IT in Industry, Vol. 9, No.3, 2021 Published Online 30-Dec-2021

Predicting Suitable Linker for Fusion Protein Using Soft Computing Techniques

U.Subhashini[1], P.Bhargavi[2], S.Jyothi[3]
[1]Research Scholar, [2]Assistant professor, [3]Professor
[1,2,3] Department of Computer Science
Sri Padmavati Mahila Visvavidyalayam Tirupati-517 502
[1]subhau7@gmail.com,[2]pbhargavi18@yahoo.co.in, [3]jyothi.spmvv@gmail.com

Abstract— **Protein is a highly complicated substance that occurs naturally, that contains amino-acid residues bounded with peptide bonds. Proteins contain numerous important biological compounds, for example, antibodies, hormones and enzymes. Also, 20 stable amino acids make up a protein. In medicine, proteins are used as antibodies to create vaccines. Initially, proteins are acquired from the cells of plants, animals and microorganisms. There is a remarkable rise in the reproduction of natural proteins through recombinant DNA (rDNA) technology. It further focused on developing "de novo" proteins, which are non-natural and are known as fusion proteins. A fusion protein is a protein formed by combining at least two types of protein domains. Fusion proteins improve bioactivities by the broad range of biotechnological and biopharmaceutical appliances. A victorious building of a recombinant fusion protein needs two indispensable elements: the linkers and the component proteins. The option of the component proteins using a preferred function of a fusion protein manufactured goods is generally relatively uncomplicated. Conversely, Linkers are small peptide sequences. Linkers do not change the operation of individual proteins to which they are attached. Straight fusion of helpful domains with no linker might guide numerous unwanted results, containing misfolding of the fusion proteins, small yield at protein manufacture, or damaged bioactivity. Prediction of a suitable linker for fusion protein is costly because of the prices related to crystallography; electron microscopy also takes more time. In this situation, further soft computing presents numerous feasibilities through creating inexpensive, high-quality results—soft computing methods utilized for fusion protein linker forecast. Here the fuzzy logic is used to forecast the suitable linker for the fusion protein. The experimental results show that the proposed soft computing based linker prediction can predict the suitable linker efficiently.**

Keywords — Peptide, Linker, Linker Flexibility, Fusion protein, Fuzzy logic.

1. INTRODUCTION

Proteins are huge, multifaceted molecules in our cells also are the quintessence of life procedure. They are the essential ingredients of the entire protoplasm, forming the living cell with its operation. Proteins are natural compounds prepared up of lesser units known as amino acids covalently connected by peptide bonds. Twenty normal amino acids could be merged to create a protein [1]. Proteins have well-known appliances at medical analysis with pharmaceuticals [2]. At first, natural proteins were taken out from plants, human sources, and animals. After that, Recombinant DNA technology commenced a novel region of investigates also apply features of biology [3]. Because then, an important raise has been observing at replicate natural proteins through Recombinant DNA technology. It further concentrated on implementing de novo proteins that do not have at nature, also are known as fusion proteins. Fusion proteins are proteins built through merging two or more various domain proteins [4]. Fusion proteins are further known as hybrid proteins or chimeric proteins. Fusion protein attains numerous operational possessions resulting from every unique protein, including biological activity [5]. Most of the examiner studies exposed that a few fusion proteins contain superior constancy with efficiency over proteins happening in nature [6]. Over the years, investigators have been utilizing the recombinant DNA technology for the building of the fusion proteins because of the broad diversity of its appliances, for example, enhancing enzyme activity [7], drug growth [8], half-life expansion, [9], biomaterial plan [10] investigation of protein-protein communications [11], [12] with tissue engineering. Victorious building of the fusion protein primarily needs the preferred proteins with its compatibility. If part domains are not friendly then it guides to misfolding [13]. The simplest technique of fusing chosen fields is an end to end genetic fusion. In a few cases, straight fusing is easy with workings most excellent where N or C terminal areas of the element proteins perform like a "bridge" to present sufficient space among protein domains for right folding [14], [15]. Though this plan fails while the N or C terminal is not supple or extended sufficient to evade steric hindrance, that decreases the amount of liberty at protein bioactivity also might provide increase to unwanted results, for example, lack of correct protein folding, small profit at protein manufacture also reduced bioactivity [16], [17], [18]. For this purpose, numerous protein bioactivity lessons produced fusing of the chosen proteins devoid of linker consequences at reduced biological activity [19], [20]. In this circumstance, the most generally utilized technique for building the fusion protein is the linker arbitrates tandem fusion technique. In this technique, the fusion protein is attained through fusing two proteins with an appropriate linker.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

