

## **Table S1. BLOCK 1 Rules for recognition and selection of a helical pair for a peptide fragment**

1. Analysis of the secondary structure of the amino acid chain is carried out using the Dictionary for Secondary Structure of Protein. Search for sections of helices of the following types:  $\alpha$ -, 310-, and  $\pi$ - helix.

2. A helical pair is a portion of the protein structure that is formed by two helices adjacent to the polypeptide chain connected by an unstructured fragment (coil) of arbitrary length and configuration.

3. The following types of helical pairs are found in proteins:  $\alpha$ - $\alpha$ -corner,  $\alpha$ - $\alpha$ -hairpin, and V- and L-structures.

4. Each helix of a pair can be inscribed in a cylinder. The cylinder axis is determined by the least squares method with the minimum deviation of the helix parameters from the ideal helix.

5. Two axes of the helices determine the spatial structure of the helical pair. The spatial structure of the helical pair is determined by the interplanar distance ( $d$ ) and the projections of the axes of the helices on the plane.

6. The polygon area ( $S$ ) of the intersection of the projections of the helices depends on the relative position of the helices (geometry) and the diameter of the cylinders of the helices. The cylinder diameter is determined by the type of helix and the average size of the side chain.

### **Distinctive properties of types of helical pairs**

#### **1. $\alpha$ - $\alpha$ corner**

- dense orthogonal packing of helices;
- the projections of the helices and the axis of the helices intersect;
- the values of the interplanar distance and the minimum distance are equal ( $d = r$ );
- the intercoil distance within  $d \approx (9-11) \text{ \AA}$ ;
- the area and perimeter of the intersection of the projections of the helices are larger than the other two- helix motifs;
- torsion  $\theta$  and flat  $\varphi$  inter helix angles have values of 70-110°;

- the length of the waist ( $N_p$ ) is usually 4-5 amino acids.

## 2. $\alpha$ - $\alpha$ -hairpin

- dense antiparallel packing of helices;
- the projections of the helices intersect, but the axes of the helices do not intersect;
- the range of the intercoil distance,  $d$ , is wider than that of the  $\alpha$ - $\alpha$  corner;
- interhelix angles  $0$ - $30^\circ$ ;
- the length of the waist ( $N_p$ ) may vary;
- usually consists of two long helices.

## 3. L-shaped structure

- dense bevelled packaging of helices;
- the projections of the helices and the axes of the helices do not intersect ( $S \approx 0 \text{ \AA}^2$ ,  $P \approx 0 \text{ \AA}$ );
- interhelix angle has a value  $\varphi \approx 90^\circ$ ;
- The waist length ( $N_p$ ) may vary.

## 4. V-shaped structure

- dense bevelled packaging of helices;
- the projections of the helices (most often) and the axes of the helices do not intersect;
- area,  $S$ , and perimeter have values close to zero;
- interhelix angle  $\varphi \approx 50^\circ$ ;
- coil ( $N_p$ ), as a rule, consists of one amino acid;
- length of the helices, as a rule, does not exceed three to four turns.

**Table S2. Data of colon cancer patients**

<b>N</b>	<b>Sex</b>	<b>Age, years</b>	<b>Diagnosis</b>	<b>Stage</b>	<b>Metastases</b>
2	Female	56	colorectal cancer	IV	in the liver
3	Female	59	colorectal cancer	IIIc	No
4	Female	28	colorectal cancer	IIa	No
5	Female	69	colorectal cancer	IV	in the ovaries
6	Female	54	colorectal cancer	IV	in the liver
8	Male	85	colorectal cancer	IV	in the liver
9	Male	63	colorectal cancer	IV	in the lymph nodes
10	Female	43	colorectal cancer	I	No
11	Female	76	colorectal cancer	IIa	No
12	Male	46	colorectal cancer	IIa	No
13	Female	64	colorectal cancer	I	No
14	Female	41	colorectal cancer	IVb	in the liver
15	Female	69	colorectal cancer	IIIa	No
16	Female	81	colorectal cancer	I	No
18	Female	32	colorectal cancer	IIIa	No
19	Male	50	colorectal cancer	I	No
20	Female	60	colorectal cancer	I	No
21	Female	31	colorectal cancer	IIa	No
22	Female	55	anal canal	I	No
23	Female	64	colorectal cancer	IIIc	No
24	Male	45	colorectal cancer	IV	in the lymph nodes
25	Male	58	colorectal cancer	IIa	No
26	Female	60	colorectal cancer	IIa	No
28	Male	93	colorectal cancer	IIa	No
30	Female	70	colorectal cancer	IIa	No
31	Female	64	colorectal cancer	IIIb	No
42	Male	81	colorectal cancer	I	No
54	Male	58	colorectal cancer	IIIb	No

