Maria Margarida dos Santos Silva
Licenciada

Assessment of Notch1 Ligands production - Key Protein Targets in Breast Cancer

Dissertação para a obtenção do grau de Mestre em
Genética Molecular e Biomedicina

Orientador: Doutor Tiago M. Bandeiras, PhD, iBET
Co-orientador: Doutora Ana Barbas, PhD, iBET

Júri:

Presidente: Prof. Paula Gonçalves, FCT-UNL

Arguente: Prof. Dr. Carlos Salgueiro, FCT-UNL

Vogal: Doutor Tiago M. Bandeiras, iBET

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro de 2014

Maria Margarida dos Santos Silva
Licenciada

Assessment of Notch1 Ligands production - Key Protein Targets in Breast Cancer

Dissertação para a obtenção do grau de Mestre em
Genética Molecular e Biomedicina

Orientador: Doutor Tiago M. Bandeiras, PhD, iBET
Co-orientador: Doutora Ana Barbas, PhD, iBET

Júri:

Presidente: Prof. Paula Gonçalves, FCT-UNL

Arguente: Prof. Dr. Carlos Salgueiro, FCT-UNL

Vogal: Doutor Tiago M. Bandeiras, iBET

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro de 2014

Assessment of Notch1 Ligands production - Key Protein Targets in Breast Cancer

Acknowledgements

I would like to thank my dearest supervisors Dr. Tiago M. Bandeiras and Dr. Ana Barbas. I am thankful for their guidance, invaluably constructive criticism and friendly advice during this work. I am sincerely grateful for the autonomy and the confidence they always had in my work, while always being there for insights and discussions. Without their supervision and trust this dissertation would not have been possible.

I am thankful to all the lab members in both labs: Ana Teresa Gonçalves, Cristiana Sousa, Inês Barbosa, Joana Galvão, Maria Jardim Cabral, Maria João Proença, Paulo Espírito Santo, Sandra Monteiro and Sara Silva. In some moment all of you contributed somehow to my thesis work and taught me something new. A special thanks goes to my team mate and friend, Micael Freitas, who taught me how to work with ÄKTA systems, and finished a protein purification while I had to suddenly leave the lab for a couple of days. I would also like to thank all the members of the Macromolecular Crystallisation Unit and the Mass Spectrometry Unit from ITQB/UNL.

I thank FCT (PTDC/SAU-ONC/121670/2010), P-CUBE (Capacities Specific Programme Research Infrastructures, Project Number 227764) and NRS/LPCC “Terry Fox” 2013/2014 grants for funding, and iBET for hosting.

I am grateful to my fiancé Filipe, for his love, support, comprehension and encouragement during the past year for me to finalize this dissertation, and most importantly for always being around to listen to me in the good, the bad, and the terrible days. Last, but not least, I would like express my gratitude to my parents and all my family for their endless love and support throughout my life.

Resumo

A comunicação celular é fundamental para os processos biológicos no desenvolvimento e na vida adulta. Um dos sistemas mais utilizados por eucariotas é a via de sinalização Notch. Existem quatro recetores e cinco ligandos em mamíferos, cujos domínios extracelulares interagem, ativando a transcrição. Foi demonstrado que a expressão dos ligandos não é detetável em tecidos normais, mas que expressão moderada a elevada foi detetada em cancro da mama. Assim, qualquer um dos ligandos da via Notch pode ser estudado como possível alvo terapêutico.

O nosso objetivo é determinar as melhores condições de expressão e purificação para os ligandos da via Notch para poder desenvolver anticorpos contra essas proteínas pela tecnologia de "Phage-Display". Queremos ainda cristalizar os ligandos do Notch e também o(s) complexo(s) com os anticorpos selecionados, revelando detalhes moleculares.

Para estudar proteínas é necessário obtê-las de forma estável em solução. *E. coli* é o hospedeiro de excelência para expressar proteínas recombinantes dada a facilidade, rápido crescimento e altas densidades celulares. No entanto, a expressão de proteínas de mamíferos em tais sistemas pode sobrecarregar a maquinaria celular bacteriana, que não possui a capacidade para modificações pós-translacionais ou compartimentos específicos para a síntese proteica. As células de mamíferos são portanto preferidas, embora as exigências técnicas e financeiras sejam maiores.

As proteínas hJag2DE3 e hDll1DE6 foram purificadas a partir de corpos de inclusão e sobrenadante de cultura de células de mamífero, respetivamente, e a pureza confirmada por SDS-PAGE (> 95%). A proteína produzida em células de mamífero mostrou-se mais estável e aparenta ter o padrão fisiológico de pontes dissulfureto, ao contrário do que foi observado na proteína "refolded". Diversas cristalizações em nano-escala foram realizadas em placas de 96 poços, sem resultados positivos.

Iremos continuar a estudar a melhor condição de expressão para os ligandos do Notch em ambos os sistemas de expressão.

Palavras-chave: Via de sinalização Notch, Cancro da mama, Expressão de proteínas recombinantes, cristalização de macromoléculas.

Abstract

Cell-to-cell communication is required for many biological processes in development and adult life. One of the most common systems utilized by a wide range of eukaryotes is the Notch signalling pathway. Four Notch receptors and five ligands have been identified in mammals that interact *via* their extracellular domains leading to transcription activation. Studies have shown that the Notch ligands expression is undetectable in normal breast tissues, but moderate to high expression has been detected in breast cancer. Thus, any of the Notch1 ligands can be studied as possible therapeutic targets for breast cancer.

To study Notch pathway proteins there is the need to obtain stable protein solutions. *E. coli* is the host of excellence for recombinant proteins for the ease of use, fast growth and high cell densities. However, the expression of mammalian proteins in such systems may overwhelm the bacterial cellular machinery, which does not possess the ability for post-translational modifications, or dedicated compartments for protein synthesis. Mammalian cells are therefore preferred, despite their technical and financial increased demands.

We aim to determine the best expression and purification conditions for the different ligand protein constructs, to develop specific function-blocking antibodies using the Phage Display technology. Moreover, we propose to crystallize the Notch1 ligands alone and in complex with the phage display selected antibodies, unveiling molecular details.

hJag2DE3 and hDll1DE6 proteins were purified from refolded inclusion bodies or mammalian cell culture supernatants, respectively, and purity was confirmed by SDS-PAGE (>95%). Protein produced in mammalian cells showed to be more stable, apparently with the physiological disulfide pattern, contrary to what was observed in the refolded protein. Several nano-scale crystallization experiments were set up in 96-well plates, but no positive result was obtained.

We will continue to pursue for the best expression for the Notch ligand constructs in both expression systems.

Keywords: Notch signalling, Breast cancer, Recombinant protein expression, Macromolecular crystallisation.

Contents

List of figures

List of tables

Abbreviations

a.a.	Aminoacid
A.U.	Arbitrary Units
ADAM	A Disintegrin and A Metalloprotease
Amp	Ampicilin
ANK	Ankyrin Repeats
BME	beta-mercaptoethanol
bp	base pair
BSA	Bovine Serum Albumin
cbEGF	Calcium-binding EGF
CD	Circular Dichroism
Cm	Chloramphenicol
CMV	Citomegalovirus
CR	Cysteine Rich Domain
CV	Column Volume
Da	Dalton
DMEM	Dulbecco's modified Eagle's medium
DNA	Deoxyribonucleic Acid
dNTPs	deoxy nucleotides tri-phosphate
DOS	Delta and OSM-11 Motif
dpt	days post transfection
DSF	Differential Scanning Fluorimetry
DSL	Delta/Serrate/LAG-2 Familiy
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
EGF	Epidermal Growth Factor-like Repeat
FBS	Fetal Bovine Serum
FCS	Fetal Calf Serum
GSH	Reduced Glutathione
GSI	gamma-Secretase Inhibitor
GSSG	Oxidised Glutathione
GST	Gluthathione S-Transferase
GuHCl	Guanidine Hydrochloride
HD	Heterodimerization Domain
HEK	Human Embrionic Kidney
hpt	hours post transfection
HRP	Horseradish Peroxidase
HVR-3C	Human Rhinovirus 3C protease
IB	Inclusion Body
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
kb	kilo basepair
kDa	kilo Dalton
LNR	Lin12-Notch Repeats
MBP	Maltose Binding Protein
MNNL	Module N-Terminal of Notch Ligands
MW	Molecular Weight
MWCO	Molecular Weight Cut-off
NECD	Notch Extracellular Domain
NICD	Notch Intracellular Domain
NLS	Nuclear Localization Signal
NRR	Negative Regulatory Region
o/n	overnight
OD_{600}	Optical density at 600nm
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffer Saline
PBS-T	PBS with 0.05% Tween-20
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEI	Polyethylenimine
PEST	Proline, Glutamate, Serine and Threonine rich Sequence
pI	Isoelectric Point
PMSF	Phenylmethylsulfonyl Fluoride
RAM	Rbp-Associated Molecule
rpm	rotations per minute
RT	Room Temperature
SDS	Sodium Dodecyl Sulphate
SEC	Size Exclusion Chromatography
SUMO	Small Ubiquitin Modifier
TACE	TNF Converting Enzyme
T-ALL	T-Cell Acute Lymphoblastic Leukemia
TGF-β	Transforming Growth Factor beta
Tm	Melting Temperature
TMAO	Trimethylamine N-oxide
tRNA	transfer Ribonucleic acid
Trx	Thioredoxin
UV	Ultraviolet
VEGF	Vascular Endothelial Growth Factor
WB	Western Blot
xg	Gravitational Force

Chapter I Introduction

The Notch Signalling Pathway is a highly conserved cell signalling system present in most multicellular organisms. Signals exchanged between neighbouring cells through the Notch receptor can amplify and consolidate molecular differences, which eventually dictate cell fates. Thus, Notch signals control how cells respond to intrinsic or extrinsic developmental cues that are necessary to unfold specific developmental programs (Artavanis-Tsakonas, Rand *et al.* 1999). The Notch signalling pathway plays a crucial role in multiple cellular processes including stem cell self-renewal, cell fate determination, epithelial cell polarity/adhesion, cell division and apoptosis (Wu, Cain-Hom *et al.* 2010). Different Notch receptors with specific functions have been identified but, interestingly, they all trigger the activation of the same downstream cascade.

I.1. Notch signalling pathway overview

The Notch signalling pathway involves direct contact of two adjacent cells with distinct molecular signatures; one cell expressing the Notch receptor and the other the Notch ligand (Figure I.1). Both Notch receptors and ligands are type I transmembrane proteins with a series of tandem Epidermal Growth Factor-like Repeats (herein referred as EGFs). For signalling to be productive, the receptor and ligand binding through the EGF domains takes place in *trans*, ie. both proteins are expressed in different cells and the polypeptide chains are inversely oriented respective to each other. When both receptor and ligand proteins are expressed in the same cell, proteins are able to bind in *cis*, albeit resulting in signalling inhibition.

At a given time, a cell can only be in a "sending" state, where predominantly the ligand is expressed on the cell surface, or in a "receiving" state, where the receptor is the most abundant (Figure I.1). A model with the two mutually exclusive states for Notch signalling was previously proposed (Sprinzak, Lakhanpal *et al.* 2010) to explain both graded *trans*-activation and sharp *cis*-inhibition: in the "signal receiving" cell the Notch receptor binds to the ligand of the "signal sending" cell, and is activated in a ligand exposition dependent manner. Consequently the Notch Intracellular Domain (NICD) migrates to the nucleus where it participates in transcription complexes; alternatively, the Notch receptor in the "signal receiving" cell also binds irreversibly to the ligand expressed in the same cell, and the complex is internalised without NICD release. This model is in agreement with the previously described requirement of Notch Extracellular Domain (NECD) *trans*-endocytosis by the ligand expressing cell for Notch receptor proteolysis events (Nichols, Miyamoto *et al.* 2007).

The very first proteolytic event of the Notch receptor polypeptide chain takes place in the Golgi apparatus by a Furin-like protease, generating a mature receptor in a heterodimeric form joined by non-covalent interactions (Kopan and Ilagan 2009). As depicted in Figure I.1, upon ligand binding Notch receptor is cleaved at two sites in a consecutive order - first by ADAM/TACE at site 2 (S2) and afterwards by γ-secretase at site 3 (S3). These three proteolytic events are necessary, in that order, for active NICD generation and in order for productive signalling to occur. When not processed by Furin in the Golgi, the Notch receptor exists in the immature form at the cell surface and, although

capable of ligand binding, it fails to generate productive signalling due to losses in receptor dissociation. ADAM proteolysis at S2 is required for NICD formation given that if ADAM proteolysis is inhibited, gamma-secretase is not able to cleave at S3, thereby releasing the active NICD to the cytoplasm (Nichols, Miyamoto *et al.* 2007).

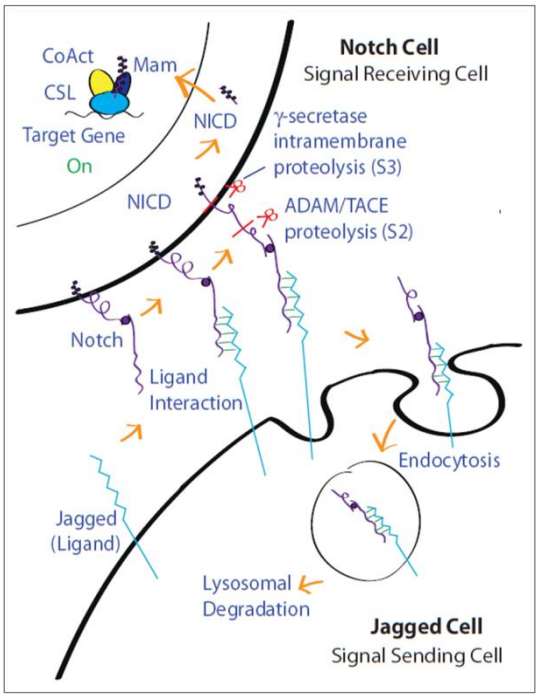


Figure I.1 - Overview of Notch canonical signalling pathway. Notch signalling is initiated by the interaction of Notch receptor with DSL ligands (Jagged and Delta in mammalian cells) *via* their extracellular domains. This interaction requires two adjacent cells, a signal sending cell (bottom corner) expressing a ligand and a signal receiving cell (top corner) expressing the receptor. Upon ligand binding, Notch receptor is proteolysed by ADAM protease, permitting a second, intramembrane cleavage event by gamma-secretase. Notch intracellular domain (NICD) is then able to translocate to the nucleus where it forms a transcription activating complex. Endocytosis of Notch extracellular domain (NECD) by the signal sending cell (*trans*-endocytosis), is thought generate the pulling forces needed to mediate a conformational change in the receptor's negative regulatory region (NRR) (Nichols, Miyamoto *et al.* 2007), where the proteolysis sites are buried, making S2 and S3 sites accessible for proteases. Figure from Wang (2011).

Notch signalling occurs through the direct interaction of one ligand protein with one receptor protein, releasing one NICD that will form one transcription activation complex. This specifies Notch signalling as a time-, space- and dose-dependent pathway (Rizzo, Osipo *et al.* 2008, South, Cho *et al.* 2012).

I.2. 100 year-lesson on Notch

Nearly a century after the description of a notched wing phenotype in *Drosophila melanogaster* by John S. Dexter in 1914, the scientific community is now starting to understand the complex and intricate Notch signalling pathway. The last 20 years of research contributed to a gradual increase on the knowledge of the roles, molecular players and interplays of Notch signalling.

Using *D. melanogaster* as a model organism, Artavanis-Tsakonas' group reported Notch receptor-ligand interactions *via* the proteins' extracellular region domains when expressed in adjacent

cells (in *trans*) and proposed that protein interactions could also occur in *cis* (when both receptor and ligand proteins are expressed in the same cell) (Fehon, Kooh *et al.* 1990).

Shortly after, Notch signalling was implicated in neural development (Heitzler and Simpson 1993) and the signal transduction mechanism assigned to its intracellular domain, whereas signal regulation would occur through the Notch extracellular domain (Greenwald 1994). Some 20-year-old research and review articles from the *era* of Notch signalling ground-breaking discoveries compiled knowledge that is still widely accepted in the field (Heitzler and Simpson 1993, Zagouras, Stifani *et al.* 1995, Robey 1997, Fleming 1998), and these reports were the first reasoning of a great amount of data in *D. melanogaster*, *C. elegans* and mammalian systems.

I.3. Notch receptor and ligand proteins

In mammals, four Receptors (Notch1-4) and five Ligands (Jagged1/2 – Jag1/2, and Delta-like protein 1, 3 and 4 - Dll1, Dll3 and Dll4) have been identified, as depicted in Figure I.2. All family members are type I transmembrane proteins that contain 6-36 EGF-like repeats in the extracellular domain.

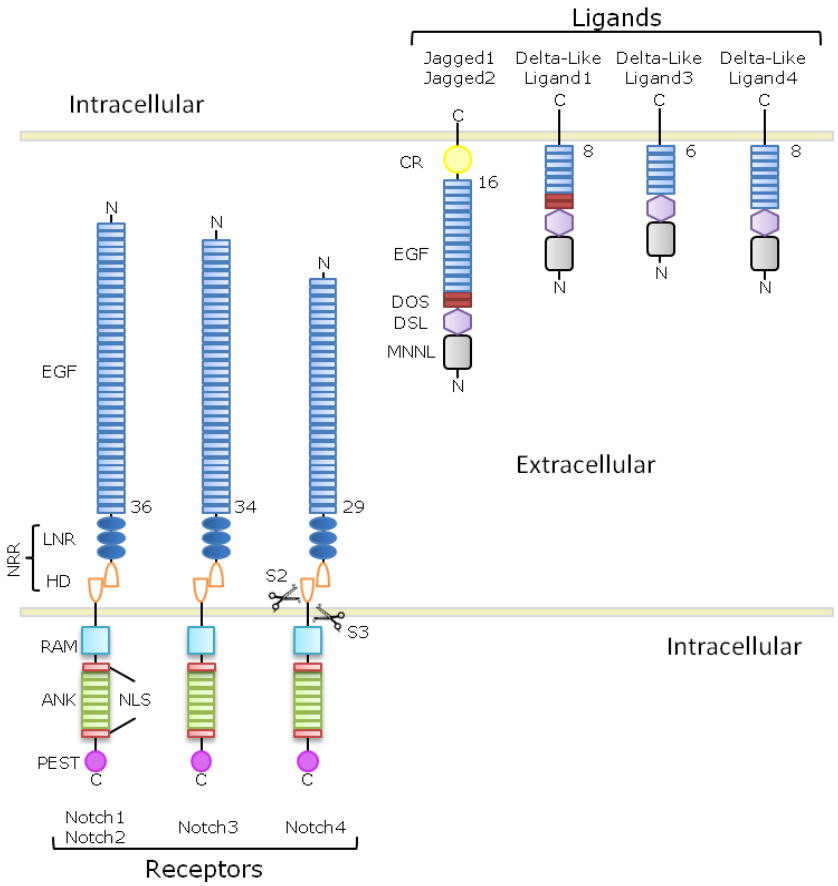


Figure I.2 – Domain organization of the canonical regulators of Notch signalling pathway. Adapted from Kopan and Ilagan (2009), Wang (2011) and South, Cho *et al.* (2012).

In addition, Notch signalling protein members have the following molecular features:

1) The receptor proteins have a Negative Regulatory Region (NRR) in the extracellular domain composed by three Lin12-Notch Repeats (LNR) and an Heterodimerization Domain (HD), harbouring

the S2 and S3 proteolytic sites. Notch Intracellular Domain (NICD) contains Nuclear Localization Signals (NLS) that target it to the nucleus, a Rbp-associated Molecule (RAM), and seven Ankyrin Repeat domains (ANK) that are capable of protein binding and transcription activation. These receptors also contain a C-terminal PEST domain that targets NICD for degradation.

2) The Notch ligands have a Module N-terminal of Notch Ligands (MNNL) followed by a Delta/Serrate/LAG-2 (DSL) domain. The ligands can be subdivided in DOS/DSL ligands, containing the Delta and OSM-Like domains (DOS) (Jag1/2 and Dll1), and DSL only ligands. Jagged1 and 2 have an additional Cysteine-Rich (CR) domain (Schofield, Pope *et al.* 2007).

As stated above, receptor-ligand interactions occur *via* the extracellular EGF-Like repeats. The minimal binding region has been assigned to EGF11 to 13 of Notch receptor, and to DSL to EGF3 of Jagged proteins (Shimizu, Chiba *et al.* 1999) by a series of cell-based and solid-phase binding assays, using combinations of mutated Jagged1 and receptor proteins, in a calcium-dependent manner. Previous studies in *D. melanogaster* had already demonstrated that Notch signalling proteins interacted *via* the extracellular domains with specificity, and that Notch EGF 11-12 are both necessary and sufficient for binding, constituting the receptor's minimal binding region (Fehon, Kooh *et al.* 1990, Rebay, Fleming *et al.* 1991). Furthermore, Notch receptor interacts with more than one *Drosophila* ligand in both cell based and *in vivo* assays (Rebay, Fleming *et al.* 1991). The importance of other Notch domains (such as EGFs 24-30, *Abruptex* region) was proposed based on *in vivo* observations in mosaic *Drosophila* models (Heitzler and Simpson 1993), expressing both *Abruptex* region mutant and wild-type Notch receptor. By analysing the cell fate decisions in the cell margins it was postulated that only the *Abruptex* mutant cells are influenced by the *wt* neighbour cells and not the other way around, reinforcing Notch role as a receptor. Recently, Dighe and colleagues demonstrated that Notch EGFs 11-15 in fact interact *in vitro* with EGFs 21-30 with a 10-fold higher affinity than for its ligands. The critical domains in *Abruptex* region for ligand binding and receptor activation was narrowed to Notch EGFs 25-26 in antibody-based experiments where a blockage of this two EGF repeats of Notch Receptor abolished ligand binding to Notch EGFs 11-13. Therefore, they proposed that in the basal state Notch receptor may be kept in an inactive form by tight interaction between these two distant regions of Notch (EGFs 25-26 and 11-13), which can be disrupted upon ligand binding, causing a conformational change permitting receptor activation (Sharma, Rangarajan *et al.* 2013).

To better understand the role of the extracellular domains of the Notch signalling components, Cordle, Johnson *et al*, (2008) solved the first X-Ray crystal structure of Jagged1 comprising the minimal binding region, ie., DSL through EGF3 domains, and recently a structure comprising also the MNNL domain was reported (Chillakuri, Sheppard *et al.* 2013) (Figure I.3). The structures reveal a near-linear arrangement of the domains, irrespective of the MNNL domain presence.

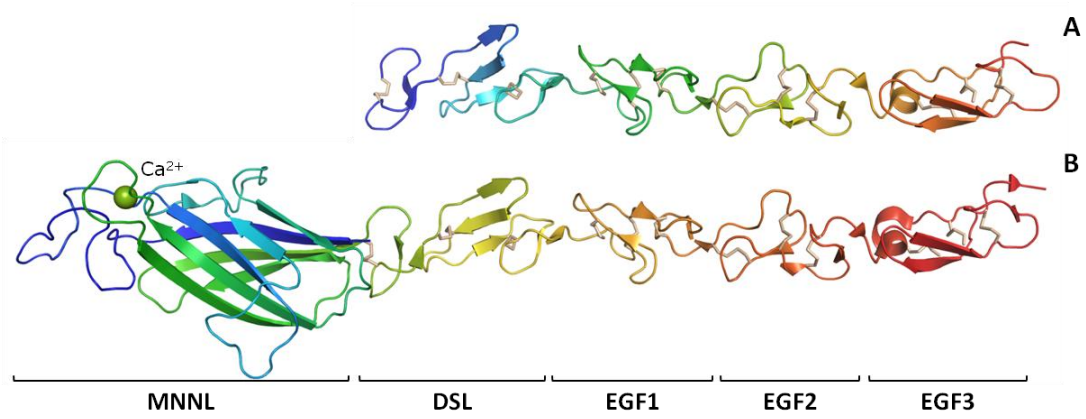


Figure I.3 – Structural models of Jagged 1. A) Jagged1 DSL-EGF3 domains, residues 187 to 335 (top, PDB: 2VJ2; (Cordle, Johnson *et al.* 2008)) and B) MNNL-EGF3 domains, residues 32 to 335 (bottom, PDB: 4CC0; (Chillakuri, Sheppard *et al.* 2013)). Cartoon representations generated using PyMOL (http://www.pymol.org) coloured from blue at the N-Terminus to red in the C-Terminus. Disulfide bonds are shown as light grey sticks.

The DSL domain has an EGF-like fold, albeit its dissulfide pattern (C1-C2, C3-C4, C5-C6) is not conserved. This domain shows a face with highly conserved residues, crucial for Notch binding (Cordle, Johnson *et al.* 2008). EGFs 1 and 2 show a truncated structure compared with other EGFs, between cysteine 3 (C3) and cysteine 4 (C4), while maintaining the EGF characteristic dissulfide pattern of C1-C3, C2-C4, C5-C6. The MNNL domain presents a C2 fold domain with a beta-sandwich of 8 β-strands. Human Jagged1 C2 or MNNL domain has high structural similarity with the Munc13-1 C2b domain with a r.m.s.d. (root-mean-square deviation) of 2.7Å despite only 10% sequence homology. As a C2 domain, MNNL retains the ability to bind phospholipids in a calcium-dependent manner, suggesting a modulation role of this domain in the signalling process (Chillakuri, Sheppard *et al.* 2013).

I.4. Notch in development and disease (cancer)

Notch signalling pathway is involved in several aspects of development, homeostasis and cell fate determination both in invertebrates and vertebrates (Robey 1997, Artavanis-Tsakonas, Rand *et al.* 1999, Wang 2011). Soon after the burst of knowledge on invertebrate Notch signalling, mammalian Notch ligands were identified in rat (Lindsell, Shawber *et al.* 1995) and mouse (Shimizu, Chiba *et al.* 1999). In the same year, Notch signalling involvement in oncogenic processes was the first normal and cell-fate-determining mechanism reported to have such role (Zagouras, Stifani *et al.* 1995), hence providing new research avenues and intricacies to the already highly populated components of this signalling pathway.

Mutations in the Notch receptors that may cause an exacerbated Notch signalling response, similar to the pathway's normal function of the given tissue, have been involved in B cell malignancies and T-cell acute lymphoblastic leukemia (T-ALL). Those mutations lead to an increased NICD half-life, and to an elevated nuclear activity by disrupting either the C-Terminal PEST sequences (responsible for NICD targeting for degradation), or the Negative Regulatory Region (that allow proteolysis events to activate Notch in a ligand-independent manner) (Figure I.2) (South, Cho *et al.* 2012). Sethi, Dai *et*

al. (2011) reported that the Jagged1 ligand, and not other pathway components, contributes to the metastatic ability of breast cancer cells towards bone through TGF-β (a pro-metastatic cytokine). Moreover, the report shows that γ-secretase inhibitors (GSIs) can inhibit breast cancer bone metastasis by disruption of Notch signalling pathway between breast cancer cells and stromal bone cells. GSIs are small molecules that are being studied as candidate therapeutic agents (Rizzo, Osipo *et al.* 2008) but are not specific neither for Notch signalling pathway or cancer treatments (Purow 2012). These small molecules already entered clinical trials but provoke mainly intestinal toxicity deregulating Notch signalling, important for intestinal epithelial cell differentiation, and causing goblet cell metaplasia (van Es, van Gijn *et al.* 2005). As such, tissue specific therapies, rather than systemic inhibition or blockage on Notch receptors, are needed that act in a target dependent manner. The development of therapies with a very specific and restricted expression pattern may result in the generation of good candidates for biological treatments, reducing mechanism based toxicities (Rizzo, Osipo *et al.* 2008).

Notch signalling is also involved in physiological angiogenesis regulation. Sprout tip cells express high levels of Delta-Like Ligand4 that activate Notch1 in adjacent cells in response to Vascular Endothelial Growth Factor (VEGF) (Garcia and Kandel 2012). Disruption of Dll4 signalling enhances angiogenic sprouting in a tumour environment, and is thought that this deregulated sprouting can cause non-productive angiogenesis, leading to tumour regression due to poor perfusion and increases in hypoxia. Ridgway, Zhang *et al.* (2006) reported that a neutralizing Dll4 antibody that blocks Notch signalling, and inhibits tumour growth in a VEGF-dependent manner, has the ability to reduce tumour size in several solid tumour models, without intestinal toxicity, as seen with the use of GSIs. This study demonstrated how tumour regression was accomplished without compromising the expression of Notch target genes, providing the proof of concept that a targeted biological treatment can in fact be more promising. Nevertheless, the same group reported recently that chronic blockade of Notch/Dll4 pathway with the neutralizing Dll4 antibody in adult mice increases the expression of several endothelium-specific genes, having a negative impact in liver homeostasis, and increasing the incidence of ulcerating subcutaneous tumours in male rats (Yan, Callahan *et al.* 2010), precluding the continuous use of this particular antibody in a treatment regimen.

Furthermore, recent studies have implicated aberrant Notch signalling in breast cancers. High level expression of Notch receptors and ligands, and their increased activation in several breast cancers and early precursors, place Notch signalling as a key player in breast cancer pathogenesis (Mittal, Subramanyam *et al.* 2009). Notch signalling genes – Notch1 and Jagged1 – have been shown to be potential prognostic markers for breast cancer (Sethi, Dai *et al.* 2011), and the expression of Notch1 ligand Jagged2 has also been correlated with the overall- and metastasis-free survival of breast cancer patients (Xing, Okuda *et al.* 2011). Mittal, Subramanyam *et al.* (2009) have also shown that Delta-like-1 expression is undetectable in normal breast tissues, but moderate to high expression has been detected in breast cancer. Thus any of the Notch1 ligands can be studied as possible therapeutic targets for breast cancer.

Structural models are very useful to allow understanding many biological pathways. As such, determination of the crystal structures of Notch ligands would allow determining the molecular basis of antagonistic activity in the Notch1 signalling pathway. However, it is first necessary to obtain a pure and homogeneous protein solution prior to crystallization experiments, biophysical characterisation and other *in vitro* research strategies. To further dissect the different roles of each of the Notch ligands domains, there is the need to obtain truncated versions of the target proteins, engineering DNA constructs bearing the desired sequences and heterologously express the respective recombinant proteins.

I.5. Recombinant Protein production strategies

Recombinant protein production requires a host organism that may not always necessarily be the original organism of the protein to be expressed.

I.5.1 Protein expression in *E. coli*

The normal first approach for recombinant protein production relies in the use of bacterial cells, namely *Escherichia coli*. This allows for a fast and cost effective production of large amounts of protein (Jonasson, Liljeqvist *et al.* 2002). The majority of structural data deposited in the Protein Data Bank (PDB) has relied in protein produced in *E. coli*, summing up to 80% in 2003 (Sorensen and Mortensen 2005).

In spite of the advantages of using *E. coli* as a host for recombinant protein production, there are several drawbacks, especially when dealing with eukaryotic (human) proteins as the codon usage between the two organisms is considerably different and rare codons tend to stall translation (Jonasson, Liljeqvist *et al.* 2002, Gustafsson, Govindarajan *et al.* 2004). To overcome this potential problem, gene synthesis of codon optimized genes or the use of host strains co-expressing rare tRNAs are the most standard strategies. Without the usage of these strains the presence of rare codons can decrease target protein quantity and quality, overwhelming of the available rare tRNA polls, and provoking translational errors (Jonasson, Liljeqvist *et al.* 2002, Gustafsson, Govindarajan *et al.* 2004, Jana and Deb 2005). Codon optimization of the desired gene sequence modifies the problematic codons present in the native gene, but that have a low frequency of usage (mostly in the third 'wobble' position), in such a way that the protein sequence of interest is maintained. The use of an optimized gene sequence can increase the recombinant protein expression up to 10-20% of total *E. coli* protein when compared to the native's gene sequence for expression (Gustafsson, Govindarajan *et al.* 2004). Whenever it is not possible, or when it is too expensive to obtain gene optimized sequences, non-optimized genes can be expressed using host strains that co-express rare tRNAs; those are commercially available with different names from a number of suppliers such as Stratagene's BL21 CodonPlus or Rosetta from Novagen, an expression host strain transformed with the pRARE plasmid.

Moreover, eukaryotic proteins may not fold correctly in *E. coli* cytoplasm due to improper redox conditions and be accumulated as inclusion bodies. In order to increase protein solubility there are various approaches to consider, either alone or in combination: 1) Reducing protein synthesis rate by:

using a weak or a moderately strong promoter, by partially inducing a strong promoter, or by decreasing the temperature; 2) Using a highly soluble fusion partner that drives the overall solubility of the fusion protein, such as Maltose-binding Protein (MBP), Small Ubiquitin Modifier (SUMO), Glutathione S-Transferase (GST) and Thioredoxin (Trx); and 3) Co-expressing molecular chaperones that drive partially or unfolded peptides to adopt the native conformation such as DnaK-DnaJ-GrpE, ClpB, GroEL-GroES, and IbpA/B (Jonasson, Liljeqvist *et al.* 2002, de Marco, Vigh *et al.* 2005, Jana and Deb 2005, Bell, Engleka *et al.* 2013). The successful application of one or a combination of the mentioned approaches is determined experimentally and in an empirical basis; in many cases though, upon over-expression, recombinant protein is stored as aggregates in Inclusion Bodies (Shimizu, Chiba *et al.* 1999). Additionally, modified strains that allow for disulfide bond formation in *E. coli* cytoplasm such as Origami and derivatives (Novagen) and SHuffle (Lobstein, Emrich *et al.* 2012) have been developed. Moreover, bacteria do not possess the cellular machineries for post-translational modifications (PTM) (Jana and Deb 2005) such as glycosylation and phosphorylation. This poses the concern when expressing mammalian/human proteins, such as the domains of Notch ligand proteins, as there are predicted sites for PTMs in their sequences (http://www.uniprot.org/; Accession numbers: hJagged1 - P78504, hJagged2 - Q9Y219, hDelta-like Ligand1 - O00548).

