Structural and functional analysis of DDR1 autoinhibition

Douglas James Sammon

Cardio-Respiratory Interface
National Heart and Lung Institute
Imperial College London

Dr Birgit Leitinger & Prof. Erhard Hohenester

A thesis submitted for the degree of
Doctor of Philosophy
December 2020

Statement of originality

I Douglas Sammon confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis and referenced appropriately.

Copyright declaration

Abstract

Discoidin domain receptor 1 (DDR1) is a collagen activated receptor tyrosine kinase (RTK) which controls cellular proliferation and migration. DDR1 plays important roles in organogenesis and wound healing. Furthermore, aberrant DDR1 signalling is implicated in the progression and poor prognosis of several diseases, including organ fibroses and cancers. DDR1 is therefore an attractive target for pharmacological intervention. However, unlike in many other RTKs, the regulatory mechanisms underpinning DDR1 signalling are poorly understood. This project investigated the regulatory function of the long intracellular juxtamembrane (JM) region of DDR1. The kinase proximal JM segment, termed JM4, is shown to be an important regulator of DDR1 kinase activity. A 2.58 Å resolution crystal structure revealed that the JM4 segment forms a hairpin which enters the kinase active site and reinforces activation loop autoinhibition. Enzymological analysis of purified DDR1 constructs demonstrated that this autoinhibition is relieved in an ordered process which begins with the rapid, in *cis*, phosphorylation of the JM4 segment (Tyr569 and Tyr586), followed by slow, in *trans*, phosphorylation of the activation loop (Tyr796). Both successive phosphorylation events are shown to have drastic activating effects on the kinase catalytic rate. Analysis of cell expressed DDR1 also revealed that JM4 Tyr mutation (DDR1-Y569F/Y586F) abolishes collagen induced receptor activation. A secondary positive role for the JM4 region in DDR1 activation is also identified through cell-based analysis. This role could be the recruitment of Src, a non-receptor tyrosine kinase, which is shown to be an activator of DDR1, but not DDR1-Y569F/Y586F, signalling. The identification of the DDR1 JM4 region as a regulator of receptor signalling provides an interesting avenue for the development of DDR1-specific kinase inhibitors.

Acknowledgements

I would like to begin by thanking my supervisors Birgit and Erhard. I am incredibly grateful for your support and mentorship throughout my PhD. I look forward to continuing our collaborations into the future.

I would like to thank all the past and present members of both labs: Yuxuan Feng, Ziteng Hao, David Corcoran, David Briggs, Marta Busse, Patricia Paracuellos Torrecilla, and Sadaf-Ahmahni Hussain. A special thank you goes to Fuad Mosis who has been an excellent friend and lab mate throughout the pandemic, and to Victoria Juskaite for her help, guidance, and friendship during the early stages of my PhD.

I am very grateful to Val Good at the Baculovirus Facility for the use of their facilities and for her help. Also, I would like to thank the OPPF for their training, and Marc Morgan and the X-ray crystallography facility for training and assistance. Thank you also to Jasmine Gratton and Athena Cheung for conducting preliminary experiments.

Thank you to the people at Bristol Zoo who accommodated me during my PIPs placement, especially Holly and Neil.

To my friends and family who have been there for me throughout the PhD, pandemic, and life in general, I cannot thank you enough. Specifically, I would like to thank Matéz, Team, Sean, and Jonny.

Finally, I would like to thank Willow. I could not ask for a more supportive partner, friend, and cat mother.

Table of contents

Table of figures

List of tables

Abbreviations

ALK - Anaplastic lymphoma kinase

A-loop – Activation loop

AMP-PNP - Adenylyl-imidodiphosphate

ANOVA – Analysis of variance

APS - Ammonium peroxodisulfate

ATP - Adenosine triphosphate

ADP – Adenosine diphosphate

AMP – Adenosine monophosphate

BSA - Bovine serum albumin

CCP - Cambridge Centre for Proteomics

CSF1R - Colony-stimulated factor 1 receptor

DDR - Discoidin domain receptor

DMSO - Dimethyl sulfoxide

DS – Discoidin

DTT - Dithiothreitol

E - Eluted

ECM – Extracellular matrix

EDTA - Ethylenediaminetetraacetic acid

EGF – Epidermal growth factor

EGFR - Epidermal growth factor receptor

ELISA - Enzyme-linked immunosorbent assay

Eph - Ephrin receptor

Erb - Homology to the erythro-blastoma viral gene product

ERK - Extracellular signal-regulated kinase

ESCRT - Endosomal sorting complex required for transport

FBS – Foetal bovine serum

FCS - Foetal Calf Serum

FRET - Fluorescence resonance energy transfer

FGFR - Fibroblast growth factor receptor

FITC - Fluorescein isothiocyanate

Flt3 - Fms like tyrosine kinase 3

FT – Flow through

GPCR – G protein coupled receptor

Grb2 - Growth factor receptor-bound protein 2

HEK293 – human embryonic kidney 293 cells

HEPES - 4-2-hydroxyethyl-1-piperazineethanesulfonic acid

IEX – Ion exchange

IF – Immunofluorescence

Ig - Immunoglobulin

IGF1-R - Insulin-like growth factor 1 receptor

IL – Interleukin

InsR - Insulin receptor

InsRR - Insulin receptor related receptor

IP – Immunoprecipitation

IRS1 - Insulin receptor substrate 1

JNK - c-Jun N-terminal kinase

LB - Lysogeny broth

LMR - Lemur tyrosine kinase

LTK - Leukocyte tyrosine kinase

mAb – Monoclonal antibody

MAPK - Mitogen activation protein kinase

MDCK - Madin-Darby Canine Kidney cells

Met - Mesenchymal-epithelial transition factor

MMP – Matrix metalloprotease

MOPS - 3-N-morpholinopropanesulfonic acid

MS – Mass spectrometry

MuSK - Muscle-specific kinase

NF-κB - Nuclear factor-κB

NGF - Nerve growth factor

NIH 3T3 - National Institute of Health 3-day transfer

Ni-NTA - Nickel-nitrilotriacetic acid

NSCLC - Non-small-cell lung carcinoma

OPPF – Oxford Protein Production Facility

PAGE - Polyacrylamide gel electrophoresis

PBS - Phosphate buffered saline

PD-1 - Programmed death protein 1

PDGFR - Platelet derived growth factor receptor

PDL-1 - Programmed death ligand 1

PI3-K - Phosphoinositide 3-kinase

PI3P - Phosphatidylinositol 3-phosphate

PIP_2 - Phosphatidylinositol-4,5-bisphosphate

PKA – Protein kinase A

PLC - Phospholipase C

PLCL-2 - Phospholipase C like-2

PLC-γ1 - Phospholipase C-γ1.

PROTACs - Proteolysis-targeting chimaeras

PTB – Phosphotyrosine binding

PTK7 - Protein tyrosine kinase 7

pY – Phosphotyrosine

Ras-GAP - Ras-GTPase activating protein.

RET - Rearranged during transfection

Ron - Recepteur d'origine Nantais

ROR - Receptor Tyr kinase Like orphan receptor 1

RTK – Receptor Tyr kinase

Ryk - Related to receptor tyrosine kinase

SDS - Sodium dodecyl sulphate

SEC-MALS - Size exclusion chromatography-multiple angle light scattering

SFKs - Src family kinases

SH2 – Src homology 2

SHIP-1/2 - SH2 containing inositol polyphosphate 5-phosphate-1/2.

SHP-1 - Src homology region 2 domain-containing phosphatase-1

Sf9 - Spodoptera frugiperda 9

STAT - signal transducer and activator of transcription

STYK1 - Ser/Thr/Tyr kinase 1

TCEP - Tris-2-carboxyethyl-phosphine hydrochloride

TEMED - Tetramethylethylenediamine

TIE1 - Tyr kinase with immunoglobulin-like and EGF-like domains

TM – Transmembrane

JM - Juxtamembrane

Trk - Tropomyosin receptor kinase B

TYRO3 - Tyr-protein kinase receptor 3

VEGFR - Vascular endothelial growth factor receptor

WB – Western blot

Chapter 1. Introduction

Adapting to changes in the extracellular environment is a fundamental cellular process (Nair et al, 2019) To effectively respond to these changes, cells express receptors which detect specific extracellular stimuli (Lemmon & Schlessinger, 2010; Nair et al, 2019; Uings & Farrow, 2000). Receptors transduce these signals across the plasma membrane triggering an intracellular response through the modulation of cell signalling pathways (Nair et al, 2019). An extraordinary array of environmental stimuli has resulted in the evolution of numerous receptor classes which have distinct ligand binding patterns and signalling outputs. These include G-protein coupled receptors (GPCRs), integrins, ligand-gated ion channels, and receptor tyrosine kinases (RTKs) (Nair et al, 2019).

GPCRs are the largest family of membrane receptors and are activated by a diverse range of stimuli; including, light, polypeptides, lipids, sugars, and proteins (Hilger et al, 2018). GPCRs consist of a single polypeptide which spans the plasma membrane 7 times (Rosenbaum et al, 2009). Ligand stimulation of a GPCR results in activation of an associated G protein. This occurs through GDP-GTP nucleotide exchange which causes dissociation of an inactive G protein heterotrimer (Rosenbaum et al, 2009). Activated G protein can then induce the production of secondary messenger molecules, such as cyclic-AMP, which propagate the GPCR signal and elicit a cellular response (Hilger et al, 2018; Rosenbaum et al, 2009). Integrins on the other hand function mechanically through interactions with the extracellular matrix (ECM), as well as biochemically by transducing extracellular signals into the cell (Kechagia et al, 2019). They exist as dimers of one α and one β subunit which have no inherent catalytic capacity (Hight-Warburton & Parsons, 2019). Therefore, integrins rely on the formation of signalling complexes on their intracellular domains to transduce signals. Ligand-gated ion channels form pores in the plasma membrane which enable ions (e.g. Na^+, K^+, Ca^{2+}, or Cl^-) to pass through the membrane in response to ligand binding (e.g. a neurotransmitters) (Barry & Lynch, 2005). Influx or efflux of ions results in a change in membrane potential which triggers a cellular response. Finally, RTKs are a family of single spanning transmembrane receptors which have an intracellular kinase domain (Lemmon & Schlessinger, 2010). Ligand binding results in kinase activation. The kinase domain then propagates the signal through

phosphorylation of downstream targets (Lemmon & Schlessinger, 2010). The following sections will focus on the activation and regulation of RTKs, with a focus in the latter half placed on a highly atypical family of RTKs, the discoidin domain receptors (DDRs).

1.1 Receptor tyrosine kinases

RTKs govern several processes which are vital to life including proliferation, differentiation, metabolism, cell cycle control, and cell migration (Lemmon & Schlessinger, 2010). There are 58 human RTKs which can be separated into 20 subfamilies (**Figure 1**, **Figure 2**) (Lemmon & Schlessinger, 2010; Trenker & Jura, 2020). The topology of RTKs is strictly conserved (Lemmon & Schlessinger, 2010; Ullrich & Schlessinger, 1990). They consist of an extracellular ligand binding domain, a single transmembrane (TM) segment, a catalytic Tyr kinase domain, and an extracellular and intracellular juxtamembrane (JM) region. RTKs sense a wide range of activating ligands, which are typically short polypeptides such as growth factors (Lemmon & Schlessinger, 2010; Trenker & Jura, 2020). In the prototypical RTK, ligand binding to the extracellular region results in dimerisation of inactive receptor monomers. This drives *trans*-phosphorylation of Tyr residues on the intracellular domains of the protein (Lemmon & Schlessinger, 2010). Phospho-Tyr residues stabilise the active conformation of the protein and act as docking sites for downstream signalling proteins which can propagate the signal (Du & Lovly, 2018; Trenker & Jura, 2020). Due to the fundamental nature of the processes governed by RTKs, their dysregulation has been linked to the progression of several diseases. These include cancers, diabetes, inflammation, atherosclerosis, and angiogenesis (Lemmon & Schlessinger, 2010; Yamaoka et al, 2018). This has prompted the development of pharmaceutical agents which can target RTKs and block their signalling (Bhullar et al, 2018; Fasano et al, 2014; Mirshafiey et al, 2014; Wang & Cole, 2014).

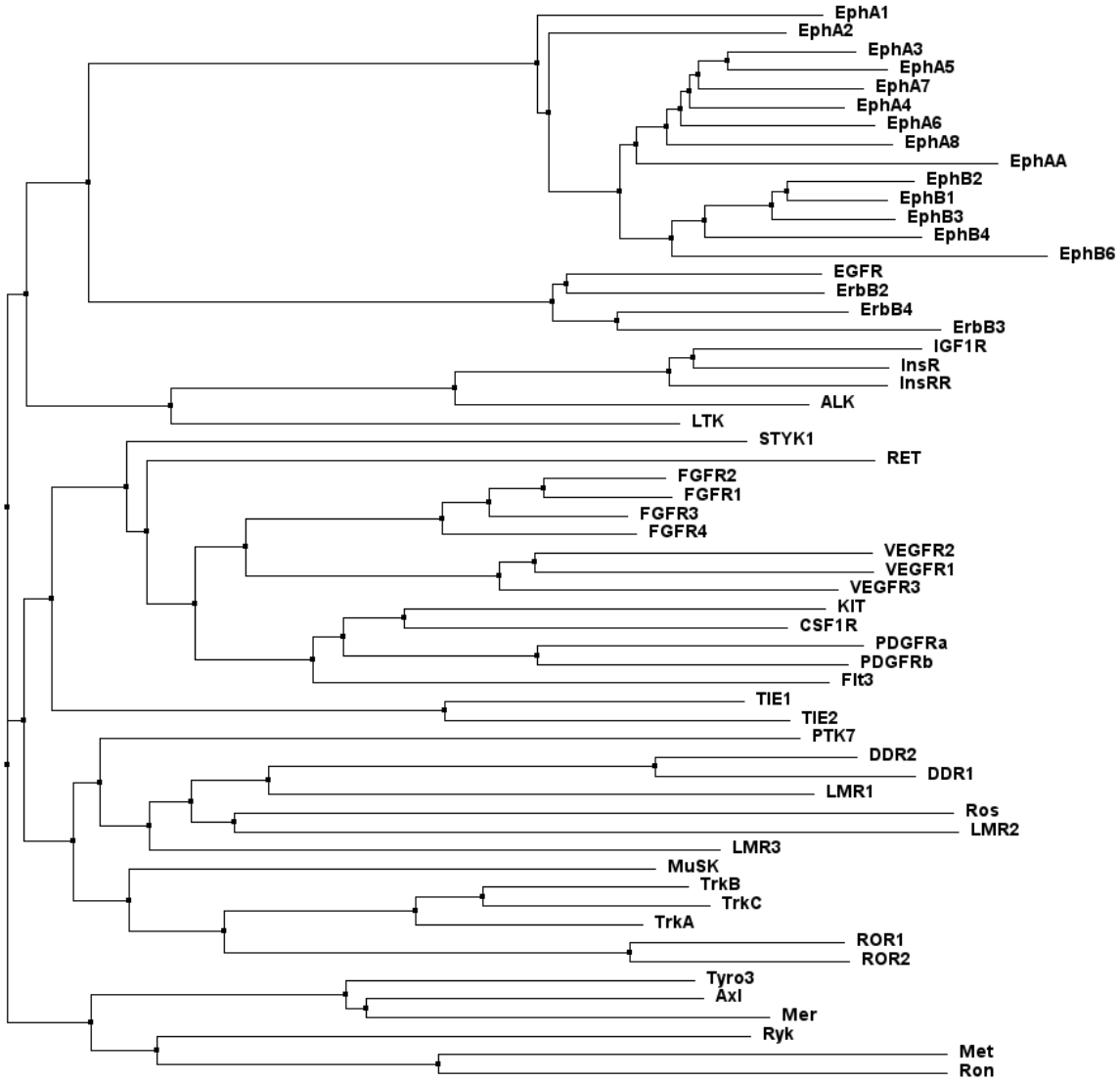

Figure 1. Human receptor tyrosine kinase family phylogram.
Full amino acid sequences were aligned using Clustal Omega Multiple Sequence Alignment (EMBL-EBI). The phylogram was then constructed using Simple Phylogeny (EMBL-EBI) and neighbour-joining clustering. The phylogram was then edited using Jalview. Evolutionary distance is drawn to scale; however, distance values have been excluded for simplicity. Eph (Ephrin receptor), EGFR (epidermal growth factor receptor), Erb (homology to the erythro- blastoma viral gene product), InsR (insulin receptor), InsRR (insulin receptor related receptor), IGF1-R (insulin-like growth factor 1 receptor), ALK (anaplastic lymphoma kinase), LTK (leukocyte tyrosine kinase), STYK1 (Ser/Thr/Tyr kinase 1), RET (rearranged during transfection), FGFR (fibroblast growth factor receptor), VEGFR (vascular endothelial growth factor receptor), CSF1R (colony-stimulated factor 1 receptor), PDGFR (platelet derived growth factor receptor), Flt3 (fms like tyrosine kinase 3), TIE1 (Tyr kinase with immunoglobulin-like and EGF-like domains), PTK7 (protein tyrosine kinase 7), DDR (discoidin domain receptor), LMR (Lemur tyrosine kinase), MuSK (muscle-specific kinase), Trk (tropomyosin receptor kinase B), ROR (Receptor Tyrosine Kinase Like Orphan Receptor 1), TYRO3 (Tyr-protein kinase receptor 3), Ryk (related to receptor tyrosine kinase), Met (mesenchymal-epithelial transition factor), Ron (recepteur d'origine Nantais).

1.1.1 Receptor tyrosine kinase molecular architecture

While RTK topology is strictly conserved there is significant structural and functional diversity in the extracellular ligand binding region (Lemmon & Schlessinger, 2010). The extracellular regions of RTKs typically contain numerous discrete domains, such as an immunoglobulin (Ig)-like, Cys-rich, epidermal growth factor (EGF)-like, and fibronectin type-III domain (**Figure 2**) (Trenker & Jura, 2020). The types and arrangement of these extracellular domains defines the 20 RTK subfamilies (Lemmon & Schlessinger, 2010; Trenker & Jura, 2020). This heterogeneity also provides a structural basis for the discrete substrate binding capacities of RTKs (Lemmon & Schlessinger, 2010). RTKs have a single membrane-spanning TM helix (Li & Hristova, 2010). The TM helix is not simply an inert hydrophobic region; instead, it is important in stabilising active receptor dimers and the formation of signalling competent structures (Li & Hristova, 2010). Crystallographic analysis of the ErbB2 TM region revealed that the TM helix can form an active and an inactive dimer, which also highlights a potential regulatory function (Bocharov et al, 2008). Indeed, I665V mutation of the ErbB2 TM region results in a stabilisation of the TM dimer and an increased risk of cancer (Bocharov et al, 2008). Most RTKs are single-spanning TM proteins, however, there are notable exceptions (Lemmon & Schlessinger, 2010). The insulin receptor, for example, forms a disulphide linked heterotetramer of 2 α and 2 β subunits (Hubbard, 2013b). Additionally, Met receptor is a disulphide linked dimer of an extracellular α, and intracellular β, subunit (Ma et al, 2003). A typically short, flexible JM region joins the TM helix to the intracellular kinase domain (Hubbard, 2001). The JM region plays important regulatory roles in kinase catalysis (Hubbard, 2004; Jura et al, 2009). It may also form functionally important interactions with the plasma membrane (**Section 1.1.3.3**) (Hedger et al, 2015). Finally, the Tyr kinase domain is the most highly conserved RTK domain (Jura et al, 2011). It is tightly regulated in the absence of ligand stimulation through autoinhibitory interactions, and the function of receptor associated phosphatases (Lemmon & Schlessinger, 2010). Upon receptor activation, the kinase domain undergoes autophosphorylation, this relieves autoinhibitory interactions and culminates in the formation of a catalytically competent kinase configuration (Du & Lovly, 2018; Hubbard & Miller, 2007). Once activated, the kinase further phosphorylates the intracellular domains of the RTK. These phosphorylated Tyr residues then act as

binding sites for SH2 (Src homology 2) and PTB (phosphotyrosine binding) domain containing proteins (Lemmon & Schlessinger, 2010; Trenker & Jura, 2020; Wagner et al, 2013). These secondary signalling proteins are then able to directly propagate the RTK signal, or act as scaffolds for the formation of signalling complexes (Belov & Mohammadi, 2012; Gual et al, 2005; Lemmon & Schlessinger, 2010).

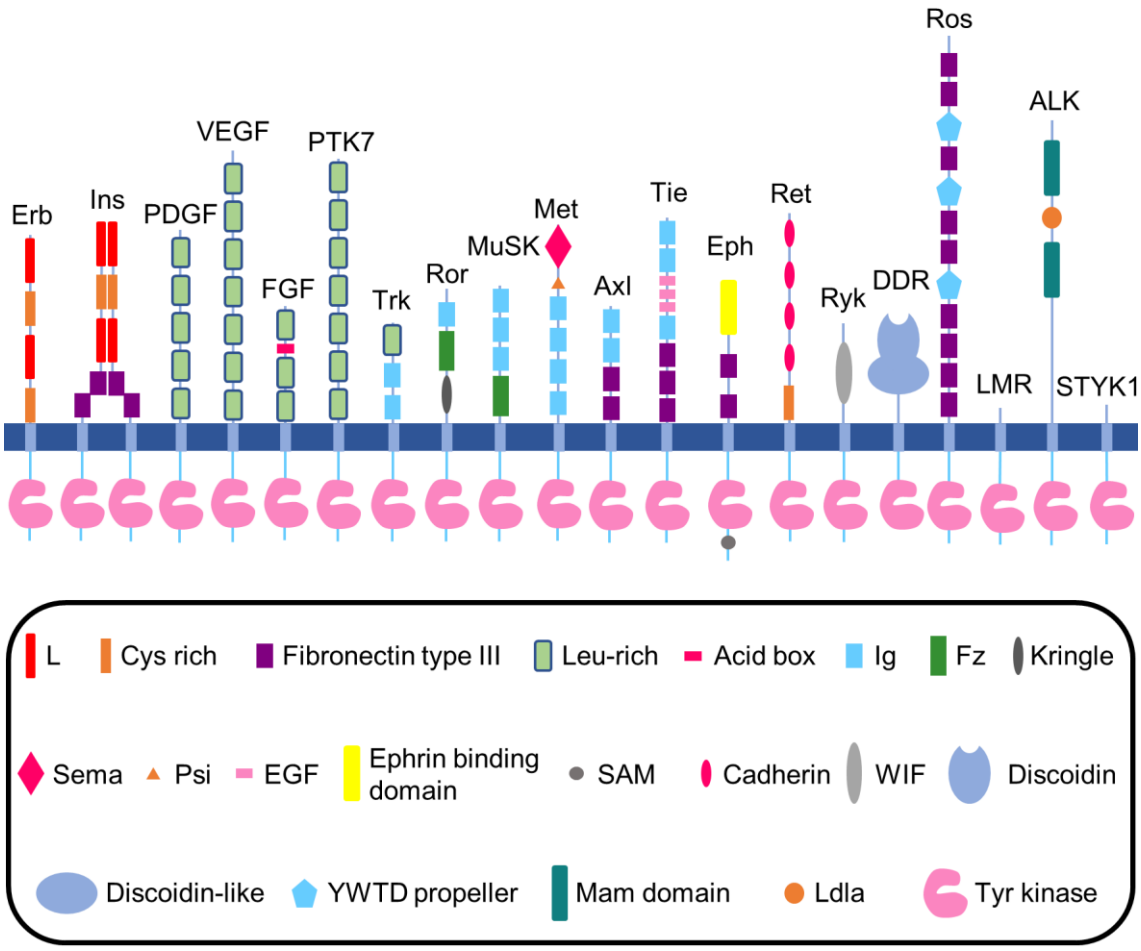

Figure 2. Receptor tyrosine kinase family molecular architecture.
The domain organisation of the 20 RTK subfamilies is depicted. Structural domains are shown and identified in the key. An adaptation of Figure 1 from Lemmon and Schlessinger, 2010.

1.1.2 Kinase domain structure and catalysis

The kinase domain is conserved not just within the RTK family, but also between the non-receptor Tyr kinases, and the Ser/Thr kinases (Endicott et al, 2012). The key catalytic residues, and structural components, are all deeply conserved throughout evolution (Endicott et al, 2012). The catalytically active conformation is also well conserved (Jura et al, 2011). However, there is diversity in the inactive conformations adopted by kinases (Jura et al, 2011). This heterogeneity in the inactive state has been exploited in the development of specific kinase inhibitors (Endicott et al, 2012). Indeed, pharmacological targeting of the active conformation of a kinase typically results in the production of more promiscuous inhibitors (Endicott et al, 2012).

All kinases adopt a bilobal architecture which consists of an N-terminal, N-lobe, which is largely β-sheet with a single α-helix (αC-helix), and a larger C-terminal, C-lobe, which is largely α-helical (Endicott et al, 2012). Both lobes are joined by a flexible linker, with the active site located in the cleft between the N- and C-lobes (catalytic cleft). The key motifs within the N-lobe include: the αC-helix, the Gly-rich loop which joins the β1 and β2 strands, and the β3 strand Lys residue. Whilst those in the C-lobe include: The activation-loop (A-loop) which is preceded by the DFG (Asp-Phe-Gly) motif, and the catalytic loop which contains the HRD (His-Arg-Asp) motif (Endicott et al, 2012).

1.1.2.1 Active kinase conformation

While kinases have specific substrates and signalling outputs, the phosphorylation reaction itself is universal, and the active kinase configuration structurally conserved (Endicott et al, 2012). Kinase activation typically leads to phosphorylation of key elements within the kinase domain and surrounding regulatory motifs (**Section 1.1.3**). These phosphorylation events facilitate the transition into an active conformation (Endicott et al, 2012). Work on the Ser/Thr kinase, Protein kinase A (PKA), and the insulin receptor in complex with an ATP analogue and peptide substrate, provided the first structural details about the active kinase configuration (Hubbard, 1997; Knighton et al, 1991). In the active state, the two kinase lobes are in a “closed” conformation which aids the positioning of ATP and the peptide (**Figure 3**). In this state, the αC-helix forms close contacts with the β-sheet of the N-lobe. This

facilitates an interaction between a conserved αC-helix Glu residue and a catalytically essential Lys residue in the β3-strand. This interaction is crucial for the Lys to coordinate the α- and β-phosphates of ATP. The N-lobe also contains a Gly-rich loop between the β1 and β2-strands which moves down towards the C-lobe in the active state to form contacts with the β- and γ-phosphates of ATP. These early structures also highlighted the importance of the C-lobe's A-loop in the active conformation. DFG is a conserved motif found at the N-terminus of the A-loop. The Asp residue of DFG coordinates a magnesium ion which is critical for ATP binding in the active site. In addition, the C-terminus of the A-loop forms contacts with the substrate peptide and ensures it is correctly oriented. Another crucial element in the C-lobe is the catalytic loop and the HRD motif found therein. The Asp of the HRD motif is the catalytic base which accepts a proton from the peptide substrate during catalysis.

In addition to these key motifs, there are also two catalytically essential hydrophobic "spines" which form upon kinase activation and ligand binding; the regulatory and catalytic spines (Jura et al, 2011). The regulatory spine is essential for peptide binding. It consists of 4 hydrophobic residues found in the β4-strand (Leu), αC-helix (Met), HRD motif (His), and the DFG motif (Phe). Upon kinase activation, the A-loop is typically phosphorylated (**Section 1.1.3.1**) (Chen et al, 2020). This leads to a reorientation of the A-loop and the αC-helix which facilitates assembly of the regulatory spine (Jura et al, 2011). The catalytic spine includes several hydrophobic residues within the N- and C-lobe. Critically however, it is completed by the adenine ring of ATP which forms a connection between the N- and C-lobes. The catalytic spine is crucial for the maintenance of an active configuration and ATP coordination (Jura et al, 2011).

1.1.2.2 Inactive kinase conformation and the DFG-flip

In contrast to the active kinase conformation, the inactive state is more open (Jura et al, 2011). Due to the flexibility of the hinge region between the two kinase lobes, the N- and C-lobes move further apart, with an upward translocation of the αC-helix and Gly-rich loop away from the C-lobe (**Figure 3**). This culminates in a loss of nucleotide binding due to a disruption of the essential Lys-Glu salt bridge (Hubbard, 2013a). Opening of the catalytic cleft also facilitates a movement of the A-loop known

as the DFG-flip (Jura et al, 2011). In the active conformation, the DFG motif is in the so-called **D**FG-IN configuration in which the DFG Asp faces into the active site and coordinates magnesium (Hubbard, 1997). Inactivation of the kinase results in a “flip” into the **D**FG-OUT conformation where the Asp and Phe exchange positions (Hubbard et al, 1994). This disrupts the regulatory spine, as the Phe is replaced by Asp. It also disrupts the catalytic spine, as ATP is lost and replaced by Phe. The **D**FG-OUT conformation also coincides with the formation of a short A-loop helix which packs against the αC-helix and supports the inactive upward positioning of the helix (Hubbard et al, 1994). The DFG-flip therefore, results in a loss of the active, tight conformation of the kinase domain and supports the adoption of a catalytically incompetent configuration. Interestingly, the rate-limiting step in kinase catalysis is the release of ADP (Grant & Adams, 1996). The **D**FG-OUT conformation is inherently more flexible and entropically favoured to **D**FG-IN (Jura et al, 2011). Therefore, the thermodynamically favourable transition to the **D**FG-OUT conformation, may be what facilitates product (ADP) release.

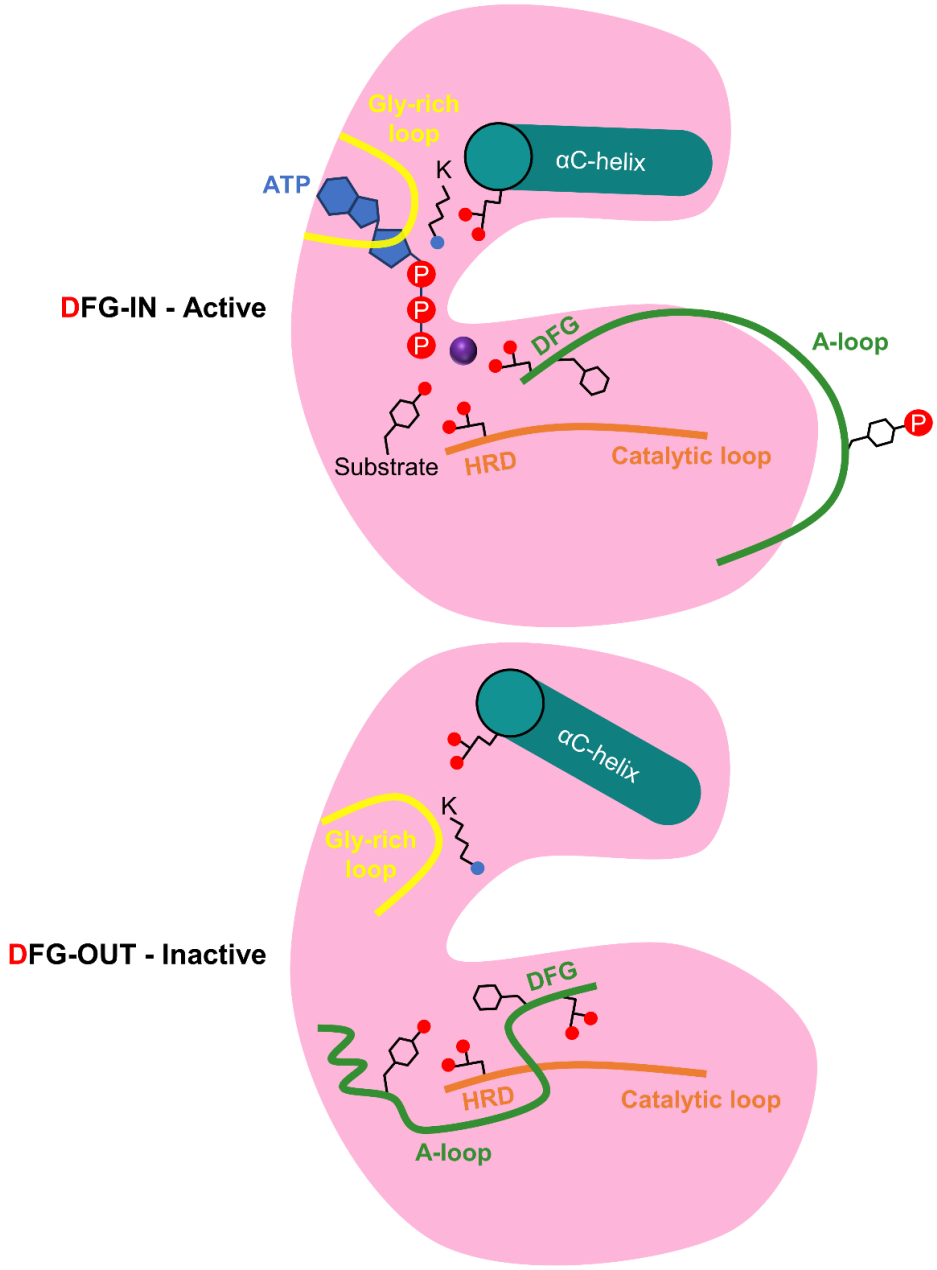

Figure 3. Kinase domain architecture.
A schematic of the active and inactive kinase conformations is shown with key regions and residues highlighted. In the active state (DFG-IN), the A-loop (green) is phosphorylated and the Asp of the DFG motif interacts with a Mg^{2+} ion (purple circle) which coordinates ATP. The Gly rich loop (yellow) also interacts with ATP, and a salt bridge forms between an αC-helix (blue) Glu and the β3-strand Lys which allows Lys to coordinate the ATP γ-phosphate. The HRD (orange) Asp residue coordinates the substrate. Upon inactivation, the A-loop is dephosphorylated and undergoes DFG-flip in which the Asp and Phe residues swap positions. The A-loop Tyr residue then forms pseudosubstrate interactions with the HRD Asp. This coincides with an upward translocation of the Gly-rich loop and the αC-helix which breaks the Lys-Glu salt bridge.

1.1.2.3 Kinase catalysis and substrate recognition

Kinases catalyse the transfer of the γ-phosphate of ATP onto a Ser, Thr, or Tyr residue (His and Arg in some bacterial and invertebrate kinases, respectively) (Kang et al, 2011; Manning et al, 2002). Kinase substrate phospho-sites are typically in flexible, disordered, regions of a protein (Iakoucheva et al, 2004). This flexibility enables the substrate to enter the catalytic site in an extended conformation and for surrounding substrate residues to interact with specificity determining pockets of the kinase (Bradley & Beltrao, 2019). The residues flanking a phosphorylation site constitute a peptide motif which defines kinase specificity (Bradley & Beltrao, 2019). In addition to the unique phosphorylation motifs, distal docking sites also promote the phosphorylation of specific targets (Endicott et al, 2012). In RTKs, for example, receptor phospho-Tyr residues help to recruit specific kinase substrates containing SH2 or PTB domains (Lemmon & Schlessinger, 2010). Due to the inherent nature of kinase catalysis, peptide substrates must form transient kinase interactions (Endicott et al, 2012). Remote binding sites however can form stable contacts, which have been shown to increase substrate affinity >1000 fold (Dar et al, 2005). These distal binding sites are therefore key determinants of specificity.

Binding of ATP and peptide substrate is not a strictly ordered process, as binding of one does not restrict binding of the other (Adams, 2001; Endicott et al, 2012). However, the peptide binding site does somewhat occlude the ATP site (Endicott et al, 2012). Therefore, in the case of larger protein substrates, ATP likely binds first, as protein binding can shield the catalytic site. Once substrates are bound and properly oriented, the catalytic reaction occurs in 2 distinct steps. 1) The kinase catalytic Asp deprotonates the peptide substrate hydroxyl group. This initiates nucleophilic attack upon the ATP γ-phosphate. 2) The phosphorylated peptide deprotonates the catalytic Asp to regenerate the native state of the enzyme (Skamnaki et al, 1999). The introduction of negative charge from the phosphate group onto the substrate peptide then triggers conformational changes which have profound effects on protein function and interactions (Lemmon & Schlessinger, 2010).

1.1.3 Regulation of receptor tyrosine kinase signalling

Dysregulated RTK signalling leads to the progression of diseases, such as cancer (Ferguson & Gray, 2018; Lahiry et al, 2010). Therefore, RTKs have evolved numerous elaborate regulatory mechanisms which help to maintain an inactive receptor state in the absence of ligand binding (Lemmon & Schlessinger, 2010). The prototypical inactive RTK is a monomer which is autoinhibited through intramolecular interactions with specific receptor peptide motifs (Bayliss et al, 2015). These can include the A-loop, C-terminal tail, or intracellular JM region (Bayliss et al, 2015). In conjunction with these intrinsic regulatory elements, RTKs also interact with phosphatases which remove activating phosphate groups (Bergeron et al, 2016). This helps maintain an inactive receptor in the absence of ligand binding. Furthermore, following ligand binding, phosphatases return RTK activation to basal levels, enabling transiency in the growth factor signal (Yao & Stagljar, 2017).

1.1.3.1 A-loop autoinhibition

The first autoinhibitory motif described for RTKs was A-loop autoinhibition in the insulin receptor (Hubbard et al, 1994). In this structure, a key A-loop Tyr, Tyr1162, forms pseudosubstrate interactions with the catalytic Asp1132 of the HRD motif. This occludes peptide and ATP access and stabilises the inactive **D**FG-OUT A-loop conformation. Following insulin binding, conformational changes in the receptor facilitate *trans*-phosphorylation of the A-loop on Tyr1162 and two other A-loop Tyr residues (Hubbard, 1997; 2013b). Phosphorylation of the A-loop leads to a loss of the inhibitory Tyr1162-Asp1132 interaction and enables the transition to the active **D**FG-IN conformation (**Figure 3**). This DFG-flip also allows the downward movement of the αC-helix and Gly-rich loop and therefore facilitates kinase activation (Hubbard, 1997; 2013b). FGFR family members also employ A-loop autoinhibition (Mohammadi et al, 1996). However, they do so through an alternative mechanism. In FGFRs, the A-loop forms kinase domain interactions that do not occlude ATP binding (Mohammadi et al, 1996). However, they do stabilise the inactive kinase conformation and block peptide substrate access (Mohammadi et al, 1996). Like in the insulin receptor, FGFR A-loop phosphorylation relieves autoinhibition (Mohammadi et al, 1996). A-loop phosphorylation is essential for the activation of most RTKs (Lemmon & Schlessinger, 2010). However, there are notable exceptions.

The EGFR family, for example, do not require A-loop phosphorylation for activation (Stamos et al, 2002; Zhang et al, 2006).

1.1.3.2 C-terminal tail autoinhibition

Another common autoinhibitory motif is the C-terminal tail. This autoinhibition was first described for the Tie2 receptor (Niu et al, 2002). However, subsequent examples have also been noted for the Ron, EGFR, and PDGFR families (Chiara et al, 2004; Gajiwala, 2013; Yokoyama et al, 2005). In the inactive Tie2 receptor, the A-loop adopts the active **D**FG-IN conformation. However, the C-terminal tail forms autoinhibitory interactions with the kinase domain that occlude substrate access. Furthermore, the tail contains phospho-sites which allow a release of autoinhibition (Niu et al, 2002).

Further complexity of C-terminal tail regulation was revealed in a recent structure of a non-receptor Tyr kinase B-Raf dimer in complex with 14-3-3 (Kondo et al, 2019). It revealed that the C-terminal tail of one B-Raf protein enters the active site of its dimer partner. This blocks substrate access to the C-terminal tail binding kinase (acceptor kinase). However, this intermolecular interaction stabilised the active conformation of the C-terminal tail donor kinase. Therefore, the C-terminal tail, whilst inhibitory for the acceptor kinase, promoted activation of the donor kinase, thereby enhancing catalysis.

1.1.3.3 Juxtamembrane autoinhibition

Autoinhibition by the JM region is a regulatory mechanism which has been reported for the PDGFR, MuSK, and Eph family receptors (Binns et al, 2000; Mol et al, 2004; Till et al, 2002). Each receptor is autoinhibited by the JM region through a unique, family-specific mechanism. However, all are relieved though autophosphorylation of Tyr residues within the JM region (Hubbard, 2004). In the PDGFR family, the JM region forms extensive contacts within the active site which reinforce the inactive conformation of the kinase (Mol et al, 2004). The Eph receptor family JM region, on the other hand, forms contacts outside the catalytic cleft (Binns et al, 2000). Interactions between the Eph receptor JM region and components of the N- and C-lobe results in a distortion of the αC-helix, breaking the catalytically essential Lys-

Glu bridge. The MuSK receptor JM region is also autoinhibitory (Till et al, 2002). However, it does not physically interact with the kinase domain. Instead, it binds to a repressor of MuSK signalling which inhibits kinase function in *trans* (Herbst & Burden, 2000; Till et al, 2002). Interestingly, the JM region can also have positive functions in receptor activation. In the EGFR family, for example, the JM region forms intermolecular interactions which promote the formation of an active asymmetric dimer (Jura et al, 2009; Red Brewer et al, 2009). In addition, the Ret receptor JM region forms intramolecular interactions with the kinase N-lobe which stabilise the active kinase conformation (Plaza-Menacho et al, 2016). Furthermore, the insulin receptor family have a dual function JM region (Cabail et al, 2015). In the inactive state of the receptor, the JM region forms autoinhibitory interactions (Cabail et al, 2015; Craddock et al, 2007). However, following receptor activation and autophosphorylation, the JM region forms intermolecular interactions which stabilise a catalytically competent dimeric kinase configuration (Cabail et al, 2015).

The JM region may also play important structural roles through interactions with the plasma membrane. Lipids, such as PIP_2 (phosphatidylinositol-4,5-bisphosphate) are important in regulating RTK activity (Michailidis et al, 2011). Computational predictions have revealed that anionic lipids, like PIP_2, cluster around RTKs through interactions with conserved basic residues in the RTK JM region (Hedger et al, 2015). It is thought that these JM region-bilayer interactions are important in regulating RTK structure and function (Hedger et al, 2015).

1.1.4 Receptor tyrosine kinase activation and downstream signalling

RTKs are typically activated through ligand-mediated dimerisation of inactive receptor monomers (Lemmon & Schlessinger, 2010). However, this is not a universal process. For example, the insulin receptor forms a covalently linked inactive heterodimer (Hubbard, 2013b). The “prototypical” RTK, EGFR, has also been shown to form inactive dimers in the absence of ligand, as well as higher order oligomers upon activation (Huang et al, 2016; Needham et al, 2016; van Lengerich et al, 2017). Furthermore, the Eph receptor family are activated through the formation of large oligomers (Ojosnegros et al, 2017). There is therefore mechanistic divergence from the classic dimerisation activation mechanism. Regardless, ligand binding must stabilise the formation of an active dimer, or oligomer. In this state, one receptor

within the complex temporarily adopts an active configuration which facilitates *trans*-phosphorylation of a neighbouring RTK (Lemmon & Schlessinger, 2010). This activates the receptor and triggers an intracellular signalling cascade (Lemmon & Schlessinger, 2010).

Receptor activation through oligomerisation may be triggered through the innate properties of the activating ligand. For example, some RTK ligands, such as the TrkA ligand, NGF (nerve growth factor), are dimeric (Wehrman et al, 2007). NGF therefore, can simultaneously bind two TrkA receptors and trigger their dimerisation (Wehrman et al, 2007). Alternatively, ligand binding can lead to conformational changes in the receptor which facilitate dimerisation. For example, EGF binding to EGFR results in a conformational change which exposes the dimerisation arm of EGFR, driving dimerisation (Ward et al, 2007). Interestingly, lower affinity ligands for EGFR, such as epiregulin and epigen, result in a less symmetrical EGFR dimer than their higher affinity ligand counterparts, such as EGF (Freed et al, 2017). These more asymmetrical receptor dimers exhibit sustained signalling and elicit cell differentiation, whilst the more symmetrical dimers resulting from EGF binding exhibit transient signalling and elicit cell proliferation (Freed et al, 2017). Insulin receptors are also activated through conformational changes but these occur within pre-formed receptor dimers following ligand binding (Hubbard, 2013b; Scapin et al, 2018; Uchikawa et al, 2019). Whatever the mechanism, ligand binding to RTKs results in activation through a highly ordered series of autophosphorylation events (Lemmon & Schlessinger, 2010). This typically begins with kinase domain A-loop phosphorylation. In the insulin receptor, Tyr1162, then Tyr1158, and finally Tyr1163 of the A-loop are phosphorylated sequentially with each event increasing catalytic rate, or substrate affinity (Favelyukis et al, 2001). In contrast, the PDGFR family must first phosphorylate their autoinhibitory JM region prior to the A-loop (Hubbard, 2004; Mol et al, 2004). These initial autophosphorylation events relieve autoinhibitory interactions and stabilise the active kinase configuration (Lemmon & Schlessinger, 2010). In a second phase, additional Tyr residues in the kinase domain and JM region are autophosphorylated. These phospho-Tyr residues recruit signalling proteins to the receptor (Lemmon & Schlessinger, 2010). RTK signalling proteins can bind directly through their SH2 and PTB domains, or through receptor associated docking proteins (Lemmon & Schlessinger, 2010; Wagner et al, 2013). These

docking proteins interact with the receptor through their SH2/PTB domains (Wagner et al, 2013). Docking proteins, such as the insulin receptor substrate 1 (IRS1) and Growth factor receptor-bound protein 2 (Grb2), have no intrinsic kinase activity but instead act as scaffolds to recruit other signalling proteins and therefore enable the formation of large signalling complexes (Belov & Mohammadi, 2012; Gual et al, 2005; Lemmon & Schlessinger, 2010). Proteins recruited to these complexes typically contain many signalling domains which allow sensing of multiple factors (Pawson, 2007). This multivalency allows spatiotemporal coincidence detection of many signalling inputs. This allows the integration of multiple inputs and hence provides a degree of specificity in the output response (Pawson, 2007).

There is significant overlap between the intracellular signalling pathways activated by RTKs (Lemmon & Schlessinger, 2010; Trenker & Jura, 2020). The MAPK (mitogen activation protein kinase) cascade, STAT (signal transducer and activator of transcription), PLC (phospholipase C), and Rac signalling pathways, for example, are all commonly used (Lemmon & Schlessinger, 2010). Despite this, RTKs can all produce specific effects in response to ligand binding. For example, in the same cell type, NGF activation of its receptor results in differentiation, whilst EGF activation of its receptor results in proliferation (Marshall, 1995; Vaudry et al, 2002). Furthermore, the expression levels of an RTK can affect the cellular response it initiates (Marshall, 1995). How RTKs produce a specific response while using a limited set of signalling pathways is currently poorly understood. But the formation of specific signalling complexes and the integration of multiple factors into a signalling network likely contributes to specificity.

Following activation, RTK signalling is modulated through a series of positive and negative feedback networks (Hu et al, 2013; Ledda & Paratcha, 2007). Positive feedback increases system sensitivity to signals through amplifying the stimulus. On the other hand, negative feedback reduces stochastic activation and returns signalling to basal levels (Purvis & Lahav, 2013). There are numerous examples of positive and negative feedback, which can be either immediate in their effect (e.g. protein-protein interaction), or delayed (e.g. gene expression) (Purvis & Lahav, 2013). Temporal differences in negative and positive feedback mechanisms can cause oscillations in the system producing a pulse like response (Purvis & Lahav,

2013). As cell fate can be determined by the intensity and duration of RTK signalling, these oscillations may be crucial regulators of the cellular response elicited by RTK signalling (Hu et al, 2013).

1.1.5 Downregulation of receptor tyrosine kinase signalling

RTK signals are typically transient, as prolonged activation can lead to malignant transformation and disease (Goh & Sorkin, 2013). This transiency is facilitated by downregulation of the receptor through internalisation and lysosomal degradation (Goh & Sorkin, 2013). Upon receptor activation and autophosphorylation, phospho-Tyr sites recruit E3 ubiquitin ligases, for example, the SH2 domain containing c-Cbl. Receptor phosphorylation of c-Cbl leads to activation of its ligase activity and RTK ubiquitination (de Melker et al, 2001; Goh & Sorkin, 2013). The ubiquitin tags added to the RTK then allow recruitment into clathrin-coated pits which invaginate and internalise the receptor (Goh & Sorkin, 2013). In the majority of cases, the receptor will be trafficked along the endocytic pathway from the early to late endosome and finally into the lysosome for degradation (Goh & Sorkin, 2013). However, the receptor can also be recycled from the early endosome to the plasma membrane, or trafficked to other intracellular compartments such as the nucleus or Golgi network (Chen et al, 2019; Podlecki et al, 1987). For example, in hepatocytes, where insulin sensing is an important process, the majority of internalised insulin receptor is recycled to the cell surface (Chen et al, 2019). While recycling of RTKs to the cell surface is a common process, dysregulation can culminate in disease. For example, ErbB2 is thought to constitutively recycle to the plasma membrane (Pietilä et al, 2019). Overexpression of ErbB2 therefore leads to high cell surface levels of the receptor, which in turn activates other EGFR family members, leading to the development of diseases, such as breast cancer (Bertelsen & Stang, 2014).

Following RTK activation, the receptor typically clusters into clathrin-coated pits which facilitate receptor internalisation (Goh & Sorkin, 2013). Recent work has shown that recruitment of RTKs to clathrin-coated pits may have additional functions in signalling. Clathrin-coated pits represent a plasma membrane nanodomain containing a concentration of specific proteins (Goh & Sorkin, 2013). They can therefore act as scaffolds which allow the assembly of transient RTK signalling complexes. Indeed, inhibition of clathrin-coated pit formation, but not endocytosis,

results in a reduction in RTK-mediated phosphorylation of Gab1 and Akt (Garay et al, 2015). Therefore, in addition to their role in internalisation, clathrin-coated pits may be important signalling nanodomains.

Following internalisation from the cell surface, ubiquitinated RTKs are sorted at the early endosome by the ESCRT (endosomal sorting complex required for transport) system (Goh & Sorkin, 2013; Sorkin & von Zastrow, 2009). It is within the early endosome, and by this system, that their fate is decided. If the receptor is to be recycled, the ligand must first dissociate, and the receptor be dephosphorylated by phosphatases (Chen et al, 2019). However, without ligand dissociation and phosphatase action, the receptor will remain active and continue signalling within the different endocytic compartments (Sugiyama et al, 2019). If targeted for degradation, the receptor will move from the early to late endosomes before fusing with the lysosome. The early and late endosomes have distinct lipid and protein compositions (Bissig & Gruenberg, 2013). The early endosome, for example, is rich in PI3P (phosphatidylinositol 3-phosphate) and Rab5 (Gillooly et al, 2000), while the late endosome is rich in lysobisphosphatidic acid and Rab7 (Kobayashi et al, 1998). These unique compositions mean that the internalised RTK interacts with a specific subset of proteins. Therefore, the receptor can form compartment specific signalling complexes (Sugiyama et al, 2019). These spatially regulated signalling nodes allow the production of specific responses and may be an important part of RTK signalling (Miaczynska, 2013; Sugiyama et al, 2019). For example, Akt is commonly activated by RTKs at different subcellular locations (plasma membrane, early endosome and late endosome) (Sugiyama et al, 2019). When activated at the early endosome, Akt phosphorylates glycogen synthase kinase 3, this results in the degradation of multiple protein targets through the endocytic pathway (Braccini et al, 2015). Therefore, whilst internalisation and degradation of RTKs is important for downregulating the signal and enabling transiency, it is also a vital process required for specific RTK signalling outputs.

1.1.6 Targeting receptor tyrosine kinases in disease

RTKs are popular targets for pharmacological intervention with dozens of kinase inhibitors currently approved as therapeutics (Bhullar et al, 2018; Ferguson & Gray, 2018). These are predominantly used for the treatment of cancers in which RTKs are

often mutated or overexpressed (Yamaoka et al, 2018). However, aberrant kinase signalling also contributes towards immunological, inflammatory, cardiovascular, and degenerative disorders (Ferguson & Gray, 2018). Expanding the use of kinase inhibitors into therapeutics for these other diseases has therefore become an interesting area of research (Ferguson & Gray, 2018).

RTK inhibitors come in the form of small molecules which target the kinase domain, and antibodies which bind to the extracellular regions of the receptor (Yamaoka et al, 2018). An excellent example of the success which can be elicited from targeting RTK signalling comes from the antibody therapy Herceptin™ (trastuzumab) (McKeage & Perry, 2002). Herceptin™ binds to the RTK, ErbB2, which is overexpressed in breast cancer. This binding triggers ErbB2 internalisation and degradation (McKeage & Perry, 2002). Herceptin™ treatment increases breast cancer survival rate and reduces recurrences (McKeage & Perry, 2002). In addition to antibodies, numerous small molecule kinase inhibitors have been developed over the past few decades (Yamaoka et al, 2018). They can be categorised into 6 groups (**Table 1**) (Roskoski, 2016). Targeting of the highly conserved kinase active site, unsurprisingly, means that many of these inhibitors are promiscuous (Hanson et al, 2019). For example, imatinib, originally developed as a Bcr-Abl inhibitor, also inhibits PDGFR, c-KIT, DDR1, and DDR2, amongst others (Hanson et al, 2019). In some cases, this promiscuity can be advantageous. Imatinib, for example, can be used in the treatment of chronic myeloid leukaemia in which Bcr-Abl is the driving kinase, and gastrointestinal stromal tumour in which c-KIT is the driving kinase (Shawver et al, 2002). However, unexpected off-target effects which result from inhibitor promiscuity can reduce the usefulness of these drugs (Zhang et al, 2008). Significant investment in pharmacological screening, and the use of allosteric inhibitors, has nevertheless produced some kinase inhibitors with good selectivity (Yamaoka et al, 2018). However, the use of these small-molecules, and of antibody therapy, is hampered by the development of drug resistance in malignancies (Yamaoka et al, 2018). Selective pressure within the tumour often results in mutations which render a drug ineffective. For example, the T766M mutation in EGFR confers resistance to gefitinib (Bean et al, 2008). Alternatively, compensation occurs, in which tumours by-pass the need for the RTK signal. For example, EGFR inhibition is overcome by an upregulation of ErbB2 (Engelman et al, 2007). The search for inhibitors which can

overcome resistance mutations, and compensation, is therefore intense (Bhullar et al, 2018).

Table 1. Small-molecule kinase inhibitor classes.
The binding mode for each of the 6 kinase inhibitor classes is shown along with an example of an inhibitor within that class and the kinase it targets in brackets.

Type	Binding mode	Example
I	ATP-competitive Active **D**FG-IN	Dasatinib (c-KIT) (Hochhaus & Kantarjian, 2013)
II	ATP-competitive Inactive **D**FG-OUT	Imatinib (PDGFR) (Yang et al, 2005)
III	Allosteric Do not occlude ATP binding site	TAK-733 (MEK) (Adjei et al, 2017)
IV	Allosteric Do not occlude ATP or peptide binding site	Everolimus (ErbB2) (André et al, 2014)
V	Bivalent Binds 2 kinase sites	NA
VI	Covalent	Ibrutinib (Bruton's tyrosine kinase) (Davids & Brown, 2014)

An interesting method for the discovery of novel kinase inhibitors is the use of covalent fragment libraries. The incorporation of electrophilic molecules into fragments enables the irreversible covalent modification of the target protein, typically on Cys or Lys residues (Keeley et al, 2020). Fragment hits which enhance target protein binding can then be optimised. Covalent modification allows the detection of compounds which bind in difficult to target regions, for example shallow pockets outside of the deep catalytic cleft (Craven et al, 2018; Uhlenbrock et al, 2019). It can therefore be useful for identifying compounds which bind space not used by existing inhibitors. Indeed, this method has been successfully used for the identification of allosteric inhibitors of several kinases, including Akt and Cdk2 (Craven et al, 2018; Uhlenbrock et al, 2019). Another example of a pharmacological technique with some promise is the use of PROTACs (proteolysis-targeting chimaeras). These compounds link a kinase binding drug to an E3-ligase binding molecule. Binding of the PROTAC to an RTK recruits an E3-ligase to the receptor, thus triggering ubiquitination and degradation of the RTK (Ferguson & Gray, 2018).

PROTACs have been developed for several kinases but perhaps the most promising avenue for PROTAC production is targeting of pseudokinases, such as ErbB3 (Ferguson & Gray, 2018). Pseudokinases have no intrinsic kinase activity, making them intractable to kinase inhibitor treatment (Kung & Jura, 2019). However, they can activate other catalytically functional family members, such as ErbB2, through heterodimerisation (Kung & Jura, 2019). PROTACs may therefore provide an opportunity to develop therapeutics against these challenging drug targets through the promotion of degradation. Another interesting avenue for RTK inhibition is through combinational therapy of kinase inhibitors with T-cell checkpoint inhibitors (Ferguson & Gray, 2018). Targeting of the T-cell checkpoint proteins, PD-1 (programmed death protein 1) and PDL-1 (programmed death ligand 1), with antibody therapies has only successfully resulted in cancer remission in some cases (Brahmer et al, 2012). There is now evidence that some RTK oncogenes promote immune evasion in T-cell checkpoint inhibitor resistant cancers (Gainor et al, 2016). Kinase inhibitors typically result in a rapid ablation of tumours; however, this is short lived (Barouch-Bentov & Sauer, 2011). Furthermore, immunotherapy resistance may be a result of strong immunosuppressive effects of well-established tumours (Phan et al, 2003). Therefore, combination of an RTK inhibitor, which rapidly reduces the tumour size and the immune escape effects of RTK activity, with T-cell checkpoint inhibitors, may be a viable therapeutic option. Indeed, combination of MEK inhibitors with anti-PDL-1 antibody results in long-lasting tumour suppression in a mouse model of colon carcinoma (Ebert et al, 2016).

In summary, RTK inhibitors have been extremely useful for controlling cancer cell growth (Ferguson & Gray, 2018). However, the rapid development of inhibitor resistance and compensation has meant that these drugs are only effective in the short term (Yamaoka et al, 2018). Compounds have therefore been developed which target inhibitor resistant mutant RTKs (Spellmon et al, 2017). However, the effectiveness of these, and any subsequent compounds, will still be hampered by the same issues. Therefore, it will be essential to use novel methods of kinase inhibition, such as PROTACs, as well as developing therapeutic strategies which will reduce the possibility for drug resistance, such as combining checkpoint inhibitor immunotherapy with kinase inhibitors.

1.2 Discoidin domain receptors

The DDRs, DDR1 and DDR2, are a structurally unique subfamily of RTKs which contain an eponymous extracellular discoidin (DS) domain (Leitinger, 2014). The DS domain was first described in the discoidin I protein of *Dictyostelium discoideum* where it functions as an N-acetylgalactosamine-binding protein (Mathieu et al, 2010; Poole et al, 1981). DS domains are also found in the blood coagulation factors V and VIII where they function as anionic phospholipid-binding proteins (Fuentes-Prior et al, 2002). In the DDRs, however, the DS domain provides the most unique aspect of DDR biology. Namely, it binds to the DDR activating ligand; collagen (Shrivastava et al, 1997; Vogel et al, 1997). RTKs are typically activated by small soluble growth factors which can be quickly removed, and therefore provide a short-lived RTK stimulus (Lemmon & Schlessinger, 2010). Collagens, however, are large, stable components of the ECM, which in many DDR expressing cell types is constitutively in contact with the cell surface (Kadler et al, 2007). Binding of collagen to the DDRs also results in extraordinarily slow kinase activation (>90 minutes), which can be sustained for days (Shrivastava et al, 1997; Vogel et al, 1997). This contrasts with typical RTKs which are activated within seconds, and then quickly downregulated (Lemmon & Schlessinger, 2010). The DDRs also exist as stable, non-covalently linked dimers in the absence of ligand stimulation (Noordeen et al, 2006). This makes the classic ligand-induced dimerisation mechanism, used by many RTKs, inapplicable to the DDRs (Lemmon & Schlessinger, 2010). These characteristics highlight the atypical nature of DDRs and hint at potentially unique activational and regulatory mechanisms.

1.2.1 DDR physiological functions

The DDRs are expressed widely during development and within adult tissue (Leitinger, 2014). DDR1 is expressed largely in epithelial cells of the brain, lung, kidney, spleen, and placenta (Di Marco et al, 1993; Leitinger, 2014). In contrast, DDR2 is expressed largely in mesenchymal cells of the skeletal and cardiac muscle, kidney, and lung (Karn et al, 1993; Lai & Lemke, 1994). Both DDRs are also expressed in immune cells (Afonso et al, 2013; Kamohara et al, 2001). Knockout studies of the DDRs in mice revealed that DDR1 is important in organogenesis of the mammary gland, kidney, and inner ear, in addition to gonadal function (Gross et al,

2004; Meyer zum Gottesberge et al, 2008; Vogel et al, 2001), while DDR2 is important in bone morphogenesis (Labrador et al, 2001). These functions appear to be elicited through DDR promoting proliferation and migration. DDR2, for example, promotes chondrocyte proliferation in bone growth (Labrador et al, 2001), whilst DDR1 deficient mice demonstrate hyper-proliferation and abnormal branching of the mammary gland (Vogel et al, 2001). In humans, there are no known developmental disorders involving DDR1. However, single nucleotide polymorphisms in the DDR1 gene have been associated with increased propensity towards schizophrenia (Gas et al, 2019). Hereditary mutations in DDR2 lead to the development of a form of dwarfism called spondylo-meta-epiphyseal dysplasia (Borochowitz et al, 1993). This occurs through mutations within the kinase domain (Y712I, I726R, and R752C) which lead to DDR2 retention within the endoplasmic reticulum, or through a mutation (E113K) within the extracellular DS domain which precludes collagen binding (Al-Kindi et al, 2014; Ali et al, 2010; Bargal et al, 2009). These mutations, therefore, abrogate DDR2 function, leading to bone morphogenic disorder. In contrast, two activating mutations within the DDR2 kinase domain (Y740C and I610P) result in a fibrotic disorder called Warburg-Cinotti syndrome. This devastating disease is characterised by corneal neovascularisation, keloid and ulcer formation, tissue wasting, and acro-osteolysis (Xu et al, 2018).

DDR function in development has been well documented; however, their function in adult tissue is less clear (Leitinger, 2014). They appear to regulate cell adhesion, migration, proliferation, and differentiation (Leitinger, 2014). However, most of our understanding of DDR contribution towards these processes comes from pathological and not physiological conditions. Nevertheless, DDR2 has been implicated in wound healing through the promotion of fibroblast proliferation (Labrador et al, 2001), whilst DDR1 has been implicated in the immune response, and collagen synthesis and remodelling (Borza et al, 2017; Zhong et al, 2019). Both DDRs are also key mediators of matrix metalloprotease (MMP) expression (Leitinger, 2014). MMPs are proteases which degrade the ECM thus facilitating cell migration and invasiveness (Van Doren, 2015). In development, this aids organogenesis; however, in adults it likely contributes towards DDR-mediated disease progression (**Section 1.2.6**). DDR1 can stimulate collagen synthesis and collagen contraction (Borza et al, 2017; Coelho et al, 2017). This, in conjunction with MMP expression,

makes DDR1 an important regulator of collagen remodelling and deposition, which has important implications during development and disease. In immunity, DDR1 activation in monocytes and T-cells promotes cell migration (Hachehouche et al, 2010; Kamohara et al, 2001), whilst in neutrophils, DDR2 promotes migration through MMP8 expression (Afonso et al, 2013). In summary, the DDRs appear to control several important cellular functions. However, despite their widespread expression, their physiological function in adult tissue has only been investigated in wound healing and the immune response. An improved understanding of DDR function in adult tissues is a vital area of research.

1.2.2 DDR molecular architecture

The DDRs, like most RTKs, are single-spanning TM proteins (Leitinger, 2014; Lemmon & Schlessinger, 2010). Their N-terminal extracellular region is composed of a collagen binding DS domain, which is closely associated with a DS-like domain (Carafoli & Hohenester, 2013). An extracellular JM segment of ~50 amino acids in DDR1, and ~30 amino acids in DDR2, links these domains to the TM region. The intracellular JM region of the DDRs is unusually long and likely flexible (Carafoli & Hohenester, 2013). Finally, the DDR catalytic domain is a typical Tyr kinase domain with significant homology to the TrkA kinase (46% sequence identity) (Leitinger, 2014). While there is only a single DDR2 isoform, 5 DDR1 isoforms exist (DDR1a-e; **Figure 4**) (Leitinger, 2014). DDR1a and DDR1b are the most abundantly expressed isoforms and differ by a 37 amino acid addition in the DDR1b intracellular JM region (Alves et al, 1995). DDR1c contains an additional 6 amino acids within the kinase domain relative to DDR1b (Playford et al, 1996). DDR1d and DDR1e are non-functional receptors which lack the entire kinase domain, or parts of the JM region and ATP binding site, respectively (Alves et al, 2001). It is feasible that these catalytically incompetent receptors could dimerise with functional DDR1 isoforms (DDR1a-c) and facilitate in *trans* activation, as is seen when the catalytically dead ErbB3 receptor complexes with ErbB2 (Kung & Jura, 2019). However, this has not been documented.

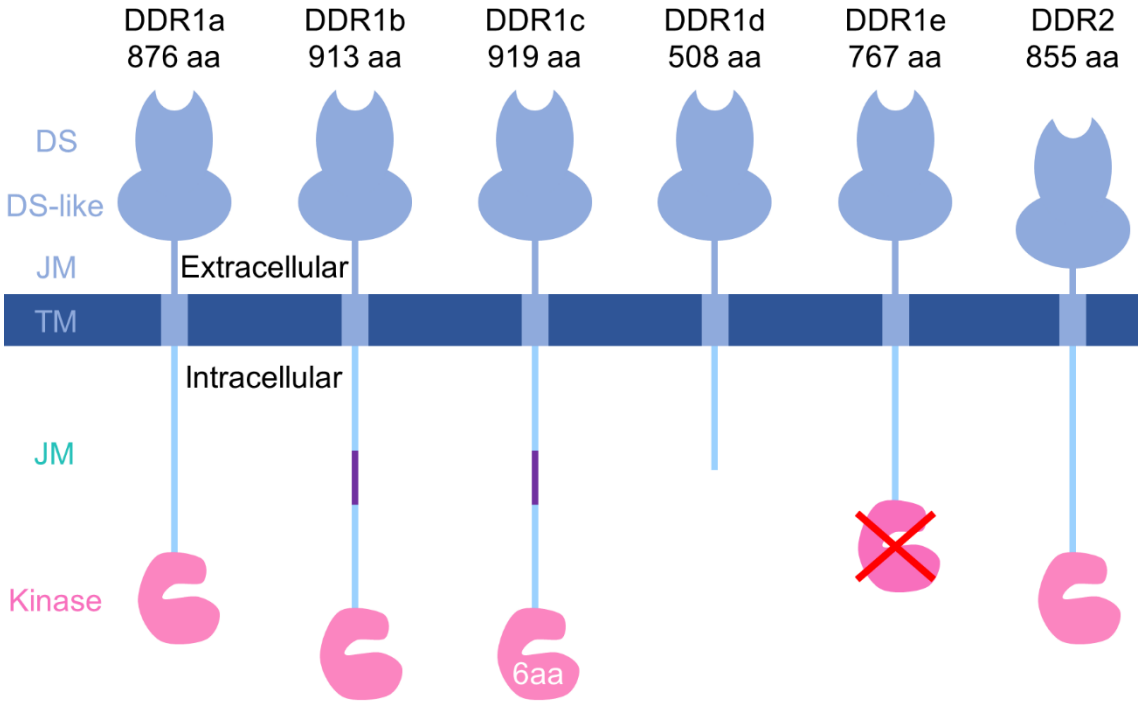

Figure 4. DDR1 and DDR2 structures.
A cartoon representation of the DDR molecular architecture. The extracellular region consists of a DS, DS-like domain, and JM region. The intracellular region consists of a JM region and kinase domain. The plasma membrane is depicted as a blue band with the TM domain shown in light blue. The 5 DDR1 isoforms (a-e) are depicted. DDR1b and DDR1c contain an additional 37 amino acid insert (purple) relative to DDR1a. DDR1c has an additional 6 amino acids (6aa) in the kinase domain relative to DDR1a and DDR1b. DDR1d does not contain a kinase domain and DDR1e contains a non-functional kinase domain (red cross). An adaptation of Figure 1 from (Leitinger, 2014).

1.2.2.1 Extracellular DDR domains

DDR1 and DDR2 have a significant degree of homology between their extracellular globular domains with 59% identity in the DS domain and 51% in the DS-like domain (Carafoli & Hohenester, 2013). The DS domain mediates collagen binding (Carafoli et al, 2009). Both DDR1 and DDR2 bind to the fibrillar type I-III collagens (Shrivastava et al, 1997; Vogel et al, 1997). However, they differ in their specificity towards the non-fibrillar collagens. For example, DDR1 is specific for collagen type IV, and DDR2 for collagen type X (Leitinger & Kwan, 2006; Shrivastava et al, 1997; Vogel et al, 1997). Interestingly, the DDRs do not bind denatured collagens, but require a native collagen fold (Shrivastava et al, 1997; Vogel et al, 1997). This fold consists of 3 polypeptide α-chains which form a triple helix with a single residue stagger (Brodsky & Persikov, 2005). This stagger means that equivalent residues in each strand of the triple helix are exposed to different environments (Carafoli & Hohenester, 2013). Each collagen chain is therefore designated as trailing, middle, or leading (Carafoli & Hohenester, 2013). Early mutagenesis studies of the DDRs demonstrated the importance of predicted loops within the DS domain in binding to collagen (Abdulhussein et al, 2004; Leitinger, 2003). Later, the minimal DDR collagen binding sequence in fibrillar collagens was identified as being GVMGFO (O being hydroxyproline) (Konitsiotis et al, 2008). The identification of this motif enabled the generation of collagen-mimetic peptides containing the DDR binding sequence (GVNleGFO; mutation of Met to Nle (norleucine) resulted in enhanced DDR2 binding) (Carafoli et al, 2009). This collagen mimetic peptide was co-crystallised with the DDR2 DS domain (**Figure 5a**) (Carafoli et al, 2009). This structure revealed that the DS domain forms an 8-stranded β-barrel composed of 2 anti-parallel β-sheets. The N- and C-termini of the DS domain are joined at the bottom of the barrel by a disulphide bond between Cys30-Cys185. The top of the barrel is characterised by 5 protruding loops, which had previously been predicted to be the collagen binding region (Abdulhussein et al, 2004; Leitinger, 2003). Indeed, the collagen triple-helical peptide was found to bind within an amphipathic trench formed by these loops. The binding motif (GVMGFO) of the leading and middle chains of the collagen-mimetic bind within this trench, whilst the trailing chain motif faces the solvent. The observation that two of the collagen chains are involved in binding demonstrates why denatured collagen cannot bind the DS domain (Shrivastava et al, 1997; Vogel et al,

1997). The residues located in the loops of the collagen binding trench are well conserved. However, those at the periphery are less well conserved. This may account for the different DDR specificities towards non-fibrillar collagens. Indeed, replacing 5 of these peripheral residues in DDR2 with their DDR1 equivalents resulted in a DDR2 DS domain which can bind to the DDR1 specific ligand, collagen type IV (Xu et al, 2011). Following the structure of the DDR2 DS domain bound to collagen, the structure of the DDR1 DS and DS-like domain pair was solved in complex with an antibody Fab fragment (Carafoli et al, 2012). Superposition of this structure on the DDR2 DS domain-collagen complex, reveals that there are few structural differences between the free and collagen bound DS domains (**Figure 5b**). This is also apparent when the ligand-free, solution structure of the DDR2 DS domain is compared with the DDR2 DS domain bound to collagen (Ichikawa et al, 2007). These structures therefore do not reveal how collagen binding triggers DDR activation.

The DS-like domain of DDR1 has a similar β-barrel fold despite only 9% sequence identity to the DS domain (**Figure 5a**) (Carafoli et al, 2012). The DS-like domain is also characterised by two N-linked glycans and a Ca^{2+} binding site. A short linker joins the DS and DS-like domains, which are related by a 90° rotation. The two domains form an interface between the bottom of the DS domain and one side of the DS-like domain. It is possible that structural rearrangement of this interface, following collagen binding, could be an important part of the DDR activation mechanism. The Fab fragment used for the crystallisation of this structure is an allosteric inhibitor of DDR activation (Carafoli et al, 2012). It binds to the DS-like domain and prevents kinase autophosphorylation without occluding collagen binding. The antibody must therefore inhibit the DDRs either through preventing structural reorientation of the ectodomains, or through preventing lateral associations of the DDRs at the cell surface (Carafoli et al, 2012). The extracellular JM region, which joins the DS-like domain to the TM region, is rich in Pro and likely unstructured. The extracellular JMs are poorly conserved between DDR1 and DDR2. However, they both contain predicted O- and N-linked glycosylation sites (Leitinger, 2014). Furthermore, in an extended configuration the JM would allow projection of the DDR globular domains up to 150 Å from the cell surface (Carafoli & Hohenester, 2013). This would enable

the DS domain to sample the available extracellular environment for collagen binding motifs.

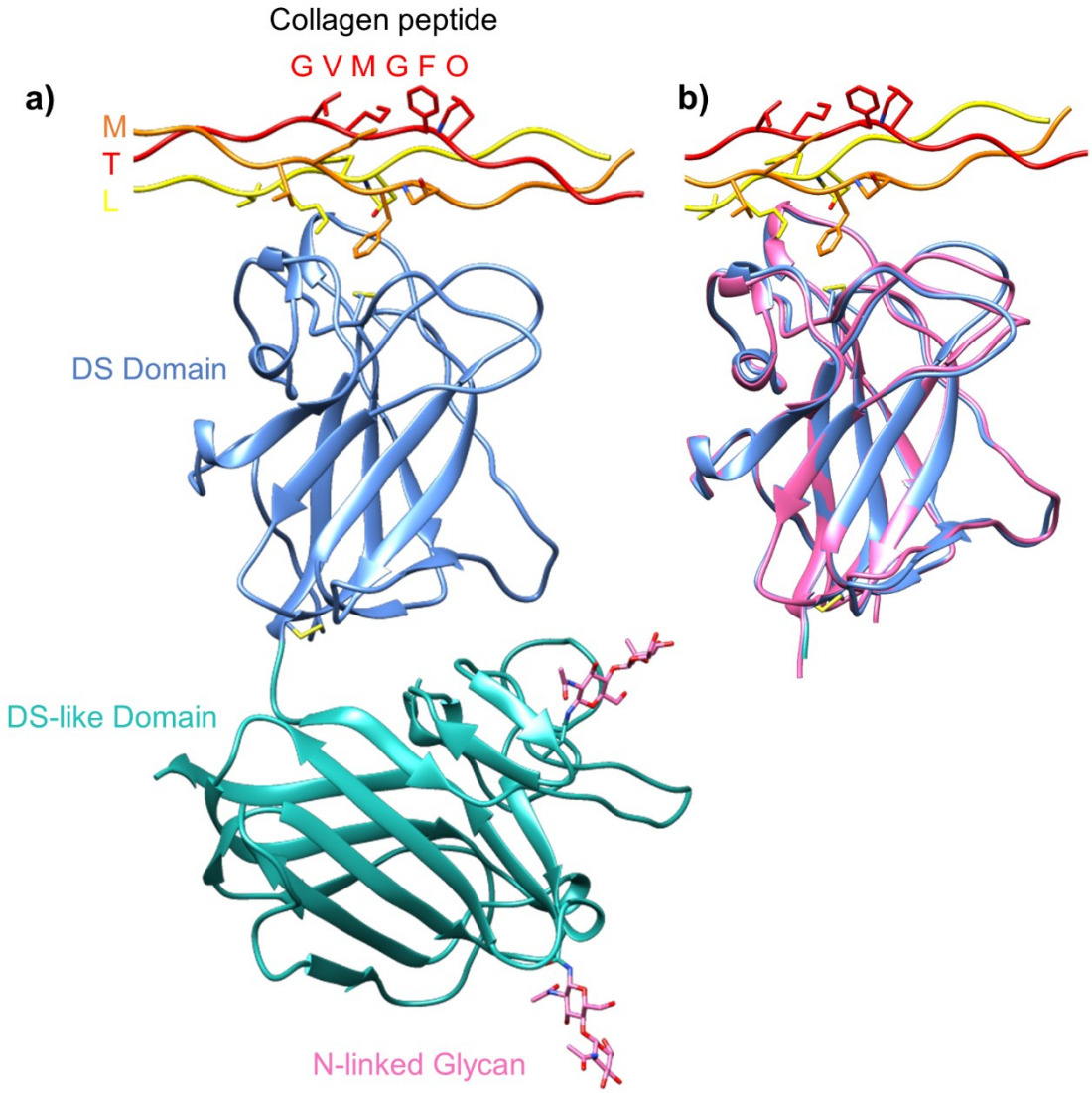

Figure 5. Crystal structure of the DDR ectodomains.
a) The DDR1 DS/DS-like domain (PDB:4AG4) was superposed onto the DDR2 DS domain in complex with a collagen peptide (PDB:2WUH) *(Carafoli et al, 2009; Carafoli et al, 2012)*. The collagen peptide from PDB:2WUH is shown with the trailing (T – red), middle (M – orange), and leading (L – yellow) chains. The GVNleGFO (GVMGFO) motif is shown for each collagen chain in atomic detail. The DDR1 DS (blue) and DS-like (light blue) domains are shown with disulphide bonds (Cys30-Cys185 and Cys177-Cys73) in yellow and N-linked glycans in pink. b) Superposition of the DDR1 (blue PDB:4AG4) and DDR2 (pink PDB:2WUH) DS domains with collagen peptide from PDB:2WUH shown with the same colouring and atomic detailing as in a).

1.2.2.2 Intracellular DDR domains

The intracellular DDR JM region is unusually long (~170 residues in DDR1b and 142 in DDR2) (Leitinger, 2014). Alignment of the DDR1 and DDR2 intracellular domains (**Figure 6**) allows segmentation of the DDR1 JM region into 4 sections (JM1-4). Sequence conservation within these segments is generally poor (**Figure 6**). However, significant homology can be seen in the kinase proximal segment, the JM4 region (89% sequence identity), which highlights this region as potentially important. The JM1 segment also shows some conservation (~48% sequence identity). At the N-terminus of the JM1 segment this conservation is dominated by positively charged residues. This high prevalence of positive amino acids at the membrane proximal JM region is common in RTKs (Hedger et al, 2015). It is believed that these residues interact with anionic phospholipid head groups on the intracellular leaflet of the cell surface (Hedger et al, 2015). These interactions may play important structural and functional roles in RTK signalling. The only other notable motif in the JM region is the DDR1b/c specific NPXY motif. NPXY motifs typically facilitate recruitment of proteins into clathrin-coated pits thereby enabling endocytosis (**Section 1.2.5**) (Bonifacino & Traub, 2003). Tyr513 in this motif is also a known DDR1 autophosphorylation site which recruits ShcA (**Figure 8**) (Leitinger, 2014; Shrivastava et al, 1997). There are 7 other potential autophosphorylation sites in the DDR1 JM region, 3 of which have been biochemically verified (Tyr484, Tyr513, Tyr520; **Figure 8**) (Iwai et al, 2013; Leitinger, 2014; Xu et al, 2014). For each of these Tyr residues, a series of potential interacting proteins have been identified (**Section 1.2.4**) (Iwai et al, 2013; Leitinger, 2014). Whether the intracellular JM region plays a functionally important role in DDR signalling, as is seen in many other RTKs (**Section 1.1.3.3**), has yet to be investigated.

Figure 6. Alignment of DDR sequences.
The intracellular and TM DDR1b and DDR2 sequences were aligned using Clustal Omega (excluding kinase domain residues 656-901 in DDR1 and 606-855 in DDR2). The TM region is shown in grey and the start of the kinase domain in pink. JM1 denotes the membrane-proximal segment. JM2 and JM3 denote the segments preceding and following the DDR1b/c-specific insert, respectively. JM4 denotes the kinase proximal JM segment and the two JM4 region Tyr residues, Tyr569 and Tyr586, are shown in red. Sequence numbers of the full-length proteins are shown to the right.

Several structures of inhibitor bound DDR1 kinase domain have been solved, along with one DDR2 kinase-inhibitor structure (Canning et al, 2014; Richter et al, 2019). They revealed that the DDR kinase domain has a typical bilobal fold with a small N-lobe and larger C-lobe (**Section 1.1.2**). The N-lobe contains the conserved Gly-rich loop, between the β1- and β2-strands (Leu616-Val629 in DDR1), and αC-helix (Lys663-Leu679 in DDR1). The DDR C-lobe contains the A-loop (Asp784-Glu813 in DDR1) which has 3 Tyr residues, and the catalytic loop (Leu760-Gly775 in DDR1). All crystal structures were solved in the inactive **D**FG-OUT conformation, regardless of the inhibitor type used (**Figure 7a**)(Hanson et al, 2019). This is the result of the DDR1 kinase domain being exceptionally stable in the inactive **D**FG-OUT conformation (Hanson et al, 2019). This characteristic also contributes towards the DDRs being highly promiscuous inhibitor binders (DDR1 binds >100 inhibitors) (Hanson et al, 2019). In this inactive conformation, the classic autoinhibitory interaction between the A-loop and the catalytic loop Asp is formed (Tyr796-Asp766; **Figure 7b**). The DDR kinase domain does not contain a C-terminal tail (Canning et

al, 2014; Leitinger, 2014). Instead, the C-terminal region is incorporated into an α-helix (αI-helix **Figure 7**). Therefore, the only regulatory motif identified to date for the DDR kinase domain is the autoinhibitory A-loop.

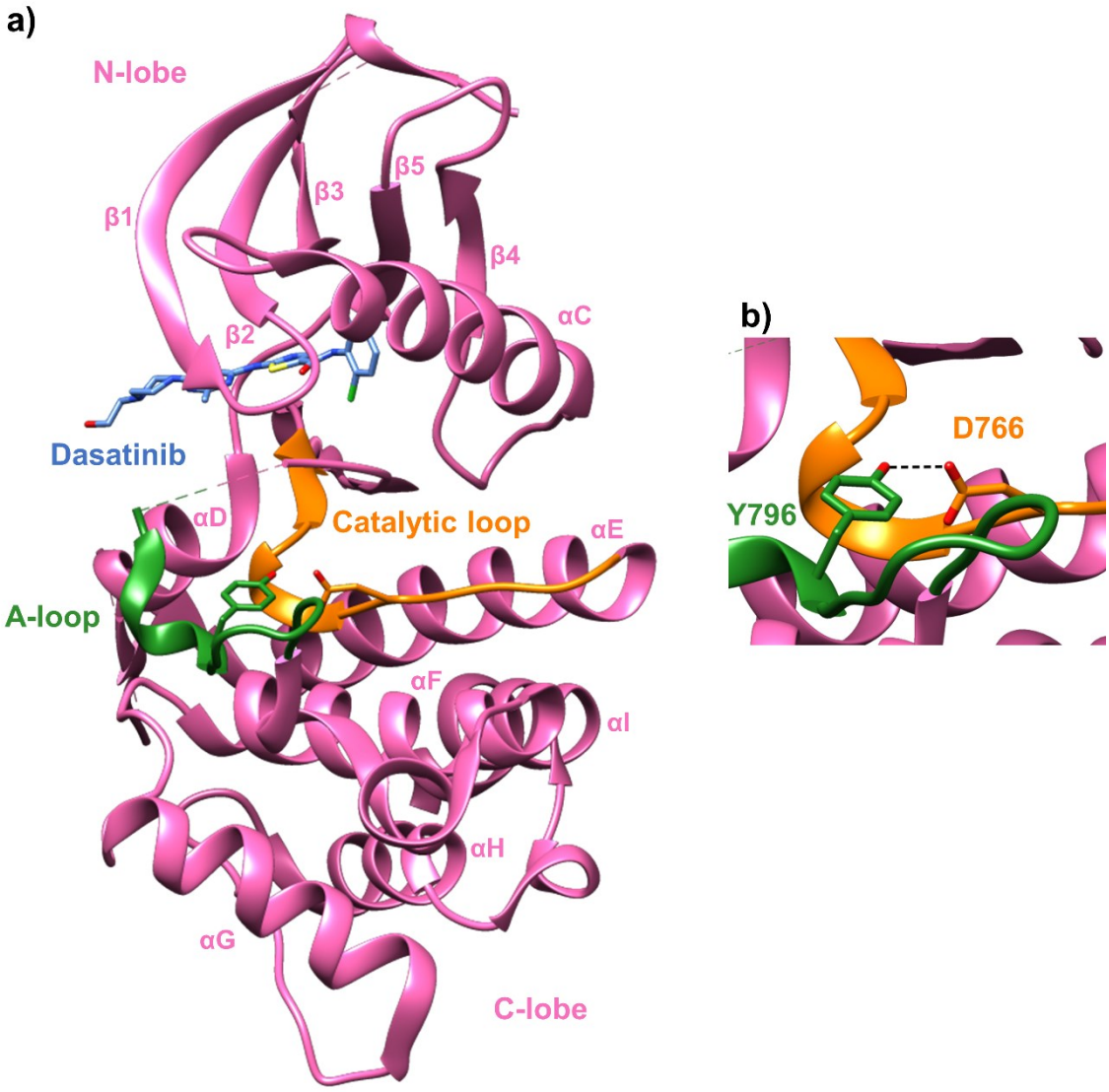

Figure 7. DDR1 kinase domain structure.
a) Crystal structure of the DDR1 kinase domain in complex with dasatinib (PDB:5BVW) *(Murray et al, 2015)*. The kinase domain is shown in pink with the A-loop in green and the catalytic loop in orange. Dasatinib is depicted in blue. The β-strands and α-helices are labelled, as are the N- and C-lobes. b) Atomic detail view of the interaction between Tyr796 (A-loop) and Asp766 (catalytic loop).

