



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:51 AM GMT

PDB ID : 3NBN
Title : Crystal structure of a dimer of Notch Transcription Complex trimers on HES1 DNA
Authors : Arnett, K.L.; Blacklow, S.C.
Deposited on : 2010-06-03
Resolution : 3.45 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

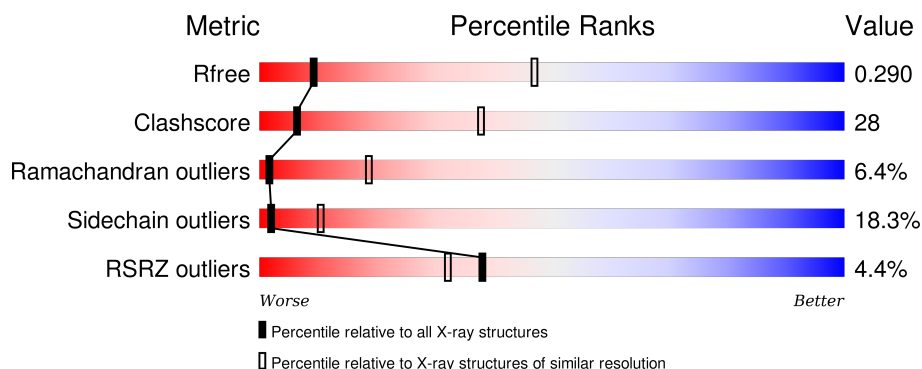
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.45 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.



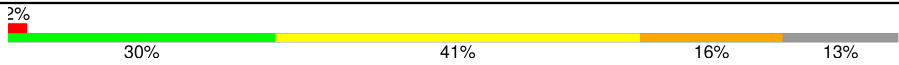


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1000 (3.56-3.36)
Clashscore	102246	1090 (3.56-3.36)
Ramachandran outliers	100387	1057 (3.56-3.36)
Sidechain outliers	100360	1058 (3.56-3.36)
RSRZ outliers	91569	1005 (3.56-3.36)

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. 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 $\leq 5\%$. 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	433	<div> <div>4%</div> <div>49%</div> <div>37%</div> <div>11%</div> <div>..</div> </div>
1	D	433	<div> <div>9%</div> <div>48%</div> <div>39%</div> <div>10%</div> <div>.</div> </div>
2	B	256	<div> <div>30%</div> <div>39%</div> <div>7%</div> <div>23%</div> </div>
2	E	256	<div> <div>%</div> <div>32%</div> <div>36%</div> <div>9%</div> <div>23%</div> </div>
3	C	63	<div> <div>37%</div> <div>37%</div> <div>14%</div> <div>13%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	63	 2% 30% 41% 16% 13%
4	X	37	 8% 49% 43%
5	Y	37	 3% 22% 32% 46%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12198 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 Recombining binding protein suppressor of hairless.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	423	Total	C	N	O	S	0	0	0
			3358	2127	576	630	25			
1	D	423	Total	C	N	O	S	0	0	0
			3358	2127	576	630	25			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	8	MET	-	EXPRESSION TAG	UNP Q06330
A	435	HIS	-	EXPRESSION TAG	UNP Q06330
A	436	HIS	-	EXPRESSION TAG	UNP Q06330
A	437	HIS	-	EXPRESSION TAG	UNP Q06330
A	438	HIS	-	EXPRESSION TAG	UNP Q06330
A	439	HIS	-	EXPRESSION TAG	UNP Q06330
A	440	HIS	-	EXPRESSION TAG	UNP Q06330
D	8	MET	-	EXPRESSION TAG	UNP Q06330
D	435	HIS	-	EXPRESSION TAG	UNP Q06330
D	436	HIS	-	EXPRESSION TAG	UNP Q06330
D	437	HIS	-	EXPRESSION TAG	UNP Q06330
D	438	HIS	-	EXPRESSION TAG	UNP Q06330
D	439	HIS	-	EXPRESSION TAG	UNP Q06330
D	440	HIS	-	EXPRESSION TAG	UNP Q06330

- Molecule 2 is a protein called Neurogenic locus notch homolog protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	196	Total	C	N	O	S	0	0	0
			1518	928	287	297	6			
2	E	196	Total	C	N	O	S	0	0	0
			1518	928	287	297	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1872	GLY	-	EXPRESSION TAG	UNP P46531
E	1872	GLY	-	EXPRESSION TAG	UNP P46531