Comprehension of linkers with their biochemical characteristics is vital at developing linkers to build various recombinant fusion proteins. Linkers are small peptide sequences that happen among protein domains. Crystallographers frequently discover domains with linkers as deciding the arrangement of multiple domain proteins.

Though the domain linkers are not found openly in various circumstances, there is a requirement to use computational techniques to determine the linkers using a protein sequence. Forecasting of appropriate linkers for fusion protein is expensive due to the prices associated with crystallography; electron microscopy further takes a long time. In this situation, soft computing presents numerous feasibilities by constructing inexpensive, high-quality results—soft computing methods for the domain-linker forecast. In this paper, the Fuzzy Logic technique is used for predicting suitable linkers for construction of fusion proteins.

The remainder of the paper is structured like follows. Section 2 explains the suitable linker prediction technique using soft computing. Followed by, section 3 discusses the experimental results. Lastly, the lessons learned with conclusions are summarized in Section 4.

2. SOFT COMPUTING BASED SUITABLE LINKER PREDICTION TECHNIQUE

Generally fusion proteins are composed of discrete proteins connected by linkers. The use of linkers in the creation of fusion protein offers unlimited possibilities. In this paper we considered non-redundant dataset of 1932 natural linkers collected from George et al. [21]. The dataset comprises PDB code, Region, Length, Dist, SA, Sequence, Secondary Structure and Description. The fuzzy logic technique used to predict suitable linker for constructing fusion proteins. Proteins are natural compounds prepared up of lesser units known as amino acids covalently connected by peptide bonds. Twenty normal amino acids could be merged to create a protein. A codon is a sequence of three RNA or DNA nucleotides corresponding to a specific amino acid during protein synthesis. Table.1 shows Codon-Amino Acid abbreviations.

Table.1: Codon-Amino Acid Abbreviations

Codon	Amino Acid	Abbreviations
AAA	Lysine	Lys
AAC	Asparagine	Asn
AAG	Lysine	Lys
AAT	Asparagine	Asn
ACA	Threonine	Thr
ACC	Threonine	Thr
ACG	Threonine	Thr
ACT	Threonine	Thr
AGA	Arginine	Arg
AGC	Serine	Ser
AGG	Arginine	Arg
AGT	Serine	Ser
ATA	Isoleucine	Ile
ATC	Isoleucine	Ile
ATG	Methionine	Met
ATT	Isoleucine	Ile
CAA	Glutamine	Gln
CAC	Histidine	His
CAG	Glutamine	Gln
CAT	Histidine	His
CCA	Proline	Pro
CCC	Proline	Pro
CCG	Proline	Pro
CCT	Proline	Pro
CGA	Arginine	Arg
CGC	Arginine	Arg
CGG	Arginine	Arg
CGT	Arginine	Arg
CTA	Leucine	Leu
CTC	Leucine	Leu
CTG	Leucine	Leu
CTT	Leucine	Leu
GAA	Glutamate	Glu
GAC	Aspartate	Asp
GAG	Glutamate	Glu
GAT	Aspartate	Asp
GCA	Alanine	Ala
GCC	Alanine	Ala
GCG	Alanine	Ala
GCT	Alanine	Ala
GGA	Glycine	Gly
GGC	Glycine	Gly
GGG	Glycine	Gly
GGT	Glycine	Gly
GTA	Valine	Val