1.2.3 DDR activation

Typical RTKs are activated through ligand-induced dimerisation; however, the DDRs form stable, non-covalent interactions in the absence of collagen (Lemmon & Schlessinger, 2010; Noordeen et al, 2006). These dimers likely form early during biosynthesis and are mediated through multiple contacts in the ectodomains, intracellular JM segment, and the TM region (Noordeen et al, 2006). However, those contacts formed within the TM region appear to be the most significant. Indeed, a bacterial TOXCAT assay demonstrated that the DDR1 TM region forms a strong helix association via a Leu-based motif (Noordeen et al, 2006; Russ & Engelman, 1999). Furthermore, this TM region self-association was shown to be the strongest of any RTK (Finger et al, 2009). The insulin receptor is also a dimeric RTK. It is activated through ligand-induced conformational changes (Hubbard, 2013b; Scapin et al, 2018; Uchikawa et al, 2019; Weis et al, 2018). However, the ligand binding DDR DS domain shows limited structural change following collagen binding (Carafoli & Hohenester, 2013). Furthermore, large flexible insertions or deletions within the extracellular JM region have little effect on DDR activation (Xu et al, 2014). Coupling of conformational changes across the TM region to the intracellular kinase domains is therefore unlikely. These data suggest that the DDRs may instead be activated through the formation of higher order oligomers. Indeed, dense clusters of DDR1 dimers are observed at the cell surface following collagen binding (Corcoran et al, 2019; Yeung et al, 2019). These clusters reach a peak density following >60 minutes of collagen stimulation, matching the DDR activation kinetics. Furthermore, it is within these clusters that the DDRs have been found to phosphorylate receptor dimers in *trans* (Juskaite et al, 2017). Collagen binding therefore triggers DDR activation through clustering of receptors into large aggregates. This aggregation may facilitate activation through increasing the local concentration of the DDR kinase domains, thereby increasing the chance of *trans*-phosphorylation.

As collagens have the potential to form multivalent interactions it is attractive to assume that this is what drives DDR clustering (Kadler et al, 2007). However, small collagen-mimetic peptides can also facilitate receptor activation (Xu et al, 2011). Collagen binding must therefore trigger clustering through an alternative mechanism. As mentioned above, no significant conformational changes are observed in the DS

domain following ligand binding. However, it is not known whether collagen binding leads to a rearrangement of the DS-DS-like interface which may then facilitate oligomerisation. Interestingly, an allosteric inhibitor of DDR2 has a predicted binding site within the DS-DS-like interface (Grither & Longmore, 2018). Mutation of two residues predicted to be important in inhibitor binding within this interface (Phe96 and Thr98 to Ala) led to a loss of DDR2 activation and extracellular domain oligomerisation (Grither & Longmore, 2018). It is possible therefore that collagen binding triggers conformational changes in the DS-DS-like interface which enable receptor clustering.

There is evidence that the DDRs may require input from other kinases. For example, complete activation of DDR2 relies upon Src (Ikeda et al, 2002; Yang et al, 2005). Likewise, DDR1 may require Src for full activation (Dejmek et al, 2003; Lu et al, 2011). DDR interactions with Src may therefore be important during receptor activation.

1.2.4 DDR downstream signalling

The DDRs, like all RTKs, signal through the recruitment of SH2 or PTB domain containing proteins to phospho-Tyr residues on their intracellular domains (**Section 1.1.4**). The intracellular domains of DDR1 contain up to 15 Tyr residues; 8 in the kinase domain and 7 in the JM region (13 in DDR1a, 15 in DDR1b and DDR1c). DDR2 contains 14; 10 in the kinase domain and 4 in the JM region. To date however, only Tyr513 of the DDR1b/c JM region, and the A-loop Tyr residues (Tyr792, Tyr796, and Tyr797 of DDR1, and Tyr736, Tyr740, and Tyr741 of DDR2) have been validated experimentally as ligand-induced phospho-sites (Iwai et al, 2013; Lemeer et al, 2012; Xu et al, 2014). Pharmacologically-induced DDR1 autophosphorylation through phosphatase inhibition has also shown that Tyr484, Tyr513, and Tyr521 of the DDR1b/c JM region can be phosphorylated (Lemeer et al, 2012). It is possible that there are other, as yet unverified, Tyr autophosphorylation sites which could act as binding sites for downstream signalling proteins.

The proteins which are recruited to the DDRs following activation are currently poorly understood. However, several potential binding partners have been identified. DDR1 for example binds to ShcA, Nck2, SHP-1 (Src homology region 2 domain-containing

phosphatase-1), SHP-2, PI3-K (phosphoinositide 3-kinase), Csk, and STAT (Abbonante et al, 2013; Dejmek et al, 2003; Faraci-Orf et al, 2006; Koo et al, 2006; Vogel et al, 1997; Yang et al, 2009). Furthermore, using phospho-peptides of the possible DDR1 phospho-Tyr residues, numerous potential DDR1 signalling partners were identified (**Figure 8**) (Lemeer et al, 2012). Even less is currently understood about DDR2 signalling. To date, ShcA, SHP-2, Nck1, Lyn, and PLCL-2 (Phospholipase C like 2) have been shown to interact with activated DDR2 (Ikeda et al, 2002; Iwai et al, 2013). The precise interplay between these different signalling proteins and the cellular outcomes which are elicited by DDR signalling are currently poorly understood.

The signalling pathways which have been described downstream of DDR activation are ligand and cell-type dependent. For example, DDR1 leads to MAPK signalling and ERK1/2 (extracellular signal-regulated kinase 1/2) activation in smooth muscle cells, mammary epithelium, and megakaryocytes (Abbonante et al, 2013; Hilton et al, 2008; Lu et al, 2011; Ongusaha et al, 2003). However, in mesangial cells, DDR1 inhibits ERK1/2 activation (Curat & Vogel, 2002). Furthermore, collagen IV stimulation of DDR1 in human embryonic kidney cells leads to activation of the MAPK signalling cascade, but collagen I stimulation of the same cell type does not (Ongusaha et al, 2003). Likewise, collagen I stimulation leads to DDR1-mediated activation of Pyk2 in pancreatic cancer cells, but collagen XV stimulation inhibits this signalling pathway (Shintani et al, 2008). Other signalling pathways induced by DDR1 activation include chemoresistant signalling through NF-κB (nuclear factor-κB), PI3-K activation in several cell lines, and JNK (c-Jun N-terminal kinase) signalling in pancreatic cancer and adipose stromal cells (Das et al, 2006; Ongusaha et al, 2003; Shintani et al, 2008). DDR2 signalling pathways are even less well understood, but DDR2 has been shown to induce MAPK signalling and activation of ERK1/2 (Zhang et al, 2013). In chondrocytes, DDR2 activates ERK2 and p38 to induce MMP13 expression, or it uses p38 and JNK to stimulate IL-12 (interleukin-12) expression (Poudel et al, 2013; Xu et al, 2007; Xu et al, 2005).

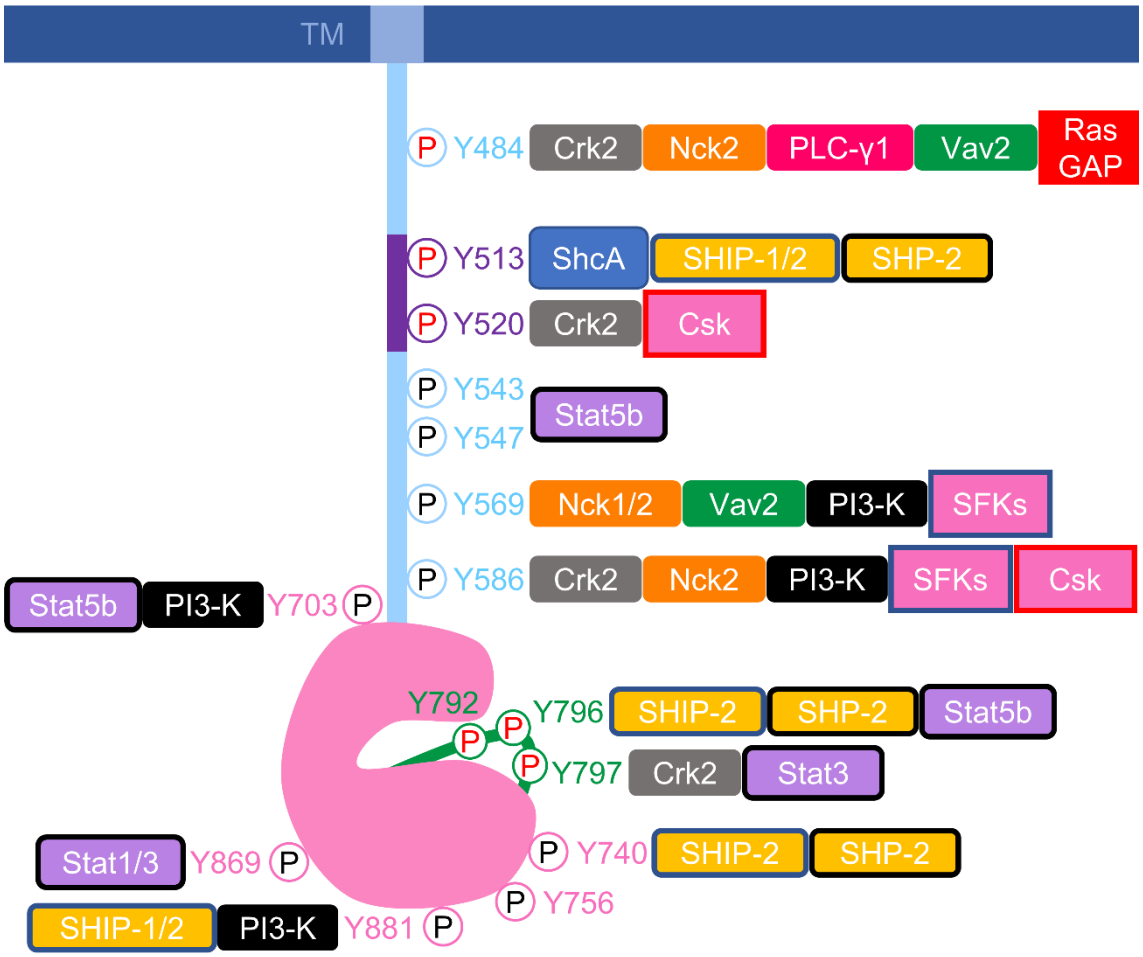

Figure 8. DDR1 phosphosite interaction map.
The residue numbering is based on the DDR1b sequence. Potential Tyr phospho-sites in the JM (blue), DDR1b/c-specific JM insert (purple), kinase domain (pink), and A-loop (green), are indicated by open circles. A red P indicates a biochemically verified phospho-site. Proteins which were found to bind at each of these sights using phospho-Tyr peptide pulldown experiments in human placental tissue are shown *(Lemeer et al, 2012)*. PLC-γ1; phospholipase C-γ1. Ras-GAP; Ras-GTPase activating protein. SHIP-1/2; SH2 containing inositol polyphosphate 5-phosphate-1/2. SHP-2; SH2 containing protein tyrosine phosphatase-2. SFKs; Src family kinases. Stat; signal transducer and activator of transcription. PI3-K; phosphoinositide 3-kinase. An adaptation of Figure 3 from (Leitinger, 2014).

Crosstalk between the DDRs and other cell surface receptors may also contribute towards signalling. For example, activation of IGF1-R and the insulin receptor leads to increased expression of DDR1 (Malaguarnera et al, 2015; Vella et al, 2017). Likewise, DDR1 overexpression leads to increased expression of IGF1-R and insulin receptor (Malaguarnera et al, 2015; Vella et al, 2017). This DDR1-insulin receptor crosstalk was shown to promote cell proliferation and invasiveness in breast cancer (Malaguarnera et al, 2015; Vella et al, 2017). Insulin and its receptors may therefore present a non-canonical activation mechanism for DDR1 (Vella et al, 2019). Another receptor class which may also signal in conjunction with the DDRs are the integrin receptor family. The crosstalk between DDRs and integrins appears to function in a cell type-dependent manner. In MDCK (Madin-Darby Canine Kidney) cells, DDR1 inhibits α2β1 mediated cell adhesion (Wang et al, 2006). However, in human embryonic kidney cells, DDR1 promoted integrin-mediated cell adhesion (Xu et al, 2012). DDR1 activation in NIH 3T3 cells also increased α2β1 and α5β1 activation in focal adhesions and promoted maturation of focal adhesions (Staudinger et al, 2013).

To summarise, several downstream signalling pathways and proteins have been implicated in DDR signalling. However, our current understanding of the ligand and cell-dependent nature of these pathways, alongside the phenotype elicited by their activation, is poor.

1.2.5 DDR regulation

As described in **Section 1.1.5,** RTKs are commonly regulated through endocytosis and lysosomal degradation following receptor activation (Goh & Sorkin, 2013). This is typically a rapid process (Goh & Sorkin, 2013). The atypically slow kinetics of DDR activation, alongside the observation that active DDR can be detected for days following collagen stimulation, makes it unlikely that the rapid RTK degradation pathway applies to the DDRs (Shrivastava et al, 1997; Vogel et al, 1997). Nevertheless, DDR1b and DDR1c both contain an NPXY motif which is a common RTK internalisation motif (**Section 1.1.5**) (Bonifacino & Traub, 2003). Indeed, internalisation of a DDR1b-YFP (yellow fluorescent protein) fusion protein was found to occur rapidly following collagen stimulation, prior to recycling of the receptor to the cell surface (Mihai et al, 2009). From these findings the authors proposed that

DDR1b becomes phosphorylated on intracellular vesicles, before being recycled to the cell surface. However, another study using untagged DDR1b demonstrated that the receptor clusters at the cell surface following stimulation, and it is within these cell surface clusters that autophosphorylation takes place (Corcoran et al, 2019). Further assessment of DDR endocytosis must therefore be conducted to clarify the disparity between these studies. While DDR1b and DDR1c contain the NPXY internalisation motif, DDR1a and DDR2 do not (**Figure 6**). This indicates a potential difference in internalisation mechanism between different DDR isoforms. Determining the precise isoform-specific function, if any, of DDR endocytosis on receptor function will be an important area for future research.

In addition to endocytosis, some cell surface receptors can be regulated through proteolytic cleavage of their extracellular domains (Clark, 2014). This ectodomain shedding regulates the pool of available competent receptor or may create functionally relevant ectodomain fragments (Clark, 2014; Higashiyama et al, 2011). Ligand stimulation of the DDRs leads to proteolytic cleavage of the ectodomain and extracellular accumulation of soluble DDR ectodomain (Slack et al, 2006; Vogel, 2002). However, constitutive DDR1 ectodomain shedding has also been reported (Alves et al, 1995). This constitutive cleavage is mediated by transmembrane MMPs, and not their homologous, secreted counterparts (Fu et al, 2013). Cleavage occurs within specific regions of the extracellular JM region (Ser387-Leu398 and Pro407-Val408). Constitutive DDR1 ectodomain cleavage likely enables clearance of DDR1 from the collagen exposed cell surface (Fu et al, 2013). However, it may also enable regulation of DDR1 signalling through cleavage of activated receptor. The precise role of DDR ectodomain shedding is currently unclear, so too is the fate of the cleaved fragments (Leitinger, 2014). Interestingly, the DDR1 DS domain still maintains collagen binding potential and may function to block collagen fibrillogenesis (Abbonante et al, 2013) in addition to full-length DDR1 signalling (Agarwal et al, 2007; Leitinger, 2003).

1.2.6 DDRs in disease

The DDRs have been implicated in numerous diseases including cancers, atherosclerosis, osteoarthritis, and numerous organ fibroses (Borza & Pozzi, 2014). These disease links have sparked an interest in the discovery of small-molecule

inhibitors of DDR activation, some of which have shown promising activity (Aguilera et al, 2017; Richter et al, 2019).

1.2.6.1 DDRs in disease - cancers

One of the most widely studied DDR disease implications is within the context of cancers (Valiathan et al, 2012). Several cancer types have been associated with the DDRs, including breast, non-small-cell lung carcinoma (NSCLC), ovarian, glioblastoma, and gastric cancers (Valiathan et al, 2012). Many of these associations have been derived from links between high DDR expression levels and cancer progression or poor prognosis. It is currently unclear what drives the overexpression of DDRs in these diseases. However, it could involve alterations in DDR suppressive miRNA expression, cancer-induced hypoxic environments, or the actions of cancer-promoting tetraspanins (Deng et al, 2017; Ren et al, 2014; Yang et al, 2017). Like many cancer associated RTKs, the DDRs can promote cancer progression through the stimulation of cell proliferation, migration, survival, and adhesion (Leitinger, 2014). They are also important regulators of collagen deposition and remodelling (Borza et al, 2017; Leitinger, 2014) (**Section 1.2.1**). It is these processes which likely lead to DDR promotion of cancer progression and metastasis (Guo et al, 2020).

High expression levels of DDR1 have been reported in NSCLC, glioblastoma, ovarian cancer, pancreatic cancer, oesophageal cancer, hepatocellular carcinomas, melanoma, and leukaemia (Couvelard et al, 2006; Ma et al, 2018; Nemoto et al, 1997; Quan et al, 2011; Reger de Moura et al, 2019; Shen et al, 2010; Tomasson et al, 2008; Yang et al, 2010). Furthermore, both DDR1 and DDR2 are overexpressed in breast cancer (Gadiya & Chakraborty, 2018; Ren et al, 2014). In all cases, these high expression levels are associated with poor patient prognosis and in some cases more aggressively metastatic phenotypes. As a result, DDR1 expression has been proposed as a potential biomarker for serous ovarian cancer, in which 63% of cancerous tissue expresses DDR1, compared with non-cancerous tissue in which DDR1 expression cannot be detected (Quan et al, 2011). DDR knockdown studies have supported these cancer associations. For example, siRNA knockdown of DDR1 reduces the malignant properties of melanomas (Reger de Moura et al, 2019). Likewise, shRNA knockdown of DDR1 in glioblastoma reduced cancer invasiveness (Ma et al, 2018). DDR overexpression therefore appears to promote cancer

progression and invasiveness. Overexpression of the DDRs likely leads to activation of the receptor. In support of this, DDR1 and DDR2 are among the most highly phosphorylated RTKs in NSCLC tumours (Rikova et al, 2007). The studies outlined above have all implicated the DDRs, and in particular DDR1, in cancer progression. However, there are contradictions in the literature. For example, high levels of DDR1 in breast cancer tissue is associated with increased metastasis and poor prognosis (Gadiya & Chakraborty, 2018). However, another study found that DDR1 expression in tissue surrounding the breast tumour suppressed cancer metastasis (Takai et al, 2018). This suggests that there may be tissue specific DDR1 functions in cancer, which could complicate pharmacological targeting of DDR1.

In addition to overexpression, DDR mutations have also been reported in cancerous cells. Numerous mutations have been reported for NSCLC (DDR1-A496S and DDR2-R105S), lung adenocarcinoma (DDR1-W385C, DDR1-F866Y, and DDR2-N456S), and acute myeloid leukaemia (DDR1-A803V, DDR1-T460M, DDR1-N502S, and DDR1-A544S) (Davies et al, 2005; Ding et al, 2008; Loriaux et al, 2008). These mutations are present throughout the DDR structure including in the DS-domain, extracellular and intracellular JM region, and the kinase domain. Few of these mutations have been followed up in functional experiments to discover their effect on DDR signalling. However, their prevalence indicates a functional role for the DDRs in these cancers. Indeed, DDR2 mutations have been detected in 4% of primary lung squamous cell cancers.

Kinase inhibitors have been successfully used in the treatment of numerous cancers (Abbaspour Babaei et al, 2016; Pottier et al, 2020; Zhang et al, 2009). Therefore, due to the extensive DDR cancer implications, there has been an increased interest in the use of small-molecule DDR inhibitors. Early studies demonstrated that the Bcr-Abl inhibitors imatinib, nilotinib, and dasatinib potently inhibit the DDRs with low nM IC_{50} values (Day et al, 2008; Hantschel et al, 2008). However, these inhibitors have poor kinase selectivity. More recently, specific DDR inhibitors have been reported: for example, molecule 7rh, DDR1-IN-1, and molecule 5N (Gao et al, 2013; Kim et al, 2013; Wang et al, 2018). Specific inhibition of DDR1 using molecule 7rh reduced tumorigenicity in pancreatic adenocarcinoma (Aguilera et al, 2017). Likewise, specific inhibition of DDR1 is beneficial in models of colorectal cancer, lung

adenocarcinoma, gastric carcinoma, and prostate cancer (Ambrogio et al, 2016; Azizi et al, 2019; Jeitany et al, 2018; Jin et al, 2018). As an alternative to small molecule inhibition of the DDRs, neutralising DDR antibodies are a promising avenue for research. Antibodies raised against the extracellular domains of DDR1 have been shown to potently inhibit DDR1 activation, likely through the prevention of receptor clustering (Carafoli et al, 2012; Corcoran et al, 2019). Interestingly, the use of neutralising DDR1 antibodies has also been shown to reduce breast cancer growth in mice (Zhong et al, 2019).

One of the biggest drawbacks in cancer management is the development of drug resistance (Guo et al, 2020). Recently, DDR1 was shown to facilitate drug resistance in glioblastoma through modulation of autophagy (Vehlow & Cordes, 2019). Inhibition of DDR1 facilitated sensitisation to chemotherapy by inducing autophagy (Vehlow & Cordes, 2019). Taken together, the studies outlined above demonstrate that DDR1 inhibition may be beneficial in numerous cancer types either through direct modulation of cancer progression, or through increasing sensitivity to chemotherapeutics.

1.2.6.2 DDRs in disease – atherosclerosis

Atherosclerosis is an inflammatory condition characterised by the formation of fibrous lipidic plaques which grow until they become vulnerable and eventually rupture (Libby et al, 2019). DDR1 has been implicated in the profibrotic and proinflammatory processes which culminate in atheroma formation (Dorison et al, 2017). DDR1 knockout mice have reduced neointimal formation following mechanical injury of carotid arteries compared with wild type animals (Hou et al, 2001). This protection in DDR1 -/- mice is a result of decreased proliferation, migration, and MMP production in smooth muscle cells (Hou et al, 2001; Hou et al, 2002). Additionally, knockdown of DDR1 in a mouse model of atherosclerosis (low-density lipoprotein receptor knock out) resulted in an accumulation of collagen in atherosclerotic plaques, but also a reduction in atherosclerotic lesion area (Franco et al, 2008). This was attributed to the decreased lipid accumulation which accompanies increased ECM production, which may have prevented plaque rupture (Franco et al, 2010; Franco et al, 2008). DDR1 knockout was also anti-inflammatory, reducing macrophage infiltration into plaques (Franco et al, 2009). In contrast, despite being expressed in smooth muscle

cells, knockout of DDR2 does not affect smooth muscle cell migration, proliferation, or MMP activity (Hou et al, 2012). This suggests that it is DDR1 and not DDR2 which functions to promote atherosclerosis.

1.2.6.3 DDRs in disease – osteoarthritis

Osteoarthritis is a common disease characterised by cartilage degeneration in joints (Xia et al, 2014). Clear indications for the involvement of DDR2 in the progression of osteoarthritis have come from the observation that DDR2 expression is increased in osteoarthritic patients and in mouse models (Xu et al, 2007). Furthermore, DDR2 expression levels have been associated with increased cartilage damage (Sunk et al, 2007). In support of this, heterozygous or homozygous knockout of DDR2 reduced cartilage degeneration in mice (Manning et al, 2016; Xu et al, 2010). DDR2 likely contributes towards osteoarthritis by increasing the expression of MMP13, a collagen II specific protease (Xu et al, 2005). Interestingly, knockout of DDR1 resulted in increased incidence of osteoarthritis and has been proposed as a potential model for the disease (Schminke et al, 2014). This suggests that DDR1 may be protective for osteoarthritis. However, DDR1 depletion was accompanied by an increased expression of DDR2 in chondrocytes (Schminke et al, 2014). Therefore, the development of osteoarthritis in DDR1 -/- mice may be a result of DDR2 function and not DDR1 depletion. These studies highlight the importance of DDR2 in osteoarthritis, but also demonstrate the need for inhibitors which can specifically target DDR2 to prevent potential off-target effects.

1.2.6.4 DDRs in disease – fibrosis

DDR1 has been implicated in kidney and lung fibrosis, whilst DDR2 has been implicated in lung and liver fibrosis (Borza & Pozzi, 2014; Moll et al, 2019). In the kidney, DDR1 knockout in mice reduces type I and type IV collagen expression in renal vasculature during angiotensin II mediated renal injury (Flamant et al, 2006). Furthermore, DDR1 knockout reduced the levels of infiltrating cells in the renal vessels and glomeruli (Flamant et al, 2006). These factors led to improved renal function and highlighted DDR1 as a major mediator of glomerular inflammation and injury. Indeed, when glomerulonephritis is induced using alloimmune sheep nephrotoxic serum, DDR1 expression is increased 17-fold in glomeruli (Kerroch et

al, 2012). The use of DDR1-specific anti-sense oligonucleotides, or a DDR1-specific inhibitor, was protective for renal function in this disease model (Kerroch et al, 2012; Moll et al, 2018). In the lung, DDR1 knockdown is protective against bleomycin induced lung injury (Avivi-Green et al, 2006). Furthermore, knockdown reduces myofibroblast expansion, apoptosis, and inflammation (Avivi-Green et al, 2006). Pharmacological inhibition of DDR1 also reduces ECM deposition, and the expression of fibrotic markers (α-smooth muscle actin, collagen, and fibronectin) in the same injury model (Wang et al, 2016). These studies demonstrate that DDR1 is also a mediator of lung inflammation and fibrosis and highlight the potential therapeutic benefits of targeting DDR1. DDR2 may also promote lung fibrosis through myofibroblast activation (Zhao et al, 2016). Indeed, DDR2-specific siRNA, or pharmacological inhibition, reduced lung fibrosis in a bleomycin injury model (Zhao et al, 2016). Finally, the role for DDR2 expression and activation in the liver is controversial. In acute liver injury, DDR2 expression is elevated resulting in hepatic stellate cell activation and proliferation (Olaso et al, 2001). Likewise, in an alcoholic liver disease model, DDR2 expression is elevated and associated with increased collagen deposition and remodelling (Luo et al, 2013). Targeting of DDR2 with siRNA in this liver disease model resulted in decreased liver damage (Luo et al, 2013). In contrast however, in a carbon tetrachloride injury model, DDR2 deletion was found to aggravate hepatic fibrosis demonstrating increased collagen deposition and hepatic stellate cell density (Olaso et al, 2011). Although the role for DDR2 in liver disease is currently unclear, the DDRs are clearly implicated in a range of organ-specific fibrotic conditions, making them attractive therapeutic targets in these diseases which currently have poor treatment options.

1.3 Hypothesis and aims

Despite the broad disease implications (**Section 1.2.6**) and significant efforts into the discovery of small molecule inhibitors, little is currently known about the processes underpinning DDR activation (Kim et al, 2013; Leitinger, 2014; Liu et al, 2017; Wang et al, 2018). Structural analysis of the DDR extracellular domains has revealed how these receptors bind to collagen but did not demonstrate how collagen binding leads to activation (Carafoli et al, 2009; Carafoli et al, 2012; Ichikawa et al, 2007; Konitsiotis et al, 2008; Xu et al, 2011). Recent work has shown that collagen stimulation leads to the formation of receptor clusters, within which DDR1 is autophosphorylated in *trans* (Corcoran et al, 2019; Juskaite et al, 2017). However, the regulatory mechanisms controlling DDR1 activation are poorly understood. This project aimed to understand these mechanisms through the structural and functional analysis of the intracellular domains of DDR1. It was hypothesised that the long intracellular DDR1 JM region, of unknown function, plays a role in regulating DDR1 signalling. The highly conserved kinase proximal JM4 segment is hypothesised to be the mediator of this regulation (**Figure 6**).

1.3.1 Chapter aims

Chapter 3: to structurally analyse the DDR1 kinase domain and JM region using X-ray crystallography.

Chapter 4: to use purified recombinant protein constructs and enzymology to functionally analyse DDR1 activation, and JM region function, *in vitro*.

Chapter 5: to assess the function of the DDR1 JM region in the full-length cell expressed receptor.

Much of the data presented in this thesis are published in PNAS (Sammon et al, 2020).

Chapter 2. Materials and methods

2.1 Chemicals, reagents, and buffers

2.1.1 Reducing agents

Tris-2-carboxyethyl-phosphine hydrochloride (TCEP) and dithiothreitol (DTT) were from ThermoFisher. Sodium bisulphite was from Sigma-Aldrich.

2.1.2 Buffering agents and salts

Trizma base (Tris), HEPES (4-2-hydroxyethyl-1-piperazineethanesulfonic acid), imidazole, Bis-Tris, MOPS (3-N-morpholinopropanesulfonic acid), potassium dihydrogen phosphate (KH_2PO_4), sodium phosphate (Na_2HPO_4), and monosodium phosphate (NaH_2PO_4) were from Sigma-Aldrich.

Sodium chloride (NaCl), manganese chloride ($MnCl_2$), magnesium chloride ($MgCl_2$), calcium chloride ($CaCl_2$), zinc chloride ($ZnCl_2$), sodium fluoride (NaF), sodium iodide (NaI), potassium chloride (KCl) were from Sigma-Aldrich.

2.1.3 Acids, bases, and alcohols

Acetic acid, hydrochloric acid, sodium hydroxide, ammonium hydroxide, ethanol, methanol, and isopropanol were from Sigma-Aldrich.

2.1.4 Molecular biology reagents

Sub-cloning Efficiency DH5α competent cells were from ThermoFisher, 5-alpha highly competent cells were from NEB, and Stellar competent cells were from Clontech.

Lysogeny broth (LB) agar plates, LB-broth, and the antibiotics; ampicillin, chloramphenicol, and kanamycin were from Sigma-Aldrich.

QIAprep spin Miniprep, Midiprep, and Maxiprep kits were from Qiagen.

CloneAmp HiFi PCR Premix, and the In-Fusion enzyme were from Takara.

T4 DNA Ligase and T4 DNA Ligase Buffer were from Invitrogen.

Sequencing and cloning primers were from Sigma-Aldrich or Invitrogen.

All restriction enzymes and buffers were from NEB.

6X purple DNA gel loading dye was from NEB. UltraPure agarose, SYBR Safe, and 1kB Plus DNA Ladder (range 0.5-10 kb) were from Invitrogen.

QIAquick Gel Extraction Kit was from Qiagen, whilst the NucleoSpin Gel and PCR Clean-up kit was from Macherey-Nagel.

TAE buffer: 40 mM Tris-HCl pH 8.5, 1.3 mM EDTA (ethylenediaminetetraacetic acid; Sigma-Aldrich), 0.1% (v/v) acetic acid.

2.1.5 Gel running, blotting, and staining reagents

Bromophenol blue, β-mercaptoethanol, tetramethylethylenediamine (TEMED), ammonium peroxodisulfate (APS), 30% acrylamide/bis-acrylamide solution, sodium dodecyl sulphate (SDS), glycine, and glycerol were from Sigma-Aldrich.

Pre-stained protein markers (190-11 and 250-11 kDa) were from NEB and were included in all SDS-PAGE (polyacrylamide gel electrophoresis) gels.

5X SDS sample buffer: 300 mM Tris-HCl pH 6.8, 10% (w/v) SDS, 50% glycerol, 0.1% bromophenol blue, 10% β-mercaptoethanol.

2X SDS sample buffer: 120 mM Tris-HCl pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 0.04% (w/v) bromophenol blue, 4% (v/v) β-mercaptoethanol.

SDS-PAGE resolving gel: 7.5-10% (v/v) acrylamide, 375 mM Tris-HCl pH 8.8, 0.1% (w/v) SDS, 0.05% (w/v) APS, 0.25% (v/v) TEMED.

SDS-PAGE stacking gel: 3% (v/v) acrylamide, 125 mM Tris-HCl pH 6.8, 0.1 % (w/v) SDS, 0.05% (w/v) APS, 0.25% (v/v) TEMED.

SDS running buffer: 25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS.

SDS blotting buffer: 25 mM Tris, 192 mM glycine, 20% methanol, 0.01% (w/v) SDS.

Phosphate buffered saline (PBS): 80 mM Na_2HPO_4 pH 7.4, 137 mM NaCl, 18 mM KH_2PO_4, 27 mM KCl.

InstantBlue Coomassie stain was from Expedeon, while Pierce Silver Staining Kit was from ThermoFisher.

2.1.6 Protein purification buffers

2.1.6.1 Buffers for small scale protein expression tests

Small scale Sf9 lysis buffer: 50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, 10 mM imidazole, 1% (v/v) Tween-20 (ThermoFisher), with 400 Kunitz U/mL DNase I (NEB), and 1 protease inhibitor cocktail tablet per 50 mL (EDTA free; Roche).

Ni-NTA wash buffer: 50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, 20 mM imidazole, 0.05% (v/v) Tween-20.

2.1.6.2 Buffers for large scale protein production

Sf9 lysis buffer: 50 mM Tris-HCl pH 7.5, 500 mM NaCl, 0.2% (v/v) Tween-20, 0.05 U/mL Benzonase (Sigma-Aldrich), and 1 protease inhibitor cocktail tablet per 50 mL (EDTA free; Roche).

HisTrap wash buffer: 50 mM Tris-HCl pH 7.5, 500 mM NaCl.

HisTrap elution buffer: 50 mM Tris-HCl pH 7.5, 500 mM NaCl, 500 mM imidazole.

C3 cleavage and dialysis buffer: 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.5 mM TCEP.

Ion exchange wash: 50 mM Tris-HCl pH 8.0, 150 mM NaCl.

Ion exchange elution: 50 mM Tris-HCl pH 8.0, 1 M NaCl.

Gel filtration buffer: 20 mM Na-HEPES pH 7.5, 200 mM NaCl, 1 mM TCEP.

2.1.7 Crystallographic reagents

96-23ll crystallisation screening plates were provided by Dr. Marc Morgan in the X-ray Crystallography Facility at Imperial College London. The screens included: Structure Screens 1 and 2, PACT premier, JCSG+, Proplex, Morpheus, PGA screen, and BCS screen from Molecular Dimensions. PEG/Ion 1 and 2, Index, SaltRx 1 and 2, and Natrix 1 and 2 from Hampton Research. Wizard-1, -2, -3 and -4, and Wizard Cryo from Rigaku Reagents/Molecular Dimensions.

Polyethylene glycol 3350, ammonium sulphate, and polyacrylic acid 5100 sodium salt were from Sigma-Aldrich. AMP-PNP (adenylyl-imidodiphosphate) was from Tocris.

2.1.8 Kinase assay buffers

Adenosine triphosphate (ATP) was from Promega (for use in ADP-Glo™), or Sigma-Aldrich.

Kinase buffer I: 25 mM Tris-HCl pH 7.5, 200 mM NaCl, 10 mM $MnCl_2$, 10 mM $MgCl_2$, 100 µM DTT, 0.1 mg/mL BSA (bovine serum albumin).

Kinase buffer II: 25 mM Tris-HCl pH 7.5, 200 mM NaCl, 20 mM $MnCl_2$, 20 mM $MgCl_2$, 100 µM DTT.

Kinase buffer III: 25 mM Tris-HCl pH 7.5, 200 mM NaCl, 5 mM $MnCl_2$, 1 mM $MgCl_2$, 100 µM DTT, 0.1 mg/mL BSA.

IP Kinase buffer: 25 mM Tris-HCl pH 7.5, 200 mM NaCl, 10 mM $MnCl_2$, and 10 mM $MgCl_2$.

2.1.9 Peptides and inhibitors

Axltide peptide substrate (sequence KKSRGDYMTMQIG) was from GenScript and is based on IRS-1 (residues 979-989). JM4 peptide mimetics were custom synthesised by GenScript (see **Section 2.15** for further details).

Dasatinib was from Selleckchem (Cat#S1021, batch S102108), and was dissolved in dimethyl sulfoxide (DMSO; Sigma-Aldrich).

2.1.10 Cell culture reagents

BSA was obtained from ThermoFisher (BP1605), or New England Biolabs (NEB; B9001S).

Collagen I (acid-soluble from rat tail; C-7661) was purchased from Sigma-Aldrich.

Human embryonic kidney (HEK) 293 cells and Cos-7 monkey fibroblast-like kidney cells were from the American tissue culture collection, Manassas, VA, Canada.

Dulbecco's modified Eagle's medium/F12 nutrient mixture (DMEM/F12; Invitrogen) was supplemented with 2 mM L-glutamine, 100 units/mL penicillin, 100 µg/mL streptomycin (ThermoFisher) and 10% (v/v) foetal bovine serum (FBS; Invitrogen).

Opti-MEM medium and Lipofectamine 3000 was from ThermoFisher.

Trypsin-EDTA solution, and Cell Dissociation Buffer (PBS) were from GibcoTechnologies.

Cell freezing medium: 90% (v/v) FBS, 10% (v/v) DMSO.

2X HEPES buffer saline (HBS): 50 mM Na-HEPES pH 7.1, 275 mM NaCl, 10 mM KCl, 1.5 mM Na_2HPO_4, 10 mM glucose.

Cell lysis buffer: 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% (v/v) Nonidet P-40, 1 mM phenylmethylsulfonyl fluoride, 50 µg/ml aprotinin, 1 mM sodium orthovanadate, and 5 mM NaF. All lysis buffer components were from Sigma-Aldrich.

ProLong Gold Antifade Mountant was from ThermoFisher.

Paraformaldehyde and Triton-X100 were from Sigma-Aldrich.

2.2 Antibodies

2.2.1 Commercial antibodies

All antibodies used in this report are highlighted in **Table 2** and **Table 3.** Rabbit polyclonal anti-DDR1 IgG (Santa Cruz; SC-532, Lot #10314), and mouse monoclonal antibody (mAb) anti-DDR1 IgG2b (Santa Cruz; SC-374618, Lot#10413), antibodies were used for the detection of total DDR1 protein levels. Mouse mAb anti-DDR1 7A9 (Carafoli et al, 2012) was used for immunoprecipitation, flow cytometry, and immunofluorescence. The following DDR1-specific anti-phosphotyrosine antibodies were used to assess DDR1 phosphorylation: rabbit mAb anti-phospho-DDR1-pY796 (anti-pY796; R&D Systems; MAB25382), and rabbit mAb anti-phospho-DDR1-pY513 (anti-pY513; Cell Signalling; 145315). These antibodies bound specifically to the DDR1 residues outlined in their name only when in the phosphorylated state; Tyr796 and Tyr513, respectively (Uniprot-Q08345, DDR1b variant amino acid numbering used throughout report). Mouse mAb anti-phosphotyrosine 4G10 (anti-pY; EMD Millipore; 05-321, Lot#3015310) was used for the general detection of phosphorylated Tyr residues. Secondary antibodies used for Western blotting were IRDye 680RD donkey anti-mouse IgG (LI-COR; 926-68070, Lot#C80619-05), and IRDye 800CW donkey anti-rabbit IgG (LI-COR; 926-32213, Lot#C80125-15). Fluorescein isothiocyanate (FITC)-conjugated goat anti-mouse IgG antibody (Sigma-Aldrich; F-9006) was used for flow cytometry. AlexaFluor 488 anti-mouse IgG1 (ThermoFisher; A21121, Lot#1964382) and AlexaFluor 555 anti-rabbit IgG (ThermoFisher; A21428, Lot#1670185) were used for immunofluorescence imaging.

2.2.2 Custom-made antibodies

As there are no commercially available JM4-specific anti-phosphotyrosine antibodies, the following antibodies were commissioned from Biomatik: rabbit polyclonal anti-phospho-DDR1-pY569 (anti-pY569), and rabbit polyclonal anti-phospho-DDR1-pY586 (anti-pY586). The peptide used for anti-pY569 production was C-QNSVPH(pY)AEDIVT, whilst that for anti-pY586 was NT(pY)AVPALPPGAVGD. Polyclonal antibodies were produced in rabbits, purified, and then their specificity for the phospho-peptide over the unphosphorylated peptide was assessed. These experiments were conducted by Biomatik using dot blot testing and ELISA (enzyme-

linked immunosorbent assay). Further analysis was conducted here through analysis of antibody reactivity with collagen I stimulated and unstimulated full-length DDR1b (**Section 2.18**). The linearity of phosphorylation detection by Western blotting using these antibodies was also assessed (**Section 2.9**).

Table 2. Primary antibody list.
Applications: WB – Western blot. IP – immunoprecipitation. Flow – flow cytometry. IF – immunofluorescence.

Primary antibody	Supplier	Source	Dilution factor	Application
Polyclonal anti-DDR1 (SC-532); C-terminal epitope	Santa-Cruz	Rabbit	1:500	WB
mAb anti-DDR1 (SC-374618); C-terminal epitope	Santa-Cruz	Mouse	1:500	WB
mAb anti-DDR1 (7A9); N-terminal epitope	(Carafoli et al, 2012)	Mouse	2-10 µg/mL	IP Flow IF
mAb anti-phospho-DDR1 (anti-pY796)	R&D systems	Rabbit	1:666	WB
mAb anti-phospho-DDR1 (anti-pY513)	Cell Signalling	Rabbit	1:500-666	WB IF
mAb anti-pY (4G10)	EMD Millipore	Mouse	1:666	WB
Polyclonal anti-phospho-DDR1 (anti-pY569)	Biomatik	Rabbit	1:666	WB
Polyclonal anti-phospho-DDR1 (anti-pY586)	Biomatik	Rabbit	1:500	WB
Polyclonal anti-Src	Cell Signalling	Rabbit	1:500	WB

Table 3. Secondary antibody list.
Applications: WB – Western blot. Flow – flow cytometry. IF – immunofluorescence.

Secondary antibody	Supplier	Source	Dilution factor	Application
IRDye 680 RD anti-mouse IgG (926-68070)	LI-COR	Donkey	1:17,500	WB
IRDye 800 CD anti-rabbit IgG (926-32213)	LI-COR	Donkey	1:17,500	WB
FITC-anti-mouse IgG (F-9006)	Sigma-Aldrich	Goat	1:500	Flow
AlexaFluor 488 anti-mouse IgG1	ThermoFisher	Goat	1:500	IF
AlexaFluor 555 anti-rabbit IgG	ThermoFisher	Goat	1:500	IF
DAPI	Sigma Aldrich	-	0.2 µg/mL	IF

2.3 DNA constructs and site directed mutagenesis

2.3.1 Plasmid DNA maintenance

A list of all plasmid constructs used throughout this thesis and their source is highlighted in **Table 4**. Plasmid cDNA constructs were transformed into Sub-cloning Efficiency DH5α competent cells. This was achieved by incubating 10 ng of DNA with 50 µL of E.*coli* for 30 minutes on ice. Cells were then heated to 42°C for 45 seconds before returning to ice for 2 minutes (heat-shock). Cells were grown overnight at 37°C on LB-agar plates containing 100 µg/mL ampicillin. Single colonies were then picked and grown in LB-broth containing 100 µg/mL ampicillin overnight at 37°C, 200 rpm (R100/TW, Luckham). Cultures were pelleted the following morning by centrifugation at 1000 xg for 20 minutes, 4°C, and QIAprep spin Miniprep (5 mL culture), Midiprep (50 mL culture), or Maxiprep (100 mL culture) kit used according to the manufacturer's instructions to purify the DNA. DNA was eluted and resuspended in EB buffer (Qiagen). DNA concentration was then recorded on a Nanodrop Spectrophotometer ND-1000 using NanoDrop ND-1000 version 3.2.1 software.

2.3.2 Restriction digestion

Plasmids were analysed through restriction enzyme digestion to confirm the expected DNA size, or to excise DNA fragments. For this, 500 ng of DNA was incubated with 10 U of restriction enzymes (the same DNA:restriction enzyme ratio was used for larger DNA amounts) and the appropriate buffer for 1 hour at 37°C. DNA was then mixed with 1X purple DNA gel loading dye (6X stock) and loaded onto a 0.8% (w/v) UltraPure agarose gel dissolved in TAE buffer (**Section 2.1.4**) containing a 1:10,000 dilution of SYBR Safe. Samples were separated at 80 V for 40-60 minutes alongside a 1kB Plus DNA Ladder (range 0.5-10 kb). DNA was then visualised using a UV transilluminator on a Gel doc and ImageLab software (BioRad).

2.3.3 DNA Sequencing

Sanger sequencing was performed by GENEWIZ UK. 100 ng of plasmid DNA was sent, along with 5 µL of 5 µM primer per reaction run. Sequence quality was checked in SnapGene (GSN Biotech LLP) and then aligned against the expected sequence using EMBOSS Needle Pairwise Sequence Alignment (EMBL-EBI).

2.3.4 Site directed mutagenesis

The mutant constructs pRK-DDR1b-Y569E, -Y586E, and -Y569E/Y586E, were produced using the QuikChange site-directed mutagenesis protocol (Stratagene). For pRK-DDR1b-Y569E/Y586E this was done in 2 steps as both mutations could not be included in a single primer. The protocol for QuikChange mutagenesis is provided below with the pRK-DDR1b-Y569E/Y586E construct as an example. Initially, the single point mutant, Y569E, was produced using the template cDNA of wild type DDR1b cloned into the pSP72 vector (pS-DDR1b - **Table 4**). To achieve this, 0.3 µM of forward and reverse primers (**Table 5**) were mixed with 50 ng of the template DNA and CloneAmp HiFi PCR Premix as per the manufacturer's instructions (25 µL reaction volume: 12.5 uL 2X PCR premix with 12.5 uL DNA/primer). Primers were designed to be 25-35 bases long, with at least 40% GC content including one or more C or G's at both ends, and a melting temperature greater than 78°C (**Table 5**). The polymerase chain reaction (PCR) cycle was then as follows: 30 secs 98°C, 15 sec 55°C, 2 min 72°C. This was repeated 25 times before a final elongation step of

72°C for 10 minutes. PCR products were then cleaved with 20 U of Dpn1 for 1 hour at 37°C. The resulting DNA constructs were transformed into NEB 5-alpha highly competent cells using the heat-shock method (**Section 2.3.1**). Cells were then grown on LB-agar plates and colonies picked and grown in LB-broth as in **Section 2.3.1**. QIAprep spin Miniprep kit was then used according to the manufacturer's instructions to purify the DNA. Constructs were subsequently sequenced by GENEWIZ UK to verify the introduction of the desired mutation (**Section 2.3.3**). The mutant construct (pS-DDR1b-Y569E) was then used as a template to produce the double mutant pS-DDR1b-Y569E/Y586E using the same methodology as above.

Following the successful introduction of mutations into the pSP72 vector the cDNA constructs were cloned into the mammalian expression vector, pRK5 (BD Biosciences). The pRK5 vector used for sub-cloning contained a truncated DDR1b sequence (pRK-MDN - **Table 5**).1 µg of insert (pS-DDR1b-Y569E/Y586E) and vector (pRK-MDN) plasmids were digested with 20 U EcoRI-HF and XbaI (NEB) for 1 hour at 37°C in the presence of CutSmart buffer (NEB). Restriction digestion products were then mixed with Gel Loading Dye and run on a 0.8% (w/v) agarose gel containing SYBR safe at 80 V for 40 minutes along with a 1kB DNA ladder (**Section 2.3.2**). DNA was excised from the gels under a low intensity UV lamp using a scalpel. The gel was weighed and purified using the QIAquick Gel Extraction Kit according to manufacturer's instructions. Inserts were ligated into the cleaved pRK5 vector at a 1:3 molar ratio (insert:vector) using 0.04 U/mL T4 DNA Ligase and T4 Ligase Buffer overnight at 16°C. Control ligation reactions were also performed in the absence of insert DNA. The ligation reactions were then transformed into NEB 5-alpha highly competent *E. coli* and minipreps performed (**Section 2.3.1**). DNA sequencing was then performed (**Section 2.3.3**) to confirm the presence of the expected mutation. Following this, the QIAfilter plasmid Maxi Kit was used to generate larger preparations of the mutant DNA constructs using the protocol described in **Section 2.3.1**. All mammalian expressions were performed using the constructs contained within the pRK5 vector.

The deletion constructs pRK5-DDR1b ΔJM1, ΔJM2, ΔJM3, ΔJM4, and ΔJM1-3, were all produced previously by Jasmine Gratton (**Table 4**) using the PCR FastCloning method (Li et al, 2011) with pSP72 DDR1b CDNA as a template. The following

residues were eliminated: Arg445 to Arg475 (DDR1-ΔJM1); Pro476 to Ser504 (DDR1-ΔJM2); Ala542 to Ser565 (DDR1-ΔJM3); Val566 to Leu591 (DDR1-ΔJM4); and Arg445 to Ser565 (DDR1-ΔJM1-3). cDNAs encoding mutant DDR1b were cloned into the mammalian expression vector pRK5 for transient transfection of HEK293 cells (as above).

2.3.5 In-Fusion cloning

For baculovirus expression of soluble DDR1 protein, constructs were cloned into a pOPINF vector, provided by Oxford protein production facility (OPPF) (**Table 4)**. The pOPINF vector allows expression of proteins in E. *coli*, mammalian cells (e.g. HEK293), and insect cells (e.g. *Spodoptera frugiperda* (Sf9)). It also encodes a His_6-tag followed by a human rhinovirus 3C protease cleavage site 5' of the DDR1 DNA insert site (**Figure 9**). This allows purification through nickel affinity chromatography, and subsequent cleavage of the His_6-tag by 3C protease. The pOPINF constructs produced were: pOPINF-DDR1-JM4K (DDR1 residues Val566-Val913), pOPINF-DDR1-Y569F/Y586F-JM4K (DDR1 residues Val566-Val913), pOPINF-Y569E/Y586E-JM4K (DDR1 residues Val566-Val913), pOPINF-12aa-JM4K (DDR1 residues Gly554-Val913), and pOPINF-Kinase (DDR1 residues Pro601-Gly913). The constructs encoding wild type DDR1 were produced by OPPF, the DDR1 mutants were produced here (**Table 4**). To achieve this cloning, the mutant pRK-DDR1b-Y569F/Y586F and pRK-DDR1b-Y569E/Y586E DNA constructs were used as templates for In-Fusion cloning (Takara). The pOPINF vector was initially linearised using HINDIII-HF and Kpn-1 in the presence of NEBuffer 2.1 prior to separation on an agarose gel, and purification as in **Section 2.3.4**. The cDNA encoding residues Val566-Val913 of DDR1 was PCR amplified from the mutant DDR1 constructs in the pRK vector using CloneAmp HiFi PCR Premix (**Section 2.3.4**). The primers used for amplification contained 18-20 base pair 5' overlap with the linearised vector (**Table 5**). PCR cycles and reagents were as outlined in **Section 2.3.4**. The PCR products were digested with 20 U of Dpn1 for 1 hour at 37°C before purification using NucleoSpin Gel and PCR Clean-up kit according to the manufacturer's instructions. 100 ng of linearised vector was then incubated with 50 ng of the PCR product and 1X In-Fusion enzyme in a 10 µL reaction volume for 15 minutes at 50°C. 2.5 µL of the reaction mixture was transformed into Stellar

competent cells, with heat-shock transformations and minipreps performed as in **Section 2.3.1**. Introduction of DDR1 DNA into the vector was subsequently confirmed through analytical digestion with Xba1 and analysis on an agarose gel prior to imaging on a Gel Doc (**Section 2.3.2**). Sequencing was then performed by GENEWIZ UK (**Section 2.3.3**).

Table 4. List of plasmid constructs used and their sources.
A. Cheung, V. Juskaite, J. Gratton, and Z. Hao are/were students in the Leitinger lab. M. Faure – SUGEN. R. Marais – Institute of Cancer Research, London.

Plasmid name	Plasmid source
pRK-DDR1b	M. Faure
pRK-MDN (truncated DDR1b)	M. Faure
pS-DDR1b	B. Leitinger
pRK-DDR1b-Y569E	D. Sammon
pRK-DDR1b-Y586E	D. Sammon
pRK-DDR1b-Y569E/Y586E	D. Sammon
pRK-DDR1b-Y569F	A. Cheung
pRK-DDR1b-Y586F	A. Cheung
pRK-DDR1b-Y569F/Y586F	A. Cheung
pRK-DDR1b-Y513F	V. Juskaite
pRK-DDR1b-K655M	V. Juskaite
pRK-DDR1b-ΔJM1	J. Gratton
pRK-DDR1b-ΔJM2	J. Gratton
pRK-DDR1b-ΔJM3	J. Gratton
pRK-DDR1b-ΔJM1-3	J. Gratton
pRK-DDR1b-ΔJM4	J. Gratton
pOPINF	OPPF
pOPINF-DDR1-12aa-JM4K	OPPF
pOPINF-DDR1-JM4K	OPPF
pOPINF-DDR1-Kinase	OPPF
pOPINF-DDR1-Y569E/Y586E-JM4K	D. Sammon
pOPINF-DDR1-Y569F/Y586F-JM4K	D. Sammon
pEF-c-Src-Myc	R. Marais
pEF-c-Src-Y527F-Myc	R. Marais

Table 5. Primer sequences used for DDR1 construct generation.
Primers for point mutations were designed based on QuikChange site-directed mutagenesis kit (Stratagene) protocol. Mutagenic bases are shown in red. Primers for In-Fusion cloning contained 18-20 base pair homology with the HindIII and KpnI cleaved pOPINF vector (homology bases shown in green).

Name	Final vector	Cloning technique	Forward primer Reverse primer
DDR1-Y569E	pRK5	QuikChange	GAACAGCGTCCCCCATGAGGCC GAGGCTGACATTG
			CAATGTCAGCCTCGGCCTCATG GGGGACGCTGTTC
DDR1-Y586E	pRK5	QuikChange	CCGGGGGCAACACCGAGGCTG TGCCTGC
			GCAGGCACAGCCTCGGTGTTGC CCCCGG
Y569E/ Y586E-JM4K	pOPINF	In-Fusion	AAGTTCTGTTTCAGGGCCCGGT CCCCCATGAGGCCGAGG
			ATGGTCTAGAAAGCTTTACACC GTGTTGAGTGCATCCTC
Y569F/ Y586F-JM4K	pOPINF	In-Fusion	AAGTTCTGTTTCAGGGCCCGGT CCCCCATTTTGCCGAGG
			ATGGTCTAGAAAGCTTTACACC GTGTTGAGTGCATCCTC

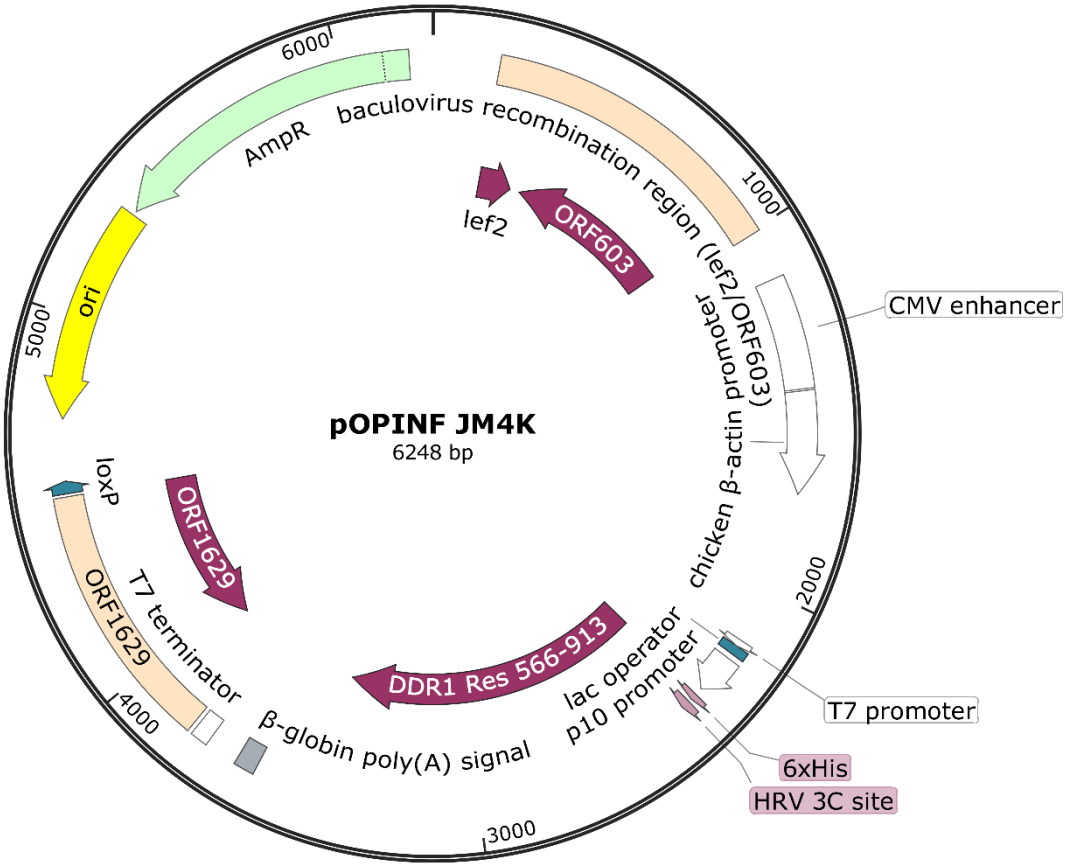

Figure 9. pOPINF vector map containing cDNA encoding JM4K.
Key components of the pOPINF vector are depicted along with the DDR1 JM4K cDNA which was inserted through In-Fusion cloning (DDR1 residues 566-913). Features include; the His_6 tag and 3C cleavage site 5' of the In-Fusion inserted DNA, and the ORF1629 baculovirus gene used for recombination. Vector map was produced using SnapGene.

2.4 Expression of soluble DDR1 kinase protein constructs

2.4.1 Bacmid purification

The bacmid used for expression of soluble DDR1 kinase constructs was provided by Prof. Ian Jones (The University of Reading). The bacmid was an *Autographa californica* multicapsid nucleopolyhedrovirus genome which lacked part of the essential gene ORF 1629. This lack of a complete ORF 1629 prevents virus replication in insect cells, while a bacterial artificial chromosome insert enables bacterial cell maintenance (Zhao et al, 2003). The bacmid was grown on LB-agar plates containing 30 µg/mL chloramphenicol and 50 µg/mL kanamycin. A colony was

picked and used to inoculate 15 mL of LB-broth containing the same antibiotics. Cells were pelleted at 1000 xg for 20 minutes at 4°C and then resuspended in 900 µL of buffer P1 from the QIAprep spin Miniprep kit (Qiagen). The resuspended cells were then split into 3 tubes and 300 µL of buffer P2 from the same kit was added and cells left for 5 minutes at room temperature. The lysate was then neutralised by the addition of 300 µL buffer P3 and cells placed on ice for 10 minutes. Samples were spun at 14,000 xg for 10 minutes at 4°C. The supernatant was removed and mixed with 800 µL isopropanol and incubated on ice for 10 minutes. DNA was then pelleted by spinning at 14,000 xg for 30 minutes at 4°C before being washed in 500 µL 70% (v/v) ethanol and pelleted again by centrifugation at 14,000 xg for 15 minutes at 4°C. The DNA was then left to dry for 5 minutes at room temperature before being resuspended overnight in 100 µL of 25 mM Tris-HCl pH 8.0 at 4°C.

2.4.2 Bacmid cleavage

Following purification, the bacmid was cleaved by restriction digestion to increase amenability to recombination. A 100 µL aliquot of bacmid produced using the method outlined in **Section 2.4.1** was incubated with 60 U Bsu36I in CutSmart buffer (NEB) for 6 hours at 37°C. The sample was heated to 72°C for 20 minutes to deactivate the restriction enzyme. DNA concentration was recorded using a NanoDrop (**Section 2.3.1**).

2.4.3 Baculovirus production and viral amplification

For baculovirus production, 1 µg of linearised bacmid and 750 ng of pOPINF-DDR1 vector (**Section 2.3.5**) were co-transfected into Sf9 cells plated at $0.5x10^6$ cells/mL in a 6-well plate (ThermoFisher) in InsectXpress medium (Lonza), using Cellfectin II (ThermoFisher) according to the manufacturer's instructions. Sf9 stocks were cultured and maintained by Val Good at the Baculovirus Facility (Imperial College London). After 7 days at 27°C in static culture, the medium containing P_0 virus was harvested from the wells. 100 µL of the medium containing P_0 virus was used to infect 10 mL of cells at $1x10^6$ cells/mL in a 125 mL Erlenmeyer flask. Cells were then incubated at 27°C, 130 rpm for 7 days in an INFORS HT Multitron Shaker, before being pelleted at 3000 xg for 15 minutes at 4°C. The supernatant containing P_1 virus was then harvested. The virus was amplified a second time by adding 400 µL of the

P_1 virus containing medium to 50 mL of Sf9 cells at $1x10^6$ cells/mL in a 250 mL Erlenmeyer flask and leaving for a further 7 days at 27°C, 130 rpm. Cells were pelleted as above and the supernatant containing P_2 virus harvested. For long term storage of virus, 2% (v/v) FCS (Foetal Calf Serum - GibcoTechnologies) was added to virus containing medium and stored in aliquots (2-5 mL) at -80°C.

2.4.4 Small scale protein expression testing

To determine whether viruses produced using the method described in **Section 2.4.3** were able to produce the desired protein constructs upon infection in Sf9 cells, a small scale expression screen was conducted using the P_2 virus. To that end, 10 mL of Sf9 cells were cultured at $1x10^6$ cells/mL in a 125 mL Erlenmeyer flask in the presence of 10 or 100 µL of the P_2 virus containing medium produced in **Section 2.4.3**. After 3 days of incubation at 27°C, 200 rpm in an INFORS HT Multitron Shaker, 1 mL of cells was pelleted at 14,000 xg for 1 minute at 4°C, and the pellets frozen at -80°C for at least 1 hour. Cells were then defrosted and resuspended in 210 µL of Small scale Sf9 lysis buffer (**Section 2.1.6.1**). Cells were then incubated for 30 minutes at room temperature, before being pelleted through centrifugation at 14,000 xg for 10 minutes. During this spin, nickel-nitrilotriacetic acid (Ni-NTA) agarose beads (Qiagen) were washed 3 times with water (spin at 3000 xg for 1 minutes, discarding the wash after each spin). Beads were then resuspended with an equal volume of PBS to create a 50% bead slurry. The cleared cell lysates were incubated with 40 µL of the 50% Ni-NTA bead solution for 30 minutes at room temperature. Beads were then washed 3 times in Ni-NTA wash buffer (**Section 2.1.6.1**) before being boiled for 5 minutes at 100°C in 20 µL 2X SDS sample buffer. Lysates were then run on a 10% SDS-PAGE gel in SDS running buffer (**Section 2.1.5**) at 100 V for 2 hours in a BioRad tank with a BioRad PowerPac Basic power supply. Gels were subsequently stained with ~20 mL InstantBlue Coomassie overnight. Gels were then washed in distilled water 3 times for 10 minutes before being imaged on an Odyssey FC imager (LI-COR) using the 700 nm channel and ImageStudio software (LI-COR). ImageStudio lite (LI-COR – V 5.2) was used for image analysis.

The viruses produced through this method allowed the expression of a number of wild type and mutant DDR1 constructs (**Table 6**). These were: **JM4K** which contained DDR1 residues 566-913 (JM4 region linked to the DDR1 kinase domain).

12aa-JM4K which contained DDR1 residues 554-913 (12 amino acids N-terminal of the JM4 region, JM4 and the DDR1 kinase domain as in JM4K). **Kinase alone** which contained DDR1 residues 601-913 (DDR1 kinase domain alone). **Y569E/Y586E-JM4K** which contained DDR1 residues 566-913 (JM4 region with kinase domain) and both JM4 Tyr residues (at positions 569 and 586) mutated to Glu. **Y569F/Y586F-JM4K** which contained DDR1 residues 566-913 (JM4 region with kinase domain) and both JM4 Tyr residues (at positions 569 and 586) mutated to Phe.

Table 6. Baculovirus expressed soluble DDR1 kinase constructs.

Construct name	Residues	Regions	Sequence
JM4K	566-913	JM4 + DDR1 kinase	Wild Type
12aa-JM4K	554-913	12 amino acids of JM3 + JM4 + DDR1 kinase	Wild Type
Kinase alone	601-913	DDR1 kinase alone	Wild Type
Y569E/Y586E-JM4K	566-913	JM4 + DDR1 kinase	Y569E/Y586E
Y569F/Y586F-JM4K	566-913	JM4 + DDR1 kinase	Y569F/Y586F

2.4.5 Large scale protein production

Large scale protein expression was performed by adding 1 µL of the P_2 virus containing media (**Section 2.4.3**) per mL of Sf9 cell suspension at $1x10^6$ cells/mL in a total volume of 0.5–3 L. Cells were incubated at 27°C, 130 rpm for 3 days in Corning Roller Bottles (500 mL/bottle) in an INFORS HT Multitron Shaker, before pelleting the cells by centrifuging at 3000 xg for 15 minutes at 4°C. Cell pellets were then frozen at -80°C for at least 1 hour. The cell pellet was then resuspended in Sf9 lysis buffer on ice (**Section 2.1.6.2**). Cell lysates were then sonicated for 10 minutes on ice (9 seconds on 9 seconds off, 50% amplitude using a Fisherbrand™ Q500 Sonicator with Probe; ThermoFisher). Lysates were then centrifuged at 30,000 xg for 1 hour at 4°C before being filtered through a 0.45 µm filter (ThermoFisher). The

cleared lysate was then loaded onto a 1 mL HisTrap HP column (GE Healthcare) using an Äkta pure chromatography system (GE Healthcare) at 4°C, 1 mL/min flow rate. Ultraviolet (UV) absorbance at 280 nm was measured by the Äkta and provided a measure of protein elution from the column. The column was washed with several column volumes (~10 mL) of HisTrap wash buffer (**Section 2.1.6.2**). HisTrap wash buffer was combined with HisTrap elution buffer (**Section 2.1.6.2**) to clean the column with 50 mM Tris-HCl pH 7.5, 500 mM NaCl, 15 mM imidazole (3% elution buffer) for several column volumes (~10 mL). Protein was then eluted in 100% HisTrap elution buffer. The His_6-tag was removed by adding 1 U of human rhinovirus 3C protease (ThermoFisher) per 0.1 mg of protein. Protein concentration was determined using a Nanodrop Spectrophotometer ND-1000 using NanoDrop ND-1000 version 3.2.1 software. The protein sample and 3C protease mixture was dialysed in 10 kDa molecular weight cut off SnakeSkin dialysis tubing (ThermoFisher) into 5 L of C3 cleavage and dialysis buffer (**Section 2.1.6.2**), overnight at 4°C. The buffer was exchanged once during dialysis (2x5L).

The following morning the dialysed solution was passed over a 1 mL HisTrap HP column at 4°C, 1 mL/min flow rate, and protein lacking the His_6-tag collected in the flow-through. Contaminating protein and the cleaved His_6-tag which bound to the column was then eluted in HisTrap elution buffer. HisTrap flow-through was then passed over a 1 mL HiTrap Q XL column (ion exchange (IEX) column; GE Healthcare) at 4°C, 1 mL/min flow rate, and the DDR1 protein collected in the flow through. The IEX column was washed of contaminating protein using ion exchange elution buffer (**Section 2.1.6.2**). The flow through from the IEX column was concentrated using a 10 kDa molecular weight cut off Vivaspin filtration device (Sartorius). The final purification step was done by size exclusion chromatography using a Superdex 200 Increase 10/300 column (GE Healthcare), with a 500 µL sample loop, and gel filtration buffer (**Section 2.1.6.2**) as the running buffer, at room temperature and a flow rate of 0.75 mL/min. Fractions containing protein were concentrated to 5–10 mg/mL, as above, and snap-frozen in liquid nitrogen. The final yield was ~2 mg of DDR1 protein per litre of cell culture.

2.5 Protein crystallography

2.5.1 DDR1-dasatinib crystallography

Attempts to crystallise the insect cell expressed soluble DDR1 JM4K (residues 566-913) and 12aa-JM4K (residues 554-913) protein constructs (see **Section 2.4.5** for expression protocol) failed when screening the free enzyme. Therefore, to stabilise protein conformation, the 12aa-JM4K protein was incubated at 8 mg/mL (200 µM) with a 1.5-fold molar excess of dasatinib. Addition of dasatinib to the protein resulted in the formation of a precipitate which disappeared following mixing at 4°C for 30 minutes. Crystal screening was done using a Mosquito robot (TTP Labtech) and a range of commercial screens (**Section 2.1.7**) in 96 well 2 Drops Sitting Drop crystallisation plates (MiTeGen). Crystals grew as thin needles at 4°C in a 400 nL sitting drop containing 200 nL of protein-inhibitor complex and 200 nL of reservoir solution consisting of 20% (w/v) polyethylene glycol 3350 and 200 mM NaI. Crystals were cryo-protected in 30% (v/v) glycerol, 20% (w/v) polyethylene glycol 3350 and 200 mM NaI and mounted on MiTeGen loops before freezing in liquid nitrogen. Diffraction data were collected at 100 K on Diamond Light Source beamline i24. A high-redundancy data set was collected from a single crystal and processed using the automatic XIA2-DIALS pipeline (Winter et al, 2013; Winter et al, 2018). The resolution limit was determined using the $CC_{1/2}$ criterion (Karplus & Diederichs, 2012) as implemented in AIMLESS (Evans & Murshudov, 2013). The structure was solved by molecular replacement using PHASER (McCoy et al, 2007) and PDB entry 5BVW (Murray et al, 2015) as a search model. As a control, dasatinib was omitted from the search model. COOT (Emsley & Cowtan, 2004) and PHENIX (Liebschner et al, 2019) were then used for model building and refinement.

2.5.1.1 Silver stain analysis of crystals

The JM4 region was not resolved in the 12aa-JM4K/dasatinib co-crystal structure, therefore, it was important to establish whether the JM4 was still present within the crystals used for X-ray diffraction data collection. To that end, crystals were fished from wells using MiTeGen loops, washed in reservoir solution (polyethylene glycol 3350 and 200 mM NaI) twice, and then boiled for 5 minutes at 100°C in SDS sample buffer. Crystals were subsequently separated by 10% SDS-PAGE at 100 V for 2

hours at room temperature in SDS running buffer (**Section 2.1.5**) using a BioRad tank and BioRad PowerPac Basic Power Supply. Gels were subsequently analysed with Pierce™ Silver Stain Kit (ThermoFisher) according to the manufacturer's instructions. Gels were imaged using a GelDoc imaging system (BioRad).

2.5.2 Y569F/Y586F-JM4K crystallography

Following the unsuccessful attempts to resolve the JM4 region in crystal structures of wild type constructs, the soluble DDR1 Y569F/Y586F-JM4K (residues 566-913) construct was used for crystallographic trials. The protein was purified as described in **Section 2.4.5** and concentrated to 9.5 mg/mL. Crystallisation trials of the free enzyme were then conducted as described in **Section 2.5.1**. Crystals used for X-ray diffraction data collection were obtained at 4°C in a 400 nL sitting drop containing 200 nL of protein and 200 nL of reservoir solution consisting of 100 mM Tris-HCl pH 8.0, 1.5 mM ammonium sulphate. Crystals were picked using MiTeGen loops and cryoprotected with 30% (v/v) glycerol, 100 mM Tris-HCl pH 8.0, 1.5 mM ammonium sulphate, before freezing in liquid nitrogen. Diffraction data were collected at 100 K on Diamond Light Source beamline i24, followed by auto-processing, molecular replacement and refinement as described in **Section 2.5.1**. COOT (Emsley & Cowtan, 2004) and PHENIX (Liebschner et al, 2019) were then used for model building and refinement. The refined structure was deposited in the Protein Data Bank (PDB code: 6Y23).

2.6 Protein molecular mass determination

To determine the oligomeric status of purified soluble DDR1 protein constructs, size exclusion chromatography-multiple angle light scattering (SEC-MALS) experiments were conducted. 500 µL of the protein sample at a concentration of 2 mg/mL was injected onto a Superdex 200 Increase 10/300 column (GE Healthcare) with a 500 µL loop attached. The column was connected to an Agilent 1260 Infinity LC system (Agilent Technologies). Gel filtration buffer (**Section 2.1.6.2**) was used as the running buffer with a flow rate of 0.07 mL/min at room temperature. Light scattering and refractive index changes were monitored using in-line Wyatt Mini Dawn and Optilab T-rEX detectors (Wyatt Technology Corp). The data were analysed with the Wyatt ASTRA V software V 6.1.1.

2.7 Differential scanning fluorimetry

Differential scanning fluorimetry experiments were conducted in gel filtration buffer (**Section 2.1.6.2**). Protein was analysed at 10 µM in the presence of 1:1000 SYPRO Orange (ThermoFisher) in an assay volume of 25 µL in 96 well PCR plates (ThermoFisher). Protein was heated from 24 to 94°C using a Stratagene Mx3005P teal-time PCR machine and MxPro software (Agilent). Data were subsequently analysed using a local macro originally obtained from ftp://ftp.sgc.ox.ac.uk/pub/biophysics and GraphPad Prism 5.0 using the Boltzmann equation. Later differential scanning fluorimetry experiments on purified DDR1 phospho-forms were conducted as above but using an Applied Biosystems 7500 fast real-time PCR machine, and Protein Thermal Shift software (ThermoFisher) for analysis.

2.8 Autophosphorylation of soluble kinase proteins

Soluble DDR1 protein constructs at 31.3–1000 nM concentration were incubated for up to 2 hours at 20°C with 1 mM ATP in kinase buffer I (**Section 2.1.8**). Alternatively, soluble DDR1 protein constructs at 100 µM concentration were incubated for up to 3 hours at 20°C with 20 mM ATP in kinase buffer II (**Section 2.1.8**). Samples were either boiled for 5 minutes at 100°C in SDS sample buffer and analysed by 10% SDS-PAGE followed by Western blotting (**Section 2.9**), or they were separated on a 7.5% native PAGE gel which was subsequently stained with InstantBlue Coomassie stain (**Section 2.10**).

2.9 Western blotting

Proteins were transferred onto nitrocellulose membrane (Amersham Protran 0.2 µm) in SDS blotting buffer (**Section 2.1.5**) on ice at 200 mA for 1 hour using a BioRad tank and BioRad PowerPac Basic Power Supply, following separation by native PAGE (**Section 2.10**) or SDS-PAGE. All SDS-PAGE gels were run at 100 V for up to 2 hours at room temperature in SDS running buffer (**Section 2.1.5)** using a BioRad tank and BioRad PowerPac Basic Power Supply, unless otherwise stated. Following transfer, membranes were dried for one hour at room temperature, before being rehydrated in distilled water and blocked for 1 hour at room temperature in Odyssey

Blocking Buffer (LI-COR). The membranes were then incubated with the indicated anti-phosphotyrosine antibody overnight at 4°C. All antibodies were diluted in Odyssey Blocking Buffer supplemented with 0.2% (v/v) Tween-20 (**Table 2**. 1:666 dilution for anti-pY569, anti-pY (4G10) and anti-pY796; 1:500 dilution for anti-pY586). Membranes were then washed 4 times for 5 minutes in PBS containing 0.1% (v/v) Tween-20. Fluorescent secondary antibodies were then incubated with membranes for 1 hour at room temperature in the dark (**Table 3**. 1:17,500 dilution). Membranes were then washed a further 4 times for 5 minutes in PBS containing 0.1% (v/v) Tween-20 in the dark. The membranes were imaged on an Odyssey FC imaging system using the 700, or 800 nm channel (LI-COR). Total DDR1 protein was detected using mouse or rabbit anti-DDR1 (**Table 2**. 1:500 dilution) with wash steps, secondary antibody incubations, and imaging carried out as above. Western blot quantification and image generation were performed using Image Studio Lite (LI-COR; V5.2).

2.9.1 Linearity of Western blot detection

To determine the linear range of detection for the primary antibodies used in Western blotting, the soluble DDR1 JM4K construct was incubated at 100 μM concentration for 6 hours at 20°C with 20 mM ATP in kinase buffer II (**Section 2.1.8**). A range of protein amounts (12.5–200 ng) was then loaded and analysed by 10% SDS-PAGE and Western blotting with each of the primary antibodies (**Table 2**). The blots were imaged and quantified as described in **Section 2.9**. Following this, the optimal amount of protein to load was found to be 35 ng for detection with the anti-pY796, anti-pY569, and anti-pY (4G10) anti-phosphotyrosine antibodies, and 50 ng for the anti-pY586 antibody. The anti-DDR1 antibodies also produced a linear response at these protein load levels.

2.10 Native gel electrophoresis

Native PAGE gels were produced with 7.5% acrylamide in the absence of SDS (**Section 2.1.5**). 2.5 μg of protein sample was mixed with sample buffer, without β-mercaptoethanol and SDS, and loaded onto gels. Gels were run at 90 V for 3 hours at 4°C in running buffer (25 mM Tris, 192 mM glycine) using a BioRad tank and BioRad PowerPac Basic Power Supply. Gels were subsequently stained with

InstantBlue Coomassie stain overnight at room temperature, and then washed 3 times for 10 minutes at room temperature in distilled water and imaged on the Odyssey FC imager (LI-COR) using the 700 nm channel and ImageStudio software. ImageStudio lite (LI-COR; version 5.2) was used for image analysis.

2.11 Phostag™ acrylamide gel electrophoresis

Phostag™ (Alpha Laboratories) is a proprietary compound which binds to phosphorylated ions. When cast into an acrylamide gel it enables separation of differently phosphorylated forms of a protein (Bekesova et al, 2015). To prepare Phostag™ containing gels, 7.5% acrylamide gels were cast containing 50 µM Phostag™, 50 µM $ZnCl_2$, and 350 mM Bis-Tris-HCl pH 6.8. 2.5 µg of protein sample was boiled for 5 minutes at 100°C in SDS sample buffer and then loaded onto gels which were run at 90 V, 4°C, for 2 hours in 100 mM Tris-HCl, 100 mM MOPS, 0.1% (w/v) SDS, and 5 mM sodium bisulphite. Gels were subsequently stained with InstantBlue Coomassie stain overnight at room temperature before being washed and imaged as in **Section 2.10.**

2.12 Ion exchange purification of DDR1 phospho-forms

The soluble DDR1 JM4K construct was incubated at 100 µM in the presence of 20 mM ATP, in kinase buffer II, for 90 minutes at 20°C. The protein sample was then dialysed overnight at 4°C into 50 mM Tris-HCl pH 8.5, 80 mM NaCl as described in **Section 2.4.5**. Protein was then loaded onto a 1-mL HiTrap Q XL column (GE Healthcare) on an Äkta pure chromatography system (GE Healthcare) at 1 mL/min, 4°C, collecting the flow-through (fraction FT). The other phospho-forms were eluted by increasing the NaCl concentration to ~160 mM (fraction E1) and then to ~230 mM (fraction E2). The fractions were dialysed into gel filtration buffer (**Section 2.4.5)** for further analysis as in. The fractions were analysed by native PAGE and Phostag™ (**Sections 2.10** and **2.11**). Fraction FT contained an intermediate phosphorylation form 1 of JM4K (JM4K-1P), and fraction E2 contained a fully phosphorylated form (JM4K-3P). Fraction E1 contained a mixture of intermediate JM4K phosphorylation forms 1 and 2 (JM4K-1P and JM4K-2P).

2.13 Tandem mass spectrometry

Liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis of DDR1 phospho-forms was conducted by the Cambridge Centre for Proteomics (CCP, University of Cambridge). For analysis, 1 µg of the JM4K-1P, or JM4K-3P, phospho-forms was separated by 10% SDS-PAGE for 2 hours, 100 V, at room temperature, and then stained with InstantBlue Coomassie for 15 minutes at room temperature. The protein was then cut from the gel using a scalpel, placed into 50 µL HPLC grade water and shipped to the CCP where the protein was cleaved with chymotrypsin, or trypsin, and analysed by LC-MS/MS. Data were analysed by the CCP.

2.14 Enzyme kinetic analysis

2.14.1 ADP-Glo™ kinase assay

For the analysis of substrate phosphorylation by purified soluble DDR1 kinase constructs, the ADP-Glo™ kinase assay was used (Promega). The substrate for phosphorylation was the Axltide IRS-1 derived peptide (KKSRGD**Y**MTMQIG – Genscript). Reactions were performed at room temperature in 96 well white assay plates (Greiner) in a 25 µL volume. Protein constructs at the indicated concentration (up to 3.2 µM) were incubated with Axltide peptide (up to 520 µM), and ATP (up to 800 µM) for the indicated time (up to 40 minutes). All reaction reagents were diluted in kinase buffer III (**Section 2.1.8)**. Following the kinase reaction, 25 µL ADP-Glo buffer (Promega) was added to plates for 40 minutes, prior to the addition of 50 µL kinase detection reagent (Promega), which was incubated for up to 60 minutes at room temperature (>500 µM ATP – 60 minutes, 100–500 µM ATP – 40 minutes, <100 µM ATP – 30 minutes). Luminescent signal was then recorded using a Tecan Spark plate reader with Spark Control V2.3 software. Pure ATP and ADP solutions (Promega) were mixed to produce an ADP standard curve (0–800 µM ADP) according to the manufacturer's instructions for each experiment. This enabled the conversion of relative light units, recorded by the Tecan Spark plate reader, into ADP concentrations produced by the kinase during the phosphorylation reaction through interpolation of the standard curve.

2.14.2 Michaelis-Menten kinetic analysis

After isolation of the different DDR1 phospho-forms (**Section 2.12**) they were analysed kinetically using the ADP-Glo™ assay. This was done either in the presence of 800 µM ATP and varying concentrations of Axltide peptide substrate (32–520µM), or in the presence of 520 µM Axltide peptide and varying concentrations of ATP (10–800 µM). The assay was performed as in **Section 2.14.1**. Luminescence was measured after 0, 10, 20, and 40 minutes from the initiation of the reaction. The rate of ADP production was linear up to 40 minutes in all reactions; initial rates in µM min^{-1} were determined by linear regression and fitted with the Michaelis-Menten equation using GraphPad Prism 8. The phosphorylated proteins (JM4K-WT-FP, JM4K-WT-2P, and K-P) were incubated at 0.1 µM concentration; partially phosphorylated (JM4K-1P) or unphosphorylated proteins (JM4K-WT-0P, K-0P, JM4K-Y569E/Y586E-0P) were incubated at 0.5 µM concentration. Specific activities (nmol min^{-1} mg^{-1}) were calculated and shown in graphs.

2.15 JM4 region peptide mimetic

A JM4 region peptide mimetic was custom-synthesised, encompassing residues Thr576-Val588 of DDR1, with Tyr586 mutated to Phe (**Table 7;** peptide 1). However, while this peptide dissolved in DMSO, it was insoluble in kinase buffer III (**Section 2.14**), making analysis challenging. Therefore, a longer peptide encompassing residues Glu571-Val588 was designed with two charged residues added at the N-terminus which are not naturally found in the JM4 sequence (**Table 7**; peptide 2). Peptide 2 dissolved to 12.5 mg/mL (6.1 mM) in 100 mM ammonium hydroxide. To test the effect of this peptide on DDR1 kinase activity it was analysed in the ADP-Glo™ assay. The soluble DDR1 kinase alone construct was incubated at 0.5 µM concentration with 25 µM ATP, 50 µM Axltide peptide, in the presence of 1.2 mM peptide 2, or 20 mM ammonium hydroxide as a control, in kinase buffer III (**Section 2.1.8)** for 30 minutes at room temperature, before following the ADP-Glo™ assay protocol described in **Section 2.14**.

Table 7. JM4 peptide mimetics.
The wild type DDR1 JM4 sequence is shown (Val566-Leu591) along with the two JM4 region peptides which were custom-synthesised. Peptide 1 encompassed Thr576-Val588 with the Y586F mutation (red). Peptide 2 encompassed Glu571-Val588 with the Y586F mutation (red). In addition, peptide 2 contained 2 charged amino acids at the N-terminus not naturally found within the DDR1 JM4 sequence (Lys and Glu) to increase peptide solubility.

JM4 Sequence	VPHYAEADIVTLQGVTGGNT**Y**AVPAL
Peptide 1	TLQGVTGGNT**F**AV
Peptide 2	**KE**EADIVTLQGVTGGNT**F**AV

2.16 Crystallography of phosphorylated DDR1

The soluble DDR1 JM4K protein construct was partially purified using a HisTrap column. The His_6-tag was cleaved overnight and the cleaved protein separated by a second round of HisTrap purification (**Section 2.4.5**). The protein was then incubated at 100 µM in kinase buffer II with 20 mM ATP for 6 hours at room temperature. Protein was then exchanged into a buffer containing 50 mM Tris-HCl pH 8.5, 60 mM NaCl using a 10 kDa molecular weight cut off Vivaspin filtration device (Sartorius). This was achieved through 6 rounds of concentration and 10X dilution of the protein sample in the outlined buffer. Phosphorylated forms of the protein were then separated by IEX chromatography as in **Section 2.12**; however, the wash buffer contained lower salt (50 mM Tris-HCl pH 8.5, 60 mM NaCl). The flow through (FT) contained nucleotides (ATP and ADP); A_{260}/A_{280} ratios recorded using a NanoDrop (**Section 2.3.1**) were > 2 indicating that the sample largely consisted of nucleotide. DDR1 JM4K phospho-forms were then eluted by increasing the NaCl concentration to ~120 mM (fraction E1) and then to ~180 mM (fraction E2). A final peak was observed at 500 mM NaCl (E3). Native PAGE and Phostag™ analysis (**Sections 2.10** and **2.11**) of samples revealed that fraction E1 contained the intermediate JM4K phosphorylation form 2 (2P), fraction E2 contained the fully phosphorylated (3P) JM4K, and E3 contained a mixture of JM4K-3P and an unidentified contaminating protein. Fraction E2 was subsequently concentrated, and gel filtered as in **Section 2.4.5**.

Following purification, the 3P JM4K phospho-form was concentrated to 9 mg/mL (**Section 2.4.5**) and incubated with 5 mM AMP-PNP (non-hydrolysable ATP

analogue) and 10 mM $MnCl_2$ for 30 minutes on ice. Initial crystallisation screening was then conducted as described in **Section 2.5**. Small needle-like crystals were found at 4°C in a 400 nL drop containing 200 nL protein/AMP-PNP/Mn^{2+} and 200 nL of solution consisting of 100 mM Na-HEPES pH 7.5, 22% (w/v) polyacrylic acid 5100 sodium salt, 20 mM $MgCl_2$. Crystals were too small for diffraction analysis; therefore, they were crushed using a CrystalProbe (Hampton Research) to generate a seed stock. A hanging drop was then set up in a 24-well hanging drop crystallisation plate (MiTeGen) which contained 0.5 µL of the protein-AMP-PNP/Mn^{2+} complex at 3, 6, or 9 mg/mL mixed with 0.5 µL of the crystal hit solution, at 17, 22, or 27% (w/v) precipitant concentration, above 1 mL of reservoir solution also containing the hit solution with the same precipitant concentration. Protein was left overnight to equilibrate, then seed stock was streaked across the drops using a seeding tool (Hampton Research). Larger crystals were observed in the 22% (w/v) precipitant condition. Crystals were cryoprotected and analysed as in **Section 2.5**. However, no diffraction data could be collected as the crystals did not diffract.

