

**MAPPING OF THE DOMAINS INVOLVED IN B-MYB-  
PARTNER PROTEIN INTERACTIONS AND  
CHARACTERISATION OF THEIR STRUCTURAL  
PROPERTIES AND FEATURES**

**Thesis submitted for the degree of**

**Doctor of Philosophy**

**at the University of Leicester**

**by**

**Nuvjeevan Singh Dosanjh BSc (Birmingham)**

**Department of Biochemistry**

**University of Leicester**

**October 2005**

UMI Number: U203777

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U203777

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.  
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against  
unauthorized copying under Title 17, United States Code.



ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

## ACKNOWLEDGEMENTS

I would like to express my immense gratitude to my supervisor Dr. Mark Carr for providing the opportunity to conduct research in his laboratory and for his continued guidance and support. I am indebted to Professor Karl-Heinz Klempnauer, Dr. David Heery and Dr. Mark Pfuhl for providing technical assistance with the project and for their useful discussions. I would also like to thank the Biotechnology and Biological Sciences Research Council (BBSRC) for funding this studentship.

I extend my thanks, in particular, to Dr. Phil Renshaw for his scholarly advice and expertise- I have learnt a great deal. I would also like to acknowledge the contribution of other members of the laboratory at the University of Leicester, Vaclav, Kirsty, Lorna, Ojay and Rebecca all of whom have been wonderful work colleagues and good friends.

To Jane Cameron, Alison Clague and Ian Johnson for their continued support and generous hospitality. A sincere thank you.

Finally, a heart felt thank you to my parents and my brother who have always been there for me and without whom, this work would not have been possible.

Nuvjeevan Singh Dosanjh, Leicester, Spring 2005

**Nuvjeevan Singh Dosanjh**

**Title:** Mapping of the domains involved in B-Myb-partner protein interactions and characterisation of their structural properties and features

### **ABSTRACT**

B-Myb is a member of the highly conserved Myb family of transcriptional activators, which is expressed in a wide variety of cell lineages and is essential for cell division. The transcriptional activity of B-Myb is significantly enhanced in S-phase of the cell cycle where it is coordinately regulated by post-translational modifications and interactions with cellular co-factors such as p300 and cyclin D1.

In the work reported here, the minimal sites of interaction between B-Myb and its partner proteins p300 and cyclin D1 were mapped, and the structural properties and features of the interacting regions characterised in detail. A transactivation-associated central region of B-Myb was conclusively shown to bind to a region of p300 corresponding to the TAZ2, zinc ion binding domain. As is common for many transcriptional activation domains, the B-Myb transactivation region is intrinsically disordered in isolation but exhibits coupled folding upon binding p300. The specific B-Myb-p300 complex formed was found to adopt a 1:1 stoichiometry with a measured dissociation constant of less than  $\sim 3.0 \times 10^{-8}$  M. An adjacent region of B-Myb containing part of exon 9a, was also found to interact with the preceding ZZ region of p300, but substantially weaker than observed for the B-Myb-TAZ2 interaction. These findings clearly show that the conserved cysteine/histidine rich region 3 (C/H3) of p300 provides an extensive binding site for B-Myb.

In parallel work, a large central part of B-Myb containing both the transactivation region and exon 9a was found to show no binding to full-length cyclin D1. Together with previous work, this suggests that residues in the N-terminal DNA binding domain of B-Myb are also involved in cyclin D1 binding.

Overall, this work suggests that the inhibition of cooperativity between B-Myb and p300 by cyclin D1 is not due to competition for a common binding site on B-Myb.

## CONTENTS

<b>Acknowledgements</b> .....	1
<b>Abstract</b> .....	2
<b>Contents</b> .....	3
<b>Abbreviations</b> .....	8
<b>Chapter 1- Introduction</b> .....	11
1.1 <i>Transcription Factors in Eukaryotic Gene Regulation</i> .....	11
1.2 <i>The Myb Family of Transcription Factors</i> .....	14
1.3 <i>Functional Regions/Domains of the Myb Family</i> .....	16
1.3.1 <i>DNA Binding Domain (DBD)</i> .....	16
1.3.2 <i>Transactivation Domain (TAD)</i> .....	19
1.3.3 <i>Negative Regulatory Domain (NRD) and the Conserved Region (CR)</i> ..	22
1.4 <i>Myb Gene Expression</i> .....	24
1.5 <i>B-Myb: a Cell-Cycle Regulated Protein</i> .....	25
1.6 <i>Regulation of B-Myb Gene Expression</i> .....	25
1.7 <i>Regulation of B-Myb Activity</i> .....	27
1.7.1 <i>Covalent Modification</i> .....	27
1.7.2 <i>Protein-Protein Interactions</i> .....	29
1.7.2.1 <i>p300/CBP</i> .....	29
1.7.2.2 <i>cyclin D1</i> .....	32
1.7.2.3 <i>Other B-Myb Interacting Proteins</i> .....	33
1.7.3 <i>Degradation of B-Myb</i> .....	35
1.8 <i>B-Myb Regulated Genes</i> .....	35
1.9 <i>B-Myb in Tumourigenesis</i> .....	36
1.10 <i>Aims of the Project</i> .....	37
<b>Chapter 2- Cloning and purification of B-Myb, p300 and cyclin D1 proteins</b> ... 39	
2.1 <i>Introduction and Aims</i> .....	39
2.2 <i>Materials and Methods</i> .....	39
2.2.1 <i>B-Myb Constructs and Cloning Strategy</i> .....	39
2.2.2 <i>p300 Constructs and Cloning Strategy</i> .....	42
2.2.3 <i>Cyclin D1 Constructs and Cloning Strategy</i> .....	44
2.2.4 <i>Agarose Gel Electrophoresis</i> .....	45
2.2.5 <i>SDS PAGE Gel Electrophoresis</i> .....	45

2.2.6	<i>Electrospray Mass Spectroscopy Sample Preparation</i> .....	46
2.2.7	<i>N-terminal Sequencing</i> .....	46
2.2.8	<i>GST/ B-Myb Protein Expression Trials</i> .....	46
2.2.9	<i>Expression and Purification of GST/ B-Myb Fusion Proteins</i> .....	47
2.2.10	<i>His-tagged p300 Protein Expression Trials</i> .....	48
2.2.11	<i>Expression and Inclusion Body Preparation of His-tagged p300 Fusion Proteins</i> .....	49
2.2.12	<i>Refolding and Purification of p300 1603-1725</i> .....	50
2.2.13	<i>Refolding and Purification of p300 1603-1839</i> .....	50
2.2.14	<i>Refolding and Purification of p300 1726-1839</i> .....	51
2.2.15	<i>His-tagged Cyclin D1 Protein Expression Trials</i> .....	52
2.2.16	<i>Expression and Inclusion Body Preparation of His-tagged Cyclin D1</i>	52
2.2.17	<i>Refolding and Purification of His-tagged Cyclin D1</i> .....	53
2.3	<i>Results</i> .....	54
2.3.1	<i>B-Myb Construct Design</i> .....	54
2.3.2	<i>PCR Amplification and Cloning of Murine B-Myb Sequences</i> .....	54
2.3.3	<i>p300 Construct Design</i> .....	57
2.3.4	<i>PCR Amplification and Cloning of Human p300 Sequences</i> .....	59
2.3.5	<i>PCR Amplification and Cloning of Human cyclin D1</i> .....	61
2.3.6	<i>GST-tagged B-Myb Protein Expression Trials</i> .....	63
2.3.7	<i>GST-tagged B-Myb Expression and Purification</i> .....	64
2.3.8	<i>PreScission Protease Cleavage and Removal of GST tag From B-Myb 275-376</i> .....	67
2.3.9	<i>Removal of Contaminating DNA from GST/B-Myb R2R3</i> .....	69
2.3.10	<i>His-tagged p300 Protein Expression Trials</i> .....	71
2.3.10.1	<i>His-tagged p300 1603-1725 Expression Trials</i> .....	71
2.3.10.2	<i>His-tagged p300 1603-1839 Expression Trials</i> .....	72
2.3.10.3	<i>His-tagged p300 1726-1839 Expression Trials</i> .....	75
2.3.11	<i>His-tagged cyclin D1 Expression Trials</i> .....	76
2.3.12	<i>p300 Protein Expression and Purification</i> .....	78
2.3.12.1	<i>p300 1603-1725 Expression and Purification</i> .....	79
2.3.12.2	<i>p300 1603-1839 Expression and Purification</i> .....	83
2.3.12.3	<i>p300 1726-1839 Expression and Purification</i> .....	86

2.3.13 Expression and Purification of His-tagged cyclin D1.....	90
2.4 Discussion.....	93
<b>Chapter 3- Structural Characterisation of B-Myb, p300 and cyclin D1</b>	
<b>Polypeptides.....</b>	<b>96</b>
3.1 Introduction and Aims.....	96
3.2 Material and Methods.....	96
3.2.1 Fluorescence Spectroscopy.....	96
3.2.2 Circular Dichroism Spectroscopy.....	96
3.2.3 Chemical Denaturation of p300 and cyclin D1 Polypeptides.....	97
3.2.4 Temperature Denaturation Studies of p300 1726-1839.....	97
3.2.5 Zn <sup>2+</sup> Binding Properties of p300 Polypeptides.....	97
3.2.6 Secondary Structure Predictions of B-Myb and p300 Polypeptides.....	98
3.3 Results.....	99
3.3.1 Structural Characterisation of B-Myb 275-376.....	99
3.3.1.1 Secondary Structure Prediction.....	99
3.3.1.2 Biophysical Characterisation of B-Myb 275-376: Secondary and Tertiary Structural Analysis.....	101
3.3.2 Structural Characterisation of p300 1603-1725.....	103
3.3.2.1 Secondary Structure Prediction.....	103
3.3.2.2 Biophysical Characterisation of p300 1603-1725: Secondary and Tertiary Structural Analysis.....	104
3.3.2.3 Chemical Denaturation of p300 1603-1725.....	106
3.3.2.4 Zn <sup>2+</sup> Binding Properties of p300 1603-1725.....	107
3.3.3 Structural Characterisation of p300 1603-1839.....	109
3.3.4 Structural Characterisation of p300 1726-1839.....	109
3.3.4.1 Secondary Structure Prediction.....	109
3.3.4.2 Biophysical Characterisation of p300 1726-1839: Secondary and Tertiary Structure Analysis.....	111
3.3.4.3 Effects of Thermal Variation on p300 1726-1839.....	112
3.3.4.4 Zn <sup>2+</sup> Binding Properties of p300 1726-1839.....	114
3.3.5 Structural Characterisation of His-tagged cyclin D1.....	115
3.3.5.1 Biophysical Characterisation of His-tagged cyclin D1: Secondary and Tertiary Structure Analysis.....	115
3.3.5.2 Chemical Denaturation of His-tagged cyclin D1.....	117

3.4 Discussion.....	119
3.4.1 B-Myb 275-376: An Intrinsically Disordered Polypeptide.....	119
3.4.2 Structural Characterisation of p300 1603-1725 and Comparison With the CBP ZZ domain.....	119
3.4.3 Structural Characterisation of p300 1726-1839 and Comparison with the CBP TAZ2 domain.....	122
3.4.4 Structural Characterisation of His-tagged cyclin D1 and Comparison with cyclin A.....	124
<b>Chapter 4-Protein-Protein Interaction Studies.....</b>	<b>128</b>
4.1 Introduction.....	128
4.1.1 Mapping of B-Myb and p300/cyclin D1 Interactions to Date and Aims.	128
4.2 Materials and Methods.....	128
4.2.1 Batch-Type Pull-Down Binding Assays.....	128
4.2.2 Column-Based Pull-Down Binding Assays.....	129
4.2.3 Western Detection of His-tagged cyclin D1 in Pull-Down Assays against GST/B-Myb proteins.....	129
4.2.4 Protease Protection Assays.....	130
4.2.5 Monitoring Protein-Protein Interactions by Intrinsic Tryptophan Fluorescence.....	130
4.2.6 Chemical Denaturation of the B-Myb 275-376•p300 1726-1839 Complex.....	131
4.3 Results.....	131
4.3.1 Interactions between p300 1603-1725 and B-Myb.....	131
4.3.1.1 Batch-Wise Assays.....	131
4.3.1.2 Column-Based Pull-Down Assays.....	132
4.3.1.3 Protease Protection Assays.....	137
4.3.2 Interactions between His-tagged p300 1603-1839 and B-Myb.....	146
4.3.3 Interactions between p300 1726-1839 and B-Myb.....	148
4.3.3.1 Column GST pull-down Assays using His-tagged p300 1726-1839.....	148
4.3.3.2 Column GST pull-down Assay using p300 1726-1839 without the His-tag.....	153
4.3.3.3 Monitoring p300 1726-1839/B-Myb 275-376 Complex Formation by Intrinsic Tryptophan Fluorescence.....	153

4.3.3.4 Structural Characterisation of the B-Myb 275-376•p300 1726-1839 Complex.....	156
4.3.4 Interactions between His-tagged cyclin D1 and B-Myb.....	157
4.4 Discussion.....	163
4.4.1 Interaction between p300 1603-1725 and B-Myb.....	163
4.4.2 Interaction of p300 1603-1839 with B-Myb.....	165
4.4.3 Interaction of p300 1726-1839 with B-Myb.....	165
4.4.4 Interaction of cyclin D1 with B-Myb.....	167
4.4.5 The protein-protein interaction strategy.....	168
<b>Chapter 5-Discussion and Conclusions.....</b>	<b>171</b>
5.1 Experimental considerations.....	171
5.1.1 The choice of protein constructs for structural and functional studies... 171	
5.1.2 The functional significance of the chosen protein constructs..... 173	
5.2 The B-Myb Transactivation Region: Rationales for Intrinsic Disorder..... 176	
5.3 The B-Myb Transactivation Region: Comparisons with c-Myb..... 178	
5.4 The B-Myb-p300 Interaction..... 180	
5.5 The ZZ domain: A Potential Protein Interaction Module..... 181	
5.6 The B-Myb 275-376•p300 1726-1839 Complex: a High Affinity Association.....	181
5.7 Comparison of B-Myb with other TAZ2 Binding Proteins..... 183	
5.8 cyclin D1 and its Interaction with B-Myb..... 186	
5.9 Conclusions..... 186	
<b>Appendix.....</b>	<b>188</b>
A.1 Reagents and Culture Media..... 188	
A.1.1 Selected Reagents..... 188	
A.1.2 Culture Media..... 189	
A.2 E. coli Strains..... 190	
A.3 Competent Cell Preparation, Transformation and Restriction Digest..... 190	
A.3.1 Preparation of Competent Cells..... 190	
A.3.2 Transformation of Competent Cells..... 191	
A.3.3 Restriction Digest Protocol..... 192	
<b>References.....</b>	<b>193</b>

## ABBREVIATIONS

### Amino Acids

Ala (A)	alanine	Leu (L)	leucine
Asn (N)	asparagine	Lys (K)	lysine
Asp (D)	aspartic acid	Met (M)	methionine
Arg (R)	arginine	Phe (F)	phenylalanine
Cys (C)	cysteine	Pro (P)	proline
Glu (E)	glutamic acid	Ser (S)	serine
Gln (Q)	glutamine	Thr (T)	threonine
Gly (G)	glycine	Trp (W)	tryptophan
His (H)	histidine	Tyr (Y)	tyrosine
Ile (I)	isoleucine	Val (V)	valine

### DNA bases

A	adenine
C	cytosine
G	guanine
T	thymine
Py	pyrimidine
U	purine

### Others

$\lambda_{\max}$	The wavelength of maximum fluorescence emission
amp	ampicillin
AMV	avian myeloblastosis virus
Bistris	2-[bis(2-hydroxyethyl)amino]-2-hydroxymethyl-propane-1,3-diol
B-Myb R2R3	minimal B-Myb DNA binding domain
bp	base pair
BSA	bovine serum albumin
CBP	creb binding protein
CD	circular dichroism
C/H1	cysteine/histidine rich region 1 in p300/CBP
C/H2	cysteine/histidine rich region 2 in p300/CBP

C/H3	cysteine/histidine rich region 3 in p300/CBP
CR	conserved region of Myb proteins
dH <sub>2</sub> O	distilled de-ionised water
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTT	dithiotreitol
EDTA	ethylene-diamine-tetra-acetic-acid
<i>E. coli</i>	<i>Escherichia coli</i>
GST	glutathione S-transferase
IPTG	isopropyl β-D thiogalactopyranoside
kDa	kilo Dalton
KID	kinase-inducible domain in p300/CBP
LB	Luria Bertani
LDS	lithium dodecyl sulphate
MBS	myb binding site
Ni-NTA	nickel-nitriloacetic acid
NMR	nuclear magnetic resonance
NRD	negative regulatory domain of Myb proteins
OD <sub>600</sub>	optical density at 600 nm wavelength
PARP	poly (ADP-ribose) polymerase
PCR	polymerase chain reaction
pET	pET28(+)
<i>Pfu</i>	<i>Pyrococcus furiosus</i>
pGEX	pGEX-6P-2
PMSF	phenylmethylsulfonyl fluoride
R1	repeat 1 from the Myb DNA binding domain
R2	repeat 2 from the Myb DNA binding domain
R3	repeat 3 from the Myb DNA binding domain
RNA	ribonucleic acid
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
SV40	simian virus 40
TAD	transactivation domain
TAE	tris-acetate

Tris	tris-(hydroxymethyl)-methylamine
v/v	volume per volume
w/v	weight per volume

# CHAPTER 1

## INTRODUCTION

### 1.1 *Transcription factors in eukaryotic gene regulation*

Large regions of the eukaryotic genome reside in a transcriptionally 'silent' state due to the inhibitory constraints imposed by the high-order packing of DNA in chromatin. Chromatin can be subdivided into two physically and functionally distinct compartments, heterochromatin and euchromatin. Heterochromatin represents a highly condensed state of chromatin that results in silenced genomic regions, whereas euchromatin is less condensed and actively transcribed (Hecht *et al.*, 1995; Ding *et al.*, 1997; Grunstein, 1998). Recent studies suggest that gene activation requires the relocation of a gene locus from one nuclear territory to another (Francastel *et al.*, 2001; Sawado *et al.*, 2001). The nucleosome represents the minimal unit of chromatin and consists of approximately 146 base pairs of DNA surrounding an octamer of histone proteins composed of a (H3-H4)<sub>2</sub> heterotetramer and two H2A-H2B dimers. The restrictive environment of nucleosomal DNA must be modified to allow transcription to begin at target enhancer and promoter regions. Consequently a major control point in eukaryotic gene expression is at the level of initiation and requires the coordinated regulation of multisubunit transcriptional assemblies containing RNA polymerase.

In order for accurate promoter targeting, the core eukaryotic RNA polymerase requires the assistance of several other proteins and protein complexes that together form the pre-initiation complex (PIC). The PIC, which contains more than 80 polypeptides, often binds to a conserved DNA sequence motif called the TATA box situated 20 to 30 nucleotides upstream of the transcription initiation site. TATA-binding protein (TBP) and several TBP-associated factors (TAF's) help nucleate the assembly of the PIC at specific core promoter elements including the TATA box, as well as the initiator element (Inr) and the downstream promoter element (DPE) elements (Butler and Kadonaga, 2002). Other components of the PIC termed the general transcription factors (GTF's) include TFIIA, TFIIB, TFIID, TFIIE, TFIIIF, and TFIIH that collectively function in promoter melting and in the transition from transcription initiation to elongation (Roeder, 1996; Orphanides and Reinberg, 2002). Several studies have identified tissue-specific homologues of the PIC including additional members of TBP and TAF's (Gross *et al.*, 1998; Losick, 1998) that clearly indicate that PIC

composition is not invariant when compared amongst different cell-types within an organism. There are two current theories for the assembly of the PIC at target gene promoters (Lemon and Tjian, 2000) which include a model where the PIC pre-assembles before being recruited by gene specific activators or alternatively transcription factor complexes recruit components of the PIC in a staged-assembly process. The second model is supported by recent kinetic studies of PIC assembly using chromatin immunoprecipitation (CHIP) assays, suggesting the process is likely to be highly regulated (Chen *et al.*, 2002; Dundr *et al.*, 2002).

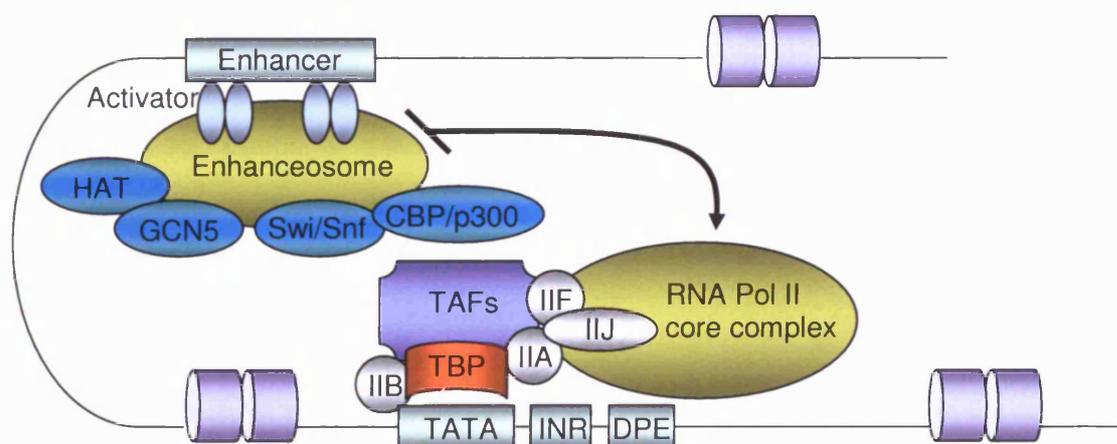
Although RNA polymerase II and the GTF's are sufficient to mediate basal transcription further sequence specific activators are required to mediate high levels of regulated transcriptional activity. Such sequence specific transcriptional activators bind to cognate DNA enhancer elements often found up to several kilobases from the core promoter. Transcriptional activators are usually composed of two essential regions; a DNA binding domain that recognises specific DNA sequences and one or more activation domains that aid the assembly of multi-protein transcriptional complexes. The DNA binding domains commonly fall into groups defined by related motifs such as the basic helix-loop-helix (HLH) motif, the basic leucine zipper (LZ) motif, the zinc finger (ZF) motif, and the helix-turn-helix (HTH) motif. Activation domains are often characterised by motifs rich in glutamine or prolines, or rich in acidic amino acids and are regulated by post-translational events such as phosphorylation, or by binding specific ligands that influence the rate of transcriptional initiation.

In order to fully augment transcriptional initiation, enhancer bound activators recruit co-activator proteins such as the ATP-dependent nucleosome remodelling factors (NURF's) SWI/SNF, and histone acetyl transferase (HAT) proteins such as p300/CBP. Such co-activators are thought to enhance gene activation by several means:

- (1) By chromatin targeted events such as the acetylation of core histone proteins within nucleosomes (by HAT's) or the removal of restrictive nucleosomal structures on DNA (by NURF's)
  
- (2) Recruiting and modulating the activity of other transcription factors by direct interaction and covalent modifications

(3) By forming a bridge between enhancer bound transcriptional activators and the promoter bound PIC. The p300/CBP coactivator contains several conserved domains that provide both structured scaffolds (TAZ1 and TAZ2 domains) and unstructured regions (SID domain) for protein interactions. The modular organisation of p300/CBP allows the potential for assembly of multi-component transcription factor complexes.

Commonly found motifs within co-activator proteins include those responsible for intrinsic DNA-stimulated ATPase activity (found in the SWI/SNF factors), histone acetyl transferase activity (HAT) (as found in GCN5, SAGA, ADA, p300/CBP and TAF proteins) and bromodomains (found in p/CAF, p300/CBP and TAF proteins). Bromodomains are thought to function by recognising and interacting with acetylated lysine residues on histone proteins (Dhalluin *et al.*, 1999; Jacobson *et al.*, 2000). Figure 1.1 summarises a current model of the proteins and protein complexes involved in eukaryotic transcription initiation.



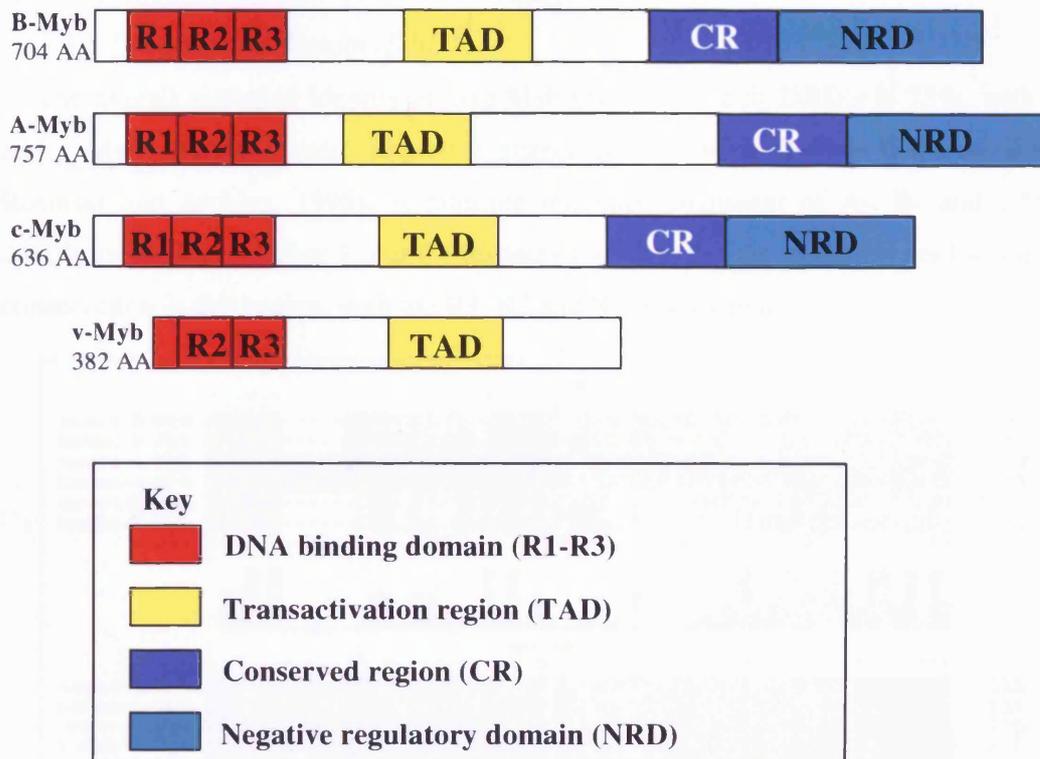
**Figure 1.1.** A schematic representation of the protein and protein complexes that are required for transcription initiation in eukaryotes (Adapted from Taajtes *et al.*, 2004). The pre-initiation complex (PIC) is primarily composed of RNA polymerase II and associated general transcription factors (GTF's) that together assemble on promoter elements including the TATA box, the initiator (INR) element and the downstream promoter element (DPE). The enhanceosome represents a collection of co-activators such as CBP/p300 and SWI/SNF that function to remodel chromatin structure and mediate signals between enhancer-bound factors and the transcriptional machinery.

Although the individual functions of the PIC, transcriptional activators and coactivators have been well characterised, the order with which these components operate to initiate gene activation are as yet unclear. In some instances transcription initiation can be triggered by the DNA-binding activator, followed by the sequential recruitment of the coactivator protein and pre-initiation complex (PIC) to the promoter (Featherstone, 2002). The IFN- $\beta$  promoter is the best understood example of such gene regulation, where four transcription factors (NF- $\kappa$ b, IRF's, ATF-2/c-Jun, and HMG1Y) bind cooperatively to enhancer DNA to form an enhanceosome (Merika and Thanos, 2001). The enhanceosome then recruits HATs and the SWI/SNF complex (Agalioti *et al.*, 2002) which modifies the nucleosomes and allows the binding of TBP and the assembly of the PIC (Lomvardas and Thanos, 2001). However other modes of gene activation exist where HAT recruitment and remodelling occur well after PIC assembly (Soutoglou and Talianidis, 2002) or where chromatin opening does not require HAT activity or nucleosome remodellers (Cirillo *et al.*, 2002) suggesting the mechanism of gene activation is not invariant.

B-Myb is an example of a transcriptional activator and is critical to the functioning of the cell cycle. It belongs to a broader family of Myb transcriptional activators that share a similar domain organisation but have varied cellular functions.

### 1.2 *The Myb family of transcription factors*

The Myb family of eukaryotic transcriptional activators are implicated in the control of cell growth, differentiation, and cancer (Sala and Watson, 1999). The highly conserved family consists of three vertebrates members, A-Myb, B-Myb and c-Myb, each of which are single chain polypeptides of a similar size (751, 704, and 636 amino acids respectively). The first of the identified vertebrates Myb proteins, c-Myb, is the cellular progenitor of the v-myb oncogene carried by the chicken retroviruses AMV and E26 (Roussel *et al.*, 1979; Klempnauer *et al.*, 1982). A Myb and B-Myb were later discovered after screening human cDNA libraries with a c-Myb probe (Nomura *et al.*, 1988). Figure 1.2 shows a representation of the functional domains found within the Myb family.



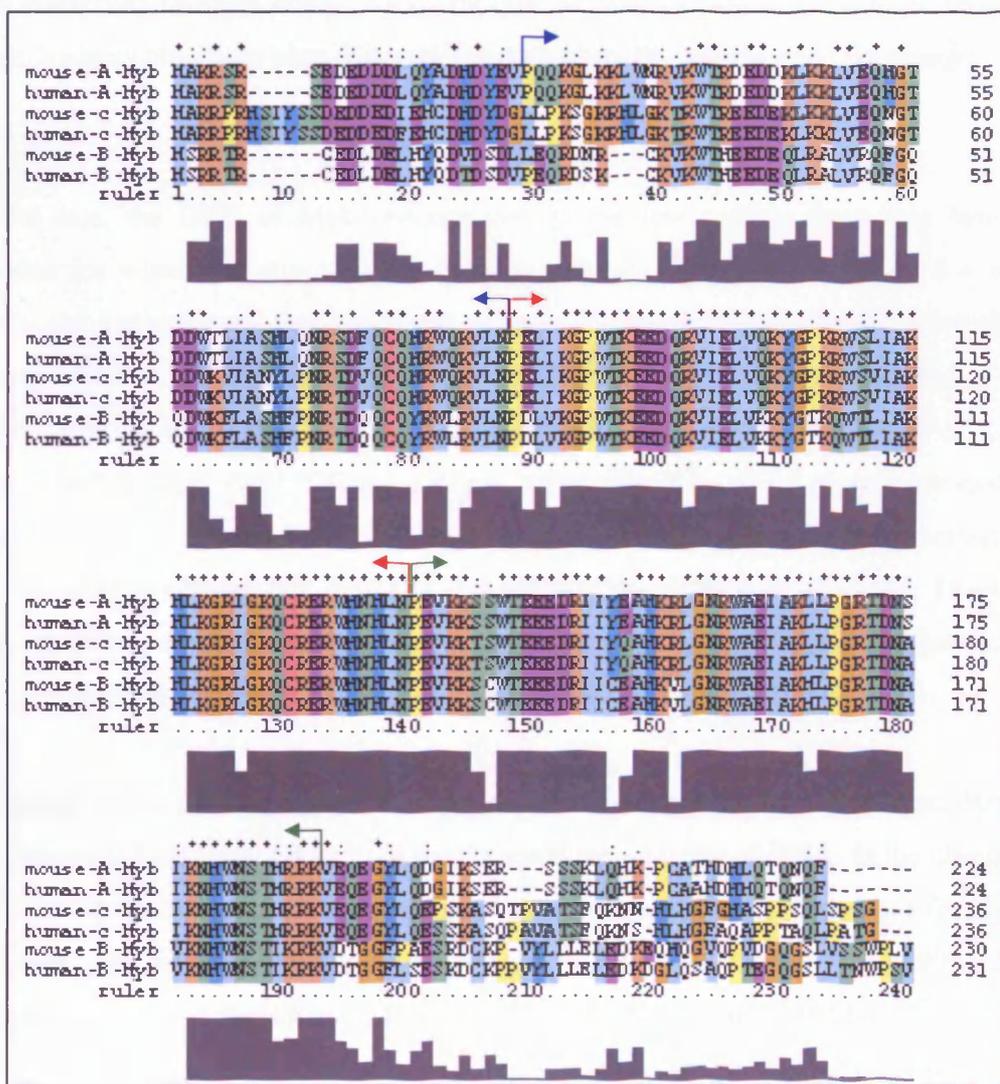
**Figure 1.2.** A schematic representation of the functional domains within the Myb family of transcription factors. The R1, R2 and R3 regions represent the 51-52 amino acid tandem repeats that make up the highly conserved DNA binding domain. The v-Myb transcription factor is truncated at both the C-terminus and in the R1 repeat, relative to its cellular progenitor c-Myb.

Each protein shares a characteristic domain organisation, with an N-terminal DNA binding domain (DBD) representing the most conserved region. In the region outside the DBD's, namely the central transactivation region and the C-terminal regulatory domain, the proteins are much more divergent. The Myb proteins are functionally distinct, in part due to their tissue specific patterns of gene expression. The transactivation potential of each protein varies with respect to cell type, and is influenced by the presence of interacting cell-type specific co-factors.

### 1.3 Functional regions/domains of the Myb family

#### 1.3.1 DNA binding domain (DBD)

The overall sequence identity among Myb proteins in their DBD's is 75%, with A- and c-Myb sharing greater sequence similarity than with B-Myb (Lipsick, 1996; Rosinski and Atchley, 1998). A multiple sequence alignment of A-, B- and c-Myb DBD's is shown in figure 1.3 and illustrates the high degree of amino acid sequence conservation in this region, with the R1, R2 and R3 repeats indicated.

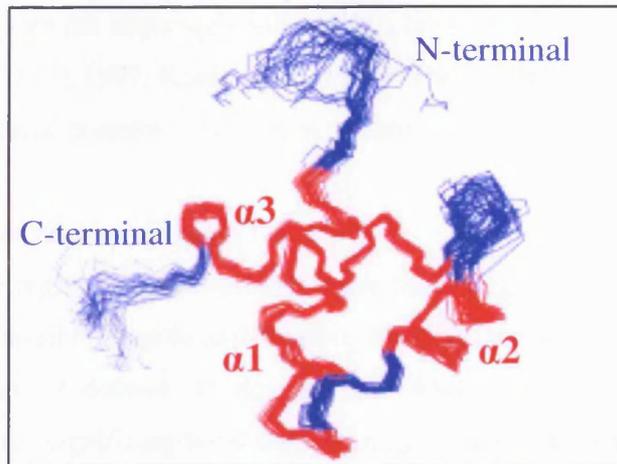


**Figure 1.3.** Amino acid sequence conservation of the A-Myb, B-Myb and c-Myb DNA binding regions. The alignment begins with the N-terminal residue (Methionine) for each of the Myb sequences, with the corresponding amino acid numbers indicated to the right of the alignment. The sequences encompassed by the blue, red, and green arrows represent the R1 (residues 23-78), R2 (residues 79-130) and R3 (residues 131-183)

repeats from the DNA binding domain of B-Myb. The multiple sequence alignment was produced using CLUSTAL X (Thompson *et al.*, 1997), with the extent of sequence conservation at each amino acid position indicated by the height of the bar in the histogram shown below the aligned sequences. The symbols denote the following: (\*) absolutely conserved residue, (:) limited to one of a closely related group of residues such as Met, Ile, Leu and Val, and (.) preference for one of a weakly related group of residues. The amino acid residues are coloured on the basis of sequence conservation, with conserved residues colour coded by type as follows: acidic pink, basic red, polar green, hydrophobic light blue, His and Tyr dark blue, Pro yellow and Gly orange.

To date, the DBD of Myb remains one of the few regions from this family of proteins for which the structure has been determined (displayed in figure 1.4 for B-Myb), and the structural and functional features of the individual domains investigated in detail (Carr *et al.*, 1996; McIntosh *et al.*, 1998). The Myb DBD is composed of three tandem repeats of 51 or 52 amino acids termed R1, R2 and R3 (Sakura *et al.*, 1989; Weston and Bishop, 1989; Saikumar *et al.*, 1990). The R2 and R3 repeats are essential for DNA binding (Howe *et al.*, 1990; Saikumar *et al.*, 1990) and specifically recognise a consensus Myb binding site (MBS) C/TAACNG (Biedenkapp *et al.*, 1988; Howe and Watson, 1991), while the R1 repeat appears to form non-specific stabilising interactions with target DNA sequences (Tanikawa *et al.*, 1993; Ogata *et al.*, 1994).

Despite the overall high sequence conservation amongst the R1-R3 repeats, dynamic differences in conformation exist in the presence and absence of DNA. In the absence of DNA R1 and R3 repeats are tightly folded in solution (shown in figure 1.4 for B-Myb R3) with each forming three  $\alpha$ -helices (Ogata *et al.*, 1992; Ogata *et al.*, 1995). R2 is comparatively less ordered in c-Myb (Myrset *et al.*, 1993) and in B-Myb (Jamin *et al.*, 1993; Myrset *et al.*, 1993; Carr *et al.*, 1996; McIntosh *et al.*, 1998). R2 does not adopt a stable, preformed recognition helix, but generates a helix-turn-helix (HTH) motif when bound to DNA. It has been proposed that the inherent conformational instability of R2 within the DNA binding region of Myb may be a mechanism of increasing the specificity of DNA binding by reducing the affinity of the protein for non-specific DNA sequences (Carr *et al.*, 1996).



**Figure 1.4.** A three-dimensional stereoview representation of the NMR solution structure derived for the B-Myb R3 repeat from the DNA-binding domain. The three helices in R3 are highlighted in red with the N- and C-termini and turn regions highlighted in blue.

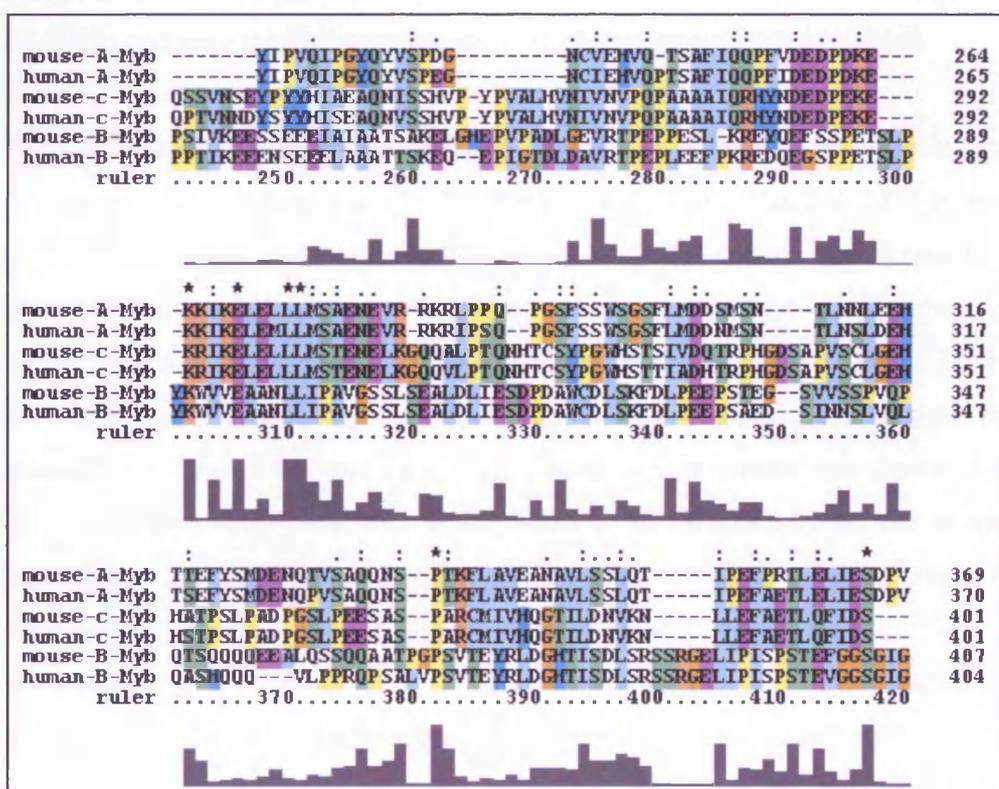
Individual Myb proteins were shown to have distinct preferences for sequences flanking the core binding site (Mizuguchi *et al.*, 1990; Howe and Watson, 1991). In particular B-Myb binds the MBS with less affinity both *in vitro* and *in vivo* than either c-Myb or A-Myb and has less tolerance to binding site variations (Bergholtz *et al.*, 2001). Interestingly, full-length B-Myb is incapable of binding the SV40 promoter MBS, yet C-terminal truncation mutants of B-Myb or B-Myb DBD fused to c-Myb were able to bind the same sequences (Watson *et al.*, 1993), suggesting the C-terminus of B-Myb is somehow imposing inhibitory constraints upon the DBD. Variations in Myb DNA binding specificity may be explained by the reported conformational instability of part of the DNA-binding motif and the proposed mechanism of DNA binding specificity this provides (Carr *et al.*, 1996).

It is of interest to note that several Myb-related proteins containing domains homologous to the Myb DBD (termed the *myb*-like domain) have been identified in mammals, yeast and drosophila (Luscher and Eisenman, 1990). Examples of such proteins include the yeast Ada-2 protein (Berger *et al.*, 1992), the telomere binding factors TRF-1 and TRF-2 (Broccoli *et al.*, 1997; Cooper *et al.*, 1997; Konig *et al.*, 1998), and the cyclin D-interacting *myb*-like protein, DMP1 (Hirai and Sherr, 1996;

Inoue and Sherr, 1998; Inoue *et al.*, 1998). The *myb*-like domain in TRF-1 and TRF-2 appears to mediate the sequence specific DNA binding of these proteins (Broccoli *et al.*, 1997; Cooper *et al.*, 1997; Konig *et al.*, 1998). The function of the *myb*-like domains in the other identified proteins is unclear at present.

### 1.3.2 Transactivation domain (TAD)

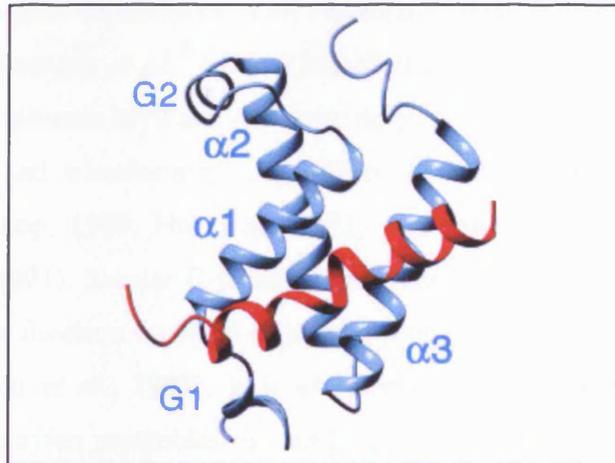
The central regions of Myb proteins have long been associated with transactivation potential (reviewed by Saville and Watson, 1998) yet the precise domain boundaries for this activity are ill defined. In this region B-Myb and A-Myb share 20% sequence identity, with no significant homology existing between the transactivation regions of B-Myb and c-Myb. A multiple sequence alignment, shown in figure 1.5, of A-, B-, and c-Myb proteins highlights the sequence diversity in this transactivation region.



**Figure 1.5.** Amino acid sequence conservation between the transactivation regions of A-Myb, B-Myb and c-Myb. A-Myb and c-Myb share greater sequence similarity in this region than with the other family member B-Myb. The individual Myb residue numbers are shown to the right of the alignment.

A characteristic feature of this region among Myb proteins is a cluster of acidic amino acids (Aspartic acid (D) and glutamic acid (E)) that are common also in other transcriptional activators (Hope and Struhl, 1986; Sakura *et al.*, 1989) For example, of the amino acids that reside between 205 and 338 in B-Myb, a high proportion of them, 26%, are acidic. The absence of any significant homology in this region of B-Myb with A-Myb and c-Myb may imply a net negative charge, rather than specific amino acid sequence appears to be important for transcriptional activation. However, contrary to this hypothesis, the deletion of 11 amino acids from within the transactivation domain of c-Myb, which had little effect on the overall charge of this region, abolished its inherent transactivation function (Weston and Bishop, 1989). A distinguishing feature of the B-Myb transactivation region which is absent in A- and c-Myb, is the high number of glutamine amino acids (6) between residues 346 and 360. Such glutamine-rich regions have been shown to function as transcriptional activators (Courey and Tjian, 1988) and may similarly contribute to B-Myb transactivation function.

The c-Myb transactivation domain had previously been mapped to a region spanning 52 residues in human c-Myb (aa 275-327) and 85 amino acids (aa 241-325) in mouse c-Myb (Sakura *et al.*, 1989; Weston and Bishop, 1989; Ibanez and Lipsick, 1990; Kalkbrenner *et al.*, 1990; Lane *et al.*, 1990). A solution structure of the mouse c-Myb transactivation domain, in complex with the CREB-binding protein (CBP) KIX domain (shown in figure 1.6), has recently been solved (Zor *et al.*, 2004). A 25 residue peptide (residues 291-315) derived from the c-Myb transactivation domain was shown to form a single  $\alpha$ -helix that binds a shallow hydrophobic groove formed by helices in the KIX domain (Zor *et al.*, 2004). The interaction formed by this short c-Myb peptide retains binding affinity for KIX similar to that of larger c-Myb peptides (Zor *et al.*, 2002), emphasising the functional relevance of this minimal interaction. It is of interest to note that the helix forming sequence deduced for c-Myb (Zor *et al.*, 2004) is almost completely conserved in A-Myb, suggesting the potential for a similar structural conformation in the A-Myb transactivation domain.



**Figure 1.6.** *The three dimensional solution structure derived for the c-Myb:KIX complex (Zor et al., 2004). The single  $\alpha$ -helix of the c-Myb transactivation domain (only the ordered parts are shown, residues 291-310) is shown in red with the backbone of the KIX domain (ordered residues 589-665) shown in blue. The KIX domain of CBP is composed of three interacting helices  $\alpha_1$  (residues 597-611),  $\alpha_2$  (residues 623-640) and  $\alpha_3$  (residues 646-664), and two short  $3_{10}$  helices  $G_1$  (residues 591-594) and  $G_2$  (617-621).*

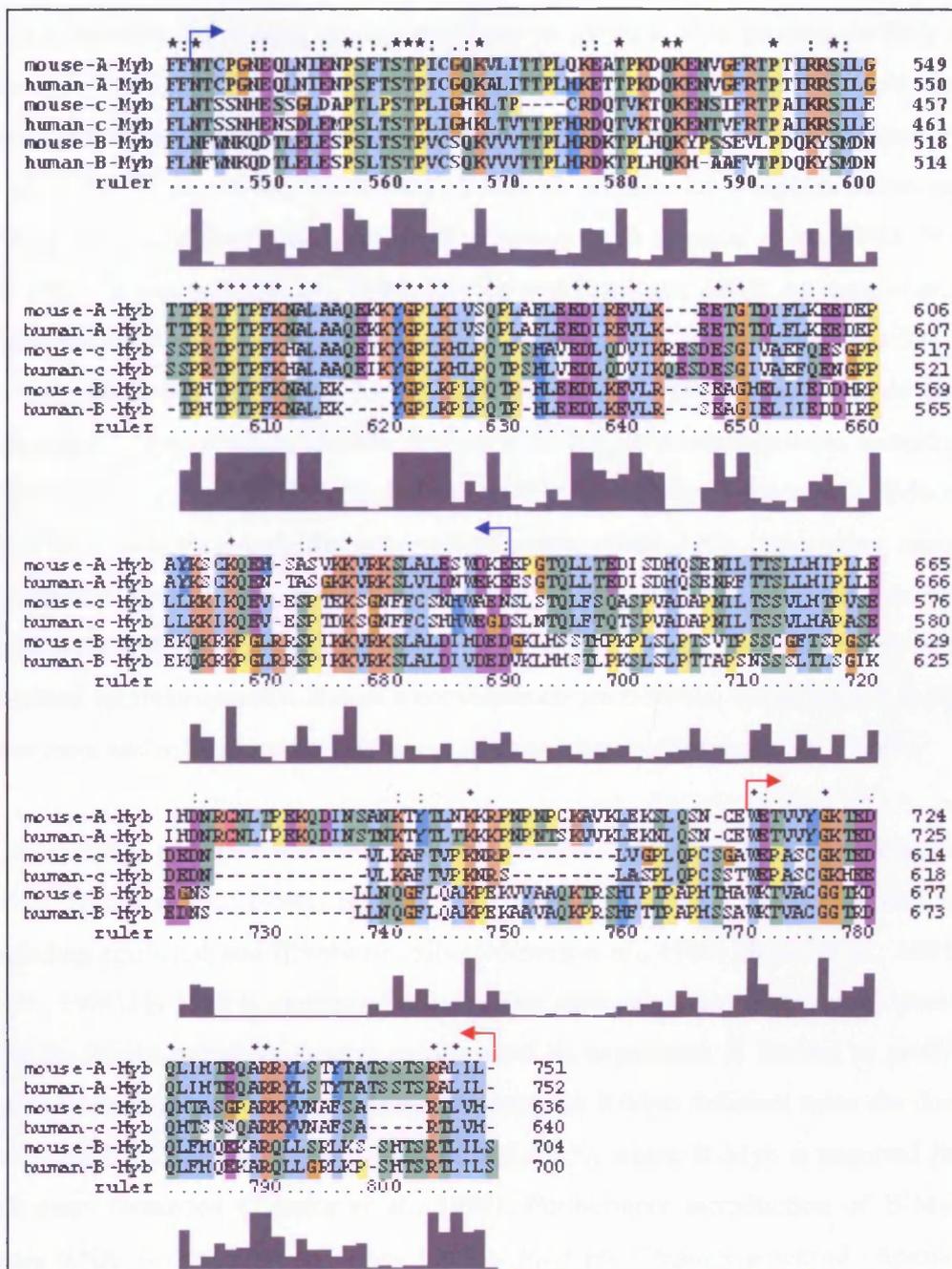
In contrast to c-Myb, the transactivation domain of B-Myb is not so well structurally defined, although multiple mapping studies have all identified similar overlapping regions. When B-Myb deletion mutants were fused to the Gal4 DNA binding domain a region comprising residues 272 to 365 was shown to transactivate promoters containing Gal4 binding sites (Ansieau *et al.*, 1997). The amino acid region identified in this study is highly conserved among vertebrate species, and overlaps a previously mapped transactivation region for B-Myb (amino acids 206-372) which was sufficient to transactivate the human c-myc promoter (Nakagoshi *et al.*, 1993). A further study using *in vitro* reporter gene assays localised the minimal B-Myb transactivation region to within residues 240-371 (Horstmann *et al.*, 2000), emphasising the functional importance of this region in conferring transcriptional activity.

### 1.3.3 Negative regulatory domain (NRD) and the conserved region (CR)

The transactivation capabilities of all vertebrates Myb proteins are modified by a C-terminal NRD (Ramsay *et al.*, 1991; Ziebold *et al.*, 1997). As its name suggests the NRD normally represses Myb activity since deletion of this domain in c-Myb increases trans-activating and transforming capabilities of the protein (Sakura *et al.*, 1989; Weston and Bishop, 1989; Hu *et al.*, 1991) and enhances its DNA binding activity (Ramsay *et al.*, 1991). Similar C-terminal deletions in B-Myb have yielded an increase in transactivation function on MBS-regulated promoters (Watson *et al.*, 1993; Lane *et al.*, 1997; Ziebold *et al.*, 1997). It is of interest to note that this domain is largely missing from the avian myeloblastosis and E26 viral encoded v-myb proteins (Roussel *et al.*, 1979; Klempnauer *et al.*, 1982).

The C-terminal regions of A-Myb, B-Myb and c-Myb share significant sequence similarity over about half their length, which is localised to two areas corresponding to residues 459 to 598 (about 45% sequence identity and 60% overall homology) and 667 to 704 (about 35% sequence identity and 55% overall homology) in B-Myb. The first of these regions of high sequence homology is known as the conserved region (CR), and has a positive regulatory role in B-Myb (Nakagoshi *et al.*, 1993; Watson *et al.*, 1993; Lane *et al.*, 1997; Oh and Reddy, 1998), whereas the CR in c-Myb and A-Myb has a negative role (Takahashi *et al.*, 1995; Oh and Reddy, 1997; Oh and Reddy, 1999). Figure 1.7 shows a multiple sequence alignment of the C-terminal region in A-, B-, and c-Myb.

The functional discrepancy between the CR's of B- and c-Myb may be explained by the structural differences in this region, namely the presence of a leucine zipper and the EVES motif found in c-Myb (Kanei-Ishii *et al.*, 1992; Dash *et al.*, 1996). Disruption of the leucine zipper (residues 375-403), by mutation of one or more leucines to proline residues, results in an increase in c-myb transactivation and transforming activity, implying a negative regulatory role for this motif in c-Myb (Kanei-Ishii *et al.*, 1992). A recent study identified the B-Myb CR as being critical for transactivation and suggested the CR is capable of co-factor binding (Tashiro *et al.*, 1995). In addition, the site of a reported dimerisation of B-Myb has been localised to the CR (Kim *et al.*, 1999).



**Figure 1.7.** Amino acid sequence conservation of the A-Myb, B-Myb and c-Myb conserved region (CR) and negative regulatory domain (NRD). The sequences encompassed by the blue and red arrows represent the CR (residues 459-598 in B-Myb) and NRD (residues 667-704 in B-Myb) regions respectively.

#### 1.4 *Myb gene expression*

A distinguishing feature among the three vertebrates Myb proteins is their distinct patterns of expression. c-Myb is predominantly expressed at high levels in immature haemopoietic cells suggesting a key role for c-Myb in blood cell development (Westin *et al.*, 1982; Golay *et al.*, 1991; Graf, 1992). Expression of c-Myb is down-regulated during terminal differentiation of these precursor cells (Gonda *et al.*, 1982; Westin *et al.*, 1982; Gonda and Metcalf, 1984; Duprey and Boettiger, 1985; Ansieau *et al.*, 1997), while constitutive expression of c-Myb in erythroleukaemia cells blocks induced differentiation (Clarke *et al.*, 1988; Todokoro *et al.*, 1988). Mutant mice lacking a functional c-Myb develop specific failure of fetal liver haematopoiesis underlining the importance of c-Myb in maintaining the proliferative state of these cells (Mucenski *et al.*, 1991). A-Myb expression is manifest in embryonic cells undergoing mitosis, the developing central nervous system, sperm cell precursors, female breast ductal epithelium and B-lymphocytes (Trauth *et al.*, 1994). Male mice lacking A-Myb have impaired spermatogenesis, and as a consequence are infertile, while female A-Myb null mice have underdeveloped breast tissue after pregnancy (Toscani *et al.*, 1997).

In contrast to the restricted tissue tropism exhibited by c-Myb and A-Myb (Graf, 1992; Trauth *et al.*, 1994), B-Myb is expressed in a wide variety of cell lineages including epithelial and fibroblast cells (Nomura *et al.*, 1988; Golay *et al.*, 1991; Reiss *et al.*, 1991). B-Myb is expressed at an earlier stage of mouse embryonic development than its family members A- and c-Myb, and its expression is limited to proliferating cells (Sitzmann *et al.*, 1996). Indeed, homozygous B-Myb deficient mice die during the early stages of murine embryogenesis (E4.5-E6.5), where B-Myb is required for inner cell mass formation (Tanaka *et al.*, 1999). Furthermore introduction of B-Myb anti-sense RNA to lymphoid cell lines inhibits their proliferation potential (Arsura *et al.*, 1992).

Further evidence for a specific role for B-Myb in cell proliferation comes from studies of B-Myb overexpression in cells. Constitutive B-Myb expression allows BALB/c 3T3 fibroblasts to grow in low serum conditions (Sala and Calabretta, 1992) and promotes progression of certain cells into S-phase (Sala *et al.*, 1996; Lane *et al.*, 1997). Such cells have increased populations in S-phase, when compared to those expressing B-Myb endogenously. A domain-swap experiment where the B-Myb DNA

binding domain was fused to a regulatable oestrogen receptor/ engrailed transcriptional repression module demonstrated B-Myb's requirement in mouse embryonic cells (Bessa *et al.*, 2001; Iwai *et al.*, 2001). In this study, the inducible dominant interfering Myb protein produced caused a G1/S phase transition defect which has been attributed to the interference of B-Myb, since this is the only member of the family which is expressed in these cells.

### 1.5 *B-Myb: a cell-cycle regulated protein*

Although ubiquitously expressed in most cell lineages, B-Myb is not a 'house-keeping gene'. Messenger RNA (mRNA) levels of both B-Myb (Golay *et al.*, 1991; Reiss *et al.*, 1991) and c-Myb (Stern and Smith, 1986; Thompson *et al.*, 1986) are growth regulated and increase at the G1/S phase boundary of the cell cycle. Furthermore, upon mitogenic stimulation of quiescent fibroblasts B-Myb expression reaches its maximum at S-phase and declines by the next G1 phase of the cell cycle (Lam *et al.*, 1992; Lam and Watson, 1993). Even in the constant presence of growth factors, B-Myb strictly conforms to this pattern of expression. In this sense, contrary to c-Myb whose transcription is determined by mitogenic stimulation, B-Myb is a true cell cycle regulated protein.

B-Myb mRNA levels were found to be undetectable in murine fibroblasts and human keratinocytes arrested by serum deprivation (Reiss *et al.*, 1991; Lam *et al.*, 1992; Lam *et al.*, 1995) and in quiescent human T and B lymphoid cells (Golay *et al.*, 1991). This study underlines how B-Myb levels are crucial in maintaining check-point control of the cell cycle. Indeed constitutive B-Myb expression prevents cell-cycle arrest in interleukin (IL)-6-induced differentiation of M1 myeloid leukaemia cells (Bies *et al.*, 1996) and was found to bypass a proliferation block induced by p53/p21 Cip1 in a human glioblastoma cell line (Lin *et al.*, 1994). Interestingly in the latter study this ability did not require B-Myb transcriptional activity. This finding is further supported in a recent study demonstrating a transcriptionally defective B-Myb mutant is able to overcome a G1 cell cycle block imposed by p107 (Joaquin *et al.*, 2002).

### 1.6 *Regulation of B-Myb gene expression*

B-Myb expression levels are controlled at the level of transcription initiation, rather than premature transcription termination (Lam and Watson, 1993). Both mouse and

human B-Myb promoters are GC rich and lack an obvious TATA box. Consistent with this type of promoter, there are multiple initiation sites for transcription, spanning 80 nucleotides in the mouse promoter (Lam and Watson, 1993). In addition, there exists an E2F transcription factor binding site, which is perfectly conserved in both the mouse and human B-Myb promoters and is critical for cell-cycle regulation of transcription (Lam and Watson, 1993; Lam *et al.*, 1995). Mutation of the B-Myb promoter E2F binding site results in constitutively high levels of B-Myb expression.

Together with Retinoblastoma-like (Rb) proteins, p107 and p130, E2F forms transcriptionally repressive complexes responsible for inhibiting activation of genes essential for entry into and progress through S-phase. In fact the first promoter identified as a target of such control was that for B-Myb, based on the observation that E1A and E7 (from adenovirus and HPV virus respectively) enhance activity of B-Myb promoters by targeting the Rb family of proteins, thus disrupting their binding to E2F (Lam and Watson, 1993).

In addition to the E2F binding site, an adjacent downstream repression site (DRS) is required for B-Myb promoter repression (Bennett *et al.*, 1996). DRS site mutations result in B-Myb promoter de-repression in G<sub>0</sub> in a similar manner to that seen with mutations in the E2F site. The DRS is thought to maintain B-Myb gene repression by specifically recruiting E2F/ p107 and E2F/p130 complexes to the adjacent E2F site. Thus, the repression of B-Myb expression involves a cooperative interaction between the E2F complex and another as yet unknown factor binding to the DRS.

The overall picture of B-Myb transcription control is of one regulated by a repression-derepression mechanism. The B-Myb promoter E2F site is occupied *in vivo* in G<sub>0</sub>/ early G<sub>1</sub> when the B-Myb promoter is repressed, but unoccupied at late G<sub>1</sub>/ S phase when B-Myb is expressed (Zwicker *et al.*, 1996; Takahashi *et al.*, 2000). B-Myb transcription de-repression occurs when E1A and E7 proteins bind to the pocket domain of p107 and p130 (Lam *et al.*, 1994), thus disrupting their interactions with E2F. B-Myb is also able to mediate such an interaction with the p107 protein, thus offering a possible mechanism of gene autoregulation (Sala *et al.*, 1996; Joaquin *et al.*, 2002). In the repressed state core histones on the B-Myb promoter are under-acetylated, where in the de-repressed state, histones H3 and H4 are specifically targeted for acetylation

(Takahashi *et al.*, 2000), suggesting chromatin structure plays a fundamental role in this mechanism also.

Once B-Myb gene expression is complete, the B-Myb transcript is subject to differential splicing yielding two mRNA species (Kamano *et al.*, 1995). An exon (referred to as exon 9a) is present in the major fraction of B-Myb mRNA, but is missing from the minor B-Myb species. Both splicing variants of B-Myb are located in the nucleus and are expressed in a large number of different cell types at similar levels (Kamano *et al.*, 1995). Recent studies suggest the shorter isoform appears to inhibit transactivation mediated by full-length B-Myb, by specifically competing for MBS's (Horstmann *et al.*, 2000).

## 1.7 Regulation of B-Myb activity

### 1.7.1 Covalent modification

Unlike A-Myb and c-Myb the transcriptional activity of B-Myb is constitutively suppressed until certain stages of the cell-cycle and so is often referred to as a repressed transactivator. B-Myb's inability to activate transcription of its cognate target genes (Ansieau *et al.*, 1997; Lane *et al.*, 1997; Ziebold *et al.*, 1997) has been attributed to inhibitory constraints imparted by the C-terminal NRD. In order to release B-Myb from its constitutively repressed state, post-translational modification in the form of phosphorylation is required. Several groups have reported a slightly larger migrating form of B-Myb with an approximate molecular weight of 112 kDa appearing in late G1/S phase, as compared to the 110 kDa form prevailing in G0/ early G1 (Robinson *et al.*, 1996; Ziebold *et al.*, 1997). The electrophoretic mobility shift witnessed in G1/S phase, is indicative of a phosphorylation event, and coincides with the timing of cyclin A and cyclin E/cdk2 activation in the cell cycle. This connection was verified when B-Myb was shown to be a substrate for cyclin A and cyclin E/cdk2 in transfected COS-7, U-2 OS, and SaOS 2 cells (Sala *et al.*, 1997; Ziebold *et al.*, 1997; Saville and Watson, 1998). Furthermore, cyclin A/cdk2 mediated phosphorylation of B-Myb at late G1 to S phase results in the substantial enhancement of its transactivation function (Ansieau *et al.*, 1997; Lane *et al.*, 1997; Sala *et al.*, 1997; Ziebold *et al.*, 1997).

Through a combination of 2D tryptic peptide mapping and point mutagenesis studies, several important cdk phosphorylation sites (with the consensus Ser/Thr-Pro) were

identified in B-Myb (Ziebold *et al.*, 1997; Saville and Watson, 1998; Bartsch *et al.*, 1999; Muller-Tidow *et al.*, 2001). Specifically Threonine's 443, 447, 490, 497, 524 and Serine 581 were shown to be *in vitro* and *in vivo* cyclin A/ cdk2 substrates (Saville and Watson, 1998; Bartsch *et al.*, 1999; Muller-Tidow *et al.*, 2001) and point mutations in these residues result in a significant reduction (40-50%) in B-Myb activity. Further B-Myb phosphorylation sites have since been identified at the CR and NRD region of B-Myb that cooperatively increase B-Myb transcriptional activity (Johnson *et al.*, 1999; Johnson *et al.*, 2002).

The general consensus is that phosphorylation of a number of specific threonine and serine residues in the C-terminal regulatory region (CR and NRD) counteract the repressive function of this domain. Interestingly, NRD deleted B-Myb mutants which have inherently high transcriptional activity cannot be further activated by phosphorylation (Ansieau *et al.*, 1997; Lane *et al.*, 1997; Ziebold *et al.*, 1997; Bessa *et al.*, 2001). Several models have been proposed to explain this phenomenon, including phosphorylation induced structural changes in B-Myb and phosphorylation enhanced nuclear localisation of B-Myb, neither of which have been fully substantiated (Robinson *et al.*, 1996). Furthermore there are conflicting reports on the effect of phosphorylation on B-Myb DNA binding, with claims that such a covalent modification has no effect (Bessa *et al.*, 2001) and increases in DNA binding reported (Johnson *et al.*, 1999). Current opinion favours a model whereby phosphorylation enhances the affinity of B-Myb for a coactivator protein or reduces its affinity for a repressor protein.

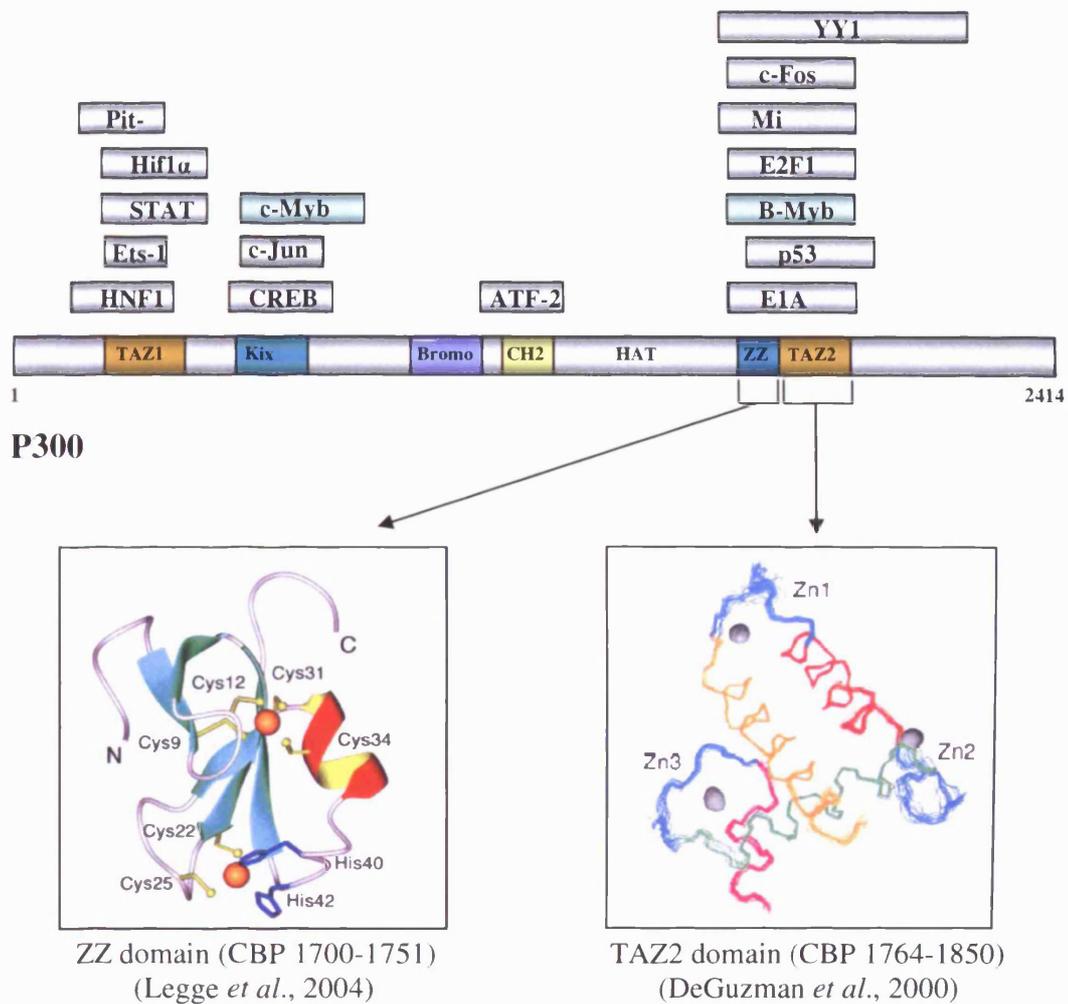
Recent studies have shown that B-Myb is subject to acetylation by the transcriptional coactivator p300 (Johnson *et al.*, 2002; Schubert *et al.*, 2004). A large central region of p300 (residues 964-1922) was shown to interact with and acetylate a central region of B-Myb containing residues 207 to 373 (Johnson *et al.*, 2002). In an independent study Schubert *et al.* report that phosphorylation of B-Myb by cyclin A/cdk2 stimulates its acetylation by p300 and that acetylation is required to disclose the full transcriptional activity of B-Myb. Although there is no evidence that p300 or CBP preferentially interacts with phosphorylated B-Myb (Bessa *et al.*, 2001; Schubert *et al.*, 2004) there is clearly cooperation of cyclin A with p300 in stimulating full activation of B-Myb.

## 1.7.2 Protein-Protein interactions

### 1.7.2.1 P300/CBP

Since a MBS is not always necessary for B-Myb transcriptional activity (Sala *et al.*, 1999; Sala and Watson, 1999) it has been argued that B-Myb could also transactivate target genes via just protein-protein interactions. As with other Myb family members B-Myb has been shown to bind the p300/CBP family of transcriptional co-activators, an interaction that results in a significant enhancement of B-Myb dependent transactivation of genes (Bessa *et al.*, 2001; Johnson *et al.*, 2002; Li and McDonnell, 2002; Schubert *et al.*, 2004).

p300 and its close relative CBP (cAMP-response element binding (CREB)-binding protein) are large, single chain polypeptides (human p300 2414 residues and human CBP 2442 residues), that contain a number of distinct functional regions (shown in figure 1.8). They both exhibit intrinsic histone acetyl transferase (HAT) activity (Bannister and Kouzarides, 1996; Ogryzko *et al.*, 1996) and function in the remodeling of chromatin by the acetylation of core histone tails within nucleosomes. This action facilitates the recruitment of RNA polymerase II and the basal transcription machinery. A second role of p300/CBP is as co-activators, which reflects their ability to form links between the basal transcription machinery and a multitude of transcription factors involved in cell-cycle control, differentiation, DNA repair and apoptosis (Janknecht and Hunter, 1996). These include CREB, c-Myb, C/EBP $\beta$ , and recently B-Myb (Chrivia *et al.*, 1993; Dai *et al.*, 1996; Schubert *et al.*, 2004) and other HAT co-factors such as pCAF, SRC-1, ACTR, and p/CIP (Yang *et al.*, 1996; Yao *et al.*, 1996; Chen *et al.*, 1997; Torchia *et al.*, 1997). Many of the proteins recruited by p300 and CBP are themselves targets for acetylation, including B-Myb (Johnson *et al.*, 2002; Schubert *et al.*, 2004).



**Figure 1.8.** A schematic representation of the p300 transcriptional coactivator with a partial list of proteins that bind the conserved domains shown. The TAZ1, KIX, Bromo, C/H2, ZZ and TAZ2 domains are indicated with the derived solution structures for the ZZ domain (Ribbon diagram, Legge *et al.*, 2004) and TAZ2 domain (Stereoview of 20 superimposed structures, De Guzman *et al.*, 2000) shown below. B-Myb is reported to interact with the C/H3 region of p300 (containing both the ZZ and TAZ2 domains) and is highlighted with c-Myb in green.

The presence of a number of highly conserved amino acid sequence motifs within p300/CBP has led to the identification and characterisation of several protein-binding regions which are depicted in figure 1.8. In particular, three cysteine/ histidine rich regions denoted C/H1, C/H2 and C/H3 (Borrow *et al.*, 1996) have been established as key protein interacting sites. The C/H2 region contains a highly conserved putative

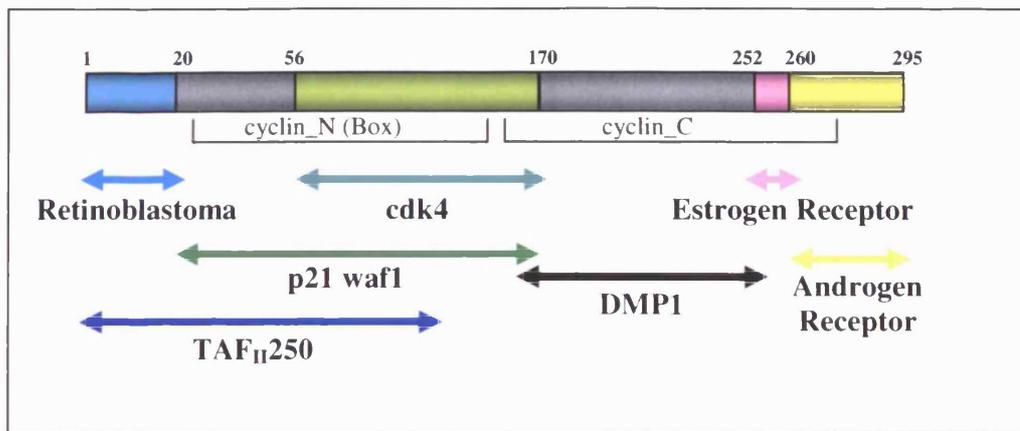
plant homeodomain type zinc finger (PHD), which forms an integral part of the enzymatic HAT domain (Kalkhoven *et al.*, 2002). C/H1 and C/H3 (spanning residues 332 to 416 and 1664 to 1806 in human p300 respectively) are distinct  $Zn^{2+}$  binding domains, which contain conserved cysteine and histidine residues essential for  $Zn^{2+}$  ion coordination. Other regions of p300 capable of protein interaction include the KIX domain (spanning residues 566 to 646 in human p300), which binds the phosphorylated kinase-inducible domain (KID) of the CREB protein (Radhakrishnan *et al.*, 1997) and the acidic transactivation domains of both c-Myb and A-Myb (Dai *et al.*, 1996; Oelgeschlager *et al.*, 1996; Facchinetti *et al.*, 1997). The bromodomain (spanning residues 1054 to 1144 in human p300) mediates the interaction of CBP with acetylated proteins such as MyoD (Polesskaya *et al.*, 2001).

