Regulation of E2F1 by PAD4 mediated Citrullination

Fatemeh Ghari Seyed Fatemeh

Department of Oncology

St Hilda's College

University of Oxford

Trinity 2015

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Regulation of E2F1 by PAD4 mediated citrullination

Fatemeh Ghari Seyed Fatemeh

St Hilda's College

Department of Oncology

University of Oxford

Trinity 2015

A thesis submitted for the degree of Doctor of Philosophy

Abstract

Peptidyl arginine deiminase (PAD) 4 is a nuclear enzyme that converts arginine residues to citrulline. PAD4 activity has been implicated in inflammatory disease and cancer, although its mechanism of action, particularly the identity of functionally relevant targets, remains unclear. Moreover, there are no known reader domains specifically recognising citrulline modifications. E2F transcription factors play a central role in regulating gene expression during cellular proliferation but in addition participate in diverse biological processes. Furthermore, it is known that arginine methylation provides an important level of control in dictating the biological consequence of E2F1 activity. Here, we show that E2F1 is citrullinated by PAD4 on functionally important arginine residues. Citrullination of E2F1 assists its chromatin association, specifically to cytokine genes in granulocyte cells, and regulates binding of the bromodomain reader BRD4 to an acetylated domain in E2F1. Accordingly, the combined inhibition of PAD4 and BRD4 impedes the chromatin association of E2F1 and the activation of cytokine gene expression. When administered as a combination therapy in the murine collagen-induced arthritis model, small molecule inhibitors of PAD4 and BRD4 provide an effective approach for preventing collagen-induced arthritis. Our results shed light on a new E2F-dependent pathway that mediates the inflammatory effect of PAD4 and, for the first time, establish the interplay between citrullination and acetylation as a regulatory interface for driving inflammatory gene expression. Moreover, the results highlight a novel therapeutic approach for treating chronic inflammatory diseases.

Acknowledgement

I would like to sincerely thank my supervisor Professor Nick La Thangue for providing me with this DPhil opportunity and patiently guiding me through it; my mentors Dr Simon Carr and Dr Shonagh Munro for their invaluable scientific insights and discussions and my parents, husband and friends for their endless love, support and kindness.

I would also like to thank Cancer Research UK for their generous funding and our collaborators for their much gratified contributions: Professor Paul Thompson provided reagents (PAD inhibitors) and performed *in vitro* citrullination of E2F1 peptides, Dr Benedikt Kessler and Dr Joanna McGouran performed mass spectrometry analysis, Dr Dilar Baban and Dr Helen Lockstone provided the Microarray service, Professor Stefan Knapp, Dr Panagis Fillipakopoulos and Dr. Sarah Picaud facilitated the SPOT array, and Professor Patrick Venables and Dr Anne-Marie Quirke collaborated on the *in vivo* experiments.

Finally I would like to thank the fourteenth century Persian poet *Hafez*, whose poems gave me perspective in the turbulence of my DPhil days.

Table of Contents

List of Figures

Abbreviations

ACPA	Anti-citrulline protein antibody
AMC	Anti-modified citrulline
ATM	Ataxia telangiectasia mutated
ATP	Adenosine triphosphate
ATR	ATM and Rad3-related
ATRA	All Trans retinoic acid
AP1	Activator protein 1
Bax	Bcl-2 associated X protein
Bcl	B-cell lymphoma 2
BET	Bromodomain and extra-terminal
Bim	Bcl-2 interacting mediator of cell death
B-Raf	B-Rapidly accelerated fibrosarcoma
BSA	Bovine serum albumin
˚ C	Centigrade
$CaCl_2$	Calcium chloride
CARM1	Coactivator associated arginine methyltransferase-1
Caspase	Cysteine aspartic protease
CBP	CREB binding protein
CDK	Cyclin dependent kinase
CDC	Cell division cycle
ChIP	Chromatin immunoprecipitation
CHK2	Checkpoint kinase
CHX	Cycloheximide
CIA	Collagen induced arthritis
CKI	Cyclin dependent kinase inhibitor
DBD	DNA binding domain
DISC	Death inducing signalling complex
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT3	DNA methyl transferase 3
Dox	Doxycycline
DP	DRTF-associated protein
DTT	Dithiothreitol
E1A	Early region 1A
E2F	E2 promoter binding factor
E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EMT	Epithelial to mesenchymal transition
FACS	Fluorescence activated cell sorting
FCS	Foetal calf serum
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase

GST	Glutathione S-transferase
HA	Hemagglutinin
HAT	Histone acetyl transferase
HCF1	Host cell factor 1
HIF-1α	Hypoxia inducible factor 1α
HP1	Heterochromatin protein 1
IL	Interleukin
INK4	Inhibitor of cyclin dependent kinase 4
kDa	kilo Dalton
LB	Luria-Bertani
LPS	Lipopolysaccharides
TNFα	Tumour necrosis factor alpha
LSD1	Lysine specific demethylase 1
LZ	Leucine zipper
MB	Marked box
MALDI-TOF	Matrix assisted laser desorption/ionization-Time of flight mass spectrometry
MCM	Mini chromosome maintenance protein
MMP	Matrix metalloproteinase
Na_3VO_3	Sodium orthovanadate
NaCl	Sodium chloride
NaF	Sodium fluoride
$NaHCO_3$	Sodium bicarbonate
NEDD8	Neural Precursor Cell Expressed, Developmentally Downregulated 8
NES	Nuclear export signal
NETs	Nuclear extracellular traps
NF-κβ	Nuclear factor kappa-light-chain-enhancer of activated B cells
NLS	Nuclear localization signal
NPM	Nucleophosmin
ON	Overnight
PAD	Peptidyl arginine deiminase
PARP	Poly (ADP-ribose) polymerase
PBS	Phosphate buffered saline
P/CAF	p300/CREB-binding protein associated factor
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
pRB	Retinoblastoma protein
PRMT	Protein arginine methyl transferase
RNA	Ribonucleic acid
R-point	Restriction point
ROS	Reactive oxygen species
RT	Room temperature
SAHA	Suberoylanilide Hydroxamic Acid
SDS	Sodium dodecyl sulphate

SETD1	SET domain containing protein
siRNA	Small interfering Ribonucleic acid
SNPs	Single nucleotide polymorphism
SOC	Super optimal broth
STAT3	Signal transducer and activator of transcription 3
TAM	Tumour associated macrophages
TIP60	HIV1 Tat interactive protein, 60 kDa
TP53	p53 protein
TPA	12-O-Tetradecanoylphorbol-13-Acetate
IPTG	Isopropyl-1-thio-beta-D-galactopyranoside
TRRAP	Transformation/Transcription Domain-Associated Protein
VEGF	Vascular endothelial growth factor

Chapter 1) Introduction

1.1 Hallmarks of cancer

According to the World Health Organization, cancer features amongst the leading causes of morbidity and mortality worldwide, with around 14 million new cases and 8 million related deaths in 2012 alone. In fact it is estimated that the incidence of cancer will increase by around 70% by the year 2030 [1]. As such cancer is and remains a huge burden and challenge at a personal, societal, economical and scientific level.

Cancer is not a single disease, but a collective term for over 200 different types of disease, which can all be characterised by several key features or hallmarks. These hallmarks originally formalised by Hanahan and Weinberg [2, 3], express the basic principles which underlie the complexity of neoplastic disease and the traits acquired in the multistep process of tumorigenesis. In summary these principles are: sustained proliferation, evasion of growth suppressive mechanisms, resistance to cell death, acquired replicative immortality, induction of angiogenesis, invasion and metastasis, metabolic reprogramming and avoidance of immune destruction. In addition to these hallmarks, two key characteristics, namely genomic instability and tumour promoting inflammation have been suggested to be crucial for the acquisition of the hallmarks mentioned above (Fig 1.1).

Sustaining proliferation: Cancer cells can acquire various mechanisms to support their chronic proliferative state. These include production of growth factor signalling molecules, which promote growth and proliferation in an autocrine manner; release of signalling molecules by cancer cells, which promote nearby stromal cells to produce and release growth promoting factors; deregulated growth factor receptor signalling, including overexpression of cell surface receptors; constitutive activation of downstream signalling molecules such as B-Raf in melanoma and defects in mechanisms which suppress growth promoting circuitries [3].

Evasion of growth suppressors: Cancer cells suppress pathways and players which negatively regulate cell proliferation, largely through suppressing and deactivating relevant tumour suppressor genes. The retinoblastoma protein (pRB) and the p53 protein (p53) are amongst the most commonly lost or deactivated tumour suppressors in cancer, although many more *bona fida* tumour suppressor genes have been identified and validated [3].

Resistance to cell death: Cancer cells encounter many stress signals and stimuli during the course of tumorigenesis which should activate the apoptotic signalling cascade. However cancer cells employ various strategies (including loss of function of the damage sensor *TP53* gene, down regulation of pro-apoptotic factors (e.g. Bax and Bim), up regulation of anti-apoptotic factors (e.g. Bcl-2 and Bcl-x_L) and overexpression of survival factors (e.g. insulin like growth factor)) to resist the cell death mechanisms [3].

Acquiring replicative immortality: Normal cells undergo limited rounds of cell division cycles before entering senescence or reaching a crisis point which leads to cell death. Cancer cells however avoid these outcomes and are considered immortalized by virtue of their acquired ability to maintain their chromosomal telomere ends, largely through overexpression of the telomerase enzyme [3].

Inducing angiogenesis: The process of neovascularization, or angiogenesis, is required to support the growing mass of tumour cells, both for continual supply of nutrients and oxygen, and for removal of metabolic waste products and carbon dioxide. As part of tumour progression the process of angiogenesis is activated to support the continual growth and sprouting of an otherwise quiescent vasculature. The prototypic proangiogenic factor VEGF-A is a frequent mediator of the angiogenic switch [3].

Invasion and metastasis: Local invasion and distant metastasis are features of particularly high grade aggressive malignancies. As part of a regulatory program referred to as the epithelial-to-mesenchymal transition (EMT), adhesion molecules that mediate cell-cell and cell-ECM contacts (e.g. E-cadherin) are downregulated and adhesion molecules that promote cell migration (N-cadherin) are upregulated. This allows malignant cells to acquire the capacity to spread beyond their local vicinity [3].

Reprogramming of metabolism: Originally noted by Warburg, it is now evident that cancer cells preferentially limit their energy metabolism to glycolysis rather than using mitochondrial oxidative phosphorylation. Although the former yields considerably less ATP per glucose molecule metabolised, it provides crucial intermediate metabolites which can be diverted into other biosynthetic pathways to support the hyper-proliferative state of the cancer cell [3].

Evading immune destruction: There are many studies to suggest that the immune system can serve as an important surveillance mechanism and barrier against tumour formation and progression. In immunocompromised mice for instance, the incidence of carcinogen induced tumours were found to be higher than immune competent control mice. Amongst the innate and adaptive immune system elements, particularly $CD8^+$ cytotoxic T lymphocytes and natural killer cells have been proposed to be important for eliminating highly immunogenic tumour cells. Selection for weakly immunogenic cancer cells, as well as modulation of the immune response by cancer cells through release of immunosuppressive factors contribute to evasion of immune destruction [3].

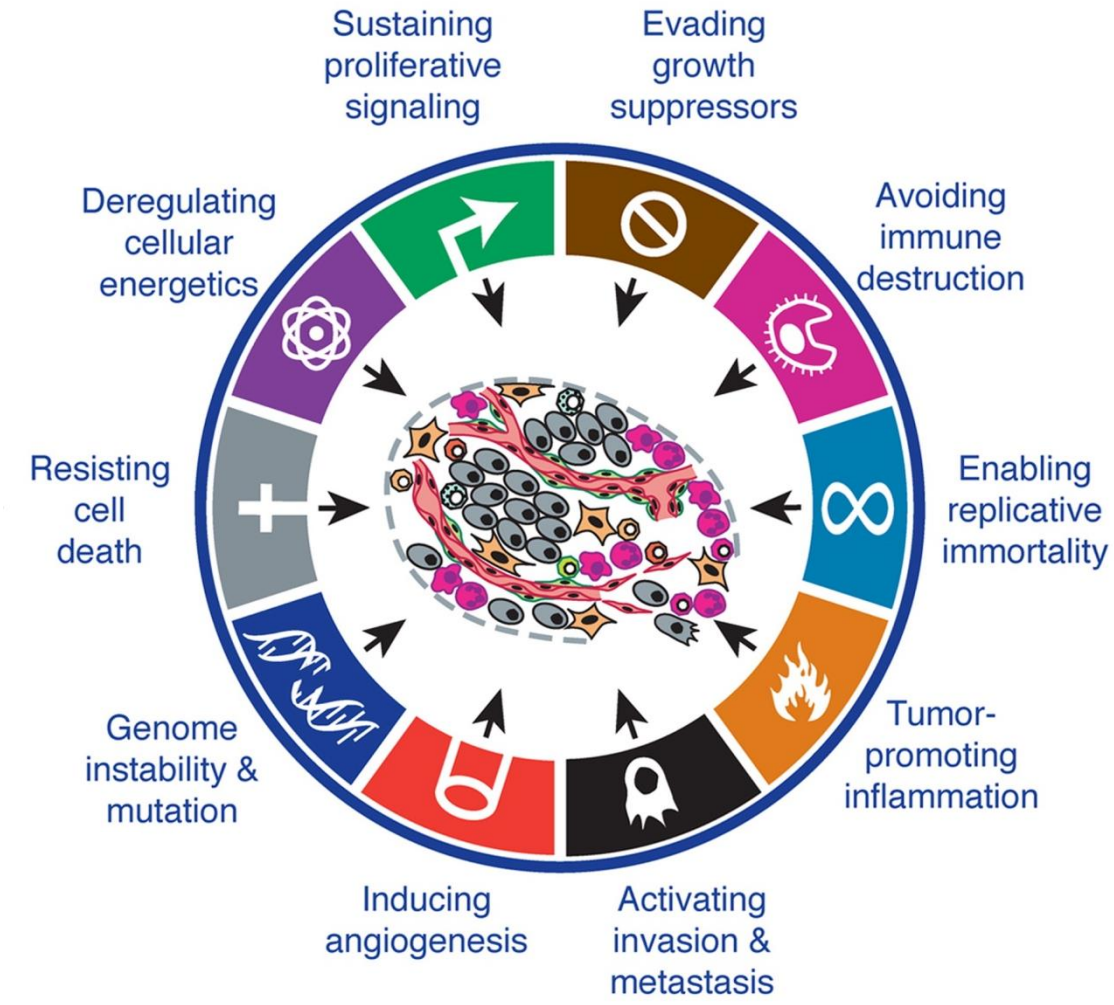


Figure 1.1) Hallmarks of cancer

The figure outlines the main hallmarks and enabling characteristics, which neoplastic cells acquire during the multistep process of tumorigenesis. Deregulated/reprogrammed cellular metabolism and escaping immune destruction are considered 'emerging hallmarks', as they still await generalisation and complete validation. Genomic instability and tumour promoting inflammation are considered to be enabling characteristics, as they assist the acquisition of other hallmarks. Image taken from 'Hallmarks of Cancer: The Next Generation'[3].

1.2 Tumour suppressors and oncogenes

Cancer cells acquire the hallmarks mentioned previously, by accumulating a range of genetic and epigenetic changes. At the heart of these are gain-of-function mutations in proto-oncogenes and loss-of-function of tumour suppressor genes [4, 5]. An oncogene has been recognised as any gene, whose encoding protein is capable of transforming cells in culture or inducing cancer in animals. Oncogenes are (generally) derived from normal cellular genes, known as proto-oncogenes. Gain-of-function changes in the proto-oncogene converts it to an oncogene, whose product then confers growth advantages to the cell. The three main gain-of-function mechanisms include: point mutations in the proto-oncogene to encode a constitutively active protein; gene amplification of the pro-oncogene, which therefore leads to overexpression of the encoded protein; and finally chromosomal translocations to bring the proto-oncogene under the expression of a different promoter that results in uncontrolled expression of the gene [6].

Tumour suppressor genes on the other hand, are those genes whose protein product could hinder cell proliferation and cancer development. These include for instance, checkpoint proteins that arrest the cell in response to DNA damage, pro-apoptotic proteins that induce apoptosis in response to cellular insults and cyclin-dependent kinase inhibitors that regulate cell cycle progression. As such, loss-of-function of these proteins, usually as a result of gene mutations or deletions, is required for unchallenged and uncontrolled cell proliferation. Unlike oncogenes however, where changes in one allele are sufficient to have a transforming effect, for most tumour suppressor genes, loss-of-function of both alleles are required to completely lose their cellular activity. Therefore tumour suppressors are generally recessive, although haploinsufficiency as well as dominant negative gain-of-function mutations may also lead to loss of functionality [7, 8].

1.3 The cell cycle

Studies into the mechanisms of the cell cycle have proved crucial to our understanding of cancer. Deregulation of many cell cycle related pathways and players have been shown to be hugely relevant to the biology and etiology of cancer, and in fact the molecular actions of most oncogenes and tumour suppressors could be understood in terms of their effect on the cell cycle [9, 10]. The cell cycle can be defined as the ordered progression of the cell through a complete round of growth and division, exemplified by the four sequential phases G1, S, G2 and M, followed by cytokinesis (i.e. division of the cell into two identical daughter cells). A cell needs to integrate many layers of signals into a simple yet crucial binary decision of whether to commit to a round of growth and division, or whether to become quiescent. Almost all normal cells (exception being early embryonic cells) need to be prompted by mitogenic growth factors to proliferate, whilst cancer cells, as outlined in the previous section acquire and employ various mechanisms to sustain their proliferation [11].

As a cell emerges from a previous division, it needs to accumulate its macromolecular content if it is to complete another round of the cell cycle. In the first gap phase (G1), cells accumulate their RNA and protein content and increase in size, before entering the synthetic phase (S), where DNA replication takes place and the genome content is doubled. The cells subsequently enter the second gap phase (G2), followed by the final mitosis phase (M), during which chromosomes align and are equally distributed into two daughter cells that form via cytokinesis (Fig 1.2) [12]. It is during G1 that cells make the crucial decision of whether or not to commit to completing the cell cycle. In fact late in G1 prior to the G1-S phase transition, at the so called Restriction point (R-point), cells make this critical decision. A cell which passes the R-point will become committed to

completing the cell cycle irrespective of the availability of mitogenic growth factors. As such, deregulation of the R-point decision making machinery underlies many cancer types [13, 14].

A cell progressing through the cell cycle deploys various surveillance mechanisms. These broadly ensure that the next phase is not commenced until the previous is successfully completed, that the cycle progresses unidirectionally and no phase is repeated, and finally if there is a problem, the cell does not progress until it is corrected. These mechanisms are collectively referred to as the cell cycle checkpoint mechanisms [15]. For instance the G1 checkpoint prevents the cell from progressing into S phase if it detects damage to the genome that requires repair. The S phase checkpoint slows or halts DNA replication in response to DNA damage. The G2 checkpoint prevents the cell from progressing into M phase if DNA replication from S phase is incomplete, and the M phase checkpoint in anaphase, halts mitosis if the chromosomes are not properly attached to the mitotic spindle [15, 16]. As cancer cells acquire extensive gene mutations and chromosomal abnormalities during the course of mutagenesis, they also disrupt and inactivate one or more of their checkpoint controls to progress through the cell cycle unhindered [17].

1.4 Cyclin-dependent kinases

The cell cycle deploys a family of kinases, known as cyclin-dependent kinases (CDKs) to orchestrate the processes of cell cycle progression. These kinases are serine-threonine kinases with roughly 40% amino acid sequence identity to one another [18]. The activity of these kinases depends on their association with their regulatory subunit, known as cyclins; so named due to their oscillating levels with the various phases of the cell cycle. In early G1, CDK4 and CDK6 associate with cyclin D (D1, D2 and D3) to drive the

cell into late in G1, whereupon CDK2 associates with cyclin E (E1 and E2) and facilitates the G1-S phase transition. As cells enter the S-phase CDK2 leaves cyclin E and instead associates with cyclin A (A1 and A2) to enable S-phase progression. Later in S-phase cyclin A associates with CDC2/CDK1 and finally in late G2, CDC2/CDK1 associates with cyclin B (B1 and B2) to trigger the events of mitosis [18, 19]. However the necessity of these CDK enzymes for cell cycle progression has been challenged in recent years. Whilst CDK1 has been suggested to be essential, the absolute requirement for other interphase CDK enzymes has been found to be cell type dependent. Similarly different cancer types have been suggested to have different requirements for the various interphase CDK enzymes [20].

The availability of cyclins is very tightly and precisely controlled to ensure orderly and timely progression through the cell cycle. The gradual accumulation of cyclins is regulated at the transcriptional level, whilst their rapid degradation is achieved by dedicated ubiquitin ligases which target the cyclins for proteasome mediated degradation [21]. In addition to cyclins, the activity of the CDK enzymes can be regulated by CDK inhibitors (CKI), of which seven have been identified. Four of these, namely $p16^{INK4A}$, $p15^{INK4B}$, $p18^{INK4C}$ and $p19^{INK4D}$ (the INK4 protein family) specifically target CDK4 and CDK6, whilst the remaining three, which include $p21^{Cip1}$, $p27^{Kip1}$ and $p57^{Kip2}$ have a broader inhibitory activity and can target all interphase CDK enzymes [22, 23].

Mitogenic signalling, particularly through the phosphatidylinositol 3-kinase (PI3K) pathway promotes cell cycle commitment and progression through both upregulation of cyclin D levels [24] and phosphorylation of CKI's to promote their nuclear exclusion [25]. On the other hand physiological stresses, such as DNA damage can strongly induce CKI levels, particularly $p21^{Cip1}$, to block CDK-cyclin activity and halt further cell cycle

progression [25]. Thus the overall balance between the various growth promoting versus growth inhibitory signals decide whether a cell should proliferate or arrest.

1.5 The pRB-E2F axis in cell cycle regulation

The retinoblastoma protein (pRB) is a key cell cycle regulator, which at the molecular level controls the execution of the Restriction point transition [26-28]. The sequential phosphorylation of pRB during G1, and its hyperphosphorylation at the R-point are necessary events leading to cell cycle commitment and G1-S phase transition. Over 10 conserved consensus phosphorylation sites have been identified on pRB in cells [29, 30].

Impairment of pRB function, either through direct mutation or deletion of the pRB gene, or deregulation of its upstream regulators, such as cyclin D, CDK4 and CDK6 are frequent occurrences in many cancer types. Indeed deregulation of pRB signalling pathways have been reported in most cancers [31, 32]. The central region of pRB, also known as the pocket domain, was originally identified as the minimal region necessary to bind viral oncoproteins such as the adenovirus E1A and the simian virus 40 large T-antigen [33]. It was later discovered that this is the region which also binds and hinders the activity of the E2F transcription factors, and thus the occlusion of this domain by viral oncoproteins essentially drives deregulated cell cycle progression by unleashing E2F activity [34].

pRB binds the transactivation domain of the E2F transcription factors via a highly conserved region in its pocket domain [35, 36]. Additional contacts are also reported between the C-terminal domain of pRB and the marked box of E2F1 specifically [37]. The association between pRB and the transactivation domain of E2F very effectively suppresses the activity of the E2F-DP dimer on target gene promoters [38]. Furthermore, pRB recruits

various chromatin and histone modifying enzymes, including HDAC and SWI/SNF remodelling complex to further suppress transcription [39-41]. The gradual and sequential phosphorylation of pRB during the course of G1 by cyclin D/CDK4/6 and subsequently cyclin E/CDK2, firstly allow for the dissociation of the chromatin remodelling factors from pRB and secondly the complete release of E2F-DP [42]. The transcriptional program initiated and executed by E2F-DP complex is then responsible for events leading to G1-S transition and S-phase progression (Fig 1.2) [43, 44]. As a cell passes through mitosis, pRB becomes dephosphorylated by protein phosphatase type I (PP1) in preparation for the subsequent G1 phase [45].

In support of the classical role of pRB in suppressing E2F activity, homozygous deletion of *Rb1* has been shown to be embryonically lethal, with the accompanying ectopic proliferation and apoptosis being partly rescued by co-deletion of the activator E2Fs [46]. However, alternative E2F independent mechanisms for pRB mediated cell cycle arrest have also been proposed. This follows two main observations. Firstly, although the other pocket proteins p107 and p130 can supress E2F activity, only pRB is frequently mutated in cancer, and secondly, there are studies to suggest that the activator E2F's are dispensable for cellular proliferation [47, 48]. Therefore, pRB has been suggested to exert its cytostatic effects within a much broader context, including for instance the stabilization of $p27^{Kip1}$ by pRB, which inhibits CDK activity to halt cell cycle progression [49].

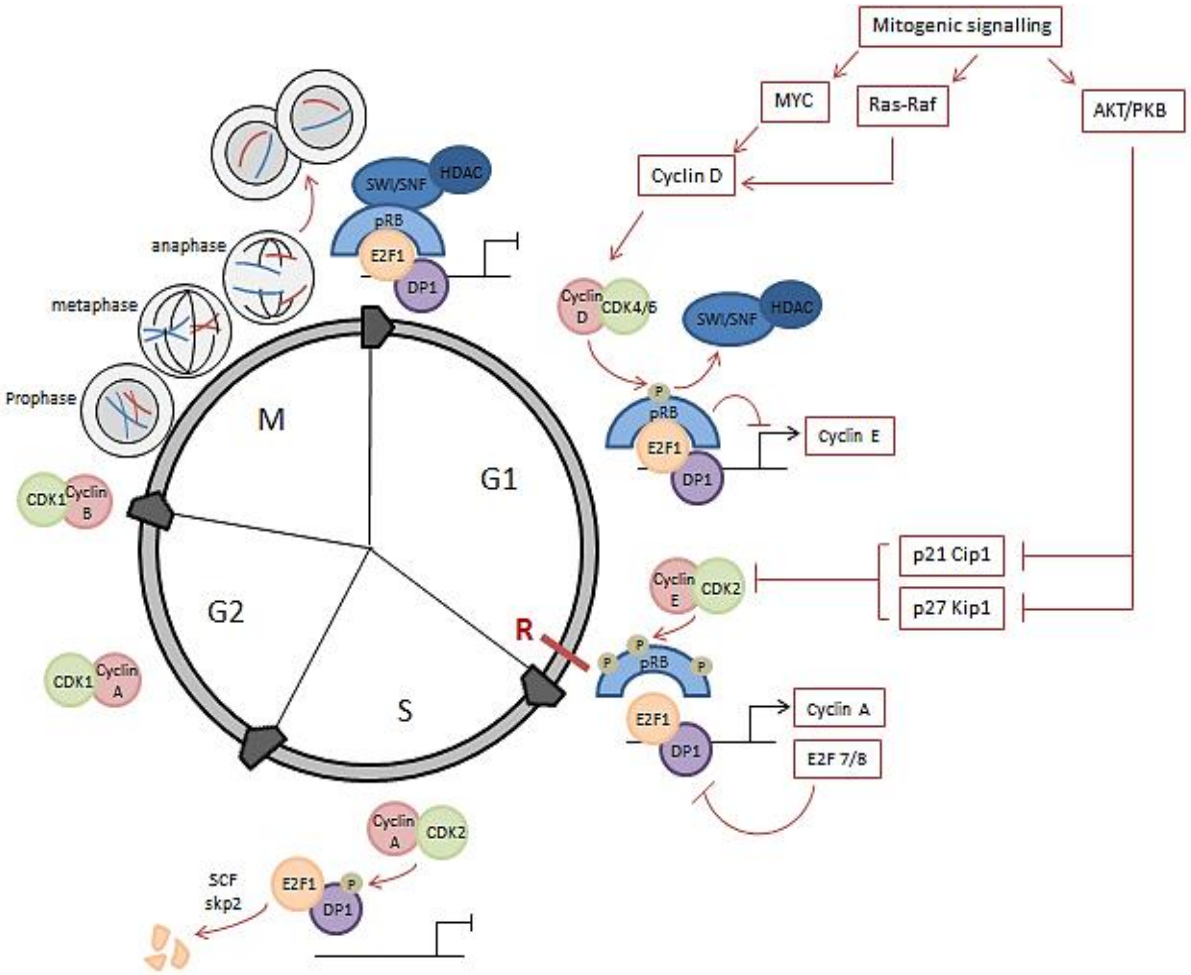


Figure 1.2) The cell cycle and the E2F-pRB axis

The diagram depicts the four phases of the cell cycle, namely G1, S, G2 and M phase. The M phase further consists of prophase, metaphase, anaphase and cytokinesis, where the cell splits into two identical daughter cells. During early G1, pRB binds to the E2F-DP complex and recruits histone modifying and chromatin remodelling factors to suppress the expression of cell cycle genes. Exposure to mitogenic growth factors induces cyclin D expression, which binds to CDK4/6 to phosphorylate pRB. This releases several pRB binding partners and allows for expression of several cell cycle genes including cyclin E. cyclin E subsequently partners with CDK2 to fully phosphorylate pRB and release E2F, which then drives G1-S phase progression [50]. During S phase, where E2F1 activity is no longer required, it is targeted for phosphorylation by cyclin A/CDK2 [51], and subsequently for ubiquitination and proteasome mediated degradation [52]. R is the restriction point. Diagram adopted from 'To live or let die-complexity within the E2F1 pathway' [53]

1.6 The E2F family of transcription factors

The E2F family of transcription factors are a group of eight members (Fig 1.3) with one or two conserved winged-helix DNA Binding Domain(s) (DBD). Based mostly on *in vitro* cells culture studies, the eight mammalian E2F transcription factors have been divided into activator or repressor members [54, 55]. The activator members (E2F1, E2F2 and longer isoform of E2F3) were sub-grouped based on their observed ability to positively regulate transcription and drive cell cycle progression, whilst the repressor E2F members (shorter isoform of E2F3 and E2F4-8) were found to partake in transcriptional repression [55-57].

The E2F members 1-6 interact with one of three dimerization partner proteins, DP1, DP2 and DP3, via the leucine zipper (LZ) and marked box (MB) domain and bind DNA as a heterodimer. E2F members 1-3 can bind pRB via their transactivation domain, whilst E2F4 can associate with all three pocket proteins, including pRB, p107 and p130. E2F5 however preferentially binds to p130 only, and the remaining E2F members do not bind any of the pocket proteins. In the amino terminal region of E2F1-3 lies a nuclear localization signal, which flanks the cyclin A binding domain. On the other hand, E2F4-5 have a nuclear export signal, allowing their cellular localization and hence activity to be regulated in a cell cycle dependent manner [56, 58].

The E2F transcription factors have been shown to exhibit extensive functional redundancy, antagonism and feedback, which complicates their study *in vitro* and *in vivo* [54]. Although classically and phenotypically, the E2F members were classified according to their ability to activate or repress genes implicated in cell cycle progression and DNA synthesis [59], studies emerging from transgenic mouse experiments have suggested a much more complex role and requirement for each of the E2F members [60, 61]. Furthermore, it has

been difficult to reduce the specific E2F transcription factors into oncogenic or tumour suppressive merely on the basis of their cell cycle effect. For instance, whilst the activator E2F members are known to promote cell proliferation, there are reports of tumours with chromosomal deletions of E2F1-3 regions [54]. On the other hand, the repressor E2F members are not frequently found to be mutated, deleted or silenced in cancer, and in fact there are reports for amplification of E2F5 in certain breast carcinomas [62, 63]. However it should also be noted that overexpression and amplification of activator E2F members have been observed in various cancers [64, 65], and in fact there are suggestions that different cancer types could acquire different preferences for specific activator E2F members for oncogenic proliferation[54]. Therefore overall, much more remains to be learnt about the different E2F members, and how each contributes to the biology of normal versus malignant cells.

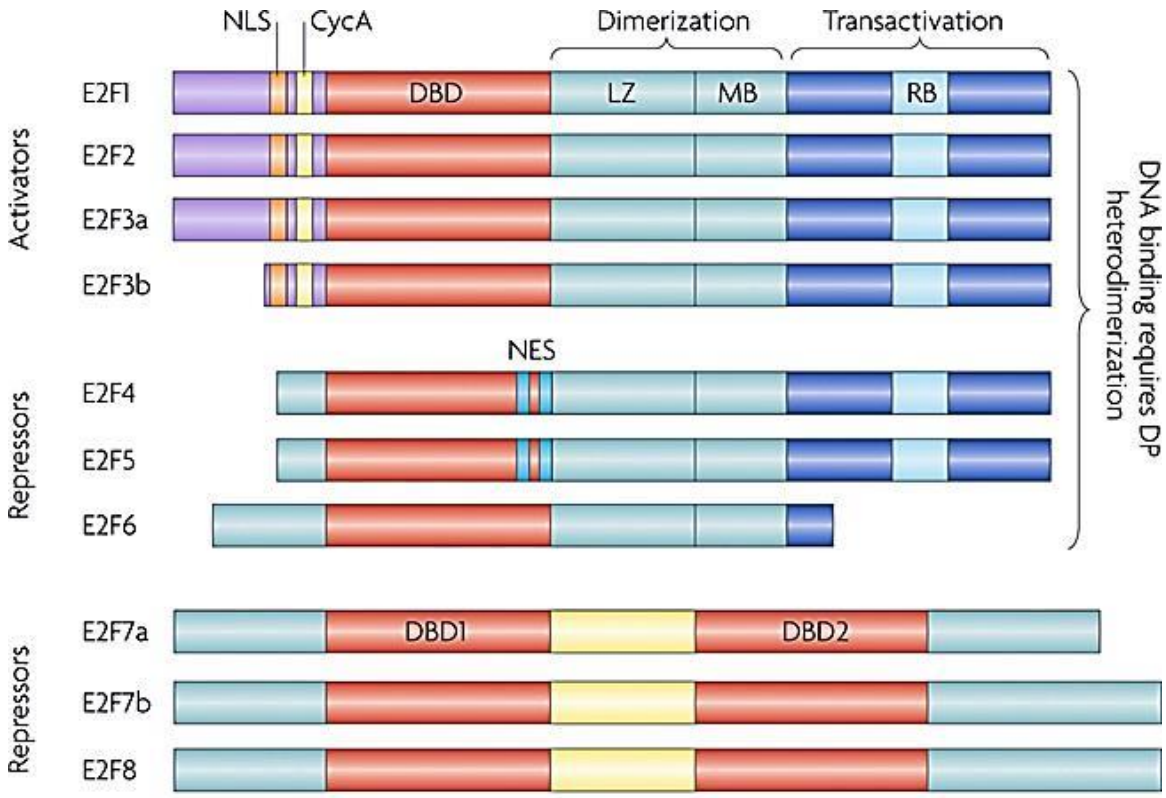


Figure 1.3) The E2F family of transcription factors

The E2F family consists of 8 members, which are broadly divided into activators and repressors. E2F1-3 are activators, as they positively control the expression of cell cycle genes and can drive cell growth and proliferation, whilst E2F4-8 are repressors, given their ability to negatively regulate cell growth and division. NLS: nuclear localization signal, CycA: cyclin A binding domain, DBD: DNA binding domain, LZ: leucine zipper, MB: marked box, RB: pRB binding domain. Figure taken from 'Emerging role of E2Fs in cancer: exit from cell cycle control' [54].

1.7 Biological functions of E2F1

The E2F1 transcription factor was originally discovered as the cellular factor required by the adenovirus E1A protein to activate its E2 promoters; hence the name **E2** promoter Binding **F**actor (E2F) [66]. Many genes associated with S phase entry have been shown to be regulated by E2F1 [67, 68], thus suggesting that it may have a strictly pro-growth and hence oncogenic activity. However, over two decades of *in vivo* studies have in fact painted a much more complicated picture. Early studies with $E2F1^{-/-}$ mice found that these animals developed a broad spectrum of tumours [69], and exhibited defects in T lymphocyte development as the thymocytes failed to undergo apoptosis during their maturation process [70]. Furthermore, loss of E2F1 was found to completely eliminate the apoptosis observed in $Rb^{-/-}$ mice [46], whilst overexpression of E2F1 in transgenic mice was found to cause testicular atrophy and sterility as a result of increased apoptosis in these tissues [71]. In combination with studies showing that loss of E2F1 can reduce the frequency of pituitary and thyroid tumours in $Rb^{-/-}$ mice [72], it is now well established in the field that E2F1 is endowed with both tumour suppressive and oncogenic properties [73]. Which aspect of E2F1 biology prevails over the other is presumably cell type and context dependent.

As already mentioned, a plethora of E2F1 target genes are required for cell cycle progression, with many of these directly implicated in DNA synthesis. In addition to cyclins and to name only a few, E2F1 regulates the expression of the MCM2-7 proteins (required to form the DNA replication pre-initiation complex), CDC6 (needed to load the MCM proteins onto DNA), PCNA (tethers DNA polymerase to DNA), DNA primase (creates RNA primer needed for DNA replication by DNA polymerase), flap endonuclease (required for maturation of Okazaki fragments), thymidylate synthase and thymidine kinase (required for thymidine synthesis, needed for DNA replication) [53]. On the other

hand, E2F1 is also known to promote and coordinate apoptosis through both p53 dependent and independent mechanisms. Consistent with *in vivo* experiments showing that ectopic E2F1 can cause tissue specific apoptosis, *in vitro* cell culture experiments revealed that E2F1 can accumulate under DNA damage with kinetics very similar to p53 [74]. E2F1 can then promote apoptosis by binding to and driving the transcription of several pro-apoptotic target genes including *APAF1*, *TP73*, *CASP3* and *CASP7* [75, 76].

However in recent years it has become apparent that E2F1 has many other roles beyond cell cycle control and apoptosis. Using global ChIP based approaches E2F1 has been shown to be recruited to thousands of genes in cells [77], expanding its repertoire of target genes to processes including differentiation [78], metabolism [79], animal development [80] and even inflammatory response [81]. Furthermore, these studies have found that the promoter recruitment of E2F1 is not necessarily restricted to regions which have a consensus E2F binding motif, but may also be facilitated or mediated by intermediate proteins and interaction partners including other transcription factors or elements of the general transcription machinery [54]. In order to understand the complex biology of E2F1, it is necessary to study its range of possible post-translational modifications and the functional and biochemical consequence of each on E2F1 activity (Fig 1.4).

1.8 E2F1 post-translational modifications

Transcription factors constitute a small fraction of the human proteome, and yet they integrate vast layers of signals and stimuli to coordinate the cellular response at the most basic level. A multitude of signalling pathways converge and manifest as different post-translational modifications on transcription factors to influence their intracellular localization, stability, target gene specificity and interaction partners [82]. Therefore the

study of post-translational modifications and their complex interplay, is critical to understanding how and when a given transcription factor regulates the cellular response. The addition of post-translational modifications is mediated by the so called 'writer' proteins. In many instances, these modifications are found to be reversible, with their removal being mediated by 'eraser' proteins. An important feature of these modifications, which also relates to their ability to regulate their target protein, is their recognition by specific and dedicated 'reader' proteins [83]. For instance, acetylation of lysine residues can be added by the histone acetyl transferase (HAT) enzyme P/CAF, and removed by the histone deacetylase enzyme HDAC. Meanwhile, the acetylated lysine residue may be recognised by a bromodomain, which could therefore target the acetylated protein to a particular pathway and outcome [84].

As previously mentioned, E2F1 accumulates under DNA damage to regulate and coordinate the apoptotic response. The DNA damage responsive kinases ATM/ATR and CHK2 have been shown to rapidly phosphorylate E2F1 on serine residues S31 and S345 respectively, to stabilize E2F1 under stress conditions [85, 86]. The 14-3-3τ protein has been reported to bind to phosphorylated S31 of E2F1 and protect it against ubiquitination and proteasome mediated degradation [87]. In addition to these phosphorylation events E2F1 is also acetylated on multiple lysine residues by several acetyl transferases including p300/CBP, P/CAF and TIP60. Acetylation of E2F1on lysine residues K117, K120 and K125 have been shown to augment its stability, and increase its DNA binding affinity for the pro-apoptotic gene promoter *TP73* [88-90]. Therefore DNA damage induced acetylation and phosphorylation of E2F1 potentiates its pro-apoptotic activity.

The lysine methyl transferase SETD7 has been reported to methylate E2F1 on K185, whilst the demethylase LSD1/KDM1A has been shown to remove it [91]. Whilst lysine methylation suppresses the pro-apoptotic activity of E2F1, by hindering acetylation and

phosphorylation at distal sites, demethylation of E2F1 under DNA damage potentiates its pro-apoptotic activity. Accordingly, it has been suggested that LSD1 helps maintain a pool of unmethylated E2F1, which can become readily acetylated and phosphorylated under DNA damage and targeted to apoptotic genes. Interestingly in a separate study, E2F1 was found to undergo NEDDylation, in a manner that was influenced by the methylation status of K185 [92]. It appeared that methylation of this lysine residue could augment the NEDDylation of E2F1, and target it for degradation. Accordingly under DNA damage, E2F1 NEDDylation levels were shown to decrease.

In a recent study E2F1 was found to undergo asymmetric arginine methylation on R109 by PRMT1, and symmetric methylation on R111 and R113 by PRMT5 [93, 94]. Interestingly these two modifications were observed to target E2F1 to distinctly different outcomes, and furthermore they occurred on E2F1 in a mutually exclusive and competitive manner. It could be shown that whilst asymmetric arginine methylation accumulates under DNA damage and augments the pro-apoptotic activity of E2F1, symmetric arginine methylation destabilizes E2F1 and suppresses apoptosis. Interestingly the study also identified a reader of the symmetric methyl marks. The tudor domain of the methyl binding protein TSN (Staphylococcal Nuclease And Tudor Domain Containing 1) was found to specifically recognise symmetrically methylated R111 and R113, thus offering insight into how these methylation marks could influence the stability and target gene specificity of E2F1. Figure 1.4 summarises the range of post-translational modifications described on E2F1.

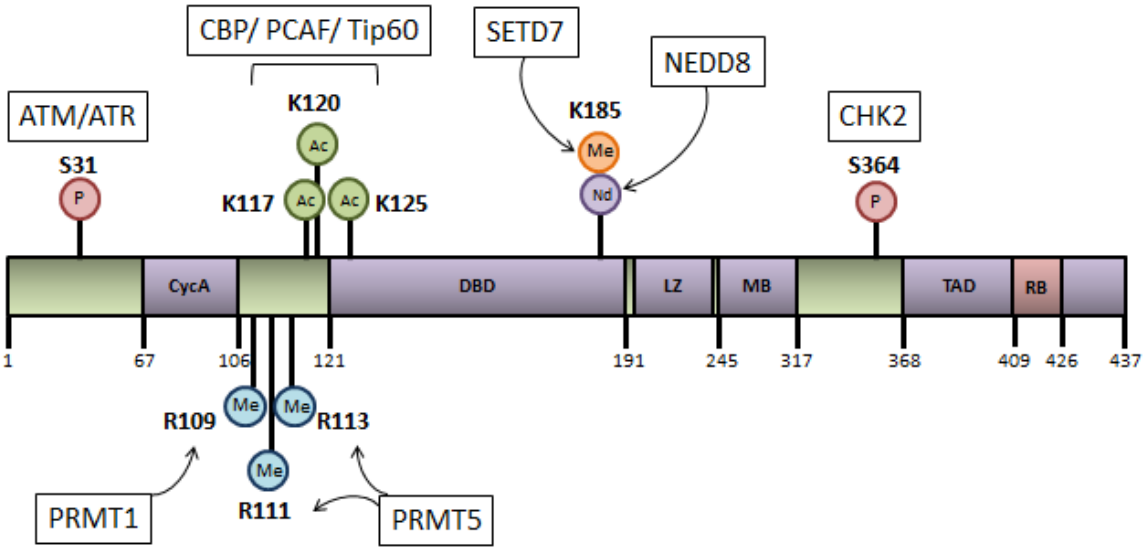


Figure 1.4) Post-translational modifications of E2F1

A schematic diagram to depict the various post-translational modifications of E2F1 and the enzyme responsible for them; the letter P denotes phosphorylation, Ac acetylation, Me methylation ad Nd NEDDylation. CycA: cyclin A binding domain, DBD: DNA binding domain, LZ: leucine zipper, MB: marked box, TAD: transactivation domain, RB: pRB binding domain

1.9 E2F1 and the epigenetic landscape

Several of the enzymes which directly modify E2F1 can also engage with it to regulate the epigenetic landscape of the cell, thus highlighting the complex interplay and feedback mechanisms between these various cellular factors [95]. The epigenetic landscape of a cell during the cell cycle has been reported to be dynamic, with various histone modifications peaking and declining at particular phases [96, 97]. E2F1 has been found to be associated with various epigenome regulators to influence the local epigenetic signatures and gene expression patterns as outlined below.

As cells progress through G1 and S phase they exhibit increased global acetylation levels, which is associated with gene activation. The cell cycle dependent recruitment of E2F1-3 to chromatin during the G1-S transition has been shown to coincide with the acquisition of activating histone marks, particularly histone proteins H3 and H4 acetylation. Accordingly, E2F1 has been reported to interact with TRRAP, which as a component of several acetyl transferases, helps recruit the TIP60 HAT to target promoters. The E2F1 dependent recruitment of TIP60 to chromatin in turn has been shown to directly promote hyperacetylation of histone proteins H3 and H4, and facilitate the timely expression of proliferation associated genes [98].

In addition to histone acetylation, certain histone lysine methylation events have also been associated with transcriptional activation, including methylation on H3K4. Accordingly, the host cell factor, HCF-1, has been shown to interact with E2F1 to recruit the H3K4 methyl transferases MLL and SETD1 to E2F target promoters during G1-S transition, and thus facilitate cell cycle progression [99]. In another study, the bromodomain containing protein ANCCA was found to interact with E2F1 and to recruit MLL1 and

MLL2 to target promoters, such that siRNA depletion of ANCAA could be shown to impair expression of cell cycle associated genes and inhibit proliferation [100].

The SWI/SNF nucleosome remodelling complex utilizes the energy from ATP to disrupt the chromatin structure. Depending on the context this can either positively or negatively affect gene transcription. The SWI/SNF complex has a core ATPase (either BRG1 or BRM), plus seven non-catalytic subunits. A variant of one of the major subunits of the complex, ARID1B, has been shown to associate with E2F1, in a manner which coincides with the timely recruitment of SWI/SNF complex to cell cycle gene promoters and their transcriptional activation. Indeed, depleting cells of ARID1B delays their re-entry into the cell cycle [101]. Overall therefore one can appreciate a complex network of players, where chromatin modifiers target E2F1 to modulate its activity, whilst they themselves become recruited to chromatin by E2F1 to further fine-tune gene expression outcomes.

1.10 Cancer and inflammation

In 1863 Rudolf Virchow observed that leucocytes are present in neoplastic tissue, and accordingly he proposed that inflammation may create an environment that promotes and sustains cancer development [102]. Inflammation is a complex process, and acute inflammation is a necessary physiological response to various pathogens and insults. However chronic inflammation, also dubbed the 'secret killer', has been associated with many pathological problems and long term damage, including rheumatoid arthritis, asthma and certain cancer types [103]. Chronic inflammation can contribute to tumorigenesis by facilitating and regulating the acquisition of several key cancer hallmarks, including sustained proliferation, evasion of apoptosis, enhanced angiogenesis and metastasis [3]. At a

molecular level, the continual supply of cytokines and chemokines, reactive oxygen species (ROS), matrix metalloproteinses (MMPs) and activation of transcription factors such as NF-κβ (nuclear factor κβ), STAT3 (signal transducer and activator of transcription 3), AP-1 (activator protein 1) and HIF-1α (hypoxia inducible factor-1α) all contribute to the development of cancer in an inflammatory milieu [103].

Over two decades ago TNFα (tumour necrosis factor alpha) was isolated as the first anti-cancer cytokine, although soon after it was paradoxically shown to have tumour promoting properties *in vivo*. TNFα is produced not only by most immune cells, but also by many cancer cells, be it solid tumours or hematopoietic cancers. Furthermore, circulating TNFα levels have been found to be higher in some cancer patients with more advanced disease, with TNFα levels directly correlating with the extent of disease [104]. The crucial evidence for the tumour promoting versus the anti-cancer activities of TNFα emerged when $TNF\alpha^{-/-}$ transgenic mice challenged with carcinogen were surprisingly found to develop fewer tumours [105]. Returning to Virchow's original postulation, TNFα produced by cancer or stromal cells is now considered to contribute to tumorigenesis by promoting inflammation in the tumour microenvironment.

At a cellular level, TNFα can promote cell proliferation and survival, transcription of pro-inflammatory genes, or cell death. Activation of NF-κβ by TNFα is largely responsible for the induction of pro-survival and pro-inflammatory genes, whilst formation of the death inducing signalling complex (DISC) downstream of the TNF receptor, can lead to caspase mediated cell death or necroptosis. Crucially also, in the tumour microenvironment, TNFα can promote the recruitment of various leukocytes, including tumour associated macrophages (TAMs), which collectively promote tumour growth and spread [106].

Various interleukins (ILs) including IL-1, IL-6, IL-8 and IL-17, have also been associated with inflammation and cancer largely through activation of NF-κβ and AP-1 signalling pathways. IL-1β for instance, shown to be produced by myeloma cells, promotes the release of IL-6 by bone marrow stromal cells and acts in an autocrine fashion to stimulate myeloma cell growth and survival [103]. IL-1β has also been demonstrated to upregulate HIF-1α via NF-κβ, and to accordingly stimulate the release of the proangiogenic factor VEGF [107].

