Table S3: Preferential cleavage sites for enzymes and chemicals over the substrate protein is given in the table below.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Enzyme name | P4 | P3 | P2 | P1 | P1' | P2' | P3' | P4' |
| Arg-C proteinase | - | - | - | R | - | - | - | - |
| Asp-N endopeptidase | - | - | - | - | D | - | - | - |
| BNPS-Skatole | - | - | - | W |  | - | - | - |
| Caspase 1 | F, W, Y, or L | - | H, A or T | D | not P, E, D, Q, K or R | - | - | - |
| Caspase 2 | D | V | A | D | not P, E, D, Q, K or R | - | - | - |
| Caspase 3 | D | M | Q | D | not P, E, D, Q, K, or R | - | - | - |
| Caspase 4 | L | E | V | D | not P, E, D, Q, K, or R | - | - | - |
| Caspase 5 | L or W | E | H | D |  | - | - | - |
| Caspase 6 | V | E | H or I | D | not P, E, D, Q, K or R | - | - | - |
| Caspase 7 | D | E | V | D | not P, E, D, Q, K or R | - | - | - |
| Caspase 8 | I or L | E | T | D | not P, E, D, Q, K or R | - | - | - |
| Caspase 9 | L | E | H | D | - | - | - | - |
| Caspase 10 | I | E | A | D | - | - | - | - |
| Chymotrypsin-high specificity (C-term to [FYW], not before P) | - | - | - | F or Y | not P | - | - | - |
| - | - | - | W | not M or P | - | - | - |
| - | - | - | F, L or Y | not P | - | - | - |
| Chymotrypsin-low specificity (C-term to [FYWML], not before P) | - | - | - | W | not M or P | - | - | - |
| - | - | - | M | not P or Y | - | - | - |
| - | - | - | H | not D, M, P or W | - | - | - |
| Clostripain (Clostridiopeptidase B) | - | - | - | R | - | - | - | - |
| CNBr | - | - | - | M | - | - | - | - |
| Enterokinase | D or E | D or E | D or E | K | - | - | - | - |
| Factor Xa | A, F, G, I, L, T, V or M | D or E | G | R | - | - | - | - |
| Formic acid | - | - | - | D | - | - | - | - |
| Glutamyl endopeptidase | - | - | - | E | - | - | - | - |
| GranzymeB | I | E | P | D | - | - | - | - |
| Hydroxylamine | - | - | - | N | G | - | - | - |
| lodosobenzoic acid | - | - | - | W | - | - | - | - |
| LysC | - | - | - | K | - | - | - | - |
| Neutrophil elastase | - | - | - | A or V | - | - | - | - |
| NTCB (2-nitro-5-thiocyanobenzoic acid) | - | - | - | - | C | - | - | - |
| Pepsin (pH1.3) | - | not H, K or R | not P not P | not R | F or L | not P | - | - |
| - | not H, K or R | not P not P | F or L | - | not P | - | - |
| Pepsin (pH>2) | - | not H, K or R | not P | not R | F, L, W or Y | not P | - | - |
| - | not H, K or R | not P | F, L, W or Y | - | not P | - | - |
| Proline-endopeptidase | - | - | H, K or R | P | not P | - | - | - |
| Proteinase K | - | - | - | A, E, F, I, L, T, V, W or Y | - | - | - | - |
| Staphylococcal peptidase I | - | - | not E | E | - | - | - | - |
| Thermolysin | - |  |  | not D or E | A, F, I, L, M or V | - | - | - |
| Thrombin | - | - | G | R | G | - | - | - |
| A, F, G, I, L, T, V or M | A, F, G, I, L, T, V, W or A | P | R | not D or E | not D or E | - | - |
| Trypsin\* |  |  |  | K or R | not P | - | - | - |
| - | - | W | K | P | - | - | - |
| - | - | M | R | P | - | - | - |
| Elastase 1 | - | - | - | A | - | - | - | - |
| Neutrophil elastase (Elastase 2) | - | - | - | V or A | - | - | - | - |
| Chymotrypsinogen B1 | - | - | - | F, L, W or Y | - | - | - | - |
| Chymotrypsin C | - | - | - | F, M, Y, W, L, N or Q | - | - | - | - |
| Pancreatic Endopeptidase E | - | - | - | A | - | - | - | - |
| Enteropeptidase | - | - | - | K | I | - | - | - |
| Prostasin | R | K | R | K | I | S | G | K |
| Gastricsin | - |  |  | Y | - | - | - | - |
| Fruit Bromelain | - | F | V | R | - | - | - | - |
| Stem Bromelain | - |  | R | R | - | - | - | - |
| Ananain | - | F | v | R | - | - | - | - |
| Papaya Proteinase | - |  |  | G | - | - | - | - |
| Chymopapain | - | - | A, V, L, I, F or W | R or L | - | - | - | - |
| Chymosin | - | - |  | S or F | M or A | - | - | - |
| Caricain | - | - | A, V, L, I, F or W | R or L | - | - | - | - |

Note: - means there is no specificity in this position. \*When the above cleavage sites do not occur, Tyrpsin will use the below cleavage sites to break the substrates.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Enzyme name | P4 | P3 | P2 | P1 | P1' | P2' | P3' | P4' |
| Trypsin | - | - | C or D | K | D | - | - | - |
| - | - | C | K | H or Y | - | - | - |
| - | - | C | R | K | - | - | - |
| - | - | R | R | H or R | - | - | - |

The amino acids on the substrate were named as P1-P4 on the N terminal of the point of cleavage and P1`-P4` on the C terminal of the point of cleavage. The cleavage process was shown as the below graph.