There is strong evidence that the acidic transactivation region of B-Myb is responsible for binding CBP/p300, with the interacting region mapped in several independent studies to residues 197 to 340 (Li and McDonnell, 2002), 207 to 373 (Johnson *et al.*, 2002), and 240 to 371 in B-Myb (Schubert *et al.*, 2004). The latter study had mapped the B-Myb binding site to within the C/H3 region of p300, a prevalent protein interaction site (Schubert *et al.*, 2004). The C/H3 region can be subdivided into smaller sequence motifs designated the ZZ and TAZ2 regions respectively (Ponting *et al.*, 1996). The TAZ2 domain along with TAZ1 (C/H1 region) was identified on the basis of sequence alignments that revealed three, near perfect Cys-X<sub>4</sub>-Cys-X<sub>8</sub>-His-X<sub>3</sub>-Cys repeats. The solution structure determined recently for the TAZ2 domain from CBP revealed a novel helical fold stabilised by the binding of three  $Zn^{2+}$  ions to HCCC motifs (DeGuzman *et al.*, 2000, shown in figure 1.8). The ZZ region (spanning residues 1664 to 1705 within human p300) is typified by the presence of a repeated Cys-X<sub>2</sub>-Cys motif and a so-called Asp-Tyr-Asp-Leu signature sequence. A ZZ motif apparent in dystrophin and dystrophin related proteins such as utrophin, forms part of a cysteine-rich region necessary to mediate interactions with the transmembrane glycoprotein,  $\beta$ -dystroglycan (Suzuki *et al.*, 1994). The solution structure determined recently for the CBP ZZ domain reveals a folded domain that binds two  $Zn^{2+}$  ions; one  $Zn^{2+}$  ion is coordinated tetrahedrally via two CXXC motifs and the second is coordinated tetrahedrally by a third CXXC motif together with a HXH motif (Legge *et al.*, 2004, shown in figure 1.8).

### 1.7.2.2 cyclin D1

In contrast to cyclin A and p300/CBP, cyclin D1 inhibits the function of B-Myb in a cdk-independent manner by directly interacting with B-Myb (Horstmann *et al.*, 2000; Schubert *et al.*, 2004). Cyclin D1 inhibits the activity of B-Myb by interfering with the interaction of B-Myb with p300 (Schubert *et al.*, 2004). This inhibitory interaction is not accompanied by changes in the phosphorylation state of B-Myb, but does result in an increased stability of B-Myb in S-phase (Horstmann *et al.*, 2000). The binding of cyclin D1 to the central region of B-Myb (amino acids 227-508) is thought to result in an accumulation of inactive B-Myb at the G1/S phase transition, with the degradation of cyclin D1 marking the point at which B-Myb escapes from its inhibitory constraints. This theory is supported by studies that show an enhancement of B-Myb transcriptional activity during the early stages of neural differentiation, which coincides with a reduced amount of associated cyclin D1 (Cesi *et al.*, 2002).

Cyclin D1 is a 295 amino acid single chain polypeptide consisting of an N-terminal cyclin\_N domain (residues 26-153) and a C-terminal cyclin\_C domain (residues 155-283, depicted in figure 1.9). The cyclin box (spanning residues 56-152 AA) found in the cyclin\_N domain of cyclin D1, is a domain of about 100 amino acids that is conserved in the different cyclins and is responsible for binding the cell-cycle kinase cdk4 (Noble *et al.*, 1997), the cell-cycle inhibitor p21waf1 (Zwicker *et al.*, 1999), and transcription coactivator TAF<sub>II</sub>250 (Adnane *et al.*, 1999). Cyclins A, B, D and E share approximately 20-30 % sequence identity with each other in the cyclin box region. The cyclin\_C domain of cyclin D1 has been linked with transcriptional co-factor binding and regulation (Zwijssen *et al.*, 1998). Binding partners include the estrogen receptor (ER) (Zwijssen *et al.*, 1997), the DMP1 (Cyclin D-interacting *myb*-like protein) family of transcription factors (Inoue *et al.*, 1998) and the androgen receptor (Reutens *et al.*, 2001). Cyclin D1 is divergent to other D-type cyclins in the cyclin\_C domain due to the presence of an acidic region (residues 272-280) and a LLXXXL motif (residues 254-259) required for binding the transcriptional coactivator NcoA/SRC1a (Zwijssen *et al.*, 1998). The B-Myb binding site on cyclin D1 has not been localised.



**Figure 1.9.** A schematic representation of the interacting domains of Cyclin D1. A partial list of interacting proteins including Retinoblastoma (Rb), Androgen Receptor (AR) and cyclin D1 interacting Myb-like protein 1 (DMP1) are shown alongside their respective binding sites. Adapted from (Coqueret, 2002).

Although a high resolution structure of cyclin D1 is unavailable at present, a crystal structure of the homologous cyclin A2 protein in isolation and in complex with cdk2 has been solved (Brown *et al.*, 1995; Jeffrey *et al.*, 1995). The two respective studies have revealed that a cyclin A2 fragment (residues 171-432 in human cyclin A2) is composed of 12  $\alpha$ -helices arranged in two domains and adopts a similar structured scaffold both in the absence and presence of its catalytic co-factor cdk2. The first domain corresponds to the cyclin box which is composed of 5  $\alpha$ -helices, with the C-terminal domain adopting a similar fold despite negligible sequence identity (13 % with the cyclin box). An  $\alpha$ -helical secondary structure content of 63% was calculated based on the cyclin A structure solved in isolation (Brown *et al.*, 1995). The authors have acknowledged that despite a divergence in sequence, cyclin D1 and other cyclins may adopt a similar cyclin Box topology.

### 1.7.2.3 Other B-Myb interacting proteins

Over the past five years there have been numerous B-Myb co-factors identified, in addition to p300 and Cyclin D1, that modulate B-Myb activity both positively and negatively. Recently discovered co-activators of B-Myb include poly (ADP-ribose) polymerase (PARP) (Cervellera and Sala, 2000), a nuclear protein of 110 kDa found associated with chromatin, which functions by transferring poly (ADP)-ribose units to acceptor proteins. PARP interacts with the B-Myb DBD, an association that does not

require PARP's Poly ADP ribosylation activity (Cervellera and Sala, 2000). A recent report identified the TATA box-binding protein (TBP) associated factor, TAF<sub>II</sub>250 as a B-Myb associating protein which enhances B-Myb mediated transactivation (Bartusel and Klempnauer, 2003). TAF<sub>II</sub>250, the largest subunit of the TFIID complex, is a major component of the basal transcription machinery and plays a crucial role in modulating both chromatin structure and associated TAF's through its intrinsic HAT activity (Albright and Tjian, 2000; Wassarman and Sauer, 2001). *In vivo* co-precipitation studies showed the B-Myb DBD is sufficient for complex formation with TAF<sub>II</sub>250 (Bartusel and Klempnauer, 2003). However these findings were not corroborated by protein interaction studies carried out *in vitro*. Therefore it is not known whether B-Myb and TAF<sub>II</sub>250 are direct interaction partners, or whether they associate indirectly through an intermediate protein.

The nuclear co-repressors N-CoR1/SMRT have recently been implicated as regulators of B-Myb transcriptional activity (Horlein *et al.*, 1995; Li and McDonnell, 2002). N-CoR1 and SMRT are homologous co-repressors that associate with nuclear receptors and a whole host of unrelated transcription factors, enabling the direct repression of the basal transcription of their respective target genes (Asahara *et al.*, 1999; Bailey *et al.*, 1999). N-CoR1 and SMRT appear to maintain B-Myb in a transcriptionally inactive state in target cells through interaction with its NRD and by recruiting histone deacetylases (HDAC's) (Li and McDonnell, 2002). A truncated B-Myb lacking the C-terminal NRD was unable to bind N-CoR1 (Li and McDonnell, 2002), and exhibited constitutive transcriptional activity. The minimal requirement for N-CoR1 binding is B-Myb residues 198-704, which includes the TAD, CR and NRD, implying the NRD itself does not comprise the complete corepressor binding domain.

To summarise B-Myb activity appears to be co-ordinately regulated by interaction with co-activators such as p300/CBP, TAF<sub>II</sub>250, and PARP in conjunction with the opposing effects of the cyclins A and D1. The nuclear co-repressors N-CoR1/SMRT binding to the NRD provide an explanation as to why under most circumstances B-Myb is unable to activate transcription of its cognate target genes. Phosphorylation and acetylation of B-Myb leads to the ablation of such repression and induces a marked increase in B-Myb activity.

### 1.7.3 Degradation of B-Myb

An alternative means of regulating B-Myb activity is through the elimination of the protein through degradation pathways. Indeed many transcription factors are short-lived and are largely degraded through the ubiquitin-proteasome pathway (Elledge and Harper, 1998; Desterro *et al.*, 2000). B-Myb has a protein half-life of 2.7 hours which is significantly reduced to 50 minutes when cyclin A is co-expressed in mammalian cells (Saville and Watson, 1998; Charrasse *et al.*, 2000). Thus, as well as enhancing B-Myb activity, phosphorylation induced by cyclin A appears to target B-Myb for ubiquitination. This model ensures that upon phosphorylation B-Myb is only transiently hyperactivated in S-phase before being directed to the proteasome.

### 1.8 B-Myb regulated genes

Several genes have been shown to be regulated at the transcription level by B-Myb. The mode of B-Myb dependent gene transactivation is somewhat varied, being in some circumstances dependent on DNA binding, and in others clearly affected by post-translational modification and interaction with potential co-activators or co-repressors. An example of an MBS-independent regulated gene is the human HSP70 gene (Foos *et al.*, 1993), which exhibits maximal expression during S-phase coinciding with that for B-Myb (Lam *et al.*, 1992; Lam *et al.*, 1995). Transactivation of HSP70 appears to be dependent on the presence of a TATA box, which implies B-Myb regulates its expression through direct or indirect interactions with the basal transcription machinery (Foos *et al.*, 1993). Further MBS-independent regulated genes include DNA polymerase  $\alpha$  (Watson *et al.*, 1993), FGF-4 (Johnson *et al.*, 2002), and B-Myb itself (Nakagoshi *et al.*, 1993). B-Myb synergises with SPI at SPI sites on the B-Myb promoter resulting in B-Myb expression (Sala *et al.*, 1999). This has led to speculation that B-Myb and SPI operate in tandem to activate the numerous genes that contain SPI sites at their promoters.

Genes which are activated by B-Myb (and c-Myb) in an MBS-dependent manner include c-myc (Nakagoshi *et al.*, 1992) and human cyclin A1 (Muller-Tidow *et al.*, 2001) which both contain several MBS's in their respective promoters. Both human DNA topoisomerase II $\alpha$  (Brandt *et al.*, 1997) and insulin-like growth factor binding protein 5 (IGFBP) are subject to B-Myb mediated gene transactivation by a single MBS site on their promoters. In general, the genes identified as B-Myb targets are intimately

involved in cell proliferation though two recently identified gene targets apolipoprotein J/ clusterin (Cervellera *et al.*, 2000; Trougakos and Gonos, 2002) and Bcl-2 (Grassilli *et al.*, 1999) have prompted speculation that B-Myb may also enhance cell survival. It must be noted however that the majority of these genes have yet to be substantiated as targets of B-Myb *in vivo*. The data relate primarily to studies using transient co-transfection/ reporter assays rather than promoters of the endogenous genes.

### 1.9 B-Myb in tumourigenesis

B-Myb is expressed at high levels in a number of tumorigenic cell lines (Nomura *et al.*, 1988; Raschella *et al.*, 1995) and is thought to maintain the proliferative state of such cells. Expression levels of B-Myb are down regulated when neuroblastoma cell lines are differentiating (Raschella *et al.*, 1995). However constitutive expression of a B-Myb transgene bypasses the differentiation block and promotes the growth and proliferation of neuroblastoma cells. The finding that B-Myb regulates expression of apolipoprotein J/ clusterin, further supports the notion that B-Myb confers proliferative or survival advantage to cancer cells. Apolipoprotein J/ clusterin actively promotes the survival of neuroblastoma cells exposed to the chemotherapeutic drug doxorubicin (Cervellera *et al.*, 2000). Interestingly high levels of B-Myb expression are associated with a poor prognosis for cancer patients, whereas patients with tumours displaying no B-myb expression have greater long-term survival rates. A recent tissue-wide expression profiling study of tumour-specific genes showed that B-Myb was upregulated in cancer patients with short survival time (Amatschek *et al.*, 2004).

### 1.10 Aims of the project

To date, the three dimensional structural characterisation of B-Myb has been limited to its highly conserved DNA binding region (B-Myb residues 79-186) (Carr *et al.*, 1996; McIntosh *et al.*, 1998). The central transactivation region has been implicated with the regulation of B-Myb transcriptional activity through interaction with cellular partner proteins, and is therefore an attractive target for structural studies. Although much work has already been done with regards to localising the transactivation domain within B-Myb, the definitive domain boundaries have yet to be determined. The main objectives of the project involved: (1) mapping the minimal sites of interaction between B-Myb and its partner proteins p300 and Cyclin D1 and (2) characterising the structural properties and features of the identified interacting regions or domains both in isolation and in complex.

(1) At the commencement of this project there were several reports of a functional interaction between B-Myb and the cell cycle protein Cyclin D1 (Horstmann *et al.*, 2000b; Cesi *et al.*, 2002), with the interaction mapped to within residues 227-508 of B-Myb. Unpublished observations had implicated the same region of B-Myb with interacting with the E1A binding region of p300 (Karl-Heinz Klempnauer). The E1A binding region of p300 is a prominent site for protein-protein interactions and contains the highly conserved ZZ and TAZ2 domains. An array of recombinant B-Myb, p300 and Cyclin D1 proteins were thus generated for use in mapping of the minimal site of interaction between B-Myb and p300/Cyclin D1. Interaction analyses of purified B-Myb, p300 and Cyclin D1 polypeptides were conducted using GST pull-down assays and protease protection assays. Using GST/B-Myb polypeptides as bait, purified p300 and Cyclin D1 polypeptides were systematically assayed for binding in GST pull-down assays. The minimal sites of interaction were thus mapped through a process of elimination. Since the formation of a specific protein-protein complex may well reduce the number of protease sensitive sites on one or both proteins, changes in the limited trypsin digest profiles of individual GST/B-Myb polypeptides in the absence and presence of a molar excess of p300 polypeptide were used to monitor the extent of protease protection. This highly sensitive method allowed the detection of relatively weak protein interactions that were not reliably detected by GST pull-down assays. Where possible, protein-protein interactions were monitored by comparing the wavelength of maximum fluorescence emission ( $\lambda_{\text{max}}$ ) of an individual polypeptide with

that of the final complex. Tryptophan residues serve as useful reporters for detecting potential interactions in fluorescence binding assays where complex formation would result in measurable perturbations in fluorescence wavelength.

(2) Biophysical techniques such as intrinsic tryptophan fluorescence spectroscopy (Jiskoot *et al.*, 1995) and far UV circular dichroism (Sreerama and Woody, 1994; Sreerama *et al.*, 1999; Sreerama and Woody, 2000) were used to assess the tertiary and secondary structural properties of individual B-Myb, p300 and cyclin D1 polypeptides. In particular resistance to chemical and thermal denaturation was monitored by following changes in circular dichroism (molar CD per residue), or wavelength of maximum fluorescence emission ( $\lambda_{\text{max}}$ ) as a function of increasing temperature or guanidine hydrochloride concentration. The structural stability of identified protein complexes were similarly assessed using fluorescence based denaturation. The zinc binding properties of individual p300 polypeptides was assessed by monitoring changes in their far UV circular dichroism spectra.

## CHAPTER 2

### CLONING AND PURIFICATION OF B-MYB, P300 AND CYCLIN D1 PROTEINS

#### 2.1 *Introduction and aims*

The aims of the work described in this chapter were to produce a range of GST/B-Myb fusion proteins spanning the central region of B-Myb (227-508) implicated in the binding of the transcriptional co-activator p300 and the repressor cyclin D1 (Schubert *et al.*, 2004). Multiple sequence alignments were used to aid the design of B-Myb and p300 constructs. The panel of GST/B-Myb proteins expressed and purified were to be used in GST pull-down assays against purified His-tagged p300 and cyclin D1 proteins in an attempt to map the minimal interacting regions. The *E. coli* produced B-Myb fusion proteins contained cleavable GST tags so that any soluble and stably produced B-Myb protein could be structurally characterised using biophysical techniques such as intrinsic tryptophan fluorescence and circular dichroism (CD) spectroscopy. The same biophysical techniques will be applied to study the structural features of stably produced p300 and cyclin D1 proteins either in isolation or in complex with B-Myb.

#### 2.2 *Materials and Methods*

##### 2.2.1 *B-Myb constructs and cloning strategy*

The expression vectors for two regions of murine B-Myb, spanning amino acids 227 to 508 (lacking the B-Myb DNA binding domain and the inhibitory C-terminal region of B-Myb), and 275 to 376 cloned as N-terminal fusions to glutathione-S-transferase (GST) were provided by Karl-Heinz Klempnauer's research group in Germany.

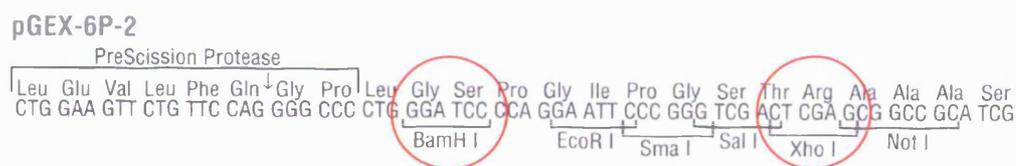
The B-Myb R2R3 DNA binding region, containing repeats 2 and 3 from the DNA binding motif of chicken B-Myb, had previously been cloned to generate a GST/B-Myb R2R3 fusion protein expression vector (Smith and Johnson, 1988). Specifically residues 79 to 186, containing close to the minimal B-Myb DNA-binding domain (Foos *et al.*, 1992) were PCR amplified and cloned into a pGEX-3X vector (Pharmacia). Further regions of murine B-Myb spanning amino acids 275 to 458, 318 to 458, 318 to 508, 369 to 458, and 369 to 508 were PCR amplified with *Pfu* DNA polymerase (Promega) using the pGST/B-Myb 227-508 vector as template, and the following synthesised primers (Proteins Nucleic Acids and Chemistry Laboratory, PNAACL- University of Leicester).

Forward (275) CGGGATCCAAGCGTGAATACCAGGAGTTC  
 Forward (318) CGGGATCCGACCCTGATGCTTGGTGCGAC  
 Forward (369) CGGGATCCCCCAGTGTGACTGAGTACCGC  
 Reverse (458) AGGGGGCCCCTACTGGGGAAGGAGAGAAGGGGAG  
 Reverse (508) AGGGGGCCCCTATACTTCTGATGATGGATACTT

The *Bam*HI and *Sma*I restriction enzyme recognition sites are underlined in the forward and reverse primer sequences respectively. The reverse primers included a stop codon to ensure translation was terminated immediately after B-Myb expression. PCR reactions were carried out in 50 µl volumes (protocol described by Promega) using a Techne (Techgene) thermal cycling system. An initial denaturation cycle for 0.5 minutes at 95 °C, was followed by thirty cycles of denaturation (0.5 minutes at 95 °C), annealing (1 minute at 60 °C), and extension (1.5 minutes at 72 °C). At the end of the final cycle there was an incubation period of 5 minutes at 72 °C (for the final extension), after which the temperature was maintained at 4 °C. PCR products were verified and quantified by 1% (w/v) agarose gel electrophoresis before recovery from gels using the GeneClean Gel Extraction Kit (Q-Bio).

Purified PCR products were 'blunt-end' ligated into the Perfectly Blunt<sup>®</sup> vector, pT7 Blue-3 (Novagen), primarily for archiving and subsequent sub-cloning. In a standard ligation reaction approximately 0.05 pmole of PCR insert was incubated with 0.02 pmole of blunt vector giving a molar ratio of insert to vector of 2.5:1. The ligation mixture was used to transform Novablue<sup>™</sup> competent cells, which were then screened for successful transformation on LB agar plates (Sambrook *et al.*, 1989) containing 150 µg/ml ampicillin. Plasmid DNA was prepared from subsequent positive transformants and successful ligation of insert verified by *Eco*RI digestion (*Eco*RI sites flank the blunt cloning site in pT7blue-3). The orientation of cloned insert was assessed by *Bam*HI restriction digests of plasmid DNA. Since a *Bam*HI site is present on both the pT7Blue-3 vector and the insert, the size of the *Bam*HI generated restriction fragment will reveal the relative orientation of the cloned insert within the pT7Blue-3 vector. The restriction digests were visualised on a 1% (w/v) agarose gel, and the presence of insert confirmed. The sequence integrity of successfully ligated pT7Blue-3 plasmids was verified by DNA sequencing (PNAFL).

Due to the poor efficiency of *Sma I* digestion of pT7Blue-3 ligated B-Myb inserts, sub-cloning into pGEX-6P2 was completed using *BamHI-Xho I* digested inserts rather than *BamHI-SmaI* as originally intended (This change had no effect on either B-Myb sequence integrity or length). pT7Blue-3 ligated B-Myb inserts to be cloned between the *BamHI* and *Xho I* sites of pGEX-6P2 (highlighted in Figure 2.1) were restriction enzyme digested with *BamHI* and *Xho I* (Promega) together, in 20 µl reaction volumes at 37 °C for 4 hours. Samples of pGEX-6P2 were digested in a similar manner. Restriction enzyme digested B-Myb inserts and vector were purified and recovered from 1% (w/v) agarose gels using the GeneClean Gel Extraction Kit (Q-Bio). Purified insert and vector were quantified by 1% (w/v) agarose gel electrophoresis by comparison to a known quantity of 500 bp marker present in the 100 bp ladder (Promega).



**Figure 2.1.** Multiple cloning site for pGEX-6P2 (Pharmacia). The B-Myb inserts were cloned between the *BamHI* and *XhoI* site allowing expression of N-terminally fused GST/B-Myb proteins. The *BamHI* and *XhoI* restriction sites are highlighted.

The ligation of B-Myb insert into pGEX-6P2 was conducted with a 3:1 ratio of insert to vector in a 10 µl reaction. Prior to the addition of T4 DNA ligase (Promega), the ligation reaction mix was heated to 42 °C for 5 minutes before being placed on ice for 5 minutes. Ligation reactions were incubated at 16 °C overnight, and then used to transform competent *E. coli* DH5α cells. Successful transformants were screened by growth on plates of LB agar containing 150 µg/ml ampicillin. Plasmid DNA was prepared from positive transformants, and the presence of insert confirmed after viewing *BamHI/XhoI* restriction digests of the plasmid on a 1 % (w/v) agarose gel. The sequence integrity of ligated insert and the flanking regions of the expression vector were verified by DNA sequencing (PNAACL).

### 2.2.2 p300 constructs and cloning strategy

Three regions of human p300 spanning amino acids 1603-1725 (containing the ZZ domain), 1603-1839 (containing the ZZ and the TAZ2 domains), and 1726-1839 (containing the TAZ2 domain) were PCR amplified with *Pfu* Polymerase (Promega) using an expression vector containing the human p300 coding sequence as the template. The template vector for human p300 (pCMV-p300CHA) containing nucleotides 1134-8329 of the published sequence (Eckner *et al.*, 1994) was provided by David Heery. Prior to PCR amplification of the desired sequences, the pCMV-p300CHA template vector was linearised and insert containing the human p300 coding sequence isolated by double restriction digestion with *Not I* and *Hind III*. Subsequent amplification from the isolated linear template as compared to the circular vector template resulted in greater PCR efficiency (Data not shown). The following primers were synthesised (PNAACL) for use in the PCR of p300 sequences.

Forward (1603) GGAATTCCATATGGGCCCTGCTGCCAACTCCCTG

Forward (1726) GGAATTCCATATGAGCCAGGCGATTCTCGCCGC

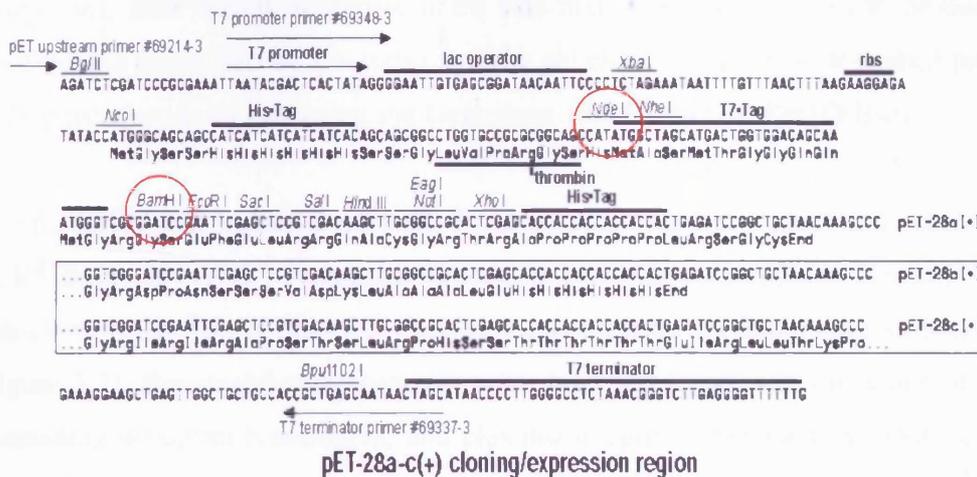
Reverse (1725) CGCGATCCCTACTGGGTGGCTGCAGCCTGCTG

Reverse (1839) CGCGATCCCTAACCAGTCCGCTGCATGCTGGC

The *Nde I* and *BamHI* restriction enzyme recognition sites are underlined in the forward and reverse primer sequences respectively. The reverse primers included a stop codon to ensure translation was terminated immediately after p300 expression. PCR reactions were carried out in 50 µl volumes using a thermal cycling sequence similar to that previously described for B-Myb amplification with the exception that an extension time of 2 minutes (at 72 °C) was allowed for p300 1603-1839 rather than the 1.5 minutes administered for the other two p300 sequences. PCR products were verified and quantified by 1% (w/v) agarose gel electrophoresis before recovery from gels using the GeneClean Gel Extraction Kit (Q-Bio).

A pET28a-based vector, expressing residues 1603-1725 and 1603-1839 from human p300 with an N-terminal hexa-histidine tag, was produced by direct sub-cloning of the appropriate PCR amplified DNA into the *NdeI* and *BamHI* sites of pET28a (Figure 2.2). The ligation of p300 insert into pET28a was completed with a similar protocol to that described for B-Myb and pGEX-6P2. Successful transformants were screened by growth on plates of LB agar containing 40 µg/ml Kanamycin. Plasmid DNA was prepared from positive transformants, and the presence of insert confirmed after viewing *NdeI/BamHI* restriction digests of the plasmid on a 1 % (w/v) agarose gel. The integrity of the ligated expression vector was confirmed by DNA sequencing (PNAACL).

The p300 1726-1839 PCR product was first 'blunt-end' cloned into a pCR<sup>®</sup>-Blunt II-TOPO<sup>®</sup> vector (Invitrogen) and the ligation mix used to transform One Shot<sup>®</sup> Competent Cells as described in the manufacturers guidelines. Successful transformants were screened by growth on plates of LB agar containing 50 µg/ml Kanamycin. The integrity of the ligated insert in the intermediate pCR<sup>®</sup>-Blunt II-TOPO<sup>®</sup> vector was confirmed by DNA sequencing (PNAACL). A subsequent *NdeI/ BamHI* double restriction digest isolated the fragment of interest which was then sub-cloned into the corresponding sites in pET28a. Again successful transformants were screened by growth on plates of LB agar containing 40 µg/ml Kanamycin, and plasmid integrity confirmed by DNA sequencing (PNAACL).



**Figure 2.2.** Multiple cloning site for pET28a (Novagen). The p300 inserts were cloned between the *NdeI* and *BamHI* site allowing expression of N-terminally fused hexahistidine tagged p300 proteins. The *NdeI* and *BamHI* restriction sites are highlighted.

### 2.2.3 Cyclin D1 constructs and cloning strategy

Full-length human cyclin D1 was PCR amplified with *Pfu* Polymerase from an *XhoI* linearised pGEX-6P2-cyclin D1 template vector (a kind gift from Karl-Heinz Klempnauer). Prior linearization of the template vector containing the human cyclin D1 coding region significantly improved PCR efficiency (Data not shown). The following primers were synthesised (PNACL) for use in the PCR of the 885 bp human cyclin D1 sequence.

Forward (1) GGAATTCCATATGATGGAACACCAGCTCCTGTGC

Reverse (885) CGCGGATCCCTAGATGTCCACGTCCCGCACGTGC

The *Nde I* and *BamHI* restriction enzyme recognition sites are underlined in the forward and reverse primer sequences respectively. The reverse primer included a stop codon to ensure translation was terminated immediately after cyclin D1 expression. PCR reactions were carried out in 50 µl volumes (protocol described by Promega) using a techne (techgene) thermal cycling system. An initial denaturation cycle for 1.5 minutes at 95 °C, was followed by thirty cycles of denaturation (0.5 minutes at 95 °C), annealing (1 minute at 60 °C), and extension (2 minutes at 72 °C). At the end of the final cycle there was an incubation period of 5 minutes at 72 °C (for the final

extension), after which the temperature was maintained at 4 °C. PCR products were verified and quantified by 1% (w/v) agarose gel electrophoresis as described previously before recovery from gels using the GeneClean Gel Extraction Kit (Q-Bio).

The full-length human cyclin D1 PCR product was first 'blunt-end' cloned into a pCR<sup>®</sup>-Blunt II-TOPO<sup>®</sup> vector (as described previously for p300 1726-1839), before sub-cloning the *Nde I* / *Bam HI* excised insert into the corresponding sites in pET28a (figure 2.2). Successful transformants were screened by growth on plates of LB agar containing 40 µg/ml Kanamycin, and plasmid integrity confirmed by DNA sequencing (PNAFL).

#### 2.2.4 Agarose gel electrophoresis

Cloning progress was continuously monitored by electrophoresis using a 1% agarose gel run in TAE buffer at a constant 100 V for 45 minutes. Typically 1 µl of PCR product was combined with 9 µl of de-ionised H<sub>2</sub>O and 2 µl of 6x Blue/Orange Loading dye (Promega) for agarose gel analysis. Staining of the gels with 0.5 µg/ ml ethidium bromide allowed the DNA to be visualised under ultraviolet light, and allowed quantitation of PCR products by comparison of band intensities with that of the 500 bp marker from the 100 bp ladder (Promega). In addition a λ DNA/ Hind III Marker (Promega) was used to reference plasmid DNA. Accurate DNA yields were estimated (after DNA was extracted from gel slices) with a Cary spectrophotometer using a 0.5 cm pathlength cell. A double stranded (ds) DNA with an absorbance of 1.0 at 260 nm corresponds to 50 µg/ ml of DNA (Sambrook *et al.*, 1989).

#### 2.2.5 SDS PAGE gel electrophoresis

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS PAGE) analysis was used to monitor progress of protein expression and purity of eluted samples. Typically 20 µl of protein sample was supplemented with 10 µl Dithiothreitol (DTT) (50 mM) and 10 µl of 4x NuPAGE<sup>®</sup> LDS sample buffer (Invitrogen). Samples were then heated for 10 minutes at 70 °C and allowed to cool prior to gel loading. For timecourse expression trials the equivalent of 125 µl of original culture was taken from cell extracts to be loaded in each lane for electrophoresis. Adjustments to volumes loaded were also made, on the basis of the OD<sub>600</sub> of the cultures at the time samples

were taken. The gel loadings equate to approximately 5-10 µg of protein per well. Prepared samples were loaded onto NU-PAGE<sup>®</sup> 4-12% Bis-Tris precast gels and run in MES SDS running buffer (Invitrogen, composition 1M MES, 1M Tris Base, 69.3 mM SDS, 5 mM EDTA) at a constant current of 250 mA for approximately thirty five minutes. Typically wide range<sup>®</sup> or broad range<sup>®</sup> molecular markers (Sigma) were loaded in each gel for protein size reference. Upon completion of electrophoresis, the gels were stained with Coomassie Brilliant Blue stain (2.5g/l in 45 % v/v methanol, 10 % v/v acetic acid) to visualise proteins, and then destained in a solution containing 10% (v/v) methanol and 5% (v/v) acetic acid solution.

#### *2.2.6 Electrospray mass spectroscopy sample preparation*

Electrospray mass spectrometry (ESMS) was performed on a micromass platform2 instrument at the Protein and Nucleic Acid Chemistry Laboratory (PNACL), University of Leicester. Protein samples required desalting prior to analysis and so were first dialysed against 2 litre volumes of de-ionised water before being flash-frozen in liquid N<sub>2</sub>. Freeze dried samples were then re-dissolved in 50 µl of 0.3% (v/v) formic acid, 50% (v/v) acetonitrile to give a final protein concentration of between 1-5mg/ml.

#### *2.2.7 N-terminal sequencing*

Using a 467 N-terminal sequencer, between 1-5 mg/ml of protein sample was applied onto a prosorb membrane (applied bioscience) for N-terminal sequencing (PNACL).

#### *2.2.8 GST/ B-Myb protein expression trials*

The pGST/ B-Myb constructs were transformed into *E. coli* BL21 (DE3) cells and typically grown in 50 ml LB medium containing 100 µg/ml ampicillin. The expression of the GST fusion proteins was induced at mid-log phase (corresponding to an absorbance at 600 nm of 0.6-0.7) by the addition of isopropyl-B-D-thiogalactopyranoside (IPTG) to a final concentration of 0.45 mM. Prior to induction, a sample of whole cells was removed from each culture for SDS PAGE analysis. Post-induction samples were removed, typically after 1, 2, 3 and 4 hours. The cells were then harvested by centrifugation at 7800 g for 15 minutes (4 °C). The cell pellets obtained were resuspended and lysed with Bug-buster HT (Novagen) to which was added 0.5 mM EDTA and 100 µM phenylmethylsulfonylfluoride (PMSF) to inhibit protease

activity. After an incubation period of 20 minutes at room temperature, a whole cell lysate sample was removed for SDS PAGE gel analysis before the cell lysates were centrifuged at 12100 g for 15 minutes at 4 °C. A lysate supernatant sample was removed for SDS PAGE gel analysis.

#### *2.2.9 Expression and Purification of GST/ B-Myb fusion proteins*

As described above GST/ B-Myb constructs were transformed into *E. coli* BL21 (DE3) cells (typically 400 mls for large scale expression) and grown in LB medium containing 100 µg/ml ampicillin. Typically cell cultures were grown to four hours post-induction with 0.45 mM IPTG before being harvested by centrifugation at 7800 g for 15 minutes (4 °C). The cell pellets obtained were resuspended and lysed with Bug-buster HT (Novagen) to which was added 0.5 mM EDTA and 100 µM phenylmethylsulfonylfluoride (PMSF) to inhibit protease activity. Cell lysates were centrifuged at 12100 g for 15 minutes at 4 °C and the soluble fractions containing the GST fusion proteins were dialysed into a 20 mM Tris and 100 mM sodium chloride pH 7.0 column running buffer for purification by affinity chromatography. Purification of the proteins was carried out using a pre-equilibrated 10 ml glutathione agarose (Sigma) column using an AKTA FPLC system. Dialysate containing the protein of interest was loaded onto the column, after which unbound protein was removed by washing with 3 column volumes of column running buffer. GST fusion proteins were eluted by the application of 10 mM reduced glutathione in a 20 mM Tris, 100 mM sodium chloride buffer at pH 7.0.

B-Myb 275-376 was produced after PreScission Protease (Pharmacia) mediated cleavage of the GST/B-Myb 275-376 fusion protein. Elution fractions containing the GST-B-Myb (275-376) fusion protein were pooled and dialysed against PreScission Protease cleavage buffer, 50 mM Tris-HCl, 150 mM sodium chloride, 1 mM EDTA, 1 mM DTT, at pH 7.0. To the retrieved dialysate was added PreScission Protease (10U per mg of fusion protein), and the mixture incubated for 5 hours at 4 °C to allow cleavage of the PreScission Protease-specific linkage between the glutathione S-transferase and B-Myb 275-376. After the cleavage step, the mixture was loaded onto a 10 ml glutathione agarose (Sigma) column, and B-Myb 275-376 collected in the flow-through. The GST and PreScission Protease remain bound to the resin making this a clean and efficient purification step.

GST/B-MybR2R3 was transformed into *E. coli* BL21 (DE3) cells, induced and grown as described above. Cells were resuspended in about 10 volumes of 10 mM Tris, 50 mM sodium chloride, 1 mM EDTA, 1 mM DTT buffer at pH 8.5 and incubated on ice in the presence of lysozyme (4.4 mg/g of cell pellet). The pH was maintained at 8.5 by the addition of 1M NaOH. Cells were typically lysed after about 40 minutes, after which Triton X-100 was added to 0.5 % and magnesium chloride to 5 mM, followed by a few  $\mu\text{g}$ 's of DNase. Once the viscosity of the mixture had returned to normal, the insoluble cell debris was removed by centrifugation at 12100 g for 15 minutes at 4 °C. The soluble material was then dialysed against a 20 mM Tris and 100 mM sodium chloride column running buffer at pH 7.0 for purification by glutathione agarose affinity chromatography using the FPLC protocol described above. Purified GST/ B-Myb R2R3 was pooled and dialysed into a 25 mM sodium phosphate, 100 mM sodium chloride, 0.1 mM EDTA, and 0.5 mM DTT buffer at pH 6.5. The final step in the purification procedure involved loading the dialysed solution onto a 10 ml previously equilibrated CM-Sepharose column and washing with several column volumes to remove contaminating nucleotides. GST/B-Myb R2R3 was eluted by applying a step gradient of sodium chloride from 100 to 500 mM.

In order to prevent possible oxidation of the free SH groups present on the protein all the solutions used in the purification were thoroughly degassed.

#### 2.2.10 *His-tagged p300 protein expression trials*

The p300 1603-1725 construct was transformed into *E. coli* BL21 (DE3) competent cells and typically grown in 50 ml LB containing 40  $\mu\text{g}/\text{ml}$  kanamycin. The expression of the His-tagged protein was induced in mid-log phase (corresponding to an absorbance at 600 nm of 0.6-0.7) by the addition of isopropyl-B-D-thiogalactopyranoside (IPTG) to a final concentration of 0.45 mM. Cells were subsequently grown supplemented with 0.25 mM  $\text{ZnCl}_2$  at both 37 °C and 20 °C (Note: culture flasks were incubated at the desired temperature for 30 minutes prior to induction with IPTG) for the duration of 4 hours with samples removed for SDS PAGE analysis as described previously for GST/B-Myb expression trials.

The p300 1603-1839 construct was transformed into several strains of *E. coli* including BL21 (DE3), BL21 pLys S, and two strains of BL21-CodonPlus<sup>®</sup> competent cells (Stratagene), BL21-CodonPlus<sup>®</sup> (DE3) RIL and BL21-CodonPlus<sup>®</sup> (DE3) RP. BL21-CodonPlus<sup>®</sup> (DE3) RIL cells contain extra copies of the *argU*, *ileY* and *leuW* tRNA genes, which specifically recognise the arginine codons AGA and AGG, the isoleucine codon AUA, and the leucine codon CUA, respectively. BL21-CodonPlus<sup>®</sup> (DE3) RP cells contain extra copies of the *argU* and *proL* genes. These genes encode tRNA's that recognise the arginine codons AGA and AGG and the proline codon CCC, respectively. The transformation of BL21-CodonPlus<sup>®</sup> competent cell strains were completed as directed in the manufacturers guidelines (Stratagene). Transformed cells were typically grown in 50 ml LB containing 40 µg/ml kanamycin to maintain the pET28a expression plasmid, with the addition of 50 µg/ml chloramphenicol to maintain the pACYC plasmid in the BL21-CodonPlus<sup>®</sup> competent cell strain. The effect of the addition of 1 µg/ml chloramphenicol on p300 1603-1839 expression in BL21 (DE3) cells was also monitored. His-tagged p300 1603-1839 was induced and expressed as previously described for p300 1603-1725 at both 37 °C and 20 °C temperatures.

The His-tagged p300 1726-1839 construct was transformed into both *E. coli* BL21 (DE3) and BL21-CodonPlus<sup>®</sup> (DE3) RP strains and grown at 37 °C and 20 °C as previously described.

### *2.2.11 Expression and Inclusion body preparation of His-tagged p300 fusion proteins*

*E. coli* cells transformed with the pET28a-based expression vectors for p300 1603-1725, 1726-1839, and 1603-1839 were grown in LB medium containing 40 µg/ml kanamycin with or without 50 µg/ml chloramphenicol and grown to 4 hours post induction with 0.45 mM IPTG in mid-log phase. The cells were harvested by centrifugation at 7800 g for 15 minutes (4 °C). Cell pellets were lysed with Bug-buster HT to which was added one EDTA-free protease inhibitor cocktail tablet (Roche) and 100 µM phenylmethylsulfonylfluoride (PMSF) to inhibit protease activity. The insoluble fraction of the cell lysate containing p300 as inclusion bodies was recovered by centrifugation (12100 g for 15 minutes at 4 °C). The inclusion bodies were then washed four times in a 50 mM Tris, 0.5% (v/v) Triton X-100 buffer adjusted to pH 8.0, with inclusion bodies recovered with centrifugation at 10000 rpm for 10 minutes at 4 °C

between subsequent washes. Washed inclusion bodies were finally solubilised in a 6 M guanidine hydrochloride, 20 mM Tris and 100 mM sodium chloride denaturant buffer at pH 8.0, to give a final p300 concentration of about 0.6 mg/ml.

#### *2.2.12 Refolding and Purification of p300 1603-1725*

The p300 protein was refolded by an Immobilised Metal Affinity Chromatography (IMAC) procedure whilst reversibly bound to a Ni-NTA affinity column as follows: The solubilised p300 1603-1725 inclusion bodies were loaded onto a 10 ml Ni-NTA superflow (Pharmacia) column pre-equilibrated with denaturant buffer and unbound protein removed after washing with three column volumes of the same buffer. The denaturant was then washed off the column by the application of five column volumes of refolding buffer (20 mM Tris, 100 mM sodium chloride, pH 8.0). Soluble His-tagged p300 was eluted after applying a linear gradient of imidazole from 0 to 300 mM, in a 20 mM Tris and 100 mM sodium chloride buffer at pH 8.0. Fractions containing His-tagged p300 1603-1725 were pooled, dialysed against a 20 mM Tris and 100 mM sodium chloride buffer at pH 7.0, and then incubated with Thrombin (Pharmacia, 10 U per mg of protein) for 16 hours at room temperature to remove the N-terminal His-tag. An SDS PAGE gel containing pre- and post- thrombin cleavage samples was run to verify correct and efficient digestion. The cleaved His-tag and thrombin were separated from the p300 1603-1725 protein by a second Ni-NTA affinity chromatography step, with bound p300 1603-1725 eluting at an imidazole concentration of about 75 mM, and the cleaved His-tag eluting at 200-300 mM imidazole concentration. The protein was finally dialysed into a filtered and degassed 20 mM Bis-tris, 100 mM sodium chloride, 20 mM Dithithreitol (DTT) buffer containing 0.5 mM ZnCl<sub>2</sub> at pH 7.0. Zn<sup>2+</sup>-free protein was produced by dialysing protein against a 20 mM Bis-tris, 100 mM sodium chloride, 20 mM Dithithreitol (DTT) buffer at pH 7.0 (i.e. omitting the ZnCl<sub>2</sub>). The purified protein was found to be over 95 % pure by SDS PAGE.

#### *2.2.13 Refolding and Purification of p300 1603-1839*

Isolated and washed p300 1603-1839 bacterial inclusion bodies were subject to a range of refolding and purification strategies. Refolding by dialysis, IMAC-based chromatography, and 'drop-wise' dilutions were all attempted. Furthermore, variations in buffer types, protein concentration, pH, reducing agent and detergent content were all investigated often in combination. The best refolding strategy required the prior

purification of His-tagged p300 from bacterial inclusion bodies by Ni-NTA affinity chromatography under denaturing conditions, before refolding by dialysis as detailed below:-

(A) Purification of p300 from bacterial inclusion bodies.

P300 1603-1839 protein inclusion bodies were 'purified' under denaturing conditions after applying the following protocol. Inclusion bodies were dissolved in 6M guanidine-hydrochloride, 20 mM Tris, and 100 mM sodium chloride at pH 8.0 (denaturant buffer), and centrifuged as previously described and loaded onto a 10 ml Ni-NTA superflow (Pharmacia) column pre-equilibrated with denaturant buffer. Unbound protein was removed after washing with three column volumes of the same denaturant buffer. P300 1603-1839 protein was then eluted under denaturing conditions after applying a linear gradient of imidazole from 0 to 300 mM, in a 6M Guanidine-Hydrochloride, 20 mM Tris and 100 mM sodium chloride buffer at pH 8.0. Elution Fractions containing purified p300 1603-1839 were pooled, and a Bradford assay performed to estimate protein concentration

(B) Refolding by dialysis

Purified p300 1603-1839 protein dissolved in 6M Guanidine-Hydrochloride, 20 mM Tris, and 100 mM Sodium Chloride at pH 7.0 was centrifuged and filtered to rid any insoluble matter. The protein concentration was then adjusted to approximately 200 µg/ml, incubated with 10 mM DTT for 30 minutes at room temperature before overnight dialysis twice against 2 litres of 20 mM Tris, 100 mM Sodium Chloride, 20 mM DTT, 0.5 mM ZnCl<sub>2</sub>, pH 7.0 Overnight dialysates were retrieved and insoluble material separated by centrifugation at 4500 g for 15 minutes. Samples from the soluble material was analysed for the presence of p300 1603-1839 protein by SDS PAGE.

*2.2.14 Refolding and Purification of p300 1726-1839*

His-tagged p300 1726-1839 inclusion bodies were washed and isolated as described in section 2.2.12 and purified under denaturing conditions as described for p300 1603-1839. Purified inclusion bodies were then adjusted to a protein concentration of approximately 200 µg/ml, incubated with 10 mM BME for 30 minutes at room temperature before overnight dialysis against the following filtered and degassed refolding buffers:

- 2 litres of a 20 mM Bis-tris, 100 mM sodium chloride, 10 mM BME, and 0.5 mM ZnCl<sub>2</sub> buffer at pH 7.0

- 2X 2 litres of a 20 mM Bis-tris and 100 mM sodium chloride buffer at pH 6.5

The overnight dialysate was centrifuged and filtered to rid insoluble matter, and a Bradford assay performed once more to estimate His-tagged p300 1726-1839 protein concentration and recovery from refolding. The His-tag was removed and p300 1726-1839 purified as previously described for p300 1603-1725 (see 2.2.12). Purified p300 1726-1839 fractions were pooled and the protein finally dialysed into a filtered and degassed 20 mM Bis-tris, 100 mM sodium chloride buffer at pH 6.5.

#### 2.2.15 *His-tagged cyclin D1 protein expression trials*

The His-tagged cyclin D1 construct was transformed into *E. coli* BL21 (DE3) competent cells and grown in 50 ml LB containing 40 µg/ml kanamycin. The expression of the His-tagged protein was induced in mid-log phase (corresponding to an absorbance at 600 nm of 0.6-0.7) by the addition of isopropyl-B-D-thiogalactopyranoside (IPTG) to a final concentration of 0.45 mM. Cells were subsequently grown at both 37 °C and 20 °C (Note: culture flasks were incubated at the desired temperature for 30 minutes prior to induction with IPTG) for the duration of 4 hours with samples removed for SDS PAGE analysis as described previously for GST/B-Myb and p300 expression trials.

#### 2.2.16 *Expression and Inclusion body preparation of His-tagged cyclin D1*

*E. coli* transformed with the pET28a-based expression vector for human cyclin D1 was grown in LB medium containing 40 µg/ml kanamycin for 4 hours after induction with IPTG. The cells were harvested by centrifugation at 7800 g for 15 minutes (4 °C). Cell pellets were lysed with Bug-buster HT to which was added 0.5 mM EDTA and 100 µM phenylmethylsulfonylfluoride (PMSF) to inhibit protease activity. The insoluble fraction of the cell lysate containing His-tagged cyclin D1 as inclusion bodies, was recovered by centrifugation (12100 g for 15 minutes at 4 °C). The inclusion bodies were then enriched by washing four times in a 50 mM Tris, 0.5% (v/v) Triton X-100, 10 mM EDTA buffer adjusted to pH 8.0, with inclusion bodies recovered with centrifugation at 10000 rpm for 10 minutes at 4 °C between subsequent washes. Washed inclusion bodies were finally solubilised in a 6 M guanidine hydrochloride containing 20 mM Tris and

100 mM sodium chloride denaturant buffer at pH 8.0 to give a final cyclin D1 concentration of typically between 0.5-1 mg/ml.

#### *2.2.17 Refolding and Purification of His-tagged cyclin D1*

His-tagged cyclin D1 bacterial inclusion bodies were readily refolded by one of two methods through either (1) purifying His-tagged cyclin D1 under denaturing conditions (IMAC) and refolding by dialysis, or by (2) directly refolding by an 'on-column' IMAC approach as previously described for His-tagged p300 1603-1725.

(1) Inclusion bodies solubilised in 6 M guanidine hydrochloride containing 20 mM Tris and 100 mM sodium chloride at pH 8.0 were centrifuged and filtered as previously described and loaded onto a 10 ml Ni-NTA superflow (Pharmacia) column pre-equilibrated with denaturant buffer. Unbound protein was removed after washing with three column volumes of the same denaturant buffer. His-tagged protein was then purified under denaturing conditions after applying a linear gradient of imidazole from 0 to 300 mM, in a 6M Guanidine-Hydrochloride, 20 mM Tris and 100 mM sodium chloride buffer at pH 8.0. Elution Fractions containing purified His-tagged cyclin D1 were pooled, and a Bradford assay performed to estimate protein concentration. Purified protein inclusion bodies were then adjusted to a protein concentration of approximately 200 µg/ ml, and dialysed extensively 3 times against 2 litres of a 25 mM NaH<sub>2</sub>PO<sub>4</sub>, 100 mM sodium chloride, 1mM DTT, 1mM EDTA buffer at pH 8.0.

(2) The His-tagged cyclin D1 protein was refolded by an Immobilised Metal Affinity Chromatography (IMAC) procedure whilst reversibly bound to a Ni-NTA affinity column as follows: The solubilised His-tagged cyclin D1 inclusion bodies were loaded onto a 10 ml Ni-NTA superflow (Pharmacia) column pre-equilibrated with denaturant buffer and unbound protein removed after washing with three column volumes of the same buffer. The denaturant was then washed off the column by the application of five column volumes of refolding buffer (20 mM Tris, 100 mM sodium chloride, pH 8.0). Soluble His-tagged cyclin D1 was eluted after applying a linear gradient of imidazole from 0 to 300 mM, in a 20 mM Tris and 100 mM sodium chloride buffer at pH 8.0. Fractions containing His-tagged cyclin D1 were identified by viewing samples on SDS PAGE and were subsequently pooled and dialysed extensively two times against 2 litres

of 25 mM NaH<sub>2</sub>PO<sub>4</sub>, 100 mM sodium chloride, 1mM DTT, 1mM EDTA buffer at pH 8.0.

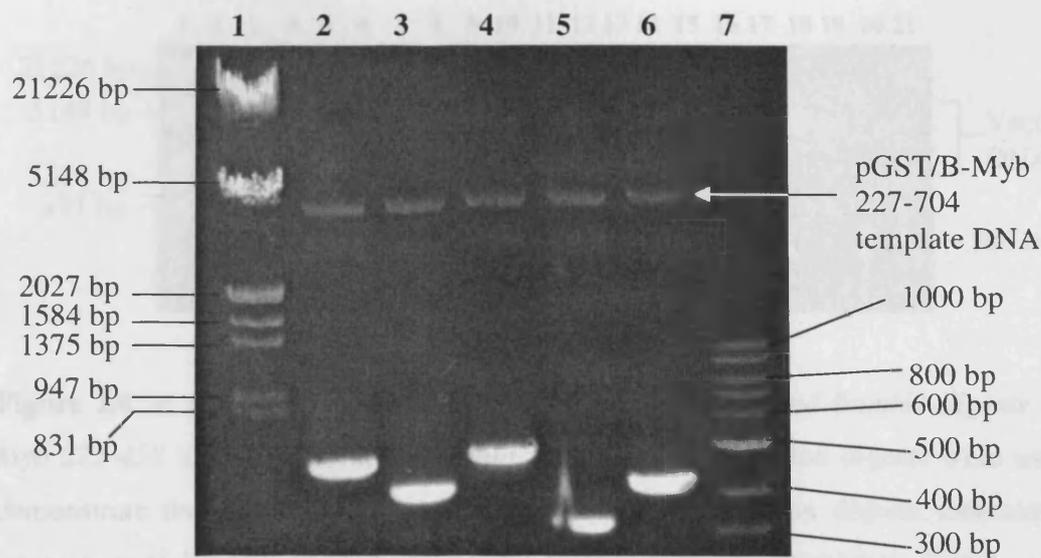
## 2.3 Results

### 2.3.1 B-Myb construct design

At the time of experimental design a central region of B-Myb (residues 227 to 508) had been shown to interact with cyclin D1 (Horstmann *et al.*, 2000b) and preliminary unpublished data implicated the same region of B-Myb with binding the C/H3 (cysteine/histidine rich region 3) region of p300, specifically residues 1515 to 1998 (Karl-Heinz Klempnauer's research group in Munster, Germany). The GST/B-Myb constructs spanning residues 227-508, 275-376 were kind gifts from Karl-Heinz Klempnauer. A further five GST/B-Myb constructs (275 to 458, 318 to 458, 318 to 508, 369 to 458, and 369 to 508) spanning the 227-508 residue region were designed primarily based on regions of sequence homology break-points which are often indicative of potential domain boundaries (alignment data not shown). The B-Myb 318-458 construct was specifically designed because it spans the exon 9a region of B-Myb which is absent in an alternatively spliced, inactive B-Myb isoform (Horstmann *et al.*, 2000a).

### 2.3.2 PCR amplification and cloning of murine B-Myb sequences

Successful amplification of B-Myb sequences was confirmed after visualising the products obtained from PCR reactions on an agarose gel shown in figure 2.3. The gel clearly shows an abundance of amplified B-Myb DNA of the expected size.



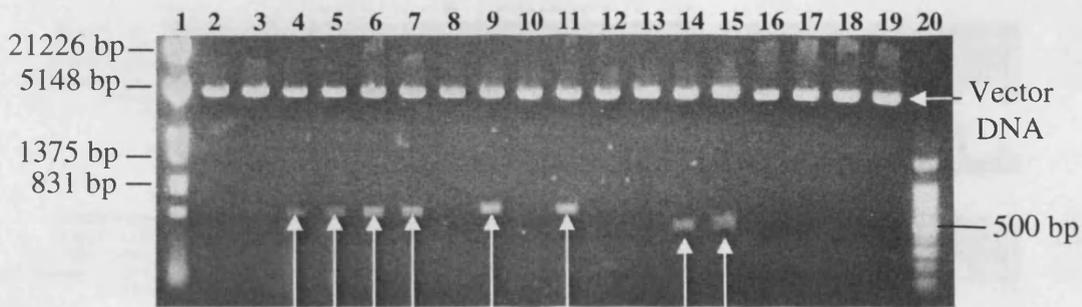
**Figure 2.3.** An ethidium bromide stained 1 % (v/v) agarose gel viewed under UV light displaying the products of B-Myb PCR amplification. Lanes 1 and 7 contain  $\lambda$ DNA/*Hind III* marker (Promega) and 100 bp DNA ladder (Promega) respectively. Lanes 2 to 7 contain PCR products for B-Myb 275-458 (569 bp), B-Myb 318-458 (440 bp), B-Myb 318-508 (590 bp), B-Myb 369-458 (287 bp), and B-Myb 369-508 (437 bp) respectively.

Ligation of B-Myb PCR products with pT7Blue-3 were confirmed after agarose gel analysis of single *EcoRI* and *BamHI* restriction digests of plasmid DNA prepared from successful ligase mix transformants (Displayed for B-Myb 275-458 PCR product in figure 2.4). The presence of insert sized fragments after *EcoRI* restriction digests of plasmid DNA firstly identified successful ligation of B-Myb PCR product, due to the existence of *EcoRI* restriction sites flanking the blunt cloning site in pT7Blue-3. Secondly, *BamHI* restriction digests both in the cloned B-Myb inserts and the pT7Blue-3 vector ensured that the relative orientation of cloned insert could be deciphered based on the size of the *BamHI* digested insert. For reference purposes a sample of PCR product control was included on agarose gels. The sequence integrity and orientation of the PCR product was further verified after DNA sequencing.



**Figure 2.4.** A 1% (w/v) agarose gel showing single *EcoRI* and *BamHI* digests of *B-Myb* 275-458 ligated *pT7Blue-3* plasmid DNA. *EcoRI* restriction digests were used to demonstrate the presence of insert, while *BamHI* restriction digests indicated the orientation of ligated insert. Successfully ligated plasmids containing *EcoRI* excised inserts of the expected size are highlighted by arrows in lanes 3, 6, 12, 15, and 18. *BamHI* excised inserts are highlighted by arrows in lanes 4, 7, and 16. The remaining lanes contain; (1)  $\lambda$ DNA/*Hind III* marker (Promega). (2, 5, 8, 11, 14, 17) Undigested *pT7Blue-3* vector. (9, 10) False positive transformants that lack ligated insert. (13, 19) Successfully ligated plasmids that lack insert sized fragment after *BamHI* digestion. (20) PCR product of *B-Myb* (275-458) as reference control. (21) 100 bp ladder (Promega). Analysis of the remaining *B-Myb* ligated *pT7Blue-3* plasmids was similarly conducted.

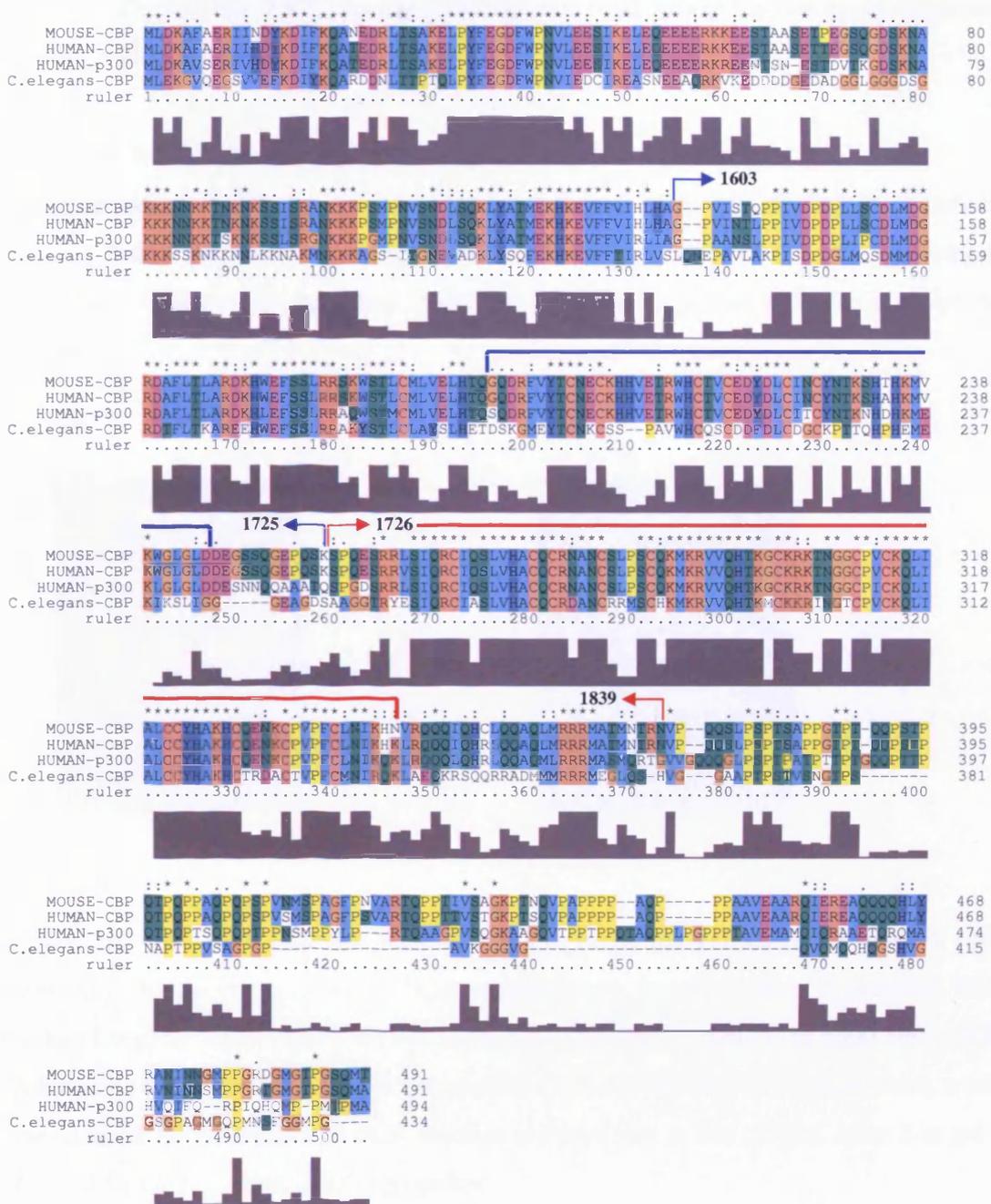
Dual restricted (*BamHI/XhoI*) *B-Myb* inserts derived from *pT7blue-3* were subsequently ligated between the same restriction sites found within the *pGEX-6P2* vector. Ligation progress was once again monitored by agarose gel electrophoresis of restricted plasmid DNA isolated from successful transformants (Displayed for *pGEX-6P2* ligated *B-Myb* 275-458 and *B-Myb* 318-458 in figure 2.5). The sequence integrity of cloned *B-Myb* inserts was verified after DNA sequencing.



**Figure 2.5.** A 1% (w/v) agarose gel showing *Bam* HI/*Xho* I dual digests of B-Myb 275-458 and B-Myb 318-458 ligated pGEX-6P2 plasmid DNA. Successfully ligated plasmids containing B-Myb 275-458 excised inserts of the expected size are highlighted by arrows in lanes 4, 5, 6, 7, 9 and 11. B-Myb 318-458 containing plasmids are evident by insert sized fragments produced in lanes 14 and 15. The remaining lanes contain; (1)  $\lambda$ DNA/*Hind* III marker (Promega), (2, 3, 8, 10, 12, 13, 16-20) False positive transformants that lack ligated insert, (20) 100 bp ladder (Promega). Analysis of the remaining B-Myb ligated pGEX-6P2 plasmids was similarly conducted.

### 2.3.3 p300 construct design

At the time of construct design a C/H3 containing region of p300 (residues 1515-1998) was implicated with binding a central region of B-Myb (residues 227 to 508) (Unpublished observations, Karl-Heinz Klempnauer). The C/H3 region, which contains the ZZ motif and TAZ2 domain, is a well documented protein interaction site associated with binding a multitude of transcription factors including E1A (Eckner *et al.*, 1994), GATA-1 (Blobel *et al.*, 1998), and E2F (Trouche *et al.*, 1996), to name but a few. Three p300 constructs were therefore designed spanning the ZZ motif (residues 1603-1725), the TAZ2 domain (residues 1726-1839) and a construct spanning both the ZZ and TAZ2 regions (1603-1839), in an attempt to define the site of interaction with B-Myb further. Construct boundaries were aided by analysing break-points in sequence homology using a multiple sequence alignment of p300 and CBP proteins (shown in figure 2.6). The amino acid boundaries for the TAZ2 domain (1726-1812) in human p300 were already known from the published structure (DeGuzman *et al.*, 2000) and were used to aid construct design also.

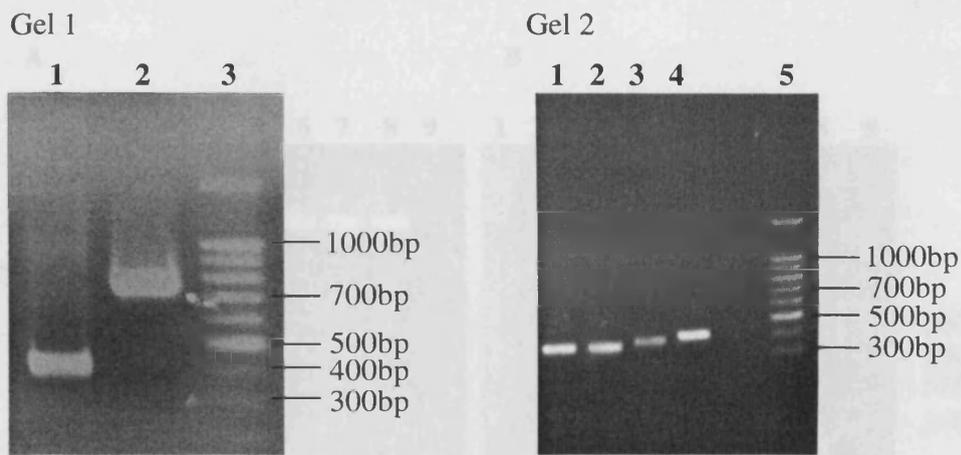


**Figure 2.6.** Multiple sequence alignment of p300/CBP sequences (residues 1515-1998 in human p300) spanning the C/H3 region used to aid the design of p300 constructs. The sequence enclosed by the blue and red arrows represents the p300 1603-1725 and p300 1726-1839 constructs respectively. A third construct spanning both of these regions was also designed (p300 1603-1839). The conserved ZZ motif (Ponting *et al.*, 1996) is shown in the sequence enclosed by the blue box (residues 1662 to 1713 in

human p300) with the TAZ2 domain (DeGuzman *et al.*, 2000) shown in the sequence enclosed by the red box (residues 1726-1812 in human p300).

#### 2.3.4 PCR amplification and cloning of human p300 sequences

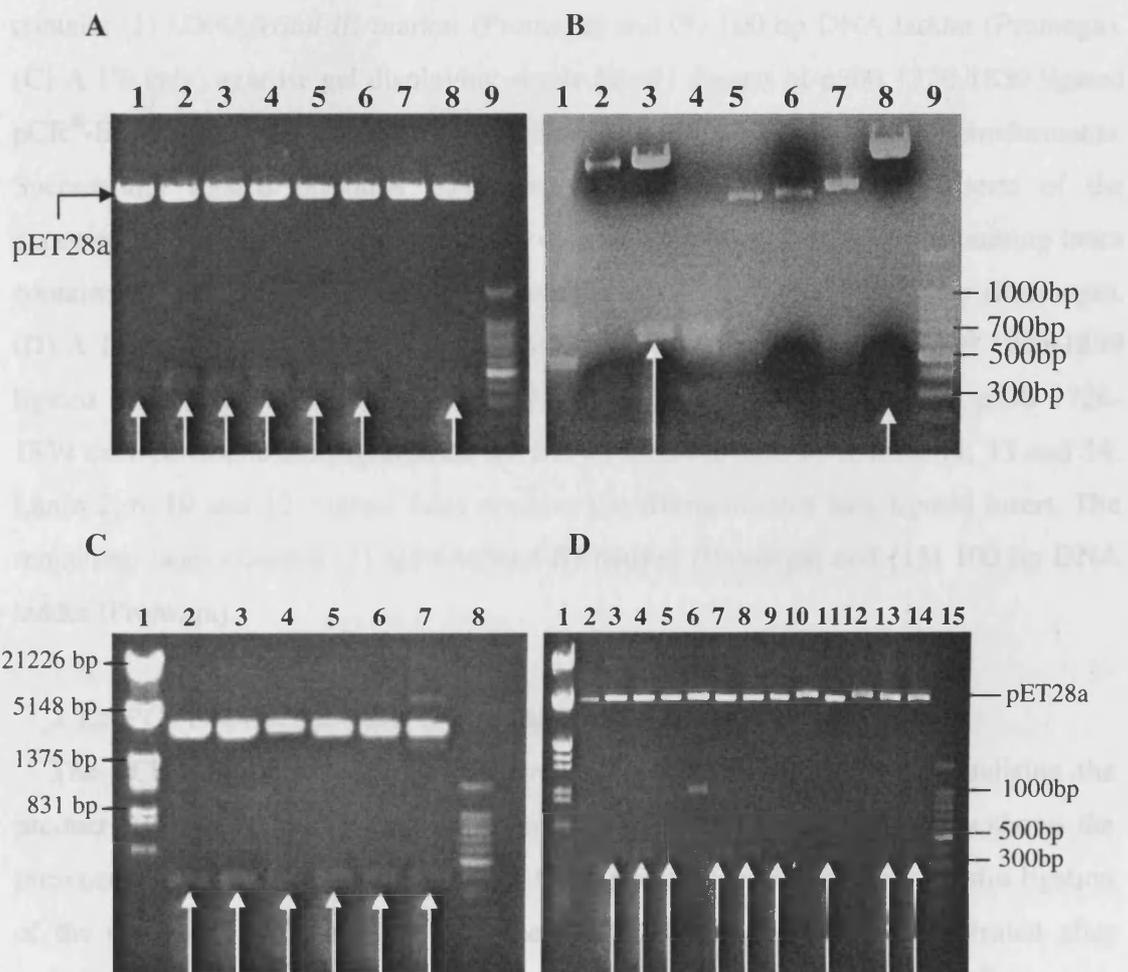
Amplification of p300 sequences was confirmed after visualising the products obtained from PCR reactions on the agarose gels shown in figure 2.7. The gel clearly shows an abundance of amplified p300 DNA of the expected size for each of the fragments.



**Figure 2.7.** Ethidium bromide stained 1 % (v/v) agarose gel viewed under UV light displaying the products of p300 PCR amplification. In gel 1, lane 1 contains PCR product for p300 1603-1725 (366 bp), lane 2 contains PCR product for p300 1603-1839 (708 bp). In gel 2 lanes 1 and 2 contain p300 1726-1839 (339 bp) PCR products while lanes 3 and 4 are products of a PCR reaction not pertinent to this project. Lane 3 in gel 1 and lane 5 in gel 2 contains 100 bp marker.

Ligations of p300 1603-1725 and p300 1603-1839 PCR products with pET28a vector were confirmed after agarose gel analysis of *NdeI* and *BamHI* dual restriction digests of plasmid DNA prepared from successful ligase mix transformants (figure 2.8A and 2.8B). A single *PstI* restriction digest of pET28a ligated p300 1603-1839 was also run to verify ligation, due to the presence of two *PstI* sites in the insert that are absent in the pET28a vector. Successful ligation of the p300 1726-1839 PCR product with the TOPO cloning vector was confirmed after viewing single *EcoRI* restriction digests of

isolated plasmid DNA (*EcoRI* sites flank the blunt cloning site in the pCR<sup>®</sup>-Blunt II-TOPO<sup>®</sup> vector, figure 2.8C). Dual restricted (*NdeI/Bam HI*) p300 1726-1839 inserts derived from the ligated TOPO vector were subsequently ligated between the same restriction sites found within pET28a. Ligation progress was once again monitored by agarose gel electrophoresis of restricted plasmid DNA (*Nde I/Bam HI*) isolated from successful transformants (figure 2.8D). The sequence integrity of both pCR<sup>®</sup>-Blunt II-TOPO<sup>®</sup> ligated and pET28a ligated vectors for all cloned p300 inserts were verified after DNA sequencing (PNAACL).



