Science and Technology

Recent Trends

Volume - II

Editors

- Jomon Thommana
- Yvonne Fernandes
- Shailesh Thakrar
- Priyank Patel
- Ashish Prasad

Science and Technology Recent Trends

Editors

Jomon Thommana

Yvonne Fernandes

Shailesh Thakrar

Priyank Patel

Ashish Prasad

CHRIST COLLEGE RAJKOT

First Published: 2026

Copyright@ Christ College, Rajkot, Gujarat, India

16th National Level Science Symposium on "Science and Technology: Recent Trends"

ISBN : 978-81-970732-2-9

No part of this book may be reproduced or transmitted in any form by any means, electronic or mechanic, including photocopy, recording or any other information storage and retrieval system, without permission in writing from the copyright owners.

DISCLAIMAER

The authors are solely responsible for the contents of the papers complied in this volume. The publishers or editors do not take any responsibility of any issue with regard to the content of the papers.

Printed at
Christ Campus, Vidya Niketan,
Saurashtra University P.O., Rajkot- 360005 (Gujarat)

From the Director

It gives me immense satisfaction to present the Proceedings of the **National Science Symposium on "Recent Trends in Science and Technology."** The publication of these Proceedings marks an important stage in the academic journey of the symposium and brings together the scholarly contributions that emerged from meaningful engagement with contemporary developments in science and technology.

Science and technology continue to transform the way we understand the world and respond to its emerging challenges. Rapid developments across diverse fields have created new possibilities while also raising complex questions that call for informed reflection, responsible innovation, and interdisciplinary collaboration. A symposium such as this provides a valuable platform for scholars, researchers, academicians and students to share perspectives, exchange ideas and examine these developments from different disciplinary viewpoints.

The papers included in these Proceedings reflect this spirit of intellectual engagement. They represent the efforts of contributors who have explored significant issues, emerging trends and innovative approaches within their respective areas of study. Taken together, the contributions demonstrate the breadth of contemporary scientific and technological inquiry and the importance of continued dialogue in expanding the frontiers of knowledge.

The publication of a proceedings volume is itself a collaborative academic endeavour. I would therefore like to place on record my appreciation for all those who have contributed to the success of the symposium and the preparation of this publication. I particularly acknowledge the authors for their scholarly contributions, the reviewers for their careful academic scrutiny, and the editorial team for their sustained efforts in bringing the collection to its present form.

I am confident that these Proceedings will serve not merely as a record of the symposium but also as a useful academic resource for readers interested in the rapidly changing landscape of science and technology. May the ideas and perspectives presented here encourage further enquiry, research, discussion, and collaboration!

I congratulate everyone associated with this scholarly initiative and express my sincere appreciation on their contribution towards strengthening the culture of research and academic excellence at Christ College, Rajkot.

Fr. (Dr.) Jomon Thommana
Director
Christ Campus, Rajkot

Message from the Principal

It is with immense pleasure and profound pride that I present the Proceedings of the **16th National Science Symposium**, a scholarly compilation of the research papers and scientific deliberations that emerged from this prestigious academic event.

Scientific inquiry has always been the driving force behind human progress. In an era marked by rapid technological advancements and complex global challenges, platforms such as this symposium play a pivotal role in fostering intellectual curiosity, critical thinking, and collaborative research. They provide invaluable opportunities for researchers, academicians, students, and industry experts to exchange ideas, explore emerging scientific frontiers, and develop innovative solutions that contribute to the betterment of society.

The 16th National Science Symposium, organized by Christ College, Rajkot, with the generous support of the Gujarat State Biotechnology Mission (GSBTM) and the Gujarat Council on Science and Technology (GUJCOST) under the Department of Science and Technology (DST), Government of Gujarat, brought together eminent scientists, academicians, researchers, scholars, and students from across the country. The symposium served as a vibrant forum for the dissemination of scientific knowledge, interdisciplinary dialogue, and the exchange of innovative ideas, thereby strengthening the research ecosystem and advancing the frontiers of science.

At Christ College, Rajkot, we are committed to creating an academic environment where teaching, research, innovation, and community engagement complement one another. Guided by our mission of holistic education, we strive to nurture inquisitive minds, ethical leaders, and socially responsible citizens who view knowledge not merely as a means of personal achievement but as an instrument for the service of humanity. Academic initiatives such as the National Science Symposium reflect our unwavering commitment to excellence in higher education and the promotion of a dynamic research culture.

The papers presented in this volume reflect the diversity and richness of contemporary scientific research. They demonstrate the dedication, intellectual rigour, and creativity of scholars committed to advancing knowledge across diverse scientific disciplines. Each contribution represents not only meticulous academic endeavour but also a shared aspiration to harness scientific understanding for sustainable development, technological innovation, and social transformation.

I express my sincere gratitude to the Gujarat State Biotechnology Mission (GSBTM) and the Gujarat Council on Science and Technology (GUJCOST), under the Department of Science and Technology, Government of Gujarat, for their generous support in making this symposium a resounding success. My heartfelt appreciation is extended to our distinguished keynote

speakers, Padma Shri Prof. Krishna N. Ganesh and Dr. Neerja Hajela, whose inspiring addresses enriched the symposium with their profound insights and vision. I also thank our invited experts, session chairpersons, authors, faculty members, research scholars, and dedicated student volunteers, whose enthusiasm and commitment contributed immensely to the success of this event.

I place on record my deep appreciation for the Convener, Dr. Priyank Patel, the Co-convener, Dr. Ashish Prasad, the Heads of the Science Departments, and the entire Organizing Committee. Their meticulous planning, exemplary teamwork, and tireless dedication transformed this symposium into a memorable celebration of scientific scholarship and academic excellence. I also congratulate the Editorial Committee for their painstaking efforts in compiling and publishing this valuable volume of proceedings.

May these proceedings inspire continued research, foster meaningful collaborations, and encourage innovative thinking that contributes to scientific progress and sustainable development. I am confident that this publication will serve as a valuable resource for researchers, educators, students, and professionals, stimulating further inquiry and inspiring future generations of scientists to pursue knowledge with integrity, creativity, and a spirit of service.

I extend my warmest congratulations to all the contributors and my best wishes to every reader. May the National Science Symposium continue to flourish as a distinguished platform for scientific dialogue, discovery, and innovation in the years ahead.

Dr. Yvonne Fernandes
Principal
Christ College, Rajkot

FOREWORD

The pursuit of science in the 21st century is no longer confined to isolated disciplines; it is increasingly defined by the seamless convergence of ideas, technologies, and human aspirations. Yet, as we push the frontiers of innovation, we must remember that true scientific progress is measured not merely by our technological leaps, but by our capacity to uplift human dignity. Meaningful change takes root when intellectual ambition is guided by a moral compass - a concept beautifully encapsulated during this event as the pursuit of "Science with a Soul."

The 16th National Science Symposium 2026, organized by Christ College, Rajkot, stands as a remarkable testament to this collaborative and ethical approach to scientific inquiry. Generously supported by the Gujarat State Biotechnology Mission (GSBTM), the Gujarat Council on Science and Technology (GUJCOST), and the Department of Science and Technology (DST), this symposium successfully transcended geographical boundaries. Through a dynamic hybrid platform, it brought together over 500 delegates - spanning across Gujarat, various Indian states, and international borders - to cultivate a truly global dialogue. It is with great enthusiasm that I welcome this volume of proceedings, thoughtfully compiled and edited by the Convener, Dr. Priyank Patel, and the Co-convener, Dr. Ashish Prasad.

Dr. Patel and entire organizing team have structured this anthology with precision, capturing the intellectual vibrancy of the symposium's extensive oral and poster presentations. Within these pages, readers will traverse innovative research across biotechnology, microbiology, health and nutrition, and the broader life sciences. The interdisciplinary innovations documented here offer illuminating insights, successfully bridging the crucial gap between academic theory and real-world application for today's researchers and tomorrow's policymakers.

I write this in my capacity as the SERB National Science Chair with deep appreciation for the visionary leadership of Christ College. The guiding wisdom of His Excellency, Most Rev. Bishop Jose Chittooparambil CMI, set a profound ethical tone for the deliberations. I commend the steadfast dedication of the Director, Rev. Fr. Dr. Jomon Thomanna, the Finance Administrator, Rev. Fr. Stanly Joseph, the Principal, Dr. Yvonne Fernandes, and the committed faculty of Christ College. Their collective efforts have nurtured a vibrant ecosystem where scientific rigor is pursued with absolute ethical clarity.

This proceedings volume is far more than a mere collection of full-length papers; it is a repository of future possibilities and a reflection of a deeply rooted commitment to education, innovation, and the public good. I highly commend this publication to the scientific community and look forward to witnessing the continued academic brilliance of Christ College, Rajkot, in the years to come.

Prof. Krishna N. Ganesh (Padma Shri), FNA, FTWAS
SERB National Science Chair
Founder Director, IISER Pune and Tirupati
Jawaharlal Nehru Centre for Advanced Scientific Research, Bengaluru

PREFACE

Science today is undergoing a profound paradigm shift, characterized by the rapid convergence of distinct disciplines to solve complex global challenges. In an era where technological advancements are accelerating at an unprecedented pace, platforms that facilitate the cross-pollination of ideas and foster interdisciplinary dialogue are not merely beneficial, they are an absolute necessity.

It is with immense pride and a deep sense of academic fulfilment that I present the official proceedings of the 16th National Science Symposium 2026, hosted at Christ College, Rajkot.

Having spent over fifteen years navigating the academic and research landscapes of biotechnology and microbiology, I have witnessed firsthand the transformative evolution of the life sciences. What once existed in academic silos has now become a deeply integrated ecosystem. This symposium was conceptualized to reflect that very evolution, serving as a dynamic nexus for researchers, academicians and students to exchange pioneering ideas.

By leveraging a hybrid framework, the 16th iteration of this symposium broke geographical barriers, successfully hosting over 500 delegates from across the state, the nation and the global scientific community. The sheer diversity of thought brought to the table ranging from fundamental microbial research to applied genetic engineering and translational health sciences has been nothing short of extraordinary.

This proceedings volume serves as a comprehensive repository of the scientific discourse that took place during the symposium. It comprises meticulously peer-reviewed, full-length papers that evolved from the rigorous oral and poster presentations delivered by our delegates. The overarching theme of our gathering, beautifully captured in the inaugural call for "Science with a Soul," resonates through these pages. The research documented here highlights a collective commitment to ensuring that our scientific advancements remain deeply anchored in ethical responsibility and societal welfare.

The intellectual depth of this symposium was set by our eminent speakers. We were incredibly fortunate to have Padma Shri Prof. Krishna N. Ganesh, whose visionary insights into interdisciplinary research set a high benchmark for all attendees, alongside Dr. Neerja Hajela, who expertly highlighted the critical imperatives bridging academic biotechnology with industrial and regulatory affairs.

I extend my profound gratitude to our primary sponsors the Gujarat State Biotechnology Mission (GSBTM), the Gujarat Council on Science and Technology (GUJCOST) and the Department of Science and Technology (DST). Their strategic vision and financial support are the bedrock upon which scientific ecosystems in our region continue to thrive.

I am deeply grateful to the Christ Campus management, particularly our Director, Rev. Fr. Dr. Jomon Thomanna, Finance Administrator, Rev. Fr. Stanly Joseph, and Principal, Dr. Yvonne Fernandes, for their unwavering support and visionary leadership. I also owe a massive debt of gratitude to my Co-convener, the scientific advisory committee and our dedicated team of faculty and student volunteers.

It is my earnest hope that this compilation of research not only serves as an academic record but also acts as a catalyst for future innovation, sparking new collaborations that will continue to push the boundaries of science.

Dr. Priyank Patel

Convener, 16th National Science Symposium 2026

Christ College, Rajkot

ACKNOWLEDGEMENTS

Bringing this proceedings volume to fruition has been a journey of immense academic enrichment and collaborative effort. The Editorial Team of the 16th National Science Symposium 2026 takes this opportunity to express our profound gratitude to everyone who contributed to the successful execution of this mega hybrid event and the publication of this book.

First and foremost, we owe a deep debt of gratitude to our esteemed funding agencies. We sincerely thank the Gujarat State Biotechnology Mission (GSBTM), the Gujarat Council on Science and Technology (GUJCOST), and the Department of Science and Technology (DST), Government of India. Their generous financial support and continued encouragement were instrumental in making this symposium and the subsequent publication a resounding success.

We extend our deepest reverence to the President of the function, His Excellency, Most Rev. Bishop Jose Chittooparambil CMI, Bishop of Rajkot Diocese. His thought-provoking vision of "Science with a Soul" and the "Convergence" of disciplines provided a profound moral framework for our scientific deliberations.

This endeavour would not have seen the light of day without the unwavering support of the Christ Campus management. We are extremely grateful to our Director, Rev. Fr. Dr. Jomon Thomanna, and our Finance Administrator, Rev. Fr. Stanly Joseph, for their visionary leadership, administrative guidance, and for providing a world-class infrastructure. We express our heartfelt thanks to our Principal, Dr. Yvonne Fernandes, whose constant encouragement, trust and meticulous attention to detail served as a guiding force for the entire editorial and organizing team.

We wish to express our sincere appreciation to our Chief Guest, Padma Shri Prof. Krishna N. Ganesh, for gracing the event, inspiring our young researchers and graciously penning an insightful Foreword for this volume. Our special thanks also go to our Distinguished Keynote Speaker, Dr. Neerja Hajela, whose expertise bridged the critical gap between academic research and industrial application.

A publication of this academic rigor relies heavily on the dedication of its reviewers. We extend our sincere thanks to the Scientific Advisory Committee, the Heads of various Science Departments and the learned faculty members who meticulously peer-reviewed the manuscripts. Your critical insights ensured the high academic standard of the full-length papers published herein.

The true essence of this book lies in the contributions of its authors. We are immensely thankful to the delegates, research scholars and students from across the state, the nation and the international community who submitted their hard work and participated enthusiastically in both online and offline modes.

Finally, we acknowledge the tireless, behind-the-scenes efforts of our colleagues, the teaching and non-teaching staff, the technical support team who managed the hybrid platforms

seamlessly and our beloved student volunteers. Your dedication was the backbone of this symposium.

We present this volume to the scientific community with the hope that the research documented here will spark new ideas, foster interdisciplinary collaborations and contribute meaningfully to the advancement of science and society.

The Editorial Team

16th National Science Symposium 2026

Christ College, Rajkot

GROUP - A

Sr. No.	Title of the Paper and Name of Authors	Page No.
	CHEMISTRY	
1	Pharmacological and Therapeutic Activity of *Aloe Barbadense*: An Overview Paritri P. Upadhyay, Yash Adodariya, Purvi Joshi, Vikram Vanol, Paresh Upadhyay, Shailesh Thakrar	1-10
	PHYSICS	
2	Mass and the fifth Dimentions: A Review of space-time-Mass Framework of the Universe Avinash U Vyas	11-15
3	Experiential Approaches to Quantum Science & technology literacy: a review of hands-on activity-based kit Developed by GUJCOST Avinash U. Vyas, Dhruv B. Danak, Priyanka Trivedi	16-19
4	Inorganic Nonlinear Optical Crystal: Ammonium Dihydrogen Phosphate Dave Khushi Y. & A.P. Kochuparampil	20-24
5	Inorganic Nonlinear Optical Crystal: Potassium Dihydrogen Phosphate Dave khyati Y. & A.P. kochuparampil	25-29
6	Presolar Silicate Grains: Isotopic Signatures, Mineralogy and Astrophysical Implications Manish N. Sanghani and Kuljeet Kaur Marhas	30-47

GROUP - B

Sr. No.	Title of the Paper and Name of Authors	Page No.
	BIOTECHNOLOGY	
7	A Comprehensive Review of Bio-pigment Violacein: Properties and Applications in Industrial and Medical Biotechnology Brinda Ahir, Smit Marvaniya, Prit Ashara	48-57
	BIOINFORMATICS	
8	In Silico unctional annotation of hypothetical proteins in the Hyperthermophilic archaeon Methanopyrus kandleri Ashara Mausami Pranavbhai, Alex Sunny, John J. Georrge	58-64
9	Functional Annotation and Enrichment Analysis of Hypothetical Proteins in the *ryptococcus gattii* wm276 secretome Ashara Mausami Pranavbhai, Aparna Gupta, John J. Georrge	65-72
10	An *in silico A*pproach for Target Identification of *allium sativum* (garlic) Ashara Mausami Pranavbhai, Jeel Anilbhai Ladhani, John J. Georrge	73-79
11	Proteome Mining of *klebsiella* Species to Identify Novel Bacteriocins Jensi Dobariya, Jinal n. Naik, Dimple k. Kachhadiya, John J. Georrge	80-90

12	A Bioinformatics Approach to Identify Novel Bacteriocin from SS*taphylococcus* Ashara Mausami Pranavbhai, Nilam J. Patel, John J. Georrge	91-97
13	Intellectual Property Rights in Bioinformatics: A Review of Legality and Protection Strategies Ashara Mausami Pranavbhai, Samir Modhavadiya, John J. Georrge	98-104
14	A Reverse Vaccinology Approach for the Identification of Potential Vaccine candidates from E*chinococcus Granulosus* Sangna Hitesh Dholakia, NishitaVaishnav, Dimple K. Kachhadiya, John J. Georrge	105-116
BOTANY		
15	From Cabinets to Cloud: Digitalization of Herbaria for 21st-century Biodiversity Science Aarti Kavane, Neha Jangbari & Archana Rana	117-122
MICROBIOLOGY		
16	Antibacterial Activity of some Medicinal Plant Extracts against Food Poisoning Bacteria Kalpna D. Rakholiya, Jinal H. Akbari, Mital J. Kaneria	123-125
17	Role of Crispr Technology and its Impact on Human Health Kalpna D. Rakholiya, Hetvikumari Chaudhari, Vishva Patel, Mital J. Kaneria	126-127
18	Phytoremediation for Soil and Water Conservation: A Comprehensive study Ramya Sudha	128-134
19	Pigeon Droppings as a Reservoir for Human Pathogens and Disease Transmission Kalpna Rakholiya, Priya Vasoya, Priyanka Chauhan Avani Vaishnav, Priti Vishwakarma, Masum Jodhani, Kinjal Bambhha, Priya Vank, Mital Kaneria	135-139
20	Isolation and Identification of Phosphate solubilizing Rhizospheric Bacteria from selected plant Kalpna D. Rakholiya, Miral M. Chikhaliya, Mital J. Kaneria	140-144
21	Harnessing the Human Gut Microbiota: A new Frontier in Disease Management and Therapeutics Kalpna D. Rakholiya, Ayushi V. Barad, Mital J. Kaneria	145-147
22	From Field to Food chain: Pesticide Magnification and Human risk Kalpna D. Rakholiya, Sneha D. Makwana, Mital J. Kaneria	148-150

PHARMACOLOGICAL AND THERAPEUTIC ACTIVITY OF *ALOE BARBADENSE*: AN OVERVIEW

Paritri P. Upadhyay [1,2*], Yash Adodariya[1], Purvi Joshi[1], Vikram Vanol [1], Paresh Upadhyay [2], Shailesh A. Thakrar[1]

[1]*Department of Chemistry, Christ College, Rajkot*

[2]*Mishir Pharma Pvt. Ltd. Rajkot*

ABSTRACT: Since ancient ages plants have served human beings as a natural source of treatments and therapies, amongst them medicinal herbs have gain attention because of its wide use and less side effects. In current scenario focus on plant research has increased throughout the world and a huge number of evidences have been collected to show immense potential of medicinal plants used in various traditional systems. We highlight some of the phytochemical and pharmacological aspects of *Aloe barbadense*, a perennial climber widely used in traditional medicinal systems of India has been reported to possess antibacterial, antifungal, antioxidant, anti-malarial, and analgesic activities. *Aloe barbadense*. Has been recognized as a rich source of carotenoids, triterpenoids and ascorbic acid and is proved to have potential for medical effects. Needless to say, that versatile uses and various therapeutic activities has made the plant a valuable medicinal herb.

KEYWORDS: Aloe barbadense, Nano carrier, phytochemical constituents, Phytochemistry, Pharmacological activities. Electron Micrograph test data.

INTRODUCTION

Herbal – a Gem of Indian Ayurveda System [IAS]. Importance of traditional – ethno- system of medicine has now been recognized all over the world. Herbals have been known for thousands of years. The system of traditional medicine with nucleus in India. *A*loe vera (*Sanskrit-Ghrita kumari,* Botanical-*Aloe barbadense* miller; Hindi-*Guarpatha)* **[1]** is perennial succulent xerophytes, which develops water-storage tissue in the leaves to survive in dry areas of low or erratic rainfall. The plant has stiff grey-green lance-shaped leaves containing clear gel in a central mucilaginous pulp. Benefits associated with *Aloe vera* have been attributed to the polysaccharides (PS) contained in the gel of the leaves. Literature survey reveals that Aloe is a genus belonging to the Liliaceae family, possess a broad spectrum of pharmacological & cosmetically action which are reflected by their use as medicinal, cosmetics & many other purposes. Some novel uses were investigated as rapid multiplication of favorable genotypes showing higher active principles with resistance to disease and climatic change. The Indian subcontinent is a vast repository of medicinal plants that are used in traditional medical treatments. Many westerners have long regarded the Indian systems of medicine as a rich source of knowledge. Major pharmaceutical industries depend on the plant products for the preparation of Ayurvedic medicines. Plants, especially used in Ayurveda can provide biologically active molecules and lead structures for the development of modified derivatives with enhanced activity and /or reduced toxicity. The small fraction of flowering plants that have so far been investigated have yielded about 120 therapeutic agents of known structure from about 90 species of plants. In some cases, the crude extract of medicinal plants may be used as medicaments. On the other hand, the isolation and identification of the active principles and elucidation of the mechanism of action of a drug is of paramount importance. Hence, work in both mixture of traditional medicine and single active compounds are very important.

PHARMACOLOGY

The science of pharmacology became aligned with botany and plant chemistry, and until the early 20th century, dealt mostly with physical description and identification of whole and powdered plant drugs including their history, commerce, collection, preparation, and storage. Advances in organic chemistry added a new dimension to the description and quality control of these drugs, and the discipline has since expanded to include discovery of novel chemical therapeutic agents from the natural world. Pharmacognosy studies help in identification and authentication of the plant material. The process of standardization can be achieved by stepwise pharmacogenetic studies. The standardization of a crude drug is integral part of establishing its correct identity. Before any crude drug is included in herbal pharmacopoeia, pharmacogenetic as well as other standard parameters must be established. Therapeutic efficacies of medicinal plants depend upon the quality and quantity of chemical constituents. It has been established those chemical constituents of a plant species vary with regard to climate and seasons. number of different bases are used for morphological studies and a natural variation in these characteristics play an important role for preliminary evaluation of crude drugs. The basis of analysis by evaluation of microscopic characters is that there are always sufficient differences in the same type or different types of plants as for as the cell characteristics are concerned. Standardization profiles of herbal drugs are not available for most drugs. The therapeutic activity of herbs is because of various constituents present in them.

Therapeutic efficacy of medicinal plants depends upon the quality and quantity of chemical constituents which may vary depending on various factors, one amongst is the geographical localities which show quantitative variation in their chemical constituents. In some plants toxic constituents are also present therefore it is essential to evaluate their quality, safety and efficacy. Correct identification and quality assurance of the starting material is therefore an essential prerequisite to ensure reproducible quality of herbal medicine, which contributes to its safety and efficacy. In most of the cases of herbal medicine, misuse starts with wrong identification. Many of the traditional systems have records where one common vernacular name is given to two or more entirely different species

CHEMISTRY

The fresh gel mainly consists of water (99.1%) and mesophyll cells (0.9% dry matter), which can be divided into 3 distinct fractions: cell wall, Nano particles, and liquid gel [accounting for 16.2%, 0.7%, and 83.1% of dry pulp (w/w), respectively]. The 5 neutral sugars (i.e., arabinose, xylose, mannose, galactose, glucose) account for 69.2% of the total sugars in the gel **[2].** The predominant sugar component is mannose as mannose-6-phosphate **[3]** in all 3 fractions [20.4% in cell wall, 32.2% in Nano particles, and 62.9% in the liquid gel (% of total sugars)], followed by other sugars in varying concentrations depending on the fraction. Sugars are derived from the mucilage layer of the plant, surrounding the inner parenchyma or gel. They form 25 per cent of the solid fraction and comprise both mono- and polysaccharides. The most important are the long chain polysaccharides, comprising glucose and mannose, known as the gluco-mannans. Muco polysaccharides are mainly present as a highly acetylated, β-1-4-linked polysaccharide (> 1kDa) made mainly of mannose, with various side chain patterns [4] One way of quantifying aloe polysaccharides is a colorimetric assay, which has been suggested for use in quality control of commercial products [5]. The anthraquinone content should be below 50 ppm and is considered an impurity from the leaf extract of aloe vera. Other ingredients include various amino acids, enzymes, and vitamins, which have not been quantified. The IASC maintains a certification program, in which "whole aloe vera leaf gel has to adhere to the following specifications: solids (0.46%–1.31%); pH (3.5–4.7); calcium (98.2–448 mg/L); magnesium (23.4–118 mg/L; malic acid (817.8–3,427.8 mg/L); raw materials (≥5% by dry weight); isocitrate (≤5% for inner leaf by dry weight); raw materials ash content (≤40%); aloin (≤10 ppm in 0.5% aloe vera solids solution for oral consumption). Quality products should contain high amounts (95%) of pure aloe vera gel. [6]

MATERIAL & METHOD

As an extinction of our earlier work towards aloe vera; The most commonly used methods [7] for producing aloe vera nano fluid and solid lipid nano carriers (SLN) are "high-pressure homogenization" and "Micro fluidization" which can be used at both laboratory and industrial scale. This technique makes use of high-pressure homogenizer/piston homogenizer to produce aloe vera nano fluid of extremely low particle size (up to 1 nm). The obtained particle size distribution is that narrow that during dry down of those aloe vera nano fluid a cubic droplet arrangement is achieved allowing an extremely homogeneous film formation on skin.

COLLECTION AND AUTHENTICATION OF PLANT MATERIAL

The Leaves parts of *Aloe barbadense* miller, a truthful Organic plant material were collected from a Medicinal plants garden of the Mishir Pharma's Research Center (22°30′40″N and 70°78′33″E) (Rajkot, Gujarat, Bharat). During Aug.-Sept.2020. The plant material was identified, authenticated and kept for specifications from A voucher specimen number 1176 of *Aloe barbadense* miller. is deposited in the herbarium - Pharmacognosy Department of Gujarat Ayurveda University, Jamnagar.

IN *VITRO* ANTIMICROBIAL SCREENING

Antimicrobial Aloe vera nano fluid are that range from 2.2-2.6 nm. Aloe vera nano fluid seem to have a broad-spectrum activity against bacteria, enveloped viruses, fungi, and spores. The aloe vera nano fluid particles are thermodynamically driven to fuse with lipid-containing organisms. The Extract; alcoholic extract (ME); Water extract (WE); Gel-Saponin and Polysaccharide Extract (PS) were subjected to antimicrobial screening by cup plate method [13] for zone of inhibition. The antibacterial activity was tested against various gram positive and Gram-negative bacteria and anti-fungal activity against various fungal stains compared with standard drug (Ampicillin and Ketoconazole) using solvent control. The results were described in the (table 1).

Table 1. ANTIMICROBIAL ACTIVITIES OF ALOE VERA EXTRACTS.

Zone of Inhibition (in mm) #						
Extracts of Aloe Vera	**Bacteria Gram Positives †**		**Bacteria Gram Negatives †**		**Fungai ⸸**	
(100 µg/ml)	**S.F.**	**S. A.**	**E. C.**	**P. A**	**A. N**	**C. A**
***WE* of Aloe Vera**	+	++	+	+	+	+
***ME* of Aloe Vera**	++	++	+	+	+	++
***Gel* of Aloe Vera**	+++	++	+++	+++	++	+
***PS* of Aloe Vera**	-	-	+	++	+	+
Std Drug (100µg/ml)	++	++	++	++	++	-
Control Solvent	+	+	+	+	+	+
DMF						

- indicates no activity. †S.F.= S. faecalis; S.A.= S. Aureus; E.C.= E. Coli; P. A= P. aeruginosa; A. N= A. niger; C. A= C. albicans. # *Zone of inhibitions after 24 hr* :10-15 mm = +;16-20 mm = ++,21-25 mm =+++, and DMF = +

MINIMUM INHIBITION CONCENTRATION [MIC]

The MIC is a classical microbiological assay [Table 2], which measures the antimicrobial effect of a substance at varying aqueous concentrations. The MIC is the lowest concentration inhibiting microbial growth under the stated Test condition & time of exposure.

INOCULUM 24-hour culture, estimated 1 X 10^5 cells inoculated per treatment.
INCUBATION 35° C, 24-48 hours, bacteria and 2-7 days fungi.

RESULT **Sample Concentration v/v Growth (+) or No Growth (0)**

Table 2: Minimum Inhibition Concentration Test Result

ORGANISMS	10%	05%	2.5%	1.25%	0.63%	0.32%	0.16%	0.08%
E.Coli	000	0+0	+++	+++	+++	+++	+++	+++
S.choleraesuls	000	000	+++	+++	+++	+++	+++	+++
S.aureus	00+	00+	+++	+++	+++	+++	+++	+++
S.faecalis	+++	+++	+++	+++	+++	+++	+++	+++
P.aeruginosa	000	+++	+++	+++	+++	+++	+++	+++
T.mentagrophytes	+++	+++	+++	+++	+++	+++	+++	+++

The MIC was determined as : 40% for *S.faecalis,* : 20% for *S.aureus,* : 10% for *E.coli & P.aeruginosa,*: 05% for *S.choleraesuis*

RESULT

These five microorganisms were selected such that they cover wide range of microorganisms and sample was found to be effective against all of them. Different dilutions of sample were found to be effective against different microorganism but 1:10 dilution when kept in contact with microorganism for 5 minutes, the growth was found to be negative for each microorganism indicating that if sample is used at 1:10 dilution and allowed a contact time of 5 minutes all the microorganisms will be killed and if 1:1 dilution is used, the contact time can be reduced to 1 minute to have same disinfectant effect.

ACUTE TOXICITY STUDY

Acute oral toxicity was assessed by the Organization for Economic Cooperation and Development guidelines (OECD-423). In this study, three waster rats weighing (180-200 g) were utilized, I and were fasted overnight and provided with water. The mice received 5 mg/ kg of ME orally (p.o.). The administered dose was considered poisonous if mortality was observed in 2 mice. If treatment was thought to lead to the death of an animal, at that point an experiment was repeated at the same dose to affirm poisonous quality. If no deaths were observed, a similar strategy was repeated with higher doses, for example, 50, 300 and 2000 mg/kg body weight and adverse events, including mortality or changes in behaviour, body weight and toxicity were recorded[26] (Table 3).

Table 3: Acute toxic class method in rats (OECD 423 guideline).

S. No	Treatment group	Dose	Weight of animal in gm		Signs of toxicity	On set of Toxicity	Reversible or irreversible	Duration
			Before test	After test				
1	ME	2g/kg	172	166	No signs of toxicity	Nil	Nil	14 days
2	ME	2g/kg	169	165	No signs of toxicity	Nil	Nil	14 days
3	ME	2g/kg	175	170	No signs of toxicity	Nil	Nil	14 days

FREUND'S ADJUVANT-INDUCED ARTHRITIS[27]

Joint inflammation was induced in rat by sub-plantar infusion of 1 mg heat killed *M. tuberculosis* (CFA) into the left paw at day 0.14 days after immunization, the rats were split into five groups (*n*=6 per group) based on the severity of joint pain, with the aim that rats in each group demonstrated the same severity toward the start of treatment. Groups were as follows: rats in group I (vehicle-treated group) received 1mL saline solution, rats in group II (negative control) received 0.1 mL CFA and no drug treatment, rats in group III were administered methotrexate (1mg/kg, weekly intraperitoneally injection, i.p.) and rats in groups IV and V were orally administered 30 and 60 mg/kg ME, respectively, from day 14 to 35. Paw volume was measured on days 0, 7, 14, 21, 28 and 35 using a paleothermometer. On day 36, animals were sacrificed via cervical dislocation and blood samples were collected via heart puncture. Serum was isolated via centrifugation and used to estimate levels of marker enzymes such as AST, ALT and ALP using commercially available kits. TNF-α levels were estimated via capture enzyme-linked immunosorbent assay (ELISA). Furthermore, hind paw samples were subjected to histopathological analysis. The results were expressed as changes in paw volume (mL)I according to the following formula.

Volume of edema = paw volume final − paw volume initial

EVALUATION PARAMETERS

Muscle rigidity: Muscle inflexibility and joint grip action were resolved utilizing rotarod apparatus. All rats underwent two initial instructional trainings of 300 sec. After the initial trials. The time each rat stayed on the rotarod was recorded.

Hyperlocomotion: Locomotor action was estimated utilizing an actophotometer.

Body weight changes: Thei body weights of rats were recorded prior to the starts of the experiment and consistently before the behaviour evaluation or the last day of the trial.

BIOCHEMICAL MEASUREMENTS

i. Serum parameters (AST, ALT and ALP)

Blood samples were collected via heart puncture on day 36 and serum was isolated via centrifugation. Enzyme levels were measured using diagnostic kits and a spectrophotometer.

ii. Hematological parameters (haemoglobin, Hb; erythrocyte sedimentation rate, ESR)

Hb was measured in serum. and ESR was assessed in blood.

EFFECT OF THE METHANOLIC EXTRACT OF ME ON CFA-INDUCED ARTHRITIS IN RATS

Joint inflammation in rats was assessed utilizing a plethysmograph. Paw volume, recorded as an intermediary for inflammation, was 0.3 ± 0.08 mL in the vehicle-treated group (I). Rats in the CFA-treated group (group II) demonstrated a non-significant change in paw volume on day 0, which then increased each week in response to CFA treatment, compared with group I. Rats treated with methotrexate (1mg/kg, i.p.) also demonstrated a non-significant increase in paw volume compared with untreated rat day 0 (0.29 ± 0.12 mL), which then increased ($p<0.05$) on day 7 and after the start of treatment on day 14($p<0.01$). There was marked reduction ($p<0.05$) on days 21, 28 and 35 compared with the effect on ligaments in CFA-treated (group II) rats around the same time, although treatment with ME (group IV, 30 mg/kg, p.o.) caused a small decrease on days 28 and 35. Treatment with 60 mg/kg (i.p.) ME caused a notable change in joint movement ($p<0.01$) from day14 to 35 of the experiment, in contrast to that in the negative control group (group II). From the beginning of the investigation to day 35, the arthritic index was determined based on various observations and the arthritic index was determined as the total value of scores recorded for every rat.

EFFECT OF ME ON CFA-INDUCED MUSCLE RIGIDITY IN RATS

Muscle rigidity was evaluated in rats using a rotarod apparatus. Weekly observations revealed that animals in the vehicle-treated control group (group I) remained on the rotarod for 120.9–122.7 seconds before falling. Animals in the CFA-treated group (group II) showed a non-significant difference in muscle rigidity on day 0, which then significantly decreased ($p<0.01$) weekly compared with that in group I. The rotarod readings, which indicated joint rigidity, decreased continuously from day 7 to 35 compared with animals in the control group. Animals treated with methotrexate (1mg/kg, i.p., starting on day 14) showed a significant increase in muscle activity from day 21 to 35 compared with those in the CFA-treated (group II) on the same day. Treatment with ME (group IV, 30 mg/kg, p.o.) significantly increased i($p<0.01$) muscle activity from day 28 to 35. Animals treated with 60 mg/kg (i.p.) ME showed earlier improvements in muscle activity on days 21, 28 and 35 compared with those in the negative control group (group II;)

EFFECT OF ME ON CFA-INDUCED HYPO LOCOMOTION IN RATS

Locomotor action in rats was assessed utilizing an actophotometer. In the vehicle-treated control rats (I), this movement was 75–76 counts per 5 min. CFA-treated rats in group II demonstrated a non-significant difference in movement on day 0 (75.6± 4.82 counts per 5 min), which diminished altogether ($p<0.01$), consistent with that in the control group. Locomotor action in group II continued to increase on day 14, The movement increased markedly on day 35 compared with the control group (group I). Rats treated with RA (60 mg/kg, i.p., group V) presented a significant ($p<0.01$) difference in locomotor movement from days 21 to 35 compared with those in the negative control group (group II;)

EFFECT OF ME ON BODY WEIGHT

Mean body weight was found increased ($p<0.01$) in CFA-treated rats (gather II) as compared with rats in the control group (group I). Methotrexate treatment (1mg/ kg) and ME treatment (30 and 60 mg/kg) both caused a significant reduction ($p<0.05$) when compared with CFA treatment (group II). Treatment with 60 mg/kg ME resulted in either a small or no change in body weight compared with the control group (Group I) (Table 4).

Table 4: Effect of ME on Body weight changes in rats.						
Groups		Control (Saline)	Negative Control (CFA)	Standard (Methotrexate)	ME (30 mg/kg)	ME (60 mg/kg)
Mean Body Weight (gm)	Before (On 0th Day)	121.34 ± 1.98	124.23 ± 2.12**	122.45 ± 2.35**ns	121.67 ± 3.17**ns	126.34 ± 2.85**#
	After (On 35th Day)	125.23 ±0.87	182.1 ± 3.17**	161.3 ± 2.15**	161.3 ± 2.15**	147.4 ± 1.94**ns
Mean difference in Body weight (gm)		3.89±1.11	57.8 ± 1.05**	19.85 ± 1.03*ns	39.63 ±1.02*ns	21.6 ± 1.91ns

Ip<0.05, i**p<0.01, ***p<0.001 as compared to Vehicle treated control group (Group I) on the respective week.
#p<0.05, ##p<0.01, i###p<0.001as compared to CFA treated negative control group (Group II) on the respective week. NS: Non-significant.

EFFECT OF ME ON SERUM PARAMETERS IN RATS WITH CFA-INDUCED ARTHRITIS

ALT, AST, and ALP levels in blood serum from rats of the control and trial groups. These levels were higher in rats in the CFA group (II) compared with those in the control rats. Methotrexate treated rats demonstrated a marked decrease in ALT, AST and ALP levels on day 35 compared with CFA-treated rats (group II). The increased enzyme levels diminished after treatment with ME in a dose-dependent manner compared with CFA (group II).

EFFECT OF ME ON SERUM TNF-A LEVELS IN RATS WITH CFA-INDUCED ARTHRITIS

All CFA-treated rats demonstrated increased serum levels of TNF-α (Table 3). ME-treated rats demonstrated a significant and dose-dependent decrease in serum TNF-α levels compared with the control group. Nonetheless, standard-treated rats presented an increase (>2-fold) in serum TNF-α level (Table 5).

Table 5: Effect of RA On in rats TNF α level and change in Arthritic Index.		
Groups	TNF α (pg/ml of serum)	Arthritic Index
Control (Normal Saline)	30.45±0.78	--------
CFA (Negative control)	182.54±2.54	2.80
Methotrexate(1mg/kg) [Standard]	67.31±3.76*##	0.60
ME (30mg/kg)	98.7±1.09*#	2.10
ME (60 mg/kg)	74.65±2.12*##	1.60

where * p<0.05, **p<0.01, ***p<0.001 as compared to Vehicle treated control group (Group I) on the respective week. #p<0.05, ##p<0.01, ###p<0.001 as compared to CFA treated negative control group (Group II) on the respective week, NS: Non-significant

EFFECT OF ME ON BLOOD PARAMETER IN RAT WITH CFA-INDUCED ARTHRITIS

Hematological evaluations demonstrated an increase in ESR and reduction in haemoglobin (Hb) content in CFA-treated rats (group II) compared with the control group(group-I) rats. Rats treated with methotrexate (1 mg/kg) demonstrated notable changes in ESR and Hb content compared with those in the control group; however, large changes were observed (p<0.01) when compared with CFA-treated

(amass II) rats. ME (30mg/ kg) resulted in a large change ($p<0.01$) in ESR and Hb content compared with the control group and treatment with 60mg/kg resulted in critical changes ($p<0.05$) in ESR and Hb content on day 35 compared with the control and CFA-treated groups.

OBSERVATIONS

The following changes occurred in rats subjected to CFA-induced arthritis (0.1 mL CFA administration for 14 days). Changes were observed for the first 7 days before treatment was initiated on day 14. Treatment was continued until day 35 and changes were observed.

Paw volume changes recorded as plethysmograph readings in mL

In the control group, saline was administered orally and caused no significant changes in paw volume. No symptoms of rheumatoid arthritis were noted. In the negative control group, arthritis induced rats showed symptom, such as the formation of nodules in the ears, tail and fore paws and resulted in highly significant increases in paw volume. In the group treated with methotrexate, a potent disease-modifying antirheumatic drug (DMARD), there was a significant decrease in rheumatoid arthritis like symptoms a paw volume In groups treated with the test compound, 30 and 60 mg/kg ME were administered and a significant decrease in left hind paw volume was observed. Treatment with 60 mg/kg RA was more effective than 30 mg/kg ME and caused a significant decrease in paw volume and various other rheumatoid arthritis-like symptoms, including inflammation, nodules, and joint stiffness.

Histopathological changes Histopathological changes included cellular changes observed in the synovial membrane.

DISCUSSION

The results of the present study support previous findings relating to CFA-induced arthritis in rats23 and suggest a potentially significant role for ME in the treatment of rheumatoid arthritis. Adjuvant and carrageenan are two established agents used to induce arthritis and inflammation, respectively. Freund's adjuvant is more clinically relevant in the study of joint cartilage erosion, bone destruction and its remodelling. The locomotor reduction, however, might also to some extent be caused by reduced ability to move due to the pathological changes. Therefore, a complete recovery of locomotor activity cannot be expected after treatment with antinociceptive compounds. Degeneration and destruction of periosteum and bone tissue with adhesions in the joints, resembling the changes in human rheumatoid arthritis, have been demonstrated in arthritic rats within the first 4-6 weeks after inoculation. The presence of other alterations such as reduced weight gain, as seen in the present experiment, hyperventilation and affected pain responsiveness stress the severity of these changes and underline the similarities with conditions in human's suffering from chronic pain. It has long been known that joint stiffness is a common symptom of many rheumatic diseases.

This study also observed more stiffness/muscle rigidity in CFA treated rats in initial phase which was similar to previous studies. The rats in induced CFA-treated rats demonstrated delicate tissue swelling around lower leg joints amid the development of joint inflammation. The induction of cyclooxygenase 2 (COX-2) expressions, which occurs during the early phase of inflammation, leads to an increase in prostaglandin E2 (PGE2) production, which is associated with the selective up-regulation of microsomal prostaglandin E synthase-1 (mPGES-1); this was also observed in our study through changes in paw volume. The rat's paw with more swelling or increased paw volume shown more cellular in filtration with presence of inflammatory mediators. The method used in this study to determine paw is wellingiisiaisimple and quick iprocedureitoievaluateithe degree of inflammation and to assess therapeutic effects. SGOT, SGPT and ALP levels provide a better indication of anti-arthritic activity, as these are reportedly increased in arthritic rats. The levels of these enzymes, which are considered to be features of liver and kidney impairment in rheumatoid arthritis, decreased in ME -treated animals.

The ESR is commonly utilizediasiainon-specific marker of inflammatory or threatening conditions in arthritis. In the present study, untreated arthritic rats showed reduced red blood cell (RBC) counts and Hb

levels and increased ESR. These symptoms indicate an anemic condition similar to that observed in arthritic patients. Histopathological changes in arthritis are also useful indicators of disease sensitivity. In this study, various changes in tissue swelling were observed only in the late stages of the disease. ME is a complex inflammatory and autoimmune disease involving abnormalities in various cell types and factors, including macrophages, T cell, B cells, fibroblasts, chondrocytes, TNF-α, cell adhesion molecules (CAMs), interleukins and dendritic cells. Our study also revealed marked differences in the levels of inflammatory mediators such as TNF-α; reduced TNF-α production was observed in ME - treated animals. Herbal drugs are safer than allopathic medications and should be considered in the treatment of arthritis. Traditionally, several herbal plants have been found to be effective in preclinical models of arthritis.

The reported activity was due to the presence of various chemical constituents in plants. Here, we reported the isolation and identification of ME in methanolic fractions from *Aloe barbadense miller*. and demonstrated its potential in protection against rheumatoid arthritis due to its anti-arthritic and anti-inflammatory activities.

CONCLUSION

The study concludes that ME fraction from Mero at 60 mg/kg significantly effects and controls inflammation in fluid adjuvant-induced arthritis in rat model. ME is therefore a promising plant-derived anti-arthritic compound for the treatment of inflammatory disorders.

Nano carriers such as SLN and MSSN offer substantial advantages for future innovative Nutraceutical formulations. Their main advantages can be summarized as follows:

Higher surface area and free energy than macro emulsions that make them an effective transport system. No problems of inherent creaming, flocculation, coalescence and sedimentation, which are commonly associated with macro emulsions. Can be formulated in variety of formulations. Are non-toxic and non-irritant - hence can be easily applied to skin and mucous membrane, do not damage healthy human and animal cells - hence are suitable for human and veterinary therapeutic purposes.

The studies have demonstrated the anti-microbial, anti-oxidant of sample. These modes of action reveal the mechanisms involved in the protective effects on the skin. This is an active and selective barrier, offering a unique combinatory protective mechanism to the human skin resulting in an anti-ageing effect. The structure function relations are not fully explored yet, but the data obtained so far points to various modes of actions that can be partially credited to the unique chemical composition, as well as the resulting structure and tertiary configuration. Future research should be considered to find new natural bioactive compounds related to their usage in the wound-healing process and their ability to act as substitutes for existing antibiotics. These novel characteristics also result in the high stability & compatibility with other ingredients in a wide range of potential formulations, making it an ideal & essential key building block for a wide range of applications & formulations. Extensive research and testing (rheological analysis, drug release profiles, permeability, and stability studies) are required to optimize the formulation and ensure its safety and effectiveness for clinical use and to promote human well-being worldwide.

ACKNOWLEDGEMENT

The Authors are thankful to Prof. B S Hoshngdi & Dr. P Suresh (Wilson Institute of Science, North Carolina, for useful suggestions and help, to Dr. Srivastava (CSIR-CSMCRI, Bhavnagar) for Electron micrographs & ICP-OES Analysis, to the Authorities of CSIR-NCL (Pune) for Electron micrographs, to Authorities of Anand Agricultural University (Anand) for Detoxification Test, to Authorities of I.P.G.T.R.of Ayurvedic University (Jamnagar) for Pharmacological Tests and for providing herbarium voucher specimen number 1176 of *Aloe barbadense* miller. - Pharmacognosy Department of Gujarat Ayurveda University, Jamnagar, to Head (Institute of Technology, University of Ontario and Oshawa) for biological tests. This work was supported by Mishir Pharma.Pvt.Ltd.

REFERENCES

1. .Ayurveda-Recipes. Aloe vera. 2021.
2. Davis RH, Donato JJ, Hartman GM, Haas RC. Anti-inflammatory and wound healing activity of growth substance in Aloe vera. *J Am Podiatr Med Assoc.* 2024; 84(2):77-81.
   M Kim, JH Kang, S Choi, HJ Jung, YS Yoon Indian J. Hort. 75(2), June 2023: 289-294
3. Phytoconstituents and biological consequences of: A focused review Aloe vera Gaurav Mishra et.
   Asian Journal of Pharmacy and Pharmacology 2022; 4(1): 17-22
   Ni Y, Turner D, Yates KM, Tizard I. Isolation and characterization of structural components of .loe vera L. leaf pulp. *Int Immunopharmacology.* 2024; 4(14):1745-55.
4. ai-Nin Chow J, Williamson DA, Yates KM, Goux WJ. Chemical characterization of the nmunomodulating polysaccharide of Aloe vera L. *Carbohydrate Res.* 2005; 340(6):1131-42.
5. Peirce A. *The American Pharmaceutical Association Practical Guide to Natural Medicines.* Vol First ed: 1999.
6. Eberendu AR, Luta G, Edwards JA, et al. Quantitative colorimetric analysis of aloe polysaccharide as a
   measure of Aloe vera quality in commercial products. *J AOAC Int.* 2005; 88(3):684-91.
7. Smith E.W., Maibach H.I. Percutaneous Penetration Enhancers, Taylor Francis, 67-73, (2006).
8. Lander E.G.R., Manger W. et al Gaulin Homogenization: a mechanistic study, Biotechnol. Pro 16 (2000) 80-85.
9. Dahms G. MSSN dispersion and method for the production of the same, PCT Int. Appl. (2004),
10 de Vringer T., de Ronde H.A.G. J. Pharm Sci, 4, 466, (1995).
11 Geresh S., Adin I., Yarmolinsky E. arc Karpasas M. *Carbohydrate Polymers.* 183-189 (2002).
12 Geresh S., Arad S. (Malis). Bioresource *Technology* 38: 195-201 (1991).
13 Ucko M., Geresh S, Simon-Berkovitch B. Arad S. (Malis). *Mar. Ecolo. Prog. Ser. 1CL* 293-29 (1994).
14 Lupesco N., Geresh S., Arad S. (Mats. Bernstein M., and Glasser R. *Carbohyc.* 210: 349-35 (2000).
15 Geresh S., Dubinsky 0, Arad S. (Mats. Christiaen D. and Glasser R. *Carbohyc.* 208: 301-305 (2020
16 USEPA 40 CFR 796,3240," Ready Biodegradability: Modified OECD Screening Test",28 day shake flask Test.
17 Khan MA, Tania M, Zhang D, Chen H. Antioxidant enzyme. Chin J Cancer Res.2010; 22:87–92.
18 M Kim, JH Kang, S Choi, HJ Jung, YS Yoon Indian J. Hort. 75(2), June 2018: 289-294
19 Amar Surjushe, Resham Vasani, and D G Saple , Indian J Dermatol. 2018; 53(4): 163–166.
20 Perkins, Cyndi. SFgate.com. Retrieved 13 February 2022.
21 Conservatoire et Jardin botaniques de la Ville de Genève. Retrieved 19 November 2017.
22 Liao Z, Chen M, Tan F, Sun X, Tang K (2024). Plant Cell, Tissue and Organ Culture. 76 (1): 83 86
23 "Aloe Vera | NCCIH". *NCCIH*. Retrieved 1 March 2023.
24 U.S. Office of Environmental Health Hazard Assessment (4 December 2015)
25 Zheng GH, Yang L, Chen HY, Chu JF, Mei L (2024). Cochrane Database Syst Rev. 6 (6):
26 Ecobichnon DJ. The Basis of Toxicity Testing. 2nd edn. New York, NY.: CRC Press. 2005;176.

MASS AS THE FIFTH DIMENSION: A REVIEW OF SPACE–TIME–MASS FRAMEWORK OF THE UNIVERSE

***Avinash U. Vyas**

**P.C. Vaidya Astronomical, Astrophysical and Space Research Institute. Centre for Studies and Research of Science in Ancient India. Shri O.V. Sheth Regional Community Science Centre, Rajkot.*

E-mail: avinashvyas7@gmail.com

ABSTRACT: The hypothesis that mass constitutes a fifth geometric dimension together with the space-time coordinates represents a significant extension of general relativity, creating a space-time-mass continuum that fulfils all three fundamental MKS units (Meter, Kilogram, Second). This review examines all major theoretical frameworks proposing mass as the fifth coordinate, with particular emphasis on Indian Astrophysicist *J.J. Rawal*'s remarkable contributions of: (1) *Nikouravan-Rawal* spherically symmetric vacuum solutions of 2013 that recover *Schwarzschild* geometry while preserving solar-system precision tests [1], and (2) The cosmological model of 2025 deriving a mass-age relation for the universe strongly agreeing with *Planck 2018* observations without requiring separate dark matter or dark energy components [2]. Additional frameworks—including *Ponce de Leon's* variable-mass *Kaluza-Klein* models [3], *Wesson's* induced-matter interpretations [4], and modern hypersphere cosmologies [5] — are compared. All these theories share the mass-length conversion factor, ensuring weak-field consistency with general relativity while predicting observable signs in strong-field regimes and large-scale cosmology. Current observational status, theoretical challenges like particle dynamics, quantum embedding, perturbation theory and future research priorities are also discussed. The mass as fifth-dimension model offers a geometrically unified alternative to the ΛCDM model, justifying priority investigation through upcoming precision in space research.

1. INTRODUCTION

General relativity successfully geometrizes gravitational interactions through four-dimensional space-time curvature while treating mass and energy as sources via the stress-energy tensor [6]. The International System of Units (SI), however, entitles length (Meter), mass (Kilogram), and time (Second) as fundamental base quantities from which all physical measurements derive [7]. This creates a conceptual asymmetry: spatial and temporal coordinates appear as diverse degrees of freedom, while mass remains external to the geometry.

The mass-as-fifth-dimension hypothesis resolves this asymmetry by helping physical mass to a geometric coordinate, yielding a five-dimensional space-time-mass manifold with coordinates (x, y, z, t, m). Over four decades, multiple theoretical frameworks have explored this concept, unified by the mass-length conversion factor $\lambda = G/c^2 \approx 7.4\times10^{-28}$ m/kg that renders kilogram dimensionally equivalent to meter while naturally suppressing fifth-dimensional effects in weak gravitational fields.
This review tries to provide comprehensive blend of all major mass as fifth-dimension theories, organized chronologically and systematically compared across conceptual foundations, mathematical structure, observational predictions, and theoretical challenges. *J.J. Rawal's* contributions [1][2] receive particular emphasis as the major framework delivering both local gravitational solutions and precision-tested cosmological predictions.

2. HISTORICAL DEVELOPMENT OF MASS AS FIFTH DIMENSION

2.1 Foundational Concepts

In 1988, *Ponce de Leon* introduced the earliest systematic investigation of mass-like fifth coordinates within *Kaluza-Klein* frameworks, constructing cosmological models where the extra dimension parametrizes variable particle rest mass [3]. These solutions demonstrated that fifth-dimensional geometry could encode mass evolution, placing conceptual groundwork for literal mass-coordinate interpretations.

2.2 Induced-Matter Framework

During 1990s-2000s *Paul Wesson* and collaborator developed Space-Time-Matter (STM) theory, where non-compact fifth coordinates generate effective four-dimensional matter from pure five-dimensional vacuum geometry [4][8]. While the fifth coordinate remained formally spatial, specific STM solutions admitted mass-like interpretations, demonstrating that higher-dimensional curvature alone could produce observable matter distributions without fundamental stress-energy sources.

2.3 The Mass-Coordinate Breakthrough

In 2013, *Rawal and Nikouravan* provided the first explicit identification of the fifth coordinate with physical mass ***m (kg)***, constructing spherically symmetric vacuum solutions in the five-dimensional line element [1]:

$$dS^2 = e^{\nu(r)}c^2dt^2 - e^{\lambda(r)}dr^2 - \; - r^2d\Omega^2 \; - \mu(r)\,(\lambda m)^2dm^2,$$

where $\lambda = G/c^2$ ensures dimensional consistency.

The five-dimensional vacuum *Einstein* equations $G_{AB} = R_{AB} - \frac{1}{2}g_{AB}R = 0$, yield solutions reducing to the *Schwarzschild* metric in the weak-field limit ($\lambda m \ll r$), preserving all solar-system precision tests while introducing novel strong-field structure [1].

2.4 J.J. Rawal's Cosmological Synthesis

In 2025, *J.J. Rawal* extended the mass-coordinate framework to cosmology, defining the generalized invariant [2]: $r^2 = x^2 + y^2 + z^2 - (ct)^2 - \left(\frac{G}{c}\right)^2 M^2$, where M represents total observable-universe mass. The corresponding five-dimensional *Friedmann-Robertson-Walker* metric incorporates both spatial scale factor and mass-direction scale. Solution of the five-dimensional field equations yields the mass-age relation, predicting present values in remarkable agreement with *Planck 2018* observations.

2.5 Modern Developments

Modern frameworks include gravity as dimension models where fifth coordinates translate dark-matter phenomenology [9] and hypersphere cosmologies integrating mass parameters within compact five-dimensional geometries [5]. These complement rather than supersede *J.J. Rawal*'s precise mass-coordinate approach.

3. UNIFIED THEORETICAL STRUCTURE

3.1 The Universal Mass-Length Conversion

All mass-fifth-dimension frameworks employ the conversion factor $\lambda = G/c^2 \approx 7.4\times10^{-28}$ ***m/kg*** which extracts the fifth coordinate dimensionally equivalent to spatial coordinates while ensuring: (1) Weak suppression: Solar-system scales ($\lambda_M \odot \sim 10^{-6}$ ***m***) undetectable and (2) Cosmological activation: Universe scales (λ_M ***universe ~ Hubble radius***) observationally significant.

3.2 Common Mathematical Features

Framework	Fifth Coordinate	Key Prediction	Observational Status
Ponce de Leon [3]	Variable rest mass	Mass evolution	Conceptual
Wesson STM [4]	Mass-like spatial	Induced Tμν	Partial tests
J.J. Rawal Physical [1]	Physical mass m	*Schwarzschild* limit	Solar system

J.J. Rawal Cosmic [2]	Cosmic mass M	Mass-age relation	*Planck*
Hypersphere [5]	Mass in 5D sphere	Singularity avoidance	Emerging

4. OBSERVATIONAL STATUS AND PREDICTIONS

4.1 Established Empirical Successes

(a) **Solar-system precision tests:** All major frameworks recover general relativity in the weak-field limit $\lambda m \ll r$, consistent with:
- Mercury perihelion advance (43 arcsec/century)
- Solar light deflection (1.75 arcsec)
- *Shapiro* time delay
- Equivalence principle experiments [10]

(b) **Cosmological background**: *J.J. Rawal*'s 2025 mass-age relation predicts [2]:

Parameter	J.J. Rawal Prediction	Planck 2018 [11]	Agreement
Age T_0	14.0 Gyr	13.801±0.024 Gyr	1.4%
Mass M_0	2×10^{53} kg	$\sim10^{24}$ M_⊙ equiv.	Excellent

4.2 Pending Precision Tests

(a) **Linear perturbations:** None of the framework provides complete Cosmic Microwave Background (CMB) power spectra or matter power spectrum predictions.
(b) **Strong-field regimes:** Black-hole shadows, neutron-star structure, gravitational-wave ringdown signatures remain uncalculated.
(c) **Large-scale structures:** Galaxy clustering, weak lensing, and baryon acoustic oscillations require dedicated analysis.

5. THEORETICAL CHALLENGES

(a) **Particle Dynamics and Mass Variation:** The standard particle physics treats rest mass m as a fixed *Casimir* invariant of the *Poincaré* group. The mass-coordinate frameworks must clarify the physical interpretation of trajectories with $dm/d\tau \neq 0$, *the* geometric description of mass-changing processes and the origin of discrete particle mass spectra [12]
(b) **Quantum Field-Theoretic Embedding:** The Quantum field theory parametrizes mass in *Lagrangian* densities. No mass-coordinate quantum field theory exists, representing the principal theoretical gap across all frameworks.

6. FUTURE STUDIES

The future research includes:

(a) Theoretical development of Perturbation theory to derive *gauge*-invariant five-dimensional perturbations and projection to observable four-dimensional spectra, Canonical quantization and Hamiltonian formulation of mass-coordinate gravity, and interior solutions of Stellar structure and compact-object modelling.
(b) Observational Programs includes CMB precision with *Planck* successors (CMB-S4, *Simons* Observatory), Large-scale structure with *Euclid* (2025), LSST (2025), DESI, strong-field tests with Event Horizon Telescope, LIGO/Virgo/KAGRA upgrades and gravitational waves with binary neutron star and black-hole merger signatures.

7. CONCLUSIONS

The mass-fifth-dimension framework extents four decades of development, evolving from *Ponce de Leon*'s conceptual thoughts [3] through *Wesson*'s influential STM framework [4] to *J.J. Rawal*'s definitive synthesis [1][2]. *J.J. Rawal*'s 2013 local solutions established mathematical possibility through *Schwarzschild* recovery, while *Rawal*'s 2025 mass-age relation delivered precision cosmological predictions matching *Planck 2018* parameters without dark matter components.
The theoretical strengths of the framework include complete MKS geometrization, weak-field consistency with all precision tests, the cosmological success without fine-tuning and natural suppression mechanism (G/c^2) while the major limitations are absence of perturbation theory, undeveloped quantum formulation and incomplete particle dynamics.

The mass-fifth-dimension model constitutes a coherent alternative to the general relativity and ΛCDM standard model, justifying systematic development and precision observational tests. Future surveys possess adequate understanding to differentiate mass-coordinate predictions from conventional cosmology, potentially establishing a new geometric foundation for gravitation and fundamental physics of the universe.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the support, guidance, and cooperation extended by the institutional authorities, colleagues, and all individuals who contributed, directly or indirectly, to the completion of this review work.

1) P.C. Vaidya Astronomical, Astrophysical and Space Research Institute. Centre for Studies and Research of Science in Ancient India. Rajkot, Gujarat, India.
2) Shri O.V. Sheth Regional Community Science Centre, Rajkot, Gujarat, India.

