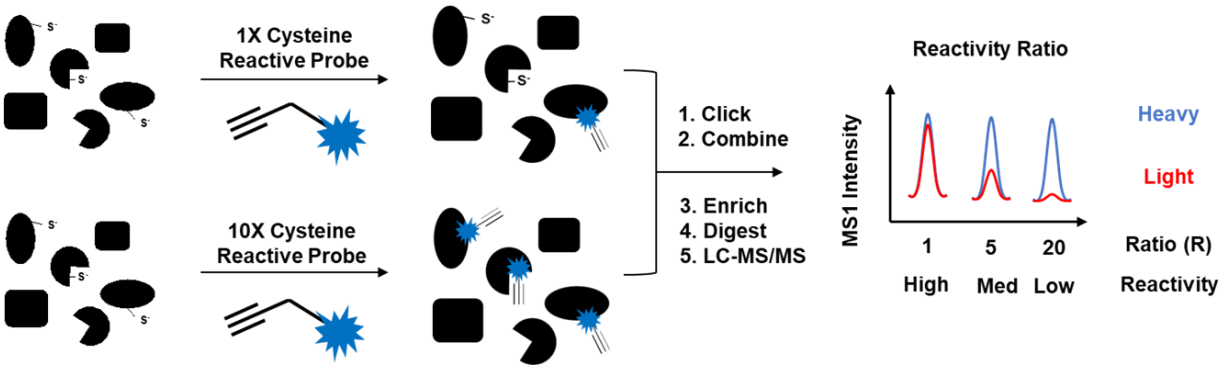

Figure S1. General chemoproteomics workflow for measuring intrinsic cysteine-reactivity towards electrophilic probes such as iodoacetamide alkyne (IAA), related to **Figure 1**. For these intrinsic reactivity studies, including those that use the Isotopic Tandem Orthogonal Proteolysis–ABPP (isoTOP–ABPP) platform, relative cysteine labeling by a high (10x) and low (1x) concentration of IAA or other probes are compared using isotopically labeled enrichment handles and MS1-based quantification. Hyper-reactive residues are those that show $R_{10:1}$ ratios close to 1, indicating saturation of labeling at the lower reagent concentration.

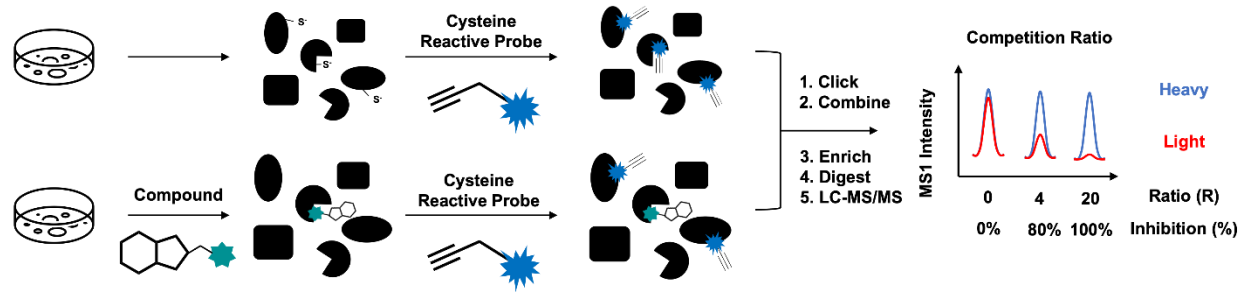

Figure S2. General chemoproteomics workflow for measuring cysteine ligandability using competitive isoTOP-ABPP and related methods, related to **Figure 1**. Proteomes are treated with electrophilic compounds or vehicle (DMSO), labeled with an iodoacetamide (IA)-alkyne probe, and conjugated to isotopically-differentiated, biotin enrichment handles by click chemistry. Treated and control samples are combined, processed, and analyzed by LC-MS/MS, where the isotopic label is used to distinguish between peptides from control and fragment-treated samples, with elevated $R_{H:L}$ ratios indicative of a liganded cysteine.

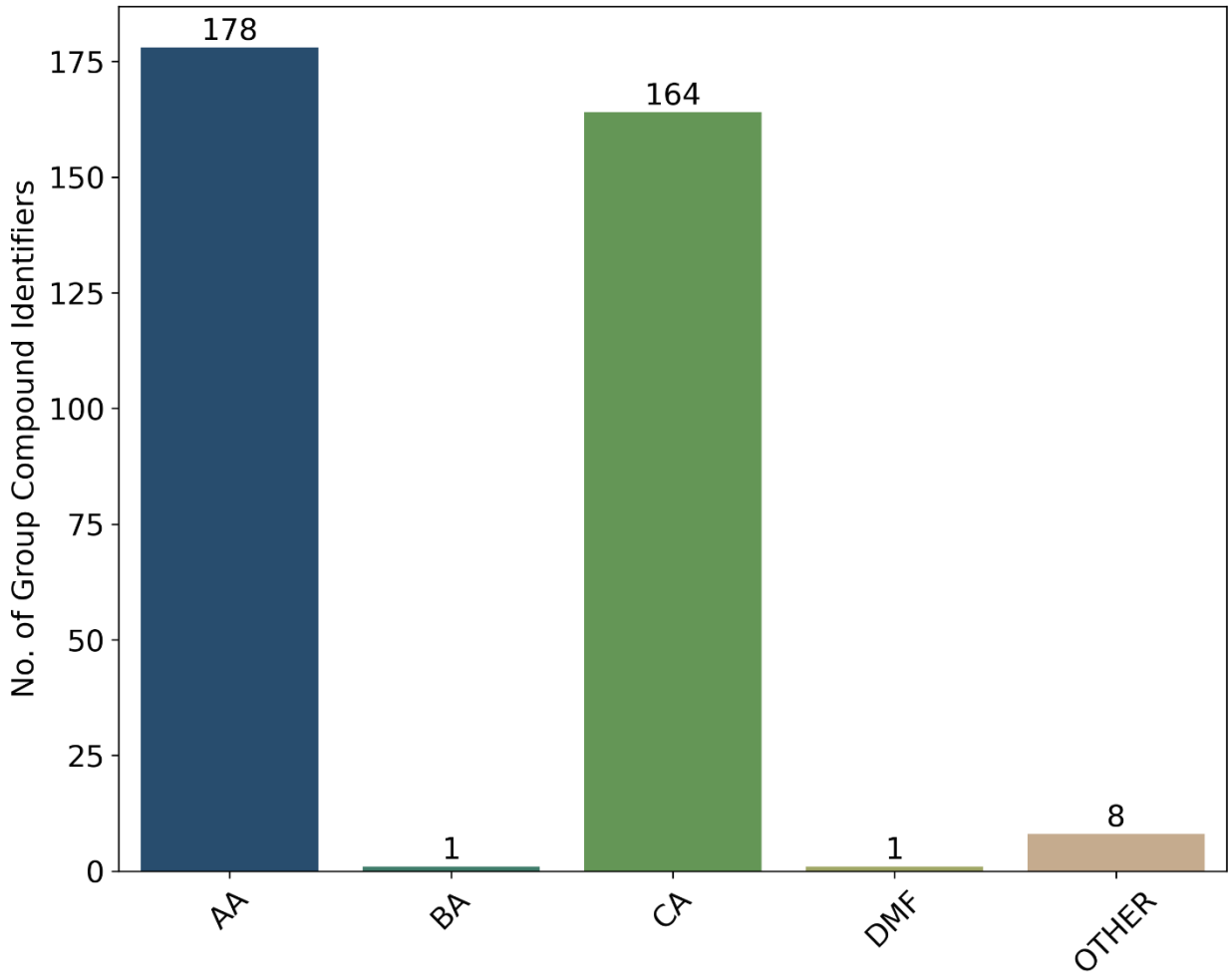

Figure S3. Total number of unique compounds for each warhead category in CysDB Lig, related to **Figure 1**: acrylamide (AA), bromoacetamide (BA), chloroacetamide (CA), dimethyl fumarate (DMF) and other (OTHER). Unique compounds were determined by SMILES strings/group compound identifiers.

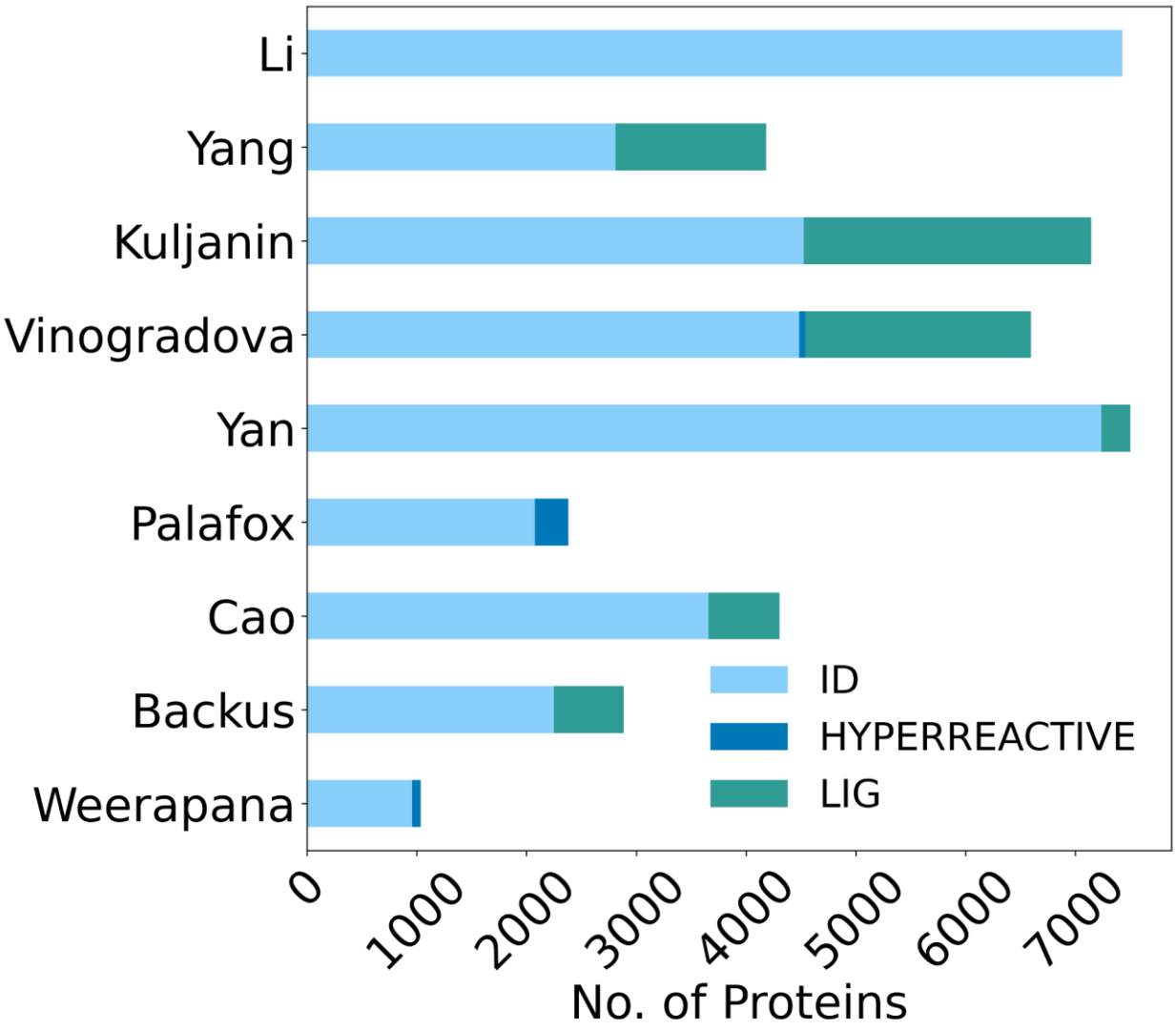

Figure S4. Total number proteins for each category in CysDB & in UniProtKB/Swiss-Prot, related to **Figure 1**.