2.17 Cell culture

HEK293 and Cos-7 cells were cultured in DMEM/F12 supplemented with 2 mM L-glutamine, 100 units/ml penicillin, 100 µg/ml streptomycin and 10% (v/v) FBS, hereafter referred to as complete medium (**Section 2.1.10**). Cells were grown at 37°C, 5% CO_2. Cell culture plates (6-, 12-, and 24-well plates), and flasks (T25 cm^2 and T75 cm^2) were purchased from ThermoFisher.

2.17.1 Cell passaging

HEK293 cells and Cos-7 cells were split 1:5 or 1:10, respectively, once they had reached 90-95% confluency. Complete medium was aspirated, cells washed in PBS, and then detached from the culture flask using trypsin-EDTA (**Section 2.1.10**) solution for 1–2 minutes at room temperature (0.5 mL into a T25, 1.5 mL into a T75). Cells were then resuspended in complete medium (4.5 mL into a T25, 13.5 mL into a T75) and split into a new flask (0.5-1 mL into a T25, 1.5-3 mL into a T75).

2.17.2 Cell freezing

Cells were grown to confluency in a T75 flask before being washed in PBS, detached with trypsin-EDTA and resuspended in medium (**Section 2.17.1**). They were subsequently pelleted through centrifuging at 200 xg for 5 minutes at 20°C. Cell pellets were then resuspended in 3 mL of cell freezing medium (**Section 2.1.10**). Cells were then transferred into Nunc Cryotube Vials at 500 µL/tube, before being put into a polystyrene box and frozen overnight at -80°C prior to transfer into liquid nitrogen for long term storage.

2.17.3 Cell thawing

A 500 µL aliquot of cells frozen using the protocol in **Section 2.17.2** was defrosted in a 37°C water bath before being diluted in 10 mL of complete medium. Cells were then spun at 200 xg for 5 minutes, supernatant removed, and cells resuspended in 5 mL of complete medium and then transferred into a T25 flask.

2.17.4 Cell seeding

Cells were detached and resuspended in complete medium as described in **Section 2.17.1**. HEK293 and Cos-7 cells were then diluted 1:5 or 1:10, respectively, in complete medium. For collagen stimulation experiments, 1 mL, or 0.5 mL, of diluted cells was added to a 12-, or 24-well plate, respectively. For immunoprecipitation and flow cytometry experiments 2 mL of cells was added to a 6-well plate. The following day cells were typically at ~50% confluency.

2.18 Collagen stimulation of DDR1 expressing cells

This assay was performed as described previously (Leitinger, 2003). Cells were transfected with the indicated full-length DDR1 constructs in the pRK5 vector using the calcium phosphate precipitation method. This was achieved by diluting DNA in $CaCl_2$ (250 mM final $CaCl_2$ concentration - **Table 8**). This mixture was then added dropwise to a 2X solution of HBS (**Section 2.1.10**), whilst vortexing. The transfection reagent was left for 30 minutes at room temperature before being added to cells at 50% confluency (**Section 2.17.4**). The final transfection reagent volume added to cells was 100 µL, 200 µL, or 400 µL, for a 24-, 12-, or 6-well plate, respectively (**Table**

8). After 24 hours, cells were serum starved for 16 hours in serum free media (SFM; complete media without FBS). Cells were subsequently stimulated with 10 µg/mL collagen I in the presence of 10 mM Na-HEPES pH 7.2 in SFM for the indicated time (up to 90 minutes) at 37°C, or incubated with 1 mM acetic acid in the presence of 10 mM Na-HEPES pH 7.2 in SFM, as a control (collagen I stock was dissolved in 0.1 M acetic acid). After stimulation, cells were place on ice and the stimulation media transferred to Eppendorf tubes also on ice. Cells were washed in cold PBS, and the washes collected in the corresponding tube. Tubes were centrifuged at 14,000 xg for 2 minutes and the supernatant discarded. Cells on plates were lysed in Cell lysis buffer (**Section 2.1.10**) for 20 minutes on ice (50, 100, or 200 µL Cell lysis buffer per well of a 24-, 12-, or 6-well plate, respectively). Cells were then scraped off plates and placed into their corresponding tubes and incubated for a further 20 minutes. Cells were then centrifuged at 14,000 xg for 10 minutes at 4°C and the supernatants transferred to fresh Eppendorf tubes which were stored at -20°C for future analysis or were analysed immediately by Western blotting. Aliquots of cell lysates were boiled for 5 minutes at 100°C in SDS sample buffer before analysis by 7.5% SDS-PAGE, followed by Western blotting (**Section 2.9**).

Table 8.Calcium phosphate precipitation transfection reagents.
The reaction mixtures per well used for calcium phosphate precipitation of HEK293 cells with pRK-DDR1b constructs are shown for cell culture plates. For the pRK-DDR1b-Y569E/Y586E, and pRK-DDR1b-ΔJM4 constructs, 3X the amount of DNA was added per well.

Reagent	24-well	12-well	6-well
DNA (µg)	1.25	2.5	5
Diluted DNA (µL)	43.75	87.5	175
2M $CaCl_2$ (µL)	6.25	12.5	25
2X HBS (µL)	50	100	200

2.19 EndoH digestion

In order to monitor the extent of oligosaccharide processing of the full-length wild type and mutant DDR1 constructs, HEK293 cells were transfected with pRK5-DDR1 constructs which were subsequently lysed (**Section 2.18**)**.** 18 µL of cell lysate was then incubated with glycoprotein denaturation buffer (NEB) for 5 minutes at 100°C. Half of this denatured sample was then mixed with G5 buffer (diluted to 1X with sample) and 50 U of Endo H (NEB) for 3 hours at 37°C, whilst the other half was incubated with G5 buffer alone for 3 hours at 37°C. Samples were then boiled for 5 minutes at 100°C in SDS sample buffer and analysed by 7.5% SDS-PAGE and Western blotting with anti-DDR1 antibodies (**Section 2.9**).

2.20 Flow cytometry

For the confirmation of cell surface expression of the full-length wild type and mutant DDR1 constructs, HEK293 cells were transfected with pRK5-DDR1 constructs in 6-well plates as described in **Section 2.18**. Cells were then incubated with 1 mL of cell dissociation buffer (**Section 2.1.10**) per well for ~2 minutes at room temperature and detached by gentle pipetting. Cells were subsequently resuspended in 5 mL of complete medium and pelleted by centrifugation at 200 xg for 5 minutes at 4°C. The supernatant was discarded, and the cell pellet resuspended in PBS containing 1% (w/v) BSA. The cells from one well of a 6-well plate were then split into two Eppendorf tubes and centrifuged at 3000 xg for 4 minutes at 4°C and the supernatant discarded. One tube of cells was then incubated with 10 µg/mL mouse anti-DDR1 7A9 in PBS containing 1% (w/v) BSA for 30 minutes on ice, whilst the other tube was incubated in PBS containing 1% (w/v) BSA alone as a control (secondary only staining condition). Cells were then washed three times in PBS containing 1% (w/v) BSA. Cells were centrifuged at 3000 xg for 4 minutes at 4°C with supernatant discarded for each wash. Following wash steps, cells were incubated with goat anti-mouse FITC (1:500) for 30 minutes on ice. They were then washed once as above, and then once in PBS, before being incubated with 1:1000 Zombie near infrared viability dye (BioLegend) in PBS for 20 minutes on ice. Cells were washed once with PBS, then with PBS containing 1% (w/v) BSA, before being fixed in PBS containing 2% (w/v) formaldehyde on ice. Data were then collected on a BS Accuri C6 Plus flow

cytometer (**Table 9**) and data analysed using BD Accuri C6 Plus software v1.0.23.1 and FlowJo version 10.6.1. Zombie negative live HEK293 cells were selected, the mean AF488 fluorescence intensities were then calculated.

Table 9. Flow cytometry settings.

2-blue 2-red flow cytometry settings	
FL1	530/30
FL2	585/40
FL3	780/60
FL4	675/25

2.21 *In vitro* autophosphorylation of full-length DDR1

HEK293 cells were transfected in 6-well plates with the indicated pRK5-DDR1 constructs (transfection as described in **Section 2.18**). After 24 hours, cells were serum starved for 16 hours in SFM. They were then lysed, and the lysate cleared through centrifugation at 14,000 xg for 10 minutes at 4°C as in **Section 2.18**. The lysates were diluted with 300 µL Cell lysis buffer (**Section 2.1.10**) to a total of 500 µL and incubated with 1 µg of anti-DDR1 7A9, overnight at 4°C with rotation. The following morning, Protein A Sepharose fast flow beads (GE-Healthcare) were washed 3 times in PBS and then resuspended to form a 50:50 solution of PBS:beads. All bead washes were performed by spinning tubes at 3000 xg for 1 minute with supernatant discarded each time. 40 µL of the 50% bead slurry was added to each tube and lysates mixed by rotation with the beads for 90 minutes at 4°C. Beads were then washed 3 times in Cell lysis buffer, before receiving 2 further washes in IP kinase buffer (**Section 2.1.8**). The beads were then incubated in 40 µL IP kinase buffer containing 1 mM ATP for the indicated time at 30°C with beads resuspended every 5 minutes by flicking the tubes. Samples were boiled for 5 minutes at 100°C in SDS sample buffer to stop reactions. Sample aliquots were then separated by 7.5% SDS-PAGE and Western blot analysis was conducted (**Section 2.9**).

2.22 Immunofluorescence imaging

Cos-7 cells were seeded in 24-well plates (**Section 2.17.4**) containing 13 mm coverslips. The following day, cells at ~50% confluency were transfected using Lipofectamine 3000 (ThermoFisher) according to the manufacturer's instructions. For each well of a 24-well plate, 500 ng of pRK5-DDR1 construct was mixed with 1 µL P3000 reagent (ThermoFisher) to a final volume of 50 µL in Opti-MEM medium (ThermoFisher). This DNA/P3000 mixture was then combined with 1.5 µL Lipofectamine 3000 reagent (ThermoFisher) diluted to 50 µL in OptiMEM medium. The transfection mixture was then incubated at room temperature for 15 minutes before being added dropwise to cells. After 24 hours, cells were serum starved for 16 hours in SFM. Cells were subsequently stimulated with 10 µg/mL collagen I in the presence of 10 mM Na-HEPES pH 7.2 in SFM for the indicated time (up to 60 minutes) at 37°C, or incubated with 1 mM acetic acid in the presence of 10 mM Na-HEPES pH 7.2 in SFM, as a control. Cells were washed in cold PBS before being incubated with 30 µL of anti-DDR1 7A9 (10 µg/mL) in 5% BSA/PBS for 1 hour on ice. Cells were then washed 3 times in PBS before being fixed in 4% paraformaldehyde in PBS (Sigma-Aldrich) for 15 minutes at room temperature. Cells were then washed 3 times in PBS, permeabilised with 0.1% Triton-X100 in PBS for 10 minutes, washed a further 3 times in PBS, and then blocked in 5% BSA/PBS for 1 hour at room temperature. Cells were then incubated with 30 µL of the DDR1 specific anti-phosphotyrosine antibody, anti-pY513 (1:500), in 5% BSA/PBS for 1 hour at room temperature. Coverslips were then washed 3 times in PBS before being incubated with AlexaFluor 488 anti-mouse IgG1 (1:500), AlexaFluor 555 anti-rabbit IgG (1:500), and DAPI (0.2 µg/mL) in 5% BSA/PBS for 1 hour at room temperature. Cells were washed 3 times in PBS and then mounted overnight using ProLong Gold Antifade Mountant (ThermoFisher).

Images were acquired using a Zeiss Axio Observer widefield microscope equipped with a 63x 1.4 numerical aperture oil objective lens and Zen pro software (Ziess) (Facility for Imaging by Light Microscopy, Imperial College London). Images were processed using ImageJ software (National Institute of Health), or Icy (Institut Pasteur).

2.23 Statistical analysis

Experiments were performed a minimum of 3 times and data presented as the mean with standard error of the mean, unless otherwise stated. Blots and other images show representative examples. GraphPad Prism 8 was used for the determination of statistical significance between experimental groups. For peptide inhibitor studies, a one-tailed unpaired t-test was performed. For experiments with more than two groups, a one-way or two-way analysis of variance (ANOVA) test, followed by a Tukey post-test was conducted. Statistical significance was set as a p-value of <0.05.

Chapter 3. The DDR1 JM4 region is autoinhibitory

3.1 Introduction

RTKs govern many vital cellular processes including, proliferation, migration, and differentiation (Lemmon & Schlessinger, 2010). To prevent unwanted signalling, kinases are tightly regulated through autoinhibitory interactions which are relieved only upon ligand binding (Hubbard & Miller, 2007). Receptor dimerisation, structural rearrangement of pre-formed dimers, or clustering of receptors into higher-order oligomers, typically occurs following ligand binding (Hubbard & Miller, 2007; Lemmon & Schlessinger, 2010). These events increase the proximity of the receptor kinase domains, facilitating *trans*-phosphorylation of Tyr residues, and consequently a release of autoinhibition (Lemmon & Schlessinger, 2010).

The main autoinhibitory motifs found in RTKs are the A-loop, C-terminal tail, and the JM region. A-loop autoinhibition is common in RTKs and is exemplified by the insulin receptor in which it was first described (Hubbard et al, 1994). A Tyr residue within the A-loop (Tyr1162) forms a pseudosubstrate interaction with Asp1132 of the HRD (His1130-Arg1131-Asp1132) motif. This interaction occludes substrate access to the active site and maintains the A-loop in the inactive **D**FG-OUT (Asp1150, Phe1151, Gly1152) conformation (**Figure 3**). *Trans*-phosphorylation of the A-loop results in release of autoinhibition, and reorientation of the A-loop into a conformation which permits substrate access and ATP coordination by Asp1150 (**D**FG-IN conformation) (Hubbard, 1997; Hubbard et al, 1994). This transition is accompanied by a reorientation of the αC-helix and Gly-rich loop towards the catalytic cleft which enables ATP binding. Autoinhibition by the C-terminal tail is found in the Tie, Ron, EGFR, and PDGFR families (Chiara et al, 2004; Gajiwala, 2013; Niu et al, 2002; Yokoyama et al, 2005). In all cases, the tail contains a Tyr autophosphorylation site and forms inhibitory interactions with the kinase domain that occlude substrate access. Finally, the JM region of some RTKs can regulate kinase activity (Hubbard, 2004). Autoinhibition by the JM region has been reported for the PDGFR, MuSK, and Eph family receptors (Binns et al, 2000; Mol et al, 2004; Till et al, 2002). This occurs through a variety of mechanisms; however, it typically involves interactions between the JM region and kinase domain which stabilise the inactive conformation, and

invariably includes a Tyr autophosphorylation site (Hubbard, 2004). However, the JM region can also have positive roles in kinase activation. In the Ret and EGF receptors, for example, the JM region functions to stabilise the active conformation of the kinase, although through different mechanisms (Jura et al, 2009; Plaza-Menacho et al, 2016). On the other hand, for the insulin receptor, the JM region is initially autoinhibitory, but following phosphorylation, can promote the formation of a catalytically competent kinase dimer (Cabail et al, 2015).

As discussed in **Chapter 1, Section 1.2**, DDR1 is an unusual RTK family member; it exists at the cell surface as a stable dimer and is activated with exceptionally slow kinetics by collagen (Shrivastava et al, 1997; Vogel et al, 1997). Upon collagen binding, DDR1 redistributes into dense clusters at the cell surface, within which phosphorylation between dimers is triggered (Corcoran et al, 2019; Juskaite et al, 2017; Yeung et al, 2019). The mechanism through which DDR1 clustering leads to kinase activation, and the regulatory motifs which maintain the inactive state, remain poorly understood.

Analysis of inhibitor bound DDR1 kinase structures has revealed that DDR1 does not contain a C-terminal tail (Canning et al, 2014). Instead, the C-terminal region is incorporated into the final α-helix (αI-helix) (Canning et al, 2014; Leitinger, 2014). It is therefore unlikely that DDR1 is regulated through C-terminal-mediated interactions. The DDR1 structure did reveal that the A-loop Tyr796 forms a pseudosubstrate interaction with the catalytic loop Asp766 (Canning et al, 2014). Autophosphorylation of Tyr796, following collagen binding, is likely to relieve this inhibition (Canning et al, 2014; Juskaite et al, 2017; Yang et al, 2005). The function of the exceptionally long DDR1 JM region is however not well characterised. There are 7 Tyr residues within the JM region of DDR1b, some of which have been identified as autophosphorylation sites that recruit downstream signalling molecules (**Figure 8**). For example, phospho-Tyr513 has been identified as a Shc1 binding site (Shrivastava et al, 1997; Vogel et al, 1997). Unlike the other highly conserved globular domains of the receptor, the JM region shows generally poor sequence conservation (**Figure 6**). However, the kinase proximal segment of the JM region (DDR1 residues 566-591), termed the JM4 region, is extremely well conserved, indicating potential functionality.

3.1.1 Hypothesis and aims

It is hypothesised that the highly conserved DDR1 JM region, termed the JM4 region, plays a regulatory role in kinase activation.

The aims were: 1) to establish whether the JM4 region plays a role in receptor activation; 2) to express and purify soluble DDR1 kinase constructs; 3) to structurally characterise the role of the JM4 region.

3.1.2 Findings

Through biochemical and biophysical analysis, the DDR1 JM4 region is shown to be an autoinhibitory motif. A crystal structure reveals that the JM4 region reinforces A-loop autoinhibition and stabilises the inactive conformation of the kinase. Structural distinctions between the JM4 autoinhibited DDR1 structure and inhibitor bound DDR1 structures are found. Furthermore, parallels between the autoinhibited DDR1 structure and that of other RTK family members are drawn.

3.2 Results

3.2.1 Identification of the DDR1 JM region as a regulator of signalling

The DDRs contain a long intracellular JM region (160 residues in DDR1b) of unknown function. The DDR JM region can be subdivided into 4 segments (JM1-4) based upon sequence conservation between DDR1 and DDR2 (excluding the DDR1b/c-specific insert – **Figure 6**). The most highly conserved segment, termed the JM4 region, is also that which lies closest to the kinase domain (**Figure 10a,b**). Sequence alignment of the human DDR1 JM4 region across multiple species revealed that it contains several invariant residues and limited sequence variation between species (**Figure 10a**). It was therefore hypothesised that this highly conserved kinase proximal JM segment may play a functional role in DDR signalling. To investigate this, the two invariant Tyr residues within the DDR1 JM4 region, Tyr569 and Tyr586, were mutated to Phe, either as single or double mutations (mutagenesis performed by Athena Cheung). DDR1-wild type (WT), or the mutant constructs; DDR1-Y569F, DDR1-Y586F, and DDR1-Y569F/Y586F, were expressed in HEK293 cells. Cells were subsequently stimulated with collagen I, a ligand known to result in robust DDR1 autophosphorylation (Juskaite et al, 2017). Collagen-stimulated activation of DDR1-WT was detected by Western blotting using the A-loop specific anti-phosphotyrosine antibody, anti-pY796. Collagen I incubation resulted in a marked increase in anti-pY796 phospho-signal compared with the unstimulated control (**Figure 10c;** WT - vs WT +). A-loop phosphorylation is an essential step in the activation of most RTKs, and has been demonstrated for DDR1 on Tyr796 (Hubbard & Miller, 2007; Juskaite et al, 2017; Lemmon & Schlessinger, 2010). Collagen stimulation of DDR1-Y569F resulted in a small reduction in A-loop phosphorylation compared with DDR1-WT. A much greater loss of collagen-induced phosphorylation was observed for the DDR1-Y586F construct, and an almost complete loss of A-loop phosphorylation was observed for the DDR1-Y569F/Y586F construct. This demonstrated that the JM4 region Tyr residues, and in particular Y586, are required for efficient collagen induced DDR1 autophosphorylation.

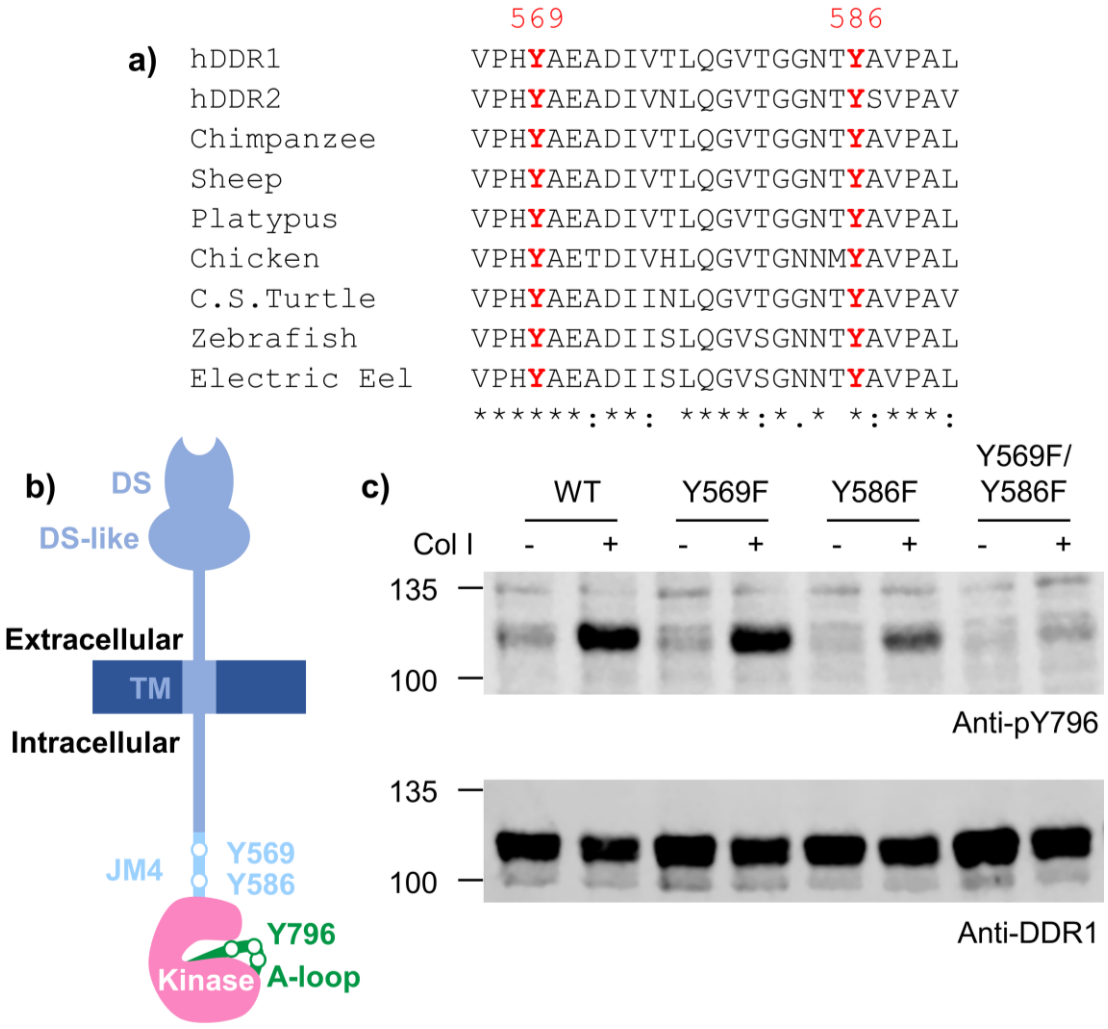

Figure 10. DDR1 JM4 region is a regulator of kinase activity.
a) Sequence alignment (Clustal Omega) of the human DDR1 (hDDR1. UniProtKB Q08345) JM4 region (residues 566-591 of DDR1b) with human DDR2 (hDDR2. Q16832), and the DDR1-JM4 sequence from other species. Chimpanzee; *Pan troglodytes* (Q7YR43). Sheep; *Ovis aries* (W5PND). Platypus; *Ornithorhynchus anatinus* (A0A6I8N581). Chicken; *Gallus gallus domesticus* (Q5F3X2). Chinese softshell (C.S) turtle; *Pelodiscus sinensis* (K7FGL7). Zebrafish; *Danio rerio* (E7F3X9). Electric eel; *Electrophorus electricus* (A0A4W4GPZ5). (*) - invariant amino acid. (:) - conservative amino acid substitution. (.) – less well conserved amino acid. **b)** Molecular architecture of DDR1. DS – discoidin. TM – transmembrane. Selected JM and kinase domain Tyr residues are depicted by open circles and labelled. The A-loop (green) of the kinase domain (pink) contains three Tyr residues (Tyr792, Tyr796, and Tyr797), while the JM4 region (light blue) has two (Tyr569 and Tyr586). **c)** Full-length DDR1-WT, or the DDR1 mutants DDR1-Y569F, -Y586F, -Y569F/Y586F, were transiently expressed in HEK293 cells. Cells were then stimulated with collagen I (+) or left unstimulated (-) for 90 minutes at 37°C. Cells were subsequently lysed, and aliquots of cell lysates analysed by SDS-PAGE and Western blotting with the A-loop-specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left.

3.2.2 Purification of soluble DDR1 kinase

To further investigate the functional role of the DDR1 intracellular JM4 region, several soluble DDR1 kinase constructs were designed and expressed in Sf9 cells using a baculovirus expression system. Initial screening of DDR1 expression in E. *coli*, mammalian (HEK293 cells), and Sf9 cells, was conducted by the Oxford Protein Production Facility (OPPF). Successful protein production was only achieved using baculovirus infection of Sf9 cells. As a result, all protein constructs used in this report were expressed using this system. The 4 protein constructs used throughout were: 12aa-JM4K (DDR1 kinase domain with the JM4 region and 12 amino acids of the JM3 region – residues 554-913), JM4K (DDR1 kinase domain with the JM4 region – residues 566-913), Y569F/Y586F-JM4K (DDR1 kinase domain with the JM4 region and both JM4 Tyr residues mutated to Phe – residues 566-913), and kinase alone (DDR1 kinase domain alone – residues 601-913) (**Figure 11**). The JM4K and kinase alone constructs were produced to determine the effect of the JM4 region on DDR1 function, whilst the 12aa-JM4K construct was produced to determine whether the 12 additional amino acids present in this construct, relative to JM4K, had any effect on stability, amenability to structural studies, or protein expression. Finally, the Y569F/Y586F-JM4K construct was produced as these mutations rendered the receptor inactive in cells (**Figure 10**). All constructs expressed to similar levels (~2-3 mg/L of culture) and the following is a description of the purification process used, with 12aa-JM4K as an example. Initially, the baculovirus was produced by co-transfection of a pOPINF vector with restriction enzyme cleaved viral bacmid. The pOPINF vector contained the DDR1 protein sequence (DDR1 residues 554-913 for the 12aa-JM4K construct), a 5' His_6 tag, and the baculovirus ORF 1629 gene (**Figure 9**). Recombination of the replication deficient cleaved bacmid with the pOPINF vector resulted in restoration of the ORF 1629 gene in the bacmid and the production of a replication proficient virus (Zhao et al, 2003). Following recovery of the recombinant bacmid from cell culture medium (P_0 virus), it was amplified to the P_1 then P_2 stage through sequential rounds of Sf9 cell infection (7-day infections). The ability of the virus to produce DDR1 protein was then assessed using the P_2 viral passage. Sf9 cells were infected with 1 or 10 µL of viral suspension per mL of cells. After 3 days of infection, cells were harvested and lysed. His_6-tagged protein was then purified using Ni-NTA beads. Following pull down, the beads were washed 3 times in a buffer

containing low concentrations (10 mM) of imidazole to remove contaminating proteins. Samples were then boiled in sample buffer and analysed by SDS-PAGE and InstantBlue Coomassie staining (**Figure 12**). Infection of Sf9 cells with the virus led to the production of a protein of the expected size (~40 kDa). Furthermore, the 1 µL virus/mL of viral suspension condition produced the highest amount of protein; therefore, this condition was used for future large-scale protein expressions. The same conditions were also applied to the production of all other soluble DDR1 protein constructs (**Table 6**).

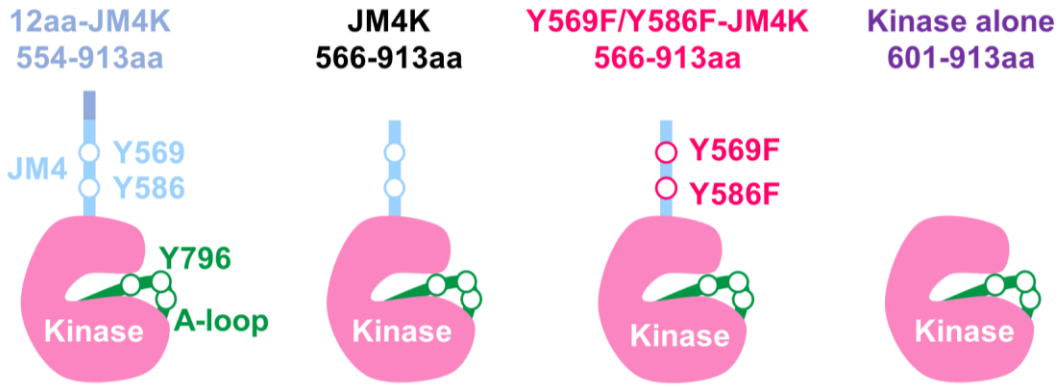

Figure 11. Soluble DDR1 kinase constructs.
A cartoon representation of the purified soluble DDR1 kinase constructs used throughout this thesis; 12aa-JM4K which contained the kinase domain, JM4 region, and 12 amino acids of the JM3 region (DDR1 residues 554-913), JM4K which contained the kinase domain and JM4 region (DDR1 residues 566-913), Y569F/Y586F-JM4K which contained the kinase domain and JM4 region and both JM4 Tyr mutated to Phe (Tyr569 and Tyr586, DDR1 residues 566-913), and kinase alone which contained only the kinase domain (DDR1 residues 566-913). Selected JM region and kinase domain Tyr residues are depicted by open circles and labelled. The A-loop (green) of the kinase domain (pink) contains three Tyr residues (Tyr792, Tyr796, and Tyr797), while the JM4 region (blue) has two (Y569 and Y586).

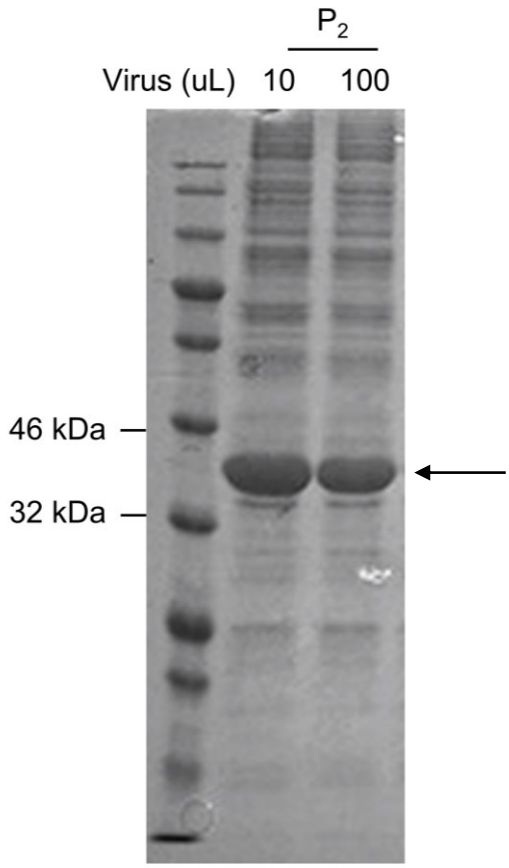

Figure 12. Expression testing of soluble DDR1-12aa-JM4K.
A pOPINF vector containing the DDR1 12aa-JM4K cDNA (encoding residues 554-913) was co-transfected with linearised bacmid into Sf9 cells. The recombinant virus (P_0) which was produced was then amplified to the P_1 then P_2 stage. Expression testing was performed for the P_2 viral passage. 10, or 100 µL of viral suspension was used to infect 10 mL of Sf9 cells at $1x10^6$ cells/mL which were incubated at 27°C, 120 rpm for 3 days. Cells were then pelleted, lysed, and the His_6-tagged protein purified using Ni-NTA agarose beads. Protein was then boiled in sample buffer and analysed by SDS-PAGE and InstantBlue Coomassie staining. The position of molecular weight markers (in kDa) are shown on the left. The position of the DDR1 protein is indicated by a black arrow.

A 4-step purification process was used for the large-scale production of all soluble DDR1 kinase constructs. This began with the infection of 0.5–3 L of SF9 cells at $1x10^6$ cells/mL with 1 µL virus/mL of cells. Following 3 days of infection, cells were harvested, lysed, centrifuged, and filtered to remove cell debris. The first step in the purification was then performed using a HisTrap HP column (**Figure 13a**). Cleared protein lysate was loaded onto the column. The column was subsequently washed with 15 mM imidazole to remove low affinity contaminants. Protein was then eluted using a 500 mM imidazole elution buffer. Protein fractions which eluted from the column were then analysed by SDS-PAGE and InstantBlue Coomassie staining (**Figure 13b**). Those fractions which contained protein of the expected molecular weight were then pooled (**Figure 13**; fractions 2-9: 12aa-JM4K Mw ~40 kDa). The protein was then cleaved overnight with C3 protease to remove the His_6-tag and dialysed into a buffer containing 150 mM NaCl, 50 mM Tris-pH 8.5, and 0.5 mM TCEP. The following morning, the solution was purified a second time using the HisTrap HP column and the flow-through (FT) containing DDR1 protein without His_6-tag collected, while contaminating protein was eluted (E) from the column in a buffer containing 500 mM imidazole (**Figure 14a,c**). The HisTrap FT was further purified by ion exchange chromatography (IEX) using a Q-XL column. Again, the DDR1 protein was collected in the flow through (FT), and further contaminating protein removed and eluted (E) from the column using a buffer containing 1 M NaCl (**Figure 14b,c**). Finally, the protein sample was purified by gel filtration. A small amount of contaminant was removed by this process (**Figure 15**, fractions 16 and 17) with pure 12aa-JM4K eluting in a sharp peak (**Figure 15**, fractions 18-20). The gel filtered fractions containing purified 12aa-JM4K (fractions 18-20) were pooled and concentrated to ~10 mg/mL. Protein samples were then analysed immediately or were snap-frozen in liquid nitrogen for future characterisation.

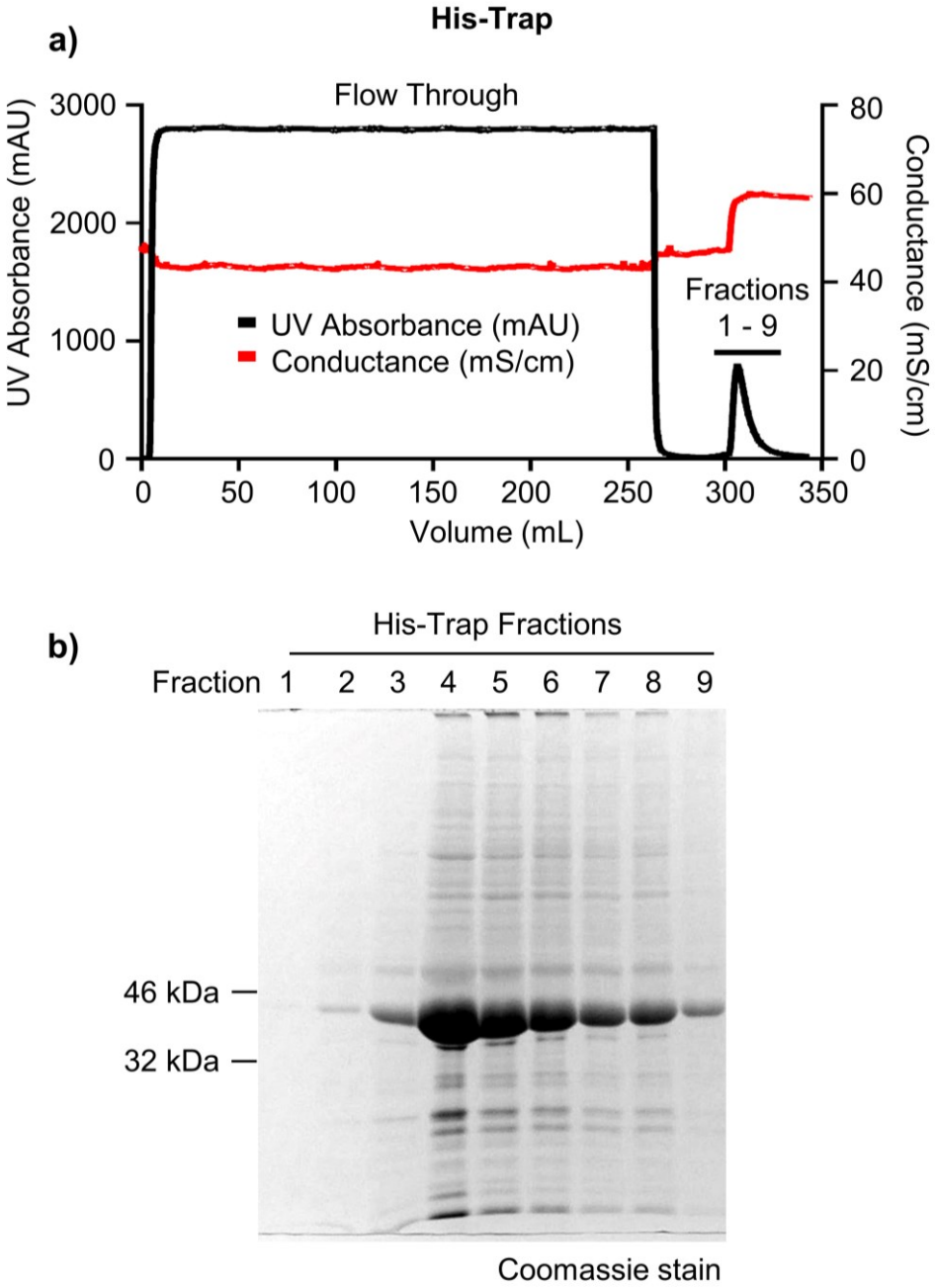

Figure 13. His-Trap purification of soluble DDR1-12aa-JM4K.
P_2 passage baculovirus was used to infect 1L of Sf9 cells at 1x10^6 cells/mL which were incubated at 27°C, 120 rpm for 3 days. Cells were then pelleted before being lysed and loaded onto a His-trap HP column. **a)** The His-trap chromatogram is shown with conductance and UV absorbance (A_{280}). The column was cleaned with 15 mM imidazole following protein loading. Following column washing, protein was eluted with 500 mM imidazole. **b)** Aliquots of fractions collected during His-trap purification (fractions 1-9 in **a)**) were boiled in sample buffer and analysed by SDS-PAGE and InstantBlue Coomassie staining. The position of molecular weight markers (in kDa) are shown on the left.

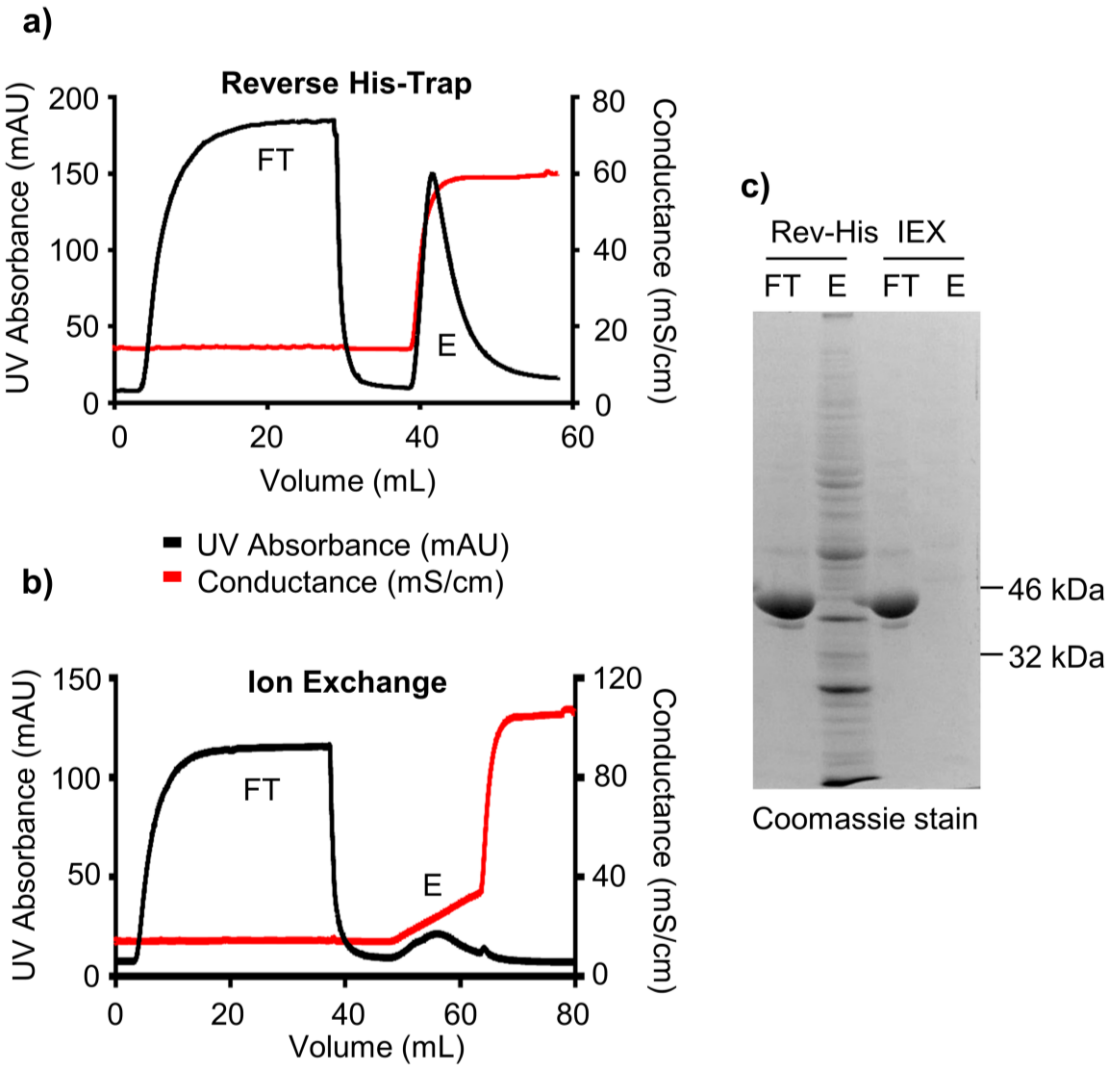

Figure 14. Reverse purification of 12aa-JM4K.
a) Chromatogram from reverse His-Trap purification with conductance and UV absorbance (A_{280}) shown. Protein which had been purified on a His-trap column was cleaved overnight with C3 protease and dialysed into 150 mM NaCl, 50 mM Tris pH 8.0, and 0.5 mM TCEP. Cleaved protein was then passed over a His-trap HP column and the flow through (FT) collected. Protein which had bound the column was then eluted (E) with 500 mM imidazole. **b)** Chromatogram from ion exchange purification with conductance and UV absorbance (A_{280}) shown. The flow through from **a)** was passed over an ion exchange Q-XL column and the flow through collected. Protein which had bound the column was then eluted (E) with 1 M NaCl, 50 mM Tris pH 8.0. **c)** Aliquots of the flow through (FT) and elution (E) peaks from **a)** and **b)** were boiled and analysed by SDS-PAGE and InstantBlue Coomassie staining. The position of molecular weight markers (in kDa) are shown on the right.

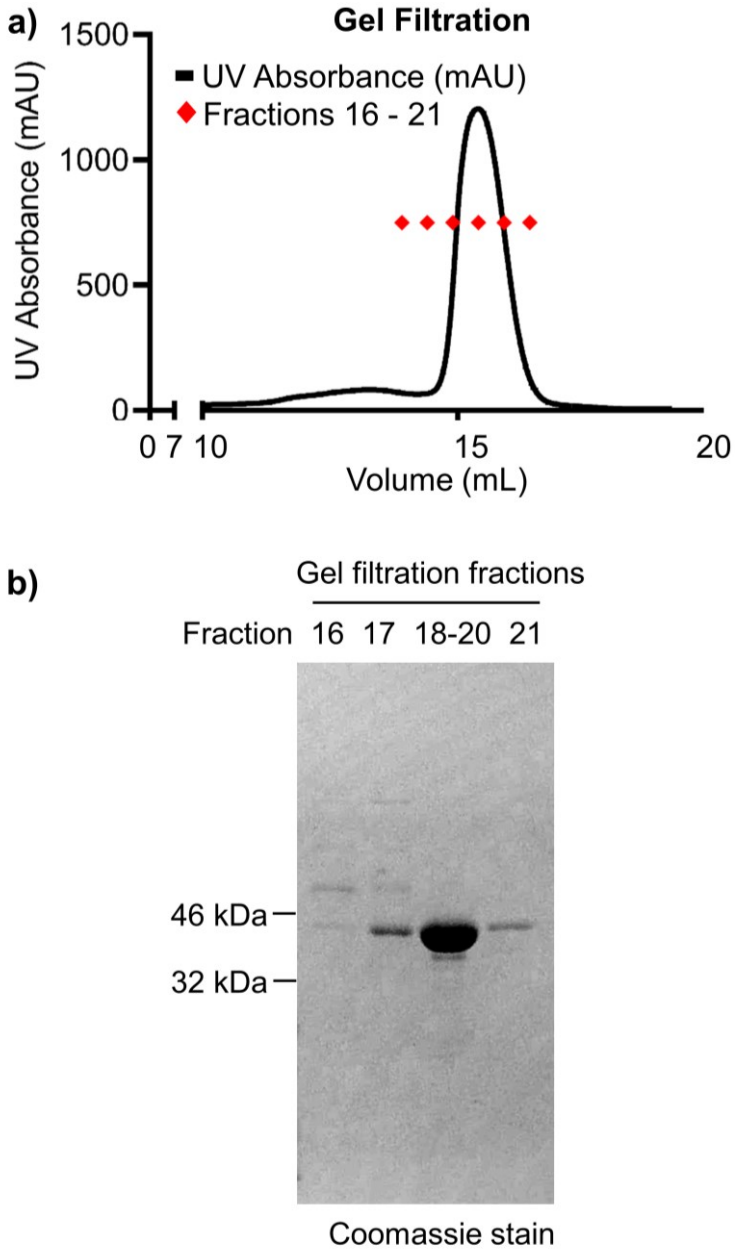

Figure 15. Gel filtration of DDR1-12aa-JM4K-WT.
Protein which had been purified by His-trap and ion exchange was further purified by gel filtration (size exclusion chromatography). **a)** Chromatogram from gel filtration with UV absorbance (A_{280}). The position of fractions which were collected for analysis (16-21) are shown. **b)** Aliquots of the indicated fractions from the gel filtration were boiled in sample buffer and analysed by SDS-PAGE and InstantBlue Coomassie staining. Fractions 18-20 were pooled prior to analysis. The position of molecular weight markers (in kDa) are shown on the left.

3.2.3 JM4 region is not resolved in a DDR1-dasatinib crystal structure

In order to understand the structural basis for the regulation of DDR1 kinase activity by the JM4 region, crystallographic analysis of soluble DDR1 kinase constructs containing the JM4 region was conducted. Initial crystallisation screening of the free DDR1 enzymes, JM4K and 12aa-JM4K, did not result in the successful production of protein crystals. However, crystals were obtained of the 12aa-JM4K construct in complex with the type-I kinase inhibitor, dasatinib, at 8 mg/mL protein concentration (**Figure 16a**).

A crystal structure was solved at a resolution of 2.26 Å (**Table 10**). A DDR1 kinase alone (residues 601-913)/dasatinib co-crystal structure had previously been solved (PDB:5BVW (Murray et al, 2015)) and was used for molecular replacement with the 12aa-JM4K/dasatinib data produced here. As an internal control, dasatinib was omitted from the search model. In support of the presented model, clear electron density was observed which corresponded to the omitted dasatinib molecule (**Figure 16b**). As reported previously (Canning et al, 2014), the DDR1 kinase domain adopts the classic bilobal architecture. This consists of a small N-terminal N-lobe which is composed largely of β-sheets with a single αC-helix, and a larger C-terminal lobe which is largely α-helical (**Figure 16c**). Comparison of the 12aa-JM4K/dasatinib structure solved here, with the previously determined DDR1 kinase alone/dasatinib structure (Murray et al, 2015), revealed that there were few structural differences (**Figure 16c**). Furthermore, no electron density was observed for the JM4 region (**Figure 16d**). Interestingly, whilst dasatinib is a type-I ATP-competitive inhibitor, which typically binds kinases in the “active” **D**FG-IN conformation (DDR1 residues Asp784, Phe785, Gly786), the DDR1 structure shown here, as well as the previously solved PDB:5BVW structure, was solved in the “inactive” **D**FG-OUT conformation (**Figure 16e**) providing further evidence of the unique nature of the DDR kinase.

Table 10. Crystallographic statistics for 12aa-JM4K-dasatinib co-crystal.

Data Collection	
Protein	12aa-JM4K (Gly554-Val913) + dasatinib
Beamline	i24
Wavelength (Å)	0.9686
Space Group	$P2_12_12_1$
Cell dimensions	
a (Å)	62.15
b (Å)	76.07
c (Å)	78.23
Resolution (Å)	48.66-2.26 (2.32-2.26)
Unique reflections	17915 (1325)
$CC_{1/2}$	0.998 (0.595)
R_{pim}	0.054 (0.719)
Mean I/σ (I)	11.4 (1.6)
Completeness	99.7 (99.5)
Multiplicity	6.3 (6.5)
Refinement	
R_{work}/R_{free}	0.252/0.304
RMSD	
Bonds (Å)	0.016
Angles (°)	1.57
Ramachandran plot	
Favoured (%)	94.4
Outliers (%)	0.75

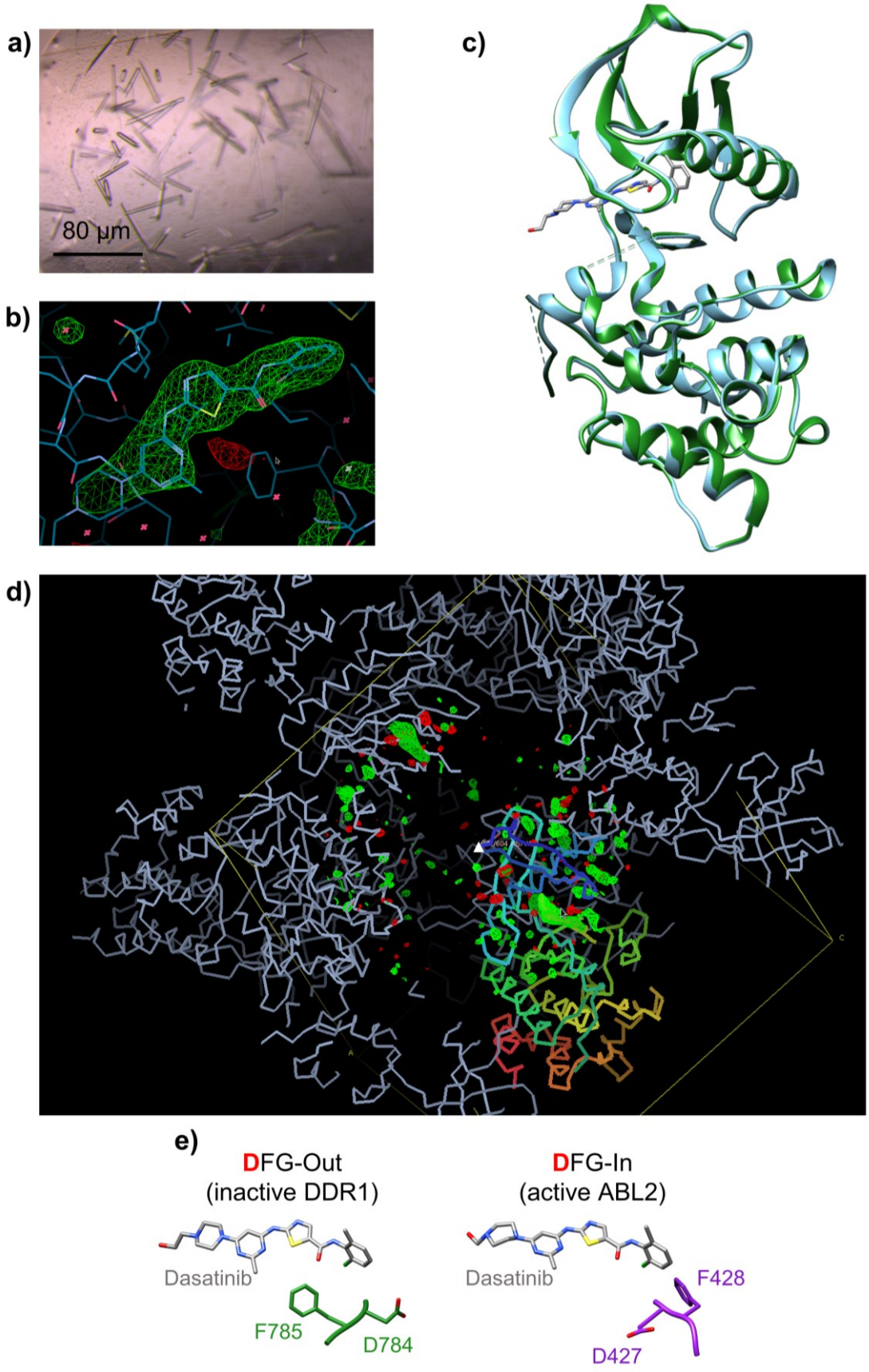

Figure 16. DDR1-dasatinib crystal structure.

a) Crystals of 12aa-JM4K/dasatinib used for X-ray diffraction data collection. Scale-bar, 80 µm. **b)** Dasatinib was omitted from the search model to provide an internal control when performing molecular replacement with PDB:5BVW (Murray et al, 2015). A weighted (Fobs-Fcalc) map contoured at 3 sigma is shown. **c)** Superimposition of 12aa-JM4K/dasatinib structure solved here (green) with a previously solved structure of DDR1-kinase alone with dasatinib (PDB:5BVW) (blue). Dasatinib is shown in grey. **d)** Electron density difference map centred on the N-terminal amino acid in the search model (D604-white triangle). The outline of the crystal unit cell is shown in yellow with symmetry-related copies shown in light blue. **e)** Left - orientation of the DFG (Asp784-Phe785-Gly-786) motif in the 12aa-JM4K/dasatinib crystal structure. Dasatinib is shown in grey. Right – Orientation of the DFG (Asp427–Phe428–Gly429) motif in the active state of the Abl2 kinase in complex with dasatinib (grey) (PDB: 4XLI; (Ha et al, 2015).

As mentioned above, there was no detectable density for the JM region within the 12aa-JM4K/dasatinib structure. Therefore, to exclude whether this result was due to cleavage of the JM region during crystal formation, crystals were removed from wells and washed in crystallisation reservoir solution. The samples were then boiled in sample buffer and analysed by SDS-PAGE and silver staining (**Figure 17**). As a control, the soluble DDR1 12aa-JM4K, and kinase alone (DDR1 residues 601-913) constructs were run on the gel. 1 or 12 ng of 12aa-JM4K and 12 ng of kinase alone was loaded. The crystallised 12aa-JM4K and soluble 12aa-JM4K were detected in the same vertical position on the gel (~40 kDa) indicating that the JM region was still present in these crystals. Therefore, the absence of JM region electron density in the resolved structure (**Figure 16**) is likely a result of disorder.

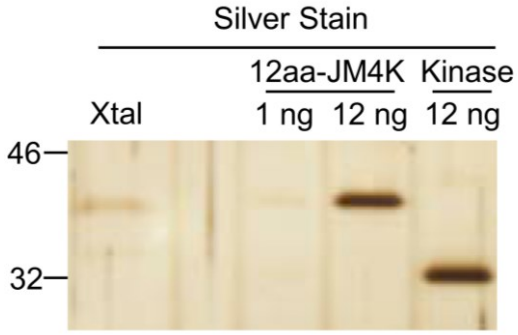

Figure 17. JM region is present in the DDR1 12aa-JM4K/dasatinib crystals.
Crystals (Xtal) were fished from the well, washed twice in crystallisation reservoir solution to remove precipitant, and then boiled in sample buffer and analysed by SDS-PAGE and silver staining. Purified 12aa-JM4K and kinase alone (DDR1 residues 601-913) were added to the gel as controls. The amount of protein loaded is indicated (in ng). The position of molecular weight markers (in kDa) are shown on the left.

3.2.4 The DDR1 kinase domain is autoinhibited by the JM4 region

Following the unsuccessful production of free enzyme (uninhibited) crystal structures of the DDR1 kinase domain and JM4 region, other constructs were assessed for their amenability to crystallographic analysis. Reasoning that the Y569F/Y586F mutation, which was inactive in the full-length receptor (**Figure 10**), might stabilise DDR1, a Y569F/Y586F-JM4K (DDR1 residues 566-913) mutant construct was produced. This construct had a thermal denaturation midpoint (T_m) of 54.7°C, when analysed by differential scanning fluorimetry, which was ~3°C higher than that found for the JM4K, or 12aa-JM4K constructs (T_m = 51.8 and 51.6 °C, respectively (**Figure 18**, **Table 11**). This indicated a promising increased stability of the mutant construct. Interestingly, the inclusion of the JM4 region (12aa-JM4K, JM4K, and Y569F/Y586F-JM4K) resulted in an increased thermal stability of the protein compared with the kinase alone (6°C increase in thermal denaturation midpoint between kinase alone and JM4K). This increased stability indicates that the JM4 region may physically interact with the kinase domain.

Crystallisation screening of the Y569F/Y586F-JM4K enzyme resulted in the production of protein crystals which were successfully used for X-ray diffraction data collection. The structure of the Y569F/Y586F-JM4K construct was solved at a resolution of 2.58 Å using PDB:5BVW (Murray et al, 2015) for molecular replacement (**Table 12**) . The asymmetric unit of the crystal contained 3 almost identical molecules (chains A (green), B (blue), and C (pink). **Figure 19b**). Chains A and C form a symmetric interface which involves the JM4 regions and αC-helices. The B chain has the same orientation as chain C and forms a crystallographically symmetric interface with a DDR1 molecule (B*) in an adjacent unit cell that resembles that of chains A and C.

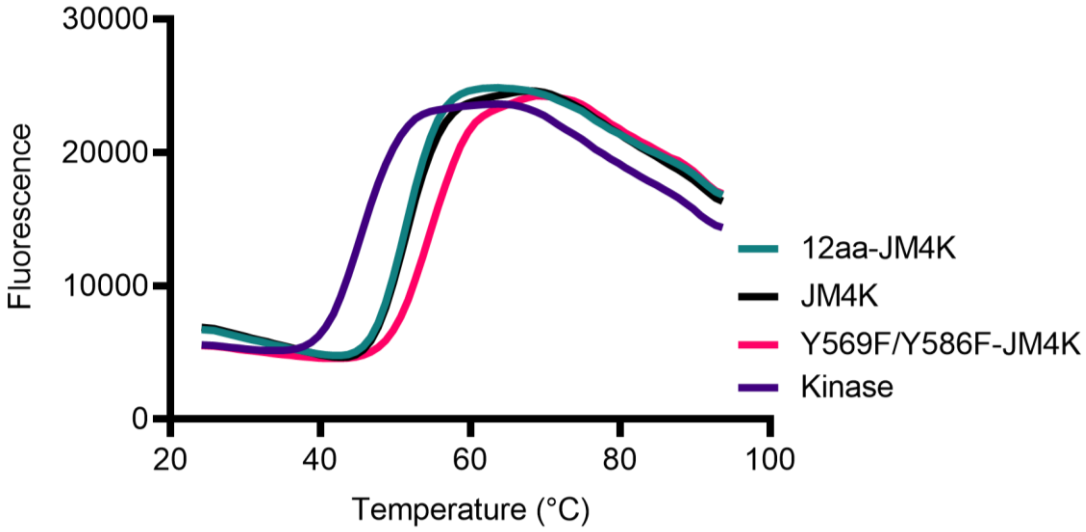

Figure 18. Differential scanning fluorimetry analysis of DDR1 constructs.
The soluble DDR1 kinase constructs 12aa-JM4K, JM4K, Y569F/Y586F-JM4K, and kinase alone were analysed by differential scanning fluorimetry. Proteins were heated from 24 to 94°C in the presence of SYPRO orange, and fluorescence detected (excitation at 492 nm; emission at 610 nm). The average fluorescence from 3 independent repeats is shown.

Table 11. Soluble DDR1 construct melting temperature midpoint.
The soluble DDR1 kinase constructs 12aa-JM4K, JM4K, Y569F/Y586F-JM4K, and kinase alone were analysed by differential scanning fluorimetry. Proteins were heated from 24 to 94°C in the presence of SYPRO orange, and fluorescence detected (excitation at 492 nm; emission at 610 nm). Data were fitted with the Boltzmann equation and the average denaturation midpoint (T_m) calculated from 3 independent experiments. Standard error of the mean is shown (±).

Protein	T_m (°C)
12aa-JM4K	51.6 (± 0.41)
JM4K	51.8 (± 0.29)
Y569F/Y586F	54.7 (± 0.58)
Kinase alone	45.8 (± 0.17)

Chain C of the asymmetric unit had the best electron density for the regions of interest; therefore, the following description is based on chain C. Clear electron density was observed for residues 575-592 of the JM4 region. The preceding residues present in the protein construct (residues 566-575) were largely invisible due to disorder. The portion of the protein connecting the JM4 segment to the kinase domain (residues 593-598) also had poor electron density. The JM4 region enters the catalytic cleft of the kinase domain between the N- and C-lobes (**Figure 19c**). The JM4 residues 577-587 form a Gly-rich hairpin between the αC-helix and the A-loop, deep within the catalytic site. The A-loop of the kinase domain is present in the inactive conformation with the catalytic Asp784 in the **D**FG-OUT conformation (**Figure 19f**). Tyr796 of the A-loop also forms a pseudosubstrate interaction with the catalytic loop Asp766 (**Figure 19g**). The wedge shaped JM4 insertion into the active site supports A-loop autoinhibition by reinforcing the stabilisation of an inactive conformation, blocking the transition to an active conformation.

There was a lack of clear electron density joining the JM4 region to the kinase domain. Furthermore, DDR1 molecules within the crystal unit cell were closely associated (C and A) around this same region of poor electron density. Therefore, it was unclear whether the JM4 region was interacting with the DDR1 active site intramolecularly (in *cis*) or intermolecularly (in *trans*). However, the protein was found to be monomeric in solution when analysed by SEC-MALS (**Figure 20, Table 13**). Therefore, the sequence of amino acids was built such that the JM4 region formed an intramolecular, in *cis,* interaction with the kinase domain.

The JM4 region was found to make contacts with all the key catalytic residues within the kinase domain (**Figure 19e**). The invariant catalytic lysine (Lys655, of the N-lobe, β3 strand), which typically coordinates the α- and β-phosphates of ATP is forced to interact with Asp784 of the A-loop in the presented structure, as a result of the positioning of Thr581 of the JM4 region. Val624 of the Gly-rich kinase loop (connects the β1 and β2 strands), which in the active structure coordinates the ATP base, instead forms contacts with the JM4 region Gly-rich hairpin. The HRD motif (His764, Arg765, Asp766) of the catalytic loop interacts with Asn584 and Thr585 of the JM4 region. The DFG motif (Asp784, F785, Gly786) is held in the inactive **D**FG-OUT conformation as a result of the interaction between Asp766 (HRD) and Tyr796 (A-

loop) which is reinforced by the JM4 region. The regulatory hydrophobic spine (Leu687 of the β4 strand, Met676 of the αC-helix, and His764 of the catalytic loop) is completed in the active conformation by the positioning of Phe785 of the DFG motif (Endicott et al, 2012). In the autoinhibited structure presented here, the mutated Phe586 residue of the JM4 region instead occupies the regulatory spine Phe785 back-pocket space. Wild-type Tyr586 likely binds in a similar manner as the mutated Phe586 used in the production of this structure. The pocket occupied by the mutant Phe residue is large enough to accommodate the Tyr hydroxyl group, with potential hydrogen bonding partners coming from the backbone carbonyl oxygens of Ile782 and Ala783.

Table 12. Crystallographic statistics for DDR1 Y569F/Y586F-JM4K.

Data Collection	
Protein	Y569F/Y586F-JM4K (Val566-Val913)
Beamline	i24
Wavelength (Å)	0.9686
Space Group	$P6_322$
Cell dimensions	
a (Å)	105.3
b (Å)	105.3
c (Å)	406.9
Resolution (Å)	58.13–2.58 (2.62–2.58)
Unique reflections	43468 (2087)
$CC_{1/2}$	0.979 (0.443)
R_{pim}	0.128 (1.13)
Mean I/σ (I)	6.5 (1.2)
Completeness	100.0 (99.1)
Multiplicity	20.3 (20.9)
Refinement	
R_{work}/R_{free}	0.225/0.270
RMSD	
Bonds (Å)	0.0027
Angles (°)	0.57
Ramachandran plot	
Favoured (%)	95.4
Outliers (%)	0

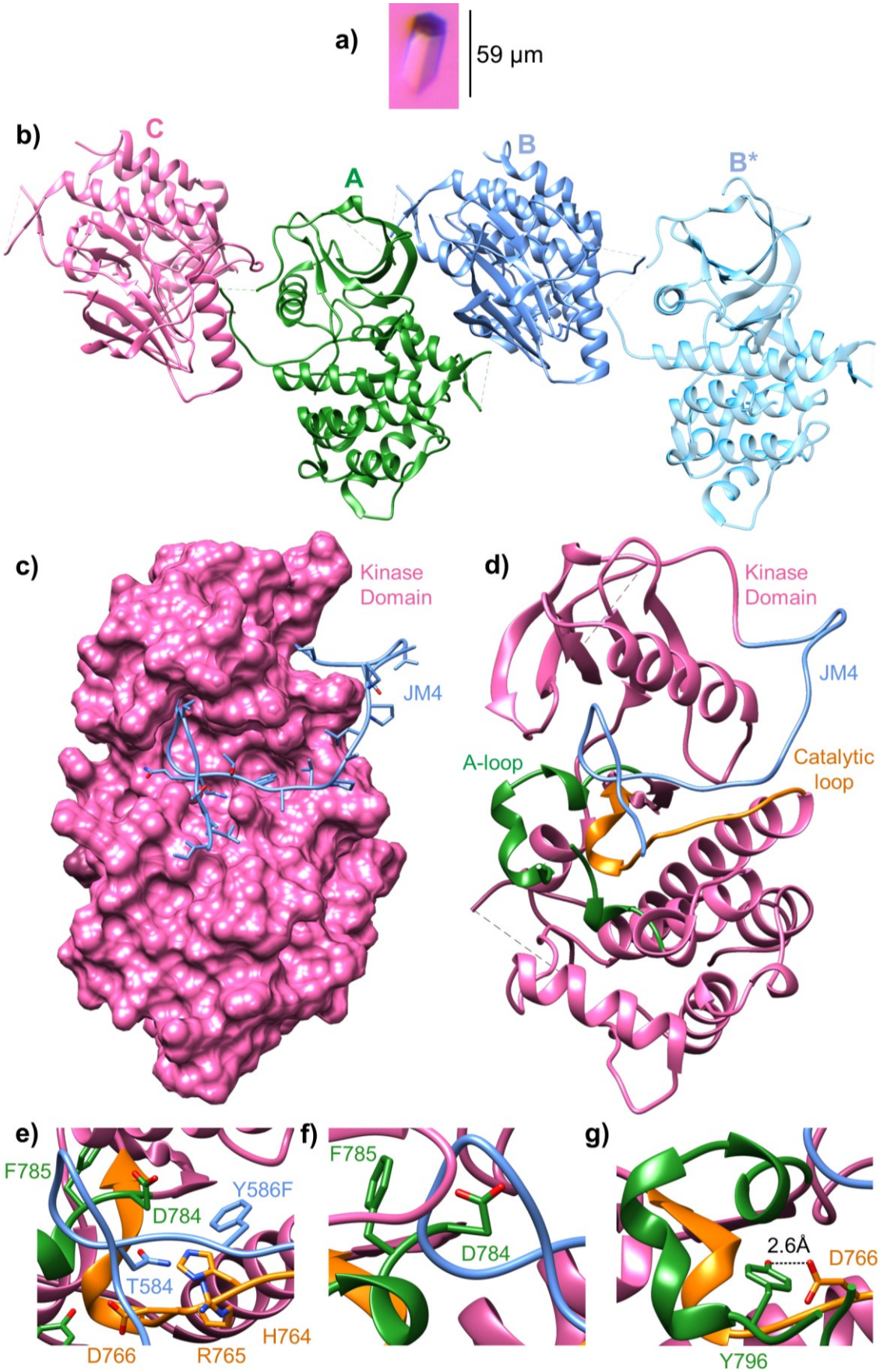

Figure 19. The DDR1 kinase is autoinhibited by the JM4 region.

a) Image of a DDR1 Y569F/Y586F-JM4K crystal. A scale-bar is shown on the right. **b)** Crystal structure of the DDR1 Y569F/Y586F-JM4K construct. The three copies of the construct within the crystal asymmetric unit are shown (chain A - green, chain B - blue, chain C - pink) with a symmetry mate shown in light blue (B*). **c)** Surface representation of the kinase domain (pink) with ribbon depiction of the crystallographically resolved JM4 region (blue; residues 574-591 and residues 592-600). **d)** Ribbon cartoon of the crystal structure in the same orientation as **c),** with kinase domain in pink, A-loop in green, catalytic loop in orange, and JM4 region in blue. **(e-f)** Close-up view of the kinase active site with residues labelled and shown in atomic detail. Atomic distance is shown in **g)** between Tyr796 and Asp766.

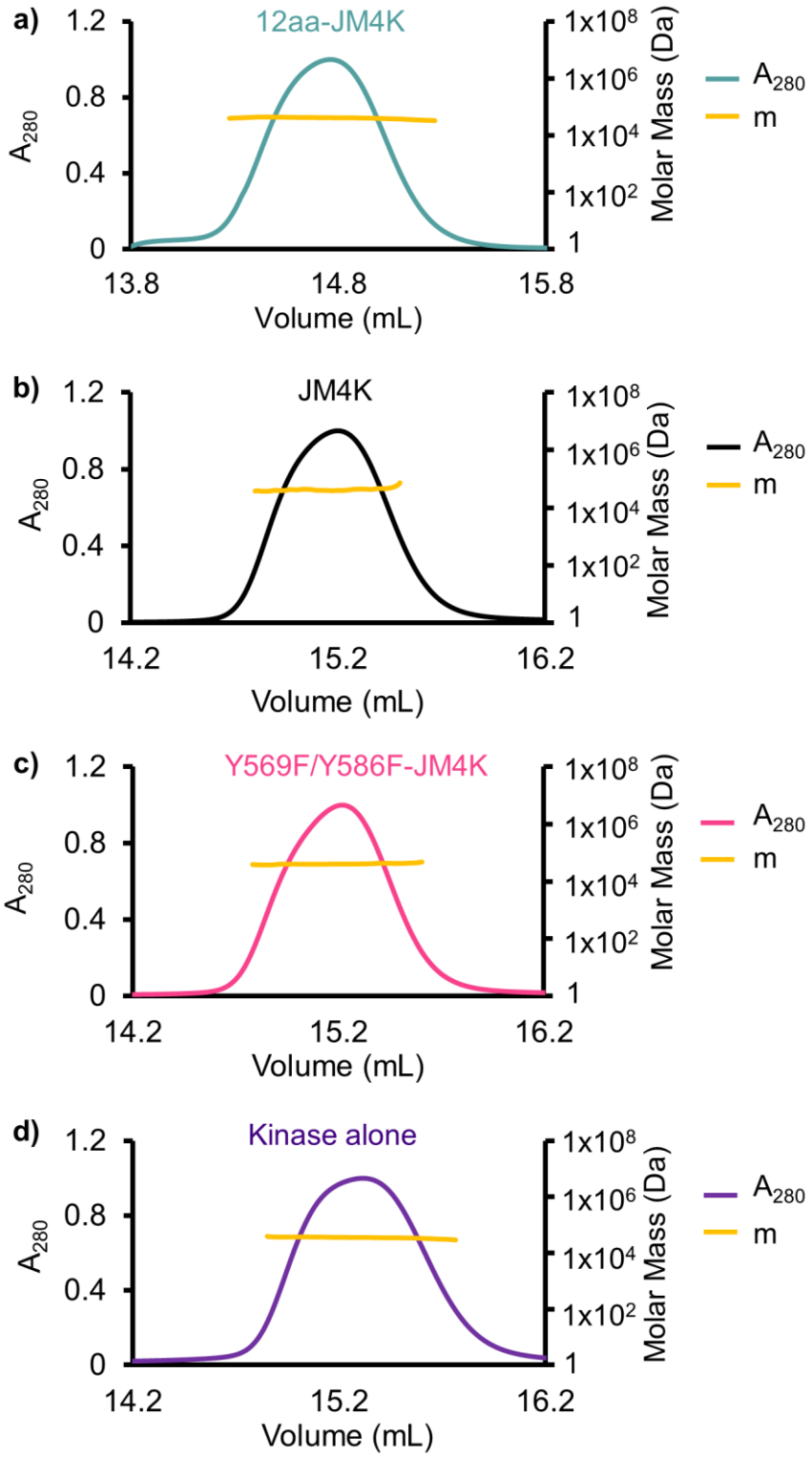

Figure 20. Molecular mass determination of soluble DDR1 kinase constructs. **(a-d)** Molecular mass was determined by SEC-MALS using the signals from absorbance (A_{280}), refractive index and light scattering detectors, for the DDR1 constructs 12aa-JM4K, JM4K, Y569F/Y586F-JM4K, and kinase alone. The UV absorbance (A_{280}) and molar mass (Da) are shown.

Table 13. Molecular mass of DDR1 kinase constructs.
The expected and SEC-MALS calculated, experimental, molecular weights for the DDR1 constructs 12aa-JM4K, JM4K, Y569F/Y586F-JM4K, and kinase alone, are shown in kDa.

SEC-MALS	12aa-JM4K	JM4K	Y569F/Y586F-JM4K	Kinase alone
Experimental	40.5	37.6	40.4	35.6
Expected	40.1	38.9	38.9	35.5

3.2.5 Inhibited and autoinhibited DDR1 kinases are structurally distinct

The Y569F/Y586F-JM4K structure has significant differences to the previously described chemically inhibited structures of the DDR1 kinase alone. Superimposition of the Y569F/Y586F-JM4K structure with DDR1/dasatinib (PDB:5BVW; (Murray et al, 2015) showed that the αC-helix, which in the active state reorients towards the active site (Hubbard, 1997), is pushed further out due to the increased size of the JM4 region hairpin compared with dasatinib (**Figure 21a**). Furthermore, the catalytically essential salt bridge between Lys655 and Glu672, which is present in the DDR1/dasatinib structure, is lost in the autoinhibited structure (**Figure 21b,c**). The JM4 region does not occlude the ATP binding space to the same extent as the inhibitors, but would still clash with the β- and γ-phosphates of ATP (**Figure 22a**; ATP superimposed from active c-KIT - PDB: 1PKG) (Mol et al, 2003). The pocket occupied by Phe586 in the presented structure does form part of the binding site for the type-II inhibitor imatinib (**Figure 22b**) (Canning et al, 2014). Interestingly, while the non-selective inhibitor, dasatinib, does not occupy any of the JM4 region binding space (**Figure 22c**) (Murray et al, 2015), an inhibitor which has shown great selectivity for the DDRs, molecule 5N, occupies both the Phe586 space and some of the N-terminal JM4 region space (**Figure 22d**) (Wang et al, 2018). This may explain why this compound has an increased specificity for the DDRs, unlike the promiscuous imatinib and dasatinib.

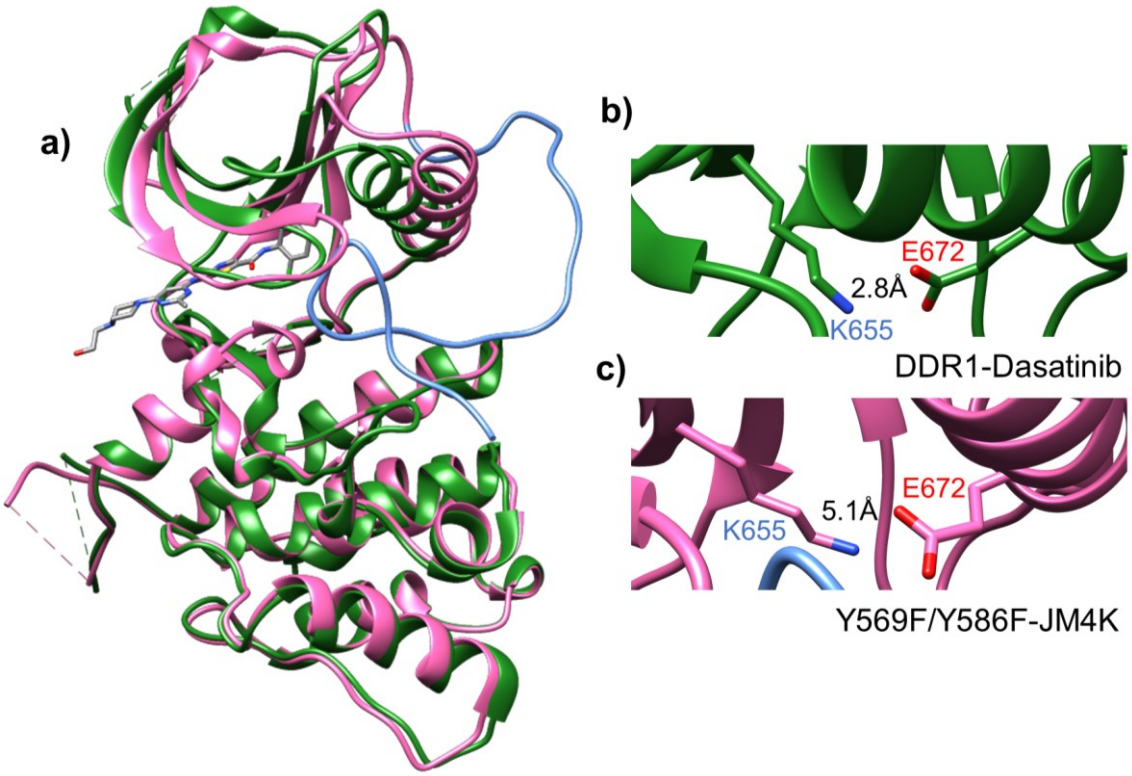

Figure 21. Comparison of Y569F/Y586F-JM4K and dasatinib inhibited DDR1. **a)** Superimposition of DDR1 Y569F/Y586F-JM4K (kinase domain in pink JM region (res 574-600) in blue), and dasatinib inhibited kinase alone (PDB:5BVW; shown in green) (Murray et al, 2015). **b,c)** Atomic details of the interaction between Lys655 and Glu672 in the dasatinib inhibited **b)** and JM4 autoinhibited **c)** DDR1 structures, with atomic distances between the two residues shown. Colouring is the same as in **a).**

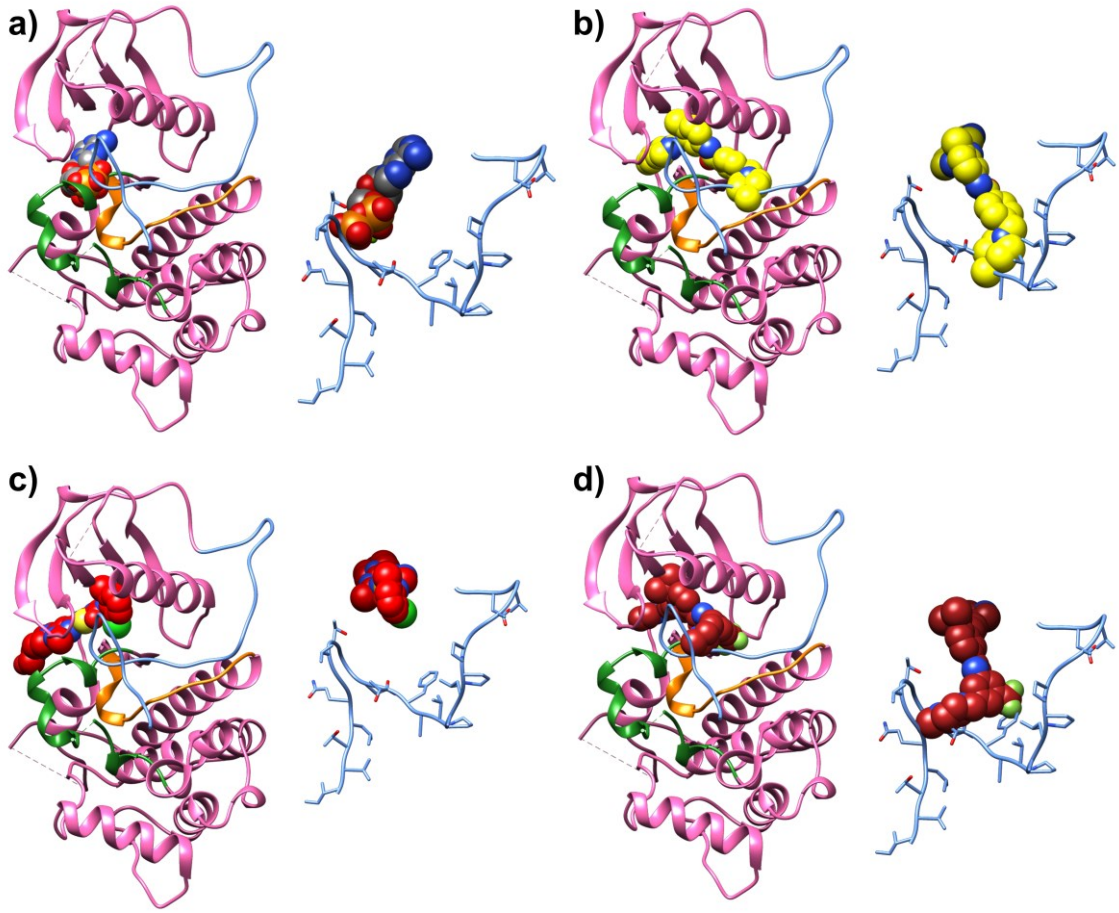

Figure 22. DDR1 inhibition and autoinhibition.
a) Left – DDR1 Y569F/Y586F-JM4K structure is shown with the kinase domain in pink, A-loop in green, and catalytic loop in orange. ATP has been superposed onto the structure from an active c-KIT structure (PDB: 1PKG; (Mol et al, 2003) and is shown in sphere depiction. Right – The JM4 region (residues 574-591) is shown with atomic detail in blue (residues 592-599 also shown in blue) and the ATP molecule superposed. **b-d)** As in **a)** but with superpositions of imatinib (yellow), dasatinib (red), and molecule 5N (brown) from PDB:4BKJ (Canning et al, 2014), PDB:5BVW (Murray et al, 2015), and PDB:6GWR (Wang et al, 2018), respectively.

3.2.6 DDR1 autoinhibition has parallels in other kinases

Structural parallels can be drawn between the JM autoinhibitory interactions reported here, with those seen in the PDGFR family. A prototypical example of this can be described for c-KIT (Mol et al, 2004). In the inactive, autoinhibited structure of c-KIT (PDB:1T45), the JM region enters the space underneath the αC-helix reinforcing the inhibited conformation of the A-loop (**Figure 23a**). The hydrophobic back-pocket (**Chapter 1, Section 1.1.2**) occupied by Phe586 in the Y569F/Y586F-JM4K structure is occupied by Trp557 in c-KIT, another bulky aromatic residue (**Figure 23a**). There are however notable differences in the autoinhibitory mechanism. The JM of c-KIT, for example, does not form a hairpin within the catalytic cleft and as a result does not penetrate as deeply into the active site as the JM4 region (**Figure 23a,b**). Furthermore, the αC-helix, and kinase Gly-rich loop (connecting the β1 and β2 strands), which must move down in the active conformation, are pushed further out in the Y569F/Y586F-JM4K structure. The phosphorylation of c-KIT JM region leads to a relief of autoinhibition and JM region expulsion from the active site (**Figure 23a,b**) (Mol et al, 2003), this provides clues to the mechanism for relief of JM4 autoinhibition in DDR1. However, the Tyr residues (Tyr568 and Tyr570) which are phosphorylated in the c-KIT JM region are not located within the kinase catalytic cleft (**Figure 23b**), like the JM4 Tyr586 residue. This indicates that there may be some mechanistic divergence in kinase activation.