- Molecule 3 is a protein called Mastermind-like protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	55	Total	C	N	O	S	0	0	0
			467	282	103	78	4			
3	F	55	Total	C	N	O	S	0	0	0
			467	282	103	78	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	12	GLY	-	EXPRESSION TAG	UNP Q92585
F	12	GLY	-	EXPRESSION TAG	UNP Q92585

- Molecule 4 is a DNA chain called DNA, HES1 promoter.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	X	37	Total	C	N	O	P	0	0	0
			770	366	150	217	37			

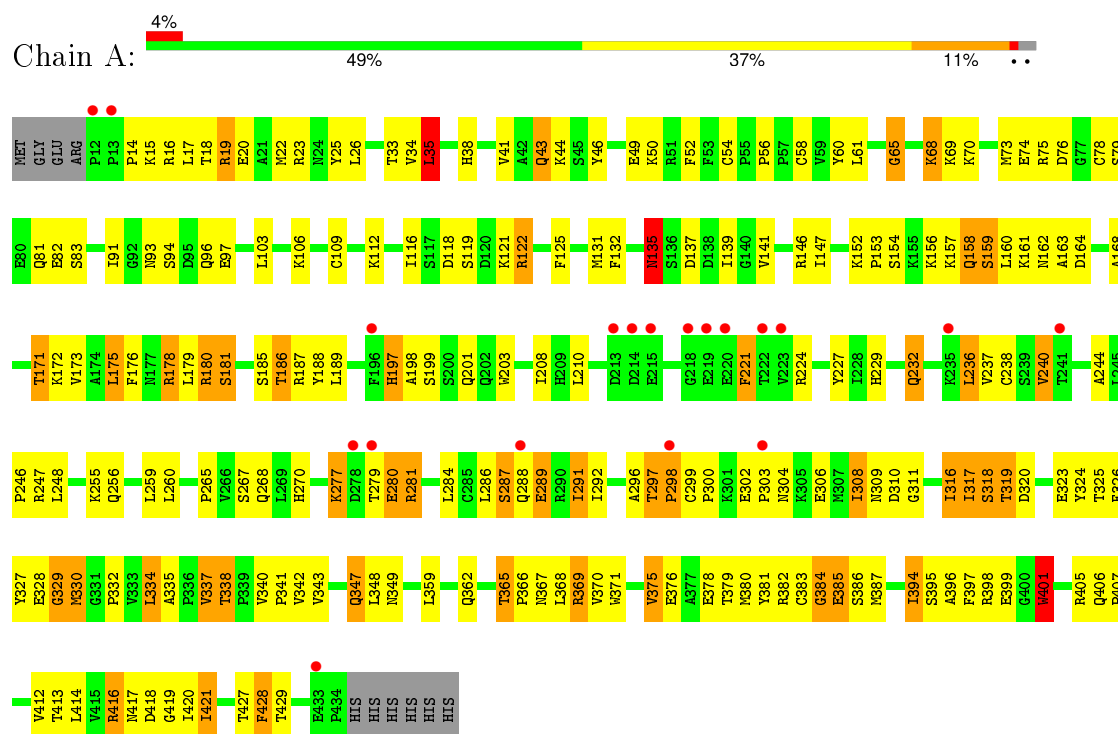
- Molecule 5 is a DNA chain called DNA, HES1 promoter.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	Y	37	Total	C	N	O	P	0	0	0
			742	358	125	223	36			