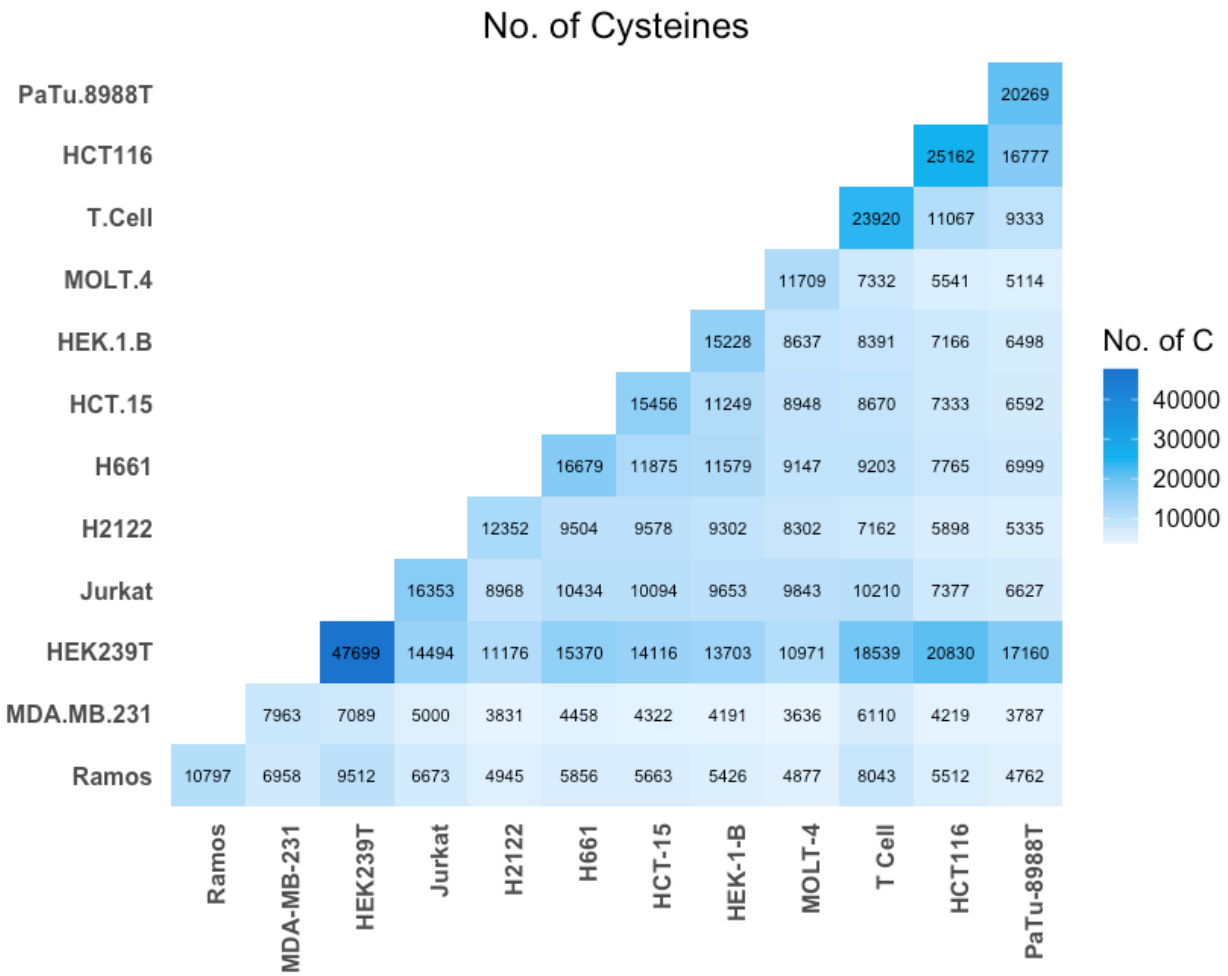

Figure S5. Number of identified cysteines shared between different cell lines, related to **Figure 1**.

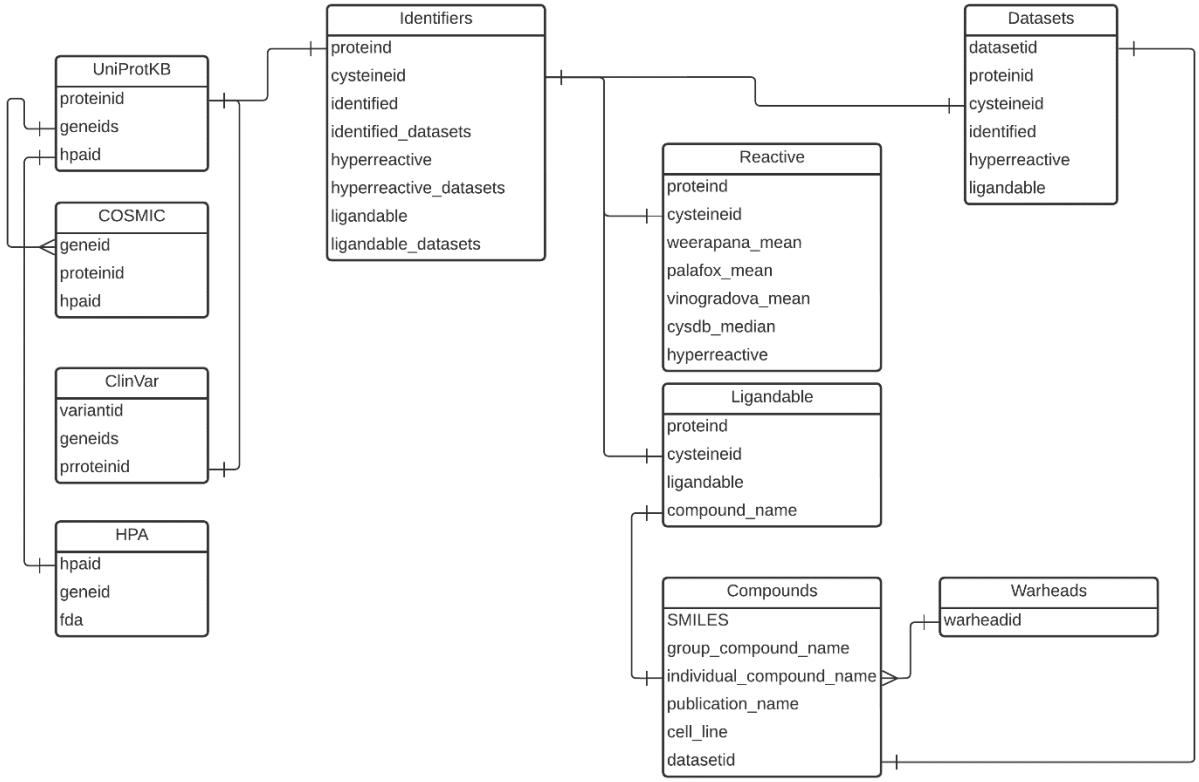

Figure S6. Entity-relationship diagram of all ten tables in CysDB and relationships with external data sources, such as UniProtKB, COSMIC, ClinVar and the Human Protein Atlas (HPA), related to **Figure 2**.

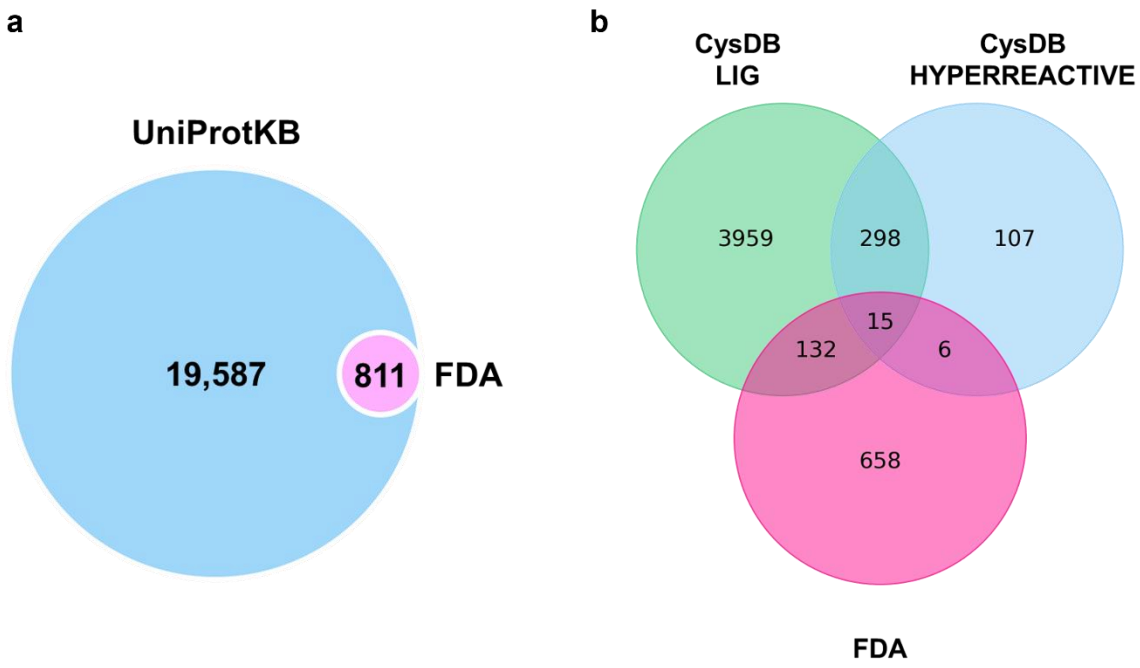

Figure S7. Total number of proteins in the human proteome from UniProtKB/Swiss-Prot and the subset targeted by FDA approved drugs (a). Total number of CysDB ligandable proteins, CysDB hyperreactive proteins and proteins targeted by FDA approved drugs (b), related to **Figure 4**.

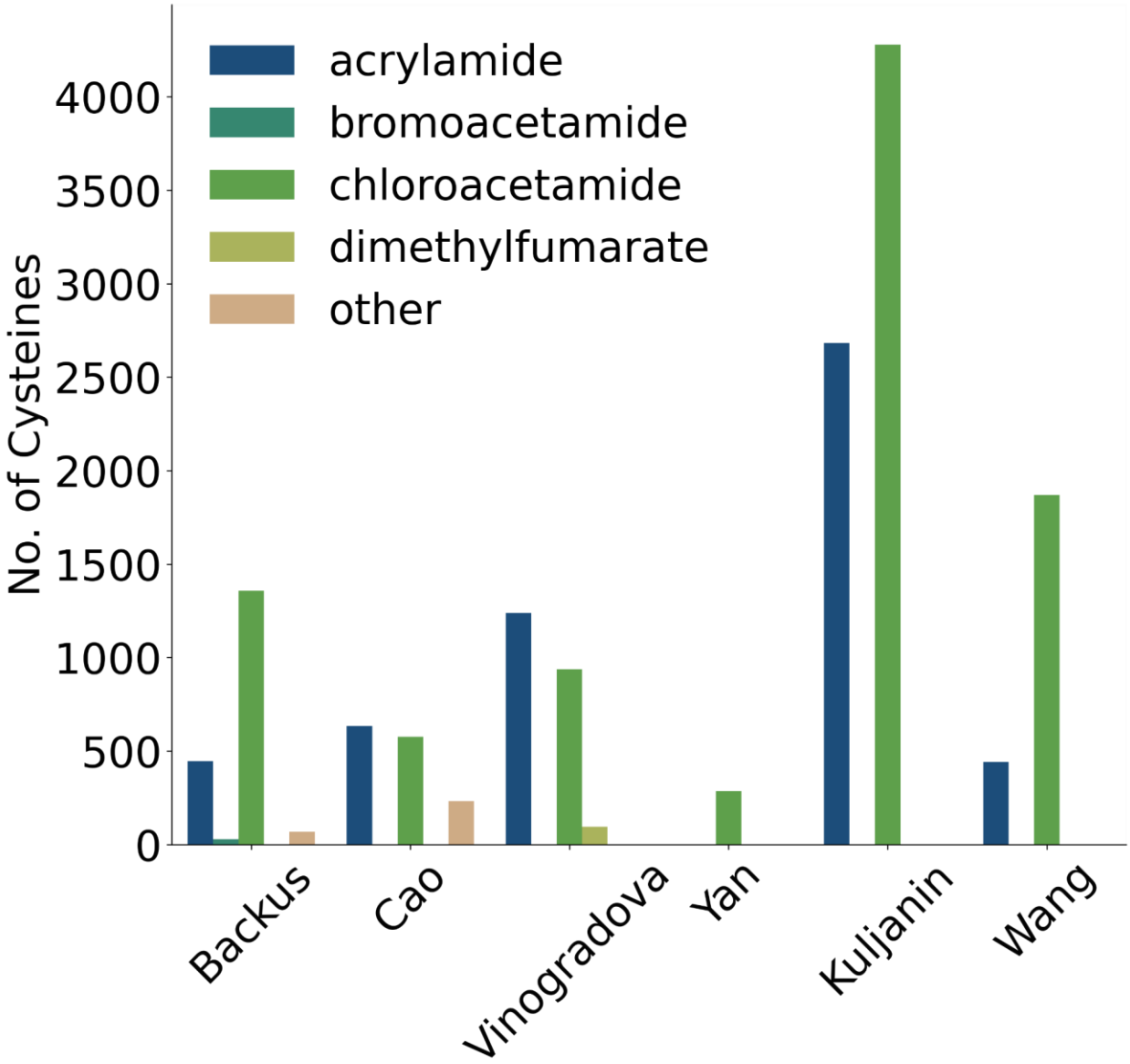

Figure S8. Total number of cysteines with an R > 4 by each warhead per dataset, in aggregate across all cell lines analyzed, related to **Figure 4**.

