



wwPDB X-ray Structure Validation Summary Report

Mar 31, 2014 – 03:23 PM BST

PDB ID : 3WS3
Title : Crystal Structure of H-2D in complex with an insulin derived peptide
Authors : Kumar, P.R.; Mukherjee, G.; Samanta, D.; DiLorenzo, T.P.; Almo, S.C.; Immune Function Network; New York Structural Genomics Research Consortium (NYSGRG)
Deposited on : 2014-02-28
Resolution : 2.33 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

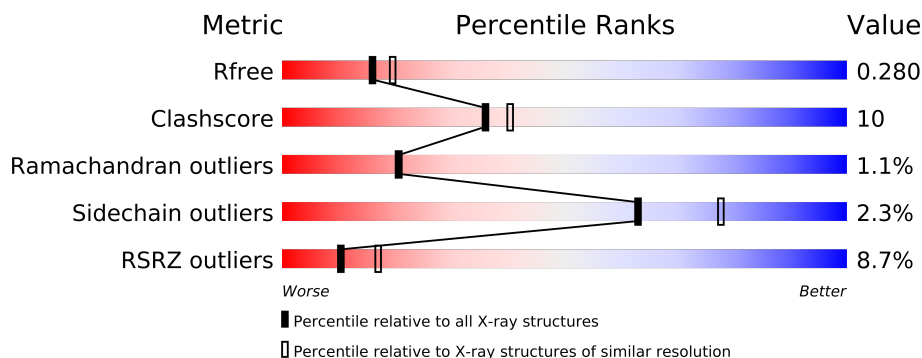
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable23004
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) : stable23004

1 Overall quality at a glance

The reported resolution of this entry is 2.33 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}	66092	4049 (2.38-2.30)
Clashscore	79885	1094 (2.36-2.32)
Ramachandran outliers	78287	1080 (2.36-2.32)
Sidechain outliers	78261	1081 (2.36-2.32)
RSRZ outliers	66119	4050 (2.38-2.30)

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	273	
1	C	273	
2	B	100	
2	D	100	
3	E	9	
3	F	9	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6355 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 H-2 class I histocompatibility antigen, D-B alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	266	Total	C	N	O	S	0	1	0
			2205	1394	388	414	9			
1	C	271	Total	C	N	O	S	0	0	0
			2233	1410	395	419	9			

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	100	Total	C	N	O	S	0	0	0
			829	529	139	153	8			
2	D	100	Total	C	N	O	S	0	0	0
			829	529	139	153	8			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	20	MET	-	EXPRESSION TAG	UNP P01887
B	105	ASP	ALA	SEE REMARK 999	UNP P01887
D	20	MET	-	EXPRESSION TAG	UNP P01887
D	105	ASP	ALA	SEE REMARK 999	UNP P01887

- Molecule 3 is a protein called Insulin derived 9-mer peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	9	Total	C	N	O	S	0	0	0
			77	49	11	16	1			
3	F	9	Total	C	N	O	S	0	0	0
			77	49	11	16	1			

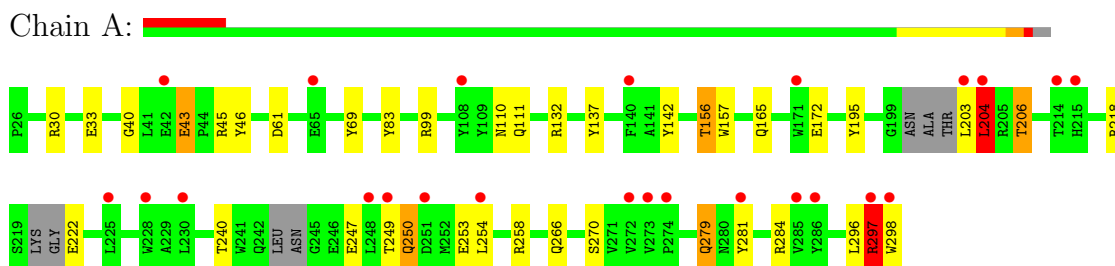
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	36	Total 36	O 36	0	0
4	B	11	Total 11	O 11	0	0
4	C	33	Total 33	O 33	0	0
4	D	23	Total 23	O 23	0	0
4	E	2	Total 2	O 2	0	0

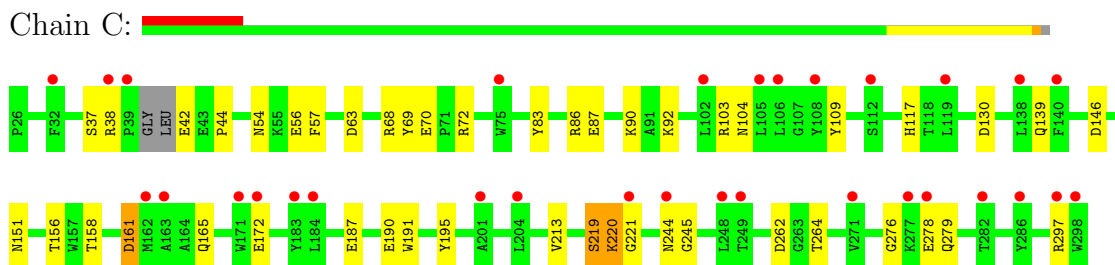
3 Residue-property plots

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.

