International Research Journal of Biochemistry and Bioinformatics Vol. 13(3) pp. 1-3, June, 2023
Available online http://www.interesjournals.org/IRJBB

Review Article

Unnatural Amino Acids into Proteins/ Protein Engineering

Hayley James*

The Skaggs Institute for Chemical Biology, United States

*Corresponding Author's E-mail: jameshayley@rediff.com

Received: 02-Jun-2023, Manuscript No. IRJBB-23-103031; Editor assigned: 05-Jun-2023, PreQC No. IRJBB-23-103031 (PQ); Reviewed: 19-Jun-2023, QC No. IRJBB-23-103031; Revised: 23-Jun-2023, Manuscript No. IRJBB-23-103031 (R); Published: 30-Jun-2023, DOI: 10.14303/2250-9941.2023.60

Abstract

An orthogonal tRNA/aminoacyl-tRNA synthetase pair may genetically integrate unnatural amino acids into proteins in living cells. Here, we report a technique to effectively produce the orthogonal tRNA and synthetase in *Saccharomyces cerevisiae*, allowing for highly effective genetic incorporation of non-natural amino acids into target proteins in yeast. We also discuss the use of a yeast strain that is defective in nonsense-mediated mRNA decay, which improves the efficiency with which an unnatural amino acid is incorporated when a stop codon is used to encode it. By using the unique features provided by synthetic amino acids, these methods will make it easier to study proteins and the biological processes that are associated to them in yeast (Elmlinger et al., 2002).

Keywords: Amino Acids, Proteins therapeutics, Protein engineering

INTRODUCTION

The majority of protein biosensors use chemical or genetic labelling of the protein, which introduces an extraneous molecule to the original, wild-type protein and frequently modifies its overall structure and characteristics. Although these labelling techniques have been effective in many instances, they also have a number of drawbacks related to how they are made and used. The addition of unnatural amino acid (UAA) analogues that can serve as labels to a protein's structure is another method for labelling proteins. While altering the immediate microenvironment, such a strategy is less taxing on the protein's overall structure. L-DOPA is a phenylalanine analogue with a catechol group that engages in a two-electron redox process that is almost reversible, making it a good candidate for use as an electrochemical label or reporter. To illustrate this detection idea, the periplasmic glucose/galactose binding protein (GBP) was used. GBP experiences a considerable conformational shift upon binding glucose, which is reflected in a change in the electrochemistry of L-DOPA. In order to measure glucose levels directly, the electroactive GBP was immobilised onto carbon electrodes that were screen-printed with caffeic acid that had been modified with gold nanoparticles (GBP-LDOPA/AuNP/PCA/SPCE). This serves as a proof-of-concept for the use of electrochemically active synthetic amino acids as the label (Soldin et al., 2005). The resulting reagent-free GBP biosensors demonstrated a highly specific and sensitive binding affinity for glucose in the micromolar range, providing the basis for a novel biosensing approach based on global incorporation of an electroactive amino acid into the protein's primary sequence for highly specific electrochemical detection of compounds of interest. The production of structurally specified protein conjugates is made possible by the site-specific integration of artificial amino acids with orthogonal chemical reactivity into proteins. Genetically encoded ketone, azide, alkyne, alkene, and tetrazine-containing amino acids can react to nonsense and frameshift codons. The physical and biological features of protein conjugates, particularly those of the next-generation protein therapies, have been optimised in a manner similar to medicinal chemistry thanks to these bio-orthogonal chemical handles, which provide precise control over the location and stoichiometry of conjugation (Owen et al., 2010).

Protein site-specific modification techniques have developed into potent tools for examining protein structure and function as well as for producing proteins with improved or novel features. In the past, electrophilic substances have

been used to modify lysine, cysteine, or nucleophilic serine side chains in protein conjugation processes. However, especially when altering big proteins like antibodies, these reactions typically result in heterogeneous mixtures of protein conjugates with unique characteristics. On the other hand, precise control of the conjugation site and stoichiometry has been made possible by the insertion of chemically orthogonal functional groups into proteins. For this reason, semi-synthetic techniques like expressed protein ligation (EPL) and enzymatic or chemical processes that change certain peptide sequences with chemical tags are frequently used (Konforte et al., 2013).