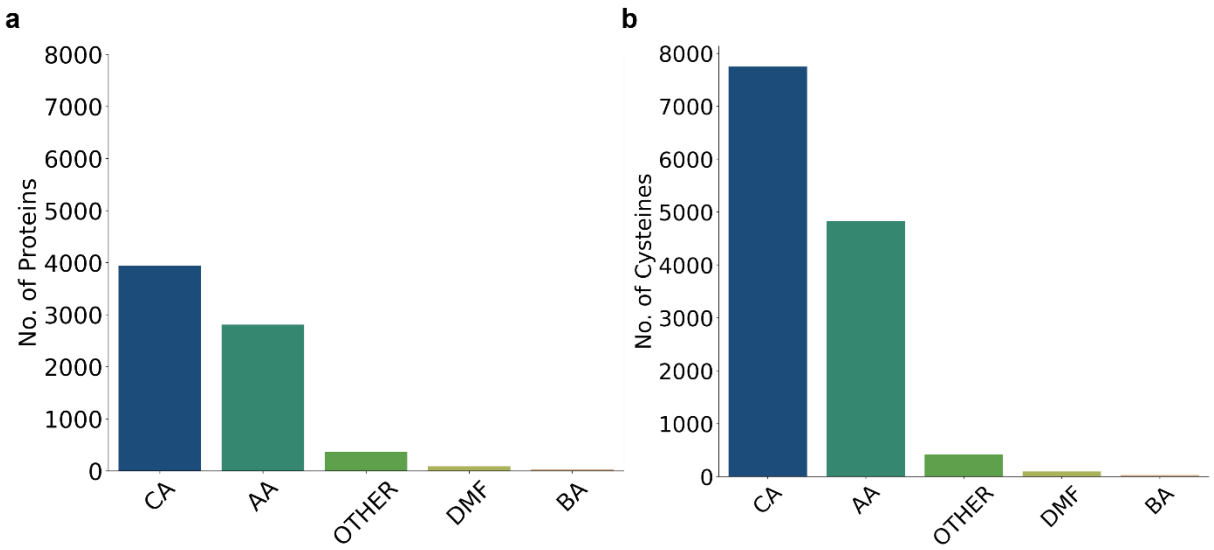

Figure S9. Total number of proteins (a) and cysteines (b) liganded by the following electrophiles, related to **Figure 4**: chloroacetamides (CA), acrylamides (AA), other (OTHER), dimethyl fumarate (DMF) and bromoacetamides (BA). Note, some proteins or cysteines were liganded by multiple warheads. Therefore, the counts in these graphs are not reflective of mutually exclusive events.

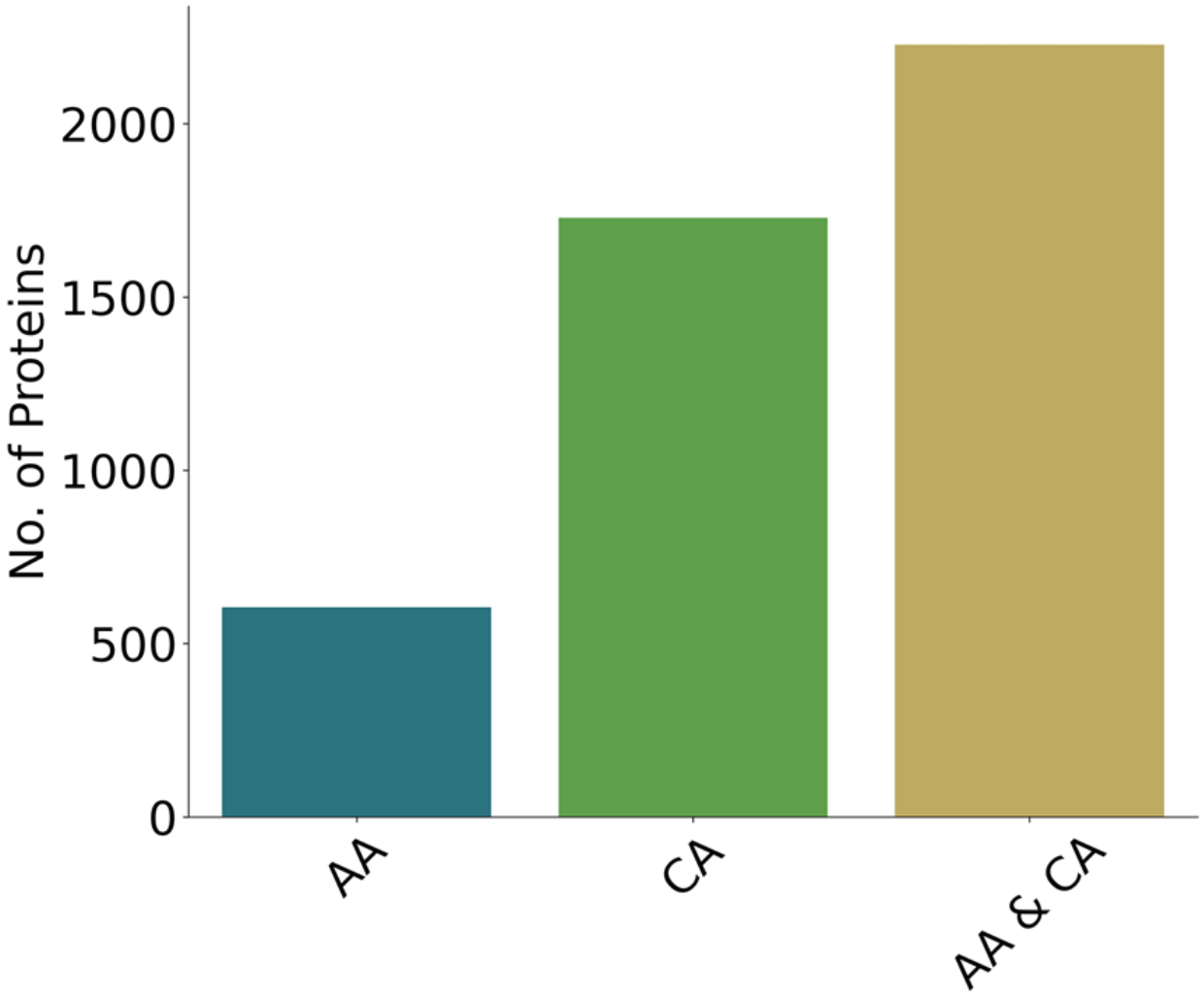

Figure S10. Total number of proteins liganded by both acrylamides and chloroacetamides, exclusively acrylamides and exclusively chloroacetamides, related to **Figure 4**.

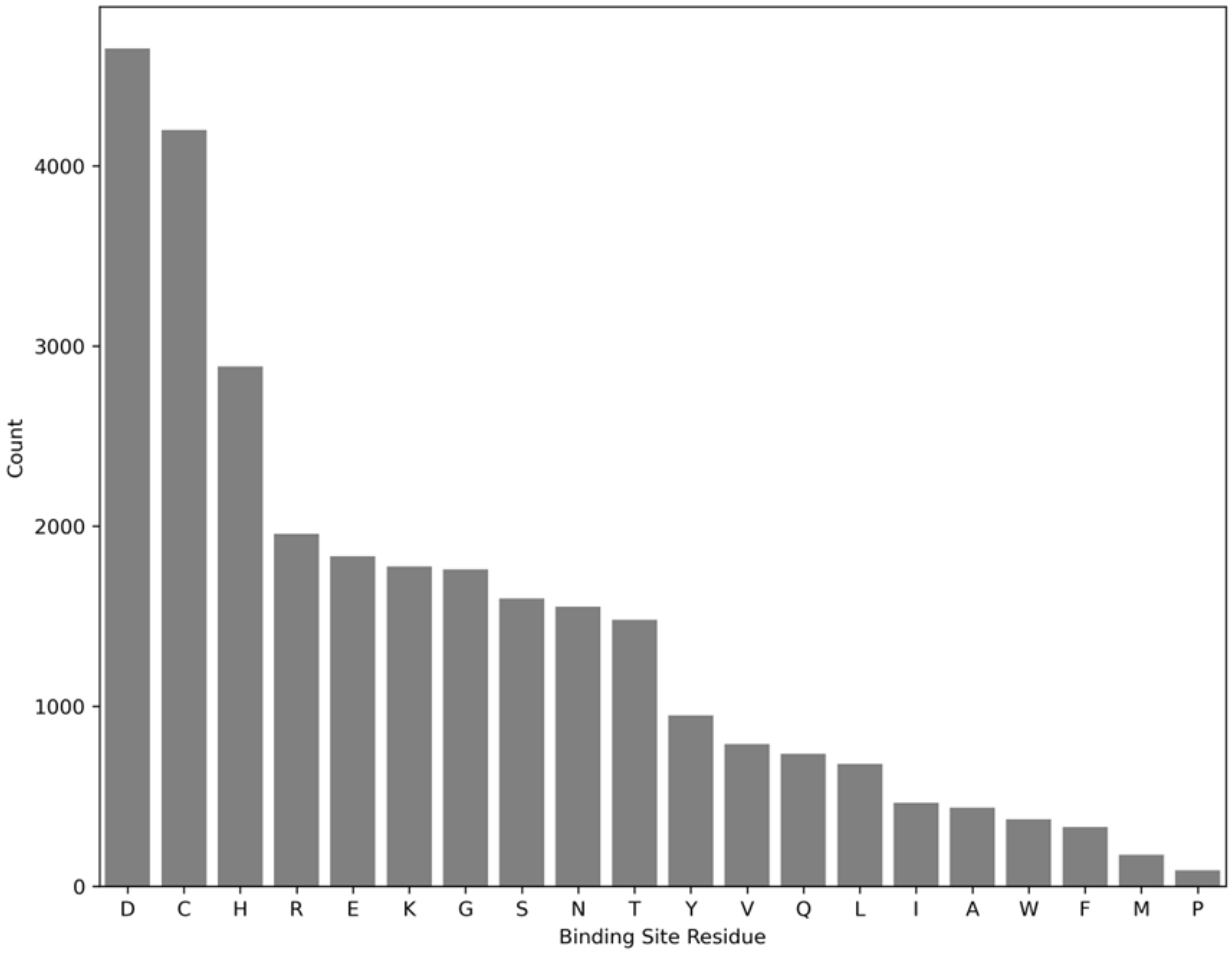

Figure S11. Distribution of amino acids annotated as binding sites in UniProtKB proteins, related to **Figure 5**.

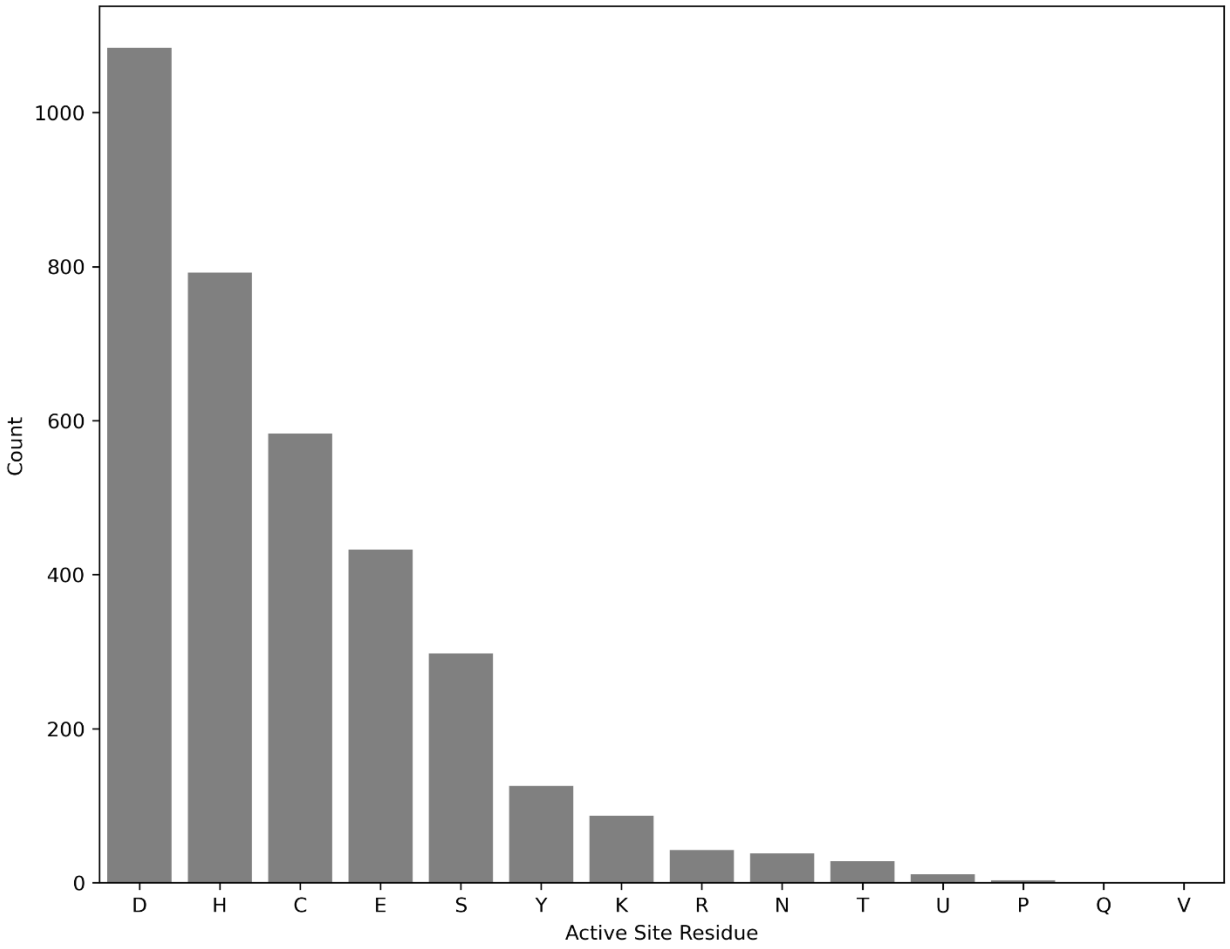

Figure S12. Distribution of amino acids annotated as active sites in UniProtKB proteins, related to **Figure 5**.

