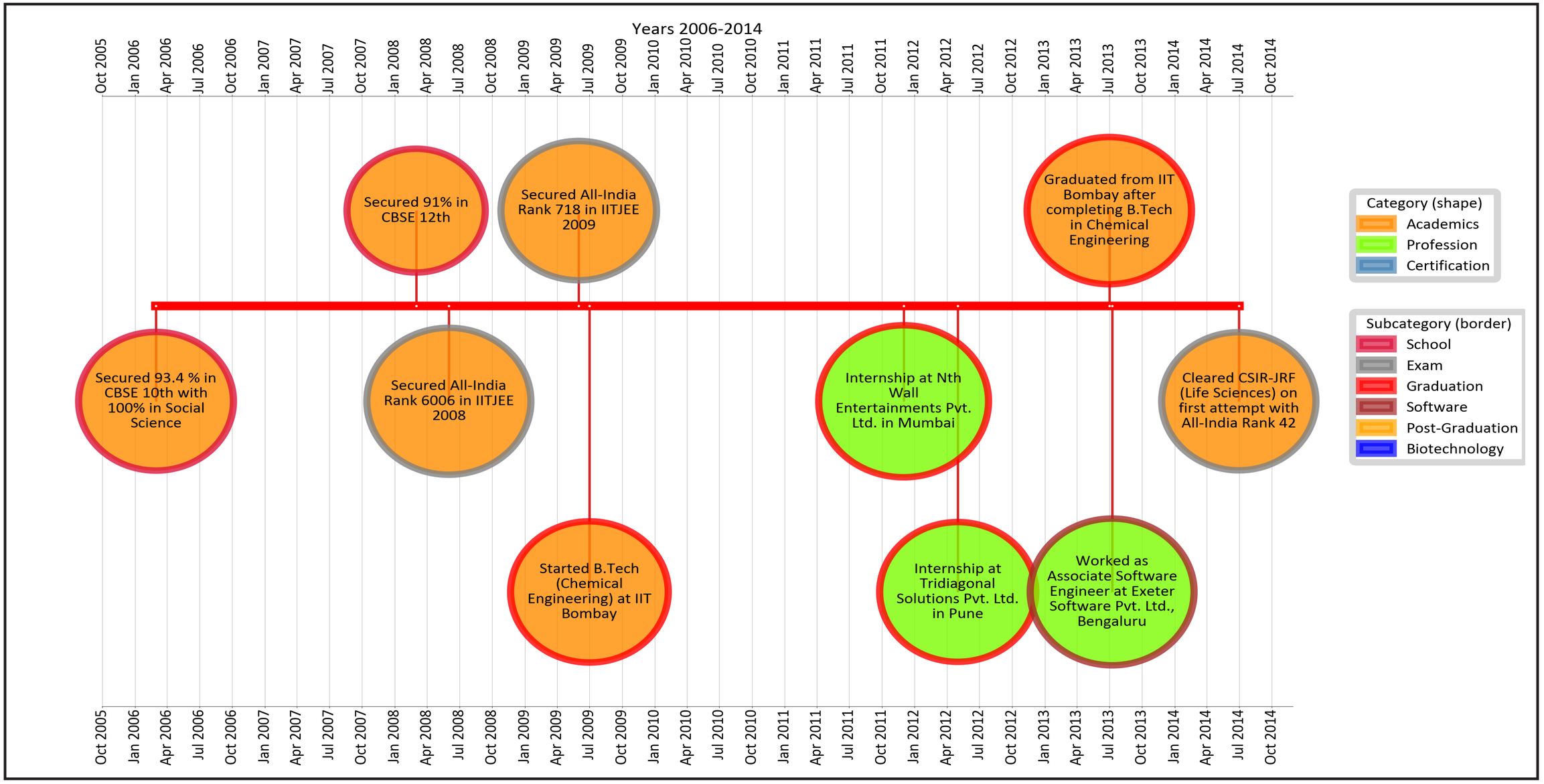
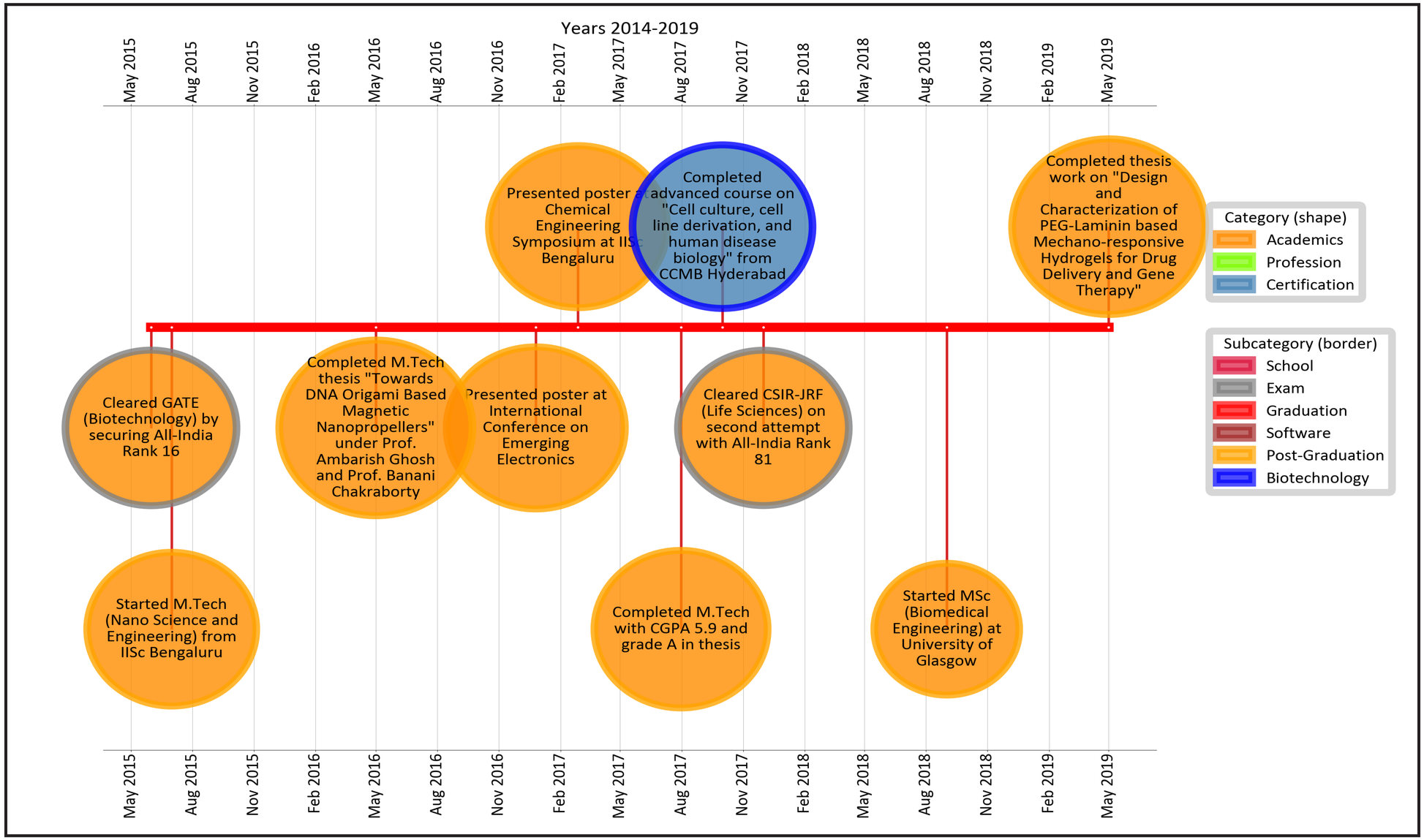
CAREER TRAJECTORY

nbahar@musingsinbiology.com



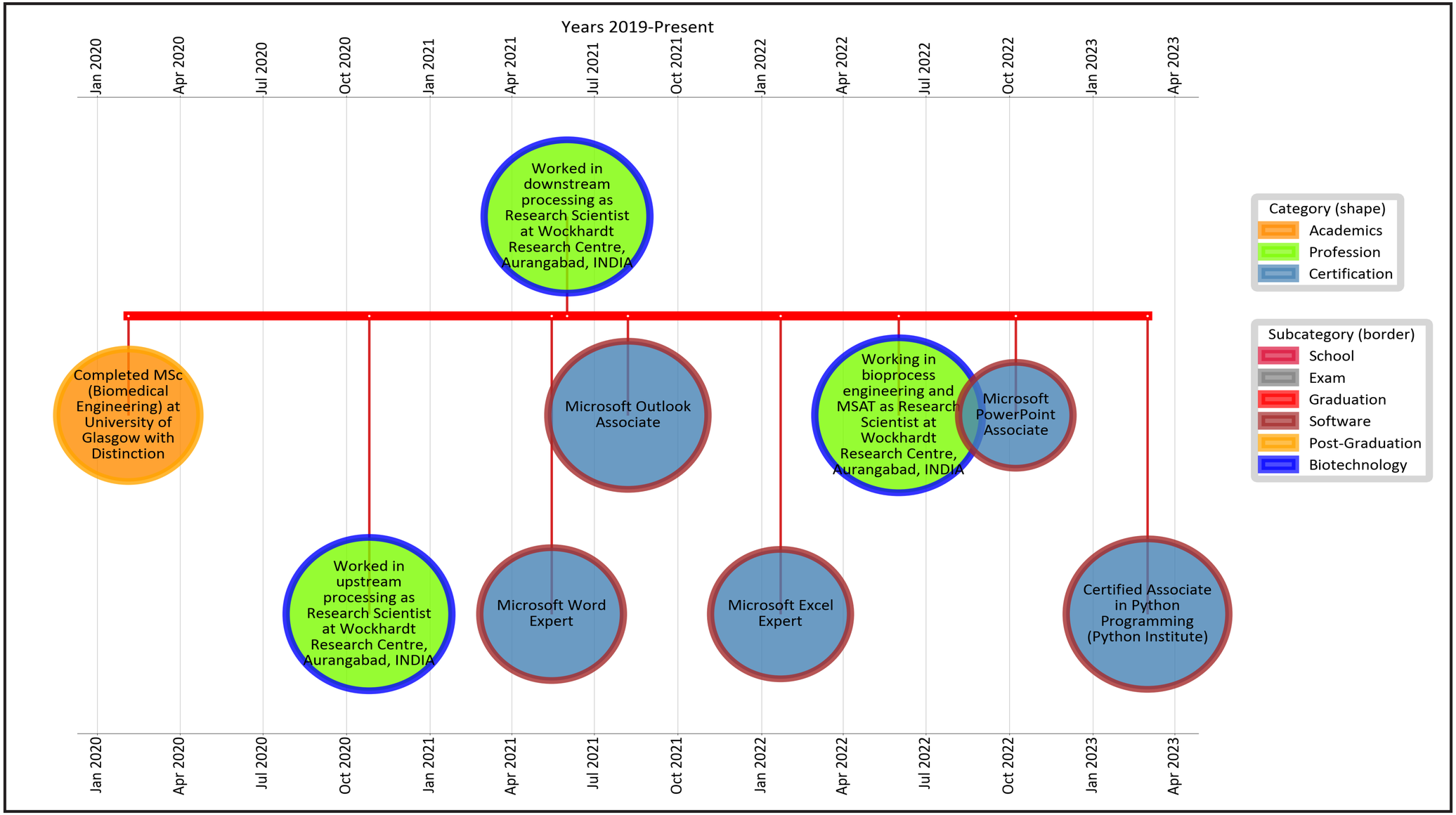
CAREER TRAJECTORY

nbahar@musingsinbiology.com



CAREER TRAJECTORY

nbahar@musingsinbiology.com



LinkedIn Learning Course Total Minutes Spent By Year and Category

Category	2020	2021	2022	2023	Grand To..
Adobe Apps	1,225	4,206			5,431
Data Analysis	270	76	180	163	689
Django				360	360
Machine Learning				248	248
MS Office Suite	843	3,829	584		5,256
Personality Development	997	265	33		1,295
Power BI				818	818
Professional Development	1,017	316	173		1,506
Python			2,089	825	2,914
Six Sigma Principles		752			752
Software	1,631		635	121	2,387
Statistics		343			343
Writing	852				852
Grand Total	6,835	9,787	3,694	2,535	22,851

Year of Date of Completion

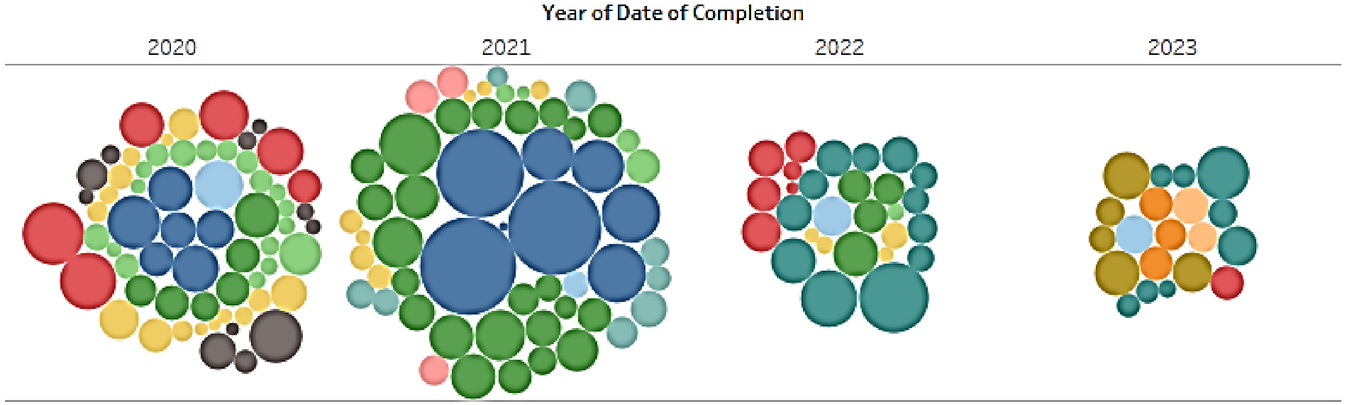
Duration in Minutes: 33 – 4,206

Sum of Duration in Minutes broken down by Date of Completion Year vs. Category. Color shows sum of Duration in Minutes. The marks are labeled by sum of Duration in Minutes. The data is filtered on Exclusions (Category,Course Name), which keeps 163 members.

I have a strong belief that learning should be a lifelong journey, and I consistently strive to apply transferable skills to my daily work. To fuel my continuous learning, I have chosen to utilize ***LinkedIn Learning*** and its wide range of courses. Over the years, this platform has significantly enhanced my knowledge and skill set, contributing to my professional development.

I have successfully completed numerous LinkedIn Learning courses, covering diverse areas such as software skills, MS Office Suite, Python, Django, Adobe applications, and personal and professional skill development. Although some courses have been retired, I have a total of 173 courses under my belt. These courses have been instrumental in expanding my expertise and enabling me to tackle various challenges in my professional endeavors.

LinkedIn Learning Course By Year and Category



Category, Course Name and Duration Hour broken down by Date of Completion Year. Color shows details about Category. Size shows sum of Duration in Minutes. The marks are labeled by Category, Course Name and Duration Hour. The view is filtered on Exclusions (Category,Course Name), which keeps 163 members.

PROJECT 1 : M.TECH THESIS

Abstract:

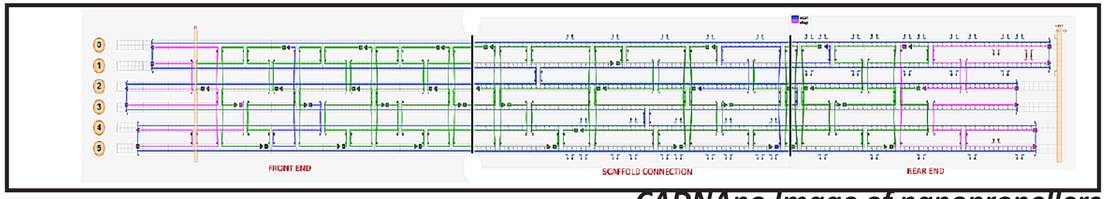
In this project, I focused on DNA origami-based magnetic nanopropellers, contributing to the design, characterization, and bioconjugation processes. The project consisted of four phases, each addressing key aspects of the research.

Phase 1: Lab Training and Orientation

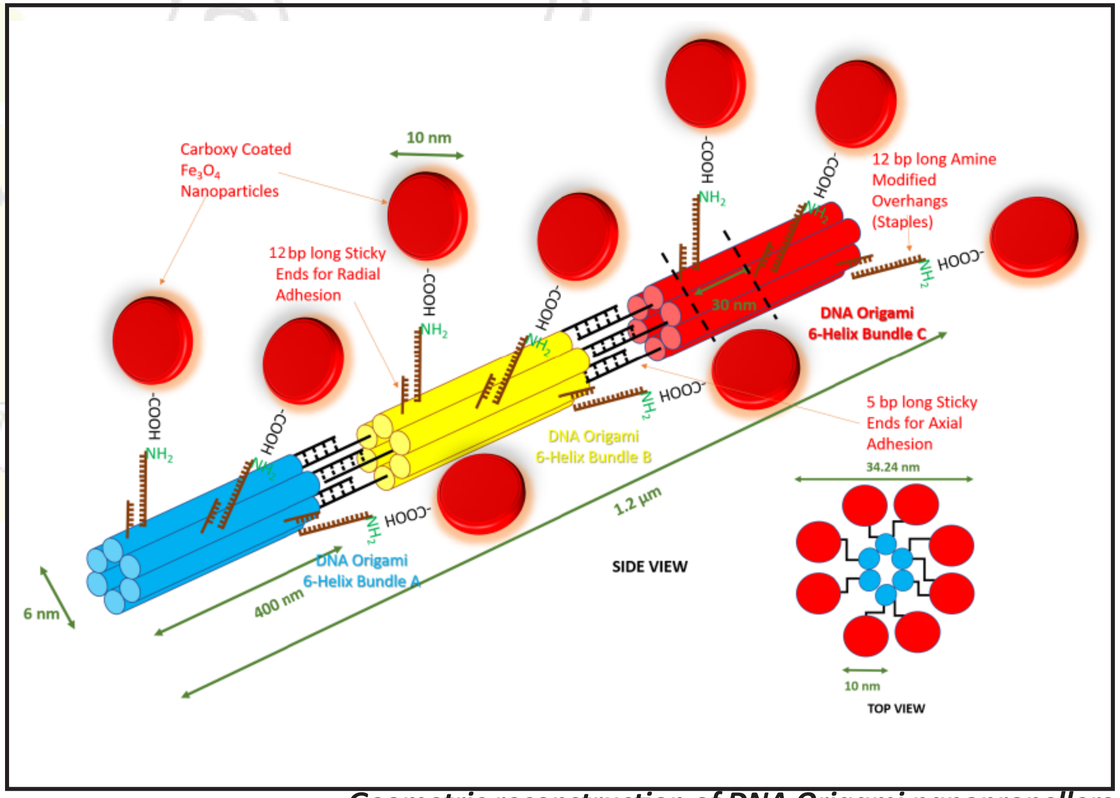
During this phase, I conducted a literature survey, received training in atomic force microscopy (AFM), and explored the optimization of iron oxide nanoparticles' size using co-precipitation methods.

Phase 2: Size Optimization of Fe3O4 Nanoparticles

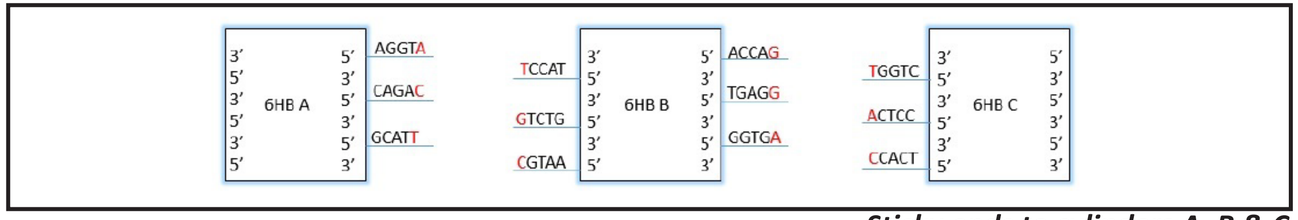
I successfully optimized the size of Fe3O4 nanoparticles, presenting my findings at a conference and submitting a paper to an international conference.



CADNAno Image of nanopropellers



Geometric reconstruction of DNA Origami nanopropellers



Sticky ends to cylinders A, B & C

nbahar@musingsinbiology.com

Phase 3: Characterization of DNA Origami Cylindrical Nanostructures

Using AFM, I characterized DNA origami cylindrical nanostructures, overcoming synthesis challenges and optimizing key factors such as storage conditions, observation time, buffer composition, and mica surface quality.

Phase 4: Gel Electrophoresis, Streptavidin-Biotin Interaction Studies, and Bioconjugation

I performed gel electrophoresis to observe DNA origami migration, studied streptavidin-biotin interactions, and successfully bioconjugated Fe3O4 nanoparticles with DNA origami structures. I presented my work in a poster and defended my thesis.

The outcomes of this project demonstrate the potential of DNA origami-based magnetic nanopropellers, which can be applied in various fields of research and development.

Towards DNA Origami Based Magnetic Nanopropellers

Nuruddin Bahar[1], Abhijith Raghuprasad[1], Ambarish Ghosh[1], Banani Chakraborty[2]

1. Centre for Nano Science and Engineering (CeNSE), IISc Bangalore

2. Department of Chemical Engineering, IISc Bangalore

Email : nbahar@cense.iisc.ernet.in

Abstract

Through DNA nanotechnology one can design very complex nanostructures via a bottom-up fabrication method. These nanostructures are now embarking in the field of application. To test the application of six-helical bundle DNA origami, we wish to mimic a silicon di-oxide-based nanopropeller, which is demonstrated using top-down nanotechnology. A two-fold approach is adopted in this respect. First is to modify the silicon dioxide nanopropeller (SiO_2-NP) with magnetic nanoparticle to observe their movement under magnetic field *in vitro*. Second is to build the nanopropeller using DNA origami through bottom-up nanotechnology where we can modify the nano-helix bundles with iron-oxide magnetic nanoparticles to be studied under a magnetic field. A couple of variants of these magnetic helices will be made for checking the movement with fluorescence microscopy and atomic force microscopy. Here, we describe a simple method for folding long, single-stranded circular genomic DNA from M13mp18 phage as the scaffold strand into cylindrical shaped 6-helix bundles (6HB). Magnetic iron oxide nanoparticles (10-20 nm) are programmed to connect to 6HB at a distance of 30 nm along the backbone of 6HB through a covalent coupling of 5' end of NH_2-modified DNA with COOH-modified magnetic nanoparticle. Aim is to generate a biocompatible propeller system, which can mimic an artificial flagellar assembly similar to an *E.coli*. Using sticky-ended cohesion, we are able to control the precise length of these 6HB to grow in multiples of 400 nm to more than 1 µm in length. Currently we are characterizing the formation of these assemblies and nanoparticles in parallel to put them together and study their behavior further.

Introduction

Maneuvering nanoscale objects in fluidic biological media in a non-invasive manner can lead to various biomedical applications. Of particular interest is the possibility of powering and controlling the motion of nanoscale objects like a cork-screw shaped magnetic nanopropeller. However, one faces various issues related to physical and chemical properties of a silicon dioxide based magnetic nanopropeller, especially biocompatibility. DNA being a biopolymer nanostructures made out of DNA such as DNA origami will be highly biocompatible. With the latest advancement of DNA Nanotechnology we can make DNA origami multiple helix bundle like structures which are proven to be robust, size-optimized and can be easily maneuvered in the presence of magnetic field upon modification with magnetic nanoparticles.

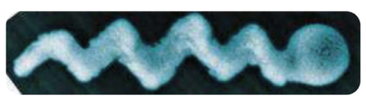


Figure 1. SiO_2 Magnetic Nanopropeller

Figure 2. DNA Origami Magnetic Nanopropeller

Schematic representation of Forming Magnetic 6 Helix Bundle

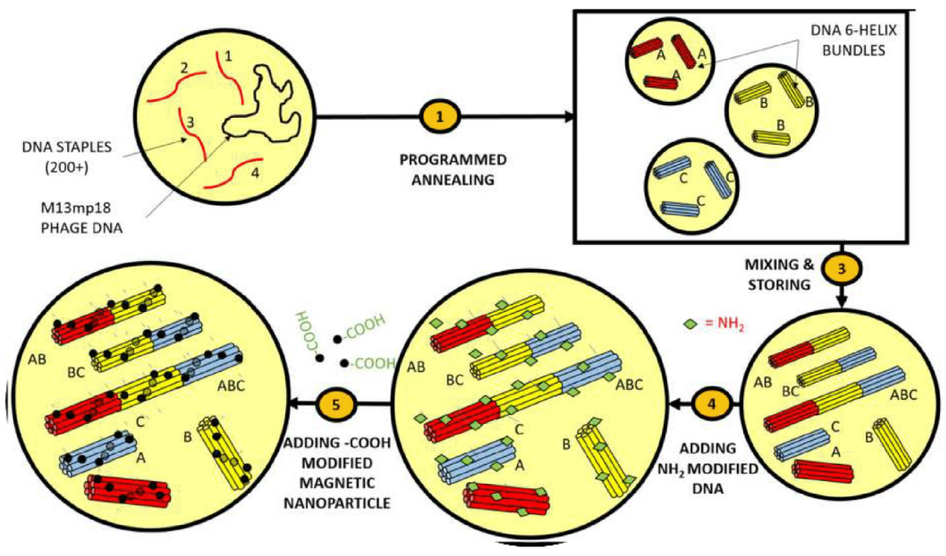


Fig 3. Schematic of synthesis of DNA Origami 6-helix Bundle Magnetic Nanopropeller

Design of individual 6HB (A,B,C) using CaDNAno

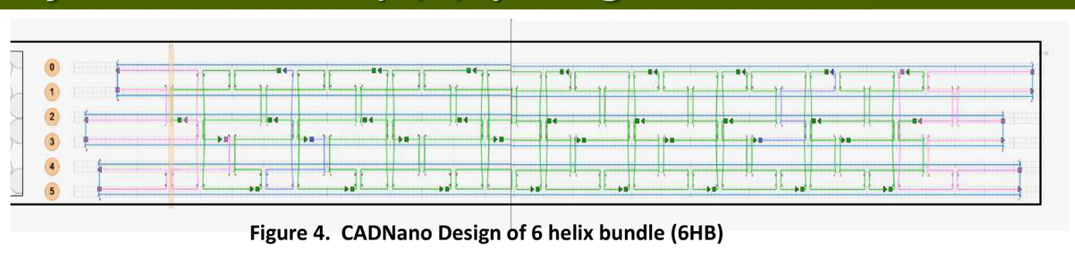


Figure 4. CADNano Design of 6 helix bundle (6HB)

Sticky ended cohesion to form A+B+C

Figure 5. Structure of 6HB showing sticky ended cohesion

AFM Visualization

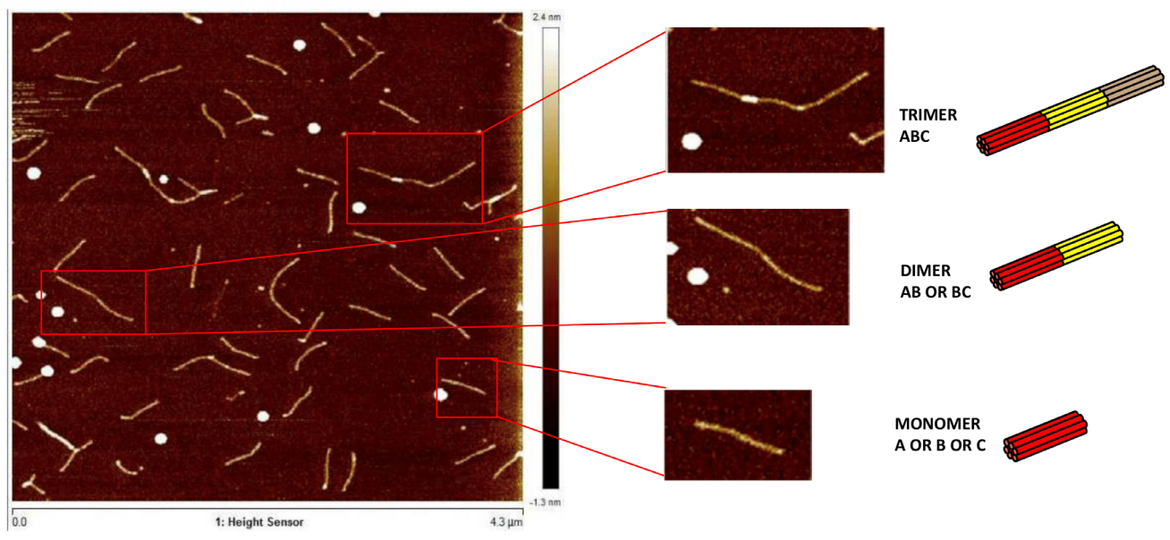
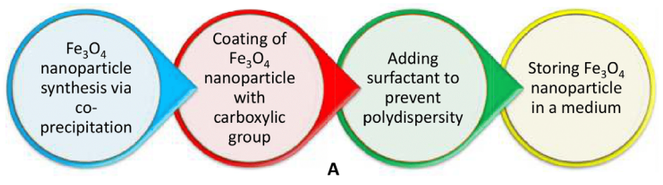


Figure 6. AFM images of non-functionalized DNA Origami 6-Helix Bundles

DNA ORIGAMI SAMPLE		MONOMER A OR B OR C		DIMER AB OR BC		TRIMER ABC	
SAMPLE #	TOTAL	#	%	#	%	#	%
1	49	36	73.5	3	6.1	10	
2	259	209	80.7	19	7.3	31	
	249	209	83.9	19	7.6	21	
			79.4		**7.0**		

Synthesis of Carboxylated Iron Oxide Nanoparticles



A

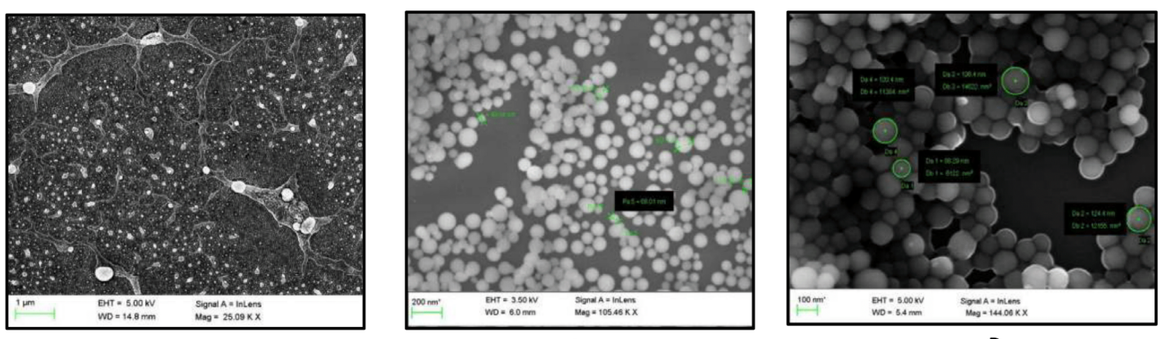


Figure 7. A. Schematic of synthesis of iron oxide nanoparticles. B. By Citric Acid method. C. By CTAB method. D. By TEPSA method

	CITRIC ACID (Method 1)	CTAB (Method 2)	TEPSA (Method 3)
Method of Synthesis	Co-Precipitation	Co-Precipitation	Co-Precipitation
Coating Agent	Citric Acid	CTAB*	TEOS@ + TEPSA$
Surfactant	PEG**	PEG	CTAB
Operating Temperature	120 °C	Room	80 °C
Storage Medium	Ethanol	Ethanol + Water	Ethanol + Water
Minimum Size Range	30 – 50 nm	20 – 65 nm	30 – 80 nm
Polydispersity	Low	Low	High

* Cetyl Trimethylammonium Bromide @ Tetraethyl Orthosilicate $ (Triethoxysilyl) Propylsuccinic Anhydride ** Polyethylene Glycol

Future Work

- To functionalize amine modified DNA 6HB with carboxylated magnetic iron oxide nanoparticles via EDC-NHS coupling

- To optimize the yield of trimer ABC formation
- To visualize the origami nanopropeller formation with fluorescent tagging using biotin modified DNA instead of amine modified DNA with further modification with Cy-3 labelled Streptavidin.
- To test the manoeuvrability of 6HB in a fluidic medium mimicking cellular environment under magnetic field

References

1. Ghosh, A., & Fischer, P. (2009). Controlled propulsion of artificial magnetic nanostructured propellers. Nano Letters, 9(6), 2243–2245.
2. Lu, A.-H., Salabas, E. L., & Schüth, F. (2007). Magnetic Nanoparticles: Synthesis, Protection, Functionalization, and Application. Angewandte Chemie International Edition, 46(8),1222–1244.
3. Kuzyk, A., Simmel, F., Schreiber, R., Fan, Z., Pardatscher, G., Roller, E.-M., Högele, A., Liedl, T. (2012). DNA-based self-assembly of chiral plasmonic nanostructures with tailored optical response. Nature, 483(7389), 311–314.

