EPB41L3

A Target Enabling Package (TEP)

Gene ID / UniProt ID	**2037** / Q9Y2J2 / -
Target Nominator	AMP-AD (Alan Levey[1])
SGC Authors	William Bradshaw, Vittorio Katis, Opher Gileadi
Target PI	Opher Gileadi (SGC Oxford)
Therapeutic Area(s)	Neurodegeneration
Disease Relevance	Expression of EPB41L3 in post-mortem brain tissues is inversely correlated with multiple markers of Alzheimer's disease progression.
Date Approved by TEP Evaluation Group	November 27th 2020
Document version	1.0
Document version date	November 2020
Citation	William Bradshaw, Vittorio Katis, & Opher Gileadi. (2020). EPB41L3; A Target Enabling Package [Data set]. Zenodo. http://doi.org/10.5281/zenodo.4429628
Affiliations	[1] Dept. of Neurology, Emory University

USEFUL LINKS

Agora Open Targets 50 YEARS OMIM Human Genetics Knowledge for the World GeneCards HUMAN GENE DATABASE

(Please note that the inclusion of links to external sites should not be taken as an endorsement of that site by the SGC in any way)

SUMMARY OF PROJECT

EPB41L3 is, with moesin (MSN), a member of the FERM family, which is highly expressed in brain as well as in other tissues (kidney, intestine, and testis). Similar to MSN, it binds to the C-terminus of CD44 and transmits signals to the cytoskeleton. EPB41L3 is inversely correlated with markers of Alzheimer's Disease (AD). We have chosen to pursue the EPB41L3-CD44 interaction as a potential key driver in AD, and as a selectivity target for FERM proteins. This TEP includes expression constructs and methods for purification of the FERM domain from recombinant *E. coli*; Fluorescence-based assays for binding its ligand; a peptide from the cytoplasmic tail of CD44; crystal structures and a soakable crystallization system; and small molecules identified from a crystal-based fragment screen. We have also expressed several related FERM domains which can be used as a selectivity panel when developing new ligands. In follow-up work, we plan to expand the fragment hits to generate chemical tools for both EPB41L3 and moesin.

SCIENTIFIC BACKGROUND

EPB41L3 (Erythrocyte Band 4.1-like protein 3) is named after its similarity to erythrocyte cytoskeletal proteins. Its functions are poorly understood. EPB41L3 is annotated as a possible tumour suppressor (also named DAL-1, for Differentially expressed in Adenocarcinoma of the Lung), inducing apoptosis in cooperation with protein arginine methyltransferases[1-3]. It interacts with transmembrane receptors such as CD44 through its FERM domain, as well as a large number of cytoskeletal proteins[4-6]. Gene knockout in mice indicate that EPB41L3 has a role in maintaining the association of neurons and glial cells, specifically the nodal arrangement of myelin; despite the ultrastructural defects, the homozygous knockout mice did not display behavioural or motor deficiencies up to 1 year of age[7, 8].

There is an urgent need to identify new targets and therapeutic modalities for treatment or prevention of Alzheimer's disease. While many groups have been focussing on a handful of genes with strong GWAS linkage, the AMP-AD consortium attempts to identify potential targets by analysis of differential gene expression (at the RNA and protein levels) in post-mortem brain samples, applying network analysis tools to suggest genes that are key drivers of expression modules. This work has generated a set of target nominations that represent therapeutic hypotheses. The aim of the series of related TEPs (MSN, EPB41L3, INPP5D) is to lay the groundwork and develop tools to test these therapeutic hypotheses in model systems.

RESULTS – THE TEP

Proteins

The FERM domain (aa 107 – 331) was expressed with an N-terminal His_6 tag that can be removed with TEV protease. The tag was removed for protein crystallisation. The uncleaved protein is used for binding assays, where it is bound by anti-His antibodies.

Structures

Highly reproducible crystallisation conditions have been determined for EPB41L3 and the structure of the protein has been determined to 1.45 Å (6IBE). The structure is a typical FERM domain, with three subdomains (**Fig 1**). The structural similarity allows modelling of the binding sites of the CD44 peptide (**Fig 1B**, based on a co-crystal structure of MSN) and of IP3 (**Fig 1C**, based on a cocrystal structure of Radixin).

To identify chemical starting points and binding pockets, we performed an X-ray fragment screen at Diamond Light Source on EPB41L3 crystals grown in the same conditions as the high-resolution structure. In total, 52 binding events were observed across 8 sites (**Fig 1D**). The details of the binding events can be seen in fragalysis; the PDB codes and further details for all structures are listed in table 1. Approximately half of the fragments bound to either a pocket near to the purported inositol phosphate binding site (Site 1) or the CD44 binding site (Site 2). It may be possible to use these fragments as starting points for the development of compounds able to disrupt binding of either of the ligands.

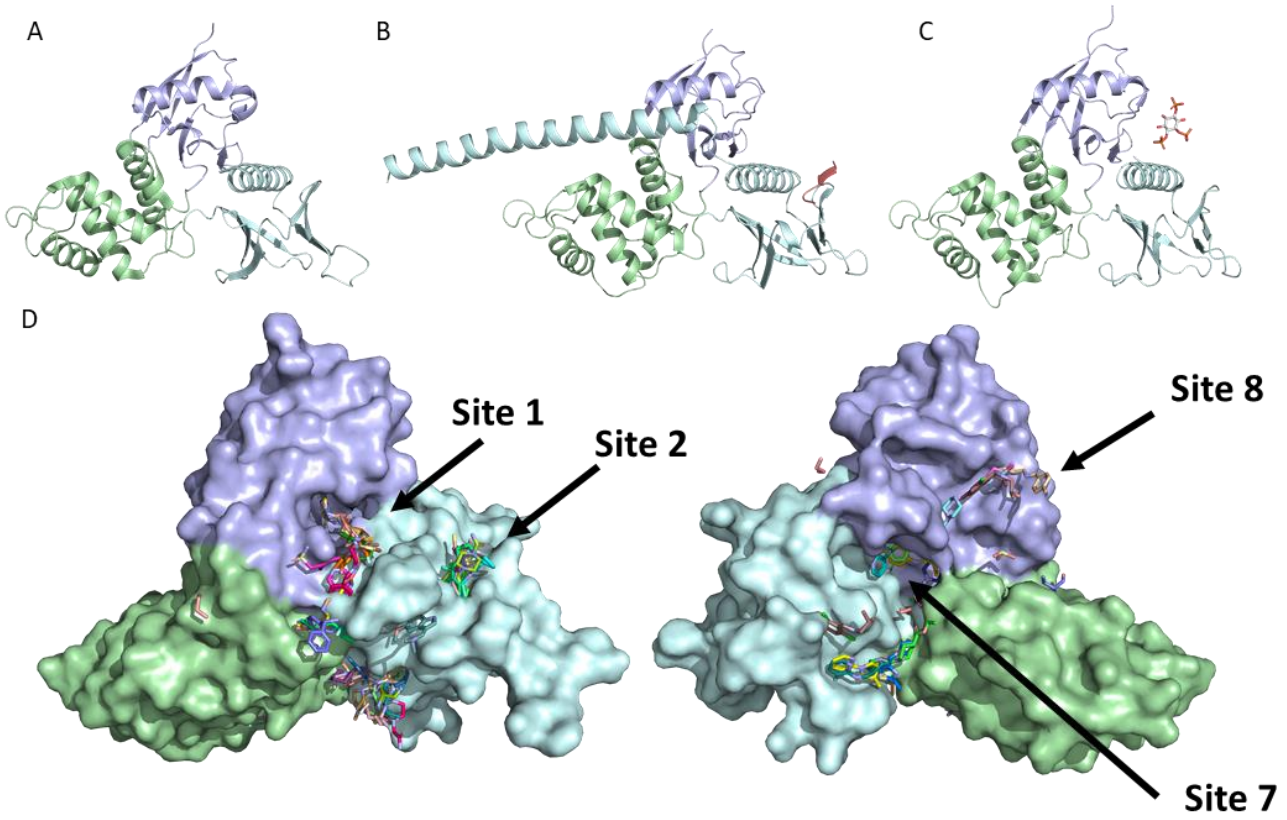


Figure 1. FERM domain structures. (A) The 1.45 Å EPB41L3 structure. FERM domains possess three subdomains – 1, 2 and 3, coloured light blue, light green and light cyan, respectively. **(B)** The CD44-bound structure of moesin (6TXS). The CD44 peptide (orange) adds a strand to a β-sheet that forms part of subdomain 3. **(C)** The IP_3 bound structure of radixin (1GC6). IP_3 is bound near to the interface between subdomains 1 and 2. **(D)** Fragment bound structures of EPB41L3. The two images are rotated 180° relative to each other. Many of the fragments bind either near to the reported inositol phosphate binding site (site 1), or at the CD44 binding site (site 2). Disrupting either of these interactions may be beneficial to modulating the functions of EPB41L3.