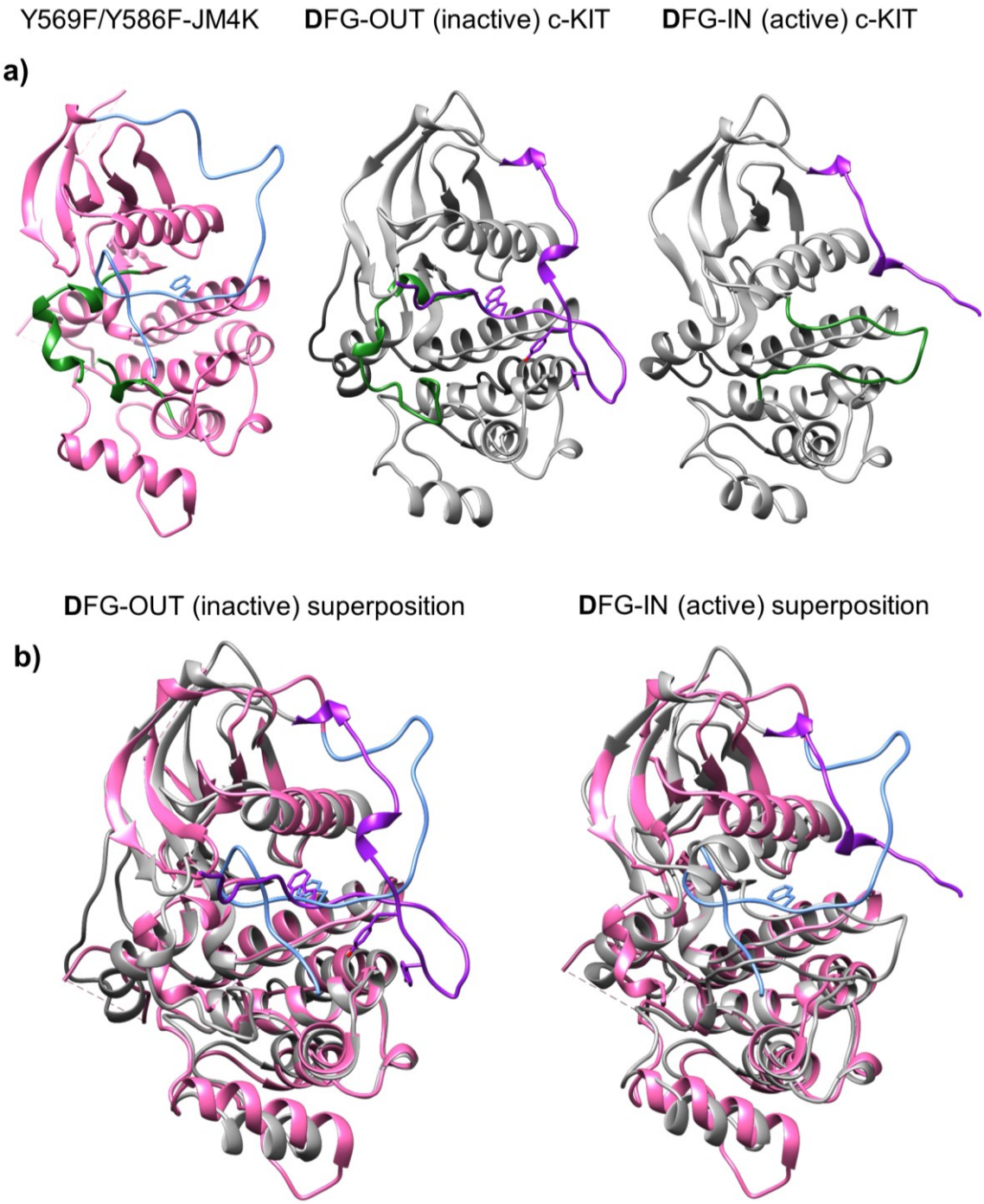

Figure 23. Superimposition of Y569F/Y586F-JM4K with c-KIT.
a) DDR1 Y569F/Y586F-JM4K structure (pink, blue JM region and green A-loop) is shown in the same orientation as structures of autoinhibited (DFG-OUT) and active (DFG-IN) c-KIT kinase (grey, with purple JM region and green A-loop (PDB: 1T45 (Mol et al, 2004) and PDB: 1PKG (Mol et al, 2003), respectively). Atomic details of Phe586 of DDR1 and Trp557, Tyr568, and Tyr570 of autoinhibited c-KIT are shown. **b)** Superimposition of the Y569F/Y586F-JM4K structure with the autoinhibited (left) and active (right) c-KIT structures. DDR1 kinase domain is shown in pink with the JM4 region in blue, c-KIT kinase is shown in grey with the JM region in purple.

3.3 Discussion

The association of aberrant RTK signalling with disease is now supported by a substantial body of clinical data (Yamaoka et al, 2018). As a result, research into the activation mechanism of this family of proteins has been extensive (Hubbard, 2013b; Lemmon & Schlessinger, 2010; Ward et al, 2007). Indeed, for many RTKs, the sequence of events which culminate in kinase activation following ligand binding are well established (Hubbard, 2013b; Lemmon & Schlessinger, 2010; Ward et al, 2007). However, despite numerous disease associations, this is not the case for the DDRs (Leitinger, 2014). Binding of collagen to DDR1 leads to the formation of large supramolecular receptor clusters, within which DDR1 becomes activated (Corcoran et al, 2019; Juskaite et al, 2017; Yeung et al, 2019). How this aggregation leads to DDR1 activation is currently unknown. Another atypical DDR feature is their long intracellular JM regions. Other than a DDR1b/c-specific NPXY motif (Tyr513), which binds to Shc upon phosphorylation, the JM region has no discernible functional motifs (Shrivastava et al, 1997; Vogel et al, 1997). In this chapter, the evolutionarily conserved, kinase-proximal, JM region, JM4, is shown to autoinhibit the kinase domain.

Initial screening of the wild type ligand-free DDR1 constructs did not lead to crystal growth. Furthermore, the JM4 region could not be resolved in a 12aa-JM4K/dasatinib crystal structure (**Figure 16**). Therefore, the Y569F/Y586F-JM4K construct was screened in crystallisation trials in the hopes that this mutant construct, which is inactive in the full-length receptor (**Figure 10**), would be more amenable to crystallisation. Screening of this construct led to the production of protein crystals. X-ray structure analysis revealed that the JM4 region forms a hairpin which penetrates the kinase catalytic cleft, reinforcing A-loop autoinhibition (**Figure 19**). Extensive contacts are made between the JM4 region and many of the catalytically essential kinase motifs. These include the HRD motif, the DFG motif, Lys655, the αC helix, and the Gly-rich loop. The lack of observable JM4 density in the dasatinib co-crystal is likely due to displacement of JM4 upon dasatinib binding. Although dasatinib does not share any of the JM4 region binding space in the kinase active site, it does cause significant rearrangement of the kinase structure. In particular, the αC-helix moves towards the active site in the dasatinib bound kinase (**Figure 21**). This movement

would cause a clash with Thr581 and Val580 of the JM4 region which would likely eject the JM4 region. Furthermore, dasatinib-kinase interactions are likely to be significantly stronger than the protein-protein interactions present at the JM4-kinase interface. Indeed, dasatinib IC50 values for DDR1 and DDR2 are 15.7 and 36.3 nM, respectively (Liu et al, 2017). Therefore, dasatinib would have competed with the JM4 region for kinase binding. Dasatinib binding would then have likely resulted in a structural rearrangement of the kinase domain and a loss of JM4 binding in the catalytic cleft. Once free from the kinase active site, the JM4 region would not be constrained in a single conformation, and consequently would become invisible by X-ray diffraction.

A challenge in building the Y569F/Y586F-JM4K model was the poor electron density for the region joining JM4 to the kinase domain (residues Pro593-Asp598). The lack of clear electron density in the model, and the close interaction of two DDR1 molecules within the crystal unit cell, meant that it was possible to build a feasible intramolecular (within the same protein) or intermolecular (between adjacent proteins) interaction between the JM4 region and kinase domain. However, the finding that the soluble DDR1 kinase constructs are monomeric in solution suggested that this interaction occurred intramolecularly (**Figure 20**). Furthermore, in all other JM region autoinhibited RTK structures, the JM region has been shown to form its interactions intramolecularly (Binns et al, 2000; Hubbard, 2004; Mol et al, 2004; Till et al, 2002). Together, this information gave confidence to complete the model as presented; with an intramolecular, in *cis*, JM4-kinase interaction.

Regulation of RTK activity by the JM region has been reported previously (Binns et al, 2000; Hubbard, 2004; Mol et al, 2004; Till et al, 2002). However, the mechanism through which this is achieved varies widely. In the Eph receptor family, the JM region forms contacts with the N- and C-lobe which leads to a distortion of the αC-helix and a disruption of the catalytically essential Lys-Glu salt bridge (Binns et al, 2000). In this mechanism, the JM region functions allosterically, making no contacts with any of the active site residues. In the MuSK receptor, the JM region recruits a cellular repressor of signalling which maintains the inactive receptor state in *trans* (Herbst & Burden, 2000; Till et al, 2002). Whilst both these mechanisms contrast significantly with DDR1 autoinhibition, surprising parallels can be drawn between the regulation

of the PDGFR family and that reported here for DDR1. The PDGFR JM region, as exemplified here with c-KIT (**Figure 23**), enters the catalytic cleft and contacts many of the kinase catalytic motifs, in much the same way as the DDR1 JM4 region (Hubbard, 2004; Mol et al, 2004). Furthermore, the JM region of both DDR1 and c-KIT, whilst extensive in their catalytic cleft contacts, do not completely occlude the ATP binding site. Despite these similarities, there are clear distinctions between the autoinhibited c-KIT and DDR1 structures. For example, the DDR1 JM4 region penetrates the space between the N- and C-lobes far deeper than the c-KIT JM region. This, in conjunction with the tight hairpin formed by JM4 region within the catalytic cleft, means that the JM4 region makes significantly more contacts with the kinase domain than the c-KIT JM region.

In all reported cases of RTK autoinhibition by the JM region, phosphorylation of Tyr residues within the JM region results in a release of autoinhibition (Binns et al, 2000; Hubbard, 2004; Mol et al, 2004; Till et al, 2002). There are two Tyr residues within the JM4 region (Tyr569 and Tyr586), both of which were mutated to Phe in the presented structure. The mutant Phe586 (Tyr586 in WT) was found to sit in the hydrophobic regulatory spine of the kinase; the space occupied by Phe785 (DFG motif) in the active conformation (Jura et al, 2011). Phosphorylation of this residue would likely result in an ejection of the JM4 region from the active site, freeing the A-loop to move into the active **D**FG-IN conformation. While JM4 region phosphorylation likely relieves autoinhibition, it is not clear whether the JM4 region has a secondary function following phosphorylation. For example, the insulin receptor JM region is autoinhibitory; however, following phosphorylation it facilitates the formation of an active signalling dimer (Cabail et al, 2015; Craddock et al, 2007). Interestingly, phosphorylated JM4 Tyr569 and Tyr586 peptides, have been shown to bind the non-receptor tyrosine kinase, Src (Lemeer et al, 2012). Src is important for DDR2 activation (Ikeda et al, 2002; Yang et al, 2005; Zhang et al, 2013); however, its function in DDR1 signalling is unclear. The role for Src upon recruitment to DDR1 is an important area for further research and may be a key secondary function of the DDR1 JM4 region.

The DDRs are essential for development (Leitinger, 2014). DDR1-null mice are deficient in mammary gland, kidney, and inner ear development (Gross et al, 2004;

Meyer zum Gottesberge et al, 2008; Vogel et al, 2001). Furthermore, DDR2-null mice have bone morphogenic disorders (Labrador et al, 2001). However, DDR function in adult tissue is less clear. DDR1 may be involved in wound healing by promoting cell migration (Hou et al, 2001). However, most of our understanding of DDR function in adults comes from disease associations (Borza & Pozzi, 2014). DDR1 has been implicated in numerous cancers, atherosclerosis, and kidney and lung fibrosis (Avivi-Green et al, 2006; Franco et al, 2008; Guerrot et al, 2011; Valiathan et al, 2012). This makes DDR1 an interesting target for pharmacological intervention. Indeed, there have been extensive small-molecule inhibitor studies of the DDRs (Day et al, 2008; Murray et al, 2015; Wang et al, 2018; Zhu et al, 2019). Early studies showed that the Bcr-Abl type I inhibitor dasatinib, and the type II inhibitor imatinib, were potent DDR inhibitors (Day et al, 2008). Some success has been made since then at generating more specific DDR inhibitors; however, all are ATP competitive (Kim et al, 2013; Wang et al, 2018; Zhu et al, 2019). As the ATP binding site is extremely well conserved between kinases, these inhibitors are fundamentally promiscuous (Hanson et al, 2019). Indeed, DDR1 is one of the most promiscuous proteins in the RTK family, binding to over 100 inhibitors (Hanson et al, 2019). While this promiscuity may be useful in some cases, for example, cancers where the targeting of multiple kinases could have therapeutic benefits, it increases the chance of deleterious off-target effects (Zhang et al, 2008). The discovery of space outside this conserved ATP binding pocket which can be targeted with inhibitors is therefore an important area of research (Wu et al, 2015a). Unlike the kinase domain, RTK JM regions are not well conserved between receptors, and the mechanism through which they regulate activity vary (Binns et al, 2000; Hubbard, 2004; Mol et al, 2004; Till et al, 2002). This heterogeneity could support the development of specific inhibitors. The space occupied by the DDR1 JM4 region could offer a tractable region for the development of novel DDR1 inhibitors. In particular, the deep pocket occupied by the tip of the JM4 region hairpin (Val580-Gly583) is an attractive candidate which is not exploited by current inhibitors. Indeed, one of the most specific DDR1 inhibitors developed to date utilises some of the JM4 region binding space (Wang et al, 2018). Alternatively, JM4-kinase interactions could be reinforced. For example, a TrkA inhibitor has been developed which formed contacts with both the kinase domain and the JM region, leading to enhanced specificity (Furuya et al, 2017). The

amenability of targeting the JM4 region with small molecules, or peptide mimetics, should be assessed in the hopes of developing novel DDR-specific inhibitors.

In summary, the structural basis for autoinhibition of DDR1 has been demonstrated. The highly conserved JM4 region reinforces A-loop autoinhibition. The unique nature of the JM4-kinase interactions makes targeting the JM4 region occupied space an interesting target for pharmacological intervention. However, an improved understanding of the functional effects of these JM4-kinase interactions, and the mechanism through which these are relieved, must be assessed.

Chapter 4. Two-step release of DDR1 autoinhibition

4.1 Introduction

Kinase inhibitors are currently in use for the treatment of several diseases (Du & Lovly, 2018; Mirshafiey et al, 2014; Patterson et al, 2014). Most of these inhibitors are ATP-competitive (Hartmann et al, 2009). The high sequence conservation of the ATP binding site makes these types of inhibitor fundamentally promiscuous (Hanson et al, 2019). Therefore, the discovery of sites outside this conserved binding pocket, which can be pharmacologically targeted, is an important area of research (Wu et al, 2015a). In the previous chapter, structural analysis revealed that the highly conserved DDR1 JM4 region (residues 566-591) forms contacts with the kinase domain which reinforce autoinhibition by the A-loop. This dually autoinhibited conformation (by the A-loop and JM4 region) is structurally unique amongst RTKs in the extent of the JM-kinase interactions, and the depth to which the JM4 region penetrates the catalytic cleft. Furthermore, while the JM4 region forms considerable active site contacts, it does not occlude the adenine base binding pocket which is typically occupied by promiscuous ATP-competitive inhibitors (Hartmann et al, 2009). Targeting of JM4 region occupied space with inhibitors could provide greater DDR1 specificity, which may have therapeutic applications. In aid of this, the functional significance of DDR1 autoinhibition by the A-loop and JM4 region, and the mechanism through which this is relieved, must be better understood.

A-loop autoinhibition is a common RTK autoinhibitory mechanism and in most receptors the A-loop is the first motif to be phosphorylated (Lemmon & Schlessinger, 2010). A-loop phosphorylation relieves autoinhibition and results in a substantial increase in kinase catalysis. In IGF-1R, for example, A-loop trans-phosphorylation results in a ~600-fold increase in k_{cat}/K_m (Li & Miller, 2006). Phosphorylation of the DDR1 A-loop has been demonstrated in the past and is predicted to relieve A-loop-mediated autoinhibition (Canning et al, 2014; Juskaite et al, 2017; Yang et al, 2005).

While autoinhibition by the JM region has been reported for a number RTKs, the mechanism through which this is achieved varies (Mol et al, 2004; Sammon et al, 2020; Till et al, 2002). However, there are significant structural parallels between the autoinhibited DDR1 and PDGFR family. In both, the JM region enters the active site

forming multiple interactions with the kinase domain which maintain an inactive conformation (Mol et al, 2004; Sammon et al, 2020). The PDGFR JM region contains a conserved YΦY motif (Φ; hydrophobic residue), the phosphorylation of which relieves autoinhibition (Hubbard, 2004; Mol et al, 2004). The unexpected structural similarities between DDR1 and PDGFR indicates that they may be activated through a similar mechanism. However, the YΦY motif, which is phosphorylated in the PDGFR JM region, is located outside the kinase catalytic cleft (**Figure 23**), in a space not occupied by the DDR1 JM4 region, this suggests that there may be some mechanistic divergence (Mol et al, 2004; Sammon et al, 2020).

4.1.1 Hypothesis and aims

It is hypothesised that the JM4 region and A-loop are autoinhibitory motifs which are relieved upon autophosphorylation. Furthermore, targeting of the JM4 region binding space with peptide mimetics is expected to inhibit kinase function.

The aims were: 1) to determine whether the A-loop and JM4 region are phosphorylated; 2) to establish whether phosphorylation of these regions relieves autoinhibition; 3) to design JM4 mimetic peptides which function as inhibitors.

4.1.2 Findings

DDR1 autoinhibition is relieved through a two-step mechanism beginning with phosphorylation of the JM4 region in *cis,* this results in an intermediate level of kinase activity. In the second step, the A-loop is phosphorylated, rendering the kinase fully active. Furthermore, a JM4 peptide mimetic functions as an effective DDR1 inhibitor. This demonstrates the potential for designing JM4 region mimetic modulators of DDR1 activity, which may have therapeutic applications.

4.2 Results

4.2.1 The JM4 region inhibits DDR1 catalysis

To understand the mechanism through which DDR1 JM4 region and A-loop autoinhibition is relieved, several soluble DDR1 kinase constructs were purified following expression in Sf9 cells. These included: JM4K (DDR1 kinase domain with the JM4 region – residues 566-913), Y569F/Y586F-JM4K (JM4K with both JM4 Tyr replaced by Phe), and kinase alone (DDR1 kinase domain alone – residues 601-913) (**Figure 11**). As discussed in **Chapter 3**, biophysical analysis revealed that the inclusion of the JM4 region in these constructs (JM4K and Y569F/Y586F-JM4K) resulted in an increased thermal stability of the protein (**Figure 18**, **Table 11**). Furthermore, the constructs were all found to be monomeric in solution (**Figure 20, Table 13**). Together these data supported the finding that the JM4 region interacts intramolecularly with the kinase domain.

4.2.1.1 Production and optimisation of tools for monitoring DDR1 activation

In addition to the soluble kinase constructs outlined above, it was important to develop tools which would allow monitoring of DDR1 activation. A DDR1-specific antibody was already available which allowed monitoring of A-loop phosphorylation (anti-pY796, commercially available (Corcoran et al, 2019; Juskaite et al, 2017)). However, no JM4-specific antibodies were commercially available. Therefore, two anti-phosphotyrosine specific antibodies were commissioned which bound to phosphorylated Tyr569 (anti-pY569) or Tyr586 (anti-pY586) of the JM4 region. Initial validation of these antibodies was conducted by the producers (Biomatik) who analysed cross-reactivity with the unphosphorylated peptides using ELISA and dot-blotting. To test if the JM4 region was phosphorylated in full-length DDR1, DDR1-WT was expressed in HEK293 cells which were subsequently stimulated with collagen I. Cells were then lysed and analysed by SDS-PAGE and Western blotting with the two JM4-specific anti-phosphotyrosine antibodies, a pan anti-phosphotyrosine antibody (anti-pY), as well as the DDR1 A-loop specific anti-phosphotyrosine antibody (anti-pY796). Stimulation of cells with collagen resulted in a substantial increase in the anti-pY569 and anti-pY586 phospho-signal relative to unstimulated controls (**Figure 24**). Furthermore, there was minimal binding of the

antibodies to other proteins on the membrane. Additionally, binding to unstimulated DDR1 was comparable to that seen using anti-pY and anti-pY796 (**Figure 24**). Therefore, these antibodies can specifically detect collagen stimulated autophosphorylation of full-length DDR1.

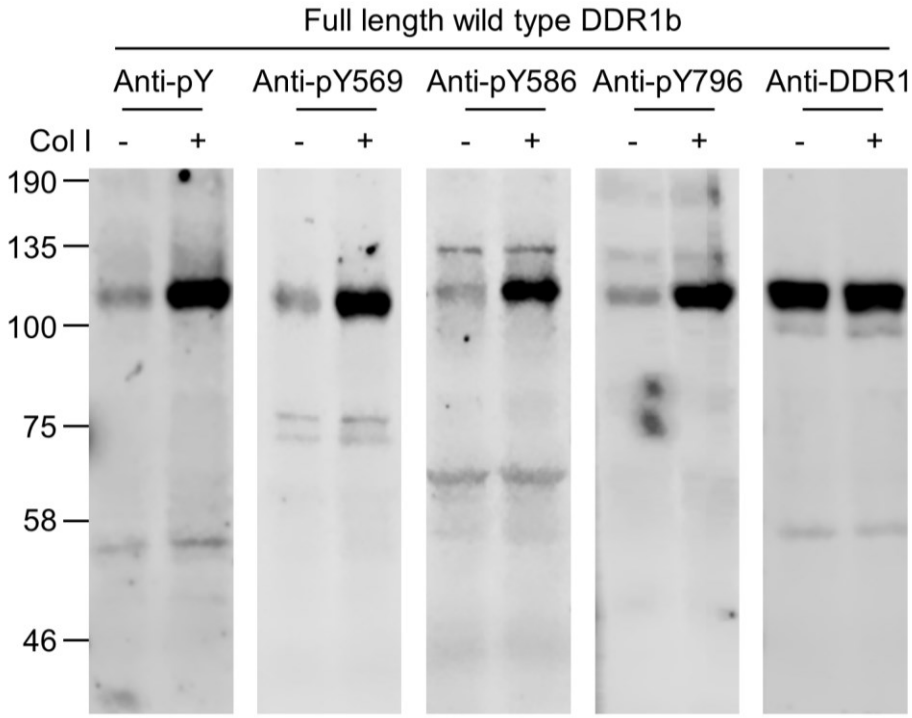

Figure 24. Collagen-induced JM4 phosphorylation in full-length DDR1.
Full-length wild type DDR1 was transiently expressed in HEK293 cells. Cells were then stimulated with collagen I (+) or left unstimulated (-) for 90 minutes at 37°C. Cells were subsequently lysed, and aliquots of cell lysates analysed by SDS-PAGE and Western blotting with the JM4-specific anti-phosphotyrosine antibodies, anti-pY569 and anti-pY586, as well as the A-loop-specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. Representative blots from 3 independent experiments are shown.

Following confirmation that the JM4-specific anti-phosphotyrosine antibodies were able to detect collagen-induced DDR1 phosphorylation, the linear range of signal detection by Western blotting was next assessed. To achieve this, the soluble DDR1 JM4K construct was incubated at 100 µM in the presence of 20 mM ATP for 6 hours to stimulate autophosphorylation. 12.5–200 ng of this protein was then separated by SDS-PAGE and analysed by Western blotting with the indicated anti-phosphotyrosine and anti-DDR1 antibodies (**Figure 25a**). All antibodies produced a linear Western blot signal between 25 and 100 ng protein (**Figure 25b**). The anti-pY586 antibody showed somewhat weaker signal at lower protein load levels compared with the other antibodies (**Figure 25a**). To ensure that the Western signal was within the linear range of signal detection, 35 ng of protein was loaded for the anti-pY796 and anti-pY569 antibodies, and 50 ng was used for the anti-pY586 antibody. 35–50 ng of protein was also within the linear range of detection for the anti-DDR1 antibodies. A second lower molecular weight band was also observed at higher protein load levels (≥50 ng) in all except the anti-pY569 blot, which likely represents a protein degradation product.

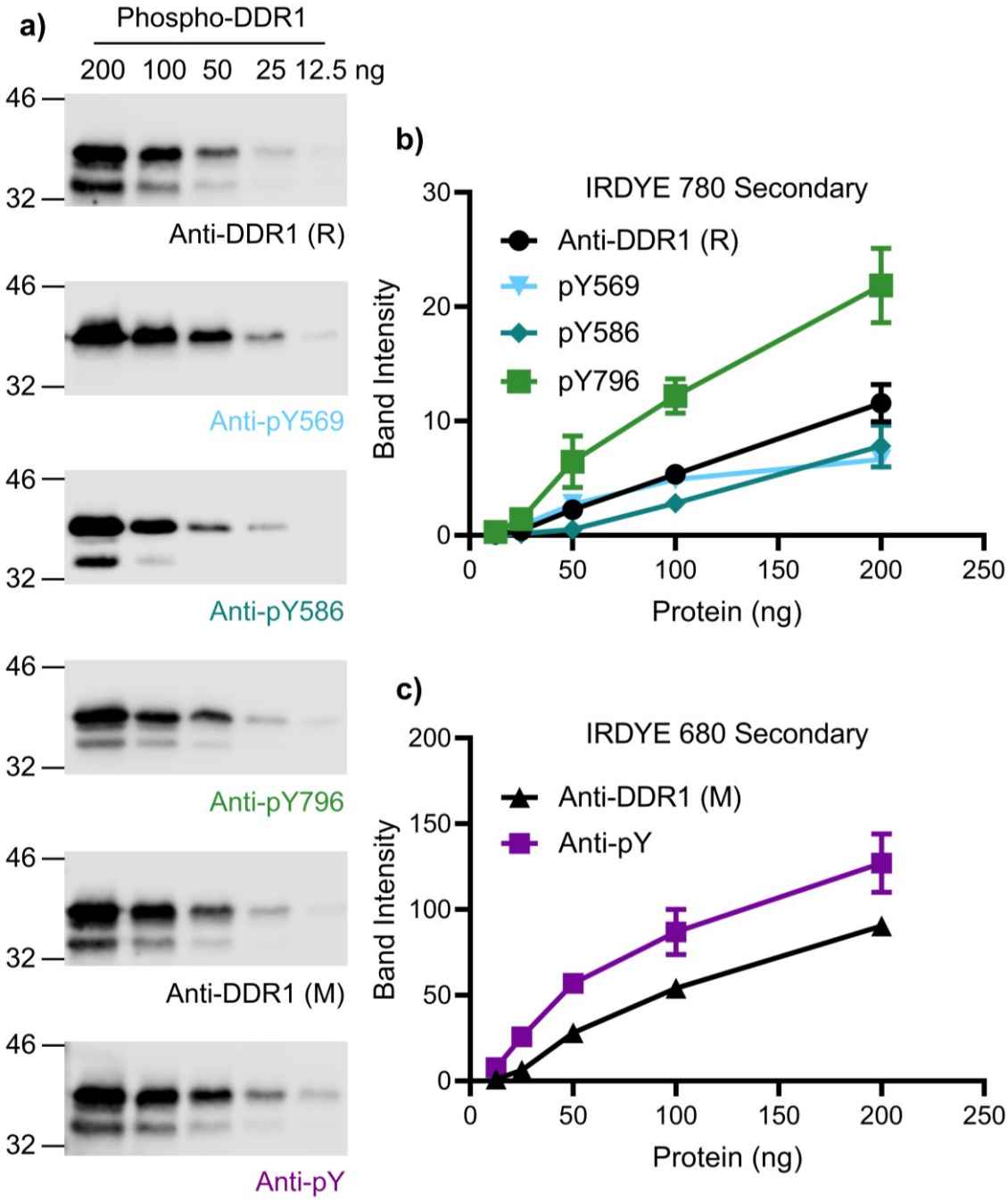

Figure 25. Relationship between Western blot signal and protein load.
a) Soluble DDR1 JM4K was incubated at 100 µM in the presence of 20 mM ATP and kinase buffer II for 6 hours, the phosphorylated protein was then titrated from 12.5–200 ng, boiled in sample buffer and then analysed by SDS-PAGE and Western blotting with the indicated antibodies. The position of molecular weight markers (in kDa) are shown on the left. Representative blots from 2 independent experiments are shown. **b)** Quantification of pY-DDR1 and total-DDR1 signals detected using IRDYE 780 secondary anti-rabbit antibody: Rabbit (R) anti-DDR1, anti-pY569, anti-pY586, and anti-pY796. **c)** Quantification of anti-pY (4G10) and mouse (M) anti-DDR1 signals detected using IRDYE 680 secondary anti-mouse antibody. *n*=2 mean and range shown in **b,c).**

4.2.1.2 JM4 is autophosphorylated intramolecularly

To assess the function of the JM4 region in kinase activation, the soluble DDR1 kinase constructs (JM4K, Y569F/Y586F-JM4K, and kinase alone) were assessed for their ability to autophosphorylate *in vitro*. Initially however, the enzyme's dependence on the divalent metal ions Mg^{2+} and Mn^{2+} was assessed. The JM4K construct was incubated at 1 µM in the presence of 1 mM ATP and 10 mM Mg^{2+}, or Mn^{2+}, or both metals, for 30 minutes at 20°C. Autophosphorylation by the kinase was then assessed using the pan anti-phosphotyrosine antibody (anti-pY - **Figure 26**). Phosphorylation of JM4K was only observed in the presence of Mn^{2+}. However, a slight increase in the phospho-signal was observed upon coincubation with Mn^{2+} and Mg^{2+}, therefore, all future kinase buffers contained both metals.

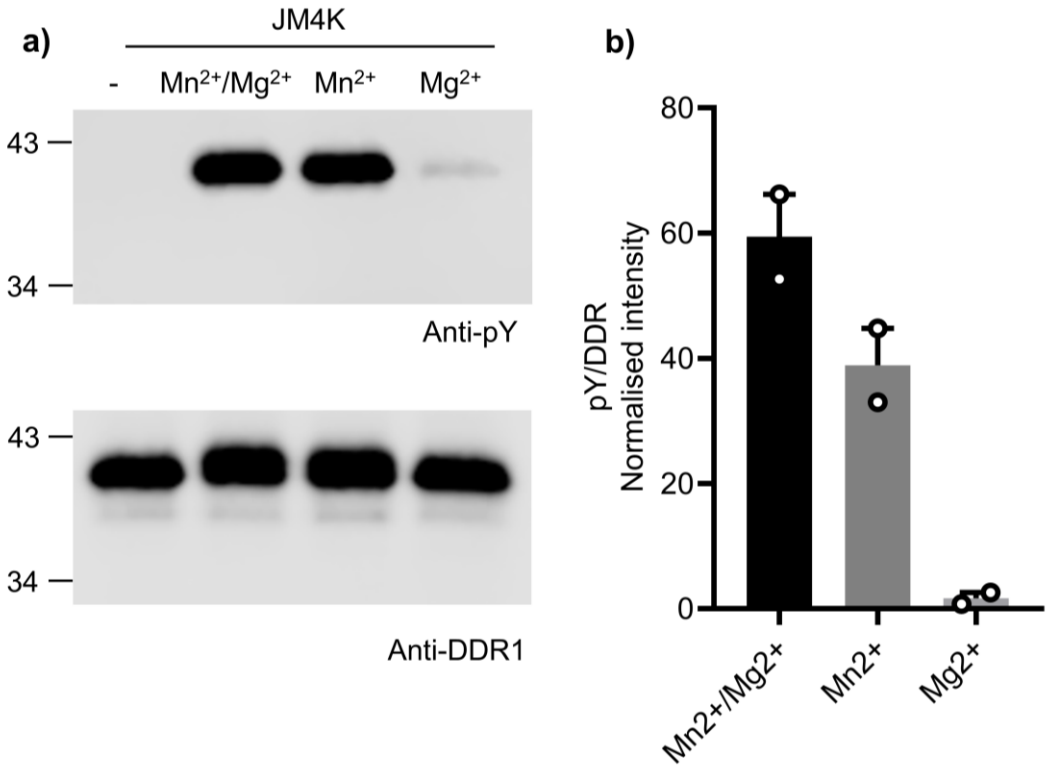

Figure 26. DDR1 autophosphorylation buffer optimisation.
a) The soluble DDR1 JM4K construct was incubated at 1 µM with 1 mM ATP in kinase buffer I in the presence or absence of 10 mM $MgCl_2$ and 10 mM $MnCl_2$ for 30 minutes at 20°C. The sample in the first lane was incubated in the absence of ATP (-). The protein was then boiled in sample buffer and analysed by SDS-PAGE and Western blotting with the pan anti-phosphotyrosine antibody 4G10 (anti-pY). Total DDR1 was then detected with an anti-DDR1 antibody. The positions of molecular weight markers (in kDa) are indicated on the left. **b)** Quantification of anti-pY signals normalised to respective anti-DDR1 signals. The anti-pY signal is expressed as a percentage of the sum of all bands on a blot, with mean and range shown (*n*=2).

Following buffer optimisation, the 3 DDR1 constructs (JM4K, Y569F/Y586F-JM4K, and kinase alone) were incubated at 1 μM in the presence of 1 mM ATP across a time course (0–120 minutes). Samples were boiled at each time point and analysed by Western blotting (**Figure 27**). Phosphorylation of the JM4K construct was observed after 5 minutes of ATP stimulation using the pan-anti-phosphotyrosine antibody (anti-pY). However, no phosphorylation was seen for the kinase alone or Y569F/Y586F-JM4K constructs (anti-pY signal, **Figure 27a**), indicating that this anti-pY signal was a result of JM4 region phosphorylation. Indeed, anti-pY569 and anti-pY586 signals appeared within 5 minutes, reaching a plateau after 30 minutes (**Figure 27a,b**). Furthermore, A-loop phosphorylation, monitored using anti-pY796, could not be detected for any construct (**Figure 27a**).

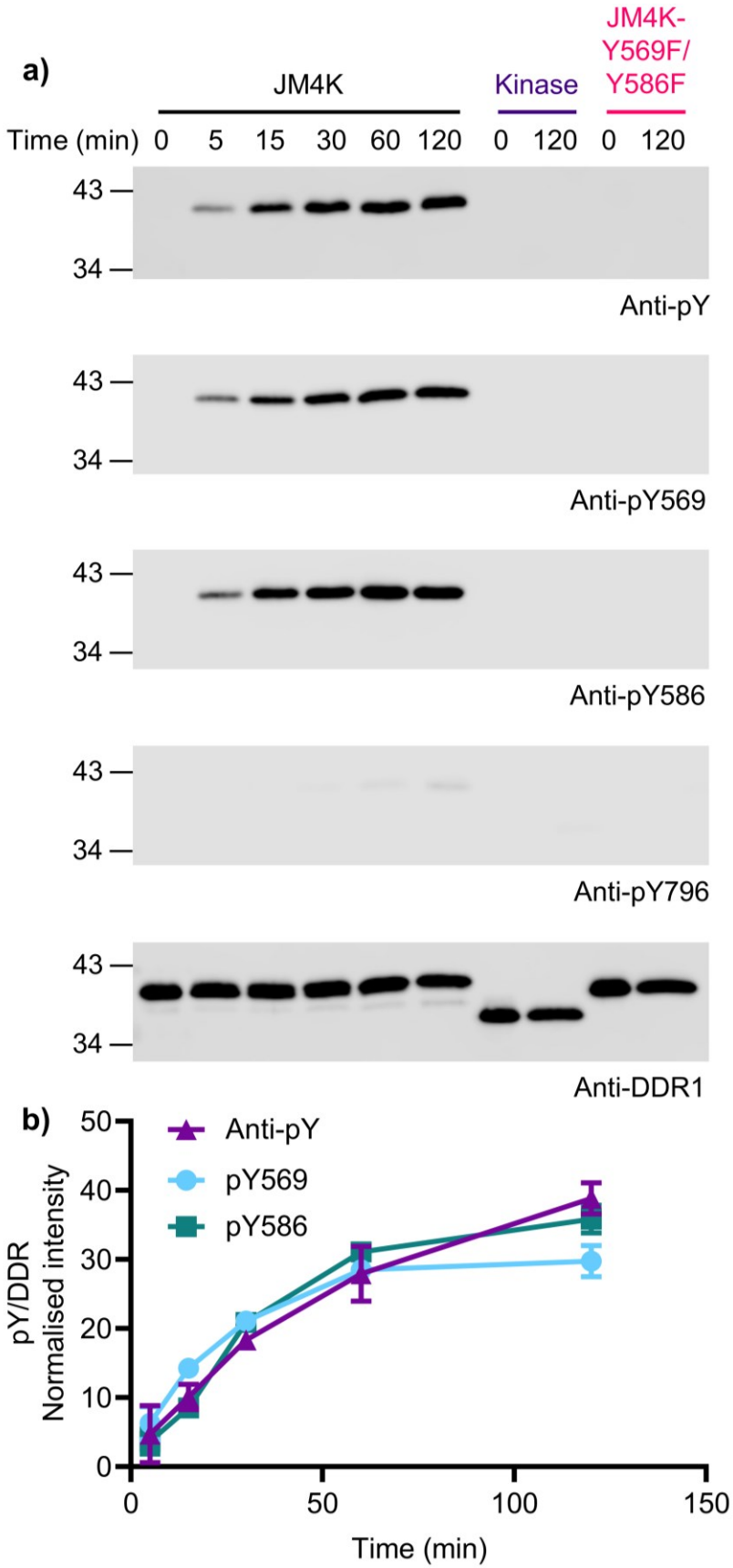

Figure 27. The JM4 region, but not the A-loop, of soluble DDR1 constructs is phosphorylated at low protein concentration.

a) The soluble DDR1 constructs, JM4K, JM4K-Y569F/Y586F, and kinase alone, were incubated at 1 μM with 1 mM ATP for the indicated times (0–120 minutes) at 20°C. Protein was then boiled in sample buffer and analysed by SDS-PAGE and Western blotting with the general anti-phosphotyrosine antibody 4G10 (anti-pY), the JM4-specific anti-phosphotyrosine antibodies, anti-pY569 and anti-pY586, as well as the A-loop-specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. **b)** Quantification of anti-phosphotyrosine signals (anti-pY, anti-pY569, and anti-pY586) normalised to respective anti-DDR1 signals. The anti-phosphotyrosine signal is expressed as a percentage of the sum of all bands on a blot, with mean and standard error of the mean shown (*n*=3).

It was next important to assess whether JM4 region autophosphorylation occurred through an intramolecular/in *cis* reaction (zero order with respect to protein concentration), or an intermolecular/in *trans* reaction (first order with respect to protein concentration). Across a 32-fold concentration range of JM4K there was no significant difference in the phospho-Tyr569 signal after 10 minutes of stimulation with 1 mM ATP (**Figure 28a,b**). This indicated that Tyr569 and Tyr586 phosphorylation occurs independently of protein concentration, and therefore occurs in *cis* (**Figure 28c**). In contrast to this independence of protein concentration, titration of ATP from 31.25–2000 nM resulted in a substantial increase in phospho-Tyr569 signal (**Figure 29a,b**). This demonstrated a clear ATP concentration dependence. Furthermore, the observation that changes in phospho-Tyr569 levels could be detected across the ATP range demonstrates that this assay is sensitive enough to detect subtle differences in autophosphorylation levels (**Figure 29b**).

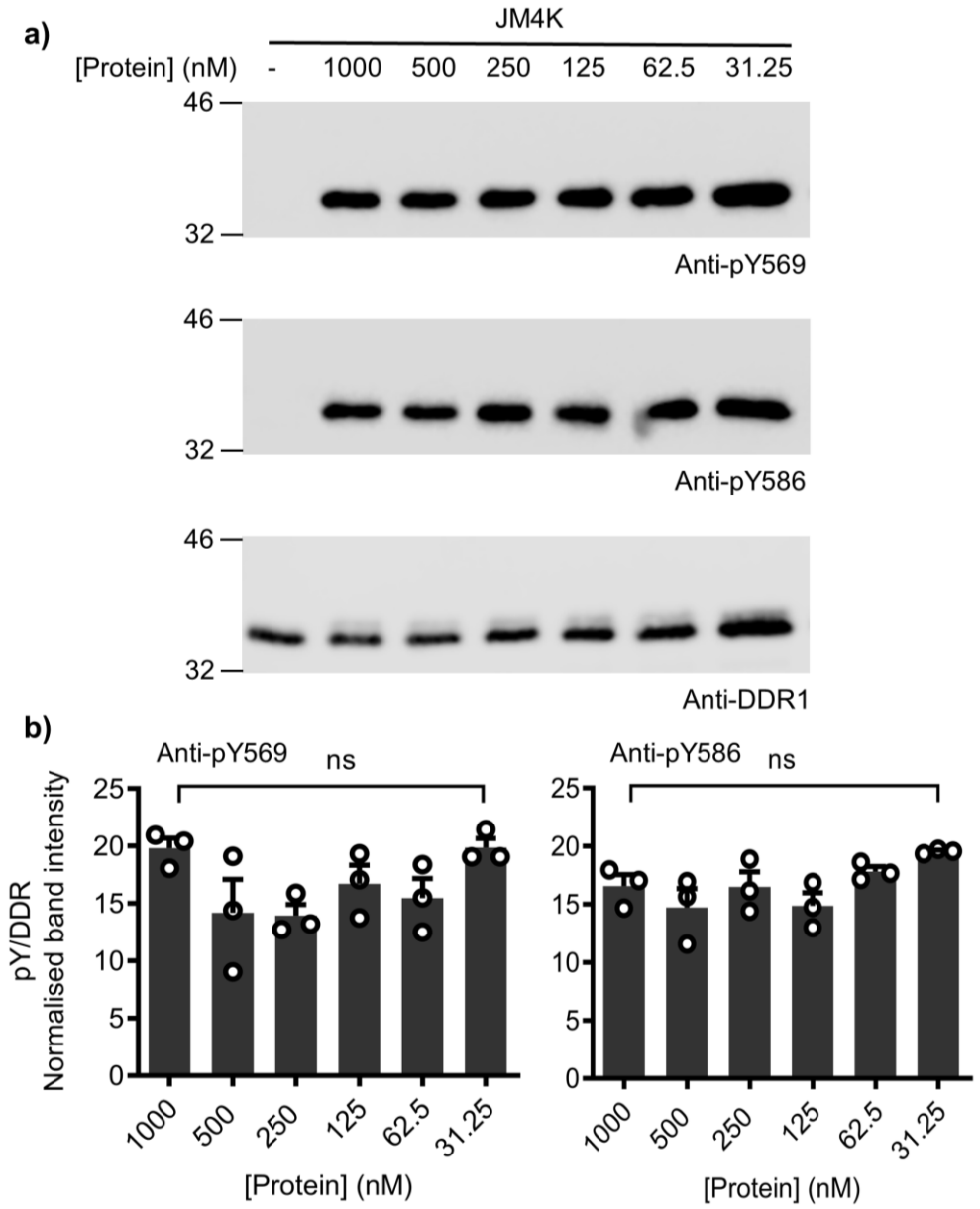

Figure 28. Protein concentration independent JM4 region autophosphorylation. **a)** The soluble DDR1 JM4K construct was titrated from 31.25–1000 nM and incubated with kinase buffer I and 1 mM ATP for 10 minutes at 20°C. The sample in the first lane was incubated in the absence of ATP (-). The protein was then boiled in sample buffer and aliquots containing equal amounts of JM4K analysed by SDS-PAGE and Western blotting with the JM4-specific anti-phosphotyrosine antibodies, anti-pY569 and anti-pY586. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. **b)** Quantification of anti-pY569 and anti-pY586 signal normalised to respective anti-DDR1 signals. The signal is expressed as a percentage of the sum of all bands on a blot, with mean and standard error of the mean shown (*n*=3). No statistically significant differences (ns) were found between any of the groups (p=0.2265 and 0.8162, 1000 nM vs 500 nM; p=0.1945 and >0.9999, 1000 nM vs 250 nM; p=0.7665 and 0.8594, 1000 nM vs 125 nM; p=0.4665 and 0.9565, 1000 nM vs 62.5 nM; p>0.9999 and 0.4013, 1000 nM vs 31.3 nM; for anti-pY569 and anti-pY586, respectively. One-way ANOVA with Tukey's post-hoc test.

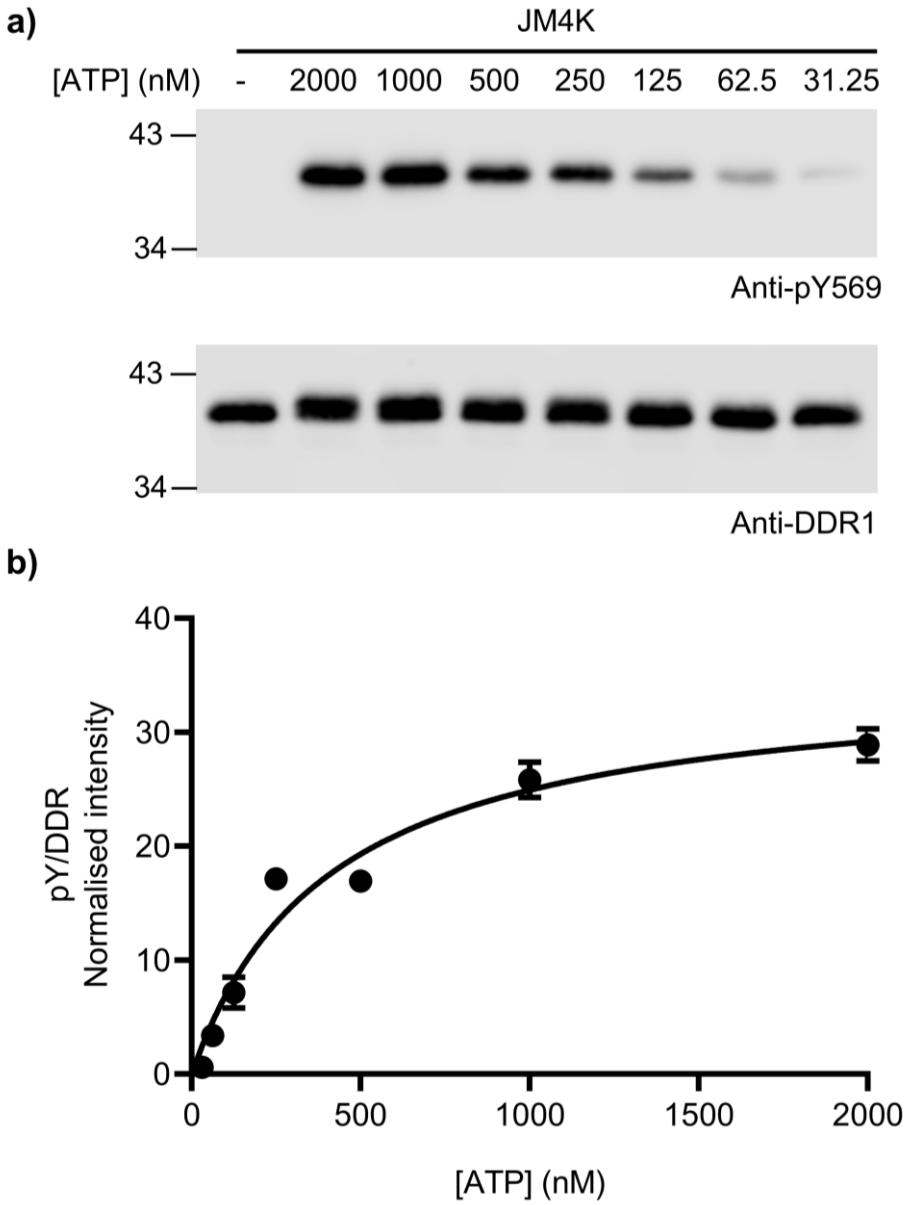

Figure 29. ATP concentration dependent JM4 region autophosphorylation.
a) The soluble DDR1 JM4K construct was incubated at 1 μM in kinase buffer I with a range of ATP concentrations (31.25–2000 nM) for 10 minutes at 20°C. The sample in the first lane was incubated in the absence of ATP (-). The protein was then boiled in sample buffer and aliquots containing equal amounts of JM4K analysed by SDS-PAGE and Western blotting with the JM4-specific anti-phosphotyrosine antibody, anti-pY569. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. **b)** Quantification of anti-pY569 signal normalised to respective anti-DDR1 signals. The anti-pY569 signal is expressed as a percentage of the sum of all bands on a blot, with mean and standard error of the mean shown (*n*=3). Fitted curve is a non-linear regression substrate response curve made using GraphPad Prism 8.

4.2.1.3 JM4 region autoinhibits substrate phosphorylation

Thus far, the DDR1 JM4 region has been shown to physically interact with the kinase domain (**Figure 19**), and the ability of the kinase to autophosphorylate the JM4 Tyr residues, Tyr569 and Tyr586, has been demonstrated. However, while these data suggest an autoinhibitory function for the JM4 region, the functional significance of this on the catalysis of substrate phosphorylation has yet to be assessed. Therefore, a commercially available kinase assay kit, ADP-Glo™, was optimised for the analysis of DDR1 phosphorylation of a peptide substrate. The substrate used was the Axltide peptide; a derivative of IRS-1 (insulin receptor substrate 1) which is a DDR2 substrate *in vitro* (Iwai et al, 2013). The ADP-Glo assay utilises two reagents: the first reagent (ADP-Glo Buffer) terminates the kinase reaction and hydrolyses any ATP remaining in the reaction mixture, which has not been used by the kinase, to AMP. A second reagent (Kinase Detection Reagent) is then added, which converts the ADP produced by the kinase into ATP. A luciferase converts this ATP into a luminescent signal which is directly proportional to the concentration of ADP produced by the kinase (Zegzouti et al, 2009). Mg^{2+} is not essential for DDR1 catalysis (**Figure 26**); however, it is required for the functioning of the ADP-Glo™ reagents. Therefore, the concentrations of Mg^{2+} and Mn^{2+} were initially optimised. Different concentrations of Mg^{2+} (5–80 mM) and Mn^{2+} (0–5 mM) were incubated with 25 µM ATP, 50 µM Axltide peptide, and 0.15 µM JM4K protein for 40 minutes at room temperature. The ADP-Glo reagents were then added and luminescence recorded (**Figure 30a**). The luminescence signal increased with decreasing Mg^{2+}, indicating inhibition of catalysis by the Mg^{2+} ions. The highest signal was observed at 5 mM Mg^{2+} with 1 mM Mn^{2+}. The 0 mM Mn^{2+} conditions produced luminescent signals around background levels (~2000 relative light units), again indicating that Mn^{2+} is essential for DDR1 catalysis. Further titration of Mg^{2+} concentration from 1–4 mM in the presence of 1 mM Mn^{2+} led to a reduction in signal compared with the 5 mM Mg^{2+} and 1 mM Mn^{2+} condition (**Figure 30b**). Therefore, future experiments were conducted with 5 mM Mg^{2+} and 1 mM Mn^{2+}.

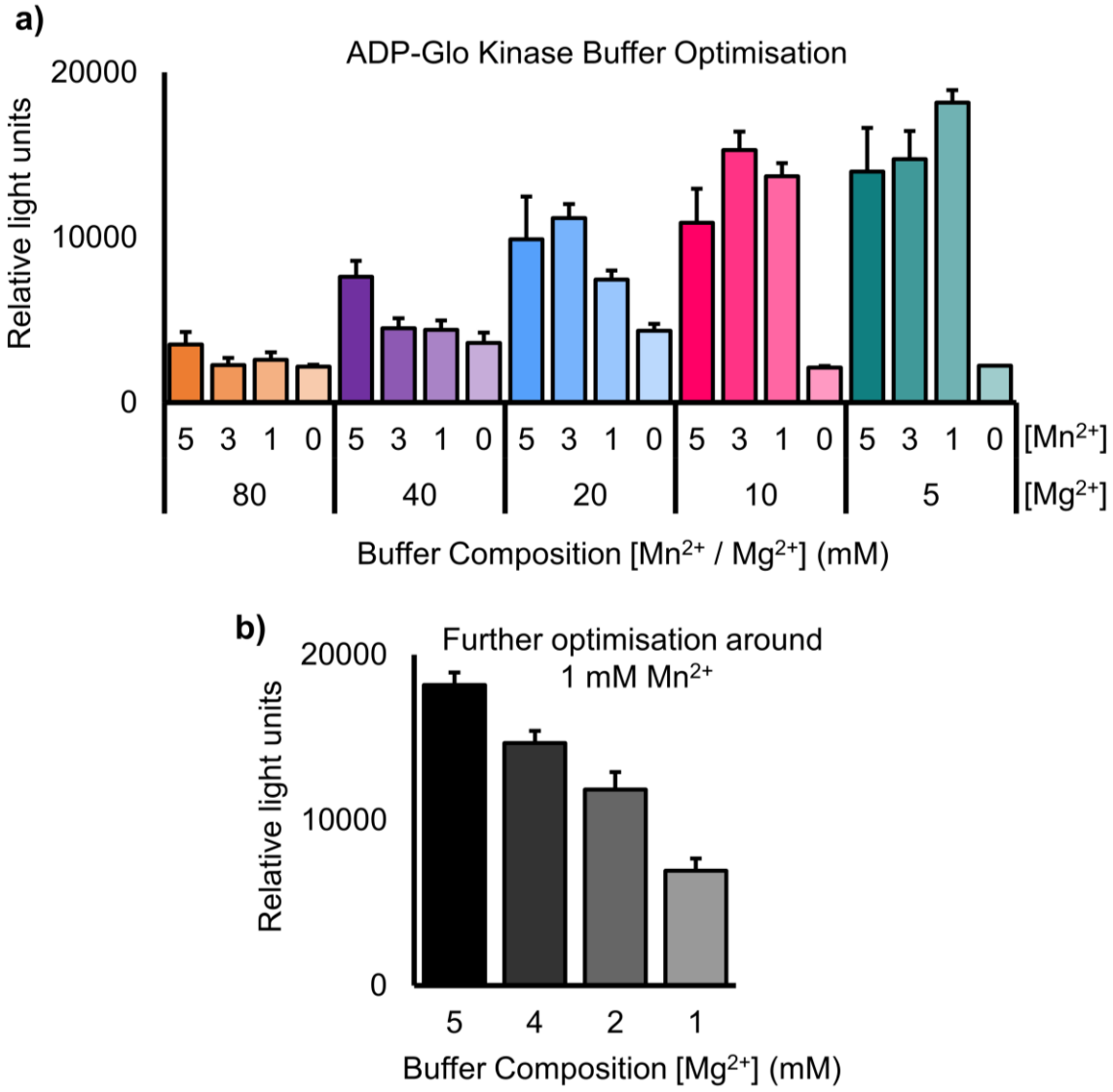

Figure 30. ADP-Glo™ kinase buffer optimisation.
a) The soluble DDR1 JM4K construct was incubated at 0.15 µM with 50 µM Axltide substrate and 25 µM ATP in the presence of kinase buffer III and a range of $MgCl_2$ (5–80 mM) and $MnCl_2$ concentrations (0–5 mM) for 40 minutes at room temperature. The ADP-Glo reagents were then added and luminescence recorded. **b)** Further optimisation of $MgCl_2$ concentrations in the ADP-Glo assay was conducted using the same conditions as in **a)** with 1 mM $MnCl_2$ and a range of $MgCl_2$ concentrations (1 - 5 mM). **a,b)**. *n* = 2 mean and range shown.

To assess the effect of the JM4 region on peptide substrate phosphorylation, the soluble DDR1 kinase constructs (JM4K, Y569F/Y586F-JM4K, and kinase alone) were titrated from 0.2–3.2 μM in the presence of 250 μM ATP and 200 μM Axltide peptide substrate and incubated for 40 minutes at room temperature. ADP-Glo™ reagents were then added and luminescence recorded (**Figure 31**). An ADP-standard curve was included which allowed conversion of relative light units recorded by the luminometer into ADP concentrations produced by the kinase through interpolation of the curve (**Figure 31a**). The kinase alone produced a significantly higher ADP concentration across the protein concentration range, indicating that it was more active than the JM4 region containing constructs (**Figure 31b**). The Y569F/Y586F-JM4K construct also produced a consistently, but not statistically significant, lower signal than the JM4K construct, indicating that it was more inactive than the wild type protein. These data demonstrated that the JM4 region inhibits substrate phosphorylation by DDR1 *in vitro*. This supports structural data which indicated an autoinhibitory JM4 region function (**Figure 19**).

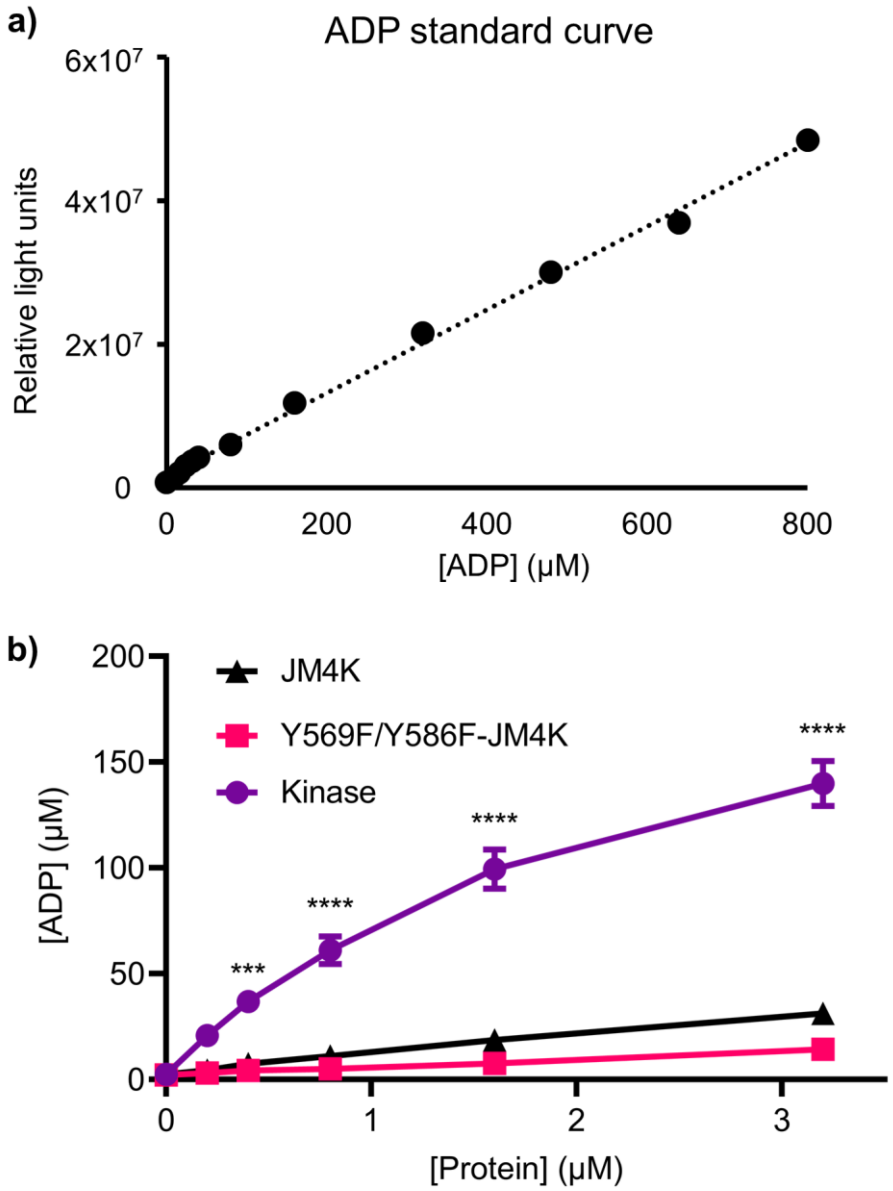

Figure 31. The JM4 region inhibits substrate phosphorylation.
a) An example ADP standard curve produced using the ADP-Glo reagents. A standard curve was included in each ADP-Glo™ analysis. **b)** The soluble DDR1 JM4K, Y569F/Y586F-JM4K, and kinase alone constructs were titrated from 0.2–3.2 μM in the presence of 250 μM ATP and 200 μM Axltide substrate in kinase buffer III for 40 minutes. The ADP-Glo™ reagents were then added and luminescence values recorded. Relative light units were then converted to [ADP] produced through interpolation of the ADP standard curve. *n* = 3, mean and standard error of the mean shown. **** - p<0.0001 statistically significant difference found between kinase alone and JM4K at 3.2, 1.6, and 0.8 μM. *** - p=0.0005 statistically significant difference found between the kinase alone and JM4K at 0.4 μM. Two-way ANOVA with a Tukey post-test was performed.

4.2.2 DDR1 autophosphorylation occurs in two distinct steps

Following the confirmation that the JM4 region is autoinhibitory (**Figure 31**), and that the JM4 region can be phosphorylated at low protein concentrations (**Figure 28**), the soluble DDR1 constructs (JM4K, Y569F/Y586F-JM4K, and kinase alone) were next incubated at higher protein concentrations to determine whether kinase domain autophosphorylation could be stimulated. All 3 proteins were incubated at 10, 40, or 100 μM in the presence of 20 mM ATP for 3 hours. The protein was then separated by native PAGE before staining with InstantBlue Coomassie (**Figure 32**). As native PAGE allows separation of proteins based on charge, as well as size and conformation, it is possible to separate phosphorylation forms (Li & Miller, 2006). Indeed, for the kinase alone it was possible to separate two different forms; an unphosphorylated form which was present in the unstimulated sample (-), and a phosphorylated form which appeared with increasing prevalence at increasing protein concentrations, indicating a protein concentration dependence for phosphorylation of the kinase domain (**Figure 32**). No shift was observed on the gel for the mutant Y569F/Y586F-JM4K construct, indicating that the protein was unable to autophosphorylate. However, the JM4K construct produced 4 distinct bands on the gel: unphosphorylated protein, two intermediate phosphorylation forms (JM4K-1P and JM4K-2P), and a fully phosphorylated form (JM4K-3P). As with the kinase alone, increased JM4K concentration led to an increased shift of the protein to the fully phosphorylated, JM4K-3P state.

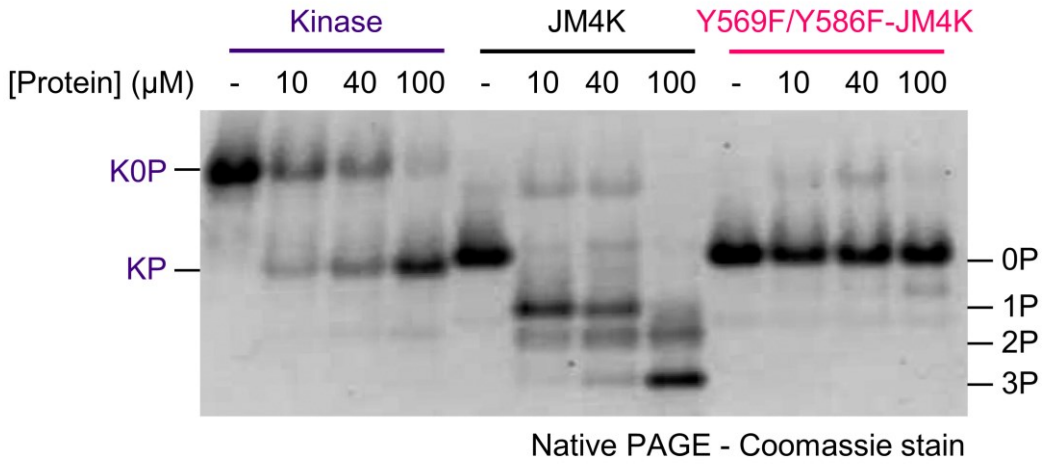

Figure 32. Kinase domain phosphorylation at higher protein concentrations. The soluble DDR1 kinase alone, JM4K, and Y569F/Y586F-JM4K constructs were incubated at 10, 40 or 100 μM in the presence of kinase buffer II and 20 mM ATP for 3 hours at 20°C, or samples were incubated in kinase buffer II in the absence of ATP (-). Protein was then separated by native PAGE and gels stained with InstantBlue Coomassie stain. Positions of the DDR1 JM4K phospho-forms, JM4K-0P, -1P, -2P, and -3P, are shown on the right of the image, while the positions of the DDR1 kinase alone phospho-forms, K0P and KP, are shown on the left. A representative image from 2 independent experiments is shown.

The time dependence for production of the phosphorylated forms of DDR1 was next assessed. The 3 DDR1 constructs (JM4K, kinase alone, and Y569F/Y586F-JM4K) were incubated at 100 μM in the presence of 20 mM ATP for 0–180 minutes. Proteins were then separated by native PAGE and stained with InstantBlue Coomassie, or boiled, separated by SDS-PAGE, and analysed by Western blotting with phosphotyrosine-specific antibodies (**Figure 33**). As in **Figure 32**, the kinase alone produced two distinct bands, whilst the JM4K construct produced 4 when assessed by native PAGE. The first JM4K intermediate phosphorylation form (JM4K-1P) appeared within 5 minutes, with very little of the unphosphorylated protein (JM4K-0P) remaining. This protein then slowly transitioned into the second intermediate phosphorylation form (JM4K-2P) which began to appear after 30 minutes, with the fully phosphorylated form (JM4K-3P) only appearing after 60 minutes. Some protein remained in the JM4K-2P state after 180 minutes of incubation, demonstrating the exceptionally slow autophosphorylation kinetics for DDR1 (**Figure 33a**). A similar slow transition to the phosphorylated form was also seen for the kinase alone, and again no shift in bands was seen for the Y569F/Y586F-JM4K construct demonstrating an almost complete loss of kinase function. Western blot analysis of

these same samples showed that the JM4 region of the JM4K construct was phosphorylated early (anti-pY569 and anti-pY586 signals, **Figure 33b**). The anti-pY569 signal did not change significantly across the time course demonstrating that phosphorylation of this residue occurs within the first few minutes of the experiment (**Figure 33c**). Furthermore, the anti-pY569 and anti-pY586 antibodies did not bind to the kinase alone, demonstrating their specificity. Kinase domain A-loop phosphorylation (anti-pY796) of the JM4K construct on the other hand was significantly slower, with very little signal detected before 60 minutes. This slow A-loop phosphorylation was also seen for the kinase alone. Native Western blotting of the JM4K time course samples (0–180 minutes) showed that JM4K-1P is phosphorylated exclusively on the JM4 region (**Figure 34**). Furthermore, A-loop phosphorylation (anti-pY796 blot) was mostly detected in the JM4K-3P state; however, some JM4K-2P phosphorylation on the A-loop could be detected (**Figure 34**). Together, these data showed that autophosphorylation of the JM4K construct occurs in distinguishable steps. In the first step, the JM4 is rapidly phosphorylated in *cis*, in the second step the kinase domain is phosphorylated on the A-loop, this likely occurs in *trans,* and at a significantly slower rate.

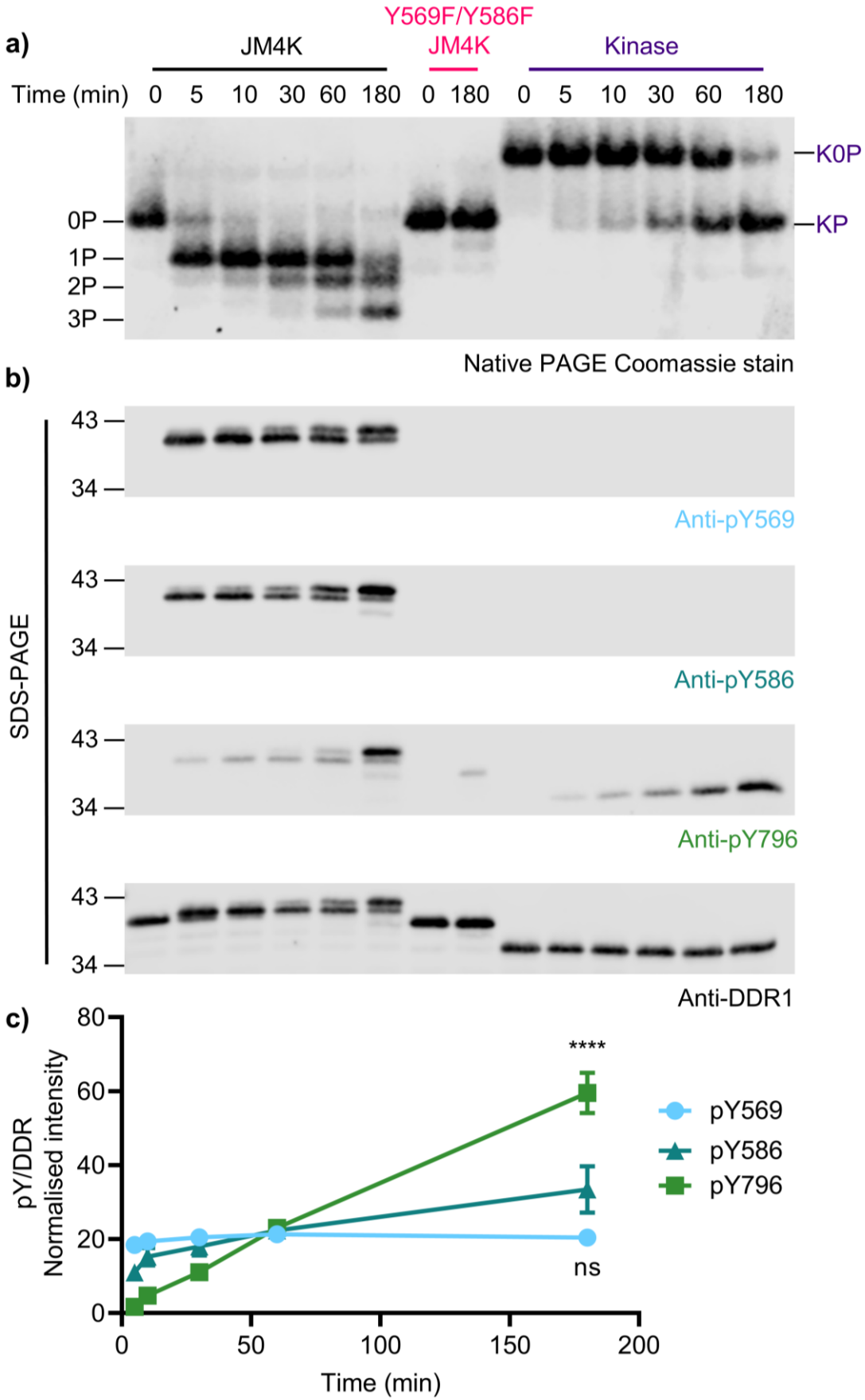

Figure 33. *In vitro* autophosphorylation of DDR1 occurs in two distinct steps.

a) The soluble DDR1 JM4K, Y569F/Y586F-JM4K, and kinase alone, constructs were incubated at 100 μM in the presence of kinase buffer II and 20 mM ATP for the indicated time course (0–180 minutes). Protein was then separated by native PAGE and gels stained with InstantBlue Coomassie stain (positions of the DDR1 JM4K phospho-forms, JM4K-0P, -1P, -2P, and -3P, are shown on the left of the image, whilst the positions of the kinase alone phospho-forms, K0P and KP, are shown on the right), or samples were boiled in sample buffer and analysed by SDS-PAGE and Western blotting with the JM4-specific anti-phosphotyrosine antibodies, anti-pY569 and anti-pY586, as well as the A-loop specific anti-phosphotyrosine antibody, anti-pY796 **b)**. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. **c)** Quantification of anti-phosphotyrosine signals (anti-pY569, anti-pY586, and anti-pY796) normalised to respective anti-DDR1 signals. The anti-phosphotyrosine signal is expressed as a percentage of the sum of all bands on a blot, with mean and standard error of the mean shown (n=4, or n=3 for anti-pY586). No statistically significant (ns) differences were found between any of the anti-pY569 time points (p=0.3222 between 5 and 180 minutes). **** - P<0.0001 statistically significant difference between the 60- and 180-minute time points for anti-pY796. One-way ANOVA with Tukey's post-test was performed for each anti-phosphotyrosine antibody.

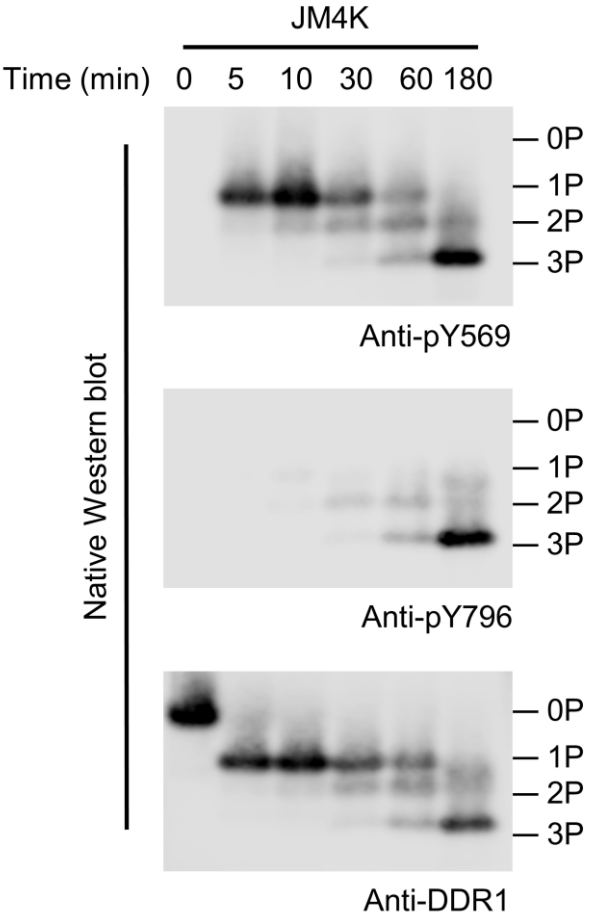

Figure 34. JM4K-3P is phosphorylated on the A-loop and JM4 region.
The soluble DDR1 JM4K construct was incubated at 100 μM in the presence of kinase buffer II and 20 mM ATP for the indicated time course (0–180 minutes). Protein was then separated by native PAGE and analysed by Western blotting with the JM4-specific anti-phosphotyrosine antibody, anti-pY569, and the A-loop specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. A representative image is shown from 2 independent experiments. Positions of the DDR1 JM4K phosphoforms (JM4K-0P, -1P, -2P, and -3P) are shown on the right.

To confirm the finding that multiple DDR1 phospho-forms are generated during autophosphorylation, Phostag™ acrylamide gel electrophoresis was used, as an alternative technique to native PAGE. Phostag™ is a compound which, when cast into acrylamide gels, allows the separation of proteins based on their phosphorylation state (Bekesova et al, 2015). The JM4K construct was incubated at 100 μM in the presence of 20 mM ATP for 0–180 minutes, as in the experiment shown in **Figure 33.** The protein was subsequently separated by native PAGE or boiled and separated using a gel containing 50 μM Phostag™ . Both the native and Phostag™ gels were then stained with InstantBlue Coomassie stain. As with the native gel, 4 distinct phosphorylation forms of the JM4K construct were detected by Phostag™ (**Figure 35**). These forms interconverted over time with similar kinetics to that observed by native PAGE. This supports the notion that there are 4 different phosphorylation states of the JM4K construct, namely; JM4K-0P (unphosphorylated), JM4K-1P (intermediate phosphorylation form 1), JM4K-2P (intermediate phosphorylation form 2), and JM4K-3P (fully phosphorylated form).

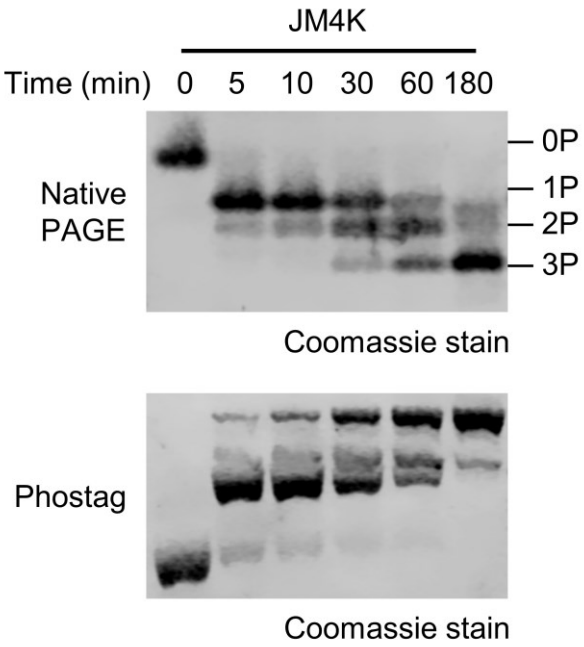

Figure 35. Soluble DDR1 JM4K produces 3 distinct phosphorylation forms. The soluble DDR1 JM4K construct was incubated at 100 µM in the presence of kinase buffer II and 20 mM ATP for the indicated time course (0–180 minutes). Protein was then separated by native PAGE and the gel stained with InstantBlue Coomassie stain (upper image), or it was boiled in sample buffer and analysed by Phostag™ SDS-PAGE and stained with InstantBlue Coomassie stain (lower image). A representative image is shown from 2 independent experiments. Positions of the DDR1 JM4K phosphoforms (0P, 1P, 2P, and 3P) are shown on the right of the upper image.

4.2.3 DDR1 phosphorylation forms can be separated by ion exchange

The significant temporal separation between JM4 region and A-loop phosphorylation led to attempts to purify these different phospho-forms for biochemical analysis. Stimulation of JM4K at 100 µM in the presence of 20 mM ATP for 90 minutes led to the production of a protein sample containing a mixture of phospho-forms (JM4K-1P, -2P, -3P) (**Figure 36b,** Input). This mix was separated by ion exchange (IEX) chromatography. Protein was found in the IEX flow through (FT), and in two elution peaks (E1 and E2) (**Figure 36a**). Native PAGE analysis of these samples revealed that the JM4K intermediate phosphorylation form 1 (JM4K-1P) was present in the FT, whilst the fully phosphorylated form (JM4K-3P) was in E2 (**Figure 36b**). E1 contained a mixture of JM4K-1P and JM4K intermediate phosphorylation form 2 (JM4K-2P), therefore this sample was not used for further analysis. Western blot analysis of the isolated JM4K phosphorylation forms (JM4K-0P, -1P and -3P), along with the kinase

alone phosphorylation forms (K0P; unphosphorylated kinase, and KP; phosphorylated kinase), revealed that both the JM4K-1P and JM4K-3P samples were phosphorylated to the same extent on the JM4 Tyr residues, Ty569 and Tyr586 (**Figure 36c**). However, only the fully phosphorylated JM4K-3P, and KP, forms showed robust A-loop phosphorylation (anti-pY796). Tandem mass spectrometry confirmed that the JM4K-1P form was phosphorylated on the JM4 region and not the A-loop, whilst the JM4K-3P form was phosphorylated on both the A-loop and the JM4 region (**Table 14**). Furthermore, mass spectrometry analysis revealed that Tyr797 of the A-loop was also phosphorylated in the JM4K-3P form. Tyr740 within the kinase domain was also found to be phosphorylated. However, phospho-Tyr740 was not consistently detected (compare JM4K-1P and 3P chymotrypsin; **Table 14**).

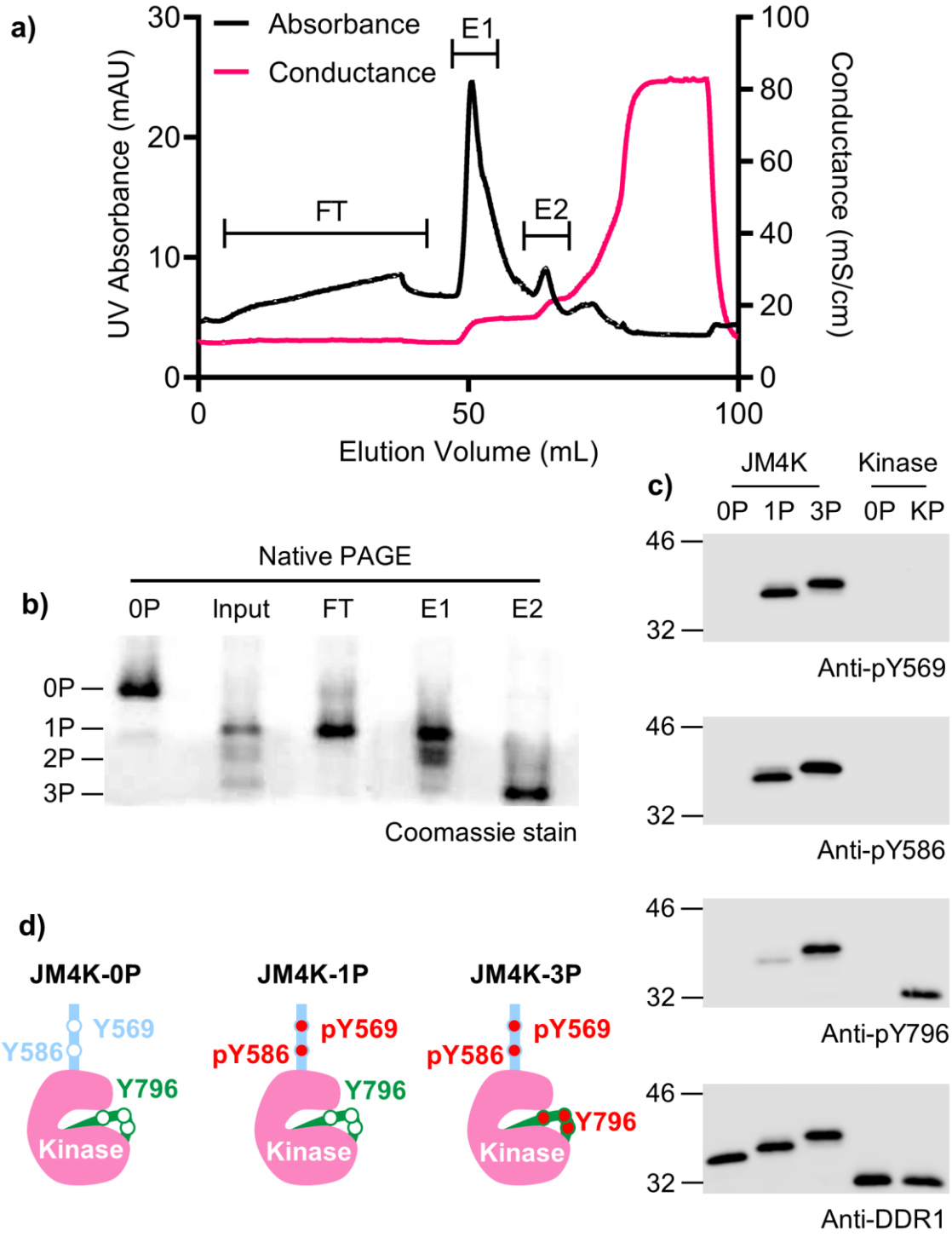

Figure 36. DDR1 phospho-forms can be separated by ion exchange.
a) The soluble DDR1 JM4K construct was incubated at 100 μM in the presence of 20 mM ATP for 90 minutes at 20°C. The reaction mixture was then dialysed into 80 mM NaCl, 50 mM Tris pH8.5 and separated by ion exchange chromatography (IEX). Protein was collected in the flow-through (FT) and in two peaks eluted with a NaCl gradient (E1 and E2). **b)** Unphosphorylated JM4K (0P), IEX input, and IEX fractions from **a)** were separated by native PAGE and stained with InstantBlue Coomassie. 0P – position of the unphosphorylated JM4K. 1P – position of the first intermediate JM4K phosphorylation form. 2P – position of the second intermediate JM4K

phosphorylation form. 3P – Fully phosphorylated JM4K. **c)** The phospho-forms of JM4K separated in **a)**, as well as the unphosphorylated (0P) and phosphorylated (KP) kinase alone, were boiled in sample buffer and analysed by SDS-PAGE and Western blotting with the JM4-specific anti-phosphotyrosine antibodies, anti-pY569 and anti-pY586, as well as the A-loop specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. **d)** Cartoon representation of the different phospho-forms separated in **a)**: JM4K-0P (unphosphorylated JM4K), JM4K-1P (intermediate JM4K phosphorylation form 1), JM4K-3P (fully phosphorylated JM4K).

Table 14. Tandem mass spectrometry analysis of JM4K phospho-forms.
1 µg of JM4K-1P (intermediate JM4K phosphorylation form 1) or JM4K-3P (fully phosphorylated JM4K) was boiled in sample buffer, separated by SDS-PAGE, stained with InstantBlue Coomassie stain, and excised from the gel. The protein was then cleaved with chymotyrpsin, or trypsin, and analysed by tandem mass spectrometry (performed by the CCP, University of Cambridge). All Tyr residues present within the JM4K structure are highlighted and those peptides which were detected by the mass spectrometer are depicted with a (+) or (-). (+); residue detected and found to be phosphorylated. (-); residue detected but found not to be phosphorylated. (ND); residue not detected. Data received from the CCP can be found in the Appendix (See **Chapter 7**; **Table 18**–**Table 21**).

Residue	JM4K-1P		JM4K-3P	
	Chymotrypsin	Trypsin	Chymotrypsin	Trypsin
Y569	+	ND	+	ND
Y586	ND	ND	+	ND
Y703	ND	ND	-	ND
Y740	+	ND	-	ND
Y756	-	-	-	-
Y792	ND	-	-	-
Y796	ND	-	ND	+
Y797	ND	-	ND	+
Y869	-	ND	-	ND
Y881	-	ND	-	ND

Differentially scanning fluorimetry analysis of the thermal stability of the different JM4K phosphorylation forms (JM4K-0P, -1P, -3P), showed that upon JM4 region phosphorylation (JM4K-1P) there was a marked decrease in the thermal denaturation midpoint compared with the unphosphorylated (JM4K-0P) protein (~3.5°C reduction - **Figure 37, Table 15**). This reduction in stability, whilst not equivalent to the removal of the JM4 region, as a whole, (~6°C drop between JM4K and kinase alone – **Table 11**), does demonstrate that JM4 region phosphorylation likely interrupts the stabilising interaction of the JM4 region with the kinase domain. Phosphorylation of the A-loop (JM4K-3P) then results in a further reduction in the thermal denaturation midpoint (~10°C reduction comparing JM4K-0P and -3P - **Table 15**). This also supports the model that A-loop phosphorylation relieves autoinhibitory interactions between the A-loop and the catalytic loop of the kinase domain.

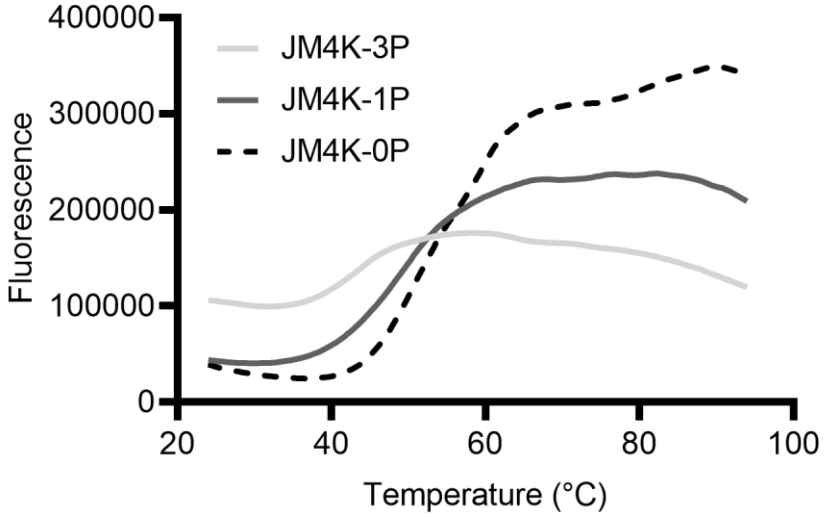

Figure 37. Thermal stability of soluble DDR1 phospho-forms.
The soluble DDR1 kinase JM4K phospho-forms JM4K-0P (unphosphorylated JM4K), JM4K-1P (intermediate JM4K phosphorylation form 1), JM4K-3P (fully phosphorylated JM4K), were analysed by differential scanning fluorimetry. Proteins were heated from 24 to 94°C in the presence of SYPRO orange, and fluorescence detected (excitation at 492 nm; emission at 610 nm). The average fluorescence from 2 independent repeats is shown. JM4K-3P – fully phosphorylated JM4K. JM4K-1P – intermediate JM4K phosphorylation form 1. JM4K-0P – unphosphorylated JM4K.

