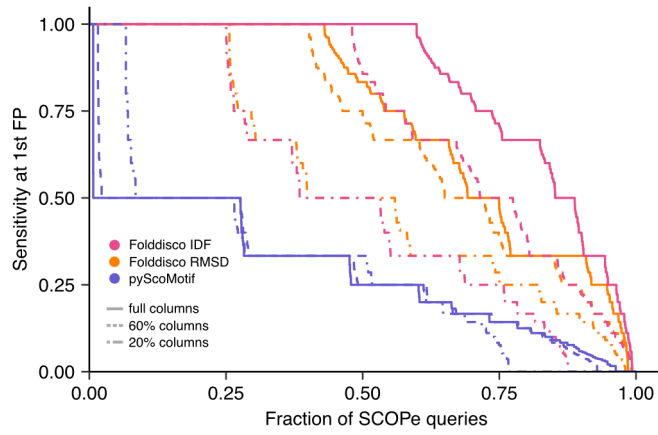
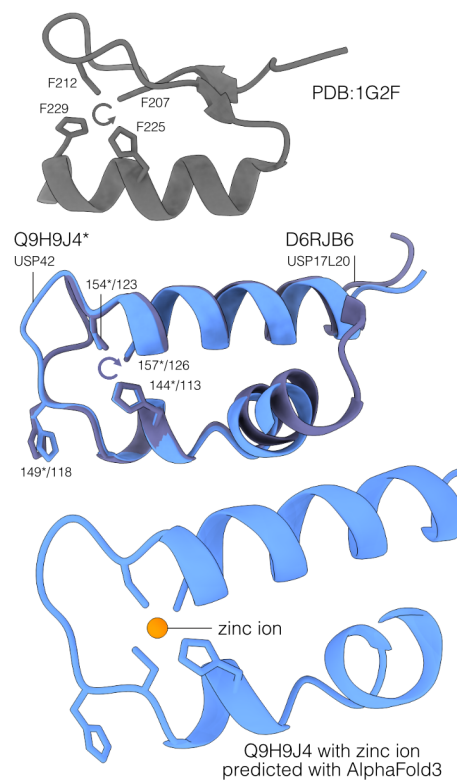


**Extended Data Fig. 1. Performance comparison of FoldDisco to pyScoMotif and RCSB** Precision, recall, F1-score and search time (in seconds) with 12 threads were evaluated for **top** serine peptidase motif query and **bottom** zinc-finger motif with 3 residues.

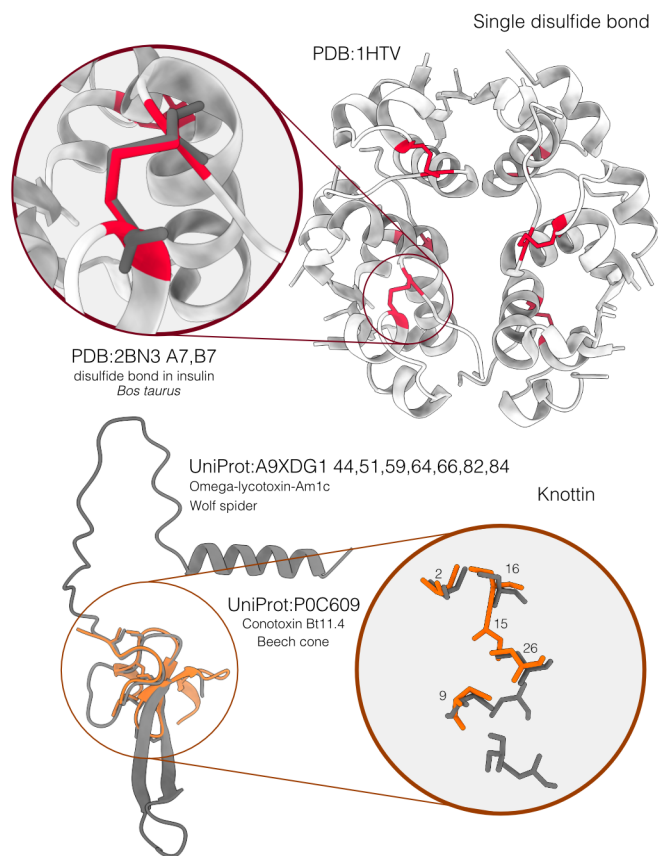


Tool	Metric	Area under curve		
		20%	60%	100%
Folddisco	IDF	0.504	<b>0.732</b>	<b>0.837</b>
Folddisco	RMSD	<b>0.540</b>	0.686	0.726
pyScoMotif	RMSD	0.300	0.290	0.285