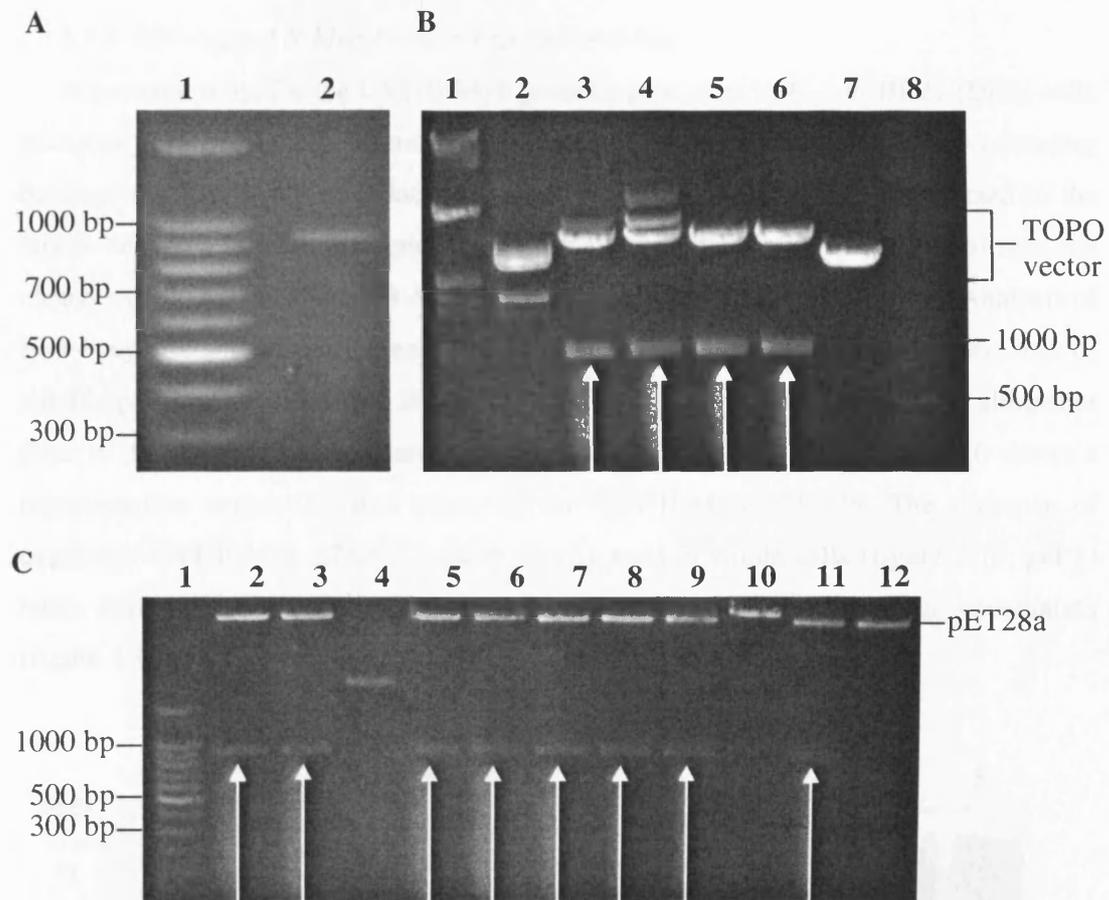
**Figure 2.8.** 1% (w/v) agarose gels showing restriction digests of p300 ligated pET28a plasmid DNA. (A) An ethidium bromide stained 1 % (v/v) agarose gel viewed under UV light displaying *NdeI/ Bam HI* dual digests of p300 1603-1725 ligated pET28a plasmid DNA. Successfully ligated plasmids containing p300 1603-1725 excised inserts of the expected size are highlighted by arrows in lanes 1 to 6 and lane 8. Lane 7 contains a

false positive transformant that lacks ligated insert. Lane 8 is the 100 bp ladder (Promega). **(B)** A 1 % (v/v) agarose gel confirming the ligation of p300 1603-1839 with pET28a. The successfully ligated pET28a vector containing the *NdeI/ BamHI* excised p300 1603-1839 insert of the expected size (708 bp) is highlighted by the arrow in lane 3. Lane 4 contains the PCR product of p300 1726-1839 as a reference control and double digested (*NdeI/ Bam HI*) pET28a is shown in lane 2. Uncut pET28a and p300 1603-1839 ligated pET28a are shown in lanes 5 and 7 respectively. Single *Pst I* digests of pET28a and p300 1603-1839 ligated pET28a are shown in lanes 6 and 8 respectively, with the *Pst I* derived insert highlighted by the arrow in lane 8. The remaining lanes contain; (1)  $\lambda$ DNA/*Hind III* marker (Promega) and (9) 100 bp DNA ladder (Promega). **(C)** A 1% (v/v) agarose gel displaying single *EcoRI* digests of p300 1726-1839 ligated pCR<sup>®</sup>-Blunt II-TOPO<sup>®</sup> plasmid DNA isolated from positive ligation transformants. Successfully ligated plasmids containing p300 1726-1839 excised inserts of the expected size (339 bp) are highlighted by arrows from lanes 2 to 7. The remaining lanes contain; (1)  $\lambda$ DNA/*Hind III* marker (Promega) and (8) 100 bp DNA ladder (Promega). **(D)** A 1% (v/v) agarose gel displaying *NdeI/ Bam HI* dual digests of p300 1726-1839 ligated pET28a plasmid DNA. Successfully ligated plasmids containing p300 1726-1839 excised inserts are highlighted by arrows in lanes 3, 5, 6, 7, 8, 9, 11, 13 and 14. Lanes 2, 6, 10 and 12 contain false positive transformants that lack ligated insert. The remaining lanes contain; (1)  $\lambda$ DNA/*Hind III* marker (Promega) and (15) 100 bp DNA ladder (Promega).

### 2.3.5 PCR amplification and cloning of human cyclin D1

The PCR amplification of human cyclin D1 was confirmed after visualising the product obtained on the agarose gels shown in figure 2.9A. The gel clearly shows the presence of amplified cyclin D1 DNA of the expected size (885 bp). Successful ligation of the cyclin D1 PCR product with the TOPO cloning vector was confirmed after viewing single *EcoRI* (*EcoRI* sites flank the blunt cloning site in the pCR<sup>®</sup>-Blunt II-TOPO<sup>®</sup> vector) restriction digests of isolated plasmid DNA (figure 2.9B). Dual restricted (*NdeI/Bam HI*) cyclin D1 inserts derived from the ligated TOPO vector were subsequently ligated between the same restriction sites found within pET28a. Ligation progress was once again monitored by agarose gel analysis of restricted plasmid DNA isolated from successful transformants (figure 2.9C). The sequence integrity of both

pCR<sup>®</sup>-Blunt II-TOPO<sup>®</sup> ligated and pET28a ligated vectors for all cloned p300 inserts were verified after DNA sequencing (PNACL).

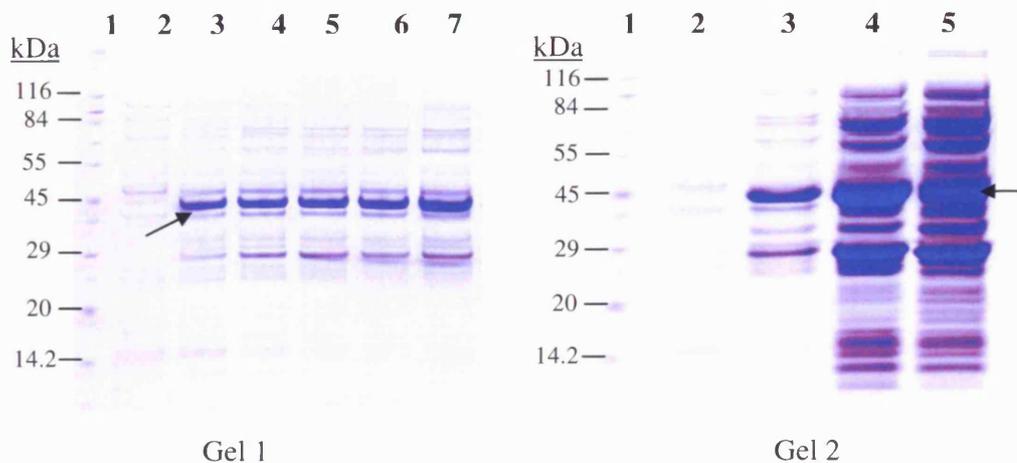


**Figure 2.9.** 1% (w/v) agarose gels detailing the cloning of cyclin D1 into pET28a plasmid DNA. (A) An ethidium bromide stained 1% (v/v) agarose gel viewed under UV light displaying the product of cyclin D1 PCR amplification (lane 2). Lane 1 contains 100 bp marker. Single *EcoRI* digests of cyclin D1 ligated pCR<sup>®</sup>-Blunt II-TOPO<sup>®</sup> plasmid DNA isolated from positive ligation transformants are displayed on the agarose gel in (B). Successfully ligated plasmids containing cyclin D1 excised (*EcoRI*) inserts of the expected size (885 bp) are highlighted by arrows from lanes 3 to 6. Lanes 2 and 7 contain false positive transformants that lack ligated cyclin D1 insert. The remaining lanes contain; (1)  $\lambda$ DNA/*Hind III* marker (Promega) and (8) 100 bp DNA ladder (Promega). (C) A 1% (v/v) agarose gel displaying *NdeI/Bam HI* dual digests of cyclin D1 ligated pET28a plasmid DNA. Successfully ligated plasmids containing cyclin D1 excised inserts of the expected size are highlighted by arrows in lanes 2, 3, 5-9, and 11.

Lanes 4, 10 and 12 contain false positive transformants that lack ligated insert. Lane 1 contains 100 bp marker.

### 2.3.6 GST-tagged B-Myb Protein Expression trials

Expression trials for the GST/B-Myb proteins generated in *E. coli* BL21 (DE3) cells revealed that they were all readily expressed with maximal production occurring between 3 to 5 hours post-induction. An induced protein product was detected in the whole cell post induction fractions for most of the GST fusion proteins but was not clearly visible for GST fused B-Myb 227-508 and B-Myb 369-508 proteins. Analysis of lysate supernatant samples revealed all the proteins were soluble. Again the presence of soluble protein for GST fused B-Myb 227-508 and B-Myb 369-508 was not altogether clear in this fraction but was later confirmed during purification. Figure 2.10 shows a representative expression trial observed for GST/B-Myb 275-376. The presence of expressed GST/B-Myb 275-376 can be clearly seen in whole cells (figure 2.10, gel 1) taken during the timecourse experiment, and is also evident in the lysate supernatant (figure 2.10, gel 2).

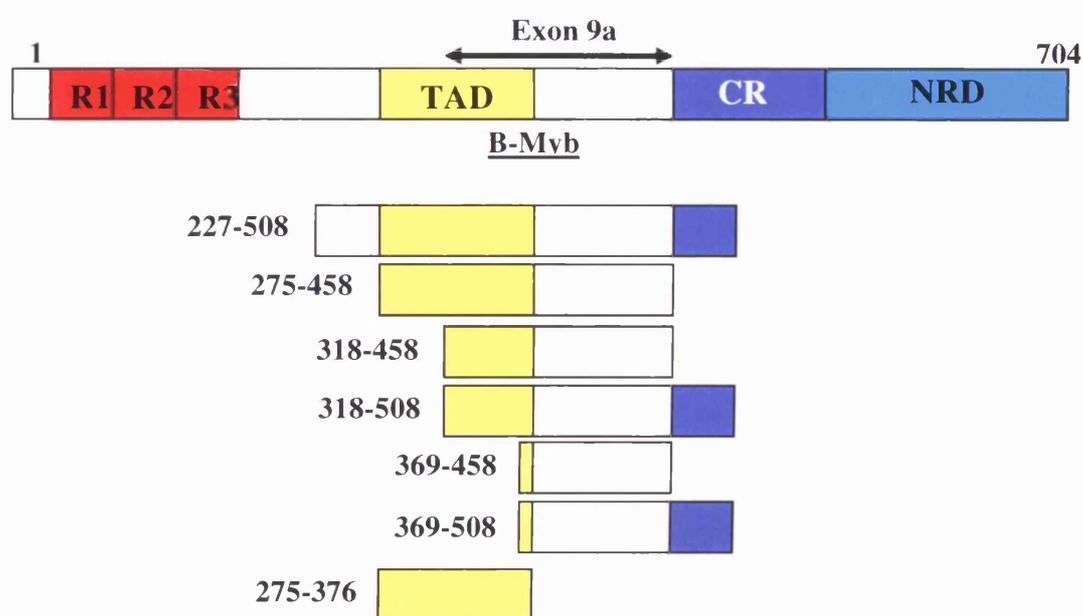


**Figure 2.10.** SDS PAGE gels showing GST-tagged B-Myb 275-376 protein readily expressed in an *E.coli*-based expression system as a soluble product. Gel 1 is representative of timecourse studies conducted for all the GST/B-Myb fusions. Expression of GST tagged B-Myb 275-376 is evident (indicated by the arrow) in the samples taken 1 to 5 hours post-induction from lanes 3 to 7 respectively. In gel 2, expression of GST-tagged B-Myb 275-376 is once again apparent in the post-induction sample in lane 3 and is also present in both the whole cell lysate and lysate supernatant

fractions in lanes 4 and 5 respectively. The remaining lanes in both gels are; lane 1, wide range molecular marker (Sigma), lane 2, pre-induction whole cell samples.

### 2.3.7 GST-tagged B-Myb Expression and Purification

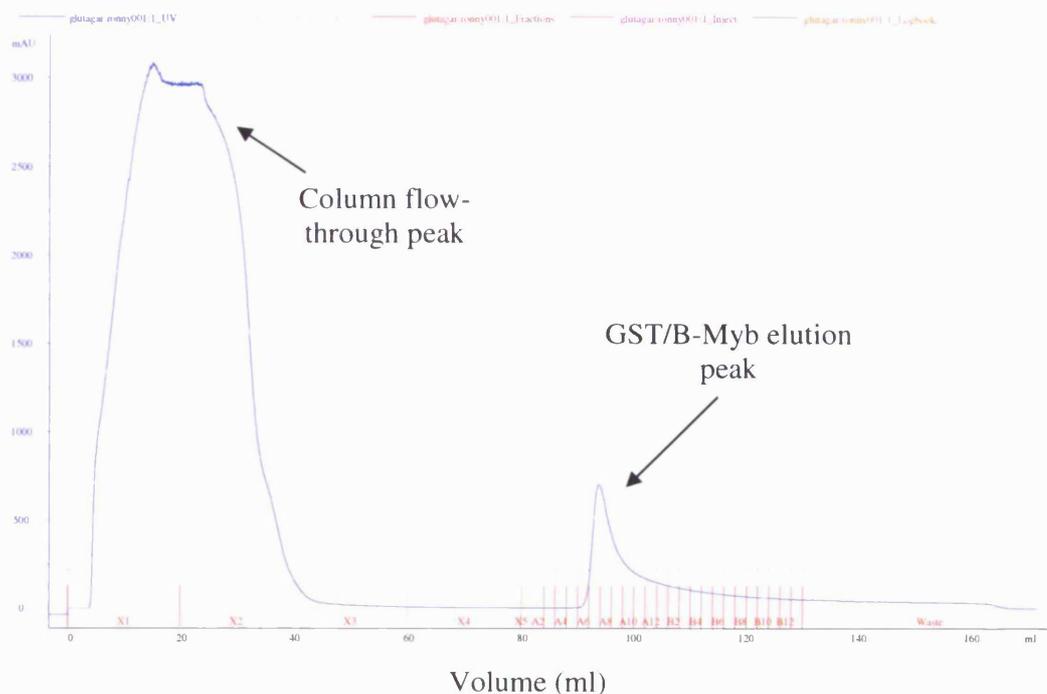
The GST-tagged B-Myb proteins, depicted in the schematic in figure 2.11, were expressed in *E. coli* and purified from lysate supernatant fractions by a single step affinity chromatography procedure using a glutathione-agarose (Sigma) column. A representative elution profile attained for a glutathione-agarose affinity purification procedure is shown for GST/B-Myb 275-376 in figure 2.12. FPLC elution profiles attained for the other GST/B-Myb proteins all have very similar traces.



**Figure 2.11.** A schematic representation of the B-Myb constructs used in this study. GST fusions were created N-terminus to the constructs. Various regions of functional importance are highlighted in red (DNA binding domain), yellow (transactivation associated region), blue (conserved region) and green (negative regulatory domain). The Exon 9a region of the protein is also highlighted by the double-headed arrow.

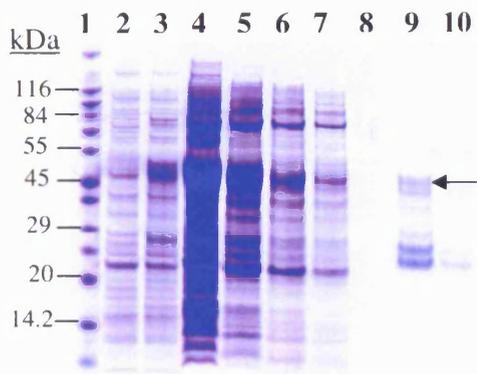
Progress of GST/B-Myb protein expression and affinity based purification was continually monitored by viewing SDS PAGE gels from samples retained at different points in the expression and purification procedure (figure 2.13). The purification of

GST/B-Myb R2R3 required an additional purification step using a CM-sepharose ion exchange column to rid associated DNA. Typical yields obtained for GST/B-Myb proteins were in the 5-10 mg/l range

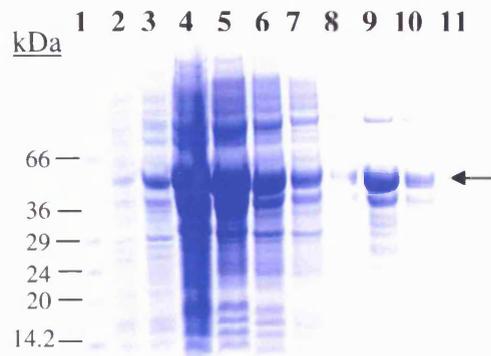


**Figure 2.12.** A typical affinity purification elution profile attained for GST/B-Myb fusion proteins from a glutathione agarose column performed on an Akta FPLC. The plot shows the collection of flow-through in fraction X1, washing of unbound material in fractions X2, X3, and X4 with base buffer (20 mM Tris, 100 mM NaCl, pH 7.0), and elution of bound GST fusion protein with a 10 mM glutathione buffer.

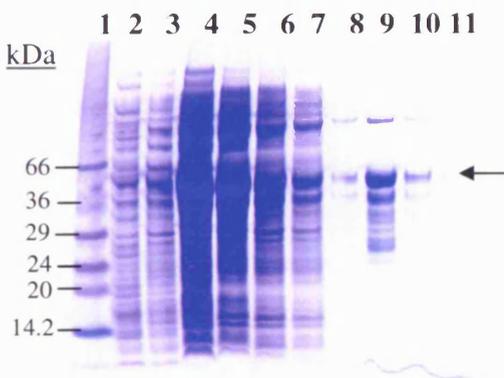
A characteristic feature of some of the final purified GST/B-Myb proteins is that they appear as a ladder on SDS PAGE, most notably GST tagged B-Myb 227 to 508, with the highest molecular weight species corresponding to that expected for the full-length protein (not in all cases). The mixture of products presumably reflects proteolysis in *E. coli* cells, leading to C-terminal truncations of GST/B-Myb fused proteins. Attempts at expressing the GST tagged B-Myb 227-508 protein at lower temperatures (20 °C) to prevent proteolysis had limited success, with a slightly greater proportion of full-length protein produced but with overall yields falling by 66 % (data not shown).



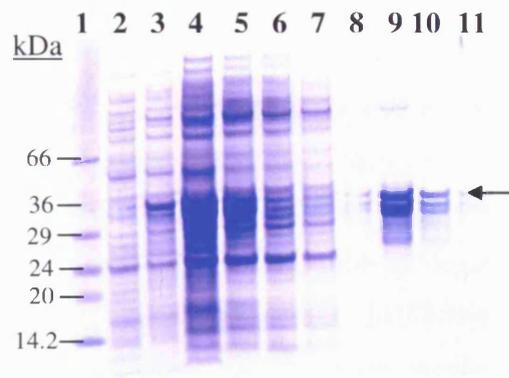
Gel 1



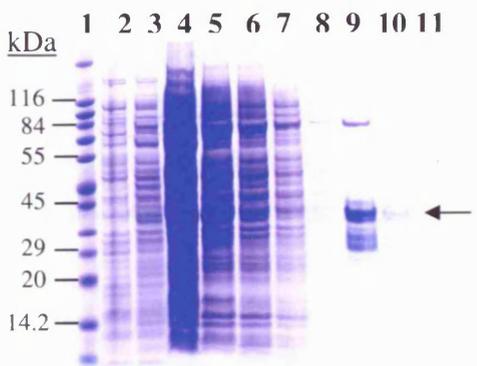
Gel 2



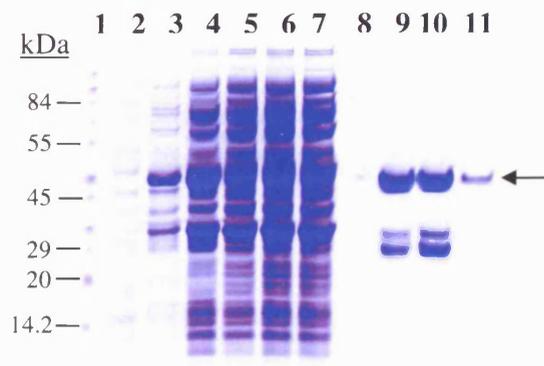
Gel 3



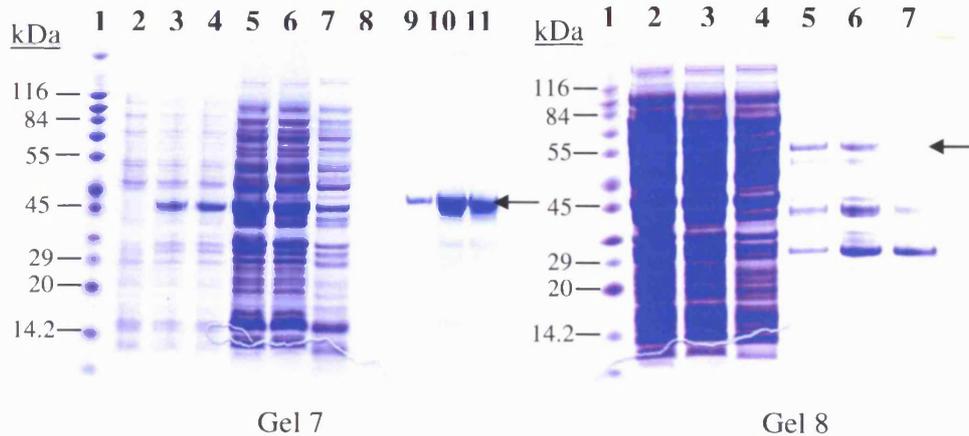
Gel 4



Gel 5



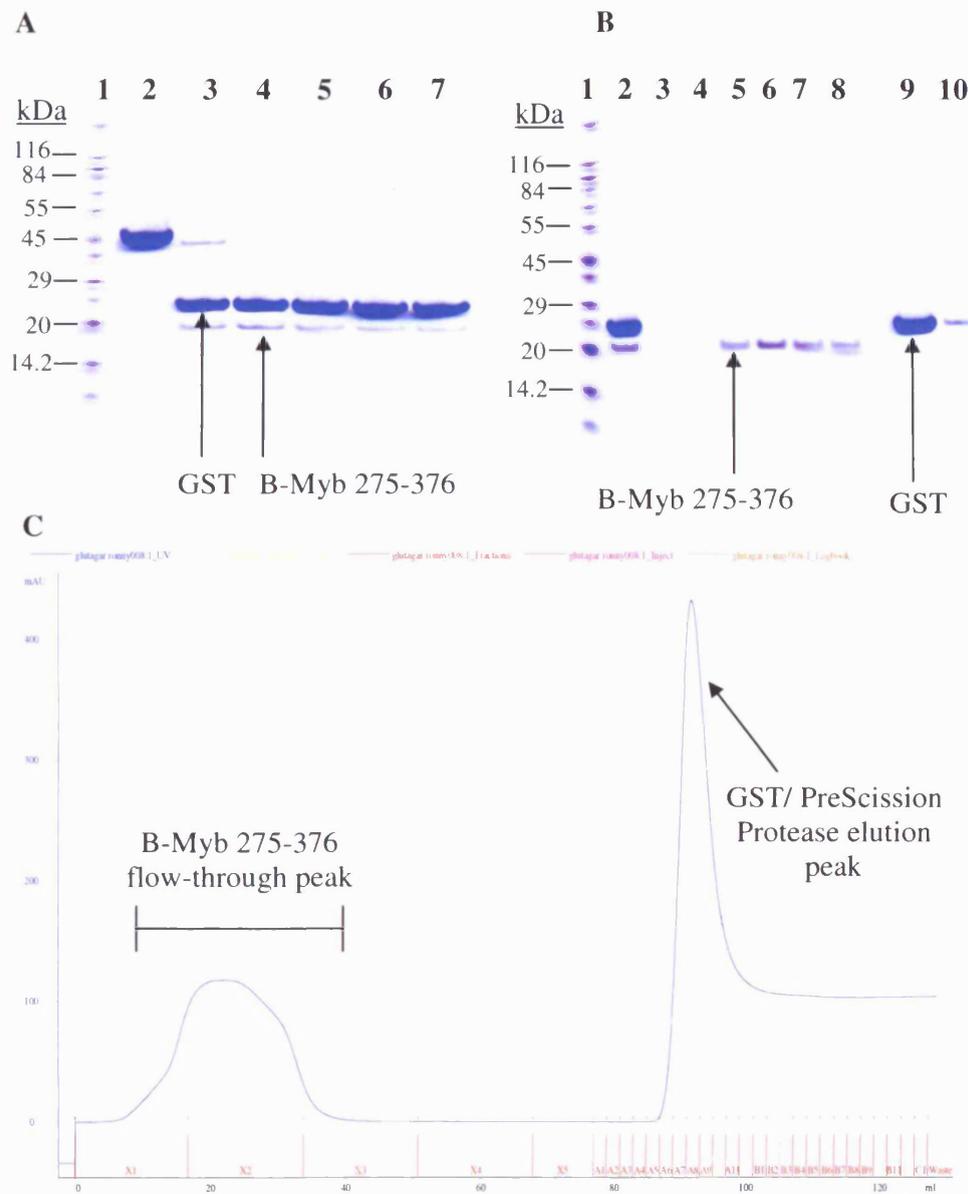
Gel 6



**Figure 2.13.** Coomassie stained SDS-PAGE gels detailing the expression and purification of GST fused B-Myb proteins. Gels 1-8 refer to (1) GST/B-Myb 275-458, (2) GST/B-Myb 318-458, (3) GST/B-Myb 318-508, (4) GST/B-Myb 369-458, (5) GST/B-Myb 369-508, (6) GST/B-Myb 275-376, (7) GST/B-Myb R2R3, and (8) GST/B-Myb 227-508 purification respectively. For each gel lane 1 represents a wide range molecular weight marker. For gels 1-6 Lanes 2 and 3 show un-induced and 3 hours post-induction whole cells, with whole cell lysates and lysate supernatants shown in lanes 4 and 5 respectively. Wash fractions collected during FPLC purification are shown in lanes 6, 7, and 8 with consecutive elution fractions occupying successive lanes from lanes 8 to 11. For gel 7 lanes 3, 4 and 5 show un-induced, 1 and 3 hours post-induction whole cells, with whole cell lysates and lysate supernatants shown in lanes 5 and 6 respectively. Wash fractions are shown in lanes 7 and 8, with elution fractions from lanes 9 to 11. For gel 8, lanes 2 and 3 show whole cell lysates and lysate supernatants, with elution fractions occupying lanes 5 to 7. The full-length GST/B-Myb protein species are highlighted with an arrow in each gel.

### 2.3.8 PreScission Protease cleavage and removal of GST tag from B-Myb 275-376

In order to isolate and purify the B-Myb transactivation region (B-Myb 275-376) the GST fusion protein was enzymatically cleaved at the PreScission Protease linker site. Upon overnight digestion (at 4 °C) of the fusion protein with PreScission Protease, B-Myb 275-376 was purified through a further glutathione agarose based affinity chromatography procedure. This ensured the efficient removal of both the cleaved GST tag and the GST tagged PreScission protease (purification detailed in figure 2.14).



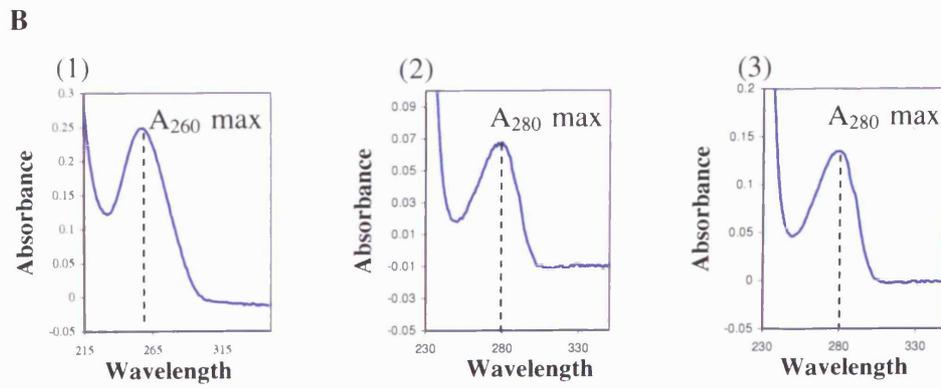
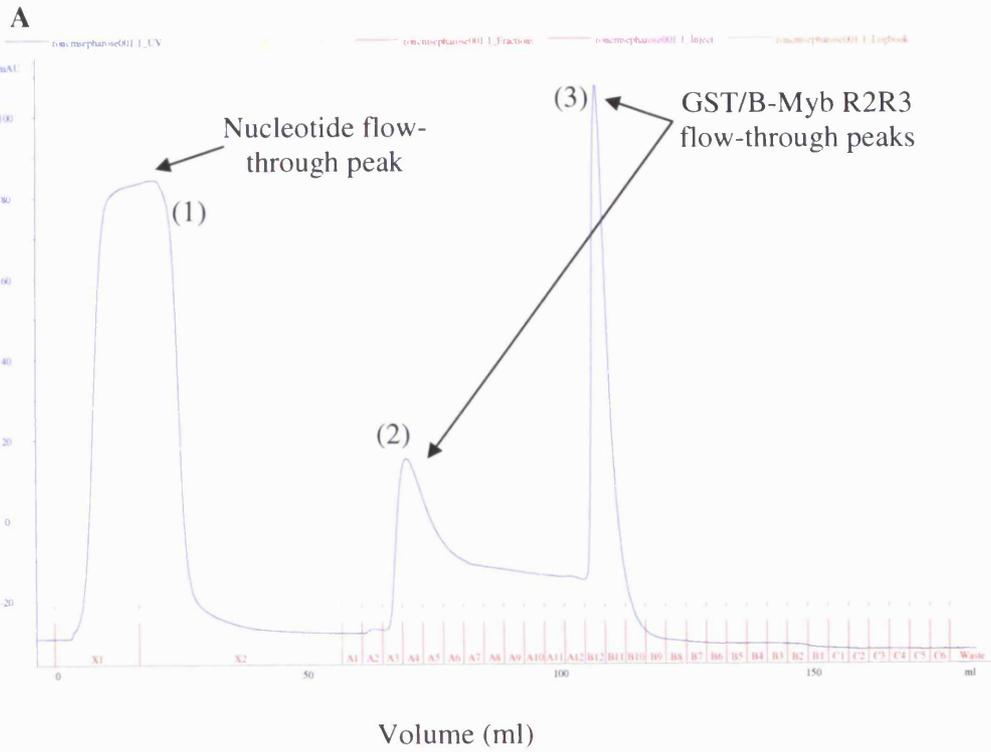
**Figure 2.14.** Purification of B-Myb 275-376. (A) An SDS PAGE detailing the progress of PreScission Protease mediated cleavage of GST/B-Myb 275-376 over 0 hr (lane 3), 1 hr (lane 4), 2 hr (lane 5), 3 hr (lane 6) and 16 hr (lane 7). Lane 1 shows a wide range molecular weight marker, and lane 2 contains undigested GST/B-Myb 275-376. Highlighted by the arrow in lane 3 is GST, where in lane 4 B-Myb 275-376 is indicated. (B) The purification of B-Myb 275-376 was monitored by SDS PAGE with lane 2 representing the column load, lanes 5 to 8 displaying the column flow-through which contains the B-Myb 275-376 protein of interest and lanes 9 and 10 containing eluted GST. The column flow-through containing B-Myb 275-376 is highlighted in the elution profile attained from the FPLC purification of B-Myb 275-376 from its GST tag using

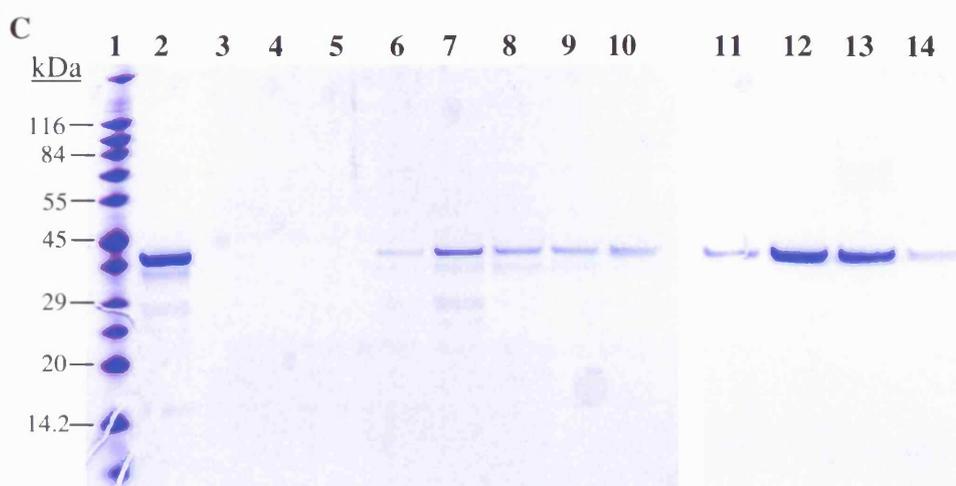
glutathione agarose based affinity chromatography (C). The plot shows the collection of flow-through in fraction X1, and washing of unbound material in fractions X2, X3, and X4 with base buffer (20 mM Tris, 100 mM NaCl, pH 7.0). The second peak represents the elution of GST and PreScission Protease after the implementation of a step containing 10 mM reduced glutathione.

Typical yields of purified B-Myb 275-376 were approximately 4 mg per litre of bacterial culture as determined by UV/ visible spectroscopy measured at 280 nm using a calculated extinction coefficient ( $15520 \text{ M}^{-1} \cdot \text{cm}^{-1}$ ) based on the number of Tryptophan and Tyrosine residues in the protein (Pace *et al.*, 1995). Both N-terminal sequencing and electrospray mass spectrometry (EMS) confirmed the purified B-Myb 275-376 protein was free of contaminants and recorded a mass of  $12089.14 \pm 1.17 \text{ Da}$  which corresponds closely to the expected 12090.2 Da mass

#### *2.3.9 Removal of contaminating DNA from GST/B-Myb R2R3*

The purification of GST/B-Myb R2R3 was found to contain a significant amount of associated DNA which is unsurprising for a DNA binding domain. In order to remove contaminating nucleotides from the glutathione agarose purified GST/B-Myb R2R3 protein, a final purification step using a CM-Sepharose column was employed. GST/B-Myb R2R3 was purified by applying a step gradient from 100 to 500 mM sodium chloride with the protein eluting at both 100 and 200 mM sodium chloride concentration steps. A combination of SDS PAGE analysis and UV/visible spectroscopy confirmed the successful purification of nucleotide free GST/ B-Myb R2R3 protein (figure 2.15).



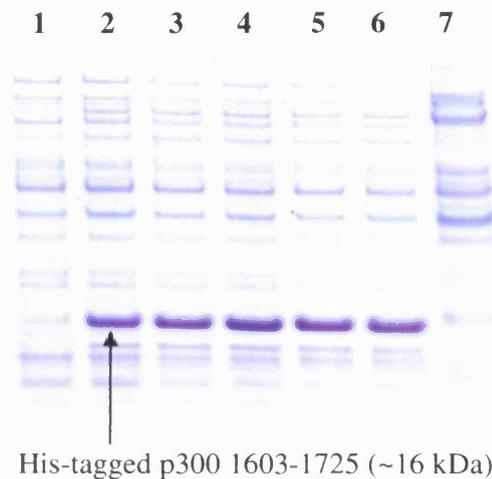


**Figure 2.15.** *The purification of nucleotide-free GST/B-Myb R2R3.* (A) The FPLC profile of the CM-Sepharose based purification procedure is displayed with highlighted regions including (1) Flow-through containing nucleotide contaminant, (2) and (3) GST/B-Myb R2R3 containing peaks eluting at 100 and 200 mM sodium chloride concentrations respectively. (B) UV/Visible spectroscopic scans taken from corresponding elution peaks (i.e. 1, 2 and 3) confirming the separation of nucleotide DNA absorbing at 260 nm from protein absorbing at 280 nm. (C) SDS PAGE analysis of fractions taken from the FPLC purification with lanes 3 to 5 taken from the nucleotide containing flow-through peak (1), lanes 6 to 10 representing the 100 mM sodium chloride protein elution peak (2) and lanes 11 to 14 corresponding to the 200 mM sodium chloride protein elution peak (3). Lane 1 contains wide range molecular marker (Sigma).

### 2.3.10 His-tagged p300 Protein Expression trials

#### 2.3.10.1 His-tagged p300 1603-1725 expression trials

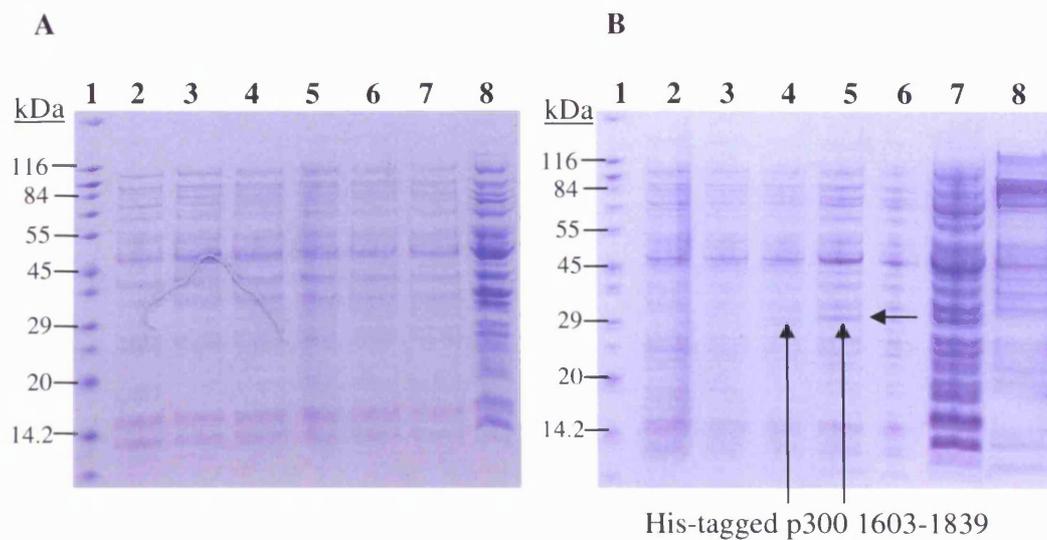
Expression trials of His-tagged p300 1603-1725 generated in *E. coli* BL21 (DE3) cells revealed the protein was readily expressed at 37 °C with maximal production occurring between 3 to 5 hours post-induction. Induced His-tagged p300 1603-1725 protein was clearly detected in whole cell post induction samples, and formed bacterial inclusion bodies in the insoluble fraction, as determined when viewing lysate supernatant samples on SDS PAGE (figure 2.16). Expression trials at the lower temperature of 25 °C failed to convert the expression of insoluble His-tagged p300 1603-1725 protein to a soluble form (Data not shown).



**Figure 2.16.** SDS PAGE gel showing His-tagged p300 1603-1725 readily expressed in an *E.coli*-based expression system as an insoluble protein. Expression of His-tagged p300 1603-1725 is evident (indicated by the arrow) in the samples taken 1 to 5 hours post-induction from lanes 2 to 6 respectively. Lane 7 shows the isolated supernatant fraction from lysed cells clearly demonstrating the absence of soluble His-tagged p300 1603-1725 protein. Lane 1 contains a sample from un-induced cells.

#### 2.3.10.2 His-tagged p300 1603-1839 expression trials

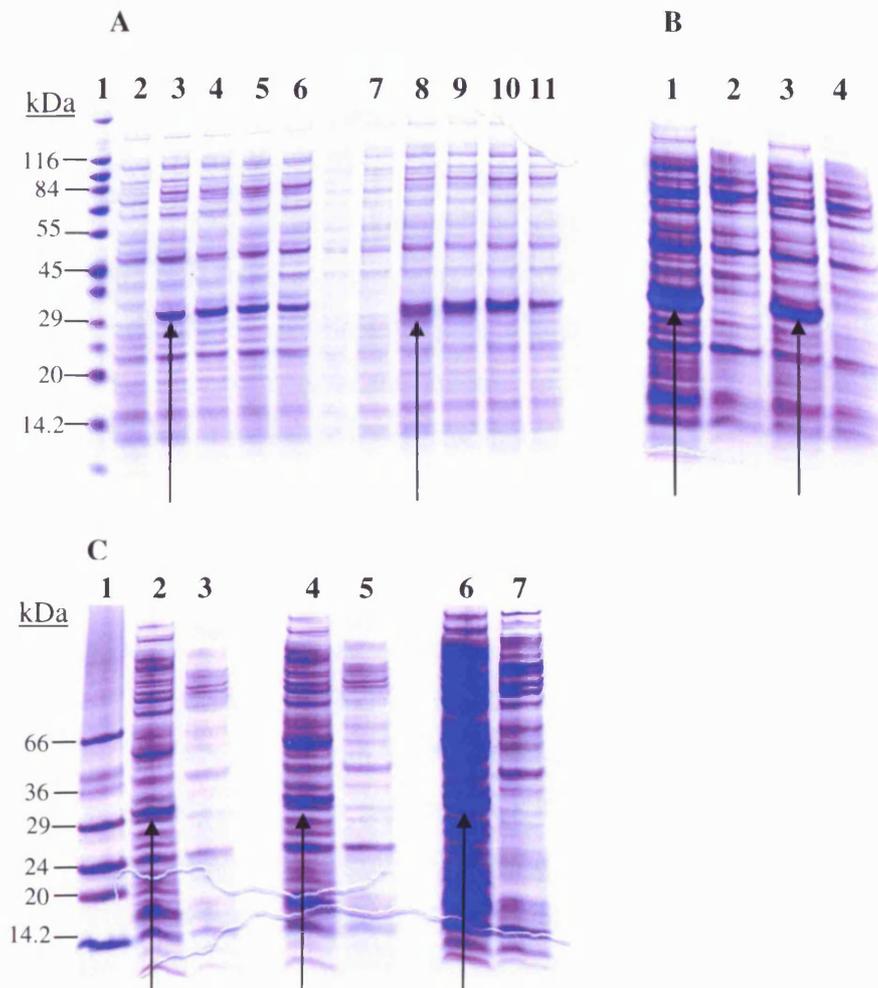
Expression trials of His-tagged p300 1603-1839 were conducted in several *E. coli* cell types including BL21 (DE3), BL21 (DE3) p(Lys) S, BL21-CodonPlus<sup>®</sup> (DE3) RIL and BL21-CodonPlus<sup>®</sup> (DE3) RP grown at both 37 °C and 20 °C. When BL21 (DE3) cells harbouring His-tagged p300 1603-1839 were grown at 37 °C, there was very little evidence of induction or expression of the protein when viewing SDS PAGE gels of expression trial samples (figure 2.17A). Upon growth at 20 °C, there was a very small level of induction and expression of His-tagged p300 1603-1939 detected in whole cell samples when viewed by SDS PAGE (figure 2.17B). Furthermore the protein appeared insoluble when comparing the whole cell lysate and supernatant samples from the growth trial. A similar observation was noted with expression trials of His-tagged p300 1603-1839 in BL21 (DE3) p(Lys) S cells. There was no visible improvement in the level of expression in this cell-type compared to that seen in BL21 (DE3) at either 37 °C or 20 °C temperatures (Data not shown).



**Figure 2.17.** Timecourse expression trials experiments for His-tagged p300 1603-1839. (A) An SDS PAGE gel displaying samples taken from a His-tagged p300 1603-1839 protein expression trial in BL21 (DE3) cells grown at 37 °C. Whole cell samples taken from 1 to 5 hours post-induction are shown in lanes 3 to 7 respectively with lane 2 showing un-induced cells. An isolated whole cell lysate sample is shown in lane 8, with lane 1 displaying the Wide range molecular marker (Sigma). (B) An SDS PAGE gel displaying samples taken from a His-tagged p300 1603-1839 protein expression trial in BL21 (DE3) cells grown at 20 °C. Whole cell samples removed at 1, 2, 4, and 6 hours post-induction are displayed in lanes 3 to 6 respectively with induction of the protein of interest highlighted by the arrows in lanes 4 and 5. Lane 1 shows un-induced cells. A sample from the whole cell lysate and lysate supernatant fractions are exhibited in lanes 7 and 8 respectively, with lane 1 displaying Wide range molecular marker (Sigma).

Expression trials from both BL21-CodonPlus<sup>®</sup> (DE3) RIL and BL21-CodonPlus<sup>®</sup> (DE3) RP transformed cells exhibit clear induction and expression of the His-tagged p300 1603-1839 protein when grown at both 37 °C and 20 °C. The SDS PAGE gels in figure 2.18A show protein expression between 1 and 6 hours post-induction in both cell types grown at 37 °C. Furthermore analyses of whole cell lysate and lysate supernatant samples from both cell types reveal the presence of His-tagged p300 1603-1839 protein in bacterial inclusion bodies in the insoluble fraction (figure 2.18B). Comparison of whole cell lysate and lysate supernatant samples from BL21 (DE3), BL21-CodonPlus<sup>®</sup> (DE3) RIL and BL21-CodonPlus<sup>®</sup> (DE3) RP cultures grown at 20 °C show there is no

visible improvement in soluble yields of His-tagged p300 1603-1839 (Figure 2.18C) at this temperature.

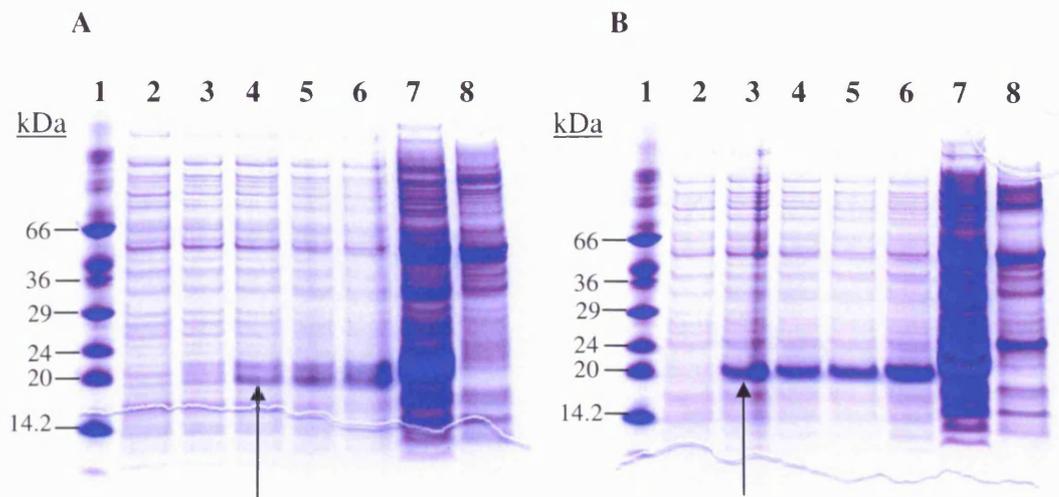


**Figure 2.18.** Timecourse expression trial experiments of His-tagged p300 1603-1839 (A) An SDS PAGE gel displaying samples taken from a His-tagged p300 1603-1839 protein expression trial in BL21-CodonPlus® (DE3) RIL and BL21-CodonPlus® (DE3) RP cells grown at 37 °C. Lanes 2 and 7 represent un-induced whole cell samples for BL21-CodonPlus® (DE3) RIL and BL21-CodonPlus® (DE3) RP cells respectively. Lanes 3 to 6 and 8 to 11 represent 1, 2, 4, and 6 hour post-induction samples for BL21-CodonPlus® (DE3) RIL and BL21-CodonPlus® (DE3) RP transformed cells respectively. The induction of His-tagged p300 1603-1839 protein expression is highlighted by the arrows in lanes 3 and 8. Lane 1 contains Wide range molecular marker (Sigma). (B) Lanes 1 and 2 represent whole cell lysate and lysate supernatant

samples for His tagged p300 1603-1839 protein expressed in BL21-CodonPlus<sup>®</sup> (DE3) RIL cells at 37 °C. Similarly lanes 3 and 4 represent whole cell lysate and lysate supernatant samples for His-tagged p300 1603-1839 protein expressed in BL21-CodonPlus<sup>®</sup> (DE3) RP cells at 37 °C. His-tagged p300 1603-1839 protein visible in the whole cell lysate samples is highlighted by the arrows in lanes 1 and 3 demonstrating the protein is largely insoluble. (C) An SDS PAGE gel displaying whole cell lysate and lysate supernatant samples from His-tagged p300 1603-1839 protein expressed in BL21 (DE3) (lanes 2 and 3), BL21-CodonPlus<sup>®</sup> (DE3) RIL cells (lanes 4 and 5), and BL21-CodonPlus<sup>®</sup> (DE3) RP (lanes 6 and 7) at 20 °C. His-tagged p300 1603-1839 protein visible in the whole cell lysate samples is highlighted by the arrows in lanes 2, 4 and 6. There is no improvement in the amount of soluble protein at the lower temperature. Lane 1 contains Wide range molecular marker (Sigma).

#### *2.3.10.3 His-tagged p300 1726-1839 expression trials*

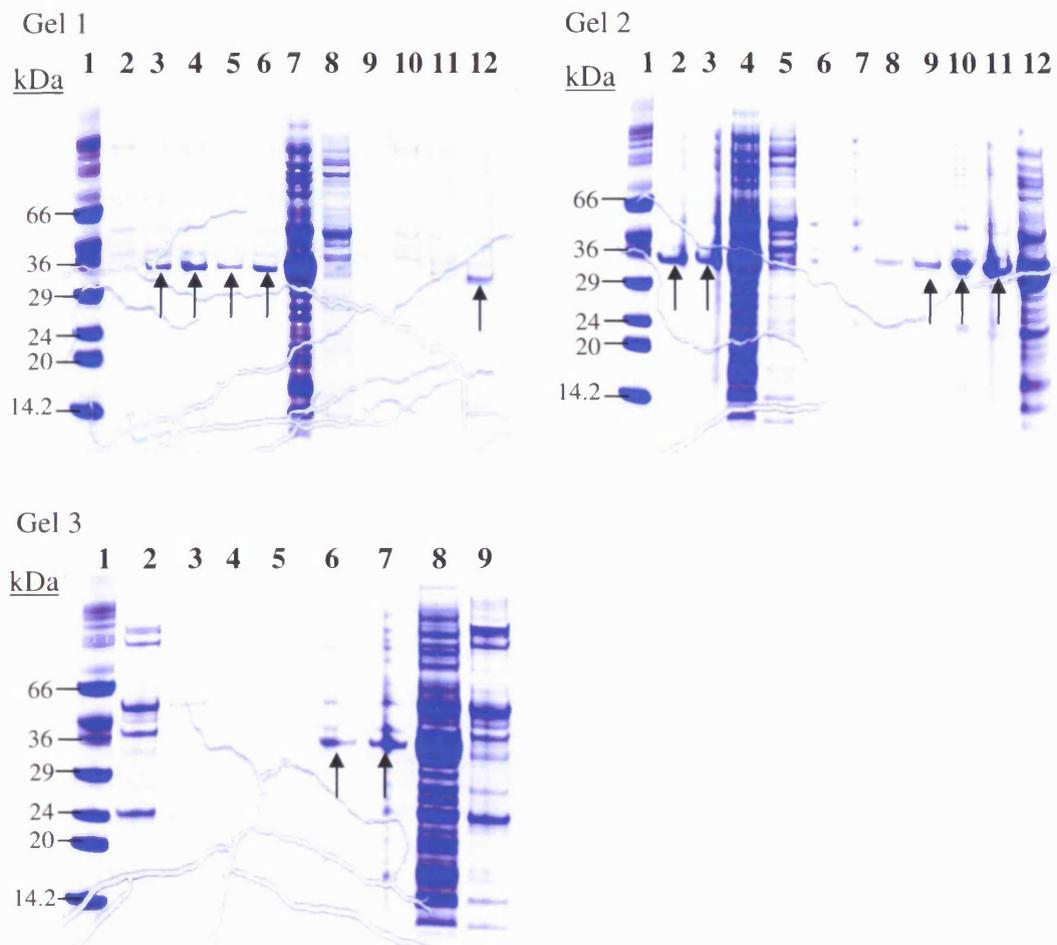
Expression trials of His-tagged p300 1726-1839 were conducted in BL21 (DE3) and BL21-CodonPlus<sup>®</sup> (DE3) RP transformed cells grown at 37 °C. In both instances expression of His-tagged p300 1726-1839 was clearly evident in post-induction samples between 1 and 4 hours (Figure 2.19). However the level of expression observed in BL21 (DE3) is slightly less than that seen in BL21-CodonPlus<sup>®</sup> (DE3) RP cells and the proteins appears to be proteolysed. His-tagged p300 1726-1839 protein generated in both expression trials were judged to be insoluble when comparing whole cell lysate and lysate supernatant samples by SDS PAGE.



**Figure 2.19.** SDS PAGE gels displaying samples taken from a His-tagged p300 1726-1839 protein expression trial in BL21 (DE3) (A) and BL21-CodonPlus<sup>®</sup> (DE3) RP cells (B) grown at 37 °C. Lanes 3 to 6 for each gel represents 1 to 4 hour post-induction samples taken from each trial with His-tagged p300 1726-1839 expression highlighted by the arrows. Un-induced protein samples are shown in lane 2, with whole cell lysate and lysate supernatant samples exhibited in lanes 7 and 8 respectively. Lane 1 contains Wide range molecular marker (Sigma).

### 2.3.11 His-tagged cyclin D1 expression trials

Expression trials of His-tagged cyclin D1 were conducted in *E. coli* BL21 (DE3) and BL21-CodonPlus<sup>®</sup> (DE3) RP cells grown at both 37 °C and 25 °C (figure 2.20). Induction and expression of the His-tagged protein was clearly evident when viewing SDS PAGE gels of post-induction samples from each cell type at both temperatures. The absence of soluble His-tagged cyclin D1 in the lysate supernatant samples from each cell type indicates the protein forms insoluble inclusion bodies. Growth at the lower 25 °C temperature failed to promote the production of soluble protein in both cell types.

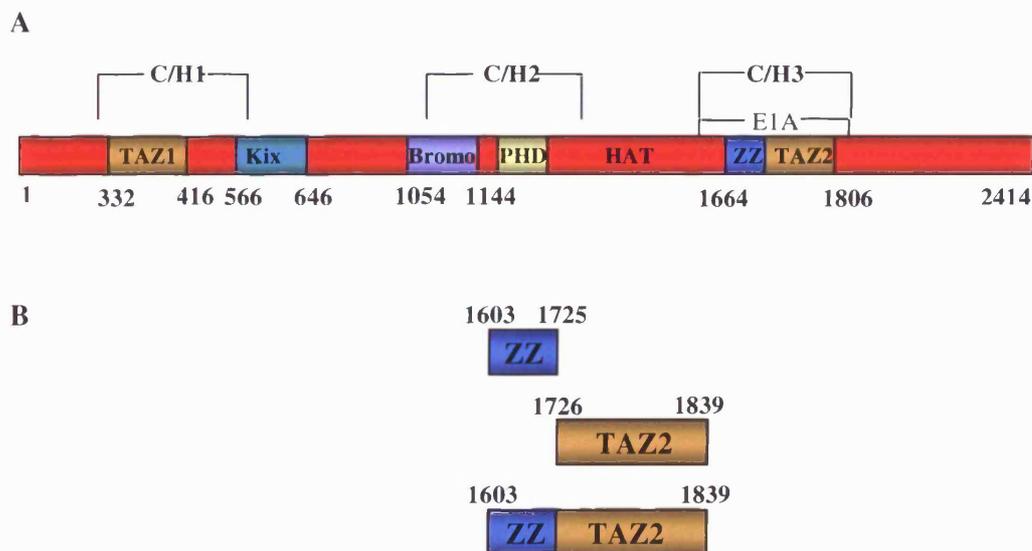


**Figure 2.20.** Coomassie stained SDS PAGE gels showing the results of the His-tagged cyclin D1 expression trials. In Gel 1, lanes 3 to 6 contain 1 to 4 hour post-induction whole cell samples respectively taken from *E. coli* BL21 (DE3) transformed cells grown at 37 °C. There is clearly induction of a ~32 kDa protein when comparing expression with that seen in un-induced whole cells in lane 2. Lane 1 contains wide range molecular marker. Lanes 11 and 12 from gel 1 and lanes 2 and 3 from gel 2 again show the expression of His-tagged cyclin D1 in 1 to 4 hour post-induction whole cell samples taken from *E. coli* BL21 (DE3) transformed cells grown at 25 °C. Lanes 4 and 5 from gel 2 contain the corresponding whole cell lysate and lysate supernatant samples indicating no significant improvement in the solubility of the protein at this temperature. A representative expression trial of His-tagged cyclin D1 expressed in BL21-CodonPlus® (DE3) RP cells at 37 °C is displayed in gel 2. Specifically, lanes 8 to 11 contain 1 to 4 hour post-induction whole cell samples respectively, with lane 7

corresponding to un-induced cells. Lane 12 of gel 2 and lane 2 of gel 3 correspond to whole cell lysate and lysate supernatant samples. Finally in gel 3 is shown an expression trial of His-tagged cyclin D1 expressed in BL21-CodonPlus<sup>®</sup> (DE3) RP cells at the lower 25 °C temperature. Lanes 4 to 7 contain 1 to 4 hour post-induction whole cell samples respectively, with lane 3 containing un-induced cells. Whole cell lysate and lysate supernatant samples for the growth trial are displayed in lanes 8 and 9. The arrows highlight induced His-tagged cyclin D1 protein.

### 2.3.12 P300 Protein Expression and Purification

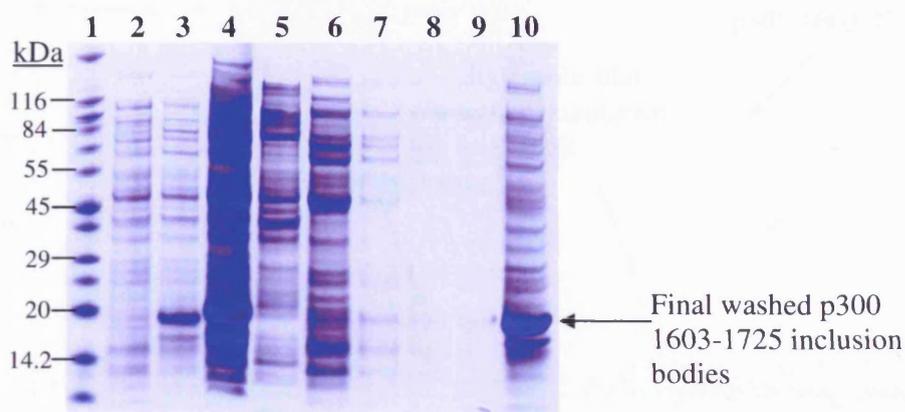
The His-tagged proteins designed and used in this study are depicted in the schematic in figure 2.21.



**Figure 2.21.** A schematic representation of the p300 constructs used in this study. **(A)** A depiction of full length p300 is shown, with various functional regions highlighted including the three cysteine/histidine rich regions (termed C/H1, C/H2 and C/H3) as well as the TAZ, KIX, Bromo, PHD, HAT and ZZ domains. **(B)** N-terminal His-tagged fusions containing the ZZ and TAZ2 regions shown were created.

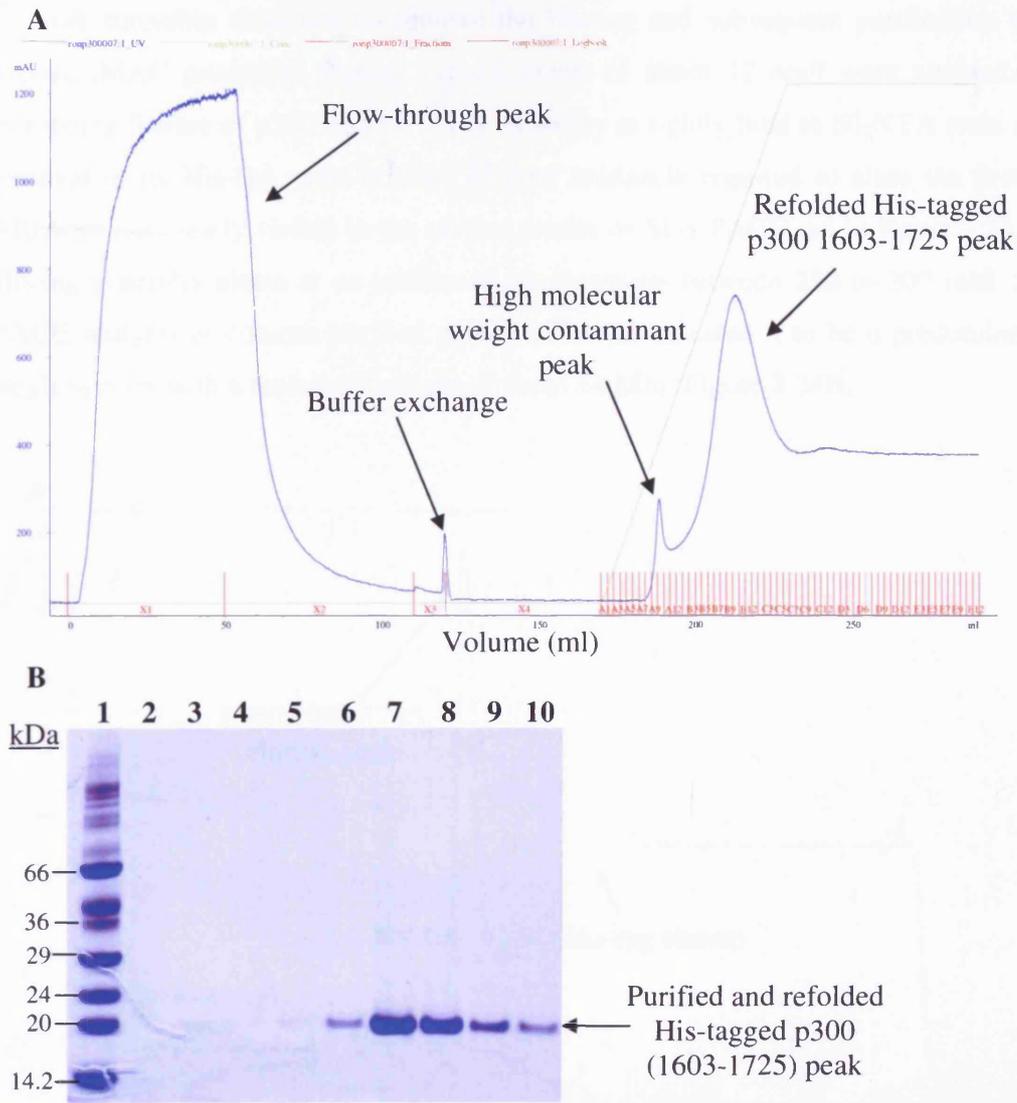
### 2.3.12.1 p300 1603-1725 Expression and Purification

The expression of His-tagged p300 1603-1725 in *E. coli* BL21 (DE3) cells resulted in the production of insoluble inclusion bodies which were subject to an inclusion body wash preparation shown in figure 2.22.



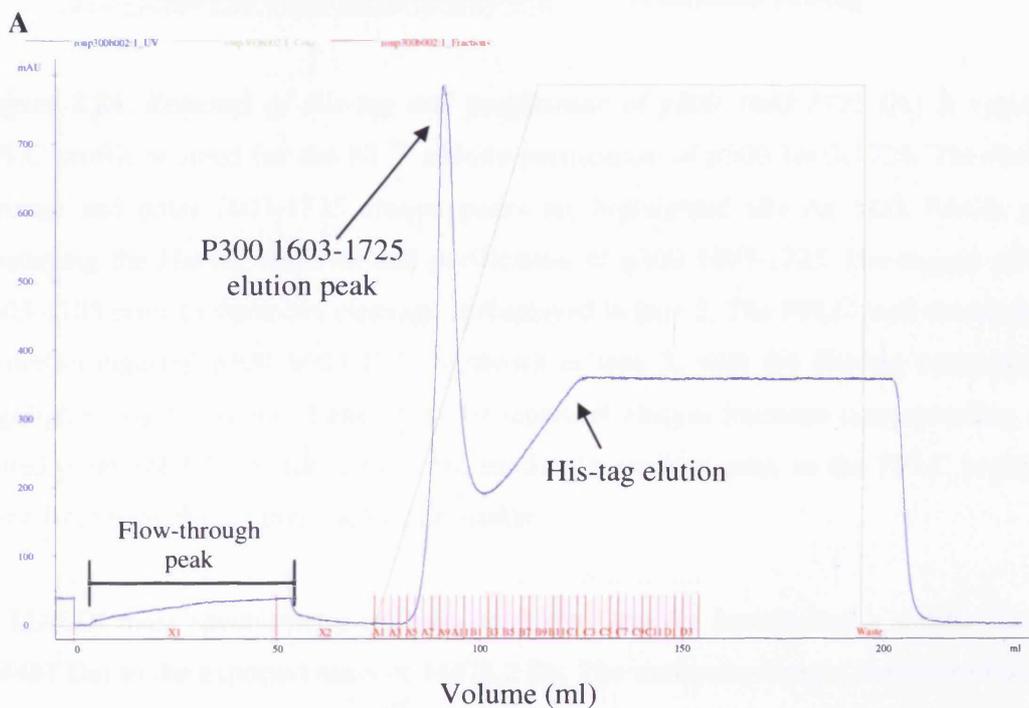
**Figure 2.22.** An SDS PAGE gel displaying the expression and inclusion body preparation of insoluble His-tagged p300 1603-1725. Lane 1, Wide range molecular marker (Sigma); Lane 2, Un-induced whole BL21 (DE3) cells; Lane 3, Whole cells four hours post-induction with IPTG; Lane 4, Whole cell lysate; Lane 5, Cell lysate supernatant sample; Lanes 6-9, consecutive inclusion body wash fractions. Lane 10 contains a sample taken from the final washed inclusion bodies containing enriched p300 1603-1725 protein.

The washed inclusion bodies were readily solubilised in 6 M guanidine hydrochloride, and the His-tagged p300 1603-1725 efficiently refolded (100 %) whilst reversibly bound to the Ni<sup>2+</sup>-NTA affinity column in an IMAC procedure (Figure 2.23A). An SDS PAGE gel of samples removed from the elution gradient revealed the presence of refolded and purified His-tagged p300 1603-1725, which was subsequently pooled and dialysed (Figure 2.23B).

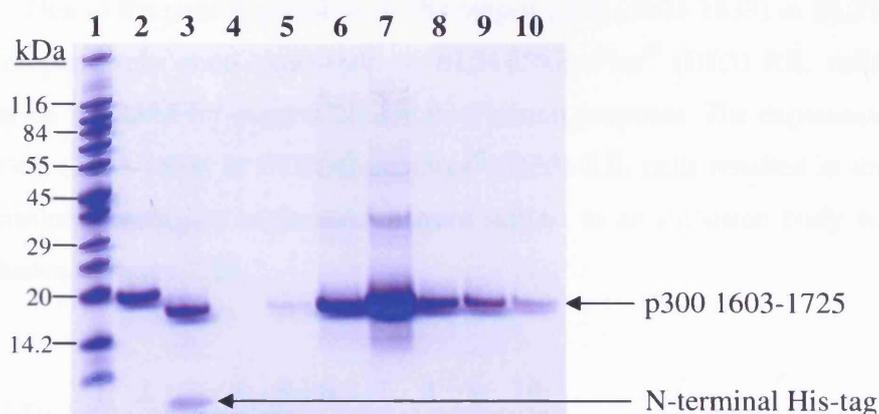


**Figure 2.23.** Purification and refolding of His-tagged p300 1603-1725 (A) A typical FPLC profile attained for the Ni<sup>2+</sup> column based IMAC refolding of His-tagged p300 1603-1725. The flow-through and buffer exchange peaks are highlighted. Elution peaks corresponding to high molecular weight contaminants and refolded protein are also highlighted. (B) An SDS PAGE gel displaying FPLC elution fractions taken from the His-tagged p300 1603-1725 refolding experiment. Lanes 2 to 5 show samples removed from the first of the elution peaks, where lanes 6 to 10 represent refolded His-tagged p300 1603-1725 taken from the second broad elution peak. Lane 1 contains broad range molecular marker (Sigma).

After thrombin treatment to remove the His-tag and subsequent purification by a second IMAC procedure (figure 2.24A), yields of about 12 mg/l were attained. An interesting feature of p300 1603-1725 is its ability to tightly bind to Ni-NTA resin after removal of its His-tag, with at least 75 mM imidazole required to elute the protein. Although not clearly visible in the elution profile or SDS PAGE gel in figure 2.24, the His-tag generally elutes at an imidazole concentration between 250 to 300 mM. SDS PAGE analysis of column purified p300 1603-1725 revealed it to be a predominantly single species with a molecular weight of about 14 kDa (Figure 2.24B).



**B**

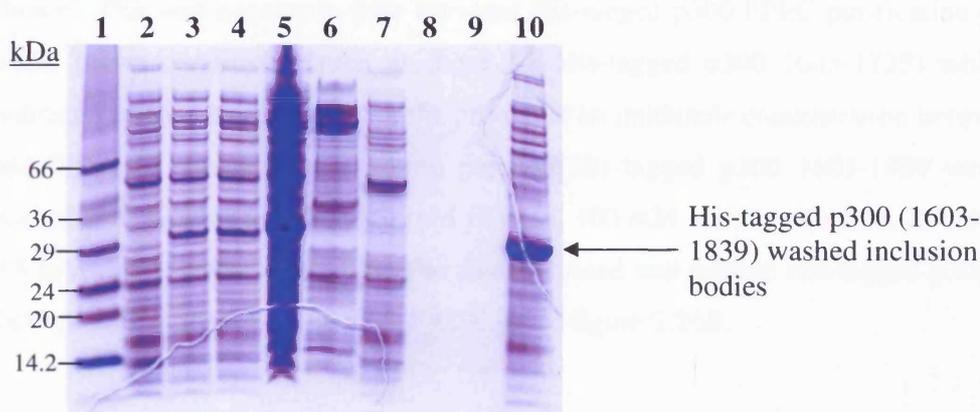


**Figure 2.24.** Removal of His-tag and purification of p300 1603-1725 (A) A typical FPLC profile attained for the Ni<sup>2+</sup> affinity purification of p300 1603-1725. The flow-through and p300 1603-1725 elution peaks are highlighted (B) An SDS PAGE gel displaying the His-tag removal and purification of p300 1603-1725. His-tagged p300 1603-1725 prior to thrombin cleavage is displayed in lane 2. The FPLC load containing thrombin digested p300 1603-1725 is shown in lane 3, with the His-tag component highlighted by the arrow. Lanes 5 to 10 represent elution fractions corresponding to eluted p300 1603-1725, taken from the imidazole gradient peak in the FPLC profile. Lane 1 contains Wide range molecular marker.

MALDI mass spectrometry analysis confirmed that the protein had a similar mass (14487 Da) to the expected mass of 14478.2 Da. The molecular weight determined was limited due to the difficulty in dissolving and preparing the protein with the MALDI Sinapinic Acid matrix. Unfortunately the poor solubility of the protein precluded the possibility of determining the mass by Electrspray Mass Spectroscopy. The protein is very rich in cysteine residues (8) and at high concentrations in the absence of a reducing agent such as DTT the protein has a propensity to form multimers through the formation of intermolecular disulphide bonds. However, comparison of reducing and non-reducing SDS PAGE gels revealed that the monomeric form of p300 (1603-1725) is stable in pH 7.0 buffers containing DTT.