Table 15. Differential scanning fluorimetry analysis of DDR1 phospho-forms.
The soluble DDR1 kinase phospho-forms JM4K-0P (unphosphorylated JM4K), JM4K-1P (intermediate JM4K phosphorylation form 1), JM4K-3P (fully phosphorylated JM4K), were analysed by differential scanning fluorimetry. Proteins were heated from 24 to 94°C in the presence of SYPRO orange, and fluorescence detected (excitation at 492 nm; emission at 610 nm). Data were then analysed using Protein Thermal Shift software. The mean average denaturation midpoint (Tm) calculated from 2 independent experiments is shown with the range (±).

Protein	T_m (°C)
JM4K – 0P	53.2 (± 2.8)
JM4K -1P	49.7 (± 0.3)
JM4K – 3P	43.7 (± 0.5)

4.2.4 DDR1 autophosphorylation enhances catalysis

The successful isolation of soluble DDR1 phospho-forms enabled the determination of the catalytic activities of these proteins using the ADP-Glo™ assay. The different JM4K and kinase alone phospho-forms (JM4K-0P, JM4K-1P, JM4K-3P, K0P, and KP) were incubated with a range of ATP concentrations (0–800 μM) in the presence of a fixed Axltide peptide concentration (520 μM). ADP production was then recorded using the ADP-Glo™ assay after 0, 10, 20, and 40 minutes of assay initiation. The reaction was linear for all proteins across the time course (**Figure 38a**). Initial rates were determined at each ATP concentration and plotted as specific activity (nmol min^{-1} mg^{-1} - **Figure 38b**). Alternatively, the same experimental procedure as above was applied using a range of Axltide concentrations (0–520 μM) in the presence of a fixed ATP concentration (800 μM). Again, the reaction was linear across the time course (**Figure 39a**) and the initial rate of reaction was calculated and plotted as specific activity (**Figure 39b**). The fully phosphorylated JM4K-3P and Kinase-P phospho-forms displayed Michalis-Menten behaviour with respect to both substrates; ATP and Axltide. Apparent K_m values were calculated to be ~150 μM for both proteins, with little difference in k_{cat} (**Table 16**). The JM4K-1P form (JM4 region phosphorylated, A-loop unphosphorylated) showed substantially lower specific activity towards both ATP and Axltide substrates compared with the fully phosphorylated forms (JM4K-3P and KP). However, JM4K-1P had far higher specific activity than the unphosphorylated JM4K-0P. Furthermore, JM4K-1P had similar activity as the unphosphorylated kinase alone (K-0P). This shows that JM4 region phosphorylation relieves JM4 region autoinhibition, leading to an intermediate state of DDR1 kinase activation. Subsequent A-loop phosphorylation (JM4K-3P and KP) relieves this second autoinhibitory motif and leads to complete activation (k_{cat} increase from 1.1 to 6.2 between JM4K-1P and 3P for ATP).

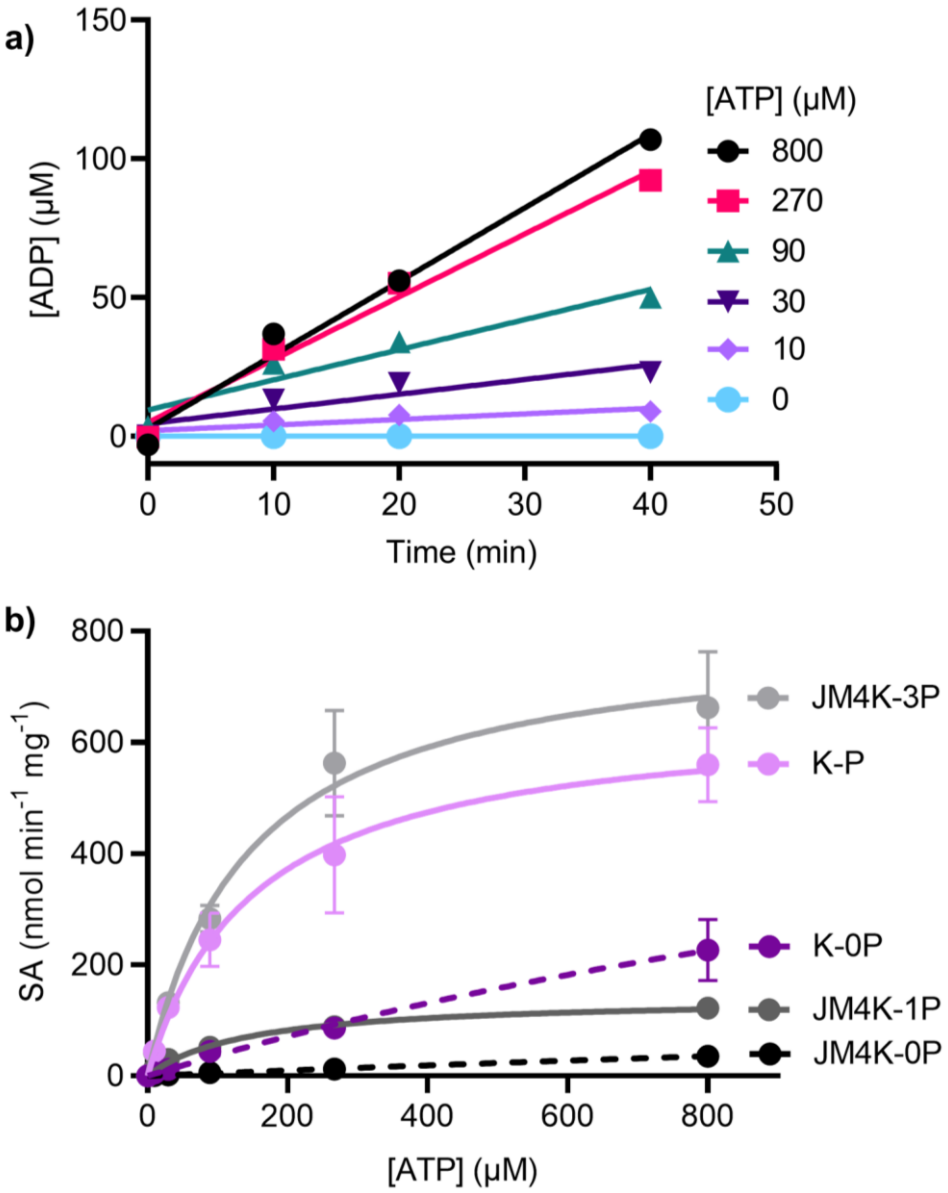

Figure 38. DDR1 autophosphorylation increases catalytic rate - ATP kinetics. *In* vitro kinase activity of the different phospho-forms of JM4K and kinase alone were assessed using the ADP-Glo™ assay. **a)** Initial velocities were determined over a range of ATP concentrations (10–800 μM) in the presence of a fixed Axltide peptide concentration (520 μM). Data for JM4K-3P are shown as an example. **b)** Data were then fitted with the Michaelis-Menten equation (n=3; mean and standard error of the mean shown, n=4 for JM4K-3P). SA, specific activity. JM4K-3P – fully phosphorylated JM4K. JM4K-1P –intermediate JM4K phosphorylation form 1. JM4K-0P – unphosphorylated JM4K. K-P – fully phosphorylated kinase alone. K-0P – unphosphorylated kinase alone.

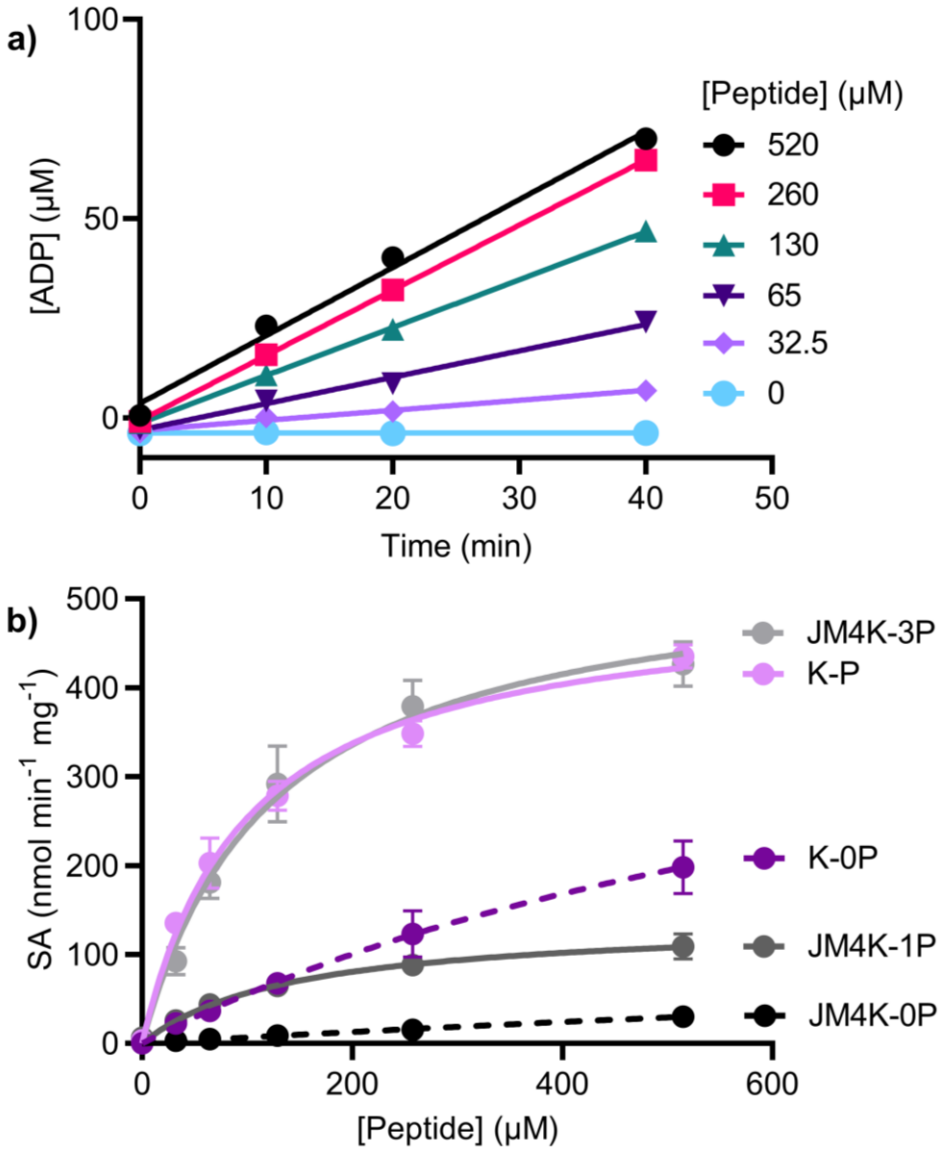

Figure 39. DDR1 autophosphorylation increases catalytic rate - Axltide kinetics. *In* vitro kinase activity of the different phospho-forms of JM4K and kinase alone were assessed using the ADP-Glo™ assay. **a)** Initial velocities were determined over a range of Axltide peptide concentrations (32.5–500 μM) in the presence of a fixed ATP concentration (800 μM). Data for JM4-3P are shown as an example. **b)** Data were then fitted with the Michaelis-Menten equation (n=3; mean and standard error of the mean shown, n=4 for JM4K-3P). SA, specific activity. JM4K-3P – fully phosphorylated JM4K. JM4K-1P –intermediate JM4K phosphorylation form 1. JM4K-0P – unphosphorylated JM4K. K-P – fully phosphorylated kinase alone. K-0P – unphosphorylated kinase alone.

Table 16. Apparent k_{cat} and K_m values for ATP and Axltide peptide.
Kinetic parameters for ATP and the Axltide peptide were calculated by fitting initial rate data (μM min^{-1}) with the Michaelis-Menten equation. The mean and 95% confidence intervals are shown for 3 independent experiments (*n*=4 for JM4K-3P). ND – indicates that the kinetic parameters were not determined due to a lack of saturation. JM4K-3P – fully phosphorylated JM4K. JM4K-1P – intermediate JM4K phosphorylation form 1. JM4K-0P – unphosphorylated JM4K. Kinase-P – fully phosphorylated kinase alone. Kinase-0P – unphosphorylated kinase alone.

Substrate	Kinetic parameter	JM4K 3P	JM4K 1P	JM4K 0P	Kinase P	Kinase 0P
ATP	k_{cat} (min^{-1})	6.2 (5–7.8)	1.1 (0.9-1.4)	ND	4.6 (3.6-6.3)	ND
	K_m (μM)	158 (84-312)	145 (79-271)	ND	151 (67-355)	ND
Peptide	k_{cat} (min^{-1})	4.2 (3.6-5)	1.1 (0.8-1.5)	ND	3.6 (3.2-4)	ND
	K_m (μM)	123.6 (80-192)	143.7 (69-309)	ND	98.54 (72-134)	ND

4.2.5 JM4 region autoinhibition is relieved by phospho-mimetic mutation

It has thus far been demonstrated that the JM4 region is autoinhibitory, and that JM4 region phosphorylation leads to a relief of this autoinhibition. Next, the effect of mimicking JM4 phosphorylation through mutating Tyr569 and Tyr586 of the JM4 region to Glu, a charged and therefore phospho-mimetic mutation, was assessed. This mutant construct (Y569E/Y586E-JM4K) along with the wild type JM4K, and Y569F/Y586F-JM4K, were incubated at 100 µM in the presence of 20 mM ATP across a time course (0–180 minutes). Protein samples were then analysed by native PAGE or Western blotting (**Figure 40**). Native PAGE analysis revealed that the Y569E/Y586E-JM4K mutant produced two phosphorylation forms compared with the four for the wild type JM4K construct, and one unphosphorylatable form of the Y569F/Y586F-JM4K construct (**Figure 40a**). This demonstrated that the intermediate phospho-forms seen for JM4K (JM4K-1P and -2P) are reliant on JM4 phosphorylation, as this step is by-passed in the mutant construct (**Figure 40a**). Western blot analysis of A-loop phosphorylation (anti-pY796 - **Figure 40b,c**) revealed that the wild type JM4K and phospho-mimetic (Y569E/Y586E-JM4K) had similar slow A-loop autophosphorylation kinetics. The mutant construct can therefore autophosphorylate and does so with similar kinetics as the wild type JM4K construct.

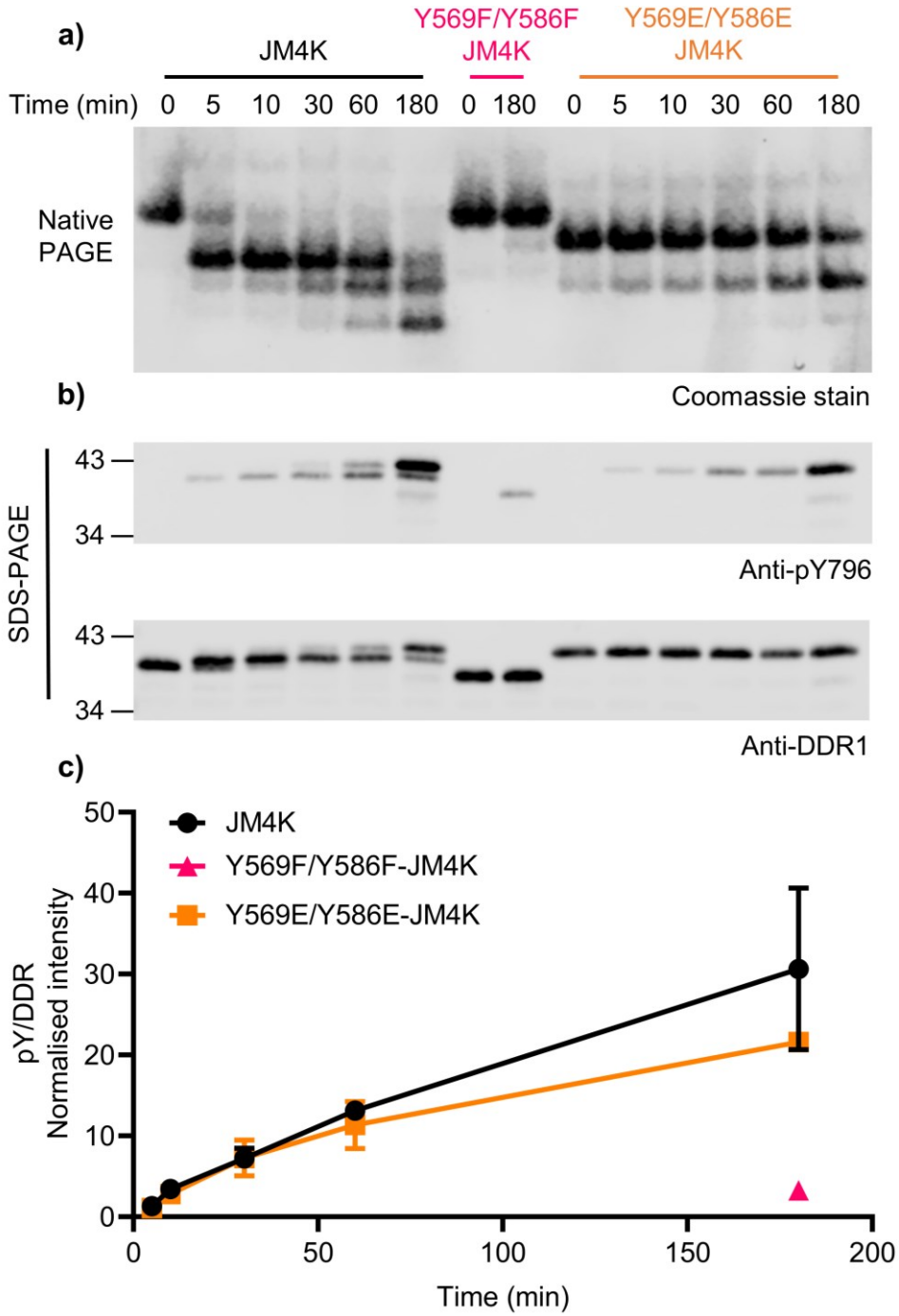

Figure 40. Autophosphorylation of the Y569E/Y586E-JM4K DDR1 construct. **a)** The soluble DDR1 JM4K, Y569F/Y586F-JM4K, and Y569E/Y586E-JM4K, constructs were incubated at 100 µM in the presence of kinase buffer II and 20 mM ATP for the indicated time course (0–180 minutes). Protein was then separated by native PAGE and gels stained with InstantBlue Coomassie stain or boiled in sample buffer and analysed by SDS-PAGE and Western blotting with the A-loop specific anti-phosphotyrosine antibody, anti-pY796. **b)**. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. **c)** Quantification of the A-loop specific, anti-pY796, signal normalised to respective anti-DDR1 signals. The anti-pY796 signal is expressed as a percentage of the sum of all bands on a blot, with mean and range shown (*n*=2).

Kinetic analysis of the Y569E/Y586E-JM4K-0P mutant using the ADP-Glo™ assay was then performed as in (**Figure 38, Figure 39**) using a range of ATP concentrations (0–800 μM) in the presence of a fixed Axltide peptide concentration (520 μM; **Figure 41a**), or a range of Axltide concentrations (0–520 μM) in the presence of a fixed ATP concentration (800 μM; **Figure 41b**). ADP production was then recorded using the ADP-Glo™ assay after 40, 20, 10, and 0 minutes of assay initiation. The reaction was linear across the full-time course. Initial rates were determined at each ATP, or peptide, concentration and plotted as specific activity (nmol min^{-1} mg^{-1}; **Figure 41**). These data revealed that the Y569E/Y586E-JM4K-0P construct had far greater activity than the unphosphorylated JM4K (JM4K-0P), and similar activity to the JM4 phosphorylated (JM4K-1P) and unphosphorylated kinase alone (K0P) constructs (**Figure 41**). This showed that autoinhibition by the JM4 region can be relieved by mimicking phosphorylation of JM4 Tyr residues, providing further support for the presented autoinhibitory model.

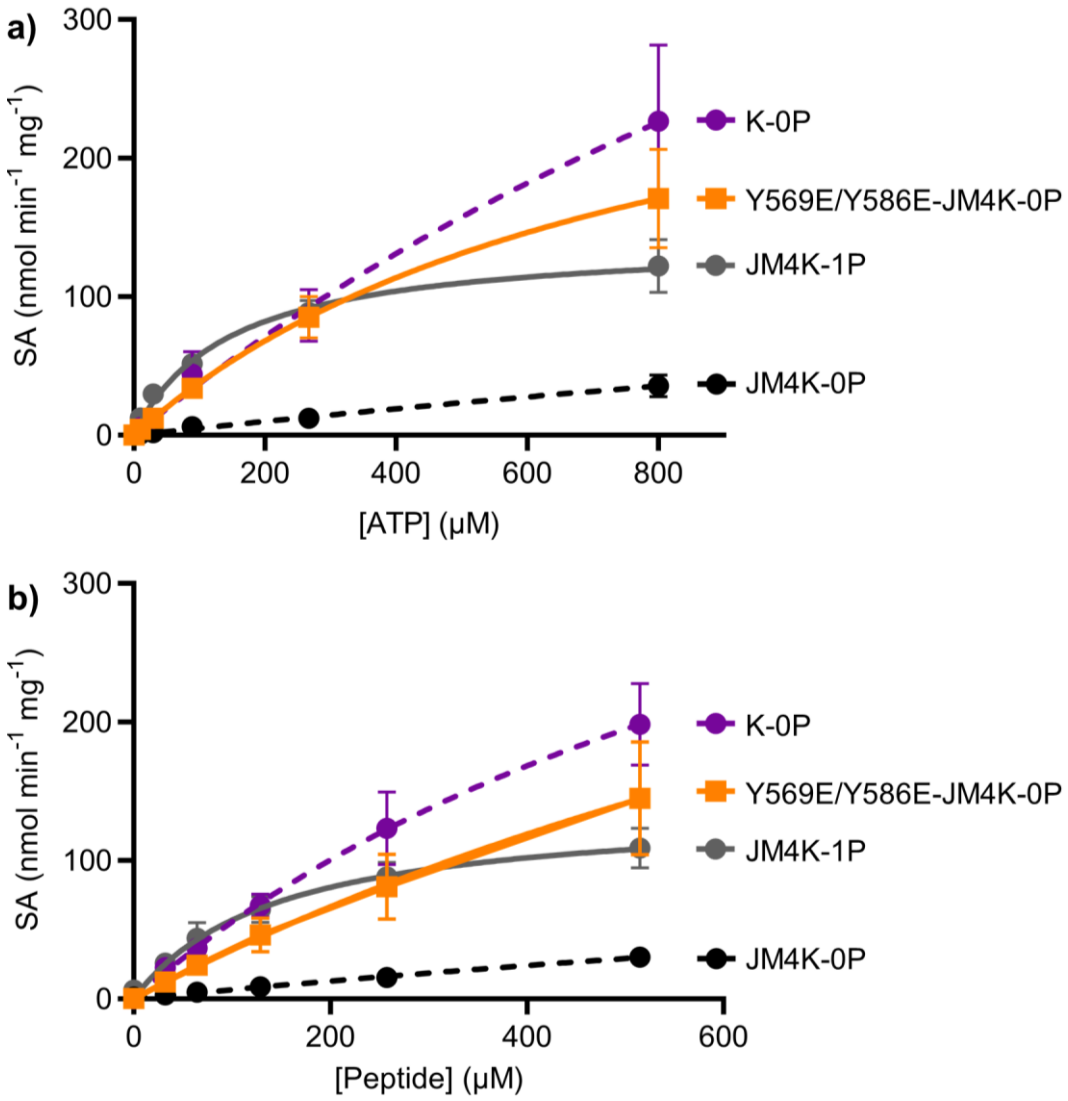

Figure 41. Y569E/Y586E-JM4K phosphorylates a peptide substrate with comparable kinetics to JM4 region phosphorylated DDR1 (JM4K-1P).
In vitro kinase activity of the phospho-mimetic Y569E/Y586E-JM4K-0P was assessed using the ADP-Glo™ assay. Initial velocities were determined over a range of ATP concentrations (10–800 μM) in the presence of a fixed Axltide peptide concentration (520 μM; **a)**, or at a range of Axltide peptide concentrations (32.5–520 μM) in the presence of a fixed ATP concentration (800 μM; **b)**. Data were then fitted with the Michaelis-Menten equation (*n*=3; mean and standard error of the mean shown). SA, specific activity. Y569E/Y586E-JM4K-0P – JM4 phosphomimetic JM4K mutant. JM4K-1P – intermediate JM4K phosphorylation form 1. JM4K-0P – unphosphorylated JM4K. K-0P – unphosphorylated kinase alone.

4.2.6 JM4 peptide mimetic has promising inhibitory potential

Most small molecule kinase inhibitors currently in use are ATP competitive (Wu et al, 2015b). This can present a problem in the generation of specific drugs, as the ATP binding site is extremely well conserved in kinases (Fasano et al, 2014; Hanson et al, 2019; Lemmon & Schlessinger, 2010). Indeed, DDR1 has been shown to bind several hundred kinase inhibitors (Hanson et al, 2019). The JM4 region occupies a space not currently utilised by DDR inhibitors. Targeting this binding pocket could therefore be beneficial in the generation of novel DDR1 inhibitors. To assess whether kinase autoinhibition could be achieved by occupying the JM4 region binding space, JM4 mimetic peptides were produced. The first peptide (peptide 1) encompassed residues Thr576-Val588 of the JM4 region with Tyr586 mutated to Phe to prevent this peptide acting as a substrate (**Figure 42a**). However, peptide 1 was not soluble in aqueous buffers making it difficult to perform inhibitor studies. Therefore, a longer peptide was designed which included residues Glu571-Val588, with Lys and Glu added to the N-terminus of the peptide to increase solubility (peptide 2 – **Figure 42a**). The incubation of peptide 2 at 1.2 mM with 0.5 μM kinase alone, 25 μM ATP, and 50 μM Axltide, resulted in a significant reduction in substrate phosphorylation by the kinase alone construct, as assessed using the ADP-Glo™ assay (**Figure 42b**). These data demonstrate the amenability of the JM4 region to the design of peptide inhibitors which may show enhanced DDR specificity over traditional ATP-competitive kinase inhibitors.

a)

```
JM4 Sequence   VPHYAEADIVTLQGVTGGNTYAVPAL
Peptide 1                TLQGVTGGNTFAV
Peptide 2          KEEADIVTLQGVTGGNTFAV
```

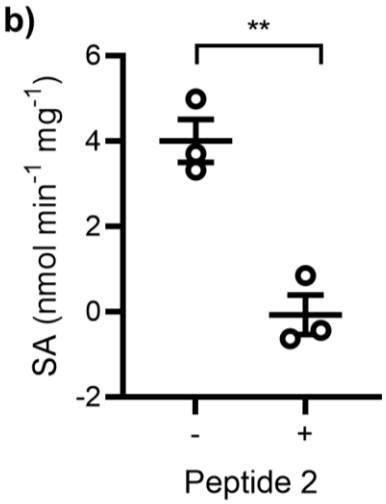

Figure 42. JM4 mimetic peptide inhibits DDR1 catalysis.
a) Synthesised peptides 1 and 2 aligned to the JM4 region sequence (DDR1 residues 566-591). Tyr586 was mutated to Phe in both peptides (red). Lys and Glu were added to the N-terminus of peptide 2 to increase peptide solubility. **b)** Peptide 2 was tested for its ability to inhibit the kinase alone construct using the ADP-Glo™ kinase assay. The kinase alone DDR1 construct was incubated in kinase buffer III at 0.5 μM in the presence of 50 μM Axltide peptide substrate, 25 μM ATP, with 1.2 mM peptide 2 (+), or 20 mM ammonium hydroxide (vector used to dissolve peptide 2) (-), for 30 minutes at room temperature. The ADP-Glo™ reagents were then added and luminescence recorded. Luminescent values were then converted to [ADP] (μM) produced by the kinase using an ADP standard curve. Specific activity (SA) is shown. ** - p=0.002. One-tailed unpaired t-test. *n*=3; mean and standard error of the mean shown.

4.2.7 Attempts to crystallise fully phosphorylated JM4K

The Y569F/Y586F-JM4K crystal structure solved in **Chapter 3** revealed important details about the inhibitory motifs present within the DDR1 structure. To better understand DDR1 activation, crystallisation trials were attempted for the fully phosphorylated form of JM4K (JM4K-3P). Protein was stimulated at 100 µM in the presence of 20 mM ATP for 8 hours. The sample was subsequently dialysed into a buffer containing 60 mM NaCl, 50 mM Tris pH 8.5, and separated by IEX chromatography (**Figure 43a**). Flow through (FT) from the column contained ATP/ADP with no protein (A_{260}/A_{280} values were >2). Increasing NaCl concentrations to ~120 mM, ~180 mM, and ~500 mM produced 3 elution peaks E1, E2, and E3, respectively. Native PAGE analysis of these samples showed that the first peak (E1) contained the intermediate JM4K phosphorylation form 2 (JM4K-2P) which had previously been unsuccessfully separated (**Figure 43b**). Peak 2 (E2) contained the fully phosphorylated (JM4K-3P) protein. The final peak (E3) contained JM4K-3P protein along with an unidentified contaminant. Phostag™ analysis of the E1 and E2 peaks confirmed that E2 contained pure JM4K-3P, whilst E1 contained an as yet uncharacterised intermediate phosphorylation form (JM4K-2P - **Figure 43c**). The purified JM4K-3P protein was gel filtered, concentrated to 9 mg/mL and incubated with 5 mM of a non-hydrolysable ATP analogue, AMP-PNP, and 10 mM Mn^{2+}, in the hope this would stabilise the protein conformation. Crystals of JM4K-3P formed over time; however, they were too small for diffraction data collection (**Figure 44 –** left image). Therefore, the crystals were used for crystal streak seeding. Streaking of microcrystals, generated from the original hit, produced larger crystals (~2x bigger - **Figure 44 –** right image). However, these were still too small for X-ray diffraction experiments.

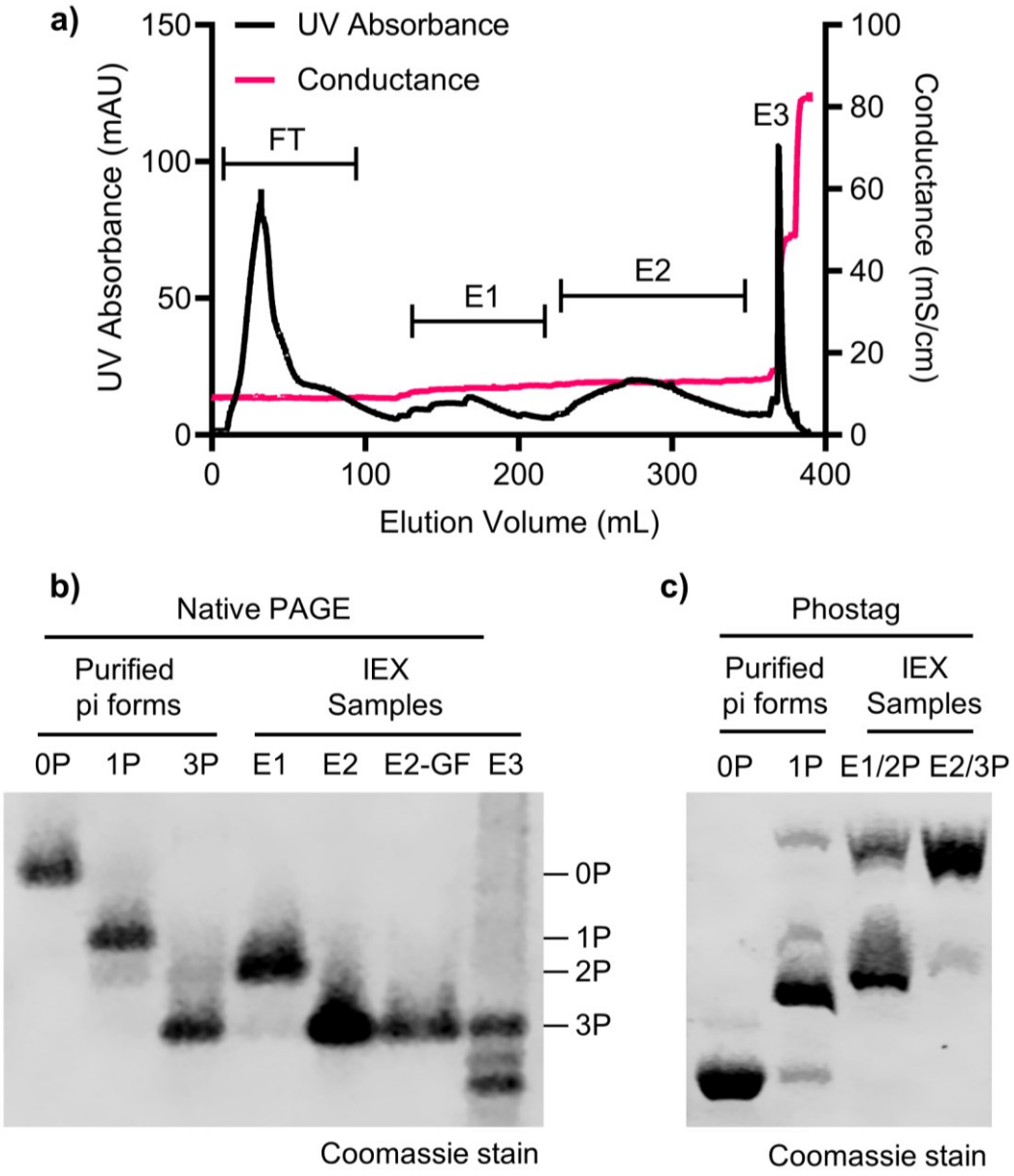

Figure 43. Purification of JM4K-3P for crystallography trials.
a) The soluble DDR1 JM4K construct was incubated at 100 µM in the presence of 20 mM ATP in kinase buffer II for 6 hours. The reaction mixture was then dialysed into 60 mM NaCl, 50 mM Tris pH 8.5, and separated by ion exchange chromatography (IEX). The flow through (FT) peak contained ATP/ADP. Protein was collected in three elution peaks E1, E2, and E3. **b)** DDR1 JM4K phospho-forms (pi forms) which had been previously purified (**Figure 36**) (0P (unphosphorylated JM4K), 1P (intermediate JM4K phosphorylation form 1), and 3P (fully phosphorylated JM4K)), along with the 3 elution peaks from **a)** (E1, E2, and E3) were separated by native PAGE (E2-GF – gel filtered IEX elution peak E2) and stained with InstantBlue Coomassie. The positions of the different phosphorylation forms of DDR1 JM4K are shown on the right (0P, 1P, 2P, and 3P). **c)** The unphosphorylated (0P), intermediate JM4K phosphorylation form 1 (1P), as well as the elution peaks E1 (intermediate JM4K phosphorylation form 2 - 2P) and E2 (fully phosphorylated JM4K - 3P) from **a)** were separated using Phostag™ SDS-PAGE before being stained with InstantBlue Coomassie.

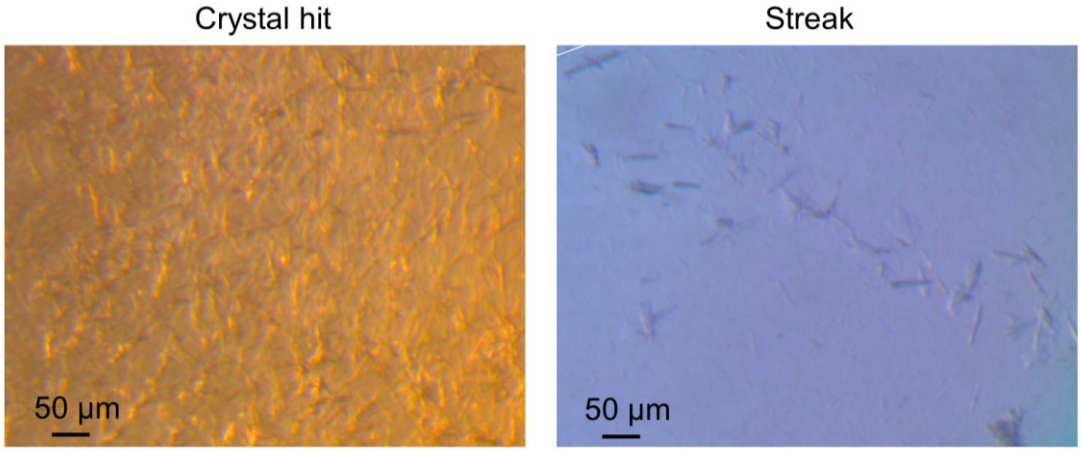

Figure 44. Attempts to crystallise JM4-3P.
JM4K-3P purified in **Figure 43** was gel filtered and then concentrated to 9 mg/mL and incubated with 5 mM AMP-PNP (non-hydrolysable ATP analogue) and 10 mM $MnCl_2$ for 30 minutes on ice. Small crystals grew in 0.02M MgCl2, 0.1M HEPES pH7.5, 22% (w/v) polyacrylic acid 5100 sodium salt (Crystal hit, left image). The crystal hit was crushed and used to streak a hanging drop containing protein and the same solution as the Crystal hit (Streak, right image). A scale-bar is shown for both images at 50 µm.

While this crystal screen was unsuccessful, it provided the opportunity to characterise the JM4K-2P phospho-form. Native and Phostag™ analysis revealed that JM4K-2P was a different phospho-form to both JM4K-3P and JM4K-1P (**Figure 43b,c**). Comparison of JM4K-2P with the other JM4K phosphoforms (JM4K-0P, -1P, and -3P) by Western blotting revealed that JM4K-2P was phosphorylated on the JM4 region (anti-pY569 and anti-pY586) to a similar extent as JM4K-1P and -3P (**Figure 45**). However, the JM4K-2P A-loop (anti-pY796) showed substantially less phosphorylated than JM4K-3P, and slightly more than JM4K-1P (**Figure 45**). Kinetic analysis of the JM4K-2P form using the ADP-Glo™ assay, as described for **Figure 38** and **Figure 39,** revealed that JM4K-2P had intermediate activity between the JM4K-1P and JM4K-3P forms: k_{cat} increased for ATP from 1.1 min^{-1} for JM4K-1P, to 3.7 min^{-1} for JM4K-2P, and then 6.2 min^{-1} for the JM4K-3P (**Figure 46, Table 17**). These data demonstrated that the 4 different phosphorylation forms of JM4K (JM4K-0P, -1P, -2P, and -3P) have differing activities.

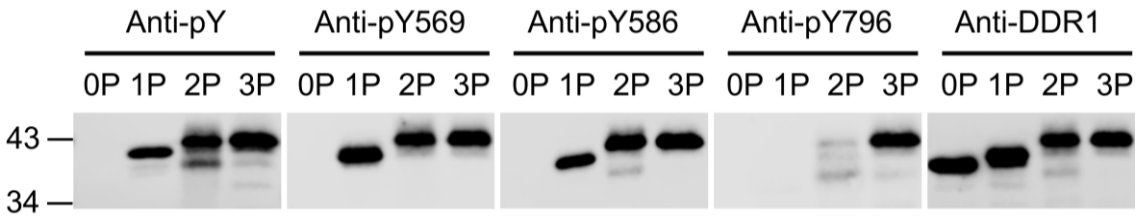

Figure 45. Phosphorylation status of JM4K-2P.
The isolated phospho-forms of JM4K (JM4K-0P, -1P, -2P, -3P) were boiled in sample buffer and analysed by SDS-PAGE and Western blotting with the JM4-specific anti-phosphotyrosine antibodies, anti-pY569 and anti-pY586, as well as the A-loop specific anti-phosphotyrosine antibody, anti-pY796, and the general anti-phosphotyrosine antibody 4G10 (anti-pY). Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left.

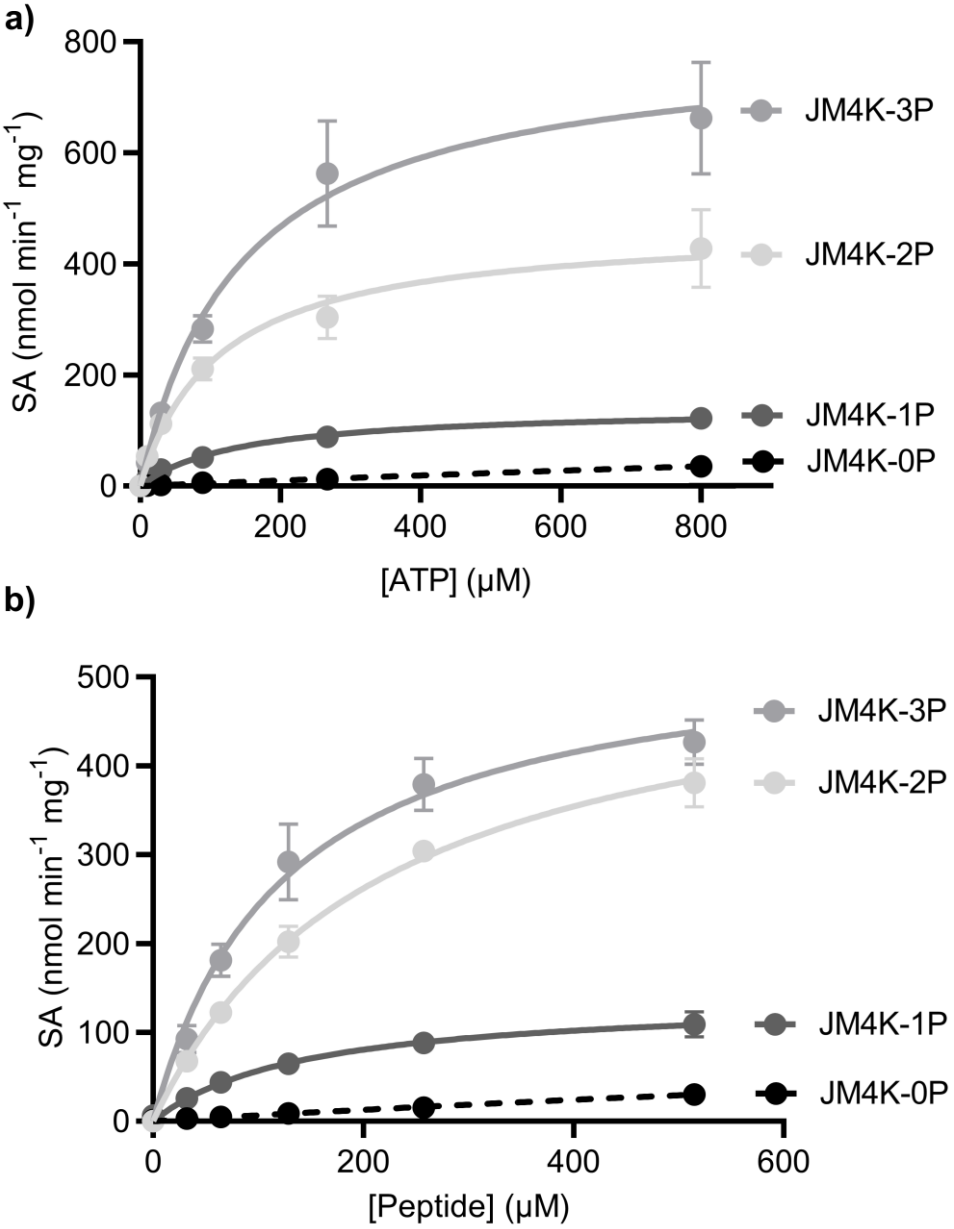

Figure 46. JM4K-2P has intermediate activity between JM4K-1P and JM4K-3P. *In* vitro kinase activity of the different phospho-forms of JM4K was assessed using the ADP-Glo™ assay. Initial velocities were determined over a range of ATP concentrations (10–800 μM) in the presence of a fixed Axltide peptide concentration (520 μM) **a)**, or a range of Axltide peptide concentrations (32.5–500 μM) in the presence of a fixed ATP concentration (800 μM) **b)**. Data were then fitted with the Michaelis-Menten equation (*n*=2; mean and range shown). SA, specific activity. JM4K-3P – fully phosphorylated JM4K. JM4K-2P – intermediate JM4K phosphorylation form 2. JM4K-1P –intermediate JM4K phosphorylation form 1. JM4K-0P – unphosphorylated JM4K. JM4-3P and -1P data are the same as those presented in **Figure 38, Figure 39**.

Table 17. Apparent k_{cat} and K_m values for ATP and Axltide peptide.
Kinetic parameters for ATP and the Axltide peptide were calculated by fitting initial rate data (μM min^{-1}) with the Michaelis-Menten equation. The mean and 95% confidence intervals are shown for 3 independent experiments (*n*=4 for JM4K-3P, *n*=2 for JM4K-2P). JM4K-3P – fully phosphorylated JM4K. JM4K-2P – intermediate JM4K phosphorylation form 2. JM4K-1P – JM4K intermediate phosphorylation form 1. JM4K-0P – unphosphorylated JM4K. JM4-3P and -1P data are the same as those presented in **Table 16.**

Substrate	Kinetic parameter	JM4K 3P	JM4K 2P	JM4K 1P
ATP	k_{cat} (min^{-1})	6.2 (5–7.8)	3.7 (3.6-6.3)	1.1 (0.9-1.4)
	K_m (μM)	158 (84-312)	151 (67-355)	145 (79-271)
Peptide	k_{cat} (min^{-1})	4.2 (3.6-5)	3.6 (3.2-4.0)	1.1 (0.8-1.5)
	K_m (μM)	124 (80-192)	99 (72-134)	144 (69-309)

4.3 Discussion

RTKs regulate several vital cellular processes including cell migration, proliferation, and differentiation (Lemmon & Schlessinger, 2010). It is unsurprising therefore that dysregulation of these receptors culminates in diseases such as cancer and fibrosis (Patterson et al, 2014; Yamaoka et al, 2018). As a result, RTKs have evolved several intricate autoinhibitory mechanisms which prevent unwanted signalling (Chiara et al, 2004; Hubbard, 2004; Lemmon & Schlessinger, 2010). In **Chapter 3**, structural analysis revealed that the DDR1 kinase domain is dually autoinhibited by the A-loop and JM4 region. In this chapter, biochemical analysis using recombinant protein constructs revealed that both these motifs have profound inhibitory effects on catalysis. Furthermore, this autoinhibition is relieved in an ordered mechanism, which begins with JM4 region phosphorylation, and culminates in A-loop phosphorylation.

Analysis of substrate phosphorylation using the ADP-Glo^{TM} assay revealed that soluble kinase constructs containing the JM4 region have significantly reduced kinase activity (**Figure 31**). This provided a functional assessment of JM4 region autoinhibition, supporting the structural data presented in **Chapter 3**. To relieve this autoinhibition the DDR1 kinase is shown to undergo a multi-step autophosphorylation mechanism (**Figure 33**). In the first step, the JM4 region is phosphorylated in *cis* (**Figure 28**), this results in a kinase of intermediate activity. Subsequently, the A-loop is phosphorylated culminating in complete kinase activation. Due to the high protein concentrations required for A-loop phosphorylation it likely occurs through in *trans* phosphorylation (**Figure 32**).The sequential autophosphorylation mechanism observed here also fits with structural data presented in **Chapter 3**. In the autoinhibited state, the JM4 region occludes the movement of the A-loop into a position which would allow *trans*-autophosphorylation of the A-loop to take place (Chen et al, 2020). It therefore follows that the JM4 region must move prior to A-loop phosphorylation. In support of this, the PDGFR family, which have a similar JM autoinhibitory mechanism, also phosphorylate the JM region prior to the A-loop (Mol et al, 2003).

RTKs are typically thought to relieve autoinhibition through in *trans* autophosphorylation (Lemmon & Schlessinger, 2010). Indeed, clustering of DDR1 facilitates in *trans* autophosphorylation (Juskaite et al, 2017). The in *cis* JM4 region

autophosphorylation demonstrated here is atypical; however, it does have precedence in RTK biology. *In vitro* analysis of the insulin receptor kinase, for example, showed that the JM region is phosphorylated in *cis* (Cann & Kohanski, 1997). Another atypical finding here was the exceptionally slow *trans*-phosphorylation of the DDR1 A-loop (**Figure 33**). A-loop phosphorylation is typically the first step in RTK activation, occurring within a few minutes of ATP or ligand stimulation (Li & Miller, 2006). The slow DDR1 activation shown here recapitulates what is seen in DDR1 expressing cells, where maximal DDR1 autophosphorylation can take several hours to be detected (Sammon et al, 2020; Shrivastava et al, 1997; Vogel et al, 1997). The A-loop of DDR1 is exceptionally stable in the inactive **D**FG-OUT conformation (Hanson et al, 2019). A crystal structure, which captured FGFR kinase in the act of *trans*-phosphorylation, demonstrated that the A-loop of the substrate kinase must adopt the active **D**FG-IN conformation (Chen et al, 2020). As transient transition into the active **D**FG-IN conformation is not thermodynamically favoured for DDR1 (Hanson et al, 2019), this may explain why A-loop phosphorylation occurs with such remarkably slow kinetics.

Collagen stimulated clustering of DDR1 at the cell surface likely facilitates activation by increasing the local concentration of the receptor (Corcoran et al, 2019). Consequently, the chance of a functional collision between two correctly oriented kinases is increased, and autoinhibition can be relieved through autophosphorylation. Here, this effect has been mimicked by increasing soluble kinase concentration *in vitro*. It would be interesting to monitor DDR1 autophosphorylation *in vitro* using artificial lipid vesicles. Sequestering DDR1 kinase constructs on vesicles would provide more biologically relevant insights into DDR1 activation, whilst still enabling tight experimental control over conditions.

The significant temporal separation between JM4 region and A-loop phosphorylation allowed isolation of the soluble DDR1 phospho-forms (**Figure 36)**. Kinetic analysis of these proteins revealed that phosphorylation (JM4K-1P) or removal (Kinase alone; K0P) of the JM4 region leads to a comparable increase in catalytic rate (**Figure 38, Figure 39**). Subsequent phosphorylation of the A-loop (JM4K-3P and KP) results in complete kinase activation and a further increase in catalytic rate (**Figure 38, Figure 39**). The was no difference in the kinetics of peptide substrate phosphorylation by

JM4K-3P or KP, indicating that the JM4 region does not affect activity of the fully phosphorylated kinase in solution. A positive role for the JM region has been described for the insulin receptor following relief of JM autoinhibition (Cabail et al, 2015). While the *in vitro* data presented here do not suggest a positive role for the JM4 region, it will be important to establish whether this is recapitulated in cells. A positive JM4 region function could be produced either through a structural interaction with adjacent DDR1 kinase domains, or through recruiting cellular potentiators of DDR1 signalling. A potential activator of DDR signalling is Src which binds phospho-Tyr569 and phospho-Tyr586 DDR1 peptides, and is essential for DDR2 activation (Iwai et al, 2013; Lemeer et al, 2012; Yang et al, 2005).

Initially only the JM4K-1P and JM4K-3P forms were successfully isolated. However, a second intermediate phospho-form, JM4K-2P, was subsequently isolated (**Figure 43**). Native gel and PhosTag^{TM} analysis revealed that JM4K-2P was a distinct phospho-form (**Figure 43**). However, the only discernible difference which could be detected was a subtle increase in phosphorylation of the A-loop (Tyr796) in JM4K-2P relative to JM4K-1P when analysed by Western blotting (**Figure 45**). Despite this, the JM4K-2P construct had significantly higher activity towards a peptide substrate relative to JM4K-1P, and somewhat lower activity relative to JM4K-3P (**Figure 46**). It is unlikely that the subtle differences in Tyr796 phosphorylation observed in the JM4K-2P construct would have accounted for this large increase in activity. Instead, it is possible that another Tyr residue within the kinase domain is being phosphorylated for which we do not currently have a phospho-specific antibody. For example, in addition to Tyr796, the DDR1 A-loop contains two other potentially phosphorylatable Tyr residues; Tyr792 and Tyr797. Mass spectrometry analysis of the JM4K-3P construct revealed that both Tyr796 and Tyr797 were phosphorylated (**Table 14**). It will be important in future to repeat this assessment using the JM4K-2P, to establish whether Tyr797 but not Tyr796 is phosphorylated in this construct. The insulin receptor also contains 3 Tyr residues within the A-loop (Hubbard et al, 1994). These Tyr residues are phosphorylated in a highly ordered sequence which begins with Tyr1162 (Tyr796 in DDR1), followed by Tyr1158 (Tyr792 in DDR1), and finally Tyr1163 (Tyr797 in DDR1) (Favelyukis et al, 2001). Each step of A-loop phosphorylation increases substrate affinity or catalytic rate. A similar sequential

mechanism may take place in DDR1, which could account for the different catalytic rates of the JM4K-2P phospho-form.

Mutational disruption of JM-kinase interactions can lead to constitutive activation of RTKs, such as the KIT-V560G JM region mutation found in gastrointestinal stromal tumours (Heinrich et al, 2003). Similarly, disruption of the DDR1 JM4-kinase interaction (Y569E/Y586E-JM4K) resulted in an intermediate level of kinase activity (**Figure 41**). This mutation was qualitatively comparable to deletion (Kinase-0P) or phosphorylation of the JM4 region (JM4K-1P). Incubation of Y569E/Y586E-JM4K with ATP stimulated A-loop phosphorylation with the same slow kinetics as the wild type JM4K protein (**Figure 40**). Conversely, the Y569F/Y586F mutation resulted in an almost complete loss of kinase function (**Figure 33**). This demonstrates that if JM4 region autoinhibition cannot be relieved through autophosphorylation, kinase catalysis will remain limited. Phe mutation replaces the Tyr residue with a similar, yet unphosphorylatable, side chain, whilst the Glu mutation mimics phosphorylation through the introduction of a negative charge. Both mutations are common in RTK experimental analysis (Chen & Cole, 2015; Ouwens et al, 1994). Nevertheless, the mutant residues clearly have distinct physiochemical properties which will affect the results. Taken together however, these mutations support a model whereby autoinhibitory interactions between the JM4 region and kinase are relieved through autophosphorylation of JM4 Tyr residues.

As discussed in **Chapter 3**, the unique nature of the JM4-kinase interactions offers an attractive target for pharmacological intervention. Mimicking the JM4 region binding space with inhibitors could produce molecules with enhanced specificity compared with traditional ATP-competitive inhibitors. Here, a JM4 mimetic peptide is shown to inhibit DDR1 substrate phosphorylation (**Figure 42**). This demonstrates the amenability of the DDR1 kinase domain to JM4 mimetic inhibitor studies. An interesting avenue for future research would be to investigate the minimal JM4 peptide sequence which can elicit an inhibitory effect. The specificity of this minimal peptide for DDR1 over other kinases could then be assessed through screening a panel of kinases. If specific, this peptide could be used to identify novel small molecules which can compete for kinase binding. This could be achieved by labelling the JM4 peptide mimetic with a fluorescent probe. A fluorescence polarisation assay

could then be developed which monitored peptide displacement from the kinase through competitive binding with a panel of small molecules (Hall et al, 2016). Identification of an appropriate small molecule panel could also be assessed through computational predictions of compounds which will bind the JM4 region autoinhibitory space (Sliwoski et al, 2014).

In summary, the JM4 region and A-loop are potent inhibitors of catalysis. Relief of this autoinhibition occurs in an ordered sequence of phosphorylation events beginning with JM4 phosphorylation. The confirmation that JM4 mimetic peptides can inhibit catalysis also provides promising avenues for the development of DDR specific inhibitors. However, the function of the JM4 region within a cellular context must first be assessed.

Chapter 5. JM4 region function in full-length DDR1

5.1 Introduction

Dual-autoinhibition by the A-loop and JM4 region has been demonstrated using soluble DDR1 kinase constructs (**Chapter 3, Chapter 4**). However, JM4 function in the context of the full-length, cell expressed receptor has not been assessed. DDR1 activation within a complex cellular environment is likely to be different to that seen using soluble kinase constructs *in vitro*. DDRs are activated through clustering at the cell surface; a phenomenon which is lost in a reductionist *in vitro* experiment (Corcoran et al, 2019; Juskaite et al, 2017; Yeung et al, 2019). It is currently unclear whether the JM4 segment, or any other part of the JM region, plays a role in this event. Furthermore, it is not known whether the JM4 region plays an additional positive role following phosphorylation, such as that seen with the insulin receptor JM (Cabail et al, 2015). In addition, a universal theme in RTK biology is the ability of the JM region to recruit cellular potentiators and repressors of signalling (Lemmon & Schlessinger, 2010). For example, the MuSK JM region produces its inhibitory effect by recruiting a cellular repressor of MuSK signalling, which inhibits the receptor in *trans* (Herbst & Burden, 2000; Till et al, 2002). The DDR1 JM region has been shown to bind phosphatases, such as SHP-2, as well as kinases, such as Src (Lemeer et al, 2012; Yang et al, 2009). Indeed, Src is important for DDR2 activation, and in DDR1 binds to phospho-peptide derivatives of phospho-Tyr569 and phospho-Tyr586 of the JM4 (Ikeda et al, 2002; Lemeer et al, 2012; Yang et al, 2005; Zhang et al, 2013). The importance of these interactions on DDR1 function has also yet to be assessed.

5.1.1 Hypothesis and aims

It is hypothesised that autoinhibition of full-length, cell expressed DDR1 is maintained by the JM4 region and A-loop. Release from this autoinhibition occurs through a two-step process beginning with JM4 phosphorylation. Furthermore, Src is predicted to be an activator of DDR1 signalling.

The aims were: 1) to assess the order of DDR1 autophosphorylation events; 2) to functionally analyse JM4 mutation and deletion within a cellular context; 3) to evaluate the role played by Src in DDR1 activation.

5.1.2 Findings

Mutation of the JM4 region results in a loss of kinase function in the full-length receptor. Furthermore, the JM4 region is shown to be autophosphorylated before the A-loop, in support of earlier *in vitro* findings (**Chapter 3**, **Chapter 4**). The finding that JM4 phospho-mimetic mutation counterintuitively leads to a loss of receptor function indicates a secondary positive function for the JM4 region in cells. This role may be the recruitment of Src, which is shown to potentiate DDR1 phosphorylation in cells; an effect which is lost upon JM4 Tyr mutation (Y569F/Y586F or Y569E/Y586E).

5.2 Results

5.2.1 JM4 is phosphorylated before the A-loop in full-length DDR1

So far, the role of the JM4 region has been explored using purified soluble DDR1 kinase constructs *in vitro*. It was therefore important to analyse the role of this region in the context of the full-length receptor. In **Chapter 4**, it was shown that the JM4 region is phosphorylated before the A-loop (**Figure 33**). To determine whether this was the case in the full-length receptor, DDR1-WT was expressed in HEK293 cells. Cells were subsequently stimulated with collagen I over a time course (0–60 minutes) at 37°C. They were then lysed and analysed by Western blotting with the JM4-specific anti-phosphotyrosine antibodies, anti-pY569 and anti-pY586, as well as the A-loop-specific anti-phosphotyrosine antibody, anti-pY796 (**Figure 47a**). No difference in the kinetics of JM4 or A-loop phosphorylation were observed (**Figure 47b**). In cells, the rate-limiting steps in receptor activation is reorganisation of DDR1 at the cell surface into large clusters (Corcoran et al, 2019). In order to remove this rate-limiting step, full-length DDR1-WT was immunoprecipitated from cell lysates using anti-DDR1 and Protein A Sepharose beads. The immunoprecipitated protein was then stimulated with ATP across a time course (0–60 minutes). Samples were boiled in sample buffer at each time point and analysed by Western blotting as above (**Figure 48a**). The JM4-specific anti-phosphotyrosine antibodies, anti-pY569 and anti-pY586, showed phospho-signals after 5 minutes of stimulation with levels plateauing after 30 minutes (**Figure 48b**). In contrast, the A-loop phospho-signal (anti-pY796) was linear across the time course indicating that complete A-loop phosphorylation was not achieved by 60 minutes (**Figure 48b**). This showed that, in full-length DDR1, the JM4 region is phosphorylated before the A-loop, in agreement with earlier *in vitro* biochemical analysis of soluble DDR1 constructs.

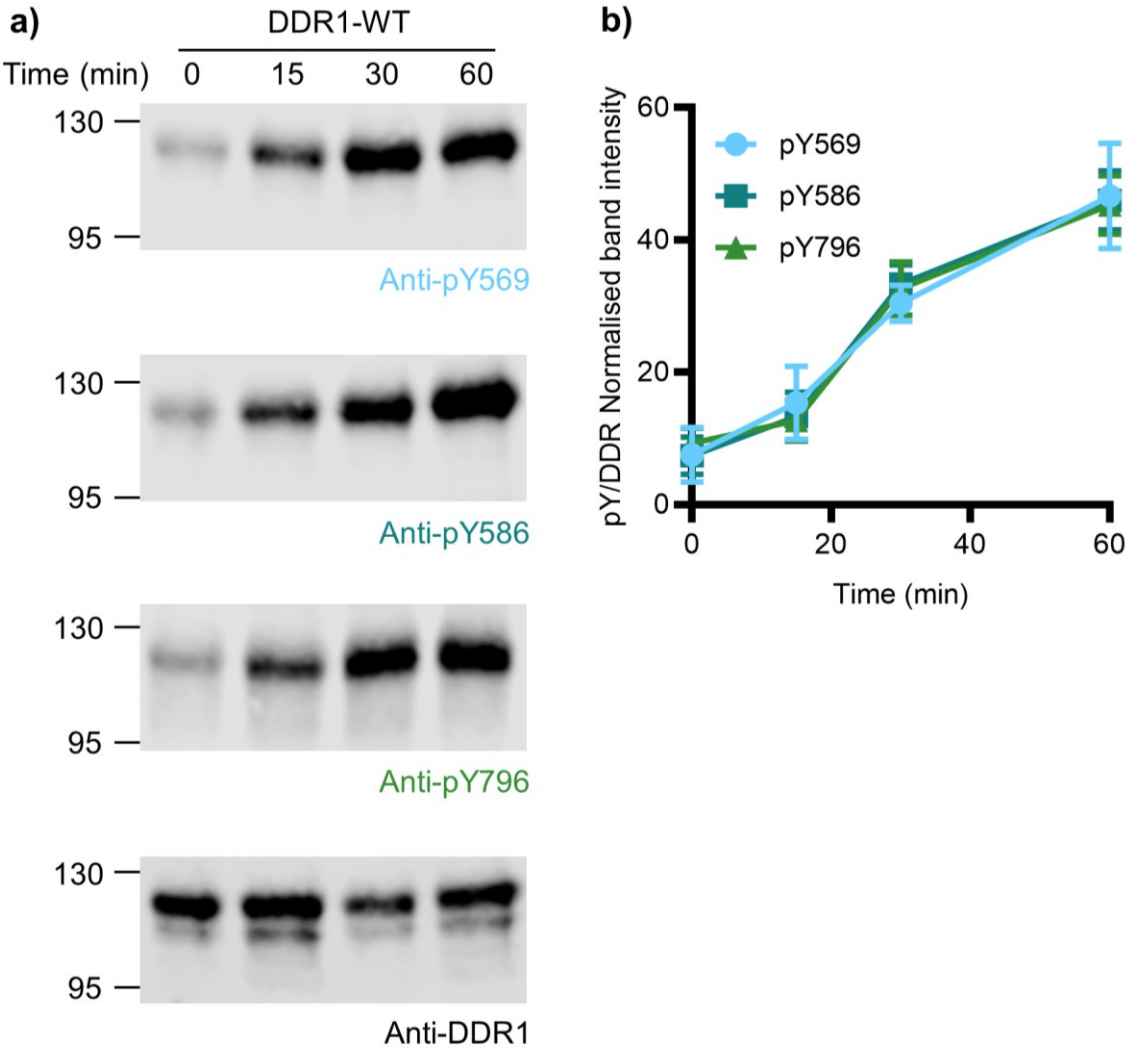

Figure 47. No kinetic difference is seen between JM4 and A-loop phosphorylation upon collagen I stimulation.
a) Full-length DDR1-wild type (WT) was transiently expressed in HEK293 cells. Cells were then stimulated with collagen I for 0–60 minutes at 37°C. Cells were subsequently lysed, and aliquots of cell lysates analysed by SDS-PAGE and Western blotting with the JM4-specific anti-phosphotyrosine antibodies, anti-pY569 and anti-pY586, or the A-loop specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. **b)** Quantification of anti-phosphotyrosine (anti-pY569, anti-pY586, and anti-pY796) signal normalised to respective anti-DDR1 signals. The anti-phosphotyrosine signal is expressed as a percentage of the sum of all bands on a blot, with mean and standard error of the mean shown (*n*=3).

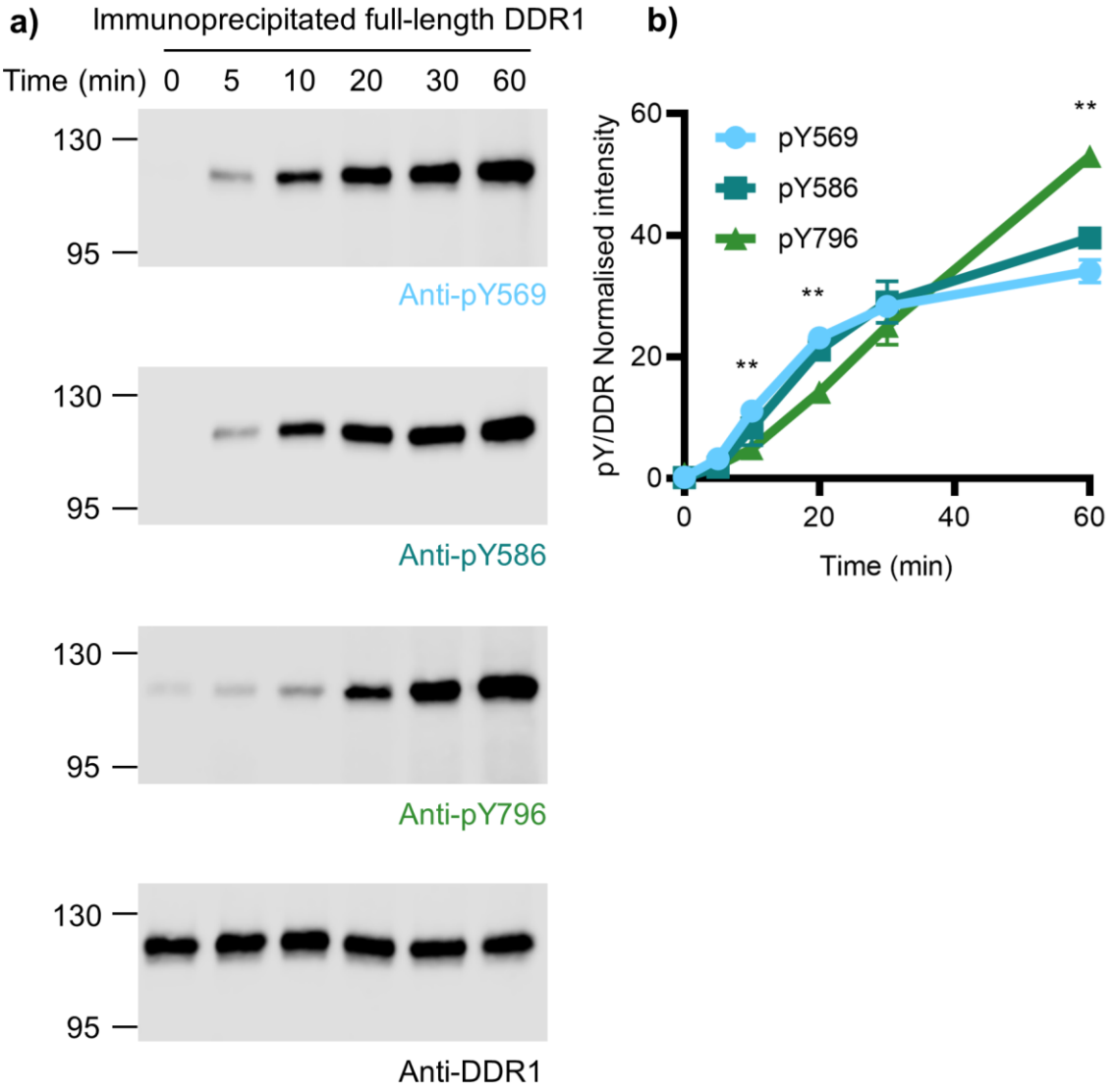

Figure 48. The JM4 region is phosphorylated first in full-length DDR1 *in vitro*. **a)** Full-length DDR1-wild type (WT) was transiently expressed in HEK293 cells. Cells were then lysed and DDR1 immunoprecipitated from cell lysates. Immunoprecipitates were then stimulated with 1 mM ATP in IP kinase buffer for the indicated time course (0–60 minutes) at 30°C. Samples were then boiled in sample buffer and analysed by SDS-PAGE and Western Blotting with the JM4 specific anti-phosphotyrosine antibodies, anti-pY569 and anti-pY586, or the A-loop-specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. **b)** Quantification of anti-pY569, anti-pY586, and anti-pY796 signals normalised to respective anti-DDR1 signals. The anti-phosphotyrosine signal is expressed as a percentage of the sum of all bands on a blot, with mean and standard error of the mean shown (*n*=3). Statistically significant differences between the anti-pY569 and anti-pY796 signals are indicated after 10, 20 and 60 minutes (p=0.0072, p=0.0032, and p=0.0034, respectively). Two-way ANOVA with Tukey's post-test was performed.

5.2.2 JM4 mutation or deletion in full-length DDR1 abrogates function

Following the confirmation that the JM4 region is phosphorylated before the A-loop, the effect of JM4 mutation on receptor function within cells was next assessed. Several full-length DDR1 JM4 mutants were produced: DDR1-Y569F, DDR1-Y586F, DDR1-Y569F/Y586F, DDR1-Y569E, DDR1-Y586E, DDR1-Y569E/Y586E. These constructs were expressed in HEK293 cells which were subsequently stimulated with collagen I for 90 minutes at 37°C. Cells were then lysed and analysed by Western blotting with the A-loop specific anti-phosphotyrosine antibody, anti-pY796. As demonstrated previously in **Figure 10**, mutation of the JM4 Tyr residues, Y569 and Y586, to Phe resulted in a significant reduction in receptor autophosphorylation (**Figure 49**). This supported *in vitro* data which showed that Y569F/Y586F mutation resulted in a loss of catalytic function (**Figure 33**). In the soluble DDR1 constructs, the Y569E/Y586E-JM4K phospho-mimetic mutant was proficient in autophosphorylation and had increased activity towards a peptide substrate when compared with the wild type JM4K (**Figure 40**, **Figure 41**). However, in the full-length receptor, this mutation resulted in a similar reduction in collagen responsiveness as the DDR1-Y569F/Y586F mutation (**Figure 49**). Furthermore, the single point mutations, DDR1-Y569E and DDR1-Y586E, were also comparable to their Tyr-to-Phe counterparts, demonstrating reduced collagen responsiveness compared with DDR1-WT. This unexpected loss of activity in the phospho-mimetic mutant (DDR1-Y569E/Y586E) indicates that there may be additional complexity to the function of the JM4 region within cells.

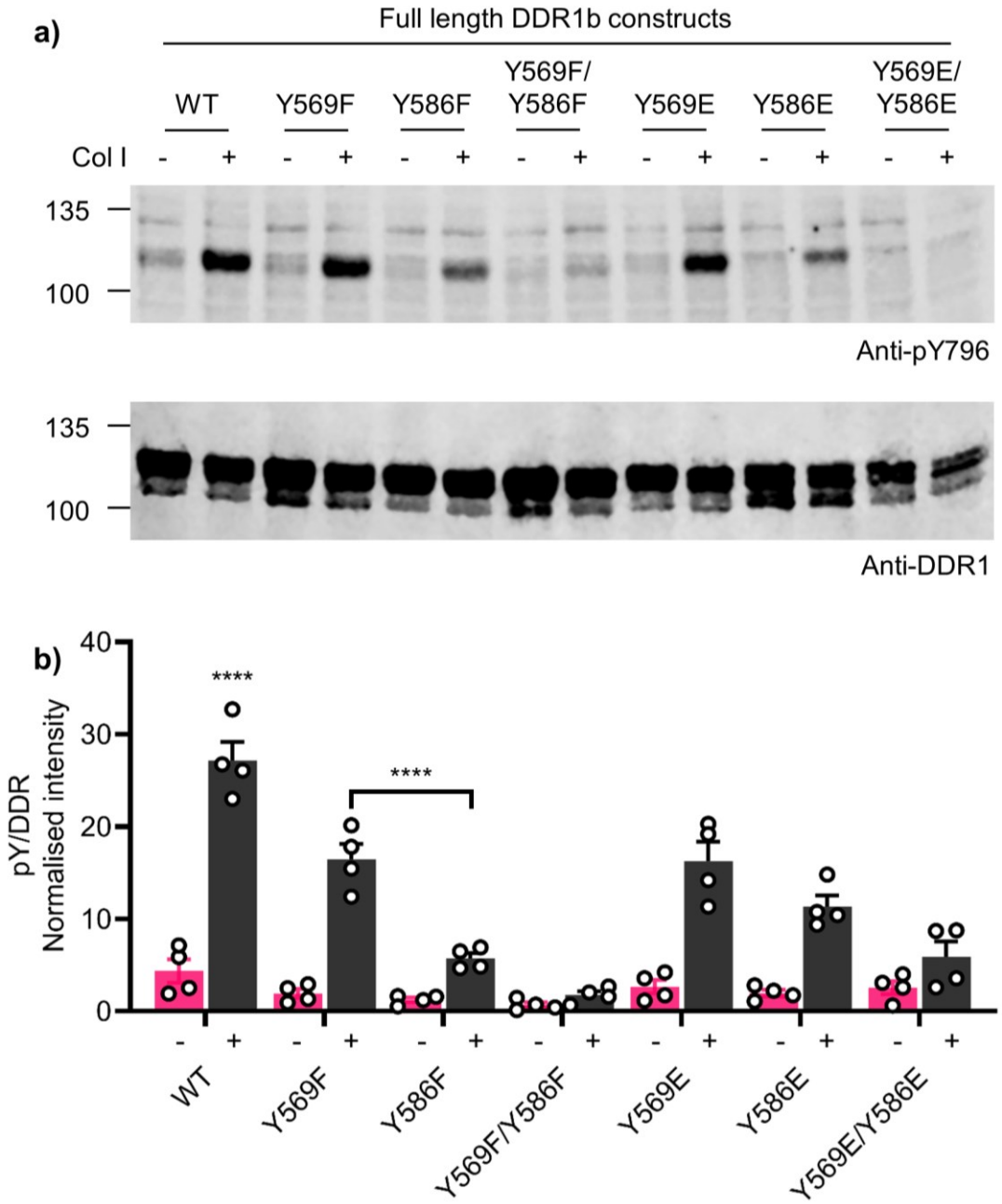

Figure 49. JM4 Tyr mutation in full-length DDR1 abrogates receptor function. **a)** Full-length DDR1-wild type (WT), or the indicated DDR1 mutants, were transiently expressed in HEK293 cells. Cells were then stimulated with collagen I (+) or left unstimulated (-) for 90 minutes at 37°C. Cells were subsequently lysed, and aliquots of cell lysates analysed by SDS-PAGE and Western blotting with the A-loop-specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. **b)** Quantification of anti-pY796 signal normalised to respective anti-DDR1 signals. The anti-pY796 signal is expressed as a percentage of the sum of all bands on a blot, with mean and standard error of the mean shown (*n*=4). **** - $p<0.0001$ statistically significant difference was found for WT (+) compared with all other collagen stimulated conditions. $p<0.0001$ statistically significant difference was also found between the Y569F (+) Col I and Y586F (+) Col I conditions. Two-way ANOVA with a Tukey post-test was performed.

Like the phospho-mimetic mutation (DDR1-Y569E/Y586E), deletion of the JM4 region was predicted to produce a receptor with increased catalytic activity. To test this, different sections of the DDR1 JM region were deleted from the full-length receptor. DDR1-ΔJM1 had Arg445-Arg475 of the JM1 region deleted, DDR1-ΔJM2 had Pro476-Ser504 of the JM2 region deleted, DDR1-ΔJM3 had Ala542-Ser565 of the JM3 region deleted, DDR1-ΔJM1-3 had Arg445-Ser565 of the JM1, 2 and 3 regions deleted, whilst DDR1-ΔJM4 had Val566-Leu591 of JM4 deleted (mutagenesis performed by Jasmine Gratton). These constructs were expressed in HEK293 cells which were subsequently stimulated with collagen I for 90 minutes at 37°C. Cells were then lysed and analysed by Western blotting with the A-loop specific anti-phosphotyrosine antibody, anti-pY796. Deletion of the JM1-3 regions, either individually (ΔJM1, ΔJM2, ΔJM3), or as a whole (ΔJM1-3), had little effect on collagen-induced receptor autophosphorylation compared with DDR1-WT (**Figure 50**). However, unexpectedly, JM4 deletion (ΔJM4) resulted in a complete loss of receptor autophosphorylation. This again demonstrated that there may be a secondary function for the JM4 region in cells which is crucial for DDR1 activation.

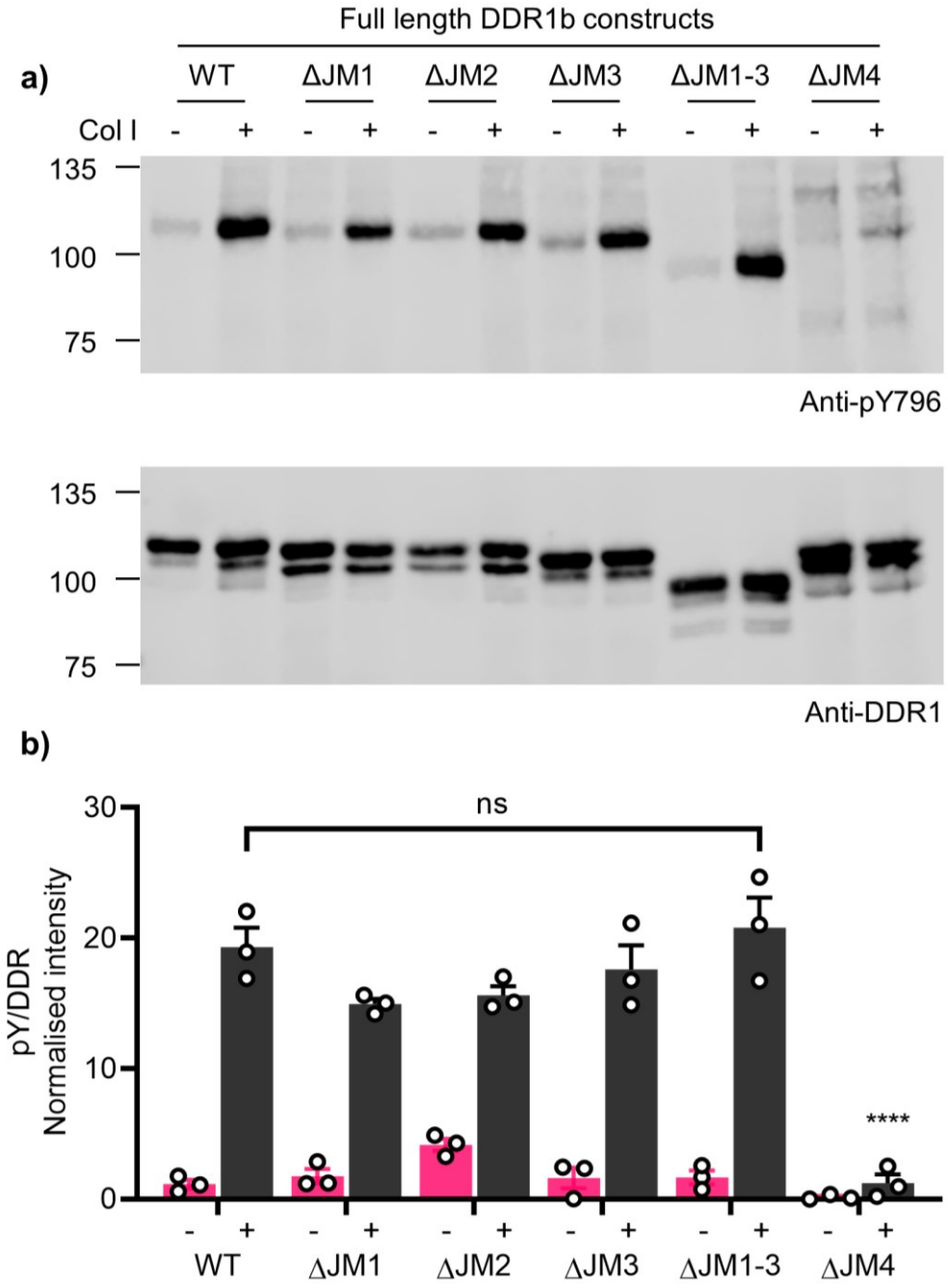

Figure 50. Deletion of JM4 in full-length DDR1 abrogates receptor function. Full-length DDR1-wild type (WT), or the indicated DDR1-JM deletion constructs were transiently expressed in HEK293 cells. Cells were then stimulated with collagen I (+) or left unstimulated (-) for 90 minutes at 37°C. Cells were subsequently lysed, and aliquots of cell lysates analysed by SDS-PAGE and Western blotting with the A-loop-specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. Two-way ANOVA with a Tukey post-test was performed. **** - $p<0.0001$ statistically significant difference found between WT (+) collagen I and ΔJM4 (+) collagen I. No statistically significant difference was found between the WT (+) collagen I condition and the other collagen I stimulated JM deletion conditions (ΔJM1 p=0.2048, ΔJM2 p= 0.4136, ΔJM3 p= 0.9896, ΔJM1-3 p= 0.9963).

5.2.3 DDR1 JM4 mutant constructs are present at the cell surface

Following the surprising discovery that phospho-mimetic mutation or deletion of the autoinhibitory JM4 region (DDR1-Y569E/Y586E and DDR1-ΔJM4; **Figure 49, Figure 50**) led to a loss of collagen-induced receptor autophosphorylation, it was next important to determine whether these mutant constructs were correctly transported to the cell surface and were therefore exposed to collagen. The DDR1-WT, DDR1-Y569F/Y586F, DDR1-Y569E/Y586E, and DDR1-ΔJM4 constructs were expressed in HEK293 cells which were subsequently stimulated with collagen I for 90 minutes at 37°C. Cells were then lysed and analysed by Western blotting with the A-loop specific anti-phosphotyrosine antibody, anti-pY796. This again revealed that there was no autophosphorylation on the A-loop of any of the mutant constructs (anti-pY796 – **Figure 51a**). Lysates from these stimulation experiments were then incubated with EndoH, and analysed by SDS-PAGE, and Western blotting with anti-DDR1. Protein which has passed through the secretory pathway, and is therefore likely cell surface localised, should be resistant to EndoH digestion due to alterations of the oligosaccharide composition of the protein during transit through the Golgi (Koide & Muramatsu, 1974; Taner et al, 2011). DDR1-WT, as expected, was mostly resistant to EndoH digestion, with only a small fraction of cleaved protein running lower on the gel following EndoH digestion (**Figure 51b;** WT +). This EndoH digested product likely represents the biosynthetic precursor of DDR1. The same banding pattern was also found for the DDR1-Y569F/Y586F construct. However, the DDR1-Y569E/Y586E and DDR1-ΔJM4 constructs had a higher proportion susceptible to EndoH cleavage, as shown by the larger lower molecular weight band seen on the SDS-PAGE gel (**Figure 51b**). There was, however, protein which was resistant to this cleavage and remained at the same molecular weight as the un-cleaved sample. This indicated that the DDR1-Y569F/Y586F protein is likely processed in the same way as the wild type receptor. However, a higher proportion of the DDR1-Y569E/Y586E and DDR1-ΔJM4 constructs was found as biosynthetic precursors, although a sizeable proportion does process successfully through the secretory pathway.

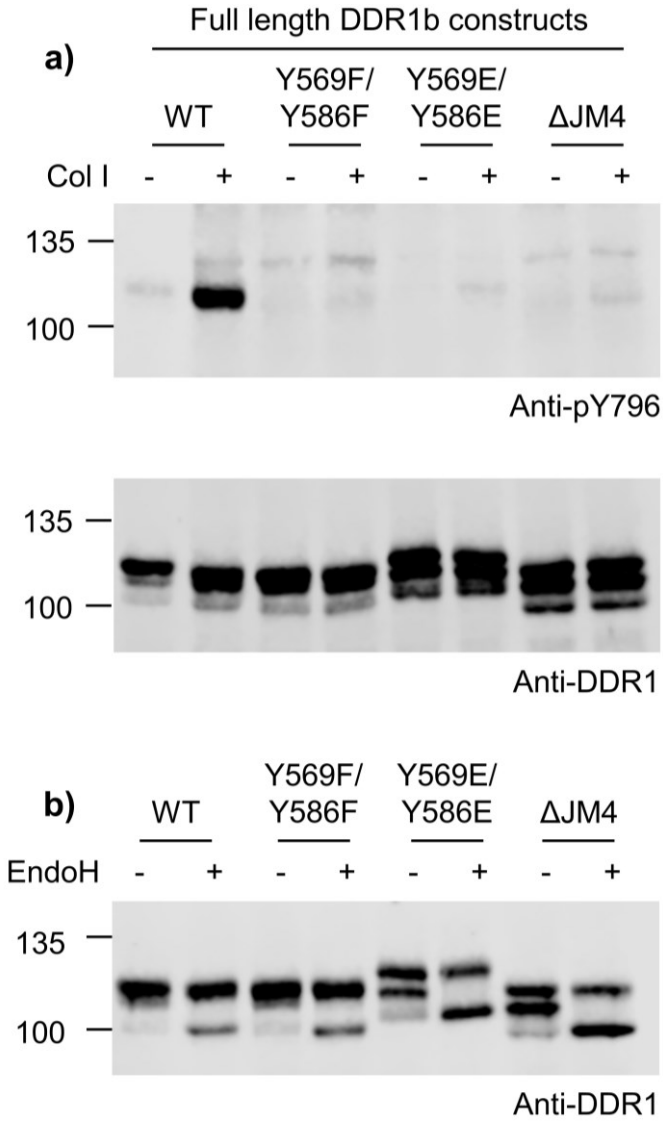

Figure 51. EndoH analysis of full-length DDR1 JM4 mutants.
a) Full-length DDR1-wild type (WT), or the DDR1 mutants Y569F/Y586F, Y569E/Y586E, and the JM4 deletion construct (ΔJM4), were transiently expressed in HEK293 cells. Cells were then stimulated for 90 minutes at 37°C with collagen I (+) or left unstimulated (-). Cells were subsequently lysed, and aliquots of cell lysates analysed by SDS-PAGE and Western blotting with the A-loop-specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. Blots are representative images from 3 independent experiments. **b)** Lysates from **a)** were cleaved with (+) EndoH or were left without (-) EndoH. Samples were subsequently boiled in sample buffer and analysed by SDS-PAGE and Western blotting with an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. This blot is a representative image from 2 independent experiments.