Acknowledgements

Department of Biotechnology
Ministry of Science & Technology, Government of India

Micro Nano Characterization Facility
Your One Source for Materials, Device Characterization and Analysis

Dr. Rahul Roy Group

PROJECT 2 : MASTER OF SCIENCE (BIOMEDICAL ENGINEERING) PROJECT

Abstract:

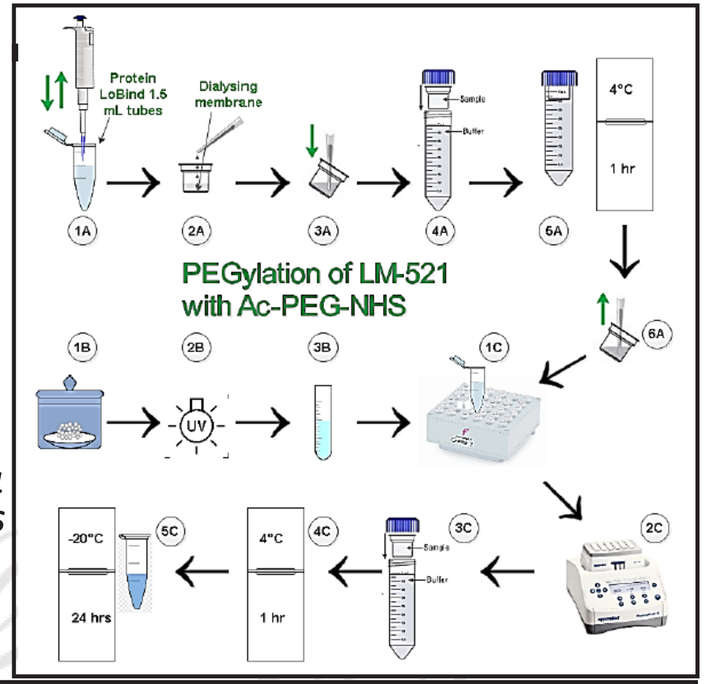
In this project, we aimed to develop mechanoresponsive hydrogels using poly (ethylene glycol)-laminin (PEG-LM) for drug delivery and gene therapy applications. These hydrogels, incorporated with human mesenchymal stem cells (hMSCs), allowed for the study of the link between mechanical properties and cellular behavior.

By varying the ratios of the cross-linker (4-arm acrylate PEG or 4-Ac) to the matrix-metalloproteinase (MMP) degradable VPM peptide, we tuned the mechanical properties of the hydrogels to facilitate tissue regeneration, including vascular and bone tissues.

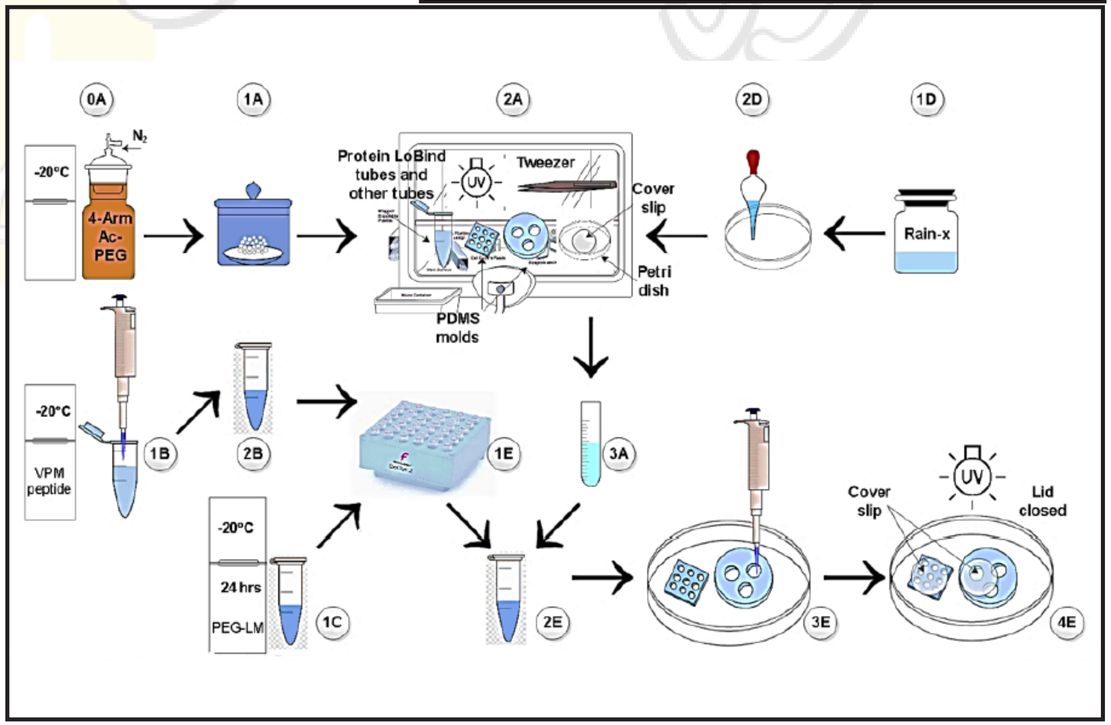
Through innovative bulk rheology measurements and a novel nanoindentation technique, we discovered that the bulk shear elastic modulus of the hydrogels strongly depended on the applied normal force and remained independent of their chemical composition.

Phase 1: Preliminary Report

During this phase, I conducted an extensive literature survey, performed a risk assessment, and developed a comprehensive experimental plan. I prepared the cell culture and established day-based measurements. Additionally, I created a characterization schedule and devised a Gantt chart to effectively manage the project.



PEGylation of LM-521 with Ac-PEG-NHS



UV-Photopolymerization reaction

Phase 2: Culture of Human Bone-Marrow Mesenchymal Stem Cells (hMSCs) and Hydrogel Preparation

In this phase, I successfully cultured human bone-marrow mesenchymal stem cells (hMSCs) and prepared the hydrogels. This involved employing PEGylation, cross-linker and peptide chemistry, and UV photo-polymerization reactions to create the desired PEG-LM hydrogels.

Phase 3: Bulk Rheology and Nanoindentation Measurements for hMSC-Laden Hydrogels

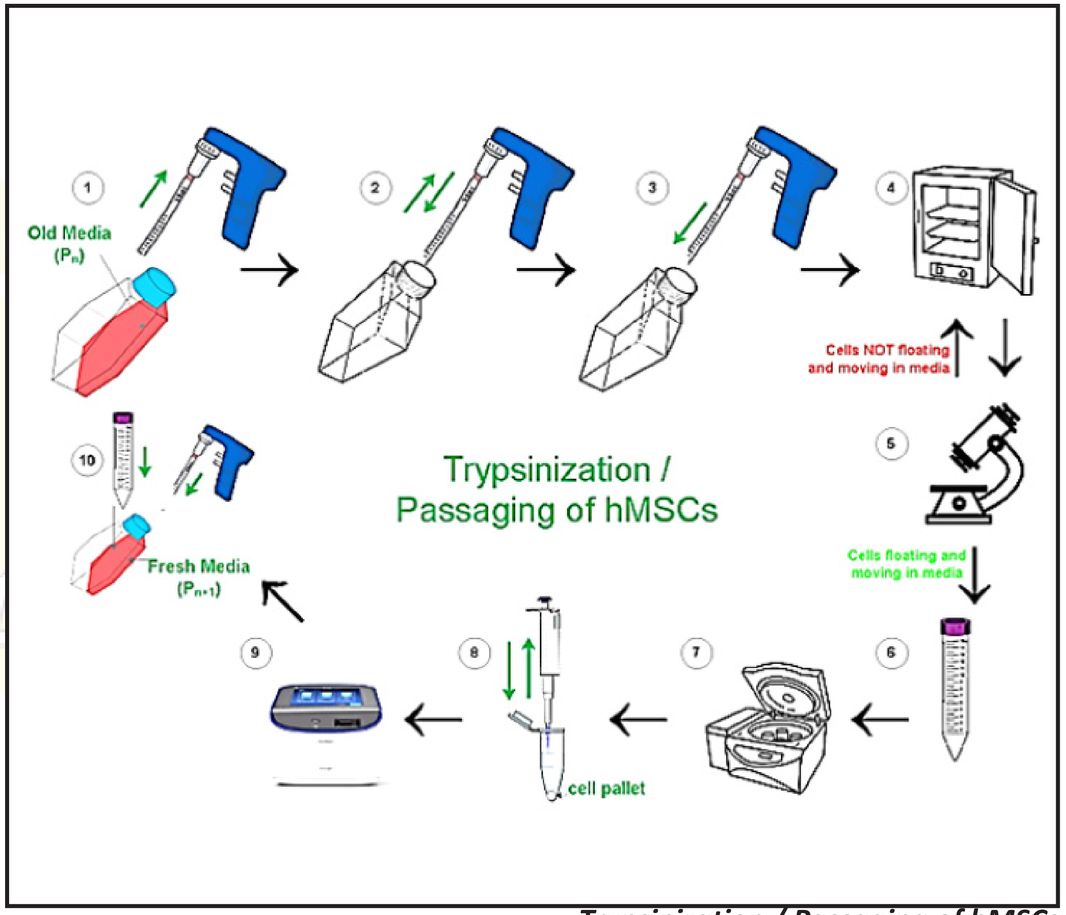
During this phase, I focused on the characterization of the hMSC-laden hydrogels. I delved into the theoretical aspects and mathematics involved in bulk rheology and nanoindentation measurements. Additionally, I made critical decisions regarding the physical aspects (geometry), instrumental aspects (measurement modes and techniques), and chemical aspects (adjusting biodegradability through cross-linker and peptide ratios) of the characterization process.

Phase 4: Statistical Data Analysis, Thesis, and Poster Submission

In the final phase, I utilized Prism (Graph Pad) for statistical data analysis. I completed the thesis, summarizing the research findings and contributions. Additionally, I prepared and submitted a poster to disseminate the key aspects of the project.

Key Results:

- Decreasing the 4-Ac concentration in the gel resulted in increased bio-degradation rates and decreased gel stiffness over time.
- Hydrogels with a 4-Ac to VPM ratio of 2:1 exhibited initial stiffening, potentially attributed to the microsphere-like capability of hMSCs.
- The presence of hMSCs reduced gel stiffness irrespective of the cross-linker-peptide ratio.
- Nanoindentation measurements revealed a bi-modality in effective Young's modulus (Eeff) across the gel area, which requires further investigation to confirm.
- By understanding the mechanoresponsive properties of PEG-LM hydrogels and their interaction with hMSCs, we contribute to the advancement of drug delivery and gene therapy strategies.



Trypsinization / Passaging of hMSCs

PROJECTS

nbahar@musingsinbiology.com

Design and Characterization of PEG-Laminin Based Mechano-responsive Hydrogels for Drug Delivery and Gene Therapy

Bahar, N., Dobre, O., Vassalli, M. and Salmeron-Sanchez, M.

Centre for the Cellular Microenvironment (CeMi) , University of Glasgow, 79-85 Oakfield Ave, Glasgow G12 8LT

Introduction

With advancement in technology and modelling techniques in stem cell-based tissue engineering, a strong link has been demonstrated between the tissue's mechanical properties and cell behavior. In this regard, hydrogels have been the preferred materials to study such a relationship. The focus is shifting from the modelling of the biomimetic system towards the establishment of epigenetic factors that mimic natural developmental processes like embryogenesis. To establish such a role, we developed 3-D Poly(ethylene-glycol)-Laminin (PEG-LM) hydrogels incorporated with human mesenchymal stem cells (hMSCs) in varying ratios of the cross-linker (4 arm acrylate PEG or 4-Ac) to the matrix-metalloproteinase (MMP)-degradable VPM peptide to tune the mechanical properties for tissue regeneration, including vascular and bone. By using an innovative bulk rheological method along with a novel nanoindentation technique, we evidentiate the fact that the bulk shear elastic modulus of hydrogels is strongly dependent on the applied normal force independent of their chemical composition. We also establish the fact that (1) decreasing the amount of 4-Ac in the gel increases the rate of degradation and decreases the stiffness of gel over time (2) 2:1 ratio shows an initial increase in the overall stiffness, mainly due to microsphere-like pore occupation capability of hMSCs (3) incorporated hMSC reduce the hydrogel stiffness substantially, regardless of the cross-linker-peptide ratio (4) a bimodality in effective Young's modulus (E_{eff}) is established via nanoindentation, although more tests are required to confirm this fact.

Materials and Methods

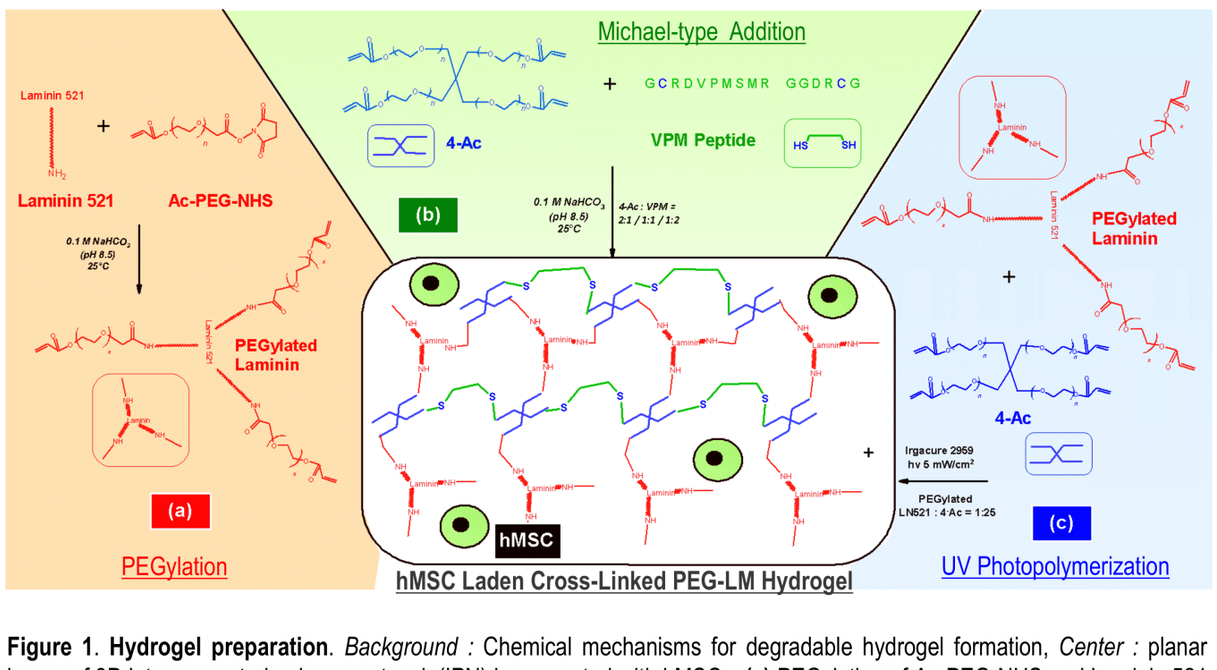


Figure 1. **Hydrogel preparation**. *Background :* Chemical mechanisms for degradable hydrogel formation, *Center :* planar image of 3D interconnected polymer network (IPN) incorporated with hMSCs. **(a)** PEGylation of Ac-PEG-NHS and Laminin 521 **(b)** Michael-type addition for forming thiol cross-linked hydrogel using 4-Ac and VPM peptide (having cysteine residues colored differently) **(c)** UV Photopolymerization reaction using photoinitiator (Irgacure 2959®) to connect 4-Ac to PEGylated LN521.

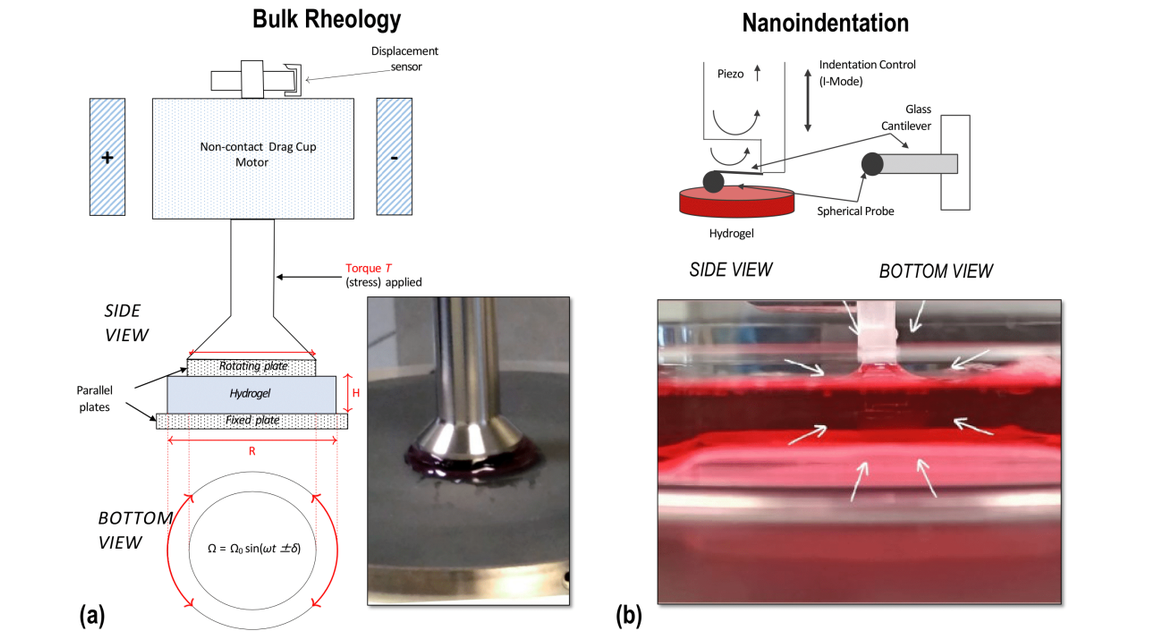


Figure 2. Characterization of hydrogels. (a) Measuring hMSC-laden hydrogel storage modulus(*G'*) using controlled shear stress (CSS) parallel-plate oscillatory rheometer (MCR302®) (*inset:* rheometer plate pressing upon hydrogel)**(b)** Measuring effective Young's modulus (E_{eff}) of hMSC-laden hydrogel using spherical probes fixed to a cantiléver (*top*) in I-mode (Chiaro Nanoindentor®) and 50µL hydrogel in DMEM(10% FBS) médium (*bottom*) set for indentation study.

Results

Bulk Rheology

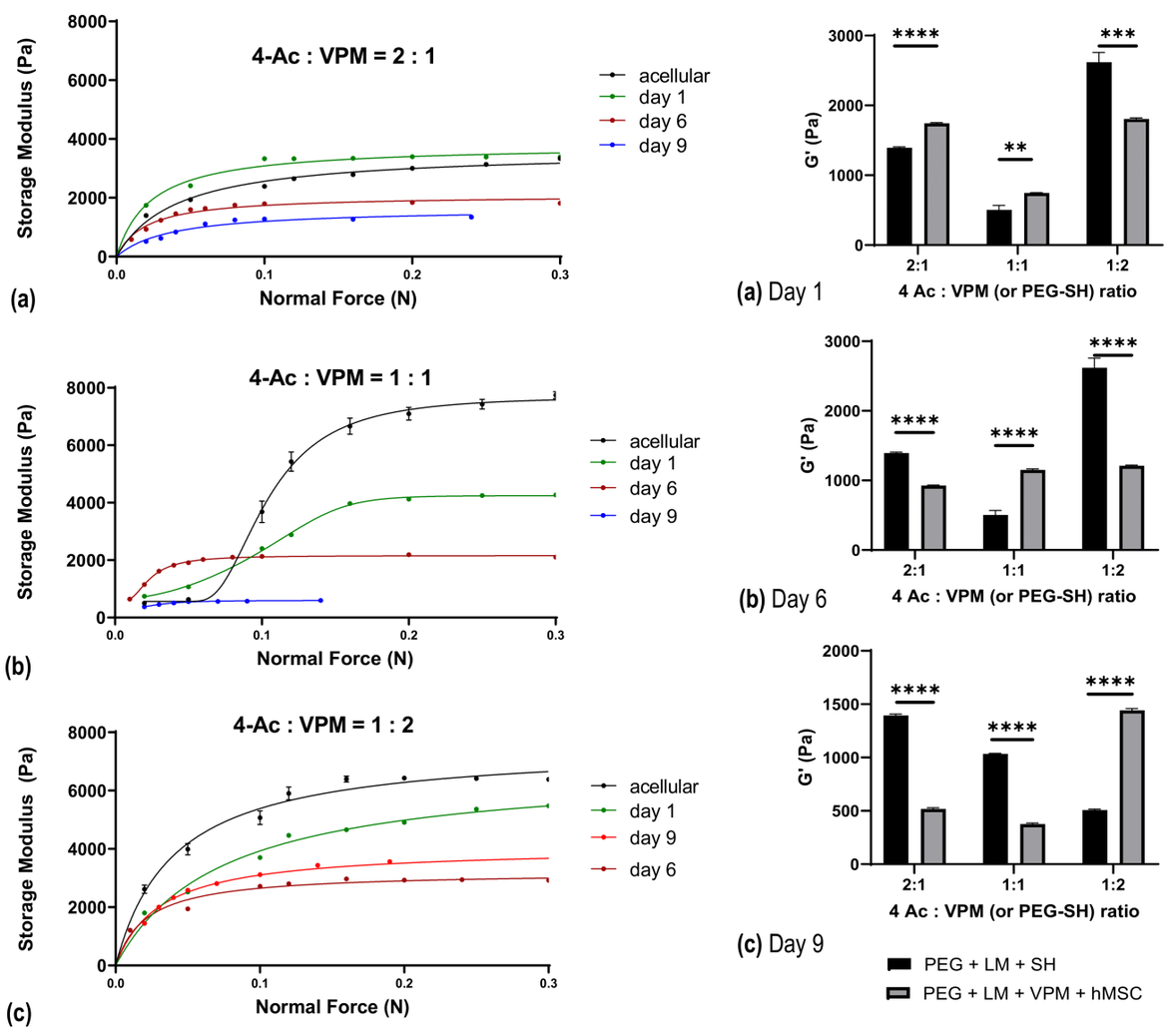


Figure 3. Temporal variation in storage modulus (Pa) with normal applied force (in N) for different 4-Ac : VPM (or PEG-SH in acellular non-degradable samples) ratios **(a)** 2:1 **(b)** 1:1 **(c)** 1:2. Note that samples marked in terms of days are all cell-laden and colored, while acellular samples are marked black. Error bars plotted for (n=5) mean ± σ.

Figure 4. (a) Day 1, **(b)** Day 6 and **(c)** Day 9 samples for bulk rheology measured at a constant normal load of 20 mN. Multiple paired t-test per row (stars represent statistical significance with **** = adjusted p value < 0.000001, *** = 0.000001, ** = 0.00004).

Nanoindentation

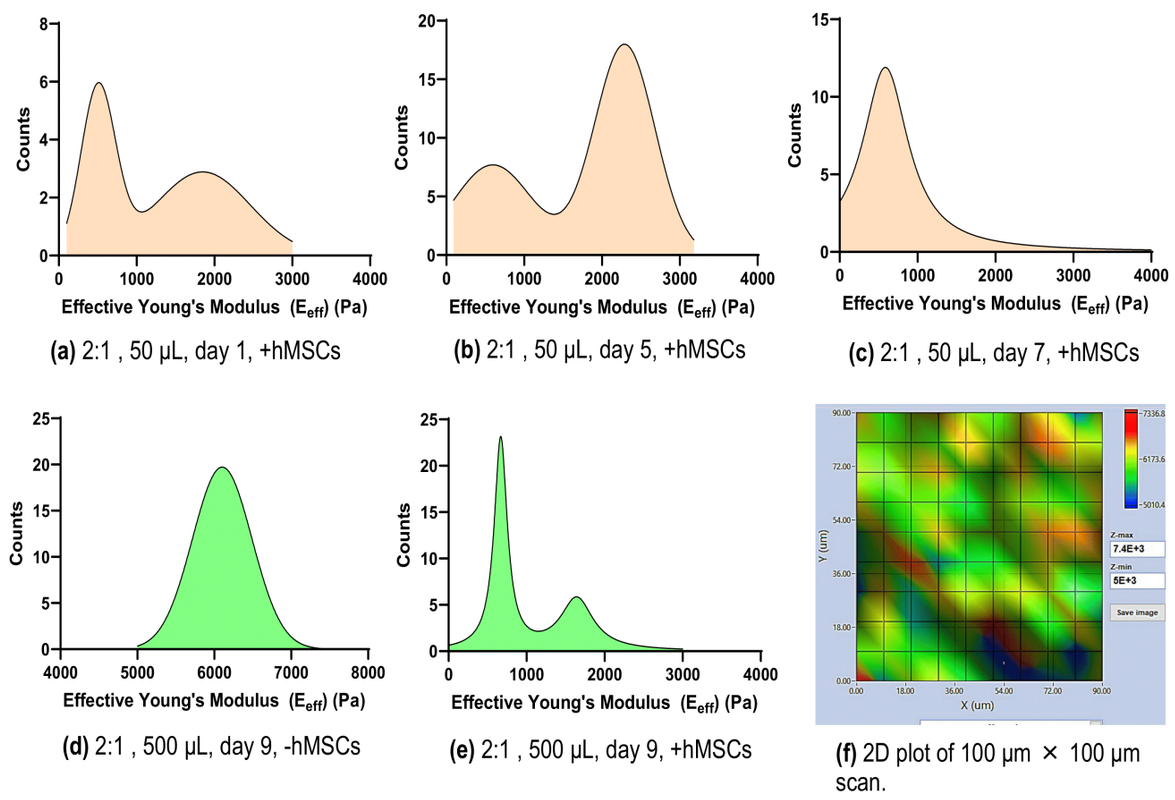


Figure 5. Temporal variability in effective Young's modulus (E_{eff}) with ratio of 4-Ac : VPM (or PEG-SH in acellular non-degradable samples) to be 2:1 and with (+) and without (-) hMSC for 50 µL and 500 µL samples using nanoindentation technique. Top row (**(a),(b)** and **(c))** represents the effective distribution curves for 50 µL samples with hMSCs measured at days d1, d5, d7; while bottom row (**(d)** and **(e))** represents the distribution curves for 500 µL samples with (+) and without (-) hMSCs. Used Sum of Gaussian (or Sum of Lorentzian) for bimodal curve-fitting and Gaussian (or Lorentzian) for unimodal curve-fitting. One sample t-test and one sample Wilcoxon signed ranked test. **(f)** represents 2D plot of the scan area showing variations in E_{eff} at a given (X,Y) co-ordinate with darker patches showing lower values.

Discussion/Conclusion

1. The range of storage modulus, G' (Pa) for 2:1 samples were found to be (**0.5 - 3.2 kPa**) which varied with normal load (0.01- 0.3N).
2. **2:1** ratio showed an initial **increase** and subsequent **decrease** in the overall stiffness, mainly due to **microsphere-like pore occupation** capability of hMSCs.
3. Due to **localization**, nanoindentation measurements showed a **bimodal** behavior in the **effective Young's modulus (E_{eff})** of hydrogels overtime.
4. **Decreasing** the amount of cross-linker (**4-Ac**) in the gel **increases** the rate of **degradation** and **decreases** the **stiffness** of gel over time.
5. Incorporated **hMSCs reduce** the hydrogel **stiffness** substantially, regardless of the cross-linker-peptide ratio.

Future Directions

1. **Further experimentation** to study the effect of various **hydrogel components**
   - Changing the concentration of **LN521** at the same 4-Ac : VPM peptide ratio.
   - Varying the % (w/v) of **4-Ac** in the hydrogel to study effect of cross-linker
2. Explore the **model** to establish a clear correlation between the **density** of cells and the effect it has on the **degradability.**
3. Comparing **nanoindentation** data with **day 1** and **day 6** measurements for both volumes to understand the temporal changes in stiffness.
4. **Fluorescence staining** of cells or their proliferation or differentiation-specific components to understand the effect on osteogenetic or adipogenetic potential of the mesenchymal stem cells *in vitro*.

Acknowledgements

EPSRC Pioneering research and skills

UK Regenerative Medicine Platform

MRC Medical Research Council

References

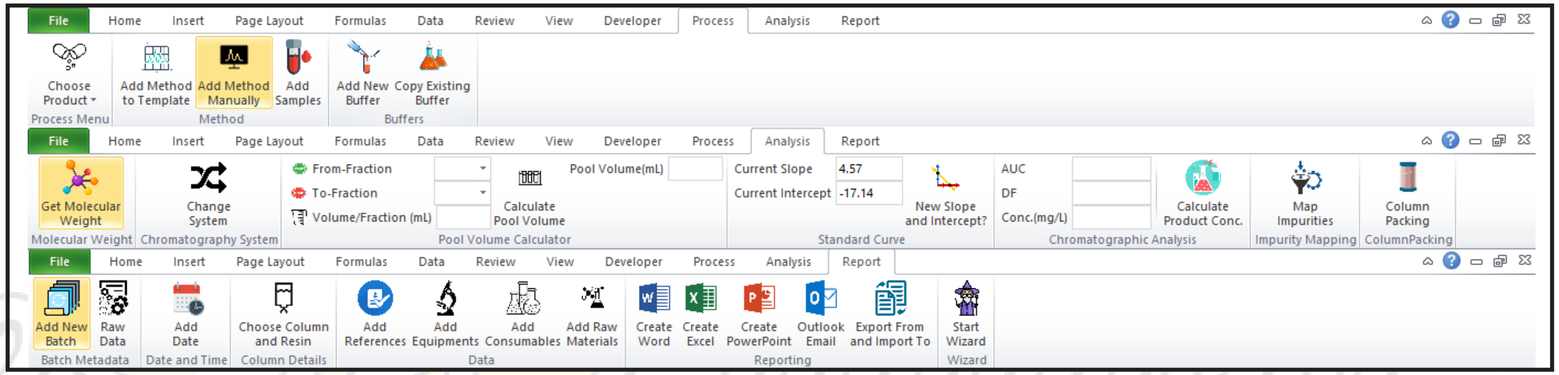
1. Ferreira.et al. Bi-directional cell-pericellular matrix interactions direct stem cell fate. Nature Communications **9**,1–12 (2018).
2. Lutolf & Hubbell, Synthesis and physicochemical characterization of end-linked poly(ethylene glycol)-co-peptide hydrogels formed by Michael-type addition. Biomacromolecules **4**,713–722 (2003).
3. Yang, T. Mechanical and Swelling properties of Hydrogels december,67(2012).