### 2.3.12.2 p300 1603-1839 Expression and Purification

Due to the poor expression of His-tagged p300 (1603-1839) in BL21 (DE3) and the comparatively good expression in BL21-CodonPlus<sup>®</sup> (DE3) RIL cells the latter cell strain was used for expression and purification purposes. The expression of His-tagged p300 (1603-1839) in BL21-CodonPlus<sup>®</sup> (DE3) RIL cells resulted in the production of insoluble inclusion bodies which were subject to an inclusion body wash preparation shown in figure 2.25.

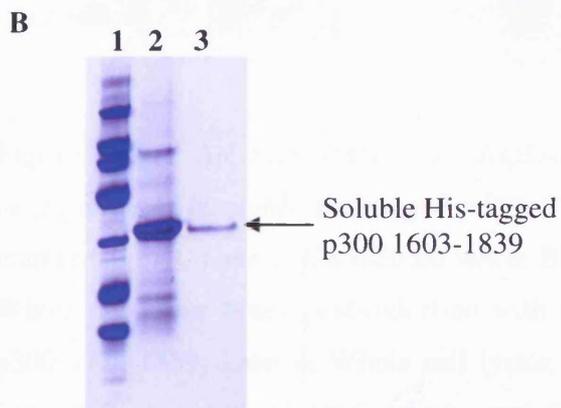
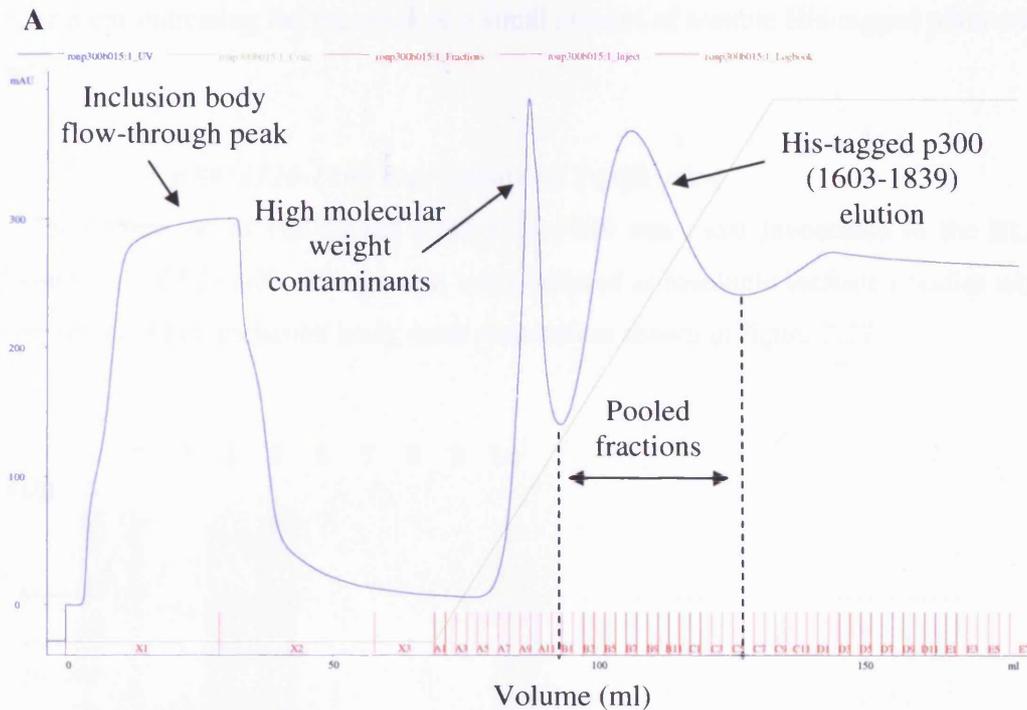


**Figure 2.25.** An SDS PAGE gel displaying the expression and inclusion body preparation of insoluble His-tagged p300 (1603-1839). Lane 1, Wide range molecular marker (Sigma); Lane 2, Un-induced whole BL21-CodonPlus<sup>®</sup> (DE3) RIL cells; Lanes 2 and 3 show post-induction whole cells after 2 and 4 hours respectively. Lane 5, Whole cell lysate; Lane 6, Cell lysate supernatant sample; Lanes 7-9, consecutive inclusion body wash fractions. Lane 10 contains a sample taken from the final washed His-tagged p300 (1603-1839) inclusion bodies.

The Ni<sup>2+</sup> affinity IMAC-based refolding procedure failed to yield soluble protein, resulting in the accumulation of precipitated protein on the Ni-NTA resin. Furthermore refolding of 'non-purified' His-tagged p300 1603-1839 inclusion bodies by dialysis under varying pH's, reducing conditions and detergents was largely unsuccessful with very little or no return of soluble protein (data not shown). The greatest success with His-tagged p300 1603-1839 refolding was achieved with prior purification of inclusion bodies under denaturing conditions, followed by dialysis to remove denaturant. Thus, the washed His-tagged p300 1603-1839 inclusion bodies were solubilised in 6 M

guanidine hydrochloride, and purified to homogeneity under denaturing conditions using a Ni<sup>2+</sup> affinity chromatography procedure (figure 2.26A). The purification of His-tagged p300 1726-1839 in 6M guanidine hydrochloride precludes the use of SDS PAGE to directly analyse samples removed from the elution gradient. However prior treatment of Guanidinium hydrochloride containing samples with TCA (as directed by Novagen) removed the denaturant, and subsequent SDS PAGE analysis revealed the first of the elution peaks in figure 2.26A contained high molecular weight contaminant and the second small and broad peak contains purified His-tagged p300 (1603-1839) (data not shown). This was consistent with Previous His-tagged p300 FPLC purification profiles under native conditions (such as those for His-tagged p300 1603-1725) which had indicated a broad elution peak for the protein at an imidazole concentration between 150 and 200 mM. Fractions containing purified His-tagged p300 1603-1839 were then pooled and refolded against a 20 mM Bis-tris, 100 mM sodium chloride, 20 mM DTT, 0.5 mM ZnCl<sub>2</sub> buffer at pH 7.0. The final refolded and soluble His-tagged p300 1603-1839 protein is shown in the SDS PAGE gel in figure 2.26B.

Unfortunately the purified His-tagged p300 1603-1839 protein had a propensity to form aggregates and precipitate over time. This was particularly apparent when concentrating the sample, precluding the structural characterisation of the protein. Therefore the protein was stored at low concentrations (0.1 mg/ml) to prevent protein aggregation and precipitation forming.

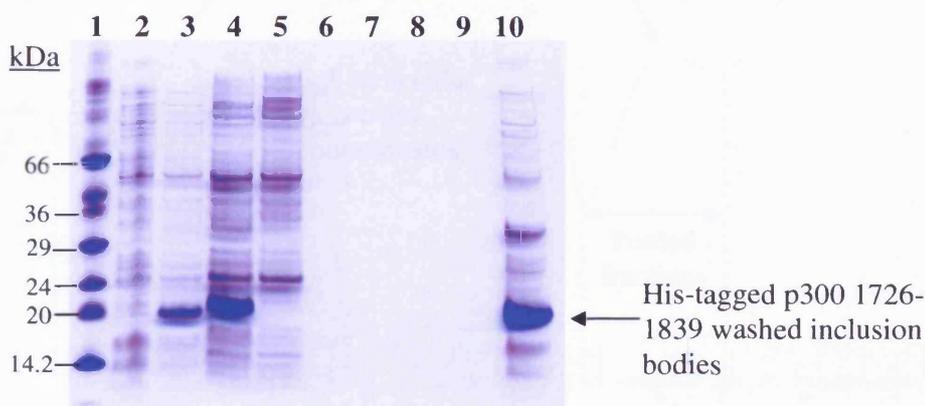


**Figure 2.26.** The purification and refolding of His-tagged p300 1603-1839 from bacterial inclusion bodies. **(A)** A typical FPLC profile attained for the Ni<sup>2+</sup> affinity purification of His-tagged p300 1603-1839 under denaturing conditions. The flow-through, contaminant and His-tagged p300 1603-1839 elution peaks are highlighted. **(B)** An SDS PAGE gel of samples taken from the refolding of His-tagged p300 1603-1839 by dialysis. Lane 2 is an aliquot of precipitated matter retrieved from the refolding/dialysis clearly showing an abundance of insoluble His-tagged p300 1603-1839. In lane 3 is displayed the soluble material retrieved from the refolding/dialysis

experiment indicating the presence of a small amount of soluble His-tagged p300 1603-1839

### 2.3.12.3 p300 1726-1839 Expression and Purification

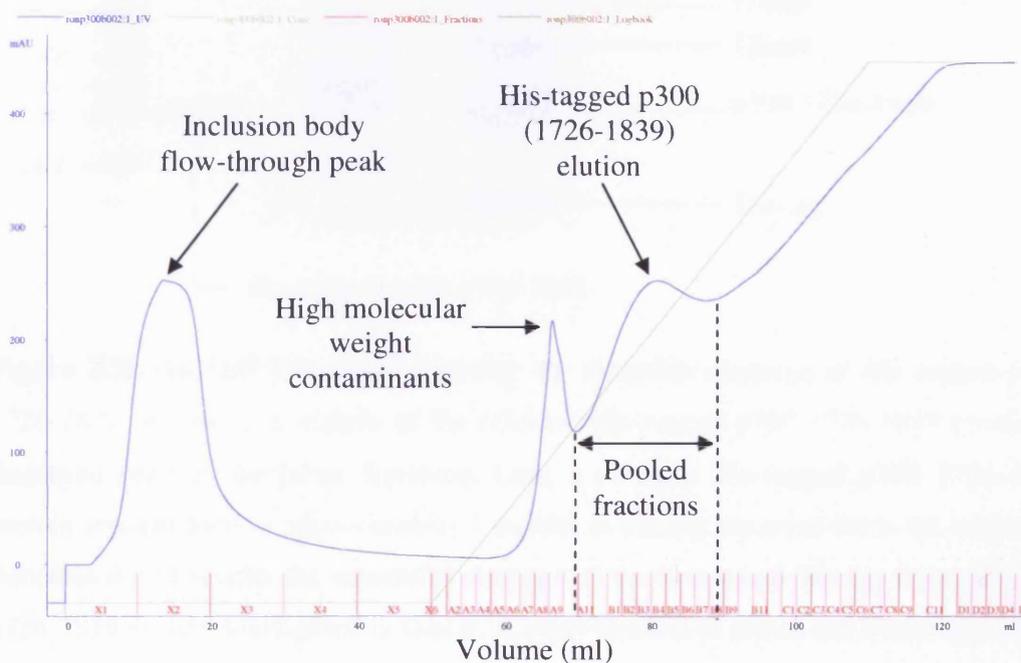
The expression of His-tagged p300 1726-1839 was most favourable in the BL21-CodonPlus<sup>®</sup> (DE3) cells. The protein was produced as insoluble inclusion bodies which were subject to an inclusion body wash preparation shown in figure 2.27.



**Figure 2.27.** An SDS PAGE gel displaying the expression and inclusion body preparation of insoluble His-tagged p300 1726-1839. Lane 1, Wide range molecular marker (sigma); Lane 2, Un-induced whole BL21-CodonPlus<sup>®</sup> (DE3) RP cells; Lane 3, Whole cells four hours post-induction with IPTG showing expression of His-tagged p300 1726-1839; Lane 4, Whole cell lysate; Lane 5, Cell lysate supernatant sample; Lanes 6-9, consecutive inclusion body wash fractions. Lane 10 contains a sample taken from the final washed His-tagged p300 (1726-1839) inclusion bodies.

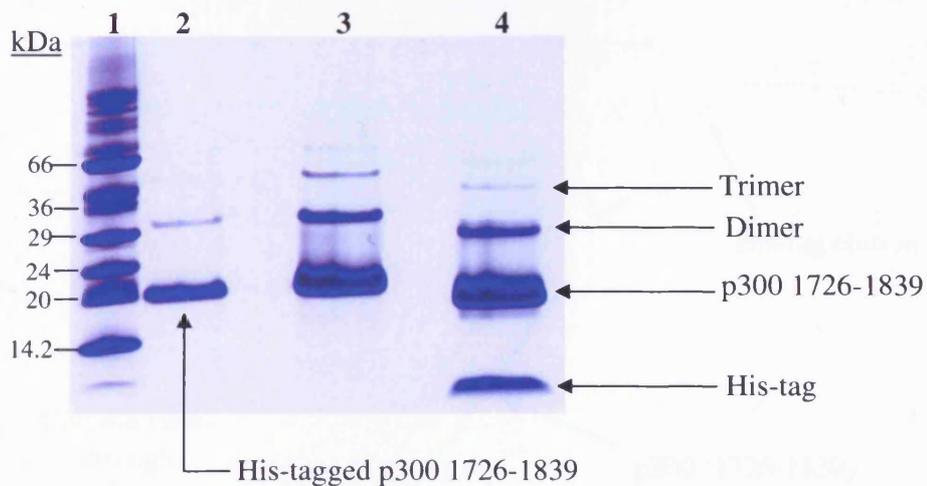
The Ni<sup>2+</sup> affinity IMAC refolding procedure used to successfully refold and purify His-tagged p300 1603-1725 failed to yield soluble protein when the same strategy was applied to His-tagged p300 1726-1839 (Data not shown). The washed His-tagged p300 1726-1839 inclusion bodies were thus solubilised in 6 M guanidine hydrochloride, and purified to homogeneity under denaturing conditions using a Ni<sup>2+</sup> affinity chromatography procedure (figure 2.28). The relatively small His-tagged p300 1726-1839 protein elution peak in the FPLC profile can be attributed to an absence of Tryptophan residues in the protein. The protein does however contain a single Tyrosine

residue which makes a small contribution to the proteins absorbance ( $1209 \text{ cm}^{-1}/\text{M}$ ) when monitored at 280 nm.



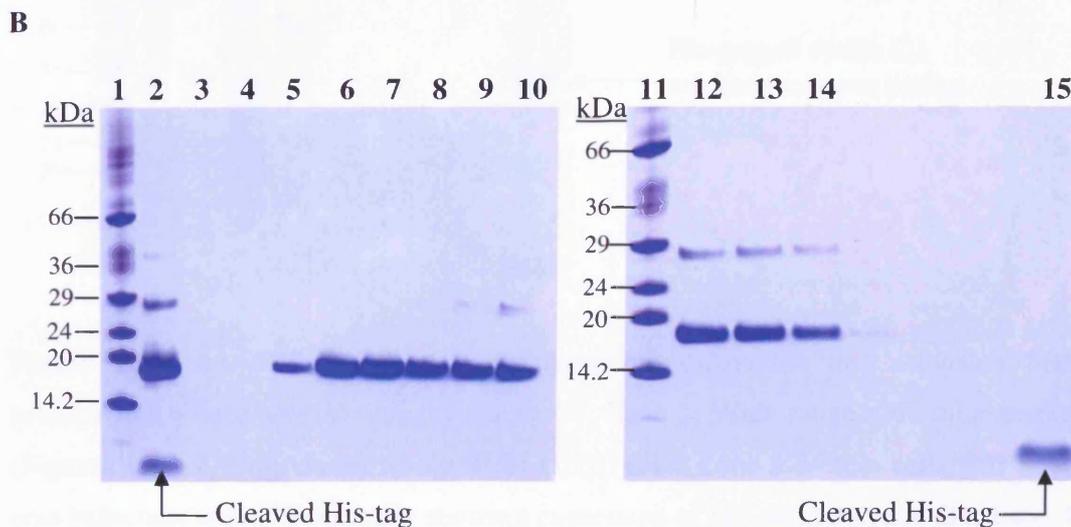
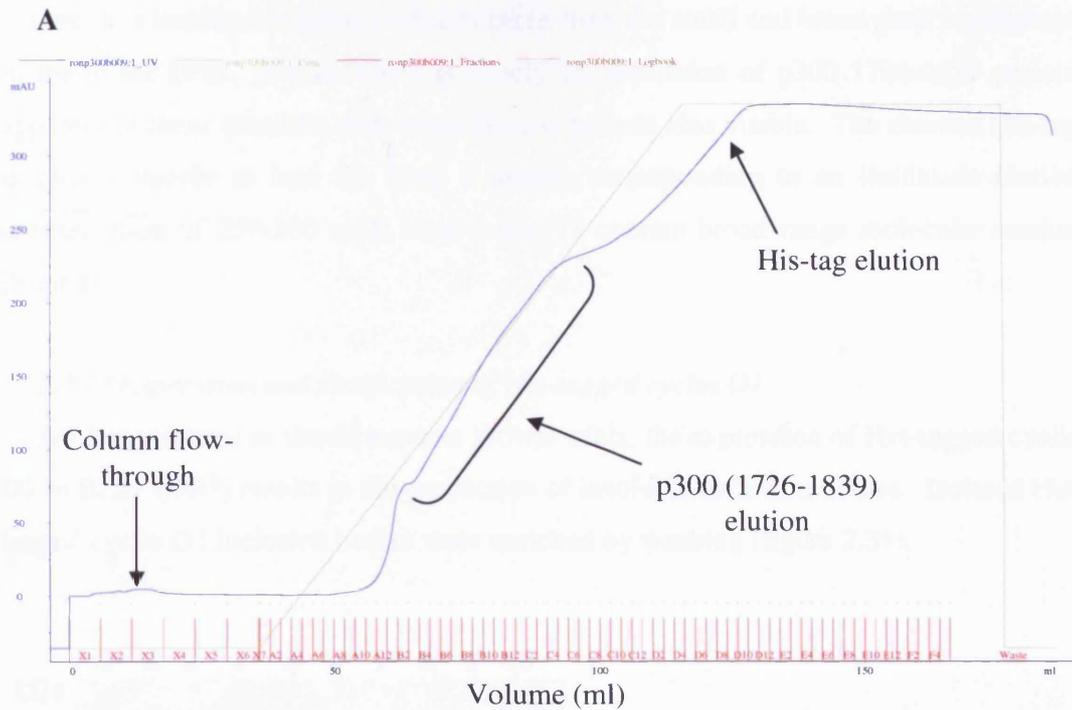
**Figure 2.28.** A typical FPLC profile attained for the  $\text{Ni}^{2+}$  affinity purification of His-tagged p300 1726-1839 under denaturing conditions. The flow-through, contaminant and His-tagged p300 1726-1839 elution peaks are highlighted.

Elution fractions containing His-tagged p300 1726-1839 protein were pooled and refolded by removal of the denaturant by several dialysis steps under reducing conditions in the presence of  $\text{Zn}^{2+}$ . A sample of the final refolded protein is shown in lane 2 of the SDS PAGE gel in figure 2.29. Upon overnight thrombin digestion to remove the N-terminal His-tag a further  $\text{Ni}^{2+}$  affinity purification procedure was implemented to purify the p300 1726-1839 protein. Protein yields of approximately 10 to 15 mg/l were attained.



**Figure 2.29.** An SDS PAGE gel depicting the thrombin cleavage of His-tagged p300 1726-1839. In lane 2, a sample of the refolded His-tagged p300 1726-1839 protein is displayed prior to thrombin digestion. Lane 3 contains His-tagged p300 1726-1839 protein concentrated to approximately 1 mg/ml. A sample removed from the overnight thrombin digest reveals the successful cleavage of the N-terminal His-tag from the p300 1726-1839 protein highlighted in lane 4. A small amount of dimer and trimer species of p300 1726-1839 are also highlighted.

As with p300 1603-1725, in the absence of its His-tag, p300 1726-1839 retains its capacity to bind tightly to Ni-NTA resin with at least 75 mM imidazole required to elute the protein. SDS PAGE gel analysis from samples removed from the FPLC purification clearly demonstrate the successful separation of the cleaved his-tag from the purified p300 1726-1839 protein (figure 2.30). The p300 (1726-1839) protein is revealed to be largely a single species with a molecular weight of about 13 kDa (Figure 2.30b).

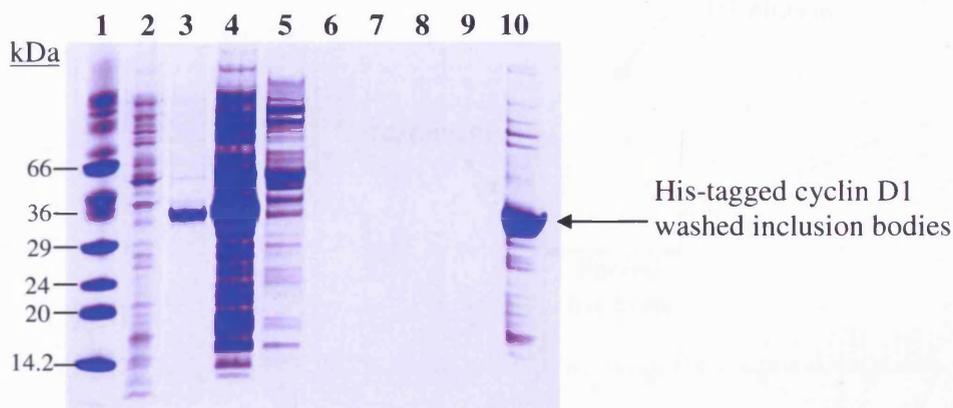


**Figure 2.30.** His-tag removal and purification of p300 1726-1839 (A) A typical FPLC profile attained for the Ni<sup>2+</sup> affinity purification of p300 1726-1839 under native conditions. The flow-through, high molecular weight contaminants and p300 1726-1839 elution peaks are all highlighted. (B) An SDS PAGE gel detailing the His-tag removal and purification of p300 1726-1839. The thrombin digested column load is shown in lane 2 comprising the p300 1726-1839 protein and the cleaved His-tag. Lanes 3 and 4 contain samples removed from the column flow-through, which unsurprisingly contains little or no visible protein. Lanes 5 to 10 and 12 to 14 represent samples removed from

consecutive imidazole elution fractions taken from the small and broad peak highlighted in the in the FPLC profile. There is clearly an abundance of p300 1726-1839 protein apparent in these fractions with some dimeric protein also visible. The cleaved His-tag is clearly visible in lane 15, from a sample corresponding to an imidazole elution concentration of 250-300 mM. Lane 1 and 11 contain broad range molecular marker (Sigma).

### 2.3.13 Expression and Purification of His-tagged cyclin D1

As demonstrated in the timecourse growth trials, the expression of His-tagged cyclin D1 in BL21 (DE3) results in the production of insoluble inclusion bodies. Isolated His-tagged cyclin D1 inclusion bodies were enriched by washing (figure 2.31).

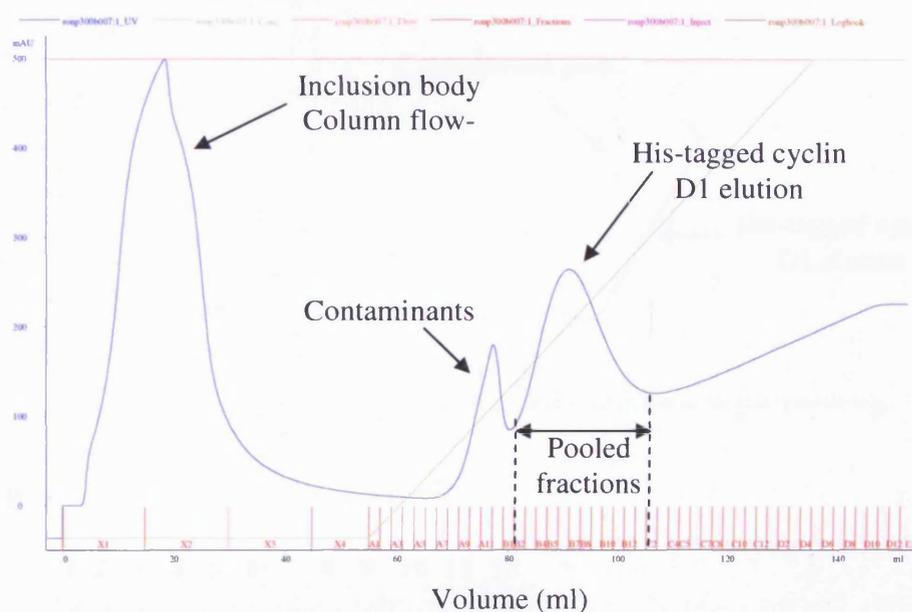


**Figure 2.31.** An SDS PAGE gel displaying the expression and inclusion body preparation of insoluble His-tagged cyclin D1. Lane 1, Wide range molecular marker (Sigma); Lane 2, Un-induced whole BL21 (DE3) cells; Lane 3, Whole cells four hours post-induction with IPTG clearly showing expression of His-tagged cyclin D1; Lane 4, Whole cell lysate; Lane 5, Cell lysate supernatant sample; Lanes 6-9, consecutive inclusion body wash fractions. Lane 10 contains a sample taken from the final washed and enriched His-tagged cyclin D1 inclusion bodies.

The washed His-tagged cyclin D1 inclusion bodies were successfully purified and refolded by one of two methods-

(1) Inclusion bodies solubilised in 6 M guanidine hydrochloride were purified to homogeneity under denaturing conditions using a Ni<sup>2+</sup> affinity chromatography

procedure (figure 2.32). Eluted His-tagged cyclin D1 protein was pooled and successfully refolded at a concentration of 0.2 mg/ml against 3 dialysis buffer changes of 25 mM NaH<sub>2</sub>PO<sub>4</sub>, 100 mM sodium chloride, 1 mM EDTA and 1 mM DTT buffer at pH 8.0. The refolding step is of very high efficiency with between 90-95 % recovery of refolded protein achieved. From a starting culture volume of 1 litre, approximately 35 to 40 mgs of His-tagged cyclin D1 was successfully purified and refolded using this protocol. The final refolded protein is shown in figure 2.33B (lane 12), demonstrating the high sample purity and refolding efficiency observed for this protein preparation.



**Figure 2.32.** A representative FPLC elution profile attained for the Ni<sup>2+</sup> affinity purification of His-tagged cyclin D1 under denaturing conditions. The flow-through peak comprising mostly contaminating protein and DNA is highlighted. Further protein contaminates elute at an imidazole concentration of about 80 mM with His-tagged cyclin D1 eluting at 150 mM.

(2) Alternatively the washed His-tagged cyclin D1 inclusion bodies solubilised in 6 M guanidine hydrochloride, were refolded whilst reversibly bound to the Ni<sup>2+</sup> affinity column in an IMAC procedure (Figure 2.33A). An SDS PAGE gel of samples removed from the elution gradient revealed the presence of refolded and purified His-tagged cyclin D1 (Figure 2.33B), which was subsequently pooled and dialysed against 3 buffer changes of 25 mM NaH<sub>2</sub>PO<sub>4</sub>, 100 mM NaCl, 1 mM EDTA and 1 mM DTT buffer at pH 8.0. Typically there is a His-tagged cyclin D1 protein recovery approaching 60 %



comprising a broad elution peak between 120 mM and 150 mM imidazole concentrations. **(B)** An SDS PAGE detailing the 'on-column' refolding and purification of His-tagged cyclin D1 from washed inclusion bodies solubilised in 6 M guanidine-HCl. Lanes 5 to 11 contain purified and refolded His-tagged cyclin D1 found within consecutive imidazole elution fractions. Lanes 2 to 4 show samples removed from the first imidazole elution peak which appears to containing mainly high molecular weight contaminants. Lane 12 contains His-tagged cyclin D1 protein refolded by dialysis after prior purification under denaturing conditions.

#### 2.4 Discussion

The results presented here clearly demonstrate it is possible to clone and effectively purify soluble B-Myb polypeptides as fusions to glutathione S-transferase (GST) in sufficient quantities for use in biochemical studies and characterisation. A common feature observed amongst the *E. coli*-expressed B-Myb fusions is the high level of proteolysis, resulting in the purification of C-terminally truncated GST/B-Myb polypeptides. The exceptions to this include the GST/B-Myb 275-376, GST/B-Myb 318-458 and GST/B-Myb R2R3 fusion proteins, where the proteolysis observed was minimal in comparison. Most cleavage appeared to be occurring in the *E. coli* cells during growth as observed in SDS PAGE gels of post-induction cell samples. The GST/B-Myb proteins proved highly susceptible to further proteolysis upon cell lysis and purification despite the use of EDTA and PMSF to inhibit proteases. The excessive level of proteolysis observed for most of the GST/B-Myb proteins precluded the full purification and investigation of the B-Myb polypeptides with their GST tags removed.

Since protease sensitive regions are often associated with flexible, non-ordered regions of proteins this would suggest that the central region of B-Myb (amino acids 227 to 508) is inherently unstructured due to the high levels of proteolysis occurring. The amount of full-length protein species purified for each GST/B-Myb fusion was deemed sufficient for use in protein-protein interaction studies since *E. coli*-expressed and purified GST/B-Myb 227-508 protein (in a proteolysed state) was successfully used in pull-down assays to demonstrate binding to cyclin D1 (Horstmann *et al.*, 2000) and p300 (Schubert *et al.*, 2004) highlighting the validity of using such proteins.

In contrast to B-Myb, recombinant p300 expression resulted in the formation of insoluble inclusion bodies, and in some cases required specific *E. coli* strains for successful expression. The presence of ten rare codons (3 Proline CCC, 4 Arginine AGG/AGA, and 3 Leucine CUA) within the p300 residues 1603 to 1839 explains the comparatively poor level of expression observed for p300 1603-1839 protein in BL21 (DE3) cells as compared to the BL21 (DE3) codon plus cells. Decreasing growth temperature appeared not to improve soluble yields for any of the p300 polypeptides. It is of interest to note that several GST/p300 fusion proteins spanning similar regions of the p300 sequence were also found to be insoluble (data not reported) highlighting the propensity of this C/H3 region to form inclusion bodies upon expression in *E. coli*. The solution structure of the CBP TAZ2 domain (amino acids 1764-1850) was solved using a protein that was refolded from inclusion bodies (DeGuzman *et al.*, 2000) though the authors fail to elaborate on the precise refolding procedure and conditions used. The CBP ZZ domain (amino acids 1700-1751) solution structure was determined from soluble protein (with cells grown at 15 °C) and thus bypassed the need for refolding (Legge *et al.*, 2004).

In this study, the IMAC-based refolding procedure was only successful for His-tagged p300 1603-1725, providing an efficient way of attaining soluble protein with good yields. Both His-tagged p300 1603-1839 and His-tagged p300 1726-1839 proteins failed to refold by this method and precipitated on the Ni-NTA resin during the chromatography procedure despite varying pH and buffer conditions. It was however possible to refold both His-tagged p300 1726-1839 and p300 1603-1839 proteins by dialysis under reducing conditions with excess Zn<sup>2+</sup> although the yields of soluble His-tagged p300 1603-1839 protein attained were extremely poor.

Upon purification and refolding, both His-tagged p300 1603-1725 and p300 1726-1839 proteins appeared more prone to precipitation and aggregation than their tag-less variants. A possible explanation for this phenomenon is that the hexa-histidine tag, a potent metal binding sequence, may have saturated free Zn<sup>2+</sup> and may have prevented the correct coordination of Zn<sup>2+</sup>. p300 1603-1725 was maintained in a reduced form with excess DTT and remained stable for periods of up to 15 days at 4 °C though was prone to aggregation (particularly in the absence of reducing agent) and formed visible fibre-like precipitates beyond this period. Over time it is highly probable that the protein

would have oxidised and the cysteine residues in p300 1603-1725 not involved in  $Zn^{2+}$  ion coordination would form disulphide bonds, promoting the assembly of multimeric protein species. In fact such dimer, trimer and high order multimeric protein species have been observed on non-reducing SDS PAGE gels of p300 1603-1725 (data not shown).

Interestingly p300 1603-1725 forms flat crystal-like matrices during concentration in a filtration device. As a result it was often difficult to achieve concentrations of p300 1603-1725 beyond 2.0 mg/ml. The resultant crystals failed to diffract in an X-ray beam thus precluding the detailed crystallographic study of the polypeptide. Crystals suited to X-ray crystallography studies are composed of tightly packed and highly coordinated unit cells which grow over an extended period from days to often months. The fast nature of p300 1603-1725 crystal formation during centrifugation does therefore not lend itself to the production of suitable crystals for X-ray diffraction. Crystal trials using purified p300 1603-1725 protein at approximately 2 mg/ml concentration were set up at 4 °C and 20 °C but failed to produce crystals under the growth conditions used (Data not reported). The low protein concentration used in the crystal trials may have been a limitation in this study, with typically 10 mg/ml required for crystal formation. Furthermore, over time protein oxidation would promote both inter- and intramolecular disulphide exchange and thus may contribute to protein aggregation which would seriously hamper crystal formation.

In contrast to p300 1603-1725, the p300 1726-1839 protein was stable to concentrations in excess of 2 mg/ml and did not require the presence of excess reducing agent in the final buffer to maintain stability. However over extended periods the protein formed small amounts of precipitate in a concentration dependent manner. The His-tagged cyclin D1 protein was stable to concentrations of approximately 2-3 mg/ml in 25 mM  $NaH_2PO_4$ , 100 mM sodium chloride buffer at pH 8.0.

To summarise, the B-Myb, p300 and cyclin D1 proteins/polypeptides produced in this study were effectively purified in sufficient quantities for use in protein-protein interaction studies and biophysical characterisation.

## CHAPTER 3

### STRUCTURAL CHARACTERISATION OF PURIFIED B-MYB, P300 AND CYCLIN D1 POLYPEPTIDES

#### *3.1 Introduction and aims*

In order to further investigate the reported interactions of a central region of B-Myb with p300 and cyclin D1 (Horstmann *et al.*, 2000b; Schubert *et al.*, 2004), a series of GST/B-Myb fusions and His-tagged p300 and cyclin D1 polypeptides were expressed and purified (see chapter 2). This chapter reports the structural characterisation of the purified polypeptides using biophysical techniques such as intrinsic tryptophan fluorescence and far UV circular dichroism spectroscopy. Of the GST/B-Myb fusion polypeptides purified only GST/B-Myb 275-376 was free of excessive proteolysis and was thus amenable to GST-tag cleavage and structural characterisation.

#### *3.2 Materials and Methods*

##### *3.2.1 Fluorescence spectroscopy*

A Perkin Elmer LS50 Biluminescence spectrometer was used to acquire intrinsic tryptophan fluorescence spectra of p300, B-Myb, and cyclin D1 proteins. The spectra were recorded at 25 °C with excitation at 280 nm and fluorescence monitored from 300 to 450 nm. Spectra were acquired using a scan rate of 150 nm per minute, with 10 accumulations averaged to give the final spectra. Typically, the spectra were acquired with 1.0 to 3.0  $\mu$ M protein samples using a 1 cm path length quartz fluorescence cuvette.

##### *3.2.2 Circular dichroism (CD) spectroscopy*

The secondary structures of isolated domains/regions from purified B-Myb, p300 and cyclin D1 proteins were determined by analysing far UV CD spectra acquired on a Jasco 715 spectrometer. Where possible the concentrations of proteins were adjusted to give an absorbance at 280 nm of typically about 1.0 for a path length of 1 cm. CD spectra were recorded from 180 to 250 nm at a scan rate of 20 nm per minute, with each spectrum representing the average of 10 accumulations. Data were acquired at 25 °C with either 0.1 mm or 2 mm path length cells. Estimates of secondary structure

composition were calculated from buffer corrected spectra using the CDPro package (Sreerama and Woody, 2000).

### *3.2.3 Chemical denaturation of p300 and cyclin D1 polypeptides*

Intrinsic tryptophan fluorescence and far UV CD techniques were used to probe the conformational stability of the p300 1603-1725 and His-tagged cyclin D1 polypeptides in the presence of guanidine hydrochloride. Fluorescence experiments with p300 1603-1725 were conducted with a series of 3  $\mu$ M samples in a 20 mM Bis-tris, 100 mM sodium chloride, 0.5 mM DTT, and 0.5 mM  $\text{ZnCl}_2$  buffer at pH 7.0, containing between 0 and 6 M guanidine hydrochloride, increasing with 0.5 M equivalents of denaturant. Fluorescence experiments with His-tagged cyclin D1 were similarly conducted with a series of 1  $\mu$ M samples in a 25 mM  $\text{NaH}_2\text{PO}_4$ , 100 mM sodium chloride buffer at pH 8.0 containing increasing equivalents (0.5 M) of denaturant. CD experiments with p300 1603-1725 were conducted with a series of 70  $\mu$ M samples in a 20 mM Bis-tris, 100 mM sodium chloride, 20 mM DTT, and 0.5 mM  $\text{ZnCl}_2$  buffer at pH 7.0, containing between 0 and 6 M guanidine hydrochloride, increasing with 0.5 M equivalents of denaturant. All samples were incubated with denaturant for 30 minutes before spectra were recorded at 25  $^\circ\text{C}$  using a 0.1 mm pathlength cell.

### *3.2.4 Temperature denaturation studies of p300 1726-1839*

Far UV CD was employed to monitor possible changes in p300 1726-1839 secondary structure upon increasing sample temperature. CD experiments were collected with 8  $\mu$ M samples in a 20 mM Bis-tris, 100 mM sodium chloride buffer at pH 6.5. CD spectra were recorded over the temperature range 5 to 85  $^\circ\text{C}$  at 10  $^\circ\text{C}$  intervals using a 2 mm pathlength cell with the temperature rising incrementally at 1  $^\circ\text{C}/\text{min}$ . An additional spectrum was recorded after the sample temperature had returned back from 85  $^\circ\text{C}$  to 5  $^\circ\text{C}$ . Changes in the maximum negative peak value at 222 nm were recorded as a function of increasing temperature.

### *3.2.5 $\text{Zn}^{2+}$ binding properties of p300 polypeptides*

Changes in the maximum negative intensity of the CD spectrum of p300 1603-1725, monitored at 212 nm, were used to characterise  $\text{Zn}^{2+}$  binding upon addition of  $\text{ZnCl}_2$ . The p300 1603-1725 polypeptide was purified and refolded in the absence of  $\text{ZnCl}_2$  in order to study the apo-protein form. CD spectra were collected for a series of samples

containing 70  $\mu\text{M}$  p300 1603-1725 with increasing molar equivalents of  $\text{ZnCl}_2$  (0.5, 1, 1.5, 2, 2.5, 5 and 10). Samples prepared in a 20 mM Bis-tris, 100 mM sodium chloride, and 20 mM DTT buffer at pH 7.0 were incubated overnight at 4 °C before the acquisition of CD data.

The dissociation constant ( $K_d$ ) for  $\text{Zn}^{2+}$  binding was estimated by fitting the following equation to the changes in CD intensity at 212 nm, which assumes a simple equilibrium between the peptide with a single bound  $\text{Zn}^{2+}$  and the free species (He *et al.*, 1991)

$$\Delta CD = \frac{\Delta CD_{\infty}}{2P} \left[ P + L + K_d - \left( \{P + L + K_d\}^2 - 4PL \right)^{1/2} \right]$$

In the equation, P and L represent the concentrations of p300 (1603-1725) and  $\text{Zn}^{2+}$  respectively,  $\Delta CD$  and  $\Delta CD_{\infty}$  are the changes in CD intensity at the ligand concentration L and saturating ligand concentration respectively, and  $K_d$  is the dissociation constant of the protein-ligand complex.

CD experiments with p300 1726-1839 were conducted with 64  $\mu\text{M}$  samples in a 20 mM Bis-tris, 100 mM sodium chloride buffer at pH 6.5 in the absence and presence of saturating  $\text{Zn}^{2+}$  (0.5 mM  $\text{ZnCl}_2$ ) and EDTA (5 mM) concentrations. Samples were incubated overnight at 4 °C before the acquisition of CD data.

### 3.2.6 Secondary structure predictions of B-Myb and p300 polypeptides

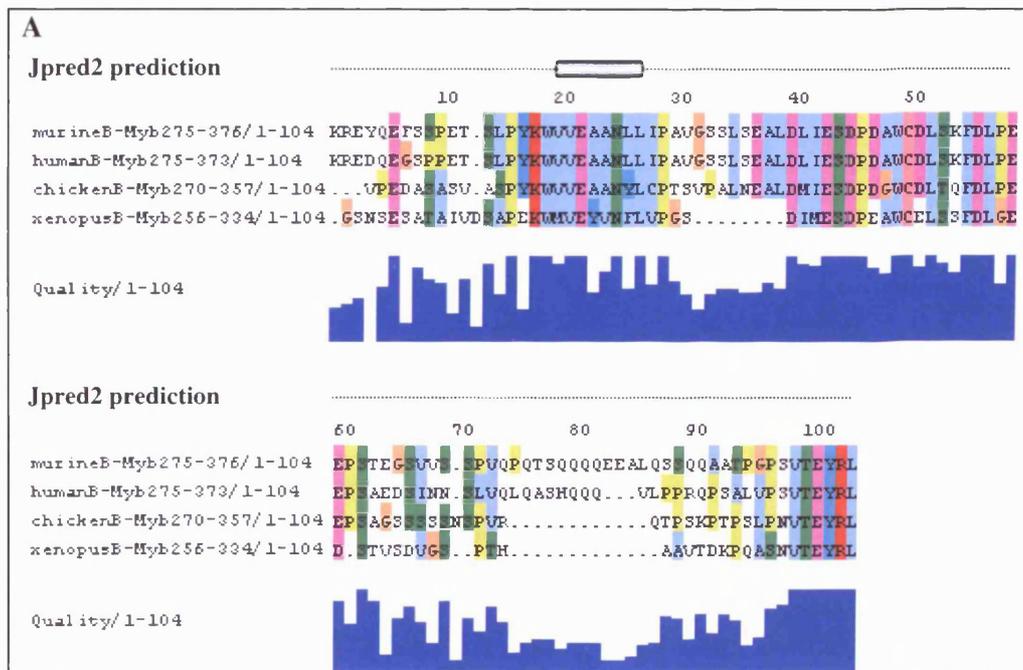
Secondary structure predictions for B-Myb 275-376, p300 1603-1725, and p300 1726-1839 regions were obtained through use of the Jpred2 package (Cuff *et al.*, 1998). Predictions were obtained by either inputting the query protein sequence alone or as part of an optimised user created protein alignment with homologous proteins. Protein alignments were created using ClustalW (available at the European Bioinformatics Institute website, [www.ebi.ac.uk](http://www.ebi.ac.uk)).

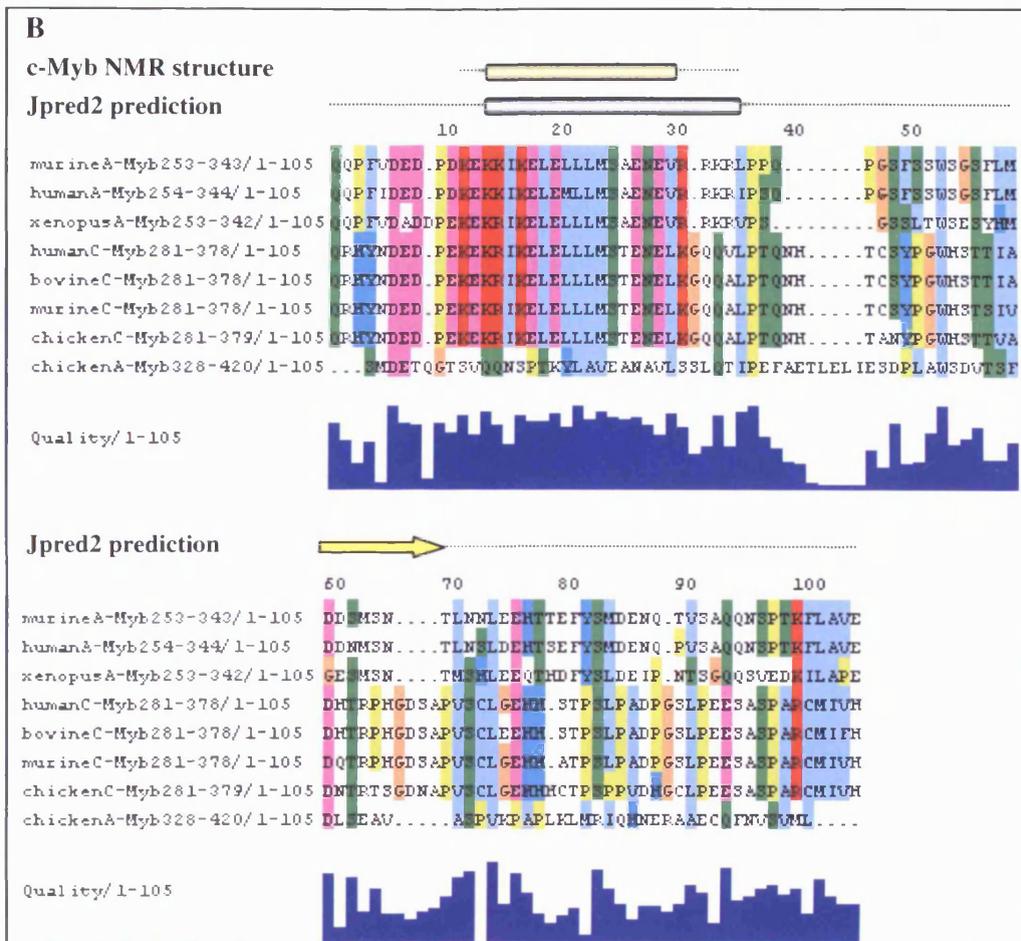
### 3.3 Results

#### 3.3.1 Structural characterisation of B-Myb 275-376

##### 3.3.1.1 Secondary structure prediction

Predictions of secondary structural content obtained for B-Myb 275-376 using the Jpred2 package is shown for an optimised alignment (using ClustalW) of Human, Mouse, Chicken and Xenopus B-Myb proteins (figure 3.1A). This region of B-Myb exhibits some sequence conservation amongst the four genera with 53 % sequence homology (defined as the sum of sequence identity and similarity). The prediction estimates very little secondary structure (7 %) for this region, comprising of a single  $\alpha$ -helix spanning the amino acids VVEAANL (residues 293 to 299) in murine B-Myb.





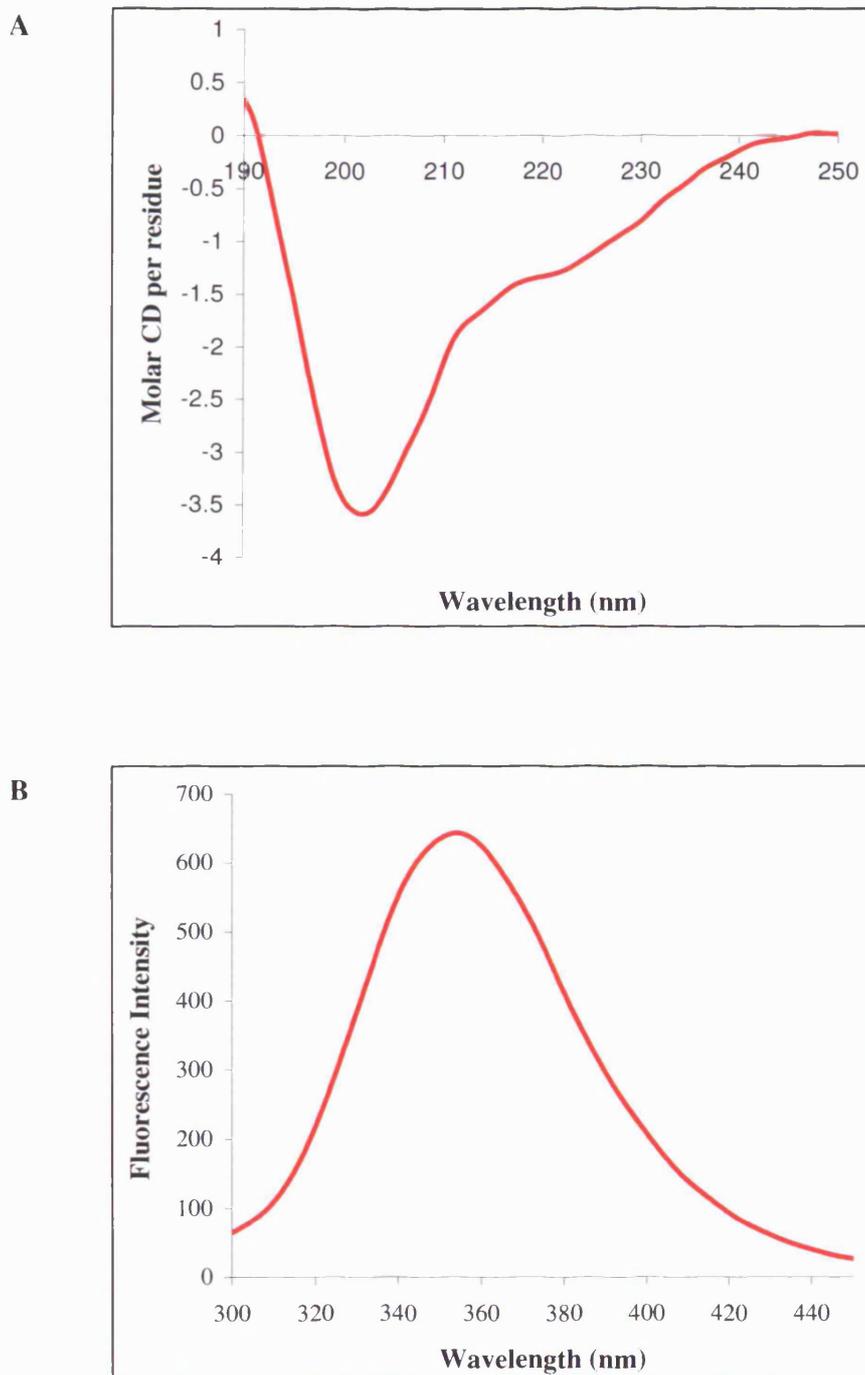
**Figure 3.1 (A)** Jpred2 secondary structure prediction obtained for B-Myb 275-376 and **(B)** the transactivation regions of c- and A-Myb. Homologous regions of B-Myb from Human, Mouse, Chicken and Xenopus were aligned using ClustalW before applying the Jpred2 prediction. The consensus prediction reveals a small amount of  $\alpha$ -helical (tube) secondary structure (7 %) for the B-Myb 275-376 region (shown in A). Comparatively, the homologous region of c- and A-Myb is predicted to adopt both  $\alpha$ -helical (22 %) and  $\beta$ -sheet (arrow) (6 %) secondary structures as shown in B. The NMR derived structure of a short c-Myb polypeptide is also shown for comparison (Zor *et al.*, 2004).

For comparison with B-Myb, a Jpred2 secondary structure prediction for a similar region of A-Myb and c-Myb is shown alongside the structure derived for a short c-Myb polypeptide in complex with the CBP KIX domain (Zor *et al.*, 2004) (see figure 3.1B). B-Myb shares very little sequence homology with the corresponding transactivation regions from A-Myb and c-Myb. Conversely A-Myb and c-Myb share significant

sequence conservation in this region (51 % sequence homology). The prediction revealed a total secondary structure content of 28 % for this region comprised of a single  $\alpha$ -helix (22 %), and a single stretch of  $\beta$ -sheet (6 %). The position of the predicted  $\alpha$ -helix for c- and A-Myb overlaps that derived from the solution structure of the c-Myb transactivation domain and a similar region estimated for B-Myb. Interestingly the small  $\beta$ -sheet region predicted for c- and A-Myb, is completely absent from the predicted B-Myb secondary structure. Although caution must be observed when predicting structural motifs based on primary amino acid sequences alone, inherent sequence homology differences and the Jpred2 secondary structure predictions suggest that the B-Myb transactivation region is likely to differ somewhat from that of c- and A-Myb. This hypothesis is supported by the differing functional roles of each of the Myb proteins as well as the differing co-factor proteins to which the respective transactivation regions bind

#### *3.3.1.2 Biophysical characterisation of B-Myb 275-376: secondary and tertiary structural analysis*

The far UV CD spectrum obtained for B-Myb 275-376 shown in figure 3.2A is typical of that observed for a protein with little or no secondary structure, characterised by the large negative peak at 200 nm. The CDPro (Sreerama and Woody, 2000) analysis of secondary structure confirmed this, estimating a secondary structure content of 9 %  $\alpha$ -helices, 24 %  $\beta$ -sheet, 17 % turns, and 50 % with no regular structure. The fluorescence emission spectrum recorded for B-Myb 275-376, used to characterise tertiary structure, is shown in figure 3.2B. B-Myb 275-376 contains two tryptophan residues (W293 and W323) that contribute to a fluorescence maximum of 355 nm, indicative of tryptophan residues fully exposed to the aqueous solvent, as expected for an unfolded random coil polypeptide.

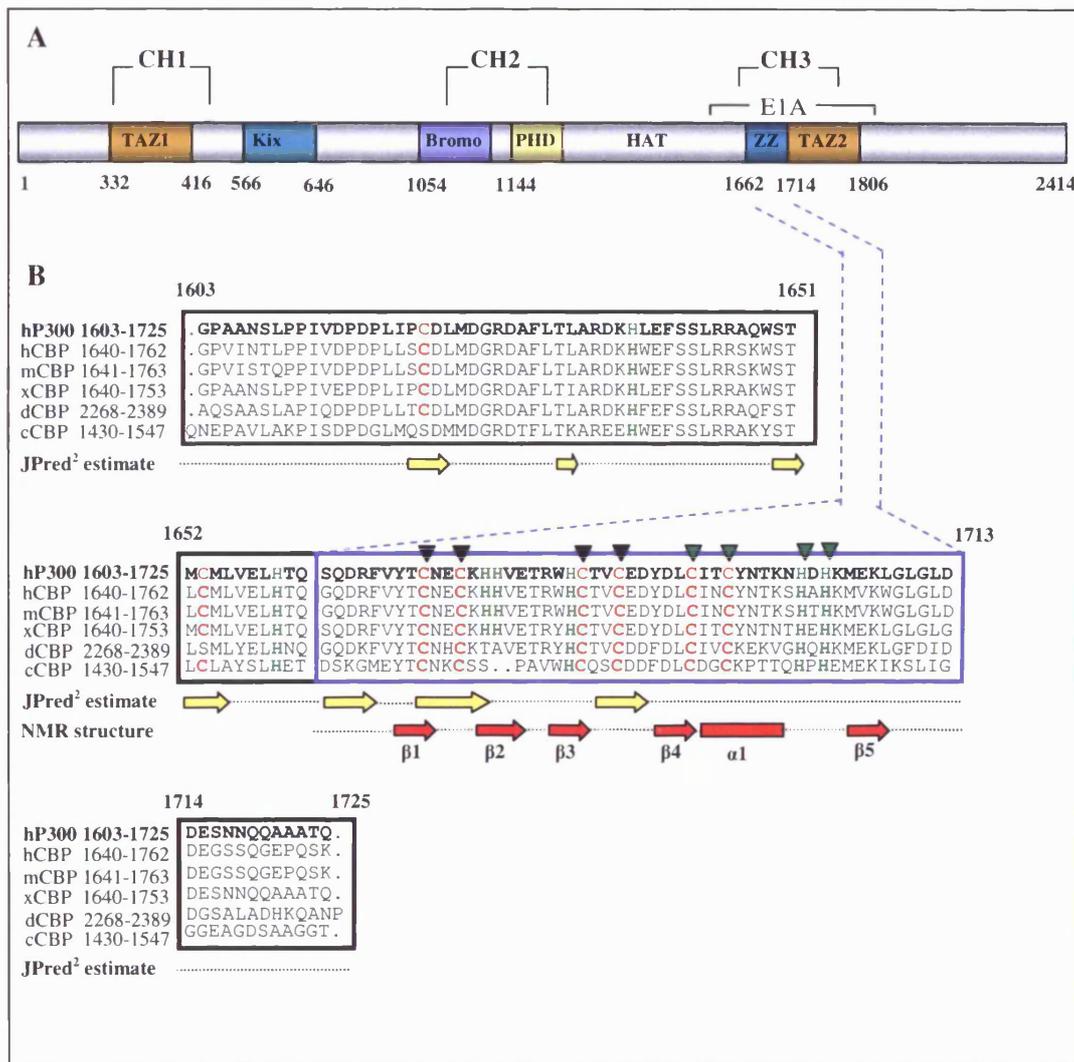


**Figure 3.2.** (A) Representative CD and (B) Intrinsic fluorescence emission spectra obtained at 25 °C for B-Myb 275-376 in a 20 mM Tris, 100 mM sodium chloride buffer at pH 7.0. The CD spectral features are suggestive of a random coil polypeptide with no evidence of a folded structure. The tryptophans of B-Myb 275-376 appear to be fully exposed to the aqueous solvent as indicated by the 355 nm maximum emission wavelength.

### 3.3.2 Structural characterisation of p300 1603-1725

#### 3.3.2.1 Secondary structure prediction

Estimates of secondary structural content obtained for the p300 1603-1725 polypeptide using the Jpred2 package is shown in figure 3.3. For comparison the secondary structure derived from the NMR structure of the murine CBP ZZ (Legge *et al.*, 2004) domain is highlighted also.



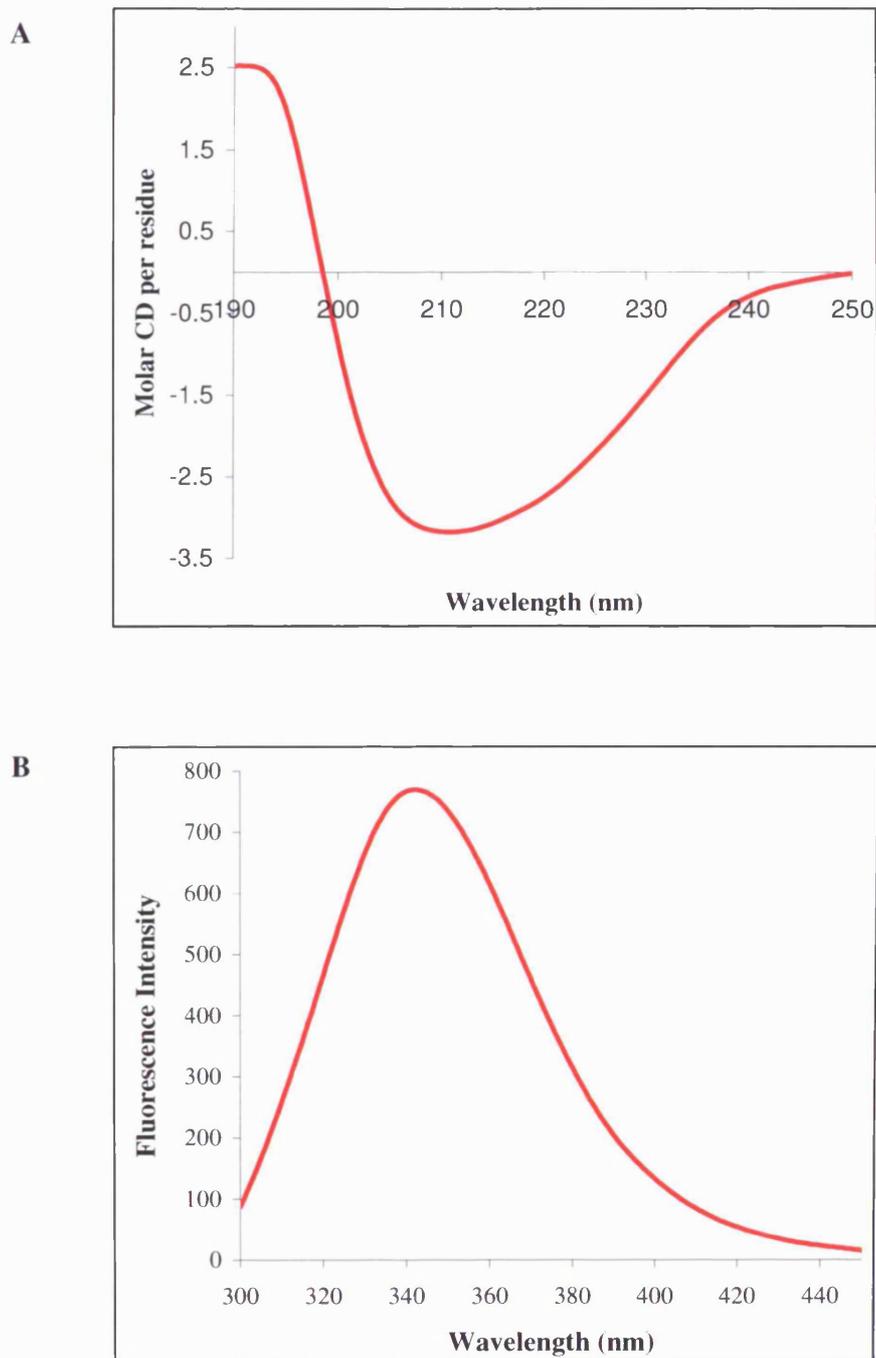
**Figure 3.3** (A) Schematic representation of the domain structure of p300 with the major functional regions characterised to date indicated (B) A multiple sequence alignment of p300/CBP sequences with predicted (Jpred2) and determined (Legge *et al.*, 2004) secondary structural elements. The 52 residue ZZ domain construct corresponding to

residues 1700-1751 of murine CBP used for determining the NMR structure (Legge *et al.*, 2004) is grouped with homologous p300 and CBP sequences in the blue box. Conserved cysteine (red) and histidine (green) residues are highlighted, with those involved as Zn<sup>2+</sup> coordination ligands highlighted with black (Zn1-CCCC) and green triangles (Zn2-CCHH). The arrows ( $\beta$ -strands) and blocks ( $\alpha$ -helices) shown below the sequence alignment indicate the positions of predicted (Jpred2, in yellow) and determined (Legge *et al.*, 2004, in red) secondary structural elements. The polypeptide sequences shown correspond to p300 and CBP from *Homo sapiens*, h; *Mus musculus*, m; *Xenopus laevis*, x; *Drosophila melanogaster*, d; and *Caenorhabditis elegans*, c. Protein accession numbers are as follows: hP300 [Q09472], hCBP [Q92793], mCBP [AAL875231], xCBP [AAH44677], dCBP [AAB53050], cCBP [P34545].

Secondary structure predictions for p300 1603-1725 obtained from the Jpred2 package used input consisting of an optimised alignment of ZZ motif sequences from several p300, CBP, dystrophin and dystrophin related proteins. p300 1603-1725 was predicted to contain six relatively short stretches of  $\beta$  sheet, constituting 24 % of the polypeptide, linked by loops, but no helical regions (figure 3.3B). The Jpred2 estimates of  $\beta$  sheet secondary structure content did not compare favourably with that derived from the solution structure and failed to predict a short helical region (residues 1693 to 1699 in human p300) within the ZZ domain.

### 3.3.2.2 Biophysical characterisation of p300 1603-1725: secondary and tertiary structural analysis

A representative far UV CD spectrum obtained for p300 1603-1725 is displayed in figure 3.4A. Analysis of the p300 1603-1725 CD spectrum using the CDPro package (Sreerama and Woody, 2000) estimated significant  $\beta$ -sheet content (30 %) and some helical secondary structure (17 %). An additional 19 % of its structure is composed of turns with 34 % attributed to unordered regions of the polypeptide. Figure 3.4B displays a representative intrinsic tryptophan fluorescence spectrum obtained for p300 1603-1725 which contains two tryptophan residues (Y1653 and Y1685). The fluorescence emission maximum of 342 nm indicates that one or both of the tryptophan residues are at least partially shielded from the aqueous solution, indicative of folding of the domain.

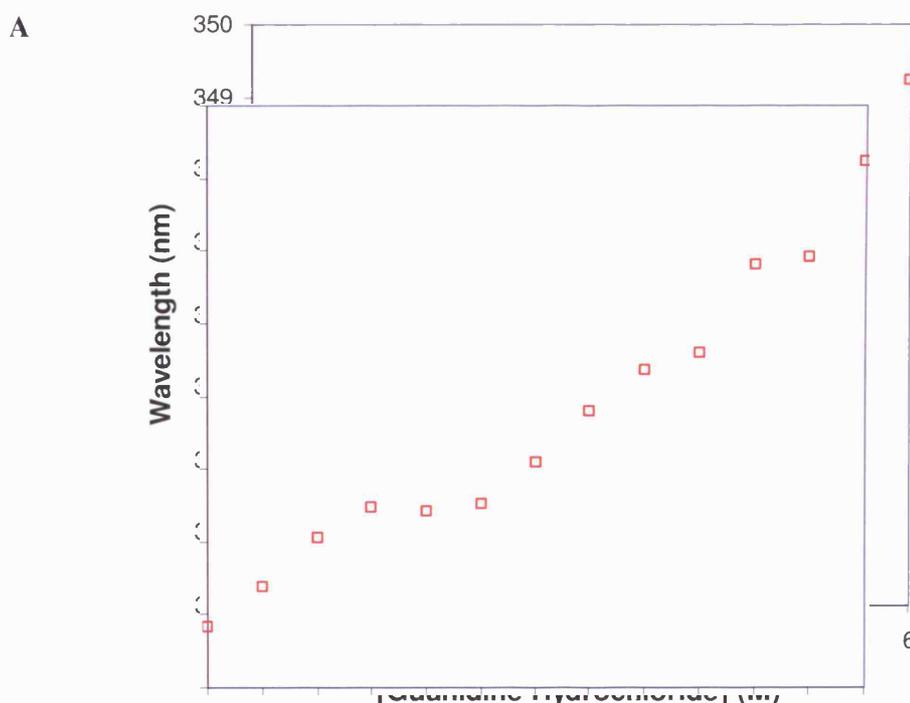


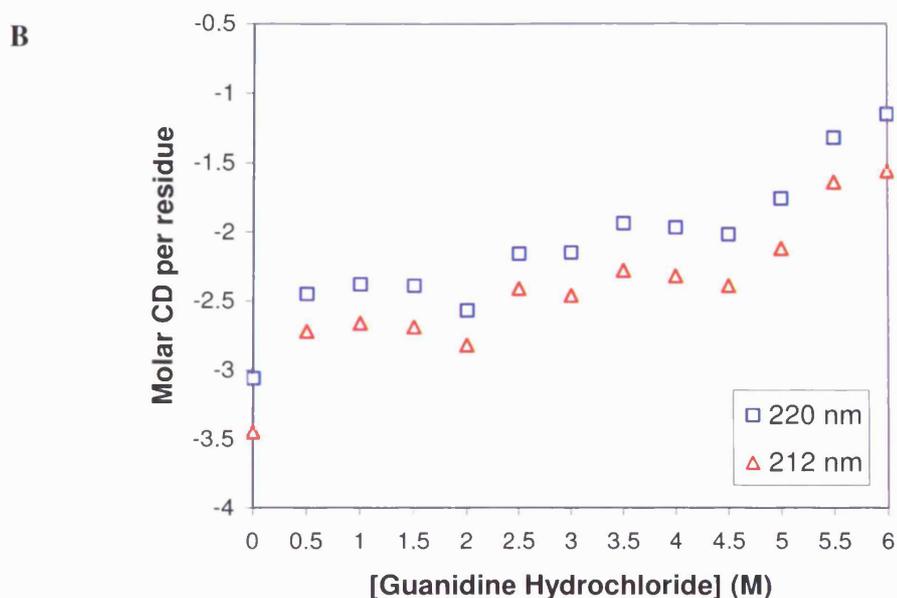
**Figure 3.4.** (A) Circular dichroism and (B) Intrinsic fluorescence emission spectra obtained for p300 1603-1725 at 25 °C. CD spectra were recorded for a 70  $\mu$ M sample in a 20 mM Bis-tris, 100 mM sodium chloride, 20 mM DTT, 0.5 mM ZnCl<sub>2</sub> buffer at pH 7.0. Fluorescence spectra were recorded for a 3  $\mu$ M sample in a 20 mM Bis-tris, 100 mM sodium chloride, 0.5 mM DTT, 0.5 mM ZnCl<sub>2</sub> buffer at pH 7.0. The CD spectral features indicate the presence of  $\beta$ -sheet secondary structure. The tryptophans of p300

1603-1725 contribute to a fluorescence emission maximum of 342 nm suggesting that at least one of the tryptophans is shielded from the solvent.

### 3.3.2.3 Chemical denaturation of p300 1603-1725

Intrinsic tryptophan fluorescence and CD studies were used to assess the effects of increasing guanidine hydrochloride concentrations on p300 1603-1725 structure. The p300 1603-1725 polypeptide contains two tryptophan residues (W1653 and W1685) which serve as useful reporters in monitoring structural transitions upon guanidine hydrochloride induced protein denaturation (figure 3.5A). The result shows that, when monitoring changes in the wavelength of maximum intrinsic tryptophan fluorescence, p300 1603-1725 does not appear to follow a cooperative sigmoidal unfolding curve. There is no evidence of structural resistance to even small amounts of denaturant suggesting the polypeptide lacks any stable tertiary structure. Similarly, when following p300 1603-1725 denaturation by far UV CD (monitoring secondary structural changes at wavelengths 220 nm and 212 nm) there is no evidence of cooperative unfolding (figure 3.5B).

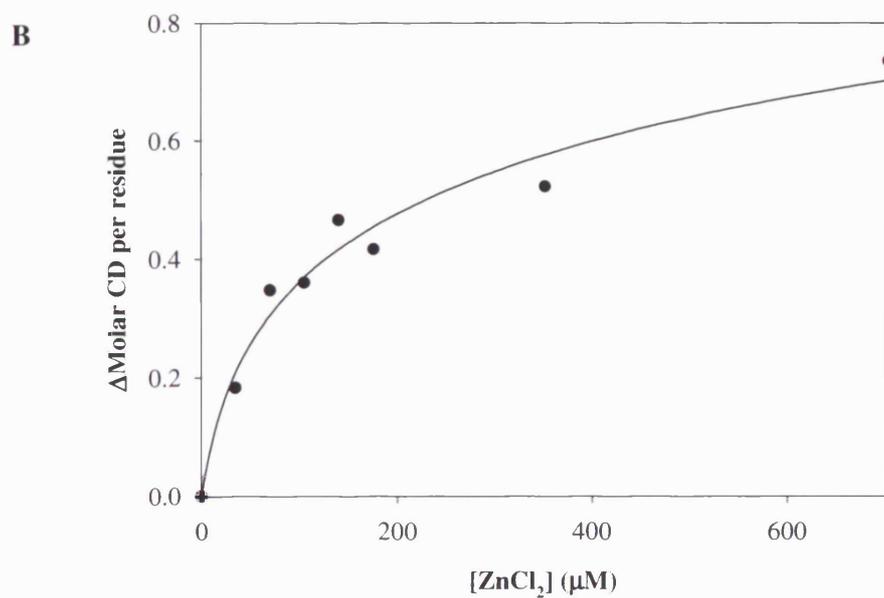
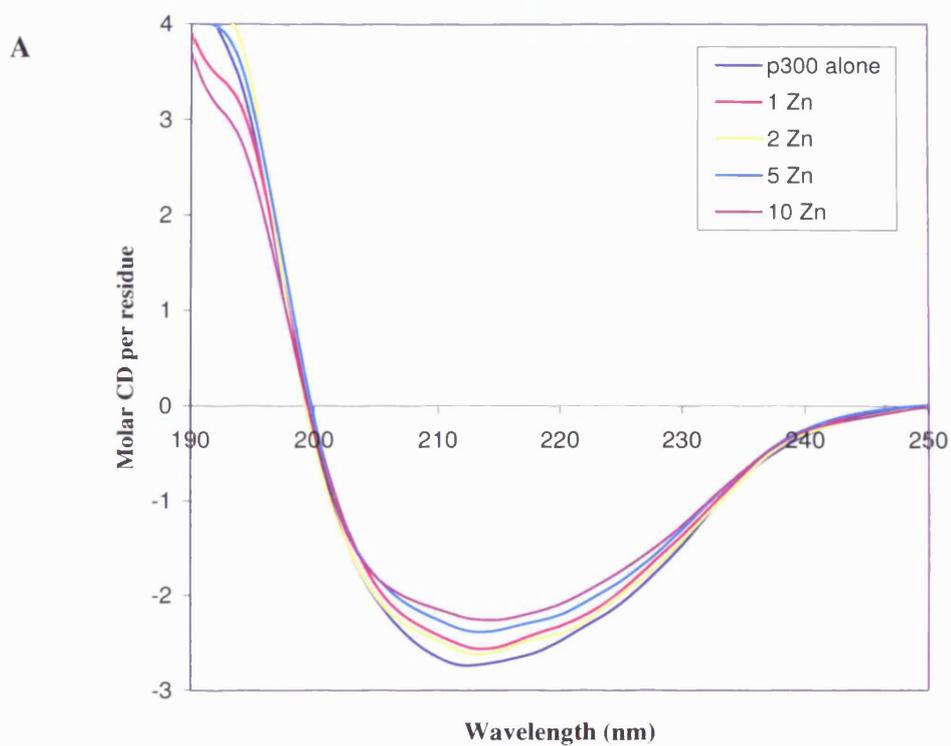




**Figure 3.5 (A)** Representative plots showing the effects of increasing concentrations of guanidine-hydrochloride on the structure of p300 1603-1725 monitored by intrinsic tryptophan fluorescence and **(B)** far UV circular dichroism. A plot of the changes in fluorescence wavelength is shown for a 3  $\mu\text{M}$  sample incubated with guanidine hydrochloride in a 20 mM Bis-tris, 100 mM sodium chloride, 0.5 mM DTT, 0.5 mM  $\text{ZnCl}_2$ , pH 7.0 buffer at 25  $^\circ\text{C}$ . CD Data was plotted as the change in Molar CD per residue intensity at the maximum negative wavelengths 212 nm and 220 nm for a 70  $\mu\text{M}$  sample incubated with guanidine hydrochloride in a 20 mM Bis-tris, 100 mM sodium chloride, 20 mM DTT, 0.5 mM  $\text{ZnCl}_2$ , pH 7.0 buffer at 25  $^\circ\text{C}$ . Together the results strongly suggest that p300 1603-1725 lacks stable tertiary structure due to the lack of resistance to chemical denaturant observed in these experiments.