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

GTC	Valine	Val
GTG	Valine	Val
GTT	Valine	Val
TAA	Termination (ochre)	Ter
TAC	Tyrosine	Tyr
TAG	Termination (amber)	Ter
TAT	Tyrosine	Tyr
TCA	Serine	Ser
TCC	Serine	Ser
TCG	Serine	Ser
TCT	Serine	Ser
TGA	Termination (opal or umber)	Ter
TGC	Cysteine	Cys
TGG	Tryptophan	Trp
TGT	Cysteine	Cys
TTA	Leucine	Leu
TTC	Phenylalanine	Phe
TTG	Leucine	Leu
TTT	Phenylalanine	Phe

As an indispensable component of recombinant fusion proteins, linkers have shown increasing significance in building stable, bioactive fusion proteins. Researchers configured linkers are generally classified into three types according to their structures: flexible Linkers, rigid Linkers and in vivo cleavable Linkers [22]. Linkers can propose numerous benefits for the construction of fusion proteins, such as the development of biological functions, increasing exposure yield and achieving desirable pharmacokinetic profiles.

For example, many properties of linkers, like length, hydrophobicity, amino acid residues, and secondary structure, are essential for appropriate linker prediction. For instance, long linkers have a lot of potentials to reveal solvent. Flexible linkers are commonly used when connected protein domains require a certain amount of movement or communication. They are usually synthesized by small, non-polar (for example, Gly) or polar (for example, Ser) amino acids. The most commonly used flexible linkers mainly display lengths of Gly and Ser residues ("GS" linker). The flexible linker has "GGGGS" at all times. To fuse Gly and Ser type proteins, Flexible Linker utilized. To fuse Hydrophilic (Arg, Asn, Asp, Gln) amino acids type proteins Flexible linker used. Flexible linkers always present good solubility and flexibility [23]. Table 2 shows necessary codons for flexible linkers.

Table 2: Codons for Flexible Linkers

Amino Acid Abbreviations	Codons
Gly	GGT, GGC, GGA, GGG
Ser	AGT, AGC, TCT, TCC, TCA, TCG
Arg	AGA, AGG
Asn	AAT, AAC
Asp	GAT, GAC
Gln	CAA, CAG

While flexible linkers have the benefit of being capable of passively connect functional domains and allow a certain amount of movement, the shortage of rigidity of these linkers may be a limitation. There are numerous examples in the literature of the use of flexible linkers published in lack of biological function or poor exposure yields. For example, a granulocyte colony promoting factor (G-CSF) fusion protein should be pronounced with a flexible (GGGGS)3 linker. In another report, the chain ability of the immunoglobulin of the protein G domain in a protein G-vercula luciferase fusion protein was not regained and then placed in a flexible GGGGS linker [24]. In these cases, the ineffectiveness of flexible linkers was due to the inefficient separation of protein domains with each other. Under these circumstances, rigid linkers have been used victoriously to retain the repair distance among domains and remain their autonomous functions [25]. The linkers that make up the alpha helix with the sequence of (EAAAK) n have been used to construct numerous recombinant fusion proteins. To fuse α-helical structures (Ala, Glu, Leu, Met, Lys) type proteins, Rigid Linker utilized. Moreover, to fuse Pro type proteins, Rigid Linker used. Table 3 shows the necessary codons for rigid linkers.

Table 3: Codons for Rigid Linkers

Amino Acid Abbreviations	Codons
Ala	GCT, GCC, GCA, GCG
Glu	GAA, GAG
Leu	CTT, CTC, CTA, CTG
Met	ATG
Lys	AAA, AAG
Pro	CCT, CCC, CCA, CCG

Therefore, the linkers examined commonly include stable peptide sequences that will not be more favorably cleaved in vivo. These stable linkers mutually connect functional domains to play as one molecule through the in

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

vivo processes. The stable linkage among available parts gives numerous advantages, such as a prolonged plasma half-life (for example, Fc or albumin fusions). But, it further has many potential drawbacks containing steric hindrance among functional domains, reduced bioactivity, and modified bio-distribution and metabolism of the protein moieties due to the interference among domains [26]. Under these situations, cleavable linkers are suggested to let out free functional domains in vivo. The plan of in vivo cleavable linker in recombinant fusion proteins is rather complicated [27][28]. Unlike the versatility of cross linking agents existing for chemical conjugation approaches, linkers in recombinant fusion proteins are needed to be oligopeptides [29]. The linkers suggested in this section take advantage of the unique in vivo processes and are cleaved below particular conditions, such as decreasing reagents or proteases [30]. This type of linker may reduce steric hindrance, enhance bioactivity, or attain autonomous actions/metabolism of own domains of recombinant fusion proteins, subsequently linker cleavage [31]. To fuse Cys type proteins, In vivo cleavable linker utilized. Table 4 shows Codons for Cleavable linkers.