I.5.2 Protein expression in Eukaryotic systems

To overcome part of the disadvantages of bacterial expression mentioned above, mammalian expression systems can be utilized for the production of human proteins. Both murine and human cells possess all the required machinery for folding, disulfide bond formation and PTM for eukaryotic protein synthesis, in dedicated cellular compartments that provide the necessary and different environments, and the machinery for each stage of protein production. The most used cell lines for protein production cited in the literature are CHO cells (Chinese Hamster Ovarian cells, a murine cell line), and HEK cells (Human Embryonic Kidney, a human cell line). The choice of cell line lies in specific needs either for productivity or fidelity in terms of molecular structure and biochemical proprieties, lying the choice between CHO or HEK cell lines, respectively (Swiech, Picanco-Castro *et al.* 2012).

Unfortunately, the use of mammalian cells as expression host poses different challenges in both technical and financial points of view when compared to bacterial expression systems. There are two strategies for protein expression using mammalian cells, Transient and Stable Gene Expression (TGE and SGE, respectively). The generation of stable expression clones for SGE requires the integration of the gene of interest, transfected as for TGE, in the cells' genome, followed by a time-consuming selection and characterization process of the generated clonal cells. For that reason, when multiple targets are to be expressed, a TGE approach is preferred (Swiech, Picanco-Castro *et al.* 2012, Diepenbruck, Klinger *et al.* 2013). Being technically less difficult than SGE, cell culture maintenance for TGE requires specialized training and a great amount of operator hours. Generally, the plasmids used for TGE have the gene of interest controlled by a constitutive promotor and for that reason, cultures must be expanded prior plasmid DNA transfection in a large scale, increasing the total amount of DNA and transfection reagents required. These are of course scaled accordingly to the final

culture volume required for each experiment. Moreover, it is imperative to use high quality DNA and transfection reagents (Zhao, Bishop *et al.* 2011), increasing the total cost per protein production.

Regardless of the technical difficulties encountered when dealing with mammalian cells as hosts for protein expression, this may me a safe option to generate human proteins amenable for purification and subsequent studies, in case bacterial expression fails.

I.6. Aim of this thesis

The present study aims to unveil the best expression and purification conditions for truncated domains of the extracellular domains of canonical Notch ligands. The major goal is to express and purify the Notch1 ligand proteins and use them as antigens to generate specific antibodies employing Phage Display technology, for target specific therapeutics. Also, and since to date there are no complete structural models of any of the Notch1 ligands, the objective is to determine the crystal structures of full-length or partial Jagged1 and 2 and Delta-Like Ligand1 proteins, alone and in complex with the generated antibodies, by X-Ray crystallography.

Chapter II Methods[1]

II.1. Molecular Biology

All primer oligonucleotides synthesis and DNA sequencing services were purchased from STABVida, Portugal. Human Jagged1 (Gene Bank Ref.IDnm_000214.2) and human Delta-like ligand1 (Gene Bank Ref.IDnm_005618.3) cDNAs were purchased from SinoBiological (Cat# HG11648-M and HG11635-M, respectively). Human Jagged2 (Gene Bank Ref.IDnm_002226.3) cDNA from nucleotide 306 to 2171 was synthesised by GeneScript.

The *E. coli* strain used to routinely transform plasmid DNAs during cloning procedures was DH5α (genotype: *fhuA2 lac(del)U169 phoA glnV44 Φ80' lacZ(del)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*) and the *E. coli* strains used for protein expression were BL21(DE3) (genotype: *E. coli B* F^- ompT gal dcm $hsdS_B(r_B^- m_B^-)$ λ(DE3)), BL21 Star (DE3) (genotype: *E. coli B* F^-*ompT* *hsd*S_B (r_B^-, m_B^-) *gal dcm rne*131 λ(DE3)) and SHuffle T7 Express (Lobstein, Emrich *et al.* 2012) (genotype: *E. coli B fhuA2 lacZ::T7 gene1* [lon] *ompT ahpC gal λatt*::pNEB3-r1-*cDsbC* (Spec^R, *lacI*^{q}) Δ*trxB sulA11 R(mcr-73::miniTn10*--Tet^S)2 [dcm] *R(zgb-210::Tn10* --Tet^S) *endA1 Δgor* Δ*(mcrC-mrr)114::IS10).*

E. coli BL21 (DE3) and BL21 (DE3) Star competent cells were generated in-house from already existing bacterial strain stocks.

II.1.1 Cloning in pET-47b(+) vector

pET-47b(+) vector (Kan^R), for expression in *E. coli*, was a kind gift from Colin E. McVey, PhD. Primers including a restriction enzyme recognition sequence were designed for each construct, as shown in Table II.1. For each construct, PCR amplification was carried out using 1.25 units of *Pfu* DNA polymerase, 0.2mM dNTPs and 0.2μM of each insert specific forward and reverse primers, 2-10ng of template DNA and 1x polymerase reaction buffer. Amplification conditions were: 1 initial denaturing cycle at 95°C for 5 minutes, 32 cycles at 95°C for 30 seconds (denaturing step), 58°C for 30 seconds (annealing step) and 72°C for 2 minutes (extension step)/ kb of amplicon, 1 final extension cycle at 72°C for 10 minutes. After amplification, insert size was confirmed by agarose gel electrophoresis, digested with 10 units of *Eco*RI and 10 units of *Hin*dIII for 1 hour at 37°C, and ligated into the destination vector, previously digested with the same restriction enzymes, under the same conditions, using 2.5 units of T4 DNA ligase and incubating overnight at 16°C. Ligation reactions were transformed into *E. coli* DH5α competent cells and plated in 2xYT agar plates supplemented with 50μg/mL Kan.

Colony PCR screening with *Taq* DNA polymerase with insert-specific primers was carried out with the same PCR conditions as for the initial amplification, to evaluate the presence of positive colonies. The colony PCR result was assessed by agarose gel electrophoresis of each PCR reaction. One positive colony for each clone was selected and incubated overnight at 37°C to saturation in liquid media (2xYT or PB) with 50μg/mL Kan. Plasmid DNA was isolated using GeneJET Plasmid Miniprep Kit according to the manufacturers' instructions. Plasmid DNA was further subjected to restriction

[1] All reagents' catalogue numbers, suppliers, and buffer recipes are listed in Chapter VI-Appendix.

enzyme digestion to verify the presence of the insert and confirm the expected molecular sizes of both the insert and the vector. DNA fragments were analysed by agarose gel electrophoresis. Sequence integrity of each clone was confirmed by sequencing at STABvida, Portugal.

Table II.1 - Primers used for Jag1, Jag2 and Dll1 cloning into pET-47b(+) bacterial expression vector. The nucleotide sequence for each primer is shown, highlighting the restriction enzyme sites in bold, the inserted stop codon in italic and the complementarity with the coding sequence in capital letters.

Protein	Target domain	Restriction enzyme	Primer Sequence
hJAG1	MNNL	EcoRI	5'- gtgtgg**gaattc**gggtCAGTTCGAGTTGG
	DSL	EcoRI	5'- gagtatca**gaattc**cGTGACCTGTGATGAC
	EGF3	HindIII	5'- gcaggc**aagctt**agc*cta*TTCACAGTTGG
	EGF9	HindIII	5'- gcagt**aagctt**tcag*tta*TGAGCAGTTCTTG
hJAG2	MNNL	EcoRI	5'- gtgca**gaattc**acgtCCGATGGGTTATTTTG
	DSL	EcoRI	5'- ctgca**gaattc**tGTTCGCTGCGATGAAAAC
	EGF3	HindIII	5'- cgtgcaggc**aagctt**cgc*tta*TTCACAGTTGC
	EGF9	HindIII	5'- cacggt**aagctt**gg*cta*GCTACAGTTTTTAC
hDLL1	MNNL	EcoRI	5'- cctt**gaattc**tCAGGTCTGGAGCTC
	DSL	EcoRI	5'- ctcaa**gaattc**ctaccgcTTCGTGTGTG
	EGF3	HindIII	5'- ctcgtc**aagctt**cag*cta*GCAGGTGGCACC
	EGF6	HindIII	5'- cgtc**aagctt**gtc*cta*ACAGTGCCTCCC

II.1.2 Cloning in pHL-sec vector

pHL-sec vector (Amp^R) (Aricescu, Lu *et al*. 2006), for mammalian expression, was a kind gift from Dr. Yuguang Zhao, through the application submitted to P-CUBE platform (Infrastructure for Protein Production Platforms, UE FP7). Primers including a restriction enzyme recognition sequence were designed, the forward with *Age*I and the reverse with *Kpn*I or *Eco*RV recognition sequences as shown in Table II.2. For each construct, PCR amplification was carried out using *Pfu* DNA polymerase as described above. Whenever necessary, primer annealing temperatures were optimized. After amplification, insert size was confirmed by agarose gel electrophoresis, digested with 10 units of *Age*I and 10 units of *Kpn*I or *Eco*RV. The destination vector was digested with *Age*I and *Kpn*I or with *Kpn*I followed by Klenow fragment and then *Age*I.

Ligations were carried out using 2.5 units of T4 DNA ligase and incubated either one hour at RT or overnight at 16°C. Ligation reactions were transformed into *E. coli* DH5α competent cells and plated in 2xYT agar plates supplemented with 100µg/mL Amp. Colony PCR screening with Taq DNA polymerase with insert-specific primers was carried out to evaluate the presence of positive colonies. One positive colony for each clone was incubated overnight in liquid media with 100µg/mL Amp and plasmid DNA was isolated using GeneJET Plasmid Miniprep Kit. Plasmid DNA was further subjected to restriction enzyme digestion to verify the presence of the insert, and confirm the expected molecular sizes of both the insert and the vector; fragments were verified by agarose gel electrophoresis. DNA sequence integrity was confirmed by sequencing.

Table II.2 - Primers used for Jag1, Jag2 and Dll1 cloning into pHL-sec mammalian expression vector. The nucleotide sequence for each primer is shown, highlighting the restriction enzyme sites in bold and the complementarity with the coding sequence in capital letters.

Protein	Target domain	Restriction enzyme	Primer Sequence
hJAG1	MNNL	AgeI	5'- ccaaggtg**accggt**GCCTCGGGTCAGTTCG
	DSL	AgeI	5'- gcatgatcaaccc**accggt**CAGTGGC
	EGF3	KpnI	5'- gcaggc**ggtacc**AGCAATTTCACAGTTGG
	EGF9	KpnI	5'- gctgcg**ggtacc**GTCTTTCAGGTGTGAGC
hJAG2	MNNL	AgeI	5'- gcactg**accggt**CAGGCACGACGTCCG
	DSL	AgeI	5'- atgatc**accggt**GAAGACCGCTGGAAATCACTGC
	EGF3	KpnI	5'- cgggttaga**ggtacc**GGCATGTTCCGC
	EGF9	EcoRV	5'- cgacaggcaccaccc**gatatC**GGTTCACG
hDLL1	MNNL	AgeI	5'- gtgtcag**accggt**AGCTCTGGGGTGTTCG
	DSL	AgeI	5'- cggtcggc**accggt**TGGTCCCAGGACC
	EGF3	KpnI	5'- ggtgacactc**ggtacc**CCCCAGCTCGC
	EGF6	KpnI	5'- gcgca**ggtacc**CTCGTTGTCGTCACAGTGC

II.1.3 Cloning in pTT22SSP4 vector

Cloning of hJagged1 domains MNNL-EGF9, human Jagged2 domains MNNL-EGF9 and human Dll1 domains MNNL-EGF6 and DSL-EGF6 was performed (as described in the previous section) in mammalian HEK expression vector pTT22SSP4 (Durocher, Perret *et al.* 2002), using *Nhe*I/*Bam*HI sites. The primers designed to amplify the cDNA containing the restriction enzymes' recognition sites are listed in Table II.3. The resulting constructs were confirmed by sequencing (STABvida).

Table II.3 - Primers used for Jag1, Jag2 and Dll1 cloning into pTT22SSP4 mammalian expression vector. The nucleotide sequence for each primer is shown, highlighting the restriction enzyme sites in bold, the inserted stop codon in italic and the complementarity with the coding sequence in capital letters.

Protein	Target Domain	Restriction Enzyme	Primer Sequence
hJAG1	MNNL	NheI	5'- ccaaggtg**gctagc**GCCTCGGGTCAGTTCG
	EGF9	BamHI	5'- gctgcg**ggatcc***tta*TTTCAGGTGTGAGC
hJAG2	MNNL	NheI	5'- gcactg**gctagc**CAGGCAGCACGTCCG
	EGF9	BamHI	5'- cgacaggcaccaccc**ggatcc***tta*TCACG
hDLL1	MNNL	NheI	5'- gtgtcag**gctagc**AGCTCTGGGGTGTTCG
	DSL	NheI	5'- cggtcggc**gctagc**TGGTCCCAGGACC
	EGF6	BamHI	5'- gcgca**ggatcc***tta*GTTGTCGTCACAGTGC

II.2. SDS-PAGE and Western Blot

Samples were adjusted with SDS-PAGE Loading Buffer (1x final concentration) and boiled for 10 minutes at 95°C; when needed, samples to be analysed in non-reducing conditions were adjusted with NuPAGE LDS sample buffer before SDS-PAGE analysis.

SDS-PAGE was carried out using Criterion XT Precast Gels or Mini-PROTEAN TGX precast Gels using MES or MOPS as a Buffer system depending on number of samples to be analysed, and on the protein size. Gels ran at 200V for 35 or 45 minutes depending on MES or MOPS buffer system,

respectively, and were stained with SimplyBlue SafeStain according to manufacturers' instructions, or transferred to nitrocellulose membranes for further analysis.

Polyacrylamide gels were transferred to nitrocellulose membranes using a Trans-Blot SD Semi-Dry Transfer Cell apparatus (BioRad) according to manufacturer's instructions: 25V for 45 minutes for large gels, and at 15V for 30 minutes for small gels, using NuPAGE Transfer Buffer.

Membranes were blocked in 3% milk powder in PBS-0.05%T for 30 minutes, washed 3 times for 5-10 minutes each in PBS-T, incubated with mouse α–His at 1µg/mL in PBS for 1 hour, washed 3 times for 5-10 minutes each in PBS-T, incubated with α-mouse IgG-HRP at 2.5µg/mL in PBS for 30 minutes and washed 5 times for 5-10 minutes each in PBS-T. Signal was developed with Western Lightning Plus-ECL Enhanced Chemiluminescence Substrate and images acquired in ChemiDoc (BioRad).

II.3. Protein quantification

Protein was quantified by Bradford assay with Bio-Rad Protein Assay Dye Reagent Concentrate. Samples were assayed in triplicate and the corresponding buffer was used for blank measurements. In a final volume of 1mL, 4µL of sample and 200µL Reagent were used. Blank-subtracted absorbance readings at 595nm were averaged and compared to a calibration curve prepared with BSA as a standard, with known concentrations.

Alternatively, protein concentration was determined by spectrophotometic methods with NanoDrop taking in account the theoretical extinction coefficient and molecular mass calculated by ProtParam tool provided by ExPASy (available online at web.expasy.org/protparam/) with the input sequence for each protein and shown in Table III.1 and III.2

II.4. Protein expression and purification using *E. coli* as host

II.4.1 Small scale expression screenings

Plasmids were transformed into competent cells and plated in 2xYT agar supplemented with the appropriate antibiotic(s). BL21 (DE3) and BL21 (DE3) star competent cells were transformed with the plasmid(s) of interest by electroporation, and transformation reactions were allowed to recover 1 hour at 37°C with 250rpm agitation; SHuffle® chemically competent cells where transformed with the plasmid of interest by heat shock, and transformation reactions were allowed to recover 1 hour at 30°C with 250rpm agitation. Of the transformation plates ten to fifteen colonies were picked and inoculated in 10mL expression media supplemented with the appropriate antibiotics, and cultures were grown to saturation overnight at 37°C and 250rpm. Cultures were diluted in the appropriate expression media (25mL) to an initial OD_{600} of 0.10-0.15 and allowed to grow at 37°C, or 30°C when using SHuffle® strain, until the chosen OD_{600} for induction; cultures were induced with IPTG and induction proceeded at the chosen temperature and duration. For BL21 strains, the OD_{600} of induction was between 1.5 and 2.0, except when co-expression of the chaperone plasmid was desired. In this case, OD_{600} of induction was between 0.5 and 0.7, and co-expression was performed in the presence of

34µg/mL Cm, and chaperone plasmid was induced with 4mg/mL L-arabinose. A summary of the tested conditions for each protein construct are depicted in Table III.3.

II.4.1.i Sample preparation for Induction and Solubility tests analysis

Samples for each test condition of small-scale screenings were collected before and after induction. Samples were standardised by harvesting the culture volume corresponding to 2 OD_{600} units. Cultures were centrifuged for 5 minutes at 13200*xg* at RT. Cell pellets corresponding to total protein before induction (NI) and total protein after induction (I) were resuspended in 1x SDS-PAGE Loading Buffer (Sambrook 2001).

To evaluate the solubility of each condition, samples were lysed either by using BugBuster Protein Extraction Reagent (Novagen) following manufacturer's instructions, home-made Lysis buffer with lysozyme and DNase for 15-30 minutes on ice, or Lysis buffer and three sonication pulses of 10 seconds each. Soluble (Sol) and Insoluble (Insol) fractions were obtained by centrifugation at 16600*xg* for 30 minutes at 4°C. SDS-PAGE Loading Buffer (4x) was added to a final concentration of 1x.

II.4.2 Protein expression and purification

Large-scale protein expression was carried out in 2.5L TunAir flasks in the best conditions determined during the small-scale expression trials. Cells were harvested by centrifugation for 15 minutes at 7500*xg* and 4°C, and cell pellet was stored at -80°C until processing.

All chromatographic steps were performed using ÄKTApurifier or ÄKTAFPLC systems unless stated otherwise. All buffers were filtered through 0.22µm membranes and degassed. Reducing agents were added freshly before use and the pH was corrected to the final desired value.

II.4.2.i Soluble fraction

Cell pellets were thawed in Lysis Buffer (Sections VI.2.i, VI.2.ii) and the lysate was disrupted three times in a French Press at 10000 psi. The lysate was clarified by centrifugation for 35 minutes at 48400*xg*, 4°C, and filtration through a 0.22µm filter. One 5-mL HisTrap column was connected to an ÄKTApurifier and equilibrated with 5 CV of MilliQ water and 5 CV of Buffer A. The sample was injected at 1mL/min and the column washed with 15 CV Buffer A, until baseline absorbance was below 100mAU. Elution was carried out with a gradient 0-0.5M Imidazole in 10 CV, followed by 0.5-1M Imidazole in 4 CV. Eluted protein was collected in 2mL fractions. Fractions were analysed by SDS-PAGE and WB.

Fractions containing the protein of interest were buffer exchanged to HRV-3C reaction buffer using a HiPrep 26/10 desalting column and cleaved overnight with a GST tagged HRV-3C protease (produced and purified in-house) at 1:100 (w/w) ratio. Protein mixture was applied to a GSTrap column connected to a HisTrap column at 1ml/min and the flow-through was collected. Protein was concentrated and injected at 1ml/min to a HiLoad 16/600 Superdex 75pg (S75) column. Protein was

fractionated in 2mL fractions; fractions were loaded in a SDS-PAGE gel to confirm the presence of the protein of interest.

II.4.2.ii Inclusion Bodies

Test conditions for refolding were prepared based on extensive refolding studies (Rudolph and Lilie 1996). Inclusion bodies pellets were resuspended on ice in ressuspension / wash buffer, clarified by centrifugation for 10 min at 30000*xg*, 8°C, and the supernatant was discarded. This washing step was performed five times. A final wash was performed where Triton X-100 was omitted from the resuspension solution. The inclusion bodies were unfolded by dissolving the washed inclusion body pellets overnight at 4°C in solubilisation buffers containing (I) 6M Guanidine Hydrochloride, or (II) 8M Urea as protein denaturants, and clarified by centrifugation for 30 min at 30000*xg*, 8°C. Supernatant corresponding to the unfolded protein was analysed by SDS-PAGE and stored at -20°C until further use.

Four Refolding Buffers were tested (Section VI.2.iii):

A - 200mM Tris, 3.7mM Cystamine, 6.5mM Cysteamine, 1M L-Arginine, pH 8.0;

B - 20mM Tris, 1mM EDTA, 1mM GSH, 0.1mM GSSG, pH 8.0;

C – Buffer A + 1M TMAO;

D – Buffer B + 1M L-Arginine and 1M TMAO.

For protein refolding by rapid dilution, unfolded protein was added dropwise at 4°C to rapidly stirring Refolding Buffer at a final concentration of 0.1mg protein per millilitre of buffer, and incubated at 4°C for ~40h. The refolded mixture was then clarified by filtration through a 0.22µm filter to remove particulate matter, concentrated in Amicon device, and desalted in batch mode using PD-10 columns according to manufacturer's instructions. Resulting samples were quantified by spectrophotometry and protein containing fractions were concentrated to ~1mg/mL in Amicon 10kDa MWCO.

For large scale purification, a 2L culture was grown using the best condition determined before, centrifuged and lysed. Inclusion bodies were isolated and protein refolding was archived using the optimal conditions determined previously from the small scale purification results scaled accordingly. Refolded protein solution was clarified by centrifugation for 25 minutes at 30100*xg* at 4°C, followed by filtration through a 0.22µm pore membrane. Protein was concentrated using Amicon Stirred Cell unit and Amicon device.

Size exclusion chromatography was performed in a S75 column connected to an ÄKTApurifier. Column was equilibrated with 1.5 CV of buffer (Section VI.2.iv) at 1mL/min, protein was injected at 0.5mL/min and fractionated. Selected fractions were analysed by SDS-PAGE.

II.5. Protein expression and purification using HEK293 cells as host

II.5.1 P-CUBE at UOXF, Oxford, United Kingdom

II.5.1.i Small-scale expression screenings

Adherent HEK293T cells were routinely cultured in DMEM supplemented with 10% FCS in a 95% air / 5% CO_2 atmosphere at 37°C. Small scale screening experiments were performed by transient transfection of pHL-sec constructs in adherent HEK293T cells, plated in 6-well plates. For each transfection experiment, a cocktail consisting of plasmid DNA:transfection reagent complex was prepared by adding 8µL Lipofectamine2000 (1mg/mL) to 250µL serum-free DMEM, and 2 to 4µg plasmid DNA to 250µL serum-free DMEM. Both mixtures were briefly vortexed and combined in a single tube followed by incubation for 20 minutes at RT. The transfection cocktail was added to cells in DMEM + 2% FCS. Cell media was collected 48 hours post-transfection for Western Blot analysis of expression.

II.5.1.ii Large-scale protein production

Large scale transfection was performed in cells grown in expanded surface roller bottles under the same conditions as described above. For each roller bottle, a transfection cocktail containing 0.5mg plasmid DNA and PEI to a DNA:PEI ratio of 1:2 was used. This was prepared by adding the pHL-sec DNA constructs to 50mL serum-free DMEM, mixing and adding 1mL PEI (1mg/mL). Transfection cocktail was incubated for 10 minutes at RT and, prior to adding to the cells, 0.375 mL Kifunensine (1mg/ml) was added to the mixture. Transfection cocktail with Kifunensine was added to cells in DMEM + 2% FCS, and incubated for 4 days. Media was harvested and clarified by centrifugation for 15 minutes at 4000rpm at 10°C, and filtration through a 0.22µm SteriTop-GP Filter Unit. Clarified media was concentrated 10 fold and partially buffer-exchanged to PBS using a QuixStand system by cross-flow filtration with hollow fiber cartridge of 10kDa MWCO, prior to purification.

II.5.1.iii Protein purification

Protein sample was purified in batch using Cobalt beads. Cobalt beads were extensively washed in distilled water followed by washing with PBS. Sample was mixed with the cobalt beads and incubated between 1 and 1 and a half hours at 16°C with 110rpm agitation. The bead:sample mixture was poured into a reusable glassware column and flow-through was collected. After all sample passed through, cobalt beads were washed with PBS, followed by a washing step of about 5 CV with wash buffer 1 (Section VI.2.v), and two washing steps of wash buffer 2 corresponding to about 3 and 2 CV, respectively. Protein was eluted with two steps of about 3 and 2 CV each of elution buffer. Eluted protein was concentrated in an Amicon 10kDa MWCO to a volume of 6mL.

A HiLoad 16/600 Superdex 200 pg column connected to an ÄKTAexplorer was equilibrated with SEC buffer, and two 3mL injections were performed by direct load method in the sample loop. Eluted

protein was fractionated in 2mL fractions and analysed by SDS-PAGE with BenchMark™ Protein Ladder protein marker.

II.5.2 Protein Expression performed at iBET/ITQB

II.5.2.i Suspension cells

Suspension culture adapted HEK293-EBNA cells were routinely maintained in FreeStyle F17 media supplemented with 20mL/L Glutamax, 10mL/L Pluronic F-68 and 0.5mL/L Gentamycin in a 93% air / 7% CO_2 atmosphere at 37°C with 110rpm agitation. Cells were sub-cultured twice a week with a starting inoculum of 0.3 to $0.4x10^6$ cells/mL. Suspension culture adapted HEK293T cells were routinely maintained in FreeStyle 293 media in a 95% air / 5% CO_2 atmosphere at 37°C and 110rpm agitation. Cells were subcultivated and transfection procedures were performed as described above for HEK293-EBNA cells.

II.5.2.ii Adherent cells

Adherent HEK293T cells were routinely maintained in DMEM supplemented with 10% FBS in a 95% air / 5% CO_2 atmosphere at 37°C, and sub-cultured when reaching 85-95% confluency. For sub-culturing, cells were rinsed once with PBS and dissociated with 0.05% Trypsin-EDTA for 5 minutes at 37°C. Trypsin was inactivated by adding serum supplemented growth media. Optimal dilution for sub-cultivation was determined by plating different dilutions of the obtained cell suspension, or different cell numbers so that sub-culturing was performed twice a week as previously described.

II.5.2.iii Small-scale expression screenings

Small-scale expression screenings were performed by transiently transfecting 20mL suspension adapted HEK cells at $1.5\text{-}1.8x10^6$ cells/mL, grown for one passage in FreeStyle F17 with Glutamax media, with 0.5 or 1mg pHL-sec or pTT22SSP4 DNA constructs per Litre of culture, using PEI transfection reagent to a DNA:PEI ratio of 1:3. Transfection cocktails were prepared by adding the plasmid DNA to FreeStyle F17 media corresponding to 1/10th of the culture volume to be transfected and mixed. PEI reagent (1mg/mL) was added to the transfection cocktail and incubated for 10 minutes at RT prior to adding to the cells. Samples were collected 4 and 5 days post-transfection, and supernatant recovered for protein expression analysis by SDS-PAGE or WB as applicable.

For adherent cells grown in 6-well plates, small-scale expression screenings were performed by transfecting 2 to 4µg pHL-sec plasmid DNA, using PEI as transfection reagent in a ratio DNA:PEI 1:3 by adding the plasmid DNA to serum-free media, mixing, adding the PEI reagent and incubating for 10 minutes prior to adding to the cells. Growth media samples were harvested 48 and/or 72 hours post transfection and analysed by SDS-PAGE or WB.

II.5.2.iv Protein production

A medium scale production was performed in the best condition obtained in the small-scale experiments, scaled up accordingly in a 3L Fernbach design erlenmeyer. A 1.5L culture of suspension adapted HEK293-Ebna cells was transfected with 1mg/L pTT22SSP4 DNA construct, PEI:DNA ratio of 3:1 and harvested at 120hpt.

II.5.2.v Protein purification

Harvested culture was clarified by centrifugation for 10 minutes at 9900*xg*, 4°C and the supernatant was recovered. Harvested media was further clarified by filtration through a 0.22µm VacuCap® filter prior to injection in one 5mL-HisTrap column o/n with a peristaltic pump set to x10, 2. Injection was finished upon connection of the column to ÄKTApurifier. HisTrap column was washed with 5 CV of buffer A (Section VI.2.vi) and 5 CV of buffer B. Protein was eluted from the column in steps of 5, 10, 15, 20, 50 and 100% of buffer C. Selected fraction were analysed by SDS-PAGE.

The fractions containing the protein of interest were pooled and concentrated using a stirred tank concentration unit with a 10kDa membrane. Protein mixture was injected in a S75 column, previously equilibrated with 1.5 CV of SEC buffer. Selected fractions were analysed by SDS-PAGE, pooled and concentrated.

II.6. Differential Scanning Fluorimetry (DSF)

DSF analysis was employed as a quality control of refolded protein, and used as a tool to evaluate the protein behaviour in different buffer compositions in several steps of protein purification.

In a typical assay for quality control, a total volume of 20µL and a dye concentration of 5 fold were used to guarantee the best signal to noise ratio. The assays were prepared by adding 2.5, 5, 10 or 20µg of protein, as applicable, to a final volume of 20µL of dye buffer solution, all prepared in the protein purification buffer. Fluorescence intensities versus temperature are used to calculate the protein melting temperature (Tm) by determining the first derivative (d(Rfu)/dT) to extract the exact transition inflection point.

For the Buffer pH / salt screenings, a protein dye mixture was prepared in 25mM Hepes pH 7.5 typically with 2µg protein, and dye 50x concentrated to a final volume of 2µl per well. This reaction mix was distributed in each one of the 96 wells of a low profile 96-well plate (Bio-Rad). Each buffer to be screened was added to a final assay volume of 20µl, thereby diluting the dye to 5x (assay concentration).

The protein melting temperature (Tm) determination was performed by monitoring protein unfolding with the fluoroprobe SYPRO Orange dye, which although completely quenched in aqueous environment, emits fluorescence upon binding to protein hydrophobic patches. Such increase in fluorescence can be measured as a function of temperature. The thermal shift assay is performed on an iCycle iQ5 Real Time PCR Detection System (Bio-Rad), equipped with a charge-coupled device (CCD) camera, and a Cy3 filter with excitation and emission wavelengths of 490 and 575nm, respectively. This equipment can simultaneously detect the fluorescence changes in 96-well low

profile plates, and thus can be used for parallel thermal stability assays. The 96-well plates are sealed with optical quality sealing tape (Bio-Rad) and centrifuged at 2500*xg* for 2 minutes immediately before the assay to remove possible air bubbles. The plates are subsequently heated from 20 to 90ºC with stepwise increments of 1ºC with 10 seconds equilibration time, followed by the fluorescence read out.

The results were analysed with iQ5 Optical System Software, version 2.1 and an Excel spreadsheet with automatic macros available in the group. This spreadsheet calculates automatically the Tm and generates normalised graphs for selected wells with the input of the raw values from the iQ5 program.

II.7. Analytical Size Exclusion Chromatography

A Superose 6 PC 3.2/30 column (GE) with a column volume of 2.4mL was equilibrated in 2 CV of milliQ water at 0.075mL/min, and 2 CV of protein buffer at 0.05mL/min. Samples were centrifuged for 15 min at 16600*xg*, 4°C prior to direct loop injection at 0.05mL/min.

II.8. Circular Dichroism (CD) spectroscopy

CD spectra in the far-UV region (far-UV CD) were recorded from 205-260nm in a JASCO J-815 spectropolarimeter. Acquisition parameters are as follows - accumulations: 10; data pitch: 0.2nm; scanning rate: 100nm/min; bandwidth: 2nm; sensitivity: standard (100mdeg); response time: 4 seconds; temperature: 20°C. Protein buffer was used as blank for baseline corrections. Samples were freshly diluted in protein buffer to 0.2mg/mL and transferred to a 1cm path length quartz cuvette for spectroscopic analysis.

Thermal ramp CD was performed with the following acquisition parameters - data pitch: 0.5°C; temperature slope: 2°C/min; bandwidth: 2nm; sensitivity: standard (100mdeg); response time: 2 seconds; temperature range: 20-70°C, monitored wavelength: 222nm; protein sample concentration: 0.2mg/mL.

II.9. Protein Crystallization

Protein crystallization experiments were performed in CrystalQuick™ 96 Well, Greiner (Hampton Research) using a Nano-Robot Cartesian Mini-Bee (Genomic Solutions, UK) for automated drop setup. Scale-up drops were performed in 24-well Cryschem Plate (Hampton Research) using sitting-drop vapour diffusion method.

For human Jagged2 DSL-EGF3 purified from inclusion bodies, besides the commercial screens, scale-up drops were performed with both protein pools based on the crystallisation condition published for human Jagged1 DSL-EGF3 (Cordle, Johnson *et al.* 2008): vapour diffusion method with reservoir solution composed by 8.5% PEG4000, 0.1M Imidazole Malate pH 7.0. A grid screen was also performed with combinations of 7.5%, 8.5%, 10% and 11% PEG4000 and 0.1 and 0.2M Imidazole Malate pH 7.0, with a 0.5:0.5 protein:crystallization solution ratio at 20°C.

Chapter III Results and Discussion

III.1. Molecular Biology

For each Notch Ligand protein in this study, human Jagged1, Jagged2 and Delta-like Ligand1, four constructs were designed spanning different domains, while maintaining the minimal binding region to the Notch receptor (DSL domain and EGF-like repeats 1 to 3) as previously defined (Shimizu, Chiba *et al.* 1999). The construct names are abbreviated using the following nomenclature: starting in MNNL domain - M; in DSL domain - D; ending in EGF3 - E3; in EGF6 - E6, in EGF9 - E9, constructions depicted in Figure III.1.

Figure III.1 - Construct domain design. Each module represents one domain. EGF as blue and cbEGF in green.

The constructs were designed and cloned successfully in the bacterial expression vector pET-47b(+) that presents an inducible T7 promoter, bearing a N-terminal 6x Histidine tag and a HRV-3C protease cleavage recognition sequence, in frame with the 5' end of the gene of interest. Construct boundaries and theoretical parameters for each construct are shown in Table III.1.

Table III.1 - Theoretical Parameters calculated by ProtParam tool (ExPASy) for proteins expressed in pET-47b(+) vector.

