possibilities. The subunits can be expressed separately and then mixed after purification, which makes their production very complex [19]. Or, the procaspase is expressed which causes autocatalytic activation. This process, however, is often not complete [20], therefore enzyme activity can vary between batches [21]. Furthermore, as caspases are active in *E. coli* they can also cleave bacterial proteins [22] and negatively influence growth and yield. In addition, the substrate specificities of both caspase-3 [23] and caspase-6 [24] have been described as rather promiscuous. They are very likely to cleave fusion proteins at undesired sites.

Therefore, there is an urgent need for an industrially applicable platform technology which enables efficient and specific tag removal to improve purification of recombinant protein products.

SUMMARY ~~OF THE INVENTION~~

In the production of recombinant proteins, processing of fusion proteins with state-of-the-art enzymes to remove tags often generates a non-authentic N- or C-terminus since these enzymes lack specificity. Such lack of specificity can also lead to unspecific cleavage or proteolytic degradation of the protein of interest.

It is the objective of the present ~~invention~~disclosure to provide an improved system for the production of recombinant proteins employing a modified caspase-2, specifically a circular permuted caspase-2.

The objective is solved by the present ~~invention~~disclosure.

Specifically provided herein is a modified caspase-2, specifically a circularly permuted caspase-2, with significantly improved P1' tolerance. The caspase-2 variants provided herein are specifically used in the production of recombinant proteins to generate a protein of interest comprising an authentic N-terminus by target specific cleavage of N-terminal tags.

According to the ~~invention~~disclosure, there is provided a single-chain circular permuted caspase-2 (cp caspase-2) comprising the following structure from N- to C-terminus:

i. a small subunit of a caspase-2, or a functionally active variant thereof; and

ii. a large subunit of a caspase-2, or a functionally active variant thereof,

wherein said cp caspase-2 comprises one or more amino acid substitutions increasing P1' tolerance of said cp caspase-2 compared to a cp caspase-2 without said amino acid substitutions.

Specifically, the cp caspase-2 provided herein is catalytically active, specifically upon dimerization. Specifically, the cp caspase-2 described herein is catalytically active and capable of catalyzing peptide bond cleavage upon dimerization. Specifically, the cp caspase-2 described herein is a single-chain caspase-2 that does not require cleavage by initiator caspases for activation.

Specifically, the caspase-2 or cp caspase-2 provided herein is a functionally active variant of wild-type caspase-2 comprising improved P1' tolerance. Specifically, said functionally active variant is capable of cleaving a substrate with high efficiency and specificity.

Specifically, provided herein is a single chain caspase-2 comprising the following structure from N- to C-terminus:

i. a small subunit of a caspase-2 comprising SEQ ID NO:3, or a functionally active variant thereof comprising SEQ ID No. 91, SEQ ID No. 94, SEQ ID No. 97, SEQ ID No. 100, SEQ ID No. 103, SEQ ID No. 106, SEQ ID No. 109, SEQ ID No. 112, SEQ ID No. 115, or SEQ ID No. 118 and optionally up to 8, 9, or 10 amino acid subsitutions, insertions and/or deletions; and

ii. a large subunit of a caspase-2 comprising SEQ ID NO:4, or a functionally active variant thereof comprising SEQ ID No. 90, SEQ ID No. 93, SEQ ID No. 96, SEQ ID No. 99, SEQ ID No. 102, SEQ ID No. 105, SEQ ID No. 108, SEQ ID No. 111, SEQ ID No. 114, or SEQ ID No. 117 and optionally up to 8, 9, or 10 amino acid subsitutions, insertions and/or deletions,

wherein said single chain caspase-2 comprises one or more amino acid substitutions increasing proteolytic activity of said single chain caspase-2 compared to a caspase-2 comprising the same sequence as said single chain caspase-2 but without said amino acid substitutions. Optionally, said single chain caspase-2 comprises one or more further amino acid substitutions, insertions or deletions.

Specifically, said variant is of animal origin, specifically of mammalian, reptile or fish origin, more specifically from human, marsupial, iguana or cartilaginous fish, ghost shark or tasman devil origin.

According to a specific ~~embodiment~~example, the cp caspase-2 provided herein comprises one or more amino acid substitutions at positions 171, 105, 172, 282, 225, 83, 185, 255, or 285 of SEQ ID No. 6 or at a position functionally equivalent to any of positions 171, 105, 172, 282, 225, 83, 185, 255, or 285 of SEQ ID No. 6 or any combination thereof.

Specifically, a cp caspase-2 comprising one or more of said amino acid substitutions, also referred to as “cp caspase-2 variant”, comprises improved P1’ tolerance compared to a cp caspase-2 not comprising said substitutions. Specifically, the cp caspase-2 variants provided herein comprise improved P1’ tolerance for at least one amino acid other than glycine.

Thus, in one aspect, the present disclosure provides a single-chain circular permuted caspase-2 (cp caspase-2) comprising the following structure from N- to C-terminus:

i. a small subunit of a caspase-2 comprising a sequence selected from the group consisting of SEQ ID No. 3, SEQ ID No. 91, SEQ ID No. 94, SEQ ID No. 97, SEQ ID No. 100, SEQ ID No. 103, SEQ ID No. 106, SEQ ID No. 109, SEQ ID No. 112, SEQ ID No. 115, SEQ ID No. 118, or a functionally active variant thereof comprising a sequence with at least 70% sequence identity to SEQ ID No. 3, SEQ ID No. 91, SEQ ID No. 94, SEQ ID No. 97, SEQ ID No. 100, SEQ ID No. 103, SEQ ID No. 106, SEQ ID No. 109, SEQ ID No. 112, SEQ ID No. 115, or SEQ ID No. 118; covalently linked, either via a linker or directly, to

ii. a large subunit of a caspase-2 comprising a sequence selected from the group consisting of SEQ ID No. 4, SEQ ID No. 90, SEQ ID No. 93, SEQ ID No. 96, SEQ ID No. 99, SEQ ID No. 102, SEQ ID No. 105, SEQ ID No. 108, SEQ ID No. 111, SEQ ID No. 114, SEQ ID No. 117, or a functionally active variant thereof comprising a sequence with at least 70% sequence identity to SEQ ID No. 4, SEQ ID No. 90, SEQ ID No. 93, SEQ ID No. 96, SEQ ID No. 99, SEQ ID No. 102, SEQ ID No. 105, SEQ ID No. 108, SEQ ID No. 111, SEQ ID No. 114, or SEQ ID No. 117,

wherein said cp caspase-2 is a protease and with the proviso that in case of a selection of SEQ ID No. 3 for the small subunit and SEQ ID No. 4 for the large subunit, then the cp caspase 2 comprises one or more amino acid substitutions at a position functionally equivalent to any of positions 171, 105, 172, 282, 225, 83, 185, 255, or 285 of SEQ ID No. 6.

According to a further specific ~~embodiment~~example, the cp caspase-2 provided herein comprises a propeptide of a small caspase-2 subunit (SS propeptide), fused to the N-terminus of the small subunit. Specifically, the SS propeptide comprises one or more amino acid substitutions at the C-terminus of the SS propeptide. Specifically, the SS propeptide of the cp caspase-2 described herein is modified to prevent cleavage at its C-terminus. Specifically, the SS propeptide comprises an amino acid substitution at position Asp^{14} of SEQ ID No. 2 or at a position functionally equivalent to Asp^{347} of SEQ ID No. 11, specifically Asp is substituted to Ala.

According to a specific ~~embodiment~~example, the SS propeptide described herein comprises the amino acid sequence of SEQ ID No. 2, wherein X can be any amino acid except D or E, specifically it is A, or a variant thereof having 1, 2, 3, 4, or 5 point mutations or deletions. Specifically, said variant is a functionally active variant. Preferably, the SS propeptide sequence comprises the amino acid sequence of SEQ ID No. 2, wherein X is not D or E.

According to a further specific ~~embodiment~~example, the cp caspase-2 provided herein comprises one or more linker sequences, specifically consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 or even more amino acid residues. Specifically, the linker can comprise more than 20 or 30 or even more amino acids, as long as the caspase retains its functional activity as described herein. Specifically, the linker sequence comprises glycine, alanine and/or serine residues. Specifically, the linker comprises at least one glycine and serine residue, more specifically the linker is GS, GSG, GGSGG, GSGSGSGS and/or GSAGSAAGSG.

Specifically, the cp caspase-2 comprises a subunit-linker sequence, which is a linker sequence between the small subunit and the large subunit of the cp caspase-2 described herein.

According to a further specific ~~embodiment~~example, the cp caspase-2 provided herein comprises one or more C-terminal or N-terminal tags, specifically selected from the

group consisting of affinity tags, solubility enhancement tags and monitoring tags. Specifically, any tag known in the art can be fused to the cp caspase-2.

Specifically, the affinity tag is selected from the group consisting of poly-histidine tag, poly-arginine tag, peptide substrate for antibodies, chitin binding domain, RNAse S peptide, protein A, ß-galactosidase, FLAG tag, Strep II tag, streptavidin-binding peptide (SBP) tag, calmodulin-binding peptide (CBP), glutathione S-transferase (GST), maltose-binding protein (MBP), S-tag, HA tag, c-Myc tag, SUMO tag, E.coli thioredoxin, NusA, chitin binding domain CBD, chloramphenicol acetyl transferase CAT, LysRS, ubiquitin, calmodulin, and lambda gpV, specifically the tag is a His tag comprising one or more His, more specifically it is a hexahistidine tag.

Specifically, the solubility enhancement tag is selected from the group consisting of T7C, T7B, T7B1, T7B2, T7B3, T7B3, T7B4, T7B5, T7B6, T7B6, T7B7, T7B8, T7B9, T7B10, T7B11, T7B12, T7B13, T7A, T7A1, T7A2, T7A3, T7A4, T7A5, T7AC, T3, N1, N2, N3, N4, N5, N6, N7, calmodulin-binding peptide (CBP), poly Arg, poly Lys, G B1 domain, protein D, Z domain of Staphylococcal protein A, DsbA, DsbC and thioredoxin.

Preferably, the solubility enhancement tag is selected from the group consisting of T7A3 tag and T7AC tag.

Specifically, the monitoring tag is selected from the group consisting of m-Cherry, GFP and f-Actin.

According to a specific ~~embodiment~~example, the cp caspase-2 described herein comprises more than one tag sequences, specifically it comprises an affinity tag and a solubility enhancement tag. Specifically, it comprises an affinity tag, a solubility enhancement tag and a monitoring tag. Specifically, it comprises more than one tag of the same functionality, specifically it comprises more than one affinity tag, more than one solubility enhancement tag and/or more than one monitoring tag, and any combination thereof. Specifically, the cp caspase-2 described herein comprises a C-terminal and an N-terminal tag, each comprising one or more tag sequences, preferably selected from affinity tag, solubility enhancement tag and monitoring tag

Specifically, the affinity tag is a hexahistidine tag and the solubility enhancement tag is a T7AC tag.

According to a further specific ~~embodiment~~example, the cp caspase-2 provided herein comprises a tag-linker sequence, which is a linker sequence between two tags or a tag and the small subunit, the large subunit or the SS propeptide of the cp caspase-2.

Specifically, the cp caspase-2 provided herein comprises one or more N-terminal tags and optionally one or more tag-linker sequences between the tags or between a tag and the N-terminus of the small subunit or the SS propeptide. Specifically, the cp caspase-2 provided herein comprises one or more C-terminal tags and optionally one or more tag-linker sequences, which are linker sequences between the tags or between a tag and the C-terminus of the large subunit.