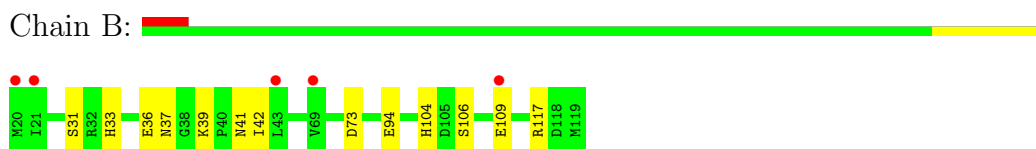
- Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain



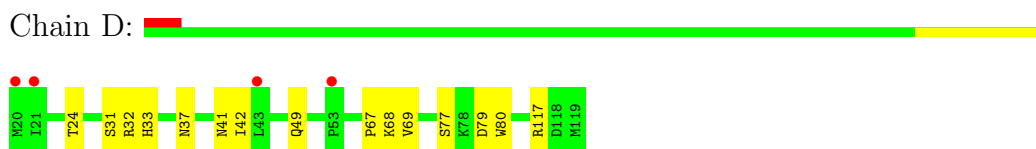
- Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain



- Molecule 2: Beta-2-microglobulin



- Molecule 2: Beta-2-microglobulin



- Molecule 3: Insulin derived 9-mer peptide



- Molecule 3: Insulin derived 9-mer peptide

Chain F: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	87.22Å 101.13Å 117.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.45 – 2.33 46.45 – 1.93	Depositor EDS
% Data completeness (in resolution range)	74.6 (46.45-2.33) 88.0 (46.45-1.93)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.19 (at 1.94Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.217 , 0.294 0.273 , 0.280	Depositor DCC
R_{free} test set	2191 reflections (5.16%)	DCC
Wilson B-factor (Å ²)	25.1	Xtriage
Anisotropy	0.519	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 48.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	12 of 69183 reflections (0.017%)	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	6355	wwPDB-VP
Average B, all atoms (Å ²)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 32.65 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 9.0861e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

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	0.55	1/2268 (0.0%)	0.77	4/3073 (0.1%)
1	C	0.49	0/2298	0.66	1/3117 (0.0%)
2	B	0.44	0/855	0.60	0/1158
2	D	0.49	0/855	0.66	0/1158
3	E	0.44	0/78	0.46	0/103
3	F	0.40	0/78	0.58	0/103
All	All	0.50	1/6432 (0.0%)	0.69	5/8712 (0.1%)

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	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	297	ARG	CG-CD	-5.42	1.38	1.51

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	204	LEU	CB-CG-CD1	8.15	124.86	111.00
1	A	204	LEU	CA-CB-CG	7.60	132.78	115.30
1	A	297	ARG	NE-CZ-NH1	-7.28	116.66	120.30
1	C	130	ASP	N-CA-C	6.04	127.29	111.00
1	A	206	THR	N-CA-C	5.14	124.88	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	203	LEU	Peptide
1	A	204	LEU	Peptide
1	A	250	GLN	Peptide

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	2205	0	10	21	0
1	C	2233	0	0	28	0
2	B	829	0	0	7	1
2	D	829	0	0	10	0
3	E	77	0	0	2	0
3	F	77	0	0	2	0
4	A	36	0	0	3	0
4	B	11	0	0	1	0
4	C	33	0	0	6	1
4	D	23	0	0	5	0
4	E	2	0	0	0	0
All	All	6355	0	10	65	1

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 10.

The worst 5 of 65 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:103:ARG:NH2	4:C:317:HOH:O	2.11	0.82
1:C:68:ARG:NH1	1:C:70:GLU:OE2	2.15	0.79
1:C:219:SER:OG	1:C:220:LYS:N	2.19	0.75
2:B:104:HIS:ND1	2:B:106:SER:OG	2.20	0.74
2:D:69:VAL:N	4:D:211:HOH:O	2.21	0.74

All (1) 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	Distance(Å)	Clash(Å)
2:B:94:GLU:OE1	4:C:317:HOH:O[2_555]	2.06	0.14

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	259/273 (95%)	239 (92%)	16 (6%)	4 (2%)	15	12
1	C	267/273 (98%)	250 (94%)	13 (5%)	4 (2%)	15	12
2	B	98/100 (98%)	95 (97%)	3 (3%)	0	100	100
2	D	98/100 (98%)	95 (97%)	3 (3%)	0	100	100
3	E	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	F	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
All	All	736/764 (96%)	691 (94%)	37 (5%)	8 (1%)	21	20

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	206	THR
1	A	249	THR
1	C	220	LYS
1	A	279	GLN
1	C	44	PRO

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	228/232 (98%)	220 (96%)	8 (4%)	48	62
1	C	231/232 (100%)	227 (98%)	4 (2%)	73	85

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	95/95 (100%)	94 (99%)	1 (1%)	84	92
2	D	95/95 (100%)	93 (98%)	2 (2%)	66	81
3	E	8/8 (100%)	8 (100%)	0	100	100
3	F	8/8 (100%)	8 (100%)	0	100	100
All	All	665/670 (99%)	650 (98%)	15 (2%)	63	78

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	254	LEU
1	A	297	ARG
1	C	190	GLU
1	A	250	GLN
1	C	139	GLN

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 ⓘ

There are no ligands in this entry.

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	266/273 (97%)	0.74	24 (9%) 10 16	32, 54, 110, 135	0
1	C	271/273 (99%)	0.72	31 (11%) 6 9	35, 65, 104, 131	0
2	B	100/100 (100%)	0.44	5 (5%) 28 39	36, 53, 76, 117	0
2	D	100/100 (100%)	0.63	4 (4%) 36 48	31, 48, 78, 137	0
3	E	9/9 (100%)	0.86	1 (11%) 6 10	41, 51, 55, 61	0
3	F	9/9 (100%)	0.77	1 (11%) 6 10	44, 51, 65, 69	0
All	All	755/764 (98%)	0.68	66 (8%) 10 17	31, 56, 103, 137	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	20	MET	9.7
1	A	203	LEU	8.7
1	C	204	LEU	7.7
1	A	204	LEU	7.3
2	B	20	MET	5.1

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

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.