#### 3.3.2.4 Zinc ion binding properties of p300 1603-1725

The addition of  $\text{Zn}^{2+}$  to solutions of p300 1603-1725 (purified in the absence of  $\text{ZnCl}_2$ ) results in small but significant changes in the protein's far UV CD spectrum, which showed a maximum change at 212 nm and were used to determine the affinity of  $\text{Zn}^{2+}$  binding. Figure 3.6 illustrates the changes in CD intensity observed at 212 nm with the addition of up to 10 molar equivalents of  $\text{Zn}^{2+}$ .



**Figure 3.6** Far UV circular dichroism study of the zinc binding properties of p300 1603-1725. **(A)** Typical spectra obtained for p300 1603-1725 alone and in the presence of increasing molar equivalents of Zn<sup>2+</sup>. **(B)** Plot of the change in CD intensity at 212 nm

induced by the titration of p300 1603-1725 with  $Zn^{2+}$ . The data is consistent with a simple ligand binding function of the form:

$$\Delta CD = \frac{\Delta CD_{\infty}}{2P} \left[ P + L + K_d - \left( \{P + L + K_d\}^2 - 4PL \right)^{1/2} \right]$$

with the curve shown obtained with a dissociation constant ( $K_d$ ) of 70  $\mu$ M, and maximum change in CD intensity ( $\Delta CD_{\infty}$ ) of 0.75.

The p300 1603-1725 polypeptide used in this study was specifically refolded in the absence of  $ZnCl_2$  to promote the purification of the apoprotein form of the ZZ domain. The CD spectrum suggests the polypeptide contains some secondary structure (CDPro estimated: 20% helix, 27%  $\beta$ -sheet, 22% turns and 31% with no regular structure) similar to that observed for protein refolded in the presence of  $ZnCl_2$  (figure 3.4A; 17% helix, 30%  $\beta$ -sheet, 19% turns and 34% with no regular structure). Upon addition of molar equivalents of  $ZnCl_2$  there are small but reproducible shifts in the CD spectra resulting in a loss of CD intensity. A perfectly reasonable fit to the data was obtained by assuming a single  $Zn^{2+}$  binding site on p300 1603-1725, with an apparent dissociation constant of about 70  $\mu$ M.

### 3.3.3 Structural characterisation of p300 1603-1839

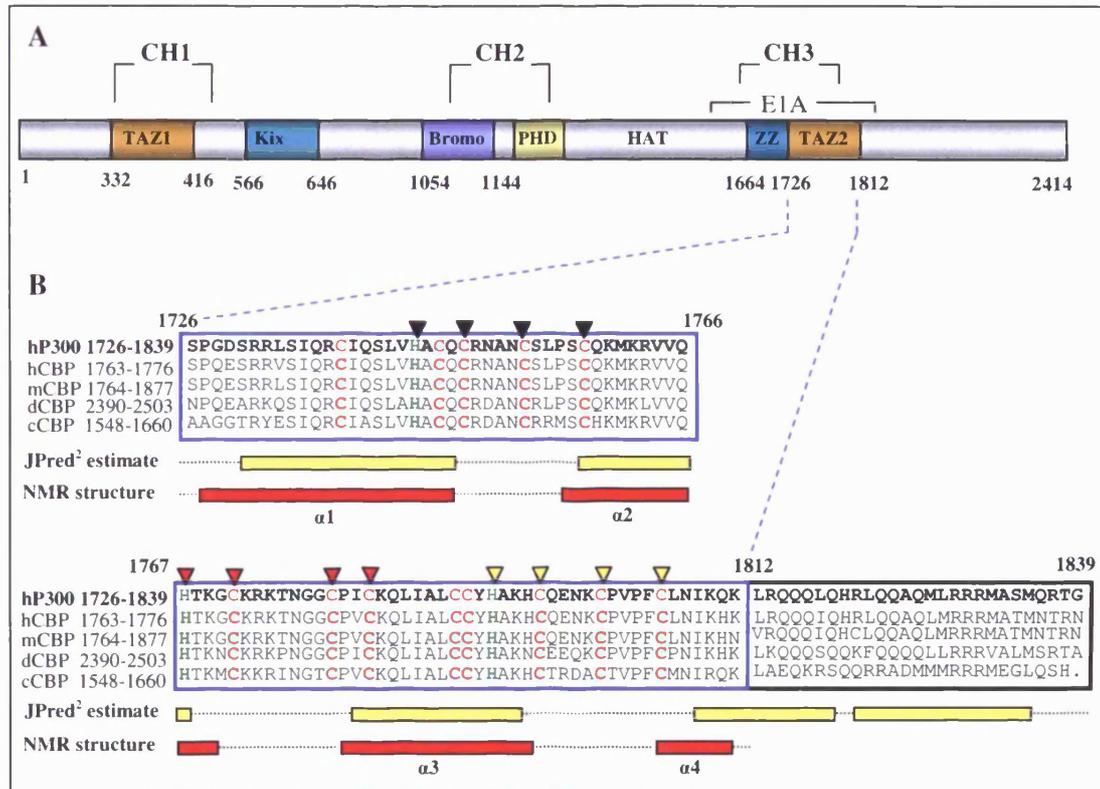
Unfortunately poor solubility and instability of refolded His-tagged p300 1603-1839 precluded an extensive structural study of the polypeptide. However preliminary far UV circular dichroism spectra of refolded and concentrated His-tagged p300 1603-1839 suggested some  $\alpha$ -helical secondary structure (data not shown). The  $\alpha$ -helical secondary structure revealed by circular dichroism is consistent with the major secondary structural component of the TAZ2 domain found within this construct (DeGuzman *et al.*, 2000)

### 3.3.4 Structural characterisation of p300 1726-1839

#### 3.3.4.1 Secondary structure prediction

The p300 1726-1839 construct used in this study to probe interactions with B-Myb incorporates the TAZ2 domain, as determined by DeGuzman *et al.*, in 2000, sharing the same N-terminal amino acid domain boundary but with an additional 27 residues at the C-terminus. Estimates of secondary structural content obtained for the p300 1726-1839 sequence using the Jpred2 package is shown in figure 3.7. For comparison secondary

structure features derived from the NMR solution structure of murine CBP TAZ2 (De Guzman *et al.*, 2000) are highlighted.



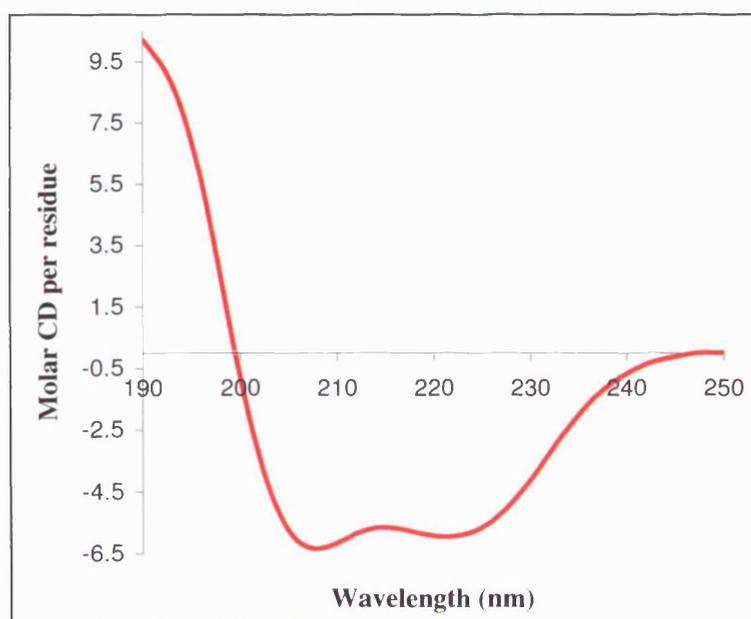
**Figure 3.7** (A) A schematic representation of the domain structure of p300 with the current characterised major functional regions indicated. (B) A multiple sequence alignment of the TAZ2 containing regions of CBP and p300 with Jpred2 and NMR derived secondary structural elements highlighted. Homologous p300 and CBP TAZ2 domain sequences are grouped in the blue box. Conserved cysteine (red) and histidine (green) residues are highlighted, with those involved in Zn<sup>2+</sup> coordination shown with black (Zn1), red (Zn2) and yellow (Zn3) triangles. The solid coloured blocks ( $\alpha$ -helices) shown below the sequence alignment indicate the positions of predicted (JPred2, in yellow) and determined (DeGuzman *et al.*, 2000, in red) secondary structural elements. The protein sequences shown correspond to p300 and CBP from *Homo sapiens*, h; *Mus musculus*, m; *Drosophila melanogaster*, d; and *Caenorhabditis elegans*, c.

Secondary structure predictions obtained using the Jpred2 package were derived from a PSI-Blast created alignment with the input consisting of a single human p300

1726-1839 sequence. The Jpred2 estimates of  $\alpha$ -helical secondary structure content and position are nearly identical to that derived from the solution structure (DeGuzman *et al.*, 2000). Furthermore, it reveals that the additional C-terminal residues 1813 to 1839 (in human p300) are predicted to be largely  $\alpha$ -helical. Interestingly the last of the  $\alpha$ -helices ( $\alpha$ 4), as designated in the CBP TAZ2 structure (DeGuzman *et al.*, 2000), is predicted to extend further, before preceding the start of another  $\alpha$ -helical stretch between residues 1821-1835. However, it must be noted that these results are based solely on sequence based estimations rather than structural data.

#### 3.3.4.2 Biophysical characterisation of p300 1726-1839: secondary and tertiary structural analysis

The far UV CD spectrum obtained for p300 1726-1839 shown in figure 3.8 is typical of a protein with significant  $\alpha$ -helical secondary structure typified by the large negative peaks at 208 nm and 222 nm. The CDPro estimated secondary structure composition for p300 1726-1839 confirms a largely  $\alpha$ -helical content of 52 % with  $\beta$ -sheet 6 %, turns 19 % and unordered regions making up 23 %. The absence of tryptophan residues in p300 1726-1839 precludes the assessment of the proteins tertiary structure by intrinsic tryptophan fluorescence.

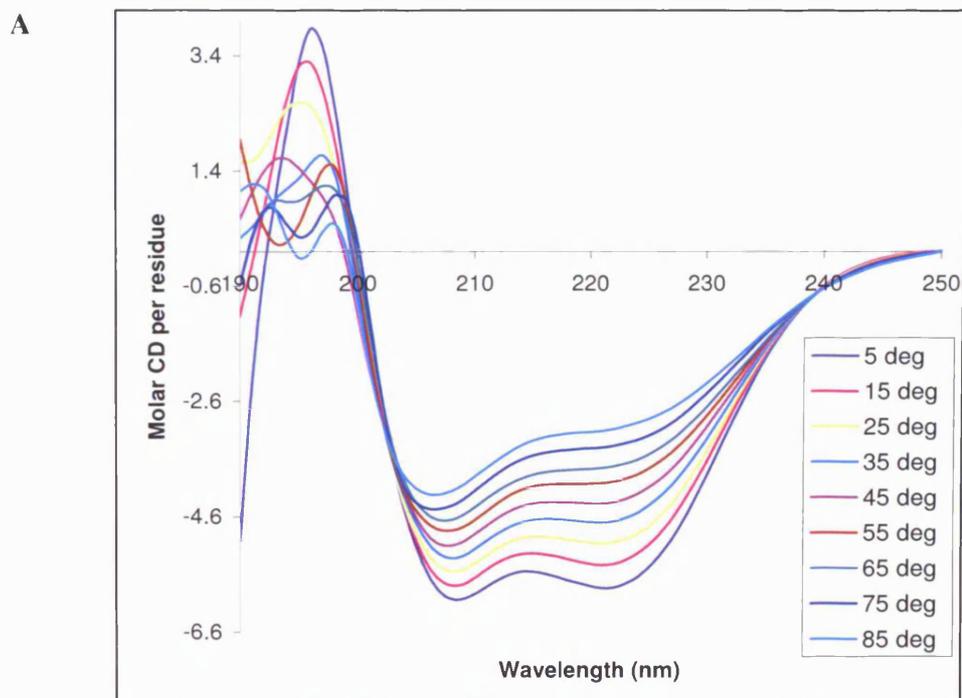


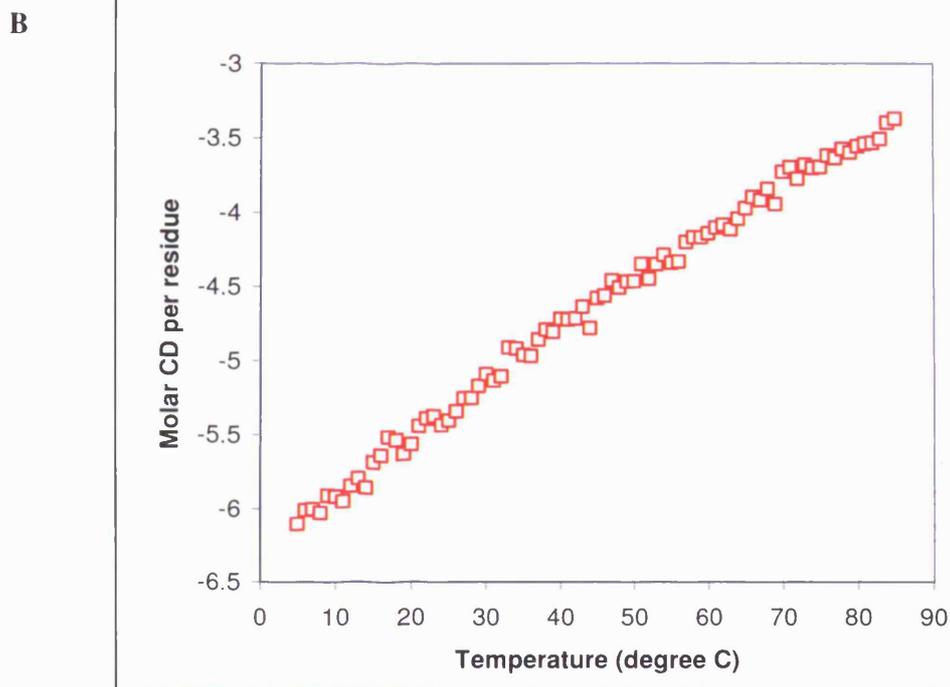
**Figure 3.8** A typical far UV CD spectrum obtained for p300 1726-1839 recorded at 25 °C. CD spectra were recorded for a 64  $\mu$ M sample in a 20 mM Bis-tris, 100 mM sodium

chloride buffer at pH 6.5. The CD spectral features indicate the polypeptide adopts a predominantly  $\alpha$ -helical secondary structure.

#### 3.3.4.3 Effects of thermal variation on p300 1726-1839

Changes in the polypeptides far UV CD spectrum were used to assess the effects of increasing temperature on p300 1726-1839 structural stability. The far UV CD spectra attained for p300 1726-1839 during thermal denaturation studies are shown in figure 3.9A, with the molar CD per residue values at the major negative 222 nm peak plotted as a function of increasing temperature from 5 to 85 °C (figure 3.9B)



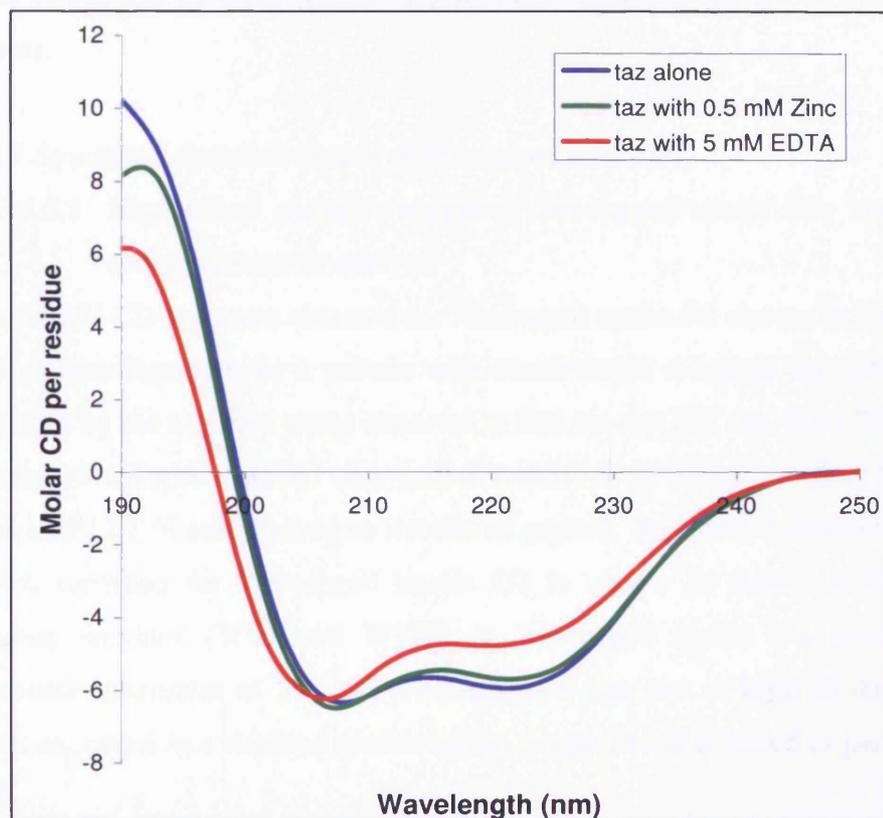


**Figure 3.9.** Far UV circular dichroism analysis of p300 1726-1839 thermal denaturation. (A) Individual CD spectra obtained for p300 1726-1839 at 5 °C, 15 °C 25 °C, 35 °C , 45 °C, 55 °C, 65 °C, 75 °C and 85 °C temperatures are shown compiled. (B) A plot of the changes in Molar CD per residue assessed at 222 nm between 5 and 85 °C in 1 °C increments. CD data was recorded with an 8 μM sample of p300 1726-1839 in a 20 mM Bis-tris, 100 mM sodium chloride, buffer at pH 6.5. The data indicates that, with increasing temperature, p300 1726-1839 exhibits a gradual linear denaturation with no obvious signs of cooperative unfolding between 5 °C and 85 °C.

The p300 1726-1839 polypeptide shows little resistance to temperature increases with a linear denaturation profile, indicating the absence of stable tertiary structure. A characteristic feature observed during the thermal study of p300 1726-1839 is the reversible nature of the melting curve observed upon cooling the sample back to 5 °C from 85 °C (Data not shown). Furthermore the protein appears to exhibit greatest structural content at the lowest temperature of 5 °C (CDPro calculated: 63 % helix, 6 % β-sheet, 14% turns and 17% with no regular structure) than at any other temperature, a feature synonymous with inherent structural instability.

#### 3.3.4.4 $Zn^{2+}$ ion binding properties of p300 1726-1839

The far UV CD spectrum of  $Zn^{2+}$  refolded p300 1726-1839 shows that the major secondary structure is helical (figure 3.8) with thermal denaturation studies suggesting the tertiary structure of the polypeptide is not stable (figure 3.9). Potential conformational changes to p300 1726-1839 structure upon adding further  $Zn^{2+}$  and excess EDTA was therefore assessed by far UV CD (figure 3.10).



**Figure 3.10.** Far UV circular dichroism spectra demonstrating the effects of zinc and EDTA on the structure of p300 1726-1839. The addition of excess  $ZnCl_2$  (0.5 mM) has very little effect on the observed p300 1726-1839 secondary structure. Incubation of the polypeptide with 5 mM EDTA however clearly perturbs the spectrum with a loss of intensity observed at 222 nm and shift of the negative peak from 208 nm to 206 nm. Data were recorded with a 64  $\mu$ M sample of p300 1726-1839 in a 20 mM Bis-tris, 100 mM sodium chloride buffer at pH 6.5.

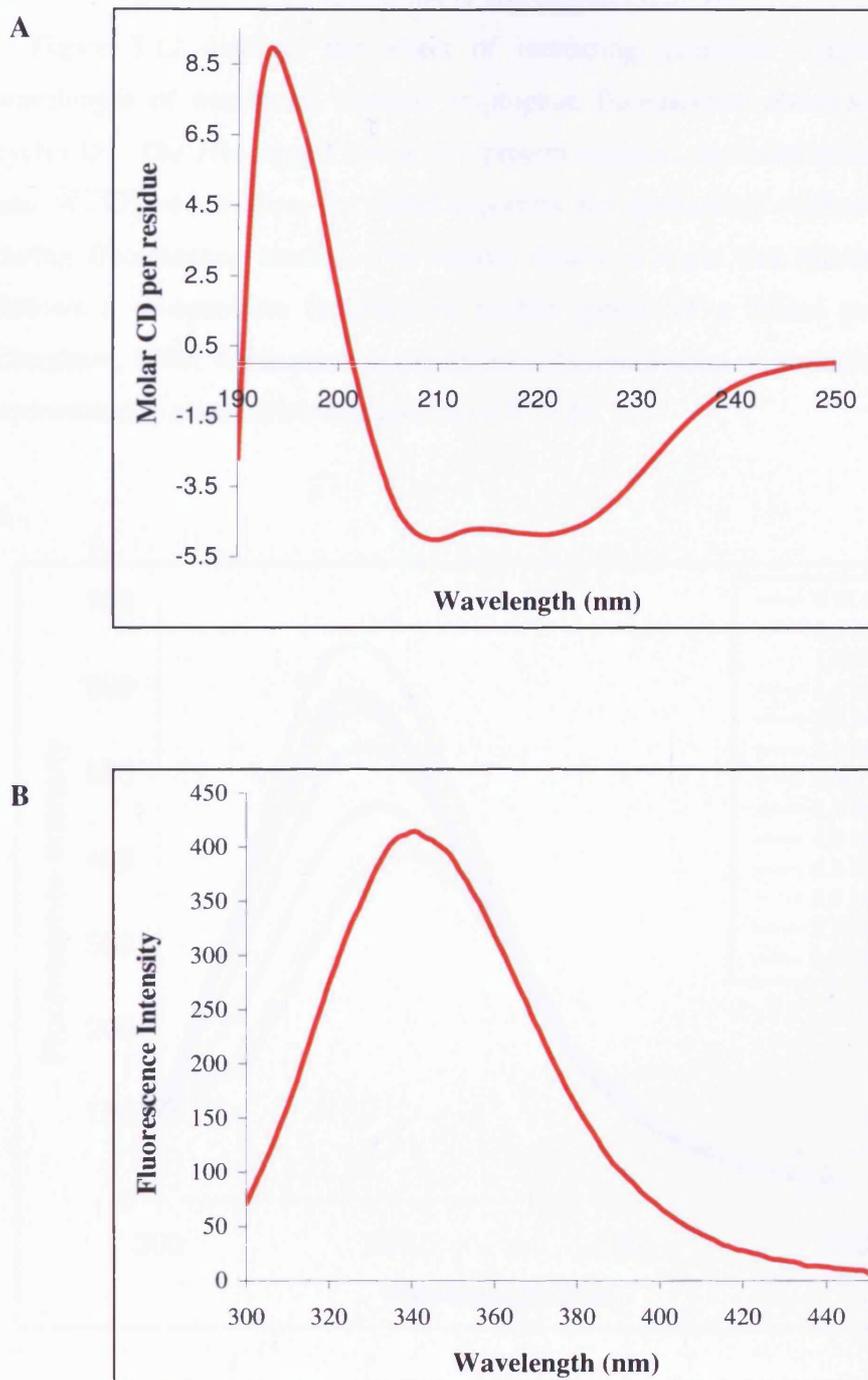
The addition of  $Zn^{2+}$  to solutions of p300 1726-1839 failed to significantly affect the polypeptides far UV CD spectrum. The spectrum in figure 3.10 displays the relatively small shift observed upon the addition of excess  $ZnCl_2$  with no significant change in

resultant secondary structure (CDPro calculated: 51 % helix, 7 %  $\beta$ -sheet, 22 % turns and 20 % with no regular structure). Conversely the addition of saturating concentrations of EDTA (5 mM) had a noticeable effect on the p300 1726-1839 CD spectrum with a loss of intensity at the negative peak at 222 nm accompanied by a shift in the negative peak at 208 nm to 206 nm (i.e. to a more random coil conformation). This correlated with a 6 % reduction in helical secondary structure from 52 % to 46 % (CDPro calculated: 46 % helix, 11 %  $\beta$ -sheet, 17 % turns and 25 % with no regular structure).

### *3.3.5 Structural characterisation of His-tagged cyclin D1*

#### *3.3.5.1 Biophysical characterisation of His-tagged cyclin D1: secondary and tertiary structure analysis*

The far UV CD spectrum obtained for His-tagged cyclin D1 shown in figure 3.11A is typical of that observed for a protein with considerable  $\alpha$ -helical secondary structure exemplified by the negative peaks observed at 208 nm and 222 nm. The CDPro estimate of secondary structure revealed an  $\alpha$ -helical content of 61 %, with 10.5 %  $\beta$ -sheet, 16.5 % turns and 12.1 % contributing to unordered regions. A typical fluorescence emission spectrum recorded for His-tagged cyclin D1 is shown in figure 3.11B. The two tryptophan residues (W63 and W150) in His-tagged cyclin D1 contribute to a fluorescence maximum of 341 nm. This suggests that one or both of the tryptophan residues are found in a shielded conformation, suggestive of a folded or partially folded protein.

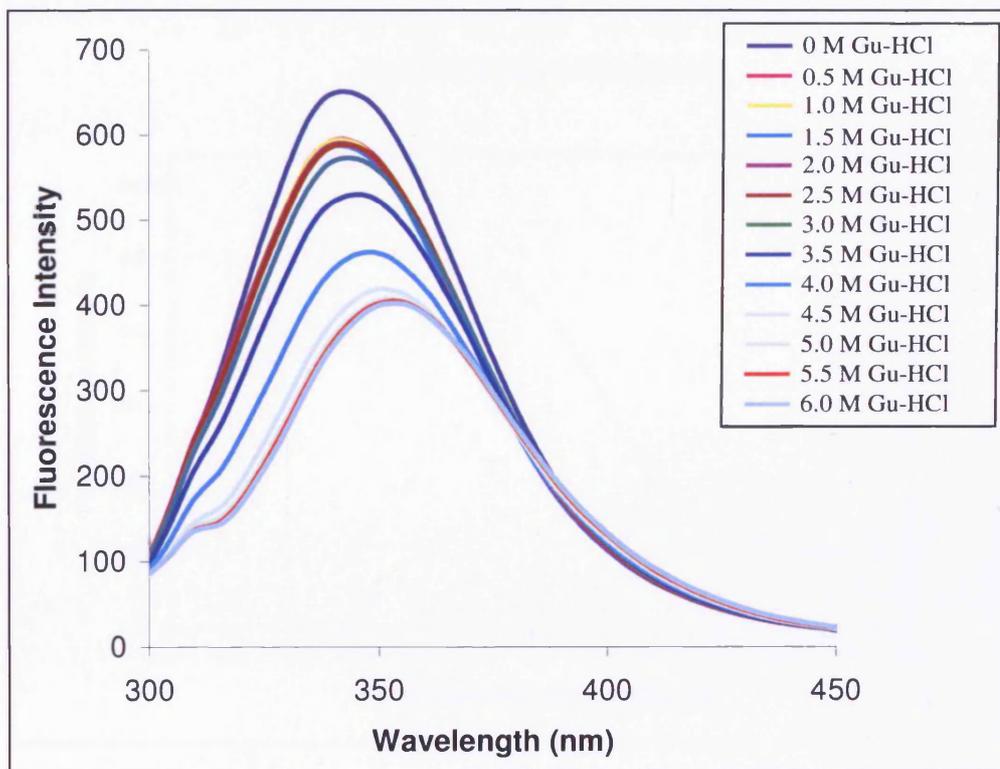


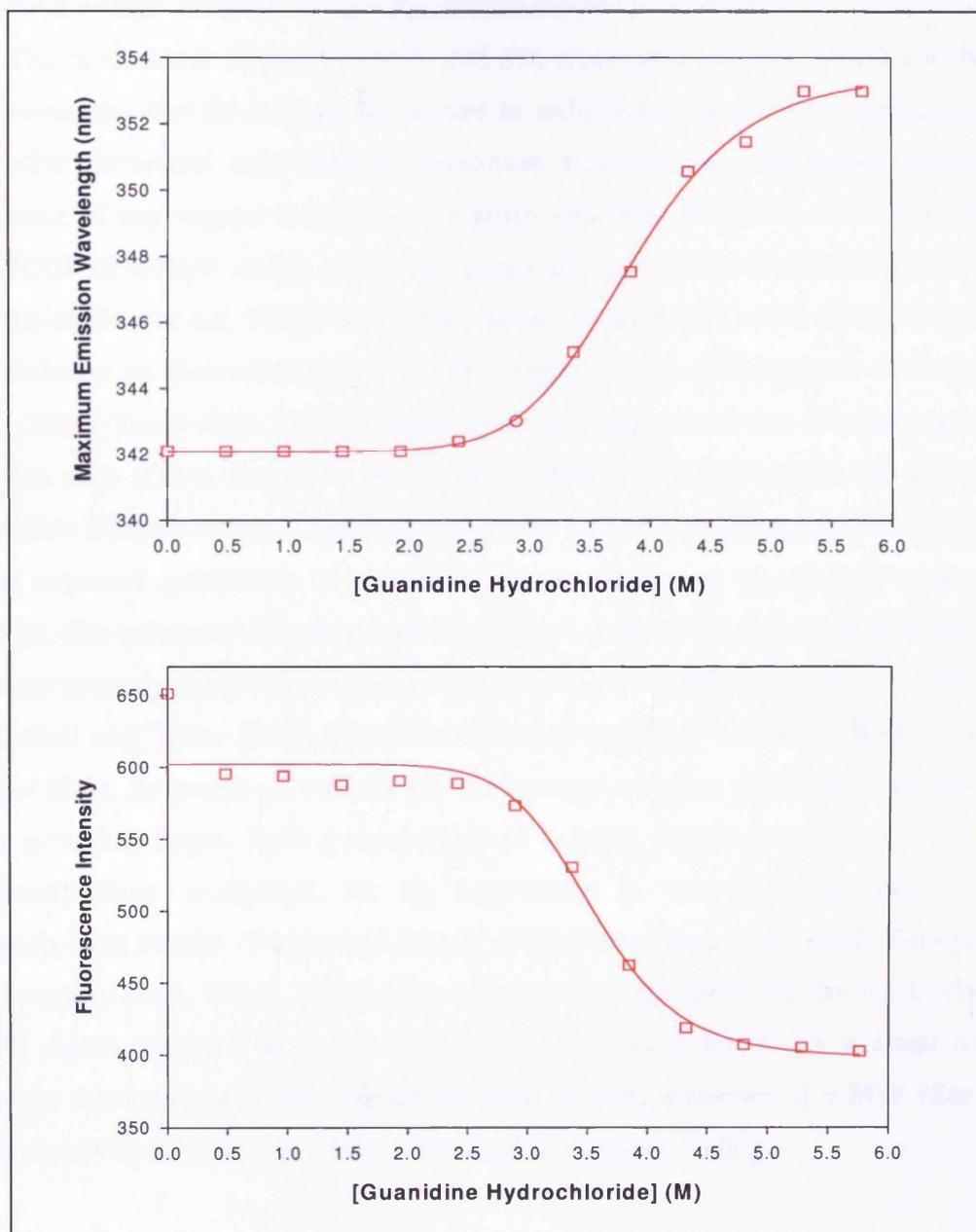
**Figure 3.11** (A) Circular dichroism and (B) Intrinsic fluorescence emission spectra obtained for His-tagged cyclin D1 at 25 °C. CD spectra were recorded for a sample in a 25 mM NaH<sub>2</sub>PO<sub>4</sub>, 100 mM sodium chloride buffer at pH 8.0. Fluorescence spectra were recorded for a 3 μM sample in the same buffer. The CD spectral features indicate the presence of α-helical secondary structure. The fluorescence emission maximum of 341 nm suggests the tryptophans adopt a partially buried conformation.

### 3.3.5.2 Chemical denaturation of His-tagged cyclin D1

Figure 3.12 displays the effect of increasing guanidine hydrochloride on the wavelength of maximum intrinsic tryptophan fluorescence observed for His-tagged cyclin D1. The His-tagged cyclin D1 protein contains two tryptophan residues (W63 and W150) which serve as useful reporters for monitoring conformational changes during fluorescence studies. The results clearly indicate that His-tagged cyclin D1 follows a co-operative denaturation profile typical of a folded protein (Pace and Creighton, 1986; Williamson *et al.*, 1996) with a mid-point of unfolding at a guanidine hydrochloride concentration approaching 3.75 M.

A



**B**

**Figure 3.12.** *Intrinsic tryptophan fluorescence studies demonstrating the effect of guanidine hydrochloride on His-tagged cyclin D1 (3  $\mu$ M) conformational stability. (A)* Individual spectra are shown compiled with a visible shift in the maximum fluorescence emission to a longer wavelength upon increasing denaturant concentrations, indicative of protein unfolding. **(B)** Plots of maximum emission wavelength (top) and fluorescence intensity (bottom), as a function of increasing denaturant concentration. Both illustrate the co-operative, sigmoidal nature of His-tagged cyclin D1 unfolding.

### 3.4 Discussion

#### 3.4.1 B-Myb 275-376: An intrinsically disordered polypeptide

The biophysical study of B-Myb 275-376 reported in section 3.3.1.2 conclusively demonstrates that the polypeptide adopts an unfolded random coil conformation. Both circular dichroism and intrinsic tryptophan fluorescence experiments indicate the absence of any regular secondary or tertiary structure. This is in contrast to residues 295-309 of c-Myb which populates a partially helical conformation (25-30 %) in isolation (Zor *et al.*, 2002) which then forms an extended stabilised helix (the helix boundaries are shown in figure 3.1B) in complex with the KIX domain of CBP (Zor *et al.*, 2004). The B-Myb 275-376 region contains a high proportion of polar and charged amino acids (Glutamine 10 %, Serine 16 %, Proline 9 %) as compared to hydrophobic residues (Phenylalanine 2 %, Tryptophan 2 %, Valine 5 %), features synonymous with low sequence complexity and intrinsic disorder (Garner *et al.*, 1998; Uversky *et al.*, 2000). The presence of such glutamine-rich and proline-rich activation domains within transcriptional regulatory proteins has long been recognised (Courey and Tjian, 1988; Mitchell and Tjian, 1989). The transactivation-associated region of B-Myb therefore most likely represents an intrinsically unstructured segment that folds upon binding to its biological target. Such a mechanism of coupled protein folding and binding has recently been recognised for its importance in transcriptional regulation and translational control (Wright and Dyson, 1999; Iakoucheva *et al.*, 2002; Tompa, 2002; Iakoucheva *et al.*, 2004). The Jpred2 secondary structure prediction for the B-Myb 275-376 region (reported in section 3.3.1.1) revealed the potential for a small  $\alpha$ -helical region that overlaps a homologous  $\alpha$ -helical forming sequence of c-Myb (Zor *et al.*, 2004), which may serve to mediate such coupled protein folding.

#### 3.4.2 Structural characterisation of p300 1603-1725 and comparison with the CBP ZZ domain

During the course of this research project, the solution structure of the ZZ domain from murine CBP was determined and published by Legge *et al.*, 2004. The 52 residue ZZ domain spanning amino acids 1700-1751 in murine CBP (the equivalent amino acids are 1661-1713 in human p300) was found to consist of two twisted antiparallel  $\beta$ -sheets and a short  $\alpha$ -helix that together co-ordinately bound two zinc ions. Thus, the human p300 1603-1725 construct used in the work presented here differs from that of

the published murine CBP ZZ domain by including an additional 58 amino acids at the N-terminus and 12 residues at the C-terminus.

Tertiary structural analysis of p300 1603-1725 using intrinsic tryptophan fluorescence, indicated the polypeptide was at least partially folded from the observation that one or both tryptophans (Y1653 and Y1685) reside in a solvent shielded environment ( $\lambda_{\text{max}} = 342 \text{ nm}$ ) described in section 3.3.2.2. Analysis of the far UV CD spectrum obtained for p300 1603-1725 indicates a predominantly  $\beta$ -sheet containing polypeptide (17% helix, 30%  $\beta$ -sheet, 19% turns and 34% with no regular structure). The circular dichroism derived secondary structure content for p300 1603-1725 cannot be directly compared with that determined from the CBP ZZ domain solution structure (Approximately 15 % helix, and 35 %  $\beta$ -sheet) since the p300 construct is larger at both the N- and C-termini. However, assuming that the regions outside the ZZ domain (as designated by Legge *et al.*, 2004) were disordered, an  $\alpha$ -helical and  $\beta$ -sheet content of 7 % and 15 % respectively would be expected for the p300 1603-1725 polypeptide. This is not consistent with the observed CD data, where both the  $\alpha$ -helical and  $\beta$ -sheet content of p300 1603-1725 is more than double this. This suggests that the regions outside the designated ZZ domain, i.e. residues 1603 to 1660 and residues 1714 to 1725 in human p300, may contain additional secondary structural elements. This evaluation is supported by the Jpred2 prediction, which suggests the potential for further  $\beta$ -sheet elements between residues 1603 to 1660 (shown in figure 3.3).

The chemical denaturation of p300 1603-1725 appears to follow a gradual linear unfolding profile as indicated by far UV circular dichroism and intrinsic tryptophan fluorescence studies described in section 3.3.2.3. Typically, a folded protein with globular tertiary structure follows a cooperative unfolding transition upon denaturation (Pace and Creighton, 1986; Williamson *et al.*, 1996), as seen for His-tagged cyclin D1 (figure 3.14). Conversely proteins with no stable tertiary structure or which reside in a molten globule state (proteins with native-like secondary structure but disordered tertiary structure) exhibit very little resistance to denaturant and unfold rapidly (Renshaw *et al.*, 2002). Thus the p300 1603-1725 denaturation profile fails to adhere to either of these protein unfolding models. A possible explanation for the observed

phenomena is that the unfolding profiles represent p300 1603-1725 dis-aggregation and bear little reflection on its true unfolding. Solvent exposed tryptophan residues in an unfolded/ denatured protein typically have a maximum fluorescence emission wavelength of around 355 nm as observed for B-Myb 275-376. Yet in the presence of 6 molar guanidine-hydrochloride the tryptophan residues in the p300 1603-1725 polypeptide contribute to a maximum fluorescence emission wavelength of around 349 nm. Since fluorescence spectroscopy gives information on the environment of the aromatic rings this suggests that, in high concentrations of denaturant, the tryptophan residues in p300 1603-1725 remain partially solvent shielded, a possible outcome of polypeptide aggregation. There were no visible signs of aggregation or precipitation during the fluorescence based unfolding experiment although the low concentrations of protein used (3  $\mu$ M) would have made this difficult to observe. However it was noted that p300 1603-1725 precipitation was encountered during the circular dichroism based unfolding studies. When samples of p300 1603-1725 were incubated with denaturant overnight before data acquisition there was visible precipitation. This problem was largely overcome by incubating the samples with denaturant for a shorter period (30 minutes) before data acquisition.

The murine CBP ZZ domain is reported to be a folded domain that binds two  $Zn^{2+}$  ions; one  $Zn^{2+}$  ion is coordinated tetrahedrally via two CXXC motifs and the second is coordinated tetrahedrally by a third CXXC motif together with a HXH motif (Legge *et al.*, 2004). Treatment of the ZZ domain with EDTA resulted in loss of the folded structure, according to CD and NMR analysis, indicating the protein fold is induced by  $Zn^{2+}$  binding (Legge *et al.*, 2004). The observation in this study that p300 1603-1725, purified in the absence of  $Zn^{2+}$ , has considerable secondary structure differs to that reported for the CBP ZZ domain. The addition of saturating concentrations of EDTA to p300 1603-1725 did not significantly alter its far UV CD spectrum or perturb its fluorescence wavelength emission (Data not shown) and thus failed to unfold the protein. This may be explained by the structural contribution of additional residues within the p300 1603-1725 polypeptide namely, amino acids 1603 to 1660 and 1714 to 1725, which were absent in the CBP ZZ domain. The additional residues account for 57 % of the polypeptide and if structured independently of  $Zn^{2+}$  would contribute significantly to the observed circular dichroism and fluorescence spectra. The Jpred2

secondary structure prediction strongly suggests the potential for  $\beta$ -sheet elements between 1600-1660.

The calculated affinity of the p300 1603-1725 polypeptide for  $\text{Zn}^{2+}$  ( $K_d \sim 7 \times 10^{-5}$  M) is much lower than that reported for several other classes of zinc finger domains found in eukaryotic transcription factors such as those typified by TFIIIA and the steroid hormone receptors ( $K_d \sim 1 \times 10^{-8}$  M) (Dudev and Lim, 2003). The affinity for  $\text{Zn}^{2+}$  binding was not reported for the CBP ZZ domain by Legge *et al.*, 2004. It is possible that the weak  $\text{Zn}^{2+}$  binding observed for p300 1603-1725 is as a result of a non-specific association with its N-terminal region. During the preparation of  $\text{Zn}^{2+}$ -free p300 1603-1725 it is possible that the ZZ domain (residues 1661-1713 in human p300) had bound  $\text{Ni}^{2+}$  during the  $\text{Ni}^{2+}$  affinity refolding and purification procedure. Indeed it was noted that the p300 1603-1725 polypeptide, in the absence of its N-terminal His-tag, still maintained the capacity to bind the Ni-NTA resin (section 2.3.10.1). Although the purified p300 1603-1725 polypeptide was dialysed extensively against metal-free buffer it is conceivable that  $\text{Ni}^{2+}$  may have remained bound to the ZZ domain, and that the observed weak  $\text{Zn}^{2+}$  binding is occurring non-specifically to conserved cysteine and histidine residues found within the N-terminal region (residues 1603 to 1660). In this region there are two conserved cysteines (residues 1621 and 1653) and two conserved histidines (residues 1638 and 1659) with the potential to act as  $\text{Zn}^{2+}$  coordination ligands.

Although there are differences in the ZZ domain amino acid content between human p300 and murine CBP, there are no changes in the highly conserved cysteine and histidine residues that constitute the characterised zinc binding ligands in CBP (Legge *et al.*, 2004) and any substitutions in other residues are conservative in nature. Therefore the structural properties observed for the ZZ domain containing p300 1603-1725 polypeptide are most likely a result of the substantial number of additional residues and the refolding and purification strategy used.

### 3.4.3 Structural characterisation of p300 1726-1839 and comparison with the CBP TAZ2 domain

Far UV CD analysis of the p300 1726-1839 polypeptide (figure 3.8) revealed a predominantly  $\alpha$ -helical secondary structure (52 %). By comparison the  $\alpha$ -helical

composition of the murine CBP TAZ2 domain 1764-1850 (residues 1726-1812 in human p300) determined from the solution structure makes up approximately 67% of the protein (DeGuzman *et al.*, in 2000). If the assumption was made that the Glutamine-rich region between residues 1813-1839 in human p300 (and the corresponding CBP region 1851-1877) was largely unstructured, then the  $\alpha$ -helical content of p300 1726-1839 would be expected to reach approximately 50 %, which compares favourably with the CD estimate. The Jpred2 prediction for the p300 1726-1839 polypeptide compares favourably with the determined CBP TAZ2 secondary structure (figure 3.7) and suggests the 1813-1839 region forms additional  $\alpha$ -helical secondary structure. Therefore, based on the Jpred2 prediction for this region, an  $\alpha$ -helical content of 68 % would be expected for the p300 1726-1839 polypeptide, suggesting that circular dichroism underestimates its  $\alpha$ -helical content or alternatively the refolded p300 1726-1839 polypeptide lacks the full complement of  $\alpha$ -helical secondary structure. All of these assumptions are based on the premise that the TAZ2 domain of p300 adopts a similar fold and structure to that reported for CBP.

The p300 1726-1839 polypeptide exhibits a non-cooperative, reversible thermal denaturation profile (figure 3.9), features that were similarly observed for the TAZ1 domain of p300 (Dial *et al.*, 2003). The highly homologous TAZ1 domain from p300 was reported, unexpectedly, to exhibit a non-cooperative, reversible melting curve characteristic of molten globule proteins. The  $Zn^{2+}$  bound TAZ1 domain only assumed a folded structure upon complex formation with the C-terminal transactivation domain from HIF-1 $\alpha$  (Dial *et al.*, 2003). The authors thus concluded that TAZ1 represents a partially structured  $Zn^{2+}$  bound domain that co-folds upon binding its transcription factor ligands. This finding is in stark contrast to a recent publication that argues the  $Zn^{2+}$  bound TAZ1 domain (from both CBP and p300), as with the  $Zn^{2+}$  bound TAZ2 domain, presents a stable structured scaffold for protein binding (De Guzman *et al.*, 2005). De Guzman *et al.*, attribute the molten globule characteristics previously reported for the TAZ1 domain (Dial *et al.*, 2003) to protein misfolding caused by the use of excess  $Zn^{2+}$  during refolding. The TAZ1 domain in the presence of stoichiometric concentrations of  $Zn^{2+}$  was shown, by CD and NMR, to adopt a well-defined structure in solution (De Guzman *et al.*, 2005). However at higher  $Zn^{2+}$  concentrations there was a visible loss of helicity as measured by CD and a loss of dispersion with line broadening resulting in 2D NMR spectra similar to that observed by Dial *et al.*

The thermal melting properties described for Zn<sup>2+</sup> bound p300 1726-1839 therefore suggests the polypeptide resides in a molten globule conformation most likely due to the misfolding of the protein. Although the p300 1726-1839 polypeptide was refolded in the presence of high concentrations of Zn<sup>2+</sup> (0.5 mM ZnCl<sub>2</sub>), excess Zn<sup>2+</sup> was removed by dialysis prior to structural characterisation (section 2.2.14). Small differences in the sequence composition of the TAZ2 domain between murine CBP and human p300 are unlikely to significantly affect the structural stability of the Zn<sup>2+</sup> binding module and are thus equally unlikely to be attributable to the structural behaviour observed for p300 1726-1839. The effect of the additional C-terminal residues 1813-1839 upon p300 TAZ2 stability is difficult to interpret but is unlikely to be structurally detrimental. This region does not contain any further conserved cysteine or histidine residues which may act to compete for Zn<sup>2+</sup> binding and thus alter correct Zn<sup>2+</sup> ion coordination. Furthermore the region is predicted to be structured with some  $\alpha$ -helical content (Jpred2 analysis, figure 3.7).

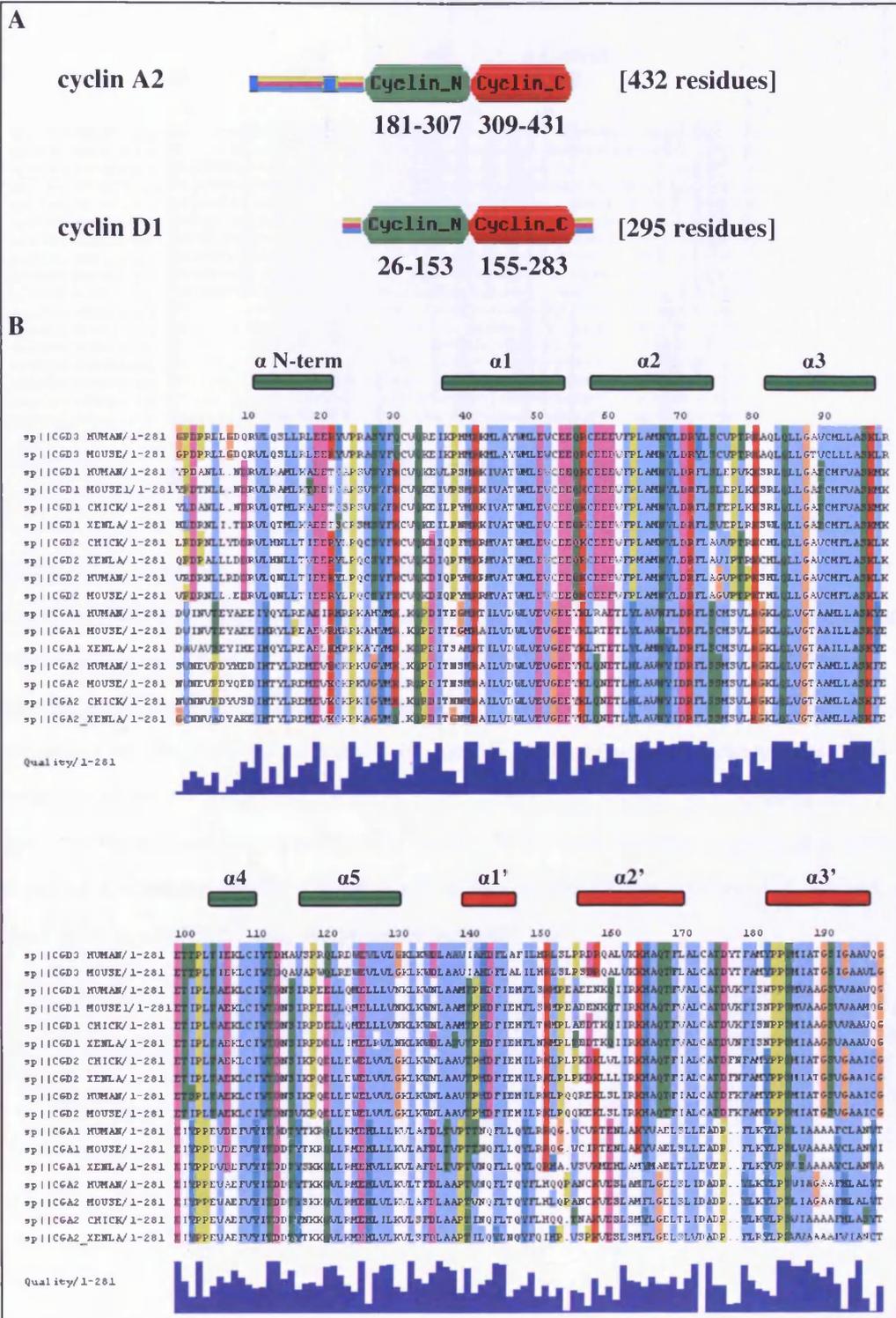
Although the addition of EDTA had visibly perturbed the p300 1726-1839 circular dichroism spectrum (figure 3.10), the level of remaining structure is not consistent with the Zn<sup>2+</sup>-free CD spectrum obtained for the CBP TAZ2 domain (DeGuzman *et al.*, 2000). Here the authors report that the Zn<sup>2+</sup>-free apoprotein form of CBP TAZ2 assumes a random coil conformation (characterised by a large negative 200 nm peak, as observed for B-Myb 275-376) that cooperatively folds into a compact globular structure upon binding to three equivalents of Zn<sup>2+</sup> (DeGuzman *et al.*, 2000). Clearly the chelation of Zn<sup>2+</sup> through the addition of excess EDTA does not substantially affect p300 1726-1839 structure, with only a small reduction in helical content observed (a 6% reduction from 52 % to 46 %). The remaining helical structure may partly be explained by the additional C-terminal 27 residues 1813-1839, not present on the CBP TAZ2 domain reported by DeGuzman *et al.*, which are predicted to be largely  $\alpha$ -helical by Jpred2 analysis (figure 3.7).

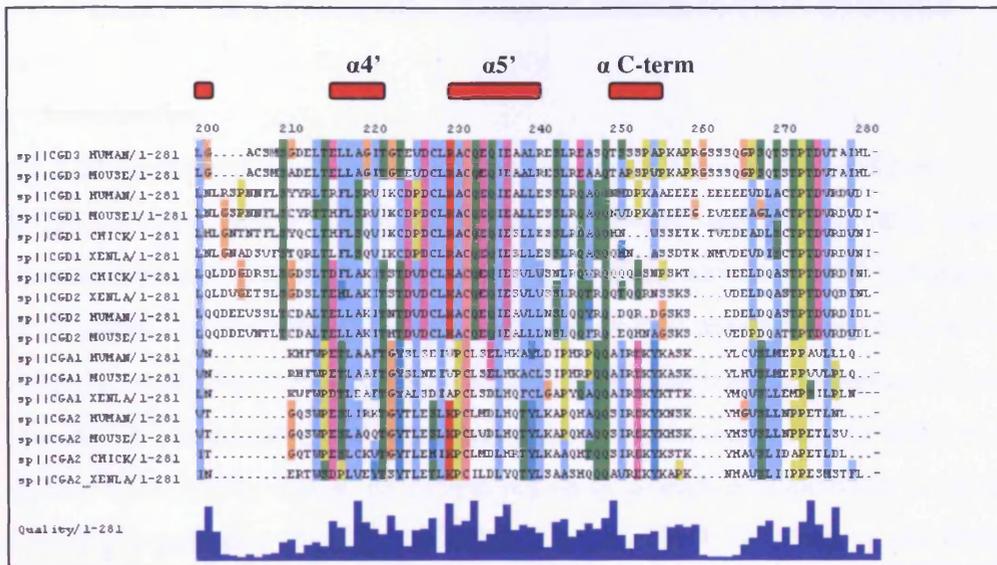
#### *3.4.4 Structural characterisation of His-tagged cyclin D1 and comparison with cyclin A.*

The biophysical studies reported in section 3.3.5 clearly demonstrate the purified His-tagged cyclin D1 protein adopts a stable tertiary fold with considerable  $\alpha$ -helical

secondary structure. His-tagged cyclin D1 follows a cooperative unfolding transition typical of a folded protein (Pace and Creighton, 1986; Williamson *et al.*, 1996) with a mid-point of unfolding at a guanidine hydrochloride concentration approaching 3.75 M. It is of interest to note that both tryptophan residues (W63 and W150) are found within the conserved cyclin\_N domain of cyclin D1. Thus the fluorescence derived unfolding curve more than likely represents the denaturation of the globular fold formed by the cyclin\_N domain, with some possible contribution from the cyclin\_C domain, unfolding (particularly at W150 found at the interface between cyclin\_N and cyclin\_C domains).

The cyclin D1  $\alpha$ -helical content derived from the CD calculation (61 %) can be directly compared with the cyclin A determined structure (Brown *et al.*, 1995; Jeffrey *et al.*, 1995) despite the inherent differences in protein size and sequence. The cyclin A fragment (residues 171-432) used to deduce the crystal structure spans the conserved cyclin\_N (cyclin box) and cyclin\_C domains found commonly in cyclin proteins including cyclin D1. The N-terminal 170 residues absent in the fragment do not have significant homology in other cyclins and are absent from cyclins D and E (Hadwiger *et al.*, 1989; figure 3.13A). Figure 3.13B displays a multiple sequence alignment of D-type cyclins together with cyclins A1 and A2 demonstrating the high degree of sequence homology apparent in both the cyclin\_N and cyclin\_C domains. The 61 %  $\alpha$ -helical secondary structure content determined for His-tagged cyclin D1 compares favourably with the 63 %  $\alpha$ -helical content of the cyclin A2 polypeptide (determined for a fragment containing residues 171-432 of human cyclin A2, Brown *et al.*, 1995) suggesting it may adopt a similar structure.





**Figure 3.13.** Multiple sequence alignment of D-type cyclins with cyclins A1 and A2 from *Homo sapiens*, *Mus musculus*, and *Xenopus laevis*. (A) A schematic representation of the domain structure of cyclin A2 and cyclin D1. (B) The alignment illustrates the high degree of sequence homology that exists between the two respective cyclin types, particularly at the N-terminal cyclin\_N domain. Helical regions determined from the crystal structure of cyclin A1 (Brown *et al.*, 1995) are shown as cylinders above the alignment with those found within the cyclin\_N domain shaded in green and those in the cyclin\_C domain shaded in red. CGD refers to the D-type cyclins 1-3; CGA1 and CGA2 refer to cyclin A1 and A2 proteins respectively.

## CHAPTER 4. PROTEIN-PROTEIN INTERACTION STUDIES

### 4.1 Introduction

#### 4.1.1 Mapping of B-Myb and p300/cyclin D1 interactions to date and aims

Recent studies report a physiologically relevant and functionally significant interaction between B-Myb and the transcriptional co-activator p300 (Johnson *et al.*, 2002; Schubert *et al.*, 2004). Both *in vitro* and *in vivo* experiments indicate the E1A binding region of p300, containing the ZZ and TAZ2 zinc binding domains, is primarily responsible for interacting with a central region (amino acids 227 to 508) of B-Myb (Schubert *et al.*, 2004). The same central region of B-Myb is responsible for binding to the cell cycle protein cyclin D1 (Horstmann *et al.*, 2000; Schubert *et al.*, 2004). The p300 1603-1725, p300 1726-1839, and p300 1603-1839 recombinant polypeptides generated in this study were therefore used to assess the importance of the ZZ and TAZ2 domains in binding B-Myb through use of GST/B-Myb pull-down assays. Similar GST pull-down assays using the array of purified GST/B-Myb proteins were conducted to map the cyclin D1 binding site within B-Myb.

### 4.2 Materials and Methods

#### 4.2.1 Batch-type pull-down binding assays

In a typical batch-type assay 100 µg of GST/B-Myb 275-376 or GST/B-Myb 227-508 was mixed with 50 µl of glutathione-agarose and incubated with constant agitation for 1 hour at room temperature to promote binding. After incubation the agarose beads were washed three times with a 20 mM Tris, 100 mM sodium chloride, 10 mM DTT, and 0.5 mM ZnCl<sub>2</sub> buffer at pH 7.0 to remove any unbound GST fusion protein. Approximately 50 µg of p300 1603-1725 was then mixed with the GST/B-Myb fusion protein loaded beads and further incubated with constant agitation for 1 hour at room temperature. The beads were then washed three times with wash buffer after which any bound proteins were eluted from the beads with buffer containing 10 mM reduced glutathione or after treatment with Lithium dodecyl sulfate (LDS) sample buffer (Invitrogen) and analysed by SDS PAGE.

#### 4.2.2 Column-based pull-down binding assays

In a typical pull-down binding assay 500 µg of purified GST/B-Myb protein was mixed and incubated with 0.5 mls of pre-equilibrated S-linked glutathione agarose (Sigma) at room temperature for 30 minutes, in order to promote binding of the fusion protein with the resin. Approximately 60-80% of GST/B-Myb protein binds the resin under these conditions. For interaction analysis with p300 proteins the glutathione agarose resin was pre-equilibrated and washed with a 20 mM Tris, 100 mM sodium chloride, 10 mM DTT, and 0.5 mM ZnCl<sub>2</sub> buffer at pH 7.0. Interaction Studies with the His-tagged cyclin D1 protein were conducted with a 20 mM Tris and 100 mM sodium chloride, equilibration/wash buffer at pH 8.0. Upon resin equilibration and GST/B-Myb protein incubation the resin was poured into a 5 ml column and then washed with five column volumes of the wash buffer. Approximately 400 µg of purified p300/cyclin D1 polypeptide was then loaded onto the column and any unbound protein removed by washing once more with five column volumes of buffer. Proteins and protein complexes still bound to the resin were finally eluted from the column in the equilibration/wash buffer containing 10 mM reduced glutathione, and the composition of the protein fractions analysed by SDS-PAGE gels (Invitrogen) stained with either coomassie or silverstain.

#### 4.2.3 Western detection of His-tagged cyclin D1 in pull-down assays against GST/B-Myb proteins.

In order to detect trace amounts of potential co-eluted His-tagged cyclin D1 during GST pull-down assays against GST/B-Myb proteins a monoclonal antibody directed towards the hexa-histidine sequence tag (Novagen) was used in a western blotting protocol. Pull-down assays involving His-tagged cyclin D1 with GST/B-Myb 227-508 and GST/B-Myb 318-508 were conducted as previously described in section 4.2.2 and loaded onto an SDS PAGE gel alongside Perfect Protein™ Western Marker (Novagen). The western transfer procedure was followed as directed in the 'NOVEX® western transfer apparatus instructions' (Invitrogen) using a NOVEX® XCell II™ Blot Module and a pre-prepared Tris-Glycine transfer buffer (composed of 12 mM Tris base, 96 mM Glycine, 20% Methanol, 0.04% (w/v) SDS). Protein was transferred to a pre-cut Hybond-P PVDF Membrane (Amersham Biosciences) by electrophoresis for 1 hour and 45 minutes at a setting of 25 volts.

The efficient transfer of protein was verified after viewing Ponceau Red (Sigma) stained membrane. The PVDF transferred membrane was then washed twice for 10 minutes each time with 15 ml of 1 X TBS (Appendix A.1 for buffer composition) and incubated overnight in Blocking solution containing BSA (5% w/v Milk, 1% w/v BSA in TBS). Blocking solution was removed and the membrane washed twice for 10 minutes each time with 20 ml of 1 X TBSTT (Appendix A.1 for buffer composition) before being washed for a further 10 minutes with 15 ml of 1 X TBS. The membrane was then incubated for 1 hour with His-Tag Monoclonal Antibody (Novagen) diluted 1:4000 in blocking solution (without added BSA). Unbound antibody was removed after washing the membrane twice for 10 minutes each time with 20 ml 1 X TBSTT, and with 15 ml of 1 X TBS for 15 minutes. The membrane was then incubated for 1 hour with 8 ml of goat (antimouse) IgG HRP conjugate secondary antibody (Novagen) diluted 1:5000 in blocking solution (without added BSA). Any unbound secondary antibody was removed from the membrane with five washes for 10 minutes each time with 20 ml of 1 X TBSTT. An ECL Plus Western blotting reagent (Amersham) was applied to the membrane (0.1 ml/cm<sup>2</sup>) and incubated for 5 minutes, before being drained. The membrane was then placed in Saran Wrap, and exposed to photographic film and developed.

#### *4.2.4 Protease protection assays*

Trypsin digest experiments were typically carried out using samples of 10  $\mu$ M GST/B-Myb protein in 20 mM Tris, 100 mM sodium chloride buffer at pH 7.0, which were incubated with approximately 30  $\mu$ M p300 1603-1725 or with B-Myb 275-376 as a negative control for 1 hour at room temperature to allow complex formation. Additional control experiments were included where the GST fusion and p300 proteins were equilibrated alone. After equilibration, trypsin (Roche) was added to the protein samples at a ratio of 1:4000 (w/w) and the digest allowed to proceed for 2 hours at room temperature, with samples removed after 0, 30, 60, 90 and 120 minutes and placed in 50 mM DTT and 1 X LDS for reaction termination and analysis by SDS-PAGE.

#### *4.2.5 Monitoring protein-protein interactions by Intrinsic Tryptophan Fluorescence*

Interactions between p300 1726-1839 and B-Myb 275-376 were followed by acquiring fluorescence emission spectra for a series of samples containing 3  $\mu$ M B-Myb 275-376 and increasing concentrations of p300 1726-1839 from 0 to 15  $\mu$ M. The protein

samples were prepared in a 20 mM Tris, 100 mM sodium chloride buffer at pH 6.5 and incubated for 1 hour at room temperature before data acquisition. The same parameters were used as described in section 2.2.19.

#### *4.2.6 Chemical denaturation of the B-Myb275-376•p300 1726-1839 complex*

Intrinsic tryptophan fluorescence was used to probe the conformational stability of the B-Myb275-376•p300 1726-1839 complex. Fluorescence experiments were conducted with a series of 1  $\mu$ M samples of 1:1 B-Myb 275-376•p300 1726-1839 complex in a 20 mM Bis-tris, 100 mM sodium chloride buffer at pH 6.5, containing between 0 and 6 M guanidine hydrochloride, increasing with 0.5 M equivalents of denaturant up to 3 M guanidine hydrochloride, and 1M equivalents between 3 M to 6 M guanidine hydrochloride concentrations. All samples were incubated with denaturant overnight at 4 °C before data acquisition at 20 °C in a 1 cm pathlength fluorescence cuvette.

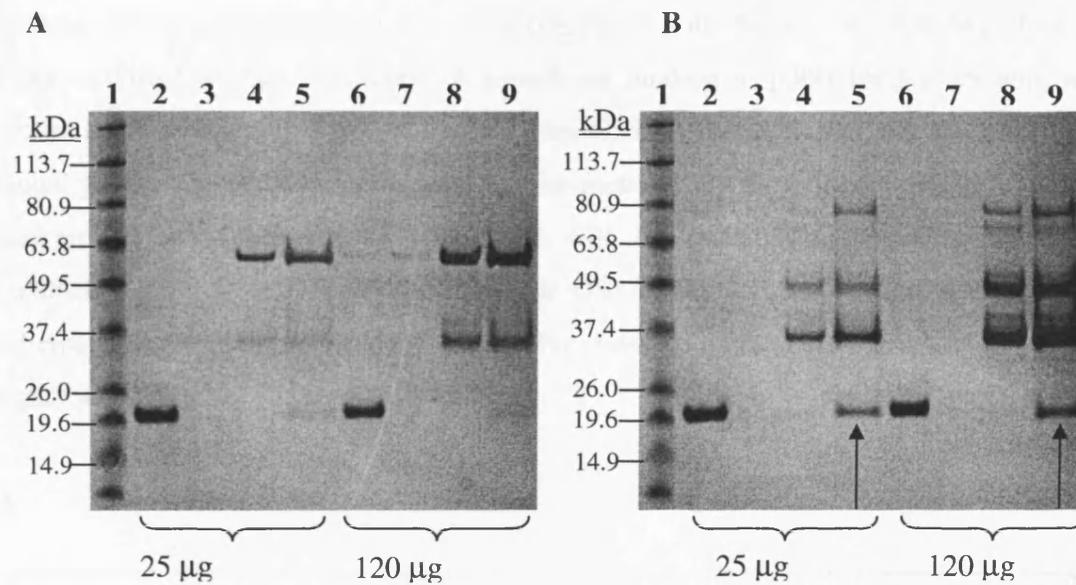
### *4.3 Results*

#### *4.3.1 Interactions between p300 1603-1725 and B-Myb*

##### *4.3.1.1 Batch-wise assays*

The results of SDS-PAGE analysis of typical batch-type pull down assays using GST/B-Myb 227-508 and GST/B-Myb 275-376 fusion proteins bound to glutathione-agarose resin as bait are shown in figures 4.1A and 4.1B. The data clearly indicate that p300 1603-1725 is not retained by beads with the shorter GST/B-Myb 275-376 fusion bound, but specifically binds to beads containing bound GST/B-Myb 227-508 fusion protein. A greater amount of GST/B-Myb protein is eluted from the resin when treated with LDS sample buffer than compared with elution from 10 mM reduced glutathione (observed when comparing lanes 4 and 5, and lanes 8 and 9 from each gel). It is very difficult to visualise co-eluted p300 1603-1725 when the complex is eluted using a 10 mM reduced glutathione buffer particularly due to the limits of detection imposed by coomassie staining the gels. Directly treating the GST/B-Myb bound beads with LDS sample buffer and heating gives a clear distinction in p300 1603-1725 binding to GST/B-Myb 227-508 rather than GST/B-Myb 275-376. Furthermore, increasing the GST/B-Myb 227-508 bait levels from 25  $\mu$ g to 120  $\mu$ g is accompanied by an increase in the amount of p300 1603-1725 retained by the beads, as expected if the two proteins

form a complex. The p300 1603-1725 protein alone was previously shown not to bind to glutathione agarose (data not shown), which clearly suggests that p300 1603-1725 binds specifically and relatively tightly to full length B-Myb 227-508 as the proteolytic products of the GST/B-Myb fusion protein are clearly smaller than GST/B-Myb 275-376 (which shows no evidence of binding to p300 1603-1725).



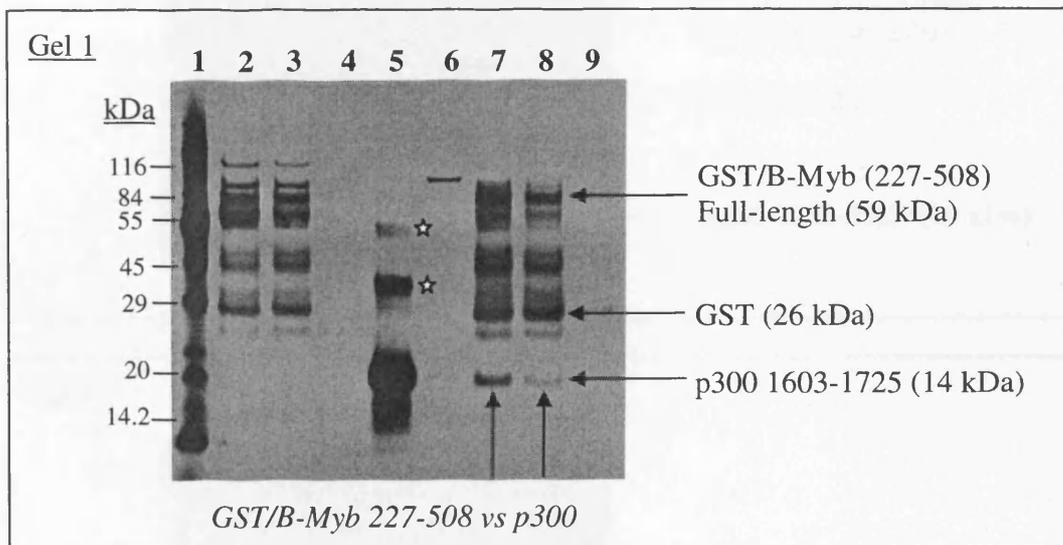
**Figure 4.1.** SDS PAGE gel analysis of batch-type GST fusion pull-down assays using (A) GST/B-Myb 275-376 and (B) GST-B-Myb 227-508 as the bait for p300 1603-1725. Lanes 2 and 6 on both gels corresponds to a sample of the 50 µg of p300 1603-1725 used in assays with about 25 µg and 120 µg of GST/B-Myb fusion proteins loaded onto the beads. Lanes 4 and 8 on both gels show proteins eluted after beads were treated with a 10 mM reduced glutathione buffer. Lanes 5 and 9 on both gels shows the proteins associated with the glutathione-agarose after treatment with LDS sample buffer and heating. Lanes 3 and 7 in both gels represents a sample removed from the final wash fraction prior to elution. Lane 1 in each of the gels contains BenchMark™ Prestained Protein Ladder (Sigma). Clearly GST/B-Myb 227-508 mediates the retention of p300 1603-1725 whereas there is no evidence for an association with GST/B-Myb 275-376.

