BRD4 (bromodomain-containing protein 4)

For fragment scientists. This report surfaces what a fragment campaign design step needs: ranked candidate pockets computed across a conformational ensemble (not a single static frame), per-residue interaction fingerprints from PLIP against representative liganded structures, and family/conservation/precedent as prioritization signal before any soaking or SPR/ITC/NMR work begins. It does not screen a fragment library, predict affinity, dock a ligand, or design a molecule — those stay explicitly out of scope.

Independently verified this session — not pipeline output

Independently checked against PDB 3MXF (BRD4 BD1 + JQ1, the acetyl-lysine mimetic that seeded the whole BET-inhibitor chemical series). The rank-1 cluster (persistence 1.0) matches 12 of 14 real JQ1 contact residues (86%) — strict top-1, including the signature Asn140 hydrogen-bond residue and the WPF-shelf residues (Trp81, Pro82, Phe83). A clean rank-1 win on the pipeline's own default ranking, not adjusted or cherry-picked. One limitation worth surfacing up front for fragment scientists specifically: Section 3 below reports no self-matching interaction fingerprints for this target. The bromodomain family (PF00439) is broad — over 40 human bromodomains across many different proteins share the fold but bind different acetyl-lysine contexts — and none of the cached ligand-bound family structures aligned closely enough to BRD4 BD1 specifically at the coverage threshold this pipeline requires. The pocket recovery above is unaffected (it comes from structural analysis, not interaction fingerprints), but per-residue PLIP contact data is not available for this report.

Protein Knowledge Report — O60885

1. Target summary

- Input type: uniprot_accession
- Label: O60885
- Length: 1362 aa
- UniProt accession: O60885
- Submitted: 2026-07-17T10:11:34Z
- Sequence:

MSAESGPGTRLRNLPVMGDGLETSQMSTTQAQAQPQPANAASTNPPPPETSNPNKPKRQTNQLQYLLRVVLKTLWKHQFAWPFQQPVDAVKLNLPDYY
KIIKTPMDMGTIKKRLENNYYWNAQECIQDFNTMFTNCYIYNKPGDDIVLMAEALEKLFLQKINELPTEETEIMIVQAKGRGRGRKETGTAKPGVSTVPNTT
QASTPPQTQTPQPNPPPVQATPHPFPAVTPDLIVQTPVMTVVPPQPLQTPPPVPPQPQPPPAPAPQPVQSHPPIIAATPQPVKTKKGVKRKADTTTPTTIDPI
HEPPSLPPEPKTTKLGQRRESSRPVKPPKKDVPDSQQHPAPEKSSKVSEQLKCCSGILKEMFAKKHAAYAWPFYKPVDVEALGLHDYCDIIKHPMDMSTIK
SKLEAREYRDAQEFGADVRLMFSNCYKYNPPDHEVVAMARKLQDVFEMRFAKMPDEPEEPVVAVSSPAVPPPTKVVAPPSSSDSSSDSSSDSDSSTDDS
EEERAQRLAELQEQLKAVHEQLAALSQPQQNKPKKKEKDKKEKKKEKHKRKEEVEENKKSKAKEPPPKKTKKNNSSNSNVSKKEPAPMKSKPPPTYESEE
EDKCKPMSYEEKRQLSLDINKLPGEKLGRVVHIIQSREPSLKNSNPDEIEIDFETLKPSTLRELERYVTSCLRKKRKPQAEKVDVIAGSSKMKGFSSSESESSS
ESSSSDSEDSETEMAPKSKKKGHPGREQKKHHHHHHQQMQQAPAPVPQQPPPPPQQPPPPPPPQQQQQPPPPPPPPSMPQQAAPAMKSSPPPFIATQV
PVLEPQLPGSVFDPIGHFTQPILHLPQPELPPHLPQPPEHSTPPHLNQHAVVSPPALHNALPQQPSRPSNRAAALPPKPARPPAVSPALTQTPLLPQPPMAQ
PPQVLLEDEEPPAPPLTSMQMQLYLQQLQKVQPPTPLLPSVKVQSQPPPPLPPPPHPSVQQQLQQQPPPPPPPQPQPPPQQQHQPPPRPVHLQPMQFSTH
IQQPPPPQGQQPPHPPPGQQPPPPQPAKPQQVIQHHHSPRHHKSDPYSTGHLREAPSPLMIHSPQMSQFQSLTHQSPPQQNVQPKKQELRAASVVQPQ
PLVVVKEEKIHSPIIRSEPFSPSLRPEPPKHPESIKAPVHLPQRPEMKPVDVGRPVIRPPEQNAPPPGAPDKDKQKQEPKTPVAPKKDLKIKNMGSWASLVQK
HPTTPSSTAKSSSDSFEQFRRAAREKEEREKALKAQAEHAEKEKERLRQERMRSREDEDALEQARRAHEEARRRQEQQQQQRQEQQQQQQQQAAAV
AAAATPQAQSSQPQSMLDQQRELARKREQERRRREAMAATIDMNFQSDLLSIFEENLF

