5BBG0206 Molecular Biology Research Skills 2021-2022

Practical Write-up Proforma

Student ID number K20041606

Family name: Popham

First name: Jude **Group: P3**

Highlight the member of staff in charge of the lab that you were working in.

- **Dr Natalie Prescott**
- **Prof Guy Tear**
- **Dr Shirley Coomber**

This Lab Report MUST BE YOUR OWN WORK. The data presented must be your own results. You will not lose marks if any step has not worked. Only the photographs (with no shared annotation or figure legends) and table of results from the ligation experiment are allowed to be identical to another student's pro forma. Any identical (or part identical) work submitted by two or more students will result in all students involved getting zero for this assessment.

Student ID (or name) of my lab partner(s) Nill Crasnacumar

The deadline for hand in the Lab Report is: 4 p.m. on Tuesday 29th March 2022

The 5BBG0206 Practical Report Pro forma must be submitted electronically to the modules KEATS site by 4 p.m. on Tuesday 29th April 2022. **This coursework MUST be submitted in electronic format as a pdf format file.** See the 'Submission of practical write-up ' section in the 5BBG0206 module handbook.

Acknowledgements

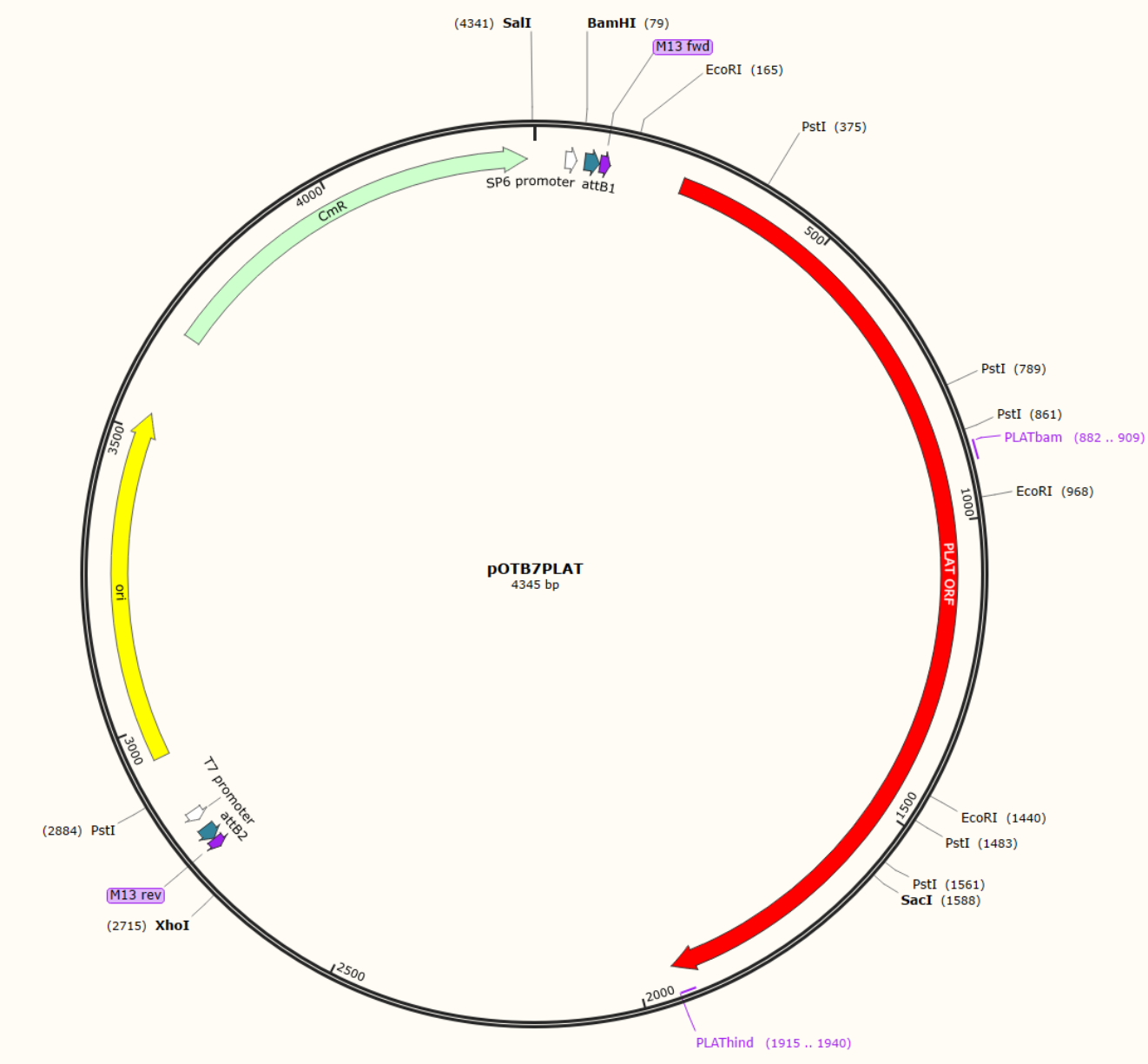
Figure 7: Sample sequencing data from 01107599 used to display the pASKPLAT sequence

Figure 8: A joint effort between Pair 1(bench next to us – Joel and Eliza) and Pair 2 (us)

DAY 1

1. Restriction enzyme characterisation of plasmid pOTB7PLAT (IMAGE 3161005)

Figure 1. Restriction enzyme map of plasmid pOTB7PLAT. The PLAT ORF has been highlighted in red and primers labelled in purple. Certain relevant restriction enzymes are visible on the plasmid map.