PROJECT 3 : VBA-BASED ROBOTIC BATCH REPORTING APPLICATION

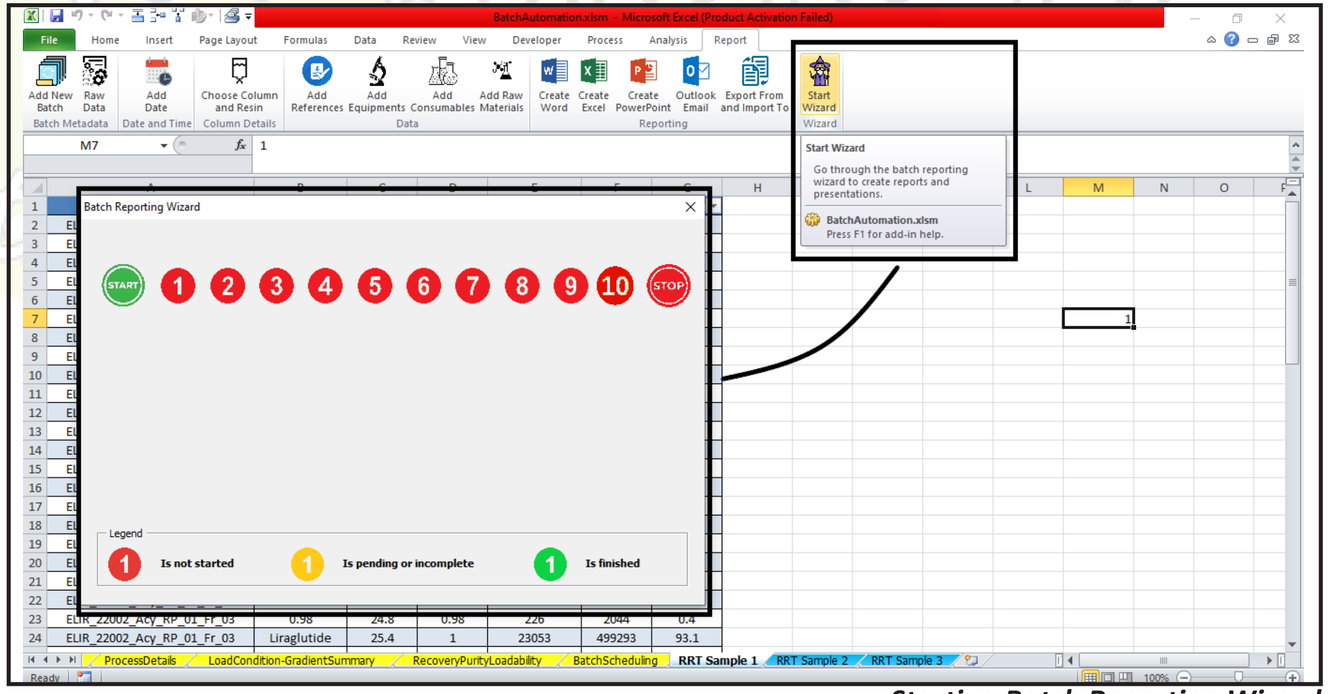
After working on multiple downstream processing batches of insulin variants, I discovered that the batch reporting process was excessively repetitive and monotonous. This inspired me to automate most of the processes using VBA (Visual Basic for Applications) in MS Excel's inline backend code. It took me over six months to grasp various VBA functionalities, implement custom UI code to create Excel buttons, develop and connect VBA userforms, and integrate the entire batch reporting process to generate reports in various formats such as Word, Excel, and PowerPoint.

I designed relevant buttons as modular functions to perform independent analysis and other tasks. For instance, I added a 'Choose Product' menu button in the custom buttons, enabling the selection of a biosimilar and its associated downstream process. Additionally, I created a 'Change System' button to facilitate inline alteration of the chromatography system.

To establish an integrated reporting process, I consolidated each module into a step within a reporting wizard, which can be accessed through the 'Start Wizard' button under the Report tab.



Custom UI buttons created in Excel

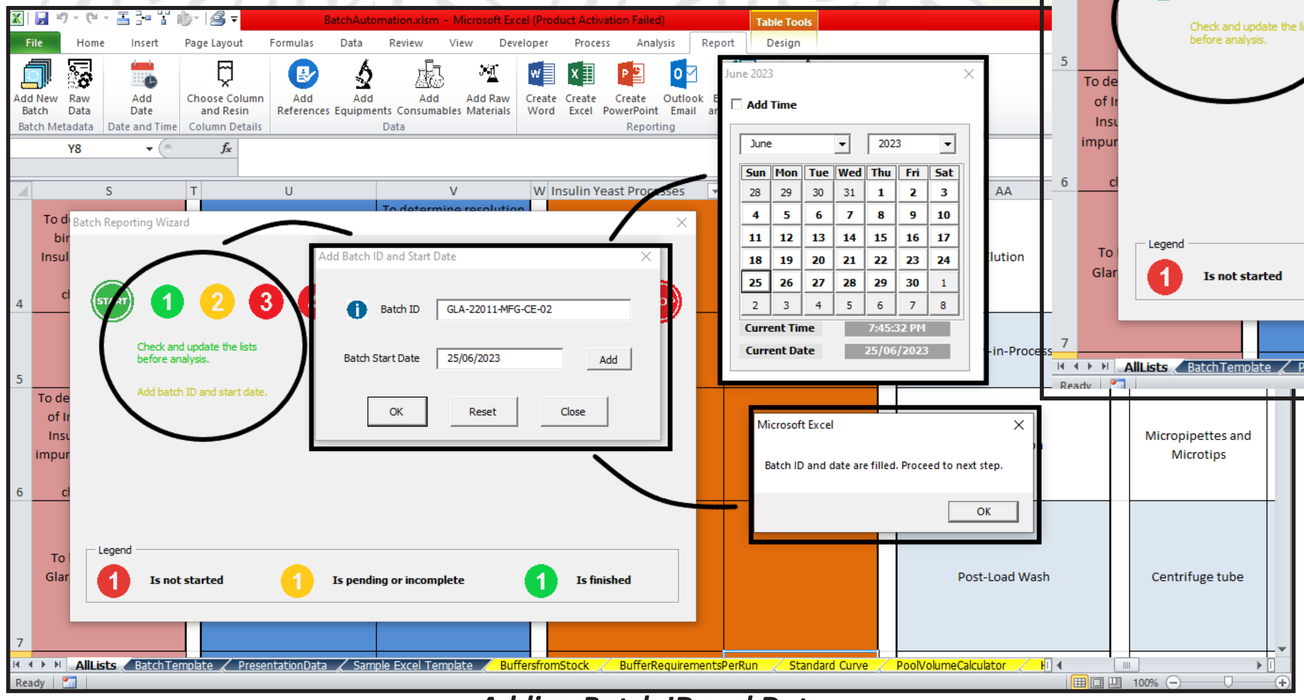


Starting Batch Reporting Wizard

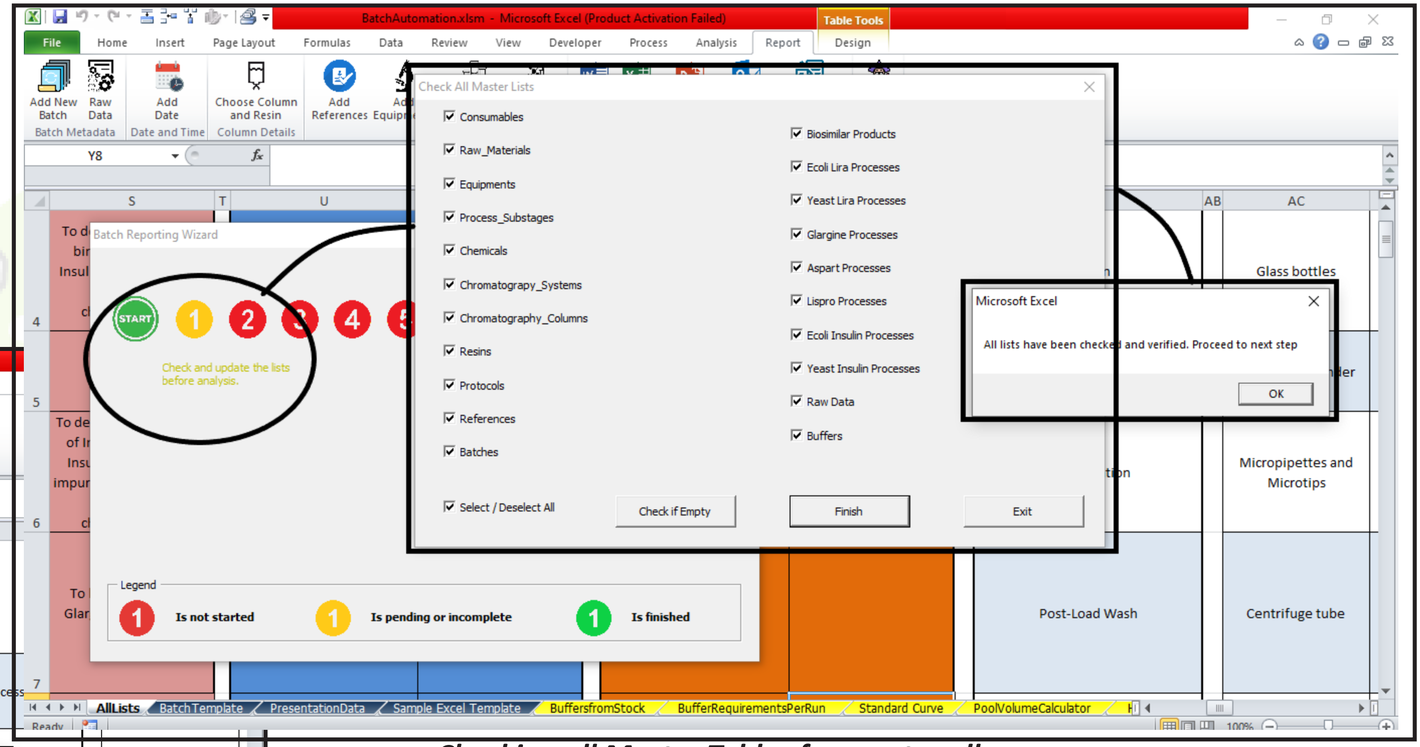
nbahar@musingsinbiology.com

The first step in the wizard verifies all master template tables for empty cells. This step ensures that any removed values are identified or replaced with new ones in the existing batch template table.

The subsequent step involves adding a batch number and the batch experiment start date using a date picker.

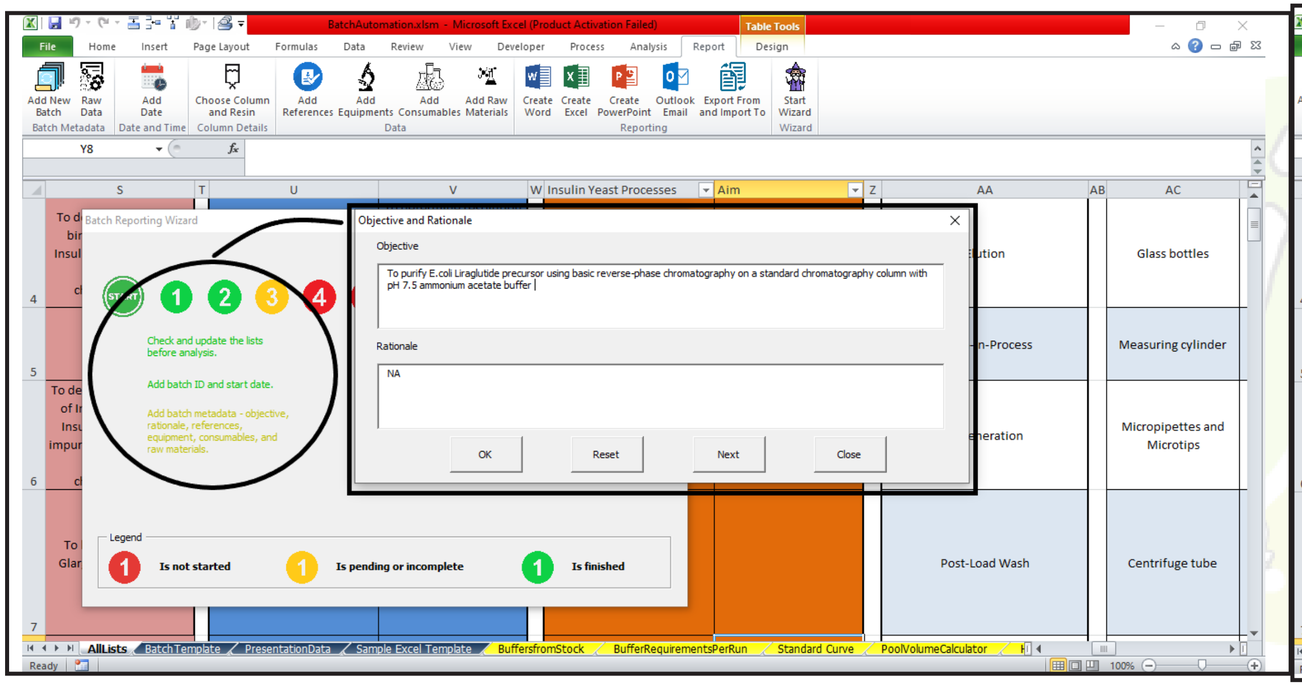


Adding Batch ID and Date

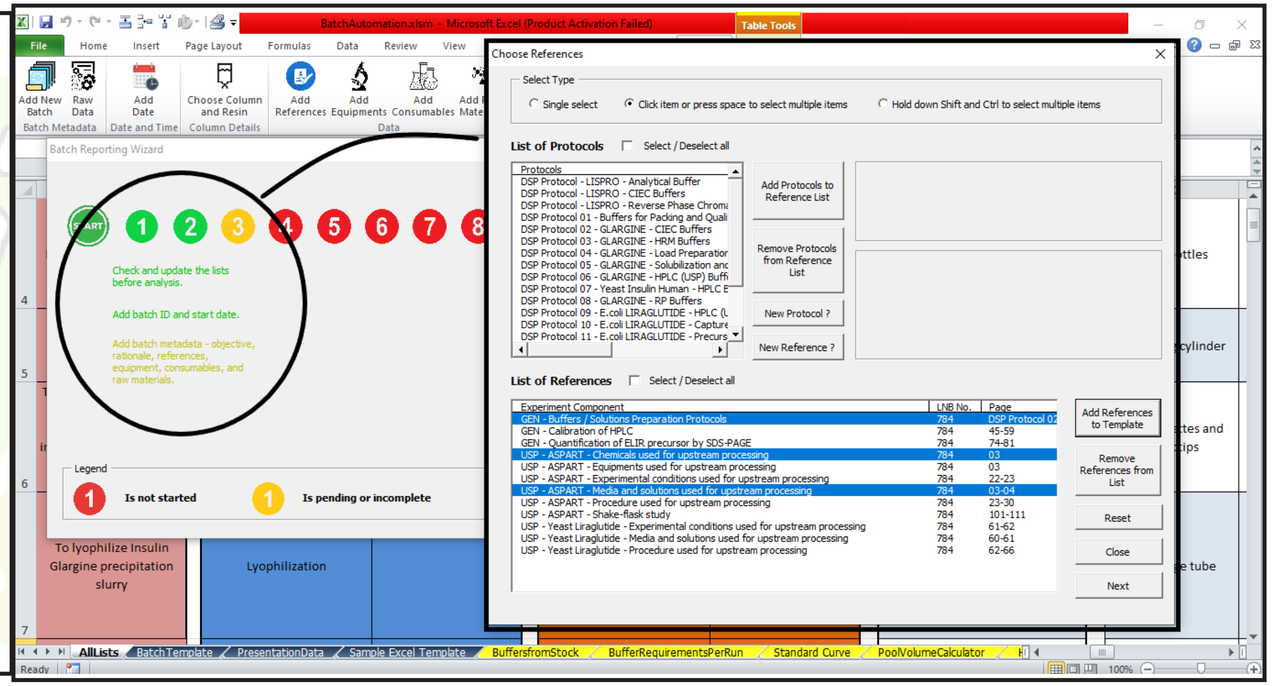


Checking all Master Tables for empty cells

Following that, a series of integrated screens prompt for relevant data, including objectives, rationale, references, equipment, consumables, and raw materials.

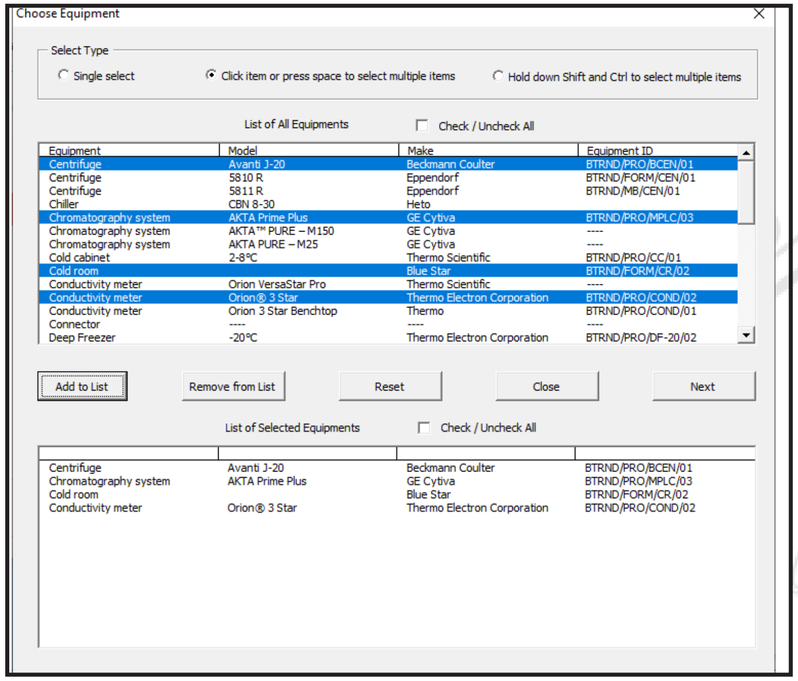


Adding Objective and Rationale

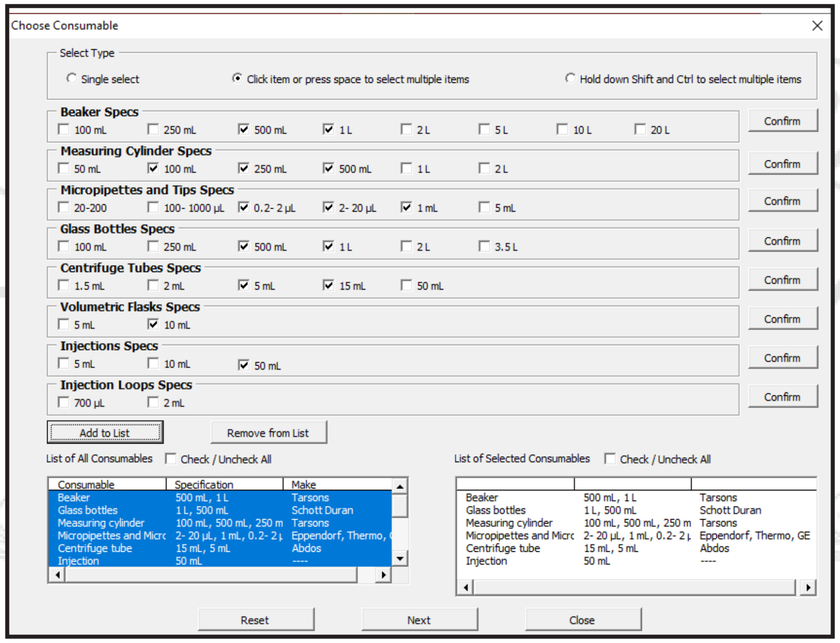


Adding References

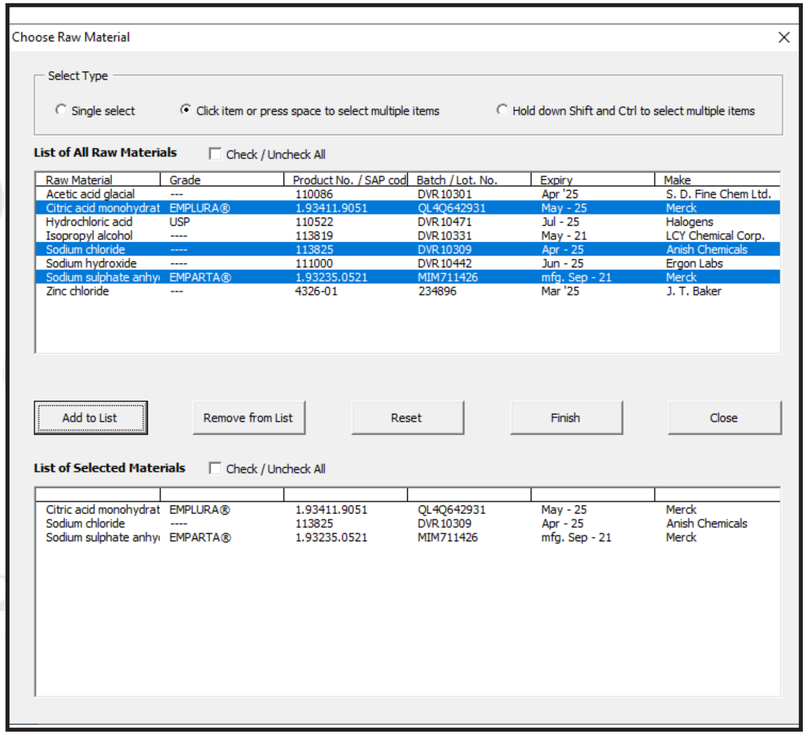
nbahar@musingsinbiology.com



Adding Equipment



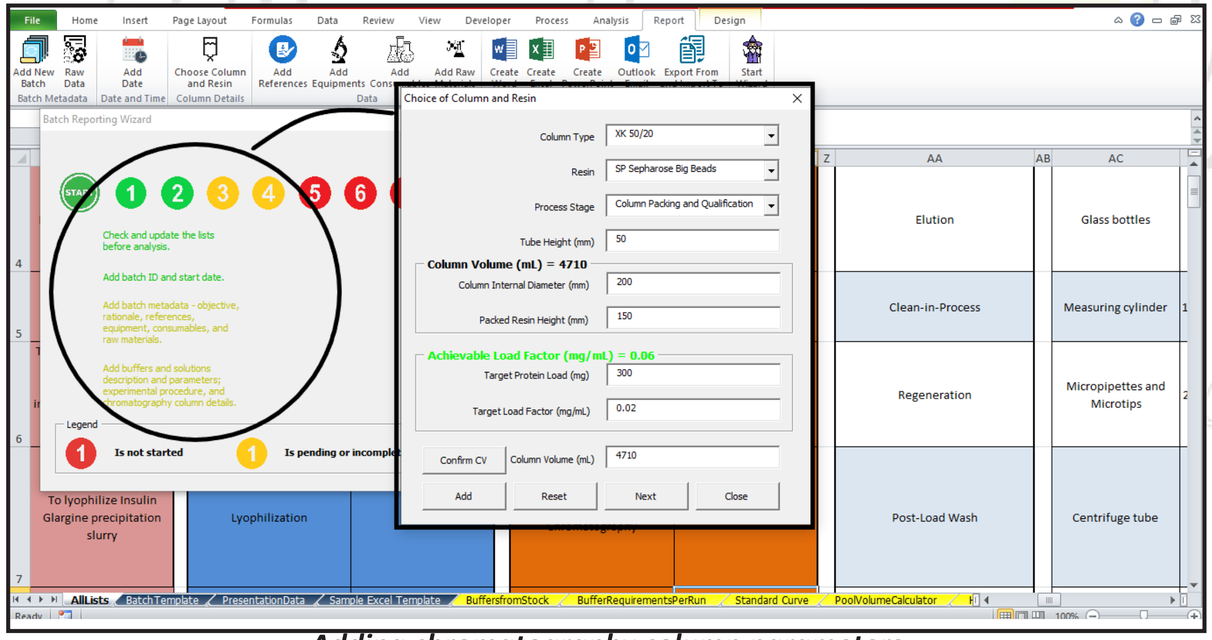
Adding Consumables



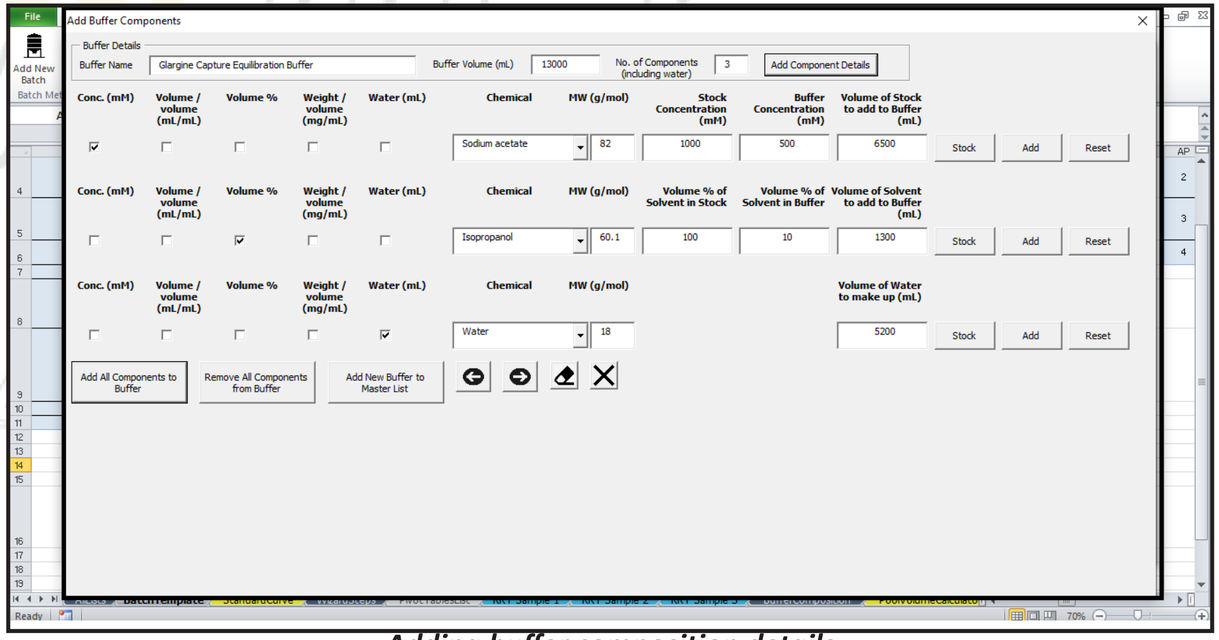
Adding Raw Materials

The next step involves adding details about the chromatography column, which also verifies the relevance of the selected column for the experiment by calculating its volume and comparing it with the target load factor achievable with that particular column.

Next, the number of buffers and their compositions are specified using check buttons that offer different composition options. At the end, a table displays the buffer amount along with its composition.



Adding chromatography column parameters

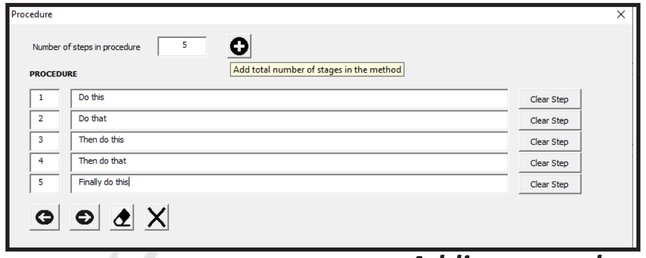


Adding buffer composition details

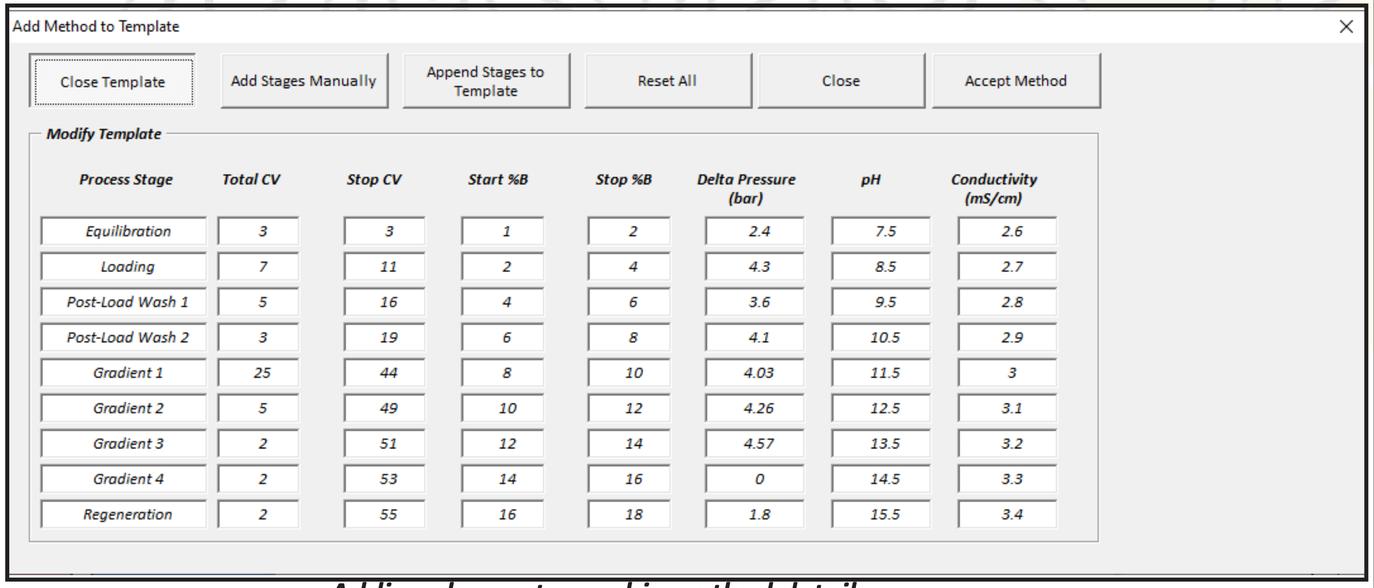
Afterwards, the experimental procedure is added in a defined number of steps.

Subsequently, the chromatographic method details are entered, including total CV, start %B, stop %B, pressure, pH, and conductivity (mS/cm). There is also an option to add all stages annually or append more stages to an existing template.

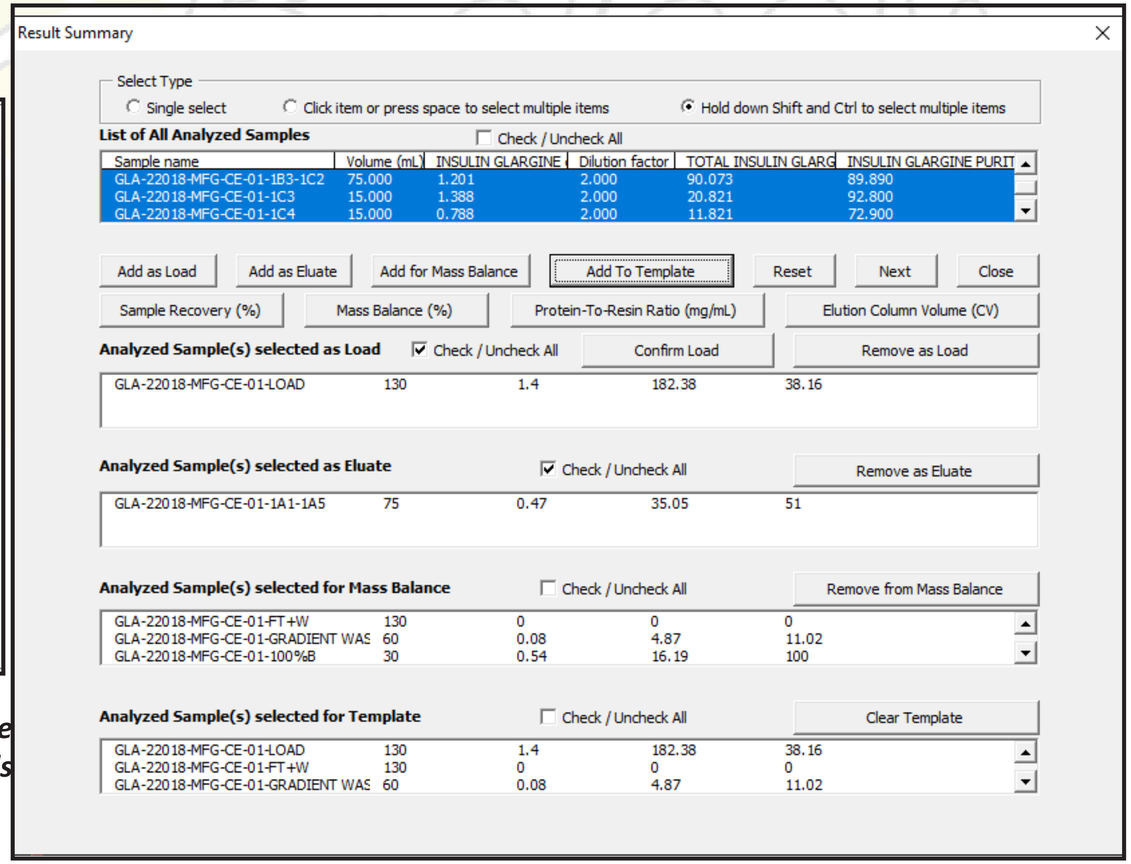
Next, sample analysis parameters are added from a template table, which includes values for volume, content, dilution factor, and chromatographic purity. The load sample, main eluate sample, and other relevant samples necessary for performing mass balance can be selected from these samples. The load sample helps determine the protein content in the input, while the eluate content indicates the overall yield. Finally, all the relevant samples are chosen to be added to the template table.