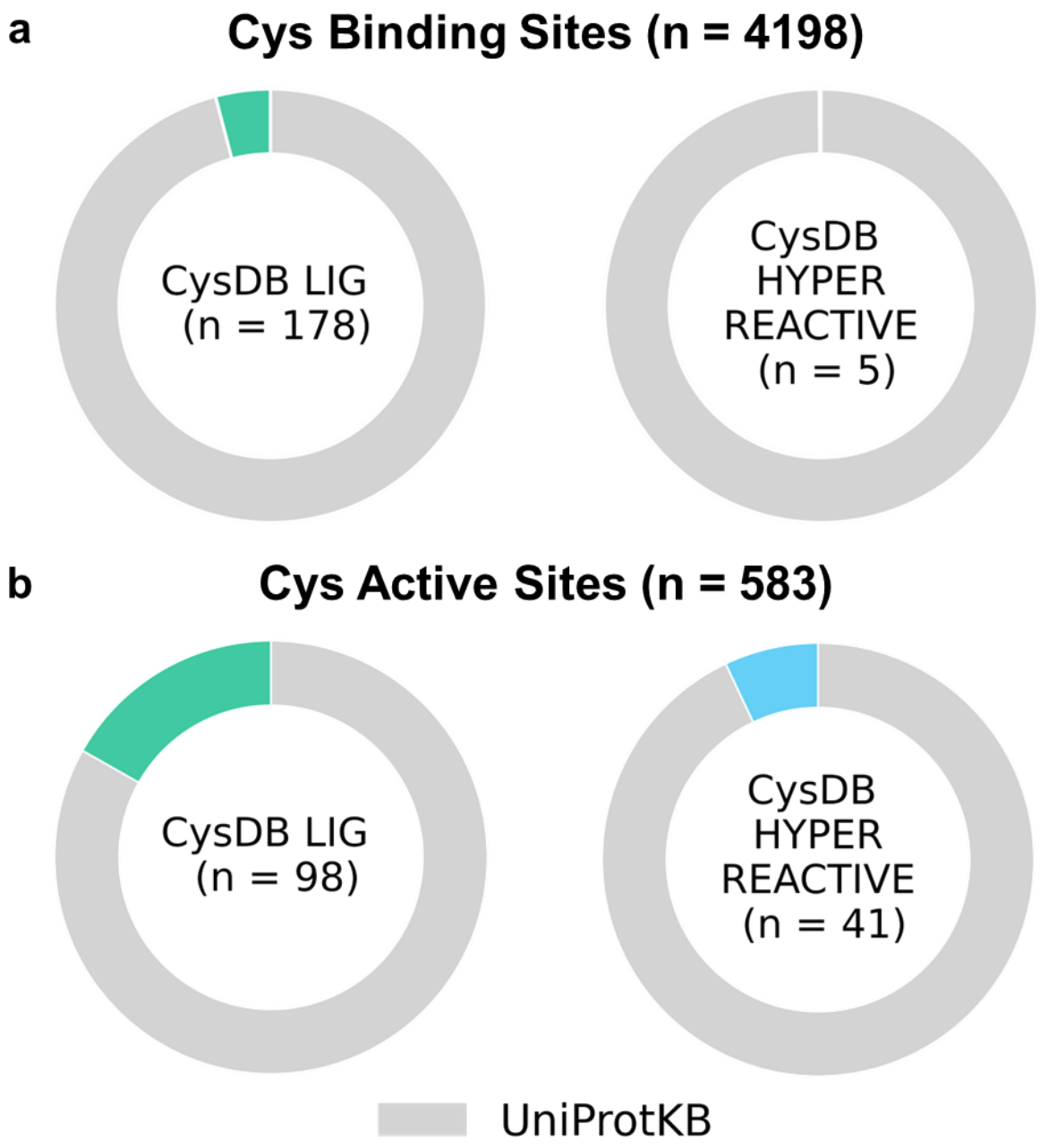

Figure S13. Distributions of ligandable (green) and hyperreactive (light blue) cysteines annotated as cysteine-specific binding sites (a) or cysteine-specific active sites (b), related to **Figure 5**. The total number of cysteines in UniProtKB annotated as binding or active sites are shown in gray.

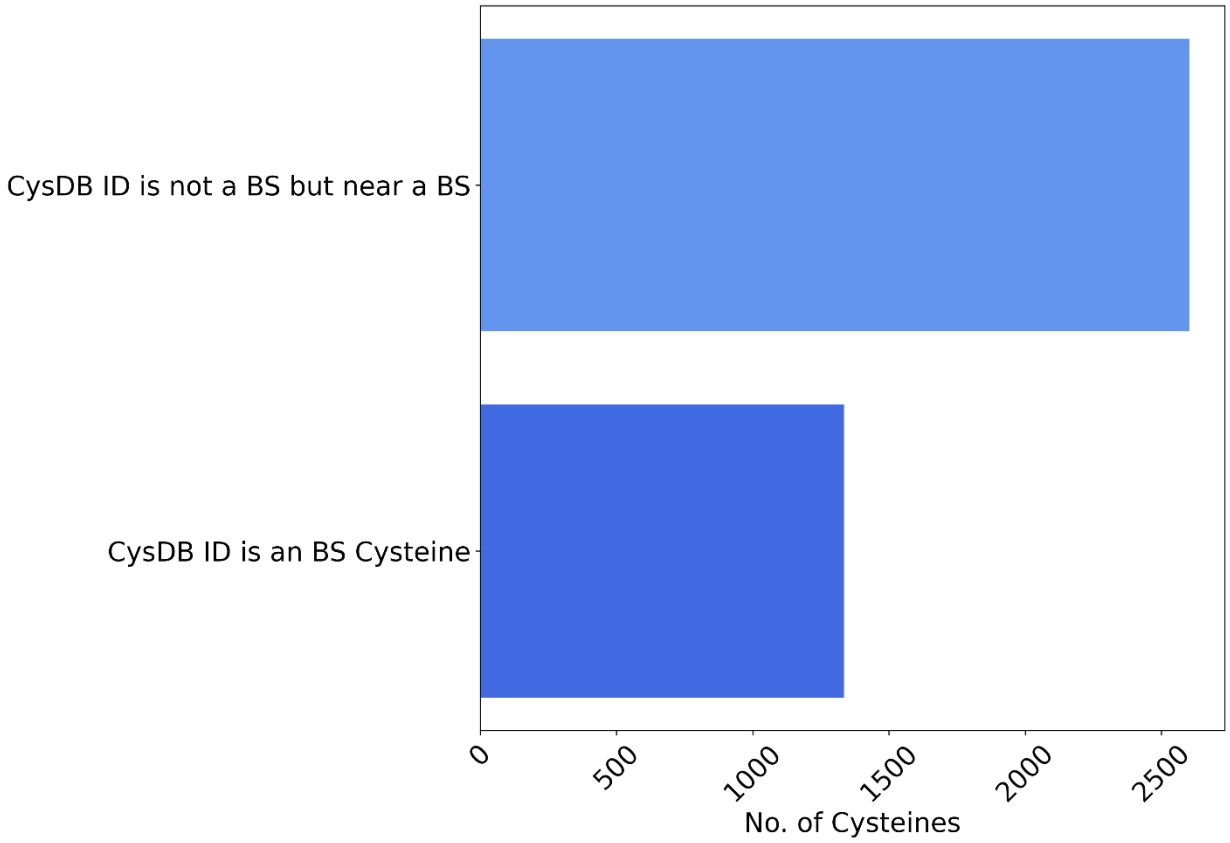

Figure S14. Number of CysDB ID cysteines that are annotated binding sites (BS) and cysteines that are not binding sites but in or near a binding site in 1D sequence, related to **Figure 5**. Primary sequences were searched +/- 10 amino acids from the location of a detected cysteine. If another binding site was within this +/- 10 amino acid window, the cysteine was considered ‘in or near’ a binding site.

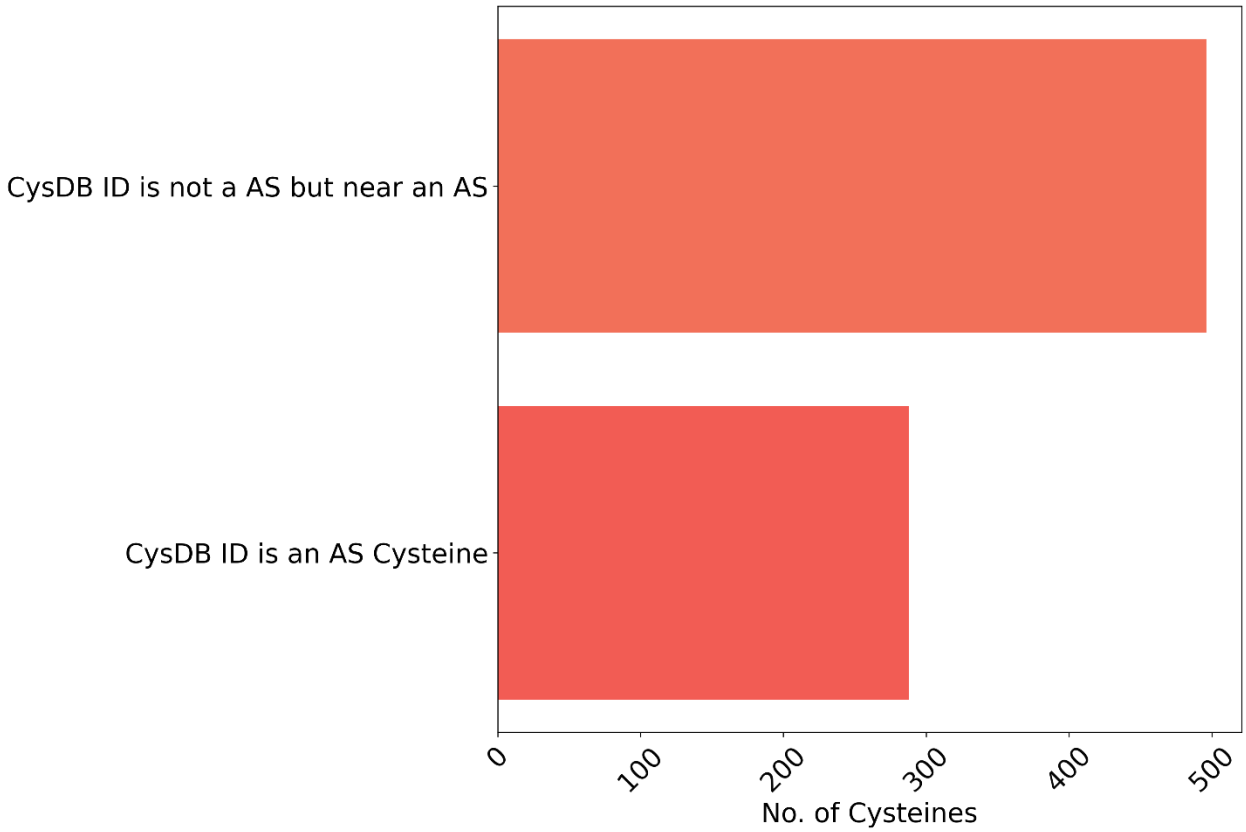

Figure S15. Number of CysDB ID cysteines that are annotated active sites (AS) and cysteines that are not active sites but in or near an active site in 1D sequence, related to **Figure 5**. Primary sequences were searched +/- 10 amino acids from the location of a detected cysteine. If another active site was within this +/- 10 amino acid window, the cysteine was considered ‘in or near’ an active site.