Unnatural amino acids (UAAs) have also been incorporated into proteins using recombinant techniques as a chemical handle for bio-orthogonal conjugation processes. The latter strategy is particularly appealing because, in theory, the UAA can be incorporated at any desired position in any protein, the structure of the wild-type protein is only slightly altered by the incorporation of the UAA because the UAA site is not specified by any specific sequence context, and site-specifically modified proteins are expressed recombinantly in high yields in bacteria, yeast, or mammalian cells and require little additional manipulation for fusing. Several approaches for incorporating UAAs with orthogonal chemical reactivity into proteins. Since its inception, protein engineering has had a significant influence on a variety of sectors and has emerged as a key method for chemical alterations, functional research, and biophysical evaluation. Since the creation of the many methods generally referred to as "directed evolution," protein engineering has become increasingly widely used in the enhancement and manipulation of protein function for medicinal and industrial reasons. Our need and desire to comprehend and utilise the power of proteins for our own goals has spurred these advancements (Yang et al., 2005). The capacity to site-specifically insert unnatural (i.e., not found among the canonical amino acids) amino acids (uAAs) into proteins is an exciting recent advancement in protein engineering. If the translational machinery has been changed to incorporate an appropriate orthogonal tRNA/aminoacyl-tRNA synthetase pair, this approach uses suppressor tRNA to enable integration of uAAs by reading over a stop codon. Conveniently, the method produces entire incorporation yields because truncated proteins that are either intractable or simple to separate from the full-length product come from competing recognition of the stop codons as termination signals (Davis et al., 2006).

By nonsense codon suppression with the designed tRNA and aminoacyl-tRNA synthetase (aaRS) pairings, a potent method created by Schultz et al. in 2001 for the study of protein structures and functions, it is possible to achieve site-specific incorporation of unnatural amino acids (UAAs) into proteins. More than 150 UAAs have been integrated into different proteins in bacterial, yeast, and mammalian cells over the past 20 years. More recently, this method has been used to transgenic mice, fruit flies, zebrafish, worms, and fruit flies. Since there are 20 canonical or natural amino acid (NAA) residues in most proteins, the NAA UAA mutation, also known as the substitution of one NAA residue with another, can give the protein new properties or altered functions, such as photo-crosslinking, biorthogonal labelling, site-specific conjugation, or enhanced enzymatic activities (Carel et al., 2009). The reported UAA-incorporated proteins could be an excellent database for research on the guidelines for UAA replacements. The evolutionarily tolerated protein sequence, the steric effects of the protein structure and the physiochemical changes brought on by the NAA UAA mutation are among the many variables that are generally thought to affect the effectiveness or outcome of UAA substitutions. Some conserved regions that are essential for how proteins work are often less tolerant of UAA changes. The protein structure is least affected when UAA substitutions are made at surface-exposed residues, and subsequent labelling is also more convenient. The condition of the substitutions relies on the location and varies for various proteins and residues. Additionally, there may be physiochemical variations between UAAs and NAAs in terms of polarity, hydrophobicity, and hydrogen bonding. Researchers seem to concur that UAAs with similar physiochemical qualities to NAAs are more likely to perform a successful substitute (Zec et al., 2012). Therefore, the UAA replacement is a multivariate process that is still difficult to forecast. One UAA's site tolerability may be predicted using machine learning, according to a recent study. For informative predictions or virtual screening of UAA-incorporated proteins, a trustworthy model or programme combining all pertinent elements is still required for the general case of any UAA replacing any position on any protein.

We began by reorganising the reported UAA-incorporated proteins into a database in order to get insight into the UAA substitution process. We identified more than 500 items by searching PubMed for publications with "unnatural/ non-canonical amino acid incorporation/substitution/ replacement" in the title or abstract. Additionally, 196 research publications that provide specific details about the target protein, the replacement site, the replaced NAA, the substituted UAA, and the result of the substitution were taken into account. Several inclusion and exclusion criteria were used during data collection, as described in the Methods section, to enhance the quality and creditability of the database. For structure predictions, the target protein, for instance, should preferably have a coordinate structure file in the Protein Data Bank (PDB) or at least a complete sequence. Manual confirmation of the protein structure or sequence revealed both the UAA replacement site and the original NAA there. UAA substitutions on linkers, peptides, or proteins with fewer than 50 residues were disregarded since these structures are often loose or have a limited number of stably interacting domains that can be used to explore steric effects. The success or failure of the UAA replacement was assessed using several experimental techniques that were presented, and the results were then categorised

as direct, indirect, or other related proofs (Chan et al., 2009). Mass spectrometry of UAA-incorporated proteins or important peptide fragments, gels or blots demonstrating the expression or absence of UAA-incorporated proteins, or several well-known optical reporters are examples of direct demonstrations. Various outcomes, such as maintaining protein functions or enzymatic activities, or the virus package of UAA-incorporated viral proteins, are examples of indirect demonstrations. UAA incorporations utilising the ligated UAA-dCA-tRNA approach or UAA replacement using auxotrophic strains are two further examples of similar evidence. The research should show that the UAA/aaRS/tRNA system is functioning elsewhere and that the failure was mostly caused by incompatible UAAs at the current sites rather than by a flawed system if failure replacements are to be successful. To supplement the results of the UAA incorporation, the UAA incorporation efficiency relative to wild-type was determined where available, and an efficiency of >0.01 was deemed successful (Elmlinger et al., 2005).

