Part A – Applicant

A.1 Applicant

Name student (initials, first name, last name, student number):	*M. (Marinda) Westerveld* *5681898*
Affiliation (university/institute + department):	*Utrecht University / Department of Chemical Biology and Drug Discovery*
Name first examiner:	*Dr. Rita Petracca*
Affiliation (university/institute + department):	*Utrecht University / Department of Chemical Biology and Drug Discovery*
Name second examiner:	*Prof. Eoin Scanlan*
Affiliation (university/institute + department):	*Trinity College Dublin / Department of Chemistry*

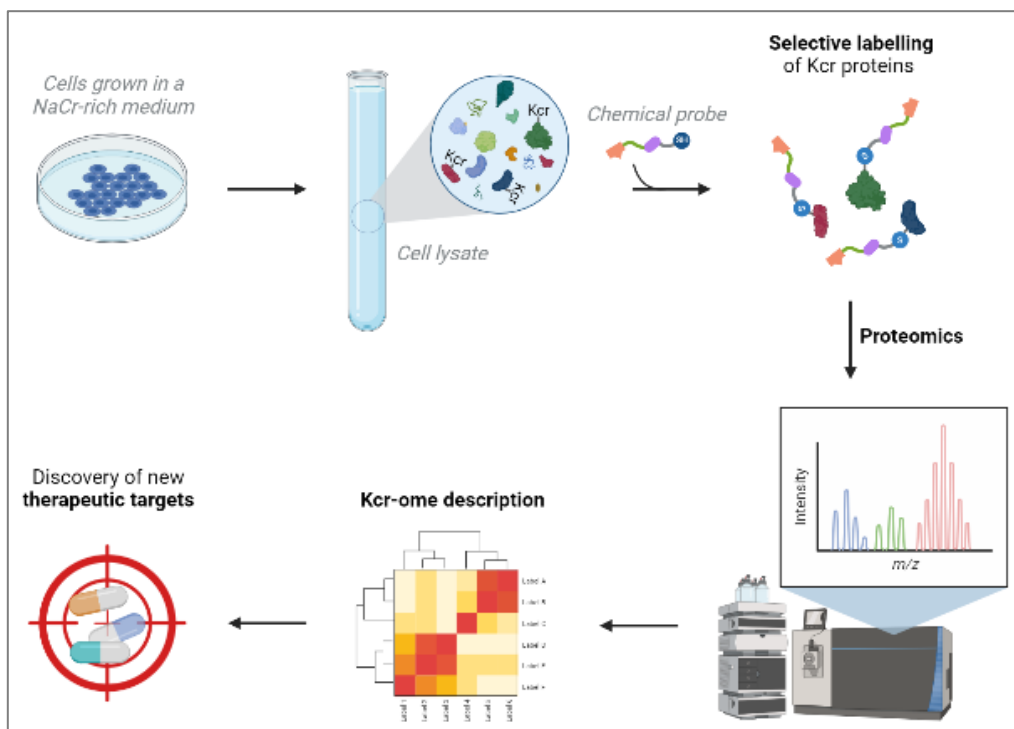## Part B – Scientific proposal

Figure 1 – Graphical abstract of research proposal.

B.1   BASIC DETAILS

B.1.1 Title
Development of an efficient photochemical method to specifically label and detect lysine-crotonylated (K-cr) proteins.

B.1.2 Abstract

Post-translational modifications (PTMs) are essential for cell survival and play an important role in regulating the biology of a cell. Lysine crotonylation (Kcr) is a recently discovered histone and non-histone PTM associated with numerous diseases, including acute kidney injury, depression, HIV latency, cardiovascular disease, and cancer. Despite the strong correlation of Kcr with multiple diseases, no chemical tools for the in-depth study of this PTM are available yet. Kcr contains an electron-rich α,β-unsaturated bond, representing an excellent selective handle for the radical-mediated thiol-ene click reaction (TEC). Here we present the development of a novel TEC-mediated photochemical approach for the identification of therapeutically relevant histone and non-histone Kcr proteins. We will use a chemoproteomics approach to detect and identify Kcr proteins in different cell lines. Initially, a selective chemical probe will be designed and synthesised; then proteomics studies will be performed for the identification and characterisation of Kcr proteins, thus leading to the first accurate description of the lysine crotonylome and aiding a better understanding of the role of crotonylation in molecular processes. Moreover, the probe will be used for the identification of non-histone Kcr proteins in a mouse model of cardiac hypertrophy, leading to the discovery of potential biomarkers or therapeutic targets for curing cardiac disease.

B.1.3 Layman's summary
Many crucial biological processes in cells are regulated by proteins. Small chemical groups can be connected to proteins to give them their shape, stability, and activity, and enable interactions with other cellular molecules. The addition of these groups is called protein post-translational modification (PTM). In 2011 a new PTM was discovered, named crotonylation. Misregulation of crotonylation is associated with numerous diseases, including acute kidney injury, depression, HIV latency, cardiovascular disease, and cancer. Despite the strong correlation between protein crotonylation and disease, there are no selective tools available yet to study crotonylation in-depth. In this project, we will develop an efficient method to discover crotonylated proteins. Initially, a chemical probe will be synthesised to selectively label crotonylated proteins in different cell lines; then the probe will be used to perform proteomics studies, a large-scale study of all proteins produced in an organism or system. The identification and characterisation of all human crotonylated proteins will provide a better understanding of the role of crotonylation in molecular processes. The probe will not only be used in healthy cells, but also in heart cells from a mouse model of cardiac disease. By comparing crotonylation patterns in cells derived from healthy or sick mice, we will identify protein crotonylation misregulation. Ultimately, this will lead to the discovery of potential biomarkers or therapeutic targets for curing cardiac disease.

B.1.4 Keywords
Post-translational modifications (PTM), lysine crotonylation (Kcr), thiol-ene click chemistry (TEC), proteomics.

B.2   SCIENTIFIC PROPOSAL

B.2.1 Research topic

B.2.1.1 Post-translational modifications

Proteins are macromolecules that play many critical roles in the body. The majority of proteins in eukaryotic cells undergo covalent modifications after ribosomal synthesis. These protein post-translational modifications (PTMs) chemically modify existing functional groups on the amino acid side chain or C-/N-terminus of the protein.[1] Due to the functional alterations, PTMs play an important role in protein folding, activity, stability, localization, and the interaction with other cellular molecules.[2] Moreover, PTMs are associated with many diseases, including cancer, Alzheimer's, Parkinson's, and cardiovascular diseases.[1,3] Different PTMs can modify various sites of one protein simultaneously. The combinatorial action of multiple PTMs on the same or different proteins is called PTM crosstalk and it is often responsible for complex phenotypic outcomes.[4,5] The importance of PTMs is underlined by a recent study, reporting 1950 known PTM-disease associations (PDAs) in 749 proteins for 23 types of PTMs and 275 types of diseases.[6] Enhancing or decreasing PTM activity represents an interesting therapeutic strategy. Consequently, many drugs targeting modification enzymes such as kinases and histone deacetylase have been approved.[7] Moreover, the majority of therapeutic proteins display one or more PTMs, greatly influencing the biochemical and therapeutic properties of the drugs.[8] The importance of an in-depth understanding of protein PTMs is underlined by the critical role protein lysine acetylation (Kac) plays in many diseases and in the development of new therapeutics.[5,9]

B.2.1.2 Protein crotonylation

In 2011, Tan *et al.* reported the discovery of a new PTM: lysine crotonylation (Kcr) on histone proteins.[10] Similar to Kac, Kcr is an acylation of the ε-amino group of the lysine side chain.[11,12] For both modifications in the context of histones, the addition of the acyl group neutralizes the positive charge on the lysine side chain, locally weakening the histone-DNA interactions. This results in a less compact chromatin structure, making the DNA more accessible to DNA binding factors, consequently enhancing gene transcription. Structurally, Kcr contains one more carbon-carbon double bond compared to Kac, making the crotonylated lysine residue larger and more hydrophobic than the acetylated form (*Figure 2a*). Moreover, the C-C π-bond in the crotonyl modification results in a rigid and planar conformation, which is unique among acyl modifications. The increase in bulk and hydrophobicity introduced by histone Kcr stimulates gene expression to a larger degree than Kac.[11,12] Following a similar mechanism as for acetylation, the crotonyl group is transferred to the substrate protein from a donor molecule, named crotonyl-CoA (*Figure 2b*).[13] The addition of the crotonyl group is mediated by histone crotonyl transferases (HCTs) and the removal by histone decrotonylases (HDCs).[11] To date, no selective HCTs or HDCs have been identified other than previously characterised histone acetyl transferases (HATs), including p300 and CBP (CREB-binding protein), and histone deacetylases (HDACs).[11]
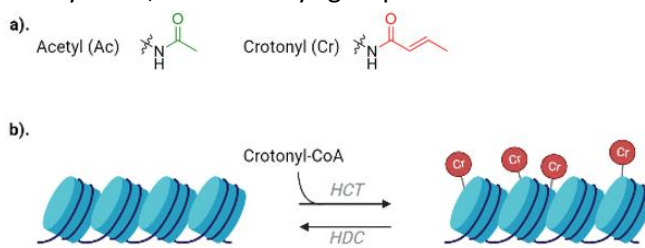

