WORLDWIDE
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:29 PM GMT

PDB ID : 1P7Q
Title : Crystal Structure of HLA-A2 Bound to LIR-1, a Host and Viral MHC Receptor
Authors : Willcox, B.E.; Thomas, L.M.; Bjorkman, P.J.
Deposited on : 2003-05-05
Resolution : 3.40 Å(reported)

The following versions of software and data (see references (i)) 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 (i)

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

The reported resolution of this entry is 3.40 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.
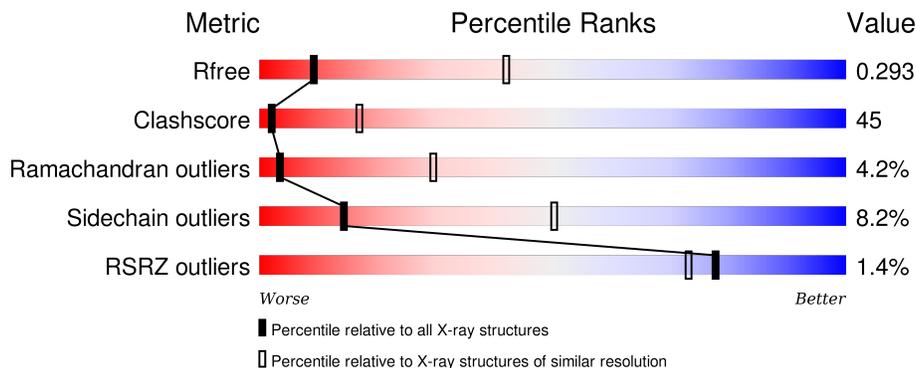
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1476 (3.50-3.30)
Clashscore	102246	1611 (3.50-3.30)
Ramachandran outliers	100387	1571 (3.50-3.30)
Sidechain outliers	100360	1571 (3.50-3.30)
RSRZ outliers	91569	1485 (3.50-3.30)

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. 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	276	38% 53% 8% •
2	B	99	23% 63% 6% • 6%
3	C	9	22% 67% 11%
4	D	197	4% 36% 47% 9% • 7%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4498 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 HLA class I histocompatibility antigen, A-2 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total 2237	C 1399	N 403	O 426	S 9	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	93	Total 775	C 496	N 129	O 147	S 3	0	0	0

- Molecule 3 is a protein called POL polyprotein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	9	Total 69	C 46	N 12	O 11	0	0	0

- Molecule 4 is a protein called leukocyte immunoglobulin-like receptor 1.

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	183	Total 1417	C 907	N 244	O 260	S 6	0	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: HLA class I histocompatibility antigen, A-2 alpha chain

Chain A:
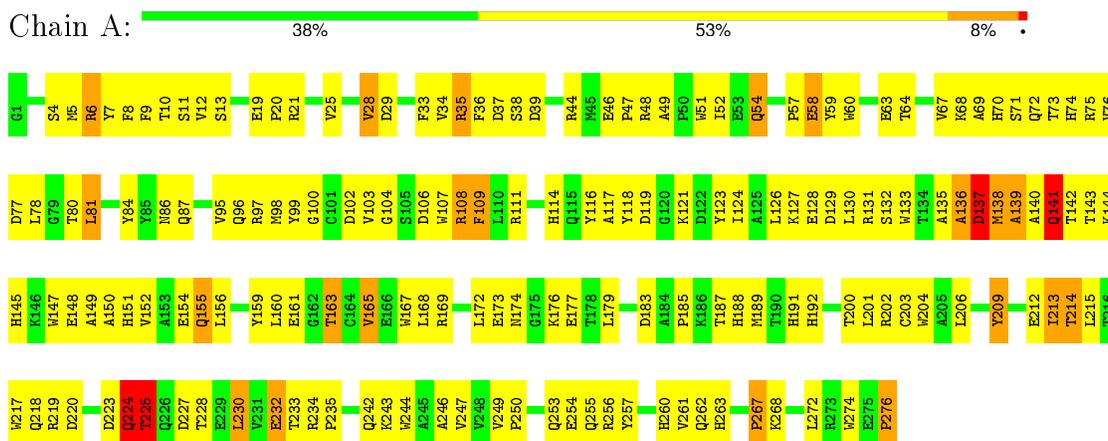
- Molecule 2: Beta-2-microglobulin

Chain B:
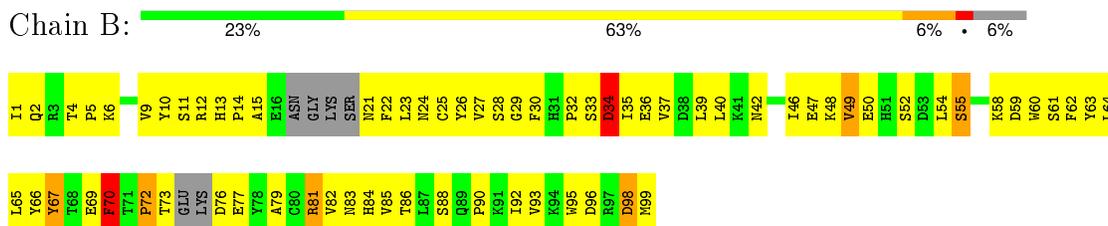
- Molecule 3: POL polyprotein

Chain C:

- Molecule 4: leukocyte immunoglobulin-like receptor 1

Chain D:

H69 A70 C75 Y76 Y77 GLY SER ASP THR ALA GLY R84 S85 E86 L91 E92 L93 V94 V95 T96 G97 I100 L104 S105 A106 Q107 P108 S109 P110 V111 V112 N113 S114 V118 T119 L120 D123 V126 F131 I132 L133 C134 K135 E136 G137 E138 ASP GLU H141 L145 N146

S147 Q148 P149 H150 A151 S154 S155 R156 A157 I158 F159 V164 S165 P166 S167 R168 R169 W170 W171 Y172 R173 C174 Y175 A176 Y177 D178 S179 N180 S181 P182 Y183 E184 L187 P188 L191 L192 E193 L194 L195 V196 L197 G198

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	113.74Å 113.74Å 89.46Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	33.11 – 3.40 33.11 – 3.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (33.11-3.40) 99.7 (33.11-3.20)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.69 (at 3.18Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.218 , 0.309 0.211 , 0.293	Depositor DCC
R_{free} test set	489 reflections (5.18%)	DCC
Wilson B-factor (Å^2)	57.1	Xtriage
Anisotropy	0.505	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 41.0	EDS
Estimated twinning fraction	0.043 for -h,-k,l	Xtriage
L-test for twinning[2]	$<|L|> = 0.47, <L^2> = 0.30$	Xtriage
Outliers	1 of 11344 reflections (0.009%)	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	4498	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.42% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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	0.49	1/2303 (0.0%)	1.36	14/3131 (0.4%)
2	B	0.43	0/796	0.69	0/1080
3	C	0.56	0/70	0.82	0/94
4	D	0.44	1/1460 (0.1%)	0.78	3/1995 (0.2%)
All	All	0.46	2/4629 (0.0%)	1.10	17/6300 (0.3%)

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	1	3
4	D	0	1
All	All	1	4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	225	THR	C-N	-10.90	1.08	1.34
4	D	52	GLN	C-N	6.28	1.48	1.34

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	224	GLN	O-C-N	-35.73	65.53	122.70
1	A	225	THR	O-C-N	-30.64	73.67	122.70
1	A	225	THR	CA-C-N	22.10	165.82	117.20
1	A	224	GLN	CA-C-N	20.79	162.94	117.20
1	A	139	ALA	N-CA-CB	-16.48	87.02	110.10

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	225	THR	CA

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	224	GLN	Mainchain,Peptide
1	A	225	THR	Mainchain
4	D	118	VAL	Mainchain

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	2237	0	2069	205	0
2	B	775	0	724	78	0
3	C	69	0	78	15	0
4	D	1417	0	1348	129	0
All	All	4498	0	4219	394	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 45.

The worst 5 of 394 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)
4:D:118:VAL:HG22	4:D:119:THR:H	0.90	1.04
4:D:118:VAL:CG2	4:D:119:THR:H	1.72	1.02
4:D:118:VAL:HG22	4:D:119:THR:N	1.74	1.01
4:D:6:PRO:HB2	4:D:27:GLN:H	1.26	1.00
4:D:37:LEU:HG	4:D:47:ILE:HD13	1.46	0.98

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	274/276 (99%)	224 (82%)	42 (15%)	8 (3%)	6	40
2	B	87/99 (88%)	71 (82%)	12 (14%)	4 (5%)	3	26
3	C	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
4	D	175/197 (89%)	145 (83%)	19 (11%)	11 (6%)	2	18
All	All	543/581 (94%)	446 (82%)	74 (14%)	23 (4%)	3	29

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	38	SER
1	A	137	ASP
4	D	63	PRO
4	D	181	SER
1	A	136	ALA

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	229/232 (99%)	212 (93%)	17 (7%)	17	55
2	B	87/94 (93%)	81 (93%)	6 (7%)	19	59
3	C	8/8 (100%)	7 (88%)	1 (12%)	6	27
4	D	152/171 (89%)	137 (90%)	15 (10%)	10	39
All	All	476/505 (94%)	437 (92%)	39 (8%)	14	50

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	34	ASP
2	B	81	ARG
4	D	191	LEU
2	B	55	SER
2	B	67	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	255	GLN
2	B	21	ASN
4	D	107	GLN
1	A	224	GLN
4	D	141	HIS

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 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(Å²)	Q<0.9
1	A	276/276 (100%)	-0.54	0	100	100	2, 26, 71, 117	0
2	B	93/99 (93%)	-0.37	0	100	100	2, 30, 82, 108	0
3	C	9/9 (100%)	-0.69	0	100	100	5, 16, 29, 46	0
4	D	183/197 (92%)	0.00	8 (4%)	38	34	8, 49, 111, 153	0
All	All	561/581 (96%)	-0.34	8 (1%)	78	73	2, 32, 89, 153	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	61	PRO	3.7
4	D	37	LEU	3.3
4	D	27	GLN	3.3
4	D	26	CYS	2.9
4	D	60	PHE	2.8

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 (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.