WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Feb 26, 2014 – 04:36 PM GMT

PDB ID : 4H5J
Title : Crystal Structure of the Guanine Nucleotide Exchange Factor Sec12 (P64 form)
Authors : McMahon, C.; Jeffrey, P.D.; Hughson, F.M.
Deposited on : 2012-09-18
Resolution : 2.60 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
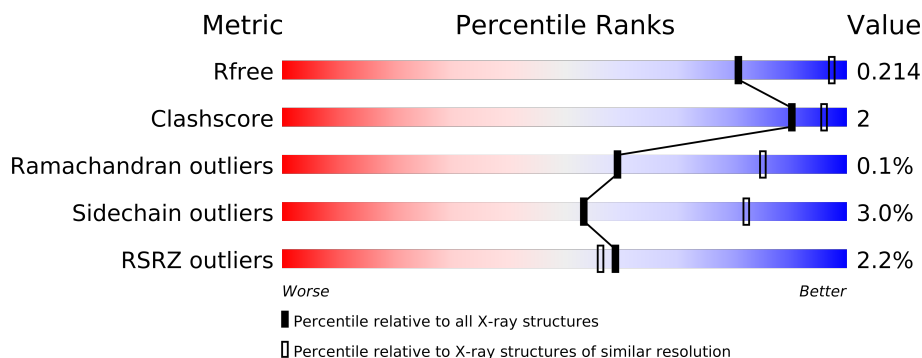A user guide is available at http://wwpdb.org/ValidationPDFNotes.html

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

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance (i)

The reported resolution of this entry is 2.60 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	Percentile Ranks	Value
Rfree		0.214
Clashscore		2
Ramachandran outliers		0.1%
Sidechain outliers		3.0%
RSRZ outliers		2.2%

Worse ————————————————— Better

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1718 (2.60-2.60)
Clashscore	79885	2154 (2.60-2.60)
Ramachandran outliers	78287	2113 (2.60-2.60)
Sidechain outliers	78261	2113 (2.60-2.60)
RSRZ outliers	66119	1718 (2.60-2.60)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	365	
1	B	365	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 5353 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Guanine nucleotide-exchange factor SEC12.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	343	Total	C	N	O	S	Se		0	0	0
			2614	1650	437	521	2	4				
1	B	347	Total	C	N	O	S	Se		0	0	0
			2647	1671	441	528	2	5				

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MSE	-	EXPRESSION TAG	UNP P11655
A	-1	ALA	-	EXPRESSION TAG	UNP P11655
A	0	SER	-	EXPRESSION TAG	UNP P11655
A	355	LEU	-	EXPRESSION TAG	UNP P11655
A	356	GLU	-	EXPRESSION TAG	UNP P11655
A	357	HIS	-	EXPRESSION TAG	UNP P11655
A	358	HIS	-	EXPRESSION TAG	UNP P11655
A	359	HIS	-	EXPRESSION TAG	UNP P11655
A	360	HIS	-	EXPRESSION TAG	UNP P11655
A	361	HIS	-	EXPRESSION TAG	UNP P11655
A	362	HIS	-	EXPRESSION TAG	UNP P11655
B	-2	MSE	-	EXPRESSION TAG	UNP P11655
B	-1	ALA	-	EXPRESSION TAG	UNP P11655
B	0	SER	-	EXPRESSION TAG	UNP P11655
B	355	LEU	-	EXPRESSION TAG	UNP P11655
B	356	GLU	-	EXPRESSION TAG	UNP P11655
B	357	HIS	-	EXPRESSION TAG	UNP P11655
B	358	HIS	-	EXPRESSION TAG	UNP P11655
B	359	HIS	-	EXPRESSION TAG	UNP P11655
B	360	HIS	-	EXPRESSION TAG	UNP P11655
B	361	HIS	-	EXPRESSION TAG	UNP P11655
B	362	HIS	-	EXPRESSION TAG	UNP P11655

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	3	Total 3	K 3	0	0
2	A	3	Total 3	K 3	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	25	Total 25	O 25	0	0
3	B	61	Total 61	O 61	0	0

3 Residue-property plots (i)

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.
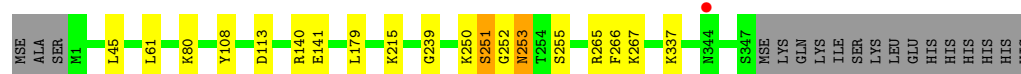
• Molecule 1: Guanine nucleotide-exchange factor SEC12