According to a further ~~embodiment~~example, herein provided is a functionally active variant of the cp caspase-2 or caspase-2, wherein

i. the small subunit of a caspase-2 comprises
   a) a first conserved region of the active center with at least 37.5 % amino acid sequence identity to SEQ ID No. 177 (1st consensus: AAMRNTKR) or 100% sequence identity to XXXRNTXX (SEQ ID No. 200), wherein X is any amino acid,
   b) a second conserved region of the active center with at least 61.5 % amino acid sequence identity to SEQ ID No. 178 (2nd consensus: EGYAPGTEFHRCK) or 100% sequence identity to EGXXPGXXXHRCK (SEQ ID No. 194), wherein X is any amino acid, and

ii. the large subunit of a caspase-2 comprises
   a) a third conserved region of the active center with at least 25.0 % amino acid sequence identity to SEQ ID No. 174 (3rd consensus: G-EKDLEFRSGGDVDH) or 100% sequence identity to X-XXXLXXRXGXXXDX (SEQ ID No. 195), wherein X is any amino acid,
   b) a fourth conserved region of the active center with at least 53.3 % amino acid sequence identity to SEQ ID No. 175 (4th consensus: LLSHGVEGGXYGVDG) or 100% sequence identity to XXSHGXXGXXYGXDG (SEQ ID No. 196), wherein X is any amino acid, and
   c) a fifth conserved region of the active center with at least 50.0 % amino acid sequence identity to SEQ ID No. 176 (5th consensus: QACRGDET) or 100% sequence identity to QACXGXXX (SEQ ID No. 197), wherein X is any amino acid.

According to a specific ~~embodiment~~example, the cp caspase-2 provided herein comprises an N-terminal and/or C-terminal truncation of at least 1, 2, 3, 4, 5 and up to 10 or even more, as long as the caspase retains its functional activity as described herein. According to a further specific ~~embodiment~~example, the cp caspase-2 provided herein comprises an N-terminal and/or C-terminal extension of at least 1, 2, 3, 4, 5 and up to 10 or even more, as long as the caspase retains its functional activity as described herein. Specifically, the cp caspase-2 may comprise a truncation and an extension.

Specifically, the small subunit of the cp caspase-2 described herein comprises the amino acid sequence of SEQ ID No. 3, SEQ ID No. 91, SEQ ID No. 94, SEQ ID No. 97, SEQ ID No. 100, SEQ ID No. 103, SEQ ID No. 106, SEQ ID No. 109, SEQ ID No. 112, SEQ ID No. 115, SEQ ID No. 118 or a functionally active variant thereof comprising at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97 or 98 % sequence identity. Specifically, the large subunit of the cp caspase-2 described herein comprises the amino acid sequence of SEQ ID No. 4, SEQ ID No. 90, SEQ ID No. 93, SEQ ID No. 96, SEQ ID No. 99, SEQ ID No. 102, SEQ ID No. 105, SEQ ID No. 108, SEQ ID No. 111, SEQ ID No. 114, SEQ ID No. 117, or a functionally active variant thereof comprising at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97 or 98 % sequence identity.

According to a specific ~~embodiment~~example, the cp caspase-2 variant provided herein comprises one or more amino acid substitutions, selected from

i. Gly^{171}, substituted with D or an amino acid selected from the group consisting of R, K, E, Q, N, A, S, T, P, H, Y;
ii. Glu^{105}, substituted with V or an amino acid selected from the group consisting of C, L, I, M, F, W, R, K, D, Q, N;
iii. Glu^{172}, substituted with V or an amino acid selected from the group consisting of C, L, I, M, F, W, R, K, D, Q, N;
iv. Asp^{282}, substituted with E or T or an amino acid selected from the group consisting of R, K, Q, N, G, A, S, P, H, Y;
v. Val^{225}, substituted with G or an amino acid selected from the group consisting of A, S, T, P, H, Y, C, L, I, M, F, W;
vi. Lys^{83}, substituted with E or an amino acid selected from the group consisting of R, D, Q, N,
vii. His^{185}, substituted with A or an amino acid selected from the group consisting of G, S, T, P, Y;

Specifically, the cp caspase-2 variant ms9 ProD comprising E105V, G171D, V225G and D282E subsitutions displays excellent P1' tolerance.

Specifically, the cp caspase-2 variant E105V G171D comprising E105V and G171D substitutions displays excellent P1' tolerance, which is increased compared to the cp caspase-2 variant ms9 ProD. Specifically, the highest tolerance of the cp caspase-2 variant E105V G171D is for the amino acid residue proline.

According to a specific ~~embodiment~~example, the cp caspase-2 described herein comprises at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97 or 98 % sequence identity to SEQ ID No. 9 (*Homo sapiens*), SEQ ID No. 64 (*Sarcophilus harrisii,* Tasmanian Devil), SEQ ID No. 66 (*Anolis carolinensisilus*), SEQ ID No. 68 (*Callorhinchus milii,* Ghost Shark), SEQ ID No. 76 or SEQ ID No. 77 (Homo sapiens) and comprises one or more amino acid substitutions at a position functionally equivalent to any of positions 171, 105, 172, 282, 225, 83, 185, 255, or 285 of SEQ ID No. 6 or any combination thereof.

Specifically, the cp caspase-2 variant described herein comprises SEQ ID No. 6 and one or more amino acid substitutions at position 171, 105, 172, 282, 225, 83, 185, 255, or 285 of SEQ ID No. 6 or at a position functionally equivalent to position 171, 105, 172, 282, 225, 83, 185, 255, or 285 of SEQ ID No. 6, or any combination thereof.

Specifically, the cp caspase-2 variant described herein comprises any one or more of amino acid substitutions G171D, E105V, E172V, D282E, D282T, V225G, K83E, H185A, V255M, D285Y and D285E, with reference to the numbering according to SEQ ID No. 6.

According to a further specific ~~embodiment~~example, the cp caspase-2 variant described herein comprises SEQ ID No. 6 or has at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97 or 98 % sequence identity with SEQ ID No. 6, and comprises amino acid substitutions E105V, G171D, V225G, D282E, and/or D285E, with reference to the numbering of SEQ ID No. 6, wherein said cp caspase-2 has increased P1' tolerance.

According to a further specific ~~embodiment~~example, the cp caspase-2 variant described herein comprises SEQ ID No. 6 or has at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97 or 98 % sequence identity with SEQ ID No. 6, and comprises amino acid substitutions K83E, E105V, E172V, V255M and/or D285Y, with reference to the numbering of SEQ ID No. 6, wherein said cp caspase-2 has increased P1' tolerance.

According to a further specific ~~embodiment~~example, the cp caspase-2 variant described herein comprises SEQ ID No. 6 or has at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96,

97 or 98 % sequence identity with SEQ ID No. 6, and comprises amino acid substitutions H185A and/or D282T, with reference to the numbering of SEQ ID No. 6, wherein said cp caspase-2 has increased P1' tolerance, specifically for branched P1' amino acid residues.

Specifically, the cp caspase-2 variant described herein comprises an amino acid sequence selected from the group consisting of SEQ ID No. 1, 13, 17, 18, 23, 24, 51, 52, 54, 70, 71, 72, 78, 79, 86, 87, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191 and 192 or an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, specifically at least 95%, specifically at least 99% sequence identity with any one of SEQ ID No. 1, 13, 17, 18, 23, 24, 51, 52, 54, 70, 71, 72, 78, 79, 86, 87, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191 and 192.

According to a specific ~~embodiment~~example, the cp caspase-2 described herein comprises a C-terminal tag and an amino acid substitution at positions 285 and 292 of SEQ ID No. 6 or at a position functionally equivalent to positions 285 and 292 of SEQ ID No. 6, specifically comprising substitutions to Glu and Ser, respectively (D285E and D292S).

Specifically, the caspase-2 or cp caspase-2 described herein is recruited by a recognition site for proteolytic cleavage, comprising 5 amino acids of the sequence P5 P4 P3 P2 P1, wherein

P1 can be any amino acid, preferably it is D or E,

P2 can be any amino acid, preferably it is A,

P3 can be any amino acid, preferably it is V,

P4 can be any amino acid, preferably it is D, and

P5 can be any amino acid, preferably it is V.

Specifically, the caspase-2 variant or cp caspase-2 variant described herein has increased specificity, specifically to the recognition site VDVAD wherein P5 is V, P4 is D, P3 is V, P2 is A and P1 is D, compared to wild-type (wt) caspase-2 comprising the amino acid sequence of SEQ ID No. 11.

Specifically, the caspase-2 or cp caspase-2 described herein recognizes or can be further modified to recognize a variety of recognition sites. According to a specific example, the caspase-2 variant or cp caspase-2 variant described herein recognizes any one or more of the recognitions sites LDESD, DVAD, DEVD, DEVE, ADVAD,

VDTTD, DTTD, DVPD, VDVPD, VDQQD, or TDTSD. Preferably, the caspase-2 or cp caspase-2 described herein recognizes and has high specifity for one recognition site.

According to a further specific example, the variants of caspase-2 or cp caspase-2 described herein recognizes the recognition site DRKD, DAVD, VKVD, DTLD, EEPD, DETD, DATD, NKVD, DALD, DSVD, NAID, DKPD, IQLD, DNAD, DVVD, ENPD, DMAD, DLID, DGAD, DVKD, GYND, ELPD, DSTD, DRQD, HAVD, QERLD, LERD, MMPD, EEPD, VESID, EAMD, EDAD, EEED, AVLD, and/or EEGD.

According to a further specific example, the variants of caspase-2 or cp caspase-2 described herein recognizes the recognition site TDTSD, LDEPD, and/or KDEVD.

Specifically, the recognition site can be selected from the group consisting of DEXD (SEQ ID No. 202) and DVXD (SEQ ID No. 203), wherein X is any amino acid.

Specifically, the recognition site comprises the sequence P5 P4 P3 P2 P1, wherein P5 is V, P4 is D, P3 is Q, P2 is Q and P1 is D. Specifically, the V on position P5 can be replaced with I, Y, L, T, N, or A, and/or the D on position P4 can be replaced with S, and/or the Q on position P3 can be replaced with V, E or T, and/or the Q on position P2 can be replaced with A, S, K, V, M, or L

Testing of a recognition site library (P4 – P1) for Caspase 2 resulted in the following predominant amino acids: position P4: D, V; position P3: V, E, T; position P2: S, T and position P1: D.

Specifically, the caspase-2 or cp caspase-2 described herein is capable of cleaving at a cleavage site P1/P1', wherein P1' can be any amino acid.

Further provided herein is a caspase-2 comprising one or more amino acid substitutions at positions 212, 431, 213, 323, 266, 409, 226, 296 or 326 of SEQ ID No. 11 or at a position functionally equivalent to any of positions 212, 431, 213, 323, 266, 409, 226, 296 or 326 of SEQ ID No. 11 or a combination thereof, also referred to as "caspase-2 variant", and wherein said amino acid substitution increases P1' tolerance compared to a caspase-2 which has the same sequence but does not comprise said substitutions. In other words, the caspase-2 that the caspase-2 variant is compared to has an identical sequence as the caspase-2 variant, except that it does not comprise any of the amino acid subsitutions at positions 212, 431, 213, 323, 266, 409, 226, 296 or 326 of SEQ ID No. 11, or at a position functionally equivalent to any of positions 212, 431, 213, 323, 266, 409, 226, 296 or 326 of SEQ ID No. 11, which increase the P1' tolerance according to the ~~invention~~disclosure.

Specifically, the caspase-2 variant provided herein comprises improved P1' tolerance for at least one amino acid other than glycine compared to a caspase-2 not comprising the respective amino acid substitution.

