An investigation of the response of a model class IIIA Adenyly Cyclase to carbon dioxide

Phillip Holliday

Dedication

For my parents, who have supported me throughout my life and given me the opportunity to accomplish this thesis.

Declaration

I confirm that the thesis conforms with the prescribed word length for the degree for which I am submitting it for examination.

I confirm that no part of the material offered has previously been submitted by me for a degree in this or in any other University. If material has been generated through joint work, my independent contribution has been clearly indicated. In all other cases material from the work of others has been acknowledged and quotations and paraphrases suitability indicated.

Acknowledgements

I would like to thank my supervisors, Dr Martin Cann, School of Biological and Biomedical Sciences and Dr David Hodgson, Department of Chemistry for their support and guidance throughout my studentship. I would also like to thank my lab colleagues, Phil Townsend, Dr Arne Hammer and Dr Stepan Fenyk for their help in the lab.

I would like to acknowledge Prof. David Parker and Dr Robert Pal, Department of Chemistry, Durham University for their kind donation of EuL^4 and the use of their Fluorolog 3-11 Spectrophotometer.

I would like to acknowledge Dr Ehmke Pohl and Dr Katja Betz, Department of Chemistry, Durham University, for their help with protein crystallisation and with the X-ray diffraction at Swiss Light Source.

I would like to acknowledge Dr Karl Coleman, Department of Chemistry, Durham University, for the use of his FTIR Spectrophotometer.

I would like to acknowledge DBS Genomics, Durham University, for sequencing DNA.

1 Abstract

Adenylyl cyclase catalyses the conversion of adenosine triphosphate into 3',5'-cyclic adenosine monophosphate and pyrophosphate. Adenylyl cyclases are grouped into six distinct Classes based on amino acid sequence similarity. Mammalian adenylyl cyclases and numerous prokaryotic adenylyl cyclases are grouped into Class III. Class III is further sub-divided into four sub-Classes, IIIa, IIIb, IIIc, and IIId, based upon amino acid polymorphisms within the active site. Class IIIa adenylyl cyclases include the mammalian G-protein regulated transmembrane adenylyl cyclases and numerous prokaryotic adenylyl cyclases such as Rv1625c from *Mycobacterium tuberculosis*. Class IIIb adenylyl cyclases include the mammalian soluble adenylyl cyclase and numerous prokaryotic adenylyl cyclase such as CyaB1 from *Anabaena* PCC 7120. Class IIIb adenylyl cyclases are stimulated by inorganic carbon. New findings dispute whether the inorganic carbon stimulation of Class IIIb adenylyl cyclases is by HCO_3^- or CO_2. Class IIIa adenylyl cyclases have previously been demonstrated to be non-responsive to HCO_3^- but have not been investigated with CO_2.

Experiments were performed using the prokaryotic Class IIIa adenylyl cyclase $Rv1625c_{204-443}$. The specific activity of $Rv1625c_{204-443}$ was obtained by measuring the conversion of [α-^{32}P]ATP into [^{32}P]cAMP. $Rv1625c_{204-443}$ was assayed for pH dependence, cation dependence, dose dependence, enzyme kinetics, and inorganic carbon-activating species in the presence or absence of inorganic carbon. The results showed that *in vitro* $Rv1625c_{204-443}$ is stimulated by CO_2 and not HCO_3^- below pH 7.5 with an apparent $E.C._{.50}$ value of 13.2 ± 0.6 mM. Experiments were performed to investigate *in vivo* CO_2 stimulation of $Rv1625c_{204-443}$ transformed into *Escherichia coli* cells. The results showed that *in vivo* $Rv1625c_{204-443}$ is stimulated by 10 % (v/v) CO_2 in air.

CO_2 stimulation of $Rv1625c_{204-443}$ was further investigated by identifying the conditions required for CO_2 to bind to $Rv1625c_{204-443}$. Numerous spectroscopic and biochemical techniques were used to identify conditions required for CO_2 binding such as circular dichroism spectroscopy, fourier transform infrared spectroscopy, HCO_3^- dependent luminescence probe

spectroscopy, CO_2 stimulation assays on $Rv1625c_{204-443}$ mutant proteins, and $^{14}CO_2$ binding assays. The results showed that metal is not required for CO_2 binding. CO_2 binds solely to the apoprotein.

As CO_2 does not require any cofactors or substrate to bind to $Rv1625c_{204-443}$, the elucidation of the CO_2 binding site on $Rv1625c_{204-443}$ was attempted. Numerous binding sites were investigated, such as amino acids in the active site with a primary or secondary amine side group (forming a potential carbamate with CO_2). Double and triple amino acid CO_2 binding sites were also investigated by identifying the amino acids that correspond to the HCO_3^- binding site from mammalian soluble adenylyl cyclase. CO_2 binding was also investigated by growing Rv1625c crystals, exposing the crystals to CO_2 and X-ray diffraction studies on the crystals to obtain a diffraction data set. Unfortunately none of these techniques was able to identify the CO_2 binding site of Rv1625c.

This surprising and novel finding that CO_2 binds and stimulates $Rv1625c_{204-443}$ suggests that other Class IIIa adenylyl cyclases, including the heterotrimeric G-protein stimulated mammalian adenylyl cyclases are also stimulated by CO_2 and further work needs to be done in this area.

1.1 Table of Contents

1.2 List of Tables

1.3 List of Figures

1.4 Abbreviations

2-HE	2-hydroxyestradiol
2-ME	2-mercaptoethanol
AC	adenylyl cyclase
ADP	adenosine diphosphate
AMPCPP	α,β-methyleneadenosine 5'-triphosphate
Arg	arginine
Asn	asparagine
Asp	aspartate
ATP	adenosine triphosphate
BSA	bovine serum albumin
CA	carbonic anhydrase
CaM	calmodulin
cAMP	3',5'-cyclic adenosine monophosphate
CB	carotid body
CBF	cilia beat frequency
CD	circular dichroism
CE	catechol oestrogen
CFTR	cystic fibrosis transmembrane conductance regulator
cGMP	3',5'-cyclic guanosine monophosphate
CHAPS	3[(3-Cholamidopropyl)dimethylammonio]-propanesulfonic acid
CHES	2-(cyclohexyamino)ethanesulfonic acid
Ci	inorganic carbon
CIAP	calf intestine alkaline phosphatase
CREB	cAMP-response element binding protein
CRP	cAMP receptor protein
CT	cholera toxin
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EEA1	early endosomal antigen 1
EF	edema factor
EPAC	exchange proteins activated by cAMP
ES	electrospray
FTIR	fourier transform infrared
GAF	cGMP-phosphodiesterase, adenylyl cyclases and formate hydrogen lyase transcriptional factor
GALT	gut-associated lymphoid tissue
GC	guanylyl cyclase
GDP	guanosine diphosphate
Gln	glutamine
GPCR	G protein-coupled receptor
GTP	guanosine triphosphate
HAMP	histidine kinase, adenyl cyclase, methyl accepting chemotaxis proteins and phosphatases
HEK	human embryonic kidney cells
HEPES	N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
His	histidine
HPLC	high performance liquid chromatography

HTH	helix-turn-helix
IPTG	isopropyl-β-D-thiogalactopyranoside
LAMP1	lysosomal membrane protein 1
LB	luria broth
LF	lethal factor
Lys	lysine
MANT	N-methylanthraniloyl
MES	2-morpholinoethanesufonic acid
mRNA	messenger ribonucleic acid
MS	mass spectroscopy
MTSCE	2-carboxyethyl methanethiosulfonate
Ni-NTA	nickel-nitrilotriacetic acid
NGF	nerve growth factor
ONP	ortho-nitrophenol
ONPG	ortho-nitrophenol-β-D-galactopyranoside
PA	protective antigen
PAGE	polyacrylamide gel electrophoresis
PAS	period clock protein, aryl hydrocarbon receptor and single-minded protein
pCO_2	partial pressure of carbon dioxide
PCR	polymerase chain reaction
PDE	phosphodiesterase
PEG	polyethylene glycol
PKA	cAMP-dependent protein kinase A
PKC	cAMP-dependent protein kinase C
PMSF	phenylmethylsulfonyl fluoride
poly-P	polyphosphates
PPi	pyrophosphates
PTS	phosphotransferase transport system
Rubisco	ribulose-1,5-bisphosphate carboxylase/oxygenase
sAC	soluble adenylyl cyclase
SDS	sodium dodecyl sulphate
Ser	serine
SOC	super optimal catabolite repression
TAE	tris-acetate-EDTA
TCP	toxin-co-regulated pilus
Thr	threonine
tmAC	transmembrane adenylyl cyclase
TOF	time of flight
Tris	tris(hydroxymethyl)aminomethane
TTSS	type III secretion system
Val	valine
V-ATPase	vacuolar H+ ATPase

2 Introduction

2.1 cAMP as an intracellular messenger

cAMP (3',5'-cyclic adenosine monophosphate) was first discovered when dog liver homogenates were incubated with ATP (adenosine triphosphate) in the presence of adrenaline and glucagon (Rall and Sutherland, 1958). Mg^{2+} ions were also found to be required for the formation of cAMP (Sutherland and Rall, 1958). Other particulate preparations from heart, skeletal muscle, and brain also formed cAMP under the same conditions (Sutherland and Rall, 1958). The enzyme responsible for the conversion of ATP into cAMP was termed adenyl cyclase (now known as adenylyl cyclase (AC) or adenylate cyclase) and was found in all animal tissues sampled (Sutherland *et al.*, 1962). Increased concentrations of cAMP in the cell diffuse through the cytosol and bind to cAMP-dependent protein kinase A (PKA) (Walsh *et al.*, 1968), Exchange Proteins Activated by cAMP (EPACs) (de Rooij *et al.*, 1998; Kawasaki *et al.*, 1998), cyclic nucleotide gated ion channels (Kaupp *et al.*, 1989; Goulding *et al.*, 1992), and the GAF (cGMP-phosphodiesterase, adenylyl cyclases and formate hydrogen lyase transcriptional factor) domains of phosphodiesterase (PDE) 10 (Gross-Langenhoff *et al.*, 2006). Activation of PKA is important for the regulation of metabolism and gene expression. PKA activates phosphorylase kinase which in turn activates phosphorylase thus increasing the breakdown of glycogen into glucose. PKA also controls gene expression by activating transcription factors such as CRE (cAMP-response element)-binding protein (CREB), which binds to consensus CRE sequences in DNA (deoxyribonucleic acid) and initiates gene transcription. This effect is hormone and tissue specific, for example adrenaline induces hydrolysis of triglycerides in adipose tissues but increases the conversion of glycogen to glucose in the liver (Sutherland, 1972).

In prokaryotic organisms, cAMP was first discovered in *Escherichia coli* where limitation of carbon and energy sources led to an increase in cAMP levels (Makman and Sutherland, 1965). *E. coli* cells grown on glucose produced low levels of cAMP, however when the glucose was exhausted cAMP levels increased and then dropped again on addition of more glucose. This together with the finding that glucose repressed the synthesis of β-

galactosidase, suggested that glucose may repress synthesis via a cAMP-dependent pathway. This hypothesis was shown to be correct by adding cAMP to *E. coli* cultures. cAMP overcame the inhibitory effects of glucose and produced β-galactosidase (Perlman and Pastan, 1968). The physiological role of cAMP was confirmed by analysis of AC deficient (*cya*) *E. coli* (Ide, 1969; Tao and Lipmann, 1969) that were unable to utilise lactose or galactose as a substrate or numerous other sugars that were also subject to glucose repression (Pastan and Perlman, 1969). The addition of cAMP restored the ability of the *E. coli cya* mutants to utilise these sugars and synthesise β-galactosidase. A cAMP receptor protein (CRP), also known as catabolite activator protein (Zubay *et al.*, 1970) and catabolite gene activator (Riggs *et al.*, 1971) was identified by the addition of cAMP to *cya E. coli* mutants that were not able to synthesise β-galactosidase. On addition of the purified CRP to the cAMP-unresponsive *cya E. coli* mutants, β-galactosidase synthesis was recovered demonstrating that CRP is required for cAMP activity (Emmer *et al.*, 1970).

2.2 Classification of ACs

ACs have an important role in amplifying primary signals. When a primary signal, such as a hormone e.g. epinephrine, binds to its receptor on the cell surface, the signal is passed to intracellular targets, such as protein kinases, ion channels or transcription factors, via ACs. Thus a cellular response is produced from the initial primary signal. ACs encompass a large group of proteins that may be the product of convergent evolution and have been categorised into six distinct classes based on their amino acid sequence similarity (Barzu and Danchin, 1994; Cotta *et al.*, 1998; Sismeiro *et al.*, 1998; Tellez-Sosa *et al.*, 2002).

Examples of some ACs from the six classes will be discussed below in the context of the organism that the ACs are found in and their function within the organism.