Overall, an important pathway or mechanism linking inflammation and cancer appears to be the NF-κβ transcription factor [108, 109]. Various cytokines and infectious agents, as well as growth factor receptors and kinases can lead to activation of NF-κβ by promoting its re-localization from the cytoplasm to the nucleus. Activated NF-κβ can then regulate the expression of a plethora of genes implicated in cell survival and proliferation, inflammation and immune regulation, angiogenesis, invasion and metastasis [103]. Interestingly, E2F1 has been shown to interact with the p65 subunit of NF-κβ, and to become co-recruited to a number of target genes. In fact, genome wide mapping of p65 binding to promoters has illustrated an enrichment of E2F1 binding motif within these loci [81]. Accordingly, E2F1 has been shown to regulate the expression of various pro-inflammatory chemokines and cytokines, notably TNFα and IL-1β, which as described previously have close ties with both cancer and inflammation. Therefore, although not explicitly explored, it is very probable that E2F1 could also contribute to tumorigenesis via its ability to promote inflammation.

1.11 Peptidyl arginine deiminases

Citrullination is a modification of arginine residues (Fig 1.5(a)). Unlike most other post-translational modifications, where a functional moiety, such as an acetyl or methyl group, is covalently added to the target amino acid, during citrullination the arginine residue is structurally converted to a citrulline. As a result, citrulline was originally considered a non-standard amino acid which is added/ created post-translationally. The family of enzymes which mediate citrullination are known as peptidyl arginine deiminases (PAD), of which five mammalian isozymes have been identified. These enzymes exhibit very high sequence homology, particularly in their C-terminal catalytic site, and yet display very specific tissue distribution. In humans, PAD1 is found in the skin epidermis, PAD2 in brain, skeletal muscle and hematopoietic cells, PAD3 in hair follicles, PAD4 in granulocytes and monocytes and PAD6 in embryonic stem cells and oocytes. All mammalian PAD enzymes have a high and specific requirement for calcium, and can only catalyse peptidyl citrullination, and not the conversion of free L-arginine to free L-citrulline (which occurs as part of the urea cycle by nitric oxide synthase) [110-113].

PAD4 is the only PAD member with a classical monopartite nuclear localization signal (NLS) (56-PPAKKKST-63), and as such is predominantly nuclear [114]. Interest in the PAD enzymes, particularly PAD4, was sparked as histones were found to be major substrates of PAD4 in the nucleus. A number of subsequent reports suggesting that PAD4 mediated citrullination could antagonize arginine methylation, fuelled excitement in the field that PAD4 may be the enzyme that reverses arginine methylation [115, 116]. This however was later dismissed as unlikely, as the rate of citrullination of methylated arginine residues *in vitro* were found to be up to 10,000 fold slower than unmethylated analogues [117]. Nonetheless, citrullination of an arginine residue does preclude and prevent

its methylation, and so these two modifications are truly antagonistic and mutually exclusive. Consequently, one would expect an interesting interplay between arginine methylation and citrullination, as already demonstrated by various groups in the context of gene regulation and histone modification [116, 118, 119].

In parallel with these studies, it was reported that PAD4 is overexpressed in multiple malignant tumours, and that high levels of PAD4 could be detected in the serum of patients with various malignancies [120, 121]. However little was known at the time about the functional significance of PAD4 and its contribution to tumorigenesis. On the basis that E2F1 can be methylated on multiple arginine residues (hence a likely interplay with citrullination) and the knowledge that E2F1 and PAD4 are both relevant to cancer biology, we set out to investigate E2F1 as a novel substrate for PAD4.

1.12 Structure of PAD4

PAD4 has an elongated structure and consists of two main domains; the N-terminal and the C-terminal. The N-terminal domain (M1-P300) consists of two immunoglobulin-like domains, and the C-terminal domain (N301-P663) contains the enzymatic active site. As mentioned, all PAD enzymes are specifically dependent on calcium for catalysis. The crystal structure of PAD4 revealed five binding sites for calcium ions, two in the C-terminal region close to the catalytic cleft, and three in the N-terminal region (Fig 1.5(b)). The enzymatic activity of PAD4 is critically dependent on the two C-terminal calcium ions, as they appear to restructure and stabilize the catalytic active site. Accordingly, almost all the residues which coordinate the binding of these two calcium ions are conserved in the PAD enzymes (except for PAD6), highlighting their importance for PAD activity. The remaining three calcium ions also appear to order and stabilise the structure

of PAD4, although they seem to have minimal effect on enzyme activity [122]. According to *in vitro* kinetics studies, the concentration of calcium that yields half-maximal activity ($K_{0.5}$) for PAD4 is in the mid to high micromolar range [117]. Crystallography studies also demonstrated that PAD4 can occur as a head-to-tail dimer (Fig 1.5(c)). The citrullination reaction mediated by PAD4 (and indeed other PAD isozymes) is hydrolytic, as the source of oxygen in the citrulline residue is from a water molecule. Mechanistically, two sequential nucleophilic attacks initiated by C645 and H471 in the active site of PAD4, mediate the hydrolytic conversion of the arginine residue to a citrulline to release ammonia as the by-product (Fig 1.6(a)) [122].

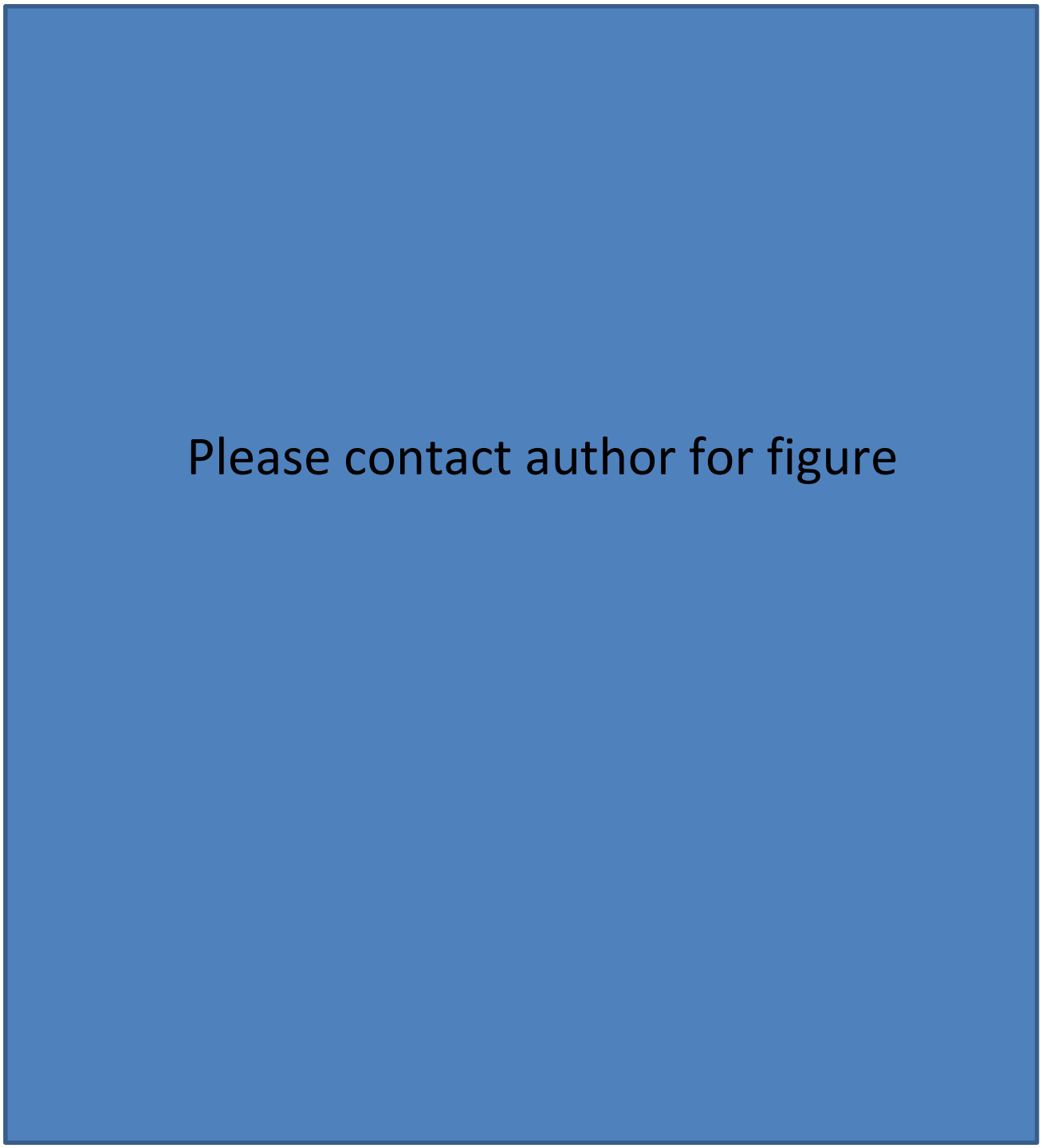


Figure 1.5) Citrullination and the structure of PAD4

(a) The citrullination reaction catalysed by the peptidyl arginine deiminase family of enzymes. The enzymes are calcium dependent. The peptidyl arginine side chain is catalysed to a peptidyl citrulline, by releasing ammonia in the process (diagram taken from [110]). (b) The crystal structure of PAD4, showing the C-terminal domain, and the two immunoglobulin-like subdomains in the N-terminal domain. There are five calcium binding sites, two in the C-terminus and 3 in the N-terminus. (c) The crystal structure of PAD4 dimer showing a head to tail contact between two PAD4 enzymes (Crystal structures taken from [122]).

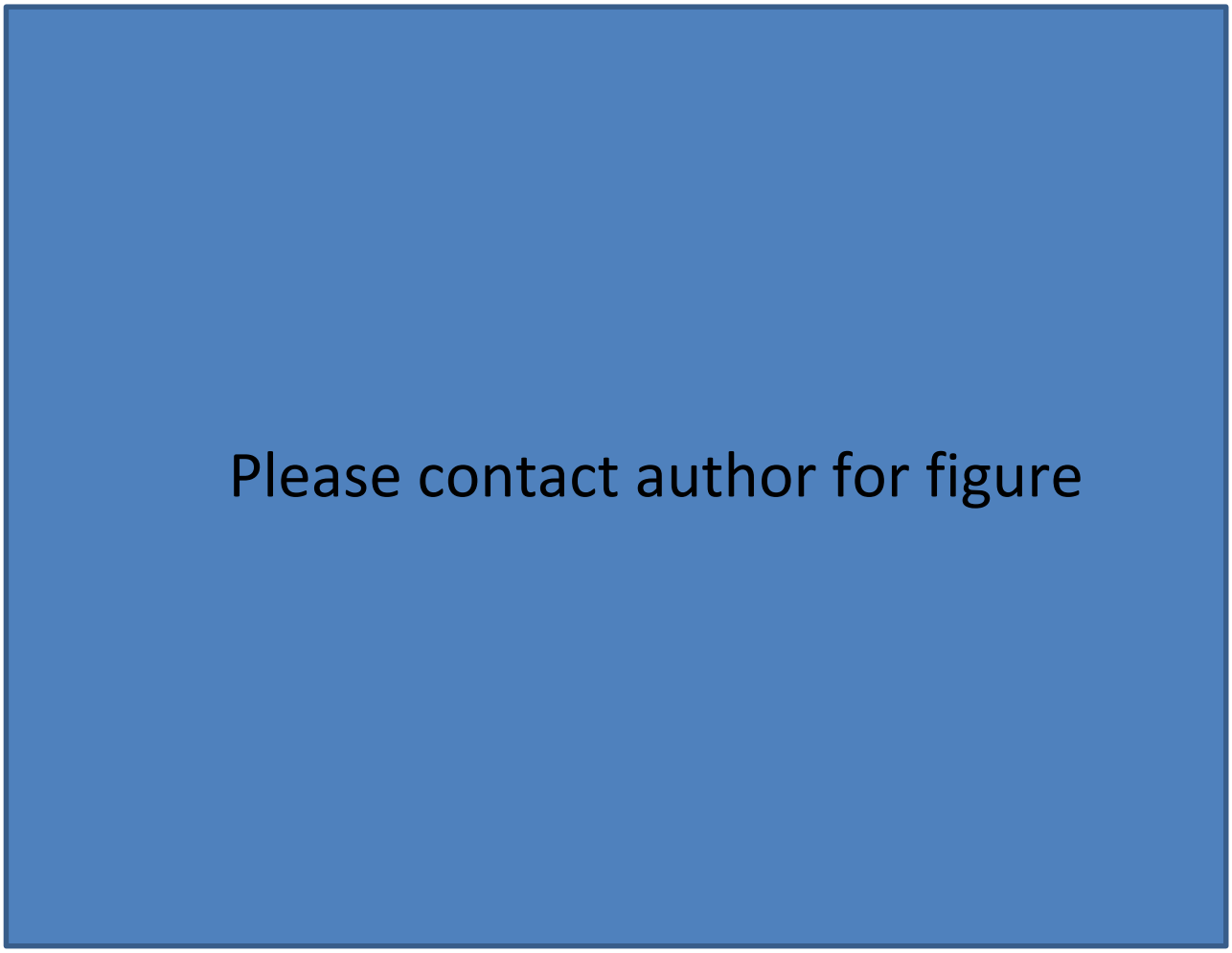


Figure 1.6) Proposed mechanism of citrullination by PAD4

(a) In the first step there is a nucleophilic attack by the thiol group of Cys645 on the C atom of peptidyl-L-arginine, with the side chain carboxyl groups of Asp350 and Asp473 forming hydrogen bonds and a salt bridge with the N atoms of the substrate. Subsequently the chemical bond between the C and N atoms is cleaved to generate ammonia. A second nucleophilic attack by water molecule, activated by the general base His471 then happens. Hydrolysis of the tetrahedral adduct releases peptidyl-L-citrulline as the final reaction product. (Reaction mechanism taken from [122]).

1.13 Biological functions of PAD4

PAD4 was originally identified in myeloid leukaemia HL60 cells differentiating to granulocytes [123], and was later shown to be abundantly expressed in peripheral blood neutrophils and eosinophils [124]. High levels of PAD4 have also been demonstrated in various malignancies [125] and pluripotent stem cells [126]. Accordingly, we shall review the physiological roles of PAD4 in the context of the immune system, tumorigenesis and pluripotency respectively.

1.11.1 PAD4 and the immune system

Neutrophils are the most abundant human leukocytes, and are essential to the innate immune system. They have a short life span of 24-48 h and continually undergo apoptosis in the absence of inflammatory stimuli. However, when challenged with various micro-organisms (e.g. bacteria, fungi or protozoa) or inflammatory stimuli (e.g. LPS, IL-8 or TNFα), they become activated. Activated neutrophils employ various tactics to tackle the invading pathogen, including phagocytosis of the invading pathogen, release of lytic enzymes, production of reactive oxygen species with antimicrobial properties and release of 'Neutrophil Extracellular Traps' (NETs) in a process known as NETosis [127, 128]. NETosis is a specialised form of cell death, during which the neutrophils decondense their chromatin, load it with various antimicrobial molecules, and release them to trap and kill the invading pathogens [128]. PAD4 has been shown to citrullinate core histones in the nucleus, including histone H3, H2A and H4 [129]. Since citrullination neutralizes the positive charge of arginine residues (hence reducing net histone charge), hypercitrullination of histones by PAD4 has been suggested to be important for the process of NETosis, as it facilitates the opening and decondensation of chromatin [130]. Inappropriate NET formation

and degradation has been described in some autoimmune diseases, including systemic lupus erythematosus (SLE), where they are known to cause vascular and organ damage. In support of the role of the PAD enzymes in NETosis, the pan-PAD inhibitor BB-Cl-amidine was shown to reduce NET formation and protect against kidney, skin and vascular damage in an *in vivo* model of SLE [131].

PAD4 expression has also been detected in haematopoietic stem cells, where it is reported to form a repressor complex with the histone deacetylase HDAC1 and the transcription factor Lymphoid Enhancer Factor 1 (LEF1) to suppress the transcription of the c-myc gene. The recruitment of PAD4 was shown to coincide with reduced promoter arginine methylation, but increased citrullination. As c-myc expression levels were higher in PAD4 deficient progenitor cells, the study therefore concluded that PAD4 is likely to negatively regulate the proliferation of multipotent progenitor cells in the bone marrow [132]. The relevance of PAD4 as a transcriptional regulator in haematopoiesis was further illustrated by a second study showing that PAD4 can serve as a transcriptional cofactor for the transcription factor Tal6 in leukaemia cells. In this study, PAD4 was shown to counteract several histone methylation marks triggered by PRMT4 and PRMT6 and to augment or repress transcription in a gene specific manner. Based on colony formation assays, and promoter binding patterns of PAD4, the authors also suggested that PAD4 is likely to play an important role in haematopoietic differentiation [119].

Inappropriate PAD4 activity has been illustrated in many autoimmune diseases, notably rheumatoid arthritis (RA) [133, 134]. Rheumatoid arthritis has a prevalence of around 1%, and is characterised by chronic inflammation of the synovial joints and infiltration of blood derived cells, particularly macrophages into the affected joints [135]. There are large numbers of circulating antibodies directed against a range of autoantigens in the blood of RA patients. The most specific of these autoantibodies are directed against citrullinated

proteins, and these can be detected in up to 80% of patients with 98% specificity [136]. Detection of these 'Anti-Citrullination Protein Antibodies' (ACPA) has been the basis of diagnosing RA for many years. These ACPA's are predominantly produced locally in the inflamed joints, as there is an abnormal presence of citrullinated proteins in the affected synovium. PAD2 containing macrophages and PAD4 containing granulocytes infiltrating the affected joint are considered the main source of deiminase activity [137]. As these cells die and rupture, the PAD enzymes released into the local vicinity are thought to be exposed to abnormally high levels of calcium. The hyperactivation of these enzymes leads to citrullination of a range of extracellular and cytoskeletal protein, such as keratin and vimentin, which then further evoke an immune response [138]. A study investigating functionally relevant polymorphisms in the PAD4 gene identified 17 single nucleotide polymorphisms (SNPs), two of which correlated very strongly with rheumatoid arthritis. Interestingly these SNPs were found to increase the stability of the PAD4 mRNA, further reinforcing the relevance of this enzyme to the etiology of RA [139]. It must be stated however, that inappropriate PAD activity and expression is not limited to RA. In fact overexpression of PAD4 has been reported in various acutely inflamed tissues such as acute gastritis [121] and multiple sclerosis (MS) [140], and antibodies to citrullinated proteins have been detected in erosive systemic lupus erythematosus [141].

Periodontitis is an oral infection that ultimately leads to destruction of the tissue supporting the teeth. Furthermore, periodontitis has been proposed to be a candidate risk factor for RA. P*orphyromonas gingivalis* is one of the major pathogens implicated in this infective disease, and of note, it is the only known prokaryote shown to express a peptidyl arginine deiminase enzyme capable of citrullination[142]. In contrast to the mammalian PAD enzymes, *P.gingivalis* PAD (PPAD) is capable of citrullinating both free and peptidyl arginine, and is furthermore, not dependent on calcium ions for activity. Citrullination by

P.gingivalis has been suggested to provide a survival advantage to the bacteria, by interfering with complement activation as part of the immune response, as well as disrupting neutrophil function[143]. However, given the strong and specific association between circulating ACPA levels and RA, PPAD has been suggested to contribute to the onset of RA by citrullinating various proteins and peptides, including human fibrinogen and α-enolase[144]. Consequently, in the study of RA, in addition to the mammalian PAD enzymes, the contribution of the prokaryotic PPAD to disease onset and severity will also need to be taken into account. This will have particular implications for drug development in RA, where inhibition of PAD enzymes is being pursued as the target.

PAD4 has also been more directly implicated in the transcriptional regulation of various cytokines, including TNFα and IL-8. PAD4 mediated citrullination of H3R8, has been shown to displace the heterochromatin protein HP1α from the adjacent trimethylated H3K9 residue, and therefore cause transcriptional de-repression of the target gene [145]. It is argued that in MS, over-expression/over-activation of PAD4 contributes to the uncontrolled immune response by upregulating cytokine expression. Accordingly, siRNA depletion of PAD4 or its chemical inhibition could be shown to reduce the expression of a number of cytokines [145]. In line with these findings, PAD4 has been shown to be activated downstream of inflammatory stimuli, such as LPS or TNFα signalling [146]. In an independent study, in MS, PAD4 was shown to be re-localised to the nucleus in response to TNFα stimulation [147]. Collectively these results suggest a role for PAD4 as a transcriptional regulator in the immune response.

1.11.2 PAD4 and tumorigenesis

PAD4 has been shown to be overexpressed in a variety of malignancies including lung, breast, oesophagus and ovaries [120]. However defining PAD4 as a tumour suppressor or oncogene, or even as a transcriptional activator or repressor, has been very difficult. As a nuclear factor, PAD4 has been shown to interact with various transcription factors, to either citrullinate them directly, or to citrullinate core histones. A number of these interactions and their consequence are discussed below:

PAD4 and ELK1 [148]: PAD4 ChIP-chip analysis in MCF7 breast cancer cell line illustrated that PAD4 is predominantly enriched near gene promoter transcription start sites (TSS). Crucially, the global genomic binding pattern of PAD4 was found to significantly correlate with various activator transcription factors including ELK1 and E2F1. By consolidating the PAD4 ChIP-chip analysis against a gene expression microarray dataset, PAD4 was found to be more than 5 fold enriched on the TSS of active genes versus inactive, under-expressed genes. Accordingly it was concluded that PAD4 mediates a positive role in gene transcription, for which the regulation of ELK1 transcription factor by PAD4 was further explored. ELK1 is a member of the Ets family of transcription factors and is a downstream target of the Ras signalling cascade. It binds to the serum response element in the sequence of the c-Fos gene to promote its transcription and expression. In this study, PAD4 was found to citrullinate Elk1 in response to growth factor stimulation, and to accordingly enhance ERK1/2 mediated phosphorylation of Elk1, histone acetylation and gene activation. As such, suppression of PAD4 levels or activity was found to diminish c-Fos expression levels.

PAD4 and DNMT3A [149]: DNMT3A is a DNA methyltransferase, with transcriptional repressive functions in cells. It has *de novo* methyltransferase activity and can target unmethylated CpG Island for methylation. PAD4 was found to associate with DNMT3A

and target it for citrullination to therefore increase its stability and augment DNA methylation. Specifically, the promoter methylation of $p21^{Cip1}$ was shown to decrease upon cellular depletion of PAD4, supporting the claim that PAD4 could regulate DNA methylation patterns and gene expression outcomes via DNMT3A.

PAD4 and GSK3β [150]*:* Depletion of PAD4 in MCF7 cells was found to lead to enhanced cell invasiveness, reduced levels of the epithelial markers E-cadherin and β-catenin and increased levels of the mesenchymal marker vimentin. Mechanistically, PAD4 was shown to citrullinate the serine-threonine kinase GSK3β at the N-terminus to increase its nuclear retention, from where it acts as a key regulator of various transcription factors implicated in tumour progression. As such, loss of PAD4 was shown to reduce nuclear levels of GSK3β, leading to epithelial-to-mesenchymal transition and increased invasiveness of breast cancer cells. This was validated *in vivo*, allowing the authors to argue for a tumour suppressive role for PAD4.

PAD4 and p53: There have been numerous studies investigating the interplay between PAD4 and p53. In one of the earlier studies, PAD4 was reported to be a transcriptional target of p53, as treatment of cells with doxorubicin induced PAD4 levels and augmented total cellular levels of citrullination in a p53 dependent manner. Under these conditions, nucleophosmin (NPM) was found to undergo citrullination and accordingly, relocate from the nucleolus to the nucleoplasm. In terms of cellular outcomes, it was suggested that whilst ectopic PAD4 reduced cell growth and expansion, its depletion could decrease the extent of apoptosis. Hence overall a pro-apoptotic function was ascribed to PAD4 [151]. In a subsequent paper, PAD4 was found to citrullinate core histones, particularly H4 in response to doxorubicin, in a p53 dependent manner. Mechanistically, this was suggested to facilitate cleavage of chromatin during the apoptotic process. Overall therefore the authors argued a p53 dependent pro-apoptotic and

tumour suppressive role for PAD4, largely based on its ability to citrullinate H4 under damage conditions [152].

In a follow-up paper however, PAD4 was reported to be a negative regulator of p53. Treatment of cancer cells with the general pan PAD inhibitor Cl-amidine was shown to reduce cell numbers and increase $p21^{Cip1}$ levels in a p53 dependent manner. Similarly, siRNA depletion of PAD4 levels resulted in reduced cell numbers and increased levels of p53 target genes, as PAD4 mediated histone citrullination was shown to suppress transcription by antagonising arginine methylation. In a follow up paper, HDAC1 was reported to interact with PAD4, and become recruited to target promoters to further suppress gene transcription. Consequently it was shown that Cl-amidine and the HDAC inhibitor SAHA acted synergistically to induce expression of p53 target genes and cause cell death [153-155].

Overall therefore it appears that the function of PAD4 and its relevance to tumorigenesis is still far from clear. Whether PAD4 is a transcriptional coactivator or repressor is probably context dependent, and whether it suppresses or augments tumour progression is in need of further studies.

1.11.3 PAD4 and Pluripotency

Given PAD4's ability to induce histone hypercitrullination and chromatin decondensation, its putative role in pluripotency was investigated on the basis that an open chromatin state is required in pluripotent cells. Indeed PAD4 was found to be induced in embryonic stem cells, with its levels and activity positively correlating with pluripotency factors. Microarray data suggested that whilst ectopic expression of PAD4 correlated with embryonic stem cell development, treatment with Cl-amidine, correlated with cellular

differentiation. Through SILAC labelling, histone H1 was identified as one of the major substrates of PAD4 in embryonic stem cells, as H1 citrullination reduced its ability to bind to nucleosomes and facilitated opening of the chromatin [126].

In conclusion, PAD4 is emerging as a very interesting nuclear enzyme with widespread involvement in many (patho)-physiological states. However, as already alluded to, PAD4's cellular and organismal functions are far from fully understood. Furthermore, with mounting interest in PAD4 as a pharmacological target, the need to understand and explore this enzyme in more depth also grows.

1.14 Research objectives

E2F1 is an interesting transcription factor and appears to be very versatile. Not only is it implicated in cell growth and division, but it also acts as a very responsive checkpoint protein, capable of coordinating and regulating DNA damage repair and apoptotic programmes in response to cellular insults. Furthermore, E2F1 continues to be implicated in diverse roles, including metabolism, differentiation and immune response regulation. Accordingly it is important to understand how E2F1 mediates these various functions, and coordinates its tumour promoting versus suppressing activities. In this respect, the study of the post-translational modifications of E2F1 is crucial to understanding its regulation and activity at the molecular level. Similar to the 'histone code' proposed for histone modifications, transcription factors are also known to undergo a vast array of modifications, which can cross-talk with each other and regulate the transcription factor in complex and subtle ways.

Previous work in our research group had illustrated that E2F1 can be methylated on R109 by PRMT1 and R111/R113 by PRMT5. It was shown that whilst methylation by PRMT1 enhances the pro-apoptotic activity of E2F1, methylation by PRMT5 suppresses it. Furthermore, an interesting cross-talk was observed between these two modifications, as they were found to occur in a mutually exclusive manner. These arginine residues and their modifications therefore appear to be very important for the regulation of E2F1 activity and stability. The peptidyl arginine deiminases are known to mediate citrullination of arginine residues. Whether these PAD enzymes could citrullinate an already methylated arginine has been disputed, but it is clear that a citrullinated arginine could not be methylated.

PAD4, as the only member with a nuclear localization signal, has been reported to be predominantly nuclear. Core histones and various other nuclear proteins have emerged as novel substrates of PAD4 in the nucleus, and PAD4 has become increasingly implicated in cancer and inflammatory disorders. Following these insights, we therefore investigated E2F1 as a novel substrate of PAD4, and explored the cellular context under which this modification might be of relevance.

In this study we have provided evidence for PAD4 mediated citrullination of E2F1 in cells and demonstrated that PAD4 can augment the transcriptional activity and DNA binding affinity of E2F1. Furthermore, we found that this modification is particularly relevant in myeloid leukaemia cells differentiated to granulocytes, as PAD4 regulated E2F1's recruitment to promoters of inflammatory genes. We uncovered a cross-talk between citrullination and nearby acetylation sites of E2F1, and demonstrated that a bromodomain reader of acetylated E2F1 can be regulated by adjacent citrullinated residues. Accordingly we took this interplay into an *in vivo* model of inflammation and illustrated an additive anti-inflammatory effect of a pan-PAD inhibitor and BET specific bromodomain inhibitor. In summary our findings suggest that E2F1 is a novel substrate for PAD4 and that citrullination of E2F1 is likely to be important for regulation of cytokine gene expression.

Chapter 2) Material and Methods

2.1 Cell culture

All cell lines used as part of this study were grown at 37 °C and 5% CO_2 in a humidified incubator. Adherent cell lines were maintained in Dulbecco's modified Eagle's medium (DMEM) (Sigma®) supplemented with 1% (v/v) Penicillin-Streptomycin (Pen-Strep) (Life Technologies Corp., USA) and 10 % (v/v) Foetal Calf Serum (FCS) (Biosera, USA). Tet-on inducible cell lines were maintained in DMEM medium supplemented with 1% Pen-Strep, 10% (v/v) tetracycline negative FCS, 100 µg/ml G418 (Clontech) and 150 µg/ml hygromycin B (Clontech). Non-adherent cell lines were maintained in RPMI-1640 (Sigma®) supplemented with 1% Pen-Strep and 20% FCS. All cell lines were passaged every 2-4 days. Adherent cell lines were passaged at a confluency of 70-90% using Trypsin (Lonza) for digestion and non-adherent cell lines at 6-7x10^6 cells per ml by spinning down cells and replacing media. Commonly used cell lines, including U2OS cells and MCF-7 cells, were re-ordered from ATCC cell bank every few years to ensure the cells do not genetically drift upon long-term use. Alternatively, the identity of the cell lines can be authenticated using polymorphic short tandem repeat (STR) loci sequencing, which gives a unique fingerprint for each cell line. Table 1 summarizes the cell lines used.

Table 1) Cell types used in this study

Cell line	Description	Type
U2OS	Bone osteosarcome	Adherent
MCF-7	Breast adenocarcinoma	Adherent
HEK293T	Human embryonic kidney cells overexpressing large T-antigen	Adherent
Tet-on inducible U2OS	Tet-On® gene expression system in osteosarcoma cell line (Clontech)	Adherent
HL60	Promyelocytic leukemia	Non-adherent

2.2 Antibodies

The various primary and secondary antibodies used in the study are listed in table 2.

Table 2) Antibodies used in this study

Antibody	Species	Supplier
E2F1 (KH95)(sc-251)	Mouse	Santa Cruz
E2F1 (C20)(sc-193)	Rabbit	Santa Cruz
pRB (IF8)(sc-102)	Mouse	Santa Cruz
DP1 (sc-53642)	Mouse	Santa Cruz
PAD4 (ab128086)	Mouse	Abcam
PAD4 (ab50332)	Rabbit	Abcam
β-Actin (A2228)	Mouse	Sigma
Anti-modified citrulline (17-437)	Rabbit	Millipore
GAPDH (sc-20357)	Goat	Santa Cruz
CDC6 (sc-9964)	Mouse	Santa Cruz
Cyclin A (sc-751)	Rabbit	Santa Cruz
BRD4 (ab128874)	Rabbit	Abcam
CD11B (ab75476)	Rabbit	Abcam
Flag (F3165)	Mouse	Sigma
HA (MMS-101R)	Mouse	Cambridge Bioscience
HRP-conjugated secondary IgG	Mouse/Rabbit	GE healthcare
Polyclonal IgG for IP	Mouse/Rabbit	Dako
HA/Flag-tagged agarose beads	Mouse	Sigma

2.3 DNA plasmid transfection

GeneJuice Transfection Reagent (Invitrogen) was added to reduced serum medium Opti-MEM® (Life Technologies) for 5 min at room temperature (RT) (1.5 µl of

transfection reagent per 1 µg plasmid DNA). The plasmid DNA was then added to the mixture, and incubated for a further 15 min at RT. The mixture was then added drop wise to cells, which were typically transfected at 40-70% confluency. The cells were harvested 36-48 hr post-transfection. The DNA plasmid used in this study are listed in table 3.

Table 3) List of expression vectors used

Plasmid	Expression
pcDNA 3.1 HA-E2F1	Mammalian
pCMV 7.1 3xFlag-E2F1	Mammalian
pCMV 7.1 3xFlag-E2F1 (1-194)	Mammalian
pCMV 7.1 3xFlag-E2F1 (194-437)	Mammalian
pCMV 7.1 3xFlag-E2F1 (1-407)	Mammalian
pcDNA 3.1 HA-PAD4	Mammalian
pcDNA 3.1 Flag-PAD4	Mammalian
pcDNA 3.1 Flag-PAD4 (C645A)	Mammalian
pcDNA3.1 vector	Mammalian
pCMV-Flag vector	Mammalian
pGEX-E2F1	Bacteria
pGEX vector	Bacteria
CDC6/p73/DHFR PGL3-Luciferase	Mammalian
β-gal	Mammalian
pTRE Flag-PAD4	Mammalian
pTRE vector	Mammalian

2.4 Small interfering RNA (siRNA) transfection

Oligofectamine Transfection Reagent (Life Technologies) and the siRNA of interest were separately incubated with reduced serum medium Opti-MEM® (Life Technologies)

for 5 min at RT. The total volume of oligofectamine added was in accordance with the manufacturer's guidelines. The two mixtures were then mixed, and incubated for a further 15 min at RT. The final mixture was added drop wise to cells, which were typically transfected at 30-50% confluency. The cells were harvested 72 hr post-transfection. The siRNA sequences used in this study are listed in table 4.

Table 4) List of siRNA sequences

siRNA	Sequence
E2F1	AAC UCC UCG CAG AUC GUC AUC
PAD4	GGU CCU GCU ACA AAC UGU UTT
Non-targeting control	Dharmacon NT2/NT3 control siRNA

2.5 Drug treatments

The various compounds, their final cellular concentrations and duration of treatment are listed in table 5. Where the compound is dissolved in DMSO, control groups are treated with equal volumes of DMSO alone. All drugs were directly added to culture medium.

Table 5) List of drugs/compounds used

Compound	Final Concentration	Length of treatment	Supplier
TDFA	10-20 µM	16 hr	Kind gift of P.Thompson
TDHA	10-20 µM	16 hr	Kind gift of P.Thompson
BB-Cl-amidine	1-20 µM	4-16 hr	Kind gift of P.Thompson
JQ1	0.1-5 µM	4-16 hr	Kind gift of S.Knapp
GSK484	10 µM	4-16 hr	Kind gift of U. Oppermann

2.6 Immunoblotting

For adherent cells, media was removed from dishes and cells were washed once with PBS. To harvest and collect cells, they were manually uplifted using cell scrapers. For non-adherent cells, the culture was spun down at 1000 rpm for 5 min to pellet cells, before removing media and washing cells once with PBS.

Cells were lysed in modified RIPA lysis buffer (150 mM NaCl, 50 mM Tris-HCl [pH 7.5], 1% Igepal CA-630, 0.1% deoxycholic acid, 1 mM EDTA, 1 mM NaF, 1 mM Na_3VO_3 and protease inhibitor cocktail [0.5 µg/ml leupeptin, 0.5 µg/ml pepstatin, 0.5 µg/ml aprotonin]), on ice for 30 min. The lysed cells were centrifuged at 13,000 rpm for 10 min at 4 °C to remove cell debris, and the supernatant was transferred into a fresh microcentrifuge tube.

Protein concentrations were measured using Bradford reagent (BioRad), with 20-80 µg of protein loaded for each experiment. The lysate were mixed with SDS loading dye (62.5 mM Tris-HCl [pH 6.8], 25% (v/v) glycerol, 2% (w/v) SDS, 5% (v/v) β-mercaptoethanol and 0.0625% (w/v) bromophenol blue), and boiled at 95 °C for 5 min. Sample were then loaded on to 8-14% SDS-polyacrylamide gels and subjected to gel electrophoresis (SDS-PAGE), before being transferred onto nitrocellulose membrane (GE). To ensure equal and efficient transfer of proteins, the nitrocellulose membranes were briefly stained with Ponceau S (Sigma). The membranes were then blocked in 5% (w/v) milk PBST (PBS + 0.1% Tween-20) for 1 hr at RT, and subsequently with primary antibody diluted in blocking solution overnight (ON) at 4 °C. Following primary incubation, the membranes were initially washed three times with PBST (10 min intervals), then incubated with secondary antibody diluted in block for 1 hr at RT, and finally washed a further three times in PBST (10 min intervals). The membranes were then

incubated with in-house made ECL (2.5 mM luminol, 396 µM coumaric acid, 100 mM Tris-HCl [pH 8] and 0.02% (v/v) hydrogen peroxide) for 2 min, and the luminescence from the membranes were detected using Fuji Medical X-ray films (Fujifilm).

To detect citrullination with the anti-modified citrulline kit (AMC) (Millipore), manufacturer's guidelines were observed. The AMC kit enables the detection of citrulline-containing proteins by Western blot analysis in a two-step process. In the first step, citrulline-containing proteins which have been transferred onto nitrocellulose or PVDF membranes are formaldehyde-cross linked to the membrane to improve their retention. In the second step, the membranes are treated with 2,3-butanedione monoxime and antipyrine in a strong acid solution, to chemically convert the citrulline residues to a ureido group adduct. This ensures detection of citrulline-containing proteins regardless of neighbouring amino acid sequences. Detection of the modified citrulline proteins is subsequently performed using a standard immunoblot protocol with the anti-Citrulline and goat anti-rabbit IgG-HRP.

2.7 Bradford assay

A calibration curve was derived by dispensing 1-10 µg of BSA (Sigma) standard in Bradford reagent and measuring absorbance using a spectrophotometer at OD_{595}. Relative reading was plotted against the concentration of standard, and was used a reference for subsequent protein concentration measurements.

2.8 Co-immunoprecipitation

Cells were harvested as previously described. 5% of cell lysates were removed as loading input, with the remaining (typically 1-3 mg of cellular lysate) used for immunoprecipitation. The lysates were incubated with primary antibody ON at 4 °C, followed by 1 hr incubation with protein A and G agarose beads (Sigma). Alternatively the lysates were incubated with HA or Flag-tagged agarose beads for 2-4 hr at 4 °C. Following incubation, the beads were spun down, and washed three times with IP wash buffer (50 mM Tris-Hcl [pH 7.4], 150 mM NaCl, 0.1% Igepal CA-630, 1 mM EDTA, 1 mM NaF, 1 mM Na_3VO_3 and protease inhibitor cocktail) to remove unbound proteins. Protein complexes bound to beads were eluted by adding SDS loading dye and boiling for 5 min at 95 °C. The elution was centrifuged at 13,000 rpm for 1 min prior to loading on to SDS-polyacrylamide gel.

2.9 Chromatin immunoprecipitation (ChIP)

Cells were cross-linked with 1% formaldehyde for 10 min at RT, with occasional gentle shaking. 125 mM glycine was then added for 10 min at RT. For adherent cells, the media was removed and the cells were washed once with PBS, before being manually harvested into collection tubes. Non-adherent cells were spun down at 1,000 rpm for 5 min, and the cell pellet was then washed once with PBS.

Cell pellets were lysed in ChIP lysis buffer I (5 mM Tris-HCl [pH8.0], 85 mM KCl, 0.5% Igepal CA-630 and protease inhibitor cocktail) for 20 min on ice. The lysates were then centrifuged at 1,000 rpm for 5 min, the supernatant discarded, and the pellet lysed in ChIP lysis buffer II (10 mM Tris-HCl [pH 7.4], 150 mM NaCl, 1 mM EDTA, 1% Igepal

CA-630, 1% DOC, 0.1% SDS and protease inhibitor cocktail) for 10 min on ice. Subsequently, the lysates were sonicated on ice for 4-6 min to shear the chromatin to fragment sizes between 200-1000 base pairs. Following sonication, the samples were centrifuged at 13,000 rpm for 10 min. The supernatant was removed, and the protein concentration was measured using Bradford reagent. Equal concentrations of lysates were pre-cleared by incubating with IgG antibodies and protein A and G agarose beads (previously blocked with 1 mg/ml BSA and 400 µg/ml sonicated salmon sperm DNA (Invitrogen)) for 1-3 hr at 4 °C. The lysates were then centrifuged, and the supernatant transferred into a fresh collection tube. 50 µl of the supernatant was removed to be used as the ChIP input control, with the remaining incubated with primary antibodies ON at 4 °C.

The following day, protein A and G agarose beads were added to the samples for 1 hr at 4 °C. The samples were then centrifuged at 4,000 rpm for 1 min, and the beads were washed twice with Buffer I (20 mM Tris-HCl [pH 8.0], 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100 and protease inhibitor cocktail), twice with Buffer II (20 mM Tris-HCl [pH 8.0], 500 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100 and protease inhibitor cocktail), twice with Buffer III (10 mM Tris-HCl [pH 8.0], 250 mM LiCl, 1 mM EDTA, 1% Igepal CA-630, 1% DOC and protease inhibitor cocktail), and once with TE buffer (10 mM Tris-HCl [pH 7.4], 1mM EDTA and protease inhibitor cocktail). Prior to last wash, the beads were transferred into a fresh micro centrifuge tube.

The protein-chromatin complexes were eluted by incubating the beads with 250 µl elution buffer (1% SDS and 0.1 M $NaHCO_3$) for 15 min at RT. The beads were centrifuged and the eluates were collected into a fresh micro centrifuge tubes. This step was repeated to yield a final elution volume of 500 µl for each sample. To the ChIP input controls 40 µg/ml Proteinase K and 20 µg/ml RNAse A were added, and to the ChIP elutions 40 µg/ml Proteinase K, 20 µg/ml RNAse A, 200 mM NaCl, 10 mM EDTA and 40

mM Tris-HCl [pH 6.5] were added. The samples were then incubated at 55 °C for 3 hr (to remove proteins with the added Proteinase K) and subsequently 65 °C ON (to reverse the cross linkage). DNA was purified from samples using the QIAGEN DNA purification kit, and was used for subsequent PCR step. The primers used for the ChIP PCR are listed in table 6.

For double ChIP, the eluate collected from the primary immunoprecipitation, was diluted 10 fold in ChIP lysis buffer III (10 mM Tris-HCl [pH 7.4], 150 mM NaCl, 1 mM EDTA, 1% Igepal CA-630 and protease inhibitor cocktail) and incubated with the second antibody ON. The beads were then added and the samples were washed, eluted and reverse cross linked as previously described.

Table 6) List of ChIP primers used

Primers	Sequence
Albumin	F: TGGGGTTGACAGAAGAGAAAAGC R: TACATTGACAAGGTCTTGTGGAG
CDC6	F: GGCCTCACAGCGACTCTAAGA R: CTCGGACTCACCACAAGC
Cyclin A	F: CTGCTCAGTTTCCTTTGGTTTACC R: CAAAGACGCCCAGAGATGCAG
p73	F: TGAGCCATGAAGATGTGCGAG R: GCTGCTTATGGTCTGATGCTTATG
Apaf1	F: CTTGGCCAGGCTGGTCTTGAAT R: GCAGCCATTCAAAATTATGACACAT
E2F1	F: AGGAACCGCCGCCGTTGTTCCCGT R: GCTGCCTGCAAAGTCCCGGCCACT
IL-1β	F: GTCATATCAATTTATAGTCCCACGCGTAAT R: CTCACACCCCAGATAAAGAGATAACTTGTT

CCL3	F: GTCTGAAACCAGCTCTCCTCTTTATAGGCA R: GGACTGACTAAGAATAGCCTTGGGTTGACA
TNFα	F: CGATGGAGAAGAAACCGAGACAGAAGGTG R: AGTTGCTTCTCTCCCTCTTAGCTGGTCCTC

2.10 Immunofluorescence

Cells were seeded onto 13 mm glass coverslips, and were cross linked with 3.7% (v/v) formaldehyde/PBS for 5 min at RT. They were then permeabilised with 0.5% (v/v) Triton X-100/PBS for 2-5 min, blocked with 5% (v/v) FCS in PBS-Tween (0.1%) for 30 min and incubated with primary antibody diluted in block for 1 hr at RT in a dark humidified box. Coverslips were subsequently washed 2-4 times with 0.025% (v/v) Tween in PBS, and incubated with secondary antibody for 30 min at RT. After 4-5 rounds of wash with 0.025% PBST, the cover slips were then mounted with using Vectashield containing DAPI (4,6-diamino-2-phenylindole).

2.11 Luciferase reporter assay

Cells were seeded in triplicates, and transfected with Luciferase reporter plasmid, β-galactosidase plasmid, and other plasmids of interest. The cells were washed once with PBS, harvested and lysed with 100 µl Reporter Lysis Buffer (Promega), following the manufacturer's guidelines. 50 µl of lysate was transferred into opaque 96-well plates and the luciferase activity was measured using a luminometer, which adds 100 µl of luciferase reagent (Promega). The luciferase readings were normalised to β-galactosidase activity to account for transfection efficiency. This was measured by mixing equal volumes of lysate and β-gal buffer (0.2 M Na_2PO_4 [pH 7.2], 2 mM $MgCl_2$, 0.7% β-mercaptoethanol and 0.44

M Ortho-nitrophenyl-β-galactopyranoside [ONPG]), and reading the absorbance using a spectrophotometer at OD_{420}.

2.12 Colony formation assay

1000 cells were seeded into 6-well plates in triplicate, and incubated for 10-14 days. Upon the emergence of colonies, media was removed from cells and they were washed once with PBS. Cells were then fixed in methanol for 10 min at RT, followed by staining with Crystal Violet stain for 10-20 min at RT. The stain was removed, and the colonies were copiously rinsed with deionised water, before counting colony numbers.

2.13 Flow cytometry (FACS)

Cells were seeded in triplicates, and harvested by Trypsin digestion. To fix cells, 75% ice cold ethanol was added drop wise to cells, and they were incubated at 4 °C for at least 1 hr. Following fixation, the cells were washed once with PBS, and were resuspended in PBS containing 125 U/ml RNAse A and 40 μg/ml propidium iodide (Sigma). Stained cells were analysed using flow cytometry (Accuri C6, BD Bioscience).

2.14 Cell fractionation

Cells were washed once with PBS and harvested into micro centrifuge tubes using a cell scraper. The cell pellets were resuspended in plasma membrane lysis buffer (PML) (10 mM HEPES [pH 7.5], 10 mM KCl, 0.1 mM EDTA, 2 mM $MgCl_2$ and protease inhibitor cocktail) and incubated on ice for 15 min, before the addition of 0.6% Igepal CA-

630 (final concentration) for a further 5 min. The lysates were briefly vortexed and centrifuged at 13,000 rpm for 1 min at 4 °C. The supernatant, containing the cytoplasmic fraction, was transferred to a fresh tube, and the pellet, containing the nuclear fraction, was lysed with nuclear membrane lysis buffer (NML) (25 mM HEPES [pH 7.5], 500 mM NaCl, 1 mM DTT, 10% glycerol, 0.1% Igepal CA-630, 5 mM $MgCl_2$ and protease inhibitor cocktail). The mixture was vortexed vigorously and sonicated at 23% amplitude to shear the chromatin. The lysate was then centrifuged at 13,000 rpm for 5 min and the supernatant containing the nuclear proteins was transferred to a fresh tube.

2.15 DNA transformation

100 ng of plasmid DNA was added to 25 µl competent *Escherichia coli* (*E.coli*) cells and incubated on on ice for 20-30 min. The bacterial cells were then heat shocked in a 42 °C water bath for 45 sec and transferred back on ice for 2 min. 150 µl Super Optimal Broth (SOC) media was added to cells, and they were incubated at 37 °C for 1 hr, to allow for metabolic recovery and expression of antibiotic resistance genes. The cultures were then spread onto agar plates containing 100 µg/ml ampicillin or 50 µg/ml kanamycin (depending on the antibiotic resistance gene of the plasmid), and incubated at 37 °C for 16 hr.

2.16 Mass spectrometry analysis of citrullination

Immunoprecipitation of Flag tagged E2F1 from cell lysates was carried out as previously described using Flag-agarose beads. To elute Flag-E2F1 from beads, the beads were incubated with 500 µg Flag peptide (Sigma) in 25 mM Tris-HCl [pH 6.5] and were

rocked at 1,000 rpm for 1 hr at 4 °C. The samples were centrifuged at 4,000 rpm for 1 min and the supernatant containing the eluted Flag-E2F1 was transferred into a fresh tube, and mixed with NuPAGE® LDS Sample Buffer and NuPAGE® Reducing Agent (Invitrogen). The samples were then run on NuPAGE® Bis-Tris gels, and silver stained using the SilverQuest™ Silver Staining Kit (Invitrogen) following manufacturer's guidelines.