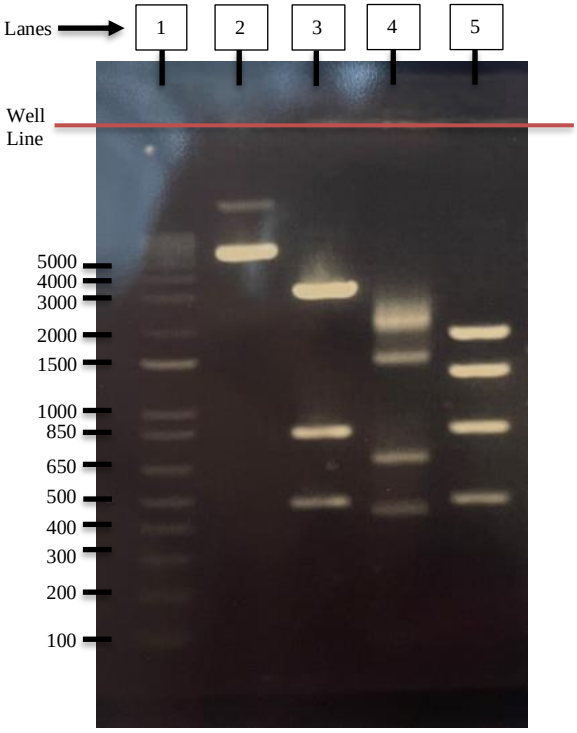


Figure 2. The lanes 1-5 of the horizontal gel electrophoresis of pOTB7PLAT restriction enzyme digests to confirm plasmid identity. **Lane 1:** Invitrogen 1kb Plus DNA Ladder, 5000bp detection limit. **Lane 2 (RE1):** Contains pOTB7PLAT with no restriction enzymes (negative control for uncut plasmid size). **Lane 3 (RE2):** Contains pOTB7PLAT + *Eco*RI restriction enymes. **Lane 4 (RE3):** Contains pOTB7PLAT + *Pst*I restriction enzymes. **Lane 5 (RE4):** Contains pOTB7PLAT and both *Eco*RI and *Pst*I restriction enzymes.

Table 1. Predicted and experimental sizes of DNA fragments produced by restriction enzyme digestion of pOTB7PLAT.

Restriction enzyme(s)	**Predicted size**	**Experimental size - measure**
***Eco*RI**	472bp, 803bp, 3070bp	475bp, 850bp, 3210bp
***Pst*I** (only if you were in FWB lab)	72bp, 78bp, 414bp, 622bp, 1323bp, 1836bp	N/a, N/a, 445bp, 660bp, 1555bp, 2215bp
***Xho*I** (only if you were in G11 or Hodgkin lab B)		
***Eco*RI and *Pst*I**	472bp, 803bp, 1275bp, 1795bp	500bp, 845bp, 1340bp, 1915bp

Results and discussion (max 150 words)

The identity of pOTB7PLAT was checked experimentally. *Pst*I's digests didn't resolve for smaller DNA fragments (<150bp) produced like 72bp and 78bp which fell outside the direction range of the Invitrogen ladder. Lane 2 (fig.2) has pOTB7PLAT DNA retarded at 5000bp, higher than 4345bp expected due to its circular form and also a nicked pOTB7PLAT band above the upper detection limit of the ladder. The supplied pOTB7PLAT plasmid is cut by the restriction enzymes *Eco*RI and *Pst*I in lanes 3-5(fig.2), producing RE digests of different lengths, shown by experimental data in Table 1.

Table 1 has some disparity among the predicted vs experimental DNA fragment sizes because it's hard to differentiate the banding resolution from 3000 onwards in the DNA ladder. The experimental results of the digests produced from the pOTB7PLAT generally resemble the predicted sizes of the fragments and it's safe to assume the correct plasmid was supplied.

Word count: 147

2. PCR amplification of *PLAT* from pOTB7PLAT

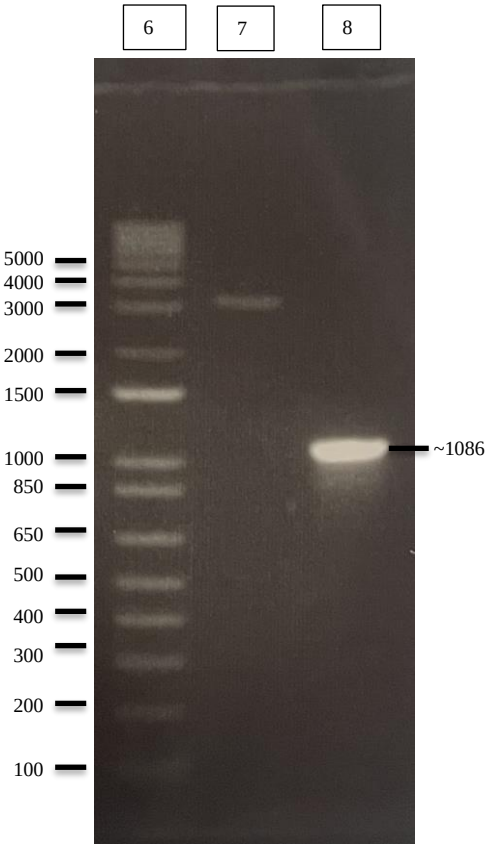


Figure 3. Lanes 6-8 of the horizontal agarose gel electrophoresis showing PCRs amplified by oligonucleotide primers and photographed by the image analyser to verify PCR sizes. **Lane 6:** 1kb plus marker for identification, 5000bp detection limit. **Lane 7:** PCR 1 – pOTB7PLAT with PLATbam and PLAThind restriction enzymes. **Lane 8:** PCR2 - pOTB7PLAT with PLATbam and PLAThind restriction enzymes.

Results and discussion (max 140 words)

Here PLAT was amplified out of pOTB7PLAT**.** The cloning strategy was the digestion of pOTB7PLAT with primers PLATbam and PLAThind to obtain the recombinant PLAT cDNA (no introns) and incorporate BamHI/HindIII restriction enzyme sites for ligation with S-Y-Q for stability (N-terminus so the 3' end matches the template for taq-polymerase). The product was amplified via PCR for 30 cycles.

PCR 1 (fig3.Lane 7) clearly failed because the band is too high, likely due to insufficient PCR1 loaded onto the gel indicated by faint banding. However, results were successful as only one PCR sample was needed to continue: the PCR 2 in lane 7 produced a clear result on the agar gel which lined up correctly with the markers in lane 6 at ~1000kb, as expected from the recombinant PLAT insert sequence information which is 1086kb including the primer groups.