**Data of healthy volunteers**

<b>Sample</b>	<b>Sex</b>	<b>Age, years</b>
1	Female	53
2	Female	38
3	Female	58
4	Female	49
5	Male	70
6	Female	41
7	Male	61

8	Female	45
9	Male	46
10	Male	41
11	Female	55
12	Female	60
13	Female	56
14	Female	52
15	Male	59
16	Male	62
17	Male	39
18	Female	40
20	Female	39
22	Male	44
24	Female	46
25	Male	55
26	Female	42
28	Male	40
29	Male	44
30	Male	40
31	Male	44
32	Male	40
35	Female	48
37	Female	39
38	Male	39
42	Male	42
43	Male	39
44	Male	34
46	Male	53
47	Female	53
50	Female	47
51	Female	52
60	Male	64
61	Female	49
62	Female	45

The recorded raw data were converted to peak lists for search engines and deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD015163.

**Table S3**

UniProt AC	Protein name	Peptide*	ID PDB	Localization**	PTM	Number of samples	Stage***
		LKCASLQK	1AO6	199	ac (K)	11	I;II;IV
ALBU	Serum albumin	ICIAEVENDEMPADLPSLAADFVESKDVC	1AO6	313	ac (K)	5	II-IV
		ADLAKYICENQDSISSK	1AO6	262	ac (K)	2	IV
		LVNEVTEFAKTCVADESAENCDK	1AO6	51	ac (K)	3	I; II; IV
		YICENQDSISSK	1AO6	263	p (Y)	2	II; IV
VTDB	Vitamin D-binding protein	VLEPTLTKSLGECCDVEDSTTCFNAK	1J78	354	ac (K)	1	II
		SCESNSPFPVHPGTAECCTKEGLER	1J78	114	ac (K)	1	I
A2MG	Alpha-2-macroglobulin	SKAIGYLNTGYQR	4ACQ	1003	ac (K)	1	III
CO4A	Complement C4-A X-ray repair cross-complementing protein 6	LLATLCSAEVCQCAEGKCPR	5JPN	1594	ac (K)	1	I
XRCC6		IISDRDLLAVVFGTEK	1JEQ	77	p (S)	1	IV
IC1	Plasma protease C1 inhibitor	LVLLNAIYLSAK	5DU3	272	p (Y)	2	I; IV

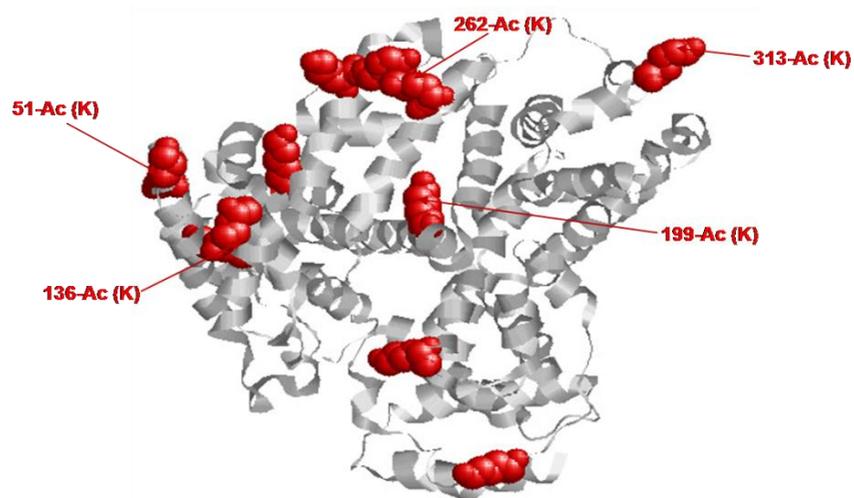
"Peptide\*" indicates tryptic peptides

"Localization\*\*" indicates locus aminoacid with PTM in the protein sequence (ID PDB)

"Stage\*\*\*" indicates the stage of development of the cancer from I to IV

**Figure S1** Three-dimensional protein structures extracted from the PDB DB with indicating spatial localizations of modified amino acids (highlighted in red and arrows in the figure). The insets show the corresponding proteins geometric of a supersecondary structure formed by two helices. (A) ALBU (PDB ID 1AO6) with 5 modified amino acids with lysine acetylation (Ac(K)) and tyrosine phosphorylation (P(Y)); (B) A2MG (PDB ID 4ACQ), (C) IC1 (PDB ID 5DU3), (D) VTDB (PDB ID 1J78), (E) CO4A (PDB ID 5JPN) with lysine acetylation; (F) XRCC-6 (PDB ID 1JEQ) with phosphorylated serine (P(S)).