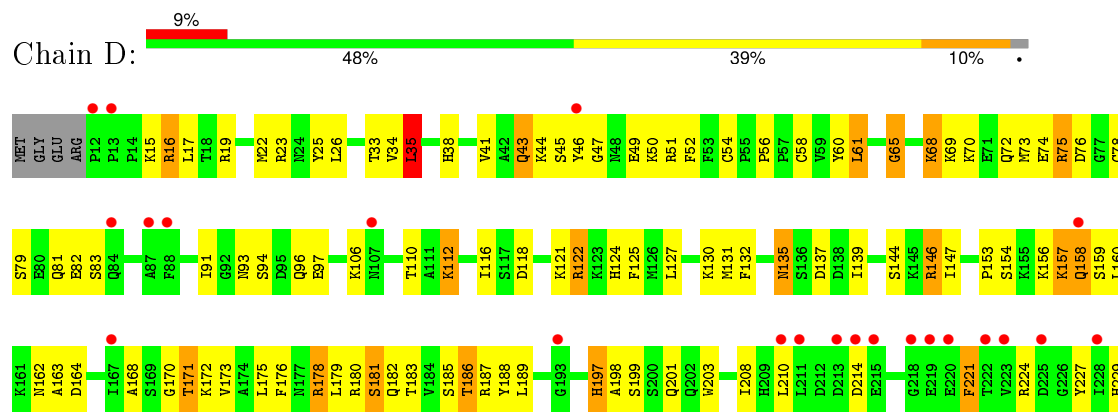
3 Residue-property plots

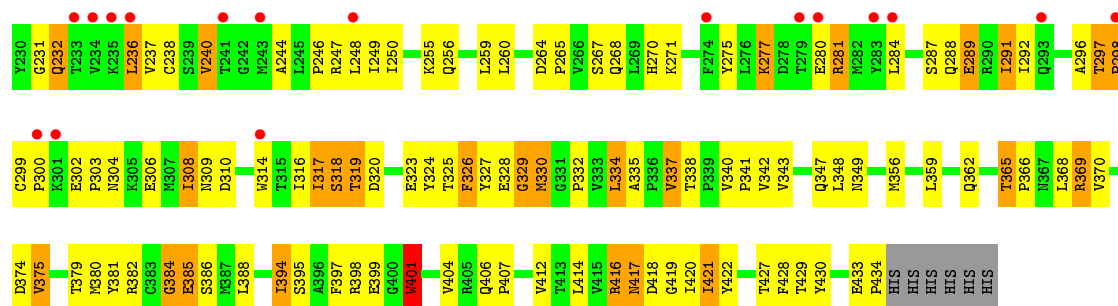
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: Recombining binding protein suppressor of hairless