Word count = 139

DAY 2

3. Restriction enzyme digestion of pASK-IBA37 plus and the PCR fragment from PLAT prior to ligation

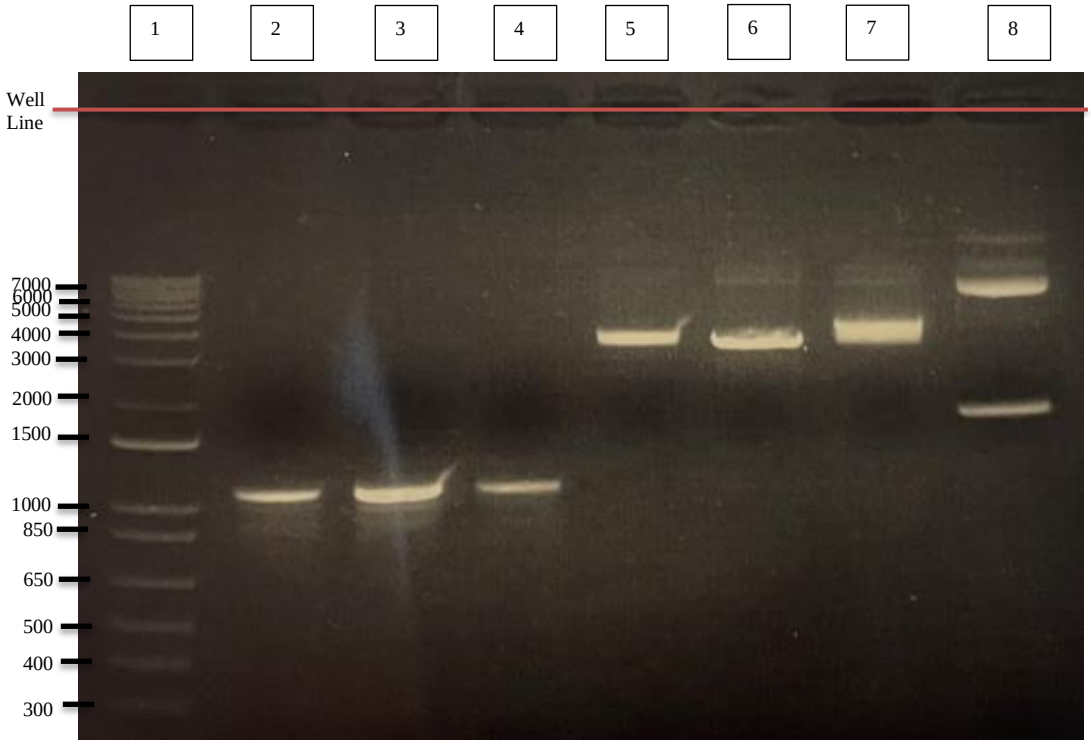


Figure 4. The analysis of restriction enzyme digests of pASK-IBA37plus and PLAT PCR fragments from pOTB7PLUS in 20ul solutions using horizontal agarose gel electrophoresis to create sticky ends for ligation and confirm plasmid the identity. **Lane 1:** 1kb plus markers, 7000bp detection limit. **Lane 2:** Purified PCR1 (PLAT cDNA amplified from pOTB7PLAT) for agarose gel. **Lane 3**: Purified PCR2 (PLAT cDNA amplified from pOTB7PLAT) for agarose gel. **Lane 4 (RE1):** *Bam*HI and *Hind*III for agarose gel with PCR1 and PCR2. **Lane 5 (RE2):** *Bam*HI and *Hind*III for agarose gel with pASK-IBA37plus. **Lane 6 (RE3):** *Bam*HI only control with pASK-IBA37plus. **Lane 7 (RE4)**: *Hind*III only control pASK-IBA37plus. **Lane 8 (RE5):** No restriction enzymes control with pASK-IBA37plus).

Results and discussion (max 140 words)

Fig.4 checks the restriction enzyme digestion of PCR and pASK-IBA37plus products with their compatible sticky ends for sub-cloning. Lane 8 has multiple bands as there are no restriction enzymes, just circular pASK-IBA37plus which encourages retards inconsistently: a bright band at ~5000kb due to nicking of the circular plasmid and the band between 1500-2000kb due to supercoiling. The faint lines at ~7000bp could be residual DNA in the well.

PCR1, PCR2 and RE1 had a band at ~1000bp, close to the expected 1074kb of the digested PLAT insert with sticky ends. The pASK-IBA37plus restriction enzyme controls (with and without restriction enzymes) and RE2 correctly show the single band of the desired plasmid product with sticky ends from *Bam*HI and *Hind*III at ~3300kb, close to the expected 3219kb position expected of the digested, linearised pASK-IBA37plus (the 51bp severed band wasn't resolved).

Word count = 140

DAY 3

4. Ligation of PLAT into pASK-IBA37plus and transformation into competent *E. coli* cells

Table 2. Colony numbers produced from the transformation into *E. coli* of the ligation product between *Hind*III *Bam*HI cut pASK-IBA37plus with *Hind*III *Bam*HI cut PLAT and the plasmid controls

	Plate 1	Plate 2	Plate 3	Plate 4	Average per plate
LIG1. plasmid + insert + ligase + *E. coli* competent cells	382	310	412	318	356
LIG2. plasmid + ligase + *E. coli* competent cells	524				
LIG3. plasmid + *E. coli* competent cells	132				
Plasmid control 10 μl	88				
Plasmid control 100 μl	1314				

Results and discussion (max 150 words)

This table is the ligation product of PLAT and pASK-IBA37plus (LIG1) transformed and grown in *E.coli*, among controls LIG2/3 to determine the percent of transformed LIG1 colonies with inserts: colony numbers of LIG2 (re-ligated and uncut plasmids) and LIG3 (uncut plasmids) should be included in LIG1 (all plasmids including ligated) with logarithmic progression 1:10:100, LIG3:LIG2:LIG1. However, the results instead propose that 0% of grown colonies had plasmid inserts as the ratio of LIG2:LIG1 with averages is 1.5:1 thus any difference is chance, and all plasmids are either uncut or re-ligated.

The transformation competency of *E.coli* can be graded by colony forming units (cfu). The 1×10^{-5} ug of plasmid pASK-IBA37plus in the control had a transformation efficiency of 8.8×10^{6} cfu/ug. This is relatively close to the manufacturer value of $0.9 – 1.5 \times 10^{9}$ under perfect conditions so ligation, not transformation efficiency, was an issue in the experiment.

Word count = 150

DAY 4

5. Restriction enzyme analysis of plasmid DNA purified from colonies resulting from the transformation of the ligation reaction between pASK-IBA37plus and PLAT into *E. coli*.