CONCLUSION

This research developed a trustworthy prediction model to define the laws of UAA substitutions by methodically compiling the reported UAA-incorporated proteins during the previous two decades into a database for machine learning. The RPDUAA programme recorded 172 UAAs in its initial release and enabled direct analysis of protein information in cif, fasta, and xml files. Users can conveniently use the RPDUAA programme to predict the likelihood of successful UAA substitutions, add their own target proteins or UAAs to the libraries, add new substitution records to the database, test the performance of the prediction model with the entire database or one of its subsets, and optimise their UAA-incorporated proteins. For instance, based on the anticipated likelihood and the whole heatmap, users can determine the best sites on a particular protein for a certain UAA to substitute (the fixed UAA method) or the best UAAs that may do so (the fixed site strategy). Prior research mostly relied on mechanical screening with time-consuming procedures to choose UAAs and replacement locations. The RPDUAA programme, which is a valuable tool for the rational design of UAA-incorporated proteins, might automate the choices based on artificial intelligence and virtual screening. The RPDUAA programme demonstrated a reliable prediction performance, as shown by the time split validations or experimental validations. Users can get the high-confidence candidates of UAA-incorporated proteins for further research by choosing a probability threshold at the appropriate cut off of the selected subset (~ 0.49 for balanced subsets, ~ 0.84 for the whole database or other subsets), or higher (such as 0.90 or 0.95).

REFERENCES

1. Elmlinger MW, Kuhnel W, Ranke MB (2002). Reference Ranges for Serum Concentrations of Lutropin (LH), Follitropin (FSH), Estradiol (E2), Prolactin, Progesterone, Sex Hormone-Binding Globulin (SHBG), Dehydroepiandrosterone Sulfate (DHEAS), Cortisol and Ferritin in Neonates, Children and Young Adults. Clin Chem Lab Med. 40: 1151-1160.
2. Soldin OP, Hoffman EG, Waring MA, Soldin SJ (2005). Pediatric reference intervals for FSH, LH, estradiol, T3, free T3, cortisol, and growth hormone on the DPC IMMULITE 1000. Clin Chim Acta. 355: 205-210.
3. Owen WE, Rawlins ML, Roberts WL (2010). Selected performance characteristics of the Roche Elecsys testosterone II assay on the Modular analytics E 170 analyzer. Clin Chim Acta. 411: 1073 -1079.
4. Konforte D, Shea JL, Kyriakopoulou L, Colantonio D, Cohen AH, et al (2013). Complex biological pattern of fertility hormones in children and adolescents: a study of healthy children from the CALIPER cohort and establishment of pediatric reference intervals. Clin Chem. 59: 1215-1227.
5. Yang L, Grey VL (2005). Pediatric reference intervals for bone markers. Clin Biochem. 39: 561-568.
6. Davis GK, Bamforth F, Sarpal A, Dicke F, Raby Y, et al (2006). B-type natriuretic peptide in pediatrics. Clin Biochem. 39: 600-605.
7. Carel JC, Eugster EA, Rogol A, Ghizzoni L, Palmert MR (2009). Consensus statement on the use of gonadotropin-releasing hormone analogs in children. Pediatrics.123: 752-762.
8. Zec I, Kučak I, Begčević I, Šimundić A, Tišlarić-Medenjak D (2012). Reference intervals for reproductive hormones in prepubertal children on the automated Roche cobas e 411 analyzer. Clinical Biochemistry. 45: 1206-1212.
9. Chan MK, Seiden-Long I, Aytekin M, Quinn F, Ravalico T, et al (2009). Canadian Laboratory Initiative on Pediatric Reference Interval Database (CALIPER): pediatric reference intervals for an integrated clinical chemistry and immunoassay analyzer, Abbott ARCHITECT ci8200. Clinical Biochemistry. 42: 885-891.
10. Elmlinger MW, Kuhnel W, Ranke MB (2005). Reference Ranges for Serum Concentrations of Lutropin (LH), Follitropin (FSH), Estradiol (E2), Prolactin, Progesterone, Sex Hormone-Binding Globulin (SHBG), Dehydroepiandrosterone Sulfate (DHEAS), Cortisol and Ferritin in Neonates, Children and Young Adults. Clin Chem Lab Med. 40: 1151-1160.