2.2.1 Class I ACs

Class I ACs are found in enterobacterial, gram-negative bacteria such as *Salmonella typhimurium* (Carter and Collins, 1974), *Vibrio cholerae* (Mekalanos *et al.*, 1983), *Vibrio vulnificus* (Tacket *et al.*, 1984), *Yersinia enterocolitica* (Cornelis *et al.*, 1998), and *E. coli* (Linder, 2008). They are involved in either catabolite repression or the development of virulence through the regulation of virulence protein production.

2.2.1.1 *S. typhimurium*

S. typhimurium, is the causative agent of food-borne infectious gastroenteritis in humans and infects the gut-associated lymphoid tissue (GALT) (Carter and Collins, 1974). The *Salmonella* pathogenicity island 1 (SPI1) encodes a TTSS that transports bacterial effecter proteins into the host cell eventually leading to internalisation of the bacterium into the cell (Hueck *et al.*, 1995). Construction of *cya* and *crp* mutants produced strains that still attached and invaded GALT but were otherwise avirulent, thus identifying cAMP as a regulator of pathogenesis (Curtiss and Kelly, 1987).

2.2.1.2 *V. cholerae*

The human pathogen that causes severe diarrhoea, *V. cholerae*, secretes a multimeric cholera toxin (CT) protein composed of one A subunit and 5 B subunits (Mekalanos *et al.*, 1983). Toxin-co-regulated pilus (TCP), which aids intestinal colonisation, is co-ordinately expressed with CT (Taylor *et al.*, 1987) and the transmembrane DNA-binding protein (ToxR) positively regulates expression of CT, TCP, and accessory colonisation factors (Peterson and Mekalanos, 1988). Insertion mutation of the *cya* and *crp* genes increased the expression of the CT promoter and thus negatively regulated the expression of ToxR-regulated virulence genes and pathogenesis (Skorupski and Taylor, 1997).

2.2.1.3 *V. vulnificus*

The pathogenic marine bacterium *V. vulnificus* is the causative agent of gastroenteritis and septicaemia in immune compromised patients (Tacket *et al.*, 1984). The organism's virulence factor haemolysin is encoded by a gene, *vvh* that has a cAMP-CRP binding promoter sequence. The addition of cAMP increased haemolysin production and the presence of glucose inhibited production. Mutants deleted for the *cya* AC gene of *V. vulnificus* displayed a significantly lower virulence in mice and a decrease in haemolysin and protease production (Kim *et al.*, 2005).

2.2.1.4 *Y. enterocolitica*

The gram-negative pathogen *Y. enterocolitica* causes gastrointestinal illnesses in mammals through a cAMP-dependent virulence pathway. Production of *cya* and *crp* mutants, which encode Cya and the CRP respectively, affected the production of virulence proteins that are transported by all three hypothesised pathogenic pathways; the Yse (Cornelis *et al.*, 1998), Ysa (Haller *et al.*, 2000), and flagellar type III secretion system (TTSS) (Young *et al.*, 1999). Complementation of the mutants with wild type genes restored protein production of each pathogenic pathway (Petersen and Young, 2002) suggesting that Cya is required for formation and transport of all virulence proteins.

2.2.1.5 *E. coli*

Growth of *E. coli* in lactose instead of glucose leads to the induction of the phosphoenolpyruvate-sugar phosphotransferase transport system (PTS) proteins which in turn activate Cya, a Class I AC, to produce cAMP (Linder, 2008). cAMP binds to the CRP which subsequently binds to and activates a catabolic promoter at a specific recognition sequence of a gene for degradation of lactose. The catalytic domain of the *E. coli* Class I AC was over expressed and purified and it displayed a V_{max} of 665 nmol mg^{-1} min^{-1} and a K_m

of 270 μM (Linder, 2008). Twelve amino acids were identified by mutagenesis as essential for catalysis, with five residues clustered in a region of the gene that is conserved in all proteins of the DNA polymerase β-like nucleotidyltransferase superfamily. From these five residues, D114 and D116 were proposed to be involved in metal co-factor binding and S103 was proposed to be involved in substrate binding.

2.2.2 Class II ACs

Class II ACs are produced by pathogenic organisms including *Bacillus anthracis* (Baillie and Read, 2001), *Bordetella pertussis* (Ladant and Ullmann, 1999), and *Pseudomonas aeruginosa* (Yahr *et al.*, 1998). All cause infections in immune compromised patients such as cancer patients, HIV^+ patients, and cystic fibrosis sufferers (Garau and Gomez, 2003). These phylogenetically unrelated bacteria all secrete toxins into the host cell using AC as part of the mechanism.

2.2.2.1 *B. anthracis*

The anthrax toxin of *B. anthracis* is composed of three proteins, protective antigen (PA), lethal factor (LF) and edema factor (EF), which form protein pairs that infect host cells to cause death (PA and LF) or edema (PA and EF) in the skin (Leppla, 1982). EF is a calmodulin (CaM) dependent AC (Guo *et al.*, 2004) that is produced in an inactive form,however it is activated by contact with the host cell. EF is composed of two functional domains, an N-terminal 30 kDa domain that enables entry into cells by binding the anthrax protective antigen (Elliott *et al.*, 2000) and a C-terminal 58 kDa AC that shares sequence homology to CyaA from *B. pertussis* and ExoY from *P. aeruginosa* (Shen *et al.*, 2002). EF has a K_m for ATP of 0.2-1 mM and a catalytic efficiency (k_{cat}/K_m) of approximately 10^7 $M^{-1}s^{-1}$, around 100-fold higher than in mammalian ACs (Guo *et al.*, 2004).

2.2.2.2 *B. pertussis*

In the case of *B. pertussis*, CyaA, a 188 kDa AC toxin translocates across the plasma membrane into the host cell by a two step internalisation process, where CyaA initially binds to the cell surface and then internalises in a temperature and Ca^{2+} dependent manner (Mock and Ullmann, 1993). Once CyaA enters the host cell's cytoplasm, it binds to CaM and produces high concentrations of cAMP, which causes cell apoptosis (Khelef *et al.*, 1993)

CyaA from *B. pertussis* is a CaM activated AC that is able to translocate across the plasma membrane of the infected cell through a process called internalisation. CyaA is constructed of an N-terminal AC domain and a C-terminal domain involved in haemolytic activity and delivery of the AC domain into the targeted cell (Bertini *et al.*, 2004). The AC domain has a high affinity for CaM (0.2 nM) which stimulates activity (k_{cat} = 2000 s^{-1}) by over 1000-fold (Ladant and Ullmann, 1999).

2.2.2.3 *P. aeruginosa*

P. aeruginosa infects through the TTSS which regulates numerous effecter proteins from the exoenzyme S regulon, such as ExoS, ExoT, ExoU, and ExoY, in response to host environmental factors (Yahr *et al.*, 1998). ExoS and ExoT are ADP (adenosine diphosphate)-ribosyltransferases, ExoU is an acute cytotoxin, and ExoY is a Class II AC (Wolfgang *et al.*, 2003). On contact with a host epithelial cell, the TTSS is activated by stimulation of a membrane-localised Class III AC, CyaB (Smith *et al.*, 2004). cAMP subsequently binds to a transcriptional regulator, Vfr, which positively regulates the expression of ExoY and the other TTSS component proteins for a secretory apparatus and the production of toxins that alter host signal transduction pathways, leading to cell death. The TTSS also activates the expression of genes for the production of pili and flagella (Wolfgang *et al.*, 2003), which are important for motility and attachment to host cells.

2.2.3 Class III ACs

Class III is the predominant class of AC found in eukaryotes and many prokaryotes, which gives rise to the class being called the Universal Class. The catalytic domains of Class III ACs form dimers, which have six canonical amino acids that are essential for catalysis. Two aspartate (Asp) residues bind the two metal cofactors; a lysine (Lys) and Asp pair selectively bind ATP; and an arginine (Arg) and asparagine (Asn) pair stabilise the transition state (reviewed by Linder and Schultz, 2003). Further details of the Class III AC catalytic domain will be discussed in Section 2.3. Most Class III ACs are multi-domain proteins that demonstrate a large variety in protein composition and therefore display a large variety of regulatory signals (Linder, 2006). Class III ACs can be subdivided into four subclasses based on the divergence of amino acid sequence of the catalytic domain, particularly among the six canonical residues. These subclasses have been designated Class IIIa, IIIb, IIIc and IIId and their distinguishing features are listed in Table 2.1.

2.2.4 Class IV, V and VI ACs

The remaining three classes of ACs have only one AC member characterised in each class and little information is known about these Classes at present. The Class IV AC was identified in the anaerobic, gram-negative bacterium, *Aeromonas hydrophila*. *A. hydrophila* is an opportunistic pathogen that can cause gastroenteritis, wound infections, septicaemia, and meningitis in humans (Janda, 1991). *A. hydrophila* contains 2 AC genes, *cyaA* encoding a Class I AC and *cyaB* encoding AC2, which is the Class IV AC that has no sequence similarity to any of the other class of ACs (Sismeiro *et al.*, 1998). AC2 has an optimum activity at a high temperature (65 °C) and at an alkaline pH (9.5); and a significant sequence similarity to proteins of *Methanobacterium thermoautotrophicum*, *Archaeglobus fulgidus* and *Methanococcus jannaschii*, all of which are hyperthermophilic archaebacteria.

Table 2.1: Distinguishing features of the subclasses of Class III ACs (Linder and Schultz, 2003).

Class IIIa AC	Signature motifs of (F/Y)XX(F/Y)D participating in formation of dimer interface and EKI**K** containing substrate-defining **K**. Arm region (essential feature for dimerisation (Tesmer *et al.*, 1997)) is 14 residues long.
Class IIIb AC	Distinguished by replacement of substrate-defining Asp with Thr/Ser and frequently the replacement of a phosphate-binding Arg with Gly/Ser. Arm region extended to 15 residues.
Class IIIc AC	Arm region only 7 to 11 residues long.
Class IIId AC	Signature motif of YEV**K**T around substrate-defining **K**. Arm region is 14 residues long.

The only Class V AC found so far is from the ruminal anaerobic bacterium *Prevotella ruminicola* (Cotta *et al.*, 1998) and encodes a 67 kDa protein that has no sequence similarity with any other protein in the GenBank database.

The Class VI AC was identified in *Rhizobium etli* by phenotypic complementation of an *E. coli cya* mutant strain with a genomic library by restoring cAMP accumulation in the *E. coli* mutant (Tellez-Sosa *et al.*, 2002). The protein product of the *cyaC* gene does not have any sequence similarity with any other known protein.

2.3 Mammalian Class III ACs

All mammalian ACs are from Class III and are represented by 10 AC isoforms encoded by distinct genes. Nine AC isoforms are from Class IIIa and are membrane-associated owing to the presence of two sets of six transmembrane domains and are termed tmACs. These ACs are hormone regulated via heterotrimeric G proteins and other activators which will be discussed later (Section 2.3.2). The other mammalian AC isoform is a soluble AC (sAC) from Class IIIb and it is more closely related to bacterial ACs than mammalian tmACs (Buck *et al.* 1999). tmACs and sAC differ structurally, molecularly, and biochemically and therefore define two separate cAMP pathways in mammalian cells (Kamenetsky *et al.*, 2006).

2.3.1 Structure of the Class IIIa ACs

Before the molecular identification of the 10 isoforms of mammalian ACs, only 3 types of ACs were biochemically characterised, a brain specific AC that is activated by $G_s\alpha$ and Ca^{2+}-CaM, a ubiquitous $G_s\alpha$ activated AC, and sAC (Sunahara *et al.*, 1996). Purification of the brain specific AC (that is activated by $G_s\alpha$ and Ca^{2+}-CaM) using forskolin affinity chromatography enabled the generation of probes to isolate cDNA for the AC which was termed type 1 (Krupinski *et al.*, 1989). Molecular cloning techniques identified 8 other

AC cDNAs (types 2-9). Type 2 AC cDNA was cloned from rat brain (Feinstein *et al.*, 1991), type 3 AC cDNA was cloned from rat olfactory tissue (Bakalyar and Reed, 1990), type 4 AC cDNA was cloned from rat testes (Gao and Gilman, 1991), type 5 AC cDNA was cloned from dog heart (Ishikawa *et al.*, 1992) and rat liver and kidney (Premont *et al.*, 1992), type 6 AC cDNA was cloned from dog heart (Katsushika *et al.*, 1992), rat heart, liver and kidney (Premont *et al.*, 1992), mouse S49 lymphoma cells (Premont *et al.*, 1992) and mouse/hamster hybrid NCB-20 cells (Yoshimura and Cooper, 1992), type 7 AC cDNA from mouse and S49 lymphoma cells (Watson *et al.*, 1994), type 8 AC cDNA from rat brain tissue (Cali *et al.*, 1994) and type 9 AC cDNA from mouse brain tissue (Premont *et al.*, 1996). The final isoform to be identified was sAC cDNA from rat testis (Buck *et al.*, 1999).