The bands of interest from the stained gel were excised using a clean scalpel, cut into smaller fragments and transferred into fresh micro centrifuge tubes. The gel pieces were incubated twice in wash solution (50% methanol, 10% acetic acid, 40% ultrapure water) at RT, initially ON and subsequently for 2 hr. The solution was then removed, and the gel pieces were dehydrated with the addition of acetonitrile for 5 min at RT. Acetonitrile was removed, and gel pieces were dried for 2 min in a vacuum centrifuge. This was followed by the addition of 10 mM Dithiothreitol (DTT) for 30 min at RT to reduce the proteins. This solution was removed and replaced with acetonitrile for 5 min, and the gel pieces were dried for 2 min in a vacuum centrifuge. Gel pieces were then rehydrated by the addition of 100 mM ammonium bicarbonate for 10 min at RT. This solution was removed and the gel pieces were dried a final time, with the addition of acetonitrile for 5 min followed by 2 min in vacuum centrifuge.

The dried gel pieces were incubated with 30 µl of 20 ng/µl sequencing grade trypsin solution (Promega) in 50 mM ammonium bicarbonate, with occasional vortexing for 10 min at RT. Samples were then centrifuged and the excess trypsin was removed, and 5 µl of 50 mM ammonium bicarbonate was added to each sample, which were incubated at 37 °c ON. The following day, 30 µl of 50 mM ammonium bicarbonate was added to each sample for 10 min at RT. The digests were centrifuged and the supernatants were transferred to a fresh micro centrifuge tube, to which 30 µl of extraction buffer (50% acetonitrile, 45% ultrapure water, 5% formic acid) was added. These samples were

incubated for 10 min at RT with the occasional gentle vortex mixing. This step was repeated and the samples were dried in a vacuum centrifuge. The dried digests were resuspended in 20 µl solution containing 2% acetonitrile and 0.1% formic acid in ultrapure water, and were submitted for analysis by tandem mass spectrometry (LC-MS/MS).

2.17 Cycloheximide half-life assay

Cells were treated with 100 µg/ml cyloheximide (CHX) (Fluka), and were harvested at indicated time points. Protein levels were subjected to immunoblotting, and band intensities were quantified using Image J software (Novell) and plotted. The half-life of the protein of interest is expressed as the time it takes for levels to decline to 50% of original levels.

2.18 Reverse transcriptase polymerase chain reaction (RT-PCR)

RNA was extracted from cells using TriZol® reagent (Life Technologies) in accordance with manufacturer's guidelines. RNA concentration was measured using NanoDrop (Thermo Fisher Scientfic), and 1 µg of total RNA was reverse transcribed into cDNA using SuperScript® III reverse transcriptase (Life Technologies) in accordance with manufacturers guidelines. cDNA was used as template in PCR (using Paq5000 DNA polymerase (Agilent Technologies)) or qPCR (using Brilliant III SYBR® Master Mix (Agilent Technologies)).

2.19 *In vitro* citrullination

Active Flag-PAD4 enzyme (or catalytically inactive C645A mutant) was purified from HEK293T cells transfected with Flag-PAD4 plasmids. The lysates were immunoprecipitated using Flag-agarose beads, and enzyme released from beads using Flag-peptide elution as previously described.

In vitro citrullination was set up by incubating 0.1-1 μg Flag-PAD4 with 1-2 μg substrate in 20 μl buffer composed of 500 mM Tris-HCl [pH 7.4], 5 mM DTT and 10 mM $CaCl_2$, at 37 ˚C for 1 hr. The reaction was stopped by the addition of SDS loading buffer, and analysed by immunoblotting using the AMC kit (Millipore).

2.20 Differentiation of HL60 cells

HL60 cells were differentiated into granulocyte like cells upon treatment with 1% DMSO (Tissue culture grade, Sigma), or 1 μM All-Trans-Retinoic Acid (ATRA) for 48-72 hr. They were differentiated into macrophage/monocyte like cells upon treatment with 10 nM 12-O-tetradecanoylphorbol-13-acetate (TPA) for 48-72 hr. The expression of CD11B cell surface marker was used to confirm differentiation.

2.21 Generating Flag-PAD4 Tet-On inducible cell lines

Flag-PAD4 was sub-cloned from pcDNA 3.1 Flag-PAD4 plasmid into pTRE3G plasmid (Clontech), using the primers listed in table 7. Flag-PAD4 doxycycline inducible cell lines were created in accordance with manufacturer's guidelines, using the Tet-

Express™ Inducible Expression System (Clontech). Expression of Flag-PAD4 was induced by treating cells with 1 μg/ml doxycycline for 24-48 hr.

Table 7) Primers used to clone Flag-PAD4 into pTRE vector

Primer	Sequence
Forward	5'-CGGGATGCTAGCCCGACGATG
Reverse	5'-TCTAGAGATATCTCAGGGCACCATGTTCCACCA

2.22 CIA induction in DBA/1 mice

In vivo work was performed by the Kennedy Institute at the University of Oxford, with all procedures conducted according to the guidelines of the Animals and Scientific Procedures Act 1986 and the approval of the UK Home Office (project licence PPL7335). Collagen induced arthritis in DBA/1 mice was used as the animal model for rheumatoid arthritis. Briefly, 8-12 week old male DBA/1OIaHsd mice (Harlan) were immunised with Complete Freund's Adjuvant (CFA) (Desiccated M. tuberculosis (avirulent strain H37RA) (Difco Laboratories), Incomplete Freund's Adjuvant (Difco Laboratories) and Bovine type II collagen (prepared in house)) to induce paw arthritis. Upon emergence of paw swelling, the mice were injected daily for 9 days, with vehicle or drug of interest, and their paw measurements taken. The treated mice were then harvested on the 10th day, with their paw(s) removed for IHC staining and blood and lymph node extracted for cytokine measurements.

2.23 Preparing drugs for *in vivo* injection

Stock solutions of BB-Cl-amidine and JQ1 were dissolved in DMSO (Sigma). Prior to injection, 5 µl of the stock solutions were diluted in 95 µl of 10% (w/v) Hydroxypropyl beta cyclodextrin (Sigma) (prepared in sterile nuclease free water (Promega)), such that the final volume injected into each mouse was 100 µl and the overall DMSO concentration 5% (v/v).

2.24 Diamino-benzidine immunostaining on wax-embedded sections

Immunostaining of sections obtained from arthritic paws of treated mice, was done in collaboration with the Kennedy Institute at the University of Oxford. The sections were deparaffinised in Xylene (2x5 min) and 100% ethanol (2x2 min). They were then rehydrated with serial washes in 95%, 85% and 70% ethanol (5 min each). The endogenous peroxidase activity in the tissue sections was suppressed by incubating the sections with 3% hydrogen peroxide in water for 20 min. Antigen retrieval from the sections was performed by immersing sections in 0.05% (v/v) Tween-20 in 0.01 M citrate buffer [pH 6], and incubating them in a 92 ˚C water bath for 10 min. The slides were brought back to room temperature, rinsed once with PBS, and serum blocked for 2 hr at RT with 10% (v/v) goat serum in PBS. They were then Avidin/Biotin blocked using the Dako Kit, following the manufacturer's guidelines. The sections were rinsed once with PBS, and incubated with primary antibody (diluted in blocking buffer at the manufacturer's recommended concentration) at 4 ˚C ON. The following day, the sections were rinsed thrice in PBS (5 min each) and incubated with secondary antibody in blocking buffer for 2 hr at RT. The sections were rinsed thrice in PBS (5 min each) and treated with a few drops of Avidin/Biotin/HRP complex reagent (Vector Laboratories) in the dark for

30 min at RT. The sections were washed once with PBS and incubated with DAB substrate (Vector laboratories) and counterstained with Mayer's Haematoxylin. The sections were rinsed with water and coverslips were mounted using Di-n-butylphthalate (DPX) mounting media (Avonchem).

2.25 Peptide interaction assay (SPOT blot analysis)

The SPOT blot analysis was performed in collaboration with the Structural Genomics Consortium at the University of Oxford as previously described[156]. Briefly, the membranes, with the peptides hybridized on them, were blocked with 5% (w/v) BSA (Sigma) in 0.3% Tween-20 PBS (PBST) for 8 hr at RT. The membranes were washed twice with PBST and once with PBS and incubated with 1 μM recombinant His tagged bromodomain at 4 ˚C ON. The next day the membranes were washed thrice with PBST (5 min each), and blocked with 5% BSA in PBST for 1 hr at RT. The membranes were washed a further three times with PBST for 5 min each, and incubated with HRP-conjugated anti-His antibody (Abcam) in 2.5% BSA in PBST for 1 hr at RT. The membranes were washed three times with PBST for 20 min each, and were incubated with Pierce™ ECL Western Blotting Substrate (Thermo Scientific) for 2 min. The luminescence from the membranes was read and captured using the ImageQuant Las-4000 camera.

2.26 Purification of GST-tagged recombinant proteins

The plasmids of interest were transformed into BL21 *E.coli* (DE3) strain as described previously. Single colonies were inoculated in Luria-Bertani (LB) growth media

containing the relevant antibiotics, and were shaken at 37 ˚C with an Orbital shaker, until an OD_{600} of 0.6-0.8 was reached, at which point 1 mM IPTG (Sigma) was added to induce protein expression. The cultures were shaken for a further 2 hr and were pelleted at 8,000 rpm for 15 min. The bacterial pellets were re-suspended in PBS, and Triton X-100 was added to a final concentration of 1% (v/v). The cultures were incubated at 4 ˚C for 1 hr, and were then sonicated at 23% amplitude to break down bacterial cell wall and shear DNA. The lysates were centrifuged at 13,000 rpm for 30 min and the supernatant was transferred to a fresh tube. Glutathione Sepharose beads (GE Healthcare) were added to the tubes, and allowed to rotate at 4 ˚C for 2 hr to bind the GST-tagged protein. The proteins were then eluted from beads using 20 mM reduced glutathione in 50 mM Tris-HCl [pH 8.0]. Excess glutathione was removed using Slide-A-Lyzer Dialysis Cassette (Thermo Scientific) with a 10 kDa molecular weight cut off immersed in dialysis buffer (50 mM Tris-HCl [pH 8.0], 100 mM NaCl, 1 mM EDTA and 10% (v/v) glycerol) at 4 ˚C ON. The purified recombinant proteins were stored at -80 ˚C.

Chapter 3) E2F1 is citrullinated *in vitro* and in cells

3.1 Introduction

Previous work in our research group had shown E2F1 to undergo methylation by the arginine methyl transferases PRMT1 and PRMT5 on arginine residues R109 and R111/R113 respectively [93, 94]. Crucially, these methylation events were found to have distinctly different consequences for E2F1 activity and properties. Given that arginine residues can be subject to citrullination, and that citrullination impedes subsequent methylation of an arginine [116], we therefore asked whether E2F1 can become citrullinated. Of the peptidyl arginine deiminase family, PAD4 has a classical NLS motif, and is predominantly nuclear [114]. Furthermore, there have been a surge of studies implicating PAD4 in cancer and regulation of transcription factors. As such, we investigated PAD4 as the putative enzyme capable of targeting E2F1 for citrullination.

3.2 E2F1 is citrullinated *in vitro*

To address the question of whether E2F1 can become citrullinated, 22-mer E2F1 peptides, with the four arginine residues R109, R111, R113 and R127 covered in the sequence (Fig 3.1(a)) were sent to our collaborator (Prof. Paul Thompson, Scripps Institute, USA), who tested the *in vitro* citrullination of this peptide. Recombinant PAD4 was incubated with the peptides, and it was shown via MALDI-TOF (Fig 3.1(c)) and MALDI-TOF-TOF (Fig 3.2(a)), that the peptides can be effectively citrullinated *in vitro*. In fact, the enzyme kinetic parameters measured under these conditions revealed that PAD4 is almost 1000-fold more effective at citrullinating the E2F1 peptide than the peptide-like positive control BAEE (Nα- benzoyl-L-arginine ethyl ester) (the $k_{cat}/K_{m,}$ which is used as a measure of catalytic efficiency, is $4.4x10^3$ $M^{-1}.sec^{-1}$ with BAEE as the substrate, and $1.1x10^6$ $M^{-1}.sec^{-1}$ with E2F1 as the substrate) (Fig 3.1(b)).

Encouraged by these findings we therefore asked whether full-length E2F1 could also become citrullinated *in vitro*. In line with this, recombinant GST-E2F1, expressed and purified from BL21 *E.coli* cells, was incubated with Flag-PAD4, overexpressed in and purified from HEK293T cells. Recombinant calf thymus histones (commercially sourced) were used as the positive control, and GST tag alone, purified from BL21 *E.coli* cells was used as the negative control. As seen in figure 3.3(a), GST-E2F1 could become effectively citrullinated by Flag-PAD4, in a calcium dependent manner (compare lane 7 and 8). Since the GST tag alone was not citrullinated (lane 5), we concluded that the citrullination signal detected for GST-E2F1 was likely to be predominantly from E2F1 itself rather than its GST tag. Furthermore, consistent with the literature, we could also observe auto-citrullination of Flag-PAD4 (second arrow from the top, lanes 4-7)[157] as well as citrullination of histones[129] (lane 6).

To confirm that the citrullination observed on E2F1 is indeed dependent on the activity of PAD4, Flag-PAD4^{WT} and its catalytically inactive mutant Flag-PAD4^{C645A} [122] were overexpressed and purified from HEK293T cells. These two enzymes were then incubated with GST-E2F1 in the presence of calcium. As expected, whilst incubation with the active PAD4 enzyme induced citrullination of GST-E2F1, the inactive PAD4 mutant failed to citrullinate GST-E2F1 and undergo auto-citrullination (Fig 3.3(b), compare lanes 6 and 7). We therefore concluded that full length E2F1 can become citrullinated *in vitro* in a PAD4 and calcium dependent manner.

(a) **E2F1-22: ^{107}PARGRGRHPGKVKSPGEKSRY^{128}**

(b) Kinetic parameters of PAD4 with peptide substrates

Peptide	K_m (mM)	k_{cat} (sec^{-1})	k_{cat}/K_m ($M^{-1} \cdot sec^{-1}$)
BAEE	1.4 ± 0.2	5.9 ± 3 x 10^{-1}	4.4 x 10^3
E2F1-22	0.007 ± 0.001	7.8 ± 3 x 10^{-1}	1.1 x 10^6

(c)

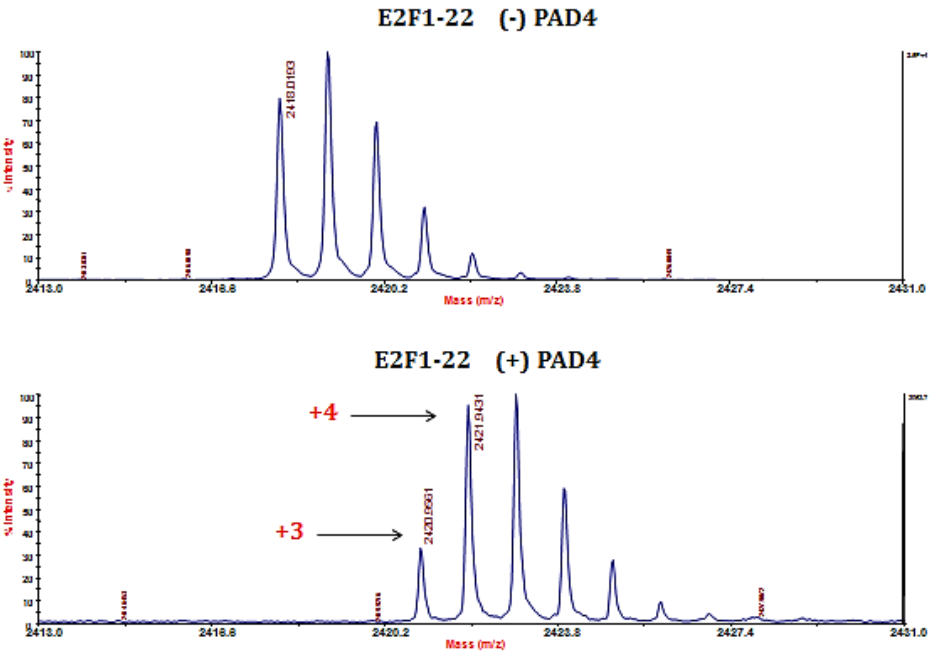


Figure 3.1) E2F1 peptides can be citrullinated *in vitro*

(a) The sequence of the E2F1 peptide used for *in vitro* citrullination assay. The arginine residues present in the sequence are highlighted red. (b) BAEE (Nα- benzoyl-L-arginine ethyl ester) or the E2F1 peptides were incubated *in vitro* with recombinant PAD4 and $CaCl_2$, and kinetic parameters were determined using a colorimetric assay. This assay is based on the specific reaction between a citrulline group and diacetyl monoxime (DAMO) and antipyrene, yielding a stable yellow complex, whose absorbance can be measured at 530 nm[158]. BAEE is a peptide like molecule that has been used as the positive control. K_m is the concentration of substrate at which the rate of catalysis is half the maximal value, k_{cat} is the turnover number of the enzyme, and k_{cat}/K_m reflects the catalytic efficiency of the enzyme[159]. (c) The un-citrullinated and citrullinated E2F1 peptides were analysed via MALDI-TOF. The shift in the peaks, highlighted by the arrows, illustrates the mass gain of the peptide following successful citrullination. (Note: These data were obtained by the laboratory of P.Thompson.)

(a)

	Observed		Predicted	MH+1 (av) 2419.7748		MH+1 (mono) 2418.3338	Predicted	Observed		
Mass Shift	(+) PAD4	(-) PAD4	b				y	(-) PAD4	(+) PAD4	Mass Shift
	140.1	140.1	140.0706	1	P	22	2376.3232		2379.5	+3
	211.1	211.1	211.1077	2	A	21	2279.2704			
+1	368.0		367.2088	3	R	20	2208.2333	2208.2		
+1	425.2		424.2303	4	G	19	2052.1322	2052.1		
		580.0	580.3314	5	R	18	1995.1107	1995.1		
+1	638.1	637.2	637.3529	6	G	17	1839.0096	1839.0		
+2	795.4	793.4	793.4540	7	R	16	1781.9882	1781.9		
+2	932.4	930.5	930.5129	8	H	15	1625.8870	1625.8	1626.7	+1
		1027.3	1027.5657	9	P	14	1488.8281	1488.8	1489.9	+1
		1084.6	1084.5871	10	G	13	1391.7754			
		1212.6	1212.6821	11	K	12	1334.7539	1334.7		
		1269.7	1269.7035	12	G	11	1206.6589			
		1368.6	1368.7720	13	V	10	1149.6375			
			1496.8669	14	K	9	1050.5691			
			1583.8990	15	S	8	922.4741			
			1680.9517	16	P	7	835.4421			
			1737.9732	17	G	6	738.3893			
		1867.0	1867.0518	18	E	5	681.3678			
		1995.1	1995.1107	19	K	4	552.3253			
			2082.1428	20	S	3	424.2303			
			2238.2439	21	R	2	337.1893			
			---	22	Y	1	181.0972			

- Observed Mass Shift - Citrullinated Arginine - Can Also Be Citrullinated

Figure 3.2) E2F1 peptides can be citrullinated *in vitro*

(a) E2F1 peptides incubated with and without recombinant PAD4 were analysed with MALDI TOF-TOF. The mass shift following incubation with PAD4 is used to determine the residues which have been citrullinated. Following conversion of an arginine to a citrulline, there is a gain of mass of 1 Da. The arginine residues highlighted in yellow, i.e. R109, R113 and R127, can be said with certainty to undergo citrullination in this assay. (Note: These data were obtained by the laboratory of P.Thompson.)

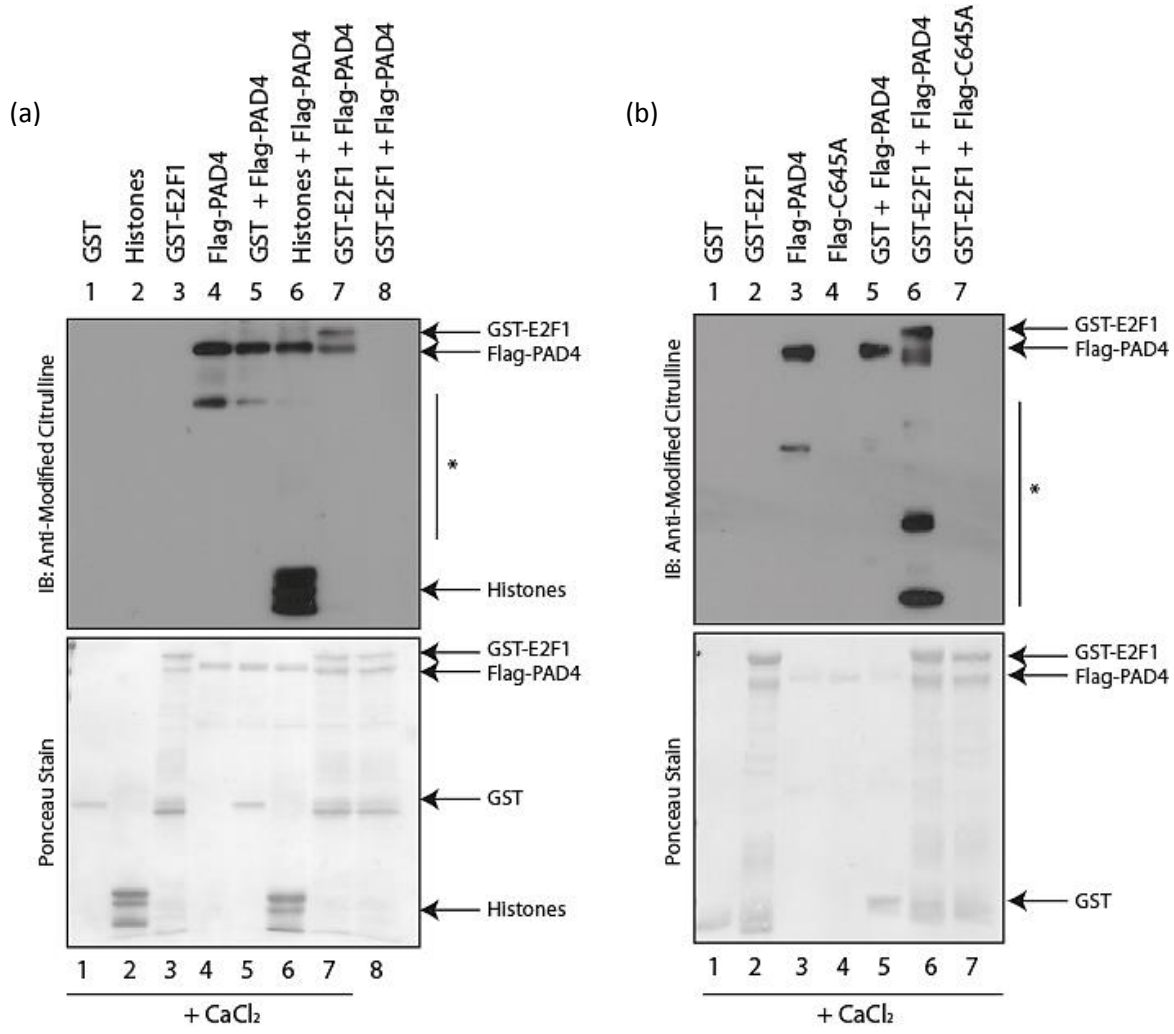


Figure 3.3) Full length recombinant E2F1 can be citrullinated *in vitro*

(a) 1 µg GST-E2F1 (purified from BL21 *E.coli* cells), 1 µg Flag-PAD4 (purified from HEK293T cells), 1 µg calf thymus histones (commercially sourced), and 1 µg GST tag alone (purified from BL21 *E.coli* cells) were incubated *in vitro*, at 37 °C in the presence of 4 mM $CaCl_2$ and 2 mM DTT (except for lane 8 which is without added $CaCl_2$). Ponceau staining of the blot illustrates equal loading of enzyme and substrates. Citrullination was detected using "Millipore Anti-Modified Citrulline kit". (n=3) (b) The assay was set up as described above. Flag-C645A is the catalytically dead PAD4 mutant, which was also purified from HEK293T cell. (n=1) * Degradation products

3.3 Mapping the sites of citrullination *in vitro*

Having established that E2F1 can be citrullinated *in vitro*, we sought to define the residues that can be targeted, and establish, whether as hypothesised, the arginine residues that have been shown to be methylated, could also be a target for citrullination. To address this question, recombinant GST-E2F1 truncation mutants, GST-ΔC E2F1 (residues 1-194), and GST-ΔN E2F1 (residues 194-437) were compared. As seen from figure 3.4(a), both domains of E2F1 were citrullinated by GST-PAD4, though GST-ΔC E2F1 (1-194), exhibited stronger citrullination than GST-ΔN E2F1 (194-437) (compare lanes 7 and 8). Nonetheless, this result suggested that other residues of E2F1, besides the arginine residues originally hypothesised, were likely to become citrullinated *in vitro*. Indeed when we compared GST-E2F1^{WT} with GST-E2F1^{KKK}, where the three arginine residues R109, R111 and R113 were mutated to lysine, we found that both constructs underwent citrullination by Flag-PAD4 (Fig 3.4(b), compare lanes 6 and 7). We therefore concluded that multiple arginine residues, both within the N-terminal and C-terminal domains of E2F1 are likely to become citrullinated *in vitro*.

In an attempt to narrow down the sites modified on E2F1, we analysed GST-E2F1, citrullinated *in vitro*, via mass spectrometry. Figure 3.5(a) is the Coomasie stained gel from which GST-E2F1 was excised, digested with Trypsin, and prepared for mass spectrometry analysis, carried out by collaborators. Figure 3.5(b) illustrates that E2F1 was successfully citrullinated in this experiment. Given that recombinant protein was digested and submitted for mass spectrometry, very high sequence coverage of E2F1 was reported back (81% of the sequence covered) (Fig 3.5(c)). Furthermore from this mass spectrometry analysis, it was found that E2F1 was citrullinated on at least 9 residues *in vitro* (personal communication). These findings suggest that *in vitro* PAD4 exhibits very

little sequence specificity, which would be consistent with previous reports[160]. It has been suggested that all PAD isotypes can citrullinate most accessible arginine residues in vitro, although certain amino acids flanking the arginine residue may influence its susceptibility to citrullination[161]. The promiscuity of PAD4 *in vitro*, may not however necessarily reflect the *in vivo* behaviour of this enzyme, particularly as in cells the calcium availability is more limited than that supplied *in vitro* and the enzyme and substrate interaction kinetics are likely very different. In a recent study it has been shown that PAD4 shows preference for the sequence it citrullinates[162], and similar selectivity could also hold valid for E2F1 in the context of the cell. From these results so far we conclude that E2F1 can be citrullinated *in vitro* on multiple arginine residues, reflective of the promiscuous enzymatic activity of PAD4.

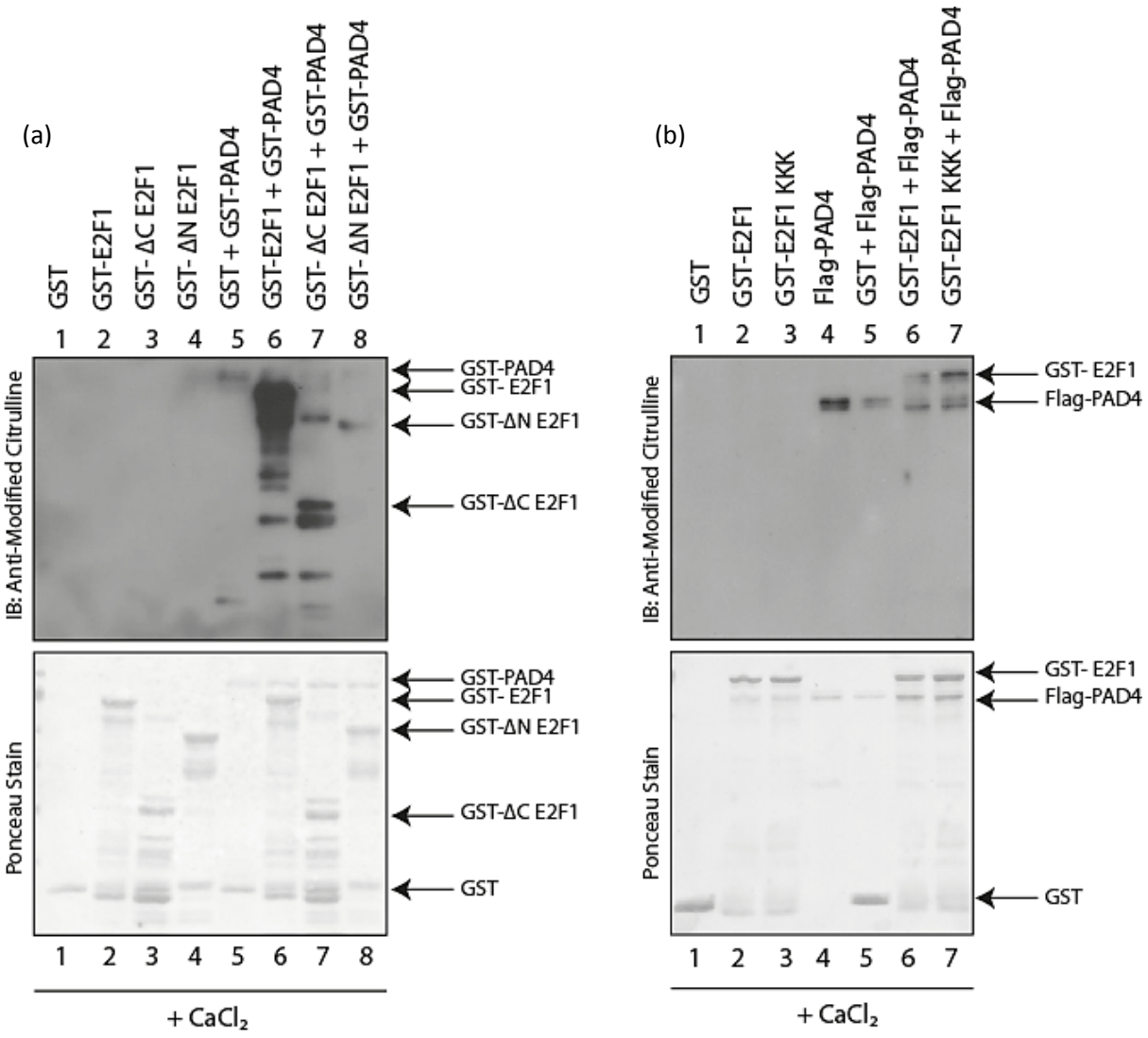
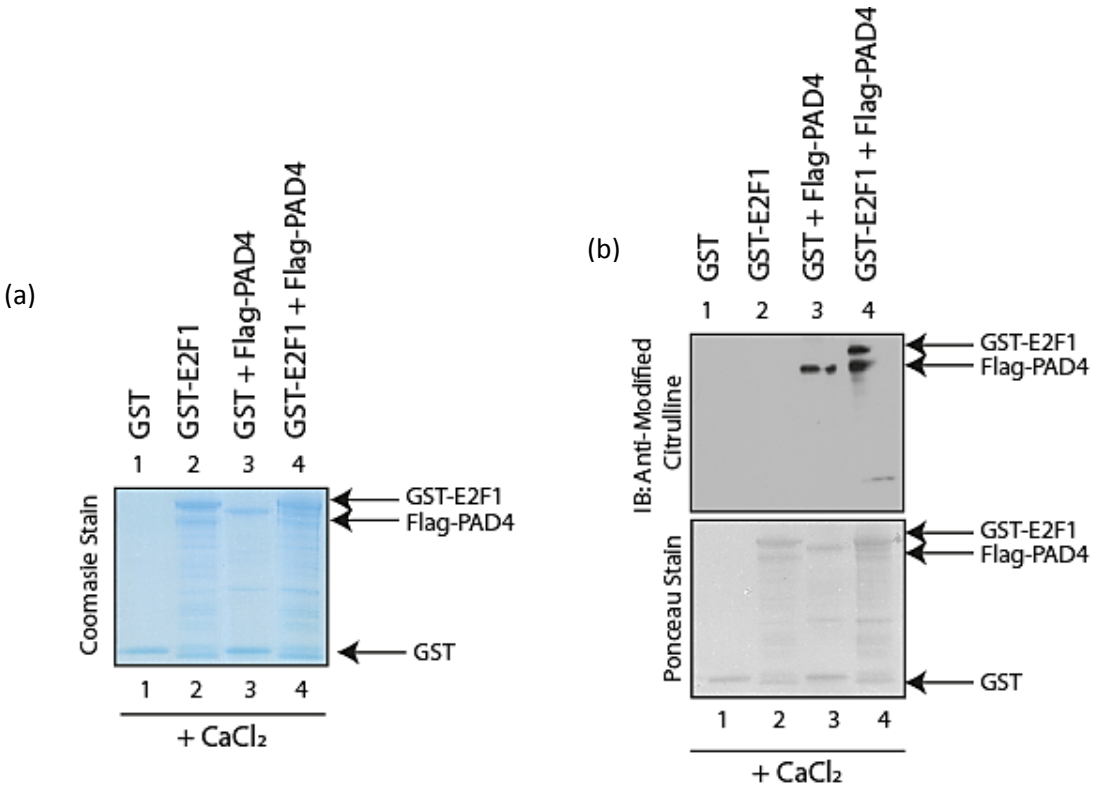


Figure 3.4) Citrullination of E2F1 mutants *in vitro*

(a) 1 µg full length GST-E2F1, GST-ΔC E2F1 (1-194), GST-ΔN E2F1 (194-437) or GST tag alone were incubated with 200 ng GST-PAD4 at 37 °C, in the presence of 4 mM $CaCl_2$ and 2 mM DTT. Ponceau staining of the blot illustrated equally loading of enzyme and substrates. Citrullination was detected using "Millipore Anti-Modified Citrulline kit". (n=2) (b) 1 µg GST-$E2F1^{WT}$, 1 µg GST-$E2F1^{KKK}$ (R109K/R111K/R113K) or GST tag alone were incubated with 1 µg Flag-PAD4 at 37 °C, in the presence of 4 mM $CaCl_2$. Ponceau staining and AMC kit were used as before. (n=2)



(c)

Protein sequence coverage: 81%

Matched peptides shown in ***bold red***.

```
  1 MALAGAPAGG PCAPALEALL GAGALRLLDS SQIVIISAAQ DASAPPAPTG
 51 PAAPAAGPCD PDLLLFATPQ APRPTPSAPR PALGRPPVKR RLDLETDHQY
101 LAESSGPARG RGRHPGKGVK SPGEKSRYET SLNLTTKRFL ELLSHSADGV
151 VDLNWAAEVL KVQKRRIYDI TNVLEGIQLI AKKSKNHIQW LGSHTTVGVG
201 GRLEGLTQDL RQLQESEQQL DHLMNICTTQ LRLLSEDTDS QRLAYVTCQD
251 LRSIADPAEQ MVMVIKAPPE TQLQAVDSSE NFQISLKSKQ GPIDVFLCPE
301 ETVGGISPGK TPSQEVTSEE ENRATDSATI VSPPPSSPPS SLTTDPSQSL
351 LSLEQEPLLS RMGSLRAPVD EDRLSPLVAA DSLLEHVRED FSGLLPEEFI
401 SLSPPHEALD YHFGLEEGEG IRDLFDCDFG DLTPLDF
```

Citrullinated residues detected: R91, R109, R127, R166, R232, R252, R361, R366, R373

Figure 3.5) Analysing *in vitro* citrullinated E2F1 using mass spectrometry

(a) 1 µg GST-E2F1 or GST tag alone were incubated with 1 µg Flag-PAD4 at 37 °C in the presence of 4 mM $CaCl_2$. A fraction of the sample was loaded on a Bis-Tris gel and Coomasie stained. The bands corresponding to GST-E2F1 and GST tag alone (from lanes 3 and 4) were excised, and processed and submitted for mass spectrometry. (b) A fraction of the above sample was loaded on a polyacrylamide gel; loading was measured using Ponceau staining and citrullination using AMC kit. (c) Sequence coverage reported for citrullinated GST-E2F1 analysed by mass spectrometry (ran and analysed by the Laboratory of Professor Benedikt Kessler), and the residues identified to be citrullinated on GST-E2F1.

3.4 E2F1 is citrullinated in cells

Given our findings *in vitro*, we asked whether E2F1 could also become citrullinated in cells. To address this, HEK293T cells were transfected with HA-PAD4 and Flag-E2F1 and stimulated with 5 μM of the calcium ionophore A23187 for 45 min (reported to increase intracellular calcium levels, and hence augment PAD activity)[129]. Flag immunoprecipitation of the lysates and immunoblotting using the Millipore AMC kit demonstrated that ectopic E2F1 could be citrullinated in these cells (Fig 3.6(a)). Similarly when U2OS cells were transfected with HA-PAD4 and lysed in the presence of 4 mM $CaCl_2$, we observed citrullination of endogenous E2F1 immunoprecipitated from cells (Fig 3.6(b)).

In order to identify the domain of E2F1 citrullinated by PAD4, U2OS cells were transfected with Flag-ΔC E2F1 (1-194) or Flag- ΔN E2F1 (194-437) and HA-PAD4 and treated with A23187. Flag-immunoprecipitation and immunoblotting for citrullination suggested that the N-terminus of E2F1 (i.e. residues 1-194) is more strongly citrullinated than the C-terminus (i.e. residues 194-437) (Fig 3.6(c)). We therefore analysed this E2F1 mutant via mass spectrometry to map the sites of citrullination.

3.5 Mapping the sites of citrullination in cells

U2OS cells were transfected with HA-PAD4 and Flag- ΔC E2F1 (1-194), treated with A23187 and immunoprecipitated with Flag antibody. Figure 3.7(a) is the Coomasie stained gel from which the band corresponding to E2F1 was excised and submitted for mass spectrometry analysis. As can be seen from figure 3.7(b) the sequence coverage reported back for E2F1 was only partial, but nonetheless two arginine residues, namely

R109 (Fig 3.7(c)) and R127 (Fig 3.7(d)) were found to be citrullinated in the E2F1 sequence covered.

To confirm these findings, site-directed mutagenesis was used to generate a E2F1 point mutant with an arginine to lysine substitution. Changing the arginine residue to lysine is a conservative change that maintains the positive charge at this position but prevents citrullination site. Comparing Flag-E2F1^{WT} (1-194) with Flag-E2F1^{R127K} (1-194), where the arginine residue in position 127 was changed to a lysine, illustrated that both forms could become effectively citrullinated in cells (Fig 3.8(a)). We therefore mutated the other arginine residues in that vicinity, including R109 as well as R111 and R113 to investigate whether this would abrogate citrullination. Indeed we observed lower citrullination levels on Flag-E2F1^{KK} (1-194) (R109K/R127K) and Flag-E2F1^{KKKK} (1-194) (R109K/R111K/R113K/R127K) compared to Flag-E2F1^{WT} (1-194) (Fig 3.8(b)). We therefore concluded that E2F1 can become citrullinated in cells upon ectopic expression of PAD4 and that the arginine residues R109, R111, R113 and R127 likely represent the main sites targeted by PAD4.

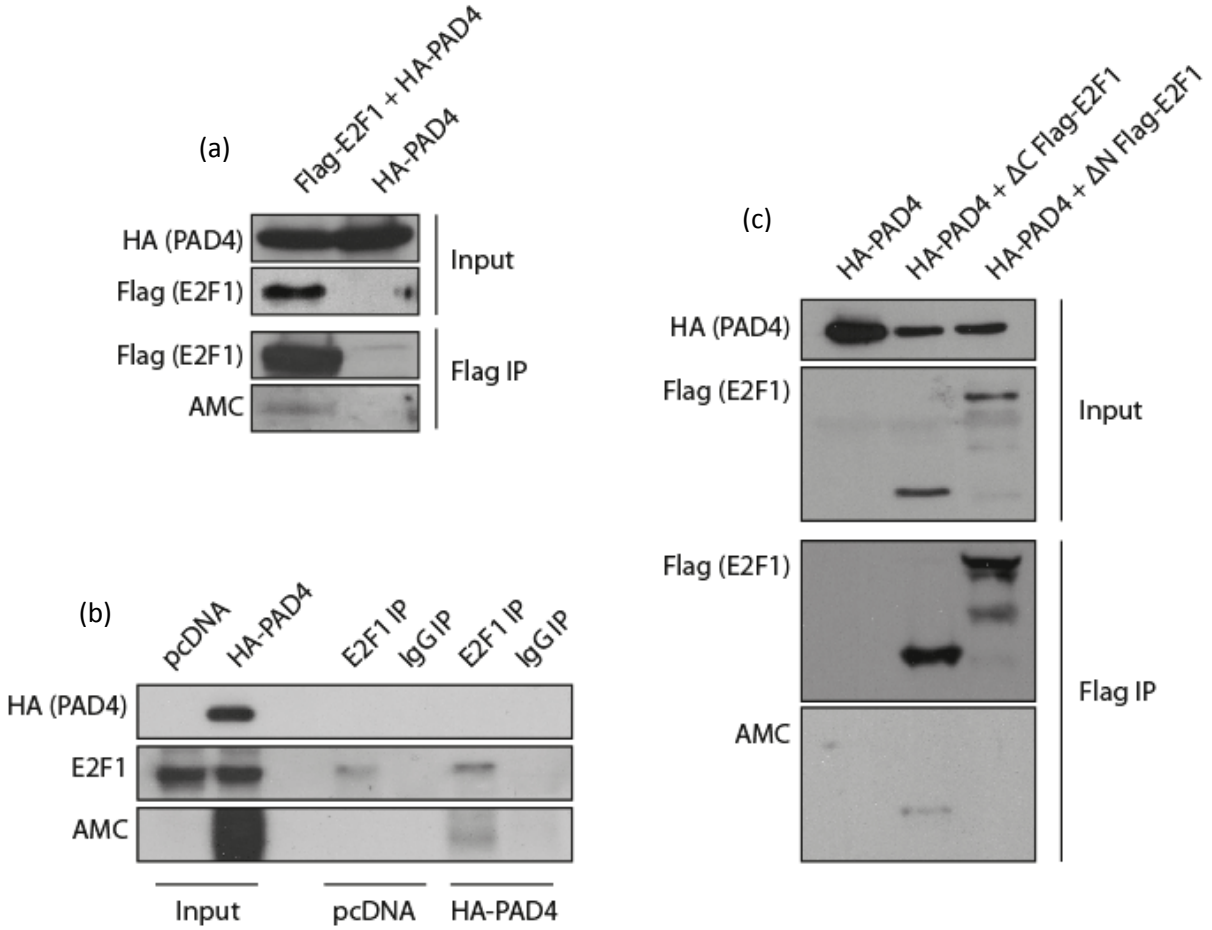


Figure 3.6) E2F1 is citrullinated in cells

(a) HEK293T transfected with 1 µg Flag-E2F1 and 2 µg HA-PAD4 were treated with 5 µM A23187 for 45 min and immunoprecipitated using Flag-agarose beads. (n=2) (b) U2OS cells transfected with 2 µg pcDNA or HA-PAD4 were treated with 4 mM $CaCl_2$ and 2mM DTT and immunoprecipitated using E2F1 (C20) or rabbit IgG control antibodies. (n=3) (c) U2OS cells transfected with 1 µg Flag-ΔC E2F1 (1-194) or 1 µg Flag-ΔN E2F1 and 2 µg HA-PAD4 were treated with 5 µM A23187 for 45 min and immunoprecipitated using Flag-agarose beads. AMC kit was used to detect citrullination. (n=3)

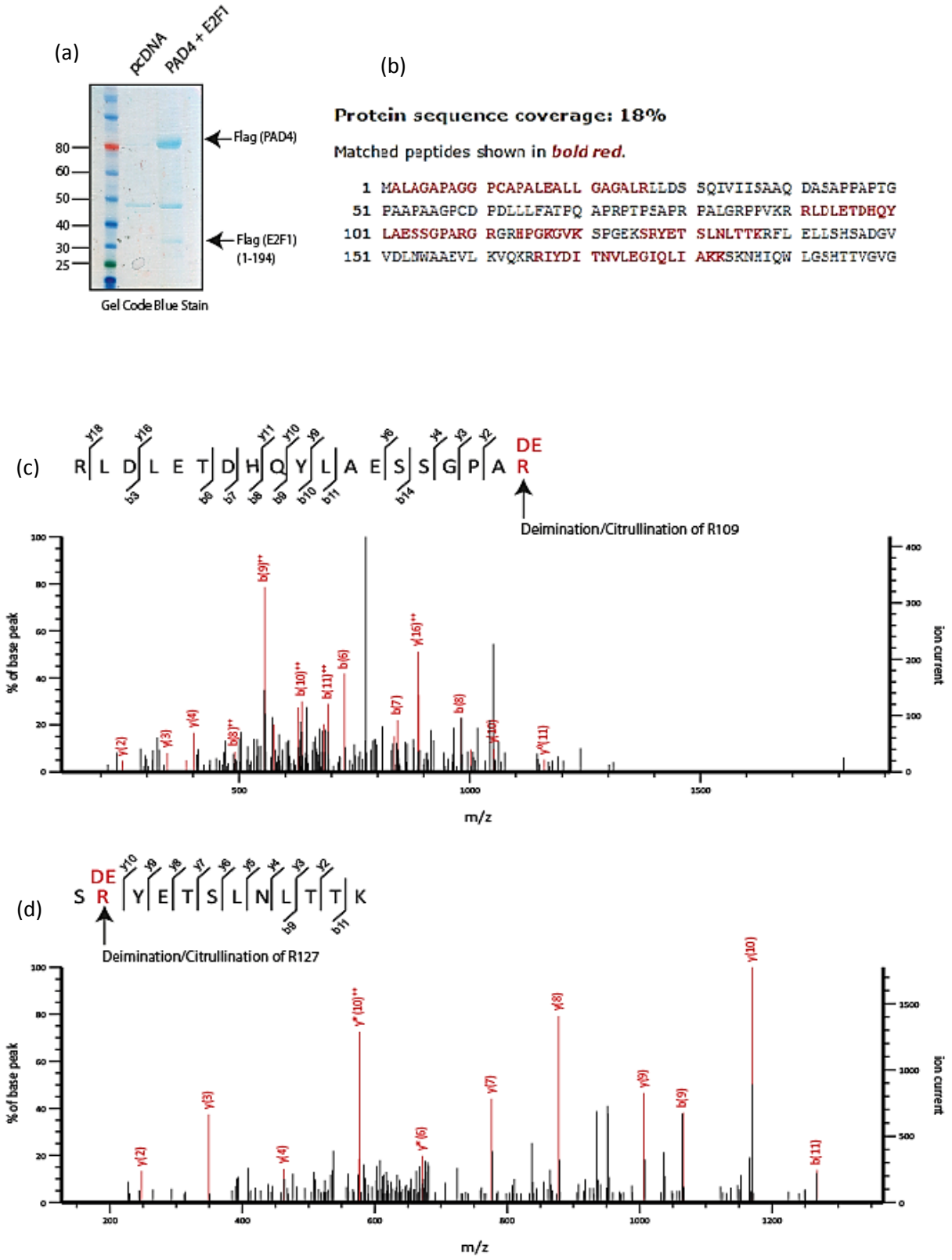


Figure 3.7) Analysing in cell citrullinated E2F1 using Mass spectrometry

(a) U2OS cells transfected with 1 µg Flag-ΔC E2F1 (1-194) and 2 µ HA-PAD4 were treated with 5 µM A23187 for 45 min, immunoprecipitated using Flag-agarose beads and eluted using Flag-peptides. The samples were loaded on a Bis-Tris gel and Coomasie stained. The band

corresponding to Flag-ΔC E2F1 (1-194) was excised, processed and submitted for analysis by LC-MS/MS mass spectrometry. (b) The sequence coverage reported for Flag-ΔC E2F1 (1-194) analysed by mass spectrometry. Note that the reported percentage sequence coverage is in relation to the full sequence of E2F1, and not the truncated N-terminal domain which was submitted for analysis. (c) Tandem mass spectra profiles illustrating citrullination of R109 (top spectra) and R127 (bottom spectra). Fragment ions are indicated as b and y ions. A mass shift of 1 Da was indicative of citrullination of an arginine residue. Mass spectrometry ran and analysed by the Laboratory of Professor Benedikt Kessler.