Figure 2 – a). Structure of an acetyl and a crotonyl group, b). Histone crotonylation.
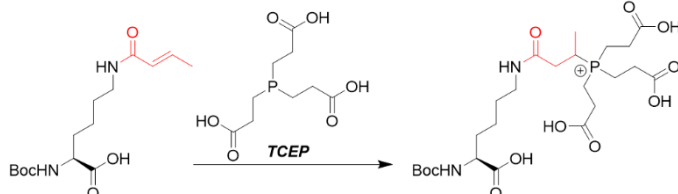
Six years after the discovery of histone Kcr, Xu *et al.* reported lysine crotonylation on non-histone proteins for the first time.[14] Since its discovery, it has been shown that non-histone crotonylation is involved in protein activity, protein localization, protein degradation, and DNA repair. Due to their large impact on protein function and gene expression, both histone and non-histone crotonylation is associated with numerous diseases, including acute kidney injury (AKI), depression, HIV latency, cardiovascular diseases, and cancer.[15,16,17,18] A recent study demonstrated that crotonylation of the GTPase Septin 2 (SEPT2) promotes cell invasion and metastasis in hepatocellular carcinoma (HCC) cells, leading to poor prognosis and a high recurrence rate in HCC patients.[19] Another study also investigated the effect of protein crotonylation on liver cancer and found that Kcr is upregulated in response to hypoxia and promotes liver cancer cell proliferation,

while it prevents cellular senescence.[20] Interestingly, in the same study a direct correlation between treatment with sodium crotonate (NaCr) and tumor size *in vivo* was highlighted. A third study showed that crotonylation of proteins required for cardiomyocyte contractility is triggered in cardiac ischemia reperfusion injury.[21] This study demonstrated that modulating site-specific Kcr of mitochondrial protein IDH3α (isocitrate dehydrogenase 3 [NAD+] alpha) and cytoskeletal protein TPM1 (tropomyosin alpha-1 chain) or enhancing general Kcr *via* NaCr supplementation not only protects cardiomyocyte from apoptosis, but also preserves postinjury myocardial function. Interestingly, not only an upregulation in global Kcr levels, as observed in thyroid, esophageal, pancreatic, and lung cancers, but also a reduction in Kcr levels is associated with diseases, as observed in liver, stomach, and kidney carcinomas, for example.[22] Despite the strong correlation between Kcr and disease, there are no selective biochemical tools available to study Kcr in-depth yet.

B.2.1.3 Current methods to study lysine crotonylation

Immunoaffinity proteomic represents the gold standard when it comes to studying protein modifications.[23] In this context, specific antibody-based protein enrichment is coupled to high resolution mass spectrometry to analyse the desired protein modification. There have been studies attempting to report the human lysine crotonylome (Kcr-ome) using anti-Kcr antibody enrichment.[24,25,26] However, immunoaffinity proteomics is not only high in costs, but it also has limited selectivity. Structural similarities can lead to unselective antibody labelling, which is the case for the commercially available anti-Kcr antibody, as it not only recognizes Kcr proteins, but also butyrylated ones.[25] Moreover, epitope occlusion, describing the situation in which adjacent modifications prevent antibody recognition, also leads to reduced selectivity in immunoaffinity assays.[27] Therefore, the use of chemoselective probes in the protein enrichment step instead of antibodies, could dramatically increase the labelling selectivity.[28]

The first chemical probe for crotonylation was discovered on serendipity by Bos *et al.*, who observed binding of tris(2-carboxyethyl)phosphine (TCEP) to the crotonyl moiety during Kcr protein synthesis (*Figure 3*).[29] They synthesised a biotinylated analogue of TCEP and showed that this TCEP-probe was able to detect crotonylation on a semisynthetic mono-nucleosome containing a single Kcr mark at position 18 on H3 (H3K18Cr) or several Kcr marks on the tail of H4 (H4pCr). The authors hypothesized that the negatively charged carboxylate groups could stabilize the formed phosphonium cation through electrostatic interactions. However, they also speculated that hydrogen bonding interactions or the involvement of a transient cyclic phosphorane species could play a role.[29] Unfortunately, the actual mechanism remains unknown as no further investigation was reported. Despite the promising use of selective chemical probes, phosphine-based probes are not selective for the crotonyl moiety, are synthetically demanding, and have poor chemical tractability.

Figure 3 - Labelling of a crotonyl moiety by TCEP, the first described chemical probe for Kcr.[29]

In this project we will develop a novel chemical probe for the discovery of histone and non-histone Kcr proteins. The probe will be based on a thiol-ene click chemistry (TEC)-mediated approach to react selectively with the unsaturated C-C bond of the crotonyl moiety. First, we will optimise the reaction conditions for the thiol-ene reaction (objective 1) using model crotonylated peptide substrates and commonly used thiols, such as cysteine and glutathione. After finding the optimal conditions, we will design, synthesise, and test our chemical probe (objective 2). We will perform proteomics studies with our novel chemical probe, which will lead to the first accurate description of the Kcr-ome (objective 3). Finally, we will use our probe for the identification of non-histone Kcr proteins in a mouse model of cardiac hypertrophy (objective 4). Together this will provide a better understanding of the role of crotonylation in molecular processes, leading to a comprehensive description of the function of protein Kcr. Moreover, the characterisation and analysis of the function of Kcr proteins in a mouse model of cardiac hypertrophy will lead to the identification of new therapeutically relevant targets. If interesting targets will be identified, a drug discovery program will be initiated, ultimately leading to the development of cardioprotective drugs with a novel mechanism of action (see *Figure 4* for objectives overview).
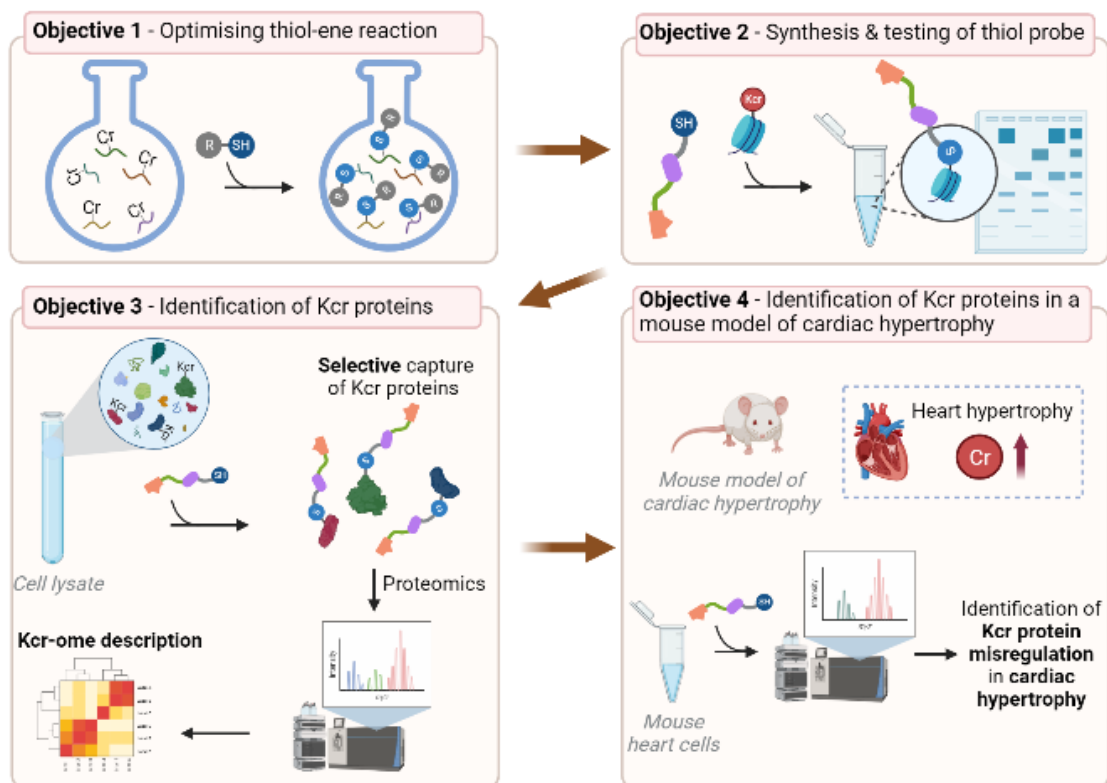
Figure 4 – Overview of the proposed research