Specifically, the caspase-2 variant described herein comprises at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97 or 98 % sequence identity to SEQ ID No. 11, SEQ ID No. 89, SEQ ID No. 92, SEQ ID No. 95, SEQ ID No. 98, SEQ ID No. 101, SEQ ID No. 104, SEQ ID No. 107, SEQ ID No. 110, SEQ ID No. 113 or SEQ ID No. 116 and comprises one or more amino acid substitutions at positions 409, 431, 212, 213, 266, 296, 226, 323 or 326 of SEQ ID No. 11 or at a position functionally equivalent to any of positions 409, 431, 212, 213, 266, 296, 323 or 326 of SEQ ID No. 11 or a combination thereof. Thus, in another aspect, the present disclosure provides a caspase-2 comprising a sequence selected from the group consisting of SEQ ID No.11, SEQ ID No. 89, SEQ ID No.92, SEQ ID No. 95, SEQ ID No. 98, SEQ ID No. 101, SEQ ID No. 104, SEQ ID No. 107, SEQ ID No. 110, 113, and SEQ ID No. 116, or a functionally active variant thereof comprising a sequence with at least 70% sequence identity to SEQ ID No.11, SEQ ID No. 89, SEQ ID No.92, SEQ ID No. 95, SEQ ID No. 98, SEQ ID No. 101, SEQ ID No. 104, SEQ ID No. 107, SEQ ID No. 110, 113, and SEQ ID No. 116, and comprising one or more amino acid substitutions at positions 409, 431, 212, 213, 266, 226, 296, 323 or 326 of SEQ ID No. 11, or at a position functionally equivalent to any of positions 409, 431, 212, 213, 266, 226, 296, 323 or 326 of SEQ ID No. 11, or a combination thereof.

Specifically, the caspase-2 described herein comprises at least a small caspase-2 subunit and a large caspase-2 subunit.

Specifically, the small subunit of the caspase-2 described herein comprises the amino acid sequence of SEQ ID No. 3, SEQ ID No. 91, SEQ ID No. 94, SEQ ID No. 97, SEQ ID No. 100, SEQ ID No. 103, SEQ ID No. 106, SEQ ID No. 109, SEQ ID No. 112, SEQ ID No. 115, SEQ ID No. 118, or a functionally active variant thereof comprising at least at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97 or 98 % sequence identity. Specifically, the large subunit of the caspase-2 described herein comprises the amino acid sequence of SEQ ID No. 4, SEQ ID No. 90, SEQ ID No. 93, SEQ ID No. 96, SEQ

ID No. 99, SEQ ID No. 102, SEQ ID No. 105, SEQ ID No. 108, SEQ ID No. 111, SEQ ID No. 114, SEQ ID No. 117, or a functionally active variant thereof comprising at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97 or 98 % sequence identity.

According to a specific ~~embodiment~~example, the caspase-2 provided herein comprises one or more linker sequences, specifically consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 or even more amino acid residues. Specifically, the linker can comprise more than 20 or 30 or even more amino acids, as long as the caspase retains its functional activity as described herein. Specifically, the linker sequence comprises glycine, alanine and/or serine residues. Specifically, the linker comprises at least one glycine and serine residue, more specifically the linker is GS, GGSGG and/or GSAGSAAGSG.

Specifically, the caspase-2 comprises a subunit-linker sequence, which is a linker sequence between the small subunit and the large subunit of the caspase-2 described herein.

According to a further specific ~~embodiment~~example, the caspase-2 provided herein comprises one or more C-terminal or N-terminal tags, specifically selected from the group consisting of affinity tags, solubility enhancement tags and monitoring tags described herein. Specifically, any tag known in the art can be fused to the caspase-2.

According to a specific ~~embodiment~~example, the caspase-2 provided herein comprises an N-terminal and/or C-terminal truncation of at least 1, 2, 3, 4, 5 and up to 10 or even more, as long as the caspase retains its functional activity as described herein. According to a further specific ~~embodiment~~example, the caspase-2 provided herein comprises an N-terminal and/or C-terminal extension of at least 1, 2, 3, 4, 5 and up to 10 or even more, as long as the caspase retains its functional activity as described herein. Specifically, the caspase-2 may comprise a truncation and an extension.

According to a specific ~~embodiment~~example, the caspase-2 variant provided herein comprises one or more amino acid substitutions, selected from

i. Gly^{212}, substituted with D or an amino acid selected from the group consisting of R, K, E, Q, N, A, S, T, P, H, Y
ii. Glu^{431}, substituted with V or an amino acid selected from the group consisting of C, L, I, M, F, W, R, K, D, Q, N
iii. Glu^{213}, substituted with V or an amino acid selected from the group consisting of C, L, I, M, F, W, R, K, D, Q, N
iv. Asp^{323}, substituted with E or T or an amino acid selected from the group consisting of R, K, Q, N, G, A, S, P, H, Y
v. Val^{266}, substituted with G or an amino acid selected from the group consisting of A, S, T, P, H, Y, C, L, I, M, F, W
vi. Lys^{409}, substituted with E or an amino acid selected from the group consisting of R, D, Q, N,
vii. His^{226}, substituted with A or an amino acid selected from the group consisting of G, S, T, P, Y,
viii. Val^{296}, substituted with M or an amino acid selected from the group consisting of C, L, I, F, W, and/or
ix. Asp^{326}, substituted with E or Y or an amino acid selected from the group consisting of R, K, Q, N, G, A, S, T, P, H,

with reference to the positions of SEQ ID No. 11, or positions functionally equivalent to positions of SEQ ID No. 11.

Specifically, the caspase-2 variant provided herein comprises amino acid substitutions at positions of SEQ ID No. 11, or at positions functionally equivalent to positions of SEQ ID No. 11, selected from

i. His^{226} and Asp^{323}, specifically comprising H226A and D323T substitutions;
ii. Glu^{431}, specifically comprising a E431V substitution;
iii. Glu^{431} and Asp^{326}, specifically comprising E431V and D326E substitutions;
iv. Glu^{431}, Gly^{212}, Val^{266} and Asp^{323}, specifically comprising E431V, G212D, V266G and D323E substitutions;
v. Glu^{431}, Gly^{212}, Val^{266}, $Asp^{323}$and Asp^{326}, specifically comprising E431V, G212D, V266G, D323E and D326E substitutions;
vi. Lys^{409}, Glu^{431}, Glu^{213}, Val^{296} and Asp^{326}, specifically comprising K409E, E431V, E213V, V296M and D326Y substitutions;
vii. Glu^{431} and Gly^{212}, specifically comprising E431V and G212D substitutions;
viii. Glu^{431} and Glu^{213}, specifically comprising E431V and E213V substitutions; and
ix. Gly^{212} and Glu^{213}, specifically comprising G212D and E213V substitutions.

Further provided herein is a method of producing a caspase variant, specifically a caspase-2, even more specifically the circular permuted caspase-2, comprising increased P1' tolerance.

Specifically, a wild-type cp caspase-2 or a cp caspase-2 variant as described herein, or a functionally active variant thereof, is produced by a method comprising the steps of

i. cloning a nucleotide sequence encoding a caspase-2, specifically a circular permuted caspase-2 into a vector, specifically said sequence is under the control of a promoter,
ii. transforming a host cell with said vector,
iii. culturing the transformed host cell under conditions wherein the caspase is expressed,
iv. isolating the caspase from the host cell culture, optionally by disintegrating the host cells, and
v. optionally purifying the caspase.

Thus, in another aspect, the present disclosure provides a method of producing a circular permuted caspase-2 (cp caspase-2) comprising the steps of

i. cloning a nucleotide sequence encoding a cp caspase-2 under the control of a promoter into an expression vector,
ii. transforming a host cell with said vector,
iii. culturing the transformed host cell under conditions wherein the cp caspase-2 is expressed,
iv. optionally isolating the cp caspase-2 from the host cell culture, optionally by disintegrating the host cells, and
v. optionally purifying the cp caspase-2.

Specifically, the nucleic acid sequence encoding the caspase-2 described herein is operably linked to a promoter. Specifically, the promoter is an inducible or a constitutive promoter. Specifically, the promoter is selected from the group consisting of

T7, lac, tac, trc, lacUV5, trp, phoA, pL, XylS/Pm regulator/promoter system, Pm-promoter, Pm-promoter variants, araBAD, T3, T5, T4, T7A1, T7A2, T7A3, hybrid promoters and the strong constitutive HCD promoter. Specifically, the promoter is associated with one or more lac operators or respective other operators or regulators or further regulation elements or the promoter is not associated with such regulatory elements. In a preferred ~~embodiment~~example, the promoter/regulator is a promoter / regulator selected from the group consisting of T7 promoter/operator, XylS/Pm regulator/promoter, functionally active variants of the Pm promoter, araBAD promoter/operator, T5, T7A1, T7A2, T7A3 promoter/operator, phoA promoter/regulator, and the trp promoter/operator system. Specifically, the promoter/regulator is of the T7 promoter/operator system.

Specifically, the host cell is a eukaryotic or a prokaryotic microbial host cell. Specifically, host cells are selected from the group consisting of bacterial cells, yeast cells, insect cells, mammalian cells and plant cells, preferably the host cells are bacterial or yeast cells selected from the group consisting of *E. coli*, *Pseudomonas sp.*, *Bacillus sp., Streptomyces sp., Saccharomyces sp., Schizosaccharomyces sp.*, *Pichia sp., Kluyveromyces sp.* and *Hansenula sp..*

Even more specifically, the host cell is of an *E. coli* B or K strain, such as but not limited to BL21 or HMS174. Specifically, the host cell has integrated in its genome a nucleotide sequence encoding the T7 RNA polymerase and is capable of constitutive of inducible expression of the caspase-2 described herein. as According to a specific example, the host cell is an *E.coli* BL21 (DE3), or HMS 174 (DE3) cell or a cell derived from BL21 (DE3), or HMS 174 (DE3) comprising a deletion of at least one essential lambda phage protein.

The caspase-2 or cp caspase-2 described herein may comprise a tag sequence, within its sequence or fused to its N- or C-terminus.

The circular permuted caspase-2 can optionally have an affinity tag, preferably fused to its N- or C-terminus, preferably the affinity tag is a 6His Tag. In a preferred ~~embodiment~~example the 6His tag is N-terminal. Specifically, in order to increase the expression of soluble cp caspase-2 the cp caspase-2 is fused with a solubility enhancement tag at its C- or N-terminus. In a preferred ~~embodiment~~example, the solubility enhancement tag is N-terminal to the cp caspase-2. Preferably the solubility tag is based on highly charged peptides of bacteriophage genes. Exemplary solubility tags and their sequences are

listed in Table 1 of US 8,535, 908 B2. Specifically, the solubility tag is selected from the group consisting of the tags, T7C, T7B, T7B1, T7B2, T7B3, T7B3, T7B4, T7B5, T7B6, T7B6, T7B7, T7B8, T7B9, T7B10, T7B11, T7B12, T7B13, T7A, T7A1, T7A2, T7A3, T7A4, T7A5, T7AC, T3, N1, N2, N3, N4, N5, N6, N7, calmodulin-binding peptide (CBP), DsbA, DsbC, poly Arg, poly Lys, G B1 domain, protein D, Z domain of Staphylococcal protein A, and thioredoxin tag, preferably it comprises a T7AC or a T7A3 tag. According to a specific example, the tag is a modified T7A3 tag, herein referred to as T7AC (SEQ ID No. 43). Preferably, one or more T7A3 (SEQ ID No. 37) and/or T7AC (SEQ ID No. 43) tags or functional variants thereof having 1-5 amino acid substitutions, additions, dilutions or the like, are used.

Specifically, the caspase produced according to the method described herein has one or more affinity tags and one or more solubility enhancement tags fused to its N-terminus with or without linker sequences between the tags or between a tag and the N-terminus of the cp caspase-2. Specifically, said caspase has a T7AC or a T7A3 tag and a 6His tag fused to its N-terminus, whereas from N- to C-terminus the 6 His Tag is the first and the T7AC or T7A3 tag is the second tag, or the T7AC or T7A3 tag is the first and the 6His tag is the second tag.

It has surprisingly been found that production of a cp caspase-2 described herein can be significantly improved using a solubility enhancement tag, such as T7A3 or T7AC, fused to the N-terminus of the enzyme. Use of such tag significantly increases the titer of the enzyme by about 2.5-fold or more.

