



HAL
open science

Epstein Barr virus EBNA2: interactions of intrinsically disordered proteins with host cell factors.

Eva-Maria Geenen

► **To cite this version:**

Eva-Maria Geenen. Epstein Barr virus EBNA2: interactions of intrinsically disordered proteins with host cell factors.. Biochemistry, Molecular Biology. Université de Grenoble, 2013. English. NNT : 2013GREN055 . tel-01555490

HAL Id: tel-01555490

<https://theses.hal.science/tel-01555490>

Submitted on 4 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE GRENOBLE



THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biologie structurale et nanobiologie**

Arrêté ministériel : 7 août 2006

Présentée par

Eva-Maria Geenen

Thèse dirigée par **Dr Darren Hart**

préparée au sein du **Laboratoire EMBL**

dans l'**École Doctorale chimie et sciences du vivant (218)**

Etude de la protéine EBNA2 du virus Epstein-Barr et ses interactions avec les facteurs de la cellule hôte

Thèse soutenue publiquement le **17 décembre 2013**,

devant le jury composé de :

Dr Wim Burmeister	UVHCI Grenoble	Président
Dr Darren Hart	EMBL Grenoble	Directeur
Dr Peter Tompa	VIB Bruxelles	Rapporteur
Dr Paul Farrell	Imperial College London	Rapporteur
Dr Imre Berger	EMBL Grenoble	Examineur
Dr François Penin	IBCP Lyon	Examineur



Université Joseph Fourier / Université Pierre Mendès France /
Université Stendhal / Université de Savoie / Grenoble INP

UNIVERSITÉ DE GRENOBLE



THESIS

To obtain the title of

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Specialisation: structural biologie and nanobiologie

Arrêté ministériel : 7 août 2006

Presented by

Eva-Maria Geenen

Thesis director **Dr Darren Hart**

Prepared at the **EMBL Grenoble**

in l'École Doctorale chimie et sciences du vivant (218)

Studies of Epstein-Barr virus EBNA2 and its interactions with host cell factors

Thesis defended in public the **17th of December 2013**,

In front of the jury compsed of :

Dr Wim Burmeister	UVHCI Grenoble	Chair
Dr Darren Hart	EMBL Grenoble	Supervisor
Dr Peter Tompa	VIB Brussels	Reviewer
Dr Paul Farrell	Imperial College London	Reviewer
Dr Imre Berger	EMBL Grenoble	Examiner
Dr François Penin	IBCP Lyon	Examiner



Université Joseph Fourier / Université Pierre Mendès France /
Université Stendhal / Université de Savoie / Grenoble INP

Contents

Résumé	7
Abstract	8
Introduction	9
Résumé	9
1.1 Herpesviridae	10
1.1.1. Epstein-Barr virus	10
1.1.1.1. Replication cycle.....	10
1.1.1.2. Infectious mononucleosis and cancer due to EBV	12
1.1.1.3. Therapy	12
1.1.1.4. Latency.....	12
1.1.1.5. EBV nuclear antigen 2 (EBNA2)	13
1.2. Signal transducer and activator of transcription 3 (STAT3)	16
1.2.1. Functions of STAT3 and possible advantages for the virus in “hijacking” STAT3.....	18
1.2.1.1. STAT3 in cell cycle control.....	19
1.2.1.2. STAT3 in immunosuppression	19
1.2.2. STAT3 in malignancies	19
1.2.3. STAT3 inhibitors	20
1.3. The nuclear receptor co-repressor 2 (NcoRII/SMRT).....	21
1.3.1. Transcriptional repression.....	21
1.3.2. Silencing mediator of retinoic acid and thyroid hormone receptor (SMRT).....	21
1.3.2.1. Functions of SMRT	22
1.4. Intrinsically disordered proteins (IDPs).....	23
1.4.1. Disorder properties of EBNA2	24
1.4.2. IDPs in malignancies	25
1.5. Project origin and objectives	26
Results	28
Résumé	28
2.1. Interaction of STAT3 with the viral transcriptional activator EBNA2	30
2.1.1. Generation of highly expressing and well behaving EBNA2 constructs.....	30
2.1.1.1. Overview ESPRIT technology.....	30
2.1.2. Expression and purification of recombinant protein.....	33
2.1.2.1. Expression and purification of the STAT3 homodimer	33

2.1.2.2.	Expression and purification of EBNA2 fragments.....	34
2.1.3.	Structural characterisation of the IDP EBNA2.....	38
2.1.3.1.	Biochemical characterisation of EBNA2.....	38
2.1.3.2.	Biophysical characterisation of EBNA2.....	38
2.1.4.	Interaction between EBNA2 and STAT3.....	39
2.1.4.1.	Binding analysis by surface plasmon resonance and pull-down experiments.....	39
2.1.4.2.	Binding analysis by high field nuclear magnetic resonance (NMR).....	42
2.1.5.	Identification of the STAT3 binding region in EBNA2.....	43
2.1.6.	Identification of EBNA2 residues involved in the binding to STAT3.....	47
2.1.7.	Identification of the EBNA2 binding region in STAT3.....	49
2.1.8.	Binding of EBNA2 D9 does not prevent DNA binding.....	50
2.1.9.	First crystallisation trials.....	51
2.1.10.	Cellular assays investigating the STAT3-EBAN2 interaction.....	52
2.1.10.1.	Generation of stable cells expressing the luciferase gene under a STAT3 promoter element.....	52
2.1.10.2.	Studying the STAT3-EBNA2 interaction by transient transfection.....	54
2.1.11.	Interaction of EBNA2 with other host cell proteins.....	58
2.1.11.1.	Binding of EBNA2 to ZMYND.....	58
2.1.11.2.	Binding of EBNA2 to Med25.....	60
2.2.	Interaction of STAT3 with the cellular co-repressor SMRT/NcoR1I.....	61
2.2.1.	Identification of the binding region using a cell based approach.....	61
2.2.2.	Expression and purification of SMRT fragment.....	62
2.2.3.	Identification of a SMRT fragment binding to STAT3.....	64
2.2.4.	Identification of the amino acids in fragment 39L23 involved in STAT3 binding.....	64
2.2.5.	Binding kinetics and mutational analysis of the STAT3-39L23 binding.....	65
Discussion and Outlook.....		67
	Résumé.....	67
3.1.	The interaction between the viral protein EBNA2 and STAT3.....	70
3.2.	The interaction between EBNA2 and other host proteins.....	73
3.2.1.	Interaction between EBNA2 and the BS69 MYND domain.....	73
3.2.2.	Interaction between EBNA2 and Med25.....	74
3.3.	The interaction between the cellular protein SMRT and STAT3.....	75
3.4.	Intrinsically disordered proteins.....	77
3.5.	The interaction triad: EBNA2 competing with SMRT for STAT3 binding.....	79

Material and Methods	82
4.1. ESPRIT technology	82
4.1.1. ESPRIT- experimental procedure	83
4.2. Protein expression and purification	85
4.2.1. Expression and purification of the STAT3 homodimer	85
4.2.2. Expression and purification of EBNA2 fragments.....	86
4.2.3. Expression and purification of SMRT fragments	88
4.2.4. Expression and purification of EBNA2-TAs-pGEX-4T and ZMYND11-pGEX-4T	89
4.3. Kinetic binding analysis	90
4.3.1. Surface plasmon resonance (SPR).....	90
4.3.2. Microscale thermophoresis (MST)	91
4.3.3. Nuclear magnetic resonance (NMR) sample preparation.....	92
4.4. Streptavidin agarose pull-down experiments.....	92
4.5. Generation of interaction mutants	93
4.5.1. Generation of GST tagged EBNA2-CR8 fragments	93
4.5.2. Site directed mutagenesis.....	94
4.5.3. Overlap extension PCR	95
4.5.4. Generation of SMRT fragment mammalian expression vectors	97
4.6. Electrophoretic mobility shift assay (EMSA).....	98
4.7. Limited proteolysis	98
4.8. Crystallisation trials.....	98
4.9. Cell based assays.....	99
4.9.1. Generation of stable, monoclonal cells expressing the firefly luciferase gene under control of a STAT3 promoter	99
4.9.2. CBF1 reporter assay.....	100
4.9.3. STAT3 reporter assay.....	100
References	101
Appendix	112
Acknowledgements	125

Résumé

Il a été montré récemment que l'interaction physiologique entre le co-répresseur nucléaire SMRT et le facteur de transcription cellulaire STAT3 pouvait être perturbée par la protéine EBNA2 du virus Epstein-Barr (EBV). Normalement, SMRT réduit l'activité transcriptionnelle de STAT3. La dissociation de SMRT médiée par EBNA2 est donc supposée augmenter la transcription des gènes de l'hôte et viraux régulés par STAT3. STAT3 régule plusieurs effets immunosuppresseur, mais aussi pro-prolifératif et anti-apoptotique. Cette activité pourrait donc être importante pour la stratégie de survie du virus. EBNA2 et SMRT sont prédites comme étant des protéines intrinsèquement désordonnées (IDPs), comme environ un tiers des protéines eucaryotes. Malgré cette abondance, notre compréhension de la relation structure-fonction des IDPs est limitée, partiellement parce qu'il est difficile de produire des échantillons stables et en grande quantité, ce qui est nécessaire pour des études structurales et biophysiques. La technologie de bibliothèque aléatoire ESPRIT a été utilisée pour résoudre ce problème et générer des fragments solubles et s'exprimant bien d'EBNA2 et de SMRT, d'une taille adaptée à des études biochimiques, biophysiques et RMN, pour étudier leurs interactions avec STAT3. Les régions d'interaction ont été cartographiées avec une résolution à l'acide aminé près, et ont été confirmées par mutagenèse. Les affinités ont été déterminées par différentes méthodes, démontrant qu'EBNA2 se lie à STAT3 avec une affinité de l'ordre du nanomolaire, et approximativement dix fois plus fort que SMRT. Des expériences utilisant la luciférase comme rapporteur ont été effectuées pour tester les interactions de mutants d'EBNA2 en cellules de mammifères en observant l'expression des gènes contrôlés par STAT3. Ces données suggèrent qu'un simple mécanisme de compétition pourrait expliquer comment EBNA2 active STAT3 lorsqu'il est réprimé, et ouvre de nouveaux horizons pour des études fonctionnelles cherchant à comprendre comment EBV établit et maintient la latence dans la cellule hôte.

Abstract

It was recently shown that the physiological interaction between the nuclear corepressor SMRT and the cellular transcription factor STAT3 could be disrupted by the Epstein-Barr virus (EBV) protein EBNA2. SMRT normally decreases the transcriptional activity of STAT3, thus EBNA2-mediated release of SMRT is expected to enhance the transcription of host and viral STAT3 regulated genes. STAT3 regulates several immunosuppressive as well as pro-proliferative and anti-apoptotic effects in the host organism and so this activity may be important for the viral survival strategy. Bioinformatics predict both EBNA2 and SMRT to be intrinsically disordered proteins (IDPs), in common with about one third of all eukaryotic proteins. Despite this abundance, our understanding of how this property relates to function is limited, in part due to difficulties in production of stable samples in the large quantities necessary for structural and biophysical studies. The ESPRIT random library technology was used to overcome these problems and generate soluble well-expressing fragments of EBNA2 and SMRT that were of suitable size for biochemical, biophysical and NMR experiments to study their interactions with STAT3. The interaction regions were mapped to individual amino acid resolution and confirmed by mutagenesis. Affinities were determined by different methods demonstrating that EBNA2 binds STAT3 with nanomolar affinity, and approximately ten-fold more tightly than SMRT. Luciferase reporter assays were performed in order to test mutants of EBNA2 binding in mammalian cells, and effects on STAT3 mediated reporter gene expression were observed. These data suggest that a simple competitive mechanism may explain how EBNA2 activates repressed STAT3, and open the way for future functional investigations into how EBV establishes and maintains latency in the host cell.

Introduction

Résumé

L'introduction présente premièrement la biologie des virus herpès en général. Les virus herpès sont des virus enveloppés dont l'ADN double brin est encapsidé dans une capsidie icosaédrique. Ils peuvent être divisés en trois sous-familles : alpha-, beta- et gammaherpesviridae. Le virus Epstein-Barr (EBV) est un des huit virus herpès pathogènes et appartient au groupe des herpès virus gamma et au genre lymphocryptovirus. C'est aussi le premier virus à avoir été associé au cancer humain. En plus de son association avec le cancer, l'infection lytique par EBV peut causer la mononucléose infectieuse (IM). D'après l'OMS, plus de 95 % de la population mondiale est séropositive pour EBV, mais à ce jour aucun traitement efficace contre EBV n'existe sur le marché. Une caractéristique spécifique d'EBV est que le virus persiste toute la vie dans l'organisme hôte et peut être réactivé à n'importe quel moment en cas de conditions immunosuppressives chez l'hôte. Pendant les phases de latence, le virus est silencieux et seulement quelques gènes viraux sont exprimés. Une des protéines exprimée pendant la latence III est l'antigène nucléaire 2 d'EBV (EBNA2). EBNA2 est un activateur transcriptionnel, la première protéine exprimée pendant l'infection des cellules B. Elle joue également un rôle important dans l'activation de cellules B au repos et stimule leur prolifération en régulant positivement l'expression de gènes viraux et en régulant négativement l'expression de gènes cellulaires. Il a été montré récemment que, lors de l'infection, EBNA2 interagit avec le facteur de transcription cellulaire STAT3. Cette interaction pourrait augmenter la transcription des gènes de l'hôte et du virus régulés par STAT3. La signalisation activée par STAT3 est pro-prolifération, anti-apoptotique, et a plusieurs effets de suppression immune. De plus, il a été proposé que STAT3 soit requise pour maintenir l'état latent. Une augmentation du niveau de STAT3 est de ce fait importante pour la stratégie de survie du virus EBV. En conditions physiologiques, STAT3 est liée à la protéine co-répresseur médiatrice pour les récepteurs aux hormones rétinoïdes et thyroïdes (SMRT). Cette interaction réduit l'activité transcriptionnelle de STAT3. EBNA2 et SMRT sont des protéines intrinsèquement désordonnées (IDPs), comme environ un tiers des protéines eucaryotes et 70 % des protéines liées au cancer. Des études structurales détaillées d'EBNA2 et SMRT n'existent pas. Cependant, d'après certains systèmes mieux compris, l'importance des IDPs dans diverses pathologies, y compris dans le cancer et les maladies neurodégénératives, est claire. Les IDPs sont fortement liées au cancer du fait de leur rôle actif dans la prolifération, l'apoptose, le contrôle du cycle cellulaires et l'angiogenèse. EBNA2 et SMRT sont des exemples de la façon dont les IDPs entrent en compétition pour influencer la transcription contrôlée par STAT3.

1.1 Herpesviridae

Herpesviridae can be divided into three subfamilies: alpha-, beta- and gammaherpesviridae. A typical herpesvirion contains a double-stranded DNA genome which is packed into the core, an icosadeltahedral capsid made up of 162 capsomeres, a tegument that surrounds the capsid and an envelope with viral glycoprotein spikes on the surface. The genome is 120-230 kbp and linear, but circularises immediately upon release from the capsid into the nucleus of the host cell. It comprises unique regions that are disrupted by internal and terminal repeated regions. Latency is a special feature of all herpesviruses whereby the virus stays quiet in the host cells and only a few viral gene products are expressed. Under immunosuppressive conditions the virus can be reactivated and enters the lytic replication cycle that is characterised by the irreversible destruction of the infected cell and the production of progeny viruses (Fields et al., 1996).

1.1.1. Epstein-Barr virus

Epstein-Barr virus (EBV) was discovered in 1964 when Anthony Epstein and Yvonne Barr isolated the virus from Burkitt's lymphoma biopsy samples and identified it by electron microscopy (Epstein et al., 1965). EBV was also the first virus to be associated with human cancer. It is one of eight human pathogenic herpesviruses and belongs to the group of gammaherpesviruses and the lymphocryptovirus genera. Different subtypes are described in the literature at which the strains P3HR-1 and B95-8 are the most studied (Bornkamm et al., 1980). The EBV genome has a size of 172 kbp and contains two direct repeats at the ends and an internal repeat that divides the genome into short and long unique regions (Fields et al., 1996).

1.1.1.1. Replication cycle

The virus binds to cells via its glycoproteins gp 220 and gp350 (Nemerow et al., 1987) to the CD21 integral membrane receptor. It is then internalised into cytoplasmic vesicles and fusion of viral and cellular membrane leads to the release of the nucleocapsid into the cytoplasm which is transported along the microtubules to the nuclear pores. The genome is released from the nucleocapsid and enters the nucleus where it circularises. Two distinct replication cycles can be distinguished: the lytic and the latent replication cycle. During the lytic phase all viral proteins are expressed whilst in latency there is expression of only a few proteins and RNAs. One of these is EBNA2 which is also referred as the transforming antigen of the virus. Viral transcription occurs in three phases: immediate early, delayed early and late transcription. After replication of the genome, which occurs via a rolling cycle mechanism, the progeny viral particles bud away from the nuclear membrane (transport budding) and then from the endoplasmic membrane (maturation budding). The viral particles are released from

the host cell following fusion with the cell membrane (Mettenleiter et al., 2009). The replication cycle of EBV is summarised in Fig. 1.1.

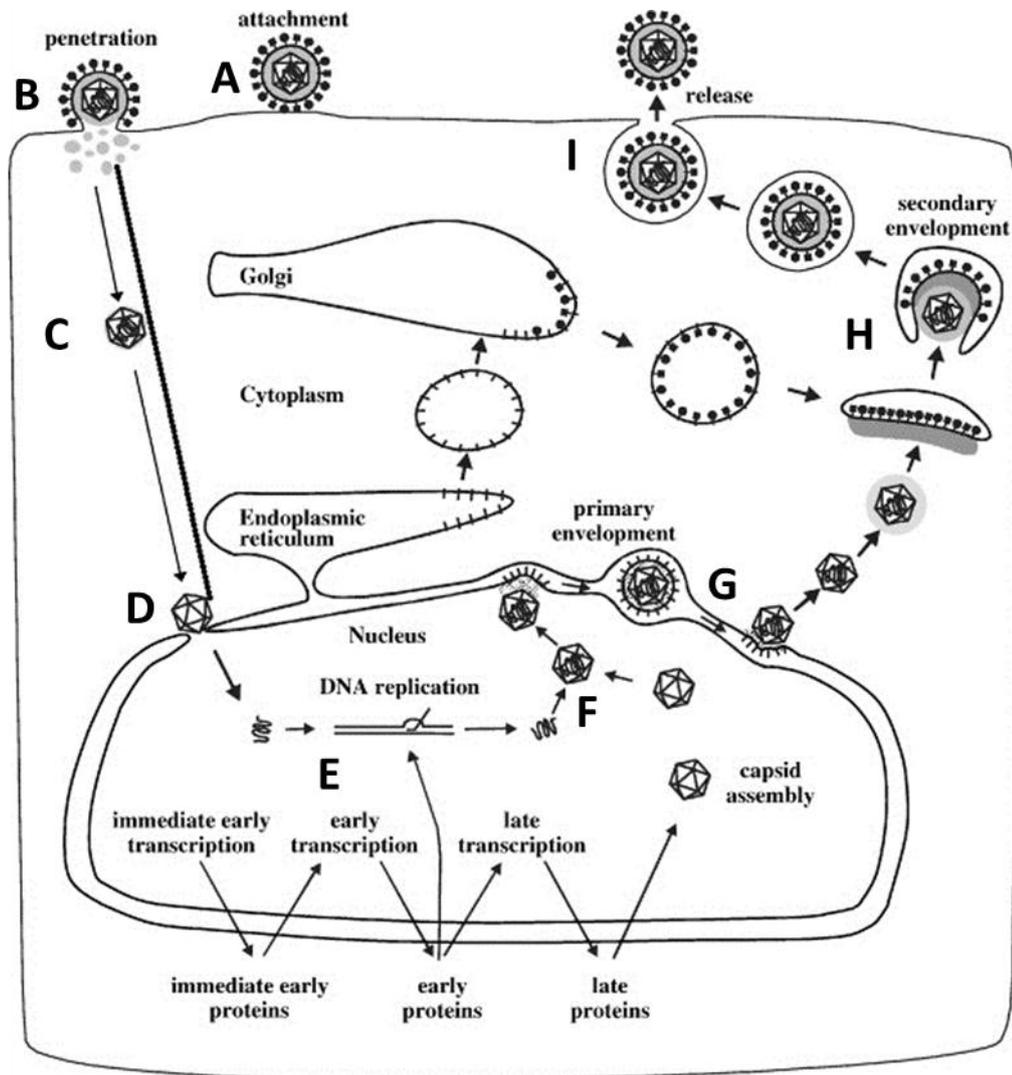


Figure 1.1: Epstein-Barr viral replication cycle. (A) The enveloped virus binds host cell surface receptors. (B) After fusion of viral and cellular membrane the capsid is released into the cytoplasm and (C) transported to the nucleus. (D) The virus docks to the nuclear pore (NP) where the viral genome is released into the nucleus and circularizes. (E) Viral genes are transcribed and the viral genome replicates. (F) Viral DNA is packed into preformed capsids and (G) leaves the nucleus (transportation budding). (H) Final maturation occurs in the cytoplasm by budding of the viral capsid into vesicles of the trans-golgi network which contain viral glycoproteins. (I) The enveloped virion is transported to the cell surface and is released into the outer cellular space by fusion. Picture modified from (Mettenleiter, 2004) and used with permission from the author.

1.1.1.2. Infectious mononucleosis and cancer due to EBV

According to the WHO over 95% of the world's population is seropositive for EBV (WHO | Viral Cancers). In most of the cases primary infection occurs during early childhood and is asymptomatic. In adults lytic infection can cause infectious mononucleosis which is marked by atypical lymphocytosis (Michelow et al., 2012). Classical symptoms are the triad of fever, oropharyngitis, and a bilateral and symmetrical lymphadenitis in the posterior triangle of the neck (Lawee, 2007). Furthermore, the virus has been associated with 1% of globally occurring cancer cases (Parkin, 2006). It is implicated in the causation of cancers like Burkitt's lymphoma, Hodgkin's lymphoma, nasopharyngeal carcinoma, primary central nervous system lymphoma (PCNSL) and posttransplant lymphoproliferative disease (PTLD) (Young and Rickinson, 2004).

1.1.1.3. Therapy

Acyclic nucleoside and nucleotide inhibitors, which are used as antiviral drugs, have a high specificity as they require phosphorylation to their triphosphorylated forms in order to be active and they are favoured by the viral kinase (Clercq and Holy, 2005; Jordheim et al., 2013). Therefore they are able to inhibit the viral polymerase but are inactive in uninfected cells. Unfortunately, they are only effective in cell culture and their use for the treatment of acute infectious mononucleosis in humans is quite limited (Young and Rickinson, 2004). Reasons for this are the long incubation time of 4-6 weeks, the location of EBV in the saliva which is difficult to access by orally administrated drugs, and the fact that most symptoms of IM are not due to cytopathic effects of the virus but rather due to immunopathic responses to EBV-infected cells.

1.1.1.4. Latency

A special feature of EBV is that the virus persists lifelong in the host organism and can be reactivated at any time under immunosuppressive conditions. During latency the virus remains silent in the cell with only a few viral genes being expressed (Young and Rickinson, 2004). Three latent membrane proteins (LMPs 1, 2A and 2B) and six EBV nuclear antigens (EBNAs 1, 2, 3A, 3B, 3C and -LP) can be detected, as well as two viral RNAs (EBER1 and EBER2). All EBV associated cancer malignancies are characterised by latent gene expression. Four latency types are characterised by different gene expression profiles and are associated with different cancers as indicated in Table 1.1 (Kuppers, 2003).

	Expressed genes	Associated disease(s)
Latency 0	EBERs, LMP2A	-
Latency I	EBERs, EBNA-1	Burkitt lymphoma, Primary effusion lymphoma (PEL)
Latency II	EBERs, EBNA-1, LMP1, LMP2A	Hodgkin lymphoma
Latency III	EBERs, EBNA-1, EBNA-2 LMP1, LMP2A, EBNA3s, EBNA-LP	Post-transplant lymphoproliferative disease (PTLD)

Table 1.1.: Different protein expression during latency 0-III is associated with different malignancies (Kuppers, 2003).

1.1.1.5. EBV nuclear antigen 2 (EBNA2)

The latent protein EBV nuclear antigen 2 (EBNA2) is a transcriptional activator, the first expressed protein during B-cell infection, and is essential for B-cell growth transformation (Young and Rickinson, 2004). It plays an important role in the activation of resting B-cells and stimulates their proliferation by upregulating the expression of viral genes, and by up- and downregulating cellular genes. EBNA2 does not regulate transcription by direct binding to DNA but by interacting with host proteins including transcription factors. EBNA2 is a highly promiscuous protein and undergoes interactions with many binding partners as confirmed by yeast two-hybrid screening of the EBV-EBV and EBV-human interactome (Calderwood et al., 2007). One EBNA2 interaction partner is the signal transducer and activator of transcription 3 (STAT3) (Muromoto et al., 2009a). In transcription assays, EBNA2 interacts with the transcription factor STAT3 and enhances its activity. STAT3 has been shown to be differentially expressed during infection by several different herpesviruses (King, 2013; Reitsma et al., 2013; Chung et al., 2004), and recently it was shown to enhance proliferation in B-cells from patient samples (Koganti et al., 2013). Generally then, it is involved in the immune response modulation of the host organism. The transcriptional activation by EBNA2 via a well characterized interaction with the transcription factor CBF1 (also called RBP-J κ) has been described (Henkel et al., 1994; Ling and Hayward, 1995) (Fig. 1.2). EBNA2 also activates transcription via CBF1 independent pathways (Grabusic et al., 2006).

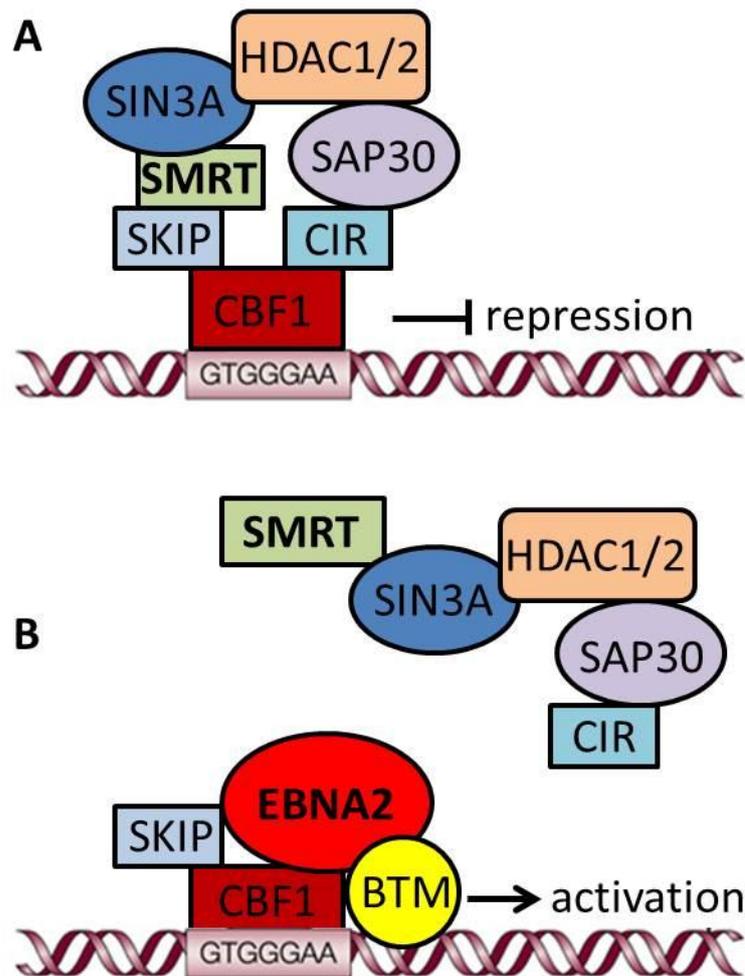


Figure 1.2: CBF1-mediated transcriptional activation by EBNA2 (A) The multi-protein complex containing SIN3A, HDAC1/2, SAP30, CIR, SKIP, SMRT and CBF1 represses transcription when bound to DNA. (B) EBNA2 competes with the HDAC-SMRT co-repressor complex for binding to CBF1 and recruits the basal transcription machinery to activate transcription. (Young and Rickinson, 2004).

EBNA2 is 487 amino acids long and bioinformatically predicted to be largely intrinsically disordered using tools such as IUPRED (Dosztányi et al., 2010). Despite the lack of three-dimensional domain structures there are still functional domains in EBNA2 which can be distinguished (Fig. 1.3) (Zimber-Strobl and Strobl, 2001). The N terminus contains a homodimerisation domain which is followed by a long poly-proline stretch (10-40 aa). The diversity region is located in the middle of the protein where the homology between the two known variants (EBNA2A and EBNA2B) is very low. The RBP-J (CBF1) binding region and an arginine-glycine stretch follow. The C terminus contains the negatively charged transactivation domain and the nuclear localisation signal.

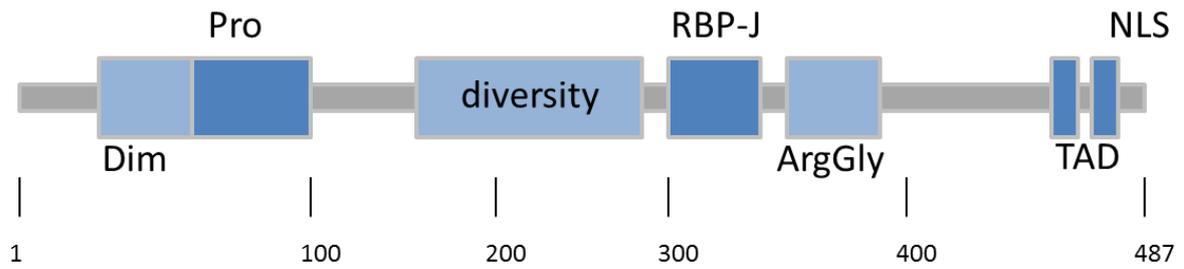


Figure 1.3: EBNA2 functional domains. Adapted from (Zimber-Strobl and Strobl, 2001). DIM = dimerisation domain, Pro = Proline stretch, RBP-J = RBP-J binding region, ArgGly = arginine glycine rich region, TAD = transactivation domain, NLS = nuclear localisation signal.

1.2. Signal transducer and activator of transcription 3 (STAT3)

STAT3 is a transcription factor that is activated due to growth factor and cytokine stimuli. The first identified cytokine that activates STAT3 signaling is IL-6 (Zhong et al., 1994). Binding of this cytokine and other growth factors to cell surface receptors leads to auto-phosphorylation of the cytoplasmic part of the receptors. Monomeric STAT3 is recruited to the phosphorylated receptor and gets phosphorylated either directly by the receptor or by associated Janus-kinases. Tyrosine-phosphorylation of the STAT3-SH2 domain causes dimerisation of STAT3 and its translocation to the nucleus where it binds to the promoter region of its target genes (Fig. 1.4). STATs use different importins for nuclear trafficking with STAT3 binding both importin- α 3 (Liu et al., 2005) and β 1 (Cimica et al., 2011). Shuttling in and out of the nucleus is presumably tyrosine phosphorylation independent (Liu et al., 2005). Besides this classical role in the nucleus, STAT3 was shown to stimulate oxidative phosphorylation in mitochondria and to support Ras-dependent oncogenic transformation (Gough et al., 2009). Another function of unphosphorylated STAT3 is to bind NF κ B and facilitate its import to the nucleus (Yang and Stark, 2008).

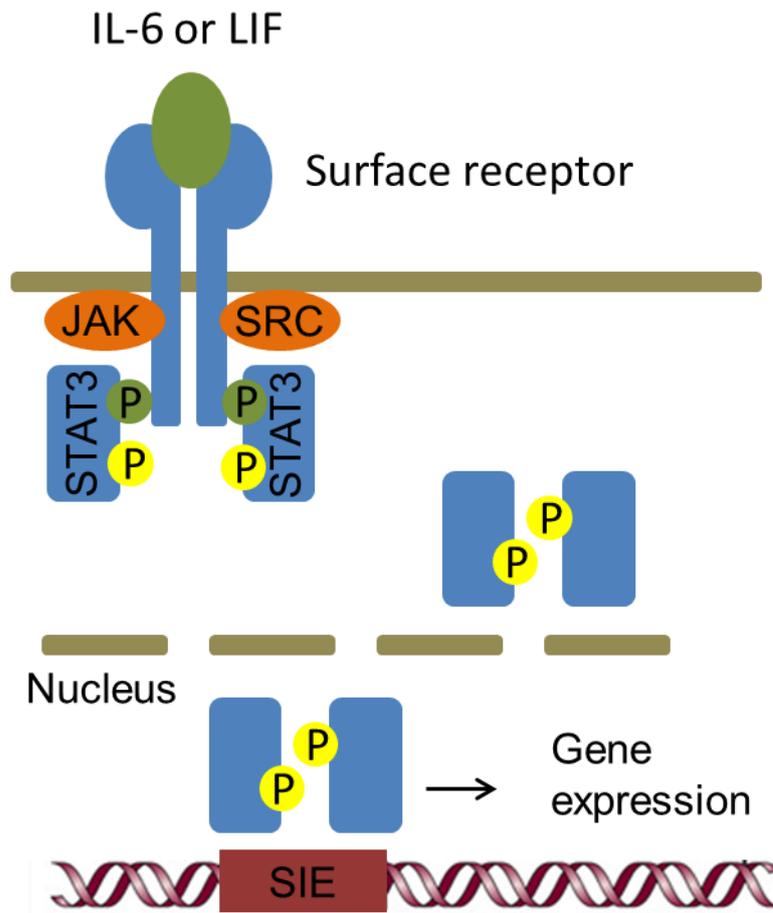


Figure 1.4: The JAK/STAT3 signaling pathway (Yu et al., 2007). IL-6 or LIF bind to the cell surface receptor leading to an auto-phosphorylation of the cytoplasmic receptor domain. STAT3 monomers are recruited to the phosphorylation site and are tyrosine-phosphorylated either directly by the receptor or by associated Janus-kinases. Phosphorylation of the STAT3-SH2 domain causes its dimerisation and translocation into the nucleus where it binds to the promoter region of its target genes

STAT3 was discovered in 1994 by two independent groups (Zhong et al., 1994; Akira et al., 1994). It is one of 7 STAT family members alongside STATs 1, 2, 4, 5a, 5b, and 6. Early functional studies (prior to the availability of structures) had defined four functional STAT3 domains: the 130 aa N-terminal domain was shown to mediate co-operative sequence specific DNA binding of multiple DNA sites (Vinkemeier et al., 1996; Xu et al., 1996). Amino acids 400-500 engage in DNA binding specificity but do not facilitate DNA binding alone (Horvath et al., 1995). The SH2 domain is located between amino acids 600-710 and mediates the dimerisation of two monomers after phosphorylation of tyrosine 705 (Shuai et al., 1994). The C terminus is associated with transcriptional activation and can be modulated by phosphorylation of serine 727 (Wen et al., 1995).

In 1998, the structure was solved of the murine STAT3 homodimer bound to DNA which shares over 99% sequence similarity with human STAT3 (Becker et al., 1998b). In the same year the structure of the N-terminal domain of STAT4, which shares a sequence similarity of about 50% with the human STAT3, was released (Vinkemeier et al., 1998a). However, solving the structure of a STAT3 monomer turned out to be a more challenging task. The murine structure was solved only in 2008 (Ren et al., 2008). Fig. 1.5 shows the crystal structures of the N-terminal domain of STAT4 (aa 1-123) and of the STAT3 β homodimer (aa 127-722) bound to DNA. The crystal structure of STAT3 shows that it is comprised of a N-terminal 4-helix bundle (in purple) followed by β -barrels (in red), a connector domain (in green) and the SH2 domain (in yellow) (Becker et al., 1998b).



Figure 1.5: Structures of N-terminal STAT4 (1-123) and C-terminal STAT3 (127-722). The STAT4 structure contains aa 1-123 (Vinkemeier et al., 1998b) and the STAT3 structure is obtained from the truncated β -isoform which is comprised of aa 127-722 (Becker et al., 1998b).

1.2.1. Functions of STAT3 and possible advantages for the virus in “hijacking” STAT3

It is not known for certain why EBV manipulates STAT3 signaling, however there are many clues in the literature relating to the role of STAT3 in cell cycle control and immunosuppression. These functions indicate that hijacking STAT3 may be to the advantage of the viral survival strategy. A notable recent publication hypothesised that STAT3 signaling might be responsible for cell proliferation and survival in EBV infected B cells (Koganti et al., 2013). Another hypothesis claiming that STAT3 is required for maintenance of the latent state upon HSV infection may also be significant (Du et al., 2013).

1.2.1.1. STAT3 in cell cycle control

The elucidation of STAT3 functions is not simple as deletion of STAT3 leads to embryonic lethality with rapid embryonic degeneration between day 6.5 and 7.5 (Takeda et al., 1997). However, extensive studies of the STAT3 signaling pathway revealed the involvement in proliferation, apoptosis, angiogenesis and immune control. STAT3 signaling enhances proliferation through upregulation of the expression of Bcl-2, Bcl-X_L (Zushi et al., 1998), Mcl-1 (myeloid cell leukaemia-1), cyclin-D₁ and c-Myc (Bromberg et al., 1999; Rahaman et al., 2002). On the other hand STAT3 was shown to inhibit apoptosis (Shen et al., 2001; Zushi et al., 1998; Aoki et al., 2003; Bhattacharya et al., 2005) and promote angiogenesis (e.g. VEGF) (Aggarwal et al., 2006). Besides the canonical pathway and gene expression regulation STAT3 functions through interaction with other transcription factors like PPAR- γ (Wang et al., 2004), β -catenin (Hao et al., 2006), NF κ B (Yu et al., 2002; Jang et al., 2004), HIF1 α (Jung et al., 2005), Pim-1 and c-myc (Shirogane et al., 1999), c-fos (Yang et al., 2003), c-jun (Yoo et al., 2001), glucocorticoid receptors (Zhang et al., 1997) and estrogen receptors (Wang et al., 2001).

1.2.1.2. STAT3 in immunosuppression

STAT3 mediates immunosuppression at many levels. On one hand it restrains the immune response by antagonising NF κ B and STAT1 mediated expression of T helper 1 (T_H1) cytokines such as IL-12 (Kortylewski et al., 2009) and interferon- γ (INF γ) (Kortylewski et al., 2005a) which are both necessary for innate and adaptive immune response, and on the other hand, STAT3 is essential for anti-inflammatory reactions mediated by IL-6 (Berg et al., 1996). The restriction of the adaptive immune response is mainly mediated by regulatory T_H17 cells. STAT3 is essentially involved in the development of T_H17 cells which requires IL-6 and TGF β and is further promoted by IL-12 and IL-23 (Gerosa et al., 2008). STAT3 functions as the major IL-6R-dependent transcription factor in the T cell and as a transcription activator of IL-23a (Kortylewski et al., 2009).

1.2.2. STAT3 in malignancies

Up-regulation of STAT3 signaling is implicated in a wide range of human cancers including e.g. breast cancer, multiple myeloma, head and neck cancer, leukemia, lymphoma and lung cancer (Bowman et al., 2000). Furthermore STAT3 was shown to be persistently tyrosine 705 phosphorylated in nasopharyngeal carcinoma (NPC) (Lui et al., 2009; Ho et al., 2013) and also in Hodgkin's lymphoma cells STAT3 was found to be constitutively active (Chen et al., 2001) which establishes the connection between EBV associated cancer and STAT3. Dominant negative mutations in the STAT3 DNA binding domain could be identified as the cause of hyper IgE syndrome (Minegishi et al., 2007; Casanova et al., 2012).

1.2.3. STAT3 inhibitors

Aberrant activation of STAT3 signaling is a common feature of many cancer types and has established STAT3 as a target for the pharmaceutical industry. Many efforts have been made to develop novel anticancer drugs targeting STAT3 signaling including: i) direct targeting of STAT3 on the protein level; ii) targeting of STAT3 on the DNA/RNA level; iii) targeting of STAT3 signaling upstream molecules; iv) inhibition of STAT3 induced growth arrest and apoptosis of tumor cells *in vitro* and tumor regression *in vivo* (Yue and Turkson, 2008). Direct STAT3 inhibitors target the SH2 domain, the DNA binding domain or the N-terminal domain. This disrupts either dimerisation, prevents binding to DNA or disturbs the transcriptional activity (Peibin Yue and James Turkson, 2008). STAT3 signaling can be targeted using various approaches: phosphopeptides (Turkson et al., 2001) and peptidomimetics (Turkson et al., 2004) which compete for binding, small molecules (Schust et al., 2006; Xu et al., 2009) blocking activity, siRNAs (Gao et al., 2005) that inhibit target gene expression and Decoy ODN (Naruya Tomita et al., 2003) which mimics the consensus sequence of the *cis* element as well as upstream inhibitors. So far, only one small molecule inhibitor and three oligonucleotide inhibitors are in the early preclinical stage (Peyser and Grandis, 2013) and more efforts will be necessary before patients can be treated with potent STAT3 inhibitors.

1.3. The nuclear receptor co-repressor 2 (NcoRII/SMRT)

1.3.1. Transcriptional repression

A crucial cellular maintenance process is the restriction of gene expression by chromatin remodeling, DNA methylation and histone modifications. Co-repressors and co-activators have a major role in histone modifications. These modifications represent the histone code and include acetylation, phosphorylation, methylation, ADP-ribosylation, deamination, proline isomerization, ubiquitylation and sumoylation (Kouzarides, 2007). Co-repressors are usually recruited by transcription factors and act as scaffold proteins to recruit chromatin remodeling enzymes like HDACs and other repressive proteins (Perissi et al., 2010). The SMRT co-repressor builds a core complex with mSIN3 and HDACs and interacts with key components of the transcriptional initiation process (Wong and Privalsky, 1998). These co-repressor complexes mediate the basal repression of genes by unbound nuclear receptors like the thyroid hormone receptor (TR) and retinoic acid receptor (RAR) (Yoon et al., 2003).

1.3.2. Silencing mediator of retinoic acid and thyroid hormone receptor (SMRT)

The nuclear receptor co-repressor 2 (NcoRII or SMRT) is a transcriptional co-repressor that plays a crucial role in many cellular processes. SMRT consists of a conserved bipartite nuclear-receptor-interaction domain (NRID) and three independent repressor domains that can actively repress a heterologous DNA-binding domain. Each NRID contains a CoRNR box motif (Jepsen and Rosenfeld, 2002a) (Fig. 1.6).

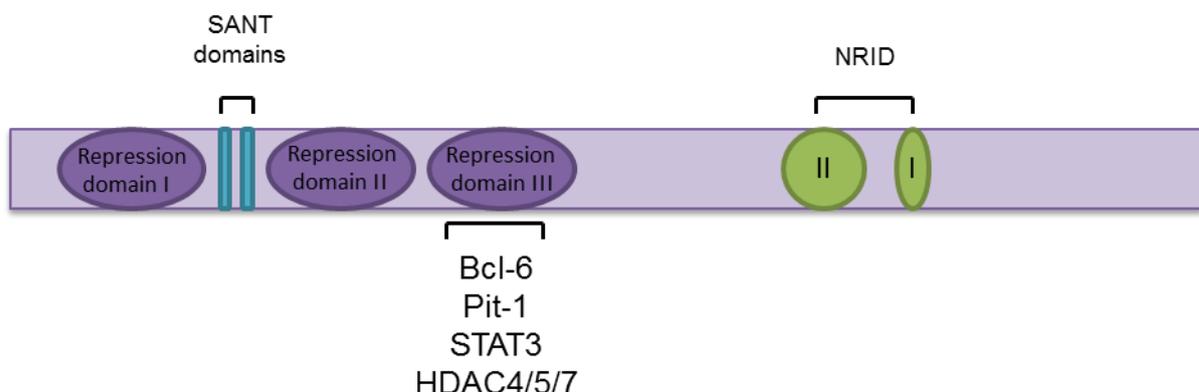


Figure 1.6: Domain localisation of the nuclear co-repressor 2 (SMRT). The three repressor domains are located in the N-terminal part of the protein. Between repression domain I and II two SANT domains are located. The C terminus contains the bipartite nuclear-receptor-interaction domain (NRID). Class II HDACs, the transcription regulator Pit-1 and the transcription factor STAT3 and Bcl-6 bind to the repressor domain III.

