Unlocking Medulloblastoma Detection: Transketolase (TKT) as a Novel Biomarker for Early Diagnosis

VARSHA MURALI

Taft Charter High School, Winnetka, CA

Published January, 2026

Medulloblastoma is an aggressive form of cancer that originates as a brain tumor in the brain's cerebellum, a portion of the brain that controls balance and regulates complex motor functions. Transketolase (TKT2) is a protein found in the cerebrospinal fluid of patients with medulloblastoma and has been identified as a potential biomarker for medulloblastoma due to its elevated levels. Antibodies are proteins that help the body fight off bacteria and viruses; each has a unique antibody that attacks it. I hypothesize that the antibody binds to the predicted binding site of the TKT protein and helps detect this protein. In this work, I have performed computational simulations to identify the antibody that can bind to the TKT protein and be used as a biomarker. I used AlphaFold 3 to retrieve the amino acid sequence for the TKT protein structure and then modeled the protein using Chimera X. I used HDOCK to carry out molecular docking. I provided the software with the TKT protein and specific antibody, and it provided us a model of how the antibody would bind to the TKT protein. I found that the 4a6y antibody binds best to the TKT protein and can be used as a biomarker to identify the TKT protein. The future application of this research would be developing an actual antibody from my model to detect this cancer in its early stages, especially in younger children.

1. INTRODUCTION

Medulloblastoma is an aggressive form of cancer that originates as a brain tumor in the brain's cerebellum, a portion of the brain that controls balance and regulates complex motor functions. [1] According to the National Cancer Institute, "Seventy percent of medulloblastoma cases occur in children and 30 percent occur in adults." [2] Symptoms of medulloblastoma when the tumor is in the cerebellum alone are balance/motor skill issues. [1] If the cancer is causing hydrocephalus, or Cerebral Spinal Fluid(CSF) buildup in cavities, symptoms include headaches, nausea, vomiting, blurred/double vision, extreme sleepiness, confusion, and seizures. [1] As the tumor spreads further, the symptoms increase. Treatment for medulloblastoma involves surgical removal of the cancer and relieving intracranial pressure due to hydrocephalus, sometimes with a shunt to drain CSF buildup. [3] Surgery is followed by radiation and chemotherapy to kill off the cancer cells that were not surgically removed, since the recurrence of medulloblastoma is high. [3] Medulloblastoma treatments have long-term effects that severely impair the quality of life of patients. [3] Common problems after medulloblastoma treatment include long-term neurological and sensory impairments, endocrine deficits, secondary tumors, and neurocognitive impairment. [3]

As shown in Figure 1, the incidence rate in children aged 0 to 14 is approximately 0.47 per 1000,000, highlighting its strong association with early childhood. [4] In contrast, medulloblastoma in adolescents and young adults (ages 15 to 39) drops significantly to around 0.10 per 100,000; it becomes exceedingly rare in individuals over 40, with an incidence of just 0.02 per 100,000. [5] This age-dependent decline in incidence underscores the importance of early diagnosis and research efforts focused on pediatric populations. [4] The diagram illustrates the role of the TKT protein in detecting medulloblastoma, a malignant brain tumor that commonly arises in the cerebellum, particularly in children.(Khanna et al., 2017) Medulloblastoma is shown as a red mass located in the posterior part of the brain, affecting the cerebellum, which is responsible for coordination and balance. [4] The image emphasizes the presence of TKT protein in the cerebrospinal fluid (CSF), suggesting its potential as a biomarker for early diagnosis. [6] By extracting a sample of CSF through a lumbar puncture, clinicians can detect elevated levels of TKT protein, which may indicate the presence of medulloblastoma. [5] This method provides a non-invasive diagnostic approach, aiding in the timely identification and management of the tumor. [4]

AlphaFold 3 is the latest advancement in computational biology, developed by DeepMind and Isomorphic Labs. [7] It represents a significant leap forward from its predecessor, AlphaFold 2, by significantly expanding the scope of molecular structure