According to a specific ~~embodiment~~example, the cp caspase-2 produced according to the method described herein, thus comprises the following elements fused to its N-terminus, in the order from N- to C-terminus:

a. affinity tag, preferably 6-His tag;
b. optionally a linker;
c. solubility enhancement tag, preferably T7AC or T7A3; and
d. cp caspase-2, wild-type or variant as described herein.

According to a further specific ~~embodiment~~example, the cp caspase-2 produced according to the method described herein, comprises the following elements fused to its N-terminus, in the order from N- to C-terminus:

a. solubility enhancement tag, preferably T7AC or T7A3;
b. optionally a linker;

c. affinity tag, preferably 6-His tag; and

d. cp caspase-2, wild-type or variant as described herein.

Specifically, the expression cassette for expression of the cp caspase-2 comprises the nucleotide sequence encoding the cp caspase-2 under control of a promoter. Optionally, the expression cassette further comprises the nucleotide sequences encoding the affinity tag, in a preferred ~~embodiment~~example the 6 His tag and/or a nucleotide sequence encoding the T7AC or T7A3 tag and nucleotide sequences encoding linker sequences between the tags and/or between a tag and the cp-caspase-2. In another ~~embodiment~~example, the expression cassette is flanked by two sequences homologous to a sequence in the genome of the host cell, preferably a microbial cell, more preferably a bacterial cell, more preferably *E. coli*, for integration of the expression cassette by homologous recombination into the genome of the host cell.

Specifically, the cell is transformed by a vector comprising the expression cassette. Specifically, the cp caspase-2, with or without tags as described herein, is expressed from one or more plasmids or from one or two copies of a nucleic acid sequence integrated in the genome of the host cell.

Specifically, the cp caspase-2 with or without tags as described herein, can be produced by cultivation of the host cell and induction of expression by addition of an inducer, such as e.g. IPTG when using the T7 promoter/operator system, in a bioreactor (fermenter).

Specifically, culturing of step (iii) of the method to procude a cp caspase-2 as described herein comprises a fed-batch phase for expression of the cp caspase-2 comprising a specific growth rate and induction of expression, preferably using IPTG.

Specifically, culturing of step (iii) of the method to procude a cp caspase-2 as described herein comprises a fed-batch phase for expression of the cp caspase-2, said fed batch phase specifically cromprising a specific growth rate, µ of about 0,01-0,1 h^{-1} , and induction of expression of the cp caspase-2 by addition of IPTG at a concentration of about 0,01 – 1,5 µmol/g of actual CDM (cell dry mass). Specifically, concentration of IPTG of µmol/ g of actual CDM means the concentration of IPTG in the fermenter at a certain time point during the feed phase related to the CDM in g at that certain time point.

According to a specific ~~embodiment~~example, growth rate µ is about 0,01-0,07 h^{-1}, preferably it is about 0,01-0,03 h^{-1} or 0,01-0,05 h^{-1} or 0,02-0,05 h^{-1} or 0,03-0,05 h^{-1} or 0,03-0,07 h^{-1} or 0,05-0,07 h^{-1}~~,~~, preferably it is any of about 0,03, 0,05 or 0,07 h^{-1}.

According to a further specific ~~embodiment~~example, the IPTG concentration is about 0,1-1,5 µmol/g or 0,1-1,3 µmol/g or 0,2-1,3 µmol/g or 0,3-1,3 µmol/g or 0,5-1,3 µmol/g of actual CDM, preferably it is about 0,5-0,9 µmol/g actual CDM or about 0,9-1,3 µmol/g actual CDM, preferably it is about 0,5, 0,9 or about 1,3 µmol/g CDM.

According to yet a further specific ~~embodiment~~example, culturing of step (ii) further comprises a first fed-batch phase for the production of biomass, prior to the fed-batch phase for the expression of the cp caspase-2, said first fed-batch phase comprising a growth rate, µ of about 0,05-0,5 h^{-1} or 0,05-0,4 h^{-1} or 0,07-0,3 h^{-1}. Specifically, the growth rate µ is about 0,1-0,3 h^{-1} or 0,1-0,2 h^{-1}, or 0.10, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19 or 0.20 h^{-1}. preferably about 0,13-0,21 h^{-1}, even more preferably about 0,16-0,18 h^{-1} and most preferably it is about 0,17 h^{-1}. Specifically, said first fed-batch phase is followed by a second fed-batch phase for expression of the recombinant protein, preferably started by addition of an inducer of expression, such as IPTG, and typically comprising a lower growth rate.

Specifically provided herein is a fusion protein comprising the following structure from N- to C-terminus:

i. a tag sequence comprising a caspase recognition site specifically recognized by the cp caspase-2 or caspase2 described herein,
ii. a cleavage site P1/P1', and
iii. a protein or polypeptide of interest (POI).

Specifically, P1' is the N-terminal amino acid of the protein of interest (POI).

Specifically, the tag sequence of the fusion protein described herein further comprises one or more tags selected from the group consisting of affinity tags, solubility enhancement tags and monitoring tags. Specifically, any tag with any function known in the art can be fused to the POI. Specifically, the fusion protein further comprises one or more linker sequences. Specifically, the fusion protein comprises a caspase recognition site comprising 5 amino acids of the sequence P5 P4 P3 P2 P1, and a cleavage site P1/P1', wherein P1' is the N-terminal amino acid of the POI.

According to a specific ~~embodiment~~example, the fusion protein provided herein comprises the cp caspase-2 or caspase 2 described herein within its sequence. Specifically, the fusion protein comprises the cp caspase-2 or caspase 2 described herein fused to the N- or the C-terminus of the fusion protein.

Specifically, such fusion protein is used to produce a POI comprising an authentic N-terminus by cleavage of the fusion protein at the N-terminus of the POI using the cp caspase-2 or caspase-2 described herein.

In one ~~embodiment~~example, the POI comprises an N-terminal tag which comprises at least a caspase recognition site, wherein the C-terminal amino acid of the recognition site, P1, is the last (C-terminal) amino acid residue of the tag and the N-terminal amino acid of the POI is the P1' residue of the caspase cleavage site. In this ~~embodiment~~example, the tag sequence including the recognition site is released from the POI through proteolytic cleavage by the caspase, generating a POI comprising an authentic N-terminus.

Further provided herein are methods of producing a protein or polypeptide of interest (POI) using the cp caspase-2 or caspase-2 described herein. Specifically, the cp caspase-2 described herein is used for the production of a POI comprising an authentic N-terminus. Thus, in another aspect, the present disclosure provides use of the cp caspase-2 as described herein for the production of a protein of interest (POI).

Specifically provided herein are methods of producing a POI comprising an authentic N-terminus, using the fusion protein described herein and the caspase-2 or the cp caspase-2 described herein.

Specifically provided herein are methods of producing a POI comprising an authentic N-terminus, using the fusion protein described herein, wherein the fusion protein comprises the caspase-2 or the cp caspase-2 described herein at its N- or C-terminus.

Specifically, the fusion protein used in the methods described herein comprises the following structure from N- to C-terminus:

i. one or more N-terminal tags,
ii. optionally one or more tag-linker sequences and
iii. a caspase recognition site comprising 5 amino acids of the sequence P5 P4 P3 P2 P1,
iv. a cleavage site P1/P1', and
v. a POI,

wherein said recognition site is specifically recognized by the caspase-2 or the cp caspase-2 described herein. Specifically, P1' is the N-terminal amino acid of the POI.

Thus, in another aspect, the present disclosure provides a method of producing a protein of interest (POI), comprising the steps of:

i. providing a fusion protein comprising from N- to C-terminus one or more tags, optionally one or more tag-linker sequences and a caspase recognition site N-terminally fused to the POI, wherein said caspase recognition site is specifically recognized by the cp caspase-2 as described herein,

ii. contacting said fusion protein with said cp caspase-2 for a period of time sufficient for said cp caspase-2 to cleave the fusion protein, and

iii. optionally purifying the POI.

Specifically provided herein is a method of producing a POI *in vivo*.

Specifically, the *in vivo* method of producing a POI comprising an authentic N-terminus comprises the steps of:

i. expressing the fusion protein comprising from N- to C-terminus optionally one or more tags, optionally one or more tag-linker sequences and a caspase recognition site N-terminally fused to the POI; and the caspase-2 or cp caspase-2 described herein specifically recognizing the recognition site of the fusion protein, in the same host cell,

ii. optionally, wherein said fusion protein and caspase-2 or cp caspase-2 are under the same promoter,

iii. culturing the host cell, wherein said caspase-2 or cp caspase-2 cleaves the fusion protein in culture, and

iv. isolating the POI from the cell and optionally purifying the POI.

Specifically, the host cell is selected from the group consisting of bacterial cells, yeast cells, insect cells, mammalian cells and plant cells, preferably the host cells are bacterial or yeast cells selected from the group consisting of *E. coli*, *Pseudomonas sp.*, *Bacillus sp., Streptomyces sp., Saccharomyces sp., Schizosaccharomyces sp.*, *Pichia sp., Kluyveromcyes sp.* and *Hansenula sp..*

According to a specific ~~embodiment~~example, the fusion protein and the caspase described herein are under transcriptional control of different promoters and the expression of the caspase is induced after expression of the fusion protein. According to a different ~~embodiment~~example, the fusion protein and the caspase are under transcriptional control of the same promoter, specifically they are expressed at the same time.

Specifically, the caspase comprises an N- or C-terminal tag, which may be used to separate the caspase and the POI.

According to a further specific ~~embodiment~~example, the caspase described herein is part of the fusion protein expressed in the host cell and cleaves the fusion protein releasing a POI comprising an authentic N-terminus in the host cell.

Specifically, the fusion protein, the POI and/or the caspase are isolated using a column, specifically a chromatography column, more specifically an immobilized metal affinity chromatography column (IMAC).

Specifically provided herein is a method of producing a POI *in vitro*.

Specifically, the *in vitro* method of producing a protein of interest (POI) comprising an authentic N-terminus comprises the steps of:

i. providing a fusion protein comprising from N- to C-terminus one or more tags, optionally one or more tag-linker sequences and a caspase recognition site N-

Specifically, the captured fusion protein is released and the N-terminal tag is removed in solution or in immobilized enzyme reactor, wherein the caspase is immobilized in a column or on a carrier.

According to a further specific ~~embodiment~~example, the fusion protein and the caspase are bound on a column.

Specifically, the method of producing a POI comprising an authentic N-terminus using a column comprises the following steps:

i. expressing a fusion protein comprising one or more N-terminal affinity tags, optionally one or more tag-linker sequences, a caspase recognition site and a cleavage site P1/P1', wherein P1' is the N- terminal amino acid of the POI, and a POI, in a host cell,
ii. isolating the fusion protein from the host cell and capturing / binding the fusion protein on a solid support using the affinity tag,
iii. providing a caspase-2 or cp caspase-2 described herein (caspase) specifically recognizing the recognition site of the fusion protein,
iv. contacting said caspase with the bound fusion protein for a period of time sufficient for said caspase to cleave the fusion protein releasing the POI from the solid support whereas tag and optionally the uncleaved fusion protein remain bound, and
v. isolating and optionally further purifying the POI.

Specifically, the caspase comprises a tag sequence, specifically an affinity tag, to allow separation of the caspase from the POI after cleavage.

In a specific ~~embodiment~~example, the caspase-2 or cp caspase-2 described herein comprises an affinity tag and is immobilized on a solid support or column and the fusion protein is brought into contact with the immobilized caspase. Specifically, the fusion protein is brought into contact with the caspase immobilized in a column by flowing the fusion protein through the column. Specifically, the cleaved tag is separated from the POI using the affinity tag in the tag sequence.

In a further specific ~~embodiment~~example, the caspase and the fusion protein comprise an identical N-terminal affinity tag, allowing immobilization of the fusion protein and the caspase on the solid support. Upon cleavage of the POI by the caspase, the POI is released from the column and the tag sequence is retained in the column as well as the caspase and optionall uncleaved fusion protein.

Specifically, the solid support is a column, specifically a chromatography column, more specifically an immobilized metal affinity chromatography column (IMAC) or an activated NHS column allowing immobilization of a polypeptide through amine coupling.