To better quantify the levels of protein at the cell surface, flow cytometry was used. Again, the DDR1-WT, DDR1-Y569F/Y586F, DDR1-Y569E/Y586E, and DDR1-ΔJM4 constructs were expressed in HEK293 cells. Cells were stained with anti-DDR1 antibody, and then a FITC-labelled secondary antibody. Flow cytometry analysis of these stained cells showed that all constructs were expressed on the cell surface to equivalent levels, although there was a slight increase in DDR1-Y569F/Y586F expression relative to DDR1-WT (**Figure 52a,b**). This indicates that the mutant constructs are present on the cell surface and should therefore have accessibility to collagen.

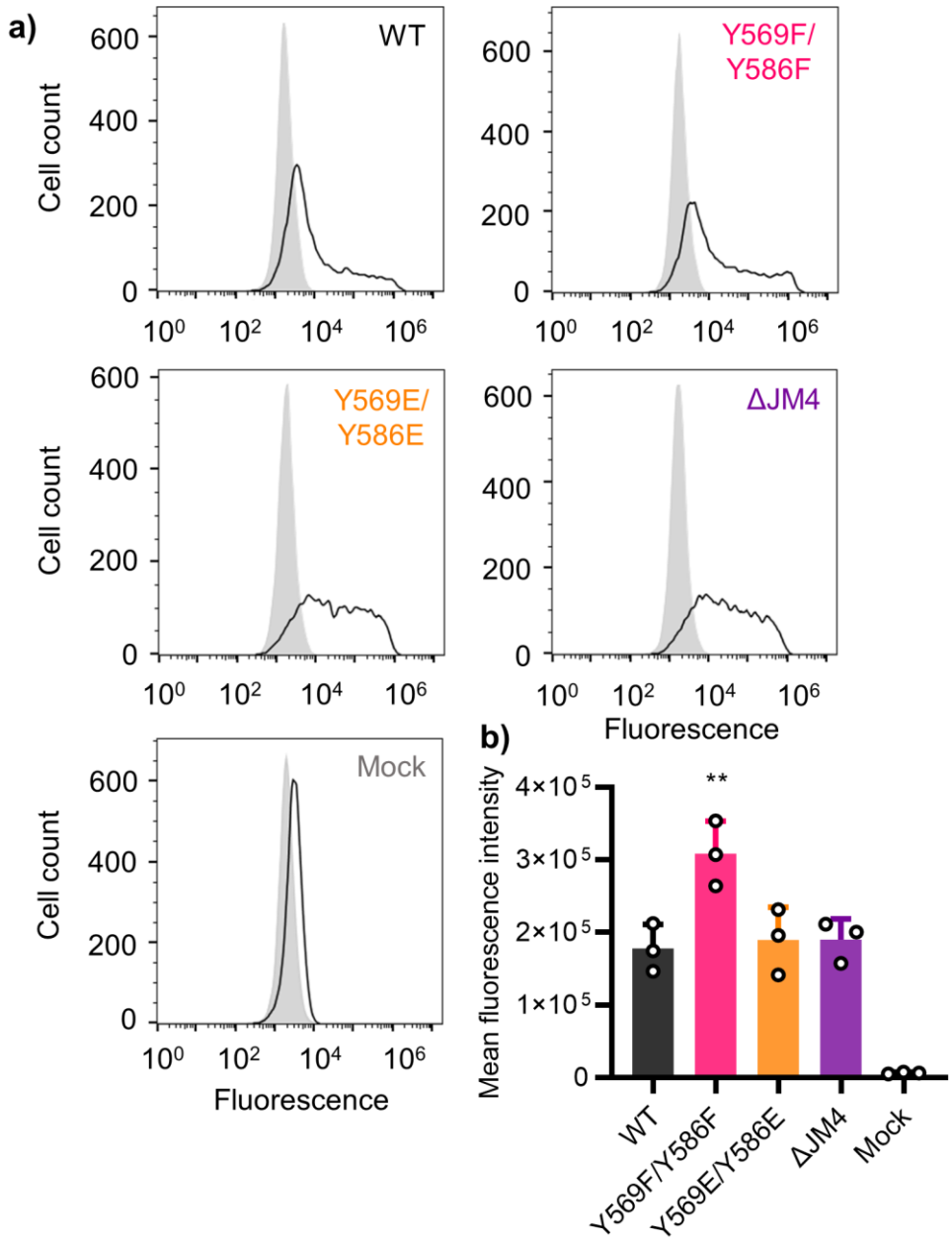

Figure 52. DDR1 JM4 mutant constructs are expressed on the cell surface. **a)** Full-length DDR1-wild type (WT), or the DDR1 mutants Y569F/Y586F, Y569E/Y586E, and the JM4 deletion construct (ΔJM4), were transiently expressed in HEK293 cells. Cells were subsequently stained with anti-DDR1 antibody against the extracellular region, followed by FITC-labelled anti-mouse-Fc antibody and analysed by flow cytometry. The filled grey histograms represent secondary only staining, while the open histograms represent anti-DDR1 and secondary antibody staining. **b)** Mean fluorescence intensity for each DDR1 construct is shown from 3 independent experiments with the standard error of the mean. ** - p=0.0065; statistically significant difference found between DDR1-WT and DDR1-Y569F/Y586F. DDR1-Y569F/Y586F signal was also significantly higher than DDR1-Y569E/Y586E and DDR1-ΔJM4 (p=0.0122 and p=0.0123, respectively). No statistically significant difference was found between DDR1-WT and DDR1-Y569E/Y586E, or DDR1-ΔJM4 constructs (p=0.9922 and p=0.9918, respectively). One-way ANOVA with Tukey post-test was performed.

5.2.4 JM4 mutants are not active in the insoluble fraction

Flow cytometry analysis of the full-length DDR1 JM4 mutants, DDR1-Y569E/Y586E and DDR1-ΔJM4, showed that the proteins were expressed at the cell surface; however, EndoH digestion indicated that they may not be processed in the same way as the wild type protein. To establish whether the smaller amount of EndoH resistant forms of the JM4 mutants seen on the Western blots was due to a loss of protein during cell lysis, the insoluble fraction produced following cell lysis was analysed. The insoluble fraction from cell lysates of DDR1-WT, DDR1-Y569F/Y586F, DDR1-Y569E/Y586E, or DDR1-ΔJM4 expressing cells was boiled in sample buffer and analysed by Western blotting (**Figure 53**). DDR1 protein was detected in the insoluble fraction for all constructs. Interestingly, stimulation with collagen resulted in an increased prevalence of DDR1-WT in the insoluble fraction, which was found to be phosphorylated (anti-pY796, right blot). This indicated that activation of DDR1 makes the protein less soluble. This is likely due to clustering of the protein into large aggregates following collagen stimulation; a process which is critical in DDR1 activation (Corcoran et al, 2019; Yeung et al, 2019). DDR1 was also present in the insoluble fraction of the lysates expressing mutant DDR1. However, this insoluble mutant protein was not found to be phosphorylated. Therefore, the reason for the absence of a phosphorylation signal following collagen stimulation (**Figure 49**) within these mutant DDR1 constructs was not because the phosphorylated protein was being lost in the insoluble fraction during cell lysis. There was a somewhat higher level of DDR1-Y569E/Y586E and DDR1-ΔJM4 (both unstimulated (-) and stimulated (+) conditions) in the insoluble fraction compared with the DDR1-WT (-) collagen and DDR1-Y569F/Y586F (+) and (-) collagen conditions, indicating that these proteins may be less soluble than the wild type receptor.

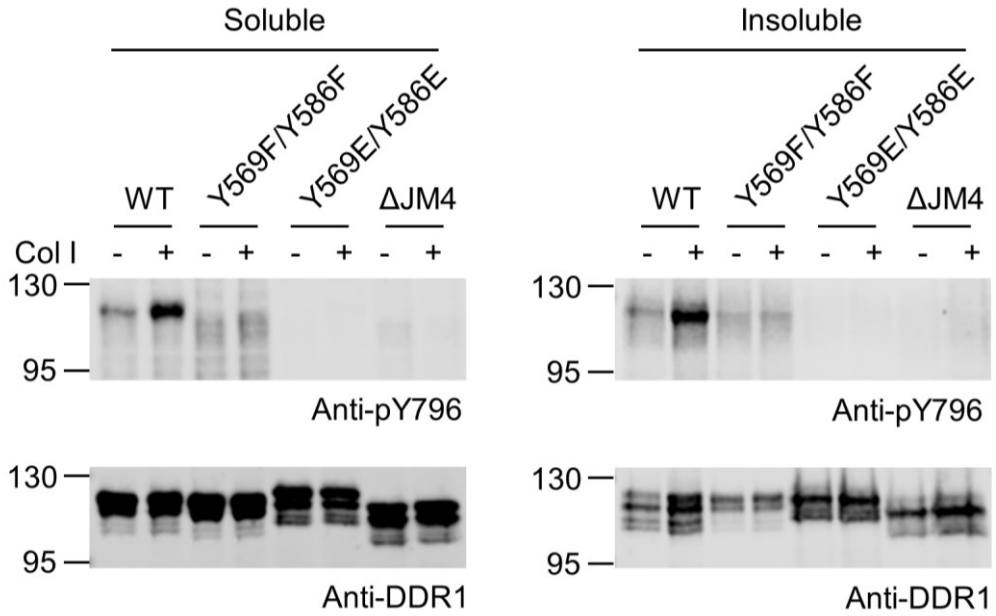

Figure 53. DDR1 JM4 mutants are not active in the insoluble fraction.
Full-length DDR1-wild type (WT), or the DDR1 mutants Y569F/Y586F, Y569E/Y586E, and the JM4 deletion construct (ΔJM4), were transiently expressed in HEK293 cells. Cells were then stimulated for 90 minutes at 37°C with collagen I (+) or left unstimulated (-). Cells were subsequently lysed. The insoluble fraction from cell lysis was then sonicated in sample buffer and boiled. Aliquots of the soluble (left) and insoluble (right) fractions were then analysed by SDS-PAGE and Western blotting with the A-loop-specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. Blots are representative images from 3 independent experiments.

5.2.5 DDR1 JM4 mutants can cluster following collagen stimulation

Thus far, it has been confirmed that the DDR1 JM4 mutants are cell surface expressed (**Figure 52**) and can therefore be exposed to collagen. However, whilst phosphomimetic mutation of the JM4 region resulted in an increased kinase activity using soluble DDR1 constructs *in vitro* (**Figure 41**), cell expressed DDR1-Y569E/Y586E was found to be inactive upon collagen stimulation (**Figure 49**). Therefore, it was next important to establish whether the mutant protein could form clusters following collagen stimulation; a process which is vital for receptor activation and *trans*-autophosphorylation (Corcoran et al, 2019; Juskaite et al, 2017; Yeung et al, 2019). To establish whether this vital process was impaired in the JM4 mutant constructs, DDR1-WT, DDR1-Y569F/Y586F, or DDR1-Y569E/Y586E, were expressed in Cos7 cells which were subsequently stimulated with collagen I for 10 or 60 minutes or left unstimulated. Live cells were then stained with anti-DDR1 before

being fixed, permeabilised and stained for phospho-DDR1 (anti-pY513). In the absence of collagen stimulation, DDR1-WT was found to occupy punctate structures (**Figure 54 –** top panel). However, following 10 minutes of collagen stimulation the receptor redistributed into clusters with a linear branching morphology (**Figure 54 –** middle panel), in agreement with previously published data (Corcoran et al, 2019). At this point some minimal levels of phospho-DDR1 signal were detected. After 60 minutes of collagen stimulation, these clusters grew more dense in some areas and robust phospho-DDR1 signal was detected (**Figure 54**). This same collagen-dependent receptor redistribution was also observed for the JM4 mutant constructs, DDR1-Y569F/Y586F (**Figure 55**), and DDR1-Y569E/Y586E (**Figure 56**). However, no DDR1 autophosphorylation was observed for these mutants by immunofluorescence imaging. This lack of autophosphorylation was confirmed by Western blotting lysates of DDR1 construct-expressing Cos7 cells with the A-loop specific anti-phosphotyrosine antibody, anti-pY796 (**Figure 57**). These data demonstrated that the JM4 mutants can bind collagen and form clusters at the cell surface. However, they are unable to catalyse autophosphorylation within these clusters.

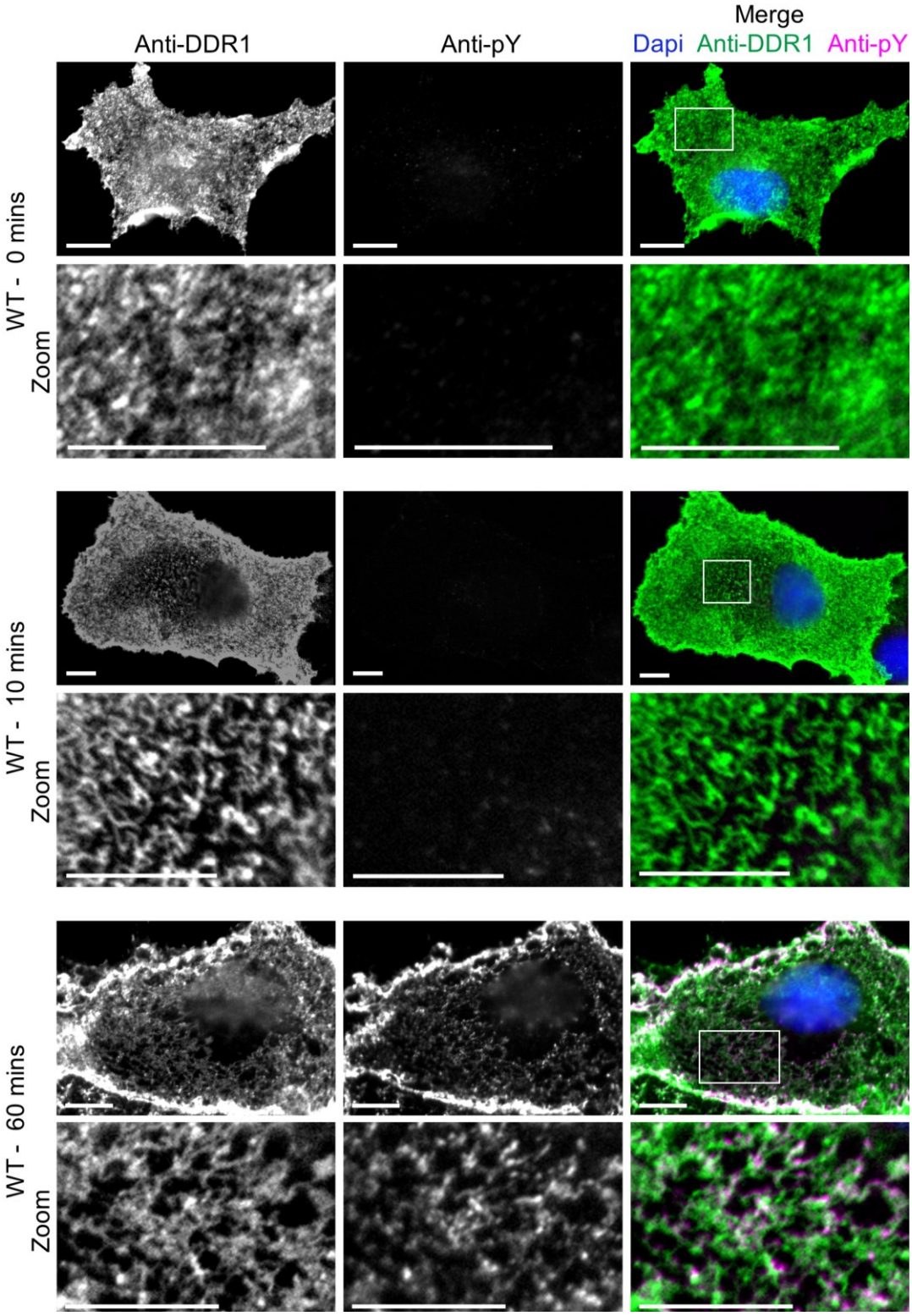

Figure 54. DDR1 clusters following collagen stimulation.
Full-length DDR1- wild type (WT) was transiently expressed in Cos7 cells grown on coverslips. Cells were then stimulated with collagen I (+) for 10 or 60 minutes or left unstimulated (-) for 60 minutes at 37°C. Cells were then stained with anti-DDR1 (green) on ice before being fixed, permeabilised and stained with anti-pY513 (magenta), and finally Dapi (blue). Scale bar = 10 μm. White box indicates magnified region of interest. Widefield micrographs using a 63x oil immersion lens.

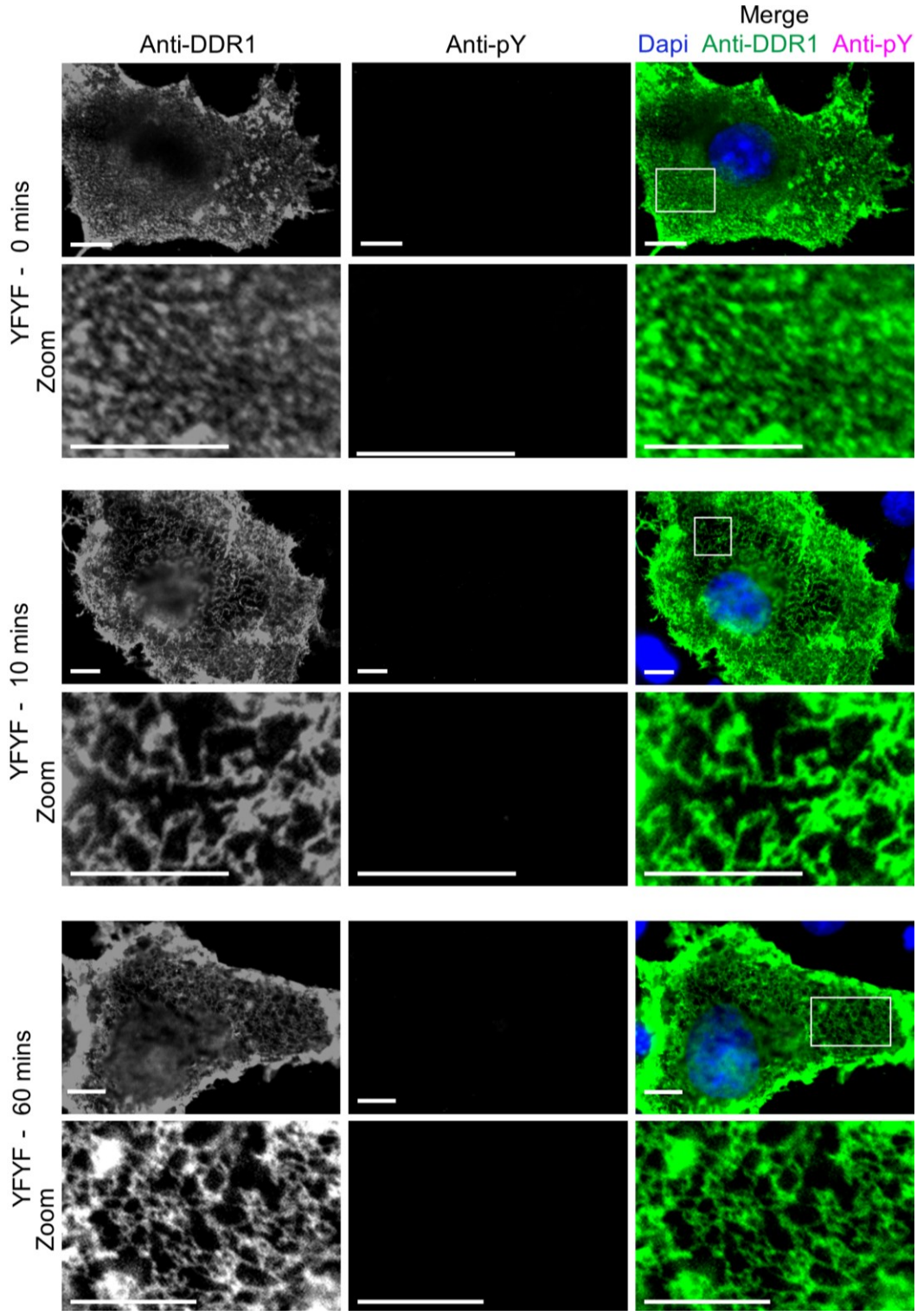

Figure 55. DDR1-Y569F/Y586F clusters following collagen stimulation.
Full-length JM4 mutant DDR1-Y569F/Y586F was transiently expressed in Cos7 cells grown on coverslips. Cells were then stimulated with collagen I (+) for 10 or 60 minutes or left unstimulated (-) for 60 minutes at 37°C. Cells were then stained with anti-DDR1 (green) on ice before being fixed, permeabilised and stained with anti-pY513 (magenta), and finally Dapi (blue). Scale bar = 10 µm. White box indicates magnified region of interest. Widefield micrographs using a 63x oil immersion lens.

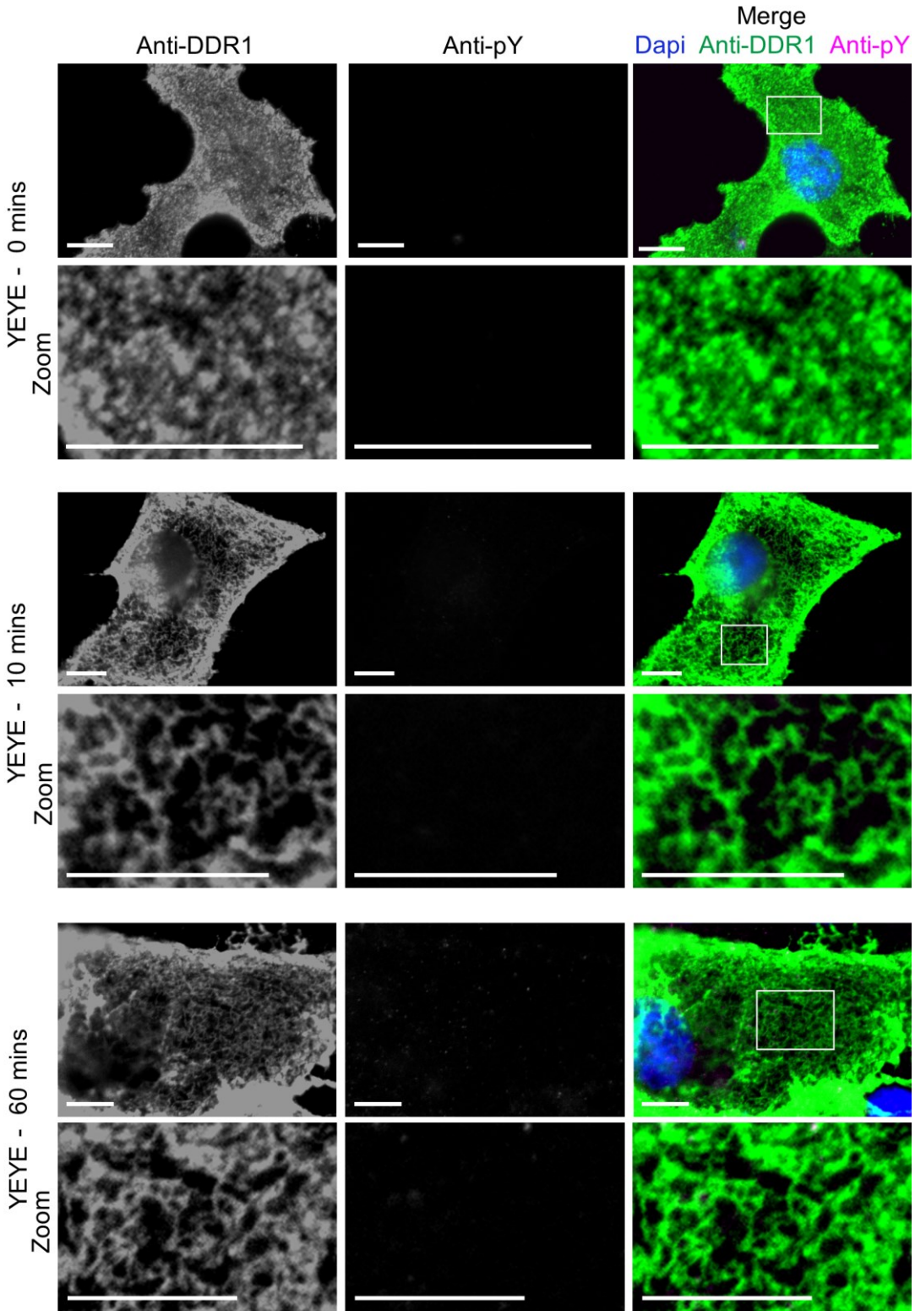

Figure 56. DDR1-Y569E/Y586E clusters following collagen stimulation.
Full-length JM4 mutant DDR1-Y569E/Y586E was transiently expressed in Cos7 cells grown on coverslips. Cells were then stimulated with collagen I (+) for 10 or 60 minutes or left unstimulated (-) for 60 minutes at 37°C. Cells were then stained with anti-DDR1 (green) on ice before being fixed, permeabilised and stained with anti-pY513 (magenta), and finally Dapi (blue). Scale bar = 10 µm. White box indicates magnified region of interest. Widefield micrographs using a 63x oil immersion lens.

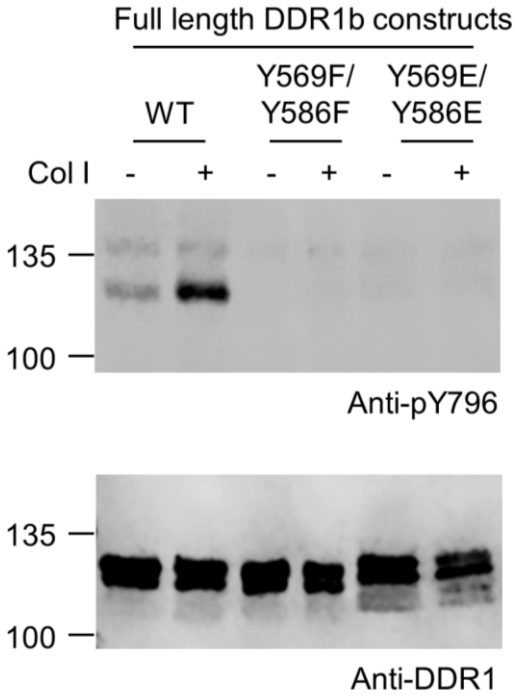

Figure 57. JM4 mutants are inactive in Cos7 cells.
Full-length DDR1-wild type (WT), or the DDR1 mutants Y569F/Y586F, Y569E/Y586E, and the JM4 deletion construct (ΔJM4), were transiently expressed in Cos7 cells. Cells were then stimulated with collagen I (+) or left unstimulated (-) for 90 minutes at 37°C. Cells were subsequently lysed, and aliquots of cell lysates analysed by SDS-PAGE and Western blotting with the activation loop-specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. Blots are representative images from 2 independent experiments.

5.2.6 JM4 mutants are substrates for *trans*-autophosphorylation

The ability of DDR1 JM4 mutant constructs to form ligand-induced clusters has been demonstrated (**Figure 55**, **Figure 56**). However, no autophosphorylation was found to occur within these clusters. It was next important to establish whether this loss of activation was because the protein could not form productive substrate interactions within these clusters.

Mutation of Tyr513 of the DDR1 JM region to Phe does not affect receptor autophosphorylation of the A-loop meaning that the protein is catalytically proficient (Juskaite et al, 2017). However, binding to the anti-pY513 anti-phosphotyrosine antibody is lost in a DDR1-Y513F mutant. Phosphorylation of Tyr513 of DDR1 has been shown to occur in *trans* within DDR1 clusters (Juskaite et al, 2017). Therefore, co-expression of DDR1-Y513F (enzyme) with an autophosphorylation deficient

DDR1 construct (substrate) should allow the determination of whether *trans*-phosphorylation has occurred between the enzyme (DDR1-Y513F) and the substrate (autophosphorylation deficient DDR1 e.g. DDR1-Y569F/Y586F) by monitoring phosphorylation at Tyr513 with anti-pY513 anti-phosphotyrosine antibody (anti-pY513 can only bind phosphorylated substrate, not DDR1-Y513F (enzyme)).

DDR1-Y513F (enzyme) was co-expressed with DDR1-Y569F/Y586F (substrate) in HEK293 cells which were subsequently stimulated with collagen I (**Figure 58**). When expressed alone, no anti-pY513 signal was detected following collagen stimulation for either construct, as expected (**Figure 58**). However, when DDR1-Y513F and DDR1-Y569F/Y586F were co-expressed, phosphorylation was detected using anti-pY513. This demonstrated that DDR1-Y569F/Y586F, whilst defective in autophosphorylation when expressed on its own, could act as a substrate for phosphorylation by DDR1-Y513F. Likewise, when DDR1-Y569E/Y586E (substrate) was co-expressed with DDR1-Y513F (enzyme) phosphorylation of DDR1-Y569E/Y586E was detected using anti-pY513 (**Figure 59**). This also demonstrated that DDR1-Y569E/Y586E can be phosphorylated in *trans.* These data supported the finding that both JM4 mutant constructs can form receptor clusters and indicate that co-clustering with DDR1-Y513F is sufficient to enable *trans*-autophosphorylation to occur.

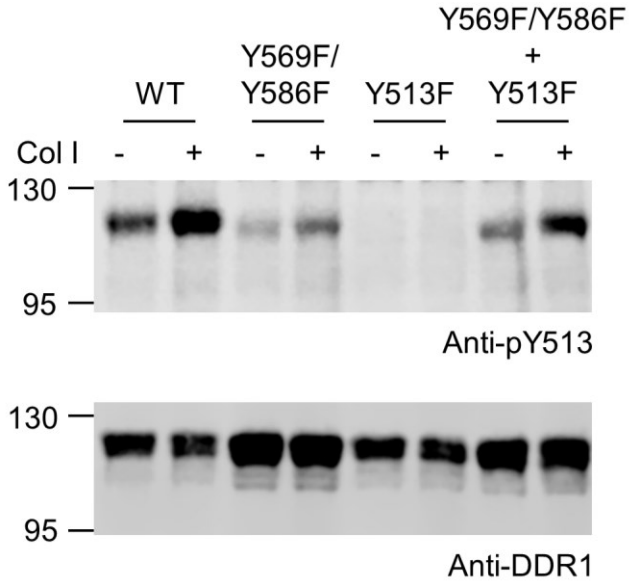

Figure 58. DDR1 Y569F/Y586F is a substrate for phosphorylation.
Full-length DDR1-wild type (WT), the DDR1 JM4 mutant Y569F/Y586F, and Y513F, were transiently expressed alone or in combination in HEK293 cells. Cells were then stimulated with collagen I (+) or left unstimulated (-) for 90 minutes at 37°C. Cells were subsequently lysed, and aliquots of cell lysates analysed by SDS-PAGE and Western blotting with the JM specific anti-phosphotyrosine antibody, anti-pY513. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. Blots are representative images from 2 independent experiments.

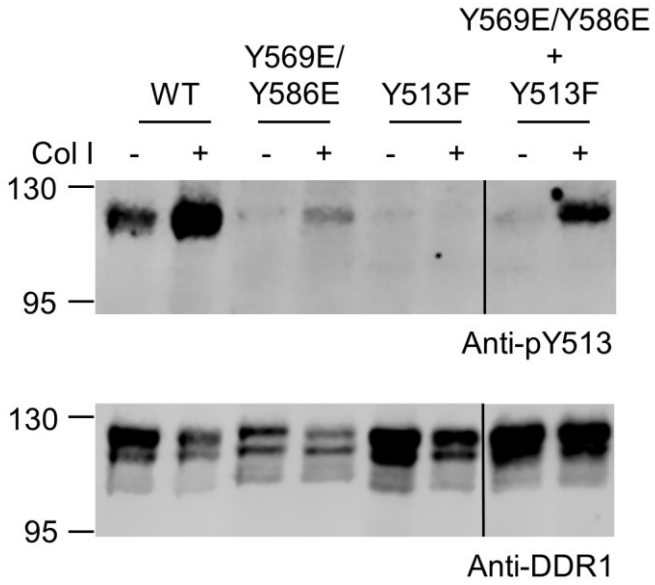

Figure 59. DDR1-Y569E/Y586E is a substrate for phosphorylation.
Full-length DDR1- wild type (WT), the DDR1 JM4 mutant (Y569E/Y586E), and Y513F, were transiently expressed alone, or in combination, in HEK293 cells. Cells were then stimulated with collagen I (+) or left unstimulated (-) for 90 minutes at 37°C. Cells were subsequently lysed, and aliquots of cell lysates analysed by SDS-PAGE and Western blotting with the JM specific anti-phosphotyrosine antibody, anti-pY513. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. The black line indicates where two lanes containing unrelated conditions have been removed from the final image for consistency with **Figure 58**. The unedited blots can be found in **Chapter 7, Figure 65**. Blots are representative images from 2 independent experiments.

5.2.7 DDR1-Y569F/Y586F has impaired catalytic ability *in vitro*

In cells, the DDR1-Y569F/Y586F, DDR1-Y569E/Y586E, and DDR1-ΔJM4 constructs all showed impaired collagen-induced receptor autophosphorylation (**Figure 49, Figure 50)**. For the latter two constructs, this was an unexpected result. To determine whether this loss of receptor activation upon collagen stimulation was due to impaired catalytic activity, or another factor within the cell, the mutant constructs, or DDR1-WT, were expressed in HEK293 cells. Cells were subsequently lysed and protein immunoprecipitated from lysates using anti-DDR1 and Protein A Sepharose beads. Immunoprecipitated protein was then stimulated with ATP for 30 minutes and proteins analysed by Western blotting with the A-loop specific anti-phosphotyrosine antibody, anti-pY796 (**Figure 60**). The DDR1-Y569F/Y586F mutant showed significantly impaired autophosphorylation compared with the DDR1-WT receptor, again supporting the notion that an inability to phosphorylate the JM4 region leads to a loss of receptor function (**Figure 60b**). In contrast, the DDR1-Y569E/Y586E and DDR1-ΔJM4 constructs had comparable autophosphorylation levels to the WT receptor (**Figure 60b**). This showed that the loss of collagen-induced receptor activation with these mutant constructs was not due to a loss of catalytic function. Furthermore, the finding that the DDR1-Y569E/Y586E construct can form clusters (**Figure 56**) and be phosphorylated within those clusters when co-expressed with autophosphorylation competent DDR1-Y513F (**Figure 59**) indicates that there may be an additional cellular factor affecting the function of this mutant receptor.

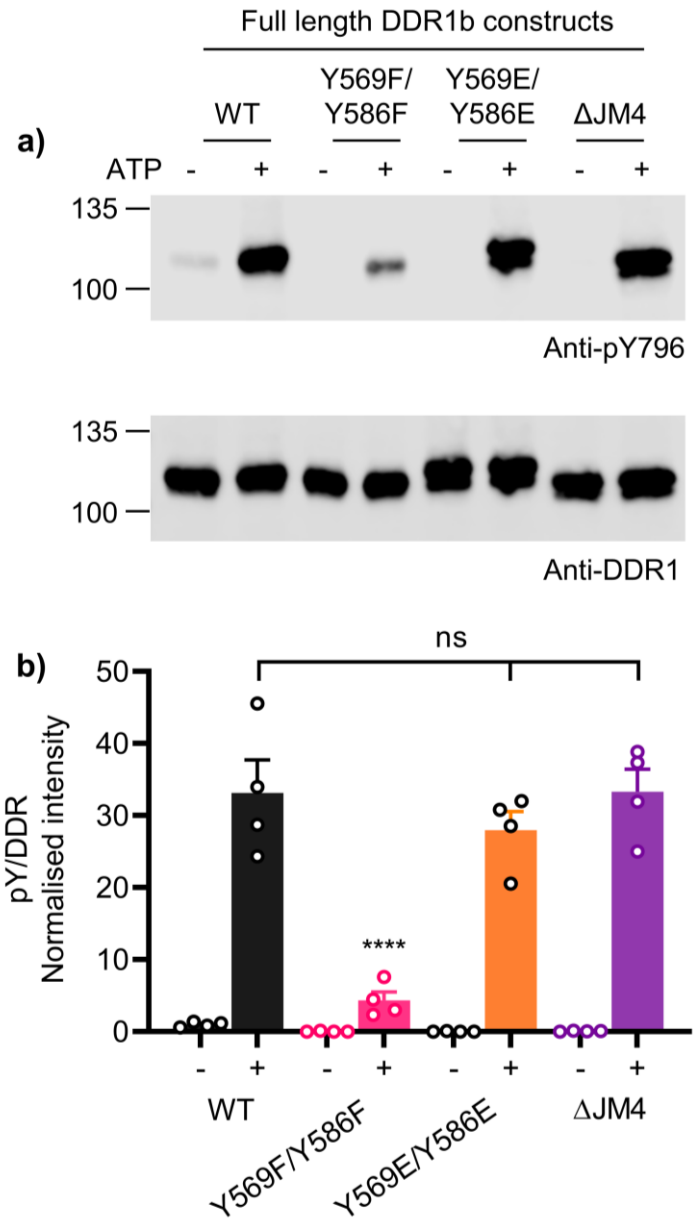

Figure 60. Full-length DDR1 Y569F/Y586F has impaired autophosphorylation. **a)** Full-length DDR1-wild type (WT) or the DDR1 mutants Y569F/Y586F, Y569E/Y586E, and the JM4 deletion construct (ΔJM4), were transiently expressed in HEK293 cells. Cells were then lysed and DDR1 immunoprecipitated from cell lysates. Immunoprecipitates were then stimulated with (+) or without (-) 1 mM ATP in IP kinase buffer for 30 minutes at 30°C. Samples were then boiled in sample buffer and analysed by SDS-PAGE and Western Blotting with the A-loop specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody. The position of molecular weight markers (in kDa) are shown on the left. **b)** Quantification of anti-pY796 signals normalised to respective anti-DDR1 signals. The anti-phosphotyrosine signal is expressed as a percentage of the sum of all bands on a blot, with mean and standard error of the mean shown (*n*=4). No significant (ns) difference was found between the WT, Y569E/Y586E, and ΔJM4 (+ ATP) conditions (p=0.7113 and p>0.9999, respectively). **** - <0.0001 statistically significant difference found between the Y569F/Y586F (+ ATP) condition and the other three + ATP stimulated conditions. Two-way ANOVA with Tukey post-test.

5.2.8 Src activation leads to DDR1 phosphorylation

Collagen binding to the phosphomimetic mutant, DDR1-Y569E/Y586E, does not lead to activation of the receptor (**Figure 49**). The observation that catalytic function (**Figure 60**) and the ability to cluster following collagen binding (**Figure 56**) is undisturbed in this mutant receptor, indicates that there may be an additional, as yet unidentified, cellular factor which is contributing to this loss of function.

Src is a non-receptor Tyr kinase which phosphorylates the DDR2 A-loop and has been found to be important in its activation in an *in vitro* study (Yang et al, 2005). Interestingly, Src also binds to phospho-peptides of Tyr569 and Tyr586 of the DDR1 JM4 region (Lemeer et al, 2012). In addition to its autoinhibitory role, the JM4 region may therefore function to recruit Src to DDR1, thereby facilitating receptor activation. To test whether Src activity can lead to DDR1 activation, DDR1-WT was expressed alone in HEK293 cells, or co-expressed with a WT or constitutively active (CA)-Src construct. Cells were then stimulated with collagen I for 1 hour at 37°C before being lysed and analysed by Western blotting with the JM4-specific anti-phosphotyrosine antibodies, anti-pY569 and anti-pY586, or the A-loop-specific anti-phosphotyrosine antibody, anti-pY796 (**Figure 61**). Endogenous Src was detected using anti-Src; overexpressed WT-Src and CA-Src constructs were detected at a higher molecular weight due to the presence of a Myc-tag. Co-expression of DDR1-WT with CA-Src resulted in phosphorylation of DDR1 at Tyr569, Tyr586, and Tyr796, in the absence of collagen stimulation (**Figure 61;** CA (-) collagen condition). This clearly indicates that Src activation leads to DDR1 phosphorylation. Interestingly, co-expression with CA-Src led to a greater enhancement of A-loop phosphorylation (anti-pY796) than JM4 phosphorylation; compare CA-Src (-) collagen and WT-Src (+) collagen conditions in the three graphs of **Figure 61b**.

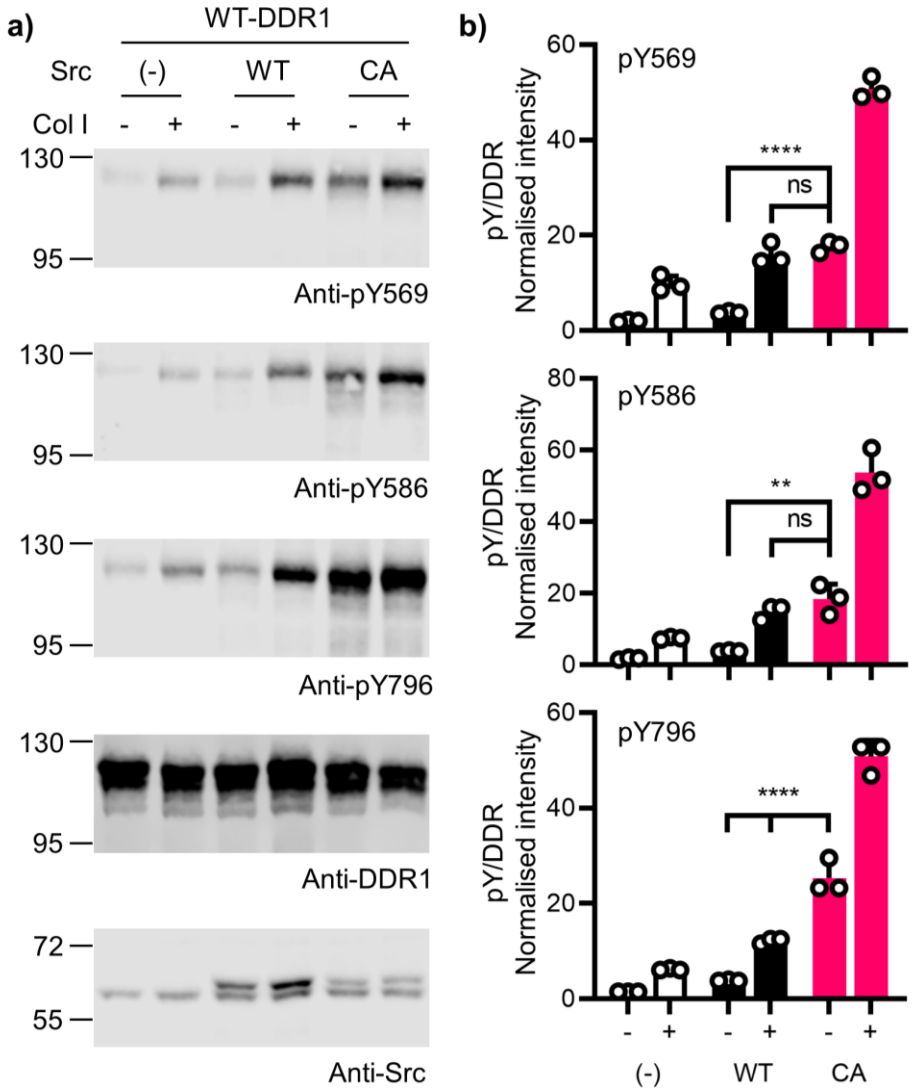

Figure 61. Src activation leads to DDR1 phosphorylation.
a) Full-length DDR1-wild type (WT) was transiently expressed alone (-), co-expressed with wild type (WT) Src, or co-expressed with constitutively active (CA) Src, in HEK293 cells. Cells were then stimulated with collagen I (+), or left unstimulated (-), for 60 minutes at 37°C. Cells were subsequently lysed, and aliquots of cell lysates analysed by SDS-PAGE and Western blotting with the JM4-specific anti-phosphotyrosine antibodies, anti-pY569 and anti-pY586, as well as the A-loop-specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody and total Src levels detected using an anti-Src antibody. The position of molecular weight markers (in kDa) are shown on the left. **b)** Quantification of anti-phosphotyrosine signals (anti-pY569, anti-pY586, and anti-pY796) normalised to respective anti-DDR1 signals. The anti-phosphotyrosine signal is expressed as a percentage of the sum of all bands on a blot, with mean and standard error of the mean shown (*n*=3). Two-way ANOVA with a Tukey post-test was performed. (-) + Col I vs WT + Col I; p=0.0043 (**), 0.1057 (ns), and 0.0317 (*) for anti-pY569, anti-pY586, and anti-pY796, respectively. WT + Col I vs CA – Col I; p=0.7910 (ns), 0.7384 (ns), <0.0001 (****) for anti-pY569, anti-pY586, and anti-pY796, respectively. CA – Col I vs CA + Col I; p=<0.0001 (****) for all antibodies. CA – Col I vs WT -; p=<0.0001 (****), 0.0011 (**), <0.0001 (****), for anti-pY569, -pY586, and -pY796, respectively.

5.2.9 Src cannot phosphorylate DDR1 JM4 mutants

To establish whether Src-mediated phosphorylation of DDR1 was affected by mutation of the JM4 region, DDR1-WT or the JM4 mutant constructs (DDR1-Y569F/Y586F and DDR1-Y569E/Y586E) were co-expressed with WT-Src or CA-Src in HEK293 cells. Cells were subsequently lysed and analysed by Western blotting with the A-loop-specific anti-phosphotyrosine antibody, anti-pY796 (**Figure 62**). Again, co-expression of DDR1-WT with CA-Src led to receptor phosphorylation in the absence of collagen stimulation. However, mutation of the DDR1 JM4 region (DDR1-Y569F/Y586F and DDR1-Y569E/Y586E) led to a loss of CA-Src mediated DDR1 phosphorylation (**Figure 62**). Furthermore, when this same Src co-expression study was repeated with a kinase dead DDR1 mutant (DDR1-K655M), CA-Src mediated phosphorylation of DDR1 was also lost (**Figure 63**). This indicated that Src mediated DDR1 phosphorylation requires a functional kinase and a wild type DDR1 JM4 region.

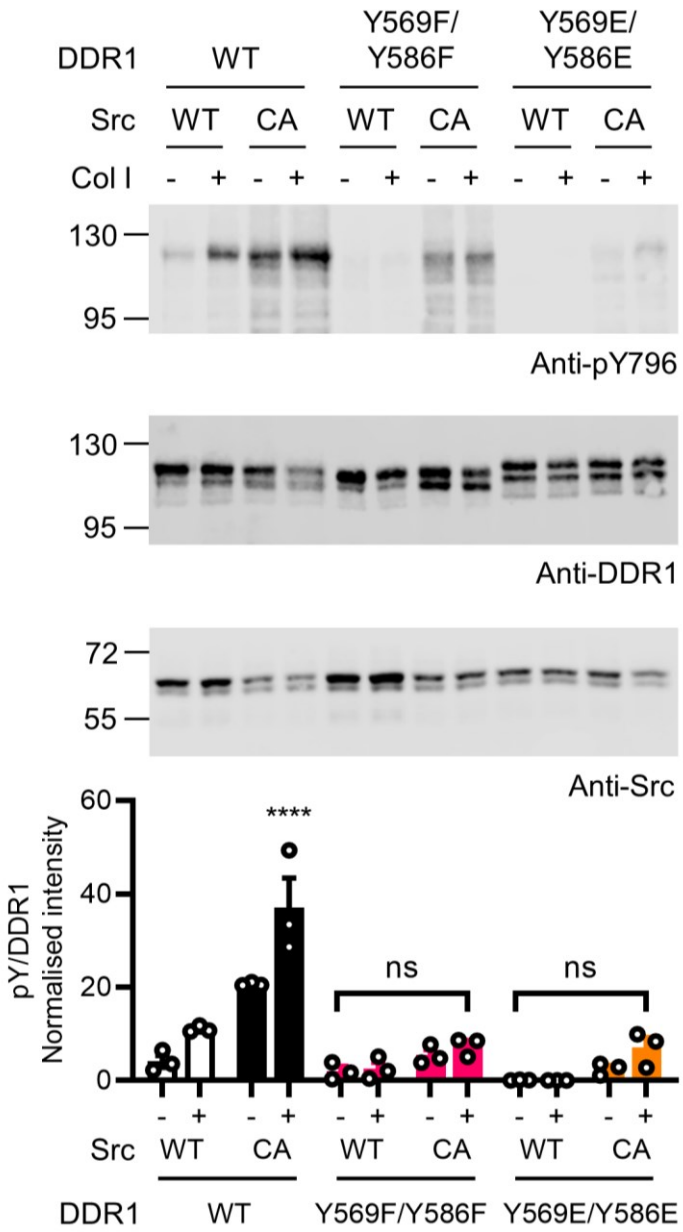

Figure 62. JM4 mutants are not phosphorylated by Src.
Full-length DDR1-wild type (WT) or the indicated DDR1 mutants were co-expressed with wild type (WT) or constitutively active (CA) Src in HEK293 cells. Cells were then stimulated with collagen I (+), or left unstimulated (-), for 60 minutes at 37°C. Cells were subsequently lysed, and aliquots of cell lysates analysed by SDS-PAGE and Western blotting with the A loop-specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody and total Src levels detected using the anti-Src antibody. The position of molecular weight markers (in kDa) are shown on the left. b) Quantification of anti-pY796 signal normalised to respective anti-DDR1 signals. The anti-phosphotyrosine signal is expressed as a percentage of the sum of all bands on a blot, with mean and standard error of the mean shown (n=3). Two-way ANOVA with a Tukey post-test was performed. $p<0.0001$ (****); DDR1-WT+CA-Src (+) Col I vs all other conditions. $p=0.7712$ (ns–not significant); DDR1-Y569F/Y586F+WT-Src (-) Col I vs DDR1-Y569F/Y586F+CA-Src (+) Col I. $p=0.9992$ (ns); DDR1-Y569E/Y586E+WT-Src (-) Col I vs DDR1-Y569F/Y586F+CA-Src (+) Col I.

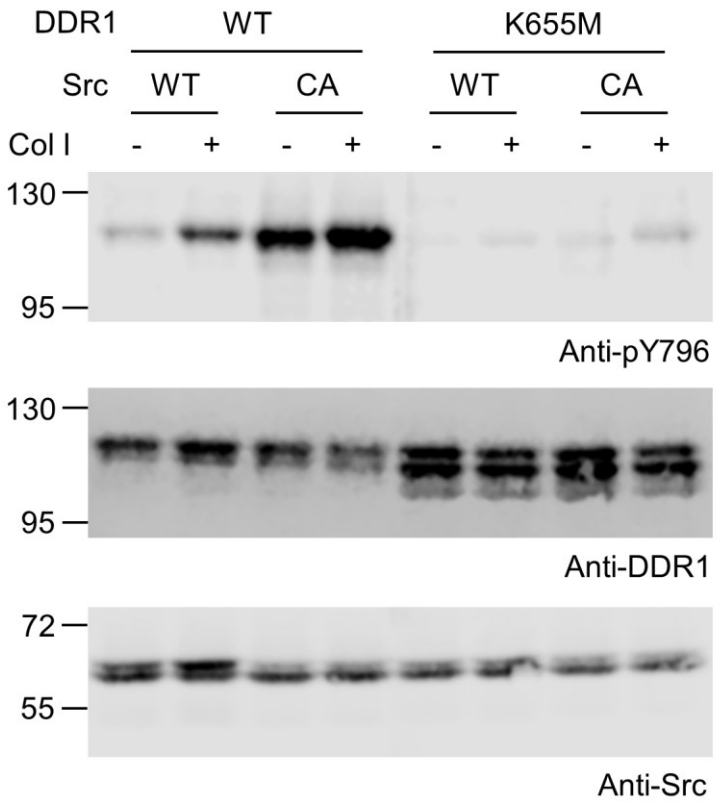

Figure 63. Kinases inactive DDR1 is not phosphorylated by Src.
Full-length DDR1-wild type (WT) or DDR1-kinase dead (K655M) was co-expressed with WT or constitutively active (CA) Src in HEK293 cells. Cells were then stimulated with collagen I (+), or left unstimulated (-), for 60 minutes at 37°C. Cells were subsequently lysed, and aliquots of cell lysates analysed by SDS-PAGE and Western blotting with the A-loop-specific anti-phosphotyrosine antibody, anti-pY796. Total DDR1 levels were then detected using an anti-DDR1 antibody and total Src levels detected using the anti-Src antibody. The position of molecular weight markers (in kDa) are shown on the left. Representative images from two independent experiments are shown.

5.3 Discussion

In **Chapter 3** and **Chapter 4**, structural and biochemical analysis of DDR1 kinase revealed that the protein exists in a dual autoinhibited state (A-loop and JM4 region) which is relieved through an ordered phosphorylation mechanism, beginning with the JM4 region. In this chapter, DDR1 activation was monitored in the full-length mammalian cell-expressed receptor. In agreement with the preceding chapters, the JM4 region is shown to be phosphorylated before the A-loop (**Figure 48**). Furthermore, mutation of the JM4 region led to a loss of receptor activation (**Figure 49**). However, in contrast to the *in vitro* data, phosphomimetic mutation of JM4 Tyr residues (DDR1-Y569E/Y586E), or deletion of the JM4 region (DDR1-ΔJM4), also led to a loss of DDR1 activation (**Figure 49**, **Figure 50**). This unexpected finding suggested that the DDR1 JM4 region may have an additional positive role in DDR1 signalling.

The use of soluble recombinant protein constructs in the preceding chapters produced a system which was easily manipulatable, but also reductionist, as it removed the complexity of interacting cellular proteins, and other DDR1 domains. It was therefore important to ensure that findings produced within this *in vitro* system could be recapitulated using the full-length protein within a cellular context. In agreement with *in vitro* data, the DDR1-Y569F/Y586F mutant created a catalytically inactive receptor (**Figure 60**). However, the phospho-mimetic mutant, and JM4 deletion constructs, which had been expected to increase activity through relief of JM4 autoinhibition, led to a similar loss of receptor activation by collagen (**Figure 49**, **Figure 50**). It is interesting to note that while JM4 deletion led to a loss of function, deletion of the preceding JM1-3 regions had little effect on receptor activation (**Figure 50**). This suggests that the long DDR1 JM1-3 region does not play a functional role in activation. The JM region of most RTKs typically comprises ~ 20 amino acids and is capable of binding to several receptor effectors following autophosphorylation (Hubbard, 2001; 2004; Lemmon & Schlessinger, 2010). Why DDR1 requires such a long intracellular JM region is unknown. There are 7 Tyr residues in the DDR1 JM region (Tyr484 in JM2 region; Tyr513 and Tyr520 in the DDR1-specific JM region. Tyr543 and Tyr547 in JM3 region. Tyr569 and Tyr586 in JM4 region), all of which have several predicted binding partners (**Figure 8**)(Leitinger,

2014; Lemeer et al, 2012). Why other RTKs can bind their intracellular signalling partners with significantly shorter JMs than DDR1 is unclear. It is possible that because DDR1 forms clusters upon activation, the long intracellular JM region allows projection of the kinase domain and other autophosphorylation sites away from dense sites of clustering (Corcoran et al, 2019). This could facilitate interactions with binding partners and the formation of signalling hubs, which may be occluded from binding closer to the TM region. It is important to note however that, unlike DDR1, most RTKs also contain a C-terminal tail which can bind to signalling proteins (Gajiwala, 2013; Lemmon & Schlessinger, 2010).

The unexpected loss of function of the phosphomimetic and JM4 deletion constructs (DDR1-Y569E/Y586E and DDR1-ΔJM4) led to attempts to characterise these proteins, to ensure they were processed correctly. Both mutant constructs were detected at the cell surface using flow cytometry at similar levels to the wild type receptor (**Figure 52**). However, EndoH analysis revealed that a higher proportion of the proteins was susceptible to cleavage and therefore did not contain mature oligosaccharides (**Figure 51**); a characteristic indicative of intracellular ER/Golgi retention (Koide & Muramatsu, 1974; Taner et al, 2011). This discrepancy between EndoH and flow cytometric analysis could be due to a loss of EndoH resistant DDR1 mutant protein in the insoluble pellet following cell lysis. Indeed, the DDR1-Y569E/Y586E and DDR1-ΔJM4 constructs were somewhat less soluble than DDR1-WT (**Figure 53**). Despite this, DDR1-Y569E/Y586E was shown to redistribute into receptor clusters following collagen binding (**Figure 56**), within which it could be phosphorylated *in trans* by a catalytically competent DDR1 construct (**Figure 59**). Furthermore, when DDR1-Y569E/Y586E and DDR1-ΔJM4 were immunoprecipitated from cell lysates they could autophosphorylate successfully (**Figure 60**). This suggested that the loss of receptor activation, following JM4 phosphomimetic mutation, was due to a cell-specific, positive role for the JM4 region.

As discussed in the preceding chapters, the JM region of RTKs is a common autoinhibitory motif which varies in its mechanism of action (Binns et al, 2000; Mol et al, 2004; Till et al, 2002). However, it can play a positive role in receptor activation. In EGFR the JM region promotes the formation of functionally competent receptor dimers (Jura et al, 2009; Red Brewer et al, 2009). Whereas in the Ret receptor, the

JM functions intramolecularly to promote kinase activity (Plaza-Menacho et al, 2016). It is possible the DDR1 JM4 region also plays a structural role in the promotion of receptor activation. Interactions between the phosphorylated JM4 region and the intracellular domains of adjacent DDR1 proteins could facilitate *trans*-autophosphorylation of the receptor. Indeed, this dual JM function has been described for the insulin receptor which is initially autoinhibited by its JM region (Craddock et al, 2007). However, following phosphorylation, the JM region promotes receptor *trans*-autophosphorylation through interactions with its dimer partners kinase domain (Cabail et al, 2015). Structural analysis of the phosphorylated DDR1 JM4-Kinase (3P) phospho-form generated in **Chapter 4** could have provided evidence of these in *trans* JM4 interactions. However, crystallographic attempts were unsuccessful. Alternatively, therefore, cryo-electron microscopy could be employed on full length DDR1 in its active state in lipid nanodiscs, as has been achieved with the insulin receptor (Weis et al, 2018). Understanding the structural rearrangements which follow DDR1 autophosphorylation and the effects this has on receptor activation and oligomerisation will be crucial in improving our understanding of DDR biology. In addition to the direct structural contributions to receptor activation seen in the EGFR, Ret, and insulin receptors (Cabail et al, 2015; Plaza-Menacho et al, 2016; Red Brewer et al, 2009), the JM region can also bind to cellular potentiators of receptor signalling, such as kinases (Bromann et al, 2004; Lemmon & Schlessinger, 2010). Src is a non-receptor tyrosine kinase which can bind to RTKs through its SH2 domain and catalyse receptor phosphorylation (Mori et al, 1993). For example, in EGFR, Src provides an EGF-independent mechanism through which the receptor can become activated (El-Hashim et al, 2017; Lombardo et al, 1995). Src is important for DDR2 signalling (Ikeda et al, 2002; Yang et al, 2005; Zhang et al, 2013). Furthermore, Src has been implicated in DDR1 signalling, and a phosphoproteomic study, using phospho-peptides, revealed that the DDR1 JM4 Tyr residues, Tyr569 and Tyr586, bind to Src (Lemeer et al, 2012; Lu et al, 2011). Here, Src activation was shown to lead to DDR1 phosphorylation in the JM4 region (Tyr569 and Tyr586) and A-loop (Tyr796; **Figure 61**). Src-mediated phosphorylation of DDR1 likely relieved autoinhibition and enabled autophosphorylation by DDR1, as is seen between Src and other RTKs like EGFR (Belsches et al, 1997; Bromann et al, 2004). Therefore, it is difficult to establish the residue which is directly phosphorylated as a result of Src activity. However, Src activation led to a greater enhancement of A-loop

phosphorylation than JM4 region phosphorylation (**Figure 61**). This suggested that the A-loop may be the site of Src-mediated DDR1 phosphorylation. In support of this, DDR2 is phosphorylated on the A-loop by Src (Yang et al, 2005). Interestingly, Src could not trigger the phosphorylation of the JM4 Tyr mutants, DDR1-Y569F/Y586F or DDR1-Y569E/Y586E, indicating that intact JM4 Tyr residues are required for Src function (**Figure 62**). Indeed, phosphorylated Tyr569 and Tyr586 are predicted to bind Src (Lemeer et al, 2012). In addition, Src was unable to trigger phosphorylation of a kinase dead DDR1 mutant, indicating that DDR1 kinase activity is also required for Src function (**Figure 63**). Together, these data lead to a model in which DDR1 autophosphorylates the JM4 region, this provides phospho-Tyr sites for Src binding (Lemeer et al, 2012). Src is then able to phosphorylate the DDR1 A-loop and potentiate DDR1 activation. An issue with this model is that DDR1 can phosphorylate its A-loop in the absence of Src *in vitro* (**Figure 33**). It will therefore be important in future to establish whether Src activity is important for DDR1 activation in cells. This could be achieved by inhibiting Src with a kinase inhibitor such as PP2, before monitoring collagen induced DDR1 activation (Wang et al, 2019). A drawback with the use of inhibitors, however, is the promiscuity of DDR1 towards kinase inhibitors (Hanson et al, 2019). Therefore, parallel experiments where Src is knocked-down (siRNA), or knocked-out (CRISPR), before monitoring collagen induced DDR1 activation, will be essential. These studies will also be important to establish whether Src represents a canonical part of the collagen mediated DDR1 activation mechanism, or if it activates DDR1 following inputs from other signalling pathways, as it does for EGFR (El-Hashim et al, 2017; Lombardo et al, 1995). For example, DDR1 has been shown to cross-talk with β1-integrins (Staudinger et al, 2013; Xu et al, 2012), which are known activators of Src signalling (Arias-Salgado et al, 2003). Furthermore, DDR1 interacts with, and is modulated by, the insulin receptor and IGF1-R (Malaguarnera et al, 2015; Vella et al, 2019; Vella et al, 2017). It is possible, therefore, that signalling through these, or other receptors, could lead to DDR1 phosphorylation through the promotion of Src activation. Finally, it is also possible that Src triggers DDR1 phosphorylation by acting on other proteins. For example, through the modulation of phosphatases bound to DDR1 such as SHP-2, a known DDR1 interacting protein (Koo et al, 2006). Alternatively, Src could activate another kinase which then phosphorylates DDR1, such as Crk2, a predicted DDR1 interacting protein (Lemeer et al, 2012). It will be important in future to establish

whether a direct interaction takes place between DDR1 and Src to discount the contribution of other cellular factors. This could be achieved through, for example, FRET (fluorescence resonance energy transfer) analysis of fluorescently tagged DDR1 and Src constructs.

In summary, the importance of the DDR1 JM4 region has been demonstrated in the full-length cell expressed receptor. However, within a cellular context, the JM4 region may have an additional positive role to play in DDR1 activation. Understanding the mechanism through which the JM4 is able to elicit this effect will be an interesting avenue for future research.

Chapter 6. Discussion

Extensive structural and functional characterisation of the extracellular domains of the DDRs has uncovered how this receptor family binds to collagen (Carafoli et al, 2009; Carafoli et al, 2012; Ichikawa et al, 2007; Konitsiotis et al, 2008; Xu et al, 2011). Recent work has also revealed that this binding event leads to DDR clustering and autophosphorylation (Corcoran et al, 2019; Juskaite et al, 2017; Yeung et al, 2019). However, the regulatory mechanisms underpinning DDR activation are still poorly characterised. Through structural and biochemical analysis, this thesis has demonstrated that the kinase proximal DDR1 JM region, JM4, is an autoinhibitory motif. It was shown that DDR1 autophosphorylation occurs in two distinct steps beginning with JM4 phosphorylation, followed by A-loop phosphorylation. These autophosphorylation events have profound activational effects on the kinase catalytic rate. In addition, cellular analysis of JM4 function revealed that it may have an additional positive role in promoting DDR1 activation.

The findings in this thesis led to a model for DDR1 activation which is depicted in **Figure 64**. In the absence of collagen, DDR1 is in a dually autoinhibited state with the A-loop and JM4 region enforcing the inactive conformation of the kinase: αC-helix is in an upward conformation, the DFG motif is in the **D**FG-OUT conformation, and the catalytically essential Lys-Glu salt bridge is broken. Collagen binding to DDR1 triggers clustering of the receptor (Corcoran et al, 2019; Juskaite et al, 2017; Yeung et al, 2019). This enables *cis*-autophosphorylation of the JM4 Tyr residues (Tyr569 and Tyr586; **Figure 64 - Step 1**). Autophosphorylation relieves JM4 autoinhibition and allows *trans*-autophosphorylation of the DDR1 A-loop (Tyr796) (**Figure 64 - Step 2a**). The JM4 phospho-Tyr residues may also act as binding sites for Src which is able to potentiate DDR1 activation (**Figure 64 - Step 2b**). Once the JM4 region and A-loop are phosphorylated, the kinase is able to *trans*-autophosphorylate other kinase and JM Tyr residues (**Figure 64 - Step 3**). These phospho-Tyr residues then act as binding sites for downstream signalling proteins which can activate intracellular signalling pathways (Lemeer et al, 2012; Lemmon & Schlessinger, 2010).

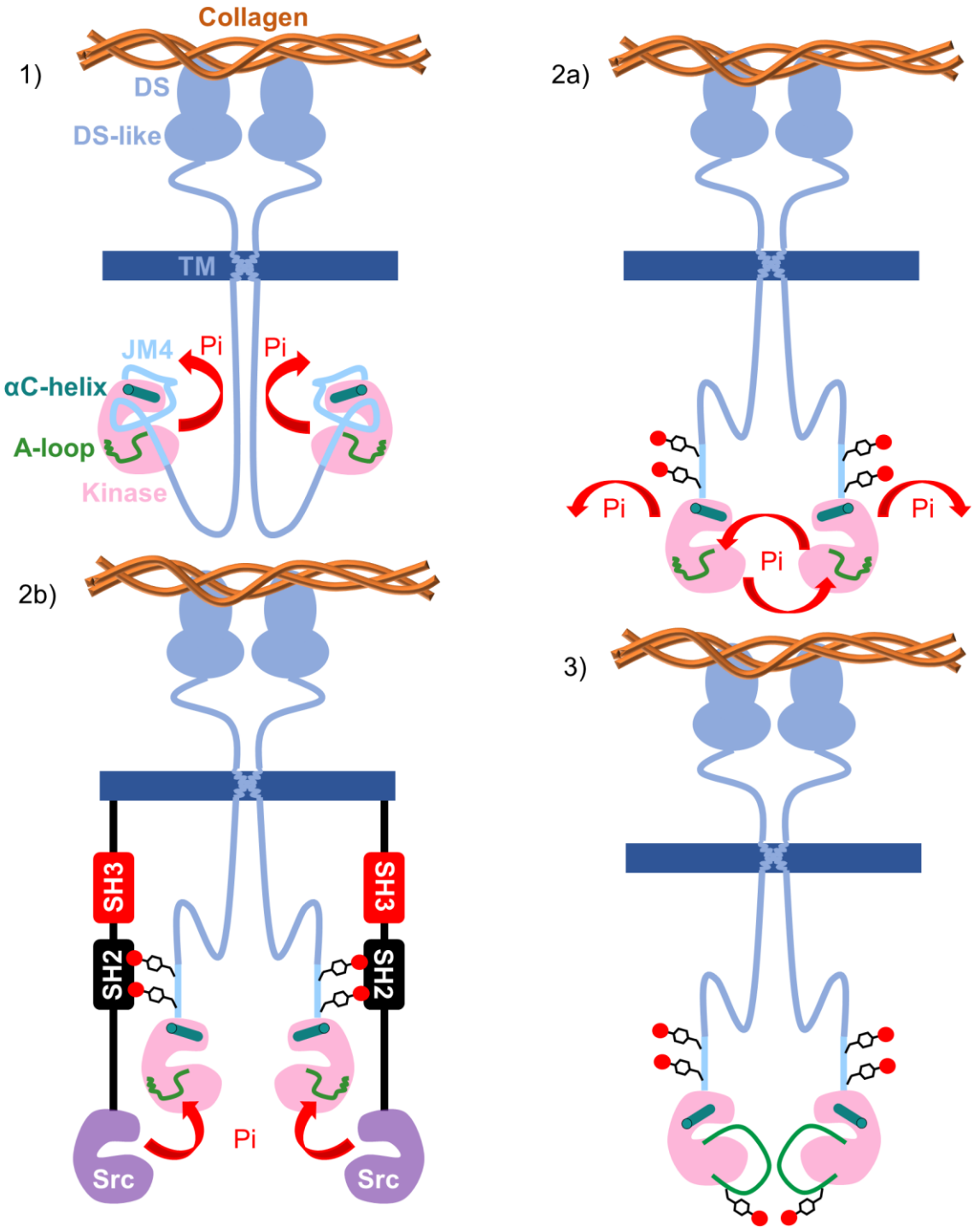

Figure 64. Proposed DDR1 activation mechanism.
1) In the absence of collagen DDR1 is a preformed dimer which is dually autoinhibited by the JM4 region and kinase A-loop. The A-loop forms a short helix and is in the **D**FG-OUT conformation. The JM4 region forms a hairpin within the active site which supports A-loop autoinhibition. JM4 and A-loop support the αC-helix in an upward, catalytically incompetent position. Following collagen stimulation, DDR1 autophosphorylates the JM4 region (Tyr569 and Tyr586) *in cis*. 2a) Following relief of JM4 autoinhibition DDR1 can phosphorylate the A-loop *in trans*. 2b) Phospho-Tyr569 and/or phospho-Tyr586 may also recruit Src via its SH2 domain which then propagates DDR1 phosphorylation. 3) Fully phosphorylated active DDR1.

6.1 JM4 function following phosphorylation

One of the key questions which remains following this thesis is the function of the JM4 region following autophosphorylation. Phosphorylation of RTK JM regions can have three major functions. The first, which is seen in RTKs with autoinhibitory JM regions, is to relieve autoinhibition (Binns et al, 2000; Hubbard, 2004; Mol et al, 2004; Sammon et al, 2020; Till et al, 2002). The second is to provide bindings sites for SH2 and PTB domain containing signalling proteins (Lemmon & Schlessinger, 2010; Trenker & Jura, 2020). Upon binding, the RTK can phosphorylate these proteins and trigger their activation and downstream signalling (Lemmon & Schlessinger, 2010). Alternatively, the recruited proteins may function to propagate receptor activation, as is seen between Src and EGFR (El-Hashim et al, 2017; Lombardo et al, 1995). The DDR1 JM4 Tyr residues, Tyr569 and Tyr586, are predicted Src binding sites following phosphorylation (Lemeer et al, 2012). Furthermore, Src is shown to amplify DDR1 phosphorylation here and mutation of the JM4 Tyr residues abrogates this function (**Figure 62**). The JM4 region may therefore also function as a promoter of DDR1 activation by binding to Src. Finally, RTK JM regions can promote the formation of a catalytically competent kinase configuration. For the insulin and Ret receptors, this positive function is dependent on JM region phosphorylation (Cabail et al, 2015; Plaza-Menacho et al, 2016). However, the EGFR JM region elicits its positive effects independent of its phosphorylation (Jura et al, 2009; Red Brewer et al, 2009). Instead, autophosphorylation of the EGFR C-terminal tail releases the JM region to form intermolecular interactions and promote the formation of an active signalling EGFR dimer (Jura et al, 2009; Red Brewer et al, 2009). Whether the DDR1 JM4 region partakes positively towards DDR1 catalysis is an important question for future analysis. It will be interesting to investigate the conformational changes which the JM4 region undergoes following phosphorylation. This could be achieved through numerous structural analyses. Crystal structure studies of phosphorylated DDR1 were unsuccessful here. However, nuclear magnetic resonance (NMR) analysis could provide interesting insights into the dynamics of the JM4 region post-phosphorylation. In depth structural analysis of the entire receptor in lipid nanodiscs by cryo-electron microscopy, as achieved for the insulin receptor (Weis et al, 2018), could also provide insights into DDR1 activation by analysing the unphosphorylated and phosphorylated receptor.

6.2 Designing novel DDR1 inhibitors

As discussed in **Chapter 3** and **Chapter 4**, targeting the JM4 binding space within the DDR1 kinase domain could offer an interesting avenue for inhibitor studies. The tractability of targeting this space was demonstrated in **Chapter 4** using a JM4 peptide mimetic which effectively inhibited DDR1 substrate phosphorylation (**Figure 42**). However, the affinity of this peptide for DDR1 was likely low, as high concentrations were required for effective inhibition. This is perhaps unsurprising as the JM4 region likely binds transiently within the kinase domain to allow movement of Tyr569 and Tyr586 into a position amenable to *in cis* autophosphorylation. Small molecules may therefore offer a higher affinity alternative to direct JM4 peptide mimetic inhibition of DDR1.

Molecule 5N is an ATP-competitive, DDR1-specific inhibitor which occupies much of the same binding space as the promiscuous kinase inhibitor imatinib (**Figure 22**) (Canning et al, 2014; Wang et al, 2018). However, unlike imatinib, 5N contains a 5-trifluoromethyl-phenyl, and a methylpiperazine, moiety which sit close to the Phe586 and Thr585 JM4 binding space, respectively (Wang et al, 2018). This portion of molecule 5N could feasibly be used as a starting point for the discovery of inhibitors which occupy the JM4 binding space. A truncated molecule 5N could be generated which contains the JM4 space occupying moieties. Extensions of this truncated molecule could then be screened for their inhibitory potential against DDR1. It will be important to use a screen which can distinguish between compounds which are directly ATP-competitive and those which occupy the JM4-binding space. This could be achieved through a FRET-based assay which monitors JM4 displacement from the active site (fluorophore labelled JM4 region and kinase domain). Similar studies have been conducted on the A-loop of Aurora kinase which provided insights into A-loop dynamics (Gilburt et al, 2017). Hits from this screen could then be analysed using the ADP-Glo^{TM} assay optimised in this thesis (**Figure 30**) and inhibitor dissociation constants evaluated. Promising compounds could then be analysed by X-ray crystallography to evaluate the binding space they occupy. Alternatively, *de novo* drug discovery could be attempted by initially screening the JM4 binding space computationally for potential compound interactions. Hits from this could be screened as above. Assessing the effectiveness of these inhibitors within the context of

disease models will be crucial to validate not only their efficacy but also the benefits of specifically targeting DDR1 in disease.

The literature is at times contradictory when it comes to the benefits of targeting DDR1, particularly in cancer therapy. For example, there are several studies which have demonstrated the link between DDR1 expression levels and poor prognosis and enhanced metastasis in cancers, including NSCLC and breast cancer (Gadiya & Chakraborty, 2018; Yang et al, 2010; Zhong et al, 2019). Contrastingly however, DDR1 expression suppressed breast tumour growth when expressed in tumour surrounding tissue (Takai et al, 2018). These cell-type dependent, contrasting effects make it difficult to predict the systemic results of DDR1 inhibition in cancer treatment. On the other hand, the link between DDR2 and poor cancer prognosis is more extensive, particularly in breast cancer (Bayer et al, 2019; Grither & Longmore, 2018; Ren et al, 2014). As the JM4 region is strictly conserved between DDR1 and DDR2 (**Figure 6**) it is likely that the JM4 region functions in the same way in both receptors. The development of JM4 region occupying molecules could therefore provide a pan-DDR tool which may be beneficial in cancer therapy. Indeed, inhibition of DDR2 is effective in the treatment of several cancer disease models including melanoma, breast cancer, and NSCLC (Grither & Longmore, 2018; Poudel et al, 2015; Xu et al, 2015). Furthermore, recent investigation into the effects of co-inhibition of immune checkpoints (anti-PD1) and DDR2, have, demonstrated promising *in vivo* effects in the treatment of multiple cancers (Tu et al, 2019).

Fibrosis may offer a more favourable therapeutic DDR1 implication than cancer (Moll et al, 2019). There is extensive pre-clinical DDR1 target validation across multiple organs, including, lung, kidney, and liver (Avivi-Green et al, 2006; Kerroch et al, 2016; Moll et al, 2019; Moll et al, 2018; Schuppan et al, 2018; Song et al, 2011). Furthermore, a DDR1-specific inhibitor has been developed which showed promise in reducing kidney fibrotic markers *in vitro* and *in vivo* (Richter et al, 2019). A similar positive effect has been elicited by inhibiting DDR1 in pulmonary and hepatic fibrotic models (Tao et al, 2018; Wang et al, 2016). While targeting DDR1 activity is attractive, DDR1 is widely expressed in the body, and has been implicated in some vital homeostatic processes (Di Marco et al, 1993; Hou et al, 2001; Laval et al, 1994; Perez et al, 1996). The generation of highly specific DDR1 inhibitors will be

invaluable in improving our understanding of the tractability of targeting DDR1 and the long-term effects of doing so in humans.

6.3 Key remaining questions in the DDR field

There are also several general questions which remain in the DDR field, including:

What function does DDR1 have in adult tissue? DDR1 has a clear function during development, yet in adult tissue has few known functions other than implications in tissue homeostasis and matrix remodelling (Borza et al, 2017; Gross et al, 2004; Meyer zum Gottesberge et al, 2008; Vogel et al, 2001). The broad tissue expression of DDR1 implies that it has a role in multiple organs (Di Marco et al, 1993; Leitinger, 2014). However, what these roles are is currently unclear (Leitinger, 2014). The *in vivo* tissue-specific knockdown of DDR1 in adult mice could provide insights into these functions.

How is DDR1 maintained in an inactive state in cells? Epithelial cells expressing DDR1 are constitutively exposed to collagen, the DDR1 activating ligand (Di Marco et al, 1993; Leitinger, 2014). How cells keep DDR1 in an inactive state in the constitutive presence of collagen is therefore an important question. It is possible that DDR1 is sequestered at intercellular epithelial junctions (Wang et al, 2009). Damage to the epithelial layer during wound formation, for example, would then expose DDR1 to collagen and trigger activation. It is also possible that collagen fibrils are not DDR activating ligands, as the DDR binding site may not be easily accessible within these fibrils (Carafoli et al, 2009; Itoh, 2018; Orgel et al, 2006). Instead, damage to collagen and the generation of triple helical collagen peptides may be the trigger for DDR activation (Itoh, 2018). Alternatively, it is possible that DDR1 is constitutively activated in some tissues. In this scenario, DDR1 signalling could be regulated through alternative mechanisms. For example, MMP mediated cleavage of the DDR1 ectodomain could provide a method of suppressing DDR1 signalling (Alves et al, 1995; Fu et al, 2013). Subsequent alterations in the levels of extracellular MMPs could then allow prolonged DDR1 activation and signalling.

What are the functional consequences of ligand induced DDR1 clustering? Clustering of DDR1 following collagen stimulation is thought to trigger receptor

activation (Corcoran et al, 2019; Juskaite et al, 2017; Yeung et al, 2019). However, several factors are still poorly understood about this process. For example, the density of receptor which is required for DDR1 activation is currently unclear. So too is the stoichiometry and structure of these clusters. Super-resolution imaging and cryo-electron microscopy analysis of DDR1 activation will be vital tools in unravelling these questions. Another key question is why such high receptor density is required for activation. In typical RTKs, dimerisation activates the receptor through increasing the local concentration of the intracellular kinase domains, thereby facilitating *trans*-autophosphorylation (Lemmon & Schlessinger, 2010; Trenker & Jura, 2020). DDR1 on the other hand requires the dense clustering of numerous receptor dimers in order to activate (Corcoran et al, 2019). It is possible that the tightly packed environment created by clusters excludes the binding of inhibitory proteins to the intracellular domains of DDR1. Expulsion of proteins such as phosphatases through steric clashes could then lead to DDR1 autophosphorylation. A similar mechanism is used by the T-cell receptor wherein the inhibitory phosphatase, CD45, is excluded from T-cell receptor clusters through later steric exclusion (James & Vale, 2012).

In conclusion, this thesis has provided important insights into DDR1 regulation, highlighting the initial steps in DDR1 activation and the effects these events have on catalysis. Future work should focus on: (1) confirming the tractability of pharmacologically targeting DDR1 and DDR2; (2) the steps which follow relief of DDR1 autoinhibition with a focus on the positive function of the JM4 region; (3) the structural and functional consequences of receptor clustering. Improved understanding in these areas will be vital if the DDRs are to be used as effective therapeutic target.

Chapter 7. Appendix

Table 18. JM4-1P chymotrypsin tandem MS data.
The data provided by the CCP when 1 µg JM4-1P was separated by SDS-PAGE extracted from the gel, cleaved with chymotrypsin, and then analysed by tandem MS is shown.