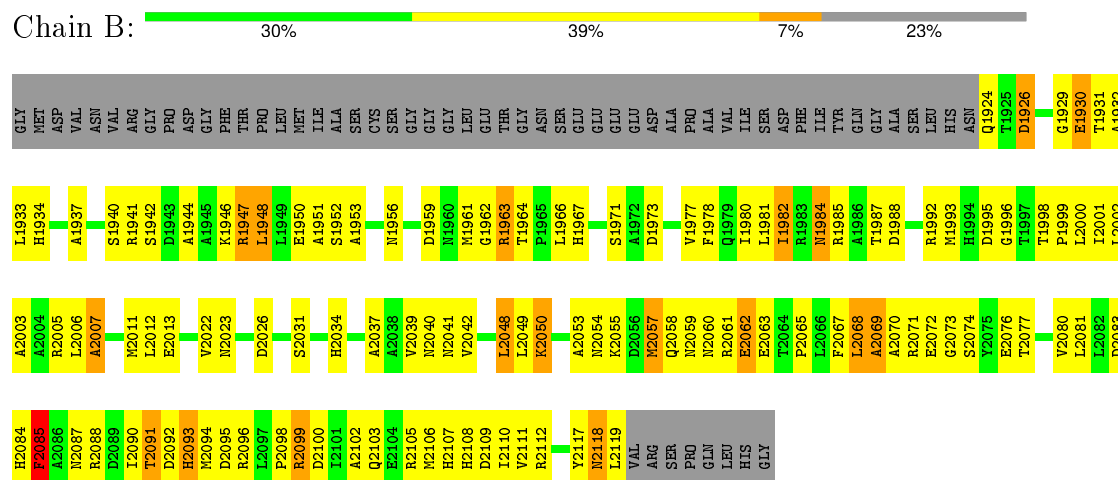


- Molecule 1: Recombining binding protein suppressor of hairless

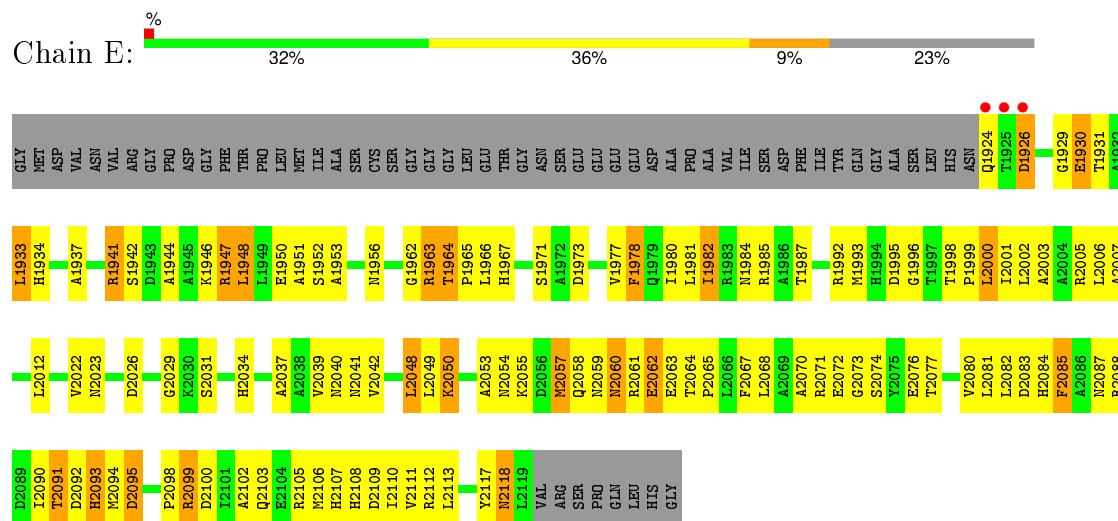




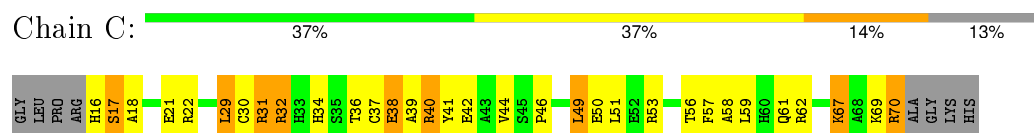
• Molecule 2: Neurogenic locus notch homolog protein 1



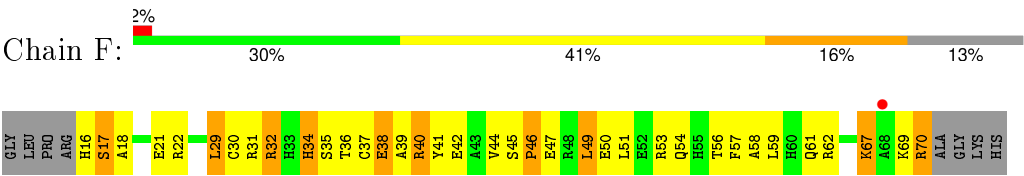
• Molecule 2: Neurogenic locus notch homolog protein 1



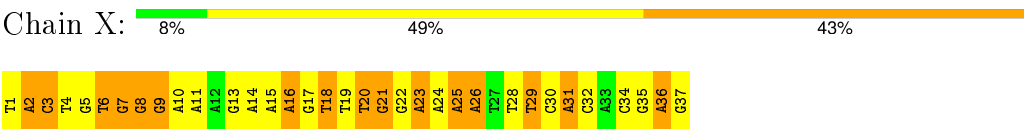
• Molecule 3: Mastermind-like protein 1



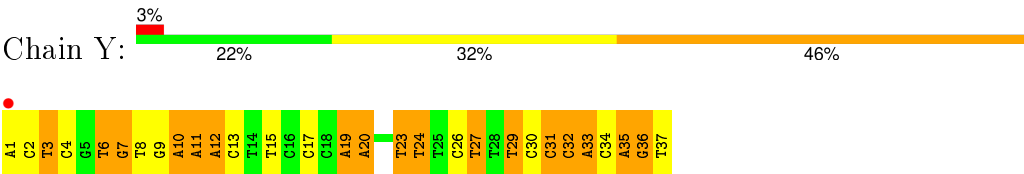
● Molecule 3: Mastermind-like protein 1



● Molecule 4: DNA, HES1 promoter



● Molecule 5: DNA, HES1 promoter



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	295.11Å 108.06Å 87.24Å 90.00° 102.52° 90.00°	Depositor
Resolution (Å)	45.02 – 3.45 45.02 – 3.45	Depositor EDS
% Data completeness (in resolution range)	99.5 (45.02-3.45) 99.5 (45.02-3.45)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.37 (at 3.48Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.254 , 0.298 0.242 , 0.290	Depositor DCC
R_{free} test set	1801 reflections (5.37%)	DCC
Wilson B-factor (Å ²)	108.1	Xtriage
Anisotropy	0.251	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 55.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 35361 reflections (0.003%)	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	12198	wwPDB-VP
Average B, all atoms (Å ²)	129.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.25% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.61	0/3433	0.76	0/4640
1	D	0.65	0/3433	0.75	0/4640
2	B	0.72	0/1538	0.85	1/2085 (0.0%)
2	E	0.76	0/1538	0.88	0/2085
3	C	0.62	0/474	0.77	1/631 (0.2%)
3	F	0.62	0/474	0.75	1/631 (0.2%)
4	X	0.98	2/867 (0.2%)	2.07	47/1338 (3.5%)
5	Y	1.06	1/828 (0.1%)	2.04	39/1273 (3.1%)
All	All	0.72	3/12585 (0.0%)	1.08	89/17323 (0.5%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	D	0	1
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	X	34	DC	C1'-N1	5.43	1.56	1.49
4	X	28	DT	C1'-N1	5.27	1.56	1.49
5	Y	4	DC	C1'-N1	5.16	1.55	1.49