SMRT is a scaffold protein with 2525 amino acids in length that builds up multi-protein complexes with several other proteins like histone deacetylases, transcription regulators and factors e.g. HDAC3 (Li et al., 2000a), Pit-1 (Xu et al., 1998) and BCL-6 (Ahmad et al., 2003). Also, the signal transducer and activator and transcription (STAT3) is bound and repressed by SMRT *in vivo* (Ikeda et al., 2009).

1.3.2.1. Functions of SMRT

The promiscuous scaffold protein SMRT is involved in many cellular processes. The kind of action is usually not determined by SMRT itself but by the variation of different proteins in the co-repressor complex. The SMRT co-repressor complex can e.g. stimulate proliferation through repression of serum response factor (SRF), activator protein-1 (AP-1) and nuclear factor- κ B (NF κ B) (Lee et al., 2000). But it can also repress transcription by interaction with MAD, MyoD and HES-related repressor proteins (HERPs) (Jepsen and Rosenfeld, 2002a). SMRT has also a role in development and cell differentiation by interaction with Pit-1, Oct-1 and the Notch-activated adapter protein Su(H)/RBP-J/CBF1 (Jepsen and Rosenfeld, 2002a). Via interaction with the transcription factor and oncoprotein BCL-6, SMRT is involved in apoptosis (Dhordain et al., 1998a). SMRT was also shown to interact with STAT5 (Nakajima et al., 2001a) and STAT3 (Ikeda et al., 2009) and therefore influences the immune response.

1.4. Intrinsically disordered proteins (IDPs)

Both EBNA2 and SMRT are predicted to lack tertiary and much secondary structure, and are therefore not amenable to classical structural biology approaches. More generally, proteins can exist in three different states: completely folded, partially folded and completely disordered. Proteins in the latter two forms are referred as intrinsically disordered proteins (IDPs). About one third of all eukaryotic proteins and even over 70% of all cancer-related proteins are either completely disordered or have large segments that are intrinsically disordered (Xie et al., 2007). Viral proteins and especially herpesviruses are very rich in IDPs (Pushker et al., 2013). It is believed that the percentage of IDPs encoded in the genome increases with the complexity of the organism (Dunker et al., 2000). An explanation for this can be that IDPs evolve more rapidly and mutations occur more often because disordered regions are not constrained by their structure. A signature of intrinsic disorder is the presence of low sequence complexity and amino acid compositional bias exhibiting a low level of bulky hydrophobic and high level of polar and charged residues (Dyson and Wright, 2005a). The lack of structure may provide a number of particular features and advantages (Gsponer et al., 2008):

- Increased interaction area
- Conformational flexibility to interact with several partner proteins
- Possibility of folding-upon-binding
- Accessible sites for posttranslational modifications
- High abundance and accessibility of short linear interaction motifs

IDPs may undergo interactions with many binding partners. The first interaction is in general unspecific and weak but leads to structural changes and a folding of the binding interface which allows a more specific, high affinity interaction. This two-step mechanism is called the “fly casting mechanism” (Sevcik et al., 2007a). The fastest and easiest way to predict intrinsic disorder is the use of disorder predictors e.g. RONN and IUPRED (Yang et al., 2005; Dosztányi et al., 2005). There are nearly 20 predictors available online and several metaservers (e.g. MeDoR (Lieutaud et al., 2008)) that aggregate the results of individual predictors to predict disorder with a very low error rate.

1.4.1. Disorder properties of EBNA2

The amino acid composition of EBNA2 perfectly reflects the sequence features commonly found in IDPs. The sequence shows a low complexity indicated in particular by the large proline stretch ranging from aa 59-100 and the following arginine and glycine stretch which lies between the two transactivation domains (Cohen and Kieff, 1991). In addition, there is a high amount of negatively and positively charged amino acids attention as well as an absence of an obvious hydrophobic core.

The disorder predictor IUPRED estimates the capacity of polypeptides to form stabilising contacts. It is assumed that structured proteins undergo a large number of inter-residue interactions, whereas IDPs do not have this capacity. The formalism used by IUPRED is based upon a 20 by 20 energy predictor matrix which involves the chemical property of one amino acid as well as the chemical properties of the surrounding amino acids to calculate the binding capacity for each amino acid. The disorder diagram for EBNA2 generated with IUPRED indicates that the protein is almost fully disordered (Fig. 7) as the probability for disorder is almost exclusively over the threshold of 50%.

A new tool linked to IUPRED is ANCHOR which calculates the probability that a region within an IDP undergoes interactions with other proteins by estimating the feasibility that a given residue can form enough favourable interactions with globular proteins upon binding. The ANCHOR prediction for EBNA2 shows that the protein is predicted to undergo multiple interactions with several interaction domains distributed over the whole protein sequence (Fig. 1.7). The prediction could be verified by yeast-two-hybrid screens performed on the EBV-EBV and EBV-Human interactome (Calderwood et al., 2007) and by already known binding partners in the literature.



Figure 1.7: Disorder prediction of EBNA2. Disorder (red line) and binding site (blue line) prediction diagram generated with IUPRED (<http://iupred.enzim.hu/>)

1.4.2. IDPs in malignancies

Neither EBNA2 nor SMRT have been well studied at the structural level, however the importance of IDPs in various pathologies including cancer and neurodegenerative diseases is clear from other better understood systems.

Disordered proteins are strongly linked to cancer due to their active roles in proliferation, apoptosis, cell cycle control and angiogenesis. The most prominent example is p53 which is one of the mostly studied oncoproteins (Muller and Vousden, 2013). The tumor suppressor protein p53 is known as the “guardian of the genome” based on its broad spectrum of activity and ability to maintain genome integrity (Xue et al., 2013). Another example is BRCA1 which is mutated in many types of breast cancer. Like p53 BRCA1 is involved in many cellular processes including cell cycle control, transcription, stress response and the DNA damage response (Foulkes and Shuen, 2013). Both proteins, p53 and BRCA1, contain structured domains but also significantly disordered regions (Uversky et al., 2008). As a typical feature of IDPs they are very promiscuous and interact with multiple binding partners. STRING, a protein-protein interaction database, shows over 100 interaction partners for BRCA1 with a confidence score of at least 0.9. For p53 there are over 300 interacting proteins suggested.

Another group of intrinsically disordered proteins is involved in neurodegenerative diseases like Alzheimer’s disease (AD), Parkinson’s disease (PD), Huntington’s disease, and the prion disease (Dunker et al., 2008). A notable example is the microtubule associated protein tau which has been investigated extensively with regard to its disordered character. In healthy humans tau is intrinsically disordered. The partial transition from random coil to β structure followed by aggregation into paired helical filaments (PHFs) causes Alzheimer’s disease (Jeganathan et al., 2008). With just 11 suggested binding partners in STRING (confidence score >0.9) tau does not undergo as many interactions as p53 and BRCA1. Reasons might be that p53 and BRCA1 are transcription factors which interact with and can be modulated by multiple other proteins or that cancer related proteins are in general under more intense investigation.

1.5. Project origin and objectives

In 2009, the same year I started my PhD work, two publications describing the competing interaction of the cellular IDP SMRT and the viral IDP EBNA2 with the cellular transcription factor STAT3 (Ikeda et al., 2009; Muromoto et al., 2009a). The authors describe how under physiological conditions SMRT bound STAT3 and suppressed its transcriptional activity. Upon transfection of cells with EBNA2 expression plasmids, the viral protein appeared to release SMRT from the complex resulting in enhanced STAT3 mediated target gene expression (Fig. 1.8). The hypothesis was that increasing STAT3 activity might be expected to enhance several immunosuppressive, pro-proliferative and anti-apoptotic effects in the host and might therefore prove to be a new and important survival strategy for Epstein-Barr virus.

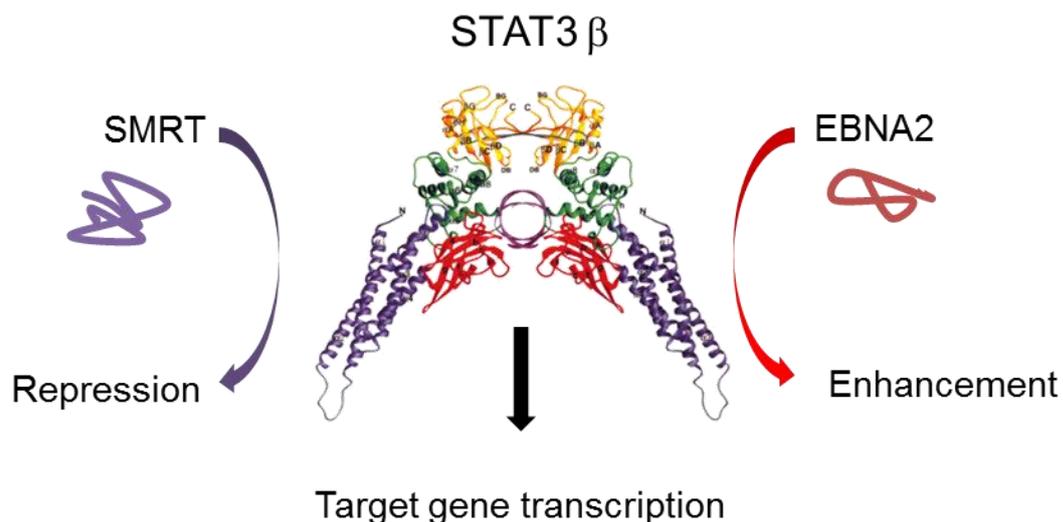


Figure 1.8: Interaction of SMRT and EBNA2 with STAT3 according to (Ikeda et al., 2009; Muromoto et al., 2009a). Binding of SMRT to STAT3 suppresses its transcriptional activity. EBNA2 releases SMRT from the complex and enhances STAT3 transcriptional activity.

The interactions were studied in limited detail by co-immunoprecipitation experiments from total cell lysate and luciferase reporter assays. As this was the first description of the interaction between these three proteins the first task was to confirm the interaction *in vitro* from purified components. For this purpose an expression library of random fragments of EBNA2 was generated, and screened for expression of purifiable fragments. The interaction of SMRT-STAT3 and EBNA2-STAT3 could be confirmed by various complementary approaches (e.g. SPR, NMR and cell based assays). Next steps involved the identification of the approximate binding region in SMRT and EBNA2 interacting with STAT3. The interacting

SMRT fragment 39L23 was assigned by NMR and 3 regions involved in the STAT3 interaction were identified. The STAT3-EBNA2 interaction was prioritised for detailed study due to its higher affinity and clearer mode of interaction, whilst the STAT3-SMRT remains to be characterised more fully in the future. The STAT3 interacting region of EBNA2 was mapped to a short span of amino acids and point mutants generated to confirm these data. The point mutants that are located within and close to conserved region 8 were tested for their effects in cell based assays. It was thus demonstrated that the point mutants not only disrupted the binding of EBNA2 to STAT3 *in vitro* (SPR, MST, NMR, pull-down and EMSA) but that the binding also disrupted in cells.

In conclusion this work confirms the interaction between SMRT-STAT3 and EBNA2-STAT3, assigns the binding regions and their affinities using various biophysical methods. The kinetics of the binding were determined and preliminary data obtained that describe the effects of the disrupted binding in cells. This work thus addresses one piece in a puzzle, shedding light on the mechanisms by which Epstein-Barr virus hijacks the host organism.

Results

Résumé

Interaction de STAT3 avec l'activateur de transcription viral EBNA2

La génération d'une banque d'expression de fragements d'EBNA2 à l'aide de la technologie ESPRIT est présentée. Les fragments obtenus couvrent les deux tiers de la partie C-terminale du gène d'EBNA2 et s'expriment bien avec des rendements de l'ordre de 10 mg/l de culture bactérienne. La méthode de purification employée pour purifier les fragments d'EBNA2 (se répartissant sur différentes régions de la protéine) et le dimer de STAT3 est décrite. La protéine EBNA2 est principalement prédite intrinsèquement désordonnée. Ainsi des approches biochimiques et biophysiques ont été choisies pour caractériser EBNA2 en tant que protéine intrinsèquement désordonnée (IDP). Le fragment D9 de EBNA2 présente des caractéristiques de protéine intrinsèquement désordonnée tels qu'une plus haute sensibilité à l'activité de protéase, un retard de migration sur gel SDS-PAGE, une vitesse de rétention augmentée en chromatographie d'exclusion de taille et une résistance à la chaleur augmentée. Le caractère d'IDP du fragment D9 a également été caractérisé grâce à des enregistrements de spectres HSQC en RMN. Ce fragment D9 de EBNA2 interagissant avec STAT3 a été identifié par des expériences de résonance plasmonique de surface (SPR) et de pull-down. L'étude de l'interaction s'est poursuivie par des expériences de RMN. La cinétique d'interaction a été déterminée à l'aide d'analyses de MST et de SPR, permettant de mesurer une constante K_D de l'ordre du nanomolaire. L'attribution de pics de RMN combiné à un alignement de séquences d'homologues d'EBNA2 suggèrent que la région conservée (CR) 8 correspond à la région d'interaction. Cette hypothèse a pu être vérifiée et confirmée grâce à des mutants de délétion. L'exacte région d'interaction a pu être définie grâce à des peptides portant une étiquette GST et de mutants (mutation ponctuelle) qui ont été soumis à des analyses de SPR. Des digestions protéasiques du complexe EBNA2-STAT3 suivies par des séquençages N-terminaux et des analyses de spectrométrie de masse ont suggéré que le site d'interaction avec l'ADN de STAT3 pouvait correspondre à la région d'interaction. Cependant des expériences EMSA ont montré que l'interaction de EBNA2 avec STAT3 n'empêchait pas l'interaction de STAT3 avec l'ADN. Des essais de cristallisation ont été effectués avec le complexe formé par STAT3 et le peptide synthétique d'EBNA2. Les cristaux obtenus diffractaient à une résolution de 6.5 Å, résolution insuffisante pour permettre de distinguer le peptide. L'interaction entre STAT3 et EBNA2 sauvage ainsi que des mutants ponctuels a été suivie en cellules de mammifère à l'aide du rapporteur luciférase de STAT3. L'activation de la voie de signalisation de STAT3 par les mutants d'EBNA2 a sévèrement été perturbée. Une perte générale de capacité de transactivation a été exclue grâce à des essais

avec le rapporteur CBF1. CBF1 est un autre facteur de transcription bien caractérisé activé par STAT3.

Interaction d'EBNA2 avec d'autres protéines cellulaires

Il a été montré que le domaine Mynd de BS69 interagissait avec la partie N-terminale d'EBNA2. De plus, deux motifs d'interaction PXLXP de EBNA2 (383-387 et 437-441) seraient les motifs impliqués dans cette interaction. Des mutations simples et doubles ont été introduites (L385A et/ou L439A). Ces mutants ont ensuite été analysés par SPR, ce qui a permis de montrer qu'une mutation simple de l'un des motifs PXLXP suffisait à totalement abolir l'interaction. Ainsi l'hypothèse publiée auparavant suggérant une redondance du motif ne pouvait pas être vérifiée par cette méthode. Il a précédemment été montré que Med25 et la protéine de *Herpes simplex* VP16 interagissaient ensemble. Des alignements structuraux de VP16 avec EBNA2 suggéraient que le site de liaison de EBNA2 était localisé entre les acides aminés 435 et 452, région couvrant la région conservée CR8. Des études d'interaction entre les fragments F3, A11, D9 d'EBNA2 et Med25 ont révélé que les fragments A11 et D9 se liaient avec Med25 avec des affinités similaires. Cela suggère que la région d'interaction se situe plutôt au sein de la région où les deux fragments se chevauchent, couvrant la région CR7.

Interaction de STAT3 avec le co-represseur cellulaire SMRT/NCoRII

SMRT est une énorme protéine, qui avec ces 2525 acides aminés est significativement plus grosse que EBNA2. Pour réussir ce challenge, le gène de SMRT a été divisé au hasard en 6 fragments qui ont été exprimés en cellules HEK293 et testés pour la suppression de la signalisation de STAT3 à l'aide du rapporteur luciférase de STAT3. Cette division au hasard était réalisable dans la mesure où il est alors possible de ne tenir compte d'aucune limite structurale. Par chance les fragments capables de suppression de signalisation de STAT3 correspondaient à la région que couvrait la bibliothèque d'expression de fragments de SMRT préexistente au laboratoire. De deux fragments avec des localisations différentes sur SMRT, un fragment, 39L23 interagissait avec STAT3. La séquence d'acides aminés de 39L23 a été assignée par RMN. Trois régions (79-89, 106-127 et 162-173) étaient principalement affectées par l'interaction. Des mutants de délétion de ces régions n'ont montré aucun effet. Seule la délétion de deux régions couvrant les acides aminés 79-127 ont montré une augmentation de l'affinité de liaison avec un changement de mode d'interaction d'un mode à 2 étapes à un mode d'interaction 1:1, la constante K_D déterminée pour l'interaction 39L23-STAT3 passant de 120 nM à 20 nM. L'interaction de 39L23 à STAT3 ne provoque pas la libération de STAT3 de l'ADN, comme montré par EMSA, mais réprime l'activité de STAT3 par le recrutement d'autres enzymes de répression comme par exemple les HDACs.

2.1. Interaction of STAT3 with the viral transcriptional activator EBNA2

Under physiological conditions STAT3 is bound by the nuclear co-repressor SMRT (Ikeda et al., 2009). Upon infection, the complex between the two cellular proteins, STAT3 and SMRT is disrupted by the Epstein-Barr virus protein EBNA2. EBNA2 was shown to somehow release SMRT from this complex and to enhance the transcriptional activity of STAT3 (Ikeda et al., 2009). EBNA2 therefore enhances the transcription of both host and viral STAT3 regulated genes. STAT3 activity has several immunosuppressive effects in the host organism (Kortylewski et al., 2005a, 2009; Berg et al., 1996; Gerosa et al., 2008) and may therefore be of importance for the viral survival strategy of Epstein-Barr virus (Koganti et al., 2013). Viruses use numerous different mechanisms to hijack host cell functions, including commonly the mimicking of natural interaction motifs (Davey et al., 2011, 2012). The IDP characteristics of both EBNA2 and SMRT suggest that this may be the mechanism EBV uses to hijack STAT3 function, thereby manipulating the host cell immune system to establish a lifelong alliance with the host organism.

This results presented here detail (1) the generation of highly expressing, recombinant protein fragments of EBNA2 (ESPRIT library), (2) characterisation of the STAT3-EBNA2 interaction, (3) mapping of the binding region to individual amino acids, and (4) effects of disrupted binding in cells.

2.1.1. Generation of highly expressing and well behaving EBNA2 constructs

2.1.1.1. Overview ESPRIT technology

To date there has been no recombinant expression of EBNA2 reported. We hypothesise that this may be because EBNA2 is an intrinsically disordered protein (IDP) and this class of proteins often cannot be expressed full length in recombinant systems. IDPs are characterised by their low sequence complexity, absence of bulky hydrophobic amino acids necessary for formation of a hydrophobic core and a high content of polar and charged amino acids (Dyson and Wright, 2005a). In order to obtain fragments of a size and yield suitable for biophysical studies a new approach was taken which promises to be an effective and time saving alternative to the classical construct design/PCR cloning approach. An expression library of random EBNA2 gene fragments encoding polypeptides from the C-

terminal two-thirds of the protein (after the polyproline stretch) was generated using the ESPRIT technology (Yumerefendi et al., 2010, 2011). ESPRIT is an approach to obtain soluble protein fragments in high quantities from poorly understood/annotated targets (Fig. 2.1) by screening for constructs that are compatible with the host organism being used for expression. Thus a design step is unnecessary as it is substituted by an experimental search. It generates random fragment libraries of about 28,000 clones which are tested for in-frame soluble expression in a fluorescent colony blot format. All clones are ranked according to signal intensity for both tags, the N-terminal His tag and the C-terminal biotin tag with the signal from the latter predicting solubility. The 96 most promising clones are test purified with NiNTA affinity chromatography and purity, degradation and expression level then assessed by SDS-PAGE. Typically, the 24 best expressing clones are sent for sequencing to determine the construct boundaries that afford stable soluble expression.

Generation of an EBNA2 expression library (see 4.2.2)

The sequence of the synthetic EBNA2 gene used in this study was derived from the laboratory strain B95-8 that has been most studied and belongs to the EBV-1 subtype group. The codon optimised gene (Geneart; appendix) was cloned into the library vector pESPRIT002 which encodes an N-terminal 6xHis tag and a C-terminal biotin acceptor peptide (BAP). The restriction enzymes *AatII/AscI* (5' end) and *NotI/NsiI* (3' end) were used to generate 5' overhangs on the gene side and the gene was digested sequentially from both ends using exonuclease III that digests 5' overhangs, but not 3'. The remaining single strand was removed using mung bean nuclease and polished by *Pfu* polymerase treatment. The plasmids were then recircularised by ligation and recovered by transformation in Mach1 cells. After transformation into expression strain BL21 AI RIL the truncated expression clones were plated on LB agar plates and around 28,000 colonies (72 x 384 well plates) picked using automated robotics. The expression of constructs bearing the 6xHis tag and BAP was analysed by colony blot (Appendix Fig A.2) and ranked according to signal intensity. The 96 best clones were analysed by purification and SDS-PAGE (Fig. A.3) and the 24 best expressing clones further analysed by DNA sequencing of the EBNA2 gene inserts.

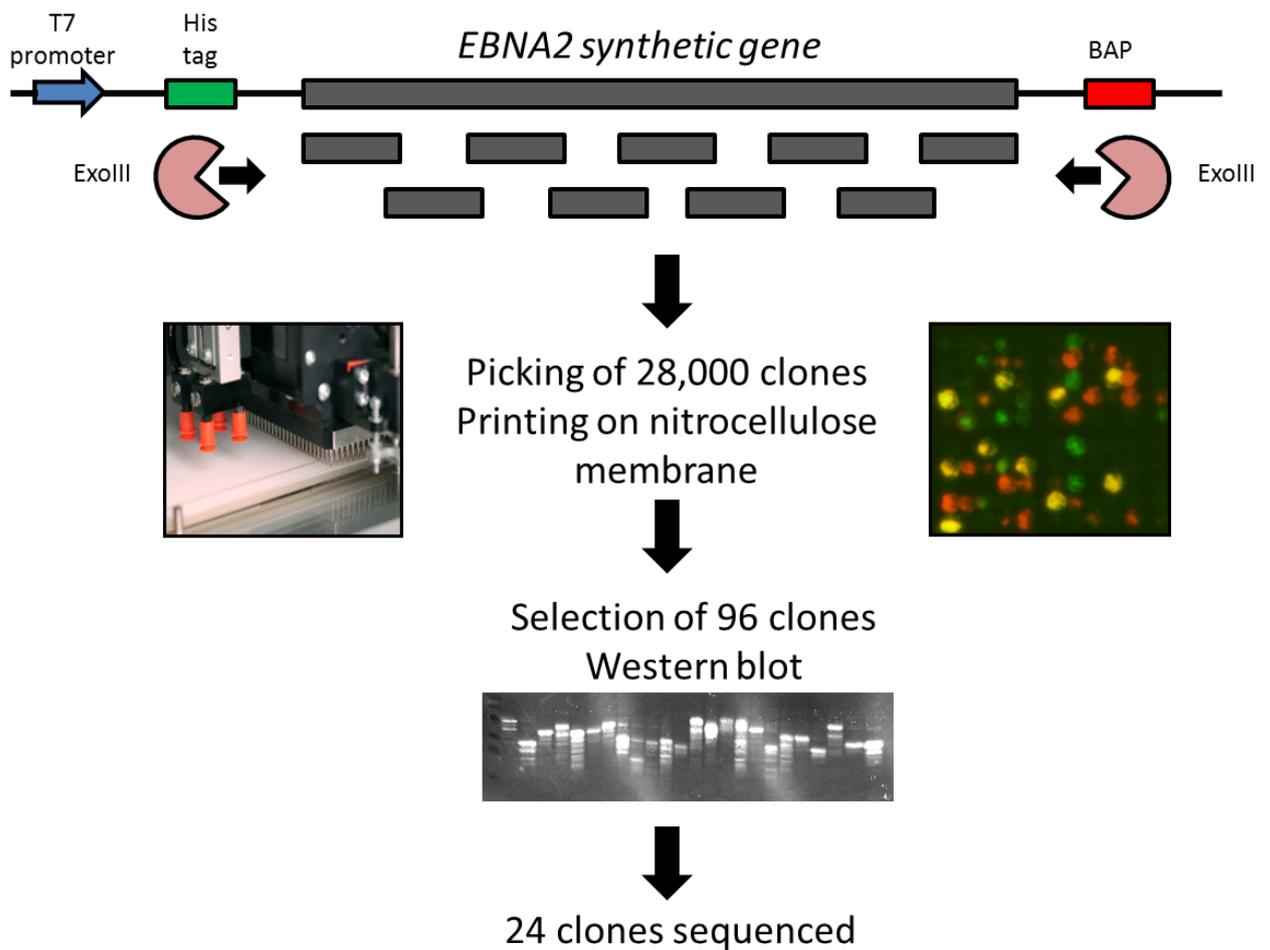


Figure 2.1: EBNA2 fragment library generation. The EBNA2 expression vector pESPRIT002 encodes an N-terminal 6xHis tag and a C-terminal biotin acceptor peptide (BAP). The restriction enzymes *AatII/Ascl* (5' end) and *NotI/NsiI* (3' end) were used to generate 5' overlapping ends and the gene was digested sequentially from both ends using exonuclease III. Around 28000 colonies were picked from LB agar plates using automated robotics. The expression of the 6xHis tag and BAP were analysed by colony blot and ranked according to signal intensity. The 96 best clones were analysed by SDS-PAGE and out of these the 24 best expressing clones were further analysed by sequencing.

Three clones were chosen according to location in the gene and soluble expression level (Fig. 2.2). Using a naming convention common in library experiments, fragments were named according to their position in the final 96-well plate that was prepared as a glycerol stock. Fragment F3 comprises aa 161-296, fragment A11 from 294-399 and the C-terminal fragment D9 344-475. The application of a random library method (ESPRIT) to IDPs is a new approach and here allowed us to obtain well-behaving sub constructs of a poorly annotated target where domain identification algorithms were of no use due to the absence of folded domains.

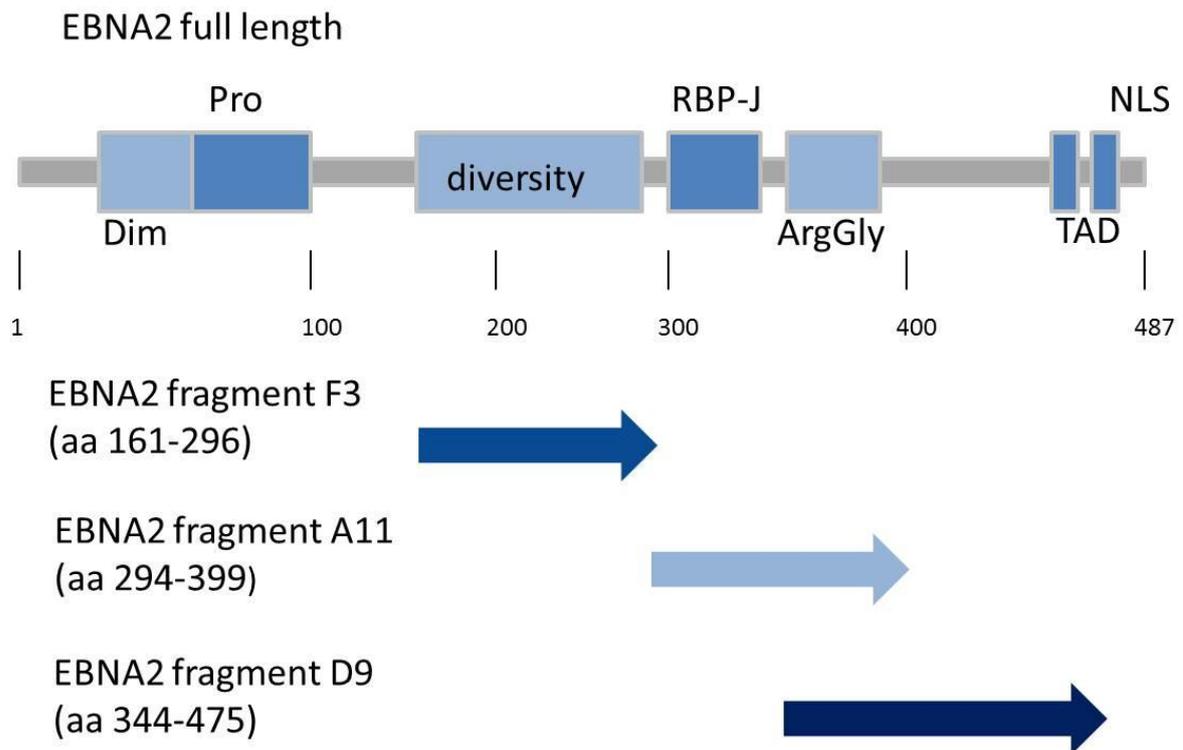


Figure 2.2: EBNA2 fragments cover the C-terminal two-thirds of the EBNA2 gene. Dim = dimerisation domain, Pro = Proline stretch, RBP-J binding region, ArgGly = arginine glycine rich region, TAD = transactivation domain, NLS = nuclear localisation signal.

2.1.2. Expression and purification of recombinant protein

2.1.2.1. Expression and purification of the STAT3 homodimer

Dimerisation of STAT3 requires phosphorylation of Tyr705 in the SH2 domain. Therefore STAT3 (127-722) was expressed in *E. coli* TKB1 cells that coexpress a tyrosine kinase which phosphorylates the SH2 domain of STAT3 (Becker et al., 1998a). As a control STAT3 was also expressed in regular BL21 cells but appeared unstable, suggesting that in the absence of phosphorylation the monomeric form does not fold into a stable form in *E. coli*. Expression and purification were performed according to the published protocol (Becker et al., 1998a). Gel filtration analysis was performed to determine if the protein was actually in the dimeric state. According to the elution profile STAT3 (127-722) was expressed as a dimer (Fig. 2.3). The purification was assessed by 12% SDS-PAGE revealing a total amount of 5 mg of at least 95% pure STAT3 dimer per litre of bacterial culture that could be concentrated to 5 mg/ml.

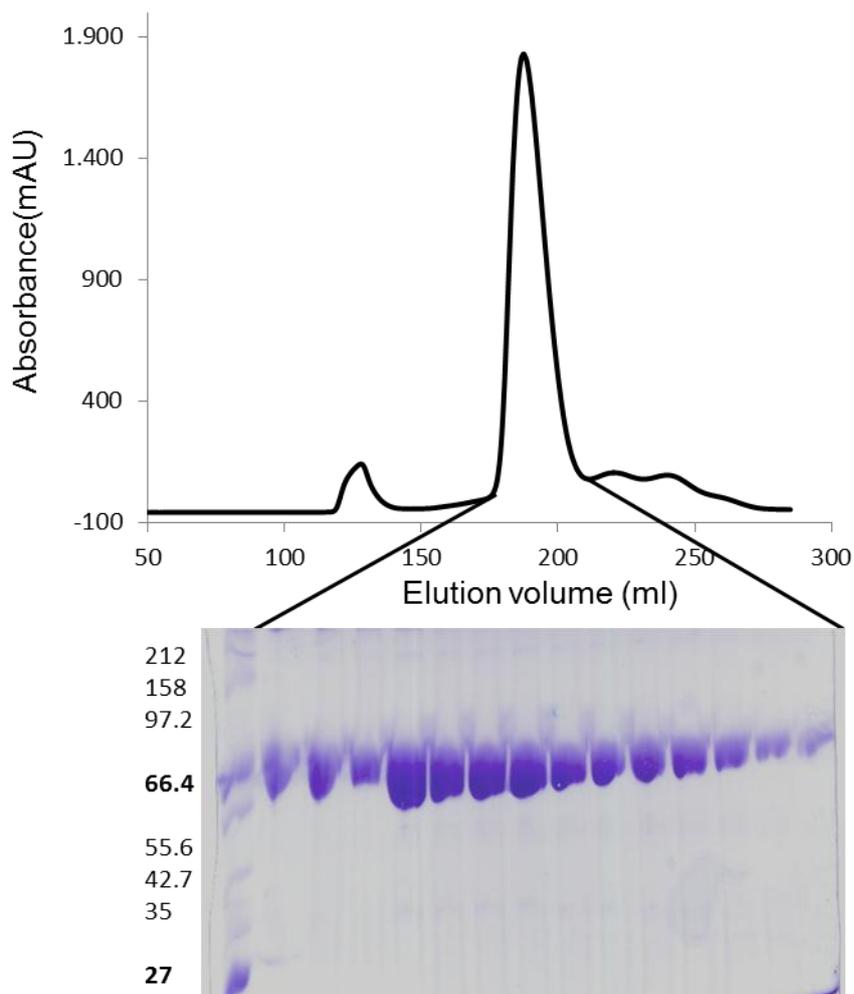


Figure 2.3: Gel filtration profile of STAT3 (127-722) expressed in *E. coli* TKB1 cells. The preparative column Superdex200 26/60 GL was used. The protein eluted at a volume of 185 ml which consistent with a 130kDa (dimer) complex. Elution fractions were loaded on a 12% SDS-PAGE.

2.1.2.2. Expression and purification of EBNA2 fragments

After deletion of the BAP that was only needed for the solubility screening procedure, the expression of the most interesting clones was scaled up. *E. coli* strain BL21 AI (Invitrogen) was used for expression and transformed as in earlier experiments. Bacteria were grown in 1 l LB medium and protein expression was induced at an OD_{600nm} of 0.6 when the expression temperature was lowered to 25 °C. Lysis was performed by ultrasonication and the soluble fraction incubated with NiNTA beads. After the first affinity chromatography purification the fractions with the highest purity were pooled and the tag was removed with TEV protease (Fig. 2.4). In order to remove the high concentration of imidazole in the elution buffer the proteins were dialyzed against washing buffer, then incubated with NiNTA beads again to remove the 6xHis tag and the 6xHis tagged TEV protease. Fragments F3 and D9 were obtained at acceptable purity using affinity chromatography but for fragment A11 it was

necessary to perform a further purification step. Size exclusion chromatography was used to remove contaminants resulting in a total purity of around 90%. The total amount of protein per litre of bacterial culture for each construct was about 10 mg (Fig. 2.5).

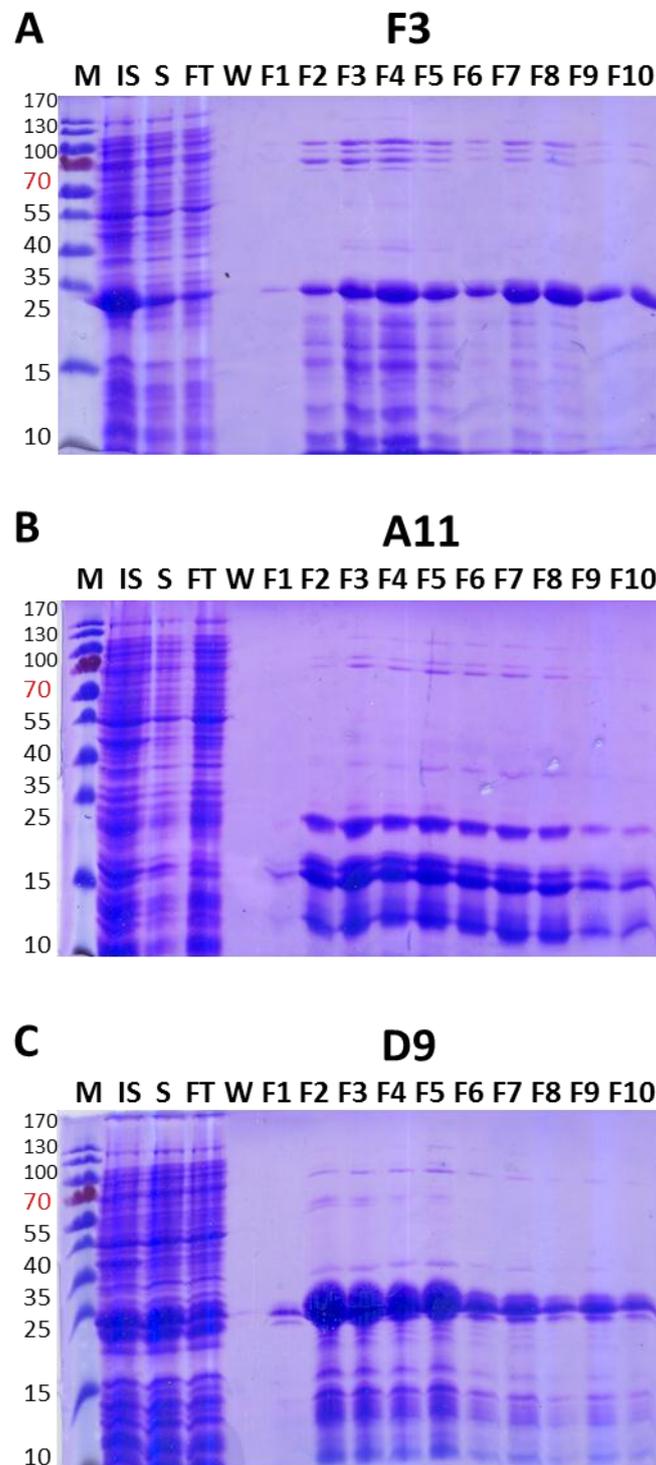
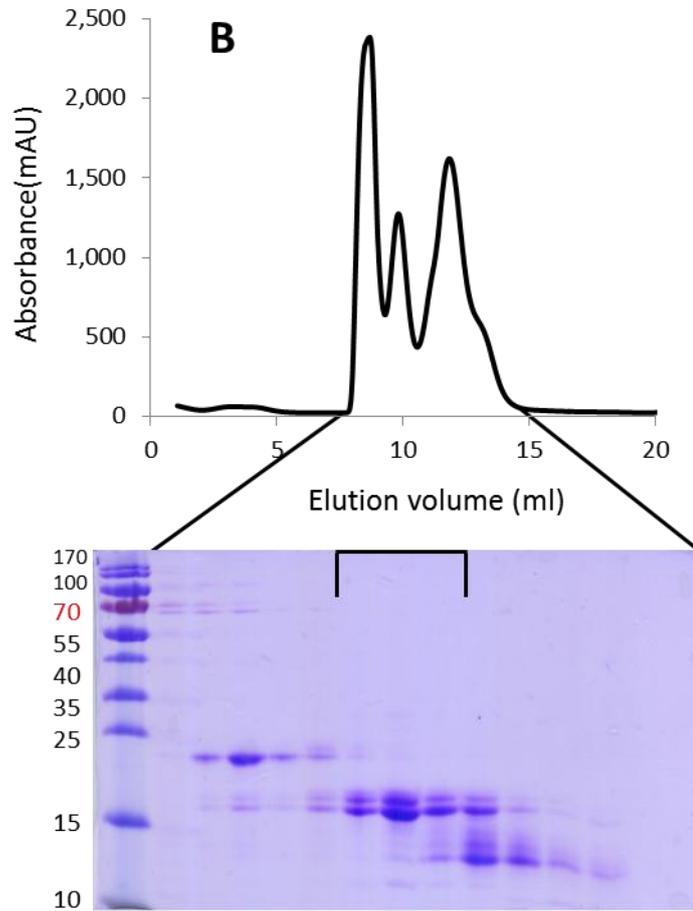
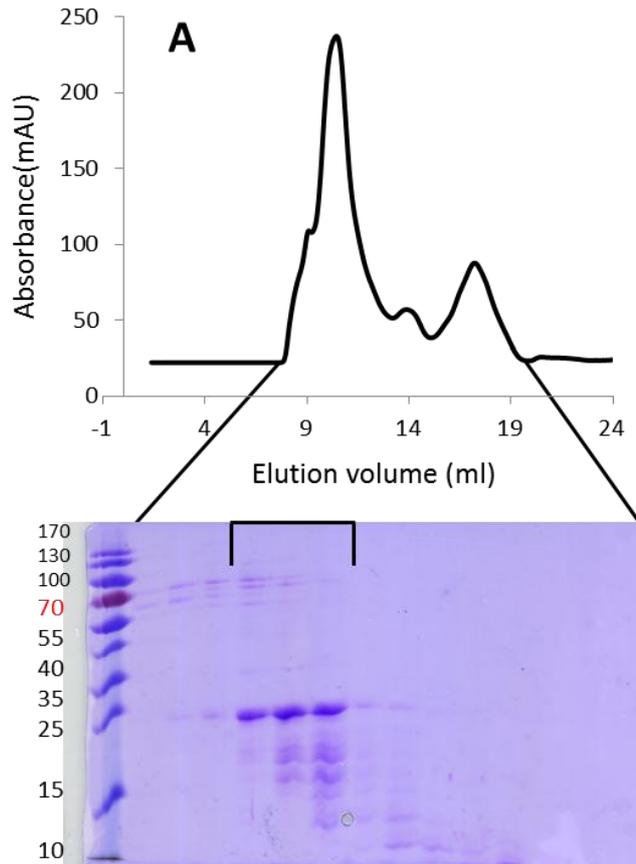


Figure 2.4: Expression and affinity purification of EBNA2 fragments A: EBNA2 fragment F3, B: EBNA2 fragment A11 and C: EBNA2 fragment D9 Abbreviations: M: Molecular weight marker, IS: Insoluble fraction, S: Soluble fraction, FT: Flow through, W: Wash, F1-10: Elution fractions 1-10.



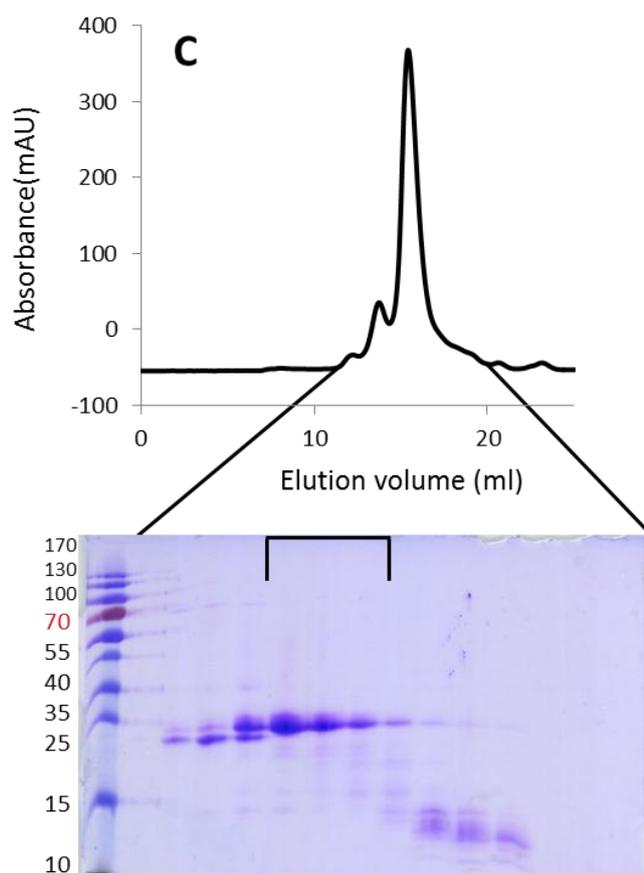


Figure 2.5: Size exclusion chromatography of EBNA2 fragments A: EBNA2 fragment F3, B: EBNA2 fragment A11 and C: EBNA2 fragment D9. Elution profile using an analytical superdex S75 10/300 column and peak fractions loaded on 15% SDS-PAGE.