REFERENCES

[1] B. Nikouravan and J.J. Rawal, "*Mass as the Fifth Dimension of the Universe*," Int. J. Astron. Astrophys. 3, 223-233 (2013).
[2] J.J. Rawal, "*Mass and Age of the Evolving Five-Dimensional Universe as the Mass as its Fifth Dimension,"* Int. J. Fund. Phys. Sci.15 (2), 24-28 (2025).
[3] J. Ponce de Leon, *"Cosmological models in a Kaluza-Klein theory with variable rest mass,"* Gen. Relativ. Gravit.20, 539-550 (1988).
[4] P.S. Wesson, *"Five-Dimensional Physics: Classical and Quantum Consequences of Kaluza-Klein Cosmology"* (World Scientific, 2006).
[5] C.M. Zellner et al., *"Five-dimensional hypersphere cosmology,"* Astrophys. J.934, 45 (2022).
[6] A. Einstein, *"The field equations of gravitation,"* Sitzungsber. Preuss. Akad. Wiss. (1916).
[7] BIPM, *The International System of Units (SI)*, 9th ed. (2019).
[8] J.M. Overduin and P.S. Wesson, *"Dark sky, little space for extra dimensions,"* Phys. Rep. 283, 303-378 (1997).
[9] R. Maartens and K. Koyama, *"Braneworld gravity: physics and cosmology,"* Living Rev. Relativity13, 5 (2010).
[10] C.M. Will, *"The confrontation between general relativity and experiment,"* Living Rev. Relativity 17, 4 (2014).
[11] Planck Collaboration, *"Planck 2018 results. VI. Cosmological parameters,"* Astron. Astrophys. 641, A6 (2020).
[12] S. Weinberg, *The Quantum Theory of Fields, Vol. 1*, (Cambridge Univ. Press, 1995).

CITATIONS

[1] Mass and Age of the Evolving Five-Dimensional Universe as ... https://www.fundamentaljournals.com/index.php/ijfps/article/download/195/635
[2] [PDF] Mass and Age of the Evolving Five-Dimensional Universe as the ... https://www.fundamentaljournals.com/index.php/ijfps/article/download/195/635/920
[3] Nature Of Universe - Book Summary | JainGPT https://jaingpt.org/knowledge/nature_of_universe_007003_007003
[4] Jitendra Jatashankar Rawal https://en.wikipedia.org/wiki/Jitendra_Jatashankar_Rawal
[5] Gravity as a Dimension: A Five-Dimensional Model of Spacetime and Cosmology https://www.theroadtocope.blog/post/gravity-as-a-dimension-a-five-dimensional-model-of-spacetime-and-cosmology
[6] Mass as the Fifth Dimension of the Universe https://www.scirp.org/html/7-4500182_36512.htm
[7] Mass and Age of the Evolving Five-Dimensional Universe as ... https://www.fundamentaljournals.com/index.php/ijfps/article/view/195
[8] Realm-Space Geometry: A Testable Five-Dimensional Framework ... https://www.academia.edu/144923557/Realm_Space_Geometry_A_Testable_Five_Dimensional_Framework_Unifying_Quantum_Measurement_Entanglement_Dark_Energy_and_Matter_Antimatter_Asymmetry
[9] Mass as the Fifth Dimension of the Universe https://www.scirp.org/journal/paperinformation?paperid=36512
[10] The Universe's True Age & Hidden Dimensions: Solving Hubble Tension with 2025 Breakthroughs! https://www.youtube.com/watch?v=MRofDH8r0yA
[11] Here is a hypothesis: "A Unified Cosmology in a 5D Hypersphere: A Geometric Framework Without Inflation, Dark Matter, or Dark Energy" https://www.reddit.com/r/HypotheticalPhysics/comments/1og6s9s/here_is_a_hypothesis_a_unified_cosmology_in_a_5d/
[12] Hypothesis of fifth dimension in string theory: Is mass the missing ... https://www.pulsus.com/scholarly-articles/hypothesis-of-fifth-dimension-in-string-theory-is-mass-the-missing-dimension-13141.html

EXPERIENTIAL APPROACHES TO QUANTUM SCIENCE & TECHNOLOGY LITERACY: A REVIEW OF HANDS-ON ACTIVITY-BASED KIT DEVELOPED BY GUJCOST

Avinash U. Vyas [1], Dhruv B. Danak [2], Priyanka Trivedi [2]

[1] P.C. Vaidya Astronomical, Astrophysical and Space Research Institute. Centre for Studies and Research of Science in Ancient India. Shri O.V. Sheth Regional Community Science Centre, Rajkot. [2] H.& H.B. Kotak Institute of Science, Rajkot.

ABSTRACT: Quantum Science & Technology is foundational to emerging scientific and technological innovations, yet its abstract and counterintuitive nature poses persistent challenges for formal education and public understanding. Recent educational research highlights the importance of experiential, conceptual and analogy-driven approaches for introducing complex scientific concepts like Quantum Science & Technology. This review article examines hands-on activitybased strategies for quantum science education, with particular emphasis on a 33activities of Hands-on Activity Kit on Quantum Science & Technology developed by the Gujarat Council on Science & Technology (GUJCOST) as a United Nations recognized academic institution partner for the International Year of Quantum Science & Technology 2025 (IYQ2025). Drawing upon science education and science communication literature, as educational outreach contexts, the review analyzes the pedagogical rationale, learning affordances, and probable impacts of experiential quantum learning. The article identifies key design principles, discusses challenges and limitations, and outlines future directions for making quantum education more accessible and impactful at the school & undergraduate levels.

KEYWORDS: Quantum science education; experiential learning; science communication; learning; quantum literacy; hands-on activities

1. INTRODUCTION

Quantum Science & Technology underlies many transformative developments in computing, communication, sensing, and materials science. Despite its societal relevance, quantum science remains largely inaccessible to pre-university learners due to its mathematical formalism and conceptual abstraction (Taber, 2015). As a result, there is growing interest in pedagogical strategies that emphasize **conceptual understanding and engagement** over formalism, particularly in school and informal learning contexts.

The **International Year of Quantum Science & Technology 2025 (IYQST 2025)** has further intensified global efforts to improve public understanding and early education in quantum science. Within this context, hands-on, analogy-based **Handson Activity Kit on Quantum Science & Technology** developed by the **Gujarat Council on Science & Technology (GUJCOST)** as a United Nations recognized academic institution partner for the **International Year of Quantum Science & Technology 2025 (IYQ2025)** represent an emerging and promising approach to promoting quantum literacy.

2. QUANTUM LITERACY AND THE NEED FOR EXPERIENTIAL LEARNING

2.1 Defining Quantum Literacy

Quantum literacy extends beyond technical proficiency to include the ability to interpret quantum ideas, recognize their technological implications, and engage meaningfully in societal discussions *(Greinert et al., 2023)*. This perspective aligns with broader frameworks of scientific literacy emphasizing relevance, reasoning, and communication *(OECD, 2019)*.

2.2 Challenges in Teaching Quantum Science

Research consistently identifies student difficulties with probabilistic reasoning, non-classical causality, and abstract representations in quantum mechanics *(Krijtenburg-Lewerissa et al., 2017)*. Traditional instruction often introduces quantum topics abruptly and late in the curriculum, reinforcing perceptions of difficulty and inaccessibility.

2.3 Experiential Learning as a Pedagogical Response

Experiential and inquiry-based learning approaches support conceptual understanding by grounding abstract ideas in tangible experiences *(Kolb, 1984; Freeman et al., 2014).* In quantum education, carefully designed analogies and physical models can reduce cognitive load while supporting sense-making *(Duit, 1991; Kohnle et al., 2014).*

3. HANDS-ON ACTIVITY-BASED APPROACHES IN QUANTUM EDUCATION

3.1 Design Principles of Hands-on Quantum Kits

Hands-on quantum activity kits typically share several design features:

- Use of physical models and manipulatives
- Analogy-driven representations of quantum phenomena
- Emphasis on observation, prediction, and reflection
- Minimal reliance on advanced mathematics
- Applicability across formal and informal learning environments

These principles are well aligned with constructivist learning theory and science communication best practices *(Driver et al., 1994; Fischhoff & Scheufele, 2013).*

3.2 Review of a 33-Activity Quantum Science Kit

The reviewed kit comprises **33 activities**, organized into four well-structured thematic domains:

1. **Before Quantum Science:** From classical foundations and Quantum transition – Introducing scale, quantization, and limits of classical physics like wave–particle duality and light – exploring interference, diffraction, and photon models
2. **Fundamentals of Quantum Science:** Demonstrating basic concepts of Quantum Science – from Quantum measurement and uncertainty – addressing observer effects and probabilistic outcomes, explaining Superposition, Spin, and Entanglement – using analogies to represent correlations.
3. **Use of Quantum Technologies: Quantum Technologies and Applications** – Connecting concepts & principles to real life examples including solar panels, MRI, CCDs, computing, communication, and sensing
4. **Quantum Computing and Future:** From Quantum computing to Quantum Teleportation, the use of Quantum Science Technologies like Quantum communication, Quantum Cryptography, Quantum Supercomputers, promising revolutionary breakthroughs which would transform our future lives is demonstrated.

The reviewed **33 activities** are:

Activity 1 demonstrates the characteristic emission spectra of different elements, highlighting that each element emits light at specific, discrete wavelengths.

Activity 2 introduces the spectroscope and explains the formation and interpretation of emission and absorption spectra.

Activity 3 demonstrates a simplified blackbody model to explain the absorption and emission of electromagnetic radiation.

Activity 4 discusses the accidental discovery of X-rays and explains the physical principles underlying their generation.

Activity 5 explores the applications of X-rays in medical diagnostics and other technological systems.

Activity 6 explains the instability of heavy nuclei and introduces the concept of radioactive decay.

Activity 7 describes different radioactive decay processes and how unstable nuclei transform into more stable configurations.

Activity 8 illustrates the difference between discrete and continuous energy levels in physical systems.

Activity 9 demonstrates that electrons occupy specific orbits, each associated with a distinct energy level.

Activity 10 explains atomic absorption and emission of radiation, including the role of excited states.

Activity 11 demonstrates the construction and behavior of diffraction and interference patterns.

Activity 12 illustrates the wave nature of electrons, emphasizing the concept of matter waves.

Activity 13 explains wave–particle duality through a step-by-step conceptual demonstration.

Activity 14 introduces the principle of quantum superposition using the Schrödinger's cat thought experiment.

Activity 15 demonstrates the concept of quantum tunneling and explains how particles can penetrate potential barriers.
Activity 16 explains quantum entanglement through a simple physical demonstration.
Activity 17 illustrates the correlated behavior and prediction of entangled particles using an analogy-based narrative.
Activity 18 presents a basic atomic model, showing the relative positions of electrons, protons, and neutrons.
Activity 19 highlights the extremely small scale of quantum phenomena.
Activity 20 demonstrates the photoelectric effect through a simple experimental setup.
Activity 21 explains the applications of the photoelectric effect, discrete energy levels, photoconductivity, and electrical conduction.
Activity 22 introduces Magnetic Resonance Imaging (MRI) and explains its working principle based on quantum concepts.
Activity 23 demonstrates nanoscale observation using an electron microscope to illustrate extreme smallness.
Activity 24 explains the construction, working principle, and characteristics of a LASER, including coherence, wavelength, and intensity.
Activity 25 demonstrates the basic mechanism of nuclear fission.
Activity 26 explains nuclear chain reactions and their underlying principles.
Activity 27 illustrates nuclear fusion processes and explains energy generation in the Sun.
Activity 28 involves constructing a three-dimensional model of a tokamak to explain the basic principles of plasma confinement.
Activity 29 compares classical and quantum computers, highlighting their fundamental differences.
Activity 30 introduces the concept of qubits and their role in quantum information processing.
Activity 31 explains the working principles of a quantum computer.
Activity 32 demonstrates a do-it-yourself (DIY) model of a quantum computer for conceptual understanding.
Activity 33 explains the concept of quantum teleportation using an analogy-based narrative.
This well-structured progression reflects best practices in scaffolding and conceptual coherence. It also includes dedicated poster of scientists contributing to Quantum Science & Technology since its discovery.

4. EDUCATIONAL AND COMMUNICATIVE IMPACT

4.1 Conceptual Understanding

These hands-on practical 33-activities indicates improved conceptual clarity regarding foundational quantum ideas such as wave–particle duality, superposition, and measurement. This align with prior research suggesting that early conceptual exposure supports long-term understanding *(Krijtenburg-Lewerissa et al., 2017).*

4.2 Engagement and Motivation

Hands-on activities will consistently increase student interest and reduce anxiety associated with advanced physics topics. Engagement is further enhanced when activities are framed within real-world technological contexts.

4.3 Skill Development and Scientific Practices

Beyond content knowledge, student's development of observational skills, reasoning, collaboration, and curiosity—core components of scientific literacy *(OECD, 2019).*

4.4 Science Communication Value

The kit functions effectively as a science communication tool, facilitating dialogue between students, educators, and the public in settings such as science fairs and outreach programs *(Falk & Dierking, 2013).*

5. CHALLENGES AND LIMITATIONS

Despite their benefits, hands-on quantum activities present challenges:

- Risk of oversimplification through analogies
- Limited scalability without teacher training
- Need for structured assessment tools
- Difficulty in addressing advanced quantum concepts

Addressing these challenges requires more careful design, facilitator support, and integration with digital resources.

6. FUTURE DIRECTIONS FOR IMPACTFUL QUANTUM EDUCATION

To enhance the effectiveness of hands-on quantum education initiatives, future efforts should consider:

- Digital simulations and hybrid learning models
- Tiered activities for diverse learner levels
- AI-supported adaptive feedback systems
- Longitudinal studies to measure sustained impact
- Professional development for educators

7. CONCLUSION

This review highlights the growing role of experiential, hands-on approaches in making Quantum Science & Technology accessible to school-level learners to undergraduates and the wider general public. The reviewed 33-activity Hands-on Activity Kit on Quantum Science & Technology developed by the Gujarat Council on Science & Technology (GUJCOST) exemplifies how structured, analogy-driven learning can promote quantum literacy, engagement, and positive attitudes toward science.

As quantum technologies increasingly influence society, such educational and communication-focused interventions are essential for fostering informed, curious, and scientifically literate communities, which primarily satisfies the goal of United Nations for the celebrations of International Year of Quantum Science & Technology 2025 (IYQ2025).

ACKNOWLEDGMENTS

The author(s) gratefully acknowledge the support, guidance, and cooperation extended by the institutional authorities, colleagues, and all individuals who contributed, directly or indirectly, to the completion of this review work.

1. Gujarat Council On Science & Technology (GUJCOST), Gandhinagar, Gujarat, India.
2. Shri O.V. Sheth Regional Community Science Centre, Rajkot, Gujarat, India.
3. P.C. Vaidya Astronomical, Astrophysical and Space Research Institute. Centre for Studies and Research of Science in Ancient India. Rajkot, Gujarat, India.
4. H.& H.B. Kotak Institute of Science, Rajkot, Gujarat, India.

REFERENCES

1. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
2. Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher*, 23(7), 5–12.
3. Duit, R. (1991). On the role of analogies and metaphors in learning science. *Science Education*, 75(6), 649–672.
4. Falk, J. H., & Dierking, L. D. (2013). *The museum experience revisited*. Routledge.
5. Fischhoff, B., & Scheufele, D. A. (2013). The science of science communication. *PNAS*, 110(Suppl. 3), 14031–14032.
6. Freeman, S., et al. (2014). Active learning increases student performance in STEM. *PNAS*, 111(23), 8410–8415.
7. Greinert, F., Müller, R., & Bitzenbauer, P. (2023). Quantum literacy: Conceptual foundations and educational implications. *Physics Education*, 58(4).
8. Kohnle, A., et al. (2014). Developing conceptual understanding of quantum mechanics. *European Journal of Physics*, 35(1).
9. Kolb, D. A. (1984). *Experiential learning*. Prentice Hall.
10. Krijtenburg-Lewerissa, K., et al. (2017). Insights into teaching quantum mechanics. *Physical Review Physics Education Research*, 13(1).
11. OECD. (2019). *PISA 2018 science framework*. OECD Publishing.
12. Taber, K. S. (2015). Learning quanta: Barriers to understanding quantum physics. *Physics Education*, 50(2), 159–168.

INORGANIC NONLINEAR OPTICAL CRYSTAL: AMMONIUM DIHYDROGEN PHOSPHATE

Dave Khushi Y*. & A.P. Kochuparampil
Department of Physics, Christ College, Rajkot
*khushidave1125@gmail.com

ABSTRACT: Ammonium dihydrogen phosphate (ADP) is a significant nonlinear optical (NLO) crystal known for its excellent optical transparency and frequency conversion efficiency. Among various crystal growth methods, slow evaporation solution growth and seeded solution growth are preferred for NLO ADP crystals. In this study, ADP crystals were grown using the slow evaporation solution method at room temperature, with variations in water quality. The crystals' growth, structure, and properties were characterized, starting with solubility observations.

KEY WORDS: - Ammonium Dihydrogen Phosphate (ADP), Nonlinear Optical Crystal (NLO), slow solvent evaporation method

INTRODUCTION

Ammonium dihydrogen phosphate (ADP) is an important inorganic nonlinear optical crystal. It is used as primary material due to its properties like optical transparency, dielectric behavior, piezoelectric response, nonlinear optical nature and frequency conversion efficiency.[1] These properties of ADP crystal show its applications in the field of optoelectronics, LASER, photonics, etc. ADP is a typical isomorph of Potassium dihydrogen phosphate (KDP). ADP has higher second and third harmonic generation efficiency thus it is used as alternative of KDP. ADP is grown by slow evaporation solution growth technique, seeded solution growth method, temperature lowering solution growth technique and gel growth method.

OBJECTIVE OF THE STUDY

1. To grow the ADP crystals.
2. To check the quality of water.
3. To find solubility and change in morphology in ADP crystal.

EXPERIMENTAL TECHNIQUE

There are four main techniques for the growth of ADP crystal namely,

1. Slow evaporation solution growth technique
2. Seeded solution growth technique
3. Temperature lowering solution growth technique
4. Gel growth technique

Among all these growth techniques, seeded solution growth method is most preferred. But for basic study and educational purposes, slow evaporation solution growth method is used.

In this research ADP crystal is grown by slow evaporation solution growth method. We have grown ADP crystal in two different solvent such as distilled water and R.O. water. Firstly, 40gm Ammonium dihydrogen phosphate is added to 100ml distilled water. Then this solution was stirred overnight (minimum 8 hours preferred for saturation). The saturated solution was filtered by Watmann filter paper no.1. This filtered solution was kept in a beaker in a dust free atmosphere with a porous lid to control evaporation process. This process is repeated for R.O. water. Here solvent is R.O. water instead of distilled water. In this case we have taken 35gm Ammonium dihydrogen phosphate in 100ml R.O. water. During the whole process, room temperature was 21ºC -26ºC. After 15-20 days, good quality, transparent, colourless crystals were harvested. Fig. 1 shows the harvested crystals of pure ADP (Distilled water and R.O. water respectively).

Fig. 1 (a) ADP crystal (Distilled Water)

Fig. 1 (b) ADP crystal (R.O. Water)

EFFECT OF TDS ON ADP CRYSTAL GROWTH, SOLUBILITY, AND WATER QUALITY

To check quality of water (both Distilled and R.O. water) we have taken this reference data and will compare it with our measured data.

TDS Range (mg/L or ppm)	Crystallization / Crystal Formation	Solubility Characteristics	Water Quality Classification
< 50	Negligible crystallization; minerals largely absent	Very high solubility; minimal ionic content	Very low mineral water (ultrapure / demineralized)
50 – 150	No visible crystallization under normal conditions	High solubility; stable ionic balance	Excellent quality drinking water
150 – 300	Low crystallization potential	High solubility of most common salts	Good quality drinking water
300 – 500	Slight crystal formation possible upon evaporation	Moderate solubility; increased ionic strength	Acceptable drinking water (upper recommended limit)
500 – 1,000	Noticeable crystallization during evaporation	Reduced solubility for some salts (e.g., $CaCO_3$)	Poor drinking water; marginal for domestic use
1,000 – 3,000	Significant crystal formation likely	Limited solubility; precipitation of salts common	Saline water; unsuitable for drinking
3,000 – 10,000	Extensive crystallization	Low solubility; high precipitation tendency	Highly saline / brackish water
> 10,000	Rapid and extensive crystal formation	Very low effective solubility	Seawater / hypersaline water

Table 1: Effects of TDS on ADP crystal growth, solubility and water quality[2]

RESULT AND DISCUSSION: -

	R.O. Water	Distilled Water
1. Quality of water	High purity water	Ultra pure
2. Solubility of ADP solution	35gm/100ml (21°C)	40gm/100ml (20°C)
3. TDS Value	47 ppm	6 ppm

Single crystals of ADP are successfully grown under controlled evaporation condition by slow evaporation solution growth method. Grown crystals are pure (undoped) and have well-defined morphology. These ADP crystals are transparent and does not have any macroscopic defect.

MORPHOLOGICAL CHANGES

To analyze crystal's morphology without any instrumental characterizations we can use methods such as macroscopic geometric measurements and optical observation.[3]
Macroscopic geometric measurements,

1) For ADP crystal in Distilled water,
   Length (L) = 8mm
   Width (W) = 4mm
   Thickness (T) = 2.5mm

$$Aspect\ ratio = \frac{L}{W}$$
$$= \frac{8}{4}$$

$$\boxed{Aspect\ ratio = 2}$$

$$Morphological\ index = \frac{L}{(W+T)/2}$$
$$= \frac{8}{(4+2.5)/2}$$
$$= \frac{8}{6.5/2}$$
$$= \frac{8}{3.25}$$

$$\boxed{Morphological\ index = 2.46}$$

2) For ADP crystal in R.O. water,
   Length (L) = 13mm
   Width (W) = 5mm
   Thickness (T) = 4mm

$$Aspect\ ratio = \frac{L}{W}$$
$$= \frac{13}{5}$$

$$\boxed{\text{Aspect ratio} = 2.6}$$

$$Morphological\ index = \frac{L}{(W+T)/2}$$

$$= \frac{13}{(5+4)/2}$$

$$= \frac{13}{9/2}$$

$$= \frac{13}{4.5}$$

$$\boxed{\text{Morphological index} = 2.88}$$

From the above calculation we can differentiate both ADP crystals. (Distilled water and R.O. water respectively)

Sample	Length (L) (mm)	Width (W) (mm)	Thickness (T) (mm)	Aspect Ratio (L/W)	Morphological Index (L/(W+T)/2)
ADP (Distilled water)	8	4	2.5	2	2.46
ADP (R.O. water)	13	5	4	2.6	2.88

OVERALL INTERPRETATION

- Morphological change can be quantitatively differentiated by measuring variation in aspect ratio and morphological index of the grown crystal.
- Significant morphological change in grown ADP crystals can be described without any advanced instrumental characterization.

FUTURE WORK TO BE DONE…

In future, we will do characterization of ADP crystals for basic structural analysis, presence of functional group, determination of optical transparency and to determine frequency conversion efficiency. [4]

The characterizations that will be performed are…

1. XRD - X ray Diffraction technique
2. FTIR - Fourier Transform Infrared spectroscopy
3. UV-Vis spectroscopy - Ultra Violet Visible spectroscopy
4. SHG - Second Harmonic Generation spectroscopy

ADVANTAGES OF ADP CRYSTAL

Larger nonlinear optical (NLO) coefficient
Higher laser damage threshold
Shorter UV transmission cut-off
Faster growth speed
Lower production cost

APPLICATIONS OF ADP CRYSTAL

Frequency conversion (SHG)
Optical switches in inertial confinement fusion[5]
Acousto-optical devices
In photonic devices
LASER optics

CONCLUSION

Ammonium dihydrogen phosphate is a crucial element in the class of inorganic nonlinear optical crystal. It is used in the field of opto-electronics, LASER, photonic devices, etc. These applications are due to its properties such as high transparency and second and third harmonic frequency generation efficiency. From the recent data it may also generates frequency of fourth harmonic[6]. These ADP crystals are grown by slow evaporation solution growth method under controlled evaporation conditions. There are two types of solvents used such as distilled water and R.O. water. Morphological changes are differentiated in both the cases by macroscopic geometric measurements and optical observations without any instrumental characterizations.

REFERENCES

[1] P. Kochuparampil, M. J. Joshi, Mod Phys Lett B, 31(2017) 1-21.
[2] D. N. Nikogosyan, Nonlinear Optical Crystals: A Complete Survey (Springer, 2001).
[3] I. Sunagawa, *Crystals: Growth, Morphology, and Perfection*. Cambridge, U.K.: Cambridge University Press, 2005.
[4] Ramachandra Raja et al., Growth and characterization of ADP crystals, Optical Materials
[5] Nalwa, H. S., Nonlinear Optics of Organic and Inorganic Materials
[6] A. A. Chernov, *Modern Crystallography III: Crystal Growth*. Berlin, Germany: Springer-Verlag, 1984.

INORGANIC NONLINEAR OPTICAL CRYSTAL: POTASSIUM DIHYDROGEN PHOSPHATE

Dave Khyati Y.* **& A.P. Kochuparampil**
Department of Physics, Christ College, Rajkot.
**khyatidave2571@gmail.com*

ABSTRACT: Inorganic nonlinear optical (NLO) crystals are vital in modern photonic and optoelectronic applications due to their ability to manipulate light through nonlinear behaviour. Potassium Dihydrogen Phosphate (KDP) is a notable example of an inorganic NLO material, boasting excellent optical transparency, high laser damage threshold, and efficient second harmonic generation (SHG). Although seeded solution growth with controlled temperature lowering is a preferred method for growing KDP crystals, the slow evaporation solvent method is employed in this work for its simplicity and educational value in lab setting.

KEY WORDS:- Nonlinear optical crystal (NLO), Potassium Dihydrogen Phosphate (KDP), slow evaporation solution growth method

INTRODUCTION

Crystals are classified into two types due their optical phenomena. They are called linear optical crystals and nonlinear optical crystals. This classification is based on crystal's response to the applied electromagnetic field i.e. light waves. Linear optical crystals are independent of light intensity while nonlinear optical crystals are dependent on light intensity and possess efficiency of second harmonic generation. Thus nonlinear optical crystals are more useful than linear optical crystals. Inorganic nonlinear optical crystals are more preferred due to high thermal, physical and chemical stability than inorganic nonlinear optical crystals. Potassium Dihydrogen Phosphate (KDP) crystal is an important example in NLO. [1] KDP has high transparency, good mechanical strength, water solubility and high laser damage threshold. Due to these properties of KDP, it is used in opto-electronics and LASERs especially in ND-YAG lasers.

OBJECTIVE OF THE STUDY

1. To grow inorganic nonlinear optical crystal KDP.
2. To check quality of water.
3. To find solubility and change in morphology in KDP crystal.

EXPERIMENTAL TECHNIQUE

There are five techniques for the growth of inorganic nonlinear optical crystal. [2]

1. Slow Evaporation Solution Growth Technique (SEST)
2. Temperature Reduction (Slow Cooling) Method
3. Seeded Solution Growth Technique
4. Rapid Growth Technique (RGT)
5. Gel Growth Method

For education and lab purposes slow evaporation solution growth method is more preferred.

In this research KDP crystal is grown by slow evaporation solution growth method. We have grown KDP crystal in two different solvent such as distilled water and R.O. water. Firstly, 25gm Potassium dihydrogen phosphate is added to 100ml distilled water. Then stirred this solution overnight (minimum 8 hours are preferred for saturation). The saturated solution was filtered by Watmann filter paper no.1. This filtered solution was kept in a beaker in a dust free atmosphere with a porous lid to control evaporation process. All this process is repeated for R.O. water. Here solvent is R.O. water instead of distilled water. During the whole process temperature room temperature was 21ºC - 26ºC.After 15-20 days good quality, transparent, colourless crystals were harvested. Fig. 1 shows the harvested crystals of pure KDP (Distilled water and R.O. water respectively).

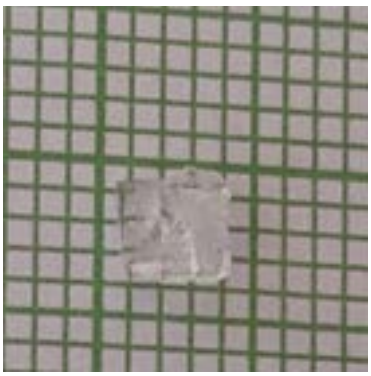


Fig 1 (a) KDP crystal (Distilled water)

Fig 1 (b) KDP crystal (R.O. water)

EFFECT OF TDS ON KDP CRYSTAL GROWTH, SOLUBILITY, AND WATER QUALITY

The table is a qualitative summary derived from established experimental observations reported in the Literature. [3]

TDS Level in Water	Water Quality	Effect on KDP Crystal Growth	Resulting Crystal Quality
< 10 ppm	Ultra-pure / Deionized water	Controlled nucleation and steady growth rate	Highly transparent, defect-free single crystals
10–50 ppm	High-purity distilled water	Slight variation in solubility due to trace ions	Good optical quality with negligible defects
50–200 ppm	Acceptable laboratory water	Increased spontaneous nucleation and uneven growth	Minor inclusions, striations may appear
200–500 ppm	Poor-quality water	Irregular growth rate and face development	Reduced transparency, lattice defects
> 500 ppm	Unsuitable for crystal growth	Uncontrolled nucleation, inhibited crystal growth	Polycrystalline growth, cracks, inclusions

Table – 1 Effect of TDS on KDP crystal growth, solubility and water quality

RESULT AND DISCUSSION

	Distilled water	R.O. water
1. Quality of water	Ultra-pure water	High-purity water
2. Solubility of KDP solution	25gm/100ml (20°C)	20gm/100ml (24°C)
3. TDS value	6 ppm	47 ppm

KDP single crystals are grown by slow evaporation solution growth method. These crystals are transparent and of well-defined shape. They have not any visible defects. Morphological change in KDP crystal in both solvent distilled water and R.O. water is differentiated by some quantitative analysis.

MORPHOLOGICAL CHANGES

For the analysis of grown KDP crystals, we do not use any instrumental characterization techniques instead we have use methods like macroscopic geometric measurements. [4]

Macroscopic geometric measurements,

1) For KDP crystal in Distilled water,

Length (L) = 3.5mm
Width (W) = 4mm
Thickness (T) = 1mm

$$Aspect\ ratio = \frac{L}{W}$$

$$= \frac{3.5}{4}$$

$$\boxed{Aspect\ ratio = 0.875}$$

$$Morphological\ index = \frac{L}{(W+T)/2}$$

$$= \frac{3.5}{(4+1)/2}$$

$$= \frac{3.5}{5/2}$$

$$= \frac{3.5}{2.5}$$

$$\boxed{Morphological\ index = 1.4}$$

2) For KDP crystal in R.O. water,

Length (L) = 16mm
Width (W) = 9mm
Thickness (T) = 5mm

$$Aspect\ ratio = \frac{L}{W}$$

$$= \frac{16}{9}$$

$$\boxed{\text{Aspect ratio} = 2.28}$$

$$Morphological\ index = \frac{L}{(W+T)/2}$$

$$= \frac{16}{(9+5)/2}$$

$$= \frac{16}{14/2}$$

$$= \frac{16}{7}$$

$$\boxed{\text{Morphological index} = 2.28}$$

From the geometrical measurements, the obtained data and values of aspect ratio and morphological index shows difference between two types of KDP crystals (based on solvent i.e. distilled water and R.O. water. Below table shows the obtained data.

Sample	**Length (L) (mm)**	**Width (W) (mm)**	**Thickness (T) (mm)**	**Aspect Ratio (L/W)**	**Morphological Index (L/(W+T)/2)**
KDP (Distilled water)	3.5	4	1	0.875	1.4
KDP (R.O. water)	16	9	5	1.8	2.28

OVERALL INTERPRETATION

- In grown KDP crystals by the method of macroscopic geometric measurements aspect ratio and morphological index is obtained which shows morphological changes in both types of crystals.
- This kind of analysis is done without any advanced instrumental characterization.

FUTURE WORK TO BE DONE…

In future, for advanced analysis of grown KDP crystals, we will do characterization. [5]

1. PXRD – Powder X ray Diffraction technique - Confirms phase purity and crystallinity
2. FTIR - Fourier Transform Infrared spectroscopy - Confirms functional groups (PO_4, OH)
3. UV-Vis spectroscopy - Confirms optical transparency and cut-off wavelength
4. SHG - Second Harmonic Generation spectroscopy - Confirms nonlinear optical activity

ADVANTAGES OF KDP CRYSTAL

Excellent nonlinear optical (NLO) properties
High optical transparency
High laser damage threshold
Good thermal stability
Electro-optic properties

APPLICATIONS OF KDP CRYSTAL

Frequency doubling (second harmonic generation)
High-power laser systems
Electro-optic modulators
Optical switching and communication devices
Laser range finding and LIDAR systems
Optical data storage and processing

CONCLUSION

Potassium dihydrogen phosphate (KDP) is a primary element in the class of nonlinear optical crystal (NLO). Its importance is due to the property of frequency conversion efficiency, high transparency, efficiency of second harmonic generation, high chemical and physical strength. Due to these properties KDP crystals are used in the field of photonics, opto-electronics, LASERs, LASER range finding and LIDAR system. KDP crystals are also important in optical data storage and processing as it has nonlinear optical property. The growth method of KDP crystals are slow evaporation solution growth method, seeded solution growth method, temperature reduction method, rapid growth technique (RGT) and gel growth method. In this research we have used slow evaporation solution growth method. It is preferred in basic analysis, educational and lab purposes over other techniques. The grown KDP crystals are transparent and have well-defined structure. Morphological change in KDP crystal is described by macroscopic geometric measurements without any advanced instrumental characterization.

REFERENCES

[1] A. P. Kochuparampil, M. J. Joshi, Mod Phys Lett B, 31(2017) 1-21.

[2] L. N. Rashkovich, KDP-Family Single Crystal (IOP Publishing Ltd, 1991).

[3] N. P. Zaitseva, J. J. De Yoreo, M. R. Dehaven, R. L. Vital, K. E. Montgomery, M. Richardson, and L. J. Atherton, "Rapid growth of large-scale (40–55 cm) KH2PO4 crystals," J. Cryst. Growth 180(2), 255–262 (1997).

[4] Sunagawa, *Crystals: Growth, Morphology, and Perfection*. Cambridge, U.K.: Cambridge University Press, 2005.

[5] N. Zaitseva and L. Carman, "Rapid growth of KDP-type crystals," Prog: Cryst. Growth Charact. 43(1), 1–118 (2001).

PRESOLAR SILICATE GRAINS: ISOTOPIC SIGNATURES, MINERALOGY AND ASTROPHYSICAL IMPLICATIONS

Manish N. Sanghani[1*] and Kuljeet Kaur Marhas[2]
[1]*P.C. Vaidya Astronomical, Astrophysical and Space Science Research Centre, Shri O.V. Sheth Regional Community Science Centre, Rajkot, India*
[2]*Planetary Science Division, Physical Research Laboratory, Ahmedabad, India*
*Corresponding Author Email: kevalgyan.manish@gmail.com

ABSTRACT: Presolar grains formed in the outflows of Red Giant Branch (RGB) and Asymptotic Giant Branch (AGB) stars, as well as in novae and supernova explosions, prior to the birth of the Sun. After their formation, they were incorporated into the molecular cloud that collapsed to form the solar system and survived subsequent energetic processing allowing their preservation in meteorites. Identified by their highly anomalous isotopic compositions, presolar grains—including SiC, graphite, nanodiamonds, Si_3N_4, oxides, and silicates record stellar nucleosynthesis, dust formation environments, and the chemical evolution of the Milky Way. Presolar silicates are identified in situ using NanoSIMS (nano Secondary Ion Mass Spectrometer) imaging and are significantly more abundant than other presolar grain types, occurring at tens to hundreds of ppm in meteorites and reaching peak abundances in cometary interplanetary dust particles (IDPs). They are predominantly ferromagnesian, commonly amorphous, and compositionally heterogeneous, reflecting formation under rapidly changing, non-equilibrium conditions in stellar outflows from low-mass RGB and AGB stars. In contrast, some crystalline silicates originating from type-II supernovae indicate formation near-equilibrium conditions. Presolar silicates commonly experienced subsequent modification through thermal metamorphism and aqueous alteration in the solar nebula and on parent bodies. The recent detection of fluorine in presolar silicates further highlights their value as tracers of both stellar environments and secondary alteration processes.

KEY WORDS: Asymptotic Giant Branch, Supernova, Silicate Grains, Meteorites, Isotopic Characterization, Stellar Nucleosynthesis, Condensation, Alteration, Parent Body Processing, Galactic Chemical Evolution

1. INTRODUCTION

1.1 Early Isotopic Evidence for Presolar Matter

Primitive meteorites, especially the least altered carbonaceous chondrites, are studied extensively because they preserve materials that experienced minimal processing since Solar System formation and thus retain records of early solar and pre-solar history. Early high-precision isotopic analyses of these meteorites were originally aimed at understanding Solar System formation under the prevailing assumption of a largely well-mixed solar nebula (Edward Anders, 1971; Gerald J. Wasserburg et al., 1977). This assumption was challenged when isotopic measurements revealed deviations from solar compositions far exceeded those expected from mass-dependent fractionation or the decay of short-lived radionuclides. Particularly compelling evidence came from noble gas studies of carbonaceous chondrites,, as noble gases are chemically inert and thus preserve distinct isotopic signatures of their sources. These studies identified isotopically exotic components such as Ne-E, strongly enriched in ^{22}Ne, and Xe-HL, characterized by correlated excesses in both light (^{124}Xe, ^{126}Xe) and heavy (^{134}Xe, ^{136}Xe) xenon isotopes (John H. Reynolds, 1964; Douglas C. Black and Robert Pepin, 1969; Roy S. Lewis et al., 1975). The distinctive isotopic pattern of Xe-HL was subsequently linked to p- and r-process nucleosynthesis, implicating supernovae sources (Donald D. Clayton, 1975). Large isotopic anomalies were later identified in light elements, including carbon, nitrogen, and hydrogen. Analyses of acid-resistant residues from meteorites such as Murchison revealed isotopic ratios indicative of stellar nucleosynthesis and interstellar inheritance (Halbout et al., 1986; Lee, 1988; Epstein et al., 1987). A critical insight from these studies was that the extreme isotopic signatures were hosted by rare

components present at parts-per-million levels or less, whereas bulk meteorites exhibited only modest deviations from solar compositions. This observation implied that presolar material survived Solar System formation as physically discrete carriers rather than as homogenized within the solar nebula (Anders and Zinner, 1993). Collectively, these observations provided strong evidence for the survival of presolar material in solid form within primitive meteorites.

This realization—that rare carriers with extreme isotopic compositions dominated bulk isotopic anomalies observed in meteorites—shifted experimental efforts toward identifying and isolating these components. Because bulk meteorites exhibited only small deviations from solar compositions, it became evident that the anomalous carriers must be physically robust phases capable of surviving both harsh stellar environments and subsequent Solar System processing. Early efforts therefore focused on chemically resistant and refractory materials, leading to the development of aggressive chemical extraction techniques to dissolve the dominant silicate matrix and concentrate the anomalous components. Among these materials, silicon carbide emerged as a prime candidate due to its high thermal stability, chemical durability, and association with anomalous noble gas components in meteoritic residues (Lewis et al., 1987). These advances ultimately enabled the isolation of individual presolar grains, beginning with SiC, and provided the first opportunity to directly analyze solid samples that formed in stellar environments prior to the birth of the Solar System.

1.2 Isolation of Presolar Grains from Meteorites: Carbon-Rich and Refractory Phases

Sequential acid dissolution techniques were developed to remove the dominant silicate matrix of primitive carbonaceous chondrites, such as Allende and Murchison, while concentrating chemically resistant residues enriched in isotopically anomalous components (Lewis et al. 1975; Anders & Zinner 1993). Application of these methods led to the first direct identification of presolar grains, beginning with nanodiamonds, typically a few nanometers in size and present at abundances of up to several hundred parts per million, which host anomalous noble gas components. Subsequent advances enabled the isolation of larger individual grains, most notably silicon carbide and graphite, with typical sizes ranging from submicrometer to several micrometers and abundances of a few parts per million or less in primitive meteorites (Lewis et al. 1987; Amari et al. 1990, 1994). These grains exhibit extreme isotopic compositions, including $^{12}C/^{13}C$ and $^{14}N/^{15}N$ ratios that differ from solar values by factors of several to orders of magnitude, as well as distinctive silicon isotopic signatures reflecting stellar nucleosynthesis. The preservation of such large isotopic anomalies provided unambiguous evidence for grain formation in stellar environments, including asymptotic giant branch stars and supernovae, and thereby established presolar grains as direct physical samples of stellar nucleosynthesis preserved within meteorites (Zinner 1998; Anders & Zinner 1993).

1.3 Presolar Silicates: A Delayed Discovery

Astronomical observations have long demonstrated that silicate dust is a dominant component of circumstellar outflows from oxygen-rich evolved stars, based on characteristic infrared emission features attributed to amorphous and crystalline silicates (Waters et al. 1996; Demyk et al. 2000). Despite this expectation, presolar silicates were not identified in meteoritic materials until the early 2000s, nearly two decades after the discovery of carbon-rich and refractory presolar grains. This delay reflects fundamental methodological limitations: silicates are readily destroyed during the acid dissolution procedures used to isolate earlier presolar phases and are mineralogically indistinguishable from the abundant silicates formed within the Solar System, preventing their recovery as a distinct chemical residue. Consequently, the identification of presolar silicates required a shift from bulk chemical extraction to high-spatial-resolution, in situ isotopic analysis. The first definitive detections were achieved through oxygen isotopic imaging using NanoSIMS, initially in interplanetary dust particles (Messenger et al. 2003) and shortly thereafter in primitive meteorites, including Acfer 094 and other carbonaceous chondrites (Mostefaoui & Hoppe 2004; Nguyen & Zinner 2004; Nagashima et al.

2004). These discoveries completed the major inventory of presolar grain types and enabled direct laboratory investigation of silicate dust formed in oxygen-rich stellar environments.

1.4 In-Situ Identification of Presolar Silicates Using NanoSIMS

Early efforts to identify presolar silicates using ion microprobe techniques included isotopic imaging with large-radius instruments such as the Cameca ims-1270 equipped with specialized detector systems (Yurimoto et al. 2003; Nagashima et al. 2004, 2005). Although these studies demonstrated the feasibility of in situ isotopic mapping, the routine identification of presolar silicates became possible only with the advent of the NanoSIMS 50 and 50L instruments, which combine high spatial resolution with high secondary ion sensitivity. In typical analyses, a finely focused Cs^+ primary ion beam with a diameter of ~100 nm is rastered over areas of approximately 10 × 10 or 20 × 20 µm² to survey large regions for rare isotopically anomalous grains. Prior to imaging, the analysis area is commonly pre-sputtered using a higher beam current to remove surface contamination and to enhance negative secondary ion yields through Cs implantation. The simultaneous collection of negative secondary ions of the three oxygen isotopes ($^{16}O^-$, $^{17}O^-$, $^{18}O^-$) enables the generation of isotope ratio images, allowing presolar silicates to be identified based on their anomalous oxygen isotopic compositions. Oxygen isotope imaging proved decisive for presolar silicate discovery because oxygen isotopic heterogeneity provides the strongest contrast between stellar condensates and solar-system silicates (Nguyen et al. 2007; Vollmer et al. 2009a).

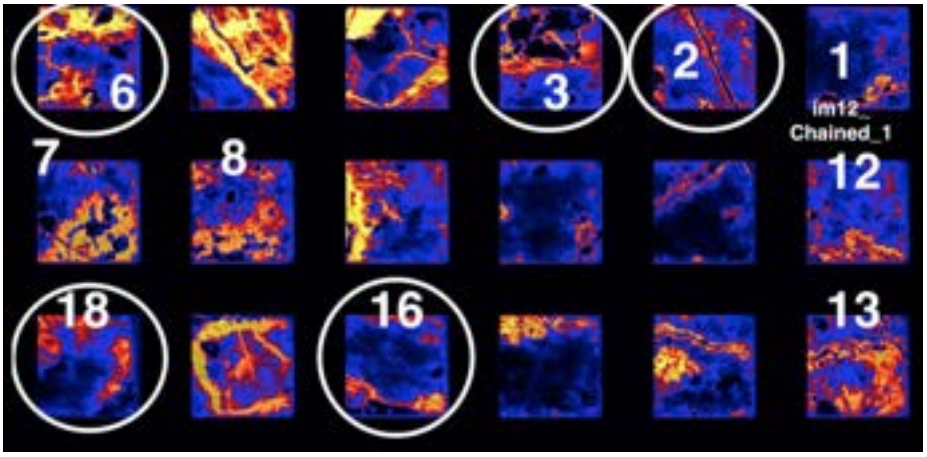


Fig.1. ^{16}O map of Isheyevo meteorite sample created by NanoSIMS 50L.

The flexibility of NanoSIMS isotopic imaging has enabled presolar silicate searches across a wide range of primitive solar-system materials prepared using diverse mounting strategies. Fine-grained matrix regions in thin sections of unequilibrated carbonaceous and ordinary chondrites have been extensively analyzed, revealing presolar silicates preserved within their petrographic context (Leitner et al. 2012a; Nittler et al. 2018). Complementary studies have examined physically disaggregated grain-size separates mounted on high-purity Au substrates, an approach that minimizes matrix effects and facilitates efficient raster imaging (Nguyen 2005). Interplanetary dust particles and Antarctic micrometeorites have likewise been investigated using NanoSIMS, either pressed into Au, mounted on TEM grids, or deposited on Si wafers, thereby extending presolar silicate searches to materials that experienced minimal parent-body processing (Busemann et al. 2009; Yada et al. 2008). In addition, presolar silicates have been identified in impact residues preserved in Al foils returned by the Stardust mission to comet Wild 2, demonstrating that circumstellar silicate grains were incorporated into cometary materials and survived hypervelocity capture (McKeegan et al. 2006; Stadermann et al. 2008). Beyond oxygen isotopes, NanoSIMS analyses commonly include the detection of secondary ions such as $^{28}Si^-$, $^{24}Mg^{16}O^-$, and $^{27}Al^{16}O^-$ on the NanoSIMS 50, while the expanded detector array of the

NanoSIMS 50L allows the simultaneous measurement of all three silicon isotopes together with oxygen, enabling more comprehensive isotopic characterization of presolar silicate grains (Nguyen et al. 2010a).

1.5 Beyond Oxygen Isotopes: Multi-Element Constraints on Presolar Silicates

Beyond oxygen isotopes, NanoSIMS has enabled measurements of additional isotopic systems in presolar silicates that provide complementary constraints on stellar sources and grain histories. Silicon isotopes were the first to be explored in detail, as they can be analyzed with relatively high spatial resolution using a Cs^+ primary beam. With the expanded magnet and detector configuration of the NanoSIMS 50L, silicon isotopes can be measured either sequentially following oxygen isotope surveys or simultaneously with oxygen, thereby improving efficiency and reducing analytical bias (Mostefaoui & Hoppe 2004; Nguyen et al. 2010a). Although Si isotopic anomalies in presolar silicates are generally smaller than those observed in presolar SiC, they nevertheless preserve diagnostically important information related to stellar nucleosynthesis and galactic chemical evolution (Hoppe et al. 2017; Nittler et al. 2018).

Magnesium and iron isotopic measurements in presolar silicates are more challenging, largely due to signal dilution caused by the relatively large beam sizes associated with O^- primary ion analyses. Modeling studies demonstrated that dilution effects are strongest for the smallest grains and can significantly mask intrinsic isotopic signatures (Nguyen et al. 2007). To mitigate these limitations, focused ion beam (FIB) techniques have been used to isolate individual grains prior to NanoSIMS analysis (Kodolányi et al. 2014), while Fe isotopes have also been measured as oxide ions using a Cs^+ primary beam to achieve improved spatial resolution despite lower ion yields (Vollmer & Hoppe 2010; Ong & Floss 2015). Together, these developments have extended NanoSIMS from a tool primarily used for oxygen-based discovery to a versatile platform for multi-isotopic characterization of presolar silicates.

2. ISOTOPIC COMPOSITIONS

2.1 Oxygen Isotopic Compositions and Classification of Presolar Silicates

Oxygen isotopic compositions provide the primary basis for identifying and classifying presolar silicate grains and for linking them to their stellar sources. As a major constituent of silicates, oxygen exhibits large isotopic variations produced by stellar nucleosynthesis, making oxygen isotope ratios a sensitive discriminator between circumstellar grains and isotopically homogeneous solar-system materials. Presolar silicates are commonly classified into distinct groups based on their $^{17}O/^{16}O$ and $^{18}O/^{16}O$ ratios, following a framework originally developed for presolar oxides and subsequently extended to silicates (Nittler et al. 1997; Messenger et al. 2003; Nittler et al. 2018). These groups reflect contributions from diverse stellar environments, including asymptotic giant branch (AGB) stars with dredge-up and extra mixing, explosive nucleosynthesis in supernovae, and, in some cases, nova outbursts (José & Hernanz 1998).

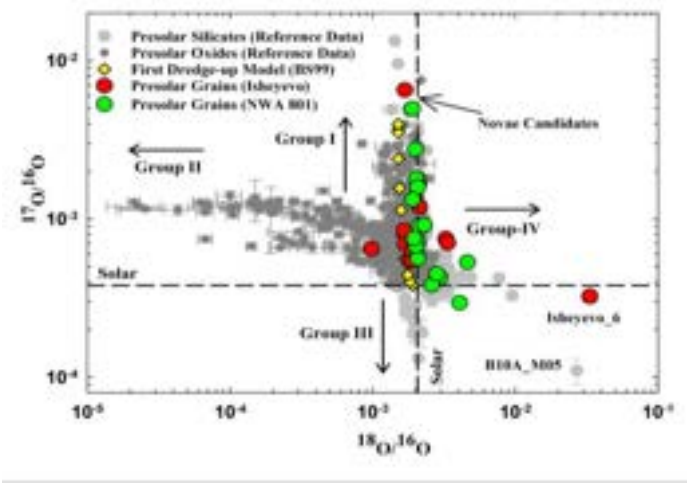


Fig.2. Oxygen isotopic plot of presolar oxide and silicate grains. Image Credit: Sanghani et al. 2021

2.1.1 Group 1 Silicates: AGB Stars and Beyond

Group 1 presolar silicates are characterized by enrichments in ^{17}O relative to solar composition, with near-solar or mildly depleted $^{18}O/^{16}O$ ratios. These isotopic signatures are broadly consistent with hydrogen burning via the CNO cycle and subsequent first and second dredge-up episodes in low-mass asymptotic giant branch (AGB) stars (Boothroyd & Sackmann 1999; Nittler et al. 1997). Observationally, Group 1 grains dominate the presolar silicate population, indicating that low-mass AGB stars were a major source of circumstellar silicate dust incorporated into primitive solar-system materials (Nittler et al. 2008; Floss & Haenecour 2016).

A subset of Group 1 presolar silicates exhibits exceptionally high $^{17}O/^{16}O$ ratios that approach or exceed the upper limits predicted for low-mass AGB stars undergoing standard dredge-up. Such "extreme" Group 1 compositions were first recognized in presolar oxides and later extended to silicates through systematic NanoSIMS surveys of primitive meteorites and interplanetary dust particles (Vollmer et al. 2007; Gyngard et al. 2010; Leitner et al. 2012a). These grains occupy a region of oxygen isotope space that overlaps with predictions for intermediate-mass AGB stars and super-AGB stars, but in some cases also intersects with compositions expected from explosive hydrogen burning in novae or from specific supernova mixing scenarios (Nittler & Hoppe 2005; José et al. 2004). Consequently, extreme Group 1 presolar silicates highlight the need for complementary isotopic constraints, particularly from Mg and Si, to reliably resolve their stellar origins. As discussed by Nguyen & Messenger (2014), nova-related mixing scenarios may involve incorporation of isotopically normal material from the binary companion star or grain formation in a companion star polluted by nova ejecta, while incomplete mixing of CNO-processed material is less favorable for silicate formation because silicates condense after carbonaceous grains (Amari et al. 2001; Helton et al. 2011).

2.1.2 Group 2 Silicates: Extra Mixing Processes

Group 2 presolar silicates are defined by enrichments in ^{17}O accompanied by pronounced depletions in ^{18}O relative to solar composition. Such isotopic signatures cannot be reproduced by standard dredge-up processes and instead require additional non-convective mixing mechanisms, commonly referred to as cool bottom processing, which transport envelope material to regions near the hydrogen-burning shell where ^{18}O is efficiently destroyed (Wasserburg et al. 1995; Boothroyd & Sackmann 1999).

Presolar silicates with Group 2 compositions are comparatively rare, but their isotopic signatures provide some of the strongest meteoritic evidence for extra mixing in low-mass AGB stars (Nittler et al. 2008). Recent surveys, including systematic searches in metal-rich carbonaceous chondrites, confirm the scarcity of Group 2 silicates while demonstrating that when present, their oxygen isotopic

compositions are robust and difficult to reconcile with alternative stellar sources (Sanghani et al. 2021). Nevertheless, recent modeling and observational work highlights that uncertainties in stellar reaction rates and dilution effects necessitate complementary isotopic constraints, particularly from Mg and Al, to fully characterize the mixing processes involved (Leitner & Hoppe 2019; Nittler 2019).

2.1.3 Group 3 Silicates: Low-Metallicity and Rare Contributors

Group 3 presolar silicates exhibit depletions in both ^{17}O and ^{18}O relative to solar values and are commonly interpreted as originating from stars with subsolar metallicities, reflecting the effects of galactic chemical evolution on initial oxygen isotopic compositions (Nittler et al. 1997). These grains are rare among presolar silicates, and in some meteorite reservoirs they may be entirely absent.

The scarcity or absence of Group 3 silicates in several well-studied samples suggests that either dust production in low-metallicity oxygen-rich stars was inefficient or that such grains were preferentially destroyed during interstellar or parent-body processing (Nittler et al. 2008). Recent NanoSIMS studies reinforce the view that oxygen isotopes alone are insufficient to distinguish between low-metallicity AGB and supernova origins for some Group 3-like compositions, further motivating the use of additional isotopic systems to resolve these ambiguities (Leitner & Hoppe 2019; Nittler 2019).

2.1.4 Group 4 Silicates: Type-II Supernova Mixing

Group 4 presolar silicates are defined by enrichments in ^{18}O relative to solar composition and are most plausibly attributed to supernova origins. Such isotopic signatures cannot be produced in asymptotic giant branch stars and instead require nucleosynthetic contributions from massive stars undergoing core-collapse. Modern supernova models demonstrate that the production and distribution of oxygen isotopes depend sensitively on progenitor mass and metallicity, leading to substantial variation in the isotopic compositions of different ejecta components (Rauscher et al. 2002).

NanoSIMS studies show that Group 4 presolar silicates span a wide range of ^{18}O enrichments and frequently overlap oxygen-isotopic fields predicted by different supernova mixing scenarios, highlighting the heterogeneous nature of supernova ejecta (Messenger et al. 2005). Detailed analyses further demonstrate that oxygen isotopic compositions alone do not uniquely constrain the nature of the supernova progenitor or the relative contributions of different ejecta components, as similar oxygen signatures can arise from distinct combinations of model parameters and mixing proportions (Nguyen & Messenger 2014, Sanghani et al. 2021).

2.2 Silicon Isotopes

Silicon isotopic compositions of presolar silicates from primitive meteorites, interplanetary dust particles (IDPs), and Antarctic micrometeorites (AMMs) have been reported as δ values in numerous studies (Mostefaoui & Hoppe 2004; Nguyen et al. 2007, 2010a; Yada et al. 2008; Vollmer et al. 2008; Leitner et al. 2012b; Nguyen & Messenger 2014), with a more recent synthesis provided by Hoppe et al. (2018). In contrast to carbon-rich presolar grains, silicon isotopes in presolar silicates are not expected to record nucleosynthetic enrichments produced during late stellar evolution. During the evolution of low-mass asymptotic giant branch (AGB) stars, enrichments in the heavy silicon isotopes (^{29}Si and ^{30}Si) occur only after the stellar envelope becomes carbon rich (Lugaro et al. 1999; Zinner et al. 2006), whereas silicates and oxides condense earlier under oxygen-rich conditions. As a result, silicon isotopic compositions of presolar silicates primarily reflect the effects of galactic chemical evolution (GCE) rather than in situ stellar nucleosynthesis.

Consistent with this expectation, the majority of Group 1 and Group 2 presolar silicates analyzed to date exhibit silicon isotopic compositions that are close to solar within analytical uncertainties. These grains closely resemble the “category B” Group 1 silicates identified by Hoppe et al. (2018), whose near-normal silicon isotopic compositions were attributed to isotope exchange processes occurring

either in the interstellar medium or on meteorite parent bodies. Such post-condensation processing may therefore partially or completely obscure original GCE signatures in many presolar silicates, complicating direct interpretation of their silicon isotopic compositions.

Despite this overall trend, a small number of presolar silicates show resolvable silicon isotopic anomalies. Several Group 1 grains exhibit >2σ enrichments in one or both heavy silicon isotopes, plotting along the best-fit correlation line defined by Group 1 silicates in silicon three-isotope space. These correlated ^{29}Si and ^{30}Si excesses, together with their oxygen isotopic compositions, are consistent with formation from stellar sources whose initial compositions were shaped by GCE, rather than by dredge-up–related silicon nucleosynthesis (Hoppe et al. 2018). In contrast, rare "extreme" Group 1 silicates characterized by large ^{17}O enrichments also show significant silicon isotopic anomalies that cannot be readily explained by GCE alone; instead, such grains may be consistent with predictions from ONe nova models, highlighting the diversity of stellar sources contributing to the presolar silicate inventory.

Silicon isotopes also provide important constraints on the origins of Group 4 presolar silicates. While most Group 4 grains show near-solar silicon isotopic compositions, some exhibit measurable depletions in ^{29}Si that can be reproduced by partial mixing of supernova ejecta. These observations support interpretations based on supernova mixing scenarios and demonstrate that, when combined with oxygen isotopes, silicon isotopic data offer critical leverage for distinguishing between competing stellar source models. Overall, silicon isotopes serve as a key complementary tracer to oxygen isotopes, recording large-scale galactic chemical evolution while, in rare cases, preserving signatures of explosive nucleosynthesis.

2.3 Magnesium Isotopes

Magnesium isotopic compositions provide critical constraints on the stellar origins and chemical histories of presolar silicates, complementing information derived from oxygen and silicon isotopes. Presolar grains condensed in the envelopes of red giant branch (RGB) and asymptotic giant branch (AGB) stars, as well as in stellar explosions such as novae and supernovae, preserve nucleosynthetic signatures of their parent stars and of galactic chemical evolution (GCE) in their isotopic compositions (Zinner 1998, 2014). Multi-isotopic studies are therefore essential for disentangling the relative contributions of stellar nucleosynthesis, GCE, and secondary processing effects that cannot be resolved using a single isotopic system (Zinner et al. 2005).

Magnesium isotopes have been investigated in presolar oxides such as corundum, spinel, and hibonite, as well as in presolar silicates from primitive meteorites, interplanetary dust particles, and Antarctic micrometeorites (Nittler et al. 1997, 2008; Choi et al. 1998; Zinner et al. 2005; Nguyen & Zinner 2004; Leitner et al. 2012a; Kodolányi et al. 2014; Nguyen & Messenger 2014; Hoppe et al. 2018; Verdier-Paoletti et al. 2019). Studies of presolar oxides demonstrated that the abundances of ^{25}Mg and ^{26}Mg relative to ^{24}Mg increase with stellar metallicity, consistent with predictions from GCE models (Zinner et al. 2005). In contrast, Mg isotopic compositions of presolar silicates are commonly close to solar within analytical uncertainties, indicating that many silicates reflect either GCE-dominated starting compositions or partial isotopic equilibration in the interstellar medium or on meteorite parent bodies (Sanghani et al. 2024).

Despite this general trend, a subset of presolar silicates exhibits resolvable Mg isotopic anomalies that cannot be explained by standard AGB nucleosynthesis models. Some grains show excesses or depletions in both heavy Mg isotopes that are inconsistent with predictions from FRUITY models for low-mass (≈1.3–3 $M_\odot$) AGB stars with near-solar metallicity (Cristallo et al. 2015) and do not follow simple GCE trends. One plausible explanation is that the initial compositions of the parent stars were influenced by incomplete mixing of supernova ejecta in the interstellar medium, leading to locally heterogeneous reservoirs enriched or depleted in ^{24}Mg prior to star formation (Nittler et al. 2008; Leitner et al. 2019).

Combined O–Mg–Si isotopic systematics provide strong support for this interpretation. Grains exhibiting correlated enrichments in ^{17}O, heavy Mg isotopes, and heavy Si isotopes are consistent with GCE-dominated trends, whereas grains with anomalous Mg isotopic compositions but normal Si isotopes require more complex histories involving supernova-processed material mixed with either ^{24}Mg-rich or ^{24}Mg-poor reservoirs (Sanghani et al. 2024; Leitner et al. 2019). These observations underscore that Mg isotopes are particularly sensitive to inherited stellar compositions and radioactive contributions from ^{26}Al, and that their interpretation requires a multi-isotopic approach.

Overall, magnesium isotopes play a key role in distinguishing between AGB nucleosynthesis, GCE effects, and supernova-influenced starting compositions in presolar silicates. When combined with oxygen and silicon isotopes, Mg isotopic data reveal a complex picture in which both stellar evolution and large-scale galactic mixing processes contribute to the isotopic diversity observed in presolar silicate grains.

2.4 Iron Isotopes

Iron isotopic compositions of presolar silicates provide potentially valuable constraints on stellar nucleosynthesis and grain origins, but their application remains limited compared to oxygen, silicon, and magnesium due to analytical challenges and the small number of grains with resolvable anomalies. Most presolar silicates measured to date exhibit $^{54}Fe/^{56}Fe$ and $^{57}Fe/^{56}Fe$ ratios that are indistinguishable from solar within uncertainties (Nguyen & Messenger 2014; Ong & Floss 2015), although a small subset of grains shows detectable Fe isotopic deviations.

Recent syntheses emphasize that Fe isotopes are produced through multiple nucleosynthetic pathways and are sensitive to both hydrostatic and explosive burning in massive stars, making them, in principle, useful tracers of supernova contributions (Liu et al. 2024). Among presolar silicates, most grains with anomalous Fe isotopic compositions belong to Group 1, typically attributed to low-mass red giant branch or asymptotic giant branch (AGB) stars. A Group 3 silicate grain exhibiting a depletion in ^{57}Fe has also been reported, although large uncertainties prevent a definitive distinction between supernova and low-metallicity stellar origins (Nguyen & Messenger 2014). Similar ^{57}Fe depletions have been observed in a small number of mainstream SiC grains of AGB origin, indicating that such anomalies are not unique to silicates (Marhas et al. 2008).