The tissue specificity of ACs has been elucidated by mRNA (messenger ribonucleic acid) studies, which have shown a large diversity of expression for the different isoforms, from being widely expressed, to only being found in specific tissue. Mammalian ACs are generally found in excitable tissue such as neurons and muscle, although they are not limited to those tissues. AC1 is found in distinct regions of the brain such as the hippocampus and cerebral cortex (Xia *et al.*, 1991) and the adrenal medulla with its possible function linked to learning and memory (Shen *et al.*, 1997). AC2 is found in the brain, heart, lungs and skeletal muscle (Suzuki *et al.*, 1998). The highest expression of AC3 is in the olfactory neuroepithelium (Bakalyar and Reed, 1990) of the brain. AC4 and AC5 are both expressed in the brain, heart, kidneys, the glomerulus of the liver (Chabardes *et al.*, 1996), lungs, uterus and brown adipose tissue (Edelhoff *et al.*, 1995), while the adrenal gland and testes express AC5 solely. AC6 and AC7 are both ubiquitous (Haber *et al.*, 1994) and are therefore found in most tissues, with AC7 being highly expressed in numerous parts of the brain where it could be involved in drug dependency (Hellevuo *et al.*, 1995). In the heart, AC6 is expressed in non-myocyte cells whereas AC5 is expressed in myocyte cells (Yu *et al.*, 1995).

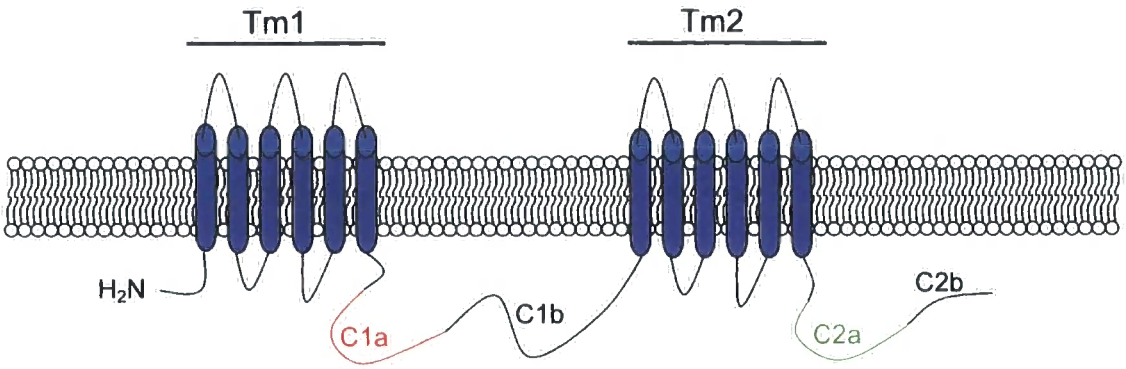
AC8 is found in the hypothalamus (Matsuoka *et al.*, 1992) with possible functions in learning and memory and it is also expressed in the heart, lungs, uterus and testes (Stengel *et al.*, 1992). AC9 is found in the brain and skeletal muscle (Hacker *et al.*, 1998).

Mammalian Class IIIa ACs contain 12 transmembrane spanning segments that can be divided into two sets of six transmembrane domains, Tm1 and Tm2. Adjacent to each transmembrane domain is a cytoplasmic domain. The two cytoplasmic domains are similar but not identical and are termed C1 and C2 and dimerise to form a pseudosymmetric heterodimer (Figure 2.1). Both cytoplasmic domains can be further characterised into two segments, C1a/C2a and C1b/C2b. C1a and C2a are highly conserved sequences that are homologous to each other and contain the catalytic residues, whereas C1b and C2b are less conserved and probably contain regulatory sites.

The cytoplasmic domains have six specific residues that are conserved in most Class III ACs. These residues are important for binding a metal cofactor or nucleotide substrate and also for the stabilisation of the catalytic transition state. This information was elucidated from sequence and mutational analyses and crystal structure analyses. The C1a domain contains two Asp residues that coordinate the first metal cofactor A, which *in vivo* is usually a Mg^{2+} ion. The catalytic mechanism is similar to that of DNA polymerase (Zimmermann *et al.*, 1998), where the hydrated metal ion A lowers the pK_a value of the hydrogen of the ribose ring 3'-OH group so that deprotonation occurs more easily. The resulting 3'-O^- nucleophilically attacks the α-phosphoryl group of ATP, eliminating PPi (pyrophosphate) (Tesmer *et al.*, 1999).

The elimination of PPi is aided by the other hydrated metal cofactor B, which lowers the pKa values of the β- and the γ-phosphoryl groups of ATP, making them a good leaving group. The C2a domain contains a Lys and Asp residue pair that gives the AC its substrate selectivity towards ATP and not to other nucleotides. The C2a domain also has an Arg and Asn pair that stabilise the transition state of the reaction (Tesmer *et al.*, 1999).

Figure 2.1: Schematic topology of the pseudoheterodimeric Class IIIa mammalian AC. The AC is subdivided into the first six transmembrane domain Tm1 (blue cylinders), the first cytoplasmic domain consisting of C1a (red) and C1b (black), the second six transmembrane domain Tm2 (blue cylinders), the second cytoplasmic domain consisting of C2a (green) and C2b (black).


Tm1
Tm2
H_2N
C1a
C1b
C2a
C2b

2.3.2 Activators of mammalian Class IIIa ACs

Mammalian tmACs are found adjacent to G protein-coupled receptors (GPCR), which are seven transmembrane receptors that mediate signal transduction through the dissociation of heterotrimeric G proteins. The binding of an extracellular signal such as a hormone or neurotransmitter to the receptor of a GPCR causes a conformational change that activates an adjacent heterotrimeric G protein through the exchanging of guanosine diphosphate (GDP) for guanosine triphosphate (GTP) on the α-subunit. This activation causes the heterotrimeric G protein to breakdown into a Gα subunit and a Gβγ subunit pair. The Gα subunit activates the adjacent tmAC in either a stimulatory ($G\alpha_s$) or inhibitory ($G\alpha_i$) fashion. The $G\alpha_s$ subunit binds between the C1 and C2 domains of AC, at the perimeter of the catalytic centre and aids the closing of the catalytic centre around the substrate (Sunahara *et al.*, 1997).

Other G protein subunits, such as Gβγ or $G\alpha_i$ selectively regulate some of the mammalian AC isoforms by either stimulating or inhibiting their activity (Sprang, 1997). Gβγ subunits inhibit AC1 and AC8 but stimulate AC2, AC4 and AC7 when $G\alpha_s$ is present. It has been suggested that Gβγ could be increasing the ability of $G\alpha_s$ to activate AC in AC2, AC4 and AC7 (Dessauer *et al.*, 1998). $G\alpha_i$ subunits inhibit AC5 and AC6 but not through competition with $G\alpha_s$ subunits as they bind to a pseudosymmetrical site to the $G\alpha_s$ binding site on the C1 domain in AC5 (Dessauer *et al.*, 1998). This binding of the $G\alpha_i$ subunit maintains the open form of the catalytic centre even in the presence of ATP.

Forskolin, a diterpene from the roots of *Coleus forskohlii* (Bhat *et al.*, 1977) was initially discovered to stimulate mammalian ACs in rat brain (Seamon *et al.*, 1981). It stimulates all mammalian AC isoforms except sAC, although AC9 is only weakly stimulated (Premont *et al.*, 1996). Forskolin binds to the non-catalytic pocket of the catalytic centre of the pseudoheterodimeric AC (Zhang *et al.*, 1997).

Ca^{2+} can also regulate AC activity. It has also been shown that CaM affects AC activity (Cali *et al.*, 1994). Ca^{2+} ions activate isoforms AC1, AC3 and AC8, with activation via CaM in AC1 and AC3 (Fagan *et al.*, 1996). The

$E.C._{50}$ value of free Ca^{2+} concentrations for stimulation of AC1 is 0.05 μM, well within physiological ranges (Choi *et al.*, 1992). Ca^{2+} ions can also inhibit the activity of all ACs when the concentration is in the millimolar range, however even at physiological concentrations, isoforms AC5 and AC6 are inhibited (Guillou *et al.*, 1999). These isoforms are localised in the brain and in excitable cell types, which may explain their sensitivity to Ca^{2+} concentrations.

cAMP signalling in mammals can also be regulated through desensitisation of G protein-coupled receptor activation and also through desensitisation of the AC. AC6 is desensitised through direct phosphorylation of S674 on the C1b region by PKA or PKC (cAMP-dependent protein kinase C) which results in a 50 % decrease in AC activity (Chen *et al.*, 1997). CO_2 has also been demonstrated to stimulate mammalian G-protein regulated ACs *in vitro*. Inorganic carbon (Ci) assays on $7C_1.2C_2$ identified that CO_2 and HCO_3^- was stimulating $7C_1.2C_2$, with a maximum 2-3-fold stimulation at 7 mM CO_2 (Townsend *et al.*, 2009). $7C_1.2C_2$ was also demonstrated to be CO_2 responsive *in vivo* in HEK 293T cells, where exposing the cells to 5 % (v/v) CO_2 whilst stimulating the recombinant AC with isoproterenol, a β-adrenergic receptor agonist, increased the accumulation of cAMP (Townsend *et al.*, 2009).

2.3.3 Inhibitors of mammalian Class IIIa ACs

The identification of a specific AC inhibitor has numerous advantages, such as identifying the processes that are involved with cAMP at the cellular level and creating inhibitors for toxic ACs such as EF produced by *B. anthracis* and CyaA by *B. pertussis*. Inhibitor specificity for AC is a problem that is gradually being overcome with the development of new AC inhibitors, which are discussed below.

Substrate analogues and the majority of nucleotide inhibitors of Class IIIa ACs are derived from adenosine (Johnson and Shoshani, 1990). Substrate analogues, such as α,β-methylene-ATP and R_p-ATPαS that inhibit tmACs, competitively bind to the unliganded enzyme (Dessauer *et al.*, 1999). All mammalian isoforms including sAC are inhibited by P-site inhibitors, which

non-competitively inhibit by binding to the enzyme when it is in a conformation that resembles the product-bound state (Dessauer and Gilman, 1997). P-site inhibitors are compounds that contain an intact purine ring, bind to the active site of an AC and inhibit the enzyme in either a non-competitive manner, usually in the presence of Mn^{2+} ions (Johnson *et al.*, 1979) or an un-competitive manner, usually in the presence of Mg^{2+} ions (Welton and Simko, 1980). The concentration of P-site inhibitor has no effect on substrate binding. The binding of a P-site inhibitor is thought to form a dead-end complex (Florio and Ross, 1983) where the inhibitor binds to the AC with PPi still bound to the AC. Thus the inhibitor binds to the enzyme-substrate complex or the enzyme-product complex and not to the free enzyme which explains why the apparent potency of the P-site inhibitor seems to increase when the AC is activated. P-site inhibitors bind to a small hydrophobic region at the C1a-C2a border near the 2 canonical Asp residues where it causes a conformational change of the active site of the AC from an open to a closed state by rotating residues on the C1 region towards the P-site binding residues on the C2 region (Dessauer *et al.*, 1999).

Another inhibitor of mammalian tmACs is Zn^{2+} which inhibits hormone and forskolin stimulation of tmACs in a time, dose and temperature dependent manner (Klein *et al.*, 2002). Zn^{2+} inhibits by decreasing the V_{max} of the AC and has an $I.C._{50}$ of 1-2 μM. Zn^{2+} inhibition was shown not to compete with Mg^{2+} binding to the active site, thus there is a distinct Zn^{2+} binding site (Klein *et al.*, 2002).

2.3.4 Mammalian soluble Class IIIb AC

sAC was initially identified in cytosolic extracts of rat testes and was proposed to be involved in spermatogenesis. Unlike mammalian tmACs, sAC preferentially utilises Mn^{2+} over Mg^{2+} for biochemical activity (Braun and Dods, 1975) and is insensitive to the tmAC activators G protein (Braun *et al.*, 1977) and forskolin (Forte, 1983). Therefore in mammalian cells, sAC is an independent system to tmACs giving two biochemically distinct cAMP pathways. A functional cDNA form of sAC was isolated (Buck *et al.*, 1999) and

gave a predicted full length protein of 187 kDa, which contained two putative catalytic domains, C1 and C2, a consensus P-loop, and a leucine zipper sequence. The C1 and C2 catalytic domains of sAC dimerise to form a pseudoheterodimer. The full length sAC mRNA can be alternatively spliced to either include or exclude exon 11 (Jaiswal and Conti, 2001). mRNA including exon 11 encoded the 187 kDa full length sAC and mRNA excluding exon 11 encoded a 48 kDa truncated sAC molecule. The 48 kDa truncated protein has more similarity to the catalytic domains of prokaryotic Class IIIb ACs than to the catalytic domains of other isoforms of mammalian ACs. This truncated version of sAC has an activity 20 times greater than the full length protein, as there is an autoinhibitory site on the full length protein that lowers V_{max} of the enzyme without affecting K_m (Chaloupka *et al.*, 2006).

sAC is activated by numerous factors. During experiments investigating the effects of Ca^{2+} on sperm capacitation and the acrosome reaction, Hyne and Garbers showed that Ca^{2+} increased concentrations of cAMP in guinea pig spermatozoa (Hyne and Garbers, 1979). This Ca^{2+}-induced increase in cAMP concentrations was further shown to be dependent on the presence of HCO_3^- (Garbers *et al.*, 1982). HCO_3^- was suggested not to bind directly to sAC due to there being no stimulation of AC by HCO_3^- in broken cell preparations. It was proposed that HCO_3^- aids in the transport of Ca^{2+} into the spermatozoa and that it is the increase in Ca^{2+} concentrations that gives rise to the stimulation of sAC.