All three EBNA2 fragments showed characteristic behaviors of IDPs regarding running size in SDS-PAGE and elution profile in size exclusion chromatography (Table 2.1), but were confirmed as correct by mass spectrometry.

EBNA2 fragment	Predicted size and size in mass spec (kDa)	Approximate size in SDS-PAGE (kDa)	Size in gel filtration (kDa)
F3	14.70 (14.70027)	25	13.4
A11	11.33 (11.33543)	18	28.6
D9	14.36 (14.36748)	25	32.9

Table 2.1: Predicted and experimental size of the EBNA2 fragments. Sizes predicted using ExPASy ProtParam tool, experimentally measured by mass spectrometry, estimated from 15% SDS-PAGE running behavior and calculated from elution volume in size exclusion chromatography using a superdex S75 10/300 column.

2.1.3. Structural characterisation of the IDP EBNA2

EBNA2 is predicted to be intrinsically disordered which applies to about one third of all eukaryotic proteins and 70% of cancer-related proteins. This characteristic is particularly common in many viral families, with the IDP content of herpesviruses being 17% (Pushker et al., 2013). These proteins are either fully disordered or have large (> 40 amino acids) segments (Galea et al., 2008) and our understanding of how these proteins function is limited.

2.1.3.1. Biochemical characterisation of EBNA2

IDPs have an increased interaction area with easily accessible modification sites and frequently interact with binding partners via short linear motifs (Tompa, 2012). This lack of structure imparts a high conformational flexibility and the possibility of folding-upon-binding, or at least the formation of local structure at interfaces. Characteristics of an IDP sample are 1) higher sensitivity to protease treatment; 2) aberrant migration in SDS-PAGE; 3) increased hydrodynamic radius resulting in rapid elution during size exclusion chromatography; 4) increased heat resistance.

All protein fragments used in this study (EBNA2 fragments F3, A11 and D9 as well as SMRT fragment 39L23) migrate aberrantly on SDS-PAGE and show a premature elution profile in size exclusion chromatography (Fig. 2.4 and 2.5). In contrast, the structured protein STAT3 runs in predicted size in SDS-PAGE and elutes in the predicted volume during size exclusion chromatography (Fig. 2.3).

EBNA2 D9 shows extreme heat resistance being stable at temperatures up to 90°C, whereas STAT3 is almost completely insolubilised at this temperature (data not shown). EBNA2 D9 is rapidly degraded during protease treatment (data not shown).

2.1.3.2. Biophysical characterisation of EBNA2

It is self-evident that the structure of an unstructured protein cannot be solved by x-ray crystallography, although in some cases a certain conformation might be favoured by crystal packing resulting in a structure with questionable biological significance. On the other hand, binding to a globular binding partner causes a local ordering of the IDP at the interface, or even acquisition of some level of fold of the IDP; this could potentially yield a biologically relevant structure that may crystallise.

A more widely used method to study the interaction of IDPs with their partner proteins is protein NMR spectroscopy (Fig. 2.8). With NMR it is not only possible to obtain structural information about a protein but also study the dynamics of its interactions. For assignment

experiments double labelled proteins are required (usually ^{15}N and ^{13}C) which allow the recording of triple resonance experiments. The obtained parameters are then used to calculate a structure (Cavanagh et al., 2007). An example of the structure determination of an IDP binding to partner proteins by NMR is the viral protein VP16 which gains structure while binding to TFIIB and PC4 (Jonker et al., 2004). The recording of a HSQC spectrum of a single labelled protein can show if it is structured or not. The chemical shifts of the amide protons on the x-axis depend directly on the secondary structure of a protein which results in a narrower peak distribution on the x-axis for unstructured proteins.

2.1.4. Interaction between EBNA2 and STAT3

EBNA2 is a small protein of 487 aa, but was found to undergo interactions with many cellular host and viral proteins in a large scale interaction screen (Calderwood et al., 2007). IDPs commonly interact with target proteins via short linear motifs (Davey et al., 2011) which can be high affinity in themselves (e.g. NLS peptides interact with low nanomolar affinities (Boivin and Hart, 2011) Alternatively, a first, relatively weak contact may occur following which the protein rearranges into a higher ordered, structured conformation to permit a strong and specific interaction (Dyson and Wright, 2005a; Sugase et al., 2007a). A series of experiments were thus performed to characterise the nature of the EBNA2-STAT3 interaction.

2.1.4.1. Binding analysis by surface plasmon resonance and pull-down experiments

All three EBNA2 fragments were analysed by surface plasmon resonance (SPR) for binding to STAT3. Because of its size and the dimeric character of STAT3, the EBNA2 fragments were immobilised, since dimeric STAT3 would be expected to dissociate leaving a chip-bound monomer. All EBNA2 proteins were immobilised in different flowcells of the same sensor chip at equal levels (~2000 RU). STAT3 was injected as analyte at two different concentration ranges: 1-100 nM and 125-1000 nM. Independent of the concentration range used fragment A11 did not bind STAT3, whereas fragments F3 and D9 interacted (Fig. 2.6A). No binding was observed to the chip surface in the underivatized flow cell. This result was then confirmed by pull-down experiments (Fig. 2.6B) where biotin-tagged EBNA2 fragments were bound to streptavidin magnetic beads and incubated with *E.coli* lysate containing recombinantly expressed STAT3. After several washing steps, proteins were eluted from the beads by boiling and analyzed by 12% SDS-PAGE. Proteins were blotted on a nitrocellulose membrane and detected by STAT3 antibody and Alexa 488-conjugated streptavidin. Fig. 6B shows that STAT3 was bound and pulled down by EBNA2 fragments F3 and D9 but not by the empty beads or EBNA2 fragment A11.

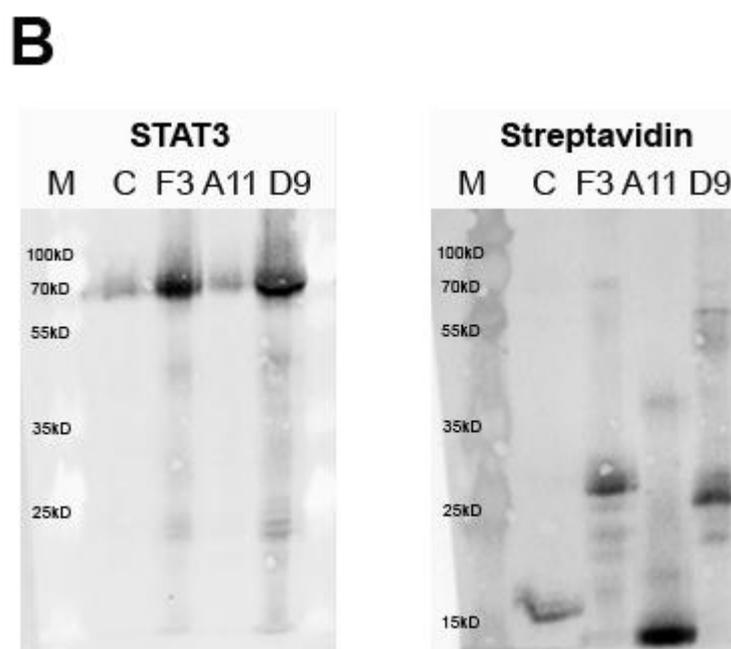
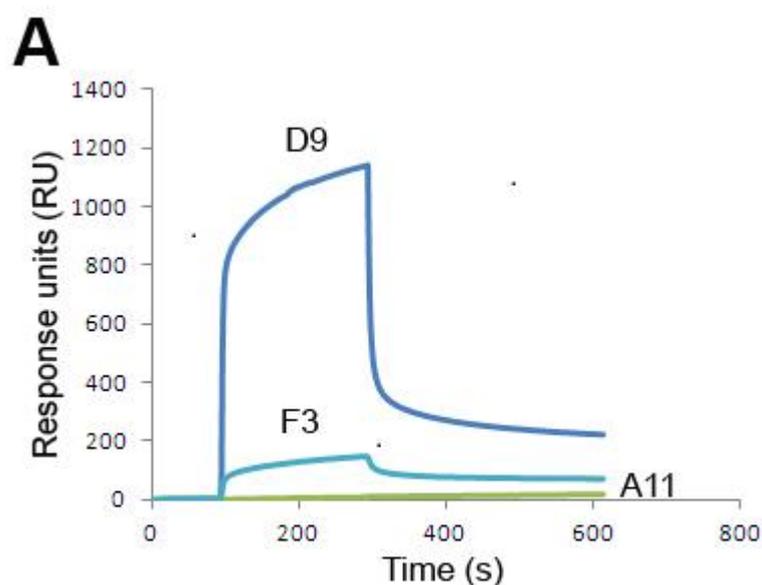


Figure 2.6: Comparative binding studies of EBNA2 fragment F3, A11 and D9 to STAT3. A: Comparison of EBNA2 fragment F3, A11 and D9 binding to STAT3 at a single injection of 1 μ M. B: Pull-down of EBNA2 fragments F3, A11 and D9 via biotin tag together with STAT3 from *E.coli* lysate. Left blot detection of STAT3 via STAT3 antibody and right loading control detection of biotin tagged proteins via streptavidin Alexa-488.

By SPR, D9 and F3 bound STAT3 with the respective affinities of $K_D=7.6$ nM (Fig. 2.7A) and $K_D=240$ nM (Fig. 2.7B). K_D values were obtained using the BIAevaluation software and sensorgrams were fitted using a two-state model (see 4.3.1). This was necessary because a 1:1 Langmuir model did not fit the data and it has been reported in other system studies that the two-state model best describes IDP interactions (Sevcik et al., 2007a) where there is an initial reversible fast encounter and then a slow rearrangement into a stable binding conformation. We thus assumed that fragments F3 and D9 bound to STAT3 via such a two-

step mechanism and may therefore conform to a “fly casting mechanism” whereby weak reversible initial interactions influence the probability of a structural rearrangement into a tight complex (Sugase et al., 2007a).

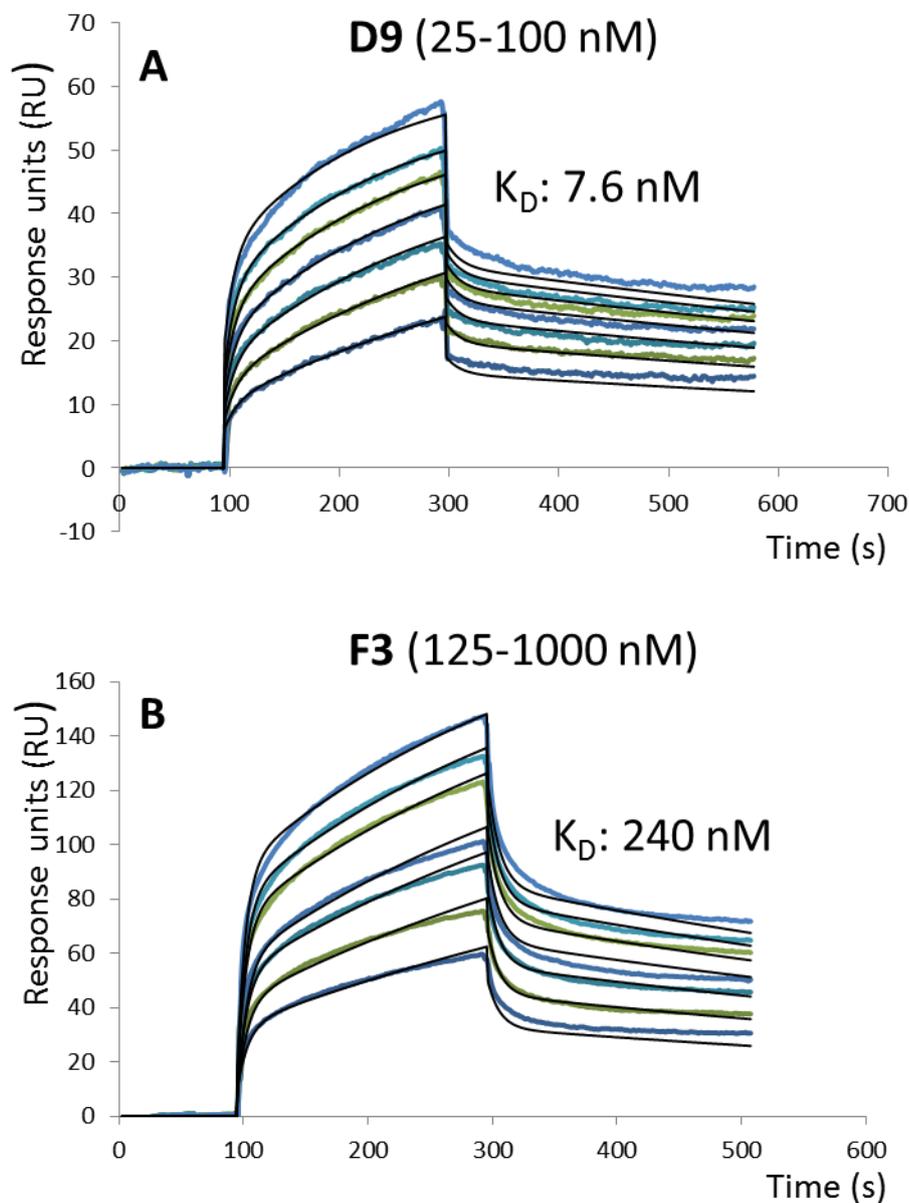


Figure 2.7: A: Sensograms with fitted curves (in black) of STAT3 binding to D9 (concentration range of injected STAT3: 25-100 nM) and B: alternative thermodynamic evaluation method applied to the same data within the linear concentration range. C: Sensograms and fitted curves (in black) of STAT3 binding to F3 (concentration range of injected STAT3: 125-1000 nM) and D: thermodynamic analysis.

2.1.4.2. Binding analysis by high field nuclear magnetic resonance (NMR)

Due to its significantly higher binding affinity, fragment D9 was studied further to confirm specific binding was occurring and characterise in detail the binding region. Heteronuclear Quantum Coherence (HSQC) spectra of ^{15}N labelled D9 alone and in complex with unlabelled STAT3 in a 1:1 ratio were recorded in collaboration with the Blackledge lab (IBS, Grenoble). HSQC spectra show the correlation between nitrogens on the y-axis and amide protons on the x-axis. Each peak in the spectrum corresponds to one amino acid with the exception of prolines that are invisible as they lack amide protons (Cavanagh et al., 2007). For D9 alone, the narrow peak distribution on the x-axis from 7-8.5 ppm indicated that this fragment was intrinsically disordered since structured proteins usually show a peak distribution from 6.5 – 9.5 ppm. This observation is in agreement with bioinformatic predictions (*IUPRED*, *RONN* and others) and further supports the choice of model used to analyse data during the SPR experiments (Fig. 2.7). The absence of β -strands and α -helices in EBNA2 fragment D9 was later confirmed by circular dichroism (data not shown).

The comparison of NMR parameters of the free and STAT3-bound state of D9 revealed significant changes. Some peaks in the complex spectrum weakened or disappeared, other peaks shifted (Fig. 2.8) which was a clear indication for STAT3 binding. Shifting peaks indicate a conformational change and vanishing peaks indicate changes in the dynamic rate of the protein. STAT3 is significantly larger than the EBNA2 fragment D9. Therefore the region in D9 binding to STAT3 tumbles in the same speed as STAT3 thus causing weaker and vanishing peaks in the spectrum.

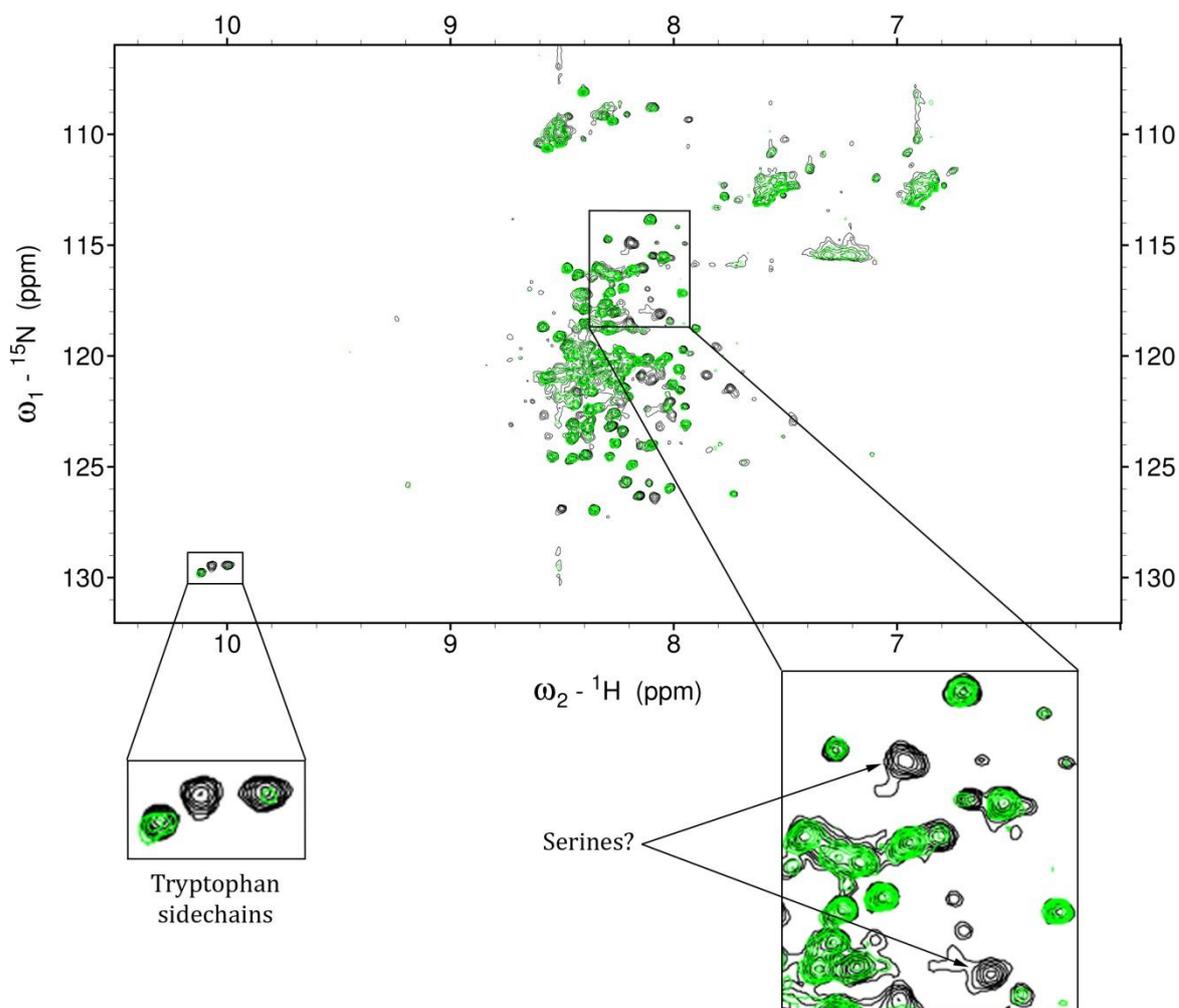


Figure 2.8: NMR spectra of D9 alone (black) and D9-STAT3 complex (green) upon titration of STAT3. The fragments were labelled with ^{15}N and measured at a concentration of above 200 μM in a volume of 250 μl . The ratio between D9 and STAT3 was 1:1.

2.1.5. Identification of the STAT3 binding region in EBNA2

The assignment of the HSQC spectrum was not performed because of the quality of the data and time constraints. Nevertheless, some specific amino acids like glycine, asparagine, glutamine, serine and tryptophan are readily distinguished in a HSQC spectrum without assignment. Therefore alignments were used to deduce the possible binding region in EBNA2 relative to conserved regions (CR). Since fragment A11, which overlaps with the N-terminal half of fragment D9, does not bind STAT3, only amino acids 400 to 475 were considered to be involved in STAT3 binding. The C-terminal sequence of D9 was aligned with the sequences of EBNA2 homologues (human EBV type 1 and 2, baboon LCV and rhesus LCV) and a region of interest defined corresponding to CR8 (Fig. 2.9). Also, two of three tryptophans within protein fragment D9 vanished and weakened respectively and two tryptophans could be located within and in close proximity to CR8. In total EBNA2 comprises

9 CRs whereas CR1-4 can be found in the N-terminal region and are associated with self-assembly (Harada et al., 2001). Separated from CR1-4 by the “diversity region” are CR5-9 which are located in the C-terminal part of the protein; these are associated with transcription regulation (Cancian et al., 2011).

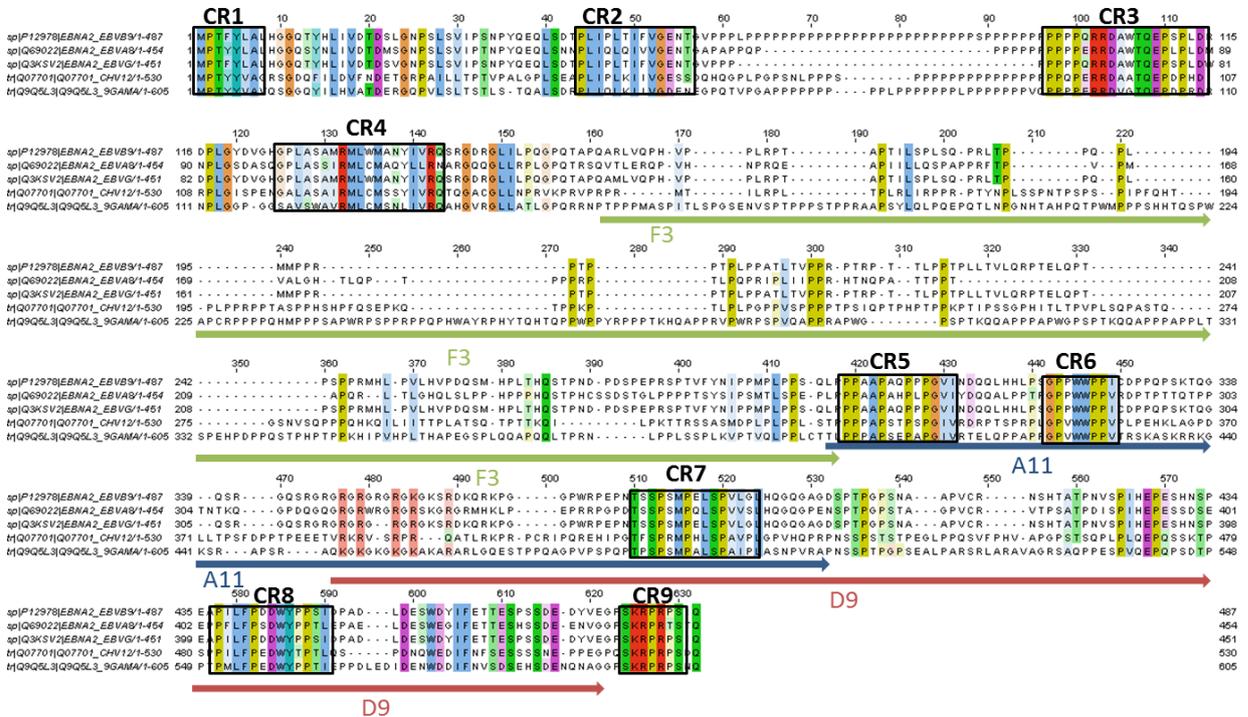


Figure 2.9: Sequence alignment of EBNA2 with indication of CR1-9 and location of the library fragments. Comparison with EBV type 1 and 2, baboon LCV and rhesus LCV reveals 2 CRs in fragment D9.

To test this hypothesis from sequence alignments a deletion mutant lacking CR8 was generated by overlap extension PCR and tested for binding to STAT3 by SPR and microscale thermophoresis (MST) (Fig. 2.10). For SPR experiment, D9 wt and D9 Δ CR8 were immobilised on a CM5 chip at equal levels (2500 RU) and STAT3 was injected at different concentrations with a maximum concentration of 1 μ M. D9 Δ CR8 did not show any binding at any given concentration.

As the binding analysis of the SPR data using the two-state evaluation model is more complicated than a simple 1:1 binding mode, it was desirable to complement this binding analysis approach with another one. MST is a relatively new method which measures motions of proteins in microscopic temperature gradients and detects thermophoretic changes in conformation, charge and size. As such, it is ideal for measuring protein-protein interactions (Jerabek-Willemsen et al., 2011). In contrast to SPR it is a free solution method comparable with ITC (Wienken et al., 2010), but with far lower sample requirements. For the MST experiment STAT3 was fluorescently labelled with an amine reactive dye (see 4.3.2 for more detail).

Both methods showed that the deletion mutant Δ CR8 abrogates binding to STAT3. The K_D obtained by MST was 70 nM; almost 10 times higher than the value obtained by SPR. This difference can be explained by the different methods of K_D calculation. The BIAevaluation software calculates the K_D using a 2 step kinetic model which takes fast and slow association/dissociation into account (see 4.3.1 for more detail) whereas MST measures mobility differences between free and saturated binding. Nevertheless, both values are in the lower nanomolar range and point to a highly specific binding event, consistent with the NMR results.

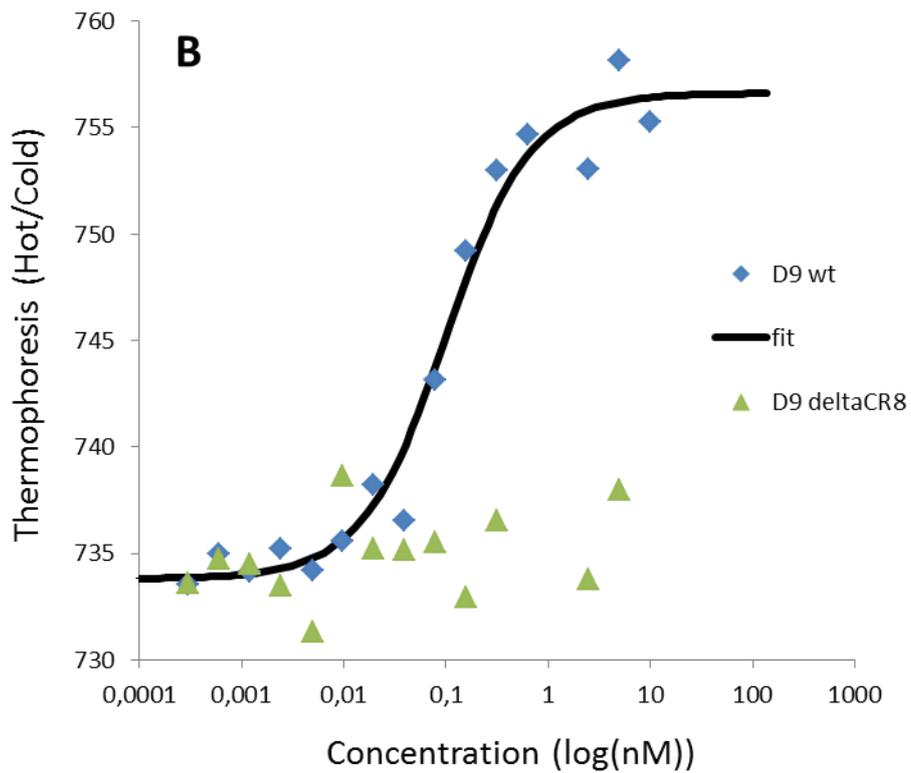
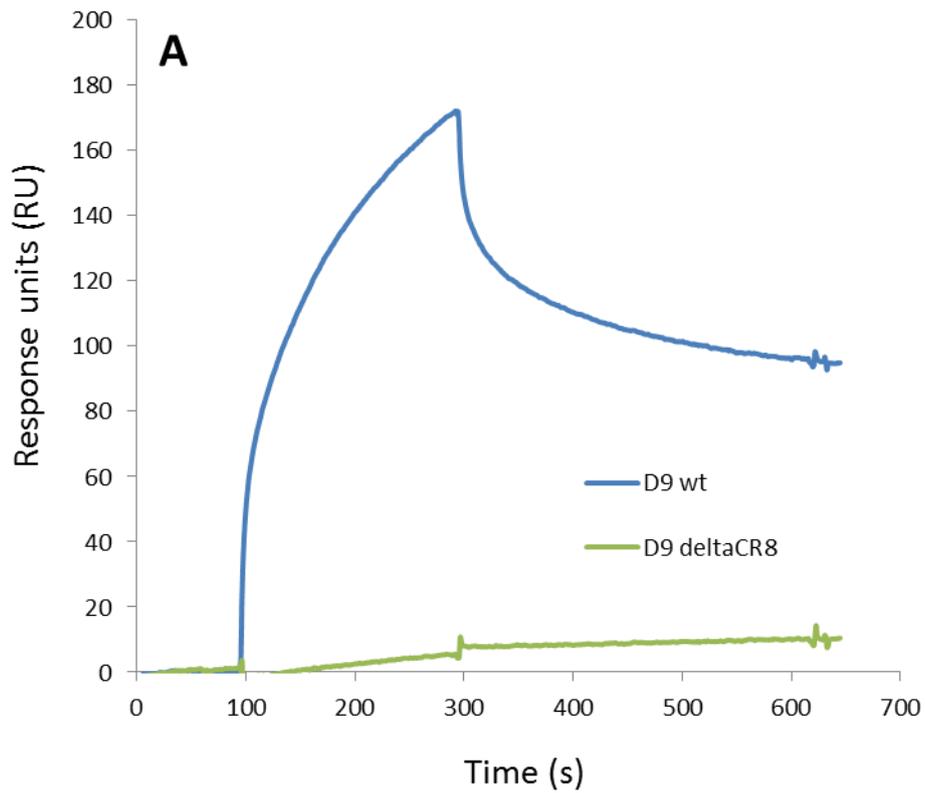


Figure 2.10: Comparison of fragment D9 wild-type and D9 Δ CR8 binding to STAT3 A: SPR analysis at a single injection of 100 nM. B: MST analysis with a STAT3 concentration range of 10 μ M to 0.3 nM.

2.1.6. Identification of EBNA2 residues involved in the binding to STAT3

In order to identify individual amino acids that mediate STAT3 binding, alanine mutants were generated using the QuikChange mutagenesis strategy (Stratagene). The alanine scan of CR8 sampled all amino acids except prolines. All mutants were tested for binding to STAT3 by SPR and MST. One mutant, Ser448 in CR8, showed a significantly reduced binding affinity. A synthetic peptide composed of the amino acid sequence of the minimal CR8 motif was ordered and tested for binding using MST (Table 2), however no binding was observed suggesting an involvement of further amino acids flanking regions of CR8. Therefore, several GST fusion peptides containing CR8 plus additional sequence N- and/or C-terminal to CR8 were cloned and expressed. To measure binding, the GST tagged peptide was captured on a Biacore CM5 chip with immobilised GST antibody (Hutsell et al., 2010) and STAT3 injected in different concentrations. The results showed that the amino acids located on the N-terminal side of CR8 are not required for binding whereas deletion of the C-terminal amino acids diminished or abolished binding, thereby demonstrating that CR8 plus 12 amino acids on the C-terminus are essential for the binding to STAT3 (Table 2.2).

Peptide name	Sequence	Binding affinity
D9TA-FL	IHEPESHNSPEA <u>PILFPDDWYPPSI</u> DPADLDESWDYIF	+++
D9TA-1	PESHNSPEA <u>PILFPDDWYPPSI</u> DPADLDESWD	+
D9TA-2	HNSPEA <u>PILFPDDWYPPSI</u> DPADLDE	-
D9TA-3	IHEPESHNSPEA <u>PILFPDDWYPPSI</u> DPADLDESWD	+
D9TA-4	IHEPESHNSPEA <u>PILFPDDWYPPSI</u> DPADLDE	-
D9TA-5	PESHNSPEA <u>PILFPDDWYPPSI</u> DPADLDESWDYIF	+++
D9TA-6	HNSPEA <u>PILFPDDWYPPSI</u> DPADLDESWDYIF	+++
CR8	<u>PILFPDDWYPPSI</u>	-

Table 2.2: Sequence and STAT3 binding affinity of GST-EBNA2 fusion peptides. EBNA2 peptide sequence (amino acids of CR8 are indicated in bold and underlined).

The alanine screen was extended to the C-terminal amino acids and all point mutants, the wild-type D9 protein as a positive control and the CR8 deletion mutant as a negative control were analysed by SPR and MST. Four mutations showed the greatest effects and a strongly reduced binding affinity: S448A, W458A, D459A and Y460A (Table 2.3).

Mutant	Binding capacity
Wild-type	+++
ΔCR8	-
I438A	+++
L439A	+++
F440A	+++
D442A	++
D443A	++
W444A	+++
Y445A	+++
S448A	+
I449A	+++
W458A	+
D459A	+
Y460A	+
F462A	+++

Table 2.3: Location of the mutation in EBNA2 fragment D9 and binding capacity. Critical amino acids are indicated in bold.

2.1.7. Identification of the EBNA2 binding region in STAT3

The STAT3 binding site of EBNA2 was thus identified and critical residues determined. To identify the complementary EBNA2 binding region on STAT3, the STAT3 homodimer, EBNA2 fragment D9 and the complex of D9 and homodimeric STAT3 were digested with trypsin, elastase and chymotrypsin. The digestion patterns were compared to identify protected regions of STAT3. After digestion with trypsin two fragments (one protected and the other one not) were visible and analysed by N-terminal sequencing and mass spectrometry (J.-P. Andrieu & L. Signor, IBS). For this, single proteins and complex were digested for 5, 15, 30 and 60 min at 21 °C. The digest was stopped by addition of Laemmli buffer to the samples and heating to 95 °C. Aliquots were electrophoresed on 12% SDS-PAGE and protein fragments were visualised by Coomassie blue staining. For N-terminal sequencing, the fragments of interest were transferred to a PVDF membrane, whilst for mass spectrometry analysis frozen aliquots containing all protein fragments plus peptidase were analysed. Both were found to have the same N-terminus but different C-termini suggesting that this is the binding region. Fig. 2.11B shows the location of the putative ENBA2 binding region of STAT3; it is located in a loop in the DNA binding domain.

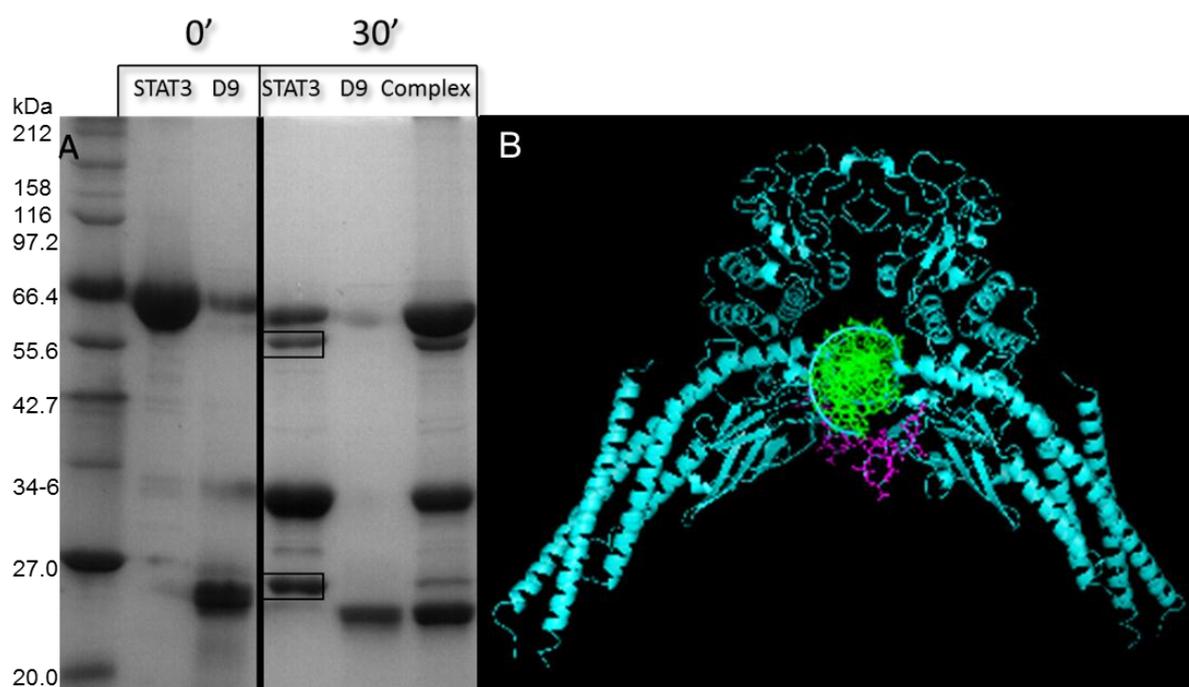


Figure 2.11: A: Digestion pattern of the STAT3-D9 complex and location of the binding region within the STAT3 dimer. Digest with 30 ng trypsin of D9 (60 µg), STAT3 (20 µg) and STAT3-D9 complex at different time points and at 21 °C. Protein fragments marked with a black box were analysed by N-terminal sequencing and mass spectrometry. B: The combination of N-terminal sequencing and mass spectrometry suggests a possible binding side for EBNA2 fragment D9 in STAT3 (pink sticks). PDB: 1BG1.

2.1.8. Binding of EBNA2 D9 does not prevent DNA binding

This raised the question of whether EBNA2 prevents DNA binding by STAT3. In order to test this hypothesis a STAT3-DNA electromobility shift assay (EMSA) was performed using fluorescent oligos. Purified STAT3, labelled target site oligo and/or wild-type or Δ CR8 D9 protein were incubated for 30 min and electrophoresed on a 6% native polyacrylamide gel (Fig. 2.12). Lanes 2-4 show that STAT3 binds DNA but EBNA2 does not. Lanes 5-7 show the STAT3-EBNA2 D9 complex binds DNA albeit with a different mobility and suggests a binding ratio of 1:2 (1 dimer STAT3: 2 D9). Protein D9 Δ CR8 does not seem to interact with STAT3 when used at the same concentrations as D9 (lanes 8-10).

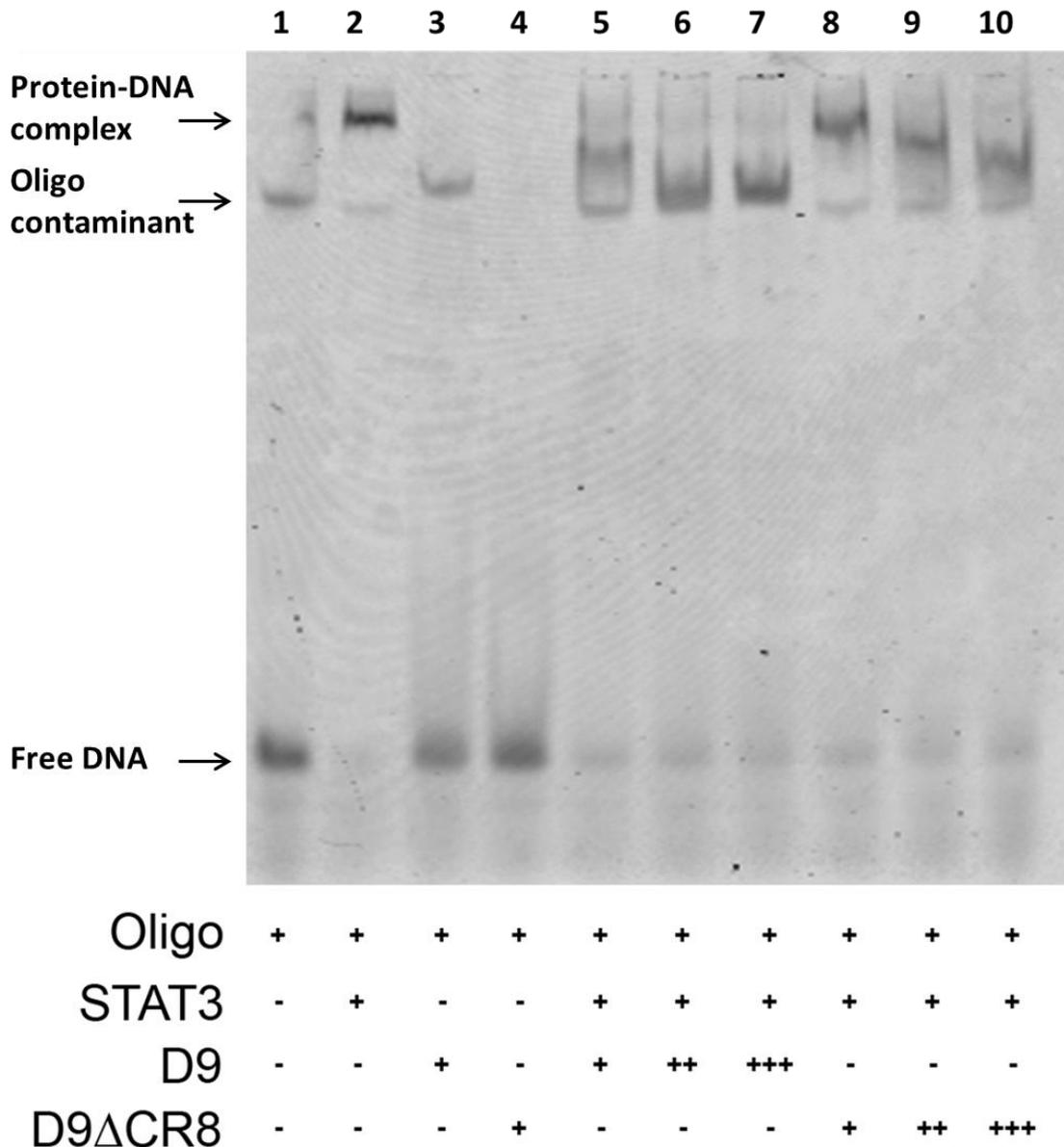


Figure 2.12: EMSA of STAT3, D9 and D9 Δ CR8 as well as STAT3-D9 and STAT3-D9 Δ CR8 complexes in different ratios bound to STAT3-binding oligo labelled with Alexa488. Protein-DNA complexes were run on a 6% native gel for 2 h at 4°C. Ratios of D9 and D9 Δ CR8 in relation to a STAT3 dimer (STAT3:D9): + = 1:1, ++ = 1:2, +++ = 1:4.

2.1.9. First crystallisation trials

First crystallisation trials were carried out in the hope of solving the crystal structure of the STAT3-EBNA2 complex. Crystallisation drops were set up robotically at the EMBL crystallisation service (HTX lab, EMBL Grenoble). STAT3 homodimer at 5 mg/ml was mixed with DNA alone and with DNA and the binding peptide D9TA-6 (see Table 2). In total, 576 crystallisation solutions were tested with drops of 100 nl volume were set robotically using the sitting drop method at room temperature. Over 70 crystallisation conditions resulted in crystals of mainly bi-pyramidal shape in the size of 40-100 μm (Fig. 2.13 A + B). The growth time ranged from 1-7 days.

Four data sets were collected at the PROXIMA I beamline (SOLEIL, Paris) with a maximum resolution of 6.5 \AA (Fig. 2.13C). The resolution would not suffice to see the peptide in the structure.