Adding procedure



Adding chromatographic method details

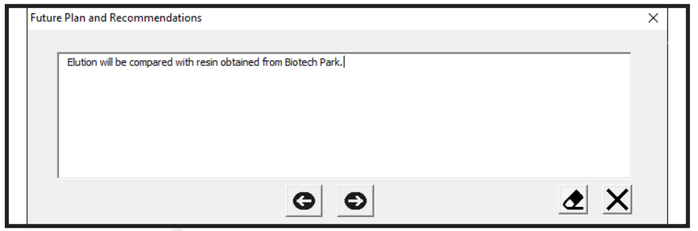


Adding sample analysis details

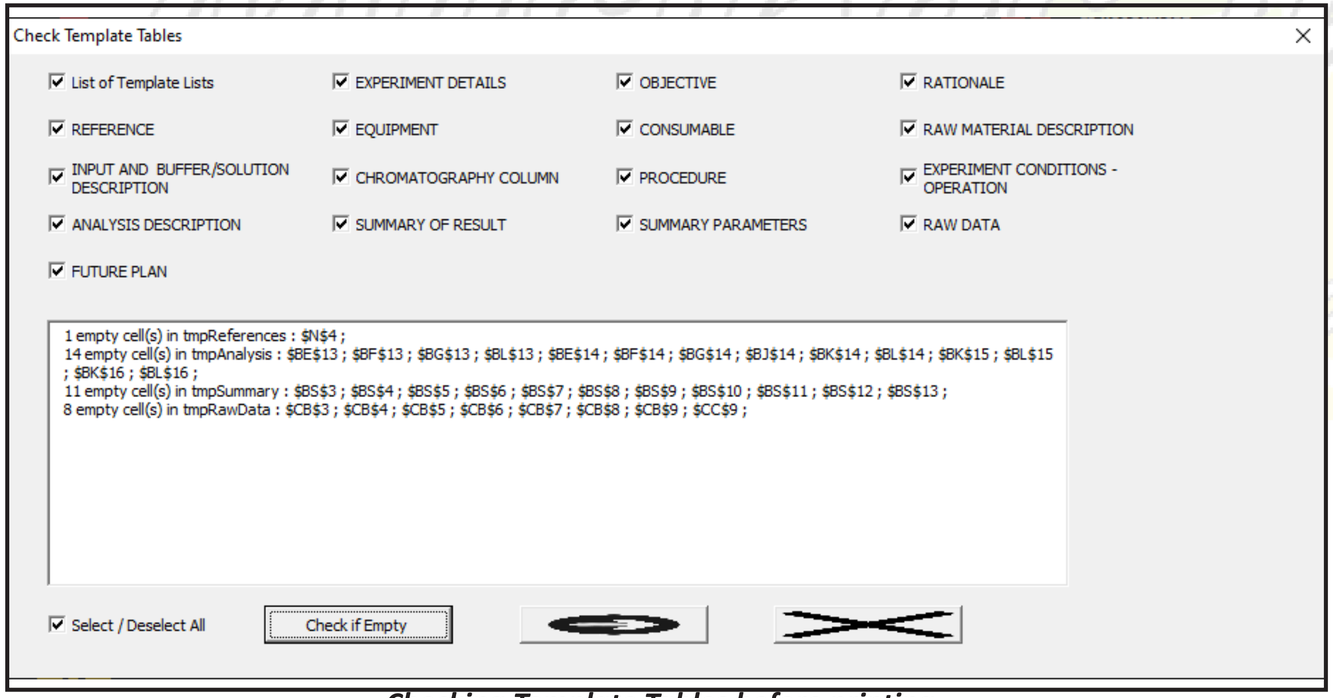
Future plans and recommendations for further experiments are then added.

Following that, all the batch template tables are confirmed before proceeding to the printing stage.

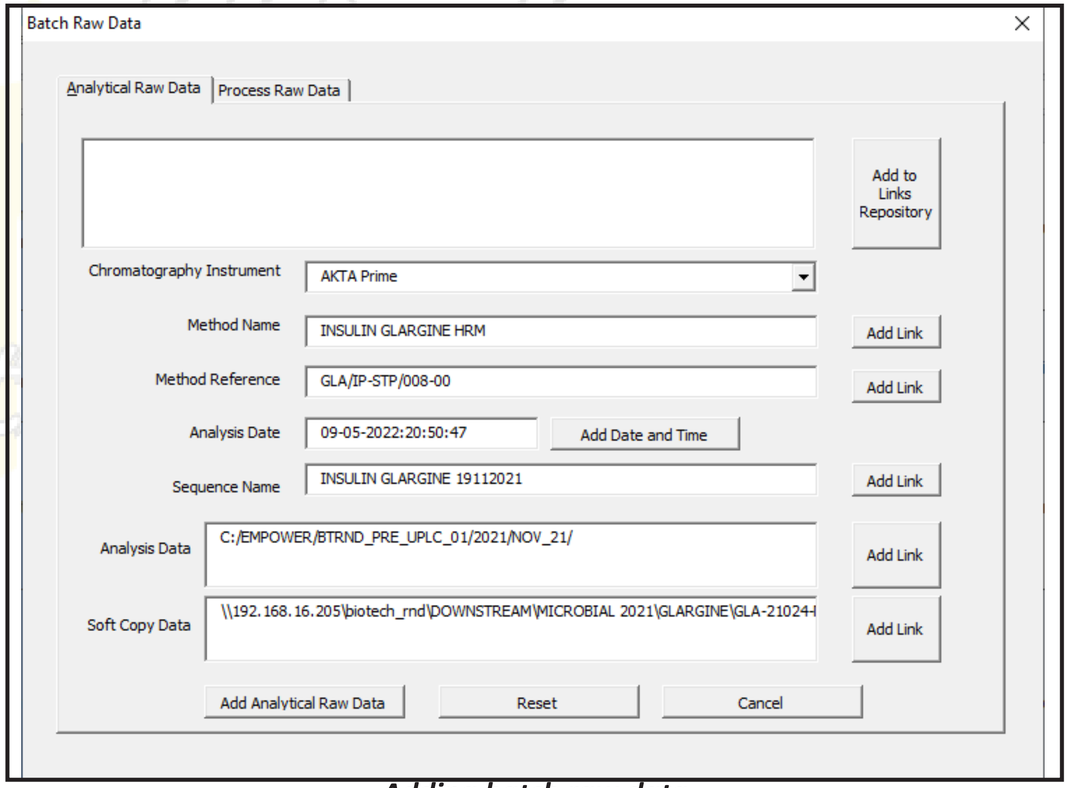
Batch raw data, such as Excel and Word files used and shared on the server, the completion date of process and analysis experiments, and any previous method reference protocols used are added next.



Adding future plans and recommendations

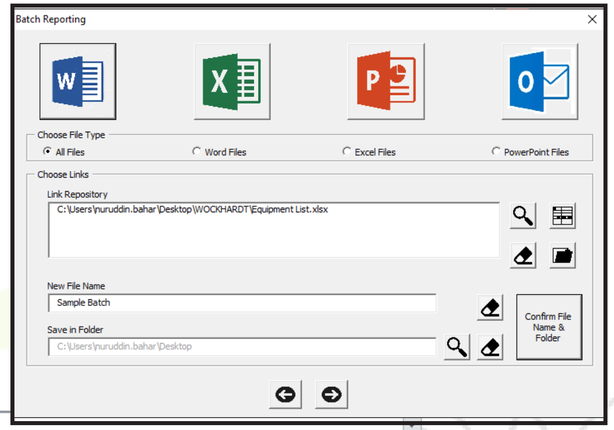


Checking Template Tables before printing

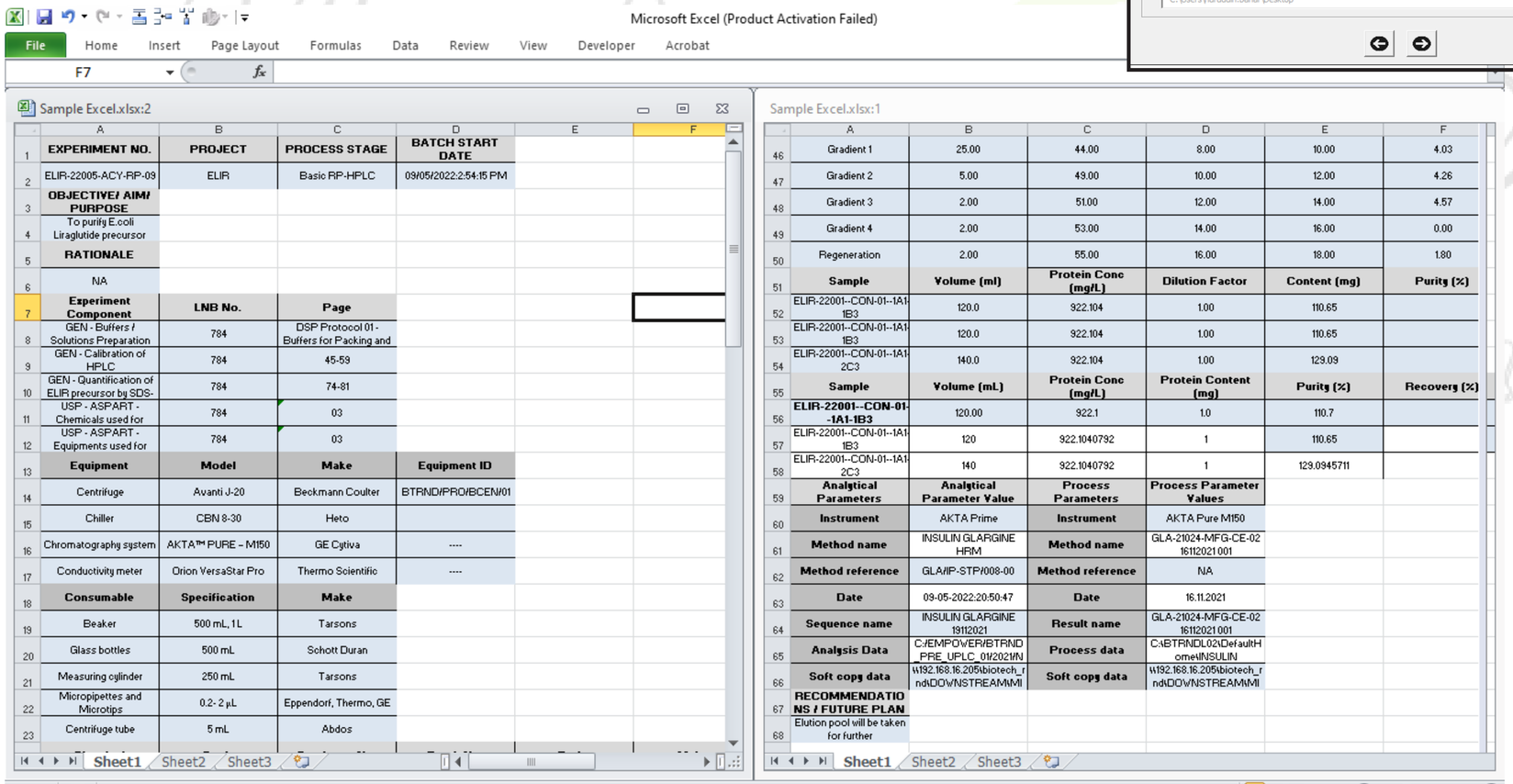


Adding batch raw data

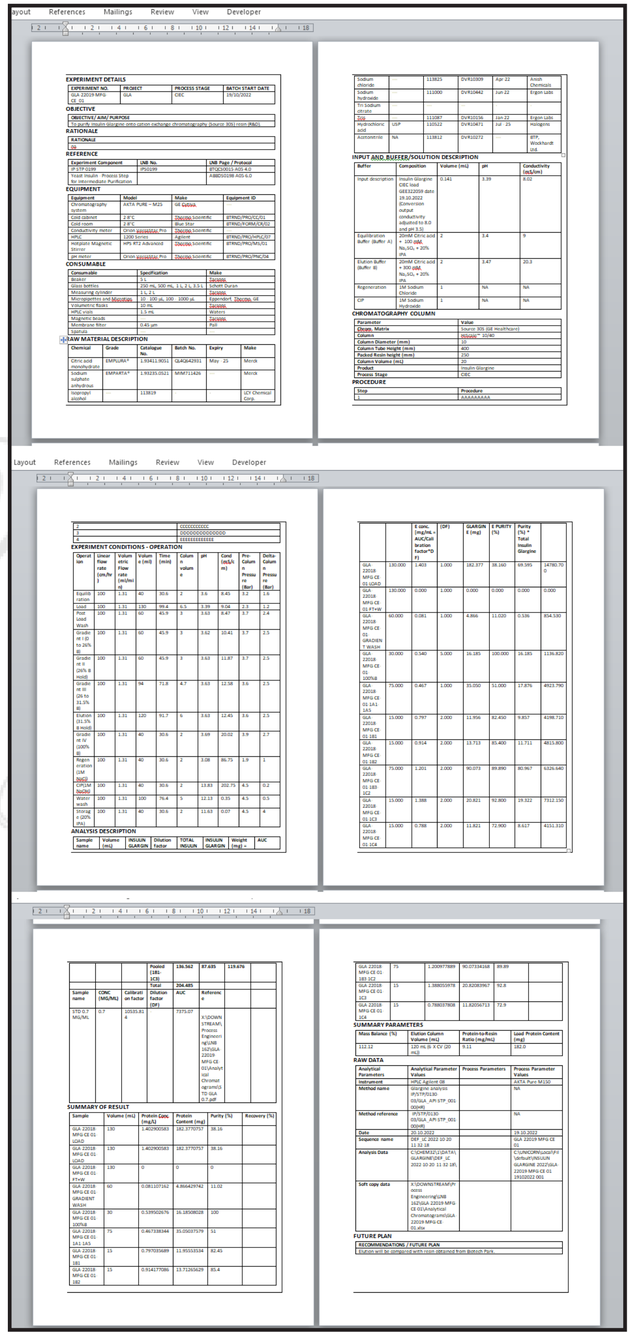
Finally, Excel, Word, or PowerPoint documents for the batch are printed.



Printing Batch



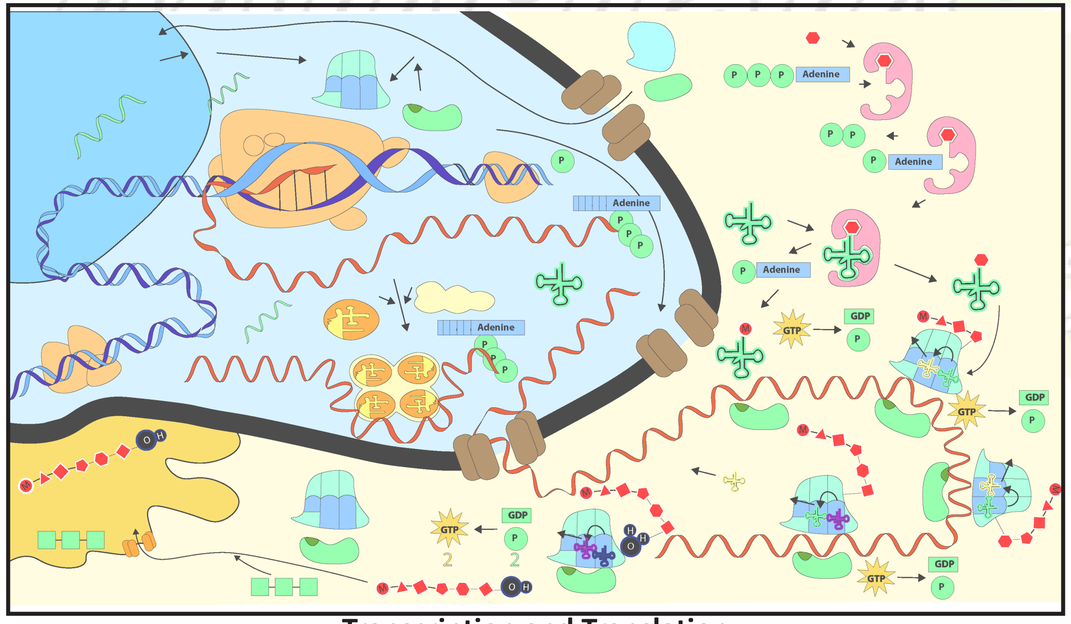
Excel File Printing



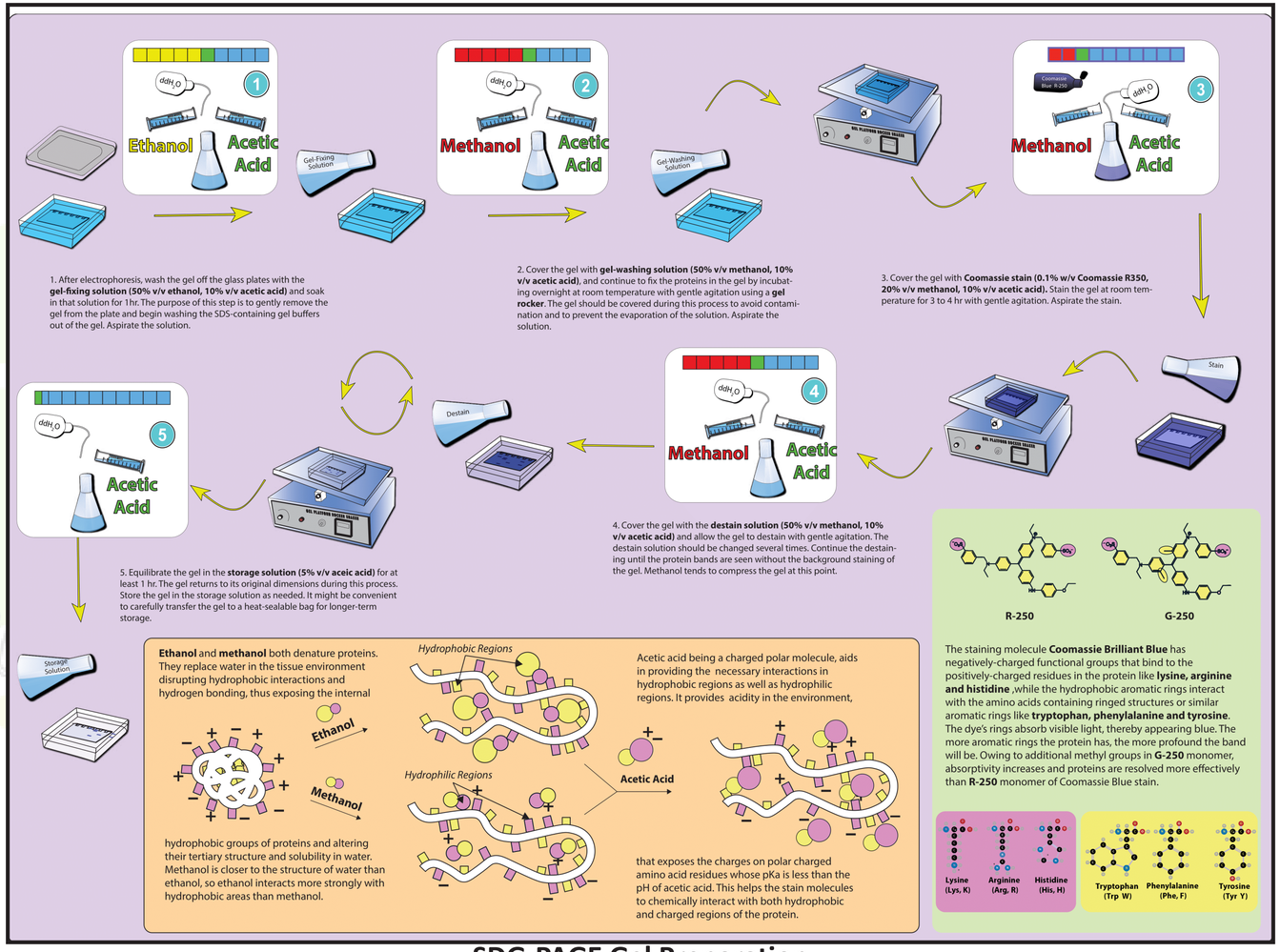
Word Document Printing

PROJECT 4 : ADOBE ILLUSTRATOR PROJECTS

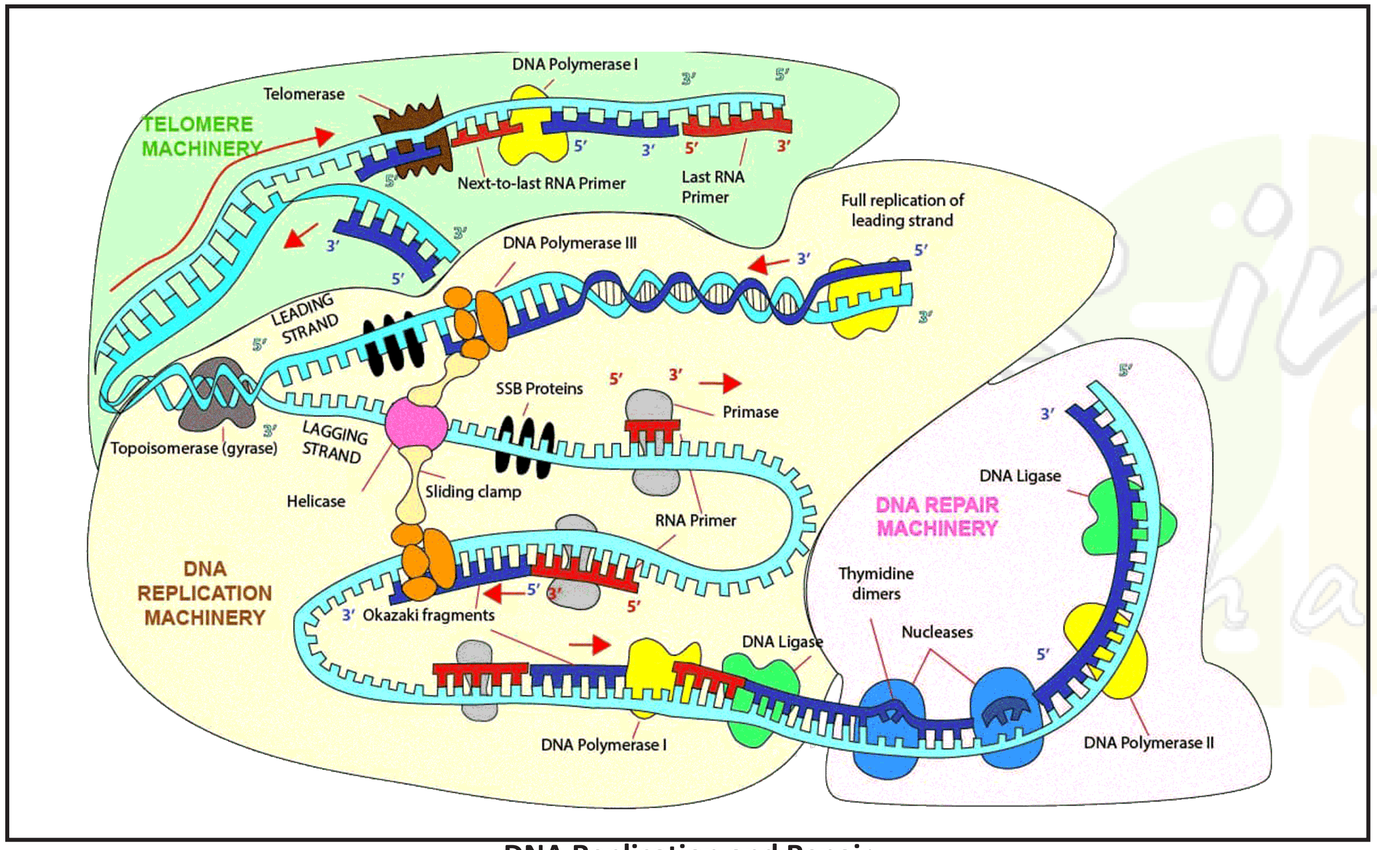
As a part of my website's (https://www.musingsinbiology.com) content, I have created numerous illustrations in biology and beyond using Adobe Illustrator. Some of my creations are shown below.



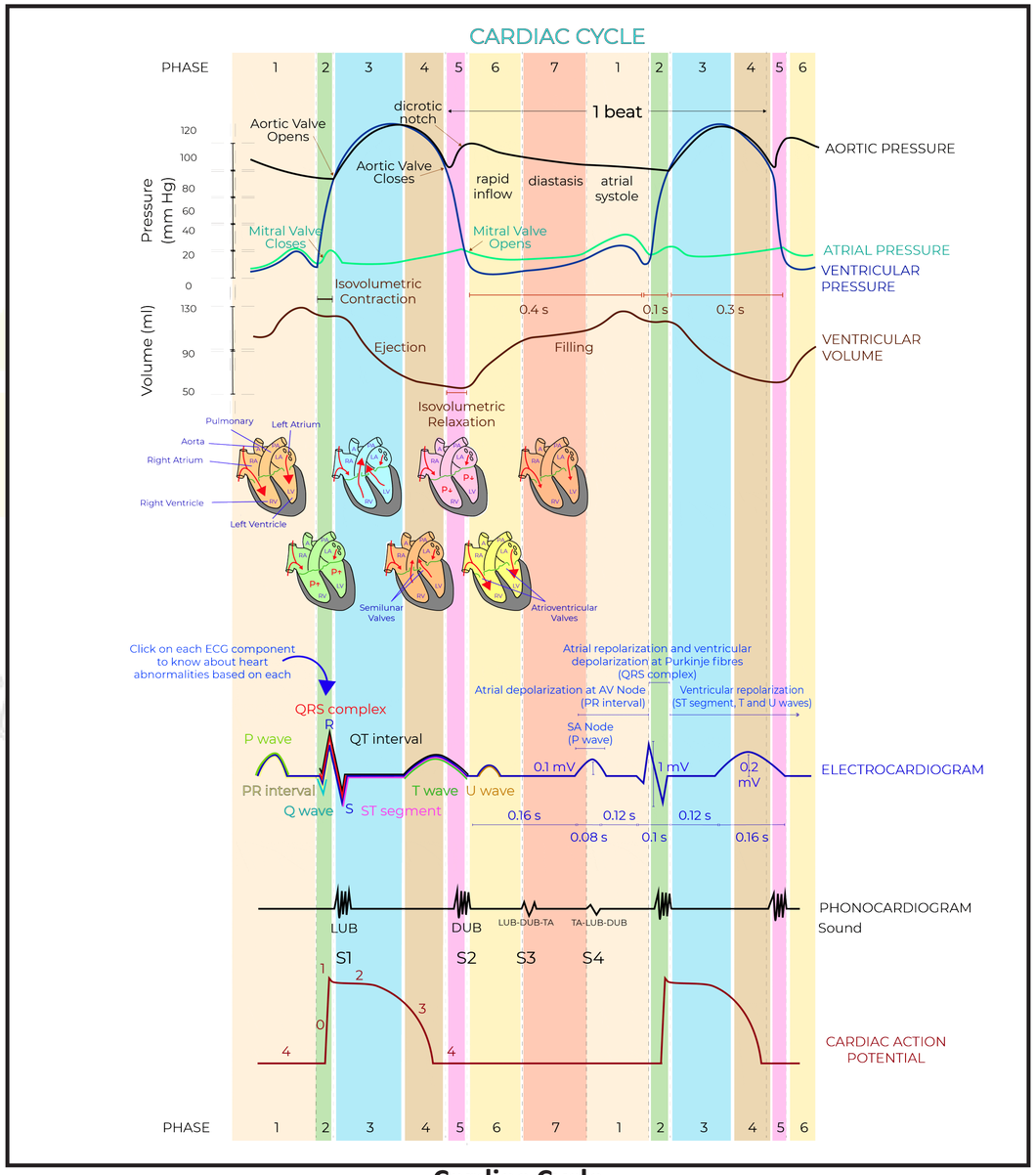
Transcription and Translation



SDG-PAGE Gel Preparation

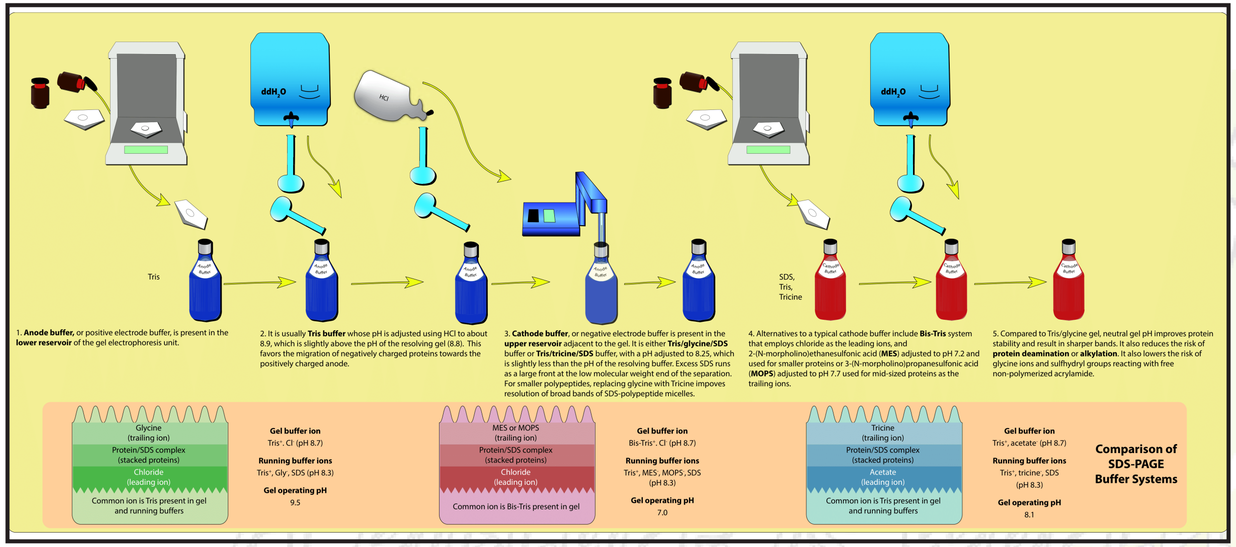


DNA Replication and Repair



Cardiac Cycle

PROJECTS ABOUT FEATURES



1. **Anode buffer,** or positive electrode buffer, is present in the **lower reservoir** of the gel electrophoresis unit.

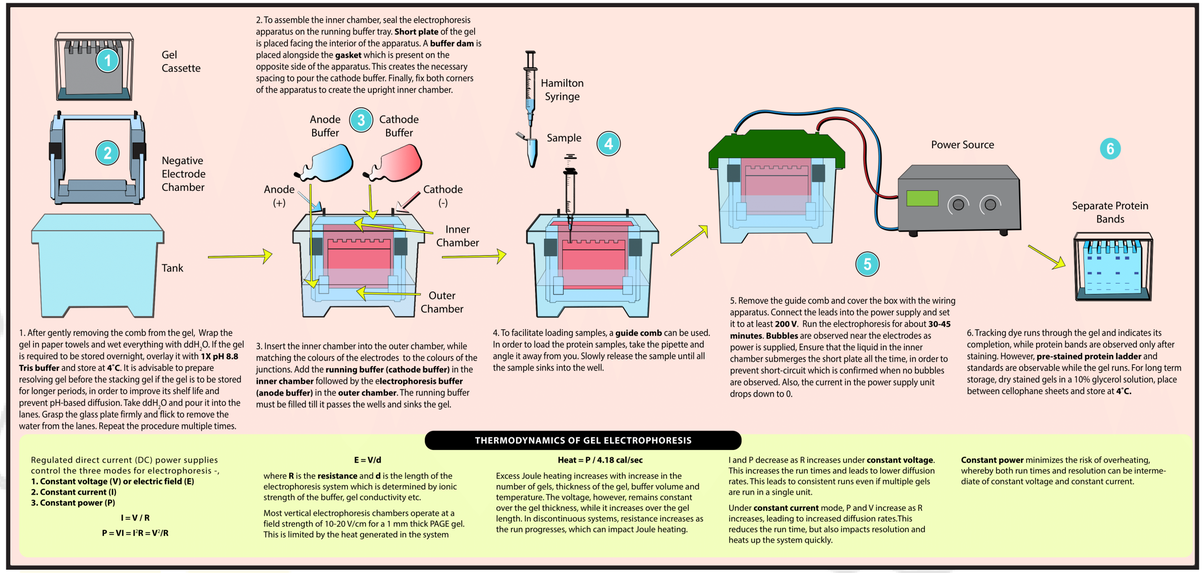
2. It is usually **Tris buffer** whose pH is adjusted using HCl to about 8.9, which is slightly above the pH of the resolving gel (8.8). This favors the migration of negatively charged proteins towards the positively charged anode.