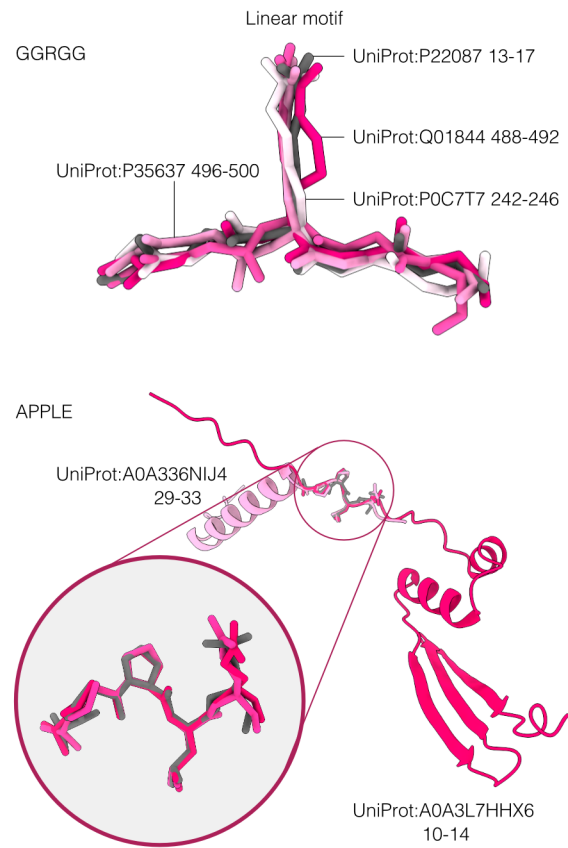
**Supplementary Figure 1. Sensitivity of Folddisco using two scoring modes and of pyScoMotif on SCOPe-constructed benchmarks.** This figure extends Fig. 1f, where the goal is to match SCOPe sequences of the same family as the query before matching a different fold, using all conserved columns ("full") or a random subsample of them (60%, 20%). Folddisco's matches were ranked either by its IDF-based pre-filter coverage score (pink) or its full pipeline's RMSD (orange), which is also used for ranking by pyScoMotif (blue). Folddisco achieves higher area under the curve values than pyScoMotif in all cases (bottom panel). Comparing Folddisco's two scoring modes, RMSD is more sensitive for ranking matches of short queries (20% benchmark), while the IDF-based coverage score yields higher area the under curve values for the longer 60% and 100% benchmarks (bottom panel).



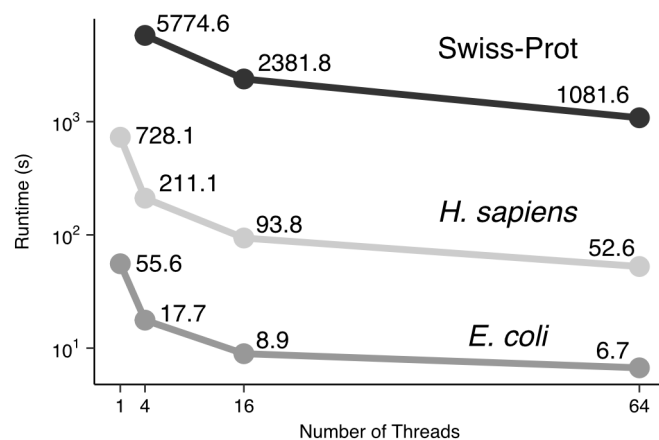
**Supplementary Figure 2. Partial zinc finger motifs identified in ubiquitin-specific peptidases.** From *H. sapiens* proteome, FoldDisco detected partial matches of the zinc finger motif (**top**) in ubiquitin-specific peptidases USP42 and USP17L20 (**middle**). Notably, the residue order is reversed in these matches. AlphaFold3 predictions confirmed zinc coordination within these motifs (**bottom**).