However Chen *et al.* showed that mammalian sAC is activated by HCO_3^- without the addition of any other factors (Chen *et al.*, 2000). HCO_3^- binds directly to sAC and activates it in a pH-independent manner. HCO_3^- was also shown to stimulate prokaryotic Class IIIb ACs from multiple prokaryotic species. It was hypothesised that the characteristic Class IIIb AC threonine (Thr) polymorph of the substrate-defining Asp and a Lys residue in the active site coordinate the binding of HCO_3^- in Class IIIb ACs (Cann *et al.*, 2003).

An alternative hypothesis for the site of HCO_3^- binding on AC was on the metal ion of the active site, displacing two H_2O molecules (Steegborn *et al.*, 2005). HCO_3^- is postulated to close the active site by inducing a 4-5 Å movement of the β7-β8 loop and a shift of the α1 helix that aids the binding of

the catalytic metal ion. HCO_3^- does not stimulate other Class IIIa ACs by this theory due to the increased negative charge caused by the substrate-defining Asp residue near the metal ion. Further work on proving that sAC is activated by HCO_3^- showed that sAC is also responsive to CO_2 (Townsend *et al.*, 2009). Ci activated specific activity of sAC was greatest at pH 8.5, suggesting that HCO_3^- is an activator, however relative Ci fold stimulation of sAC was greatest at pH 6.5, suggesting that CO_2 is an activator. An assay performed under conditions of Ci disequilibrium proved that sAC is activated by both HCO_3^- and CO_2 (Townsend *et al.*, 2009).

2.3.5 sAC function *in vivo*

sAC was demonstrated to be an important factor for sperm maturation when a sAC knockout mutant was generated. The female knockout mutants had normal fertility whereas the male knockout mutants had normal sperm morphology but drastically impaired sperm motility that was restored by cAMP loading (Esposito *et al.*, 2004). As sAC is expressed in numerous somatic tissues (see Sections 2.3.5.1 to 2.3.5.5) the knock out mutant mice should have displayed somatic phenotypes. Farrell *et al.* used immunological and molecular methods to identify the existence of another somatic isoform of sAC in mice that utilises a unique start site at exon 5 which escapes the knockout allele that removed exons 1 - 5 (Farrell *et al.*, 2008). Hess *et al.* (Hess *et al.*, 2005) demonstrated that rescued sperm from knockout males were not hyperactive and could fertilise an egg *in vitro*. Using a specific sAC inhibitor, they identified that the induction of protein tyrosine phosphorylation and motility are sAC dependent events whereas acrosomal endocytosis is not sAC dependent.

2.3.5.1 sAC and V-ATPase

HCO_3^- responsive sAC was shown to be involved in the recycling of vacuolar H^+ ATPase (V-ATPase) in the male reproductive tract. An increase in the pH of the epididymal lumen elevates the HCO_3^- concentrations that

activate sAC and increase cAMP concentrations. This in turn leads to the accumulation of V-ATPase at the apical surface of the epididymal and vas deferens clear cells (Pastor-Soler *et al.*, 2003). The accumulated V-ATPase pumps H^+ into the epididymal lumen to maintain spermatozoa in a quiescent state (Pastor-Soler *et al.*, 2003). As sAC is expressed in numerous other tissues including the kidney and choroids plexus, it may play an important part in a cAMP-signal transduction pathway by indirectly sensing pH via HCO_3^- concentrations. Indeed, further studies found sAC localised to numerous cellular compartments that contain cAMP targets such as the nuclei (Zippin *et al.*, 2003).

2.3.5.2 sAC and CFTR channels

sAC has been demonstrated to be involved in the regulation of the HCO_3^--transporting cystic fibrosis transmembrane conductance regulator (CFTR), the defective HCO_3^-/Cl^- channel in airway epithelial cells of CF sufferers. Studies performed in Calu-3 cells showed that HCO_3^--increased cAMP levels and the addition of a sAC specific inhibitor (2-hydroxyestradiol (2-HE)), but not a tmAC inhibitor (SQ22536) blocked increases in cAMP, suggesting that sAC is functional in Calu-3- cells (Wang *et al.*, 2005). HCO_3^- also increased the conductance of CFTR to Cl^- and the increase in conductance was blocked by 2-HE and a CFTR channel inhibitor, glibenclamide (Wang *et al.*, 2005). These findings demonstrate that levels of HCO_3^-/CO_2 regulate CFTR through the sAC pathway.

2.3.5.3 sAC and airway cilia

sAC has been demonstrated to be expressed in the cilia of human tracheal cells using RT-PCR, Western blotting and immunofluorescence (Schmid *et al.*, 2007). HCO_3^- was observed to increase the accumulation of cAMP in human tracheal cells and cAMP accumulation was blocked in the presence of a sAC specific inhibitor (KH7) and a PKA specific inhibitor but not by a tmAC specific inhibitor (SQ22536) (Schmid *et al.*, 2007). The addition of

HCO_3^-/CO_2 to tracheal cells caused a reduction of intracellular pH and decreased the cilia beat frequency (CBF). However, cAMP production increased CBF via the sAC pathway as demonstrated by a greater decrease in CBF seen in the presence of KH7 (Schmid *et al.*, 2007). sAC was therefore shown to be located in airway cilia and regulated by HCO_3^-.

2.3.5.4 sAC and nerve growth factor

Nerve growth factor (NGF) and cAMP are both involved in neuronal differentiation through accumulation of cAMP that activates Rap1 and neuritogenesis (Heidemann *et al.*, 1985). However a link between NGF and cAMP was unknown. Stessin *et al.* demonstrated, in PC12 cells, that NGF accumulation of cAMP was dependent on intracellular Ca^{2+}, which in turn activated Rap1. The addition of a sAC specific inhibitor (KH7) rather than a tmAC specific inhibitor blocked the activation of Rap1 (Stessin *et al.*, 2006). This finding identified sAC as involved in NGF signalling through a distinct cAMP-dependent pathway to tmACs.

2.3.5.5 sAC and axonal outgrowth

Netrins are a family of signalling molecules that stimulate axonal outgrowth and growth cone attraction or repulsion. These in turn are required for the development of the nervous system and axonal regeneration after an injury. Addition of netrin-1 leads to the accumulation of cAMP in growth cones and inhibition of PKA reduces the effect of netrin-1 (Hopker *et al.*, 1999). Wu *et al.* demonstrated that netrin-1 accumulation of cAMP in growth cones was blocked with the addition of sAC specific inhibitors, KH7 and 2-HE (Wu *et al.*, 2006). Thus in response to netrin-1, sAC was shown to be important in mediating axonal outgrowth.

2.3.5.6 Inhibitors of sAC

The metabolites of the steroid hormone oestrogen, catechol oestrogens (CEs), inhibit the production of cAMP in sAC rich cell extracts (Braun, 1990). CEs bind non-competitively to a hydrophobic patch near the active site of the enzyme, distinct from the binding site of P-site inhibitors (Steegborn *et al.*, 2005). CEs form a chelating interaction with the Mg^{2+} in the active site, distorting the shape of the active site and fixing the enzyme in a non-productive conformation.

N-Methylanthraniloyl (MANT)-GTP, a fluorescence probe that was used to study conformational changes in G proteins and AC toxins (Sarfati *et al.*, 1990; Jameson and Eccleston, 1997) inhibited the Class II AC EF from *B. anthracis* and CyaA from *B. pertussis* with a K_i of approximately 10 μM (Sarfati *et al.*, 1990). It was hypothesised that the MANT group interacts with a hydrophobic pocket in the active site (Sunahara *et al.*, 1998) although a model has yet to be developed. MANT-GTP and other MANT-nucleotides showed discrimination between tmACs and sAC, whereby tmACs were inhibited but sAC was surprisingly insensitive to inhibition (Gille *et al.*, 2004).

A sAC specific inhibitor was discovered by a screen of combinatorial chemical compound library and was termed KH7 (Hess *et al.*, 2005). KH7 has an $I.C._{50}$ of approximately 3-10 μM and was used to determine that sperm capacitation and *in vitro* fertilisation are sAC dependent (Hess *et al.*, 2005).

2.4 *Mycobacterium tuberculosis* Class III ACs

2.4.1 The mechanism of tuberculosis infection

Tuberculosis is a disease that has been documented extensively throughout history. The gram-positive bacterium *Mycobacterium tuberculosis* is the causative agent of tuberculosis and, interestingly, 30% of its genome is devoted to either lipid synthesis or metabolism (Cole *et al.*, 1998). In 2000 the World Health Organisation reported that 2 billion people are infected with *M. tuberculosis* and that there are approximately 8 million new cases of infection a

year (Russell, 2003). Tuberculosis is second only to HIV as the main killer among infectious diseases and because of this there is a large amount of research into its mechanism of infection, with the hope that ultimately there will be the production of an effective anti-tuberculosis therapy.

Inhaled *M. tuberculosis* bacilli target the macrophages of the alveoli and are phagocytosed into phagosomes inside the macrophage. As the body's first line of defence, macrophages phagocytose foreign particles and follow a well-defined sequence of steps leading to acidification and fusion of phagosomes to lysosomes. However, *M. tuberculosis* is able to survive dormant in the phagosome until the host is immuno-compromised, by blocking phagosomal acidification and lysosomal fusion. During normal phagosome maturation, the internalised early phagosome binds markers such as early endosomal antigen 1 (EEA1) and the monomeric G-protein, Rab5. As the phagosome matures to a late phagosome, cathepsin D, lysosomal membrane protein 1 (LAMP1), V-ATPase and Rab7 also bind to the phagosome and causes the transportation of the phagosome into the late endosome and fusion with the lysosome (Rink *et al.*, 2005). Work done on *Mycobacterium avium* showed that infected phagosomes failed to acidify below pH 6.3 due to inhibition of fusion with V-ATPases, however, phagosomal acquisition of LAMP1 still occurred, (Sturgill-Koszycki *et al.*, 1994) suggesting that phagosome activity was not affected, but that maturation of an early phagosome into a late phagosome is somehow blocked. This hypothesis was backed up by the detection of Rab5 on *M. tuberculosis* infected phagosomes and the conspicuous absence of Rab7 (Via *et al.*, 1997). Thus the bacterium is able to survive dormant inside the phagosome while living off the nutrients endocytosed by the cell. *M. tuberculosis* infected phagosomes also inhibit the assembly of actin in the membrane, which is a requirement for phagosomal maturation and membrane fusion, whereas avirulent or dead mycobacteria infected phagosomes nucleated actin (Anes *et al.*, 2003). Anes *et al.* also showed that the addition of certain lipids such as ceramide and sphingosine reversed the actin assembly inhibition, suggesting a possible mode of anti-tuberculosis therapy.

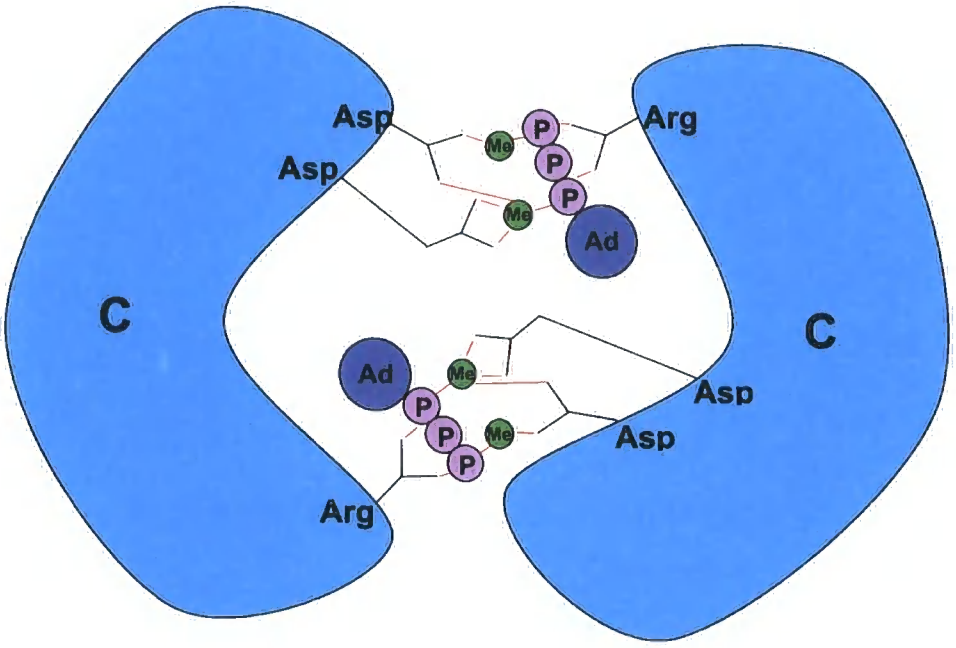
Further work on actin assembly regulation found that increased levels of cAMP inhibited actin assembly and conversely decreased levels of cAMP stimulated actin assembly (Kalamidas *et al.*, 2006). The authors argue that the

phagosome cAMP-PKA system behaves as a molecular switch to regulate actin assembly and phagosome maturation. High intracellular levels of cAMP have been identified from work done on virulent and avirulent *Mycobacterium microti* infected macrophages as a factor that inhibits the fusion of phagosome and lysosome (Lowrie *et al.*, 1975). It was further shown that *Mycobacterium bovis* BCG and *Mycobacterium lepraemurium* also inhibited phagolysosome formation by increasing cAMP levels (Lowrie *et al.*, 1979). There are only two systems in the cell that can elevate cAMP levels, the first is activation of AC and the second is inactivation of a PDE.