Table 4: Codons for Cleavable Linkers

Amino Acid Abbreviations	Codons
Cys	TGT, TGC

Fuzzy logic based suitable linker prediction for constructing fusion protein showed in below Algorithm. This algorithm first predicts the suitable linker for two protein sequences that are fused (Step 1 - 26). Followed by, it finds the candidate linkers (Steps 27 to 41). After that, it predicts suitable linkers (Step 42 - 62).

Algorithm: Predicting suitable linkers for fusion protein using Fuzzy Logic

Input : Protein Sequence 1 (PS1), Protein Sequence 2 (PS2), Linkers Dataset (LD), Protein Sequences Dataset (PSD)
Output : Suitable Linkers
Step 1 Set PS1flcount=0, PS2flcount=0, PS1rlcount=0, PS2rlcount=0, PS1clcount=0, PS2clcount=0
Step 2 : For each codon C from codons for Flexible Linker **// Table 2**
Step 3 : If PS1 contains C, Then
Step 4 : PS1flcount++
Step 5 : If PS2 contains C, Then
Step 6 : PS2flcount++
Step 7 : End For
Step 8 : For each codon C from codons for Rigid Linker **// Table 3**
Step 9 : If PS1 contains C, Then
Step 10 : PS1rlcount++
Step 11 : If PS2 contains C, Then
Step 12 : PS2rlcount++
Step 13 : End For
Step 14 : For each codon C from codons for Cleavable Linker **// Table 4**
Step 15 : If PS1 contains C, Then
Step 16 : PS1clcount++
Step 17 : If PS2 contains C, Then
Step 18 : PS2clcount++
Step 19 : End For
Step 20 : linkerType = "" **// Predict Suitable Linker Type**
Step 21 : If (PS1flcount > PS1rlcount) && (PS2flcount > PS2rlcount) Then
Step 22 : linkerType="Flexible"
Step 23 : Else If (PS1flcount < PS1rlcount) && (PS2flcount < PS2rlcount) Then
Step 24 : linkerType="Rigid"
Step 25 : Else If (PS1clcount != 0) && (PS2clcount != 0) Then
Step 26 : linkerType="Cleavable"
Step 27 : candidateLinkers = { }, i = 0
Step 28 : For each Linker L from LD
Step 29 : If linkerType is equal to "Flexible", Then
Step 30 : If L contains "GGGGS", Then
Step 31 : candidateLinkers[i] = L

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

```
Step 32 :                 i++
Step 33 :         Else If linkerType is equal to “Rigid”, Then
Step 34 :             If L contains “EAAAK”, Then
Step 35 :                 candidateLinkers[i] = L
Step 36 :                 i++
Step 37 :         Else If linkerType is equal to “Cleavable”, Then
Step 38 :             If L contains “TGT”, Then
Step 39 :                 candidateLinkers[i] = L
Step 40 :                 i++
Step 41 : End For
Step 42 : Min with Max = Find Minimum and Maximum protein sequences lengths
          using MinMaxNormalization from PSD
Step 43 : X = Length of PS1 + Length of PS2
Step 44 : MMN = (X – Min) / (Max - Min)
Step 45 : Min = 0, Max = 0, i = 0
Step 46 : For each Linker L from candidateLinkers
Step 47 :     If i == 0, then
Step 48 :         Min = Length of L
Step 49 :         Max = Length of L
Step 50 :     Else If (Length of L < Min), Then
Step 51 :         Min = Length of L
Step 52 :     Else If (Length of L > Max), Then
Step 53 :         Max = Length of L
Step 54 : End For
Step 55 : suitableLinkers = { }, i = 0
Step 56 : For each Linker L from candidateLinkers
Step 57 :     X = Length of L
Step 58 :     MMNi = (X – Min) / (Max - Min)
Step 59 :     If(((MMN - 0.1) > MMNi) || ( MMNi >= (MMN + 0.1))) Then
Step 60 :         suitableLinkers[i] = L
Step 61 :         i++
Step 62 : End For
```