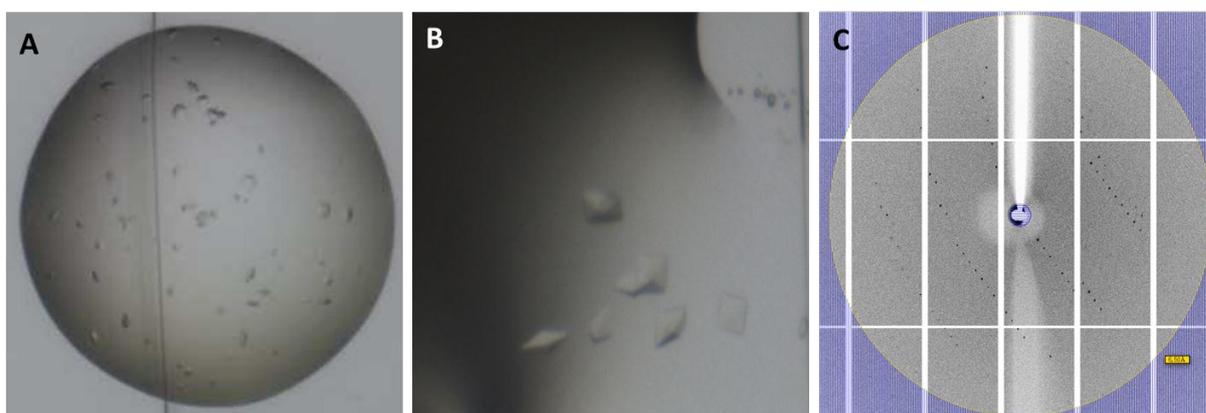


Figure 2.13: Crystals and diffraction pattern of STAT3 homodimer and EBNA2 binding peptide. Crystals obtained in A: 0.8 M ammonium sulphate and 0.1M MES pH6 and B: 25% ethylene glycol, C: Diffraction pattern with 6.5 \AA resolution.

2.1.10. Cellular assays investigating the STAT3-EBNA2 interaction

Prior to this work, the interaction between STAT3 and EBNA2 had been described using luciferase-STAT3 promoter assays and co-immunoprecipitation (Ikeda et al., 2009; Muromoto et al., 2009a) with no indication of binding regions or affinities. Here we were able to obtain recombinant proteins to study the interaction *in vitro* using various biophysical methods and determine binding affinities, the binding region and crucial amino acids. Furthermore, EBNA2 could be characterised structurally by NMR. It was thus desirable to study the interaction again in cell based assays to see whether the point mutations of EBNA2 that affected the STAT3 interaction *in vitro* show effects *in vivo*.

2.1.10.1. Generation of stable cells expressing the luciferase gene under a STAT3 promoter element

Lentiviral particles are widely used tools to generate stable cell lines (Matrai et al., 2010) and permit the gene of interest to be stably inserted into the cellular genome. A STAT3-luciferase HEK293 cell line was generated using lentiviral particles carrying the firefly luciferase gene under a STAT3 promoter. After transfection cells were kept under puromycin selection pressure and single cells isolated using the limiting dilution method to obtain a stable, monoclonal cell line. The generation of monoclonal cell lines is required to exclude the possibility that observed changes might be caused by the genome integration and not directly by modulation of the STAT3-EBNA2 interaction. Leukaemia inhibitory factor (LIF), an activator of STAT3-signaling, was used to challenge the cells and to stimulate STAT3 signaling.

Cells were transfected with the KG172 full-length EBNA2 expression plasmid in increasing concentrations to confirm a dose-dependent effect and identify the concentration of plasmid that could be used to test mutants (Fig. 2.14A). Consequently, the S448A and Δ CR8 EBNA2 mutants were analysed at a constant concentration (Fig. 2.14B) and in the context of the full-length protein. Dose dependent activation of STAT3 signalling by EBNA2 expression was observed (Fig. 2.14A), consistent with previous data using transient transfections (Muromoto et al., 2009a). Differences in the activation potential of STAT3 between EBNA2 wild-type and mutants were measurable (Fig. 2.14B).

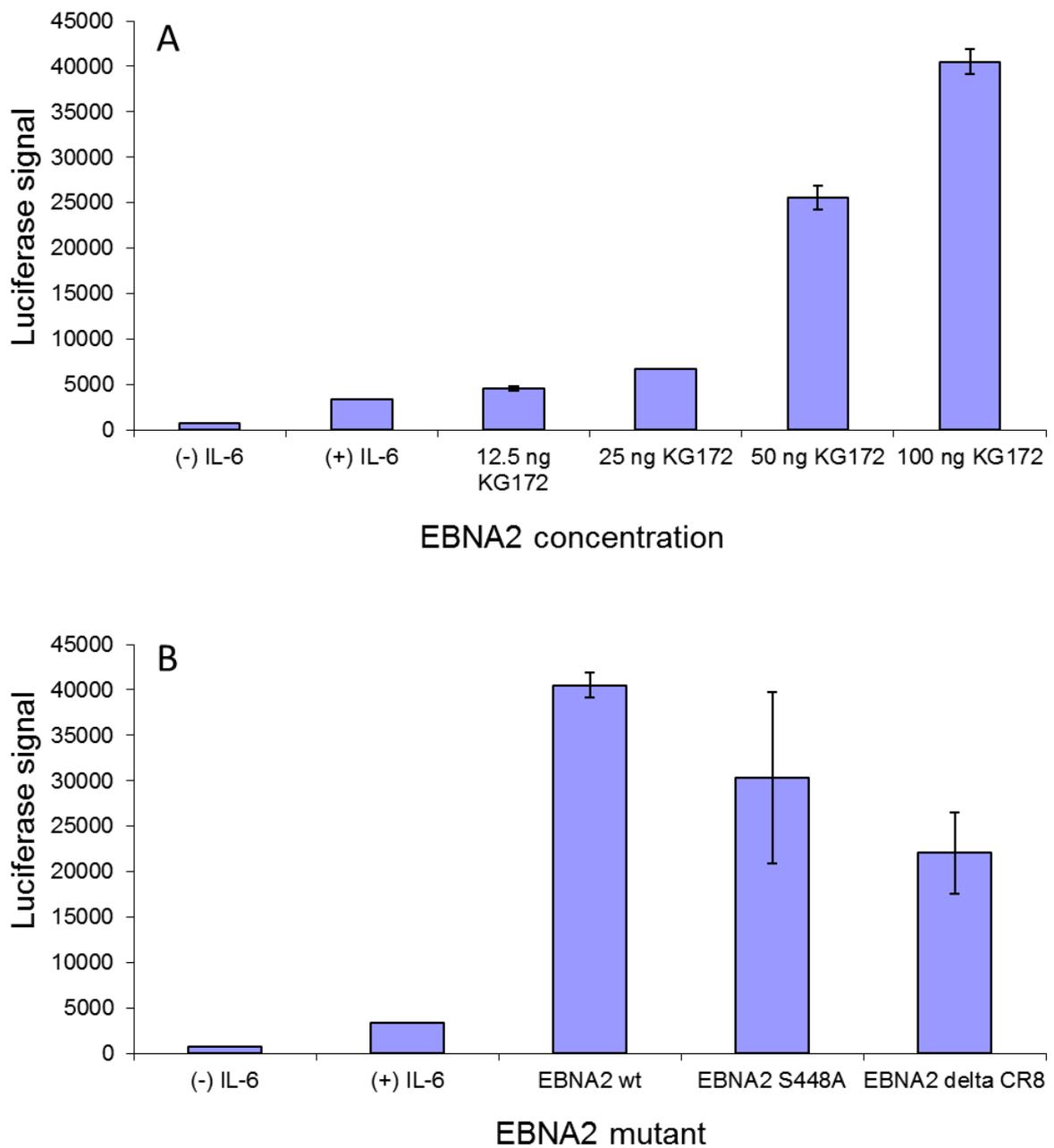


Figure 2.14: Luciferase assay of EBNA2 and EBNA2 binding mutants. A: Stably transfected HEK293-STAT3-Luc cells were co-transfected with increasing concentration of full-length EBNA2 expression plasmid (KG172). After 24 h, cells were treated with 100 ng/ml of LIF for 6 h. Luciferase signals were detected using a Wallac luminometer. B: Using the same method cells were transfected with a concentration of 100 ng wild-type EBNA2, S448A and Δ CR8.

Deletion of CR8 in EBNA2 reduced STAT3 activation by around 50% (p-value for KG172wt compared with KG172 Δ CR8 < 0.01). Substitution S448A in CR8 reduced the activation of STAT3 by around 25% (p-value for KG172wt compared with KG172 S448A < 0.36). These results confirmed the importance of CR8 for binding to STAT3 but the standard deviation values, especially for the most interesting *in vitro* mutant S448A, were rather high; however despite repeated attempts the data remained similar.

2.1.10.2. Studying the STAT3-EBNA2 interaction by transient transfection

In order to obtain better statistics and to study the STAT3-EBNA2 interaction in a B cell line environment which is closer to the natural host of EBV, a transient transfection approach was performed in DG75 cells, a human B lymphocyte cell line. The following experiments were performed during an EMBO short term fellowship visit to the laboratory of B. Kempkes, Helmholtz institute, Munich.

Initially, the expression of the different EBNA2 expression constructs in DG75 cells was confirmed by western blot (Fig. 2.15). All constructs were expressed but at different levels with EBNA2 mutants Y460A-W458A-D459A (triple mutant) and S448A having the highest expression level. It is not clear what causes the difference in running behavior (the constructs were all sequence verified). A possible explanation might be that the point mutants might affect certain migration-influencing posttranslational modifications or that the deletion of CR8 might change the charge of the protein leading to aberrant migration. However, because of time constraints the repetition of the experiment was not possible (but is planned).

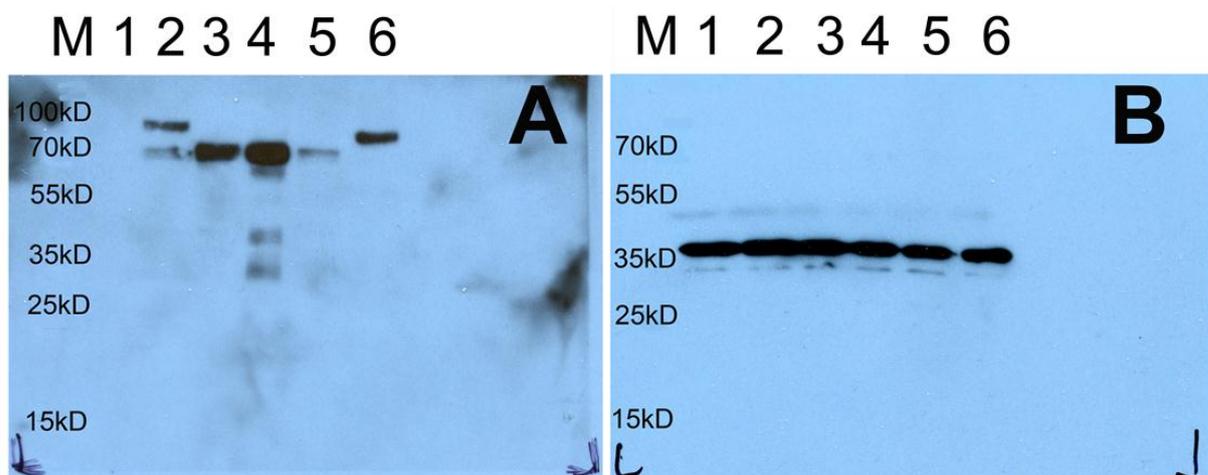


Figure 2.15: Expression of EBNA2 wild-type and mutants in DG75 cells. A: Detection of HA tagged EBNA2 wild-type and mutants using HA antibody. B: Detection of GAPDH in total cell lysate as a loading control. Legend: M: Ladder, 1: Empty vector, 2:EBNA2 wt, 3: EBNA2 S448A, 4: EBNA2 Y460A-W458A-D459A, 5: EBNA2 S448A-W458A-D459A, 6: EBNA2 Δ CR8.

To measure the transactivation activity of EBNA2 mutants, DG75 cells were transfected with a plasmid containing the luciferase gene under control of a STAT3 promoter element in a similar manner to that performed in the lentivirus experiment (Fig. 2.14). The luciferase signal reports the activation of STAT3 by EBNA2. A concentration series of EBNA2 wild-type plasmids was used to determine the sensitive range of the assay, and then a constant concentration of point mutant expression plasmids was used to co-transfect cells.

After performing the experiment several times with varying LIF and IL-6 concentrations we concluded that the particular DG75 cell line used in the lab was unexpectedly insensitive to LIF or IL-6 treatment; thus STAT3 signaling could not be activated in these cells. We therefore switched back to HEK293 cells that are highly LIF and IL-6 sensitive but kept the transient transfection approach. Cells were co-transfected with a STAT3-firefly luciferase construct and a control renilla luciferase construct employed for transfection efficiency normalisation. Cells were stimulated with 100 ng/ml of LIF 20 h post- transfection and incubated for 6 h. The results show that STAT3 activity was enhanced by EBNA2 in a concentration dependent manner (Fig. 2.16A) and that the EBNA2 point mutants reduced this effect (Fig. 2.16B). These results are in agreement with the results obtained with the stably transfected cells (Fig.2.14A and B).

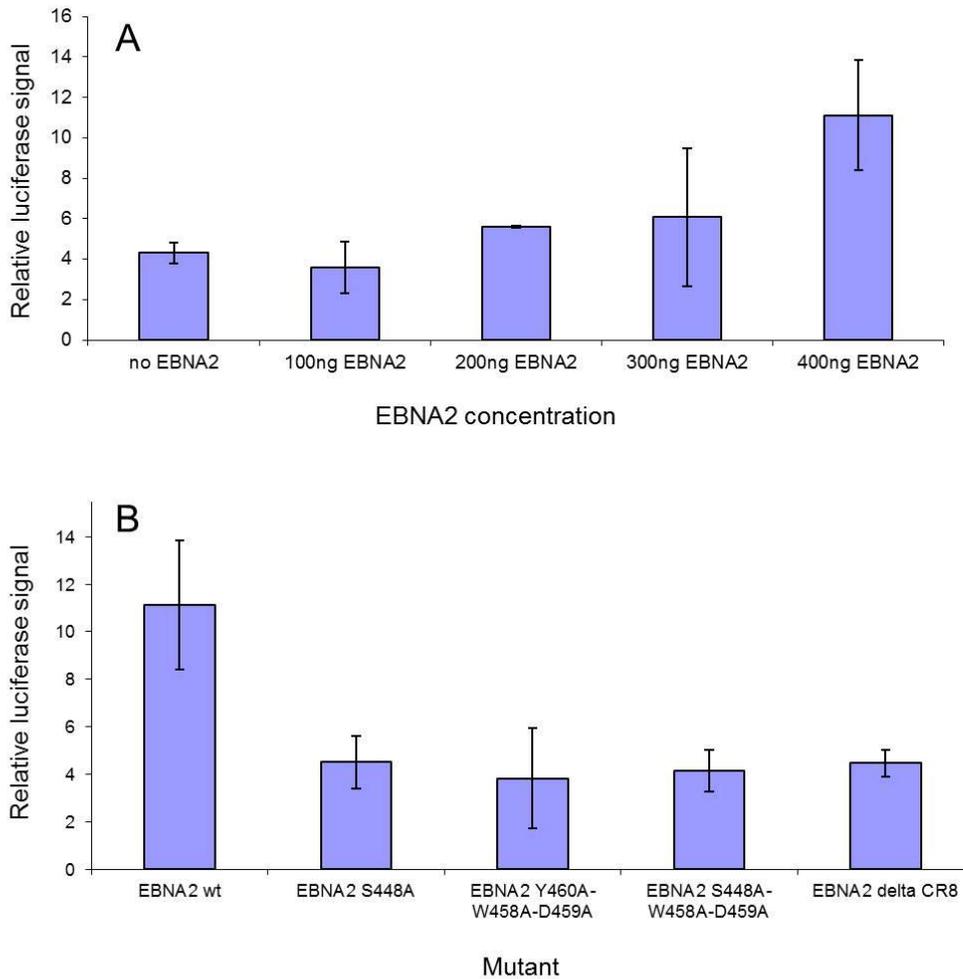


Figure 2.16: Transient transfection of luciferase reporter construct together with EBNA2 and EBNA2 binding mutants. 2.5×10^5 HEK293 cells were transfected (using lipofectamine). (A) various concentrations of EBNA2 (100, 200, 300 and 400 ng) and 1 μ g reporter plasmid and control plasmid mix (40:1). (B) constant concentration (400 ng) of EBNA2 wt and point mutant constructs and 1 μ g reporter plasmid and control plasmid mix (40:1). Experiments were performed in triplicates. Luciferase signal was measured using the Promega dual luciferase reporter assay kit according to manufacturer's instructions.

We then confirmed that the general transactivation activity of EBNA2 was not impaired by the STAT3 interaction mutations. For this purpose, we measured the CBF1-EBNA2 interaction via a well established CBF1 reporter assay (Hayward, 2004; Maier et al., 2005). The DG75 cells were transfected with a CBF1-firefly luciferase plasmid, a β -galactosidase plasmid for transfection normalisation and a constant concentration of EBNA2 expression plasmid. The transactivation levels of both wild-type and mutant EBNA2 constructs were measured 2 days post electroporation. The results show that the general transactivation ability of EBNA2 was not impaired by the point mutations that disrupted STAT3 binding (Fig.

2.17). Only the EBNA2 deletion mutant Δ CR8 was unable to activate CBF1 driven transcription. This was expected as this mutant lacked the whole conserved region 8 which is a large fragment of the transactivation domain of EBNA2. The EBNA2 mutant S448A seemed even to increase CBF1 mediated transcriptional activity, perhaps explained by the difference in expression level between EBNA2 wild-type and S448A (Fig. 2.15).

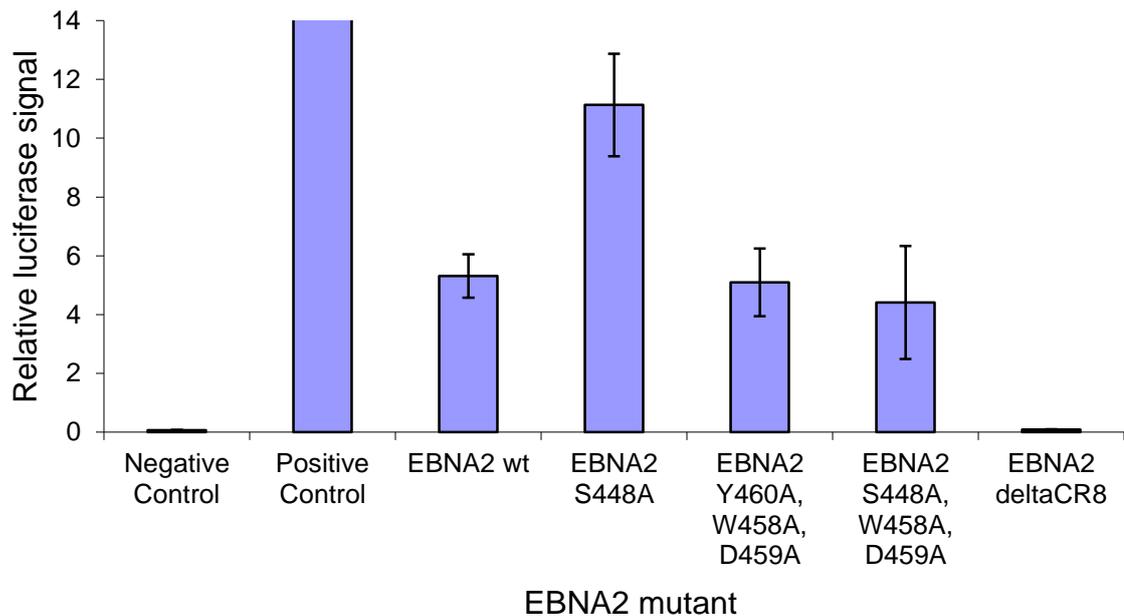


Figure 2.17: CBF1 mediated transcription activation by EBNA2 wild-type and STAT3 binding mutants. 5×10^6 DG75 cells were transfected (using electroporation) with $1 \mu\text{g}$ EBNA2 wild-type, point mutant construct or control vector, $5 \mu\text{g}$ CBF1 reporter plasmid and $3 \mu\text{g}$ β -galactosidase. Experiments were performed in triplicates. Negative control: Luciferase vector with mutated CBF1 binding site, Positive control: Luciferase vector with constantly active CMV promoter element.

2.1.11. Interaction of EBNA2 with other host cell proteins

As side projects the interaction of EBNA2 with two partner proteins other than STAT3 were investigated and analyzed in regard of binding affinities and binding mutants. The BS69-Mynd domain has been suggested to be involved in repression through N-CoR recruitment (Masselink and Bernards, 2000) and in cell growth control (Ansieau and Leutz, 2002). Med25 is a subunit of the mediator complex which forms a bridge between transcription factors and the general transcription machinery (Borggreffe and Yue, 2011).

2.1.11.1. Binding of EBNA2 to ZMYND

EBNA2 was identified as a BS69 Mynd domain binding partner by yeast two hybrid screen of an EBV-immortalised B cell line using the BS69 Mynd domain as a bait. The screen identified the C-terminal part of EBNA2 as a binding partner. Sequence alignment with the adenoviral oncoprotein E1A, another already known binding partner of the BS69 Mynd domain, suggested two *PXLXP* (amino acids 383-387 and 437-441) motifs in EBNA2 to be the binding motifs. Pull-down experiments revealed that mutation of both motifs abrogates binding, whereas mutation of single motifs had no effect suggesting a redundancy of EBNA2 *PXLXP* motifs (Ansieau and Leutz, 2002). Both *PXLXP* motifs identified in this study can be found in EBNA2 fragment D9. Wild-type and two *PXLXP* motif mutants (EBNA2 D9 L385A and EBNA2 D9 L439A) were subjected to SPR analysis. EBNA2 fragments were immobilised on a CM5 chip and Mynd domain protein was injected in a concentration range of 0.25 to 5 nM. Binding curves of D9 wt to Mynd could be fitted using the BIAevaluation software and the 1:1 Langmuir binding model. The resulting K_D is 450 nM (Chi value 0.25). The single *PXLXP* motif mutants did not show binding anymore (Fig. 2.18).

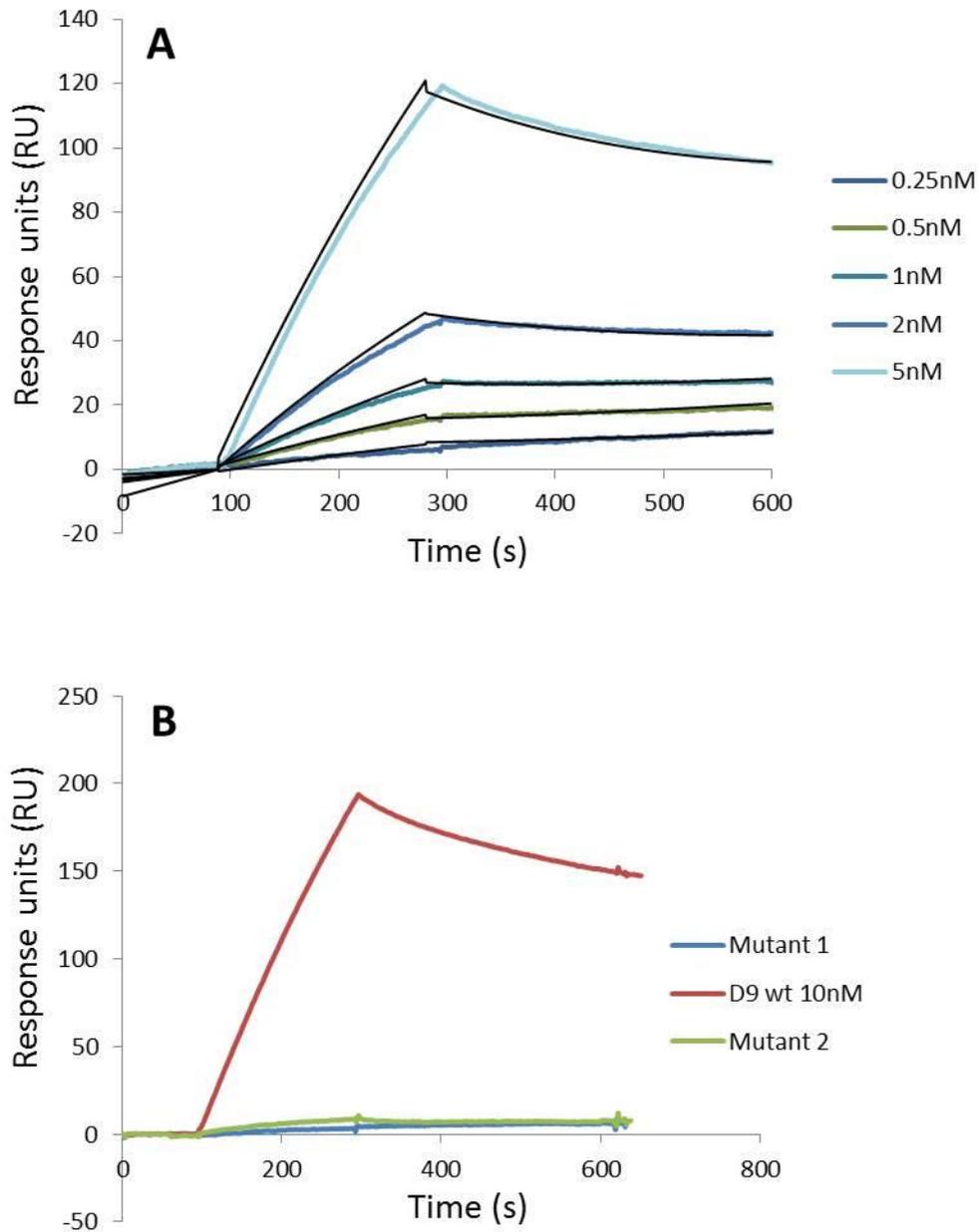


Figure 2.18: SPR sensograms of EBNA2 wild-type and *PXLXP* motif mutants binding to EBNA2 fragment D9. A: Sensogram of EBNA2 fragment D9 binding to Mynd. Curves are fitted with a 1:1 langmuirmodel. Concentration range of Mynd is 0.25 to 5 nM. B: Comparison of EBNA2 D9 wild-type, EBNA2 D9 L385A (mutant 1) and EBNA2 D9 L339A (mutant 2) binding to Mynd at a single concentration of 10 nM.

The binding of the C-terminal domain of EBNA2 to Mynd domains was therefore confirmed together with the role of the *PXLXP* binding motif. Single *PXLXP* motif mutations abrogated binding in contrast to the pull-down experiments (Ansieau and Leutz, 2002, 69).

2.1.11.2. Binding of EBNA2 to Med25

The Med25 activator interaction domain (ACID) has been shown to interact with the viral protein VP16 of herpes simplex virus I by NMR including a structural assignment (Vojnic et al., 2011). Due to structural and functional similarity between EBNA2 and VP16 and interaction of Med25 ACID and EBNA2 is presumably CR8. Med25 ACID protein was kindly provided by Alexis Verger (Université de Lille) and tested for EBNA2 binding by SPR. EBNA2 fragments F3, A11 and D9 were immobilised on a CM5 chip up to equal levels and Med25 was injected in the concentration range of 1-128 nM (Fig. 2.19). EBNA2 fragments A11 and D9 bound to Med25 ACID domain with the respective affinities of 2.63 nM for A11 (Chi 0.56) and 4.52 nM for D9 (Chi 0.64). Values were obtained by global curve fitting using the two-state binding model. EBNA2 fragment F3 did not bind to Med25. Due to the overlap of A11 and D9 and the very similar binding affinities CR7 is suggested to be the binding region in EBNA2, however this project was not continued further so the significance or validity of these data awaits future confirmation.

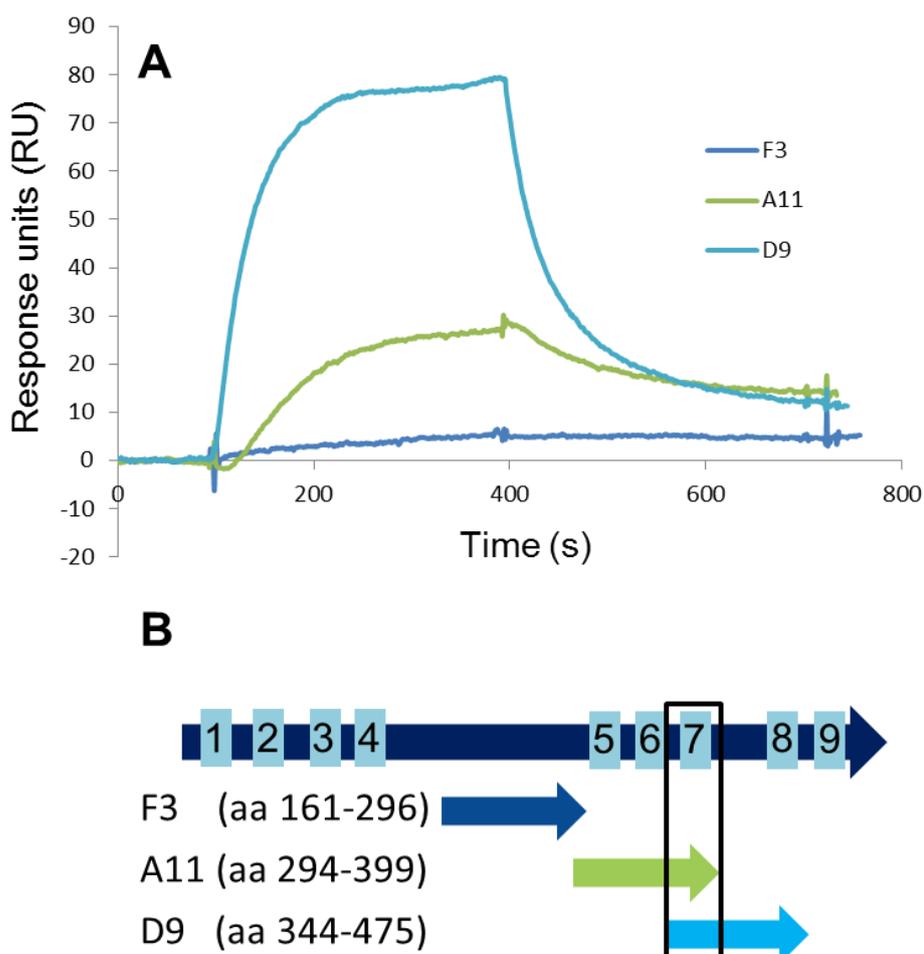


Figure 2.19: SPR sensogram of EBNA2 fragments binding to Med25 ACID and suggested binding site. A: Comparison of EBNA2 fragments F3, A11 and D9 binding to Med25 ACID at a single concentration of 128 nM. B: Sequence alignment of EBNA2 full length with the EBNA2 fragments suggest a possible binding region between amino acids 344 and 399 (black box) containing CR7. Light blue boxes indicate CRs.

2.2. Interaction of STAT3 with the cellular co-repressor SMRT/NcoRII

Under physiological conditions some of the cellular STAT3 is complexed by the nuclear co-repressor SMRT (Ikeda et al., 2009). In transfection assays, it was demonstrated that this complex could be disrupted by EBNA2 (Muromoto et al., 2009a). In contrast to EBNA2 which activates STAT3 mediated transcription, binding of SMRT to STAT3 decreases its transcriptional activity. EBNA2 somehow releases SMRT from this complex (Ikeda et al., 2009) with the probable, but not yet characterised, enhancement of transcription of both host and viral STAT3 regulated genes. STAT3 activity has several immunosuppressive effects in the host organism (Kortylewski et al., 2005b, 2009; Berg et al., 1996; Gerosa et al., 2008) and may therefore have importance in the viral survival strategy of Epstein-Barr virus (Koganti et al., 2013). Furthermore, EBNA2 interacts with STAT3 with significantly higher affinity and is therefore a further example of how viruses hijack the host cell immune system (Davey et al., 2011; Pushker et al., 2013). As a complement to the data on EBNA2-STAT3 interactions, we performed similar studies on the interaction of SMRT and STAT3 with the aim of understanding both sides of this potentially antagonistic relationship.

2.2.1. Identification of the binding region using a cell based approach

The region of the 2525aa SMRT to which STAT3 binds is unknown so, in order to identify the binding site, an *in vivo* approach similar to that applied to EBNA2 was used. The SMRT gene was randomly divided into six fragments of equal size and expressed in HEK293-STAT3-Luc cells. This cell line was generated using lentiviral particles carrying the firefly luciferase gene under control of a specific STAT3 promoter. Cells were kept under puromycin selection pressure and isolated using the limiting dilution method to obtain stable, monoclonal cells. Interleukin-6, a well-known activator of STAT3 signalling was used to challenge cells after transfection with the SMRT fragment plasmids. The cell based assay shows a significant reduction in luciferase signal for fragments 3, 4 and 5 corresponding to a region spanning amino acids 820-1920 (Fig. 2.20). The next step was to further narrow down the binding region using recombinant protein and biophysical methods.

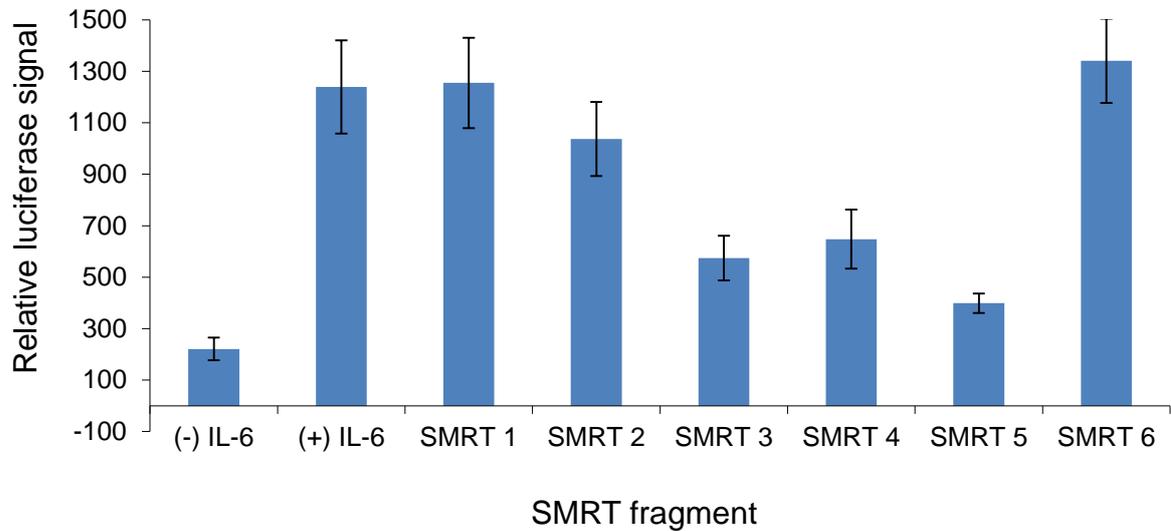


Figure 2.20: Luciferase signal induced by STAT3 stimulation with IL-6. 10^5 cells/well of STAT3-Luc HEK293 cells were plated into a 96-well plate in quintuplicates. After 24 h cells were transfected with pcDNA3-myc-SMRT1-6 (1-420, 421-820, 821-1260, 1261-1680, 1681-1920 and 1921-2525) and pCMV-SPORT- β -Gal. Cells were treated with 50 ng/ml of IL-6. After 12 h luciferase and β -galactosidase signals were detected.

2.2.2. Expression and purification of SMRT fragment

An ESPRIT expression library of random SMRT fragments covering amino acids 1089 to 1993 of the SMRT protein had been already generated from which two purifiable constructs were isolated: 39L23 (aa 1256-1455) and 27M12 (aa 1785-1994) (D. Desravines and P. Mas, described in thesis of D. Desravines “Structural, functional and inhibition studies of human histone deacetylase 7”, 2010).

The *E.coli* strain BL21 AI (RIL) was chosen for expression of these SMRT fragments and transformed by electroporation. Bacteria were grown in 1 l LB medium and protein expression was induced at an $OD_{600\text{ nm}}$ of 0.6 at 25°C. Lysis was performed by microfluidisation. The soluble fraction was then incubated in a water bath with 95°C temperature for 10 min since a particular feature of the SMRT fragments studied was their extreme heat resistance, indicative of an absence of structure characteristic of IDPs. Lysate heating provided an efficient purification step since cellular proteins precipitated at these temperatures. The boiled lysate was cleared by centrifugation and remaining proteins further purified by affinity purification on NiNTA beads. After the first affinity purification, the fractions

with the highest purity were pooled and the 6xHis tag was cleaved with TEV protease. Proteins were diluted in wash buffer and incubated with NiNTA beads again to remove the cleaved 6xHis tag and the 6xHis tagged TEV protease. The efficiency of TEV cleavage was checked on SDS-PAGE (Fig. 2.21) and the proteins concentrated to the desired concentration. Using this protocol, a total purity of around 90% could be achieved with the total amount of protein per litre of bacterial culture approaching 10 mg.

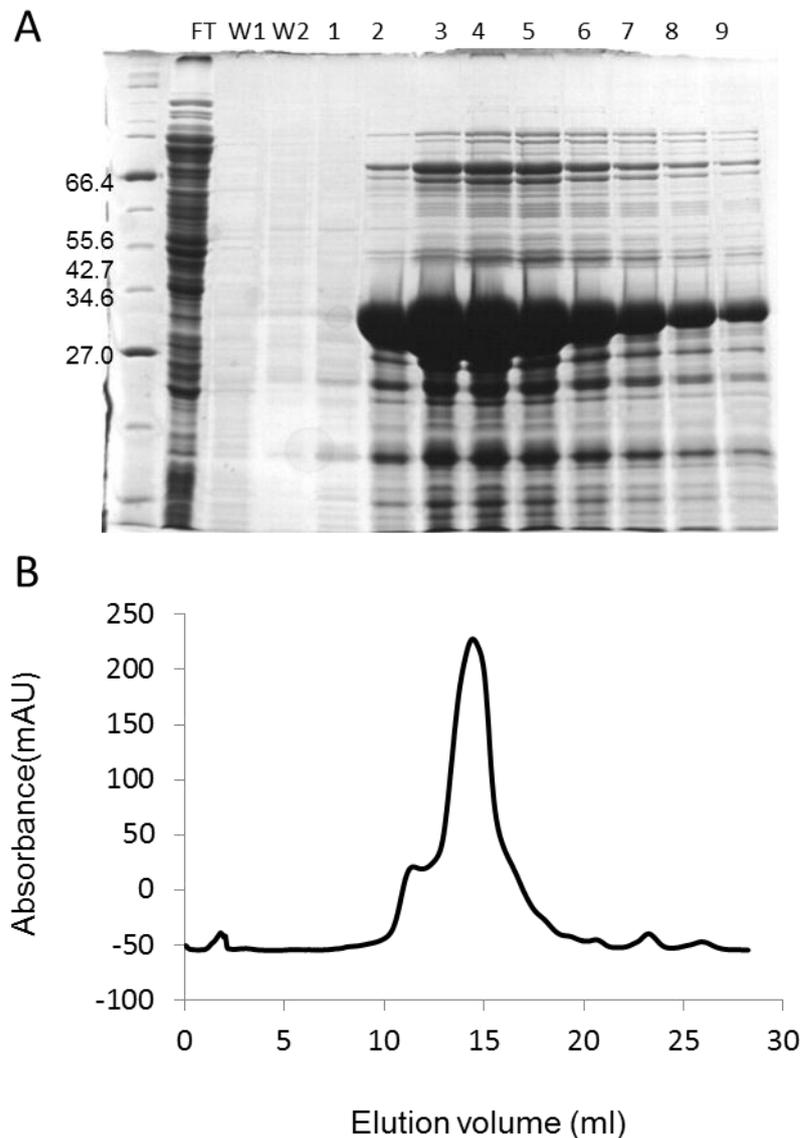


Figure 2.21: Expression of SMRT fragment 39L23. A: SMRT fragment 39L23 after 1st affinity purification (FT: Flow through), W1 and W2: Wash 1 and 2, 1-9: Elution fractions 1-9, B: Size exclusion chromatography diagram for SMRT fragment 39L23.

2.2.3. Identification of a SMRT fragment binding to STAT3

Two SMRT fragments, 39L23 (aa 1256-1455) and 27M12 (aa 1785-1994) were chosen and analysed by $^1\text{H}^{15}\text{N}$ HSQC NMR for STAT3 binding. Although the chosen fragments don't cover the whole region identified in the cell based assay they still cover large parts of it (Fig. 1). For this purpose the two SMRT fragments were ^{15}N -labelled and spectra of the SMRT fragment in free form and in complex with various concentrations of STAT3 (3:1, 1:1, 1:3 and 1:5) were recorded in collaboration with the Blackledge lab (IBS, Grenoble). One of these fragments, 39L23 shows binding to STAT3 indicated by peak intensity changes and shifts, whereas the other fragment, 27M12 does not bind to STAT3 (Fig. 2.22).

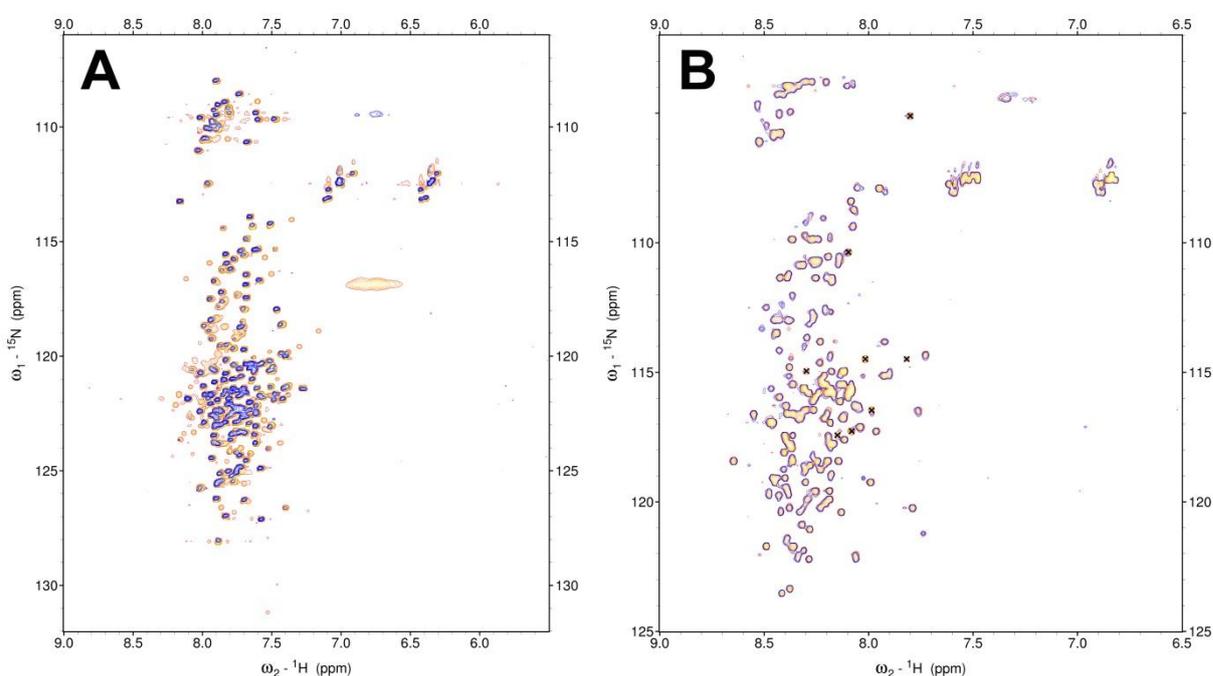


Figure 2.22: NMR spectra of 39L23 (A) and 27M12 (B) upon titration of STAT3. The fragments were labelled with ^{15}N and measured at a concentration of above 200 μM in a volume of 250 μl . The ratio between SMRT fragment and STAT3 is 1:1. Spectra recorded by R.Schneider (Blackledge lab, IBS Grenoble).

2.2.4. Identification of the amino acids in fragment 39L23 involved in STAT3 binding

The 39L23 SMRT fragment was further investigated by NMR assignment (collaboration R. Schneider, Blackledge lab, IBS Grenoble). For this purpose the protein was expressed in minimal medium containing ^{15}N -ammonium chloride and ^{13}C -glucose. The assignment mapped each peak to the corresponding amino acid and allowed peak intensity change for each residue to be quantified upon addition of STAT3. The assignment showed that three regions within 39L23 are mainly affected by the binding: 79-89, 106-127 and 162-173

corresponding to residues 1336-1344, 1362-1383 and 1418-1429 in the full-length sequence (Fig. 2.23).