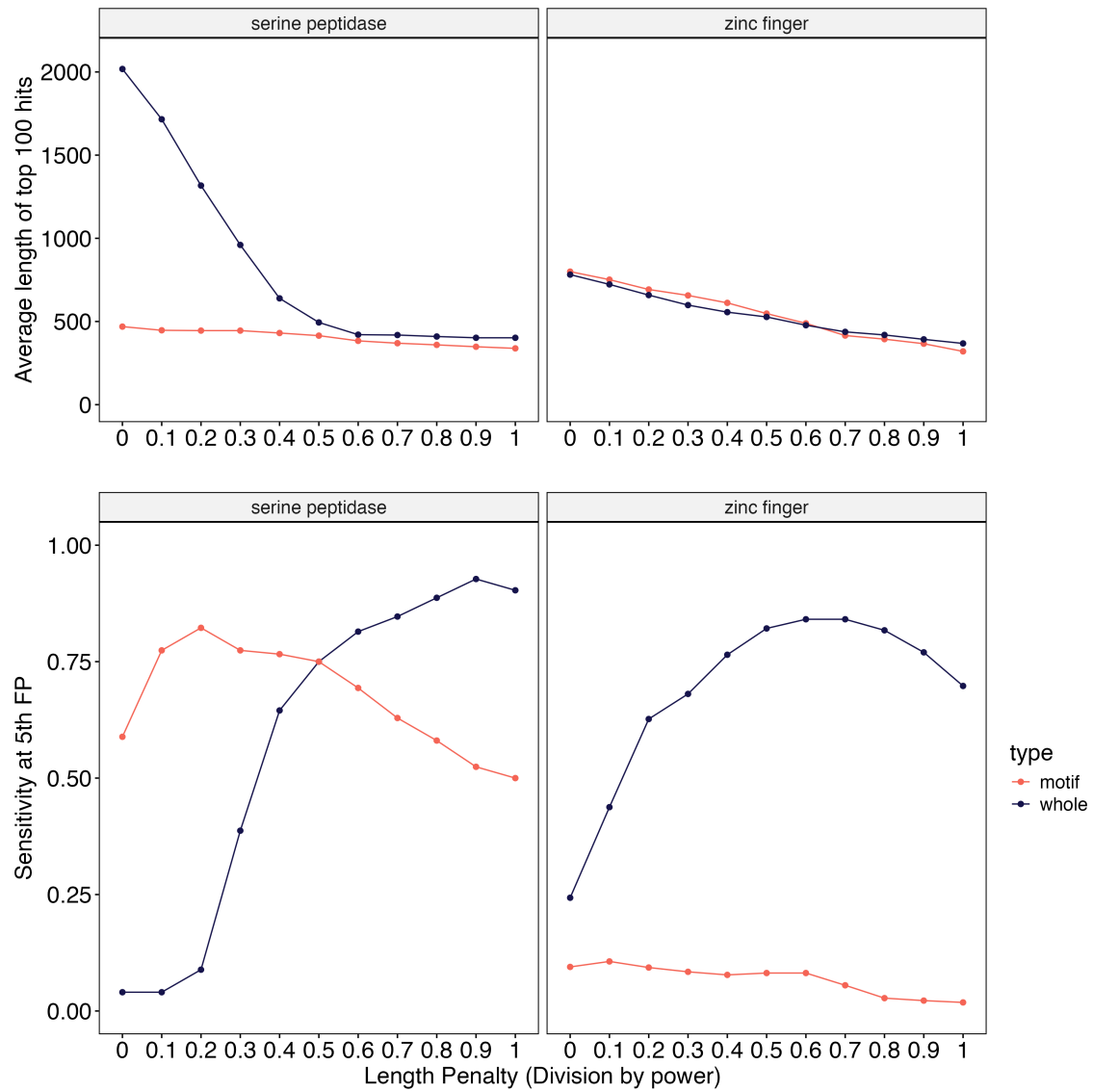
**Supplementary Figure 3. Detection of single disulfide bonds and knottin motifs** FoldDisco identified single disulfide bonds (red) in insulin (**top**) and complex knottin motifs formed by multiple disulfide bonds (**bottom**). From search of a knottin motif in a spider toxin, a conotoxin was retrieved with a partial match.



**Supplementary Figure 4. Identification of short linear motifs by FoldDisco.** FoldDisco searched and detected the known “GGRGG” motif (**top**) and an arbitrarily generated “APPLE” motif (**bottom**). The query “APPLE” motif structure (gray) was predicted using ColabFold.



**Supplementary Figure 5. Runtime benchmarking of Folddisco index construction across databases and CPU cores.** Index building time (in seconds) was measured for three databases — Swiss-Prot, human and *E. coli* subset of the AFDB-proteome — using 1, 4, 16, and 64 CPU cores.



**Supplementary Figure 6. Tuning the length-penalty exponent.** For motif-only (coral) and full-structure (navy) queries, the average length of the top-100 hits (top) and sensitivity at the 5th false positive (bottom) are plotted against the length-penalty exponent. The length penalty value selected for Folddisco—an exponent of 0.5—removes long serine-peptidase false-positive hits while maintaining zinc-finger sensitivity and is used as the default.