Models predict that the largest Fe isotopic modifications in low-mass stars occur during the AGB phase, when neutron-capture nucleosynthesis in the He intershell enhances the heavier Fe isotopes (Cristallo et al. 2011). Grains forming early, under oxygen-rich conditions, are therefore expected to record only modest Fe isotopic anomalies, consistent with the near-solar Fe isotopic compositions observed for most presolar silicates. However, the observed depletions in ^{57}Fe are not reproduced by current AGB models and remain unexplained.

Overall, Fe isotopic data for presolar silicates remain sparse and difficult to interpret, but they provide an important complementary constraint when combined with O–Si–Mg isotopic systematics. As highlighted in recent reviews, future progress will depend on improved analytical precision, better grain isolation strategies, and coordinated multi-isotopic and microstructural analyses to fully exploit the diagnostic potential of Fe isotopes (Liu et al. 2024).

Taken together, oxygen, silicon, magnesium, and iron isotopic systematics demonstrate that presolar silicates record a wide range of stellar sources and nucleosynthetic histories, but also reveal significant degeneracies arising from galactic chemical evolution, inherited stellar compositions, and secondary processing. In many cases, isotopic information alone is insufficient to uniquely constrain grain formation environments or evolutionary pathways. Morphological features, internal textures, and phase associations therefore provide essential complementary constraints, linking isotopic signatures to specific condensation sequences, growth mechanisms, and alteration histories. The physical

characteristics of presolar silicates thus form a critical bridge between isotopic evidence and the astrophysical environments in which these grains formed and evolved.

3. MORPHOLOGY AND TEXTURAL DIVERSITY OF PRESOLAR SILICATES

The morphology and textures of presolar silicate grains provide important constraints on their formation in stellar environments and their subsequent processing in the interstellar medium (ISM) and on meteorite parent bodies. Because presolar silicates are typically submicrometer in size, their physical characteristics must be interpreted together with isotopic and chemical data. Variations in grain shape, surface features, and internal complexity reflect a combination of primary condensation processes and secondary modification histories, and in many cases help resolve ambiguities inherent in isotopic interpretations alone.

3.1 Morphology of Presolar Silicates in Different Extraterrestrial Samples

Presolar silicates exhibit a wide range of morphologies, including elliptical, irregular, and aggregate forms, across meteorites, interplanetary dust particles, and micrometeorites. Elliptically shaped grains are commonly interpreted as single grains and may preserve primary condensation features. Laboratory experiments and astronomical observations suggest that elongated silicate grains can form through growth or evaporation processes that favor elongation along crystallographic axes, particularly the *c*-axis of olivine (Fabian et al. 2001; Takigawa et al. 2008). Similar morphologies observed in presolar silicates may therefore reflect analogous formation mechanisms in circumstellar environments, provided the grains have not been significantly altered after condensation (Vollmer et al. 2009).

In contrast, most presolar silicates identified in meteorites exhibit irregular shapes, a trend consistently reported in multiple studies (Vollmer et al. 2009; Sanghani et al. 2021). Such morphologies may result from nonequilibrium condensation, fragmentation during grain–grain collisions in the ISM (Jones et al. 1996), or alteration in the solar nebula or on meteorite parent bodies. Surface features such as voids and rough textures observed on some grains are consistent with low-energy ion irradiation in the ISM, which can modify grain surfaces without erasing isotopic signatures. Grain morphology thus reflects the cumulative effects of stellar and post-stellar processing.

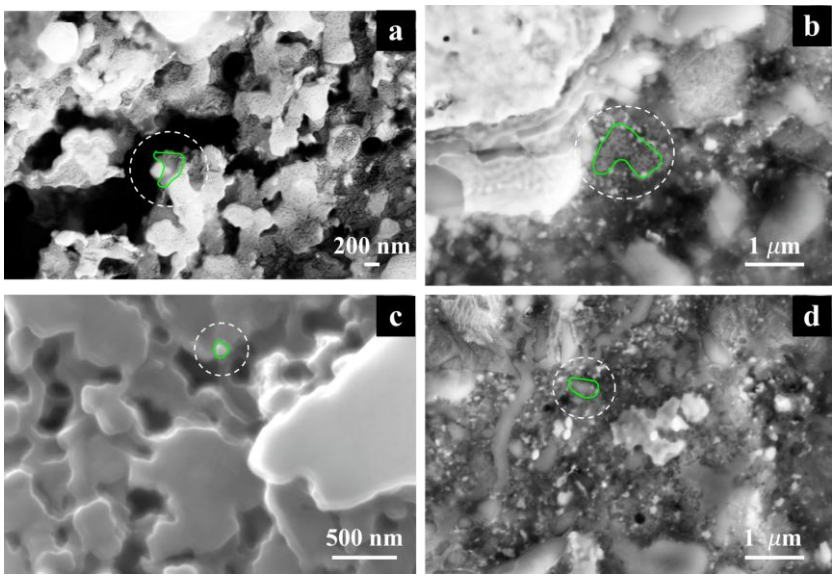


Fig.3. Various morphological features, e.g. (a) an irregular shape (b) with surface voids (c) near-circular shape and (d) an elliptical shape of presolar silicates. Image Credit: Sanghani M.N.

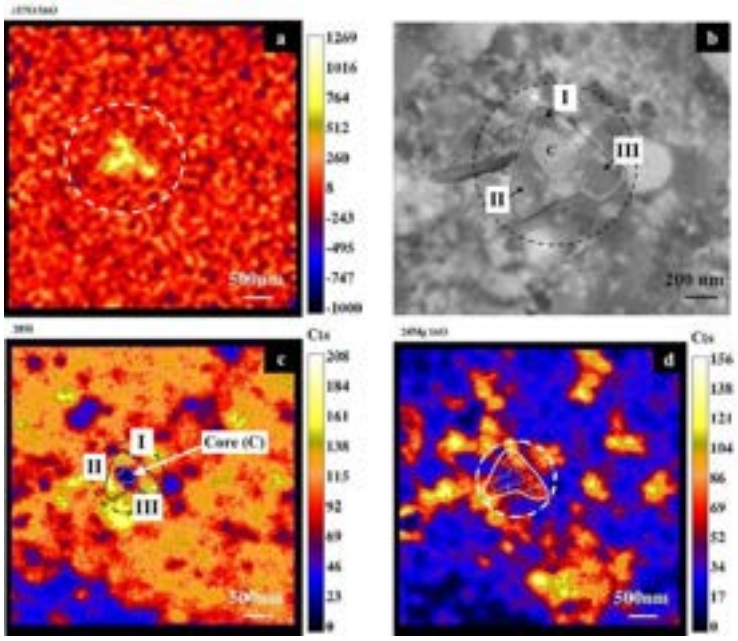


Fig.4. (a) $\delta^{17}O/^{16}O$ image (b) secondary electron image (c) ^{28}Si image and (d) $^{24}Mg^{16}O$ image of a composite presolar silicate grain. Image Credit: Sanghani et al. 2021

3.2 Complex and Composite Grains

A subset of presolar silicates exhibits complex or composite textures, in which silicate material is physically associated with refractory oxide phases. Such "complex grains" have been identified in several studies and provide direct evidence for multistage condensation and growth processes in stellar outflows (Vollmer et al. 2006, 2009; Bose et al. 2012; Floss & Stadermann 2012; Nguyen et al. 2014; Leitner et al. 2012b, 2018a). These grains commonly display core–rim architectures, multiphase assemblages, or aggregates of chemically distinct subgrains that nevertheless share homogeneous oxygen isotopic compositions within analytical uncertainties.

The presence of oxide–silicate associations is consistent with equilibrium and kinetic condensation models, in which refractory oxides such as Al_2O_3 condense at higher temperatures and earlier evolutionary stages, followed by silicate condensation under cooler conditions and higher mass-loss rates (Gail & Sedlmayr 1999; Demyk et al. 2000; Maldoni et al. 2008; Sloan et al. 2010). In this framework, early-formed oxide grains may act as nucleation sites for subsequent silicate growth, producing composite grains with oxide cores and silicate shells. Studies of complex grains have shown that oxygen isotopic compositions are often homogeneous across different subregions, despite clear chemical and structural contrasts, indicating that isotopic and elemental compositions can be decoupled during grain formation (Bose et al. 2012; Floss & Stadermann 2012; Nguyen et al. 2014).

3.3 Implications for Alteration and Processing Histories

The morphological and textural diversity of presolar silicates places important constraints on their alteration histories. Fragmentation of composite grains, partial loss of silicate rims, and surface roughening may result from irradiation and collisional processing in the ISM, as inferred for several complex grains reported in meteorites (Vollmer et al. 2009; Sanghani et al. 2021). Parent-body processes, including aqueous alteration and terrestrial weathering, can further modify grain textures through the development of cracks, veins, and secondary products without necessarily homogenizing isotopic compositions.

Aggregates composed of chemically distinct but isotopically homogeneous subgrains suggest that grain assembly may occur after condensation, potentially within dynamically evolving environments such as supernova ejecta or the ISM, where aggregation may be favored over growth of large single crystals (Messenger et al. 2005; Vollmer et al. 2008). Overall, presolar silicates record a complex interplay between primary condensation, aggregation, irradiation, and parent-body processing. Integrating morphological observations with isotopic, chemical, and structural data is therefore essential for reconstructing the full evolutionary histories of presolar silicate grains.

4. CHEMICAL AND STRUCTURAL CHARACTERIZATION OF PRESOLAR SILICATES

Isotopic measurements identify presolar silicate grains and constrain their stellar sources, but they provide limited information on the physical and chemical conditions of grain formation and evolution. Presolar silicates are submicrometer in size and commonly display chemical heterogeneity, variable crystallinity, and complex internal textures, reflecting nonequilibrium condensation and subsequent processing in circumstellar environments, the interstellar medium, and meteorite parent bodies (Keller & Messenger 2011). In the following subsections, we explore the mineralogical and chemical compositions of presolar silicates and discuss their implications for condensation pathways and alteration histories.

4.1 Conceptual Overview of Analytical Techniques

Transmission electron microscopy (TEM) and scanning transmission electron microscopy (STEM) are the primary tools for resolving the nanoscale mineralogy and internal structure of presolar silicates. These techniques provide direct imaging and diffraction information at nanometer to sub-nanometer resolution, allowing identification of crystalline, nanocrystalline, and amorphous phases, as well as defects, inclusions, and phase boundaries that record grain growth and processing histories (Stroud et al. 2008; Vollmer et al. 2009; Keller & Messenger 2011). Because many presolar silicates are poorly ordered or amorphous, electron microscopy is often the only means of determining their structural state.

Preparation of presolar silicates for TEM analysis typically relies on focused ion beam (FIB) techniques, which enable site-specific extraction of isotopically identified grains from meteorite thin sections or mounts while preserving internal structure and spatial context. Correlative NanoSIMS–FIB–TEM/STEM workflows have therefore become standard for directly linking isotopic compositions with nanoscale chemical and structural properties (Floss et al. 2008; Keller & Messenger 2011; Nguyen et al. 2016).

Electron energy loss spectroscopy (EELS), commonly coupled with STEM, provides information on local bonding environments and oxidation states. In presolar silicates, EELS measurements of Fe $L_{2,3}$ edges have been particularly valuable for determining Fe^{2+}/Fe^{3+} ratios and constraining redox conditions during grain formation and subsequent alteration (Stroud et al. 2009; Bose et al. 2010; Leitner et al. 2020; Sanghani et al. 2022). Auger electron spectroscopy (AES) complements these approaches by offering high-spatial-resolution surface compositions and efficient identification of O-rich presolar grains prior to more invasive analyses (Floss et al. 2008; Vollmer et al. 2009).

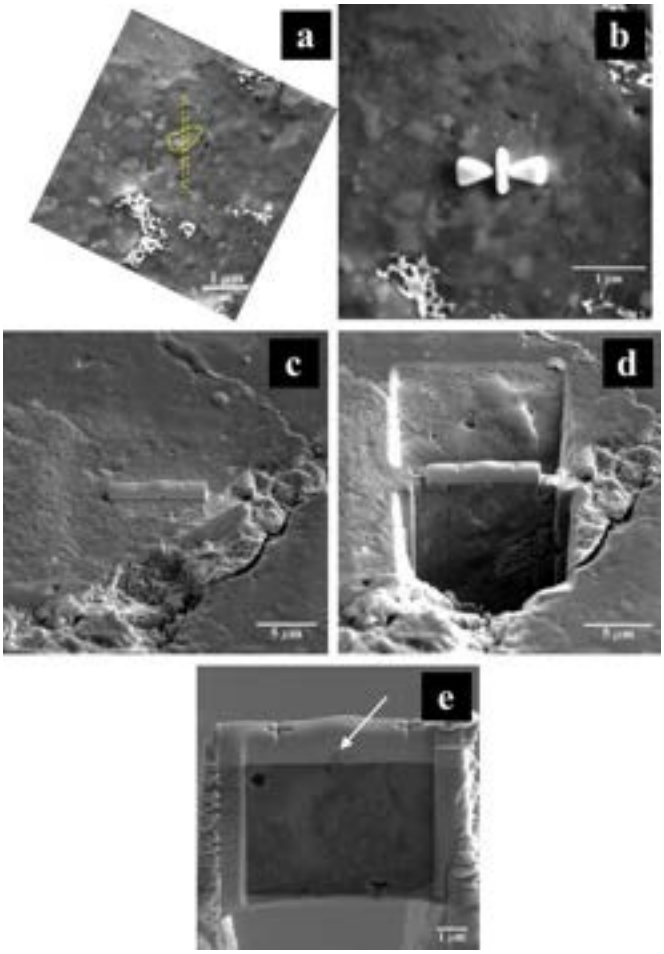


Fig.4. Preparation of sample grain for TEM analysis using FIB. Image Credit: Sanghani et al. 2022

4.2 Types and Chemical Compositions of Presolar Silicates

Most presolar silicates are ferromagnesian silicates broadly belonging to the olivine–pyroxene compositional system, with variable Mg/Fe ratios and frequent deviations from ideal stoichiometry (Keller & Messenger 2011; Bose et al. 2010). STEM-EDS analyses reveal a wide range of compositions, including Mg-rich, Fe-rich, and chemically heterogeneous grains, reflecting nonequilibrium condensation and evolving gas compositions during formation (Vollmer et al. 2009; Sanghani et al. 2022). Minor and trace elements such as Ca, Al, Mn, S, and, in some cases, F are variably incorporated and provide additional constraints on condensation environments and precursor phases (Bose et al. 2010; Leitner et al. 2020; Sanghani et al. 2022).

A subset of presolar silicates exhibits textures and compositions resembling GEMS (glass with embedded metal and sulfides), characterized by amorphous silicate matrices containing nanometer-scale metal or sulfide inclusions (Stroud et al. 2008; Keller & Messenger 2011). Although the origin of most GEMS remains debated, the identification of isotopically anomalous, presolar GEMS-like grains demonstrates that at least some formed in stellar or interstellar environments rather than exclusively in the solar nebula (Stroud et al. 2008; Keller & Messenger 2011).

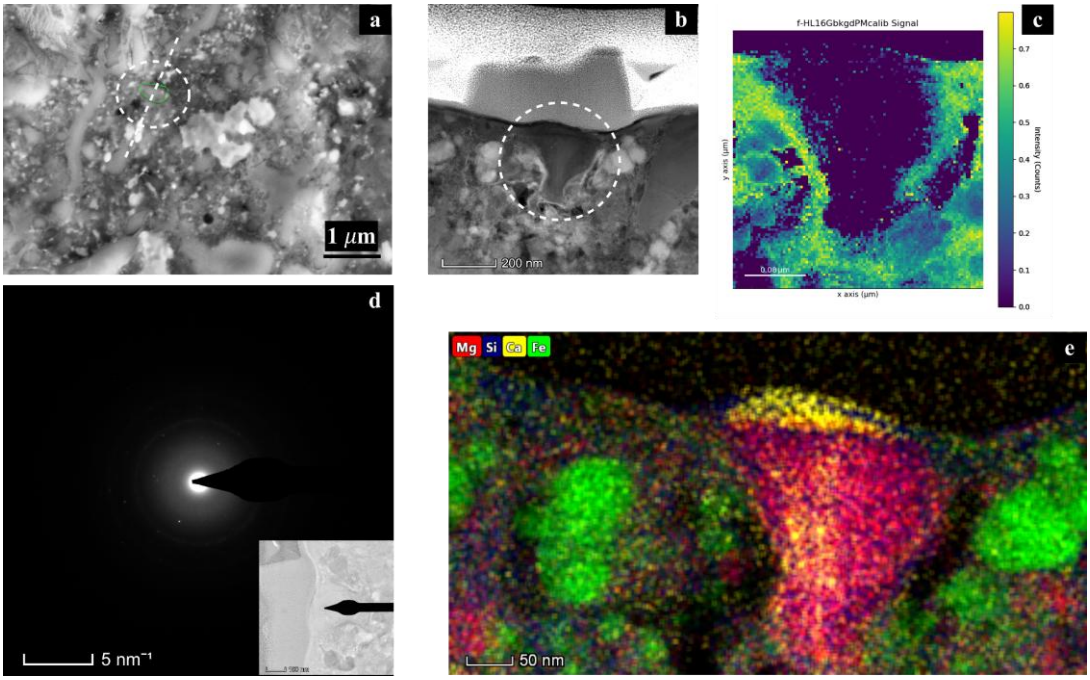


Fig.5. (a) SE image (b) STEM-HAADF image (c) $Fe^{3+}/\sum Fe$ map (d) Selected Area Electron Diffractogram (e) STEM-EDS map of a complex presolar silicate. Image Credit: Lajaunie et al. 2021

Chemical zoning, Fe-rich rims, and core–rim architectures are common features among presolar silicates and indicate multistage growth histories. Such textures may reflect initial condensation of Mg-rich phases followed by Fe enrichment during later stages of stellar outflows or chemical modification during residence in the ISM or incorporation into parent bodies (Vollmer et al. 2009; Bose et al. 2010; Leitner et al. 2020). Aggregate grains composed of chemically distinct but isotopically homogeneous subgrains further suggest assembly of presolar silicates from multiple components formed under varying physical conditions.

An important extension of this complexity is the discovery of a CaS–enstatite complex grain with a confirmed presolar origin (Sanghani et al. 2022). The observed phase relationship indicates condensation of CaS at higher temperatures, followed by enstatite condensation at lower temperatures, consistent with predicted condensation sequences in chemically evolving stellar outflows. This grain represents the first documented sulfide–silicate composite grain among presolar silicates and provides direct mineralogical evidence that sulfide phases can act as precursors to silicate growth.

4.3 Structural Characteristics and Alteration Histories

Structurally, presolar silicates occur predominantly as amorphous or poorly ordered phases, with nanocrystalline and fully crystalline grains also present (Stroud et al. 2008; Keller & Messenger 2011). The dominance of amorphous and nanocrystalline textures suggests rapid cooling and nonequilibrium condensation in stellar outflows, while the presence of crystalline domains may reflect localized annealing or slower cooling in specific environments (Floss et al. 2008; Vollmer et al. 2009).

Post-formation processing plays a major role in modifying presolar silicate structures. Ion irradiation in the ISM can amorphize initially crystalline silicates and generate defects, disordered rims, or surface voids without erasing isotopic anomalies (Stroud et al. 2008; Floss et al. 2008). Thermal metamorphism and aqueous alteration on meteorite parent bodies can further modify grain structures, leading to redistribution of Fe, oxidation–reduction reactions, and partial recrystallization (Vollmer et al. 2009; Nguyen et al. 2016; Leitner et al. 2020). EELS measurements revealing variable Fe^{2+}/Fe^{3+} ratios within

and among grains provide direct evidence for such redox modification (Stroud et al. 2009; Bose et al. 2010; Sanghani et al. 2022).

Importantly, many of these chemical and structural modifications occur at length scales smaller than those probed by isotopic measurements, allowing isotopic signatures to be preserved despite substantial nanoscale alteration.

4.4 New Tracers: Fluorine in Presolar Silicates

The detection of fluorine in presolar silicates introduces a new tracer for probing both stellar nucleosynthesis and parent-body processing. Fluorine production in astrophysical environments is poorly constrained and has been attributed to α-capture reactions in asymptotic giant branch and Wolf–Rayet stars, as well as neutrino-induced nucleosynthesis in type-II supernovae (Woosley & Haxton 1988; Lugaro et al. 2004), while proton- and α-capture reactions readily destroy ^{19}F, making its survival highly sensitive to local conditions. Recent STEM-EELS analyses have confirmed the presence of fluorine in presolar silicates from both putative AGB and supernova sources (Sanghani et al. 2022). In several grains, fluorine enrichment is spatially correlated with oxygen- or iron-depleted regions, suggesting redistribution during aqueous alteration on meteorite parent bodies, although primary incorporation during grain formation or early chemical processing cannot be excluded. The occurrence of F-bearing phases in carbonaceous chondrites further highlights the sensitivity of fluorine to low-temperature, fluid-mediated processes (Brearley 1990; Greenwood et al. 2010). These observations demonstrate that fluorine can survive in presolar silicates and that its distribution provides new constraints on volatile element behavior across stellar, interstellar, and parent-body environments.

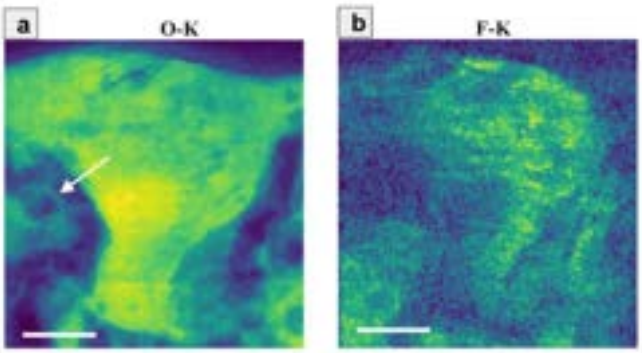


Fig.6. STEM-EELS map of (a) oxygen and (b) fluorine of a presolar silicate grain. Image Credit: Sanghani et al. 2022

5. ABUNDANCES OF PRESOLAR SILICATES

Presolar silicate abundances vary widely among extraterrestrial materials, spanning more than two orders of magnitude, and this variability primarily reflects the degree and nature of secondary processing rather than differences in stellar dust production. Because silicates are less refractory and more susceptible to destruction than presolar oxides, their survival is strongly influenced by thermal metamorphism, aqueous alteration, and shock-related processes experienced by their host materials (Nguyen et al. 2007; Floss & Stadermann 2009; Leitner et al. 2016). Consequently, comparisons of presolar silicate abundances across meteorites, interplanetary dust particles (IDPs), micrometeorites, and returned samples provide important constraints on both astrophysical dust preservation and Solar System processing histories.

Among all sample types, the highest presolar silicate abundances are consistently observed in IDPs, where values commonly reach several hundred ppm. These high abundances are attributed to minimal

parent-body alteration and preservation in cold, outer Solar System reservoirs, likely of cometary origin (Messenger et al. 2003; Floss et al. 2006, 2010; Busemann et al. 2009). Antarctic micrometeorites generally show lower, but still substantial, presolar silicate abundances, reflecting a combination of relatively primitive parent bodies and partial modification during atmospheric entry heating (Yada et al. 2008; Busemann et al. 2009).

In meteorites, presolar silicate abundances are systematically lower than in IDPs, indicating more extensive secondary processing even in the most primitive samples. The highest meteoritic abundances are found in carbonaceous chondrites that experienced minimal thermal metamorphism and limited aqueous alteration, particularly CR3 and CO3 chondrites and ungrouped carbonaceous chondrites such as Acfer 094, where abundances can reach up to ~200 ppm (Floss & Stadermann 2009; Nguyen et al. 2010; Leitner et al. 2016). A defining feature of these meteorites is the presence of abundant fine-grained, Fe- and Si-rich amorphous matrix material, which appears to enhance the preservation of fragile presolar silicates (Brearley 1993; Greshake 1997; Abreu & Brearley 2010). In contrast, CM chondrites typically exhibit much lower presolar silicate abundances, consistent with extensive aqueous alteration that promotes the formation of phyllosilicates and preferentially destroys presolar silicates (Leitner et al. 2020).

Within individual meteorites, presolar silicate abundances can vary significantly depending on petrographic context, illustrating the importance of localized processing. Studies of fine-grained rims (FGRs) and interchondrule matrix (ICM) in CR chondrites have shown that these components can record different alteration histories, with presolar silicate abundances varying accordingly (Leitner et al. 2016). In moderately to extensively altered CR2 chondrites, FGRs often preserve higher presolar silicate abundances than surrounding ICM regions, possibly due to early compaction that reduced fluid access during later alteration. In contrast, the least-altered CR chondrites can show the opposite trend, highlighting the sensitivity of presolar silicate preservation to both the timing and intensity of alteration.

Results from the CH3/CB_b3 chondrite Isheyevo and the CR2 chondrite Northwest Africa 801 provide illustrative examples of these effects. In Isheyevo, presolar silicate abundances differ between anhydrous and hydrated clasts, consistent with variable aqueous alteration histories (Bonal et al. 2010; van Kooten et al. 2017; Sanghani et al. 2021). Similarly, in NWA 801, presolar silicate abundances differ between fine-grained rims and interchondrule matrix, reflecting localized processing and possible preservation effects associated with rim formation. These case studies underscore that presolar silicate abundances are controlled not only by bulk meteorite classification, but also by small-scale lithologic and textural variations.

The ratio of presolar silicates to presolar oxides provides an additional metric for assessing processing histories. Theoretical models predict high silicate-to-oxide abundance ratios for dust condensed around AGB stars, consistent with ratios observed in IDPs and the most primitive CR chondrites (Floss & Stadermann 2009; Leitner et al. 2012a). Lower ratios observed in aqueously altered carbonaceous chondrites, such as CM chondrites and some CR2 chondrites, indicate preferential destruction of silicates relative to more refractory oxides during parent-body or nebular processing (Leitner et al. 2012b, 2015). However, interpretation of silicate-to-oxide ratios requires caution when grain statistics are limited, as small numbers of oxide grains can lead to large uncertainties.

Overall, presolar silicate abundances and silicate-to-oxide ratios together provide complementary constraints on the preservation and modification of circumstellar dust in Solar System materials. High abundances and high silicate-to-oxide ratios reflect minimal secondary processing and likely approximate the original presolar dust inventory, whereas reduced abundances and lower ratios record progressive destruction of silicates during thermal and aqueous alteration. Integrating abundance systematics with isotopic, chemical, and structural data is therefore essential for disentangling primary stellar signatures from secondary Solar System processes.

6. SYNTHESIS AND OUTLOOK

Presolar silicates provide a uniquely direct record of stellar nucleosynthesis, dust formation, and subsequent processing in the interstellar medium and the early Solar System. Advances in high-spatial-resolution isotopic imaging have transformed presolar silicates into a statistically meaningful grain population, enabling systematic classification based on oxygen isotopes and increasingly robust constraints from additional isotopic systems such as Si, Mg, and Fe. These data demonstrate that presolar silicates originate from diverse stellar sources, including low-mass AGB stars, supernovae, and, more rarely, novae, while also highlighting the need for multi-isotopic approaches to overcome ambiguities inherent in single-isotope interpretations.

Chemical and structural investigations reveal that presolar silicates are mineralogically and texturally complex, recording nonequilibrium condensation, multistage growth, and extensive post-formation modification. The prevalence of amorphous and nanocrystalline silicates, composite and aggregate grains, and variable redox states reflects the dynamic environments of stellar outflows as well as subsequent irradiation, thermal processing, and aqueous alteration. Correlative NanoSIMS–FIB–TEM/STEM–EELS studies show that isotopic signatures can survive substantial nanoscale chemical and structural modification, underscoring the importance of interpreting isotopic data within a mineralogical and alteration framework.

Abundance systematics provide a complementary constraint on presolar silicate preservation. Large variations in abundances among meteorites, IDPs, micrometeorites, and returned samples primarily reflect differences in parent-body and nebular processing rather than stellar dust production. High presolar silicate abundances and high silicate-to-oxide ratios in IDPs and the most primitive carbonaceous chondrites likely approximate the original presolar dust inventory, whereas reduced abundances and lower ratios record preferential destruction of silicates during thermal or aqueous alteration. Petrographically resolved studies further demonstrate that preservation can vary significantly on small spatial scales within individual samples.

Despite substantial progress, important challenges remain. For many grains, stellar source assignments are still limited by analytical uncertainties, signal dilution, and incomplete isotopic coverage, particularly for Mg, Fe, and heavier elements. Distinguishing Galactic chemical evolution effects from local stellar nucleosynthesis signatures remains difficult, and the extent to which parent-body processes modify elemental and isotopic records is not yet fully constrained.

Future progress will rely on continued improvements in analytical capabilities and sample availability. Advances in NanoSIMS instrumentation, combined with integrated FIB–TEM/STEM–EELS and Auger analyses, will enable more comprehensive multi-isotopic and multi-element characterization of individual grains. Sample-return missions from primitive asteroids and comets offer exceptional opportunities to study presolar silicates in well-characterized contexts with minimal contamination. Expanding grain inventories and improving statistical robustness, together with closer integration of laboratory data and astrophysical models, will further establish presolar silicates as quantitative probes of stellar dust production and Galactic chemical evolution.

REFERENCES

Abreu, N. M., & Brearley, A. J. 2010, *Geochim. Cosmochim. Acta*, 74, 1146
Anders, E. 1971, *Annu. Rev. Astron. Astrophys.*, 9, 1
Anders, E., & Zinner, E. 1993, *Meteoritics*, 28, 490
Black, D. C., & Pepin, R. O. 1969, *Earth Planet. Sci. Lett.*, 6, 395
Bonal, L., Quirico, E., Bourot-Denise, M., & Montagnac, G. 2010, *Geochim. Cosmochim. Acta*, 74, 6590
Boothroyd, A. I., & Sackmann, I.-J. 1999, *Astrophys. J.*, 510, 232
Bose, M., Floss, C., Stadermann, F. J., & Stroud, R. M. 2010, *Astrophys. J.*, 714, 1624

Bose, M., Floss, C., Stadermann, F. J., & Stroud, R. M. 2012, *Astrophys. J.*, 755, 21
Brearley, A. J. 1990, *Geochim. Cosmochim. Acta*, 54, 831
Brearley, A. J. 1993, *Geochim. Cosmochim. Acta*, 57, 1521
Busemann, H., Nguyen, A. N., Cody, G. D., et al. 2009, *Earth Planet. Sci. Lett.*, 288, 44
Choi, B.-G., Wasserburg, G. J., & Huss, G. R. 1998, *Astrophys. J.*, 498, 879
Clayton, D. D. 1975, *Astrophys. J.*, 199, 765
Cristallo, S., Piersanti, L., Straniero, O., et al. 2011, *Astrophys. J. Suppl. Ser.*, 197, 17
Cristallo, S., Straniero, O., Piersanti, L., & Gobrecht, D. 2015, *Astrophys. J. Suppl. Ser.*, 219, 40
Demyk, K., Dartois, E., Wiesemeyer, H., Jones, A. P., & d'Hendecourt, L. 2000, *Astron. Astrophys.*, 364, 170
Epstein, S., Krishnamurthy, R. V., Cronin, J. R., Pizzarello, S., & Yuen, G. U. 1987, *Nature*, 326, 477
Fabian, D., Henning, T., Jäger, C., Mutschke, H., & Dorschner, J. 2001, *Astron. Astrophys.*, 378, 228
Fenner, Y., Gibson, B. K., Lee, H.-C., et al. 2003, *Publ. Astron. Soc. Aust.*, 20, 189
Floss, C., & Stadermann, F. J. 2009, *Geochim. Cosmochim. Acta*, 73, 2415
Floss, C., & Stadermann, F. J. 2012, *Geochim. Cosmochim. Acta*, 76, 31
Floss, C., Stadermann, F. J., & Bose, M. 2008, *Astrophys. J.*, 672, 1266
Floss, C., Stadermann, F. J., Bradley, J. P., Dai, Z. R., Bajt, S., & Graham, G. 2006, *Astrophys. J.*, 642, 1190
Floss, C., Stadermann, F. J., Kearsley, A. T., Burchell, M. J., Ong, W. J., & Snead, C. J. 2013, *Astrophys. J.*, 763, 140
Gail, H.-P., & Sedlmayr, E. 1999, *Astron. Astrophys.*, 347, 594
Greenwood, R. C., Lee, M. R., Hutchison, R., & Barber, D. J. 2010, *Geochim. Cosmochim. Acta*, 74, 1684
Greshake, A. 1997, *Geochim. Cosmochim. Acta*, 61, 437
Halbout, J., Robert, F., Javoy, M., & Clayton, R. N. 1986, *Geochim. Cosmochim. Acta*, 50, 2343
Hoppe, P., Leitner, J., Kodolányi, J., et al. 2017, *Geochim. Cosmochim. Acta*, 221, 182
Jones, A. P., Tielens, A. G. G. M., & Hollenbach, D. J. 1996, *Astrophys. J.*, 469, 740
José, J., & Hernanz, M. 1998, *Astrophys. J.*, 494, 680
Keller, L. P., & Messenger, S. 2011, *Geochim. Cosmochim. Acta*, 75, 5336
Kodolányi, J., Hoppe, P., & Vollmer, C. 2014, *Geochim. Cosmochim. Acta*, 127, 377
Lajaunie, L., Botton, G. A., & Couillard, M. 2021, *Microsc. Microanal.*, 27, 1225
Lee, T. 1988, *Geochim. Cosmochim. Acta*, 52, 2567
Lewis, R. S., Srinivasan, B., & Anders, E. 1975, *Science*, 190, 1251
Lewis, R. S., Amari, S., & Anders, E. 1987, *Nature*, 326, 160
Leitner, J., & Hoppe, P. 2019, *Nat. Astron.*, 3, 725
Leitner, J., Vollmer, C., Hoppe, P., & Zipfel, J. 2012a, *Astrophys. J.*, 745, 38
Leitner, J., Hoppe, P., Floss, C., & Zipfel, J. 2012b, *Meteorit. Planet. Sci.*, 47, 1996
Leitner, J., Hoppe, P., Floss, C., & Zipfel, J. 2015, *Meteorit. Planet. Sci.*, 50, 1318
Leitner, J., Hoppe, P., Floss, C., & Zipfel, J. 2016, *Geochim. Cosmochim. Acta*, 190, 38
Leitner, J., Hoppe, P., Floss, C., & Zipfel, J. 2018b, *Geochim. Cosmochim. Acta*, 221, 255
Leitner, J., Hoppe, P., Floss, C., & Zipfel, J. 2020, *Geochim. Cosmochim. Acta*, 269, 91
Liu, N., Lugaro, M., Leitner, J., Meyer, B. S., & Schönbächler, M. 2024, *Space Sci. Rev.*, 220, 88
Lugaro, M., Gallino, R., Amari, S., et al. 1999, *Astrophys. J.*, 527, 369
Lugaro, M., Karakas, A. I., Stancliffe, R. J., & Rijs, C. 2004, *Astrophys. J.*, 615, 934
Maldoni, M. M., Ireland, T. R., & Hoppe, P. 2008, *Meteorit. Planet. Sci.*, 43, 1243
Marhas, K. K., Hoppe, P., & Ott, U. 2008, *Meteorit. Planet. Sci.*, 43, 725
McKeegan, K. D., Aléon, J., Bradley, J. P., et al. 2006, *Science*, 314, 1724
Messenger, S., Keller, L. P., Stadermann, F. J., Walker, R. M., & Zinner, E. 2003, *Science*, 300, 105
Messenger, S., Keller, L. P., Nguyen, A. N., & Zinner, E. 2005, *Astrophys. J.*, 627, 643
Mostefaoui, S., & Hoppe, P. 2004, *Astrophys. J. Lett.*, 613, L149
Nagashima, K., Krot, A. N., & Yurimoto, H. 2004, *Nature*, 428, 921
Nagashima, K., Krot, A. N., & Yurimoto, H. 2005, *Geochim. Cosmochim. Acta*, 69, 5417
Nguyen, A. N. 2005, *Astrophys. J.*, 632, 584

Nguyen, A. N., & Messenger, S. 2014, *Astrophys. J.*, 784, 149
Nguyen, A. N., & Messenger, S. 2016, *Astrophys. J.*, 827, 12
Nguyen, A. N., Stadermann, F. J., Zinner, E., et al. 2007, *Astrophys. J.*, 656, 1223
Nguyen, A. N., & Zinner, E. 2004, *Science*, 303, 1496
Nguyen, A. N., Nakamura-Messenger, K., Messenger, S., Keller, L. P., & Klöck, W. 2014, *Lunar Planet. Sci. Conf.*, 45, 2351
Nguyen, A. N., Nittler, L. R., Stadermann, F. J., et al. 2010, *Astrophys. J.*, 719, 166
Nittler, L. R., Alexander, C. M. O'D., Gao, X., Walker, R. M., & Zinner, E. 1997, *Astrophys. J.*, 483, 475
Nittler, L. R., Alexander, C. M. O'D., Gallino, R., et al. 2018, *Astrophys. J.*, 856, 58
Nittler, L. R., & Hoppe, P. 2005, *Astrophys. J.*, 631, L89
Nittler, L. R. 2019, *Geochem. Perspect.*, 8, 1
Nittler, L. R., Hoppe, P., Alexander, C. M. O'D., et al. 2008, *Astrophys. J.*, 682, 1450
Ong, W. J., & Floss, C. 2015, *Meteorit. Planet. Sci.*, 50, 1192
Rauscher, T., Heger, A., Hoffman, R. D., & Woosley, S. E. 2002, *Astrophys. J.*, 576, 323
Reynolds, J. H. 1964, *Phys. Rev. Lett.*, 12, 417
Sanghani, M. N., Marhas, K. K., Hsiao, S. S.-Y., Peeters, Z., Shang, H., Lee, D.-C., & Bizzarro, M. 2021, *Astrophys. J. Suppl. Ser.*, 253, 41
Sanghani, M. G., Marhas, K. K., Hsiao, S. S. Y., et al. 2022, *Astrophys. J.*, 931, 105
Sanghani, M. N., Marhas, K. K., Hsiao, S. S. Y., Shang, H., Lee, D. C., Lee, T., Verdier-Paoletti, M., Nittler, L. R., & Bizzarro, M. 2024, in *National Space Science Symposium*, India (abstract)
Sloan, G. C., Lagadec, E., Zijlstra, A. A., et al. 2010, *Astrophys. J.*, 719, 1274
Stadermann, F. J., Floss, C., & Hoppe, P. 2008, *Meteorit. Planet. Sci.*, 43, 299
Stroud, R. M., Chisholm, M. F., Heck, P. R., Alexander, C. M. O'D., & Nittler, L. R. 2009, *Astrophys. J.*, 696, 115
Stroud, R. M., Nittler, L. R., & Alexander, C. M. O'D. 2008, *Astrophys. J.*, 675, 394
Takigawa, A., Tachibana, S., Nagahara, H., & Ozawa, K. 2008, *Astrophys. J. Lett.*, 688, L29
van Kooten, E. M. M. E., Wielandt, D., Schiller, M., et al. 2017, *Proc. Natl. Acad. Sci.*, 114, 6716
Verdier-Paoletti, M. J., Floss, C., Bose, M., & Haenecour, P. 2019, *Meteorit. Planet. Sci.*, 54, 1850
Vollmer, C., & Hoppe, P. 2010, *Astrophys. J. Lett.*, 719, L172
Vollmer, C., Hoppe, P., & Brenker, F. E. 2006, *Astrophys. J.*, 645, 608
Vollmer, C., Hoppe, P., Stadermann, F. J., et al. 2008, *Astrophys. J.*, 684, 611
Vollmer, C., Hoppe, P., Stadermann, F. J., et al. 2009, *Astrophys. J.*, 700, 774
Wasserburg, G. J., Lee, T., & Papanastassiou, D. A. 1977, *Geophys. Res. Lett.*, 4, 299
Waters, L. B. F. M., Molster, F. J., de Jong, T., et al. 1996, *Astron. Astrophys.*, 315, L361
Woosley, S. E., & Haxton, W. C. 1988, *Nature*, 334, 45
Yada, T., Floss, C., Stadermann, F. J., et al. 2008, *Meteorit. Planet. Sci.*, 43, 1287
Yurimoto, H., Itoh, S., Zinner, E., et al. 2003, *Appl. Surf. Sci.*, 203–204, 793
Zinner, E. 1998, *Annu. Rev. Earth Planet. Sci.*, 26, 147
Zinner, E. 2014, in *Treatise on Geochemistry*, 2nd ed., Vol. 1, ed. A. M. Davis (Oxford: Elsevier), 181
Zinner, E., Nittler, L. R., Gallino, R., et al. 2006, *Astrophys. J.*, 650, 350

A COMPREHENSIVE REVIEW OF BIO-PIGMENT VIOLACEIN: PROPERTIES AND APPLICATIONS IN INDUSTRIAL AND MEDICAL BIOTECHNOLOGY

Brinda Ahir[1], Smit Marvaniya[1], Prit Ashara[*]

[1]*Semester 6 students, Biotechnology Engineering Department, VVP Engineering College, Rajkot, Gujarat-360005*

[*]*Assistant Professor, Biotechnology Engineering Department, VVP Engineering College, Rajkot, Gujarat-360005 (Email for correspondence: prit.ashara.bt@vvpedulink.ac.in)*

ABSTRACT: The increasing adverse effects associated with synthetic dyes have shifted the research attention towards natural pigment and pigment – producing organisms, collectively known as Bio-pigments. Among the different bacteria that produce pigments, *Chromobacterium violaceum* has received considerable attention due to its deep violet pigment and the exceptional bioactive attributes of the violacein pigment. *Chromobacterium violaceum* is a Gram-negative, facultative anaerobic, rod-shaped species commonly found in tropical and subtropical regions soil and water. It produces bright violet pigment – Violacein, which exhibits a wide range of biological activities, including anticancer, antimicrobial, antifungal, antioxidant, antiparasitic, antiviral, antibiofilm, and anti-inflammatory properties. Owing to these multifunctional characteristics, violacein has gained considerable interest for applications in textiles, medicine, and natural dyeing industries. This paper reviews about characteristics of violacein, its diverse application and properties, its production and optimization strategies, and future scope.

KEYWORDS: bio-pigment, *Chromobacterium violaceum,* violacein, antimicrobial, anticancer, textile.

INTRODUCTION

Chromobacterium violaceum is a gram negative, facultative anaerobic, rod-shaped species, which commonly are saprophyte found in soil and water of tropical and subtropical regions. These species are generally considered as non-pathogenic to humans and animals [1]. C*hromobacterim violaceum* belongs to family Neisseriaceace, order Neisseriales, class Betaproteobacteria, phylum Pseudomonadota. If we see there are currently 11 species known in this genus [2]. For the first time it was extracted from wet rice paste in the year 1882. The species give specific deep violet pigment – violacein, the production of the pigment occurs in the aerobic condition. The violacein is involved in the regulation of production of Tryptophan; the increased level of amino acid is destructive for the bacteria [3]. The genus *Chromobacterim* comprises a group of bacilli producing on ordinary media a characteristic pigment violacein which is soluble in ethanol but not in water or chloroform, possess both polar monotrichous and peritrichous flagella. They produce some amount of acidity from carbohydrates and utilize citrate [4]. Violacein is a bis-indole pigment synthesized as a secondary metabolite by a number of bacterial species [5]. Current literature suggests that it possess the antitumoral, antibacterial, antiulcerogenic, antileishmanial and antiviral activities [6]. Other research also suggests that it may have immunomodulatory properties as well as suppressive and inhibitory activity against *Staphylococcus aureus* and other multidrug-resistant pathogens [5]. Further studies show that this organism has been characterised as an opportunistic pathogen causing severe and fatal indication in immunocompromised hosts. Wooley first reported its pathogenic characteristics in 1905, observing septicemia in water buffaloes caused by these bacteria [7].

Violacein production is control by operon which consist of mainly five different genes, collectively named as vioABCDE.

vioA: Catalyse second molecules Tryptophan by oxidative deamination

vioB: This protein is similar to polyketide synthase – able to catalyse nonribosomal peptide bond.

vioC: Catalyse intermediate violacein by oxidation, also in deoxyviolacein.

vioD: Catalyse first Tryptophan molecule hydroxylation.

vioE: Rearrange the dimer into prodeoxyviolacein. Its improper functioning causes yellow colouration of colonies.

vioA, vioD, vioC are the nucleotide dependent monooxygenase activity [3,8][9]

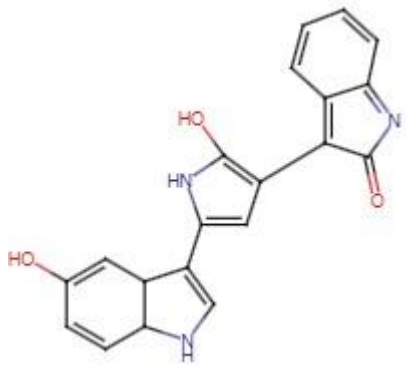


Fig 1. Violacein structure

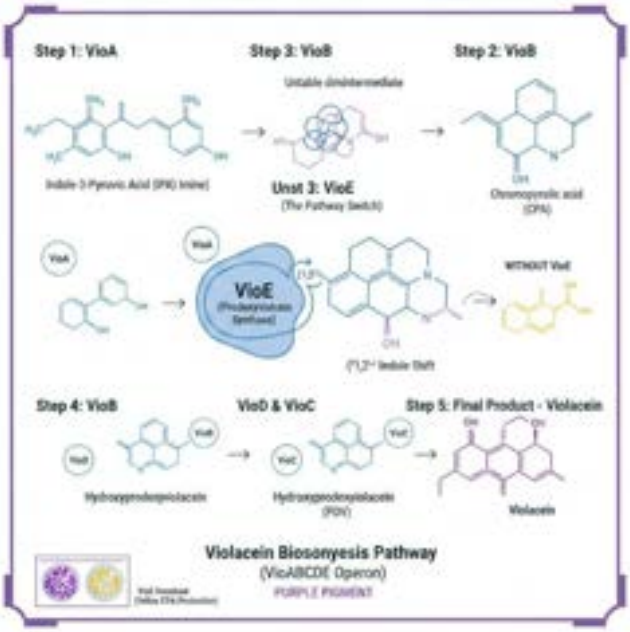


Fig 2. Violacein biosynthesis pathway

Tryptophan is the precursor for violacein production, all carbon, hydrogen and nitrogen atoms of violacein comes from the tryptophan. Hydroxylation of tryptophan to 5-hydroxy tryptophan, suggest that molecular oxygen participates in oxidation reaction confirming that aerobic conditions are necessary for violacein production [3,9]

Table1. violacein production from various organism.

Strains	Comments
Alteromonas luteoviolacea	Marine bactéria, Scotland
Chromobacterium violaceum ATCC 553	Collection
Chromobacterium violaceum	Bacillus violaceum, USA
Chromobacterium violaceum	Lowland river, England
Chromobacterium violaceum B78	From Amazon River, Manaus, Brazil
Chromobacterium violaceum CCT 3496	Collection
Chromobacterium violaceum	Agricultural waste, Malaysia
Citrobacter freundii/ pCOM10vio	Recombinant strains
Chromobacterium violaceum (MTCC 2656)	Indian Collection
Duganella sp.	Agricultural soils (olive), Spain
Duganella violaceinigra	Forest soil
Duganella sp. B2	From China
Escherichia coli MG1655-Vio4	Recombinant strains
Escherichia coli BL21(DE3)/ pET32avio	Recombinant strains
Escherichia coli Bl21(DE3) B2/pED + pVio	Recombinant strains

Janthinobacterium lividum S9601	Collection
Janthinobacterium lividum strain DSM1522	Collection
Janthinobacterium lividum	Glacier, China
Janthinobacterium svalbardensis	Glacier, Slovenia
Pseudoalteromonas DSM 13623	Marine sediment bacterium Germany Patent
Pseudoalteromonas sp.	Deep sea Waters, Japan
Psychotropic bacterium RT102	Close to *J. lividum*
Pseudoalteromonas luteoviolacea	Marine sponge, China
Psychrotrophic bacterium, XT1	Close to *J. lividum*

PROPERTIES

ANTIMICROBIAL / ANTIBACTERIAL

There are very few gram-negative species which are susceptible to violacein, some studies claim that violacein exhibits low MIC or growth inhibitory activities with Gram-negative strains. In contrast, the Gram-positive bacterial strain activity is well established, including *Staphylococcus*, *Bacillus* and *Streptococcus*, also to acid-fast microbes which are *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*, though *Enterococcus faecalis* was not affected by the violacein. While research it was found that multidrug-resistant *Streptococcus aureus* clinical isolate with resistance to seven different antibiotics was found to be susceptible to violacein. Violacein rapidly insert into the membranes of the microbe and disrupt their integrity, leading to ATP and protein leakage. Recently modelled interaction between violacein and bacterial membrane suggested that this compound does not embed very deeply within the lipid bilayer. The other studies show that *E. coli* cellular membranes are just likely to be attacked by violacein and that its inherent resistance to violacein stems from the protective nature of the outer membrane, which absorbs this antibiotic and prevents its access to the cytoplasmic membrane. Doing research from different perspective, asking how violacein acts as an antibiotic in nature if it is hydrophobic and remains embedded primarily within the membrane of the strain that produced it [6]. Liposomes made by commercial and bacterial phospholipids were disrupted after direct interaction with violacein, leading to their increased permeability.

ANTIFUNGAL

A study report shows action of violacein against phytopathogenic fungus *Rosellinia necatrix*, the causative agent of white root rot of mulberry, suggesting a possible application as a fungicide [10]. Plant pathogen *Rhizoctonia solani* and *Batrachochytrium dendrobatidis*, a fungus that is lethal to amphibians are susceptible to violacein. In the other case study, the violacein producing bacterium, *J. lividium*, present on skin of the black-backed salamander or frog provided protection against *B. dendrobatidis* [6]. The other studies have coupled silver nanoparticles to violacein and evaluated the pharmacological profile of this new molecule against different fungal species, such as *Aspergillus tamari, A. tubingensis* and *F. proliferatum*, as an attempt to solve the violacein solubility issue. Results have shown increased efficiency in antifungal activity. Thus, nanotechnology is an important tool used to help improving the antifungal activity of new or already used compounds in therapies [11].

ANTIPARASITIC

The main interest in using violacein as antiparasitic molecule is based on its reported leishmanicidal, trypanocidal, antioxidant and antiprotozoal activities, among others. Comparative violacein and dexoyviolacein toxicity analysis applied to *P. falciparum* and red blood cells has shown approximately 5 and 20 times more sensitivity to the parasite than the sensitivity recorded for mammalian cells, respectively. Another research studies focused on investigating violacein's effect on *Acanthamoeba castellanii*, a protozoan species capable of causing severe diseases, such as amebic keratitis and

granulomatous amebic encephalitis – has shown that 1 μM of violacein was capable of producing 100% amoeba lysis. Studies shows that main cause of increased nematodes mortality rate is accumulation of violacein in the intestine, and also violacein inhibits the *C. elegans* larval development [11]. The antimalarial activity of violacein was demonstrated in-vitro and in-vivo, shown that micromolar concentrations of violacein efficiently killed chloroquine sensitive and resistant *P. falciparum* strains in-vitro [10]. And in-vivo activity against *P. chabaudi chabaudi*, trypanocidal activity was reported against *Trypanosoma cruzi* [12]. The antileishmanial activity was tested against promastigote forms of *Leishmania amazonensis*, the result showed an effect that was comparable with that of pentamidine – the second drug of choice in the treatment [13].

ANTIVIRAL

Violacein was reported to shoe antiviral activity against herps simplex virus and poliovirus after infection of HeLa cells. At concentrations of about 0.25 μg/ml, violacein inhibited HSV 62%, and, at 0.063 μg/ml, it inhibited 56 % of poliovirus-infected HeLa cells. These results indicate a modest antiviral activity for the compound [13]. Other research employed three experimental approaches to study the antiviral and cytotoxic activities of violacein, with concluding results that with those previously published by other studies [10].

ANTIOXIDANT

The antioxidant activity was tested through ferric-reducing activity, were approximately 20% that of ascorbic acid which is well characterised, potent antioxidant. The antioxidant activity of violacein increased considerably in the presence of ascorbic acid; i.e. 138.4% relative to ascorbic acid alone [14]. Using density functional theory method, it was found that violacein had a strong electron donating ability, but a weak hydrogen atom transfer ability. The coplanar isomers had the strong hydrogen atom transfer ability as well as electron donating ability, and then the N7–H7 band plays an important role in its antioxidant activity. Different experimental approach demonstrated that violacein prevented the strand-breaking action of Methylene Blue on plasmid DNA [10].

ANTICANCER

A study reported violacein – induction of apoptosis in myeloid leukemia HL60 cells. This study was followed by the studies in the less differentiated TF1 leukemia cells, which are known to be resistant to PCD (programmed cell death). Violacein cytotoxicity against breast cancer cells (MCF7 cell line), for example, was investigated, aiming to clarify its potential effect on cancer invasion and metastasis**.** The studies reported the ability of violacein at a non-cytotoxic concentration (1 $\mu mol \cdot dm^{-3}$) to downregulate the chemokine/receptor CXCL12/CXCR4 interaction in breast cancer cells. This chemokine axis is involved in the adhesion and migration of metastasis-initiating cancer cells and has been implicated in the pathogenesis of breast and pancreatic cancer and leukemias. These results are of particular importance considering the roles of CXCL12/CXCR4 not only in cell-tumor microenvironment interaction, leading to tumor growth and metastasis, but also in angiogenesis. Interestingly, the study also reported a decrease in TNF-α and TGF-β-induced MMP2/9 expression by breast cancer cells. The other study reported the cytotoxicity and mechanisms of action of violacein in three different cells line – glioblastoma(U87), breast cancer (MCF7) and lung cancer(A549). Breast and lung cancer are two of the most common cancer types in men and women, frequently showing a high incidence of brain metastasis. The authors also investigated the ability of violacein to inhibit migration of glioblastoma cells. Using a crystal violet proliferation assay, it was observed that 1 $\mu mol \cdot dm^{-3}$ violacein inhibited proliferation of all cell lines. Cell viability experiments indicated a significant sensitivity of glioblastoma and lung cancer cells to violacein, with an IC50 value around 1 $\mu mol \cdot dm^{-3}$ [12]. Study has shown that this pigment is non-genotoxic. It is known that cytotoxic agent often induce cell death, partly due to reactive oxygen species generation. A investigated study was carried out to know whether violacein affects cellular redox status, the observation were that violacein was cytotoxic to CHO-K1

and MRC5 (non-tumor), and to HeLa cell lines (tumor), in this study no association between increased oxidative stress and cell death induction was observed, however enhanced mitochondrial membrane potential was observed. The poor violacein solubility in water which can compromise its bioavailability and, consequently, lead to low effect on biological systems – some studies have suggested using delivery devices located with violacein [11].

ANTIPYRETIC

Hyperthermia was induced in Rats by injecting brewer's yeast subcutaneously. Initial rectal temperature was measured. after 18h, Rats showed increased in rectal temperature when thermometer inserted into rectum about 3 cm. Violacein was given to rats in different concentration and Paracetamol as a reference drug. Decline in temperature was observed but not as much as paracetamol. This experiment supports the Antipyretic activity of Violacein [15]

IMMUNOMODULATORY

Violacein has antipyretic, analgesic and immunomodulatory reactions when orally administered to rats. In ulcer rat models, violacein relieved inflammation of the gastrointestinal tract, possibly working through the COX-1 mediated pathways while other study suggesting violacein immunomodulatory effects by regulating cytokine production: production: it down regulated the expression of IL-6 and TNF-α but induced expression of IL-10, on direct injecting into the intraperitoneal cavity. Unlike the above study that reported violacein inhibits TNF-α expression, TNF-α expression was elevated in HL60, and TNF receptor 1 signaling was also activated when this cell line was exposed to violacein . These differences may be due to the experimental protocols, though, as the above studies were performed in vitro, i.e., violacein treatment directly into human or murine cell cultures, rather than in vivo, i.e., the oral administration or injection of violacein into the digestive tract or intraperitoneal cavity [16]

USES / APPLICATIONS

Violacein uses in dyeing of fabric:

Pigment dyeing property was tested on cotton, khadi, synthetic cloth. Two different methods were used for dyeing fabrics – first method in which Simultaneous Fermentation and Dyeing (SFD) was done. In the second method, Dyeing After Fermentation and Sonication (DAFS) was done. Checking the stability of dyed fabric: In order to mimic the human skin and body conditions, the dyed fabric was dipped in tap water adjusted to pH 3.5 with concentrated HCl and was further incubated. Further, effect of different detergents 0.1% SLS (anionic), Tween 80 (neutral) and normal detergents was checked on the dyed fabric and the light fastness of the dyed fabric was checked directly by exposing the fabric to the sunlight for 1hr and the intensity of the sunlight was measured with the lux meter.

Results: out of the two different methods of dyeing, SFD gave the best dyeing results. Dyed fabric showed good stability in acidic conditions and at 37°C. Pigment sustained its colour, no fading of the colour was observed on exposing to the direct sunlight of the summer afternoon. Further, on washing with different detergents, tween showed best results with no colour fading followed by SLS with little fading and with normal detergents it showed considerable fading. This indicates the stability of the dyed fabric to suit human body temperature, skin pH and climatic conditions. Further parameters like abrasion resistance, color fastness need to be checked [17].

Violacein uses in cosmeceuticals

Three commercially available sunscreens with stated sunscreens protection factor (SPFs) of 15, 24 and 40 were purchased from local market of India. Plant extracts were prepared by mixing 10g of *Aloe vera* leaf and *Cucumis sativus* (cucumber) fruit, followed by filtration through muslin cloth. Violacein was added at 4% [w/w] to samples of the commercial sunscreens and plant extracts. Commercial sunscreens, plant extracts, pigment-supplemented sunscreens, and pigment-supplemented plant extracts were used

for empirical determinations of SPF. Purified pigments, commercial sunscreens, and their combinations were subjected to spectroscopic and chromatographic analysis. The pigments, individually and in combination with commercial sunscreen (SPF 15), were analysed by FTIR spectroscopy using a Testscan Shimadzu FTIR 8400 at 4000 to 400 cm^{-1}.
Results: Enhanced UV absorption within range 290 to 320 nm indicates that the pigment have the potential to increase SPFs when combined with sunscreen formulations. The SPFs of the three commercial sunscreens assayed were also determined empirically. The stated SPFs were 15,24 and 40, the empirically determined values were 16.46±0.83, 26.64±0.58 and 39.48±0.73, respectively. Addition of violacein to each of these sunscreens caused an increase of SPF to 19.66±0.6, 32.63±0.89, 43.79±0.86, respectively, showed 19, 22 and 10% increases in SPF, respectively. Pigments alone did not show considerable SPF. Extracts of *A. vera* and *C. sativus*, which are known to contain substances that protect the plant against UV has SPF values of 0.1±0.02 and 0.2±0.03, respectively. Addition of violacein to *A. vera* extract caused SPF-increase to 3.5±0.59 and upon addition to the *C. sativus* extract, caused SPF increase to 1.54±0.25. Collectively, the SPF values for pigments alone, their antioxidant activities, and their ability to increase the SPF of plant extracts constitutes evidence that the increase in sunscreen SPF upon addition of pigments relates to the photo-protective activities of the latter [14]

Violacein uses in Biosensor

In study the biosynthesis pathway of violacein was constructed heterogeneously – genes vioA, vioB, vioC, vioD, vioE from *C. violaceum* under the Pb(II)-dependent PbrR metalloregulator and its promoter, derived from the pbr operon of *Cupriavidus metallidurans*. This setup replaces traditional fluorescent or enzymatic reporters with violacein accumulation, enabling naked-eye detection via purple color change in cell pellets at high Pb(II) levels and quantification at 490 nm after butanol extraction. The high-copy plasmid pET21a with T7 lac promoter elements enhances sensitivity.
The biosensor shows high selectivity for Pb(II) over ions like Cd(II), Cu(II), Zn(II), and Hg(II) at 5 μM, even in competitive mixtures, due to PbrR's specificity. It detects Pb(II) from 0.1875–1.5 μM linearly (quantified post-extraction), outperforming prior fluorescent (LOD >10 μM) and enzymatic (>1 μM) Pb(II) sensors. Colorimetric sensitivity reaches below 0.1875 μM visually in pellets, though extraction is needed for precision.
Tested in distilled, tap, and lake water spiked with Pb(II) (0.375–1.5 μM), it reliably quantifies bioavailable Pb(II) without matrix interference. Violacein production is IPTG-inducible in a dose-response manner, confirming pathway functionality, and offers advantages like stability, antibacterial properties against Gram-positives, and simple spectrophotometry. This supports ecotoxicological monitoring, detecting below USEPA freshwater criteria (0.396 μM) but above WHO drinking water limits (0.048 μM) [18].

PRODUCTION

Production of violacein using various agricultural waste:
Agricultural wastes included sugarcane bagasse (SCB, dried and cut to 1.2-1.5 cm), solid pineapple waste (SPW), molasses (MOL), and brown sugar (BS). SCB/SPW used at 1-5% (w/v); BS at 0.4-4 g/L; MOL at 0.1-1% (v/v), all supplemented with 100 mg/L filter-sterilized L-tryptophan (pH 7.0).
Active cultures (10% v/v) were inoculated into 500-mL flasks with 50 mL medium and incubated at 30°C, 200 rpm for 24 h. For immobilized cells, SCB-packed columns (27 cm^3) were perfused with culture then L-tryptophan at 3.5 mL/min down-flow.