Specifically, a flow-through reactor is used comprising immobilized caspase-2, cp caspase-2 or fusion protein described herein. Specifically, the flow-through reactor is a plug flow reactor.

Further provided herein is an isolated nucleotide sequence encoding the caspase-2 or cp caspase-2 described herein. In another aspect, the present disclosure provides an isolated nucleotide sequence, preferably comprised in a vector or in an expression cassette, encoding the cp caspase-2 as described herein.

Further provided herein is a vector comprising the isolated nucleotide sequence described herein, specifically said vector is a bacterial expression vector. More specifically the vector is a plasmid. In another ~~embodiment~~example, the vector is a linear vector flanked with homology regions for homologous integration of the nucleotide sequence encoding the caspase-2 or cp caspase-2 described herein into the chromosome of the host cell.

Further provided herein is an expression cassette comprising the nucleotide sequence operably linked to regulatory elements such as promoter, operator, terminator and the like. Specifically, said regulatory elements are one or more promoters or expression enhancing elements.

~~Further provided herein is~~In another aspect, the present disclosure provides a host cell or a host cell line expressing the caspase-2 or cp caspase-2 described herein, wherein the host cells are selected from the group consisting of bacterial cells, yeast cells, insect cells, mammalian cells and plant cells, preferably the host cells are bacterial or yeast cells selected from the group consisting of *E. coli*, *Bacillus sp., Streptomyces sp., Saccharomyces sp., Schizosaccharomyces sp., Kluyveromyces sp.* and *Pichia sp..*

According to a specific ~~embodiment~~example, the expression cassette and the host cell or host cell line described herein are comprised in an expression system. Further provided herein is an expression system comprising the expression cassette and the host cell or host cell line described herein.

Further described herein is the use of the caspase-2 or cp caspase-2 described herein for the *in vivo* cleavage of a substrate in a non-human organism. Specifically, the non-human organism is a prokaryotic organism, specifically it is *E. coli*. Thus, in another aspect, the present disclosure provides use of the cp caspase-2 as described herein for the *in vitro* cleavage of a substrate in a non-human organism, preferably wherein the non-human organism is a prokaryotic organism, especially preferred *E. coli.*

Further provided herein is a kit, comprising

i. the caspase-2 or cp caspase-2 described herein, and

ii. an expression vector, optionally comprising an affinity tag, preferably a 6His tag, a linker sequence, and/or a nucleotide sequence coding for a recognition site, preferably VDVAD.

The kit may optionally further comprise chromatography material for affinity chromatography, preferably an IMAC (immobilized metal affinity chromatography) material, preferably Ni-NTA (Ni- Nitrilotriacetic acid) chromatography material, preferably pre-packed in a chromatography column.

Further optionally comprised in the kit is a plasmid comprising the nucleotide sequence from 5´ to 3´ encoding an affinity tag, preferably a 6His tag, optionally a linker sequence and a nucleotide sequence coding for the recognition site, preferably VDVAD. Via a multiple cloning site, a DNA sequence encoding a POI can be inserted into the plasmid directly fused to the nucleotide sequence encoding the recognition site.

Further described herein is a pharmaceutical composition comprising the caspase-2 or cp caspase-2 provided herein and optionally one or more excipients.

Further described herein is use of the caspase-2 or cp caspase-2 provided herein for preparing a pharmaceutical composition.

Specifically, the caspase-2 or cp caspase-2 described herein is provided for use in the treatment of a disease. Specifically, the caspase-2 or cp caspase-2 described herein is provided for use in the treatment of cancer, osteoporosis, Alzheimer's disease, Parkinson's disease, inflammatory disease, or auto-immune diseases, specifically via proteolytically attacking respective disease relevant proteins. Thus, in another aspect, the present disclosure provides the cp caspase-2 as described herein, for use in the treatment of a disease, preferably for use in the treatment of cancer, Alzheimer's disease, Parkinson's disease or inflammatory disease.

Specifically, the caspase-2 or cp caspase-2 described herein is provided for the manufacture of a medicament for the treatment of cancer, Alzheimer's disease, Parkinson's disease or inflammatory disease.Further provided herein is a protein tag for enhanced expression of a POI, comprising a solubility enhancement tag and the amino acid sequence VDVAD (SEQ ID NO:45). Specifically, the sequence VDVAD is at the C-terminus of the protein tag described herein, specifically directly linked to the N-terminus of the POI.

It has been surprisingly found that by including the amino acid sequence VDVAD in a protein tag, the expression of a protein of interest fused to the protein tag can be significantly increased. Importantly, this increase in expression persists, despite addition of a histidine tag sequence to the protein tag. Using a histidine affinity tag, such as 6-His, typically decreases the expression rate of a protein of interest significantly. The

inventors have surprisingly found that by including the sequence VDVAD in the protein tag this effect is reversed, and increased expression titers can be provided.

According to a specific ~~embodiment~~example, the solubility enhancement tag is selected from the group consisting of T7C, T7B, T7B1, T7B2, T7B3, T7B3, T7B4, T7B5, T7B6, T7B6, T7B7, T7B8, T7B9, T7B10, T7B11, T7B12, T7B13, T7A, T7A1, T7A2, T7A3, T7A4, T7A5, T3, N1, N2, N3, N4, N5, N6, N7, T7AC, calmodulin-binding peptide (CBP), DsbA, DsbC, poly Arg, poly Lys, G B1 domain, protein D, Z domain of Staphylococcal protein A, and thioredoxin tag. Specifically, the solubility enhancement tag is T7AC or T7A3.

According to a further specific ~~embodiment~~example, the protein tag described herein further comprises a histidine tag sequence, preferably comprising 1-20 histidine residues, even more preferably it is a 1-His, 2-His, 3-His, 4-His, 5-His, 6-His, 7-His, 8-His, 9-His, 10-His, 11-His, 12-His, 13-His, 14-His, 15-His, 16-His, 17-His, 18-His, 19-His or 20-Histag sequence.

Specifically, the solubility enhancement tag is located at the N-terminus of the protein tag described herein.

Specifically, the histidine tag sequence is located at the N-terminus of the protein tag described herein.

According to a specific ~~embodiment~~example, the protein tag described herein further comprises one or more linker sequences comprising one or more amino acid residues. Specifically, said linker sequence comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more amino acid residues. Specifically, the linker can comprise more than 20 or 30 or even more amino acids, as long as the caspase retains its functional activity as described herein. Specifically, the one or more amino acid residues of the linker sequence are any of the naturally occurring amino acids or derivatives thereof, preferably selected from the group consisting of G, S, T, N, A. Specifically, the linker sequence comprises glycine, alanine and/or serine residues. Specifically, the linker comprises at least one glycine and serine residue, more specifically the linker is GS, GSG, GGSGG, GSGSGSG and/or GSAGSAAGSG.

Specifically, said one or more linker sequences are located between the VDVAD sequence and the solubility enhancement tag or the histidine tag sequence.

According to a specific ~~embodiment~~example, the protein tag described herein further comprises a signal peptide at its N-terminus. Signal peptides are known to the person

l. 6-His– T7A3 – GSG - VDVAD.

Specifically, the protein tag described herein comprises one of the following structures from N- to C-terminus:

a. ompA signal peptide – T7AC – 6-His – VDVAD;
b. ompA signal peptide – T7A3 – 6-His – VDVAD;
c. ompA signal peptide – T7AC – 6-His – GSG – VDVAD;
d. ompA signal peptide – T7A3 – 6-His – GSG – VDVAD;
e. ompA signal peptide – T7AC – 6-His – GSGSGSG – VDVAD;
f. ompA signal peptide – T7A3 – 6-His – GSGSGSG – VDVAD;
g. ompA signal peptide – 6-His – T7AC – VDVAD;
h. ompA signal peptide – 6-His – T7A3 – VDVAD;
i. ompA signal peptide – 6-His– T7AC – GSG - VDVAD;
j. ompA signal peptide – 6-His– T7A3 – GSG - VDVAD;
k. ompA signal peptide – 6-His– T7AC- GSGSGSG – VDVAD;
l. ompA signal peptide – 6-His– T7A3 – GSGSGSG - VDVAD.

Further provided herein is a fusion protein comprising the protein tag described herein and a POI. Specifically, the N-terminus of the POI is fused to the C-terminus of said protein tag. Even more specifically, the N-terminus of the POI is directly fused to the C-terminus of the protein tag, which C-terminus is the sequence VDVAD, i.e., the N-terminal amino acid of the POI is directly linked to the C-terminal D of the VDVAD sequence of the protein tag.

With regard to the POI there is no limitation. The POI may be any polypeptide, including e.g. the caspases described herein.

Further also provided herein is a method of producing a POI, comprising the steps of:

i. providing the fusion protein described herein comprising the protein tag described herein comprising a POI,
ii. contacting said fusion protein with a circular permuted caspase-2 (cp caspase-2) for a period of time sufficient for said cp caspase-2 to cleave the fusion protein thereby releasing the POI, and
iii. optionally purifying the POI.

According to a specific ~~embodiment~~example, the method of producing a POI as described herein comprises the following steps:

CLAIMS

1. A single-chain circular permuted caspase-2 (cp caspase-2) comprising the following structure from N- to C-terminus:

   i. a small subunit of a caspase-2 comprising a sequence selected from the group consisting of SEQ ID No. 3, SEQ ID No. 91, SEQ ID No. 94, SEQ ID No. 97, SEQ ID No. 100, SEQ ID No. 103, SEQ ID No. 106, SEQ ID No. 109, SEQ ID No. 112, SEQ ID No. 115, SEQ ID No. 118, or a functionally active variant thereof comprising a sequence with at least 70% sequence identity to SEQ ID No. 3, SEQ ID No. 91, SEQ ID No. 94, SEQ ID No. 97, SEQ ID No. 100, SEQ ID No. 103, SEQ ID No. 106, SEQ ID No. 109, SEQ ID No. 112, SEQ ID No. 115, or SEQ ID No. 118; ~~and~~covalently linked, either via a linker or directly, to

   ii. a large subunit of a caspase-2 comprising a sequence selected from the group consisting of SEQ ID No. 4, SEQ ID No. 90, SEQ ID No. 93, SEQ ID No. 96, SEQ ID No. 99, SEQ ID No. 102, SEQ ID No. 105, SEQ ID No. 108, SEQ ID No. 111, SEQ ID No. 114, SEQ ID No. 117, or a functionally active variant thereof comprising a sequence with at least 70% sequence identity to SEQ ID No. 4, SEQ ID No. 90, SEQ ID No. 93, SEQ ID No. 96, SEQ ID No. 99, SEQ ID No. 102, SEQ ID No. 105, SEQ ID No. 108, SEQ ID No. 111, SEQ ID No. 114, or SEQ ID No. 117,

wherein said cp caspase-2 is a protease and with the proviso that in case of a selection of SEQ ID No. 3 for the small subunit and SEQ ID No. 4 for the large subunit, then the cp caspase 2 comprises one or more amino acid substitutions at a position functionally equivalent to any of positions 171, 105, 172, 282, 225, 83, 185, 255, or 285 of SEQ ID No. 6 ~~increasing P1' tolerance of said cp caspase-2 compared to a cp caspase-2 without said amino acid substitutions~~.

2. The cp caspase-2 of claim 1 comprising one or more amino acid substitutions at positions 171, 105, 172, 282, 225, 83, 185, 255, or 285 of SEQ ID No. 6 or at a position functionally equivalent to any of positions 171, 105, 172, 282, 225, 83, 185, 255, or 285 of SEQ ID No. 6 or any combination thereof.

3. The cp caspase-2 of claim 1 or 2, comprising a propeptide of a small caspase-2 subunit (SS propeptide), fused to the N-terminus of the small subunit, preferably wherein the SS propeptide comprises an amino acid substitution at position Asp^{14} of SEQ ID No. 2 or at a position functionally equivalent to Asp^{347} of SEQ ID No. 11, especially preferred wherein Asp is substituted to Ala.