JM4-1P – Chymotrypsin tandem MS data											
Query	Observed	Mr (expt)	Mr (calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide	
702	460.7748	919.5351	919.5352	-0.05	0	21	0.0087	1	U	L.NVRKGHPL.L	
829	473.2481	944.4816	944.4815	0.16	2	24	0.0041	1	U	F.LAEDALNTV.-	
950	484.7981	967.5816	967.5814	0.22	0	(32)	0.00069	1	U	L.KDPNIIRL.L	
951	484.7983	967.5821	967.5814	0.65	0	33	0.0005	1	U	L.KDPNIIRL.L	
1227	511.7493	1021.4840	1021.4837	0.31	1	21	0.0082	1	U	L.MLCRAQPF.G	
1296	517.3166	1032.6186	1032.6192	-0.58	1	25	0.0035	1	U	L.NVRKGHPLL.V	
1303	518.2772	1034.5398	1034.5396	0.12	1	(53)	5.3e-06	1	U	F.KEKLGEGQF.G	
1304	518.2773	1034.5400	1034.5396	0.35	1	54	4e-06	1	U	F.KEKLGEGQF.G	
1417	531.7243	1061.4340	1061.4335	0.47	0	23	0.0045	1	U	L.CEVDSPQDL.V	
1461	536.2830	1070.5514	1070.5509	0.46	2	38	0.00015	1	U	L.HRFLAEDAL.N	
1471	537.2472	1072.4798	1072.4825	-2.53	1	21	0.0078	1	U	L.GEGQFGEVHL.C + Deamidated (NQ)	
1509	541.3400	1080.6655	1080.6655	0.01	1	(41)	7.7e-05	1	U	L.KDPNIIRLL.G	
1510	541.3403	1080.6660	1080.6655	0.45	1	58	1.4e-06	1	U	L.KDPNIIRLL.G	
1512	541.7744	1081.5341	1081.5339	0.27	0	34	0.0004	1	U	L.SRPPACPQGL.Y	
1514	542.2241	1082.4337	1082.4339	-0.17	1	(34)	0.00036	1	U	Y.MENGDLNQF.L + Oxidation (M)	
1515	542.2242	1082.4339	1082.4339	0.06	1	(28)	0.0017	1	U	Y.MENGDLNQF.L + Oxidation (M)	
1520	542.7161	1083.4176	1083.4179	-0.29	1	43	5.3e-05	1	U	Y.MENGDLNQF.L + Deamidated (NQ); Oxidation (M)	
1713	560.3316	1118.6486	1118.6481	0.46	0	44	3.9e-05	1	U	L.KEVKIMSRL.K + Oxidation (M)	
1947	584.7964	1167.5782	1167.5785	-0.24	1	(26)	0.0025	1	U	F.FRDQGRQVY.L	
1948	390.2001	1167.5786	1167.5785	0.05	1	35	0.0003	1	U	F.FRDQGRQVY.L	
1949	390.2001	1167.5786	1167.5785	0.05	1	(35)	0.00035	1	U	F.FRDQGRQVY.L	
1950	390.2004	1167.5794	1167.5785	0.77	1	(31)	0.00084	1	U	F.FRDQGRQVY.L	
2026	396.2119	1185.6140	1185.6142	-0.19	2	33	0.00053	1	U	L.ATLNFVHRDL.A + Deamidated (NQ)	
2079	598.3162	1194.6179	1194.6179	-0.03	1	27	0.0021	1	U	Y.LSRPPACPQGL.Y	

2280	616.8041	1231.5936	1231.5945	-0.75	0	(29)	0.0012	1	U	W.SRESEQRPPF.S	
2282	411.5386	1231.5940	1231.5945	-0.45	0	(26)	0.0026	1	U	W.SRESEQRPPF.S	
2283	411.5387	1231.5944	1231.5945	-0.13	0	(33)	0.00052	1	U	W.SRESEQRPPF.S	
2285	411.5389	1231.5948	1231.5945	0.23	0	(32)	0.00062	1	U	W.SRESEQRPPF.S	
2286	411.5389	1231.5950	1231.5945	0.38	0	35	0.00031	1	U	W.SRESEQRPPF.S	
2287	616.8049	1231.5952	1231.5945	0.54	0	(31)	0.00072	1	U	W.SRESEQRPPF.S	
2288	616.8050	1231.5954	1231.5945	0.73	0	(25)	0.003	1	U	W.SRESEQRPPF.S	
2289	411.5406	1231.5999	1231.5945	4.40	0	(29)	0.0013	1	U	W.SRESEQRPPF.S	
2293	616.8734	1231.7323	1231.7322	0.06	1	42	6.3e-05	1	U	F.LKEVKIMSRL.K + Oxidation (M)	
2294	411.8659	1232.5759	1232.5785	-2.14	0	(33)	0.00053	1	U	W.SRESEQRPPF.S + Deamidated (NQ)	
2366	623.3062	1244.5979	1244.5972	0.55	1	20	0.0097	1	U	L.SRPPACPQGLY.E	
2542	640.8071	1279.5996	1279.5979	1.31	1	42	6e-05	1	U	L.ATRNCLVGENF.T	
2548	641.2991	1280.5836	1280.5819	1.30	1	(41)	8.6e-05	1	U	L.ATRNCLVGENF.T + Deamidated (NQ)	
2549	641.2998	1280.5850	1280.5819	2.44	1	(22)	0.0069	1	U	L.ATRNCLVGENF.T + Deamidated (NQ)	
2551	427.8941	1280.6605	1280.6626	-1.64	2	29	0.0014	1	U	F.FRDQGRQVYL.S	
2743	440.5556	1318.6451	1318.6452	-0.09	0	(33)	0.00052	1	U	L.HVAAQIASGMRY.L + Oxidation (M)	
2748	660.8220	1319.6294	1319.6292	0.13	0	38	0.00016	1	U	L.HVAAQIASGMRY.L + Deamidated (NQ); Oxidation (M)	
2826	669.8610	1337.7074	1337.7092	-1.34	2	(46)	2.3e-05	1	U	L.RFKEKLGEGQF.G	
2827	446.9104	1337.7093	1337.7092	0.09	2	51	8.2e-06	1	U	L.RFKEKLGEGQF.G	
3075	468.9037	1403.6892	1403.6906	-1.00	0	(28)	0.0017	1	U	L.RPDATKNARNDF.L	
3078	468.9047	1403.6922	1403.6906	1.14	0	(25)	0.0034	1	U	L.RPDATKNARNDF.L	
3083	469.2325	1404.6757	1404.6746	0.81	0	38	0.00017	1	U	L.RPDATKNARNDF.L + Deamidated (NQ)	
3084	469.2335	1404.6788	1404.6746	3.03	0	(35)	0.00033	1	U	L.RPDATKNARNDF.L + Deamidated (NQ)	
3182	478.2507	1431.7303	1431.7292	0.72	1	32	0.00056	1	U	L.HVAAQIASGMRYL.A + Oxidation (M)	
3365	741.3859	1480.7572	1480.7562	0.66	1	(37)	0.00021	1	U	-.GPVPHYAEADIVTL.Q	
3366	741.3860	1480.7574	1480.7562	0.83	1	(44)	4.1e-05	1	U	-.GPVPHYAEADIVTL.Q	
3460	506.5989	1516.7747	1516.7746	0.07	1	35	0.00034	1	U	L.RPDATKNARNDFL.K	
3464	506.9269	1517.7588	1517.7586	0.13	1	(29)	0.0012	1	U	L.RPDATKNARNDFL.K + Deamidated (NQ)	
3465	506.9270	1517.7593	1517.7586	0.42	1	(20)	0.0091	1	U	L.RPDATKNARNDFL.K + Deamidated (NQ)	
3547	780.8911	1559.7677	1559.7692	-0.98	1	(23)	0.0053	1	U	W.SRESEQRPPFSQL.H	
3548	520.9306	1559.7700	1559.7692	0.50	1	(22)	0.0064	1	U	W.SRESEQRPPFSQL.H	

3551	520.9313	1559.7720	1559.7692	1.79	1	(21)	0.0086	1	U	W.SRESEQRPPFSQL.H	
3553	781.3688	1560.7231	1560.7225	0.39	1	(40)	0.00011	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3554	781.3690	1560.7235	1560.7225	0.62	1	(45)	3.2e-05	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3555	781.3691	1560.7236	1560.7225	0.70	1	47	1.8e-05	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3556	781.3691	1560.7236	1560.7225	0.70	1	(38)	0.00016	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3557	781.3691	1560.7236	1560.7225	0.70	1	(42)	5.7e-05	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3558	781.3692	1560.7238	1560.7225	0.85	1	(23)	0.0047	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3559	781.3692	1560.7238	1560.7225	0.85	1	(39)	0.00013	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3560	781.3693	1560.7240	1560.7225	0.93	1	(42)	6.2e-05	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3561	781.3693	1560.7241	1560.7225	1.01	1	(26)	0.0023	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3562	781.3694	1560.7242	1560.7225	1.08	1	(34)	0.00039	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3563	781.3695	1560.7243	1560.7225	1.17	1	(33)	0.00053	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3565	781.3696	1560.7246	1560.7225	1.33	1	(24)	0.0042	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3566	781.3696	1560.7246	1560.7225	1.33	1	(33)	0.00049	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3567	781.3698	1560.7251	1560.7225	1.63	1	(38)	0.00015	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3568	781.3703	1560.7260	1560.7225	2.26	1	(38)	0.00017	1	U	-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)	pY569
3570	781.3864	1560.7583	1560.7532	3.24	1	34	0.00037	1	U	W.SRESEQRPPFSQL.H + Deamidated (NQ)	
3595	524.2796	1569.8170	1569.8151	1.20	2	57	2e-06	1	U	F.KEKLGEGQFGEVHL.C	
3660	799.3626	1596.7107	1596.7090	1.04	1	53	5.5e-06	1	U	F.GEVHLCEVDSPQDL.V	
3803	825.3873	1648.7600	1648.7580	1.18	1	26	0.0023	1	U	F.GQLTDEQVIENAGEF.F	
5541	1047.1724	3138.4953	3138.4928	0.79	2	(53)	5.3e-06	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPML.L.H + Oxidation (M)	
5542	1047.1727	3138.4964	3138.4928	1.15	2	(64)	4.4e-07	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPML.L.H + Oxidation (M)	
5543	1047.1727	3138.4964	3138.4928	1.15	2	(64)	4.4e-07	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPML.L.H + Oxidation (M)	
5544	1047.1728	3138.4967	3138.4928	1.26	2	(53)	5.2e-06	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPML.L.H + Oxidation (M)	
5546	1047.1732	3138.4978	3138.4928	1.61	2	64	4.3e-07	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPML.L.H + Oxidation (M)	
5548	785.8815	3139.4968	3139.4768	6.37	2	(38)	0.00017	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPML.L.H + Deamidated (NQ); Oxidation (M)	
5579	1073.8175	3218.4307	3218.4591	-8.83	2	26	0.0024	1	U	F.LSAHQLEDKAAEGAPGDGQAAQGPTISYPML.L + Oxidation (M); Phospho_STY (STY)	

Table 19. JM4-1P trypsin tandem MS data.
The data provided by the CCP when 1 µg JM4-1P was separated by SDS-PAGE extracted from the gel, cleaved with trypsin, and then analysed by tandem MS is shown.

JM4K-IP - Trypsin tandem MS data										
Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide
754	448.7172	895.4199	895.4222	-2.56	0	(21)	0.0084	1	U	K.IADFGMSR.N
755	448.7172	895.4199	895.4222	-2.56	0	(24)	0.004	1	U	K.IADFGMSR.N
760	448.7181	895.4215	895.4222	-0.71	0	(22)	0.0063	1	U	K.IADFGMSR.N
761	448.7181	895.4216	895.4222	-0.64	0	(22)	0.0066	1	U	K.IADFGMSR.N
763	448.7181	895.4217	895.4222	-0.50	0	(23)	0.0052	1	U	K.IADFGMSR.N
765	448.7182	895.4218	895.4222	-0.44	0	(22)	0.0064	1	U	K.IADFGMSR.N
766	448.7182	895.4218	895.4222	-0.44	0	(25)	0.0029	1	U	K.IADFGMSR.N
768	448.7182	895.4218	895.4222	-0.37	0	41	7.9e-05	1	U	K.IADFGMSR.N
770	448.7183	895.4221	895.4222	-0.10	0	(22)	0.0061	1	U	K.IADFGMSR.N
771	448.7183	895.4221	895.4222	-0.04	0	(25)	0.0031	1	U	K.IADFGMSR.N
772	448.7184	895.4222	895.4222	0.03	0	(33)	0.00052	1	U	K.IADFGMSR.N
774	448.7184	895.4223	895.4222	0.10	0	(25)	0.0035	1	U	K.IADFGMSR.N
854	456.7138	911.4130	911.4171	-4.49	0	(31)	0.00073	1	U	K.IADFGMSR.N + Oxidation (M)
856	456.7150	911.4155	911.4171	-1.75	0	(35)	0.00031	1	U	K.IADFGMSR.N + Oxidation (M)
857	456.7151	911.4156	911.4171	-1.61	0	(31)	0.0008	1	U	K.IADFGMSR.N + Oxidation (M)
871	456.7157	911.4168	911.4171	-0.34	0	(20)	0.0092	1	U	K.IADFGMSR.N + Oxidation (M)
872	456.7157	911.4168	911.4171	-0.28	0	(35)	0.00032	1	U	K.IADFGMSR.N + Oxidation (M)
875	456.7158	911.4170	911.4171	-0.14	0	(35)	0.00031	1	U	K.IADFGMSR.N + Oxidation (M)
876	456.7158	911.4170	911.4171	-0.08	0	(20)	0.009	1	U	K.IADFGMSR.N + Oxidation (M)
878	456.7159	911.4172	911.4171	0.12	0	(32)	0.00062	1	U	K.IADFGMSR.N + Oxidation (M)
880	456.7160	911.4174	911.4171	0.38	0	(32)	0.00065	1	U	K.IADFGMSR.N + Oxidation (M)
882	456.7161	911.4176	911.4171	0.58	0	(35)	0.00031	1	U	K.IADFGMSR.N + Oxidation (M)
885	456.7163	911.4180	911.4171	1.00	0	(33)	0.00049	1	U	K.IADFGMSR.N + Oxidation (M)
886	456.7163	911.4181	911.4171	1.06	0	(32)	0.00068	1	U	K.IADFGMSR.N + Oxidation (M)
887	456.7163	911.4181	911.4171	1.06	0	(32)	0.00067	1	U	K.IADFGMSR.N + Oxidation (M)
888	456.7163	911.4181	911.4171	1.06	0	(32)	0.00069	1	U	K.IADFGMSR.N + Oxidation (M)
897	457.2762	912.5379	912.5392	-1.51	0	(29)	0.0012	1	U	K.ILRPDATK.N
907	457.2765	912.5384	912.5392	-0.96	0	(21)	0.0073	1	U	K.ILRPDATK.N
910	457.2765	912.5385	912.5392	-0.76	0	31	0.00088	1	U	K.ILRPDATK.N
912	457.2766	912.5387	912.5392	-0.63	0	(24)	0.0039	1	U	K.ILRPDATK.N
914	457.2767	912.5389	912.5392	-0.37	0	(30)	0.00094	1	U	K.ILRPDATK.N
916	457.2768	912.5390	912.5392	-0.24	0	(30)	0.00092	1	U	K.ILRPDATK.N
917	457.2768	912.5391	912.5392	-0.17	0	(31)	0.00088	1	U	K.ILRPDATK.N
919	457.2769	912.5393	912.5392	0.05	0	(25)	0.0032	1	U	K.ILRPDATK.N
920	457.2772	912.5398	912.5392	0.57	0	(30)	0.0011	1	U	K.ILRPDATK.N
923	457.2774	912.5403	912.5392	1.19	0	(26)	0.0024	1	U	K.ILRPDATK.N
1326	484.7965	967.5785	967.5814	-3.07	1	(26)	0.0025	1	U	R.LKDPNIIR.L

1333	484.7975	967.5803	967.5814	-1.12	1	(26)	0.0024	1	U	R.LKDPNIIR.L
1345	484.7978	967.5811	967.5814	-0.30	1	(26)	0.0027	1	U	R.LKDPNIIR.L
1348	484.7979	967.5812	967.5814	-0.24	1	27	0.0022	1	U	R.LKDPNIIR.L
1352	484.7980	967.5814	967.5814	0.01	1	(26)	0.0027	1	U	R.LKDPNIIR.L
1360	484.7981	967.5817	967.5814	0.28	1	(26)	0.0024	1	U	R.LKDPNIIR.L
1370	485.2890	968.5634	968.5654	-2.06	1	(23)	0.005	1	U	R.LKDPNIIR.L + Deamidated (NQ)
1586	496.7743	991.5340	991.5338	0.20	1	35	0.00032	1	U	R.NDFLKEVK.I
2035	531.3458	1060.6771	1060.6757	1.34	1	22	0.0062	1	U	R.KGHPLLVAVK.I
2213	546.7806	1091.5466	1091.5499	-3.00	0	(52)	5.7e-06	1	U	R.FLAEDALNTV.-
2215	546.7812	1091.5478	1091.5499	-1.88	0	(34)	0.00043	1	U	R.FLAEDALNTV.-
2216	546.7816	1091.5486	1091.5499	-1.20	0	70	1.1e-07	1	U	R.FLAEDALNTV.-
2217	546.7818	1091.5490	1091.5499	-0.76	0	(32)	0.00066	1	U	R.FLAEDALNTV.-
2218	546.7818	1091.5490	1091.5499	-0.76	0	(23)	0.0046	1	U	R.FLAEDALNTV.-
2220	546.7819	1091.5492	1091.5499	-0.65	0	(24)	0.0038	1	U	R.FLAEDALNTV.-
2222	546.7819	1091.5493	1091.5499	-0.54	0	(22)	0.0063	1	U	R.FLAEDALNTV.-
2223	546.7820	1091.5494	1091.5499	-0.43	0	(28)	0.0018	1	U	R.FLAEDALNTV.-
2224	546.7820	1091.5494	1091.5499	-0.43	0	(22)	0.0059	1	U	R.FLAEDALNTV.-
2226	546.7820	1091.5494	1091.5499	-0.43	0	(26)	0.0028	1	U	R.FLAEDALNTV.-
2227	546.7821	1091.5496	1091.5499	-0.21	0	(46)	2.6e-05	1	U	R.FLAEDALNTV.-
2228	546.7821	1091.5496	1091.5499	-0.21	0	(36)	0.00027	1	U	R.FLAEDALNTV.-
2229	546.7822	1091.5499	1091.5499	0.02	0	(42)	6.1e-05	1	U	R.FLAEDALNTV.-
2230	546.7822	1091.5499	1091.5499	0.02	0	(32)	0.00063	1	U	R.FLAEDALNTV.-
2231	546.7822	1091.5499	1091.5499	0.02	0	(37)	0.00022	1	U	R.FLAEDALNTV.-
2232	546.7823	1091.5500	1091.5499	0.13	0	(23)	0.0054	1	U	R.FLAEDALNTV.-
2233	546.7823	1091.5501	1091.5499	0.24	0	(29)	0.0012	1	U	R.FLAEDALNTV.-
2234	546.7825	1091.5504	1091.5499	0.46	0	(38)	0.00018	1	U	R.FLAEDALNTV.-
2235	546.7825	1091.5505	1091.5499	0.57	0	(41)	7.9e-05	1	U	R.FLAEDALNTV.-
2236	546.7825	1091.5505	1091.5499	0.57	0	(31)	0.00074	1	U	R.FLAEDALNTV.-
2237	546.7825	1091.5505	1091.5499	0.57	0	(23)	0.005	1	U	R.FLAEDALNTV.-
2238	546.7825	1091.5505	1091.5499	0.57	0	(23)	0.0053	1	U	R.FLAEDALNTV.-
2239	546.7827	1091.5509	1091.5499	0.90	0	(33)	0.00056	1	U	R.FLAEDALNTV.-
2240	546.7827	1091.5509	1091.5499	0.90	0	(27)	0.0019	1	U	R.FLAEDALNTV.-
2241	546.7827	1091.5509	1091.5499	0.90	0	(26)	0.0027	1	U	R.FLAEDALNTV.-
2242	546.7829	1091.5512	1091.5499	1.25	0	(31)	0.0008	1	U	R.FLAEDALNTV.-
2243	546.7829	1091.5512	1091.5499	1.25	0	(27)	0.0022	1	U	R.FLAEDALNTV.-
2244	546.7830	1091.5515	1091.5499	1.47	0	(35)	0.00034	1	U	R.FLAEDALNTV.-
2252	547.2747	1092.5349	1092.5339	0.91	0	(45)	3e-05	1	U	R.FLAEDALNTV.- + Deamidated (NQ)
2253	547.2751	1092.5356	1092.5339	1.59	0	(54)	4e-06	1	U	R.FLAEDALNTV.- + Deamidated (NQ)
2254	547.2780	1092.5415	1092.5339	6.95	0	(35)	0.0003	1	U	R.FLAEDALNTV.- + Deamidated (NQ)
2562	567.7625	1133.5105	1133.5141	-3.24	0	(22)	0.0069	1	U	R.NLYAGDYYR.V
2579	567.7639	1133.5133	1133.5141	-0.76	0	(28)	0.0017	1	U	R.NLYAGDYYR.V
2583	567.7644	1133.5142	1133.5141	0.09	0	(21)	0.0084	1	U	R.NLYAGDYYR.V
2584	567.7645	1133.5144	1133.5141	0.21	0	(30)	0.0011	1	U	R.NLYAGDYYR.V
2587	567.7645	1133.5145	1133.5141	0.32	0	(28)	0.0016	1	U	R.NLYAGDYYR.V
2588	567.7646	1133.5146	1133.5141	0.43	0	(30)	0.001	1	U	R.NLYAGDYYR.V
2589	567.7646	1133.5146	1133.5141	0.43	0	(25)	0.003	1	U	R.NLYAGDYYR.V

2590	567.7646	1133.5146	1133.5141	0.43	0	(20)	0.0099	1	U	R.NLYAGDYYR.V
2594	567.7647	1133.5149	1133.5141	0.64	0	(24)	0.0038	1	U	R.NLYAGDYYR.V
2603	567.7650	1133.5155	1133.5141	1.17	0	(26)	0.0024	1	U	R.NLYAGDYYR.V
2604	567.7650	1133.5155	1133.5141	1.17	0	(22)	0.0065	1	U	R.NLYAGDYYR.V
2616	568.2587	1134.5029	1134.4982	4.19	0	32	0.00058	1	U	R.NLYAGDYYR.V + Deamidated (NQ)
3401	617.3398	1232.6650	1232.6666	-1.28	0	(22)	0.0071	1	U	R.YLATLNFVHR.D
3402	617.3398	1232.6650	1232.6666	-1.28	0	(21)	0.0082	1	U	R.YLATLNFVHR.D
3407	617.3399	1232.6652	1232.6666	-1.08	0	(39)	0.00014	1	U	R.YLATLNFVHR.D
3408	617.3399	1232.6652	1232.6666	-1.08	0	(21)	0.0081	1	U	R.YLATLNFVHR.D
3409	617.3400	1232.6654	1232.6666	-0.97	0	(21)	0.008	1	U	R.YLATLNFVHR.D
3410	617.3400	1232.6655	1232.6666	-0.87	0	(23)	0.0047	1	U	R.YLATLNFVHR.D
3411	617.3400	1232.6655	1232.6666	-0.87	0	(23)	0.0054	1	U	R.YLATLNFVHR.D
3420	411.8961	1232.6664	1232.6666	-0.11	0	(21)	0.0071	1	U	R.YLATLNFVHR.D
3428	617.3410	1232.6674	1232.6666	0.70	0	(29)	0.0013	1	U	R.YLATLNFVHR.D
3445	617.3421	1232.6696	1232.6666	2.49	0	39	0.00013	1	U	R.YLATLNFVHR.D
3964	647.8234	1293.6323	1293.6387	-4.96	0	(22)	0.0069	1	U	R.NCLVGENFTIK.I
3965	647.8247	1293.6347	1293.6387	-3.07	0	(22)	0.0061	1	U	R.NCLVGENFTIK.I
3967	647.8254	1293.6362	1293.6387	-1.94	0	(23)	0.0056	1	U	R.NCLVGENFTIK.I
3968	647.8255	1293.6364	1293.6387	-1.75	0	(31)	0.00083	1	U	R.NCLVGENFTIK.I
3970	647.8257	1293.6369	1293.6387	-1.38	0	(22)	0.0071	1	U	R.NCLVGENFTIK.I
3971	647.8258	1293.6371	1293.6387	-1.28	0	(31)	0.00075	1	U	R.NCLVGENFTIK.I
3972	647.8260	1293.6374	1293.6387	-1.00	0	(30)	0.00096	1	U	R.NCLVGENFTIK.I
3973	647.8261	1293.6375	1293.6387	-0.90	0	(31)	0.00087	1	U	R.NCLVGENFTIK.I
3976	647.8262	1293.6379	1293.6387	-0.63	0	(33)	0.00048	1	U	R.NCLVGENFTIK.I
3978	647.8264	1293.6381	1293.6387	-0.44	0	(36)	0.00023	1	U	R.NCLVGENFTIK.I
3981	647.8266	1293.6386	1293.6387	-0.05	0	(29)	0.0011	1	U	R.NCLVGENFTIK.I
3982	647.8267	1293.6389	1293.6387	0.13	0	(22)	0.0071	1	U	R.NCLVGENFTIK.I
3983	647.8268	1293.6390	1293.6387	0.22	0	(26)	0.0025	1	U	R.NCLVGENFTIK.I
3984	647.8268	1293.6391	1293.6387	0.32	0	(20)	0.0096	1	U	R.NCLVGENFTIK.I
3985	647.8269	1293.6392	1293.6387	0.41	0	(23)	0.0056	1	U	R.NCLVGENFTIK.I
3986	647.8269	1293.6392	1293.6387	0.41	0	(35)	0.00031	1	U	R.NCLVGENFTIK.I
3988	647.8270	1293.6394	1293.6387	0.52	0	(22)	0.0067	1	U	R.NCLVGENFTIK.I
3989	647.8270	1293.6395	1293.6387	0.61	0	(23)	0.0054	1	U	R.NCLVGENFTIK.I
3990	647.8271	1293.6396	1293.6387	0.70	0	48	1.5e-05	1	U	R.NCLVGENFTIK.I
3992	647.8271	1293.6396	1293.6387	0.70	0	(26)	0.0023	1	U	R.NCLVGENFTIK.I
3993	647.8273	1293.6400	1293.6387	0.98	0	(21)	0.0074	1	U	R.NCLVGENFTIK.I
3994	647.8273	1293.6401	1293.6387	1.07	0	(39)	0.00012	1	U	R.NCLVGENFTIK.I
3995	647.8274	1293.6402	1293.6387	1.17	0	(39)	0.00013	1	U	R.NCLVGENFTIK.I
3996	647.8274	1293.6403	1293.6387	1.26	0	(31)	0.00088	1	U	R.NCLVGENFTIK.I
3997	647.8275	1293.6405	1293.6387	1.35	0	(21)	0.0075	1	U	R.NCLVGENFTIK.I
3998	647.8276	1293.6406	1293.6387	1.46	0	(32)	0.00059	1	U	R.NCLVGENFTIK.I
3999	647.8278	1293.6410	1293.6387	1.74	0	(29)	0.0012	1	U	R.NCLVGENFTIK.I
4000	647.8279	1293.6412	1293.6387	1.92	0	(25)	0.0034	1	U	R.NCLVGENFTIK.I
4001	647.8279	1293.6412	1293.6387	1.92	0	(25)	0.0034	1	U	R.NCLVGENFTIK.I
4002	647.8279	1293.6413	1293.6387	2.02	0	(33)	0.0005	1	U	R.NCLVGENFTIK.I
4003	647.8282	1293.6418	1293.6387	2.40	0	(21)	0.0087	1	U	R.NCLVGENFTIK.I

4008	648.3187	1294.6229	1294.6227	0.14	0	(29)	0.0013	1	U	R.NCLVGENFTIK.I + Deamidated (NQ)
4009	648.3190	1294.6234	1294.6227	0.51	0	(32)	0.0006	1	U	R.NCLVGENFTIK.I + Deamidated (NQ)
4010	648.3190	1294.6235	1294.6227	0.60	0	(43)	5.3e-05	1	U	R.NCLVGENFTIK.I + Deamidated (NQ)
4018	648.8091	1295.6037	1295.6067	-2.32	0	(31)	0.00074	1	U	R.NCLVGENFTIK.I + 2 Deamidated (NQ)
4019	648.8097	1295.6048	1295.6067	-1.47	0	(31)	0.00073	1	U	R.NCLVGENFTIK.I + 2 Deamidated (NQ)
4724	712.8297	1423.6447	1423.6451	-0.21	0	(42)	6.5e-05	1	U	R.WMAWECILMGK.F
4725	712.8298	1423.6450	1423.6451	-0.04	0	54	4.4e-06	1	U	R.WMAWECILMGK.F
4799	720.8271	1439.6396	1439.6400	-0.24	0	(28)	0.0018	1	U	R.WMAWECILMGK.F + Oxidation (M)
4800	720.8271	1439.6397	1439.6400	-0.15	0	(36)	0.00025	1	U	R.WMAWECILMGK.F + Oxidation (M)
4802	720.8274	1439.6402	1439.6400	0.18	0	(20)	0.0093	1	U	R.WMAWECILMGK.F + Oxidation (M)
4869	728.8235	1455.6325	1455.6349	-1.61	0	(38)	0.00016	1	U	R.WMAWECILMGK.F + 2 Oxidation (M)
4871	728.8238	1455.6331	1455.6349	-1.20	0	(22)	0.0062	1	U	R.WMAWECILMGK.F + 2 Oxidation (M)
4873	728.8240	1455.6334	1455.6349	-1.03	0	(27)	0.0018	1	U	R.WMAWECILMGK.F + 2 Oxidation (M)
4874	728.8240	1455.6335	1455.6349	-0.93	0	(38)	0.00016	1	U	R.WMAWECILMGK.F + 2 Oxidation (M)
4876	728.8247	1455.6349	1455.6349	-0.01	0	(25)	0.003	1	U	R.WMAWECILMGK.F + 2 Oxidation (M)
6770	1198.5786	2395.1427	2395.1444	-0.74	0	(52)	5.8e-06	1	U	R.AQPFGQLTDEQVIENAGEFFR.D
6771	1198.5787	2395.1429	2395.1444	-0.63	0	(25)	0.0032	1	U	R.AQPFGQLTDEQVIENAGEFFR.D
6772	1198.5789	2395.1432	2395.1444	-0.53	0	(26)	0.0028	1	U	R.AQPFGQLTDEQVIENAGEFFR.D
6774	1198.5791	2395.1436	2395.1444	-0.33	0	(31)	0.00076	1	U	R.AQPFGQLTDEQVIENAGEFFR.D
6780	1198.5800	2395.1454	2395.1444	0.39	0	(30)	0.001	1	U	R.AQPFGQLTDEQVIENAGEFFR.D
6781	1198.5801	2395.1456	2395.1444	0.49	0	63	5e-07	1	U	R.AQPFGQLTDEQVIENAGEFFR.D
6783	1198.5804	2395.1463	2395.1444	0.79	0	(46)	2.6e-05	1	U	R.AQPFGQLTDEQVIENAGEFFR.D
6787	1198.5811	2395.1475	2395.1444	1.30	0	(31)	0.00072	1	U	R.AQPFGQLTDEQVIENAGEFFR.D
6788	1198.5812	2395.1478	2395.1444	1.41	0	(48)	1.5e-05	1	U	R.AQPFGQLTDEQVIENAGEFFR.D
6798	1198.5823	2395.1500	2395.1444	2.32	0	(22)	0.0058	1	U	R.AQPFGQLTDEQVIENAGEFFR.D
7058	1301.6480	2601.2813	2601.2760	2.05	0	82	6.2e-09	1	U	K.FTTASDVWAFGVTLWEVLMLCR.A
7059	1301.6482	2601.2818	2601.2760	2.23	0	(69)	1.2e-07	1	U	K.FTTASDVWAFGVTLWEVLMLCR.A
7088	1309.6426	2617.2706	2617.2709	-0.12	0	(43)	5.5e-05	1	U	K.FTTASDVWAFGVTLWEVLMLCR.A + Oxidation (M)

Table 20. JM4-3P trypsin tandem MS data.
The data provided by the CCP when 1 µg JM4-3P was separated by SDS-PAGE extracted from the gel, cleaved with trypsin and then analysed by tandem MS is shown.

JM4K-3P - Trypsin tandem MS data											
Query	Observed	Mr (expt)	Mr (calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide	
720	448.7173	895.4201	895.4222	-2.34	0	(23)	0.0047	1	U	K.IADFGMSR.N	
721	448.7175	895.4204	895.4222	-1.93	0	(37)	0.00021	1	U	K.IADFGMSR.N	
722	448.7180	895.4214	895.4222	-0.91	0	(20)	0.0093	1	U	K.IADFGMSR.N	
723	448.7180	895.4215	895.4222	-0.77	0	(23)	0.0046	1	U	K.IADFGMSR.N	
724	448.7181	895.4215	895.4222	-0.71	0	(21)	0.0071	1	U	K.IADFGMSR.N	
725	448.7181	895.4216	895.4222	-0.64	0	(23)	0.0046	1	U	K.IADFGMSR.N	
729	448.7182	895.4218	895.4222	-0.44	0	(35)	0.00034	1	U	K.IADFGMSR.N	
732	448.7183	895.4220	895.4222	-0.17	0	(38)	0.00017	1	U	K.IADFGMSR.N	
733	448.7183	895.4220	895.4222	-0.17	0	(28)	0.0016	1	U	K.IADFGMSR.N	
734	448.7183	895.4221	895.4222	-0.10	0	(29)	0.0012	1	U	K.IADFGMSR.N	
735	448.7183	895.4221	895.4222	-0.10	0	(30)	0.0011	1	U	K.IADFGMSR.N	
736	448.7183	895.4221	895.4222	-0.04	0	(30)	0.00095	1	U	K.IADFGMSR.N	
738	448.7184	895.4222	895.4222	0.03	0	(22)	0.0062	1	U	K.IADFGMSR.N	
812	456.7136	911.4126	911.4171	-4.97	0	(24)	0.004	1	U	K.IADFGMSR.N + Oxidation (M)	
815	456.7152	911.4159	911.4171	-1.28	0	(35)	0.00031	1	U	K.IADFGMSR.N + Oxidation (M)	
819	456.7154	911.4163	911.4171	-0.89	0	(35)	0.00031	1	U	K.IADFGMSR.N + Oxidation (M)	
829	456.7158	911.4170	911.4171	-0.14	0	(35)	0.00031	1	U	K.IADFGMSR.N + Oxidation (M)	
831	456.7158	911.4170	911.4171	-0.08	0	(35)	0.0003	1	U	K.IADFGMSR.N + Oxidation (M)	
836	456.7158	911.4171	911.4171	0.05	0	38	0.00015	1	U	K.IADFGMSR.N + Oxidation (M)	
840	456.7159	911.4173	911.4171	0.19	0	(30)	0.00091	1	U	K.IADFGMSR.N + Oxidation (M)	
841	456.7159	911.4173	911.4171	0.25	0	(35)	0.0003	1	U	K.IADFGMSR.N + Oxidation (M)	
846	456.7161	911.4176	911.4171	0.58	0	(23)	0.0052	1	U	K.IADFGMSR.N + Oxidation (M)	
847	456.7162	911.4178	911.4171	0.73	0	(26)	0.0028	1	U	K.IADFGMSR.N + Oxidation (M)	
849	456.7163	911.4179	911.4171	0.93	0	(28)	0.0015	1	U	K.IADFGMSR.N + Oxidation (M)	
851	456.7165	911.4185	911.4171	1.59	0	(32)	0.00068	1	U	K.IADFGMSR.N + Oxidation (M)	
865	457.2765	912.5385	912.5392	-0.83	0	(34)	0.0004	1	U	K.ILRPDATK.N	
873	457.2768	912.5390	912.5392	-0.30	0	35	0.00033	1	U	K.ILRPDATK.N	
875	457.2768	912.5391	912.5392	-0.17	0	(31)	0.00075	1	U	K.ILRPDATK.N	
877	457.2769	912.5393	912.5392	0.05	0	(31)	0.00086	1	U	K.ILRPDATK.N	
878	457.2769	912.5393	912.5392	0.05	0	(30)	0.00089	1	U	K.ILRPDATK.N	
880	457.2770	912.5394	912.5392	0.18	0	(34)	0.00041	1	U	K.ILRPDATK.N	
883	457.2771	912.5396	912.5392	0.44	0	(26)	0.0023	1	U	K.ILRPDATK.N	
968	467.2965	932.5784	932.5807	-2.49	0	(28)	0.0015	1	U	K.GHPLLVAVK.I	
974	467.2968	932.5791	932.5807	-1.72	0	(31)	0.00086	1	U	K.GHPLLVAVK.I	
982	467.2975	932.5804	932.5807	-0.33	0	32	0.00056	1	U	K.GHPLLVAVK.I	
984	467.2976	932.5807	932.5807	-0.01	0	(27)	0.0022	1	U	K.GHPLLVAVK.I	
1294	484.7969	967.5792	967.5814	-2.30	1	(21)	0.0073	1	U	R.LKDPNIIR.L	
1299	484.7971	967.5797	967.5814	-1.74	1	(21)	0.0077	1	U	R.LKDPNIIR.L	
1303	484.7973	967.5801	967.5814	-1.37	1	(26)	0.0026	1	U	R.LKDPNIIR.L	

1304	484.7975	967.5803	967.5814	-1.12	1	(24)	0.0044	1	U	R.LKDPNIIR.L	
1308	484.7976	967.5807	967.5814	-0.73	1	(22)	0.0069	1	U	R.LKDPNIIR.L	
1312	484.7978	967.5810	967.5814	-0.48	1	27	0.0018	1	U	R.LKDPNIIR.L	
1318	484.7979	967.5813	967.5814	-0.17	1	(21)	0.0074	1	U	R.LKDPNIIR.L	
1319	484.7979	967.5813	967.5814	-0.17	1	(21)	0.0084	1	U	R.LKDPNIIR.L	
1322	484.7981	967.5816	967.5814	0.16	1	(21)	0.008	1	U	R.LKDPNIIR.L	
1329	484.7993	967.5841	967.5814	2.74	1	(27)	0.0019	1	U	R.LKDPNIIR.L	
1332	485.2886	968.5627	968.5654	-2.81	1	(25)	0.0029	1	U	R.LKDPNIIR.L + Deamidated (NQ)	
1333	485.2892	968.5638	968.5654	-1.69	1	(21)	0.0073	1	U	R.LKDPNIIR.L + Deamidated (NQ)	
2005	531.3458	1060.6770	1060.6757	1.20	1	24	0.0039	1	U	R.KGHPLLVAVK.I	
2174	546.7805	1091.5465	1091.5499	-3.11	0	(29)	0.0012	1	U	R.FLAEDALNTV.-	
2175	546.7811	1091.5477	1091.5499	-1.99	0	(30)	0.001	1	U	R.FLAEDALNTV.-	
2176	546.7813	1091.5481	1091.5499	-1.66	0	(37)	0.0002	1	U	R.FLAEDALNTV.-	
2177	546.7814	1091.5482	1091.5499	-1.55	0	(61)	8.1e-07	1	U	R.FLAEDALNTV.-	
2178	546.7814	1091.5483	1091.5499	-1.44	0	(26)	0.0027	1	U	R.FLAEDALNTV.-	
2179	546.7817	1091.5488	1091.5499	-0.98	0	(32)	0.00057	1	U	R.FLAEDALNTV.-	
2180	546.7818	1091.5490	1091.5499	-0.76	0	(33)	0.00045	1	U	R.FLAEDALNTV.-	
2181	546.7819	1091.5492	1091.5499	-0.65	0	(28)	0.0015	1	U	R.FLAEDALNTV.-	
2182	546.7819	1091.5493	1091.5499	-0.54	0	(42)	6.1e-05	1	U	R.FLAEDALNTV.-	
2183	546.7819	1091.5493	1091.5499	-0.54	0	61	7.5e-07	1	U	R.FLAEDALNTV.-	
2184	546.7819	1091.5493	1091.5499	-0.54	0	(26)	0.0028	1	U	R.FLAEDALNTV.-	
2185	546.7819	1091.5493	1091.5499	-0.54	0	(29)	0.0012	1	U	R.FLAEDALNTV.-	
2186	546.7819	1091.5493	1091.5499	-0.54	0	(35)	0.00029	1	U	R.FLAEDALNTV.-	
2187	546.7820	1091.5494	1091.5499	-0.43	0	(26)	0.0025	1	U	R.FLAEDALNTV.-	
2188	546.7820	1091.5495	1091.5499	-0.32	0	(33)	0.00049	1	U	R.FLAEDALNTV.-	
2189	546.7820	1091.5495	1091.5499	-0.32	0	(27)	0.0019	1	U	R.FLAEDALNTV.-	
2190	546.7822	1091.5498	1091.5499	-0.09	0	(37)	0.00018	1	U	R.FLAEDALNTV.-	
2191	546.7822	1091.5499	1091.5499	0.02	0	(29)	0.0012	1	U	R.FLAEDALNTV.-	
2192	546.7823	1091.5500	1091.5499	0.13	0	(40)	9.6e-05	1	U	R.FLAEDALNTV.-	
2193	546.7824	1091.5503	1091.5499	0.35	0	(25)	0.0029	1	U	R.FLAEDALNTV.-	
2194	546.7825	1091.5504	1091.5499	0.46	0	(47)	1.9e-05	1	U	R.FLAEDALNTV.-	
2195	546.7825	1091.5505	1091.5499	0.57	0	(42)	6.2e-05	1	U	R.FLAEDALNTV.-	
2196	546.7825	1091.5505	1091.5499	0.57	0	(21)	0.0071	1	U	R.FLAEDALNTV.-	
2197	546.7826	1091.5506	1091.5499	0.68	0	(28)	0.0015	1	U	R.FLAEDALNTV.-	
2198	546.7826	1091.5506	1091.5499	0.68	0	(40)	0.00011	1	U	R.FLAEDALNTV.-	
2199	546.7826	1091.5507	1091.5499	0.79	0	(23)	0.0051	1	U	R.FLAEDALNTV.-	
2201	546.7827	1091.5509	1091.5499	0.90	0	(29)	0.0013	1	U	R.FLAEDALNTV.-	
2202	546.7828	1091.5510	1091.5499	1.03	0	(50)	9e-06	1	U	R.FLAEDALNTV.-	
2203	546.7828	1091.5511	1091.5499	1.14	0	(38)	0.00017	1	U	R.FLAEDALNTV.-	
2204	546.7830	1091.5515	1091.5499	1.47	0	(27)	0.002	1	U	R.FLAEDALNTV.-	
2213	547.2740	1092.5335	1092.5339	-0.31	0	(36)	0.00028	1	U	R.FLAEDALNTV.- + Deamidated (NQ)	
2214	547.2747	1092.5349	1092.5339	0.91	0	(52)	7e-06	1	U	R.FLAEDALNTV.- + Deamidated (NQ)	
2215	547.2764	1092.5383	1092.5339	4.04	0	(42)	6.5e-05	1	U	R.FLAEDALNTV.- + Deamidated (NQ)	
2216	547.2785	1092.5424	1092.5339	7.83	0	(31)	0.00082	1	U	R.FLAEDALNTV.- + Deamidated (NQ)	
2558	567.7635	1133.5124	1133.5141	-1.51	0	(25)	0.0032	1	U	R.NLYAGDYYR.V	
2559	567.7635	1133.5124	1133.5141	-1.51	0	(22)	0.0065	1	U	R.NLYAGDYYR.V	

2560	567.7636	1133.5125	1133.5141	-1.41	0	(27)	0.002	1	U	R.NLYAGDYYR.V	
2561	567.7636	1133.5127	1133.5141	-1.30	0	(29)	0.0013	1	U	R.NLYAGDYYR.V	
2566	567.7640	1133.5135	1133.5141	-0.54	0	34	0.00036	1	U	R.NLYAGDYYR.V	
2567	567.7642	1133.5139	1133.5141	-0.23	0	(24)	0.0041	1	U	R.NLYAGDYYR.V	
3204	**607.7459**	**1213.4771**	**1213.4805**	**-2.74**	**0**	**(24)**	**0.0041**	**1**	**U**	**R.NLYAGDYYR.V + Phospho_STY (STY)**	**A-loop**
3205	**607.7459**	**1213.4773**	**1213.4805**	**-2.64**	**0**	**(20)**	**0.0091**	**1**	**U**	**R.NLYAGDYYR.V + Phospho_STY (STY)**	**A-loop**
3214	**607.7469**	**1213.4792**	**1213.4805**	**-1.02**	**0**	**(28)**	**0.0015**	**1**	**U**	**R.NLYAGDYYR.V + Phospho_STY (STY)**	**A-loop**
3222	**607.7475**	**1213.4804**	**1213.4805**	**-0.02**	**0**	**(23)**	**0.0047**	**1**	**U**	**R.NLYAGDYYR.V + Phospho_STY (STY)**	**A-loop**
3233	**607.7481**	**1213.4817**	**1213.4805**	**0.99**	**0**	**(22)**	**0.0062**	**1**	**U**	**R.NLYAGDYYR.V + Phospho_STY (STY)**	**A-loop**
3322	617.3395	1232.6644	1232.6666	-1.76	0	(35)	0.00031	1	U	R.YLATLNFVHR.D	
3323	617.3395	1232.6645	1232.6666	-1.66	0	(25)	0.0031	1	U	R.YLATLNFVHR.D	
3325	617.3397	1232.6648	1232.6666	-1.47	0	(37)	0.00018	1	U	R.YLATLNFVHR.D	
3328	617.3399	1232.6652	1232.6666	-1.08	0	49	1.2e-05	1	U	R.YLATLNFVHR.D	
3330	617.3400	1232.6654	1232.6666	-0.97	0	(46)	2.4e-05	1	U	R.YLATLNFVHR.D	
3336	617.3404	1232.6662	1232.6666	-0.29	0	(24)	0.0043	1	U	R.YLATLNFVHR.D	
3339	617.3405	1232.6665	1232.6666	-0.07	0	(33)	0.00046	1	U	R.YLATLNFVHR.D	
3354	617.3409	1232.6673	1232.6666	0.61	0	(33)	0.0005	1	U	R.YLATLNFVHR.D	
3373	617.3420	1232.6694	1232.6666	2.29	0	(33)	0.00054	1	U	R.YLATLNFVHR.D	
3376	617.8336	1233.6526	1233.6506	1.60	0	(33)	0.0005	1	U	R.YLATLNFVHR.D + Deamidated (NQ)	
3377	617.8337	1233.6529	1233.6506	1.90	0	(47)	1.9e-05	1	U	R.YLATLNFVHR.D + Deamidated (NQ)	
3378	617.8345	1233.6544	1233.6506	3.08	0	(24)	0.0041	1	U	R.YLATLNFVHR.D + Deamidated (NQ)	
3380	617.8359	1233.6573	1233.6506	5.46	0	(21)	0.0072	1	U	R.YLATLNFVHR.D + Deamidated (NQ)	
3381	412.2267	1233.6581	1233.6506	6.11	0	(20)	0.0091	1	U	R.YLATLNFVHR.D + Deamidated (NQ)	
3608	419.2418	1254.7035	1254.7044	-0.75	1	22	0.0064	1	U	K.ILRPDATKNAR.N + Deamidated (NQ)	
3917	**647.7316**	**1293.4487**	**1293.4468**	**1.48**	**0**	**(24)**	**0.004**	**1**	**U**	**R.NLYAGDYYR.V + 2 Phospho_STY (STY)**	**A-loop**
3918	**647.7321**	**1293.4496**	**1293.4468**	**2.15**	**0**	**(22)**	**0.0064**	**1**	**U**	**R.NLYAGDYYR.V + 2 Phospho_STY (STY)**	**A-loop**
3920	647.8247	1293.6347	1293.6387	-3.07	0	(33)	0.00047	1	U	R.NCLVGENFTIK.I	
3921	647.8249	1293.6353	1293.6387	-2.61	0	(25)	0.0032	1	U	R.NCLVGENFTIK.I	
3922	647.8254	1293.6362	1293.6387	-1.94	0	(29)	0.0012	1	U	R.NCLVGENFTIK.I	
3923	647.8257	1293.6368	1293.6387	-1.48	0	(23)	0.005	1	U	R.NCLVGENFTIK.I	
3924	647.8258	1293.6371	1293.6387	-1.28	0	(32)	0.00063	1	U	R.NCLVGENFTIK.I	
3925	647.8259	1293.6373	1293.6387	-1.09	0	(21)	0.0077	1	U	R.NCLVGENFTIK.I	
3926	647.8259	1293.6373	1293.6387	-1.09	0	(32)	0.00065	1	U	R.NCLVGENFTIK.I	
3927	647.8260	1293.6374	1293.6387	-1.00	0	(21)	0.0086	1	U	R.NCLVGENFTIK.I	
3929	647.8261	1293.6377	1293.6387	-0.81	0	(28)	0.0015	1	U	R.NCLVGENFTIK.I	
3930	647.8262	1293.6379	1293.6387	-0.63	0	(32)	0.00069	1	U	R.NCLVGENFTIK.I	
3932	647.8263	1293.6380	1293.6387	-0.53	0	(42)	5.7e-05	1	U	R.NCLVGENFTIK.I	
3936	647.8267	1293.6389	1293.6387	0.13	0	(32)	0.00066	1	U	R.NCLVGENFTIK.I	
3937	647.8268	1293.6391	1293.6387	0.32	0	(36)	0.00026	1	U	R.NCLVGENFTIK.I	
3938	647.8269	1293.6392	1293.6387	0.41	0	(23)	0.0045	1	U	R.NCLVGENFTIK.I	
3939	647.8269	1293.6392	1293.6387	0.41	0	(33)	0.00052	1	U	R.NCLVGENFTIK.I	
3941	647.8269	1293.6392	1293.6387	0.41	0	(30)	0.001	1	U	R.NCLVGENFTIK.I	
3942	647.8270	1293.6394	1293.6387	0.52	0	(23)	0.0045	1	U	R.NCLVGENFTIK.I	
3944	647.8271	1293.6397	1293.6387	0.80	0	(24)	0.0043	1	U	R.NCLVGENFTIK.I	
3946	647.8273	1293.6400	1293.6387	0.98	0	(25)	0.0031	1	U	R.NCLVGENFTIK.I	
3947	647.8273	1293.6400	1293.6387	0.98	0	(39)	0.00013	1	U	R.NCLVGENFTIK.I	

3948	647.8273	1293.6401	1293.6387	1.07	0	(22)	0.0067	1	U	R.NCLVGENFTIK.I	
3949	647.8273	1293.6401	1293.6387	1.07	0	(35)	0.00034	1	U	R.NCLVGENFTIK.I	
3950	647.8274	1293.6403	1293.6387	1.26	0	(27)	0.0019	1	U	R.NCLVGENFTIK.I	
3951	647.8275	1293.6405	1293.6387	1.35	0	(30)	0.0011	1	U	R.NCLVGENFTIK.I	
3952	647.8275	1293.6405	1293.6387	1.35	0	(31)	0.00084	1	U	R.NCLVGENFTIK.I	
3953	647.8276	1293.6407	1293.6387	1.55	0	(44)	3.6e-05	1	U	R.NCLVGENFTIK.I	
3954	647.8277	1293.6408	1293.6387	1.65	0	(29)	0.0012	1	U	R.NCLVGENFTIK.I	
3955	647.8278	1293.6411	1293.6387	1.83	0	(24)	0.0038	1	U	R.NCLVGENFTIK.I	
3956	647.8278	1293.6411	1293.6387	1.83	0	(31)	0.00081	1	U	R.NCLVGENFTIK.I	
3957	647.8279	1293.6413	1293.6387	2.02	0	(42)	6.7e-05	1	U	R.NCLVGENFTIK.I	
3958	647.8284	1293.6422	1293.6387	2.68	0	(24)	0.0044	1	U	R.NCLVGENFTIK.I	
3959	647.8284	1293.6422	1293.6387	2.68	0	(30)	0.001	1	U	R.NCLVGENFTIK.I	
3965	648.3185	1294.6225	1294.6227	-0.15	0	(27)	0.002	1	U	R.NCLVGENFTIK.I + Deamidated (NQ)	
3966	648.3190	1294.6234	1294.6227	0.51	0	(36)	0.00027	1	U	R.NCLVGENFTIK.I + Deamidated (NQ)	
3967	648.3198	1294.6251	1294.6227	1.82	0	(34)	0.00039	1	U	R.NCLVGENFTIK.I + Deamidated (NQ)	
3968	648.3204	1294.6262	1294.6227	2.67	0	(32)	0.00062	1	U	R.NCLVGENFTIK.I + Deamidated (NQ)	
3976	648.8104	1295.6062	1295.6067	-0.44	0	55	3e-06	1	U	R.NCLVGENFTIK.I + 2 Deamidated (NQ)	
3977	648.8109	1295.6073	1295.6067	0.41	0	(36)	0.00025	1	U	R.NCLVGENFTIK.I + 2 Deamidated (NQ)	
4704	712.8303	1423.6460	1423.6451	0.64	0	(35)	0.00034	1	U	R.WMAWECILMGK.F	
4709	712.8312	1423.6479	1423.6451	2.02	0	(34)	0.00039	1	U	R.WMAWECILMGK.F	
4710	712.8312	1423.6479	1423.6451	2.02	0	(23)	0.0051	1	U	R.WMAWECILMGK.F	
4769	720.8277	1439.6408	1439.6400	0.61	0	(30)	0.00099	1	U	R.WMAWECILMGK.F + Oxidation (M)	
4771	720.8286	1439.6427	1439.6400	1.87	0	(31)	0.00071	1	U	R.WMAWECILMGK.F + Oxidation (M)	
4819	728.8232	1455.6319	1455.6349	-2.03	0	(25)	0.0029	1	U	R.WMAWECILMGK.F + 2 Oxidation (M)	
4821	728.8239	1455.6333	1455.6349	-1.11	0	(28)	0.0015	1	U	R.WMAWECILMGK.F + 2 Oxidation (M)	
4824	728.8251	1455.6357	1455.6349	0.56	0	45	3.2e-05	1	U	R.WMAWECILMGK.F + 2 Oxidation (M)	
6649	1198.5787	2395.1429	2395.1444	-0.63	0	(24)	0.0039	1	U	R.AQPFGQLTDEQVIENAGEFFR.D	
6651	1198.5791	2395.1436	2395.1444	-0.33	0	(44)	4.1e-05	1	U	R.AQPFGQLTDEQVIENAGEFFR.D	
6653	1198.5796	2395.1446	2395.1444	0.08	0	71	7.6e-08	1	U	R.AQPFGQLTDEQVIENAGEFFR.D	
6654	1198.5796	2395.1446	2395.1444	0.08	0	(63)	5.4e-07	1	U	R.AQPFGQLTDEQVIENAGEFFR.D	
6660	1198.5802	2395.1458	2395.1444	0.59	0	(47)	1.9e-05	1	U	R.AQPFGQLTDEQVIENAGEFFR.D	
6664	1198.5806	2395.1466	2395.1444	0.90	0	(24)	0.0042	1	U	R.AQPFGQLTDEQVIENAGEFFR.D	
6665	1198.5806	2395.1466	2395.1444	0.90	0	(52)	6.2e-06	1	U	R.AQPFGQLTDEQVIENAGEFFR.D	
6679	1198.5824	2395.1502	2395.1444	2.42	0	(39)	0.00012	1	U	R.AQPFGQLTDEQVIENAGEFFR.D	
6680	1198.5825	2395.1505	2395.1444	2.53	0	(63)	4.9e-07	1	U	R.AQPFGQLTDEQVIENAGEFFR.D	
6929	1301.6455	2601.2765	2601.2760	0.17	0	(39)	0.00013	1	U	K.FTTASDVWAFGVTLWEVLMLCR.A	
6930	1301.6464	2601.2782	2601.2760	0.83	0	(60)	9.9e-07	1	U	K.FTTASDVWAFGVTLWEVLMLCR.A	
6959	1309.6453	2617.2760	2617.2709	1.92	0	70	9.9e-08	1	U	K.FTTASDVWAFGVTLWEVLMLCR.A + Oxidation (M)	
6960	1309.6462	2617.2779	2617.2709	2.67	0	(41)	8.3e-05	1	U	K.FTTASDVWAFGVTLWEVLMLCR.A + Oxidation (M)	

Table 21. JM4-3P chymotrypsin tandem MS data.
The data provided by the CCP when 1 µg JM4-3P was separated by SDS-PAGE extracted from the gel, cleaved with chymotrypsin, and then analysed by tandem MS is shown.

JM4K 3P Chymotrypsin tandem MS data											
Query	Observed	Mr (expt)	Mr (calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide	
113	389.6872	777.3598	777.3592	0.74	0	(23)	1.2	1	U	L.CRAQPF.G	
115	389.6874	777.3602	777.3592	1.31	0	24	1	1	U	L.CRAQPF.G	
205	397.2149	792.4152	792.4170	-2.26	2	31	0.17	1		W.AFGVTLW.E	
206	397.2151	792.4156	792.4170	-1.78	2	(29)	0.27	1		W.AFGVTLW.E	
334	404.2342	806.4539	806.4538	0.19	0	(34)	0.059	1		F.TIKIADF.G	
335	404.2343	806.4541	806.4538	0.34	0	(37)	0.03	1		F.TIKIADF.G	
336	404.2343	806.4541	806.4538	0.41	0	39	0.019	1		F.TIKIADF.G	
337	404.2344	806.4542	806.4538	0.51	0	(35)	0.039	1		F.TIKIADF.G	
338	404.2344	806.4542	806.4538	0.51	0	(27)	0.25	1		F.TIKIADF.G	
339	404.2345	806.4544	806.4538	0.81	0	(35)	0.043	1		F.TIKIADF.G	
340	404.2346	806.4546	806.4538	0.96	0	(34)	0.06	1		F.TIKIADF.G	
341	404.2346	806.4547	806.4538	1.11	0	(32)	0.082	1		F.TIKIADF.G	
343	404.2348	806.4551	806.4538	1.63	0	(29)	0.16	1		F.TIKIADF.G	
344	404.2349	806.4552	806.4538	1.70	0	(21)	1.2	1		F.TIKIADF.G	
345	404.2349	806.4553	806.4538	1.85	0	(34)	0.06	1		F.TIKIADF.G	
423	410.7559	819.4972	819.4966	0.71	1	21	0.41	2	U	L.RFKEKL.G	
424	410.7561	819.4976	819.4966	1.15	1	(20)	0.51	1	U	L.RFKEKL.G	
564	420.7054	839.3962	839.3959	0.31	1	(23)	1.6	1		F.GMSRNLY.A	
565	420.7054	839.3963	839.3959	0.45	1	30	0.34	1		F.GMSRNLY.A	
866	448.7091	895.4036	895.4036	0.03	0	22	2.2	1		L.QGVTGGNTY.A	
867	448.7093	895.4040	895.4036	0.43	0	(22)	2.6	1		L.QGVTGGNTY.A	
923	451.2294	900.4443	900.4454	-1.17	1	21	2.6	1		L.NFVHRDL.A + Deamidated (NQ)	
924	451.2297	900.4448	900.4454	-0.61	1	(20)	2.7	1		L.NFVHRDL.A + Deamidated (NQ)	
1049	460.7732	919.5319	919.5352	-3.57	0	(20)	1.2	2	U	L.NVRKGHPL.L	
1050	460.7734	919.5323	919.5352	-3.16	0	24	0.47	1	U	L.NVRKGHPL.L	
1051	460.7751	919.5355	919.5352	0.41	0	(21)	1.1	1	U	L.NVRKGHPL.L	
1052	460.7751	919.5356	919.5352	0.50	0	(23)	0.68	1	U	L.NVRKGHPL.L	
1168	473.2478	944.4811	944.4815	-0.37	2	(28)	0.43	1	U	F.LAEDALNTV.-	
1169	473.2479	944.4812	944.4815	-0.25	2	(55)	0.00082	1	U	F.LAEDALNTV.-	
1170	473.2480	944.4814	944.4815	-0.03	2	(36)	0.065	1	U	F.LAEDALNTV.-	
1171	473.2481	944.4816	944.4815	0.16	2	63	0.00015	1	U	F.LAEDALNTV.-	
1177	473.7377	945.4608	945.4655	-4.95	2	(25)	1.1	1	U	F.LAEDALNTV.- + Deamidated (NQ)	
1178	473.7395	945.4645	945.4655	-1.02	2	(33)	0.16	1	U	F.LAEDALNTV.- + Deamidated (NQ)	
1328	484.7972	967.5799	967.5814	-1.62	0	24	0.18	1	U	L.KDPNIIRL.L	
1496	499.2451	996.4757	996.4772	-1.57	1	32	0.24	1	U	W.ECILMGKF.T	
1498	499.2464	996.4783	996.4772	1.06	1	(32)	0.23	1	U	W.ECILMGKF.T	
1543	501.7316	1001.4486	1001.4488	-0.23	0	63	0.00022	1	U	L.GVCVQDDPL.C	
1544	501.7317	1001.4489	1001.4488	0.07	0	(41)	0.029	1	U	L.GVCVQDDPL.C	
1545	501.7326	1001.4507	1001.4488	1.90	0	(47)	0.0073	1	U	L.GVCVQDDPL.C	

1546	501.7328	1001.4510	1001.4488	2.20	0	(47)	0.0082	1	U	L.GVCVQDDPL.C	
1576	507.2431	1012.4716	1012.4722	-0.57	1	(26)	0.91	1	U	W.ECILMGKF.T + Oxidation (M)	
1634	511.2627	1020.5109	1020.5101	0.81	0	22	2.5	1	U	F.RDQGRQVY.L	
1647	511.7485	1021.4825	1021.4837	-1.18	1	(23)	2	2	U	L.MLCRAQPF.G	
1648	511.7488	1021.4830	1021.4837	-0.71	1	29	0.5	1	U	L.MLCRAQPF.G	
1733	518.2763	1034.5381	1034.5396	-1.52	1	41	0.025	1		F.KEKLGEGQF.G	
1734	518.2773	1034.5400	1034.5396	0.35	1	(36)	0.078	1		F.KEKLGEGQF.G	
1866	531.7249	1061.4353	1061.4335	1.62	0	24	1.4	1	U	L.CEVDSPQDL.V	
1897	534.2248	1066.4350	1066.4389	-3.68	1	(29)	0.36	1		Y.MENGDLNQF.L	
1898	534.2252	1066.4359	1066.4389	-2.87	1	(27)	0.6	1		Y.MENGDLNQF.L	
1899	534.2253	1066.4360	1066.4389	-2.76	1	(40)	0.033	1		Y.MENGDLNQF.L	
1900	534.2260	1066.4375	1066.4389	-1.39	1	(24)	1.2	1		Y.MENGDLNQF.L	
1907	534.7159	1067.4173	1067.4230	-5.27	1	(42)	0.021	1		Y.MENGDLNQF.L + Deamidated (NQ)	
1908	534.7168	1067.4190	1067.4230	-3.66	1	(31)	0.28	1		Y.MENGDLNQF.L + Deamidated (NQ)	
1909	534.7180	1067.4215	1067.4230	-1.38	1	(33)	0.17	1		Y.MENGDLNQF.L + Deamidated (NQ)	
1922	536.2834	1070.5522	1070.5509	1.27	2	(25)	0.86	1	U	L.HRFLAEDAL.N	
1923	536.2836	1070.5526	1070.5509	1.60	2	31	0.19	1	U	L.HRFLAEDAL.N	
1930	536.7559	1071.4972	1071.4985	-1.28	1	47	0.0077	1		L.GEGQFGEVHL.C	
1932	536.7564	1071.4983	1071.4985	-0.25	1	(47)	0.0078	1		L.GEGQFGEVHL.C	
1978	541.3386	1080.6626	1080.6655	-2.71	1	(31)	0.026	1	U	L.KDPNIIRLL.G	
1979	541.3387	1080.6628	1080.6655	-2.49	1	45	0.00088	1	U	L.KDPNIIRLL.G	
1980	541.3389	1080.6633	1080.6655	-2.03	1	(24)	0.13	1	U	L.KDPNIIRLL.G	
1983	541.3397	1080.6649	1080.6655	-0.56	1	(34)	0.011	1	U	L.KDPNIIRLL.G	
1985	541.3398	1080.6651	1080.6655	-0.34	1	(35)	0.0099	1	U	L.KDPNIIRLL.G	
1986	541.3406	1080.6667	1080.6655	1.14	1	(27)	0.058	1	U	L.KDPNIIRLL.G	
1987	541.3408	1080.6671	1080.6655	1.47	1	(25)	0.11	1	U	L.KDPNIIRLL.G	
1990	541.7742	1081.5339	1081.5339	0.05	0	26	1.1	1	U	L.SRPPACPQGL.Y	
1991	541.7744	1081.5343	1081.5339	0.38	0	(23)	1.9	1	U	L.SRPPACPQGL.Y	
1992	542.2226	1082.4306	1082.4339	-2.97	1	(31)	0.28	1		Y.MENGDLNQF.L + Oxidation (M)	
1993	542.2227	1082.4309	1082.4339	-2.75	1	(29)	0.43	1		Y.MENGDLNQF.L + Oxidation (M)	
1994	542.2230	1082.4314	1082.4339	-2.31	1	(44)	0.012	1		Y.MENGDLNQF.L + Oxidation (M)	
1995	542.2236	1082.4327	1082.4339	-1.07	1	47	0.007	1		Y.MENGDLNQF.L + Oxidation (M)	
1996	542.2237	1082.4329	1082.4339	-0.85	1	(27)	0.71	1		Y.MENGDLNQF.L + Oxidation (M)	
1997	542.2240	1082.4334	1082.4339	-0.39	1	(40)	0.032	1		Y.MENGDLNQF.L + Oxidation (M)	
1998	542.2244	1082.4342	1082.4339	0.30	1	(36)	0.079	1		Y.MENGDLNQF.L + Oxidation (M)	
1999	542.2244	1082.4342	1082.4339	0.30	1	(35)	0.11	1		Y.MENGDLNQF.L + Oxidation (M)	
2214	558.2731	1114.5317	1114.5329	-1.05	1	51	0.0029	1	U	L.LGVCVQDDPL.C	
2215	558.2739	1114.5333	1114.5329	0.38	1	(48)	0.0067	1	U	L.LGVCVQDDPL.C	
2551	584.7964	1167.5782	1167.5785	-0.24	1	(22)	2.5	1	U	F.FRDQGRQVY.L	
2552	390.2001	1167.5786	1167.5785	0.05	1	27	0.91	1	U	F.FRDQGRQVY.L	
2616	590.7673	1179.5201	1179.5230	-2.46	2	41	0.031	1		Y.MENGDLNQFL.S	
2617	590.7675	1179.5205	1179.5230	-2.14	2	(36)	0.095	1		Y.MENGDLNQFL.S	
2618	590.7692	1179.5238	1179.5230	0.66	2	(35)	0.12	1		Y.MENGDLNQFL.S	
2620	591.2593	1180.5040	1180.5070	-2.55	2	(40)	0.035	1		Y.MENGDLNQFL.S + Deamidated (NQ)	
2710	598.3163	1194.6180	1194.6179	0.07	1	34	0.15	1	U	Y.LSRPPACPQGL.Y	
2711	598.3171	1194.6197	1194.6179	1.51	1	(28)	0.46	1	U	Y.LSRPPACPQGL.Y	

2712	598.7660	1195.5174	1195.5179	-0.42	2	(32)	0.27	1		Y.MENGDLNQFL.S + Oxidation (M)	
2713	598.7662	1195.5178	1195.5179	-0.12	2	(33)	0.22	1		Y.MENGDLNQFL.S + Oxidation (M)	
2714	598.7665	1195.5185	1195.5179	0.50	2	(37)	0.092	1		Y.MENGDLNQFL.S + Oxidation (M)	
2715	598.7668	1195.5190	1195.5179	0.90	2	(24)	1.9	1		Y.MENGDLNQFL.S + Oxidation (M)	
2983	616.8029	1231.5912	1231.5945	-2.73	0	(21)	3.8	1	U	W.SRESEQRPPF.S	
2984	616.8032	1231.5919	1231.5945	-2.14	0	(30)	0.47	1	U	W.SRESEQRPPF.S	
2985	616.8046	1231.5947	1231.5945	0.15	0	(29)	0.6	1	U	W.SRESEQRPPF.S	
2988	616.8048	1231.5951	1231.5945	0.44	0	(22)	2.7	1	U	W.SRESEQRPPF.S	
2989	616.8049	1231.5952	1231.5945	0.54	0	(31)	0.35	1	U	W.SRESEQRPPF.S	
2990	616.8051	1231.5957	1231.5945	0.93	0	(40)	0.044	1	U	W.SRESEQRPPF.S	
2993	616.8054	1231.5962	1231.5945	1.33	0	(38)	0.062	1	U	W.SRESEQRPPF.S	
2995	411.5395	1231.5966	1231.5945	1.72	0	(25)	1.5	1	U	W.SRESEQRPPF.S	
3005	617.3024	1232.5902	1232.5785	9.50	0	43	0.021	1	U	W.SRESEQRPPF.S + Deamidated (NQ)	
3070	623.3048	1244.5951	1244.5972	-1.70	1	(36)	0.1	1	U	L.SRPPACPQGLY.E	
3073	623.3057	1244.5969	1244.5972	-0.22	1	(35)	0.16	1	U	L.SRPPACPQGLY.E	
3074	623.3061	1244.5976	1244.5972	0.36	1	(35)	0.13	1	U	L.SRPPACPQGLY.E	
3076	623.3065	1244.5984	1244.5972	0.95	1	43	0.023	1	U	L.SRPPACPQGLY.E	
3077	623.3070	1244.5993	1244.5972	1.74	1	(28)	0.68	1	U	L.SRPPACPQGLY.E	
3293	640.8054	1279.5962	1279.5979	-1.37	1	(23)	2.3	1	U	L.ATRNCLVGENF.T	
3295	640.8065	1279.5985	1279.5979	0.45	1	(23)	2.4	1	U	L.ATRNCLVGENF.T	
3296	640.8066	1279.5986	1279.5979	0.54	1	(32)	0.35	1	U	L.ATRNCLVGENF.T	
3297	640.8071	1279.5996	1279.5979	1.31	1	(28)	0.81	1	U	L.ATRNCLVGENF.T	
3298	640.8079	1279.6012	1279.5979	2.54	1	(31)	0.43	1	U	L.ATRNCLVGENF.T	
3303	641.2968	1280.5791	1280.5819	-2.23	1	39	0.073	1	U	L.ATRNCLVGENF.T + Deamidated (NQ)	
3309	641.3383	1280.6620	1280.6626	-0.47	2	21	2.7	1	U	F.FRDQGRQVYL.S	
3516	660.3279	1318.6412	1318.6452	-3.02	0	53	0.0024	1	U	L.HVAAQIASGMRY.L + Oxidation (M)	
3519	660.3364	1318.6582	1318.6452	9.90	0	(45)	0.014	1	U	L.HVAAQIASGMRY.L + Oxidation (M)	
3616	446.9104	1337.7094	1337.7092	0.15	2	(22)	1.7	1	U	L.RFKEKLGEGQF.G	
3617	669.8631	1337.7116	1337.7092	1.85	2	37	0.052	1	U	L.RFKEKLGEGQF.G	
3618	669.8637	1337.7129	1337.7092	2.77	2	(36)	0.061	1	U	L.RFKEKLGEGQF.G	
3659	676.3040	1350.5935	1350.5939	-0.33	0	31	0.41	1	U	L.TDEQVIENAGEF.F	
3660	676.3045	1350.5944	1350.5939	0.38	0	(22)	3.1	1	U	L.TDEQVIENAGEF.F	
3680	679.8466	1357.6787	1357.6812	-1.88	2	49	0.0053	1	U	Y.LSRPPACPQGLY.E	
3681	679.8467	1357.6789	1357.6812	-1.71	2	(38)	0.066	1	U	Y.LSRPPACPQGLY.E	
3682	679.8467	1357.6789	1357.6812	-1.71	2	(42)	0.024	1	U	Y.LSRPPACPQGLY.E	
3683	679.8470	1357.6795	1357.6812	-1.25	2	(46)	0.0097	1	U	Y.LSRPPACPQGLY.E	
3684	679.8476	1357.6806	1357.6812	-0.44	2	(22)	2.8	1	U	Y.LSRPPACPQGLY.E	
3685	679.8479	1357.6812	1357.6812	0.00	2	(27)	0.84	1	U	Y.LSRPPACPQGLY.E	
3686	679.8479	1357.6812	1357.6812	0.00	2	(27)	0.77	1	U	Y.LSRPPACPQGLY.E	
3687	679.8480	1357.6815	1357.6812	0.18	2	(36)	0.11	1	U	Y.LSRPPACPQGLY.E	
3690	679.8481	1357.6817	1357.6812	0.36	2	(24)	1.7	1	U	Y.LSRPPACPQGLY.E	
3695	680.3276	1358.6407	1358.6653	-18.05	2	(38)	0.07	1	U	Y.LSRPPACPQGLY.E + Deamidated (NQ)	
3701	681.3174	1360.6203	1360.6181	1.67	1	32	0.36	1	U	L.CEVDSPQDLVSL.D	
3702	681.3180	1360.6215	1360.6181	2.57	1	(31)	0.44	1	U	L.CEVDSPQDLVSL.D	
3841	702.8519	1403.6893	1403.6906	-0.89	0	(21)	3.9	1	U	L.RPDATKNARNDF.L	
3842	468.9040	1403.6901	1403.6906	-0.29	0	(21)	3.5	1	U	L.RPDATKNARNDF.L	

3843	468.9040	1403.6902	1403.6906	-0.23	0	(24)	2.1	1	U	L.RPDATKNARNDF.L	
3844	702.8524	1403.6903	1403.6906	-0.19	0	30	0.46	1	U	L.RPDATKNARNDF.L	
3848	703.3438	1404.6729	1404.6746	-1.15	0	(23)	2.8	1	U	L.RPDATKNARNDF.L + Deamidated (NQ)	
3850	469.2334	1404.6785	1404.6746	2.78	0	(30)	0.49	1	U	L.RPDATKNARNDF.L + Deamidated (NQ)	
4107	741.3845	1480.7544	1480.7562	-1.23	1	(34)	0.18	1	U	-.GPVPHYAEADIVTL.Q	
4108	741.3848	1480.7550	1480.7562	-0.82	1	(33)	0.17	1	U	-.GPVPHYAEADIVTL.Q	
4109	741.3853	1480.7559	1480.7562	-0.17	1	(48)	0.0059	1	U	-.GPVPHYAEADIVTL.Q	
4111	741.3858	1480.7570	1480.7562	0.58	1	52	0.0027	1	U	-.GPVPHYAEADIVTL.Q	
4160	749.8380	1497.6615	1497.6624	-0.59	1	38	0.082	1	U	L.TDEQVIENAGEFF.R	
4161	749.8383	1497.6620	1497.6624	-0.26	1	(37)	0.13	1	U	L.TDEQVIENAGEFF.R	
4338	780.8915	1559.7684	1559.7692	-0.51	1	28	0.7	1	U	W.SRESEQRPPFSQL.H	
4344	**781.3663**	**1560.7181**	**1560.7225**	**-2.83**	**1**	**(23)**	**2.9**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4347	**781.3671**	**1560.7196**	**1560.7225**	**-1.88**	**1**	**(31)**	**0.43**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4349	**781.3680**	**1560.7214**	**1560.7225**	**-0.71**	**1**	**(23)**	**2.8**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4350	**781.3680**	**1560.7215**	**1560.7225**	**-0.63**	**1**	**(23)**	**2.7**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4351	**781.3682**	**1560.7218**	**1560.7225**	**-0.48**	**1**	**(22)**	**3.7**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4352	**781.3683**	**1560.7220**	**1560.7225**	**-0.31**	**1**	**(31)**	**0.42**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4353	**781.3684**	**1560.7223**	**1560.7225**	**-0.16**	**1**	**(29)**	**0.81**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4354	**781.3687**	**1560.7229**	**1560.7225**	**0.22**	**1**	**(28)**	**0.9**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4355	**781.3690**	**1560.7235**	**1560.7225**	**0.62**	**1**	**(30)**	**0.53**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4357	**781.3693**	**1560.7240**	**1560.7225**	**0.93**	**1**	**(21)**	**4.9**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4359	**781.3695**	**1560.7243**	**1560.7225**	**1.17**	**1**	**(23)**	**2.9**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4360	**781.3695**	**1560.7245**	**1560.7225**	**1.25**	**1**	**(36)**	**0.14**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4361	**781.3697**	**1560.7248**	**1560.7225**	**1.48**	**1**	**(23)**	**2.9**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4362	**781.3699**	**1560.7253**	**1560.7225**	**1.79**	**1**	**(36)**	**0.15**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4363	**781.3701**	**1560.7257**	**1560.7225**	**2.03**	**1**	**(26)**	**1.5**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4364	**781.3735**	**1560.7324**	**1560.7225**	**6.32**	**1**	**(33)**	**0.28**	**1**	**U**	**-.GPVPHYAEADIVTL.Q + Phospho_STY (STY)**	**pY569**
4384	524.2794	1569.8163	1569.8151	0.74	2	(35)	0.12	1		F.KEKLGEGQFGEVHL.C	
4385	785.9154	1569.8163	1569.8151	0.75	2	(50)	0.0035	1		F.KEKLGEGQFGEVHL.C	
4386	524.2795	1569.8168	1569.8151	1.08	2	(24)	1.4	1		F.KEKLGEGQFGEVHL.C	
4387	785.9164	1569.8183	1569.8151	2.06	2	56	0.00085	1		F.KEKLGEGQFGEVHL.C	
4491	799.3608	1596.7070	1596.7090	-1.25	1	51	0.0051	1	U	F.GEVHLCEVDSPQDL.V	
4492	799.3613	1596.7081	1596.7090	-0.56	1	(48)	0.0092	1	U	F.GEVHLCEVDSPQDL.V	
4628	825.3842	1648.7539	1648.7580	-2.52	1	(58)	0.001	1	U	F.GQLTDEQVIENAGEF.F	
4629	825.3852	1648.7558	1648.7580	-1.34	1	91	4.4e-07	1	U	F.GQLTDEQVIENAGEF.F	
4630	825.3854	1648.7562	1648.7580	-1.11	1	(76)	1.6e-05	1	U	F.GQLTDEQVIENAGEF.F	
4631	550.5931	1648.7574	1648.7580	-0.38	1	(53)	0.0028	1	U	F.GQLTDEQVIENAGEF.F	
4632	825.3860	1648.7574	1648.7580	-0.37	1	(87)	1.2e-06	1	U	F.GQLTDEQVIENAGEF.F	
4633	825.3861	1648.7577	1648.7580	-0.23	1	(70)	6.3e-05	1	U	F.GQLTDEQVIENAGEF.F	
4634	825.3864	1648.7583	1648.7580	0.14	1	(63)	0.00033	1	U	F.GQLTDEQVIENAGEF.F	
4635	825.3868	1648.7591	1648.7580	0.66	1	(40)	0.067	1	U	F.GQLTDEQVIENAGEF.F	
4636	825.3869	1648.7592	1648.7580	0.73	1	(25)	1.9	1	U	F.GQLTDEQVIENAGEF.F	
4637	550.5938	1648.7594	1648.7580	0.84	1	(57)	0.0014	1	U	F.GQLTDEQVIENAGEF.F	
4638	825.3873	1648.7601	1648.7580	1.25	1	(91)	4.9e-07	1	U	F.GQLTDEQVIENAGEF.F	
4640	825.3879	1648.7612	1648.7580	1.92	1	(55)	0.002	1	U	F.GQLTDEQVIENAGEF.F	
4642	825.3882	1648.7619	1648.7580	2.36	1	(47)	0.012	1	U	F.GQLTDEQVIENAGEF.F	

4643	825.8776	1649.7407	1649.7421	-0.83	1	(86)	1.6e-06	1	U	F.GQLTDEQVIENAGEF.F + Deamidated (NQ)	
4644	825.8783	1649.7420	1649.7421	-0.00	1	(86)	1.8e-06	1	U	F.GQLTDEQVIENAGEF.F + Deamidated (NQ)	
5018	898.9198	1795.8250	1795.8265	-0.79	2	80	6.8e-06	1	U	F.GQLTDEQVIENAGEFF.R	
5019	898.9199	1795.8252	1795.8265	-0.72	2	(33)	0.33	1	U	F.GQLTDEQVIENAGEFF.R	
5020	898.9206	1795.8266	1795.8265	0.09	2	(36)	0.16	1	U	F.GQLTDEQVIENAGEFF.R	
5021	898.9213	1795.8281	1795.8265	0.92	2	(68)	9.7e-05	1	U	F.GQLTDEQVIENAGEFF.R	
5033	901.3751	1800.7357	1800.7369	-0.67	1	(68)	9.1e-05	1	U	L.GVCVQDDPLCMITDY.M + Oxidation (M)	
5034	901.3773	1800.7401	1800.7369	1.77	1	110	6.7e-09	1	U	L.GVCVQDDPLCMITDY.M + Oxidation (M)	
5086	925.8741	1849.7336	1849.7321	0.81	2	(28)	0.85	1	U	L.CMITDYMENGDLNQF.L	
5087	925.8745	1849.7345	1849.7321	1.28	2	(39)	0.072	1	U	L.CMITDYMENGDLNQF.L	
5088	926.3649	1850.7152	1850.7161	-0.50	2	(56)	0.0012	1	U	L.CMITDYMENGDLNQF.L + Deamidated (NQ)	
5089	926.3653	1850.7160	1850.7161	-0.03	2	(70)	5.2e-05	1	U	L.CMITDYMENGDLNQF.L + Deamidated (NQ)	
5111	933.8669	1865.7192	1865.7270	-4.19	2	83	2.9e-06	1	U	L.CMITDYMENGDLNQF.L + Oxidation (M)	
5112	933.8702	1865.7259	1865.7270	-0.59	2	(56)	0.0014	1	U	L.CMITDYMENGDLNQF.L + Oxidation (M)	
5113	933.8716	1865.7286	1865.7270	0.85	2	(79)	6.2e-06	1	U	L.CMITDYMENGDLNQF.L + Oxidation (M)	
5114	933.8727	1865.7309	1865.7270	2.09	2	(51)	0.004	1	U	L.CMITDYMENGDLNQF.L + Oxidation (M)	
5158	948.9533	1895.8921	1895.8935	-0.76	2	40	0.058	1	U	F.GEVHLCEVDSPQDLVSL.D	
5159	948.9535	1895.8924	1895.8935	-0.57	2	(30)	0.6	1	U	F.GEVHLCEVDSPQDLVSL.D	
5160	632.9724	1895.8954	1895.8935	1.00	2	(23)	3.6	1	U	F.GEVHLCEVDSPQDLVSL.D	
5404	1058.4686	2114.9227	2114.9215	0.56	2	60	0.00063	1	U	L.GEGQFGEVHLCEVDSPQDL.V	
5405	1058.4689	2114.9232	2114.9215	0.79	2	(52)	0.0041	1	U	L.GEGQFGEVHLCEVDSPQDL.V	
5623	868.4167	2602.2282	2602.2221	2.37	1	21	4.8	1	U	L.EDKAAEGAPGDGQAAQGPTISYPMLL.H + Oxidation (M)	
6090	1041.8407	3122.5003	3122.4978	0.78	2	(52)	0.0033	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPMLL.H	
6091	1041.8414	3122.5025	3122.4978	1.48	2	(64)	0.00019	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPMLL.H	
6097	1047.1650	3138.4733	3138.4928	-6.21	2	(68)	8e-05	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPMLL.H + Oxidation (M)	
6098	1047.1682	3138.4828	3138.4928	-3.18	2	(58)	0.00081	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPMLL.H + Oxidation (M)	
6099	1047.1696	3138.4869	3138.4928	-1.88	2	(68)	7.4e-05	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPMLL.H + Oxidation (M)	
6100	1047.1711	3138.4916	3138.4928	-0.37	2	85	1.6e-06	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPMLL.H + Oxidation (M)	
6101	1047.1711	3138.4916	3138.4928	-0.37	2	(68)	7e-05	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPMLL.H + Oxidation (M)	
6102	1047.1726	3138.4960	3138.4928	1.03	2	(72)	3.1e-05	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPMLL.H + Oxidation (M)	
6103	1047.1731	3138.4975	3138.4928	1.50	2	(65)	0.00015	1	U	L.SAHQLEDKAAEGAPGDGQAAQGPTISYPMLL.H + Oxidation (M)	
6183	**1133.5702**	**3397.6887**	**3397.6933**	**-1.33**	**1**	**(22)**	**1.8**	**1**	**U**	**L.QGVTGGNTYAVPALPPGAVGDGPPRVDFPRSRL.R + Phospho_STY (STY)**	**pY586**
6184	1133.5707	3397.6902	3397.6933	-0.90	1	41	0.027	1	U	L.QGVTGGNTYAVPALPPGAVGDGPPRVDFPRSRL.R + Phospho_STY (STY)	

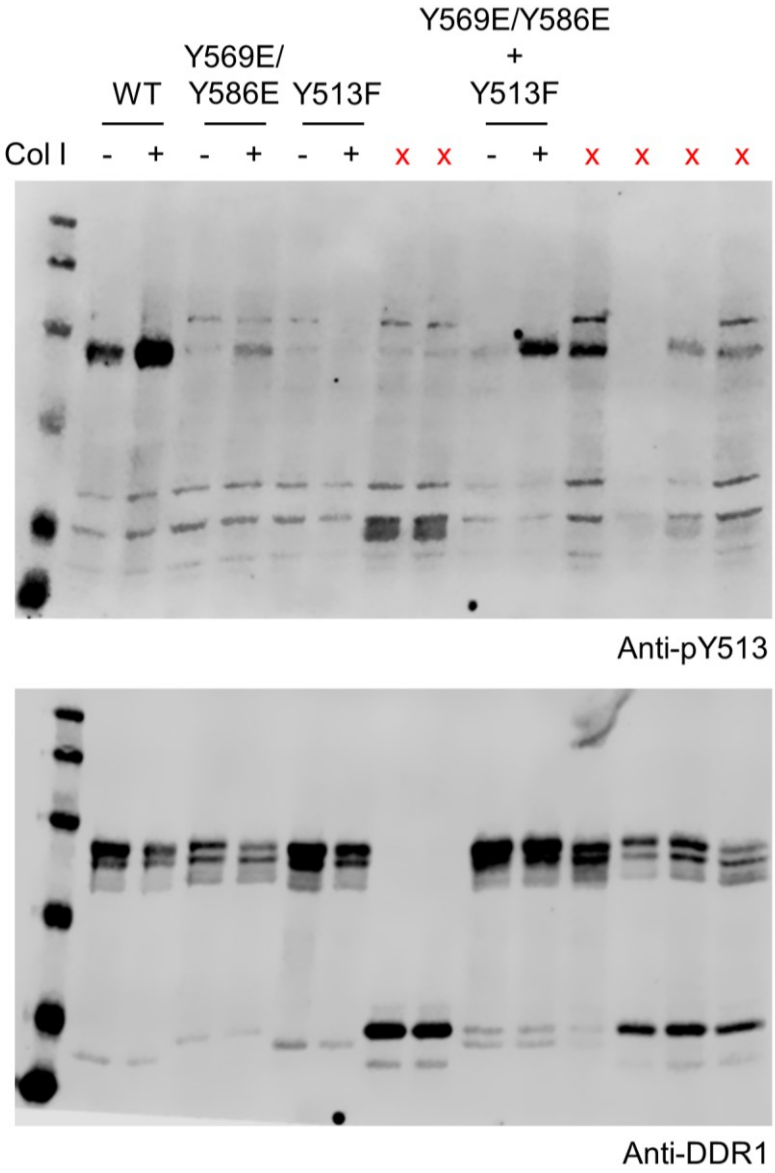

Figure 65. Uncropped and unedited image of Figure 59.

Reference List

Abbaspour Babaei, M., Kamalidehghan, B., Saleem, M., Huri, H. Z. & Ahmadipour, F. (2016) Receptor tyrosine kinase (c-Kit) inhibitors: a potential therapeutic target in cancer cells. *Drug Des Devel Ther*, 10, 2443-59.

Abbonante, V., Gruppi, C., Rubel, D., Gross, O., Moratti, R. & Balduini, A. (2013) Discoidin domain receptor 1 protein is a novel modulator of megakaryocyte-collagen interactions. *J Biol Chem*, 288(23), 16738-46.

Abdulhussein, R., McFadden, C., Fuentes-Prior, P. & Vogel, W. F. (2004) Exploring the collagen-binding site of the DDR1 tyrosine kinase receptor. *J Biol Chem*, 279(30), 31462-70.

Adams, J. A. (2001) Kinetic and catalytic mechanisms of protein kinases. *Chem Rev*, 101(8), 2271-90.

Adjei, A. A., LoRusso, P., Ribas, A., Sosman, J. A., Pavlick, A., Dy, G. K., Zhou, X., Gangolli, E., Kneissl, M., Faucette, S., Neuwirth, R. & Bózon, V. (2017) A phase I dose-escalation study of TAK-733, an investigational oral MEK inhibitor, in patients with advanced solid tumors. *Invest New Drugs*, 35(1), 47-58.

Afonso, P. V., McCann, C. P., Kapnick, S. M. & Parent, C. A. (2013) Discoidin domain receptor 2 regulates neutrophil chemotaxis in 3D collagen matrices. *Blood*, 121(9), 1644-50.

Agarwal, G., Mihai, C. & Iscru, D. F. (2007) Interaction of discoidin domain receptor 1 with collagen type 1. *J Mol Biol*, 367(2), 443-55.

Aguilera, K. Y., Huang, H., Du, W., Hagopian, M. M., Wang, Z., Hinz, S., Hwang, T. H., Wang, H., Fleming, J. B., Castrillon, D. H., Ren, X., Ding, K. & Brekken, R. A. (2017) Inhibition of Discoidin Domain Receptor 1 Reduces Collagen-mediated Tumorigenicity in Pancreatic Ductal Adenocarcinoma. *Mol Cancer Ther*, 16(11), 2473-2485.

Al-Kindi, A., Kizhakkedath, P., Xu, H., John, A., Sayegh, A. A., Ganesh, A., Al-Awadi, M., Al-Anbouri, L., Al-Gazali, L., Leitinger, B. & Ali, B. R. (2014) A novel mutation in DDR2 causing spondylo-meta-epiphyseal dysplasia with short limbs and abnormal calcifications (SMED-SL) results in defective intra-cellular trafficking. *BMC Med Genet*, 15, 42.