Medium	Concentration	Violacein Yield (g/L)
SCB 3% + 100 mg/L Trp	3 g SCB	0.82 (free cells)
NB (rich medium)	-	0.15
SCB-immobilized	Column	0.15
SPW 3% + 100 mg/L Trp	3% SPW	0.07
BS 4 g/L + 100 mg/L Trp	4 g/L BS	0.08

MOL 0.1% + 100 mg/L Trp	0.1% MOL	0.05

SCB at 3% (w/v) with L-tryptophan yielded the highest violacein (0.82 g/L), far exceeding rich media, due to lignocellulosic components (cellulose 40%, hemicellulose 27%) serving as carbon sources [19,20].
Since 1882. When C. violaceum was first reported as an isolate from wet rice paste, its most notable characteristic has been the production of deep violet pigment named violacein [18,21]

Production of violacein using soyabean meal waste

In the study, the bacterial strain *Chromobacterium violaceum* MTCC 2656 was utilized for violacein production. The soybean meal medium was prepared by grinding soybean chunks into a fine powder, adding it to deionized water at various concentrations (w/v), and sterilizing it via autoclaving.
Different commercial complex media, namely MacConkey broth, nutrient broth, soyabean casein digest medium and leuria broth, were used to compare violacein production. Varying concentration of soyabean meal medium (1%, 2% and 3% w/v) with and without supplementation of 100 mg/L L-tryptophan were prepared for production of violacein.
For extraction of crude violacein, the withdrawn culture samples were centrifuged at 10,400 g for 10 min at 37°C. The supernatant was discarded and the pellet was resuspended in an equal volume of methanol. Samples were then sonicated for 15 min, vortexed for 2 min and again centrifuged at 10,000 g for 10 min at 37°C. The absorbance of cell-free extract containing crude violacein was measured at 570 nm. The crude violacein concentration was calculated using the standard curve of violacein based upon absorbance at 570 nm wavelength
To maximize production, researchers employed the Box–Behnken design (BBD) to optimize four critical variables: pH, incubation time, substrate concentration, and L-tryptophan concentration. The study identified the ideal parameters as pH 6.09, an incubation time of approximately 114 hours and 42 minutes, a soybean meal concentration of 1.07% w/v, and 101 mg/L of L-tryptophan. Under these optimized conditions, the researchers achieved 1504.5 mg/L of crude violacein. This is noted as the highest recorded yield to date using agro-industrial waste as a substrate in a shake flask with minimal supplementation [22].

DISCUSSION AND CONCLUSION

This article discusses about the replacement of synthetic dye by the usage of bio-pigment (natural dye), seeing the wide range of micro-flora producing pigment, out of which violacein a dark violet coloured pigment with multifaceted biological activities and potential applications makes it more fascinating to work on, the data presented reveal that violacein is not merely a metabolic byproduct but a bioactive compound with significant therapeutic and industrial promise. Around 5-6 organism are there producing the pigment such as *Chromobacterium violaceum, Janthinobacterium lividum, Pseudoalteromonas species, Duganella species and Collimonas species.* From the above *C. violaceum* is most suitable for the production of the pigment, as it produces the pigment without requiring complex metabolic engineering of foreign genes. It efficiently utilizes L-tryptophan as a precursor, converting it into violacein. *C. violaceum* produces violacein via a CviI/R quorum sensing system, which means production is activated based on cell density. This regulation makes it efficient for, and responsive to, fermentation optimizations. The listed properties of violacein are Antimicrobial, Antifungal, Antiparasitic, Antiviral, Anticancer, Antioxidant, Antipyretic, Immunomodulatory.
The antimicrobial activity of violacein, especially against Gram-positive bacteria and multi-drug resistant bacteria such as *Staphylococcus aureus*, emphasizes its use as an alternative to conventional antibiotics. The mechanism of action of violacein, which involves the disruption of microbial membrane integrity, implies a reduced risk of resistance development compared to target-specific drugs. Nevertheless, its reduced efficacy against Gram-negative bacteria due to the outer membrane barrier

poses a challenge to its use as an antimicrobial agent, although this may be addressed by advances in formulation technology.

Violacein possesses significant antifungal activity against diverse range of pathogenic fungi, its efficacy against phytopathogens *Rhizoctonia solani* suggest that violacein could be developed as biocontrol agent for crop protection, offering an eco-friendly alternative to synthetic fungicides. The protective role of violacein producing *Janthinobacterium lividum* on amphibian skin against *Batrachochytrium dendrobatidis* highlights a remarkable ecological application. Studies showing combining violacein with silver nanoparticles have shown enhanced antifungal performance against *Aspergillus tamari* and *Fusarium proliferatum.* This nano-formulation approach addresses solubility limitations.

Violacein's antiparasitic activity against neglected tropical disease pathogens such as *Trypanosoma cruzi* and *Leishmania amazonensis*—with efficacy comparable to established drugs like pentamidine—highlights its potential as a candidate for treating diseases with limited therapeutic options. The observed effects on nematodes, including intestinal accumulation and inhibition of larval development in *C. elegans*, suggest additional potential applications in helminth control. The observed selectivity – greater toxicity toward parasites than mammalian cells – is particularly promising for drug development.

Violacein's selective cytotoxicity has been observed in different cancer cell lines, including glioblastoma, breast, and lung cancers, in the field of oncology. The compound's capacity to suppress metastasis-related signaling pathways, including the CXCL12/CXCR4 signaling pathway and MMP expression, indicates its potential in inhibiting the progression of cancer. Notably, its non-genotoxic properties and apoptosis induction in drug-resistant leukemia cells confirm its safety and efficacy as an anticancer agent.

In addition to its biomedical applications, violacein has great potential in industrial and environmental applications. Its stability and color make it an ideal dye for textile industry applications, having shown fastness to both physiological and environmental factors. In the cosmeceutical industry, violacein has been shown to increase the sun protection factor of commercial sunscreens, taking advantage of its ability to absorb UV radiation and act as an antioxidant. Additionally, its use in biosensors for the detection of lead is a great example of using biological pigments for environmental monitoring, offering a cost-effective and visible detection method

Production optimization using agricultural waste, such as sugarcane bagasse and soybean meal, underscores the economic and sustainable potential of violacein biosynthesis. The achievement of high yields under minimal supplementation aligns with green chemistry principles and circular economy goals.

Despite these advances, challenges remain. Violacein's poor water solubility limits its bioavailability and formulation in therapeutic contexts. Future research should focus on nano-delivery systems, prodrug strategies, or chemical modifications to enhance solubility and targeted delivery. Additionally, *in vivo* studies and clinical trials are necessary to validate preclinical findings and assess pharmacokinetics, toxicity, and therapeutic windows.

REFERENCES

[1] Rettori D, Dura N. Production, extraction and puri®cation of violacein: an antibiotic pigment produced by Chromobacterium violaceum. n.d.

[2] Deryabin DG, Inchagova KS, Nikonorova ER, Karimov IF, Duskaev GK. Quorum Sensing in Chromobacterium subtsugae ATCC 31532 (Formerly Chromobacterium violaceum ATCC 31532): Transcriptomic and Genomic Analyses. Microorganisms 2025;13. https://doi.org/10.3390/microorganisms13051021.

[3] Vasconcellos Antônio R, Creczynski-Pasa TB. Violacein biosynthesis in Chromobacterium violaceum Genetic analysis of violacein biosynthesis by Chromobacterium violaceum. vol. 3. 2004.

[4] H A Sneath BP. Cultural and Biochemical Characteristics of the Genus Chromobacterium. vol. 15. 1958.

[5] Nemer G, Louka N, Rabiller Blandin P, Maroun RG, Vorobiev E, Rossignol T, et al. Purification of Natural Pigments Violacein and Deoxyviolacein Produced by Fermentation Using Yarrowia lipolytica. Molecules 2023;28. https://doi.org/10.3390/molecules28114292.

[6] Gallardo MJ, Staforelli JP, Meza P, Bordeu I, Torres S. Characterization of Chromobacterium violaceum pigment through a hyperspectral imaging system. vol. 4. 2014.

[7] Alisjahbana B, Debora J, Susandi E, Darmawan G. Chromobacterium violaceum: A review of an unexpected scourge. Int J Gen Med 2021;14:3259–70. https://doi.org/10.2147/IJGM.S272193.

[8] Dimitrova PD, Damyanova T, Paunova-Krasteva T. Chromobacterium Violaceum: A Model for Evaluating the Anti-Quorum Sensing Activities of Plant Substances. Sci Pharm 2023;91. https://doi.org/10.3390/scipharm91030033.

[9] Bilsland E, Tavella TA, Krogh R, Stokes JE, Roberts A, Ajioka J, et al. Antiplasmodial and trypanocidal activity of violacein and deoxyviolacein produced from synthetic operons. BMC Biotechnol 2018;18. https://doi.org/10.1186/s12896-018-0428-z.

[10] Durán M, Ponezi AN, Faljoni-Alario A, Teixeira MFS, Justo GZ, Durán N. Potential applications of violacein: A microbial pigment. Medicinal Chemistry Research 2012;21:1524–32. https://doi.org/10.1007/s00044-011-9654-9.

[11] Durán N, Nakazato G, Durán M, Berti IR, Castro GR, Stanisic D, et al. Multi-target drug with potential applications: violacein in the spotlight. World J Microbiol Biotechnol 2021;37. https://doi.org/10.1007/s11274-021-03120-4.

[12] Durán N, Justo GZ, Durán M, Brocchi M, Cordi L, Tasic L, et al. Advances in Chromobacterium violaceum and properties of violacein-Its main secondary metabolite: A review. Biotechnol Adv 2016;34:1030–45. https://doi.org/10.1016/j.biotechadv.2016.06.003.

[13] Durán N, Justo GZ, Ferreira C V., Melo PS, Cordi L, Martins D. Violacein: properties and biological activities. Biotechnol Appl Biochem 2007;48:127–33. https://doi.org/10.1042/ba20070115.

[14] Suryawanshi RK, Patil CD, Borase HP, Narkhede CP, Stevenson A, Hallsworth JE, et al. Towards an understanding of bacterial metabolites prodigiosin and violacein and their potential for use in commercial sunscreens. Int J Cosmet Sci 2015;37:98–107. https://doi.org/10.1111/ics.12175.

[15] Antonisamy P, Ignacimuthu S. Immunomodulatory, analgesic and antipyretic effects of violacein isolated from Chromobacterium violaceum. Phytomedicine 2010;17:300–4. https://doi.org/10.1016/j.phymed.2009.05.018.

[16] Choi SY, Lim S, Yoon K hye, Lee JI, Mitchell RJ. Biotechnological Activities and Applications of Bacterial Pigments Violacein and Prodigiosin. J Biol Eng 2021;15. https://doi.org/10.1186/s13036-021-00262-9.

[17] Kanade YB, Abhyankar PS, Patwardhan RB. Studies on Violacein Extracted from Chromobacterium violaceum with Its Application in Textile Dyeing. vol. 13. 2024.

[18] Hui CY, Guo Y, Liu L, Zhang NX, Gao CX, Yang XQ, et al. Genetic control of violacein biosynthesis to enable a pigment-based whole-cell lead biosensor. RSC Adv 2020;10:28106–13. https://doi.org/10.1039/d0ra04815a.

[19] Abu A, Malaysian H, Board R. Microbial Surfactant for Preservation of Natural Rubber Latex. n.d.

[20] Ahmad WA, Yusof NZ, Nordin N, Zakaria ZA, Rezali MF. Production and characterization of violacein by locally isolated chromobacterium violaceum grown in agricultural wastes. Appl. Biochem. Biotechnol., vol. 167, 2012, p. 1220–34. https://doi.org/10.1007/s12010-012-9553-7.

[21] De León ME, Wilson HS, Jospin G, Eisen JA. Genome sequencing and multifaceted taxonomic analysis of novel strains of violacein-producing bacteria and non-violacein-producing close relatives. Microb Genom 2023;9. https://doi.org/10.1099/MGEN.0.000971.

[22] Gohil N, Bhattacharjee G, Gayke M, Narode H, Alzahrani KJ, Singh V. Enhanced production of violacein by Chromobacterium violaceum using agro-industrial waste soybean meal. J Appl Microbiol 2022;132:1121–33. https://doi.org/10.1111/jam.15277.

IN SILICO FUNCTIONAL ANNOTATION OF HYPOTHETICAL PROTEINS IN THE HYPERTHERMOPHILIC ARCHAEON *METHANOPYRUS KANDLERI*

Ashara Mausami Pranavbhai[1,2], Alex Sunny[2], John J. Georrge[3*]
[1] *Department of Botany, Bioinformatics & CCIM, Gujarat University, Gujarat-380009, India*
[2] *Department of Bioinformatics, Christ College, Rajkot, Gujarat, India*
[3] *Department of Bioinformatics, University of North Bengal, District-Darjeeling, West Bengal-734013, India*
**Corresponding author: johnjgeorrge@gmail.com*

ABSTRACT: *Methanopyrus kandleri* AV19 is a phylogenetically distinct hyperthermophilic archaeon adapted to extreme environments; although it has a complete genome sequence, a significant portion of its proteome remains functionally uncharacterized. Thus, by functionally characterizing *M. kandleri's* unannotated proteome. To assign possible functions to previously hypothetical proteins, domain-enhanced sequence analysis, distant homology identification, and structural fold modelling can help with annotation. This method enabled the discovery of enzymes and structural motifs relevant to genome stability and thermotolerance, including proteins involved in topoisomerase activity and in cyclic 2,3-diphosphoglycerate-mediated stabilization. Analyses of secretion potential and subcellular localization also showed metabolic architecture consistent with extreme environmental adaptation. Taken as a whole, our results enhance the proteomic framework of M. kandleri and illuminate the evolution of archaea in hyperthermophilic environments.
KEYWORDS: *Methanopyrus kandleri* AV19, hypothetical proteins, in silico functional annotation, structural fold modelling, thermotolerance, and genome stability

1. INTRODUCTION

Methanopyrus kandleri AV19 is a hyperthermophilic archaeon adapted to extreme environments such as deep-sea hydrothermal vents. It grows optimally at temperatures between 80-100°C under anaerobic conditions using hydrogen and carbon dioxide. The organism is rod-shaped, with polar flagella, and possesses a unique membrane composed of terpenoid lipids that contributes to its thermal stability. It also maintains high intracellular levels of cyclic 2,3-diphosphoglycerate, a molecule known to stabilize proteins under extreme heat. The structural morphology is rod-shaped, and M. kandleri AV19 is adapted for survival in extreme environments. The genome of *M. kandleri* consists of a single circular chromosome of approximately 1.69 Mb with high GC content, reflecting adaptation to harsh conditions. It encodes around 1,692 protein-coding genes and 39 RNA genes. A significant portion of these proteins is classified as hypothetical, indicating that their functions are still unknown. This highlights the gap between genome sequencing and functional understanding (Jensen et al., 2003; Kurr et al., 1991; Saralov, 2019).

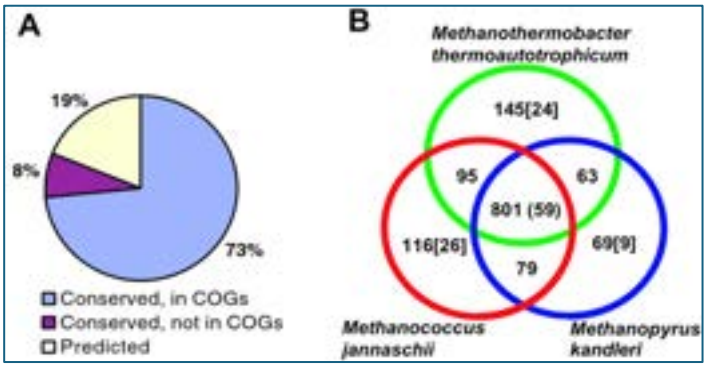


Figure 1: Distribution of conserved and unique proteins in archaeal methanogens

Comparative genome analysis shows that while many proteins belong to conserved functional groups, a considerable number lack homologs in existing databases. These uncharacterized proteins are often grouped in specific regions rather than being randomly distributed, suggesting their involvement in specialized

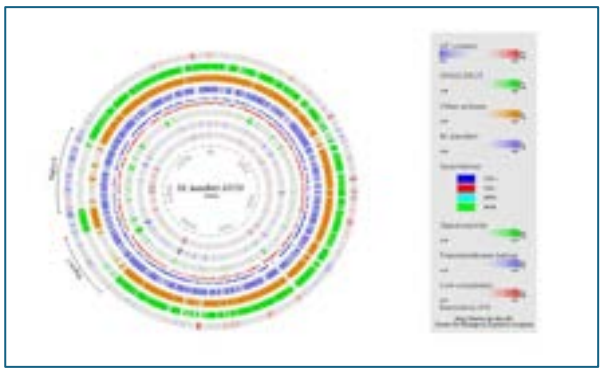


Figure 2: Atlas of the entire M. kandleri genome

biological processes. Figure 1 (Slesarev et al., 2002) indicates a Venn diagram showing the shared distribution of conserved and unique protein families among archaeal methanogens, highlighting the proportion of uncharacterized proteins in *M. kandleri*. Genome organization further supports this observation, where clusters of unknown genes may represent novel pathways or protein complexes linked to adaptation in extreme environments. Figure 2 shows the overall genomic architecture, a circular genome map showing gene organization and clustering of functionally unknown regions and gene distribution across the chromosome (Jensen et al., 2003). At the cellular level, *M. kandleri* exhibits several adaptations that support survival under extreme conditions. Its membrane contains archaeal lipids that maintain stability at high temperatures. It also possesses a unique histone-like protein (HMK) that helps in DNA packaging under stress conditions (Pavlov, Cherny, Nazimov, Slesarev, & Subramaniam, 2002). Enzyme activity in this organism is strongly influenced by intracellular salt concentration, which contributes to protein stability and function. Additionally, structural components such as pseudomurein provide resistance against osmotic stress (Pavlov, Cherny, Jovin, & Slesarev, 2002; Schacherl, Waltersperger, & Baumann, 2013; Slesarev, Belova, Kozyavkin, & Lake, 1998).

Metabolically, *M. kandleri* is a strict anaerobic methanogen that produces methane from hydrogen and carbon dioxide. It relies entirely on inorganic carbon sources, classifying it as a chemolithotrophic autotroph. These metabolic and structural features allow it to survive in high-temperature and high-pressure environments. The presence of topoisomerase V, a unique enzyme known for its ability to function under extreme conditions. Its high stability and efficiency make it useful in DNA-related applications, especially in sequencing technologies. Such adaptations highlight the biotechnological potential of extremophiles (Rajan, Osterman, Gast, & Mondragón, 2014).

A large fraction of proteins in *M. kandleri* are classified as hypothetical proteins, meaning their existence is predicted, but their functions are not experimentally confirmed. These proteins are typically identified during genome annotation when open reading frames lack similarity to known proteins. Despite this, many contain conserved domains, suggesting possible biological roles. Understanding these hypothetical proteins is essential, as experimental characterization of all genes (Majumder, Patra, & Mandal, 2006).

2. MATERIALS AND METHODS

An *in silico* pipeline was used to functionally annotate hypothetical proteins from *Methanopyrus kandleri* AV19. Protein sequences labelled as “hypothetical” were retrieved from the genome dataset and used as input for downstream analysis. Sequence similarity search was first performed using DELTA-BLAST (Boratyn et al., 2012) to identify homologous proteins. Each protein sequence was submitted individually, and hits with significant similarity were selected based on alignment score, query coverage, and E-value thresholds. This step helped in identifying distant homologs that are often missed by conventional similarity searches. To identify conserved domains and functional motifs, protein sequences were analysed using Pfam (Mistry et al., 2021), InterProScan (Jones et al., 2014), and CDD (Marchler-Bauer et al., 2015). FASTA sequences were submitted, and predicted domains were selected based on confidence scores and consensus across multiple databases. Only domains consistently identified by more than one tool were considered for further interpretation.

Structural characterization was carried out using Phyre2 (Powell, Islam, David, & Sternberg, 2025). Protein sequences were submitted in FASTA format, and 3D models were generated using homology-based fold recognition. Predicted models with high confidence scores (typically >90%) and good alignment coverage were selected. Secondary-structure elements and disordered regions were also noted during this step. To validate structural similarity, predicted sequences were analysed using PDB-BLAST. This allowed comparison with experimentally resolved protein structures, providing additional

support for structural and functional predictions. Functional annotation was further refined using ProtFun (Talo & Bozdag, 2025) and PFP (Ibtehaz & Kihara, 2025). These tools were used to assign probable biological roles, enzyme classification, and Gene Ontology (GO) terms based on sequence features and similarity patterns. Only high-confidence predictions were considered.

Subcellular localization of proteins was predicted using CELLO by submitting amino acid sequences and selecting archaeal classification settings. Additionally, TMHMM (Cuthbertson, Doyle, & Sansom, 2005) was used to predict transmembrane helices, helping to classify proteins as membrane-bound or soluble. The combined use of multiple tools ensured cross-validation of results and improved the reliability of predicted functions. Figure 3 shows the Stepwise workflow used to functionally annotate hypothetical proteins.

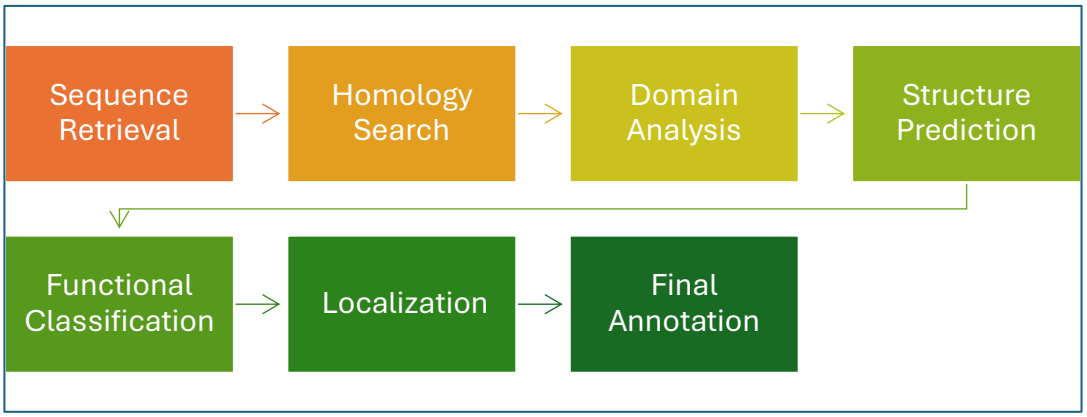


Figure 3: Workflow for in silico functional annotation of hypothetical proteins

3. RESULTS

The 629 hypothetical proteins identified in *M. kandleri* AV19 were analysed using ten different bioinformatics tools, and the results are presented here. The initial analysis using DELTA-BLAST showed that 628 out of 629 proteins produced significant matches (99.8%), as illustrated in Figure 4. This high recovery indicates that almost all proteins show detectable similarity to previously reported sequences, suggesting that most are evolutionarily conserved, even if their exact functions are not yet clearly defined. This step provided a strong base for further annotation.

Figure 4: DELTA-BLAST graphical representation of high homology detection

Figure 6: Distribution of protein domain families

Figure 5: conserved motifs and protein functional signatures

Figure 7: Conserved domain detection with CDD

these domains. Further validation using InterProScan identified functional signatures in 308 proteins, as shown in Figure 6. Although the earlier reported percentage needs correction, the overall result still confirms that a significant subset of proteins carries integrated annotations from multiple databases. This strengthens confidence in proteins supported by overlapping evidence.

Figure 8: Comparison with experimentally solved protein structures

alone is insufficient. The results show that fold-recognition methods can extend functional prediction beyond traditional sequence-based limitations. In support of this, PDB-BLAST also identified structural homologs for 628 proteins (99.8%), as shown in Figure 9. This suggests that nearly all proteins can be associated with known three-dimensional structures, adding another layer of confidence to the annotation process. Functional prediction tools ProtFun and PFP generated annotations for all 629 proteins, indicating full coverage, which is shown in Figure 10. The PFP annotates the proteins as having functional attributes, including sulfotransferase

In the next step, conserved domain identification using Pfam detected domains in 448 proteins (71.2%), as shown in Figure 5. Compared to DELTA-BLAST, the number is lower because domain-based methods require well-defined conserved regions. However, the presence of domains in a large proportion of proteins indicates that many of them contain structurally and functionally important conserved units. Pfam also identified 82 predicted active sites with a bit score of 93.8, supporting the functional relevance of

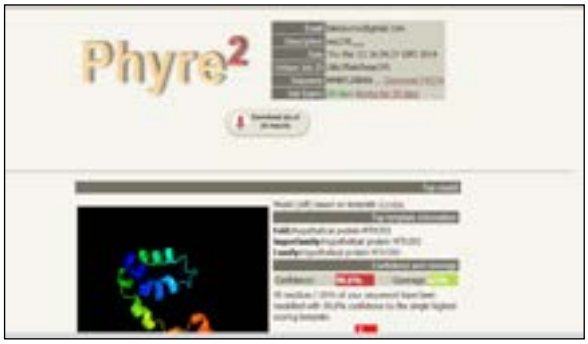


Figure 7: 3D structural modelling showing complete coverage of all proteins

Similarly, conserved domain detection with CDD identified 442 annotated proteins (70.2%), as shown in Figure 7. The close agreement between CDD and Pfam results suggests consistency across domain-based approaches and indicates that many proteins possess well-preserved functional regions. Structural prediction using Phyre2 provided models for all 629 proteins, achieving complete coverage of the dataset (Figure 8). This highlights the importance of structure-based approaches, especially for proteins where sequence similarity

Figure 9: Gene Ontology-based functional prediction using PFP

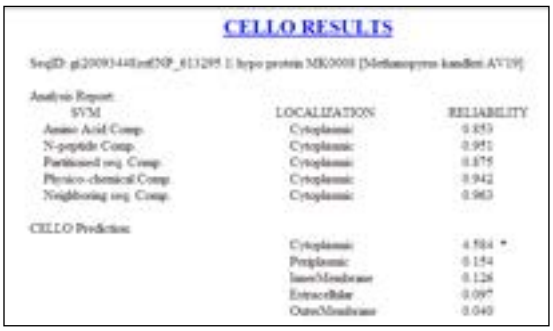


Figure 10: subcellular localization prediction using CELLO

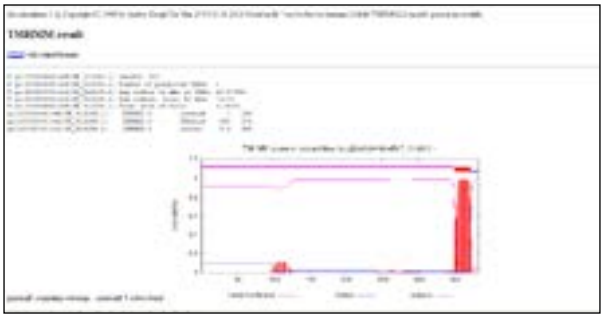


Figure 11: transmembrane topology prediction using TMHMM

activity, DNA-binding activity, and DNA-directed RNA polymerase activity, as molecular function terms. These tools were particularly useful for proteins lacking strong domain or homology support, as they rely on sequence-derived properties to suggest possible functions. Subcellular localization and membrane topology analysis using CELLO (Figure 11) and TMHMM (Figure 12) also covered the entire dataset. While these tools do not directly assign function, they provide important biological context by indicating whether proteins are soluble or membrane-associated and by suggesting potential roles in transport or membrane-related processes. Overall, the combined results from all ten tools demonstrate that no single method is sufficient to fully characterize hypothetical proteins. Instead, each tool contributed complementary information. Sequence-based methods identified evolutionary relationships, domain-based tools refined functional regions, structural prediction expanded annotation coverage, and localization analysis provided biological context.

This integrated approach significantly improved the overall annotation quality and enabled the transition of several hypothetical proteins into putative functional proteins. When the outputs from all tools were combined, confidence scores were assigned to evaluate the reliability of functional predictions. Out of the 629 proteins, 70 proteins showed more than 50% an notation support and were considered for putative functional assignment. These proteins were grouped into four categories based on confidence levels.

The highest confidence group (80%) included five proteins: DRSE-like protein, predicted membrane protein (DUF2104), NADH dehydrogenase I subunit N-related protein, Cna protein B-type domain, and a 2TM domain-containing protein (Table 1). These proteins showed strong and consistent support across multiple tools, and most of them are associated with membrane functions or energy-related processes, indicating well-conserved structural and functional features.

Table 1: High-confidence functional annotation (≥80% score

Sr no	Ac no	Final Annotation
1.	gi|20093448	DRSE like protein
2.	gi|20093904	Predicted membrane protein (DUF2104)
3.	gi|20093908	NADH dehydrogenase I, subunit N related protein
4.	gi|20094551	Cna protein B-type domain
5.	gi|20094559	2TM domain

The 70% confidence group included proteins such as phosphoribosyl transferases, CobN/magnesium chelatase, tRNA pseudo uridine synthase D, type II secretion system protein F, TRAM domain-containing protein, CrcB-like protein, RAMP superfamily protein, and calcineurin-like phosphoesterase. These proteins suggest involvement in nucleotide metabolism, cofactor biosynthesis, RNA modification, transport systems, and cellular regulation. Proteins in the 60% confidence category showed more diverse functions, including glycosyl hydrolases, Sec61beta family proteins, FecCD transporters, amidohydrolases, phospho-sulfolactate synthase (ComA), fructose-1,6-bisphosphatase, arginine decarboxylase, macro domain-containing proteins, nucleotidyltransferases, and OB-fold nucleic acid-binding proteins. These annotations indicate roles in metabolism, transport, enzymatic activity, and nucleic acid interactions. The largest group was the 50% confidence category, which

included proteins such as methyltransferases, YcaO-like proteins, archease family proteins, exosortase-related proteins, PUA-domain proteins, NurA-domain proteins, radical SAM proteins, BacA-related proteins, PrsW protease family proteins, and tRNA synthetase-related proteins. Although supported at a moderate level, these annotations still provide useful insights into possible roles in RNA metabolism, cofactor-dependent reactions, protein processing, and archaeal cellular functions.

The graphical representation of these results (Figure 13) shows that the number of proteins increases as the confidence threshold decreases. Only five proteins fall into the highest-confidence category, whereas larger numbers are observed in lower-confidence groups. This trend is expected in large-scale annotation studies, where only a small fraction of proteins receive strong consensus support. However, the identification of 70 proteins with more than 50% confidence represents a significant improvement in the functional understanding of the proteome. Figure 14 compares the performance of different tools, highlighting variations in annotation coverage and prediction consistency. Structure- and sequence-based tools showed higher coverage, while domain-based tools provided more specific but conservative annotations. So, these results confirm that a combined, multi-tool strategy is essential for reliable and comprehensive annotation of hypothetical proteins

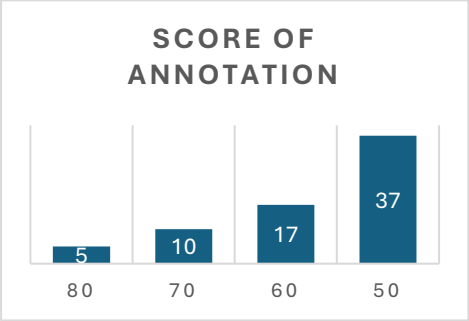


Figure 13: Graphical representation of functional annotation based on score values

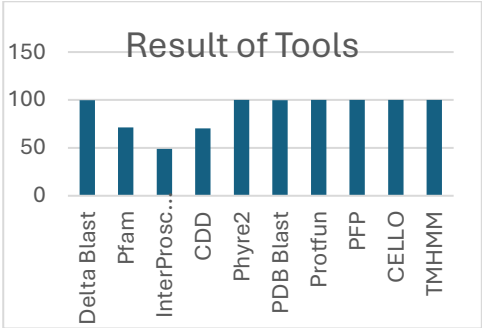


Figure 12: Comparative analysis of functional annotation across multiple tools values

4. Conclusion

This study demonstrates that integrated bioinformatics approaches can effectively assign putative functions to previously uncharacterized proteins in *Methanopyrus kandleri* AV19. Using a multi-tool annotation pipeline, a subset of proteins was confidently annotated, with over 50% of proteins functionally predicted. The identified proteins are associated with key processes such as metabolism, membrane transport, and genome stability, reflecting adaptations to extreme environments. These findings enhance the current understanding of the organism's proteome and provide a reliable foundation for future experimental validation and functional studies.

5. REFERENCES

Boratyn, G. M., Schäffer, A. A., Agarwala, R., Altschul, S. F., Lipman, D. J., & Madden, T. L. (2012). Domain enhanced lookup time accelerated BLAST. *Biology Direct, 7*(1), 12. doi:10.1186/1745-6150-7-12

Cuthbertson, J. M., Doyle, D. A., & Sansom, M. S. P. (2005). Transmembrane helix prediction: a comparative evaluation and analysis. *Protein Engineering, Design and Selection, 18*(6), 295-308. doi:10.1093/protein/gzi032 %J Protein Engineering, Design and Selection

Ibtehaz, N., & Kihara, D. (2025). Predicting Protein Functions with Function-Aware Domain Embeddings Using Domain-PFP. In L. Kurgan & D. Kihara (Eds.), *Protein Function Prediction: Methods and Protocols* (pp. 151-160). New York, NY: Springer US.

Jensen, L. J., Skovgaard, M., Sicheritz-Pontén, T., Jørgensen, M. K., Lundegaard, C., Pedersen, C. C., . . . Ussery, D. (2003). Analysis of two large functionally uncharacterized regions in the Methanopyrus kandleri AV19 genome. *BMC Genomics, 4*(1), 12. doi:10.1186/1471-2164-4-12

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., . . . Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics, 30*(9), 1236-1240. doi:10.1093/bioinformatics/btu031 %J Bioinformatics

Kurr, M., Huber, R., König, H., Jannasch, H. W., Fricke, H., Trincone, A., . . . Stetter, K. O. (1991). Methanopyrus kandleri, gen. and sp. nov. represents a novel group of hyperthermophilic methanogens, growing at 110°C. *Archives of Microbiology, 156*(4), 239-247. doi:10.1007/BF00262992

Majumder, S., Patra, M., & Mandal, C. (2006). Search for fucose binding domains in recently sequenced hypothetical proteins using molecular modeling techniques and structural analysis. *Glycoconjugate Journal, 23*(3), 251-257. doi:10.1007/s10719-006-7930-6

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., . . . Bryant, S. H. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Research, 43*(D1), D222-D226. doi:10.1093/nar/gku1221 %J Nucleic Acids Research

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, Gustavo A., Sonnhammer, E. L. L., . . . Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research, 49*(D1), D412-D419. doi:10.1093/nar/gkaa913 %J Nucleic Acids Research

Pavlov, N. A., Cherny, D. I., Jovin, T. M., & Slesarev, A. I. (2002). Nucleosome-like Complex of the Histone from the Hyperthermophile Methanopyrus Kandleri (MkaH) with Linear DNA. *Journal of Biomolecular Structure and Dynamics, 20*(2), 207-214. doi:10.1080/07391102.2002.10506836

Pavlov, N. A., Cherny, D. I., Nazimov, I. V., Slesarev, A. I., & Subramaniam, V. (2002). Identification, cloning and characterization of a new DNA-binding protein from the hyperthermophilic methanogen Methanopyrus kandleri. *Nucleic Acids Research, 30*(3), 685-694. doi:10.1093/nar/30.3.685 %J Nucleic Acids Research

Powell, H. R., Islam, S. A., David, A., & Sternberg, M. J. E. (2025). Phyre2.2: A Community Resource for Template-based Protein Structure Prediction. *Journal of Molecular Biology, 437*(15), 168960. doi:https://doi.org/10.1016/j.jmb.2025.168960

Rajan, R., Osterman, A. K., Gast, A. T., & Mondragón, A. (2014). Biochemical Characterization of the Topoisomerase Domain of Methanopyrus kandleri Topoisomerase V*. *Journal of Biological Chemistry, 289*(42), 28898-28909. doi:https://doi.org/10.1074/jbc.M114.590711

Saralov, A. I. (2019). Adaptivity of Archaeal and Bacterial Extremophiles. *Microbiology, 88*(4), 379-401. doi:10.1134/S0026261719040106

Schacherl, M., Waltersperger, S., & Baumann, U. (2013). Structural characterization of the ribonuclease H-like type ASKHA superfamily kinase MK0840 from Methanopyrus kandleri. *Acta Crystallographica Section D, 69*(12), 2440-2450. doi:doi:10.1107/S0907444913022683

Slesarev, A. I., Belova, G. I., Kozyavkin, S. A., & Lake, J. A. (1998). Evidence for an early prokaryotic origin of histones H2A and H4 prior to the emergence of eukaryotes. *Nucleic Acids Research, 26*(2), 427-430. doi:10.1093/nar/26.2.427 %J Nucleic Acids Research

Slesarev, A. I., Mezhevaya, K. V., Makarova, K. S., Polushin, N. N., Shcherbinina, O. V., Shakhova, V. V., . . . Kozyavkin, S. A. (2002). The complete genome of hyperthermophile Methanopyrus kandleri AV19 and monophyly of archaeal methanogens. *99*(7), 4644-4649. doi:doi:10.1073/pnas.032671499

Talo, M., & Bozdag, S. (2025). ProtFun: a protein function prediction model using graph attention networks with a protein large language model. *Bioinformatics Advances, 5*(1). doi:10.1093/bioadv/vbaf245

FUNCTIONAL ANNOTATION AND ENRICHMENT ANALYSIS OF HYPOTHETICAL PROTEINS IN THE *CRYPTOCOCCUS GATTII* WM276 SECRETOME

Ashara Mausami Pranavbhai[1], Aparna Gupta[2], John J. Georrge[2,3*]

[1] *Department of Botany, Bioinformatics & CCIM, Gujarat University, Gujarat-380009, India*
[2] *Department of Bioinformatics, Christ College, Rajkot, Gujarat, India*
[3] *Department of Bioinformatics, University of North Bengal, District-Darjeeling, West Bengal-734013, India*
**Corresponding author: johnjgeorrge@gmail.com*

ABSTRACT: *Cryptococcus gattii* is an opportunistic fungal pathogen responsible for severe pulmonary and neurological diseases. Despite being fully sequenced, a substantial portion of the WM276 proteome remains classified as hypothetical, leaving a critical gap in information about its biological significance. Extracellular proteins have a crucial role in the pathophysiology of *C. gattii*; however, the secretome remains incomplete due to the presence of hypothetical proteins (HPs). This study addresses this knowledge deficit by developing a pipeline to annotate functional identities for these unannotated proteins. By evaluating conserved domains, structural folds, biochemical characteristics, and downstream functional diversity, as well as protein–protein interaction (PPI) network analysis, we cluster the putative proteins into specified functional categories. The study identifies previously unknown enzymes and metabolic regulators whose interaction networks and pathway linkages indicate coordinated involvement in cellular and extracellular activities and are likely crucial to the organism's physiological resilience.

KEYWORDS: *Cryptococcus gattii* WM276 secretome, hypothetical protein functional annotation, conserved domain and structural fold analysis, protein-protein interaction network analysis, enzymes, and metabolic regulators

1.INTRODUCTION

Cryptococcus is a fungal genus that has approximately 30 species, among which *Cryptococcus neoformans*, *Cryptococcus gattii*, *Cryptococcus albidus*, and *Cryptococcus uniguttulatus* are pathogenic. *Cryptococcus gattii* is typically observed as single or budding yeast cells with round, oval, or sometimes cylindrical morphology enclosed within a prominent polysaccharide capsule. This organism, formerly known as C. neoformans var. gattii, is now recognized as a distinct pathogenic species that causes cryptococcosis. It is generally found in Asia, Africa, Australia, South America, and parts of North America. It has significant clinical importance because of its capacity to infect in conditions like immunocompetent and immunocompromised (Ballard et al., 2025; Cao, 2025), including recipients of organ transplants and patients with hematological malignancies. The infection can also affect organs such as the skin, lungs, and central nervous system (CNS), often leading to lesions and chronic disease symptoms. It has a tendency to be more virulent in nature and easily infect healthy hosts than *C. neoformans* (Byrnes, Bartlett, Perfect, & Heitman, 2011).

Cryptococcus species are characterized by a polysaccharide capsule composed of glycoproteins and exhibiting a gelatinous consistency. This capsule is of paramount importance for the cells' environmental persistence and their interactions with the host. The organism can survive in environmental reservoirs such as soil and trees, where it produces infectious propagules, including desiccated yeast cells and spores. There are more observations that C. gattii is found in different parts of the world, both in the environment and in clinical environments. This means that the organism is spreading more than it used to. In addition, *C. gattii* employs several mechanisms to evade host immune responses, including interference with neutrophil migration (Chaturvedi & Chaturvedi, 2011), thereby facilitating infection even in otherwise healthy individuals. Figure 1 shows how structural and clinical features are contained, how the life cycle of infectious diseases is transmitted, and common characteristics (Dixit, Carroll, & Qureshi, 2009; Peltak & Steen, 2025).

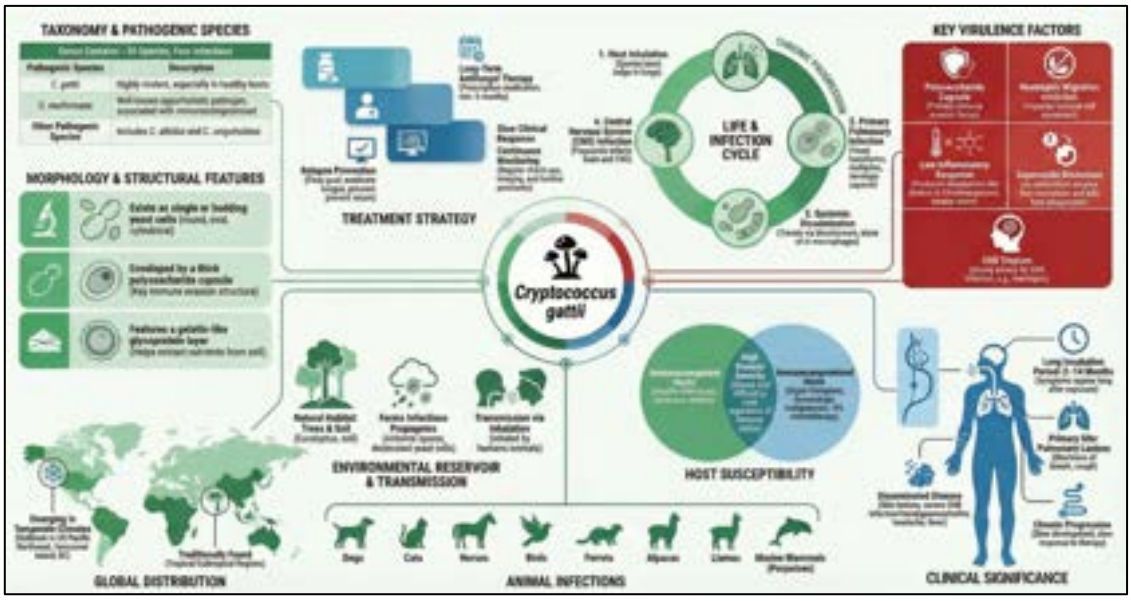


Figure 1: *Overview of morphology, taxonomy, pathogenicity, virulence, host susceptibility, clinical features, and treatment of Cryptococcus gattii*

2. MATERIALS AND METHODS

2.1 Retrieving of Hypothetical Proteins: The proteome of *C. gattii* WM276 was retrieved from the NCBI RefSeq database (O'Leary et al., 2024). The database is a comprehensive, non-redundant collection of reference sequences that includes genomic, transcript, and protein data. RefSeq also provides high-quality, standardized annotations, making it a reliable resource for whole-organism comparative and functional genomics studies. Protein sequences associated with *C. gattii* WM276 were obtained and subsequently filtered to exclude those designated as hypothetical proteins. This was accomplished by removing entries whose functional descriptions included the phrase "hypothetical protein".

2.2 Sequence Similarity Search: To find homologous sequences and deduce possible functional relationships, Position-Specific Iterated BLAST (PSI-BLAST) (Ding, Li, Shi, & Yan, 2014) was used to perform sequence similarity analysis of the screened hypothetical proteins. By calculating with a position-specific scoring matrix (PSSM) from initial alignments and iteratively improving it to find distant homologs, PSI-BLAST increases sensitivity. Using default parameters, searches were performed against the non-redundant (nr) protein database, and iterations were carried out until convergence. To identify significant matches, we used E-value thresholds and sequence identity. This approach helped us identify conserved evolutionary relationships and enabled preliminary functional annotations.

2.3 Domain Analysis: Further functional domain identification could be done using InterProScan (Tiittanen, Holm, & Törönen, 2022) and the Conserved Domain Database (CDD) (Wang et al., 2022) to detect conserved motifs and structural domains within the hypothetical proteins. InterProScan integrates multiple databases, such as Pfam (James, Nelson, Masel, & evolution, 2023), SMART (Letunic & Bork, 2025), and PROSITE (Sigrist et al., 2025), which provide annotations of protein signatures. CDD was used to validate and complement domain predictions by identifying conserved functional regions based on curated alignments. So, protein classification into families and functional groups was made possible by integrating these tools.

2.4 Secondary Structure Prediction: Secondary structure prediction was performed using the Phyre2 (Powell, Islam, David, & Sternberg, 2025) server, which uses advanced homology detection and fold recognition methods. So, this approach is based on profile-profile matching combined with secondary-structure prediction to generate reliable structural models. This analysis will interpret into α-helices, β-

sheets, and coil regions, which are critical for understanding protein stability and function, especially in the absence of experimentally resolved structures.

2.5 Three-Dimensional Homology Modeling: 3D homology modeling was performed by using BLAST searches against the Protein Data Bank (PDB) (Singh & Raghava, 2016). It used to identify structurally characterized homologs against the PDB. Templates with significant sequence similarity were selected to infer structural features of the hypothetical proteins. The alignment of query sequences with known structures enabled the prediction of tertiary conformations, which are essential for understanding protein function, ligand-binding potential, and structural conservation.

2.6 Protein Function Prediction: The functional annotation can be done using Pfam, ProtFun (Talo & Bozdag, 2025), and Protein Function Prediction (PFP) (Khan, Wei, Chapman, Kc, & Kihara, 2015). Pfam uses hidden Markov models to identify protein families and domains that are conserved across many species. ProtFun used sequence features and machine learning to predict functional classes, including enzyme activity, cellular functions, and metabolic roles. Additionally, PFP assigned Gene Ontology (GO) terms by combining sequence similarity information with functional relationships.

2.7 Virulence Factor Prediction: The virulence factor of the annotated proteins was predicted using VirulentPred (Sharma, Garg, Ramana, & Gupta, 2023), VICMpred (Saha, Raghava, & bioinformatics, 2006). These tools were developed using support vector machine (SVM)- based methods and other machine learning algorithms, trained on datasets experimentally validated to distinguish virulent from non-virulent proteins. Predictions were derived from sequence composition, physicochemical characteristics, and conserved patterns linked to pathogenicity. The analysis helped to identify proteins likely involved in interactions between the host and the pathogen, as well as the progression of the disease.

2.8 Subcellular Localization Prediction: It was predicted using TMHMM (Lu et al., 2004), WoLF PSORT (Horton et al., 2007), and PSORTdb (Lau et al., 2020) to identify the probable cellular compartments/location of the proteins. TMHMM uses hidden Markov models to predict transmembrane helices, thereby providing information about membrane-associated proteins. WoLF PSORT predicted localization by analyzing sorting signals and amino acid composition, whereas PSORTdb provided curated localization data to enhance accuracy. These predictions are essential for correlating protein location with functional roles for secretome analysis.

2.9 Functional Clustering and Categorization: Functional clustering of the annotated proteins was executed based on InterProScan-derived annotations and corresponding to functional annotations. Proteins were grouped into specific categories such as enzymes, transporters, regulatory proteins, and hypothetical conserved proteins based on shared domains and predicted biological roles. This classification provides a systematic understanding of the functional diversity within the dataset and facilitates downstream enrichment analysis.

2.10 Pathway Enrichment Analysis: It was done with the annotated proteins using the KEGG database (Joshi, Sharma, MacKinnon, & Masakapalli, 2020) to identify their involvement in biological pathways and cellular processes. The functionally annotated proteins were mapped to KEGG Orthology (KO) identifiers, thereby allowing their classification into specific metabolic and regulatory pathways. The analysis focused on identifying pathways associated with core cellular functions, including metabolism, genetic information processing, and signal transduction. Pathway enrichment, including amino acid metabolism, energy metabolism, and stress-response signaling, was evaluated to understand the biological significance of the annotated proteins and to examine their associations with conserved fungal pathways, providing insights into their evolutionary and functional relevance. This approach enabled systematic characterization of the functional roles of hypothetical proteins and facilitated the identification of key pathways potentially involved in cellular regulation, environmental adaptation, and pathogenicity in *C. gattii* WM276.

3. RESULTS AND DISCUSSION

The complete proteome of *C. gattii* WM276 was analyzed, and 1000 randomly selected hypothetical proteins were subjected to a systematic functional annotation pipeline focused on identifying secretome-relevant components. Functional predictions were analyzed, and each was assigned an equal contribution toward a cumulative annotation score. A confidence threshold of ≥50%, corresponding to agreement among at least five tools, was used to define reliable functional assignments. Using this consensus-based methodology, 235 proteins were annotated with high confidence; of these, 24 proteins showed complete concordance across all tools, achieving a 100% annotation score. The distribution of annotation confidence, as illustrated in Figure 2, indicates that a subset of proteins achieved high reliability through multi-tool agreement, reflecting robustness in prediction accuracy. Furthermore, the comparative analysis presented in Figure 3, conducted before and after annotation, shows a significant decrease in uncharacterized proteins, demonstrating the effectiveness of the pipeline used in resolving previously unannotated proteome regions.

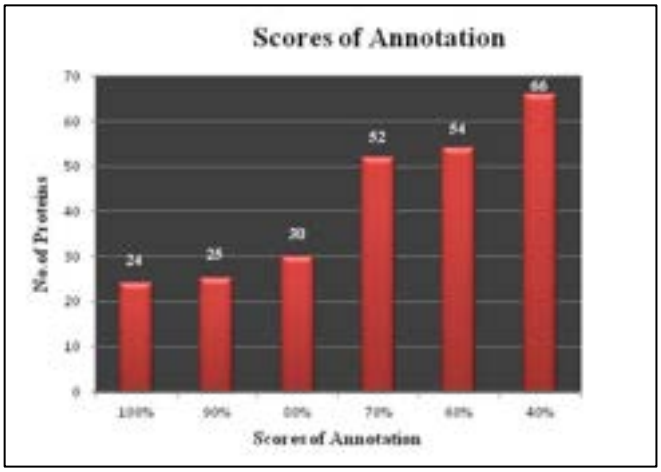


Figure 4: *Annotated proteins scoring (%) - 24 proteins achieved 100% agreement across 10 tools*

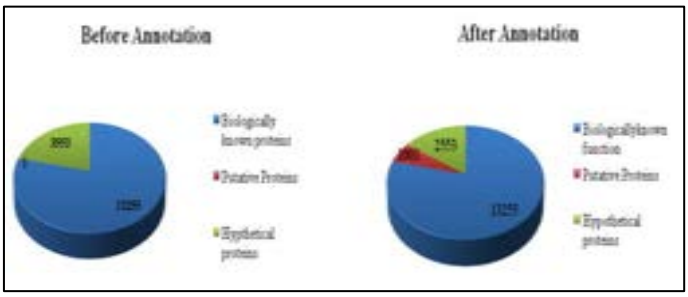


Figure 3: *Comparative statistical analysis of protein dataset before and after functional annotation*

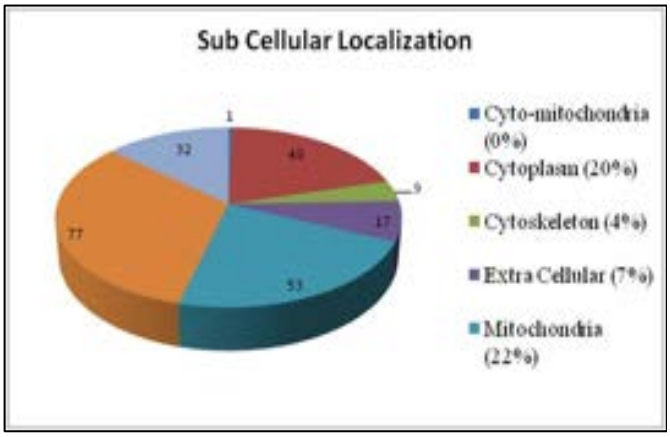


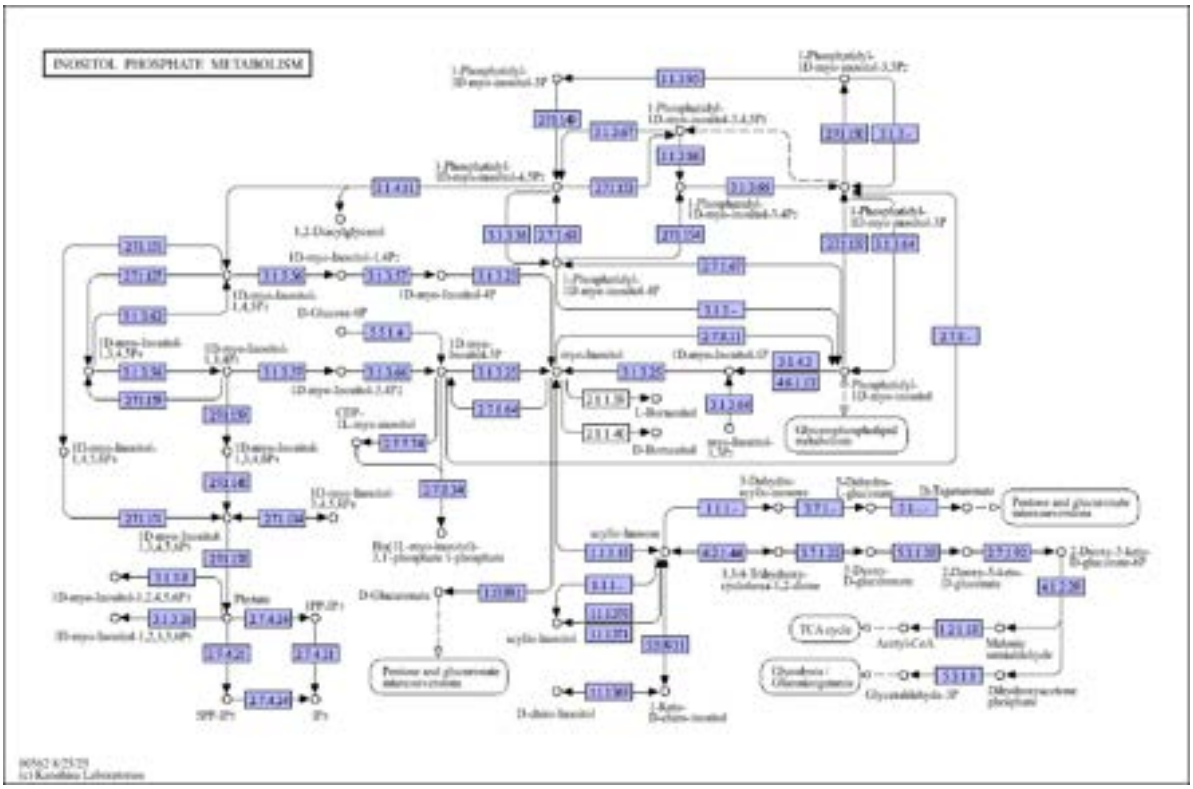
Figure 5: *Subcellular localization profile of annotated hypothetical proteins*

The functional classification of the annotated proteins statistics is illustrated in Figure 4, indicating that metabolic processes predominated (41%), followed by genetic and environmental information processing. This distribution suggests that a significant proportion of hypothetical proteins contribute to essential cellular functions, including biosynthesis, energy metabolism, and regulatory mechanisms. Domain and structural feature analyses further supported these findings, highlighting the presence of enzyme-associated domains, including peptidases, dehydrogenases, and kinases, as well as conserved structural folds validated by homology-based approaches. Additionally, DNA/RNA-binding motifs were identified, indicating major roles in transcriptional regulation and genome maintenance.

Protein subcellular localization profiling is shown in Figure 5, indicating that cytoplasmic proteins constitute the dominant class, consistent with their involvement in core metabolic activities. However, a substantial proportion of membrane-bound and secretory proteins was also identified, suggesting functional roles in transport, signaling, and extracellular interactions. The presence of secretory proteins is particularly significant in the context of the *C. gattii* secretome, as these proteins are likely to participate in host-pathogen interactions, environmental sensing, and pathogenic processes. Nuclear-localized proteins further indicate regulatory functions associated with gene expression and cellular adaptation.

Figure 6: *MAPK Pathway - Stress Response and Secretome Regulation*

Virulence-associated analysis was done with multiple prediction tools, which identified proteins linked to its pathogenicity. Notably, the major proteins are metalloproteases, which have been identified as potential contributors to host invasion. The stress response and survival-related proteins indicate adaptive mechanisms under hostile conditions. Additionally, transporter proteins associated



with pathogenic fitness were detected, suggesting their role in nutrient acquisition and survival within the host environment. The consensus observed across multiple virulence prediction platforms enhances confidence in these proteins as potential virulence determinants. From a relational perspective, the pathway enrichment analysis provided further information on the biological roles of the annotated proteins.

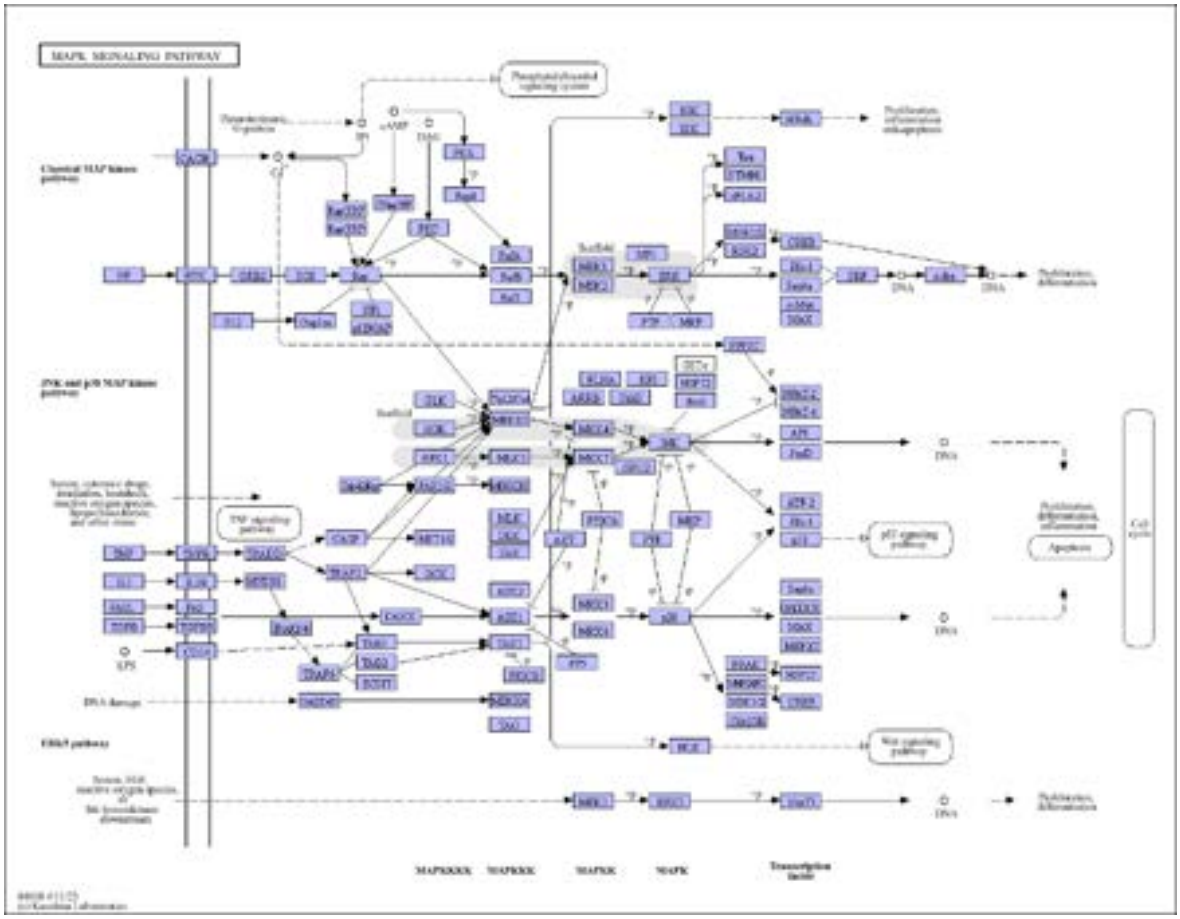


Figure 7: Inositol Phosphate Metabolism - Cellular Signaling and Membrane Dynamics

Proteins were identified as significantly associated with amino acid metabolism pathways, indicating their contribution to biosynthetic and energy-producing processes essential for cellular maintenance. The predominance of energy metabolism pathways further supports the involvement of these proteins in sustaining cellular activity and growth. In addition, proteins were mapped to conserved fungal pathways, reflecting evolutionary conservation and functional significance. As a result, the signaling-related pathways, such as the inositol phosphate metabolism pathway and the MAPK signaling pathway (Figure 6), were identified. The inositol phosphate metabolic pathway (Figure 7) is known to regulate intracellular signaling and membrane-associated processes, thereby coordinating cellular responses and maintaining membrane dynamics. The MAPK signaling pathway is a key regulatory system that maintains stress responses and environmental adaptation, enabling the organism to respond effectively to external stimuli. The identification of these pathways is particularly relevant to the secretome. Accordingly, they may regulate the expression and secretion of extracellular proteins involved in host interactions and pathogenicity. Among the annotated proteins, Carboxypeptidase Y, Peptidase M48, and Ornithine cyclodeaminase were identified as key functional examples. Carboxypeptidase Y is involved in proteolytic processing and regulatory functions, contributing to protein maturation and cellular homeostasis. Peptidase M48, a metalloprotease, is associated with metal-dependent catalytic activity and is implicated in virulence-related mechanisms, particularly in host interaction and protein degradation. Ornithine cyclodeaminase catalyzes the conversion of L-ornithine to L-proline via an NAD^+-dependent reaction, playing a role in amino acid metabolism and stress adaptation. The consistent prediction of these proteins across multiple tools highlights their functional importance and reliability. Overall, this study significantly reduced the burden of hypothetical proteins within *the C. gattii* WM276 proteome and enabled the identification of novel functional and virulence-associated proteins. The integration of domain analysis, structural validation, and pathway enrichment provides insight into protein function within the context of the secretome. Furthermore, the identification of key metabolic and signaling pathways, along with virulence-

associated factors, establishes a strong foundation for future experimental validation and functional characterization.