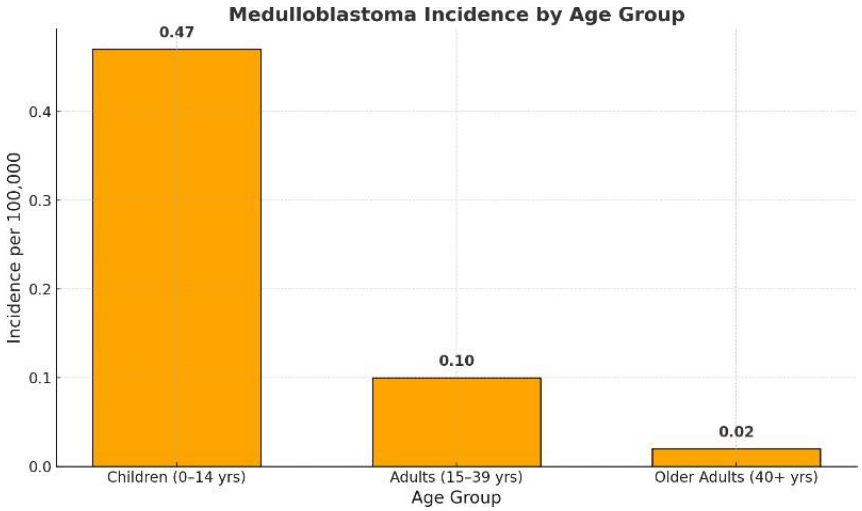

Fig. 1. Bar graph showing the incidence of medulloblastoma across different age groups. The tumor is most prevalent in children, while its occurrence significantly decreases in adolescents and young adults.

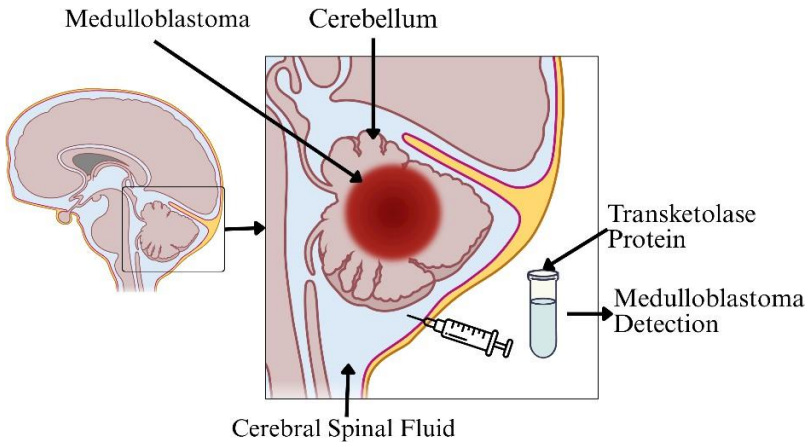

Fig. 2. Medulloblastoma, a malignant brain tumor, originates in the cerebellum and releases biomarkers into the cere- brospinal fluid (CSF). TKT protein detection in CSF is a potential diagnostic method for medulloblastoma.

prediction. [8] Unlike AlphaFold 2, which focused primarily on predicting the 3D structure of single protein sequences, AlphaFold 3 can accurately model individual proteins and complexes involving DNA, RNA, small molecules (ligands), ions, and antibodies. [7] This allows scientists to study a broader range of biological interactions using a unified, AI-powered platform. AlphaFold 3 achieves this by integrating physical and geometric principles with deep learning techniques, making it a powerful tool for drug discovery, understanding cellular machinery, and designing novel therapeutics. In this research, I have performed computational modeling of the TKT protein, marking the first time this protein has been linked to medulloblastoma detection. Our simulations demonstrate its potential for use in detecting this protein, and thereby, in the early cancer diagnosis.