Protein	Domains	Construct Boundaries (a.a.)	Molecular mass (Da)	Extinction coefficient (M^{-1} . cm^{-1})
Jagged1	MNNL-EGF3	33 - 334	36378.5	61235
	MNNL-EGF9	33 - 561	61076.9	91905
	DSL-EGF3	185 - 334	19580.7	38400
	DSL-EGF9	185 - 561	44279.2	69070
Jagged2	MNNL-EGF3	26 - 345	38158.5	68225
	MNNL-EGF9	26 - 572	62154.2	95915
	DSL-EGF3	196 - 345	19661.8	38400
	DSL-EGF9	196 - 572	43675.5	66090
Delta-Like Ligand1	MNNL-EGF3	18 - 325	36889.2	59745
	MNNL-EGF6	18 - 440	48908.5	69810
	DSL-EGF3	176 - 325	19476.6	36910
	DSL-EGF6	176 - 440	31486.0	46975

The constructs were also cloned successfully in the mammalian expression vector pHL-sec (Aricescu, Lu *et al.* 2006). This vector has a CMV enhancer and a constitutive chicken beta-actin promoter, followed by a secretion leader peptide in frame with the 5' end of the gene of interest. At the C-terminal there is a 6x Histidine tag sequence in frame with the 3' end of the gene of interest. The constructs emcompassing the largest fragments for each protein, and the hDLL construct that showed the best expression profile at OPPF (Section III.3, page 52), were also cloned in the mammalian expression vector pTT22SSP4 (Durocher, Perret *et al.* 2002). pTT22SSP4 plasmid has a human CMV promoter followed by an intronic sequence. ORF starts with a signal peptide in frame with a 8x Histidine Tag and a TEV protease site, after which the gene of interest can be cloned using *Nhe*I

restriction enzyme. Construct boundaries and theoretical parameters for each construct are shown in Table III.2.

Table III.2 - Theoretical Parameters calculated by ProtParam tool (ExPASy) for proteins expressed in mammalian expression vectors.

Protein	Domains	Construct Boundaries (a.a.)	Molecular mass (Da)	Extinction coefficient (M^{-1} . cm^{-1})
Jagged1	MNNL-EGF3	31 - 336	34986.9	59745
	MNNL-EGF9	31 - 565	59994.7	90415
	DSL-EGF3	166 - 336	20303.6	43900
	DSL-EGF9	166 - 565	45311.4	74570
Jagged2	MNNL-EGF3	22 - 350	37330.6	66735
	MNNL-EGF9	22 - 577	61368.4	94425
	DSL-EGF3	175 - 350	20948.3	42410
	DSL-EGF9	175 - 577	44986.1	70100
Delta-Like Ligand1	MNNL-EGF3	21 - 328	34994.1	52755
	MNNL-EGF6	21 - 444	47147.5	62820
	DSL-EGF3	159 - 328	20166.3	43900
	DSL-EGF6	159 - 444	32319.8	53965

III.2. Bacterial Expression

III.2.1 Small-scale expression screenings with BL21 (DE3) strains

The aim of the small scale expression screenings was to find the best conditions to express soluble Notch ligand proteins constructs using *E. coli* as host. For that purpose, different conditions known to affect soluble protein expression were tested such as: different host strains, growth media, induction OD_{600}, time and temperature, IPTG concentration and the co-expression with molecular chaperones groEL and groES (encoded by the pGro7 plasmid, TaKaRa). Details of each particular expression screening are detailed in Section III.2, Table III.3.

Since the only structural models available to date regarding Notch ligands are a short construct of human Jagged1 comprising domains DSL to EGF3 (Cordle, Johnson *et al.* 2008), and a smaller construct comprising exon 6 (Pintar, Guarnaccia *et al.* 2009), the first small-scale expression screenings performed were those using a similar construct from human Jagged2, as follows.

III.2.1.i His_6-human Jagged2 DSL-EGF3

Since this construct was the first to be analysed, a comprehensive screening matrix was tested: two strains (BL21 (DE3) and BL21 Star (DE3)), two growth media (TB and PB), two IPTG concentrations (0.1 and 0.5mM), two temperatures and three induction durations (37°C for 2 and 4 hours and 20°C o/n). Growth in the absence of IPTG was performed as a negative control.

Table III.3 - Summary of small-scale expression screenings for each construct. For each experiment, OD for induction, Induction temperature (T), time (t), concentration of inductor ([IPTG]), expression media, host strain and co-expression of pGro7 are indicated by (x).

	OD_{600}		T (ºC)				t (h)				[IPTG] (mM)		Media		Host strain			pGro7
	0.5 -0.7	1.5 -2.0	37	30	27	18	2	4	6	o/n	0.1	0.5	TB	PB	BL21	BL21 Star	SHuffle®	
hJ2 DE3		x	x				x				x		x		x			
		x	x					x			x		x		x			
		x	x				x					x	x		x			
		x	x					x				x	x		x			
		x				x				x	x		x		x			
		x				x				x		x	x		x			
		x	x				x				x			x	x			
		x	x					x			x			x	x			
		x	x				x					x		x	x			
		x	x					x				x		x	x			
		x				x				x	x			x	x			
		x				x				x		x		x	x			
		x	x				x				x		x			x		
		x	x					x			x		x			x		
		x	x				x					x	x			x		
		x	x					x				x	x			x		
		x				x				x	x		x			x		
		x				x				x		x	x			x		
		x	x				x				x			x		x		
		x	x					x			x			x		x		
		x	x				x					x		x		x		
		x	x					x				x		x		x		
		x				x				x	x			x		x		
		x				x				x		x		x		x		
	x		x				x				x			x	x			x
hJ2 ME3	x		x				x				x			x	x			x
hJ2 ME9	x		x				x				x			x	x			x
	x		x					x				x		x	x			x
	x		x				x				x			x	x			x
	x		x					x				x		x	x			x
	x			x			x				x			x	x			x
	x			x				x				x		x	x			x
	x					x	x				x			x	x			x
	x					x		x				x		x	x			x
		x	x					x			x			x	x			
		x	x				x					x		x	x			
		x		x				x			x			x	x			
		x		x					x			x		x	x			
		x				x				x	x			x	x			
		x				x				x		x		x	x			
		x			x					x	x			x			x	
		x			x					x		x		x			x	

	OD_{600}		T (ºC)				t (h)				[IPTG] (mM)		Media		Host strain			pGro7
	0.5 -0.7	1.5 -2.0	37	30	27	18	2	4	6	o/n	0.1	0.5	TB	PB	BL21	BL21 Star	SHuffle®	
hJ2 ME9		x				x				x	x			x			x	
		x				x				x		x		x			x	
hJ2 DE9	x		x				x				x			x	x			x
	x		x				x					x		x	x			x
	x		x					x			x			x	x			x
	x		x					x				x		x	x			x
	x			**x**				**x**			**x**			**x**	**x**			**x**
	x			x				x				x		x	x			x
	x			**x**					**x**		**x**			**x**	**x**			**x**
	x			x					x			x		x	x			x
	x					x				x	x			x	x			x
	x					x				x		x		x	x			x
hD1 ME3		x		x					x			x		x	x			
		x				x				x		x		x	x			
hD1 DE3		x		x					x			x		x	x			
		x				x				x		X		x	x			
hD1 ME6		x		x					x			X		x	x			
		x				x				x		X		x	x			
hD1 DE6		x		x					x			X		x	x			
		x				x				x		X		x	x			

The total protein extracts, corresponding to the test conditions shaded in green in Table III.3, were analysed by SDS-PAGE and the expected size for the protein (19.6KDa) is marked with a blue arrow in Figure III.2-A and -B.

As one can observe, for hJag2 DE3 construct the small-scale screenings suggest that the best induction conditions are achieved when using BL21 (DE3) as host strain and PB as growth media (Figure III.2-A, right half of golden box). In this particular case, a similar intensity band is observed in all conditions where protein expression was induced, irrespective of IPTG induction concentration, time or temperature of induction. When using TB as growth media a band was also detected corresponding to the expression of the protein of interest, however the intensity is lower than that achieved when using PB as growth media (Figure III.2-A, left half of golden box).

The results obtained when expressing this construct using BL21 Star (DE3) as host strain failed to deliver any positive result (Figure III.2-B, red rectangle). Neither of the tested conditions showed to have any expression of hJag2 DE3 upon induction since the band profile observed was similar to the non-induced control sample (comparing the lanes where IPTG concentration is 0 and 0.1 or 0.5mM). This result was surprising since the use of this strain had the purpose of augmenting recombinant protein expression. This strain has the *rne131* mutation and the cells are defective on RNaseE, decreasing mRNA degradation throughout expression which enhances protein expression as the expression rate depends on mRNA stability (Sorensen and Mortensen 2005). Possibly the tested expression conditions using BL21 Star (DE3) as host strain were not the most suitable with the combination of media, induction OD and expression time for this particular construct.

A total absence of protein expression using BL21 Star (DE3) was not expected, but based on these observations the use of this strain was discarded for further studies.

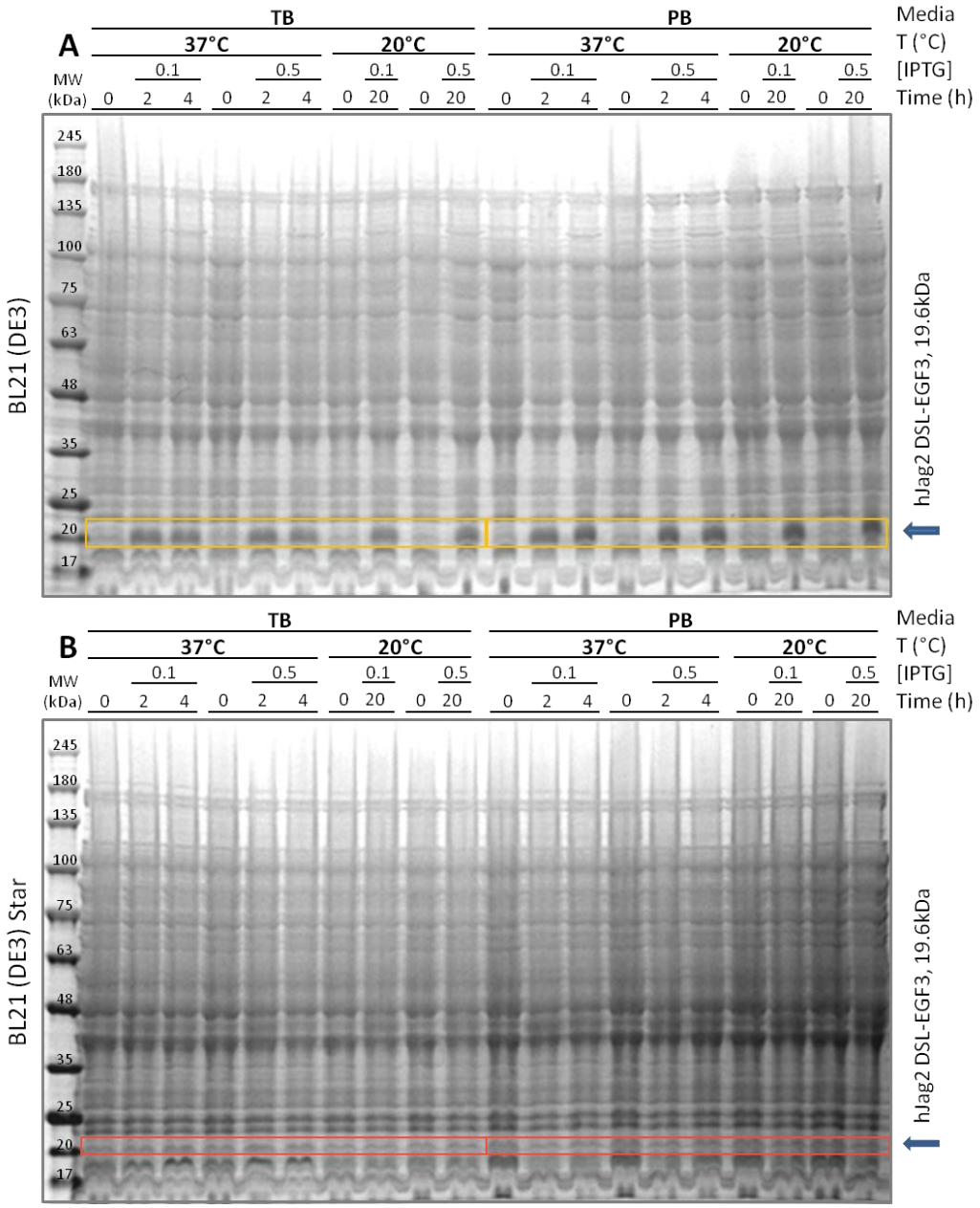


Figure III.2 - SDS-PAGE of the initial small-scale expression screening results for human Jagged2 DSL-EGF3. Panel A – expression using BL21 (DE3) as host strain; Panel B – expression using BL21 (DE3) Star as host strain. Growth media, Induction Temperature, IPTG concentration in millimolar and duration of induction in hours are indicated in the figure. Expected size of the protein of interest highlighted with the blue arrow on the right in both panels. Red rectangle in Panel B highlights the absence of expression using BL21 (DE3) Star as host strain.

Bearing in mind the best results for protein expression for the human Jagged2 DSL-EGF3 construct that were obtained using BL21 (DE3) as host strain and PB as growth media (as shown in Figure III.2-A), the best induction conditions for this construct were evaluated for protein solubility. BugBuster reagent was used to disrupt the cells, and the results were analyzed by Western Blot (Figure III.3). Western blot is a far more sensitive technique to allow for a more accurate evaluation of the ratio of soluble/insoluble protein of interest.

Samples taken from cultures after induction for 2 hours at 37°C show the highest percentage of soluble *vs.* insoluble protein of interest, more than the samples from cultures induced o/n at 20°C. We estimated about 10-fold more soluble protein by comparing the protein bands corresponding to the soluble fraction upon induction at 37°C compared to that at 20°C (Figure III.3).

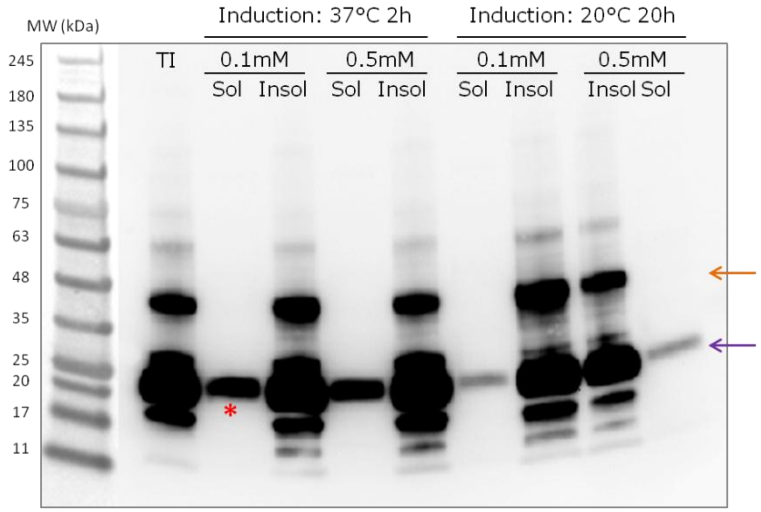


Figure III.3 - Western Blot analysis of the solubility tests for human Jagged2 DSL-EGF3 construct using BL21(DE3) as host strain and PB as growth media. Epi-white image of the molecular markers was superposed to the chemiluminescence image acquired after 51 seconds. TI - Total induced sample, Sol - Soluble fraction, Insol - Insoluble fraction. The sample marked by a red asterisk was used as reference in Figure III.4 and Figure III.5.

As one can observe in Figure III.3, only around 10% of the total protein of interest is soluble in the conditions tested so far. This result was not considered good enough to proceed for large scale production based on the conditions already obtained. Our results somewhat disagree with what is commonly reported, that induction with lower inducer concentration for longer periods of time at reduced temperatures increases protein solubility (de Marco, Vigh *et al.* 2005). It is possible that this hJag2 construct is toxic to bacterial cells, impeding their growth. It may also be the case that, upon a longer expression time, the protein concentration becomes higher than the bacterial cell can endure. Those reasons may reflect that most of the produced protein is stored in the form of inclusion bodies for cellular protection. Possibly, a shorter pulse of high expression rate can be resulting in a higher percentage of soluble protein for this hJag2 construct.

As such, a new set of screenings was performed with the aim of improving protein solubility. To increase the soluble expression of our protein, co-expression of Jag2 construct together with the molecular chaperones groEL and groES encoded by pGro7 was pursued. Increased recombinant protein solubility can be achieved in the presence of groEL and groES co-expression with lower inducer concentrations as been reported previously (de Marco, Vigh *et al.* 2005).

The conditions for these tests were based on previous optimizations, however the use of pGro7 implies several alterations such as lowering the induction OD_{600} to 0.5-0.7 and adding L-arabinose as an inducer for chaperone expression. The influence of the lysis method in the percentage of soluble protein obtained while co-expressing the molecular chaperone was also evaluated using three lysis methods: BugBuster commercial reagent, a home-made lysis buffer supplemented with lysozyme and

DNase and a combination of the lysis buffer and mechanical disruption by sonication. The results were analysed by SDS-PAGE and WB (Figure III.4).

As one can observe in Figure III.4-A, all lysis conditions and methods tested were efficient in terms of cell lysis as we could obtain a wide range of soluble proteins. However, no intense band with the expected molecular weight of hJag2 DE3 (19.6kDa, Table III.1), indicated by the orange arrow, was detected; moreover, an intense band in all the insoluble fraction samples was observed with the expected size for hJag2 DE3, as indicated by the orange box.

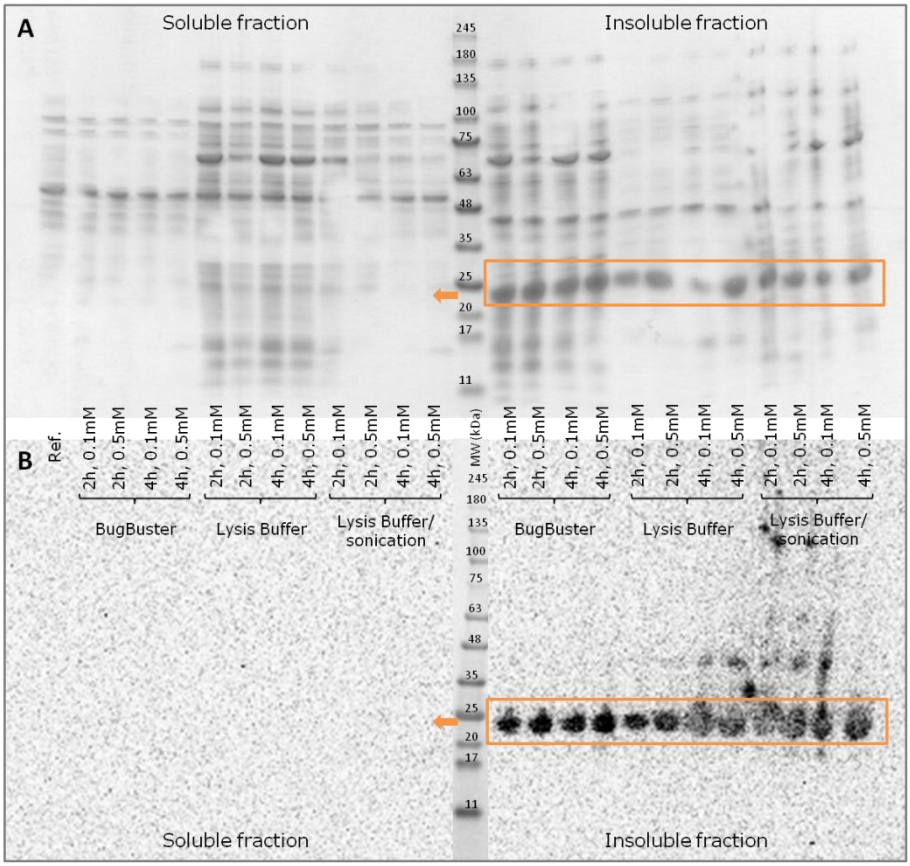


Figure III.4 - Expression and lysis conditions screening for human Jagged2 DSL-EGF3 co-expressed with pGro7. Panel A – SDS-PAGE, Panel B – Western Blot, exposition time is 300 seconds. Ref. stands for reference sample (marker with a red asterisk in Figure III.3). Orange arrows and box mark the expected size and the bands corresponding to this protein (19.6kDa, Table III.1). Samples are grouped according to fraction after lysis, soluble on the left and insoluble on the right, and lysis method as indicated in the figure.

These observations are confirmed by the WB result in Figure III.4-B, where a protein band with the size corresponding to the His-tagged human Jagged2 DSL-EGF3 protein appears only for the insoluble fraction samples. As no protein was detected in the soluble fraction samples, no conclusions could be drawn from this result regarding the best conditions for obtaining soluble protein. However, when using the home-made lysis buffer supplemented with lysozyme and DNase we obtained the best lysis condition since the extracted proteins show a wider range in terms of molecular weight than the two other lysis methods employed.

The Soluble fraction samples analysed in Figure III.4 were subjected to a second analysis by Western Blot, where a larger volume of sample was analysed in order to increase the total protein

amount per lane, and to allow evaluating more accurately the solubility profile of the tested expression and lysis conditions. The results are presented in Figure III.5.

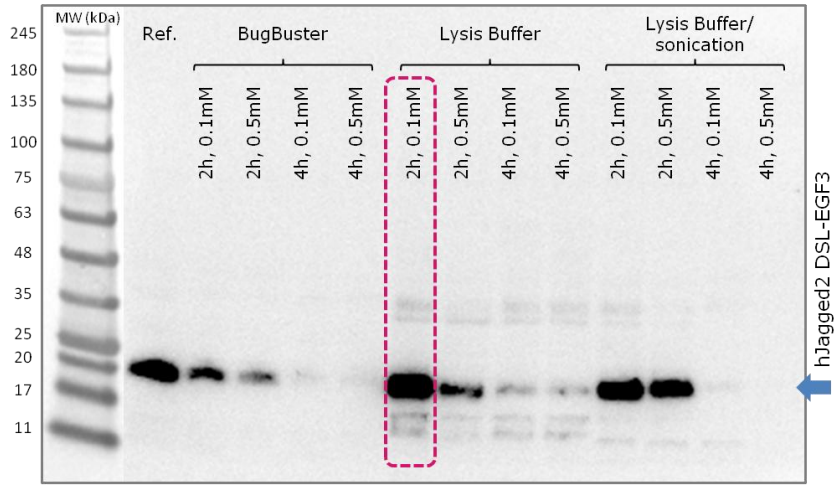


Figure III.5 - Western Blot analysis of soluble fraction samples from Figure III.4. Exposition time is 160 seconds. Ref. stands for reference sample (marker with a red asterisk in Figure III.3). Blue arrow marks the expected size for this protein (19.6kDa, Table III.1).

The sample from Figure III.3, marked with a red asterisk (soluble fraction obtained when inducing expression for 2 hours at 37°C with 0.1mM IPTG without co-expressing the molecular chaperones where about 10% soluble protein could be obtained) was used as reference sample for Figure III.5. The results obtained showed to be promising: protein expression under the conditions of induction with 0.1mM IPTG for 2 hours at 37°C and co-expression with the chaperone plasmid results in a higher amount of soluble protein (Figure III.5 – fuchsia box) compared to the protein previously obtained. Also, induction with lower concentration of IPTG also results in higher solubility in each induction time tested. This result is in agreement with the afore mentioned hypothesis that a short pulse of intense protein expression can increase the percentage of recoverable soluble protein.

We could then estimate to obtain 20 to 25% soluble protein when inducing protein expression in the presence of the groEL and groES chaperones at 37°C for 2 hours with 0.1mM IPTG, and lysing the cells with the home-made lysis buffer supplemented with lysozyme and DNase (as there is a ~2 fold increase in band intensity between the reference sample, previously estimated for 10% soluble protein, and the described condition).

Therefore, we determined that for large-scale protein production and purification for this construct of human Jagged2 DSL-EGF3 we would use the following conditions: BL21 (DE3) host strain, PB as growth media, co-expression of the protein of interest with pGro7, induction of protein expression at OD_{600} 0.5-0.7 at 37°C for 2 hours with 0.1mM IPTG and 4mg/mL L-arabinose followed by lysis with home-made lysis buffer.

III.2.1.ii His_6-human Jagged2 MNNL-EGF3, MNNL-EGF9 and DSL-EGF9

In order to evaluate if there was a different construct of human Jagged2 that is amenable for expression of soluble protein, the remaining three hJagged2 constructs were expressed in an initial

small scale expression test based on the best expression conditions set for the human Jagged2 DSL to EGF3 construct reported above in Section III.2.1.i - Induction with 0.1mM IPTG at OD_{600} 0.5 to 0.7 for 2 hours at 37°C with co-expression of groEL and groES chaperones, using BL21 (DE3) and PB. The results were analysed by SDS-PAGE and WB (Figure III.6).

In the SDS-PAGE analysis of hJagged2 constructs MNNL-EGF3, MNNL-EG9 and DSL-EGF9 expression (Figure III.6-A) no soluble protein was detected. Nevertheless, the presence of the said proteins with the correct expected sizes could be observed in the lanes corresponding to the total induced and insoluble fractions. The Western blot analysis in Figure III.6-B confirmed the expression of the proteins with the correct sizes in the Insoluble fractions, and also showed that there was no soluble protein expressed, as expected by visualizing the SDS-PAGE in panel A.

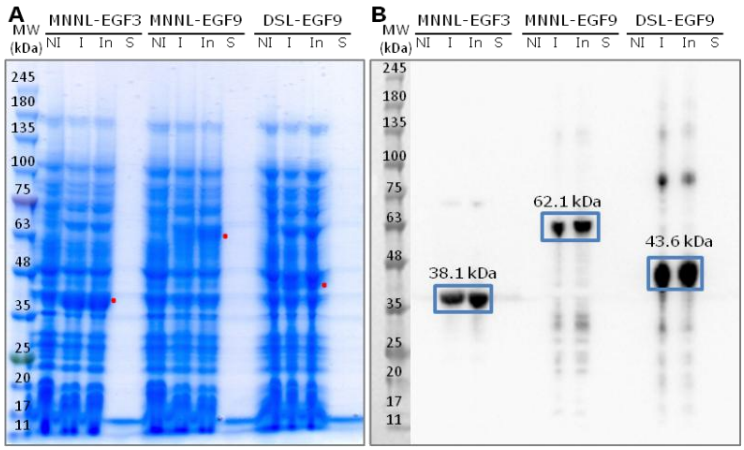


Figure III.6 - Initial small scale expression test for human Jagged2 constructs MNNL-EGF3 (38.1kDa), MNNL-EGF9 (62.1kDa) and DSL-EGF9 (43.6kDa). Panel A – SDS-PAGE. Red dots mark the expected sizes for each construct; Panel B – Western Blot. Exposition time is 220 seconds. NI – Non-induced culture; I – Total induced culture; In – Insoluble fraction; S – Soluble fraction.

As such, a more comprehensive expression and solubility screening was performed for the human Jagged2 constructs bearing domains MNNL-EGF9 and DSL-EGF9 by testing different temperatures, times and IPTG concentrations, while maintaining BL21 (DE3), PB media and co-expression with pGro7 (Figure III.7). We chose to pursue with these two MNNL-EGF9 and DSL-EGF9 constructs of Jagged2 as they comprise more domains and could provide us more structural and biophysical information.

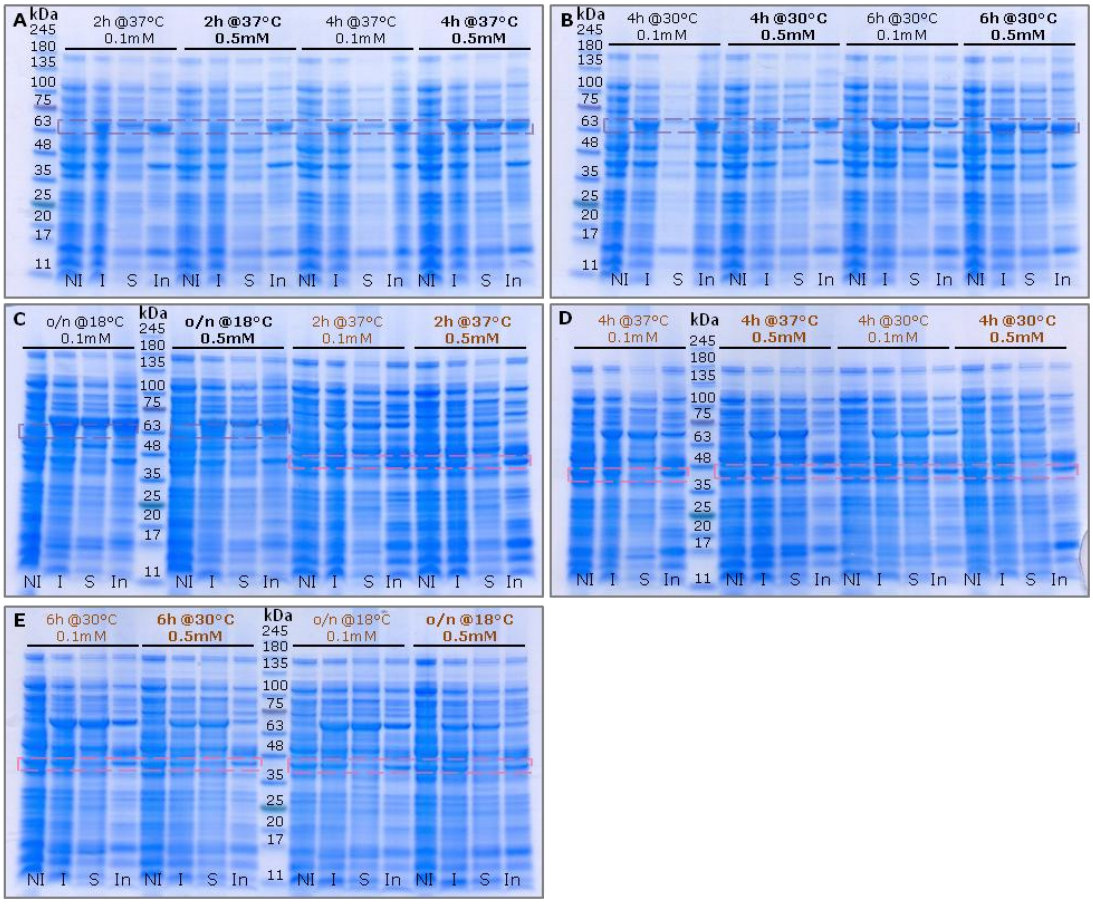


Figure III.7 - SDS.PAGE analysis of the small-scale expression screening for fine tuning for hJagged2 expression conditions. Results for hJagged2 MNNL-EGF9 indicated in black and for hJagged2 DSL-EGF9 in brown. Induction duration, temperature and IPTG concentration depicted above each condition. Purple and pink dashed rectangles mark the expected size of hJagged2 MNNL-EGF9 and hJagged2 DSL-EGF9, respectively. NI – Non-induced culture; I – Total induced culture; S – Soluble fraction; In – Insoluble fraction.

From the results presented in the SDS-PAGE in Figure III.7, no bands could be detected in the respective lanes of the soluble fractions for neither proteins in each test condition. These expression tests were performed with co-expression of pGro7 (enconding GroEL/ES molecular chaperones, with ~60-65kDa), and the expression of these protein is observed in all analysed soluble fractions' lanes. It is especially evident in Fig. III.7-C the presence of two distinct bands in total induced fractions, corresponding to the molecular chaperones and hJagged2 MNNL-EGF9, but not in the lane corresponding to the soluble fraction.

We could conclude however that the lysis efficiency was greater than in the previous tests since a broader range of proteins could be observed in the lanes corresponding to the soluble fractions (Figure III.6). The soluble fractions analysed by SDS-PAGE in Figure III.7 were subject to a more thorough analysis by Western Blot as shown in Figure III.8.

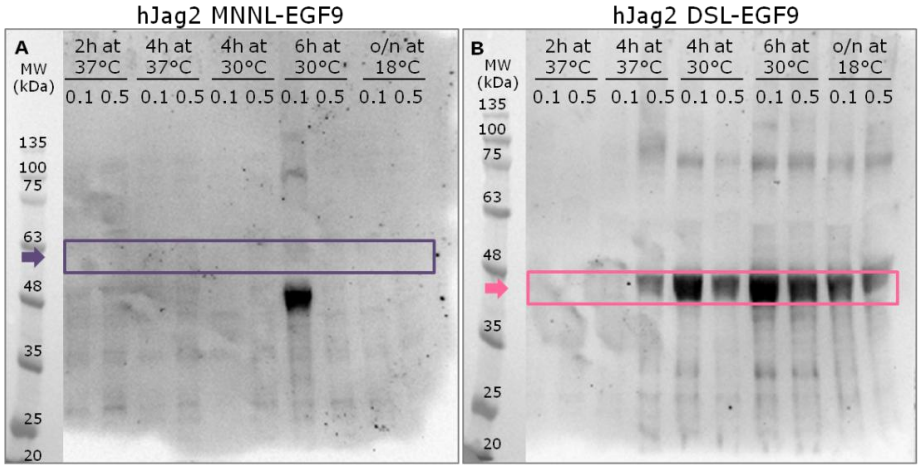


Figure III.8 - Western Blot analysis of soluble fractions from Figure III.7. Panel A – hJagged2 MNNL-EGF9 (62kDA); Panel B - hJagged2 DSL-EGF9 (43.6kDa) (Exposition time is 300 seconds). Induction conditions and IPTG concentration (in millimolar) for induction as depicted above each lane.

The results show that human Jagged2 MNNL-EGF9 (Figure III.8-A) was not expressed in any of the conditions tested in the soluble fraction, and the only observed band does not have the expected size for this protein (62kDa, Table III.1) but we cannot exclude possible proteolysis. The fact that neither of the conditions tested gave positive results for this protein construct could be due to its higher molecular weight and complexity. The higher complexity of Jag2 MNL-EGF9 protein, with 30 predicted disulfide bonds (1 DSL and 9 EGF domains with 3 disulfides per domain), can be overwhelming for the cellular machinery causing not only translational stalling, but also can influence the redox potential of the bacteria's cytoplasm, aggravating the overall tendency of expressed protein insolubility, as observed for the previous constructs of human Jagged2.

