W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 17, 2020 – 10:39 am BST

PDB ID	:	3PGD
Title	:	Crystal Structure of HLA-DR1 with CLIP106-120, canonical peptide orientation
Authors	:	Gunther, S.; Schlundt, A.; Sticht, J.; Roske, Y.; Heinemann, U.; Wiesmuller, K.-H.; Jung, G.; Falk, K.; Rotzschke, O.; Freund, C.
Deposited on	:	2010-11-01
Resolution	:	2.72 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 following versions of software and data (see references (i)) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

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 2.72 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.
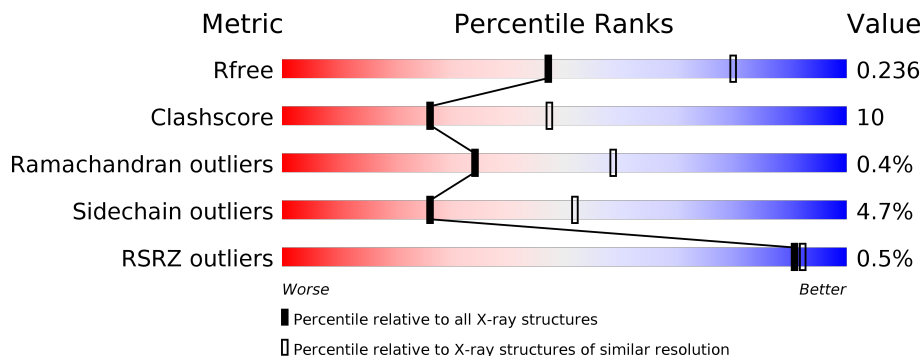
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3359 (2.74-2.70)
Clashscore	141614	3686 (2.74-2.70)
Ramachandran outliers	138981	3622 (2.74-2.70)
Sidechain outliers	138945	3623 (2.74-2.70)
RSRZ outliers	127900	3276 (2.74-2.70)

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 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	193	71% 21% · 8%
1	D	193	70% 21% · 7%
2	B	199	% 62% 29% · 7%
2	E	199	% 66% 24% · 7%
3	C	15	93% 7%
3	F	15	87% 13%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6324 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 II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	178	Total 1450	C 941	N 234	O 270	S 5	0	0	0
1	D	179	Total 1459	C 947	N 236	O 271	S 5	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	EXPRESSION TAG	UNP P01903
D	0	MET	-	EXPRESSION TAG	UNP P01903

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DRB1-1 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	185	Total 1503	C 946	N 267	O 283	S 7	0	0	0
2	E	185	Total 1507	C 948	N 267	O 285	S 7	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	EXPRESSION TAG	UNP P04229
E	0	MET	-	EXPRESSION TAG	UNP P04229

- Molecule 3 is a protein called HLA class II histocompatibility antigen gamma chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	15	Total 116	C 75	N 20	O 17	S 4	0	0	0
3	F	15	Total 116	C 75	N 20	O 17	S 4	0	0	0