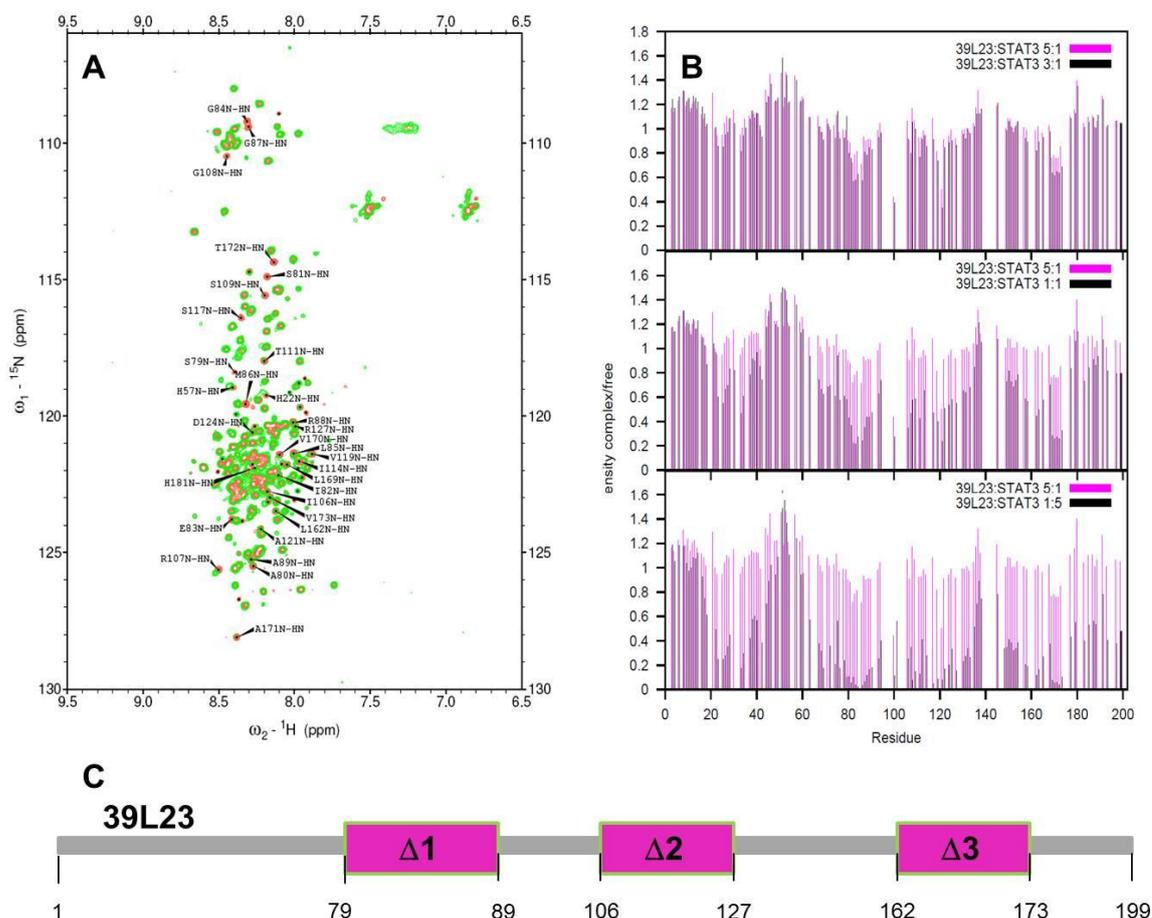


Figure 2.23: Assignment of SMRT fragment 39L23. A: Peak intensity changes upon STAT3 titration from 1- to 5-fold access. B: Amino acids involved in STAT3 binding. Each bar of the histogram corresponds to one residue of fragment 39L23. C: Location of the STAT3 binding region within 39L23 from the NMR data. Assignment performed by R. Schneider (Blackledge lab, IBS Grenoble).

2.2.5. Binding kinetics and mutational analysis of the STAT3-39L23 binding

Deletion mutants were generated for the regions showing the highest peak intensity changes (Fig 2.23c): $\Delta 79-89$, $\Delta 106-127$ and $\Delta 162-173$ as well as one longer deletion mutant $\Delta 79-127$. The mutants and wild-type 39L23 were tested for binding to STAT3 by SPR (Fig. 2.24). There were no significant differences between the single deletion mutants and the wild-type detectable. Only the mutant lacking both region 1 and 2 showed a significant change in binding mode and binding affinity.

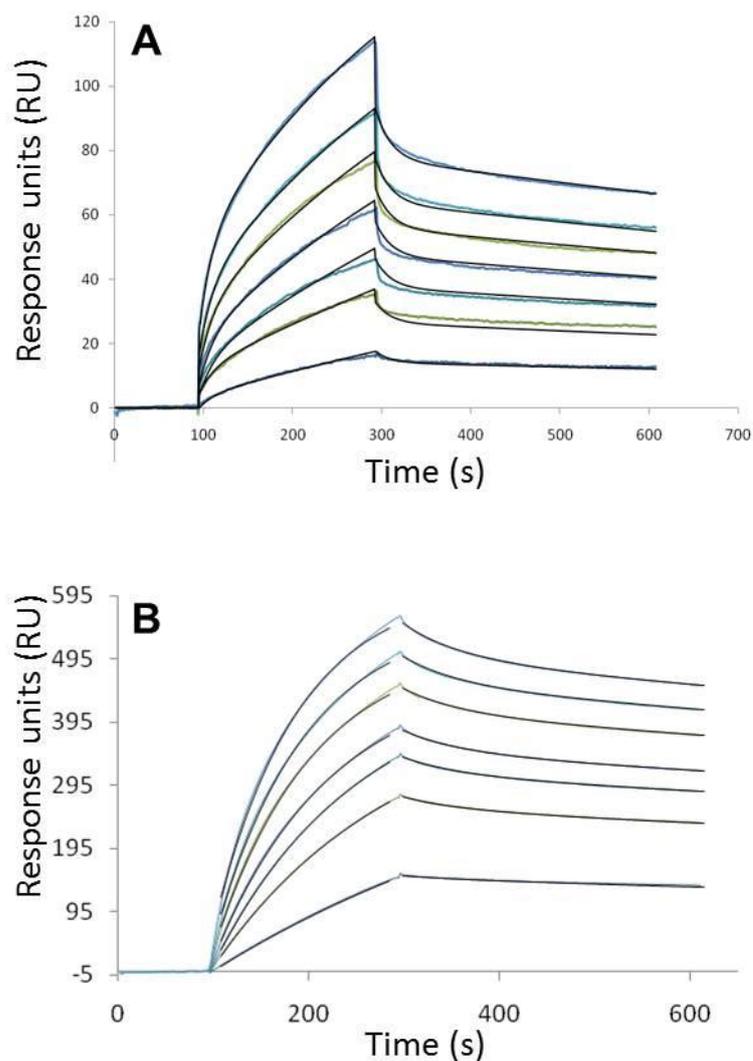


Figure 2.24: Sensograms and fitted curves of immobilised 39L23 wild-type (A) and $\Delta 1+2$ (B) binding to STAT3 injected from 125 nM to 1 μ M in low salt buffer conditions (10mM MgCl_2). Data recorded together with M. Roth (Trainee, Hart team, EMBL Grenoble).

The wild-type 39L23 fragment bound to STAT3 resulting in sensograms best fitted to a two-step model with fast and slow association/dissociation which has been used to describe IDP interactions by SPR previously (Sevcik et al., 2007a). Using this model, the data could be fitted with high confidence (χ^2 2.947) to give a K_D of 121 nM (Fig. 2.24A). The higher response values of mutant $\Delta 79-127$ indicated a significantly higher level of binding, but could no longer be fitted by the two-state model. Instead, the data was better described by the most simple 1:1 one-state model predicting an increased affinity of 19 nM (Fig. 2.24B).

Discussion and Outlook

Résumé

Interactions entre EBNA2 et STAT3

Le transactivateur viral EBNA2 a été décrit à maintes reprises dans la littérature à propos de sa fonction, cependant sa structure est faiblement caractérisée. EBNA2 interagit avec beaucoup de virus ainsi que des protéines cellulaires de l'hôte. Beaucoup d'interacteurs ont été identifiés, néanmoins sa caractérisation biophysique et ses propriétés tant biochimiques que structurales sont piètrement mise en évidence. Surtout, les interactions d'EBNA2 liées à ses particularités intrinsèquement désordonnées restent un mystère. L'interaction avec STAT3 fut étudiée en détail de manière à comprendre l'un des nombreux aspects de la stratégie de survie de l'EBV, à savoir l'efficace échappatoire au système immunitaire de la cellule hôte et du contournement de la régulation cellulaire. Les résultats qui incluent la création d'une banque de constructions aléatoire (ESPRIT) d'EBNA2, l'expression et la purification de fragments d'EBNA2 ainsi que le dimère de STAT3; montrent la cinétique d'interaction ainsi que les sites impliqués à travers celle-ci. Il sera également discuté d'autres expériences à réaliser qui impliquent principalement de la culture cellulaire.

Interactions entre EBNA2 et d'autres protéines hôte

EBNA2 a été extensivement étudiée et il est clair que ce dernier est impliqué de façon fonctionnelle dans beaucoup de processus cellulaires. À travers une étude de double hybride à haut débit, il a été découvert que cette protéine interagit avec 16 autres partenaires cellulaires. L'interaction d'EBNA2 avec BS69 Mynd et Med25 fut étudiée préalablement. Ici, ces interactions furent confirmées par SPR et des sites de fixation suggérés.

Interactions entre SMRT et STAT3

Trois régions de SMRT (1336-1344, 1362-1383 and 1418-1429) eurent un effet sur l'interaction de STAT3. D'une façon surprenante, aucune des délétions générées par mutation n'ont affecté individuellement de façon significative l'interaction. Cependant, la délétion de deux régions de fixation (aa 80-127) a drastiquement modifié le mode de fixation. Ce dernier fut modifié d'un simple 1:1 mécanisme de Langmuir et l'affinité de fixation fut augmentée d'environ 10 fois à 19 nM. Les possibles explications de cette observation inhabituelle pourraient être que la région de fixation 3 (non abolie) est initialement auto-inhibée par les régions 1 et 2, peut être masquant un motif d'interaction à STAT3. La délétion des régions 1 et 2 pourrait mener à des interactions plus fortes à STAT3. Alternativement, SMRT est constitué par une grande organisation de type "échafaudage" qui serait engagé avec

STAT3 à travers des régions additionnelles en dehors du fragment 39L23. Ces résultats furent mis en évidence par des expériences de “cell based assay” ou des fragments de SMRT (AA821 a 1920) ont diminué la transcription régulée par STAT3. L’hypothèse est que SMRT supprime l’activité de transcription de STAT3 par un recrutement d’autres protéines telles que HDACs qui inhibe activement la transcription.

Protéines intrinsèquement désordonnées (IDPs)

EBNA2 et SMRT sont prédites comme étant des protéines intrinsèquement désordonnées. Des études structurales de cette protéine sont peu abondante en comparaison des protéines ayant un repliement, même si la majorité des oncoprotéines et la moitié des autres protéines cellulaires sont prédites comme étant des IDPs. Ces dernières sont localement structurées après leur fixation à des protéines partenaires, parfois a travers des motifs linéaires ou de courtes hélices alpha, mais aussi a travers des événements plus prononcés de repliement après interaction. A ce moment, des liaisons électrostatiques faibles sont suivies par un changement conformationnel sont le résultat d’une seconde interaction plus spécifique ainsi qu’une plus grande affinité. Ce mécanisme a été décrit comme “fly casting” et complète le dogme longtemps établi du système clé serrure qui correspond a la relation structure fonction. Des défis spécifiques existent avec les IDPs, surtout la sensibilité aux protéases et la difficulté de concevoir des constructions pour de l’expression. Ces deux difficultés ont été résolues en utilisant ESPRIT. Néanmoins après isolement, ces protéines se sont comportées de façon remarquable quant a leur facilité de manipulation due a leur absence de structure à dénaturer. Plusieurs idées ont été avancées à propos de l’avantage biologique de ces IDPs et sont discute ici. En raison de leur rôle clé dans beaucoup de procédés cellulaires et de maladies mais aussi du fait de leur abondance dans le protéome, les IDPs sont des cibles très intéressante en Recherche.

La triade d’interaction: EBNA2 en compétition avec SMRT pour l’interaction de STAT3

Dans des conditions physiologiques, le corépresseur SMRT se fixe a beaucoup d’adapteurs et de facteurs de transcriptions (par exemple, STAT3) et concilie la suppression de la transcription a travers le recrutement d’autres protéines de répression tels que les HDACs qui causent une deacetylation des gènes cibles afin d’inhiber leur transcription. Lors de l’infection de l’EBV d’une cellule, EBNA2 relargue SMRT du complexe. Ceci permettrait d’activer la transcription régulée de STAT3 qui se produit sensiblement par recrutement d’enzymes d’activation telle que la machine de transcription basale. STAT3 active est pro prolifératif, anti apoptotique et montre plusieurs activités suppressive sur l’immunité. De plus, des études sur HSV montrent l’hypothèse que STAT3 est requis pour maintenir la latence. Dans une autre étude, il est décrit qu’EBNA2 régule l’activation de STAT3 qui serait responsable de la prolifération et de l’oncogenèse. Aussi, une aberrante activation de STAT3

est retrouvée dans d'autres virus herpétiques. Ceci renforce la preuve que STAT3 est un facteur déterminant de la prolifération cellulaire et la survie des cellules B infectée par l'EBV.

3.1. The interaction between the viral protein EBNA2 and STAT3

The viral transactivator EBNA2 interacts with many viral and host cell proteins (Calderwood et al., 2007). Although many interactors have been identified, there exists little characterisation beyond this with respect to their biophysical, biochemical and structural properties. Notably the characterisation of EBNA2 interactions from the perspective of its disorder properties remains almost unstudied. One explanation for this lack of study might be that EBNA2, as an IDP, cannot be expressed full length in high yielding recombinant systems (at least we can find no reports of this in the literature). The predicted absence of three-dimensional domains means that the generation of expression constructs via the classical domain-focused design approach is not really possible. To address this, we present a new approach for obtaining fragments of a size and yield suitable for biophysical studies. The generation of a random expression library (ESPRIT) proved to be an effective and time-saving alternative to the classical PCR approach yielding multiple constructs that expressed milligrams of purifiable proteins in just a few weeks. The library efficiency was good with most of the 96 positive clones giving some level of detectable protein by SDS-PAGE. From these, three were chosen according to their location and expression level. The obtained EBNA2 fragments named F3 (aa 161-296), A11 (aa 294-399) and D9 (aa 344-475) were well-behaving and yielded up to 10 mg per litre of bacterial culture. Biochemical and biophysical analyses of these fragments showed that they were intrinsically disordered as evidenced by unusual thermal stability, sensitivity to protease digest, retarded running behavior in SDS-PAGE, accelerated elution in size exclusion chromatography and narrow peak distribution in HSQC NMR spectra.

The EBNA2 constructs show promise for studying interactions as evidenced by preliminary results with Mynd and Med25, and in more detail with STAT3. The interaction with STAT3 was studied in detail in order to understand an important aspect of the EBV survival strategy namely the effective escape from the host cell immune system and circumvention of cell regulation mechanisms (e.g. pro-proliferative and anti-apoptotic effects).

The interaction between EBNA2 and STAT3 was initially shown by pull-down experiments and luciferase assays in 2009 (Muromoto et al., 2009a); here we have validated this further using a number of biophysical and biochemical assays. The interacting EBNA2 fragment D9 was identified as the major STAT3 interacting region by streptavidin pull-down experiments (Fig. 2.6B) and SPR (Fig. 2.6A). The EBNA2 fragment F3 also interacted with STAT3 but with much lower affinity (Fig. 2.7). Therefore it was decided that the major binding region must lie within EBNA2 fragment D9, perhaps complemented by a weaker upstream motif in F3.

Further characterisation by NMR (Fig. 2.8) revealed the interaction was mediated by a limited number of interacting residues only i.e. seemed to be specific. The NMR spectrum in combination with the knowledge of the CRs within EBNA2 (Fig. 2.9) indicated that the binding region covered CR8, one out of 9 CRs within EBNA2, and which is located close to the C terminus. CRs in proteins are often indicative of functional importance. In folded domains, the sequence may have structural importance, but in IDPs it is perhaps indicative of linear interaction motifs. CR8 is located between aa 437-449 and therefore lies right in the middle of the acidic transactivation domain (TAD) that is located between amino acids 426 and 483 (Tong et al., 1995).

The critical amino acids for the interaction with STAT3 were determined via site directed mutagenesis and SPR, further validating that binding was specific. These mutations were also tested in cellular assays to see if the same effects could be measured under more physiological conditions. Initial results indicate that this is the case and that inhibition of STAT3 binding does not affect the overall transactivation function of this region. Further work is underway in collaboration with the Kempkes lab in Munich and will involve the following experiments:

1. The STAT3-luciferase assay was performed repeatedly with transient and stably transfected cells and resulted each time in a decrease in STAT3 mediated transcription activation for the mutants compared with wild-type EBNA2. However, the mutants tested did not completely inhibit, so it remains necessary to repeat the work to get better statistical parameters.
2. Also, our collaborators will repeat the CBF1 reporter assay, or perform a Gal4 transactivation assay, which monitors the general transactivation ability of EBNA2. In this assay EBNA2 will be fused to the Gal4 DNA binding domain and the transactivation levels of both wild-type and mutant EBNA2 constructs will be measured by luciferase assay. This is significant because one of the mutated amino acids within our triple mutant constructs (Trp 454) was identified previously as crucial for transactivation and subsequent transformation of B cells (Cohen, 1992). Surprisingly, neither of these mutant EBNA2 forms affected activation of CBF1-mediated transcription suggesting that the normal transactivation function was unaffected (Fig. 2.17).
3. The differences in EBNA2 function between wild-type and point mutants could be studied with a recent approach based (Kempkes et al., 1995; Lucchesi et al., 2008) on EBV infected cells that lack the EBNA2 gene locus and which are engineered to express EBNA2 fused to the hormone-binding domain of the oestrogen receptor. These cells are maintained in growth medium supplemented with oestrogen whereupon EBNA2 is imported in the nucleus. Upon removal of oestrogen, EBNA2

becomes localised in the cytoplasm and cells undergo apoptosis due to the abolished EBNA2 activity. By transient expression of unfused EBNA2 the cells can be rescued. Thus the system can be used to study the function of EBNA2 mutants. Furthermore, in order to identify host genes that are affected by the EBNA2-STAT3 interaction, RT-PCR and western blot analyses could be used as a read-out in this system. Based upon our preliminary data, we would expect that EBNA2 point mutants would affect the expression of a number of STAT3-dependent host genes that could then be compared to those identified via chromatin precipitation studies (Snyder et al., 2008; Vallania et al., 2009).

4. Beyond this cellular assay format, a recombinant EBV containing EBNA2 point mutations that prevent the STAT3 interaction could be generated and tested in B cells. Infected cells could then be analysed regarding their infection rate, transformation efficiency, viability of the cells and spontaneous lytic reactivation. Our hypothesis is that viruses with wild-type EBNA2 will infect and transform cells more efficiently than mutants that precisely disrupt STAT3 binding.

Collectively these experiments seek to show how EBV uses EBNA2 to subvert STAT3 signalling, and by comparison with interaction mutants, will reveal the effects of this host-virus interaction on cellular and viral functions.

The biophysical and biochemical results presented would ideally be complemented with a crystal structure of STAT3 in complex with an EBNA2 binding peptide. Different short EBNA2 binding constructs fused to a GST tag were analyzed for STAT3 binding. One 32 aa peptide was identified as the minimal binding region and synthesised chemically. This was added to the STAT3-DNA complex and small crystals were obtained in around 70 conditions. Six data sets were collected at the SOLEIL synchrotron (Proxima I beamline), but diffracted only to 6.5 Å resolution and it was not clear if the peptide co-crystallised. An intensive effort is now required to obtain bigger and better diffracting crystals.

3.2. The interaction between EBNA2 and other host proteins

EBNA2 has been well studied (~900 papers in PubMed) and it is clear that it is functionally involved in many cellular processes. In one high-throughput yeast two-hybrid study it was found to bind sixteen different cellular partners (Calderwood et al., 2007). The interaction of EBNA2 with BS69 Mynd and Med25 was previously studied (Ansieau and Leutz, 2002; Milbradt et al., 2011, 25; Vojnic et al., 2011, 25), but the availability of highly purified EBNA2 fragments permitted further quantitative analyses to be performed.

3.2.1. Interaction between EBNA2 and the BS69 MYND domain

A yeast two-hybrid screen of an EBV-immortalised B cell line using the BS69 Mynd domain as bait identified the C-terminal part of EBNA2 as a binding partner (Ansieau and Leutz, 2002, 69). Sequence alignment and comparison with the adenoviral oncoprotein E1A, another known binding partner of the BS69 Mynd domain, suggested two binding motifs in EBNA2: *PXLXP* (aa 383-387 and 437-441). Pull-down experiments revealed that mutation of both motifs in EBNA2 abrogated binding, whereas mutation of single motifs had no effect suggesting a redundancy of EBNA2 *PXLXP* motifs (Ansieau and Leutz, 2002). EBNA2 fragment D9 also contains both motifs. Single and double mutations were introduced (L385A and L439A) and subjected to binding analysis via SPR. The published hypothesis for E1A, suggesting a redundancy of the motif, could be not confirmed for EBNA2 where a single mutation of just one of the *PXLXP* motifs abolished binding completely (Fig. 2.18). However, the previously published data were based on pull-down experiments in which the BS69 (411–561) Mynd domain was expressed as a GST-fusion protein together with single and double mutants of EBNA2 (L385A and L439A). Pull-down experiments can yield false positive signals due to unspecific binding to the beads and from the data in the publication it is not clear whether this was the case. Another reason might be that the protein lysates used for pull-down experiments contain many different proteins and binding of the single mutants of EBNA2 to BS69 Mynd could have been facilitated by additional unknown binding partners. In contrast, SPR experiments are more specific since they use highly purified proteins, but it is also a highly artificial system that only permits evaluation of the interaction between ligand and analyte.

3.2.2. Interaction between EBNA2 and Med25

The structure of the Med25 acidic domain was solved by NMR (Bontems et al., 2011). In 2011 the solution structures of the Med25 VP16 binding domain (Med25 VBD) and the VP16 transactivation domain (VP16 TAD) (Milbradt et al., 2011) were published and the structural and functional interaction between these two proteins was detailed (Vojnic et al., 2011). Structural sequence alignment of VP16 and EBNA2 suggested that the binding site in EBNA2 is located between aa 435-452 covering CR8. Binding analyses of EBNA2 fragments F3, A11 and D9 with Med25 revealed that fragment A11 and D9 interact with Med25 with similar affinities (2.63 nM for A11 and 4.52 nM for D9). This suggests that the binding region is located within the overlap of both fragments which contains just one conserved region, CR7. Thus CR7 may be the binding site in EBNA2 for Med25 or, alternatively, EBNA2 may contain at least two binding sites for Med25. One of them might be within CR8 as suggested by sequence alignment whilst the other one is located within the N-terminal part of fragment A11 that contains CR5 and 6. Deletion mutants for CR5, 6, 7 and 8, assayed for interaction with Med25, would confirm which hypothesis is the right one.

3.3. The interaction between the cellular protein SMRT and STAT3

SMRT is significantly bigger than EBNA2, so in order to narrow down the STAT3 binding region within the 2525 aa long protein 6 SMRT fragments were expressed in stable STAT3-Luc reporter cells and the effects were assessed by luciferase signal (Fig. 2.20). The SMRT fragments showing a decrease in STAT3 mediated transcription activation were found within the region aa 820-1920. Two recombinant SMRT fragments covering most of this region, 39L23 (aa 1256-1455) and 27M12 (aa 1785-1994) were selected from a previous ESPRIT screen and analysed by HSQC NMR for evidence of STAT3 binding (Fig. 2.22). Fragment 39L23 bound STAT3 and the SMRT amino acids involved were identified following assignment of the peaks (Fig. 2.23). Three individual regions exhibited changes upon STAT3 binding; these were located between aa 80-88, 106-127 and 162-173, corresponding to SMRT residues (1336-1344, 1362-1383 and 1418-1429). Three deletion mutants were then constructed and tested separately for STAT3 binding using SPR. Surprisingly, none of the deletion mutations individually affected the binding significantly. In contrast, the deletion of two binding regions (aa 80-127) significantly changed the binding behavior (Fig. 2.24). The two-state binding mechanism of STAT3-SMRT (1256-1455) that has been reported previously to describe IDP interactions (Sevcik et al., 2007b), with both fast and slow association and dissociation steps, changed to a simple 1:1 Langmuir binding mechanism and the binding affinity increased nearly 10-fold to 19 nM (Fig. 2.24). Possible explanations for this unusual observation may be that a binding site in region 3 (not deleted) is initially autoinhibited by regions 1 and 2, perhaps masking a STAT3 interaction motif. Deletion of regions 1 and 2 may lead to a tighter interaction with STAT3. Alternatively, SMRT as a large scaffold protein may bind STAT3 via additional regions outside fragment 39L23, in possible agreement with the results from the cell based assay where SMRT fragments corresponding to aa 821 to 1920 decreased STAT3 mediated transcription (Fig. 2. 20). Compared to 39L23 which covers aa 1256-1455 this is a significantly bigger region. Further work will be necessary to test these ideas.

The STAT3 residues involved in SMRT binding are not known, however other observations may hint at two possible binding sites on STAT3. SMRT was reported to bind the coiled-coil domain of STAT5 (Nakajima et al., 2001) while other groups reported a binding of a YXXQ motif to the SH2 domain of STAT3 (Shao et al., 2004). The second hypothesis was tested by adding SMRT (1256-1455) at saturating concentrations to a STAT3-DNA gel-shift assay since this motif is present in 39L23, however results were inconclusive (data not shown). No release of STAT3 from DNA was observed suggesting SMRT did not compete with DNA

binding under the conditions tested. An alternative and perhaps stronger hypothesis is that SMRT binds to the coiled-coil domain of STAT3 (as shown for STAT5) and suppresses STAT3 transcriptional activity through recruitment of other proteins e.g. HDACs that actively inhibit transcription.

3.4. Intrinsically disordered proteins

Both the cellular co-repressor SMRT and the viral trans-activator EBNA2 are predicted to be highly intrinsically disordered. Structural studies of this protein type are relatively rare when compared to folded proteins, even though the majority of oncoproteins and half of all cellular proteins are predicted IDPs. Also, certain viruses are characterised by a high amount of disorder within their proteomes. Herpes viruses were found to possess 17,9% of disorder (Pushker et al., 2013). The determination of the structure of something unstructured might look illogical at the first glance, but IDPs become locally structured upon binding partner proteins, often via linear motifs or short α -helices, but sometimes via more pronounced folding-upon-binding events (Dyson and Wright, 2005b; Tompa, 2005; Sugase et al., 2007b). Specific challenges exist when working on IDPs, notably protease sensitivity and the difficulty in designing meaningful expression constructs; both of these have been addressed by the use of ESPRIT. However, once obtained, such proteins can be surprisingly easy to handle since they have no structure to denature. One demonstration of this is extreme heat resistance that was exploited in this work as an enrichment step after *E. coli* cell lysis, prior to affinity purification.

Some insights into how EBNA2 binds STAT3 may be obtained from the transactivation domain (TAD) of the herpes simplex protein VP16, which shares functional and structural features with EBNA2 (Ansieau and Leutz, 2002, 16). This was shown to undergo conformational changes from random coil to significant α -helical content when binding partner proteins (Jonker et al., 2004; Uesugi et al., 1997). It was furthermore shown that the first interaction between VP16 TAD and binding partner was electrostatically driven and induced structural changes around hydrophobic residues important for the adoption of an α -helical conformation. In this work, EBNA2 and SMRT binding to STAT3 were studied by SPR in order to gain kinetic and mechanistic insights. Previous analyses of IDP interactions by SPR, such as the binding of a monoclonal antibody to the IDP Tau (Sevcik et al., 2007b), suggest a two-step mechanism can be used for fitting data (see chapter 2.1.3.1 and 2.2.5). This model best described the experimental data and is in line with a first weak electrostatic interaction followed by conformational change resulting in a second higher affinity specific binding event. This mechanism has been termed “fly-casting” (Shoemaker et al., 2000a) and complements the long established “lock and key” hypothesis relating structure with function.

Several ideas have been proposed regarding biological advantages of IDPs: first, cell cycle regulation may require a fast turnover of proteins which is facilitated by the fact that IDPs are more sensitive to proteolytic degradation (Kriwacki et al., 1996) (see chapter 2.3.2). Second, the increased interaction area of IDPs may allow an easier capture of binding partners (Shoemaker et al., 2000b) and also an increased interaction speed when folding-upon-

binding, rather than prior to binding, was suggested as it leads to a significant reduction of the free-energy barrier (Huang and Liu, 2009). It has been also suggested that IDPs confer stability to complex regulatory networks as they are less sensitive to environmental changes e.g. temperature increase (Lee et al., 2001). Post-translational modification and alternative splicing sites are very often located within disordered regions conferring a high variability and adaptability upon a single primary sequence (Dunker et al., 2008). This variability as well as the possibility to fold-upon-binding allows many specific interactions to be made with multiple partners, perhaps explaining the apparent promiscuity of many IDPs.

Both SMRT and EBNA2, are known to undergo interactions with many binding partners. SMRT is a huge hub-protein with a length of 2525 amino acids and so the high number of interaction partners described (Dhordain et al., 1998b; Li et al., 2000b; Nakajima et al., 2001b; Jepsen and Rosenfeld, 2002b) is to be expected. But EBNA2, at 487 aa, is quite small for its number of interaction partners (Tong et al., 1995; Ansieau and Leutz, 2002, 69; Kwiatkowski et al., 2004; Calderwood et al., 2007; Muromoto et al., 2009b) and perhaps this can now be explained by its lack of order and high flexibility.

Knowledge of the structure of the binding interface between an IDP and its partner may ultimately lead to the development of new anti-cancer drugs that act by inhibiting complex formation. Because of their key roles in many cell processes and in many diseases, as well as their abundance in the proteome, IDPs are clearly interesting research targets. The study of IDPs only began around 15-years ago but it is becoming a very active area. Bioinformatics has played a key role in identifying IDPs, and high throughput studies in cataloging their interactions; now more effort should be made to validate these interactions and their mechanisms. In addition to investigating EBV biology, the work presented here has sought to address the areas of sample production and analysis in a manner that should be applicable to IDPs more generally.

3.5. The interaction triad: EBNA2 competing with SMRT for STAT3 binding

Under physiological conditions the co-repressor SMRT binds to many adaptors and transcription factors (e.g. STAT3) and mediates transcriptional suppression via recruitment of other repressing proteins such as HDACs that cause deacetylation of target genes in order to inhibit their transcription (Battaglia et al., 2010). Upon EBV infection of the cell, EBNA2 was shown to release SMRT from the complex (Muromoto et al., 2009a). This might be expected to activate STAT3 mediated transcription and happens presumably through recruitment of activating enzymes e.g. the basal transcription machinery (Tong et al., 1995) (Fig. 3.1), although currently no studies have shown this. Activated STAT3 signaling is pro-proliferative (Zushi et al., 1998; Bromberg et al., 1999; Rahaman et al., 2002), anti-apoptotic (Zushi et al., 1998; Shen et al., 2001; Aoki et al., 2003; Bhattacharya et al., 2005) and has several immune suppressive effects (Kortylewski et al., 2005a, 2009; Berg et al., 1996; Gerosa et al., 2008) (see 1.2.2 for details). Furthermore, it was hypothesised for HSV that STAT3 is required for the maintenance of the latent state as interference with STAT3 functions leads to reactivation of the latent virus in ganglia (Du et al., 2013). Because of the close similarity of both viruses this might be transferable to EBV. It was found recently that B cells containing a dominant negative mutation of STAT3 are resistance to EBV induced cell outgrowth (Koganti et al., 2013). It was therefore hypothesised that EBNA2 mediated STAT3 activation is highly connected with oncogenesis. Abberant STAT3 activation could be also found in other herpes viruses e.g. HCMV (Reitsma et al., 2013), KSHV (King, 2013) and *Herpesvirus saimiri* (Chung et al., 2004). This reinforces the evidence that STAT3 is a crucial factor for cell proliferation and survival of EBV infected B cells.

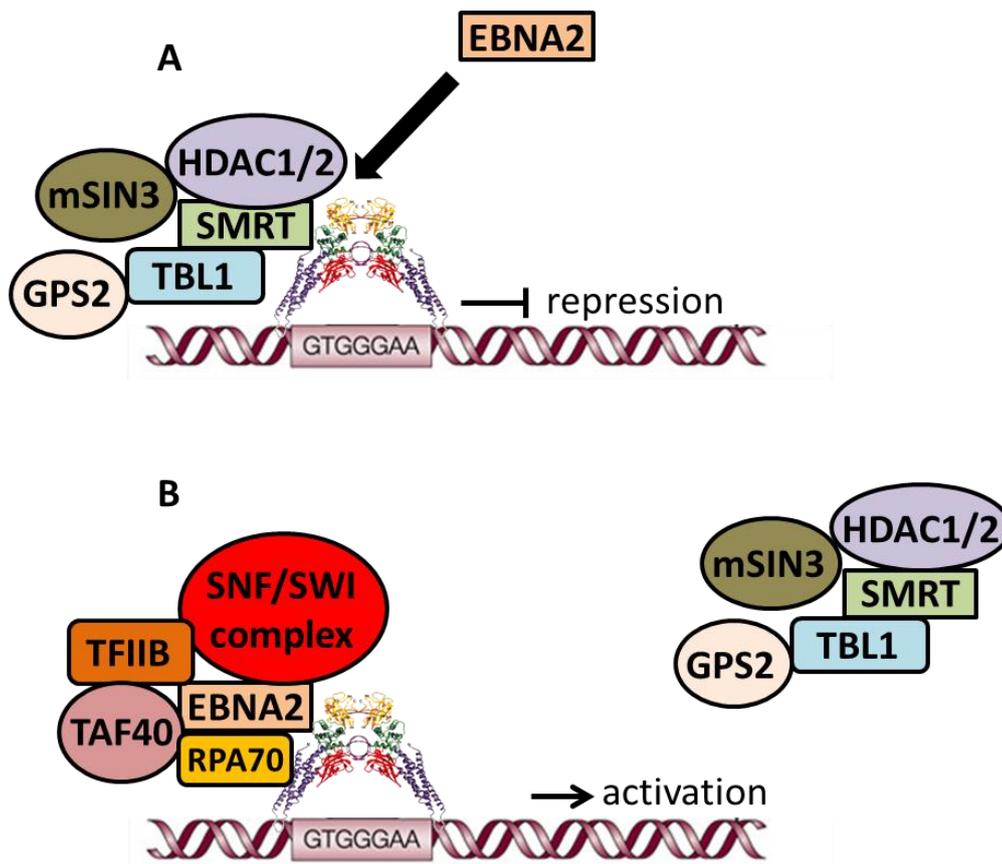


Figure 3.1: Proposed mechanism of STAT3 activation by EBNA2. A: SMRT binds to STAT3 and suppresses its transcriptional activity via recruitment of other repressing proteins. B: Upon EBV infection of the cell, EBNA2 releases SMRT from the complex and causes STAT3 mediated transcription activation through recruitment of activating.

A model is hereby proposed (Fig. 3.1) that takes into account the measurements made during this work. EBNA2 and SMRT compete for binding to STAT3 and EBNA2 is able to displace SMRT through a higher affinity for STAT3. The K_D values as measured by SPR are 7 nM for EBNA2 (Fig. 2.7) compared with 120 nM for SMRT (Fig. 2.24). This is in agreement with the observation that many viruses use short non-globular interaction interfaces that are adapted from host protein binding interfaces but which exhibit a higher affinity for the partner protein (Davey et al., 2011). In the context of IDPs, host linear motifs that bind with a medium affinity necessary for their function can be mimicked easily by viruses with adaptation of the sequence for high affinity.

The EBNA2-SMRT-STAT3 triad represents only one aspect of how EBV hijacks the host cell processes and establishes a lifelong alliance with the host cell. Many other processes clearly occur, but are beyond the scope of this work that focuses on interactions of recombinant purified proteins. In addition to the proposed cell based work which aims to study the biological significance of the binding it would be interesting to study and compare the dynamics of the viral and physiological interaction. A new tool to study protein dynamics is

in-cell NMR (Selenko and Wagner, 2007; Serber et al., 2007; Binolfi et al., 2012; Tochio, 2012). It permits the study of protein dynamics at atomic resolution in living cells. In order to record a protein spectrum inside a cell the protein of interest needs to be selectively isotope-labelled and highly overexpressed if *E. coli* is the host, or microinjected into mammalian cells which might have negative effects on biological function (Li and Liu, 2013) but the technique is still in its infancy and improvements are ongoing. Nevertheless, the *in vitro* methods used here may soon be complemented by more *in vivo* type measurements in the crowded milieu of the cellular cytoplasm/nucleoplasm.

Material and Methods

4.1. ESPRIT technology

The synthetic gene sequence of EBNA2 from EBV strain B95-8 (encoding a protein identical to UniProt P12978 EBNA2_EBVB9) was purchased from Genart with codon optimisation for *E. coli* expression. This process recodes the gene sequence to employ codons preferred by *E. coli*, eliminates internal ribosome binding sites and reduces mRNA secondary structure.

The expression vector pESPRIT002 contains a N-terminal 6xHis tag followed by a TEV protease cleavage site (MGHHHHHDYDIPTTENLYFQG) and a C-terminal biotin acceptor peptide with linker (SNNGSGGGLNDIFEAKIEWHE). In order to direct the exonuclease III digest the gene within pESPRIT002 is flanked by two restriction sites on each end of the gene. *AatII/AscI* are located on the 5' end of the gene and *NsiI/NotI* are on the 3' end.

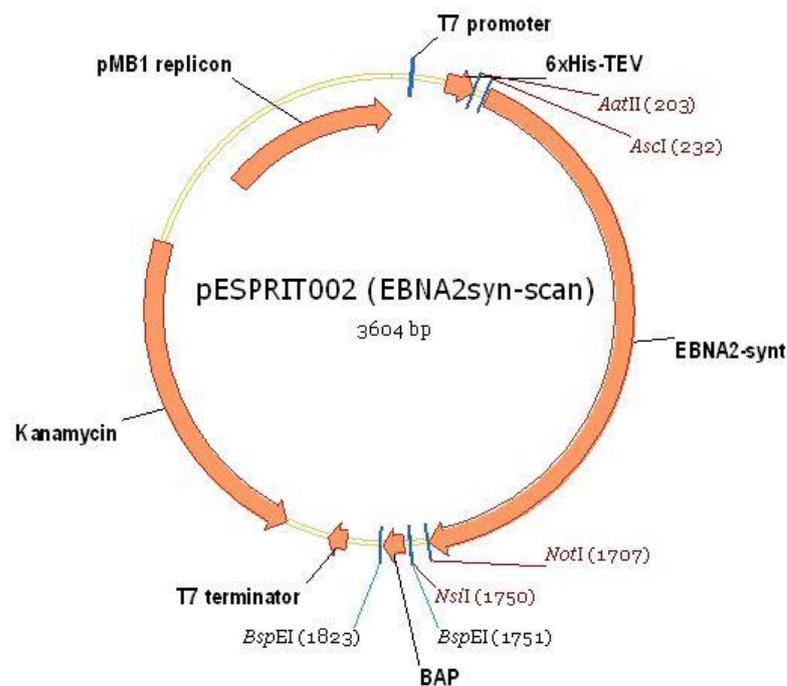


Figure 4.1: Vector map pESPRIT002 backbone vector. Restriction sites used in ESPRIT library synthesis are shown. The vector is derived from pET9a.

4.1.1. ESPRIT- experimental procedure

Manipulation of DNA

Mach1 *E. coli* cells (Invitrogen) were chemically transformed with pESPRIT002 containing the EBNA2 gene and DNA was isolated from 200 ml overnight culture by alkaline lysis and phenol chloroform extraction and ammonium acetate precipitation. The plasmid was further purified with a DNA miniprep kit (Qiagen) in order to remove residual RNA and salts. Ten micrograms of plasmid were digested with *Aat*II and *As*cl for 5' truncation, and *Not*I and *Ns*I for 3' truncation for 4 h in 200 µl reactions. Four micrograms of linearised plasmid were used for the truncation with 400 U of exonuclease III (NEB) at 22 °C in 120 µl total volume. The kinetics of the truncation reaction were adjusted to the EBNA2 DNA size by addition of 50 mM NaCl. Two microlitre of the digestion mixture were removed every 60 s and quenched into a single tube containing 200 µl of 3 M NaCl. After 60 min the exonuclease was heat inactivated at 70°C for 20 min. DNA was purified with a Nucleospin Extract kit (Macherey Nagel) and eluted in 35 µl of elution buffer. The single stranded DNA extension was removed with Mung Bean nuclease (NEB) (5 U enzyme, 30°C, 30 min) and *Pfu* polymerase (Stratagene) (72 °C, 20 min, 25 mM dNTPs). The unidirectionally truncated DNA was electrophoresed at 5 V/cm on an ethidium bromide-containing 0.5% w/v agarose gel at 4°C. The smear visible on the gel was cut and gel purified with Nucleospin Extraction kit (Macherey Nagel) and purified DNA recircularised by ligation with a Rapid Ligation kit (Roche) and recovered by transformation of Mach1 cells. All colonies were recovered as a single pellet from LB agar plates and plasmid DNA extracted. The protocol was repeated to truncate the 3' end of the gene. This time fractions of 0-500 bp, 500-1000 bp and 1000-1500 bp were isolated from the agarose gel (Fig. A.1) and DNA extracted, recircularised by ligation Mach1 cells transformed. Approximately 30,000 colonies were pooled from LB agar plates and plasmids containing randomly bidirectionally truncated EBNA2 inserts were extracted using a midiprep kit (Qiagen) and used to transform electrocompetent BL-21 AI (DE3) RIL cells (Invitrogen).

Screening of the generated diversity of library clones

Transformed cells were spread on Qtrays (Genetix) containing 300 ml LB agar supplemented with kanamycin (50 µg/ml) and chloramphenicol (50 µg/ml). A picker-gridder robot (KBiosystems) isolated around 28000 clones into 384 well plates (Genetix) containing 80 µl of TB with 10% final glycerol. Cultures were grown overnight at 37 °C in a HiGro incubator (GeneMachines) and frozen. Freshly replicated copies of these plates were robotically arrayed onto a nitrocellulose membrane (Amersham) laid on a Qtray containing 300 ml LB agar with antibiotics. Colonies were grown overnight at room temperature and the nitrocellulose membrane was transferred to a fresh Qtray containing LB agar, antibiotics, 50

μ M biotin and 0.2% L-arabinose. Protein induction was performed for 4 h at 30 °C. The membrane was transferred onto filter paper soaked with denaturing solution (500 mM NaOH, 1.5 M NaCl) for 10 min to lyse the colonies and then neutralised with neutralisation solution pH7.5 (1M Tris HCl, 1.5 M NaCl) twice for 5 min. Finally, the nitrocellulose array was incubated in 2 \times SSC solution pH7 (300 mM NaCl, 35 mM Na Citrate) for 15 min to remove cell debris. The activated membrane was placed in a roller incubator (Techne) and unspecific binding sites blocked overnight with 50 ml of SuperBlock (Pierce). Then the membrane was incubated with mouse anti-his-tag antibody (dilution 1/3125, Amersham) in 0.1% Tween-PBS for 1 h at 4 °C. The membrane was washed 3 x 5 min with 0.1% Tween-PBS. Streptavidin labelled with Alexa488 (dilution 1/5000, Molecular probes) and secondary antibody anti-mouse Alexa532 (dilution 1/1000) were incubated with the membrane for 45 min at 4 °C. Three final washing cycles with 0.1% Tween-PBS and one washing cycle with water followed. The membrane was scanned using a Typhoon imager (GE Healthcare) and analysed using VisualGrid software (GPC Biotech). Data were transferred to MS Excel for analysis and clones exhibiting strong signals for both histidine and biotin tags identified.