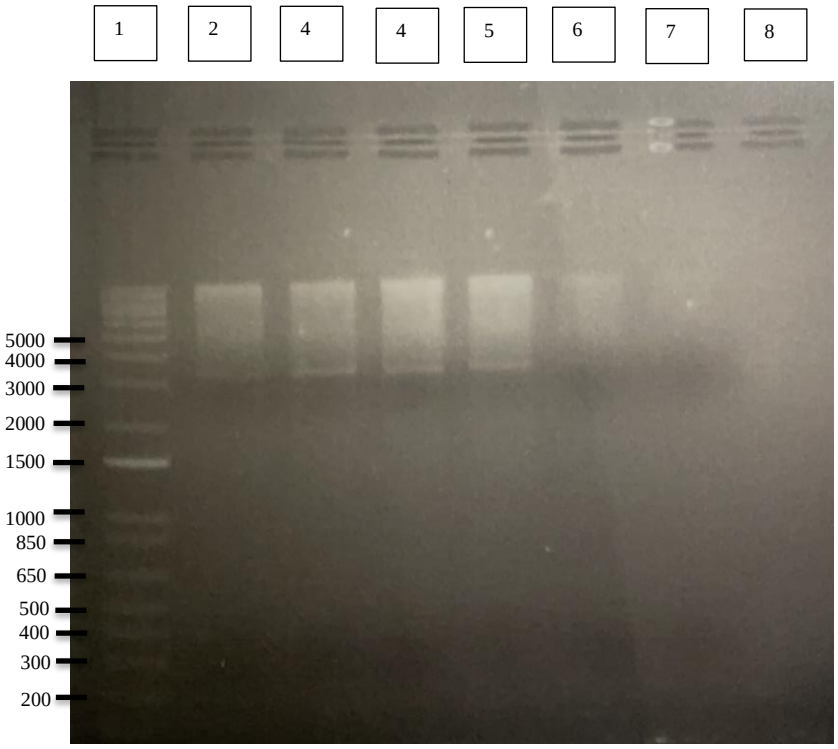


Figure 5. Image showing the 1% w/V horizontal agarose gel electrophoresis of purified LIG1 colony plasmids analysing their restriction enzyme digests.
Lane 1: 1kb plus DNA ladder markers (bp)
Lane 2 (RE1): *Bam*HI and *Hind*III digested plasmid DNA from LIG1 Colony 1. **Lane 3(RE2):** BamHI and *Hind*III digested plasmid DNA from LIG1 Colony 2. **Lane 4 (RE3):** *Bam*HI and *Hind*III digested plasmid DNA from LIG1 Colony 3. **Lane 5 (RE4):** *Bam*HI and *Hind*III digested plasmid DNA from LIG1 Colony 4. **Lane 6 (RE5):** *Bam*HI and *Hind*III digested plasmid DNA from LIG1 Colony 5. **Lane 7 (RE6):** The control *E.coli* with *Bam*HI and *Hind*III digested pASKPLAT **Lane 8:** Empty Lane.

Results and discussion (max 150 words)

Six plasmid colony samples were electrophoresed from LIG1. From Day 3, no plasmids with the insert, PLAT, are expected, evident from no bands around the 1074bp region of PLAT in colony lanes. A smear of DNA is visible from the gel top to 3200bp, likely from RE samples 1-6 of pASK-IBA37plus plasmid being contaminated with protein/salt or excess DNA from the poor lysate purification with the Gel-Elute kit, or from too much colony DNA smearing in excess.

The line at 3000 is distinct and represents 3219bp pASK-IBA37plus linearised by *Bam*HI and *Hind*III - the region removed is only 51 base pairs which is too small to be resolved. The control *E.coli* pASKPLAT was not visible, likely due too human error loading and pipetting as little volume was left in the final tube. Partial digestion in all lanes could be due to an impure plasmid supplied, but the ligation failed nonetheless.

Word count = 150

6. DNA sequence analysis of plasmid pASKPLAT

Figure 6. Map of plasmid pASKPLAT. The truncated PLAT insert (reteplase) is visible in green between the *Bam*HI and *Hind*III sites in the multiple cloning section of the expression vector pASK-IBA37plus. The full reading frame from methionine is visible in grey.