~~4. The cp caspase-2 of claim 3, wherein the SS propeptide comprises one or more amino acid substitutions at the C-terminus of the SS propeptide.~~

~~5. The cp caspase-2 of claim 3 or 4, wherein the SS propeptide comprises an amino acid substitution at position Asp^{14} of SEQ ID No. 2 or at a position functionally equivalent to Asp^{347} of SEQ ID No. 11, specifically Asp is substituted to Ala.~~

~~6.~~4. The cp caspase-2 of any one of claims 1 to ~~5~~3, ~~further comprising one or more linker sequences, specifically consisting~~wherein the linker consists of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 amino acid residues, preferably wherein the linker sequence comprises glycine and/or serine residues, more preferred the linker is GS, GGSGG, GSAGSAAGSG, $(GS)_n$, GSG or G_4S.

~~7. The cp caspase-2 of claim 6, wherein the linker sequence comprises glycine and/or serine residues, more specifically the linker is GS, GGSGG, GSAGSAAGSG, $(GS)_n$, GSG or G_4S.~~

~~8. The cp caspase-2 of claims 6 or 7, wherein the linker sequence is a subunit-linker sequence between the small subunit and the large subunit.~~

~~9.~~5. The cp caspase-2 of any one of claims 1 to ~~8~~4, comprising one or more C-terminal or N-terminal tags~~, specifically~~ selected from the group consisting of affinity tags, solubility enhancement tags and monitoring tags, preferably wherein the affinity tag is selected from the group consisting of poly-histidine tag, poly-arginine tag, peptide substrate for antibodies, chitin binding domain, RNAse S peptide, protein A, ß-galactosidase, FLAG tag, Strep II tag, streptavidin-binding peptide (SBP) tag, calmodulin-binding peptide (CBP), glutathione S-transferase (GST), maltose-binding protein (MBP), S-tag, HA tag, c-Myc tag, SUMO tag, E.coli thioredoxin, NusA, chitin binding domain CBD, chloramphenicol acetyl transferase CAT, LysRS, ubiquitin, calmodulin, and lambda gpV, preferably the tag is a His tag comprising one or more His, more preferably it is a hexahistidine tag, preferably wherein the solubility enhancement tag is selected from the group consisting of T7C, T7B, T7B1, T7B2, T7B3, T7B3, T7B4, T7B5, T7B6, T7B6, T7B7, T7B8, T7B9, T7B10, T7B11, T7B12, T7B13, T7A, T7A1,

T7A2, T7A3, T7A4, T7A5, T3, N1, N2, N3, N4, N5, N6, N7, T7AC, calmodulin-binding peptide (CBP), DsbA, DsbC, poly Arg, poly Lys, G B1 domain, protein D, Z domain of Staphylococcal protein A, and thioredoxin, preferably wherein the monitoring tag is selected from the group consisting of m-Cherry, GFP and f-Actin.

~~10. The cp caspase-2 of claim 9, wherein the affinity tag is selected from the group consisting of poly-histidine tag, poly-arginine tag, peptide substrate for antibodies, chitin binding domain, RNAse S peptide, protein A, ß-galactosidase, FLAG tag, Strep II tag, streptavidin-binding peptide (SBP) tag, calmodulin-binding peptide (CBP), glutathione S-transferase (GST), maltose-binding protein (MBP), S-tag, HA tag, c-Myc tag, SUMO tag, E.coli thioredoxin, NusA, chitin binding domain CBD, chloramphenicol acetyl transferase CAT, LysRS, ubiquitin, calmodulin, and lambda gpV, specifically the tag is a His tag comprising one or more His, more specifically it is a hexahistidine tag.~~

~~11. The cp caspase-2 of claim 9, wherein the solubility enhancement tag is selected from the group consisting of T7C, T7B, T7B1, T7B2, T7B3, T7B3, T7B4, T7B5, T7B6, T7B6, T7B7, T7B8, T7B9, T7B10, T7B11, T7B12, T7B13, T7A, T7A1, T7A2, T7A3, T7A4, T7A5, T3, N1, N2, N3, N4, N5, N6, N7, T7AC, calmodulin-binding peptide (CBP), DsbA, DsbC, poly Arg, poly Lys, G B1 domain, protein D, Z domain of Staphylococcal protein A, and thioredoxin.~~

~~12. The cp caspase-2 of claim 9, wherein the monitoring tag is selected from the group consisting of m-Cherry, GFP and f-Actin.~~

~~13.~~6. The cp caspase-2 of ~~any one of claims 9 to 12~~claim 5, comprising more than one tags, ~~specifically~~ preferably comprising an affinity tag and a solubility enhancement tag, optionally comprising a tag linker sequence linking two tags or linking a tag and the small subunit, the large subunit or the SS propeptide of the cp caspase-2.

~~14.~~7. The cp caspase-2 of claim ~~13~~6, wherein the affinity tag is a hexahistidine tag and the solubility enhancement tag is a T7AC or a T7A3 tag.

~~15. The cp caspase-2 of any one of claims 6 to 14, wherein the linker sequence is a tag-linker sequence, linking two tags or linking a tag and the small subunit, the large subunit or the SS propeptide of the cp caspase-2.~~

~~16. The cp caspase-2 of any one of claims 1 to 15, comprising one or more N-terminal tags and optionally one or more tag-linker sequences between the tags or between a tag and the N-terminus of the small subunit or the SS propeptide.~~

~~17. The cp caspase-2 of any one of claims 1 to 16, comprising one or more C-terminal tags and optionally one or more tag-linker sequences, which are linker sequences between the tags or between a tag and the C-terminus of the large subunit.~~

~~18.~~8. ~~A functionally active variant of the~~The cp caspase-2 of any one of claims ~~1-17~~1 to 7, wherein

i. the small subunit of a caspase-2 comprises

a) a first conserved region of the active center with at least 37.5 % amino acid sequence identity to SEQ ID No. 177 (1st consensus: AAMRNTKR) or 100% sequence identity to XXXRNTXX (SEQ ID No. 200), wherein X is any amino acid,

b) a second conserved region of the active center with at least 61.5 % amino acid sequence identity to SEQ ID No. 178 (2nd consensus: EGYAPGTEFHRCK) or 100% sequence identity to EGXXPGXXXHRCK (SEQ ID No. 194), wherein X is any amino acid, and

ii. the large subunit of a caspase-2 comprises

a) a third conserved region of the active center with at least 25.0 % amino acid sequence identity to SEQ ID No. 174 (3rd consensus: G-EKDLEFRSGGDVDH) or 100% sequence identity to X-XXXLXXRXGXXXDX (SEQ ID No. 195), wherein X is any amino acid,

b) a fourth conserved region of the active center with at least 53.3 % amino acid sequence identity to SEQ ID No. 175 (4th consensus: LLSHGVEGGXYGVDG) or 100% sequence identity to XXSHGXXGXXYGXDG (SEQ ID No. 196), wherein X is any amino acid, and

c) a fifth conserved region of the active center with at least 50.0 % amino acid sequence identity to SEQ ID No. 176 (5th consensus: QACRGDET) or 100% sequence identity to QACXGXXX (SEQ ID No. 197), wherein X is any amino acid.

~~19. A functionally active variant of the cp caspase-2, comprising at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97 or 98 % sequence identity to the cp caspase-2 of any one of claims 1 to 18.~~

~~20.~~9. The ~~functionally active variant of claim 19~~ cp caspase-2 of any one of claims 1 to 8, comprising an amino acid sequence comprising at least 70~~, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97 or 98~~ % sequence identity to SEQ ID No. 9, 6, 14, 15, 16, 80, 88, 25, 26, 27, 28, 29, 30, 35, 39, 41, 64, 66, 68, 73, 74, 75, 76, 77, 81, 82, 83, 84, or 85.

~~21. The cp caspase-2 of any one of claims 1 to 20, wherein the~~

~~i. small subunit is selected from the group consisting of SEQ ID No. 3, SEQ ID No. 91, SEQ ID No. 94, SEQ ID No. 97, SEQ ID No. 100, SEQ ID No. 103, SEQ ID No. 106, SEQ ID No. 109, SEQ ID No. 112, SEQ ID No. 115, SEQ ID No. 118 or functionally active variants thereof having at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97 or 98 % sequence identity, and/or~~

~~ii. the large subunit is selected from the group consisting of SEQ ID No. 4, SEQ ID No. 90, SEQ ID No. 93, SEQ ID No. 96, SEQ ID No. 99, SEQ ID No. 102, SEQ ID No. 105, SEQ ID No. 108, SEQ ID No. 111, SEQ ID No. 114, SEQ ID No. 117, or functionally active variants thereof having at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97 or 98 % sequence identity.~~

~~22. The cp caspase-2 of any one of claims 1 to 21, comprising~~

~~i. an N-terminal and/or C-terminal truncation, and/or~~

~~ii. an N-terminal and/or C-terminal extension.~~

~~23.~~10. The cp caspase-2 of any one of claims 1 to ~~22~~9, comprising one or more amino acid substitutions, selected from

i. Gly^{171}, substituted with D, or an amino acid selected from the group consisting of R, K, E, Q, N, A, S, T, P, H, Y
ii. Glu^{105}, substituted with V, or an amino acid selected from the group consisting of C, L, I, M, F, W, R, K, D, Q, N
iii. Glu^{172}, substituted with V, or an amino acid selected from the group consisting of C, L, I, M, F, W, R, K, D, Q, N
iv. Asp^{282}, substituted with E, or T, or an amino acid selected from the group consisting of R, K, Q, N, G, A, S, P, H, Y
v. Val^{225}, substituted with G, or an amino acid selected from the group consisting of A, S, T, P, H, Y, C, L, I, M, F, W

vi. Lys^{83}, substituted with E, or an amino acid selected from the group consisting of R, D, Q, N,

vii. His^{185}, substituted with A, or an amino acid selected from the group consisting of G, S, T, P, Y,

viii. Val^{255}, substituted with M, or an amino acid selected from the group consisting of C, L, I, F, W, and/or

ix. Asp^{285}, substituted with E, or Y, or an amino acid selected from the group consisting of R, K, Q, N, G, A, S, T, P, H,

with reference to the positions of SEQ ID No. 6, or positions functionally equivalent to positions of SEQ ID No. 6.

~~24.~~11. The cp caspase-2 of any one of claims 1 to ~~22~~ 10 comprising amino acid substitutions at positions of SEQ ID No. 6, or at positions functionally equivalent to positions of SEQ ID No. 6, selected from the group consisting of

i. His^{185} and Asp^{282}, ~~specifically~~ preferably comprising H185A and D282T substitutions;

ii. Glu^{105} and Asp^{285}, ~~specifically~~ preferably comprising E105V and D285E substitutions;

iii. Glu^{105}, Gly^{171}, Val^{225} and Asp^{282}, ~~specifically~~ preferably comprising E105V, G171D, V225G and D282E substitutions;

iv. Glu^{105}, Gly^{171}, Val^{225}, Asp^{282} and Asp^{285}, ~~specifically~~ preferably comprising E105V, G171D, V225G, D282E and D285E substitutions;

v. Lys^{83}, Glu^{105}, Glu^{172}, Val^{255} and Asp^{285}, ~~specifically~~ preferably comprising K83E, E105V, E172V, V255M and D285Y substitutions;

vi. Glu^{105} and Gly^{171}, ~~specifically~~ preferably comprising E105V and G171D substitutions;

vii. Glu^{105} and Glu^{172}, ~~specifically~~ preferably comprising E105V and E172V substitutions; and

viii. Gly^{171} and Glu^{172}, ~~specifically~~ preferably comprising G171D and E172V substitutions,

~~wherein said cp caspase-2 has increased P1' tolerance compared to a cp caspase-2 without the respective amino acid substitution, optionally~~preferably wherein said cp caspase-2 comprises an SS propeptide comprising an amino acid substitution

to Ala at position Asp^{14} of SEQ ID No. 2 or at a position functionally equivalent to position Asp^{347} of SEQ ID No. 11.