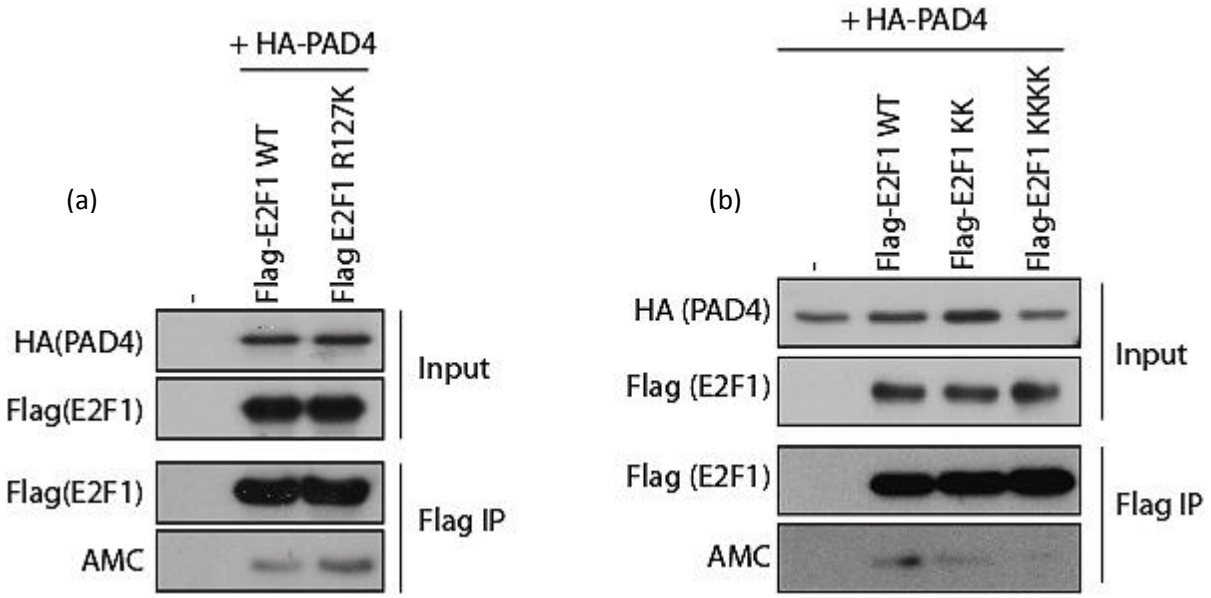


Figure 3.8) Citrullination of E2F1 mutants in cells

(a) U2OS cells transfected with 1 µg Flag-E2F1^{WT} (1-194) or 1 µg Flag-E2F1^{R127K} (1-194) and 2 µg HA-PAD4 were treated with 5 µM A23187 for 45 min and immunoprecipitated using Flag-agarose beads. (b) U2OS cells transfected with 1 µg Flag-E2F1^{WT} (1-194), 1 µg Flag-E2F1^{KK} (R109K/R127K) (1-194) or Flag-E2F1^{KKKK} (R109K/R111K/R113K/R127K) (1-194) and 2 µg HA-PAD4 were treated with 5 µM A23187 for 45 min and immunoprecipitated using Flag-agarose beads. AMC kit was used to detect citrullination. (n=2)

3.6 Tet-On inducible Flag-PAD4.pTRE cell line

With the observations that E2F1 can become citrullinated by PAD4, we took advantage of a Tet-On gene expression system, to investigate the regulation of E2F1 by PAD4. Such a system enables one to effectively induce the expression of a gene of interest (such as PAD4 in this instance) by the addition of doxycycline to the culture medium, in a fast, reproducible and dose dependent manner[163-165]. It therefore offers a valuable tool for studying the function and downstream effects of a protein on interest.

The Tet-On system consists of two elements, a regulatory plasmid and a response plasmid. The regulatory plasmid, which is under the control of a CMV (cytomegalovirus) promoter, encodes a 37 kDa hybrid protein known as rtTA (reverse tetracycline controlled trans-activator). rtTA is derived from the fusion of the N-terminus (1-207) of the *E.coli* protein TetR (Tet repressor protein) with the Herpes Simplex virus VP16 activation domain. Four amino acid substitutions in the N-terminal TetR domain of this fusion protein then yields a doxycycline responsive transcription factor, which binds its target sequence in the presence of doxycycline only. The second element, i.e. the response plasmid (also known as the TRE plasmid, or pTRE), expresses the gene of interest, under the control of a tetracycline response element (TRE) and a minimal CMV promoter. In the presence of doxycycline, the rtTA protein binds to the TRE, to induce expression of the gene of interest.

In addition, the regulatory plasmid encodes a neomycin selection gene, and the pTRE encodes a hygromycin selection gene. Therefore to make the Tet-On cell line a double-stable cell line expressing both regulatory and response plasmid is required. Parental U2OS Tet-On cell lines, stably expressing the regulatory plasmid, were used to create the Flag-PAD4.pTRE inducible cell line.

Flag-PAD4 was cloned into pTRE vector (Fig 3.9(a)), and transfected into U2OS Tet-On cell lines. Selection with both G418 and hygromycin ensured that only colonies with both the regulatory plasmid and Flag-PAD4.pTRE can grow. Figure 3.9(b) illustrates that the selected cell lines, could be successfully induced with doxycycline to express Flag-PAD4. Furthermore as expected the Flag-PAD4 induced in these cells was mainly localised to the nucleus (Fig 3.9(c)). These cell lines were therefore used for subsequent functional studies.

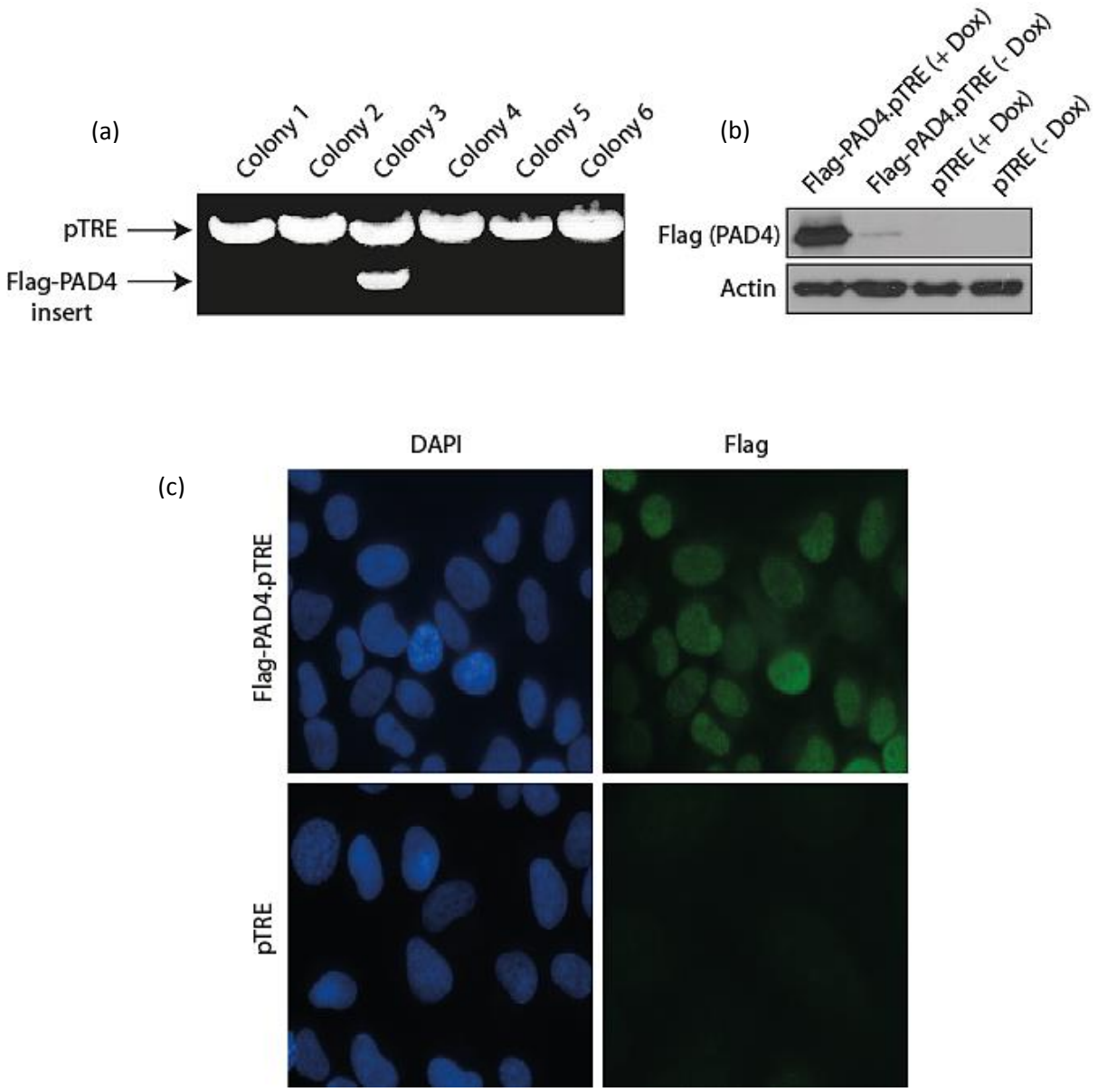


Figure 3.9) Generating Tet-On inducible Flag-PAD4.pTRE cell lines

(a) Flag-PAD4 was cloned into pTRE vector, transformed into DH5α *E.coli* cells and spread onto ampicillin containing agar plates to allow for successfully transformed bacterial cells to be selected for. Several colonies were picked and cultured and the expressed plasmid was purified. The plasmids were digested and analysed on ethidium bromide stained agarose gels to pick the colony with successful insertion of Flag-PAD4 into the pTRE vector (colony 3 in this instance). (b) Parental Tet-On U2OS cells were transfected with 5 µg Flag-PAD4.pTRE or empty pTRE plasmid and were subjected to double selection with G418 and hygromycin. Flag-PAD4.pTRE and pTRE U2OS Tet-On cells were cultured in the presence (+ Dox) or absence (- Dox) of 1 µg/ml doxycycline for 36 h and immuno-blotted using Flag antibody. Treatment of Flag-PAD4.pTRE cell line with doxycycline illustrates strong expression of Flag-PAD4. (c) Flag-PAD4.pTRE and pTRE U2OS Tet-On cells were seeded on cover slips and treated with 1 µg/ml doxycycline for 36 h. Flag antibody (1:200 dilution) was used for the primary immunofluorescence staining. The cells were stained with DAPI to visualise the nuclei.

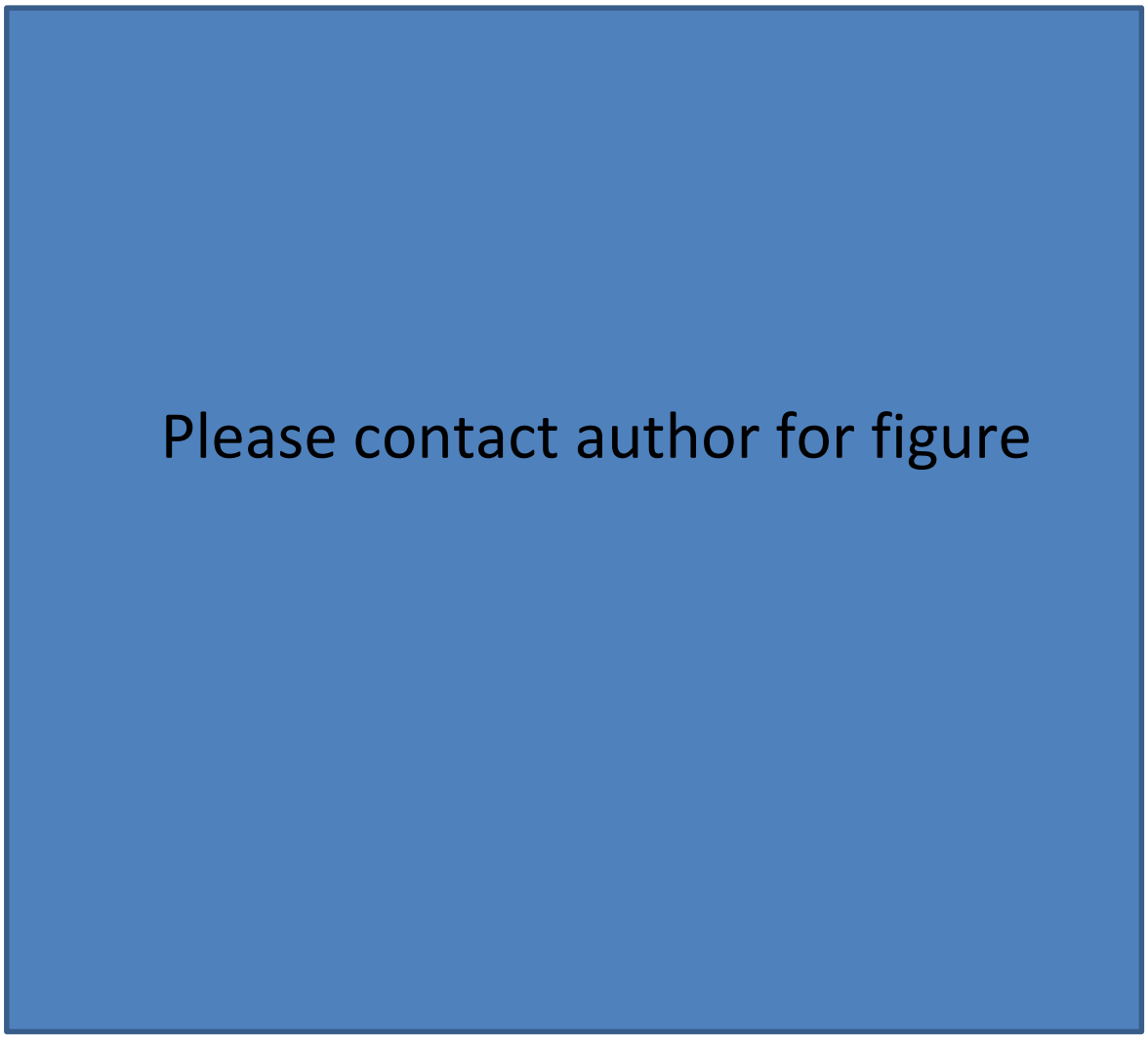
3.7 Chapter Summary

In this chapter we reported that E2F1 can be citrullinated *in vitro* and in cells in a PAD4 and calcium dependent manner. Our attempt to map the main residues targeted by PAD4, implied that multiple residues are likely modified *in vitro*, though in cells four arginine residues, which lie in close proximity, emerged as the main target sites. Interestingly, three of these residues, namely R109, R111 and R113, are the sites which we had previously reported to be methylated in cells.

Although PAD4 has been frequently reported to show little sequence specificity particularly *in vitro*, there are examples in the literature to show that it can target certain substrates at more defined residues in cells, such as ING4[166] and GSK3B[150]. Furthermore, in a recent study comparing human PAD2 and PAD4 for their target sequence preference, it was found that the two enzymes show different specificities, with PAD4 the more selective of the two[162]. Analysis of citrullinated proteins in cells, and an enrichment of the residues relative to the target arginine, enabled the authors to map the 'preferred' versus 'disfavoured' residues. Residues which were overrepresented by at least two fold were considered 'preferred', and those which were underrepresented by at least two fold, were considered 'disfavoured' (Fig 3.10(a)).

Although not absolute, comparing the results from this study with ours, illustrates that the E2F1 residues which we found to be citrullinated in cells are generally flanked by residues which are considered to be 'preferred' (Fig 3.10(b)). It is therefore conceivable, that despite the limitations of our approach to exactly and unequivocally map the sites of citrullination on E2F1, under more physiological conditions E2F1 could be citrullinated in a more defined and controlled manner and the residues we have identified could be the main targets.

On the basis that PAD4 targets E2F1 for citrullination, we generated a Tet-On inducible Flag-PAD4.pTRE cell line, alongside its control empty vector pTRE cell line. The aim was to utilize this system to achieve strong and reproducible expression of PAD4 and study its downstream effect on E2F1 as explored in the following chapter.



(b)

R109 R111 R113 R127

SGPARGRGRHPGKGVKSPGEKSRYETS

Figure 3.10) Schematic diagram depicting a proposed target sequence for PAD4

(a) Based on in cell citrullinated proteins, residues flanking the target arginine site were analysed and given an enrichment score. Residues which were overrepresented were considered preferred, whilst residues that were underrepresented were considered disfavoured. The Diagram adopted from [162]. (b) The E2F1 sequence showing the relative positions of R109, R111, R113 and R127. The arrows indicate residues whose relative positions can be considered favourable or preferred based on the diagram depicted above.

Chapter 4) Regulation of E2F1 by PAD4

4.1 Introduction

Citrullination has been known to have important consequences for the properties and functional behaviour of a given protein. Although the gain of mass is very small and only of the magnitude of 1 Da relative to the unmodified arginine, the biochemical changes associated with citrullination are significant, as the positively charged arginine is neutralised by citrullination. Therefore one can expect changes in protein-protein interactions, DNA binding, protein folding, stability and even sub-cellular localization of the target protein[111]. With the observations that E2F1 can become citrullinated by PAD4, we set out to investigate the consequences this entails for the properties and cellular activities of E2F1.

4.2 E2F1 and PAD4 interact in cells

As E2F1 is a substrate for PAD4, it was also likely that the two proteins would interact in cells, though one may not necessarily expect to detect a stable complex, as the interaction may be a transient enzyme-substrate association. Nonetheless, when Flag-PAD4.pTRE or the empty vector pTRE control cell lines were transfected with HA-E2F1, and immunoprecipitated using Flag antibody, a specific interaction could be observed between E2F1 and PAD4 (Fig 4.1(a)). Furthermore, in MCF7 cells transfected with HA-PAD4, we could detect an interaction between endogenous E2F1 and the ectopically transfected PAD4 (Fig 4.1(b)).

To map the region of interaction, U2OS cells were transfected with Flag-tagged E2F1 truncation mutants and HA-PAD4, and immunoprecipitated using Flag antibody. The region of E2F1 which could interact with ectopically expressed HA-PAD4 in cells

was found to lie within the N-terminus of E2F1. This was since Flag-ΔC E2F1 (1-194) (Fig 4.1(c)) and the shorter mutant Flag-ΔC E2F1 (1-110) (Fig 4.1(d)) could both pull down HA-PAD4 from cells, whilst the C-terminal mutants, Flag-ΔN E2F1 (194-437) and Flag-ΔN E2F1 (110-437), both failed to do so. Thus we conclude that E2F1 can interact with PAD4 in cells, largely through its N-terminal domain.

4.3 PAD4 augments E2F1 transcriptional activity

A very effective and quantitative assay to assess the activity of a transcription factor such as E2F1 is the luciferase reporter assay. In this system, the regulatory region of the 'gene of interest' is cloned upstream of the luciferase gene in an appropriate expression vector. In this instance, luciferase reporter constructs were used, whose upstream promoter elements were derived from known E2F1 target genes, such as CDC6, p73 and DHFR. As such, the relative expression of the luciferase genes was used as a measure of the transcriptional activity of E2F1.

As can be seen from figure 4.2(a), ectopic HA-E2F1 transfected into U2OS cells could induce the expression of the CDC6 and p73 luciferase genes by approximately 20 and 40 fold respectively. Crucially however, co-transfection of Flag-PAD4 could further increase E2F1 activity by at least 50%. This was whilst Flag-PAD4 transfected in the absence of HA-E2F1 failed to stimulate the expression of the luciferase gene, suggesting that the ability of PAD4 to augment transcription in this context was dependent on the presence of E2F1.

In order to understand whether the effect observed is dependent on the catalytic activity of PAD4, U2OS cells were treated with the specific PAD4 inhibitor TDFA, and its

inactive analogue TDHA[167]. Interestingly we observed E2F1 transcriptional activity to decrease in a dose-dependent manner with the active PAD4 inhibitor, implying that the activity of PAD4 is at least partially important for its ability to enhance transcription (Fig 4.2(b)). Further evidence for the importance of PAD4 catalytic activity in promoting E2F1 transcriptional activity came from comparing HA-E2F1^{WT} with the mutant HA-E2F1^{KKKK}, which we found to be less efficiently citrullinated in cells. Indeed whilst Flag-PAD4 could enhance the activity of HA-E2F1^{WT}, it was less effective at enhancing the activity of the mutant HA-E2F1^{KKKK} (Fig 4.2(c)). Although we have not investigated whether both forms of E2F1 could equally interact with and recruit PAD4, one potential implication of these results is that citrullination of E2F1 by PAD4 is at least part of the mechanism through which PAD4 regulates E2F1 activity. We therefore conclude that PAD4 can augment E2F1 transcriptional activity, and that citrullination of E2F1 is important for this effect.

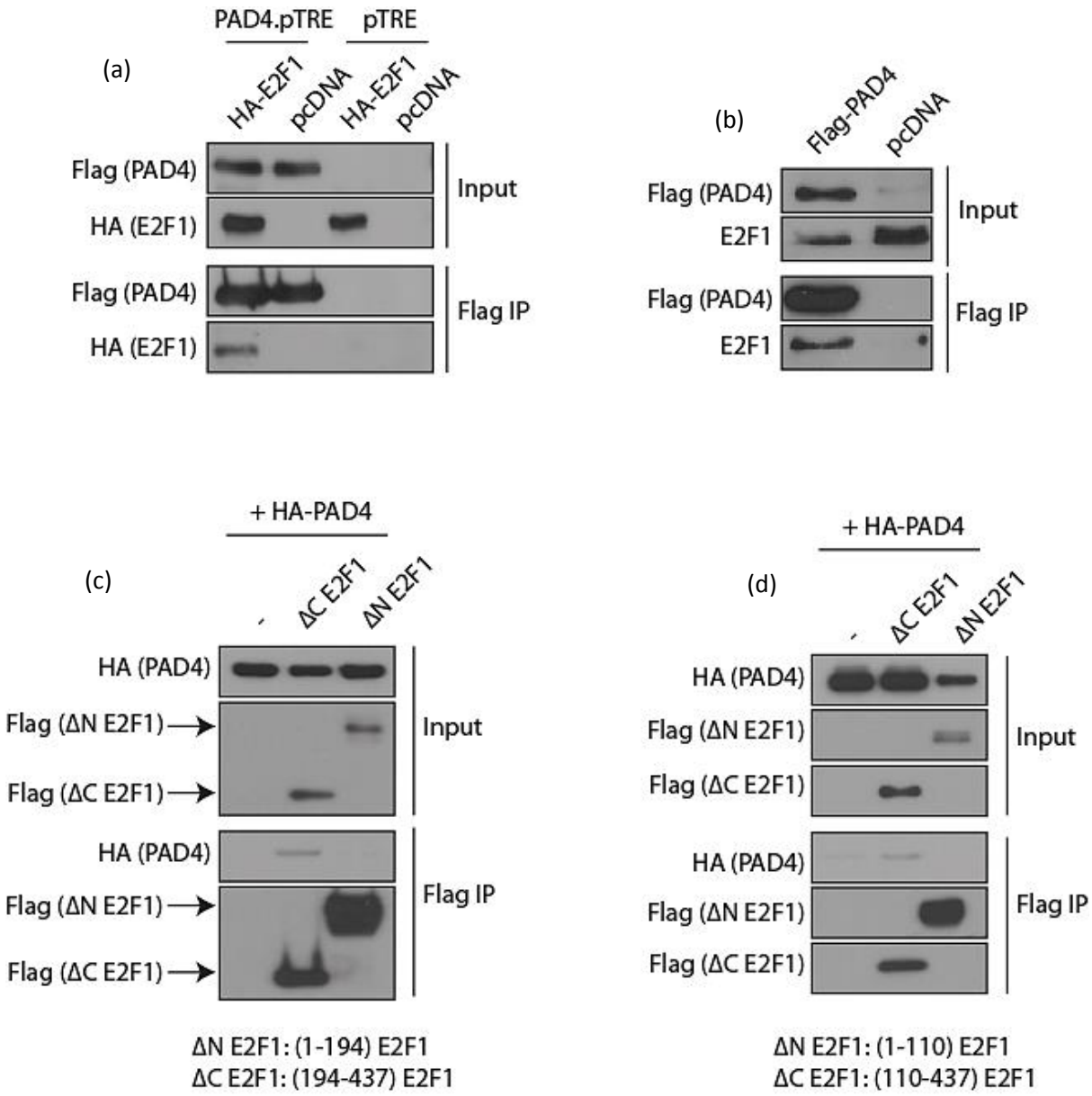


Figure 4.1) E2F1 and PAD4 interact in cells

(a) Flag-PAD4.pTRE and pTRE cell lines transfected with 1 µg HA-E2F1 or pcDNA were immunoprecipitated using Flag-agarose beads. (n=3) (b) MCF-7 cells transfected with 2 µg Flag-PAD4 were immunoprecipitated using Flag-agarose beads. (n=2) (c) U2OS cells transfected with 1 µg Flag-ΔC E2F1 (1-194) or Flag-ΔN E2F1 (194-437) and 2 µg HA-PAD4 were immunoprecipitated using Flag-agarose beads. (n=3) (d) U2OS cells transfected with 1 µg Flag-ΔC E2F1 (1-110) or Flag-ΔN E2F1 (110-437) and 2 µg HA-PAD4 were immunoprecipitated using Flag-agarose beads. (n=2)

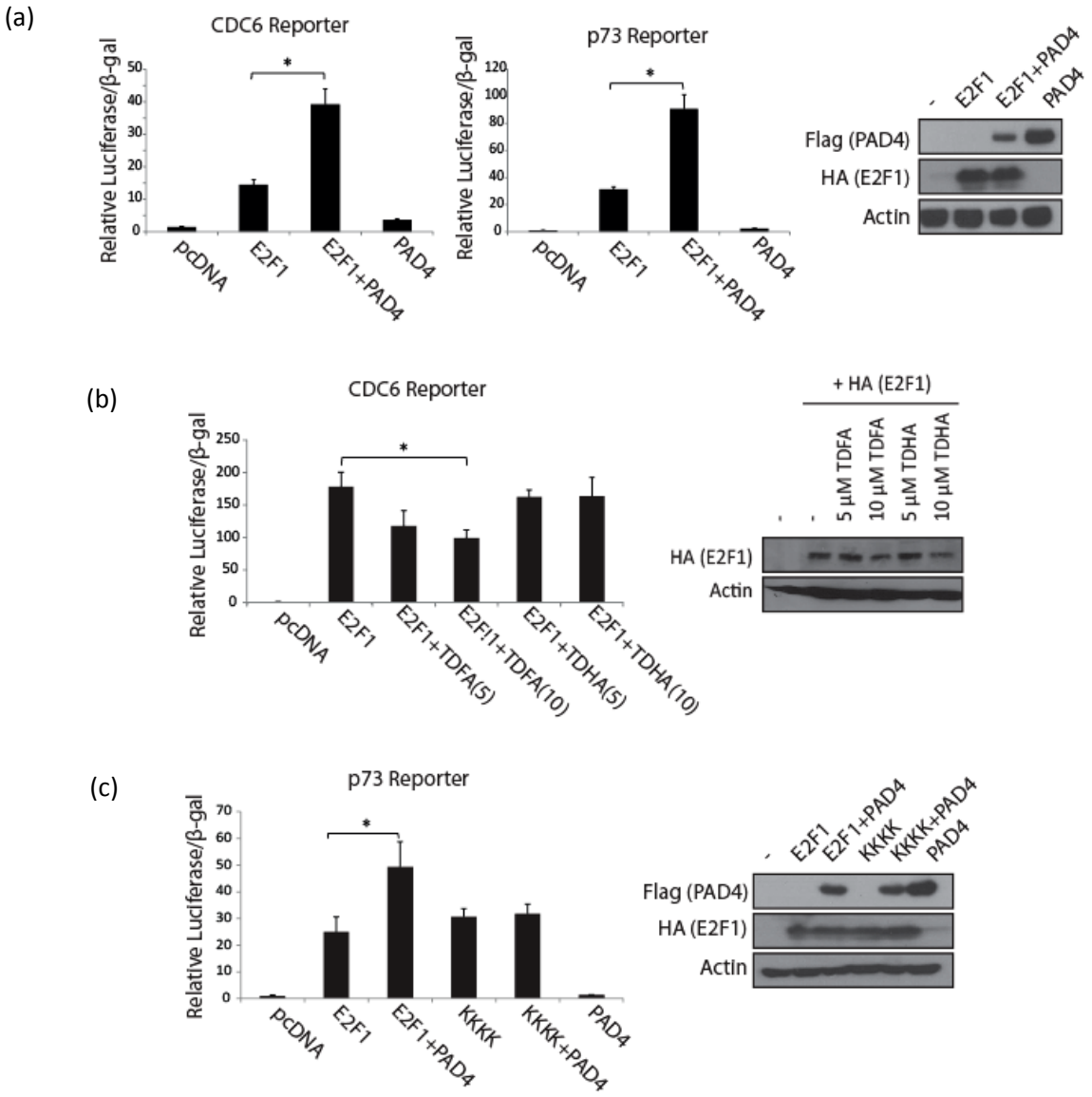


Figure 4.2) PAD4 augments E2F1 transcriptional activity

(a) U2OS cells were transfected with 100 ng HA-E2F1, 500 ng Flag-PAD4, 100 ng CDC6/p73 luciferase reporter plasmid and 150 ng β-gal for 36 h. (n=5) (b) U2OS cells were transfected with 100 ng HA-E2F1, 100 ng CDC6 luciferase reporter plasmid and 150 ng β-gal for 36 h and were treated with 5 or 10 μM TDFA or TDHA where indicated for 16 h. (n=2) (c) U2OS cells were transfected with 100 ng HA-E2F1 WT or KKKK mutant (R109K/R111K/R113K/R127K), 500 ng Flag-PAD4, 100 ng CDC6/p73 luciferase reporter plasmid and 150 ng β-gal for 36 h. In the above reporter assays Luciferase measurements were taken using a luminometer and were normalised to β-gal values. (n=2) ± S.D * $p < 0.05$

4.4 PAD4 increases E2F1 DNA binding

With the observation that PAD4 increases E2F1 transcriptional activity, we asked whether at the genomic level this correlated with enhanced E2F1 DNA binding. We approached this question by measuring E2F1 binding to target gene promoters using ChIP (chromatin immunoprecipitation). Transfection of U2OS cells with Flag-E2F1, followed by Flag antibody mediated ChIP, illustrated strong binding of E2F1 to several of its target gene promoters, including TK and CDC6. Crucially, when HA-PAD4 was co-transfected with Flag-E2F1, we observed on average a 30-40% increase in E2F1 binding to these promoter regions (Fig 4.3(a)), a finding which is consistent with the observations from the luciferase reporter assays.

The ability of PAD4 to augment E2F1 binding to target gene promoters appeared to be dependent on its ability to citrullinate E2F1. This is based on the observation that whilst HA-PAD4 transfected into U2OS cells could increase the promoter occupancy of the truncated E2F1 mutant Flag-E2F1^{WT} (1-194), it did not affect the recruitment of the Flag-E2F1^{KKKK} (1-194) mutant (Fig 4.3(b)). Therefore, we concluded that the ability of PAD4 to regulate E2F1 was at least partly mediated through its ability to citrullinate E2F1, although we could not rule out the importance of other indirect effects, including citrullination of core histones as a mechanism of increased E2F1 DNA binding and activity.

Whilst ectopic PAD4 could increase E2F1 binding to target gene promoters, siRNA mediated depletion of PAD4 in U2OS cells, could reduce binding of endogenous E2F1 to several of its target promoters, including CDC6 and E2F1, when compared to control siRNA transfected cells (Fig 4.3(c), compare lanes 4 and 7). Therefore, from the

luciferase reporter assays and ChIP analysis, we concluded that PAD4 can serve as a transcriptional co-activator of E2F1.

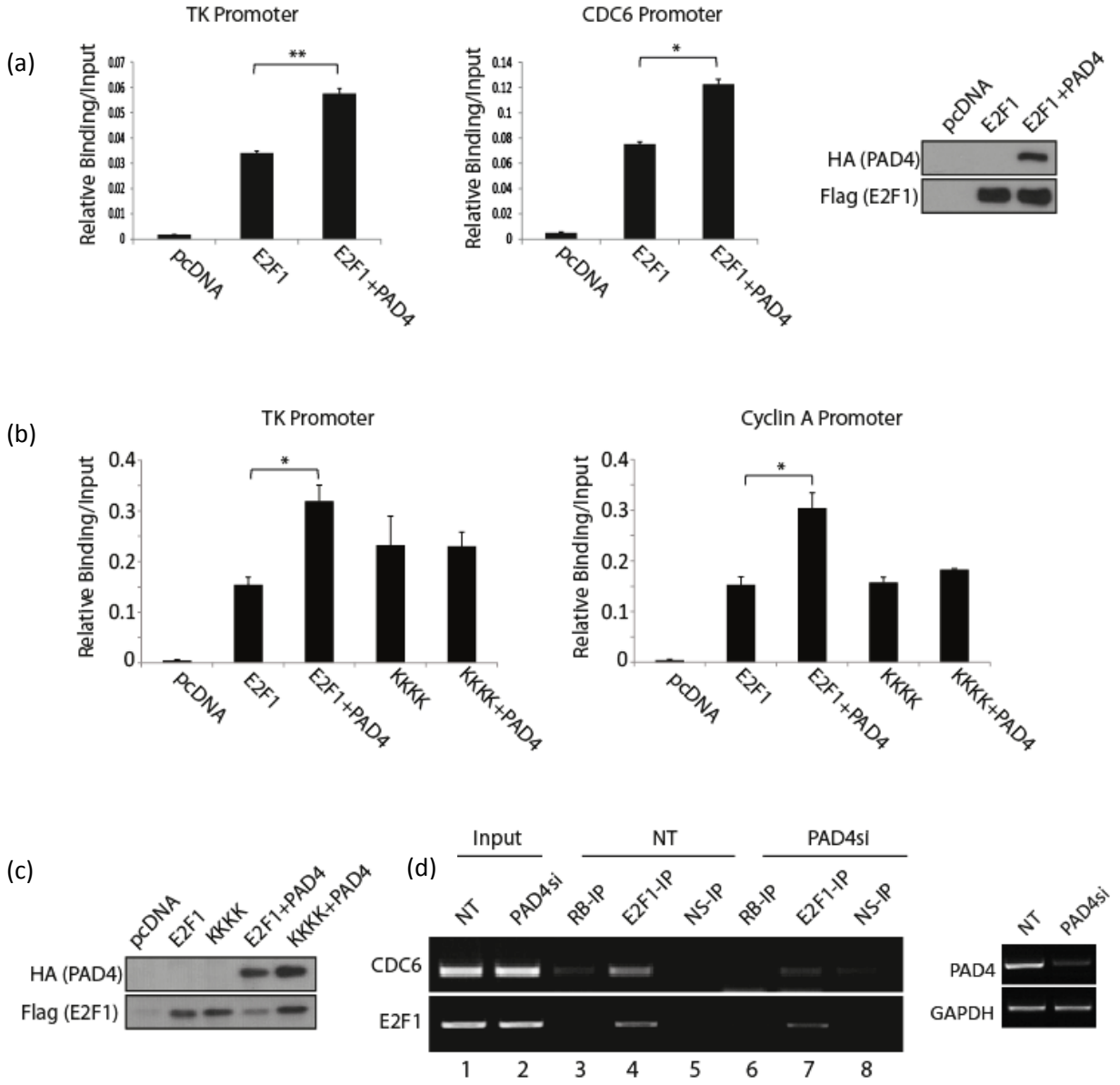


Figure 4.3) PAD4 increases E2F1 DNA binding

(a) Quantitative ChIP PCR from U2OS cells transfected with 1 µg Flag-E2F1 and 2 µg HA-PAD4 and immunoprecipitated using Flag antibody. (n=5) (b) Quantitative ChIP PCR from U2OS cells transfected with 1 µg Flag-ΔC E2F1^{WT} (1-194) or Flag-ΔC E2F1^{KKKK} (1-194) and 2 µg HA-PAD4 and immunoprecipitated using Flag antibody, with corresponding immunoblot (n=3) (c). (d) Non-quantitative ChIP PCR from U2OS cells transfected with 30 nM Non-targeting (NT) siRNA or PAD4 siRNA and immunoprecipitated using E2F1 (C20), pRB (IF8) or IgG antibodies. Successful depletion of PAD4 was confirmed by analysing PAD4 mRNA levels versus GAPDH levels (ethidium bromide gel on the left). NS: non-specific ChIP, ± S.D * $p< 0.05$, ** $p< 0.01$

4.5 PAD4 is recruited to E2F target genes in an E2F1-dependent manner

There are numerous studies which have shown PAD4 being recruited to gene promoter elements to increase or decrease target gene expression. In one study, investigating PAD4 as a transcriptional co-regulator of p53, it was reported that the recruitment of PAD4 to p53 target gene promoters is critically dependent on p53[155]. This is perhaps unsurprising, as PAD4 lacks a defined DNA binding element, and thus its recruitment to chromatin may need to be mediated by other proteins, including transcription factors. Furthermore as previously noted, a ChIP-chip analysis performed in MCF7 cells, analysing the global binding pattern of PAD4, discovered that PAD4's promoter occupancy closely follows that of other transcription factors including E2F1[148].

Given our observations that PAD4 can enhance E2F1 activity and DNA binding, and that the two proteins can interact in cells, it was very plausible to suspect PAD4's presence on E2F1 target gene promoters. Accordingly, we performed Flag antibody mediated ChIP in Flag-PAD4.pTRE cell lines, which were either untreated or treated with doxycycline. As expected we observed binding of PAD4 to several E2F1 target promoters, including CDC6 and E2F1 (Fig 4.4(a), lane 5). Crucially however, the presence of PAD4 on these elements could be shown to be dependent on E2F1, since transient siRNA mediated depletion of E2F1 abolished the promoter binding of PAD4 (Fig 4.4(a), compare lanes 5 and 7). This is whilst the protein levels of Flag-PAD4, as confirmed via immunoblotting, were unaffected by the E2F1 siRNA transfection (fig 4.4(b)). Thus we conclude that PAD4 can be present on E2F target genes in an E2F1-dependent manner.

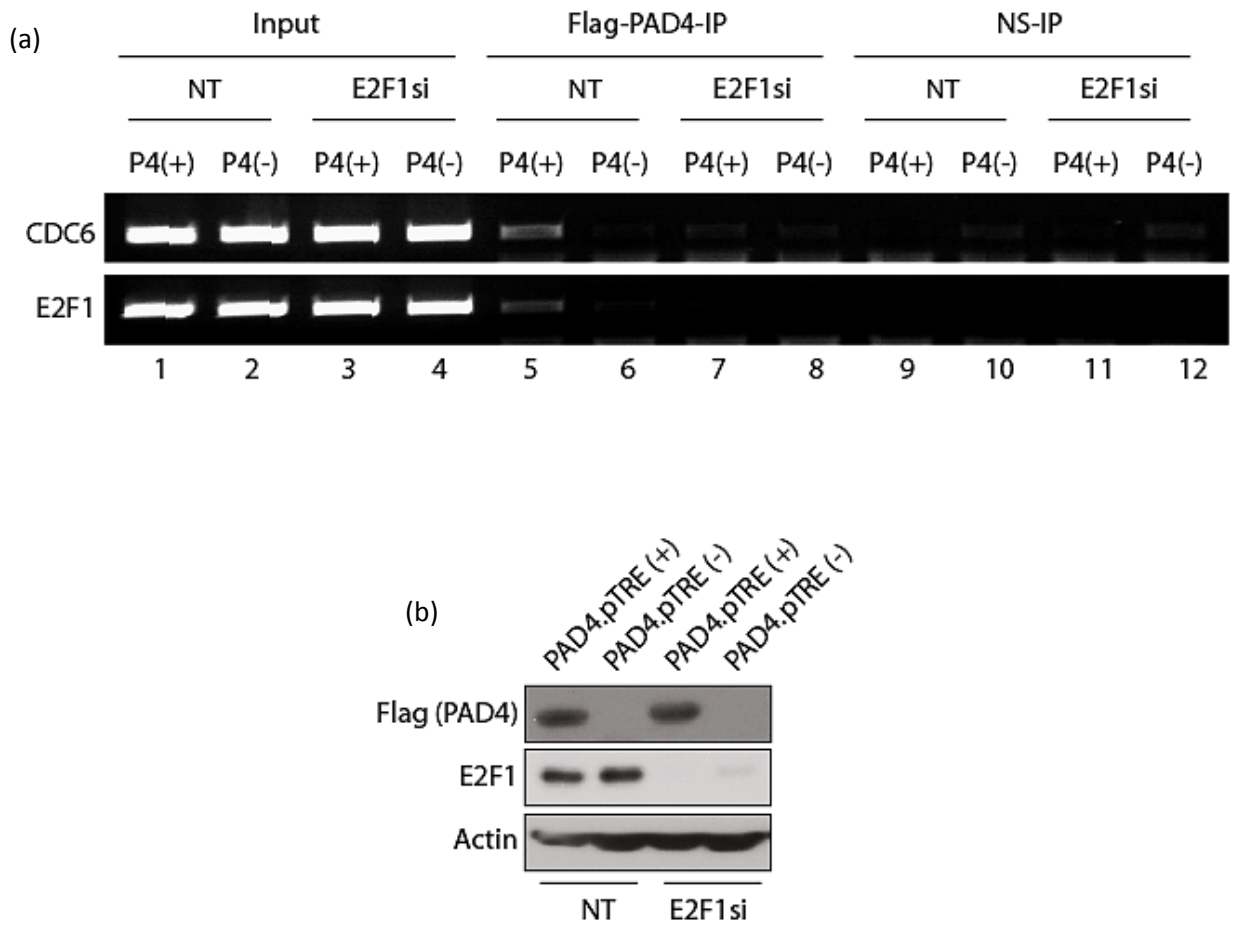


Figure 4.4) PAD4 is recruited to E2F target genes in an E2F1-dependent manner

(a) Non-quantitative ChIP PCR from Flag-PAD4.pTRE cells transfected with 30 nM non-targeting (NT) or E2F1 siRNA for 72 h, treated with (P4 (+)) or without (P4 (-)) 1 µg/ml doxycycline for 36 h, and immunoprecipitated using Flag antibody. (n=2) (b) Corresponding immunoblot to illustrate induction of Flag-PAD4 and depletion of E2F1.

4.6 PAD4 and pRB compete for E2F1

The retinoblastoma protein (pRB) is the main negative regulator of the E2F1 transcription factor. By binding predominantly to the C-terminus of E2F1, pRB effectively hinders the transcriptional activation domain of E2F1, also located in the C-terminus, from engaging with the general transcriptional machinery. Hence pRB can very efficiently suppress the activity of E2F1[47]. Given that PAD4 could augment the activity of E2F1, we investigated whether there could be any interplay between PAD4 and pRB.

In HEK293T cells transfected with HA-E2F1 we could detect an interaction between endogenous pRB and ectopic E2F1. However when Flag-PAD4 was co-transfected into these cells, the interaction between HA-E2F1 and pRB was found to decrease (Fig 4.5(a)). In addition, in Flag-PAD4.pTRE cells we observed that induction of Flag-PAD4 expression coincided with reduced pRB binding to E2F1 target promoters (Fig 4.5(b), compare lanes 3 and 4). These findings were consistent with the observed ability of PAD4 to augment E2F1 activity, although they do not necessarily suggest that PAD4 directly competes with pRB for E2F1 binding, or that citrullination disrupts the E2F1/pRB interaction.

On the other hand, we found that the truncated Flag-E2F1 (residues 1-407) mutant, which lacks the pRB binding domain, interacts with HA-PAD4 more readily than full length Flag-E2F1; potentially suggesting that pRB also impedes the association between E2F1 and PAD4 (Fig 4.5(c)). To confirm that pRB binding to E2F1 could hinder its association with PAD4, we generated the $E2F1^{Y411C}$ point mutant (by substituting tyrosine 411 in the pRB binding domain of E2F1 to a cytosine). This mutant had been reported to maintain transcriptional activity, whilst loosing pRB binding[168]. However unlike the truncation mutant, we did not observe improved binding of the Y411C point mutant to

PAD4 (Fig 4.5(d)). This may be due to the fact that unlike the E2F1 (1-407) truncation mutant which completely lost pRB binding, the $E2F1^{Y411C}$ point mutant maintained residual binding to pRB, which may have been sufficient to competitively preclude PAD4 binding. On the other hand, we cannot exclude the possibility that deletion of the thirty most C-terminal amino acids in E2F1 (1-407) could have induced structural changes which facilitated interaction with PAD4. Alternatively, there may be other important interactors of this region, whose presence are antagonistic to PAD4's binding. Nonetheless, we tentatively conclude that E2F1's recruitment to pRB or PAD4 or protein complexes thereof are mutually exclusive, consistent with the ability of pRB to supress E2F1 and PAD4 to activate it.

4.7 PAD4 increases E2F1 stability

We investigated whether PAD4 could regulate the biochemical properties of E2F1, by assessing how it influences the protein stability of E2F1 in a Cycloheximide time course assay. Cycloheximide can block *de novo* protein synthesis by interfering with mRNA translation, and as such it can be used to monitor the half-life of a given protein in the cell[169]. Comparing Flag-PAD4.pTRE cells induced with doxycycline with the un-induced counterpart demonstrated that the half-life of E2F1 was almost doubled in the presence of PAD4 (half-life increases from 2 hr to 4 hr upon induction of PAD4), suggesting that PAD4 can enhance E2F1 stability (Fig 4.6(a)).

The increased stability of E2F1 upon Flag-PAD4 expression suggested a decrease in its extent of polyubiquitination under these conditions. During the S/G2 phase transition, E2F1 has been shown to be targeted for polyubiquitination and proteasomal degradation[170]. Polyubiquitination is mediated by an ubiquitin system of three enzymes,

E1, E2 and E3. The first of this series is an ubiquitin activating enzyme, which transfers activated ubiquitin to the ubiquitin-conjugating enzyme E2. E2 subsequently transfers the ubiquitin to the ubiquitin ligase enzyme, E3, which catalyses the transfer of the ubiquitin to the substrate. Polyubiquitinated substrates are then targeted to the proteasome for degradation[171]. The extent of this modification on a given substrate could therefore reflect its rate of turnover. In U2OS cells transfected with HA-E2F1 and His-ubiquitin and treated with the proteasome inhibitor MG132, we observed reduced polyubiquitination of E2F1 when PAD4 was co-transfected into cells (Fig 4.6(b)); reinforcing the claim that PAD4 enhances E2F1 stability. MG132 is a potent and cell permeable proteasome inhibitor that reduces the degradation of His-ubiquitin conjugated proteins, and therefore transient treatment of cells with this compound can allow the accumulation and detection of ubiquitin conjugated proteins, which otherwise have a rapid turnover in cells.

These two observations collectively suggest that PAD4 can augment E2F1 stability, which could in turn account for the increased transcriptional activity and promoter binding described earlier. However, these experiments do not address the importance of citrullination *per se*. Therefore, it remains to be established in future works, whether the citrullination of E2F1 directly hinders ubiquitination on proximal and distal lysine residues, or whether the mechanisms involved are an indirect consequence of PAD4overexpression.

4.8 Interplay between PAD4 and cyclin A

Cyclin A, one of the first cyclins to be cloned, and an E2F1 target gene, is expressed during S phase of the cell cycle, where it partners with Cdk2 to facilitate S phase progression[172]. E2F1 possesses a cyclin A binding domain in its N-terminal domain,

stretching from alanine 67 to alanine 108. During S phase, the active cyclin A/Cdk2 complex interacts with E2F1, via its cyclin A binding domain, to phosphorylate both DP1 and E2F1. As a consequence, the DNA binding activity of the E2F1/DP1 dimer is reduced, and its activity suppressed[51]. From earlier results we found that PAD4 preferentially binds to the N-terminus of E2F1, particularly residues 1-110. Although we did not map this interaction any further, we speculated that PAD4 binding to this region may hinder cyclin A binding. Such an interplay was made more probable as the citrullination sites on E2F1 were also close to the cyclin A binding domain.

In U2OS cells, we found that the interaction between Flag-ΔC E2F1 (1-194) and cyclin A was reduced when PAD4 was overexpressed (Fig 4.7(a)). Similarly, endogenous E2F1 immunoprecipitated from Flag-PAD4.pTRE cells exhibited lower binding to cyclin A compared to E2F1 immunoprecipitated from pTRE cells (Fig 4.7(b)), again suggesting that PAD4 can disrupt the interaction between E2F1 and cyclin A.

By implication, we therefore speculated that PAD4 may reduce the phosphorylation of DP1 and E2F1 mediated by Cdk2. Although we did not have specific antibodies against Cdk2 phosphorylated sites on E2F1 and DP1, when we probed with a pan-Cdk phospho antibody, we observed reduced phosphorylation on both E2F1 (Fig 4.7(c)) and DP1 (Fig 4.7(d)) immunoprecipitated from Flag-PAD4.pTRE cells compared to the empty vector pTRE control. We therefore conclude that PAD4 can decrease the binding between cyclin A and E2F1, and accordingly the phosphorylation of DP1 and E2F1 by Cdk2.