Chain A:
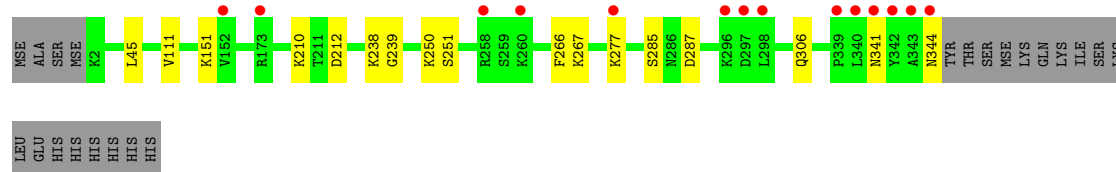
• Molecule 1: Guanine nucleotide-exchange factor SEC12

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, α, β, γ	189.47Å 189.47Å 53.42Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.37 – 2.60 47.37 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.37-2.60) 99.9 (47.37-2.60)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.26 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1144)	Depositor
R, R_{free}	0.179 , 0.216 0.180 , 0.214	Depositor DCC
R_{free} test set	1727 reflections (5.06%)	DCC
Wilson B-factor (Å^2)	28.5	Xtriage
Anisotropy	0.112	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 24.9	EDS
Estimated twinning fraction	0.028 for h,-h-k,-l	Xtriage
L-test for twinning	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Outliers	0 of 34141 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5353	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.13% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: K

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	0.35	0/2647	0.59	0/3572
1	B	0.41	0/2681	0.62	0/3618
All	All	0.38	0/5328	0.60	0/7190

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2614	0	0	5	0
1	B	2647	0	0	6	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
3	A	25	0	0	1	0
3	B	61	0	0	2	0
All	All	5353	0	0	11	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

All (11) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:210:LYS:NZ	1:A:212:ASP:OD1	2.29	0.66
1:B:239:GLY:O	1:B:267:LYS:NZ	2.36	0.58
1:A:306:GLN:N	3:A:507:HOH:O	2.37	0.57
1:B:80:LYS:NZ	1:B:141:GLU:OE1	2.39	0.56
1:A:341:ASN:OD1	1:A:344:ASN:ND2	2.43	0.52
1:A:285:SER:OG	1:A:287:ASP:OD1	2.28	0.52
1:B:337:LYS:NZ	3:B:543:HOH:O	2.44	0.49
1:B:215:LYS:NZ	3:B:511:HOH:O	2.46	0.49
1:A:239:GLY:O	1:A:267:LYS:NZ	2.49	0.45
1:B:252:GLY:O	1:B:253:ASN:ND2	2.50	0.44
1:B:108:TYR:OH	1:B:113:ASP:OD2	2.38	0.42

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	341/365 (93%)	331 (97%)	10 (3%)	0	100	100
1	B	345/365 (94%)	335 (97%)	9 (3%)	1 (0%)	50	77
All	All	686/730 (94%)	666 (97%)	19 (3%)	1 (0%)	59	85

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	251	SER

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	294/308 (96%)	286 (97%)	8 (3%)	57	85
1	B	298/308 (97%)	288 (97%)	10 (3%)	49	78
All	All	592/616 (96%)	574 (97%)	18 (3%)	53	82

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	45	LEU
1	A	111	VAL
1	A	151	LYS
1	A	238	LYS
1	A	250	LYS
1	A	251	SER
1	A	266	PHE
1	A	277	LYS
1	B	45	LEU
1	B	61	LEU
1	B	140	ARG
1	B	179	LEU
1	B	250	LYS
1	B	251	SER
1	B	253	ASN
1	B	255	SER
1	B	265	ARG
1	B	266	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3　RNA　ⓘ

There are no RNA chains 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 ⓘ

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 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	343/365 (93%)	-0.23	14 (4%)	35	32	21, 38, 68, 89	0
1	B	347/365 (95%)	-0.64	1 (0%)	91	92	12, 23, 46, 84	0
All	All	690/730 (94%)	-0.43	15 (2%)	59	56	12, 31, 62, 89	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	344	ASN	3.2
1	A	342	TYR	3.1
1	A	260	LYS	2.7
1	A	298	LEU	2.7
1	A	343	ALA	2.5
1	A	297	ASP	2.5
1	A	258	ARG	2.4
1	B	344	ASN	2.3
1	A	173	ARG	2.3
1	A	341	ASN	2.2
1	A	340	LEU	2.2
1	A	296	LYS	2.1
1	A	152	VAL	2.1
1	A	339	PRO	2.1
1	A	277	LYS	2.0

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 carbohydrates in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å²)	Q<0.9
2	K	A	402	1/1	0.15	1.41	41,41,41,41	0
2	K	A	401	1/1	0.11	0.18	30,30,30,30	0
2	K	A	403	1/1	0.12	-0.20	41,41,41,41	0
2	K	B	402	1/1	0.11	-0.27	24,24,24,24	0
2	K	B	401	1/1	0.10	-0.86	21,21,21,21	0
2	K	B	403	1/1	0.08	-2.69	38,38,38,38	0

6.5 Other polymers ⓘ

There are no such residues in this entry.