~~25.~~12. The cp caspase-2 of any one of claims 1 to ~~22~~11, comprising SEQ ID No. 6 ~~and~~ with one or more amino acid substitutions at position 171, 105, 172, 282, 225, 83, 185, 255, or 285 of SEQ ID No. 6; or comprising SEQ ID No. 64, 66 or 68 with one or more amino acid substitutions at a position functionally equivalent to position 171, 105, 172, 282, 225, 83, 185, 255, or 285 of SEQ ID No. 6, or any combination thereof, preferably comprising any one or more of amino acid substitutions G171D, E105V, E172V, D282E, D282T, V225G, K83E, H185A, V255M, D285Y and D285E.

~~26. The cp caspase-2 of claim 25, comprising any one or more of amino acid substitutions G171D, E105V, E172V, D282E, D282T, V225G, K83E, H185A, V255M, D285Y and D285E.~~

~~27~~13. The cp caspase-2 of any one of claims 1 to ~~26~~12, comprising an amino acid sequence selected from the group consisting of SEQ ID No. 1, 13, 17, 18, 23, 24, 51, 52, 54, 70, 71, 72, 78, 79, 86, 87, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191 and 192, or an amino acid sequence having at least 70%~~, 75%, 80%, 85%, 90%, specifically at least 95%, specifically at least 99%~~ sequence identity with any one of SEQ ID No. 1, 13, 17, 18, 23, 24, 51, 52, 54, 70, 71, 72, 78, 79, 86, 87, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191 and 192.

~~28~~14. The cp caspase-2 of any one of claims 1 to ~~27~~13, comprising a C-terminal tag and an amino acid substitution at positions 285 and 292 of SEQ ID No. 6 or at a position functionally equivalent to positions 285 and 292 of SEQ ID No. 6, specifically comprising substitutions to Glu and Ser (D285E and D292S).

~~29~~15. The cp caspase-2 of any one of claims 1 to ~~28~~14, wherein said cp caspase-2 is recruited by a recognition site for proteolytic cleavage, comprising 5 amino acids of the sequence P5 P4 P3 P2 P1, wherein

P1 can be any amino acid, preferably it is D or E,
P2 can be any amino acid, preferably it is A,
P3 can be any amino acid, preferably it is V,
P4 can be any amino acid, preferably it is D, and
P5 can be any amino acid, preferably it is V.

~~30~~16. A method of producing a circular permuted caspase-2 (cp caspase-2) comprising the steps of

i. cloning a nucleotide sequence encoding a cp caspase-2 under the control of a promoter into an expression vector,
ii. transforming a host cell with said vector,
iii. culturing the transformed host cell under conditions wherein the cp caspase-2 is expressed,
iv. optionally isolating the cp caspase-2 from the host cell culture, optionally by disintegrating the host cells, and
v. optionally purifying the cp caspase-2.

~~31~~17. The method of claim ~~30~~16, wherein the cp caspase-2 is the cp caspase-2 of any one of claims 1 to ~~29~~15.

~~32~~18. The method of claim ~~30 or 31~~16 or 17, wherein the promoter is selected from the group consisting of T7 promoter/operator, XylS/Pm regulator/promoter or variants of the Pm promoter, araBAD promoter/operator, T5, T7A1, T7A2, T7A3 promoter/operator, phoA promoter/regulator and the trp promoter/operator system.

~~33~~19. The method of any one of claims ~~30 to 32~~16 to 18, wherein the cp caspase-2 comprises a solubility enhancement tag, ~~selected from the group consisting of T7C, T7B, T7B1, T7B2, T7B3, T7B3, T7B4, T7B5, T7B6, T7B6, T7B7, T7B8, T7B9, T7B10, T7B11, T7B12, T7B13, T7A, T7A1, T7A2, T7A3, T7A4, T7A5, T7AC, T3, N1, N2, N3, N4, N5, N6,N7, calmodulin-binding peptide (CBP), DsbA, DsbC, poly Arg, poly Lys, G B1 domain, protein D, Z domain of Staphylococcal protein A, and thioredoxin tag,~~ preferably it comprises a T7AC or a T7A3 tag.

~~34~~20. The method of any one of claims ~~30 to 33~~16 to 19, wherein the cp caspase-2 comprises an affinity tag, preferably a His tag, and even more preferably a 6-His tag.

~~35~~21. The method of any one of claims ~~30 to 34~~16 to 20, wherein the host cell is ~~a eukaryotic or prokaryotic host cell, preferably a yeast cell or a bacterial cell, and even more preferably~~ an *E. coli* cell.

~~36. The method of any one of claims 30 to 35, wherein the cp caspase-2 comprises an N-terminal tag comprising an affinity tag, preferably a His-tag and even more preferably a 6-His tag, and a solubility enhancement tag, preferably T7AC or T7A3.~~

~~37. The method of claim 36, wherein the cp caspase-2 further comprises a linker between the affinity tag and the solubility enhancement tag.~~

~~38. The method of claim 36 or 37, wherein the cp caspase-2 comprises the following elements fused to its N-terminus, in the order from N- to C-terminus:~~

~~a. an affinity tag, preferably 6-His tag;~~
~~b. optionally a linker;~~
~~c. a solubility enhancement tag, preferably T7AC or T7A3;~~
~~d. optionally a linker; and~~
~~e. a cp caspase-2, comprising the small subunit N-terminal to the large subunit.~~

~~39. The method of claim 36 or 37, wherein the cp caspase-2 comprises the following structure, in the order from N- to C-terminus:~~

~~f. a solubility enhancement tag, preferably T7AC or T7A3;~~
~~g. optionally a linker;~~
~~h. an affinity tag, preferably 6-His tag;~~
~~i. optionally a linker; and~~
~~j. a cp caspase-2, comprising the small subunit N-terminal to the large subunit.~~

~~40~~22. The method of any one of claims ~~30 to 39~~16 to 21, wherein culturing of step (iii) comprises a fed-batch phase for expression of the cp-caspase-2, said fed batch phase specifically cromprising a growth rate µ of about 0,01-0,1 h^{-1}, and induction of expression of the cp caspase-2 by addition of IPTG at a concentration of about 0,01 – 1,5 µmol/g actual CDM (cell dry mass).

~~41~~23. The method of claim ~~40~~22, wherein the growth rate µ is about 0,03-0,07 h^{-1}, preferably it is about 0,05-0,07 h^{-1} or 0,03-0,05 h^{-1}, preferably it is any of about 0,03, 0,05 or 0,07 h^{-1}.

~~42~~24. The method of claim ~~40 or 41~~22 or 23, wherein the IPTG concentration is about 0,5-1,3 µmol/g CDM, preferably it is about 0,5-0,9 µmol/g CDM or about 0,9-1,3 µmol/g CDM, preferably it is about 0,5, 0,9 or about 1,3 µmol/g CDM.

~~43~~25. The method of any one of claims ~~40 to 42~~22 to 24, wherein culturing of step (iii) further comprises a first fed-batch phase for the production of biomass, prior to the fed-batch phase for the expression of the cp caspase-2, said first fed-batch phase comprising a growth rate, µ of about 0,07-0,3 h^{-1}.

~~44~~26. The method of claim ~~43~~25, wherein the growth rate µ is about 0,1-0,2 h^{-1}, preferably about 0,13-0,21 h^{-1}, even more preferably about 0,16-0,18 h^{-1} and most preferably it is about 0,17 h^{-1}.

~~45. The method of any one of claims 30 to 44, wherein the cp caspase-2 is purified using affinity chromatography, preferably IMAC.~~

~~46. A cp caspase-2 obtained by the method of any one of claims 30 to 45.~~

~~47~~27. A method of producing a protein of interest (POI) ~~comprising an authentic N-terminus~~, comprising the steps of:

i. providing a fusion protein comprising from N- to C-terminus one or more tags, optionally one or more tag-linker sequences and a caspase recognition site N-terminally fused to the POI, wherein said caspase recognition site is specifically recognized by the cp caspase-2 of any one of claims 1 to ~~29~~15,
ii. contacting said fusion protein with said cp caspase-2 for a period of time sufficient for said cp caspase-2 to cleave the fusion protein, and
iii. optionally purifying the POI.

~~48~~28. ~~A method of producing a protein of interest (POI) comprising an authentic N-terminus~~The method of claim 27, comprising the steps of:

i. expressing the fusion protein comprising from N- to C-terminus optionally one or more tags, optionally one or more tag-linker sequences and a caspase recognition site N-terminally fused to the POI, wherein said caspase recognition site is specifically recognized by the cp

caspase-2 of any one of claims 1 to ~~29~~15; and the cp caspase-2 of any one of claims 1 to ~~29~~ 15 specifically recognizing the recognition site of the fusion protein, in the same host cell,

ii. optionally, wherein said fusion protein and cp caspase-2 are under the same promoter,

iii. cultivating the host cell, wherein said cp caspase-2 cleaves the fusion protein in vivo in the cell, and

iv. optionally isolating the POI from the cell and optionally purifying the POI.

~~49~~29. The method of claim ~~47 or 48~~27 or 28, wherein the fusion protein comprises a caspase recognition site comprising 5 amino acids of the sequence P5 P4 P3 P2 P1, and a cleavage site P1/P1', wherein P1' is the N-terminal amino acid of the POI.

~~50~~30. The method of claim ~~47 or 48~~27 or 28, wherein the fusion protein and the cp caspase-2 are under transcriptional control of different promoters and wherein the expression of the cp caspase-2 is induced after expression of the fusion protein.

~~51~~31. The method of any one of claims ~~47 to 50~~27 to 30, wherein the fusion protein comprises the cp caspase-2 of any one of claims 1 to ~~29~~15, specifically wherein the fusion protein comprises ~~the~~said cp caspase-2 ~~of any one of claims 1 to 29~~ at its N- or C-terminus, and wherein the fusion protein comprises the following structure from N- to C-terminus:

i. one or more N-terminal tags,

ii. optionally one or more tag-linker sequences and

iii. a caspase recognition site comprising 5 amino acids of the sequence P5 P4 P3 P2 P1,

iv. a cleavage site P1/P1',

v. a POI, and

wherein P1' is the N-terminal amino acid of said POI and said cp caspase-2 specifically recognizes said recognition site.

~~52~~32. The method of any one of claims ~~47 to 51~~27 to 31, comprising the steps of:

i. expressing a fusion protein in a host cell comprising the following structure from N- to C-terminus:

a. an N-terminal affinity tag,

b. optionally a linker sequence,

c. a caspase recognition site,

d. a cleavage site P1/P1', and

e. a POI,

wherein P1' is the N-terminal amino acid of the POI, and wherein said recognition site is specifically recognized by the cp caspase-2 of any one of claims 1 to ~~29~~15,

ii. isolating said fusion protein

iii. purifying said fusion protein using the N-terminal affinity tag,

iv. providing a cp caspase-2 of any one of claims 1 to ~~29~~15, specifically recognizing the recognition site of the fusion protein,

v. contacting said fusion protein with said cp caspase-2 for a period of time sufficient for said cp caspase-2 to cleave the fusion protein,

vi. optionally removing the cleaved affinity tag, and optionally the non-cleaved fusion protein using the affinity tag and the cp caspase-2, and

vii. optionally further purifying the POI.

~~53~~33. The method of claim ~~52~~32, wherein the cp caspase-2 comprises at its N- or C-terminus an affinity tag identical to the affinity tag of the fusion protein and wherein the cp caspase-2 is removed in step vi. using said affinity tag.