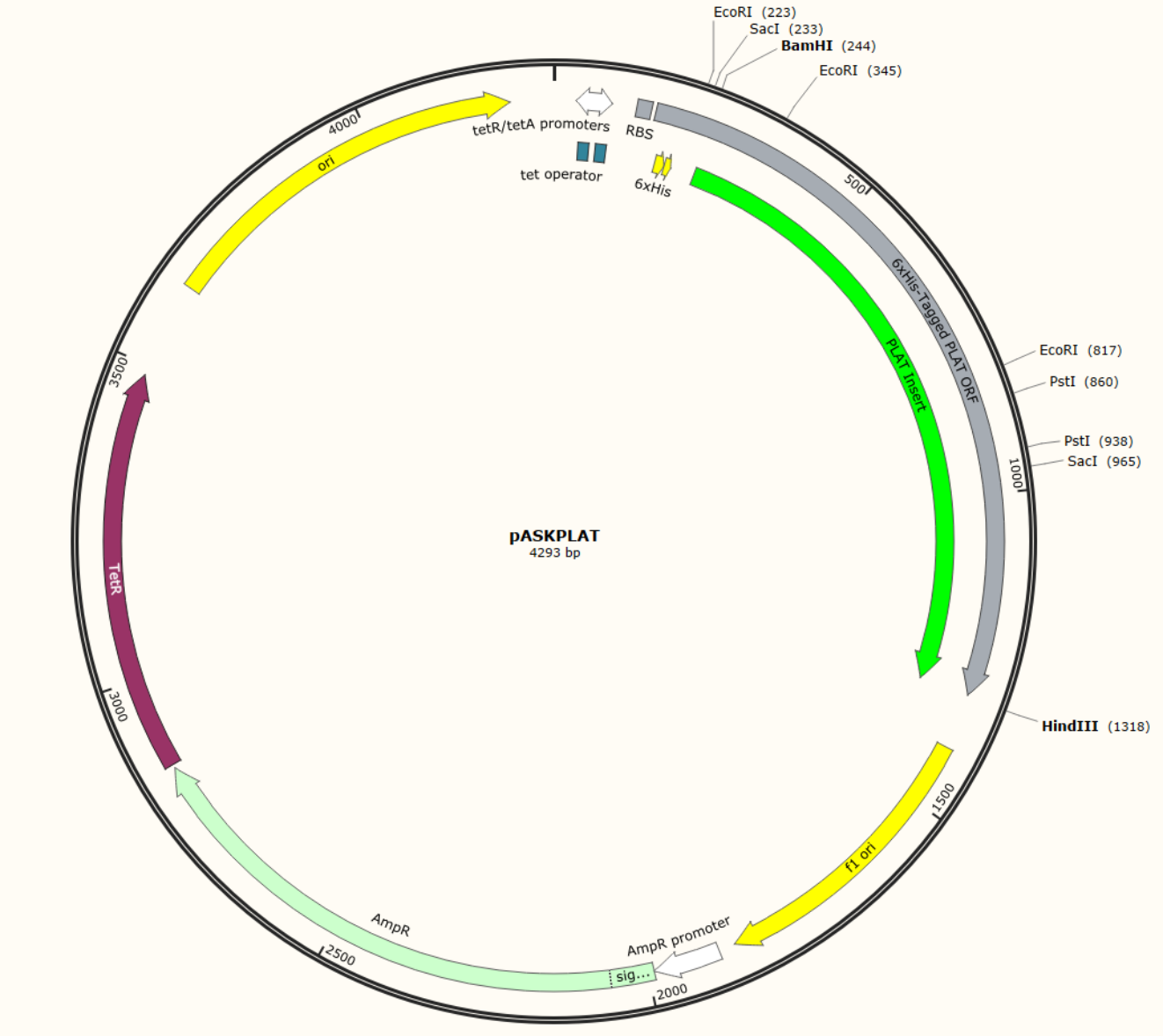


Figure 7. Partial DNA (T replaced for U) sequence of the PLAT insert in pASKPLAT translated to show the three forward reading frames. DNA sequence from sample data no.*01107599* is highlighted in different colours. Grey: The translated region of pASKPLAT upstream of nucleotides from the reteplase ORF. Yellow: 6x Histidine tag. Pink: The Factor Xa site sequence. Green: The recombinant PLAT (reteplase) ORF.

```
1 AGUUUUGGUUGACACUCUAUCAUUGAUAGAGUUAUUUUACCACUCCCUAUCAGUGAUAGA 60

      S  F  G  *  H  S  I  I  D  R  V  I  L  P  L  P  I  S  D  R
       V  L  V  D  T  L  S  L  I  E  L  F  Y  H  S  L  S  V  I  E
        F  W  L  T  L  Y  H  *  *  S  Y  F  T  T  P  Y  Q  *  *  R

 61 GAAAAGUGAAAUGAAUAGUUCGACAAAAAUCUAGAAAUAAUUUUGUUUAACUUUAAGAAG 120

      E  K  *  N  E  *  F  D  K  N  L  E  I  I  L  F  N  F  K  K
       K  S  E  M  N  S  S  T  K  I  *  K  *  F  C  L  T  L  R  R
        K  V  K  *  I  V  R  Q  K  S  R  N  N  F  V  *  L  *  E  G

121 GAGAUAUACAAAUGGCUAGCAGAGGAUCGCAUCACCAUCACCAUCACAUCGAAGGGCGCC 180

      E  I  Y  K  W  L  A  E  D  R  I  T  I  T  I  T  S  K  G  A
       R  Y  T  N  G  *  Q  R  I  A  S  P  S  P  S  H  R  R  A  P
        D  I  Q  M  A  S  R  G  S  H  H  H  H  H  H  I  E  G  R  R

181 GAGACCGCGGUCCCGAAUUCGAGCUCGGUACCCGGGGAUCCUCUUACCAAGGAAACAGUG 240

      E  T  A  V  P  N  S  S  S  V  P  G  D  P  L  T  K  E  T  V
       R  P  R  S  R  I  R  A  R  Y  P  G  I  L  L  P  R  K  Q  *
        D  R  G  P  E  F  E  L  G  T  R  G  S  S  Y  Q  G  N  S  D

241 ACUGCUACUUUGGGAAUGGGUCAGCCUACCGUGGCACGCACAGCCUCACCGAGUCGGGUG 300

      T  A  T  L  G  M  G  Q  P  T  V  A  R  T  A  S  P  S  R  V
       L  L  L  W  E  W  V  S  L  P  W  H  A  Q  P  H  R  V  G  C
        C  Y  F  G  N  G  S  A  Y  R  G  T  H  S  L  T  E  S  G  A

301 CCUCCUGCCUCCCGUGGAAUUCCAUGAUCCUGAUAGGCAAGGUUUACACAGCACAGAACC 360

      P  P  A  S  R  G  I  P  *  S  *  *  A  R  F  T  Q  H  R  T
       L  L  P  P  V  E  F  H  D  P  D  R  Q  G  L  H  S  T  E  P
        S  C  L  P  W  N  S  M  I  L  I  G  K  V  Y  T  A  Q  N  P

361 CCAGUGCCCAGGCACUGGGCCUGGGCAAACAUAAUUACUGCCGGAAUCCUGAUGGGGAUG 420

      P  V  P  R  H  W  A  W  A  N  I  I  T  A  G  I  L  M  G  M
       Q  C  P  G  T  G  P  G  Q  T  *  L  L  P  E  S  *  W  G  C
        S  A  Q  A  L  G  L  G  K  H  N  Y  C  R  N  P  D  G  D  A

421 CCAAGCCCUGGUGCCACGUGCUGAAGAACCGCAGGCUGACGUGGGAGUACUGUGAUGUGC 480