Selection of the best protein clones

The 96 best clones were transferred into a 96 deep-well plate (Qiagen) and grown overnight in the HiGro plate shaker (37 °C at 220 rpm). Protein expression trials were performed in 4 ml TB medium in 24 deep-well plates (Qiagen). Cultures were grown at 37 °C and induced at OD 600 0.6-0.8 with 0.2% arabinose for 16 h at 25 °C. Cells were harvested by centrifugation at 3700 rpm for 10 min in a swinging rotor and the cell pellets resuspended with 4 ml of resuspension buffer containing 20 mM Tris pH 8, 250 mM NaCl, 20% sucrose and 1 mg/ml lysosyme and incubated 15 min in cold room on a rocking platform. The resulting sphaeroplasts were pelleted at 3700 rpm for 15 min. Pellets were kept at -80 °C for at 1 h and then resuspended in 700 μ l of lysis buffer containing 10 mM Tris pH 8, 0.5% Brij58, 1/1000 dilution of benzonase, 5 mM β -mercaptoethanol and protease inhibitor cocktail. Protein purifications were performed using using a TECAN robot: proteins were incubated with 300 μ l of NiNTA agarose resin (Qiagen) for 30 min at 4 °C for batch binding and washed with 10 volumes of washing buffer (300 mM NaCl, 50 mM Na₂HPO₄/NaH₂PO₄ pH 7.5, 10 mM imidazole, 5 mM β - mercaptoethanol). Proteins were eluted into 70 μ l of elution buffer (300 mM NaCl, 50 mM Na₂HPO₄/NaH₂PO₄ pH 7.5, 300 mM imidazole, 5 mM β -mercaptoethanol). The expression level was analysed by SDS-PAGE and western blot analysis. The BAP tag of the three best clones was removed by *Bsp*EI digestion and relegation with Rapid DNA ligation kit (Roche).

4.2. Protein expression and purification

4.2.1. Expression and purification of the STAT3 homodimer

The STAT3 β expression vector is a pET32a vector (Novagen) which contains the STAT3 β gene truncated at the N-Terminus, covering amino acids 127-722 (provided by C. Müller, EMBL, Heidelberg) who published the expression protocol and crystal structure of the STAT3 β dimer in 1998 (Becker et al., 1998a, 1998b).

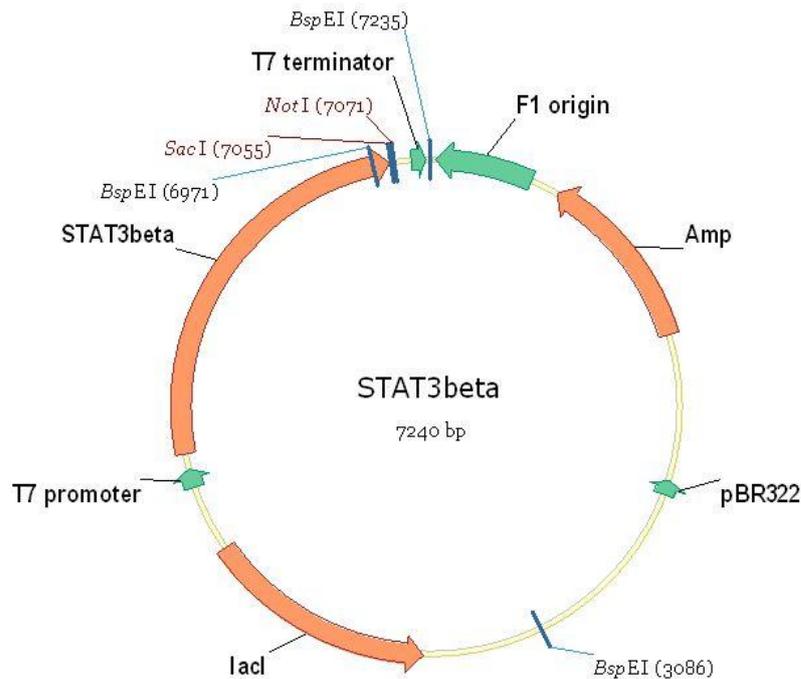


Figure 4.2: Vector map STAT3 β expression vector. Strong expression of the STAT3 beta gene is via a T7 promoter.

TKB1 cells were chemically transformed with the STAT3 plasmid (Stratagene). These engineered *E. coli* cells express the Elk receptor kinase domain which is required for the SH2 domain phosphorylation and dimerisation of STAT3 (Letwin et al., 1988). Cells were grown in 1 l bottles containing LB medium supplemented with 50 μ g/ml ampicillin and 12.5 μ g/ml tetracycline at 37 $^{\circ}$ C. At OD (600 nm) 0.4 the temperature was switched to 21 $^{\circ}$ C and at OD 600 0.6 expression of STAT3 was induced with 1 mM IPTG (Thermo Scientific) overnight. The next day medium was changed to kinasing medium (M9 medium supplemented with 1 mM MgSO₄, 11mM glucose, 0.1% casamino acids, 1.5 mM thiamine-HCl, 53 mM 3 β -indoleacrylic acid, supplemented with 50 mg/ml ampicillin and 12.5 mg/ml tetracycline) and cells were grown for additional 2.5 h at 37 $^{\circ}$ C. Cells were harvested via centrifugation at 7000 rpm for 20 min and pellets resuspended in extraction buffer (20 mM

HEPES-HCl, pH 7.6, 0.1 M KCl, 10% glycerol, 1 mM EDTA, 10mM MnCl₂, 20 mM DTT, protease inhibitor cocktail). Cells were lysed using a microfluidiser (M110-L; Microfluidics Corporation, USA) and the lysate was cleared by centrifugation at 27000 g for 45 min at 4 °C. Nucleic acids and associated proteins were removed by addition of 0.1% polyethyleneimine final concentration to the ice cooled, stirred supernatant, incubated for 15 min and centrifuged at 27000 g for 20 min. STAT3 was precipitated with ammonium sulfate at 35% saturation and the protein pellet was dissolved in Buffer D (20 mM HEPES-HCl, pH 7.0, 200 mM NaCl, 10 mM MgCl₂, 5 mM DTT, protease inhibitor) and dialysed over night against Buffer D in order to remove remaining ammonium sulfate. After dialysis STAT3 was further purified by size exclusion chromatography (Pharmacia Superose 12 HR 10/30 column, GE Healthcare). The purity was evaluated by SDS-PAGE electrophoresis followed by Coomassie blue staining. Fractions with highest purity were pooled and stored at -80 °C.

4.2.2. Expression and purification of EBNA2 fragments

The EBNA2 expression clones F3, A11 and D9 were generated by ESRIT technology and have a pESPRIT002 backbone vector.

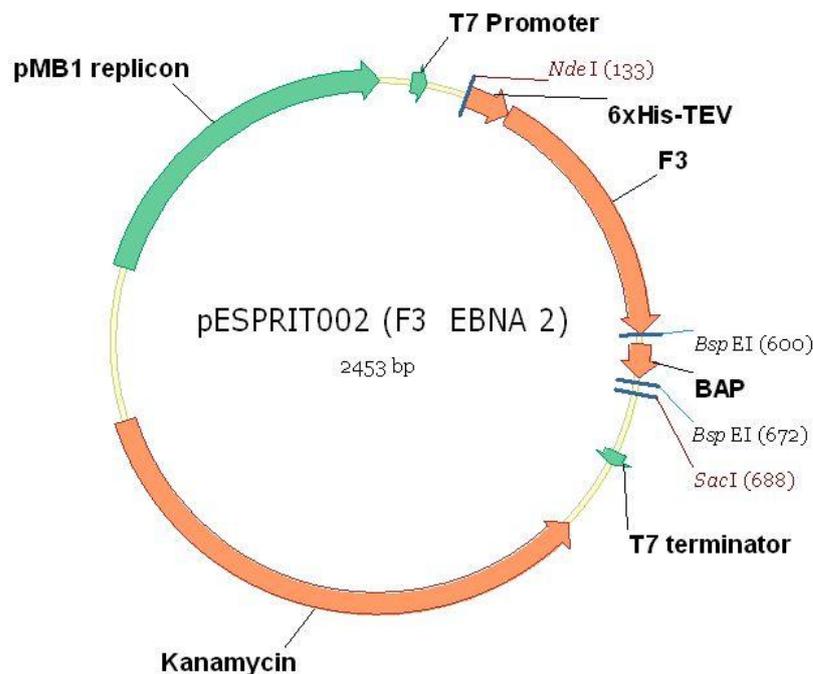


Figure 4.3: Vector map pESPRIT002 EBNA2 F3. Expression vector for EBNA2 fragment F3 with the biotin acceptor peptide (BAP) included.

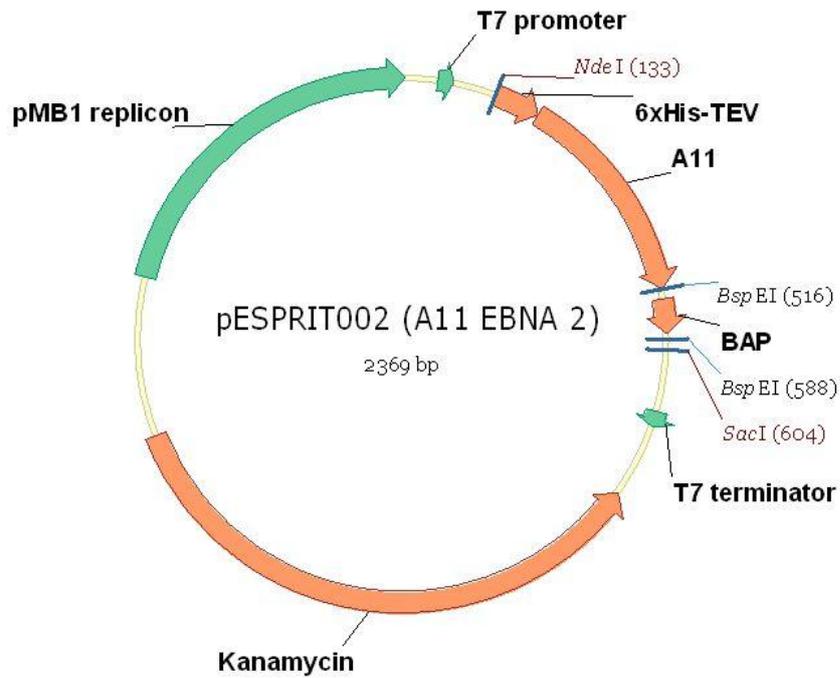


Figure 4.4: Vector map pESPRIT002 EBNA2 A11. Expression vector for EBNA2 fragment A11 with the biotin acceptor peptide (BAP) included.

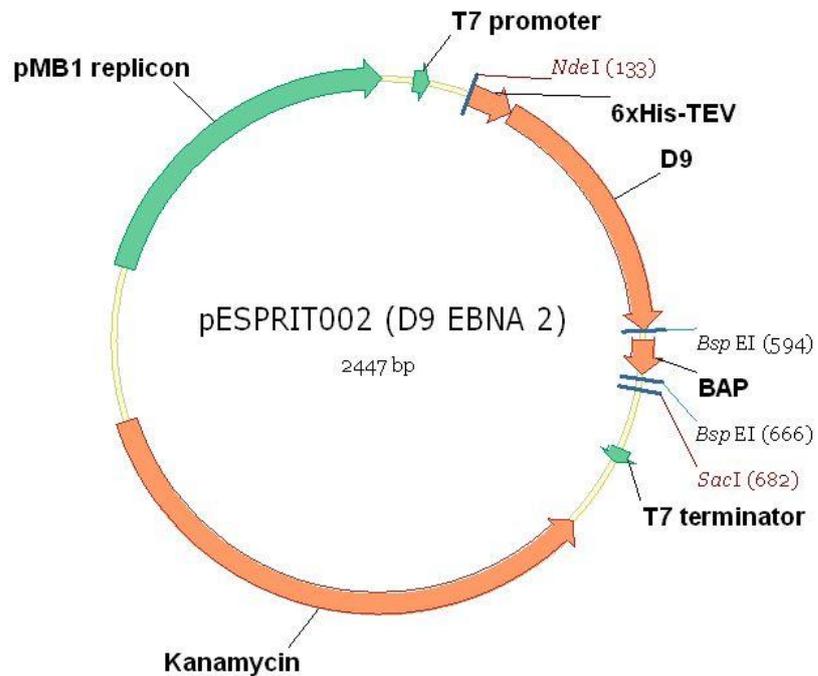


Figure 4.5: Vector map pESPRIT002 EBNA2 D9. Expression vector for EBNA2 fragment D9 with the biotin acceptor peptide (BAP) included.

BL-21 AI (DE3) RIL cells (Invitrogen) were transformed with the EBNA2 expression plasmids. Cells were grown in 1 l bottles containing LB medium supplemented with 50 µg/ml kanamycin and 50 µg/ml chloramphenicol at 37 °C. At OD (600 nm) 0.6 the temperature was switched to 25 °C and protein expression was induced with 0.2% L-arabinose for 16 h. Cells were harvested by centrifugation at 7000 rpm for 20 min and the pellet was resuspended in lysis buffer (25 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10 mM Imidazole, 5 mM β-mercaptoethanol, protease inhibitor cocktail (Roche), 1 % Tween and Benzonase). Cells were lysed by sonication and cell debris removed by centrifugation at 27000g for 45 min. The supernatant was incubated for 2 h with Ni²⁺-NTA agarose resin (Qiagen) at 4 °C and then loaded onto a plastic column. This was washed with 100ml of washing buffer (25 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole and 5 mM β-mercaptoethanol) and eluted in elution buffer (50 mM Tris pH8.0, 300 mM KCl, 5mM β-mercaptoethanol, and 250 mM imidazole). The purity was evaluated by SDS-PAGE and Coomassie blue staining. Fractions with highest purity were pooled, concentrated with Amicon Ultra concentrator (Millipore) and subjected to gel filtration (Superdex S75 10/300 column, GE Healthcare). Peak fractions were collected and pooled. Then the hexahistidine tag was removed by incubation with hexahistidine tagged tobacco etch virus protease (TEV) overnight at 4 °C. The cleaved his tag was removed by reversed affinity chromatography with Ni²⁺-NTA agarose resin (Qiagen). Cleaved EBNA2 fragment were concentrated and stored at -80 °C.

4.2.3. Expression and purification of SMRT fragments

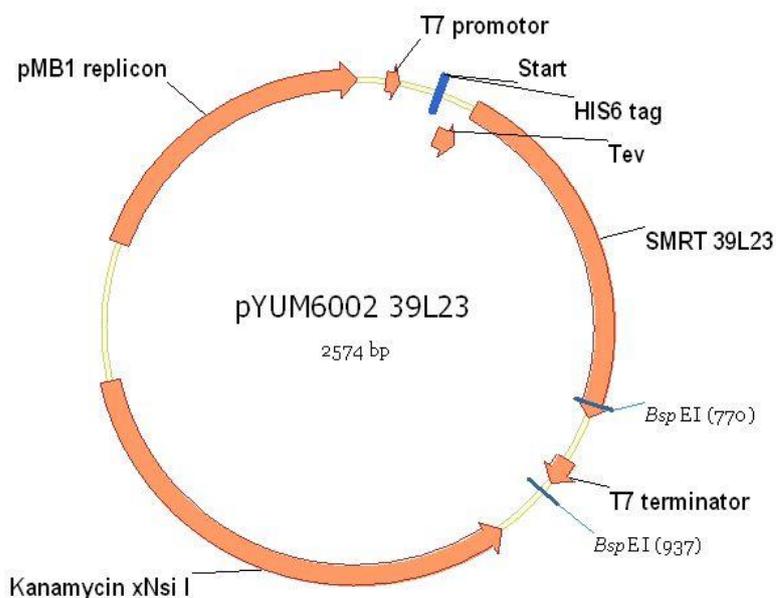


Figure 4.6: Vector map pYUM6002 2 39L23. Expression vector of SMRT fragment 39L23 which was kindly provided by D. Desravines (Hart team , EMBL, Grenoble).

The 39L23 expression construct was generated previously using ESPRIT technology (D. Desravines and P. Mas).

BL-21 AI (DE3) RIL cells (Invitrogen) were transformed with the 39L23 expression plasmid by electroporation. Cells were grown in 1 l bottles containing LB medium supplemented with 50 µg/ml kanamycin and 50 µg /ml chloramphenicol at 37 °C. At OD (600 nm) 0.6 the temperature was shifted to 25 °C and protein expression was induced with 0.2% L-arabinose overnight. The next day cells were harvested by centrifugation at 7000 rpm for 20 min and the pellet was resuspended in lysis buffer (25 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10 mM imidazole, 5 mM β- mercaptoethanol, protease inhibitor cocktail (Roche), 1 % Tween and benzonase). Cells were lysed by sonication and cell debris removed by centrifugation at 27000 g for 45 min. The cleared supernatant was then incubated at 80 °C for 10 min until the supernatant appeared milky. Aggregated proteins were removed by centrifugation at 27000 g for 20 min. The supernatant was incubated for 2 h with Ni²⁺-NTA agarose resin (Qiagen) at 4 °C and then loaded onto a plastic column. This was washed with 100ml of washing buffer (25 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole and 5 mM β- mercaptoethanol) and proteins were eluted in elution buffer (50 mM Tris pH8.0, 300 mM KCl, 5mM β- mercaptoethanol, and 250 mM imidazole). The hexahistidine tag was removed by incubation with hexahistidine tagged tobacco etch virus protease (TEV) overnight at 4 °C. The cleaved his tag was removed by reversed affinity chromatography with Ni²⁺-NTA agarose resin (Qiagen). The purity was evaluated by SDS-PAGE electrophoresis followed by Coomassie blue staining. Fractions with highest purity were pooled and stored at -80 °C after concentration to the desired concentration by Amicon Ultra concentrator (Millipore).

4.2.4. Expression and purification of EBNA2-TAs-pGEX-4T and ZMYND11-pGEX-4T

BL-21 AI (DE3) RIL cells (Invitrogen) were transformed with ZMYND-pGEX-4T and EBNA2-TAs-pGEX-4T by electroporation. Single colony cells were grown in 1 l bottles containing LB medium supplemented with 50 µg/ml ampicillin at 37 °C. At OD (600 nm) 0.6 the temperature was shifted to 25 °C and protein expression was induced with 1 mM IPTG (Thermo Scientific) for 16 h. Cells were harvested by centrifugation at 7000 rpm for 20 min and the pellet was resuspended in lysis buffer (25 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10 mM Imidazole, 5 mM β- mercaptoethanol, protease inhibitor cocktail (Roche), 1 % Tween and benzonase). Cells were lysed by sonication and cell debris removed by centrifugation at 27000 g for 45 min. The supernatant was incubated with glutathione sepharose (GE Healthcare) overnight at 4 °C and then loaded onto a plastic column. Proteins were washed with 100ml of PBS 0.05%

Tween and eluted in elution buffer (10 mM reduced glutathione in 50 mM Tris-HCl, pH 8.0). The GST tag was removed by incubation with Thrombin Cleave Clean kit (Sigma) according to the manufacturer's protocol. The cleaved GST tag was removed by reversed affinity chromatography with glutathione sepharose beads. The purity was evaluated by SDS-PAGE and Coomassie blue staining. Fractions with highest purity were pooled and stored at -80 °C after concentration to the desired concentration by Amicon Ultra concentrator (Millipore).

4.3. Kinetic binding analysis

4.3.1. Surface plasmon resonance (SPR)

SPR measurements were performed on a Biacore 3000 instrument (Biacore AB, Uppsala, Sweden) using four flow cells. The flow cell temperature was 25 °C in all experiments. EBNA2 and SMRT fragments were immobilised on CM5 sensorchips by conventional amine reactive chemistry. The surface of the carboxy-methylated dextran CM5 chip (Biacore AB, Uppsala, Sweden) was activated with 70 µl of a 1:1 mixture of 200 mM EDC (N-ethyl-N'-(3-diethylaminopropyl) and 50 mM NHS (N-hydroxysuccinimide). After surface activation the ligand which was diluted in 10 mM sodium acetate buffer pH 6.5 was injected until an immobilisation level of about 2000-2500 RU (response units) was reached. Free active esters were blocked by injection of 70 µl of 1M ethanolamine/hydrochloride at pH 8.5. Immobilisation flow rate was 10µl/ min The analyte STAT3 was diluted in Buffer B, pH 6.5 (20 mM HEPES-HCl, pH 7.0, 200 mM NaCl, 10 mM MgCl₂, 5 mM DTT, 0.05% Tween-P20) and injected at various concentration according to the affinity of the binding. The concentration ranged from 500 pM for high affinity binding to 10 µM for fragments that did not bind. The flow rate of binding experiments was 30 µl/min After each injection bound proteins were removed by injection of 30 µl of 1M MgCl₂. Binding curves were analyzed using both BIAevaluation software (Biacore AB, Uppsala, Sweden) and manually using saturation binding values. The simple 1:1 Langmuir model and the more complex two-state model were both used to fit the binding data. The two-state model assumes that analyte (A) and ligand (L) form a first complex (AL) which is followed by a conformational change and the formation of a second more stable complex (AL*). From the two association and dissociation rate constants an apparent affinity constant can be derived. The 1:1 Langmuir model is based on the assumption of a simple 1:1 binding event without conformational change or multiple binding sites. It can be used to obtained dissociation and association rate constants as well as the equilibrium dissociation constant.

1:1 Langmuir model

Constant	Description	Unit
k_a	Association rate for formation of $A + L \rightarrow AL$	$M^{-1}s^{-1}$
k_d	Dissociation rate for $AL \rightarrow A + L$	s^{-1}
KD	$KD = \frac{kd}{ka}$	M

Two-state model

Constant	Description	Unit
k_{a1}	Association rate for formation of $A + L \rightarrow AL$	$M^{-1}s^{-1}$
k_{a2}	Association rate for formation of $AL \rightarrow AL^*$	s^{-1}
k_{d1}	Dissociation rate for $AL \rightarrow A + L$	s^{-1}
k_{d2}	Dissociation rate for $AL^* \rightarrow AL$	s^{-1}
KD	$KD = \frac{kd1 \times kd2}{ka1(ka2 + kd2)}$	M

4.3.2. Microscale thermophoresis (MST)

For the MST experiment 100 μ l of STAT3 at 20 μ M were fluorescently labelled with a Red fluorescent dye (NT-647) using Monolith NT™ Protein Labeling Kit RED (NanoTemper). EBNA2 fragments were diluted in Buffer B , pH 6.5 (20 mM HEPES-HCl, pH 7.0, 200 mM NaCl, 10 mM MgCl₂, 5 mM DTT, 0.05% Tween-P20) and 16 dilutions (1:1) were pipetted. The concentration ranged from 2 μ M to 62.5 pM in 10 μ l volume. The concentration of STAT3 was set constantly to 20 nM in all experiment. The proteins mixture was centrifuged at 27000 g for 15 min and soaked into standard capillaries (NanoTemper). Mobility changes were measured using the Monolith NT.115 instrument (NanoTemper) and binding generated. The binding affinity was calculated using the NanoTemper evaluation software.

4.3.3. Nuclear magnetic resonance (NMR) sample preparation

EBNA2 fragment D9 and SMRT fragments were isotopically labelled by expression in M9 minimal medium supplemented with ammonium ¹⁵N chloride. The adaption to minimal medium required two steps by transfer of the 10 ml LB starting culture first in 40 ml minimal medium and then in 1 l. After purification proteins were dialysed in 50 mM potassium phosphate buffer with 150 mM of KCl, pH 6.5. Before starting the measurement samples were supplemented with 10% D₂O (v/v). All NMR spectra were recorded by R. Schneider (Blackledge NMR group, IBS, Grenoble). All experiments were recorded at 25°C using 600 MHz or 800 MHz Varian spectrometers and analyzed using NMRPipe and Sparky software. The concentration of labelled proteins was 250 µM alone or decreased to nearly 50 µM when in complex with STAT3. Spectra of EBNA2 D9 in complex with STAT3 were recorded at a 1:1 ratio and spectra of the complex between SMRT 39L23 and STAT3 at 1:5, 1:3, 1:1, 3:1 and 5:1 ratios.

4.4. Streptavidin agarose pull-down experiments

EBNA2 fragments F3, A11 and D9 with BAP were expressed in LB medium supplemented with 50 µM biotin (Sigma). Magnetic streptavidin beads (Dynabeads, Invitrogen) were washed three times with PBS-Tween (0.5%), then blocked in 5% fetal calf serum (FCS, Sigma) for 1 h. The FCS solution was aspirated and beads were incubated with EBNA2 fragment lysate for 1 h. Beads were washed three times with PBS-Tween (0.5%), then STAT3 lysate was added and incubated for 30 min. Beads were again washed three times for 10 min. All incubations were performed at 4 °C with agitation. Proteins were eluted by addition of Laemmli buffer at 95 °C for 5 min. Eluted proteins were loaded on a 12% SDS-PAGE (running time 1 h, 180 V) and then transferred to a nitrocellulose membrane by semi-dry western blot transfer (Biorad) (20 V, 30 min). Proteins were detected using a rabbit STAT3 polyclonal antibody (Cell Signaling) and an anti-rabbit antibody-ALEXA488, and Streptavidin-ALEXA633. Immunofluorescence signals were detected using a Typhoon scanner (GE Healthcare).

4.5. Generation of interaction mutants

4.5.1. Generation of GST tagged EBNA2-CR8 fragments

The different EBNA2 fragments were amplified from EBNA2-D9-pESPRIT002 by polymerase chain reaction (PCR).

Primer

ATGCATGAATTCCCAATTCATGAACCGGAAAGCCATAATTCA	(D9TA-FL forward)
ATGCATGAGCTCTAGAAAGATATAATCCCAGCTTTCATCCAGATCTGC	(D9TA-FL reverse)
ATGCATGAATTCCCACCGGAAAGCCATAATTCACCGGAAGCA	(D9TA-1 forward)
ATGCATGAGCTCTAGATCCCAGCTTTCATCCAGATCTGCCGG	(D9TA-1 reverse)
ATGCATGAATTCCCACATAATTCACCGGAAGCACCGATTCTG	(D9TA-2 forward)
ATGCATGAGCTCTAGCAGAAATCGGTGCTTCCGGTGAATTATG	(D9TA-2 reverse)
ATGCATGAATTCCCAATTCATGAACCGGAAAGCCATAATTCA	(D9TA-3 forward)
ATGCATGAGCTCTAGATCCCAGCTTTCATCCAGATCTGCCGG	(D9TA-3 reverse)
ATGCATGAATTCCCAATTCATGAACCGGAAAGCCATAATTCA	(D9TA-4 forward)
ATGCATGAGCTCTAGCAGAAATCGGTGCTTCCGGTGAATTATG	(D9TA-4 reverse)
ATGCATGAATTCCCACCGGAAAGCCATAATTCACCGGAAGCA	(D9TA-5 forward)
ATGCATGAGCTCTAGAAAGATATAATCCCAGCTTTCATCCAGATCTGC	(D9TA-5 reverse)
ATGCATGAATTCCCACATAATTCACCGGAAGCACCGATTCTG	(D9TA-6 forward)
ATGCATGAGCTCTAGAAAGATATAATCCCAGCTTTCATCCAGATCTGC	(D9TA-6 reverse)

Template vector (50 ng) were supplemented with 500 μ M dNTPs, 1x Pfu reaction buffer, 200 nM forward and reverse primer and 2.5 U Pfu polymerase (Stratagene). The volume was adjusted to 25 μ l with water. The reaction was performed in a ThermoCycler (Eppendorf) according to the following cycling parameters:

	Temperature ($^{\circ}$ C)	Time (sec)	Cycles
Denaturation	95	60	1
Denaturation	95	30	30
Annealing	55	30	
Elongation	72	60	
Final elongation	72	300	1

The fragments were purified using a NucleoSpin DNA Extraction Kit (Macherey Nagel). Amplified PCR products and receiving vector pGEX-4T were digested with *Eco*RI and *Sac*I in buffer 1 (NEB) for 1 h. The digested vector was purified from an ethidium bromide-containing 1% w/v agarose gel using a NucleoSpin Extraction kit and dephosphorylated by shrimp

alkaline phosphatase (SAP, Roche) for 30 min at 37 °C. Vector and insert were ligated using the Rapid ligation kit (Roche) and transformed into NEB-10 *E. coli* cells (NEB). Positive clones were verified by DNA sequencing (Macrogen).

4.5.2. Site directed mutagenesis

The different EBNA2 fragments were amplified from EBNA2-D9-pESPRIT002 by polymerase chain reaction (PCR).

Primer

ATGCATGAATTCCCAATTCATGAACCGGAAAGCCATAATTCA	(D9TA-FL forward)
ATGCATGAGCTCTAGAAAGATATAATCCCAGCTTTCATCCAGATCTGC	(D9TA-FL reverse)
ATGCATGAATTCCCACCGGAAAGCCATAATTCACCGGAAGCA	(D9TA-1 forward)
ATGCATGAGCTCTAGATCCCAGCTTTCATCCAGATCTGCCGG	(D9TA-1 reverse)
ATGCATGAATTCCCACATAATTCACCGGAAGCACCGATTCTG	(D9TA-2 forward)
ATGCATGAGCTCTAGCAGAAATCGGTGCTTCCGGTGAATTATG	(D9TA-2 reverse)
ATGCATGAATTCCCAATTCATGAACCGGAAAGCCATAATTCA	(D9TA-3 forward)
ATGCATGAGCTCTAGATCCCAGCTTTCATCCAGATCTGCCGG	(D9TA-3 reverse)
ATGCATGAATTCCCAATTCATGAACCGGAAAGCCATAATTCA	(D9TA-4 forward)
ATGCATGAGCTCTAGCAGAAATCGGTGCTTCCGGTGAATTATG	(D9TA-4 reverse)
ATGCATGAATTCCCACCGGAAAGCCATAATTCACCGGAAGCA	(D9TA-5 forward)
ATGCATGAGCTCTAGAAAGATATAATCCCAGCTTTCATCCAGATCTGC	(D9TA-5 reverse)
ATGCATGAATTCCCACATAATTCACCGGAAGCACCGATTCTG	(D9TA-6 forward)
ATGCATGAGCTCTAGAAAGATATAATCCCAGCTTTCATCCAGATCTGC	(D9TA-6 reverse)

Template vector (50 ng) were supplemented with 500µM dNTPs, 1x *Pfu* reaction buffer, 200 nM forward and reverse primer and 2.5 U *Pfu* polymerase (Stratagene). The volume was adjusted to 25 µl with water. The reaction was performed in a ThermoCycler (Eppendorf) according to the following cycling parameters:

	Temperature (°C)	Time (sec)	Cycles
Denaturation	95	60	1
Denaturation	95	30	30
Annealing	55	30	
Elongation	72	60	
Final elongation	72	300	1

The fragments were purified using a NucleoSpin DNA Extraction Kit (Macherey Nagel). Amplified PCR products and receiving vector pGEX-4T were digested with *EcoRI* and *SacI* in buffer 1 (NEB) for 1 h. The digested vector was purified from a ethidium bromide-containing 1% w/v agarose gel using a NucleoSpin Extraction kit and dephosphorylated by shrimp alkaline phosphatase (SAP, Roche) for 30 min at 37 °C. Vector and insert were ligated using the Rapid ligation kit (Roche) and transformed into NEB-10 *E. coli* cells (NEB). Positive clones were verified by DNA sequencing (Macrogen).

4.5.3. Overlap extension PCR

EBNA2 point and deletion mutations in the mammalian expression vectors pAG155 and KG172 were generated by overlapping extension PCR. Therefore two fragments were amplified which are overlapping for a sufficient amount of base pairs.

Primers pAG155

GAGCCATCACCTCTTGATAGGGATCCGCTAGGATATGAC (BamHI forward)
 GAGCGAGGAAGCGGAAGAGTCTAGAGTCGACCAGAC (XbaI reverse)

GATGATTGGTATCCTCCAGCTATAGACCCCGCAGACTTAGAC (S448A forward)
 GTCTAAGTCTGCGGGTCTATAGCTGGAGGATACCAATCATC (S448A reverse)

CCCGCAGACTTAGACGAAAGTGCGGCTTACATTTTTGAGACAACAG (W458A, D459A forward)
 CTGTTGTCTCAAAAATGTAAGCCGCACTTTCGTCTAAGTCTGCGGG (W458A, D459A reverse)

CCCGCAGACTTAGACGAAAGTGCGGCTGCCATTTTTGAGACAACAG (W458A, D459A, Y460Afor)
 CTGTTGTCTCAAAAATGGCAGCCGCACTTTCGTCTAAGTCTGCGGG (W458A, D459A, Y460Arev)

AGCCTCTGGGCTATTATGGGACTCCGGTTCATGGTCTAAGTCTGCGGGGTC (ΔCR8 for)
 CATAATAGCCCAGAGGCTGACCCCGCAGACTTAGACGAAAGTTGGGATTACATTTTTGAG (ΔCR8re)

Primers KG172

CTTGATAGGGATCCGCTAGGATATGACGTCGGGCAT (BamHI forward)
 CAGCGACCCACCGCGGCCCGCCGCTGCC (NotI reverse)

CCTCCAGCTATAGACCCCGCAGACTTAGACGAAAGTTGG (S448A forward)
 CCAACTTTCGTCTAAGTCTGCGGGTCTATAGCTGGAG (S448A reverse)

AATAGCCCAGAGGCTGACCCCGCAGACTTAGACGAAAGTTGGGATTACATTTTTGAG (ΔCR8 for)
 TAAGTCTGCGGGGTCAGCCTCTGGGCTATTATGGGACTCCGGTTCATGTATTG (ΔCR8 rev)

Template vector at 100 ng (pAG155 or KG172 EBNA2 vector), 500 μ M dNTPs, 1x *Pfu* reaction buffer, 200 nM forward and reverse primer and 2.5U *Pfu* polymerase (Stratagene) were adjusted with water to 25 μ l. The first reaction used the parameters:

	Temperature ($^{\circ}$ C)	Time (sec)	Cycles
Denaturation	95	60	1
Denaturation	95	30	30
Annealing	55	30	
Elongation	72	120	
Final elongation	72	300	1

The amplified fragments were gel purified from a 1% ethidium bromide agarose gel using a NucleoSpin Extraction kit (Macherey Nagel) and both fragments were mixed in equal molar ratio supplemented with 500 μ M dNTPs, 1x *Pfu* reaction buffer and 2.5U *Pfu* polymerase. A second reaction cycle followed in order to generate one longer fragment out of the two shorter fragments. The second reaction followed these parameters:

	Temperature ($^{\circ}$ C)	Time (sec)	Cycles
Denaturation	95	60	1
Denaturation	95	30	20
Annealing	55	30	
Elongation	72	120	
Final elongation	72	300	1

Both, fragments and pAG155 EBNA2 vector were digested with *Bam*HI and *Xba*I/*Not*I in buffer 4 (*Xba*I) or 3 (*Not*I) supplemented with BSA for 2 h at 37 $^{\circ}$ C. Inserts and vector were gel purified from a 1% ethidium bromide agarose gel using a NucleoSpin Extraction kit (Macherey Nagel) and the vector was dephosphorylated by shrimp alkaline phosphatase (SAP, Roche) for 30 min at 37 $^{\circ}$ C. Vector and insert were ligated using the Rapid ligation kit (Roche) and transformed into Mach1 *E.coli* cells (Invitrogen). Positive clones were verified by DNA sequencing (Macrogen).

4.5.4. Generation of SMRT fragment mammalian expression vectors

The different SMRT fragments were amplified from SMRT iso 1 (1-1195) and SMRT iso 2 (1012-2075) by polymerase chain reaction (PCR). Fragments were inserted into pcDNA3-myc.

Primer

ATGCATGAATTCAAGAAGAAGCGTAAGGTAATGTCGGGATCCACACAGCCTG (SMRT1 forward)
 ATGCATGGATCCTCAGACCTGGCGGTCTTTGTACACCTTC (SMRT1 reverse)

ATGCATGAATTCAAGAAGAAGCGTAAGGTAATGAACATGTGGAGTGAGCAGGAGAAG (SMRT2 for)
 ATGCATGGATCCTCAGGGAAGGCCGAGGAGCCC (SMRT2 rev)

ATGCATGAATTCAAGAAGAAGCGTAAGGTAGTCAAGAGCGAGTGCACGGAGG (SMRT3 forward)
 ATGCATGGATCCTCATTCTTGCCTTCGTAGATGACGTGG (SMRT3 reverse)

ATGCATGAATTCAAGAAGAAGCGTAAGGTAGGGCCACGTCTTGTCTTATGAGG (SMRT4 forward)
 ATGCATGGATCCTCATGATGTAGTCATTGATGATGGTCTGCC (SMRT4 reverse)

ATGCATGAATTCAAGAAGAAGCGTAAGGTACCTCGCAGCAGATGCACCACA (SMRT5 forward)
 ATGCATGGATCCTCACCGCAGGTGTGGGAGGTGG (SMRT5 reverse)

ATGCATGAATTCAAGAAGAAGCGTAAGGTACCGCTGCCTGAGAGCCAGC (SMRT6 forward)
 ATGCATGGATCCTCATCACTCGCTGTCCGAGAGTGTCTC (SMRT6 reverse)

Fifty nanograms of template vector were supplemented with 500 μ M dNTPs, 1x Pfu reaction buffer, 200 nM forward and reverse primer and 2.5 U *Pfu* polymerase (Stratagene). The volume was adjusted to 25 μ l with water. The reaction was performed in a ThermoCycler (Eppendorf) according to the following cycling parameters:

	Temperature (°C)	Time (sec)	Cycles
Denaturation	95	60	1
Denaturation	95	30	30
Annealing	50	60	
Elongation	72	120	
Final elongation	72	300	1

The fragments were purified using a NucleoSpin DNA Extraction Kit (Macherey Nagel). Amplified PCR products and receiving vector pcDNA3-myc were digested with *EcoRI* and *BamHI* in buffer 3 (NEB) for 1 h. The digested vector was purified from an ethidium bromide-

containing 1% w/v agarose gel using a NucleoSpin Extraction kit and dephosphorylated by shrimp alkaline phosphatase (SAP, Roche) for 30 min at 37 °C. Vector and insert were ligated using the Rapid ligation kit (Roche) and transformed into NEB-10 *E.coli* cells (NEB). Positive clones were verified by DNA sequencing (Macrogen).

4.6. Electrophoretic mobility shift assay (EMSA)

An electromobility shift assay (EMSA) was performed using ALEXA488 fluorescently labelled STAT3 binding oligos (GAATCCTAAGTGCATTTCCCGTAAATCTTGAAGTCGCG). 1 μM of STAT3 binding oligo was incubated with the STAT3 homodimer (20 μM) alone or together with wild-type or ΔCR8 D9 EBNA2 proteins in different ratios (1:2, 1:1 and 2:1) for 30 min in the dark. Sample buffer was buffer B, pH 6.5 (20 mM HEPES-HCl, pH 7.0, 200 mM NaCl, 10 mM MgCl₂, 5 mM DTT, 0.05% Tween-P20). Afterwards oligo alone and protein-oligo complexes were electrophoresed on a 6% native polyacrylamide gel at 4 °C for 2 h. The membrane was analysed using a Typhoon scanner (GE Healthcare).

4.7. Limited proteolysis

STAT3 (20 μg), EBNA2-D9 (60 μg) and the complex of both were digested for 5, 15, 30 and 60 min at room temperature with 20 ng of protease (trypsin, elastase and chymotrypsin). The digest was terminated by adding Laemmli buffer to the samples and boiling at 95 °C or freezing and storage at -20 °C. Aliquots were electrophoresed on 12% SDS-PAGE and Coomassie stained. For N-terminal sequencing the fragments of interest were transferred onto a PVDF membrane and then submitted to the N-terminal sequencing facility at the IBS, Grenoble (J.-P. Andrieu). For mass spectrometry analysis, frozen aliquots containing all protein fragments plus peptidase were submitted to the mass spectrometry facility at the IBS, Grenoble (L. Signor).

4.8. Crystallisation trials

Purified STAT3 was concentrated to 5 mg/ml with an Amicon Ultra concentrator (Millipore) and CR8 synthetic peptide was added in 3 time molar ratio. The protein complex was then submitted to the high-throughput nanodrop crystallisation facility at EMBL, Grenoble. Over 70 crystallisation conditions resulted in positive hits. 12 crystals have been tested and 4 complete datasets have been collected at Proxima I (SOLEIL, Paris). Best resolution obtained was 6.5 Å (Space Group I4₁, a= b=174.0, c=79.4).

Screening kits used by the HTX lab:

Screen contents	Supplier
The Classics	Qiagen/Nextal
Crystal Screen Lite & PEG/Ion	Hampton Research
MembFac & Natrix	Hampton Research
QuickScreen & Grid screens Ammonium Sulfate, Sodium Malonate - Sodium Formate	Hampton Research - Home made
Grid screens PEG 6K, PEG/LiCl, MPD - Screen Mme	Hampton Research - Home made
Index Screen	Hampton Research

4.9. Cell based assays

HEK293 cells (ATCC-CLR 1573) were maintained in DMEM medium supplemented with 10% fetal bovine serum (FCS). Cells were harvested twice a week by trypsination and were split 1:10. DG75 cells (ATCC-CLR 2625) were maintained in RPMI medium supplemented with 10% FCS, 2% L-glutamine and 1% pen-strep. Cells were split 1:1 twice a week.

4.9.1. Generation of stable, monoclonal cells expressing the firefly luciferase gene under control of a STAT3 promoter

HEK293 cells were transduced using lentiviral particles (Signal Lenti STAT3 Reporter (luc) Kit: CLS-6028L) containing the firefly luciferase gene under control of a STAT3 promoter region. 2×10^5 cells were transduced with one vial of lentiviral particles (MOI=10). The next day the virus was removed and cells were washed twice. Five days after transduction 500 ng/ml of puromycin were added to the growth medium for selection. Two weeks after selection, cells were pipetted in 96 well plates in a limiting dilution series in order to generate monoclonal cells. Stably transfected HEK293-STAT3-Luc cells were co-transfected with increasing concentration of KG172 EBNA2 expression plasmid (12.5-100 ng), constant concentration of mutants (S448A and Δ CR8 at 100 ng) or SMRT fragment expression vector. Twenty-four h later cells were treated with 100 ng/ml of LIF for 6 h. Luciferase signals were detected using a Wallac luminometer.

4.9.2. CBF1 reporter assay

CBF1-luciferase reporter plasmid (5 μ g), β -galactosidase normalisation vector (3 μ g) and 1 μ g of target vector (pAG155 empty vector, pAG155 EBNA2 wild-type or point mutant expression vector) were transfected into 5×10^6 DG75 cells by electroporation at 0.22 kV and 950 μ F. Experiments were performed in triplicates. After transfection cells were grown in 5 ml growth medium for 2 days. Then cells were harvested by centrifugation at 300 g for 10 min and washed with 1 ml of PBS. Pellets were lysed in 100 μ l ice cool extraction buffer (50% (w/v) glycerine, 5% (w/v) Triton x-100, 10 mM EDTA, 125 mM Tris-HCl pH7.8 and 2 mM DTT) and the supernatant was cleared in a microcentrifuge at 15300 rpm for 15 min. Supernatant was stored at -80 °C before being assayed. Ten μ l of cell lysate were pipetted into white 96 well plates in duplicate for each assay. For the β -galactosidase assay 100 μ l of assay buffer (100 mM Na-P pH 8.0, 1% Galacton Plus, 0.1 mM $MgCl_2$) were added to each well and incubated for 15 min. Luciferase and β -galactosidase assays were performed in an automated luminometer that injects 50 μ l of either luciferase assay buffer (20 mM Tricin, 1.07 mM $MgCO_3$, 2.67 $MgSO_4$, 0.1 mM EDTA, 33.3 mM DTT, 270 μ M acetyl-CoA, 470 μ M Luciferin and 530 μ M ATP) or β -galactosidase enhancer buffer (0.2 M NaOH and 9% Emerald enhancer) to each well shortly before measuring the luciferase or β -galactosidase signal.

