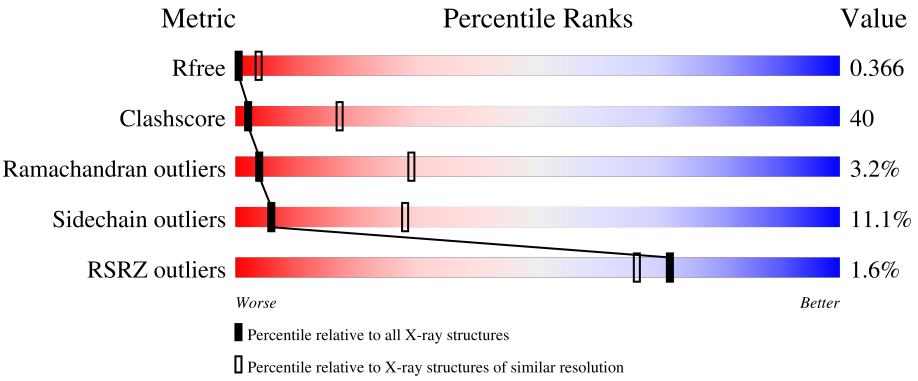
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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))
R_{free}	130704	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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.

Mol	Chain	Length	Quality of chain
1	A	349	% 30% 41% • 27%
1	C	349	3% 31% 39% • 27%
2	B	564	% 31% 34% 5% 30%
2	D	564	% 35% 30% 6% 30%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9689 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 Nucleoporin SEH1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	255	Total	C	N	O	S	0	0	0
			1869	1198	304	358	9			
1	C	255	Total	C	N	O	S	0	0	0
			1872	1201	304	358	9			

- Molecule 2 is a protein called Nucleoporin NUP85.

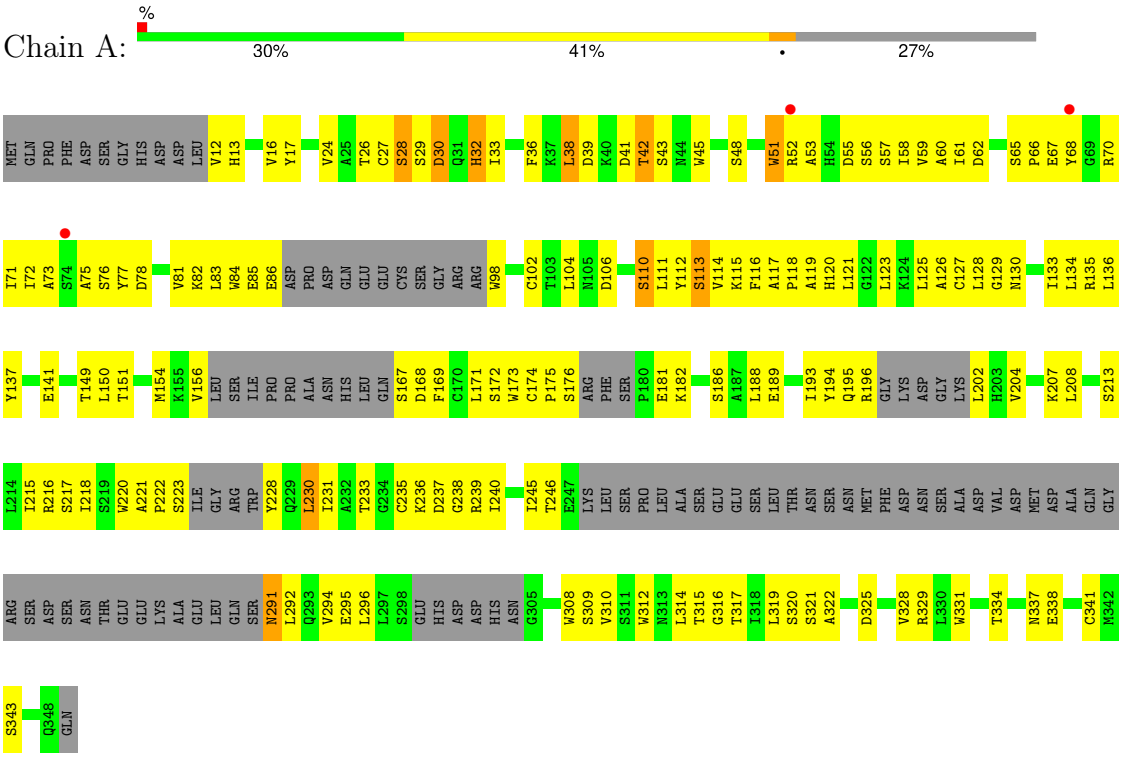
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	395	Total	C	N	O	S	0	0	0
			2974	1941	452	560	21			
2	D	395	Total	C	N	O	S	0	0	0
			2974	1941	452	560	21			

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.