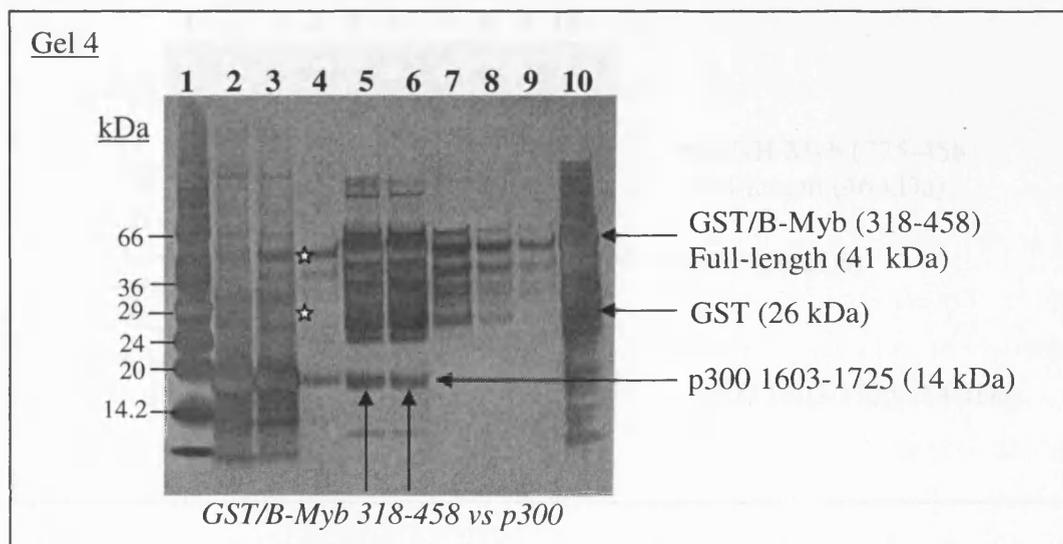
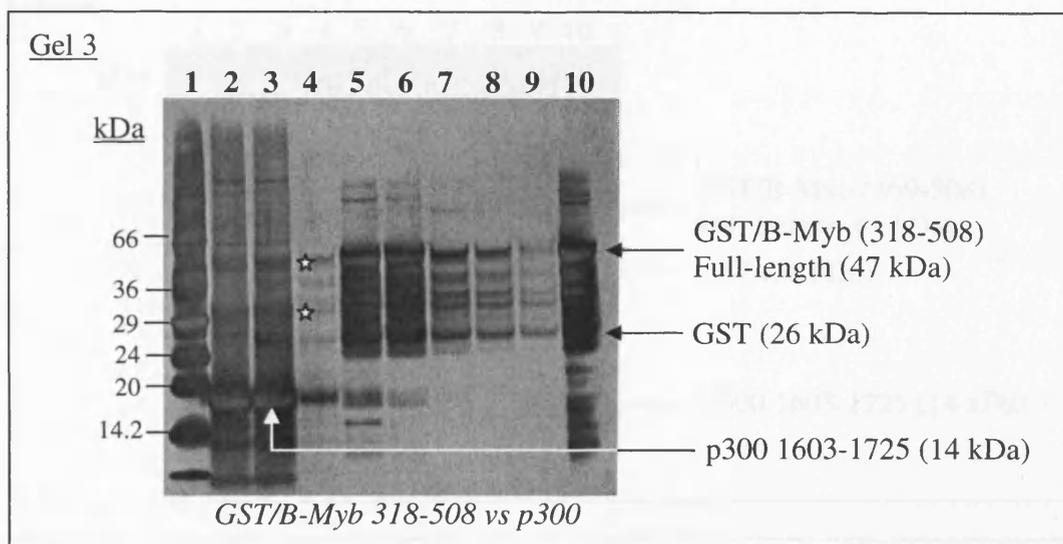
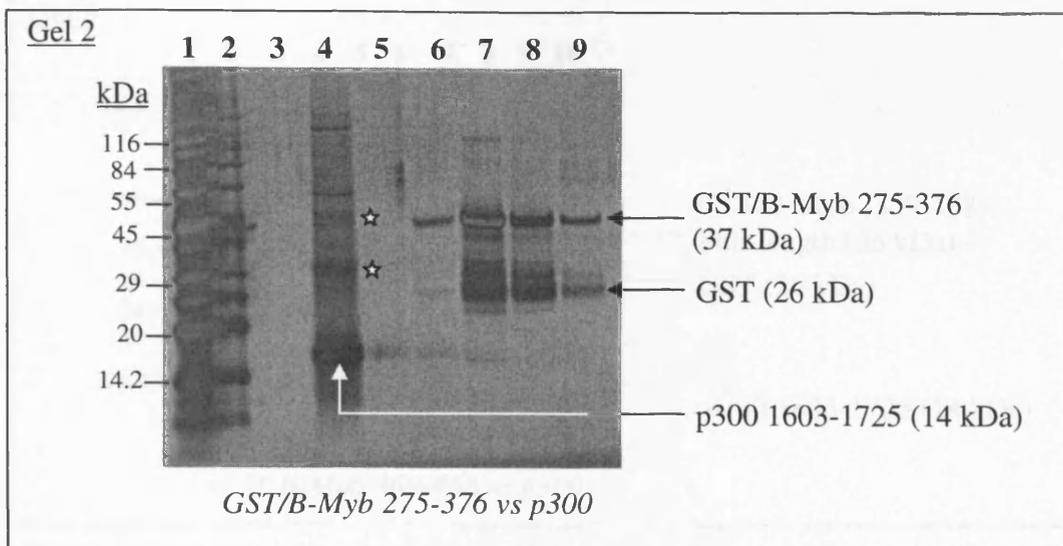
#### 4.3.1.2 Column-based pull-down assays

Column-based pull down assays using GST/B-Myb 227-508 and GST/B-Myb 275-376 proteins confirmed the binding of p300 1603-1725 by GST/B-Myb 227-508 seen in the batch-wise approach. The silver-stained SDS-PAGE gels shown in figure 4.2 (A)

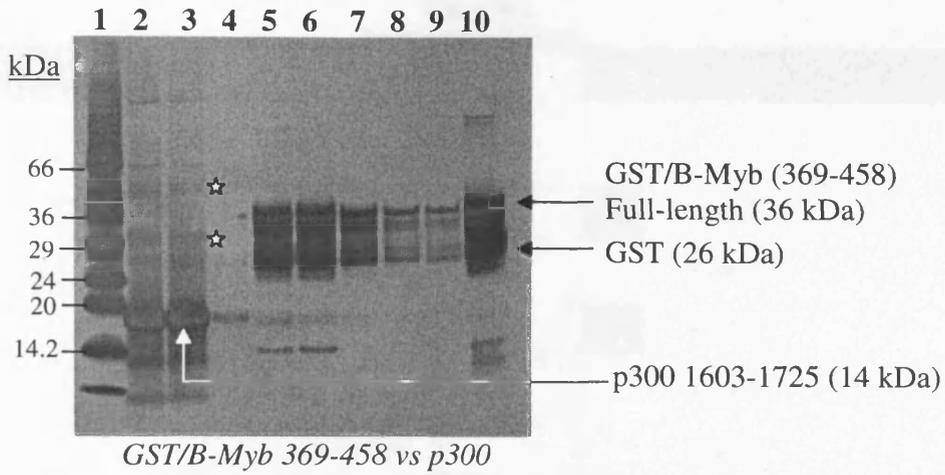
Gel 1 and Gel 2 illustrate typical results obtained from column-based pull down assays. The p300 1603-1725 protein is clearly specifically co-eluted with GST/B-Myb 227-508 but fails to bind to the shorter GST/B-Myb 275-376 fusion protein. The very small amounts of p300 1603-1725 detected in the eluted GST/B-Myb 275-376 fractions clearly correspond to background levels seen in the final wash and show no correlation with the quantities of eluted B-Myb fusion protein, which would be indicative of binding. Column based pull down assays completed with the other GST/B-Myb fusions however failed to show evidence of significant binding to p300 1603-1725 with any reproducibility. As with the binding assay for GST/B-Myb 275-376, the amount of co-eluted p300 1603-1725 observed in the experiments is not significantly greater than the background levels seen in the final wash with the possible exception of the assay conducted with GST/B-Myb 318-458 and GST/B-Myb 318-508 (Gel 3 and 4). A schematic representation of the GST/B-Myb fusions used in this study is shown in figure 4.2 (B).

A

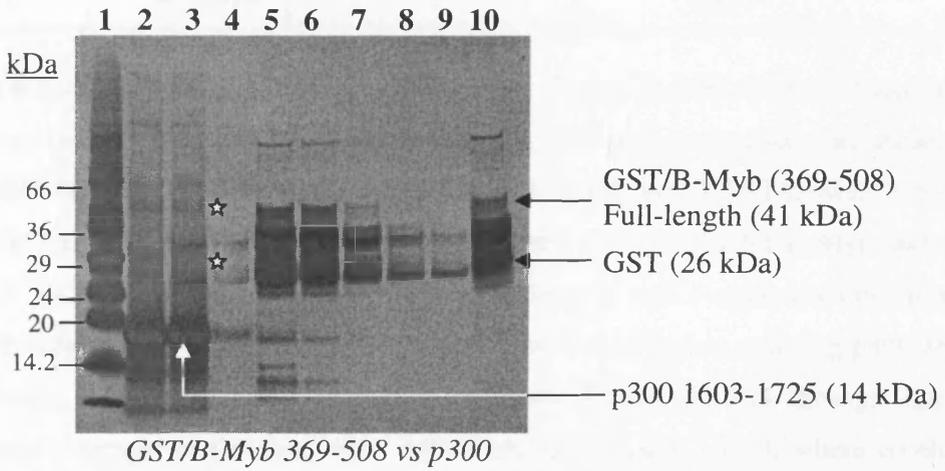




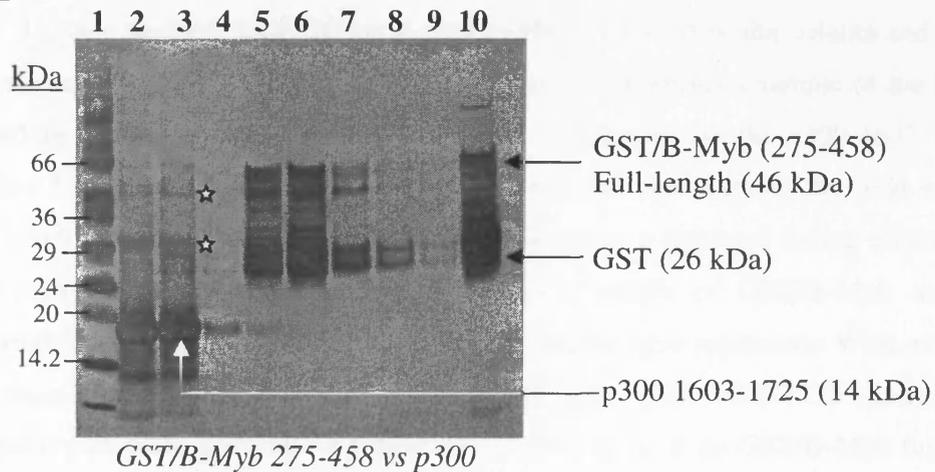
Gel 5

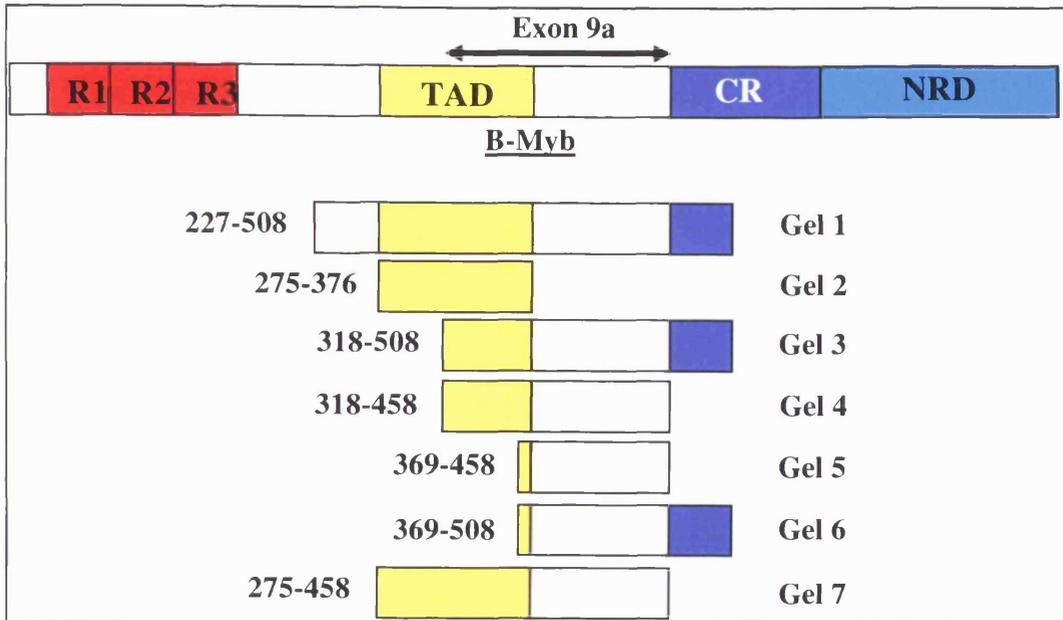


Gel 6



Gel 7



**B**

**Figure 4.2.** Silver-stained SDS PAGE gel analysis of typical p300 1603-1725 pull down assays using GST-tagged B-Myb bait proteins (A) GST pull-down assays are shown for Gel 1-GST/B-Myb 227-508, Gel 2-GST/B-Myb 275-376, Gel 3-GST/B-Myb 318-508, Gel 4-GST/B-Myb 318-458, Gel 5-GST/B-Myb 369-458, Gel 6-GST/B-Myb 369-508, and Gel 7-GST/B-Myb 275-458. In Gel 1 both lanes 2 and 3 represent GST/B-Myb 227-508 column flow-through, lane 4 shows the final wash before applying p300 1603-1725 to the column and lane 5 shows a sample taken from the flow-through. Lane 6 represents a sample of the final wash before elution in lanes 7 to 9, where co-eluted p300 1603-1725 is highlighted by the vertical arrows. In Gel 2, lane 2 contains GST/B-Myb 275-376 column flow-through but is obscured somewhat from the marker in lane 1. Lane 3 shows the final wash before applying p300 1603-1725 to the column and lane 4 shows a sample taken from the flow-through. Lane 5 represents a sample of the final wash before elution in lanes 6 to 9. In Gels 3 to 7, lane 2 shows the p300 1603-1725 load, lane 3 shows the p300 1603-1725 flow-through and lane 4 shows the final wash. Lanes 5 to 9 correspond to samples taken from consecutive fractions during elution of bound proteins by glutathione. Lane 10 shows a sample of GST/B-Myb at the concentration used in the assay. Lane 1 for each of the gels represents Wide range/ Broad range Molecular marker (Sigma). Dimer and trimer p300 1603-1725 species are highlighted with white stars. (B) A schematic representation of the GST/B-Myb fusions

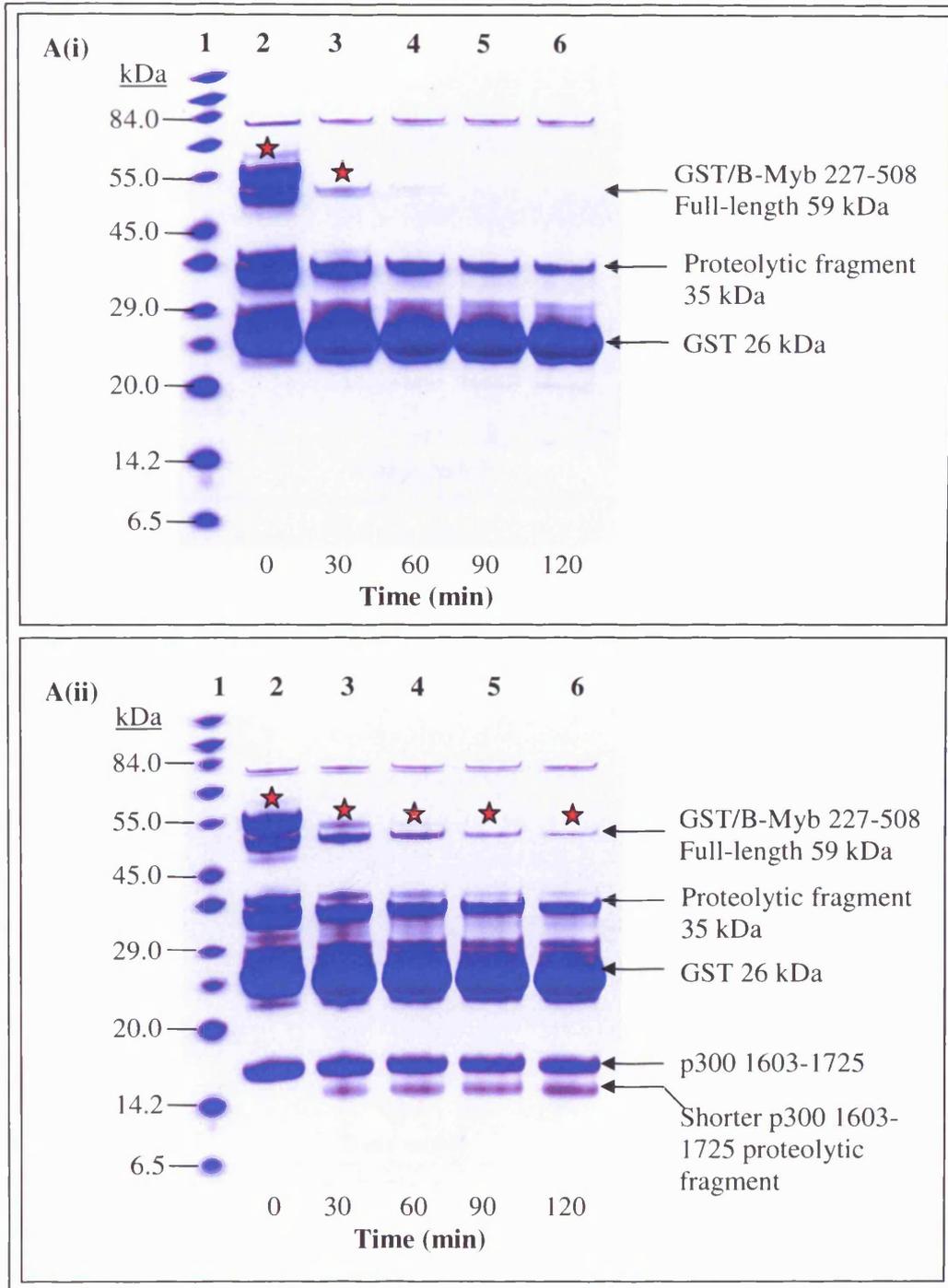
used in this study with the corresponding gel numbers relating to the GST pull-down assays labelled in (A) are indicated.

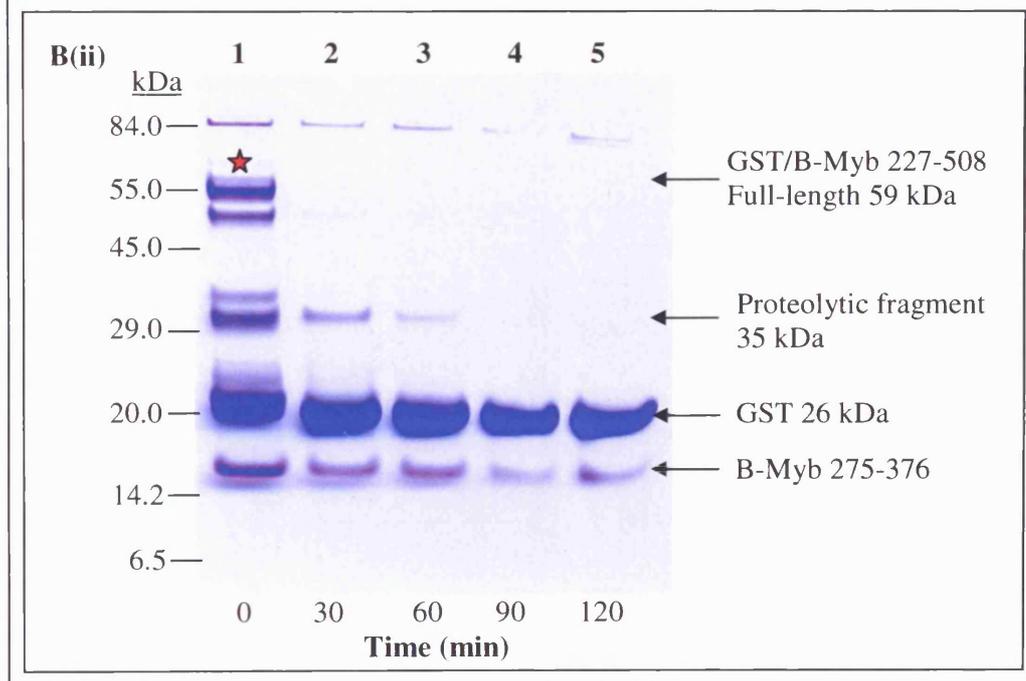
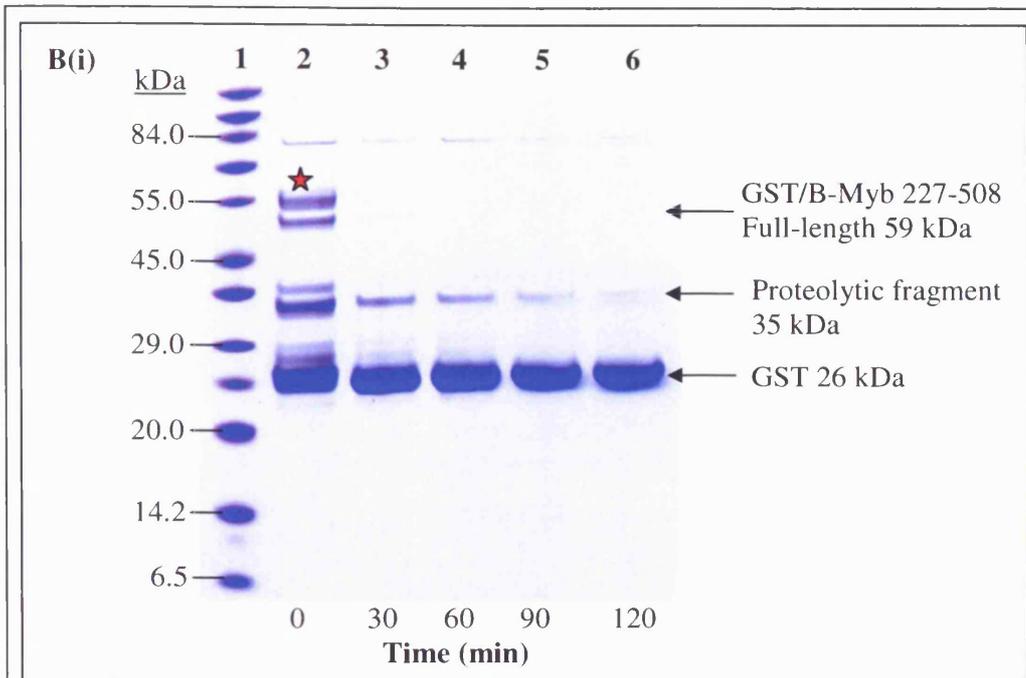
#### *4.3.1.3 Protease protection assays*

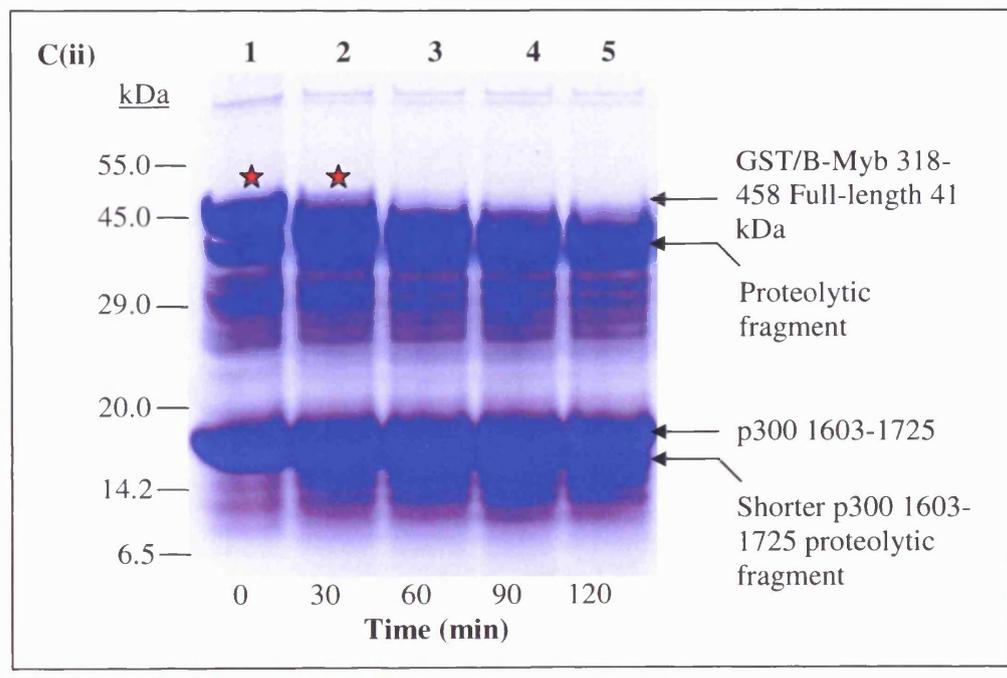
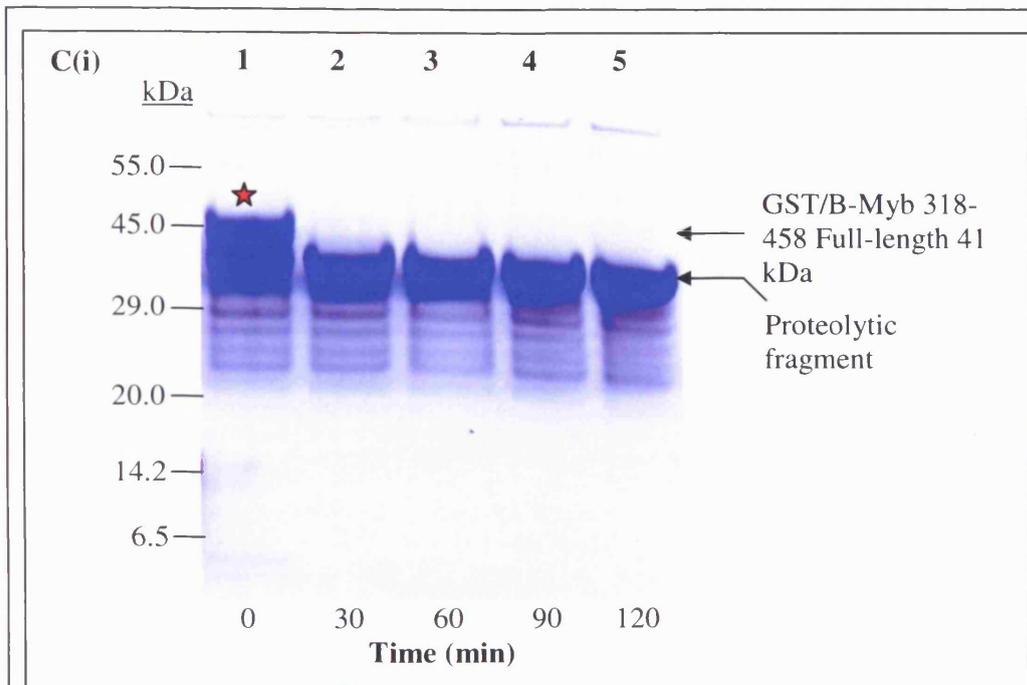
Since the formation of a specific protein-protein complex may well reduce the number of protease sensitive sites on one or both proteins, trypsin was used to monitor the extent of protease protection of GST/B-Myb and p300 1603-1725 on forming a complex. In a series of experiments the patterns of limited trypsin digestion of GST/B-Myb proteins were compared in both the presence and absence of a molar excess of p300 1603-1725. Figure 4.3 shows the results of SDS PAGE analysis of the limited trypsin digests. A comparison of the SDS PAGE gels shown for experiments with GST/B-Myb 227-508 (figure 4.3A (i and ii)), clearly indicates that GST/B-Myb 227-508 is protected from trypsin digestion in the presence of p300 1603-1725. This protection from trypsin digestion is manifest by the reduced rate of cleavage of the full-length GST/B-Myb 227-508 species highlighted by the red stars. In contrast, when p300 1603-1725 is replaced by B-Myb 275-376 (figures 4.3B), no protection is observed, underlining the specificity of the interaction between B-Myb 227-508 and p300 1603-1725.

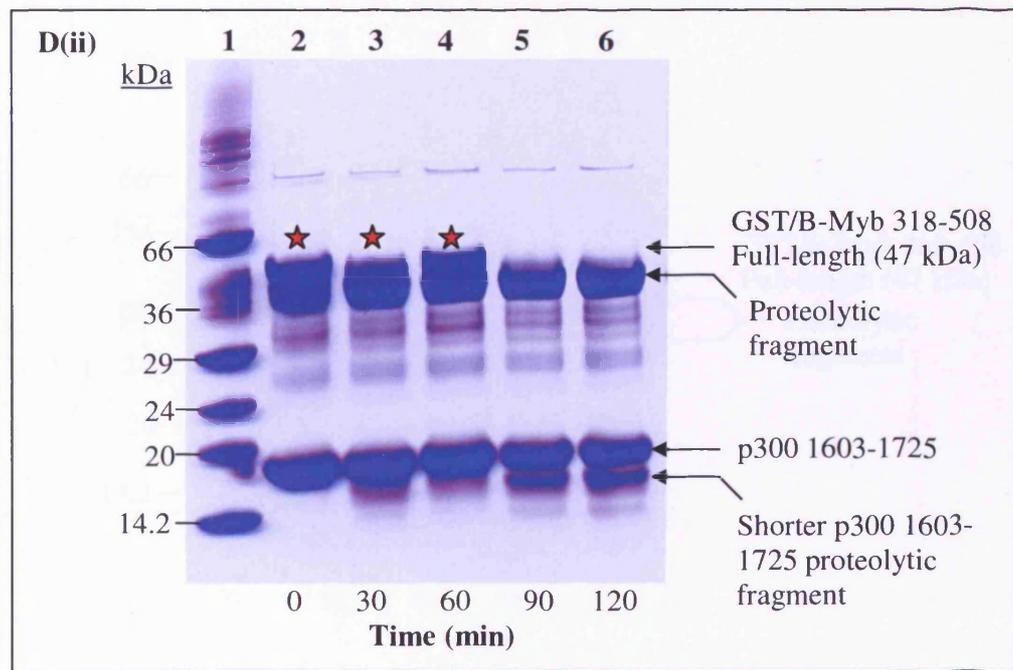
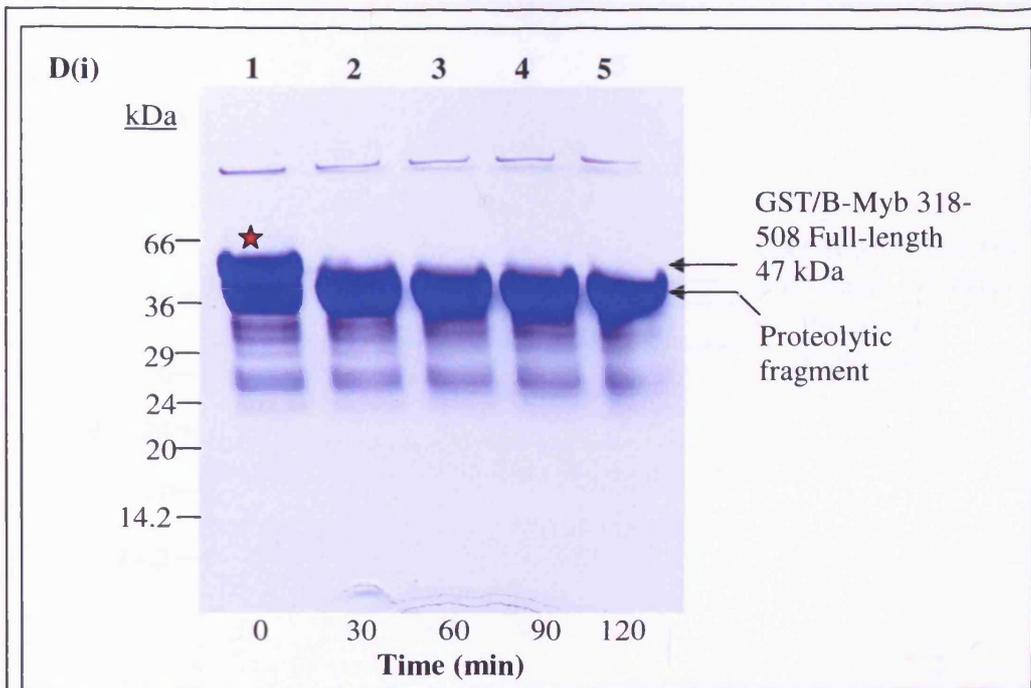
Similarly a degree of protease protection is observed for assays conducted with GST/B-Myb 318-458, 318-508 and 369-458 proteins when comparing SDS PAGE gels in the absence and presence of a molar excess of p300 1603-1725 (shown in figures 4.3C, D and E respectively). In each case the rate of cleavage of the full-length (or slightly smaller) GST/B-Myb fusion is noticeably reduced in the presence of p300 1603-1725 (highlighted by the red stars). By contrast little or no protection is observed for GST/B-Myb R2R3 and GST/B-Myb 275-376 proteins when digested with trypsin in the presence of p300 1603-1725 suggesting these regions of B-Myb do not interact with p300 1603-1725. GST/B-Myb 275-376 is composed of two major protein species; the full-length GST/B-Myb fusion, and a proteolysed GST component. During the digest experiment the fusion protein does not give rise to any further proteolytic fragments, but the GST protein component accumulates. In the presence of p300 1603-1725 there is no evident changes in this pattern of digest suggesting that there is little or no interaction between B-Myb 275-376 and p300 1603-1725. The same conclusion can be drawn from the digest experiment involving GST/B-Myb R2R3 in figure 4.3G where there is little

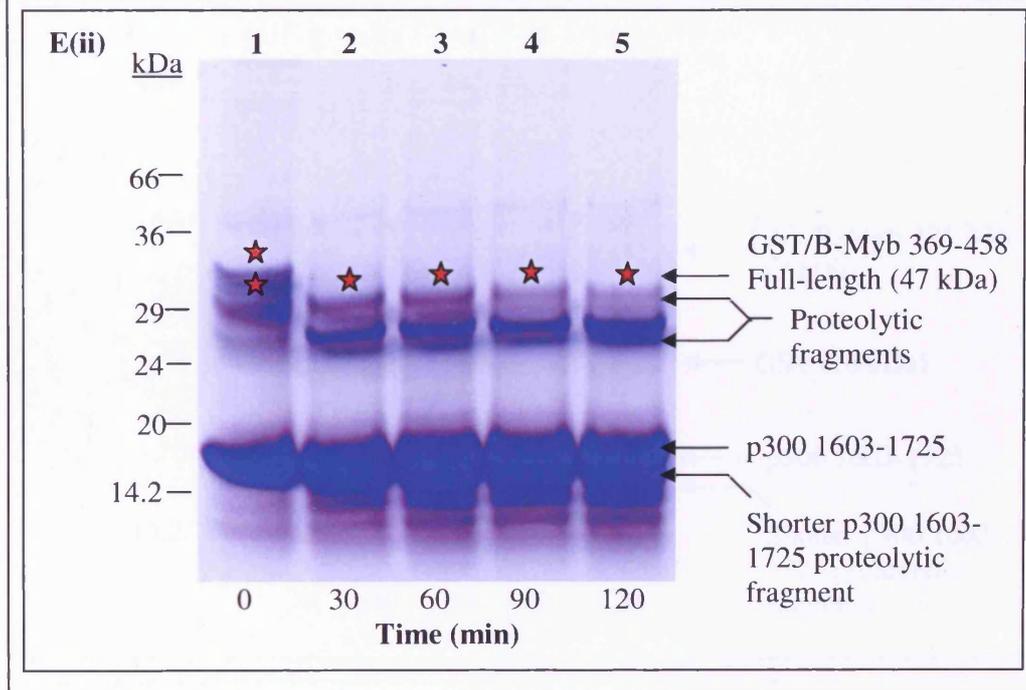
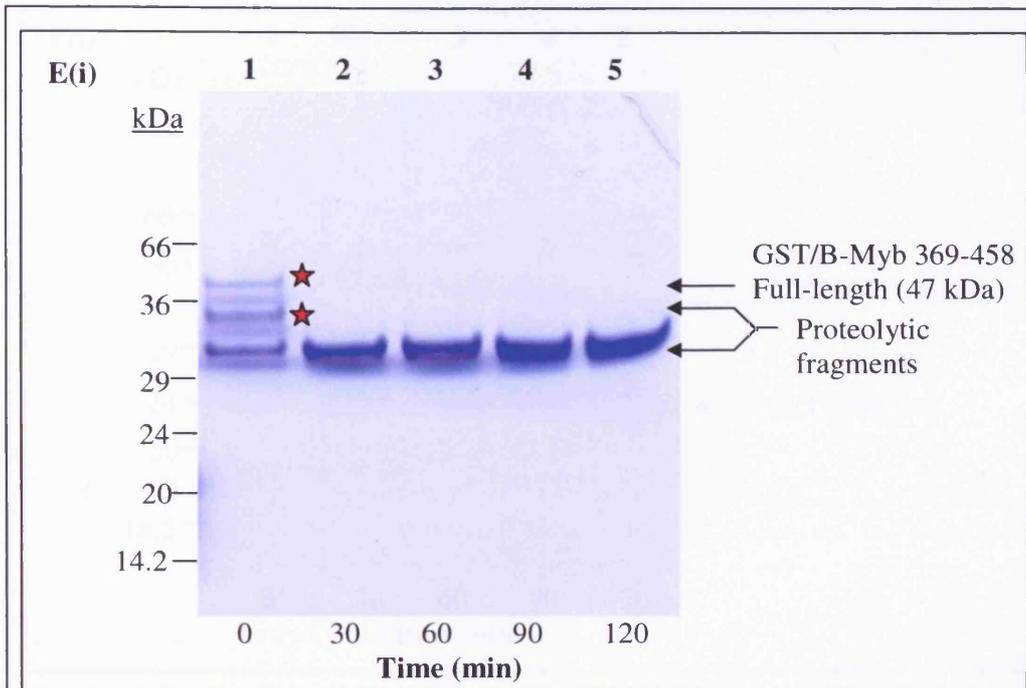
evidence of protease protection. The accumulation of a low molecular weight GST/B-Myb R2R3 proteolytic fragment (highlighted in figure 4.3G) during trypsin digestion is unaffected in the presence of p300 1603-1725.

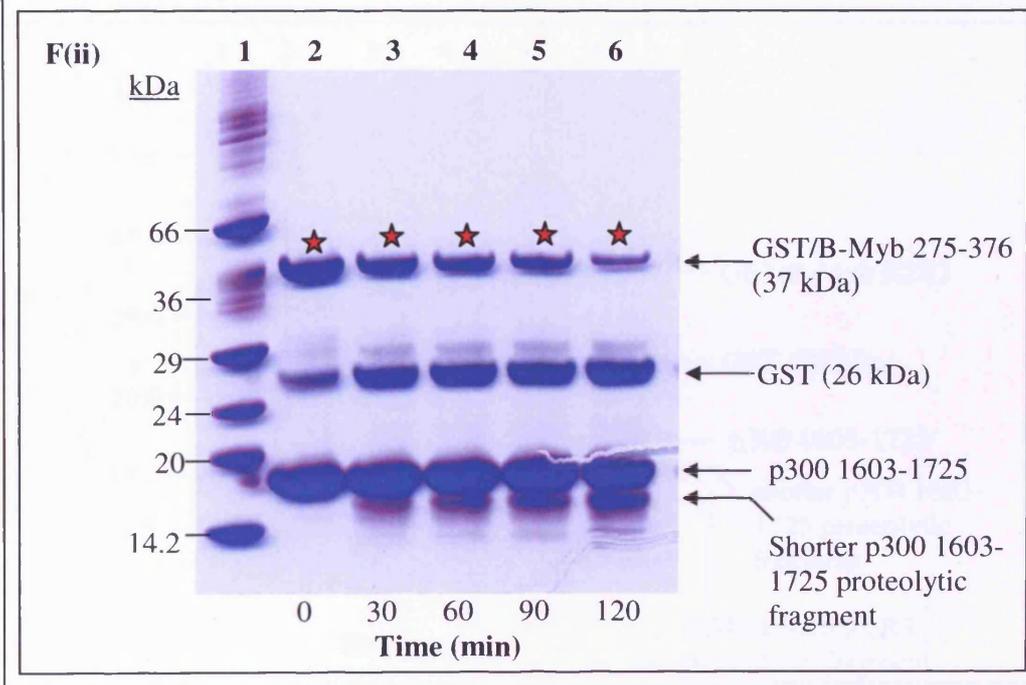
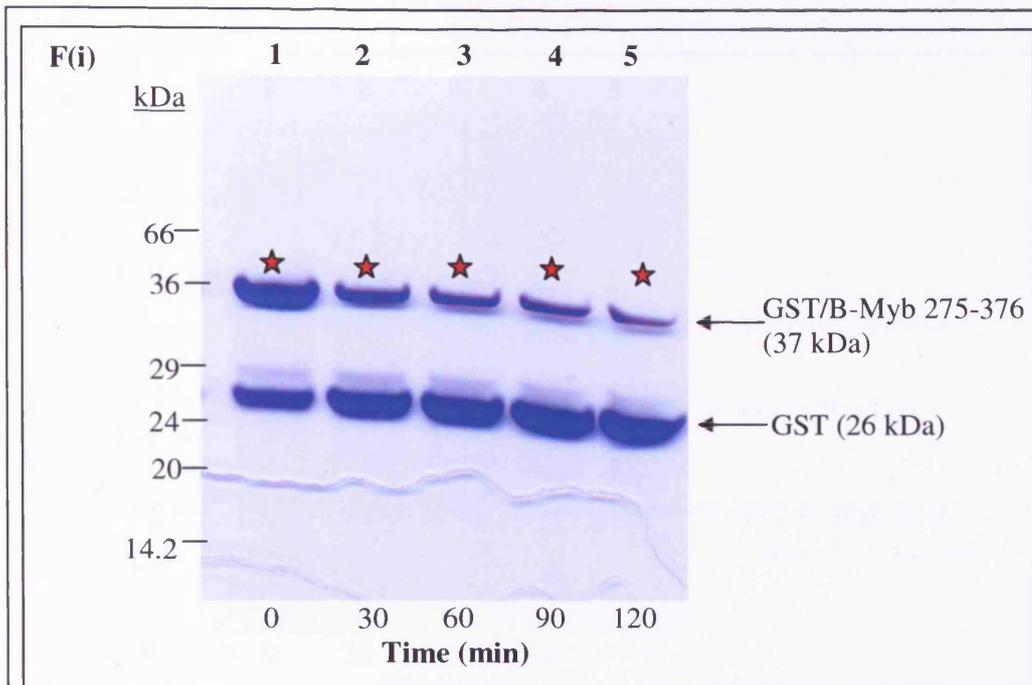


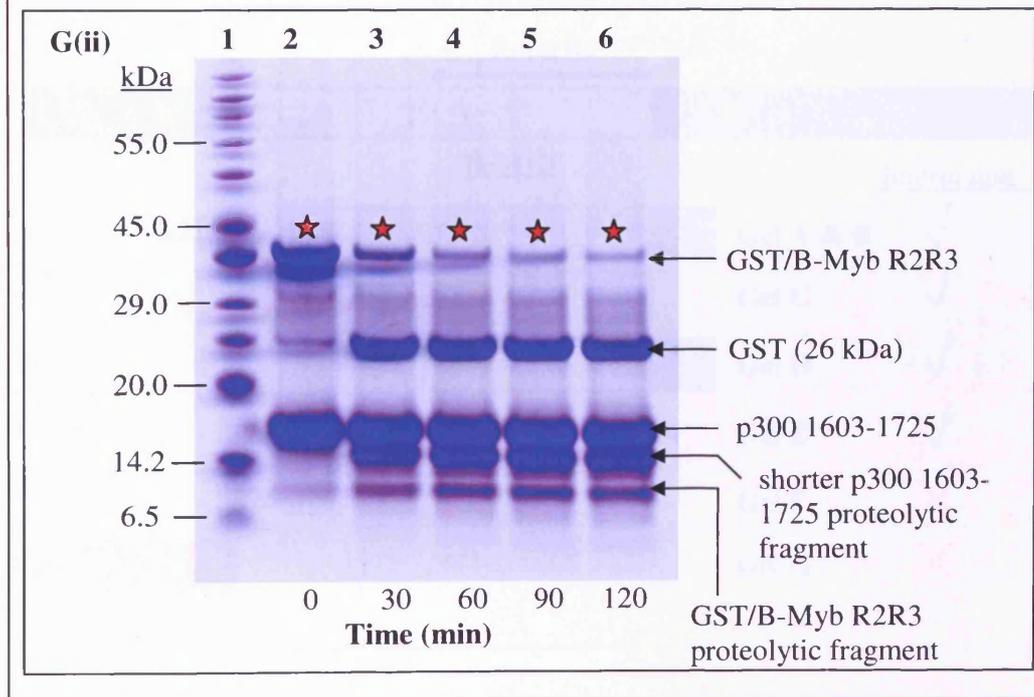
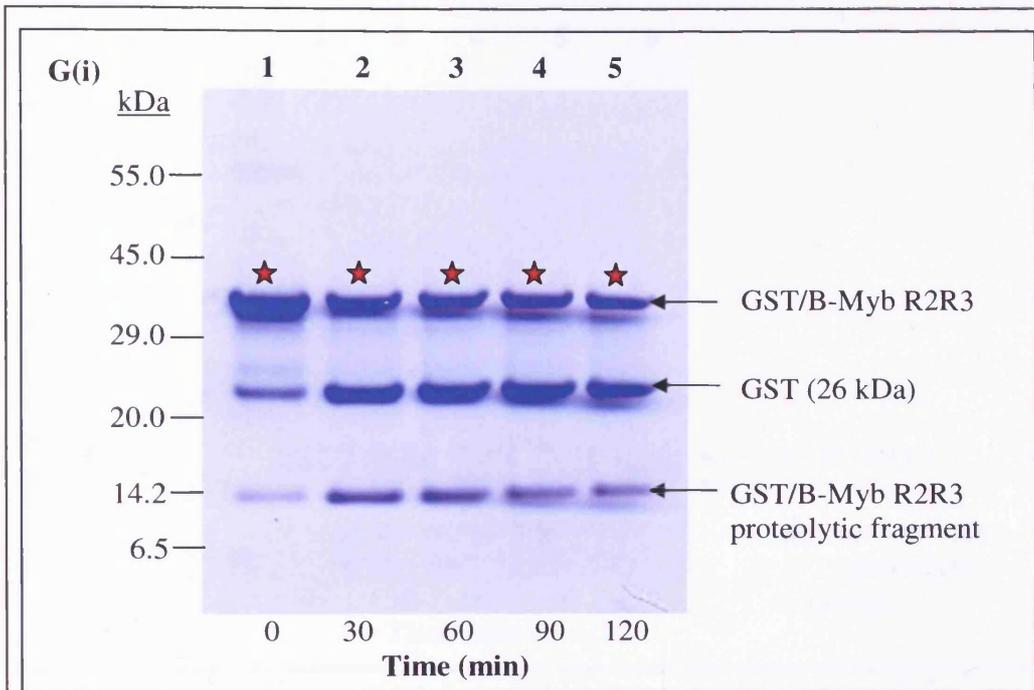


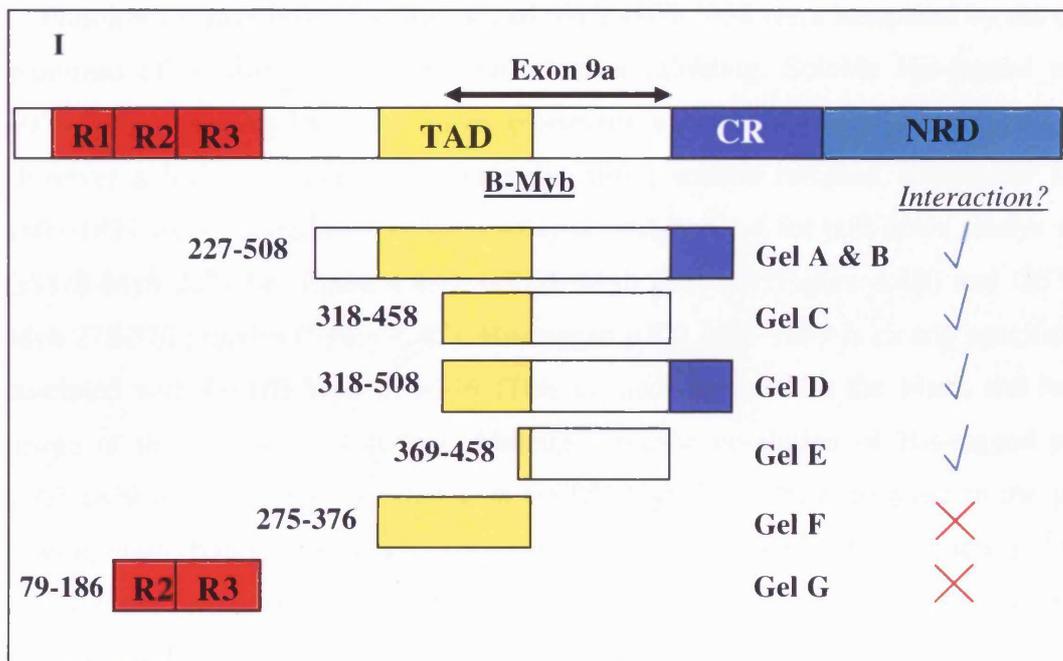
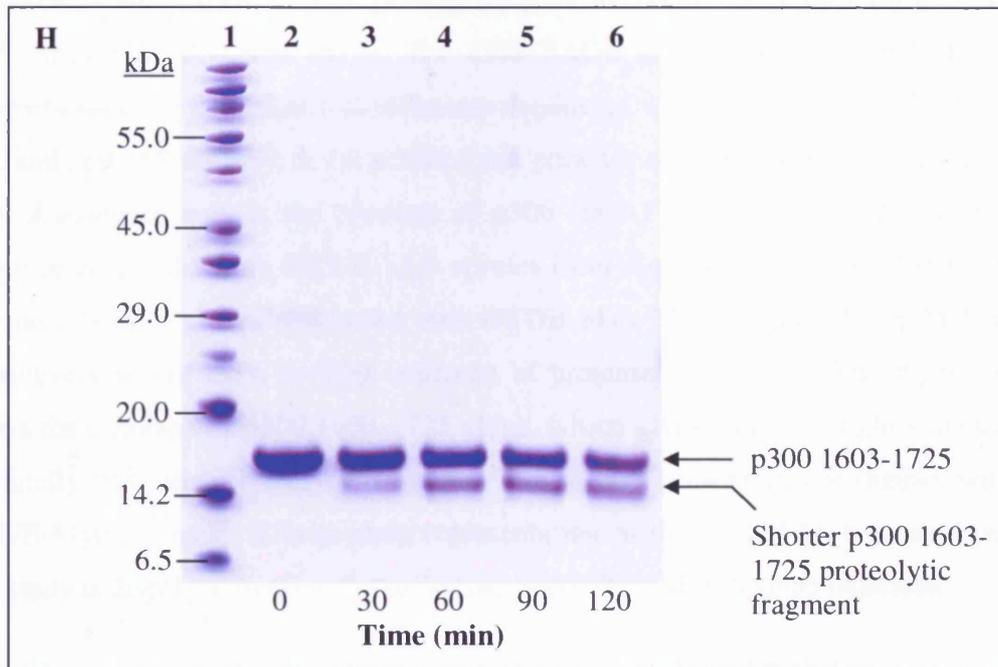










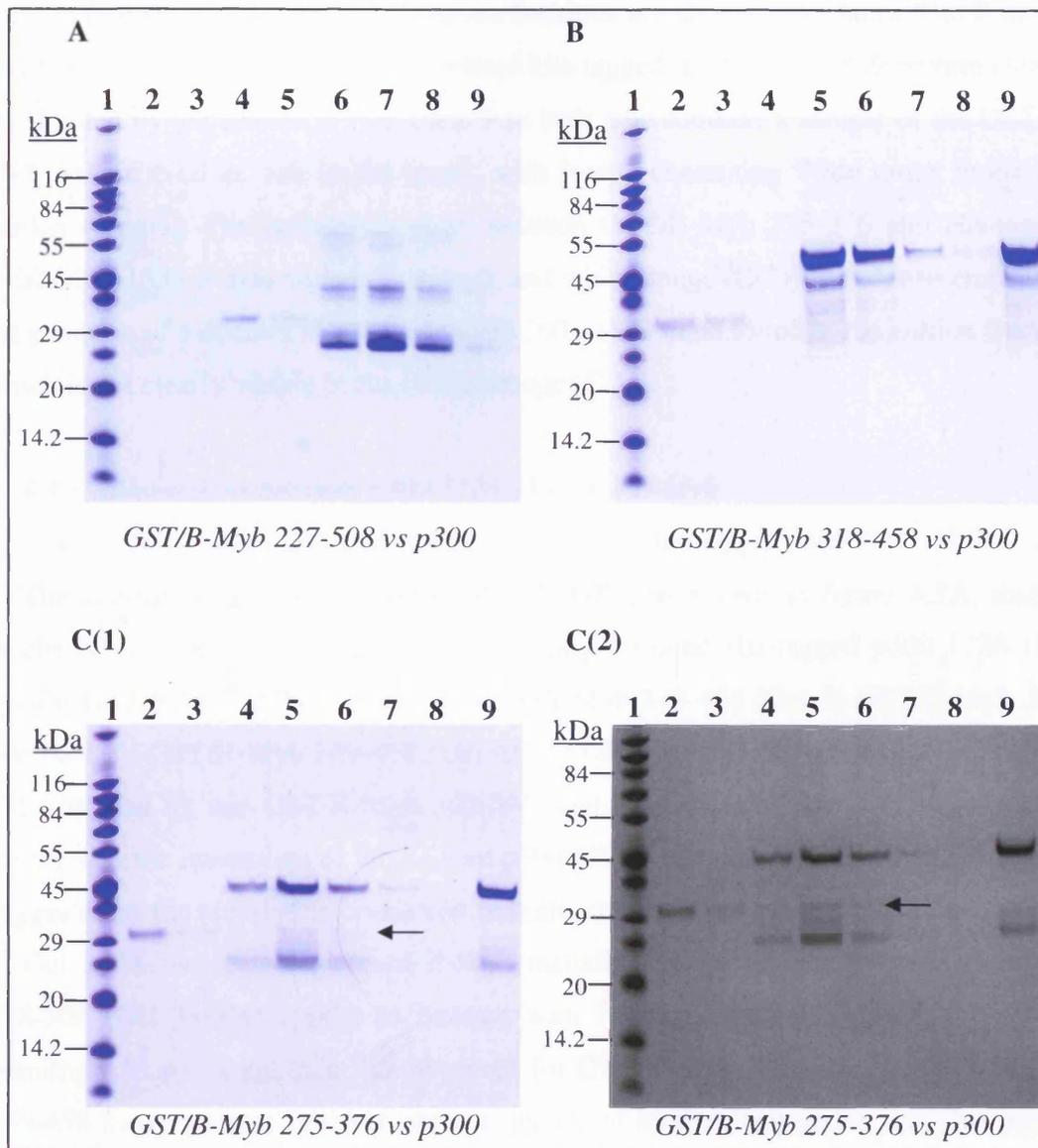


**Figure 4.3.** SDS PAGE gels of trypsin mediated protease protection assays involving GST/B-Myb proteins and p300 1603-1725. SDS PAGE gel analysis of trypsin digestion of GST/B-Myb 227-508 alone A(i) and in the presence of p300 1603-1725 A(ii) clearly indicates specific protection of full-length GST/B-Myb 227-508 by p300 1603-1725. The red stars highlight full-length or near full-length species of the GST/B-Myb protein.

When the assay is repeated with p300 1603-1725 substituted by B-Myb 275-376 there is no indication of protease protection suggesting the interaction is specific for p300 1603-1725 (shown in B(i) and B(ii)). The SDS PAGE gels in (C), (D) and (E) show comparisons of trypsin digest experiments conducted with GST/B-Myb 318-458, 318-508, and 369-458 proteins in the absence and presence of p300 1603-1725. In each case there is evidence that, in the presence of p300 1603-1725, there is protection of full-length or near full-length GST/B-Myb species from digestion. The SDS PAGE gels in (F) and (G) show assays conducted with GST/B-Myb 275-376 and GST/B-Myb R2R3 respectively where there is little evidence of protease protection. The digest in (H) shows the cleavage of p300 1603-1725 alone, which gives rise to a slightly smaller C-terminally truncated proteolytic fragment that is also present in the digests with the GST/B-Myb proteins. A schematic representation of the GST/B-Myb fusions used in this study is displayed in (I) with the interacting GST/B-Myb fusions indicated.

#### *4.3.2 Interactions between His-tagged p300 1603-1839 and B-Myb*

Pull-down assays involving His-tagged p300 1603-1839 were hampered by the poor quantities of soluble polypeptide attained after refolding. Soluble His-tagged p300 1603-1839 was short lived due to its propensity to precipitate and form aggregates. However a limited number of experiments using soluble refolded His-tagged p300 1603-1839 were completed and are displayed in figure 4.4 for pull-down assays with GST/B-Myb 227-508 (figure 4.4A), GST/B-Myb 318-458 (Figure 4.4B) and GST/B-Myb 275-376 proteins (figure 4.4C). His-tagged p300 1603-1839 is clearly specifically co-eluted with GST/B-Myb 275-376 (This is more apparent in the black and white image of the gel, figure 4.4C(ii)). Although specific co-elution of His-tagged p300 1603-1839 is not clearly observed with GST/B-Myb 227-508, a decrease in the p300 flow-through (figure 4.4A, lane 5) suggests that the proteins may be interacting. There is no clear indication of an interaction with GST/B-Myb 318-458 for the assay displayed in figure 4.4B.



**Figure 4.4.** SDS PAGE gel analysis of GST pull-down assays involving His-tagged p300 1603-1839 with GST-tagged B-Myb 227-508 (A), B-Myb 318-458 (B) and B-Myb 275-376 (C) bait proteins. (A) Lane 2 contains a sample from the GST/B-Myb 227-508 column flow-through, and lane 3 represents a subsequent wash fraction (to rid unbound GST fusion protein). Lane 4 contains a sample of the His-tagged p300 1603-1839 column load, with a sample from the column flow-through shown in lane 5. Consecutive elution fractions are shown from lanes 6 to 9. Lane 1 contains Wide range molecular marker (Sigma). For the SDS PAGE gels in (B) and (C) lane 2 contains a sample of His-tagged p300 1603-1839 column load, with a sample from the column

flow-through shown in lane 3. Lane 4 in (B) contains a sample taken from the final wash before elution. Consecutive elution fractions are shown from lanes 5 to 8 in (B) and from lanes 4 to 8 in (C) with co-eluted His-tagged p300 1603-1839 protein clearly highlighted by the arrows in (C). Lane 9 in both gels contains a sample of the GST/B-Myb protein used as bait in the assay, with lane 1 containing Wide range molecular marker (Sigma). The pull-down assay between GST/B-Myb 275-376 and His-tagged p300 1603-1839 is also shown as a black and white image (C(2)) in order to emphasise the presence of a distinct His-tagged p300 1603-1839 band found in the elution fraction which is not clearly visible in the colour image (C(1)).

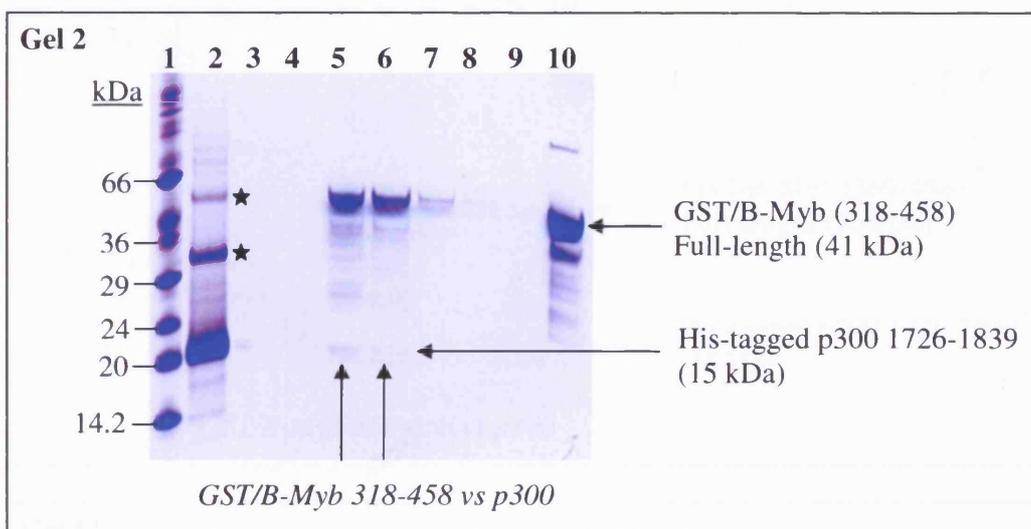
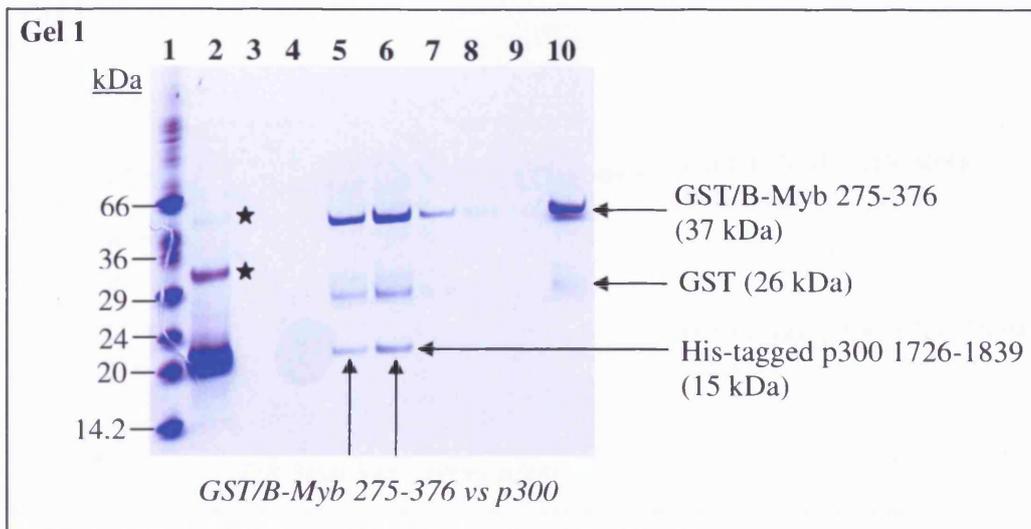
#### *4.3.3 Interactions between p300 1726-1839 and B-Myb*

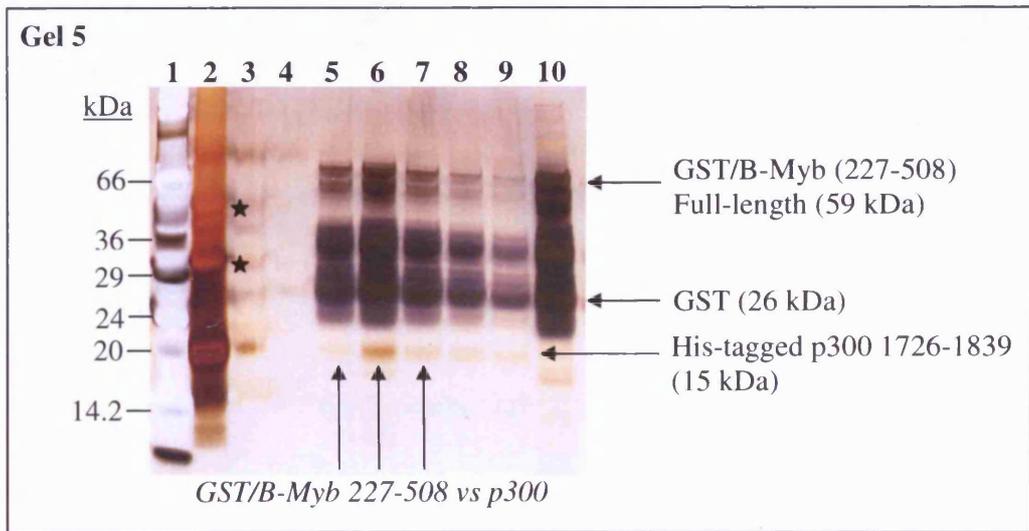
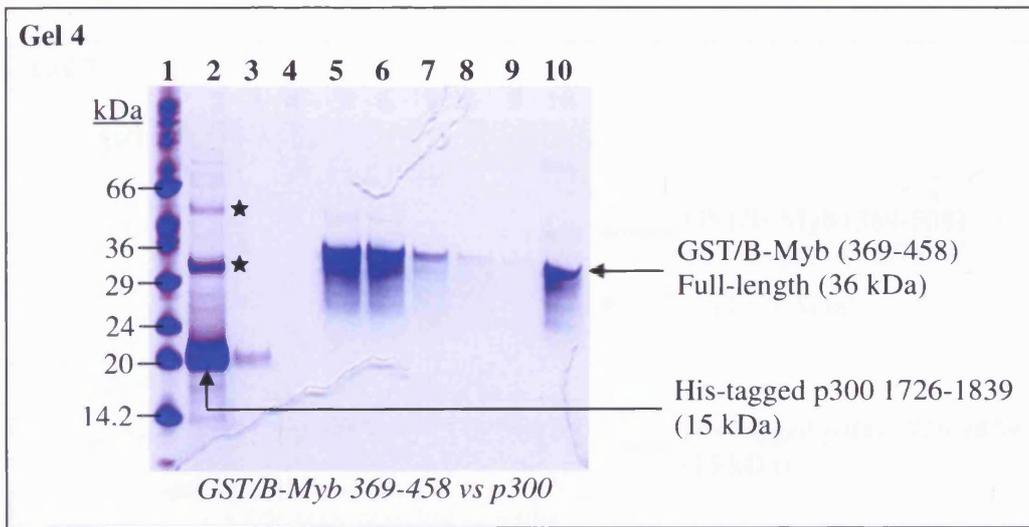
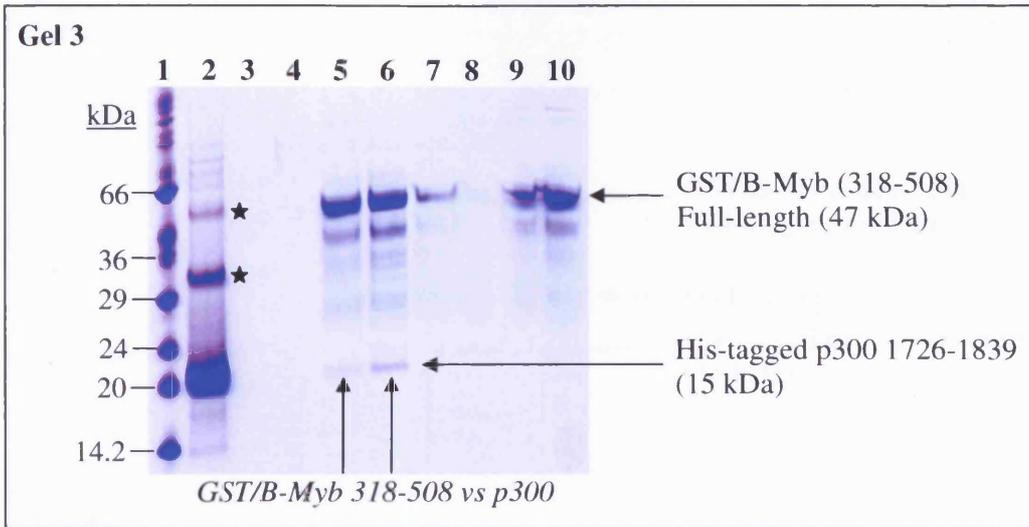
##### *4.3.3.1 Column GST pull-down assays using His-tagged p300 1726-1839*

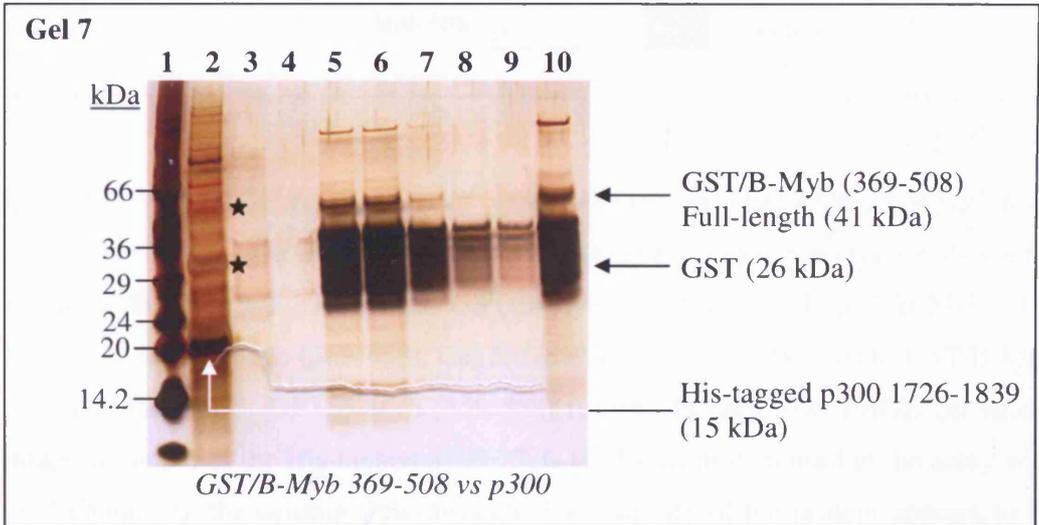
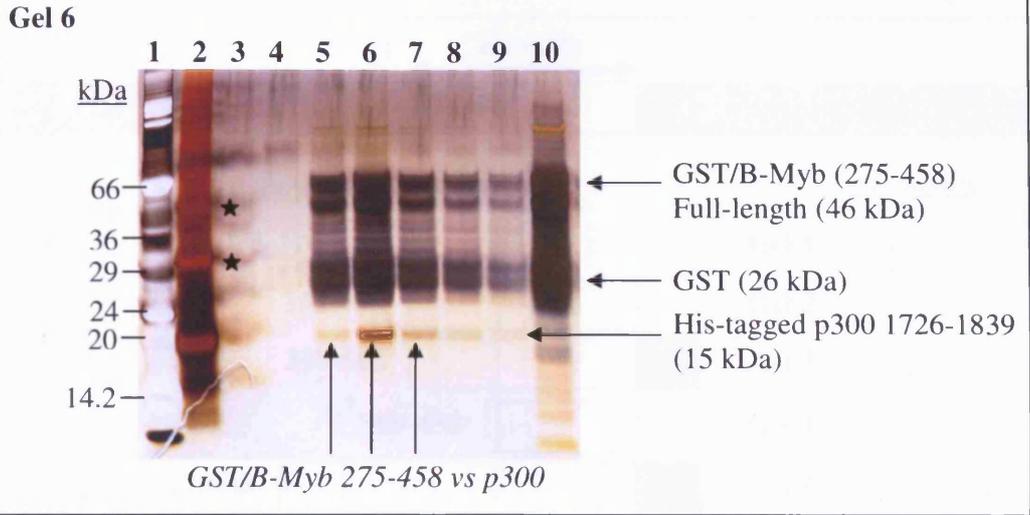
The coomassie and silver-stained SDS-PAGE gels shown in figure 4.5A, display results of column GST pull-down assays using refolded His-tagged p300 1726-1839 against GST/B-Myb 275-376 (Gel 1), GST/B-Myb 318-458 (Gel 2) GST/B-Myb 318-508 (Gel 3), GST/B-Myb 369-458 (Gel 4), GST/B-Myb 227-508 (Gel 5), GST/B-Myb 275-458 (Gel 6), and GST/B-Myb 369-508 (Gel 7) proteins. There is strong evidence for the specific interaction of His-tagged p300 1726-1839 with GST/B-Myb 275-376 as suggested by the correlating co-elution of both proteins observed in the elution fractions of Gel 1. The adjacent regions of B-Myb including amino acids 318-458 (Gel 2) and 318-508 (Gel 3) also appear to interact with His-tagged p300 1726-1839 but to a seemingly lesser extent than that observed for GST/B-Myb 275-376. The GST/B-Myb 369-458 fusion protein failed to show evidence of binding suggesting that this region does not provide the required interaction surface for p300 1726-1839 (Gel 4).

Further assays were visualised by silver-staining SDS PAGE gels for greater detection sensitivity of potentially co-eluted protein. Both GST/B-Myb 227-508 and GST/B-Myb 275-458 fusions exhibit binding to His-tagged p300 1726-1839 as shown in the silver-stained gels (Gels 5 and 6 respectively). The GST/B-Myb 369-508 fusion however showed no evidence of binding (Gel 7). A lack of binding was also confirmed for GST/B-Myb 369-458 by silver-staining (Data not shown). From the differential binding affinities exhibited by each of the GST/B-Myb fusions, it can be deduced that the B-Myb region from amino acids 227 to 369 provides the principal binding site for p300 1726-1839 (summarised in figure 4.5B).