Assays

One known biochemical function of EPB41L3 is binding to cytoplasmic tails of transmembrane receptors such as CD44. We chose to focus on this binding, and its displacement by small-molecule ligands, as our primary assay of EPB41L3 activity.

The assay (**Fig 2**) is based on Time-resolved Fluorescence energy transfer. An anti-His antibody conjugated to europium serves as the fluorescence donor, bound to the His6 tag of recombinant EPB41L3. A peptide derived from CD44 (672-691), biotinylated at the N-terminal, is bound to streptavidin conjugated to XL665, a phycobiliprotein pigment. When the peptide is bound to EPB41L3, excitation of the donor at 320 nm results in some fluorescence from the acceptor dye. As a control, we performed a competition of the biotinylated peptide with increasing concentrations of an unmodified peptide; this results in dose-dependent decrease of energy transfer.

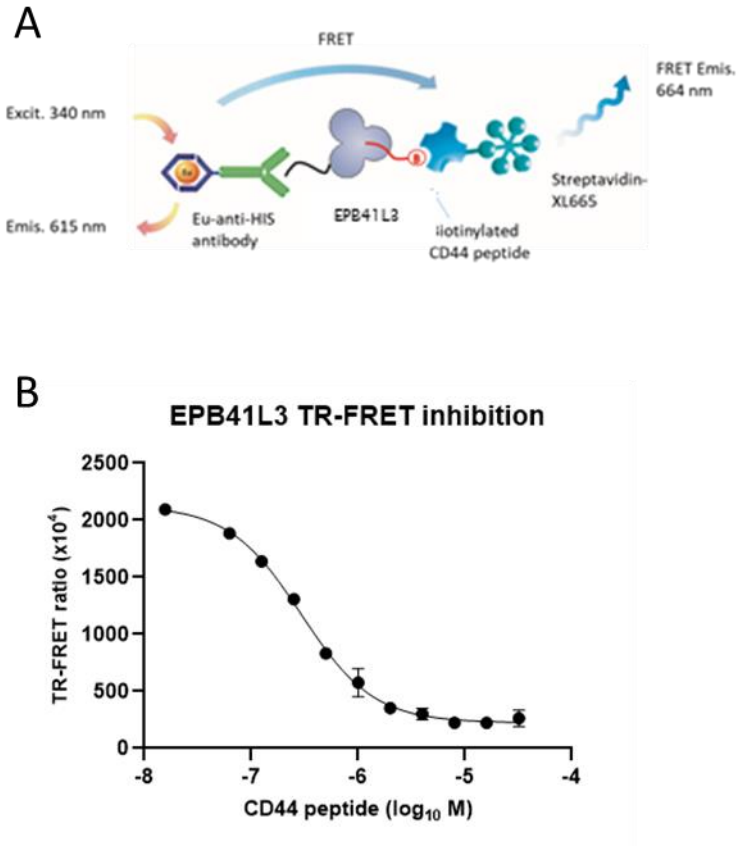


Figure 2. TR-FRET assay of CD44-EPB41L3 binding. **(A)** Diagram of the FRET assay: Donor and acceptor dyes are bound to EPB41L3 and CD44 through anti-his antibodies and streptavidin, respectively. **(B)** Competition of the biotinylated probe peptide with the same peptide lacking a biotin. Effective competition can be seen at sub-micromolar concentration. This assay can be adapted for high-throughput screening.

Chemical matter

Fragment screening

Screening of several hundred compounds was performed by soaking individual crystals of the EPB41L3 FERM domain with high concentrations of small molecule fragments. The crystals were then cryo-cooled and screened on beamline I04-1 at Diamond Light Source. Data was processed by PanDDA to reveal low-occupancy ligand binding.

We were able to identify and refine 51 binding events clustered in 8 sites, as listed in table 1. **Fig 3** shows composite images of some of the compounds bound in sites 1, 2, 7, and 8. Only a subset of bound fragments are shown. Sites 1, 7 and 8 are deep pockets, which may be used for allosteric modulation or for targeting by PROTACs. Molecules bound in site 2 could directly compete with binding of the natural peptide ligands. It has been our experience that even with excellent ligand efficiencies, the small fragments (200-250 Da) used in these experiments do not bind with measurable affinity to the protein in solution; a round of fragment expansion would be needed to achieve sub-millimolar binding.