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	40	Total 40	O 40	0	0
4	B	34	Total 34	O 34	0	0
4	C	13	Total 13	O 13	0	0
4	D	49	Total 49	O 49	0	0
4	E	31	Total 31	O 31	0	0
4	F	6	Total 6	O 6	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 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.
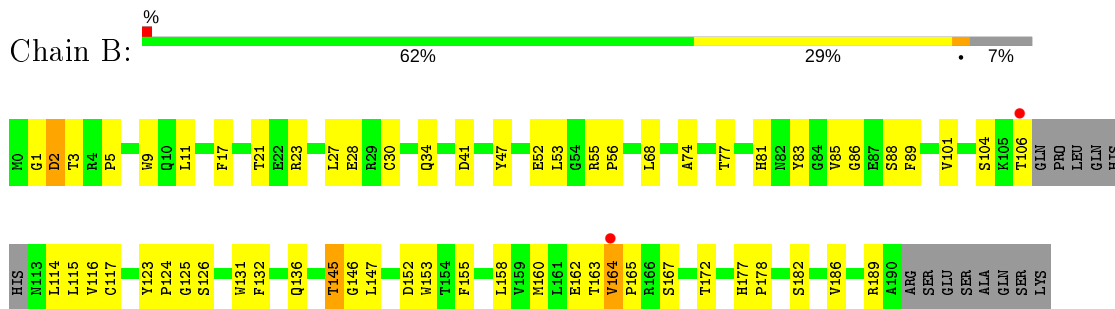
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain
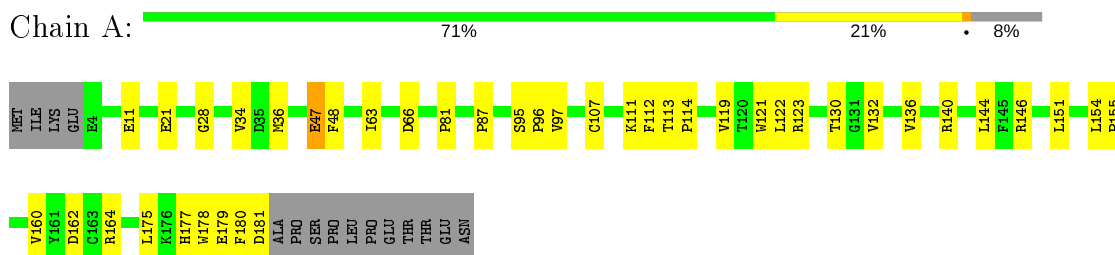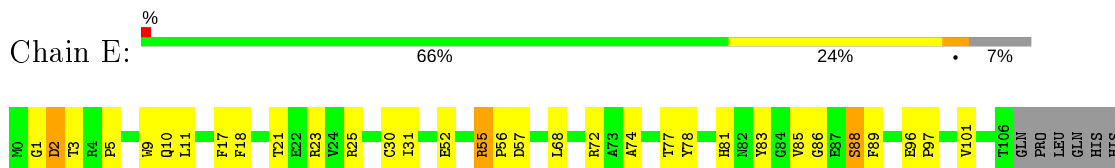

- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain
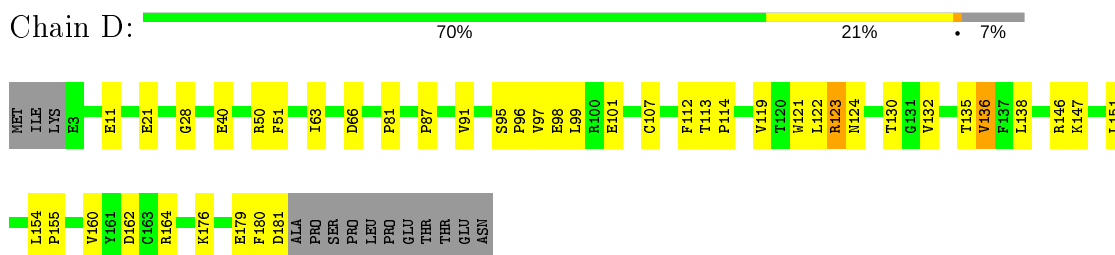
- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

N113 L114 L115 V116 S126 T145 G146 L147 D152 W153 L158 V159 M160 T163 V164 P165 R166 S167 G168 E169 V170 Y171 T172 S182 V186 E187 W188 R189 A190 ARG SER GLU SER ALA GLN SER LYS

• Molecule 3: HLA class II histocompatibility antigen gamma chain

Chain C:

93%   7%

K106 M120

• Molecule 3: HLA class II histocompatibility antigen gamma chain

Chain F:

