Table 3: Predicted pentapeptides in 8 contact regions of other proteins not included in the dataset

PDB Code	Protein Name	pentapeptide	Class
1abr	Abrin-A	R I **G** G R	β
1got*	Gt-alpha/gu-alpha chimera gt-beta gt-gamma	A T **G** S D	β
1p04	Alpha-lytic protease	D S **G** G S	β
3sil	Sialidase	R S **D** I S	β
1b2n*	Methanol dehydrogenase	G S **G** N P	α/β
1agn*	Human sigma alcohol dehydrogenase	A G **A** S R	α/β
1ysc	Serine carboxypeptidase	A G **D** K D	α/β
2ctc*(156)	Carboxypeptidase	A G **A** S S	α/β
2ctc*(252)	Carboxypeptidase	A S **G** G S	α/β
2tys*	Tryptophan synthase	A G **A** I S	α/β
1fdi	Formate dehydrogenase	K S **A** A I	α/β
1gso	Glyanamide ribonucleotide synthease	V I **G** N G	α/β
1ixh*	Phosphate binding protein	G I **G** S S	α/β
1rus	Rubisco	G A **S** G I	α/β
2pia*	Pthalate deoxygenase reductase	V A **G** G I	α/β

* denotes the pentapeptide in the protein is close to the active/binding site. Within the braces is indicated the residue number of the residue occurring in 8 contact of 2ctc.