B.2.1.4 Thiol-ene click chemistry

Click reactions are extensively exploited for biochemical and in-cell applications, as they are high yielding with (almost) no by-products, are regio- and stereospecific, insensitive to oxygen or water, and proceed under mild, solventless or aqueous reaction conditions.[30] A promising click reaction for the labelling of crotonyl moieties is the reaction of thiols with reactive carbon-carbon double bonds ("enes"), named thiol-ene click (TEC).[30] TEC was discovered over a hundred years ago and has many applications in synthetic chemistry.[31,32,33] The thiol-ene coupling reaction can proceed through two main mechanisms: the catalysed Michael addition and the radical-mediated thiol-ene addition.[30] The Michael addition proceeds by an anionic chain, starting with the formation of a thiolate by the abstraction of a hydrogen from a thiol in the presence of a nucleophile (*Figure 5a*).[34] The thiolate anion can react with an electron deficient carbon-carbon double bond in an anti-Markovnikov addition in which the thiolate anion adds to the least substituted carbon atom, forming a carbon-centred anionic intermediate. Abstraction of a hydrogen yields the thioether product. The Michael addition reaction only proceeds between a thiol and an electron deficient carbon-carbon double bond. The terminal methyl of a crotonyl moiety is electron donating, leading to electron enrichment of the carbon-carbon double bond, making crotonyl unreactive as Michael acceptor.

The radical-mediated thiol-ene reaction on the other hand is most efficient with electron rich carbon-carbon double bonds.[30] The reaction can be initiated by light, heat, and/or radical initiators, forming a thiyl radical species (*Figure 5b*). Thiyl radicals are easily formed due to their low bond dissociation energies (87 kcal mol^{-1}), however photochemical radical initiators are often required for the initial radical formation.[35] The thiyl radical can propagate with a carbon-carbon double bond to form a carbon-centred radical *via* an anti-Markovnikov addition. The carbon-centred radical can abstract a hydrogen from another thiol, forming the thioether product and a new thiyl radical, which can subsequently participate in another propagation step. Any non-sterically hindered terminal -ene is able to participate in radical-mediated TEC, with strained or electron rich- enes reacting more rapidly compared to electron deficient ones.[30] The electron rich crotonyl carbon-carbon double bond makes crotonyl groups an excellent substrate for the radical-mediated thiol-ene reaction.
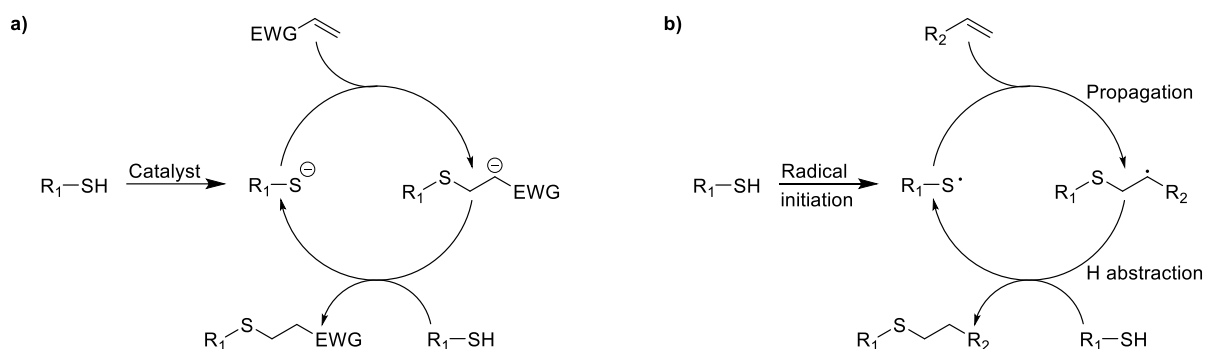
Figure 5 – Overview of the two main mechanisms of TEC: a). Catalysed Michael addition, b). Radical-mediated thiol-ene addition.

The radical-mediated thiol-ene reaction has click chemistry characteristics as it proceeds rapidly under mild conditions, in the presence of water under ambient atmosphere, giving a single regioselective product with almost quantitative yields.[30] The use of single electron radical chemistry provides superior selectivity in the presence of unprotected amino acid functional groups, including the presence of both nucleophilic and electrophilic groups prone to side reactions in non-radical reactions, making this reaction highly compatible with peptides and proteins.[32]

A potential drawback of the radical-mediated thiol-ene reaction in bioconjugation is the need of a photoinitiator. Photoinitiators can be subdivided into two categories: type I and type II.[36] Irradiation of type I photoinitiators with UV light results in homolytic bond cleavage, generating two highly reactive radical species (*Figure 6a*). Irradiation of type II photoinitiators does not lead to the direct formation of two radical species, they require a hydrogen donor, like alcohols or amines, to react. UV irradiation induces an excited state in the type II photoinitiator, after which the photoinitiator abstracts a hydrogen atom from the hydrogen donor, leading to the formation of two radicals (*Figure 6b*). In addition to UV-activated free-radical initiators, visible light and thermally activated radical initiators also exist.[37,38] All these initiators have different chemical properties, different limitations, and different efficiencies. This wide variety provides the possibility to test for and select the optimal free radical initiator for the development of a selective photochemical method to target Kcr.

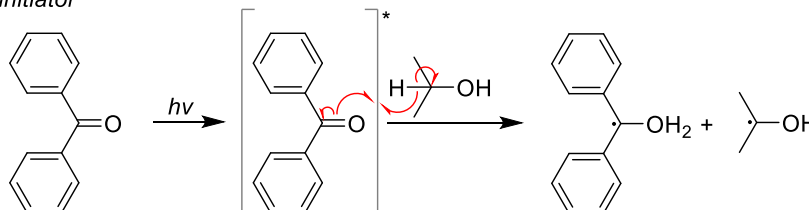**a)** *Type I photoinitiator*
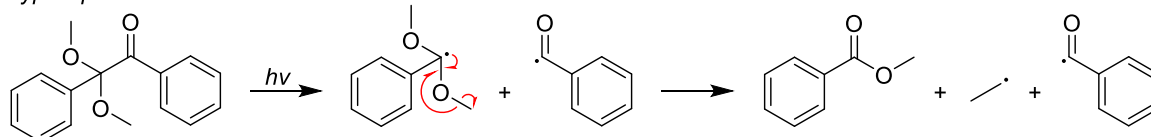
b) *Type II photoinitiator*

Figure 6 – Mechanism of radical formation after UV irradiation of a). 2,2-dimethoxy-2phenylacetophenone (DPAP), a type I photoinitiator and b). benzophenone (BP), a type II photoinitiator.