Ali, B. R., Xu, H., Akawi, N. A., John, A., Karuvantevida, N. S., Langer, R., Al-Gazali, L. & Leitinger, B. (2010) Trafficking defects and loss of ligand binding are the underlying causes of all reported DDR2 missense mutations found in SMED-SL patients. *Hum Mol Genet*, 19(11), 2239-50.

Alves, F., Saupe, S., Ledwon, M., Schaub, F., Hiddemann, W. & Vogel, W. F. (2001) Identification of two novel, kinase-deficient variants of discoidin domain receptor 1: differential expression in human colon cancer cell lines. *Faseb j*, 15(7), 1321-3.

Alves, F., Vogel, W., Mossie, K., Millauer, B., Höfler, H. & Ullrich, A. (1995) Distinct structural characteristics of discoidin I subfamily receptor tyrosine kinases and complementary expression in human cancer. *Oncogene*, 10(3), 609-18.

Ambrogio, C., Gómez-López, G., Falcone, M., Vidal, A., Nadal, E., Crosetto, N., Blasco, R. B., Fernández-Marcos, P. J., Sánchez-Céspedes, M., Ren, X., Wang, Z., Ding, K., Hidalgo, M., Serrano, M., Villanueva, A., Santamaría, D. & Barbacid, M. (2016) Combined inhibition of DDR1 and Notch signaling is a therapeutic strategy for KRAS-driven lung adenocarcinoma. *Nat Med*, 22(3), 270-7.

André, F., O'Regan, R., Ozguroglu, M., Toi, M., Xu, B., Jerusalem, G., Masuda, N., Wilks, S., Arena, F., Isaacs, C., Yap, Y. S., Papai, Z., Lang, I., Armstrong, A., Lerzo, G., White, M., Shen, K., Litton, J., Chen, D., Zhang, Y., Ali, S., Taran, T. & Gianni, L. (2014) Everolimus for women with trastuzumab-resistant, HER2-positive, advanced breast cancer (BOLERO-3): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet Oncol*, 15(6), 580-91.

Arias-Salgado, E. G., Lizano, S., Sarkar, S., Brugge, J. S., Ginsberg, M. H. & Shattil, S. J. (2003) Src kinase activation by direct interaction with the integrin beta cytoplasmic domain. *Proc Natl Acad Sci U S A*, 100(23), 13298-302.

Avivi-Green, C., Singal, M. & Vogel, W. F. (2006) Discoidin domain receptor 1-deficient mice are resistant to bleomycin-induced lung fibrosis. *Am J Respir Crit Care Med*, 174(4), 420-7.

Azizi, R., Salemi, Z., Fallahian, F. & Aghaei, M. (2019) Inhibition of didscoidin domain receptor 1 reduces epithelial-mesenchymal transition and induce cell-cycle arrest and apoptosis in prostate cancer cell lines. *J Cell Physiol*, 234(11), 19539-19552.

Bargal, R., Cormier-Daire, V., Ben-Neriah, Z., Le Merrer, M., Sosna, J., Melki, J., Zangen, D. H., Smithson, S. F., Borochowitz, Z., Belostotsky, R. & Raas-Rothschild, A. (2009) Mutations in DDR2 gene cause SMED with short limbs and abnormal calcifications. *Am J Hum Genet*, 84(1), 80-4.

Barouch-Bentov, R. & Sauer, K. (2011) Mechanisms of drug resistance in kinases. *Expert Opinion on Investigational Drugs*, 20(2), 153-208.

Barry, P. H. & Lynch, J. W. (2005) Ligand-gated channels. *IEEE Trans Nanobioscience*, 4(1), 70-80.

Bayer, S. V., Grither, W. R., Brenot, A., Hwang, P. Y., Barcus, C. E., Ernst, M., Pence, P., Walter, C., Pathak, A. & Longmore, G. D. (2019) DDR2 controls breast tumor stiffness and metastasis by regulating integrin mediated mechanotransduction in CAFs. *Elife*, 8.

Bayliss, R., Haq, T. & Yeoh, S. (2015) The Ys and wherefores of protein kinase autoinhibition. *Biochim Biophys Acta*, 1854(10 Pt B), 1586-94.

Bean, J., Riely, G. J., Balak, M., Marks, J. L., Ladanyi, M., Miller, V. A. & Pao, W. (2008) Acquired resistance to epidermal growth factor receptor kinase inhibitors associated with a novel T854A mutation in a patient with EGFR-mutant lung adenocarcinoma. *Clin Cancer Res*, 14(22), 7519-25.

Bekesova, S., Komis, G., Krenek, P., Vyplelova, P., Ovecka, M., Luptovciak, I., Illes, P., Kucharova, A. & Samaj, J. (2015) Monitoring protein phosphorylation by acrylamide pendant Phos-Tag in various plants. *Front Plant Sci*, 6, 336.

Belov, A. A. & Mohammadi, M. (2012) Grb2, a double-edged sword of receptor tyrosine kinase signaling. *Sci Signal*, 5(249), pe49.

Belsches, A. P., Haskell, M. D. & Parsons, S. J. (1997) Role of c-Src tyrosine kinase in EGF-induced mitogenesis. *Front Biosci*, 2, d501-18.

Bergeron, J. J., Di Guglielmo, G. M., Dahan, S., Dominguez, M. & Posner, B. I. (2016) Spatial and Temporal Regulation of Receptor Tyrosine Kinase Activation and Intracellular Signal Transduction. *Annu Rev Biochem*, 85, 573-97.

Bertelsen, V. & Stang, E. (2014) The Mysterious Ways of ErbB2/HER2 Trafficking. *Membranes (Basel)*, 4(3), 424-46.

Bhullar, K. S., Lagarón, N. O., McGowan, E. M., Parmar, I., Jha, A., Hubbard, B. P. & Rupasinghe, H. P. V. (2018) Kinase-targeted cancer therapies: progress, challenges and future directions. *Mol Cancer*, 17(1), 48.

Binns, K. L., Taylor, P. P., Sicheri, F., Pawson, T. & Holland, S. J. (2000) Phosphorylation of tyrosine residues in the kinase domain and juxtamembrane region regulates the biological and catalytic activities of Eph receptors. *Mol Cell Biol*, 20(13), 4791-805.

Bissig, C. & Gruenberg, J. (2013) Lipid sorting and multivesicular endosome biogenesis. *Cold Spring Harb Perspect Biol*, 5(10), a016816.

Bocharov, E. V., Mineev, K. S., Volynsky, P. E., Ermolyuk, Y. S., Tkach, E. N., Sobol, A. G., Chupin, V. V., Kirpichnikov, M. P., Efremov, R. G. & Arseniev, A. S. (2008) Spatial structure of the dimeric transmembrane domain of the growth factor receptor ErbB2 presumably corresponding to the receptor active state. *J Biol Chem*, 283(11), 6950-6.

Bonifacino, J. S. & Traub, L. M. (2003) Signals for sorting of transmembrane proteins to endosomes and lysosomes. *Annu Rev Biochem*, 72, 395-447.

Borochowitz, Z., Langer, L. O., Jr., Gruber, H. E., Lachman, R., Katznelson, M. B. & Rimoin, D. L. (1993) Spondylo-meta-epiphyseal dysplasia (SMED), short limb-hand type: a congenital familial skeletal dysplasia with distinctive features and histopathology. *Am J Med Genet*, 45(3), 320-6.

Borza, C. M. & Pozzi, A. (2014) Discoidin domain receptors in disease. *Matrix Biol*, 34, 185-92.

Borza, C. M., Su, Y., Tran, T. L., Yu, L., Steyns, N., Temple, K. J., Skwark, M. J., Meiler, J., Lindsley, C. W., Hicks, B. R., Leitinger, B., Zent, R. & Pozzi, A. (2017) Discoidin domain receptor 1 kinase activity is required for regulating collagen IV synthesis. *Matrix Biol*, 57-58, 258-271.

Braccini, L., Ciraolo, E., Campa, C. C., Perino, A., Longo, D. L., Tibolla, G., Pregnolato, M., Cao, Y., Tassone, B., Damilano, F., Laffargue, M., Calautti, E., Falasca, M., Norata, G. D., Backer, J. M. & Hirsch, E. (2015) PI3K-C2γ is a Rab5 effector selectively controlling endosomal Akt2 activation downstream of insulin signalling. *Nat Commun*, 6, 7400.

Bradley, D. & Beltrao, P. (2019) Evolution of protein kinase substrate recognition at the active site. *PLoS Biol*, 17(6), e3000341.

Brahmer, J. R., Tykodi, S. S., Chow, L. Q., Hwu, W. J., Topalian, S. L., Hwu, P., Drake, C. G., Camacho, L. H., Kauh, J., Odunsi, K., Pitot, H. C., Hamid, O., Bhatia, S., Martins, R., Eaton, K., Chen, S., Salay, T. M., Alaparthy, S., Grosso, J. F., Korman, A. J., Parker, S. M., Agrawal, S., Goldberg, S. M., Pardoll, D. M., Gupta, A. & Wigginton, J. M. (2012) Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med*, 366(26), 2455-65.

Brodsky, B. & Persikov, A. V. (2005) Molecular structure of the collagen triple helix. *Adv Protein Chem*, 70, 301-39.

Bromann, P. A., Korkaya, H. & Courtneidge, S. A. (2004) The interplay between Src family kinases and receptor tyrosine kinases. *Oncogene*, 23(48), 7957-68.

Cabail, M. Z., Li, S., Lemmon, E., Bowen, M. E., Hubbard, S. R. & Miller, W. T. (2015) The insulin and IGF1 receptor kinase domains are functional dimers in the activated state. *Nat Commun*, 6, 6406.

Cann, A. D. & Kohanski, R. A. (1997) Cis-autophosphorylation of juxtamembrane tyrosines in the insulin receptor kinase domain. *Biochemistry*, 36(25), 7681-9.

Canning, P., Tan, L., Chu, K., Lee, S. W., Gray, N. S. & Bullock, A. N. (2014) Structural mechanisms determining inhibition of the collagen receptor DDR1 by selective and multi-targeted type II kinase inhibitors. *J Mol Biol*, 426(13), 2457-70.

Carafoli, F., Bihan, D., Stathopoulos, S., Konitsiotis, A. D., Kvansakul, M., Farndale, R. W., Leitinger, B. & Hohenester, E. (2009) Crystallographic insight into collagen recognition by discoidin domain receptor 2. *Structure*, 17(12), 1573-81.

Carafoli, F. & Hohenester, E. (2013) Collagen recognition and transmembrane signalling by discoidin domain receptors. *Biochim Biophys Acta*, 1834(10), 2187-94.

Carafoli, F., Mayer, M. C., Shiraishi, K., Pecheva, M. A., Chan, L. Y., Nan, R., Leitinger, B. & Hohenester, E. (2012) Structure of the discoidin domain receptor 1 extracellular region bound to an inhibitory Fab fragment reveals features important for signaling. *Structure*, 20(4), 688-97.

Chen, L., Marsiglia, W. M., Chen, H., Katigbak, J., Erdjument-Bromage, H., Kemble, D. J., Fu, L., Ma, J., Sun, G., Zhang, Y., Liang, G., Neubert, T. A., Li, X., Traaseth, N. J. & Mohammadi, M. (2020) Molecular basis for receptor tyrosine kinase A-loop tyrosine transphosphorylation. *Nat Chem Biol*, 16(3), 267-277.

Chen, Y., Huang, L., Qi, X. & Chen, C. (2019) Insulin Receptor Trafficking: Consequences for Insulin Sensitivity and Diabetes. *Int J Mol Sci*, 20(20).

Chen, Z. & Cole, P. A. (2015) Synthetic approaches to protein phosphorylation. *Curr Opin Chem Biol*, 28, 115-22.

Chiara, F., Bishayee, S., Heldin, C. H. & Demoulin, J. B. (2004) Autoinhibition of the platelet-derived growth factor beta-receptor tyrosine kinase by its C-terminal tail. *J Biol Chem*, 279(19), 19732-8.

Clark, P. (2014) Protease-mediated ectodomain shedding. *Thorax*, 69(7), 682-684.

Coelho, N. M., Arora, P. D., van Putten, S., Boo, S., Petrovic, P., Lin, A. X., Hinz, B. & McCulloch, C. A. (2017) Discoidin Domain Receptor 1 Mediates Myosin-Dependent Collagen Contraction. *Cell Rep*, 18(7), 1774-1790.

Corcoran, D. S., Juskaite, V., Xu, Y., Gorlitz, F., Alexandrov, Y., Dunsby, C., French, P. M. W. & Leitinger, B. (2019) DDR1 autophosphorylation is a result of aggregation into dense clusters. *Sci Rep*, 9(1), 17104.

Couvelard, A., Hu, J., Steers, G., O'Toole, D., Sauvanet, A., Belghiti, J., Bedossa, P., Gatter, K., Ruszniewski, P. & Pezzella, F. (2006) Identification of potential therapeutic targets by gene-expression profiling in pancreatic endocrine tumors. *Gastroenterology*, 131(5), 1597-610.

Craddock, B. P., Cotter, C. & Miller, W. T. (2007) Autoinhibition of the insulin-like growth factor I receptor by the juxtamembrane region. *FEBS Lett*, 581(17), 3235-40.

Craven, G. B., Affron, D. P., Allen, C. E., Matthies, S., Greener, J. G., Morgan, R. M. L., Tate, E. W., Armstrong, A. & Mann, D. J. (2018) High-Throughput Kinetic Analysis for Target-Directed Covalent Ligand Discovery. *Angew Chem Int Ed Engl*, 57(19), 5257-5261.

Curat, C. A. & Vogel, W. F. (2002) Discoidin domain receptor 1 controls growth and adhesion of mesangial cells. *J Am Soc Nephrol*, 13(11), 2648-56.

Dar, A. C., Dever, T. E. & Sicheri, F. (2005) Higher-order substrate recognition of eIF2alpha by the RNA-dependent protein kinase PKR. *Cell*, 122(6), 887-900.

Das, S., Ongusaha, P. P., Yang, Y. S., Park, J. M., Aaronson, S. A. & Lee, S. W. (2006) Discoidin domain receptor 1 receptor tyrosine kinase induces cyclooxygenase-2 and promotes chemoresistance through nuclear factor-kappaB pathway activation. *Cancer Res*, 66(16), 8123-30.

Davids, M. S. & Brown, J. R. (2014) Ibrutinib: a first in class covalent inhibitor of Bruton's tyrosine kinase. *Future Oncol*, 10(6), 957-67.

Davies, H., Hunter, C., Smith, R., Stephens, P., Greenman, C., Bignell, G., Teague, J., Butler, A., Edkins, S., Stevens, C., Parker, A., O'Meara, S., Avis, T., Barthorpe, S., Brackenbury, L., Buck, G., Clements, J., Cole, J., Dicks, E., Edwards, K., Forbes, S., Gorton, M., Gray, K., Halliday, K., Harrison, R., Hills, K., Hinton, J., Jones, D., Kosmidou, V., Laman, R., Lugg, R., Menzies, A., Perry, J., Petty, R., Raine, K., Shepherd, R., Small, A., Solomon, H., Stephens, Y., Tofts, C., Varian, J., Webb, A., West, S., Widaa, S., Yates, A., Brasseur, F., Cooper, C. S., Flanagan, A. M., Green, A., Knowles, M., Leung, S. Y., Looijenga, L. H., Malkowicz, B., Pierotti, M. A., Teh, B. T., Yuen, S. T., Lakhani, S. R., Easton, D. F., Weber, B. L., Goldstraw, P., Nicholson, A. G., Wooster, R., Stratton, M. R. & Futreal, P. A. (2005) Somatic mutations of the protein kinase gene family in human lung cancer. *Cancer Res*, 65(17), 7591-5.

Day, E., Waters, B., Spiegel, K., Alnadaf, T., Manley, P. W., Buchdunger, E., Walker, C. & Jarai, G. (2008) Inhibition of collagen-induced discoidin domain receptor 1 and 2 activation by imatinib, nilotinib and dasatinib. *Eur J Pharmacol*, 599(1-3), 44-53.

de Melker, A. A., van der Horst, G., Calafat, J., Jansen, H. & Borst, J. (2001) c-Cbl ubiquitinates the EGF receptor at the plasma membrane and remains receptor associated throughout the endocytic route. *J Cell Sci*, 114(Pt 11), 2167-78.

Dejmek, J., Dib, K., Jönsson, M. & Andersson, T. (2003) Wnt-5a and G-protein signaling are required for collagen-induced DDR1 receptor activation and normal mammary cell adhesion. *Int J Cancer*, 103(3), 344-51.

Deng, Y., Zhao, F., Hui, L., Li, X., Zhang, D., Lin, W., Chen, Z. & Ning, Y. (2017) Suppressing miR-199a-3p by promoter methylation contributes to tumor aggressiveness and cisplatin resistance of ovarian cancer through promoting DDR1 expression. *J Ovarian Res*, 10(1), 50.

Di Marco, E., Cutuli, N., Guerra, L., Cancedda, R. & De Luca, M. (1993) Molecular cloning of trkE, a novel trk-related putative tyrosine kinase receptor isolated from normal human keratinocytes and widely expressed by normal human tissues. *J Biol Chem*, 268(32), 24290-5.

Ding, L., Getz, G., Wheeler, D. A., Mardis, E. R., McLellan, M. D., Cibulskis, K., Sougnez, C., Greulich, H., Muzny, D. M., Morgan, M. B., Fulton, L., Fulton, R. S., Zhang, Q., Wendl, M. C., Lawrence, M. S., Larson, D. E., Chen, K., Dooling, D. J., Sabo, A., Hawes, A. C., Shen, H., Jhangiani, S. N., Lewis, L. R., Hall, O., Zhu, Y., Mathew, T., Ren, Y., Yao, J., Scherer, S. E., Clerc, K., Metcalf, G. A., Ng, B., Milosavljevic, A., Gonzalez-Garay, M. L., Osborne, J. R., Meyer, R., Shi, X., Tang, Y., Koboldt, D. C., Lin, L., Abbott, R., Miner, T. L., Pohl, C., Fewell, G., Haipek, C., Schmidt, H., Dunford-Shore, B. H., Kraja, A., Crosby, S. D., Sawyer, C. S., Vickery, T., Sander, S., Robinson, J., Winckler, W., Baldwin, J., Chirieac, L. R., Dutt, A., Fennell, T., Hanna, M., Johnson, B. E., Onofrio, R. C., Thomas, R. K., Tonon, G., Weir, B. A., Zhao, X., Ziaugra, L., Zody, M. C., Giordano, T., Orringer, M. B., Roth, J. A., Spitz, M. R., Wistuba, I. I., Ozenberger, B., Good, P. J., Chang, A. C., Beer, D. G., Watson, M. A., Ladanyi, M., Broderick, S., Yoshizawa, A., Travis, W. D., Pao, W., Province, M. A., Weinstock, G. M., Varmus, H. E., Gabriel, S. B., Lander, E. S., Gibbs, R. A., Meyerson, M. & Wilson, R. K. (2008) Somatic mutations affect key pathways in lung adenocarcinoma. *Nature*, 455(7216), 1069-1075.

Dorison, A., Dussaule, J. C. & Chatziantoniou, C. (2017) The Role of Discoidin Domain Receptor 1 in Inflammation, Fibrosis and Renal Disease. *Nephron*, 137(3), 212-220.

Du, Z. & Lovly, C. M. (2018) Mechanisms of receptor tyrosine kinase activation in cancer. *Mol Cancer*, 17(1), 58.

Ebert, P. J. R., Cheung, J., Yang, Y., McNamara, E., Hong, R., Moskalenko, M., Gould, S. E., Maecker, H., Irving, B. A., Kim, J. M., Belvin, M. & Mellman, I. (2016) MAP Kinase Inhibition Promotes T Cell and Anti-tumor Activity in Combination with PD-L1 Checkpoint Blockade. *Immunity*, 44(3), 609-621.

El-Hashim, A. Z., Khajah, M. A., Renno, W. M., Babyson, R. S., Uddin, M., Benter, I. F., Ezeamuzie, C. & Akhtar, S. (2017) Src-dependent EGFR transactivation regulates lung inflammation via downstream signaling involving ERK1/2, PI3Kδ/Akt and NFκB induction in a murine asthma model. *Scientific Reports*, 7(1), 9919.

Emsley, P. & Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, 60(Pt 12 Pt 1), 2126-32.

Endicott, J. A., Noble, M. E. & Johnson, L. N. (2012) The structural basis for control of eukaryotic protein kinases. *Annu Rev Biochem*, 81, 587-613.

Engelman, J. A., Zejnullahu, K., Mitsudomi, T., Song, Y., Hyland, C., Park, J. O., Lindeman, N., Gale, C. M., Zhao, X., Christensen, J., Kosaka, T., Holmes, A. J., Rogers, A. M., Cappuzzo, F., Mok, T., Lee, C., Johnson, B. E., Cantley, L. C. & Jänne, P. A. (2007) MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science*, 316(5827), 1039-43.

Evans, P. R. & Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr*, 69(Pt 7), 1204-14.

Faraci-Orf, E., McFadden, C. & Vogel, W. F. (2006) DDR1 signaling is essential to sustain Stat5 function during lactogenesis. *J Cell Biochem*, 97(1), 109-21.

Fasano, M., Della Corte, C. M., Califano, R., Capuano, A., Troiani, T., Martinelli, E., Ciardiello, F. & Morgillo, F. (2014) Type III or allosteric kinase inhibitors for the treatment of non-small cell lung cancer. *Expert Opin Investig Drugs*, 23(6), 809-21.

Favelyukis, S., Till, J. H., Hubbard, S. R. & Miller, W. T. (2001) Structure and autoregulation of the insulin-like growth factor 1 receptor kinase. *Nat Struct Biol*, 8(12), 1058-63.

Ferguson, F. M. & Gray, N. S. (2018) Kinase inhibitors: the road ahead. *Nat Rev Drug Discov*, 17(5), 353-377.

Finger, C., Escher, C. & Schneider, D. (2009) The single transmembrane domains of human receptor tyrosine kinases encode self-interactions. *Sci Signal*, 2(89), ra56.

Flamant, M., Placier, S., Rodenas, A., Curat, C. A., Vogel, W. F., Chatziantoniou, C. & Dussaule, J. C. (2006) Discoidin domain receptor 1 null mice are protected against hypertension-induced renal disease. *J Am Soc Nephrol*, 17(12), 3374-81.

Franco, C., Ahmad, P. J., Hou, G., Wong, E. & Bendeck, M. P. (2010) Increased cell and matrix accumulation during atherogenesis in mice with vessel wall-specific deletion of discoidin domain receptor 1. *Circ Res*, 106(11), 1775-83.

Franco, C., Britto, K., Wong, E., Hou, G., Zhu, S. N., Chen, M., Cybulsky, M. I. & Bendeck, M. P. (2009) Discoidin domain receptor 1 on bone marrow-derived cells promotes macrophage accumulation during atherogenesis. *Circ Res*, 105(11), 1141-8.

Franco, C., Hou, G., Ahmad, P. J., Fu, E. Y., Koh, L., Vogel, W. F. & Bendeck, M. P. (2008) Discoidin domain receptor 1 (ddr1) deletion decreases atherosclerosis by accelerating matrix accumulation and reducing inflammation in low-density lipoprotein receptor-deficient mice. *Circ Res*, 102(10), 1202-11.

Freed, D. M., Bessman, N. J., Kiyatkin, A., Salazar-Cavazos, E., Byrne, P. O., Moore, J. O., Valley, C. C., Ferguson, K. M., Leahy, D. J., Lidke, D. S. & Lemmon, M. A. (2017) EGFR Ligands Differentially Stabilize Receptor Dimers to Specify Signaling Kinetics. *Cell*, 171(3), 683-695.e18.

Fu, H. L., Sohail, A., Valiathan, R. R., Wasinski, B. D., Kumarasiri, M., Mahasenan, K. V., Bernardo, M. M., Tokmina-Roszyk, D., Fields, G. B., Mobashery, S. & Fridman, R.

(2013) Shedding of discoidin domain receptor 1 by membrane-type matrix metalloproteinases. *J Biol Chem*, 288(17), 12114-29.

Fuentes-Prior, P., Fujikawa, K. & Pratt, K. P. (2002) New insights into binding interfaces of coagulation factors V and VIII and their homologues lessons from high resolution crystal structures. *Curr Protein Pept Sci*, 3(3), 313-39.

Furuya, N., Momose, T., Katsuno, K., Fushimi, N., Muranaka, H., Handa, C., Ozawa, T. & Kinoshita, T. (2017) The juxtamembrane region of TrkA kinase is critical for inhibitor selectivity. *Bioorg Med Chem Lett*, 27(5), 1233-1236.

Gadiya, M. & Chakraborty, G. (2018) Signaling by discoidin domain receptor 1 in cancer metastasis. *Cell Adh Migr*, 12(4), 315-323.

Gainor, J. F., Shaw, A. T., Sequist, L. V., Fu, X., Azzoli, C. G., Piotrowska, Z., Huynh, T. G., Zhao, L., Fulton, L., Schultz, K. R., Howe, E., Farago, A. F., Sullivan, R. J., Stone, J. R., Digumarthy, S., Moran, T., Hata, A. N., Yagi, Y., Yeap, B. Y., Engelman, J. A. & Mino-Kenudson, M. (2016) EGFR Mutations and ALK Rearrangements Are Associated with Low Response Rates to PD-1 Pathway Blockade in Non-Small Cell Lung Cancer: A Retrospective Analysis. *Clin Cancer Res*, 22(18), 4585-93.

Gajiwala, K. S. (2013) EGFR: tale of the C-terminal tail. *Protein Sci*, 22(7), 995-9.

Gao, M., Duan, L., Luo, J., Zhang, L., Lu, X., Zhang, Y., Zhang, Z., Tu, Z., Xu, Y., Ren, X. & Ding, K. (2013) Discovery and optimization of 3-(2-(Pyrazolo[1,5-a]pyrimidin-6-yl)ethynyl)benzamides as novel selective and orally bioavailable discoidin domain receptor 1 (DDR1) inhibitors. *J Med Chem*, 56(8), 3281-95.

Garay, C., Judge, G., Lucarelli, S., Bautista, S., Pandey, R., Singh, T. & Antonescu, C. N. (2015) Epidermal growth factor-stimulated Akt phosphorylation requires clathrin or ErbB2 but not receptor endocytosis. *Mol Biol Cell*, 26(19), 3504-19.

Gas, C., Canales-Rodríguez, E. J., Radua, J., Abasolo, N., Cortés, M. J., Salvadó, E., Muntané, G., Alemán-Gómez, Y., Julià, T., Marsal, S., Sanjuan, J., Guitart, M., Costas, J., Martorell, L., Pomarol-Clotet, E. & Vilella, E. (2019) Discoidin domain receptor 1 gene variants are associated with decreased white matter fractional anisotropy and decreased processing speed in schizophrenia. *Journal of Psychiatric Research*, 110, 74-82.

Gilburt, J. A. H., Sarkar, H., Sheldrake, P., Blagg, J., Ying, L. & Dodson, C. A. (2017) Dynamic Equilibrium of the Aurora A Kinase Activation Loop Revealed by Single-Molecule Spectroscopy. *Angewandte Chemie International Edition*, 56(38), 11409-11414.

Gillooly, D. J., Morrow, I. C., Lindsay, M., Gould, R., Bryant, N. J., Gaullier, J. M., Parton, R. G. & Stenmark, H. (2000) Localization of phosphatidylinositol 3-phosphate in yeast and mammalian cells. *Embo j*, 19(17), 4577-88.

Goh, L. K. & Sorkin, A. (2013) Endocytosis of receptor tyrosine kinases. *Cold Spring Harb Perspect Biol*, 5(5), a017459.

Grant, B. D. & Adams, J. A. (1996) Pre-steady-state kinetic analysis of cAMP-dependent protein kinase using rapid quench flow techniques. *Biochemistry*, 35(6), 2022-9.

Grither, W. R. & Longmore, G. D. (2018) Inhibition of tumor–microenvironment interaction and tumor invasion by small-molecule allosteric inhibitor of DDR2 extracellular domain. *Proceedings of the National Academy of Sciences*, 115(33), E7786-E7794.

Gross, O., Beirowski, B., Harvey, S. J., McFadden, C., Chen, D., Tam, S., Thorner, P. S., Smyth, N., Addicks, K., Bloch, W., Ninomiya, Y., Sado, Y., Weber, M. & Vogel, W. F. (2004) DDR1-deficient mice show localized subepithelial GBM thickening with focal loss of slit diaphragms and proteinuria. *Kidney Int*, 66(1), 102-11.

Gual, P., Le Marchand-Brustel, Y. & Tanti, J. F. (2005) Positive and negative regulation of insulin signaling through IRS-1 phosphorylation. *Biochimie*, 87(1), 99-109.

Guerrot, D., Kerroch, M., Placier, S., Vandermeersch, S., Trivin, C., Mael-Ainin, M., Chatziantoniou, C. & Dussaule, J. C. (2011) Discoidin domain receptor 1 is a major mediator of inflammation and fibrosis in obstructive nephropathy. *Am J Pathol*, 179(1), 83-91.

Guo, J., Zhang, Z. & Ding, K. (2020) A patent review of discoidin domain receptor 1 (DDR1) modulators (2014-present). *Expert Opinion on Therapeutic Patents*, 30(5), 341-350.

Ha, B. H., Simpson, M. A., Koleske, A. J. & Boggon, T. J. (2015) Structure of the ABL2/ARG kinase in complex with dasatinib. *Acta Crystallogr F Struct Biol Commun*, 71(Pt 4), 443-8.

Hachehouche, L. N., Chetoui, N. & Aoudjit, F. (2010) Implication of discoidin domain receptor 1 in T cell migration in three-dimensional collagen. *Mol Immunol*, 47(9), 1866-9.

Hall, M. D., Yasgar, A., Peryea, T., Braisted, J. C., Jadhav, A., Simeonov, A. & Coussens, N. P. (2016) Fluorescence polarization assays in high-throughput screening and drug discovery: a review. *Methods Appl Fluoresc*, 4(2), 022001.

Hanson, S. M., Georghiou, G., Thakur, M. K., Miller, W. T., Rest, J. S., Chodera, J. D. & Seeliger, M. A. (2019) What Makes a Kinase Promiscuous for Inhibitors? *Cell Chem Biol*, 26(3), 390-399 e5.

Hantschel, O., Rix, U. & Superti-Furga, G. (2008) Target spectrum of the BCR-ABL inhibitors imatinib, nilotinib and dasatinib. *Leuk Lymphoma*, 49(4), 615-9.

Hartmann, J. T., Haap, M., Kopp, H. G. & Lipp, H. P. (2009) Tyrosine kinase inhibitors - a review on pharmacology, metabolism and side effects. *Curr Drug Metab*, 10(5), 470-81.

Hedger, G., Sansom, M. S. & Koldso, H. (2015) The juxtamembrane regions of human receptor tyrosine kinases exhibit conserved interaction sites with anionic lipids. *Sci Rep*, 5, 9198.

Heinrich, M. C., Corless, C. L., Demetri, G. D., Blanke, C. D., von Mehren, M., Joensuu, H., McGreevey, L. S., Chen, C. J., Van den Abbeele, A. D., Druker, B. J., Kiese, B., Eisenberg, B., Roberts, P. J., Singer, S., Fletcher, C. D., Silberman, S., Dimitrijevic, S.

& Fletcher, J. A. (2003) Kinase mutations and imatinib response in patients with metastatic gastrointestinal stromal tumor. *J Clin Oncol*, 21(23), 4342-9.

Herbst, R. & Burden, S. J. (2000) The juxtamembrane region of MuSK has a critical role in agrin-mediated signaling. *Embo j*, 19(1), 67-77.

Higashiyama, S., Nanba, D., Nakayama, H., Inoue, H. & Fukuda, S. (2011) Ectodomain shedding and remnant peptide signalling of EGFRs and their ligands. *J Biochem*, 150(1), 15-22.

Hight-Warburton, W. & Parsons, M. (2019) Regulation of cell migration by alpha4 and alpha9 integrins. *Biochem J*, 476(4), 705-718.

Hilger, D., Masureel, M. & Kobilka, B. K. (2018) Structure and dynamics of GPCR signaling complexes. *Nature Structural & Molecular Biology*, 25(1), 4-12.

Hilton, H. N., Stanford, P. M., Harris, J., Oakes, S. R., Kaplan, W., Daly, R. J. & Ormandy, C. J. (2008) KIBRA interacts with discoidin domain receptor 1 to modulate collagen-induced signalling. *Biochim Biophys Acta*, 1783(3), 383-93.

Hochhaus, A. & Kantarjian, H. (2013) The development of dasatinib as a treatment for chronic myeloid leukemia (CML): from initial studies to application in newly diagnosed patients. *J Cancer Res Clin Oncol*, 139(12), 1971-84.

Hou, G., Vogel, W. & Bendeck, M. P. (2001) The discoidin domain receptor tyrosine kinase DDR1 in arterial wound repair. *J Clin Invest*, 107(6), 727-35.

Hou, G., Vogel, W. F. & Bendeck, M. P. (2002) Tyrosine kinase activity of discoidin domain receptor 1 is necessary for smooth muscle cell migration and matrix metalloproteinase expression. *Circ Res*, 90(11), 1147-9.

Hou, G., Wang, D. & Bendeck, M. P. (2012) Deletion of discoidin domain receptor 2 does not affect smooth muscle cell adhesion, migration, or proliferation in response to type I collagen. *Cardiovasc Pathol*, 21(3), 214-8.

Hu, H., Goltsov, A., Bown, J. L., Sims, A. H., Langdon, S. P., Harrison, D. J. & Faratian, D. (2013) Feedforward and feedback regulation of the MAPK and PI3K oscillatory circuit in breast cancer. *Cell Signal*, 25(1), 26-32.

Huang, Y., Bharill, S., Karandur, D., Peterson, S. M., Marita, M., Shi, X., Kaliszewski, M. J., Smith, A. W., Isacoff, E. Y. & Kuriyan, J. (2016) Molecular basis for multimerization in the activation of the epidermal growth factor receptor. *Elife*, 5.

Hubbard, S. R. (1997) Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *Embo j*, 16(18), 5572-81.

Hubbard, S. R. (2001) Theme and variations: juxtamembrane regulation of receptor protein kinases. *Mol Cell*, 8(3), 481-2.

Hubbard, S. R. (2004) Juxtamembrane autoinhibition in receptor tyrosine kinases. *Nat Rev Mol Cell Biol*, 5(6), 464-71.

Hubbard, S. R. (2013a) Structural biology: Insulin meets its receptor. *Nature*, 493(7431), 171-2.

Hubbard, S. R. (2013b) The insulin receptor: both a prototypical and atypical receptor tyrosine kinase. *Cold Spring Harb Perspect Biol*, 5(3), a008946.

Hubbard, S. R. & Miller, W. T. (2007) Receptor tyrosine kinases: mechanisms of activation and signaling. *Curr Opin Cell Biol*, 19(2), 117-23.

Hubbard, S. R., Wei, L., Ellis, L. & Hendrickson, W. A. (1994) Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature*, 372(6508), 746-54.

Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z. & Dunker, A. K. (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res*, 32(3), 1037-49.

Ichikawa, O., Osawa, M., Nishida, N., Goshima, N., Nomura, N. & Shimada, I. (2007) Structural basis of the collagen-binding mode of discoidin domain receptor 2. *Embo j*, 26(18), 4168-76.

Ikeda, K., Wang, L. H., Torres, R., Zhao, H., Olaso, E., Eng, F. J., Labrador, P., Klein, R., Lovett, D., Yancopoulos, G. D., Friedman, S. L. & Lin, H. C. (2002) Discoidin domain receptor 2 interacts with Src and Shc following its activation by type I collagen. *J Biol Chem*, 277(21), 19206-12.

Itoh, Y. (2018) Discoidin domain receptors: Microenvironment sensors that promote cellular migration and invasion. *Cell Adh Migr*, 12(4), 378-385.

Iwai, L. K., Payne, L. S., Luczynski, M. T., Chang, F., Xu, H., Clinton, R. W., Paul, A., Esposito, E. A., Gridley, S., Leitinger, B., Naegle, K. M. & Huang, P. H. (2013) Phosphoproteomics of collagen receptor networks reveals SHP-2 phosphorylation downstream of wild-type DDR2 and its lung cancer mutants. *Biochem J*, 454(3), 501-13.

James, J. R. & Vale, R. D. (2012) Biophysical mechanism of T-cell receptor triggering in a reconstituted system. *Nature*, 487(7405), 64-9.

Jeitany, M., Leroy, C., Tosti, P., Lafitte, M., Le Guet, J., Simon, V., Bonenfant, D., Robert, B., Grillet, F., Mollevi, C., El Messaoudi, S., Otandault, A., Canterel-Thouennon, L., Busson, M., Thierry, A. R., Martineau, P., Pannequin, J., Roche, S. & Sirvent, A. (2018) Inhibition of DDR1-BCR signalling by nilotinib as a new therapeutic strategy for metastatic colorectal cancer. *EMBO Mol Med*, 10(4).

Jin, H., Ham, I. H., Oh, H. J., Bae, C. A., Lee, D., Kim, Y. B., Son, S. Y., Chwae, Y. J., Han, S. U., Brekken, R. A. & Hur, H. (2018) Inhibition of Discoidin Domain Receptor 1 Prevents Stroma-Induced Peritoneal Metastasis in Gastric Carcinoma. *Mol Cancer Res*, 16(10), 1590-1600.

Jura, N., Endres, N. F., Engel, K., Deindl, S., Das, R., Lamers, M. H., Wemmer, D. E., Zhang, X. & Kuriyan, J. (2009) Mechanism for activation of the EGF receptor catalytic domain by the juxtamembrane segment. *Cell*, 137(7), 1293-307.

Jura, N., Zhang, X., Endres, N. F., Seeliger, M. A., Schindler, T. & Kuriyan, J. (2011) Catalytic control in the EGF receptor and its connection to general kinase regulatory mechanisms. *Mol Cell*, 42(1), 9-22.

Juskaite, V., Corcoran, D. S. & Leitinger, B. (2017) Collagen induces activation of DDR1 through lateral dimer association and phosphorylation between dimers. *Elife*, 6.

Kadler, K. E., Baldock, C., Bella, J. & Boot-Handford, R. P. (2007) Collagens at a glance. *J Cell Sci*, 120(Pt 12), 1955-8.

Kamohara, H., Yamashiro, S., Galligan, C. & Yoshimura, T. (2001) Discoidin domain receptor 1 isoform-a (DDR1alpha) promotes migration of leukocytes in three-dimensional collagen lattices. *Faseb j*, 15(14), 2724-6.

Kang, L., Shi, H., Liu, X., Zhang, C., Yao, Q., Wang, Y., Chang, C., Shi, J., Cao, J., Kong, J. & Chen, K. (2011) Arginine kinase is highly expressed in a resistant strain of silkworm (Bombyx mori, Lepidoptera): Implication of its role in resistance to Bombyx mori nucleopolyhedrovirus. *Comp Biochem Physiol B Biochem Mol Biol*, 158(3), 230-4.

Karn, T., Holtrich, U., Bräuninger, A., Böhme, B., Wolf, G., Rübsamen-Waigmann, H. & Strebhardt, K. (1993) Structure, expression and chromosomal mapping of TKT from man and mouse: a new subclass of receptor tyrosine kinases with a factor VIII-like domain. *Oncogene*, 8(12), 3433-40.

Karplus, P. A. & Diederichs, K. (2012) Linking crystallographic model and data quality. *Science*, 336(6084), 1030-3.

Kechagia, J. Z., Ivaska, J. & Roca-Cusachs, P. (2019) Integrins as biomechanical sensors of the microenvironment. *Nature Reviews Molecular Cell Biology*, 20(8), 457-473.

Keeley, A., Petri, L., Ábrányi-Balogh, P. & Keserű, G. M. (2020) Covalent fragment libraries in drug discovery. *Drug Discov Today*, 25(6), 983-996.

Kerroch, M., Alfieri, C., Dorison, A., Boffa, J. J., Chatziantoniou, C. & Dussaule, J. C. (2016) Protective effects of genetic inhibition of Discoidin Domain Receptor 1 in experimental renal disease. *Sci Rep*, 6, 21262.

Kerroch, M., Guerrot, D., Vandermeersch, S., Placier, S., Mesnard, L., Jouanneau, C., Rondeau, E., Ronco, P., Boffa, J. J., Chatziantoniou, C. & Dussaule, J. C. (2012) Genetic inhibition of discoidin domain receptor 1 protects mice against crescentic glomerulonephritis. *Faseb j*, 26(10), 4079-91.

Kim, H. G., Tan, L., Weisberg, E. L., Liu, F., Canning, P., Choi, H. G., Ezell, S. A., Wu, H., Zhao, Z., Wang, J., Mandinova, A., Griffin, J. D., Bullock, A. N., Liu, Q., Lee, S. W. & Gray, N. S. (2013) Discovery of a potent and selective DDR1 receptor tyrosine kinase inhibitor. *ACS Chem Biol*, 8(10), 2145-50.

Knighton, D. R., Zheng, J. H., Ten Eyck, L. F., Ashford, V. A., Xuong, N. H., Taylor, S. S. & Sowadski, J. M. (1991) Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science*, 253(5018), 407-14.

Kobayashi, T., Stang, E., Fang, K. S., de Moerloose, P., Parton, R. G. & Gruenberg, J. (1998) A lipid associated with the antiphospholipid syndrome regulates endosome structure and function. *Nature*, 392(6672), 193-7.

Koide, N. & Muramatsu, T. (1974) Endo-beta-N-acetylglucosaminidase acting on carbohydrate moieties of glycoproteins. Purification and properties of the enzyme from Diplococcus pneumoniae. *J Biol Chem*, 249(15), 4897-904.

Kondo, Y., Ognjenovic, J., Banerjee, S., Karandur, D., Merk, A., Kulhanek, K., Wong, K., Roose, J. P., Subramaniam, S. & Kuriyan, J. (2019) Cryo-EM structure of a dimeric B-Raf:14-3-3 complex reveals asymmetry in the active sites of B-Raf kinases. *Science*, 366(6461), 109-115.

Konitsiotis, A. D., Raynal, N., Bihan, D., Hohenester, E., Farndale, R. W. & Leitinger, B. (2008) Characterization of high affinity binding motifs for the discoidin domain receptor DDR2 in collagen. *J Biol Chem*, 283(11), 6861-8.

Koo, D. H., McFadden, C., Huang, Y., Abdulhussein, R., Friese-Hamim, M. & Vogel, W. F. (2006) Pinpointing phosphotyrosine-dependent interactions downstream of the collagen receptor DDR1. *FEBS Lett*, 580(1), 15-22.

Kung, J. E. & Jura, N. (2019) Prospects for pharmacological targeting of pseudokinases. *Nature Reviews Drug Discovery*, 18(7), 501-526.

Labrador, J. P., Azcoitia, V., Tuckermann, J., Lin, C., Olaso, E., Mañes, S., Brückner, K., Goergen, J. L., Lemke, G., Yancopoulos, G., Angel, P., Martínez, C. & Klein, R. (2001) The collagen receptor DDR2 regulates proliferation and its elimination leads to dwarfism. *EMBO Rep*, 2(5), 446-52.

Lahiry, P., Torkamani, A., Schork, N. J. & Hegele, R. A. (2010) Kinase mutations in human disease: interpreting genotype–phenotype relationships. *Nature Reviews Genetics*, 11(1), 60-74.

Lai, C. & Lemke, G. (1994) Structure and expression of the Tyro 10 receptor tyrosine kinase. *Oncogene*, 9(3), 877-83.

Laval, S., Butler, R., Shelling, A. N., Hanby, A. M., Poulsom, R. & Ganesan, T. S. (1994) Isolation and characterization of an epithelial-specific receptor tyrosine kinase from an ovarian cancer cell line. *Cell Growth Differ*, 5(11), 1173-83.

Ledda, F. & Paratcha, G. (2007) Negative Regulation of Receptor Tyrosine Kinase (RTK) Signaling: A Developing Field. *Biomark Insights*, 2, 45-58.

Leitinger, B. (2003) Molecular analysis of collagen binding by the human discoidin domain receptors, DDR1 and DDR2. Identification of collagen binding sites in DDR2. *J Biol Chem*, 278(19), 16761-9.

Leitinger, B. (2014) Discoidin domain receptor functions in physiological and pathological conditions. *Int Rev Cell Mol Biol*, 310, 39-87.

Leitinger, B. & Kwan, A. P. (2006) The discoidin domain receptor DDR2 is a receptor for type X collagen. *Matrix Biol*, 25(6), 355-64.

Lemeer, S., Bluwstein, A., Wu, Z., Leberfinger, J., Muller, K., Kramer, K. & Kuster, B. (2012) Phosphotyrosine mediated protein interactions of the discoidin domain receptor 1. *J Proteomics*, 75(12), 3465-77.

Lemmon, M. A. & Schlessinger, J. (2010) Cell signaling by receptor tyrosine kinases. *Cell*, 141(7), 1117-34.

Li, C., Wen, A., Shen, B., Lu, J., Huang, Y. & Chang, Y. (2011) FastCloning: a highly simplified, purification-free, sequence- and ligation-independent PCR cloning method. *BMC Biotechnol*, 11, 92.

Li, E. & Hristova, K. (2010) Receptor tyrosine kinase transmembrane domains: Function, dimer structure and dimerization energetics. *Cell Adh Migr*, 4(2), 249-54.

Li, W. & Miller, W. T. (2006) Role of the activation loop tyrosines in regulation of the insulin-like growth factor I receptor-tyrosine kinase. *J Biol Chem*, 281(33), 23785-91.

Libby, P., Buring, J. E., Badimon, L., Hansson, G. K., Deanfield, J., Bittencourt, M. S., Tokgözoğlu, L. & Lewis, E. F. (2019) Atherosclerosis. *Nature Reviews Disease Primers*, 5(1), 56.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol*, 75(Pt 10), 861-877.

Liu, L., Hussain, M., Luo, J., Duan, A., Chen, C., Tu, Z. & Zhang, J. (2017) Synthesis and biological evaluation of novel dasatinib analogues as potent DDR1 and DDR2 kinase inhibitors. *Chem Biol Drug Des*, 89(3), 420-427.

Lombardo, C. R., Consler, T. G. & Kassel, D. B. (1995) In vitro phosphorylation of the epidermal growth factor receptor autophosphorylation domain by c-src: identification of phosphorylation sites and c-src SH2 domain binding sites. *Biochemistry*, 34(50), 16456-66.

Loriaux, M. M., Levine, R. L., Tyner, J. W., Fröhling, S., Scholl, C., Stoffregen, E. P., Wernig, G., Erickson, H., Eide, C. A., Berger, R., Bernard, O. A., Griffin, J. D., Stone, R. M., Lee, B., Meyerson, M., Heinrich, M. C., Deininger, M. W., Gilliland, D. G. & Druker, B. J. (2008) High-throughput sequence analysis of the tyrosine kinome in acute myeloid leukemia. *Blood*, 111(9), 4788-96.

Lu, K. K., Trcka, D. & Bendeck, M. P. (2011) Collagen stimulates discoidin domain receptor 1-mediated migration of smooth muscle cells through Src. *Cardiovasc Pathol*, 20(2), 71-6.

Luo, Z., Liu, H., Sun, X., Guo, R., Cui, R., Ma, X. & Yan, M. (2013) RNA interference against discoidin domain receptor 2 ameliorates alcoholic liver disease in rats. *PLoS One*, 8(2), e55860.

Ma, P. C., Maulik, G., Christensen, J. & Salgia, R. (2003) c-Met: structure, functions and potential for therapeutic inhibition. *Cancer Metastasis Rev*, 22(4), 309-25.

Ma, Y. S., Wu, Z. J., Bai, R. Z., Dong, H., Xie, B. X., Wu, X. H., Hang, X. S., Liu, A. N., Jiang, X. H., Wang, G. R., Jiang, J. J., Xu, W. H., Chen, X. P., Tan, G. H., Fu, D., Liu, J. B. & Liu, Q. (2018) DRR1 promotes glioblastoma cell invasion and epithelial-mesenchymal transition via regulating AKT activation. *Cancer Lett*, 423, 86-94.

Malaguarnera, R., Nicolosi, M. L., Sacco, A., Morcavallo, A., Vella, V., Voci, C., Spatuzza, M., Xu, S. Q., Iozzo, R. V., Vigneri, R., Morrione, A. & Belfiore, A. (2015) Novel cross talk between IGF-IR and DDR1 regulates IGF-IR trafficking, signaling and biological responses. *Oncotarget*, 6(18), 16084-105.

Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. (2002) The protein kinase complement of the human genome. *Science*, 298(5600), 1912-34.

Manning, L. B., Li, Y., Chickmagalur, N. S., Li, X. & Xu, L. (2016) Discoidin Domain Receptor 2 as a Potential Therapeutic Target for Development of Disease-Modifying Osteoarthritis Drugs. *Am J Pathol*, 186(11), 3000-3010.

Marshall, C. J. (1995) Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell*, 80(2), 179-85.

Mathieu, S. V., Aragão, K. S., Imberty, A. & Varrot, A. (2010) Discoidin I from Dictyostelium discoideum and Interactions with oligosaccharides: specificity, affinity, crystal structures, and comparison with discoidin II. *J Mol Biol*, 400(3), 540-54.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr*, 40(Pt 4), 658-674.

McKeage, K. & Perry, C. M. (2002) Trastuzumab: a review of its use in the treatment of metastatic breast cancer overexpressing HER2. *Drugs*, 62(1), 209-43.

Meyer zum Gottesberge, A. M., Gross, O., Becker-Lendzian, U., Massing, T. & Vogel, W. F. (2008) Inner ear defects and hearing loss in mice lacking the collagen receptor DDR1. *Lab Invest*, 88(1), 27-37.

Miaczynska, M. (2013) Effects of membrane trafficking on signaling by receptor tyrosine kinases. *Cold Spring Harb Perspect Biol*, 5(11), a009035.

Michailidis, I. E., Rusinova, R., Georgakopoulos, A., Chen, Y., Iyengar, R., Robakis, N. K., Logothetis, D. E. & Baki, L. (2011) Phosphatidylinositol-4,5-bisphosphate regulates epidermal growth factor receptor activation. *Pflugers Arch*, 461(3), 387-97.

Mihai, C., Chotani, M., Elton, T. S. & Agarwal, G. (2009) Mapping of DDR1 distribution and oligomerization on the cell surface by FRET microscopy. *J Mol Biol*, 385(2), 432-45.

Mirshafiey, A., Ghalamfarsa, G., Asghari, B. & Azizi, G. (2014) Receptor Tyrosine Kinase and Tyrosine Kinase Inhibitors: New Hope for Success in Multiple Sclerosis Therapy. *Innov Clin Neurosci*, 11(7-8), 23-36.

Mohammadi, M., Schlessinger, J. & Hubbard, S. R. (1996) Structure of the FGF receptor tyrosine kinase domain reveals a novel autoinhibitory mechanism. *Cell*, 86(4), 577-87.

Mol, C. D., Dougan, D. R., Schneider, T. R., Skene, R. J., Kraus, M. L., Scheibe, D. N., Snell, G. P., Zou, H., Sang, B. C. & Wilson, K. P. (2004) Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. *J Biol Chem*, 279(30), 31655-63.

Mol, C. D., Lim, K. B., Sridhar, V., Zou, H., Chien, E. Y., Sang, B. C., Nowakowski, J., Kassel, D. B., Cronin, C. N. & McRee, D. E. (2003) Structure of a c-kit product complex reveals the basis for kinase transactivation. *J Biol Chem*, 278(34), 31461-4.

Moll, S., Desmouliere, A., Moeller, M. J., Pache, J. C., Badi, L., Arcadu, F., Richter, H., Satz, A., Uhles, S., Cavalli, A., Drawnel, F., Scapozza, L. & Prunotto, M. (2019) DDR1 role in fibrosis and its pharmacological targeting. *Biochim Biophys Acta Mol Cell Res*, 1866(11), 118474.

Moll, S., Yasui, Y., Abed, A., Murata, T., Shimada, H., Maeda, A., Fukushima, N., Kanamori, M., Uhles, S., Badi, L., Cagarelli, T., Formentini, I., Drawnel, F., Georges, G., Bergauer, T., Gasser, R., Bonfil, R. D., Fridman, R., Richter, H., Funk, J., Moeller, M. J., Chatziantoniou, C. & Prunotto, M. (2018) Selective pharmacological inhibition of DDR1 prevents experimentally-induced glomerulonephritis in prevention and therapeutic regime. *J Transl Med*, 16(1), 148.

Mori, S., Ronnstrand, L., Yokote, K., Engstrom, A., Courtneidge, S. A., Claesson-Welsh, L. & Heldin, C. H. (1993) Identification of two juxtamembrane autophosphorylation sites in the PDGF beta-receptor; involvement in the interaction with Src family tyrosine kinases. *Embo j*, 12(6), 2257-64.

Murray, C. W., Berdini, V., Buck, I. M., Carr, M. E., Cleasby, A., Coyle, J. E., Curry, J. E., Day, J. E., Day, P. J., Hearn, K., Iqbal, A., Lee, L. Y., Martins, V., Mortenson, P. N., Munck, J. M., Page, L. W., Patel, S., Roomans, S., Smith, K., Tamanini, E. & Saxty, G. (2015) Fragment-Based Discovery of Potent and Selective DDR1/2 Inhibitors. *ACS Med Chem Lett*, 6(7), 798-803.

Nair, A., Chauhan, P., Saha, B. & Kubatzky, K. F. (2019) Conceptual Evolution of Cell Signaling. *Int J Mol Sci*, 20(13).

Needham, S. R., Roberts, S. K., Arkhipov, A., Mysore, V. P., Tynan, C. J., Zanetti-Domingues, L. C., Kim, E. T., Losasso, V., Korovesis, D., Hirsch, M., Rolfe, D. J., Clarke, D. T., Winn, M. D., Lajevardipour, A., Clayton, A. H. A., Pike, L. J., Perani, M., Parker, P. J., Shan, Y., Shaw, D. E. & Martin-Fernandez, M. L. (2016) EGFR oligomerization organizes kinase-active dimers into competent signalling platforms. *Nature Communications*, 7(1), 13307.

Nemoto, T., Ohashi, K., Akashi, T., Johnson, J. D. & Hirokawa, K. (1997) Overexpression of protein tyrosine kinases in human esophageal cancer. *Pathobiology*, 65(4), 195-203.

Niu, X. L., Peters, K. G. & Kontos, C. D. (2002) Deletion of the carboxyl terminus of Tie2 enhances kinase activity, signaling, and function. Evidence for an autoinhibitory mechanism. *J Biol Chem*, 277(35), 31768-73.

Noordeen, N. A., Carafoli, F., Hohenester, E., Horton, M. A. & Leitinger, B. (2006) A transmembrane leucine zipper is required for activation of the dimeric receptor tyrosine kinase DDR1. *J Biol Chem*, 281(32), 22744-51.

Ojosnegros, S., Cutrale, F., Rodríguez, D., Otterstrom, J. J., Chiu, C. L., Hortigüela, V., Tarantino, C., Seriola, A., Mieruszynski, S., Martínez, E., Lakadamyali, M., Raya, A. & Fraser, S. E. (2017) Eph-ephrin signaling modulated by polymerization and condensation of receptors. *Proc Natl Acad Sci U S A*, 114(50), 13188-13193.

Olaso, E., Arteta, B., Benedicto, A., Crende, O. & Friedman, S. L. (2011) Loss of discoidin domain receptor 2 promotes hepatic fibrosis after chronic carbon tetrachloride through altered paracrine interactions between hepatic stellate cells and liver-associated macrophages. *Am J Pathol*, 179(6), 2894-904.

Olaso, E., Ikeda, K., Eng, F. J., Xu, L., Wang, L. H., Lin, H. C. & Friedman, S. L. (2001) DDR2 receptor promotes MMP-2-mediated proliferation and invasion by hepatic stellate cells. *J Clin Invest*, 108(9), 1369-78.

Ongusaha, P. P., Kim, J. I., Fang, L., Wong, T. W., Yancopoulos, G. D., Aaronson, S. A. & Lee, S. W. (2003) p53 induction and activation of DDR1 kinase counteract p53-mediated apoptosis and influence p53 regulation through a positive feedback loop. *Embo j*, 22(6), 1289-301.

Orgel, J. P., Irving, T. C., Miller, A. & Wess, T. J. (2006) Microfibrillar structure of type I collagen in situ. *Proc Natl Acad Sci U S A*, 103(24), 9001-5.

Ouwens, D. M., van der Zon, G. C., Pronk, G. J., Bos, J. L., Moller, W., Cheatham, B., Kahn, C. R. & Maassen, J. A. (1994) A mutant insulin receptor induces formation of a Shc-growth factor receptor bound protein 2 (Grb2) complex and p21ras-GTP without detectable interaction of insulin receptor substrate 1 (IRS1) with Grb2. Evidence for IRS1-independent p21ras-GTP formation. *J Biol Chem*, 269(52), 33116-22.

Patterson, H., Nibbs, R., McInnes, I. & Siebert, S. (2014) Protein kinase inhibitors in the treatment of inflammatory and autoimmune diseases. *Clin Exp Immunol*, 176(1), 1-10.

Pawson, T. (2007) Dynamic control of signaling by modular adaptor proteins. *Curr Opin Cell Biol*, 19(2), 112-6.

Perez, J. L., Jing, S. Q. & Wong, T. W. (1996) Identification of two isoforms of the Cak receptor kinase that are coexpressed in breast tumor cell lines. *Oncogene*, 12(7), 1469-77.

Phan, G. Q., Yang, J. C., Sherry, R. M., Hwu, P., Topalian, S. L., Schwartzentruber, D. J., Restifo, N. P., Haworth, L. R., Seipp, C. A., Freezer, L. J., Morton, K. E., Mavroukakis, S. A., Duray, P. H., Steinberg, S. M., Allison, J. P., Davis, T. A. & Rosenberg, S. A. (2003) Cancer regression and autoimmunity induced by cytotoxic T lymphocyte-associated antigen 4 blockade in patients with metastatic melanoma. *Proc Natl Acad Sci U S A*, 100(14), 8372-7.

Pietilä, M., Sahgal, P., Peuhu, E., Jäntti, N. Z., Paatero, I., Närvä, E., Al-Akhrass, H., Lilja, J., Georgiadou, M., Andersen, O. M., Padzik, A., Sihto, H., Joensuu, H., Blomqvist, M., Saarinen, I., Boström, P. J., Taimen, P. & Ivaska, J. (2019) SORLA regulates endosomal trafficking and oncogenic fitness of HER2. *Nature Communications*, 10(1), 2340.

Playford, M. P., Butler, R. J., Wang, X. C., Katso, R. M., Cooke, I. E. & Ganesan, T. S. (1996) The genomic structure of discoidin receptor tyrosine kinase. *Genome Res*, 6(7), 620-7.

Plaza-Menacho, I., Barnouin, K., Barry, R., Borg, A., Orme, M., Chauhan, R., Mouilleron, S., Martinez-Torres, R. J., Meier, P. & McDonald, N. Q. (2016) RET Functions as a Dual-Specificity Kinase that Requires Allosteric Inputs from Juxtamembrane Elements. *Cell Rep*, 17(12), 3319-3332.

Podlecki, D. A., Smith, R. M., Kao, M., Tsai, P., Hueckstead, T., Brandenburg, D., Lasher, R. S., Jarett, L. & Olefsky, J. M. (1987) Nuclear translocation of the insulin receptor. A possible mediator of insulin's long term effects. *J Biol Chem*, 262(7), 3362-8.

Poole, S., Firtel, R. A., Lamar, E. & Rowekamp, W. (1981) Sequence and expression of the discoidin I gene family in Dictyostelium discoideum. *J Mol Biol*, 153(2), 273-89.

Pottier, C., Fresnais, M., Gilon, M., Jérusalem, G., Longuespée, R. & Sounni, N. E. (2020) Tyrosine Kinase Inhibitors in Cancer: Breakthrough and Challenges of Targeted Therapy. *Cancers (Basel)*, 12(3).

Poudel, B., Ki, H. H., Lee, Y. M. & Kim, D. K. (2013) Induction of IL-12 production by the activation of discoidin domain receptor 2 via NF-κB and JNK pathway. *Biochem Biophys Res Commun*, 434(3), 584-8.

Poudel, B., Lee, Y. M. & Kim, D. K. (2015) DDR2 inhibition reduces migration and invasion of murine metastatic melanoma cells by suppressing MMP2/9 expression through ERK/NF-κB pathway. *Acta Biochim Biophys Sin (Shanghai)*, 47(4), 292-8.

Purvis, J. E. & Lahav, G. (2013) Encoding and decoding cellular information through signaling dynamics. *Cell*, 152(5), 945-56.

Quan, J., Yahata, T., Adachi, S., Yoshihara, K. & Tanaka, K. (2011) Identification of receptor tyrosine kinase, discoidin domain receptor 1 (DDR1), as a potential biomarker for serous ovarian cancer. *Int J Mol Sci*, 12(2), 971-82.

Red Brewer, M., Choi, S. H., Alvarado, D., Moravcevic, K., Pozzi, A., Lemmon, M. A. & Carpenter, G. (2009) The juxtamembrane region of the EGF receptor functions as an activation domain. *Mol Cell*, 34(6), 641-51.

Reger de Moura, C., Battistella, M., Sohail, A., Caudron, A., Feugeas, J. P., Podgorniak, M. P., Pages, C., Mazouz Dorval, S., Marco, O., Menashi, S., Fridman, R., Lebbé, C., Mourah, S. & Jouenne, F. (2019) Discoidin domain receptors: A promising target in melanoma. *Pigment Cell Melanoma Res*, 32(5), 697-707.

Ren, T., Zhang, W., Liu, X., Zhao, H., Zhang, J., Zhang, J., Li, X., Zhang, Y., Bu, X., Shi, M., Yao, L. & Su, J. (2014) Discoidin domain receptor 2 (DDR2) promotes breast cancer cell metastasis and the mechanism implicates epithelial-mesenchymal transition programme under hypoxia. *J Pathol*, 234(4), 526-37.

Richter, H., Satz, A. L., Bedoucha, M., Buettelmann, B., Petersen, A. C., Harmeier, A., Hermosilla, R., Hochstrasser, R., Burger, D., Gsell, B., Gasser, R., Huber, S., Hug, M. N., Kocer, B., Kuhn, B., Ritter, M., Rudolph, M. G., Weibel, F., Molina-David, J., Kim, J. J., Santos, J. V., Stihle, M., Georges, G. J., Bonfil, R. D., Fridman, R., Uhles, S., Moll,

S., Faul, C., Fornoni, A. & Prunotto, M. (2019) DNA-Encoded Library-Derived DDR1 Inhibitor Prevents Fibrosis and Renal Function Loss in a Genetic Mouse Model of Alport Syndrome. *ACS Chem Biol*, 14(1), 37-49.

Rikova, K., Guo, A., Zeng, Q., Possemato, A., Yu, J., Haack, H., Nardone, J., Lee, K., Reeves, C., Li, Y., Hu, Y., Tan, Z., Stokes, M., Sullivan, L., Mitchell, J., Wetzel, R., Macneill, J., Ren, J. M., Yuan, J., Bakalarski, C. E., Villen, J., Kornhauser, J. M., Smith, B., Li, D., Zhou, X., Gygi, S. P., Gu, T. L., Polakiewicz, R. D., Rush, J. & Comb, M. J. (2007) Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell*, 131(6), 1190-203.

Rosenbaum, D. M., Rasmussen, S. G. & Kobilka, B. K. (2009) The structure and function of G-protein-coupled receptors. *Nature*, 459(7245), 356-63.

Roskoski, R., Jr. (2016) Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. *Pharmacol Res*, 103, 26-48.

Russ, W. P. & Engelman, D. M. (1999) TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc Natl Acad Sci U S A*, 96(3), 863-8.

Sammon, D., Hohenester, E. & Leitinger, B. (2020) Two-step release of kinase autoinhibition in discoidin domain receptor 1. *Proc Natl Acad Sci U S A*, 117(36), 22051-22060.

Scapin, G., Dandey, V. P., Zhang, Z., Prosise, W., Hruza, A., Kelly, T., Mayhood, T., Strickland, C., Potter, C. S. & Carragher, B. (2018) Structure of the insulin receptor-insulin complex by single-particle cryo-EM analysis. *Nature*, 556(7699), 122-125.

Schminke, B., Muhammad, H., Bode, C., Sadowski, B., Gerter, R., Gersdorff, N., Bürgers, R., Monsonego-Ornan, E., Rosen, V. & Miosge, N. (2014) A discoidin domain receptor 1 knock-out mouse as a novel model for osteoarthritis of the temporomandibular joint. *Cell Mol Life Sci*, 71(6), 1081-96.

Schuppan, D., Ashfaq-Khan, M., Yang, A. T. & Kim, Y. O. (2018) Liver fibrosis: Direct antifibrotic agents and targeted therapies. *Matrix Biol*, 68-69, 435-451.

Shawver, L. K., Slamon, D. & Ullrich, A. (2002) Smart drugs: tyrosine kinase inhibitors in cancer therapy. *Cancer Cell*, 1(2), 117-23.

Shen, Q., Cicinnati, V. R., Zhang, X., Iacob, S., Weber, F., Sotiropoulos, G. C., Radtke, A., Lu, M., Paul, A., Gerken, G. & Beckebaum, S. (2010) Role of microRNA-199a-5p and discoidin domain receptor 1 in human hepatocellular carcinoma invasion. *Mol Cancer*, 9, 227.

Shintani, Y., Fukumoto, Y., Chaika, N., Svoboda, R., Wheelock, M. J. & Johnson, K. R. (2008) Collagen I-mediated up-regulation of N-cadherin requires cooperative signals from integrins and discoidin domain receptor 1. *J Cell Biol*, 180(6), 1277-89.

Shrivastava, A., Radziejewski, C., Campbell, E., Kovac, L., McGlynn, M., Ryan, T. E., Davis, S., Goldfarb, M. P., Glass, D. J., Lemke, G. & Yancopoulos, G. D. (1997) An orphan receptor tyrosine kinase family whose members serve as nonintegrin collagen receptors. *Mol Cell*, 1(1), 25-34.

Skamnaki, V. T., Owen, D. J., Noble, M. E., Lowe, E. D., Lowe, G., Oikonomakos, N. G. & Johnson, L. N. (1999) Catalytic mechanism of phosphorylase kinase probed by mutational studies. *Biochemistry*, 38(44), 14718-30.

Slack, B. E., Siniaia, M. S. & Blusztajn, J. K. (2006) Collagen type I selectively activates ectodomain shedding of the discoidin domain receptor 1: involvement of Src tyrosine kinase. *J Cell Biochem*, 98(3), 672-84.

Sliwoski, G., Kothiwale, S., Meiler, J. & Lowe, E. W., Jr. (2014) Computational methods in drug discovery. *Pharmacol Rev*, 66(1), 334-95.

Song, S., Shackel, N. A., Wang, X. M., Ajami, K., McCaughan, G. W. & Gorrell, M. D. (2011) Discoidin domain receptor 1: isoform expression and potential functions in cirrhotic human liver. *Am J Pathol*, 178(3), 1134-44.

Sorkin, A. & von Zastrow, M. (2009) Endocytosis and signalling: intertwining molecular networks. *Nat Rev Mol Cell Biol*, 10(9), 609-22.

Spellmon, N., Li, C. & Yang, Z. (2017) Allosterically targeting EGFR drug-resistance gatekeeper mutations. *J Thorac Dis*, 9(7), 1756-1758.

Stamos, J., Sliwkowski, M. X. & Eigenbrot, C. (2002) Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. *J Biol Chem*, 277(48), 46265-72.

Staudinger, L. A., Spano, S. J., Lee, W., Coelho, N., Rajshankar, D., Bendeck, M. P., Moriarty, T. & McCulloch, C. A. (2013) Interactions between the discoidin domain receptor 1 and β1 integrin regulate attachment to collagen. *Biol Open*, 2(11), 1148-59.

Sugiyama, M. G., Fairn, G. D. & Antonescu, C. N. (2019) Akt-ing Up Just About Everywhere: Compartment-Specific Akt Activation and Function in Receptor Tyrosine Kinase Signaling. *Front Cell Dev Biol*, 7, 70.

Sunk, I. G., Bobacz, K., Hofstaetter, J. G., Amoyo, L., Soleiman, A., Smolen, J., Xu, L. & Li, Y. (2007) Increased expression of discoidin domain receptor 2 is linked to the degree of cartilage damage in human knee joints: a potential role in osteoarthritis pathogenesis. *Arthritis Rheum*, 56(11), 3685-92.

Takai, K., Drain, A. P., Lawson, D. A., Littlepage, L. E., Karpuj, M., Kessenbrock, K., Le, A., Inoue, K., Weaver, V. M. & Werb, Z. (2018) Discoidin domain receptor 1 (DDR1) ablation promotes tissue fibrosis and hypoxia to induce aggressive basal-like breast cancers. *Genes Dev*, 32(3-4), 244-257.

Taner, S. B., Pando, M. J., Roberts, A., Schellekens, J., Marsh, S. G., Malmberg, K. J., Parham, P. & Brodsky, F. M. (2011) Interactions of NK cell receptor KIR3DL1*004 with chaperones and conformation-specific antibody reveal a functional folded state as well as predominant intracellular retention. *J Immunol*, 186(1), 62-72.

Tao, J., Zhang, M., Wen, Z., Wang, B., Zhang, L., Ou, Y., Tang, X., Yu, X. & Jiang, Q. (2018) Inhibition of EP300 and DDR1 synergistically alleviates pulmonary fibrosis in vitro and in vivo. *Biomed Pharmacother*, 106, 1727-1733.

Till, J. H., Becerra, M., Watty, A., Lu, Y., Ma, Y., Neubert, T. A., Burden, S. J. & Hubbard, S. R. (2002) Crystal structure of the MuSK tyrosine kinase: insights into receptor autoregulation. *Structure*, 10(9), 1187-96.

Tomasson, M. H., Xiang, Z., Walgren, R., Zhao, Y., Kasai, Y., Miner, T., Ries, R. E., Lubman, O., Fremont, D. H., McLellan, M. D., Payton, J. E., Westervelt, P., DiPersio, J. F., Link, D. C., Walter, M. J., Graubert, T. A., Watson, M., Baty, J., Heath, S., Shannon, W. D., Nagarajan, R., Bloomfield, C. D., Mardis, E. R., Wilson, R. K. & Ley, T. J. (2008) Somatic mutations and germline sequence variants in the expressed tyrosine kinase genes of patients with de novo acute myeloid leukemia. *Blood*, 111(9), 4797-808.

Trenker, R. & Jura, N. (2020) Receptor tyrosine kinase activation: From the ligand perspective. *Curr Opin Cell Biol*, 63, 174-185.

Tu, M. M., Lee, F. Y. F., Jones, R. T., Kimball, A. K., Saravia, E., Graziano, R. F., Coleman, B., Menard, K., Yan, J., Michaud, E., Chang, H., Abdel-Hafiz, H. A., Rozhok, A. I., Duex, J. E., Agarwal, N., Chauca-Diaz, A., Johnson, L. K., Ng, T. L., Cambier, J. C., Clambey, E. T., Costello, J. C., Korman, A. J. & Theodorescu, D. (2019) Targeting DDR2 enhances tumor response to anti-PD-1 immunotherapy. *Sci Adv*, 5(2), eaav2437.

Uchikawa, E., Choi, E., Shang, G., Yu, H. & Bai, X. C. (2019) Activation mechanism of the insulin receptor revealed by cryo-EM structure of the fully liganded receptor-ligand complex. *Elife*, 8.

Uhlenbrock, N., Smith, S., Weisner, J., Landel, I., Lindemann, M., Le, T. A., Hardick, J., Gontla, R., Scheinpflug, R., Czodrowski, P., Janning, P., Depta, L., Quambusch, L., Müller, M. P., Engels, B. & Rauh, D. (2019) Structural and chemical insights into the covalent-allosteric inhibition of the protein kinase Akt. *Chem Sci*, 10(12), 3573-3585.

Uings, I. J. & Farrow, S. N. (2000) Cell receptors and cell signalling. *Mol Pathol*, 53(6), 295-9.

Ullrich, A. & Schlessinger, J. (1990) Signal transduction by receptors with tyrosine kinase activity. *Cell*, 61(2), 203-12.

Valiathan, R. R., Marco, M., Leitinger, B., Kleer, C. G. & Fridman, R. (2012) Discoidin domain receptor tyrosine kinases: new players in cancer progression. *Cancer Metastasis Rev*, 31(1-2), 295-321.

Van Doren, S. R. (2015) Matrix metalloproteinase interactions with collagen and elastin. *Matrix Biol*, 44-46, 224-31.

van Lengerich, B., Agnew, C., Puchner, E. M., Huang, B. & Jura, N. (2017) EGF and NRG induce phosphorylation of HER3/ERBB3 by EGFR using distinct oligomeric mechanisms. *Proc Natl Acad Sci U S A*, 114(14), E2836-E2845.

Vaudry, D., Stork, P. J., Lazarovici, P. & Eiden, L. E. (2002) Signaling pathways for PC12 cell differentiation: making the right connections. *Science*, 296(5573), 1648-9.

Vehlow, A. & Cordes, N. (2019) DDR1 (discoidin domain receptor tyrosine kinase 1) drives glioblastoma therapy resistance by modulating autophagy. *Autophagy*, 15(8), 1487-1488.

Vella, V., Malaguarnera, R., Nicolosi, M. L., Morrione, A. & Belfiore, A. (2019) Insulin/IGF signaling and discoidin domain receptors: An emerging functional connection. *Biochim Biophys Acta Mol Cell Res*, 1866(11), 118522.

Vella, V., Malaguarnera, R., Nicolosi, M. L., Palladino, C., Spoleti, C., Massimino, M., Vigneri, P., Purrello, M., Ragusa, M., Morrione, A. & Belfiore, A. (2017) Discoidin domain receptor 1 modulates insulin receptor signaling and biological responses in breast cancer cells. *Oncotarget*, 8(26), 43248-43270.

Vogel, W., Gish, G. D., Alves, F. & Pawson, T. (1997) The discoidin domain receptor tyrosine kinases are activated by collagen. *Mol Cell*, 1(1), 13-23.

Vogel, W. F. (2002) Ligand-induced shedding of discoidin domain receptor 1. *FEBS Lett*, 514(2-3), 175-80.

Vogel, W. F., Aszodi, A., Alves, F. & Pawson, T. (2001) Discoidin domain receptor 1 tyrosine kinase has an essential role in mammary gland development. *Mol Cell Biol*, 21(8), 2906-17.

Wagner, M. J., Stacey, M. M., Liu, B. A. & Pawson, T. (2013) Molecular mechanisms of SH2- and PTB-domain-containing proteins in receptor tyrosine kinase signaling. *Cold Spring Harb Perspect Biol*, 5(12), a008987.

Wang, C. Z., Su, H. W., Hsu, Y. C., Shen, M. R. & Tang, M. J. (2006) A discoidin domain receptor 1/SHP-2 signaling complex inhibits alpha2beta1-integrin-mediated signal transducers and activators of transcription 1/3 activation and cell migration. *Mol Biol Cell*, 17(6), 2839-52.

Wang, C. Z., Yeh, Y. C. & Tang, M. J. (2009) DDR1/E-cadherin complex regulates the activation of DDR1 and cell spreading. *Am J Physiol Cell Physiol*, 297(2), C419-29.

Wang, D., Liu, Y., Zhao, W., Wang, R., Tong, X. & Jiang, G. (2019) [SRC kinase inhibitor PP2 inhibits invasion and metastasis of lung cancer A549 cells by upregulating connexin43 expression]. *Nan Fang Yi Ke Da Xue Xue Bao*, 39(7), 797-803.

Wang, Z., Bian, H., Bartual, S. G., Du, W., Luo, J., Zhao, H., Zhang, S., Mo, C., Zhou, Y., Xu, Y., Tu, Z., Ren, X., Lu, X., Brekken, R. A., Yao, L., Bullock, A. N., Su, J. & Ding, K. (2016) Structure-Based Design of Tetrahydroisoquinoline-7-carboxamides as Selective Discoidin Domain Receptor 1 (DDR1) Inhibitors. *J Med Chem*, 59(12), 5911-6.

Wang, Z. & Cole, P. A. (2014) Catalytic mechanisms and regulation of protein kinases. *Methods Enzymol*, 548, 1-21.

Wang, Z., Zhang, Y., Pinkas, D. M., Fox, A. E., Luo, J., Huang, H., Cui, S., Xiang, Q., Xu, T., Xun, Q., Zhu, D., Tu, Z., Ren, X., Brekken, R. A., Bullock, A. N., Liang, G., Ding, K. & Lu, X. (2018) Design, Synthesis, and Biological Evaluation of 3-(Imidazo[1,2-a]pyrazin-3-ylethynyl)-4-isopropyl- N-(3-((4-methylpiperazin-1-yl)methyl)-5-(trifluoromethyl)phenyl)benzamide as a Dual Inhibitor of Discoidin Domain Receptors 1 and 2. *J Med Chem*, 61(17), 7977-7990.

Ward, C. W., Lawrence, M. C., Streltsov, V. A., Adams, T. E. & McKern, N. M. (2007) The insulin and EGF receptor structures: new insights into ligand-induced receptor activation. *Trends Biochem Sci*, 32(3), 129-37.

Wehrman, T., He, X., Raab, B., Dukipatti, A., Blau, H. & Garcia, K. C. (2007) Structural and mechanistic insights into nerve growth factor interactions with the TrkA and p75 receptors. *Neuron*, 53(1), 25-38.

Weis, F., Menting, J. G., Margetts, M. B., Chan, S. J., Xu, Y., Tennagels, N., Wohlfart, P., Langer, T., Müller, C. W., Dreyer, M. K. & Lawrence, M. C. (2018) The signalling conformation of the insulin receptor ectodomain. *Nature Communications*, 9(1), 4420.

Winter, G., Lobley, C. M. & Prince, S. M. (2013) Decision making in xia2. *Acta Crystallogr D Biol Crystallogr*, 69(Pt 7), 1260-73.

Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K. & Evans, G. (2018) DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr D Struct Biol*, 74(Pt 2), 85-97.

Wu, P., Clausen, M. H. & Nielsen, T. E. (2015a) Allosteric small-molecule kinase inhibitors. *Pharmacol Ther*, 156, 59-68.

Wu, P., Nielsen, T. E. & Clausen, M. H. (2015b) FDA-approved small-molecule kinase inhibitors. *Trends Pharmacol Sci*, 36(7), 422-39.

Xia, B., Di, C., Zhang, J., Hu, S., Jin, H. & Tong, P. (2014) Osteoarthritis pathogenesis: a review of molecular mechanisms. *Calcif Tissue Int*, 95(6), 495-505.

Xu, C., Buczkowski, K. A., Zhang, Y., Asahina, H., Beauchamp, E. M., Terai, H., Li, Y. Y., Meyerson, M., Wong, K. K. & Hammerman, P. S. (2015) NSCLC Driven by DDR2 Mutation Is Sensitive to Dasatinib and JQ1 Combination Therapy. *Mol Cancer Ther*, 14(10), 2382-2389.

Xu, H., Abe, T., Liu, J. K., Zalivina, I., Hohenester, E. & Leitinger, B. (2014) Normal activation of discoidin domain receptor 1 mutants with disulfide cross-links, insertions, or deletions in the extracellular juxtamembrane region: mechanistic implications. *J Biol Chem*, 289(19), 13565-74.

Xu, H., Bihan, D., Chang, F., Huang, P. H., Farndale, R. W. & Leitinger, B. (2012) Discoidin domain receptors promote α1β1- and α2β1-integrin mediated cell adhesion to collagen by enhancing integrin activation. *PLoS One*, 7(12), e52209.

Xu, H., Raynal, N., Stathopoulos, S., Myllyharju, J., Farndale, R. W. & Leitinger, B. (2011) Collagen binding specificity of the discoidin domain receptors: binding sites on collagens II and III and molecular determinants for collagen IV recognition by DDR1. *Matrix Biol*, 30(1), 16-26.

Xu, L., Jensen, H., Johnston, J. J., Di Maria, E., Kloth, K., Cristea, I., Sapp, J. C., Darling, T. N., Huryn, L. A., Tranebjaerg, L., Cinotti, E., Kubisch, C., Rodahl, E., Bruland, O., Biesecker, L. G., Houge, G. & Bredrup, C. (2018) Recurrent, Activating Variants in the Receptor Tyrosine Kinase DDR2 Cause Warburg-Cinotti Syndrome. *Am J Hum Genet*, 103(6), 976-983.

Xu, L., Peng, H., Glasson, S., Lee, P. L., Hu, K., Ijiri, K., Olsen, B. R., Goldring, M. B. & Li, Y. (2007) Increased expression of the collagen receptor discoidin domain receptor 2

in articular cartilage as a key event in the pathogenesis of osteoarthritis. *Arthritis & Rheumatism*, 56(8), 2663-2673.

Xu, L., Peng, H., Wu, D., Hu, K., Goldring, M. B., Olsen, B. R. & Li, Y. (2005) Activation of the discoidin domain receptor 2 induces expression of matrix metalloproteinase 13 associated with osteoarthritis in mice. *J Biol Chem*, 280(1), 548-55.

Xu, L., Servais, J., Polur, I., Kim, D., Lee, P. L., Chung, K. & Li, Y. (2010) Attenuation of osteoarthritis progression by reduction of discoidin domain receptor 2 in mice. *Arthritis Rheum*, 62(9), 2736-44.

Yamaoka, T., Kusumoto, S., Ando, K., Ohba, M. & Ohmori, T. (2018) Receptor Tyrosine Kinase-Targeted Cancer Therapy. *Int J Mol Sci*, 19(11).

Yang, G., Li, Q., Ren, S., Lu, X., Fang, L., Zhou, W., Zhang, F., Xu, F., Zhang, Z., Zeng, R., Lottspeich, F. & Chen, Z. (2009) Proteomic, functional and motif-based analysis of C-terminal Src kinase-interacting proteins. *Proteomics*, 9(21), 4944-61.

Yang, J. C., Zhang, Y., He, S. J., Li, M. M., Cai, X. L., Wang, H., Xu, L. M. & Cao, J. (2017) TM4SF1 Promotes Metastasis of Pancreatic Cancer via Regulating the Expression of DDR1. *Sci Rep*, 7, 45895.

Yang, K., Kim, J. H., Kim, H. J., Park, I. S., Kim, I. Y. & Yang, B. S. (2005) Tyrosine 740 phosphorylation of discoidin domain receptor 2 by Src stimulates intramolecular autophosphorylation and Shc signaling complex formation. *J Biol Chem*, 280(47), 39058-66.

Yang, S. H., Baek, H. A., Lee, H. J., Park, H. S., Jang, K. Y., Kang, M. J., Lee, D. G., Lee, Y. C., Moon, W. S. & Chung, M. J. (2010) Discoidin domain receptor 1 is associated with poor prognosis of non-small cell lung carcinomas. *Oncol Rep*, 24(2), 311-9.

Yao, Z. & Stagljar, I. (2017) Multiple functions of protein phosphatases in receptor tyrosine kinase signaling revealed by interactome analysis. *Mol Cell Oncol*, 4(3), e1297101.

Yeung, D. A., Shanker, N., Sohail, A., Weiss, B. A., Wang, C., Wellmerling, J., Das, S., Ganju, R. K., Miller, J. L. C., Herr, A. B., Fridman, R. & Agarwal, G. (2019) Clustering, Spatial Distribution, and Phosphorylation of Discoidin Domain Receptors 1 and 2 in Response to Soluble Collagen I. *J Mol Biol*, 431(2), 368-390.

Yokoyama, N., Ischenko, I., Hayman, M. J. & Miller, W. T. (2005) The C terminus of RON tyrosine kinase plays an autoinhibitory role. *J Biol Chem*, 280(10), 8893-900.

Zegzouti, H., Zdanovskaia, M., Hsiao, K. & Goueli, S. A. (2009) ADP-Glo: A Bioluminescent and homogeneous ADP monitoring assay for kinases. *Assay Drug Dev Technol*, 7(6), 560-72.

Zhang, J., Yang, P. L. & Gray, N. S. (2009) Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer*, 9(1), 28-39.

Zhang, K., Corsa, C. A., Ponik, S. M., Prior, J. L., Piwnica-Worms, D., Eliceiri, K. W., Keely, P. J. & Longmore, G. D. (2013) The collagen receptor discoidin domain receptor 2 stabilizes SNAIL1 to facilitate breast cancer metastasis. *Nat Cell Biol*, 15(6), 677-87.

Zhang, X., Crespo, A. & Fernández, A. (2008) Turning promiscuous kinase inhibitors into safer drugs. *Trends Biotechnol*, 26(6), 295-301.

Zhang, X., Gureasko, J., Shen, K., Cole, P. A. & Kuriyan, J. (2006) An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell*, 125(6), 1137-49.

Zhao, H., Bian, H., Bu, X., Zhang, S., Zhang, P., Yu, J., Lai, X., Li, D., Zhu, C., Yao, L. & Su, J. (2016) Targeting of Discoidin Domain Receptor 2 (DDR2) Prevents Myofibroblast Activation and Neovessel Formation During Pulmonary Fibrosis. *Mol Ther*, 24(10), 1734-1744.

Zhao, Y., Chapman, D. A. & Jones, I. M. (2003) Improving baculovirus recombination. *Nucleic Acids Res*, 31(2), E6-6.

Zhong, X., Zhang, W. & Sun, T. (2019) DDR1 promotes breast tumor growth by suppressing antitumor immunity. *Oncol Rep*, 42(6), 2844-2854.

Zhu, D., Huang, H., Pinkas, D. M., Luo, J., Ganguly, D., Fox, A. E., Arner, E., Xiang, Q., Tu, Z. C., Bullock, A. N., Brekken, R. A., Ding, K. & Lu, X. (2019) 2-Amino-2,3-dihydro-1H-indene-5-carboxamide-Based Discoidin Domain Receptor 1 (DDR1) Inhibitors: Design, Synthesis, and in Vivo Antipancreatic Cancer Efficacy. *J Med Chem*, 62(16), 7431-7444.