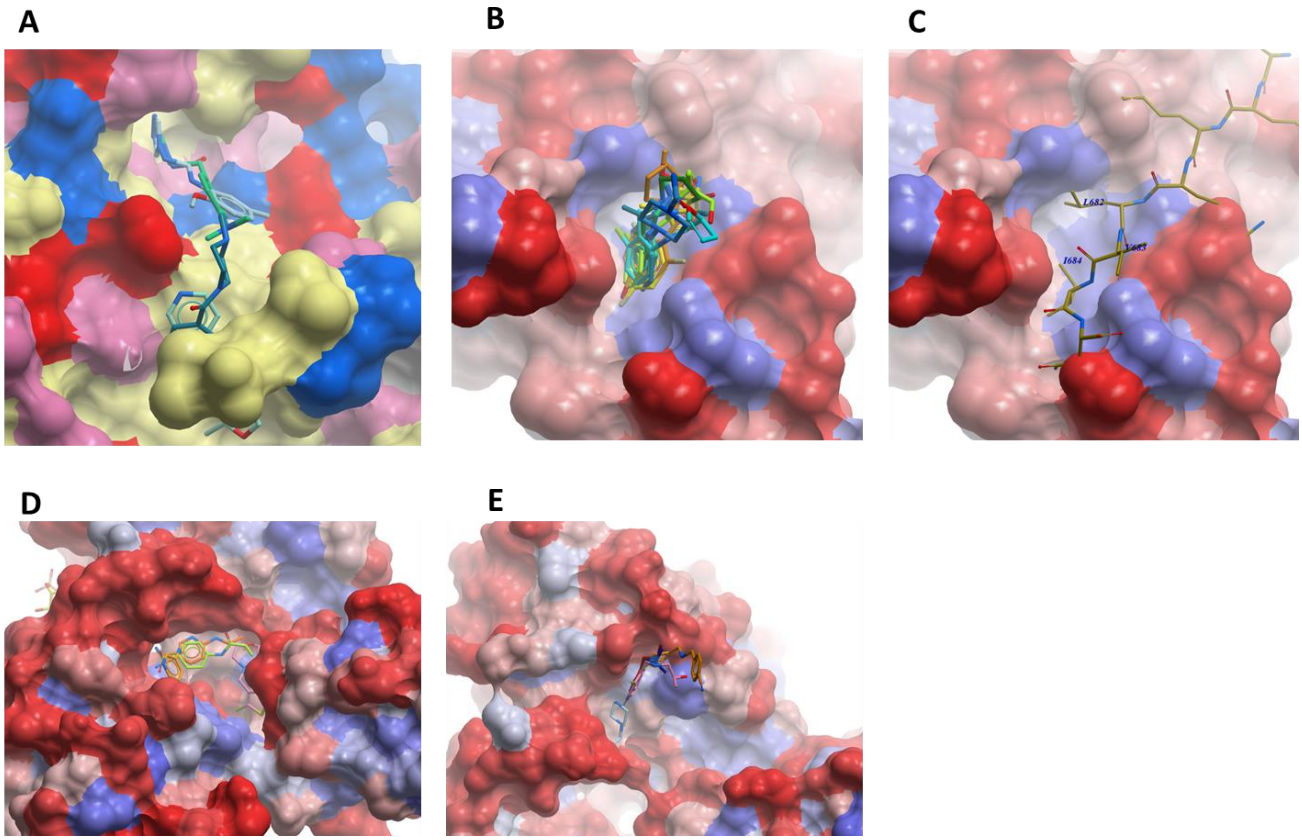
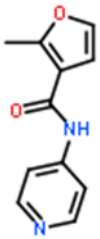
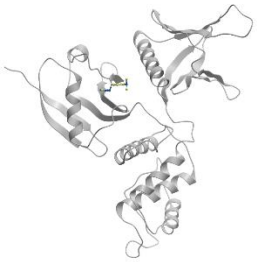
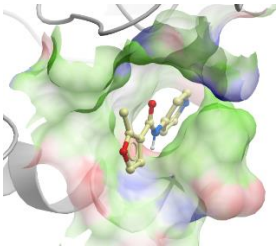
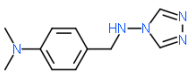
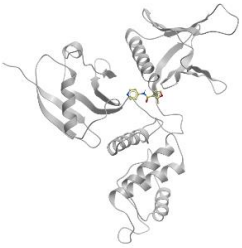
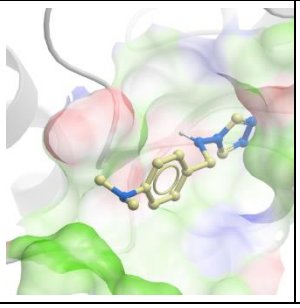
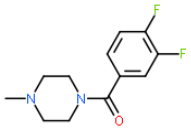
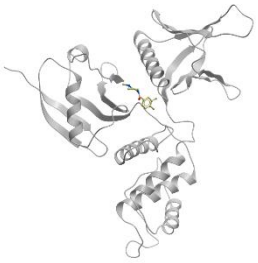
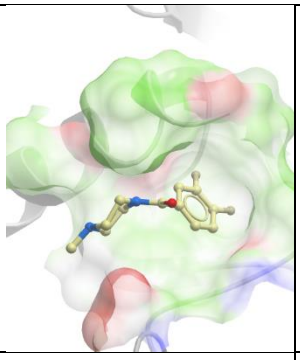
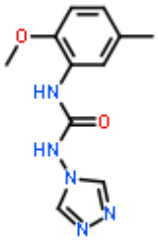
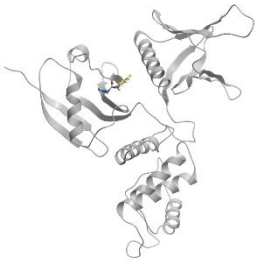
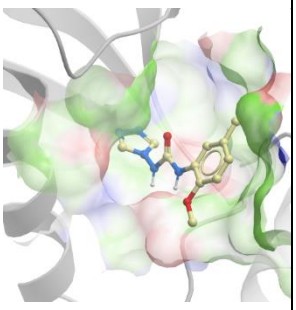
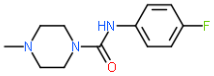
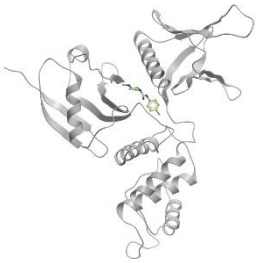
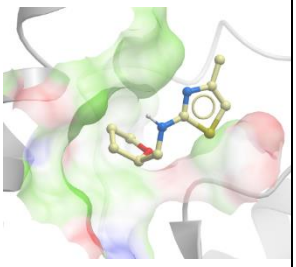
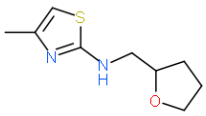
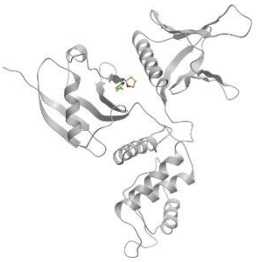
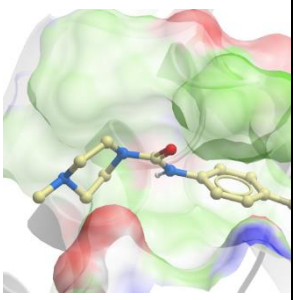
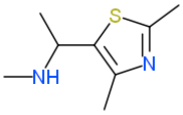
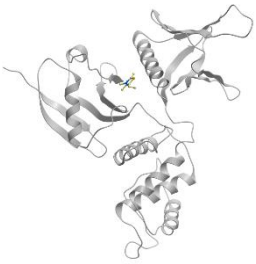
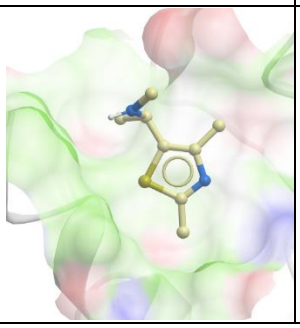
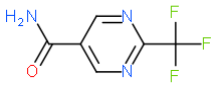
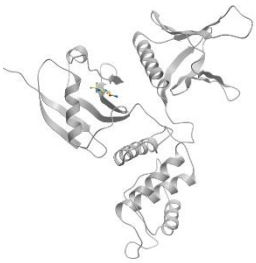
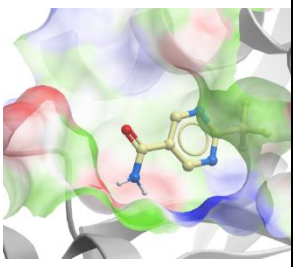
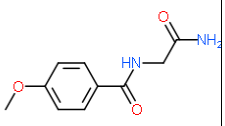
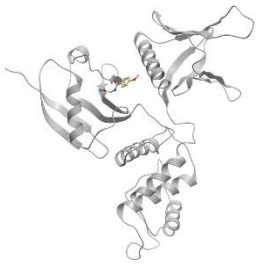
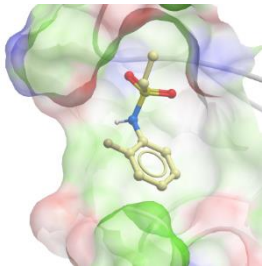
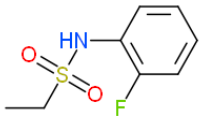
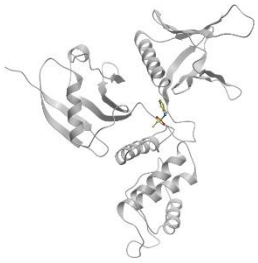
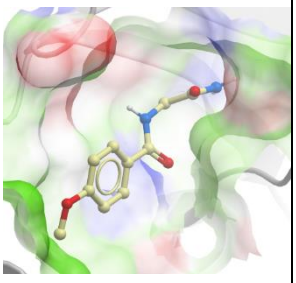
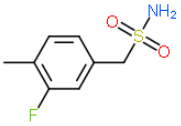
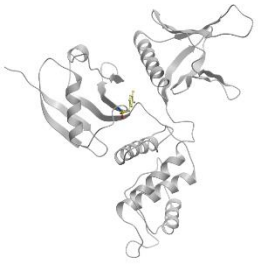
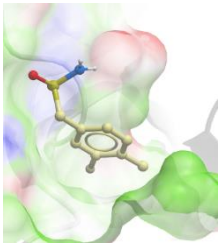
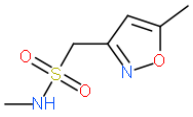
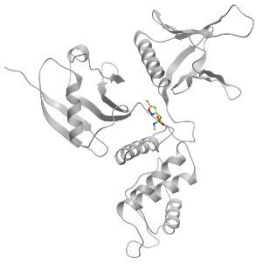
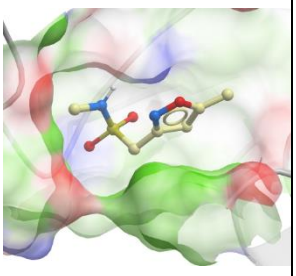


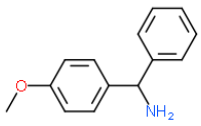
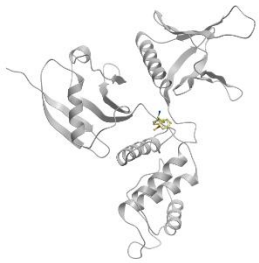
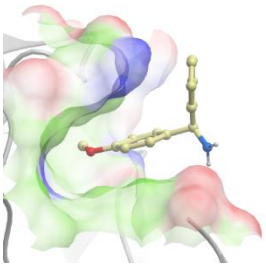
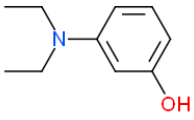
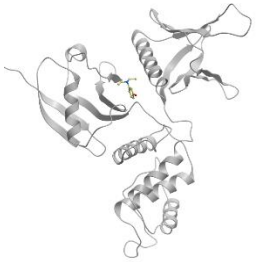
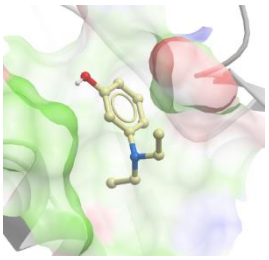
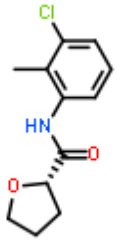
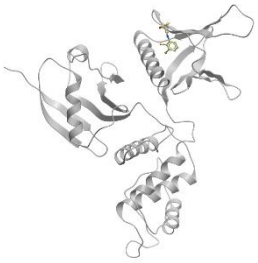
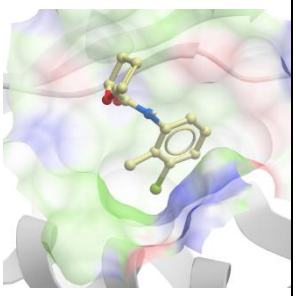
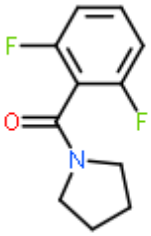
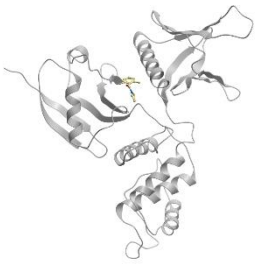
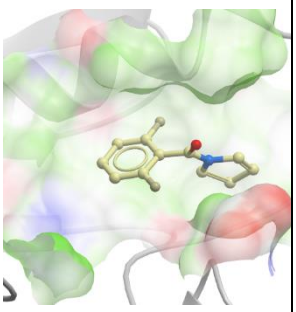
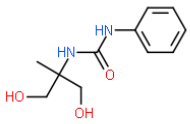
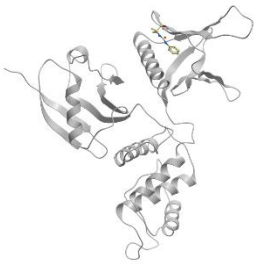
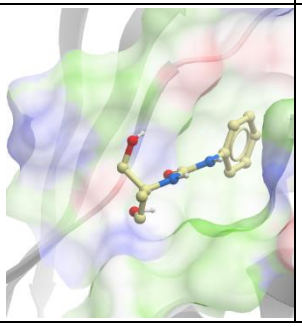
Figure 3. Clusters of bound fragments from crystal-based fragment screens. **(A)** Bound molecules in site 1. **(B)** Bound molecules in site 2. **(C)** Model of a CD44 peptide bound in site 2 (based on the co-crystal structure of MSN, PDB: 6TXS). **(D)** Bound molecules in site 7. **(E)** Bound molecules in site 8.