To confirm the suitability of TEC for the identification of Kcr proteins, preliminary experiments were performed (*Table 1*). The radical mediated thiol addition was tested employing ε-NH-Kcr as model substrate and unprotected cysteine as thiol source. 2,2-dimethoxy-2-phenylacetophenone (DPAP) was used as radical initiator. The reaction was monitored by liquid chromatography mass spectrometry (LC-MS). UV irradiation (365 nm) resulted in full conversion of ε-NH-Kcr to the desired thioether ligated product within half an hour (*Table 1, entry 1*). Longer reaction times did not change the outcome of the reaction, nor did it lead to the formation of side products (*Table 1, entry 2*). Leaving the reaction in the dark for 2 hours did not result in any

product formation, indicating that the crotonyl moiety and thiol have no reactivity towards each other on its own (*Table 1, entry 3*). In the absence of a photoinitiator also no product formation was observed after 2 hours, showing the necessity of this to initiate the radical reaction (*Table 1, entry 4*). Overall, these results show that TEC is a promising reaction for selectively labelling crotonyl modified proteins.
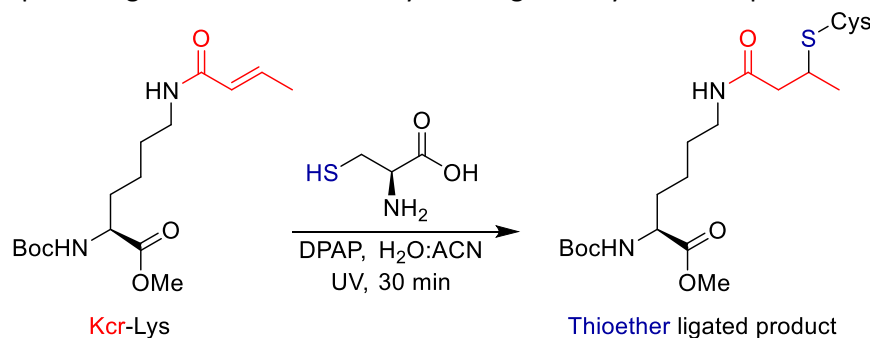
Entry	DPAP	Irradiation	Time	Conversion*
1	1%	365 nm	30 min	100%
2	1%	365 nm	2 h	100%
3	1%	-	2 h	-
4	-	365 nm	2 h	-

*Table 1 – Preliminary experiments of TEC reaction employing ε-NH-Kcr (1 equiv.) as model substrate, cysteine (4 equiv.) as thiol source and DPAP (0.01 equiv.) as photoinitiator. *Conversion calculated based on LC-MS*

B.2.1.5 Design of our novel chemical probe

Inspired by the promising preliminary results, a TEC-mediated photochemical ligation strategy will be exploited for the selective labelling of crotonyl modified proteins. A thiol containing chemical probe will be designed and used for proteomics studies on cell lysates to identify histone and non-histone Kcr proteins. Upon irradiation, the thiol of the chemical probe will selectively react with the double bond of Kcr proteins in the presence of a photoinitiator. After labelling, the chemical probe will enable the capturing, detection, and characterisation of the crotonylated proteins, leading to the first accurate description of the Kcr-ome.

Mass spectrometry, which provides a valuable analysis method to investigated crotonylated proteins; the read-out is sensitive and fast, the required sample amount is very low, and it allows for the identification of the precise protein modification site.[39] Protein detection yield and sensitivity are constantly improving by the development of new fragmentation methods and identification software.[40,41] Tandem MS is often used in the analysis of biomolecules such as proteins and peptides. Generally, tandem MS is performed on the most abundant species in the full-scan mass spectra, consequently excluding low abundant species. However, for the identification and characterisation of histone and non-histone Kcr proteins it is important that all crotonylated proteins can be detected. The use of computational pattern-recognition has the potential to increase detection of low abundance species. Bertozzi *et al.* found that a dibromide motif could be detected using a specific pattern-recognition algorithm (IsoStamp), as a result of the natural abundance of the stable isotopes, [79]Br and [81]Br (1:1).[42] By designing a selective chemical probe bearing a dibromide motive, only crotonylated peptides displaying the specific isotopic signature will be computationally detected by the algorithm. Based on the detected peptides an inclusion list can be made. In a follow up tandem MS experiment only fragmentation of the peptide ions on this inclusion list will be selected, allowing for the selection of all crotonylated peptides, including the low abundance ones.

In addition to a dibromide motif, it is important that a chemical probe contains a moiety that enables enrichment and purification of the labelled proteins. A biotin tag is commonly used for enrichment on immobilized streptavidin.[43] The biotin-streptavidin interaction is very strong, consequently requiring harsh conditions for protein elution. These harsh conditions not only break the biotin-streptavidin interaction, but also elute non-specifically bound proteins and might modify the capped proteins. To avoid harsh cleaving conditions, a cleavable linker will be incorporated between the reactive handle and biotin. Such a cleavable linker should be cleaved in conditions mild enough to avoid the release of streptavidin or proteins non-

specifically bound to the resin. Moreover, cleaving should not modify proteins in any way and allow for efficient protein recovery compatible with downstream analytical procedures. Tirrel *et al.* tested five cleavable probes and found a probe constructed around a dialkoxydiphenylsilane linker to be most promising for use in biomolecular labelling and proteomics studies.[43] Cleavage of the silane linker can be achieved in mild conditions by treatment with 10% formic acid for 0.5 h, leaving only a small mass tag (143 Da) on the labelled protein. Bertozzi *et al.* used the silane cleavable linker and biotin tag in combination with the dibromide motif in the design of a chemical probe to describe the isotope-targeted glycoproteomics (IsoTag).[44]

Altogether, our novel chemical probe will consist of four moieties: a thiol group as reactive handle for the selective reaction with the crotonyl moiety *via* TEC, a dibromide motif to enhance the sensitivity of the proteomics method, a silane cleavable linker to enable elution of the proteins under mild conditions and to improve the analysis of the MS data, and a biotin tag for enrichment on Streptavidin. These features will ensure optimal detection and characterisation of Kcr proteins. The chemical probe will be used for proteomic studies on cell lysates to identify histone and non-histone Kcr proteins, leading to the first accurate description of the Kcr-ome. An in-depth understanding of the role of specific proteins is expected to enhance the understanding of the impact of crotonylation on these cellular processes as well as the discovery of novel therapeutically relevant targets.

B.2.1.6 Identification of non-histone Kcr proteins in a mouse model of cardiac hypertrophy
In addition to the identification of Kcr proteins in cell lysates for the description of the Kcr-ome, we want to identify non-histone Kcr proteins in a mouse model of cardiac hypertrophy. Cardiac hypertrophy is the abnormal thickening of the heart muscle, often due to hypertension or valvular heart disease.[45] Applying our photochemical approach to this medically relevant system will provide a more in-depth understanding in the function of Kcr in cardiac disease and allows for the identification of Kcr proteins potentially involved in cardiovascular disease generation and progression.

Dr. Burgoyne, our collaborator at King's College London (KCL), and his team developed a mouse model of cardiac hypertrophy.[46] They performed transverse aortic constriction (TAC) surgery in adult mice to model pressure overload and subsequent progression to heart failure. As a control, the same surgical procedure was performed on a second group of mice, but without the constructive suture (Sham mice). 7 days after surgery, hearts from TAC and Sham mice were harvested and analysed. The levels of H3Kcr were found to be significantly higher in TAC mice compared to Shams, suggesting that protein crotonylation is involved in cardiac hypertrophy. To gain more insights in the signalling processes regulating H3Kcr, the Burgoyne lab performed studies using neonatal rat ventricular myocytes (NRVMs). NRVMs have proven to be an ideal research model for cardiac disease.[47] The Burgoyne lab induced cardiomyocyte hypertrophy by treating the NRVMs with angiotensin II (AngII), phenylephrine (PE) or endothelin-1 (ET-1) which are known to stimulates a cardiac hypertrophic response. They found that H3Kcr is mainly enhanced at K18 (*Unpublished results*). Together this indicates a strong correlation between H3 crotonylation and cardiac disease.

Our selective photochemical ligation method will be used to study the function of non-histone Kcr in this mouse model of cardiac hypertrophy, providing a better understanding in the role of specific proteins and the impact of crotonylation on these cellular processes, ultimately, leading to the discovery of novel therapeutically relevant targets for cardiovascular disease.

B.2.2 Approach

The herein proposed research workplan is divided into four objectives. We will start by optimising the reaction conditions for the thiol-ene click (underline objective 1), which is expected to take one year performed by a PhD student. Afterwards we will design and test our novel chemical probe (underline objective 2), which will take up to two PhD years. In this context, to minimize the risks of the project, we will investigate an additional strategy, performed in parallel by a post-doctoral researcher. Once we obtain our optimal chemical probe, we will perform proteomics studies to identify histone and non-histone Kcr proteins during an expected period of one year (underline objective 3). Finally, we will use our novel probe for the identification of non-histone Kcr proteins in a mouse model of cardiac hypertrophy (underline objective 4). This research will be performed in the fourth year by the PhD and the post-doctoral researcher.