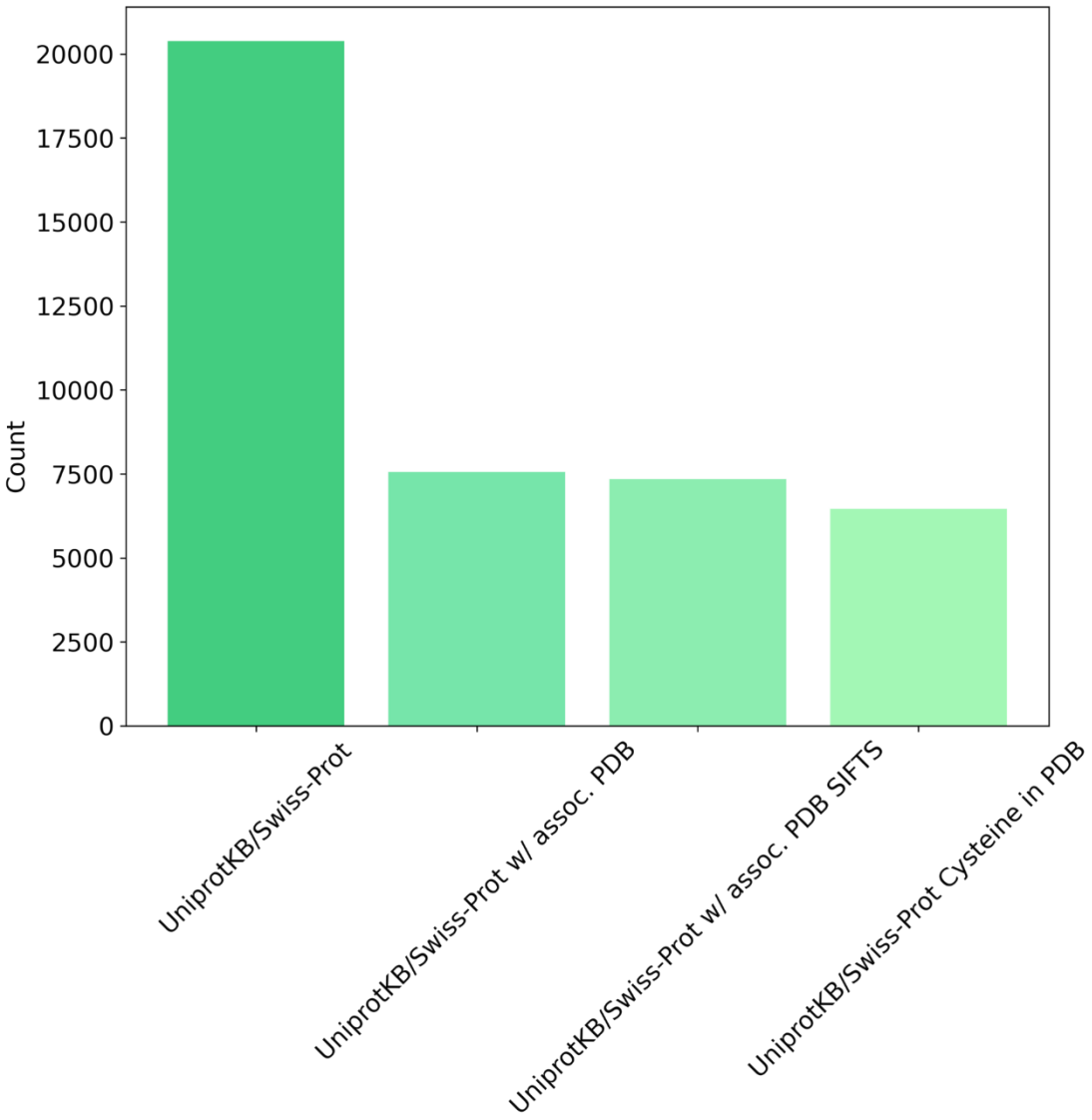

Figure S16. Number of UniProtKB proteins in the human proteome, with an associated PDB structure, residue mapped SIFTS file and with a cysteine resolved in the corresponding associated PDB, related to **Figure 5**.

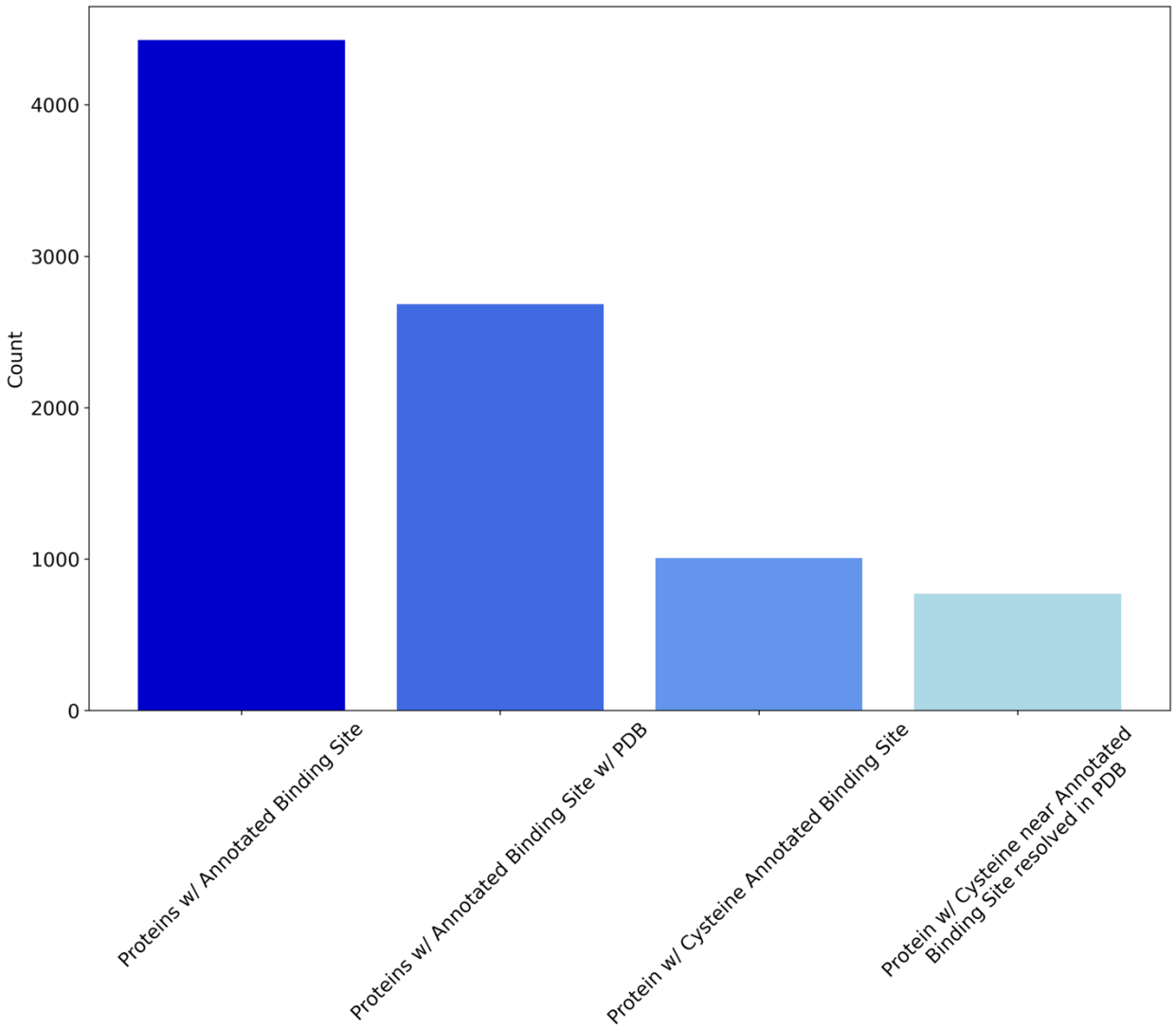

Figure S17. Number of UniProtKB proteins with an annotated binding site, associated PDB structure, with an annotated cysteine binding site and with cysteines near an annotated binding site in an associated PDB structure, related to **Figure 5** and see **STAR Methods**. The distance from the sulfur atom of each cysteine to an annotated binding site residue was calculated. Cysteines within 10 Angstroms of the annotated binding site residue were considered as cysteines ‘in or near’ binding sites.

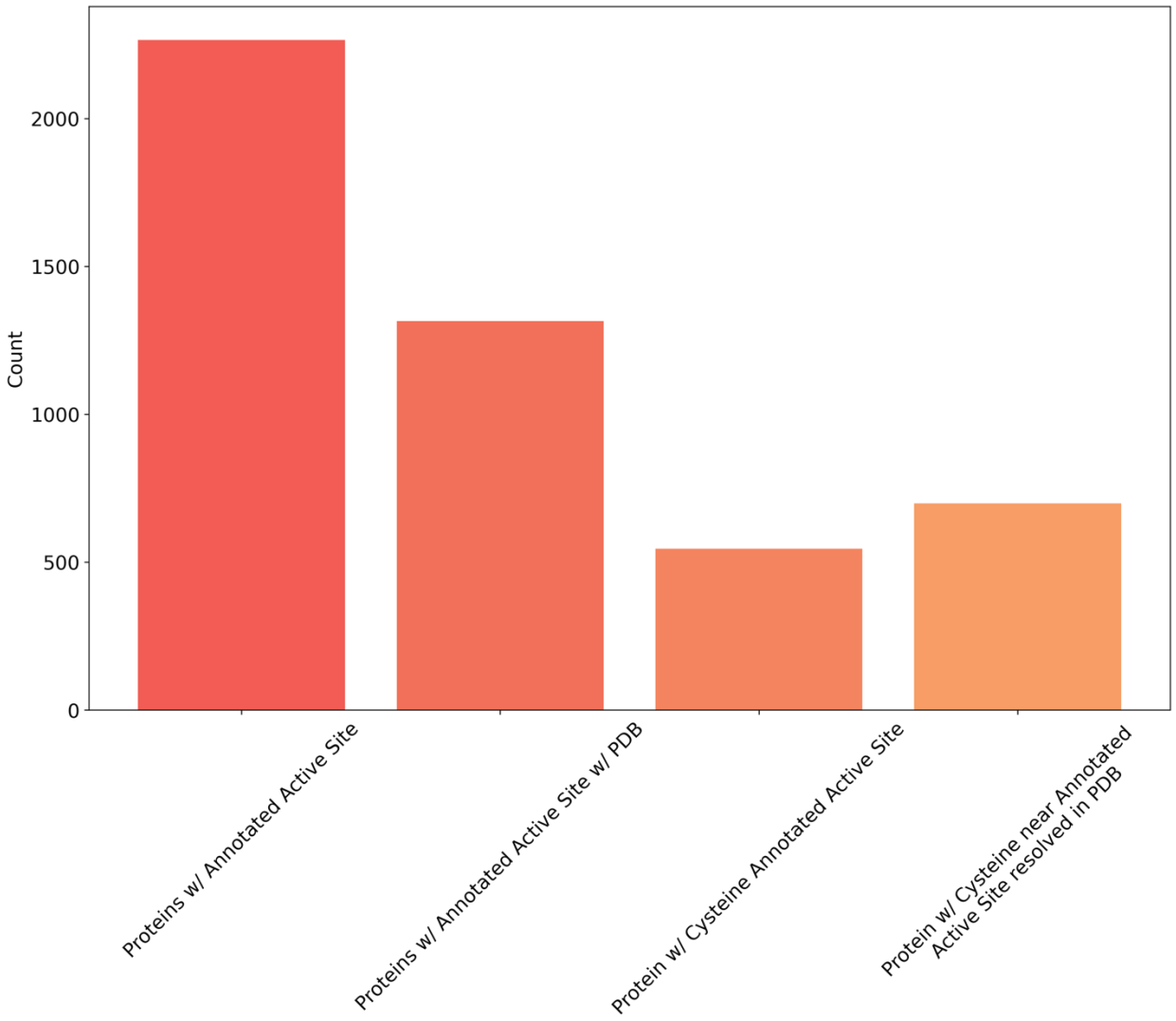

Figure S18. Number of UniProtKB human proteins with an annotated active site, associated PDB structure, with an annotated as cysteine active site and with cysteines near an annotated active site in an associated PDB structure, related to **Figure 5** and see **STAR Methods**. The distance from the sulfur atom of each cysteine to an annotated active site residue was calculated. Cysteines within 10 Angstroms of the annotated active site residue were considered as cysteines 'in or near' active sites.