3. RESULTS & DISCUSSIONS

The successful construction of a fusion protein requires component proteins and the suitable linkers. The selection of the component proteins is primarily based on the desired functions of the fusion protein and selection of a proper linker to join the chosen proteins. Generally predicting the suitable linker is a complicated task. Therefore, prediction of the suitable linker for the construction of the fusion protein is important and it has a significant effect on the structure of a fusion protein. This section provides the results attained by simulate suitable linker prediction for fuse various types of proteins using soft computing methods, especially Fuzzy Logic. Java is utilized for simulation to appropriate linker prediction using Fuzzy Logic.

Example: Consider the following protein sequences
Protein Sequence of 100D: CCGGCGCCGG
Protein Sequence of 103D: GTGGAATGGAAC
Protein Sequence of 102D: CGCAAATTTGCG

Case Study 1:
Protein Structure Id 1: 100D
Protein Structure Id 2: 103D
Figure 3.1 shows the predicted suitable linkers for fuse proteins 100D and 103D.

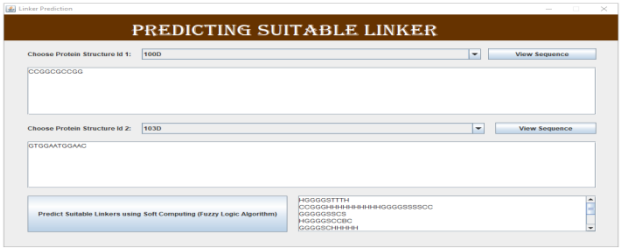

Figure 3.1: Predicted suitable linkers for fuse proteins 100D and 103D

Predicted Linker Type: Flexible

Candidate Linkers:
HGGGGSTTTH
CCGGGHHHHHHHHHHGGGGSSSSCC
GGGGGSSCS
HGGGGSCCBC
GGGGSCHHHHH
GGGGS
GGGGSGGGGS
GGGGSGGGGSGGGGS
GGGGSGGGGSGGGGSGGGGS
GGGGSGGGGSGGGGSGGGGSGGGGS

Suitable Linkers:
HGGGGSTTTH
CCGGGHHHHHHHHHHGGGGSSSSCC
GGGGGSSCS
HGGGGSCCBC
GGGGSCHHHHH
GGGGSGGGGS

Case Study 2:
To fuse two protein structures,

Protein Structure Id 1: 100D
Protein Structure Id 2: 102D

Protein Sequence of 100D: CCGGCGCCGG
Protein Sequence of 102D: CGCAAATTTGCG

Figure 3.2 shows the predicted suitable linkers for fuse proteins 100D and 102D.