B.2.2.1 Optimisation of the thiol-ene reaction
In the first part of this project, we will optimise the reaction conditions of the thiol-ene reaction. In preliminary studies, we successfully reacted ε-NH-Kcr with cysteine *via* a radical-mediated thiol-ene mechanism in the presence of DPAP (1%) as radical initiator and half an hour of UV irradiation (365 nm). These conditions led to full conversion into the desired thioether ligated product and will therefore be used to test different alkene and thiol substrates. Short histone 3 (H3) Kcr peptides, known to be crotonylated *in vivo*, will be synthesised and used as alkene substrates (*Figure 7a*).[10] Fmoc-Cr-K will be chemically synthesised and used as building block in Fmoc solid phase peptide synthesis (SPPS) for the synthesis of the different H3 Kcr substrates. Cysteine, glutathione (GSH), and thiol alkanes of different lengths will be used as thiol source (*Figure 7b*). We will test and select the highest yielding substrates. Moreover, other factors influencing the ligation reaction will be investigated, such as the steric effects of the amino acids adjacent to Kcr and the effect of more than one crotonylated site on the peptide.

Initially DPAP will be used as radical initiator as it gives the highest yields in TEC. However, DPAP is not water soluble and, as all type I radical initiators, it generates reactive benzoyl radicals during bond cleavage that can react to the double bond, making DPAP less suitable for TEC in biological settings. Therefore, we aim to investigate the effect of different photoinitiators. We will test water soluble photoinitiators such as 2,2'-azobis[2-(2-imidazolin-2-yl)propane]-dihydrochloride (VA044) or lithium phenyl-2,4,6-trimethylbenzoylphosphinate (LAP) (*Figure 7c*). To avoid radical cross reactions, type II photoinitiators, such as benzophenone (BP) will also be tested (*Figure 7c*). Finally, photoinitiators activated by visible light, such as Eosin Y, will be investigated (*Figure 7c*). With these different alkene substrates, thiol sources, and photoinitiators we aim to find optimal conditions for TEC in which the ligated thioether product is obtained in high yield, with the minimal use of a water soluble photoinitiator and short irradiation times.
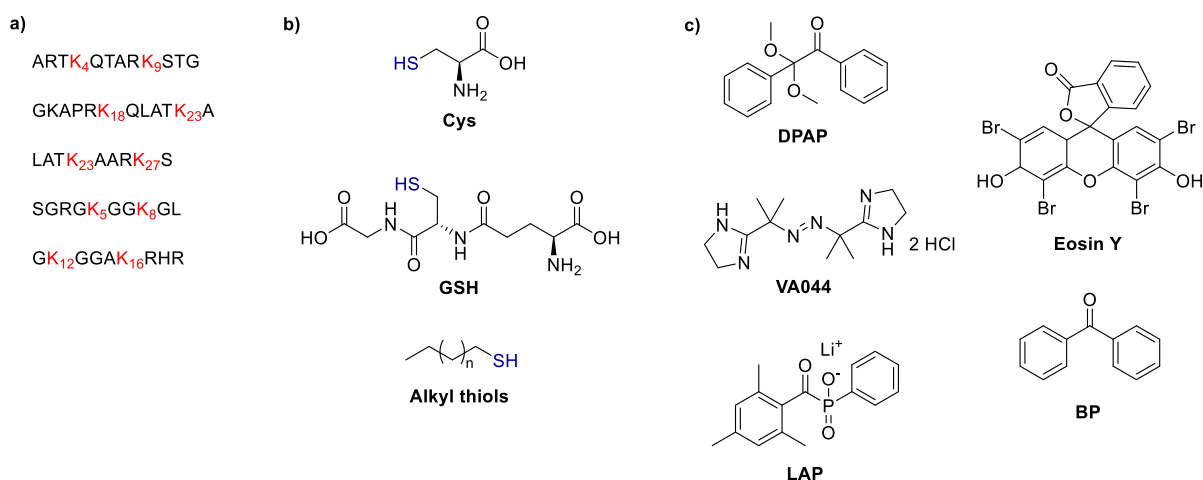

Figure 7 – Compounds for the optimisation of the thiol-ene reaction. a). Alkene substrates: H3 Kcr peptide sequences, b). Thiol substrates, c). Different photoinitiators.

After optimising the thiol-ene reaction with short peptides, we will perform the ligation with Kcr proteins. Recombinant human H3K27cr nucleosomes (dNuc) will be purchased from EpyCypher. The same thiol sources, cysteine, GSH, and thiol alkanes of different lengths, will be used (*Figure 7b*). We will investigate the effect of UV irradiation and photoinitiator on the protein and optimise pH, salt/detergent concentration in the protein buffer and the reaction time. The ligation reaction will be monitored and analysed by three different methods:

- High resolution MS will be used to confirm the formation of the desired thioether. Peptide mapping will be used investigate the Kcr modification site.
- Western Blotting (WB) analysis using anit-H3K27cr recombinant antibody (ThermoFisher, Rabbit Monoclonal Antibody (RM401)) will indicate the formation of the thioether product as the formation of the thioether product blocks protein recognition by the antibody.
- In-gel fluorescence will be used to check the TEC reaction with rhodamine-PEG-thiol (Nanocs, PG2-RBTH-3k)

With this first part of the project, we establish the optimal experimental conditions (i.e. type and concentration of photoinitiator, thiol concentration, buffer composition, pH, salt/detergent concentration, and irradiation time) for TEC using a recombinant Kcr protein.

B.2.2.2 Synthesis and testing of our novel selective chemical probes

After optimisation of the thiol-ene reaction, thiol-containing chemical probes will be synthesised and tested on cell lysates for their ability to selectively capture Kcr proteins. The first probe we aim to synthesise is a **b**iotin-**c**leavable **di-br**ominated **thiol** probe, named **Bc2Br-SH** (**1**, *Figure 8*). The design of our probe is based on the probe Bertozzi *et al.* reported while describing the isotope targeted glycoproteomics (IsoTaG).[44] Their probe contains a biotin moiety, spacer, silane cleavage site, dibromo motif, and an alkyne as a *click* handle (**2**, *Figure 8*). The described probe was synthesised from three building blocks: biotin-PEG$_4$ tertiary alcohol (**4**, *Figure 8*), dichlorodiphenylsilane as silane cleavage site (**5**, *Figure 8*), and the dibromo alcohol motif conjugated to the alkyne group (**6**, *Figure 8*). Instead of an alkyne as *click* handle, our Bc2Br-SH probe will contain an alkyl thiol. Synthetic procedures for the introduction of a thiol group directly next to the dibromo motif are not published and seem to be challenging as the nucleophilic nature of the thiol group will probably lead to side reactions. Instead of replacing the alkyne handle with a thiol group, we will use the alkyne to introduce thiol-PEG$_3$-azide (**3**, *Figure 8*) *via* a copper catalysed click reaction.[43,44] In the first objective, thiol alkanes with different chain lengths will be tested, based on these results, we will consider different chain lengths for the thiol-PEG linker. If necessary, we will synthesise thiol-PEG$_n$-azide with a specific linker length.
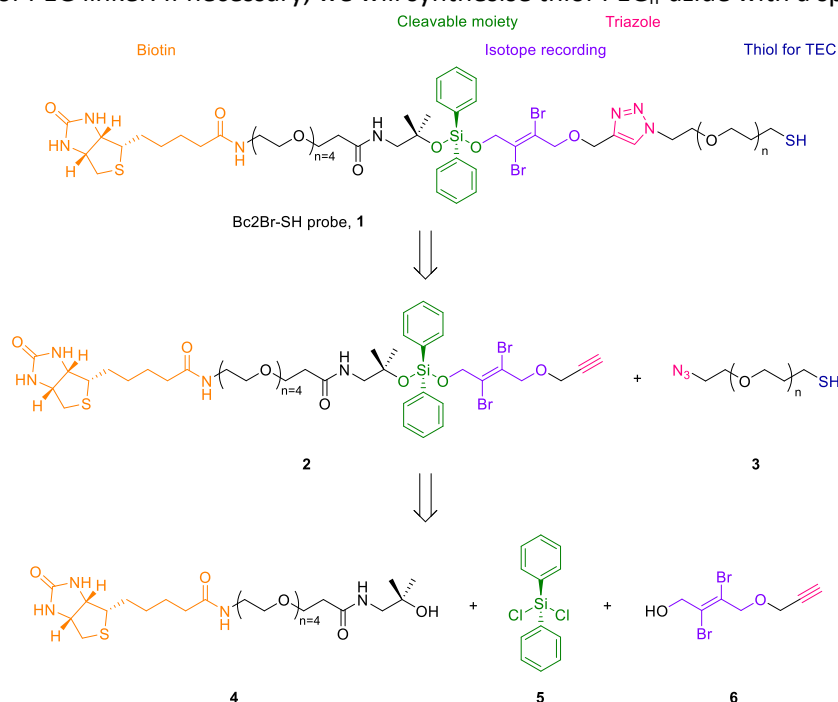

Figure 8 – Retrosynthesis of novel Bc2Br-SH probe (1).