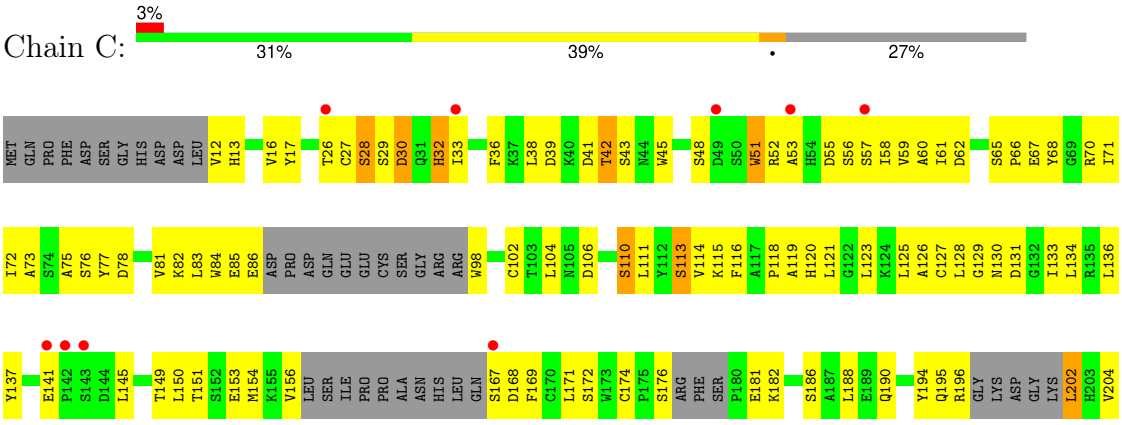
• Molecule 1: Nucleoporin SEH1

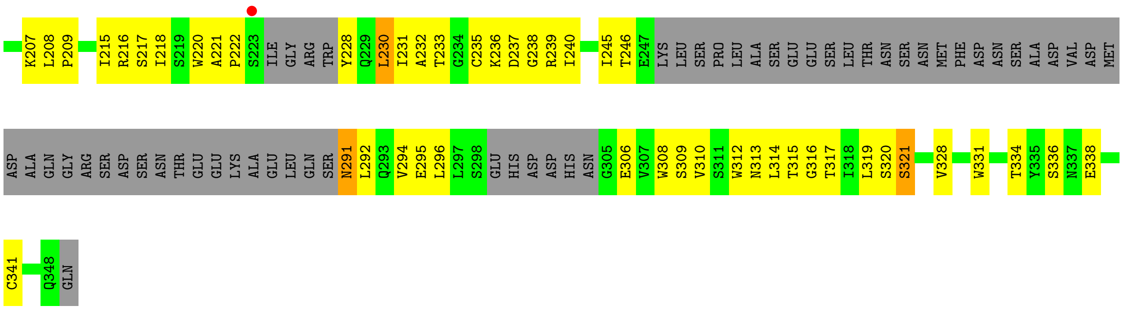
Chain A:



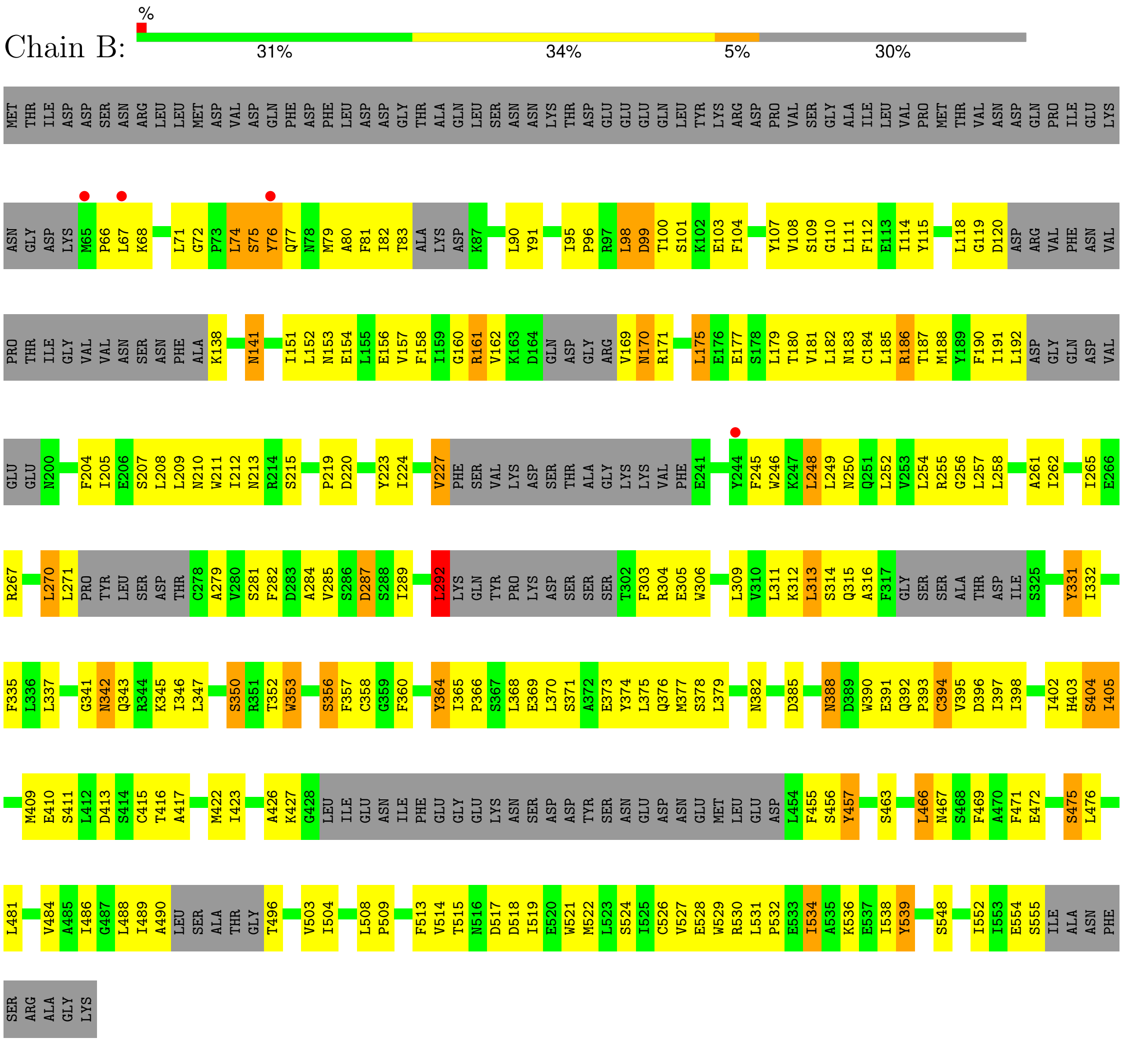
• Molecule 1: Nucleoporin SEH1

Chain C:

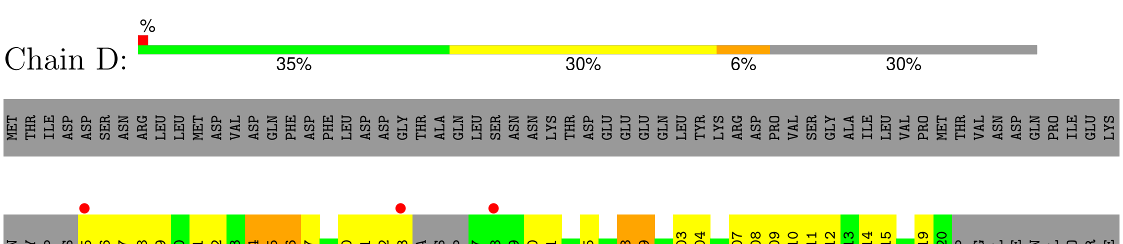


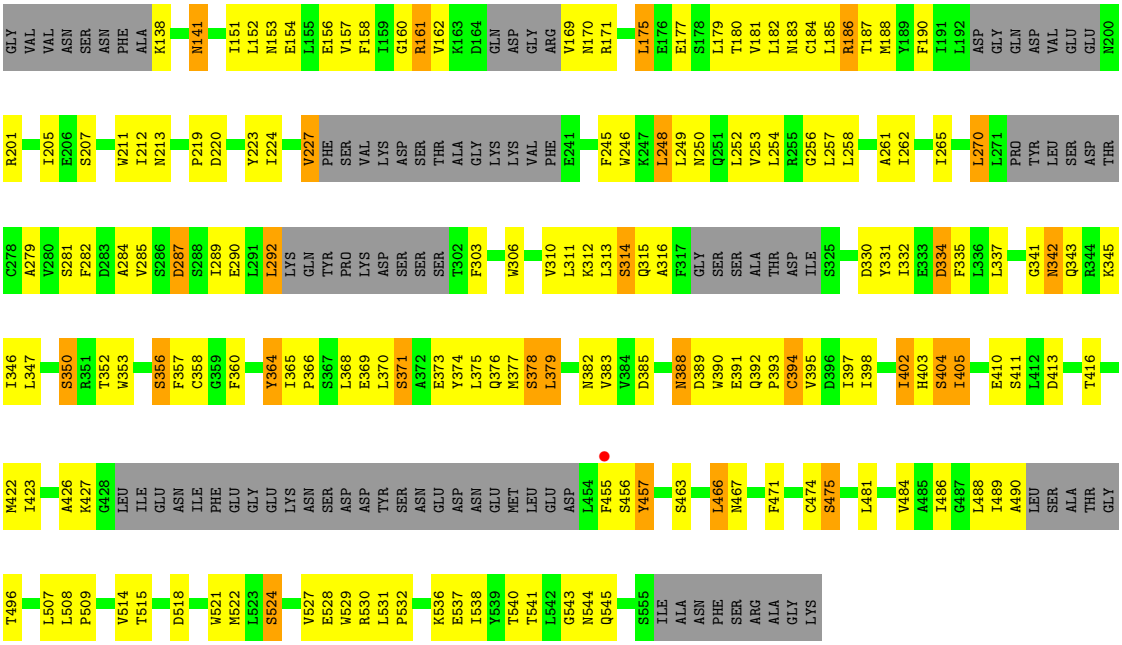


- Molecule 2: Nucleoporin NUP85



- Molecule 2: Nucleoporin NUP85




GLY VAL VAL ASN SER ASN PHE ALA K138 M141 I151 L152 N153 E154 L155 E156 V157 F158 I159 G160 R161 V162 K163 D164 GLN ASP GLY ARG V169 N170 R171 L175 E176 E177 S178 L179 T180 V181 L182 N183 C184 L185 R186 T187 M188 Y189 F190 I191 L192 ASP GLY GLN ASP VAL GLU GLU N200
R201 I205 E206 S207 W211 I212 N213 P219 D220 Y223 I224 V227 PHE SER VAL LYS ASP SER THR ALA GLY LYS LYS VAL PHE E241 F245 W246 K247 L248 L249 N250 Q251 L252 V253 L254 R255 G256 L257 L258 A261 I262 I265 L270 L271 PRO TYR LEU SER ASP THR
C278 A279 V280 S281 F282 D283 A284 V285 S286 D287 S288 I289 E290 L291 L292 LYS GLN TYR PRO LYS ASP SER SER SER T302 F303 W306 V310 L311 K312 L313 S314 Q315 A316 F317 GLY SER SER ALA THR ASP ILE S325 D330 Y331 I332 E333 D334 F335 L336 L337 G341 N342 Q343 R344 K345
I346 L347 S350 R351 T352 W353 S356 F357 C358 G359 F360 Y364 I365 P366 S367 L368 E369 L370 S371 A372 E373 Y374 L375 Q376 M377 S378 L379 N382 V383 V384 D385 N388 D389 W390 E391 Q392 P393 C394 V395 D396 I397 I398 I402 H403 S404 I405 E410 S411 L412 D413 T416
M422 I423 A426 K427 G428 LEU ILE GLU ASN ILE PHE GLU GLY GLU LYS ASN SER ASP ASP TYR SER ASN GLU ASP ASN GLU MET LEU GLU ASP L454 F455 S456 Y457 S463 L466 N467 F471 C474 S475 L481 V484 A485 I486 G487 L488 I489 A490 LEU SER ALA THR GLY
T496 L507 L508 P509 V514 T515 D518 W521 M522 L523 S524 V527 E528 W529 R530 L531 P532 K536 E537 I538 Y539 T540 T541 L542 G543 N544 Q545 S555 ILE ALA ASN PHE SER ARG ALA GLY LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	112.56Å 112.56Å 350.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.75 – 3.50 39.80 – 3.13	Depositor EDS
% Data completeness (in resolution range)	96.4 (29.75-3.50) 95.4 (39.80-3.13)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	1.17 (at 3.12Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.326 , 0.369 0.315 , 0.366	Depositor DCC
R_{free} test set	1955 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	117.7	Xtriage
Anisotropy	0.133	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 102.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	9689	wwPDB-VP
Average B, all atoms (Å^2)	156.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.63% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] 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

