Awards issued prior to January 20, 2025, were funded under previous Administrations and may not reflect the priorities and policies of the current Administration.

United States Department of Agriculture
Project Initiation

Project Data for Accession Number 1032168

Project Title

Using Artificial Intelligence and Robust Protein Design Methods for Highly Efficient Production of Milk Proteins via Bacterial Fermentation

Project Details

Sponsoring Institution National Institute of Food and Agriculture	**Project Status** ACTIVE
Funding Source NIFA Non Formula	**Grant Year** 2024
Grant No. 2024-67017-42678	**Cumulative Award Amt.** $300,000.00
Proposal No. 2023-10521	**Multistate No.** (N/A)
Project Start Date Jul 15, 2024	**Project End Date** Jul 14, 2027
Program Code [A1364] Innovations in Food Manufacturing Technologies	

Project Director

Love, J.

Recipient Organization

SAN DIEGO STATE UNIVERSITY FOUNDATION

5250 CAMPANILE DR

SAN DIEGO, CALIFORNIA 92182

Performing Department

(N/A)

Non Technical Summary

Global weather patterns have been changing unpredictably as the average planet surface temperature has been steadily increasing over the last 50 to 60 years. Global warming is increasing the intensity of dangerous weather events, such as tornadoes and hurricanes and inducing climate change, which has resulted in more floods and droughts. Our research functions to reduce the greenhouse gas (GHG) emissions associated with dairy farming. Sources of GHG emissions on dairy farms include diesel-powered machinery, methane emissions from cow digestive systems, nitrous oxide emissions from fertilized soils, as well as the significant energy used for mass production of fertilizers. We believe it is crucial that everything possible should be done in order to reduce our collective carbon footprint.We are not proposing the elimination of dairy farming, which has historically provided high-quality nutrition to generations of Americans. What we are proposing is to engineer an alternative method of producing certain milk

proteins, which can be utilized to produce different types of cheeses and other nutritious food products. Our method has a lower carbon footprint as it entails the use of precision bacterial fermentation to produce milk casein proteins. For much of human history, we have been perfecting the use of microorganism fermentation, which utilizes beneficial bacteria and yeast to produce many foods such as cheeses, yogurts, beer, wine, and many types of breads. Our research will result in enhanced bacterial fermentation by combining artificial intelligence with robust biotechnology to produce nutritious milk proteins more sustainably and with a lower carbon footprint.

Animal Health Component

Animal Health Component: 0%

Research Effort Categories

Basic	80%
Applied	20%
Developmental	0%

Classification

Knowledge Area (KA)	Subject of Investigation (SOI)	Field of Science (FOS)	Percent
502	5010	1000	100%

Knowledge Area

[502] New and Improved Food Products

Subject Of Investigation

[5010] Food (not readily associated with specific plant and animal products)

Field Of Science

[1000] Biochemistry and biophysics

Keywords

Bacterial fermentation, Cheese production, Milk protein (caseins), Protein Design

Goals / Objectives

Overarching Research Goal: The overarching goal of this research project is to engineer a rapid, effective, and sustainable protein expression/purification system used to produce nutritious milk proteins (caseins) via precision bacterial fermentation. The system will enable casein protein production with a lower carbon footprint relative to producing milk proteins via traditional dairy farming.Immediate Research Goal: The immediate research goal is to engineer increased, yet transient solubility/stability into the ?S1-casein milk protein. To achieve this, small stable proteins, referred to as transient tertiary structure (TTS) tags, will be expressed as N-terminal and C-terminal fusions with the ?S1-casein milk protein. The goal of this form of protein engineering is to impart greater solubility on the ?S1-casein and thus render this protein more straightforward to produce, isolate, and purify. Following high-level expression, the TTS tags will be catalytically removed from the casein proteins. We will continue to use the AI