2. METHOD

To obtain the amino acid sequence of the TKT protein, UniProt, a free database of protein sequences and functional information, is utilized. [9] I used this database to obtain the amino acid sequence of TKT to model the protein. I then uploaded the amino acid sequence into the AlphaFold 3 server to obtain the protein 3D structure. I then opened these files in ChimeraX-1.9, a program that allows us to visualize and analyze the molecular structure of the protein. [10] PrankWeb (https://prankweb.cz/) predicts ligand-binding sites on protein structures using

machine learning, highlighting likely binding pockets on a 3D model. [11] PROTTER (https://wlab.ethz.ch/protter/start/) visualizes membrane topology, showing transmembrane regions and domains from protein sequences. [12] The Human Protein Atlas (https://www.proteinatlas.org/) provides detailed data on protein expression, localization, and disease relevance across human tissues. [13]

3. RESULTS

Structural and Electrostatic Characterization of TKT Protein: To identify potential ligand-binding regions and understand structural organization, I performed a comprehensive structural analysis of the TKT protein (Figure 3). The top five predicted binding pockets were visualized using molecular surface mapping (Figure 3a), with the highest-ranking binding site (Score: 11.68) marked in red, indicating a strong probability for ligand interaction. The quaternary structure of TKT revealed a dimeric conformation, where each monomer was distinctly visualized in pink and purple (Figure 3b). This dimeric state represents the biologically active form of the protein. Electrostatic sur- face analysis (Figure 3c) showed distinct charge distributions, with negatively charged (red), positively charged (blue), and neutral (white) regions. These electrostatic patterns are critical for understanding potential molecular recognition and binding interactions.

Conservation of Antibody Binding Sites Across TKT Isoforms: Structural superimposition of antibody-bound TKT protein complexes from various PDB entries revealed a conserved pattern of active site and antibody-binding interfaces (Figure 4). The conserved regions were predominantly colored in pink (active sites) and grey (binding interfaces), suggesting that antibody detection methods targeting these epitopes could be broadly applicable across different isoforms. These findings validate the potential of these antibodies for consistent recognition of TKT in cerebrospinal fluid (CSF), an important step for diagnostic applications.

Topology and Subcellular Localization of TKT: Topology analysis revealed that the TKT protein spans multiple domains, traversing the periplasmic space, membrane, and cytoplasmic regions (Figure 5a). Key post-translational modifications, including PTMs and disulfide bonds, were mapped along the sequence, indicating structural and functional regulation across cellular compartments. Subcellular localization analysis confirmed that TKT is predominantly detected in the nucleoplasm (Figure 5b), pointing to a significant role in nuclear metabolic pathways and potentially linking it to transcriptional regulation or DNA-associated enzymatic functions.

Binding Affinity of Antibodies to TKT: The binding energies of various antibodies interacting with TKT were assessed through docking simulations (Figure 6). Antibodies such as 1fl5 and 1li1 exhibited the strongest affinities, with binding energies below -11 kcal/mol, suggesting robust interactions. In contrast, antibodies like 1a5f and 1nlb showed relatively weaker binding affinities. These findings can inform the selection of optimal antibody candidates for both therapeutic and diagnostic development targeting TKT.

4. DISCUSSION

Figure 6 represents the annotated secondary structure of a protein, combining its amino acid sequence with structural motifs such as alpha-helices, beta-strands, and turn regions. The protein sequence is displayed using single-letter amino acid codes, with each segment color-coded—red, blue, or green—to indicate different domains or chains. Purple coils represent alpha-helices (labeled H1 to H34), while pink arrows indicate beta-strands. Turns and loops are labeled with Greek letters like γ or $\beta\gamma$, marking transitions between structural elements. Various symbols, such as colored triangles and orange boxes, highlight important residues, including possible active sites, binding motifs, or mutation points. Sequence positions are clearly numbered for reference. This type of annotated map is commonly used in structural biology to understand protein folding, domain organization, and functional regions.

The Ramachandran plot shown in Figure 7 represents the dihedral angle distribution (phi and psi) of amino acid residues in the protein structure 1ck. Each blue point corresponds to the torsion angles of an individual residue, revealing conformational preferences within the protein. The plot is divided into regions based on the steric feasibility of angle combinations: the dark red and brown areas indicate the most favored conformations, typically associated with right-handed α-helices and β-sheets, while the yellow areas are allowed but less favorable, and the white regions represent disallowed conformations due to steric clashes. The dense clustering of residues in the favored regions suggests the protein has a well-folded and stereochemically valid structure. This plot is a critical tool in protein validation, offering insights into the overall stability and correctness of the modeled or experimentally derived structure.