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	0.41	0/1909	0.60	0/2610
1	C	0.39	0/1912	0.59	0/2614
2	B	0.41	0/3033	0.59	2/4137 (0.0%)
2	D	0.42	0/3033	0.60	0/4137
All	All	0.41	0/9887	0.59	2/13498 (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	B	313	LEU	CA-CB-CG	5.25	127.38	115.30
2	B	292	LEU	CA-CB-CG	5.11	127.06	115.30

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	1869	0	1709	171	0
1	C	1872	0	1718	169	0
2	B	2974	0	2765	228	0
2	D	2974	0	2765	210	0
All	All	9689	0	8957	750	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 40.

The worst 5 of 750 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:239:ARG:HD3	1:A:296:LEU:HD11	1.17	1.15
1:C:239:ARG:HD3	1:C:296:LEU:HD11	1.25	1.09
2:B:279:ALA:HA	2:B:282:PHE:HD2	1.22	1.03
1:C:118:PRO:HB2	1:C:120:HIS:HD2	1.23	1.03
1:C:154:MET:CB	1:C:202:LEU:HD13	1.89	1.02

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	239/349 (68%)	201 (84%)	33 (14%)	5 (2%)	7 38
1	C	239/349 (68%)	204 (85%)	30 (13%)	5 (2%)	7 38
2	B	373/564 (66%)	293 (79%)	64 (17%)	16 (4%)	2 22
2	D	373/564 (66%)	298 (80%)	62 (17%)	13 (4%)	3 27
All	All	1224/1826 (67%)	996 (81%)	189 (15%)	39 (3%)	4 29

5 of 39 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	130	ASN
2	B	270	LEU
2	B	364	TYR
2	B	403	HIS
1	C	130	ASN

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	188/305 (62%)	172 (92%)	16 (8%)	10	39
1	C	189/305 (62%)	173 (92%)	16 (8%)	10	39
2	B	304/504 (60%)	266 (88%)	38 (12%)	4	23
2	D	304/504 (60%)	265 (87%)	39 (13%)	4	22
All	All	985/1618 (61%)	876 (89%)	109 (11%)	6	28

5 of 109 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	42	THR
2	D	98	LEU
2	D	404	SER
1	C	51	TRP
1	C	246	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	170	ASN
2	D	382	ASN
2	D	544	ASN
2	D	388	ASN
2	D	376	GLN

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

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^{th} 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^2)	Q<0.9
1	A	255/349 (73%)	-0.17	3 (1%) 79 73	113, 154, 229, 288	0
1	C	255/349 (73%)	0.03	10 (3%) 39 35	124, 191, 259, 320	0
2	B	395/564 (70%)	-0.48	4 (1%) 82 77	106, 137, 205, 278	0
2	D	395/564 (70%)	-0.51	4 (1%) 82 77	97, 134, 201, 267	0
All	All	1300/1826 (71%)	-0.33	21 (1%) 72 66	97, 148, 235, 320	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	142	PRO	7.6
2	B	76	TYR	3.9
2	B	65	MET	3.4
1	C	141	GLU	3.3
1	C	49	ASP	3.2

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

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.