Supporting Information

PDZ domains as drug targets

*Nikolaj R. Christensen, Jelena Čalyševa, Eduardo F. A. Fernandez, Susanne Lüchow, Louise S, Clemmensen, Linda M. Haugaard-Kedström, Kristian Strømgaard**

Supporting methodology

All human PDZ domains were collected from UniProt database and filtered to remove redundancy. To produce a sequence alignment, Chimera MatchMaker structural alignment tool was used on all available human PDZ domain structures bound to a canonical ligand; the parts of the structure aligned poorly, mostly loops, were aligned using Clustal Omega sequence alignment tool. All human PDZ domains were then aligned to this reference alignment using Clustal Omega. The domain starting and ending positions were adjusted according to the sequence alignment. Sequences for GORASP1 and GORASP2 PDZ domains were adjusted to follow the secondary structure of other PDZ domains: first and second halves of the sequences were switched. The resulting alignment was then visualized using Jalview software. After careful manual curation of the resulting alignment and secondary structure prediction using Jpred software, several patterns in PDZ domains were visible. A strong Ψ-G-Ψ binding site motif, described above, was present in most of the proteins. For more than half of the PDZ domains, the motif followed the canonical LGF sequence. Throughout the alignment, noticeable conservation was showing for some regions, while others were similar for some PDZ domains and completely different for others. The overall similarity of the PDZ domain sequences lead to a very conserved domain structure and similar molecular function; the subtle differences, however, must contribute to the specificity of particular domains and their differences.

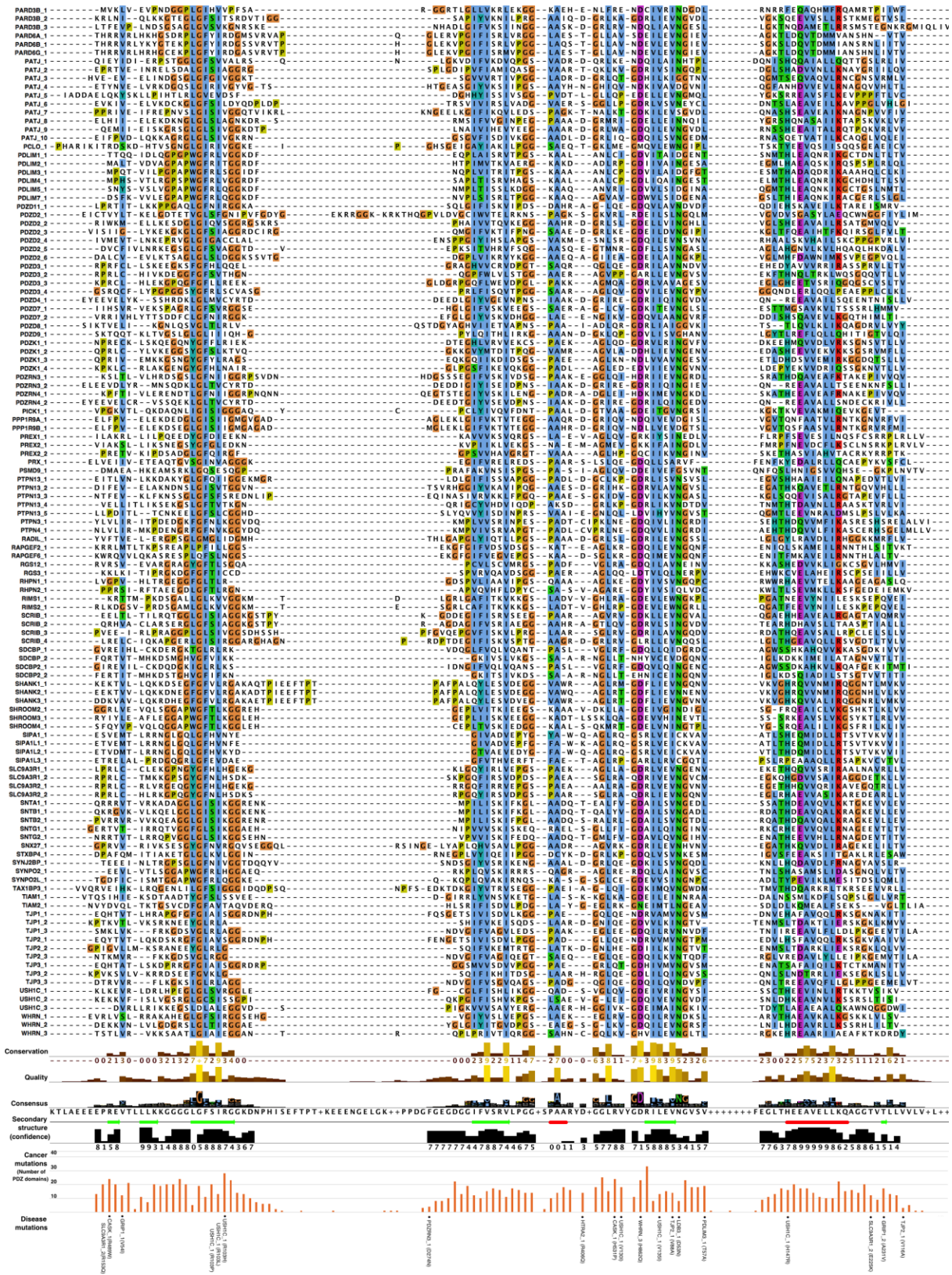
Conservation
Quality
Consensus
Secondary structure (confidence)
Cancer mutations (Number of PDZ domains)
Disease mutations