Figure 8 illustrates key molecular interactions within a protein-ligand complex: In panel (a), the schematic shows how two protein chains—Chain A (92 amino acids) and Chain B (88 amino acids)—interact with each other. Multiple non-covalent interactions stabilize the interface between them. Specifically, there are 11 salt bridges (electrostatic interactions between oppositely charged residues), 56 hydrogen bonds, and many non-bonded contacts (598) such as van der Waals forces. These interactions are essential for maintaining the protein dimer's structural integrity and proper function. No disulfide bonds (covalent links between cysteine residues) are observed in this interface.

In panel (b), a detailed interaction map is shown for Tpp 682(A), a ligand or small molecule, binding to the protein. The diagram identifies specific amino acid residues that interact with the ligand through hydrogen bonds (green dashed lines), electrostatic interactions, and hydrophobic interactions (represented by orange arcs). Key residues involved include His 69, His 263, and Asn 187, among others. These interactions are critical for ligand recognition, binding stability, and potentially the protein's biological activity.

Therapeutic candidates (TCs) related to TKT and its isoform TKTL1, highlighting their targets, associated diseases, mechanisms of action, and current development status.

5. CONCLUSION

This study demonstrates the potential of TKT as a novel biomarker for the early detection of medulloblastoma, a highly aggressive pediatric brain tumor. Through computational modeling and molecular docking simulations, the antibody 4a6y was identified as having the strongest binding affinity to the TKT protein, suggesting its utility in a diagnostic assay. By employing state-of-the-art tools such as AlphaFold 3, ChimeraX, HDOCK, and PRODIGY, I successfully predicted the structural binding interactions between TKT and various antibodies. These in silico results support the hypothesis that antibody-based detection of

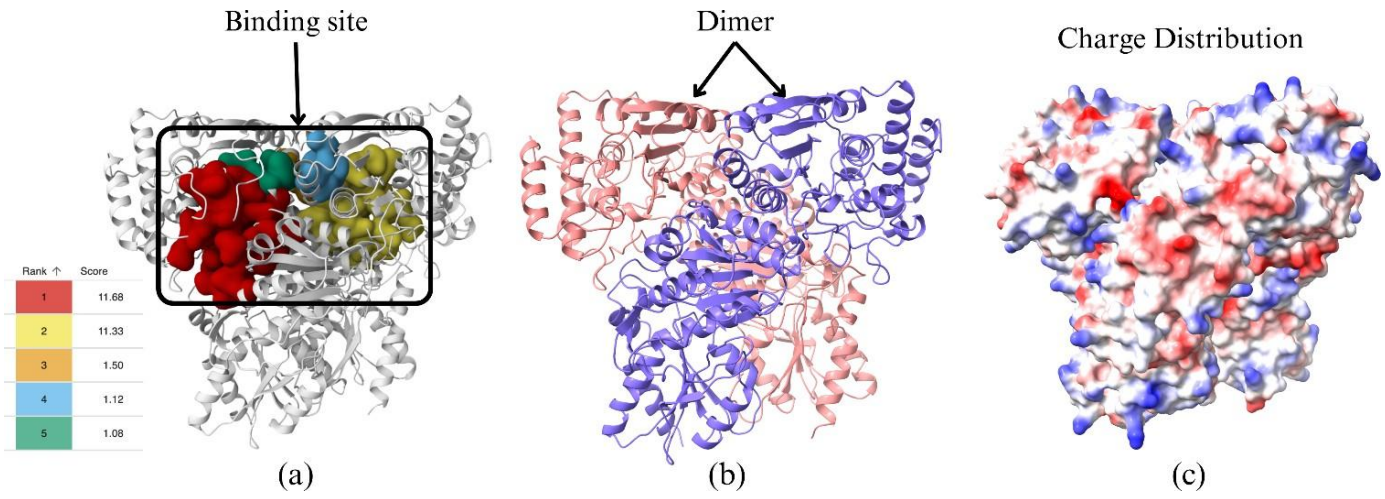

Fig. 3. Structural and Electrostatic Characterization of the Target Protein. (a) Predicted Binding Site: Top five ranked binding pockets are visualized on the protein surface, with the highest-ranked site highlighted in red (Score: 11.68), indicating the most probable ligand binding region; (b) Dimeric Architecture: The protein is shown in its biologically active dimeric form, with individual monomers colored pink and purple to demonstrate the quaternary structure; and (c) Electrostatic Surface Potential: The surface charge distribution is mapped, with red representing negatively charged regions, blue denoting positively charged regions, and white indicating neutral areas—important for understanding molecular recognition and interaction specificity.