The complete genome of the *M. tuberculosis* H37Rv strain contains approximately 4,000 genes (Cole *et al.*, 1998) with 15 genes putatively assigned as ACs based on sequence homology (McCue *et al.*, 2000). It was suggested that due to this large number of ACs, *M. tuberculosis* may sense intracellular and extracellular signals via the cAMP pathway. As elevated cAMP levels halt phagosome–lysosome fusion, pathogenesis in *M. tuberculosis* may involve AC, such as with the AC CyaA from *B. pertussis*, which is part of the edema factor that causes whooping cough (Mock and Ullmann, 1993).

So far, of the 15 putative *M. tuberculosis* AC genes, 10 have been characterised as ACs and the remaining 5 genes have yet to be characterised. The characterised ACs all have catalytic domains that dimerise to form a homodimer with two catalytic centres (Figure 2.2).

Figure 2.2: Structure of a typical prokaryotic Class IIIb AC forming a homodimer with two catalytic centres from the two cytoplasmic domains, C. Ad – adenosine, P – phosphate group, Me – divalent metal cofactor.


Asp
Asp
Arg
P
P
P
Me
Me
Ad
C
C
Ad
Me
P
P
P
Me
Asp
Asp
Arg

2.4.2 Rv0386

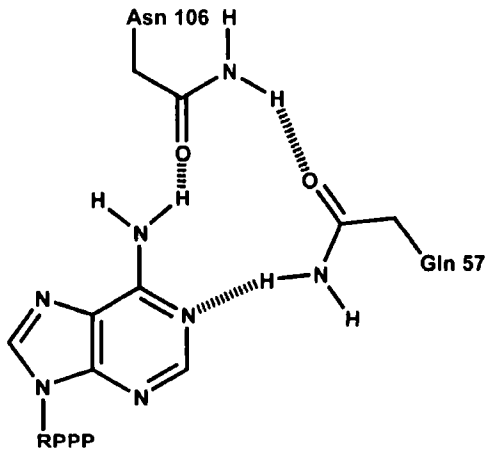
The *rv0386* gene encodes Rv0386, classified as a Class IIIc AC due to a shortened 'arm' region (Castro *et al.*, 2005). The AC catalytic domain is at the N-terminus, followed by an AAA-ATPase domain, a tetratricopeptide repeat-like region and a helix-turn-helix (HTH) DNA-binding domain at the C-terminus. The AAA-ATPase domain is similar to the transcriptional regulator of *Streptomyces coelicolor* (Horinouchi *et al.*, 1990) and the HTH DNA-binding domain suggest that Rv0386 may be linked to transcriptional regulators. Rv0386 has glutamine (Gln) 57 and Asn 106 as the substrate defining residues instead of Lys and Asp. Mutational analyses of the substrate defining residues showed that both Q57 and N106 are required for substrate binding. Rv0386 also has a 20% guanylyl cyclase (GC) activity compared to AC activity, the highest in all the characterised *M. tuberculosis* ACs. Rv0386 must use a novel substrate binding mechanism to bind both ATP and GTP and it was proposed that it was an amide switch of both Q57 and N106 (Figure 2.3) (Castro *et al.*, 2005).

2.4.3 Rv1900c

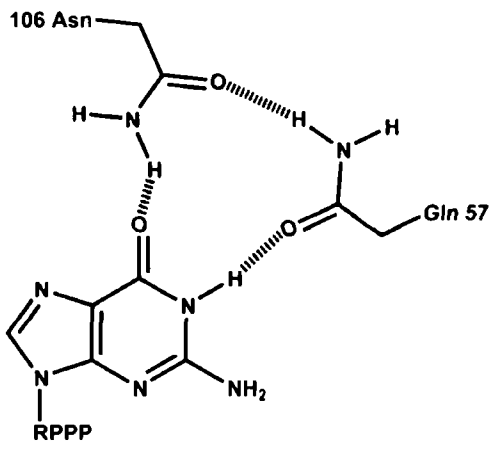
Rv1900c is a Class IIIc AC composed of a N-terminal α/β-hydrolase domain and a C-terminal AC catalytic domain (Sinha *et al.*, 2005). Purified Rv1900c has a specific activity of 1.3 μmol mg^{-1} min^{-1} in the presence of ATP and Mn^{2+} and does not accept Mg^{2+} as a metal cofactor. This could be due to the metal coordination sphere of the active site being too large for the smaller Mg^{2+} ion. The *M. tuberculosis* genome encodes a Mn^{2+} transporter which may sustain sufficiently high concentrations of intracellular Mn^{2+} for activity (Cole *et al.*, 1998). Rv1900c also has a high GC side activity of 7% of AC activity. This is probably due to the substrate specific residues in the active site, N342 and D395, not specifying adenine. This was shown to be the case by mutation of the residues, which did not alter activity. Therefore there may be other factors that influence nucleotide specificity such as electrostatic gradients, shape complementarity, and size constraints.

Figure 2.3: Hypothesised binding of ATP and GTP to the active site of Rv0386. Gln 57 and Asn 106 act as an amide switch that allows Rv0386 to bind to either ATP (A) or GTP (B).

A


Asn 106
H
N
H
O
O
H
H
N
O
Gln 57
N
H
N
N
H
N
N
N
RPPP

B


106 Asn
O
H
N
H
H
N
H
O
O
Gln 57
H
N
N
N
N
NH2
RPPP

The transition state stabilising residue, H402, the equivalent of the mammalian asparagine, does not appear to bind substrate, as seen from an X-ray crystal structure of Rv1900c with the nonreactive ATP analogue, α,β-methyleneadenosine 5'-triphosphate (AMPCPP). However, the lack of transition state stabilisation does not affect the affinity or rate of catalysis of Rv1900c for ATP. The crystal structures of unliganded and AMPCPP-bound Rv1900c ACs show that only one of the two active sites is used for catalysis. The other active site is kept in an inactive conformation by the catalytic residues forming a network of hydrogen bonds. Rv1900c is thus intrinsically asymmetric, having two identical active sites in sequence but not in structure. This asymmetry may have led to the formation of pseudoheterodimeric mammalian ACs that use one active site for catalysis and the other as a regulatory site (Sinha *et al.*, 2005).

2.4.4 Rv1647

Rv1647 is a Class IIIc AC that has greater sequence similarity to fungal and protist ACs than to eukaryotic ACs (Shenoy *et al.*, 2005). Purified Rv1647 from amino acids 97 to 328 has a 200-fold greater specific activity in the presence of Mn^{2+} than Mg^{2+}. The optimum pH for activity was 8.5, the optimum temperature was 37 °C, the activation energy was 74.4 ± 4.8 kJ mol^{-1}, and there was no GC side-activity detected. The catalytic activity of the enzyme was increased 2–2.5-fold in the presence of detergents such as CHAPS (3[(3-Cholamidopropyl)dimethylammonio]-propanesulfonic acid) and specific activity also increased in the presence of 500 mM NaCl suggesting that dimerisation may be dependent on hydrophobic interactions. Rv1647 was shown to be a homodimer through the generation of active site mutants.

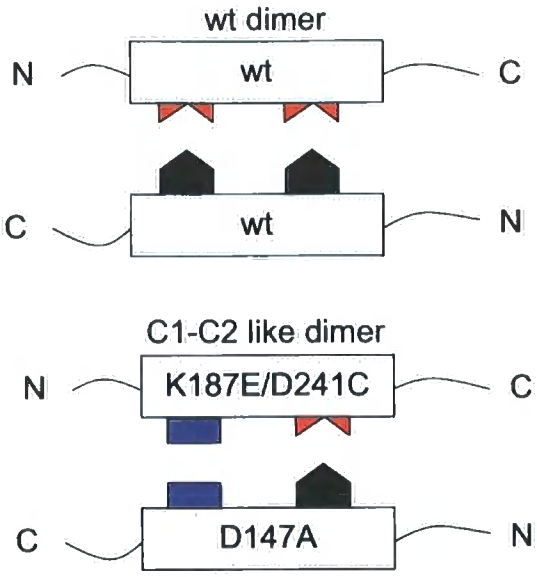
C1-like mutants were generated by mutation of the substrate binding residues, K187 and D241, and a C2-like mutant was generated by mutation of a metal binding residue, D147. These mutants were barely active by themselves, but when reconstituted to form a C1/C2-like AC, specific activity was high demonstrating that Rv1647 acts as a homodimer and that the catalytic residues are similar to the catalytic residues of mammalian ACs

(Figure 2.4). Mutations to GC substrate specific residues did not give Rv1647 GC activity. Rv1647 has an orthologue from *Mycobacterium leprae*, ML1399 that has also been cloned, expressed, purified and assayed. *ML1399* is a functional gene in the *M. leprae* genome, despite being flanked by pseudogenes, as shown by Western blot analysis of cell-free extracts. Thus Rv1647 and its conserved orthologues may be important and essential ACs in mycobacteria.

2.4.5 Rv1318c, Rv1319c, Rv1320c and Rv3645

M. tuberculosis H37Rv has four Class IIIb ACs, Rv1318c, Rv1319c, Rv1320c and Rv3645, characterised by the substitution of the substrate specific Asp for a Thr and an arm region extended by one residue compared to Class IIIa ACs (Linder and Schultz, 2003). The four *M. tuberculosis* Class IIIb ACs all have an N-terminal membrane anchor consisting of six transmembrane regions, then a HAMP domain (histidine kinase, adenylyl cyclase, methyl accepting chemotaxis proteins and phosphatases) followed by the C-terminal AC domain. HAMP domains contain two amphiphilic helices of about 50 amino acids each joined by a short linker. HAMP domains are speculated to function as autonomous switches between two signalling states (Aravind and Ponting, 1999). All four enzymes have specific activity in the presence of Mn^{2+} with a small GC side-activity detected in Rv1319c and Rv3645 (Linder *et al.*, 2004). The HAMP domain in Rv1319c reduced activity by 70% compared to catalytic domain of Rv1319c only, whereas in Rv3645 the activity was increased 21-fold.

Figure 2.4: Schematic representation of C_1-C_2 like complementation of Rv1647 proteins. Red block represents substrate binding residue, black block represents metal binding residue and blue block represents mutated residue. Wild type dimer forms two catalytic pockets with functional substrate binding and metal binding residues. C_1-like $Rv1647_{K187E/D241C}$ substrate binding double mutant protein was complemented with the C_2-like $Rv1647_{D147A}$ metal binding single mutant protein demonstrating that Rv1647 has a similar catalytic mechanism to other known Class III ACs.


wt dimer
N
wt
C
C
wt
N
C1-C2 like dimer
N
K187E/D241C
C
C
D147A
N

2.4.6 Rv1264

Rv1264 is a Class IIIc AC that has an N-terminal domain followed by a linker and then the C-terminal AC domain. The N-terminus is similar to six other N-terminal AC domains, one from *Brevibacterium liqufaciens*, two from *Streptomyces*, Rv2212 from *M. tuberculosis,* and two DNA sequences from *M. smegmatis*, which are probably homologues of the *M. tuberculosis* ACs. Both the holoenzyme and the isolated catalytic domain (amino acids 211-397) were expressed, purified, and assayed (Linder *et al.*, 2002). $Rv1264_{1-397}$ and $Rv1264_{211-397}$ were both active in the presence of Mn^{2+} with Mg^{2+} not supporting activity and there was no GC side-activity detected. The temperature optimum for specific activity was 37 °C and the optimum pH was 7.3. The specific activity of the holoenzyme was 300-fold less than $Rv1264_{211-397}$, with V_{max} decreased by 3 % and $K_{m(app)}$ decreased 6-fold. This suggests that the N-terminal domain inhibits the catalytic activity of Rv1264. Studies of the holoenzyme AC activity at different pH values showed that there was a 40-fold increase in specific activity from pH 8.0 to pH 6.0 and a 12-fold increase in V_{max} (Tews *et al.*, 2005). Assays of $Rv1264_{211-397}$ gave uniform AC activity between pH 8.0 and 5.5. Thus pH sensitivity is mediated by the N-terminal domain of Rv1264.