Figure S1.

The sequence alignment of all human PDZ domains. Sequence alignment and conservation and alignment quality measures, the best estimate of the sequence consensus, secondary structure prediction (Jpred), numbers of cancer-causing mutations and the disease causing mutations. Alignment characteristics calculated and visualised using Jalview software. To produce the alignment, Chimera MatchMaker structural alignment tool was used on all available human PDZ domain structures bound to a canonical ligand; the parts of the structure aligned poorly, mostly loops, were aligned using Clustal Omega sequence alignment tool. All human PDZ domains were then aligned to this reference alignment using Clustal Omega. The domain starting and ending positions were adjusted according to the sequence alignment. Sequences for GORASP1 and GORASP2 PDZ domains were adjusted to follow the secondary structure of other PDZ domains: first and second halves of the sequences were switched.

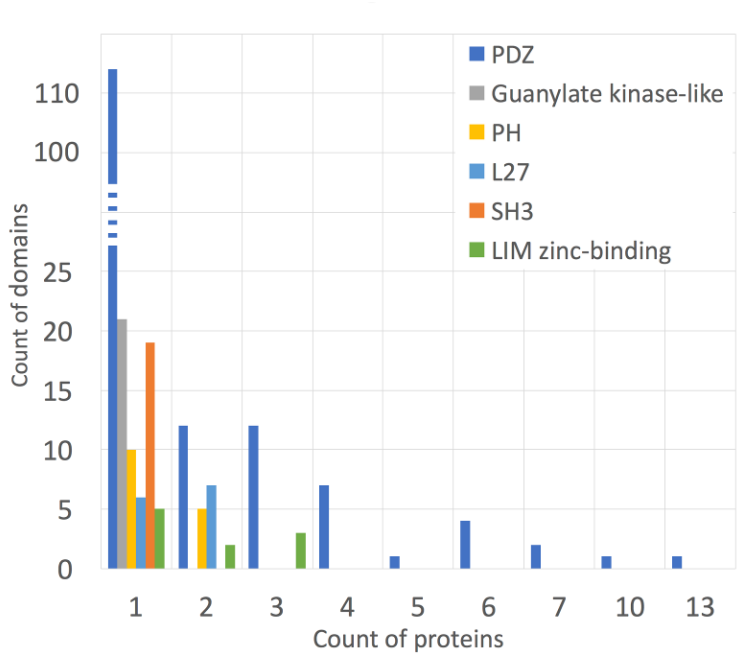

Figure S2.

Domain composition of the human PDZ-containing proteins. For each of the human PDZ domain-containing proteins, the number of different domains was calculated. Only data for the 6 most common domains within these proteins shown.