4. CONCLUSION

The study addresses a significant gap in proteome characterization by providing a methodical functional annotation of putative proteins in the Cryptococcus gattii WM276 secretome. 235 out of 1000 proteins were accurately annotated using a multi-tool consensus approach, including 24 with perfect agreement, indicating high confidence in the predictions. The majority of annotated proteins were correlated with stress-responsive signaling pathways, metabolism, and genetic processing. Amino acid and energy metabolism, as well as important regulatory pathways like inositol phosphate metabolism. Also, MAPK signaling was highlighted by pathway enrichment analysis, indicating roles in secretory regulation and cellular adaptation. Ornithine cyclodeaminase, Peptidase M48, and Carboxypeptidase Y were identified as functionally significant or principal contributors to virulence expression. Using all available methods, this study reduces reliance on hypothetical proteins. As a result, it provides a strong foundation for future experimental validation and the development of new treatments.

5. REFERENCES

Ballard, S., Montgomery, A., Rose, I., Lockhart, S., DeBess, E., & Bermudez, L. E. (2025). Epidemiological Study of Cryptococcus gattii Complex Infection in Domestic and Wild Animals in Oregon. *12*(2), 185.

Byrnes, E. J., Bartlett, K. H., Perfect, J. R., & Heitman, J. (2011). Cryptococcus gattii: an emerging fungal pathogen infecting humans and animals. *Microbes and Infection, 13*(11), 895–907. doi:https://doi.org/10.1016/j.micinf.2011.05.009

Cao, C. (2025). The emerging fungal pathogen Cryptococcus gattii: Epidemiology, pathogenesis, immunomodulatory attributes, and drug susceptibility. *PLOS Neglected Tropical Diseases, 19*(7), e0013245. doi:10.1371/journal.pntd.0013245

Chaturvedi, V., & Chaturvedi, S. (2011). Cryptococcus gattii: a resurgent fungal pathogen. *Trends in Microbiology, 19*(11), 564-571. doi:https://doi.org/10.1016/j.tim.2011.07.010

Ding, S., Li, Y., Shi, Z., & Yan, S. (2014). A protein structural classes prediction method based on predicted secondary structure and PSI-BLAST profile. *Biochimie, 97*, 60-65. doi:https://doi.org/10.1016/j.biochi.2013.09.013

Dixit, A., Carroll, S. F., & Qureshi, S. T. (2009). Cryptococcus gattii: An Emerging Cause of Fungal Disease in North America. *2009*(1), 840452. doi:https://doi.org/10.1155/2009/840452

Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., & Nakai, K. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Research, 35*(suppl_2), W585-W587. doi:10.1093/nar/gkm259 %J Nucleic Acids Research

James, J. E., Nelson, P. G., Masel, J. J. M. b., & evolution. (2023). Differential retention of pfam domains contributes to long-term evolutionary trends. *40*(4), msad073.

Joshi, C., Sharma, S., MacKinnon, N., & Masakapalli, S. K. (2020). Efficient System Wide Metabolic Pathway Comparisons in Multiple Microbes Using Genome to KEGG Orthology (G2KO) Pipeline Tool. *Interdisciplinary Sciences: Computational Life Sciences, 12*(3), 311-322. doi:10.1007/s12539-020-00375-7

Khan, I. K., Wei, Q., Chapman, S., Kc, D. B., & Kihara, D. J. G. (2015). The PFP and ESG protein function prediction methods in 2014: effect of database updates and ensemble approaches. *4*(1), s13742-13015-10083-13744.

Lau, Wing Yin V., Hoad, G. R., Jin, V., Winsor, G. L., Madyan, A., Gray, K. L., . . . Brinkman, F. S. L. (2020). PSORTdb 4.0: expanded and redesigned bacterial and archaeal protein subcellular localization database incorporating new secondary localizations. *Nucleic Acids Research, 49*(D1), D803-D808. doi:10.1093/nar/gkaa1095 %J Nucleic Acids Research

Letunic, I., & Bork, P. (2025). SMART v10: three decades of the protein domain annotation resource. *Nucleic Acids Research, 54*(D1), D499-D503. doi:10.1093/nar/gkaf1023 %J Nucleic Acids Research

Lu, Z., Szafron, D., Greiner, R., Lu, P., Wishart, D. S., Poulin, B., . . . Eisner, R. (2004). Predicting subcellular localization of proteins using machine-learned classifiers. *Bioinformatics, 20*(4), 547-556. doi:10.1093/bioinformatics/btg447 %J Bioinformatics

O'Leary, N. A., Cox, E., Holmes, J. B., Anderson, W. R., Falk, R., Hem, V., . . . Schneider, V. A. (2024). Exploring and retrieving sequence and metadata for species across the tree of life with NCBI Datasets. *Scientific Data, 11*(1), 732. doi:10.1038/s41597-024-03571-y

Peltak, S. N., & Steen, T. Y. (2025). An Emerging Infectious Yeast to Watch: Cryptococcus gattii Species Complex. *11*(12), 858.

Powell, H. R., Islam, S. A., David, A., & Sternberg, M. J. E. (2025). Phyre2.2: A Community Resource for Template-based Protein Structure Prediction. *Journal of Molecular Biology, 437*(15), 168960. doi:https://doi.org/10.1016/j.jmb.2025.168960

Saha, S., Raghava, G. J. G., proteomics, & bioinformatics. (2006). VICMpred: an SVM-based method for the prediction of functional proteins of Gram-negative bacteria using amino acid patterns and composition. *4*(1), 42-47.

Sharma, A., Garg, A., Ramana, J., & Gupta, D. J. P. S. (2023). VirulentPred 2.0: an improved method for prediction of virulent proteins in bacterial pathogens. *32*(12), e4808.

Sigrist, C. J. A., Cuche, B. A., de Castro, E., Coudert, E., Redaschi, N., & Bridge, A. (2025). The PROSITE database for protein families, domains, and sites. *Nucleic Acids Research, 54*(D1), D451-D458. doi:10.1093/nar/gkaf1188 %J Nucleic Acids Research

Singh, H., & Raghava, G. P. S. (2016). BLAST-based structural annotation of protein residues using Protein Data Bank. *Biology Direct, 11*(1), 4. doi:10.1186/s13062-016-0106-9

Talo, M., & Bozdag, S. (2025). ProtFun: a protein function prediction model using graph attention networks with a protein large language model. *Bioinformatics Advances, 5*(1). doi:10.1093/bioadv/vbaf245

Tiittanen, H., Holm, L., & Törönen, P. (2022). Optimizing InterProScan representation generates a surprisingly good protein function prediction method. *bioRxiv*, 2022.2008.2010.503467. doi:10.1101/2022.08.10.503467

Wang, J., Chitsaz, F., Derbyshire, M. K., Gonzales, N. R., Gwadz, M., Lu, S., . . . Marchler-Bauer, A. (2022). The conserved domain database in 2023. *Nucleic Acids Research, 51*(D1), D384-D388. doi:10.1093/nar/gkac1096 %J Nucleic Acids Research

AN *IN SILICO* APPROACH FOR TARGET IDENTIFICATION OF *ALLIUM SATIVUM* (GARLIC)

Ashara Mausami Pranavbhai[1], Jeel Anilbhai Ladhani[2], John J. Georrge[3*]

[1] *Department of Botany, Bioinformatics & CCIM, Gujarat University, Gujarat-380009, India*

[2] *G H Patel Post Graduate Department of Computer Science and Technology (GDCST), Sardar Patel University, V V Nagar, Gujarat*

[3] *Department of Bioinformatics, University of North Bengal, District-Darjeeling, West Bengal-734013, India*

**Corresponding author: johnjgeorrge@gmail.com*

ABSTRACT: Medicinal plants have been a good source of new pharmacologically active molecules. *Allium sativum* (garlic) is an aromatic herbaceous plant that is consumed worldwide as food. This species and its active components are reported to reduce the risk of diabetes and cardiovascular diseases, protect against infections by activating the immune system, and exhibit antimicrobial, antifungal, anticancer, antiviral, antidiabetic, anti-atherosclerotic, antioxidant, and anti-inflammatory activities, as confirmed by epidemiological data from human clinical studies. The diverse therapeutic effects of molecules and their targets that underlie their pharmacological activity remain incompletely understood. In the context of increasing drug resistance and the limitations of single-target therapeutics, computational strategies offer an effective approach to systematically resolving these interactions. An integrated *in silico* approach was applied to analyse garlic-derived phytochemicals using chemical clustering, followed by target fishing and reverse docking approaches based on pharmacophore models. This strategy enabled the identification of high-affinity protein targets and functional interactions, which were further refined by assessing drug-likeness and metabolic stability.

KEYWORDS: Allium sativum phytochemicals, in silico target identification, reverse docking, pharmacophore-based target fishing, drug likeness, and metabolic stability

1. INTRODUCTION

Allium sativum (garlic) is a globally consumed medicinal plant known for both nutritional and therapeutic properties. Originating in Central Asia, it has been utilised since ancient times to treat infections, cardiovascular diseases, metabolic disorders, and inflammatory conditions. The research shows that garlic has antioxidant, antibacterial, anti-inflammatory, anticancer, antiviral, and immunomodulatory effects. These effects are primarily mediated by reduced oxidative stress, inhibition of lipid peroxidation, modulation of immunological signalling, and suppression of pro-inflammatory cytokines (Donma & Donma, 2020). Their phytochemicals demonstrate protective effects across numerous organ systems, including the liver, brain, kidney, and heart. Also known to lower cholesterol and triglyceride levels, limit platelet aggregation, and improve cardiovascular function. Additionally, garlic has broad-spectrum antibacterial activity, including effects against multidrug-resistant organisms (Magryś et al., 2021), and has shown synergistic interactions with conventional antibiotics (Donma & Donma, 2020).

The specific molecular targets responsible for these pharmacological activities remain poorly characterised, limiting their systematic application in drug discovery. Botanically, garlic belongs to the genus *Allium* (family *Amaryllidaceae*) and is a diploid, bulb-forming perennial plant, comprised of numerous cloves encased in a tunic. It is farmed worldwide due to its adaptability to varied environmental conditions. Figure 1 illustrates the structural characteristics of garlic, including the bulb (cloves), flat leaves, and flowering stem (Bin et al., 2020), emphasising important morphological elements for the storage and extraction of phytochemicals (Dorrigiv et al., 2020). It is widely used as both a culinary ingredient and a traditional remedy. It has been used to treat respiratory infections, gastrointestinal disorders, hypertension, and skin conditions. Clinical and experimental evidence indicate its role in lowering blood pressure, preventing atherosclerosis, inhibiting microbial growth, and

suppressing tumour progression through mechanisms such as free radical scavenging, enzyme modulation, and regulation of apoptosis (Magryś et al., 2021).

Figure 1: Morphology of Allium sativum

The therapeutic potential of garlic is primarily attributed to its diverse phytochemical composition. These bioactive compounds act synergistically, contributing to their wide-ranging pharmacological effects. Major classes include flavonoids, terpenoids, organosulfur compounds, phenolic acids, and other specialised metabolites (Singh & Kumar, 2017). Table 1 summarises representative compounds from each phytochemical class along with their primary biological roles, providing a consolidated overview of garlic's chemical diversity (El-Saber Batiha et al., 2020).

Table 1: Major phytochemical classes of Allium sativum and their key biological activities

Phytochemical Class	**Representative Compounds**	**Key Biological Activities**
Flavonoids	Quercetin, Apigenin, Luteolin	Antioxidant, anti-inflammatory, neuroprotective
Terpenoids & Monoterpenoids	Limonene, Citral, Nerolidol	Antimicrobial, anti-inflammatory
Organosulfur Compounds	Allicin, Alliin, Ajoene	Anticancer, antimicrobial, cardioprotective
Phenolic Acids	Coumaric acid, Vanillic acid	Antioxidant, anti-inflammatory
Polyphenols	Caffeic acid	Anti-inflammatory, anticancer
Other metabolites	Allixin, Ferulic acid	Antioxidant, regulatory functions

Among these, organosulfur compounds are the most significant contributors to garlic's biological activity. Compounds such as allicin and ajoene regulate important cellular processes involved in oxidative stress, inflammation, apoptosis, and cancer growth (Pandey et al., 2023). Flavonoids and phenolic acids increase antioxidant protection and provide neuroprotective effects (Mondal & Rahaman, 2020). Terpenoids contribute to antibacterial and anti-inflammatory responses (Zhang & Hong, 2020). The combined presence of several chemicals exhibits a multi-target mode of action that traditional single-target investigations cannot fully explain. A total of 53 compounds reported from garlic were selected for computational analysis in this study. Their 2D chemical structures were retrieved from the PubChem database (Kim & Bolton, 2024) and used for downstream in silico screening.

Although garlic is safe at nutritional levels, adverse effects may occur at higher doses. These include gastrointestinal discomfort, a sense of disorientation, hypotension, and allergic responses. Excessive ingestion has been related to liver damage and anaemia, while topical use may cause skin irritation or burns. Thus, examining both efficacy and safety is critical for the development of treatments. Given the complexity of garlic's phytochemicals and structural features, and the lack of well-defined molecular targets, there is a need for systematic methodologies to elucidate its mechanism of action (Colomer et al., 2020). In silico technologies offer a powerful solution by enabling large-scale

analysis of compound target interactions. In this study, an integrated computational approach comprising chemical clustering, pharmacophore-based target fishing, reverse docking, and drug-likeness evaluation was employed to identify putative protein targets and uncover the multi-target pharmacological profile of garlic phytochemicals (Volcho et al., 2018).

2. MATERIALS AND METHODS

An *in silico* workflow was applied to identify potential molecular targets of *A. sativum* phytochemicals. The pipeline involved phytochemical data mining, chemical clustering, target prediction using multiple computational approaches, reverse docking, and ADMET evaluation. Phytochemical data were collected from the KNApSAcK (Nakamura et al., 2014), a comprehensive database containing species-metabolite relationships. Chemical structures were retrieved from the PubChem and ChEMBL databases (Zdrazil et al., 2024). A total of 65 phytochemicals associated with Allium sativum were initially compiled. These compounds were standardised and prepared for further computational analysis. Figure 2 illustrates the complete computational pipeline, including phytochemical data collection, structure retrieval, chemical clustering, target identification via target fishing and reverse docking, and ADMET evaluation.

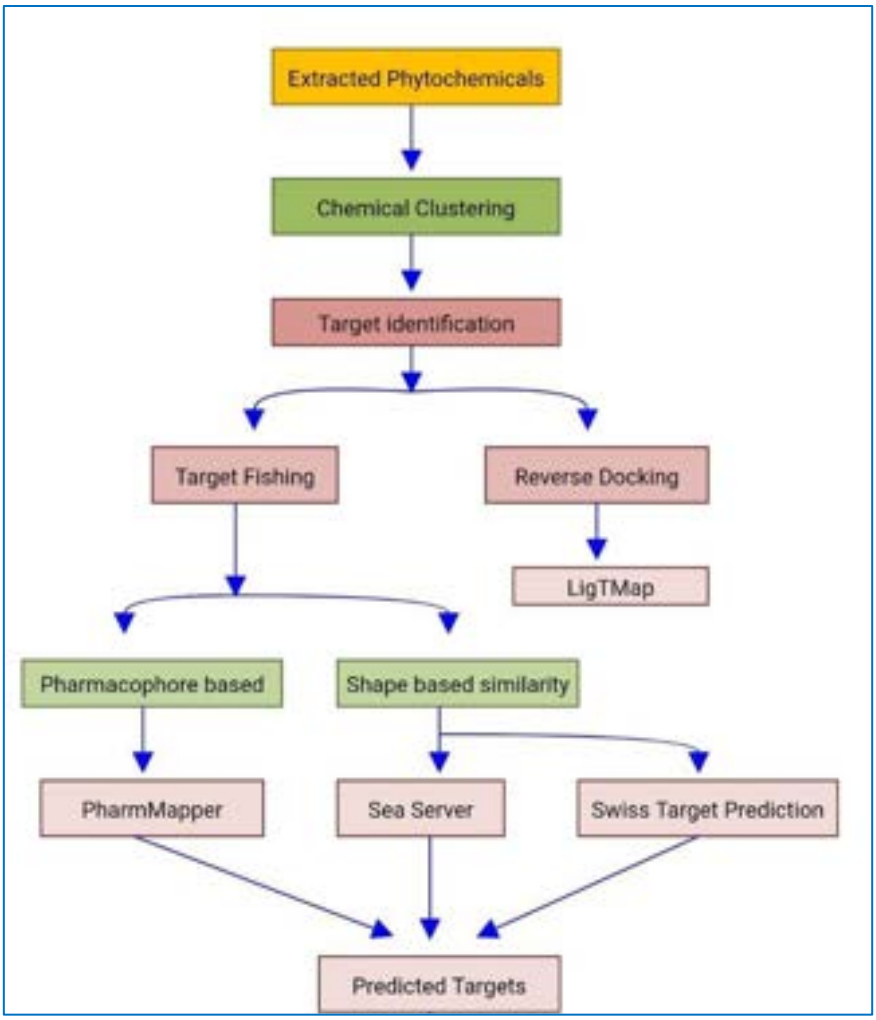


Figure 2: Workflow of target identification and computational analysis

To refine the dataset, chemical clustering was performed to identify structurally relevant and representative compounds. Clustering was carried out using the Canvas module of the Schrödinger suite (Bell et al., 2012), by hierarchical clustering based on molecular fingerprints and similarity metrics. Various linkage methods are available, and clustering can group compounds based on structural similarity and physicochemical properties. From the initial 65 compounds, 53 phytochemicals were selected as potential candidates based on clustering results.

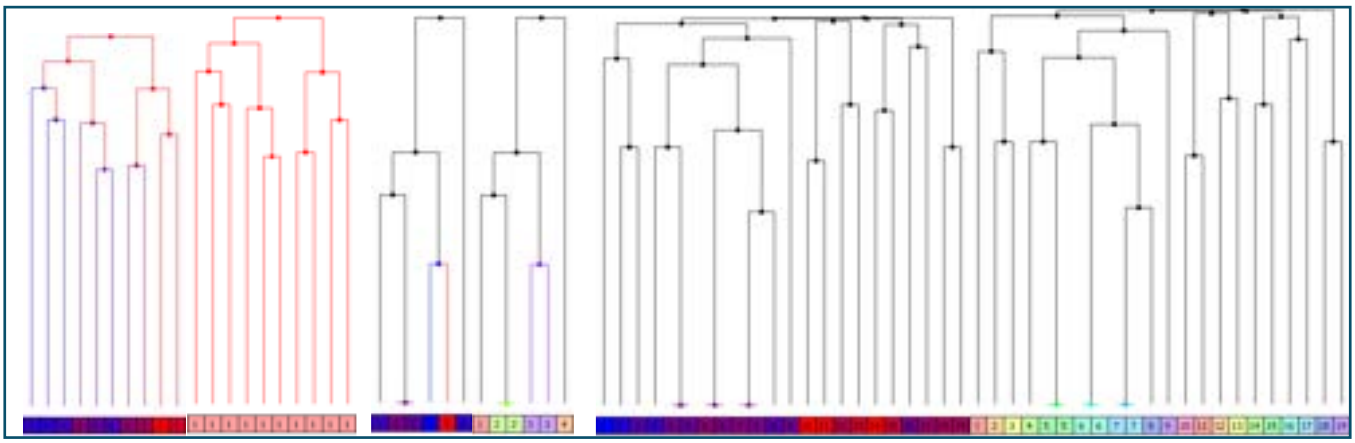
Then, target identification was performed using two complementary approaches: target fishing and reverse docking, enabling both ligand-based and structure-based predictions. For target fishing, multiple computational tools were used to improve prediction reliability. PharmMapper (Wang et al., 2017) was applied to pharmacophore-based mapping, in which each phytochemical was screened against a large database of pharmacophore models. Parameters were set to generate up to 300 conformations per molecule, and the top targets were selected based on fit scores.

The SEA (Similarity Ensemble Approach) server (Wang et al., 2016) was used to identify targets based on chemical similarity to known ligands, using SMILES as input and comparing them against the ChEMBL database. Targets were ranked based on statistical significance (p-value). PASS Targets was used to predict biological activity spectra via structure-activity relationships, generating probability scores for compound-target interactions. Swiss Target Prediction (Gfeller et al., 2014) was used to further validate predicted targets based on combined 2D and 3D similarity with known bioactive molecules. For each phytochemical, the top five targets from each tool were selected based on their respective scoring criteria to ensure consistency and reduce false positives.

Reverse docking was performed to validate and refine the predicted targets. LigTMap server (Shaikh et al., 2021) was used as it integrates ligand similarity, molecular docking, and binding mode comparison. Each phytochemical (submitted as SMILES) was docked against a set of protein targets, and results were ranked based on combined docking and similarity scores. The top five targets for each compound were selected based on binding affinity and docking scores. To evaluate pharmacokinetic properties, ADMET prediction was performed using the QikProp (Bell et al., 2012) module of Schrödinger. This analysis provided key descriptors related to absorption, distribution, metabolism, and excretion, including human oral absorption, lipophilicity, and reactivity parameters. All 53 phytochemicals were assessed, and their drug-likeness profiles were analysed based on comparison with known drug standards.

3. RESULTS AND DISCUSSION

The identification of potential bioactive phytochemicals and their relevant molecular targets, as determined through chemical clustering, target prediction, reverse docking, and ADMET analysis, is presented here. Chemical clustering was conducted to eliminate redundancy and discover structurally relevant molecules. Among the 65 initially identified phytochemicals, clustering based on molecular similarity and energy factors yielded 53 representative compounds. Classes with higher chemical density, including flavonoids, monoterpenoids, and organosulfur compounds, displayed unique clustering patterns. Molecules located in lower-energy regions were prioritised due to their greater structural stability and probable biological relevance. Figure 3 represents hierarchical clustering of (3a) flavonoids, (3b) monoterpenoids, and (3c) organosulfur compounds based on molecular similarity and energy distribution, where blue indicates low-energy stable conformations and red indicates higher energy states.



(3a) flavonoids *(3b) monoterpenoids* *(3c) organosulfur*

Figure 3: Energy-based chemical clustering of major phytochemical classes

Table 2 lists key phytochemicals from different classes after clustering, representing structurally diverse and biologically relevant compounds. ADMET analysis using QikProp revealed significant variation in pharmacokinetic properties across the phytochemicals. Most monoterpenoids and organosulfur compounds demonstrated high human oral absorption (close to 100%), indicating favourable bioavailability. In contrast, larger molecules such as glutathione and soyacerebroside showed poor absorption, suggesting limited drug-likeness. CNS activity prediction indicated that several small, lipophilic compounds, such as alpha-pinene, limonene, and gamma-terpinene, can cross the blood-brain barrier, highlighting their potential relevance in neurological disorders.

Table 2: Representative phytochemicals selected after clustering

Class	Representative Compounds
Flavonoid	Apigenin
Organosulfur	Allicin, Ajoene
Terpenoid	Nerolidol, Alpha-pinene
Monoterpenoid	Limonene, Gamma-terpinene
Phenolic acid	Coumaric acids, Vanillic acid
Others	Allixin, Ferulic acid, Caffeic acid

Compounds such as allicin and apigenin exhibited balanced ADMET profiles, with moderate absorption and acceptable drug-like properties. Table 3 highlights key ADMET parameters, including human oral absorption and CNS activity, and indicates the drug-likeness potential of the selected compounds. Target identification using multiple computational tools gave a wide range of potential protein targets associated with major disease pathways.

Table 3: Summary of ADMET properties of selected phytochemicals

Category	Observation
High oral absorption	Monoterpenoids, organosulfur compounds
Moderate absorption	Flavonoids, phenolic acids
Low absorption	Large molecules (e.g., glutathione)
CNS penetration	Small lipophilic compounds

Reverse docking using LigTMap provided binding interactions between phytochemicals and key disease-related proteins. Several compounds showed high affinity towards enzymes and receptors involved in metabolic, inflammatory, and cancer pathways. *Apigenin* showed strong interaction with Tankyrase-2 and kinase-related proteins, for potential anticancer activity. *Coumaric acid derivatives* bind carbonic anhydrase and PPAR-γ, indicating roles in metabolic and cardiovascular regulation. *Allixin and caffeic acid* interacted with enzymes such as acetylcholinesterase and MAP kinases, suggesting possible neuroprotective and anti-inflammatory effects. *Organosulfur compounds*, which show fewer docking hits, are known to act through reactive mechanisms rather than stable binding, which is reflected in the limited representation in docking results. The combined analysis of all prediction tools revealed that garlic phytochemicals are associated with multiple disease pathways, including cardiovascular diseases (*PPAR-γ, carbonic anhydrase*), cancer (*kinases, transferases, apoptosis-related proteins*), neurodegenerative disorders (*acetylcholinesterase, MAPK pathway*), and infectious diseases (*microbial enzymes and viral proteins*). This study provided that its activity is not limited to a single pathway but involves network-level modulation of biological systems. The results evaluation provides biologically relevant targets. Also, potential phytochemicals can act as multi-target agents and provide a computational basis for further experimental validation.

4. CONCLUSION

An integrated in silico pipeline applied in this study efficiently classified bioactive phytochemicals from *A. sativum* and identified their potential protein targets using combined clustering, target fishing, and reverse docking. The multi-tool consensus strategy improved target prediction reliability, while ADMET profiling enabled the selection of pharmacologically relevant candidates. The findings provide

a focused set of phytochemical-target interactions that can be directly advanced for structure-based validation and drug discovery.

5. REFERENCES

Bell, J., Cao, Y., Gunn, J., Day, T., Gallicchio, E., Zhou, Z., Levy, R., & Farid, R. (2012). PrimeX and the Schrödinger computational chemistry suite of programs. *International Tables for Crystallography F. ch. 18.10*, 534-538. https://doi.org/10.1107/97809553602060000864

Colomer, J. P., Traverssi, M., & Oksdath-Mansilla, G. (2020). Oxidation of organosulfur compounds promoted by continuous-flow chemistry. *Journal of Flow Chemistry*, *10*(1), 123-138. https://doi.org/10.1007/s41981-019-00066-5

Donma, M. M., & Donma, O. (2020). The effects of allium sativum on immunity within the scope of COVID-19 infection. *Medical Hypotheses*, *144*, 109934. https://doi.org/10.1016/j.mehy.2020.109934

Dorrigiv, M., Zareiyan, A., & Hosseinzadeh, H. (2020). Garlic (Allium sativum) as an antidote or a protective agent against natural or chemical toxicities: A comprehensive update review. *Phytotherapy Research*, *34*(8), 1770-1797. https://doi.org/10.1002/ptr.6645

El-Saber Batiha, G., Magdy Beshbishy, A., G. Wasef, L., Elewa, Y. H. A., A. Al-Sagan, A., Abd El-Hack, M. E., Taha, A. E., M. Abd-Elhakim, Y., & Prasad Devkota, H. (2020). Chemical Constituents and Pharmacological Activities of Garlic (Allium sativum L.): A Review. *Nutrients*, *12*(3), 872. https://doi.org/10.3390/nu12030872

Gfeller, D., Grosdidier, A., Wirth, M., Daina, A., Michielin, O., & Zoete, V. (2014). SwissTargetPrediction: a web server for target prediction of bioactive small molecules. *Nucleic Acids Research*, *42*(W1), W32-W38. https://doi.org/10.1093/nar/gku293

Kim, S., & Bolton, E. E. (2024). PubChem: A Large-Scale Public Chemical Database for Drug Discovery. In *Open Access Databases and Datasets for Drug Discovery* (pp. 39-66). https://doi.org/10.1002/9783527830497.ch2

Magryś, A., Olender, A., & Tchórzewska, D. (2021). Antibacterial properties of Allium sativum L. against the most emerging multidrug-resistant bacteria and its synergy with antibiotics. *Archives of Microbiology*, *203*(5), 2257-2268. https://doi.org/10.1007/s00203-021-02248-z

Mondal, S., & Rahaman, S. (2020). Flavonoids: A vital resource in healthcare and medicine. *Pharmacy & Pharmacology International Journal*, *8*(2), 91-104. https://doi.org/10.15406/ppij.2020.08.00285

Nakamura, Y., Mochamad Afendi, F., Kawsar Parvin, A., Ono, N., Tanaka, K., Hirai Morita, A., Sato, T., Sugiura, T., Altaf-Ul-Amin, M., & Kanaya, S. (2014). KNApSAcK Metabolite Activity Database for Retrieving the Relationships Between Metabolites and Biological Activities. *Plant and Cell Physiology*, *55*(1), e7-e7. https://doi.org/10.1093/pcp/pct176

Pandey, P., Khan, F., Alshammari, N., Saeed, A., Aqil, F., & Saeed, M. (2023). Updates on the anticancer potential of garlic organosulfur compounds and their nanoformulations: Plant therapeutics in cancer management [Review]. *Frontiers in Pharmacology*, *14*, 1154034. https://doi.org/10.3389/fphar.2023.1154034

Shaikh, F., Tai, H. K., Desai, N., & Siu, S. W. (2021). LigTMap: ligand and structure-based target identification and activity prediction for small molecular compounds. *Journal of Cheminformatics*, *13*(1), 44. https://doi.org/10.1186/s13321-021-00523-1

Singh, V., & Kumar, R. (2017). Study of phytochemical analysis and antioxidant activity of Allium sativum of Bundelkhand region. *International Journal of Life Sciences Research*, *3*(6), 1451-1458. https://doi.org/10.21276/ijlssr.2017.3.6.4

Volcho, K. P., Laev, S. S., Ashraf, G. M., Aliev, G., & Salakhutdinov, N. F. (2018). Application of monoterpenoids and their derivatives for treatment of neurodegenerative disorders. *Current Medicinal Chemistry*, *25*(39), 5327-5346. https://doi.org/10.2174/0929867324666170112101837

Wang, X., Shen, Y., Wang, S., Li, S., Zhang, W., Liu, X., Lai, L., Pei, J., & Li, H. (2017). PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. *Nucleic Acids Research*, *45*(W1), W356-W360. https://doi.org/10.1093/nar/gkx374

Wang, Z., Liang, L., Yin, Z., & Lin, J. (2016). Improving chemical similarity ensemble approach in target prediction. *Journal of Cheminformatics*, *8*(1), 20. https://doi.org/10.1186/s13321-016-0130-x

Zdrazil, B., Felix, E., Hunter, F., Manners, E. J., Blackshaw, J., Corbett, S., de Veij, M., Ioannidis, H., Lopez, D. M., Mosquera, Juan F., Magarinos, Maria P., Bosc, N., Arcila, R., Kizilören, T., Gaulton, A., Bento, A P., Adasme, Melissa F., Monecke, P., Landrum, Gregory A., & Leach, Andrew R. (2024). The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Research*, *52*(D1), D1180-D1192. https://doi.org/10.1093/nar/gkad1004

Zhang, C., & Hong, K. (2020). Production of Terpenoids by Synthetic Biology Approaches [Mini Review]. *Frontiers in Bioengineering and Biotechnology*, *8*, 347. https://doi.org/10.3389/fbioe.2020.00347

PROTEOME MINING OF *KLEBSIELLA* SPECIES TO IDENTIFY NOVEL BACTERIOCINS

Jensi Dobariya[1], Jinal N. Naik[2], Dimple K. Kachhadiya[1], John J. Georrge[3*]

[1] Department of Microbiology, Atmiya University, Rajkot, Gujarat, India
[2] *G H Patel Post Graduate Department of Computer Science and Technology (GDCST), Sardar Patel University, V V Nagar, Gujarat*
[3] *Department of Bioinformatics, University of North Bengal, District-Darjeeling, West Bengal-734013, India*
**Corresponding author: johnjgeorrge@gmail.com*

ABSTRACT: Currently, most research focuses on developing drugs, including antimicrobial agents and antibiotics, for various infectious diseases. Antibiotic use has increased over time, but antibiotics have adverse effects on the body. Researchers are developing new approaches to identify antimicrobial agents beyond chemical entities. Bacteriocins are peptides produced by bacteria that exhibit antimicrobial activity against pathogens. Bacteriocins are present in most of the Gram-negative as well as Gram-positive bacteria. A limited number of known biological bacteriocins are available from *Klebsiella*. This is the time to predict a novel bacteriocin from *Klebsiella*. Protein sequences of *Klebsiella* species were retrieved from NCBI, and various tools were used to predict bacteriocins, including extracellular and secretome predictions. The interacting proteins of the bacteriocin were predicted using the STRING database, which provides information on its use in inhibiting the growth of other pathogens. The three-dimensional structures of screened bacteriocins and interacting proteins were predicted using I-TASSER. To confirm interactions between the bacteriocin and proteins from other pathogens, protein-protein docking was performed using ClusPro, and the docked model was visualised in UCSF Chimaera. Extensive computational power and manual analysis are required to conclude. Finally, 3 proteins were predicted as novel bacteriocins from *Klebsiella* species that inhibit the growth of other pathogens.

KEYWORDS: *Klebsiella*, Novel bacteriocins, Proteome mining, Antimicrobial activity, Protein-protein docking

1. INTRODUCTION

The genus *Klebsiella* consists of ubiquitous, Gram-negative bacteria that thrive across a remarkably wide range of environments, from surface water and soil to the mucosal linings of mammals. This dual lifestyle gives *Klebsiella* an ecological flexibility shared by genera like *Enterobacter*. Still, it sets it apart from more host-restricted pathogens like *Escherichia coli* (Bagley et al., 1978). While these bacteria often live as harmless commensals in the human gut, species such as *K. pneumoniae*, *K. oxytoca*, and *K. rhinoscleromatis* turn into dangerous opportunistic pathogens the moment they breach sterile body sites (Nordmann et al., 2009). Armed with a diverse array of capsular (K) and lipopolysaccharide (O) antigens that actively shield them from host immunity, they cause up to 14% of nosocomial pneumonias and roughly 8% of all hospital-acquired infections in the US and Europe (Nordmann et al., 2009). In hospital settings, standard medical devices like ventilators and catheters frequently act as conduits for transmission, sparking severe outbreaks among vulnerable, immunocompromised patients.

Treating these infections has become a major clinical hurdle due to the rapid rise of Antimicrobial Resistance (AMR). Even when a strain appears vulnerable *in vitro* to classic beta-lactams or fluoroquinolones (Chan et al., 2009), using these drugs long-term is often complicated by severe patient toxicities (http://sideeffects.embl.de/). Far more alarming, though, is the global spread of carbapenemases and extended-spectrum beta-lactamases (ESBLs). These enzymes essentially dismantle our last-line antibiotics, leaving clinicians with very few working options. When you combine this enzymatic resistance with structural defences, like specific capsular serotypes that physically block the host's complement system from killing the cell, it is easy to see why mortality rates are soaring. We desperately need a new class of non-traditional antimicrobials.

Bacteriocins offer a highly promising way forward. These are ribosomally synthesised antimicrobial peptides that bacteria naturally produce to eliminate competitors. Unlike standard, broad-spectrum antibiotics, bacteriocins have a narrow, laser-focused spectrum of activity, killing closely related strains by disrupting their membrane potential or punching lethal pores through their walls (Nes et al., 2007). They fall into several distinct structural groups, ranging from small, heat-stable Class I (antibiotics) and Class II peptides to much larger, heat-labile Class III bacteriolysins (Heng et al., 2007). However, finding these compounds through traditional wet-lab screening is incredibly slow and tedious. That is where *in silico* computational biology and proteome mining change the game (Andersson & Hughes, 2010). Because bacteriocin operons are highly organised and usually sit right next to two-component signalling systems, we can use bioinformatics to rapidly scan massive genomic datasets (Dufour et al., 2007). By combining this genomic scanning with structural tools like CASTp to map out active binding pockets, we can predict exactly how these peptides will behave (Dundas et al., 2006). This study uses these targeted proteome mining strategies to systematically uncover novel bacteriocins within the genus *Klebsiella*, offering a direct path to bypass established multi-drug resistance.

2. MATERIALS AND METHODS

The objective of this research is to predict bacteriocins from Klebsiella other than the bacteriocin mentioned above. The following method is used to predict a novel bacteriocin.

2.1. Literature mining: A comparative search of all eligible studies on Bacteriocins, Klebsiella, and their pathogenic properties was conducted by searching the electronic literature (PubMed database) for relevant reports and by manually searching the reference lists of articles on this topic.

2.2. Retrieval of sequences: The protein sequences of *Klebsiella* were retrieved from the National Centre for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov).

2.3. Multidimensionality reduction: CD-HIT was used to cluster the collected protein FASTA sequences. CD-HIT is a clustering algorithm for databases with high identity and tolerance.CD-HIT uses a greedy incremental clustering algorithm. This program reduces the overall size of the database without removing any sequence information by removing only redundant sequences. Therefore, it provides a non-redundant database (www.bioinformatics.org/cd-hit/).

2.4. Protein localisation: The Protein cluster was generated from CD-HIT. That result is used as input for protein localisation. Protein localisation was performed using PSORTB. It is the most precise bacterial protein subcellular localisation (SCL) predictor. The PSORTB program analysed several features at once. It uses an algorithm that assigns a probable localisation site to a protein. It has predictive capability for Gram-negative and Gram-positive bacteria (http://www.psort.org/psortb/index.html).

2.5. Transmembrane Helix Prediction: Transmembrane proteins were predicted using TMHMM for outer membrane proteins based on the obtained PSORTB results. TMHMM is based on a hidden Markov model. TMHMM can differentiate soluble from membrane proteins with specificity and sensitivity up to 99%. TMHMM can predict transmembrane segments and their in- and out-of-membrane orientations (Krogh et al., 2001).

2.6. Secretory Protein prediction: A local BLAST (Blast+) was performed for Secretory protein prediction. The input for local BLAST is the extracellular sequences of *Klebsiella*. A local BLAST was performed using the Secretome p Database. This database was created locally by retrieving Secretory protein sequences from different Secretory protein databases.SPD has an improved secreted protein prediction approach and comprehensive data sources, including Swiss-Prot, TrEMBL, RefSeq, Ensembl and CBI-Gene (Chen et al., 2005). An online BLAST was carried out using three databases: Lactic Acid Bacteria (LAB), Fungal Secretome Database (FSD), and Fungal Secretome Knowledgebase (FunSecKB). The Fungal Secretome Database (FSD) has been developed to collect and analyse fungal secretomes from 61 species. In the FSD, there are 107859 putative proteins with signal peptides (Park et al., 2007). FSD is freely available at (http://fsd.riceblast.snu.ac.kr/index.php?a=view). FunSecKB, the Fungal Secretome knowledgebase, provides a resource of secreted fungal proteins identified from all available fungal protein data in the NCBI RefSeq database (http://proteomics.ysu.edu/secretomes/fungi.php). The Secretome P tool was also employed to predict non-classical secreted proteins using an ab initio method.

2.7. Antimicrobial peptide and Bacteriocin Prediction: The chosen Sequences after the secretory protein prediction were subjected to Blast+ against collected AMPs using 10 databases of AMPs and Bacteriocins, with different databases and algorithms, as shown in Tables 2 and 3. Amongst

them, AMPer is a database and an automated discovery tool for predicting antimicrobial peptides. It uses a Hidden Markov Model and has 99% accuracy; it is available both as a standalone and online (http://marray.cmdr.ubc.ca/cgi-bin/amp.pl). CAMP is a manually curated database and tool. Sequence information, protein definition, accession numbers, activity, the taxonomy of the source organism, target organisms with MIC values, hemolytic activity of the peptide, and links to external databases is included here. It follows Random Forests (RF), Support Vector Machines (SVM), and Discriminant Analysis, achieving 93.2%, 91.5%, and 87.5% accuracy, respectively. APD, BACTIBASE is a new, automated, web-accessible database for bacteriocin characterisation. The database provides detailed descriptions of bacteria producing various Bacteriocins, and, accordingly, predicted sequences were analysed using online tools such as CAMP and Antibp2 for AMP prediction and Bagel2 for Bacteriocin prediction. Thus, these tools and databases are supportive and will, in turn, help in the accurate identification of antimicrobial peptides.

Table 2: Database of AMPs

SR NO	Name of Database	URL
1	AMPer	http://marray.cmdr.ubc.ca/cgi-bin/amp.pl
2	APD	http://aps.unmc.edu/AP/main.html
3	CAMP	http://www.bicnirrh.res.in/antimicrobial/
4	Defensins knowledgebase	http://defensins.bii.a-star.edu.sg/
5	RAPD	http://faculty.ist.unomaha.edu/chen/rapd/index.php
6	PhytAMP	http://phytamp.pfba-lab-tun.org/main.php
7	PenBase	http://penbase.immunaqua.com/infos.php
8	DAMPD	http://www.biosharing.org/biodbcore-000019
9	YADAMP	http://www.yadamp.unisa.it/default.aspx
10	AMSDB	http://www.bbcm.units.it/~tossi/amsdb.html

Table 3: Database of bacteriocins

SR NO	Name of Database	URL
1	BACTIBASE	http://bactibase.pfba-lab-tun.org/main.php
2	BAGEL	http://bagel2.molgenrug.nl/

2.8. Pattern Construction: Pattern construction for the resultant filtered protein sequences was performed using the Pratt tool (http://web.expasy.org/pratt/). Pratt is a tool that allows users to search for patterns conserved across a set of protein sequences. Thus, three different patterns were reported for three classes of Bacteriocins, like Bacteriocin class 1(L-x (2,4)-G-[ACDGNSTV]), Bacteriocin class 2 (A-x (1, 2)-G-x (0, 1)-G), Bacteriocin class 3 (G-x (3, 5)-A-x(3)-[AGSV]).

2.9. Pattern Searching: Pattern searching was performed using the previously constructed patterns of three bacteriocin classes with the ScanProsite tool (http://prosite.expasy.org/scanprosite/). SCANPROSITE is a server used to determine the function of uncharacterised proteins. It allows scanning protein sequences against the PROSITE database. This tool also allows searching for hits by a specific motif in a protein sequence.

2.10. Blast against Human: Human Blast was performed to identify Klebsiella proteins that matched humans.

2.11. Virulent factor prediction: Virulent Factor prediction was performed for non-matching sequences with humans using different tools and algorithms, such as Virulent Factor Prediction (Virulent Pred). Virulent Pred is a bacterial virulent protein prediction method based on a bi-layer cascade support vector machine (http://203.92.44.117/virulent/index.html). Other tools used for virulent factor prediction were VFDB (http://www.mgc.ac.cn/VFs/main.htm) and VIMC pred (http://www.imtech.res.in/raghava/vicmpred/index.html).VFDB provides virulence factor

prediction searches for pathogenic bacteria. The objective of the VICMpred server is to classify the pathogenic microbial proteins into four broad functional classes.

2.12. Finding interacting protein partners: *Klebsiella* protein sequences were screened for interacting protein partners using the STRING database. STRING is a database for known and predicted protein interactions. The interaction includes direct (physical) or indirect (functional) associations. This database provides four different sources that help users derive interactions. These sources include genomic context, high-throughput experiments, coexpression, and previous knowledge bases (http://string-db.org/).

2.13. Structure predictions: Klebsiella proteins and their respective binding partners were predicted using the I-TASSER tool. The I-TASSER server is an internet service for protein structure and function prediction. It allows academic users to automatically generate high-quality 3D protein structures and biological functions of protein molecules from amino acid sequences using an ab initio modelling method.

2.14. Structure Validation: Predicted structures were validated using ANOLEA and PROCHECK. ANOLEA is the atomic empirical mean force potential (Melo, F. and E. Feytmans, 1998), used to assess the packing quality of the model. The program performs energy calculations on a protein chain, evaluating the "Non- Local Environment" (NLE) of each heavy atom in the molecule. This program provides a graphical result for each model protein. The y-axis of the plot represents the energy for each amino acid of the protein chain. Negative energy values (in green) represent a favourable energy environment. In contrast, positive values (in red) represent an unfavourable energy environment for a given amino acid. Procheck is secondary-structure validation software that generates a Ramachandran plot for each protein.

2.15. Structure refinement: The Predicted Structure was submitted for refinement by Mod refiner if it failed to meet all secondary structure validation conditions. Mod refiner software is used to refine the predicted model structure. It is an algorithm for atomic-level, high-resolution protein structure refinement that can start from a c-alpha trace, a main-chain model, or a fully automatic model (http://zhanglab.ccmb.med.umich.edu/ModRefiner/).

2.16. Docking: Docking was performed for the already-predicted structures of Klebsiella Bacteriocins and their interacting protein partners using the ClusPro tool. ClusPro software is used for protein-protein docking. It is widely used for protein-protein Docking (http://cluspro.bu.edu/home.php).

2.17. Docking Structure visualisation: Docking models predicted by ClusPro were visualised and analysed using UCSF Chimaera. It is a highly extensible program for interactive visualisation and analysis of molecular structure and related data, including density maps, submolecular assemblies, sequence alignments, docking results, trajectories, and conformational ensembles. (www.cgl.ucsf.edu/chimera/download.html).

3. RESULTS

In the first step, a total of 2, 23,539 *Klebsiella* protein sequences were collected from NCBI. The CD-HIT protein clustering resulted in a cluster of 14,754 sequences, which was given to PSORTb for Bacterial subcellular localisation. PSORTb provided 424 outer membrane and 215 extracellular proteins. The TMHMM result provided information about the outer membrane protein region, which is important for bacteriocin activity. TMHMM identified 24 sequences as outer membrane proteins.

Further Data mining was carried out for 239 extracellular proteins, and secretory proteins, antimicrobial peptides, and bacteriocins were predicted; 38 sequences were filtered out. Thus, a total of 38 sequences were identified as potentially suitable for further research in bacteriocin prediction. The pattern-matching step identified 36 sequences with bacteriocin patterns related to three classes of bacteriocins. Comparison of all 36 sequences with the TMHMM result showed that all 36 sequences are present outside the cell. The next step was a BLAST against the human database; none of the 36 sequences matched human. Virulent factor prediction was provided for 28 sequences (Table 4). All 28 sequences were screened to identify interacting protein partners, and the String database was used. 11 proteins had optimal binding partners, with a significant prediction of bacteriocin. Of those 11 proteins, 3 were optimal, having interacting protein partners from different pathogenic bacteria (Table 5).

Table 4: Filtration of proteins during each step

Steps performed	**Number of proteins in the result**
Protein Sequence of *Klebsiella* retrieved from NCBI	2,23,539
CD-HIT(clustering the sequence)	14,754
PSORTb	Outer membrane -424 Extracellular-215
TMHMM for outer membrane	Outside protein-24
Outside by TMHMM + extracellular	239
Antimicrobial peptide, secretory protein, bacteriocin prediction using respective tools	38
Comparison of the SCAN-PROSITE result with TMHMM	36
Blast against human	36
Virulence factor prediction	28
Finding interacting proteins by STRING database	11

All 11 Klebsiella proteins were used to predict bacteriocins. Still, only three were considered because they had interacting partners from other species. Once Binding partners were found, Structure prediction was undertaken using the I-TASSER tool for all 11 Bacteriocins, and three were considered.

Table 5: Final bacteriocin and their binding partners

1.	gi|428148488	subsp. pneumoniae ST512-K30BO	1. *Shigella sonnei* **Protein:** RNA polymerase sigma factor FecI 2. *Shigella dysenteriae* **Protein:** RNA polymerase sigma factor FecI 3. *Photorhabdus asymbiotica* **Protein:** probable RNA polymerase sigma factor FecI
2	gi|152970529	subsp. pneumoniae MGH 78578	1. *Citrobacter koseri* **Protein:** 1. Anhydro-N-acetylmuramic acid kinase. 2. Transcriptional regulator SlyA. 2. *Legionella pneumophila* Paris **Protein:** 1. Cytochrome C biogenesis protein 2. Cytochrome C-type biogenesis protein CcmF 3. Cytochrome c-type biogenesis protein CcmE 3. *Salmonella enterica* CVM19633 **Protein:** 1. pili assembly chaperone 2. fimbrial chaperone protein
3.	gi|238895706	subsp. pneumoniae NTUH-K2044	1. *Salmonella enterica* Gallinarium **Protein:** 1. Possible outer membrane protein 2. Probable secreted protein 3. Possible transcriptional regulator 4. ATP synthase protein I 5. Cytochrome c-type biogenesis protein 6. RNA polymerase-associated protein RapA 2. *Legionella pneumophila* Paris **Protein:** 1. cytochrome C biogenesis protein 2. Cytochrome C-type biogenesis protein CcmF 3. Cytochrome c-type biogenesis protein CcmE 4. Cytochrome c-type biogenesis protein CcmE 5. heme exporter protein CcmD 6. heme exporter protein CcmC

Table 6: Final bacteriocins' patterns detail

Bacteriocin	Bacteriocin class-1	Bacteriocin class-2	Bacteriocin class-3
gi|428148488	YES		
gi|152970529	YES		
gi|238895706	YES	YES	YES

Table 6 gives an idea of which pattern class the bacteriocin belongs to. So, we can identify the putative functional mechanism of that bacteriocin against its binding partners.

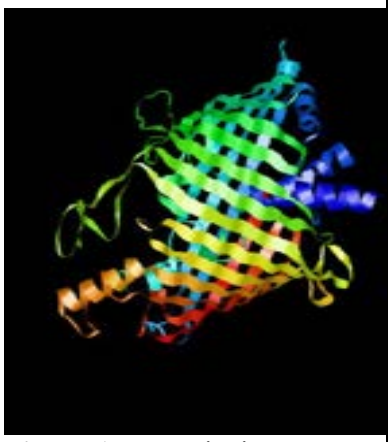


Figure 1: Bacteriocin structure of *Klebsiella pneumoniae* Sub spp. Pneumoniae NTUH K2044 (GI|28895706) predicted by I-TASSER

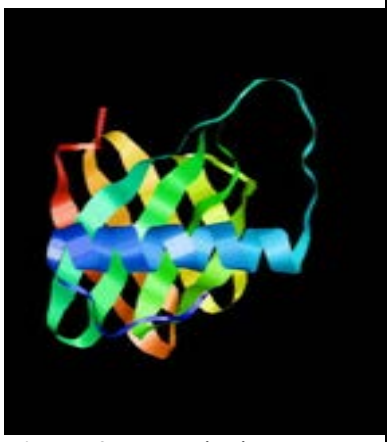


Figure 2: Bacteriocin structure of *Klebsiella pneumoniae* subsp. Pneumoniae MGH 78578 (gi|152970529) predicted by I-TASSER.

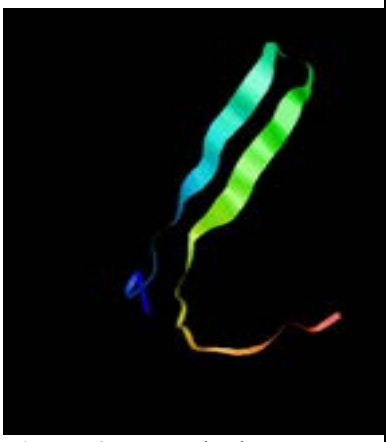


Figure 3: Bacteriocin structure of *Klebsiella pneumoniae* SUBSP. PNEUMONIAE ST512-K30BO (gi|428148488) predicted by I-TASSER.

These Structures were used for protein-protein Docking of Bacteriocins and their respective binding partners using ClusPro (Figures 1 to 3). ClusPro produced a cluster list for a single docking model. Docking models were screened, and the lowest-energy model was selected. Docking models finalised in the previous step were visualised for hydrogen bonds in the UCSF Chimera tool, which also provided the number of hydrogen bonds for each docking model (Figures 4 to 13).

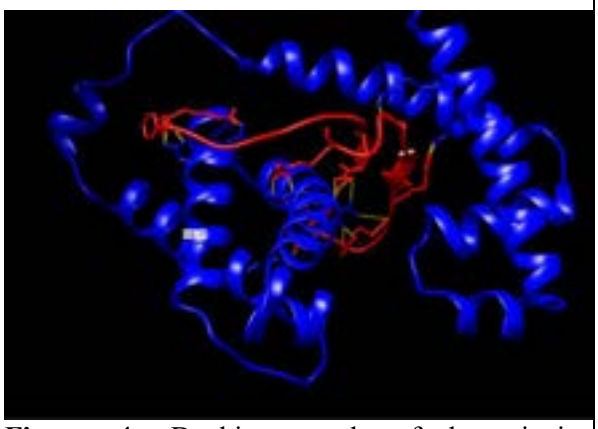


Figure 4: Docking result of bacteriocin (gi|428148488) with its binding partner RNA Polymerase protein of *Shigella dysenteriae*.

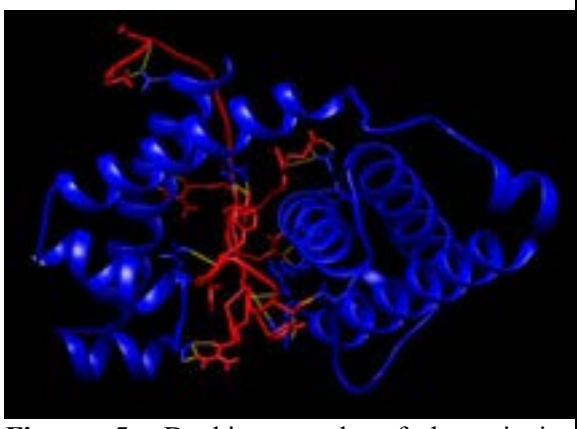


Figure 5: Docking result of bacteriocin (gi|428148488) with its binding partner RNA polymerase protein sigma factor fec1 of *Photorhabdus asymbiotica*.

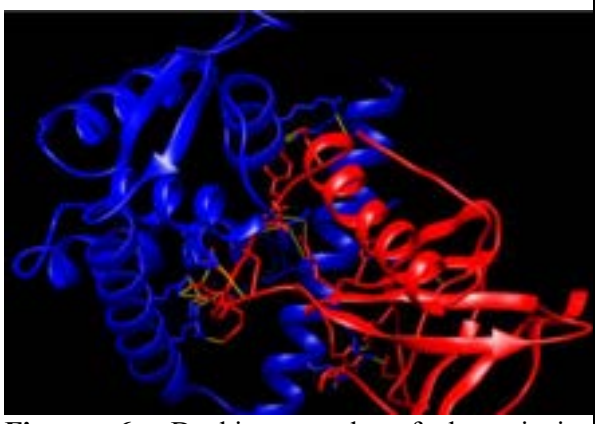


Figure 6: Docking result of bacteriocin (gi|152970529) with its binding partner transcriptional regulator SlyA of *Citrobacter koseri*.

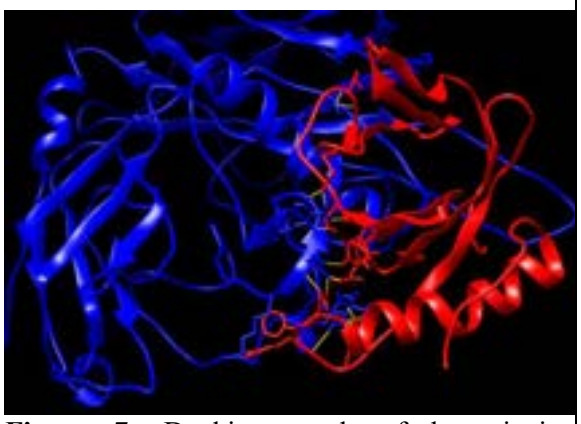


Figure 7: Docking result of bacteriocin (gi|152970529) with its binding partner cytochrome c type biogenesis protein CcmF of *Legionella pneumophila* Paris.

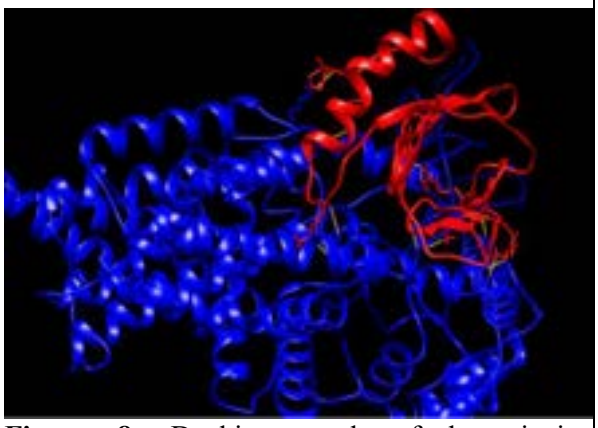


Figure 8: Docking result of bacteriocin (gi|152970529) with its binding partner Pili assembly chaperone of S*almonella enterica* cvm19633.

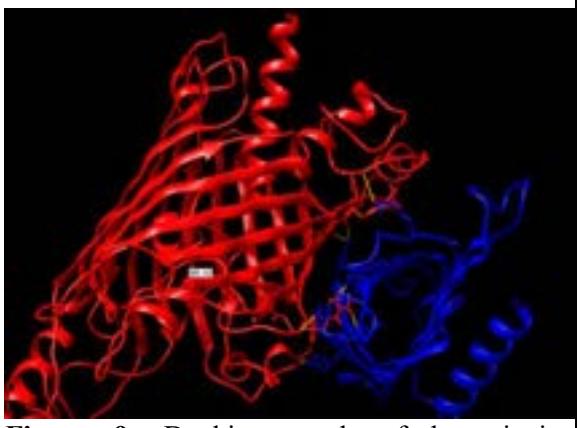


Figure 9: Docking result of bacteriocin (gi|238895706) with its binding partner, the possible outer membrane protein of S*almonella enterica* Gallinarium.

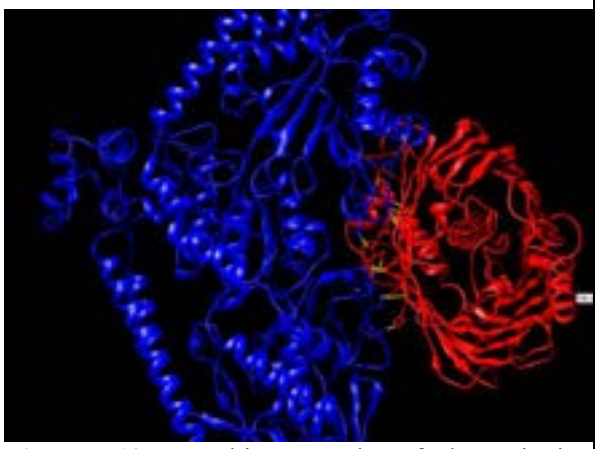


Figure 10: Docking result of bacteriocin (gi|238895706) with its binding partner RNA Polymerase protein of S*almonella enterica* Gallinarium.

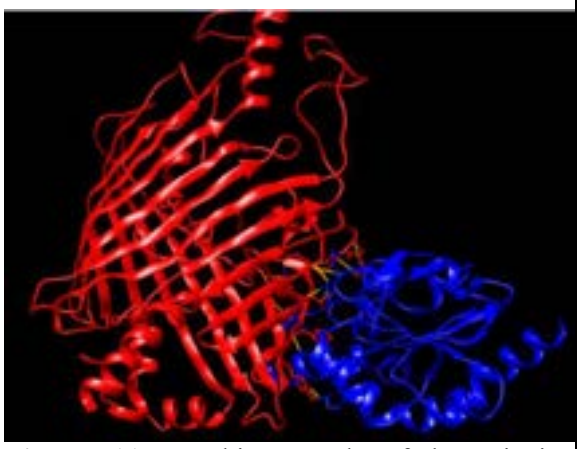


Figure 11: Docking result of bacteriocin (gi|238895706) with its binding partner, the cytochrome c biogenesis protein of *Legionella Pneumophila* Paris.

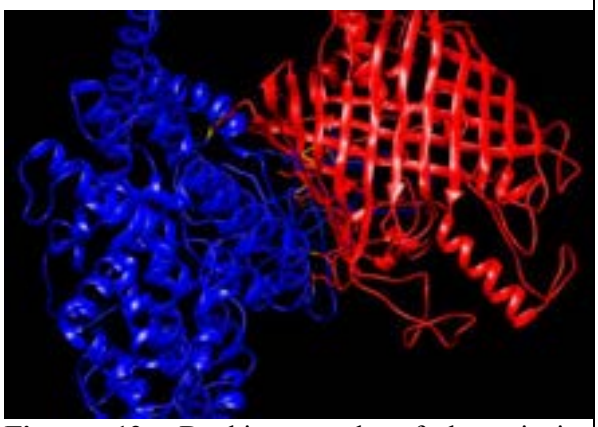


Figure 12: Docking result of bacteriocin (gi|238895706) with its binding partner cytochrome c biogenesis protein CcmF of *Legionella Pneumophila* Paris

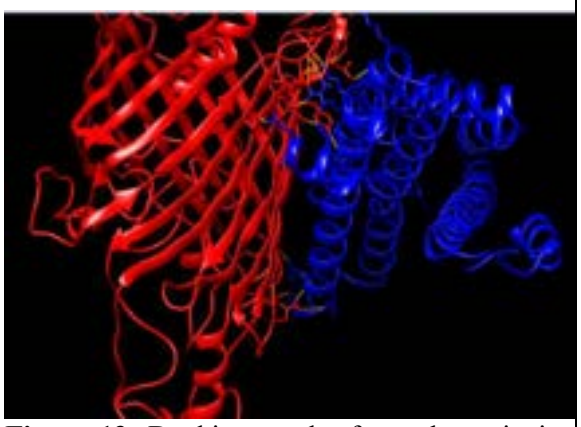


Figure 13: Docking result of bacteriocin (gi|238895706) with its binding partner heme exporter protein CcmC of *Legionella Pneumophila* Paris

Finally, after performing all this analysis, the conclusion is that three potential Klebsiella proteins out of 14754 (TABLE-5) indicate the final three Bacteriocins and their significant binding partners.