Table 1.Therapeutic candidates (TCs) related to TKT and its isoform TKTL1, highlighting their targets, associated diseases and mechanisms of action.

TC / Drug	Target	Related Disease	Mechanism	Status
Oxythiamine	TKT / TKTL1	Cancer (colorectal, lung)	Inhibits transketolase function	Preclinical / Experimental
Benfotiamine	TKT cofactor (B1)	Diabetic complications	Enhances TKT activity	Clinical use (supplement)
B-OT (Benfo-Oxythiamine)	TKTL1	Prostate Cancer, HCC	Dual effect: block & modulate PPP	Case studies, experimental
Thiamine (Vitamin B1)	Cofactor for TKT	Wernicke's, TRMA, Alzheimer's	Restores enzymatic activity	Widely used supplement
RNAi / CRISPR TKTL1	TKTL1 gene	Drug-resistant tumors	Gene knockdown to reduce tumor growth	Research / Preclinical

TKT could facilitate early diagnosis of medulloblastoma. Moving forward, the development of an antibody-based screening test—validated by experimental techniques like ELISA—may significantly improve diagnostic accuracy and patient outcomes, particularly in children. Overall, this research provides a promising foundation for future translational applications in cancer diagnostics.

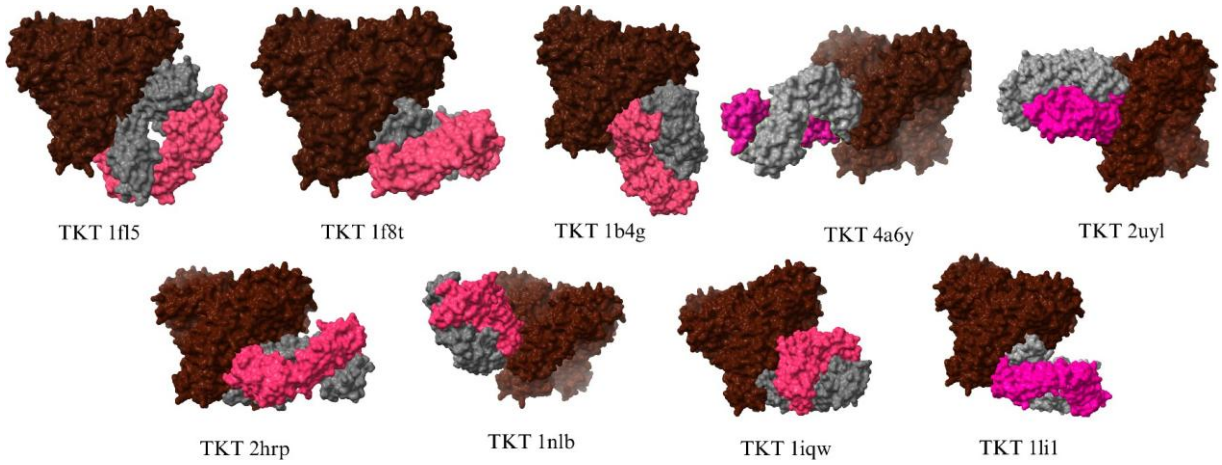

Fig. 4. Molecular docked structure of TKT protein bound antibodies that can be used in the detection of this protein in CSF. Structural comparison of TKT proteins from various PDB entries highlights conserved active sites (pink) and binding interfaces (grey). The consistent spatial arrangement suggests functional conservation across TKT isoforms and species.