3. **Cathode buffer**, or negative electrode buffer is present in the **upper reservoir** adjacent to the gel. It is either **Tris/glycine/SDS** buffer or **Tris/tricine/SDS** buffer, with a pH adjusted to 8.25, which is slightly less than the pH of the resolving buffer. Excess SDS runs as a large front at the low molecular weight end of the separation. For smaller polypeptides, replacing glycine with Tricine impoves resolution of broad bands of SDS-polypeptide micelles.

4. Alternatives to a typical cathode buffer include **Bis-Tris** system that employs chloride as the leading ions, and 2-(N-morpholino)ethanesulfonic acid (**MES**) adjusted to pH 7.2 and used for smaller proteins or 3-(N-morpholino)propanesulfonic acid (**MOPS**) adjusted to pH 7.7 used for mid-sized proteins as the trailing ions.

5. Compared to Tris/glycine gel, neutral gel pH improves protein stability and result in sharper bands. It also reduces the risk of **protein deamination** or **alkylation**. It also lowers the risk of glycine ions and sulfhydryl groups reacting with free non-polymerized acrylamide.

Various intermediate stages in SDS-PAGE



1. After gently removing the comb from the gel, Wrap the gel in paper towels and wet everything with ddH_2O. If the gel is required to be stored overnight, overlay it with 1**X pH 8.8 Tris buffer** and store at **4˚C**. It is advisable to prepare resolving gel before the stacking gel if the gel is to be stored for longer periods, in order to improve its shelf life and prevent pH-based diffusion. Take ddH_2O and pour it into the lanes. Grasp the glass plate firmly and flick to remove the water from the lanes. Repeat the procedure multiple times.

2. To assemble the inner chamber, seal the electrophoresis apparatus on the running buffer tray. **Short plate** of the gel is placed facing the interior of the apparatus. A **buffer dam** is placed alongside the **gasket** which is present on the opposite side of the apparatus. This creates the necessary spacing to pour the cathode buffer. Finally, fix both corners of the apparatus to create the upright inner chamber.

3. Insert the inner chamber into the outer chamber, while matching the colours of the electrodes to the colours of the junctions. Add the **running buffer (cathode buffer)** in the **inner chamber** followed by the **electrophoresis buffer (anode buffer)** in the **outer chamber**. The running buffer must be filled till it passes the wells and sinks the gel.

4. To facilitate loading samples, a **guide comb** can be used. In order to load the protein samples, take the pipette and angle it away from you. Slowly release the sample until all the sample sinks into the well.

5. Remove the guide comb and cover the box with the wiring apparatus. Connect the leads into the power supply and set it to at least **200 V**. Run the electrophoresis for about **30-45 minutes**. **Bubbles** are observed near the electrodes as power is supplied, Ensure that the liquid in the inner chamber submerges the short plate all the time, in order to prevent short-circuit which is confirmed when no bubbles are observed. Also, the current in the power supply unit drops down to 0.

6. Tracking dye runs through the gel and indicates its completion, while protein bands are observed only after staining. However, **pre-stained protein ladder** and standards are observable while the gel runs. For long term storage, dry stained gels in a 10% glycerol solution, place between cellophane sheets and store at **4˚C.**

THERMODYNAMICS OF GEL ELECTROPHORESIS

Regulated direct current (DC) power supplies control the three modes for electrophoresis -,
1. **Constant voltage (V) or electric field (E)**
2. **Constant current (I)**
3. **Constant power (P)**

$$I = V / R$$

$$P = VI = I^2R = V^2/R$$

$$E = V/d$$

where **R** is the **resistance** and **d** is the length of the electrophoresis system which is determined by ionic strength of the buffer, gel conductivity etc.

Most vertical electrophoresis chambers operate at a field strength of 10-20 V/cm for a 1 mm thick PAGE gel. This is limited by the heat generated in the system

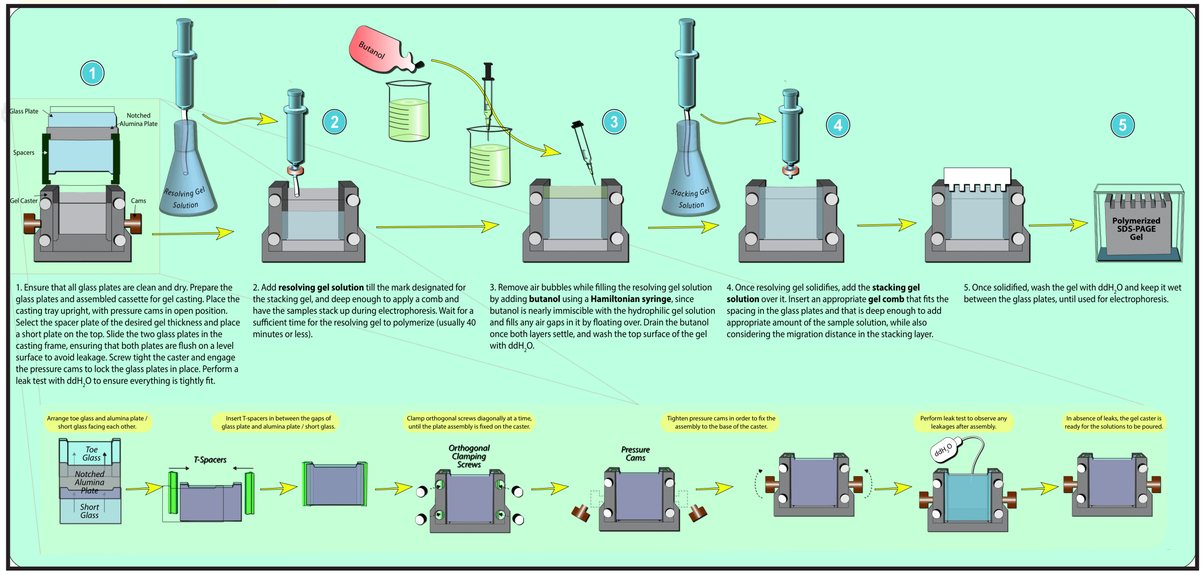
$$Heat = P / 4.18\ cal/sec$$

Excess Joule heating increases with increase in the number of gels, thickness of the gel, buffer volume and temperature. The voltage, however, remains constant over the gel thickness, while it increases over the gel length. In discontinuous systems, resistance increases as the run progresses, which can impact Joule heating.

I and P decrease as R increases under **constant voltage**. This increases the run times and leads to lower diffusion rates. This leads to consistent runs even if multiple gels are run in a single unit.

Under **constant current** mode, P and V increase as R increases, leading to increased diffusion rates.This reduces the run time, but also impacts resolution and heats up the system quickly.

Constant power minimizes the risk of overheating, whereby both run times and resolution can be interme-diate of constant voltage and constant current.



1. Ensure that all glass plates are clean and dry. Prepare the glass plates and assembled cassette for gel casting. Place the casting tray upright, with pressure cams in open position. Select the spacer plate of the desired gel thickness and place a short plate on the top. Slide the two glass plates in the casting frame, ensuring that both plates are flush on a level surface to avoid leakage. Screw tight the caster and engage the pressure cams to lock the glass plates in place. Perform a leak test with ddH_2O to ensure everything is tightly fit.

2. Add **resolving gel solution** till the mark designated for the stacking gel, and deep enough to apply a comb and have the samples stack up during electrophoresis. Wait for a sufficient time for the resolving gel to polymerize (usually 40 minutes or less).

3. Remove air bubbles while filling the resolving gel solution by adding **butanol** using a **Hamiltonian syringe**, since butanol is nearly immiscible with the hydrophilic gel solution and fills any air gaps in it by floating over. Drain the butanol once both layers settle, and wash the top surface of the gel with ddH_2O.

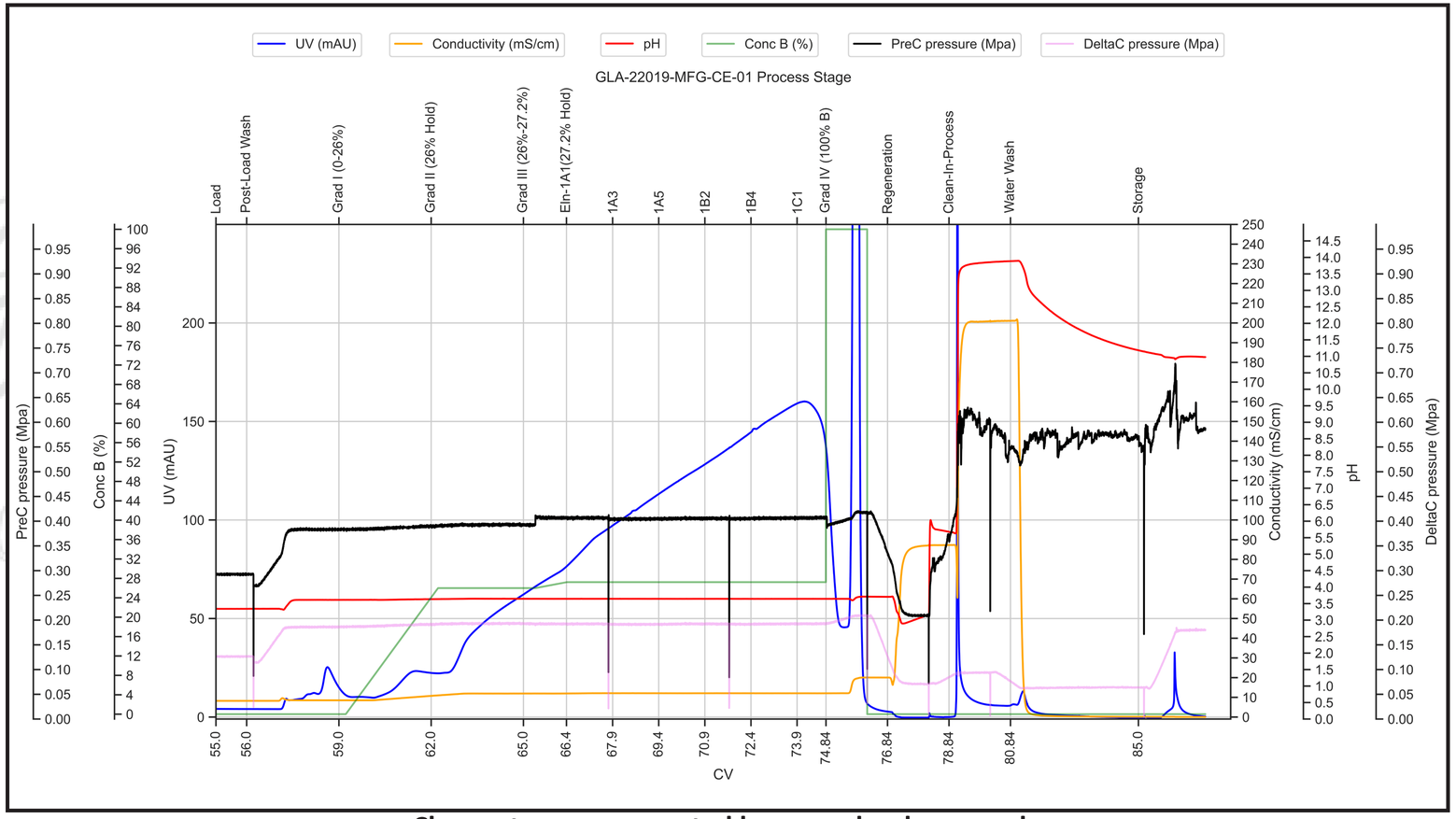
4. Once resolving gel solidifies, add the **stacking gel solution** over it. Insert an appropriate **gel comb** that fits the spacing in the glass plates and that is deep enough to add appropriate amount of the sample solution, while also considering the migration distance in the stacking layer.

5. Once solidified, wash the gel with ddH_2O and keep it wet between the glass plates, until used for electrophoresis.

PROJECT 5 : PYTHON BASED CHROMATOGRAM VISUALIZATION TOOL

This project introduces a Python tool designed to facilitate the visualization of chromatogram data. Chromatography, a technique widely used in biotechnology, generates data that can be complex and time-consuming to analyze. The tool aims to simplify this process by offering an interactive and user-friendly approach.

A number of challenges are faced by process scientists while working with chromatography data, such as the need to create multiple graphs and manually correlate various parameters. Python can be used for generating comprehensive plots that include all relevant parameters without the need for multiple graphs, thus saving time and effort.



Chromatogram generated by procedural approach.

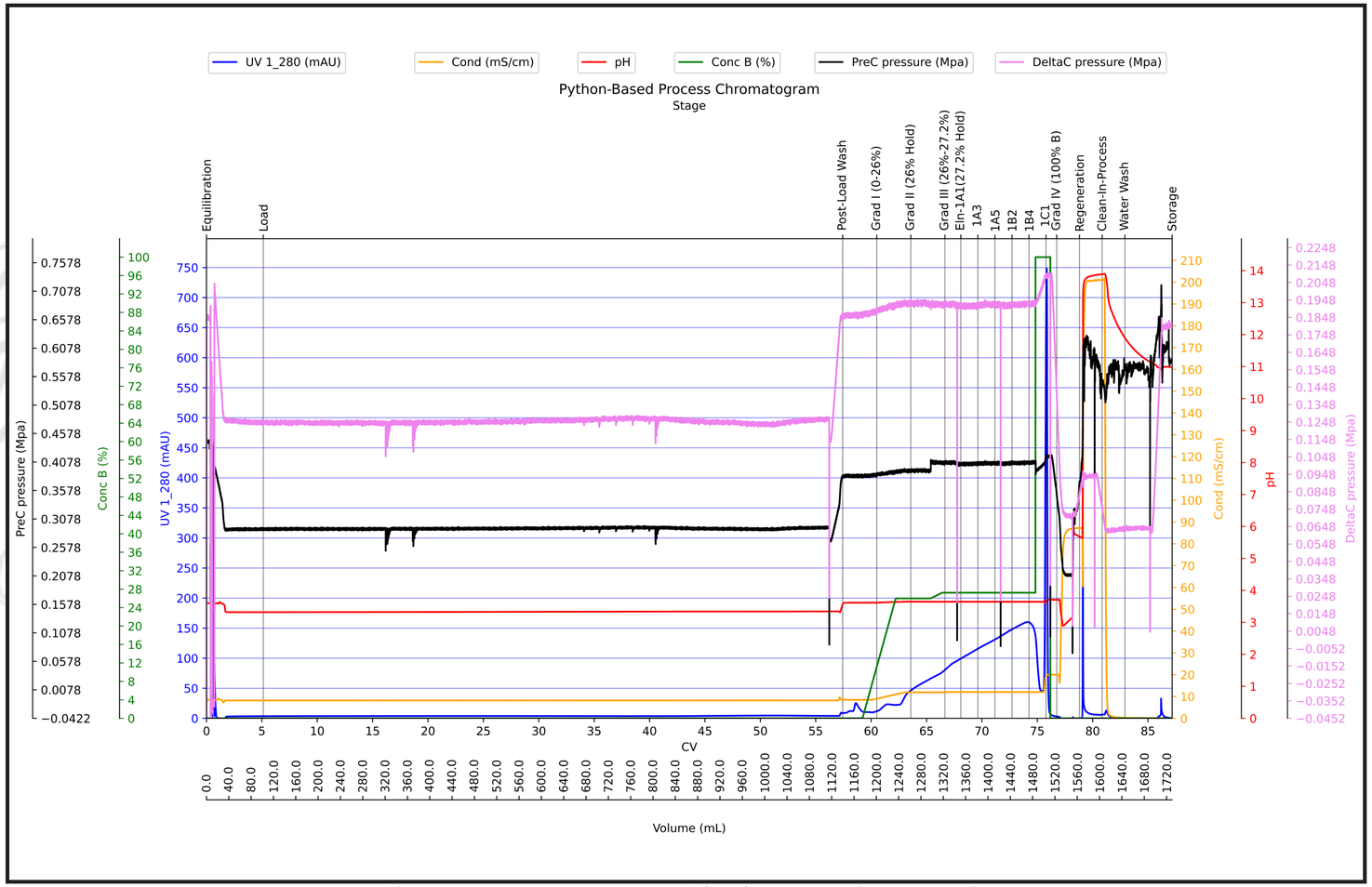
The process of generating chromatogram plots using Python is outlined in the following steps:

1. Obtain cleaned Excel or CSV files containing the chromatograph data. Organize the data correctly, ensuring proper calculations, and labeling of the parameters.

2. Set up a Jupyter Notebook, which involves installing Python and the Anaconda Data Science Platform. The instructions guide the user on creating a new Jupyter Notebook file and installing the necessary Python libraries.

3. Write the relevant Python code within the Jupyter Notebook. It covers reading the data from the Excel or CSV files using the pandas library and creating separate DataFrames for different data sources. The text emphasizes the importance of matching and renaming column names to ensure consistency across DataFrames.

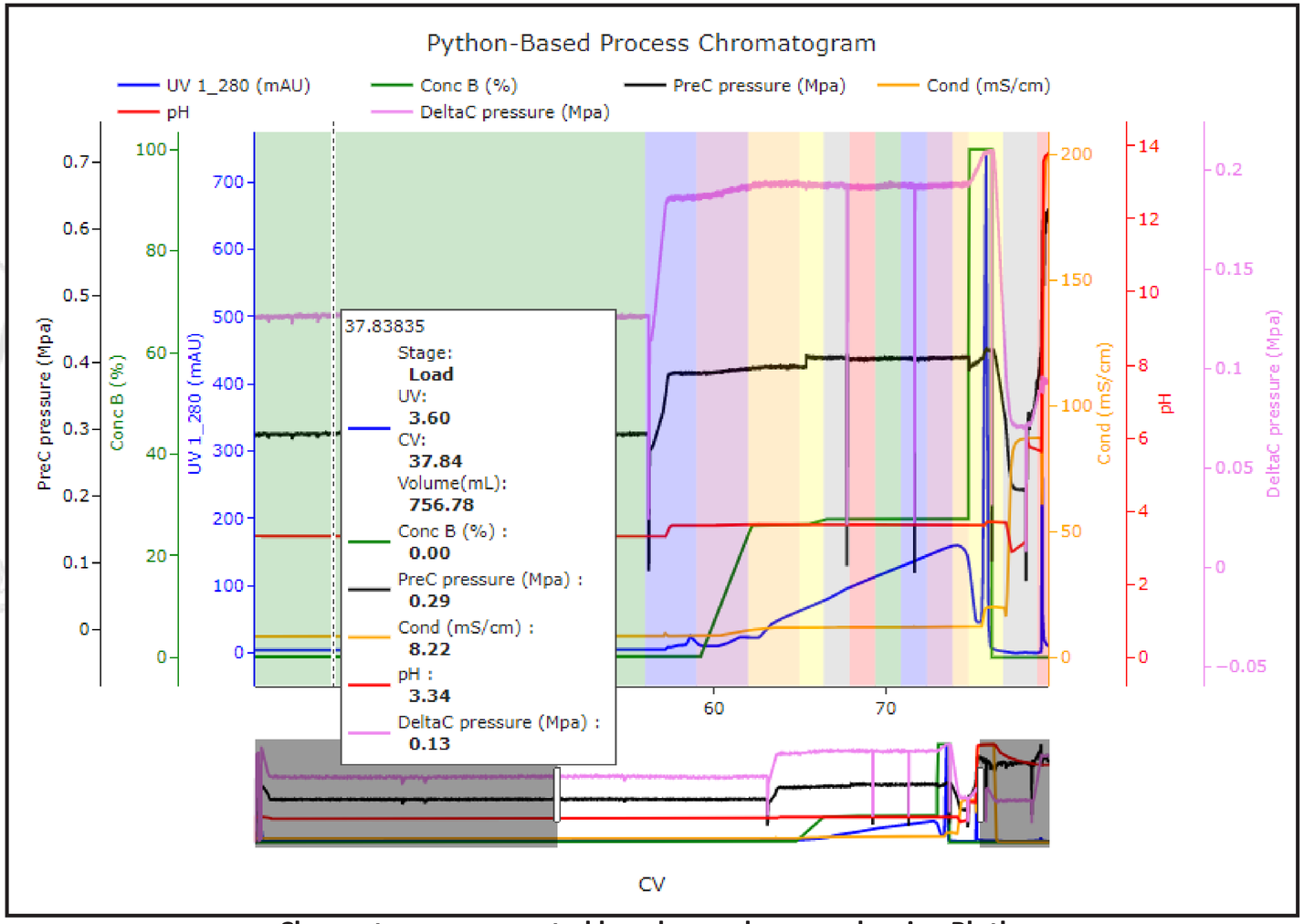
Various data manipulations and analyses can then be performed using the created DataFrames. The text explains how to calculate column volume (CV) values and incorporate additional information, such as process stages, into the main DataFrame.



Chromatogram generated by functional approach.

I explain three different approaches to generate chromatogram plots: ***procedural***, ***functional***, and ***advanced using Plotly.*** Each approach is compared to a relevant analogy to aid understanding. The procedural approach involves hard-coding plot parameters and offers the most flexibility and control but may require changing the code for specific plot requirements. The functional approach focuses on creating reusable functions that perform specific tasks and strike a balance between flexibility and control. Finally, the advanced approach utilizing Plotly is highlighted as the most flexible and convenient method, offering a wide range of customization options for chromatogram plots.

In conclusion, the Python tool presented in the text aims to simplify the visualization of chromatogram data by providing an interactive and efficient approach. It guides users through the process of data preparation, setting up a Jupyter Notebook, writing Python code, and selecting an appropriate approach for generating chromatogram plots.



Chromatogram generated by advanced approach using Plotly.

PROJECT 6 : STATISTICAL ANALYSIS PROJECTS

Project 6.1 : Statistical analysis of alternate vendor experiments for TBEE used in conversion of pro-insulin to insulin ester.

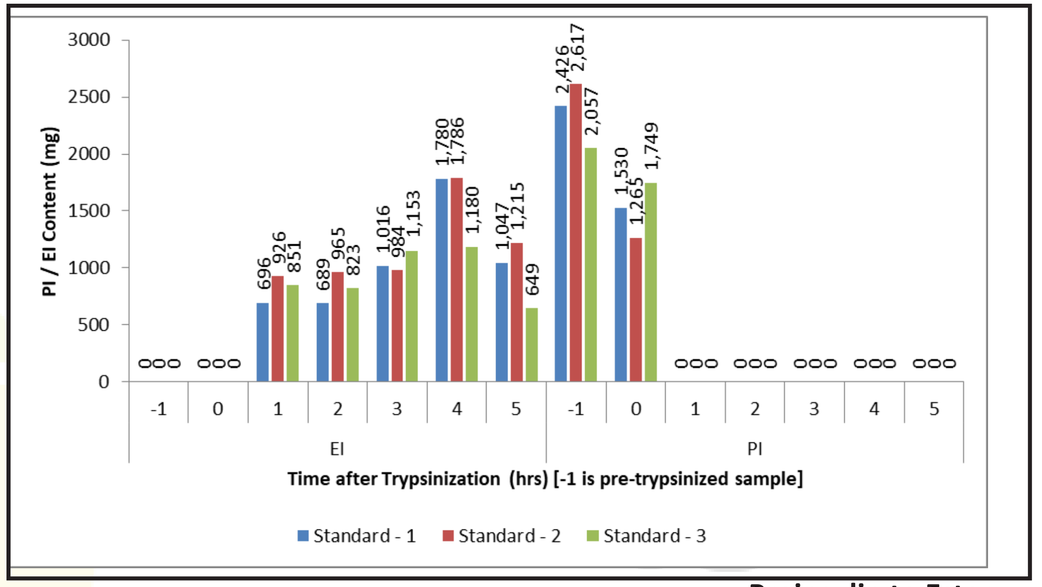
Four experiments were conducted using pro-insulin wet samples obtained from Biotech Park. The experiments involved testing different samples of TBEE (Tris-HCl buffer with Bovine Serum Albumin and EDTA) and analyzing the conversion efficiency of pro-insulin to insulin ester. The experiments aimed to compare the results of using standard TBEE samples with alternate vendor samples and recovered TBEE samples. The data was analyzed using statistical modeling and profiling techniques.

The results showed that experiments 2 and 4 achieved the highest conversion efficiency (98.5%) using standard TBEE, as predicted by the modeling. Experiments 3 and 4 exhibited comparable growth rates (-0.44 and -0.40 respectively) with minimal standard error, indicating consistency. Experiment 4 demonstrated the maximum consistency compared to the average mean.

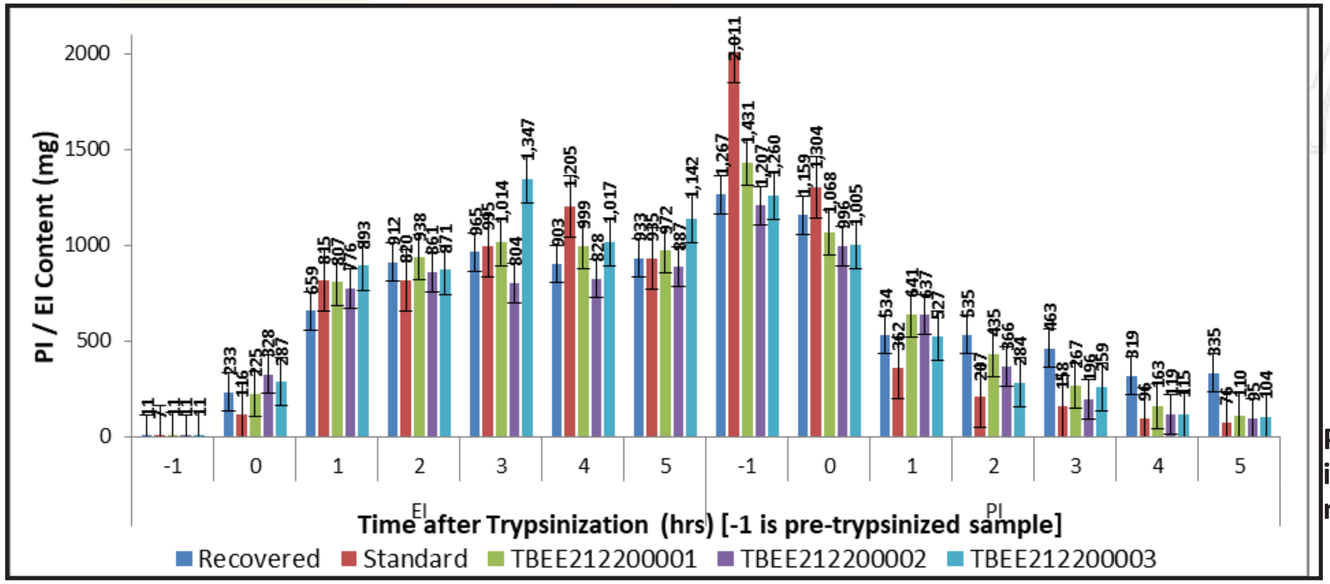
Equivalence tests were conducted, and experiment 4 served as the reference. The tests indicated equivalence between the estimates of experiments 3 and 4, confirming consistency between the two.

The prediction profiler was used to optimize the conversion reaction efficiency, and experiment 2 achieved a maximum conversion efficiency of 98.51% at 5 hours post-trypsinization for standard TBEE samples. Experiments 3 and 4 were averaged and predictively modeled, yielding high correlations between the real and predicted models.

Based on the Lorentzian Peak model, sample TBEE212200002 contained the minimum amount of pro-insulin (145.67 mg) at approximately 5 hours post-trypsinization.



Proinsulin to Ester insulin content for standard experiment.

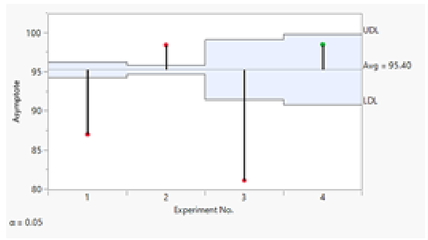


Proinsulin to Ester insulin content for main experiment.

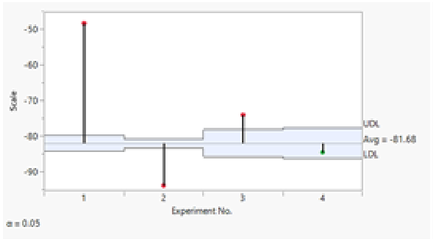
Sample TBEE212200003 produced the maximum amount of insulin ester (1.33 g) with a maximum desirability of 0.885.

The Cell Growth 4-parameter model provided the best fit to the conversion efficiency data, with a high correlation (R-square value of 99.91%) between the actual and predicted models. Sample TBEE212200003 exhibited the maximum conversion efficiency (91.93%) at 2.08 hours post-trypsinization.

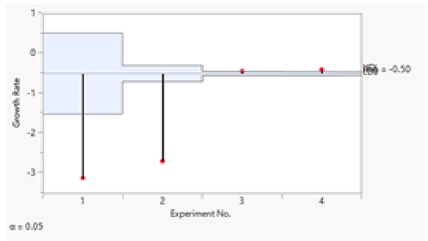
In conclusion, the conversion efficiency achieved with TBEE from the alternate vendor was similar to that of the in-house vendor. The alternate vendor samples generally showed a greater increase in conversion efficiency compared to standard TBEE, while the comparison with recovered TBEE was inconclusive. Specifically, sample TBEE212200003 exhibited a faster rise in conversion efficiency with an overall increase in insulin ester content. It is important to consider human and experimental errors when interpreting these results.