Table 7: Final bacteriocins and their binding partners with docking energy

Sr. no	Bacteriocin name& accession no	Binding partner's name		Lowest docking energy	Number of hydrogen bonds
1.	gi|428148488 subsp. Pneumoniae ST512-K30BO	1. *Shigella dysenteriae* RNA polymerase sigma factor FecI		-1010.8	19
		2. *Photorhabdus asymbiotica* probable RNA polymerase sigma factor FecI		-1057.2	22
2.	gi|152970529 *Klebsiella pneumoniae* subsp. Pneumoniae MGH 78578	1. *Citrobacter koseri* Transcriptional regulator SlyA		-1012	23
		2. *Legionella pneumophila* Paris Cytochrome C-type biogenesis protein CcmF		-898.6	18
		3. *Salmonella enterica* CVM19633 pili assembly chaperone		-1424.9	14
3.	gi|238895706 subsp. Pneumoniae NTUH-K2044	1. *Salmonella enterica* Gallinarium	Possible outer membrane protein	-1932.3	9
			RNA polymerase-associated protein RapA	-1444.7	18
		2. *Legionella pneumophila* Paris	cytochrome C biogenesis protein	-1606.4	13
			Cytochrome C-type biogenesis protein CcmF	-2085.5	11
			Heme exporter protein CcmC	-1882.3	16

4. DISCUSSION

The objective behind this research is to predict bacteriocins. Bacteriocin prediction requires analysing all properties needed to determine potential bacteriocins from the query sequences (2, 23, and 237

Klebsiella proteins), which were collected in the first step from NCBI. The Cd-HIT result was clustered protein sequences and reduced the number of proteins to 14754. This step is important because Cd-HIT removes redundant sequences and clusters proteins. Hence, the resultant protein sequence set is non-redundant. In the next step, filtering for outer membranes and extracellular proteins generated 424 outer membrane proteins and 215 extracellular proteins. This separation is necessary because the main purpose of this research is to predict bacteriocins to inhibit other proteins that cause infection in humans. A bacteriocin is an inhibitor of a pathogen only when it is outside the cell.

Some outer membrane proteins are transmembrane, and the outside domain of those transmembrane proteins is used for bacteriocin production. Transmembrane helix prediction was performed using TMHMM, and the outside domain of each membrane protein was identified. After that, only extracellular and outer membrane protein sequences totalling 639 were used for further analysis. If a protein can become a bacteriocin, it might have potential antimicrobial activity or a secretory nature. Proteins with any of these properties were used in bacteriocin prediction for this purpose. The Secretory proteins, antimicrobial peptides, and Bacteriocins prediction step was performed. That step generated 38 potential bacteriocin sequences. Of these, 38 still had only a 50% chance of being bacteriocins.

To obtain a more accurate Bacteriocin pattern, construction was performed. This pattern was constructed based on the conserved part in the input protein set. It helps remove sequences that were not found to have a pattern by ScanProsite. For more precise results, the ScanProsite output was matched with the TMHMM output for the outer membrane. The outside domain of a transmembrane protein may be significant for becoming a bacteriocin. That outside domain must resemble one of the mentioned bacteriocin patterns. Thus, pattern searching was important and yielded 36 proteins with positive results. The bacteriocin pattern is important because each class of bacteriocin has a significant functional mechanism, and research is ongoing in this area. Still, literature studies provide basic ideas about their functions.

Bacteriocin class 1 has the property of binding to the cell wall of target bacteria and disrupting cell wall production, as well as several physiological functions. Thus, it can inhibit bacteria from surviving in the human body.

Class IIa bacteriocins are all cationic and display anti-listeria activity, killing target cells by permeabilising the cell membrane. Class IIb bacteriocins have two distinct peptides for antimicrobial activity. They are Enterocin and Lactococcin. These peptides act as pore-forming toxins that create cell membrane channels through a barrel-stave mechanism, thus allowing ionic molecules into the cell. Class IIc circular bacteriocins have a range of effects on the membrane and pheromone actions of the target cell. This suggests that class II bacteriocins may be able to enter the membrane and inhibit the target cell.

Class IIIa bacteria kill bacteria by degrading their cell walls, thereby causing cell lysis. Class IIIb bacteria kill target bacteria by disrupting membrane potential, leading to ATP efflux. Researchers also concluded that class III bacteriocins may have the capacity to enter bacteria.

The objective of bacteriocin prediction was to protect humans against other pathogens that can cause infection. A BLAST search against humans was performed to identify proteins that are non-homologous to humans. If a bacteriocin is homologous with a human protein, it may have the potential to inhibit a human protein that serves a specific purpose and that creates certain problems in humans.

From these 38 sequences, some may have both bacteriocin and virulence factor properties, because some Klebsiella proteins may also be able to cause infections. Therefore, virulence factor prediction was performed using three different tools, and virulent proteins were excluded.

Bacteriocin prediction must have a significant purpose in the human body against infection. For that, it is necessary to identify interacting protein partners that play an infectious role in humans. So bacteriocins can inhibit them. For that, protein sequences were screened using the STRING database, and interacting partners were found based on function and their pathogenicity. Only 3 bacteriocins and their binding partners were considered for further analysis (Table 5). The pattern of each bacteriocin was identified (Table 6).

After testing all three Klebsiella proteins, they were finalised as bacteriocins (Table 5**).** The first *Klebsiella* protein (gi|4281484488) of *Klebsiella pneumoniae* subsp. Pneumoniae ST512-K30BO was interacting with proteins of *Shigella sonnei*, *Shigella dysenteriae*, *Photorhabdus asymbiotica* (Table 5). These three are infectious to humans. First, *Shigella sonnei* and *Shigella dysenteriae* belong to the same genus, *Shigella,* and they are antibiotic-resistant (Jain et al., 2005)**.** They are responsible for 90% of shigellosis, or Marlow syndrome, in humans. *Shigella* causes 90 million cases of severe infection,

resulting in 100000 deaths. So, bacteriocin may inhibit this disease because Klebsiella proteins interact with the sigma factor of RNA polymerase in Shigella sonnei and Shigella dysenteriae, and can inhibit their transcription factor activity. If RNA synthesis of the organism stops, it can't synthesise proteins, and its growth will be inhibited.

Thus, bacteriocin may cure shigellosis. RNA polymerase of *Photorhabdus asymbiotica* also interacts with the same bacteriocin. This bacterium causes locally invasive soft tissue disease in the United States and Australia, and it is difficult to stop. The bacteriocin acts as an inhibitor against this disease. This bacteriocin belongs to the type one bacteriocin class (Table 6). It binds to the cell membrane and causes cell lysis. However, research in this area is ongoing, so it is still unclear which bacteriocin can enter the cell. If Type-1 can enter the cell, then this bacteriocin is capable of inhibiting the replication of all three bacterial replication proteins.

The second bacteriocin (gi|152970529) interacts with *Citrobacter koseri*, *Legionella pneumophila* Paris, and *Salmonella enterica* CVM19633, as shown in Table 3. All three cause various infections and diseases in humans. First, *Citrobacter koseri* causes neonatal brain abscesses, having a high rate of mortality and complications; therefore, neonates are usually left with several residual permanent damages, so it is dangerous for the neonate.

This second bacteriocin interacts with the Anhydro-N-acetylmuramic acid kinase of Citrobacter koseri, which functions in carbohydrate metabolism, kinase transferase activity, ATP binding, peptidoglycan turnover, etc. Based on function, this protein may be a membrane protein, and bacteriocin belongs to type 1 (Table 6); therefore, it has the capacity to interact with the membrane and cause cell lysis. Thus, this bacteriocin interacts with the Citrobacter koseri protein and inhibits its function, such as metabolism and ATP binding, so it cannot survive or function in our body. This same bacteriocin also interacts with *Legionella pneumophila,* the causative agent of legionellosis. They are a common cause of hospital- and community-acquired pneumonia. The bacteriocin interacts with three distinct types of cytochrome c proteins, which function in heme transport, disulfide oxidoreductase activity, and related processes **(**Table 5**).** They are also present on the membrane. Thus, this bacteriocin can inhibit their function and may prevent diseases caused by *Legionella pneumophila*.

This second bacteriocin also interacts with the pili and fimbrial chaperone protein of *Salmonella enterica*. Pilus assembly via the chaperone-usher pathway can mediate microbial attachment and is useful in the establishment of infection. The fimbrial protein is involved in pilus biogenesis. Both proteins play a significant role in causing infection and are key proteins that establish infection and are membrane-bound. This bacteriocin belongs to the type one class and may have the potential to inhibit these two proteins, Kl.

The Third bacteriocin (gi|238895706) (Table 5) interacts with *Legionella pneumophila*'s cytochrome c proteins, as described above, and with *Salmonella enterica serovar Gallinarum's* possible outer membrane proteins, ATP-binding proteins, transcriptional regulator proteins, cytochrome c proteins, etc. **(**Table 5). The bacteriocin pattern belongs to all three classes (Table 6). So, based on the hypothesis that it can inhibit proteins, whether inside or outside the cell, and inhibit the function of interacting partners' proteins.

Docking was performed for the three bacteriocins (Table 5) using previously predicted structures (Figures 4 to 13). ClusPro gave several docked models for a single interaction; the lowest-energy model from that cluster was considered. Docking with the lowest energy is preferred.

UCSF Chimera was used to visualise the docking model and identify hydrogen bonds. This suggests that interactions between bacteriocins and binding partners with more hydrogen bonds reflect stronger binding. A model with sufficient hydrogen bonds and the lowest energy is very promising. Bacteriocins have great potential to prevent or inhibit diseases caused by their respective binding partners.

After analysing all the facts, three Klebsiella bacteriocins were identified as antimicrobial agents based on their interacting protein partners (Table 7). These *Klebsiella* bacteriocins help prevent diseases caused by their binding partners.

5. CONCLUSION

Prediction of multifunctional antimicrobial peptides using a variety of bioinformatics tools and databases is a major focus after the availability of numerous completed genomes. Several tools and databases have shown accurate results in predicting these multifunctional peptides. As a result, docking generated three significant bacteriocins.

6. REFERENCES

Andersson, D. I., & Hughes, D. (2010). Antibiotic resistance and its cost: is it possible to reverse resistance? *Nature Reviews Microbiology*, *8*(4), 260-271.

Bagley, S. T., Seidler, R. J., Talbot Jr, H. W., & Morrow, J. (1978). Isolation of Klebsielleae from within living wood. *Applied and environmental microbiology*, *36*(1), 178-185.

Chan, Y. R., Liu, J. S., Pociask, D. A., Zheng, M., Mietzner, T. A., Berger, T., Mak, T. W., Clifton, M. C., Strong, R. K., & Ray, P. (2009). Lipocalin 2 is required for pulmonary host defense against Klebsiella infection. *The Journal of Immunology*, *182*(8), 4947-4956.

Chen, Y., Zhang, Y., Yin, Y., Gao, G., Li, S., Jiang, Y., Gu, X., & Luo, J. (2005). SPD—a web-based secreted protein database. *Nucleic acids research*, *33*(suppl_1), D169-D173.

Dufour, A., Hindré, T., Haras, D., & Le Pennec, J.-P. (2007). The biology of lantibiotics from the lacticin 481 group is coming of age. *FEMS microbiology reviews*, *31*(2), 134-167.

Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., & Liang, J. (2006). CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic acids research*, *34*(suppl_2), W116-W118.

Heng, N. C., Wescombe, P. A., Burton, J. P., Jack, R. W., & Tagg, J. R. (2007). The diversity of bacteriocins in Gram-positive bacteria. In *Bacteriocins: ecology and evolution* (pp. 45-92). Springer.

Jain, S. K., Gupta, A., Glanz, B., Dick, J., & Siberry, G. K. (2005). Antimicrobial-resistant Shigella sonnei: limited antimicrobial treatment options for children and challenges of interpreting in vitro azithromycin susceptibility. *The Pediatric infectious disease journal*, *24*(6), 494-497.

Krogh, A., Larsson, B., Von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology*, *305*(3), 567-580.

Nes, I. F., Diep, D. B., & Holo, H. (2007). Bacteriocin diversity in Streptococcus and Enterococcus. *Journal of bacteriology*, *189*(4), 1189-1198.

Nordmann, P., Cuzon, G., & Naas, T. (2009). The real threat of Klebsiella pneumoniae carbapenemase-producing bacteria. *The Lancet infectious diseases*, *9*(4), 228-236.

Park, J., Park, B., Jung, K., Jang, S., Yu, K., Choi, J., Kong, S., Park, J., Kim, S., & Kim, H. (2007). CFGP: a web-based, comparative fungal genomics platform. *Nucleic acids research*, *36*(suppl_1), D562-D571.

A BIOINFORMATICS APPROACH TO IDENTIFY NOVEL BACTERIOCIN FROM *STAPHYLOCOCCUS*

Ashara Mausami Pranavbhai[1], Nilam J. Patel[2], John J. Georrge[3*]

[1] *Department of Botany, Bioinformatics & CCIM, Gujarat University, Gujarat-380009, India*

[2] *G H Patel Post Graduate Department of Computer Science and Technology (GDCST), Sardar Patel University, V V Nagar, Gujarat*

[3] *Department of Bioinformatics, University of North Bengal, District-Darjeeling, West Bengal-734013, India*

**Corresponding author: johnjgeorrge@gmail.com*

ABSTRACT: Bacteriocins are produced by bacteria and inhibit the growth of closely related strains. Bacteriocins are classified in several ways, including by mode of production, common resistance mechanisms, and mechanisms of action. Staphylococcus is a Gram-positive bacterium. It is harmless and normally occurs on the skin and mucous membranes of humans and other organisms. Staphylococcus can cause a variety of diseases in humans and other animals by producing toxins. Various antibiotics have different side effects. Antimicrobial peptides are a vital component of the innate immune response. They are found among all classes of plant and animal life. Gram-positive bacteria contain different types of secretion systems for the transport of proteins into or across the cytoplasmic membrane. Studies of the subcellular localisation of components of secretion systems and their substrates have shown that these components and their substrates can be distributed across multiple cellular locations. Pattern construction: discover patterns that are conserved in a set of protein sequences. Prosite is a database of protein families and domains. Important for the function of a protein and for the maintenance of its three-dimensional structure. Virulence is a quantitative measure of pathogenicity, or the likelihood of causing disease. The development of all organisms is mainly determined by molecular interactions, e.g., DNA-protein, protein-protein, and protein-small-molecule interactions. Protein-protein interactions appear to be critical in regulating key biological processes, including cell cycle control, cell type differentiation, protein folding, signalling, transcription, translation, post-translational modifications, and transport. The string database identifies a protein's binding partners. 3D structures were predicted for bacteriocin and binding partners. ClusPro is a protein docking program that docks two proteins. Chimera is a highly extensible program for interactive visualisation and analysis of molecular structures and related data, including density maps, supramolecular assemblies, sequence alignments, docking results, trajectories, and conformational ensembles. The docking results from chimera studies identified several proteins as novel Staphylococcus bacteriocins.

KEYWORDS: Staphylococcus bacteriocins, bioinformatics screening, antimicrobial peptides, protein-protein interaction analysis, protein docking, and 3D structure prediction

1. INTRODUCTION

Staphylococcus is a Gram-positive, non-motile, non-spore-forming bacterium that typically appears in grape-like clusters. It is commonly part of normal human skin and mucosal flora, although certain species can act as opportunistic pathogens. Among over 40 identified species, *Staphylococcus aureus* and *Staphylococcus* epidermidis are the most extensively studied due to their clinical relevance. The cell wall of *Staphylococcus* is structurally robust, primarily composed of peptidoglycan and teichoic acids. These components not only maintain cell integrity, but they also play a role in host interaction, including adhesion to tissues. Although a normal commensal, *S. aureus* can cause a wide range of infections, from minor skin conditions to severe diseases such as pneumonia and sepsis (Al-Attwani, 2014). The pathogenicity of *S. aureus* is driven by a wide array of virulence factors, including toxins, enzymes, and surface adhesion proteins. These include cytotoxins (e.g., alpha and beta toxins), superantigens such as enterotoxins and toxic shock syndrome toxin (Anné et al., 2017), and adhesion

molecules known as MSCRAMMs that facilitate binding to host tissues, such as collagen and fibronectin (Milani et al., 2023). Additionally, enzymes such as proteases, lipases, and coagulase support tissue invasion and immune evasion (Patti et al., 1994). The ability to produce multiple virulence factors simultaneously makes *S. aureus* a highly adaptable and dangerous pathogen, capable of causing both localised and systemic infections (Foster, 2019; Milani et al., 2023).

One of the major concerns associated with *S. aureus* is its rapid development of antibiotic resistance. Strains such as Methicillin-Resistant *S. aureus* (MRSA) have significantly reduced treatment options (Milani et al., 2023). Resistance mechanisms include enzymatic degradation of antibiotics (e.g., β-lactamases), Modification of drug targets (e.g., PBP2a encoded by the mecA gene), and reduced drug-binding affinity. Over time, resistance has emerged to multiple antibiotics, including penicillin, methicillin, and even advanced drugs such as vancomycin (Ambade et al., 2023; Lee et al., 2018). This has led to increased treatment complexity, higher healthcare costs, and significant mortality rates. Although several antibiotics, such as vancomycin, linezolid, and daptomycin, are used to treat resistant infections, their effectiveness is limited by factors including rising resistance, Toxic side effects (e.g., nephrotoxicity, ototoxicity), and High treatment costs. These challenges highlight the urgent need for alternative antimicrobial strategies (Roach & Donovan, 2015).

In Gram-positive bacteria such as Staphylococcus, proteins are transported via specialised secretion systems. Many of these secreted proteins contribute directly to virulence, host interaction, and immune evasion. Identifying these proteins is essential for understanding bacterial pathogenicity and discovering new therapeutic targets. Cell-penetrating peptides (CPPs) and antimicrobial peptides (AMPs) are particularly important in this context. CPPs facilitate the transport of molecules across cell membranes, while AMPs are naturally occurring defence molecules with broad-spectrum antimicrobial activity (Jiang et al., 2026). These peptides play a key role in innate immunity and offer promising alternatives to traditional antibiotics (Goormaghtigh & Van Bambeke, 2024).

Bacteriocins are ribosomally synthesised antimicrobial peptides produced by bacteria to inhibit competing microorganisms. Unlike conventional antibiotics, they are often narrow-spectrum and target closely related species (Bahrami et al., 2024). Based on structure and function, bacteriocins are broadly classified into: Class-I: Lantibiotics (post-translationally modified peptides), Class-II: Small, heat-stable peptides, Class-III: Larger, heat-labile proteins. Due to their specificity and reduced toxicity, bacteriocins are gaining attention as potential therapeutic agents and natural preservatives (Mokoena, 2017; Wang, 2024). This study aims to identify novel bacteriocin candidates from Staphylococcus using an integrated bioinformatics pipeline. By combining sequence analysis, structural prediction, and protein interaction studies, the research aims to identify potential antimicrobial peptides as alternatives to conventional antibiotics (Aljohani et al., 2023).

2. MATERIALS AND METHODS

A structured bioinformatics pipeline was employed to identify potential bacteriocin candidates from *Staphylococcus*. Protein sequences were retrieved from NCBI (Jenuth, 1999) and UniProt (Consortium, 2019), and redundancy was reduced to obtain representative datasets. Figure 1 indicates the overall workflow for the identification of novel bacteriocins from Staphylococcus species, including sequence collection, localisation prediction, secretome and antimicrobial screening, motif analysis, virulence filtering, interaction analysis, structure prediction, docking, and visualisation. Supporting datasets for antimicrobial peptides, bacteriocins, and secretory proteins were compiled from curated databases such as APD (Wang & Wang, 2004), CAMP (Waghu et al., 2016), DAMPD (Seshadri Sundararajan et al., 2012), YADAMP (Piotto et al., 2012), and BACTIBASE (Hammami et al., 2010), along with secretome-specific resources, including SPD (Chen et al., 2005), FSD (Choi et al., 2010), and FunSecKB (Lum & Min, 2011). Screening began with subcellular localisation analysis using PSORTb (Peabody et al., 2016), retaining only extracellular and membrane-associated proteins, as these are more likely to exhibit antimicrobial activity. TMHMM was used to confirm membrane topology and exclude irrelevant proteins (Cuthbertson et al., 2005). Secretory potential was further evaluated using BLAST against secretome datasets and validated using SecretomeP and SecretP 2.0 (Yu et al., 2010), ensuring the selection of proteins involved in secretion pathways (Cecconi et al., 2021).

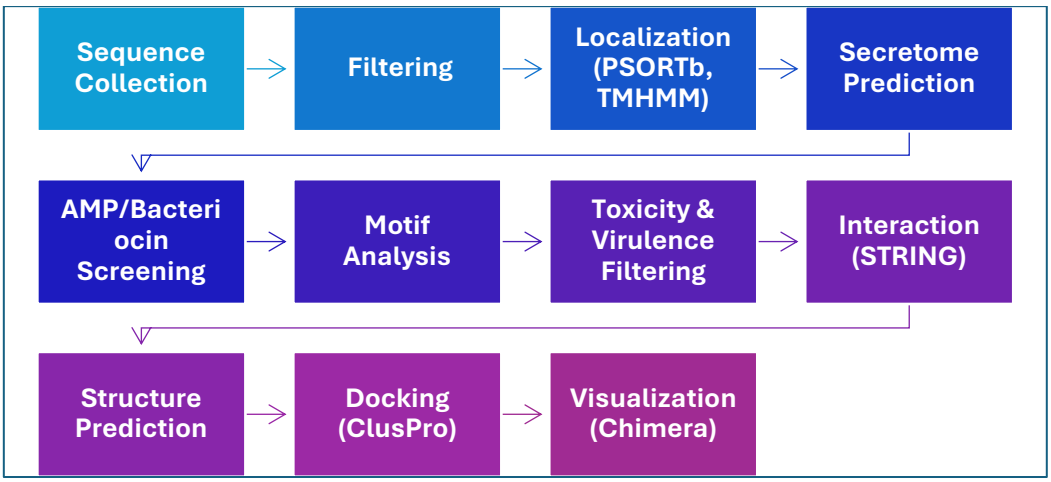


Figure 1: Overall computational workflow for identification and characterisation of bacteriocins from Staphylococcus spp

To identify antimicrobial and bacteriocin-like properties, sequence similarity searches were performed against AMP and bacteriocin datasets. Predictions were refined using CAMP, AntiBP3 (Bajiya et al., 2023), and BAGEL2 (de Jong et al., 2010). Conserved bacteriocin-specific motifs were identified through PRATT (Revett et al., 2007) and validated with PROSITE (Sigrist et al., 2010) patterns, enabling classification into major bacteriocin classes. To improve specificity, candidate proteins were screened against the human proteome using BLAST, and those showing significant similarity were excluded. Virulence factor analysis was then performed using VFDB (Zhou et al., 2025), VICMpred (Saha & Raghava, 2006), and SVM-based prediction to remove pathogenic proteins and retain safer candidates. Functional relevance was assessed using protein–protein interaction analysis in STRING (Szklarczyk et al., 2025), which identified key interacting partners. Selected candidates were then subjected to structural modelling using I-TASSER (Xu & Zhang, 2011), followed by refinement with ModRefiner (Yang & Zhang, 2015). Finally, protein–protein docking was performed using ClusPro (Kozakov et al., 2017) and interaction stability was analysed in Chimera software using binding energies and hydrogen-bonding patterns.

3. RESULTS AND DISCUSSION

A total of 29,151 *Staphylococcus* protein sequences were retrieved and clustered to remove redundancy. Subcellular localisation screening identified 835 extracellular and 424 cell wall proteins. Further refinement using transmembrane topology analysis reduced the cell wall dataset to 101 proteins with outward-facing domains. Secretome prediction identified 698 extracellular and 93 cell-wall proteins, indicating a substantial pool of proteins potentially involved in external interactions. Screening for antimicrobial and bacteriocin properties using multiple prediction tools identified 190 extracellular and 69 cell-wall candidates. Motif-based filtering based on conserved bacteriocin patterns further reduced this dataset to 133 proteins. Subsequent similarity analysis against the human proteome ensured removal of homologous sequences, improving specificity and reducing potential toxicity concerns. Virulence filtering across three independent tools identified 112 non-virulent proteins, which were considered suitable for downstream analysis. Protein-protein interaction analysis using STRING revealed that 12 proteins exhibited cross-genus interactions, suggesting broader antimicrobial potential. Among these, two candidates demonstrated the most stable and biologically relevant interactions, as assessed by docking energy and structural analysis. Table 1 presents the two highest-ranking bacteriocins identified by docking analysis, including their protein IDs, sequence lengths, target organisms, interacting proteins, functional roles, binding energies, and numbers of hydrogen bonds. Lower energy and stronger hydrogen bonding indicate stronger, more stable interactions. The first candidate (gi|242371730) showed the strongest interaction, with the lowest docking energy (-1219.8) and the highest number of hydrogen bonds (30), indicating a highly stable complex. Its interaction with a cell-surface protein precursor suggests a mechanism that disrupts membrane-associated processes, consistent with known bacteriocin activity. The high binding stability further supports its potential as a strong antimicrobial

agent. The second candidate (gi|418444278) also demonstrated significant interaction stability, with a docking energy of -1153.3 and 13 hydrogen bonds. Its interaction with a cAMP-binding regulatory protein suggests interference with cellular signalling pathways, potentially contributing to antimicrobial effects by disrupting essential regulatory mechanisms. Figure 2 shows the protein–protein docking complex of clumping factor A with the cell surface protein precursor. Figure 3 shows the docking complex between the LPXTG-domain-containing protein and the cAMP-binding protein, highlighting a strong binding interaction supported by low binding energy and a high number of hydrogen bonds, indicating structural stability.

Protein ID	Length	Target Organism	Binding Protein	Function	Energy	H-bonds
gi|242371730	547	*Lactobacillus plantarum*	Cell surface protein precursor	Transport-related	-1219.8	30
gi|418444278	237	*Hahella chejuensis*	cAMP-binding protein	Regulatory protein	-1153.3	13

Table 1: Top predicted bacteriocin candidates and their interacting protein partners

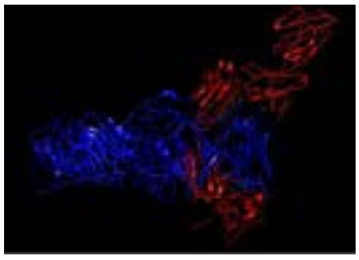


Figure 2: Docking interaction of clumping factor A with cell surface protein precursor

Figure 3: Docking interaction of LPXTG-domain-containing protein with cAMP-binding protein

Structural modelling and refinement showed that both candidates have stable shapes that are conducive to interaction with other molecules. Docking analysis showed that they bind effectively to each other. Visualisation further emphasised robust interaction interfaces, proving their functional significance. The integrated analysis narrowed a large protein dataset to a few high-confidence candidates. Finding these two proteins as possible bacteriocins shows how effectively combining sequence screening, structural modelling, and interaction analysis works. Their strong binding affinity, non-virulent nature, and ability to interact with other species suggest they could be good replacements for traditional antibiotics, especially against strains resistant to them.

4. CONCLUSION

A stepwise in silico screening strategy reduces an immense dataset of Staphylococcus proteins to a narrow selection of high-confidence bacteriocin candidates via localisation filtering, secretome prediction, antimicrobial screening, motif analysis, and virulence exclusion. The top two candidates among these exhibited the most stable protein–protein interactions, with very good binding energies and strong hydrogen-bonding patterns. Identifying non-virulent, extracellularly accessible proteins with stable docking profiles shows that they could be effective antimicrobial agents. These outcomes affirm that combining sequence-level filtering with structural and interaction-based validation significantly enhances the accuracy of bacteriocin discovery. The selected candidates provide a targeted basis for subsequent experimental validation and prospective therapeutic advancement against antibiotic-resistant pathogens.

5. REFERENCES

Al-Attwani, J. H. (2014). *The effect of probiotics on bacterial human skin pathogens,* University of Plymouth.

Aljohani, A. B., Al-Hejin, A. M., Shori, A. B. J. F. S., & Technology. (2023). Bacteriocins as promising antimicrobial peptides, definition, classification, and their potential applications in cheeses. *Food Science and Technology*, *43*, e118021. https://doi.org/10.1590/fst.118021

Ambade, S. S., Gupta, V. K., Bhole, R. P., Khedekar, P. B., & Chikhale, R. V. (2023). A Review on Five and Six-Membered Heterocyclic Compounds Targeting the Penicillin-Binding Protein 2 (PBP2A) of Methicillin-Resistant Staphylococcus aureus (MRSA). *Molecules*, *28*(20), 7008. https://doi.org/10.3390/molecules28207008

Anné, J., Economou, A., & Bernaerts, K. (2017). Protein Secretion in Gram-Positive Bacteria: From Multiple Pathways to Biotechnology. In F. Bagnoli & R. Rappuoli (Eds.), *Protein and Sugar Export and Assembly in Gram-positive Bacteria* (pp. 267-308). Springer International Publishing. https://doi.org/10.1007/82_2016_49

Bahrami, S., Andishmand, H., Pilevar, Z., Hashempour-baltork, F., Torbati, M., Dadgarnejad, M., Rastegar, H., Mohammadi, S. A., & Azadmard-Damirchi, S. (2024). Innovative perspectives on bacteriocins: advances in classification, synthesis, mode of action, and food industry applications. *Journal of Applied Microbiology*, *135*(11). https://doi.org/10.1093/jambio/lxae274

Bajiya, N., Choudhury, S., Dhall, A., & Raghava, G. P. S. (2023). AntiBP3: A hybrid method for predicting antibacterial peptides against gram-positive/negative/variable bacteria. *bioRxiv*, 2023.2007.2025.550443. https://doi.org/10.1101/2023.07.25.550443

Cecconi, D., Di Carlo, C., & Brandi, J. (2021). Protein Secretion Prediction Tools and Extracellular Vesicles Databases. In D. Cecconi (Ed.), *Proteomics Data Analysis* (pp. 213-227). Springer US. https://doi.org/10.1007/978-1-0716-1641-3_13

Chen, Y., Zhang, Y., Yin, Y., Gao, G., Li, S., Jiang, Y., Gu, X., & Luo, J. (2005). SPD—a web-based secreted protein database. *Nucleic Acids Research*, *33*(suppl_1), D169-D173. https://doi.org/10.1093/nar/gki093

Choi, J., Park, J., Kim, D., Jung, K., Kang, S., & Lee, Y.-H. (2010). Fungal Secretome Database: Integrated platform for annotation of fungal secretomes. *BMC Genomics*, *11*(1), 105. https://doi.org/10.1186/1471-2164-11-105

Consortium, T. U. (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research*, *47*(D1), D506-D515. https://doi.org/10.1093/nar/gky1049

Cuthbertson, J. M., Doyle, D. A., & Sansom, M. S. P. (2005). Transmembrane helix prediction: a comparative evaluation and analysis. *Protein Engineering, Design and Selection*, *18*(6), 295-308. https://doi.org/10.1093/protein/gzi032

de Jong, A., van Heel, A. J., Kok, J., & Kuipers, O. P. (2010). BAGEL2: mining for bacteriocins in genomic data. *Nucleic Acids Research*, *38*(suppl_2), W647-W651. https://doi.org/10.1093/nar/gkq365

Foster, T. J. (2019). The MSCRAMM Family of Cell-Wall-Anchored Surface Proteins of Gram-Positive Cocci. *Trends in Microbiology*, *27*(11), 927-941. https://doi.org/10.1016/j.tim.2019.06.007

Goormaghtigh, F., & Van Bambeke, F. (2024). Understanding Staphylococcus aureus internalisation and induction of antimicrobial tolerance. *Expert Review of Anti-infective Therapy*, *22*(1-3), 87-101. https://doi.org/10.1080/14787210.2024.2303018

Hammami, R., Zouhir, A., Le Lay, C., Ben Hamida, J., & Fliss, I. (2010). BACTIBASE second release: a database and tool platform for bacteriocin characterisation. *BMC Microbiology*, *10*(1), 22. https://doi.org/10.1186/1471-2180-10-22

Jenuth, J. P. (1999). The NCBI. In S. Misener & S. A. Krawetz (Eds.), *Bioinformatics Methods and Protocols* (pp. 301-312). Humana Press. https://doi.org/10.1385/1-59259-192-2:301

Jiang, J., Qi, Y., & Ma, S. (2026). Host–Pathogen Interactions and Peptide-Based Therapeutics in Intracellular Bacterial Infections. *ACS Infectious Diseases*, *12*(3), 897-912. https://doi.org/10.1021/acsinfecdis.5c00858

Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., Beglov, D., & Vajda, S. (2017). The ClusPro web server for protein–protein docking. *Nature Protocols*, *12*(2), 255-278. https://doi.org/10.1038/nprot.2016.169

Lee, A. S., de Lencastre, H., Garau, J., Kluytmans, J., Malhotra-Kumar, S., Peschel, A., & Harbarth, S. (2018). Methicillin-resistant Staphylococcus aureus. *Nature Reviews Disease Primers*, *4*(1), 18033. https://doi.org/10.1038/nrdp.2018.33

Lum, G., & Min, X. J. (2011). FunSecKB: the Fungal Secretome KnowledgeBase. *Database*, *2011*. https://doi.org/10.1093/database/bar001

Milani, M., Curia, R., Shevlyagina, N. V., & Tatti, F. (2023). Staphylococcus aureus. In *Bacterial Degradation of Organic and Inorganic Materials: Staphylococcus aureus Meets the Nanoworld* (pp. 3-20). Springer International Publishing. https://doi.org/10.1007/978-3-031-26949-3_1

Mokoena, M. P. (2017). Lactic Acid Bacteria and Their Bacteriocins: Classification, Biosynthesis and Applications against Uropathogens: A Mini-Review. *Molecules*, *22*(8), 1255. https://doi.org/10.3390/molecules22081255

Patti, J. M., Allen, B. L., McGavin, M. J., & Hook, M. (1994). MSCRAMM-mediated adherence of microorganisms to host tissues [Article]. *Annual Review of Microbiology*, *48*, 585+. https://doi.org/10.1146/annurev.mi.48.100194.003101

Peabody, M. A., Laird, M. R., Vlasschaert, C., Lo, R., & Brinkman, F. S. L. (2016). PSORTdb: expanding the bacteria and archaea protein subcellular localisation database to better reflect diversity in cell envelope structures. *Nucleic Acids Research*, *44*(D1), D663-D668. https://doi.org/10.1093/nar/gkv1271

Piotto, S. P., Sessa, L., Concilio, S., & Iannelli, P. (2012). YADAMP: yet another database of antimicrobial peptides. *International Journal of Antimicrobial Agents*, *39*(4), 346-351. https://doi.org/10.1016/j.ijantimicag.2011.12.003

Revett, K., Gorunescu, F., & Gorunescu, M. (2007). Mining an animal toxin database: characterising protein folds. Proceedings of the International Conference on Knowledge Engineering, Principles and Techniques, KEPT2007, Cluj-Napoca (Romania).

Roach, D. R., & Donovan, D. M. (2015). Antimicrobial bacteriophage-derived proteins and therapeutic applications. *Bacteriophage*, *5*(3), e1062590. https://doi.org/10.1080/21597081.2015.1062590

Saha, S., & Raghava, G. (2006). VICMpred: an SVM-based method for the prediction of functional proteins of Gram-negative bacteria using amino acid patterns and composition. *Genomics Proteomics Bioinformatics*, *4(1)*, 42-47. https://doi.org/10.1016/S1672-0229(06)60015-6

Seshadri Sundararajan, V., Gabere, M. N., Pretorius, A., Adam, S., Christoffels, A., Lehväslaiho, M., Archer, J. A. C., & Bajic, V. B. (2012). DAMPD: a manually curated antimicrobial peptide database. *Nucleic Acids Research*, *40*(D1), D1108-D1112. https://doi.org/10.1093/nar/gkr1063

Sigrist, C. J. A., Cerutti, L., de Castro, E., Langendijk-Genevaux, P. S., Bulliard, V., Bairoch, A., & Hulo, N. (2010). PROSITE, a protein domain database for functional characterisation and annotation. *Nucleic Acids Research*, *38*(suppl_1), D161-D166. https://doi.org/10.1093/nar/gkp885

Szklarczyk, D., Nastou, K., Koutrouli, M., Kirsch, R., Mehryary, F., Hachilif, R., Hu, D., Peluso, M. E., Huang, Q., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., & von Mering, C. (2025). The STRING database in 2025: protein networks with directionality of regulation. *Nucleic Acids Research*, *53*(D1), D730-D737. https://doi.org/10.1093/nar/gkae1113

Waghu, F. H., Barai, R. S., Gurung, P., & Idicula-Thomas, S. (2016). CAMPR3: a database on sequences, structures and signatures of antimicrobial peptides. *Nucleic Acids Research*, *44*(D1), D1094-D1097. https://doi.org/10.1093/nar/gkv1051

Wang, G. (2024). Bacteriocin Classification, Prediction and Applications. In S. Ray, P. Kumar, & M. Mandal (Eds.), *Antimicrobial Peptides from Lactic Acid Bacteria: Diversity, Biosynthesis and Applications* (pp. 97-114). Springer Nature Singapore. https://doi.org/10.1007/978-981-97-3413-9_4

Wang, Z., & Wang, G. (2004). APD: the Antimicrobial Peptide Database. *Nucleic Acids Research*, *32*(suppl_1), D590-D592. https://doi.org/10.1093/nar/gkh025

Xu, D., & Zhang, Y. (2011). Improving the Physical Realism and Structural Accuracy of Protein Models by a Two-Step Atomic-Level Energy Minimization. *Biophysical Journal*, *101*(10), 2525-2534. https://doi.org/10.1016/j.bpj.2011.10.024

Yang, J., & Zhang, Y. (2015). I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Research*, *43*(W1), W174-W181. https://doi.org/10.1093/nar/gkv342

Yu, L., Guo, Y., Zhang, Z., Li, Y., Li, M., Li, G., Xiong, W., & Zeng, Y. (2010). SecretP: A new method for predicting mammalian secreted proteins. *Peptides*, *31*(4), 574-578. https://doi.org/10.1016/j.peptides.2009.12.026

Zhou, S., Liu, B., Zheng, D., Chen, L., & Yang, J. (2025). VFDB 2025: an integrated resource for exploring anti-virulence compounds. *Nucleic Acids Research*, *53*(D1), D871-D877. https://doi.org/10.1093/nar/gkae968

INTELLECTUAL PROPERTY RIGHTS IN BIOINFORMATICS: A REVIEW OF LEGALITY AND PROTECTION STRATEGIES

Ashara Mausami Pranavbhai [1,2], Samir Modhavadiya[2], John J. Georrge[2, 3*]

[1] *Department of Botany, Bioinformatics & CCIM, Gujarat University, Gujarat-380009, India*
[2] *Department of Bioinformatics, Christ College, Rajkot, Gujarat, India*
[3] *Department of Bioinformatics, University of North Bengal, District-Darjeeling, West Bengal-734013, India*
**Corresponding author: johnjgeorrge@gmail.com*

ABSTRACT: Bioinformatics is the intersection of computer science and the life sciences, enabling the transformation of biological data into proprietary assets under intellectual property rights (IPR). It involves evaluating the global IPR landscape with a focus on protecting valuable biotechnological innovations alongside diverse computational architectures. Also investigate cross-jurisdictional differences in patentability standards, particularly the "technical effect" of bioinformatics algorithms and the eligibility of modified industrial organisms. The review further delineates the categories of protection for bioinformatics assets, including patents for unique diagnostic techniques, copyrights for original software code, and trade secrets for proprietary datasets. By combining these internationally recognised legal norms, this study creates an exhaustive basis for addressing the challenges of ownership, execution, and compatibility within the international scientific ecosystem.

KEYWORDS: Intellectual property rights, bioinformatics patents, algorithm patentability, copyright protection, trade secrets for datasets

1. INTRODUCTION

Intellectual Property Rights (IPR) are the legal protections granted to creations of the human intellect. These include inventions, software, artistic works, and scientific discoveries. IPR can be broadly classified into two categories: copyright and industrial property rights, which include trademarks, patents, and geographical indications. This allows creators to gain recognition and financial benefits from their work. To promote ongoing scientific and technological advancement, the system is designed to strike a balance between rewarding innovation and ensuring public access.

A concept of intellectual property protection is not new. In ancient Greece, innovators were granted exclusive rights to their inventions as early as 500 BC, when IPR first emerged. This concept developed into official legal systems over time. The Paris Convention (1883), which enabled the protection of inventions across national borders, and the Berne Convention (1886), which extended protection to literary and artistic works, are two significant turning points. The World Intellectual Property Organisation (WIPO), which currently supervises international IPR frameworks, was founded in 1967 as a result of these developments (Gopalan, 2009; Shilpa Gupta, 2018).

Among different forms of IPR, patents play a central role in scientific and technological fields. **A patent** grants an inventor exclusive rights for a limited period (generally 20 years), preventing others from making, using, or selling the invention without permission. To qualify for a patent, an invention must be novel, non-obvious, and industrially applicable. However, specific criteria may vary across countries. For instance, jurisdictions differ in how they define patentable subject matter, particularly in areas such as software and biotechnology. **Copyright**, on the other hand, protects original expressions such as software code, written content, and digital outputs. Unlike patents, copyright protection is automatic in most countries and lasts for the creator's lifetime plus several years after death. It safeguards both economic rights (financial benefits) and moral rights (recognition and integrity of the work).

With advances in the life sciences, a new interdisciplinary field, bioinformatics, has emerged. Bioinformatics applies computational tools to manage and analyse biological data such as DNA, RNA, and protein sequences. The growth of large-scale projects, such as genome mapping, has significantly increased the commercial and scientific value of this data. As a result, protecting these assets through IPR has become increasingly important (Chang & Zhu, 2010; Fernandez & Chow, 2005).

Bioinformatics assets can be broadly classified into three categories:

A. **Biological sequences** (DNA, RNA, proteins)
B. **Databases** storing biological information
C. **Software tools and algorithms** used for analysis

Nevertheless, not all of these elements are protected equally. For instance, modified biological materials or computational tools may be eligible for patents under specific circumstances, but raw biological sequences are typically not. While software is mainly protected by copyright, with patents applicable in certain jurisdictions, databases are frequently protected by copyright or special database rights. Table 1 illustrates the IP protection strategy's applicability in accordance with the component categories (Singh, 2014).

Table 1: *IPR Protection in Bioinformatics Components*

Component	Patent Protection	Copyright Protection	Trade Secret Protection
DNA, RNA, Protein Sequences	Limited or conditional; raw sequences are generally not patentable	Not applicable	Applicable if kept confidential
Databases	Conditional; patentable when linked to a technical process	Applicable for structure and arrangement	Applicable
Software and Tools	Conditional; depends on the technical effect and jurisdiction	Applicable to source code	Applicable

Patent inventions include Protein modelling, sequence modelling, computer-aided sequence visualisation, automation, databases, and data mining tools. Raw products, such as gene sequences, DNA, and RNA that copy genetic information and are transmitted from generation to generation, are not patentable. Microorganism peptides can be patented if they are isolated from nature or synthesised. Over twenty years ago, the US Supreme Court ruled that anything under the sun produced by man is patentable subject matter, although restrictions vary by nation. Table 2 illustrates gene-related patentability in India, the US, and Europe (Arora, 2022; Harrison, 2003; Maschio & Kowalski, 2001).

Table 2*: Legal Status of Patentability for Genetically Modified Biological Entities Across Major Jurisdictions*

Subject Matter	India	United States	Europe
Genetically modified unicellular organisms	Patentable	Patentable	Patentable
Genetically modified multicellular organisms	Patentable	Patentable	Patentable
Genetically modified animals (excluding humans)	Not patentable	Patentable	Patentable
Genetically modified plants	Not patentable	Patentable	Patentable
Genetically modified humans	Not patentable	Not patentable	Not patentable
Gene therapy	Not patentable	Patentable	Not patentable

2. MATERIALS AND METHODS

2.1 Patent Search and Prior Art Analysis

A structured patent search was conducted as the first step of this study to identify existing inventions and assess the bioinformatics novelty landscape. Patent searches are essential for locating prior art, which includes any publicly available information, such as patents, publications, or technical disclosures, that existed before the filing date of a patent application. The search process was designed not only to retrieve relevant patents but also to understand technological trends and identify innovation gaps. By analysing it, it becomes possible to evaluate how new bioinformatics tools or methods differ from existing solutions, particularly in terms of technical contribution and application relevance. This step is crucial in determining patentability, especially in fields like bioinformatics, where incremental improvements and algorithmic variations are common.

2.2 Patent Classification Framework

To ensure systematic analysis, retrieved patents were interpreted using established patent classification systems. These systems categorise patent documents based on their technical features, enabling efficient organisation, comparison, and retrieval of domain-specific innovations. Patent classification plays a key role in identifying technological clusters, tracking innovation trends, and supporting detailed prior art analysis. An overview of patent classification systems is shown in Table 3.

***Table 3**: Overview of Patent Classification Systems*

System	Description	Relevance in This Study
Cooperative Patent Classification (CPC)	Developed jointly by the European Patent Office and the United States Patent and Trademark Office	Used as the primary framework due to its detailed and updated classification structure
International Patent Classification (IPC)	Globally accepted hierarchical classification system used by over 100 patent offices	Supported cross-database comparison and standardisation
US Patent Classification (USPC)	The legacy classification system was previously used in the United States	Referenced for older patents where CPC classification was not available

The increasing shift toward CPC reflects the need for more granular classification, especially for interdisciplinary fields like bioinformatics (Chatterjee, 2017; Mago, Deshpande, & Hirwani, 2017; Park, 2012).

2.3 Data Sources and Search Strategy

Patent data were collected from multiple authoritative databases to ensure comprehensive coverage across jurisdictions. Databases Accessed: These databases were selected because they collectively cover international (PCT), regional, and national patent filings, providing a broad and balanced dataset (Figure 1). The Workflow of Patent Retrieval is shown in Figure 2.

WIPO: World Intellectual Property Organisation is one of the 15 specialised agencies of the United Nations (UN). WIPO (Birkbeck, 2016) was created in 1967 to encourage creative activity and promote the protection of intellectual property worldwide. The terms DNA Bioinformatics, RNA Bioinformatics, Protein/peptide Bioinformatics and Software Bioinformatics were used to retrieve bioinformatics patents from the PATENTSCOPE (Lerner, 1994) portal (https://patentscope.wipo.int/search/en/search.jsf) of the WIPO database, including all patents published in the PCT format. The search was in the FP (search criteria entered in this field are searched on the document's front page) field.

USPTO: The same terms were used in the USPTO database (Marco, Myers, Graham, D'Agostino, & Apple, 2015) (https://www.uspto.gov). Under PatFT (The Patent Full-Text Database), hyperlinks are provided to all United States Patents issued from 1790 to the present. Earlier patents (1790 – 1975) are only accessible by patent number, issue date, or current US classification) patent search option. To retrieve patents, DNA, RNA, proteins, and software, term 1 was set to "DNA, RNA, protein, and software," and term 2 to "Bioinformatics," with the AND operator. The field selection in each term was Abstract AND Title, Title AND Abstract, Title AND Title, Abstract AND Abstract, with the select years set from 1976 to the present (Wallace, 2016).

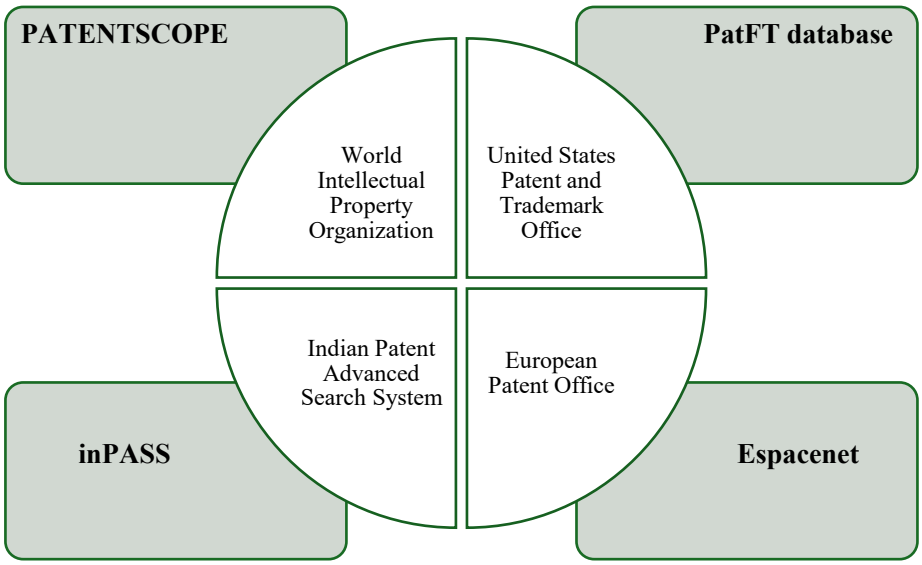


***Figure 1**: The search database details the jurisdictions*

EPO: The terms DNA Bioinformatics, RNA Bioinformatics, Protein/peptide Bioinformatics and Software Bioinformatics were used to retrieve bioinformatics patents from EPO's Espacenet patent search (https://worldwide.espacenet.com) with the AND operator in smart search (Barbieri & Andreoni, 2023).

Indian Patent Database (inPASS): The terms DNA Bioinformatics, RNA Bioinformatics, Protein/peptide Bioinformatics, and Software Bioinformatics were used to retrieve bioinformatics patents from inPASS (https://ipindiaservices.gov.in/publicsearch/). The publication type was set to published patents. Terms DNA and Bioinformatics, RNA and Bioinformatics, Protein and Bioinformatics and software and Bioinformatics used with the field selection of Abstract AND Title, Title AND Abstract, Title AND Title, Abstract AND Abstract (Mishra & Bonde, 2022).

Search Methodology: A keyword-based search strategy was adopted to retrieve patents related to bioinformatics. The following terms were used:

- DNA Bioinformatics
- RNA Bioinformatics
- Protein/Peptide Bioinformatics
- Software Bioinformatics

To improve relevance, keywords were combined using Boolean operators (primarily AND) and applied across multiple fields, including title, abstract, and front-page content. This multi-field approach helped capture both highly specific and broader patent records. Where applicable, searches were limited to full-text records from 1976 to the present, ensuring access to detailed patent descriptions and claims. For databases such as PATENTSCOPE, searches were refined with front-page filters to improve precision.

3. RESULT AND DISCUSSION

Bioinformatics patents retrieved from various databases were further classified by year from 2000 to 2019, across 9 countries: the Republic of China, North America (US and Canada), Switzerland, Australia, Germany, the United Kingdom, France, Japan, and India. The patent distribution by category across major patent offices is shown in Figure 3. The overall patent distribution is shown in Figure 4.

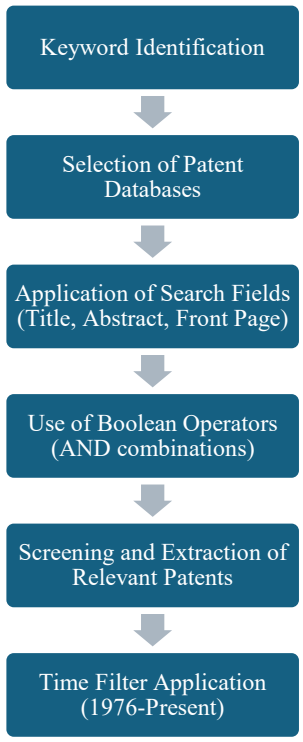


***Figure 2**: Workflow of Patent Retrieval*

USPTO: A total of 48 Bioinformatics patents were found from 2001 to 2019, using the specified parameters, and among them, 33 are patents classified under biological compounds, and 15 patents were found under the software category.

1. **DNA patents**: 22 were found in the database using the chosen parameters.
2. **RNA patents:** Only one patent was retrieved from the database using the chosen parameters.
3. **Protein Patents:** A total of 11 patents were found from the database using the chosen parameters.
4. **Software Patent:** A total of 15 patents were found from the database using the chosen parameters.

WIPO: A total of 627 Bioinformatics patents were found from 2001 to 2019, using the specified parameters, and among them, 522 are patents classified under biological compounds, and 105 patents were found under the software category.

1. **DNA Patents:** A total of 193 patents were found from the database using the chosen parameters.
2. **RNA patents:** 66 patents were found in the database using the chosen parameters.
3. **Protein Patent:** A total of 263 patents were found from the database using the chosen parameters.
4. **Software Patents:** A total of 105 patents were found from the database using the chosen parameters.

Indian Patent Database (inPASS): A total of 04 Bioinformatics patents were identified from 2001 to 2019 using the specified parameters; among them, 04 were classified under biological compounds, and 0 were classified under software.

1. **DNA Patents:** A total of 01 patent was found from the database using the chosen parameters.
2. **RNA Patents:** 0 Patents were found in the database using the chosen parameters.
3. **Protein Patent:** 3 patents were found in the database using the chosen parameters.
4. **Software Patent:** 0 patents were found in the database using the chosen parameters.

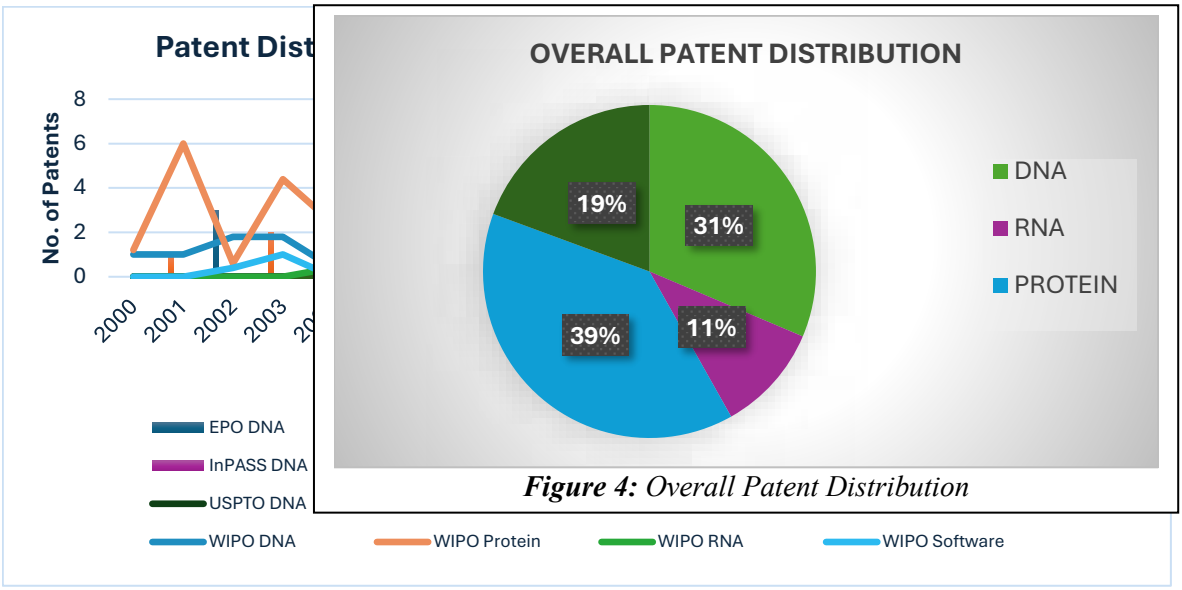


Figure 4: *Overall Patent Distribution*

Figure 3*: Patent Distribution by Category Across Major Patent Offices*

EPO: A total of 54 Bioinformatics patents were found from the year 2001 to 2019, using the specified parameters, and among them, 44 are patents classified under biological compounds, and 10 patents were found under the software category.

1. **DNA PATENT:** A total of 24 patents were found from the database using the chosen parameters.
2. **RNA Patent:** A total of 11 patents were found from the database using the chosen parameters.
3. **Protein patent:** 9 patents were found in the database using the chosen parameters.
4. **Software patent:** 10 patents were found in the database using the chosen parameters.

4. CONCLUSION

After collecting all retrieved patents, a total of 650 Bioinformatics patents were found, of which 524 are categorised as biological patents and 126 as software patents. Among all the patents, Protein has the highest number, at 38.78%, followed by DNA (31.38%), software (19.38%), and RNA (10.46%).

5. REFERENCES

Arora, D. (2022). Biotechnology Patenting: Comparison between USA and India. *Indian Journal of Integrated Research in Law, II*(V), 1. Retrieved from https://ijirl.com/wp-content/uploads/2022/09/BIOTECHNOLOGY-PATENTING-COMPARISON-BETWEEN-USA-AND-INDIA-.pdf

Barbieri, M., & Andreoni, G. (2023). Carbon Allotrope-Based Textile Biosensors: A Patent Landscape Analysis. *Engineering Proceedings, 58*(1), 107. Retrieved from https://www.mdpi.com/2673-4591/58/1/107

Birkbeck, C. D. (2016). The world intellectual property organisation (WIPO): A reference guide. In *The World Intellectual Property Organization (WIPO)*: Edward Elgar Publishing.

Chang, J., & Zhu, X. (2010). Bioinformatics databases: Intellectual property protection strategy. *Journal of Intellectual property rights, 15*(6), 447-454. Retrieved from

http://docs.manupatra.in/newsline/articles/Upload/DD3C8F8C-78E5-4B1B-96D3-B81C1436ADA0.pdf

Chatterjee, T. (2017). Patenting Bioinformatics Innovations: Emerging Trends and Challenges in the United States. *International Journal of Advanced Research (IJAR), 5*(5), 2219-2231. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2984980

Fernandez, D., & Chow, M. (2005). Intellectual property strategy in bioinformatics and biochips. *Biosensors and Bioelectronics, 21*(1), 197-200. doi:10.1016/j.bios.2004.09.014

Gopalan, R. (2009). Bioinformatics: scope of intellectual property protection. *Journal of Intellectual property rights, 14*(1), 46-51. Retrieved from https://docs.manupatra.in/newsline/articles/Upload/4277A947-FC79-4F6D-B213-82A67D76754B.pdf

Harrison, R. (2003). Protecting Innovation in Bioinformatics and In-Silico Biology. *BioDrugs, 17*(4), 227-231. doi:10.2165/00063030-200317040-00001

Lerner, J. (1994). The importance of patent scope: an empirical analysis. *The RAND Journal of Economics*, 319-333. doi:10.2307/2555833

Mago, N., Deshpande, N., & Hirwani, R. R. (2017). A landscape of bioinformatics patents - Garnering of IPR in the field of bioinformatics. *World Patent Information, 51*, 66-78. doi:10.1016/j.wpi.2017.11.007

Marco, A. C., Myers, A., Graham, S. J., D'Agostino, P., & Apple, K. (2015). The USPTO patent assignment dataset: Descriptions and analysis. *USPTO Economic Working Paper 2015-2*. doi:10.2139/ssrn.2849634

Maschio, T., & Kowalski, T. (2001). Bioinformatics – a patenting view. *Trends in Biotechnology, 19*(9), 334-339. doi:10.1016/S0167-7799(01)01702-4

Mishra, B., & Bonde, G. V. (2022). Patent Searching. In V. A. Saharan (Ed.), *Computer Aided Pharmaceutics and Drug Delivery: An Application Guide for Students and Researchers of Pharmaceutical Sciences* (pp. 473-503). Singapore: Springer Nature Singapore.

Park, H.-S. (2012). Preliminary Study of Bioinformatics Patents and Their Classifications Registered in the KIPRIS Database. *Genomics Inform, 10*(4), 271-274. doi:10.5808/GI.2012.10.4.271

Shilpa Gupta, U. (2018). Intellectual Property Rights and Bioinformatics: An Introduction. In *Bioinformatics: Sequences, Structures, Phylogeny* (pp. 1-14): Springer.

Singh, K. K. (2014). Intellectual property protection in bioinformatics and open bio development. *Asian Biotechnology Development Review, 16*(3), 25-45. Retrieved from https://ssrn.com/abstract=2611626

Wallace, M. K. (2016). Patents, Trademarks, and Intellectual Property. In A. L. S. Cassandra J. Hartnett, Eric J. Forte (Ed.), *Fundamentals of Government Information. Mining, Finding, Evaluating, and Using Government Resources* (2nd ed., pp. 335). Chicago, IL: Neal- Schuman.

A REVERSE VACCINOLOGY APPROACH FOR THE IDENTIFICATION OF POTENTIAL VACCINE CANDIDATES FROM *ECHINOCOCCUS GRANULOSUS*

Sangna Hitesh Dholakia[1], NishitaVaishnav[2], Dimple K. Kachhadiya[1], John J. Georrge[2,3*]

[1] *Department of Microbiology, Atmiya University, Rajkot, Gujarat, India*
[2] *Department of Bioinformatics, Christ College, Rajkot, Gujarat, India*
[3] *Department of Bioinformatics, University of North Bengal, District-Darjeeling, West Bengal-734013, India*
**Corresponding author: johnjgeorrge@gmail.com*

ABSTRACT: Infectious diseases, caused by the invasion of pathogenic microbes, are a major focus of drug design and vaccine development. This can be achieved through host immune responses or drug administration, but, for several reasons, the host immune response may be insufficient to control pathogenicity. In such cases, the drugs are administered, but these chemical entities are endowed with copious fatal side effects as well. Thus, there is a pressing need to prevent microbial infections rather than eliminate them after entry, leading to the development of vaccines to combat these diseases. The impending approach of computational vaccinology along with reverse vaccinology is employed in the work presented, which abets in reducing the wet lab work to a huge extent and to provide highly specific vaccine candidates, which can be further confirmed through wet lab experiments; the work aims at predicting vaccine candidates from a cestode, *Echinococcus granulosus*, a causative agent of Echinococcosis, using assorted bioinformatics tools viz., MHCPred: Prediction of MHC class I and II epitopes, BCPREDS: Prediction of B cell epitopes, EpiJen: Prediction of T cell epitopes, EMBOSS antigenic: Confirming the antigenicity of vaccine candidates.