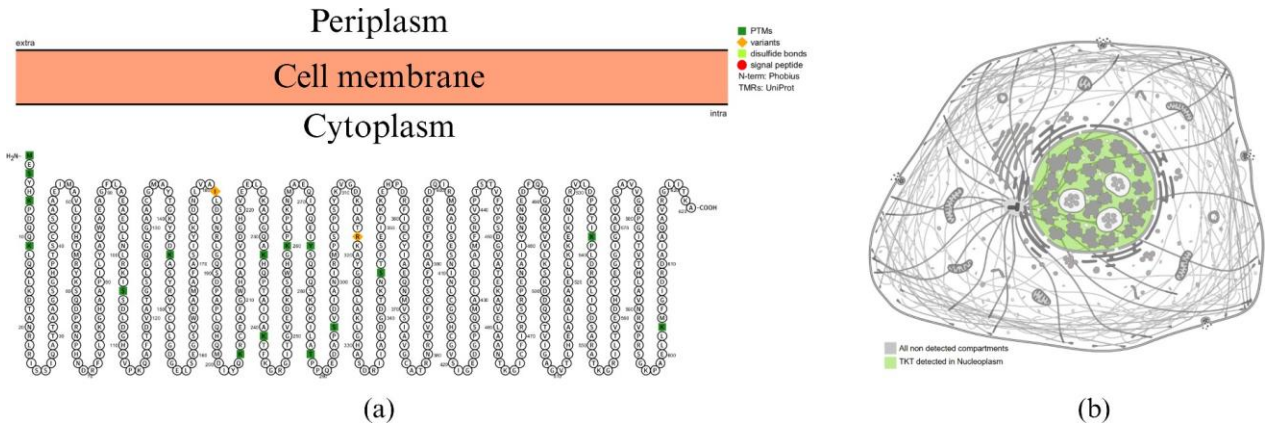

Fig. 5. (a) TKT topology reveals its transmembrane regions and post-translational modifications across the periplasm, membrane, and cytoplasm; and (b) Subcellular localization indicates TKT is predominantly detected in the nucleoplasm, suggesting nuclear metabolic involvement.

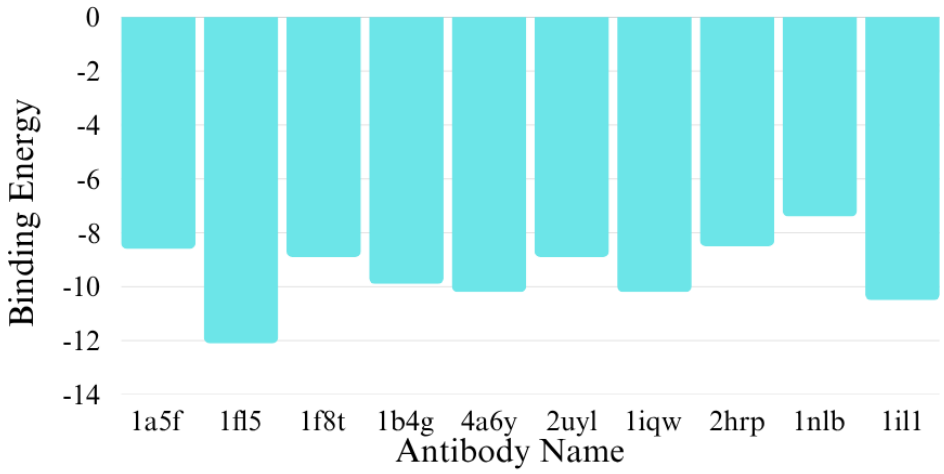

Fig. 6. Binding energies of various antibodies to the TKT protein. Lower values indicate stronger binding affinity.