      P  S  P  G  A  T  C  *  R  T  A  G  *  R  G  S  T  V  M  C
       Q  A  L  V  P  R  A  E  E  P  Q  A  D  V  G  V  L  *  C  A
        K  P  W  C  H  V  L  K  N  R  R  L  T  W  E  Y  C  D  V  P

481 CCUCCUGCUCCACCUGCGGCCUGAGACAGUACAGCCAGCCUCAGUUUCGCAUCAAAGGAG 540

      P  P  A  P  P  A  A  *  D  S  T  A  S  L  S  F  A  S  K  E
       L  L  L  H  L  R  P  E  T  V  Q  P  A  S  V  S  H  Q  R  R
        S  C  S  T  C  G  L  R  Q  Y  S  Q  P  Q  F  R  I  K  G  G
```

```
541 GGCUCUUCGCCGACAUCGCCUCCCACCCCUGGCAGGCUGCCAUCUUUGCCAAGCACAGGA 600

      G  S  S  P  T  S  P  P  T  P  G  R  L  P  S  L  P  S  T  G
       A  L  R  R  H  R  L  P  P  L  A  G  C  H  L  C  Q  A  Q  E
         L  F  A  D  I  A  S  H  P  W  Q  A  A  I  F  A  K  H  R  R

601 GGUCGCCCGGAGAGCGGUUCCUGUGCGGGGGCAUACUCAUCAGCUCCUGCUGGAUUCUCU 660

      G  R  P  E  S  G  S  C  A  G  A  Y  S  S  A  P  A  G  F  S
       V  A  R  R  A  V  P  V  R  G  H  T  H  Q  L  L  L  D  S  L
         S  P  G  E  R  F  L  C  G  G  I  L  I  S  S  C  W  I  L  S

661 CUGCCGCCCACUGCUUCCAGGAGAGGUUUCCGCCCCACCACCUGACGGUGAUCUUGGGCA 720

      L  P  P  T  A  S  R  R  G  F  R  P  T  T  *  R  *  S  W  A
       C  R  P  L  L  P  G  E  V  S  A  P  P  P  D  G  D  L  G  Q
         A  A  H  C  F  Q  E  R  F  P  P  H  H  L  T  V  I  L  G  R

721 GAACAUACCGGGUGGUCCCUGGCGAGGAGGAGCAGAAAUUUGAAGUCGAAAAAUACAUUG 780

      E  H  T  G  W  S  L  A  R  R  S  R  N  L  K  S  K  N  T  L
       N  I  P  G  G  P  W  R  G  G  A  E  I  *  S  R  K  I  H  C
         T  Y  R  V  V  P  G  E  E  E  Q  K  F  E  V  E  K  Y  I  V

781 UCCAUAAGGAAUUCGAUGAUGACACUUACGACAAUGACAUUGCGCUGCUGCAGCUGAAAU 840

      S  I  R  N  S  M  M  T  L  T  T  M  T  L  R  C  C  S  *  N
       P  *  G  I  R  *  *  H  L  R  Q  *  H  C  A  A  A  A  E  I
         H  K  E  F  D  D  D  T  Y  D  N  D  I  A  L  L  Q  L  K  S

841 CGGAUUCGUCCCGCUGUGCCCAGGAGAGCAGCGUGGUCCGCACUGUGUGCCUUCCCCCGG 900

      R  I  R  P  A  V  P  R  R  A  A  W  S  A  L  C  A  F  P  R
       G  F  V  P  L  C  P  G  E  Q  R  G  P  H  C  V  P  S  P  G
         D  S  S  R  C  A  Q  E  S  S  V  V  R  T  V  C  L  P  P  A

901 CGGACCUGCAGCUGCCGGACUGGACGGAGUGUGAGCUCUCCGGCUACGGCAAGA 954

      R  T  C  S  C  R  T  G  R  S  V  S  S  P  A  T  A  R
       G  P  A  A  A  G  L  D  G  V  *  A  L  R  L  R  Q
         D  L  Q  L  P  D  W  T  E  C  E  L  S  G  Y  G  K
```