87%   13%

K106 M115 M120

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	94.45Å 94.45Å 275.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.39 – 2.72 33.52 – 2.72	Depositor EDS
% Data completeness (in resolution range)	99.9 (33.39-2.72) 100.0 (33.52-2.72)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.44 (at 2.72Å)	Xtriage
Refinement program	PHENIX 1.6.2_432	Depositor
R, R_{free}	0.196 , 0.242 0.192 , 0.236	Depositor DCC
R_{free} test set	1725 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	44.6	Xtriage
Anisotropy	0.289	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 48.1	EDS
L-test for twinning[2]	$<|L|> = 0.46, <L^2> = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6324	wwPDB-VP
Average B, all atoms (Å^2)	44.0	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 71.56 % 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 2.5696e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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 (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	0.45	0/1495	0.61	0/2042
1	D	0.46	0/1504	0.61	0/2053
2	B	0.45	0/1539	0.64	2/2089 (0.1%)
2	E	0.43	0/1543	0.65	3/2094 (0.1%)
3	C	0.44	0/117	0.60	0/155
3	F	0.42	0/117	0.60	0/155
All	All	0.45	0/6315	0.63	5/8588 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	23	ARG	NE-CZ-NH2	6.46	123.53	120.30
2	E	55	ARG	NE-CZ-NH2	-5.96	117.32	120.30
2	E	55	ARG	NE-CZ-NH1	5.85	123.22	120.30
2	B	189	ARG	NE-CZ-NH1	-5.51	117.54	120.30
2	E	189	ARG	NE-CZ-NH1	5.26	122.93	120.30

There are no chirality outliers.

There are no planarity outliers.

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	1450	0	1375	23	0
1	D	1459	0	1388	32	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1503	0	1433	37	0
2	E	1507	0	1437	37	0
3	C	116	0	133	0	0
3	F	116	0	133	1	0
4	A	40	0	0	1	0
4	B	34	0	0	0	0
4	C	13	0	0	0	0
4	D	49	0	0	4	0
4	E	31	0	0	1	0
4	F	6	0	0	0	0
All	All	6324	0	5899	125	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 10.

The worst 5 of 125 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)
2:E:52:GLU:HG2	2:E:55:ARG:HH22	1.25	1.00
2:E:52:GLU:HG2	2:E:55:ARG:NH2	1.78	0.99
2:B:52:GLU:HG2	2:B:55:ARG:HH22	1.46	0.81
2:E:163:THR:O	2:E:165:PRO:HD3	1.82	0.79
2:B:163:THR:O	2:B:165:PRO:HD3	1.84	0.77

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	176/193 (91%)	165 (94%)	11 (6%)	0	100	100
1	D	177/193 (92%)	170 (96%)	6 (3%)	1 (1%)	25	48

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	181/199 (91%)	170 (94%)	10 (6%)	1 (1%)	25	48
2	E	181/199 (91%)	173 (96%)	7 (4%)	1 (1%)	25	48
3	C	13/15 (87%)	13 (100%)	0	0	100	100
3	F	13/15 (87%)	12 (92%)	1 (8%)	0	100	100
All	All	741/814 (91%)	703 (95%)	35 (5%)	3 (0%)	34	58

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	2	ASP
2	E	2	ASP
1	D	136	VAL

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	160/177 (90%)	153 (96%)	7 (4%)	28	54
1	D	161/177 (91%)	157 (98%)	4 (2%)	47	75
2	B	164/179 (92%)	153 (93%)	11 (7%)	16	35
2	E	165/179 (92%)	157 (95%)	8 (5%)	25	51
3	C	13/13 (100%)	12 (92%)	1 (8%)	13	29
3	F	13/13 (100%)	12 (92%)	1 (8%)	13	29
All	All	676/738 (92%)	644 (95%)	32 (5%)	26	52

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	145	THR
2	B	182	SER
2	E	167	SER
2	B	167	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	C	120	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	156	GLN
2	E	156	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 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	178/193 (92%)	-0.34	0	100	100	19, 38, 74, 107	0
1	D	179/193 (92%)	-0.41	0	100	100	18, 37, 74, 107	0
2	B	185/199 (92%)	-0.30	2 (1%)	80	82	21, 44, 82, 103	0
2	E	185/199 (92%)	-0.32	2 (1%)	80	82	21, 43, 80, 102	0
3	C	15/15 (100%)	-0.20	0	100	100	25, 33, 64, 86	0
3	F	15/15 (100%)	-0.24	0	100	100	25, 33, 64, 86	0
All	All	757/814 (92%)	-0.34	4 (0%)	91	92	18, 41, 76, 107	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	106	THR	4.1
2	E	164	VAL	3.2
2	B	164	VAL	3.1
2	E	168	GLY	2.4

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.