2. Family classification

- Family: Bromodomain (PF00439) — Bromodomain
- E-value: 3.10e-45 Bit score: 152.8
- Model span: 4-83 of model (70-151 of query)
- Passes family's curated gathering threshold: yes

Source: HMMER hmmscan vs Pfam-A (local)

3. Known ligand precedent

- Total structures in this family: 2187
- Distinct ligand scaffolds curated below: 8

PDB ID	Ligand(s)	Resolution (Å)	Why selected
5IG6	6B3	0.91	highest-resolution structure...
4QB3	30M	0.94	highest-resolution structure...
6FO5	DZH	0.95	highest-resolution structure...
7QYO	GKI	0.98	highest-resolution structure...
7Z9W	BYZ	0.99	highest-resolution structure...
4ZC9	4MW	0.99	highest-resolution structure...
9P34	A1CGY	0.99	highest-resolution structure...
6I7Y	H7E	1.00	highest-resolution structure...

scanned the 100 highest-resolution of 2187 total structures in this family (RCSB result capped for query cost) — the curated examples below are drawn from that subset, not from the full family

Source: RCSB PDB REST API, curated top-N by distinct ligand scaffold + resolution

Interaction fingerprints (Evidence Integration Layer, Phase 1)

8 ligand-bound structure(s) exist for this target's Pfam family (PF00439), but none checked aligned to the query at >=50% sequence coverage — likely other members of the same broad family, not this specific protein. Interaction fingerprints require a structure of the query protein itself.

Source: PLIP (Protein-Ligand Interaction Profiler), run on representative structures from known ligand precedent

4. Conservation

- Family: PF00439
- Seed alignment size: 41 sequences
- Mean pairwise identity (seed alignment): 17.3%
- Note: computed over a sample of 40/41 seed sequences

Source: Pfam seed alignment (via InterPro API), mean pairwise identity across aligned members at matched-region columns

5. Structural analysis

Not available (error). no structure provider could resolve this target: all structure providers exhausted: AlphaFold DB: AlphaFold DB metadata fetch failed: HTTP Error 429: Too Many Requests; ESMFold: not available for this target

Source: StructureProvider + PocketDetector abstraction (see api/knowledge_protein/structural.py)

6. Confidence flags

1 caution flag(s) — discount confidence accordingly:

- [Structural analysis] structural_analysis_unavailable: Structural analysis did not complete (error): no structure provider could resolve this target: all structure providers exhausted: AlphaFold DB: AlphaFold DB metadata fetch failed: HTTP Error 429: Too Many Requests; ESMFold: not available for this target.

Source: rules-based checks in api/knowledge_protein/confidence_flags.py — fixed thresholds on numbers already computed elsewhere in this report, no learned calibration, no combined/weighted score.

7. Similar known proteins

Sequence-based (Pfam family co-membership, top 5 by identity):

PDB ID	Sequence identity to query	Ligand(s)
5IG6	0.0%	6B3

4QB3	0.0%	30M
5IBN	0.0%	(none recorded)
7QYO	0.0%	GKI
7Z9W	0.0%	BYZ

Structure-based (Foldseek vs. full PDB structural index, Evidence Integration Layer Phase 2):

Not computed. no structure provider could resolve this target: all structure providers exhausted: AlphaFold DB: AlphaFold DB metadata fetch failed: HTTP Error 429: Too Many Requests; ESMFold: not available for this target

Source: sequence-based: Pfam family co-membership among cached PDB reference structures (sequence-based); Foldseek vs local PDB structural database (structure-based)

8. Limitations & data provenance

Auto-populated from each section's own status — not hand-maintained.

Section	Status	Detail
Family classification	returned data	
Known ligand precedent	returned data	scanned the 100 highest-resolution of …
Interaction fingerprints	no self-matching ligand-bound structur…	8 ligand-bound structure(s) exist for …
Functional context (UniProt + ClinVar)	returned data	
Conservation	returned data	computed over a sample of 40/41 seed s…
Structural analysis	errored	no structure provider could resolve th…
Similar known proteins	returned data	
Similar known proteins (structure-base…	not computed this run	no structure provider could resolve th…

3 of 8 sections did not return data this run: Interaction fingerprints, Structural analysis, Similar known proteins (structure-based).