On the other hand, the results obtained for the human Jagged2 construct bearing domains DSL-EGF9 (43.6kDa, Table III.1) showed to be promising with soluble protein expressed under induction with 0.1mM IPTG at 30°C for 4 or 6 hours as one can observe in Figure III.8. As such, the expression conditions chosen for a large scale expression for this construct were induction at OD_{600} 0.5-0.7 with 0.1mM IPTG for 4 or 6 hours at 30°C, using BL21 (DE3), PB media and co-expression with pGro7 plasmid.

III.2.1.iii His_6-human Delta-Like Ligand1

In order to assess the expression and solubility profiles for the human Delta-like Ligand1 constructs, a small-scale expression screening based on our previous results was set up, with BL21 (DE3), PB media, and protein induction at OD_{600} 1.5-2.0 at 30°C for 6 hours or 18°C o/n with 0.5mM IPTG. The results are presented in Figure III.9.

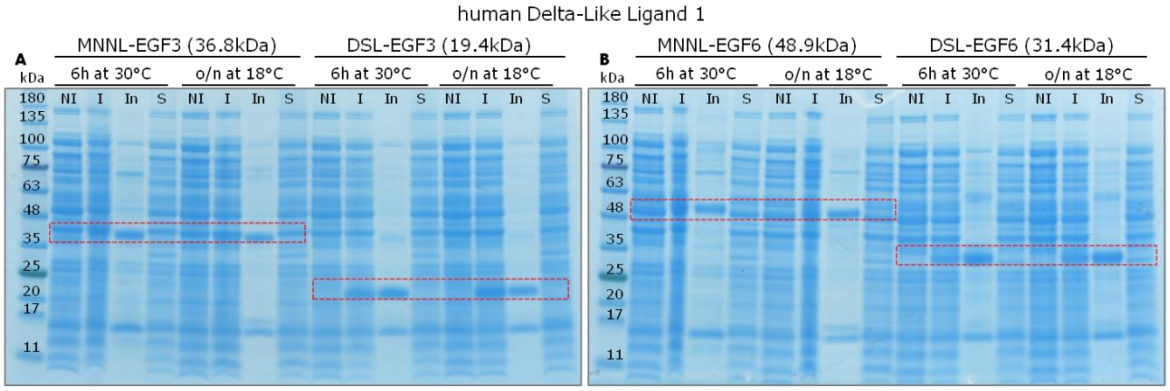


Figure III.9 – SDS-PAGE analysis of human DLL1 constructs' small scale expression screenings. Expression conditions are indicated above each set of four lanes. Red rectangles mark the expected sizes for each protein. NI – Non-induced culture; I – Total induced culture; In – Insoluble fraction; S – Soluble fraction.

As one can observe in Figure III.9, all four constructs show good expression profiles: there is an evident band in the lanes corresponding to total induced extract (I, red rectangles), compared to the total non-induced extract (NI). However, in terms of solubility, the proteins were only detected in the insoluble fraction (Figure III.9, In). In all constructs, irrespective of induction condition tested, no clear band corresponding to the expected size of the proteins is seen in the soluble fraction by SDS-PAGE (Figure III.9, S). We therefore concluded that with the tested expression conditions little or no soluble protein could be obtained. Perhaps a Western blot analysis could reveal with greater detail the solubility profile of these constructs. Taking the knowledge that by that time we had acquired, we did not further analyse these samples. The purification tests of hJag2 DSL-EGF3 (Section III.2.3.i) and DSL-EGF9 (Section III.2.3.ii) constructs were performed prior to this analysis and we already had a hint that, without a visible band corresponding to the protein of interest as visualised by SDS-PAGE, there would be no interest in pursuing an expression condition that yields low amounts of soluble protein.

In summary, we observe that although all the expression experiments using BL21 (DE3) strain can deliver good results in terms of total protein expression almost all the amount of the recombinant protein produced was insoluble and trapped into the inclusion bodies.

III.2.2 Small-scale expression screenings with SHuffle® strain

Since that the use of BL21 (DE3) strains as host for protein expression failed to fulfil the need of fair amounts of soluble protein amenable for protein purification, we sought for a modified strain capable of disulfide formation in *E. coli* cytoplasm as discussed in Chapter I. We chose SHuffle® T7 Express (Lobstein, Emrich *et al.* 2012) since this strain has both the advantages of *E. coli* cytoplasmic expression systems, combined with the cytoplasmic expression of periplasmic chaperones. Moreover, SHuffle® has cellular reductases deletions that allow for a more oxidizing cytoplasmic environment tolerating disulfide bond formation in the cytoplasm, and it is adequate to use with T7 promoter-based vectors such as pET-47b(+).

As such, a simple small-scale screening was carried out using the conditions described by Lobstein, Emrich *et al.* (2012), where we tested the human Jagged2 constructs bearing the smaller and largest number of domains - DSL-EGF3 and MNNL-EGF9. The construct comprising domains DSL-EGF3 had been already extensively studied so it was screened using SHuffle® as a quick control for changes in expression and solubility behaviour. The results are presented in Figure III.10.

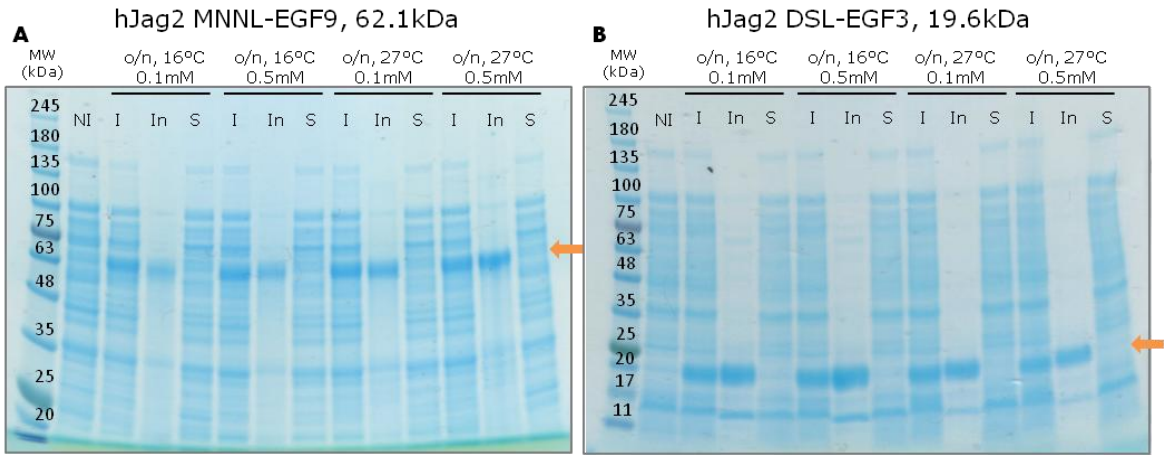


Figure III.10 - Small scale expression screenings for human Jagged2 MNNL-EGF9 (panel A) and DSL-EGF3 (Panel B) using SHuffle® as host strain. Induction conditions are depicted above each set of lanes. NI – Non-induced culture; I – Total induced culture; In – Insoluble fraction; S – Soluble fraction. Blue arrows point to the expected molecular weight for each protein.

Using SHuffle® strain as expression host we could obtain a good expression of both human Jagged2 constructs screened as there is an evident band in the total extracts of induced cultures (I) compared to the culture before induction (NI). Regarding solubility, it is again apparent for both constructs that no soluble protein could be produced. There could be several reasons for our results: the expression conditions used can be sub-optimal for SHuffle® cells to properly form the disulfide bonds. Moreover, Lobstein, Emrich *et al.* (2012) concluded that this strain's effect in disulfide bond formation in the cytoplasm is protein-specific, which may not be the case of this study.

Taken together our results with both BL21 and SHuffle® strains, we were not able to generate more than 20 to 25% estimated soluble protein. However we chose the human Jagged2 DSL-EGF3 and DSL-EGF9 constructs to carry out purification trials of the protein available in the soluble fraction.

III.2.3 Protein Purification from soluble fraction

III.2.3.i Human Jagged2 - His6-hJAG2 DSL-EGF3

Bacterial cell pellet of *E. coli* BL21 (DE3) expressing human Jagged2 DSL-EGF3 (19.6kDa, Table III.1) was lysed and the clarified soluble fraction (sample A) was injected in a HisTrap column connected to an ÄKTApurifier system. The insoluble fraction was stored at -80°C. The bound protein was eluted with a continuous gradient of imidazole and protein fractions were collected.

The elution chromatogram is showed in Figure III.11. One single protein peak was eluted with concentrations between 90 and 235mM imidazole. Fractions of the peak were analysed by SDS-PAGE for the presence of hJag2 DSL-EGF3 (Figure III.12), and a protein pool consisting of fractions 7 to 11 (right side of the peak) was desalted to HRV-3C protease cleavage reaction Buffer. Total protein was determined by Bradford Protein assay to be 9.53mg.

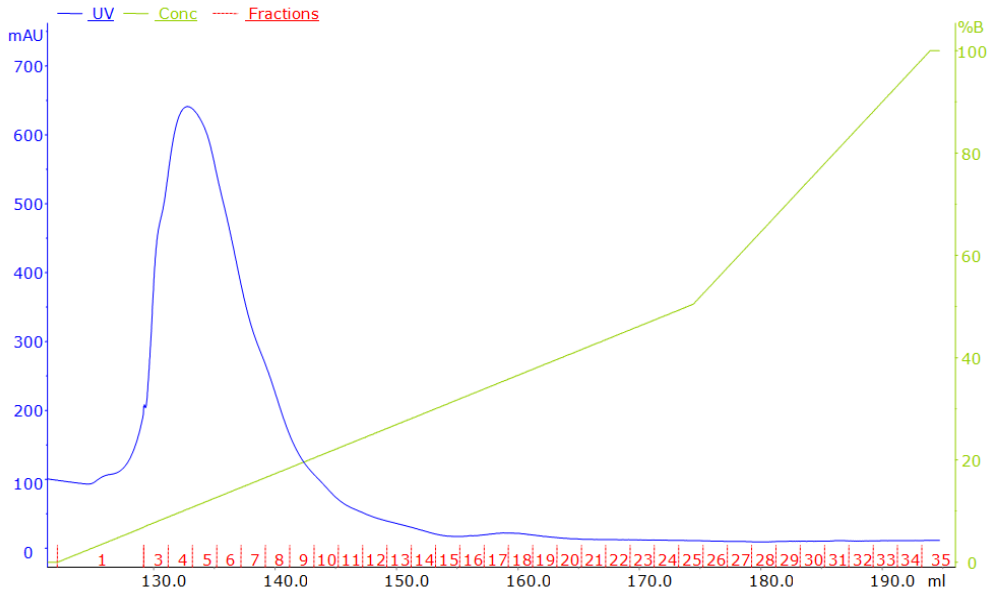


Figure III.11 - Elution chromatogram of HisTrap column for the purification of human Jagged2 DSL-EGF3. Bound protein was eluted with a continuous gradient of Buffer B (in green, right vertical axis). Eluted protein was monitored by absorbance at 280nm and recorded (in blue, left vertical axis). The eluate´s fraction numeration is shown in red, in the bottom.

After His-tag cleavage with HRV-3C protease, the protein mixture was separated in a tandem column setting consisting of a GSTrap column connected directly to a HisTrap column, allowing that both the protease (that was engineered to have a GST-tag), uncleaved protein and cleaved tag are separated from the untagged protein, which will not bind to the column and therefore will be collected in the flow-through. GSTrap and HisTrap columns were eluted and samples from each key step were analysed by SDS-PAGE to access cleavage efficiency (Figure III.13).

As shown in Figure III.13, His-tag cleavage did not appear to be efficient as the percentage of our protein of interest in the purified cleaved protein sample was very low. As depicted in the figure, both uncleaved (Red dot) and cleaved protein (Green dot) are present in the samples corresponding to the tag cleavage reaction and the purified protein after cleavage. Moreover, there is also a large percentage of uncleaved hJag2 DE3 in the eluate of HisTrap column.

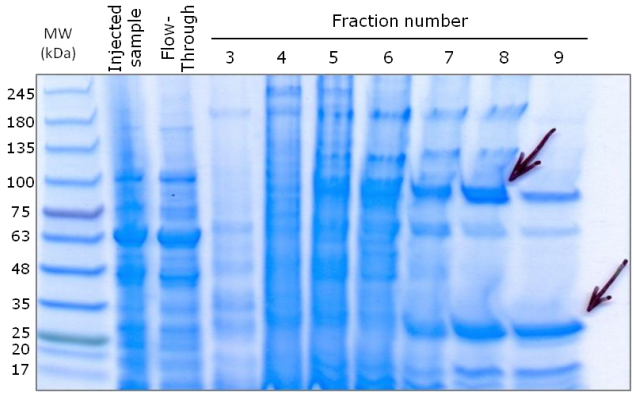


Figure III.12 - SDS-PAGE analysis of the fraction resulting from the first chromatographic step of hJag2 DE3 purification. Arrows point to the presence of hJag2 DE3 possibly in both monomeric (bottom arrow) and tetrameric (top arrow) forms.

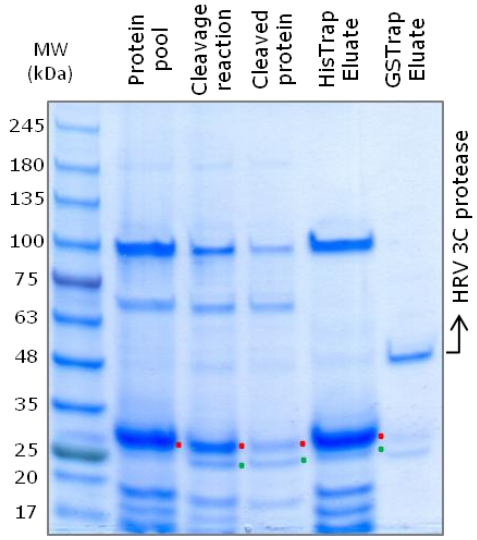


Figure III.13 - SDS-PAGE analysis of hJag2 DE3 His-tag cleavage with HRV-3C protease. Red dots mark uncleaved protein. Green dots mark cleaved protein

As depicted in Figure III.12, the last analysed fraction still contains a great amount of our protein of interest. In order to assess if the target protein was also present in the remaining fractions of the first chromatographic step (eluted with high imidazole concentrations), western blot analysis was performed. As one can observe in Figure III.14 in fact, all the sampled fractions contained the protein of interest.

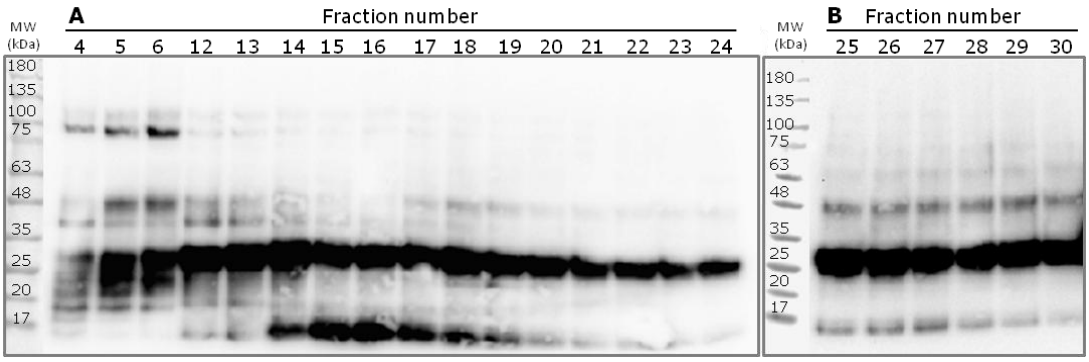


Figure III.14 - Western Blot analysis of the first chromatographic step of hJag2 DE3 purification. All fractions of the chromatogram were sampled, except those already pooled in Sample B (see above in the text).

Hence, since HRV-3C protease cleavage was not 100% efficient and the protein of interest was present in all eluted fractions of the chromatogram a second protein pool was prepared combining also the remaining fractions and the protein eluted from the HisTrap (see Figure III.13). This second protein pool was then subjected to the same procedure as Sample B, namely, desalting to HRV-3C protease cleavage buffer, protein concentration determination (428.4mg), tag cleavage overnight, and cleaved protein purification using the same column tandem setup described earlier. Flow-through was collected and the efficiency of this purification was accessed by SDS-PAGE and WB (Figure III.15).

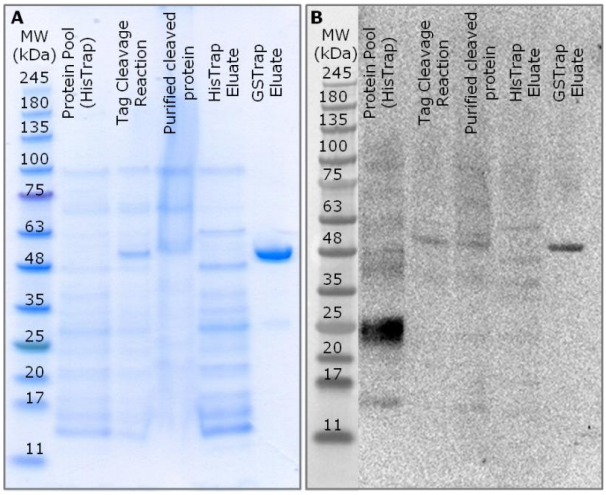


Figure III.15 - SDS-PAGE (Panel A) and WB (Panel B) analysis of second hJag2 DE3 His-Tag cleavage with HRV-3C protease.

From the analysis of Figure III.15-A, it is clear that the second pool contains less protein of interest than the first pool (Figure III.13, lane 1), and therefore no intense band corresponding to hJag2 DE3 could be visualised by SDS-PAGE in neither samples therein analysed. However, the WB analysis of the same samples revealed the presence of the target protein in this second pool (Figure III.15-B, lane 1). We assumed here that the cleavage reaction was more effective than the previous one, since no protein band was detected by WB upon protease addition, namely in the tag cleavage reaction, in purified cleaved protein and HisTrap elution samples (Figure III.15-B). Still, the protein mixture contained several contaminant proteins.

The protein mixture was concentrated and injected in a SEC column in order to separate our protein of interest from the remaining contaminants. The resulting chromatogram is shown in Figure III.16. Selected fractions were analysed by SDS-PAGE (Figure III.17).

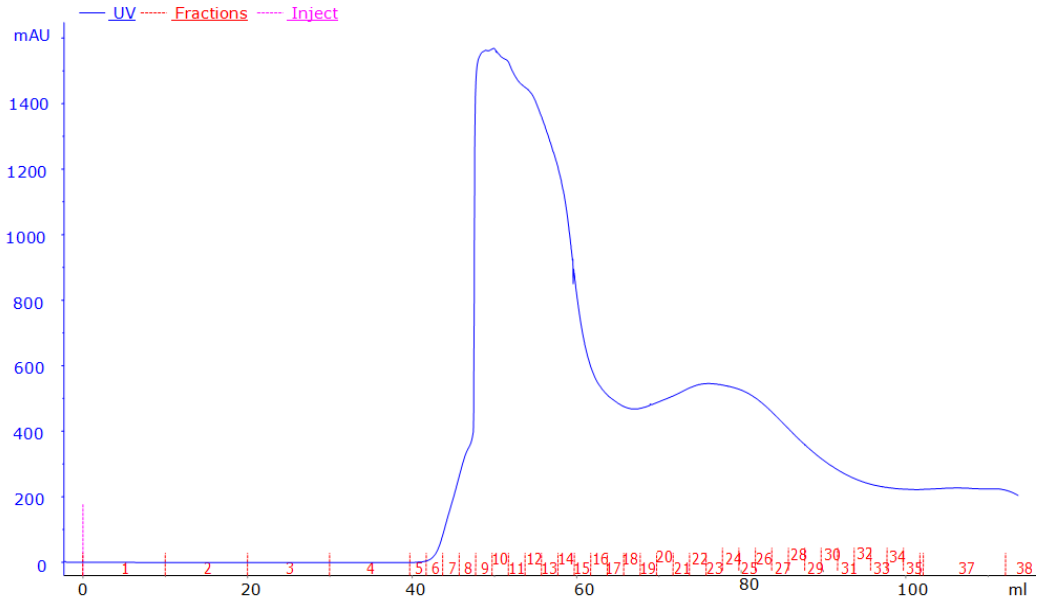


Figure III.16 - Size Exclusion Chromatogram of hJag2 DSL-EGF3. Injection is marked in pink and volume was corrected to zero at that point. Protein was monitored by absorbance at 280nm and recorded (in blue, left vertical axis). The eluted fraction numeration is shown in red, in the bottom.

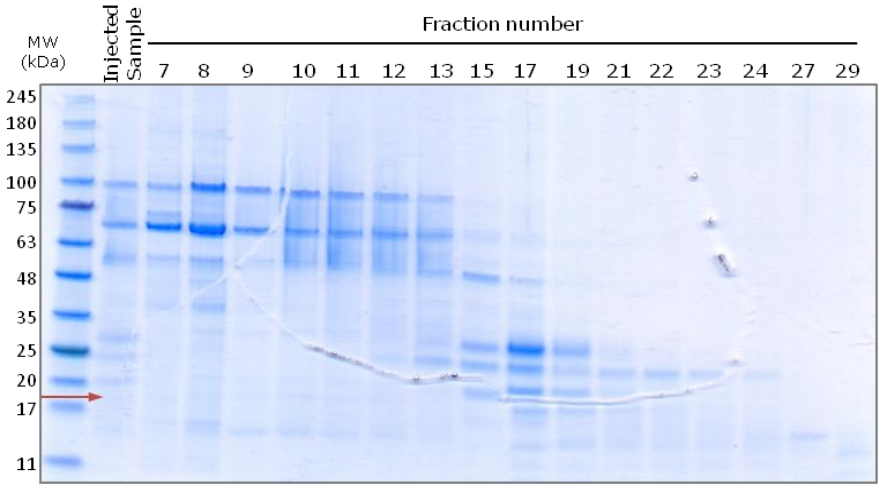


Figure III.17 - SDS-PAGE analysis of SEC fractions for the purification of human Jagged2 DSL-EGF3. Red arrow marks the expected size for human Jagged2 DSL-EGF3 after tag removal (17.6KDa).

Rationalizing the chromatogram with the SDS-PAGE in Figure III.17, the first peak corresponds to the proteins not resolved by this column, i.e., proteins with higher molecular weight than the upper limit of resolution for this column; the broad peak corresponds to an unresolved mixture of lower molecular weight proteins (Figure III.17, fractions 17 to 29). We conclude that all lanes have multiple proteins and no intense band corresponding in the expected size for human Jagged2 DSL-EGF3 protein, which is 17.6kDa after tag removal, was evident by SDS-PAGE analysis (Figure III.15, lane 3 and Figure

III.17, lane 1). It would have been crucial to finish the first chromatographic step with a more pure sample in order to facilitate the identification of our protein of interest by SDS-PAGE following the SEC step. Moreover, in Figure III.17 any of the myriad of bands could point to the presence of different oligomers as one may find bands corresponding to the expected size of dimers, trimers and tetramers.

Therefore, a new purification strategy was pursued, using the inclusion bodies from the expression experiments since, by observing the western blots from Figure III.3 and Figure III.4, a large portion of the protein remains insoluble after lysis when expressed using the stated conditions.

III.2.3.ii Human Jagged2 - His6-hJAG2 DSL-EGF9

Bacterial cell pellet of *E. coli* BL21 (DE3) expressing human Jagged2 DSL-EGF9 (43.6kDa, Table III.1) was lysed and the clarified soluble fraction (sample A) was injected in a HisTrap column. The bound protein was eluted with steps of imidazole corresponding to a 2% increase in Buffer B in each step between 70 and 210mM imidazole. Afterwards, a continuous gradient was applied between 210 and 500mM imidazole in 10 CV, and between 500 and 1000mM imidazole in 2 CV. Eluted protein fractions were collected (Fig. III.18).

For this purification, the elution was performed with steps rather than a gradient in order to achieve a better protein separation in this first chromatographic step as discussed in the previous section for human Jagged2 DSL-EGF3 purification. We opted for 20mM imidazole steps in order to generate, by discrete increments in imidazole concentration, a fine-tuned elution profile. Several peaks were eluted with each step of increasing concentration of imidazole. Fractions corresponding to each peak were analysed by SDS-PAGE and WB for the presence of hJag2 DE9 protein (Fig. III.19).

Figure III.18 - Elution Chromatogram from HisTrap column for the purification of human Jagged2 DSL-EGF9. Bound protein was eluted with a combination of steps with increasing concentration and a continuous gradient of Buffer B (in green, right vertical axis). Eluted protein was monitored by absorbance at 280nm and recorded (in blue, left vertical axis). The numeration of the analysed fractions is shown in red, in the bottom.

As one can observe in Fig.III.19, no band with the expected molecular mass for hJag2 DE9 (43.6kDa, Table III.1) is evident neither by SDS-PAGE (Panel A) or WB (Panel B) analysis. The SDS-PAGE analysis reveals low molecular weight contaminants spread in all tested fractions and one 100kDa protein in fractions 17 to 29, corresponding to two consecutive peaks, eluted with 90 and 110mM imidazole, respectively. Moreover, the WB analysis reveals the presence of multiple proteins throughout the analysed fractions. Those proteins are not visible in the SDS-PAGE analysis, and most likely are those *E. coli* contaminants that also react with anti-His-tag antibody besides interacting with the HisTrap column.

The possibility that our protein could be forming stable oligomers was discarded as in Fig. III.19-B there is not a positive band for His-tagged proteins corresponding to any possible size of hJag2 DE9 as a monomer or oligomer. Therefore, we concluded that a new strategy design to produce and purify this protein was also needed as for the purification of human Jagged2 DSL-EGF3.

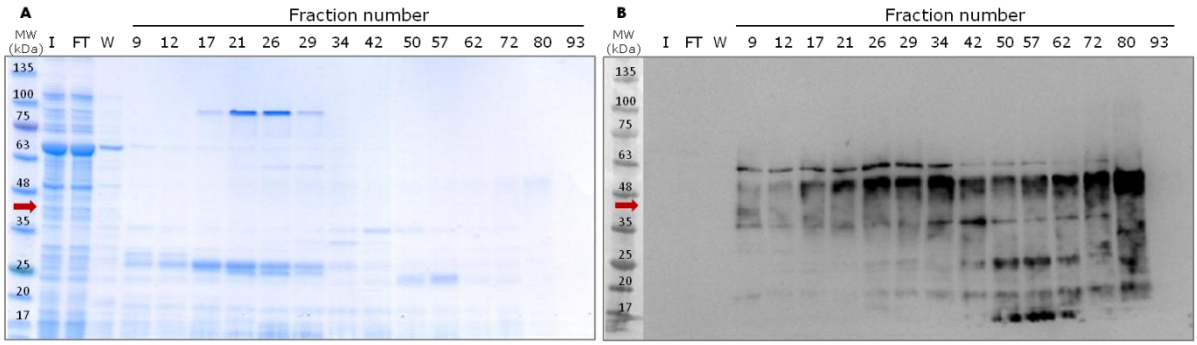


Figure III.19 - Analysis of the HisTrap Chromatographic step fractions for hJagged2 DSL-EGF9 (44kDa). Panel A - SDS-PAGE; Panel B - WB, exposition time is 240 seconds. Blue arrows mark the expected size for this protein. I - Injected sample, FT - Flow-through, W - Wash, Numbers correspond to fraction numbers in the chromatogram in Fig. III.18.

III.2.4 Purification of His6-hJAG2 DSL-EGF3 from inclusion bodies

Inclusion bodies are frequently formed upon recombinant protein expression in *E. coli*. These inactive protein aggregates accumulate in the bacterial cells due to the high metabolic burden of recombinant protein expression and/or due to toxicity issues. Several studies were conducted using model proteins for *in vitro* refolding studies, but highly disulfide-bonded protein still pose enormous challenges (Rudolph and Lilie 1996). In this article, the authors go through the parameters affecting both inclusion body solubilisation and refolding such as the use of different reagents. The strategy followed was to purify the His6-hJAG2 DSL-EGF3 from inclusion bodies. The protocols followed were based on that information and in the protocols by Marko Hyvonen's group (University of Cambridge, Department of Biochemistry), available online (http://www-cryst.bioc.cam.ac.uk/groups/hyvonen/methods, February 2013). In search for the best refolding conditions for this protein construct, several different solubilisation (GuHCl and Urea) and refolding (oxido-shuffling pairs, buffer strengths, additives) reagents were tested (see Section II.4.2.ii for the exact conditions).

III.2.4.i Small-scale refolding tests

The small-scale expression tests for human Jagged2 DSL-EGF3 construct (Figure III.3; Figure III.4) showed that a large portion of the produced protein is insoluble. Therefore, the insoluble fraction was used for the initial inclusion bodies solubilisation and refolding tests. We used two of the most used solubilisation agents, GuHCl and Urea and two oxido shuffling pairs GSH/GSSH and cysteamine/cystamine (Rudolph and Lilie 1996). Also, L-arginine was included in the refolding buffers to evaluate if its presence could improve refolding and prevent the formation of non-soluble intermediates (Chen, Liu *et al.* 2008).

Two denaturation agents in a buffered solution, and four Refolding Buffers were tested, as mentioned previously in Section II.4.2.ii, to a total of eight different test conditions, summarised here briefly in the purpose of contextualisation:

(I) – 6 M Guanidine Hydrochloride,
(II) – 8M Urea;
A – 200mM Tris, 3.7mM Cystamine, 6.5mM Cysteamine, 1M L-Arginine, pH 8.0,
B – 20mM Tris, 1mM EDTA, 1mM GSH, 0.1mM GSSG, pH 8.0,
C – Buffer A + 1M TMAO,
D – Buffer B + 1M L-Arginine and 1M TMAO.

After solubilisation, protein was quantified by UV spectrophotometry in Nanodrop using the determined theoretical extinction coefficient (Table III.1). The process of inclusion bodies isolation and solubilisation was assessed by SDS-PAGE and the purification and solubilisation of inclusion bodies was considered to be successful with both GuHCl and Urea (Figure III.20).

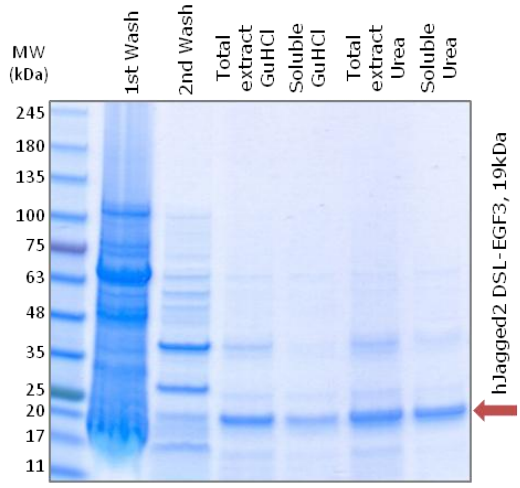


Figure III.20 - SDS-PAGE analysis of human Jagged2 DSL-EGF3 inclusion bodies isolation and purification steps.

The refolding conditions were prepared as described in Chapter II, Section II.4.2.ii using two oxido shuffling systems, reduced and oxidised glutathione GSH-GSSH (Pintar, Guarnaccia *et al.* 2009) and cystamine-cysteamine pairs. The presence of protein precipitation was clearly observed immediately after setting up and during the refolding experiment duration. All refolding solutions

appeared to be cloudy, indicative of protein precipitation occurring during the refolding process. Refolding solution B with protein solubilised with GuHCl (Shimizu, Chiba *et al.* 1999) and refolding solution C (IC and IIC) appeared to cause more protein precipitation.

After refolding, protein solutions were clarified by centrifugation, desalted using PD-10 disposable columns and protein concentration was determined by spectrophotometry (Table III.4 §). The fractions containing protein were pooled and concentrated using Amicon concentrator devices, and protein concentration was determined by Bradford Protein Assay (Table III.4 ¤). Refolding conditions (I)B and (II)C showed no protein in the quantification and were not further analysed. Remaining protein samples obtained after refolding were analysed by SDS-PAGE (Figure III.21) and all presented >90% purity.

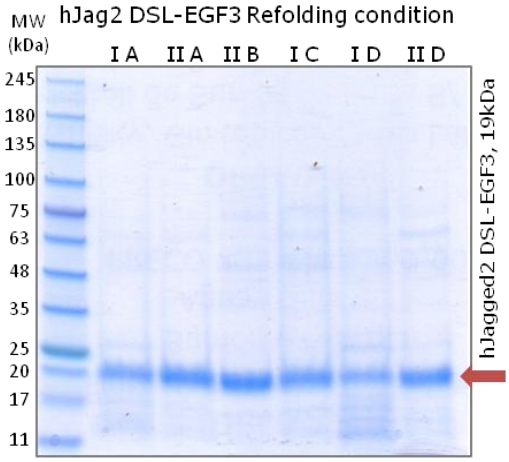


Figure III.21 - SDS-PAGE of refolded human Jagged2 DSL-EGF3 in each refolding condition. I - Denaturation of inclusion bodies in GuHCl containing buffer; II - Denaturation with Urea containing buffer; A - Refolding with cystamine-cysteamine pair and arginine; B - Refolding with GSH-GSSG; C - Refolding as in A supplemented with TMAO; D - Refolding as in B, supplemented with arginine and TMAO. 2ug of each protein sample was applied in the gel.

To assess if the refolding process was successful, Differential Scanning Fluorimetry (DSF) assays were performed. DSF allows the determination of the protein's melting temperature (Tm) by monitoring protein thermal unfolding and requires minimal amounts of protein. The thermal denaturation curves for each assay were analyzed, and the melting temperatures determined by the calculation of the first derivative, which corresponds to the midpoint temperature of the cooperative protein-unfolding transition. We employed this strategy for the reason that if the refolding process was successful, our protein of interest would be folded and would probably show a signal in the DSF.

The DSF result (Figure III.22) shows that human Jagged2 DSL-EGF3 purified from inclusion bodies, solubilised with GuHCl and refolded with Buffer D (20mM Tris, 1mM EDTA, 1mM GSH, 0.1mM GSSG, 1M Arginine, 1M TMAO, pH 8.0) was the only protein sample where thermal denaturation yielded Tms (around 50°C). However, protein refolded in Buffer A and C with protein extracted from inclusion bodies solubilised with GuHCl generated low intensity denaturation curves (Figure III.22-A in green and -B in light blue). In all denaturation curves obtained, one can notice that the first reading corresponds to the maximum fluorescence value obtained, suggesting that all the protein samples have denatured protein.