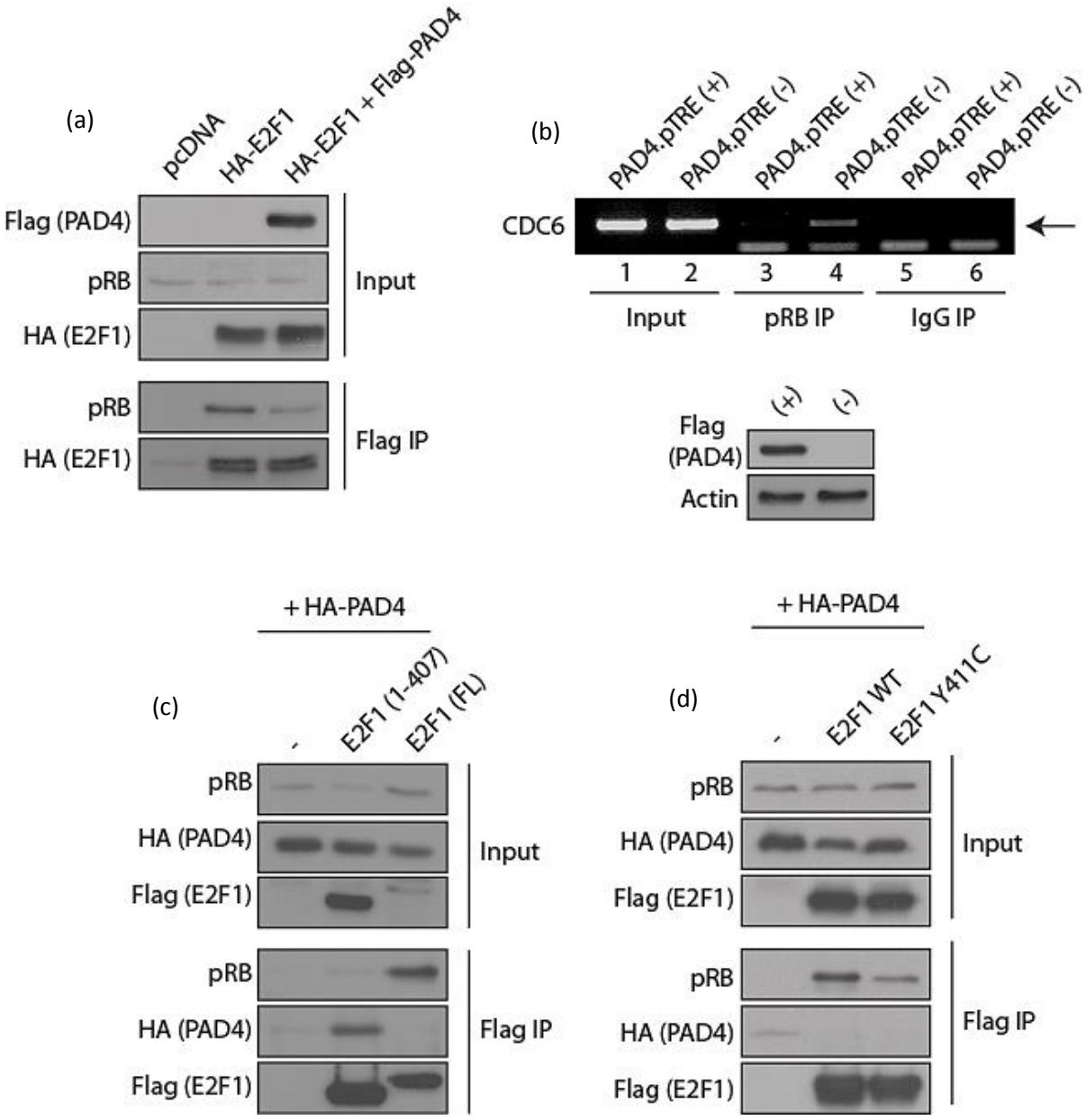


Figure 4.5) Interplay between PAD4 and pRB

(a) HEK293T cells transfected with 1 µg HA-E2F1 and 2 µg Flag-PAD4 were immunoprecipitated using Flag-agarose beads. (n=3) (b) Non-quantitative ChIP PCR from Flag-PAD4.pTRE cell lines treated (+) or not treated (-) with 1 µg/ml doxycycline for 36 h and immunoprecipitated using pRB (IF8) or mouse IgG antibodies. Below the ethidium bromide gel is the corresponding immunoblot to illustrate induction of Flag-PAD4. (n=2) (c) U2OS cells transfected with 1 µg full length (FL) Flag-E2F1 or truncated Flag-E2F1 (1-407) and 2 µg HA-PAD4 were immunoprecipitated using Flag-agarose beads. (d) U2OS cells transfected with 1 µg Flag-E2F1^{WT} or Flag-E2F1^{Y411C} and 2 µg HA-PAD4 were immunoprecipitated using Flag-agarose beads.

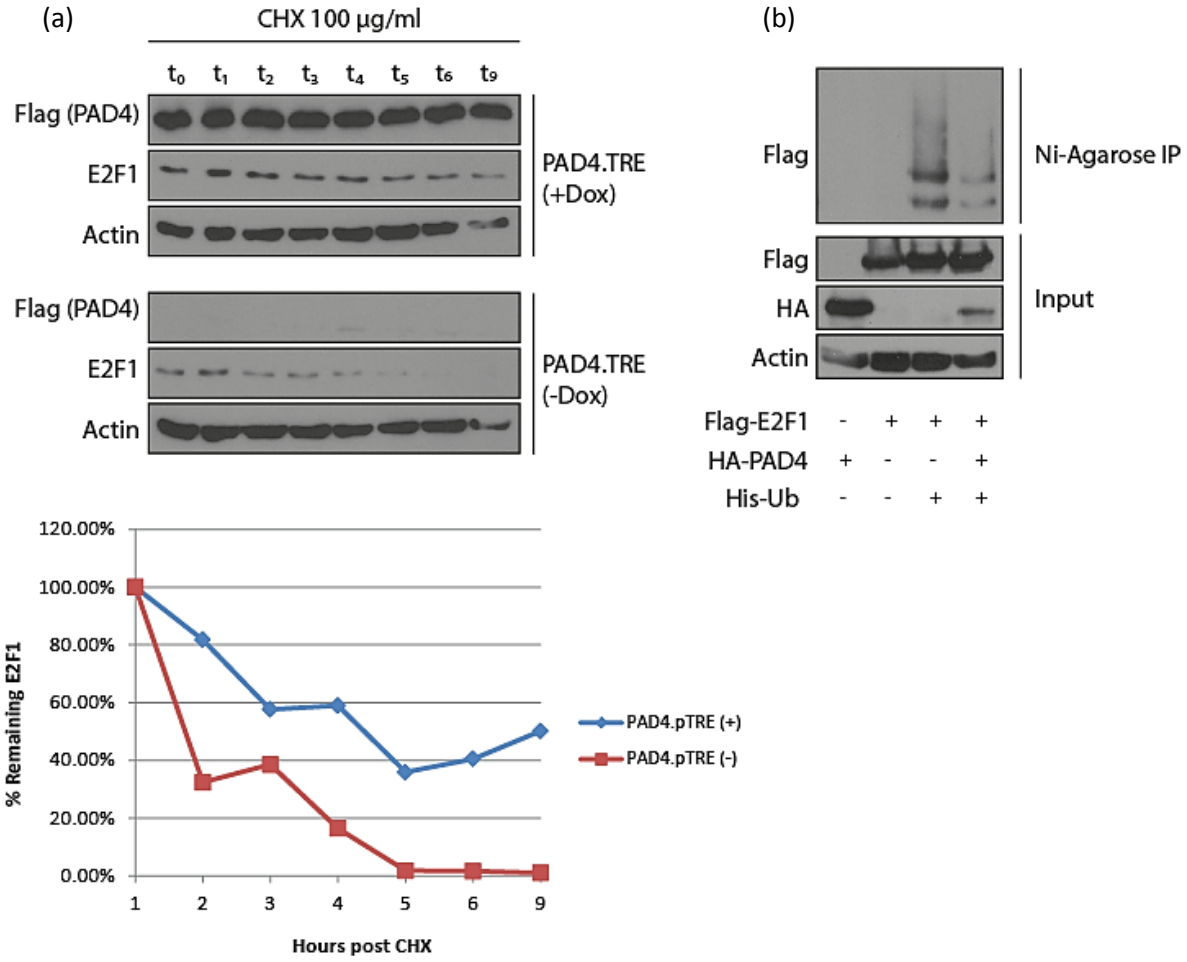


Figure 4.6) PAD4 increases E2F1 stability

(a) Flag-PAD4.pTRE cell lines were treated with (top panel) or without (bottom panel) 1 µg/ml doxycycline for 36 h. All cells were then treated with 100 µg/ml Cycloheximide, and were harvested at the time points indicated (t_0 = 0 h post-CHX, t_1 = 1 h post-CHX, etc). Flag and E2F1 (KH95) antibodies were used to immunoblot for Flag-PAD4 and E2F1 respectively. The E2F1 signal intensity was quantified using Image J Software and normalised to actin loading control. The relative amount of E2F1 at t_0 was arbitrarily assigned to 100%. (n=3) (b) U2OS cells transfected with 1 µg His-ubiquitin, 1 µg Flag-E2F1 and 2 µg HA-PAD4 were treated with 4 µM MG132 for 4 h and immunoprecipitated using Nickel agarose beads. Immunoblotting using Flag antibody in the IP lane illustrates the extent of E2F1 ubiquitination. (n=2)

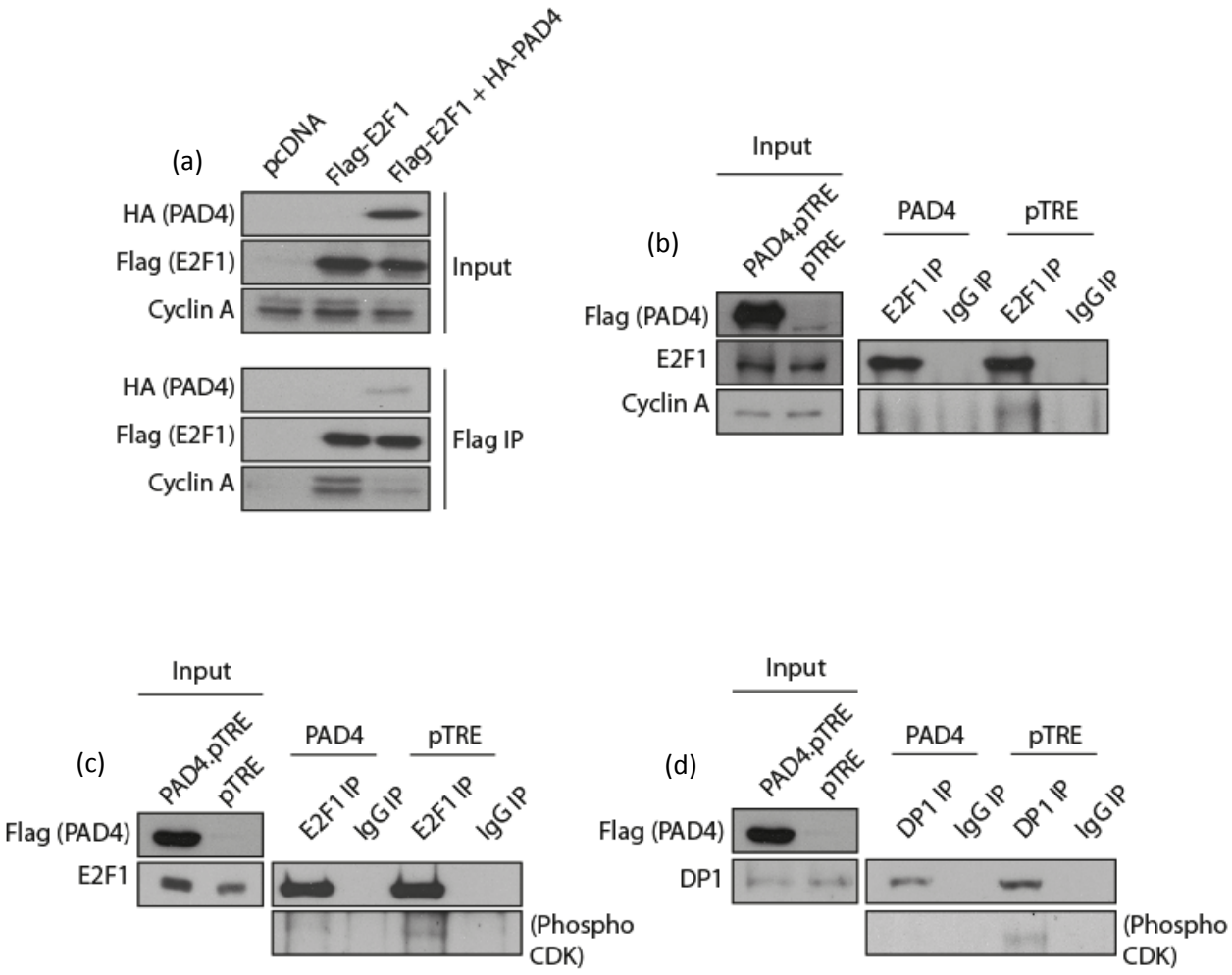


Figure 4.7) Interplay between PAD4 and cyclin A

(a) U2OS cells transfected with 1 µg Flag-ΔC E2F1 (1-194) and 2 µg HA-PAD4 were immunoprecipitated using Flag-agarose beads. (b) Flag-PAD4.pTRE or pTRE cell lines treated with 1 µg/ml doxycycline for 36 h were immunoprecipitated using E2F1 (C20) or rabbit IgG control antibodies. (n=2) (c) and (d) Flag-PAD4.pTRE or pTRE cell lines treated with 1 µg/ml doxycycline for 36 h were immunoprecipitated using E2F1 (C20) (For experiment c) or DP1 (SCB, Rb) (For experiment d) or rabbit IgG control antibodies. The pan-phospho Cdk (Cell Signalling) antibody was used to immunoblot Cdk-mediated phosphorylation. (n=2)

4.9 Chapter Summary

From this chapter we conclude that PAD4 could be a positive regulator of E2F1 activity and stability. Using reporter assays we find that PAD4 can augment E2F1 activity, in a manner that may depend on effective citrullination of E2F1. Furthermore we find this to coincide with increased DNA recruitment of E2F1, reduced pRB association and increased half-life. These results are consistent with previous ChIP-chip analysis demonstrating that PAD4 is largely enriched near TSS (transcription start site) of active genes, and that PAD4 binding pattern closely follows that of E2F1. We suspect that in cancer cells PAD4 could augment cellular growth rate and proliferation through E2F1, and this may be part of the mechanism for the cytotoxic and cytostatic effects of PAD inhibitors. We have not however explored the apoptotic activities of E2F1 in any depth, and it remains to be established how PAD4 could regulate the pro-apoptotic capacity of E2F1.

Chapter 5) The E2F1-PAD4 axis and inflammation

5.1 Introduction

In the early 1990's peptidyl arginine deiminase activity was reported in mouse and rat hematopoietic cells, including peritoneal macrophages, neutrophils and monocytes[173]. In line with these early studies, a novel PAD enzyme was later identified and cloned from cultured myeloid leukaemia HL60 cells, as they differentiated into granulocytes or monocytes[123]. This PAD enzyme, now known to be PAD4, was also subsequently found to be abundantly expressed in human peripheral blood cells, particularly neutrophils and eosinophils[124]. Since the discovery that PAD4 expression and activity can be induced in differentiating HL60 cells, this system has provided a powerful model to investigate the function and regulation of PAD4. We therefore utilized this cellular model to investigate the regulation of E2F1 by PAD4.

5.2 PAD4 is induced in differentiating HL60 cells

HL60 cells are myeloid leukemic cells, which can be differentiated to granulocytes with dimethyl sulfoxide (DMSO) or all-trans retinoic acid (ATRA), to macrophages with 12-O-Tetradecanoylphorbol 13-acetate (TPA) or to monocytes with 1α, 25-dihydroxy vitamin D_3[174]. PAD4 expression has been demonstrated in these cells as they differentiate into granulocytes or monocytes, but not macrophages[123]. We could recapitulate these reports, by treating HL60 cells with 1% DMSO, 1 μM ATRA or 10 ng/ml TPA and immunoblotting for PAD4, whereupon we observed robust expression of PAD4 following treatment with DMSO or ATRA, but not TPA (Fig 5.1(a)). DMSO is a small polar molecule, whose cytostatic and pro-differentiating effects on cancer cells, led to the discovery of suberoylanilide hydroxamic acid (SAHA), which is a HDAC inhibitor[175].

Although the mechanistic basis for DMSO induced differentiation of HL60 cells have not been fully unravelled, it is plausible that inhibition of HDAC enzymes, and the associated altered patterns of acetylation on histones and non-histone proteins, could be mediating the effects of DMSO. Expression of the CD11B cell surface protein was used to confirm differentiation, as this is a marker which is expressed on the surface of several innate immune system leukocytes, including granulocytes, monocytes and macrophages[176]. Since PAD4 levels were more dramatically induced with DMSO treatment, and given the reports that a higher proportion of HL60 cells are resistant to differentiation by ATRA (40%) than with DMSO (10%)[174], we thus used DMSO for subsequent experiments.

In a time course assay, comparing mock versus DMSO treated HL60 cells, we found that PAD4 levels are upregulated early in the treatment (day 2), with maximal induction achieved by day 6 (Fig 5.1(b)). Intriguingly, we found that upon differentiation of HL60 cells, E2F1 shifts to a higher migrating band. A different pattern of post-translational modifications is likely to case this shift, although we do not believe PAD4 to be responsible for this. This is because a similar band shift was observed when HL60 cells are treated with TPA or doxorubicin in the absence of PAD4 expression (Fig 5.1(c)). Since maximal induction of PAD4 on day 6 (post-DMSO treatment) coincided with loss of E2F1 protein levels, for subsequent experiments investigating the E2F1/PAD4 interplay, we limited the DMSO treatment to 2-3 days.

We confirmed that both E2F1 and PAD4 are localised to the nucleus by performing a standard cell fractionation protocol, where the cytoplasmic fraction was separated by means of a hypotonic buffer to exclude the nuclear fraction, which was then disrupted using a more chaotropic high salt buffer. By analysing these two fractions using immunoblotting, with the exclusively cytoplasmic protein GAPDH as a marker for the cytoplasmic fraction and the nuclear protein lamin B as a marker for the nuclear fraction,

we found that both E2F1 and PAD4 are indeed localised to the nucleus before and after differentiation (Fig 5.1(d)). We also confirmed an interaction between endogenous E2F1 and PAD4 in Mg132for E2F1 (Fig 5.1(e)). We therefore used this system to determine the molecular events underlying the E2F1/PAD4 interplay.

5.3 E2F1 is citrullinated in differentiated HL60 cells

Having observed that PAD4 becomes expressed in differentiated HL60 cells, and that it can interact with E2F1, we asked whether under these conditions, E2F1 could be targeted for citrullination. To this end, we immunoprecipitated E2F1 from undifferentiated and differentiated HL60 cells, incubated with $CaCl_2$ and DTT, and immunoblotted using the Millipore AMC kit. As hypothesised, we found endogenous E2F1 to become citrullinated in this cellular system as they underwent differentiation with DMSO (Fig 5.2(a)). Furthermore we found that citrullination of E2F1 could be prevented by pre-treating differentiated cells with the pan-PAD inhibitor BB-Cl-amidine[131] prior to incubating the lysate with $CaCl_2$ and DTT (Fig 5.2(b)). Although the inhibitor does not distinguish between PAD2 and PAD4, it does nonetheless support the observation that E2F1 can become citrullinated in differentiated HL60 cells in a PAD-dependent manner.

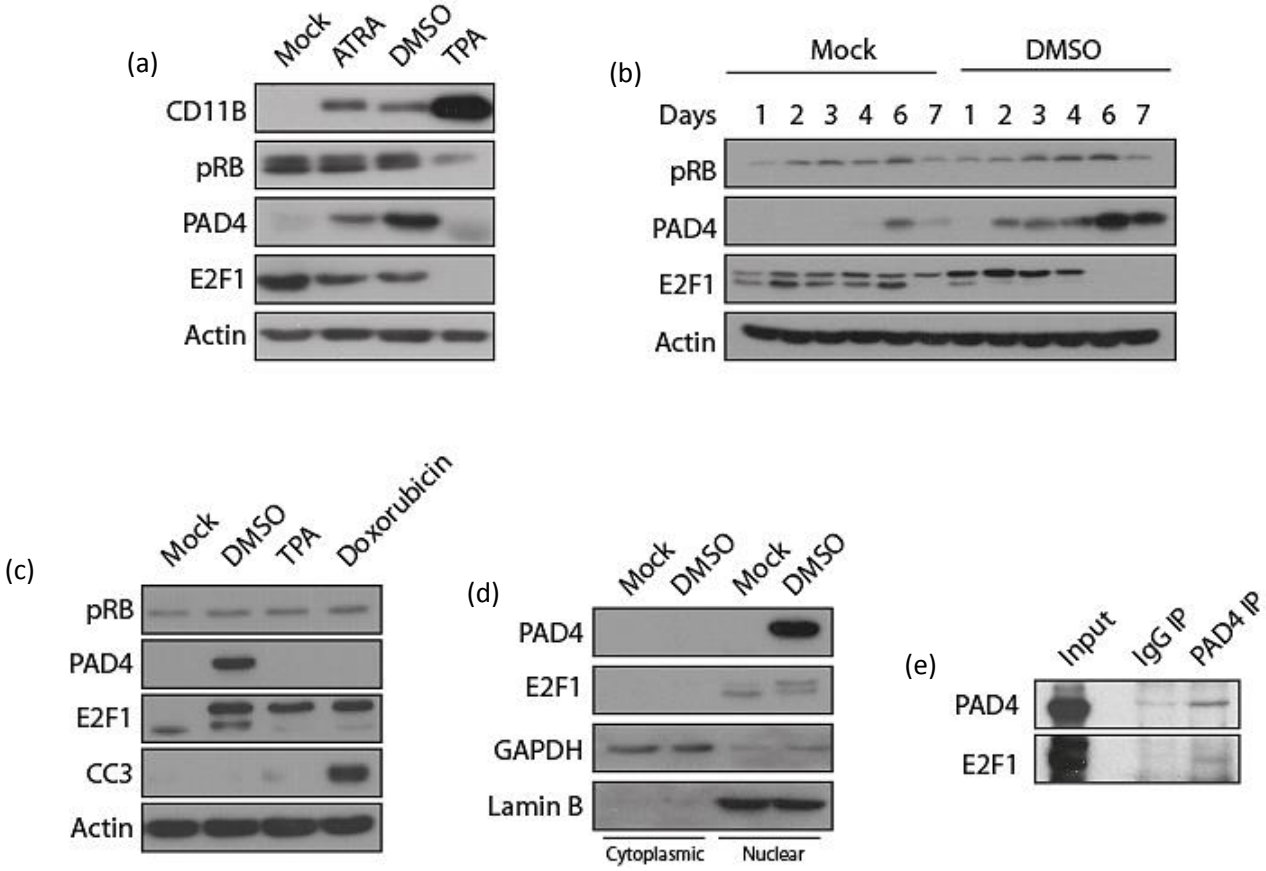


Figure 5.1) PAD4 is upregulated in differentiating myeloid leukemic HL60 cells

(a) HL60 cells were treated with 1% DMSO, 1 µM ATRA or 10 ng/ml TPA (72 h) (b) HL60 cells were treated with 1% DMSO or left untreated (mock) and were harvested every 24 h for 7 days. (c) HL60 cells were treated with 1% DMSO (48 h), 10 ng/ml TPA (48 h) or 2 µM doxorubicin (24 h). (d) HL60 cells were treated with 1% DMSO (48 h) or left untreated (mock) and were fractionated to separate the cytoplasmic and nuclear compartment. GAPDH was used as a marker for the cytoplasmic fraction and lamin B as a marker for the nuclear fraction.(n=3) (e) HL60 cells were treated with 1% DMSO (48 h) and immunoprecipitated using PAD4 or rabbit IgG antibodies. (n=2)

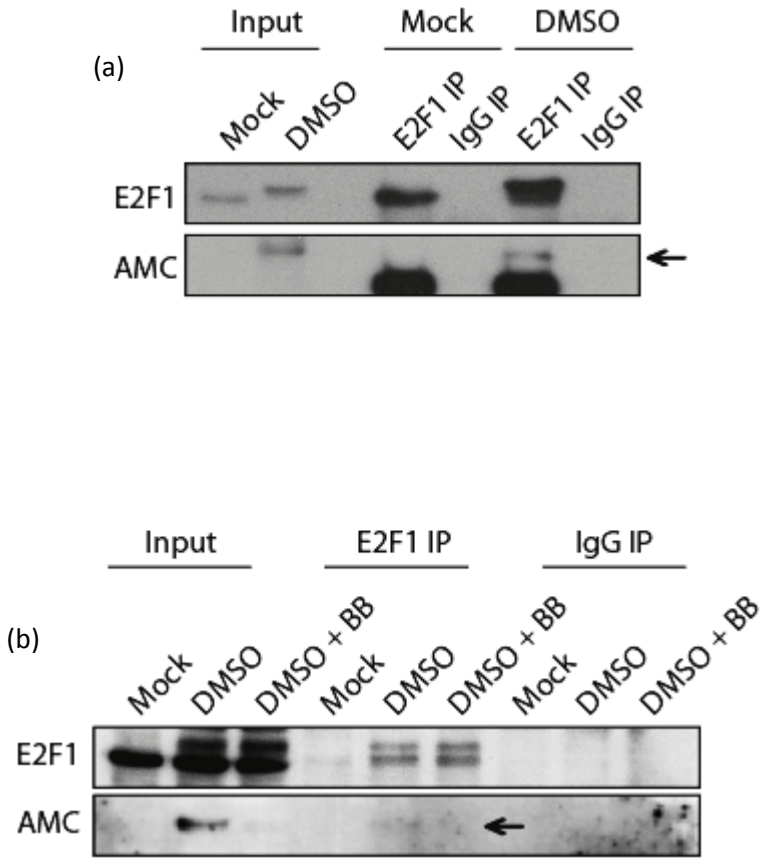


Figure 5.2) E2F1 is citrullinated in granulocyte differentiated HL60 cells

(a) HL60 cells were treated with 1% DMSO (48 h) or left untreated, lysed in the presence of 4mM $CaCl_2$ and 2 mM DTT and immunoprecipitated using E2F1 (C20) or rabbit IgG antibodies. Millipore AMC kit was used to probe for citrullination. (n=2) (b) HL60 cells were treated with 1% DMSO for 48 h or left untreated, lysed in the presence of 100 μM BB-Cl-amidine where indicated and incubated with 4 mM $CaCl_2$and 2 mM DTT. E2F1 immunoprecipitation and immunoblotting for citrullination was performed as above. (n=1)

5.4 Cell cycle changes in differentiated HL60 cells

DMSO induced differentiation of HL60 cells has been shown to be associated with significant qualitative and quantitative changes in gene expression patterns. In one study, microarray analysis was performed to compare transcript levels between undifferentiated HL60 cells versus DMSO differentiated cells, which as a result, acquire phenotypic hallmarks of mature neutrophils. Unsurprisingly, significant down-regulation of many cell cycle related genes were amongst the changes documented as the cells differentiate[177]. Consistent with these reports, we also observed HL60 cells arresting in G1 (Fig 5.3(a)) as they underwent differentiation. The levels of cyclin A, a known E2F1 target gene associated with cell cycle progression, could also be seen to decrease, whilst that of the Cdk inhibitor $p27^{Kip1}$ was found to increase (Fig 5.3(b)). Furthermore, at the transcript level, E2F1 mRNA was downregulated (Fig 5.3(c)), and at the protein level, the association between E2F1 and pRB was augmented in DMSO differentiated cells (Fig 5.3 (d) and (e)).

Collectively, these observations are consistent with the profile of a cell which is exiting the cell cycle and is undergoing differentiation. In relation to the regulation of E2F1 by PAD4, they also suggest a potentially different interplay than what we observed and reported in the rapidly proliferating undifferentiated cellular systems explored in the previous chapter. We therefore looked beyond the classical functions of E2F1 in cell cycle control in trying to understand the role of citrullinated E2F1 in this cellular system.

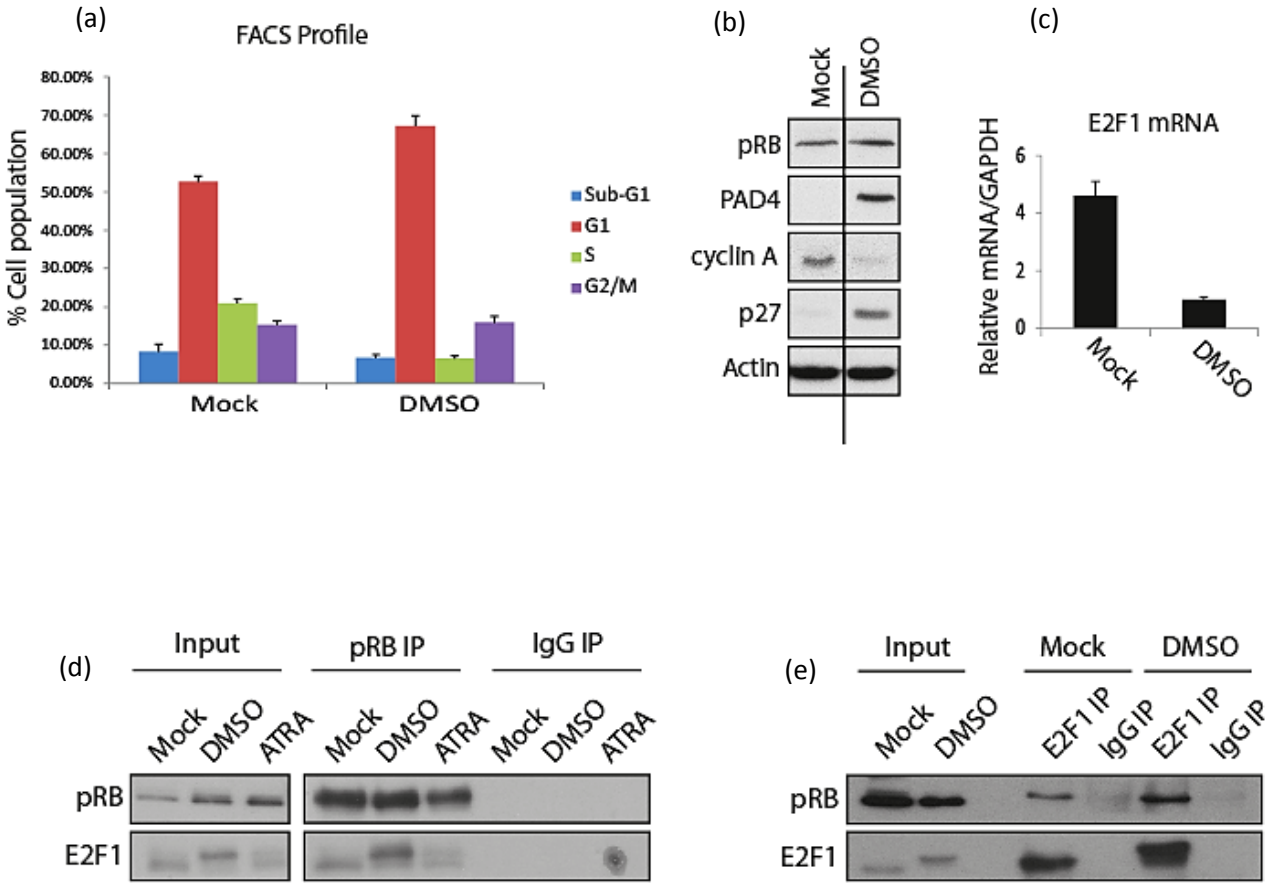


Figure 5.3) Cell cycle changes in differentiating HL60 cells

(a) HL60 cells were untreated (mock) or treated with 1% DMSO (48 h), and the cell profiles were analysed by FACS. (b) Representative immunoblot illustrating changes in the levels of cyclin A, $p27^{Kip1}$, PAD4 and pRB in mock versus DMSO treated HL60 cells. (c) Quantitative real time PCR from mock versus DMSO treated HL60 cells. Relative E2F1 mRNA levels were normalised to housekeeping GAPDH mRNA levels. (d) HL60 cells were untreated, or treated with 1% DMSO or 1 µM ATRA (48 h) and immunoprecipitated using pRB (IF8) or mouse IgG antibodies. (e) HL60 cells were untreated, or treated with 1% DMSO and immunoprecipitated using E2F1 (C20) or rabbit IgG antibodies.

5.5 E2F1 is recruited to cytokine promoters

In the same microarray study where many cell cycle genes were shown to be downregulated in DMSO differentiated cells, several cytokines were found to be significantly upregulated[177]. This was of interest, since in an independent study in THP1 macrophages challenged with lipopolysaccharides (LPS), E2F1 was reported to be recruited to several cytokine gene promoters[81]. A number of the cytokines regulated by E2F1, including chemokine (C-C) motif 3 (CCL3) and interleukin-1β (IL-1β), were also amongst the genes reported to be transcriptionally upregulated in HL60 cells as they differentiate with DMSO.

Furthermore, various inflammatory stimuli, including LPS and TNFα, have been shown to enhance PAD activity in differentiated HL60 cells and promote citrullination[146]. Accordingly, we asked whether E2F1could be regulating cytokine expression in this cellular system, and whether PAD4 could impact on this. We focused on the three cytokine CCL3, IL-1β and TNFα, as these have been shown to be transcriptional targets of E2F1.

In a ChIP experiment, E2F1, pRB and PAD4 were immunoprecipitated from untreated versus DMSO treated HL60 cells, and binding to the promoter regions of IL-1β, CCL3 and TNFα were assessed using previously published primers[81]. Interestingly, we observed strong recruitment of E2F1 to the promoter regions of these cytokine genes in DMSO differentiated HL60 cells (Fig 5.4(a), compare lanes 3 and 7), whilst binding to the cell cycle gene TK was unaffected. Coincident with this, we observed reduced pRB (Fig 5.4(a), compare lanes 4 and 8) yet increased PAD4 binding (Fig 5.4(a), compare lanes 5 and 9) to the cytokine genes. Interestingly this is whilst pRB binding to the TK promoter appeared to increase under DMSO treatment (Fig 5.4(a)). These observations would be

consistent with the presence of a transcriptionally active E2F1/PAD4 complex on cytokine promoters, and an inactive E2F1/pRB complex on cell cycle genes in differentiated cells.

As previously mentioned, TNFα has been reported to augment PAD4 enzymatic activity in DMSO-treated HL60 cells[146]. We confirmed these reports by using the pan-citrullinated histone 3 (H3) antibody as a surrogate for PAD4 activity, whereupon we observed increasing citrullination of H3 as HL60 cells were differentiated and treated with 10 ng/ml TNFα (Fig 5.4(b)). In a similar manner, we observed increasing binding of E2F1 to the TNFα promoter in cells co-treated with DMSO and 10 ng/ml TNFα (Fig 5.4(c), lane 11).; i.e. conditions that coincided with maximal PAD4 expression and activation. We therefore conclude that E2F1 can be recruited to promoters of several cytokine genes, and speculated that PAD4 may influence the binding of E2F1 to these genes.

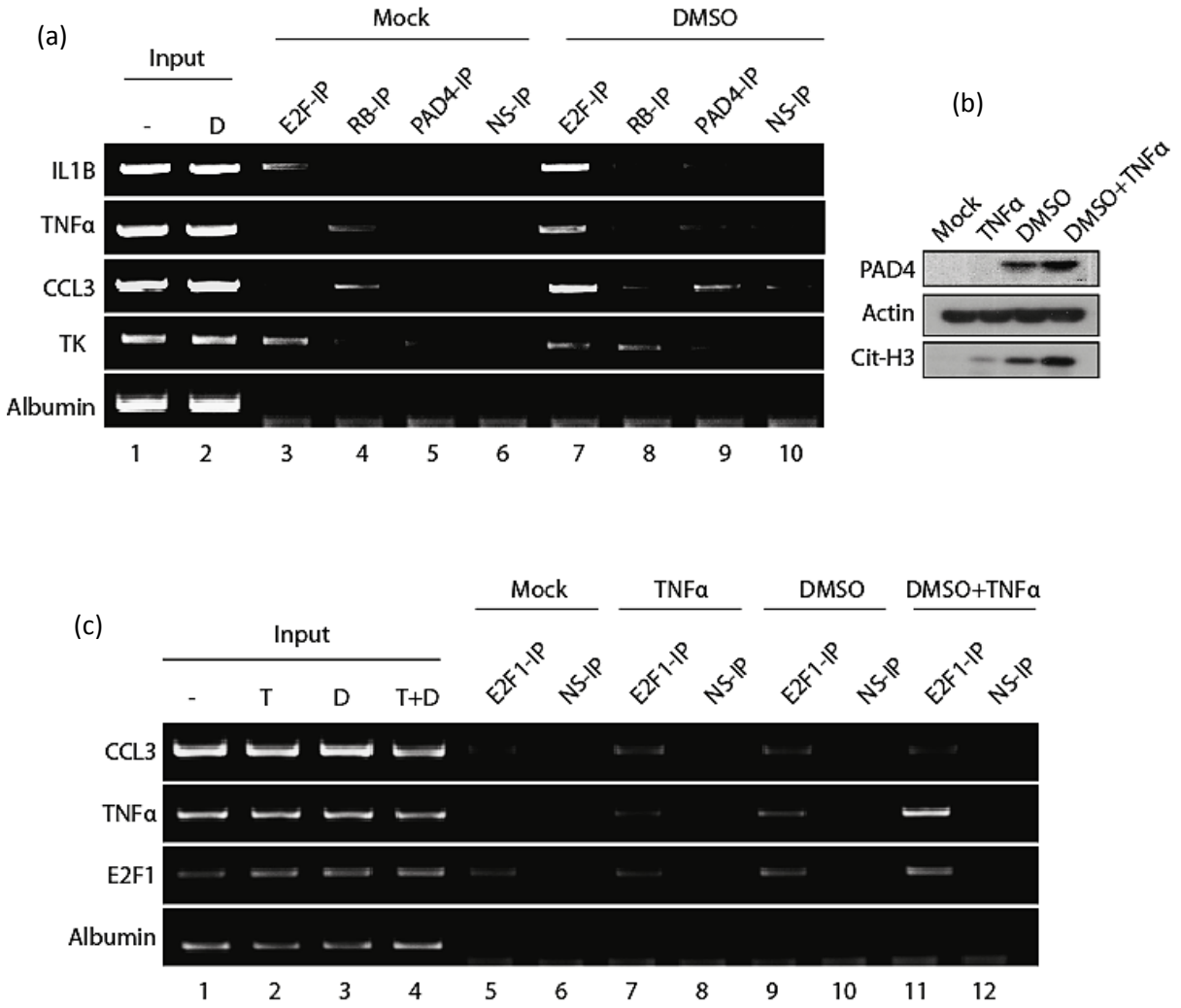


Figure 5.4) E2F1 is recruited to cytokine promoters

(a) Non-quantitative ChIP PCR from HL60 cells untreated or treated with 1% DMSO (48 h) and immunoprecipitated using E2F1 (C20), pRB (IF8), PAD4 or IgG antibodies. (b) Where indicated HL60 cells were treated with 1% DMSO (48 h) and/or 10 ng/ml TNFα (3 h). (n=3) (c) Non-quantitative ChIP PCR from HL60 cells treated with 1% DMSO (48 h) and/or 10 ng/ml TNFα (3 h) where indicated and immunoprecipitated using E2F1 (C20) or IgG antibodies. (n=2) T: TNFα, D: DMSO, NS: non-specific ChIP

5.6 E2F1 recruitment to promoters of cytokine genes is PAD4-dependent

To investigate whether PAD4 could influence the recruitment of E2F1 to these cytokine gene promoters, we took advantage of the pan-PAD inhibitor BB-Cl-amidine, and the selective PAD4 inhibitor, GSK484[178]. DMSO differentiated HL60 cells were challenged with LPS, treated with the pan-PAD inhibitor BB-Cl-amidine and subjected to ChIP using E2F1 antibody. Consistent with the literature[81], recruitment of E2F1 to IL-1β and TNFα promoters was seen to increase upon LPS treatment (Fig 5.5(a)). Crucially, in BB-Cl-amidine treated cells, the relative promoter occupancy of E2F1 was reduced (Fig 5.5(a)), therefore suggesting that PAD activity is important for targeting E2F1 to these genes.

We could specifically confirm the importance of PAD4 in relation to this observation by using the PAD4 specific inhibitor GSK484. Treating DMSO differentiated HLSO cells with GSK484 coincided with reduced levels of E2F1 on the promoters of select cytokines (Fig 5.5(b)). Overall we therefore conclude that E2F1 can be present on several cytokine promoters in a manner that is dependent on PAD4 activity, since inhibition of this enzyme using broad or selective inhibitors could abrogate the promoter occupancy of E2F1.

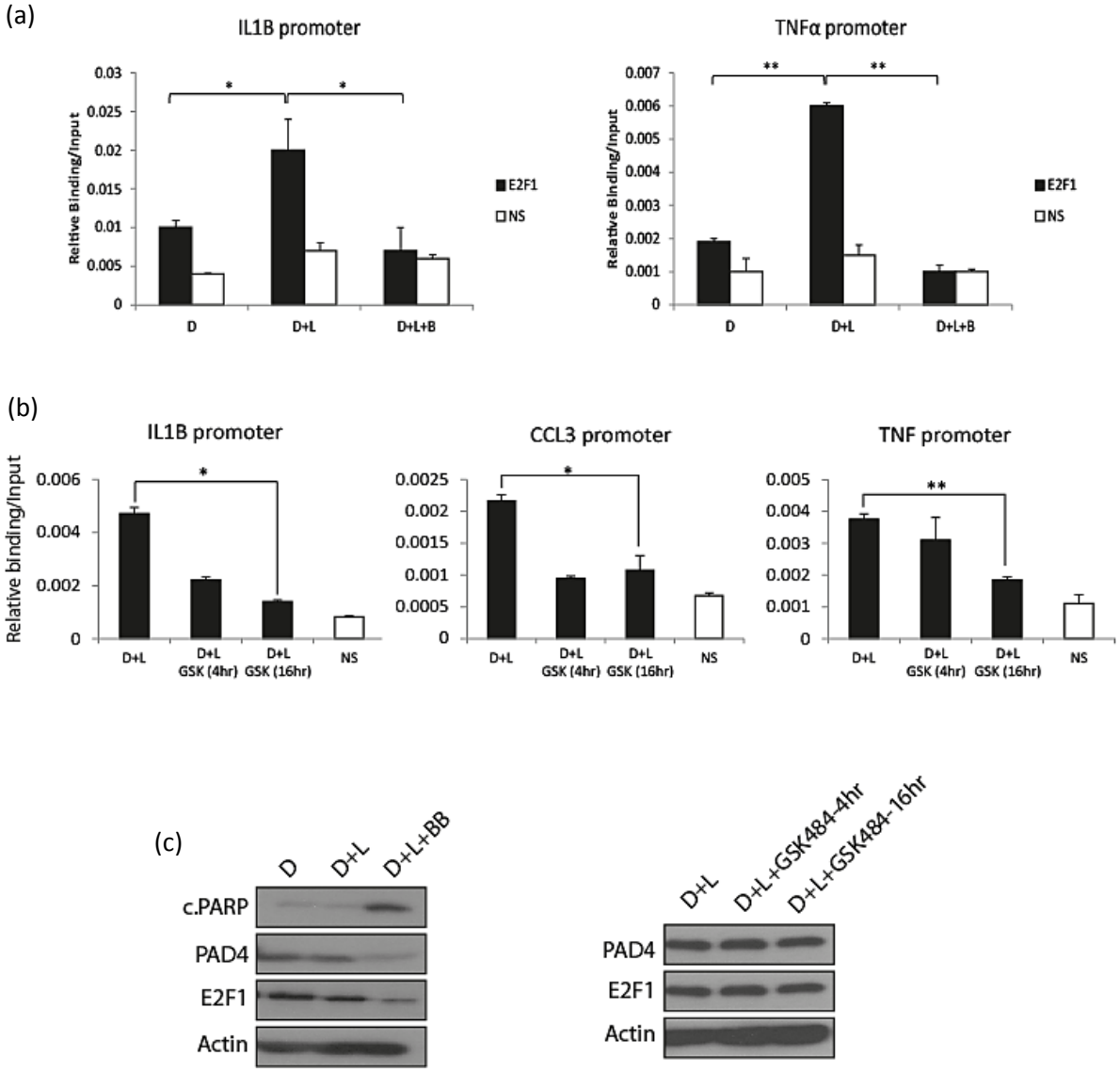


Figure 5.5) E2F1 recruitment to cytokine promoters is PAD4 dependent

(a) Quantitative ChIP PCR from HL60 cells treated with 1% DMSO (48 h), 2.5 μM BB-Cl-amidine (16 h) and 100 ng/ml LPS (3 h). E2F1 (C20) and rabbit IgG antibodies were used for ChIP immunoprecipitation. (n=3) (b) Quantitative ChIP PCR from HL60 cells treated with 1% DMSO (48 h), 10 μM GSK484 (4 or 16 h) and 100 ng/ml LPS (3 h). E2F1 (C20) and rabbit IgG antibodies were used for ChIP immunoprecipitation. (n=3) (c) Blots corresponding to the above ChIP experiments. D: DMSO, L: LPS, BB: BB-Cl-amidine, NS: non-specific ChIP, ±S.D * $p < 0.05$, ** $p < 0.01$

5.7 PAD4 regulates cytokine expression levels

Previous reports have shown PAD4 to regulate the expression of several cytokine genes. One study suggested that PAD4 mediated citrullination of arginine 8 of histone 3 (H3R8) could hinder the binding of the transcriptional repressor HP1 to the adjacent residue, trimethylated lysine 9 of histone 3 ($H3K9(me)_3$) [145]. Displacement of HP1 from the promoter regions of these cytokine genes then results in their transcriptional reactivation and expression.

Since HL60 cells are suspension cells and difficult to transfect, we were not successful at manipulating PAD4 levels in these cells through transfection of ectopic plasmids or siRNA. We therefore used U2OS cells, which can be more readily transfected, to confirm that manipulation of PAD4 levels can indeed affect the transcript levels of cytokine genes. We assessed the expression of the cytokines of interest by measuring their mRNA levels using quantitative real time PCR. When we overexpressed PAD4 in U2OS cells, we found that there is a moderate, but nonetheless statistically significant increase in the transcript levels of IL-1β and TNFα (Fig 5.6(a)). Likewise when PAD4 mRNA was transiently downregulated using siRNA transfection, we observed a correlated decrease in the levels of IL-1β and TNFα transcript levels (Fig 5.6(b). These observations support a positive role for PAD4 in augmenting the expression of pro-inflammatory genes.

In a genome-wide transcript analysis performed early in this project, we obtained further evidence for the involvement of both PAD4 and E2F1 in immune response regulation, as described in the following section.

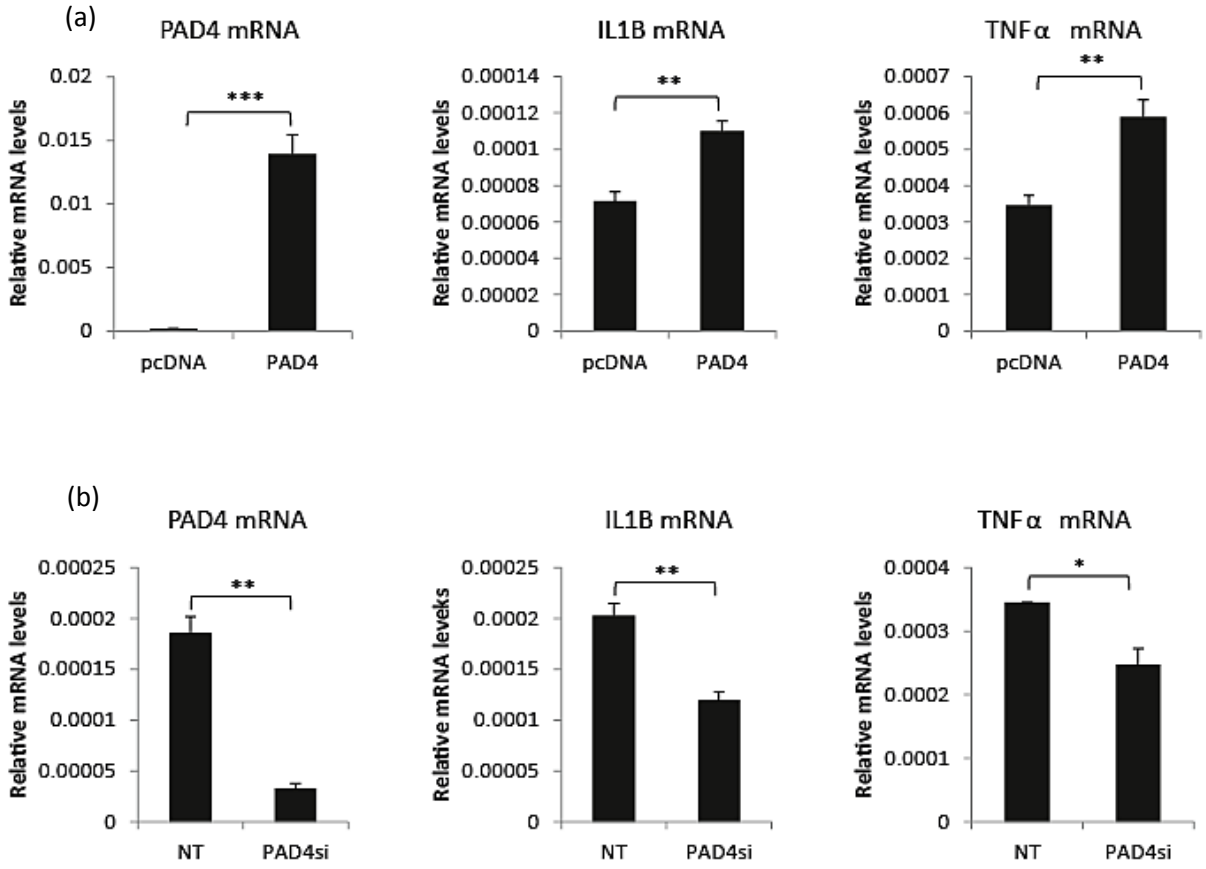


Figure 5.6) PAD4 augments expression of cytokines

(a) Quantitative real time PCR from U2OS cells transfected with 2 µg pcDNA or HA-PAD4 and treated with 100 ng/ml LPS (3 h). Relative mRNA levels were normalised to housekeeping GAPDH mRNA levels. (n=2) (b) Quantitative real time PCR from U2OS cells transfected with 30 nM non-targeting or PAD4 siRNA and treated with 100 ng/ml LPS (3 h). Relative mRNA levels were normalised to housekeeping GAPDH mRNA levels. (n=2) ±S.D * $p < 0.05$, ** $p < 0.01$, ***$p < 0.001$

5.8 Microarray analysis of U2OS cells

In order to gain better insight into the cellular functions of E2F1 and PAD4, we performed a genome-wide gene expression microarray analysis in U2OS cells, where we transiently ablated E2F1 and PAD4 mRNA levels using siRNA transfection (Fig 5.7) and analysed the samples using the Illumina Human HT12v4.0 gene expression array. We then analysed the gene expression datasets using the publicly available GSEA (Gene Set Enrichment Analysis) software[179]. GSEA is an analytical and bioinformatics tool for analysing gene expression patterns, and assigning statistically enriched pathways to a given gene set.