(A) ALBU



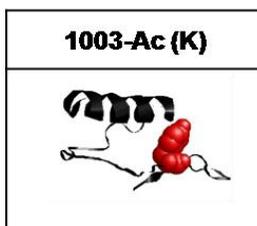
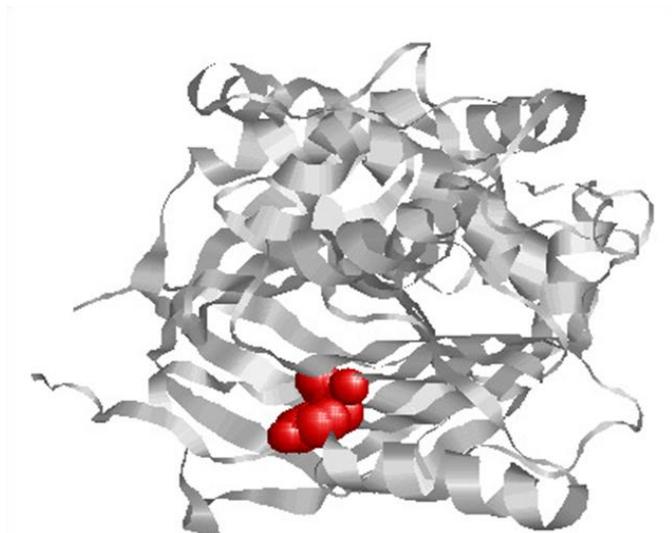
51-Ac (K)	263-P (Y)	199-Ac (K)	313-Ac (K)	262-Ac (K)

(B) A2MG

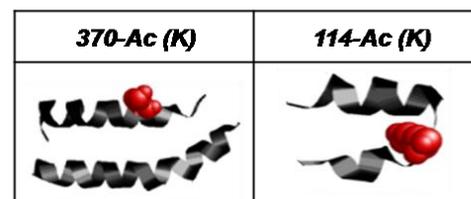
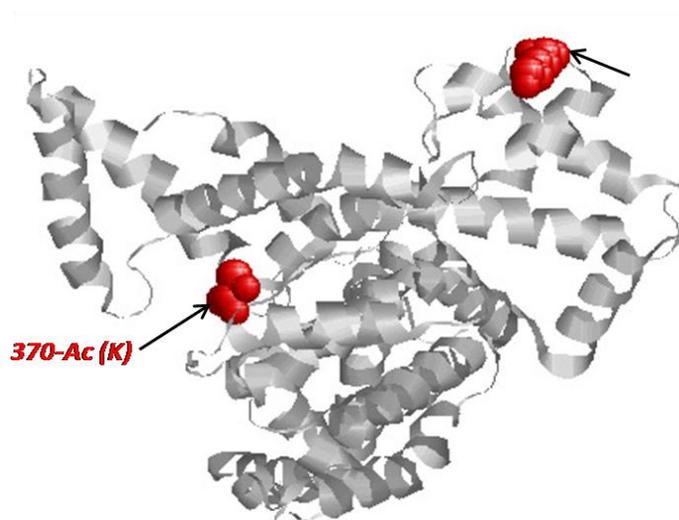


1003-Ac (K)

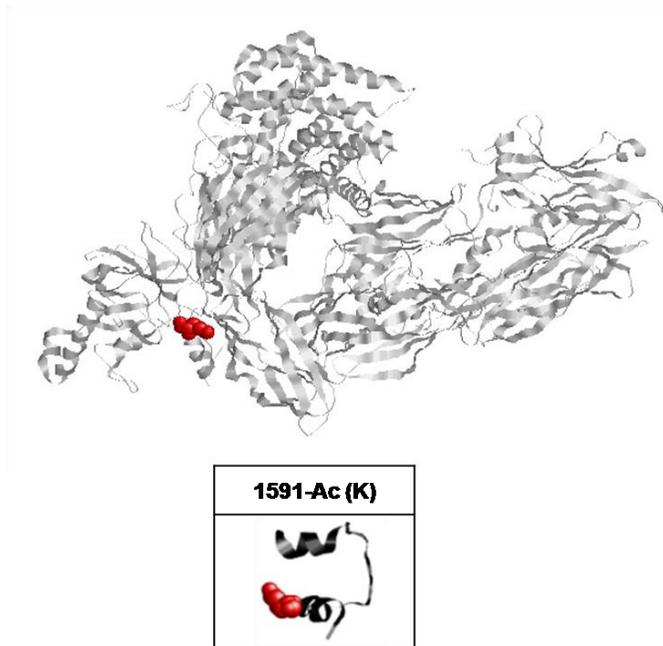
(C) IC1



(D) VTDB



(E) CO4A



(F) XRCC-6