In addition to the Bc2Br-SH probe, appropriate controls will be synthesised. A simple biotin thiol will be synthesised and tested in the thiol-ene ligation reaction. Bc2Br-SH either without the silane cleavage site or the dibromo motif will be synthesised to test the advantages of these moieties in the proteomics workflow and protein purification. Finally, a Bc2Br-SH version without the terminal thiol will be synthesised and used as a negative control in the labelling experiments. The thiol-ene reaction will be tested with H3K27cr using the experimental conditions optimised in the first objective. The general workflow will be as following: a stock solution of Bc2Br-SH and photoinitiator will be added to the protein sample. UV irradiation will initiate the radical-mediated thiol-ene reaction. The reaction will be analysed by high resolution MS to confirm the formation of the desired thioether product. In addition, the samples will be run on a polyacrylamide gel. WB analysis will be performed using Neutravidin-HRP to visualize the ligated proteins.

To ensure efficient identification of Kcr proteins and minimise the risks of this research plan, a second probe will be synthesised and tested in parallel. For this probe, TEC will be performed in solution between a biotin thiol and crotonyl-CoA, forming a **cro**tonyl-CoA **thio**ether **b**iotin probe, named **CroSB** (*Figure 9*). CroSB has a different mode of action compared to Bc2Br-SH and has great potential in case TEC shows low labelling efficiency *in vitro*. The crotonyl-CoA moiety of CroSB will enable metabolic incorporation of the probe in cells. Proteins normally crotonylated within the cell will instead incorporate a biotin tag. The advantage of this one-step selective exoenzymatic labelling process is that no side reaction or other limitations associated with the bio-orthogonal TEC reaction will occur, potentially increasing the labelling yield. A potential limitation of CroSB is that the crotonyl transferase activity of the *writer* enzymes needs to be retained to enable metabolic incorporation. We will test HCT activity using a recombinant HCT. CroSB is not only a substrate for crotonyl transferases but could also be a substrate for acetyl transferases. To guarantee identification of Kcr proteins, we will use a p300 mutant with a specific mutation (I1395G). This mutation leads to deactivation of histone acetyl transferase activity, while maintaining HCT activity.[48] Modifying the isoleucine residue 1395 changes the size of the binding pocket, enabling the accommodation of acyl-CoA bearing longer alkyl chains than acetyl. Therefore, we expect this p300 mutant to be able to catalyse the transfer of the CroSB acyl group. We will test the ability of the p300 mutant to transfer the acyl group of CroSB using an *in vitro* HCT assay. Once confirmed, HEK293 cells will be transfected with the plasmid corresponding to the recombinant p300 mutant and treated with the CroSB probe. The corresponding cell lysates will be analysed by WB using a Neutravidin-HRP antibody.
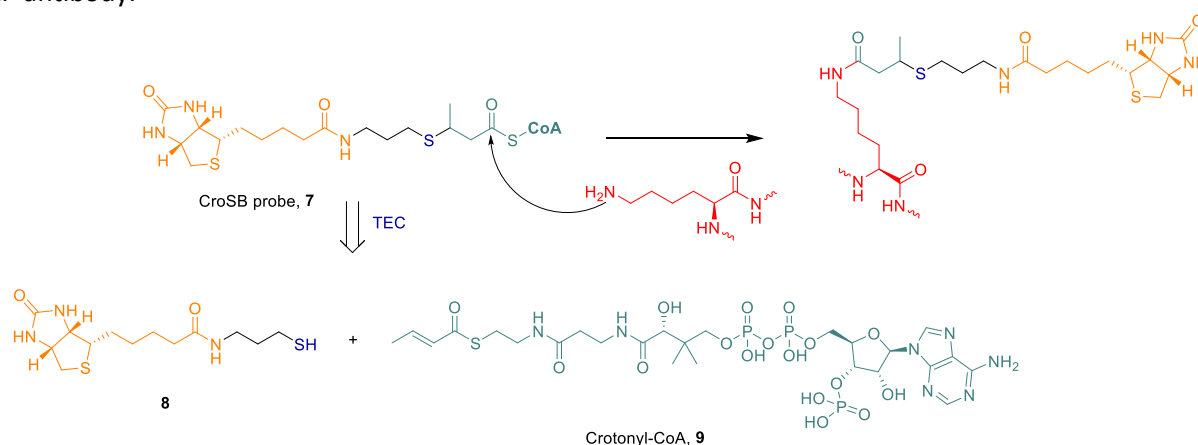

*Figure 9 – Retrosynthesis of CroSB probe (**7**) and mechanism of incorporation.*

In this part of the project, efficient biotinylated thiol-containing probes are synthesised. Their ability to selectively capture Kcr proteins in cell lysates will be analysed, providing us with the optimal probe for the identification of Kcr proteins.

B.2.2.3 Proteomics studies for the description of the Kcr-ome

Proteomics studies will be performed using the newly synthesised chemical probe for the first accurate description of the Kcr-ome. These studies will be conducted in collaboration with Prof. A. Heck, who is a world leader in identification of protein phosphorylation.[49] The initial experiments will be performed with cell lysates from HEK293 cells grown in the presence of sodium crotonate (NaCr). Addition of NaCr is performed as it has been shown to increase the total level of protein crotonylation.[50] The optimal conditions tested in the previous workplan will be used to label Kcr proteins in freshly lysed cells with the Bc2Br-SH probe. To maximize labelling efficiency, we will carefully consider the concentration and incubation time of the probe and photoinitiator with the cell lysates. The TEC-mediated ligation reaction will be analysed by WB using Neutravidin-HRP and anti-crotonyl antibodies. Labelling experiments with Bc2Br-SH will be performed in three biological replicates.

The general proteomics workflow (*Figure 10*) starts by incubating freshly obtained cell lysates with a stock solution of Bc2Br-SH and the photoinitiator. The sample will be mixed shortly and afterwards UV irradiated. After the ligation reaction, the excess of the probe will be removed by precipitation and the proteins will be resuspended in buffer. Streptavidin Sepharose affinity chromatography will be performed to capture the labelled biotin tagged Kcr proteins from the cell lysates. Peptides that are not covalently conjugated to Bc2Br-SH will be removed by on-bead digestion with trypsin. The captured Kcr proteins are then released upon cleavage of the silane linker. As a final step before analysis, the trypsin digest and cleavage fraction will be concentrated and desalted. The recovered peptides will be separated by reversed-phase liquid chromatography and analysed by mass spectrometry. The dibromide motif of the probe will enable us to process the full scan MS spectra using IsoStamp, allowing us to build an inclusion list of isotope recorded peptides.[42] This inclusion list will be used in a follow up tandem MS experiment run in data dependent acquisition (DDA) mode. MaxQuant or Byonic will be used to perform DDA data processing and quantification.[41] In case neutral losses or signature fragments are consistently observed during DDA by relatively quick higher-energy collisional dissociation (HCD), then these could be used as a mass trigger for more in-depth dissociation methods such as stepping-energy HCD or electron-transfer dissociation/HCD hybrid methodologies. This would result in a higher number of specific fragmentation events of Kcr sites, which ultimately leads to better peptide sequence coverage and Kcr localization.
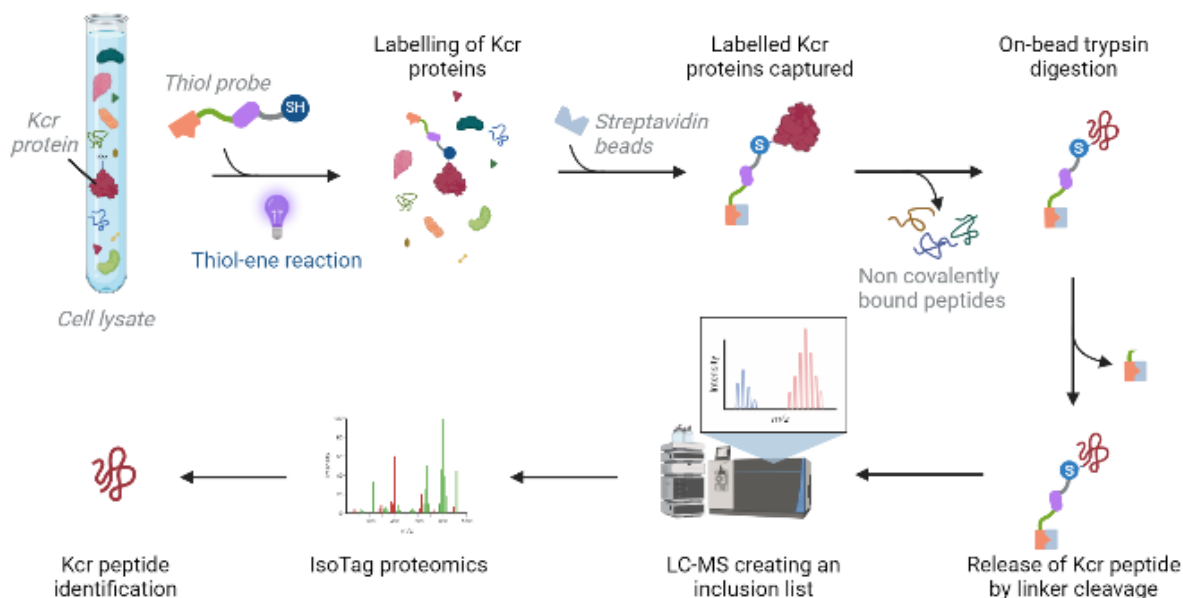

Figure 10 – General proteomics workflow.