program AlphaFold, in combination with rational design, to generate fusion variants of ?S1-casein using designed G?1 TTS tags. Relying on this approach, we have already generated preliminary data indicating that this research approach is highly feasible.Objective 1:Produce ?S1-casein fusions with increased solubility and higher levels of bacterial expression.Thus far we have engineered one variant with a wild-type G?1 TTS tag expressed as a fusion to the N-terminus of ?S1-casein. This variant exhibited favorable solubility properties relative to the wild-type protein and therefore we intend to express WT G?1 TTS tags as fusions to both the N- and C-terminus of ?S1-casein. Increases in solubility of the associated inclusion bodies for these variants will be analyzed using standard solubility tests that include but are not limited to variable centrifugation/gel electrophoresis, circular dichroism, and NMR.Milestone: Generate 2 to 3 ?S1-casein-TTS tag fusion variants that are straightforward to extract from inclusion bodies and have a good level of solubility in the context of the fused protein construct.Timeline: Protein expression and testing, Months 1 - 4.Objective 2: Continue to use the AI program AlphaFold to engineer variants of the G?1 TTS tags with the goal of increasing the solubility and expression levels of ?S1-casein. Rational design, based on human intuition and molecular visualization (Pymol), results in a series of candidate mutant proteins that are rapidly 'filtered' and critically assessed using the program AlphaFold. The resulting mutant G?1 TTS tags, which provide greater transient tertiary interactions, will be expressed as fusions with ?S1-casein.Milestone: Generate 2 to 3 engineered ?S1-casein-TTS tag fusions using the described AlphaFold-based methods.Timeline: AlphaFold design: Months 1-2. Gene design, protein expression, and assessment: Months 2 - 6.Objective 3: Design additional TTS tags based on the G?1 variants that have been previously engineered to dimerize upon binding metals. These G?1 variants dimerize upon binding divalent transition metals and are referred to as Metal-controlled dimer (MCD) proteins. The MCDs afford unprecedented chemical and temporal control over dimerization. The MCDs will be expressed as N- and C-terminal fusions of ?S1-casein for enhanced solubilization from inclusion bodies during purification. Upon expression, the divalent transition metal Zn2+ will be added to bacterial cultures to drive MCD dimerization and reduce strong intermolecular interactions that cause insoluble inclusion bodies. During inclusion body workup, the metal chelating chemical EDTA will be added to the inclusion body solution to dissociate the dimers and enhance ?S1-casein fusion folding and solubility.Milestone: Express MCD TTS tags as N- and C-terminus fusions with ?S1-casein and test the addition of different metal concentrations to enhance inclusion body purification.Timeline: Gene design, protein expression, and assessment: Months 5 - 8.Objective 4: Cleave the fusion proteins produced from Objectives 1 thru 3with the TEV protease and induce precipitationin a manner similar to cheese curds. The precipitated curds will be formulated into different types of cheese products by varying ingredient components such as water, fat, salt, lactic acid, stabilizers, and flavors. The rheological properties (elastic modulus, viscous modulus, complex modulus, etc.), textural properties (firmness, stickiness, meltability, stretchability, etc.), color and browning, microscopic structures, and volatile profiles of the cheese products will be characterized and compared to dairy cow produced ?S1-casein.Milestone: Cleave successful variants from Objectives 1 thru 3 and analyze the textural and structural properties of the produced material.Timeline: Cleavage and cheese curd quality assessment, Months 7 - 12.Objective 5: The fusion proteins produced from Objectives 1 thru 3 will be tested for phosphorylation by the casein Kinase I and 2 enzymes displayed on the surface of E. coli using the Bacterial Surface Display (BSD) system. Catalysis will be tested in the context of the G?1-?S1-casein fusions both before and after cleavage with the TEV protease. Functional and nutritional properties of cheese made with the phosphorylated casein will be evaluated as described in Objective 4.Milestone: Subclone the gene for casein Kinase I and 2 enzymes into the BSD system. Phosphorylate ?S1-casein and test for degree and pattern of phosphorylation. Compare degree of phosphorylation to verify similarity to the same protein from standard milk. Test functional properties of the phosphorylated casein.Timeline: Subcloning of the gene for casein Kinase I enzyme into the BSD system, test degree and pattern of phosphorylation, test casein functional properties. Months 12 - 20. Objective 6:This objectiveis a proof-of-concept study that aims to develop a novel expression system for robust surface expression (and possibly secretion) of target proteins in a methane-consuming host (the 20ZR methanotrophic bacterial strain) and demonstrate its applicability for production of G?1-?S1-casein variants. We will focus on surface (S)-layer proteins as a target fusion protein for surface level expression of G?1-?S1-casein variants.Milestone: Subclone the gene(s) for select G?1-?S1-casein variants into 20ZR engineered plasmids so they are expressed as S-layer protein fusions, or possibly as secretion proteins.Timeline: Subcloning of the gene for G?1-?S1-casein variants, workup and assessment, Months 16 - 24.