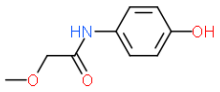
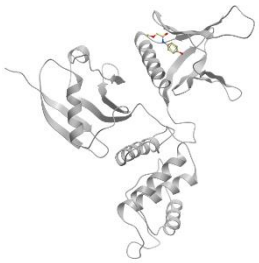
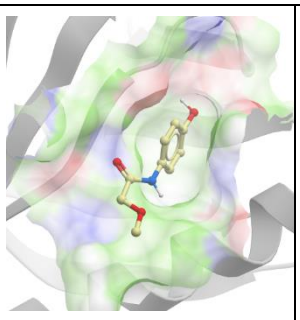
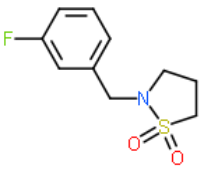
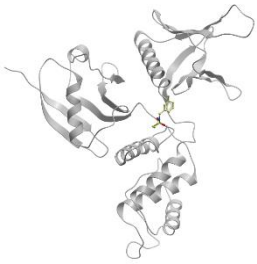
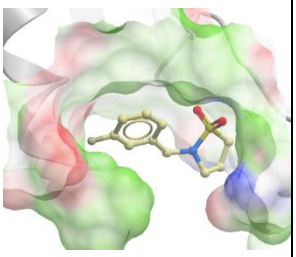
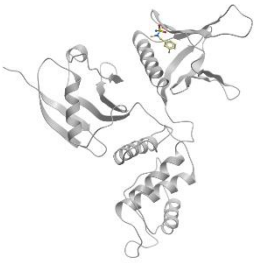
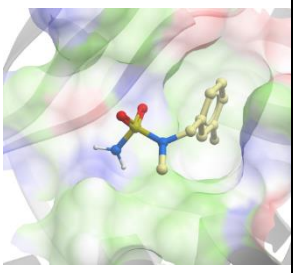
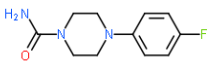
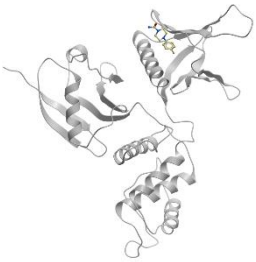
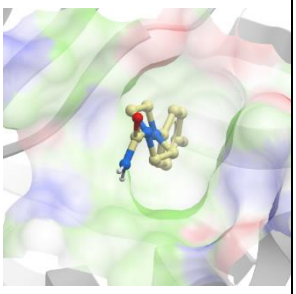
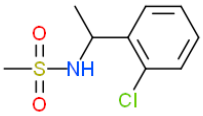
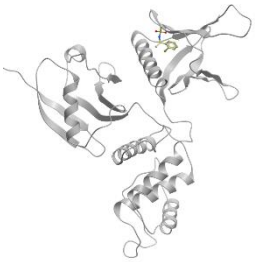
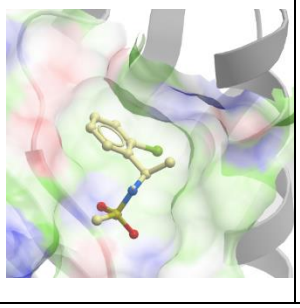
Table 1. Complete list of binding events from fragment screens

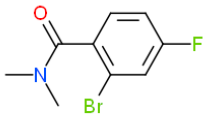
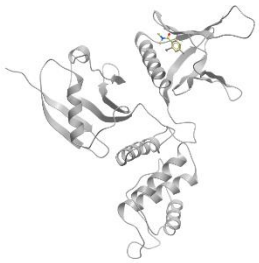
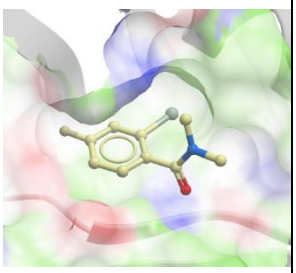
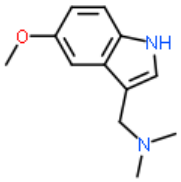
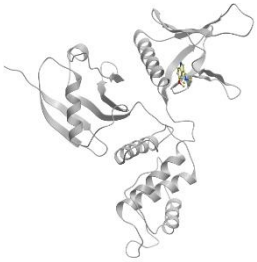
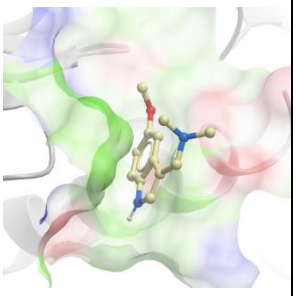
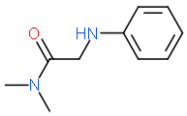
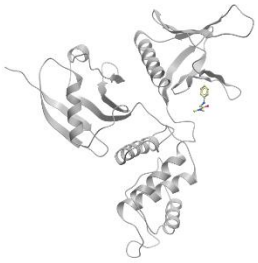
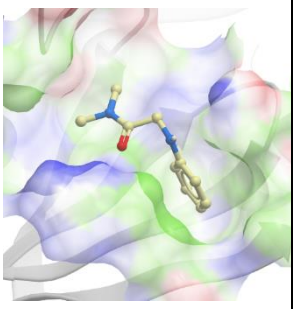
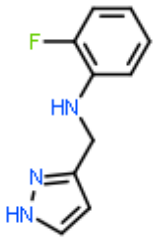
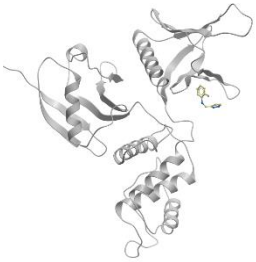
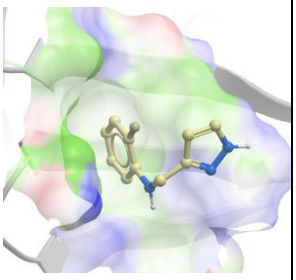
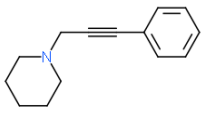
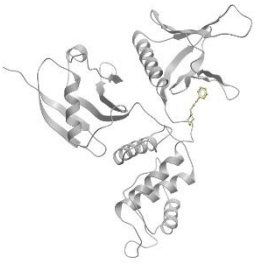
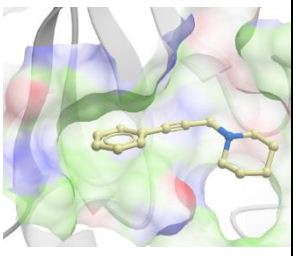
PDB ID	Ligand	Binding Location	Binding pocket	Resolution (Å)
5RYO	x0105 	Site 1 		1.58
5RYQ	x0150 	Site 1 		1.64