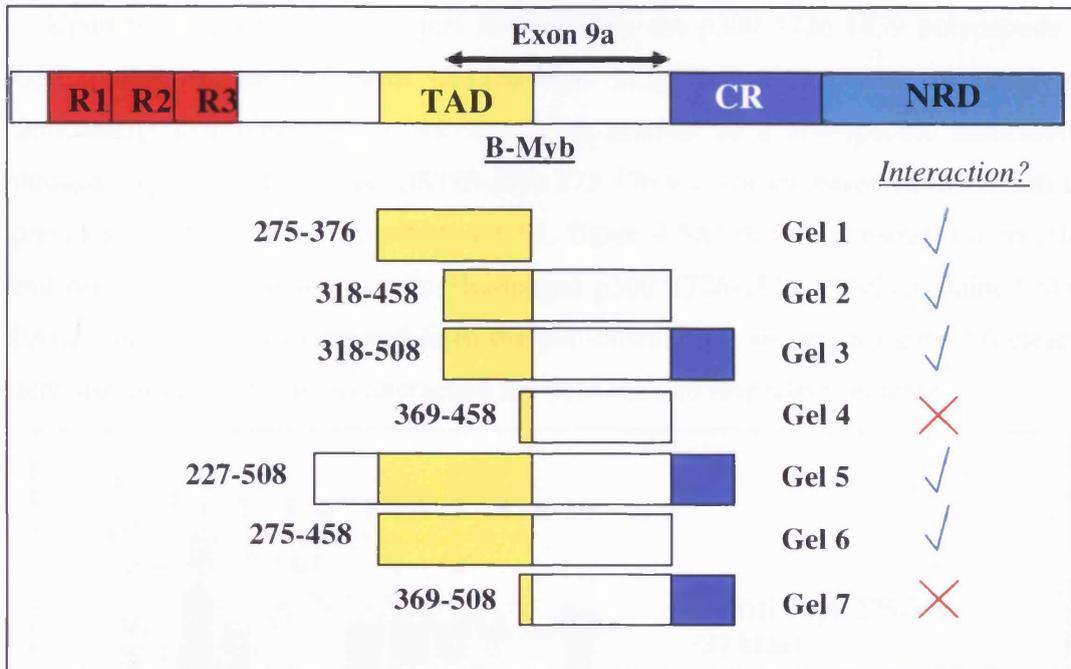
A







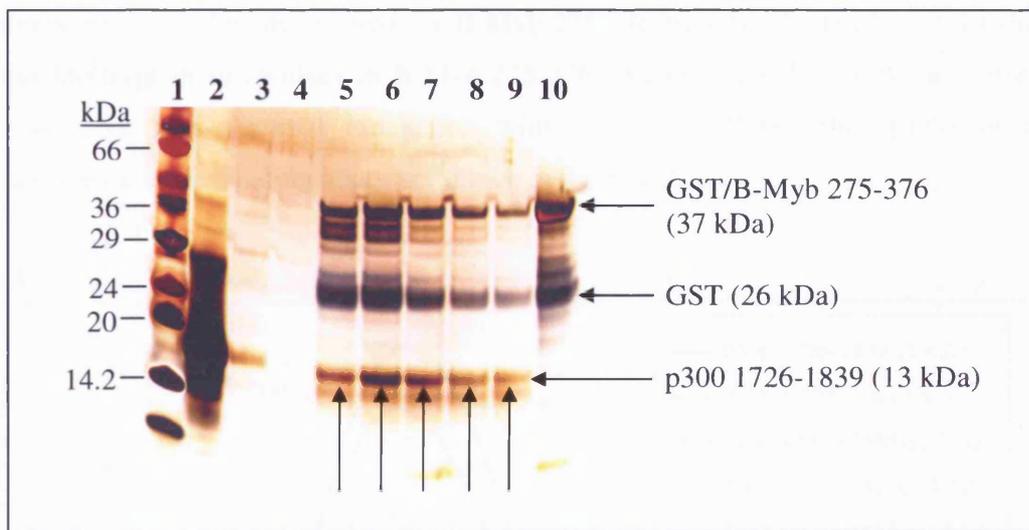
B



**Figure 4.5.** SDS PAGE gel analysis of typical His-tagged p300 1726-1839 pull down assays using GST-tagged B-Myb bait proteins. (A) GST pull-down assays are shown for Gel 1- GST/B-Myb (275-376), Gel 2- GST/B-Myb (318-458), Gel 3- GST/B-Myb (318-508), Gel 4- GST/B-Myb (369-458), Gel 5- GST/B-Myb (227-508), Gel 6- GST/B-Myb (275-458), and Gel 7- GST/B-Myb (369-508) fusions. In each SDS PAGE gel lane 2 contains a sample of the His-tagged p300 1726-1839 load protein used in the assay with lane 3 containing the column flow-through. The majority of the protein appears to be monomeric with a small amount of dimer and trimer species present, highlighted by the black stars in lane 2 of each gel. Lane 4 in each gel represents a sample taken from the final wash fraction prior to elution, with lanes 5 to 9 displaying consecutive column elution fractions. Co-elution of His-tagged p300 1726-1839 is clearly observed for the pull-down assays conducted with GST/B-Myb (275-376) (Gel 1), GST/B-Myb (318-458) (Gel 2) GST/B-Myb (318-508) (Gel 3), GST/B-Myb (227-508) (Gel 5), and GST/B-Myb (275-458) (Gel 6) fusions and is highlighted by the vertical arrows in the elution lanes. Lane 10 in each gel shows a sample of the GST/B-Myb fusion used in the assay. Lane 1 in each gel contains broad range molecular marker (Sigma). (B) A schematic representation of the GST/B-Myb fusions used in this study with the interacting GST/B-Myb fusions indicated.

#### 4.3.3.2 Column GST-Pull down assay using p300 1726-1839 without the His-tag

Upon removal of the N-terminal His-tag from the p300 1726-1839 polypeptide a GST pull-down assay against GST/B-Myb 275-376 was repeated in order to demonstrate that binding was not merely an artefact of a non-specific association through the hexa-histidine tag. GST/B-Myb 275-376 was chosen based on the results of previous pull down assays (section 4.3.3.1, figure 4.5A) that demonstrated a specific and high affinity interaction with His-tagged p300 1726-1839. A silver-stained SDS PAGE gel of fractions removed from the pull-down assay shown in figure 4.6 clearly demonstrates evidence of an interaction between the two respective proteins.

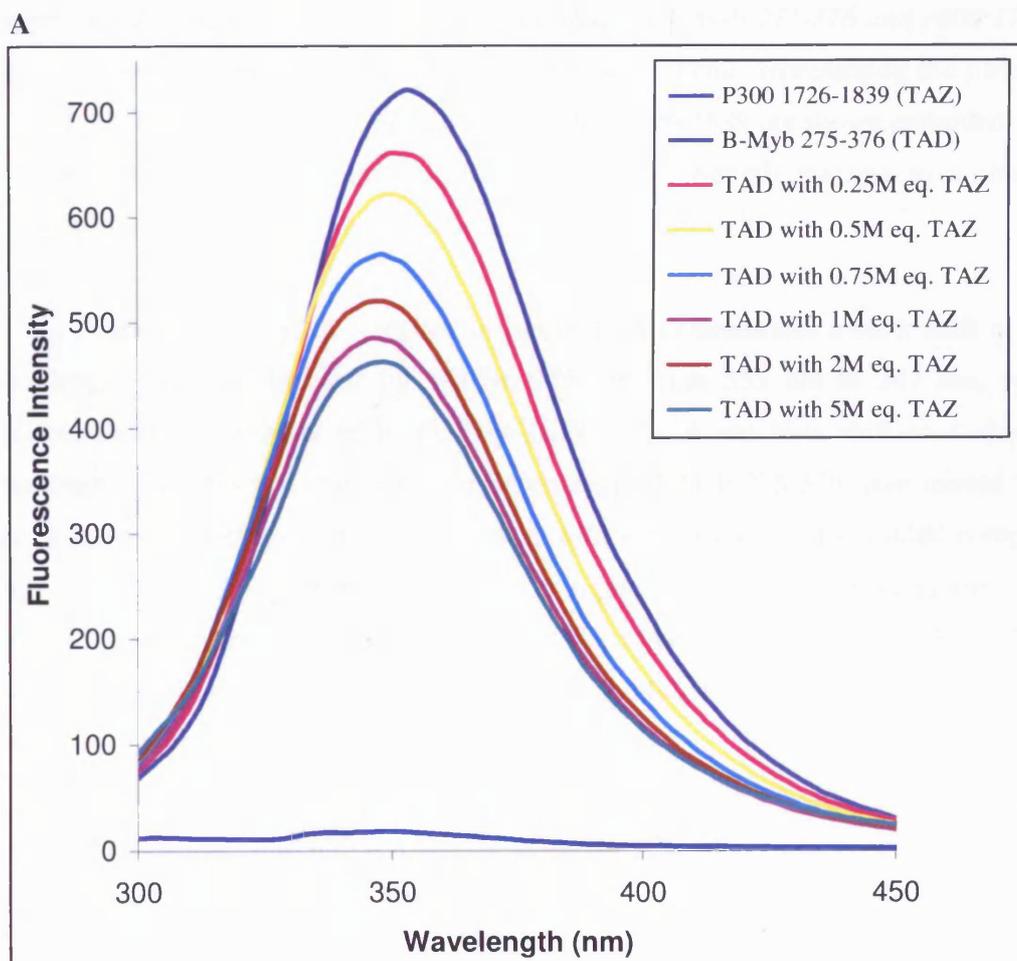


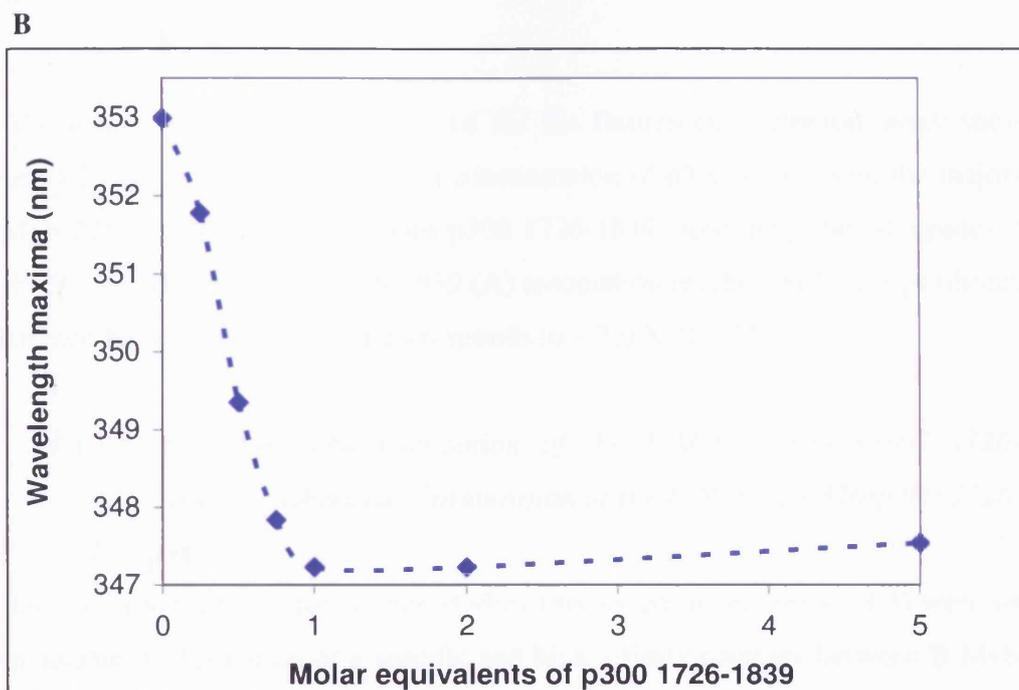
**Figure 4.6.** SDS PAGE gel analysis of a GST pull-down assay using p300 1726-1839 protein (without the His-tag) against GST/B-Myb (275-376) bait protein. Lane 2 contains a sample of the p300 1726-1839 load protein with lane 3 showing a sample from the resultant column flow-through. Lane 4 represents a sample taken from the final wash fraction prior to elution with lanes 5 to 9 displaying consecutive column elution fractions. Co-eluted p300 1726-1839 protein is clearly observed in the elution fractions highlighted by the vertical arrows from lanes 5 to 9. Lane 10 shows a sample of the GST/B-Myb 275-376 fusion protein used in the assay. Lane 1 contains broad range molecular marker (Sigma).

#### 4.3.3.3 Monitoring p300 1726-1839/B-Myb 275-376 complex formation by intrinsic tryptophan fluorescence.

The described GST pull-down assays clearly demonstrate that the central region of B-Myb, particularly amino acids 275 to 376, form a complex with the TAZ2 domain

containing region of p300, spanning amino acids 1726 to 1839. The GST/B-Myb 275-376 fusion is one of the few B-Myb deletion mutants purified in this study that does not undergo significant proteolysis. Furthermore, the GST tag is efficiently removed to yield a soluble polypeptide with random coil characteristics, as determined by the intrinsic tryptophan fluorescence studies reported in section 3.2.1.2. The absence of tryptophan residues in p300 1726-1839 precludes the direct assessment of its tertiary structure by intrinsic tryptophan fluorescence but does however provide a useful means of studying its complex formation with B-Myb 275-376 (which contains two tryptophan residues). If the two respective proteins do interact, a measurable perturbation in the fluorescence wavelength maxima for B-Myb 275-376 may be observed upon binding. Thus the tryptophan residues in B-Myb 275-376 (W293 and W323) serve as a useful reporter for any potential interaction with p300 1726-1839. The results of the fluorescence based binding assay are shown in figure 4.7.





**Figure 4.7.** *Intrinsic tryptophan fluorescence study of B-Myb 275-376 and p300 1726-1839 complex formation.* (A) Individual fluorescence spectra representing the titration of B-Myb 275-376 with molar equivalents of p300 1726-1839 are shown compiled. (B) A titration plot displaying a general shift in the wavelength maxima to a shorter wavelength upon increasing molar additions of p300 1726-1839.

The Fluorescence spectra compiled in figure 4.7A demonstrates a clear shift in the wavelength maxima observed for B-Myb 275-376 from 353 nm to 347 nm, upon titrating molar equivalents of p300 1726-1839. The 6 nm blue shift to a shorter wavelength indicates that one or both tryptophans in B-Myb 275-376 have moved to a less solvent exposed environment consistent with the formation of a folded complex with p300 1726-1839. The titration curve in figure 4.7B, derived from plotting the change in wavelength maxima at differing molar equivalents of titrated p300 1726-1839, indicates that complex association between B-Myb 275-376 and p300 1726-1839 assumes a 1:1 stoichiometry.

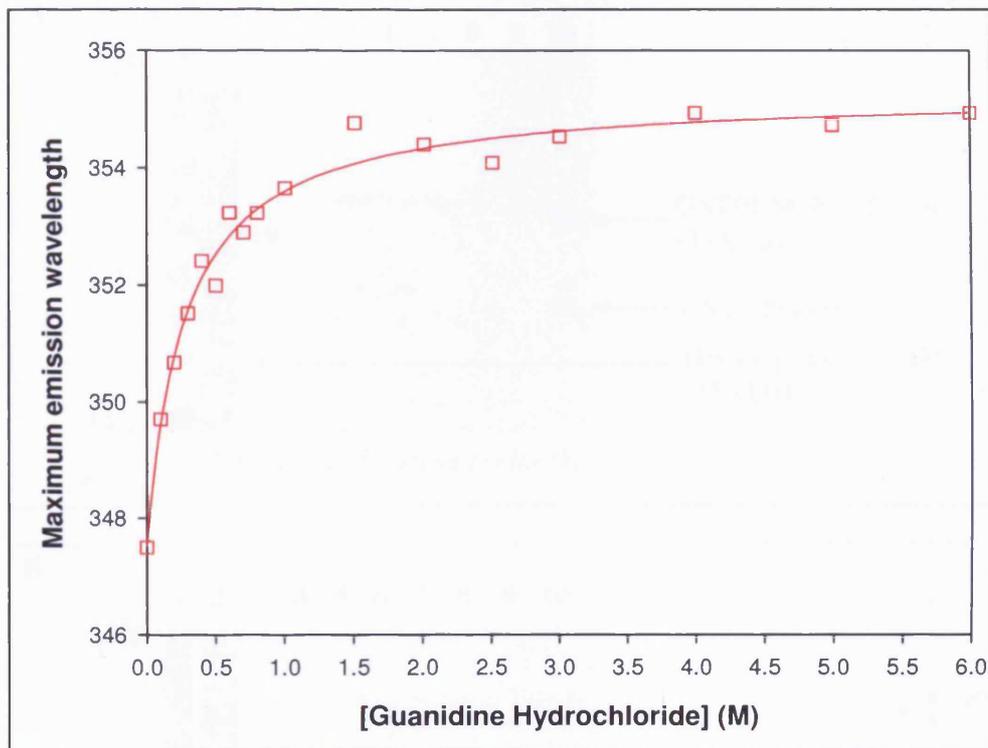
The dissociation constant ( $K_d$ ) for the observed interaction can be defined as the affinity between B-Myb 275-376 (P) and p300 1726-1839 (A) for the binding reaction equilibrium as:

$$K_d = \frac{[P][A]}{[P \bullet A]}$$

The nature of the binding observed for the fluorescence titration assay shown in figure 4.7 suggests that, at any given concentration of p300 1726-1839, the majority of B-Myb 275-376 is in complex with p300 1726-1839. Assuming the occupancy of B-Myb 275-376 (P) and p300 1726-1839 (A) association reaches 90 % at equilibrium, the calculated  $K_d$  for the interaction corresponds to  $\sim 3.0 \times 10^{-8}$  M.

#### *4.3.3.4 Structural characterisation of the B-Myb 275-376•p300 1726-1839 complex: Chemical denaturation of the B-Myb 275-376•p300 1726-1839 complex*

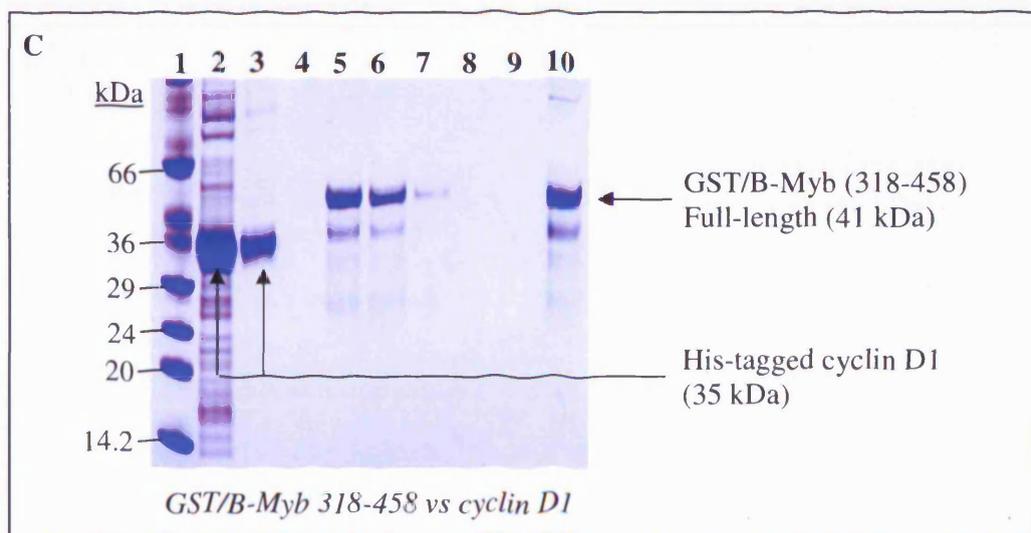
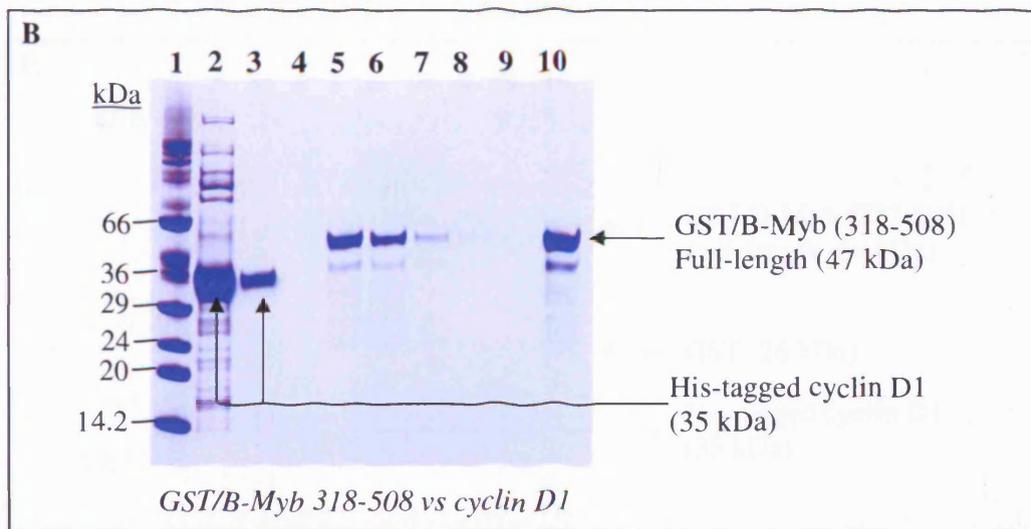
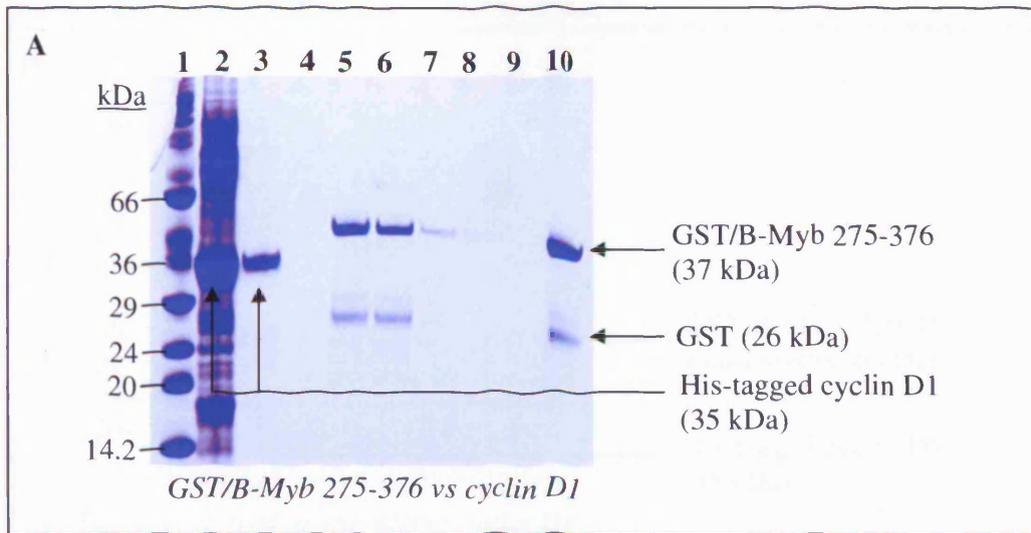
Intrinsic tryptophan fluorescence studies (discussed in section 4.3.3.3) were used to demonstrate the formation of a specific and high affinity complex between B-Myb 275-376 and p300 1726-1839. The structural stability of the complex was assessed by monitoring the maximum emission wavelength of the complex in response to increasing guanidine hydrochloride concentrations (shown in figure 4.8). The data clearly indicate that the B-Myb 275-376•p300 1726-1839 complex shows little resistance to denaturant exhibiting a non co-operative unfolding curve. The maximum emission wavelength observed for the complex is red shifted to a longer wavelength upon denaturant titration, with a 0.75 M Guanidine Hydrochloride approximate midpoint in unfolding. The fully denatured complex assumes a maximum emission wavelength of 355 nm, which is similar to that observed for B-Myb 275-376 in isolation (figure 3.2), indicating the polypeptide's tryptophan residues (W293 and W323) are solvent exposed in a random coil conformation. This type of behaviour suggests the complex rapidly dissociates and is unlikely to have assumed a stable tertiary fold. Typically, complex dissociation and unfolding occur together in a cooperative manner for a folded protein complex with defined tertiary structure.

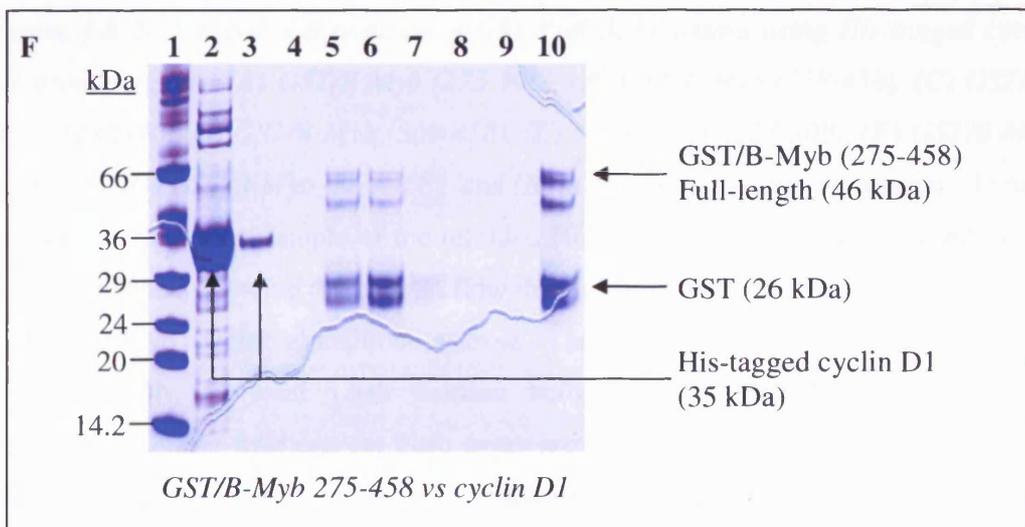
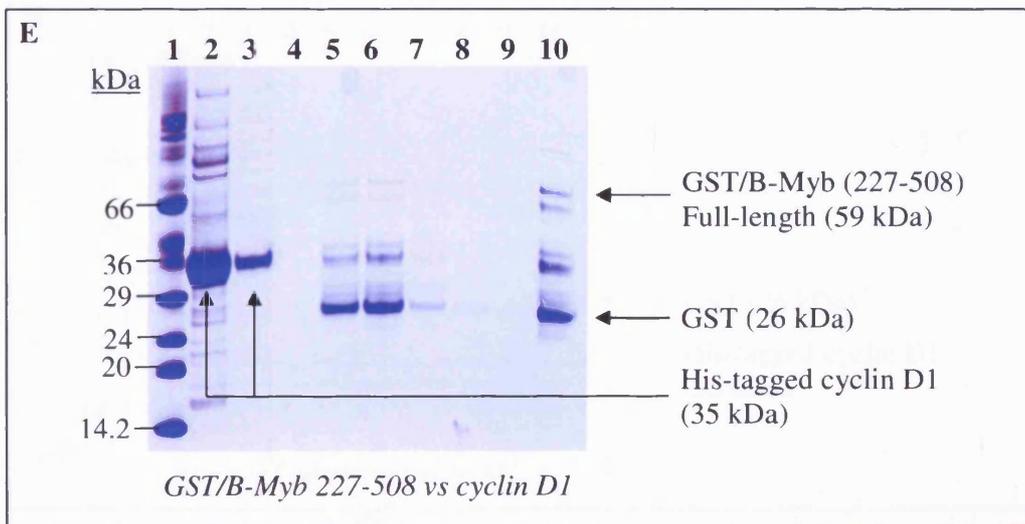
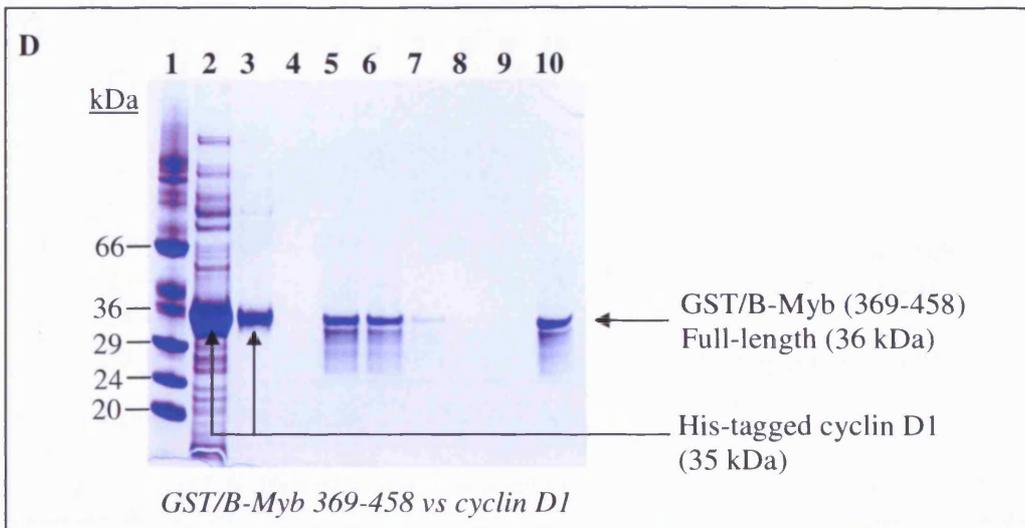


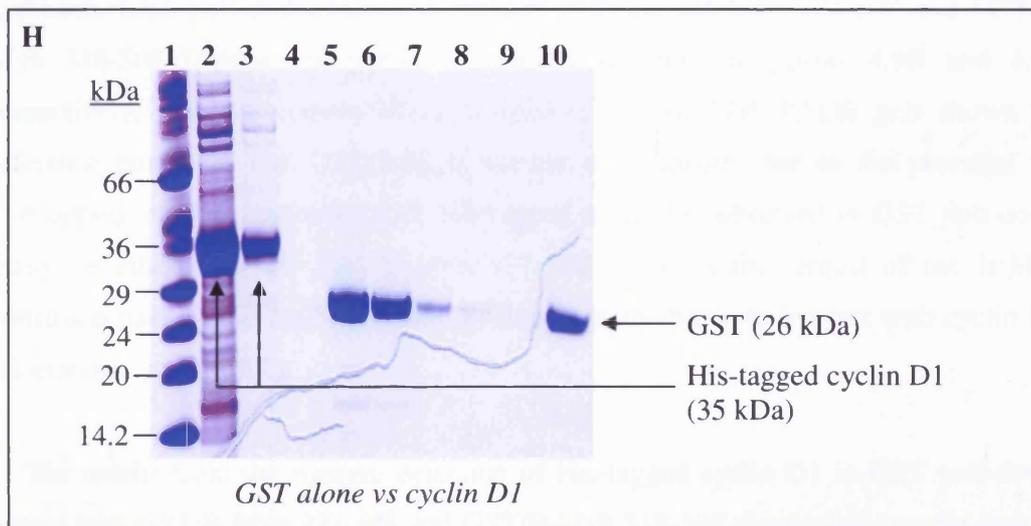
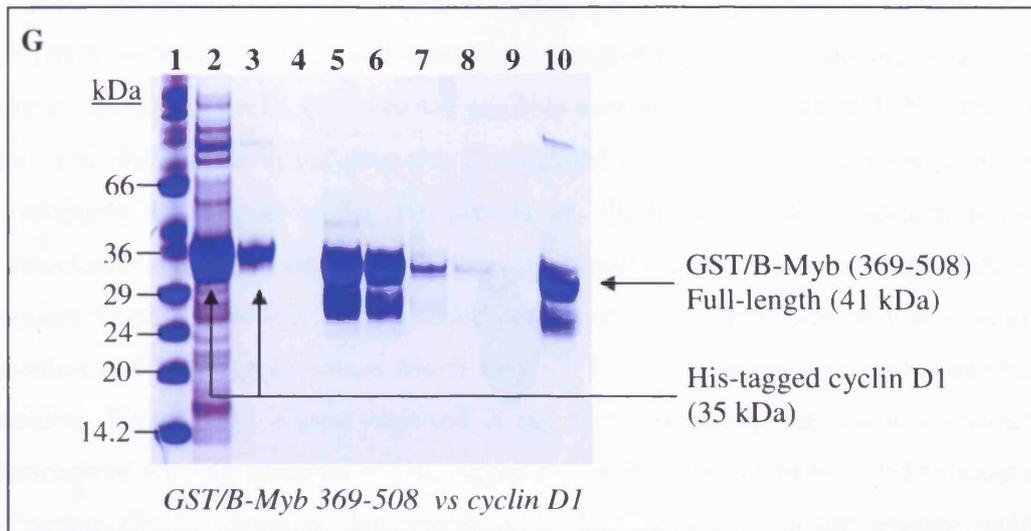
**Figure 4.8.** Guanidine hydrochloride induced denaturation curve for the 1:1 B-Myb 275-376•p300 1726-1839 complex ( $3 \mu\text{M}$ ). The stability of the protein complex was assessed by monitoring shifts in the maximum emission wavelength as a function of increasing denaturant concentration. The lack of resistance to even small amounts of denaturant suggests the complex lacks stable tertiary structure.

#### 4.3.4 Interactions between His-tagged cyclin D1 and B-Myb

SDS PAGE gels displaying the results of GST pull-down assays between His-tagged cyclin D1 and various GST/B-Myb fusions as bait protein bound to glutathione agarose are shown in figure 4.8. There is little evidence for an interaction with the selected GST/B-Myb fusion proteins used in this study. The absence of co-eluted His-tagged cyclin D1 observed in all of the assays, coupled with an abundance of protein in the column flow-through (lane 3 of each gel) is indicative of a lack of binding.



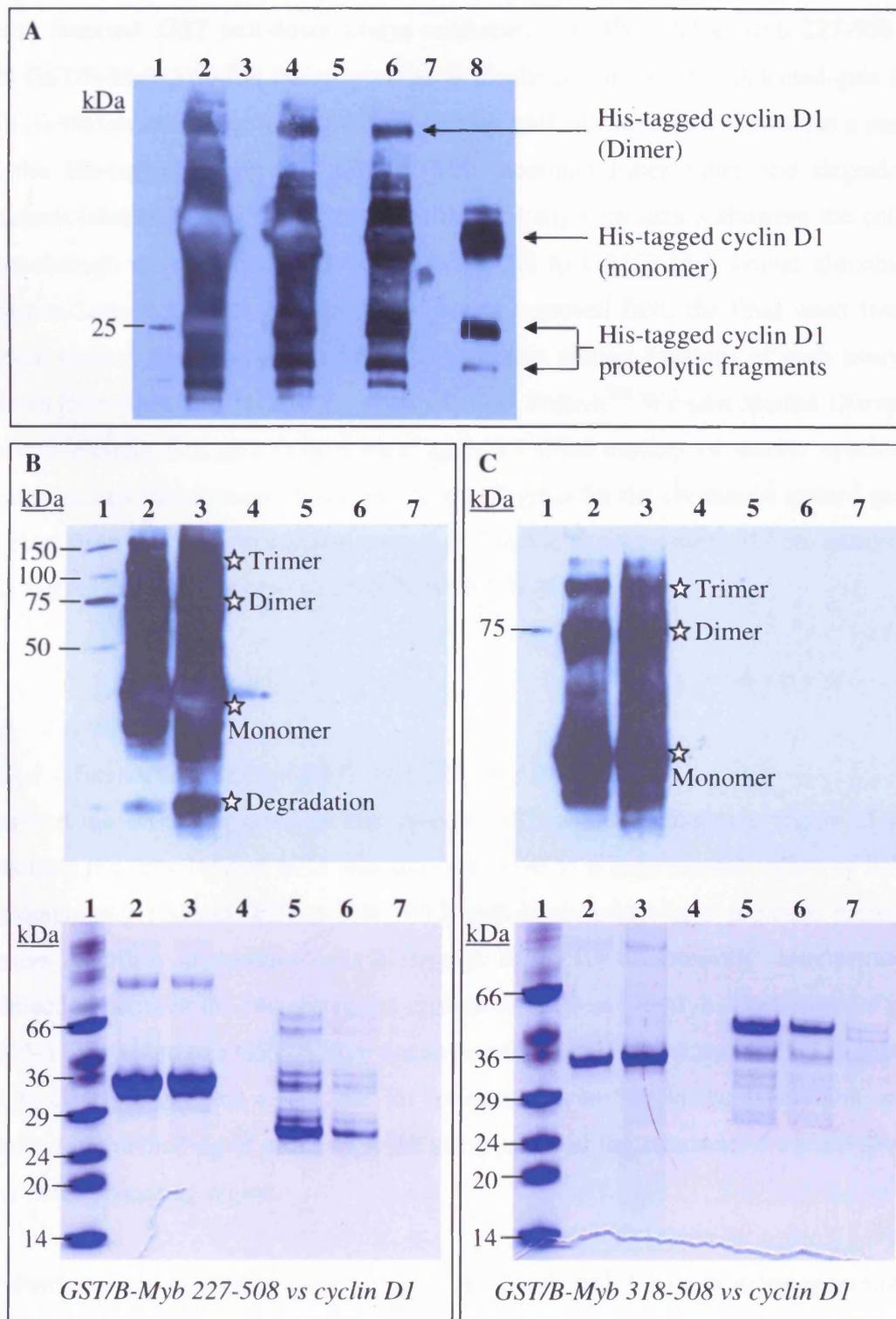




**Figure 4.8.** SDS PAGE gel analysis of GST-Pull down assays using His-tagged cyclin D1 protein against (A) GST/B-Myb (275-376), (B) GST/B-Myb (318-458), (C) GST/B-Myb (318-508), (D) GST/B-Myb (369-458), (E) GST/B-Myb (227-508), (F) GST/B-Myb (275-458), (G) GST/B-Myb (369-508), and (H) GST alone fusion bait proteins. Lane 2 in each gel contains a sample of the refolded His-tagged cyclin D1 protein used in the assay with lane 3 showing the column flow-through after applying His-tagged cyclin D1 to GST/B-Myb bound glutathione-agarose. Lane 4 in each gel contains a sample removed from the final wash fraction before elution. Samples removed from consecutive elution fractions for each assay are shown from lanes 5 to 9. Lane 10 in each gel contains a sample of the GST/B-Myb fusion protein used in the assay. Broad range molecular marker is shown in lane 1 for each gel.

From the coomassie stained gels in figure 4.8 it is apparent that several of the GST/B-Myb fusion proteins and their respective proteolytic fragments are of a similar size to His-tagged cyclin D1 (with the possible exception of GST/B-Myb 275-458) and thus may obscure the visual detection of co-eluted protein. In order to view potentially overlapping His-tagged cyclin D1 protein an alternative mode of detection using monoclonal antibodies directed towards the hexa-histidine tag was employed. This western blotting approach has the added benefit of being highly sensitive to nanogram quantities of His-tagged protein which may not be observed through coomassie based staining. Figure 4.9A shows exposed X-ray film displaying the chemi-luminescent detection of varying amounts of His-tagged cyclin D1, with as little as 500 nanograms of protein clearly observed, demonstrating the high sensitivity of the western blotting approach. GST pull-down assays conducted with GST/B-Myb 227-508 and GST/B-Myb 318-508 fusions are shown in the X-ray films in figures 4.9B and 4.9C respectively, with the corresponding coomassie stained SDS PAGE gels shown for reference purposes. The GST/B-Myb fusions were chosen due to the potential for overlapping elution fragments with His-tagged cyclin D1 observed in GST pull-down assays. Furthermore, the GST/B-Myb 227-508 fusion is the largest of the B-Myb constructs used in this study and had previously been shown to interact with cyclin D1 (Horstmann *et al.*, 2000).

The results from the western detection of His-tagged cyclin D1 in GST pull-down assays with GST/B-Myb 227-508 and GST/B-Myb 318-508 clearly indicate the protein does not interact with these regions under the assay conditions used. The presence of small quantities of dimer and trimer species as well as degradation/ proteolytic products of His-tagged cyclin D1 give rise to a largely unresolved appearance in lanes 2 and 3 of the western blot in figure 4.9 (by contrast the coomassie stained gel shows predominantly monomeric His-tagged cyclin D1 protein with no visible degradation products further exemplifying the greater detection sensitivity of the western blot). Furthermore the X-ray film had been deliberately over-exposed as to ensure the detection of any potential co-eluted protein. The small amount of His-tagged cyclin D1 observed in lane 4 of the GST/B-Myb 227-508 Western blot is most likely to be due to 'overspill' in the loading of lane 3 of the SDS PAGE gel.



**Figure 4.9.** Western detection of His-tagged cyclin D1 in GST/B-Myb pull-down assays. (A) The western detection of varying quantities of His-tagged cyclin D1: lane 2- 12  $\mu$ g; lane 4- 6  $\mu$ g; lane 6- 3  $\mu$ g; lane 8- 500 ng. Lanes 3, 5 and 7 are empty, with lane 1

containing Perfect Protein™ Western Marker (Novagen) which unfortunately was not easily detected. GST pull-down assays conducted with (B) GST/B-Myb 227-508 and (C) GST/B-Myb 318-508 fusion proteins is displayed for western detected gels (top) and coomassie stained gels (bottom). For all the gels shown, lane 2 represents a sample of the His-tagged cyclin D1 protein (with monomer/dimer/trimer and degradation products labelled with a white star) used in the assay with lane 3 showing the column flow-through after applying His-tagged cyclin D1 to GST/B-Myb bound glutathione-agarose. Lane 4 in each gel contains a sample removed from the final wash fraction before elution. Samples removed from consecutive elution fractions of each assay are shown from lanes 5 to 7. Lane 1 contains Perfect Protein™ Western Marker (Novagen) for the Western blot gels (where once again a limited number of marker species are detected), and Broad range molecular marker (Sigma) for the coomassie stained gels. It is clear from viewing the elution lanes 5 to 7 in the Western blots of both assays that His-tagged cyclin D1 fails to co-elute with GST/B-Myb protein.

#### 4.4 Discussion

##### 4.4.1 Interaction between p300 1603-1725 and B-Myb

From the results reported in this chapter a ZZ domain containing region of p300 spanning residues 1603 to 1725 was shown to bind to a large central region of B-Myb encompassing residues 227 to 508. GST pull-down assays and protease protection assays described in sections 4.3.1.2 through to 4.3.1.4 conclusively demonstrate an interaction between the two respective regions of p300 and B-Myb. The failure of p300 1603-1725 to bind to a GST/B-Myb fusion protein containing residues 275-376 from B-Myb, clearly highlights a key role for residues 376 to 508 in the interaction, either directly in the binding of p300, or in the stabilisation of the structure of a p300 binding site in the preceding region.

Further mapping studies using an array of GST/B-Myb fusion proteins spanning the 227 to 508 amino acid region proved inconclusive due to the irreproducible nature of some of the binding observed. In most instances the amount of co-eluted p300 1603-1725 (prey protein) observed in GST pull-down experiments was not significantly greater than the background levels observed in the final wash. Such high background levels of p300 1603-1725 in the final wash fractions may be explained by a potentially

high 'on-off' rate of binding to the GST/B-Myb fusion caused by a relatively weak interaction between the proteins. The protease protection assays provided a much less stringent but more sensitive method of detecting protein-protein interaction and was successful in mapping the interaction of p300 1603-1725 to a short 369-458 fragment in B-Myb. The greater sensitivity provided by this method is due to the potential for proteins to reach an equilibrium binding state in solution before limited proteolysis. In GST pull-down assays an equilibrium of binding is never achieved between the bait and prey protein. Thus, the segment containing residues 369-458, found within the Exon 9a region of B-Myb appears to provide a minimal interacting surface with the p300 1603-1725 polypeptide. The absence of clear binding observed in GST pull-down assays involving the p300 1603-1725 and B-Myb 369-458 polypeptide more than likely suggests the interaction is weak or transient in nature. Interestingly, this region of B-Myb contains a large proportion of Exon 9a (residues 318 to 458) which is absent in an alternatively spliced form of B-Myb that is inactive as a transcriptional activator (Kamano *et al.*, 1995). Horstmann *et al.* proposed that the Exon 9a region may contain a crucial element of the B-Myb transactivation domain (Horstmann *et al.*, 2000a). The finding that residues within this region are important for the binding of the transcriptional coactivator p300 provides a possible explanation for the importance of Exon 9a in B-Myb mediated gene activation.

The limited proteolysis results presented in figures 4.3A-G, together with identical data from control experiments carried out in the absence of GST/B-Myb protein (figure 4.3H), also reveal that p300 1603-1725 is very resistant to trypsin digestion compared to an unstructured polypeptide such as B-Myb 275-376, with cleavage occurring at just a single site to produce a 2 to 3 kDa smaller fragment (Figure 4.3H). N-terminal sequencing of the p300 1603-1725 cleavage product revealed that proteolysis was occurring towards the C-terminus of the protein, with the probable cleavage site therefore corresponding to either Lysine 1704 or Lys 1707. It is of interest to note that p300 1603-1725 contains five potential trypsin cleavage sites in the highly conserved N-terminal region preceding the ZZ motif (residues 1603-1661) and one at the boundary between the regions, which are clearly not susceptible to trypsin digestion. In addition the region preceding the ZZ motif is not suggestive of a disordered linker between independent domains as evidenced by its prediction to form part of a  $\beta$ -sheet (section 3.2.2.1, figure 3.3A).

#### 4.4.2 Interaction of p300 1603-1839 with B-Myb

Due to the poor refolding efficiency and solubility of the p300 1603-1839 polypeptide an extensive protein-protein mapping study with B-Myb was not possible. However, as described in section 4.3.2, a limited number of GST pull-down assay experiments using refolded His-tagged p300 1603-1839 were completed and clearly demonstrate an interaction with B-Myb residues 275-376. The same region of B-Myb shows no evidence of binding the p300 1603-1725 polypeptide suggesting that residues 1726 to 1839 (containing the TAZ2 domain) are important for this interaction. This is consistent with the previously reported interaction of a TAZ2 containing region of human p300 (spanning residues 1710-1891) with B-Myb (Schubert *et al.*, 2004) demonstrated by use of *in vitro* reporter gene assays and GST pull-down assays, and *in vivo* co-immunoprecipitation studies. Schubert *et al.* mapped the p300 1710-1891 interacting site to within residues 240-371 which spans a very similar region to the B-Myb 275-376 polypeptide used in this study.

#### 4.4.3 Interaction of p300 1726-1839 with B-Myb

The results presented in section 4.3.3 identify the central region of B-Myb as a key binding site for the p300 1726-1839 polypeptide with interactions observed to several overlapping B-Myb constructs in GST pull-down assays including regions 227-508, 275-376, 275-458, 318-458 and 318-508 with the exception of regions 369-458 and 369-508 (summarised in figure 4.5B). In particular the binding of p300 1726-1839 to B-Myb 275-376 appeared consistently stronger than with any other region of B-Myb suggesting that this may sufficiently represent the minimal interaction site. This finding supports previous mapping studies reported in this thesis which show a tight and specific interaction between B-Myb 275-376 and p300 1603-1839 (section 4.3); an interaction that is absent with the p300 1603-1725 polypeptide. The result further emphasises the importance of the TAZ2 domain region of p300 in mediating an interaction with B-Myb as first identified by Schubert *et al.*, 2004.

The fluorescence based binding assay reported in section 4.3.3.3 provides conclusive evidence for a highly specific complex formed between B-Myb 275-376 and p300 1726-1839. The complex clearly assumes a 1:1 stoichiometry of binding with B-Myb 275-376, a random coil polypeptide in isolation, exhibiting a tryptophan wavelength fluorescence shift from 353 nm to 347 nm upon addition of equimolar amounts of p300

1726-1839. This is consistent with the tryptophan residues in the B-Myb 275-376 polypeptide adopting a less polar environment as would be expected during complex formation. The calculated  $3.0 \times 10^{-8}$  M (30 nM)  $K_d$  of the complex suggests a high affinity association between B-Myb 275-376 and p300 1726-1839 comparable to that observed for the ACTR-SID (CBP) complex,  $3.4 \times 10^{-8}$  M (34 nM) (Demarest *et al.*, 2002) and the reported Hif-1 $\alpha$ -TAZ1 (CBP) complex,  $7 \times 10^{-9}$  M (7 nM) (Dames *et al.*, 2002).

The B-Myb 275-376•p300 1726-1839 complex, unexpectedly, showed little resistance to denaturant, exhibiting a rapid unfolding transition (section 4.3.3.4). The tryptophan residues in the B-Myb 275-376 polypeptide readily assume a solvent exposed environment (from 347 nm to 355 nm) suggesting the complex rapidly dissociates and is unlikely to have assumed a stable tertiary fold. Typically, complex dissociation and unfolding occur together in a cooperative manner for a folded protein complex with defined tertiary structure (Dyson and Wright, 2002).

This finding, together with the results attained from the structural characterisation of p300 1726-1839 has several implications for the nature of the interaction observed for the B-Myb 275-376•p300 1726-1839 complex. The inherent structural instability of the B-Myb 275-376•p300 1726-1839 complex may be a reflection of the unstable nature of the p300 1726-1839 polypeptide, deduced from thermal melting profiles reported in section 3.2.4.2. Although the characterised p300 1726-1839 polypeptide contained significant  $\alpha$ -helical secondary structure it lacked clearly defined tertiary structure, features synonymous with a molten globule protein state (Dobson, 1994; Ptitsyn, 1995). Therefore the observed complex instability may be a direct result of the unstable p300 1726-1839 component of the complex. It is likely then, that an interaction formed between a fully folded p300 1726-1839 polypeptide, which should provide a structured scaffold for protein binding (DeGuzman *et al.*, 2000), and B-Myb 275-376 will result in the formation of a complex with stable tertiary structure.

A further explanation for the structural behaviour of the B-Myb 275-376•p300 1726-1839 complex is that additional residues from either B-Myb or p300 may be required to achieve structural stability. Recent mapping studies, have identified residues 1710 to 1891 from human p300 and residues 240-371 from murine B-Myb as providing the

minimal site of a functional interaction between the two respective proteins (Schubert *et al.*, 2004). The additional B-Myb residues 240-274 found in the Schubert *et al.* construct (absent in the B-Myb 275-376 polypeptide used to characterise interactions with p300 1726-1839) is aspartic acid rich (30%) and exhibits very poor sequence homology amongst the vertebrates Myb family. A B-Myb construct spanning these residues was not produced (apart from the large B-Myb 227-508 construct) in the mapping study described in this thesis making conjectures of its importance in the interaction with p300 difficult to comment on. A cursory study of the additional regions present in the p300 construct used by Schubert *et al.* reveal a linker region between the ZZ domain and TAZ2 domain consisting of residues 1710 to 1725 and a proline /glutamine rich region from residues 1840 to 1891. The residues 1710-1725 formed the C-terminus of the p300 1603-1725 polypeptide and showed no evidence of binding to B-Myb 275-376 (section 4.3.1). The p300 region from residues 1840 to 1891 contains many prolines (15) and glutamines (9) which are commonly found within the transactivation regions of transcriptional regulatory proteins (Courey and Tjian, 1988). At present there is no evidence to suggest that this region functions as a transactivation domain or directly participates in protein-protein interactions. In fact many of the C/H3 binding proteins including E1A, E2F, Mi and p53 have been shown to interact with a short highly conserved peptide sequence (TRAM) found within the TAZ2 domain (Trouche *et al.*, 1996; Sato *et al.*, 1997; O'Connor *et al.*, 1999 DeGuzman *et al.*, 2000) further emphasising the importance of this domain in mediating protein-protein interactions.

#### 4.4.4 Interaction of cyclin D1 with B-Myb

From the results presented in section 4.3.2 there is no evidence for an interaction between cyclin D1 and a large central region of B-Myb spanning residues 227-508. The purified His-tagged cyclin D1 protein used in this study failed to bind an array of B-Myb proteins spanning the 227-508 region in GST pull-down assays assessed by both coomassie stained SDS PAGE gels and chemi-luminescent detection. This finding is in stark contrast to the numerous reports of a functional interaction between the two respective proteins (Horstmann *et al.*, 2000; Cesi *et al.*, 2002; Schubert *et al.*, 2004). Horstmann *et al.* demonstrated the proteins interact *in vivo* by co-immunoprecipitation assays and *in vitro* by GST pull-down assays. Specifically, GST/B-Myb 227-704 and GST/B-Myb 227-508 fusion proteins immobilised on sepharose beads were shown to

interact with *in vitro* translated cyclin D1 (Horstmann *et al.*, 2000). Furthermore cyclin D1 and p300 were shown to actively compete for binding with the central region of B-Myb deduced by *in vivo* coimmunoprecipitation studies and *in vitro* GST pull-down assays (Schubert *et al.*, 2004).

The GST/B-Myb 227-508 fusion construct used in the GST pull-down assays reported in this chapter was a kind gift from Karl-Heinz Klempnauer and was previously used in the Horstmann *et al.* 2000b study. The GST/B-Myb fusion protein was bacterially expressed using similar conditions to that reported by Horstmann *et al.*, resulting in the purification of soluble protein (discussed in section 2.3.5). The bacterially expressed and purified His-tagged cyclin D1 protein (purification discussed in section 2.3.11) used in the binding assays was shown to adopt a stable tertiary fold with considerable  $\alpha$ -helical secondary structure (discussed in section 3.3.4). Therefore there is no obvious explanation as to why the proteins failed to interact in the *in vitro* GST pull-down assays reported in this chapter. The N-terminal His-tag incorporates an additional 15 residues to the cyclin\_N domain of cyclin D1 which is reported to mediate interactions with several proteins including TAF<sub>II</sub>250, p21 and cdk4 (reviewed in Coqueret *et al* 2002). If the cyclin\_N domain represents the minimal interaction surface with B-Myb there is the possibility that the N-terminal His-tag may interfere with this association. Initial attempts at removing the thrombin cleavable His-tag from cyclin D1 however proved largely unsuccessful due to the non-specific digestion of the protein (data not shown).

#### 4.4.5 *The protein-protein interaction strategy*

The pull-down strategy reported in this thesis was based on the purification of an array of recombinant B-Myb, p300 and cyclin D1 polypeptides in *E. coli* for use in mapping the minimal sites of interaction between B-Myb and p300/cyclin D1. The findings of previous interaction studies involving B-Myb with p300 and cyclin D1 (Horstmann *et al.*, 2000) together with analysis of protein sequence conservation were used to aid the design of the constructs. The mapping study was based on the regions involved in interactions between B-Myb and partner proteins where the functional significance had already been demonstrated by *in vivo* functional and immunoprecipitation assays. The interactions detected in GST pull-down assays were therefore not used to imply functional significance. Using GST/B-Myb polypeptides as bait,

purified p300 and cyclin D1 polypeptides were systematically assayed for binding in GST pull-down assays; GST-pull-down assays are commonly used to probe interactions between a fusion protein and a known protein of interest (Kaelin *et al.*, 1991; Grgurvichet *et al.*, 1999; Hunter *et al.*, 1999; Posern *et al.*, 1999). The results of individual assays were assessed by SDS PAGE using coomassie, silver-stain and chemiluminescent detection of proteins. The minimal sites of interaction were thus delineated based on the differential binding affinities observed by SDS PAGE for the assayed polypeptides. The long-term aim of the mapping study was to characterize the structural properties and features of the identified interacting regions in isolation and as part of a complex.

The main advantage conferred by using an *in vitro* GST pull-down strategy with purified polypeptides is that the observed binding is unambiguously a result of a direct interaction between the two respective proteins rather than through an intermediate accessory protein. Protein complexes identified by other approaches such as co-immunoprecipitation or by using *in vitro* translated proteins, may contain intermediate or accessory binding proteins. Thus for a mapping interaction study between two proteins of interest, it is pertinent that no other potential binding partner be present in the assay. The GST pull-down approach is well suited to defining the specific domains involved in protein-protein interactions, i.e screening for the interacting regions, which can then be followed by more quantitative methods such as fluorescence, surface plasmon resonance (SPR, Biacore) and isothermal calorimetry (ITC), techniques that are not well suited to the mapping phase. Isothermal titration calorimetry can provide a complete thermodynamic profile of a protein-protein interaction including the kD, reaction stoichiometry, enthalpy and entropy. SPR is an optical resonance technique that provides a means for collecting kinetic data using only microgram amounts of protein. Such studies can be used to further validate the results attained from the GST pull-down assays and also allows for a more quantitative measure of protein interactions as described for the interaction between B-Myb 275-376 and p300 1726-1839 (see section 4.3.3.3) deduced by intrinsic tryptophan fluorescence.

The limitations of using *E. coli* expressed polypeptides for GST pull-down assays include the potential for low expression, synthesis of fusion proteins with degradation or proteolysis (as observed for several of the GST/B-Myb polypeptides) and insolubility

(as observed for the p300 polypeptides). The protocol described in this thesis for GST pull down assays produced results that are qualitative rather than quantitative. An alternative approach would be to use radiolabelled proteins in GST pull-down assays which allow for a quantitative measure of protein binding. For example, Matsuda *et al.* produced *in vitro* translated [<sup>35</sup>S] methionine-labeled protein for use in GST pull-down assays which were then analysed by a phosphorimager. Although the interaction affinities (kd's) were not reported, the differential binding affinities of several proteins in competition binding assays were quantified by this method (Matsuda *et al.*, 2004).

Further approaches to mapping protein-protein interactions include the use of *in vivo* co-immunoprecipitation assays which often serve to substantiate interactions observed *in vitro*. Co-immunoprecipitation is commonly used to demonstrate and identify physiologically relevant protein-protein interactions from cell extracts and can be used to test an interaction between two known proteins. At the start of this research project, collaborators in Germany (Karl-Heinz Klempnauer group) had demonstrated that a large central region of B-Myb (residues 227 to 508) co-immunoprecipitated as a complex with cyclin D1 (Horstmann *et al.*, 2001) and p300 (at the time unpublished observations, later published Schubert *et al.*, 2004) in QT6 cells. These initial *in vivo* studies served as the premise for the GST pull-down mapping studies reported in this thesis.

Yeast or bacterial two-hybrid assays can also be used to identify or map protein interactions between known proteins. This *in vivo* assay is based on the reconstitution of a functional transcriptional activator (often Gal4 in yeast) composed of a DNA-binding domain (DBD) and an activation domain each fused to the proteins of interest. Reporter genes containing binding sites for the DBD confer nutritional selection or  $\beta$ -galactosidase expression which allows for a quantitative measure of interaction affinities. A yeast-two hybrid study was initiated to map the interactions of B-Myb with p300 and cyclin D1 but unfortunately failed to yield any positive interaction results (data not reported in this thesis). The same amino acid regions used in the GST pull-down assay were produced as B-Myb fusions to Gal4 DBD and p300/cyclin D1 fusions to the Gal4 activation domain. The Gal4 DBD-B-Myb fusion protein was not detected in western blots of yeast extracts indicating a lack of expression, and providing a plausible explanation for the lack of positive interaction data.

## CHAPTER 5

### DISCUSSION AND CONCLUSIONS

#### 5.1 *Experimental considerations*

##### 5.1.1 *The choice of protein constructs for structural and functional studies*

At the time of experimental design, a large central region of B-Myb (residues 227-508) was shown to be essential for an inhibitory interaction with cyclin D1 (Horstmann *et al.*, 2000). Further unpublished studies by the same research group implicated the same region of B-Myb with an association with the p300 co-activator. The 227 to 508 amino acid region lacks both the N-terminal DNA binding domain, and a C-terminal region associated with negative regulation of B-Myb activity; c-terminal truncations of B-Myb up to amino acid 508 greatly enhance B-Myb transcriptional activity (Ziebold *et al.*, 1997). Although there is little structural information for the 227 to 508 residue region of B-Myb, several groups have identified a functional region responsible for transcriptional activity that spans the residues 272 to 365 (Ansieau *et al.*, 1997), 206 to 372 (Nakagoshi *et al.*, 1993), and 240 to 371 (Horstmann *et al.*, 2000). In order to further map and localize the region of B-Myb responsible for mediating the reported interaction with cyclin D1 and p300, a series of B-Myb constructs were designed that spanned the 227 to 508 amino acid region of B-Myb. The primary aim of such a study was to structurally characterize the properties of the isolated proteins or protein complexes identified in the mapping study. Thus a major consideration when designing B-Myb constructs was to select amino acid boundaries that would potentially give rise to autonomously folded, independent functional domains or regions that would fold upon binding its co-factor.

A commonly used approach in deducing potential domain boundaries within a protein includes a limited proteolysis study of the protein. Such a study would identify protease sensitive unstructured regions from those regions that are structured and resistant to proteolysis. The 227 to 508 amino acid region of B-Myb, when expressed in *E. coli* as a GST-fusion was notable for its susceptibility to proteolysis (Shown in figure 2.13, gel 8). Trypsin mediated limited digestion of purified GST/B-Myb 227-508 leads to the rapid degradation of the protein and does not give rise to any stable proteolytic fragments (Data not shown) and thus proved largely unsuccessful in determining potential domain boundaries.

Multiple sequence alignments of Myb proteins in the region of interest were used to aid the design of B-Myb constructs. Comparison of Myb amino acid sequences from different vertebrates such as mouse and human allowed the identification of several break-points of sequence conservation including amino acids 275 and 369. It was therefore reasonable to assume that the amino acids 275 and 369 may represent the boundary points of a functional domain. The same two break-points in sequence conservation were previously identified by an alignment compiled by Horstmann *et al.*, and were used in that study to produce the Gal4/B-Myb fusions D2 (spanning residues 275-354) and D3 (spanning residues 369-502) (Horstmann *et al.*, 2000). Further construct boundaries included residues 318 and 458 that, although did not coincide with break-points in sequence homology, were chosen because they span the exon 9a region of B-Myb which is absent in an alternatively spliced, transcriptionally inactive B-Myb isoform (Horstmann *et al.*, 2000).

At the time of p300 construct design the cysteine/histidine rich region 3 (C/H3) of p300 was implicated with binding B-Myb (Unpublished observations, Karl-Heinz Klempnauer). The C/H3 region is a well documented interaction site for a multitude of diverse transcription factors including E1A (Eckner *et al.*, 1994) and E2F (Trouche *et al.*, 1996). This region is notable for containing two distinct zinc-binding motifs called the ZZ and TAZ2 domains. Therefore, in an attempt to define the interaction site with B-Myb further, three p300 constructs were designed spanning the ZZ motif (residues 1603-1725), the TAZ2 domain (residues 1726-1839), and a construct spanning both the ZZ and TAZ2 regions (1603-1839). The amino acid boundaries for the TAZ2 domain in p300 were already known from the published structure and were thus used to aid construct design. Analysis of break-points in sequence conservation using a multiple sequence alignment of p300 and CBP proteins (shown in figure 2.6) were also used to aid construct design, particularly for the ZZ domain due to the absence of a structure for that motif at the time.

### 5.1.2 *The functional significance of the chosen protein constructs*

Functional units or domains of proteins are often characterized by their innate capacity to form autonomously folded structures. Such functional domains are either folded independently or fold upon binding a co-factor. It is clear from the biophysical studies conducted with B-Myb 275-376 that the polypeptide adopts an unfolded, random coil conformation. Upon interaction with its p300 binding partner (p300 1726-1839) the B-Myb 275-376 polypeptide adopts a more structured conformation, consistent with a 'folding upon binding' model as observed with other transcription factor complexes (Radhakrishnan *et al.*, 1997; Zor *et al.*, 2002). It is likely therefore that the B-Myb 275-376 polypeptide is an intact functional unit due to its specific and high affinity ( $K_d \sim 30$  nM) association with p300.

The other B-Myb constructs include amino acid regions 227 to 508, 275 to 458, 318 to 458, 369 to 458 and 369 to 508. Most of the aforementioned polypeptides were characterized by the high levels of proteolysis observed upon purification and were thus not amenable to the biophysical studies afforded with the B-Myb 275-376 polypeptide. Therefore it is difficult to assess directly whether these portions are likely to be intact functional units capable of folding. The high levels of proteolysis would intuitively suggest that the regions are likely to be unstructured, random coil regions. However this does not exclude the possibility that these portions are indeed functional units, given the right context (i.e. potential binding partners). The B-Myb 318-458 (exon 9a) polypeptide, which showed evidence of binding p300 1603-1725, has a well documented positive functional role in B-Myb transcriptional activity (Horstmann *et al.*, 2000) and so is highly likely to be an intact functional unit. The p300 constructs used in this study contain well documented and characterized zinc binding domains (ZZ and TAZ2) that form independent structured units (DeGuzman *et al.*, 2001; Legge *et al.*, 2004). The TAZ2 domain forms a structured scaffold for binding many transcription factors including E1A (Eckner *et al.*, 1994) and p53 (Avantiatti *et al.*, 1997). Although a function has yet to be assigned for the ZZ domain it is thought to be involved in protein-protein interactions (Legge *et al.*, 2004).

One may test the functionality of the B-Myb constructs used in this study by performing *in vitro* reporter gene assays using Myb responsive promoters. Such studies have previously been conducted to localize the region of B-Myb responsible for

conferring transcriptional activity (Ansieau *et al.*, 1997; Nakogoshi *et al.*, 1993; Horstmann *et al.*, 2000). The functional relevance of the exon 9A region (residues 318 to 458) was addressed by comparing the ability of full-length B-Myb and B-Myb lacking the exon 9A to activate the Myb-dependent reporter gene pG13-3xATk-Luc (contains three copies of a Myb-binding site fused to the thymidine kinase promoter) in mammalian cell lines such as Cos7 and QT6 (Horstmann *et al.*, 2000). The full-length B-Myb strongly stimulated the reporter gene, while the shorter form of B-Myb had little effect. Furthermore it was found that B-Myb lacking exon 9A was an inhibitor of B-Myb mediated transactivation by competing with full-length B-Myb for the same binding site. Further functional regions of B-Myb have been identified by analyzing transcriptional activity of B-Myb deletion mutants on the Myb responsive reporter pA10CAT6MBS-1 in which the SV40 early promoter and six copies of a Myb-binding site are linked to the bacterial CAT gene (Nakagoshi *et al.*, 1993). Ansieau *et al.*, mapped a functional B-Myb transactivation region to within residues 272 to 365 after measuring the activity of a series of Gal4 DNA binding domain fused B-Myb deletion mutants on a Gal4 reporter gene (Ansieau *et al.*, 1997).

There are several methods to test the physiological or functional relevance of the identified *in vitro* interactions between B-Myb and p300 reported in this thesis. Indeed many of the reported functional *in vivo* interactions of B-Myb with p300/CBP involve similar regions to that identified in this thesis. To show that p300 acts as a co-activator for B-Myb, Schubert *et al.*, co-transfected a pG13-3xATk-Luc reporter gene with different combinations of expression vectors for full-length and partially deleted B-Myb, p300 and cyclin D1 proteins (Schubert *et al.*, 2004). They showed that B-Myb dependent transactivation is stimulated by p300 on these promoters and is abrogated by transfection with cyclin D1, suggesting that cyclin D1 inhibits the cooperativity of B-Myb with p300. Similar studies showed that B-Myb and p300 synergise to activate transcription on FGF-4 promoter/reporter constructs and that co-transfection with E1A inhibited this cooperativity (Johnson *et al.*, 2002). Another commonly used technique to study *in vivo* interactions between domains of proteins is *in vivo* coimmunoprecipitation. This technique has been widely used to study the interaction of endogenous levels of protein and of various deletion mutants of B-Myb with p300 and cyclin D1, with each study identifying similar interacting regions of B-Myb (Horstmann *et al.*, 2000; Schubert *et al.*, 2004; Li and McDonnell, 2002; Johnson *et al.*, 2002).

Finally, a yeast two hybrid assay was used to demonstrate that B-Myb, specifically a region spanning residues 197 to 340, was able to interact with CBP (Li and McDonnell, 2002).

In order to test whether the identified binding regions involving B-Myb and p300 were not only sufficient but necessary for interactions *in vivo*, competition binding assays can be employed. For example Schubert *et al.*, concluded that the C/H3 region of p300 was necessary for an *in vivo* association with B-Myb since the levels of co-immunoprecipitated B-Myb /p300 complex were severely reduced in the presence of cyclin D1 (Schubert *et al.*, 2004). Such studies can be corroborated with similar competition assays using the described *in vitro* reporter gene assays. For example, a dominant negative p300 mutant would be expected to diminish the observed synergistic activation of a reporter gene by wild-type B-Myb and p300. An alternative approach would be to introduce point-mutations to the regions thought to be responsible in the interaction of B-Myb with p300 and then to repeat the co-immunoprecipitation study. This way the functional importance of individual residues within the complex can be assessed. However care must be taken to ensure that the introduced mutations do not affect the overall structure of the protein regions involved, since a reduction in binding may be a result of a change in local or global domain structure.

## 5.2 *The B-Myb transactivation region: rationales for intrinsic disorder*

The B-Myb 275-376 polypeptide structurally characterised in this thesis contains a considerable proportion of the transactivation region as mapped in several independent studies (Nakogoshi *et al.*, 1993; Horstmann *et al.*, 2000; Ansieau *et al.*, 1997). The polypeptide was conclusively shown to adopt an unfolded, random coil conformation (discussed in section 3.2.1.2).

Disordered proteins appear to be very common in nature particularly in higher eukaryotes (Dunker *et al.*, 2002; Uversky, 2002) with an increasing number of intrinsically disordered domains found within proteins involved in cellular signalling, transcription and translation (Wright *et al.*, 1999; Tompa, 2002; (Namba, 2001). The transcriptional activation domains of many transcription factors are either unstructured or partly structured including the kinase-inducible activation domain (KID) of CREB (cAMP-response element binding protein) (Radhakrishnan *et al.*, 1997), the transactivation domain of p53 (Kussie *et al.*, 1996), and the acidic activation domain of herpes simplex virus VP16 (Uesugi *et al.*, 1997). They are characterised by amino acid compositional bias with low sequence complexity (Uversky *et al.*, 2000) similar to that shown for the B-Myb transactivation region (discussed in section 3.3.1).

The intrinsic lack of structure shown for the B-Myb transactivation region (residues 275-376) may confer several thermodynamic and functional advantages including the ability to potentially bind in different conformations to several different targets. It also allows for a mechanism of inducibility through post-translational modifications; the phosphorylation and acetylation of B-Myb towards its C-terminus results in a substantial increase in its transcriptional activity (Bessa *et al.*, 2001; Schubert *et al.*, 2004). A further proposed rationale for the abundance of unstructured transactivation regions in nature involves the greater 'capture radius' provided for binding co-factors, compared to compact folded proteins (Shoemaker *et al.*, 2000). Shoemaker *et al.*, report that a relatively unstructured protein molecule can have a greater 'capture radius' for a specific binding site as compared to the folded state with its dynamically restricted conformation. In a binding mechanism termed 'fly-casting' the unfolded protein binds weakly at a relatively large distance with simultaneous folding and binding ensuing rapidly (Shoemaker *et al.*, 2000). Coupled folding and binding has an energy penalty as the enthalpic contribution of hydrogen bond and salt bridge formation is countered by

the entropic cost associated with the process (Wright and Dyson, 1999). Such a mechanism provides a means of complex formation with reduced affinity whilst retaining high specificity.

The targeted degradation of proteins containing intrinsically unstructured regions has been suggested as a key regulatory mechanism in transcriptional control (Elledge and Harper, 1998; Desterro *et al.*, 2000; Salghetti *et al.*, 2001). The relative instability of unstructured transactivation regions allows for their targeted degradation by the ubiquitin-proteasome system, providing a further mechanism of switching 'off' transcriptional processes. Indeed B-Myb is actively targeted for ubiquitination and degradation (Charrasse *et al.*, 2000). B-Myb has a protein half-life of 2.7 hours which is significantly reduced to 50 minutes upon phosphorylation by the cyclin A/ cdk2 complex which targets the protein for ubiquitination. This mechanism ensures that B-Myb is only transiently hyperactivated in S-phase before being directed to the proteasome.

Based on the structural properties of the B-Myb 275-376 polypeptide in isolation it is more than likely that this region would adopt a more structured conformation upon binding a co-factor. There are several examples of transcriptional activation domains that are either unstructured or partially structured that co-fold upon interaction with target proteins. These include CREB (cAMP-response element binding protein) and the c-Myb transactivation domain that both fold upon binding to the KIX domain of CBP (CREB-binding protein) (Radhakrishnan *et al.*, 1997; Zor *et al.*, 2002), the transactivation domain of p53 binding to the cellular oncoprotein MDM2 (Kussie *et al.*, 1996) and the activation domain of VP16 binding to human TAF<sub>II</sub>31 (Uesugi *et al.*, 1997). It is of interest to note that high resolution structural studies have shown that the p53 and VP16 activation domains are unstructured in the context of the full-length proteins (Ayed *et al.*, 2001; Grossmann *et al.*, 2001) emphasising the physiological importance of coupled folding and binding in transcriptional activation.