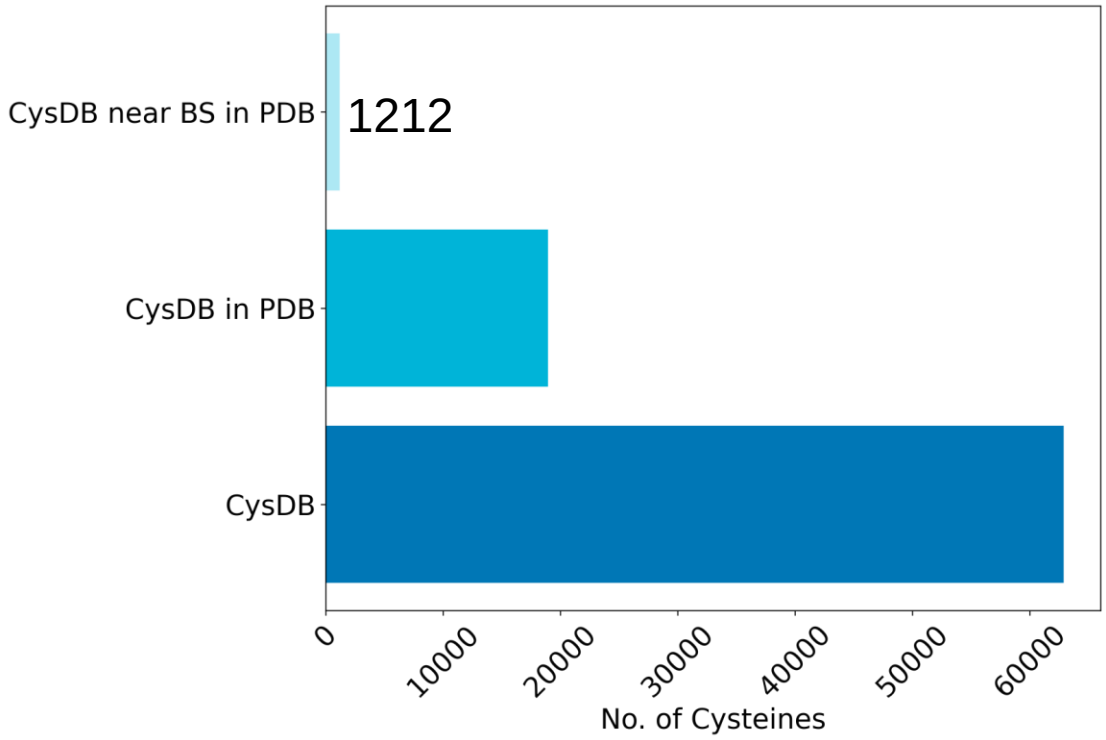

Figure S19. Number of CysDB ID cysteines identified by chemoproteomics, resolved in an associated PDB and CysDB ID cysteines that are not annotated binding sites but are 'in or near' an annotated binding site in 3D space, related to **Figure 5**. Proteins with an annotated binding site, annotated as a binding site resolved in an associated PDB structure and with cysteines 'in or near 'an annotated binding site. The distance from the sulfur atom of each cysteine to an annotated binding site residue was calculated. Cysteines within 10 Angstroms of the annotated binding site residue were considered as cysteines 'in or near' binding sites.

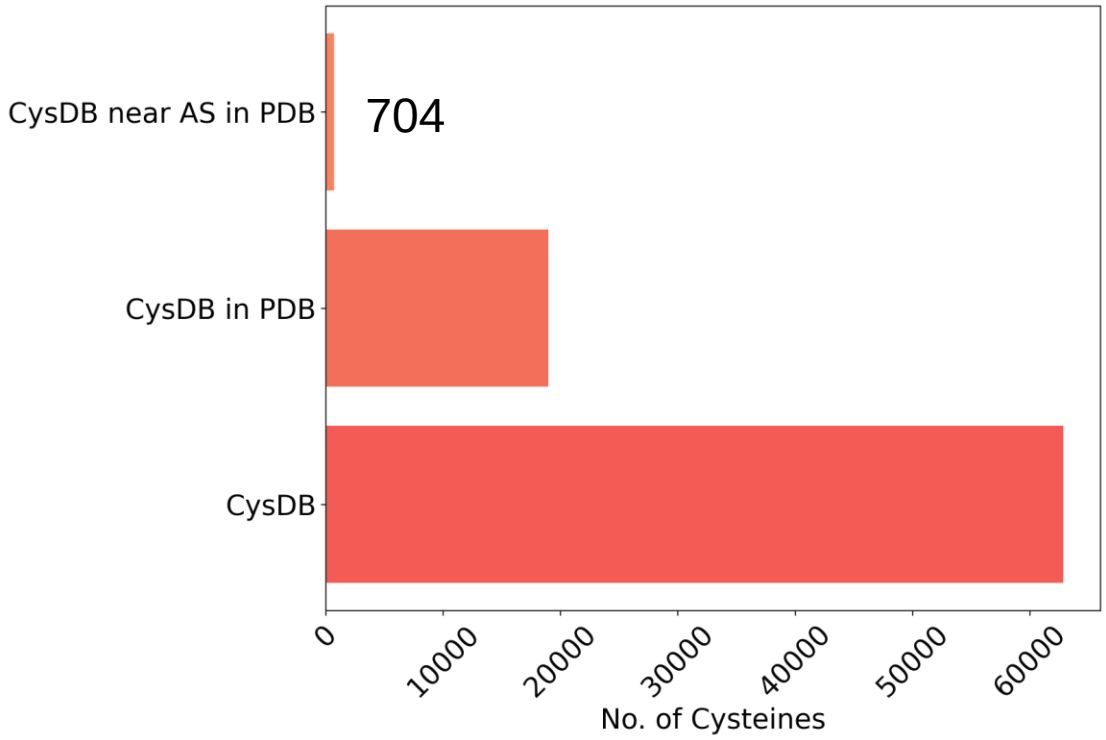

Figure S20. Number of CysDB ID cysteines identified, resolved in an associated PDB and CysDB ID cysteines that are not annotated active sites but are 'in or near' an annotated active site in 3D space, related to **Figure 5**. Proteins with an annotated active site, annotated as an active site resolved in an associated PDB structure and with cysteines 'in or near' an annotated binding site. The distance from the sulfur atom of each cysteine to an annotated active site residue was calculated. Cysteines within 10 Angstroms of the annotated active site residue were considered as cysteines 'in or near' active sites.

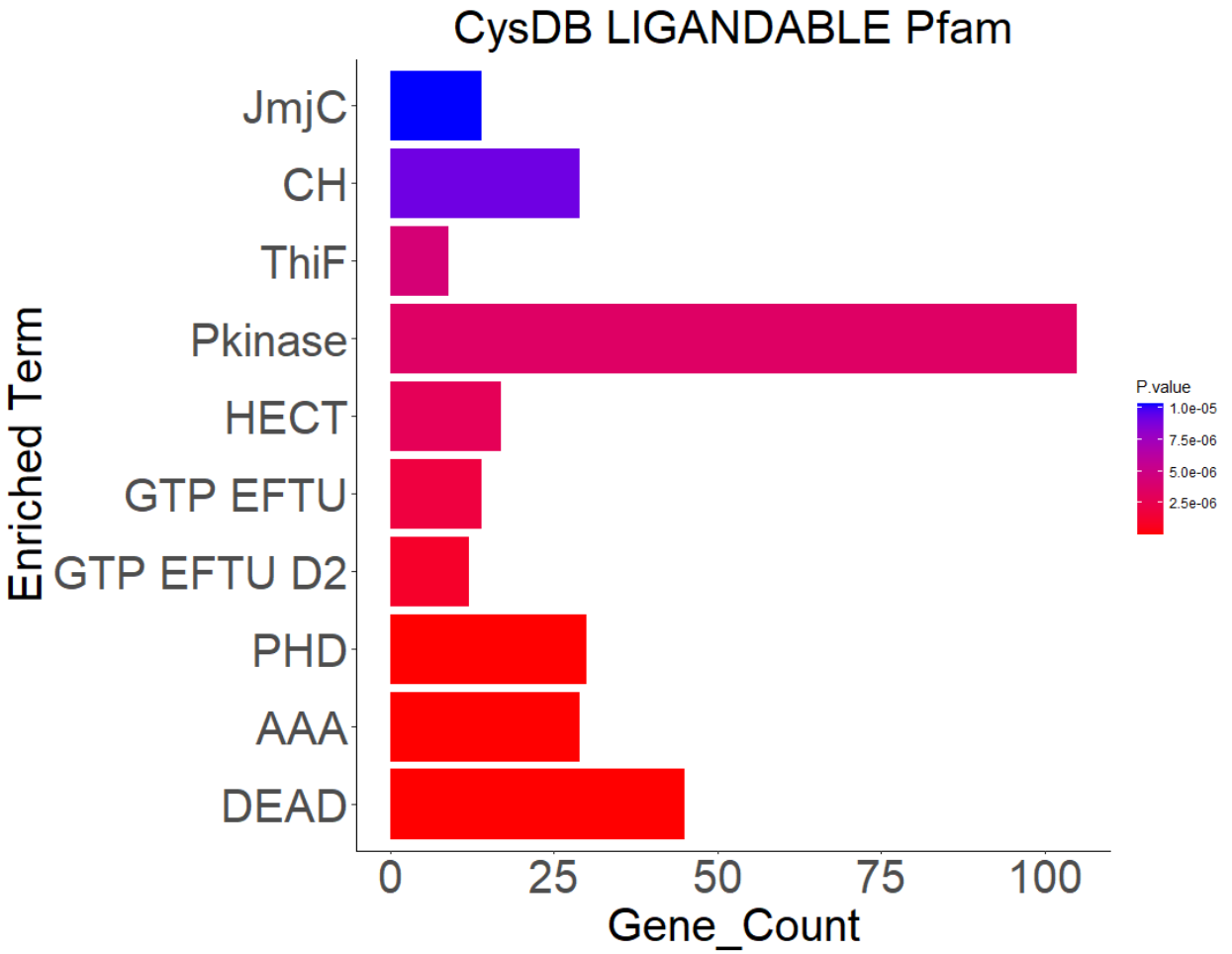

Figure S21. Top-10 enriched protein domains from Pfam-term enrichment analysis of liganded proteins with gene counts, related to **Figure 6**.

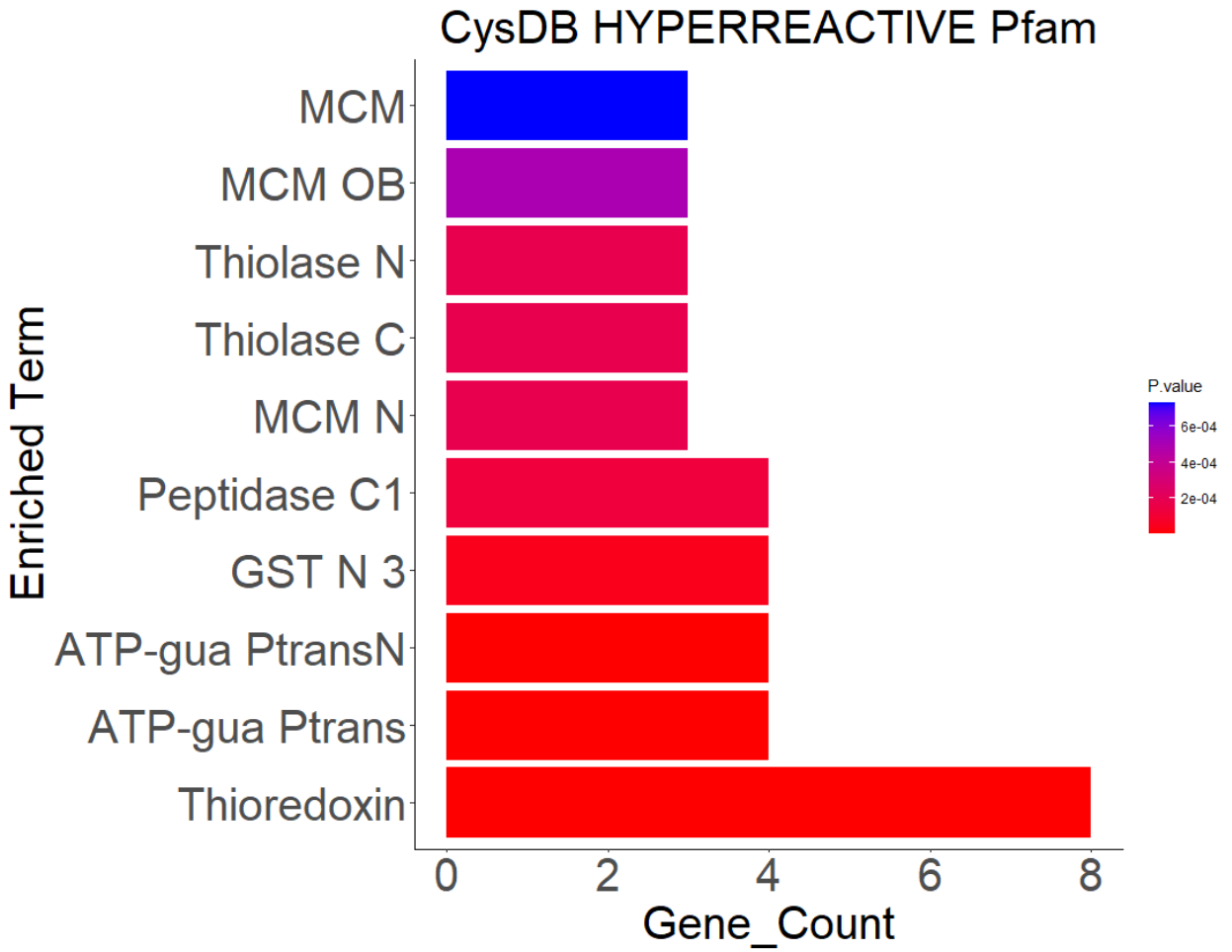

Figure S22. Top-10 enriched protein domains from Pfam-term enrichment analysis of hyper-reactive proteins with gene counts, related to **Figure 6**.