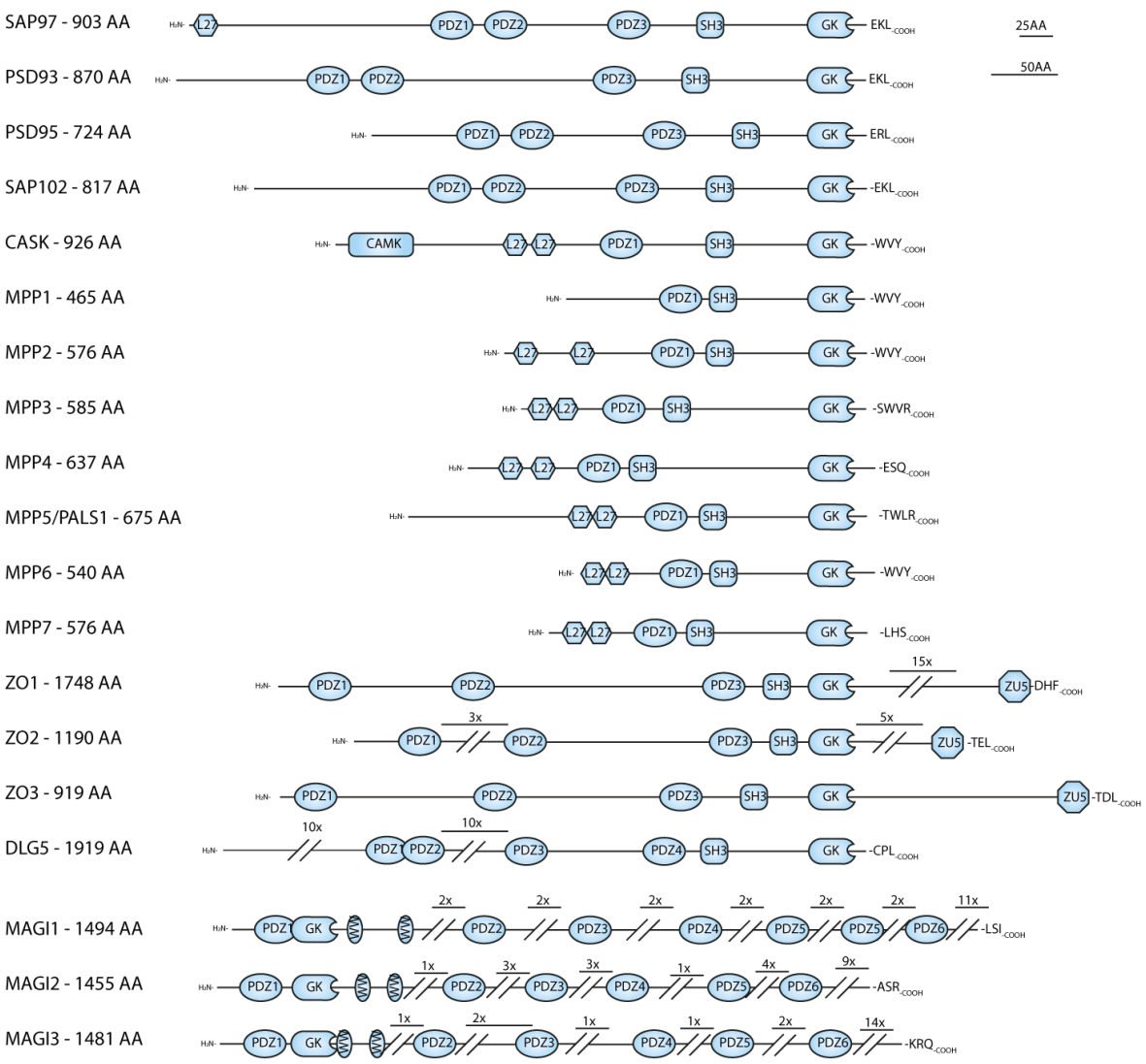

Figure S3. Domain organization of selected MAGUK proteins.

Table S1.

Table S1_JC_all_pdz_index_Structure details.xlsx

The summary of all human PDZ domains collected from UniProt database and filtered to remove redundancy. PDB codes for available structures are included. For the PDZ domains without available structures, the structures with highest sequence similarity are included.

Table S2.

Table S2_reactome_JC.xlsx

Extensive summary of Reactome pathway participation of PDZ-containing proteins

Table S3.

Table S3 diseases_JC.xlsx

The summary of human disease associations with PDZ-containing proteins and the genes encoding them and disease-causing mutations within the PDZ domains collected from gene/protein-disease association databases

Table S4.

Table S4 BindingDB_binders_sup.xlsx

PDZ-containing proteins in the BindingDB database of measured binding affinities for small drug-like molecules