wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:45 PM GMT

PDB ID	:	1QHM
Title	:	ESCHERICHIA COLI PYRUVATE FORMATE LYASE LARGE DOMAIN
Authors	:	Leppanen, V.-M.; Merckel, M.C.; Ollis, D.L.; Wong, K.K.; Kozarich, J.W.; Goldman, A.
Deposited on	:	1999-05-19
Resolution	:	2.80 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 (i) symbol.

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

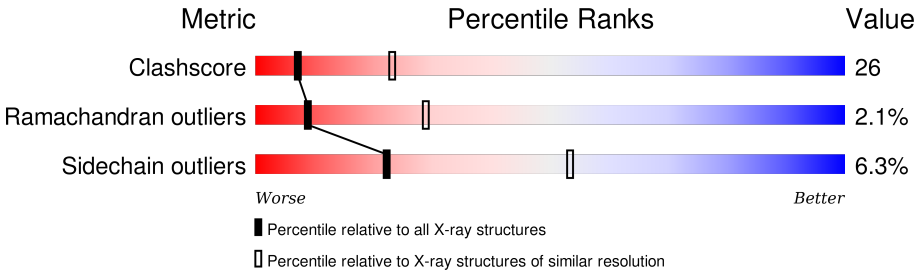
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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 2.80 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))
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	624	60% 33% 5% •
1	B	624	58% 35% 5% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9810 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 PYRUVATE FORMATE-LYASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	612	Total	C	N	O	S	0	0	1
			4794	3034	810	920	30			
1	B	612	Total	C	N	O	S	0	0	1
			4782	3026	809	917	30			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	125	Total	O	0	0
			125	125		
2	B	109	Total	O	0	0
			109	109		

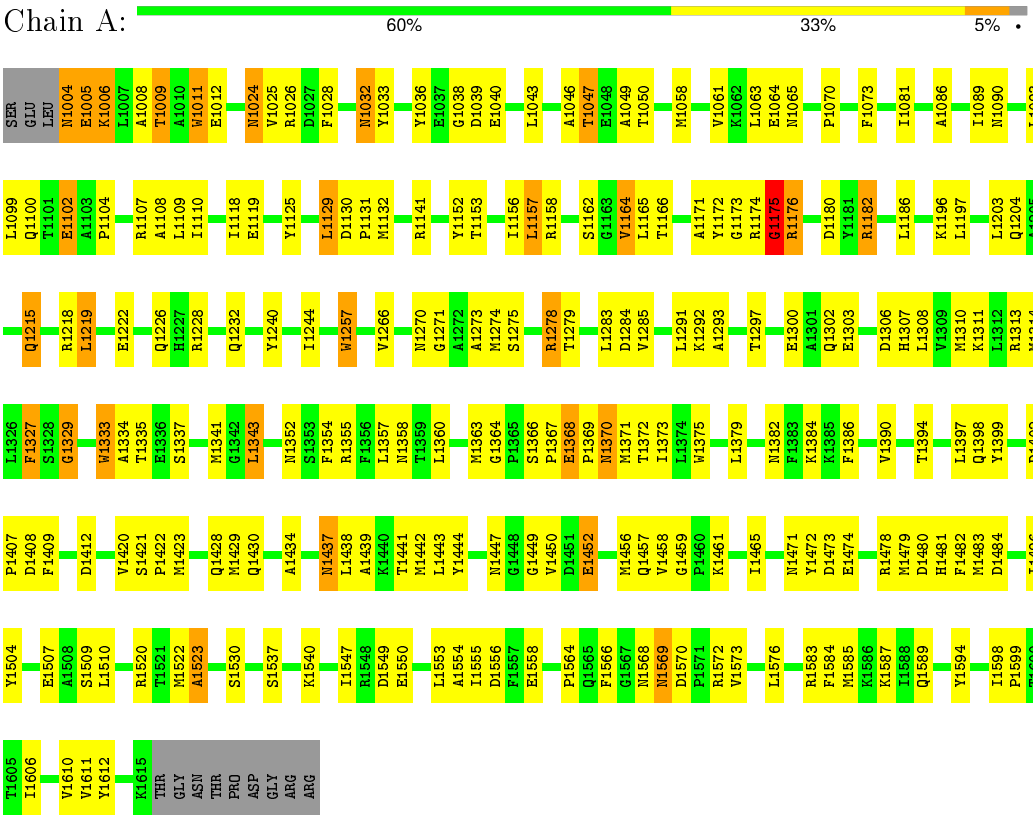
3 Residue-property plots (i)

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.