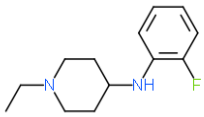
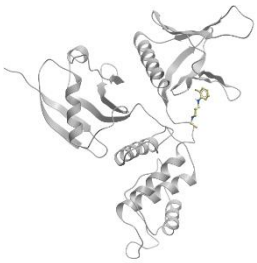
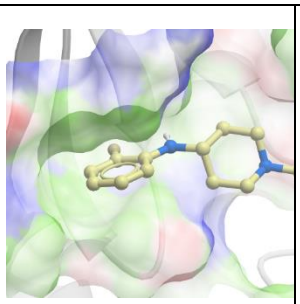
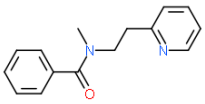
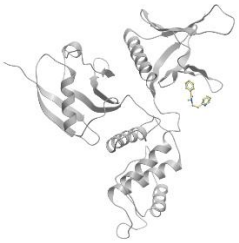
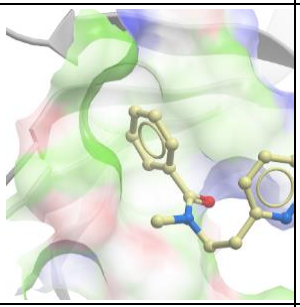
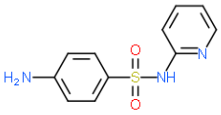
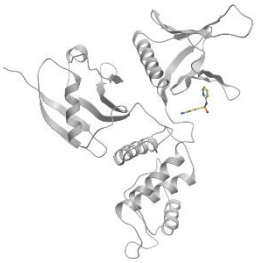
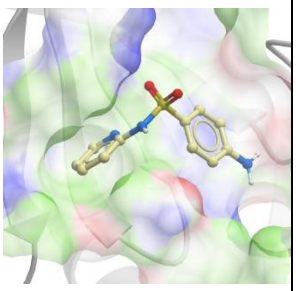
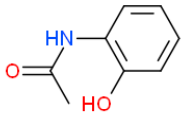
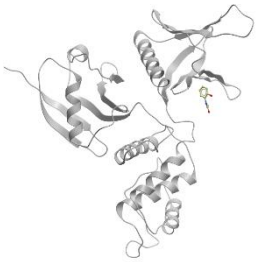
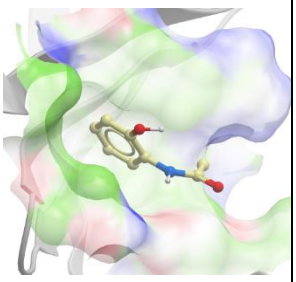
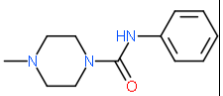
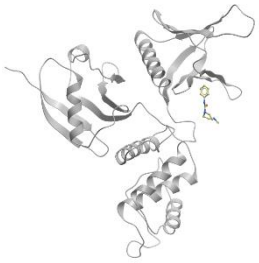
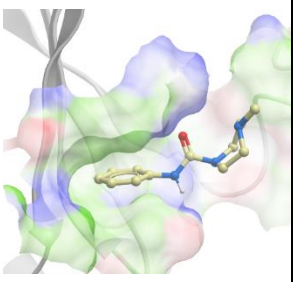
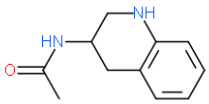
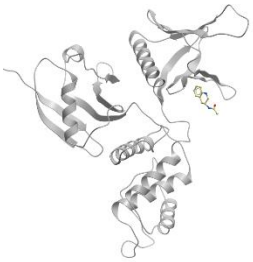
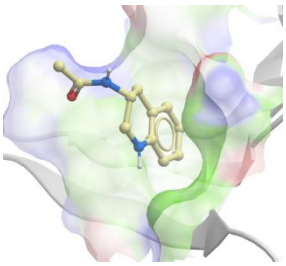
5RYW	x0182 	Site 1 		1.66
5RYX	x0202 	Site 1 		1.63
5RZ6	x0305 	Site 1 		1.64
5RZI	x0463 	Site 1 		2.09
5RZK	x0475 	Site 1 		1.84

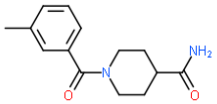
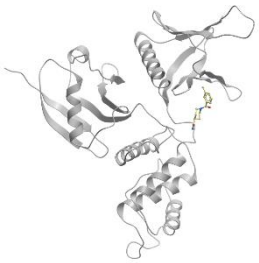
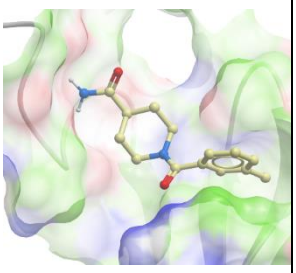
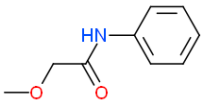
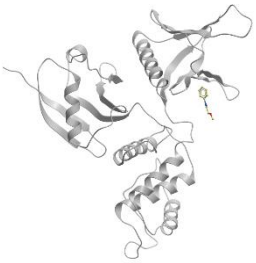
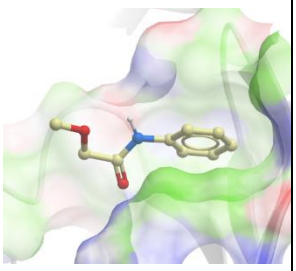
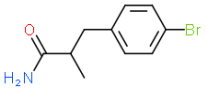
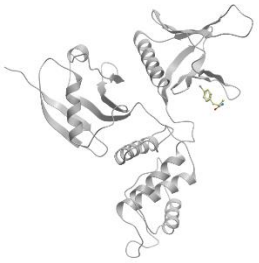
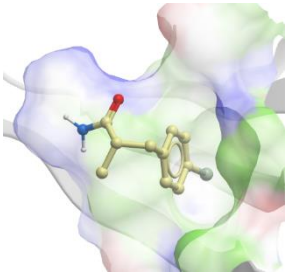
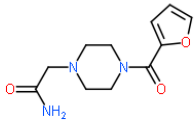
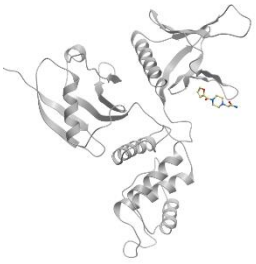
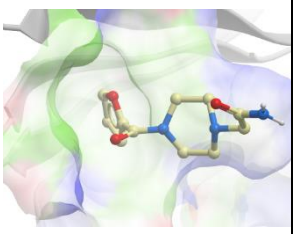
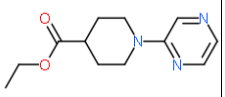
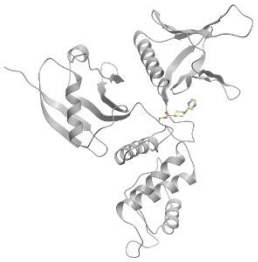
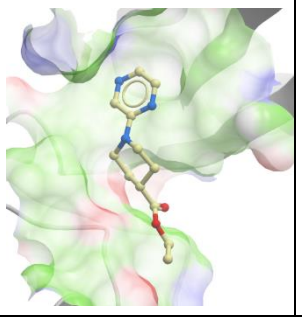
5RZN	x0509 	Site 1 		1.83
5RZP	x0527 	Site 1 		1.70
5RZR	x0544 	Site 1 		1.78
5RZS	x0547 	Site 1 		1.69
5RZT	x0550 	Site 1 		1.79

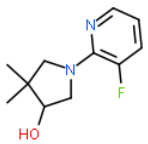
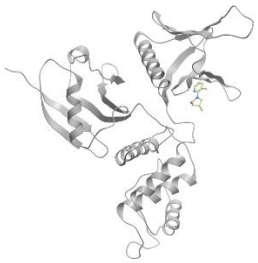
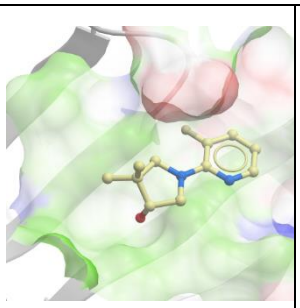
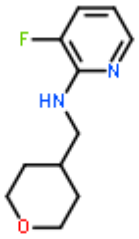
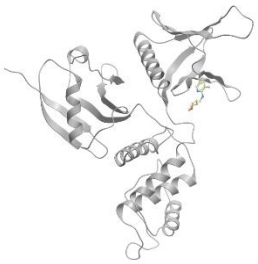
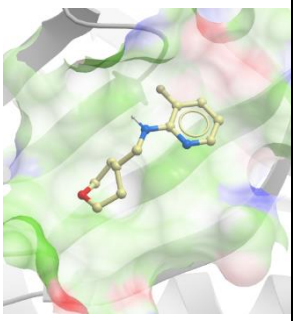
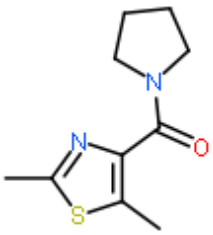
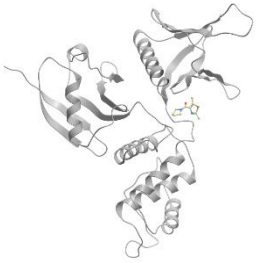
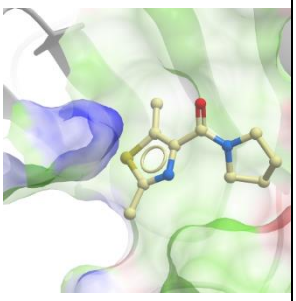
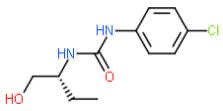
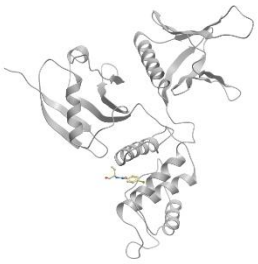
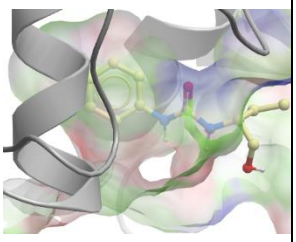
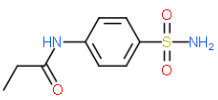
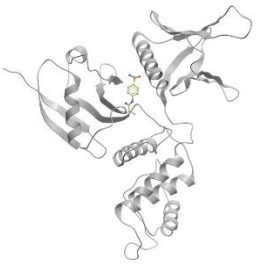
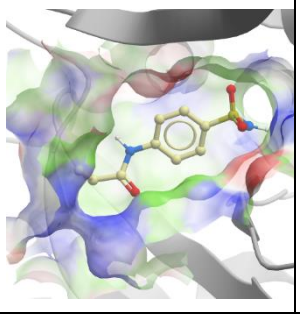
5RZW	x0575 	Site 1 		1.62
5RZX	x0598 	Site 1 		1.74
5RYN	x0104 	Site 2 		1.88
5RYT	x0171 	Site 2 		1.72
5RZ7	x0306 	Site 2 		1.76