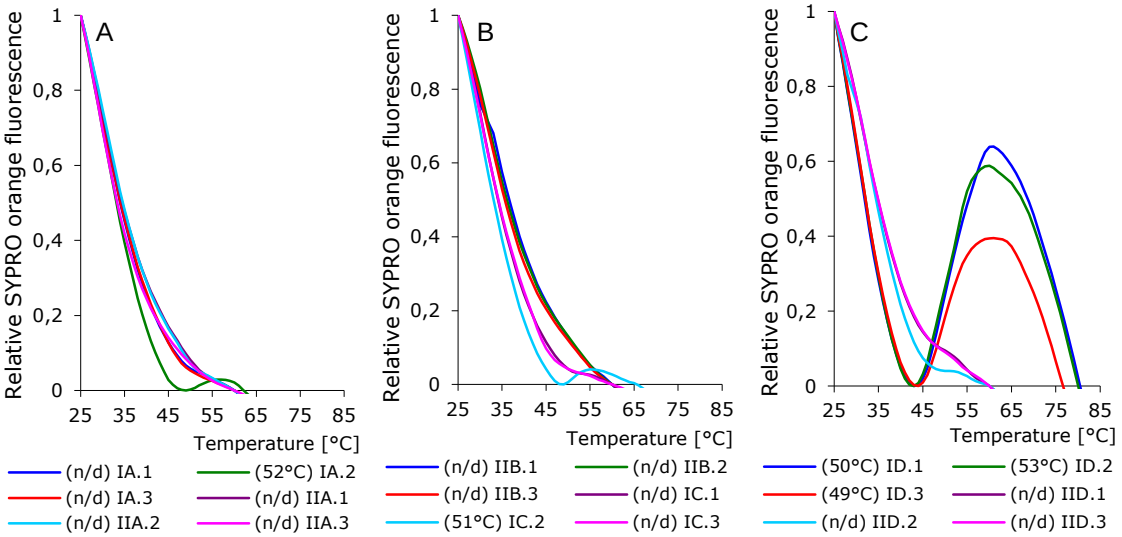


Figure III.22 - DSF analysis of human Jagged2 DSL-EGF3 refolded from Inclusion Bodies. Thermal denaturation curves, normalised such as the highest value corresponds to 1 and the lowest value to zero. Panel A - Protein solubilised with GuHCl and Urea and refolded in Refolding Buffer A; Panel B - Protein solubilised with Urea and refolded in Refolding Buffer B, Protein solubilised with GuHCl and refolded in Refolding Buffer C; Panel C - Protein solubilised with GuHCl and Urea and refolded in Refolding Buffer D (see appendix for buffer composition).

The results obtained for the small-scale refolding tests are summarised in Table III.4.

Table III.4 - Summary of refolding tests for human Jagged2 DSL-EGF3.

Solubilisation Buffer	Solubilised Protein	Refolding Buffer	Precipitation during refolding	Refolded protein (µg) §	Refolded protein (µg) ¤	DSF readout
GuHCl (I)	37.4mg; 22mg/mL	A	++	1060	824	#
		B	++++ *	n/a	n/a	n/d
		C	+++	357	88	#
		D	**++**	**107**	**101**	**Yes**
Urea (II)	66.9mg; 35.22mg/mL	A	+	448	222	No
		B	+++	229	170	No
		C	++++	362	n/a	n/d
		D	+	369	231	No

n/a - not applicable; n/d - not determined; § - initial protein quantity determination; ¤ - final protein quantity determination; ***** - all protein precipitated in the refolding experiment; # - only one triplicate with denaturation curve in the DSF experiment.

We found that when using Urea in the solubilisation buffer more protein could be extracted from the inclusion bodies (66.9mg) when compared to Guanidine Hydrochloride (37.4mg). Regarding the results obtained when using GuHCl as denaturation agent, the amount of protein obtained after refolding with Refolding solution A was the highest, rendering 1060 and 824µg. Nevertheless, human Jagged2 DSL-EGF3 refolded under this condition failed to generate denaturation curves in the DSF assay, indicating that either no folding was aquired, or that protein was not stable. Using condition (I)C produced the same outcome as (I)A in the DSF result, apart of giving much less protein. On the other hand, and altough yielding less protein at the first quantification, the output of Refolding condition (I)D was the only protein sample that generated denaturation curves, consistently, in all triplicates of the

DSF assay. This was also the only protein sample that whitstand during the buffer exchange and concentration steps, having minimal (~5.5%) losses. All protein samples obtained after refolding urea-solubilised protein failed to generate denaturation curves in the DSF assay, regardless of the fact that more protein could be solubilised in the start of the process.

As the buffers used for the refolding assays were based on the conditions for other proteins (reviewed by Rudolph and Lilie (1996) and in the protocols from Hyvonen's group) and may not be the best suited for our protein, we tested the influence of different buffers and salt concentrations on its the denaturation profile by DSF. DSF-based buffer screens have been used for identify buffer formulations that stabilize proteins (for example for structural studies (Ericsson, Hallberg *et al.* 2006, Santos, Bandeiras *et al.* 2012)). The protein sample that presented denaturation curves in DSF (ID, Figure III.22-C) was used to perform a screen where different buffer formulations with pH values between 5.0 and 10, in combination with four different salt conditions were tested in a 96-well plate format (Section VI.3, Table VI.2). The theoretical Isoelectric Point (pI) for this protein is 5.82 (calculated with ProtParam Tool online), and the overall protein solubility tends to increase with higher solvation; hence, 74 buffer-salt conditions were tested. A first reference assay with 25mM Hepes, pH 7.5 was set up, as this buffer shows to have the least pH variation upon heating, and a second reference with the protein's buffer was also used. In Figure III.23 the best curves are presented.

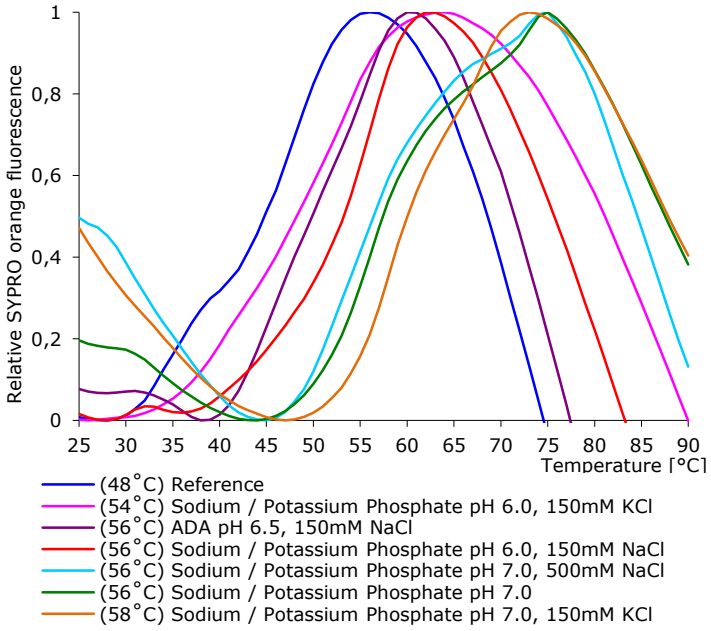


Figure III.23 - DSF Buffer Screen result for refolded hJag2DE3. The represented curves were chosen according to the parameters stated in the text. hJag2 DE3 Tm is also shown. Sodium / Potassium Phosphate pH 7.0, 150mM KCl, in orange, was chosen as the protein buffer.

There were several factors to consider while analysing the results: the initial fluorescence, the intensity of the denaturation curves, corresponding to the overall availability of protein to denaturate, the sharpness of the first derivative curve and the relative increase in Tm compared to the reference sample.

As one can observe, Sodium / Potassium Buffer is represented in five of the tested conditions showing a positive effect on the stability of human Jagged2 DSL-EGF3. Moreover, this buffer system at pH 7.0 seems to increase hJag2 DSL-EGF3 stability as, regardless of salt conditions, the curves in light blue, green and orange shown in Figure III.22 have a higher Tm shift compared to the reference curve, in blue. These buffer/salt conditions are increasing the protein's stability as observed by the Tm increase. We chose to pursue a large scale protein purification and refolding using Sodium / Potassium Buffer pH 7.0 with 150mM KCl (Tm = 58°C) in the composition of all buffers used.

III.2.4.ii Large scale refolding and purification of His6-hJAG2 DSL-EGF3

Inclusion bodies were isolated from 2L culture volume, (expressed in the best conditions determined for the highest protein expression as in Sections II.4.2.i and III.2.1.i), solubilised in Solubilisation Buffer and protein was quantified by spectrophotometry (55.99mg/mL, 587.93mg total). The purity of the preparation was estimated by SDS-PAGE to be 95% (Figure III.24-A). Protein was refolded by rapid dilution to 0.125mg/mL in Refolding Buffer for 40 hours at 4ºC. subsequently, a 50-fold concentration step was followed and protein concentration was determined by Bradford assay to be 1.1mg/mL, yielding 105.7mg total protein. Refolded protein was desalted and protein concentration was determined by NanoDrop to be 0.480mg/mL, 69.57mg total, confirming a loss of around 30% of total protein.

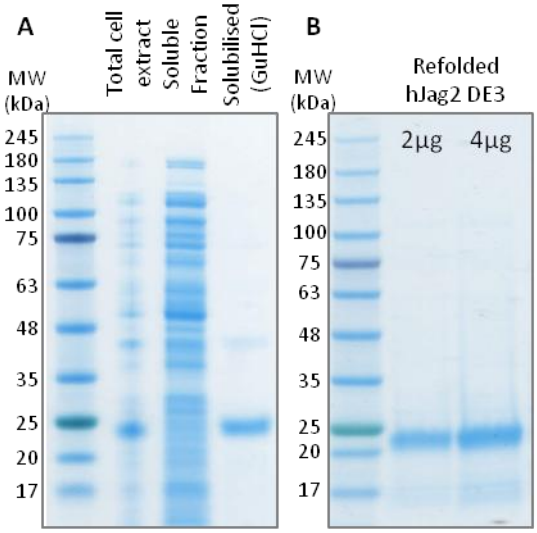


Figure III.24 - Evaluation of human Jagged2 DSL-EGF3 Large-scale purification from inclusion bodies. Panel A - Isolation and Solubilisation steps; Panel B - Refolded protein after the desalting step.

Protein stability was evaluated by DSF, as before (Figure III.25). Surprisingly, the obtained Tms in this assay are comparable to those obtained previously (49, 50 and 53°C in the small scale test and 48, 51 and 52°C in this large-scale purification). However, the initial fluorescence intensity was in this case lower, suggesting the presence of less denaturated protein in solution, and in agreement the denaturation curves are more intense indicating the presence of more properly folded protein in solution. Thus, comparing the overall DSF results, the applied modifications in buffer formulation may

have improved the refolding process. Nevertheless, the initial fluorescence was still a matter of concern and the obtained Tm values could be improved, thus a DSF Buffer and Salt screen was performed (Figure III.26).

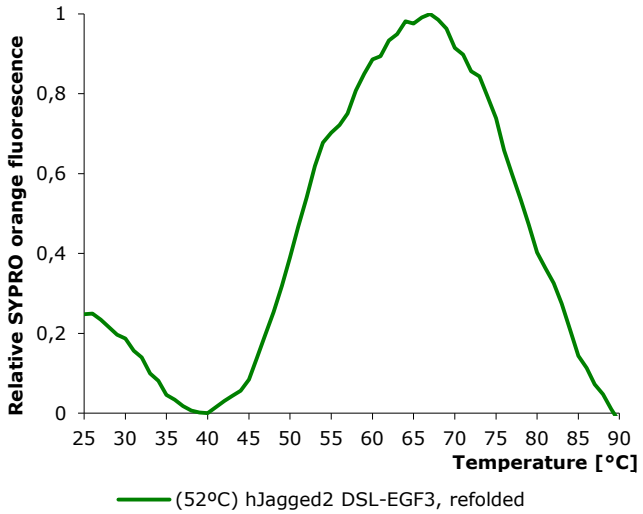


Figure III.25 - DSF evaluation of hJag DE3 protein after refolding. The thermal denaturation curve represents one triplicate of the assay.

As in the previous DSF buffer screen, the best buffer for our target protein is Sodium/Potassium Phosphate in which the protein showed the best denaturation curves and the highest Tm. Jagged2 stability was also shown to be salt dependent, with the preferred salt conditions to be 500mM NaCl, as there is a clear shift of 2°C, compared to lower salt concentrations assayed.

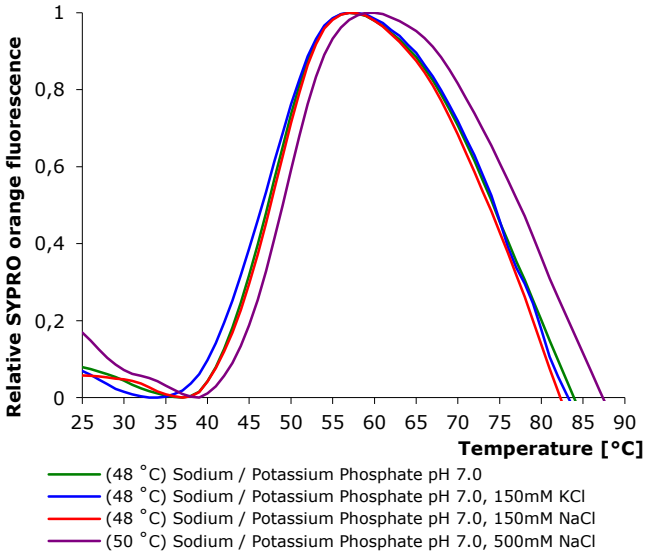


Figure III.26 - DSF buffer screen result after large scale refolding of human Jagged2 DSL-EGF3 from inclusion bodies. The best thermal denaturation curves are shown, chosen with the previously described criteria.

As such, the protein buffer was exchanged for Sodium /Potassium Phosphate, 300mM NaCl, pH 7.0. We decided to lower the salt concentration for the sake of crystallisation experiments (discussed

below). Also, envisioning the crystallisation trials, there was the need to assure that the protein sample was monodisperse.

For those reasons, a SEC was performed with the optimized buffer and therefore both requirements would be accomplished simultaneously (Fig. III.27). Two injections were performed, corresponding to 16.6mg of protein, and showing a similar elution profile. According to the calibration curve, the estimated molecular masses corresponding to the obtained elution volumes were 73.3, 38.2, 18.6 and 6.9kDa, respectively. The peak assigned to have 18.6kDa most likely corresponds to hJag2 DSL-EGF3 (19.6kDa, Table III.1). Nevertheless, fractions corresponding to all the peaks were analysed by SDS-PAGE (Figure III.28).

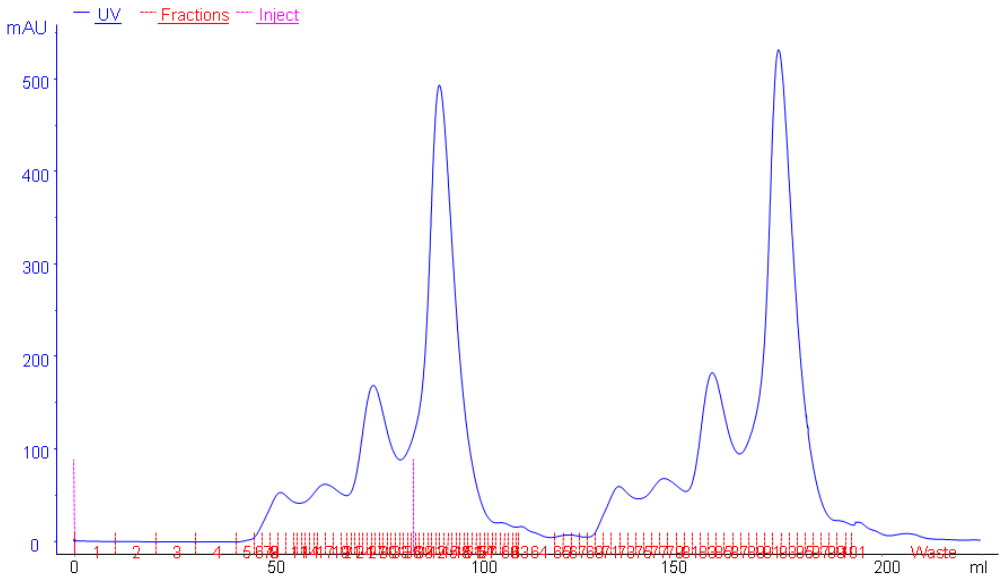


Figure III.27 - Size Exclusion Chromatogram for the purification of human Jagged2 DSL-EGF3 from inclusion bodies. Injection is marked in pink and volume was corrected to zero at that point. Protein was monitored by absorbance at 280nm and recorded (in blue, left vertical axis). The eluted fraction numeration is shown in red, in the bottom.

As one can observe in the SDS-PAGE (Figure III.28), the first and second peaks (fractions 7 to 21) correspond to high molecular weight contaminants; the third peak (27-34) corresponds indeed to our protein of interest but surprisingly, the fourth peak also corresponds to hJag2 DE3 (fractions 37 to 54). Since both peaks contain the protein with a high degree of purity but were eluted separately, two pools were prepared and treated separately thereafter - the third SEC peak was termed Pool 1 and the fourth Pool 2.

The fact that Jagged2 may adopt a pencil-like shape, as for Jagged1 structural data presented by Cordle, Johnson *et al.* (2008), can explain why both protein peaks contain our target protein with elution volumes so discrepantly different than those calculated for globular proteins. We hypothesize that Pool1 may correspond to a dimeric and Pool2 to a monomeric form of Jag2 DE3.

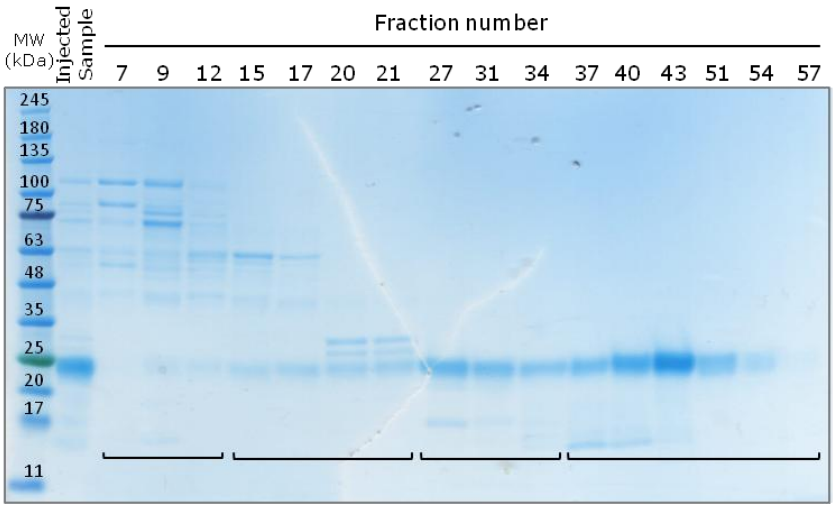


Figure III.28 - SDS-PAGE analysis of the SEC fractions for the purification of human Jagged2 DSL-EGF3 from inclusion bodies. Sampled fractions are marked with the respective numbers on top; bottom lines delimit each of the four peaks obtained in the chromatogram from Fig. III.27.

Both protein pools were concentrated and protein quantified by Bradford assay. Pool1 was concentrated to a final volume of 87µL and protein concentration was determined to be 7.5mg/mL, 0.63mg total; Pool2 was concentrated to a final volume of 733µL and protein concentration was determined to be 7.9mg/mL, 5.7mg total.

The results obtained for the of hJag2 DSL-EGF3 large scale protein refolding and purification are summarised in Table III.4.

Table III.4 – Summary of hJagged2 DSL-EGF3 Large Scale Protein Refolding and purification.

	Volume (mL)	Quantity (mg)	Concentration (mg/mL)	Loss (%)
Initial Refolding Solution	4500	560	0.125	
Final, Clarified Refolding Solution	95	105.7	1.11	81.1
Refolded Protein after Buffer exchange	145	69.57	0.48	34.2
Concentrated Protein Prior to SEC	24	56.96	2.37	18.1
Resulting Protein Pools from SEC of 16.6mg	n/a	10.14	n/a	38.9
Resulting Pool 1	0.42	2.30	5.5	77.3
Concentrated Pool 1	0.087	0.63	7.5	72.6*****
Resulting Pool 2	2.80	7.84	2.8	22.68
Concentrated Pool 2	0.733	5.7	7.9	27.3*****

***** - Relative to the obtained protein sample

In order to determine the differences between both protein pools obtained in the SEC, DSF (Figure III.29), analytical SEC (Fig. III.32) and CD analysis were employed. Circular dichroism spectra unveils protein secondary structure elements in solution. Both tested hJag2 DE3 protein pools showed no signal significantly different from the baseline as well as no thermal ramp denaturation profile,

possibly indicating that the protein samples did not contain measurable amounts of secondary structure elements. The results are therefore not presented.

As one can observe in Figure III.29-A, pools 1 and 2 have significantly different melting temperatures, 56 and 67°C on average, respectively. Also, the curve intensities are incredibly disparate (Figure III.29-B). These facts corroborate that there must exist differences between the two protein pools in terms of structural elements and percentage of folded protein that can account for both the different melting temperatures and curve intensities.

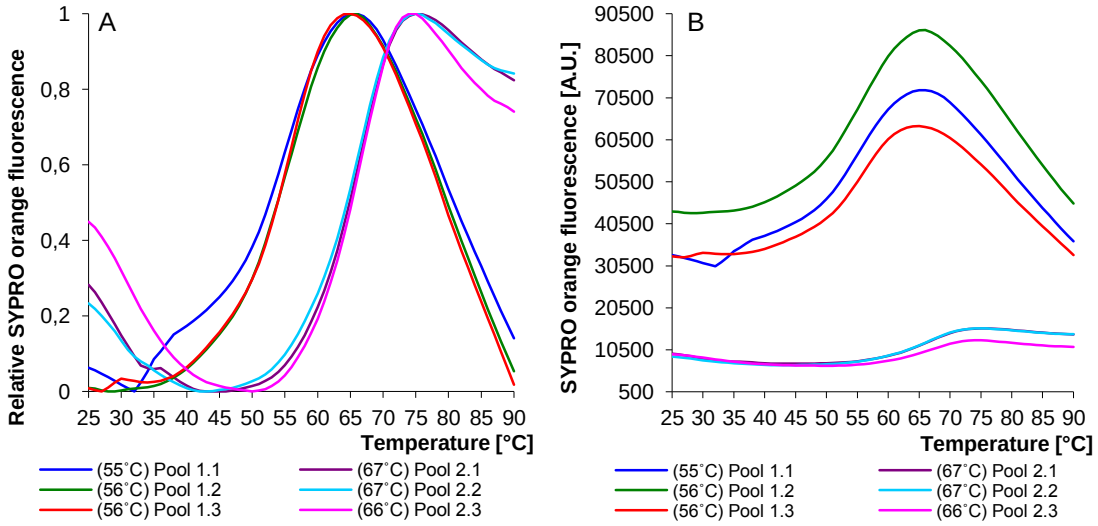


Figure III.29 - DSF analysis of pure human Jagged2 DSL-EGF3 protein after SEC from the inclusion bodies purification. Panel A - Normalised curves, Panel B - Non-normalised curves. Curve colour code is the same for both panels.

In order to further characterise the protein pools and understand the differences observed in the DSF, an SDS-PAGE with both reduced and non-reduced samples was performed. For a non-reducing SDS-PAGE, the sample is prepared without reducing agent and boiling, in a SDS rich environment, allowing the peptide chains to be charged negatively and therefore separated based on the molecular mass, while maintaining the disulfide bonds. As one can observe in the SDS-PAGE (Figure III.30), the non-reduced protein samples show an apparently lower molecular weight in the non-reducing conditions than in the reducing conditions. This can be due to the more compact conformation imposed by the disulfide bonds in the non-reduced samples, compared to the denatured peptide chain in the reduced samples (Figure III.30, NR and R, respectively). The protein from the inclusion body solubilisation shows one single band whereas the protein sample injected in the SEC column, after the refolding, has two. This can possibly indicate that we are in the presence of mixed protein populations after refolding (possibly refolded with different combinations of disulfide pairs), or this double band can represent protein degradation products. Respecting the obtained protein pools after SEC, it is very clear that Pool 2 has two distinct bands, clearly indicative of, at least, two distinct protein populations, possibly with different arrangements in the disulfide bonds. Pool1 does not show the apparently higher molecular weight band as Pool2, but seems to have two bands as well.

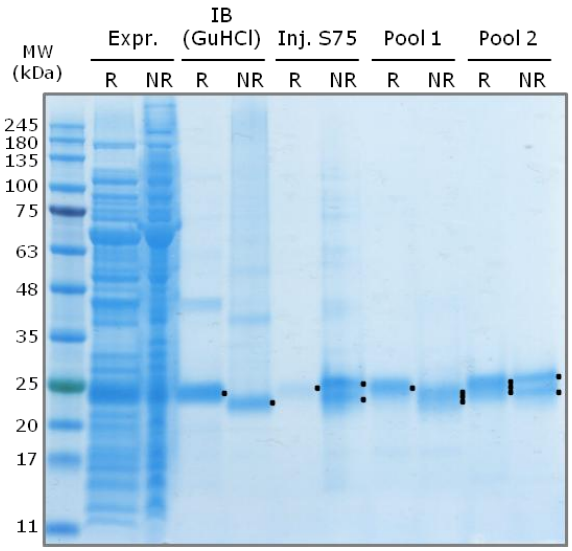


Figure III.30 - SDS-PAGE analysis of the hJag2 DE3 refolding and purification process. Bacterial cell expression (Expr.), Inclusion body purification (IB (GuHCl)) and SEC – Injected (Inj. S75) and resulting protein samples (Pool 1 and Pool 2) were analysed in reduced (R) and non-reduced (NR) environment.

Similarly to human Jagged1, Jagged2 has twelve predicted disulfide bonds in DSL to EGF3 domains (Figure I.3), which seem to be important in maintaining its folding and structure. Reducing agents disrupt intermolecular disulfide bonds, that are usually solvent exposed, and it is thought that the intramolecular disulfides remain unchanged. However, the structure of human Jagged1 (Figure I.3) enlightens that the intramolecular disulfides are, in these proteins, solvent exposed. Therefore, if protein refolding had failed or protein folded incorrectly, giving rise to an abnormal protein, there could be a completely random arrangement of disulfide bonds and possibly no secondary structure elements would be present in the resulting protein. In a DSF assay, we would expect total protein denaturation upon reducing agent exposition, with curves similar to those presented in Figure III.22-A and -B, high initial fluorescence that decays in function of temperature, indicative of protein:dye complexes precipitation at the starting point of the assay. On the contrary, if correct protein refolding had been successful, even if the pattern of disulfides is not C1-C3, C2-C4, C5-C6 (Knott, Downing *et al.* 1996), some structural elements would be present and a shift to lower Tm upon reduction is expected.

In order to assess the influence of reducing agents in human Jagged2 DSL-EGF3 refolded from inclusion bodies, DSF was employed to monitor, if any, changes in protein Tm. The results from this assay are shown in Figure III.31. For both protein pools assayed, a thermal denaturation curve is observed even after incubation with reducing agents. However, we observe a shift to a lower Tm in the presence of both tested reducing agents (Figure III.31, purple and red curves) compared to the protein solutions without reducing agents (Figure III.31, blue curve). We therefore conclude that both protein samples present secondary structure elements and the predicted disulfide bonds. We cannot exclude however, the possibility of having a protein molten globule (Leal and Gomes 2007).

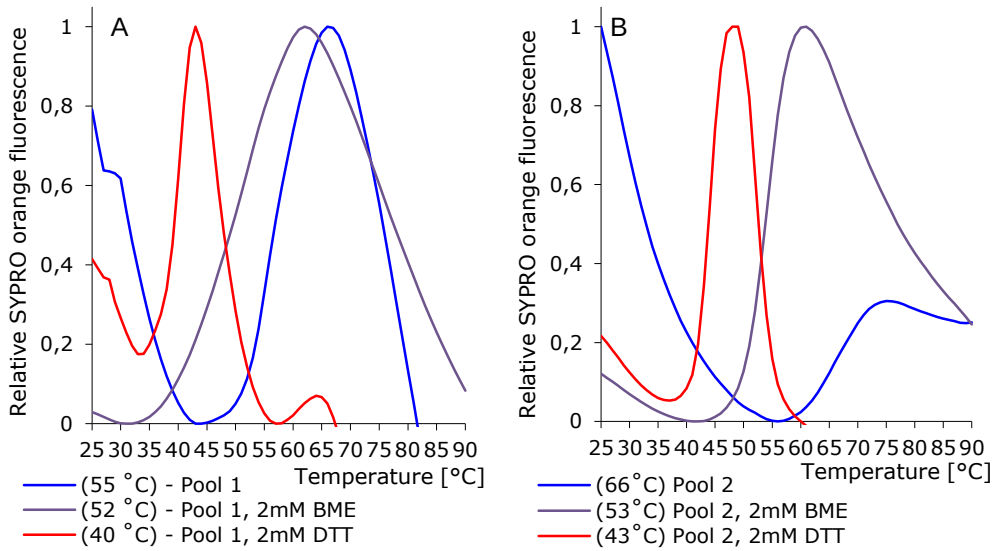


Figure III.31 - Influence of reducing agents in human refolded Jagged2 DSL-EGF3 protein pools assessed by DSF. Panel A - Pool 1 from SEC, Panel B - Pool 2 from SEC. Colour code of reducing agents is the same in both panels

III.2.4.iii Analytical size exclusion chromatography

To verify if the human Jagged2 DSL-EGF3 protein pools prepared after the SEC (as shown in Fig. III.27) were stable and homogenous, a protein aliquot of each protein pool was applied on an analytical SEC column. Analytical columns such as Superose 6 PC3.2/30 are designed to archive a highly resolving separation of biomolecules for the high theoretical plates combined with low flow rates.

As one can observe in Fig. III.32, the elution volume of both protein samples is different, maintaining the relative differences observed in the initial preparative SEC (Fig. III.27), pool 1 having a higher elution volume than pool 2. This result confirms that each protein sample is stable and does not re-equilibrate to other species in solution. Moreover, it confirms that each protein sample contains one and only one species, i.e., that the separation of fractions to prepare each pool did not contribute for sample heterogeneity.

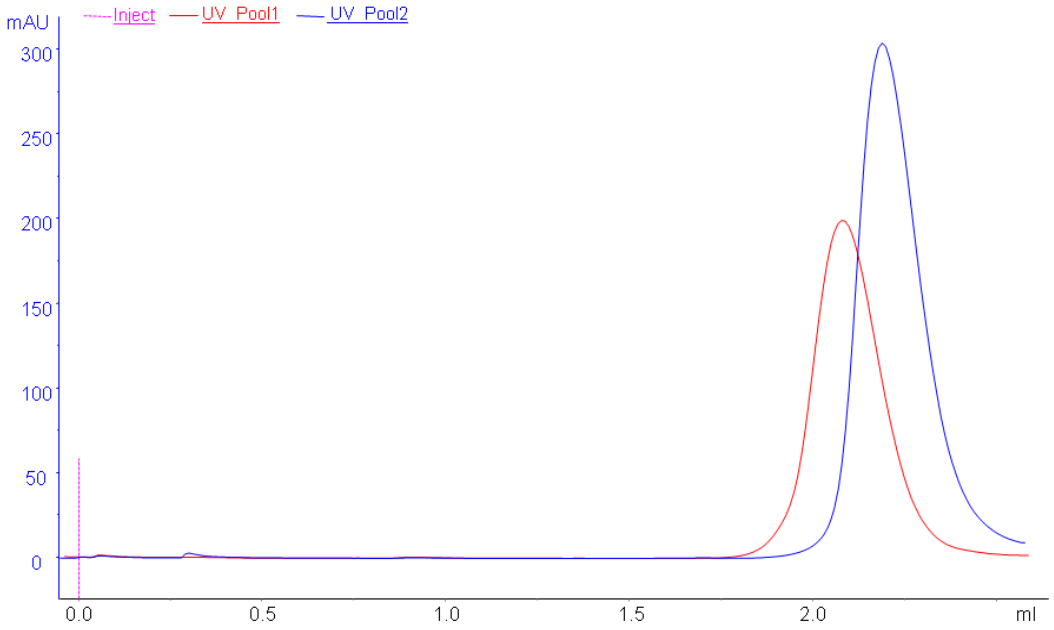


Figure III.32 - Analytical Size Exclusion Chromatograms for human Jagged2 DSL-EGF3 purified from inclusion bodies. Injection is marked in pink and volume was corrected to zero at that point. Protein was monitored by absorbance at 280nm and recorded (left vertical axis) in red for Pool1 and in blue for Pool2. Each Protein sample was run independently and the chromatograms superposed in UNICORN software.

We conclude that these constructs are not amenable for bacterial expression nor for soluble protein purification, as only around 20% soluble protein, as estimated by Western Blot analysis, could be obtained, and the target proteins could not be purified from the bacterial contaminants after tag cleavage. We conclude also that, even when purifying human Jagged2 DSL-EGF3 refolded from inclusion bodies with high yields, as reported previously for a similar construct of hJag1 (Cordle, Johnson *et al.* 2008), we could not obtain a homogenous sample in terms of disulfide pattern, which can compromise significantly the crystallization attempts (Section III.5).

III.3. Mammalian Expression System

In order to assess the expression of each protein construct in mammalian cell expression system, the Notch ligands domains previously cloned in pET47b were also successfully cloned in pHL-SEC vector (Aricescu, Lu *et al.* 2006). As part of the European project P-CUBE, the constructs were sent to OPPF and the small scale expression screenings performed as follows.