<u>Results and discussion (max 150 words) – *Sample data: 01107599-ASKtetpro.ab1*</u>

Fig.6 visualises pASKPLAT. Fig.7 is the partial DNA sequence of a *sample data* pASKPLAT (ligation failed, fig.5) to ensure correct localisation of PLAT from experimental data via BLAST. Lack of alignment with homo-sapiens PLAT begins at nucleotide 958 of the sequence due to a drop in Sanger sequencing fluorescence and accuracy so it's stopped; the first bases are also unreliable as there aren't enough short ddNTP terminated DNA fragments to fluoresce and produce distinct peaks. SYQ was incorporated as it's after the primer restriction enzyme site, unlike aaa nucleotides.

This figure shows the 6xHis-tag is in frame with the 3rd frame of the PLAT insert which suggests that the ligation and transformation was successful. There appear to be no mutations and premature stop codons in this data, but the full sequence of the PLAT insert sequenced with a reverse primer (as well as ASKtetpro) can identify potential problems throughout the ORF.

Word count = 150

DAY 5

7. Tetracycline induction of *E. coli* pASKPLAT and E. coli pASK-IBA37plus (no insert control)

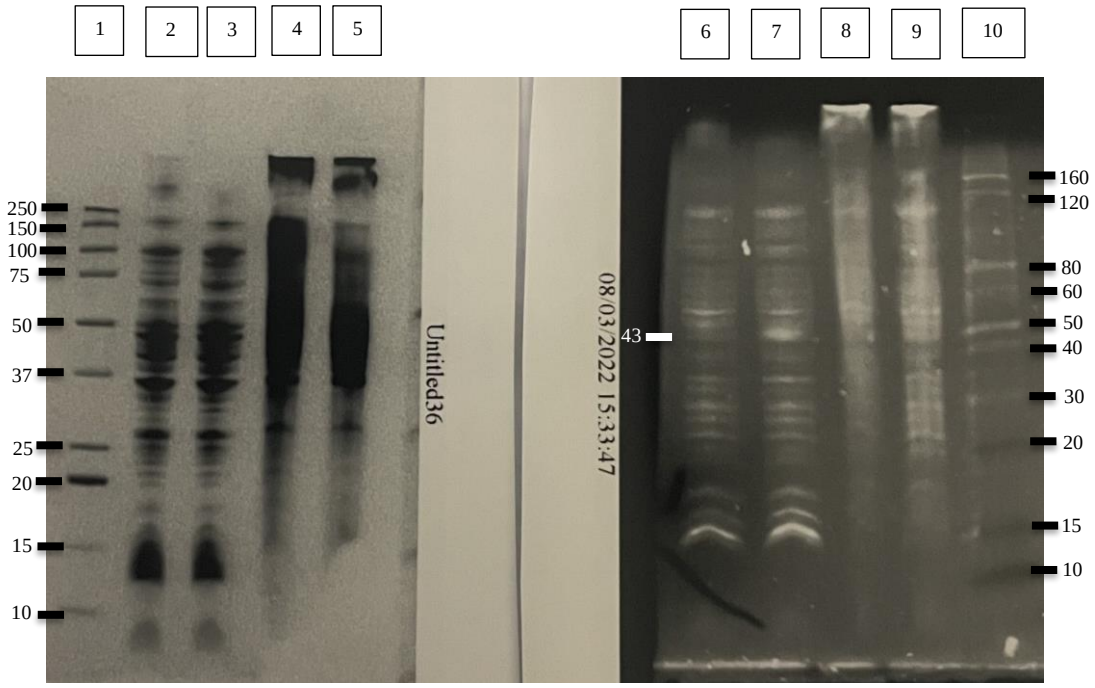


Figure 8. The analysis of restriction enzyme digests of pASKIBA37-plus and PLAT PCR fragment in 20ul solutions using SDS-PAGE to identify the His-tagged reteplase protein. Lane 1: BioRad Marker (kD). **Lane 2 (PAIR 1):** Tube CL1 (*E.coli* pASK-IBA37plus tet induction - no insert control) with Biosafe Coomassie stain. **Lane 3 (PAIR 1)** Tube CL2 (*E.coli* pASKPLAT tet induction – with insert) with Biosafe Coomassie stain . **Lane 4 (PAIR 2):** Tube CL1 with Biosafe Coomassie stain. **Lane 5 (PAIR 2):** Tube CL2 with Biosafe Coomassie stain. **Lane 6 (PAIR 1):** Tube CL1 with Invitrogen Invision His-tag In-gel stain. **Lane 7 (PAIR 1)**: Tube CL2 with Invitrogen Invision His-tag In-gel stain. **Lane 8 (PAIR 2):** Tube CL1 with Invitrogen Invision His-tag In-gel stain. **Lane 9 (PAIR 2):** Tube CL2 with Invitrogen Invision His-tag In-gel stain. **Lane 10:** Invitrogen His Marker (kDa)

Results and discussion (max 140 words) ***(using PAIR 2 as well as our own (PAIR 1))***

Fig.8 shows the protein expression of PAIR 1 and 2 via SDS-PAGE. PAIR 2 lanes have smearing due to proteins and salt from bad colony purification and no expression as the ligation for pASKPLAT failed. Coomassie was used (Lane 1-5) to show all proteins and Invision His-tag in-gel stain (Lane 6-10) for His-rich proteins. The SDS detergent denatures proteins in CL1 and CL2 to be sized properly on the polyacrylamide gel. The expected size of the 6x His-tagged recombinant tPA protein is 43 kDa and the white label shows that in fig.8. The band can't be distinguished amongst proteins in the Coomassie stain.

As expected, this band was only visible in Lane 7(CL2, Pair 1) and not Lane 6(CL1, Pair 1) as, although the tetracycline promotor is induced in both, CL1 is pASK-IBA37plus which cannot produce the full 43kDa His-tagged tPA protein.

Word count = 140

8. Future work (see Pro Forma Guidelines document for information on this section)

Purification of 6xHis tagged PLAT (max 400 words)

The recombinant PLAT gene (r-tPA) was isolated from pOTB7PLAT, inserted into pASK-IBA37plus and transformed into *E.coli* forced to express 6xHis-tagged PLAT as reteplase via tet promotors. IMAC (immobilised metal-ion-affinity chromatography) is specialised to separate His-tagged proteins via immobilisation of His residues on an Ni^{2+} chelating surface resin [Renatus *et al*., 1997]. *E.coli* produces reteplase on a large scale relative to mammalian cultures so a moderate throughput Ni-NTA spin kit of up to 600ug lysate is used. Due to reteplase's expression in soluble form and the inefficient refolding process of its several disulphide bonds, it will be purified under native conditions [Long *et al*., 2015]

The *E.coli* lysate is purified by 3 native buffers with 50mM NaH_2PO_4, 300mM NaCl but varying amounts of imidazole: a lysate buffer with 10mM, wash buffer with 20mM and elution buffer with 500mM. Buffers are below PH 8.4 so the Ni^{2+}-NTA silica is inert. Purification begins with 5ml of native protein culture mixed in lysis buffer and Benzonase© nuclease to degrade residual nucleic acids and meet recombinant protein requirements [*SigmaAldrich*, 2022]. 600µl lysate is mixed with lysis buffer of 10mM imidazole that displaces single His residues from native *E.coli* proteins but not 6xHis-tagged proteins which outcompete and bind the chelated Ni^{2+} affinity column resin. Spinning for supernatant (contaminating protein) aspiration must not exceed 1600rpm or kinetics will reduce the resin-6xHis interactions, and a closed centrifuge lid reduces flow rate for binding to the folded protein (denatured proteins are better attracted). Wash buffer (20mM imidazole) displaces stronger bound single or double His residues in the remaining lysate solution and centrifugation at 2900rpm is used to aspirate the next supernatant. After 3 washes, reteplase is competitively eluted with elution buffer (500mM imidazole) at 2900rpm in 300µl fractions. The 'wash' supernatants and final 600µl product obtained can be compared with SDS-page [*Qiagen*, 2008].

If the aim after IMAC is to obtain a high, moderate purity yield (>80%) of r-tPA, a desalting buffer exchange column is run to remove imidazole and toxic Ni^{2+} salts before sample volume reduction. A protein of highest purity (>99%) is needed from a 3-step protocol that ensures our reteplase is suitable for protein structural studies *in-vitro* and *in-vivo*, regardless of less yield. This will first require a step of IEX (ion exchange chromatography) that requires a specific IEX buffer exchange, followed by size exclusion chromatography for a moderate yield of 99% purity r-tPA, characterised by elution peaks [*Cytiva*, 2022].

Words count: 403

Testing 6xHis tagged PLAT suitability as a drug *in vitro* and *in vivo* (max 600 words)

The 99% purified recombinant 6xHis-tagged tPA (r-tPA) was obtained after the final step of size exclusion chromatography. The relationship between elution peaks and corresponding molecular weights infers r-tPA (reteplase) as a monomer for its structural and active unit for thrombolysis in physiological solution [Long *et al*., 2015].

R-tPA lasts longer than tPA in the plasma. The active folding of r-tPA and its basic function as a serine protease tissue-type plasminogen activator can be assessed with fibrin plate assays *in vitro* with plasminogen substrate: administered r-tPA will bind and cleave the plasminogen, producing plasmin which will work on the fibrin meshwork bound to the plate and produce a lysed zone. A positive control of urokinase can be used to compare activity empirically [Long *et al*., 2015]. Human blood in 1ml paediatric Penrose tubes can extend testing *in vitro*; clots will be induced by coagulation factors factor FXIII, fibrinogen and factor V and the thrombolysis in the presence of r-tPA can be measured quantitively as a relative increase in suspended clot fragment (D-dimer) concentration formed by fibrin degradation. A value moderately greater than a coagulated control (~0.5 without r-tPA) will gauge the efficiency of r-tPA as a plasminogen activator [*Cleveland Clinic*, 2021] [Frenkel *et al.,* 2006].