Note EDS was not executed.

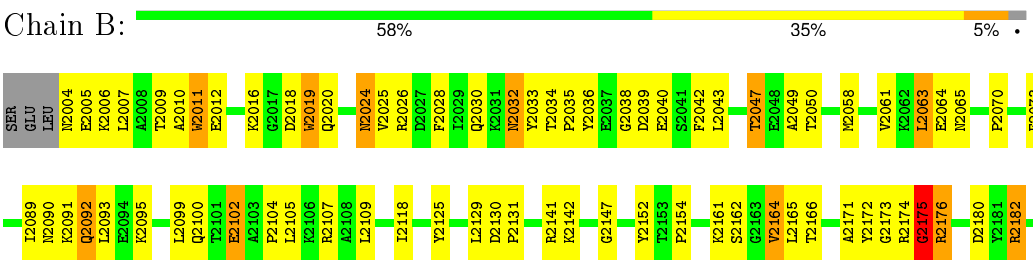
- Molecule 1: PYRUVATE FORMATE-LYASE

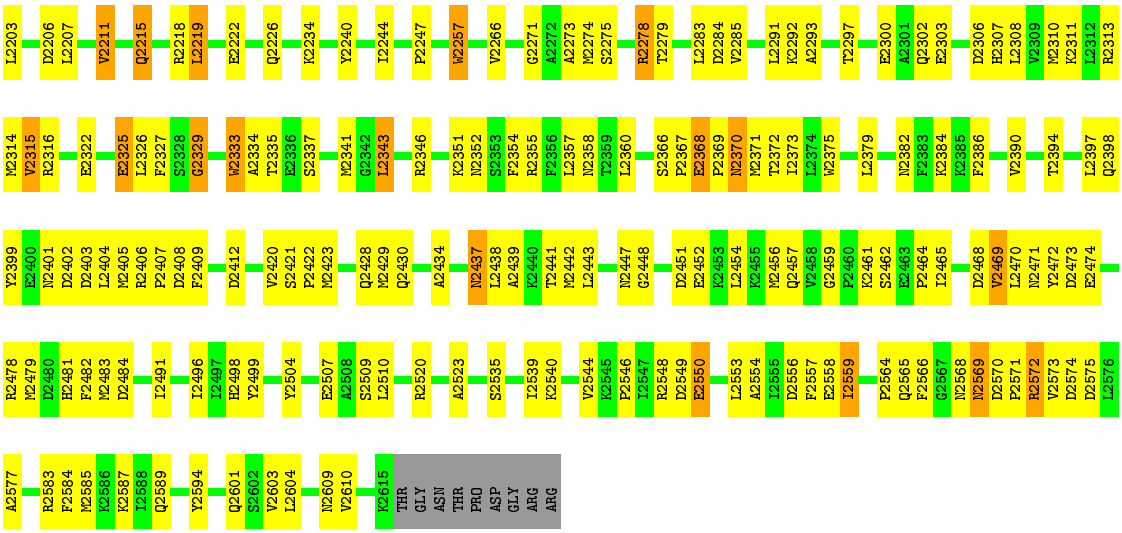
Chain A:



- Molecule 1: PYRUVATE FORMATE-LYASE

Chain B:




L2203 D2206 L2207 V2211 Q2215 R2218 L2219 E2222 Q2226 K2234 Y2240 I2244 P2247 W2257 V2266 G2271 A2272 A2273 M2274 S2275 R2278 T2279 L2283 D2284 V2285 L2291 K2292 A2293 T2297 E2300 A2301 Q2302 E2303 D2306 H2307 L2308 V2309 M2310 K2311 L2312 R2313
M2314 V2315 R2316 E2322 E2325 L2326 F2327 S2328 G2329 W2333 A2334 T2335 E2336 S2337 M2341 G2342 L2343 R2346 K2351 N2352 S2353 F2354 R2355 F2356 L2357 N2358 T2359 L2360 S2366 P2367 E2368 P2369 N2370 M2371 T2372 I2373 L2374 W2375 L2379 N2382 F2383 K2384 K2385 F2386 V2390 T2394 L2397 Q2398
Y2399 E2400 N2401 D2402 D2403 L2404 M2405 R2406 P2407 D2408 F2409 D2412 V2420 S2421 P2422 M2423 Q2428 M2429 Q2430 A2434 N2437 L2438 A2439 K2440 T2441 M2442 L2443 N2447 G2448 D2451 E2452 K2453 L2454 K2455 M2456 Q2457 V2458 G2459 P2460 K2461 S2462 E2463 P2464 I2465 D2468 V2469 L2470 N2471 Y2472 D2473 E2474
R2478 M2479 D2480 H2481 F2482 M2483 D2484 I2491 I2496 I2497 H2498 Y2499 Y2504 E2507 A2508 S2509 L2510 R2520 A2523 S2535 I2539 K2540 V2544 K2545 P2546 I2547 R2548 D2549 E2550 L2553 A2554 I2555 D2556 F2557 E2558 I2559 P2564 Q2565 F2566 G2567 N2568 N2569 D2570 P2571 R2572 V2573 D2574 D2575 L2576
A2577 R2583 F2584 M2585 K2586 K2587 I2588 Q2589 Y2594 Q2601 S2602 V2603 L2604 N2609 V2610 K2615 THR GLY ASN THR PRO ASP GLY ARG ARG

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	140.80Å 140.80Å 215.90Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.80	Depositor
% Data completeness (in resolution range)	99.0 (20.00-2.80)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	5.20	Depositor
Refinement program	CNS 0.4	Depositor
R, R_{free}	0.228 , 0.253	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9810	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

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	0.46	0/4891	0.66	2/6623 (0.0%)
1	B	0.46	0/4879	0.67	2/6610 (0.0%)
All	All	0.46	0/9770	0.67	4/13233 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1175	GLY	N-CA-C	6.33	128.93	113.10
1	B	2175	GLY	N-CA-C	6.09	128.32	113.10
1	A	1176	ARG	N-CA-C	-5.68	95.66	111.00
1	B	2176	ARG	N-CA-C	-5.52	96.09	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	4794	0	4659	235	0
1	B	4782	0	4636	261	0
2	A	125	0	0	3	0
2	B	109	0	0	5	0
All	All	9810	0	9295	493	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 26.

The worst 5 of 493 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:B:2465:ILE:CA	1:B:2478:ARG:HH12	1.42	1.33
1:B:2465:ILE:CG1	1:B:2478:ARG:NH1	1.96	1.28
1:B:2465:ILE:HG13	1:B:2478:ARG:NH1	1.57	1.14
1:B:2465:ILE:CA	1:B:2478:ARG:NH1	2.11	1.14
1:B:2465:ILE:N	1:B:2478:ARG:HH12	1.49	1.11

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	610/624 (98%)	545 (89%)	54 (9%)	11 (2%)	11 34
1	B	610/624 (98%)	539 (88%)	56 (9%)	15 (2%)	7 24
All	All	1220/1248 (98%)	1084 (89%)	110 (9%)	26 (2%)	9 29

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1175	GLY
1	A	1569	ASN
1	B	2175	GLY
1	B	2327	PHE
1	B	2329	GLY

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	502/528 (95%)	470 (94%)	32 (6%)	22 52
1	B	499/528 (94%)	468 (94%)	31 (6%)	23 54
All	All	1001/1056 (95%)	938 (94%)	63 (6%)	22 53

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1443	LEU
1	B	2047	THR
1	B	2473	ASP
1	A	1452	GLU
1	A	1550	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1457	GLN
1	B	2024	ASN
1	B	2457	GLN
1	A	1471	ASN
1	A	1498	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 (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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.