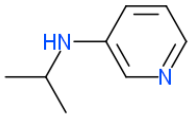
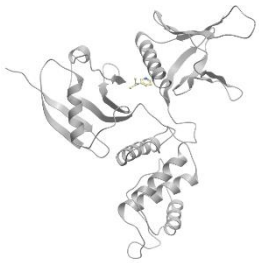
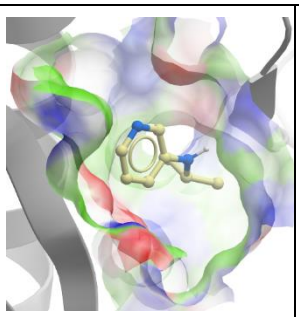
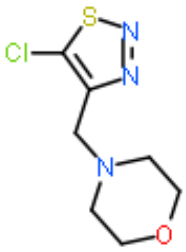
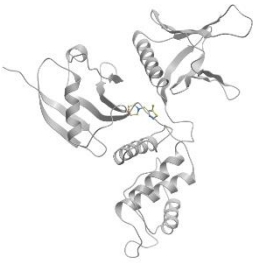
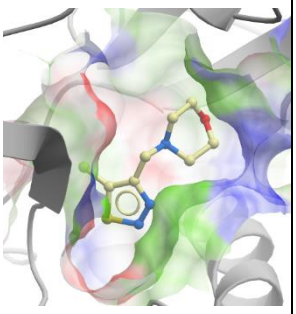
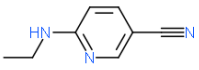
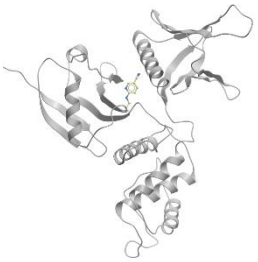
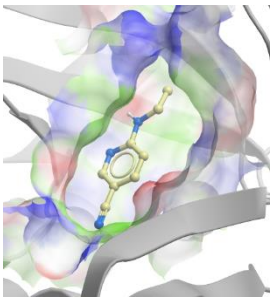
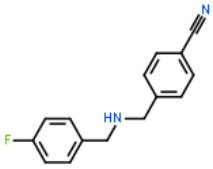
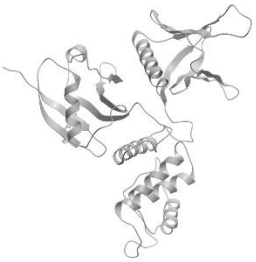
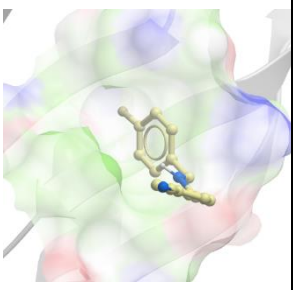
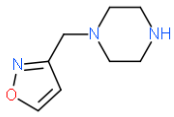
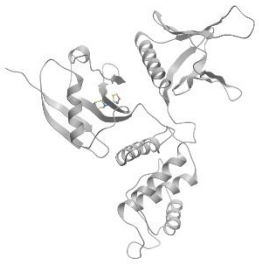
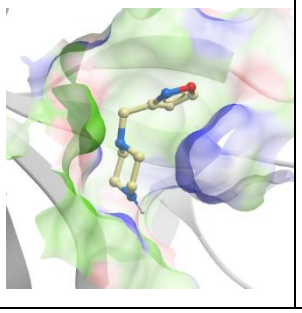
5RZ9	x0325 	Site 2 		1.79
5RZA	x0337 	Site 2 		1.89
5RZC	x0403 	Site 2 		1.75
5RZD	x0413 	Site 2 		1.81
5RZG	x0455 	Site 2 		1.70

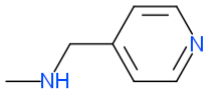
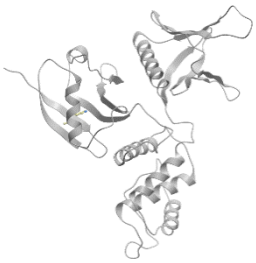
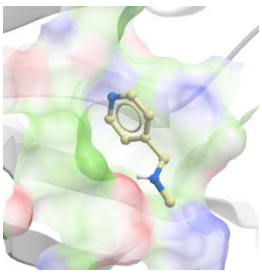
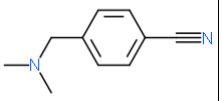
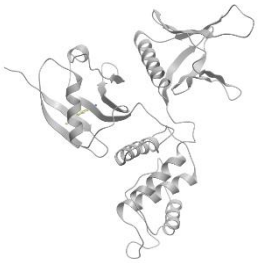
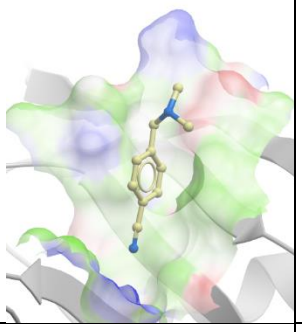
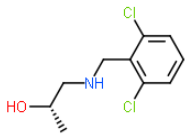
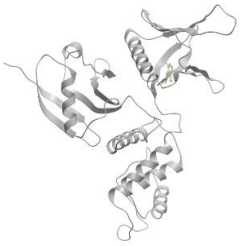
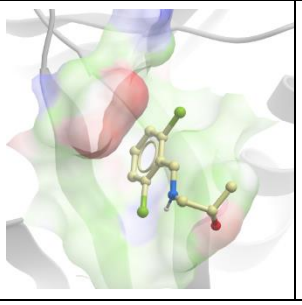
5RZQ	x0538 	Site 2 		1.88
5RZJ	x0473 	Site 3 		1.68
5RYR	x0161 	Site 4 		1.87
5RYU	x0175 	Site 4 		1.63
5RYY	x0209 	Site 4 		1.69

5RYZ	x0212 	Site 4 		1.61
5RZ2	x0244 	Site 4 		1.77
5RZB	x0401 	Site 4 		1.59
5RZF	x0429 	Site 4 		1.76
5RZH	x0457 	Site 4 		1.88
5RZL	x0485 	Site 4 		1.71

5RZU	x0556 	Site 4 		1.66
5RZV	x0558 	Site 4 		1.75
5RZZ	x0608 	Site 4 		1.76
5S00	x0630 	Site 4 		1.77
5RYM	x0076 	Site 5 		1.64

5RYP	x0141 	Site 5 		1.63
5RYV	x0179 	Site 5 		1.69
5RZ0	x0224 	Site 5 		1.74
5RYS	x0162 	Site 6 		1.75
5RZ1	x0235 	Site 7 		1.62

5RZ3	x0249 	Site 7 		1.74
5RZO	x0520 	Site 7 		1.97
5RZY	x0601 	Site 7 		1.75
5RZ5	x0274 	Site 8 		1.63
5RZ8	x0316 	Site 8 		1.66

5RZE	x0419 	Site 8 		1.69
5RZM	x0494 	Site 8 		1.71
5RZ4	x0267 	Sites 5, 8 		1.61

IMPORTANT: Please note that the existence of small molecules within this TEP indicates only that chemical matter might bind to the protein in potentially functionally relevant locations. The small molecule ligands are intended to be used as the basis for future chemistry optimisation to increase potency and selectivity and yield a chemical probe or lead series. As such, the molecules within this TEP should not be used as tools for functional studies of the protein, unless otherwise stated, as they are not sufficiently potent or well-characterised to be used in cellular studies.

CONCLUSION

There is an urgent need to identify new targets and therapeutic modalities for the treatment or prevention of Alzheimer's disease. While many groups have been focussing on a handful of genes with strong GWAS linkage, the AMP-AD consortium has looked at target nominations based on disease-linked RNA and protein expression patterns and network analysis. EPB41L3 is one of the genes predicted to be key drivers on Alzheimer-linked expression modules. The TEP contains novel reagents and data to critically test the utility of EPB41L3 and its associated proteins as targets for treatment of Alzheimer.