Analyses of the crystal structure of the active and inhibited states of Rv1264 provide a mechanism for pH sensing. The transition from the inhibited state to the active state is marked by a change of tertiary structure. The two catalytic domains are rotated 55° and translated by 6 Å with respect to the N-termini. This movement forms two catalytic sites with the catalytic residues moving by 25 Å. Further X-ray crystallography work on the regulatory N-terminal domain gave a 1.6 Å crystal structure of the protein, which has a helical bundle surrounded by a disk-shaped dimer (Findeisen *et al.*, 2007). The regulatory domain must also be a defined distance from the catalytic domain to mediate pH sensitivity of the AC. The regulatory domain could also act as a sensor of the changing lipid environment due to the discovery that oleic acid binds to a hydrophobic pocket on the regulatory domain.

As *M. tuberculosis* must survive in the acidic phagosome during host invasion, it is feasible that Rv1264 acts as a pH sensor in the phagosome and results in adaptation of *M. tuberculosis*. An Rv1264 homologue from *S. coelicolor* (*cya*) was linked with a physiological role in the acid stress response. The *cya* mutant of *S. coelicolor* did not neutralise an acidic environment, unlike the wildtype, and stopped growth and sporulation (Susstrunk *et al.*, 1998). Thus the *rv1264* mutant *M. tuberculosis* was characterised for virulence. However both the wild type and mutant infected mouse organs showed characteristic mycobacterial growth, therefore deletion of Rv1264 does not affect the course of infection of *M. tuberculosis* in mouse organs (Dittrich *et al.*, 2006). The authors do note that although the results suggest that Rv1264 is not a virulence factor, one of the other 14 mycobacterial ACs may be compensating for the loss of Rv1264 activity.

2.4.7 Rv2212

Rv2212 is a Class IIIc AC that has the same domain composition as the pH sensing Rv1264, with an N-terminus that is 21% identical to that of Rv1264 and a C-terminal catalytic domain that is 29% identical (Linder *et al.*, 2002). Both the holoenzyme and the isolated catalytic domain (amino acids 212-388) were expressed, purified, and assayed (Abdel Motaal *et al.*, 2006). Both recombinant ACs were active in the presence of Mn^{2+} but not Mg^{2+}, there was no GC side-activity detected and the optimum pH was 6.5. The specific activities of $Rv2212_{1\text{-}388}$ and $Rv2212_{212\text{-}388}$ were similar demonstrating that the N-terminal domain is not autoinhibitory as in Rv1264. The N-terminal domain of Rv2212 also does not act as a pH sensor since the activity at pH 6.5 is only five-fold greater than at pH 9.0, whereas in Rv1264 the fold activity from pH 9.0 to pH 6.0 is 110-fold (Tews *et al.*, 2005). This suggests that the two similar N-termini are probably connected differently or have additional regulatory inputs. An AC from *B. liquefaciens* has an identical domain composition as Rv1264 and Rv2212 and was shown to be strongly stimulated by pyruvate, α-ketocarbonic acids, glycine, alanine, and lactate (Takai *et al.*, 1974). Various metabolites were assayed with Rv2212 and fatty acids such as palmitic, oleic,

linoleic, linnolenic, and arachidonic acids were found to cause a large stimulation. The stimulation was caused by an increase in substrate affinity with V_{max} unaffected and it was specific to the Rv2212 holoenzyme. The unsaturated fatty acids also increased pH sensitivity in Rv2212 from 5-fold to 20-fold. The stimulation was dependent on the presence of the N-terminal domain. The stimulation was not seen in the presence of detergents and so is not wholly due to the amphipathic nature of Rv2212.

2.4.8 Rv1625c

Rv1625c is the most studied of the *M. tuberculosis* ACs due to it corresponding to a half of a mammalian AC. At the N-terminus of Rv1625c there is a six transmembrane domain followed by the catalytic domain at the C-terminus. A soluble catalytic AC construct was generated from the end of the transmembrane domain, D204, to the C-terminus and this construct formed a homodimer with two catalytic centres. $Rv1625c_{204-443}$ had K_m values of 150 µM in the presence of 2 mM Mn^{2+} and 417 µM in the presence of 20 mM Mg^{2+}. V_{max} values were 2100 nmol cAMP mg^{-1} min^{-1} with Mn^{2+} and 27 nmol cAMP mg^{-1} min^{-1} with Mg^{2+}. The optimum pH for specific activity was 7.2, optimum temperature was 37 °C, and the activation energy was 67.5 kJ/mol. $Rv1625c_{204-443}$ has no GC activity and forskolin did not activate the enzyme (Guo *et al.*, 2001). The Rv1625c holoenzyme was expressed successfully in both *E. coli* and HEK (human embryonic kidney) 293 cells, with a specific activity of 1.2 µmol cAMP mg^{-1} min^{-1} using 75 µM Mn-ATP for *E. coli* expressed protein and 65 nmol cAMP mg^{-1} min^{-1} using 65 µM Mn-ATP for HEK 293 expressed protein. The residues K296 and D365 are required for ATP binding and when they were mutated to GTP substrate binding residues, along with a phenylalanine residue to form a K296E/F363R/D365C triple mutant, the mutant protein did not demonstrate GC activity but did show impaired AC activity (Shenoy *et al.*, 2003). This loss of activity was due to the triple mutant existing as a monomer, thus suggesting that the substrate defining residues are also required for dimerisation to occur. The monomeric triple mutant

protein was also used to form crystals for X-ray diffraction at a resolution of 3.4 Å (Ketkar *et al.*, 2004).

To study the structural relationship between Rv1625c and mammalian ACs, point mutations were inserted into Rv1625c to form mammalian C_1- and C_2-like domains. Rv1625c$_{N372A}$ and Rv1625c$_{R376A}$ mutants correspond to mammalian C_1 domains due to the lack of one transition-state stabilising residue. Rv1625c$_{D256A}$ and Rv1625c$_{D300A}$ correspond to mammalian C_2 domains due to the lack one metal-binding Asp. Combination of the C_1-like domain Rv1625c$_{N372A}$ with the C_2-like domain Rv1625c$_{D300A}$ resulted in AC activity of 2.5 µmol cAMP mg^{-1} min^{-1}, comparable to wild type activity (Linder *et al.*, 2004). An attempt to create a forskolin binding site equivalent to that found in mammalian ACs types I-VII was made by creating an Rv1625c$_{D300S}$/Rv1625c$_{N372T}$ heterodimer. The heterodimer was active at 75 µM ATP but was not activated in the presence of 100 µM forskolin. Chimerisation of the C_2-like Rv1625c$_{D300A}$ with mammalian AC C_1 domains did not give functional reconstitutions, however chimerisation with the *Paramecium* GC C_2 domain gave an AC activity of 6.8 nmol cAMP mg^{-1} min^{-1} using Mn^{2+}.

Work on the role of the membrane domain of Rv1625c showed that the membrane domains stabilise the enzyme, lowering the K_m value to 20 µM Mn^{2+}-ATP. Production of a mammalian-like Rv1625c holoenzyme, with a head-to-tail sequence, gave a K_m value of 16 µM Mn^{2+} and a V_{max} value of 3.5 µmol cAMP mg^{-1} min^{-1} (Guo *et al.*, 2005). Thus the mammalian-like tandem sequence is kinetically favourable compared to Rv1625c monomer. Replacement of the Rv1625c membrane domains with membrane domain 1 or 2 from rabbit type V AC formed chimeric proteins that were active, demonstrating that the membrane domains of Rv1625c and rabbit type V AC are compatible.

The role of Rv1625c in pathogenesis *in vivo* has also been investigated. An ΔRv1625c mutant of *M. tuberculosis* H37Rv was infected into mice and the infection was observed through mycobacterial growth in the organs and granuloma formation in the lungs. However the ΔRv1625c mutant was indistinguishable from the wild type strain and so it can be considered not a virulence factor on its own (Guo *et al.*, 2005). The authors suggest that

Rv1625c may affect a different stage of infection or that one of the other 14 ACs from *M. tuberculosis* may compensate for the loss of Rv1625c.

Cellular regulation of Rv1625c has been hypothesised to work via the bacterial stress response. Polyphosphates (poly-P) of a mean chain length of 72 residues were demonstrated to inhibit not only Rv1625c but other Class III ACs as well (Guo *et al.*, 2009). Interestingly, Class I ACs are not inhibited by poly-P, suggesting that poly-P binds to the catalytic site and also allow bacteria with different classes of ACs to have different levels of regulation.

2.5 Inorganic carbon and cAMP

Ci plays an important role in the physiology of all organisms. The predominate forms of Ci that are biologically active are HCO_3^- and CO_2. These Ci forms are involved in numerous physiological roles in bacteria, plants and animals, for example as a substrate in photosynthetic carbon fixation in plants (Falkowski, 1997), maintaining pH homeostasis in all organisms (Roos and Boron, 1981), as a metabolite in carbon metabolism in all organisms (Smith and Ferry, 2000), as an activator of virulence in pathogenic organisms (Bahn and Muhlschlegel, 2006), as an activator of sperm maturation (Esposito *et al.*, 2004), and as an alarmone in *Drosophila* (Jones *et al.*, 2007; Kwon *et al.*, 2007). Owing to the physiological importance of Ci, considerable work has been performed to elucidate the molecular mechanism of Ci sensing in organisms. The two seven transmembrane receptors, Gr21a and Gr63a in *Drosophila* confer CO_2 responsiveness in *Drosophila* neurons (Jones *et al.*, 2007; Kwon *et al.*, 2007). CO_2 sensitivity is also observed in mice GC-D^+ expressing olfactory neurons that allow mice to detect CO_2 at near-atmospheric concentrations (approximately 0.038 %) (Hu *et al.*, 2007). Carbonic anhydrase (CA) is required for CO_2 sensing and the authors hypothesise that CO_2 is converted into HCO_3^- which in turn increases cGMP (3', 5'-cyclic guanosine monophosphate) levels and the opening of cGMP-sensitive cyclic nucleotide-gated channels. CO_2 avoidance in *Caenorhabditis* involves CO_2 sensing via the cGMP-activated channel subunits TAX-2 and

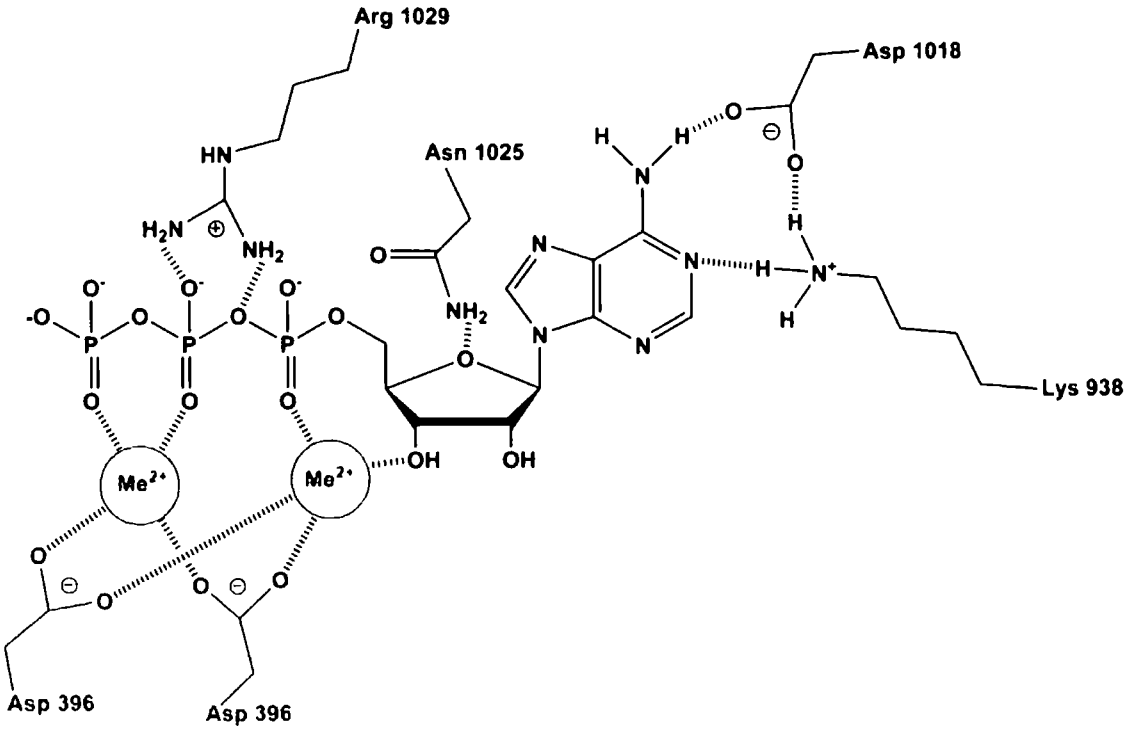
TAX-4, however a CO_2 receptor has yet to be identified (Bretscher *et al.*, 2008; Hallem and Sternberg, 2008).