The proteomics workflow will be optimised and adjusted in case problems are encountered. Solubility problems may arise during proteomics experiments due to the hydrophobic nature of the Kcr moiety. Increasing the acetonitrile concentration potentially overcomes this problem. We will pay attention to the possible cross reactivity of the chemical probe with unsaturated fatty acids. Despite the preferred addition *via* TEC to terminal alkenes, it is known that thiols can react with unsaturated C-C bonds *via* a radical-catalysed addition.[51] Peptides bearing long alkyl chain fatty acids are expected to precipitate and are

therefore not expected to interfere with Kcr proteins identification. We will consider adjustments in the experimental procedure when treating the cell lysates or perform a specific MS search for the identification and exclusion of peptides with long alkyl chains, to exclude this cross reactivity entirely.

Proteomics studies with our novel chemical probe will lead to the first accurate description of the Kcr-ome, providing a better understanding of the role of crotonylation in molecular processes.

B.2.2.4 Identification of non-histone Kcr proteins in a mouse model of cardiac hypertrophy

The newly developed photochemical method for the capture and identification of Kcr proteins will be applied to medically relevant biological samples. Dr. Burgoyne and his team will provide samples derived from mice with cardiac disease.[46] Our first aim is to validate the efficacy of TEC by replicating the data obtained by the Burgoyne group. Method validation will first be performed following the established proteomics workflow using neonatal rat ventricular myocytes (NRVMs). After validation of the efficacy of TEC and method optimalisation, we will use our chemical probe on tissue samples derived from both TAC and Sham operated mice, following the same general workflow. Samples will be incubated with a stock solution of the Bc2Br-SH probe and photoinitiator, UV irradiated, and precipitated. The labelled Kcr proteins will be captured by Streptavidin Sepharose affinity chromatography. We will obtain all non-conjugated peptides by on-bead trypsin digestion and release all labelled peptides by cleaving of the silane linker. The obtained peptides will be analysed by MS following the same methods as described for the proteomics workflow in section 2.2.3. This will enable us to evaluate the differences in crotonylation levels for non-histone proteins in TAC and Sham mice derived samples and consequently to identify Kcr misregulation.

By applying our TEC-mediated photochemical ligation approach to samples derived from a mouse model of cardiac hypertrophy, we will identify non-histone Kcr proteins in a medically relevant setting. This will provide a more in-depth understanding of the function of Kcr in cardiac disease and lead to the identification of new therapeutic targets.

B.2.3 Feasibility / Risk assessment

The leading researcher on this project, Dr. Rita Petracca, is highly experienced with organic and peptide chemistry, and protein labelling techniques. In addition, she is an expert on performing TEC on small molecules and short peptides. The preliminary studies showed promising results. Together, the risks for <u>objective 1</u> are low. In the unlikely event that TEC does not give the expected high yields, more photoinitiators and reaction conditions will be evaluated. In case the thiol-ene reaction is not efficient with recombinant human H3K27cr, TEC will be performed using different commercially available recombinant Kcr proteins, such as H3K9cr or H3K18cr. The help of Master/Bachelor students will speed up the search for the optimal reaction conditions.

Low efficiency of TEC in cell lysates could represent a problem when using our Bc2Br-SH probe, caused by the increased biological complexity or the crotonyl moieties might not be readily available. As a fallback strategy for <u>objective 2</u>, we develop in parallel a different approach based on metabolic incorporation of the probe. In case cross reactivity of the thiol probe with double bonds of fatty acids occurs, we will consider adjustments in the experimental procedure when treating the cell lysates or perform a specific MS search for the identification and exclusion of peptides with long alkyl chains to prevent this cross reactivity. The proteomics studies for <u>objective 3</u> will be performed in collaboration with the Heck group to assure the successful identification of Kcr proteins. For <u>objective 4</u> we will work with tissues from an established mouse model. Mice tissues will be provided by Dr. Burgoyne and his team, who will provide us with additional knowledge and help when working with the cardiac tissues. In the unlikely event that TEC will not work with the mice tissues, we will only use neonatal rat ventricular myocytes (NRVMs) for the investigations.

B.2.4 Scientific and societal impact

PTMs are essential for cell survival and play an important role in regulating the biology of a cell. Over the last decade major advances have been made in the field of proteomics. Despite, there remains a need for new selective methods to understand the biological mechanism of PTMs. Here we aim to develop a selective photochemical conjugation method for the identification of Kcr, a recently discovered PTM. Our novel photochemical approach provides a highly selective alternative to the current crotonylation labelling methods and will be added to the existing chemical biology toolbox. With our probe it will be possible to distinguish crotonylation from similar modifications like acetylation. Proteomics studies will uncover the role of crotonylation in molecular processes and lead to the first accurate description of the Kcr-ome. When applied to biologically relevant samples, our novel probe will lead to the identification of valuable targets for the discovery of new therapeutics for different diseases. The identification of proteins showing Kcr misregulation could be essential for an in-depth understanding of cardiac hypertrophy generation and heart failure progression. These protein targets identified with our photochemical ligation method represent potential biomarkers or therapeutic targets for curing cardiac disease.

B.2.5 Ethical considerations

This project involves the use of HEK293 (CRL-1573) cells, human cells commercially available from the American Type Culture Collection (ATCC). Unfortunately, it is not possible to gain full understanding of the role of crotonylation in cardiac hypertrophy (objective 4) only using cell systems. To be closest to physiological conditions, we will use mouse tissues provided by Dr. Burgoyne, our collaborator in King's College London (KCL).