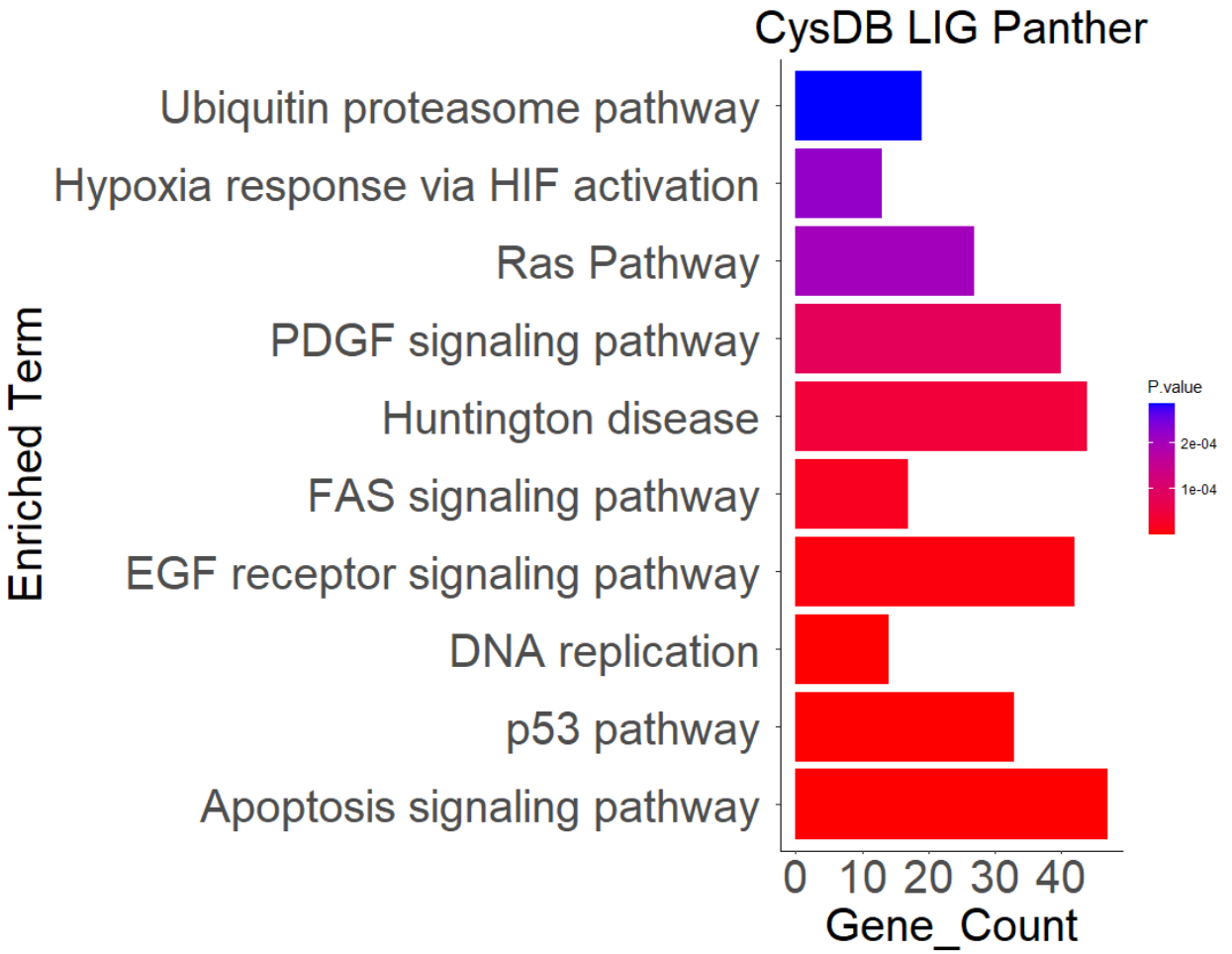

Figure S23. Top-10 enriched pathways from Panther-term enrichment analysis of liganded proteins with gene counts, related to **Figure 6**.

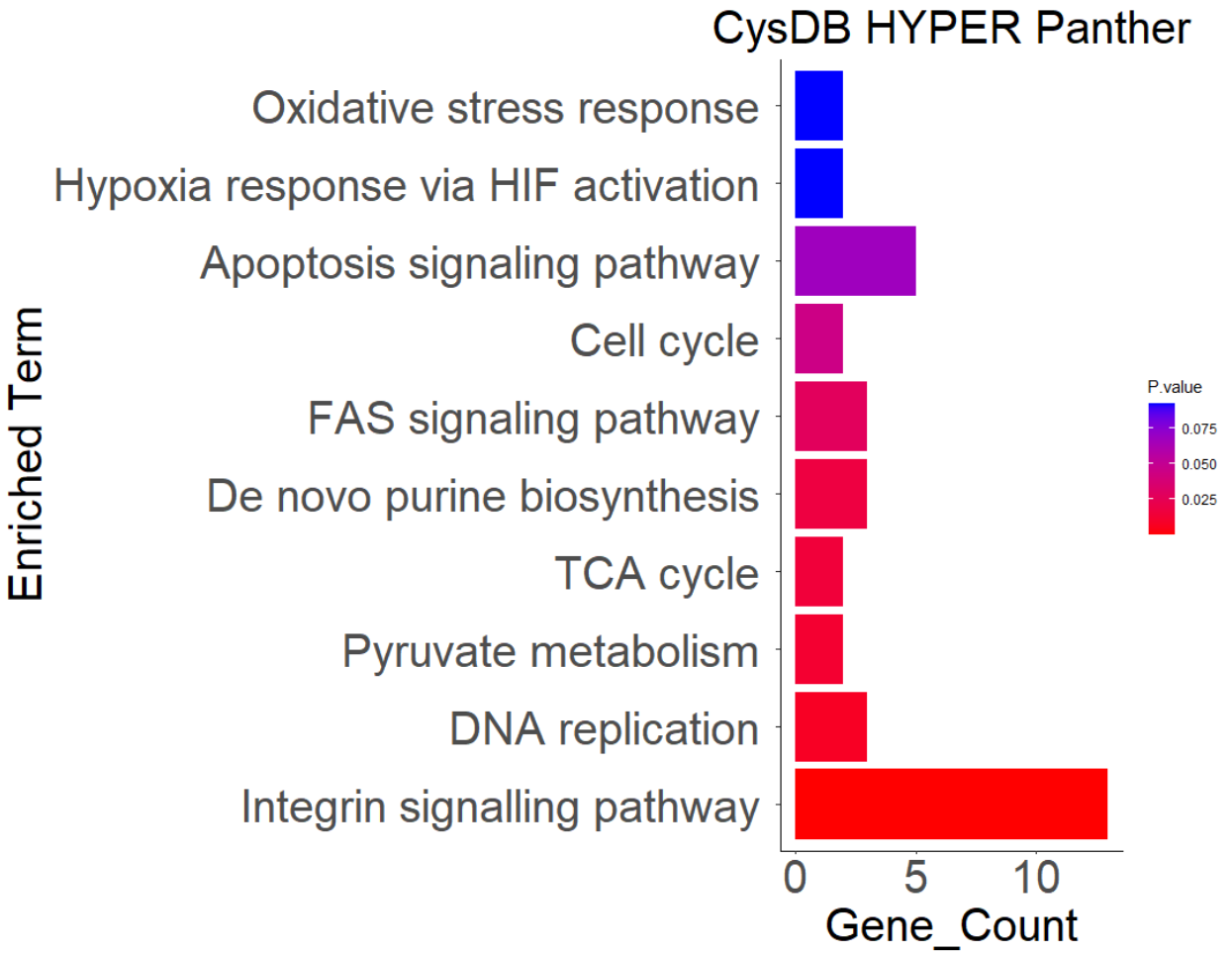

Figure S24. Top-10 enriched pathways from Panther-term enrichment analysis of hyperreactive proteins with gene counts, related to **Figure 6**.

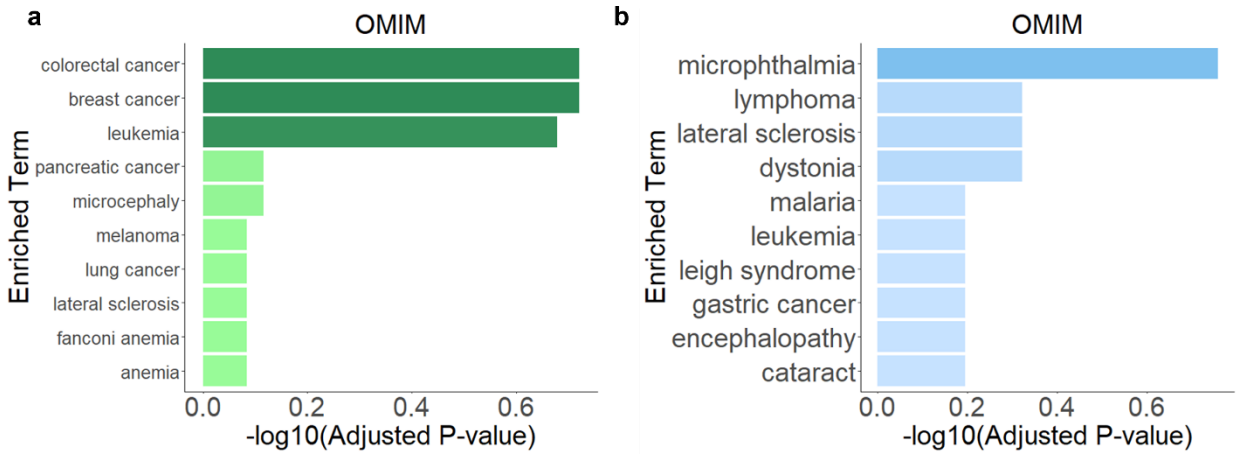

Figure S25. Top-10 enriched pathways from OMIM-term enrichment analysis of ligandable proteins (a) and hyperreactive proteins (b), related to **Figure 6**.

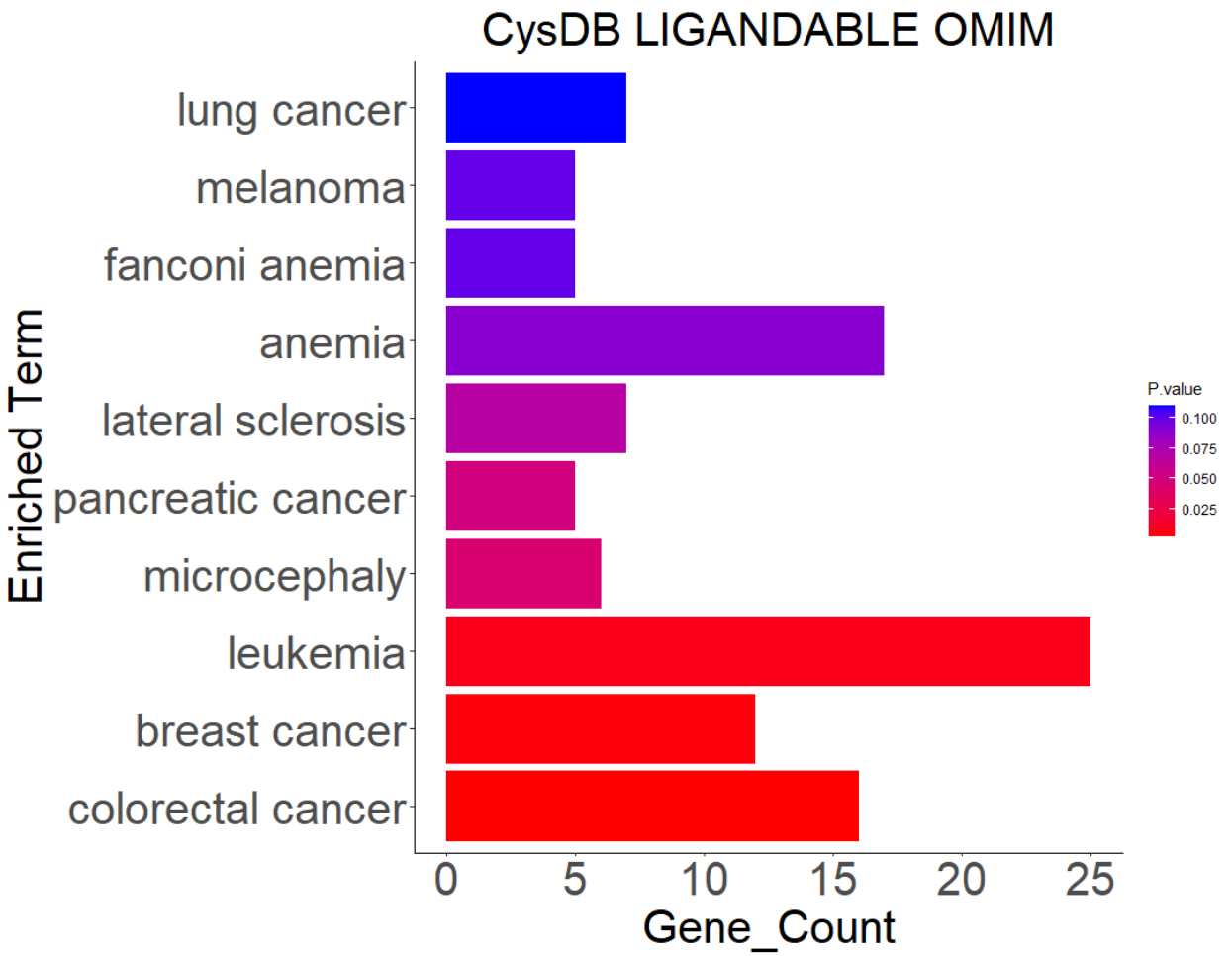

Figure S26. Top-10 enriched pathways from OMIM-term enrichment analysis of ligandable proteins with gene counts, related to **Figure 6**.