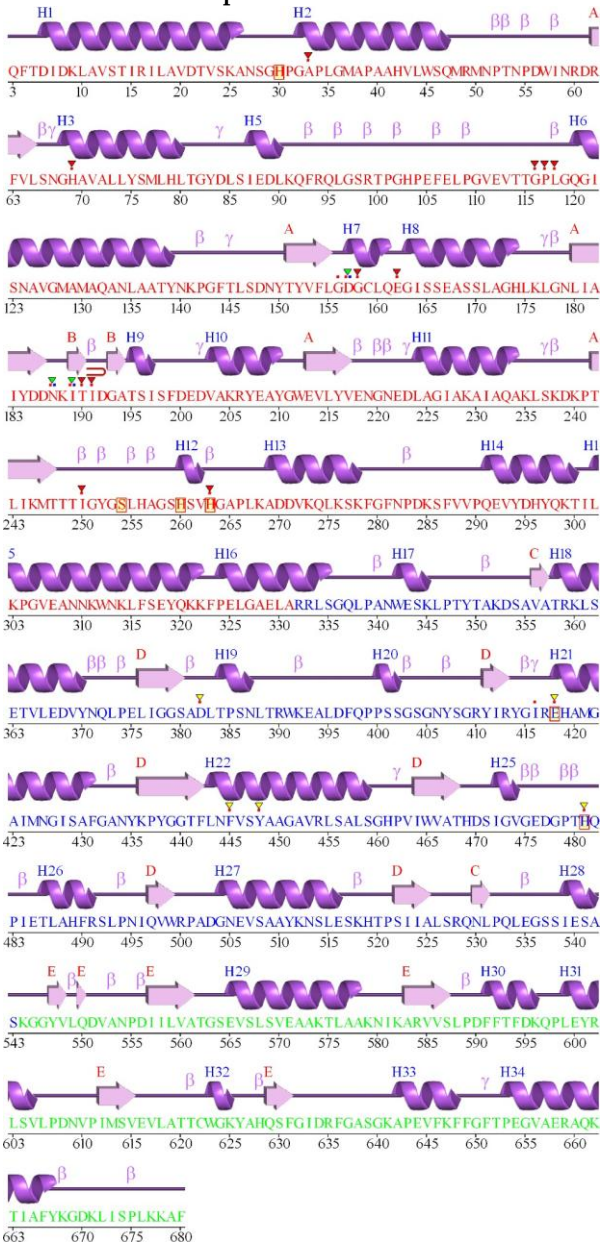

Fig. 7. Annotated secondary structure of a protein showing helices (H1–H34), beta strands, and functional motifs along the amino acid sequence.

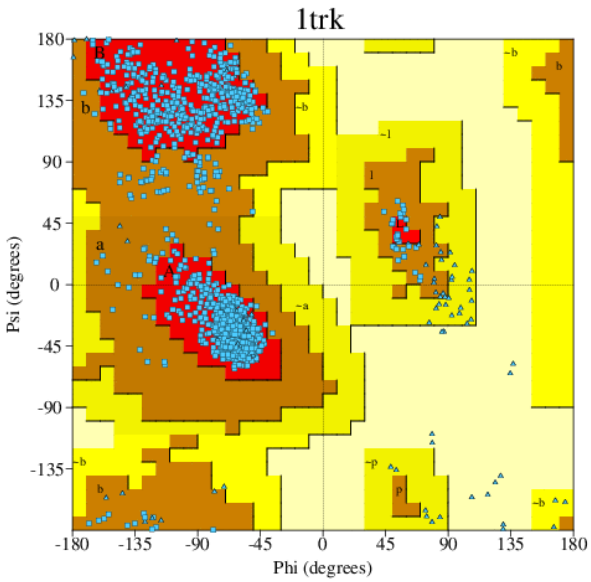

Fig. 8. Ramachandran plot of protein 1ck showing the distribution of phi (Φ) and psi (Ψ) dihedral angles, with most residues clustered in favored *α*-helix and *β*-sheet regions.