4.9.3. STAT3 reporter assay

HEK293 were harvested with Trypsin/EDTA (0.05%) (Invitrogen) and 2.5×10^5 cells were seeded into each well of a 6 well plate. The next day 1 μ g of STAT3 firefly luciferase reporter and constitutively active renilla luciferase plasmid mixtures (40:1, SABioscience) were co-transfected with increasing concentration of pAG155 EBNA2 expression plasmid (100 ng, 200 ng, 300 ng and 400 ng) or constant concentration of pAG155 EBNA2 point mutants via lipofectamine (Invitrogen) transfection protocol. Twenty h later, cells were stimulated with 100 ng/ml of leukemia inhibitory factor (LIF) for 6 h. Cells were harvested by scraping and centrifugation at 300 g for 10 min. Then cells were washed with 1 ml of PBS and pellets were lysed in 100 μ l ice cool 1x extraction buffer (dual luciferase reporter assay kit, Promega) and the supernatant cleared by microcentrifugation at 15300 rpm for 15 min. The firefly and renilla luciferase signal were detected using commercially available assay reagents (dual luciferase reporter assay kit, Promega) according to the manufacturer's protocol and signals were recorded with a luminometer.

References

- Aggarwal, BB, Sethi, G, Ahn, KS, Sandur, SK, Pandey, MK, Kunnumakkara, AB, Sung, B, Ichikawa, H. 2006. Targeting Signal-Transducer-and-Activator-of-Transcription-3 for Prevention and Therapy of Cancer. *Ann. N. Y. Acad. Sci.* 1091: 151–169.
- Ahmad, KF, Melnick, A, Lax, S, Bouchard, D, Liu, J, Kiang, C-L, Mayer, S, Takahashi, S, Licht, JD, Privé, GG. 2003. Mechanism of SMRT Corepressor Recruitment by the BCL6 BTB Domain. *Mol. Cell* 12: 1551–1564.
- Akira, S, Nishio, Y, Inoue, M, Wang, X-J, We, S, Matsusaka, T, Yoshida, K, Sudo, T, Naruto, M, Kishimoto, T. 1994. Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related transcription factor involved in the gp130-mediated signaling pathway. *Cell* 77: 63–71.
- Ansieau, S, Leutz, A. 2002. The conserved Mynd domain of BS69 binds cellular and oncoviral proteins through a common PXLXP motif. *J Biol Chem* 277: 4906–10.
- Aoki, Y, Feldman, GM, Tosato, G. 2003. Inhibition of STAT3 signaling induces apoptosis and decreases survivin expression in primary effusion lymphoma. *Blood* 101: 1535–1542.
- Battaglia, S, Maguire, O, Campbell, MJ. 2010. Transcription factor co-repressors in cancer biology: roles and targeting. *Int. J. Cancer* 126: 2511–2519.
- Becker, S, Corthals, GL, Aebersold, R, Groner, B, Müller, CW. 1998a. Expression of a tyrosine phosphorylated, DNA binding Stat3[beta] dimer in bacteria. *FEBS Lett.* 441: 141–147.
- Becker, S, Groner, B, Muller, CW. 1998b. Three-dimensional structure of the Stat3[beta] homodimer bound to DNA. *Nature* 394: 145–151.
- Berg, DJ, Davidson, N, Kühn, R, Müller, W, Menon, S, Holland, G, Thompson-Snipes, L, Leach, MW, Rennick, D. 1996. Enterocolitis and colon cancer in interleukin-10-deficient mice are associated with aberrant cytokine production and CD4(+) TH1-like responses. *J. Clin. Invest.* 98: 1010–1020.
- Bhattacharya, S, Ray, RM, Johnson, LR. 2005. STAT3-mediated transcription of Bcl-2, Mcl-1 and c-IAP2 prevents apoptosis in polyamine-depleted cells. *Biochem. J.* 392: 335–344.
- Binolfi, A, Theillet, F, Selenko, P. 2012. Bacterial in-cell NMR of human α -synuclein: a disordered monomer by nature? *Biochem Soc Trans* 40: 950–4.
- Boivin, S, Hart, DJ. 2011. Interaction of the Influenza A Virus Polymerase PB2 C-terminal Region with Importin α Isoforms Provides Insights into Host Adaptation and Polymerase Assembly. *J. Biol. Chem.* 286: 10439–10448.
- Bontems, F, Verger, A, Dewitte, F, Lens, Z, Baert, J-L, Ferreira, E, Launoit, Y de, Sizun, C, Guittet, E, Villeret, V, Monté, D. 2011. NMR structure of the human Mediator MED25 ACID domain. *J. Struct. Biol.* 174: 245–251.
- Borggreffe, T, Yue, X. 2011. Interactions between subunits of the Mediator complex with gene-specific transcription factors. *Semin. Cell Dev. Biol.* 22: 759–768.

- Bornkamm, GW, Delius, H, Zimmer, U, Hudewentz, J, Epstein, MA. 1980. Comparison of Epstein-Barr virus strains of different origin by analysis of the viral DNAs. *J. Virol.* 35: 603–618.
- Bowman, T, Roy, G, Turkson, J, Jove, R. 2000. STATs in oncogenesis. *Oncogene* 19: 2474–2488.
- Bromberg, JF, Wrzeszczynska, MH, Devgan, G, Zhao, Y, Pestell, RG, Albanese, C, Darnell Jr, JE. 1999. Stat3 as an Oncogene. *Cell* 98: 295–303.
- Calderwood, MA, Venkatesan, K, Xing, L, Chase, MR, Vazquez, A, Holthaus, AM, Ewence, AE, Li, N, Hirozane-Kishikawa, T, Hill, DE, Vidal, M, Kieff, E, Johannsen, E. 2007. Epstein-Barr virus and virus human protein interaction maps. *Proc Natl Acad Sci U S A* 104: 7606–11.
- Cancian, L, Bosshard, R, Lucchesi, W, Karstegl, CE, Farrell, PJ. 2011. C-terminal region of EBNA-2 determines the superior transforming ability of type 1 Epstein-Barr virus by enhanced gene regulation of LMP-1 and CXCR7. *PLoS Pathog* 7: e1002164.
- Casanova, J-L, Holland, SM, Notarangelo, LD. 2012. Inborn Errors of Human JAKs and STATs. *Immunity* 36: 515–528.
- Cavanagh, J, Fairbrother, W, Palmer III, A, Rance, M, Skelton, N. 2007. *Protein NMR Spectroscopy (Second Edition)*. Elsevier.
- Chen, H, Lee, JM, Zong, Y, Borowitz, M, Ng, MH, Ambinder, RF, Hayward, SD. 2001. Linkage between STAT Regulation and Epstein-Barr Virus Gene Expression in Tumors. *J. Virol.* 75: 2929–2937.
- Chung, Y-H, Cho, N, Garcia, MI, Lee, S-H, Feng, P, Jung, JU. 2004. Activation of Stat3 Transcription Factor by Herpesvirus Saimiri STP-A Oncoprotein. *J. Virol.* 78: 6489–6497.
- Cimica, V, Chen, HC, Iyer, JK, Reich, NC. 2011. Dynamics of the STAT3 transcription factor: nuclear import dependent on Ran and importin-beta1. *PLoS One* 6: e20188.
- Clercq, ED, Holy, A. 2005. Acyclic nucleoside phosphonates: a key class of antiviral drugs. *Nat Rev Drug Discov* 4: 928–940.
- Cohen, JI. 1992. A region of herpes simplex virus VP16 can substitute for a transforming domain of Epstein-Barr virus nuclear protein 2. *Proc. Natl. Acad. Sci.* 89: 8030–8034.
- Cohen, JI, Kieff, E. 1991. An Epstein-Barr virus nuclear protein 2 domain essential for transformation is a direct transcriptional activator. *J Virol* 65: 5880–5.
- Davey, N, Travé, G, Gibson, T. 2011. How viruses hijack cell regulation. *Trends Biochem Sci* 36: 159–69.
- Davey, NE, Van Roey, K, Weatheritt, RJ, Toedt, G, Uyar, B, Altenberg, B, Budd, A, Diella, F, Dinkel, H, Gibson, TJ. 2012. Attributes of short linear motifs. *Mol. Biosyst.* 8: 268–281.
- Dhordain, P, Quief, S, Lantoine, D, Kerckaert, J-P, Albagli, O, Lin, RJ, Evans, RM. 1998a. The LAZ3(BCL-6) oncoprotein recruits a SMRT/mSIN3A/histone deacetylase containing complex to mediate transcriptional repression. *Nucleic Acids Res.* 26: 4645–4651.

- Dhordain, P, Quief, S, Lantoine, D, Kerckaert, J-P, Albagli, O, Lin, RJ, Evans, RM. 1998b. The LAZ3(BCL-6) oncoprotein recruits a SMRT/mSIN3A/histone deacetylase containing complex to mediate transcriptional repression. *Nucleic Acids Res.* 26: 4645–4651.
- Dosztányi, Z, Csizmok, V, Tompa, P, Simon, I. 2005. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21: 3433–3434.
- Dosztányi, Z, Mészáros, B, Simon, I. 2010. Bioinformatical approaches to characterize intrinsically disordered/unstructured proteins. *Brief. Bioinform.* 11: 225–243.
- Du, T, Zhou, G, Roizman, B. 2013. Modulation of reactivation of latent herpes simplex virus 1 in ganglionic organ cultures by p300/CBP and STAT3. *Proc. Natl. Acad. Sci.* 110: E2621–E2628.
- Dunker, A, Obradovic, Z, Romero, P, Garner, E, Brown, C. 2000. Intrinsic protein disorder in complete genomes. *Genome Inf. Ser Work. Genome Inf.* 11: 161–71.
- Dunker, AK, Silman, I, Uversky, VN, Sussman, JL. 2008. Function and structure of inherently disordered proteins. *Catal. Regul. Proteins* 18: 756–764.
- Dyson, HJ, Wright, PE. 2005a. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6: 197–208.
- Dyson, HJ, Wright, PE. 2005b. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6: 197–208.
- Epstein, MA, Henle, G, Achong, BG, Barr, YM. 1965. Morphological and Biological Studies on a Virus in Cultured Lymphoblasts from Burkitt's Lymphoma. *J Exp Med* 121: 761–70.
- Fields, BN, Knipe, DM, Howley, PM. 1996. *Fields virology*. Philadelphia: Lippincott-Raven Publishers.
- Foulkes, WD, Shuen, AY. 2013. In Brief: BRCA1 and BRCA2. *J. Pathol.* 230: 347–349.
- Galea, CA, Wang, Y, Sivakolundu, SG, Kriwacki, RW. 2008. Regulation of Cell Division by Intrinsically Unstructured Proteins: Intrinsic Flexibility, Modularity, and Signaling Conduits†. *Biochemistry (Mosc.)* 47: 7598–7609.
- Gao, L, Zhang, L, Hu, J, Li, F, Shao, Y, Zhao, D, Kalvakolanu, DV, Kopecko, DJ, Zhao, X, Xu, D-Q. 2005. Down-Regulation of Signal Transducer and Activator of Transcription 3 Expression Using Vector-Based Small Interfering RNAs Suppresses Growth of Human Prostate Tumor In vivo. *Clin. Cancer Res.* 11: 6333–6341.
- Gerosa, F, Baldani-Guerra, B, Lyakh, LA, Batoni, G, Esin, S, Winkler-Pickett, RT, Consolaro, MR, De Marchi, M, Giachino, D, Robbiano, A, Astegiano, M, Sambataro, A, Kastelein, RA, Carra, G, Trinchieri, G. 2008. Differential regulation of interleukin 12 and interleukin 23 production in human dendritic cells. *J. Exp. Med.* 205: 1447–1461.
- Gough, DJ, Corlett, A, Schlessinger, K, Wegrzyn, J, Lerner, AC, Levy, DE. 2009. Mitochondrial STAT3 Supports Ras-Dependent Oncogenic Transformation. *Science* 324: 1713–1716.

- Grabusic, K, Maier, S, Hartmann, A, Mantik, A, Hammerschmidt, W, Kempkes, B. 2006. The CR4 region of EBNA2 confers viability of Epstein–Barr virus-transformed B cells by CBF1-independent signalling. *J. Gen. Virol.* 87: 3169–3176.
- Gsponer, J, Futschik, ME, Teichmann, SA, Babu, MM. 2008. Tight Regulation of Unstructured Proteins: From Transcript Synthesis to Protein Degradation. *Science* 322: 1365–1368.
- Hao, J, Li, T-G, Qi, X, Zhao, D-F, Zhao, G-Q. 2006. WNT/ β -catenin pathway up-regulates Stat3 and converges on LIF to prevent differentiation of mouse embryonic stem cells. *Dev. Biol.* 290: 81–91.
- Harada, S, Yalamanchili, R, Kieff, E. 2001. Epstein-Barr Virus Nuclear Protein 2 Has at Least Two N-Terminal Domains That Mediate Self-Association.
- Hayward, SD. 2004. Viral interactions with the Notch pathway. *Notch Signal. Cancer* 14: 387–396.
- Henkel, T, Ling, P, Hayward, S, Peterson, M. 1994. Mediation of Epstein-Barr virus EBNA2 transactivation by recombination signal-binding protein J kappa. *Science*: 92–95.
- Ho, Y, Tsao, S-W, Zeng, M, Lui, VWY. 2013. STAT3 as a therapeutic target for Epstein-Barr virus (EBV) – associated nasopharyngeal carcinoma. *Cancer Lett.* 330: 141–149.
- Horvath, CM, Wen, Z, Darnell, JE. 1995. A STAT protein domain that determines DNA sequence recognition suggests a novel DNA-binding domain. *Genes Dev.* 9: 984–994.
- Huang, Y, Liu, Z. 2009. Kinetic Advantage of Intrinsically Disordered Proteins in Coupled Folding–Binding Process: A Critical Assessment of the “Fly-Casting” Mechanism. *J. Mol. Biol.* 393: 1143–1159.
- Hutsell, SQ, Kimple, RJ, Siderovski, DP, Willard, FS, Kimple, AJ. 2010. High-affinity immobilization of proteins using biotin- and GST-based coupling strategies. *Methods Mol Biol* 627: 75–90.
- Ikedo, O, Togi, S, Kamitani, S, Muromoto, R, Sekine, Y, Nanbo, A, Fujimuro, M, Matsuda, T. 2009. Silencing mediator of retinoic acid and thyroid hormone receptor regulates enhanced activation of signal transducer and activator of transcription 3 by epstein-barr virus-derived epstein-barr nuclear antigen 2. *Biol Pharm Bull* 32: 1283–5.
- Jang, HD, Yoon, K, Shin, YJ, Kim, J, Lee, SY. 2004. PIAS3 Suppresses NF- κ B-mediated Transcription by Interacting with the p65/RelA Subunit. *J. Biol. Chem.* 279: 24873–24880.
- Jeganathan, S, von Bergen, M, Mandelkow, E-M, Mandelkow, E. 2008. The Natively Unfolded Character of Tau and Its Aggregation to Alzheimer-like Paired Helical Filaments†. *Biochemistry (Mosc.)* 47: 10526–10539.
- Jepsen, K, Rosenfeld, MG. 2002a. Biological roles and mechanistic actions of co-repressor complexes. *J Cell Sci* 115: 689–698.
- Jepsen, K, Rosenfeld, MG. 2002b. Biological roles and mechanistic actions of co-repressor complexes. *J Cell Sci* 115: 689–698.
- Jerabek-Willemsen, M, Wienken, C, Braun, D, Duhr, S. 2011. Molecular interaction studies using microscale thermophoresis. *Assay Drug Dev Technol* 9: 342–53.

- Jonker, HRA, Wechselberger, RW, Boelens, R, Folkers, GE, Kaptein, R. 2004. Structural Properties of the Promiscuous VP16 Activation Domain†. *Biochemistry (Mosc.)* 44: 827–839.
- Jordheim, LP, Durantel, D, Zoulim, F, Dumontet, C. 2013. Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. *Nat Rev Drug Discov* 12: 447–464.
- Jung, JE, Lee, H-G, Cho, I-H, Chung, DH, Yoon, S-H, Yang, YM, Lee, JW, Choi, S, Park, J-W, Ye, S-K, Chung, M-H. 2005. STAT3 is a potential modulator of HIF-1-mediated VEGF expression in human renal carcinoma cells. *FASEB J*.
- Kempkes, B, Spitkovsky, D, Jansen-Durr, P, Ellwart, JW, Kremmer, E, Delecluse, HJ, Rottenberger, C, Bornkamm, GW, Hammerschmidt, W. 1995. B-cell proliferation and induction of early G1-regulating proteins by Epstein-Barr virus mutants conditional for EBNA2. *EMBO J* 14: 88–96.
- King, CA. 2013. Kaposi's Sarcoma-Associated Herpesvirus Kaposin B Induces Unique Monophosphorylation of STAT3 at Serine 727 and MK2-Mediated Inactivation of the STAT3 Transcriptional Repressor TRIM28. *J. Virol.* 87: 8779–8791.
- Koganti, S, de la Paz, A, Freeman, AF, Bhaduri-McIntosh, S. 2013. B lymphocytes from patients with hypomorphic mutation in STAT3 resist Epstein-Barr virus-driven cell proliferation. *J. Virol.*
- Kortylewski, M, Kujawski, M, Wang, T, Wei, S, Zhang, S, Pilon-Thomas, S, Niu, G, Kay, H, Mule, J, Kerr, WG, Jove, R, Pardoll, D, Yu, H. 2005a. Inhibiting Stat3 signaling in the hematopoietic system elicits multicomponent antitumor immunity. *Nat Med* 11: 1314–21.
- Kortylewski, M, Kujawski, M, Wang, T, Wei, S, Zhang, S, Pilon-Thomas, S, Niu, G, Kay, H, Mule, J, Kerr, WG, Jove, R, Pardoll, D, Yu, H. 2005b. Inhibiting Stat3 signaling in the hematopoietic system elicits multicomponent antitumor immunity. *Nat Med* 11: 1314–21.
- Kortylewski, M, Xin, H, Kujawski, M, Lee, H, Liu, Y, Harris, T, Drake, C, Pardoll, D, Yu, H. 2009. Regulation of the IL-23 and IL-12 balance by Stat3 signaling in the tumor microenvironment. *Cancer Cell* 15: 114–23.
- Kouzarides, T. 2007. Chromatin Modifications and Their Function. *Cell* 128: 693–705.
- Kriwacki, RW, Hengst, L, Tennant, L, Reed, SI, Wright, PE. 1996. Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proc. Natl. Acad. Sci.* 93: 11504–11509.
- Kuppers, R. 2003. B cells under influence: transformation of B cells by Epstein-Barr virus. *Nat Rev Immunol* 3: 801–812.
- Kwiatkowski, B, Chen, SYJ, Schubach, WH. 2004. CKII Site in Epstein-Barr Virus Nuclear Protein 2 Controls Binding to hSNF5/Ini1 and Is Important for Growth Transformation. *J. Virol.* 78: 6067–6072.
- Lawee, D. 2007. Mild infectious mononucleosis presenting with transient mixed liver disease: Case report with a literature review. *Can. Fam. Physician* 53: 1314–1316.

- Lee, L, Stollar, E, Chang, J, Grossmann, JG, O'Brien, R, Ladbury, J, Carpenter, B, Roberts, S, Luisi, B. 2001. Expression of the Oct-1 Transcription Factor and Characterization of Its Interactions with the Bob1 Coactivator†. *Biochemistry (Mosc.)* 40: 6580–6588.
- Lee, S-K, Kim, J-H, Lee, YC, Cheong, J, Lee, JW. 2000. Silencing Mediator of Retinoic Acid and Thyroid Hormone Receptors, as a Novel Transcriptional Corepressor Molecule of Activating Protein-1, Nuclear Factor- κ B, and Serum Response Factor. *J. Biol. Chem.* 275: 12470–12474.
- Letwin, K, Yee, S, Pawson, T. 1988. Novel protein-tyrosine kinase cDNAs related to *fps/fes* and *epf* cloned using anti-phosphotyrosine antibody. *Oncogene*.
- Li, C, Liu, M. 2013. Protein dynamics in living cells studied by in-cell NMR spectroscopy. *Many Faces Proteins* 587: 1008–1011.
- Li, J, Wang, J, Wang, J, Nawaz, Z, Liu, JM, Qin, J, Wong, J. 2000a. Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. *EMBO J* 19: 4342–4350.
- Li, J, Wang, J, Wang, J, Nawaz, Z, Liu, JM, Qin, J, Wong, J. 2000b. Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. *EMBO J* 19: 4342–4350.
- Lieutaud, P, Canard, B, Longhi, S. 2008. MeDor: a metaserver for predicting protein disorder. *BMC Genomics* 9.
- Ling, PD, Hayward, SD. 1995. Contribution of conserved amino acids in mediating the interaction between EBNA2 and CBF1/RBPJk. *J. Virol.* 69: 1944–1950.
- Liu, L, McBride, KM, Reich, NC. 2005. STAT3 nuclear import is independent of tyrosine phosphorylation and mediated by importin- α 3. *Proc. Natl. Acad. Sci. U. S. A.* 102: 8150–8155.
- Lucchesi, W, Brady, G, Dittrich-Breiholz, O, Kracht, M, Russ, R, Farrell, PJ. 2008. Differential Gene Regulation by Epstein-Barr Virus Type 1 and Type 2 EBNA2. *J. Virol.* 82: 7456–7466.
- Lui, VWY, Wong, EYL, Ho, Y, Hong, B, Wong, SCC, Tao, Q, Choi, GCG, Au, TCC, Ho, K, Yau, DMS, Ma, BBY, Hui, EP, Chan, AS-K, Tsang, C-M, Tsao, S-W, Grandis, JR, Chan, AT-C. 2009. STAT3 activation contributes directly to Epstein-Barr virus-mediated invasiveness of nasopharyngeal cancer cells in vitro. *Int. J. Cancer* 125: 1884–1893.
- Maier, S, Santak, M, Mantik, A, Grabusic, K, Kremmer, E, Hammerschmidt, W, Kempkes, B. 2005. A Somatic Knockout of CBF1 in a Human B-Cell Line Reveals that Induction of CD21 and CCR7 by EBNA-2 Is Strictly CBF1 Dependent and that Downregulation of Immunoglobulin M Is Partially CBF1 Independent. *J. Virol.* 79: 8784–8792.
- Masselink, H, Bernards, R. 2000. The adenovirus E1A binding protein BS69 is a corepressor of transcription through recruitment of N-CoR.
- Matrai, J, Chuah, MK, VandenDriessche, T. 2010. Recent Advances in Lentiviral Vector Development and Applications. *Mol Ther* 18: 477–490.
- Mettenleiter, TC. 2004. Budding events in herpesvirus morphogenesis. *Mech. Envel. Virus Release* 106: 167–180.

- Mettenleiter, TC, Klupp, BG, Granzow, H. 2009. Herpesvirus assembly: An update. *Virus Res.* - 25th Anniv. Issue 143: 222–234.
- Michelow, P, Wright, C, Pantanowitz, L. 2012. A Review of the Cytomorphology of Epstein-Barr Virus-Associated Malignancies. *Acta Cytol.* 56: 1–14.
- Milbradt, AG, Kulkarni, M, Yi, T, Takeuchi, K, Sun, Z-YJ, Luna, RE, Selenko, P, Näär, AM, Wagner, G. 2011. Structure of the VP16 transactivator target in the Mediator. *Nat Struct Mol Biol* 18: 410–415.
- Minegishi, Y, Saito, M, Tsuchiya, S, Tsuge, I, Takada, H, Hara, T, Kawamura, N, Ariga, T, Pasic, S, Stojkovic, O, Metin, A, Karasuyama, H. 2007. Dominant-negative mutations in the DNA-binding domain of STAT3 cause hyper-IgE syndrome. *Nature* 448: 1058–1062.
- Muller, PAJ, Vousden, KH. 2013. p53 mutations in cancer. *Nat Cell Biol* 15: 2–8.
- Muromoto, R, Ikeda, O, Okabe, K, Togi, S, Kamitani, S, Fujimuro, M, Harada, S, Oritani, K, Matsuda, T. 2009a. Epstein-Barr virus-derived EBNA2 regulates STAT3 activation. *Biochem. Biophys. Res. Commun.* 378: 439–443.
- Muromoto, R, Ikeda, O, Okabe, K, Togi, S, Kamitani, S, Fujimuro, M, Harada, S, Oritani, K, Matsuda, T. 2009b. Epstein-Barr virus-derived EBNA2 regulates STAT3 activation. *Biochem. Biophys. Res. Commun.* 378: 439–443.
- Nakajima, H, Brindle, PK, Handa, M, Ihle, JN. 2001a. Functional interaction of STAT5 and nuclear receptor co-repressor SMRT: implications in negative regulation of STAT5-dependent transcription. *EMBO J* 20: 6836–6844.
- Nakajima, H, Brindle, PK, Handa, M, Ihle, JN. 2001b. Functional interaction of STAT5 and nuclear receptor co-repressor SMRT: implications in negative regulation of STAT5-dependent transcription. *EMBO J* 20: 6836–44.
- Naruya Tomita, Toshio Ogihara, Ryuichi Morishita. 2003. Transcription Factors as Molecular Targets: Molecular Mechanisms of Decoy ODN and their Design. *Curr. Drug Targets* 4: 603–608.
- Nemerow, GR, Mold, C, Schwend, VK, Tollefson, V, Cooper, NR. 1987. Identification of gp350 as the viral glycoprotein mediating attachment of Epstein-Barr virus (EBV) to the EBV/C3d receptor of B cells: sequence homology of gp350 and C3 complement fragment C3d. *J. Virol.* 61: 1416–1420.
- Parkin, DM. 2006. The global health burden of infection-associated cancers in the year 2002. *Int. J. Cancer* 118: 3030–3044.
- Peibin Yue, James Turkson. 2008. Targeting STAT3 in cancer: how successful are we? *Expert Opin. Investig. Drugs* 18: 45–56.
- Perissi, V, Jepsen, K, Glass, CK, Rosenfeld, MG. 2010. Deconstructing repression: evolving models of co-repressor action. *Nat Rev Genet* 11: 109–123.
- Peyser, N, Grandis, J. 2013. Critical analysis of the potential for targeting STAT3 in human malignancy. *OncoTargets Ther.* 2013:6: 999–1010.
- Pushker, R, Mooney, C, Davey, N, Jacque, J, Shields, D. 2013. Marked variability in the extent of protein disorder within and between viral families. *PLoS ONE* 8.

- Rahaman, SO, Harbor, PC, Chernova, O, Barnett, GH, Vogelbaum, MA, Haque, SJ. 2002. Inhibition of constitutively active Stat3 suppresses proliferation and induces apoptosis in glioblastoma multiforme cells. *Oncogene* 21: 8404–13.
- Reitsma, JM, Sato, H, Nevels, M, Terhune, SS, Paulus, C. 2013. Human Cytomegalovirus IE1 Protein Disrupts Interleukin-6 Signaling by Sequestering STAT3 in the Nucleus. *J. Virol.* 87: 10763–10776.
- Ren, Z, Mao, X, Mertens, C, Krishnaraj, R, Qin, J, Mandal, PK, Romanowski, MJ, McMurray, JS, Chen, X. 2008. Crystal structure of unphosphorylated STAT3 core fragment. *Biochem. Biophys. Res. Commun.* 374: 1–5.
- Schust, J, Sperl, B, Hollis, A, Mayer, A, Berg, T. 2006. Stattic: A Small-Molecule Inhibitor of STAT3 Activation and Dimerization. *Chem. Biol.:* 1235–1242.
- Selenko, P, Wagner, G. 2007. Looking into live cells with in-cell NMR spectroscopy. *Struct. Anal. Supramol. Assem. Hybrid Methods* 158: 244–253.
- Serber, Z, Selenko, P, Hansel, R, Reckel, S, Lohr, F, Ferrell, JE, Wagner, G, Dotsch, V. 2007. Investigating macromolecules inside cultured and injected cells by in-cell NMR spectroscopy. *Nat Protoc.* 1: 2701–2709.
- Sevcik, J, Skrabana, R, Dvorsky, R, Csokova, N, Iqbal, K, Novak, M. 2007a. X-ray structure of the PHF core C-terminus: Insight into the folding of the intrinsically disordered protein tau in Alzheimer's disease. *FEBS Lett.* 581: 5872–5878.
- Sevcik, J, Skrabana, R, Dvorsky, R, Csokova, N, Iqbal, K, Novak, M. 2007b. X-ray structure of the PHF core C-terminus: Insight into the folding of the intrinsically disordered protein tau in Alzheimer's disease. *FEBS Lett.* 581: 5872–5878.
- Shao, H, Xu, X, Mastrangelo, MA, Jing, N, Cook, RG, Legge, GB, Tweardy, DJ. 2004. Structural requirements for signal transducer and activator of transcription 3 binding to phosphotyrosine ligands containing the YXXQ motif. *J Biol Chem* 279: 18967–73.
- Shen, Y, Devgan, G, Darnell, JE, Bromberg, JF. 2001. Constitutively activated Stat3 protects fibroblasts from serum withdrawal and UV-induced apoptosis and antagonizes the proapoptotic effects of activated Stat1. *Proc. Natl. Acad. Sci.* 98: 1543–1548.
- Shirogane, T, Fukada, T, Muller, JMM, Shima, DT, Hibi, M, Hirano, T. 1999. Synergistic Roles for Pim-1 and c-Myc in STAT3-Mediated Cell Cycle Progression and Antiapoptosis. *Immunity* 11: 709–719.
- Shoemaker, BA, Portman, JJ, Wolynes, PG. 2000a. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc. Natl. Acad. Sci.* 97: 8868–8873.
- Shoemaker, BA, Portman, JJ, Wolynes, PG. 2000b. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc. Natl. Acad. Sci.* 97: 8868–8873.
- Shuai, K, Horvath, CM, Huang, LHT, Qureshi, SA, Cowburn, D, Darnell Jr., JE. 1994. Interferon activation of the transcription factor Stat91 involves dimerization through SH2-phosphotyrosyl peptide interactions. *Cell* 76: 821–828.
- Snyder, M, Huang, X-Y, Zhang, JJ. 2008. Identification of Novel Direct Stat3 Target Genes for Control of Growth and Differentiation. *J. Biol. Chem.* 283: 3791–3798.

- Sugase, K, Dyson, HJ, Wright, PE. 2007a. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447: 1021–5.
- Sugase, K, Dyson, HJ, Wright, PE. 2007b. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447: 1021–5.
- Takeda, K, Noguchi, K, Shi, W, Tanaka, T, Matsumoto, M, Yoshida, N, Kishimoto, T, Akira, S. 1997. Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. *Proc. Natl. Acad. Sci.* 94: 3801–3804.
- Tochio, H. 2012. Watching protein structure at work in living cells using NMR spectroscopy. *Mech. Aesthetics• Mol. Imaging* 16: 609–613.
- Tompa, P. 2012. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* 37: 509–516.
- Tompa, P. 2005. The interplay between structure and function in intrinsically unstructured proteins. *Bp. Spec. Issue Proteins Pept.* 579: 3346–3354.
- Tong, X, Wang, F, Thut, C, Kieff, E. 1995. The Epstein-Barr Virus Nuclear Protein 2 Acidic Domain Can Interact with TFIIB, TAF40, and RPA70 but Not with TATA-Binding Protein. *J VIROL* 69: 585–8.
- Turkson, J, Kim, JS, Zhang, S, Yuan, J, Huang, M, Glenn, M, Haura, E, Sebt, S, Hamilton, AD, Jove, R. 2004. Novel peptidomimetic inhibitors of signal transducer and activator of transcription 3 dimerization and biological activity. *Mol. Cancer Ther.* 3: 261–269.
- Turkson, J, Ryan, D, Kim, JS, Zhang, Y, Chen, Z, Haura, E, Laudano, A, Sebt, S, Hamilton, AD, Jove, R. 2001. Phosphotyrosyl Peptides Block Stat3-mediated DNA Binding Activity, Gene Regulation, and Cell Transformation. *J. Biol. Chem.* 276: 45443–45455.
- Uesugi, M, Nyanguile, O, Lu, H, Levine, AJ, Verdine, GL. 1997. Induced α Helix in the VP16 Activation Domain upon Binding to a Human TAF. *Science* 277: 1310–1313.
- Uversky, VN, Oldfield, CJ, Dunker, AK. 2008. Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept. *Annu. Rev. Biophys.* 37: 215–246.
- Vallania, F, Schiavone, D, Dewilde, S, Pupo, E, Garbay, S, Calogero, R, Pontoglio, M, Provero, P, Poli, V. 2009. Genome-wide discovery of functional transcription factor binding sites by comparative genomics: The case of Stat3. *Proc. Natl. Acad. Sci.* 106: 5117–5122.
- Vinkemeier, U, Cohen, S, Moarefi, I, Chait, B, Kuriyan, J, Darnell, JJ. 1996. DNA binding of in vitro activated Stat1 alpha, Stat1 beta and truncated Stat1: interaction between NH2-terminal domains stabilizes binding of two dimers to tandem DNA sites. *EMBO J:* 5616–26.
- Vinkemeier, U, Moarefi, I, Darnell, JE, Kuriyan, J. 1998a. Structure of the Amino-Terminal Protein Interaction Domain of STAT-4. *Science* 279: 1048–1052.
- Vojnic, E, Mourão, A, Seizl, M, Simon, B, Wenzek, L, Larivière, L, Baumli, S, Baumgart, K, Meisterernst, M, Sattler, M, Cramer, P. 2011. Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nat Struct Mol Biol* 18: 404–409.
- Wang, LH, Yang, XY, Mihalic, K, Xiao, W, Li, D, Farrar, WL. 2001. Activation of Estrogen Receptor Blocks Interleukin-6-inducible Cell Growth of Human Multiple Myeloma

- Involving Molecular Cross-talk between Estrogen Receptor and STAT3 Mediated by Co-regulator PIAS3. *J. Biol. Chem.* 276: 31839–31844.
- Wang, LH, Yang, XY, Zhang, X, Huang, J, Hou, J, Li, J, Xiong, H, Mihalic, K, Zhu, H, Xiao, W, Farrar, WL. 2004. Transcriptional Inactivation of STAT3 by PPAR γ Suppresses IL-6-Responsive Multiple Myeloma Cells. *Immunity* 20: 205–218.
- Wen, Z, Zhong, Z, Darnell Jr., JE. 1995. Maximal activation of transcription by stat1 and stat3 requires both tyrosine and serine phosphorylation. *Cell* 82: 241–250.
- WHO | Viral Cancers.
- Wienken, CJ, Baaske, P, Rothbauer, U, Braun, D, Duhr, S. 2010. Protein-binding assays in biological liquids using microscale thermophoresis. *Nat Commun* 1: 100.
- Wong, C-W, Privalsky, ML. 1998. Transcriptional Repression by the SMRT-mSin3 Corepressor: Multiple Interactions, Multiple Mechanisms, and a Potential Role for TFIIB. *Mol. Cell. Biol.* 18: 5500–5510.
- Xie, H, Vucetic, S, Iakoucheva, LM, Oldfield, CJ, Dunker, AK, Uversky, VN, Obradovic, Z. 2007. Functional Anthology of Intrinsic Disorder. 1. Biological Processes and Functions of Proteins with Long Disordered Regions. *J. Proteome Res.* 6: 1882–1898.
- Xu, L, Lavinsky, RM, Dasen, JS, Flynn, SE, McInerney, EM, Mullen, T-M, Heinzl, T, Szeto, D, Korzus, E, Kurokawa, R, Aggarwal, AK, Rose, DW, Glass, CK, Rosenfeld, MG. 1998. Signal-specific co-activator domain requirements for Pit-1 activation. *Nature* 395: 301–306.
- Xu, X, Kasembeli, M, Tweardy, B, Tweardy, D. 2009. Chemical Probes that Competitively and Selectively Inhibit Stat3 Activation. *PLoS One* 4 (3).
- Xu, X, Sun, Y-L, Hoey, T. 1996. Cooperative DNA Binding and Sequence-Selective Recognition Conferred by the STAT Amino-Terminal Domain. *Science* 273: 794–797.
- Xue, B, Brown, CJ, Dunker, AK, Uversky, VN. 2013. Intrinsically disordered regions of p53 family are highly diversified in evolution. *Biochim. Biophys. Acta BBA - Proteins Proteomics* 1834: 725–738.
- Yang, E, Lerner, L, Besser, D, Darnell, JE. 2003. Independent and Cooperative Activation of Chromosomal c-fos Promoter by STAT3. *J. Biol. Chem.* 278: 15794–15799.
- Yang, J, Stark, GR. 2008. Roles of unphosphorylated STATs in signaling. *Cell Res* 18: 443–451.
- Yang, ZR, Thomson, R, McNeil, P, Esnouf, RM. 2005. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics* 21: 3369–3376.
- Yoo, J-Y, Wang, W, Desiderio, S, Nathans, D. 2001. Synergistic Activity of STAT3 and c-Jun at a Specific Array of DNA Elements in the α 2-Macroglobulin Promoter. *J. Biol. Chem.* 276: 26421–26429.
- Yoon, H-G, Chan, DW, Huang, Z-Q, Fondell, JD, Qin, J, Wong, J. 2003. Purification and functional characterization of the human N-CoR complex: the roles of HDAC3, TBL1 and TBLR1. *EMBO J*: 1336–1346.

- Young, LS, Rickinson, AB. 2004. Epstein-Barr virus: 40 years on. *Nat Rev Cancer* 4: 757–768.
- Yu, H, Kortylewski, M, Pardoll, D. 2007. Crosstalk between cancer and immune cells: role of STAT3 in the tumour microenvironment. *Nat Rev Immunol* 7: 41–51.
- Yu, Z, Zhang, W, Kone, BC. 2002. Signal transducers and activators of transcription 3 (STAT3) inhibits transcription of the inducible nitric oxide synthase gene by interacting with nuclear factor kappaB. *Biochem. J.* 367: 97–105.
- Yue, P, Turkson, J. 2008. Targeting STAT3 in cancer: how successful are we? *Expert Opin. Investig. Drugs* 18: 45–56.
- Yumerefendi, H, Desravines, DC, Hart, DJ. 2011. Library-based methods for identification of soluble expression constructs. *Methods San Diego Calif* 55: 38–43.
- Yumerefendi, H, Tarendeau, F, Mas, PJ, Hart, DJ. 2010. ESPRIT: An automated, library-based method for mapping and soluble expression of protein domains from challenging targets. *J. Struct. Biol.* 172: 66–74.
- Zhang, Z, Jones, S, Hagood, JS, Fuentes, NL, Fuller, GM. 1997. STAT3 Acts as a Co-activator of Glucocorticoid Receptor Signaling. *J. Biol. Chem.* 272: 30607–30610.
- Zhong, Z, Wen, Z, Darnell, JE. 1994. Stat3: a STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and interleukin-6. *Science* 264: 95–8.
- Zimber-Strobl, U, Strobl, LJ. 2001. EBNA2 and Notch signalling in Epstein–Barr virus mediated immortalization of B lymphocytes. *Semin. Cancer Biol.* 11: 423–434.
- Zushi, S, Shinomura, Y, Kiyohara, T, Miyazaki, Y, Kondo, S, Sugimachi, M, Higashimoto, Y, Kanayama, S, Matsuzawa, Y. 1998. STAT3 mediates the survival signal in oncogenic ras-transfected intestinal epithelial cells. *Int. J. Cancer* 78: 326–330.

Appendix

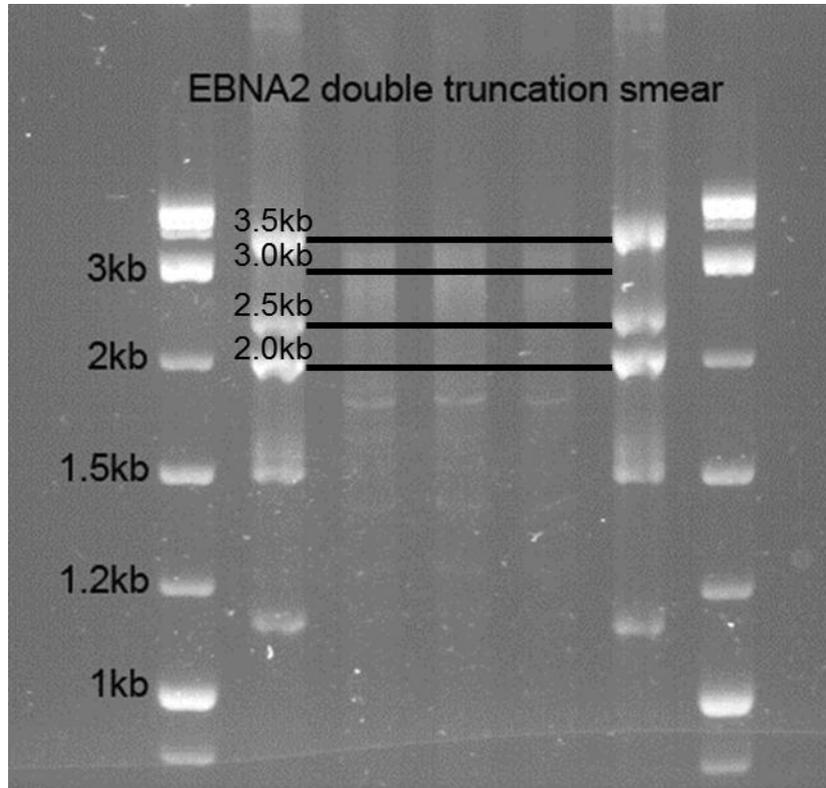


Figure A.1: pESPRIT002-EBNA2 double truncated DNA smear. Fractions of 3.5kb to 3kb, 3kb to 2.5kb and 2.5kb to 2 kb were isolated from a 0.5% w/v agarose gel.