### 5.3 *The B-Myb transactivation region: comparisons with c-Myb*

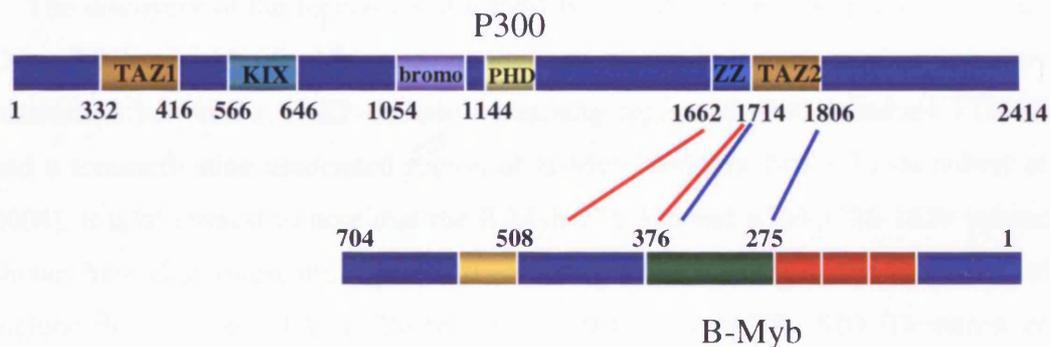
In contrast to B-Myb 275-376, the c-Myb transactivation domain (residues 291 to 315) in isolation assumes a partially helical conformation estimated at 25-30% helical content (Zor *et al.*, 2004), compared to the 9% reported here for the B-Myb 275-376 polypeptide (section 3.2.1.2). A comparison of far UV CD spectra attained for each shows a small 222 nm negative peak for c-Myb (Parker *et al.*, 1999) indicative of helical content which is absent from the B-Myb 275-376 spectrum (section 3.2.1.2). The binding of the KIX domain from CBP to the c-Myb peptide results in the stabilisation and extension of a single  $\alpha$ -helix from residues 293 to 309, with the regions outside the helix boundaries, residues 291-292 and 311-315 remaining unstructured (Zor *et al.*, 2004). Although biophysical studies indicate the absence of structure for the B-Myb 275-376 polypeptide, JPred2 secondary structure predictions suggest the potential for a short  $\alpha$ -helix spanning the amino acids VVEAANL (residues 293 to 299) in murine B-Myb (described in section 3.2.1.1). Despite the lack of significant sequence homology a cursory study of conserved residues between the helix forming region of c-Myb and the predicted helical region of B-Myb highlights the spacing of critical hydrophobic residues in both (figure 5.1A). Helical wheel analysis of the B-Myb 293 to 301 residue region (VVEAANLLI) indicates that this peptide has the potential to form an amphipathic helix (figure 5.1B). Interestingly the three functionally important hydrophobic residues (Ile 295, Leu 298, and Leu 302) in c-Myb that interact with the hydrophobic groove in the KIX domain (Parker *et al.*, 1999; Zor *et al.*, 2004) are similarly clustered on one face of the helix in B-Myb (Val 293, Ala 296, and Leu 300). There are some differences in the amino acid composition of the helices; the putative B-Myb amphipathic helix presents more of an extensive hydrophobic face and lacks a charged residue in position 5 (as denoted in the helical wheel) present in c-Myb (E299). Several mapping studies have excluded the KIX domain of CBP/ p300 as an interaction site for the B-Myb transactivation region (Johnson *et al.*, 2002; Schubert *et al.*, 2004).



blue denote hydrophobic amino acids, red denote charged residues with green representing polar amino acids.

#### 5.4 The B-Myb-p300 interaction

The work described in this thesis builds on previous mapping studies for the two proteins and leads to the conclusion that regions of p300 encompassing both the ZZ domain and the TAZ2 domain contribute to the interaction with B-Myb. A central region of B-Myb containing part of exon 9a (residues 369 to 458) appears to specifically bind to a ZZ domain containing polypeptide (p300 1603-1725), with the adjacent transactivation associated region of B-Myb (residues 275-376) specifically binding a TAZ2 containing polypeptide (discussed in chapter 4). On the basis of the GST pull-down assay results and protease protection assays employed, the B-Myb interaction involving the ZZ domain region of p300 was substantially weaker than that observed for the TAZ2 domain of p300. This suggests that the principal site of interaction exists between the transactivation associated region of B-Myb and the p300 TAZ2 domain, with the adjacent exon 9a region forming a stabilising interaction with the ZZ domain region of p300. The relative orientation of the B-Myb-p300 interaction based on the results obtained is depicted in figure 5.2.



**Figure 5.2** Schematic representation of the interacting regions of B-Myb and p300 based on the binding data reported in this thesis. The interacting regions between the blue lines depict the association between B-Myb 275-376 and p300 1726-1839, with the regions between the red lines showing the association between B-Myb 369-458 and p300 1603-1725.

### **5.5 The ZZ domain: A potential protein interaction module**

The binding data described here strongly implicate a region containing the ZZ domain as an interacting site with B-Myb. The ZZ domain is a recently discovered zinc-binding domain (Ponting *et al.*, 1996), with a proposed function as a protein interaction module (Legge *et al.*, 2004). At present there are no reported interaction partners for the CBP/p300 ZZ domain; the C/H3 binding proteins E1A and p53 have been shown unequivocally to interact with the TAZ2 domain, not the ZZ domain (DeGuzman *et al.*, 2000; Legge *et al.*, 2004). However, a ZZ domain apparent in dystrophin and dystrophin related proteins such as utrophin, forms part of a cysteine-rich region necessary to mediate interactions with the transmembrane glycoprotein,  $\beta$ -dystroglycan (Suzuki *et al.*, 1994). Deletion of the region containing the ZZ motif in utrophin diminishes its interaction with  $\beta$ -dystroglycan (Tommasi di Vignano *et al.*, 2000) and an analogous deletion in dystrophin leads to a slight reduction in its binding to  $\beta$ -dystroglycan (Rentschler *et al.*, 1999). The isolated ZZ region from both dystrophin and utrophin shows binding to calmodulin, but a functional role for this interaction has not been ascribed (Anderson *et al.*, 1996; Tommasi di Vignano *et al.*, 2000). The ZZ domain of p62 mediates direct interaction with the death domain kinase RIP (Sanz *et al.*, 2000) further exemplifying its role as a protein interaction module.

### **5.6 The B-Myb 275-376•p300 1726-1839 complex: a high affinity association**

The discovery of the formation of a tight B-Myb 275-376•p300 1726-1839 complex ( $K_d \sim 3.0 \times 10^{-8}$  M, 30 nM) supports previous reported mapping studies that show an interaction between a TAZ2 domain containing region of p300 (residues 1710-1891) and a transactivation associated region of B-Myb (residues 240-371) (Schubert *et al.*, 2004). It is of interest to note that the B-Myb 275-376 and p300 1726-1839 interaction shown here is amongst the higher affinity complexes reported in the literature, which include the Hif-1  $\alpha$  - TAZ1 (Dames *et al.*, 2000) and ACTR- SID (Demarest *et al.*, 2002) complexes, all with  $K_d$ 's in the nanomolar range. There are several examples of known interaction affinities for transactivation domains in complex with their target binding proteins, which are summarised in table 5.1.

<b>Transactivation domain (TAD)</b>	<b>Interacting domain</b>	<b>K<sub>d</sub> Affinity</b>	<b>Reference</b>
KID (101-160) (CREB) (Unphosphorylated)	KIX 586-672 (CBP)	108 X 10 <sup>-6</sup> M (108 μM)	Zor <i>et al.</i> , 2002
pKID (101-160) (CREB) (Phosphorylated)	KIX 586-672 (CBP)	0.7 X 10 <sup>-6</sup> M (0.7 μM)	Zor <i>et al.</i> , 2002
c-Myb (291-315)	KIX 586-672 (CBP)	15 X 10 <sup>-6</sup> M (15 μM)	Zor <i>et al.</i> , 2002
p53 (14-28)	TAZ2 1764-1850 (CBP)	300 X 10 <sup>-6</sup> M (300 μM)	De Guzman <i>et al.</i> , 2000
p53 (15-29)	MDM2 17-125	0.6 X 10 <sup>-6</sup> M (0.6 μM)	Kussie <i>et al.</i> , 1996
Hif-1 α (776-826)	TAZ1 354-439 (CBP)	7 X 10 <sup>-9</sup> M (7 nM)	Dames <i>et al.</i> , 2000
ACTR (1018-1088)	SID 2059-2117 (CBP)	34 X 10 <sup>-9</sup> M (34 nM)	Demarest <i>et al.</i> 2002
<i>B-Myb (275-376)</i>	<i>TAZ2 1726-1839 (p300)</i>	<i>3.0 X 10<sup>-8</sup> M (30 nM)</i>	<i>Reported in this thesis (Unpublished)</i>

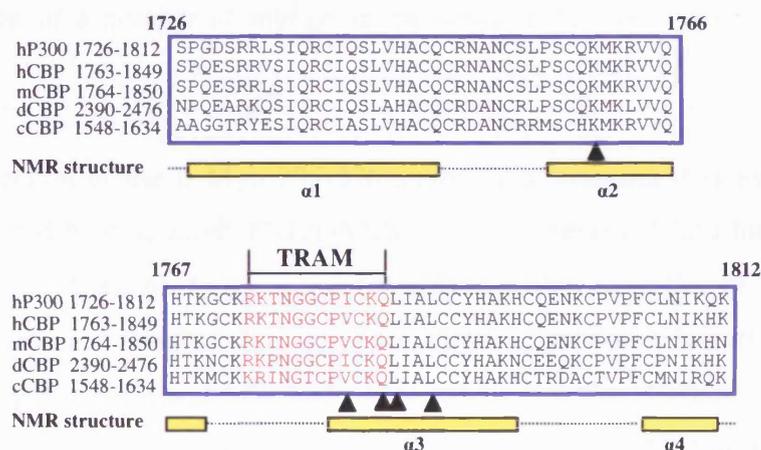
**Table 5.1** A summary of the known interaction affinities of transcriptional activation domains in complex with their target co-factor domains. The specific residues required for complex formation and binding K<sub>d</sub>'s are indicated.

It is noticeable that most of the reported interactions involving transcriptional activation domains involve domains from the CBP/p300 co-activator. The KIX, TAZ1 and TAZ2 domains all provide a structured scaffold for protein binding, with the SID domain representing an intrinsically disordered region that co-folds upon binding. The

highly specific and high affinity interaction deduced for the B-Myb 275-376•p300 1726-1839 complex is in stark contrast to the low affinity binding measured for the complex formed between the c-Myb transactivation domain and the KIX domain from CBP ( $K_d \sim 15 \times 10^{-6}$ , 15  $\mu\text{M}$ ) (Zor *et al.*, 2002). The different interaction affinities observed for the B-Myb and c-Myb transactivation regions may reflect diversity in the structural composition of these regions. In particular the B-Myb 275-376 peptide may confer a more extensive binding surface (when in complex with p300 1726-1839) than that for the comparatively shorter c-Myb peptide (residues 291-315) in complex with the KIX domain (Zor *et al.*, 2002; Zor *et al.*, 2004).

### 5.7 Comparison of B-Myb with other TAZ2 binding proteins

The C/H3 region of p300/CBP has been implicated with the binding of a multitude of diverse transcription factors including E1A (Eckner *et al.*, 1994), p53 (Avantaggiati *et al.*, 1997), E2F (Trouche *et al.*, 1996), the Microphthalmia transcription factor (Mi) (Sato *et al.*, 1997), c-Fos (Bannister and Kouzarides, 1995), YY1 (Lee *et al.*, 1995) and recently B-Myb (Schubert *et al.*, 2004). In many cases the TAZ2 domain is responsible for the interaction through a recently identified Transcriptional Adaptor motif (TRAM) (O'Conner *et al.*, 1999). The TRAM motif is composed of a 12 amino acid sequence (RKTNGGCPVCKQ in human CBP) conserved in all members of the p300/CBP family of proteins and spans part of the  $\alpha 3$  helix within the TAZ2 domain (shown in figure 5.2).



**Figure 5.2.** A multiple sequence alignment of the 87 amino acid TAZ2 domain from CBP and p300 homologues, with  $\alpha$ -helical secondary structural elements highlighted in yellow and the conserved TRAM motif residues shaded in red. The black triangles

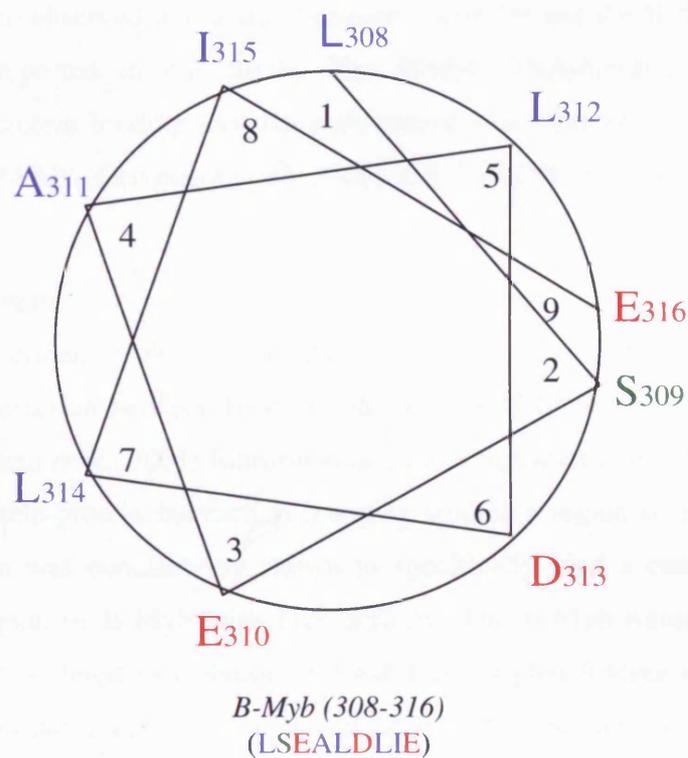
highlight residues (K1798, V1819 and Q1822, L1823 and L1826 in mouse CBP) within the TAZ2 domain that showed the largest chemical shifts upon titration with the p53 transactivation domain (DeGuzman *et al.*, 2000).

The importance of the TRAM was first recognised when studying its interaction with the N-terminal region of E1A, where a series of CBP deletion mutants abrogating the conserved motif diminished its binding to E1A. Furthermore, the p53, E2F1, and TFIIB transcription factors were also shown to bind the TRAM motif in an interaction that could be competitively displaced by an E1A peptide (O'Conner *et al.*, 1999). Residues within the TRAM sequence of TAZ2 were shown to be greatly affected by the addition of a p53 peptide when monitoring NMR chemical shifts (DeGuzman *et al.*, 2000). The study revealed that V1819 and Q1822 were among the residues from mouse CBP that were substantially perturbed upon association with p53 (shown in figure 5.2).

The minimal region of E1A, p53 and E2F1 required for the interaction with the TAZ2 domain was mapped to amino acids that reside within a conserved FXE/DXXXL sequence, which when mutated in each protein, resulted in abolished binding to CBP (Hagemeier *et al.*, 1993; Wang *et al.*, 1995; Gu *et al.*, 1997). Similar sequences to those in E1A, p53 and E2F have been shown to be necessary for the binding of MyoD (Sartorelli *et al.*, 1997) and YY1 (O'Connor *et al.*, Unpublished observations, 1998) by CBP. In MyoD, the sequence is known as the FYD motif and is present in the N-terminal region of a number of myogenic transcription factors also (Nakajima *et al.*, 1997).

Upon inspection of the B-Myb 275-376 amino acid sequence it is evident that the conserved TRAM binding motif FXE/DXXXL found in several TAZ2 binding proteins is absent. The Ets-1 and Mi transcription factors are further examples of TAZ2 binding proteins that lack a conserved FXE/DXXXL motif. A common feature found in the activation domains of all the TAZ2 binding proteins so far characterised is the potential to form amphipathic helices. The activation domain from the Mi transcription factor shares significant sequence homology with the activation domains of E1A and E2f which are all predicted to form amphipathic helices (Sato *et al.*, 1997; Trouche *et al.*, 1996). A short segment within the activation domain of p53, required for binding the TAZ2 domain, was shown to form a single helix when bound to the MDM2 oncoprotein

(Kussie *et al.*, 1996). The B-Myb 275-376 region contains a sequence (VVEAANLLI) with the potential to form an amphipathic  $\alpha$ -helix similar to that found in the c-Myb transactivation domain (discussed in section section 5.2). Furthermore an adjacent leucine rich sequence (LSEALDLIE) in the B-Myb 275-376 region (residues 308 to 316) also has the capacity to form an amphipathic helix (figure 5.3) which was not predicted by the Jpred2 analysis (section 3.2.1.1). The helical wheel analysis clearly shows the potential of the sequence to adopt an extensive hydrophobic face for potential protein binding. Similar leucine-rich sequences, including the canonical LXXLL motif (Heery *et al.*, 1997; Heery *et al.*, 2001; Matsuda *et al.*, 2004), are known to form amphipathic helices that mediate protein-protein interactions (Livengood *et al.*, 2002). Thus the B-Myb 275-376 transactivation region contains at least two potential helix forming sequences that may be responsible for the interaction with the TAZ2 domain.



**Figure 5.3.** Helical wheel analysis of the leucine rich B-Myb (308-316) sequence. Residues shaded in blue denote hydrophobic amino acids, red denote charged residues with green representing polar amino acids. An extensive hydrophobic face demonstrates the capacity of this sequence to form an amphipathic helix.

### 5.8 *cyclin D1 and its interaction with B-Myb*

The His-tagged cyclin D1 protein purified and characterised in this study was shown to adopt a stable folded tertiary structure but failed to bind a central region of B-Myb spanning residues 227 to 508. This is in direct contrast to the reported functional interaction between B-Myb and cyclin D1 (Cesi *et al.*, 2002; Horstmann *et al.*, 2000b; Schubert *et al.*, 2004). The latter two studies demonstrate the interaction is mediated specifically between the two respective proteins rather than through an associated co-factor and that binding to the central region of B-Myb (residues 227 to 508) was abrogated by p300. In a recent correspondence with Professor Karl-Heinz Klempnauer's group in Germany, they have disclosed that the region of B-Myb encompassing residues 227 to 508 may not alone be sufficient to mediate an interaction with cyclin D1. Further work in their laboratory has implicated the involvement of the B-Myb DNA-binding domain with the binding of cyclin D1 (Unpublished observations). This may explain why there is no observed interaction between cyclin D1 and the B-Myb 227-508 amino acid region reported in this thesis. The B-Myb DNA-binding domain is a well documented protein binding module with interactions reported for poly (ADP-ribose) polymerase (PARP) (Cervellera *et al.*, 2000) and TAF<sub>II</sub>250 (Bartusel *et al.*, 2003).

### 5.9 *Conclusions*

The work described in this thesis adds to an increasing body of data relating to a functional interaction between B-Myb and the p300/CBP family of transcriptional co-activators (Bessa *et al.*, 2001; Johnson *et al.*, 2002; Schubert *et al.*, 2004). As a result of extensive protein-protein interaction mapping studies a region of p300 containing the TAZ2 domain was conclusively shown to specifically bind a central, transactivation associated region of B-Myb with high affinity. The B-Myb transactivation region is intrinsically disordered in isolation and exhibits coupled folding upon binding p300. The tight binding measured for the B-Myb-p300 complex is comparable to the interaction affinities reported for numerous transcriptional activation domains in complex with their binding partners. An adjacent region of B-Myb containing part of exon 9a was shown to specifically interact with a ZZ domain containing region of p300 in a substantially weaker association than that observed for the B-Myb-TAZ2 interaction. Taken together this suggests that the C/H3 region of p300 provides an

extensive interaction surface for B-Myb with the TAZ2 domain representing the principal binding site, and a ZZ domain region serving to stabilise this interaction.

Contrary to the findings of previous published data (Horstmann *et al.*, 2000b; Schubert *et al.*, 2004), a central region of B-Myb spanning residues 227 to 508 failed to bind full-length cyclin D1 protein. The recombinant His-tagged cyclin D1 protein used in this study was shown conclusively to adopt a stable tertiary structure. The absence of binding suggests that the B-Myb 227-508 amino acid region alone is not sufficient to mediate an interaction with cyclin D1. It also implies the inhibitory function of cyclin D1 with respect to B-Myb transcriptional activity is not as a result of competition with p300 for a common binding site within B-Myb.

## APPENDIX

### *A.1 Reagents and culture media*

#### *A.1.1 Selected reagents*

##### 1 % (w/v) Agarose

1 g agarose, made up to 100 ml with 1x TAE, with the addition of 0.5 µg/ml ethidium bromide.

##### LDS Sample buffer

4.0 g glycerol, 0.682 g Tris base, 0.666 g Tris HCl, 0.800 g lithium dodecyl sulphate (LDS), 0.006 g EDTA, Serva Blue G250 0,75 ml of 1 % solution, Phenol Red 0.25 ml of 1 % solution in Ultrapure water to 10 ml (Invitrogen).

##### MES SDS running buffer: (20x stock solution, Invitrogen).

97.6 g 2-(N-morpholino) ethanesulphonic acid (MES), 60.6 g Tris base, 10.0 g sodium dodecyl sulphate (SDS), 3.0 g EDTA in Ultrapure water to 500 ml

##### 6x Blue Orange-dye

0.4 % orange G, 0.03 % bromophenol blue, 0.03 % xylene cyanol FF, 15 % Ficoll<sup>®</sup> 400, 10 mM Tris-HCl (pH 7.5) and 50 mM EDTA (pH 8.0) (Promega).

##### Coomassie blue stain

50 % (v/v) Methanol, 10 % (v/v) Acetic Acid, 0.2 % (w/v) Coomassie brilliant blue R250

Transfer buffer (1 litre)

Glycine 2.9g (39 mM), Tris base 5.8g (48 mM) 0.04 % (w/v) SDS, 20 % (v/v)

Methanol (200 ml)

TBS buffer (1 litre)

Tris-HCl (20 mM), NaCl (137 mM)

TBST buffer

TBS buffer + 0.1 % (v/v) Tween-20

Blocking buffer (Prepared fresh on the day of use)

TBST buffer + 5 % (w/v) nonfat dry milk + 1 % (w/v) Bovine serum albumin (BSA).

*A.1.2 Culture media*

Luria Bertani (LB) broth (1 litre)

10g tryptone, 5g yeast extracts, and 10g NaCl

LB agar

Using the same recipe for LB broth add agar to 15 g/L. Autoclave agar and allow to cool before addition of desired antibiotic.

## APPENDIX

### A.2 *E. coli* strains

#### DH5 $\alpha$

*F<sup>-</sup>  $\phi$ 80lacZ $\Delta$ M15  $\Delta$  (lacZYA-argF)U169 deoR recA1 endA1 hsdR17(*rk<sup>-</sup>*, *m<sub>k</sub><sup>+</sup>*) phoA supE44 thi-1 gyrA96 relA1  $\lambda<sup>-</sup>$*

#### BL21 (DE3)

*F<sup>-</sup> ompT hsdS<sub>B</sub>(*r<sub>B</sub><sup>-</sup>* *m<sub>B</sub><sup>-</sup>*) gal dcm (DE3)*

#### BL21 (DE3) CodonPlus-RIL (Stratagene)

*F<sup>-</sup> ompT hsdS(*r<sub>B</sub><sup>-</sup>* *m<sub>B</sub><sup>-</sup>*) dcm<sup>+</sup> Tet<sup>r</sup> gal endA Hte [argU ileY leuW Camr]*

#### BL21 (DE3) CodonPlus-RP (Stratagene)

*F<sup>-</sup> ompT hsdS(*r<sub>B</sub><sup>-</sup>* *m<sub>B</sub><sup>-</sup>*) dcm<sup>+</sup> Tet<sup>r</sup> gal endA Hte [argU proL Camr]*

### A.3 *Competent cell preparation, transformation and restriction digest*

#### A.3.1 *Preparation of competent cells*

Prepare LB agar plates without antibiotic and streak with bacterial strain of interest.

Incubate overnight at 37 °C.

Prepare SOB medium (100 ml) as follows:

Tryptone 2 g, Yeast extracts 0.5 g, NaCl 0.05 g.

Add 95 ml-deionised water; stir until mixture dissolves completely. Add 1 ml of 250 mM KCl, adjust pH to 7.0 and bring volume up to 100 ml. Aliquot 50 ml each into two 250 ml flasks and autoclave. Add 0.5 ml sterile (by autoclaving) 2 M MgCl<sub>2</sub> just before use. Prepare 50 mM CaCl<sub>2</sub> and filter sterilize by passage through a 0.2  $\mu$ m filter into a previously autoclaved Schott bottle.

#### Methodology:

Inoculate a single colony from the culture plates into 10 ml LB broth and incubate at 37 °C with shaking at 200rpm overnight. Use 1 ml of the overnight culture to inoculate 50 ml SOB and incubate at 37 °C with shaking at 200rpm until OD<sub>600</sub> is 0.5 (1cm path length cell). Harvest the cells by centrifugation at 6000g at 4 °C for 15 minutes. Resuspend cells very gently in 25 ml ice-cold 50 mM CaCl<sub>2</sub> and leave on ice for 30 minute. Pellet cells at 6000g at 4 °C for 15 minutes and resuspend very gently in 2 ml ice-cold 50 mM CaCl<sub>2</sub>, keeping cells on ice. Add sterile ice-cold 50 % glycerol to 15 % and mix gently by inversion. Aliquot 200 µl cells into cryogen tubes, flash freeze in liquid nitrogen and store at -80°C until required.

#### A.3.2 Transformation of Competent Cells

Prepare LB agar plates with the appropriate antibiotic and LB broth without antibiotic. Working on ice, aliquot 50 µl competent cells into 1.5 ml eppendorf tubes and add between 1-4 µl of the plasmid. Mix gently and leave on ice for 30 minutes. Immediately transfer the cells to a water bath set at 42 °C and heat shock the cells for 45 seconds (this time is critical). Transfer the cells to ice and leave for 5 minutes.

Add 300 µl LB broth (no antibiotic) and incubate for 45 minutes at 37 °C at 200rpm with the eppendorfs placed horizontally in the incubator.

Spread 50 and 150 µl aliquots of cells onto LB agar plates with the appropriate antibiotic and incubate at 37 °C overnight.

### *A.3.3 Restriction Digest Protocol*

All restriction digests were carried out as recommended by Promega (the supplier of all restriction enzymes and reagents used for this purpose). Plasmid and PCR restriction digests were carried out in 20  $\mu$ l reaction volumes containing 2  $\mu$ l buffer, 0.2  $\mu$ l BSA (0.1 mg/ml), plasmid DNA or PCR product to 1  $\mu$ g, 1  $\mu$ l of appropriate enzyme/s (2-10 u) and deionised water to 20  $\mu$ l.

The reaction mixture was mixed gently and centrifuged briefly for a few seconds and incubated at 37 °C for 4 hours (unless stated otherwise). Samples were stored at – 20 °C until required.

## REFERENCES

- Adnane, J., Shao, Z. and Robbins, P.D.: Cyclin D1 associates with the TBP-associated factor TAF(II)250 to regulate Sp1-mediated transcription. *Oncogene* **18** (1999) 239-47.
- Agalioti, T., Chen, G. and Thanos, D.: Deciphering the transcriptional histone acetylation code for a human gene. *Cell* **111** (2002) 381-92.
- Albright, S.R. and Tjian, R.: TAFs revisited: more data reveal new twists and confirm old ideas. *Gene* **242** (2000) 1-13.
- Amatschek, S., Koenig, U., Auer, H., Steinlein, P., Pacher, M., Gruenfelder, A., Dekan, G., Vogl, S., Kubista, E., Heider, K.H., Stratowa, C., Schreiber, M. and Sommergruber, W.: Tissue-wide expression profiling using cDNA subtraction and microarrays to identify tumor-specific genes. *Cancer Res.* **64** (2004) 844-56.
- Anderson, J.T., Rogers, R.P. and Jarrett, H.W.: Ca<sup>2+</sup>-calmodulin binds to the carboxyl-terminal domain of dystrophin. *J. Biol. Chem.* **271** (1996) 6605-10.
- Ansieau, S., Kowenz-Leutz, E., Dechend, R. and Leutz, A.: B-Myb, a repressed transactivating protein. *J. Mol. Med.* **75** (1997) 815-9.
- Arsura, M., Introna, M., Passerini, F., Mantovani, A. and Golay, J.: B-myb antisense oligonucleotides inhibit proliferation of human hematopoietic cell lines. *Blood* **79** (1992) 2708-16.
- Asahara, H., Dutta, S., Kao, H.Y., Evans, R.M. and Montminy, M.: Pbx-Hox heterodimers recruit coactivator-corepressor complexes in an isoform-specific manner. *Mol. Cell. Biol.* **19** (1999) 8219-25.
- Avantaggiati, M.L., Ogryzko, V., Gardner, K., Giordano, A., Levine, A.S. and Kelly, K.: Recruitment of p300/CBP in p53-dependent signal pathways. *Cell* **89** (1997) 1175-84.
- Ayed, A., Mulder, F.A., Yi, G.S., Lu, Y., Kay, L.E. and Arrowsmith, C.H.: Latent and active p53 are identical in conformation. *Nat. Struct. Biol.* **8** (2001) 756-60.
- Bailey, P., Downes, M., Lau, P., Harris, J., Chen, S.L., Hamamori, Y., Sartorelli, V. and Muscat, G.E.: The nuclear receptor corepressor N-CoR regulates differentiation: N-CoR directly interacts with MyoD. *Mol. Endocrinol.* **13** (1999) 1155-68.
- Bannister, A.J. and Kouzarides, T.: The CBP co-activator is a histone acetyltransferase. *Nature* **384** (1996) 641-3.

- Bannister, A.J. and Kouzarides, T.: CBP-induced stimulation of c-Fos activity is abrogated by E1A. *Embo. J.* **14** (1995) 4758-62.
- Bartsch, O., Horstmann, S., Toprak, K., Klempnauer, K.H. and Ferrari, S.: Identification of cyclin A/Cdk2 phosphorylation sites in B-Myb. *Eur. J. Biochem.* **260** (1999) 384-91.
- Bartusel, T. and Klempnauer, K.H.: Transactivation mediated by B-Myb is dependent on TAF(II)250. *Oncogene* **22** (2003) 2932-41.
- Bennett, J.D., Farlie, P.G. and Watson, R.J.: E2F binding is required but not sufficient for repression of B-myb transcription in quiescent fibroblasts. *Oncogene* **13** (1996) 1073-82.
- Berger, S.L., Pina, B., Silverman, N., Marcus, G.A., Agapite, J., Regier, J.L., Triezenberg, S.J. and Guarente, L.: Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains. *Cell* **70** (1992) 251-65.
- Bergholtz, S., Andersen, T.O., Andersson, K.B., Borrebaek, J., Luscher, B. and Gabrielsen, O.S.: The highly conserved DNA-binding domains of A-, B- and c-Myb differ with respect to DNA-binding, phosphorylation and redox properties. *Nucleic Acids Res.* **29** (2001) 3546-56.
- Bessa, M., Joaquin, M., Tavner, F., Saville, M.K. and Watson, R.J.: Regulation of the cell cycle by B-Myb. *Blood Cells Mol. Dis.* **27** (2001) 416-21.
- Bessa, M., Saville, M.K. and Watson, R.J.: Inhibition of cyclin A/Cdk2 phosphorylation impairs B-Myb transactivation function without affecting interactions with DNA or the CBP coactivator. *Oncogene* **20** (2001) 3376-86.
- Biedenkapp, H., Borgmeyer, U., Sippel, A.E. and Klempnauer, K.H.: *Nature* **335** (1988) 835-837.
- Bies, J., Hoffman, B., Amanullah, A., Giese, T. and Wolff, L.: B-Myb prevents growth arrest associated with terminal differentiation of monocytic cells. *Oncogene* **12** (1996) 355-63.
- Blobel, G.A., Nakajima, T., Eckner, R., Montminy, M. and Orkin, S.H.: CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proc. Natl. Acad. Sci. U S A* **95** (1998) 2061-6.
- Borrow, J., Stanton, V.P., Jr., Andresen, J.M., Becher, R., Behm, F.G., Chaganti, R.S., Civin, C.I., Disteché, C., Dube, I., Frischauf, A.M., Horsman, D., Mitelman, F.,

- Volinia, S., Watmore, A.E. and Housman, D.E.: The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein. *Nat. Genet.* **14** (1996) 33-41.
- Brandt, T.L., Fraser, D.J., Leal, S., Halandras, P.M., Kroll, A.R. and Kroll, D.J.: c-Myb trans-activates the human DNA topoisomerase IIalpha gene promoter. *J. Biol. Chem.* **272** (1997) 6278-84.
- Broccoli, D., Smogorzewska, A., Chong, L. and de Lange, T.: Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat. Genet.* **17** (1997) 231-5.
- Brown, N.R., Noble, M.E., Endicott, J.A., Garman, E.F., Wakatsuki, S., Mitchell, E., Rasmussen, B., Hunt, T. and Johnson, L.N.: The crystal structure of cyclin A. *Structure* **3** (1995) 1235-47.
- Butler, J.E. and Kadonaga, J.T.: The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev* **16** (2002) 2583-92.
- Carr, M.D., Wollborn, U., McIntosh, P.B., Frenkiel, T.A., McCormick, J.E., Bauer, C.J., Klempnauer, K.H. and Feeney, J.: Structure of the B-Myb DNA-binding domain in solution and evidence for multiple conformations in the region of repeat-2 involved in DNA binding: implications for sequence-specific DNA binding by Myb proteins. *Eur. J. Biochem.* **235** (1996) 721-35.
- Cervellera, M.N. and Sala, A.: Poly(ADP-ribose) polymerase is a B-MYB coactivator. *J. Biol. Chem.* **275** (2000) 10692-6.
- Cesi, V., Tanno, B., Vitali, R., Mancini, C., Giuffrida, M.L., Calabretta, B. and Raschella, G.: Cyclin D1-dependent regulation of B-myb activity in early stages of neuroblastoma differentiation. *Cell Death Differ.* **9** (2002) 1232-9.
- Charrasse, S., Carena, I., Brondani, V., Klempnauer, K.H. and Ferrari, S.: Degradation of B-Myb by ubiquitin-mediated proteolysis: involvement of the Cdc34-SCF(p45Skp2) pathway. *Oncogene* **19** (2000) 2986-95.
- Chen, H., Lin, R.J., Schiltz, R.L., Chakravarti, D., Nash, A., Nagy, L., Privalsky, M.L., Nakatani, Y. and Evans, R.M.: Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell* **90** (1997) 569-80.
- Chrivia, J.C., Kwok, R.P., Lamb, N., Hagiwara, M., Montminy, M.R. and Goodman, R.H.: Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **365** (1993) 855-9.

- Cirillo, L.A., Lin, F.R., Cuesta, I., Friedman, D., Jarnik, M. and Zaret, K.S.: Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol Cell* **9** (2002) 279-89.
- Clarke, M.F., Kukowska-Latallo, J.F., Westin, E., Smith, M. and Prochownik, E.V.: Constitutive expression of a c-myc cDNA blocks Friend murine erythroleukemia cell differentiation. *Mol. Cell. Biol.* **8** (1988) 884-92.
- Cooper, J.P., Nimmo, E.R., Allshire, R.C. and Cech, T.R.: Regulation of telomere length and function by a Myb-domain protein in fission yeast. *Nature* **385** (1997) 744-7.
- Coqueret, O.: Linking cyclins to transcriptional control. *Gene* **299** (2002) 35-55.
- Courey, A.J. and Tjian, R.: Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* **55** (1988) 887-98.
- Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M. and Barton, G.J.: JPred: a consensus secondary structure prediction server. *Bioinformatics* **14** (1998) 892-3.
- Dai, P., Akimaru, H., Tanaka, Y., Hou, D.X., Yasukawa, T., Kanei-Ishii, C., Takahashi, T. and Ishii, S.: CBP as a transcriptional coactivator of c-Myb. *Genes Dev.* **10** (1996) 528-40.
- Dames, S.A., Martinez-Yamout, M., De Guzman, R.N., Dyson, H.J. and Wright, P.E.: Structural basis for Hif-1 alpha /CBP recognition in the cellular hypoxic response. *Proc. Natl. Acad. Sci. U S A.* **99** (2002) 5271-6.
- Dash, A.B., Orrico, F.C. and Ness, S.A.: The EVES motif mediates both intermolecular and intramolecular regulation of c-Myb. *Genes Dev.* **10** (1996) 1858-69.
- De Guzman, R.N., Liu, H.Y., Martinez-Yamout, M., Dyson, H.J. and Wright, P.E.: Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP. *J Mol. Biol.* **303** (2000) 243-53.
- De Guzman, R.N., Wojciak, J.M., Martinez-Yamout, M.A., Dyson, H.J. and Wright, P.E.: CBP/p300 TAZ1 domain forms a structured scaffold for ligand binding. *Biochemistry* **44** (2005) 490-7.
- Demarest, S.J., Chung, J., Dyson, H.J. and Wright, P.E.: Assignment of a 15 kDa protein complex formed between the p160 coactivator ACTR and CREB binding protein. *J. Biomol. NMR.* **22** (2002) 377-8.
- Desterro, J.M., Rodriguez, M.S. and Hay, R.T.: Regulation of transcription factors by protein degradation. *Cell Mol. Life Sci.* **57** (2000) 1207-19.

- Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K. and Zhou, M.M.: Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399** (1999) 491-6.
- Dial, R., Sun, Z.Y. and Freedman, S.J.: Three conformational states of the p300 CH1 domain define its functional properties. *Biochemistry* **42** (2003) 9937-45.
- Ding, H.F., Bustin, M. and Hansen, U.: Alleviation of histone H1-mediated transcriptional repression and chromatin compaction by the acidic activation region in chromosomal protein HMG-14. *Mol Cell Biol* **17** (1997) 5843-55.
- Dobson, C.M.: Protein folding. Solid evidence for molten globules. *Curr. Biol.* **4** (1994) 636-40.
- Dudev, T. and Lim, C.: Principles governing Mg, Ca, and Zn binding and selectivity in proteins. *Chem. Rev.* **103** (2003) 773-88.
- Dundr, M., Hoffmann-Rohrer, U., Hu, Q., Grummt, I., Rothblum, L.I., Phair, R.D. and Misteli, T.: A kinetic framework for a mammalian RNA polymerase in vivo. *Science* **298** (2002) 1623-6.
- Dunker, A.K., Brown, C.J., Lawson, J.D., Iakoucheva, L.M. and Obradovic, Z.: Intrinsic disorder and protein function. *Biochemistry* **41** (2002) 6573-82.
- Duprey, S.P. and Boettiger, D.: Developmental regulation of c-myb in normal myeloid progenitor cells. *Proc. Natl. Acad. Sci. U S A* **82** (1985) 6937-41.
- Dyson, H.J. and Wright, P.E.: Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **12** (2002) 54-60.
- Eckner, R., Ewen, M.E., Newsome, D., Gerdes, M., DeCaprio, J.A., Lawrence, J.B. and Livingston, D.M.: Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes Dev.* **8** (1994) 869-84.
- Elledge, S.J. and Harper, J.W.: The role of protein stability in the cell cycle and cancer. *Biochim. Biophys. Acta.* **1377** (1998) M61-70.
- Facchinetti, V., Loffarelli, L., Schreek, S., Oelgeschlager, M., Luscher, B., Introna, M. and Golay, J.: Regulatory domains of the A-Myb transcription factor and its interaction with the CBP/p300 adaptor molecules. *Biochem. J.* **324** (Pt 3) (1997) 729-36.
- Featherstone, M.: Coactivators in transcription initiation: here are your orders. *Curr Opin Genet Dev* **12** (2002) 149-55.

- Foos, G., Natour, S. and Klempnauer, K.H.: TATA-box dependent trans-activation of the human HSP70 promoter by Myb proteins. *Oncogene* **8** (1993) 1775-82.
- Foos, G., Grimm, S. and Klempnauer, K.H.: Functional antagonism between members of the myb family: B-myb inhibits v-myb-induced gene activation. *Embo J.* **11** (1992) 4619-29.
- Francastel, C., Magis, W. and Groudine, M.: Nuclear relocation of a transactivator subunit precedes target gene activation. *Proc Natl Acad Sci U S A* **98** (2001) 12120-5.
- Garner, E., Cannon, P., Romero, P., Obradovic, Z. and Dunker, A.K.: Predicting Disordered Regions from Amino Acid Sequence: Common Themes Despite Differing Structural Characterization. Genome Inform Ser Workshop *Genome Inform.* **9** (1998) 201-213.
- Golay, J., Capucci, A., Arsura, M., Castellano, M., Rizzo, V. and Introna, M.: Expression of c-myb and B-myb, but not A-myb, correlates with proliferation in human hematopoietic cells. *Blood* **77** (1991) 149-58.
- Gonda, T.J. and Metcalf, D.: Expression of myb, myc and fos proto-oncogenes during the differentiation of a murine myeloid leukaemia. *Nature* **310** (1984) 249-51.
- Gonda, T.J., Sheiness, D.K. and Bishop, J.M.: Transcripts from the cellular homologs of retroviral oncogenes: distribution among chicken tissues. *Mol. Cell. Biol.* **2** (1982) 617-24.
- Graf, T.: Myb: a transcriptional activator linking proliferation and differentiation in hematopoietic cells. *Curr. Opin. Genet. Dev.* **2** (1992) 249-55.
- Gross, C.A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J. and Young, B.: The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harb Symp Quant Biol* **63** (1998) 141-55.
- Grossmann, J.G., Sharff, A.J., O'Hare, P. and Luisi, B.: Molecular shapes of transcription factors TFIIB and VP16 in solution: implications for recognition. *Biochemistry* **40** (2001) 6267-74.
- Grunstein, M.: Yeast heterochromatin: regulation of its assembly and inheritance by histones. *Cell* **93** (1998) 325-8.
- Gu, W., Shi, X.L. and Roeder, R.G.: Synergistic activation of transcription by CBP and p53. *Nature* **387** (1997) 819-23.

- Hadwiger, J.A., Wittenberg, C., Richardson, H.E., de Barros Lopes, M. and Reed, S.I.: A family of cyclin homologs that control the G1 phase in yeast. *Proc. Natl. Acad. Sci. U S A.* **86** (1989) 6255-9.
- Hagemeier, C., Cook, A. and Kouzarides, T.: The retinoblastoma protein binds E2F residues required for activation in vivo and TBP binding in vitro. *Nucleic Acids Res.* **21** (1993) 4998-5004.
- Hecht, A., Laroche, T., Strahl-Bolsinger, S., Gasser, S.M. and Grunstein, M.: Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast. *Cell* **80** (1995) 583-92.
- Heery, D.M., Hoare, S., Hussain, S., Parker, M.G. and Sheppard, H.: Core LXXLL motif sequences in CREB-binding protein, SRC1, and RIP140 define affinity and selectivity for steroid and retinoid receptors. *J. Biol. Chem.* **276** (2001) 6695-702.
- Heery, D.M., Kalkhoven, E., Hoare, S. and Parker, M.G.: A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387** (1997) 733-6.
- Hirai, H. and Sherr, C.J.: Interaction of D-type cyclins with a novel myb-like transcription factor, DMP1. *Mol. Cell. Biol.* **16** (1996) 6457-67.
- Hope, I.A. and Struhl, K.: Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* **46** (1986) 885-94.
- Horlein, A.J., Naar, A.M., Heinzl, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Soderstrom, M., Glass, C.K. and et al.: Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature* **377** (1995) 397-404.
- Horstmann, S., Ferrari, S. and Klempnauer, K.H.: An alternatively spliced isoform of B-Myb is a transcriptional inhibitor. *Oncogene* **19** (2000) 5428-34.
- Horstmann, S., Ferrari, S. and Klempnauer, K.H.: Regulation of B-Myb activity by cyclin D1. *Oncogene* **19** (2000) 298-306.
- Howe, K.M., Reakes, C.F. and Watson, R.J.: Characterization of the sequence-specific interaction of mouse c-myb protein with DNA. *Embo. J.* **9** (1990) 161-9.
- Howe, K.M. and Watson, R.J.: *Nucleic Acids Res.* **19** (1991) 3913-3919.
- Hu, Y.L., Ramsay, R.G., Kanei-Ishii, C., Ishii, S. and Gonda, T.J.: Transformation by carboxyl-deleted Myb reflects increased transactivating capacity and disruption of a negative regulatory domain. *Oncogene* **6** (1991) 1549-53.

- Iakoucheva, L.M., Brown, C.J., Lawson, J.D., Obradovic, Z. and Dunker, A.K.: Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* **323** (2002) 573-84.
- Iakoucheva, L.M., Radivojac, P., Brown, C.J., O'Connor, T.R., Sikes, J.G., Obradovic, Z. and Dunker, A.K.: The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32** (2004) 1037-49.
- Ibanez, C.E. and Lipsick, J.S.: trans activation of gene expression by v-myb. *Mol. Cell. Biol.* **10** (1990) 2285-93.
- Inoue, K. and Sherr, C.J.: Gene expression and cell cycle arrest mediated by transcription factor DMP1 is antagonized by D-type cyclins through a cyclin-dependent-kinase-independent mechanism. *Mol. Cell. Biol.* **18** (1998) 1590-600.
- Inoue, K., Sherr, C.J. and Shapiro, L.H.: Regulation of the CD13/aminopeptidase N gene by DMP1, a transcription factor antagonized by D-type cyclins. *J. Biol. Chem.* **273** (1998) 29188-94.
- Iwai, N., Kitajima, K., Sakai, K., Kimura, T. and Nakano, T.: Alteration of cell adhesion and cell cycle properties of ES cells by an inducible dominant interfering Myb mutant. *Oncogene* **20** (2001) 1425-34.
- Jacobson, R.H., Ladurner, A.G., King, D.S. and Tjian, R.: Structure and function of a human TAFII250 double bromodomain module. *Science* **288** (2000) 1422-5.
- Jamin, N., Gabrielsen, O.S., Gilles, N., Lirsac, P.N. and Toma, F.: Secondary structure of the DNA-binding domain of the c-Myb oncoprotein in solution. A multidimensional double and triple heteronuclear NMR study. *Eur. J. Biochem.* **216** (1993) 147-54.
- Janknecht, R. and Hunter, T.: Transcription. A growing coactivator network. *Nature* **383** (1996) 22-3.
- Jeffrey, P.D., Russo, A.A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J. and Pavletich, N.P.: Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* **376** (1995) 313-20.
- Jiskoot, W., Hlady, V., Naleway, J.J. and Herron, J.N.: Application of fluorescence spectroscopy for determining the structure and function of proteins. *Pharm. Biotechnol.* **7** (1995) 1-63.
- Joaquin, M., Bessa, M., Saville, M.K. and Watson, R.J.: B-Myb overcomes a p107-mediated cell proliferation block by interacting with an N-terminal domain of p107. *Oncogene* **21** (2002) 7923-32.

- Johnson, L.R., Johnson, T.K., Desler, M., Luster, T.A., Nowling, T., Lewis, R.E. and Rizzino, A.: Effects of B-Myb on gene transcription: phosphorylation-dependent activity and acetylation by p300. *J. Biol. Chem.* **277** (2002) 4088-97.
- Johnson, T.K., Schweppe, R.E., Septer, J. and Lewis, R.E.: Phosphorylation of B-Myb regulates its transactivation potential and DNA binding. *J. Biol. Chem.* **274** (1999) 36741-9.
- Kalkbrenner, F., Guehmann, S. and Moelling, K.: Transcriptional activation by human c-myb and v-myb genes. *Oncogene* **5** (1990) 657-61.
- Kalkhoven, E., Teunissen, H., Houweling, A., Verrijzer, C.P. and Zantema, A.: The PHD type zinc finger is an integral part of the CBP acetyltransferase domain. *Mol. Cell. Biol.* **22** (2002) 1961-70.
- Kamano, H., Burk, B., Noben-Trauth, K. and Klempnauer, K.H.: Differential splicing of the mouse B-myb gene. *Oncogene* **11** (1995) 2575-82.
- Kanei-Ishii, C., MacMillan, E.M., Nomura, T., Sarai, A., Ramsay, R.G., Aimoto, S., Ishii, S. and Gonda, T.J.: Transactivation and transformation by Myb are negatively regulated by a leucine-zipper structure. *Proc. Natl. Acad. Sci. U S A* **89** (1992) 3088-92.
- Kim, T., Jung, H., Min, S., Kim, K.T. and Ha, H.: B-myb proto-oncogene products interact in vivo with each other via the carboxy-terminal conserved region. *FEBS Lett.* **460** (1999) 363-8.
- Klempnauer, K.H., Gonda, T.J. and Bishop, J.M.: Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene. *Cell* **31** (1982) 453-63.
- Konig, P., Fairall, L. and Rhodes, D.: Sequence-specific DNA recognition by the myb-like domain of the human telomere binding protein TRF1: a model for the protein-DNA complex. *Nucleic Acids Res.* **26** (1998) 1731-40.
- Kussie, P.H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A.J. and Pavletich, N.P.: Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **274** (1996) 948-53.
- Lam, E.W., Bennett, J.D. and Watson, R.J.: Cell-cycle regulation of human B-myb transcription. *Gene* **160** (1995) 277-81.
- Lam, E.W., Morris, J.D., Davies, R., Crook, T., Watson, R.J. and Vousden, K.H.: HPV16 E7 oncoprotein deregulates B-myb expression: correlation with targeting of p107/E2F complexes. *Embo J.* **13** (1994) 871-8.

- Lam, E.W., Robinson, C. and Watson, R.J.: Characterization and cell cycle-regulated expression of mouse B-myb. *Oncogene* **7** (1992) 1885-90.
- Lam, E.W. and Watson, R.J.: An E2F-binding site mediates cell-cycle regulated repression of mouse B-myb transcription. *Embo J.* **12** (1993) 2705-13.
- Lane, S., Farlie, P. and Watson, R.: B-Myb function can be markedly enhanced by cyclin A-dependent kinase and protein truncation. *Oncogene* **14** (1997) 2445-53.
- Lane, T., Ibanez, C., Garcia, A., Graf, T. and Lipsick, J.: Transformation by v-myb correlates with trans-activation of gene expression. *Mol. Cell. Biol.* **10** (1990) 2591-8.
- Lee, J.S., Galvin, K.M., See, R.H., Eckner, R., Livingston, D., Moran, E. and Shi, Y.: Relief of YY1 transcriptional repression by adenovirus E1A is mediated by E1A-associated protein p300. *Genes Dev.* **9** (1995) 1188-98.
- Legge, G.B., Martinez-Yamout, M.A., Hambly, D.M., Trinh, T., Lee, B.M., Dyson, H.J. and Wright, P.E.: ZZ domain of CBP: an unusual zinc finger fold in a protein interaction module. *J. Mol. Biol.* **343** (2004) 1081-93.
- Lemon, B. and Tjian, R.: Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev* **14** (2000) 2551-69.
- Li, X. and McDonnell, D.P.: The transcription factor B-Myb is maintained in an inhibited state in target cells through its interaction with the nuclear corepressors N-CoR and SMRT. *Mol. Cell. Biol.* **22** (2002) 3663-73.
- Lin, D., Fiscella, M., O'Connor, P.M., Jackman, J., Chen, M., Luo, L.L., Sala, A., Travali, S., Appella, E. and Mercer, W.E.: Constitutive expression of B-myb can bypass p53-induced Waf1/Cip1-mediated G1 arrest. *Proc. Natl. Acad. Sci. U S A* **91** (1994) 10079-83.
- Lipsick, J.S.: derick. *Oncogene* **13** (1996) 223-235.
- Livengood, J.A., Scoggin, K.E., Van Orden, K., McBryant, S.J., Edayathumangalam, R.S., Laybourn, P.J. and Nyborg, J.K.: p53 Transcriptional activity is mediated through the SRC1-interacting domain of CBP/p300. *J. Biol. Chem.* **277** (2002) 9054-61.
- Lomvardas, S. and Thanos, D.: Nucleosome sliding via TBP DNA binding in vivo. *Cell* **106** (2001) 685-96.
- Losick, R.: Summary: three decades after sigma. *Cold Spring Harb Symp Quant Biol* **63** (1998) 653-66.

- Luscher, B. and Eisenman, R.N.: New light on Myc and Myb. Part II. Myb. *Genes Dev.* **4** (1990) 2235-41.
- Marhamati, D.J. and Sonenshein, G.E.: B-Myb expression in vascular smooth muscle cells occurs in a cell cycle-dependent fashion and down-regulates promoter activity of type I collagen genes. *J. Biol. Chem.* **271** (1996) 3359-65.
- Matsuda, S., Harries, J.C., Viskaduraki, M., Troke, P.J., Kindle, K.B., Ryan, C. and Heery, D.M.: A Conserved alpha-helical motif mediates the binding of diverse nuclear proteins to the SRC1 interaction domain of CBP. *J. Biol. Chem.* **279** (2004) 14055-64.
- McIntosh, P.B., Frenkiel, T.A., Wollborn, U., McCormick, J.E., Klempnauer, K.H., Feeney, J. and Carr, M.D.: Solution structure of the B-Myb DNA-binding domain: a possible role for conformational instability of the protein in DNA binding and control of gene expression. *Biochemistry* **37** (1998) 9619-29.
- Merika, M. and Thanos, D.: Enhanceosomes. *Curr Opin Genet Dev* **11** (2001) 205-8.
- Mitchell, P.J. and Tjian, R.: Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245** (1989) 371-8.
- Mizuguchi, G., Nakagoshi, H., Nagase, T., Nomura, N., Date, T., Ueno, Y. and Ishii, S.: DNA binding activity and transcriptional activator function of the human B-myb protein compared with c-MYB. *J. Biol. Chem.* **265** (1990) 9280-4.
- Mucenski, M.L., McLain, K., Kier, A.B., Swerdlow, S.H., Schreiner, C.M., Miller, T.A., Pietryga, D.W., Scott, W.J., Jr. and Potter, S.S.: A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell* **65** (1991) 677-89.
- Muller-Tidow, C., Wang, W., Idos, G.E., Diederichs, S., Yang, R., Readhead, C., Berdel, W.E., Serve, H., Saville, M., Watson, R. and Koeffler, H.P.: Cyclin A1 directly interacts with B-myb and cyclin A1/cdk2 phosphorylate B-myb at functionally important serine and threonine residues: tissue-specific regulation of B-myb function. *Blood* **97** (2001) 2091-7.
- Myrset, A.H., Bostad, A., Jamin, N., Lirsac, P.N., Toma, F. and Gabrielsen, O.S.: DNA and redox state induced conformational changes in the DNA-binding domain of the Myb oncoprotein. *Embo. J* **12** (1993) 4625-33.
- Nakagoshi, H., Kanei-Ishii, C., Sawazaki, T., Mizuguchi, G. and Ishii, S.: Transcriptional activation of the c-myc gene by the c-myb and B-myb gene products. *Oncogene* **7** (1992) 1233-40.

- Nakagoshi, H., Takemoto, Y. and Ishii, S.: Functional domains of the human B-myb gene product. *J. Biol. Chem.* **268** (1993) 14161-7.
- Namba, K.: Roles of partly unfolded conformations in macromolecular self-assembly. *Genes Cells* **6** (2001) 1-12.
- Noble, M.E., Endicott, J.A., Brown, N.R. and Johnson, L.N.: The cyclin box fold: protein recognition in cell-cycle and transcription control. *Trends Biochem. Sci.* **22** (1997) 482-7.
- Nomura, N., Takahashi, M., Matsui, M., Ishii, S., Date, T., Sasamoto, S. and Ishizaki, R.: Isolation of human cDNA clones of myb-related genes, A-myb and B-myb. *Nucleic Acids Res.* **16** (1988) 11075-89.
- O'Connor, M.J., Zimmermann, H., Nielsen, S., Bernard, H.U. and Kouzarides, T.: Characterization of an E1A-CBP interaction defines a novel transcriptional adapter motif (TRAM) in CBP/p300. *J. Virol.* **73** (1999) 3574-81.
- Oelgeschlager, M., Janknecht, R., Krieg, J., Schreek, S. and Luscher, B.: Interaction of the co-activator CBP with Myb proteins: effects on Myb-specific transactivation and on the cooperativity with NF-M. *Embo J.* **15** (1996) 2771-80.
- Ogata, K., Hojo, H., Aimoto, S., Nakai, T., Nakamura, H., Sarai, A., Ishii, S. and Nishimura, Y.: Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. *Proc. Natl. Acad. Sci. U S A* **89** (1992) 6428-32.
- Ogata, K., Morikawa, S., Nakamura, H., Hojo, H., Yoshimura, S., Zhang, R., Aimoto, S., Ametani, Y., Hirata, Z., Sarai, A. and et al.: Comparison of the free and DNA-complexed forms of the DNA-binding domain from c-Myb. *Nat. Struct. Biol.* **2** (1995) 309-20.
- Ogata, K., Morikawa, S., Nakamura, H., Sekikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S. and Nishimura, Y.: *Cell* **79** (1994) 639-648.
- Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H. and Nakatani, Y.: The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87** (1996) 953-9.
- Oh, I.H. and Reddy, E.P.: Murine A-myb gene encodes a transcription factor, which cooperates with Ets-2 and exhibits distinctive biochemical and biological activities from c-myb. *J. Biol. Chem.* **272** (1997) 21432-43.

- Oh, I.H. and Reddy, E.P.: The C-terminal domain of B-Myb acts as a positive regulator of transcription and modulates its biological functions. *Mol. Cell. Biol.* **18** (1998) 499-511.
- Oh, I.H. and Reddy, E.P.: The myb gene family in cell growth, differentiation and apoptosis. *Oncogene* **18** (1999) 3017-33.
- Orphanides, G. and Reinberg, D.: A unified theory of gene expression. *Cell* **108** (2002) 439-51.
- Pace, C.N., Vajdos, F., Fee, L., Grimsley, G. and Gray, T.: How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4** (1995) 2411-23.
- Pace, C.N. and Creighton, T.E.: The disulphide folding pathway of ribonuclease T1. *J. Mol. Biol.* **188** (1986) 477-86.
- Parker, D., Rivera, M., Zor, T., Henrion-Caude, A., Radhakrishnan, I., Kumar, A., Shapiro, L.H., Wright, P.E., Montminy, M. and Brindle, P.K.: Role of secondary structure in discrimination between constitutive and inducible activators. *Mol. Cell. Biol.* **19** (1999) 5601-7.
- Poleskaya, A., Naguibneva, I., Duquet, A., Bengal, E., Robin, P. and Harel-Bellan, A.: Interaction between acetylated MyoD and the bromodomain of CBP and/or p300. *Mol. Cell. Biol.* **21** (2001) 5312-20.
- Ponting, C.P., Blake, D.J., Davies, K.E., Kendrick-Jones, J. and Winder, S.J.: ZZ and TAZ: new putative zinc fingers in dystrophin and other proteins. *Trends Biochem. Sci.* **21** (1996) 11-13.
- Ptitsyn, O.B.: How the molten globule became. *Trends Biochem. Sci.* **20** (1995) 376-9.
- Radhakrishnan, I., Perez-Alvarado, G.C., Parker, D., Dyson, H.J., Montminy, M.R. and Wright, P.E.: Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. *Cell* **91** (1997) 741-52.
- Ramsay, R.G., Ishii, S. and Gonda, T.J.: Increase in specific DNA binding by carboxyl truncation suggests a mechanism for activation of Myb. *Oncogene* **6** (1991) 1875-9.
- Raschella, G., Negroni, A., Sala, A., Pucci, S., Romeo, A. and Calabretta, B.: Requirement of b-myb function for survival and differentiative potential of human neuroblastoma cells. *J. Biol. Chem.* **270** (1995) 8540-5.
- Reiss, K., Travali, S., Calabretta, B. and Baserga, R.: Growth regulated expression of B-myb in fibroblasts and hematopoietic cells. *J. Cell. Physiol.* **148** (1991) 338-43.

- Renshaw, P.S., Panagiotidou, P., Whelan, A., Gordon, S.V., Hewinson, R.G., Williamson, R.A. and Carr, M.D.: Conclusive evidence that the major T-cell antigens of the Mycobacterium tuberculosis complex ESAT-6 and CFP-10 form a tight, 1:1 complex and characterization of the structural properties of ESAT-6, CFP-10, and the ESAT-6\*CFP-10 complex. Implications for pathogenesis and virulence. *J. Biol. Chem.* **277** (2002) 21598-603.
- Rentschler, S., Linn, H., Deininger, K., Bedford, M.T., Espanel, X. and Sudol, M.: The WW domain of dystrophin requires EF-hands region to interact with beta-dystroglycan. *Biol. Chem.* **380** (1999) 431-42.
- Reutens, A.T., Fu, M., Wang, C., Albanese, C., McPhaul, M.J., Sun, Z., Balk, S.P., Janne, O.A., Palvimo, J.J. and Pestell, R.G.: Cyclin D1 binds the androgen receptor and regulates hormone-dependent signaling in a p300/CBP-associated factor (P/CAF)-dependent manner. *Mol. Endocrinol.* **15** (2001) 797-811.
- Robinson, C., Light, Y., Groves, R., Mann, D., Marias, R. and Watson, R.: Cell-cycle regulation of B-Myb protein expression: specific phosphorylation during the S phase of the cell cycle. *Oncogene* **12** (1996) 1855-64.
- Roeder, R.G.: The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci* **21** (1996) 327-35.
- Rosinski, J.A. and Atchley, W.R.: Molecular evolution of the Myb family of transcription factors: evidence for polyphyletic origin. *J. Mol. Evol.* **46** (1998) 74-83.
- Roussel, M., Saule, S., Lagrou, C., Rommens, C., Beug, H., Graf, T. and Stehelin, D.: Three new types of viral oncogene of cellular origin specific for haematopoietic cell transformation. *Nature* **281** (1979) 452-5.
- Saikumar, P., Murali, R. and Reddy, E.P.: *Proc. Natl. Acad. Sci. U S A* **87** (1990) 8452-8456.
- Sakura, H., Kanei-Ishii, C., Nagase, T., Nakagoshi, H., Gonda, T.J. and Ishii, S.: *Proc. Natl. Acad. Sci. U S A* **86** (1989) 5758-5762.
- Sala, A. and Calabretta, B.: Regulation of BALB/c 3T3 fibroblast proliferation by B-myb is accompanied by selective activation of cdc2 and cyclin D1 expression. *Proc. Natl. Acad. Sci. U S A* **89** (1992) 10415-9.
- Sala, A., De Luca, A., Giordano, A. and Peschle, C.: The retinoblastoma family member p107 binds to B-MYB and suppresses its autoregulatory activity. *J. Biol. Chem.* **271** (1996) 28738-40.

- Sala, A., Kundu, M., Casella, I., Engelhard, A., Calabretta, B., Grasso, L., Paggi, M.G., Giordano, A., Watson, R.J., Khalili, K. and Peschle, C.: Activation of human B-MYB by cyclins. *Proc. Natl. Acad. Sci. U S A* **94** (1997) 532-6.
- Sala, A., Saitta, B., De Luca, P., Cervellera, M.N., Casella, I., Lewis, R.E., Watson, R. and Peschle, C.: B-MYB transactivates its own promoter through SP1-binding sites. *Oncogene* **18** (1999) 1333-9.
- Sala, A. and Watson, R.: B-Myb protein in cellular proliferation, transcription control, and cancer: latest developments. *J. Cell. Physiol.* **179** (1999) 245-50.
- Salghetti, S.E., Caudy, A.A., Chenoweth, J.G. and Tansey, W.P.: Regulation of transcriptional activation domain function by ubiquitin. *Science* **293** (2001) 1651-3.
- Sambrook, J., Fritsch, E. and Maniatis, T.: Molecular Cloning: A laboratory manual, Second edition ed. Cold Spring Harbor Laboratory Press, New York, 1989.
- Sanz, L., Diaz-Meco, M.T., Nakano, H. and Moscat, J.: The atypical PKC-interacting protein p62 channels NF-kappaB activation by the IL-1-TRAF6 pathway. *Embo J.* **19** (2000) 1576-86.
- Sartorelli, V., Huang, J., Hamamori, Y. and Kedes, L.: Molecular mechanisms of myogenic coactivation by p300: direct interaction with the activation domain of MyoD and with the MADS box of MEF2C. *Mol. Cell. Biol.* **17** (1997) 1010-26.
- Sato, S., Roberts, K., Gambino, G., Cook, A., Kouzarides, T. and Goding, C.R.: CBP/p300 as a co-factor for the Microphthalmia transcription factor. *Oncogene* **14** (1997) 3083-92.
- Saville, M.K. and Watson, R.J.: B-Myb: a key regulator of the cell cycle. *Adv. Cancer. Res.* **72** (1998) 109-40.
- Saville, M.K. and Watson, R.J.: The cell-cycle regulated transcription factor B-Myb is phosphorylated by cyclin A/Cdk2 at sites that enhance its transactivation properties. *Oncogene* **17** (1998) 2679-89.
- Sawado, T., Igarashi, K. and Groudine, M.: Activation of beta-major globin gene transcription is associated with recruitment of NF-E2 to the beta-globin LCR and gene promoter. *Proc Natl Acad Sci U S A* **98** (2001) 10226-31.
- Schubert, S., Horstmann, S., Bartusel, T. and Klempnauer, K.H.: The cooperation of B-Myb with the coactivator p300 is orchestrated by cyclins A and D1. *Oncogene* **23** (2004) 1392-404.

- Shoemaker, B.A., Portman, J.J. and Wolynes, P.G.: Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc. Natl. Acad. Sci. U S A.* **97** (2000) 8868-73.
- Sitzmann, J., Noben-Trauth, K., Kamano, H. and Klempnauer, K.H.: Expression of B-Myb during mouse embryogenesis. *Oncogene* **12** (1996) 1889-94.
- Smith, D.B. and Johnson, K.S.: Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67** (1988) 31-40.
- Soutoglou, E. and Talianidis, I.: Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation. *Science* **295** (2002) 1901-4.
- Sreerama, N., Venyaminov, S.Y. and Woody, R.W.: Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy. *Protein. Sci.* **8** (1999) 370-80.
- Sreerama, N. and Woody, R.W.: Protein secondary structure from circular dichroism spectroscopy. Combining variable selection principle and cluster analysis with neural network, ridge regression and self-consistent methods. *J. Mol. Biol.* **242** (1994) 497-507.
- Sreerama, N. and Woody, R.W.: Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287** (2000) 252-60.
- Stern, J.B. and Smith, K.A.: Interleukin-2 induction of T-cell G1 progression and c-myc expression. *Science* **233** (1986) 203-6.
- Suzuki, A., Yoshida, M., Hayashi, K., Mizuno, Y., Hagiwara, Y. and Ozawa, E.: Molecular organization at the glycoprotein-complex-binding site of dystrophin. Three dystrophin-associated proteins bind directly to the carboxy-terminal portion of dystrophin. *Eur. J. Biochem.* **220** (1994) 283-92.
- Takahashi, T., Nakagoshi, H., Sarai, A., Nomura, N., Yamamoto, T. and Ishii, S.: Human A-myb gene encodes a transcriptional activator containing the negative regulatory domains. *FEBS Lett.* **358** (1995) 89-96.
- Takahashi, Y., Rayman, J.B. and Dynlacht, B.D.: Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression. *Genes Dev.* **14** (2000) 804-16.

- Tanaka, Y., Patestos, N.P., Maekawa, T. and Ishii, S.: B-myb is required for inner cell mass formation at an early stage of development. *J. Biol. Chem.* **274** (1999) 28067-70.
- Tanikawa, J., Yasukawa, T., Enari, M., Ogata, K., Nishimura, Y., Ishii, S. and Sarai, A.: Recognition of specific DNA sequences by the c-myb protooncogene product: role of three repeat units in the DNA-binding domain. *Proc. Natl. Acad. Sci. U S A* **90** (1993) 9320-4.
- Tashiro, S., Takemoto, Y., Handa, H. and Ishii, S.: Cell type-specific trans-activation by the B-myb gene product: requirement of the putative cofactor binding to the C-terminal conserved domain. *Oncogene* **10** (1995) 1699-707.
- Thompson, C.B., Challoner, P.B., Neiman, P.E. and Groudine, M.: Expression of the c-myb proto-oncogene during cellular proliferation. *Nature* **319** (1986) 374-80.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G.: The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25** (1997) 4876-82.
- Todokoro, K., Watson, R.J., Higo, H., Amanuma, H., Kuramochi, S., Yanagisawa, H. and Ikawa, Y.: Down-regulation of c-myb gene expression is a prerequisite for erythropoietin-induced erythroid differentiation. *Proc. Natl. Acad. Sci. U S A* **85** (1988) 8900-4.
- Tompa, P.: Intrinsically unstructured proteins. *Trends Biochem Sci* **27** (2002) 527-33.
- Torchia, J., Rose, D.W., Inostroza, J., Kamei, Y., Westin, S., Glass, C.K. and Rosenfeld, M.G.: The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. *Nature* **387** (1997) 677-84.
- Toscani, A., Mettus, R.V., Coupland, R., Simpkins, H., Litvin, J., Orth, J., Hatton, K.S. and Reddy, E.P.: Arrest of spermatogenesis and defective breast development in mice lacking A-myb. *Nature* **386** (1997) 713-7.
- Trauth, K., Mutschler, B., Jenkins, N.A., Gilbert, D.J., Copeland, N.G. and Klempnauer, K.H.: Mouse A-myb encodes a trans-activator and is expressed in mitotically active cells of the developing central nervous system, adult testis and B lymphocytes. *Embo J* **13** (1994) 5994-6005.
- Trouche, D., Cook, A. and Kouzarides, T.: The CBP co-activator stimulates E2F1/DP1 activity. *Nucleic. Acids. Res.* **24** (1996) 4139-45.
- Trougakos, I.P. and Gonos, E.S.: Clusterin/apolipoprotein J in human aging and cancer. *Int. J. Biochem. Cell. Biol.* **34** (2002) 1430-48.

- Uesugi, M., Nyanguile, O., Lu, H., Levine, A.J. and Verdine, G.L.: Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science* **277** (1997) 1310-3.
- Uversky, V.N., Gillespie, J.R. and Fink, A.L.: Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* **41** (2000) 415-27.
- Uversky, V.N.: Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* **11** (2002) 739-56.
- Wang, H.G., Moran, E. and Yaciuk, P.: E1A promotes association between p300 and pRB in multimeric complexes required for normal biological activity. *J. Virol.* **69** (1995) 7917-24.
- Wassarman, D.A. and Sauer, F.: TAF(II)250: a transcription toolbox. *J. Cell. Sci.* **114** (2001) 2895-902.
- Watson, R.J., Robinson, C. and Lam, E.W.: Transcription regulation by murine B-myb is distinct from that by c-myb. *Nucleic Acids Res.* **21** (1993) 267-72.
- Westin, E.H., Gallo, R.C., Arya, S.K., Eva, A., Souza, L.M., Baluda, M.A., Aaronson, S.A. and Wong-Staal, F.: Differential expression of the amv gene in human hematopoietic cells. *Proc. Natl. Acad. Sci. U S A* **79** (1982) 2194-8.
- Weston, K. and Bishop, J.M.: *Cell* **58** (1989) 85-93.
- Williamson, R.A., Natalia, D., Gee, C.K., Murphy, G., Carr, M.D. and Freedman, R.B.: Chemically and conformationally authentic active domain of human tissue inhibitor of metalloproteinases-2 refolded from bacterial inclusion bodies. *Eur. J. Biochem.* **241** (1996) 476-83.
- Wright, P.E. and Dyson, H.J.: Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293** (1999) 321-31.
- Yang, B.S., Hauser, C.A., Henkel, G., Colman, M.S., Van Beveren, C., Stacey, K.J., Hume, D.A., Maki, R.A. and Ostrowski, M.C.: Ras-mediated phosphorylation of a conserved threonine residue enhances the transactivation activities of c-Ets1 and c-Ets2. *Mol. Cell. Biol.* **16** (1996) 538-47.
- Yao, T.P., Ku, G., Zhou, N., Scully, R. and Livingston, D.M.: The nuclear hormone receptor coactivator SRC-1 is a specific target of p300. *Proc. Natl. Acad. Sci. U S A* **93** (1996) 10626-31.
- Ziebold, U., Bartsch, O., Marais, R., Ferrari, S. and Klempnauer, K.H.: Phosphorylation and activation of B-Myb by cyclin A-Cdk2. *Curr. Biol.* **7** (1997) 253-60.

- Zor, T., De Guzman, R.N., Dyson, H.J. and Wright, P.E.: Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb. *J. Mol. Biol.* **337** (2004) 521-34.
- Zor, T., Mayr, B.M., Dyson, H.J., Montminy, M.R. and Wright, P.E.: Roles of phosphorylation and helix propensity in the binding of the KIX domain of CREB-binding protein by constitutive (c-Myb) and inducible (CREB) activators. *J. Biol. Chem.* **277** (2002) 42241-8.
- Zwicker, J., Liu, N., Engeland, K., Lucibello, F.C. and Muller, R.: Cell cycle regulation of E2F site occupation in vivo. *Science* **271** (1996) 1595-7.
- Zwijnsen, R.M., Buckle, R.S., Hijmans, E.M., Loomans, C.J. and Bernards, R.: Ligand-independent recruitment of steroid receptor coactivators to estrogen receptor by cyclin D1. *Genes Dev.* **12** (1998) 3488-98.
- Zwijnsen, R.M., Wientjens, E., Klompaker, R., van der Sman, J., Bernards, R. and Michalides, R.J.: CDK-independent activation of estrogen receptor by cyclin D1. *Cell* **88** (1997) 405-15.