Analysis of Means for Asymptote Estimates



Analysis of Means for Scale Estimates



Analysis of Means for Growth Rate Estimates

Figure 5.3.1-2: Analysis of Means (ANOM) report for experiments based on predictive polynomial modelling for standard TBEE samples

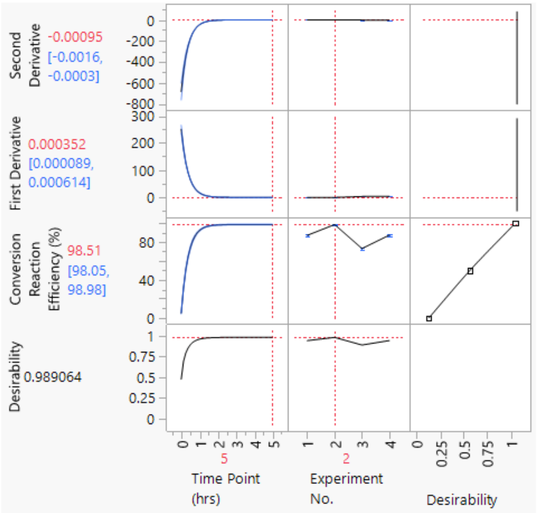
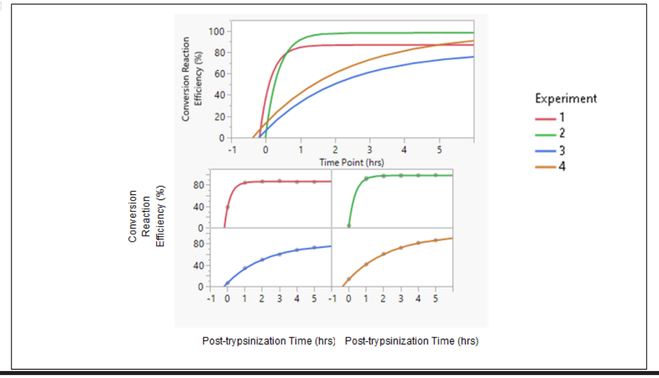


Figure 5.3.1-4: Prediction Profiler for the exponential 3P model for standard TBEE samples

Table 5.3.1-1: Model comparison for standard TBEE samples

Model	AICc	BIC	SSE	MSE	RMSE	R-Square
Exponential 3P	113.10382	117.14412	14.963241	0.6234684	0.7896001	0.9995858



Statistical analysis performed using JMP software

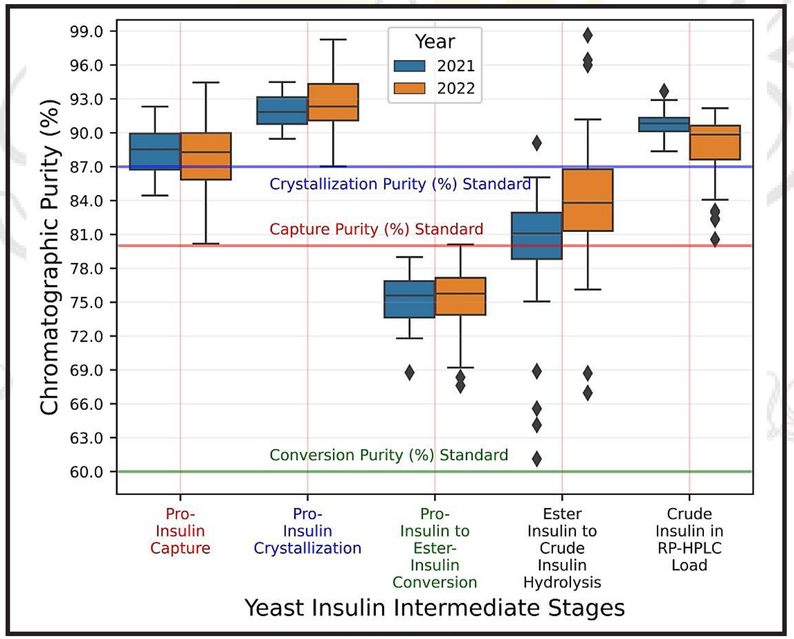
PROJECT 6 : STATISTICAL ANALYSIS PROJECTS

Project 6.2 : Statistical analysis of production batch data of Insulin Human (yeast)

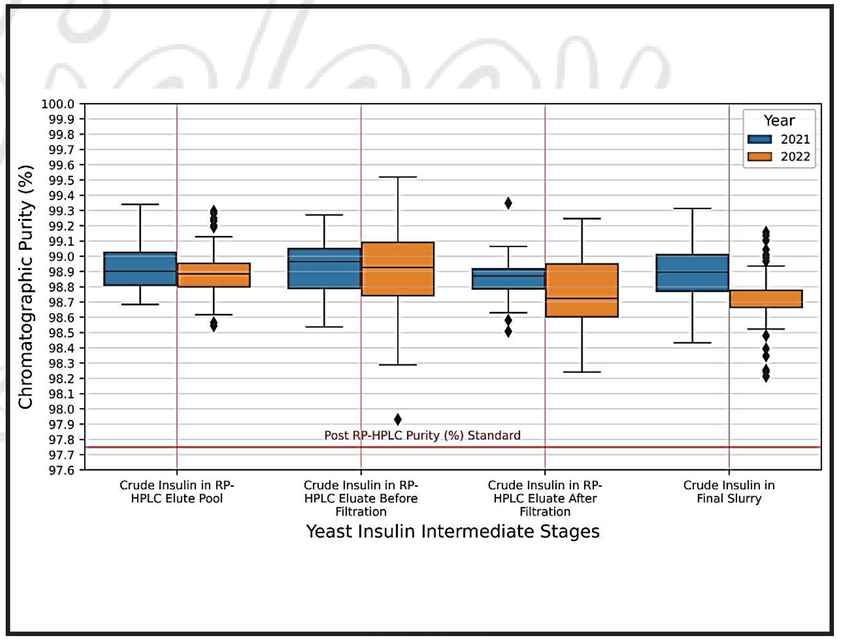
A comparison was made between the chromatographic purity at different intermediate stages in the years 2021 and 2022 for the Human Insulin production process.

Box plots were used to analyze the trends in purity profiles before and after RP-HPLC stages. The results showed that the average purity remained similar in the intermediate stages until EI isolation. However, the average purity increased in the EI to CI hydrolysis stage in the current year. Conversely, the crude insulin (CI) isolation stage exhibited lower average purity in 2022 compared to the previous year. Batch-to-batch variability was higher in all stages in 2022.

Comparisons between stages revealed that the conversion of pro-insulin (PI) to EI exhibited the lowest chromatographic purity (around 75%) among all stages, while the PI crystallization step showed the highest purity among pre-RP-HPLC steps. In the post-RP-HPLC stages, lower average purity was observed in solubilized CI and final CI slurry in 2022 batches compared to 2021 batches. Batch-to-batch variability was also higher in 2022 for post-RP-HPLC stages.



Pre RP-HPLC Stage



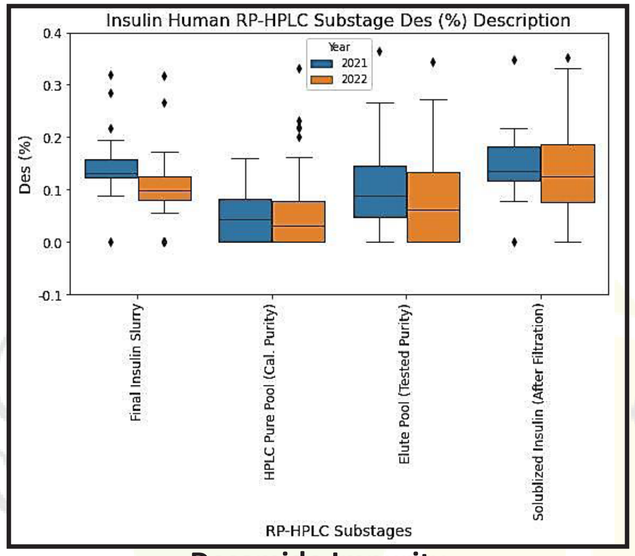
Post RP-HPLC Stage

nbahar@musingsinbiology.com

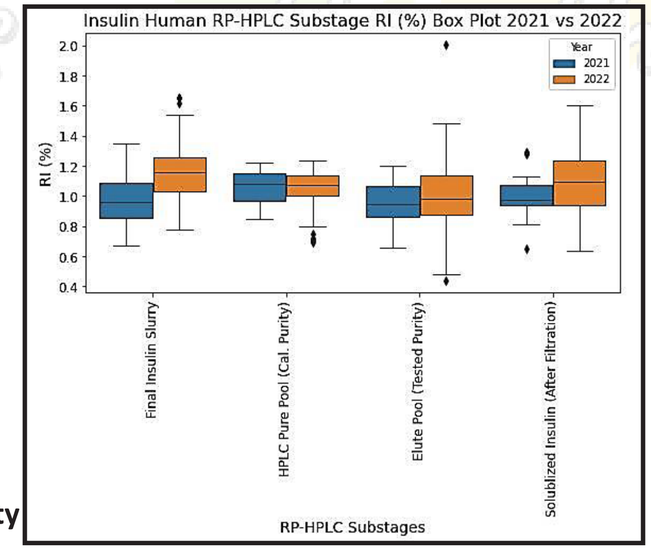
Desamido impurity data for post-RP-HPLC stages indicated a reduction in impurity levels in 2022 compared to 2021 in solubilized zinc insulin and final insulin slurry. However, the overall RI content in the post-RP-HPLC stages showed increased levels in solubilized zinc insulin and final insulin slurry, with higher batch-to-batch variability in 2022.

In RP-HPLC Load impurity analysis, related impurities (RI) increased in 2022 compared to 2021. While the RI composition at RRT 0.47 remained relatively constant, the RRT 0.70 impurity decreased, and the RRT 0.85 impurity increased in 2022. The appearance of additional RI impurities contributed to the overall increment in RI.

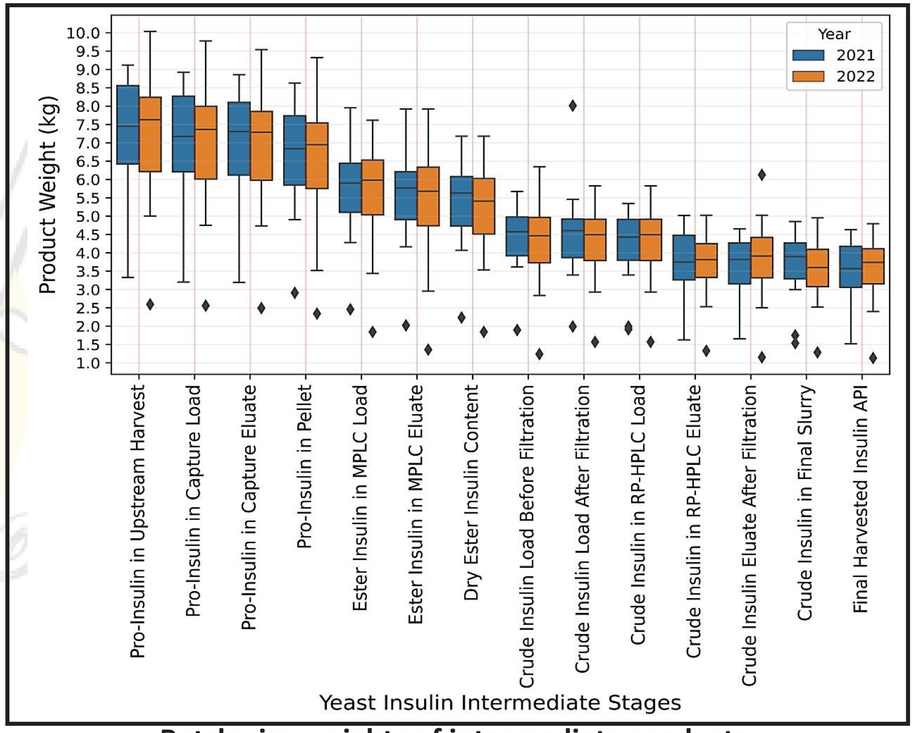
Batchwise weights of intermediate products (Pro-insulin, Ester-insulin, Crude-insulin) were averaged to compare yearly data. The analysis showed an average increase in intermediate product content from PI capture to PI to EI conversion in 2022, while EI and CI isolation stages exhibited a decline, suggesting possible improvement strategies for the later stages.



Desamido Impurity



RI Impurity



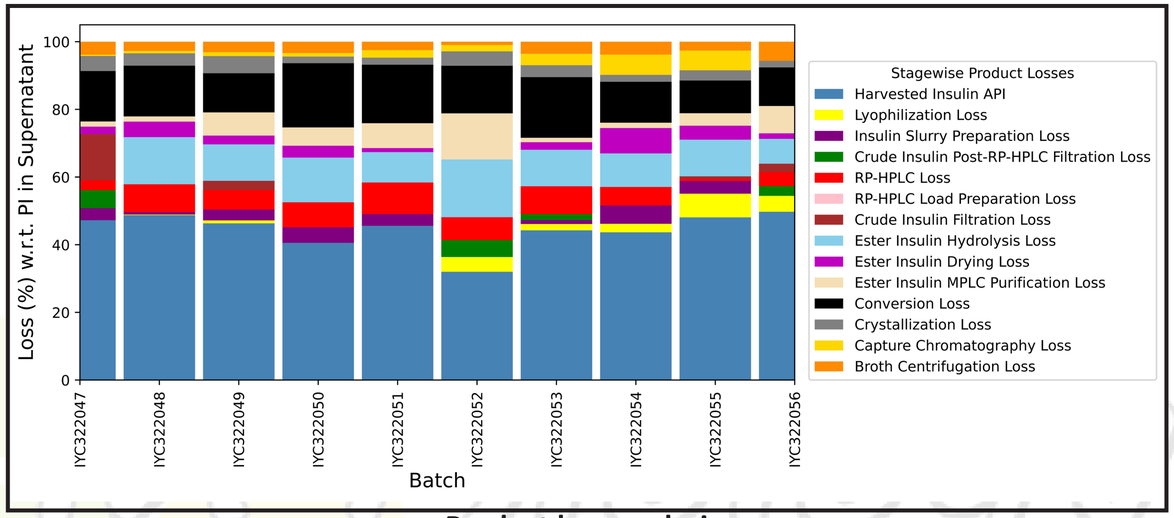
Batchwise weights of intermediate products.

nbahar@musingsinbiology.com

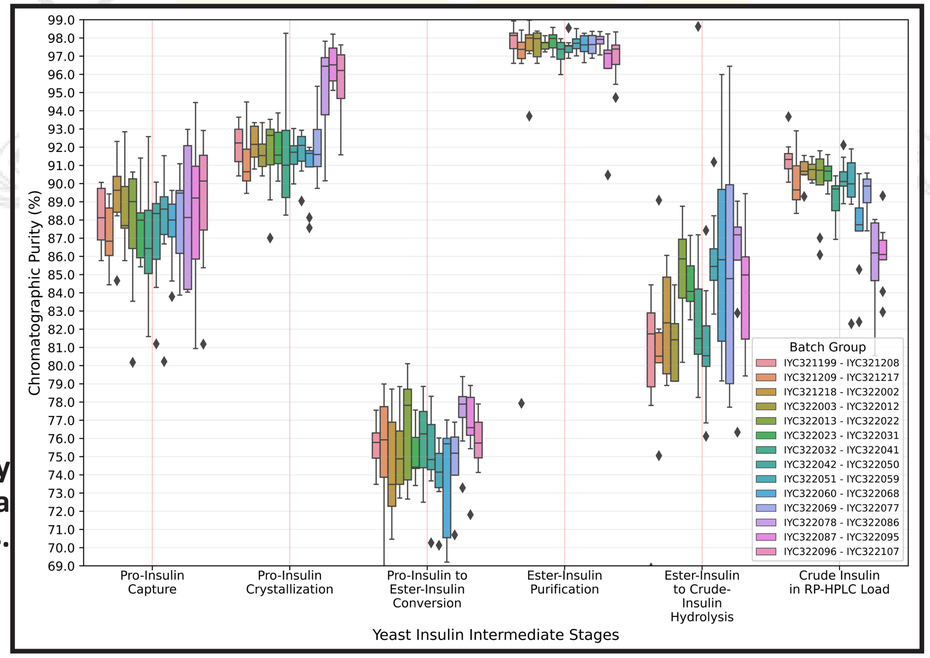
The product loss analysis indicated that the maximum loss occurred during the conversion of pro-insulin to ester-insulin, followed by hydrolysis of ester-insulin to crude insulin. Recent batches also showed increased loss during ester insulin MPLC purification. Losses during RP-HPLC decreased, but losses during lyophilization and capture chromatography increased.

Overall API yield ranged between 45-50% compared to pro-insulin from supernatant. Bootstrapping was performed to determine confidence intervals for mean values, and 90%, 95%, and 99% confidence intervals were calculated.

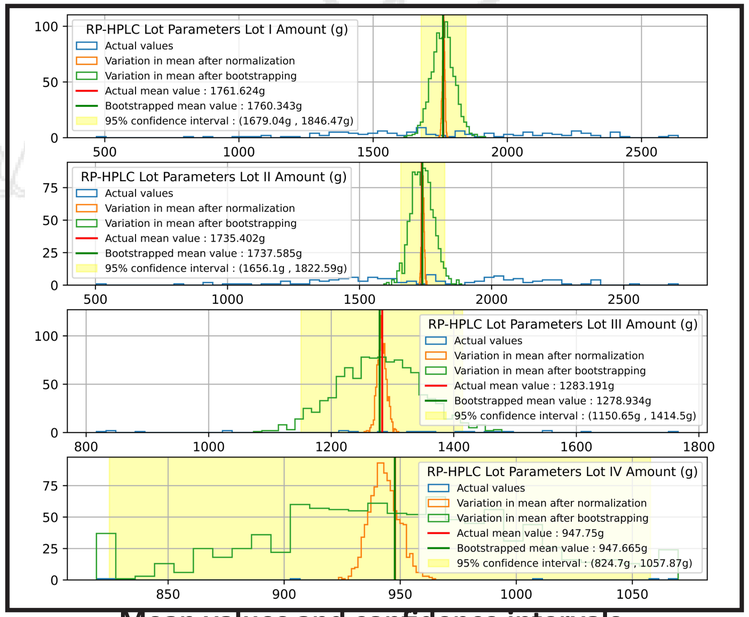
OLS regression analysis revealed that the reduction in chromatographic purity in post-RP-HPLC stages was attributed to an increase in related impurities rather than Desamido content. There was a slight negative correlation between RI at RRT 0.85 and overall RI in HPLC load, requiring further investigation.



Product loss analysis.



Chromatographic Purity grouped by batches of 10 at a time over past 2 years.



Mean values and confidence intervals.

PROJECT 6 : STATISTICAL ANALYSIS PROJECTS

Project 6.3 : Statistical evaluation of mock pools during fraction collection in Glargine CIEC and Yeast Insulin RP-HPLC stages

Statistical analysis was performed using Python on 52 batches of the RP-HPLC stage in Yeast Insulin Human from 2022 and 2023. A total of 114 lots, including multiple and resampling lots, were analyzed to determine the count of batches within specified pool ranges.

The following steps were followed:

1. Data from different Excel files was consolidated into a single Excel workbook with separate worksheets for each batch and lot.

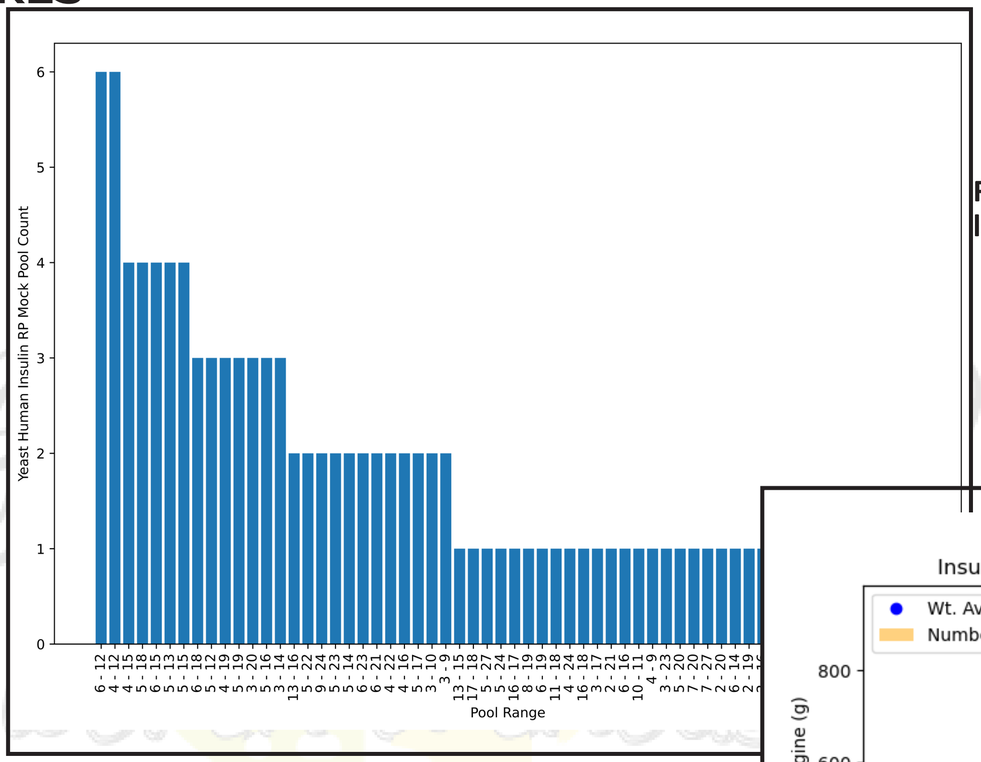
2. A data table was created using the data from each Excel worksheet, organized batch-wise.

3. The fraction column was split into start and end fraction columns to specify the range.

4. For each batch, the minimum start fraction and maximum end fraction were determined and recorded in a pool range column. These ranges were based on the weighted average purity of Yeast Insulin Human, which had to meet specific purity (NLT 97.75%), desamido impurity (NMT 1.0%), and RI (NMT 1.25%) criteria.

5. The counts for different pool ranges were calculated and used for plotting.

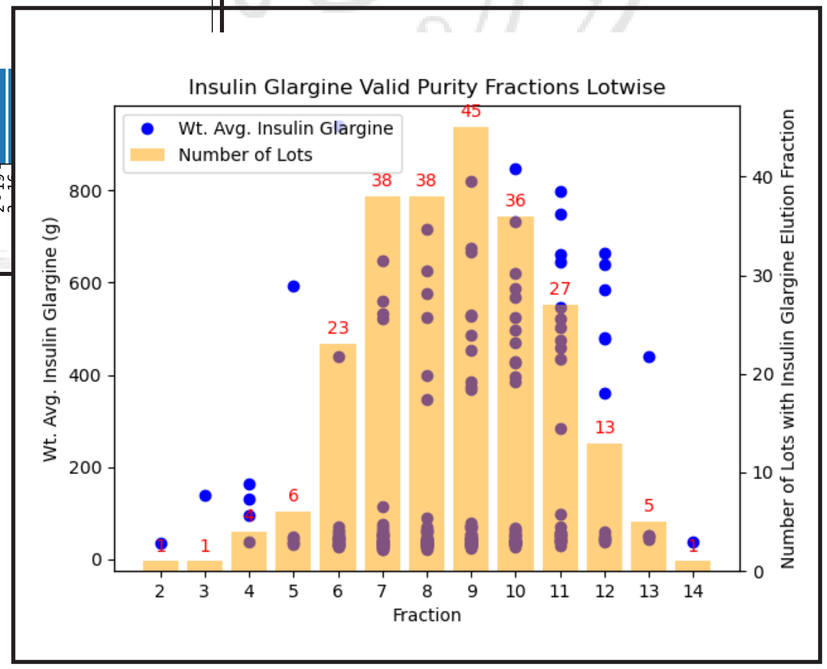
Based on this data, pool range 6 to 12 appears to be a suitable starting point for creating mock pools. Additional fractions can then be added from leading or trailing fractions to refine the pool composition.



A similar mock pool analysis is ongoing for Insulin Glargine CIEC stage.

Frequency of pool ranges by number of lots they occured in.

Average weight of Insulin Glargine per fraction and their histogram of counts.



PROJECT 7 : IMPURITY ANALYSIS OF ANALYTICAL DATA

Project 7.1 : Pivot Table Based Impurity Analysis

This project analyzes analytical impurity data using pivot tables in Excel, Python, and Power BI. By using these tools, I intend to simplify the identification and visualization of impurity data present in analytical samples.

The project involves several steps:

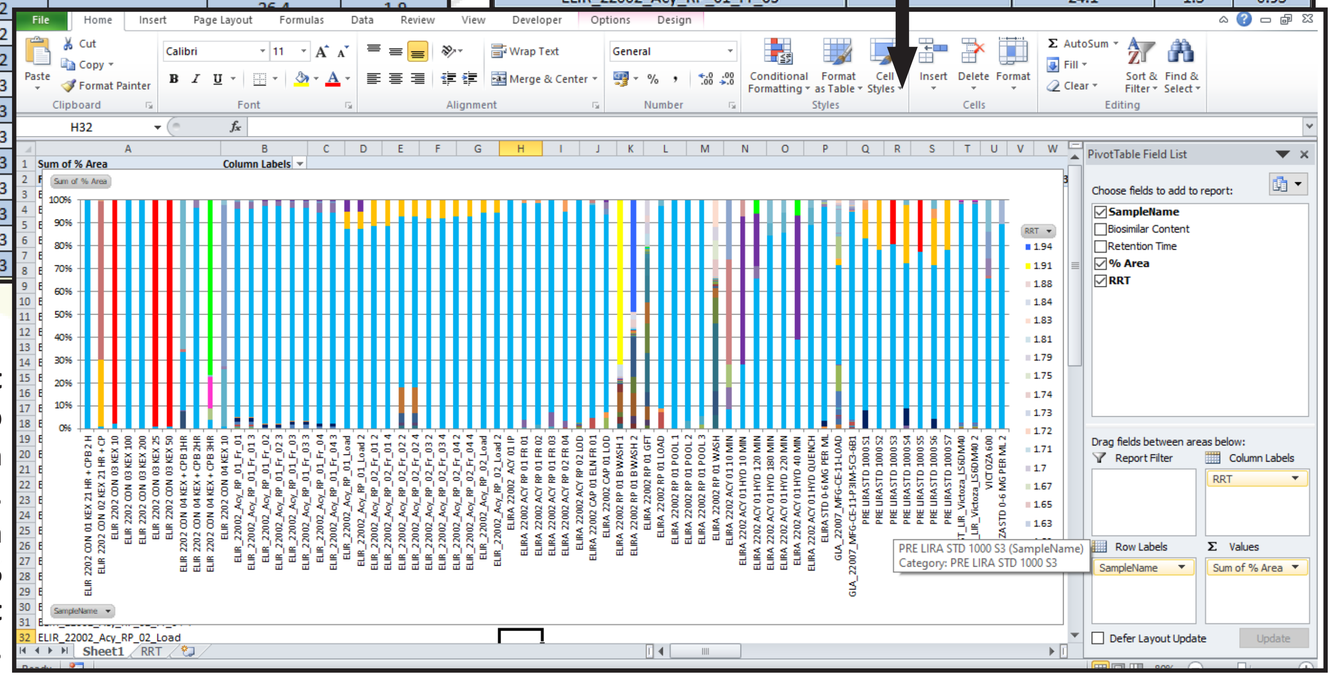
1. Collecting Relevant Data: The relevant data, including sample names, biosimilar content, retention times, and area under the curve (AUC), is obtained from the chromatography / analytical system and transferred to an Excel file.

2. Cleaning the Data: The data is processed and cleaned to remove anomalies, errors, or inconsistencies. Washable components are excluded and duplicate sample names are handled appropriately. The cleaned data is organized as an Excel table.

3. Determining Relative Retention Times: Python programming is used to calculate the relative retention times (RRT) of impurities. Each impurity's retention time

SampleName	Biosimilar Content	Retention Time	% Area
ELIR_22002_Acy_RP_01_Fr_01		21.7	0.5
ELIR_22002_Acy_RP_01_Fr_01		22.4	0.7
ELIR_22002_Acy_RP_01_Fr_01		23.5	1.5
ELIR_22002_Acy_RP_01_Fr_01		24.1	1.7
ELIR_22002_Acy_RP_01_Fr_01		24.8	0.4
ELIR_22002_Acy_RP_01_Fr_01	Liraglutide	25.4	91
ELIR_22002_Acy_RP_01_Fr_01		26.4	2.9
ELIR_22002_Acy_RP_01_Fr_01		27.3	0.3
ELIR_22002_Acy_RP_01_Fr_01		29.6	0.4
ELIR_22002_Acy_RP_01_Fr_02		23.6	0.6
ELIR_22002_Acy_RP_01_Fr_02		24.2	1.3
ELIR_22002_Acy_RP_01_Fr_02		24.8	0.4
ELIR_22002_Acy_RP_01_Fr_02	Liraglutide	25.4	95.3
ELIR_22002_Acy_RP_01_Fr_02			
ELIR_22002_Acy_RP_01_Fr_02			
ELIR_22002_Acy_RP_01_Fr_02			
ELIR_22002_Acy_RP_01_Fr_03			
ELIR_22002_Acy_RP_01_Fr_03			
ELIR_22002_Acy_RP_01_Fr_03			
ELIR_22002_Acy_RP_01_Fr_03			
ELIR_22002_Acy_RP_01_Fr_03			
ELIR_22002_Acy_RP_01_Fr_03			
ELIR_22002_Acy_RP_01_Fr_03			

SampleName	Biosimilar Conten	Retention Tim	% Are	RRT
ELIR_22002_Acy_RP_01_Fr_01		26.4	2.9	1.04
ELIR_22002_Acy_RP_01_Fr_01		27.3	0.3	1.07
ELIR_22002_Acy_RP_01_Fr_01		29.6	0.4	1.17
ELIR_22002_Acy_RP_01_Fr_02		23.6	0.6	0.93
ELIR_22002_Acy_RP_01_Fr_02		24.2	1.3	0.95
ELIR_22002_Acy_RP_01_Fr_02		24.8	0.4	0.98
ELIR_22002_Acy_RP_01_Fr_02	Liraglutide	25.4	95.3	1
ELIR_22002_Acy_RP_01_Fr_02		26.4	1.9	1.04
ELIR_22002_Acy_RP_01_Fr_02		27.3	0.3	1.07
ELIR_22002_Acy_RP_01_Fr_02		29.5	0.3	1.16
ELIR_22002_Acy_RP_01_Fr_03		21.8	0.2	0.86
ELIR_22002_Acy_RP_01_Fr_03		22.3	0.2	0.88
ELIR_22002_Acy_RP_01_Fr_03		23.6	1.3	0.93
ELIR_22002_Acy_RP_01_Fr_03		24.1	1.5	0.95



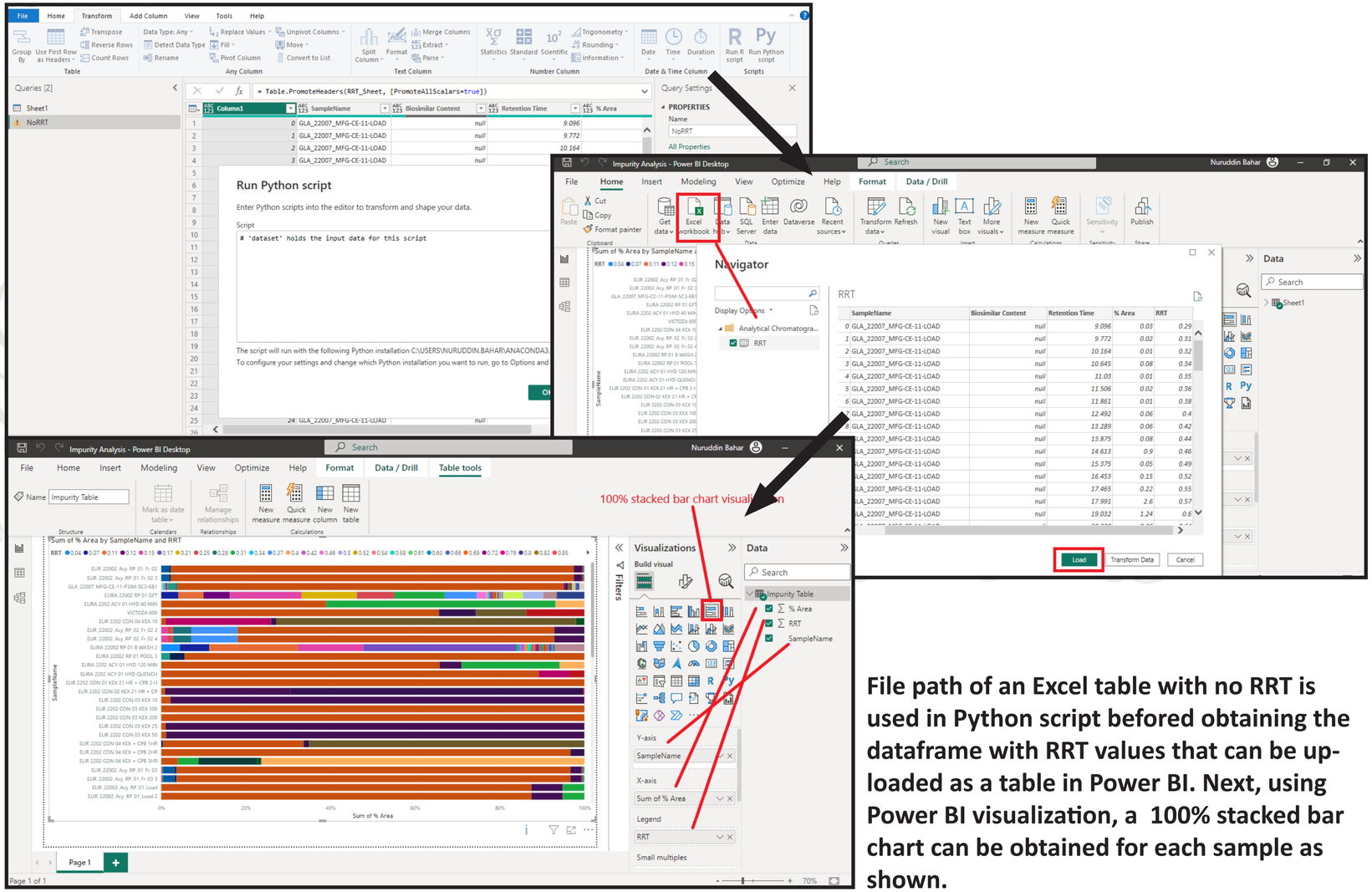
Excel table without RRT is converted to an Excel table with RRT using Python. The table is then used to create 100% stacked pivot chart as shown.

is divided by the retention time of the component of interest to obtain the RRT values.