The worst 5 of 89 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	Y	23	DT	O4'-C1'-N1	12.40	116.68	108.00
4	X	32	DC	O4'-C1'-N1	12.16	116.51	108.00
4	X	36	DA	C1'-O4'-C4'	-12.13	97.97	110.10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
4	X	18	DT	O4'-C1'-N1	11.73	116.21	108.00
4	X	26	DA	O4'-C1'-N9	11.27	115.89	108.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	54	CYS	Peptide
1	D	54	CYS	Peptide

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	3358	0	3335	173	0
1	D	3358	0	3335	163	0
2	B	1518	0	1486	121	0
2	E	1518	0	1486	117	0
3	C	467	0	471	36	0
3	F	467	0	471	32	0
4	X	770	0	417	42	0
5	Y	742	0	421	33	0
All	All	12198	0	11422	650	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 28.

The worst 5 of 650 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:238:CYS:SG	1:A:240:VAL:HG13	1.83	1.18
2:B:2040:ASN:HB2	2:B:2074:SER:HB2	1.27	1.13
1:A:122:ARG:HH21	1:A:125:PHE:HB2	0.95	1.12
1:D:238:CYS:SG	1:D:240:VAL:HG13	1.90	1.10
5:Y:35:DA:H2"	5:Y:36:DG:C8	1.86	1.09

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	421/433 (97%)	320 (76%)	75 (18%)	26 (6%)	2	19
1	D	421/433 (97%)	315 (75%)	81 (19%)	25 (6%)	2	21
2	B	194/256 (76%)	146 (75%)	36 (19%)	12 (6%)	2	19
2	E	194/256 (76%)	146 (75%)	36 (19%)	12 (6%)	2	19
3	C	53/63 (84%)	36 (68%)	12 (23%)	5 (9%)	1	10
3	F	53/63 (84%)	35 (66%)	12 (23%)	6 (11%)	0	7
All	All	1336/1504 (89%)	998 (75%)	252 (19%)	86 (6%)	2	19

5 of 86 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	146	ARG
1	A	181	SER
1	A	329	GLY
2	B	1926	ASP
2	B	2099	ARG

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	373/382 (98%)	303 (81%)	70 (19%)	2	9

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	373/382 (98%)	303 (81%)	70 (19%)	2	9
2	B	157/204 (77%)	134 (85%)	23 (15%)	4	21
2	E	157/204 (77%)	135 (86%)	22 (14%)	4	23
3	C	49/54 (91%)	37 (76%)	12 (24%)	1	4
3	F	49/54 (91%)	34 (69%)	15 (31%)	0	2
All	All	1158/1280 (90%)	946 (82%)	212 (18%)	2	10

5 of 212 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	32	ARG
1	D	121	LYS
2	E	2093	HIS
3	C	42	GLU
1	D	26	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	34	HIS
1	D	135	ASN
2	E	2107	HIS
1	D	38	HIS
1	D	229	HIS

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 carbohydrates 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, 95th 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	423/433 (97%)	0.38	19 (4%) 37 31	81, 135, 182, 193	1 (0%)
1	D	423/433 (97%)	0.44	39 (9%) 11 11	81, 135, 182, 193	1 (0%)
2	B	196/256 (76%)	-0.22	0 100 100	69, 89, 129, 157	0
2	E	196/256 (76%)	-0.19	3 (1%) 76 70	69, 89, 129, 157	0
3	C	55/63 (87%)	-0.23	0 100 100	77, 122, 167, 172	0
3	F	55/63 (87%)	-0.06	1 (1%) 71 64	77, 122, 167, 172	0
4	X	37/37 (100%)	-0.50	0 100 100	113, 169, 226, 233	0
5	Y	37/37 (100%)	-0.46	1 (2%) 58 51	123, 171, 220, 225	0
All	All	1422/1578 (90%)	0.15	63 (4%) 38 32	69, 121, 182, 233	2 (0%)

The worst 5 of 63 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	223	VAL	6.0
1	D	235	LYS	5.3
2	E	1925	THR	4.4
1	D	218	GLY	4.3
1	D	234	VAL	4.1

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 carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.