In our analysis, the E2F1si and PAD4si groups were normalised to the non-targeting (NT) control group, and as such, the adjusted p-value and relative fold change of each transcript was expressed relative to the control sample. These lists (i.e. E2F1si vs. NT and PAD4si vs. NT) were then analysed using GSEA, which ranks each gene on the merit of its fold change and statistical significance, to determine the biologically meaningful pathways which are enriched in the given sample. Interestingly we found that depletion of both E2F1 (Fig 5.7(a)) and PAD4 (Fig 5.7(b)) correlated with down-regulation of immune system process, with a normalised enrichment scores of -2.43 and -1.93, respectively.

We also analysed the common set of transcripts between the two groups (i.e. E2F1si vs. NT and PAD4si vs. NT) using the gene list analysis tool PANTHER (PANTHER Overrepresentation Test, version 9.0 Released 2014-01-24)[180] (Fig 5.8(a)). Crucially, we found that within this common set of genes, the inflammatory response was enriched and statistically significant (Fig 5.8(b)) (pathways with an enrichment score above 1 are considered to be overrepresented and a p-value less than 0.05 is considered significant). Many of the genes that overlapped between these two groups were cytokines

and cytokine receptors (Fig 5.8(c)), although we note that the genes which we analysed in HL60 cells (i.e. CCL3, TNFα and IL-1β) did not feature within this category, which may reflect differences between the two cellular systems. Nevertheless the genome-wide gene expression microarray analyses suggested that E2F1 and PAD4 can affect expression of various cytokines.

5.9 Chapter Summary

In this chapter we described a novel context under which we propose PAD4 to be regulating E2F1. We observed that as the myeloid leukemic HL60 cells differentiate to granulocytes, PAD4 levels are upregulated, and E2F1 becomes recruited to a number of cytokine promoters. Crucially we found that PAD4 activity was necessary for this, since broad and selective PAD4 inhibitors could abrogate E2F1 binding to these promoter elements.

In the previous chapter, we found that PAD4 can positively regulate the transcriptional activity of E2F1. Comparing the wild-type and citrullination-defective mutants of E2F1, led us to conclude that the ability of PAD4 to augment E2F1 activity and DNA binding is at least partly mediated through citrullination of E2F1. Integrating those findings with the results from this chapter led us to conclude that in a different cellular context, such as granulocytes, PAD4 could be targeting E2F1 to cytokine promoters. Indeed, microarray gene expression data obtained from U2OS cells supported a role for both E2F1 and PAD4 in the regulation of the immune response.

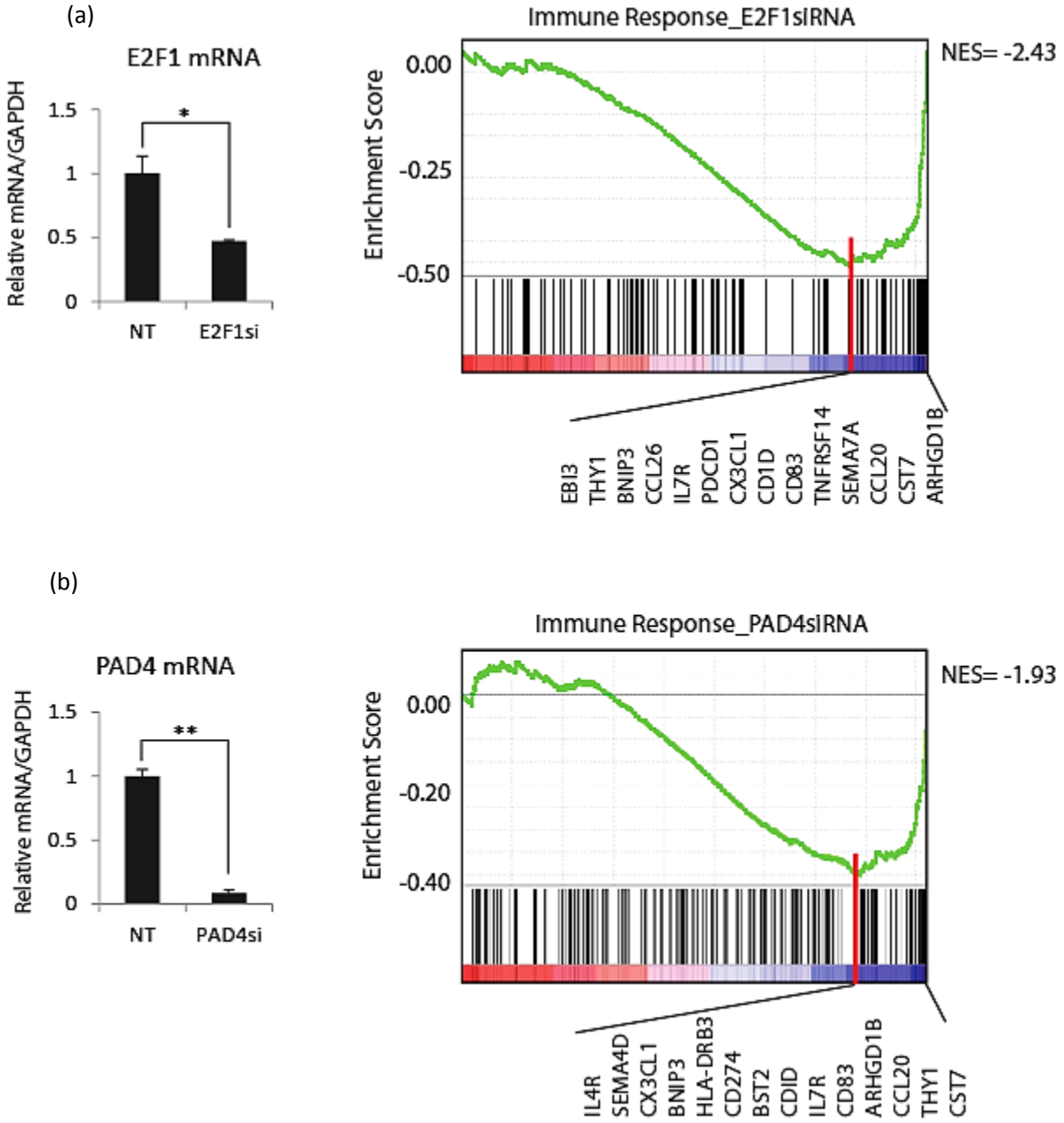


Figure 5.7) siRNA depletion of E2F1 and PAD4 in U2OS cells suppress immune response

(a) and (b) U2OS cells were transfected with 50 nM non-targeting (NT), E2F1 or PAD4 siRNA (72 h). Total RNA were extracted using Trizol reagent, and were submitted to the Welcome Trust for Human Genetics (University of Oxford). Gene expression patterns were analysed using Illumina Human HT12v4.0 gene expression array. Set of transcripts in the E2F1si and PAD4si groups were normalised to the NT group, and analysed using GSEA. The immune system was enriched and downregulated (illustrated by the green line) for both groups analysed, i.e. NT vs. E2F1si and NT vs. PAD4siRNA, with a normalised enrichment score (NES) of -2.43 and -1.93 respectively. The black bars below the green line represent the genes that contribute to the plot, with the far right bars accumulated under the valley of the green plot signifying the most

downregulated genes. P-value (Pval) <0.01 and False rate discovery rate (FDR) <0.05 are considered significant. ±S.D *p<0.05, **p<0.01 (The microarray was performed and analysed by the Wellcome Trust of Human Genetics)

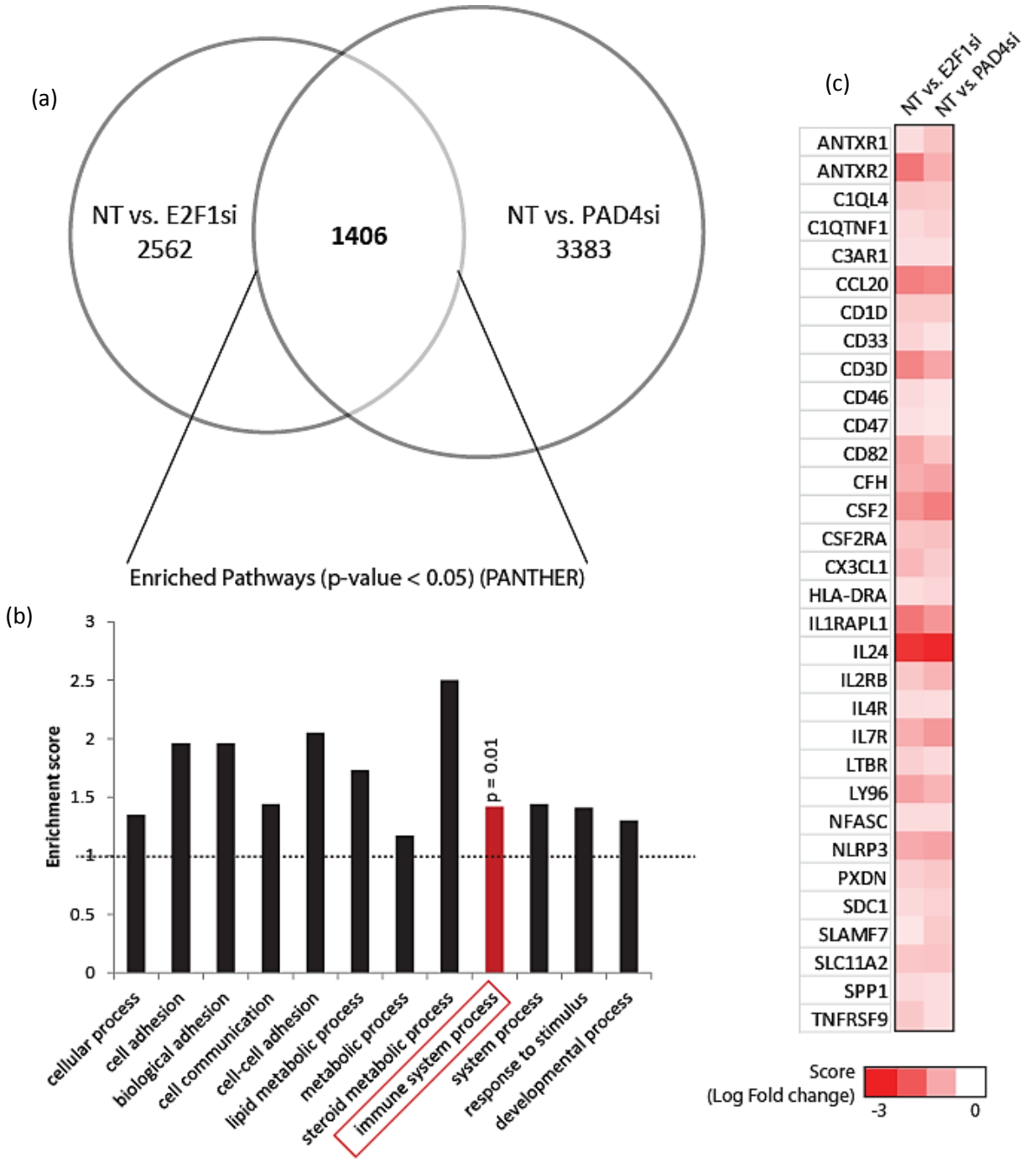


Figure 5.8) Analysis of enriched pathways which overlap between E2F1si and PAD4si treatments

(a) A Venn diagram showing the number of transcripts which are altered in both groups, and whose p-value is less than 0.05. (b) The common set of genes between the two groups was determined and analysed using PANTHER gene list classifier. The immune system process was amongst the pathways enriched. (c) List of transcripts implicated in the immune system process, which were downregulated upon depletion of either E2F1 or PAD4.

Chapter 6) Bromodomain recognition of E2F1

6.1 Introduction

In the first chapter we identified the four most probable sites of citrullination on E2F1 as R109, R111, R113 and R127. These sites lie in an intensely modified region of E2F1, as both arginine methylation and lysine acetylation have been proposed to occur in the same short stretch (Fig 6.1(a)). The close proximity of these residues and their modifications could imply potential interplay between them. We had previously shown that symmetric and asymmetric methylation marks of E2F1 on R111/R113 and R109 respectively, occur in a mutually exclusive manner, with the occurrence of one hindering the incidence of the other. In addition to such direct interplay, one can envisage a scenario where recognition of one of these marks by a reader protein becomes influenced by the presence or absence of a nearby modification.

Indeed, an important consequence of protein post-translational modifications is the subsequent downstream recognition of that mark by a dedicated 'reader' protein. Post-translational modifications can be examined in relation to a cycle of writers, readers and erasers. The 'writers' are the enzymes which establish the modification, such as arginine methyl transferases, the 'readers' are those protein domains which specifically recognise and bind a given modification, such as Tudor domain recognition of methyl-arginine, and the 'erasers' are those enzyme which remove the modification[181]. The recognition of a modification by a reader protein can have very important functional consequences for the protein modified, and as such an active endeavour for any new modification discovered is to determine the reader domain that recognises it.

To date no readers of a peptidyl citrulline mark have been identified. We therefore decided to screen E2F1 against a library of bromodomain proteins for three main reasons. Firstly, a citrullinated arginine resembles an acetylated lysine (Fig 6.1(b)). Since

bromodomains recognise acetylated lysines, by similarity we hypothesised that they may also recognise a citrullinated arginine. Secondly, although E2F1 was discovered to be acetylated on K117, K120 and K125 over a decade ago, there have been no screens to identify the bromodomain readers of these marks. Finally, a screen which identifies novel bromodomain readers of acetylated or citrullinated E2F1 could also enable us to probe the interplay between these two modifications, which as mentioned earlier are in very close proximity. We therefore, collaborated with the Structural Genomics Consortium (SGC), at Oxford University, to screen E2F1 against the library of the bromodomain family of proteins.

6.2 Bromodomain screen identified readers of acetylated E2F1

Bromodomains are protein interaction modules which recognise and bind ε-*N*-acetylated lysine residues. These domains are present in a diverse range of proteins which are broadly implicated in regulation of gene transcription and chromatin organisation. Overall, 61 different bromodomains have been described in humans, which occur in 41 predominantly nuclear proteins. These bromodomains have been ascribed to 8 main families (group I-VIII) according to structure based alignment studies (Fig 6.2), and are present in diverse cellular proteins including histone acetyltransferases, ATP-dependent chromatin-remodelling complexes, helicases, methyl-transferases, transcriptional coactivators and mediators, nuclear-scaffolding proteins, and the BET family proteins[156, 182, 183].

To perform the screen, four different E2F1 peptides, with a range of modifications, in the form of arginine residue methylation, lysine residue acetylation and arginine residue citrullination, were printed on membranes as distinct spots and incubated with

recombinant His-tagged bromodomain modules as described previously[156]. Immunoblotting the membranes with His-antibody revealed the peptides (each represented as a spot on the membrane) bound to that particular His-bromodomain. The intensities of the signals were reflective of the strength of binding between the peptide and the bromodomain (Fig 6.1(c)). The signal intensity for each peptide and each bromodomain was quantified, normalised and represented as a heatmap (Fig 6.3). Each spot on the heatmap gives information about the affinity of the given peptide for that particular bromodomain, with the most intensely red spots representing the strongest peptide-bromodomain interactions.

From the screen we note that acetylated E2F1 can be recognized by several bromodomains, with the fully acetylated peptide typically demonstrating the greatest affinity for the bromodomain (Fig 6.3, peptide c). We also note that the BET family (group II) are readers of acetylated E2F1, and particularly both bromodomains of BRD4 and BRD2 can be seen to bind acetylated E2F1 with high affinity. From the screen we also find that certain bromodomains appear to directly recognise citrullinated E2F1, particularly double citrullination on R109 and R111 (Fig 6.3, peptide a, see ATAD2, PB1, TIF1, etc). We also note that there is a complex interplay between citrullination and acetylation. A combination of two citrulline marks on a peptide generally seems to disrupt interaction of a bromodomain with multiply acetylated peptide (Fig 6.3, peptide b). However a single citrulline mark in proximity to a single acetyl mark appears to augment binding to several of the bromodomains (including BRD4) when compared to a single acetyl alone (Fig 6.3, peptide b). This is particularly notable with BRD4, which could be expected, since crystal structures have shown it to have preference for two nearby acetyl marks rather than one[184], and thus a combination of a citrulline and an acetyl could be mimicking the double acetylation pattern. Therefore it appears that the overall impact of

citrullination on bromodomain recognition of acetylated E2F1 would depend on the extent as well as exact positions of acetylation and citrullination marks respectively.

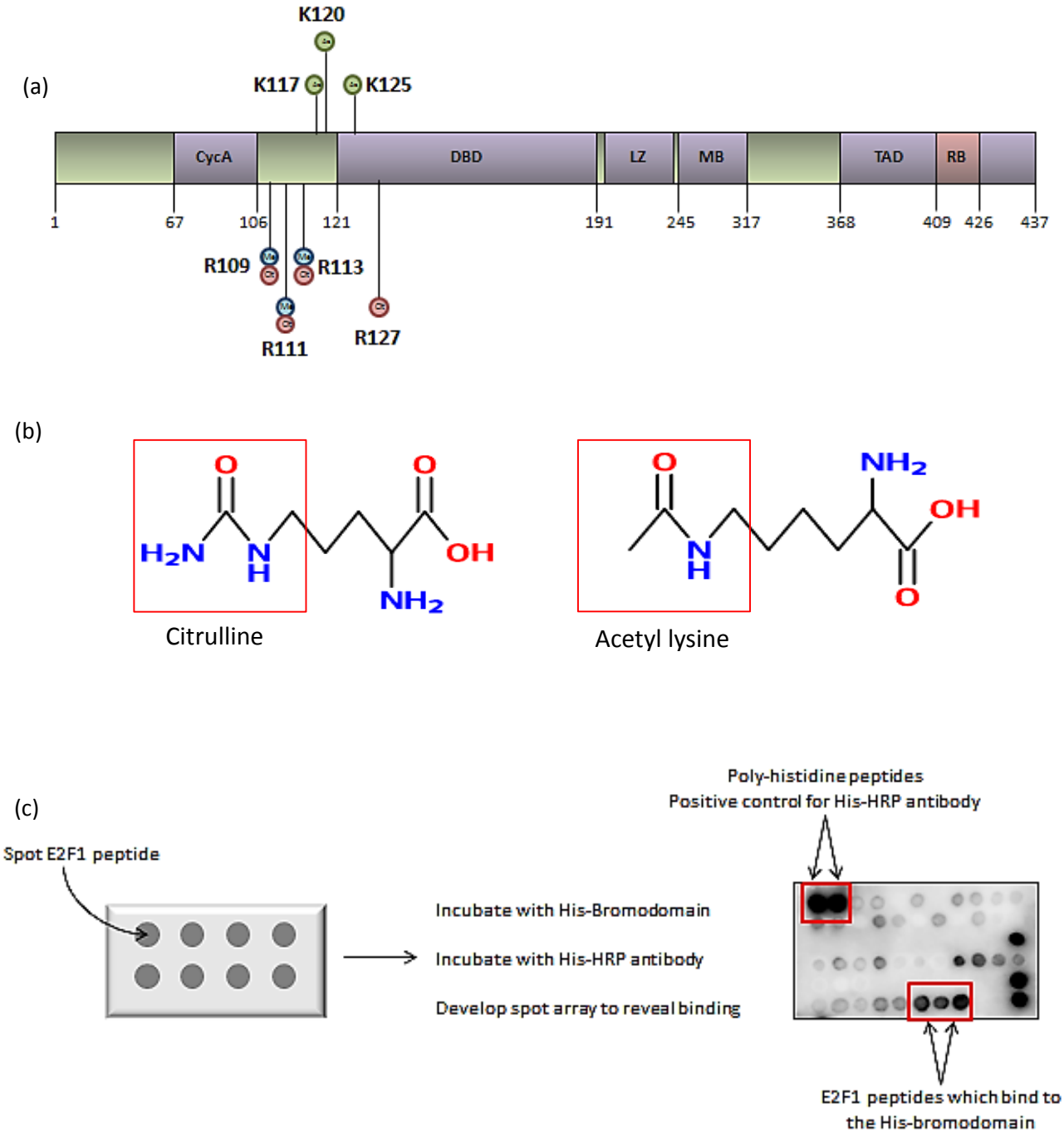


Figure 6.1) N-terminal post-translational modifications of E2F1

(a) Schematic diagram of E2F1 showing the N-terminal arginine and lysine residues which can be methylated, citrullinated and acetylated. Methylated arginine is represented as blue circle, citrullinated arginine as red, and acetyl lysine as green. (b) Structure of citrullinated arginine (left) and acetylated lysine (right). (c) Schematic diagram illustrating the process of SPOT array for screening E2F1 peptides against recombinant His-tagged bromodomains. Each spot on the membrane is composed of a single type of peptide, with defined post-translational modification(s). The membrane was incubated with a given His-bromodomain, followed by His-HRP antibody to visualise the interactions.

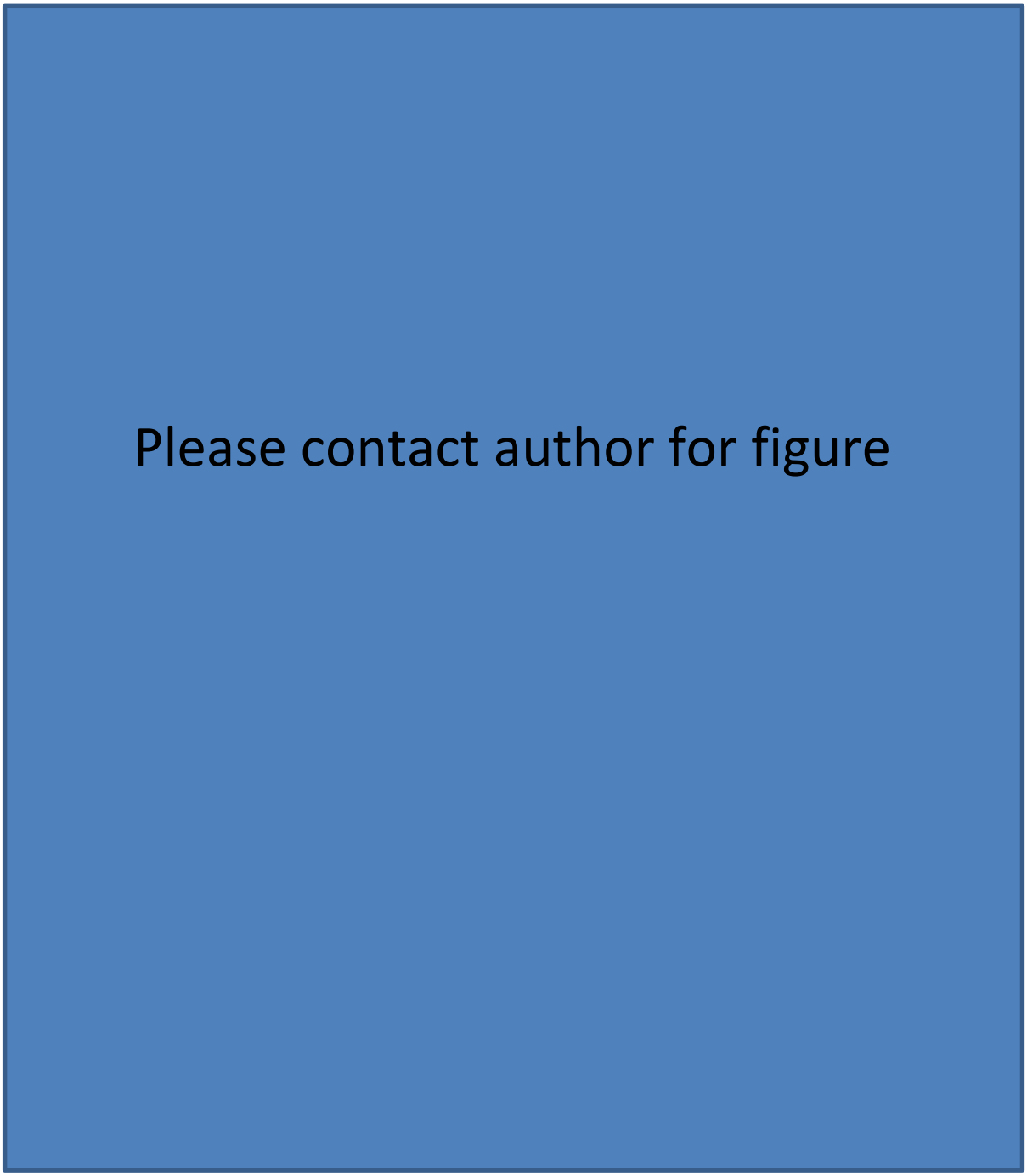


Figure 6.2) Bromodomain family of proteins

(a) Phylogenetic tree of human bromodomains, which based on structure alignments are categorised into 8 main families (group I-VIII). Diagram taken from [185]

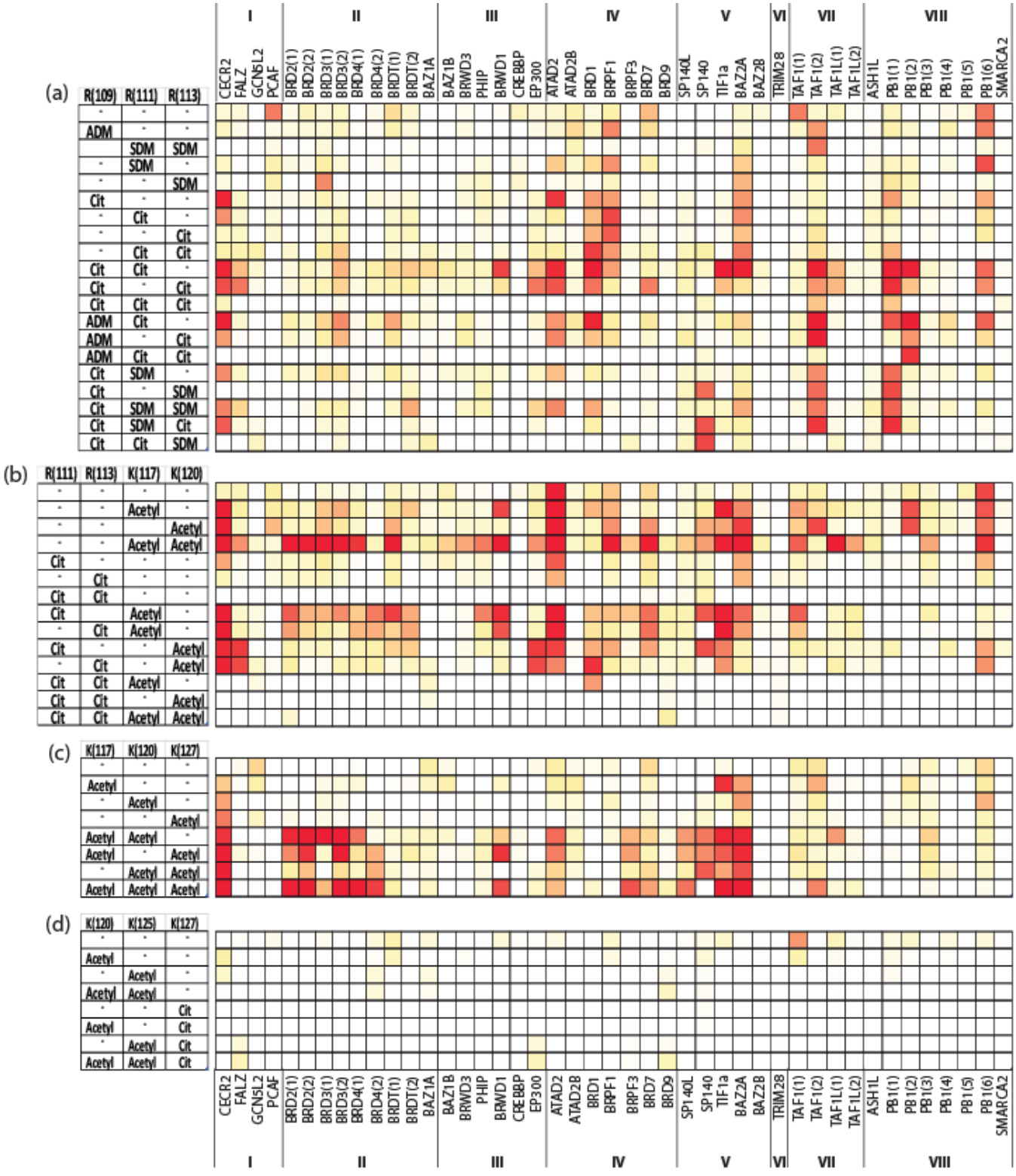

(a)
R(109) R(111) R(113)
(b)
R(111) R(113) K(117) K(120)
(c)
K(117) K(120) K(127)
(d)
K(120) K(125) K(127)
I
II
III
IV
V
VI
VII
VIII
CECR2
FALZ
GCN5L2
PCAF
BRD2(1)
BRD2(2)
BRD3(1)
BRD3(2)
BRD4(1)
BRD4(2)
BRDT(1)
BRDT(2)
BAZ1A
BAZ1B
BRWD3
PHIP
BRWD1
CREBBP
EP300
ATAD2
ATAD2B
BRD1
BRPF1
BRPF3
BRD7
BRD9
SP140L
SP140
TIF1a
BAZ2A
BAZ2B
TRIM28
TAF1(1)
TAF1(2)
TAF1L(1)
TAF1L(2)
ASH1L
PB1(1)
PB1(2)
PB1(3)
PB1(4)
PB1(5)
PB1(6)
SMARCA2
% Binding
0
50
100
(a) 105-SGPARGRGRHPGKGV-119
(b) 110-GRGRHPGKGVKSPGE-124
(c) 113-RHPGKGVKSPGEKSR-127
(d) 118-GVKSPGEKSRYETSLNL-134

Figure 6.3) Bromodomain heatmap illustrating the binding affinity of E2F1 peptides for different bromodomains

The four different E2F1 peptides are represented as sequences (a), (b), (c) and (d). For each peptide, the various possible modifications, individually and in combination were made on the array (illustrated in the tables). The relative binding of each peptide on the array to a given bromodomain was quantified and normalised relative to the intensity of a poly-histidine control spot. On the y-axis, the four peptides and their post-translational modifications are presented in tabular form. On the x-axis, the bromodomains are listed, from group I of bromodomains, to group VIII. ADM: asymmetric dimethyl arginine, SDM: symmetric dimethyl arginine, Cit: citrullinated arginine, Acetyl: acetylated lysine (The screen was performed in collaboration with the Structural Genomics Consortium with Dr Panagis Fillipakopoulos and Dr. Sarah Picaud)

6.3 E2F1 interacts with BRD4

BRD4 is a ubiquitously expressed nuclear protein, with two N-terminal tandem bromodomains. Particularly via its first bromodomain, BRD4 recognises and associates with acetylated histones, from where it is involved with chromatin organisation and transcriptional regulation [186, 187]. BRD4 has a PID domain (P-TEFb Interacting Domain) within its C-terminal region, which interacts with the Positive translation elongation factor-b (P-TEFb) and recruits it to chromatin[188]. P-TEFb, composed of the two main subunits Cdk9 and cyclin T, subsequently phosphorylates RNA Polymerase II (RNPII) on Serine 2 within its C-terminal domain (CTD). This step is deemed critical for transition of RNPII from paused to elongating mode, and hence for effective transcription [189]. As such, BRD4 has been recognised as a critical transcriptional regulator, as it controls the processivity of RNPII by recruiting P-TEFb to proximal promoter regions.

Accordingly, BRD4 has been found to play a critical role in the immune response, and expression of inducible genes, such as cytokines [190-192]. Following LPS or TNFα stimulation, BRD4 becomes recruited to promoters of NF-κB inducible genes by binding to acetylated histone marks as well as acetylated p65 subunit of NF-κB itself, from where it recruits P-TEFb to induce gene transcription. Indeed, siRNA depletion of BRD4, or treatment of cell with the BET inhibitor JQ1[193, 194], have been shown to dampen the expression of cytokines in cells challenged with inflammatory stimuli.

From the bromodomain screen we performed, we noted that both bromodomains of BRD4 can bind acetylated E2F1 peptides. Interestingly, the region of E2F1 proposed to be acetylated on the three lysine residues K117, K120 and K125, has a sequence very similar to the N-terminus of histone 4 (Fig 6.4(a)), which has been previously shown to bind BRD4 with high affinity[191]. Given that both BRD4 and E2F1 were implicated in immune

response regulation, we therefore sought to verify this interaction in cells and investigate it in relation to PAD4.

To confirm the interaction between E2F1 and the two bromodomains of BRD4 in cells, we transfected HEK293T cells with HA-E2F1 and Flag-BD1 or Flag-BD2 (bromodomains 1 and 2 of BRD4 respectively), and performed Flag agarose immunoprecipitation. Here we observed an interaction between E2F1 and both bromodomains of BRD4 (Fig 6.4(b)). Furthermore we found that this interaction also occurs between ectopic E2F1 and endogenous BRD4, and is not mediated via pRB. This is since both full length Flag-E2F1, as well as the truncated Flag-E2F1 (1-407), which lacks pRB binding domain, immunoprecipitated endogenous BRD4 (Fig 6.4(c)). An interaction between endogenous E2F1 and BRD4 was also be demonstrated in HEK293T (Fig 6.4(d)) as well as DMSO differentiated HL60 cells (Fig 6.4(e)). We therefore conclude that E2F1 can interact with BRD4 in cells.

JQ1 is a potent inhibitor of the bromodomains of the BET family, including BRD4. With the assumption that the interaction between E2F1 and BRD4 is via bromodomain recognition of acetylated E2F1, we used JQ1 in cells to examine whether inhibition of the bromodomains of BRD4 can disrupt the interaction between E2F1 and BRD4. Indeed we found that in HEK293T cells transfected with HA-E2F1, the interaction between the immunoprecipitated E2F1 and endogenous BRD4 is reduced in JQ1 versus mock-treated cells (Fig 6.5(a)). Furthermore, the HA-E2F1^{RRR} mutant, where lysine residues K117, K120 and K125 were substituted to an arginine, illustrated weaker interaction with both His-tagged BD1 (Fig 6.5(b)) and full length endogenous BRD4 (Fig 6.5(c)) when compared to wild type HA-E2F1. We therefore conclude that the interaction between E2F1 and BRD4 is at least partly mediated through bromodomain recognition of an acetylated domain in E2F1.

In conclusion, our data suggest that E2F1 can bind to BRD4 in cells, and that both bromodomains could in effect contribute to this interaction. However, we cannot rule out the contribution of other regions of BRD4 to this interaction, nor can we conclude with certainty that the interaction observed in cells is in fact a direct association, particularly as the p65 subunit of NF-κB has been shown to interact with both E2F1 and BRD4, and thus a scenario where p65 is mediating this interaction is not unimaginable.

6.4 E2F1 is in complex with BRD4 on cytokine promoters

With the observation that E2F1 and BRD4 can interact in cells, we enquired whether the two could also associate on promoters, particularly of cytokine genes. We addressed this in HL60 cells differentiated to neutrophil like cells, as these were conditions where we observed recruitment of E2F1 to cytokine promoters. As such, we performed double ChIP in DMSO treated HL60 cells challenged with LPS, and used BRD4 antibody for the primary ChIP followed by E2F1 secondary ChIP (Fig 6.6(a)). Consistent with previous reports we found that BRD4 is present on the promoters of TNFα and CCL3 genes. However we did not detect any binding to the promoter region of the E2F1 gene, whose expression we would expect to be limited in differentiated cells (as we find transcript levels to decline in differentiated cells). Furthermore, we found that E2F1 was associated with BRD4 on the promoters of TNFα and CCL3, but not E2F1 promoter itself (Fig 6.6(a)). This confirms that the two proteins can interact or be present in the same complex on certain promoters, which in the context of differentiated HL60 cells includes transcriptionally active cytokine genes.

Importantly we observed that the interaction between BRD4 and E2F1 on promoters could be disrupted with JQ1. In mock versus JQ1 treated HL60 cells (treated

with DMSO and LPS), we observed reduced association between BRD4 and E2F1 on select cytokine promoters, with either E2F1 (Fig 6.6(b)) or BRD4 (Fig 6.6(c)) immunoprecipitated in the primary ChIP. These findings complement the earlier immunoprecipitation experiments, and support the claim that E2F1 and BRD4 can associate on target gene promoters, in a bromodomain dependent manner. However we cannot distinguish between JQ1 mediated disruption of a direct interaction between E2F1 and BRD4, or the release of BRD4 from acetylated histone to indirectly disrupt the E2F1:BRD4 association. Nonetheless we can conclude that a consequence of JQ1 treatment in cells is a decrease in the extent of association between E2F1 and BRD4.

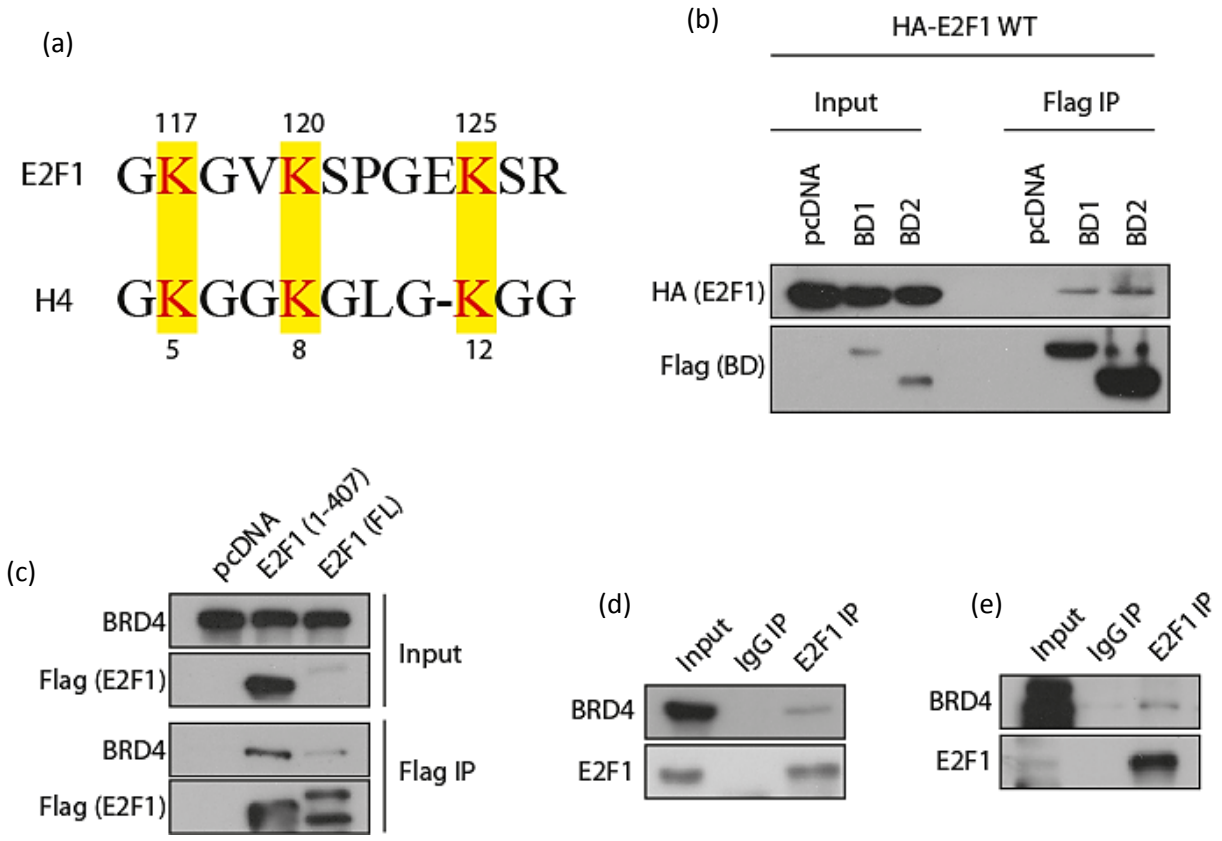


Figure 6.4) E2F1 and BRD4 interact in cells

(a) Sequence alignment of E2F1 and Histone H4, to show similarity of region proposed to bind BRD4. (b) HEK293T cells transfected with 1 µg HA-E2F1 and 2 µg Flag-BD1 or Flag-BD2, immunoprecipitated using Flag-agarose beads. (n=2) (c) HEK293T cells transfected with 1 µg Flag-E2F1 (full length, FL) or Flag-E2F1 (1-407) immunoprecipitated using Flag-agarose beads. (d) HEK293T cell lysates immunoprecipitated using E2F1 (C20) or rabbit IgG antibodies (e) DMSO differentiated HL60 cell lysates were immunoprecipitated using E2F1 (C20) or rabbit IgG antibodies.

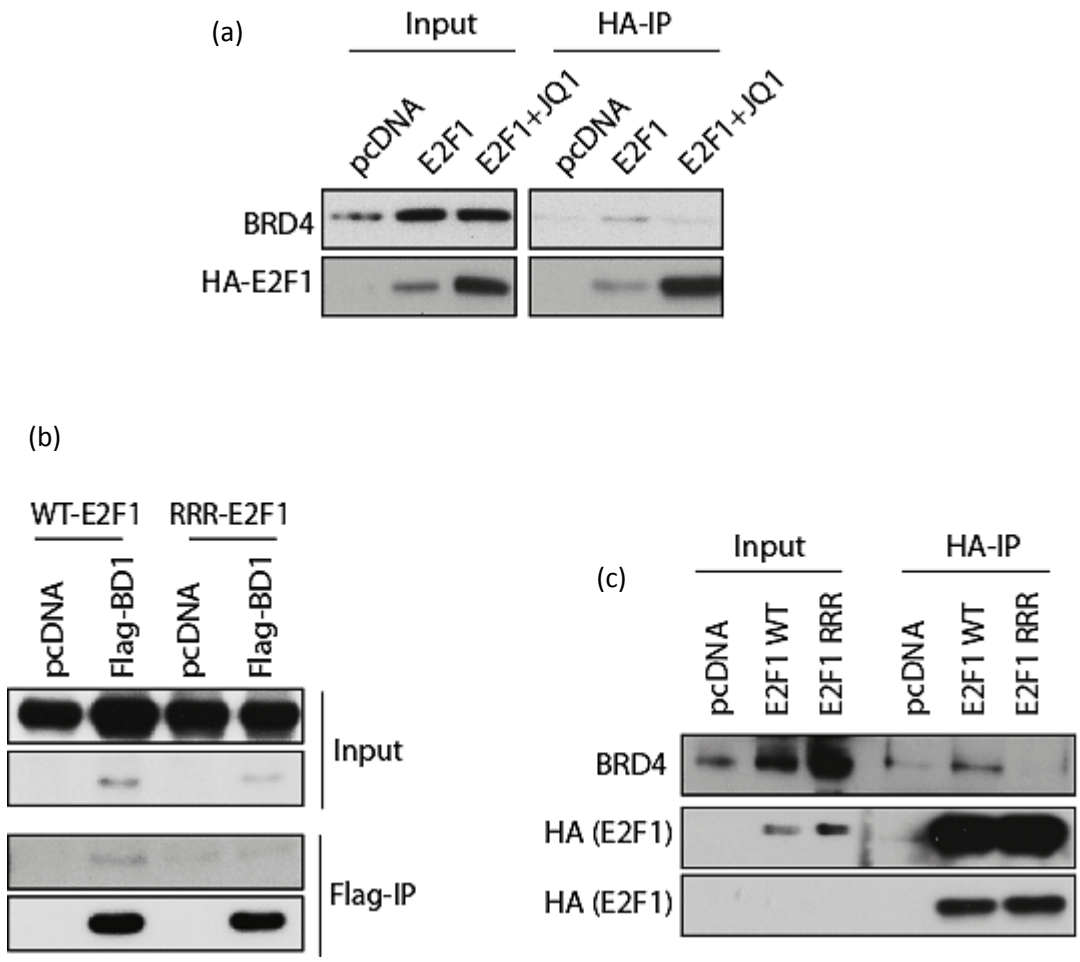


Figure 6.5) Interaction between E2F1 and BRD4 is bromodomain dependent

(a) HEK293T cells were transfected with 1 µg HA-E2F1, treated with 5 µM JQ1 and immunoprecipitated using HA-agarose beads. (n=3) (b) HEK293T cells were transfected with 1 µg HA-E2F1 WT or HA-E2F1 RRR mutant (K117R/K120R/K125R) and 2 µg Flag-BD1 and immunoprecipitated using Flag agarose beads. (c) HEK293T cells were transfected with 1 µg HA-E2F1 WT or HA-E2F1 RRR mutant (K117R/K120R/K125R) and immunoprecipitated using HA agarose beads.

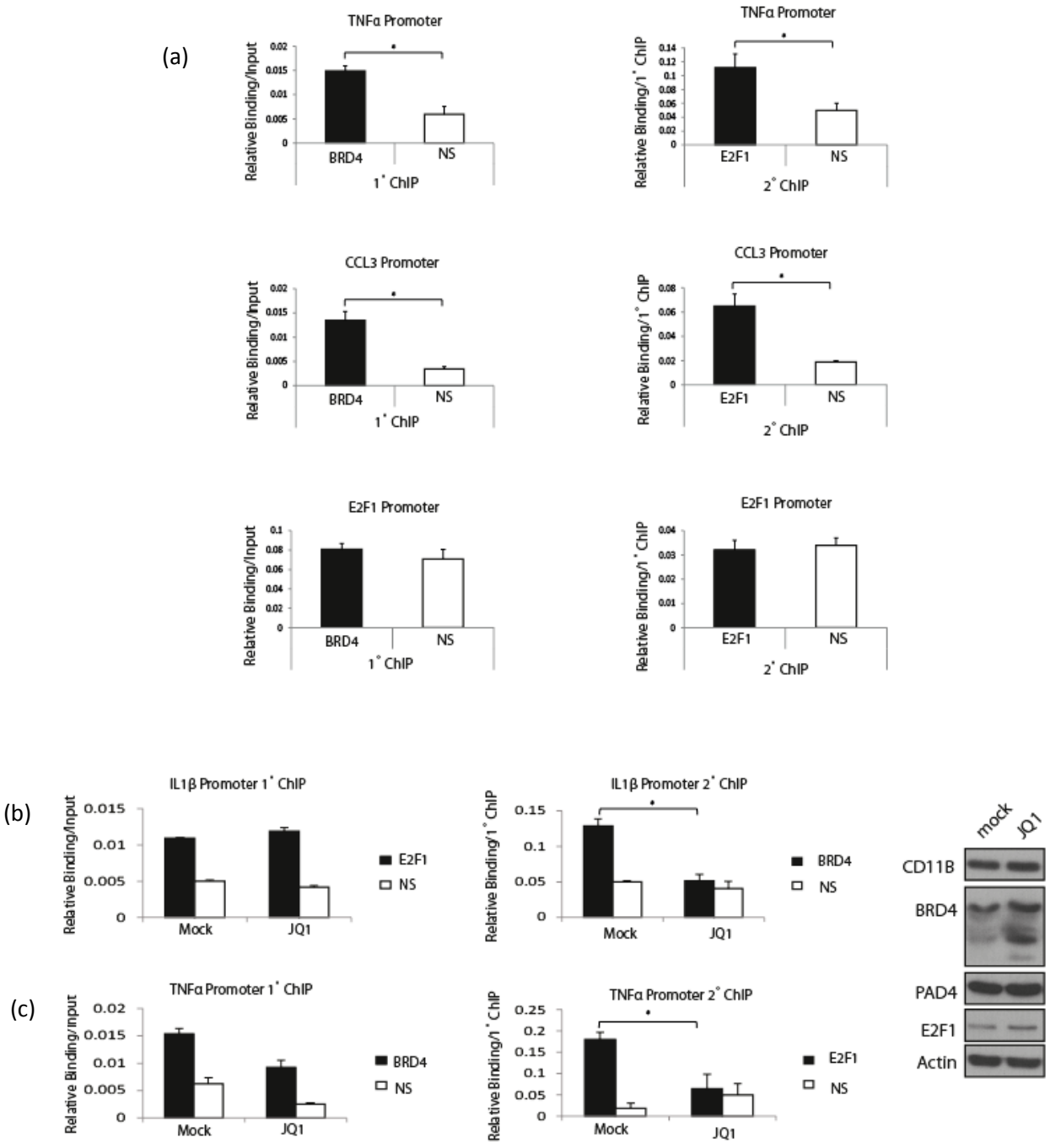


Figure 6.6) E2F1 and BRD4 interact on cytokine promoters

(a) Quantitative ChIP PCR from DMSO differentiated HL60 cells challenged with 100 ng/ml LPS for 3 hr. BRD4 antibody was used for the primary ChIP immunoprecipitation, and E2F1 (C20) antibody for the secondary ChIP. (n=3) (b) Quantitative ChIP PCR from DMSO differentiated HL60 cells challenged with 100 ng/ml LPS for 3 hr and 5 μM JQ1 for 4 h. E2F1 (C20) antibody was used for the primary ChIP and BRD4 antibody for the secondary ChIP. (n=2) (c) Experiment set up as (b), with BRD4 antibody used for primary ChIP and E2F1 (C20) antibody for secondary ChIP. (d) Blot corresponding to figures (b) and (c). ± S.D *p<0.05

6.5 PAD4 regulates the E2F1-BRD4 interaction

As mentioned earlier, an important aim of the screen was to investigate possible interplay between acetylation and citrullination. From the heat map we concluded that there is likely to be a possible interplay, depending on the extent of each modification and their relative positions. Figure 6.7(a) is taken from the heatmap, and is focused on the BET family of bromodomain containing proteins. It illustrates the complex interplay between citrullination on R111 and R113 and acetylation on K117 and K120. Generally, one can see that a single citrulline flanking a single acetyl can marginally improve binding to the bromodomain (Fig 6.7(a), black arrow), possibly by mimicking a double acetyl state. However when more than one citrulline residue is found in the vicinity of the acetylated lysine residues, the interaction with the bromodomain is lost (Fig 6.7(a), green arrow). The arginine residues flanking the double acetyl marks may be required to make electrostatic contacts with the bromodomain pocket, and thus their neutralization by citrullination could disrupt an interaction.