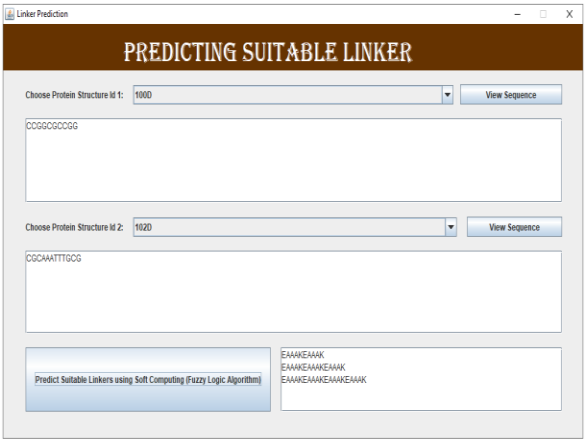

Figure 3.2: Predicted suitable linkers for fuse proteins 100D and 102D

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

Predicted Linker Type: Rigid

Candidate Linkers:
EAAAK
EAAAKEAAAK
EAAAKEAAAKEAAAK
EAAAKEAAAKEAAAKEAAAK
EAAAKEAAAKEAAAKEAAAKEAAAK

Suitable Linkers:
EAAAKEAAAK
EAAAKEAAAKEAAAK
EAAAKEAAAKEAAAKEAAAK

4. SUMMARY:

This paper proposed soft computing based suitable linker prediction technique for generating fusion protein. Soft computing provides several possibilities by developing good low-cost solutions. However, the prediction of a suitable linker for fusion protein is costly because of the prices related to crystallography; electron microscopy also takes more time. In this situation, soft computing presents numerous feasibilities by constructing inexpensive, high-quality results—soft computing methods for the domain-linker forecast. This paper uses the Fuzzy Logic technique for soft computing to forecast suitable linkers for given protein sequences. Extensive experimental results indicate that the proposed fuzzy logic technique predicts the suitable linkers for given protein sequences efficiently.

References:

[1] Blaskovich, M. A. (2016). Unusual amino acids in medicinal chemistry. Journal of medicinal chemistry, 59(24), 10807-10836.

[2] Bornscheuer, U. T., Hauer, B., Jaeger, K. E., & Schwaneberg, U. (2019). Directed evolution empowered the redesign of natural proteins for the sustainable production of chemicals and pharmaceuticals. Angewandte Chemie International Edition, 58(1), 36-40.

[3] Khan, F. A. (2018). Recombinant DNA technology. In Biotechnology Fundamentals (pp. 125-177). CRC Press.

[4] Strohl, W. R. (2015). Fusion proteins for half-life extension of biologics as a strategy to make biobetters. BioDrugs, 29(4), 215-239.

[5] Jafari, R., M Zolbanin, N., Rafatpanah, H., Majidi, J., & Kazemi, T. (2017). Fc-fusion proteins in therapy: an updated view. Current medicinal chemistry, 24(12), 1228-1237.

[6] Krishnamoorthy, S., Liu, T., Drager, D., Patarroyo-White, S., Chhabra, E. S., Peters, R., ... & Jiang, H. (2016). Recombinant factor VIII Fc (rFVIIIFc) fusion protein reduces immunogenicity and induces tolerance in haemophilia A mice. Cellular immunology, 301, 30-39.

[7] Chen, S., & Hall, E. A. (2020). A biosilification fusion protein for a 'self-immobilising'sarcosine oxidase amperometric enzyme biosensor. Electroanalysis, 32(4), 874-884.

[8] Lagassé, H. D., Alexaki, A., Simhadri, V. L., Katagiri, N. H., Jankowski, W., Sauna, Z. E., & Kimchi-Sarfaty, C. (2017). Recent advances in (therapeutic protein) drug development. F1000Research, 6.

[9] Xi, Y., Yu, D., Yang, R., Zhao, Q., Wang, J., Zhang, H., ... & Gong, D. W. (2019). Recombinant Fc-Elabela fusion protein has extended plasma half-life and mitigates post-infarct heart dysfunction in rats. International journal of cardiology, 292, 180-187.

[10] Abe, S., & Ueno, T. (2015). Design of protein crystals in the development of solid biomaterials. RSC Advances, 5(27), 21366-21375.

[11] Xing, S., Wallmeroth, N., Berendzen, K. W., & Grefen, C. (2016). Techniques for the analysis of protein-protein interactions in vivo. Plant Physiology, 171(2), 727-758.

[12] Zhou, M., Li, Q., & Wang, R. (2016). Current experimental methods for characterizing protein-protein interactions. ChemMedChem, 11(8), 738.

[13] Bersuker, K., Brandeis, M., & Kopito, R. R. (2016). Protein misfolding specifies recruitment to cytoplasmic inclusion bodies. Journal of Cell Biology, 213(2), 229-241.

[14] Davis, C. M., Deutsch, J., & Gruebele, M. (2020). An in vitro mimic of in-cell solvation for protein folding studies. Protein Science, 29(4), 1046-1054.

[15] Xu, C., & Ng, D. T. (2015). Glycosylation-directed quality control of protein folding. Nature reviews Molecular cell biology, 16(12), 742-752.