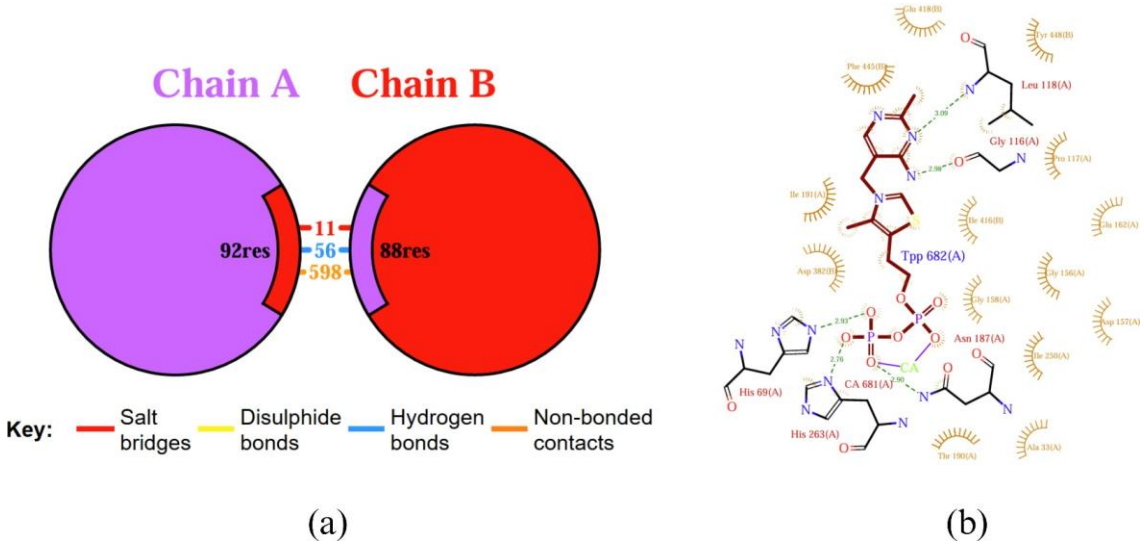

Fig. 9. (a) Chain Interaction Map: Interaction between Chain A and B showing 11 salt bridges, 56 hydrogen bonds, and 598 non-bonded contacts; and (b) Ligand Binding Site: TRP 682(A) interacts with nearby residues via hydrogen bonds, electrostatic, and hydrophobic contacts.

REFERENCES

1. P. A. Northcott, G. W. Robinson, C. P. Kratz, D. J. Mabbott, S. L. Pomeroy, S. C. Clifford, . . Pfister, and S. M, Nature Reviews Disease Primers 5, 11–11 (2019).
2. C. H. Rickert and W. Paulus, “Epidemiology of central nervous system tumors in childhood and adolescence based on the new WHO classification,” Nervous System 17, 503–511 (2001).
3. J. V. Dyk, T. Jenkin, R. D. Leung, P. M. K. Cunningham, and J. R, “Medulloblastoma: Treatment technique and radiation dosimetry,” International Journal of Radiation Oncology*Biology*Physics 2, 993–1005 (1977).
4. E. V. Dressler, T. A. Dolecek, M. Liu, and J. L. Villano, “Demographics, patterns of care, and survival in pediatric medulloblastoma,” Journal of Neuro-Oncology 132, 497–506 (2017).
5. P. A. Northcott, A. Korshunov, S. M. Pfister, and M. D. Taylor, “The clinical implications of medulloblastoma subgroups,” Nature Reviews Neurology 8, 340–351 (2012).
6. P. A. Northcott, A. Korshunov, H. Witt, T. Hielscher, C. G. Eberhart, S. Mack, . . Taylor, and M. D, “Medulloblastoma Comprises Four Dis tinct Molecular Variants,” Journal of Clinical Oncology 29, 1408–1414 (2010).
7. J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, . . Jumper, and J. M, “Accurate structure prediction of biomolecular interactions with AlphaFold 3,” Nature 630, 493–500 (2024).
8. J. Skolnick, M. Gao, H. Zhou, and S. Singh, “AlphaFold 2: Why It Works and Its Implications for Understanding the Relationships of Protein Sequence, Structure, and Function,” Journal of Chemical Information and Modeling pp. 4827–4831 (2021).
9. T. U. Consortium, “UniProt: a hub for protein information,” Nucleic acids research 43 (2014).
10. C. E. Meng, D. T. Goddard, F. E. Pettersen, S. G. Couch, J. Z. Pearson,

11. H. J. Morris, and E. T. Ferrin, UCSF ChimeraX: Tools for structure building and analysis pp. 32–32 (2023).
12. R. Krivák and D. Hoksza, “P2Rank: machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure,” Journal of Cheminformatics 10, 39–39 (2018).
13. U. Omasits, C. H. Ahrens, S. Müller, and B. Wollscheid, “Protter: inter- active protein feature visualization and integration with experimental proteomic data,” Bioinformatics 30, 884–886 (2013).
14. F. Pontén, K. Jirström, and M. Uhlen, “The Human Protein Atlas-a tool for pathology,” The Journal of Pathology 216, 387–393 (2008).