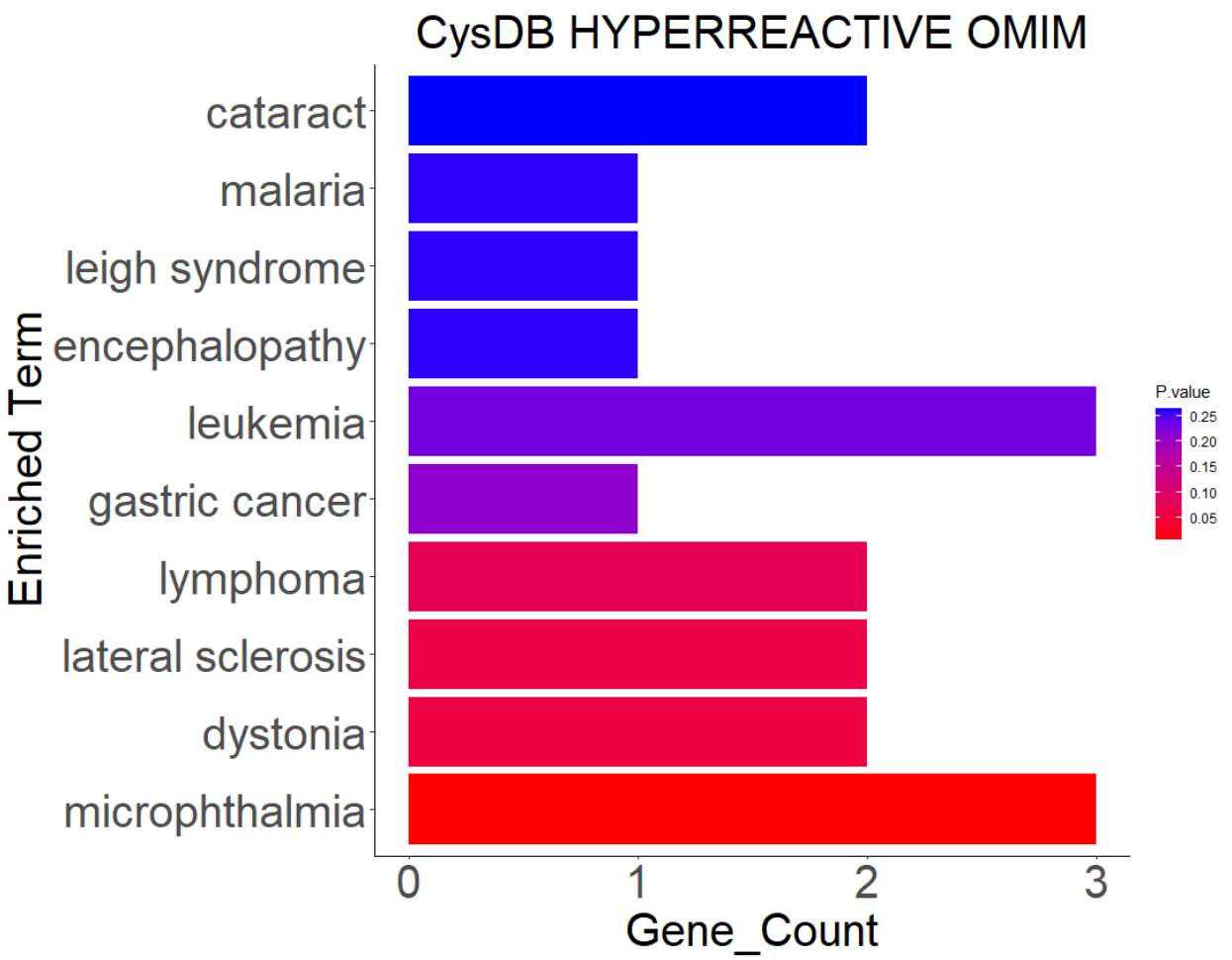

Figure S27. Top-10 enriched pathways from OMIM-term enrichment analysis of hyperreactive proteins with gene counts, related to **Figure 6**.

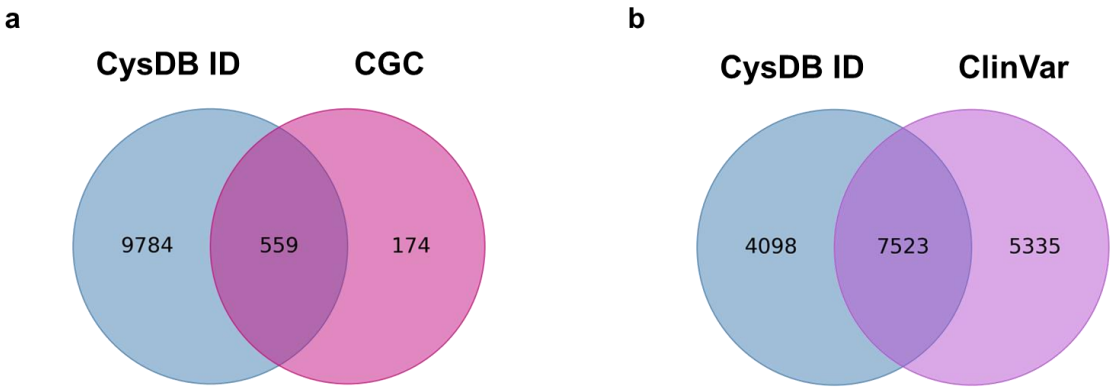

Figure S28. Overlap between the number of genes associated with CysDB identified proteins and Cancer Gene Census (CGC) genes (a), related to **Figure 6**. Overlap between the number of CysDB identified proteins and proteins associated with ClinVar variants (b).

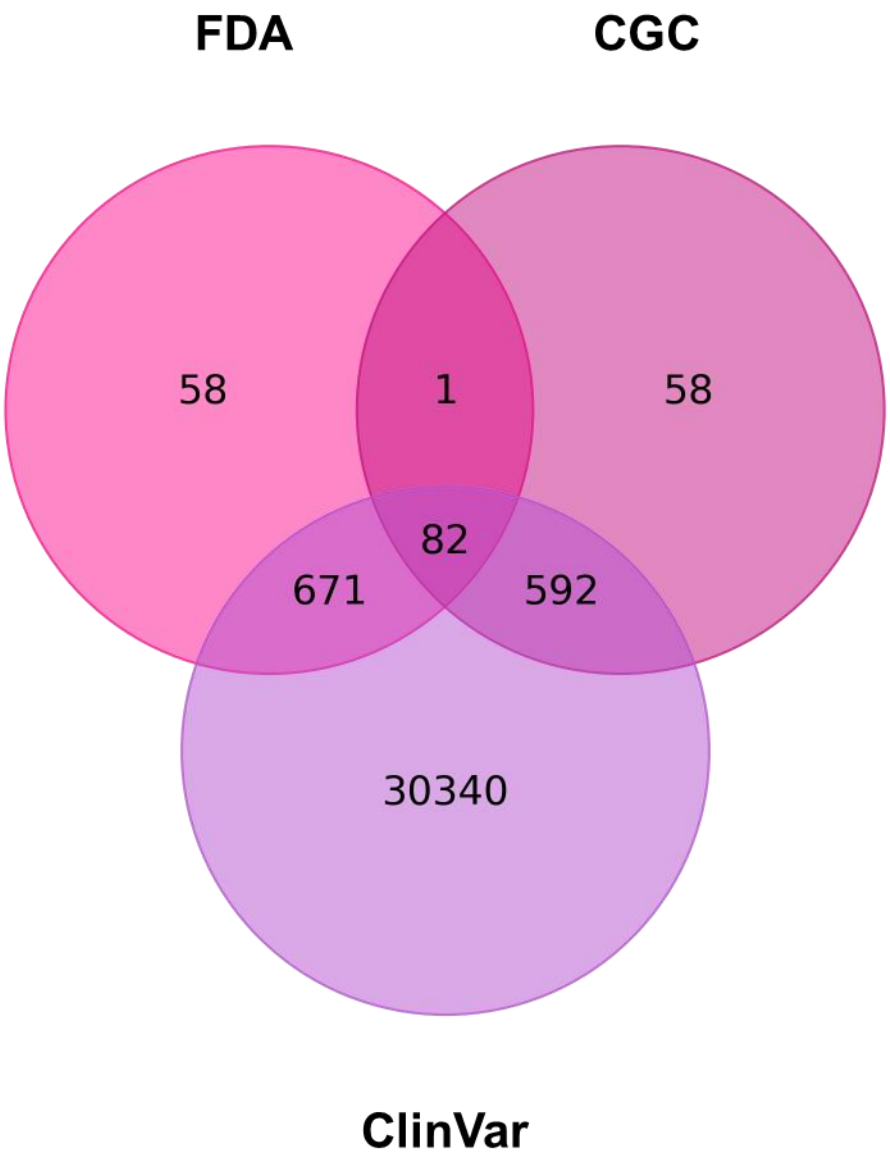

Figure S29. Overlap between the number of FDA targeted genes, Cancer Gene Census (CGC) genes and genes associated with ClinVar variants, related to **Figure 6**.

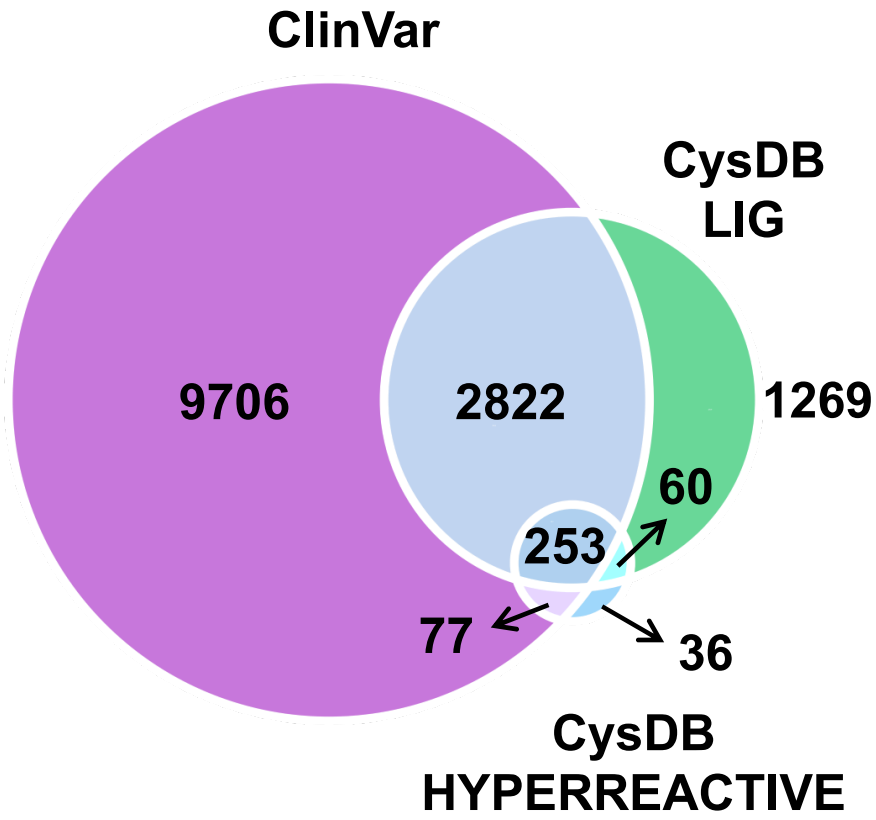

Figure S30. Overlap between the number of CysDB LIG, CysDB HYPERREACTIVE proteins and proteins associated with ClinVar variants, related to **Figure 6**.

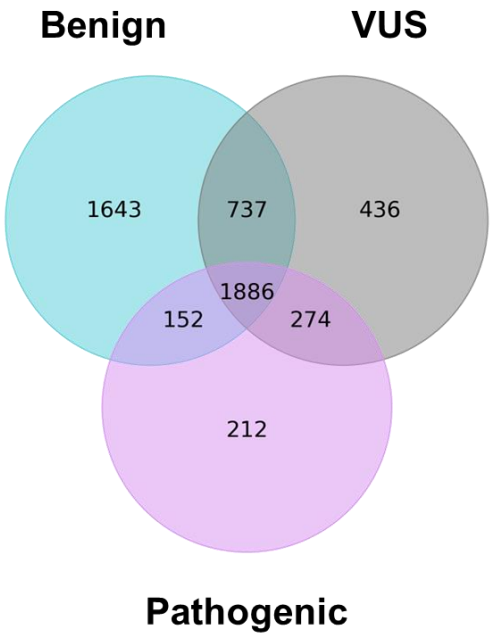

Figure S31. Overlap between the number of benign, variants of unknown significance (VUS) and pathogenic ClinVar missense variants for CysDB ID proteins, related to **Figure 6**.