~~54~~34. The method of claim ~~52 or 53~~32 or 33, comprising the steps of

i. expressing a fusion protein comprising one or more N-terminal affinity tags, optionally one or more tag-linker sequences, a caspase recognition site and a cleavage site P1/P1', wherein P1' is the N-terminal amino acid of the POI, and a POI, wherein said recognition site is specifically recognized by the cp caspase-2 of any one of claims 1 to ~~29~~15, in a host cell, and

ii. isolating the fusion protein and binding / capturing the fusion protein on a solid support using the affinity tag,

iii. providing a cp caspase-2 of any one of claims 1 to ~~29~~15, specifically recognizing the recognition site of the fusion protein,

iv. contacting said cp caspase-2 with the bound / captured fusion protein for a period of time sufficient for said cp caspase-2 to cleave the fusion protein,

v. releasing the POI from the solid support, and

vi. isolating and optionally further purifying the POI.

~~55. The method of claim 54, wherein the cp caspase-2 and the fusion protein comprise an identical affinity tag, allowing binding of the fusion protein and the caspase on the solid support and release of the POI upon cleavage by the caspase.~~

~~56. The method of claim 54 or 55, wherein the solid support is a column, specifically a chromatography column, more specifically an immobilized metal affinity chromatography column (IMAC).~~

~~57. The method of any one of claims 47 to 56, wherein a flow-through reactor comprising immobilized cp caspase-2 of any one of claims 1 to 29 is used.~~

~~58~~35. An isolated nucleotide sequence, preferably comprised in a vector or in an expression cassette, encoding the cp caspase-2 of any one of claims 1 to ~~29~~15.

~~59. A vector comprising the nucleotide sequence of claim 58, specifically it is a bacterial expression vector.~~

~~60. An expression cassette comprising the nucleotide sequence of claim 58 operably linked to regulatory elements.~~

~~61~~36. A host cell or a host cell line expressing the cp caspase-2 of any one of claims 1 to ~~29~~15, wherein the host cells are selected from the group consisting of bacterial cells, yeast cells, insect cells, mammalian cells and plant cells, preferably the host cells are bacterial or yeast cells selected from the group consisting of *E. coli*, *Pseudomonas sp.*, *Bacillus sp., Streptomyces sp., Saccharomyces sp., Schizosaccharomyces sp.*, *Pichia sp., Kluyveromyces sp.* and *Hansenula sp.*.

~~62. An expression system comprising the vector of claim 59 or the expression cassette of claim 60 and a host cell of claim 61.~~

~~63~~37. Use of the cp caspase-2 of any one of claims 1 to ~~29~~ 15 for the *in ~~vivo~~ vitro* cleavage of a substrate in a non-human organism, preferably wherein the non-human organism is a prokaryotic organism, especially preferred *E. coli*.

~~64. The use of claim 63, wherein the non-human organism is a prokaryotic organism, specifically it is *E. coli*.~~

~~65~~38. Use of the cp caspase-2 of any one of claims 1 to ~~29~~ 15 for the production of a protein of interest (POI).

~~66. The use of claim 65, wherein the POI comprises an authentic N-terminus.~~

~~67. A fusion protein comprising the following structure from N- to C-terminus:~~

~~i. a tag sequence comprising a caspase recognition site comprising 5 amino acids of the sequence P5-P4-P3-P2-P1, specifically recognized by the cp caspase-2 of any one of claims 1 to 29,~~

~~ii. a cleavage site P1/P1', wherein P1' is the N-terminal amino acid of the protein of interest (POI), and~~

~~iii. a POI.~~

~~68. The fusion protein of claim 67, wherein the tag sequence further comprises one or more tags selected from the group consisting of affinity tags, solubility enhancement tags and monitoring tags.~~

~~69. The fusion protein of claim 67, further comprising one or more tag-linker sequences.~~

~~70. A kit comprising the caspase of claim 74 and/or the cp caspase-2 of any one of claims 1 to 29, specifically for cleaving a fusion protein of any one of claims 67 to 69 or claim 90 or 91.~~

~~71. The kit of claim 70, comprising an expression vector comprising a polynucleotide encoding the protein tag of claims 75 to 89.~~

~~72~~39. The cp caspase-2 of any one of claims 1 to ~~29~~15, for use in the treatment of a disease, preferably for use in the treatment of cancer, Alzheimer's disease, Parkinson's disease or inflammatory disease.

~~73. The cp caspase-2 of any one of claims 1 to 29, for use in the treatment of cancer, Alzheimer's disease, Parkinson's disease or inflammatory disease.~~

~~74~~40. A caspase-2 comprising a sequence selected from the group consisting of SEQ ID No.11, SEQ ID No. 89, SEQ ID No.92, SEQ ID No. 95, SEQ ID No. 98, SEQ ID No. 101, SEQ ID No. 104, SEQ ID No. 107, SEQ ID No. 110, 113, and SEQ ID No. 116, or a functionally active variant thereof comprising a sequence with at least 70% sequence identity to SEQ ID No.11, SEQ ID No. 89, SEQ ID No.92, SEQ ID No. 95, SEQ ID No. 98, SEQ ID No. 101, SEQ ID No. 104, SEQ ID No. 107, SEQ ID No. 110, 113, and SEQ ID No. 116, and comprising one or more amino acid substitutions at positions 409, 431, 212, 213, 266, 226, 296, 323 or 326 of SEQ ID No. 11, or at a position functionally equivalent to any of positions 409, 431, 212, 213, 266, 226, 296, 323 or 326 of SEQ ID No. 11, or a combination thereof~~, wherein said amino acid substitution increases P1' tolerance compared to a caspase-2 comprising the same sequence but not comprising said amino acid substitutions~~.

~~75. A protein tag for enhanced expression of a POI, comprising a solubility enhancement tag and the amino acid sequence VDVAD (SEQ ID NO:45), wherein the sequence VDVAD is located at the C-terminus of the protein tag.~~

~~76. The tag of claim 75, wherein the solubility enhancement tag is selected from the group consisting of T7C, T7B, T7B1, T7B2, T7B3, T7B3, T7B4, T7B5, T7B6, T7B6, T7B7, T7B8, T7B9, T7B10, T7B11, T7B12, T7B13, T7A, T7A1, T7A2, T7A3, T7A4, T7A5, T3, N1, N2, N3, N4, N5, N6, N7 and T7AC.~~

~~77. The tag of claim 76, wherein the solubility enhancement tag is T7AC or T7A3.~~

~~78. The tag of any one of claims 75 to 77, further comprising a histidine tag sequence, preferably comprising 1-20 histidine residues, even more preferably it is a 3-His, 6-His or 9-His tag sequence.~~

~~79. The tag of any one of claims 75 to 78, wherein the solubility enhancement tag is located at the N-terminus of said protein tag.~~

~~80. The tag of any one of claim 78, wherein the histidine tag sequence is located at the N-terminus of said protein tag.~~

~~81. The tag of any one of claims 75 to 80, further comprising one or more linker sequences comprising one or more amino acid residues.~~

~~82. The tag of claim 81, wherein said one or more linker sequences are located between the VDVAD sequence and the solubility enhancement tag and/or the histidine tag sequence.~~

~~83. The tag of claim 81 or 82, wherein the one or more amino acid residues of the linker sequence are any of the naturally occurring amino acids or derivatives thereof, preferably selected from the group consisting of G, S, A, T and N.~~

~~84. The tag of any one of claims, 81 to 83, wherein the linker sequence is GSG.~~

~~85. The tag of any one of claims 81 to 83, wherein the linker sequence is GSGSGSG.~~

~~86. The tag of any one of claims 75 to 85, further comprising a signal peptide at the N-terminus of said protein tag.~~

~~87. The tag of claim 86, wherein the signal peptide is selected from the group consisting of ompA (outer membrane protein A), DsbA (Thiol:disulfide interchange protein), MalE (maltose-binding protein), PelB (pectate lyase B) from Erwinia carotovora, PhoA (alkaline phosphatase), OmpC (outer-membrane protein C), OmpF (outer-membrane protein F), OmpT (protease VII), Endoxylanase from Bacillus sp., LamB (λ receptor protein), Lpp (murein lipoprotein), LTB (heat-labile enterotoxin subunit B), PhoE (outer-membrane pore protein E), and StII (heat-stable enterotoxin 2).~~

~~88. The tag of any one of claims 75 to 87, wherein the tag comprises one of the following structures from N- to C-terminus:~~

~~k. T7AC – 6-His – VDVAD;~~
~~l. T7A3 – 6-His – VDVAD;~~
~~m. T7AC – 6-His – GSG – VDVAD;~~
~~n. T7A3 – 6-His – GSG – VDVAD;~~
~~o. T7AC – 6-His – GSGSGSG – VDVAD;~~
~~p. T7A3 – 6-His – GSGSGSG – VDVAD;~~

q. 6-His – T7AC – VDVAD;

r. 6-His – T7A3 – VDVAD;

s. 6-His – T7AC – GSG – VDVAD;

t. 6-His – T7A3 – GSG – VDVAD;

u. 6-His – T7AC – GSGSGSG – VDVAD;

v. 6-His – T7A3 – GSGSGSG – VDVAD.

89. The tag of claim 86 or 87, wherein the tag comprises one of the following structures from N-to C-terminus:

w. ompA signal peptide – T7AC – 6-His – VDVAD;

x. ompA signal peptide – T7A3 – 6-His – VDVAD;

y. ompA signal peptide – T7AC – 6-His – GSG – VDVAD;

z. ompA signal peptide – T7A3 – 6-His – GSG – VDVAD;

aa. ompA signal peptide – T7AC – 6-His – GSGSGSG – VDVAD;

bb. ompA signal peptide – T7A3 – 6-His – GSGSGSG – VDVAD;

cc. ompA signal peptide – 6-His – T7AC – VDVAD;

dd. ompA signal peptide – 6-His – T7A3 – VDVAD;

ee. ompA signal peptide – 6-His – T7AC – GSG – VDVAD;

ff. ompA signal peptide – 6-His – T7A3 – GSG –VDVAD;

gg. ompA signal peptide – 6-His – T7AC – GSGSGSG –VDVAD;

hh. ompA signal peptide – 6-His – T7A3 – GSGSGSG – VDVAD.

90. A fusion protein comprising the protein tag of any one of claims 75 to 89 and a POI, wherein the N-terminus of the POI is fused to the C-terminus of said protein tag.

91. The fusion protein of claim 90, wherein the N-terminus of the POI is directly fused to the C-terminus of the protein tag, which C-terminus is the sequence VDVAD.

92. A method of producing a POI, comprising the steps of:

i. providing the fusion protein of claim 90 or 91 comprising a POI,

ii. contacting said fusion protein with a circular permuted caspase-2 (cp caspase-2) for a period of time sufficient for said cp caspase-2 to cleave the fusion protein thereby releasing the POI, and

iii. optionally purifying the POI.

~~93. The method of claim 92, further comprising the following steps:~~

~~iv. cloning a nucleotide sequence encoding the fusion protein of claim 90 or 91, under the control of a promoter into an expression vector,~~

~~v. transforming a host cell with said vector,~~

~~vi. culturing the transformed host cell under conditions wherein said fusion protein is expressed,~~

~~vii. optionally isolating said fusion protein from the host cell culture, optionally by disintegrating the host cells, and~~

~~viii. purifying said fusion protein using IMAC chromatography,~~

~~ix. contacting said fusion protein with a circular permuted caspase-2 (cp caspase-2) for a period of time sufficient for said cp caspase-2 to cleave the fusion protein thereby releasing the POI, and~~

~~x. optionally further purifying the POI,~~

~~xi. optionally modifying the POI and~~

~~xii. optionally formulating the POI.~~

~~94. The method of claim 92 or 93, wherein the promoter is selected from the group consisting of T7 promoter/operator, XylS/Pm regulator/promoter or variants of the Pm promoter, araBAD promoter/operator, T5, T7A1, T7A2, T7A3 promoter/operator, phoA promoter/regulator and the trp promoter/operator system.~~

~~95. The method of claims 65 to 66, wherein the host cell is a eukaryotic or prokaryotic host cell, preferably a yeast or a bacterial cell, preferably it is an *E. coli* cell.~~