III.3.1 Small-scale expression screenings and Protein purification at OPPF

All plasmids were transiently transfected in adherent HEK293T cells encoding each protein, preceded with a secretion signal, and with a C-terminal His-tag. Supernatant samples harvested 48h post-transfection were analysed by WB using an anti His-tag antibody (Figure III.33). Three constructs did not show expression of the respective proteins (hJag1 MNNL-EGF3, hJag2 MNNL-EGF9 and hJag2 DSL-EGF9). This may be due to a premature stop before the translation of the C-terminal His-tag. The former construct sequencing result confirmed that the DNA sequence was as expected but the later constructs' result failed to confirm the presence of the correct sequence at the 3' end. This can explain that the lack of result for hJag2 MNNL-EGF9 and hJag2 DSL-EGF9 protein expression might be due to a DNA level error. All the other tested constructs gave a positive result with intense bands corresponding to the expected size of each protein except for hJag1 MNNL-EGF9 with one higher and hDll1 DSL-EGF3 with one lower molecular weight contaminant protein.

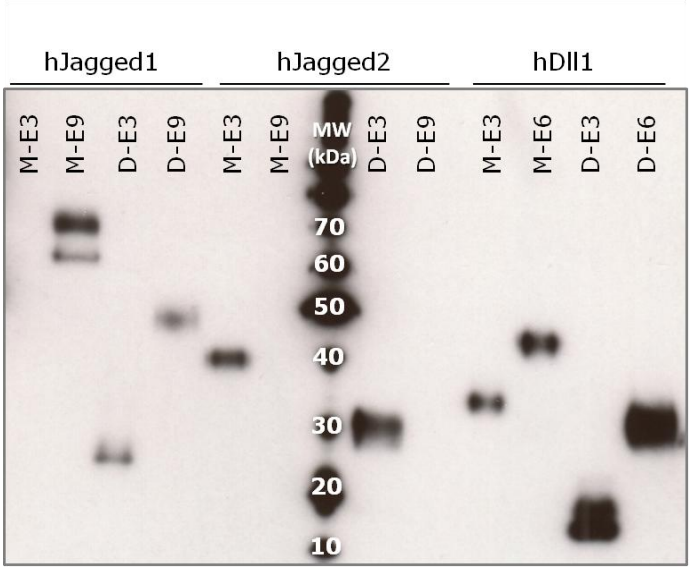


Figure III.33 – Western Blot analysis of small scale expression screenings in adherent HEK293T cells with anti His-tag antibody. Exposure time: 10 seconds. M – MNNL domain; E – EGF domain; D – DSL domain.

The construct encoding hDll1 DSL-EGF6 showed the more intense band of all the tested constructs. As such, a large-scale protein production and purification was pursued using the standard conditions used in OPPF.

III.3.2 Purification of human Delta-like Ligand1 DSL-EGF6 at OPPF

Six litres of media were harvested from 24 expanded surface roller bottles, clarified, diafiltered and applied into cobalt beads for metal affinity purification. Eluted protein was concentrated and two injections were performed in a S200 SEC column. The resulting chromatogram is shown in Figure III.34.

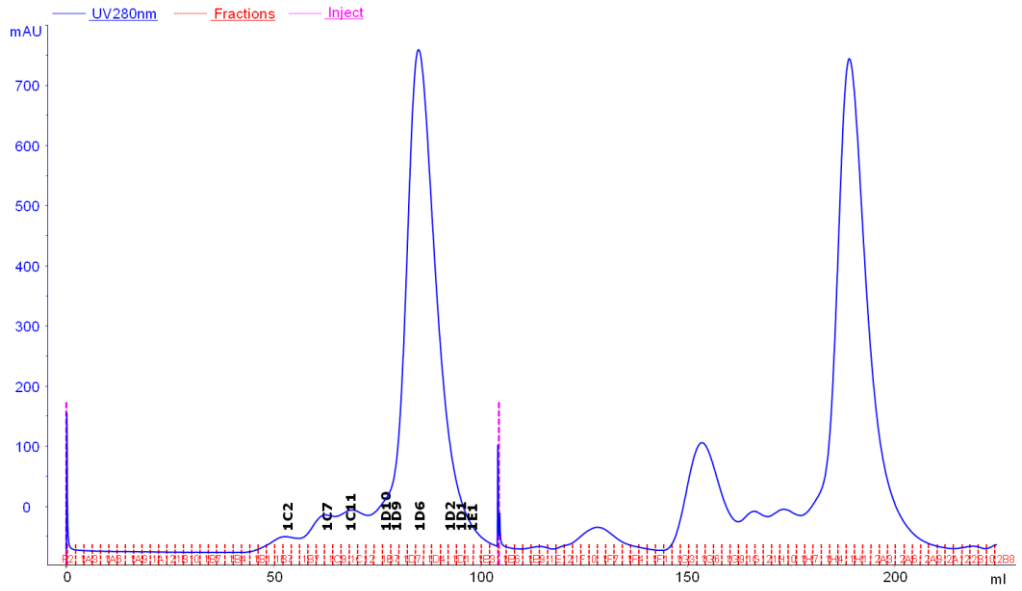


Figure III.34 – Size exclusion chromatogram for the purification of human Dll1 DSL-EGF6. Injections are marked in pink. Protein was monitored by absorbance at 280nm and recorded (in blue, left vertical axis). The eluted fractions are shown in red in the bottom and those analysed in the SDS-PAGE in Figure III.35 in black.

Both injections showed the same profile, with four peaks having the same elution volume. The first peak was more intense in the second injection, which may indicate protein aggregation with time. The peaks resulting from the first infection were analysed by SDS-PAGE in as well as the previous step's samples (Figure III.35).

The samples from the batch metal affinity purification reveal that some protein was lost during the washing steps, but the amount was negligible when comparing with the amount of contaminants removed from the chromatography media (Figure III.35). Nevertheless, some degree of protein contamination was present, in the sample injected in the SEC column. Regarding the samples from the SEC column (Figure III.34), as expected, the first three peaks correspond to higher molecular weight contaminant proteins as one can observe in the lanes corresponding to Fractions 1C1, 1C7 and 1C11. The fourth and most intense peak corresponds to human Dll1 DSL-EGF6 with a molecular weight of around 32kDa, as in Table III.2.

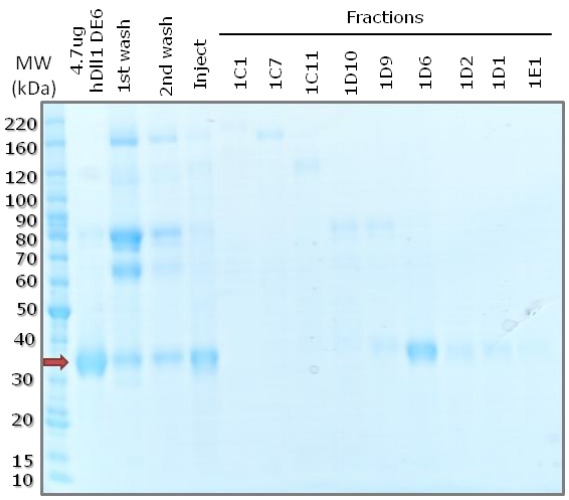


Figure III.35 – SDS-PAGE analysis of human Delta-like1 DSL-EGF6 purification. Purification using cobalt beads (in batch) and SEC fractions were analysed.

The fractions corresponding to each peak were pooled and concentrated separately. The first peak was pooled and readily concentrated as the second injection was still running. As some precipitation occurred during the first chromatogram peak concentration, 300mM NaCl was added to pool 1 in order to stabilize the protein due to the increase in ionic strength. As this did not appear to solve the precipitation issues, we further added 5% glycerol as a stabilizer. As such, pool1 buffer composition is 10mM Hepes, 300mM NaCl, 5% Glycerol, pH 7.5; the protein pool1 was concentrated to 4.7mg/mL and resulted in 2.35mg of protein.

With the previous knowledge taken from Pool 1 concentration, and envisioning the crystallization trials (Section III.5 - Protein Crystallisation), we added 650mM NaCl, omitting glycerol addition, prior to the concentration steps and this prevented precipitation events. Pool2 was concentrated to 4.6mg/mL, 1.27mg of protein in 10mM Hepes, 650mM NaCl, pH 7.5. Both protein pools were used to set up crystallization trials at OPPF/HTX and the remaining protein samples were shipped to IBET.

III.3.3 Biophysical characterisation of human Dll1 DSL-EGF6 at IBET

To evaluate the purified Dll1 DSL-EGF6 protein expressed in mammalian cells, DSF assays were employed. EGF4 and EGF7 are two putative calcium binding EGFs present in hDll1 protein as predicted by UniProtKB (http://www.uniprot.org, accession O00548). As such, it would be expected that the produced protein, encompassing EGF4, could be influenced by the presence or absence of calcium ions. As the protein could have calcium ions bound or they could have been lost during the purification, each protein pool was assayed using the respective buffer and with equimolar EDTA, a chelating agent, or two molar fold $CaCl_2$ incubation.

The Dll1 thermal denaturation curves (Figure III.36-A, solid blue and green lines) exhibit high initial fluorescence (HIF), and low intensity denaturation curves with essentially the same melting temperatures, 65-66°C. This may suggest that part of the purified protein may be in an unfolded but soluble state, and only a small portion is able to denaturate i.e. folded. This result suggests that the chosen purification buffers are not optimal for this protein, and a buffer screen could elucidate a better

option for the next purification. The HIF could also be protein dependent in the case that this protein has more exposed hydrophobic areas.

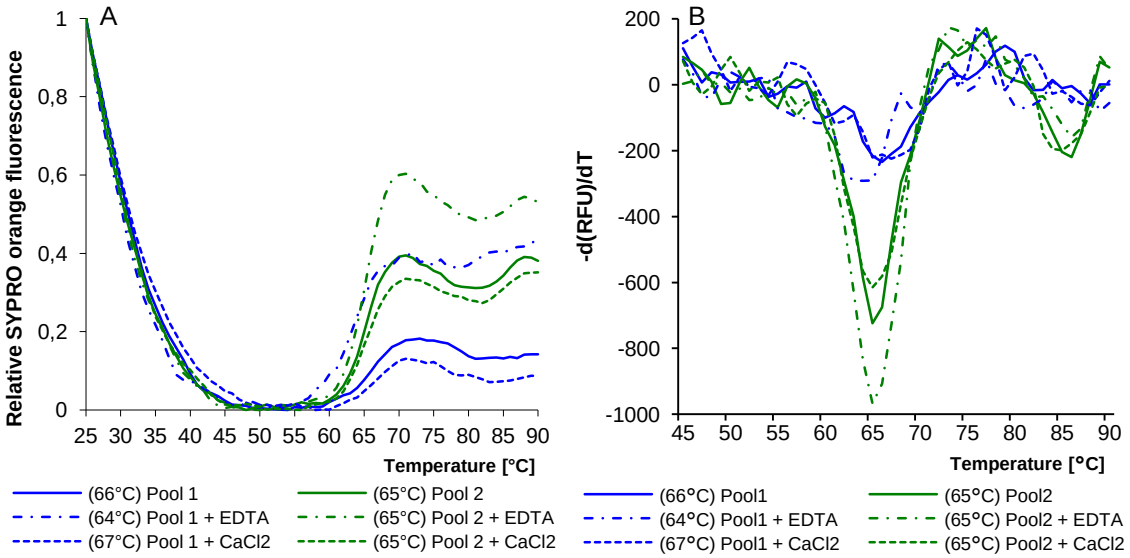


Figure III.36 – DSF assays for human Dll1 DSL-EGF6 produced in mammalian cells and purified at OPPF. Panel A – Representative denaturation curves for each protein pool and the results obtained upon incubation with equimolar EDTA or two molar fold $CaCl_2$; Panel B – First derivative of the data presented in Panel A. Curves for Pool1 in blue and for Pool2 in green.

Regarding the assays with $CaCl_2$ and EDTA, both protein pools appear to behave in a similar manner: calcium incubation causing a slight decrease in the denaturation curves intensity, and EDTA increasing it. Our results regarding the influence of EDTA in protein stability do not show significant differences in protein Tm (-1°C), possibly indicating that the protein does not have calcium ions bound, although addition of $CaCl_2$ to the DSF assay does show an improvement of 1°C in protein stability. This can possibly indicate that either the incubation time was not sufficient for calcium ion incorporation or that the putative calcium sites are not actually present.

From the DSF analysis of the first derivative (Figure III.36-B), pool 2 shows sharper curves than pool 1, maybe indicating that the transition between the native and unfolded protein species is more cooperative. Comparing the sharpness of the first derivative between pool 1 and pool 2 we can conclude that hDll1 DSL-EGF6 is more stable in high salt concentrations. This result in the DSF assay may also indicate that pool 2 sample can be more prone to crystallize, due to its higher homogeneity (Santos, Bandeiras *et al.* 2012).

In order to screen for a buffer formulation in which human Dll1 DSL-EGF6 is more stable, a DSF buffer screen was performed (Figure III.37). Two sets of denaturation curves were chosen based on the denaturation profile, showing one (panel A) or two melting transitions (panel B).

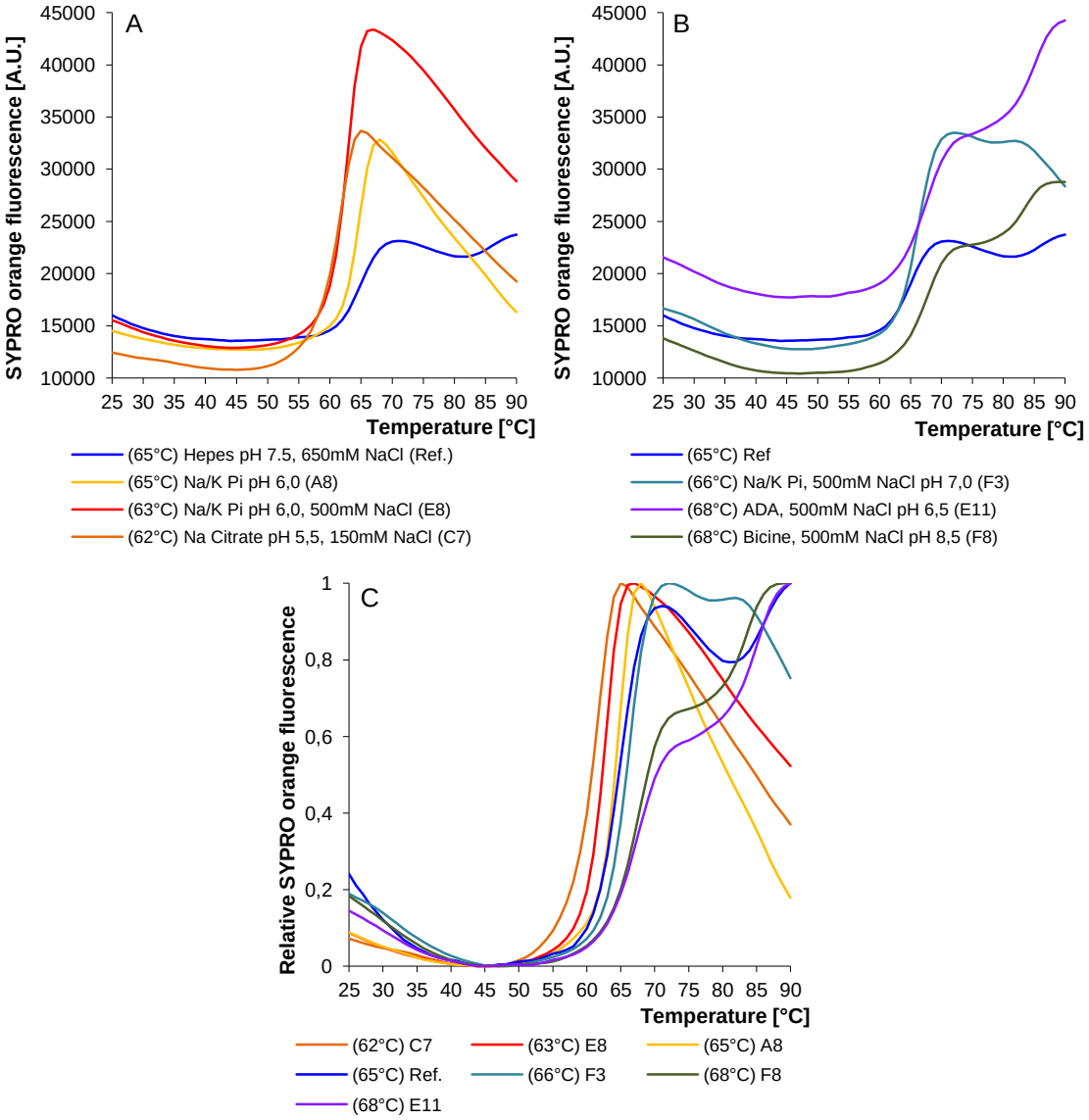


Figure III.37 - DSF buffer screen for human Delta-Like Ligand1 DSL-EGF6. Panel A – Non-normalized denaturation curves with one transition compared to the Reference; Panel B – Non-normalized two-transition denaturation curves; Panel C - Normalized denaturation curves of the selected condition presented in Panels A and B. Colour code for each Buffer system is the same across panels.

Overall, the thermal denaturation curves with a single melting transition show lower Tms. Despite the lower Tm, those buffers present a single cooperative protein denaturation profile which may suggest the presence of a more homogenous sample (Santos, Bandeiras *et al.* 2012). All buffers showing two melting transitions display a higher Tm than the reference sample. Based on the fact that Na/K phosphate pH 6.0, 500mM NaCl with 63°C and ADA pH 6.5, 500mM NaCl with 68°C display the lowest initial fluorescence, sharpest melting transition and highest Tm, we believe these two buffers formulations can be good options.

For protein crystallization purposes the best buffer option is one that provides an homogeneous sample and as such, Sodium / Potassium phosphate pH 6.0, 500mM NaCl is chosen for the next purification batch of hDll1 DSL-EGF6.

III.4. Mammalian expression at iBET

III.4.1 Expression of Notch ligands constructs in pHL-sec vector

In order to reproduce the small scale expression results obtained at OPPF, to validate the expression protocol at our lab and to generate more hDelta-Like Ligand1 DSL-EGF6 protein, all four human Delta-Like Ligand1 constructs were transfected into HEK293T adherent cells and conditioned media harvested to western blot analysis (Figure III.38).

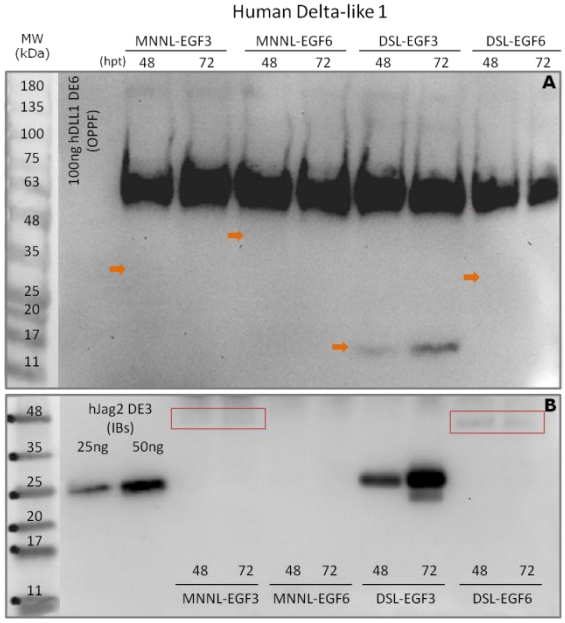


Figure III.38 - Western Blot analysis of hDll1 constructs in pHL-sec transfected into HEK293T adherent cells. The protein domains and time of media harvest (in hours) are depicted above (panel A) or below (panel B) the corresponding lanes. A - Initial result obtained with the protocol from OPPF (see Figure III.33); B - Result obtained after loading ten times the sample volume. Previously purified proteins were also loaded as controls. Faint bands are highlighted by the red squares.

Our results showed great differences regarding the expression profiles when compared to those obtained at OPPF using the same plasmid constructs and the same cells as expression host. Moreover we detected the presence of an unspecific band, assigned to be a contribution from the serum proteins. Nevertheless, we could obtain a positive signal for the presence of hDll1 DE3 protein (Figure III.38-A) and, in order to better evaluate the expression profiles of the remaining constructs, a new western blot analysis was performed. This assay was pursued with a ten-fold increase in the amount of each sample, and the membrane was clipped below the 63kDa protein in the molecular weight marker to prevent the unspecific band to react with the antibodies (as in panel A). Also, for size comparison and a rough estimate of the amount of protein, proteins purified previously were also loaded in the gels, as indicated. We could observe faint bands indicating low amounts of hDll1 ME3 and DE6 proteins with this new analysis (Figure III.38-B). Comparing the results obtained using OPPF protocol to those obtained at OPPF (Figure III.33) we could not obtain similar expression results. As such, the expression profiles of these constructs were also assessed in suspension adapted HEK293T cells (Figure III.39).

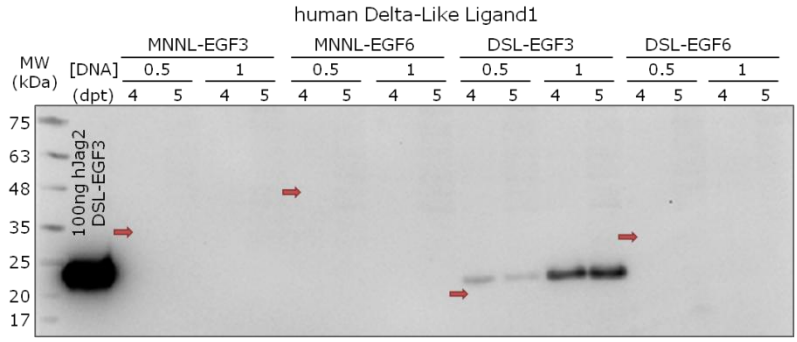


Figure III.39 – Western blot analysis of human Dll1 constructs' expression upon DNA transfection into suspension adapted HEK293T cells. Exposure time is 120 seconds. The expected molecular weight for each protein is marked with red arrows.

Human Dll1 DSL-EGF3 was, once again, the only construct for which protein expression was detected in the Western Blot analysis. As hJag2 DE3 protein was used as both an antibody control and a quantification estimative, we could observe that the expression levels were still very low. Moreover, the proteins' apparent molecular weight seems to be slightly higher than expected, as indicated by the arrows across figures hereafter. The results obtained thus far did not reveal to be promising using HEK293T cells.

HEK293T and HEK293-EBNA cells were engineered to express SV40 large-T antigen or Epstein-Barr virus (EBV) nuclear antigen-1 protein, allowing the episomal replication of plasmid DNA with SV40 or EBV origin of replication sequences, respectively ((Pear, Nolan *et al.* 1993) and Invitrogen's manual Catalog no. R620-07, Version C/102810/25-0347 available online at http://tools.lifetechnologies.com/content/sfs/manuals/293ebna_man.pdf as for September 2014). Taking this knowledge in consideration, we pursued expression trials with HEK293-EBNA cells (Figure III.40) in order to increase protein production yields as reported for Antibodies and Ab-fusion proteins (Jager, Bussow *et al.* 2013). As far as we know pHL-sec plasmid originated from a series of modifications of pCAβ-EGFP plasmid, where SV40 origin was removed (Yaneza, Gilthorpe *et al.* 2002). pHL-sec thus can possibly be maintained in both HEK293 cell types. We therefore sought to investigate whether in HEK293-EBNA cells the expression of our constructs would lead to secretion of correct molecular weight proteins.

The human Dll1 constructs in the pHL-sec vector backbone were transfected in suspension adapted HEK293-EBNA cells as in Section II.5.2.iii. The result was evaluated by Western Blot (Figure III.40).

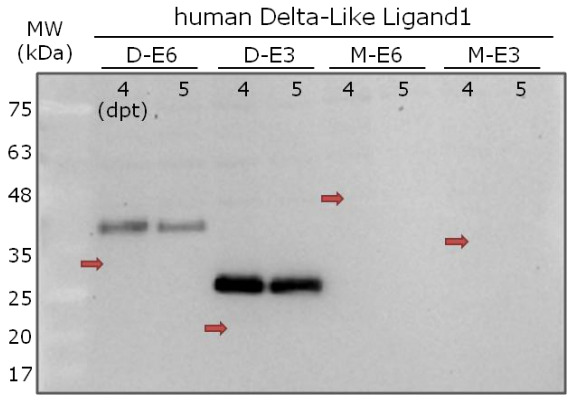


Figure III.40 – Western blot analysis of human Dll1 constructs' expression upon DNA transfection into suspension adapted HEK293-EBNA cells. Exposure time is 300 seconds. The expected molecular weight for each protein is marked with red arrows.

The hDll1 DE3 protein shows the highest expression, consistent with all the previous western blot analysis. The construct encompassing DSL to EGF6 domains (D-E6), that gave the best results in the OPPF small scale screenings (see Figure III.33), generated only modest protein expression when compared to the former construct.

Aiming to increase cell viability and expression rate, a second small scale expression trial was performed as before with an additional step of 1% FBS feeding 24 hours post transfection (Figure III.41). This strategy was followed since serum feeding post-transfection has been shown to improve cell viability, contributing for higher viable cell densities and consequently higher productivities (Pham, Perret *et al.* 2005).

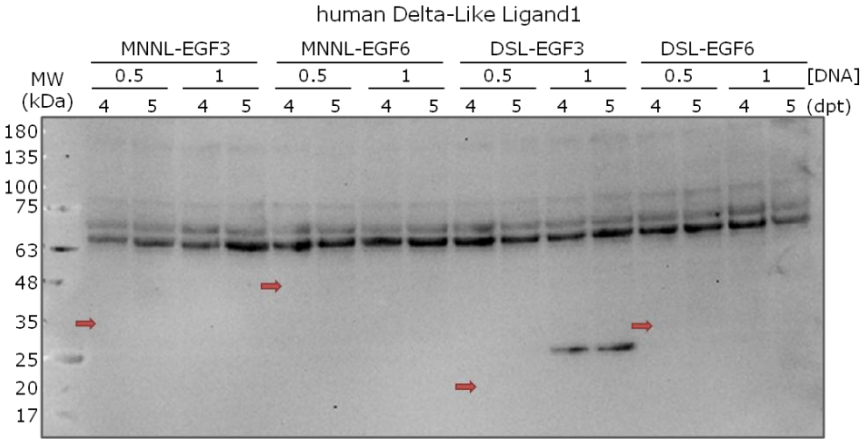


Figure III.41 – Western blot analysis of human Dll1 constructs' expression upon DNA transfection into suspension adapted HEK293-Ebna cells feeded with 1% FBS at 24hpt. Exposure time is 300 seconds. The expected molecular weight for each protein is marked with red arrows.

The results obtained with post transfection feeding were still not substantial, with hDll1 DE3 construct having the only detectable expression. Once again, the apparent molecular weight of each protein ranged far from expected. It is possible that the signal peptide encoded by the vector is not fully processed in the HEK293-EBNA cells, as it seems to be in HEK293T cells used at OPPF, causing the molecular weight shift. It is also possible that the unprocessed signal peptide is precluding the correct export across the plasma membrane, reducing the levels of detectable protein in culture

supernatants. Surprisingly as well, the protein expressions obtained are not comparable to those obtained without serum feeding (Figure III.40). This can perhaps be due to the unspecific antibody binding to FBS in the Western Blot procedure, precluding the detection to the heterologous expressed proteins in lower levels. Also, the images acquired with higher exposure time (not presented) do not show protein expression different than the presented figure, possibly due to limitations of the detection device.

Striving to generate more protein using mammalian cells as expression host, we pursued the use of a different vector backbone for transfection.

III.4.2 Expression of Notch ligands constructs in pTT22SSP4 vector

The cDNAs encoding human Jagged1 ME9, Jagged2 ME9 and Dll1 ME6 and DE6 domains were cloned in pTT22SSP4 vector, which has been successfully used for TGE in HEK293-EBNA cells (Durocher, Perret *et al.* 2002). We followed with these protein domain constructs as they correspond to the largest fragments of each protein in this study plus the construct that was already purified at OPPF.

The expression profiles were assessed in HEK293-EBNA suspension adapted cells transfected with each plasmid in both secreted and intracellular forms (Figure III.42).

Each of the tested proteins shows an intense intracellular expression (Figure III.42, right half). All of the tested pTT22SSP4 constructs give rise to proteins with molecular weights similar to the predicted ones (Table III.2).

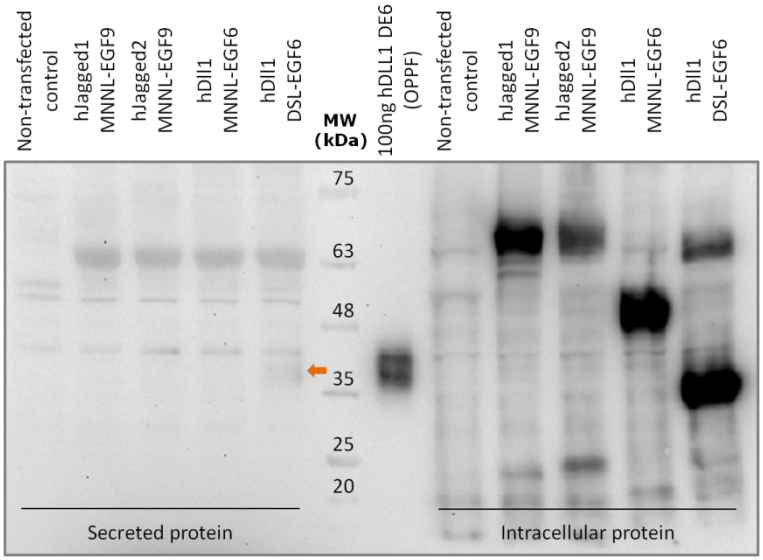


Figure III.42 - Western blot analysis of Notch ligand constructs in pTT22SSP4. Normalised amounts of condition media and cell lysate were evaluated for expression with anti-His antibody. Exposition time: 300 seconds.

We found that hDll1 DSL-EGF6, the protein produced at OPPF, gives rise to the best result obtained in secreted form in this particular experiment, being the only slightly detected protein in secreted form, as indicated by the orange arrow (Figure III.42, left half). Since some assessments and characterisation were already performed in hDll1 DE6 protein from OPPF, it would be of interest to generate higher amounts of this protein that could be purifiable at our lab.

Several additives are being studied to improve recombinant protein production yields in mammalian cells such as DNA methyl transferase (iDNMTs) and histone deacetylase (iHDACs) inhibitors and protein hydrolysates. The former act by inhibiting plasmid DNA methylation or association with deacetylated histones, increasing the plasmid's half-life in the transfected cells' nucleus, promoting transcription of extra-chromosomal DNA prior to physical loss of the plasmid (Backliwal, Hildinger *et al.* 2008). On the other hand protein hydrolysates, or peptones, counteract the nutrient depletion in the high density cultures, may improve gene expression through regulation of transcriptional and translational machinery and provide an additional source of aminoacids after transfection (Pham, Perret *et al.* 2005).

We carried out a small scale expression screening for the construct encompassing hDelta-Like Ligand1 DSL-EGF6 domains using different feeding strategies. The cultures were transfected and, 24 hours later, fed with Valproic Acid (VPA) and Pancreatic digest of Casein (Tryptone) as these additives showed the best results in those studies mentioned above (Figure III.43).

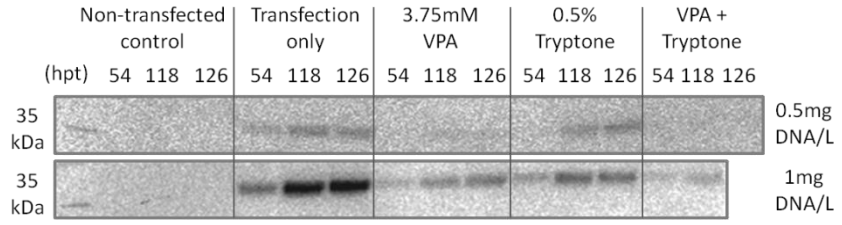


Figure III.43 - Western blot analysis of hDll1 DE6 expression after feeding with reported enhancers of protein expression in mammalian cells. Both 0.5 (top) and 1mg (bottom panel) of DNA per culture litre were tested for protein expression in the absence and presence of VPA and/or Tryptone feeding post transfection. Exposition time: 300 seconds.

As expected, the non-transfected control showed no band and the transfected cultures showed increasing amounts of hDll1 DE6 with time for both DNA concentrations tested. The use of 1mg of DNA per liter condition (Figure III.43, bottom panel) rendered a higher protein expression than 0.5mg/L (top panel). Regarding our results with both feeding additives for protein expression enhancement, we saw no increase in protein expression compared to the transfected only cultures, rather a decrease of it. Feeding with Tryptone proved more efficacious than VPA either in cultures transfected with 0.5 or 1mg/L, although the levels of secreted recombinant protein did not supplant those obtained without feeding additives.

It may be that the concentration-dependent VPA action in stabilizing transgene's mRNA levels is not followed by a stable increase in transcriptional activity as reported by Wulhfard (2009). In concentrations of 3.75mM VPA and above, cells' viability is severely compromised within seven days post-transfection (Wulhfard 2009), although this concentration was reported as having the highest

positive effect in expression enhancement (Backliwal, Hildinger *et al.* 2008). Moreover, hystone deacetylase inhibitors have been shown to decrease cell viability and to have a profound anti growth activity *in vivo* (Cinatl, Cinatl *et al.* 1997, Takai, Desmond *et al.* 2004). For its ability to modulate expression profile of transgenes, one can hypothesise that VPA modulation is not restricted to exogenous transcripts but may also interfere with the cells' physiological pathways. In fact, Greenblatt, Vaccaro *et al.* (2007) reported that VPA activates Notch-1 signalling, increasing its expression at the cells' surface. Therefore, it is possible to conceive that if the pathway is activated by VPA in the expressing cells, more interactions between the endogenous over-expressed receptor and our heterologously expressed ligand (that contains the minimal Notch binding domain) could occur, sequestering the recombinant protein and diminishing its availability.

Pham, Perret *et al.* (2005) reported a series of peptone feedings in HEK293E cells and found that some (namely CPN3, C1, GPN2 and GPN3) had little or negative effects in augmenting recombinant SEAP protein expression. Regarding the results obtained with Tryptone, our results can possibly reflect one of such cases, in the event of a system dependence of peptone feedings. We did not carry a comprehensive analysis of several peptones and only pursued with the protein hydrolysate reported as a better enhancer. Perhaps the effects of protein hydrolysate feedings can be system-dependent and what is true for a particular system may not be of suitable use for another.