In mammalian systems it has long been known that increased levels of CO_2 affect cellular cAMP levels. In rabbit carotid bodies (CB), a high partial pressure of CO_2 (pCO_2) and low pH gave an increase in cAMP levels (Perez-Garcia *et al.*, 1990). This rise in cAMP arose because of AC activation as shown by the treatment of the CBs with forskolin, which also increased cAMP levels and released dopamine. Further experiments on changes in rat spleen in response to CO_2 showed that on administration of CO_2 the spleen contracted and cAMP levels increased. This was due to CO_2 stimulating the α-adrenoreceptors which caused the contraction of the spleen and the stimulation of the β-adrenoreceptors which increased cAMP levels (Lehotay and Murphy, 1974). The addition of β-adrenoreceptor inhibitors and agonists was separately shown to stimulate the adrenoreceptors. Elevated pCO_2 does not always cause an increase in cAMP levels. Work done on coronary ligation in dogs induced an increase in pCO_2 of the ischemic tissues post-ligation, which also severely decreased levels of cAMP but did increase cGMP levels (Stewart *et al.*, 1978). A possible explanation of the decrease in cAMP levels could be due to the arrest of blood flow to the heart which would diminish the delivery of adrenergic hormones.

The first biochemical evidence of a link of Ci to ACs was in studies designed to study Ca^{2+} induced stimulation of cAMP in guinea pig spermatozoa (Garbers *et al.*, 1982). It was found that HCO_3^- was required for Ca^{2+} stimulation to occur and that HCO_3^- increased cAMP levels by 25-fold. The first AC demonstrated to be stimulated by HCO_3^- was mammalian sAC (Section 2.3.4). A knockout strain in mice caused male sterility that impaired sperm motility but not spermatogenesis (Esposito *et al.*, 2004). This suggested a link with HCO_3^- which is required for sperm maturation and mobilisation. The presence of HCO_3^- caused a 7–30-fold stimulation in specific activity of the truncated sAC. A crystal structure of sAC has not been obtained, however the Ci mechanism of regulation was studied with a prokaryotic Class IIIb AC that is also stimulated 2.5-fold by HCO_3^- (Chen *et al.*, 2000), CyaC from *Spirulina platensis*. A crystal structure with ATP analogues in the presence and absence

of HCO_3^- was solved, however the addition of HCO_3^- caused the CyaC crystals to dissolve so the mechanism of HCO_3^- stimulation was suggested from flash soaking of HCO_3^--induced conformational changes of the crystals (Steegborn *et al.*, 2005). The suggested mechanism of HCO_3^- stimulation of CyaC is through the recruitment of a second metal ion into the active site (see Figure 2.5). The authors suggested that the addition of HCO_3^- dissolved the crystals due to HCO_3^- causing a structural change (Steegborn *et al.*, 2005). Thus HCO_3^- was not detected in the CyaC crystal structure and the HCO_3^- binding site is unsolved.

Figure 2.5: Active site of a typical mammalian Class IIIa AC (Tesmer *et al.*, 1999). ATP is in black, metal ions are in green, hydrogen bonds are in red, C1a residues are in pink and C2a residues are in blue. Amino acid labelling is from their positions in canine type V AC for C1a and rat type II AC for C2a.


Arg 1029
Asp 1018
Asn 1025
Lys 938
Asp 396
Asp 396

2.5.1 Ci stimulated Prokaryotic Class IIIb ACs

Regulation of ACs with HCO_3^- had been thought to be limited to Class IIIb ACs where there is a Thr or Ser (serine) in the substrate defining position of the active site, for example in mammalian sAC (Chen *et al.*, 2000) and the prokaryotic ACs, CyaC from *S. platensis* (Chen *et al.*, 2000), CyaB1 from *Anabaena* PCC 7120 (Cann *et al.*, 2003) and Slr1991 from *Synechocystis* PCC 6803 (Masuda and Ono, 2004).

However more recent work on HCO_3^- stimulation of Class IIIb ACs has identified that the Ci species that is stimulating AC is not HCO_3^- but CO_2 (Hammer *et al.*, 2006). An assumption was made in previous work that the Ci activating species for Class IIIb ACs was HCO_3^- and not CO_2 as the ionic Ci form was more likely to bind to the active site (Hammer *et al.*, 2006).

The final proof that CO_2 was the Class IIIb AC activating form of Ci was from assays conducted under conditions of Ci disequilibrium. The Ci disequilibrium was achieved by reducing the temperature of the assay, which slowed the acquisition of the Ci equilibrium, and shortening the assay period, such that the Ci form added to the assay (either HCO_3^- or CO_2) remained mostly in that form added. The shortened assay period was elucidated by measuring the acquisition of the Ci equilibrium in a weakly buffered solution with the addition of HCO_3^- or CO_2 and in the presence or absence of CA. CO_2 was demonstrated to be the Ci form that stimulated a model Class IIIb AC, Slr1991 of *Synechocystis* PCC 6803. This experiment was repeated with another Class IIIb AC, CyaB1 from *Anabaena* PCC 7120, previously thought to be stimulated by HCO_3^-, and this enzyme was also proven to be stimulated by CO_2 solely (Hammer *et al.*, 2006). These two enzymes are the first identified signalling molecules that respond directly to CO_2.

Thus in all previous work on Ci stimulation of Class IIIb ACs quoted in this section, HCO_3^- was assumed to be the activating form of Ci. For clarity, from now on in this thesis, the term Ci will be used instead, as the stimulating forms of Ci were not identified in these earlier studies.

2.5.1.1 CyaC from *S. platensis*

Secreted cAMP was found in media containing *S. platensis*, a filamentous cyanobacterium (Sakamoto *et al.*, 1991). The increase in cAMP concentrations were observed to encourage the formation of algal mats (Ohmori *et al.*, 1992). *S. platensis* has at least three genes that code for ACs, *cyaA, cyaB,* and *cyaC* (Yashiro *et al.*, 1996). The CyaC AC is a Class IIIb AC that contains an N-terminal receiver domain of the R1 response regulator protein, a transmitter domain of the sensory kinase protein, a receiver domain of the R2 response regulator protein, and a C-terminal AC domain (Kasahara *et al.*, 1997). CyaC was assayed for Ci stimulation and the addition of Ci stimulated CyaC 2.5-fold with an $E.C._{50}$ value of 18.8 ± 1.6 mM (Chen *et al.*, 2000).

2.5.1.2 CyaB1 from *Anabaena* PCC 7120

The cyanobacterium *Anabaena* PCC 7120 has six different genes that code for ACs in its genome, *cyaA, cyaB1, cyaB2, cyaC, cyaD,* and *cyaE* (Katayama and Ohmori, 1997). The AC from *cyaB1* is a Class IIIb AC and forms a homodimer. At the N-terminus, there are two GAF domains, GAF-A and GAF-B, followed by a PAS (period clock protein, aryl hydrocarbon receptor and single-minded protein) domain and then the catalytic AC domain (Kanacher *et al.*, 2002). GAF domains are associated with mammalian cyclic nucleotide PDEs and GAF domains specifically bind and hydrolyse cyclic nucleotides, whereas PDEs degrade other cyclic nucleotides (Soderling and Beavo, 2000). In the CyaB1 AC, GAF-B specifically binds to cAMP and activates the AC, thus acting as an autoregulatory switch. This switch is important in the control of homeostasis at limiting Na^+ (Cann, 2007).

Ci stimulates the CyaB1 AC domain with a maximal 2-fold stimulation and an $E.C._{50}$ value of 9.6 mM (Cann *et al.*, 2003). This occurs because Ci lowers substrate affinity but increases reaction velocity. Ci was hypothesised to stimulate Class IIIb ACs that have a substrate defining Thr by mimicking the carboxyl group of the substrate defining Asp found in Class IIIa ACs. This was

thought to occur through Ci binding to the substrate defining Lys (K646), which is conserved in Class IIIb ACs. This theory was supported by a point mutation of Lys-646 to Ala, which removed the residue that hypothetically bound Ci and completely ablated Ci sensitivity. A point mutation of Thr-721 to Asp, which re-introduced the carboxyl group to mimic a Class IIIa AC also ablated Ci stimulation. Ci stimulation assays on other Class IIIb ACs such as CyaB from *Stigmatella* and Rv1319c from *Mycobacterium* demonstrated a fold stimulation which supported this hypothesis (Cann *et al.*, 2003). ACs that contained the substrate defining Asp polymorph (Class IIIa ACs) were shown to be unresponsive to Ci at the assayed conditions such as the mammalian tmACs (Chen *et al.*, 2000) and Rv1625 from *Mycobacterium* (Cann *et al.*, 2003) which again supported the hypothesis that Class IIIb ACs were the only group of ACs that were responsive to Ci.

2.5.1.3 Slr1991 from *Synechocystis* PCC 6803

Synechocystis PCC 6803 is an unicellular organism that moves by twitching in response to light which is dependent upon an increase in intracellular cAMP levels (Terauchi and Ohmori, 2004). Motility occurs due to the sigma factor SigF, which is responsible for biogenesis of pili and the type IV pili structural component PilA1 (Bhaya *et al.*, 2000). Slr1991 (also known as Cya1) consists of an N-terminal Forkhead-associated domain and a C-terminal Class IIIb AC (Terauchi and Ohmori, 1999). It is hypothesised that cAMP is involved in light transduction in *Synechocystis*, whereby cAMP produced by Slr1991 binds to the SYCRP1 protein to form the cAMP-SYCRP1 complex, which controls the biogenesis of pili (Yoshimura *et al.*, 2002) for motility. Slr1991 has a K_m for ATP of 2.2 ± 0.3 μM at pH 7.5 and V_{max} was 66.3 ± 2.2 pmol cAMP nmol protein^{-1} min^{-1} (Masuda and Ono, 2004).

Interestingly, Slr1991 has been demonstrated to be inhibited by Ci. Assays conducted at pH 7.5 showed that Ci increased K_m 15-fold to 33.9 ± 8.1 μM and V_{max} decreased two-fold to 40.5 ± 3.2 pmol cAMP nmol protein^{-1} min^{-1} (Masuda and Ono, 2005). Slr1991 was the first identified Class IIIb AC to be inhibited by Ci and it was hypothesised that Ci suppressed Slr1991 activity by

lowering ATP affinity to the active site and decreasing ATP turnover rate. It was suggested that Ser-173 and Tyr-178 are the possible residues responsible for the Ci inhibition, which sit adjacent to the substrate defining Lys that confers Ci sensitivity (Masuda and Ono, 2005) although mutational studies have so far not been conducted.

However, a repeat analysis of the Ci dependence assay for Slr1991 by Hammer *et al.* in the presence of 20 mM Ci, with an assay time of 40 minutes was shown to stimulate Slr1991 2-fold by 1.2 mM CO_2 (Hammer *et al.*, 2006). This stimulation of Slr1991 was contrary to the findings of Masuda *et al.* but was explained by the authors as probably due to the use of a shorter assay time that did not allow Ci activation of the enzyme (Hammer *et al.*, 2006). CO_2 gave an apparent E.C.$_{50}$ of 52.7 ± 1.0 mM and increased K_m from 11.4 ± 0.7 µM to 16.2 ±1.2 µM but also increased V_{max} from 0.74 ± 0.01 nmol cAMP mg^{-1} min^{-1} to 1.13 ± 0.03 nmol cAMP mg^{-1} min^{-1}, thus increasing k_{cat} (Hammer *et al.*, 2006). To further prove that CO_2 was the activating Ci species, assays with Slr1991 were conducted with 40 mM Ci with a pH range of 6.5 to 8.5, thus increasing the CO_2 from 0.3 mM at pH 8.5 to 15.5 mM at pH 6.5. Fold stimulation of Slr1991 increased from 1.1 at pH 8.5 to 2.4 at pH 6.5 (Hammer *et al.*, 2006). CO_2 stimulation was definitively proved by assaying Slr1991 under conditions of Ci disequilibrium. A fold stimulation was only observed with the addition of 20 mM CO_2 (Hammer *et al.*, 2006) (see Section 2.5.1 for details of Ci disequilibrium assays).

2.5.2 Are Class IIIa ACs stimulated by CO_2?

The new finding that CO_2 and not HCO_3^- is the activating Ci form for Class IIIb ACs raises intriguing questions about previous results on ACs that demonstrated Ci unresponsiveness. The mammalian tmACs were postulated to be unresponsive to Ci as compared to basal activity based on the insensitivity of type V Class IIIa AC (Chen *et al.*, 2000). The prokaryotic Class IIIa ACs were also postulated to be unresponsive to Ci as compared to basal activity based on the insensitivity of Rv1625c from *Mycobacterium* (Cann *et al.*, 2003). The common factor from these Ci assays was that the assays were

conducted at pH 7.5 and thus their Ci insensitivity may be due to the low CO_2 concentrations at pH 7.5 that were unable to activate the AC, rather than unresponsiveness to Ci. New investigation into this area would elucidate whether Class IIIa ACs are stimulated by CO_2. A model Class IIIa AC for a CO_2 stimulation experiment would be Rv1625c from *M. tuberculosis* H37Rv (Section 2.4.8), which was previously shown to be unresponsive to Ci as compared to its basal activity (Cann *et al.*, 2003). Rv1625c is a good model for investigating CO_2 stimulation in mammalian Class IIIa ACs for several reasons. Firstly, mammalian tmACs are difficult to express in prokaryotic systems as they are not post-translationally modified, unlike Rv1625c which expresses well in *E. coli*. Secondly, mammalian tmACs are difficult to assay as they function as heterodimers and thus require two catalytic domains to be active whereas Rv1625c is a homodimer and therefore only needs one catalytic domain. Thirdly, mammalian tmAC activating factors such as $G_s\alpha$ (Sunahara *et al.*, 1997) or forskolin (Seamon *et al.*, 1981) are required to observe any activity of mammalian tmACs, whereas Rv1625c does not require any activating factors. The identification of a CO_2 sensor in *M. tuberculosis* would aid in the understanding of the possible role for CO_2 in pathogenesis, even if the process of pathogenesis is not identified.