United States Department of Agriculture
Project Initiation

Project Methods

The research methods employed to sustainably produce casein milk proteins through precise bacterial fermentation will follow a systematic approach based on standard scientific methodology, with unique aspects tailored to the specific objectives of this project:Experimental Design and Protein Engineering: The project will begin with the design of casein fusion proteins engineered for increased solubility/stability. This will involve the engineering of transient tertiary structure (TTS) tags designed to increase the solubility and bacterial expression of milk casein proteins. Unique to this project is the precision required for effective or protein engineering to ensure high level, consistent, and sustainable production of the targeted casein milk proteins.Protein Production Optimization: Once the casein fusion proteins have been successfully engineered, protein production will be optimized to maximize yield and purity. This will involve controlling factors such as solubility, temperature, pH, nutrient availability, and oxygen levels.Protein Purification and Characterization: Following protein expression, the produced proteins will be subjected to purification to remove impurities and isolate the target casein milk proteins. Various chromatographic techniques will be tested and employed for purification. The purified proteins will be characterized using analytical methods such as SDS-PAGE, circular dichroism, solubility tests, NMR, and other spectroscopic techniques used to confirm their increased solubility and assess their characteristics relative to milk proteins obtained from dairy cows. Assessment will entail characterizing rheological properties (elastic modulus, viscous modulus, complex modulus, etc.), textural properties (firmness, stickiness, meltability, stretchability, etc.), color and browning, microscopic structures, and volatile profiles of the cheese products.Data Analysis and Interpretation: The results obtained following protein production and characterization will be analyzed statistically to identify optimal conditions for increased production and high purity. Deviations from expected outcomes will be carefully examined and interpreted to refine experimental procedures and enhance protein yield and quality.Knowledge Dissemination and Outreach: Efforts will be made to disseminate the findings of the research to both scientific and non-scientific audiences. This will include publishing research articles in relevant journals, presenting findings at conferences and seminars, and potentially engaging with industry stakeholders.Impact Evaluation: The scientific impact of this project will be assessed through metrics such as publications, presentations, collaborations, and potential adoption of the developed methods by the scientific community. Additionally, the project's potential to contribute to sustainable food production practices and reduce reliance on traditional dairy farming will be evaluated.By following this systematic approach, the project aims to advance scientific knowledge in the field of protein design, while also contributing to the development of sustainable and innovative solutions for casein milk protein production.The scientific methods employed to achieve the research goals of this project include standard biochemical and molecular biology methods and technologies. These include, but are not limited to the following:Recombinant DNA technologiesBiochemical techniques and methodologies including buffer preparation, pH titration, bacterial media preparation, etc.Autoclaving of bacterial media and agar plates for antibiotic selection.Buffer preparations for optimal bacterial growth and protein production.PCR based sub-cloning, DNA amplification, and mutagenesis methods.Bacterial transformation with recombinant DNA plasmids.IPTG induced bacterial protein expression.T7-based bacterial protein expression and purification.Bacterial cell lysis via sonication or freeze/thaw methodsPAGE and agar gel electrophoresisHPLC and FPLC column chromatographyStructure-based Rational Protein DesignComputational chemistry methods and programs including the molecular visualization program Pymol, the AI-based program AlphaFold, and Molecular Dynamics calculations run using the programs Amber and GROMACS.