For this particular construct of human Dll1 DSL-EGF6 in pTT vector, we therefore determined that a medium scale production of 1.5L culture, performed with 1mg/L DNA tranfection with PEI, harvested at 118 to 126hpt would be performed. The purification scheme was as HisTrap column affinity purification (Figure III.44) followed by SEC (Figure III.46).

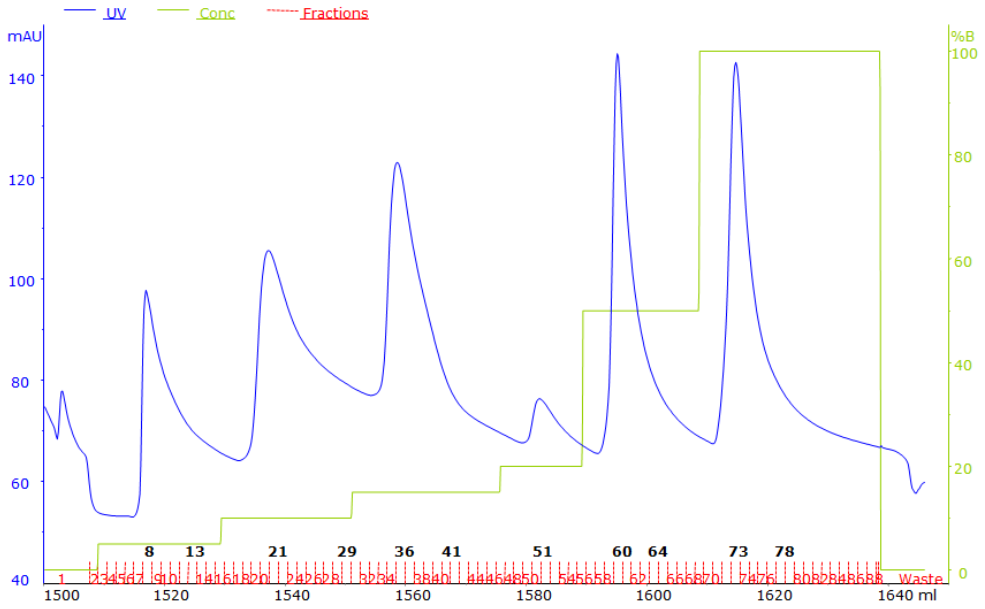


Figure III.44 - Elution chromatogram of HisTrap affinity purification of hDll1 DE6 (in pTT22SSP4) from HEK293-Ebna cells' conditioned media. Bound protein was eluted with steps of increasing concentrations of imidazole (buffer B, right axis) monitored by absorbance at 280nm (in blue, left axis. The analysed fractions (Figure III.45) are shown in black.

After conditioned media injection and column wash, bound protein was eluted in steps of imidazole. In each step, a single peak was eluted (Figure III.44). Selected fractions were analysed by SDS-PAGE in order to discern in which peak hDll1 DE6 was eluted (Figure III.45).

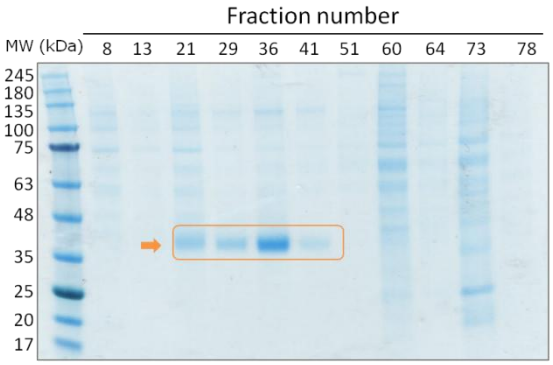


Figure III.45 - SDS-PAGE analysis of hDll1 DE6 HisTrap affinity purification fractions from Figure III.44. Highlighted in orange are the fractions containing hDll1 DE6 protein (32kDa).

By analysing the SDS-PAGE in Figure III.45, we can observe the presence of an intense band in the samples from the eluted fractions 21 trough 41, highlighted in orange. Those samples match to the second and third peaks of the elution chromatogram (Figure III.44), corresponding to an elution with imidazole concentrations of 100 and 150mM respectively.

In order to confirm the identity of the eluted protein, a MS/peptide fingerprinting was performed and the protein identity shown the protein to be in fact human Delta-Like Ligand1 DSL to EGF6 domains (Section VI.4).

The fractions containing hDll1 DE6 were pooled and injected in a S75 SEC column in order to separate our protein of interest from the remaining contaminants (Figure III.46).

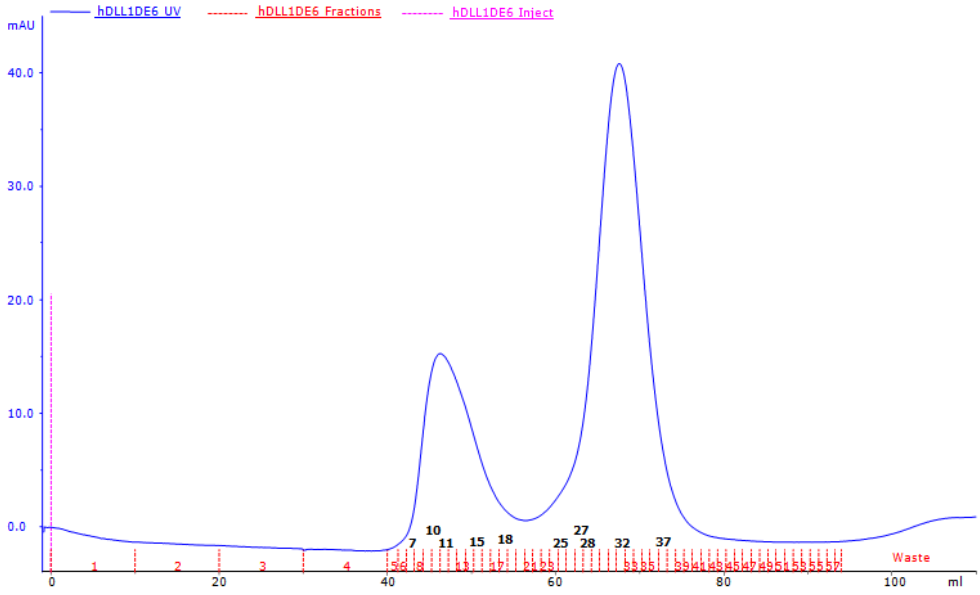


Figure III.46 - Size exclusion chromatogram of hDll1 DE6 purification. Injection is marked in pink. Protein was monitored by absorbance at 280nm (in blue, left axis). Fraction numeration is shown in the bottom, being those analysed in Figure III.47 in black.

Selected fractions from the eluted protein were analysed by SDS-PAGE (Figure III.47, in black). As one can observe in the gel analysis, the first peak corresponds to high molecular weight contaminants and the second to our target, human Dll1 DSL-EGF6, with high purity.

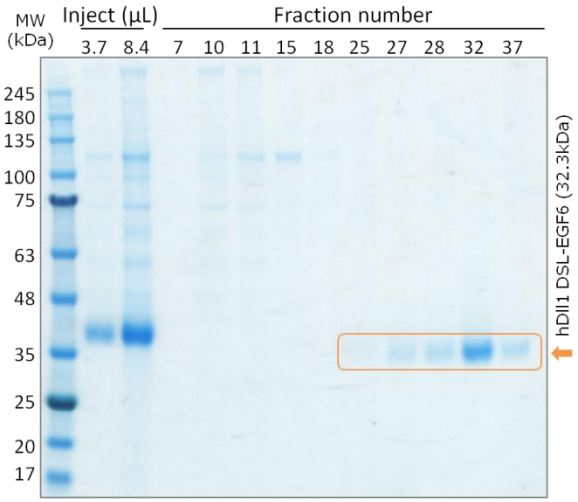


Figure III.47 - SDS-PAGE analysis of hDll1 DE6 SEC fractions. Highlighted in orange are the fractions containing the protein of interest (32kDa).

The fractions from the second peak were pooled and concentrated to a final volume of 230µL and protein concentration, as determined by Bradford assay, to be 0.9mg/mL, 207µg total.

Here, we finally found that with this TGE system using HEK293-EBNA cells and pTT vector-based constructs we can express our proteins of interest that can be purified with a two-column purification scheme. Although the technical difficulties are higher than in bacterial systems, we could obtain a pure protein sample.

The protein obtained was used for a study of another Master student Thesis, Maria Jardim Cabral. The aim of Maria's work is to select specific Fab's against a target protein that intervenes in the Notch1 signalling pathway – the DLL1 ligand, using the phage display technology. As such a panning campaign was carried out using the purified protein.

III.5. Protein Crystallisation

III.5.1 hJagged2 DSL-EGF3 protein purified from inclusion bodies

Protein crystallisation experiments were performed before all the characterizations of hJag2 DE3 described above were completed. As such, both protein pools were used in a grid screen around the published crystallization condition of hJagged1 DSL-EGF3 (Cordle, Johnson *et al.* 2008): sitting drops by the vapour diffusion technique at 4.6mg/ml, 0.2µl of protein and 0.2µl of mother liquor in 8.5% (w/v) PEG 4000, 100mM Imidazole/Malate, pH 7.0. Precipitant was varied between 7.5 and 11% PEG 4000 in 100mM and 200mM Imidazole/Malate, pH 7.0.

Neither condition tested produced a positive result nor something useful for a follow up experiment.

Consequently, crystallisation space was sampled with the commercially available screens as depicted in Table III.5.

Table III.5 – Nano-scale crystallisation commercial screens performed with human Jagged2 DSL-EGF3 purified from Inclusion bodies.

Crystallisation Screen	Temperature (°C)	Protein Pool	Protein conc. (mg/mL)
Stura FootPrint	20	1	5,5
Structure 1+2	20	1	5,5
		1	7,5
		2	7,9
Index	20	1	7,5
		2	7,9
PactPremier	20	1	7,5
		2	7,9
SaltRx	20	1	7,5
		2	7,9
	4	1	7,5
		2	7,9
MPD	20	1	7,5
		2	7,9
	4	1	7,5
		2	7,9
Cryos	20	1	7,5
		2	7,9
	4	1	7,5
		2	7,9

The screens were performed in chronological order as presented in Table III.6. The Stura Footprint together with Structure 1+2 screens were used for an initial assessment of the protein precipitant solubility curve and for sampling the crystallization space with a classical sparse matrix screen, respectively. hJag2 DSL-EGF3 protein at 5.5mg/mL was used for this initial assessment since the hJag1 DE3 protein was crystallised at a similar concentration (Cordle, Johnson *et al.* 2008).

After visual inspection of the crystallization screens, we could observe that more than 90% of the drops were clear, with no signs of precipitation. This indicates that most of the tested conditions are below the saturation curve of the phase diagram (Fig. III.48). Given that Footprint screen explores the protein solubility curve, we could conclude that the protein samples needed to be further concentrated in order to drive the system towards saturation.

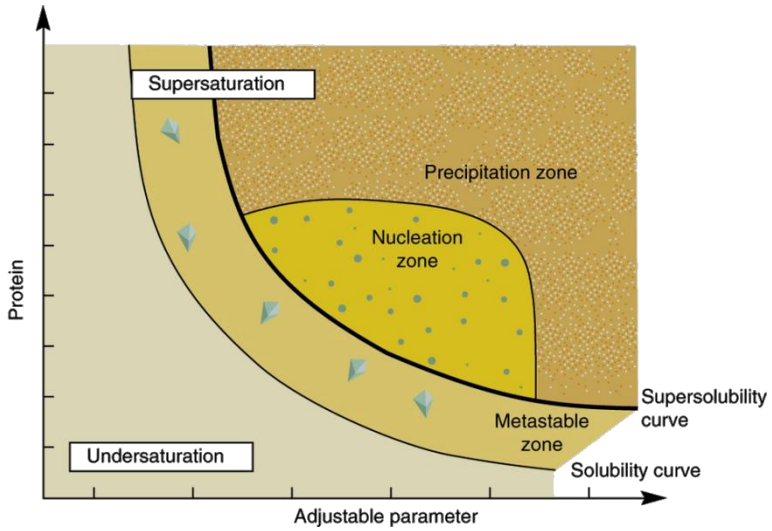


Figure III.48 - Schematic illustration of a protein crystallization phase diagram. The adjustable parameter can be precipitant or additive concentration, pH, temperature and so on. The solubility is defined as the concentration of protein in the solute that is in equilibrium with crystals. Figure from Khurshid, Saridakis *et al.* (2014)

After a second round of protein concentration, both protein pools after hJag2 DE3 refolding (Section III.2.4.ii) were assayed with Structure 1+2 screen. The visual inspection over the first days revealed that the newly obtained protein solutions were more prone to be in the concentration range appropriated for crystallization. A more even distribution of clear and precipitated drops was obtained.

Therefore, the remaining screens presented in Table III.6 were performed. Amongst the available screens in our group, these were selected based on the variety of different reagents' combinations, widely sampling crystallization space. Index is a data-driven biased sparse matrix and grid seen. This screen followed the initial experiments for its broad range in crystallization space sampling and for its data-driven construction, i.e., a combination of crystallization solution that were reported to produce crystals in the literature. PactPremier screen is a systematic grid screen that could provide invaluable clues on the protein behaviour in the crystallization plate. On the other hand, SaltRX screen is a pH and salt matrix screen, sampling a variety of different salts as precipitating agents in a wide pH range, two of the factor that are reported to be crucial for macromolecular crystallisation (McPherson 1990).

Ultimately, MPD and Cryos screens were performed. Both screens have the advantage of including cryoprotectant agents in all solutions. MPD is, as Index, a data-driven grid screen. MPD (2-methyl-2,4-pentandiol) is a precipitant and a cryoprotectant. The rationale for the use of this screen is that MPD lowers the chemical activity of water, thereby reducing the electrostatic interactions between the solvent molecules and the protein in solution, promoting crystallization. Cryos is a sparse matrix screen and has glycerol as a cryoprotectant.

In spite of the amount of crystallization conditions tested and the vast sampling of the crystallization space, no positive results were obtained. All the drops were either clear or precipitated after plate set up. Few clear drops evolved, and those that evolved were to a slight precipitate. Some drops provided crystals that were unbreakable by touch, meaning that those were salt crystals.

The fact that none of the protein samples assayed in the crystallization experiments was homogeneous in terms of dissulfide pattern (Fig. III.30) may contribute to the lack of success in this stage. It is absolutely crucial to have a pure and homogeneous sample for crystallization (McPherson 1990), having the later not been accomplished.

III.5.2 hDelta-Like Ligand1 DSL-EGF6 purified from mammalian cell culture at OPPF

Straight after the last protein purification step of hDll1 DE6, all the screens listed in Table III.7 were performed. This list encompasses the most utilised screens at OPPF, those that generally produce crystals of the variety of proteins produced and purified by the applicants of the European project P-CUBE.

Table III.6 - Nano-scale crystallisation commercial screens performed with human Delta-Like Ligand1 DSL-EGF6 purified from mammalian cells at OPPF.

Crystallisation Screen	Temperature (°C)	Protein Pool	Protein Concentration
Hampton Crystal Screen 1&2	20	1	4.7
		2	4.6
PEG/Ion Screen, Grid Screen P6K, Grid AmSO_4	20	1	4.7
		2	4.6
Salt Rx	20	1	4.7
		2	4.6
Index	20	1	4.7
		2	4.6
	4	1	4.7
		2	4.6
Morpheus	20	1	4.7
		2	4.6
Pact Premier	20	1	4.7
		2	4.6

Crystallization experiments at OPPF were remotely monitored through *xtal*PIMS, which is part of the Protein Information Management System (PIMS) which covers crystallization. Drops were monitored by the automated hotels with integrated imaging system both at 4 and 20°C. All drops were imaged with the defined OPPF schedule: immediately after setup, 4-5 hours later, every day for one week, weekly for a month, and monthly thereafter until the cessation of the experiment. At each inspection, an automated e-mail was received with the link for the new set of images.

All images acquired by the automated system were readily visualised and no positive result was obtained after 222 days (which is the limit of the experiment in the OPPF-HTX facility) in any of the 1344 tested crystallisation conditions. However, one crystal was **obtained (Figure III.49) but it was not** possible to clarify on the salt or protein nature of such crystal.

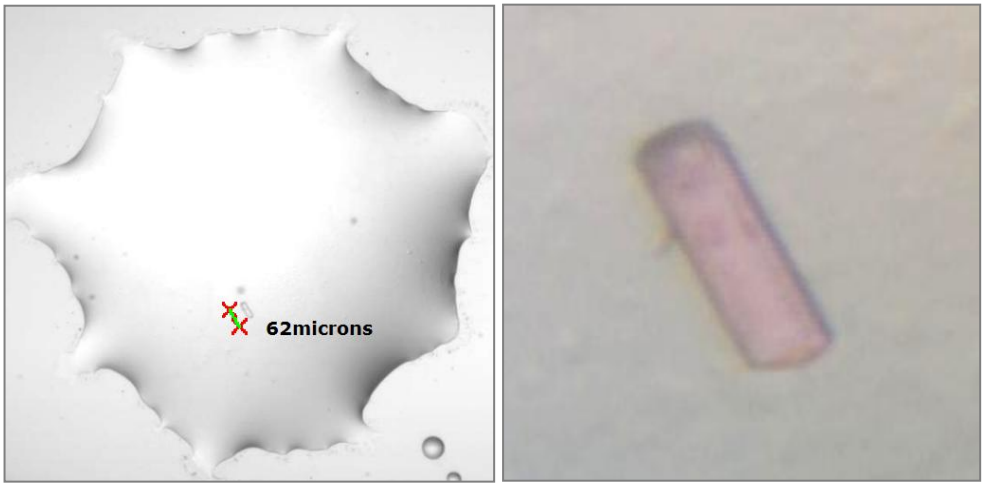


Figure III.49 - Crystal obtained at OPPF-HTX facility. A - Automated image of the drop (crystallisation condition: 20% (w/v) Polyethylene Glycol 3350, 0.2M Sodium Fluoride). B - Manually acquired image of the crystal shown in A with a polarised filter.

Chapter IV Conclusions and Future Perspectives

With the work presented in this thesis, we opened new views for the expression and purification of Notch ligand extracellular domain.

Although bacterial expression systems are widely the first choice host organisms for protein production, we can conclude that because of the nature of our target proteins they may not be suitable. These ECD portions of Notch ligands are hard targets to express and purify, due to their pencil-like shape conformation and disulfide bond content that overwhelm the bacterial cell machinery. As for the purification yields, we could obtain a large amount of hJag2 DSL-EGF3 refolded protein (26.8mg per culture litre) although the quality of the final product could not be ascertain as excellent. With the mammalian expression system we obtained a much lower amount of hDll1 DSL-EGF6, and this was due to the low expression yield (603µg/L at OPPF, and 138µg/L at iBET). Nevertheless, to accomplish the main goal of this project – function-blocking antibodies that can exert their effect *in vivo* – quality is more important than quantity in our opinion.

As such, we can now begin to understand how the Notch ligand domains behave as recombinantly expressed proteins in the test tube. With the knowledge here generated, a more informed choice of expression organism, vector system, and expression conditions may be followed in order to obtain pure and homogeneous protein samples for the extracellular domain portions of Notch ligands.

In face of the knowledge brought by the presented work, we aim to continue the efforts in expressing and purifying Notch Ligand protein domains. The work developed in this thesis did not totally answer all the questions that were raised throughout the entire experimental work. However, this thorough study lead us to a broader knowledge on the behaviour of these types of proteins, and this know-how will be critical for the next steps this project.

We must express and purify more hDll1 DE6 protein, since this is essential to extend the crystallization experiments using higher protein concentration and different crystallization screens. Once the expression in HEK293 cells and protein purification protocols are optimized, and we obtain more protein sample, the continuation of the crystallisation experiments will be pursued. Since the panning campaign has started using human Dll1 DSL-EGF6 protein purified herein, we are a step closer to obtain antibody fragment candidates to pursue for protein expression and purification.

The hDll1 DSL-EGF6 protein will also be used to characterise the binding affinities of the selected antibody fragments after their purification. We also aim to co-crystallise the antibody:protein complex with the most promising antibody fragments, those that show a higher binding affinity for example. We hope that the latter serves as well as a crystallisation scaffold. With the knowledge that will be obtained from the protein structural model, alone and in complex with its Fab counterpart, we hope to be contributing to the elucidation of the molecular features that a function blocking antibody for Notch ligands may require, and understand why and how the Notch1 inhibition pathways functions via blockage of the Dll1 association.

In case the attempts with mammalian expression fail for the remaining constructs, we can always try to increase the percentage of soluble protein obtained with bacterial expression systems. We believe that, even though the quantities are low, an upscale of the whole upstream process would allow a fair amount of protein to be recovered and used for further studies. In such a case, we would however recommend no tag cleavage attempts to be performed so that we would always have an easy method to detect our target protein by Western blot, using an anti-tag antibody. Also, more efforts can be put together in exploring other expression conditions in SHuffle® strain than those presented. Other *E. coli* modified strains for cytoplasmic disulfide formation such as Origami and derivatives could also be tested attempting to screen for the strain with the modification(s) most suitable for the correct expression of our target protein domains. Yet other option could be the periplasmic expression of our target proteins using different vector backbones, for example pET39b or pET40b that harbour the coding sequence of disulfide oxidoreductase DsbA and disulfide isomerase DsbC respectively.

With the generation of different protein samples, we hope to generate more antibody fragments and characterize their binding affinities carrying out assays using SPR, Biacore or Octet. The selected antibody fragments will be ranked according to binding affinities and kinetics. The top ranked fragments that show the highest binding affinity will be subsequently used in the crystallisation experiments, cellular assays and elucidation of the molecular features of the function blocking antibodies.

We hope that we can select a couple of good lead candidates that will be further tested in cellular assays, for Notch pathway inhibition. For example, cell proliferation assays will be conducted to determine the effect of the isolated anti-Dll1 on the growth of breast cancer cells with moderate or high level expression of Dll1.

With the increased knowledge generated by the work presented in this thesis and the future work ahead, our goal is to find a highly specific and active function blocking antibody for Notch ligands, and this will help elucidating the molecular basis of anti-Dll1 antagonistic activity in the Notch1 signaling pathway.

Chapter V References

Aricescu, A. R., W. Lu and E. Y. Jones (2006). "A time- and cost-efficient system for high-level protein production in mammalian cells." Acta Crystallogr D Biol Crystallogr **62**(Pt 10): 1243-1250.

Artavanis-Tsakonas, S., M. D. Rand and R. J. Lake (1999). "Notch signaling: cell fate control and signal integration in development." Science **284**(5415): 770-776.

Backliwal, G., M. Hildinger, S. Chenuet, S. Wulhfard, M. De Jesus and F. M. Wurm (2008). "Rational vector design and multi-pathway modulation of HEK 293E cells yield recombinant antibody titers exceeding 1 g/l by transient transfection under serum-free conditions." Nucleic Acids Res **36**(15): e96.

Backliwal, G., M. Hildinger, I. Kuettel, F. Delegrange, D. L. Hacker and F. M. Wurm (2008). "Valproic acid: a viable alternative to sodium butyrate for enhancing protein expression in mammalian cell cultures." Biotechnol Bioeng **101**(1): 182-189.

Bell, M. R., M. J. Engleka, A. Malik and J. E. Strickler (2013). "To fuse or not to fuse: What is your purpose?" Protein Sci **22**(11): 1466-1477.

Chen, J., Y. Liu, Y. Wang, H. Ding and Z. Su (2008). "Different effects of L-arginine on protein refolding: suppressing aggregates of hydrophobic interaction, not covalent binding." Biotechnol Prog **24**(6): 1365-1372.

Chillakuri, C. R., D. Sheppard, M. X. Ilagan, L. R. Holt, F. Abbott, S. Liang, R. Kopan, P. A. Handford and S. M. Lea (2013). "Structural analysis uncovers lipid-binding properties of Notch ligands." Cell Rep **5**(4): 861-867.

Cinatl, J., Jr., J. Cinatl, P. H. Driever, R. Kotchetkov, P. Pouckova, B. Kornhuber and D. Schwabe (1997). "Sodium valproate inhibits in vivo growth of human neuroblastoma cells." Anticancer Drugs **8**(10): 958-963.

Cordle, J., S. Johnson, J. Z. Tay, P. Roversi, M. B. Wilkin, B. H. de Madrid, H. Shimizu, S. Jensen, P. Whiteman, B. Jin, C. Redfield, M. Baron, S. M. Lea and P. A. Handford (2008). "A conserved face of the Jagged/Serrate DSL domain is involved in Notch trans-activation and cis-inhibition." Nat Struct Mol Biol **15**(8): 849-857.

de Marco, A., L. Vigh, S. Diamant and P. Goloubinoff (2005). "Native folding of aggregation-prone recombinant proteins in Escherichia coli by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones." Cell Stress Chaperones **10**(4): 329-339.

Diepenbruck, C., M. Klinger, T. Urbig, P. Baeuerle and R. Neef (2013). "Productivity and quality of recombinant proteins produced by stable CHO cell clones can be predicted by transient expression in HEK cells." Mol Biotechnol **54**(2): 497-503.

Durocher, Y., S. Perret and A. Kamen (2002). "High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells." Nucleic Acids Res **30**(2): E9.

Ericsson, U. B., B. M. Hallberg, G. T. Detitta, N. Dekker and P. Nordlund (2006). "Thermofluor-based high-throughput stability optimization of proteins for structural studies." Anal Biochem **357**(2): 289-298.

Fehon, R. G., P. J. Kooh, I. Rebay, C. L. Regan, T. Xu, M. A. Muskavitch and S. Artavanis-Tsakonas (1990). "Molecular interactions between the protein products of the neurogenic loci Notch and Delta, two EGF-homologous genes in Drosophila." Cell **61**(3): 523-534.

Fleming, R. J. (1998). "Structural conservation of Notch receptors and ligands." Semin Cell Dev Biol **9**(6): 599-607.

Garcia, A. and J. J. Kandel (2012). "Notch: a key regulator of tumor angiogenesis and metastasis." Histol Histopathol **27**(2): 151-156.

Greenblatt, D. Y., A. M. Vaccaro, R. Jaskula-Sztul, L. Ning, M. Haymart, M. Kunnimalaiyaan and H. Chen (2007). "Valproic acid activates notch-1 signaling and regulates the neuroendocrine phenotype in carcinoid cancer cells." Oncologist **12**(8): 942-951.

Greenwald, I. (1994). "Structure/function studies of lin-12/Notch proteins." Curr Opin Genet Dev **4**(4): 556-562.

Gustafsson, C., S. Govindarajan and J. Minshull (2004). "Codon bias and heterologous protein expression." Trends Biotechnol **22**(7): 346-353.

Heitzler, P. and P. Simpson (1993). "Altered epidermal growth factor-like sequences provide evidence for a role of Notch as a receptor in cell fate decisions." Development **117**(3): 1113-1123.

Jager, V., K. Bussow, A. Wagner, S. Weber, M. Hust, A. Frenzel and T. Schirrmann (2013). "High level transient production of recombinant antibodies and antibody fusion proteins in HEK293 cells." BMC Biotechnol **13**(1): 52.
Jana, S. and J. K. Deb (2005). "Strategies for efficient production of heterologous proteins in Escherichia coli." Appl Microbiol Biotechnol **67**(3): 289-298.
Jonasson, P., S. Liljeqvist, P. A. Nygren and S. Stahl (2002). "Genetic design for facilitated production and recovery of recombinant proteins in Escherichia coli." Biotechnol Appl Biochem **35**(Pt 2): 91-105.
Khurshid, S., E. Saridakis, L. Govada and N. E. Chayen (2014). "Porous nucleating agents for protein crystallization." Nat Protoc **9**(7): 1621-1633.
Knott, V., A. K. Downing, C. M. Cardy and P. Handford (1996). "Calcium binding properties of an epidermal growth factor-like domain pair from human fibrillin-1." J Mol Biol **255**(1): 22-27.
Kopan, R. and M. X. Ilagan (2009). "The canonical Notch signaling pathway: unfolding the activation mechanism." Cell **137**(2): 216-233.
Leal, S. S. and C. M. Gomes (2007). "Studies of the molten globule state of ferredoxin: structural characterization and implications on protein folding and iron-sulfur center assembly." Proteins **68**(3): 606-616.
Lindsell, C. E., C. J. Shawber, J. Boulter and G. Weinmaster (1995). "Jagged: a mammalian ligand that activates Notch1." Cell **80**(6): 909-917.
Lobstein, J., C. A. Emrich, C. Jeans, M. Faulkner, P. Riggs and M. Berkmen (2012). "SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm." Microb Cell Fact **11**: 56.
McPherson, A. (1990). "Current approaches to macromolecular crystallization." Eur J Biochem **189**(1): 1-23.
Mittal, S., D. Subramanyam, D. Dey, R. V. Kumar and A. Rangarajan (2009). "Cooperation of Notch and Ras/MAPK signaling pathways in human breast carcinogenesis." Mol Cancer **8**: 128.
Nichols, J. T., A. Miyamoto, S. L. Olsen, B. D'Souza, C. Yao and G. Weinmaster (2007). "DSL ligand endocytosis physically dissociates Notch1 heterodimers before activating proteolysis can occur." J Cell Biol **176**(4): 445-458.
Pear, W. S., G. P. Nolan, M. L. Scott and D. Baltimore (1993). "Production of high-titer helper-free retroviruses by transient transfection." Proc Natl Acad Sci U S A **90**(18): 8392-8396.
Pham, P. L., S. Perret, B. Cass, E. Carpentier, G. St-Laurent, L. Bisson, A. Kamen and Y. Durocher (2005). "Transient gene expression in HEK293 cells: peptone addition posttransfection improves recombinant protein synthesis." Biotechnol Bioeng **90**(3): 332-344.
Pintar, A., C. Guarnaccia, S. Dhir and S. Pongor (2009). "Exon 6 of human JAG1 encodes a conserved structural unit." BMC Struct Biol **9**: 43.
Purow, B. (2012). "Notch inhibition as a promising new approach to cancer therapy." Adv Exp Med Biol **727**: 305-319.
Rebay, I., R. J. Fleming, R. G. Fehon, L. Cherbas, P. Cherbas and S. Artavanis-Tsakonas (1991). "Specific EGF repeats of Notch mediate interactions with Delta and Serrate: implications for Notch as a multifunctional receptor." Cell **67**(4): 687-699.
Ridgway, J., G. Zhang, Y. Wu, S. Stawicki, W. C. Liang, Y. Chanthery, J. Kowalski, R. J. Watts, C. Callahan, I. Kasman, M. Singh, M. Chien, C. Tan, J. A. Hongo, F. de Sauvage, G. Plowman and M. Yan (2006). "Inhibition of Dll4 signalling inhibits tumour growth by deregulating angiogenesis." Nature **444**(7122): 1083-1087.
Rizzo, P., C. Osipo, K. Foreman, T. Golde, B. Osborne and L. Miele (2008). "Rational targeting of Notch signaling in cancer." Oncogene **27**(38): 5124-5131.
Robey, E. (1997). "Notch in vertebrates." Curr Opin Genet Dev **7**(4): 551-557.
Rudolph, R. and H. Lilie (1996). "In vitro folding of inclusion body proteins." FASEB J **10**(1): 49-56.
Sambrook, J. (2001). Molecular cloning: a laboratory manual (3-volume set). 3rd Edition.
Santos, S. P., T. M. Bandeiras, A. F. Pinto, M. Teixeira, M. A. Carrondo and C. V. Romao (2012). "Thermofluor-based optimization strategy for the stabilization and crystallization of Campylobacter jejuni desulforubrerythrin." Protein Expr Purif **81**(2): 193-200.
Schofield, D. J., A. R. Pope, V. Clementel, J. Buckell, S. Chapple, K. F. Clarke, J. S. Conquer, A. M. Crofts, S. R. Crowther, M. R. Dyson, G. Flack, G. J. Griffin, Y. Hooks, W. J. Howat, A. Kolb-Kokocinski, S. Kunze, C. D. Martin, G. L. Maslen, J. N. Mitchell, M. O'Sullivan, R. L. Perera, W. Roake, S. P. Shadbolt, K. J. Vincent, A. Warford, W. E. Wilson, J. Xie, J. L. Young and J. McCafferty (2007). "Application of phage display to high throughput antibody generation and characterization." Genome Biol **8**(11): R254.
Sethi, N., X. Dai, C. G. Winter and Y. Kang (2011). "Tumor-derived JAGGED1 promotes osteolytic bone metastasis of breast cancer by engaging notch signaling in bone cells." Cancer Cell **19**(2): 192-205.