TEP IMPACT

The structures, assays and fragment hits provide starting points for the development of small-molecule modulators of EPB41L3 activity.

FUNDING INFORMATION

The work performed at the SGC has been funded by grants from NIH 1RF1AG057443-01 and 1U54AG065187-01 and Wellcome [106169/ZZ14/Z].

ADDITIONAL INFORMATION

Structure Files

PDB ID	Structure Details
6IBE	1.45 Å structure of EPB41L3 (apo)

Materials and Methods

Cloning and expression

Expression constructs

1. **EPB41L3A-c021 (FERM domain, 107-390).** Vector : pNIC28-Bsa4[9] (Genbank EF198106.1, Kanamycin -resistance, IPTG-inducible)

Protein sequence (Tag sequence underlined; * TEV protease cleavage site)

*MHHHHHHSSGVDLGTENLYFQ*SM*PKSMQCKVILLDGSEYTCDVEKRSRGQVLFDKVCEHLNLLEKDYFGLTYRDAENQKN
WLDPAKEIKKQVRSGAWHFSFNVKFYPPDPAQLSEDITRYYLCLQLRDDIVSGRLPCSFVTLALLGSYTVQSELGDYDPDECGS
DYISEFRFAPNHTKELEDKVIELHKSHRGMTPAEAEMHFLENAKKLSMYGVDLHHAKDSEGVEIMLGVCASGLLIYRDRLRIN
RFAWPKVLKISYKRNNFYIKIRPGEFEQFESTIGFKLPNHRAAKRLWKVCVEHHTFFRLL
Predicted mass: 35809.0.

Cleaved protein used in crystallisation:
SMPKSMQCKVILLDGSEYTCDVEKRSRGQVLFDKVCEHLNLLEKDYFGLTYRDAENQKNWLDPAKEIKKQVRSGAWHFSFN
VKFYPPDPAQLSEDITRYYLCLQLRDDIVSGRLPCSFVTLALLGSYTVQSELGDYDPDECGSDYISEFRFAPNHTKELEDKVIELH
KSHRGMTPAEAEMHFLENAKKLSMYGVDLHHAKDSEGVEIMLGVCASGLLIYRDRLRINRFAWPKVLKISYKRNNFYIKIRPG
EFEQFESTIGFKLPNHRAAKRLWKVCVEHHTFFRLL

Predicted mass: 33343.3.

DNA sequence (ORF):

ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCATGCCTAAAAGCAT
GCAGTGCAAAGTGATACTTCTCGATGGATCAGAATATACCTGTGATGTAGAGAAACGCTCCAGAGGACAAGTGCTGTT
TGATAAAGTGTGTGAACACTTGAACTTGCTAGAGAAAGACTACTTTGGGCTTACGTATCGAGATGCTGAAAACCAGAA
GAATTGGTTGGACCCTGCTAAGGAAATAAAAAAACAGGTTCGAAGTGGTGCTTGGCACTTTTCATTTAATGTGAAATTT
TATCCACCAGACCCTGCCCAACTATCTGAAGATATCACCAGGTACTACCTCTGCTTGCAGTTGCGAGATGACATCGTGTC
CGGAAGGCTGCCCTGCTCCTTTGTTACCCTGGCCTTGCTGGGCTCCTACACTGTCCAGTCAGAGCTCGGAGACTATGAC
CCAGATGAATGTGGGAGCGATTACATTAGTGAGTTCCGCTTTGCACCAAACCACACTAAAGAACTGGAAGACAAAGTG
ATCGAGCTGCACAAGAGCCACAGAGGAATGACGCCAGCAGAAGCAGAGATGCATTTCTTGGAAAATGCCAAAAAATTA
TCAATGTATGGGGTAGATTTACATCATGCTAAGGACTCAGAAGGGGTAGAAATTATGTTAGGAGTTTGTGCAAGTGGT
CTGTTGATATATCGCGACCGGCTGCGAATAAACAGATTTGCCTGGCCCAAGGTTCTAAAGATTTCATACAAACGGAACA
ACTTTTACATTAAGATCCGGCCGGGAGAGTTTGAACAATTTGAAAGCACCATTGGGTTTAAGCTGCCAAACCATCGAGC
TGCCAAGCGTTTATGGAAAGTATGTGTTGAGCATCATACATTTTTCAGACTACTGTGA

2. **EPB41L3A-c001 (FERM domain and N-terminal, 1-390).** Vector : pNIC28-Bsa4 (Genbank EF198106.1, Kanamycin -resistance, IPTG-inducible)

Protein sequence: (Tag sequence underlined; * TEV protease cleavage site)

*MHHHHHHSSGVDLGTENLYFQ*SM*TTESGSDSESKPDQEAEPQEAAGAQGRAGAPVPEPPKEEQQQALEQFAAAAAHST
PVRREVTDKEQEFAARAAKQLEYQQLEDDKLSQKSSSSKLSRSPLKIVKKPKSMQCKVILLDGSEYTCDVEKRSRGQVLFDKVC
EHLNLLEKDYFGLTYRDAENQKNWLDPAKEIKKQVRSGAWHFSFNVKFYPPDPAQLSEDITRYYLCLQLRDDIVSGRLPCSFV
TLALLGSYTVQSELGDYDPDECGSDYISEFRFAPNHTKELEDKVIELHKSHRGMTPAEAEMHFLENAKKLSMYGVDLHHAKD
SEGVEIMLGVCASGLLIYRDRLRINRFAWPKVLKISYKRNNFYIKIRPGEFEQFESTIGFKLPNHRAAKRLWKVCVEHHTFFRLL
Predicted mass: 47167.4

Protein sequence after tag cleavage:
SMTTESGSDSESKPDQEAEPQEAAGAQGRAGAPVPEPPKEEQQQALEQFAAAAAHSTPVRREVTDKEQEFAARAAKQLEY
QQLEDDKLSQKSSSSKLSRSPLKIVKKPKSMQCKVILLDGSEYTCDVEKRSRGQVLFDKVCEHLNLLEKDYFGLTYRDAENQKN
WLDPAKEIKKQVRSGAWHFSFNVKFYPPDPAQLSEDITRYYLCLQLRDDIVSGRLPCSFVTLALLGSYTVQSELGDYDPDECGS
DYISEFRFAPNHTKELEDKVIELHKSHRGMTPAEAEMHFLENAKKLSMYGVDLHHAKDSEGVEIMLGVCASGLLIYRDRLRIN
RFAWPKVLKISYKRNNFYIKIRPGEFEQFESTIGFKLPNHRAAKRLWKVCVEHHTFFRLL

Predicted mass: 44701.7

DNA sequence (ORF):
ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCATGACGACCGAATC
TGGATCAGACTCGGAATCCAAGCCGGACCAGGAGGCCGAGCCCCAGGAGGCGGCGGGGGCGCAGGGGCGCGCGGG
GGCGCCCGTGCCGGAGCCGCCCAAGGAGGAGCAGCAGCAGGCCCTGGAGCAGTTCGCCGCCGCTGCAGCGCACAGCA
CCCCGGTGCGGAGGGAGGTCACTGACAAGGAACAGGAGTTTGCTGCCAGGGCTGCAAAACAGCTCGAATATCAGCAA
TTAGAAGACGATAAACTTTCTCAGAAATCATCTAGCAGTAAACTCTCTCGGTCTCCATTAAAGATTGTCAAAAAGCCTAA
AAGCATGCAGTGCAAAGTGATACTTCTCGATGGATCAGAATATACCTGTGATGTAGAGAAACGCTCCAGAGGACAAGT
GCTGTTTGATAAAGTGTGTGAACACTTGAACTTGCTAGAGAAAGACTACTTTGGGCTTACGTATCGAGATGCTGAAAAC
CAGAAGAATTGGTTGGACCCTGCTAAGGAAATAAAAAAACAGGTTCGAAGTGGTGCTTGGCACTTTTCATTTAATGTGA
AATTTTATCCACCAGACCCTGCCCAACTATCTGAAGATATCACCAGGTACTACCTCTGCTTGCAGTTGCGAGATGACATC
GTGTCCGGAAGGCTGCCCTGCTCCTTTGTTACCCTGGCCTTGCTGGGCTCCTACACTGTCCAGTCAGAGCTCGGAGACT
ATGACCCAGATGAATGTGGGAGCGATTACATTAGTGAGTTCCGCTTTGCACCAAACCACACTAAAGAACTGGAAGACA
AAGTGATCGAGCTGCACAAGAGCCACAGAGGAATGACGCCAGCAGAAGCAGAGATGCATTTCTTGGAAAATGCCAAA
AAATTATCAATGTATGGGGTAGATTTACATCATGCTAAGGACTCAGAAGGGGTAGAAATTATGTTAGGAGTTTGTGCAA
GTGGTCTGTTGATATATCGCGACCGGCTGCGAATAAACAGATTTGCCTGGCCCAAGGTTCTAAAGATTTCATACAAACG
GAACAACTTTTACATTAAGATCCGGCCGGGAGAGTTTGAACAATTTGAAAGCACCATTGGGTTTAAGCTGCCAAACCAT
CGAGCTGCCAAGCGTTTATGGAAAGTATGTGTTGAGCATCATACATTTTTCAGACTACTGTGA

Protein expression and Purification (Both constructs)

Protein Expression

The plasmids were transformed in the *E. coli* strain BL21(DE3)-R3-pRARE, a phage-resistant variant of Rosetta 2 (MSD). Then, plated on LB-agar plates containing kanamycin (50 µg/ml) and chloramphenicol, (34 µg/ml). Several colonies were picked together, and used to inoculate liquid cultures in the same medium; after overnight incubation, this was stored at -80°C after the addition of 15% (v/v) glycerol.