KEYWORDS: Infection, Reverse vaccinology, *In silico*, Immunoinformatics

1. INTRODUCTION

Infection is the infiltration of host tissues by pathogenic microbes, including bacteria, protozoans, encapsulated viruses, parasites, and fungi, after multiplication into pathogenic progeny, accompanied by abnormal host responses due to toxin release. Such pathogenic microbes are transmitted from person to person, from animal to person, and vice versa, or via nosocomial means (UNICEF & Organization, 2013). Zoonotic infections, such as Echinococcosis, cysticercosis, taeniasis, trematodosis, and trichinellosis, imply that transmission of pathogens from animals to humans has increased considerably. A global death rate of 2-3 million is attributed to Echinococcosis, an infection caused by Echinococcus granulosus, and its prevalence has increased over time (Bristow et al., 2012). The current scenario makes it clear that an alternative, safer way to treat Echinococcus infection is not only desired but also cannot be ignored.

E. granulosus is a member of the genus Echinococcus, branching from the Cestoda class of the Eukaryota superkingdom, infecting the definitive hosts, canids, from which the eggs are relocated to intermediate hosts like sheep, goats, etc. The eggs develop into the embryonic stage, further infecting humans through physical contact with infected intermediate hosts (NCBI-Taxonomy, 2012) (Figure 1). Treatment of such a zoonotic infection utilises a set of anti-echinococcosis benzimidazole drugs like Albendazole and Mebendazole, with a decline in neurological disorders. This results in a need for better candidates to hamper the infection (Knox et al., 2010).

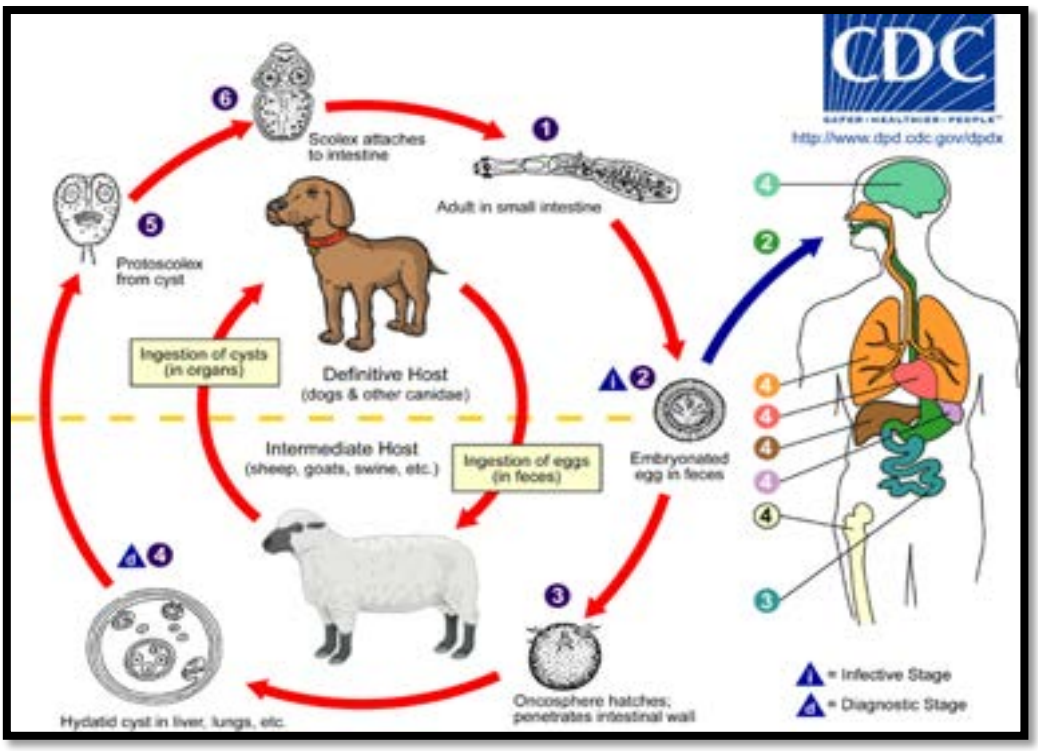


Figure 1: Life cycle of *Echinococcus granulosus*

Pathogens are also outsmarting human intelligence by developing multidrug resistance (MDR), making it unwise to rely on a single drug. Refractory cases of such infections are difficult to treat, necessitating high drug doses that increase sensitivity and requiring immediate withdrawal of treatment (Sood & Kapil, 2008), even at standard doses. Alternatives to antimicrobial drugs are vaccines, which, in case of infection, target pathogenic factors, i.e., surface proteins of microbes that can attach to host cells and subsequently release toxins; thus, these surface proteins are targets for drug and vaccine design. This is of greatest significance, as vaccines train the immune system by mimicking a natural infection, thus providing a basis for predicting vaccine candidates from *E. granulosus* (Bøhle et al., 2011).

Also, Shiga toxin-producing *E. coli*, *lpfA, paa, iha,* and *toxB* genes, as well as the genes *tsol16, to16, tsol18, and tsa18* of cestodes like Taenia ovina, Taenia solium, and Taenia saginata, encode pathogenic adhesins *that* have been targeted for vaccine development (Gauci & Lightowlers, 2003; Gomes et al., 2011). There are several types of vaccines in use: live or attenuated, inactivated, subunit, toxoid, conjugate, DNA vaccine, and recombinant vector vaccine. Still, they are burdened by major limitations, such as prolonged development timelines and high costs, which hinder vaccine design. A new approach to vaccine informatics highlights the expansion and relevance of bioinformatics techniques in preclinical vaccine development, demonstrating a breakthrough. Also, recurrent inflow of genomic information from various sequencing centres and private laboratories provides bulk raw data that must be analysed to be fabricated into useful knowledge. Various immunoinformatics algorithms have been developed for this to predict MHC class, T- and B-cell epitopes for epitope vaccine development. This is attributed to the reverse vaccinology approach (Major et al., 2011). Mathematical models of host-pathogen interactions that can be computationally simulated also serve in vaccine development protocols. Alongside vaccine production, these protocols are utilised for diverse purposes, including the development of immunisation registries, the appraisal of vaccine safety and value, and immunisation modelling (He et al., 2010). Successful vaccine development through reverse vaccinology is explicitly provided in Tables 1 and 2.

To undertake vaccine development, a range of information technology modules, including databases and tools (Tables 3 and 4), is available, enabling the development of a highly precise protocol for this procedure.

Table 1: Vaccines developed through the reverse vaccinology approach

Sr. No.	Vaccines	References
1.	*Anaplasma marginale*	PMID: 22066488
2.	*Schistosoma spp*	PMID: 21458582
3.	*Neisseria meningitides*	PMID: 21993656
4.	*Listeria monocytogenes*	PMID: 21791233
5.	*Trypanosoma brucei*	PMID: 21349321
6.	*Corynebacteriumpseudo tuberculosis*	PMID: 23201561
7.	*Stryptococcus spp*	PMID: 23181284
8.	*Herpes simplex virus type 2*	PMID: 22649465
9.	*Mycobacterium tuberculosis*	PMID: 21756938
10.	*Human adenovirus type 55*	PMID: 23087107
11.	*Pasteurella multocida*	PMID: 22540951
12.	*Rhipicephalus microplus*	PMID: 22521592
13.	*Brachyspira hyodysenteriae*	PMID: 19179021
14.	*Leptospira borgpetersenii*	PMID: 23176980
15.	*Pasteurella multocida*	PMID: 22792202
16.	*Cryptosporidium spp*	PMID: 21918117
17.	*Ehrlichia ruminantium*	PMID: 20566221
18.	*Teladorsagia circumcincta*	PMID: 23282110

Table 2: Vaccine candidate prediction through reverse vaccinology approach

Sr. No.	Vaccine	References
1.	*Leishmania spp.*	PMID: 22434357
2.	*Campylobacter fetus*	PMID: 22890137
3.	*E coli 536*	PMID: 22493535
4.	*Plasmodium vivax*	PMID: 21713006
5.	*Mannheimia haemolytica*	PMID: 20336679
6.	*human influenza H1N1*	PMID: 22450272
7.	*Helicobacter pylori*	PMID: 22130156
8.	*Mycoplasma genitalium*	PMID: 22057004
9.	*Human papillomavirus 16 E6*	PMID: 21699918

Table 3: Immunoinformatics Databases

Name	Purpose	URL
AntigenDB	Sequence, structure, and other data on pathogen antigens	http://www.imtech.res.in/raghava/antigendb/index.html
AntiJen	Quantitative binding data for peptides and proteins of immunological interest	http://www.ddg-pharmfac.net/antijen/AntiJen/antijenhomepage.htm
BCIpep	Warehouses information of all experimentally determined B-cell epitopes of antigenic proteins.	http://bioinformatics.uams.edu/mirror/bcipep

dbMHC	An open, publicly accessible platform for DNA and clinical data related to the human Major Histocompatibility Complex (MHC).	http://www.ncbi.nlm.nih.gov/gv/mhc/main.fcgi?cmd=init
DIGIT	Database of ImmunoGlobulin sequences and Integrated Tools.	http://www.biocomputing.it/digit4
GPX-Macrophage Expression Atlas	Database for expression profiles of macrophages challenged with a variety of pro-inflammatory, anti-inflammatory, benign and pathogen insults.	http://gpxmea.gti.ed.ac.uk
HPTAA	Database of potential tumour-associated antigens that uses expression data from a range of expression platforms, including vigilantly selected overtly accessible microarray expression data, GEO SAGE data and Unigene expression data.	http://www.bioinfo.org.cn/hptaa
IEDB	Contains data related to antibody and T-cell epitopes for humans, non-human primates, rodents, and other animal species.	http://www.iedb.org
IEDB-3D	Structural data within the Immune Epitope Database.	http://www.iedb.org/bb_structure.php
IL2Rgbase	X-linked severe combined immunodeficiency mutations	http://www.ncbi.nlm.nih.gov/lovd/home.php?select_db=IL2RG
IMGT®	High-quality integrated knowledge resource specialised in immunoglobulins (IG) or antibodies, T-cell receptors (TR), major histocompatibility (MH) of human and other vertebrate species, and in the immunoglobulin (IgSF) and MH superfamily (MhSF) and related proteins of the immune system (RPI) of vertebrates and invertebrates.	http://www.imgt.org
IMGT_GENE-DB	The IMGT® comprehensive genome database for IG and TR genes from human and mouse (in development, from other vertebrate species).	http://www.imgt.org/IMGT_GENE-DB/GENElect?livret=0
IMGT/HLA	Provides a specialist database for sequences of the human major histocompatibility complex (HLA)	http://www.ebi.ac.uk/ipd/imgt/hla
IMGT/LIGM-DB	The first and the largest database of IMGT®, the international ImMunoGeneTics information system®, the high-quality integrated knowledge resource specialised in IG, TR, MHC of humans/other vertebrate species, and related proteins of the immune system (RPI) that belong to the IgSF and MHC superfamily (MhcSF).	http://www.imgt.org/cgi-bin/IMGTlect.jv
IFN-Stimulated Gene Database	Contains information on >400 microarray-based IFN-stimulated genes.	http://www.lerner.ccf.org/labs/williams/xchip-html.cgi
IPD-ESTDAB	Permits search for HLA-typed, immunologically-characterised tumour cells available from a central bank.	http://www.ebi.ac.uk/ipd/estdab

IPD-HPA - Human Platelet Antigens	Provides a centralised repository for data defining human platelet antigens (HPA).	http://www.ebi.ac.uk/ipd/hpa
IPD-KIR - Killer-cell immunoglobu-lin-like Receptors	Provides a centralised repository of human KIR sequences. Killer cell immunoglobulin-like receptors (KIRs) are highly polymorphic at the allelic and haplotypic levels.	http://www.ebi.ac.uk/ipd/kir
IPD-MHC	Provides a centralised repository of Major Histocompatibility Complex (MHC) sequences from multiple species.	http://www.ebi.ac.uk/ipd/mhc
MHCBN	A curated database consisting of detailed information about MHC-binding/Non-binding peptides and T-cell epitopes	http://www.imtech.res.in/raghava/mhcbn
MHCPEP	Data repository bridging the gap between immunological and computer science/machine learning communities by providing pre-processed, scaled immunological datasets for use in machine learning applications.	http://bio.dfci.harvard.edu/DFRMLI
MUGEN	MUGEN Mouse Database (MMdb) is a fully searchable database of murine models of immune processes and immunological diseases.	http://bioit.fleming.gr/mugen/mde.jsp
Protegen	The host's acquired immune response specifically targets protective antigens and can induce protection against infectious and non-infectious diseases.	http://www.violinet.org/protegen
SuperHapten	Current database version compiles 2D/3D structures, physicochemical properties, and references for about 7500 haptens and 25,000 synonyms.	http://bioinformatics.charite.de/superhapten
VBASE2	Integrative database of germ-line variable genes from the immunoglobulin loci of human and mouse.	http://www.vbase2.org
Epitome	Database of all known antigenic residues and antibodies that interact with them, including a detailed description of residues involved in the interaction and their sequence/structure environments.	http://rostlab.org/services/epitome

Table 4: Immunoinformatics Tools

Name	**Function**	**Algorithm**	**Accuracy**	**Stand-alone/Web-based**	**URL**
VaxiJen	First server for alignment-independent prediction of protective antigens.	Auto cross covariance (ACC).	70%-89%	Web-based	http://www.jenner.ac.uk/VaxiJen.

PSortb	Used for bacterial protein sub-cellular localisation prediction	Support Vector Machine (SVM)	80–95%	Stand-alone and web-based	http://www.psort.org/psortb
WoLF PSORT	Converts protein amino acid sequences to numerical localisation features.	k-nearest neighbour	70%	Stand-alone (Apache and Linux) and Web-based	wolfpsort.org
BEST	Predicts B-cell epitopes from antigen sequences.	SVM	74.5%	Stand-alone	http://biomine.ece.ualberta.ca/BEST/
SVMTriP	Capable of recognising viral peptides from a human protein sequence background	SVM	80.1%	Web-based	http://sysbio.unl.edu/SVMTriP
BCPREDS	Tool for predicting linear B-cell epitopes	SVM	74.5%	Web-based	http://ailab.cs.iastate.edu/bcpreds/.
TMHMM	Used for transmembrane helix prediction	Hidden Markov Model (HMM)	77%	Stand-alone (UNIX) and Web-based	http://www.cbs.dtu.dk/services/TMHMM/
LocateP	Pipeline designed to mimic protein targeting and secretion processes. Used to predict sub-cellular localisation of proteins	TMHMM 2.0, SignalP 3.0, PrediSi, Phobius, CELLO and Psortb	90%	Web-based	http://www.cmbi.ru.nl/locatep-db/cgi-bin/locatepdb.py
ProPred	Predicts MHC II binding regions in antigenic protein sequences	Matrix-based prediction algorithm	80%	Web-based	http://www.imtech.res.in/raghava/propred/
T-Epitope Designer	Facilitates HLA-peptide binding prediction.	Uses 3D models of MHC peptides and Q matrix	60%	Web-based	http://www.bioinformation.net/ted/
Pepitope	Predicts discontinuous epitopes	PepSurf/ Mapitope	69%	Web-based	http://pepitope.tau.ac.il/
MimoPro	Used for epitope prediction	Dynamic programming and branch and bound optimisation	54%	Web-based	http://informatiin.nenu.edu.cn/MimoPro

ProPred 1	Online tool for prediction of peptide binding to MHC Class-I alleles.	Matrix-based approach	80%	Web-based	http://www.imtech.res.in/raghava/propred1/
NetChop	Combines predictors of proteasomal processing, TAP trans-port, and MHC binding to produce overall score for each peptide's intrinsic potential of being a T-cell epitope	Neural Network	70%	Web-based	http://tools.immuneepitope.org/stools/netchop/netchop.do?app=netchop

2. MATERIALS AND METHODS

2.1. Data acquisition and scaling

The proteome (10241 sequences) of Echinococcus granulosus was downloaded from the Wellcome Trust Sanger Institute, where research builds an understanding of gene function in health and disease and constructs resources of lasting value to biomedical research. (Wellcome Trust Sanger Institute, 2013). The CD-HIT tool, based on a greedy algorithm, was utilised to process the downloaded sequences in the Multidimensionality Scaling (MDS) step, with a 70% cut-off. This resulted in reference and non-redundant sequences in the form of unique sequence clusters (Huang et al., 2010).

2.2. BLAST and Sub-cellular localisation

The resulting sequences were then processed through the SubCellular Localisation predictor (CELLO) (Yu et al., 2004), which predicted localisation. Those qualifying as "plasma membrane proteins" were selected, as these proteins play a major role in host-pathogen interactions, making them potential vaccine candidates. The next tool, TMHMM (Krogh et al., 2001), was used to predict transmembrane helices, with CELLO results as input. From the resultant sequences, only those with TM = 1 were chosen for the following steps, as more than 1 transmembrane domain in proteins poses an obstacle to cloning and expression *in vitro* (Pizza et al., 2000). The conclusion of CELLO and TMHMM were searched for homology against human and mouse proteomes, subsequently, using organism-specific Basic Local Alignment Search Tool (BLASTP), a highly used tool to find exact alignments that optimise and assess local similarity, giving the maximal segment pair (MSP) score (Altschul et al., 1990). Sequences non-homologous to human and mouse were selected for further processing in the protocol.

2.3. MHC class I and II prediction

Further, for the prediction of MHC Class I and Class II binding epitopes, MHCPred, a freely available online tool implementing partial least squares-based and robust multivariate statistical methods, was used. The significance of the current step lies in the involvement of MHC Class I and II molecules in the immune response for recognition and presentation of epitopes to B-cells and T-cells (Guan et al., 2003). The sequences with a confidence value of 1 were selected for the following steps.

2.4. B-Cell epitopes prediction

The protocol advances with the prediction of B-cell epitopes, which are key molecules in the generation of the humoral immune response, using BCPREDS. It is an online tool with a 74.5% prediction accuracy. Sequences with a score less than 1 were eliminated.

2.5. T-Cell epitopes prediction

Sequences predicted from BCPREDS were used as input to EpiJen to predict T-cell epitopes. Prediction of T-cell epitopes is important due to their role in eliciting cell-mediated immune responses. The tool is based on a quantitative matrix derived from an additive method and provides an accuracy of 85% (Doytchinova et al., 2006). The processed sequences with an IC50 value less than 50 were selected as

final probable vaccine candidates. IC50 values show an inverse relationship with the protein's binding affinity across varied HLA alleles, implying that lower IC50 values indicate higher binding affinity (Rai et al., 2012).

2.6. Antigenic protein prediction

The final vaccine candidates were further confirmed for their antigenicity; the same Antigenic module of the EMBOSS package was used, which is potent enough to predict essentially antigenic regions of the protein with an accuracy of 75% (Kolaskar & Tongaonkar, 1990). The predicted antigenic sequences were provided with a range of tools for function predictions.

2.7. Function Prediction

PSI-BLAST predicted all antigenic proteins as hypothetical (Altschul et al., 1990); subsequently these were annotated using variety of tools namely, InterproScan (Syed & Upton, 2006), CDD (Marchler-Bauer et al., 2012), COGnitor (Tatusov et al., 2000), Phyre2 (Bennett-Lovsey et al., 2008), PDB-BLAST, PFAM (Finn et al., 2010), Protfun (Jensen et al., 2002) and PFP (Hawkins et al., 2009). Further, I-TASSER was used to model the proteins and predict function; the modrefine module of I-TASSER was also employed to refine the modelled structures (Zhang, 2008). The templates used by I-TASSER were structurally compared with antigenic protein structures using DaliLite, a program for pairwise structure comparison and structure database searching, and its web interface provides 3D superimposition and multiple alignment views. Even function prediction was carried out by ProFunc server (Laskowski et al., 2005).

2.8. Epitope prediction

To predict the antigenic portion from the 3D structure of the antigenic protein, the Pepitope server was used with the default parameters. Pepitope is based on the Mapitope algorithm, which is designed to predict at most 3 epitope candidates at the atomic level from the antigenic protein structure (Bublil et al., 2007).

2.9. HLA alleles and paratope mining

HLA alleles were obtained from SYFPEITHI, a database comprising more than 7000 peptide sequences known to bind with class I and class II MHC molecules. HLA alleles crossing the score of 20 were selected to be further studied in the protocol (Rammensee et al., 1999). The 3D structures of these HLA alleles were retrieved from the PDB (PDB ID: 1DLH) (Bennett-Lovsey et al., 2008). The paratope residues of HLA alleles were also obtained from the SYFPEITHI database. Their position was confirmed using the Swiss-PDB-Viewer (SPDBV), which allows superimposition of proteins to construct structural alignments and evaluate their active sites and other relevant features (Guex & Peitsch, 1997).

2.10. Protein- Protein Docking

Protein–protein docking is carried out between antigenic proteins and HLA alleles to confirm the affinity of the antigen to bind with the HLA allele and initiate the immune response in the host. The ClusPro tool was employed for the same, using the Fast Fourier Transform (FFT) algorithm (Comeau SR, 2004), and has been selected as the best-performing server in the latest rounds of the CAPRI (Critical Assessment of PRediction of Interactions).

3. RESULTS

The proteome of *Echinococcus granulosus* was retrieved from the Wellcome Trust Sanger Institute and consisted of 10,241 protein sequences. These were clustered using CD-HIT, and 9889 unique sequence clusters were obtained. CELLO and TMHMM were performed on the resulting sequences, yielding 7128 and 1511 plasma membrane sequences, respectively. BLASTP against Human and Mouse produced outputs containing 170 and 168 non-homologous sequences, respectively, of which 168 sequences were used as input to MHCPred. 134 sequences with MHC class I and II binding epitopes were predicted and provided to BCPREDS, yielding 80 promising B-cell epitope-containing sequences. Further, T cell epitope prediction by EpiJen identified the 11 most promising vaccine candidates. EMBOSS antigenic was further employed to assess the antigenicity of the predicted vaccine candidates and provided satisfactory results, predicting that all 11 candidates were antigenic. Even the tools used for functional annotation gave indistinguishable results. I-TASSER and ProFunc also furnished indistinct results. DaliLite confirmed no significant structural similarity between the antigenic protein and the human protein, but did identify only a few single-amino-acid matches between the target and

template. Thus, proving that the predicted antigenic proteins are novel and lack any clear, promising functions. As a preliminary step, the Pepitope server for antigenic portion prediction from the modelled 3D structure provided satisfactory results, including the residues involved in antigenicity and their number (Figures 2 and 3). Sequences EgrG_000087000.1, EgrG_000101700.1, EgrG_000155300.1 and EgrG_002067700.1 each contained one epitope candidate, while the remaining protein sequences had 2-3 predicted candidates each. The SYFPEITHI database identified HLA-DR-B as the most potent binder of the given antigenic sequences, with the paratope region (Table 5). The modelled antigenic structure and HLA-DR-B were docked using ClusPro, and the most optimally docked complexes (i.e., the 0th complexes) with energy values are shown in Table 6 (Comeau SR, 2004).

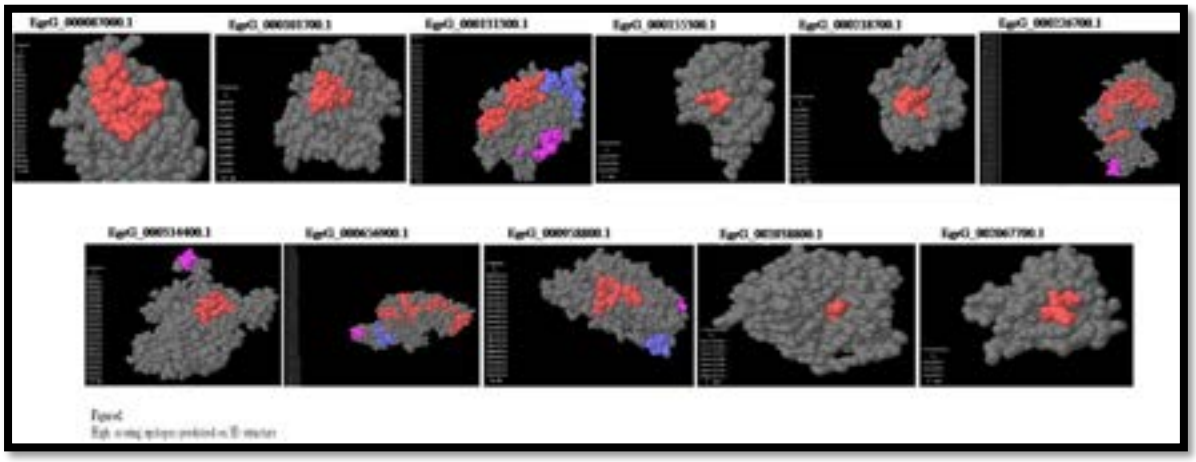


Figure 2: All antigenic epitopes on the 3D antigen structure

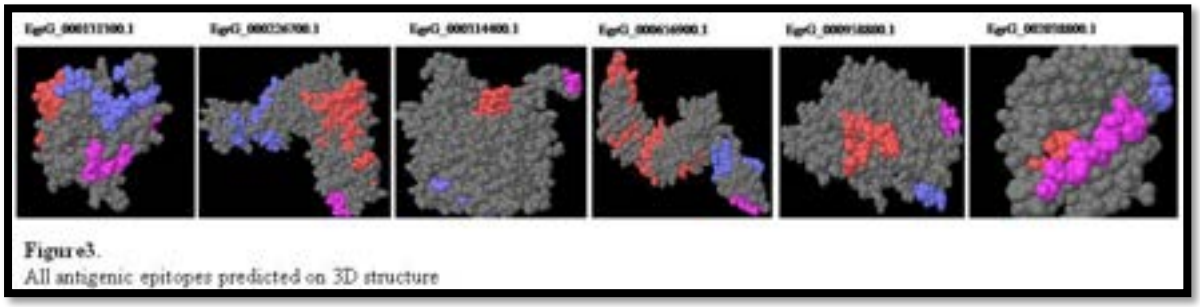


Figure 3: High-scoring antigenic epitopes on 3D antigenic structure

4. DISCUSSION

The prediction of highly antigenic proteins from the retrieved proteome of *E. granulosus* was the foremost endeavour of the presented work. The former use of CD-HIT and subsequently protein subcellular localisation was of utmost importance, as for a protein to be considered as a potent vaccine candidate, it must be present in the outer membrane and protruding out (Chen et al., 2012); for the same, predicted outer membrane proteins were selected for further substantiation of their antigenicity. The BLAST search of the proteome against human and mouse helped remove homologous sequences, because homology between the E. granulosus protein and those of human and mouse leads directly to autoimmune reactions during vaccine administration (John et al., 2012). Further, the analysis of MHC class I and II binding epitopes, along with B cell and T cell epitopes, on the predicted antigenic proteins yielded reasonable results, proving the selected proteins as the most promising and probable vaccine candidates. The function prediction of all vaccine candidates did not provided distinct results, subsequently, antigenic module of EMBOSS was used which generated unique sequence clusters based on sequence identity and potent antigenic sequence from the provided protein sequences, abetting in providing highly precise and needful data also the task was surpassed by the antigenic portion prediction along with residues and its number, paratopic region of HLA allele via Pepitope server, SYFPEITHI and SPDBV respectively, of the antigenic protein structures which were modelled using I-TASSER.

Protein_ID	Epitope (Antigen)	Paratope (MHC classes)
EgrG_000087000.1	H6, S7, A8, N9, T10, R11 A12, H13, P14, H15, A16, R17, T18, H19, A20, G21, T22, H23, A24, H25, F26, H254, F256	C107, T113, Y150, L151, Y161, A182, W61, L67, A73, I127, P165, Q169, G174, S179, H5
EgrG_000101700.1	N78, C79, G80, Q81, P82, N83, I84, C85, I86, V99	M23, R44, P81, F108, M160, S167, L184, I7, S19, L45, L60, S77, T80, V85, L99
EgrG_000131300.1	L136, V138, T145, T147, T147, S151, I152, T155, T156, P157, L161, Y162, T163, G164, L165, K166, M167, Q168, S169, Y170, Y171, G172, L173, I180, D181, N182, T183	N15, R44, N62, V65, L70, C107, Y121, M160, G168, L184, L314, A317, H5, I7, Y43
EgrG_000155300.1	G103, E104, Q105	P16, Q18, S19, Y79, G125, T129, A182, N20, N33, R39, N62, V75, L109, C117, Y123
EgrG_000218700.1	I49, I50, V51, I52, I53, I54, I55, I56, Y57, C58, C59	T41, Q57, T74, F108, G125, V128, D142, S156, R164, V165, A182, Y9, C15, P56, T77
EgrG_000226700.1	T370, A381, W383, S384, C387, T388, T389, A390, A398, S409, A410, T411, T413, T414, A415, F416, A417, A418, T419, A418, T419, A420, T421, T422, T423, T424, T425, T426, S427, L429, S430, S431, S432, I433, K434, T435, T436, E437, T438, A439	P16, R44, G49, V65, I72, R76, L92, N94, T113, S133, F137, M160, S167, V170, E176
EgrG_000314400.1	E13, D422, G447, E485, D506, G507, E508, V509, M510, N511, E512, G513, E514, E515, E516, D517, G518, I519, P520, E521, E522, G523	L14, H33, E40, L45, E46, F48, A52, S53, T74, T80, L92, V97, L99, F108, F137
EgrG_000656900.1	L7, I8, S9, L10, L11, S12, V14, L15, A16, A18, L19, E26, E27, L34, S35, A36, D37, Q38, E39, L41, V54, D59, L60, N62, D63, L77, P78, E79, G80, K82, D83, L97, L102, E103, L104, E120, L122, Q123, V124, G125, T126	P16, G20, T41, Y79, T80, Y150, L154, L170, A182, D41, S42, S63, N82, I127, T145
EgrG_000958800.1	N323, T324, A325, I326, A347, H348, S350, Q352, P353, L354, Y355, T356, T357, D358, G359, S360, L371, Y372	S19, I31, L45, L60, A61, L70, L92, F108, G125 S133, T135, S156, I31, A59, F89
EgrG_002038800.1	P145, P146, S147, P148, S149, S150, F151	T41, A52, F54, F108, A182, T3, W9, S37, S42, L53, H81, M160, D17, P81, L92
EgrG_002067700.1	G93, E94, E95	I7, I8, H33, L45, P86, P87, L92, A182, A49, R55, A58, V75, L109, V116, L147

Also, the binding affinity of the antigen with the HLA allele was confirmed using ClusPro, which yielded satisfactory low-energy values. Leading to the achievement of the endeavour by predicting most potential vaccine candidates for a virulent strain of *Echinococcus* i.e. *Echinococcus granulosus*, causing a fatal disorder in the intermediate host, i.e. *Homo sapiens*, of the severe disease called Echinococcosis, it also aims at a sooner development of the peptide-based vaccine for the disease and its rapid use for the eradication the disease, which has started showing its spread world-wide.

Table 5: Epitopes and Paratopes

5. CONCLUSION

A range of Bioinformatics tools provides a completely novel protocol for improving vaccine development strategies. It has led to advances in computational vaccinology. Vaccine candidates predicted from *Echinococcus granulosus* using a computational reverse vaccinology approach appear to be most potent. Still, substantiation of these proteins in wet-lab settings is of utmost importance for their initiation in vaccination programmes. In the presented work, the 11 most potent vaccine candidates were selected from a bulk proteome containing 10,241 sequences using a wide range of immunoinformatics approaches, aiming to reduce wet-lab studies.

Table 6: Docking Score

Protien_ID	Lowest energy values
EgrG_000087000.1	-1109.9
EgrG_000101700.1	-1142.2
EgrG_000131300.1	-1195.6
EgrG_000155300.1	-906.6
EgrG_000218700.1	-1163.8
EgrG_000226700.1	-1368.4
EgrG_000314400.1	-1273.3
EgrG_000656900.1	-1177.4
EgrG_000958800.1	-1219.2
EgrG_002038800.1	-1418.7
EgrG_002067700.1	-1084

6. REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3), 403-410.

Bennett-Lovsey, R. M., Herbert, A. D., Sternberg, M. J., & Kelley, L. A. (2008). Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins: Structure, Function, and Bioinformatics*, *70*(3), 611-625.

Bøhle, L. A., Riaz, T., Egge-Jacobsen, W., Skaugen, M., Busk, Ø. L., Eijsink, V. G., & Mathiesen, G. (2011). Identification of surface proteins in Enterococcus faecalis V583. *BMC genomics*, *12*(1), 135.

Bristow, B. N., Lee, S., Shafir, S., & Sorvillo, F. (2012). Human echinococcosis mortality in the United States, 1990–2007. *PLoS Neglected Tropical Diseases*, *6*(2), e1524.

Bublil, E. M., Freund, N. T., Mayrose, I., Penn, O., Roitburd-Berman, A., Rubinstein, N. D., Pupko, T., & Gershoni, J. M. (2007). Stepwise prediction of conformational discontinuous B-cell epitopes using the Mapitope algorithm. *Proteins: Structure, Function, and Bioinformatics*, *68*(1), 294-304.

Chen, X., Xu, Z., Li, L., Chen, H., & Zhou, R. (2012). Identification of conserved surface proteins as novel antigenic vaccine candidates of Actinobacillus pleuropneumoniae. *Journal of microbiology*, *50*(6), 978-986.

Doytchinova, I. A., Guan, P., & Flower, D. R. (2006). EpiJen: a server for multistep T cell epitope prediction. *BMC bioinformatics*, *7*(1), 131.

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., & Forslund, K. (2010). The Pfam protein families database. *Nucleic acids research*, *38*(suppl_1), D211-D222.

Gauci, C., & Lightowlers, M. W. (2003). Molecular cloning of genes encoding oncosphere proteins reveals conservation of modular protein structure in cestode antigens. *Molecular and biochemical parasitology*, *127*(2), 193-198.

Gomes, T. A., Hernandes, R. T., Torres, A. G., Salvador, F. A., Guth, B. E., Vaz, T. M., Irino, K., Silva, R. M., & Vieira, M. A. (2011). Adhesin-encoding genes from Shiga toxin-producing Escherichia coli are more prevalent in atypical than in typical enteropathogenic E. coli. *Journal of Clinical Microbiology*, *49*(9), 3334-3337.

Guan, P., Doytchinova, I. A., Zygouri, C., & Flower, D. R. (2003). MHCPred: bringing a quantitative dimension to the online prediction of MHC binding. *Applied bioinformatics*, *2*(1), 63-66.

Guex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *electrophoresis*, *18*(15), 2714-2723.

Hawkins, T., Chitale, M., Luban, S., & Kihara, D. (2009). PFP: Automated prediction of gene ontology functional annotations with confidence scores using protein sequence data. *Proteins: Structure, Function, and Bioinformatics*, *74*(3), 566-582.

He, Y., Rappuoli, R., De Groot, A. S., & Chen, R. T. (2010). Emerging vaccine informatics. *BioMed Research International*, *2010*(1), 218590.

Huang, Y., Niu, B., Gao, Y., Fu, L., & Li, W. (2010). CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics*, *26*(5), 680-682.

Jensen, L. J., Gupta, R., Blom, N., Devos, D., Tamames, J., Kesmir, C., Nielsen, H., Stærfeldt, H. H., Rapacki, K., & Workman, C. (2002). Prediction of human protein function from post-translational modifications and localization features. *Journal of molecular biology*, *319*(5), 1257-1265.

John, L., John, G. J., & Kholia, T. (2012). A reverse vaccinology approach for the identification of potential vaccine candidates from Leishmania spp. *Applied biochemistry and biotechnology*, *167*(5), 1340-1350.

Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., & Neveu, V. (2010). DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic acids research*, *39*(suppl_1), D1035-D1041.

Kolaskar, A. S., & Tongaonkar, P. C. (1990). A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS letters*, *276*(1-2), 172-174.

Laskowski, R. A., Watson, J. D., & Thornton, J. M. (2005). ProFunc: a server for predicting protein function from 3D structure. *Nucleic acids research*, *33*(suppl_2), W89-W93.

Major, M., Moss, S., & Gold, R. (2011). From genes to vaccine: A breakthrough in the prevention of meningococcal group B disease. *Paediatrics & Child Health*, *16*(8), e61-e64.

Marchler-Bauer, A., Zheng, C., Chitsaz, F., Derbyshire, M. K., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., & Lanczycki, C. J. (2012). CDD: conserved domains and protein three-dimensional structure. *Nucleic acids research*, *41*(D1), D348-D352.

Pizza, M., Scarlato, V., Masignani, V., Giuliani, M. M., Arico, B., Comanducci, M., Jennings, G. T., Baldi, L., Bartolini, E., & Capecchi, B. (2000). Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science*, *287*(5459), 1816-1820.

Rai, J., Lok, K. I., Mok, C. Y., Mann, H., Noor, M., Patel, P., & Flower, D. R. (2012). Immunoinformatic evaluation of multiple epitope ensembles as vaccine candidates: E coli 536. *Bioinformation*, *8*(6), 272.

Rammensee, H.-G., Bachmann, J., Emmerich, N. P. N., Bachor, O. A., & Stevanović, S. (1999). SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics*, *50*(3), 213-219.

Sood, S., & Kapil, A. (2008). An update on Trichomonas vaginalis. *Indian Journal of Sexually Transmitted Diseases and AIDS*, *29*(1), 7-14.

Syed, A., & Upton, C. (2006). Java GUI for InterProScan (JIPS): a tool to help process multiple InterProScans and perform ortholog analysis. *BMC bioinformatics*, *7*(1), 462.

Tatusov, R. L., Galperin, M. Y., Natale, D. A., & Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic acids research*, *28*(1), 33-36.

UNICEF, & Organization, W. H. (2013). *Monitor evaluate improve: TDR results 2012 report: measuring results*.

Yu, C. S., Lin, C. J., & Hwang, J. K. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein science*, *13*(5), 1402-1406.

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC bioinformatics*, *9*(1), 40.

FROM CABINETS TO CLOUD: DIGITALIZATION OF HERBARIA FOR 21ST-CENTURY BIODIVERSITY SCIENCE

[1*]Aarti Kavane, [2]Neha Jangbari & [2]Archana Rana
[1] **Assistant Professor, Botany Dept. Christ College, Rajkot*
[2] *Assistant Professor, Microbiology Dept. Christ College, Rajkot*
**Corresponding Author Email: dr.aarti@christcollegerajkot.edu.in*

ABSTRACT: Herbaria preserve plant specimens that document biodiversity, distribution, morphology, and historical change, but the value of these collections increases greatly when specimens are converted into accessible digital records. Digitalization of herbarium specimens generally involves specimen preparation, high-resolution imaging, label transcription, georeferencing, metadata capture, database entry, and long-term archiving. Large-scale initiatives such as iDigBio (Digitization workflows for biodiversity collections), GBIF (Global Biodiversity Information Facility), Botanical Survey of India (BSI) and the Royal Botanic Gardens, Kew have shown that digitized collections can be made widely accessible for taxonomy, ecology, conservation, phenology, climate-change studies, and education. Workflow-based digitization is especially important because herbarium projects often vary in funding, staff size, and collection scope; modular systems allow institutions to combine imaging, data capture, and quality control in flexible ways. Studies on herbarium digitization also show that optical character recognition, barcode-based specimen tracking, and machine-learning methods can improve throughput and unlock new analytical possibilities. At the same time, digitization creates challenges such as incomplete label data, image quality control, standardization of metadata, and the need for sustainable curation of digital files. This review article summarizes the main practices, outcomes, and limitations associated with herbarium digitalization and highlights its importance in modern biodiversity science. The growing availability of specimen images through online platforms has already transformed how researchers study species distributions and historical patterns, while also making collections more visible to students, curators, and citizen scientists. As collections continue to scale up, digital herbaria are becoming essential infrastructure for biodiversity informatics and conservation decision-making.

INTRODUCTION

Herbaria are among the most important repositories of plant diversity, preserving pressed and dried specimens that document where plants were found, when they were collected, and how they were identified. Major herbaria such as Kew's herbarium contain millions of specimens and are used in taxonomy, conservation, climate research, and historical studies. However, the physical nature of these collections traditionally limits access to users who can visit the herbarium in person. Digitalization addresses this limitation by transforming specimen information into searchable online resources (Heberling et al 2019; Graves. (2017)

The need for digitization has become more urgent as biodiversity research increasingly depends on large, interoperable datasets. GBIF (Global Biodiversity Information Facility) promotes open access to biodiversity data through shared standards such as Darwin Core, which helps make specimen records reusable across platforms (GBIF 2021). Likewise, iDigBio (Digitization workflows for biodiversity collections) has documented digitization workflows to support the herbarium community, showing that successful projects require coordination among specimen preparation, imaging, transcription, and data management. (iDigBio 2016; Figg 2024).

Digitized herbarium specimens are now being used in studies of species identification, phenology, ecological change, and conservation planning. A recent systematic review found that digitized herbarium images are already supporting computer vision and machine-learning applications, with transfer learning being a common approach in the literature. This shows that digital herbaria are no longer only archival tools; they are active research infrastructure. (Hussein et al 2022)

MATERIALS AND METHODS

This article is a review-based synthesis of published digitization workflows, institutional documentation, and recent literature on herbarium digitalization. Sources were selected from authoritative collections and biodiversity platforms, including iDigBio, GBIF, Kew, and peer-reviewed studies describing herbarium digitization workflows and analytical applications. The review focused on four components: specimen workflow, data workflow, image workflow, and downstream use of digitized records. (Haston 2012; iDigBio 2016; Table 2)

The workflow model described in the literature begins with pre-digitization curation, including checking specimen condition, arranging folders, and ensuring that names and labels are legible. Specimens are then imaged using scanners or cameras, usually with barcode-based tracking to link each sheet to a unique record. After imaging, label data are transcribed manually, semi-automatically, or with OCR support. Curatorial data and supplementary data may also be captured, and all records are then stored in a collection database and shared through standardized formats such as Darwin Core Archive. (Ellwood et al 2015; GBIF 2017)

Biodiversity Information Standards (TDWG)—commonly known as **TDWG**—is an international, non-profit organization that develops and promotes standards for sharing biodiversity data globally. It plays a central role in ensuring that biodiversity information collected by researchers, museums, herbaria, and databases can be easily exchanged, understood, and reused (TDWG 2010). The method used here also considers the role of image management and quality control. In large digitization projects, image files are created in archival and web-friendly formats, checked for clarity and completeness, and then archived for long-term preservation. This modular approach allows institutions to digitize collections in phases according to available funding and staffing (Smithsonian Institution Digitization Program Office 2016).

RESULT AND DISCUSSION

The reviewed sources show that herbarium digitization produces three major outcomes: increased accessibility, improved research utility, and better integration with biodiversity informatics systems. Kew reports that millions of its herbarium and fungarium specimens are now available online, and that digitized collections support research on climate change, extinction, agriculture, medicines, and ecosystems. Similar benefits are described by GBIF, whose standards and data-sharing framework are designed to make biodiversity records widely reusable. (Soltis 2017)

A major strength of digitization is the ability to build scalable workflows. The Royal Botanic Garden Edinburgh model separates specimen, data, and image workflows so that digitization can continue even when funding or project goals differ. This modular design is practical for herbaria that must balance ongoing curation, research access, and the need to create digital records at high speed. The iDigBio workflow papers similarly emphasize that successful digitization depends on clearly defined tasks and community-vetted procedures. (Page et al. 2015)

Digitization also expands the possibilities for computational analysis. The systematic review on computer vision and machine learning found that deep learning with transfer learning is commonly used for digitized herbarium specimens, and that collaboration between computer scientists and biologists is strong. These methods are already helping with plant identification and may improve phenological and morphological studies in the future. (Nelson, et al 2012)

At the same time, the literature highlights several challenges. Many herbarium sheets still contain incomplete, messy, or handwritten labels, which slows transcription and affects data quality. Another issue is long-term maintenance: digital files, metadata, and specimen records must remain connected as systems evolve Table 1 & 2). The need for ongoing curation is therefore as important as initial

digitization. The modular workflow approach helps address this problem, but it also requires continued standardization and coordination across institutions. (Willis et al 2017)

Overall, the evidence suggests that herbarium digitalization is not just a technical upgrade but a transformation in how botanical collections are used. Instead of being limited to local visitors, specimens become global scientific resources that support research, teaching, and conservation. (Younis et al 2018; Kew Gardens)

Table 1: Major organizations for conserving Herbarium in India

Herbarium / Organization	Country	Acronym	Approx. No. of Specimens	Online/Digitized Records Available	Remarks
Botanical Survey of India (Multiple Herbaria)	India	BSI	>3,000,000	Partial (GBIF + BSI portals)	Network of regional herbaria across India (Botanical Survey of India)
Central National Herbarium, Howrah	India	CAL	~2,000,000	Limited online	Oldest herbarium in India (Botanical Survey of India)
Forest Research Institute, Dehradun	India	DD	~350,000	Very limited	Forestry-focused collection (Botanical Survey of India)
National Botanical Research Institute, Lucknow	India	LWG	~260,000	Partial	Important for economic botany (Botanical Survey of India)
Blatter Herbarium, Mumbai	India	BLAT	~200,000	Limited	One of largest private herbaria in India (Botanical Survey of India)
BSI Southern Regional Centre	India	MH	~280,000	Limited	Regional flora (South India)
BSI Eastern Regional Centre	India	ASSAM	~260,000	Limited	North-East flora
BSI Western Regional Centre	India	BSI (Pune)	~175,000	Limited	Western India flora
BSI Northern Regional Centre	India	BSD	~110,000	Limited	Himalayan flora

Table 1: Global organizations for conserving Herbarium

Herbarium / Organization	Country	Acronym	Approx. No. of Specimens	Online/Digitized Records Available	Remarks
Royal Botanic Gardens, Kew Herbarium	UK	K	~7,000,000	~1.2+ million records online	One of world's largest herbaria (Royal Botanic Gardens, Kew)
Muséum National d'Histoire Naturelle, Paris	France	P	~8,000,000+	Large digital database	Major European herbarium
New York Botanical Garden Herbarium	USA	NY	~7,800,000	Extensive online access (NYBG portal)	Strong tropical collections
Smithsonian National Museum of Natural History	USA	US	~5,000,000	Large digitization via NMNH	Linked to global biodiversity research
Chinese National Herbarium (CAS)	China	PE	~5,000,000+	Increasing digitization	Largest in Asia
Komarov Botanical Institute	Russia	LE	~6,000,000	Limited online	Historical collections
Australian National Herbarium	Australia	CANB	~1,300,000	Strong digital access	Linked with AVH database
Naturalis Biodiversity Center	Netherlands	L	~5,000,000	High digital availability	Integrated biodiversity datasets

Herbarium specimen downloaded from GBIF and BSI

SUMMARY

Digitalization of herbaria converts traditional specimen collections into accessible scientific databases. The process includes imaging, transcription, metadata standardization, and archival storage. The literature shows that modular workflows, OCR, barcode systems, and Darwin Core standards improve efficiency and interoperability. Digitized herbarium specimens are now supporting taxonomy, climate studies, conservation work, and machine-learning applications.

CONCLUSION

Herbarium digitalization is essential for modern biodiversity science. It increases access to specimens, strengthens data sharing, and opens new avenues for research and conservation. Although challenges remain in transcription quality, workflow sustainability, and long-term curation, the benefits are substantial. Future progress will depend on continued collaboration among botanists, curators, data managers, and computer scientists. Digital herbaria represent a durable bridge between classical botany and contemporary biodiversity informatics.

REFERENCES

Biodiversity Information Standards (TDWG). (2010). *Guidelines for the capture and management of digital images of biological specimens*. Available at: https://www.tdwg.org/standards/

Ellwood, E. R., et al. (2015). Digitization workflows for flat sheets and packets of plants, algae, and fungi. *PhytoKeys*.

Figg P. (2024). *How herbarium specimens hold the key to combating climate change*. Royal Botanic Gardens, Kew.

GBIF (Global Biodiversity Information Facility). (2021). Darwin Core Standards for Biodiversity Data Sharing. Available at: https://www.gbif.org

Global Biodiversity Information Facility (GBIF). (2017). *GBIF best practices for data digitization*. Copenhagen: GBIF Secretariat. https://www.gbif.org

Global Biodiversity Information Facility (n.d.). *Data standards*. Retrieved April 29, 2026, from https://www.gbif.org/standards

Graves E. (2017). *Revealing the treasures of Kew's herbarium to a digital world*. Royal Botanic Gardens, Kew.

Haston, E., Cubey, R., Pullan, M., Atkins, H., & Harris, D. J. (2012). Developing integrated workflows for the digitisation of herbarium specimens using a modular and scalable approach. *ZooKeys*, 209, 93–102.

Heberling, J. M., Prather, L. A., & Tonsor, S. J. (2019). The changing uses of herbarium data in an era of global change. *American Journal of Botany*, 106(9), 1–9.

Hussein, B. R., Malik, O. A., Ong, W.-H., & Slik, J. W. F. (2022). Applications of computer vision and machine learning techniques for digitized herbarium specimens: A systematic literature review. *Ecological Informatics*, 69, 101641.

iDigBio. (2016). *Digitization workflows for biodiversity collections*. Gainesville, FL: University of Florida. Available at: https://www.idigbio.org/wiki/index.php/Digitization_Workflows

Kew Gardens: Royal Botanic Gardens, Kew. Digitising Kew's Collections.

Nelson, G., Paul, D., Riccardi, G., & Mast, A. R. (2012). Five task clusters that enable efficient and effective digitization of biological collections. *ZooKeys*, 209, 19–45.

Page, L. M., et al. (2015). Digitization of biodiversity collections reveals biggest data on biodiversity. *BioScience*, 65(9), 841–842.

Smithsonian Institution. (2012). *Digitization Program Office: Best practices for digitization*. Washington, DC. https://dpo.si.edu

Soltis, P. S. (2017). Digitization of herbaria enables novel research. *American Journal of Botany*, 104(9), 1281–1284.

Willis, C. G., et al. (2017). Old plants, new tricks: Phenological research using herbarium specimens. *Trends in Ecology & Evolution*, 32(7), 531–546.

Younis, S., Schmidt, M., & Weiland, C. (2018). Taxon and trait recognition from digitized herbarium specimens using deep convolutional neural networks. *Botany Letters*, 165(3–4), 377–383.

ANTIBACTERIAL ACTIVITY OF SOME MEDICINAL PLANT EXTRACTS AGAINST FOOD POISONING BACTERIA

Kalpna D. Rakholiya[1]*, Jinal H. Akbari[1], Mital J. Kaneria[2]
All authors have equal contribution
[1]*Department of Microbiology, Harivandana College, Rajkot-360005, Gujarat, India*
[2]*Department of Biosciences (UGC-CAS), Saurashtra University, Rajkot – 360005, Gujarat*
kalpna.rakholiya@gmail

ABSTRACT: The discovery and development of antibiotics are among the most powerful and successful achievements of modern science and technology for the control of infectious diseases. However, the increasing antimicrobial resistance emergence and dissemination among bacteria reduced the efficacy of treatment success of many drugs. The ethanol extracts from medicinal leaves were assessed for *in vitro* antibacterial activity against six bacteria (three Gram positive bacteria, three Gram negative bacteria). From the result highest antibacterial activity was observed against Gram negative bacteria. The further research is ongoing.

1. INTRODUCTION

Food-borne diseases remain a major public health concern worldwide, particularly in developing countries, due to the consumption of contaminated food and water (WHO, 2015). Pathogenic bacteria such as *Staphylococcus aureus*, *Salmonella typhimurium*, *Escherichia coli*, *Klebsiella pneumoniae*, *Bacillus subtilis*, and *Corynebacterium rubrum* are commonly associated with food poisoning and gastrointestinal infections (Ray and Bhunia, 2014). The increasing incidence of antibiotic resistance among these food-borne pathogens has necessitated the search for alternative, safe, and effective antimicrobial agents (Ventola, 2015).

Medicinal plants have been traditionally used as therapeutic agents owing to their rich content of bioactive phytochemicals such as tannins, flavonoids, alkaloids, phenolics, and terpenoids, which exhibit significant antimicrobial properties (Cowan, 1999). Among them, species of the genus *Terminalia* (family: Combretaceae) are well known in traditional medicine systems for their antimicrobial, antioxidant, anti-inflammatory, and gastroprotective activities (Kokate et al., 2014).

Terminalia arjuna, *Terminalia catappa*, and *Terminalia bellirica* have been reported to possess diverse pharmacological properties, including antibacterial and antioxidant activities (Anand et al., 2011; Parekh and Chanda, 2007). However, systematic evaluation of their antibacterial potential against food-poisoning bacteria using standardized bacterial strains remains limited. Therefore, the present study was undertaken to evaluate the antibacterial activity of ethanol extracts of leaves of selected *Terminalia* species against major food-borne bacterial pathogens using the agar well diffusion method.

2. MATERIALS AND METHODS

2.1 Collection and Preparation of Plant Material

Fresh and healthy leaves of *Terminalia arjuna, Terminalia catappa*, and *Terminalia bellirica* were collected from the local area. The plant materials were washed thoroughly with running tap water followed by distilled water to remove dust and debris. The leaves were shade-dried at room temperature and ground into a fine powder using a sterile grinder. The powdered samples were stored in airtight containers until further use.

2.2 Preparation of Ethanolic Extracts

Ethanolic extracts of the plant leaves were prepared using the cold percolation method as described by Parekh and Chanda (2007). The dried extracts were stored at 4 °C until antibacterial analysis.

2.3 Test Microorganisms

The antibacterial activity of the plant extracts was evaluated against the following standard food-poisoning bacterial strains: SA: *Staphylococcus aureus* ATCC 25923, BC: *Bacillus subtilis* ATCC 6633, CR: *Corynebacterium rubrum* ATCC 14898, ST: *Salmonella typhimurium* ATCC 23564, KP:

Klebsiella pneumoniae NCIM 2719, EC: *Escherichia coli* ATCC 25922. All bacterial cultures were obtained from a recognized culture collection and maintained on nutrient agar slants at 4 °C.

2.3.1 Preparation of Inoculum
A loopful of each bacterial culture was inoculated into sterile nutrient broth and incubated at 37 °C for 18–24 hours. The turbidity of the cultures was adjusted to match 0.5 McFarland standard to obtain a uniform bacterial inoculum.

2.3.2 Antibacterial Assay by Agar Well Diffusion Method
The antibacterial activity of the ethanolic leaf extracts was assessed using the agar well diffusion method as described by Perez et al. (1990) and Parekh and Chanda (2007).

2.4 Statistical Analysis
All experiments were performed in triplicate, and the results were expressed as mean ± standard deviation.

3. RESULT AND DISCUSSION

The antibacterial activity of ethanolic leaf extracts of *Terminalia arjuna*, *Terminalia catappa*, and *Terminalia bellirica* was evaluated against selected food-borne bacterial pathogens using the agar well diffusion method. The results were expressed as mean zone of inhibition (mm), as shown in Table 1.

Among the tested plant extracts, all three *Terminalia* species exhibited notable antibacterial activity against both Gram-positive and Gram-negative bacteria. The zones of inhibition ranged from 17 mm to 25mm, indicating moderate to strong antibacterial efficacy. The highest antibacterial activity was observed against *Escherichia coli*, where *Terminalia bellirica* showed a maximum zone of inhibition (25 mm), followed by *T. arjuna* (20 mm) and *T. catappa* (19 mm). Against *Staphylococcus aureus*, inhibition zones ranged between 18 mm and 19 mm, indicating effective activity against Gram-positive cocci.

Moderate antibacterial activity was observed against *Salmonella typhimurium* and *Klebsiella pneumoniae*, with inhibition zones ranging from 18 ± 0.6 mm to 20 ± 0.8 mm. In general, *Terminalia bellirica* exhibited comparatively higher antibacterial activity than the other two species across most tested organisms, as clearly depicted in the bar graph.

The present study confirms the antibacterial efficacy of ethanolic leaf extracts of *Terminalia* species against selected food-borne bacteria. The observed zones of inhibition indicate that the plant extracts contain bioactive compounds capable of suppressing bacterial growth. The higher inhibition observed against *Escherichia coli* (25 ± 0.9 mm) suggests that Gram-negative bacteria can be susceptible to plant-derived phytochemicals when extracted using polar solvents such as ethanol. This may be due to the presence of phenolic compounds and tannins that disrupt membrane permeability and interfere with essential cellular functions (Cowan, 1999).

Moderate antibacterial activity observed against other tested organisms indicates differential sensitivity among bacterial species. Such variation may be attributed to differences in cell wall structure, metabolic pathways, and resistance mechanisms. Gram-positive bacteria generally show consistent susceptibility due to the absence of an outer membrane, allowing easier penetration of antimicrobial compounds (Parekh and Chanda, 2007).

The overall antibacterial activity observed in this study supports earlier reports on the antimicrobial potential of *Terminalia* species and validates their traditional use in treating infectious diseases (Anand et al., 2011). The findings suggest that ethanolic leaf extracts of *Terminalia* species can serve as promising sources of natural antibacterial agents and may be further explored for applications in food safety and pharmaceutical formulations.

Table 1: Antibacterial activity of terminalia sp. against food poisoning bacteria

Plant Extract	***SA***	***BC***	***CR***	***ST***	***KP***	***EC***
T. arjuna	14 ±0.6	13 ±0.5	12 ±0.4	18 ±0.7	17 ±0.6	19 ±0.8
T. catappa	16 ±0.7	15 ±0.6	14 ±0.5	20 ±0.8	19 ±0.7	**25 ±0.9**
T. bellirica	15 ±0.6	14 ±0.5	13 ±0.4	19 ±0.7	18 ±0.6	20 ±0.8
Ethanol control	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Ciprofloxacin (10 µg)	25	24	19	28	21	22

4. CONCLUSION AND FUTURE PERSPECTIVE

The study concludes that ethanolic leaf extracts of *Terminalia* species exhibit effective antibacterial activity against major food-borne pathogens. The observed zones of inhibition indicate their potential as natural antibacterial agents, with notable activity against both Gram-positive and Gram-negative bacteria. These findings support the traditional use of *Terminalia* plants in treating microbial infections. Future studies should focus on identifying active phytochemicals, determining MIC and MBC values, and evaluating toxicity and in vivo efficacy. The development of plant-based antimicrobial formulations may offer safe alternatives to synthetic antibiotics for controlling food-borne diseases.

5. ACKNOWLEDGEMENT

The authors sincerely thank the management of Harivandana College for providing excellent research facilities and constant support throughout this study.

6. REFERENCES

1. Anand, A. V., Divya, N., & Kotti, P. P. (2011). An updated review of *Terminalia arjuna*: a miracle herb. *Asian Journal of Pharmaceutical Sciences*, 6(1), 1–14.
2. Atlas, R. M. (2010). Handbook of Microbiological Media. 4th Edition. CRC Press, Boca Raton, USA.WHO (2015). *WHO estimates of the global burden of foodborne diseases*. World Health Organization, Geneva.
3. Cappuccino, J. G., & Sherman, N. (2014). Microbiology: A Laboratory Manual. 10th Edition. Pearson Education, New York, USA.
4. Clinical and Laboratory Standards Institute (CLSI). (2018). Performance Standards for Antimicrobial Susceptibility Testing. CLSI Supplement M100, 28th ed. Wayne, PA, USA.
5. Cowan, M. M. (1999). Plant products as antimicrobial agents. *Clinical Microbiology Reviews*, 12(4), 564–582.
6. Kokate, C. K., Purohit, A. P., & Gokhale, S. B. (2014). *Pharmacognosy*. 50th Edition. Nirali Prakashan, Pune, India.
7. Parekh, J., & Chanda, S. (2007). In vitro antimicrobial activity and phytochemical analysis of some Indian medicinal plants. *Turkish Journal of Biology*, 31, 53–58.
8. Perez, C., Pauli, M., & Bazerque, P. (1990). An antibiotic assay by the agar well diffusion method. Acta Biologiae et Medicinae Experimentalis, 15, 113–115.
9. Ray, B., & Bhunia, A. (2014). *Fundamental Food Microbiology*. 5th Edition. CRC Press, Boca Raton, USA.
10. Ventola, C. L. (2015). The antibiotic resistance crisis: Part 1—causes and threats. *Pharmacy and Therapeutics*, 40(4), 277–283.

ROLE OF CRISPR TECHNOLOGY AND ITS IMPACT ON HUMAN HEALTH

Kalpna D. Rakholiya*, Hetvikumari Chaudhari, Vishva Patel, Mital J. Kaneria[2]
All authors have equal contribution
[1]Department of Microbiology, Harivandana College, Rajkot-360005, Gujarat, India
[2]Department of Biosciences (UGC-CAS), Saurashtra University, Rajkot – 360005, Gujarat
kalpna.rakholiya@gmail

ABSTRACT: CRISPR technology has emerged as a revolutionary genome-editing tool with profound implications for human health. It enables precise, efficient, and cost-effective modification of DNA, allowing targeted correction of disease-causing genetic mutations. CRISPR-based approaches are being widely explored in the diagnosis and treatment of genetic disorders, cancer, infectious diseases, and inherited metabolic conditions. Beyond therapeutics, CRISPR is also used in developing rapid diagnostic platforms and advancing personalized medicine. Despite its transformative potential, ethical concerns, off-target effects, and regulatory challenges must be addressed to ensure safe clinical applications. Overall, CRISPR technology represents a significant advancement in modern healthcare.