Strokes are caused by blood clots which affect the brain blood supply. Serine protease tissue-type plasminogen activator (tPA) is the only FDA-approved treatment for ischemic stroke which causes intracerebral haemorrhage. The recommended dose is 0.9mg/kg with an IV infusion but maximum human dose is 90 mg [Katz *and* Tadi., 2021]. It would be tested *in vivo* on mice by *in situ* injection of thrombin into the middle cerebral artery until 60% of blood supply is restricted at baseline. It is important to add r-tPA no later than 3 hours after stroke onset: its use 3-4.5 hours after has not been studied. The effect of r-tPA administered mice can be compared to urokinase and saline control mice using MRI after 24 hours stroke onset: T2-weighted MRI images and ImageJ software can be used to quantify sizes of lesions, located by the mouse stereotaxic atlas. The weighted mean difference in between the r-tPA and control groups tumours in mm^3 can be calculated to gauge the effect of tissue plasminogen on thrombolysis of cerebral clots *in vivo* [Orset *et al.,* 2016].

If the r-tPA administered is inactive *in vivo*, this could be due overexpression of the principal serine protease inhibitor, plasminogen activator inhibitor-1 (PAI-1). An MTT assay can assess metabolic activity of endothelial cells from an affected mouse to quantify their PAI-1 production calorimetrically against a control [*SigmaAldrich*, 2021]. The activity of the drug *in vivo* can be improved with the addition of a C-terminally fused RGDS integrin binding motif which improves r-tPA affinity to plasminogen [Long *et al*., 2015]. The His-tag is generally inert but may interfere with the downstream r-tPA function causing inactivity; the incorporated Factor Xa Protease site easily removes the 6x-His tag with a Prescission Protease (PSP) without denaturation risks [Hui *et al.,* 2017].

If r-tPA is inactive *in vitro*, there are no external factors and failure is likely due to protein misfolding into insoluble inclusion bodies: 5 distinct structural domains and 17 disulphide bonds in r-tPA mean this is common in the reductive *E.coli* environment. New techniques are being developed obtaining active r-tPA via secretion into the oxidising (human-like) periplasm of *E.coli* suitable for disulphide bond formation [Qui *et al.,* 1998] [Zhan *et al.,* 1999]. A mammalian protocol may have to be considered with superior eukaryotic post-translational processing mechanisms but a smaller expression scale of 0.01 recombinant protein to 1% of total [*Qiagen*, 2003].

Word count: 600

Reference list

Renatus, M., Bode, W., Huber, R., Stürzebecher, J., Prasa, D., Fischer, S., Kohnert, U., & Stubbs, M. T. (1997). Structural mapping of the active site specificity determinants of human tissue-type plasminogen activator. Implications for the design of low molecular weight substrates and inhibitors. *The Journal of biological chemistry*, *272*(35), 21713–21719.

Long, X., Gou, Y., Luo, M., Zhang, S., Zhang, H., Bai, L., Wu, S., He, Q., Chen, K., Huang, A., Zhou, J., & Wang, D. (2015). Soluble expression, purification, and characterization of active recombinant human tissue plasminogen activator by auto-induction in E. coli. *BMC biotechnology*, *15*, 13.

SigmaAldrich. (2022). *Benzonase© Nuclease*. Retrieved from sigmaaldrich.com: https://www.sigmaaldrich.com/deepweb/assets/sigmaaldrich/product/documents/177/968/benzonase-nuclease-mk.pdf

Qiagen. (2008). *Ni-NTA Spin Handbook.* Retrieved from Qiagen.com: https://www.qiagen.com/gb/resources/resourcedetail?id=3fc8c76d-6d21-4887-9bf8-f35f78fcc2f2&lang=en

Cytiva. (2022). *Purifying histidine-tagged protein.* Retrieved from Cytvialifesciences.com: https://www.cytivalifesciences.com/en/us/solutions/protein-research/knowledge-center/protein-purification-methods/how-to-combine-chromatography-techniques/his-tagged-protein-purification-protocols

ClevelandClinic. (2021). *D-Dimer test.* Retrieved from Clevelandclinic.com: https://my.clevelandclinic.org/health/diagnostics/22045-d-dimer-test

Frenkel, V., Oberoi, J., Stone, M. J., Park, M., Deng, C., Wood, B. J., Neeman, Z., Horne, M., 3rd, & Li, K. C. (2006). Pulsed high-intensity focused ultrasound enhances thrombolysis in an in vitro model. *Radiology*, *239*(1), 86–93.

Katz JM, Tadi P. Physiology, Plasminogen Activation. [Updated 2021 Oct 1]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan-.

Orset, Cyrille et al. “Efficacy of Alteplase in a Mouse Model of Acute Ischemic Stroke: A Retrospective Pooled Analysis.” *Stroke* vol. 47,5 (2016): 1312-1318.

SigmaAldrich. (2021). *MTT assay protocl*. Retrieved from sigmaaldrich.com: https://www.sigmaaldrich.com/GB/en/technical-documents/protocol/cell-culture-and-cell-culture-analysis/cell-counting-and-health-analysis/cell-proliferation-kit-i-mtt

Hui Chin Goh, Radoslaw M. Sobota, Farid J. Ghadessy, Saurabh Nirantar,
Going native: Complete removal of protein purification affinity tags by simple modification of existing tags and proteases, Protein Expression and Purification, Volume 129, 2017, Pages 18-24, ISSN 1046-5928.

Qiu, J et al. “Expression of active human tissue-type plasminogen activator in Escherichia coli.” *Applied and environmental microbiology* vol. 64,12 (1998): 4891-6. doi:10.1128/AEM.64.12.4891-4896.1998.

Qiagen. (2003). *QIAexpressionist.* Retrieved from Qiagen.com: https://www.qiagen.com/gb/resources/resourcedetail?id=79ca2f7d-42fe-4d62-8676-4cfa948c9435&lang=en

Zhan X, Schwaller M, Gilbert HF, Georgiou G: Facilitating the formation of disulfide bonds in the Escherichia coli periplasm via coexpression of yeast protein disulphide isomerase. Biotechnol Prog. 1999, 15: 1033-1038.