For expression, we inoculated an overnight culture of LB+kan+chlp at 37°C. Then, 10 ml of the overnight culture was used to inoculate a 1L culture containing Terrific Broth (TB) with kanamycin only. The cultures were grown at 37°C with vigorous aeration in 2.5L Tunair flasks until reaching OD600 of between 1.5-3. Shift the cultures to 18°C; after 30 minutes, add 0.3 mM IPTG (from a 1.0M stock) and continued incubation for 16 hours at 18°C.

Cells were harvested by centrifugation (JLA8.1000 rotor, 4000 RPM, 25 min). The medium was safely discarded and the cell pellets scraped and collected with a rubber spatula into 50-ml tubes, which were frozen and kept at -80°C.

Purification

Buffers

- Lysis buffer: 50 mM HEPES (pH 7.5), 500 mM NaCl, 10 mM imidazole, 5% glycerol, 1 mM TCEP
- W30 Buffer: 50 mM HEPES (pH 7.5), 500 mM NaCl, 30 mM imidazole, 5% glycerol, 1 mM TCEP
- Elution Buffer (EB): 50 mM HEPES (pH 7.5), 500 mM NaCl, 300 mM imidazole, 5% glycerol, 1 mM TCEP
- SEC buffer: 50 mM HEPES (pH 7.5), 500 mM NaCl, 5% glycerol, 1 mM TCEP
- Ni-sepharose beads, equilibrated in Lysis buffer

Procedure

1. Thaw the cell pellet and suspend in 80 ml/Litre of culture of Lysis buffer. Lyse the cells by sonication on ice (20 min, 5s on, 10s off, 35% amplitude) with occasional stirring.
2. Centrifuge the lysate (25 min, 67000 g). Decant the supernatant carefully. Save 0.1 ml for analysis,
3. Add 0.6 ml of Ni-sepharose beads in 50-ml falcon tubes. Mix by rotation for 1 hr at <7°C.
4. Spin 700g/5 min/4°C. Decant lysate (FT) and wash beads with 100 ml LB. Spin, decant (W1), and wash pellets with 50 ml lysis buffer. Spin, decant (W2), and add 1 ml LB. Transfer beads to gravity column in cold room.
5. Wash column with 20 ml W30 (keep W30 eluate).
6. Elute protein with 3x 10 ml EB (E1, E2, E3).
7. If the protein is to be used for crystallisation, the N-terminal tag is cleaved using TEV protease. Skip this step if the tag is to be retained. The protein is combined in a dialysis tube with His-tagged TEV protease at a 1:20 mass ratio (TEV : EPB41L3) and placed in 1-2 litres of SEC buffer, at 4°C overnight. Then, pass the protein solution through a Ni-Sepharose column equilibrated with SEC buffer using gravity flow. Wash the beads successively with 10 ml each of Lysis buffer, W30 buffer, and Elution buffer and collect each effluent. Analyze by gel electrophoresis and/or intact MS to locate the cleaved protein.
8. Concentrate the protein (cleaved or tagged) to <1 ml using a centrifugal concentrator with MWCO of 30 kDa.
9. Purify the protein further by Size-exclusion chromatography (SEC) on a HiLoad Superdex S200 HR 16/60 column in SEC buffer at 1 ml/min. Identify the fractions containing pure EPB41L3 protein by SDS-PAGE, pool and concentrates as required. Snap-freeze in thin-walled PCR tubes in liquid N2, and store at -80°C.

Crystallisation

Structure 6IBE: 20% PEG Smear Medium -- 0.1M MES pH 6.0 -- 0.15M ammonium nitrate -- 5%(v/v) ethylene glycol

Fragment screen: 0.1M MES pH 6.0 -- 0.075M ammonium nitrate -- 26% PEG Smear Medium -- 10% ethylene glycol with seeding

Soaks for X-ray fragment screens were performed at Diamond Light Source using the standard X-Chem protocol with fragments at 100 mM (20% DMSO). The crystals were harvested after 2.5 to 5 hours.

Data collection and processing: Data were collected to 1.45 Å on beamline I03 at Diamond Light Source. The data were integrated with Dials and scaled with Aimless. The structure was determined by molecular replacement with Phaser using an earlier EPB41L3 structure **2HE7** as a model and refined using Refmac to R / R_{free} of 15.4/20.0%

Assays

Reagents

Protein: EPB41L3-c001 (14.9 mg/ml; 316 µM)
Peptide: biotin-SRRRCGQKKKLVINSGNGAVEDY (b-CD44[672-691];10 mM in water)
competitor peptide: SRRRCGQKKKLVINSGNGAVEDY (CD44[672-691])

Assay Buffer (AB): 25 mM HEPES (pH 7.5), 100 mM NaCl, 0.1% BSA, 0.05% Tween-20. Filter.
Donor Reagent: LANCE Eu-W1024 Anti-6xHis (Perkin Elmer AD0205; 10ug) at 0.625 µM.

Acceptor Reagent: Streptavidin-XL665 (Cisbio 610SAXLF; 1000 tests) at 200 ug/ml or 3.3 µM.

Procedure for titration of competitor, unlabelled peptide

1. Dilute protein to 4x concentration in AB (8 nM):
2. Dilute non-biotinylated CD44(672-691) to 4x concentration in AB. Do a 16-point dilution series from 32 µM to 1 nM final.
3. Dilute b-CD44(672-691) to 4x concentration in AB (240 nM for 60nM final).
4. Dilute Donor and Acceptor to 4x concentration in AB: Eu-anti-6His to 1 nM final, SA-XL665 to 10 nM final.
5. Add 5 µl of protein to wells.
6. Add 5 µl of CD44(672-691) to wells. Incubate 30 min at RT.
7. Add 5 µl of b-CD44(672-691) to wells. Incubate 30 min at RT.
8. Add 5 µl of diluted donor and acceptor (4x). Spin plate. Incubate for 2 h.
9. Read on PheraStar FSX at 620 and 660 nm.

References

1. Jiang, W. and I.F. Newsham, *The tumor suppressor DAL-1/4.1B and protein methylation cooperate in inducing apoptosis in MCF-7 breast cancer cells.* Mol Cancer, 2006. **5**: p. 4.
2. Singh, V., et al., *DAL-1/4.1B tumor suppressor interacts with protein arginine N-methyltransferase 3 (PRMT3) and inhibits its ability to methylate substrates in vitro and in vivo.* Oncogene, 2004. **23**(47): p. 7761-71.
3. Tran, Y.K., et al., *A novel member of the NF2/ERM/4.1 superfamily with growth suppressing properties in lung cancer.* Cancer Res, 1999. **59**(1): p. 35-43.
4. Baines, A.J., H.C. Lu, and P.M. Bennett, *The Protein 4.1 family: hub proteins in animals for organizing membrane proteins.* Biochim Biophys Acta, 2014. **1838**(2): p. 605-19.
5. Cifuentes-Diaz, C., et al., *Protein 4.1B contributes to the organization of peripheral myelinated axons.* PLoS One, 2011. **6**(9): p. e25043.
6. Mori, T., et al., *Structural basis for CD44 recognition by ERM proteins.* J Biol Chem, 2008. **283**(43): p. 29602-12.
7. Buttermore, E.D., et al., *The cytoskeletal adaptor protein band 4.1B is required for the maintenance of paranodal axoglial septate junctions in myelinated axons.* J Neurosci, 2011. **31**(22): p. 8013-24.
8. Einheber, S., et al., *The 4.1B cytoskeletal protein regulates the domain organization and sheath thickness of myelinated axons.* Glia, 2013. **61**(2): p. 240-53.
9. Savitsky, P., et al., *High-throughput production of human proteins for crystallization: the SGC experience.* J Struct Biol, 2010. **172**(1): p. 3-13.

We respectfully request that this document is cited using the DOI value as given above if the content is used in your work.