Reflective of the heatmap, immunoprecipitation experiments probing the effect of PAD4 on the E2F1: BRD4 interaction yielded variable results. In all three immunoprecipitation experiments in figure 6.7, HA-PAD4 is transfected into HEK293T cells, and either ectopically expressed or endogenous E2F1 are immunoprecipitated from cells, and immunoblotted for BRD4. All experiments have also been treated with sodium butyrate (to augment acetylation) and the calcium ionophore A23187 to augment citrullination. Figure 6.7(b) illustrates an example were forced expression of PAD4 does not affect the interaction between E2F1 and BRD4, whilst figure 6.7(c) and 6.7(d) represent cases where the expression of PAD4 decreases or increases the interaction between E2F1 and BRD4 respectively. With these experiments it is difficult to probe the

extent and pattern of acetylation and citrullination on E2F1, but one can envisage that the variability observed is due to the vast array of possible arrangements of these modifications. Particularly with ectopic experiments it is difficult to control for these variabilities and mimic the physiological context under which these modifications are likely to occur.

In addition to ectopically expressing PAD4 in cells, we took advantage of the broad PAD inhibitor BB-Cl-amidine to investigate its effect on the E2F1: BRD4 association. By incubating mock, JQ1 or BB-Cl-amidine treated HEK293T cell lysates with Flag-BD1, we observed that both drug treatments could reduce the interaction between E2F1 and BRD4 (Fig 6.8a). We furthermore performed double ChIP in HL60 cells, to investigate the impact of BB-Cl-amidine treatment on the E2F1: BRD4 interaction on promoters (Fig 6.8(c)). HL60 cells were differentiated with DMSO and challenged with LPS, and either treated with mock or BB-Cl-amidine. We then performed double ChIP, with BRD4 immunoprecipitated in the primary ChIP, and E2F1 in the secondary. Using the TNFα promoter as a representative example, we found that BB-Cl-amidine could reduce the levels of BRD4 on promoters, and accordingly reduce E2F1 levels associated with it. This is an interesting finding, which suggests that PAD inhibition could not only disrupt the E2F1: BRD4 association on promoters, but also release BRD4 from chromatin itself.

It is important to note however that BB-Cl-amidine also reduces E2F1 levels even at sub-therapeutic concentrations. Therefore, it is difficult to distinguish between absence of citrullination marks versus loss of E2F1 protein levels in accounting for these observations. In future experiments, pre-treating cells with MG132, to stabilize E2F1 levels, may help repeat these experiments with less variability in E2F1 protein levels.

Overall we conclude that the interaction between E2F1 and BRD4 is likely to be affected by PAD4 and by citrullination at multiple layers. The extent of each modification and their respective positioning is likely to prove decisive, although with the current tools it is difficult to address this in cells. Nonetheless we predict that given the relevance of both PAD4 and BRD4 to regulation of cytokine expression, and the likely interplay of the two via E2F1, targeting the two in cells should yield an additive anti-inflammatory effect.

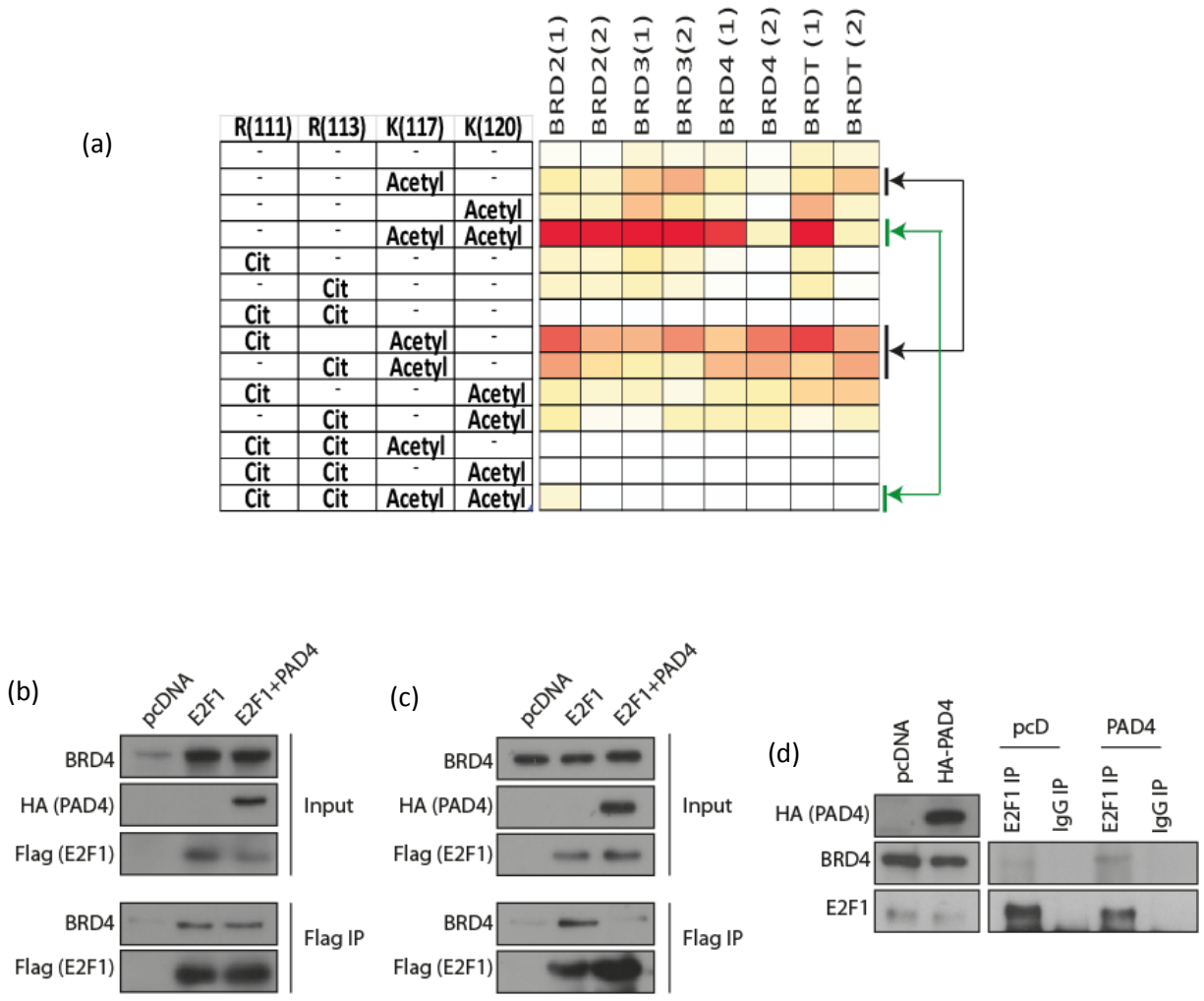


Figure 6.7) The extent of E2F1 citrullination may affect its interaction with BRD4

(a) Heatmap focused on the BET family to illustrate the complex interplay between acetylation and citrullination. The black arrow represents a single citrulline combined with an acetyl improving binding to the bromodomains, whilst the green arrow illustrates two citrulline residues flanking the double acetyl marks disrupting binding to the bromodomains. (b) and (c) HEK293T cells were transfected with 1 Flag-E2F1 and 2 µg HA-PAD4 and immunoprecipitated using Flag agarose beads. (d) HEK293T cells were transfected 2 µg HA-PAD4 and immunoprecipitated using E2F1 (C20) or rabbit IgG antibodies.

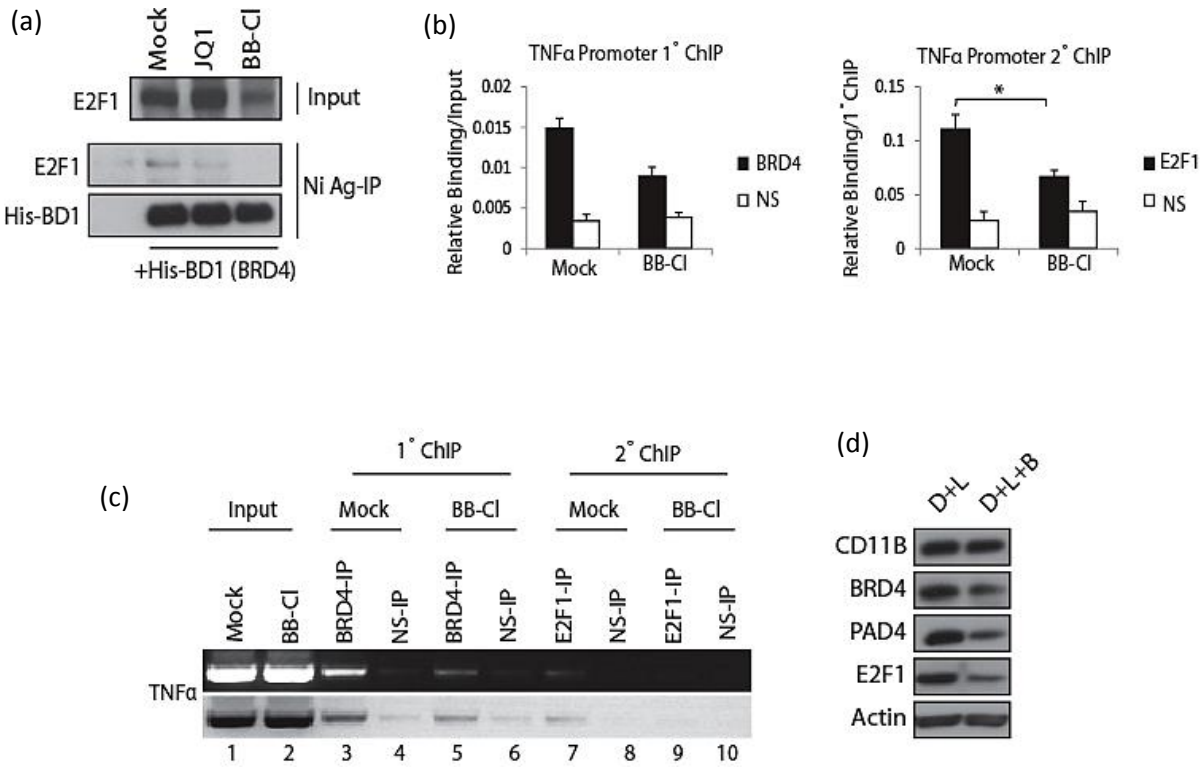


Figure 6.8) The pan-PAD inhibitor disrupts the E2F1:BRD4 association

(a) HEK293T cells were treated with 5 μM JQ1 or 5 μM BB-Cl-amidine for 16 h, incubated with His tagged BD1 (bromodomain 1) of BRD4 and subject to Ni-agarose immunoprecipitation. (n=2) (b) Quantitative and (c) non-quantitative ChIP PCR from DMSO differentiated HL60 cells challenged with 100 ng/ml LPS for 3 h and 2.5 μM BB-Cl-amidine for 16 h. BRD4 antibody was used for primary ChIP and E2F1 (C20) antibody for secondary ChIP.(n=3) (d) Blot corresponding to figures (c) and (d). NS: non-specific ChIP. ±S.D $p<0.05$

6.6 JQ1 and BB-Cl-amidine additively dampen cytokine expression

We hypothesised that co-inhibition of PAD4 and BRD4 could have a synergistic or additive anti-inflammatory effect, on the basis that both could regulate cytokine expression via E2F1. In other words we speculated that since blocking two arms of a pathway could have a stronger phenotype than either alone, targeting both PAD4 and BRD4, could more effectively limit the transcriptional activity of E2F1 on cytokine promoters. We approached this in two ways, firstly by looking at E2F1 promoter occupancy and secondly by measuring the mRNA levels of selected cytokines under single versus combination treatments. HL60 cells were differentiated with DMSO and treated with sub-toxic levels of JQ1 and BB-Cl-amidine and challenged with LPS. Relative binding of E2F1 to TNFα and CCL3 promoters was measured using quantitative ChIP PCR. We found that whilst single agent treatment with JQ1 or BB-Cl-amidine could reduce E2F1 promoter occupancy, the combination of the two drugs did not have a more pronounced additive effect (Fig 6.9(a)). However when we measured the total transcript levels of CCL3, TNFα and IL-1β using qRT-PCR, we observed that whilst both JQ1 and BB-Cl-amidine could suppress expression of these cytokines, the combination of the two compounds could reduce their expression to almost basal value. This was more pronounced with CCL3 and IL-1β than TNFα (Fig 6.9(c)).

The results obtained from measuring cytokine mRNA levels suggest that the two compounds could indeed exert a more pronounced anti-inflammatory effect. The promoter occupancy of E2F1 however did not follow the same trend, suggesting that the combination may not necessarily cause displacement of E2F1 from target promoters. In fact one could envisage that the combination of JQ1 and BB-Cl-amidine works by partly dissociating E2F1 from promoters, and partly by disrupting the association with BRD4.

Therefore at a functional level, there may be a meaningful and important interplay between PAD4 and BRD4. Given these results we were therefore interested to investigate whether similar effects are relevant *in vivo.*

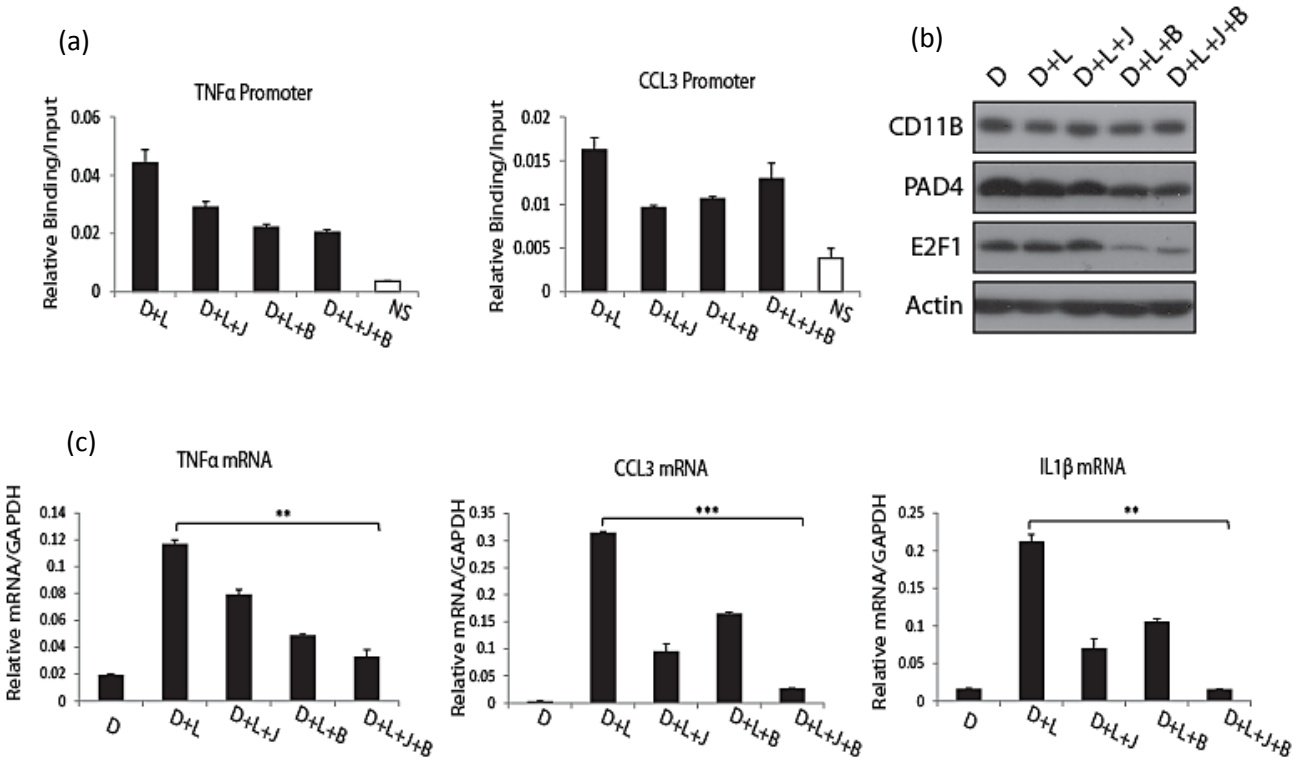


Figure 6.9) BB-Cl-amidine and JQ1 can additively suppress cytokine expression levels

(a) Quantitative ChIP PCR from DMSO differentiated HL60 cells challenged with 100 ng/ml LPS for 3 h. Where indicated cells were treated with 2.5 μM BB-Cl-amidine and/or 100 nM JQ1 for 16 h. E2F1 (C20) was used for ChIP immunoprecipitation. (n=4) (b) Immunoblot corresponding to the qPCR outlined above. (c) Quantitative real time PCR from DMSO differentiated HL60 cells challenged with 100 ng/ml LPS for 3 h. Relative mRNA normalised to housekeeping GAPDH mRNA levels. Where indicated cells were treated with 2.5 μM BB-Cl-amidine and/or 100 nM JQ1 for 16 h. (n=4) D: DMSO, L: LPS, J: JQ1, B: BB-Cl-amidine, NS: non-specific ChIP, ±S.D *p<0.05, **p<0.01, ***p<0.001

6.7 Combining JQ1 and BB-Cl-amidine in an *in vivo* model of inflammation

To understand whether the combination of JQ1 and BB-Cl-amidine is additive in an *in vivo* model of inflammation, we selected the DBA/1 mouse model of collagen induced arthritis (CIA) to investigate the combination regimen. We were interested in this model since PAD4 has been known to play a role in the development and etiology of rheumatoid arthritis (RA). Furthermore, the BB-Cl-amidine compound is a new generation PAD inhibitor which had not been tested in an *in vivo* model of RA in the past. Therefore there was interest in the community to investigate the efficacy of this new compound *in vivo*.

In order for us to establish whether the two drugs have an additive effect, it was necessary for us to combine the two drugs at their sub-therapeutic concentrations. Therefore, for the first phase of the study, a dose titration was carried out, where we tested three different doses of JQ1 and three of BB-Cl-amidine, to identify the therapeutic and sub-therapeutic doses of each compound.

Arthritis was induced in DBA/1 mice as described in chapter 2. For dosing purposes, we administered JQ1 at 5 mg/kg, 10 mg/kg or 20 mg/kg, and BB-Cl-amidine at 0.5 mg/kg, 1 mg/kg or 10 mg/kg. The mice were treated with either drug or vehicle upon the onset of their symptoms for 9 consecutive days. They were then harvested on the 10^{th} day (Fig 6.10(a)). The paws of the affected mice were measured daily, and a clinical score was assigned depending on the severity of swelling. Each paw could be assigned a score of 0-3 (0 being normal and 3 representing extreme swelling) (Fig 6.10(b)). Therefore for a given mouse, a score of 0-12 was possible.

As can be seen from the figure 6.11(a), the highest dose of BB-Cl-amidine (10 mg/kg) and the two higher doses of JQ1 (10 mg/kg and 20 mg/kg) were extremely effective at

controlling disease severity and progression (Fig 6.11(b) and (c)). High doses of JQ1 had been demonstrated in previous studies to effectively control paw swelling in CIA afflicted mice[195]. The results obtained with BB-Cl-amidine were novel.

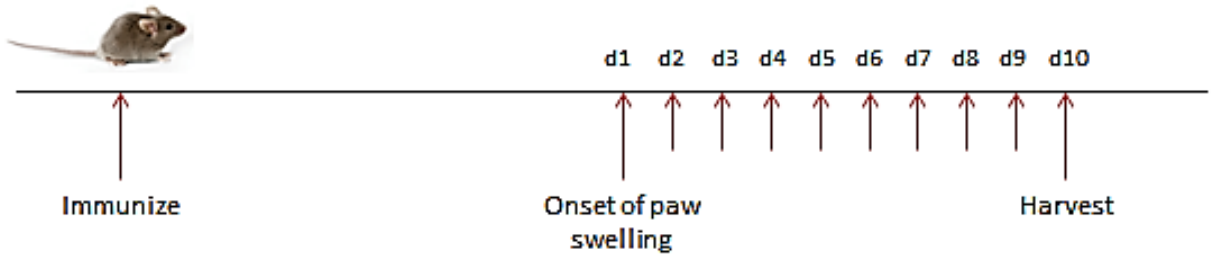
After the mice were harvested, the affected paw(s) were removed for histological examination. By applying HandE (Hematoxylin and Eosin) staining on the paws, it is possible to visualise the integrity of the joints and therefore assess whether the treatments have been effective at controlling the disease. A representative HandE stain of paw specimen from vehicle treated mouse (Fig 6.12 (a)) revealed visible bone degradation and heavy infiltration of joint spaces (indicated by the arrow). These features are typical characteristics of inflamed arthritic paw/joint[196]. In contrast, the HandE stains corresponding to arthritic mice treated with high doses of BB-Cl-amidine (10 mg/kg) (Fig 6.12(b)) or JQ1 (10 mg/kg) (Fig 6.12(c)), demonstrated well defined joint spaces, with minimal immune cell infiltration and minimal bone degradation. These histological examinations therefore confirmed the clinical scores and verified that high doses of JQ1 and BB-Cl-amidine are effective at controlling disease progression and severity in this CIA model.

Furthermore, we used these slides to stain for E2F1 protein in the affected paws with IHC (immunohistochemistry). We compared vehicle with JQ1 (10 mg/kg) or BB-Cl-amidine (10 mg/kg) treated paws (Fig 6.13). Interestingly we observed stronger E2F1 staining in vehicle treated mice (Fig 6.13(a)) compared to drug treated mice (Fig 6.13(b) and (c)). Although this may merely reflect the higher cell numbers seen in the vehicle treated paw, it nonetheless illustrates that treating these mice with an effective dose of the drug, correlates with reduced levels of E2F1 in the infiltrated paws.

For the second phase of the *in vivo* studies, we investigated the combined effect of BB-Cl-amidine at 1mg/kg with JQ1 at either 5 mg/kg or 1 mg/kg. Indeed we observed that the combination of low doses of JQ1 and BB-Cl-amidine are more effective at controlling disease severity than either compound alone (Fig 6.14(a), compare blue and purple single dose lines, with the black combination lines). This was an exciting finding, since obtaining novel and meaningful combinations of drugs is an important clinical ambition. In a disease such as cancer, this can equate to increased efficacy and reduced resistance, and in a chronic disease such RA, combinations could allow each drug to be used at a reduced dose hence minimising adverse side effects. Similar synergy was previously reported for methotrexate and anti-TNF alpha regimens[197].

Histological examinations were also performed on paws from vehicle and combination treated mice, and a score was assigned to each paw examined to reflect the integrity of joint spaces and the extent of bone degradation. An average of these scores demonstrated that the mice which had undergone combination treatment had a more favourable histological score compared to vehicle treated mice (Fig 6.14(b)). This confirmed that arthritic paw swelling was controlled with the combination regimen. We also used the spleen as a surrogate tissue, to assess the relative promoter occupancy of E2F1 on the TNFα promoter and to measure the transcript levels of TNFα. Interestingly we found that the relative enrichment of E2F1 on TNFα promoter was reduced in the combination treatment relative to vehicle (Fig 6.14(c)) and correlated with this, the relative TNFα mRNA levels were also lower (Fig 6.14(d)). Since TNFα is an important driver of RA pathology[135], these results suggest that the combination of JQ1 and BB-Cl-amidine could be an effective treatment for RA by exerting an additive anti-inflammatory effect.

(a)



Please contact author for figure

Figure 6.10) CIA protocol and scoring system

(a) Diagram depicting the CIA schedule. Mice were immunised at 8-12 weeks, and monitored until onset of paw swelling. Once paw swelling was detected (d1), the affected mouse was treated for 9 consecutive days, and was harvested on day 10. The paw size of the affected mouse was measured daily and a clinical score was assigned depending on severity of swelling. (b) Diagram taken from[198], to illustrate the scoring system. The top paw on the left is a healthy paw with a score of 0, the top on the right is mild swelling with a score of 1, the bottom paw on the left is moderate to severe swelling with a score of 2, and the bottom paw on the right is severe swelling with a maximum score of 3.

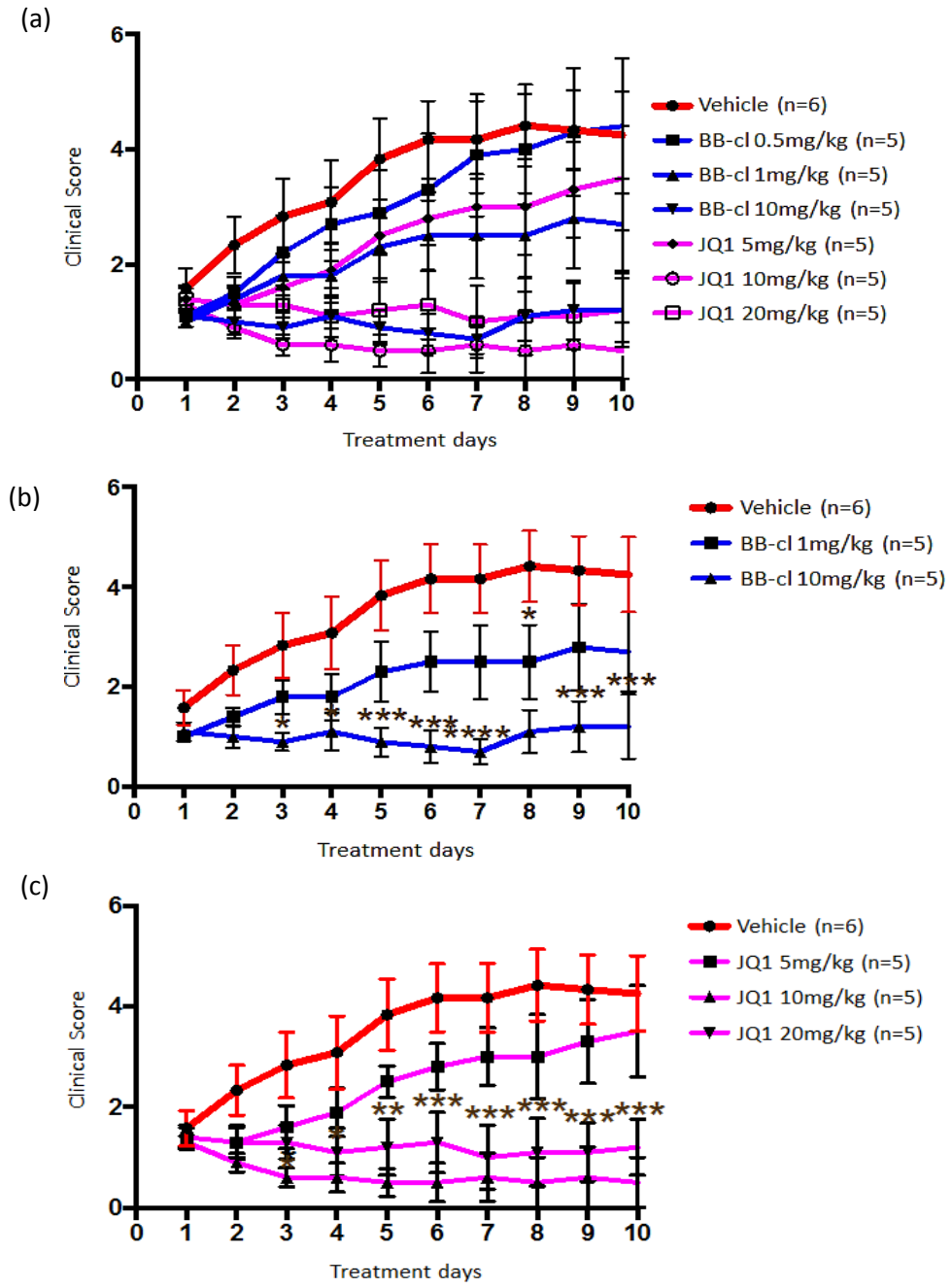


Figure 6.11) Dose responses and clinical scores obtained with titrating doses of JQ1 and BB-Cl-amidine

(a) The clinical scores obtained for each time point were averaged for a given drug. Five to six animals were treated and scored for each drug. The vehicle treated mice are represented by the red line, the three different doses of BB-Cl-amidine by the blue lines, and the three different doses of JQ1 by the purple lines. (b) A closer view of graph (a), to illustrate the efficacy of the two higher doses of BB-Cl-amidine compared to vehicle treated mice. (c) A closer view of graph (a), to illustrate the efficacy of three different doses of JQ1 compared to vehicle treated mice. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001 (treated mice vs vehicle by two-way analysis of variance (ANOVA) with the Bonferroni post hoc test) (The *in vivo* work was performed at the Kennedy Institute for Rheumatology in collaboration with Professor Patrick Venables.)

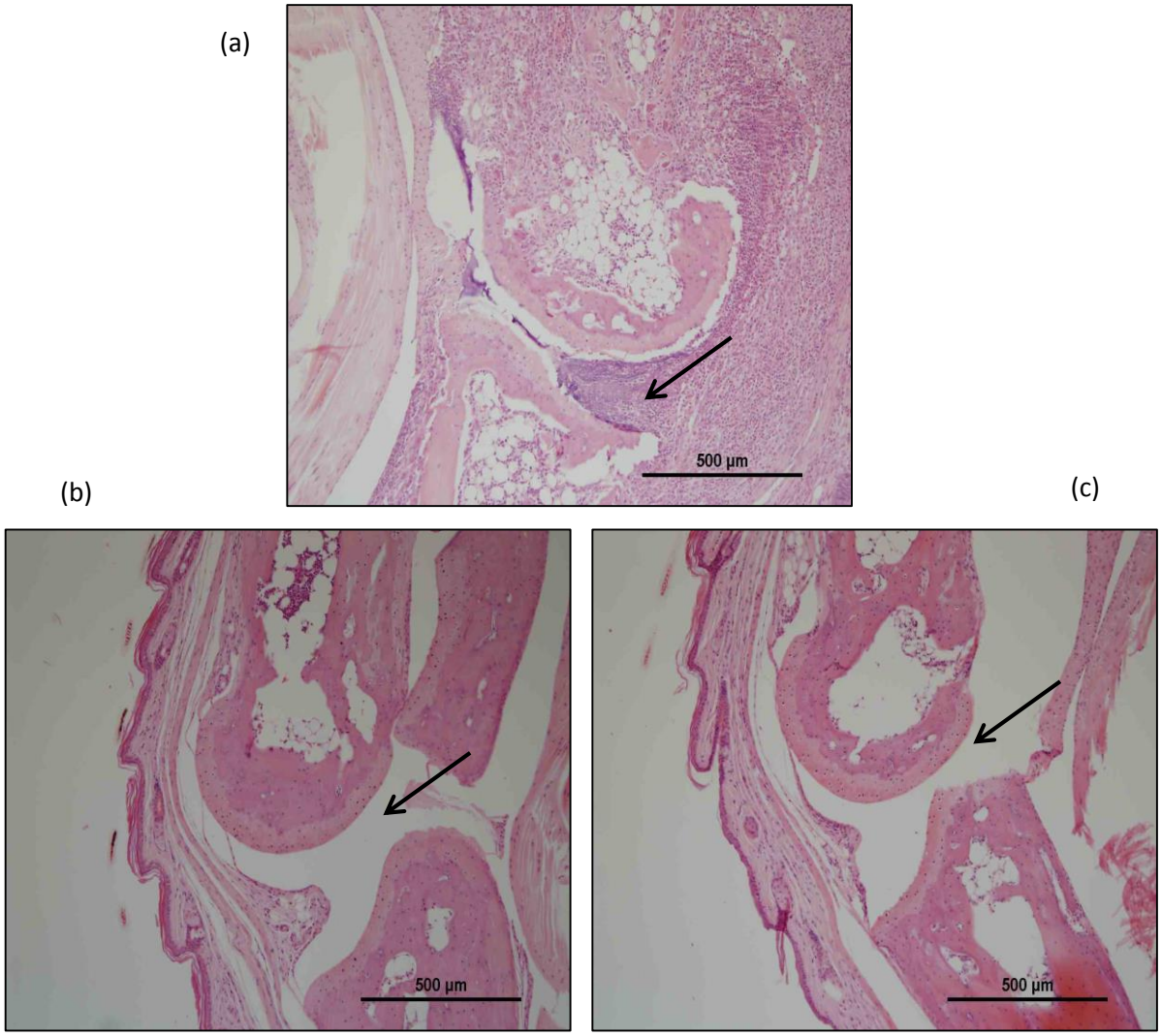


Figure 6.12) HandE staining of the affected paws

HandE staining of affected paws harvested 10 days after treatment. Representative staining slides for vehicle (a), 10 mg/kg BB-Cl-amidine (b) and 10 mg/kg JQ1 (c) treated mice are presented here. The arrows illustrate the joint space.

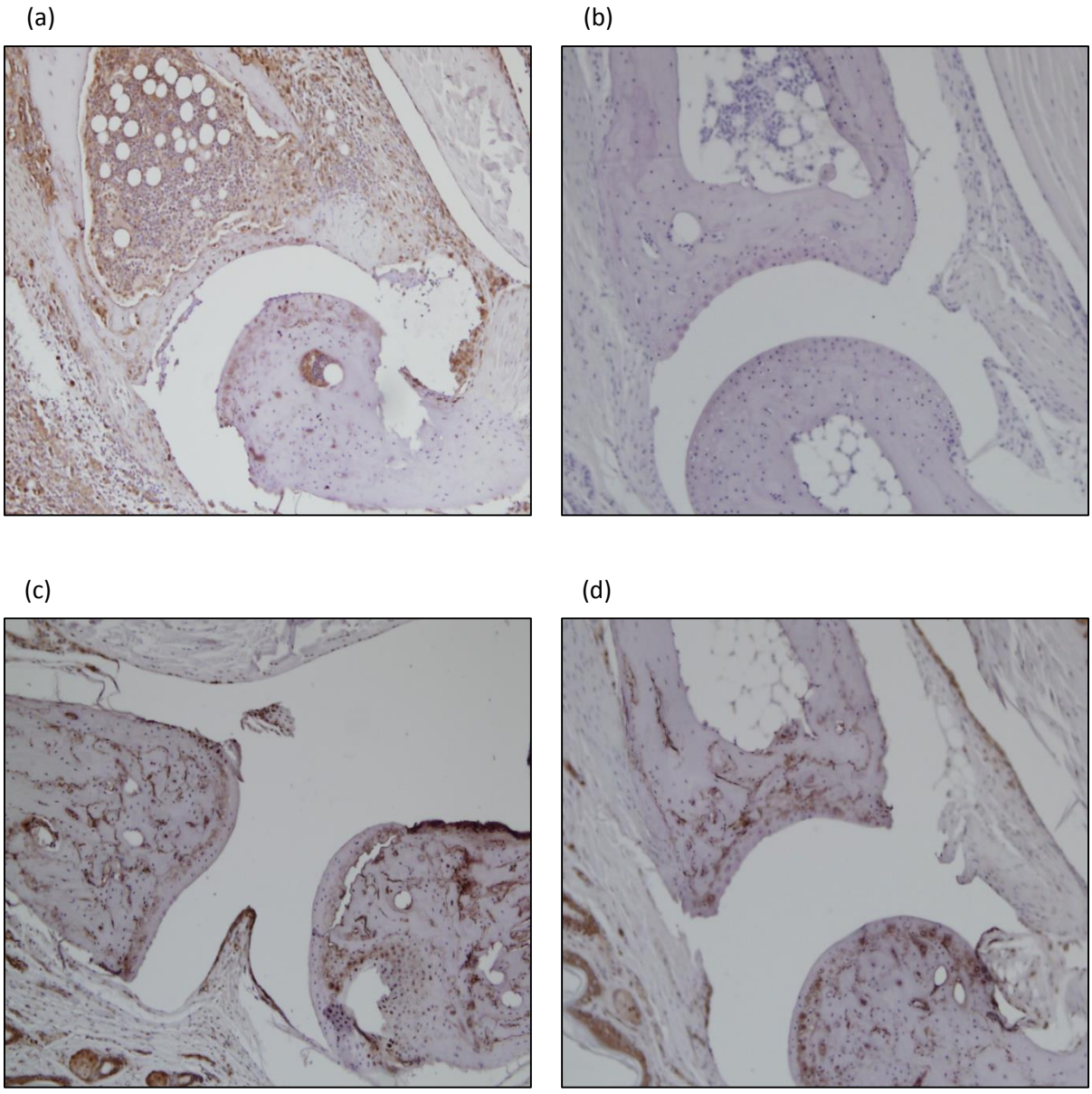


Figure 6.13) E2F1 IHC staining in the affected paws

Representative paw slides from vehicle (a), 10 mg/kg BB-Cl-amidine (c) and 10 mg/kg JQ1 (d) treated mice IHC stained for E2F1. (b) IgG negative control IHC staining in a randomised slide.

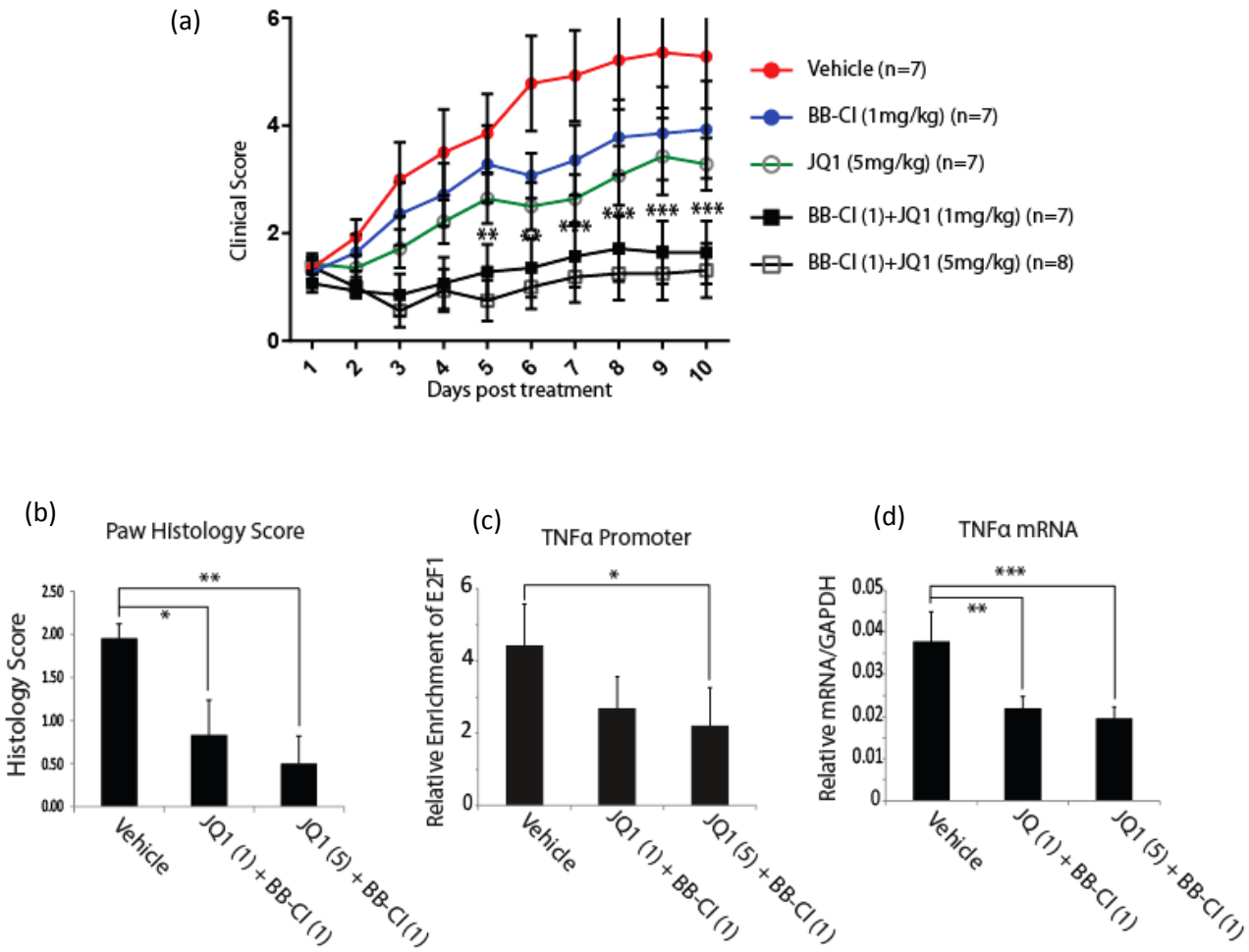


Figure 6.14) CIA clinical scores for single and combination doses of JQ1 and BB-Cl-amidine

(a) The clinical scores obtained at each time point were averaged for a given drug. Seven to eight animals were treated and scored for each drug. The vehicle treated mice are represented by the red line, the single doses of BB-Cl-amidine and JQ1 by the blue and green lines respectively, and the combination treatments by the back lines. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001 (treated mice vs vehicle by two-way analysis of variance (ANOVA) with the Bonferroni post hoc test) (b) HandE stained paw slides were scored from 0 to 3 reflective of the extent of bone degradation and immune cell infiltration within the joints. The average scores presented for vehicle and combination treatments. (c) Quantitative ChIP PCR from spleen harvested from mice. E2F1 (C20) antibody ChIP was normalised to IgG antibody ChIP to assess relative promoter enrichment of E2F1. (d) Quantitative real time PCR from spleen harvested from mice. Relative TNFα mRNA levels normalised to housekeeping GAPDH mRNA levels. JQ1 (1): JQ1 (1 mg/kg), JQ1 (5): JQ1 (5 mg/kg), BB-Cl (1): BB-Cl-amidine (1 mg/kg), ± S.D * p<0.05, **p<0.01, ***p<0.001

6.8 Chapter Summary

In this chapter, by employing a bromodomain screen, we identified the BET family member BRD4 as a novel interaction partner of E2F1, and provided evidence to suggest that this association may be important in the context of cytokine regulation. We found that PAD4 could regulate the association between E2F1 and BRD4 in complex and subtle ways, but one which could translate into an additive anti-inflammatory effect of PAD4 and BRD4 inhibitors. By exploring the *in vivo* efficacy of the PAD inhibitor BB-Cl-amidine and BET inhibitor JQ1, we obtained preliminary evidence to suggest that this combination may be a novel and effective approach for controlling RA.

Chapter 7) Discussion

7.1 Summary of Findings

In this study we demonstrated that E2F1 is a novel substrate for PAD4, and used mass spectrometry and site directed mutagenesis to identify the residues targeted by PAD4 to arginine residues 109, 111, 113 and 127. Using the U2OS osteosarcoma cancer cell line as our model system, we illustrated that PAD4 can increase the transcriptional activity, DNA binding affinity and stability of E2F1. Consistent with these findings, we found that PAD4 can displace pRB from E2F1 target gene promoters and reduce the interaction between cyclin A and E2F1. I*n vitro* binding assays in future experiments could help establish whether these observations are a direct consequence of E2F1 citrullination. To investigate the impact of citrullination under more physiological conditions, we took advantage of HL60 myeloid leukaemia cells, which express PAD4 as they differentiate into granulocytes. Although differentiation of these cells (as expected) was accompanied by downregulation of many classical E2F1 cell cycle target genes, we observed PAD4-dependent recruitment of E2F1 to various cytokine genes promoters. We therefore propose that PAD4 is a transcriptional co-activator of E2F1 and that it regulates the immune response by citrullinating E2F1.

Pertinent to any post-translation modification is identifying the reader protein that recognises it, as this provides invaluable insight into how that given modification could regulate the activity of its target. Given the similarity between a citrulline and an acetyl moiety, and furthermore the close proximity of citrullination and acetylation on E2F1, we screened E2F1 against a library of bromodomains. Various bromodomains were found to bind to acetylated E2F1 peptides, with a certain degree of interplay and cross-talk with nearby citrullination residues. Importantly, we identified BRD4 as a reader of acetylated E2F1, which was noteworthy, as BRD4 had been extensively implicated in the

transcriptional activation of cytokine genes [191]. We therefore speculated that BRD4 could regulate/mediate E2F1's ability to control the expression of cytokines, and that PAD4 could impact on this.

Accordingly, we demonstrated a specific interaction between E2F1 and BRD4 in cells, and found that both JQ1 and BB-Cl-amidine could loosen this association. Prompted by these findings we explored the combined immunosuppressive effects of PAD4 and BRD4 inhibition, with the rational that if both PAD4 and BRD4 target E2F1 to cytokine promoters, or regulate its transcriptional activity, then inhibiting both should have a more pronounced impact on cytokine expression than either alone. Supportive of our prediction, in an *in vivo* model of inflammation, the combination of sub-therapeutic doses of JQ1 and BB-Cl-amidine were found to be more efficacious than either drug alone. From the regulation of E2F1 by PAD4 and its interaction with BRD4, we therefore arrived at a novel and clinically effective combination of drugs, which we believe could be a valuable treatment for RA.

7.2 E2F1-PAD4 axis in cell proliferation versus immune response regulation

As mentioned previously, PAD4 has been implicated in various pathways pertaining to either cancer or immunity. Interestingly, the same holds true for BRD4, with the BET inhibitor JQ1 shown to be effective not only in autoimmune and inflammatory disorders, but also in certain cancer types [193]. E2F1 has largely been associated with cancer, either as an oncogene or tumour suppressor, though there have been several reports implying a role for E2F1 in immune response regulation [81].

The ability of PAD4 to target E2F1 to cell cycle genes and increase its transcriptional activity is consistent with a potentially oncogenic role for PAD4. Having said so, the E2F-PAD4 axis was not explored under conditions of cellular stress and DNA damage, and it is possible that PAD4 may also modulate the pro-apoptotic activity of E2F1. Methylation of E2F1 on R109 by PRMT1 is known to target it to apoptotic target genes, and is important for E2F1 mediated cell death. Therefore one may predict that citrullination of E2F1 on R109 under DNA damage, could hinder subsequent methylation, and therefore suppress its apoptotic potential. Validation of this prediction, coupled with the proliferative effects of PAD4 would likely support an oncogenic role for this enzyme in the context of cancer. This would also offer some mechanistic insight into the cytostatic and cytotoxic effects of PAD inhibitors in various tumour cell lines [110, 155, 199].

In HL60 cells differentiated into granulocytes, despite downregulation of cell cycle genes, including E2F1 transcript itself, we observed PAD4-dependent recruitment of E2F1 to cytokine and chemokine genes. Previously, PAD4 had been illustrated to regulate the expression of various cytokines through displacing HP1 from target promoters. Our study suggested that PAD4 may regulate expression of these target genes via a novel mechanism, i.e. through targeting E2F1 to their promoters. In a study by C.A Lim[81], E2F1 was found to become enriched on various cytokine and chemokine gene promoters in a LPS-responsive and NF-κβ dependent manner [81]. siRNA depletion of E2F1 could be shown to effectively suppress the expression of these genes, amongst them CCL3, IL-1β and TNFα. Furthermore, $E2F1^{-/-}$ mice challenged with LPS were found to have a dampened immune response, with lower circulating levels of TNFα, IL-1β and IL-6. Macrophages depleted from E2F1 also failed to produce these cytokines effectively [200-202]. Collectively, these results suggest that E2F1 plays an important role in modulating the

immune response, and from our studies we suggest that PAD4 mediated citrullination helps to regulate this.

It is important to note however, that PAD4$^{-/-}$ mice challenged with LPS do not appear to phenocopy the E2F1$^{-/-}$ mice despite our suggestion that PAD4 acts upstream of E2F1 to target it to pro-inflammatory genes. Compared to wild type mice, PAD4$^{-/-}$ mice with LPS induced septic shock had similar circulating levels of the pro-inflammatory cytokine IL-6, although interestingly the levels of the anti-inflammatory and immunomodulatory IL-10 were found to be higher [203]. The higher IL-10 levels in PAD$^{-/-}$ mice were suggested to be protective against septic shock. These results suggest that *in vivo*, PAD4 does not promote inflammation, but modulates the immune response, and therefore the relevance of PAD4 for immune response regulation appears more intricate and complex than appreciated. Further studies are needed to understand the role of PAD4 in immune response regulation *in vivo*. It will be particularly interesting to investigate how PAD4 controls the expression of pro- and anti-inflammatory cytokines, and whether these effects are cell type dependent.