4. Performing Impurity Analysis: Impurity analysis is performed using either a Python library or Excel Pivot Table. This analysis provides information about the types and quantities of impurities present in the samples, facilitating informed decisions regarding purification processes.

The process involves importing relevant Python libraries, reading the Excel file into a pandas DataFrame, creating lists of biosimilar peaks and sample names, mapping samples to their retention times, calculating RRT values, and saving the updated data to a new Excel file.

Three options for performing impurity analysis are discussed. The first option involves creating a Pivot Table and Pivot Chart in Excel, which can be customized according to user preferences. The second option involves using the pivottablejs library in Python to generate an interactive pivot table and HTML/JavaScript code. The third option uses Power BI to import the cleaned and processed data, perform additional calculations or aggregations if necessary, and create insightful visualizations to explore and analyze the impurity data. These visualizations can include charts, graphs, tables, and other interactive components that provide a comprehensive understanding of the impurity profiles in the chromatographic data.



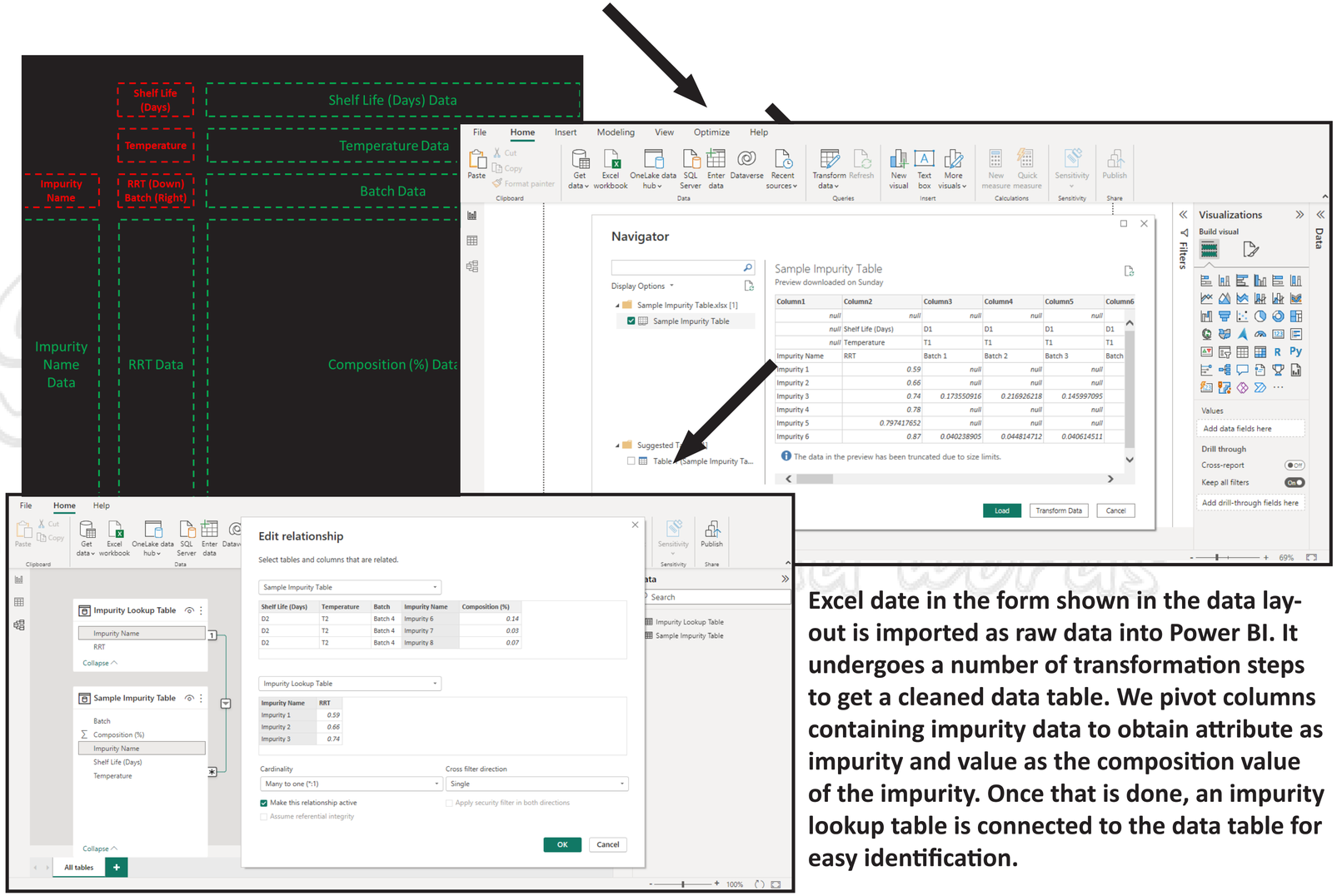
File path of an Excel table with no RRT is used in Python script befored obtaining the dataframe with RRT values that can be uploaded as a table in Power BI. Next, using Power BI visualization, a 100% stacked bar chart can be obtained for each sample as shown.

PROJECT 7 : IMPURITY ANALYSIS OF ANALYTICAL DATA

Project 7.2 : Creating Power BI Impurity Dashboard and Analysis

In this project, the focus is on utilizing Power BI to create an Impurity Dashboard that enables data analysts to gain actionable insights and make data-driven decisions. The Impurity Dashboard is designed to handle a large Excel data table containing information about impurities across various experiments with different temperature and shelf-life conditions.

The Power BI visualizations provide powerful tools for customizing and presenting the data in a way that enhances the understanding of the experiments. Researchers can easily address important questions such as identifying impurities with the highest average composition across experiments, analyzing how impurity compositions change with different shelf-life durations for specific impurities, exploring correlations between different impurities, understanding the general trends of impurity composition by temperature, and identifying impurities that exhibit significant increases with longer shelf-life durations.

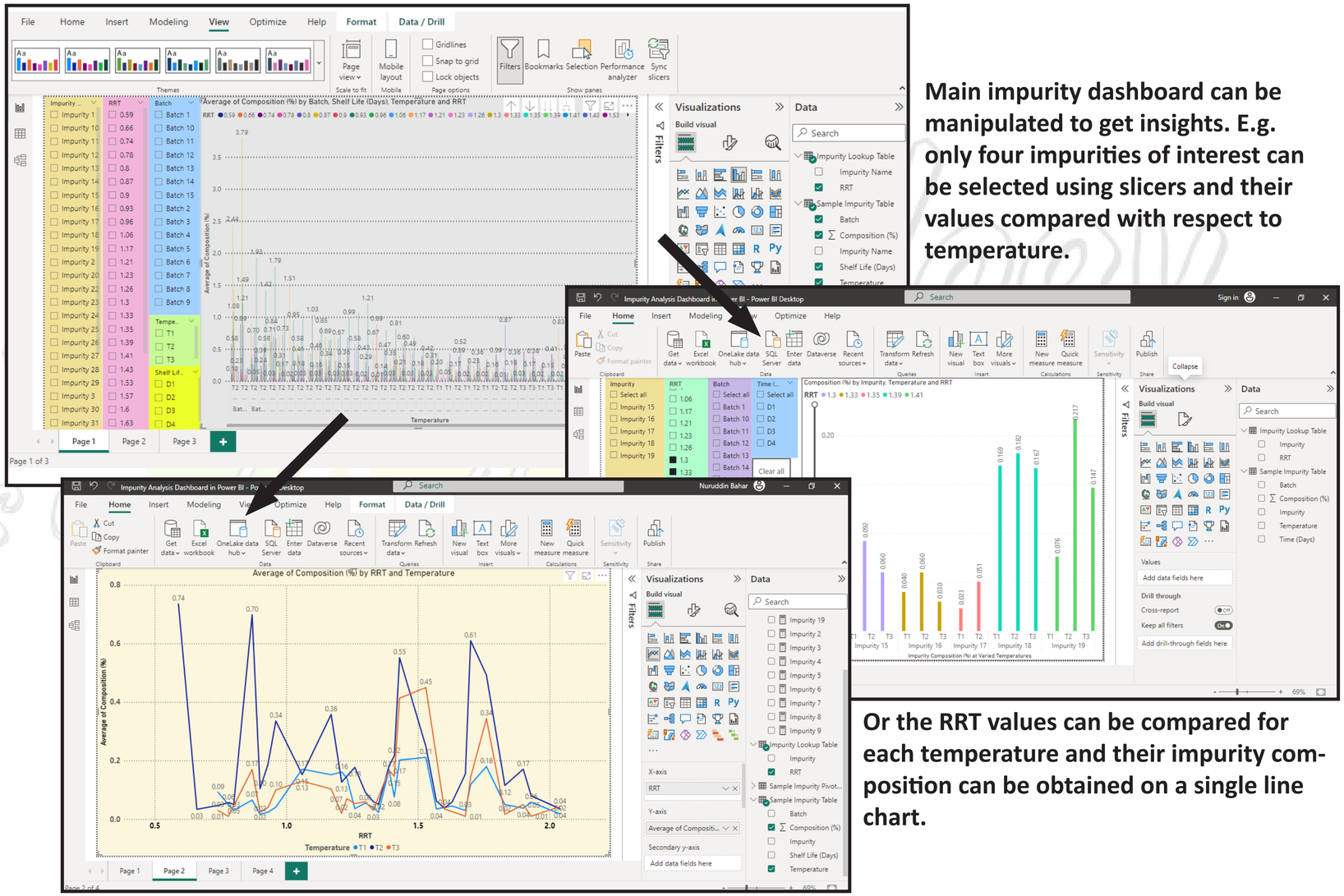


Excel date in the form shown in the data lay-out is imported as raw data into Power BI. It undergoes a number of transformation steps to get a cleaned data table. We pivot columns containing impurity data to obtain attribute as impurity and value as the composition value of the impurity. Once that is done, an impurity lookup table is connected to the data table for easy identification.

To achieve this, the article outlines the steps to restructure the data in Power BI, including loading the data table into Power Query editor, modifying and organizing the table to ensure optimal visualization, and creating the necessary relationships between tables.

Once the data is properly structured, researchers can create an Impurity Dashboard in Power BI. The dashboard includes slicers for easy data filtering, a clear all slicers button to reset the selections, and a clustered column chart to visualize the impurity data.

To answer the specific business questions, various visualizations are utilized. For example, a Q&A visualization is employed to identify the impurity with the highest average composition across experiments. Line charts help analyze the composition changes of impurities for different shelf-life durations.



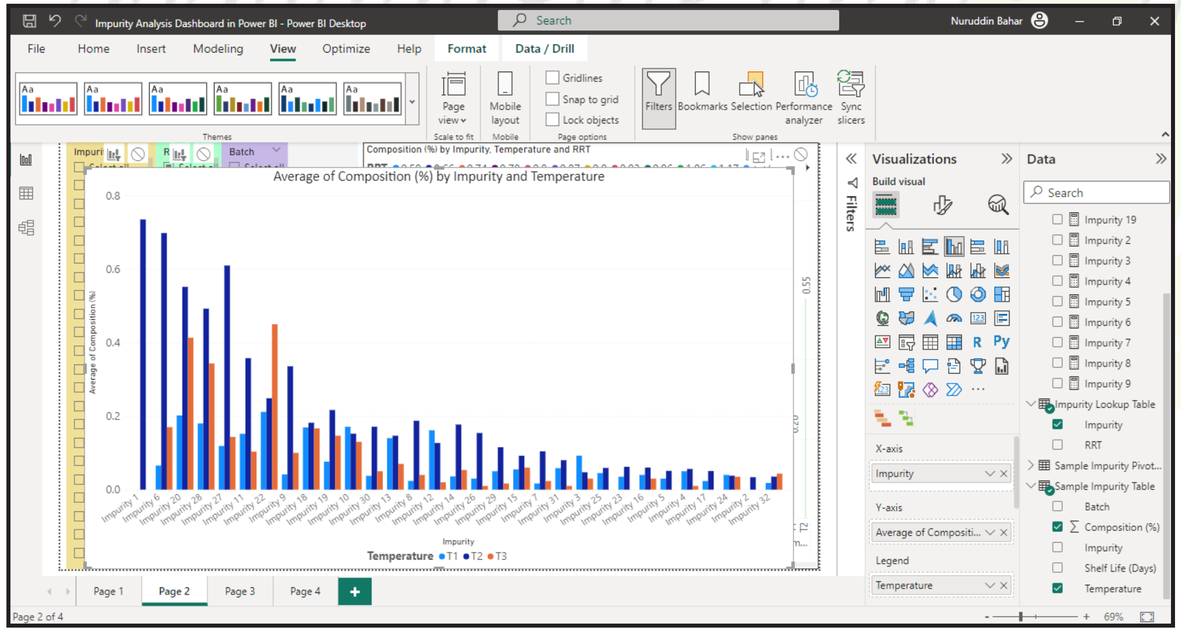
Main impurity dashboard can be manipulated to get insights. E.g. only four impurities of interest can be selected using slicers and their values compared with respect to temperature.

Or the RRT values can be compared for each temperature and their impurity composition can be obtained on a single line chart.

PROJECTS

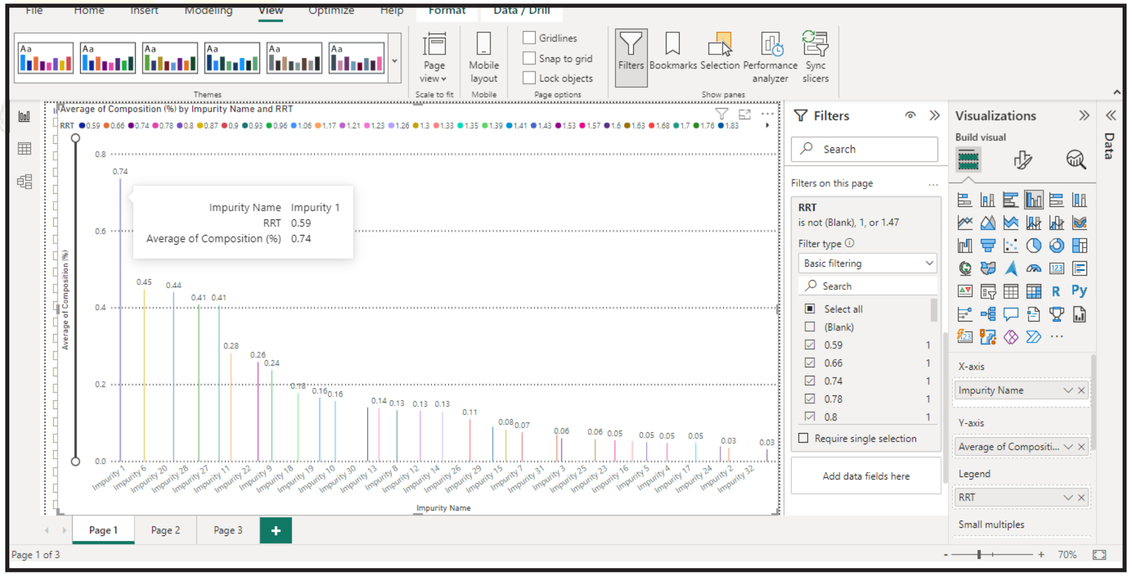
A correlation plot visualizes the relationships between impurities, providing insights into correlations and groupings. Clustered column charts and line plots display the general trends of impurity composition by temperature and identify impurities that experience significant increases with longer shelf-life.

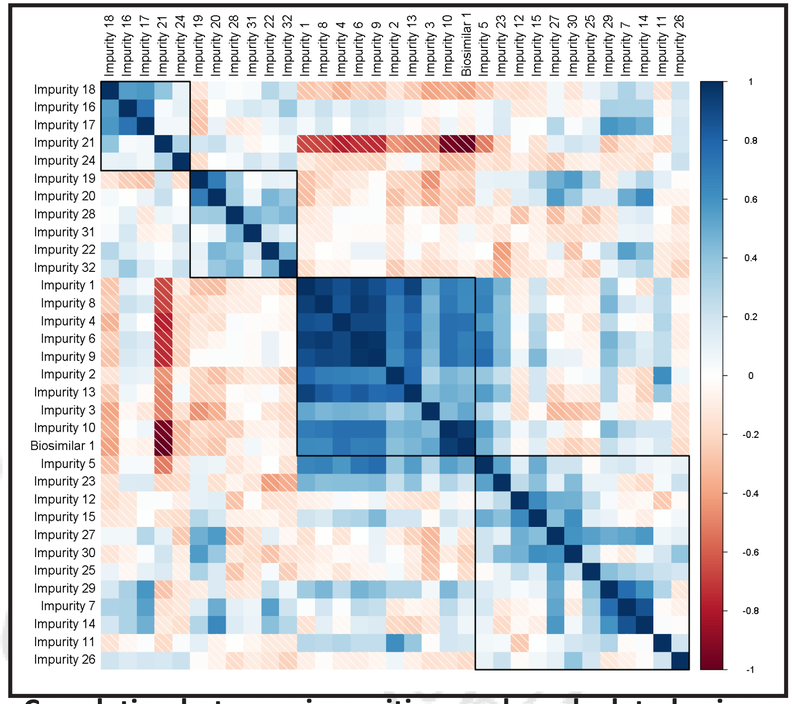
Overall, by leveraging machine learning, data visualization, and analysis techniques in Power BI, researchers can gain meaningful insights from the impurity data and enhance their decision-making process based on the actionable information obtained.



Variation in impurity composition can be measured with respect to the shelf life (days).

Average impurity composition across all batches can be plotted.





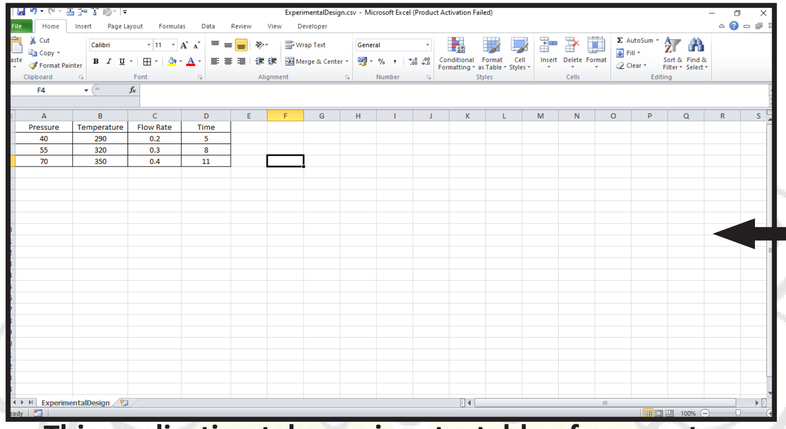
Correlation between impurities can be calculated using a correlation plot.

PROJECT 8 : DESIGN OF EXPERIMENTS GUI

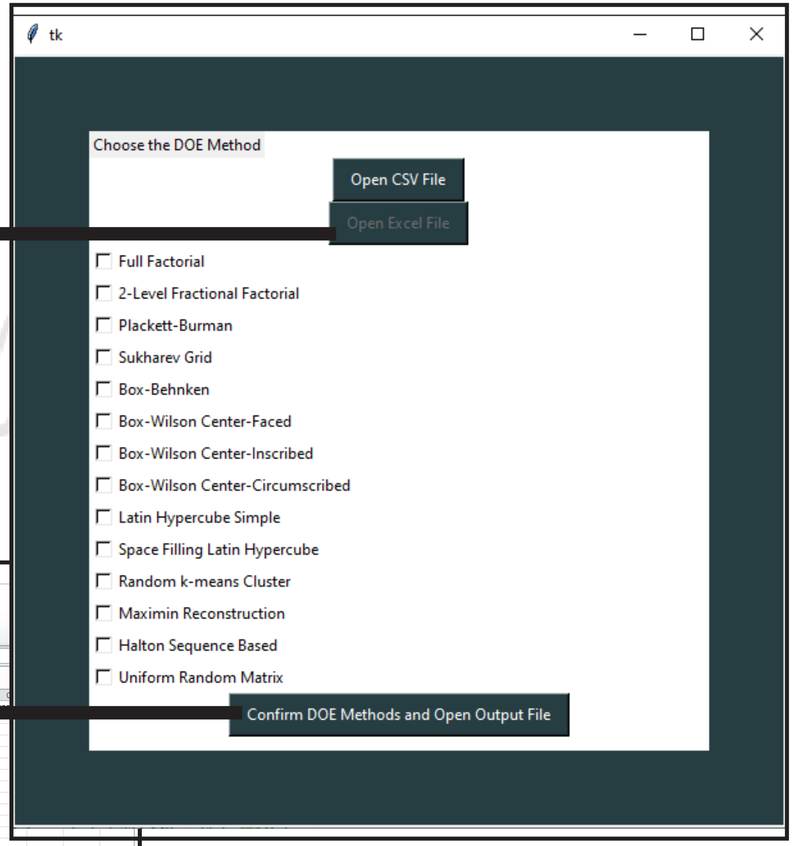
This project includes writing a Python script that creates a graphical user interface (GUI) application for performing Design of Experiments (DOE) using various libraries and modules.

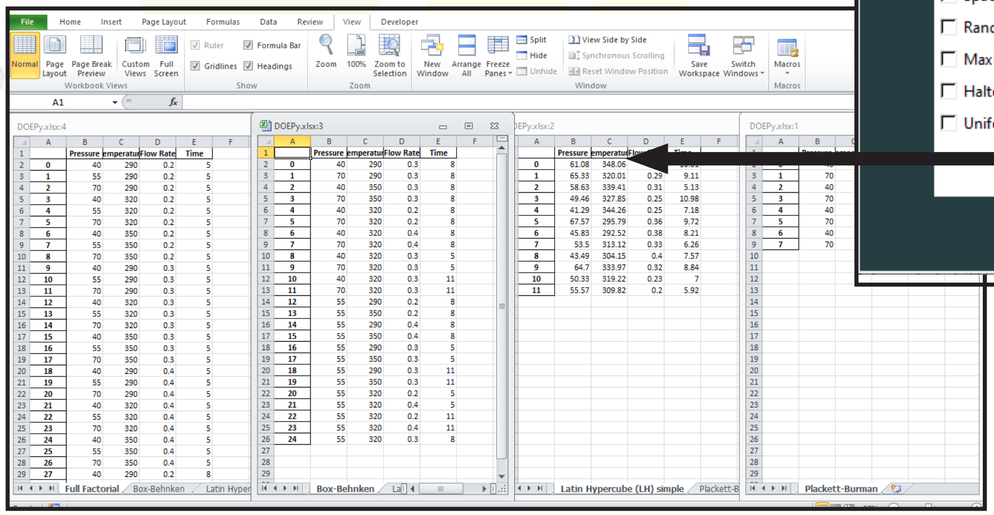
Here's a summary of what the code does:

- It imports the necessary libraries, including pandas, tkinter, os, load_workbook from openpyxl, and norm from scipy.stats.distributions.
- The script defines a dictionary called data_dict, which contains factor names as keys and corresponding lists of values as values. This dictionary represents the factors used in the experimental design.
- The script defines a class called DOEApp, which represents the GUI application. This class contains methods that correspond to different DOE methods, such as full factorial, fractional factorial, Plackett-Burman, Sukharev Grid, Box-Behnken, Latin Hypercube, etc.
- The GUI application is created using the tkinter library. It consists of a canvas, a frame, and various buttons and checkboxes that allow the user to select different DOE methods.



This application takes as input a table of parameters (headers) with end-point and/or central values provided as values under the repective columns.





Once the parameter table is loaded, we can choose among different DOE algorithms (multiple choices allowed). On clicking 'Confirm DOE Methods and Open Output File' is clicked, the DOE table with optimal number of experiments and designated parametric values is obtained.

- When a DOE method is selected, the corresponding method in the DOEApp class is called. This method generates the experimental design using the chosen method and saves the results in an Excel file.
- The application also provides functionality to open a CSV file to load the factor values and an Excel file to save the output.
- The main function creates an instance of the DOEApp class and starts the GUI application.

In addition, the code has been enhanced by utilizing PyInstaller to convert the Python script into a standalone executable. The PyInstaller conversion process involves installing PyInstaller on the system, analyzing the script, resolving dependencies, and generating an executable file. This executable file contains the bundled code, the Python interpreter, and the necessary dependencies. By using PyInstaller, the GUI application can be easily distributed and executed on other machines without the need for a separate Python installation or external dependencies.

Overall, current project is a Python script that has been transformed into a user-friendly GUI application for DOE, and further enhanced by utilizing PyInstaller to convert it into a standalone executable for easy distribution and execution.

Python code for DOE GUI

```
import pandas as pd
import os
import tkinter as tk
from tkinter import filedialog

from openpyxl import load_workbook
from doepy import build, read_write

class DOEApp:
    def __init__(self, master):
        self.data_dict = {}
        self.excelfilepath = “”
        self.writer = None

        canvas = tk.Canvas(master, height=600, width=600,
bg=”#263D42”)
        canvas.pack()

        frame = tk.Frame(master, bg=’white’)
        frame.place(relwidth=0.8, relheight=0.8, relx=0.1, rely=0.1)

        self.label = tk.Label(frame, text=”Choose the DOE Method”)
        self.label.pack(anchor=’w’)

        self.openCSVFile = tk.Button(frame, text=”Open CSV File”,
padx=10, pady=5, fg=”white”, bg=”#263D42”,
                                     command=self.open_csv_file)
        self.openCSVFile.pack()

        self.openExcelFile = tk.Button(frame, text=”Open Excel File”,
padx=10, pady=5, fg=”white”, bg=”#263D42”,
                                       command=self.open_excel_file)
        self.openExcelFile.pack()
        self.openExcelFile[‘state’] = ‘disabled’

        methods = [
            (“Full Factorial”, self.full_factorial),
            (“2-Level Fractional Factorial”, self.two_level_fractional_facto-
rial),
            (“Plackett-Burman”, self.plackett_burman),
            (“Sukharev Grid”, self.sukharev_grid),
            (“Box-Behnken”, self.box_behnken),
            (“Box-Wilson Center-Faced”, self.box_wilson_center_faced),
            (“Box-Wilson Center-Inscribed”, self.box_wilson_center_in-
scribed),
            (“Box-Wilson Center-Circumscribed”, self.box_wilson_
center_circumscribed),
            (“Latin Hypercube Simple”, self.latin_hypercube_simple),
            (“Space Filling Latin Hypercube”, self.latin_hypercube_space_
filling),
            (“Random k-means Cluster”, self.random_k_means_cluster),
            (“Maximin Reconstruction”, self.maximin_reconstruction),
            (“Halton Sequence Based”, self.halton_sequence_based),
            (“Uniform Random Matrix”, self.uniform_random_matrix)
        ]

        for text, command in methods:
            tk.Checkbutton(frame, text=text, bg=’white’, com-
mand=command).pack(anchor=’w’)

        tk.Button(frame, text=”Confirm DOE Methods and Open Output
File”, padx=10, pady=5, fg=”white”, bg=”#263D42”,
                  command=self.confirm).pack()

    def open_csv_file(self):
        csvfilepath = filedialog.askopenfilename(title=”Choose Source
CSV File”,
                                                 filetypes=((“Comma-separated val-
ues (CSV) files”, “*.csv”),
                                                            (“All files”, “*.*”)))
        self.data_dict = read_write.read_variables_csv(csvfilepath)
        self.openExcelFile[‘state’] = ‘active’

    def open_excel_file(self):
        self.excelfilepath = filedialog.askopenfilename(title=”Choose
Target Excel File”,
                                                        filetypes=((“Excel files”,
“*.xlsx”), (“All files”, “*.*”)))
```

```
        self.writer = pd.ExcelWriter(self.excelfilepath, engine='xlsxwriter')

    def execute_doe(self, doe_name, doe_function):
        if not self.data_dict:
            return

        doe_results = round(doe_function(self.data_dict), 2)
        df_doe = pd.DataFrame(doe_results)
        df_doe.to_excel(self.writer, sheet_name=doe_name)

    def full_factorial(self):
        self.execute_doe("Full Factorial", build.full_fact)

    def two_level_fractional_factorial(self):
        self.execute_doe("2-L Frac Fact", build.frac_fact_res)

    def plackett_burman(self):
        self.execute_doe("Plackett-Burman", build.plackett_burman)

    def sukharev_grid(self):
        sample_no = int(input("Enter number of samples: "))
        self.execute_doe("Sukharev Grid", lambda data: build.sukharev(data,
num_samples=sample_no))

    def box_behnken(self):
        self.execute_doe("Box-Behnken", build.box_behnken)

    def box_wilson_center_faced(self):
        self.execute_doe("Box-Wilson Center-Faced", lambda data: build.cen-
tral_composite(data, face='ccf'))

    def box_wilson_center_inscribed(self):
        self.execute_doe("Box-Wilson Center-Inscribed", lambda data: build.
central_composite(data, face='cci'))

    def box_wilson_center_circumscribed(self):
        self.execute_doe("Box-Wilson Center-Circumscribed", lambda data:
build.central_composite(data, face='ccc'))

    def latin_hypercube_simple(self):
        self.execute_doe("Latin Hypercube (LH) simple", lambda data: build.
lhs(data, num_samples=12))

    def latin_hypercube_space_filling(self):
        self.execute_doe("Space Filling LH", lambda data: build.space_filling_
lhs(data, num_samples=12))

    def random_k_means_cluster(self):
        self.execute_doe("Random k-means Cluster", lambda data: build.
random_k_means(data, num_samples=12))

    def maximin_reconstruction(self):
        self.execute_doe("Maximin Reconstruction", lambda data: build.max-
imin(data, num_samples=12))

    def halton_sequence_based(self):
        self.execute_doe("Halton Sequence Based", lambda data: build.hal-
ton(data, num_samples=12))

    def uniform_random_matrix(self):
        self.execute_doe("Uniform Random Matrix", lambda data: build.uni-
form_random(data, num_samples=12))

    def confirm(self):
        if self.writer is not None:
            self.writer.save()
            self.writer.close()
            if self.excelfilepath:
                os.startfile(self.excelfilepath)

def main():
    root = tk.Tk()
    app = DOEApp(root)
    root.mainloop()

if __name__ == "__main__":    main()
```

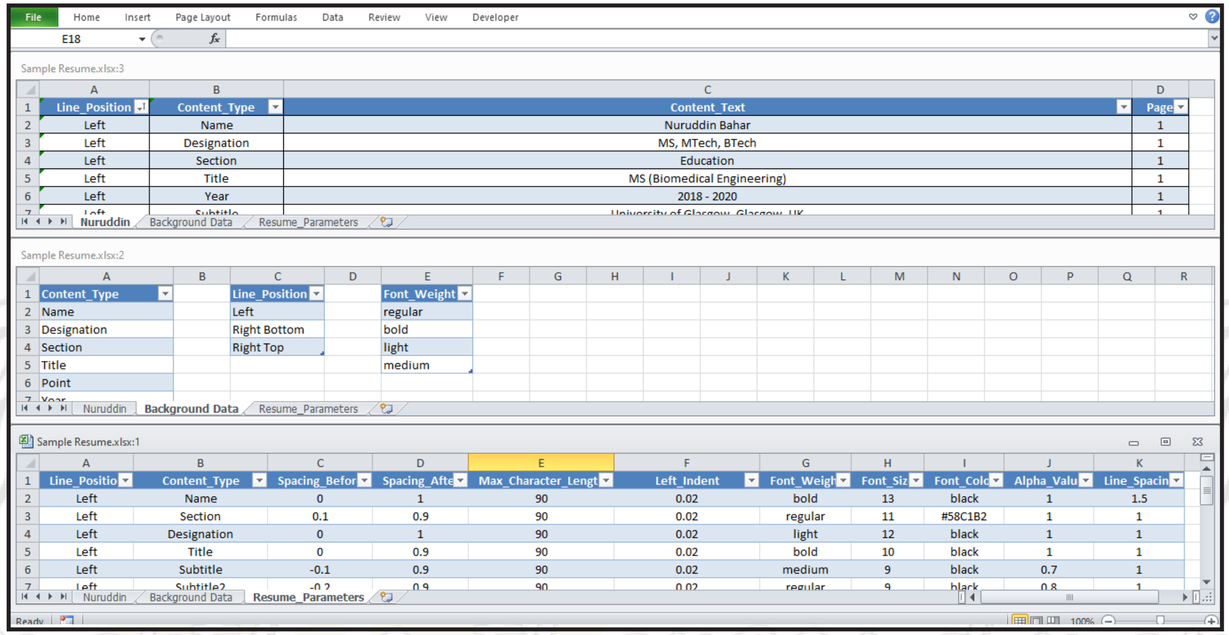
nbahar@musingsinbiology.com

PROJECT 9 : RESUME MAKER TOOL

This project describes the process of creating a resume using Python coding and an Excel table. For a cost-effective solution for creating a resume, we can use Python programming.