[16] Elleuche, S. (2015). Bringing functions together with fusion enzymes—from nature's inventions to biotechnological applications. Applied microbiology and biotechnology, 99(4), 1545-1556.

[17] Huang, Z., Zhang, C., & Xing, X. H. (2021). Design and construction of chimeric linker library with controllable flexibilities for precision protein engineering. Methods Enzymol, 647, 23-49.

[18] Kamalirousta, M., & Pilehchian, R. (2018). A new bifunctional fusion protein of vaccine strains Clostridium perfringens type D and Clostridium septicum epsilon-alpha toxin genes. Microbiology Research Journal International, 1-9.

[19] Tsai, C. J., Aslam, K., Drendel, H. M., Asiago, J. M., Goode, K. M., Paul, L. N., ... & Hazbun, T. R. (2015). Hsp31 is a stress response chaperone that intervenes in the protein misfolding process. Journal of Biological Chemistry, 290(41), 24816-24834.

[20] Takeuchi, T., & Nagai, Y. (2017). Protein misfolding and aggregation as a therapeutic target for polyglutamine diseases. Brain sciences, 7(10), 128.

[21] George RA. and Heringa J. (2002) An analysis of protein domain linkers: their classification and role

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

in protein folding. Protein Engineering 15, 871-879.
[22] Liu, C., Chin, J. X., & Lee, D. Y. (2015). SynLinker: an integrated system for designing linkers and synthetic fusion proteins. Bioinformatics, 31(22), 3700-3702.
[23] Yu, K., Liu, C., Kim, B. G., & Lee, D. Y. (2015). Synthetic fusion protein design and applications. Biotechnology advances, 33(1), 155-164.
[24] Shamriz, S., Ofoghi, H., & Moazami, N. (2016). Effect of linker length and residues on the structure and stability of a fusion protein with malaria vaccine application. Computers in biology and medicine, 76, 24-29.
[25] Li, G., Huang, Z., Zhang, C., Dong, B. J., Guo, R. H., Yue, H. W., ... & Xing, X. H. (2016). Construction of a linker library with widely controllable flexibility for fusion protein design. Applied microbiology and biotechnology, 100(1), 215-225.
[26] Shen, X., Li, Q., Wang, F., Bao, J., Dai, M., Zheng, H., & Lao, X. (2018). Generation of a novel long-acting thymosin alpha1-Fc fusion protein and its efficacy in inhibiting breast cancer in vivo. Biomedicine & Pharmacotherapy, 108, 610-617.
[27] Yang, J., Sun, Y., Bao, R., Zhou, D., Yang, Y., Cao, Y., ... & Zhong, M. (2017). The second-generation flagellin-rPAc fusion protein, KFD2-rPAc, shows high protective efficacy against dental caries with low potential side effects. Scientific reports, 7(1), 1-13.
[28] Ameghi, A., Baradaran, B., Aghaiypour, K., Barzegar, A., Pilehvar-Soltanahmadi, Y., Moghadampour, M., ... & Zarghami, N. (2015). Generation of new M2e-HA2 fusion chimeric peptide to development of a recombinant fusion protein vaccine. Advanced pharmaceutical bulletin, 5(Suppl 1), 673.
[29] Sper, R. B., Koh, S., Zhang, X., Simpson, S., Collins, B., Sommer, J., ... & Piedrahita, J. A. (2017). Generation of a stable transgenic swine model expressing a porcine histone 2B-eGFP fusion protein for cell tracking and chromosome dynamics studies. PloS one, 12(1), e0169242.
[30] Niesler, N., Arndt, J., Silberreis, K., & Fuchs, H. (2020). Generation of a soluble and stable apoptin-EGF fusion protein, a targeted viral protein applicable for tumour therapy. Protein Expression and Purification, 175, 105687.
[31] Ongaro, T., Matasci, M., Cazzamalli, S., Gouyou, B., De Luca, R., Neri, D., & Villa, A. (2019). A novel anti-cancer L19-interleukin-12 fusion protein with an optimized peptide linker efficiently localizes in vivo at the site of tumours. Journal of biotechnology, 291, 17-25.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731