1. INTRODUCTION

CRISPR–Cas technology has rapidly evolved into one of the most powerful tools in modern biomedical science, fundamentally transforming approaches to disease diagnosis, treatment, and prevention. Its simplicity, precision, and versatility have accelerated translational research and opened new avenues in human health applications.

2. CRISPR in Genetic Disease Therapy

One of the most impactful applications of CRISPR in human health is the correction of monogenic disorders. Diseases caused by single-gene mutations are ideal targets for CRISPR-based interventions. Hemoglobinopathies: CRISPR-Cas9 has been successfully used to treat sickle cell anemia and β-thalassemia by editing the BCL11A gene to restore fetal hemoglobin expression. These therapies have progressed to advanced clinical trials with long-term therapeutic benefits. Inherited Muscular and Metabolic Disorders: CRISPR shows promise in Duchenne muscular dystrophy, phenylketonuria, and hereditary tyrosinemia by repairing or disabling defective genes. Innovative Editing Strategies: Base editing and prime editing enable precise single-nucleotide corrections without inducing double-strand DNA breaks, improving safety and therapeutic efficiency (Davis and Yeddula 2024)

3. CRISPR in Cancer Diagnosis and Therapy

CRISPR has revolutionized cancer research by enabling identification of oncogenes, tumor suppressor genes, and drug-resistance pathways. Cancer Immunotherapy: CRISPR-engineered T cells and CAR-T cells exhibit enhanced tumor recognition, reduced immune rejection, and improved persistence in patients. Multiplex gene editing allows simultaneous modification of multiple immune checkpoints. Targeted Cancer Therapy: CRISPR screens help identify novel therapeutic targets and personalize treatment strategies based on tumor genetic profiles. Early Cancer Detection: CRISPR-based diagnostics detect cancer-associated mutations and circulating tumor DNA with high sensitivity (Wang et al., 2023).

4. CRISPR in Infectious Disease Management

CRISPR technology is being explored for both therapeutic and diagnostic purposes in infectious diseases (Far et al., 2025). Antiviral Therapies: CRISPR-Cas systems have been investigated to target viral genomes of HIV, hepatitis B virus (HBV), herpesviruses, and SARS-CoV-2, offering potential curative strategies. **Rapid Diagnostics:** Platforms such as SHERLOCK and DETECTR enable rapid, low-cost detection of viral and bacterial pathogens, antimicrobial resistance genes, and emerging infections. **Antimicrobial Resistance Control**: CRISPR can selectively eliminate antibiotic-

resistance genes in pathogenic bacteria, offering a novel strategy against multidrug-resistant infections (Shen and Mukhopadhyay, 2024).

5. CRISPR in Neurological and Neurodegenerative Disorders

Neurological diseases pose significant therapeutic challenges due to complex genetics and delivery barriers. Neurogenetic Disorders: CRISPR is being investigated for Huntington's disease, spinal muscular atrophy, and amyotrophic lateral sclerosis (ALS) by targeting disease-causing mutations. Alzheimer's and Parkinson's Diseases: CRISPR helps identify disease-associated genes and molecular pathways, facilitating drug discovery and biomarker development. Brain-Specific Editing: Advances in viral and non-viral delivery systems are improving CRISPR accessibility to neural tissues (Akbar et al., 2025).

6. CRISPR in Precision Medicine and Drug Discovery

CRISPR supports precision medicine by linking genetic variation to disease susceptibility and treatment response. Functional Genomics: Genome-wide CRISPR screens identify genes involved in disease pathways and drug resistance. Personalized Therapeutics: CRISPR enables tailoring therapies based on individual genetic profiles. High-Throughput Drug Screening: CRISPR-modified cell lines improve the accuracy of preclinical drug testing (Chanchal et al., 2024, Mace et al., 2010).

7. Ethical, Safety, and Regulatory Considerations

Despite its promise, CRISPR raises ethical concerns, particularly regarding germline editing and long-term safety. Off-Target Effects: Ongoing improvements in guide RNA design and editing enzymes aim to enhance precision. Germline Editing: Strict ethical and regulatory oversight is required to prevent misuse. Equitable Access: Ensuring global accessibility to CRISPR-based therapies remains a critical challenge.

8. CONCLUSION

CRISPR technology has profoundly reshaped human health research and clinical practice. Its applications span genetic disease therapy, cancer treatment, infectious disease control, regenerative medicine, and precision diagnostics. Continued innovation in editing accuracy, delivery systems, and ethical governance will be essential to fully realize CRISPR's potential as a cornerstone of future healthcare.

9. REFERENCES

1. Mace F.C.; Critchfield T.S.; Translational research in behavior analysis: Historical traditions and imperative for the future. J Exp Anal Behav 2010,93(3),293-312
2. Davis DJ, Yeddula SGR. CRISPR Advancements for Human Health. Mo Med, 121(2):170-176.
3. Wang M, Chen M, Wu X, Huang X, Yu B. CRISPR applications in cancer diagnosis and treatment. Cell Mol Biol Lett. 2023, 6;28(1):73.
4. Far, B. F., Akbari, M., Habibi, M. A., Katavand, M., & Nasseri, S. (2025). CRISPR Technology in Disease Management: An Updated Review of Clinical Translation and Therapeutic Potential. *Cell Proliferation*, *58*(11), e70099.
5. Debmitra Sen, Poulami Mukhopadhyay, Antimicrobial resistance (AMR) management using CRISPR-Cas based genome editing, Gene and Genome Editing, 7:2024.
6. Akbar A, Haider R, Agnello L, Noor B, Maqsood N, Atif F, Ali W, Ciaccio M, Tariq H. CRISPR in Neurodegenerative Diseases Treatment: An Alternative Approach to Current Therapies. Genes (Basel). 2025 Jul 22;16(8):850
7. Chanchal DK, Chaudhary JS, Kumar P, Agnihotri N, Porwal P. CRISPR-Based Therapies: Revolutionizing Drug Development and Precision Medicine. Curr Gene Ther. 2024;24(3):193-207.
8. Carey N.; Hacking the Code of Life: How gene editing will rewrite our futures Icon Books 2019

PHYTOREMEDIATION FOR SOIL AND WATER CONSERVATION: A COMPREHENSIVE STUDY

V. Ramya Sudha
ZPP High School, Rangampeta,
East Godavari Dt., Andhra Pradesh,
Email: Akasapuriona@gmail.com

ABSTRACT: Phytoremediation is an emerging multidisciplinary technology that utilizes plants and their associated microorganisms to reduce, remove, or immobilize contaminants from soil and water systems. With escalating industrialization, urban growth, and agricultural expansion, soil degradation and water pollution have become critical global challenges. This study investigates the potential of phytoremediation as an eco-friendly, cost-effective strategy for restoring and conserving soil and water quality, focusing on heavy metals, organic pollutants, and nutrient overload. Using both qualitative and quantitative research methods, this paper synthesizes existing literature, ex-amines field case studies, and presents empirical evidence from surveys conducted in three contaminated sites across southern India: industrial effluent discharge zones, agricultural runoff areas, and municipal wastewater channels. The methodology involved soil/water sampling, plant species selection based on tolerance and uptake efficiency, monitoring pollutant concentration changes, and community surveys to assess local perceptions of phytoremediation effectiveness. Data were analyzed using statistical tools and visualized through graphs to evaluate pollutant reductions and plant growth correlations."The findings reveal that certain plant species such as Brachiaria mutica, Vetiveria zizanioides, Phragmites australis, and Helianthus annuus exhibited significant contaminant uptake, reducing heavy metals (lead, cadmium) by up to 65% in soils and organic pollutants (pesticides, hydrocarbons) by up to 58% in water within 6–9 months. Survey results indicate high community acceptance but highlight knowledge gaps regarding long-term maintenance and scalability. This research underscores phytoremediation's role in sustainable environmental management and proposes integrative frameworks involving policy support, community participation, and technological innovations. The study concludes with recommendations for future research on genetic enhancement of hyperaccumulator species and large-scale implementation models.

KEYWORDS: Phytoremediation, Soil conservation, Water quality management, Heavy metal uptake, Sustainable remediation

INTRODUCTION:

Environmental degradation in soil and water systems has intensified due to anthropogenic pressures such as industrial effluents, agricultural chemical use, and urban waste discharge. Contaminants like heavy metals, persistent organic pollutants (POPs), and excessive nutrients not only disrupt ecological balance but also pose serious health risks to humans, livestock, and biodiversity. Traditional remediation techniques such as excavation, chemical treatments, and soil replacement are often expensive, energy-intensive, and disruptive to the ecosystem. In this context, phytoremediation has emerged as a promising green alternative that harnesses the natural abilitiesof plants to clean contaminated environments. "Phytoremediation encompasses various mechanisms phytoextraction, phytostabilization, phytodegradation, rhizodegradation, and phytovolatilization each targeting specific pollutant classes. For example,

hyperaccumulator plants can absorb heavy metals from soil into harvestable biomass, while rhizosphere microorganisms assist in degrading organic pollutants. The effectiveness of phytoremediation depends on plant traits, environmental conditions, and pollutant characteristics.“This research explores phytoremediation’s role in soil and water con-servation, emphasizing ecological outcomes, plant species efficacy, and integration with community-based environmental management.

OBJECTIVES

The study aims to:

1. Assess the efficacy of selected plant species in reducing soil and water contaminants, including heavy metals and organic pollutants.

2. Evaluate changes in soil and water quality over time using phytoremediation.

3. Understand community perceptions and knowledge about phytoremediation practices in target areas.

4. Compare phytoremediation outcomes across different environmental contexts (industrial, agricul-tural, municipal).

5. Provide recommendations for scalable and sustainable phytoremediation applications within re-gional conservation strategies.

LITERATURE REVIEW

3.1 Definitions and Scope of Phytoremediation

Phytoremediation is defined as the use of plants to remediate contaminated soil and water by accumulating, degrading, or stabilizing pollutants through biological, chemical, and physical pro-cesses. It is part of a broader suite of bioremediation technologies that leverage biological systems for environmental cleanup.

3.2 Mechanisms of Phytoremediation

a. Phytoextraction

Plants uptake contaminants from soil through their root systems and accumulate them in aboveground tissues. Hyperaccumulator species, characterized by their ability to tolerate and con-centrate pollutants such as cadmium (Cd) and lead (Pb), are particularly useful for this process.

b. Phytostabilization

Plants immobilize contaminants in the rhizosphere, reducing bioavailability and preventing leaching, erosion, or dispersion.

c. Phytodegradation and Rhizodegradation

Plants and their rhizosphere microbes degrade organic contaminants into less toxic compounds through enzymatic processes.

d. Phytovolatilization

Certain plants uptake volatile contaminants and release them into the atmosphere in less harmful forms.

3.3 Plant Species Used in Phytoremediation

Studies have identified numerous plant species with phytoremediation potential. Vetiveria zizanioides has been widely documented for heavy metal removal and sediment stabilization. Phragmitesaustralis is effective for nutrient and organic pollutant uptake in water systems.

Sunflower (Helianthus annuus) and Indian mustard (Brassica juncea) are recognized for phytoextraction of heavy metals.

3.4 Factors Affecting Phytoremediation Efficiency

Environmental variables such as soil pH, temperature, pollutant concentration, and bioavailability influence plant uptake and degradation rates. Microbial synergy in the rhizosphere enhances organic pollutant breakdown.

3.5 Gaps in Research

Despite extensive studies, challenges remain in large-scale application, long remediation times, genetic limitations of plants, and economic analyses of cost-benefit outcomes.

4. METHODOLOGY

4.1 Research Design

This study adopted a mixed-methods approach, combining field experiments, laboratory analy-ses, and social surveys to evaluate both environmental and socio-economic aspects of phytoremediation.

4.2 Site Selection

Three sites in southern India were chosen: Site A: Industrial effluent discharge area
Site B: Agricultural runoff zone Site C: Municipal wastewater canal

4.3 Plant Selection and Cultivation

Plant species were selected based on literature and primary screening:

1. Vetiveria zizanioides
2. Phragmites australis
3. Helianthus annuus
4. Brachiaria mutica
5. Plants were transplanted in designated plots with similar soil/water conditions and monitored over 9 months.

4.4 Sampling Protocol

Soil and water samples were collected monthly from control (non-planted) and treatment (planted) plots. Parameters measured included:

1. Heavy metals (Pb, Cd, Cr)
2. Organic pollutants (pesticides, hydrocarbons)
3. pH and electrical conductivity
4. Nutrients (N, P, K)

4.5 Survey Instrument

A structured questionnaire with Likert-scale and open-ended questions was administered to 150 local residents (50 per site) to understand awareness, perceptions, and acceptance of phytoremediation.

5. CASE STUDY:

5.1 Site A: Industrial Effluent Zone

Background Site A receives effluents from nearby tanneries and chemical plants. Historical records indicate elevated levels of Pb and Cr.

Implementation

Vetiver grass plots were established along drainage channels. Outcomes
Heavy metal concentration in soil reduced significantly over 9 months (details in Tables 1 & 2).

5.2 Site B: Agricultural Runoff Zone

Background

Site B suffers from high nutrient loads due to fertilizer runoff. Implementation
Phragmites and Brachiaria were planted in riparian strips. Outcomes
Nutrient concentrations decreased, improving water quality (Tables 3 & 4).

5.3 Site C: Municipal Wastewater Canal

Background

Municipal wastewater carries organic pollutants and suspended solids.
Implementation
Sunflowers were grown along canal banks. Outcomes
Organic pollutant levels dropped, with significant biomass accumulation.

6. SURVEY RESULTS AND ANALYSIS

Table 1: Awareness Levels of Phytoremediation (n = 150)

Awareness Level	Frequency	Percentage
High	25	16.7
Moderate	60	40%
Low	65	43.3

Table 2 : Perceptions of Effectiveness

Perception	Mean Score
Improves soil quality	4.1
Improves Water quality	3.9

7. DATA ANALYSIS

7.1 : Reduction of Lead (Pb) Concentration in Soil (mg/kg)

Month	Control Plot	Phytoremedia	Reduction
1	120	120	0
2	118	105	11.0
3	117	95	18.8
4	119	85	28.6
5	121	75	38.0
6	120	65	45.8
7	122	55	54.9
8	121	48	60.3
9	123	42	65.9

ANALYSIS STATEMENT

A Continuous decline in lead concentration was observed in phytoremediated soil, achieving nearly 66% reduction within nine months, while the control plot showed no significant change.

7.2 : Reduction of Nitrogen Concentration in Water (mg/L)

Month	Control Water	Treated Water	Reduction(%)
1	30	30	0
2	31	27	12.9
3	32	25	21.9
4	33	22	33.3
5	34	20	41.2
6	35	18	48.6
7	36	16	55.6
8	36	15	58.3
9	37	14	62.2

ANALYSIS STATEMENT

Nitrogen levels in phytoremediated water decreased substantially due to nutrient uptake by aquatic plants, whereas nutrient accumulation was evident in control water bodies.

7.3 : Cadmium (Cd) Reduction in Soil (mg/kg)

Month	Control Plot	Phytoremediated Plot	Reduction(%)
1	4.5	4.5	0
2	4.6	4.0	13.0
3	4.7	3.6	23.4
4	4.6	3.1	32.6
5	4.8	2.7	43.8
6	4.9	2.3	53.1
7	5.0	1.9	62.0
8	5.1	1.6	68.6
9	5.2	1.4	73.1

ANALYSIS STATEMENT

Cadmium Concentration showed a maximum reduction of 73% conforming the efficiency of selected plant species for heavy metal phytoextraction.

7.4 : Community Perception Survey Results (n=150)

Perameter	Mean score(1-5)	Interpretation
Awareness of Phytoremediation	3.6	Moderate
Improvement in Soil quality	4.1	High
Improvement in Water quality	3.9	High
Environmental Safety	4.2	Very High
Willingness to adopt	3.8	positive

7.5 : Site-Wise Community perception scores

Study area	Mean score(1-5)
Industrial Area	3.6
Agricultural Area	4.2
Muncipal Area	4.0

ANALYSIS STATEMENT

Community acceptance was highest in agricultural areas due to visible improvements in crop productivity and water quality.

8. DISCUSSION

8.1 Environmental Impact

The study confirms that phytoremediation can significantly reduce contaminants in both soil and water. Heavy metals decreased more rapidly in vegetated plots, corroborating with earlier research that root uptake and rhizosphere interactions enhance pollutant removal.

8.2 Species Performance Comparison

Vetiver and Phragmites showed superior performance for metal and nutrient reduction, respectively. Sunflower performed moderately for organic pollutants, indicating species-specific capabilities.

8.3 Socio-Economic Considerations

Survey findings suggest limited initial awareness but strong positive perception after explanation. This highlights the importance of community engagement and education for successful implementation.

8.4 Limitations

Challenges include long remediation timelines, variable pollutant bioavailability, and the need for peri-odic plant biomass management.

9. CONCLUSION

Phytoremediation proves to be a viable, eco-friendly approach to soil and water conservation. With proper species selection, site planning, and community involvement, it can substantially reduce pollutants. While not a immediate solution, its long-term benefits and low costs make it a sustainable choice in environ-mental restoration portfolios.

10. FUTURE PLAN

1. Genetic enhancement programs to develop superior hyperaccumulator plants.“Pilot projects at larger scales, integrating phytoremediation with constructed wetlands.
2. Policy frameworks to incentivize adoption by local governments and industries.
3. Capacity building for rural communities to apply and maintain phytoremediation systems.
4. Longitudinal monitoring frameworks to evaluate ecological effects over decades.

REFERENCES

1. Ali, H., Khan, E., & Sajad, M.A. (2013). Phytoremediation of heavy metals — Concepts and applica-tions. Chemosphere, 91(7), 869–881.
2. Pilon-Smits, E.A.H. (2005). Phytoremediation. Annual Review of Plant Biology, 56, 15–39.
3. Salt, D.E., Smith, R.D., & Raskin, I. (1998). Phytoremediation. Annual Review of Plant Physiology and Plant Molecular Biology, 49, 643–668.
4. Yadav, S., & Kumar, A. (2018). Use of vetiver grass (Vetiveria zizanioides) in soil and water remediation. Environmental Science and Pollution Research, 25(25), 25017–25029.
5. Zhang, F., & Wang, Q. (2019). Advances in constructed wetlands coupled with phytoremediation. Journal of Environmental Management, 237, 179–187.

PIGEON DROPPINGS AS A RESERVOIR FOR HUMAN PATHOGENS AND DISEASE TRANSMISSION

Kalpna Rakholiya[1]*, Priya Vasoya[1], Priyanka Chauhan Avani Vaishnav, Priti Vishwakarma, Masum Jodhani, Kinjal Bambhha, Priya Vank, Mital Kaneria[2]
All authors have equal contribution
[1]*Department of Microbiology, Harivandana College, Rajkot-360005, Gujarat, India*
[2]*Department of Biosciences (UGC-CAS), Saurashtra University, Rajkot – 360005, Gujarat*
kalpna.rakholiya@gmail

ABSTRACT: Pigeon droppings are increasingly recognized as an important environmental reservoir of human pathogens, posing significant public health concerns, particularly in urban and peri-urban settings. This comprehensive review examines the microbial diversity associated with pigeon excreta and its role in the transmission of infectious diseases to humans. Pigeon droppings provide a nutrient-rich medium that supports the survival and proliferation of a wide range of pathogenic microorganisms, including bacteria (*Chlamydia psittaci, Salmonella spp., Escherichia coli*), fungi (*Cryptococcus neoformans, Histoplasma capsulatum*), viruses, and parasitic agents. Human exposure occurs primarily through inhalation of aerosolized dried droppings, direct contact, or contamination of water and food sources. Immunocompromised individuals, children, and the elderly are particularly susceptible to severe infections. This review also highlights occupational risks faced by sanitation workers, construction workers, and individuals involved in bird handling. Environmental persistence of pathogens in pigeon droppings facilitates long-term transmission potential, especially in densely populated areas. Preventive strategies such as effective waste management, public awareness, structural deterrents to pigeon nesting, and routine surveillance are essential to mitigate health risks. Understanding the epidemiological significance of pigeon droppings as a pathogen reservoir is crucial for developing integrated public health interventions and reducing zoonotic disease burden associated with urban wildlife.

1. INTRODUCTION

Feral pigeons are ubiquitous in cities globally due to their adaptability to urban ecosystems and frequent proximity to humans. These birds defecate persistently, leading to accumulations of droppings on roofs, monuments, streets, and other public areas. Pigeon feces are rich in nitrogenous compounds and minerals, creating environments conducive to microbial survival and growth. Numerous studies indicate that pigeon droppings harbor microorganisms—including bacteria and fungi—that are capable of causing disease in humans. Their droppings thus function as ecological reservoirs for potential pathogens that may infect susceptible individuals through direct contact or environmental contamination (Esakkiammal & Rajakumari, 2025).

The aim of this review is to provide a comprehensive synthesis of current scientific evidence on the role of pigeon droppings in harboring and transmitting human pathogens, detailing the microbes identified, routes of exposure, clinical implications, and public health considerations.

2. MICROBIAL DIVERSITY IN PIGEON DROPPINGS

Pigeon droppings have been identified as a potential vector for the transmission of multiple human diseases. These include fungal infections such as cryptococcosis and histoplasmosis, bacterial infections like psittacosis and salmonellosis, and even viral infections (Table 1).

Exposure to pigeon droppings has been associated with several diseases that can affect humans. The clinical manifestations of these diseases vary depending on the specific pathogen involved. Below is an overview of the primary diseases linked to pigeon droppings and their associated clinical signs in humans:

Pigeon droppings serve as a significant reservoir for various pathogenic microorganisms, posing potential public health risks. The accumulation of these excreta in urban environments creates favorable conditions for the growth and dissemination of bacteria, fungi, and viruses that can cause zoonotic infections in humans. Among the most commonly associated pathogens are Cryptococcus neoformans, Histoplasma capsulatum, Chlamydia psittaci, and Escherichia coli, which have been

implicated in respiratory illnesses, systemic infections, and gastrointestinal diseases (Haag-Wackernagel & Moch, 2004; Soltani et al., 2013).

The presence of *Cryptococcus neoformans* in pigeon droppings is particularly concerning due to its association with cryptococcosis, a severe fungal infection that predominantly affects immunocompromised individuals (Casadevall & Perfect, 1998). Additionally, Histoplasma capsulatum, the causative agent of histoplasmosis, thrives in nitrogen-rich bird feces and can become airborne, leading to respiratory infections upon inhalation (Gugnani, 2003). Bacterial pathogens such as Chlamydia psittaci, responsible for psittacosis, have also been detected in pigeon excreta, with transmission occurring through inhalation of contaminated aerosols (Smith et al., 2005).

Beyond microbial hazards, pigeon droppings contribute to environmental contamination and infrastructure deterioration, exacerbating public health concerns (Haag-Wackernagel & Spiewak, 2004). Understanding the microbiological composition and epidemiological impact of these droppings is crucial for developing effective control measures to mitigate associated health risks. This review aims to provide a comprehensive analysis of pigeon droppings as a reservoir for human pathogens, exploring their role in disease transmission, associated risk factors, and potential prevention strategies.

Table 1: Review on Microorganisms present in pigeon dropping and associated disease

No.	Microorganism	Disease	Microorganism	References
1	*Cryptococcus neoformans*	Cryptococcosis	Fungi	Adengiro, 2021
2	*Histoplasma Capsulatum*	Histoplasmosis	Fungi	Lenhart et al., 2004
3	*Candida albicans*	Candidiasis	Fungi	Pinto et al., 2019
4	*Aspergillus Fumigatus*	Aspergillosis	Fungi	Rizwan et al., 2023
5	*Salmonella spp.*	Salmonellosis	Bacteria	Rocha-e-sillva et al., 2014
6	Diarrheagenic *E. coli*	Diarrhea	Bacteria	Vania L Silva et al., 2014
7	*Campylobacter Jejuni*	Diarrhea	Bacteria	Moises Cabllero et al., 2015
8	*Chlamydophilla psittacosis*	Flu like	Bacteria	Magnino et al., 2009
9	Ectoparsasite	Allergic	Parasite	Wackerngel, 2007
10	Pegion Paramyxovirus serotype-1	Newcastle Disease	Virus	Giza et al., 2023

3. MECHANISMS OF HUMAN EXPOSURE AND TRANSMISSION

3.1 Inhalation of Aerosolized Particles

The airborne route is perhaps the most significant mode of pathogen transmission from pigeon droppings to humans. As droppings dry, they become friable and can fragment into fine dust particles. These particles may contain infectious bacteria or fungal spores that can be inhaled, leading to respiratory or systemic infection. For example, inhalation of aerosolized C. psittaci organisms from dried fecal dust is a recognized pathway for psittacosis. Similarly, airborne spores of Cryptococcus spp. from contaminated droppings can lead to cryptococcosis in susceptible individuals.

3.2 Direct Contact and Environmental Contamination

Direct contact with pigeon droppings, particularly without protective equipment, can result in the transfer of pathogens to hands and subsequently to mucous membranes or food. Indirect transmission can occur when droppings contaminate surfaces, water sources, or soil. In urban environments, rain runoff from roofs or pavements laden with droppings can introduce microbes into water systems or recreational areas, increasing the likelihood of human exposure.

4. CLINICAL IMPLICATIONS AND PUBLIC HEALTH SIGNIFICANCE

4.1 Psittacosis and Respiratory Disease

Psittacosis caused by *C. psittaci* can range from mild flu-like illness to severe pneumonia and systemic disease. Although more commonly associated with parrots and psittacine birds, psittacosis cases linked to pigeon exposure have been documented, especially among individuals who work closely with bird populations or encounter large flocks of feral pigeons.

4.2 Fungal Infections

Cryptococcosis resulting from *Cryptococcus* spp. is a global health concern, particularly in immunocompromised populations. Inhalation of spores from pigeon guano has been identified as a key risk factor, with pulmonary and central nervous system involvement posing severe outcomes. While healthy individuals may clear low levels of exposure, patients with compromised immune systems face significant disease risk.

4.3 Gastrointestinal Illness

Bacterial pathogens such as Campylobacter spp., Salmonella spp., and E. coli, which have been detected in pigeon feces, can cause gastrointestinal symptoms in humans following ingestion. While direct evidence of transmission from feral pigeons to humans remains limited compared with traditional sources (e.g., contaminated food), the presence of these organisms in droppings underscores the potential for risk, especially in scenarios involving poor hygiene or food contamination.

5. RISK FACTORS AND VULNERABLE POPULATIONS

Although the general public's risk from pigeon droppings is often considered low, certain populations are at higher risk:

Immunocompromised individuals: Those with weakened immune systems (e.g., HIV/AIDS, cancer patients, transplant recipients) are particularly susceptible to opportunistic infections such as cryptococcosis and psittacosis.

Occupational exposure: Workers involved in building maintenance, pest control, or environments with heavy pigeon infestations face increased exposure to airborne particles and contaminated surfaces.

Children and elderly: Age groups with weaker or developing immune systems may be more vulnerable to infection from environmental pathogens.

6. PREVENTION AND CONTROL MEASURES

Environmental Management

Effective management of pigeon populations and droppings is crucial in reducing pathogen reservoirs. Strategies include:

Habitat modification: Reducing nesting and roosting sites through architectural design and deterrents.

Clean-up protocols: Regular removal of droppings using wet cleaning methods to prevent aerosolization, combined with the use of PPE for workers.

Public awareness: Educating communities on risks associated with feeding pigeons and proper cleanliness practices.

Occupational Safety

For individuals with occupational exposure, implementing safety protocols—such as respiratory protection and hygiene measures—can minimize risk, particularly when cleaning areas with heavy accumulations of pigeon feces.

7. CONCLUSIONS

Pigeon droppings in urban environments function as reservoirs for diverse microbial populations, including bacteria and fungi with zoonotic potential. While disease transmission to humans from pigeon feces is relatively infrequent in healthy populations, the risk increases for immunocompromised individuals and those with heavy exposure. Continued surveillance of microbial carriage in pigeon populations, coupled with targeted public health interventions and environmental management, is essential to mitigate potential disease transmission risks.

8. REFERENCES

1. Aden Giro (2021). Review on Cryptococcus Disease.Journal of Tropical Disease and Public Health.volume-9: page no1-6.
2. animals". University of california press on behalf of the society for the study of social
3. Casadevall, A., & Perfect, J. R. (1998). Cryptococcus neoformans. ASM Press.
4. CDC — Psittacosis and related occupational risks. NIOSH/CDC. 2024.
5. Chlamydia psittaci in Faecal Samples of Feral Pigeons in Urban Areas of Lublin city, Poland. Current Microbiology. 2022.
6. Dimitri Giunchi, Albores-Barajas.V.Y, Baldaccini.E.B, Vanni.L, &Soldatini.(2012)." pigeons: problems, Dynamics and contrl", Intech.volume-3 page no 215
7. Do pigeons carry disease? The Environmental Literacy Council. 2024.
8. Esakkiammal M, Rajakumari DV. Pigeon Droppings as a Cryptococcus Reservoir: A Review. Uttar Pradesh Journal of Zoology. 2025;46(6):169-177.
9. Gugnani, H. C. (2003). Role of Histoplasma capsulatum in causing infections in humans and animals in the tropical and subtropical regions of the world. Mycoses, 46(1-2), 33-36.
10. Haag-Wackernagel, D., & Spiewak, R. (2004). Human pigeon contact zones as infection risk areas. International Journal of Hygiene and Environmental Health, 207(4), 325-331.
11. Introduction of Chlamydia psittaci into a hospital area by feral pigeons: a cross-sectional study. PubMed. 2024.
12. Jerolmack (2008)." How pigeons become Rat: The cultural-sapital Logic of problem
13. Pasmans, F., Baert, K., Martel, A., Bousquet-Melou, A., Lanckriet, R., De Boever, S., Van Immerseel, F., Eeckhaut, V., De Backer, P., & Haesbrouck, F. (2008). Induction of the carrier state in pigeons infected with *Salmonella enterica* subspecies enterica serovar typhimurium PT99 by treatment with florfenicol: A matter of pharmacokinetics. Antimicrobial Agents and Chemotherapy, 52(3), 954–961.

14. Smith, K. A., Bradley, K. K., Stobierski, M. G., & Tengelsen, L. A. (2005). Compendium of measures to control Chlamydophila psittaci infection among humans (psittacosis) and pet birds. Journal of the American Veterinary Medical Association, 226(4), 532-539.
15. Soltani, M., Pakshir, K., & Zare Shahrabadi, M. (2013). Isolation and molecular identification of Cryptococcus neoformans and other opportunistic fungi from pigeon droppings. Mycopathologia, 175(3-4), 199-210.
16. Steven W. Lenhart, M illie P. Schafer, M itchell Singal, Rana A. Hajjeh, (2004) "Histoplasmosisb Protecting Workers at Risk "National Institute for Occupational Safety and Health volume-6 :page no-1-26
17. Weber, Walter, Pigeon associated people diseses(1979)Bird Control Seminars Proceedings. volume 21.page 156-158.
18. World Health Organisation (2017). The burden of foodborne diseases in the who European region. : World Health Organisation.
19. Zhou, Y., Huang, Y., Yang, C., Zang, X., Deng, H., Liu, J., … Xue, X. (2024). The pathways and the mechanisms by which *Cryptococcus* enters the brain. *Mycology*, *15*(3), 345–359. https://doi.org/10.1080/21501203.2023.2295409

ISOLATION AND IDENTIFICATION OF PHOSPHATE SOLUBILIZING RHIZOSPHERIC BACTERIA FROM SELECTED PLANT

Kalpna D. Rakholiya[1]*, Miral M. Chikhaliya[1], Mital J. Kaneria[2]
All authors have equal contribution
[1]*Department of Microbiology, Harivandana College, Rajkot-360005, Gujarat, India*
[2]*Department of Biosciences (UGC-CAS), Saurashtra University, Rajkot – 360005, Gujarat*
kalpna.rakholiya@gmail

ABSTRACT: Increasing crop production to feed a growing population has driven the use of mineral fertilizers to ensure nutrients availability and fertility of agricultural soils. After nitrogen, phosphorus (P) is the second most important nutrient for plant growth and productivity. However, P availability in most agricultural soils is often limited because P strongly binds to soil particles and divalent cations forming insoluble P-complexes. Therefore, there is a constant need to sustainably improve soil P availability. In the present investigation, phosphate solubilization bacteria were isolated from medicinal plants. In order to study their *in vitro* characteristics, SU 11 are able to solubilize phosphate (Pi) on Pikovskaya agar plates 24 mm (5 DAI). The results strongly suggest that the PGPR strains isolated under present investigation may be useful as a novel biofertilizers for crop production.

KEYWORDS: Rhizospheric bacteria; phosphate solubilization; Agricultural applications, PVK agar

1. INTRODUCTION

Phosphate solubilizing rhizospheric bacteria (PSRB) are vital for improving plant growth by transforming insoluble phosphate compounds in soil into forms readily available for plant uptake. Phosphorus, a key macronutrient, is essential for processes such as energy transfer, signal transduction, and biosynthesis of macromolecules (Vance et al., 2003). Despite its importance, much of the soil phosphorus exists in insoluble forms, restricting its accessibility and limiting crop productivity (Richardson & Simpson, 2011). PSRB, inhabiting the rhizosphere, promote phosphate solubilization primarily through secretion of organic acids, enzymes, and chelating agents that mobilize phosphate ions from mineral complexes (Khan et al., 2007).

Isolation and identification of PSRB from the rhizosphere of plants like pearl millet are crucial to understanding their diversity, functional mechanisms, and potential application as biofertilizers in sustainable agriculture. These bacteria not only enhance phosphorus availability but also improve soil health and plant tolerance to various stresses (Gyaneshwar et al., 2002). Characterization of PSRB involves culture-dependent isolation on selective media such as Pikovskaya's agar, followed by morphological, biochemical, and molecular analyses including 16S rRNA gene sequencing for accurate identification and phylogenetic study (Chen et al., 2006).

This study aims to isolate and identify PSRB from the rhizospheric soil of pearl millet to evaluate their phosphate solubilizing efficiency and explore their potential role in enhancing soil fertility and crop yield. Understanding the diversity and functional traits of PSRB associated with pearl millet can guide the development of targeted biofertilizer formulations, promoting environmentally sustainable agricultural practices (Rodríguez & Fraga, 1999).

2. MATERIAL AND METHODS

2.1 Isolation of Phosphate Solubilizing Bacteria (PSB)

Rhizospheric soil samples were collected from the root zone of healthy pearl millet plants following the procedure described by Sandhya et al. (2009). Soil tightly adhering to the roots was carefully separated and used for bacterial isolation. Serial dilutions of the soil samples were prepared in sterile saline

solution. Dilutions were plated on Pikovskaya's agar (PVK medium) and National Botanical Research Institute's phosphate growth medium (NBRIP medium), both containing insoluble phosphate sources, to selectively culture phosphate solubilizing bacteria. Plates were incubated at 28–30°C for 3–7 days. Colonies exhibiting clear halo zones indicating phosphate solubilization were selected for further analysis.

2.2 Soil Analysis

Physicochemical properties of the rhizospheric soil were analyzed to assess parameters influencing microbial activity and phosphate availability. Standard soil analysis methods were employed to determine pH, electrical conductivity, organic carbon content.

2.3 Biochemical Characterization

Selected PSB isolates were subjected to biochemical tests as per the protocols outlined by Cappuccino and Sherman (1992). Tests included Gram staining, catalase activity, oxidase test, citrate utilization, indole production, and other standard assays to characterize the metabolic and enzymatic profiles of the isolates.

3. RESULT AND DISCUSSION

Assessment of physical properties of soil sample

Soil samples A, B, and C were brown in color with a slightly alkaline pH (7.50). Low EC (1.2 dS m^{-1}) indicated non-saline soil. Moderate TSS (800 mg L^{-1}) and organic carbon (0.58%) suggested good soil health and moderate fertility suitable for cultivation. (Table 1).

Table 1: Physical properties of soil sample

Sample Id	Soil color	Soil analysis parameter			
		pH	EC (dSm^{-1})	TSS (mgL^{-1})	OC (%)
A, B, C	Brown color	7.50	1.2	800	0.58

Phosphate solubilization

Phosphate solubilization activity varied among the bacterial isolates. Isolate SU-11 showed the highest solubilization efficiency, producing a clear halo zone of 24 mm with a solubilization index of 5.0, and was graded as excellent (Table 2). Other isolates exhibited moderate to low phosphate solubilization activity, indicating variability in their ability to mobilize insoluble phosphate.

The superior performance of isolate SU-11 suggests strong phosphate solubilizing potential, possibly due to the production of organic acids that lower pH and chelate cations bound to insoluble phosphate. Such activity enhances phosphorus availability to plants and is a desirable trait for biofertilizer development. Similar observations have been reported for efficient phosphate solubilizing bacteria belonging to genera such as *Pseudomonas*, *Bacillus*, and *Rhizobium*, which significantly improve soil phosphorus availability and plant growth.

The superior performance of isolate SU-11 suggests strong phosphate solubilizing potential, possibly due to the production of organic acids that lower pH and chelate cations bound to insoluble phosphate. Such activity enhances phosphorus availability to plants and is a desirable trait for biofertilizer development. Similar observations have been reported for efficient phosphate solubilizing bacteria belonging to genera such as *Pseudomonas*, *Bacillus*, and *Rhizobium*, which significantly improve soil phosphorus availability and plant growth.

Table 2: Phosphate solubilization activity of isolates

Isolate Code	Colony Diameter (mm)	Halo Zone (mm)	Solubilization Index (SI)	Observation
SU-1	6	18	4.0	Moderate
SU-3	7	20	3.85	Good
SU-5	5	16	4.2	Moderate
SU-7	6	22	4.7	Very good
SU-11	**6**	**24**	**5.0**	**Excellent**
SU-14	7	19	3.7	Good
SU-18	5	15	4.0	Moderate
SU-21	6	17	3.8	Moderate
SU-23	7	21	4.0	Good
SU-25	6	18	4.0	Moderate
SU-27	5	16	4.2	Moderate

3.3 Morphological and Biochemical Characterization of Isolate SU-11

The biochemical tests performed showed characteristic reactions consistent with *Pseudomonas* species. The isolate was oxidase positive and catalase positive, indicating the presence of cytochrome oxidase and the ability to decompose hydrogen peroxide, respectively. The organism was negative for indole production, methyl red, Voges–Proskauer, and urease tests.

The isolate showed oxidative utilization of glucose in the oxidation–fermentation (O/F) test, confirming its strictly aerobic and oxidative metabolism. Nitrate reduction test was positive, indicating the ability to reduce nitrate to nitrite. The isolate did not ferment lactose or sucrose and showed no gas production. Based on the collective morphological features and biochemical reactions, the isolate was identified as *Pseudomonas* sp. (Table 3).

Table 3: Morphological and Biochemical Characterization of Isolate SU-11

Character	Result
Gram reaction	Gram-negative
Cell shape	Rod-shaped
Motility	Positive
Catalase	Positive
Oxidase	Positive
Indole	Negative
Citrate utilization	Positive
Starch hydrolysis	Positive
Probable genus	*Pseudomonas* sp.

4. CONCLUSION AND FUTURE PERSPECTIVES

The present study highlights the effective phosphate solubilization ability of the bacterial isolates, with isolate SU-11 exhibiting excellent solubilization efficiency. These findings indicate the potential of phosphate solubilizing bacteria as eco-friendly biofertilizers to enhance phosphorus availability and improve soil fertility.

Future studies should focus on molecular identification of efficient isolates, optimization of phosphate solubilization under field conditions, and evaluation of their impact on crop growth and yield. Development of stable formulations and large-scale application may contribute to sustainable agricultural practices.

5. ACKNOWLEDGEMENT

The authors sincerely thank the management of Harivandana College for providing excellent research facilities and constant support throughout this study.

6. REFERENCES

1. Bergottini VM, Herve V, Sosa DA, Otegui MB, Zapata PD, Junier P (2017). Exploring the diversity of the root-associated microbiome of *Ilex paraguariensis* St. Hil. (Yerba Mate). Appl. Soil Ecol. 109: 23-31.
2. Cappuccino JC, Sherman N (1992). Microbiology. A Laboratory Manual, 3rd Edition. Benjamin/Cummings Publ. Co., New York.
3. Chen, Y. P., Rekha, P. D., Arun, A. B., Shen, F. T., Lai, W. A., & Young, C. C. (2006). Phosphate solubilizing bacteria from subtropical soil and their tricalcium phosphate solubilizing abilities. Applied Soil Ecology, 34(1), 33-41.
4. Dinesh R, Anandaraj M, Kumar A, Bini YK, Subila KP, Aravind R (2015). Isolation, characterization, and evaluation of multi-trait plant growth promoting rhizobacteria for their growth promoting and disease suppressing effects on ginger. Microbiol. Res. 173: 34-43.
5. Gouzou L, Burtin G, Philippy R, Bartoli F, Heulin T (1993). Effect of inoculation with *Bacillus polymyxa* on soil aggregation in the wheat Rhizosphere: preliminary examination. Geoderma 56: 479-490.
6. Gyaneshwar, P., Kumar, G. N., Parekh, L. J., & Poole, P. S. (2002). Role of soil microorganisms in improving P nutrition of plants. Plant and Soil, 245(1), 83-93.
7. Gyaneshwar, P., Kumar, G. N., Parekh, L. J., & Poole, P. S. (2002). Role of soil microorganisms in improving P nutrition of plants. *Plant and Soil*, 245, 83–93.
8. Khan, M. S., Zaidi, A., & Wani, P. A. (2007). Role of phosphate-solubilizing microorganisms in sustainable agriculture—a review. Agronomy for Sustainable Development, 27(1), 29-43.
9. Lorck H (1948). Production of hydrocyanic acid by bacteria. Plant Physiol. 1: 142-146.
10. Majeed A, Abbasi MK, Hameed S, Imran A, Rahim N (2015). Isolation and characterization of plant growth-promoting rhizobacteria from wheat rhizosphere and their effect on plant growth promotion. Front. Microbiol. doi: 10.3389/fmicb.2015.00198
11. Natarajan A, Kumar K, Madhuri K, Usharani GK (2016). Isolation and characterization of salt tolerant plant growth promoting rhizobacteria from plants grown in tsunami affected regions of Andaman and Nicobar Islands. Geomicrobiol. J. DOI: 10.1080/01490451.2015.1128994
12. Pikovskaya RI (1948). Mobilization of phosphorus in soil in connection with vital activity of some microbial species. Mikrobiologya 17: 362–370.
13. Rakholiya KD, Kaneria MJ, Singh SP, Vora VD, Sutaria GS (2017). Biochemical and proteomics analysis of the plant growth-promoting rhizobacteria in stress conditions. In: Singh RP, Kothari R, Koringa PG, Singh SP, editors. Understanding Host-Microbiome Interactions - An Omics Approach. Springer, pp. 227-246.
14. Richardson, A. E., & Simpson, R. J. (2011). Soil microorganisms mediating phosphorus availability update on microbial phosphorus. Plant Physiology, 156(3), 989-996.

15. Rodríguez, H., & Fraga, R. (1999). Phosphate solubilizing bacteria and their role in plant growth promotion. Biotechnology Advances, 17(4-5), 319-339.
16. Rodríguez, H., & Fraga, R. (1999). Phosphate solubilizing bacteria and their role in plant growth promotion. *Biotechnology Advances*, 17, 319–339.
17. Sandhya S, Ali SK Z, Grover M, Reddy G, Venkateswarlu B (2009). Alleviation of drought stress effects in sunflower seedlings by the exopolysaccharides producing *Pseudomonas putida* strain GAP-P45. Biol. Fertil. Soils 46: 17-26.
18. Schwyn B, Neilands JB (1987). Universal chemical assay for the detection and determination of siderophores. Anal. Biochem. 160: 47-56.
19. Soussi A, Ferjani R, Marasco R, Guesmi A, Cherif H, Rolli E, Mapelli F, Ouzari HI, Daffonchio D, Cherif A (2016). Plant-associated microbiomes in arid lands: diversity, ecology and biotechnological potential. Plant Soil 405: 357-370.
20. Vance, C. P., Uhde-Stone, C., & Allan, D. L. (2003). Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. New Phytologist, 157(3), 423-447.
21. Vassilev, N., Vassileva, M., & Nikolaeva, I. (2006). Simultaneous P-solubilizing and biocontrol activity of microorganisms. *Applied Microbiology and Biotechnology*, 71, 137–144.

HARNESSING THE HUMAN GUT MICROBIOTA: A NEW FRONTIER IN DISEASE MANAGEMENT AND THERAPEUTICS

Kalpna D. Rakholiya*, Ayushi V. Barad, Mital J. Kaneria
All authors have equal contribution
[1]*Department of Microbiology, Harivandana College, Rajkot-360005, Gujarat, India*
[2]*Department of Biosciences (UGC-CAS), Saurashtra University, Rajkot – 360005, Gujarat*
kalpna.rakholiya@gmail

ABSTRACT: The human gut microbiota plays a pivotal role in maintaining health and has emerged as a key regulator in the prevention and management of numerous diseases. Comprising trillions of microorganisms, the gut microbiome influences host metabolism, immune modulation, nutrient absorption, and protection against pathogenic invasion. Alterations in gut microbial composition, known as dysbiosis, have been strongly associated with metabolic disorders, inflammatory bowel diseases, cardiovascular diseases, allergies, neurological conditions, and certain cancers. Recent advances highlight the therapeutic potential of modulating gut microbiota through probiotics, prebiotics, synbiotics, dietary interventions, fecal microbiota transplantation, and microbiome-targeted drugs. These approaches have shown promise in restoring microbial balance, enhancing immune responses, reducing inflammation, and improving treatment outcomes. Personalized microbiome-based strategies are gaining attention as innovative tools in precision medicine. This review emphasizes the critical roles of the human gut microbiota in disease management and underscores its potential as a novel target for preventive and therapeutic interventions, offering new perspectives for improving human health.

1. INTRODUCTION

The gut microbiota consists of a vast consortium of bacteria, archaea, fungi, viruses, and protozoa residing primarily in the gastrointestinal tract. These microbial communities contribute substantially to human health, performing metabolic functions, promoting immune homeostasis, and influencing distant organ systems through biochemical communication networks such as the gut-brain axis. Dysbiosis—an imbalance in microbial composition and function—has been implicated in numerous diseases, prompting interest in harnessing the microbiota as a therapeutic target (Ding et al., 2025). Recent advances have expanded this field beyond observational studies to translational and therapeutic applications, including engineered probiotics and microbiota modulation strategies.

2. PHYSIOLOGICAL ROLES OF THE GUT MICROBIOTA

The gut microbiota contributes to human physiology through: Metabolic functions: Microbes ferment non-digestible carbohydrates into short-chain fatty acids (SCFAs), which are crucial energy sources for colonocytes and regulators of host metabolism. Immune modulation: Microbial metabolites interact with host immune pathways, promoting immunological tolerance and shaping systemic immune responses. Drug interactions: The gut microbiota influences drug metabolism and efficacy, contributing to inter-individual variability in treatment responses (Prakash et al., 2011).

Bidirectional organ communication: Signaling along axes such as the gut-brain and gut-liver pathways demonstrates how gut microbial changes can affect distant physiological systems. Together, these functions highlight the microbiota's systemic influence and therapeutic potential (Abavisani et al., 2025).

3. DYSBIOSIS AND DISEASE PATHOGENESIS

Disruptions in microbial diversity and composition—dysbiosis—are linked to a variety of disorders:

3.1 Gastrointestinal Diseases

Gut dysbiosis has strong associations with inflammatory bowel disease (IBD), irritable bowel syndrome (IBS), celiac disease, colorectal cancer, and recurrent Clostridioides difficile infections. Microbiota

imbalance contributes to mucosal inflammation, barrier dysfunction, and immune dysregulation (Shen et al., 2025).

3.2 Metabolic Disorders

Altered microbiota composition is observed in obesity, type 2 diabetes, and non-alcoholic fatty liver disease, likely involving mechanisms such as SCFA imbalance, increased intestinal permeability, and systemic low-grade inflammation.

3.3 Extra-intestinal Conditions

Emerging evidence links gut dysbiosis with neuropsychiatric conditions (via the gut-brain axis), autoimmune diseases, and inflammatory disorders beyond the gut. Although mechanisms remain under investigation, interventional studies show gut microbiota modulation can influence outcomes in these systemic conditions.

4. MICROBIOTA-TARGETED THERAPEUTIC STRATEGIES

4.1 Probiotics, Prebiotics, and Synbiotics

Probiotics (live beneficial microbes), prebiotics (substrates that promote beneficial microbes), and synbiotics (combination products) aim to restore microbial balance. Clinical studies report improved outcomes for some gastrointestinal symptoms and metabolic biomarkers, though strain specificity and individual responses vary (Yassine et al., 2025).

4.2 Fecal Microbiota Transplantation (FMT)

FMT involves transferring stool from healthy donors to patients to reconstitute gut microbial communities. It is highly effective for *recurrent C. difficile* infection and has shown promise in IBD, metabolic syndrome, and neurological conditions, although evidence quality varies by indication and methodology. Clinical trials also suggest a role for FMT in alleviating immunosuppressant-associated diarrhea and recurrent infections in transplant recipients, expanding its potential clinical utility (Panah et al., 2026).

4.3 Diet and Lifestyle Interventions

Dietary modulation of the microbiota offers a non-invasive therapeutic approach. Personalized nutrition strategies, informed by microbiota profiles, may improve metabolic outcomes and complement pharmacological treatments (Conlon et al., 2014).

5. CHALLENGES AND FUTURE DIRECTIONS

Despite rapid advances, several challenges remain: Inter-individual variability: The highly individualized nature of microbiota composition complicates generalizable therapeutic approaches. Mechanistic gaps: While associations between microbiota and disease are strong, causal mechanisms require further elucidation. Safety and regulation: For FMT and engineered microbial therapies, donor selection, long-term outcomes, and regulatory frameworks need refinement to ensure safety and efficacy. Precision medicine integration: Combining multi-omics data with clinical phenotypes will be essential for tailoring microbiota therapies to individual patients. Future research should prioritize controlled clinical trials, molecular mechanistic studies, and standardized protocols to harness the gut microbiota fully in clinical practice.

6. CONCLUSIONS

The human gut microbiota represents a transformative frontier in disease management and therapeutics. From conventional probiotics to cutting-edge engineered microbial drugs and microbiota transplantation, these strategies demonstrate significant potential across a spectrum of diseases. Realizing this potential will require integrated multidisciplinary research, robust clinical evidence, and thoughtful translation into precision therapies that account for individual microbiota landscapes.

7. **REFERENCES**

1. Abavisani M, Sobhan Karbas Foroushan, Prashant Kesharwani, Amirhossein Sahebkar, The frontier of health: Exploring therapeutic potentials of the microbiome, PharmaNutrition, 31: 2025, 100435
2. Cao Z, Gao T, et al. Fecal microbiota transplantation-current perspective on human health. Front Med. 2025.
3. Conlon MA, Bird AR. The impact of diet and lifestyle on gut microbiota and human health. Nutrients. 2014 Dec 24;7(1):17-44.
4. Ding W, Cheng Y, Liu X, Zhu Z, Wu L, Gao J, Lei W, Li Y, Zhou X, Wu J, Gao Y, Ling Z, Jiang R. Harnessing the human gut microbiota: an emerging frontier in combatting multidrug-resistant bacteria. Front Immunol. 2025 Mar 17;16:1563450.
5. Jonkers DMAE, et al. Fecal microbiota transplantation for gastrointestinal disorders: A review. PubMed. 2017.
6. Panah, F.M., Støving, R.K., Sjögren, M. *et al.* Impact of a single fecal microbiome transplantation in adult women with anorexia nervosa: an open-label feasibility pilot trial. *Nat Commun* (2026). https://doi.org/10.1038/s41467-026-68455-8
7. Prakash S, Rodes L, Coussa-Charley M, Tomaro-Duchesneau C. Gut microbiota: next frontier in understanding human health and development of biotherapeutics. Biologics. 2011;5:71-86.
8. Shen Y, Fan N, Ma SX, Cheng X, Yang X, Wang G. Gut Microbiota Dysbiosis: Pathogenesis, Diseases, Prevention, and Therapy. MedComm (2020). 2025 Apr 18;6(5):e70168.
9. SienabioACTIVE, et al. Fecal Microbiota Transplantation as new therapeutic avenue for human diseases. J Clin Med. 2022.
10. Yang L, et al. Harnessing the human microbiome for innovative drug delivery systems. PubMed. 2023.
11. Yassine F, Najm A and Bilen M (2025) The role of probiotics, prebiotics, and synbiotics in the treatment of inflammatory bowel diseases: an overview of recent clinical trials. *Front. Syst. Biol.* 5:1561047.
12. Zhang D, Luo Y, et al. Gut microbiome dysbiosis and regulation by fecal microbiota transplantation: umbrella review. Front Microbiol. 2023.

FROM FIELD TO FOOD CHAIN: PESTICIDE MAGNIFICATION AND HUMAN RISK

Kalpna D. Rakholiya*, Sneha D. Makwana, Mital J. Kaneria[2]
All authors have equal contribution
[1]Department of Microbiology, Harivandana College, Rajkot-360005, Gujarat, India
[2]Department of Biosciences (UGC-CAS), Saurashtra University, Rajkot – 360005, Gujarat
kalpna.rakholiya@gmail

ABSTRACT: Pesticides are widely used in agriculture to improve crop yield and control pests, but their persistence in the environment poses significant risks to ecosystems and human health. Biomagnification refers to the progressive increase in concentration of certain pesticides as they move through different trophic levels of the food chain. From agricultural fields, pesticide residues enter soil and water systems, are absorbed by plants, accumulate in herbivores, and become highly concentrated in top consumers, including humans. Persistent pesticides, especially organochlorines and some organophosphates, are of particular concern due to their lipophilic nature and resistance to degradation. Long-term human exposure to biomagnified pesticides has been linked to neurological disorders, endocrine disruption, reproductive toxicity, immune system impairment, and increased cancer risk. Children and pregnant women are especially vulnerable to these effects. This presentation emphasizes the need for safer pesticide use, sustainable farming practices, effective regulatory measures, and public awareness to reduce health risks and ensure food safety.

1. INTRODUCTION

Pesticides are applied extensively to control pests and improve crop yields. However, most of the applied chemicals do not reach their target organisms; instead, they disperse into soil, water, and air, contaminating ecosystems. Persistent pesticides like DDT, dieldrin, and endosulfan are lipophilic and resistant to degradation, making them prone to bioaccumulation in fatty tissues of organisms and biomagnification in food chains. Humans, being top predators, are particularly susceptible to chronic exposure through dietary intake. (Ray & Shaju, 2023 ; Zhou et al., 2025).

2. PESTICIDE TRANSFER AND ENVIRONMENTAL BEHAVIOR

After application, pesticides undergo volatilization, runoff, leaching, and spray drift, contaminating soil, surface water, and the atmosphere. Agricultural runoff introduces pesticides into aquatic systems, where they can accumulate in sediments and aquatic organisms. Plants can absorb pesticides from soil and water, leading to residues in crops. Livestock ingest contaminated feed and water, transferring residues to meat and dairy products. These processes facilitate the entry of pesticides into human food systems. (Muñoz-Bautista et al., 2025, Shekhar et al., 2024)

3. BIOACCUMULATION AND BIOMAGNIFICATION

Bioaccumulation occurs when an organism absorbs a pesticide faster than it can eliminate it, resulting in accumulation in tissues. Lipophilic organochlorine pesticides are particularly prone to storage in fatty tissues. Biomagnification refers to the progressive increase of pesticide concentrations along the food chain, from primary producers to apex consumers. Humans, as apex consumers, often accumulate the highest concentrations, with chronic exposure leading to toxicological consequences. (Carvalho 2017, Shekhar et al., 2024).

4. HUMAN HEALTH RISKS

Humans are exposed to pesticides primarily through dietary intake of contaminated crops, meat, dairy, and fish, with additional exposure via inhalation and dermal contact in agricultural settings. Chronic low-dose exposure is associated with neurodegenerative diseases (e.g., Parkinson's), endocrine disruption, reproductive toxicity, and cancer. Acute exposure can cause headaches, nausea, respiratory distress, and neurological symptoms, particularly in occupational settings. Vulnerable populations such as children, pregnant women, and immunocompromised individuals face heightened risks. (Hua & Liu, 2024 ; Bell, 2021).

5. MONITORING AND MITIGATION STRATEGIES

Monitoring pesticide residues in food and environmental samples is essential to assess human risk. Regulatory approaches such as Maximum Residue Limits (MRLs) and Integrated Pest Management (IPM) help reduce exposure. Bioremediation, phytoremediation, and development of biodegradable pesticides represent sustainable approaches to minimize environmental persistence and entry into the food chain. (Pathak et al., 2022; Verdadero et al., 2025)

6. CONCLUSIONS

Pesticide contamination from fields to food chains poses significant public health challenges. Persistent and lipophilic pesticides undergo bioaccumulation and biomagnification, increasing human exposure risks, especially through dietary intake. Comprehensive monitoring, adherence to regulations, and adoption of sustainable pest management strategies are essential to reduce human health hazards. Future research should focus on cumulative exposure assessment, low-dose chronic effects, and innovative mitigation technologies.

7. REFERENCES

1. Ray, S., Shaju, S. T. Bioaccumulation of pesticides in fish resulting toxicities in humans through food chain and forensic aspects. Environ Anal Health Toxicol 38, e2023017 (2023).
2. Shekhar C, Khosya R, Thakur K, Mahajan D, Kumar R, Kumar S, Sharma AK. A systematic review of pesticide exposure, associated risks, and long-term human health impacts. Toxicol Rep. 2024 Nov 30;13:101840.
3. Zhou, W., Li, M., Achal, V. A comprehensive review on environmental and human health impacts of chemical pesticide usage. Environ Chall 11, 100410 (2025).
4. Shekhar, C., Khosya, R., Thakur, K. et al. A systematic review of pesticide exposure, associated risks, and long-term human health impacts. Toxicol Rep 13, 101840 (2024).
5. Muñoz-Bautista, J. M., et al. Environmental and health impacts of pesticides: toxicity, oxidative stress and genotoxic damage. Agronomy 15, 1878 (2025).
6. Carvalho, F. P. Pesticides, environment, and food safety. Food Energy Security 6, 48–60 (2017).
7. Bell, J. Impacts of pesticide exposure on human health: a systematic review. J Public Health Environ 4, 71–83 (2021).
8. Hua, Y., Liu, G. Food pesticide residues monitoring and health risk assessment. Foods 13, 474 (2024).
9. Pathak, V. M., et al. Current status of pesticide effects on environment, human health and clean-up strategies. Front Microbiol (2022).
10. Verdadero, F. X. D., Agarap, A. Z., Macatingrao, C. N. E., Ordonez, I. A., Jr., Tavu, L. E. J., Pires, D., Balendres, M. A. O. Pesticides in the environment: benefits, harms, and detection methods. Sci 7, 171 (2025).
11. Hoh, E., Hites, R. A. Atmospheric organochlorine pesticides in the western Canadian Arctic. Environ Sci Technol 38, 345–352 (2004).
12. Singh, B. K., Walker, A. Microbial degradation of organophosphorus compounds. FEMS Microbiol Rev 30, 428–471 (2006).

13. US EPA. Human health issues related to pesticides. U.S. Environmental Protection Agency. (2026).
14. Intestinal absorption and biomagnification of organic contaminants in fish, wildlife, and humans. Environ Sci Technol (2004).
15. Aguilar, A., et al. Bioaccumulation of organochlorine pesticides in marine food webs and human health implications. Mar Pollut Bull 130, 150–160 (2018).
16. Tang, J., et al. Risk of global pesticide pollution and food chain contamination. Discover Agriculture 2, 101–115 (2021).
17. Sharma, A. Worldwide pesticide usage and its impacts on ecosystem. SN Appl Sci 1, 144 (2019).

Science and Technology

Recent Trends

Volume - II

The 21st century is defined by an unprecedented convergence of scientific disciplines, driving transformative advances that continuously reshape our understanding of the natural world. From pioneering breakthroughs in biotechnology and microbiology to interdisciplinary innovations addressing global health and environmental challenges, the current scientific landscape demands both intellectual rigor and ethical responsibility a pursuit of "Science with a Soul." The 16th National Science Symposium brought together a global community of scholars, researchers, educators and students in a dynamic hybrid platform to reflect upon and contribute to this rapidly evolving narrative.

This volume encompasses a curated collection of full-length research papers categorized into vital domains, including Innovations in Life Sciences, Emerging Technologies in Physical and Chemical Sciences, Environmental Challenges and Sustainable Solutions, omputational Advances and Interdisciplinary Approaches in Modern Research. Emerging from the rigorous oral and poster presentations at the 16th National Science Symposium, organized by Christ College, Rajkot, on February 1, 2026, these papers offer critical insights and novel perspectives on pressing scientific developments.

Generously supported and sponsored by the Gujarat State Biotechnology Mission (GSBTM), the Gujarat Council on Science and Technology (GUJCOST) and the Department of Science and Technology (DST), Government of India, this event reflects the institution's steadfast commitment to fostering a vibrant research culture and igniting scientific curiosity among young minds.

This publication stands as a testament to the enduring spirit of inquiry and the collaborative pursuit of knowledge, capturing the essence of contemporary science and its profound potential to shape a sustainable and technologically advanced future.

Christ College, Rajkot
Vidya Niketan, Saurashtra University PO
Rajkot 360005 [Gujarat] India.
Phone: +919427164732
www.christcollegerajkot.edu.in
e-mail: info@christcollegerajkot.edu.in

ISBN 9788197073229