The following steps describe the process of creating a homemade resume maker:

- ***Idea***: I searched for online resumes made using Python and found various implementations. I decided to use matplotlib to create a plot of the resume and utilize the annotate functionality to position the text effectively.
- ***Resume Excel Table***: I created an Excel workbook with three worksheets: Data table, Background data, and Resume parameters. The data table contained information such as line position, content type, content text, and page number.
- ***Resume Parameters***: The resume parameters table included various parameters for each content type and line position combination, such as spacing before and after, maximum character length, left indentation, font weight, font size, font color, alpha value, and line spacing.



Sample Resume.xlsx:3

Line_Position	Content_Type	Content_Text	Page
Left	Name	Nuruddin Bahar	1
Left	Designation	MS, MTech, BTech	1
Left	Section	Education	1
Left	Title	MS (Biomedical Engineering)	1
Left	Year	2018 - 2020	1
Left	Subtitle	University of Glasgow, Glasgow, UK	1

Sheets: Nuruddin | Background Data | Resume_Parameters

Resume content table

Sample Resume.xlsx:2

Content_Type	Line_Position	Font_Weight
Name	Left	regular
Designation	Right Bottom	bold
Section	Right Top	light
Title		medium
Point		

Sheets: Nuruddin | Background Data | Resume_Parameters

Background data

Sample Resume.xlsx:1

Line_Positio	Content_Type	Spacing_Befor	Spacing_Afte	Max_Character_Lengt	Left_Indent	Font_Weigl	Font_Siz	Font_Colc	Alpha_Valu	Line_Spacin
Left	Name	0	1	90	0.02	bold	13	black	1	1.5
Left	Section	0.1	0.9	90	0.02	regular	11	#58C1B2	1	1
Left	Designation	0	1	90	0.02	light	12	black	1	1
Left	Title	0	0.9	90	0.02	bold	10	black	1	1
Left	Subtitle	-0.1	0.9	90	0.02	medium	9	black	0.7	1
Left	Subtitle2	-0.2	0.9	90	0.02	regular	9	black	0.8	1

Sheets: Nuruddin | Background Data | Resume_Parameters

Resume parameters

	Line_Position	Content_Type	Content_Text	Page	Spacing_Before	Spacing_After	Max_Character_Length	Left_Indent	Font_Weight	Font_Size	Font_Color	Al
0	Left	Name	Nuruddin Bahar	1	0.0	1.0	90	0.02	bold	13	black	
1	Left	Designation	MS, MTech, BTech	1	0.0	1.0	90	0.02	light	12	black	
2	Left	Section	Education	1	0.1	0.9	90	0.02	regular	11	#58C1B2	
3	Left	Title	MS (Biomedical Engineering)	1	0.0	0.9	90	0.02	bold	10	black	
4	Left	Year	2018 - 2020	1	-0.9	1.0	90	0.02	light	9	black	
5	Left	Subtitle	University of Glasgow, Glasgow, UK	1	-0.1	0.9	90	0.02	medium	9	black	

Merged pandas DataFrame that merges the resume content table along with parameters table based on content type.

- Pythonic Resume Maker: I imported relevant libraries and defined functions to create the resume. I created a matplotlib plot with the size of an A4 sheet and added vertical and horizontal lines as space fillers. The data table and resume parameters were read and merged to create a content table.
- Generating the Resume: I initialized y-coordinate parameters for each section and defined variables for the previous page and total number of lines. I used a loop to filter the data based on the page and assigned values from each row.

The resume plot and spacing variables were updated for each new page. An annotation function was used to annotate each row of text, and text wrapping was implemented for longer lines. The code handled different line positions and annotated the right top and bottom sidebars. The plots were saved as PDFs, and the PDF files were combined using the PDFMerger object.

- Saving Multiple Styles: I have also demonstrated how to save multiple styled PDFs by assigning different font styles in a loop and calling the create_resume() function.

NURUDDIN BAHAR

MS, MTech, BTech

EDUCATION

MS (Biomedical Engineering) 2018 - 2020

University of Glasgow, Glasgow, UK

- Qualified with Distinction
- THESIS: Design and characterization of poly(ethylene glycol) laminin (PEG LM) based mechano responsive hydrogels for drug delivery and gene therapy Grade A4 (19/22)

MTech (Nanotechnology) 2015 - 2017

Indian Institute of Science, Bangalore, India

- Qualified with CGPA 5.9/8.0
- THESIS: Towards DNA Origami Based Magnetic Nanopropellers Grade A (7/8)

BTech (Chemical) 2009 - 2013

Indian Institute of Technology Bombay, Mumbai, India

- Qualified with CPI 7.53/10.0

WORK EXPERIENCE

Research Scientist

Wockhardt Research Centre, Aurangabad, India

Bioprocess Engineering and MSAT 10/2020 - Present

- In process of creating a Django website to display relevant projects
- Created a VBA-based robotic batch process recording application for ease of reporting
- Statistically calculated mock pools for CIEC stage of Insulin Glargine and RP stage of Yeast Human Insulin using pandas and matplotlib
- Created a Python-based design of experiments user interface for creating experimental plans for given parameters
- Statistically analyzed 2021-22 batches of Yeast Insulin, applied Python-based data visualization, inferential statistics and descriptive statistics
- Built Gantt charts, timeline charts, chromatograph visualization tool using Python

Associate Software Engineer

Exeter Software Pvt. Ltd., Bengaluru, India 07/2013 - 01/2014

- Learned essentials of Javascript, SQL, HTML, CSS and Salesforce as a part of training
- Certified with Oracle Policy Automation 10 Rule Developer Essentials 1Z0-534
- Created an online passport application portal using Salesforce

TRAINING AND INTERNSHIPS

Skill Development Trainee

Centre for Cellular and Molecular Biology, Hyderabad, India 10/2017 - 01/2018

Aurangabad (MAH), INDIA
nbahar2@gmail.com
(+91) 701-436-8446
www.musingsinbiology.com

Scholastic Achievements & Awards

- AIR 718 - IITJEE 2009
- AIR 16 - GATE (Biotechnology) 2015
- CSIR-JRF (Life Sciences) 2015, 2018
- CBSE 10th merit certificate for 100% in social science 2006

Certifications

- Certified Associate in Python Programming (Python Institute)
- MS Excel Expert (Microsoft)
- MS Word Expert (Microsoft)
- MS Outlook Associate (Microsoft)
- MS PowerPoint Associate (Microsoft)

Resume Page 1

NURUDDIN BAHAR

MS, MTech, BTech

R & D Intern

Tridiagonal Solutions Pvt. Ltd., Pune, India 05/2012 - 07/2012

- Developed and optimied CSTR and tank-in-series models for a chemical plant simulation
- Explored process modelling applications like cape open to cape open (COCO) and differential algebraic equation (DAE) tools

PROJECTS

Django Website

- Currently creating a portfolio website using Django, Python, TinyMCE, HTML, CSS and JS
- Created topic-wise list of projects, including data analysis, data visualization, python tools and biology
- Learned various methods of deploying website (AWS, Heroku, Azure, Digital Ocean)

Statistical Report for Production Batches

- Cleaned and visualized data using pandas/matplotlib libraries in Python
- Applied inferential statistics (bootstrapping, correlation, confidence intervals) and descriptive statistics (mean, median, box plots)

Downstream Processes Batch Template

- Used macro-based approach in Excel, Word and VBA
- Created an impurity analysis module and process batch scheduler

Design of Experiments GUI

- Used Openpyxl and Tkinter libraries in Python
- Created choice-based experimental plans for user to select

Productivity Tools

- Created a process chromatogram visualization tool using matplotlib
- Created a categorial Gantt chart and timeline chart using Excel data table

Biology Interactive Content

- Added interactive biology content in Musings in Biology website
- In process of using HTML, CSS, and JavaScript to shift the content on Django website

CONFERENCES ATTENDED

Chemical Engineering Association Symposium (ChEA), IISc Bangalore 2017

- Presented poster on 'DNA Origami Based Magnetic Nanopropellers'

International Conference on Emerging Electronics (ICEE), IIT Bombay 2016

- Presented paper on 'Two-tier size optimization of monodispersed iron-oxide nanoparticles'

Aurangabad (MAH), INDIA
nbahar2@gmail.com
(+91) 701-436-8446
www.musingsinbiology.com

Skills

- MS Office Suite
- Python
- Django
- Power BI
- HTML & CSS
- Machine Learning
- Data Visualization
- Data Analysis
- VBA
- LaTeX
- JMP
- Illustrator
- SQL, PostgreSQL

Other Certifications

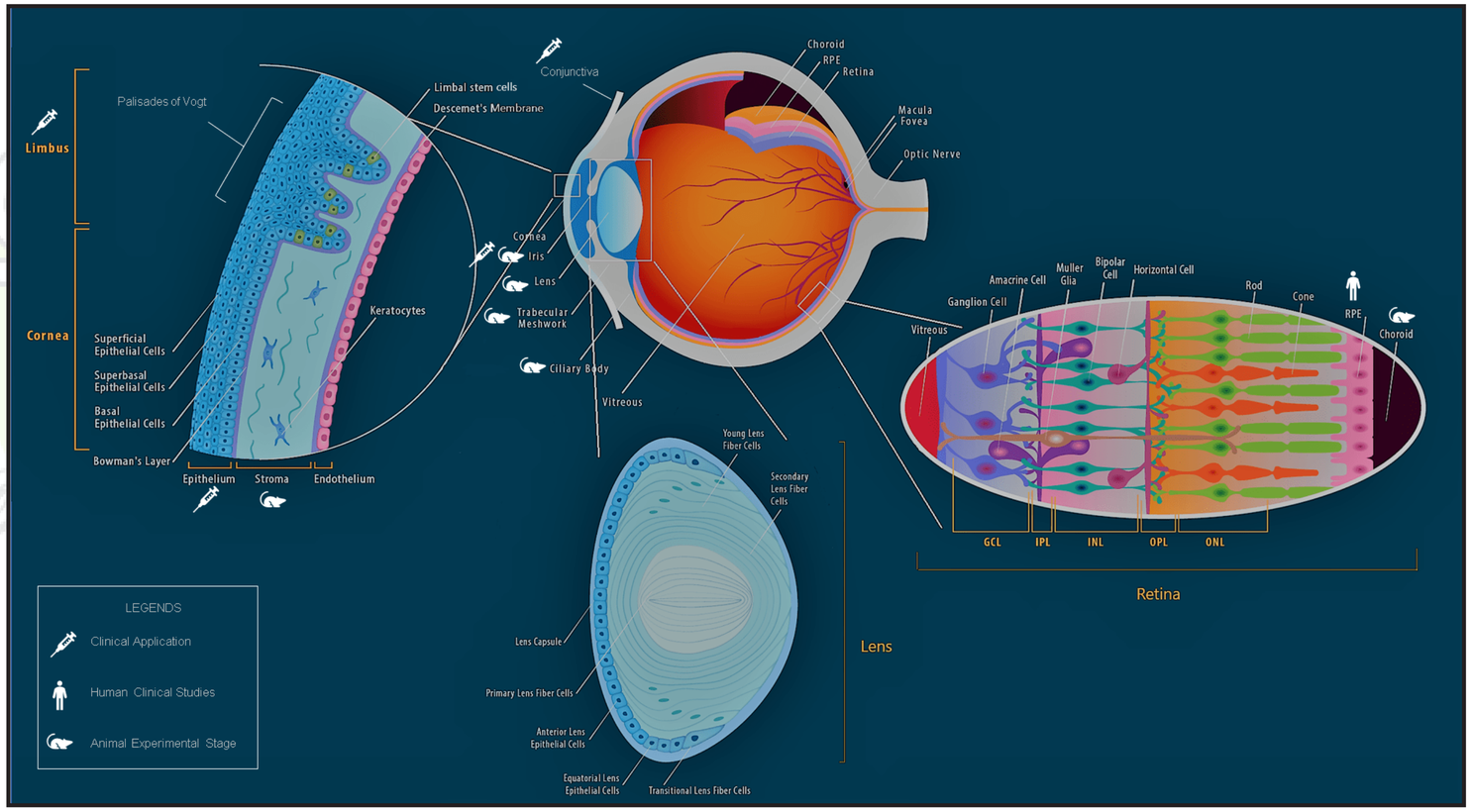
- LaTeX for Professional Publications

Resume Page 2

PROJECT 10 : SCIENTIFIC PAPERS

PAPER 1 : Sight Restoration via Ocular Tissue Engineering

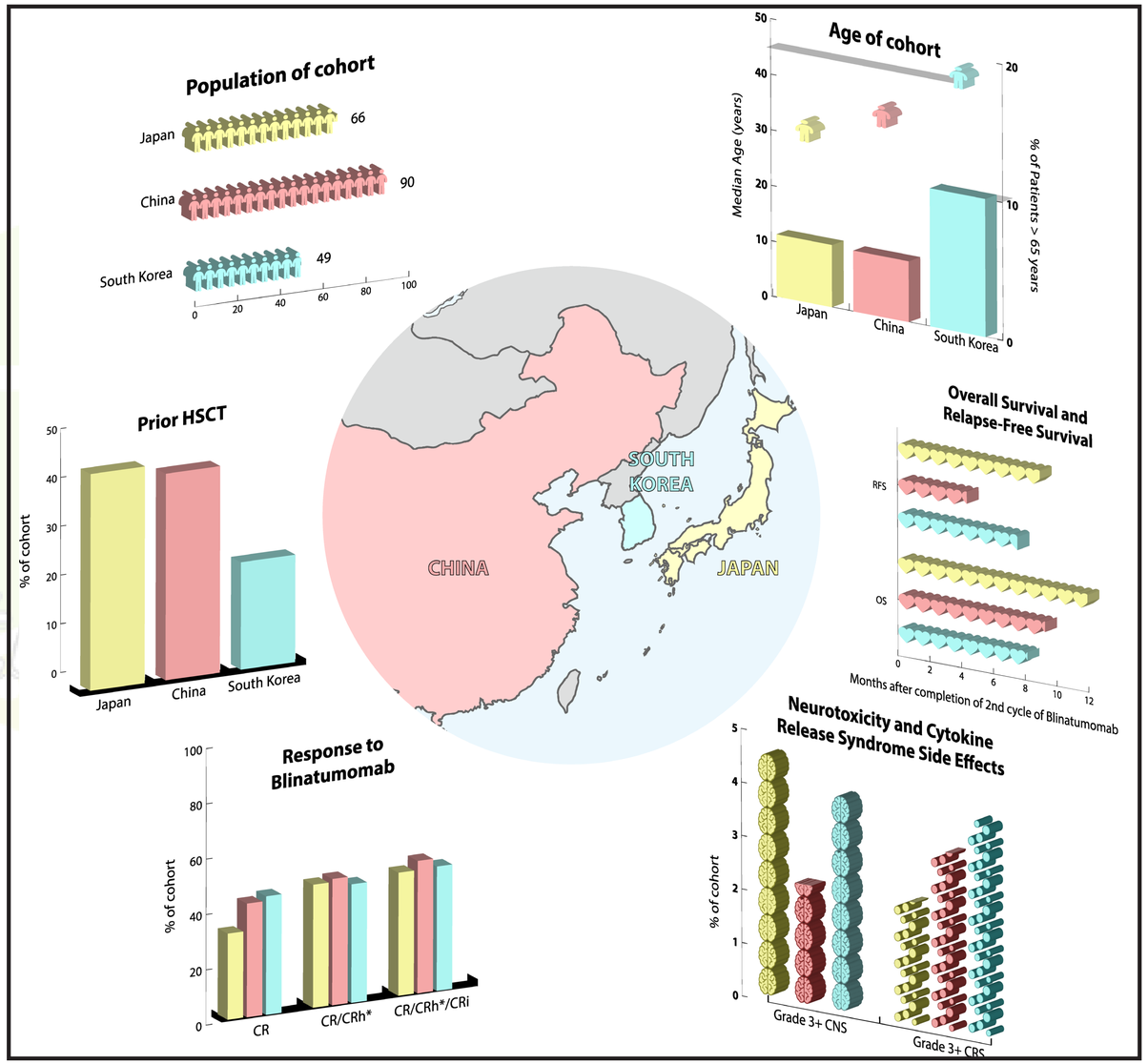
The eyes, often regarded as the windows to the soul, play a pivotal role in visual perception and understanding. Vision loss can pose significant physical and emotional challenges, motivating individuals to seek ways to regain or improve their sight. Ocular regenerative therapy, encompassing stem cell-based and scaffold-based techniques, has emerged as a promising approach for addressing various degenerative eye conditions. This review explores the interpretation of light by the eye and examines common degenerative eye disorders affecting different eye components. It delves into the different types of ocular and extra-ocular stem cells, their phenotypic markers, functional significance, and associated research. The utilization of scaffold-based tissue techniques in vivo is also discussed. Furthermore, insights into approved, marketable, and phase-tested stem cell methods with potential medical success are provided. Challenges encountered throughout the research process are highlighted, culminating in a discussion of the potential research problems and future prospects of ocular regenerative therapy. The integration of stem cell-based and scaffold-based approaches holds promise for revolutionizing ocular therapeutics and paving the way for vision restoration.



PROJECT 10 : SCIENTIFIC PAPERS

PAPER 2 : Use of Blnatumomab in Asian Patients with Relapsed/Refractory B-cell Precursor Acute Lymphoblastic Leukemia - a Review

Blinatumomab is an FDA-approved biopharmaceutical used for the treatment of acute lymphoblastic leukemia (ALL). It is a bispecific T-cell engager that targets CD19 on lymphoblasts and CD3 on cytotoxic T cells. This report focuses on recent clinical trials conducted in Asian countries, including Japan, South Korea, and China, to evaluate the efficacy of blinatumomab in relapsed or refractory B-cell ALL. The trials included patients who were primary refractory or relapsed after previous treatments. Blinatumomab was administered over several cycles, and the response was measured based on complete remission rates and minimal residual disease. The trials demonstrated positive outcomes, including increased overall survival and favorable hematologic remission rates compared to standard chemotherapy. However, blinatumomab treatment also carried risks of cytokine release syndrome and neurological toxicities. The report emphasizes the need for globally conducted clinical trials to further validate the efficacy of blinatumomab as an anti-cancer treatment. The results from these Asian trials contribute to the understanding of blinatumomab's effectiveness in diverse populations and provide valuable data for future research in ALL treatment.



PROJECT 10 : SCIENTIFIC PAPERS

PAPER 3 : Two-tier Size Optimization of Monodispersed Iron Oxide Magnetic Nanoparticles

Magnetic iron oxide nanoparticles (IONPs) promise a wide range of biomedical applications, including in-vivo cancer theranostics. In this paper, synthesis of monodispersed IONPs using co-precipitation of $FeCl_3$ and $FeCl_2$ in basic medium has been considered. Analysis of size range has been subjected to differences in two parameters simultaneously viz. pH of solution buffer storing IONPs and concentration of NH_4OH in the initial reaction. Scanning electron microscopy (SEM), X-Ray Photoelectron Spectroscopy (XPS), X-Ray Diffraction (XRD) and Dynamic Light Scattering (DLS) were utilized to characterize the morphology, size-distribution and functionalization of the synthesized nanoparticles. It was observed that very high pH (12) and low pH (7,8) basic buffer solutions and a typical concentration of NH_4OH (2M) displayed higher monodispersity and smaller sizes of nanoparticles.

ICEE 2016 Poster

Two-tier Size Optimization of Monodispersed Iron Oxide Nanoparticles

Nuruddin Bahar, Abhijith Raghuprasad, Ambarish Ghosh, Banani Chakraborty

Abstract

Magnetic iron oxide nanoparticles (IONPs) promise a wide range of biomedical applications, including in-vivo cancer theranostics. In this paper, synthesis of monodispersed IONPs using co-precipitation of $FeCl_3$ and $FeCl_2$ in basic medium has been subjected to differences in two parameters simultaneously viz. pH of solution buffer storing IONPs and concentration of NH_4OH in the initial reaction in order to determine nanoparticle size variability.

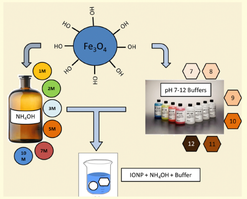


Figure 1: Experiment overview.

KEYWORDS - iron oxide nanoparticles (IONPs), co-precipitation, buffer pH, base concentration

Materials and Methods

Co-Precipitation Reaction

$$Fe^{2+} + 2\,Fe^{3+} + 8\,OH^- \xrightarrow{120\,^\circ C,\,N_2,\,1\,atm} Fe(OH)_2 + 2\,Fe(OH)_3 \longrightarrow Fe_3O_4 + 4\,H_2O$$

Experiment

- Mix 2M $FeCl_3 \cdot 6\,H_2O$ and 1M $FeCl_2 \cdot 4\,H_2O$ in 100 mL deionized water and add 0.85 mL of conc. HCl
- Purge in continuous N_2 gas at 900 RPM and heat to 120 °C in a reflux column
- Add 0.5 mL of 'x'M NH_4OH where x=1,2,3,5,7 & 10. Black precipitate of IONPs is obtained
- Separate IONPs by magnetic decantation
- Coat IONPs with citric acid by mixing 1.5 g IONPs with 2 mL citric acid
- Suspend 0.5 mL of solution mixture to 1 mL of various buffer solutions as described in TABLE 1.

pH	Solute			Solvent			Buffer Type
	Compound	Vol.(mL)	Conc.(M)	Compound	Vol.(mL)	Conc.(M)	
7	NaH_2PO_4	39	1	Na_2HPO_4	61	1	Neutral
8	NaH_2PO_4	6.8	1	Na_2HPO_4	93.2	1	Basic
9	Tris [1]	89.8	0.1	HCl	10.2	0.1	Basic
10	Borax [2]	73.2	0.1	NaOH	26.8	0.1	Basic
11	$NaHCO_3$	68.8	0.05	NaOH	31.2	0.1	Basic
12	Na_2HPO_4	65	0.05	NaOH	35	0.1	Basic

Table 1: Protocols for various buffer solutions (net volume = 100 mL).

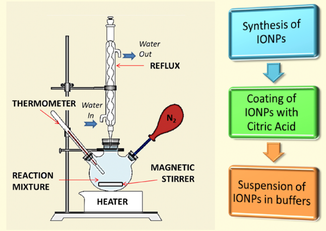


Figure 2: Experimental Setup.

Experimental Conditions

N_2 gas supply prevents oxidation of Fe_3O_4 to γ-Fe_2O_3 from ambient O_2 while high temperature creates enough turbulence to prevent nanoparticles from polydispersity

Results

Characterization

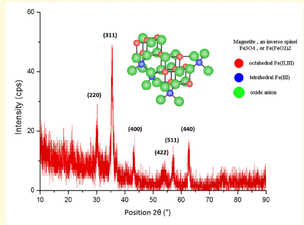


Figure 3: XRD spectrum of Fe_3O_4 nanoparticles.

Crystal Structure of Magnetite

Magnetite, Fe_3O_4, is an inverse spinel with a formula of $(Fe^{3+})_T[Fe^{2+}Fe^{3+}]_OO_4$, as indicated by strongest reflection that proceeds from the (311) plane.

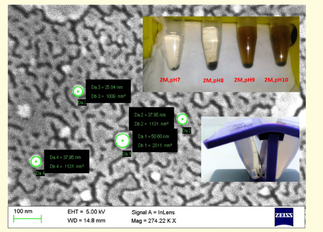


Figure 4: SEM image of citric acid coated IONPs (CA-IONPs) from 2M NH_4OH stored in buffer pH 8. Inset : Samples of IONPs created in 2M NH_4OH stored in various buffer solutions (top-right) and magnetic separation of IONPs (bottom-right).

IONP Size & Distribution

- Most CA-IONPs were of size range 30-50 nm.
- A few bigger but rounded agglomerates were possibly formed due to anisotropic nucleation events.
- Smaller and stable CA-IONPs were present in pH 7, 8 and 12.
- NH_4OH concentration of 2M seemed to possess smaller, uniformly-sized and less-distributed nanoparticles, irrespective of pH.
- An oculiform variation is observed in dR/R graph.

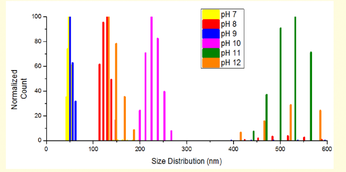


Figure 5: Size distribution of 1M NH_4OH CA-IONPs using Dynamic Light Scattering(DLS) at varied pH.

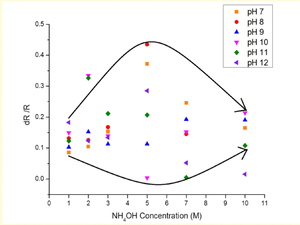


Figure 6: $\frac{dR}{R}$ values of CA-IONPs using DLS.

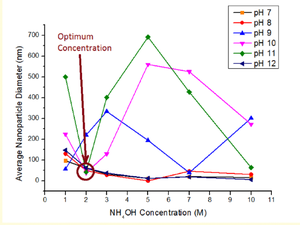


Figure 7: $D_{average}$ values of CA-IONPs using DLS.

Conclusions

- Variation in pH is dominant over base concentration, in that low pH solutions have smaller CA-IONPs
- Intermediate basic pH increases polydispersity, while higher and lower pH reduces it
- 2M NH_4OH is the locus for increased monodispersity and reduced size under given conditions

References

- R. Massart, "Preparation of aqueous magnetic liquids in alkaline and acidic media",*IEEE Transactions on Magnetics*, vol. 17, no. 2, pp. 1247-1248, 1981.
- A.-H. H. Lu, E. L. Salabas, F. Schüth, and F. Schüth, "Magnetic nanoparticles: Synthesis, protection, functionalization, and application", *Angewandte Chemie - International Edition*, vol. 46, no. 8, pp. 1222-1244, feb 2007.

nbahar@musingsinbiology.com

PROJECT 11 : DJANGO WEBSITE

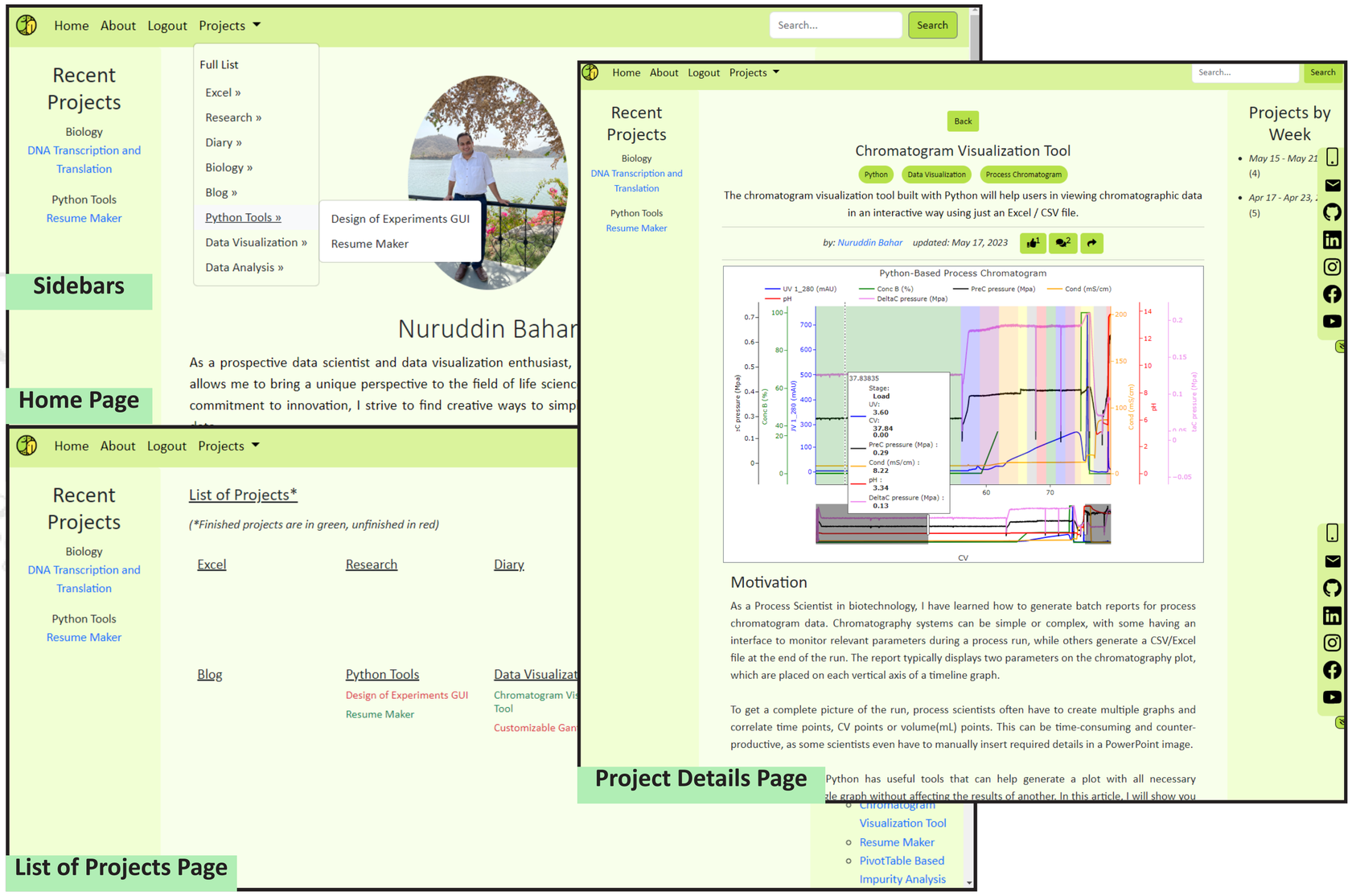
This Django website is a platform that allows authenticated users to create, publish, and share articles related to data analysis, data visualization, biology, and Python tools.

Here is a summary of its features:

Frontend and Backend Development: The website is built using Django, a Python web framework, and allows users to add HTML, CSS and JS to customize the appearance.

Article Management: Users can create, edit, and publish articles on the website. Each article consists of various components, including the main content, comments section, a like button for users to show appreciation, and social media buttons to share the article on different platforms.

Project Listing: The website includes a list of projects related to data analysis, data visualization, biology, and Python tools. Users can browse these projects, view details, and access relevant resources.



PROJECTS ABOUT FEATURES

Search Functionality: The website provides a search page where users can enter keywords to search for specific articles or projects. This helps users quickly find relevant content.

User Authentication: The website offers a login page and a signup page for users to create accounts and authenticate themselves. This allows users to personalize their experience, such as saving favorite articles or participating in discussions.

Sidebar Components: The website includes sidebars displaying recent projects and weekly projects. These sections highlight new and noteworthy content for users to explore.

Overall, this Django website serves as a platform for sharing knowledge and resources in the fields of data analysis, data visualization, biology, and Python tools. It offers article management features, project listings, search functionality, user authentication, and various components to enhance user engagement and interaction.