Sharma, A., A. Rangarajan and R. R. Dighe (2013). "Antibodies against the extracellular domain of human Notch1 receptor reveal the critical role of epidermal-growth-factor-like repeats 25-26 in ligand binding and receptor activation." Biochem J **449**(2): 519-530.
Shimizu, K., S. Chiba, K. Kumano, N. Hosoya, T. Takahashi, Y. Kanda, Y. Hamada, Y. Yazaki and H. Hirai (1999). "Mouse jagged1 physically interacts with notch2 and other notch receptors. Assessment by quantitative methods." J Biol Chem **274**(46): 32961-32969.
Sorensen, H. P. and K. K. Mortensen (2005). "Advanced genetic strategies for recombinant protein expression in Escherichia coli." J Biotechnol **115**(2): 113-128.
South, A. P., R. J. Cho and J. C. Aster (2012). "The double-edged sword of Notch signaling in cancer." Semin Cell Dev Biol **23**(4): 458-464.
Sprinzak, D., A. Lakhanpal, L. Lebon, L. A. Santat, M. E. Fontes, G. A. Anderson, J. Garcia-Ojalvo and M. B. Elowitz (2010). "Cis-interactions between Notch and Delta generate mutually exclusive signalling states." Nature **465**(7294): 86-90.
Swiech, K., V. Picanco-Castro and D. T. Covas (2012). "Human cells: new platform for recombinant therapeutic protein production." Protein Expr Purif **84**(1): 147-153.
Takai, N., J. C. Desmond, T. Kumagai, D. Gui, J. W. Said, S. Whittaker, I. Miyakawa and H. P. Koeffler (2004). "Histone deacetylase inhibitors have a profound antigrowth activity in endometrial cancer cells." Clin Cancer Res **10**(3): 1141-1149.
van Es, J. H., M. E. van Gijn, O. Riccio, M. van den Born, M. Vooijs, H. Begthel, M. Cozijnsen, S. Robine, D. J. Winton, F. Radtke and H. Clevers (2005). "Notch/gamma-secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells." Nature **435**(7044): 959-963.
Wang, M. M. (2011). "Notch signaling and Notch signaling modifiers." Int J Biochem Cell Biol **43**(11): 1550-1562.
Wu, Y., C. Cain-Hom, L. Choy, T. J. Hagenbeek, G. P. de Leon, Y. Chen, D. Finkle, R. Venook, X. Wu, J. Ridgway, D. Schahin-Reed, G. J. Dow, A. Shelton, S. Stawicki, R. J. Watts, J. Zhang, R. Choy, P. Howard, L. Kadyk, M. Yan, J. Zha, C. A. Callahan, S. G. Hymowitz and C. W. Siebel (2010). "Therapeutic antibody targeting of individual Notch receptors." Nature **464**(7291): 1052-1057.
Wulhfard, S. (2009). Transient Recombinant Protein Expression in Mammalian Cells: the Role of mRNA Level and Stability. PhD, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE.
Xing, F., H. Okuda, M. Watabe, A. Kobayashi, S. K. Pai, W. Liu, P. R. Pandey, K. Fukuda, S. Hirota, T. Sugai, G. Wakabayshi, K. Koeda, M. Kashiwaba, K. Suzuki, T. Chiba, M. Endo, Y. Y. Mo and K. Watabe (2011). "Hypoxia-induced Jagged2 promotes breast cancer metastasis and self-renewal of cancer stem-like cells." Oncogene **30**(39): 4075-4086.
Yan, M., C. A. Callahan, J. C. Beyer, K. P. Allamneni, G. Zhang, J. B. Ridgway, K. Niessen and G. D. Plowman (2010). "Chronic DLL4 blockade induces vascular neoplasms." Nature **463**(7282): E6-7.
Yaneza, M., J. D. Gilthorpe, A. Lumsden and A. S. Tucker (2002). "No evidence for ventrally migrating neural tube cells from the mid- and hindbrain." Dev Dyn **223**(1): 163-167.
Zagouras, P., S. Stifani, C. M. Blaumueller, M. L. Carcangiu and S. Artavanis-Tsakonas (1995). "Alterations in Notch signaling in neoplastic lesions of the human cervix." Proc Natl Acad Sci U S A **92**(14): 6414-6418.
Zhao, Y., B. Bishop, J. E. Clay, W. Lu, M. Jones, S. Daenke, C. Siebold, D. I. Stuart, E. Y. Jones and A. R. Aricescu (2011). "Automation of large scale transient protein expression in mammalian cells." J Struct Biol **175**(2): 209-215.

Chapter VI Appendix

VI.1. Reagents List

Table VI.1 - Reagents and equipments used in this study, organized by alphabetical order.

Reagent / Equipment	Supplier	Catalogue number
24-well Cryschem Plate	Hampton Research	HR3-158
2xYT media	Applichem	A0981
3L Fernbach design erlenmeyer	Corning	431252
96-well low profile plates	Bio-Rad	MLL9651
Agar	Nzytech	MB02902
Agarose	SeaKem	50004
AgeI	NEB	R3552S
Amicon 10kDa MWCO	Millipore	UFC901024
Amicon Stirred Cell unit	Millipore	8050
Ampicillin	Sigma	A9518
*Bam*HI	Fermentas	ER0055
BenchMark™ Protein Ladder	Invitrogen	10747-012
Benzonase	Novagen	70746
Beta-mercaptoethanol	Carl Roth GmbH	4227.1
Bio-Rad Protein Assay Dye Reagent Concentrate	Bio-Rad	500-0006
BugBuster Protein Extraction Reagent	Novagen	70584
Calcium chloride	Calbiochem	3000
Chloramphenicol	Sigma	C0378
Criterion XT Precast Gel	Bio-Rad	345-0124 / 345-0125
CrystalQuick™ 96 Well, Greiner	Hampton Research	HR3-281
Cystamine	Sigma	30050
Cysteamine	Sigma	M6500
DMEM	Gibco	41966
DMEM	Sigma	D6546
dNTP mix	Nzytech	MB08604
*Eco*RI	Fermentas	ER0271
*Eco*RV	Fermentas	ER0305
EDTA	Carl Roth GmbH	8040.2
FCS	Gibco	16250
	Invitrogen	10270
FreeStyle 293 media	Gibco	12338
FreeStyle F17 media	Gibco	A13835
GeneJET Plasmid Miniprep Kit	Fermentas	K0502
Gentamycin	Gibco	15710
Glutamax	Gibco	35050
Glycerol	Calbiochem	4760
Goat α-mouse IgG-HRP conjugated antibody	Sigma	A9917
GSH	Sigma	G4251
GSSG	Sigma	G4376
GSTrap column	GE Healthcare Life Sciences	17-5281-01

Reagent / Equipment	Supplier	Catalogue number
Hepes	Carl Roth GmbH	9105.3
HiLoad 16/600 Superdex 200 pg column	GE Healthcare Life Sciences	28-9893-35
HiLoad 16/600 Superdex 75pg column	GE Healthcare Life Sciences	28-9893-33
*Hind*III	Fermentas	ER0501
HiPrep 26/10 desalting column	GE Healthcare Life Sciences	17-5087-01
HisTrap column	GE Healthcare Life Sciences	17-5248-02
Imidazole	Merck	1.04716.1000
IPTG	Nzytech	MB02602
Kanamycin	Sigma	K1876
Kifunensine	Cayman Chemical	10009437
Klenow fragment	Fermentas	EP0051
*Kpn*I	Fermentas	ER0521
L-arabinose	Sigma	A3256
L-Arginine	Sigma	W381918
Lipofectamine2000	Invitrogen	11668-019
MES Running Buffer	Bio-Rad	161-0789
Milk Powder	Molico	n/a
Mini-PROTEAN TGX precast Gels	Bio-Rad	456-1083 / 456-1086
MOPS Running Buffer	Bio-Rad	161-0788
Mouse α–His antibody	GE Healthcare Life Sciences	27-4710-01
*Nhe*I	Fermentas	ER0975
Nitrocellulose membranes	GE Healthcare Life Sciences	RPN203D
NP-40	Sigma	NP40S
NuPAGE LDS sample buffer	Invitrogen	NP0007
NuPAGE Transfer Buffer	Life Technologies	NP0006-1
NZYColour Protein Marker II	Nzytech	MB090
Optical quality sealing tape	Bio-Rad	MSB-1001
PBS	Sigma	D5773
	Gibco	14190
	Lonza	BE17-516F
PD-10 columns	GE Healthcare Life Sciences	17-0851-01
PEI	Sigma	40872-7
	PolySciences	23966
Peristaltic pump P-1	GE Healthcare Life Sciences	18-1110-91
pET47b(+) (kan^R)	Novagen	71461-3
Pfu DNA polymerase	Fermentas	EP0501
pGro7 (Cm^R)	TaKaRa	3340
Pluronic F-68	Gibco	24040
Potassium phosphate dibasic	Prolabo	26931.263
Potassium phosphate monobasic	Prolabo	26936.293
Power Broth (PB)	AthenaES	MD12-106-1
Protease inhibitor cocktail	Roche	11873580001
SHuffle® T7 express competent cells	NEB	C3029
SimplyBlue SafeStain	Life Technologies	LC6060
Sodium chloride	Carl Roth GmbH	9265.1

Reagent / Equipment	Supplier	Catalogue number
Sodium phosphate dibasic	Merck	1.06586
Sodium phosphate monobasic	Merck	1.06346
SteriTop-GP Filter Unit	Millipore	SCGPT10RE
Superose 6 PC 3.2/30 column	GE Healthcare Life Sciences	17-0673-01
SYPRO Orange	Molecular Probes	S6650
T4 DNA ligase	Fermentas	EL0011
TALON metal Affinity resin	Clontech	635504
Taq DNA polymerase	Fermentas	EP0402
Terrific Broth (TB)	MoBio	12105-1
TMAO	Sigma	92277
Tris	Merck	108387
Trypan blue 0.4% solution	Gibco	15250-061
Trypsin-EDTA	Gibco	25300
TunAir full baffle flasks	Sigma	Z710822
Tween-20	Sigma	P7949
VacuCap Filter Unit 0.22µm filter	PALL Life Sciences	4628
Western Lightning Plus-ECL Enhanced Chemiluminescence Substrate	Perkin-Elmer	NEL104001EA

VI.2. Buffers

1X SDS-PAGE Loading Buffer: 50mM Tris-Cl pH 6.8, 2% SDS, 0.1% Bromophenol Blue, 10% Glycerol, 100mM BME

KPi: Potassium Phosphate buffer

Na/KPi: Sodium / Potassium Phosphate buffer

Home-made Lysis Buffer: 50mM KPi, 150mM NaCl, 1% NP-40, 1mM PMSF, 5% glycerol, pH 7.5

VI.2.i His_6-hJAG2 DSL-EGF3 purification from soluble faction

HisTrap Buffer A: 50mM Potassium Phosphate buffer, 500mM NaCl, 0.5% NP-40, 5mM$MgCl_2$, 20mM Imidazole, 5% Glycerol, 5mM BME, pH 7.5.

Lysis Buffer: HisTrap Buffer A freshly supplemented with Protease inhibitor cocktail and 250 units of Benzonase.

HisTrap Buffer B: 50mM Potassium Phosphate buffer, 150mM NaCl, 0.5% NP-40, 5mM$MgCl_2$, 1M Imidazole, 5% Glycerol, 5mM BME, pH 7.5.

HRV-3C protease buffer: 20mM Tris, 150mM NaCl, 0.5mM EDTA, 10% Glycerol, 5mM BME, pH 8.0.

SEC Buffer: 20mM Tris, 150mM NaCl, 10% Glycerol, 5mM BME, pH 8.0.

VI.2.ii His_6-hJAG2 DSL-EGF9 purification from soluble fration

HisTrap Buffer A: 50mM Potassium Phosphate buffer, 500mM NaCl, 0.5% NP-40, 5mMmgCl2, 50mM Imidazole, 5% Glycerol, 5mM BME, pH 7.5.

Lysis Buffer: HisTrap Buffer A freshly supplemented with Protease inhibitor cocktail and 250 units of Benzonase.

HisTrap Buffer B: 50mM Potassium Phosphate buffer, 150mM NaCl, 5mMmgCl2, 1M Imidazole, 5% Glycerol, 5mM BME, pH 7.5.

VI.2.iii hJagged2 DSL-EGF3 inclusion bodies refolding tests

Ressuspension / wash buffer: 50mM Tris, 100mM NaCl, 0.5% Triton X-100, pH 8.0.

Solubilization buffer (I): 6M GuHCl, 50mM Tris, 100mM NaCl, 10mM DTT, pH 8.0
Solubilization buffer (II): 8M Urea, 50mM Tris, 100mM NaCl, 10mM DTT, pH 8.0

Refolding buffer A: 200mM Tris, 3.7mM Cystamine, 6.5mM Cysteamine, 1M L-Arginine, pH 8.0,
Refolding buffer B: 20mM Tris, 1mM EDTA, 1mM GSH, 0.1mM GSSG, pH 8.0,
Refolding buffer C: 200mM Tris, 3.7mM Cystamine, 6.5mM Cysteamine, 1M L-Arginine, 1M TMAO, pH 8.0;
Refolding buffer D: 20mM Tris, 1mM EDTA, 1mM GSH, 0.1mM GSSG, 1M L-Arginine and 1M TMAO, pH 8.0.

VI.2.iv hJagged2 DSL-EGF3 purification from inclusion bodies

Solubilisation Buffer: 6M GuHCl, 50mM Na/K Pi, 150mM KCl, 10mM DTT, pH 7.0
Refolding Buffer: 50mM Na/K Pi, 150mM KCl, 1mM EDTA, 1mM GSH, 0.1mM GSSG, 1M L-Arginine, pH 7.0
Desalting buffer: 50mM Tris, 300mM NaCl, pH 7.5
SEC Buffer: 50mM Na/KPi, 300mM NaCl, pH 7.0

VI.2.v hDelta-Like Ligand1 DSL-EGF6 purification at Oxford

Cobalt beads wash buffer 1: PBS, 5mM Imidazole
Cobalt beads wash buffer 2: PBS, 10mM Imidazole
Cobalt beads elution buffer1: PBS, 350mM Imidazole
SEC Buffer: 10mM Hepes, 150mM NaCl, pH 7.5

VI.2.vi hDelta-Like Ligand1 DSL-EGF6 purification at iBET

HisTrap Buffer A: 50mM Tris, 5mM $CaCl_2$, pH 7.5
HisTrap Buffer B: 50mM Tris, 5mM $CaCl_2$, 300mM NaCl, 10mM Imidazole, pH 7.5
HisTrap Buffer C: 50mM Tris, 5mM $CaCl_2$, 1M Imidazole, pH 7.5
SEC Buffer: 50mM Tris, 5mM $CaCl_2$, 150mM NaCl

VI.3. DSF Buffer screen plate layout

Table VI.2 - DSF Buffer and salt screen in a 96-well plate layout.

	1	2	3	4	5	6	7	8	9	10	11	12
A	Glycine pH 3.0	Citric Acid pH 3.2	PIPPS pH 3.7	Citric Acid pH 4.0	Sodium Acetate pH 4.5	Na / K Pi pH 5.0	Sodium Citrate pH 5.5	Na / K Pi pH 6.0	Bis-Tris pH 6.0	MES pH 6.2	ADA pH 6.5	Bis-Tris Propane pH 6.5
B	Ammonium Acetate pH 7.0	MOPS pH 7.0	Na / K Pi pH 7.0	HEPES pH 7.5	Tris pH 7.5	EPPS pH 8.0	Imidazole pH 8.0	Bicine pH 8.5	Tris pH 8.5	CHES pH 9.0	CHES pH 9.5	CAPS pH 10.0
C	Glycine pH 3.0, 150mM NaCl	Citric Acid pH 3.2, 150mM NaCl	PIPPS pH 3.7, 150mM NaCl	Citric Acid pH 4.0, 150mM NaCl	Sodium Acetate pH 4.5, 150mM NaCl	Na / K Pi pH 5.0, 150mM NaCl	Sodium Citrate pH 5.5, 150mM NaCl	Na / K Pi pH 6.0, 150mM NaCl	Bis-Tris pH 6.0, 150mM NaCl	MES pH 6.2, 150mM NaCl	ADA pH 6.5, 150mM NaCl	Bis-Tris Propane pH 6.5, 150mM NaCl
D	Ammonium Acetate pH 7.0, 150mM NaCl	MOPS pH 7.0, 150mM NaCl	Na / K Pi pH 7.0, 150mM NaCl	HEPES pH 7.5, 150mM NaCl	Tris pH 7.5, 150mM NaCl	EPPS pH 8.0, 150mM NaCl	Imidazole pH 8.0, 150mM NaCl	Bicine pH 8.5, 150mM NaCl	Tris pH 8.5, 150mM NaCl	CHES pH 9.0, 150mM NaCl	CHES pH 9.5, 150mM NaCl	CAPS pH 10.0, 150mM NaCl
E	Glycine pH 3.0, 500mM NaCl	Citric Acid pH 3.2, 500mM NaCl	PIPPS pH 3.7, 500mM NaCl	Citric Acid pH 4.0, 500mM NaCl	Sodium Acetate pH 4.5, 500mM NaCl	Na / K Pi pH 5.0, 500mM NaCl	Sodium Citrate pH 5.5, 500mM NaCl	Na / K Pi pH 6.0, 500mM NaCl	Bis-Tris pH 6.0, 500mM NaCl	MES pH 6.2, 500mM NaCl	ADA pH 6.5, 500mM NaCl	Bis-Tris Propane pH 6.5, 500mM NaCl
F	Ammonium Acetate pH 7.0, 500mM NaCl	MOPS pH 7.0, 500mM NaCl	Na / K Pi pH 7.0, 500mM NaCl	Hepes pH 7.5, 500mM NaCl	Tris pH 7.5, 500mM NaCl	EPPS pH 8.0, 500mM NaCl	Imidazole pH 8.0, 500mM NaCl	Bicine pH 8.5, 500mM NaCl	Tris pH 8.5, 500mM NaCl	CHES pH 9.0, 500mM NaCl	CHES pH 9.5, 500mM NaCl	CAPS pH 10.0, 500mM NaCl
G	Glycine pH 3.0, 150mM KCl	Citric Acid pH 3.2, 150mM KCl	PIPPS pH 3.7, 150mM KCl	Citric Acid pH 4.0, 150mM KCl	Sodium Acetate pH 4.5, 150mM KCl	Na / K Pi pH 5.0, 150mM KCl	Sodium Citrate pH 5.5, 150mM KCl	Na / K Pi pH 6.0, 150mM KCl	Bis-Tris pH 6.0, 150mM KCl	MES pH 6.2, 150mM KCl	ADA pH 6.5, 150mM KCl	Bis-Tris Propane pH 6.5, 150mM KCl
H	Ammonium Acetate pH 7.0, 150mM KCl	MOPS pH 7.0, 150mM KCl	Na / K Pi pH 7.0, 150mM KCl	HEPES pH 7.5, 150mM KCl	Tris pH 7.5, 150mM KCl	EPPS pH 8.0, 150mM KCl	Imidazole pH 8.0, 150mM KCl	Bicine pH 8.5, 150mM KCl	Tris pH 8.5, 150mM KCl	CHES pH 9.0, 150mM KCl	CHES pH 9.5, 150mM KCl	CAPS pH 10.0, 150mM KCl

VI.4. MS/peptide fingerprinting report[2]

Methodology:

The protein band was destained, reduced, alkylated and digested with trypsin (Promega, 6.7ng/µl) overnight at 37°C. The tryptic peptides were desalted and concentrated using POROS R2 (Applied Biosystems) and eluted directly onto the MALDI plate using 0.6µl of 5mg/ml CHCA (alpha-cyano-4-hydroxycinnamic acid, Sigma) in 50% (v/v) acetonitrile and 5% (v/v) formic acid.

The data was acquired in positive reflector MS and MS/MS modes using a 4800*plus* MALDI-TOF/TOF (AB Sciex) mass spectrometer and using 4000 Series Explorer Software v.3.5.3 (Applied Biosystems). External calibration was performed using Pepmix1 (Laser BioLabs).

The fifteen most intense precursor ions from the MS spectra were selected for MS/MS analysis. The MS and MS/MS data were analyzed in combined search mode (MS+MS/MS) using GPS Explorer Software v.3.6 (Applied Biosystems) and MASCOT search engine. The search parameters were as follows: monoisotopic peptide mass values were considered, maximum precursor mass tolerance (MS) of 50ppm and a maximum fragment mass tolerance (MS/MS) of 0.3Da. The searches were performed against a custom protein database (containing the Sponsor's protein sequences and the SwissProt protein database). A maximum of one missed cleavage was allowed. Carboxyamidomethylation of cysteines, oxidation of methionines, and deamidation of asparagines and glutamine were set as variable modifications.

Protein identification was only accepted when significant protein homology scores were obtained ($p<0.05$) and at least one peptide was fragmented with a significant individual ion score ($p<0.05$).

[2] Data provided/obtained by the Mass Spectrometry Unit (UniMS), ITQB/iBET, Oeiras, Portugal"

Results:

Protein sequence coverage:

MASCOT SEARCH RESULTS

PROTEIN VIEW

Match to: sp|Dll1| Score: 499 Expect: 6.5e-045

Dll1 DSL-EGF6 WITH SIGNAL PEPTIDE FROM VECTOR

Nominal mass (M_r): 36229; Calculated pI value: 5.85

NCBI BLAST search of sp|Dll1| against nr

Unformatted sequence string for pasting into other applications

Variable modifications: Carbamidomethyl (C),Deamidated (NQ),Gln->pyro- Glu (N-term Q),Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P Sequence Coverage: 34%

Matched peptides shown in Red

```
  1 MGELLLLLLL GLRLQLSLGA GAPGSSTGHH HHHHHHGSTG ENLYFQGASW
 51 SQDLHSSGRT DLKYSYRFVC DEHYYGEGCS VFCRPRDDAF GHFTCGERGE
101 KVCNPGWKGP YCTEPICLPG CDEQHGFCDK PGECKCRVGW QGRYCDECIR
151 YPGCLHGTCQ QPWQCNCQEG WGGLFCNQDL NYCTHHKPCK NGATCTNTGQ
201 GSYTCSCRPG YTGATCELGI DECDPSPCKN GGSCTDLENS YSCTCPPGFY
251 GKICELSAMT CADGPCFNGG RCSDSPDGGY SCRCPVGYSG FNCEKKIDYC
301 SSSPCSNGAK CVDLGDAYLC RCQAGFSGRH CDDN
```

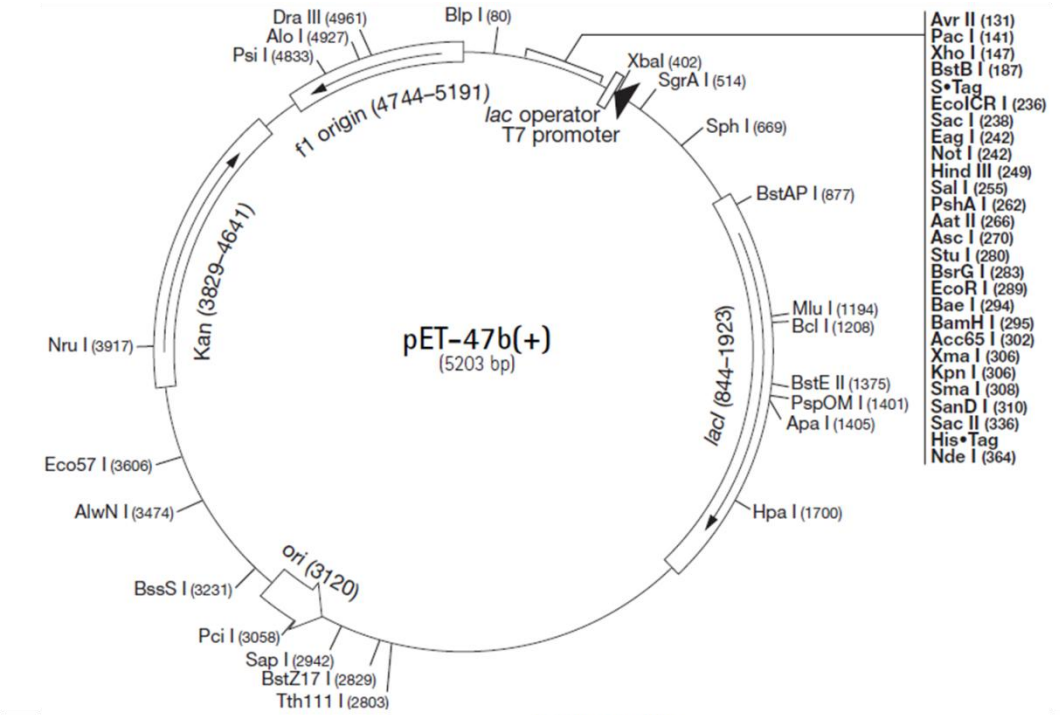
Delta-like protein 1 OS=Homo sapiens GN=DLL1 PE=2 SV=2	sp|O00548|DLL1_HUMAN	486	100	478	100	78003.6	0	8	52.311		

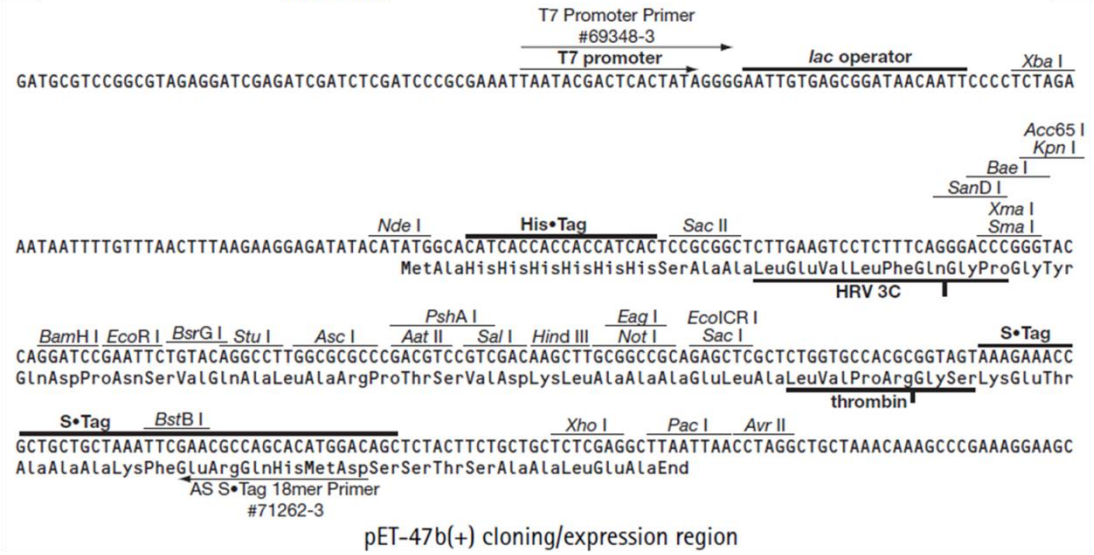
Peptide Information

Calc. Mass	Obsrv. Mass	± da	± ppm	Start Seq.	End Seq.	Sequence	Ion Score	C. I. %	Modification	Rank	Result Type
882.3887	882.3589	-0.0298	-34	431	438	CQAGFSGR			Carbamidomethyl (C)[1]		Mascot
1015.3972	1015.3843	-0.0129	-13	253	259	YCDECIR	30	0	Carbamidomethyl (C)[2,5]		Mascot
1015.3972	1015.3843	-0.0129	-13	253	259	YCDECIR			Carbamidomethyl (C)[2,5]		Mascot
1341.5927	1341.5757	-0.017	-13	420	430	CVDLGDAYLCR	83	100	Carbamidomethyl (C)[1,10]		Mascot
1341.5927	1341.5757	-0.017	-13	420	430	CVDLGDAYLCR			Carbamidomethyl (C)[1,10]		Mascot
1411.5696	1411.55	-0.0196	-14	196	207	DDAFGHFTCGER	101	100	Carbamidomethyl (C)[9]		Mascot
1411.5696	1411.55	-0.0196	-14	196	207	DDAFGHFTCGER			Carbamidomethyl (C)[9]		Mascot
1417.5875	1417.569	-0.0185	-13	393	404	CPVGYSGFNCEK			Carbamidomethyl (C)[1,10]		Mascot
2115.8716	2115.8408	-0.0308	-15	362	380	ICELSAMTCADGPCFNGGR			Carbamidomethyl (C)[2,9,14]		Mascot
2116.8555	2116.8225	-0.033	-16	362	380	ICELSAMTCADGPCFNGGR	78	99.998	Carbamidomethyl (C)[2,9,14], Deamidated (NQ)[16]		Mascot
2116.8555	2116.8225	-0.033	-16	362	380	ICELSAMTCADGPCFNGGR			Carbamidomethyl (C)[2,9,14], Deamidated (NQ)[16]		Mascot
2132.8506	2132.8396	-0.011	-5	362	380	ICELSAMTCADGPCFNGGR			Carbamidomethyl (C)[2,9,14], Deamidated (NQ)[16], Oxidation (M)[7]		Mascot
2438.0112	2437.9646	-0.0466	-19	177	195	FVCDEHYYGEGCSVFCRPR	92	100	Carbamidomethyl (C)[3,12,16]		Mascot
2438.0112	2437.9646	-0.0466	-19	177	195	FVCDEHYYGEGCSVFCRPR			Carbamidomethyl (C)[3,12,16]		Mascot
3211.3047	3211.2551	-0.0496	-15	218	244	GPYCTEPICLPGCDEQHGFCDKPGECK	93	100	Carbamidomethyl (C)[4,9,13,20,26]		Mascot
3211.3047	3211.2551	-0.0496	-15	218	244	GPYCTEPICLPGCDEQHGFCDKPGECK			Carbamidomethyl (C)[4,9,13,20,26]		Mascot

VI.5. Cloning vector maps

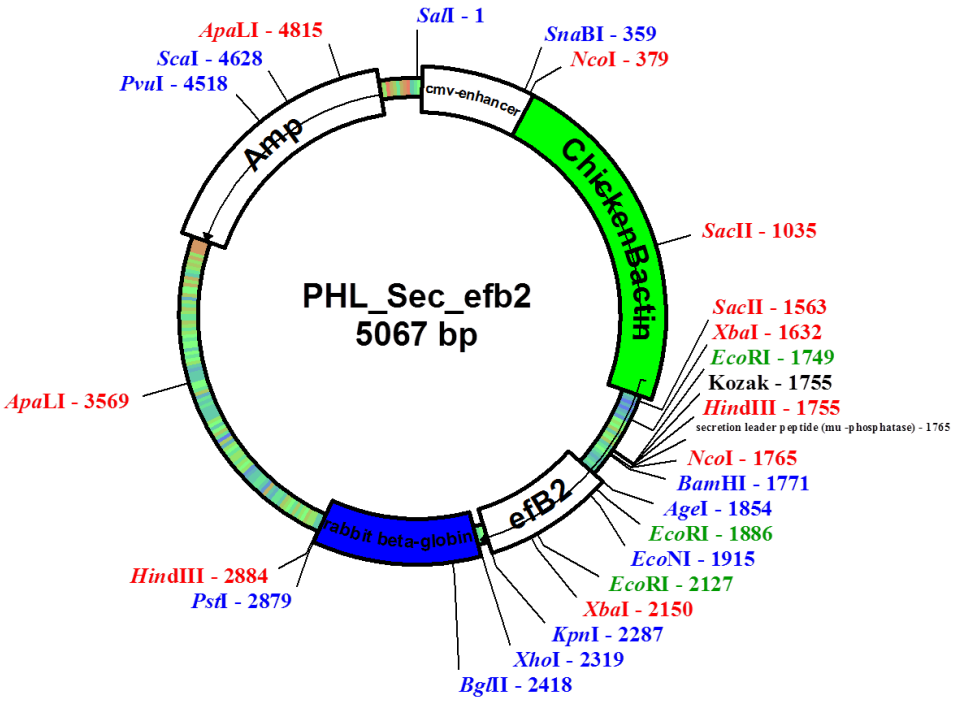
VI.5.1 pET-47b





pET-47b(+) cloning/expression region

VI.5.2 pHL-sec



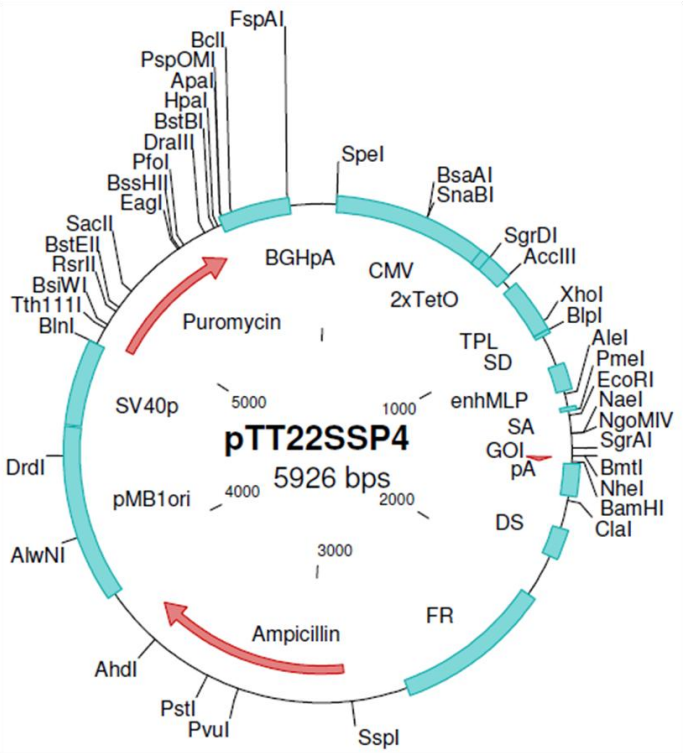

```
EcoRI HindIIIKozak M  G  I  L  P  S  P  G  M  P  A  L  L  S
GAATTCAAGCTTGCCACCATGGGGATCCTTCCCAGCCCTGGGATGCCTGCGCTGCTCTCC

 L  V  S  L  L  S  V  L  L  M  G  C  V  A  E  T  G
CTCGTGAGCCTTCTCTCCGTGCTGCTGATGGGTTGCGTAGCTGAAACCGGT...insert...

 G  T  K  H  H  H  H  H  H  *  *    XhoI
GGTACCAAGCACCACCATCACCATCACTAATGATCACTCGAG
```

NOTE: The vector was supplied containing the 428bp efB2 insert to facilitate cloning. Adapted from Aricescu, Lu *et al.* (2006).

VI.5.3 pTT22SSP4




```
1349  atg gga gaa ctg ctg ctg ctc ctt ctg ctg ggg ctg cgg ctt cag ctg agt ctt gga gcc
                                    signal peptide
        m   g   e   l   l   l   l   l   l   l   g   l   r   l   q   l   s   l   g   a

1409  ggc gct cct ggc tcc tcc acc ggc cac cac cat cac cat cat cac cac ggc tcc acc ggc
                  link                                H8G
        g   a   p   g   s   s   t   g   h   h   h   h   h   h   h   h   g   s   t   g

                                    NheI                            BamHI
                                  --+-----                        --+-----
1469  gag aac ctg tac ttt cag ggc gct agc nnn nnn nnn nnn nnn nnn gga tcc
                  TEV
        e   n   l   y   f   q   g   a   s   ?   ?   ?   ?   ?   ?   g   s
                                    >>........GOI........>>
```