B.2.6 References

1. Chou, K. C. Progresses in Predicting Post-translational Modification. *Int. J. Pept. Res. Ther.* **26**, 873–888 (2020).
2. Duan, G. & Walther, D. The Roles of Post-translational Modifications in the Context of Protein Interaction Networks. *PLoS Comput. Biol.* **11**, 1–23 (2015).
3. Deribe, Y. L., Pawson, T. & Dikic, I. Post-translational modifications in signal integration. *Nat. Struct. Mol. Biol.* **17**, 666–672 (2010).
4. Leutert, M., Entwisle, S. W. & Villén, J. Decoding post-translational modification crosstalk with proteomics. *Mol. Cell. Proteomics* **20**, 0–11 (2021).
5. Yang, X. J. & Seto, E. Lysine Acetylation: Codified Crosstalk with Other Posttranslational Modifications. *Mol. Cell* **31**, 449–461 (2008).
6. Xu, H. *et al.* PTMD: A Database of Human Disease-associated Post-translational Modifications. *Genomics, Proteomics Bioinforma.* **16**, 244–251 (2018).
7. Meng, F., Liang, Z., Zhao, K. & Luo, C. Drug design targeting active posttranslational modification protein isoforms. *Med. Res. Rev.* **41**, 1701–1750 (2021).
8. Walsh, G. Post-translational modifications of protein biopharmaceuticals. *Drug Discov. Today* **15**, 773–780 (2010).
9. Xia, C., Tao, Y., Li, M., Che, T. & Qu, J. Protein acetylation and deacetylation: An important regulatory modification in gene transcription (Review). *Exp. Ther. Med.* **20**, 2923–2940 (2020).
10. Tan, M. *et al.* Resource Identification of 67 Histone Marks and Histone Lysine Crotonylation as a New Type of Histone Modification. *Cell* **146**, 1016–1028 (2011).
11. Ntorla, A. & Burgoyne, J. R. The Regulation and Function of Histone Crotonylation. *Cell Dev. Biol.* **9**, 1–11 (2021).
12. Jiang, G., Li, C., Lu, M. & Lu, K. Protein lysine crotonylation : past , present , perspective. *Cell Death Dis.* **12**, 1–11 (2021).
13. Sabari, B. R. *et al.* Intracellular Crotonyl-CoA Stimulates Transcription through p300-Catalyzed Histone Crotonylation. *Mol. Cell* **58**, 203–215 (2015).
14. Xu, W. *et al.* Global profiling of crotonylation on non-histone proteins. *Cell Res.* **27**, 946–949 (2017).
15. Wang, S. *et al.* The Function and related Diseases of Protein Crotonylation. *Int. J. Biol. Sci.* **17**, 3441–3455 (2021).
16. Ruiz-andres, O., Sanchez-nin, M. D., Ortiz, A., Sanz, A. B. & Egido, J. Histone lysine crotonylation during acute kidney injury in mice. *Dis. Model. Mech.* **9**, 633–645 (2016).
17. Liu, Y. *et al.* Chromodomain Y-like Protein – Mediated Histone Crotonylation Regulates Stress-Induced Depressive Behaviors. *Biol. Psychiatry* **85**, 635–649 (2019).
18. Jiang, G. *et al.* HIV latency is reversed by ACSS2-driven histone crotonylation. *J. Clin. Invest.* **128**, 1190–1198 (2018).
19. Zhang, X. SEPT2 Protein Crotonylation Promotes Metastasis and Recurrence Through AKT Pathway in Hepatocellular Carcinoma and is Associated with Poor Survival. *Cell Biosci.* **13**, 63 (2023).
20. Zhang, D. *et al.* Global crotonylome reveals hypoxia-mediated lamin A crotonylation regulated by HDAC6 in liver cancer. *Cell Death Dis.* **13**, 717 (2022).
21. Cai, W. *et al.* Modulating Lysine Crotonylation in Cardiomyocytes Improves Myocardial Outcomes. *Circ. Res.* **131**, 456–472 (2022).
22. Wan, J. & Liu, H. Biomedicine & Pharmacotherapy Lysine crotonylation is involved in hepatocellular carcinoma progression. *Biomed. Pharmacother.* **111**, 976–982 (2019).
23. Zhao, Y. & Ã, O. N. J. Modification-specific proteomics : Strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* **9**, 4632–4641 (2009).
24. Lv, Y. *et al.* Global Profiling of the Lysine Crotonylome in Different Pluripotent States. *Genomics, Proteomics Bioinforma.* **19**, 80–93 (2021).
25. Wei, W. *et al.* Large-Scale Identification of Protein Crotonylation Reveals Its Role in Multiple Cellular Functions. *J. Proteome Res.* **16**, 1743–1752 (2017).
26. Huang, H., Wang, D. L. & Zhao, Y. Quantitative Crotonylome Analysis Expands the Roles of p300 in the Regulation of Lysine Crotonylation Pathway. *Proteomics* **18**, (2018).
27. Clayton, A. L., Hazzalin, C. A. & Mahadevan, L. C. Enhanced Histone Acetylation and Transcription: A Dynamic Perspective. *Mol. Cell* **23**, 289–296 (2006).

28. Drewes, G. & Knapp, S. Chemoproteomics and Chemical Probes for Target Discovery. *Trends Biotechnol.* **36**, 1275–1286 (2018).
29. Bos, J. & Muir, T. W. A Chemical Probe for Protein Crotonylation. *J. Am. Chem. Soc.* **140**, 4757–4760 (2018).
30. Hoyle, C. E. & Bowman, C. N. Thiol-ene click chemistry. *Angew. Chemie - Int. Ed.* **49**, 1540–1573 (2010).
31. Posner, T. Zum Schluss darf ich es nicht unterlausen, nieiuen Aasisteuteii Hrn. *Ber. Dtsch. Chem. Ges.* **38**, 646–657 (1905).
32. Nolan, M. D. & Scanlan, E. M. Applications of Thiol-Ene Chemistry for Peptide Science. *Front. Chem.* **8**, 1–21 (2020).
33. Nolan, M. D., Shine, C., Scanlan, E. M. & Petracca, R. Thioether analogues of the pituitary neuropeptide oxytocin via thiol-ene macrocyclisation of unprotected peptides. *Org. Biomol. Chem.* (2022) doi:10.1039/d2ob01688e.
34. Nair, D. P. *et al.* The Thiol-Michael addition click reaction: A powerful and widely used tool in materials chemistry. *Chem. Mater.* **26**, 724–744 (2014).
35. Dénès, F., Pichowicz, M., Povie, G. & Renaud, P. Thiyl radicals in organic synthesis. *Chem. Rev.* **114**, 2587–2693 (2014).
36. Uygun, M., Tasdelen, M. A. & Yagci, Y. Influence of type of initiation on thiol-ene 'click' Chemistry. *Macromol. Chem. Phys.* **211**, 103–110 (2010).
37. Taylor, N. C. & McGouran, J. F. Investigating eosin Y as a photocatalyst for the radical-dependent activity-based probing of deubiquitinating enzymes. *Org. Biomol. Chem.* **19**, 2177–2181 (2021).
38. Murale, D. P., Hong, S. C., Haque, M. M. & Lee, J. S. Photo-affinity labeling (PAL) in chemical proteomics: A handy tool to investigate protein-protein interactions (PPIs). *Proteome Sci.* **15**, 1–34 (2017).
39. Gingras, A. C., Gstaiger, M., Raught, B. & Aebersold, R. Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell Biol.* **8**, 645–654 (2007).
40. Li, X., Wang, W. & Chen, J. Recent progress in mass spectrometry proteomics for biomedical research. *Sci. China Life Sci.* **60**, 1093–1113 (2017).
41. Bern, M., Kil, Y. J. & Becker, C. Byonic: advanced peptide and protein identification software, Curr Protoc Bioinformatics. *Curr Protoc Bioinforma.* *Chapter* *13* Unit13 (2012) doi:10.1002/0471250953.bi1320s40.Byonic.
42. Palaniappan, K. K. *et al.* Isotopic signature transfer and mass pattern prediction (IsoStamp): An enabling technique for chemically-directed proteomics. *ACS Chem. Biol.* **6**, 829–836 (2011).
43. Szychowski, J. *et al.* Cleavable biotin probes for labeling of biomolecules via azide-alkyne cycloaddition. *J. Am. Chem. Soc.* **132**, 18351–18360 (2010).
44. Woo, C. M., Iavarone, A. T., Spiciarich, D. R., Palaniappan, K. K. & Bertozzi, C. R. Isotope-targeted glycoproteomics (IsoTaG): A mass-independent platform for intact N- and O-glycopeptide discovery and analysis. *Nat. Methods* **12**, 561–567 (2015).
45. Shimizu, I. & Minamino, T. Physiological and pathological cardiac hypertrophy. *J. Mol. Cell. Cardiol.* **97**, 245–262 (2016).
46. Ranieri, A., Kemp, E., Burgoyne, J. R. & Avkiran, M. β -Adrenergic regulation of cardiac type 2A protein phosphatase through phosphorylation of regulatory subunit B56 δ at S573. *J. Mol. Cell. Cardiol.* **115**, 20–31 (2018).
47. Yang, D. *et al.* A new method for neonatal rat ventricular myocyte purification using superparamagnetic iron oxide particles. *Int. J. Cardiol.* **270**, 293–301 (2018).
48. Liu, X. *et al.* MOF as an evolutionarily conserved histone crotonyltransferase and transcriptional activation by histone acetyltransferase-deficient and crotonyltransferase-competent CBP/p300. *Cell Discov.* **3**, (2017).
49. Leijten, N. M., Heck, A. J. R. & Lemeer, S. Histidine phosphorylation in human cells ; a needle or phantom in the haystack? **19**, 827–828 (2022).
50. Hou, J. Y., Zhou, L., Li, J. L., Wang, D. P. & Cao, J. M. Emerging roles of non - histone protein crotonylation in biomedicine. *Cell Biosci.* **11**, 1–16 (2021).
51. Chatgilialoglu, C., Altieri, A. & Fischer, H. The kinetics of thiyl radical-induced reactions of monounsaturated fatty acid esters. *J. Am. Chem. Soc.* **124**, 12816–12823 (2002).