The ability of PAD4 to regulate cancer cell proliferation and inflammatory response via E2F1 raises an important implication, and that is the link between inflammation and cancer. Chronic inflammation has been long associated with cancer. Long standing inflammation, such as colitis, or inflammation secondary to infection or irritation, such as H.pylori infection or smoking, have been shown to promote and support cancer [103, 204]. Among the many mechanisms, infiltrating macrophages and granulocytes are known to facilitate neoplastic growth. This is primarily through generating reactive oxygen species that damage cellular DNA and cause genomic instability, producing proangiogenic and proliferative cytokines that support cell growth and tumour vascularization, and by releasing extracellular matrix modifying enzymes that facilitate tumour invasion and

metastasis. Therefore, innate and adaptive elements of the immune system, in the long term can support tumorigenesis and the acquisition of many cancer hallmarks [3]. The ability of PAD4 to augment inflammation via E2F1, and promote cellular proliferation, further reinforces this protein as potentially oncogenic. On one hand, in neutrophils and macrophages, PAD4 can promote the expression and secretion of various cytokines, which favour neoplastic growth and survival, and on the other hand in cancer cells it helps modulate the cell cycle machinery via the E2F network. NF-κβ mediated signalling has been shown to be an important pathway linking inflammation and cancer [109], and in this study we propose a similar role for the E2F1-PAD4 pathway.

7.3 E2F1 recognition by BRD4

BRD4 has been shown to regulate the expression of various cytokine genes [191]. In cells challenged with inflammatory stimuli, BRD4 has been shown to redistribute from enhancer elements of resting-state housekeeping genes to promoters of inflammatory genes by NF-κβ [205, 206]. On these promoters, BRD4 recruits p-TEFb (positive translation elongation factor b), which phosphorylates the C-terminal domain of RNA polymerase II to enable gene transcription [186, 188, 189]. BRD4 has been reported to interact with acetylated histones on these promoters (particularly acetylated H4K5/8/12) [191], as well as with the RelA subunit of NF-κβ acetylated on lysine 310 [207, 208]. In this study we propose that E2F1 could also interact with BRD4 on these promoter elements, though it remains to be established whether E2F1 would act in concert with NF-κβ to recruit BRD4 from resting state enhancer regions to inflammatory genes. In a previous study it was shown that NF-κβ recruits E2F1 to cytokine gene promoters, such that siRNA depletion of E2F1 could downregulate the expression of several of these genes, although the promoter recruitment of NF-κβ was unaffected [81]. Therefore, E2F1 may be required to assist with the

recruitment of BRD4 and the formation of an active transcriptional complex on these gene elements. In future studies it will be interesting to investigate whether siRNA depletion of E2F1 could compromise the recruitment of BRD4 to cytokine promoters, and whether it could affect the association between BRD4 and NF-κβ.

Importantly in this study we suggest that the interaction between E2F1 and BRD4 can be regulated by PAD4. In this study, we propose that a single acetyl-lysine residue in proximity to a single citrulline residue on E2F1 may mimic a diacetyl state, and become recognised by BRD4. Therefore, where E2F1 is minimally acetylated and citrullinated, inhibiting PAD4 activity and removing the citrulline mark could disrupt the association between E2F1 and BRD4. In this study we did not investigate the acetylation status of E2F1 on cytokine promoters, and it remains to be established how acetylation versus citrullination defective mutants associate with BRD4. It will be particularly interesting to establish whether a singly acetylated lysine is the predominant acetylated form of E2F1 on cytokine promoters. Histone acetyl transferases including p300/CBP and P/CAF have been shown to be recruited to cytokine promoters [209] in activated immune cells. It is therefore conceivable that these HATs could also target E2F1 for acetylation on these promoters, which coupled with PAD4 mediated citrullination, could then facilitate the recruitment of BRD4.

Future studies will also need to address the impact of histone citrullination and their cross-talk with acetylation for gene transcription. As previously mentioned, the citrullination of H3R8 has been shown to weaken HP1α binding to the adjacent trimethylated H3K9 residue on target promoters [145]. From our heatmap analysis (Fig 6.3, peptide b, line 14) we conclude loss of binding of bromodomains when multiple citrulline residues were flanking the acetylated lysine residues. By analogy, one may therefore argue that citrullination of histones may also prevent BRD4 from binding to acetylated core

histones, and that PAD4 could in fact act as a transcriptional repressor. Based on this proposition, and our observations regarding E2F1 citrullination, it could be said that the cellular effect of PAD4 are largely dependent on the extent of citrullination, with this in turn influenced by the total protein levels of PAD4 and the cellular availability of calcium ions.

With excess calcium availability, such as stimulated neutrophils undergoing NETosis or dying immune cells in inflamed arthritic joints, PAD4 may be seen to hypercitrullinate its substrates. In relation to transcription factors this could unfold their structure, and disrupt all interactions, and in relation to histones, this could lead to chromatin opening and complete loss of transcription. These conditions as previously mentioned coincide with the ability of PAD4 to regulate NETosis. Under more limiting calcium levels however, PAD4 may target its substrates more selectively in a manner that is conducive to precise transcriptional regulation. In HL60 cells undergoing differentiation with DMSO, one can see gradual accumulation of PAD4 levels, with day 6 coinciding with maximal PAD4 levels yet loss of E2F1 (Fig 5.1(b), day6). During this time course, reflective of its expression levels, PAD4 may be destined to different cellular functions. In other words, during earlier time points, PAD4 may be important for fine-tuning gene transcription outcomes, but in later time points it may be required to mediate cell death and NETosis. There are very few studies which have investigated the upstream regulators of PAD4. For future studies it will be interesting to explore the activity and functional capacity of PAD4 in relation to local and cellular levels of calcium, and how this integrates with the cellular phenotype.

7.4 Clinical significance of interplay between PAD4 and BRD4

From the interplay and cross-talk between citrullination and acetylation, and furthermore the ability of both PAD4 and BRD4 to regulate the transcriptional activity of E2F1 on cytokine promoters, we predicted a synergy between PAD4 and BRD4 inhibitors in targeting inflammation. Indeed in an *in vivo* model of RA we demonstrated that the combined sub-therapeutic doses of BB-Cl-amidine and JQ1 are more effective than single doses. This is a significant finding, even though we did not unequivocally show E2F1 to be the mediator of this effect.

Long-term use of immune suppressive drugs in chronic conditions such as RA raises real concerns regarding their safety and side effects. This is particularly true of targeting BRD4, as it is ubiquitously expressed in the body, and serves many important functions, including cell cycle control [210]. Therefore a synergistic drug combination, which allows one to reduce the dose of individual components whilst achieving therapeutic efficacy is of considerable value. JQ1 had previously been shown to work in synergy with other drugs, particularly for the management of cancer [211]. Use of complementary drug combinations is especially important in cancer as it helps avoid or overcome drug resistance, whereas for a chronic inflammatory disease such as RA, limiting side effects is probably of greater clinical priority. Mechanistically we speculate that regulation of the PAD4-E2F1-BRD4 axis would (partly) account for the efficacy of the drug combination, though it would be interesting to assess how $E2F1^{-/-}$ mice would respond in future studies.

As mentioned previously, arthritis in humans is characterised by high titres of anti-citrulline protein antibodies (ACPA) in the serum, and autoimmunity to citrullinated autoantigens is considered a driver of arthritis. Therefore it is conceivable that blocking PAD activity in this context may ameliorate disease intensity and progression. However,

according to previous reports, circulating ACPAs do not feature in collagen induced arthritis in DBA mice [212], which were also used in our study. Nonetheless we found that at high doses, and in combination with JQ1, BB-Cl-amidine can very effectively control and even reverse arthritic paw swelling in the disease model. This therefore suggests that PAD4 (and PAD2) contribute to inflammation and autoimmunity in ways other than simply creating citrullinated autoantigens. It also supports our premise that PAD4 can regulate the immune response in more intricate ways involving transcription factor signalling networks.

7.5 Conclusion

In conclusion we demonstrate that E2F1 is a novel substrate for PAD4. In relation to cancer cell proliferation we suggest that PAD4 can augment the proliferative capacity of E2F1 by increasing its activity and DNA binding affinity. Crucially, with respect to immune response regulation, we suggest that PAD4 can target E2F1 to cytokine promoters and modulate the immune response via the E2F1 signalling network (Fig 7.1). We find that E2F1 can interact with the BET bromodomain BRD4 on cytokine promoters, and that PAD4 could regulate this association. Accordingly we arrive at a novel combination of two drugs, namely PAD and BRD4 inhibitors, for the treatment of inflammatory disorders and illustrate this combination to be effective in an *in vivo* model of RA.

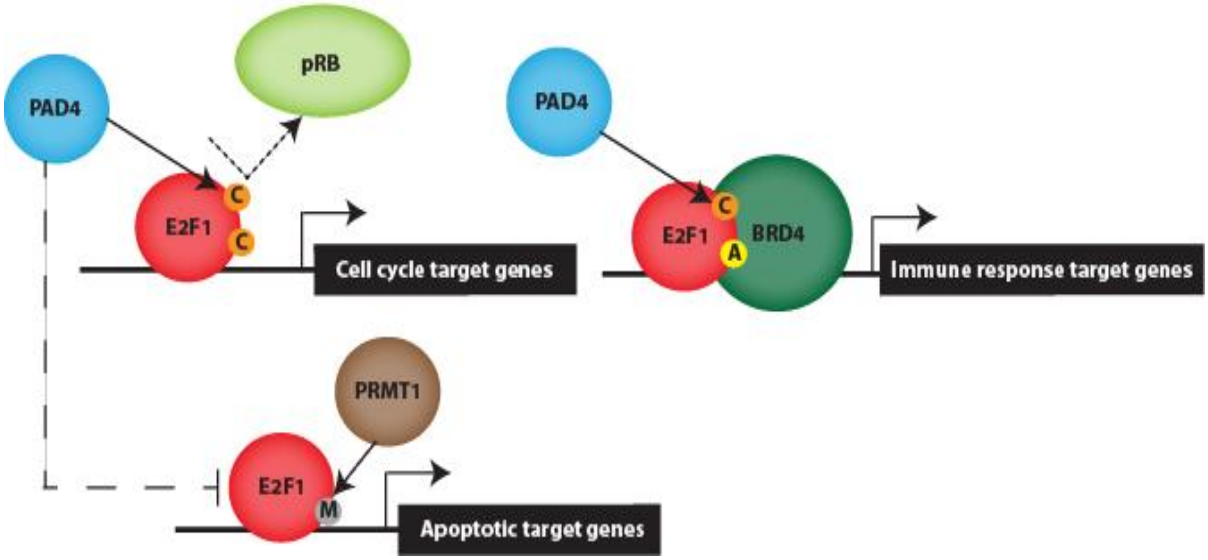


Figure 7.1) Model: Regulation of E2F1 by PAD4 mediated citrullination

During proliferation we propose that PAD4 citrullinates E2F1 to increase its transcriptional activity and binding to several target gene promoters. This is associated with release of pRB from target genes and cell cycle progression. We also speculate (hence the dashed lines) that PAD4 could block methylation of E2F1 by PRMT1 by targeting the same residues for citrullination, and accordingly could suppress E2F1 mediated apoptosis. In immune cells on the other hand E2F1 is recruited to cytokine promoters in a PAD4 dependent manner, where it interacts with BRD4 to facilitate the expression of inflammatory genes.

References

References

1. WHO | Cancer. *WHO* (2015).
2. Hanahan, D. & Weinberg, R.A. The hallmarks of cancer. *Cell* **100**, 57-70 (2000).
3. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).
4. Lodish, H. *et al.* Proto-Oncogenes and Tumor-Suppressor Genes. (2000).
5. You, J.S. & Jones, P.A. Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell* **22**, 9-20 (2012).
6. Crose, C.M. Oncogenes and Cancer — NEJM. (2015).
7. Little, J.B. Holland-Frei Cancer Medicine. 6th edition. Molecular Mechanisms. (2003).
8. Payne, S.R. & Kemp, C.J. Tumor suppressor genetics. *Carcinogenesis* **26**, 2031-2045 (2005).
9. Kopnin, B.P. Targets of oncogenes and tumor suppressors: key for understanding basic mechanisms of carcinogenesis. *Biochemistry (Mosc)* **65**, 2-27 (2000).
10. Vogelstein, B. & Kinzler, K.W. Cancer genes and the pathways they control. *Nature Medicine* **10**, 789-799 (2004).
11. Evan, G.I. & Vousden, K.H. Proliferation, cell cycle and apoptosis in cancer. *Nature* **411**, 342-348 (2001).
12. Lodish, H. *et al.* Overview of the Cell Cycle and Its Control. (2000).
13. Johnson, A. & Skotheim, J.M. Start and the restriction point. *Curr Opin Cell Biol* **25**, 717-723 (2013).
14. Blagosklonny, M.V. & Pardee, A.B. The restriction point of the cell cycle. *Cell Cycle* **1**, 103-110 (2002).
15. Lodish, H. *et al.* Checkpoints in Cell-Cycle Regulation. (2000).
16. Lukas, J., Lukas, C. & Bartek, J. Mammalian cell cycle checkpoints: signalling pathways and their organization in space and time. *DNA Repair (Amst)* **3**, 997-1007 (2004).
17. Kastan, M.B. & Bartek, J. Cell-cycle checkpoints and cancer. *Nature* **432**, 316-323 (2004).
18. Malumbres, M. Cyclin-dependent kinases, in *Genome Biol*, Vol. 15 122 (2014).
19. Malumbres, M. & Barbacid, M. Cell cycle kinases in cancer. *Curr Opin Genet Dev* **17**, 60-65 (2007).
20. Malumbres, M. & Barbacid, M. Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer* **9**, 153-166 (2009).
21. Murray, A.W. Recycling the cell cycle: cyclins revisited. *Cell* **116**, 221-234 (2004).
22. Pavletich, N.P. Mechanisms of cyclin-dependent kinase regulation: structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors. *J Mol Biol* **287**, 821-828 (1999).
23. Obaya, A.J. & Sedivy, J.M. Regulation of cyclin-Cdk activity in mammalian cells. *Cell Mol Life Sci* **59**, 126-142 (2002).
24. Eric A. Kelin, R.K.A. Transcriptional regulation of the cyclin D1 gene at a glance. *Journal of Cell Science*, 3853-3857 (2008).
25. Duronio, R.J. & Xiong, Y. Signaling Pathways that Control Cell Proliferation. (2013).
26. Herrera, R.E. *et al.* Altered cell cycle kinetics, gene expression, and G1 restriction point regulation in Rb-deficient fibroblasts. (1996).
27. Weinberg, R.A. The retinoblastoma protein and cell cycle control. *Cell* **81**, 323-330 (1995).
28. Yao, G., Lee, T.J., Mori, S., Nevins, J.R. & You, L. A bistable Rb-E2F switch underlies the restriction point. *Nat Cell Biol* **10**, 476-482 (2008).
29. Knudsen, E.S. & Wang, J.Y. Differential regulation of retinoblastoma protein function by specific Cdk phosphorylation sites. *J Biol Chem* **271**, 8313-8320 (1996).
30. Adams, P.D. Regulation of the retinoblastoma tumor suppressor protein by cyclin/cdks. *Biochim Biophys Acta* **1471**, M123-133 (2001).

31. Di Fiore Riccardo, D.A.A., Tesoriere Giovanni, Vento Renza RB1 in cancer: Different mechanisms of RB1 inactivation and alterations of pRb pathway in tumorigenesis. *Journal of Cellular Physiology* **228**, 1676-1687 (2015).
32. Giacinti, C. & Giordano, A. RB and cell cycle progression. *Oncogene* **25**, 5220-5227 (2006).
33. Kaelin, W.G., Jr., Ewen, M.E. & Livingston, D.M. Definition of the minimal simian virus 40 large T antigen- and adenovirus E1A-binding domain in the retinoblastoma gene product. *Mol Cell Biol* **10**, 3761-3769 (1990).
34. Bandara, L.R. & La Thangue, N.B. Adenovirus E1a prevents the retinoblastoma gene product from complexing with a cellular transcription factor. *Nature* **351**, 494-497 (1991).
35. Chellappan, S.P., Hiebert, S., Mudryj, M., Horowitz, J.M. & Nevins, J.R. The E2F transcription factor is a cellular target for the RB protein. *Cell* **65**, 1053-1061 (1991).
36. Qin, X.Q., Chittenden, T., Livingston, D.M. & Kaelin, W.G., Jr. Identification of a growth suppression domain within the retinoblastoma gene product. *Genes Dev* **6**, 953-964 (1992).
37. Dick, F.A. & Dyson, N. pRB contains an E2F1-specific binding domain that allows E2F1-induced apoptosis to be regulated separately from other E2F activities. *Mol Cell* **12**, 639-649 (2003).
38. Qin, X.Q. *et al.* The transcription factor E2F-1 is a downstream target of RB action. *Mol Cell Biol* **15**, 742-755 (1995).
39. Brehm, A. *et al.* Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature* **391**, 597-601 (1998).
40. Trouche, D., Le Chalony, C., Muchardt, C., Yaniv, M. & Kouzarides, T. RB and hbrm cooperate to repress the activation functions of E2F1. *Proc Natl Acad Sci U S A* **94**, 11268-11273 (1997).
41. Zhang, H.S. & Dean, D.C. Rb-mediated chromatin structure regulation and transcriptional repression. *Oncogene* **20**, 3134-3138 (2001).
42. Harbour, J.W., Luo, R.X., Dei Santi, A., Postigo, A.A. & Dean, D.C. Cdk phosphorylation triggers sequential intramolecular interactions that progressively block Rb functions as cells move through G1. *Cell* **98**, 859-869 (1999).
43. Stevaux, O. & Dyson, N.J. A revised picture of the E2F transcriptional network and RB function. *Curr Opin Cell Biol* **14**, 684-691 (2002).
44. Dyson, N. The regulation of E2F by pRB-family proteins. *Genes Dev* **12**, 2245-2262 (1998).
45. Ludlow, J.W., Glendening, C.L., Livingston, D.M. & DeCarprio, J.A. Specific enzymatic dephosphorylation of the retinoblastoma protein. *Mol Cell Biol* **13**, 367-372 (1993).
46. Saavedra, H.I. *et al.* Specificity of E2F1, E2F2, and E2F3 in mediating phenotypes induced by loss of Rb. *Cell Growth Differ* **13**, 215-225 (2002).
47. Dick, F.A. & Rubin, S.M. Molecular mechanisms underlying RB protein function. *Nature Reviews Molecular Cell Biology* **14**, 297-306 (2013).
48. Chen, D. *et al.* Division and apoptosis of E2f-deficient retinal progenitors. *Nature* **462**, 925-929 (2009).
49. Wang, H. *et al.* Skp2 is required for survival of aberrantly proliferating Rb1-deficient cells and for tumorigenesis in Rb1+/- mice. *Nat Genet* **42**, 83-88 (2010).
50. Bertoli, C., Skotheim, J.M. & de Bruin, R.A. Control of cell cycle transcription during G1 and S phases. *Nat Rev Mol Cell Biol* **14**, 518-528 (2013).
51. Xu, M., Sheppard, K.A., Peng, C.Y., Yee, A.S. & Piwnica-Worms, H. Cyclin A/CDK2 binds directly to E2F-1 and inhibits the DNA-binding activity of E2F-1/DP-1 by phosphorylation. *Mol Cell Biol* **14**, 8420-8431 (1994).
52. Campanero, M.R. & Flemington, E.K. Regulation of E2F through ubiquitin-proteasome-dependent degradation: stabilization by the pRB tumor suppressor protein. *Proc Natl Acad Sci U S A* **94**, 2221-2226 (1997).
53. Roworth A.P., G.F., La Thangue N.B To live or let die – complexity within the E2F1 pathway. *Molecular & Cellular Oncology* **2** (2015).

54. Chen, H.Z., Tsai, S.Y. & Leone, G. Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nat Rev Cancer* **9**, 785-797 (2009).
55. DeGregori, J. & Johnson, D.G. Distinct and Overlapping Roles for E2F Family Members in Transcription, Proliferation and Apoptosis. *Curr Mol Med* **6**, 739-748 (2006).
56. Trimarchi, J.M. & Lees, J.A. Sibling rivalry in the E2F family. *Nat Rev Mol Cell Biol* **3**, 11-20 (2002).
57. Iaquinta, P.J. & Lees, J.A. Life and death decisions by the E2F transcription factors. *Curr Opin Cell Biol* **19**, 649-657 (2007).
58. Dimova, D.K. & Dyson, N.J. The E2F transcriptional network: old acquaintances with new faces. *Oncogene* **24**, 2810-2826 (2005).
59. Takahashi, Y., Rayman, J.B. & Dynlacht, B.D. Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression. *Genes Dev* **14**, 804-816 (2000).
60. Chong, J.L. *et al.* E2f1-3 switch from activators in progenitor cells to repressors in differentiating cells. *Nature* **462**, 930-934 (2009).
61. Tsai, S.Y. *et al.* Mouse development with a single E2F activator. *Nature* **454**, 1137-1141 (2008).
62. Polanowska, J. *et al.* Human E2F5 gene is oncogenic in primary rodent cells and is amplified in human breast tumors. *Genes Chromosomes Cancer* **28**, 126-130 (2000).
63. Umemura, S. *et al.* Overexpression of E2F-5 correlates with a pathological basal phenotype and a worse clinical outcome. *Br J Cancer* **100**, 764-771 (2009).
64. Eymin, B., Gazzeri, S., Brambilla, C. & Brambilla, E. Distinct pattern of E2F1 expression in human lung tumours: E2F1 is upregulated in small cell lung carcinoma. *Oncogene* **20**, 1678-1687 (2001).
65. Reimer, D. *et al.* Expression of the E2F family of transcription factors and its clinical relevance in ovarian cancer. *Ann N Y Acad Sci* **1091**, 270-281 (2006).
66. Kovesdi, I., Reichel, R. & Nevins, J.R. Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. *Proc Natl Acad Sci U S A* **84**, 2180-2184 (1987).
67. Helin, K. Regulation of cell proliferation by the E2F transcription factors. *Curr Opin Genet Dev* **8**, 28-35 (1998).
68. Attwooll, C., Lazzerini Denchi, E. & Helin, K. The E2F family: specific functions and overlapping interests. *Embo j* **23**, 4709-4716 (2004).
69. Yamasaki, L. *et al.* Tumor induction and tissue atrophy in mice lacking E2F-1. *Cell* **85**, 537-548 (1996).
70. Field, S.J. *et al.* E2F-1 functions in mice to promote apoptosis and suppress proliferation. *Cell* **85**, 549-561 (1996).
71. Holmberg, C., Helin, K., Sehested, M. & Karlstrom, O. E2F-1-induced p53-independent apoptosis in transgenic mice. *Oncogene* **17**, 143-155 (1998).
72. Yamasaki, L. *et al.* Loss of E2F-1 reduces tumorigenesis and extends the lifespan of Rb1(+/-)mice. *Nat Genet* **18**, 360-364 (1998).
73. Stevens, C. & La Thangue, N.B. E2F and cell cycle control: a double-edged sword. *Arch Biochem Biophys* **412**, 157-169 (2003).
74. Blattner, C., Sparks, A. & Lane, D. Transcription factor E2F-1 is upregulated in response to DNA damage in a manner analogous to that of p53. *Mol Cell Biol* **19**, 3704-3713 (1999).
75. Biswas, A.K. & Johnson, D.G. Transcriptional and nontranscriptional functions of E2F1 in response to DNA damage. *Cancer Res* **72**, 13-17 (2012).
76. Wu, Z., Zheng, S. & Yu, Q. The E2F family and the role of E2F1 in apoptosis. *Int J Biochem Cell Biol* **41**, 2389-2397 (2009).
77. Bieda, M., Xu, X., Singer, M.A., Green, R. & Farnham, P.J. Unbiased location analysis of E2F1-binding sites suggests a widespread role for E2F1 in the human genome. *Genome Res* **16**, 595-605 (2006).

78. McClellan, K.A. & Slack, R.S. Specific in vivo roles for E2Fs in differentiation and development. *Cell Cycle* **6**, 2917-2927 (2007).
79. Wells, J., Graveel, C.R., Bartley, S.M., Madore, S.J. & Farnham, P.J. The identification of E2F1-specific target genes. *Proc Natl Acad Sci U S A* **99**, 3890-3895 (2002).
80. Dimova, D.K., Stevaux, O., Frolov, M.V. & Dyson, N.J. Cell cycle-dependent and cell cycle-independent control of transcription by the Drosophila E2F/RB pathway. *Genes Dev* **17**, 2308-2320 (2003).
81. Lim, C.A. *et al.* Genome-wide mapping of RELA(p65) binding identifies E2F1 as a transcriptional activator recruited by NF-kappaB upon TLR4 activation. *Mol Cell* **27**, 622-635 (2007).
82. Filtz, T.M., Vogel, W.K. & Leid, M. Regulation of transcription factor activity by interconnected, post-translational modifications. *Trends Pharmacol Sci* **35**, 76-85 (2014).
83. Beltrao, P., Bork, P., Krogan, N.J. & van Noort, V. Evolution and functional cross-talk of protein post-translational modifications. *Mol Syst Biol* **9**, 714 (2013).
84. You, L., Nie, J., Sun, W.J., Zheng, Z.Q. & Yang, X.J. Lysine acetylation: enzymes, bromodomains and links to different diseases. *Essays Biochem* **52**, 1-12 (2012).
85. Lin, W.C., Lin, F.T. & Nevins, J.R. Selective induction of E2F1 in response to DNA damage, mediated by ATM-dependent phosphorylation. *Genes Dev* **15**, 1833-1844 (2001).
86. Stevens, C., Smith, L. & Thangue, N.B.L. Chk2 activates E2F-1 in response to DNA damage. *Nature Cell Biology* **5**, 401-409 (2003).
87. Wang, B., Liu, K., Lin, F.T. & Lin, W.C. A role for 14-3-3 tau in E2F1 stabilization and DNA damage-induced apoptosis. *J Biol Chem* **279**, 54140-54152 (2004).
88. Ianari, A., Gallo, R., Palma, M., Alesse, E. & Gulino, A. Specific role for p300/CREB-binding protein-associated factor activity in E2F1 stabilization in response to DNA damage. *J Biol Chem* **279**, 30830-30835 (2004).
89. Van Den Broeck, A., Nissou, D., Brambilla, E., Eymin, B. & Gazzeri, S. Activation of a Tip60/E2F1/ERCC1 network in human lung adenocarcinoma cells exposed to cisplatin. *Carcinogenesis* **33**, 320-325 (2012).
90. Martinez-Balbas, M.A., Bauer, U.M., Nielsen, S.J., Brehm, A. & Kouzarides, T. Regulation of E2F1 activity by acetylation. *Embo j* **19**, 662-671 (2000).
91. Kontaki, H. & Talianidis, I. Lysine methylation regulates E2F1-induced cell death. *Mol Cell* **39**, 152-160 (2010).
92. Loftus, S.J., Liu, G., Carr, S.M., Munro, S. & La Thangue, N.B. NEDDylation regulates E2F-1-dependent transcription. *EMBO Rep* **13**, 811-818 (2012).
93. Cho, E.C. *et al.* Arginine methylation controls growth regulation by E2F-1. *EMBO J* **31**, 1785-1797 (2012).
94. Zheng, S. *et al.* Arginine methylation-dependent reader-writer interplay governs growth control by E2F-1. *Mol Cell* **52**, 37-51 (2013).
95. Blais, A. & Dynlacht, B.D. E2F-associated chromatin modifiers and cell cycle control. *Curr Opin Cell Biol* **19**, 658-662 (2007).
96. Bou Kheir, T. & Lund, A.H. Epigenetic dynamics across the cell cycle. *Essays Biochem* **48**, 107-120 (2010).
97. Brown, S.E., Fraga, M.F., Weaver, I.C., Berdasco, M. & Szyf, M. Variations in DNA methylation patterns during the cell cycle of HeLa cells. *Epigenetics* **2**, 54-65 (2007).
98. Taubert, S. *et al.* E2F-dependent histone acetylation and recruitment of the Tip60 acetyltransferase complex to chromatin in late G1. *Mol Cell Biol* **24**, 4546-4556 (2004).
99. Tyagi, S., Chabes, A.L., Wysocka, J. & Herr, W. E2F activation of S phase promoters via association with HCF-1 and the MLL family of histone H3K4 methyltransferases. *Mol Cell* **27**, 107-119 (2007).

100. Revenko, A.S., Kalashnikova, E.V., Gemo, A.T., Zou, J.X. & Chen, H.W. Chromatin loading of E2F-MLL complex by cancer-associated coregulator ANCCA via reading a specific histone mark. *Mol Cell Biol* **30**, 5260-5272 (2010).
101. Nagl, N.G., Wang, X., Patsialou, A., Van Scoy, M. & Moran, E. Distinct mammalian SWI/SNF chromatin remodeling complexes with opposing roles in cell-cycle control. *EMBO J* **26**, 752-763 (2007).
102. Balkwill, F. & Mantovani, A. Inflammation and cancer: back to Virchow? *Lancet* **357**, 539-545 (2001).
103. Sethi, G., Shanmugam, M.K., Ramachandran, L., Kumar, A.P. & Tergaonkar, V. Multifaceted link between cancer and inflammation. *Biosci Rep* **32**, 1-15 (2012).
104. Michalaki, V., Syrigos, K., Charles, P. & Waxman, J. Serum levels of IL-6 and TNF-alpha correlate with clinicopathological features and patient survival in patients with prostate cancer. *Br J Cancer* **90**, 2312-2316 (2004).
105. Moore, R.J. *et al.* Mice deficient in tumor necrosis factor-alpha are resistant to skin carcinogenesis. *Nat Med* **5**, 828-831 (1999).
106. Waters, J.P., Pober, J.S. & Bradley, J.R. Tumour necrosis factor and cancer. *J Pathol* **230**, 241-248 (2013).
107. Jung, Y.J., Isaacs, J.S., Lee, S., Trepel, J. & Neckers, L. IL-1beta-mediated up-regulation of HIF-1alpha via an NFkappaB/COX-2 pathway identifies HIF-1 as a critical link between inflammation and oncogenesis. *Faseb j* **17**, 2115-2117 (2003).
108. Pikarsky, E. *et al.* NF-kappaB functions as a tumour promoter in inflammation-associated cancer. *Nature* **431**, 461-466 (2004).
109. Karin, M. & Greten, F.R. NF-kappaB: linking inflammation and immunity to cancer development and progression. *Nat Rev Immunol* **5**, 749-759 (2005).
110. Mohanan, S. *et al.* Potential role of peptidylarginine deiminase enzymes and protein citrullination in cancer pathogenesis. *Biochem Res Int* **2012**, 895343 (2012).
111. Vossenaar, E.R., Zendman, A.J., van Venrooij, W.J. & Pruijn, G.J. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *Bioessays* **25**, 1106-1118 (2003).
112. Wang, S. & Wang, Y. Peptidylarginine deiminases in citrullination, gene regulation, health and pathogenesis. *Biochim Biophys Acta* **1829**, 1126-1135 (2013).
113. Curis, E. *et al.* Almost all about citrulline in mammals. *Amino Acids* **29**, 177-205 (2005).
114. Nakashima, K., Hagiwara, T. & Yamada, M. Nuclear localization of peptidylarginine deiminase V and histone deimination in granulocytes. *J Biol Chem* **277**, 49562-49568 (2002).
115. Wang, Y. *et al.* Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science* **306**, 279-283 (2004).
116. Cuthbert, G.L. *et al.* Histone deimination antagonizes arginine methylation. *Cell* **118**, 545-553 (2004).
117. Kearney, P.L. *et al.* Kinetic characterization of protein arginine deiminase 4: a transcriptional corepressor implicated in the onset and progression of rheumatoid arthritis. *Biochemistry* **44**, 10570-10582 (2005).
118. Lee, Y.H., Coonrod, S.A., Kraus, W.L., Jelinek, M.A. & Stallcup, M.R. Regulation of coactivator complex assembly and function by protein arginine methylation and demethylimination. *Proc Natl Acad Sci U S A* **102**, 3611-3616 (2005).
119. Kolodziej, S. *et al.* PADI4 acts as a coactivator of Tal1 by counteracting repressive histone arginine methylation, in *Nat Commun*, Vol. 5 3995 (England; 2014).
120. Chang, X. & Han, J. Expression of peptidylarginine deiminase type 4 (PAD4) in various tumors. *Mol Carcinog* **45**, 183-196 (2006).
121. Chang, X. *et al.* Increased PADI4 expression in blood and tissues of patients with malignant tumors. *BMC Cancer* **9**, 40 (2009).

122. Arita, K. *et al.* Structural basis for Ca(2+)-induced activation of human PAD4. *Nat Struct Mol Biol* **11**, 777-783 (2004).
123. Nakashima, K. *et al.* Molecular Characterization of Peptidylarginine Deiminase in HL-60 Cells Induced by Retinoic Acid and 1α,25-Dihydroxyvitamin D3. (1999).
124. Asaga, H., Nakashima, K., Senshu, T., Ishigami, A. & Yamada, M. Immunocytochemical localization of peptidylarginine deiminase in human eosinophils and neutrophils. *J Leukoc Biol* **70**, 46-51 (2001).
125. Chang, X. & Fang, K. PADI4 and tumourigenesis. *Cancer Cell Int* **10**, 7 (2010).
126. Christophorou, M.A. *et al.* Citrullination regulates pluripotency and histone H1 binding to chromatin. *Nature* **507**, 104-108 (2014).
127. Mantovani, A., Cassatella, M.A., Costantini, C. & Jaillon, S. Neutrophils in the activation and regulation of innate and adaptive immunity. *Nature Reviews Immunology* **11**, 519-531 (2011).
128. Remijsen, Q. *et al.* Dying for a cause: NETosis, mechanisms behind an antimicrobial cell death modality. *Cell Death Differ* **18**, 581-588 (2011).
129. Hagiwara, T., Nakashima, K., Hirano, H., Senshu, T. & Yamada, M. Deimination of arginine residues in nucleophosmin/B23 and histones in HL-60 granulocytes. *Biochem Biophys Res Commun* **290**, 979-983 (2002).
130. Wang, Y. *et al.* Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation. *J Cell Biol* **184**, 205-213 (2009).
131. Knight, J.S. *et al.* Peptidylarginine deiminase inhibition disrupts NET formation and protects against kidney, skin and vascular disease in lupus-prone MRL/lpr mice. *Ann Rheum Dis* (2014).
132. Nakashima, K. *et al.* PAD4 regulates proliferation of multipotent haematopoietic cells by controlling c-myc expression. *Nat Commun* **4**, 1836 (2013).
133. Foulquier, C. *et al.* Peptidyl arginine deiminase type 2 (PAD-2) and PAD-4 but not PAD-1, PAD-3, and PAD-6 are expressed in rheumatoid arthritis synovium in close association with tissue inflammation. *Arthritis Rheum* **56**, 3541-3553 (2007).
134. Valesini, G. *et al.* Citrullination and autoimmunity. *Autoimmun Rev* **14**, 490-497 (2015).
135. Feldmann, M., Brennan, F.M. & Maini, R.N. Rheumatoid arthritis. *Cell* **85**, 307-310 (1996).
136. van Boekel, M.A., Vossenaar, E.R., van den Hoogen, F.H. & van Venrooij, W.J. Autoantibody systems in rheumatoid arthritis: specificity, sensitivity and diagnostic value, in *Arthritis Res*, Vol. 4 87-93 (2002).
137. Vossenaar, E.R. *et al.* Expression and activity of citrullinating peptidylarginine deiminase enzymes in monocytes and macrophages. *Ann Rheum Dis* **63**, 373-381 (2004).
138. Sakkas, L.I., Bogdanos, D.P., Katsiari, C. & Platsoucas, C.D. Anti-citrullinated peptides as autoantigens in rheumatoid arthritis-relevance to treatment. *Autoimmun Rev* **13**, 1114-1120 (2014).
139. Suzuki, A. *et al.* Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat Genet* **34**, 395-402 (2003).
140. Wood, D.D. *et al.* Myelin localization of peptidylarginine deiminases 2 and 4: comparison of PAD2 and PAD4 activities. *Lab Invest* **88**, 354-364 (2008).
141. Mediwake, R., Isenberg, D.A., Schellekens, G.A. & van Venrooij, W.J. Use of anti-citrullinated peptide and anti-RA33 antibodies in distinguishing erosive arthritis in patients with systemic lupus erythematosus and rheumatoid arthritis. *Ann Rheum Dis* **60**, 67-68 (2001).
142. Koziel, J., Mydel, P. & Potempa, J. The link between periodontal disease and rheumatoid arthritis: an updated review. *Curr Rheumatol Rep* **16**, 408 (2014).
143. Gabarrini, G. *et al.* The peptidylarginine deiminase gene is a conserved feature of Porphyromonas gingivalis. *Sci Rep* **5**, 13936 (2015).

144. Konig, M.F., Paracha, A.S., Moni, M., Bingham, C.O., 3rd & Andrade, F. Defining the role of Porphyromonas gingivalis peptidylarginine deiminase (PPAD) in rheumatoid arthritis through the study of PPAD biology. *Ann Rheum Dis* **74**, 2054-2061 (2015).
145. Sharma, P. *et al.* Citrullination of histone H3 interferes with HP1-mediated transcriptional repression. *PLoS Genet* **8**, e1002934 (2012).
146. Neeli, I., Khan, S.N. & Radic, M. Histone deimination as a response to inflammatory stimuli in neutrophils. *J Immunol* **180**, 1895-1902 (2008).
147. Mastronardi, F.G. *et al.* Increased citrullination of histone H3 in multiple sclerosis brain and animal models of demyelination: a role for tumor necrosis factor-induced peptidylarginine deiminase 4 translocation. *J Neurosci* **26**, 11387-11396 (2006).
148. Zhang, X. *et al.* Genome-wide analysis reveals PADI4 cooperates with Elk-1 to activate c-Fos expression in breast cancer cells. *PLoS Genet* **7**, e1002112 (2011).
149. Deplus, R. *et al.* Citrullination of DNMT3A by PADI4 regulates its stability and controls DNA methylation. *Nucleic Acids Res* **42**, 8285-8296 (2014).
150. Stadler, S.C. *et al.* Dysregulation of PAD4-mediated citrullination of nuclear GSK3β activates TGF-β signaling and induces epithelial-to-mesenchymal transition in breast cancer cells. *Proc Natl Acad Sci U S A* **110**, 11851-11856 (2013).
151. Tanikawa, C. *et al.* Regulation of protein Citrullination through p53/PADI4 network in DNA damage response. *Cancer Res* **69**, 8761-8769 (2009).
152. Tanikawa, C. *et al.* Regulation of histone modification and chromatin structure by the p53-PADI4 pathway. *Nat Commun* **3**, 676 (2012).
153. Yao, H. *et al.* Histone Arg modifications and p53 regulate the expression of OKL38, a mediator of apoptosis. *J Biol Chem* **283**, 20060-20068 (2008).
154. Li, P. *et al.* Regulation of p53 target gene expression by peptidylarginine deiminase 4. *Mol Cell Biol* **28**, 4745-4758 (2008).
155. Li, P. *et al.* Coordination of PAD4 and HDAC2 in the regulation of p53-target gene expression. *Oncogene* **29**, 3153-3162 (2010).
156. Filippakopoulos, P. *et al.* Histone recognition and large-scale structural analysis of the human bromodomain family, in *Cell*, Vol. 149 214-231 (2012 Elsevier Inc, United States; 2012).
157. Andrade, F. *et al.* Autocitrullination of human peptidyl arginine deiminase type 4 regulates protein citrullination during cell activation. *Arthritis Rheum* **62**, 1630-1640 (2010).
158. Hensen, S.M.M. & Pruijn, G.J.M. Methods for the Detection of Peptidylarginine Deiminase (PAD) Activity and Protein Citrullination*, in *Mol Cell Proteomics*, Vol. 13 388-396 (2014).
159. Berg, J.M., Tymoczko, J.L. & Stryer, L. The Michaelis-Menten Model Accounts for the Kinetic Properties of Many Enzymes. (2002).
160. Kubilus, J., Waitkus, R.F. & Baden, H.P. Partial purification and specificity of an arginine-converting enzyme from bovine epidermis. *Biochim Biophys Acta* **615**, 246-251 (1980).
161. Nomura, K. Specificity and mode of action of the muscle-type protein-arginine deiminase. *Arch Biochem Biophys* **293**, 362-369 (1992).
162. Assohou-Luty, C. *et al.* The human peptidylarginine deiminases type 2 and type 4 have distinct substrate specificities. *Biochim Biophys Acta* **1844**, 829-836 (2014).
163. Clontech (2015).
164. Blau, H.M. & Rossi, F.M. Tet B or not tet B: advances in tetracycline-inducible gene expression. *Proc Natl Acad Sci U S A* **96**, 797-799 (1999).
165. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A* **89**, 5547-5551 (1992).
166. Guo, Q. & Fast, W. Citrullination of inhibitor of growth 4 (ING4) by peptidylarginine deminase 4 (PAD4) disrupts the interaction between ING4 and p53. *J Biol Chem* **286**, 17069-17078 (2011).
167. Jones, J.E. *et al.* Synthesis and screening of a haloacetamidine containing library to identify PAD4 selective inhibitors. *ACS Chem Biol* **7**, 160-165 (2012).

168. Korah, J., Falah, N., Lacerte, A. & Lebrun, J.J. A transcriptionally active pRb-E2F1-P/CAF signaling pathway is central to TGFB-mediated apoptosis. *Cell Death & Disease* **3** (2012).
169. Schneider-Poetsch, T. *et al.* Inhibition of Eukaryotic Translation Elongation by Cycloheximide and Lactimidomycin. *Nat Chem Biol* **6**, 209-217 (2010).
170. Marti, A., Wirbelauer, C., Scheffner, M. & Krek, W. Interaction between ubiquitin-protein ligase SCFSKP2 and E2F-1 underlies the regulation of E2F-1 degradation. *Nat Cell Biol* **1**, 14-19 (1999).
171. Kerscher, O., Felberbaum, R. & Hochstrasser, M. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* **22**, 159-180 (2006).
172. Yam, C.H., Fung, T.K. & Poon, R.Y. Cyclin A in cell cycle control and cancer. *Cell Mol Life Sci* **59**, 1317-1326 (2002).
173. Nagata, S. & Senshu, T. Peptidylarginine deiminase in rat and mouse hemopoietic cells. *Experientia* **46**, 72-74 (1990).
174. Birnie, G.D. The HL60 cell line: a model system for studying human myeloid cell differentiation. *Br J Cancer Suppl* **9**, 41-45 (1988).
175. Marks, P.A. & Breslow, R. Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nature Biotechnology* **25**, 84-90 (2007).
176. Li, Z. The alphaMbeta2 integrin and its role in neutrophil function. *Cell Res* **9**, 171-178 (1999).
177. Mollinedo, F., Lopez-Perez, R. & Gajate, C. Differential gene expression patterns coupled to commitment and acquisition of phenotypic hallmarks during neutrophil differentiation of human leukaemia HL-60 cells. *Gene* **419**, 16-26 (2008).
178. Lewis, H.D. *et al.* Inhibition of PAD4 activity is sufficient to disrupt mouse and human NET formation, in *Nat Chem Biol*, Vol. 11 189-191 (United States; 2015).
179. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).
180. Mi, H., Muruganujan, A. & Thomas, P.D. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res* **41**, D377-386 (2013).
181. Yun, M., Wu, J., Workman, J.L. & Li, B. Readers of histone modifications, in *Cell Res*, Vol. 21 564-578 (2011).
182. Muller, S., Filippakopoulos, P. & Knapp, S. Bromodomains as therapeutic targets, in *Expert Rev Mol Med*, Vol. 13 (
183. Sanchez, R. & Zhou, M.M. The role of human bromodomains in chromatin biology and gene transcription. *Curr Opin Drug Discov Devel* **12**, 659-665 (2009).
184. Jung, M. *et al.* Affinity Map of Bromodomain Protein 4 (BRD4) Interactions with the Histone H4 Tail and the Small Molecule Inhibitor JQ1. (2014).
185. Müller, S. & Knapp, S. Discovery of BET bromodomain inhibitors and their role in target validation. *Med. Chem. Commun.* **5**, 288-296 (2014).
186. Wu, S.Y. & Chiang, C.M. The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. *J Biol Chem* **282**, 13141-13145 (2007).
187. Junwei, S. & Vakoc, C.R. The mechanisms behind the therapeutic activity of BET bromodomain inhibition. *Mol Cell* **54**, 728-736 (2014).
188. Jang, M.K. *et al.* The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* **19**, 523-534 (2005).
189. Chen, R., Yik, J.H., Lew, Q.J. & Chao, S.H. Brd4 and HEXIM1: multiple roles in P-TEFb regulation and cancer. *Biomed Res Int* **2014**, 232870 (2014).
190. Nicodeme, E. *et al.* Suppression of inflammation by a synthetic histone mimic. *Nature* **468**, 1119-1123 (2010).

191. Hargreaves, D.C., Horng, T. & Medzhitov, R. Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell* **138**, 129-145 (2009).
192. Prinjha, R.K., Witherington, J. & Lee, K. Place your BETs: the therapeutic potential of bromodomains. *Trends Pharmacol Sci* **33**, 146-153 (2012).
193. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067-1073 (2010).
194. Belkina, A.C., Nikolajczyk, B.S. & Denis, G.V. BET protein function is required for inflammation: Brd2 genetic disruption and BET inhibitor JQ1 impair mouse macrophage inflammatory responses. *J Immunol* **190**, 3670-3678 (2013).
195. Mele, D.A. *et al.* BET bromodomain inhibition suppresses TH17-mediated pathology, in *J Exp Med*, Vol. 210 2181-2190 (2013).
196. Knoerzer, D.B., Donovan, M.G., Schwartz, B.D. & Mengle-Gaw, L.J. Clinical and histological assessment of collagen-induced arthritis progression in the diabetes-resistant BB/Wor rat. *Toxicol Pathol* **25**, 13-19 (1997).
197. Lipsky, P.E. *et al.* Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. *N Engl J Med* **343**, 1594-1602 (2000).
198. Inglis, J.J., Šimelyte, E., McCann, F.E., Criado, G. & Williams, R.O. Protocol for the induction of arthritis in C57BL/6 mice. *Nature Protocols* **3**, 612-618 (2008).
199. Wang, Y. *et al.* Anticancer peptidylarginine deiminase (PAD) inhibitors regulate the autophagy flux and the mammalian target of rapamycin complex 1 activity. *J Biol Chem* **287**, 25941-25953 (2012).
200. Warg, L.A. *et al.* The role of the E2F1 transcription factor in the innate immune response to systemic LPS, in *Am J Physiol Lung Cell Mol Physiol*, Vol. 303 L391-400 (United States; 2012).
201. Ren, Z. *et al.* E2F1 renders prostate cancer cell resistant to ICAM-1 mediated antitumor immunity by NF-kappaB modulation. *Mol Cancer* **13**, 84 (2014).
202. Yang, I.V. *et al.* Novel regulators of the systemic response to lipopolysaccharide. *Am J Respir Cell Mol Biol* **45**, 393-402 (2011).
203. Martinod, K. *et al.* PAD4-deficiency does not affect bacteremia in polymicrobial sepsis and ameliorates endotoxemic shock. *Blood* **125**, 1948-1956 (2015).
204. Rakoff-Nahoum, S. Why Cancer and Inflammation? *Yale J Biol Med* **79**, 123-130 (2006).
205. Brown, J.D. *et al.* NF-kappaB directs dynamic super enhancer formation in inflammation and atherogenesis. *Mol Cell* **56**, 219-231 (2014).
206. Xu, Y. & Vakoc, C.R. Brd4 is on the move during inflammation. *Trends Cell Biol* **24**, 615-616 (2014).
207. Huang, B., Yang, X.D., Zhou, M.M., Ozato, K. & Chen, L.F. Brd4 coactivates transcriptional activation of NF-kappaB via specific binding to acetylated RelA, in *Mol Cell Biol*, Vol. 29 1375-1387 (United States; 2009).
208. Zou, Z. *et al.* Brd4 maintains constitutively active NF-kappaB in cancer cells by binding to acetylated RelA. *Oncogene* **33**, 2395-2404 (2014).
209. Medzhitov, R. & Horng, T. Transcriptional control of the inflammatory response. *Nat Rev Immunol* **9**, 692-703 (2009).
210. Dey, A., Chitsaz, F., Abbasi, A., Misteli, T. & Ozato, K. The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc Natl Acad Sci U S A* **100**, 8758-8763 (2003).
211. Costa, D.D. *et al.* BET inhibition as a single or combined therapeutic approach in primary paediatric B-precursor acute lymphoblastic leukaemia. *Blood Cancer Journal* **3** (2013).
212. Cantaert, T., Teitsma, C., Tak, P.P. & Baeten, D. Presence and role of anti-citrullinated protein antibodies in experimental arthritis models. *Arthritis Rheum* **65**, 939-948 (2013).