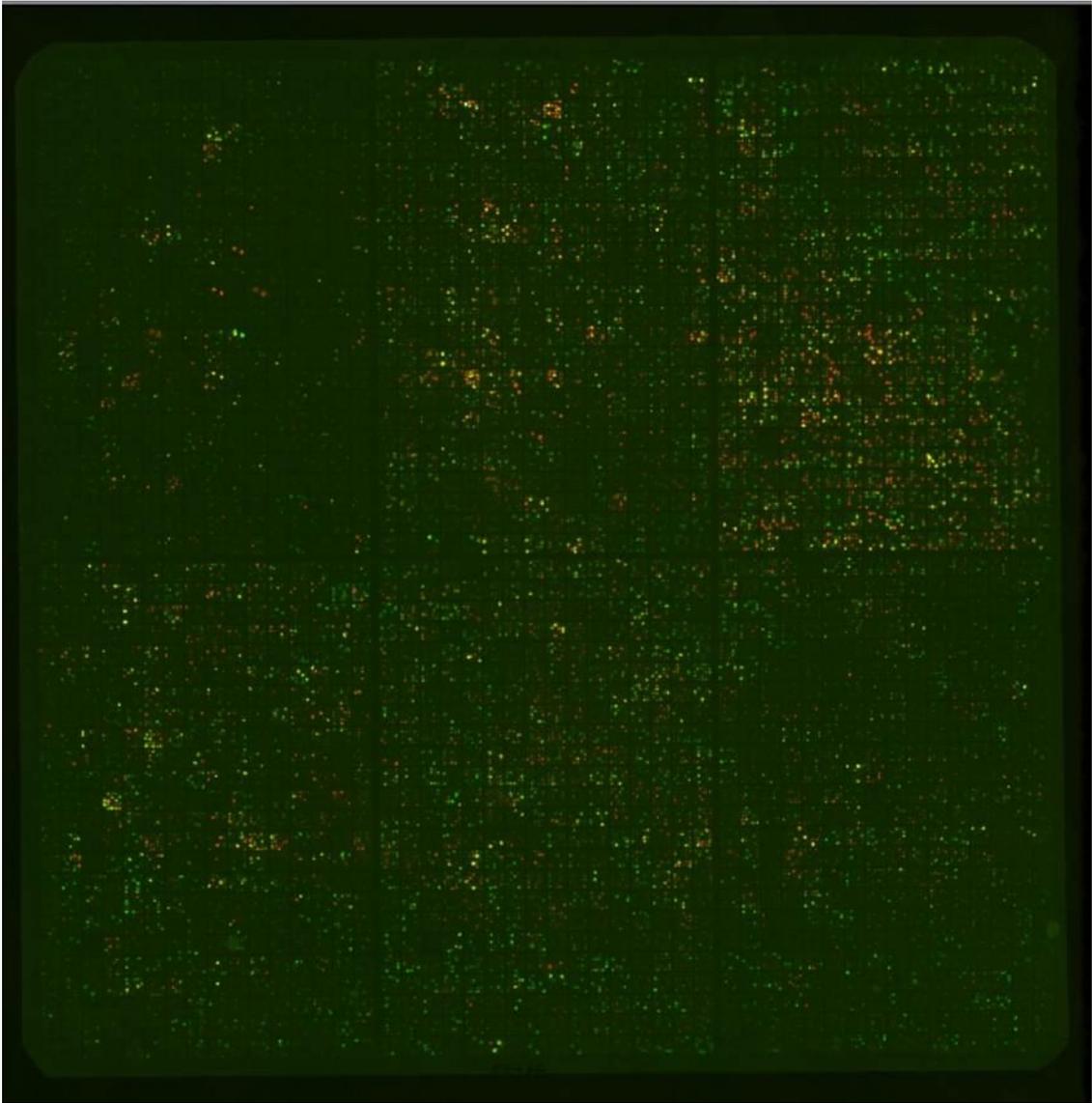


Figure A.2: Colony blot of EBNA2 expression library clones. 27648 clones were printed on a nitrocellulose membrane and signals for the 6xHis tag (red signal) and BAP(green signal) were detected with Streptavidin- Alexa488 and His tag primary antibody together with secondary antibody anti-mouse Alexa532



Figure A.3: Western blot of the 96 best expressing EBNA2 expression clones. NiNTA-purified proteins were detected via the BAP using Streptavidin- Alexa488.

Green: EBNA2-D9 ^{15}N : STAT3b unlabeled, 1:1, both proteins $\sim 71 \mu\text{M}$
Black: EBNA2-D9 free, $\sim 460 \mu\text{M}$
SOFAST HSQC
800 MHz, 25C, pH 6.5
9.12.11

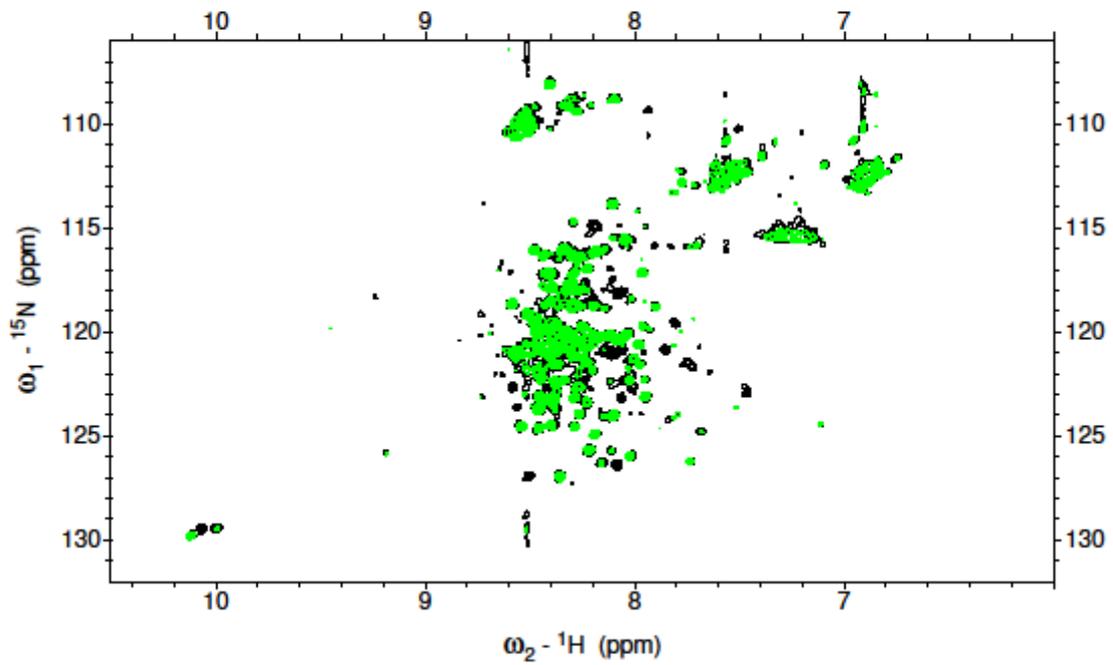
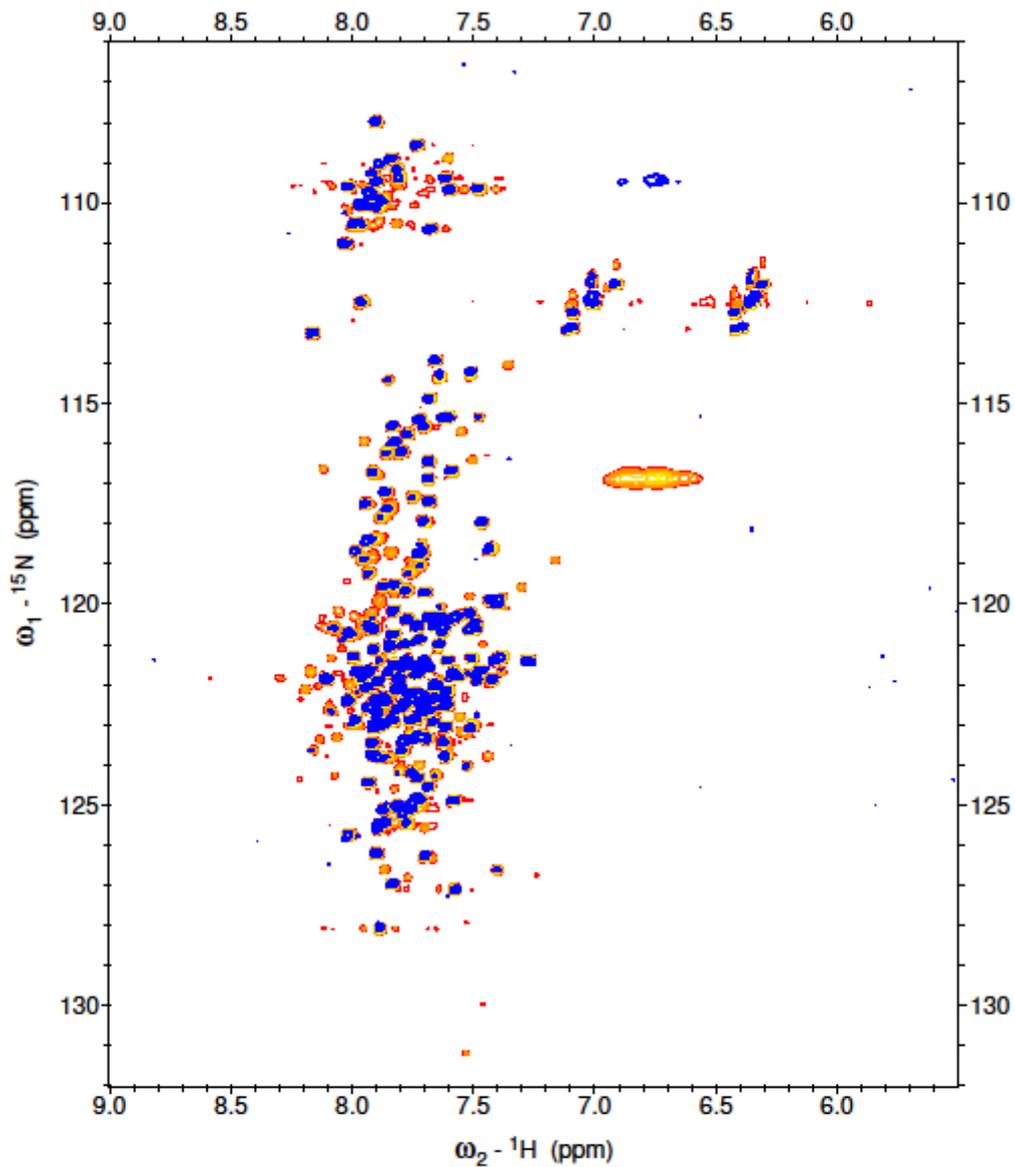


Figure A.4: HSQC spectrum of EBNA2-D9 in complex with STAT3. D9 alone (black) and D9-STAT3 complex (green). The ratio between D9 and STAT3 was 1:1. Spectra recorded by R.Schneider (Blackledge lab, IBS Grenoble).

39L23 ^{15}N , boiled, with biotin tag
15N HSQC
pH 6.5, 25C, 600 MHz

red: free, 925 μM (23.09.2010)

blue: with 1:1 STAT3b, both proteins 54 μM (05.11.2010)



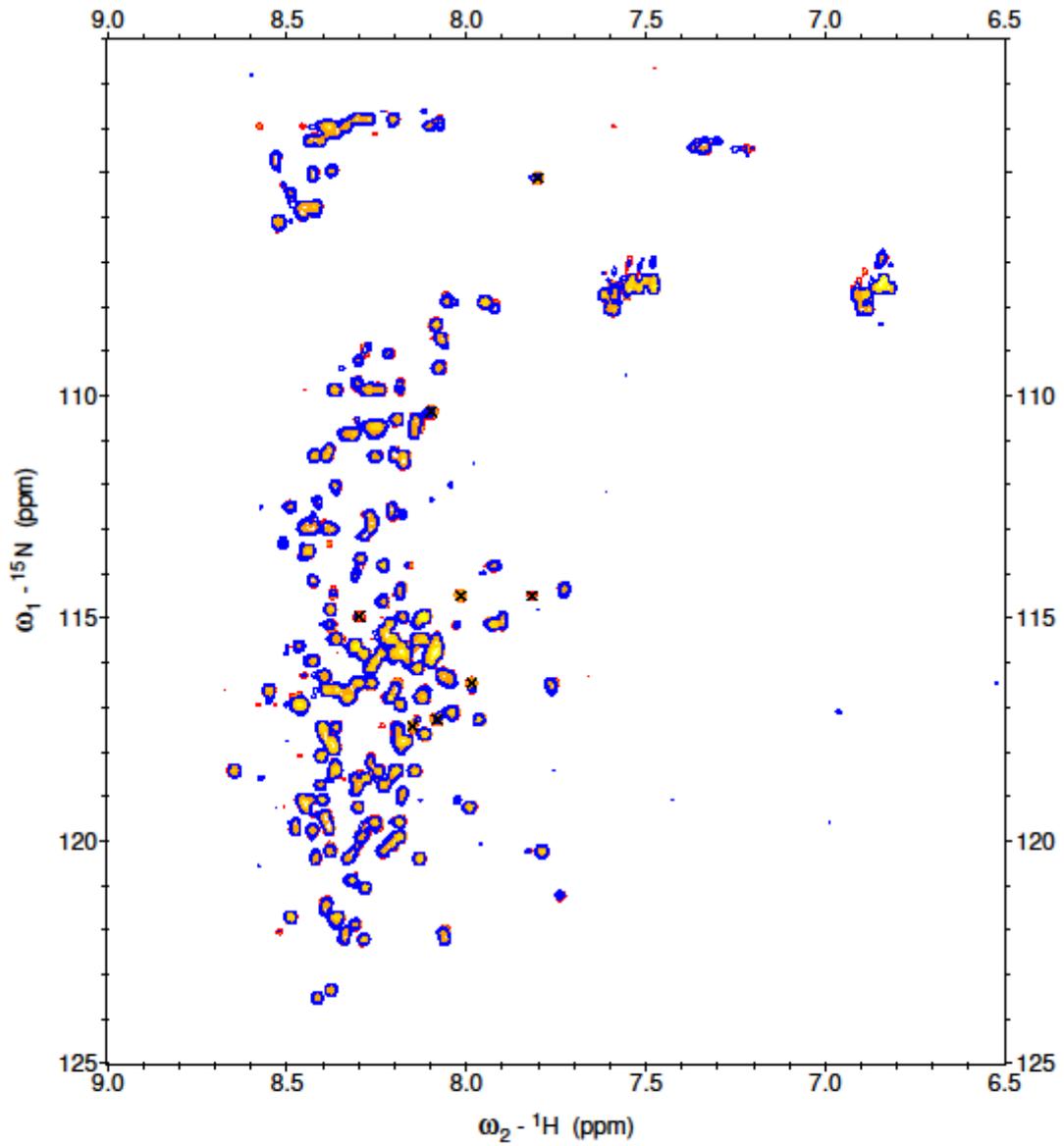
39L23-boil_free+STAT3b-101105_Nhsqc_600.pdf

Figure A.5: HSQC spectrum of SMRT-39L23 in complex with STAT3. Orange peaks indicate free 39L23 and blue peaks indicate 39L23 in complex with STAT3. The protein ratio was 1:1. Spectra recorded by R.Schneider (Blackledge lab, IBS Grenoble).

27M12 15N with biotin tag
15N HSQC, 25C, 800 MHz, phosphate buffer pH 6.5

red: free
blue: with 1:1 STAT3b

Crosses mark peaks that get weaker or vanish



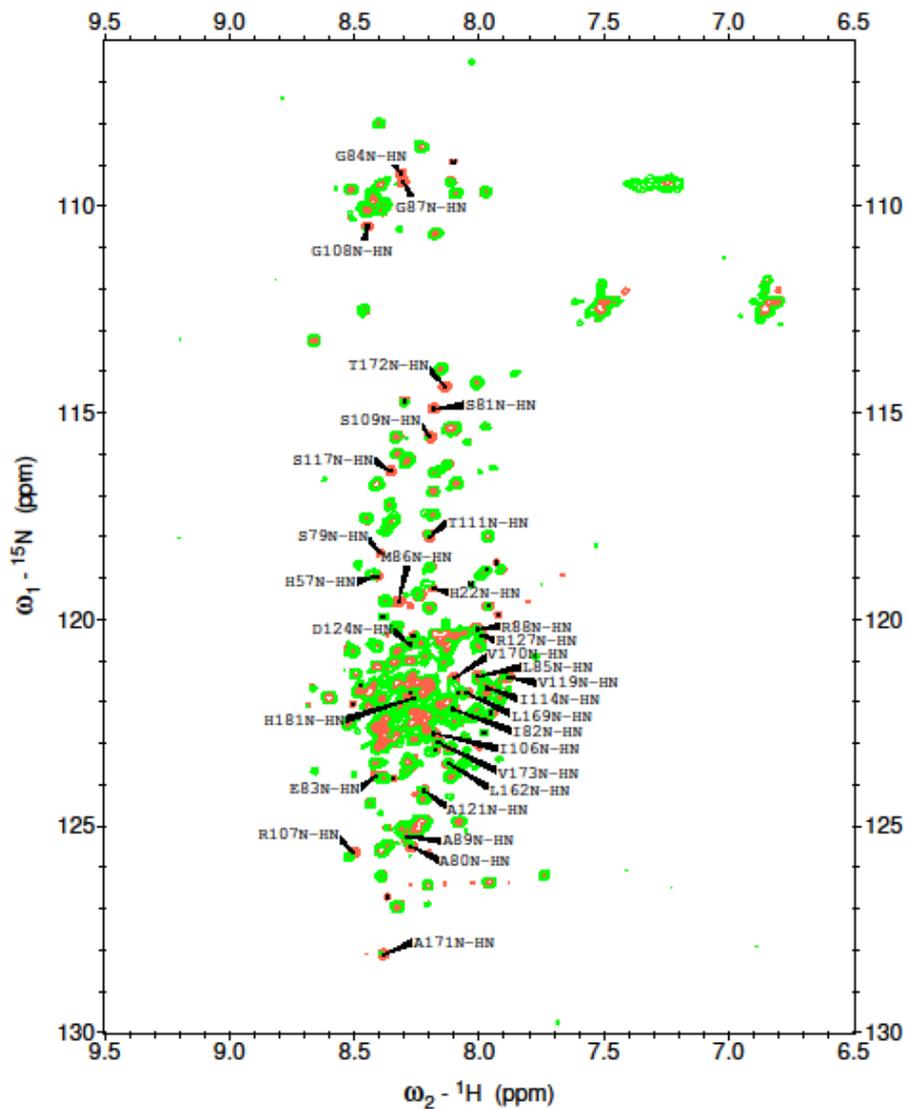
27M12_free+STAT3b_Nhsqc_25c_800_101124.pdf

Figure A.6: HSQC spectrum of SMRT-27M12 in complex with STAT3. . Orange peaks indicate free 27M12 and blue peaks indicate 27M12 in complex with STAT3. The protein ratio was 1:1. Spectra recorded by R.Schneider (Blackledge lab, IBS Grenoble).

39L23 15N no tags (sample of Feb 4, 2011)
15N HSQC, 25C, pH 6.5, 600 MHz

red: free (204 μM)
(39L23_Nhsqc_25C_600cr_110210.save)

green: with 5x STAT3b (39L23: 56 μM , STAT3b: 283 μM)
(39L23+STAT3b_1to5_Nhsqc_25C_600cr_110210.save)



39L23_free+STAT3b-5to1_Nhsqc_25C_600cr_110210_assigned.pdf

Figure A.7: HSQC spectrum of the assigned SMRT-39L23 fragment. Orange peaks indicate free 39L23, green peaks indicate 39L23 in complex with 5x access of STAT3.

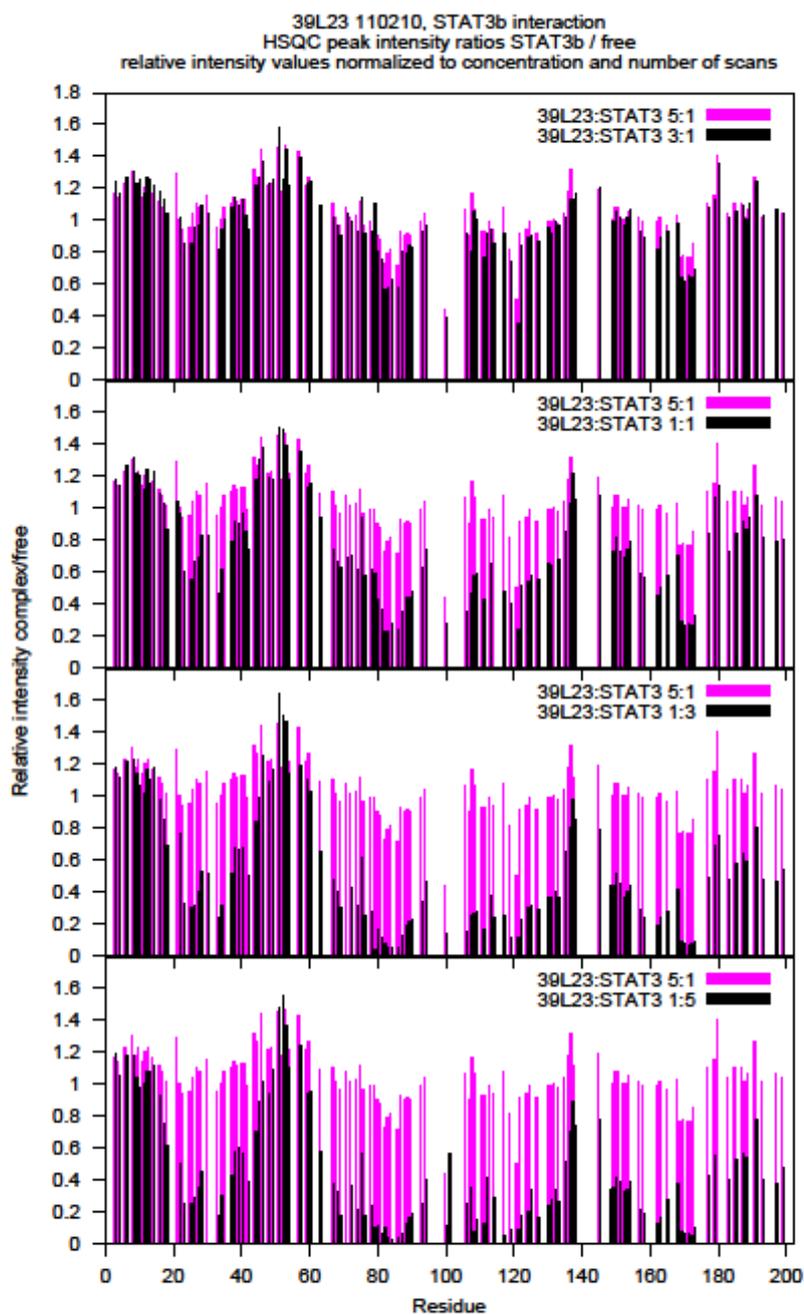


Figure A.8: Peak intensity changed of the amino acids in 39L23 effected by STAT3 binding. Each lane corresponds to one residue of fragment 39L23 starting from 1-199. Changes occur upon STAT3 titration from 1- to 5-fold excess. Assignment performed by R. Schneider (Blackledge lab, IBS Grenoble).

EBNA2 synthetic gene sequence:

1 ATGCCGACCT TTTATCTGGC ACTGCATGGT GGTCAGACCT ATCATCTGAT
51 TGTTGATACC GATAGCCTGG GTAATCCGAG CCTGAGCGTT ATTCCGAGCA
101 ATCCGTATCA AGAACAGCTG AGCGATACAC CGCTGATTCC GCTGACCATT
151 TTTGTTGGTG AAAATACCGG TGTTCCGCCT CCGCTGCCTC CTCCTCCACC
201 ACCTCCGCCT CCGCCACCGC CTCCCCCACC TCCACCACCG CCCCCACCCC
251 CACCTCCCCC TCCGAGTCCG CCTCCACCGC CACCCCCTCC ACCGCCGCCT
301 CAACGTCGTG ATGCCTGGAC CCAAGAACCG AGTCCGCTGG ATCGTGATCC
351 GCTGGGTTAT GATGTTGGTC ATGGTCCGCT GGCAAGCGCA ATGCGTATGC
401 TGTGGATGGC AAACATATATT GTTCGTCAGA GCCGTGGTGA TCGTGGTCTG
451 ATTCTGCCGC AGGGTCCGCA GACCGCACCG CAAGCACGTC TGGTTCAGCC
501 GCATGTTCCG CCTCTGAGAC CGACCGCACC GACCATTCTG AGTCCGCTGA
551 GCCAGCCTAG ACTGACACCG CCTCAGCCGC TGATGATGCC TCCGCGTCCG
601 ACACCTCCGA CTCCTCTGCC TCCGGCAACC CTGACCGTTC CGCCTCGTCC
651 GACCCGTCCG ACAACCTTAC CGCCTACACC GCTGCTGACC GTTCTGCAGC
701 GTCCGACAGA ACTGCAGCCG ACCCCGAGCC CTCCGCGTAT GCACTTACCG
751 GTTCTGCATG TGCCGGATCA GAGCATGCAC CCGCTGACCC ATCAGAGCAC
801 CCCGAATGAT CCTGATTAC CGGAACCGCG TAGCCCGACC GTGTTTTATA
851 ACATTCCGCC TATGCCGCTG CCTCCGAGCC AGTTACCGCC TCCTGCAGCA
901 CCGGCACAGC CTCCGCCTGG TGTTATTAAT GATCAGCAGC TGCATCATCT
951 GCCGAGCGGT CCGCCTTGGT GGCCTCCGAT TTGTGATCCT CCGCAGCCGA
1001 GCAAACCCA GGGTCAGAGT CGTGGTCAGT CACGTGGTCG TGGTCGCGGT
1051 CGTGGTAGAG GTCGCGGTAA AGGTAANAAG CGTGATAAAC AGCGTAAACC
1101 GGGTGGTCCG TGGCGTCCCTG AACCGAATAC CAGCAGCCCG AGCATGCCGG
1151 AACTGAGTCC GGTTCGTTGGT CTGCATCAGG GTCAGGGTGC CGGTGATAGC
1201 CCGACACCGG GTCCGAGCAA TGCAGCACCG GTTTGTGCGTA ATAGCCATAC
1251 CGCAACCCCG AATGTTAGCC CGATTTCATGA ACCGGAAAGC CATAATTAC
1301 CGGAAGCACC GATTCTGTTT CCAGATGATT GGTATCCGCC TAGCATTGAT
1351 CCGGCAGATC TGGATGAAAG CTGGGATTAT ATCTTTGAAA CCACCGAAAG
1401 CCCGAGCAGT GATGAAGATT ATGTTGAAGG TCCGAGCAA CGTCCGCGTC
1451 CGAGCATTCA GTAA

EBNA2 amino acid sequence:

1 MPTFYLLALHG GQTYHLIVDT DSLGNPSLSV IPSNPYQEQ L SDTPLIPLTI
51 FVGENTGVPP PLPPPPPPPP PPPPPPPPP PPPPPPPSP PPPPPPPPP
101 QRRDAWTQEP SPLDRDPLGY DVGHGPLASA MRMLWMANYI VRQSRGDRGL
151 ILPQGPQTAP QARLVQPHVP PLRPTAPTIL SPLSQPRLTP PQPLMMPRP
201 TPPTPLPPAT LTVPPRPTRP TTLPPTPLLT VLQRPTLQ PTPSPRMHLP
251 VLHVPDQSMH PLTHQSTPND PDSPEPRSPT VFYNIPPMPL PPSQLPPPAA
301 PAQPPPGVIN DQQLHHLPSG PPWWPPICDP PQPSKTQGQS RGQSRGRGRG
351 RGRGRGKGS RDKQRKPGGP WRPEPNTSSP SMPELSPVLG LHQQGAGDS
401 PTPGPSNAAP VCRNSHTATP NVSPIHEPES HNSPEAPILF PDDWYPPSID
451 PADLDESWDY IFETTESPSS DEDYVEGPSK RPRPSIQ

Sequence of synthetic codon optimized DNA encoding EBNA2 C-terminal region

(from BamH1 to XbaI restriction site)

```
1    gaattccatc atgcctacat tctatcttgc gttacatggg ggacaaacat
51   atcatcta atgttgacacg gatagtcttg gaaaccgctc actctcagta
101  attccctcga atccctacca ggaacaactg tcagacactc cattaattcc
151  actaacaatc tttggtgggg aaaacacggg ggtgccccca cactccccac
201  ccccccccc accaccacc ccaccacccc caccaccccc accacccccca
251  ccacccccac cacctccacc accttcacca ccacccccgc ccccaccacc
301  cccaccacct cagcgcaggg atgcctggac acaagagcca tcacctcttg
351  atagggatcc gctaggatat gacgtcgggc atggacctct agcatctgct
401  atgcgaatgc tttggatggc taattatatt gtaagacaat cacgggggtga
451  ccggggcctt attttgccac aaggcccaca aacagcccct caggccaggt
501  tggtcagacc acatgtcccc cctctacgcc cgacagcacc caccatthttg
551  tcacctctgt cacaaccgag gcttaccctt ccacaaccac tcatgatgcc
601  accaaggcct acccctccta ccctctgccc acctgcaaca ctaacgggtgc
651  caccaaggcc taccgctcct accactctgc caccacacc actactcacg
701  gtactacaaa ggccctaccga acttcaaccc acaccatcac caccacgcat
751  gcatctccct gtcttgcatg tgccagacca atcaatgcac cctcttactc
801  atcaaagcac cccaaatgat ccagatagtc cagaaccacg gtccccgact
851  gtatthttata acattccacc tatgccatta cccccctcac aattgccacc
901  accagcagca ccagcacagc cacctccagg ggtcatcaac gaccaacaat
951  tacatcatct accctcgggg ccaccatggt ggccaccat ctgcgacccc
1001 ccgcaaccct ctaagactca aggccagagc cggggacaga gcagggggag
1051 gggcaggggc aggggcaggg gcaggggcaa gggcaagtcc agggacaagc
1101 aacgcaagcc cgggtggacct tggagaccag agccaaacac ctccagtcct
1151 agcatgcctg aactaagtcc agtcctcggg ctcatcaggg gacaaggggc
1201 tggggactca ccaactcctg gcccatccaa tgccgcccc gtttgtagaa
1251 attcacacac ggcaaccctt aacgthttcac caatacatga accggagtcc
1301 cataatagcc cagaggctcc cattctcttc cccgatgatt ggtatcctcc
1351 atctatagac cccgcagact tagacgaaag ttgggattac atthtttgaga
1401 caacagaatc tcctagctca gatgaagatt atgtggaggg acccagtaaa
1451 agacctgcc cctccatcca gccatacgat gttccagatt acgctagcta
1501 aagatcttat taaagcagaa cttgthttatt gcagcttata atggttacaa
1551 ataaagcaat agcatcacia atthcaciaa taaagcattt thttcactgc
1601 attctagttg tggthttgtcc aaactcatca atgtatctta tcatgtctgg
1651 tcgactctag a
```

Sequence of SMRT synthetic codon optimized DNA

```
1    atgtcgggat ccacacagcc tgtggcacag acgtggaggg cactgagacc
51   ccgtaccceg cccacacagcc thttcctacc agtgcagatc gcccggacgc
101  acacggacgt cgggctcctg gagtaccagc accactcccg cgactatgcc
151  tcccacctgt cgcccggctc catcatccag cccagcggc ggaggccctc
201  cctgctgtct gagttccagc ccgggaatga acggtcccag gagctccacc
251  tgcggccaga gtcccactca tacctgcccg agctggggaa gtcagagatg
301  gagttcattg aaagcaagcg ccctcggcta gagctgctgc ctgacccccct
351  gctgcgaccg tcaccctgc tggccacggg ccagcctgcg ggatctgaag
401  acctaccaa ggaccgtagc ctgacgggca agctggaacc ggtgtctccc
451  cccagcccc cgcaactga ccctgagctg gagctggtgc cgccacggct
501  gtccaaggag gagctgatcc agaacatgga ccgcgtggac cgagagatca
```

551 ccatggtaga gcagcagatc tctaagctga agaagaagca gcaacagctg
601 gaggaggagg ctgccaaacc gcccgagcct gagaagcccg tgtcaccgcc
651 gcccatcgag tcgaagcacc gcagcctggg gcagatcatc tacgacgaga
701 accggaagaa ggctgaagct gcacatcgga ttctggaagg cctggggccc
751 caggtggagc tgccgctgta caaccagccc tccgacaccc ggcagtatca
801 tgagaacatc aaaataaacc aggcgatgcg gaagaagcta atcttgtact
851 tcaagaggag gaatcacgct cggaaacaat gggagcagaa gttctgccag
901 cgctatgacc agctcatgga ggctggggag aagaaggtgg agcgcacoga
951 gaacaacccc cggcggcggg ccaaggagag caaggtgcgc gactactacg
1001 agaagcagtt ccctgagatc cgcaagcagc gcgagctgca ggagcgcagc
1051 cagagcaggg tgggcccagcg gggcagtggg ctgtccatgt cggcccggcg
1101 cagcagcac gaggtgtcag agatcatcga tggcctctca gagcaggaga
1151 acctggagaa gcagatgccc cagctggccc tgatcccgcc catgctgtac
1201 gacgctgacc agcagcgcac caagttcatc aacatgaacg ggcttatggc
1251 cgaccccatg aaggtgtaca aagaccgcca ggtcatgaac atgtggagtg
1301 agcaggagaa ggagacctc cgggagaagt tcatgcagca tcccaagaac
1351 tttggcctga tcgcatcatt cctggagagg aagacagtgg ctgagtgcgt
1401 cctctattac tacctgacta agaagaatga gaactataag agcctgggtga
1451 gacggagcta tcggcgcccgc ggcaagagcc agcagcagca acaacagcag
1501 cagcagcagc agcagcagca gcagcagcag cccatgccc gcagcagcca
1551 ggaggagaaa gatgagaagg agaaggaaa ggaggcggag aaggaggagg
1601 agaagccgga ggtggagaac gacaaggaag acctcctcaa ggagaagaca
1651 gacgacacct caggggagga caacgacgag aaggaggctg tggcctccaa
1701 aggccgcaa actgccaaca gccagggaag acgcaaaggc cgcacacccc
1751 gctcaatggc taatgaggcc aacagcaggg aggccatcac ccccagcag
1801 agcggcagc tggcctccat ggagctgaat gagagtctc gctggacaga
1851 agaagaatg gaaacagcca agaaaggtct cctggaacac ggccgcaact
1901 ggtcggccat cgcccggatg gtgggctcca agactgtgtc gcagtgtaag
1951 aacttctact tcaactaca gaagaggcag aacctcgatg agatcttgca
2001 gcagcacaag ctgaagatgg agaaggagag gaacgcgcgg aggaagaaga
2051 agaaagcggc ggcggcggcc agcagaggagg ctgcattccc gcccggtgtg
2101 gaggatgagg agatggaggc gtcgggctgt agcggaaatg agggaggagat
2151 ggtggaggag gctgaagcca ctgtcaacaa cagctcagac accgagagca
2201 tcccctctcc tcacactgag gccgccaagg acacagggca gaatggggcc
2251 aagccccag ccaccctggg cgcgacggg ccacccccag ggccaccac
2301 cccaccaccg gaggacatcc cggccccac tgagcccacc cggcctctg
2351 aagccaccgg agcccctacg cccccaccag cacccccac gcccctctgca
2401 cctcctctctg tgggtcccaa ggaggagaag gaggaggaga ccgcagcagc
2451 gccccagtg gaggaggggg aggagcagaa gcccccgcg gctgaggagc
2501 tggcagtgga cacaggaag gccgaggagc ccgtcaagag cgagtgcacg
2551 gaggaagccg agggagggcc ggccaagggc aaggacgcgg aggccgctga
2601 ggccacggcc gagggggcgc tcaaggcaga gaagaaggag ggccggagcg
2651 gcagggccac cacagccaag agctcgggcg cccccagga cagcgactcc
2701 agtgctacct gcagtgcaga cgaggtgat gaggccgagg gcggcgacaa
2751 gaaccggctg ctgtcccaa ggcccagcct cctcaccg actggcgacc
2801 cccgggcaa tgccctaccc cagaagccac tggacctgaa gcagctgaag
2851 cagcagcgg ctgccatccc cccatccag gtcaccaaaag tccatgagcc
2901 ccccgggag gacgcagctc ccaccaagcc agctcccca gcccaccgc
2951 caccgcaaaa cctgcagccg gagagcgagc cccctcagca gcctggcagc
3001 agccccggg gcaagagcag gagcccggca cccccggcg acaaggaggc
3051 agagaagcct gtgttcttcc cagccttcgc agccgaggcc cagaagctgc
3101 ctggggacc cccttgctgg acttcggcc tgccctccc cgtgcccccc
3151 cgtgaggtga tcaaggcctc cccgcatgcc ccggaccct cagccttctc

3201 ctacgctcca cctggtcacc cactgcccct gggcctccat gacactgccc
3251 ggcccgtcct gccgcgcccc cccaccatct ccaaccgcgc tcccctcatc
3301 tcctctgcca agcaccacag cgtcctcgag aggcaaatag gtgccatctc
3351 ccaaggaatg tcgggtccagc tccacgtccc gtactcagag catgccaagg
3401 ccccgggtggg ccctgtcacc atggggctgc ccctgcccac ggacccccaaa
3451 aagctggcac ccttcagcgg agtgaagcag gagcagctgt ccccacgggg
3501 ccaggctggg ccaccggaga gcctgggggt gccacagcc caggaggcgt
3551 ccgtgctgag agggacagct ctgggctcag ttccgggcgg aagcatcacc
3601 aaaggcattc ccagcacacg ggtgccctcg gacagcgcca tcacataccg
3651 cggctccatc acccacggca cgccagctga cgtcctgtac aaggggacca
3701 tcaccaggat catcggcgag gacagcccga gtcgcttggg ccgcggccgg
3751 gaggacagcc tgcccagggg ccacgtcatc tacgaaggca agaaggcca
3801 cgtcttgtcc tatgaggggt gcatgtctgt gaccagtgcc tccaaggagg
3851 acggcagaag cagctcagga ccccccatg agacggccgc cccaagcgc
3901 acctatgaca tgatggaggg ccgcgtgggc agagccatct cctcagccag
3951 catcgaaggt ctcatggggc gtgccatccc gccggagcga cacagcccc
4001 accacctcaa agagcagcac cacatccgcg ggtccatcac acaagggatc
4051 cctcggctcct acgtggaggc acaggaggac tacctgcgtc gggaggccaa
4101 gtcctaaag cgggagggca cgcctccgc ccaccgccc tcacgggacc
4151 tgaccgaggc ctacaagacg caggccctgg gccccctgaa gctgaagccg
4201 gcccatgagg gcctgggtggc cacggtgaag gaggcgggcc gctccatcca
4251 tgagatcccg cgcgaggagc tgcggcacac gcccgagctg cccctggccc
4301 cgcggccgct caaggagggc tccatcacgc agggcacccc gctcaagtac
4351 gacaccggcg cgtccaccac tggctccaaa aagcacgacg tacgctccct
4401 catcggcagc cccggccgga cgttcccacc cgtgcaccgc ctggatgtga
4451 tggccgacgc ccgggcaactg gaacgtgct gctacgagga gagcctgaag
4501 agccggccag ggaccgcccag cagctcgggg ggctccattg cgcgcggcgc
4551 cccggtcatt gtgcctgagc tgggtaagcc gcggcagagc cccctgacct
4601 atgaggacca cggggcaccc tttgcgggcc acctcccacg aggttcgccc
4651 gtgaccacgc gggagcccac gccgcgctg caggagggca gccttctgctc
4701 cagcaaggca tcccaggacc gaaagctgac gtcgacgcct cgtgagatcg
4751 ccaagtcccc gcacagcacc gtgcccagac accaccaca ccccatctcg
4801 ccctatgagc acctgcttcg gggcgtgagt ggcgtggacc tgtatcgcag
4851 ccacatcccc ctggccttcg accccacctc catacccgcg ggcatecctc
4901 tggacgcagc cgctgcctac tacctgcccc gacacctggc ccccaacccc
4951 acctaccgcg acctgtacc accctacctc atccgcggct accccgacac
5001 ggcggcgcctg gagaaccggc agaccatcat caatgactac atcacctcgc
5051 agcagatgca ccacaacgcg gccaccgcca tggcccagcg agctgatatg
5101 ctgagggggc tctcgccccg cgagtcctcg ctggcactca actacgctgc
5151 ggggtccccga ggcacatcgc acctgtccca agtgccacac ctgcctgtgc
5201 tcgtgccccg gacaccaggc accccagcca ccgcatgga ccgccttgcc
5251 tacctccccg ccgcgccccg gcccttcagc agccgcccaca gcagctcccc
5301 actctccccg ggaggtccaa cacacttgac aaaaccaacc accacgtcct
5351 cgtccgagcg ggagcgagac cgggatcgag agcgggaccg ggatcgggag
5401 cgggaaaagt ccacctcac gtccaccacg acggtggagc acgcacccat
5451 ctggagacct ggtacagagc agagcagcgg cagcagcggc agcagcggcg
5501 ggggtggggg cagcagcagc cgccccgcct cccactcca tgcccaccag
5551 cactcgccca tctcccctcg gaccaggat gccctccagc agagaccag
5601 tgtgcttcac aacacaggca tgaagggtat catcaccgct gtggagccca
5651 gcacgcccac ggtcctgagg tccacctcca cctcctcacc cgttcgccccg
5701 gctgccacat tcccacctgc caccactgc cactggggcg gcaccctoga
5751 tggggtctac cctaccctca tggagcccgt cttgctgccc aaggaggccc
5801 cccgggtcgc ccggccagag cggccccgag cagacaccgg ccatgccttc

5851 ctcgccaagc ccccagcccc ctccgggctg gagccccgct cctccccag
5901 caagggctcg gagccccggc ccctagtgcc tctgtctct ggccacgcca
5951 ccatcgcccg caccctgcyg aagaacctcg cacctacca cgcagccccg
6001 gaccgcccgg cgcacactgc ctccggcctcg gaccgcacc gggaaaagac
6051 tcaaagtaaa cccttttcca tccaggaact ggaactccgt tctctgggtt
6101 accacggcag cagctacagc cccgaagggg tggagccgt cagccctgtg
6151 agctcaccca gtctgaccca cgacaagggg ctccccaaag acctggaaga
6201 gctcgacaag agccacctgg agggggagct gcggcccaag cagccaggcc
6251 ccgtgaagct tggcggggag gccgccacc tcccacacct gcggccgctg
6301 cctgagagcc agccctcgtc cagcccgtg ctccagaccg ccccaggggt
6351 caaaggtcac cagcgggtgg tcaccctggc ccagcacatc agtgagggtca
6401 tcacacagga ctacaccggg caccaccac agcagctcag cgcaccctg
6451 cccgcccccc tctactcctt ccctggggcc agctgccccg tcttgacct
6501 ccgcccacca cccagtgacc tctacctccc gccccggac catggtgcc
6551 cggcccgtgg ctccccccac agcgaagggg gcaagaggtc tccagagcca
6601 aacaagacgt cgggtcttggg tgggtgtgag gacggtattg aacctgtgtc
6651 cccaccggag ggcatgacgg agccagggca ctcccggagt gctgtgtacc
6701 cgctgctgta ccgggatggg gaacagacgg agcccagcag gatgggctcc
6751 aagtctccag gcaacaccag ccagccgcca gccttcttca gcaagctgac
6801 cgagagcaac tccgccatgg tcaagtccaa gaagcaagag atcaacaaga
6851 agctgaacac ccacaaccgg aatgagcctg aatacaatat cagccagcct
6901 gggacggaga tcttcaatat gcccgccatc accggaacag gccttatgac
6951 ctatagaagc caggcgggtgc aggaacatgc cagcaccac atggggctgg
7001 aggccataat tagaaaggca ctcatgggtg gcggcgggaa ggccaaggtc
7051 tctggcagac ccagcagccg aaaagccaag tccccggccc cgggcctggc
7101 atctggggac cggccaccct ctgtctcctc agtgcactcg gagggagact
7151 gcaaccgccc gacgccgctc accaaccgcyg tgtgggagga caggccctcg
7201 tccgcagggt ccacgccatt cccctacaac ccctgatca tgcggctgca
7251 ggcgggtgtc atggcttccc cacccccacc gggcctccc gcgggcagcg
7301 ggcccctcgc tggccccac cacgcctggg acgaggagcc caagccactg
7351 ctctgctcgc agtacgagac actctccgac agcgagtga

Acknowledgements

First of all, I like to thank my supervisor Darren Hart for giving me the opportunity to conduct this work and also for his scientific guidance and support.

I like to thank the EMBL International PhD programme for financial support. I would also like to thank EMBO for rewarding me a short term fellowship which allowed me to spend 1 month in the Kempkes lab in Munich.

I wish to thank the members of my Thesis Advisory Committee Toby Gibson, Ramesh Pilai and Wim Burmeister for their efforts and good advices during my PhD.

I would also like to thanks the members of my thesis jury Darren Hart, Peter Tompa, Paul Farrell, Imre Berger, Wim Burmeister and François Penin for their time and kind revision of my thesis work.

I thank Bettina Kempkes and her lab members for welcoming me in their lab and for the nice working atmosphere.

Furthermore, I wish to thank Robert Schneider and Martin Blackledge for devoting some working and magnet time for my project.

I am grateful to all the Hart team lab members which accompanied me during my time in Grenoble. I always enjoyed coming to the lab and spending time with all my lab mates. I especially would like to thank Mareike Roth, Philippe Mas and Danielle Desravines for their huge support.

I also like to thank Darren Hart, Helke Hillebrand and everybody else involved for the optimal support which allowed me taking care of my baby daughter and finishing my thesis in the most comfortable conditions.

I also thank my family for their constant support and love.