2.6 Aims of this thesis

All Class IIIb ACs studied to date that are activated by Ci were previously thought to be regulated by HCO_3^- but have more recently shown to actually be regulated by CO_2. HCO_3^- and CO_2 are in a pH-dependent equilibrium (the Ci equilibrium). All Class IIIa ACs that were assayed for Ci sensitivity were insensitive to HCO_3^-. However, since Class IIIb ACs have been demonstrated to be regulated by CO_2 and not HCO_3^-, Class IIIa ACs may be Ci sensitive with CO_2 as the regulatory form of Ci.

This thesis investigates the hypothesis that Class IIIa ACs, like Class IIIb ACs are regulated by CO_2. This thesis aims to investigate the following points:

- Is $Rv1625c_{204-443}$, a model bacterial Class IIIa AC, , regulated by Ci in the form of CO_2 *in vitro*?

- If $Rv1625c_{204-443}$ is stimulated by CO_2 *in vitro*, is that CO_2 stimulation also observed *in vivo*?

- What are the assay conditions that are required for CO_2 to bind to $Rv1625c_{204-443}$?

- Which amino acid(s) of the catalytic domain of $Rv1625c_{204-443}$ is/are the CO_2 binding site?

3 Methods and Materials

3.1 Molecular Biology techniques

3.1.1 Isolation of plasmid DNA

Plasmid DNA containing the recombinant protein gene can be isolated from bacteria to allow digestion of the recombinant protein gene, mutagenesis of the gene, ligation into another plasmid and transformation into bacteria.

E. coli were grown in 5 mL luria broth (LB) containing appropriate antibiotics at 30 °C overnight. The culture was harvested by centrifugation at 5,445 g for 5 min at 4 °C and the DNA was extracted using a Qiagen Mini-prep kit (Qiagen). The bacterial pellet was resuspended in 250 µL Buffer P1 and transferred to a microcentrifuge tube. 250 µL Buffer P2 was added and the mixture was inverted six times. 350 µL Buffer N3 was then added and the mixture was inverted six times. The mixture was centrifuged at 15,700 g for 10 min. The supernatant was removed and added to a QIAprep spin column. The spin column was centrifuged at 15,700 g for 1 min and the flow-through was discarded. 0.5 mL Buffer PB was added to the spin column, the column was centrifuged at 15,700 g for 1 min and the flow-through was discarded. 0.75 mL Buffer PE was added to the spin column, the column was centrifuged at 15,700 g for 1 min and the flow-through was discarded. The spin column was centrifuged a second time at 15,700 g for 1 min and the residual flow-through discarded. The spin column was placed into a clean microcentrifuge tube. 50 µL Buffer EB was added to the centre of the resin and left to stand for 1 min. The spin column was then centrifuged at 15,700 g for 1 min and the eluted DNA was stored at –18 °C.

3.1.2 Digestion of DNA with restriction enzymes

Digestion of DNA with restriction enzymes allows the isolation at specific recognition sites. A plasmid digested with the same restriction enzymes would allow ligation of the DNA sequence into the plasmid.

DNA was digested with the appropriate restriction enzymes. The appropriate 10x buffer was used with 3 U of restriction enzyme (Fermentas).

H_2O was added to make the total volume of the restriction digest such that enzyme solution constitutes no more than 10 % of total volume. Digests were incubated at 37 °C for 30 min.

3.1.3 Removal of 3' phosphate from DNA

Phosphatase treatment removes 5' and 3' phosphate groups from the vector DNA which prevents the cut vector from recircularisation.

1 U calf intestine alkaline phosphatase (CIAP) (Fermentas) was added to cut vector DNA and the mixture was incubated at 37 °C for 30 min. CIAP was then deactivated by heating the mixture at 75 °C for 10 min. DNA was then separated by electrophoresis (Section 3.1.5) on a 1.5 % agarose gel and the appropriate bands were excised.

3.1.4 Ligation of insert DNA and vector DNA

Ligation of the insert DNA and the vector DNA allows the recombinant protein gene to be inserted into a vector for expression of recombinant protein.

A 3:1 molar ratio of cut insert DNA and cut vector DNA was incubated with 5 µL of 2x rapid ligation buffer and 3 U (1 µL) of T4 DNA ligase (Promega). The ligation mixture was incubated at 25 °C for 60 min.

3.1.5 DNA electrophoresis

Agarose gel electrophoresis separates DNA by size. An electric field moves the negatively charged DNA through an agarose matrix.

A 1.5 % (w/v) agarose gel was produced by adding 0.75 g agarose (Sigma Aldrich) to 50 mL Tris- (tris(hydroxymethyl)aminomethane) acetate-EDTA (TAE) electrophoresis buffer (40 mM Tris-acetate, 1 mM EDTA (ethylenediaminetetraacetic acid)) and boiling the mixture until the agarose dissolved. 1.5 µM ethidium bromide was added to the liquid agarose mixture when cool. The agarose mixture was poured into a Biorad Mini-Sub Cell GT.

6x Orange G loading buffer (0.1 % (w/v) Orange G, 15 % (w/v) Ficoll 400, 1x TAE) was added to DNA samples which were subsequently loaded onto the agarose gel. 2.5 µL of a 1 kb DNA ladder was added to a lane and the gel was electrophoresed at 20 V/cm until the DNA bands were separated.

3.1.6 DNA purification from agarose gel

DNA separated by agarose gel electrophoresis can be purified from the agarose gel to obtain the DNA band of interest for ligation into a plasmid.

The DNA band of interest was excised from the agarose gel using a scalpel and the gel piece was placed into a microcentrifuge tube. 600 µL binding buffer (6 M sodium perchlorate, 50 mM Tris pH 8.0, 10 mM EDTA) (Sigma Aldrich) was added and the agarose was melted by heating at 37 °C. 5 µL of silica matrix (Promega) was added to the mixture which was then incubated for 10 min at room temperature. The mixture was centrifuged at 15,700 g for 1 min and the supernatant was discarded. The pellet was resuspended in 125 µL binding buffer. The mixture was centrifuged at 15,700 g for 1 min and the supernatant was discarded. The pellet was resuspended in 1.5 mL wash buffer (20 mM Tris pH 8.0, 2 mM EDTA, 400 mM NaCl, 50 % (v/v) ethanol). The mixture was centrifuged at 15,700 g for 1 min and the supernatant was discarded. The pellet was again resuspended in 1.5 mL wash buffer, centrifuged at 15,700 g for 1 min and the supernatant was discarded. Any remaining liquid was removed with a paper towel. The pellet was resuspended in 5 µL ultrapure water and incubated for 10 min at 37 °C. The mixture was centrifuged at 15,700 g for 1 min and the supernatant containing the purified DNA was transferred into a new microcentrifuge tube.

3.1.7 Pfu DNA polymerase PCR

PCR is a technique that amplifies a DNA fragment of interest using a pair of primers that complementarily bind to the start and end of the DNA region that is to be amplified, and using thermal cycling, melts the DNA double strands to form the template for replication and elongates using DNA

polymerase, which forms two identical DNA strands from the original strand. This process can be repeated numerous times to amplify the amount of DNA. Pfu DNA polymerase is from the hyperthermophilic bacterium *Pyrococcus furiosus* and has superior proofreading properties compared to Taq DNA polymerase due to a 3' to 5' exonuclease proofreading activity. This allows for higher fidelity of DNA amplification, however the DNA products are blunt-ended.

The polymerase chain reaction (PCR) mixture was made up of 5 ng template DNA, 2.5 µL 10x buffer, 200 µM dNTPs, 25 pmol forward primer, 25 pmol reverse primer, 1 U Pfu DNA polymerase (all Promega) and H_2O added to make a total volume of 25 µL.

The PCR process was performed on an Eppendorf Mastercycler Gradient PCR machine using the following programme: Initial denaturation at 95 °C for 2 min; 30x cycle of denaturation at 95 °C for 15 s, annealing at 55 °C for 30 s and extension at 72 °C for 2 min per kb of DNA; and a final extension at 72 °C for 5 min. 6x Orange G DNA loading dye (Promega) was added, the PCR products were separated using a 1.5 % (w/v) agarose gel and the appropriate bands were excised.

3.1.8 Taq DNA polymerase PCR

Taq DNA polymerase is from the thermophilic bacterium *Thermus aquaticus* that has low replication fidelity, however has a faster amplification of DNA compared with Pfu DNA polymerase and the DNA products of Taq DNA polymerase have an adenine overhang at the 3' end, which is required for TOPO Cloning (Section 3.1.12).

The PCR mixture was made up of 5 ng template DNA, 2.5 µL 10x buffer, 1.25 µL 20x $MgCl_2$, 200 µM dNTPs, 25 pmol forward primer, 25 pmol reverse primer, 1 U Taq DNA polymerase (all Promega) and H_2O added to make a total volume of 25 µL.

The PCR process was performed on an Eppendorf Mastercycler Gradient PCR machine using the following programme: Initial denaturation at 95 °C for 2 min; 30x cycle of denaturation at 95 °C for 15 s, annealing at 55 °C

for 2 min and extension at 72 °C for 2 min per kb DNA; and a final extension at 72 °C for 5 min. The PCR products were separated using a 1.5 % (w/v) agarose gel and the correct bands were excised.

3.1.9 Taq DNA polyermase Adenylation of DNA

Adenylation of DNA fragments is required for TOPO cloning (Section 3.1.12). DNA products of Taq DNA polymerase have a 3' end adenine overhang. A short extension of DNA amplified using Pfu DNA polymerase allows for high fidelity DNA with a 3' end adenine overhang.

The reaction was made up of 500 ng purified blunt-end DNA, 2.5 µL 10x buffer, 1.25 µL 20x $MgCl_2$, 200 µM dATP, 1 U Taq DNA polymerase (all Promega) and H_2O added to make a total volume of 25 µL.

The PCR process was performed on an Eppendorf Mastercycler Gradient PCR machine using the following programme: Extension at 72 °C for 10 min. The PCR products were separated using a 1.5 % (w/v) agarose gel and the appropriate bands were excised.

3.1.10 DNA mutagenesis

To mutate a particular amino acid of a recombinant protein, the mutation can be inserted into the isolated and restricted DNA by two PCRs, one with the forward primer and a reverse mutagenic primer and one with the reverse primer and a forward mutagenic primer. The mutagenic primers contain complementary base pairing and a codon for adenine for the specific mutation. The two PCR products are rejoined by PCR with the forward and reverse primers.

Two PCR experiments were set up, PCR 1 with 25 pmol forward primer and 25 pmol reverse mutagenic primer and PCR 2 with 25 pmol forward mutagenic primer and 25 pmol reverse primer, 5 ng DNA, 200 µM dNTPs and 1 U Pfu DNA polymerase enzyme (all Promega). PCR products were run on a 1.5 % (w/v) agarose gel and the appropriate bands were excised and purified.

Purified DNA from PCR 1 and 2 were used to set up another PCR using 25 pmol forward primer and 25 pmol reverse primer, 200 µM dNTPs and 1 U Pfu DNA polymerase enzyme. PCR product was separated using a 1.5 % (w/v) agarose gel and the appropriate band was excised and purified.

3.1.11 Preparing competent cells for transformation

To allow for the transformation into *E. coli* cells of a DNA vector containing a cloned DNA product of the desired recombinant protein, the *E. coli* cells must be made competent to be able to take up extracellular DNA.

E. coli BL21 cells were grown in 2 mL LB containing 100 µg/mL ampicillin and 50 µg/mL kanamycin at 30 °C overnight. 0.25 mL of overnight culture was grown in 25 mL LB containing appropriate antibiotics at 30 °C until OD_{600} = 0.4. The culture was cooled to 0 °C then the cells were harvested by centrifugation at 2,700 g for 10 min at 4 °C. The supernatant was discarded and the pellet was resuspended in 15 mL solution of 80 mM $MgCl_2$ and 20 mM $CaCl_2$. The cells were again harvested by centrifugation at 2,700 g for 10 min at 4 °C. The supernatant was discarded and the competent cells were resuspended in 1 mL 0.1 M $CaCl_2$.

3.1.12 TOPO Cloning of PCR products and Transformation into Mach-Ti competent cells

TOPO cloning allows the cloning of DNA fragments into the TOPO vector without the addition of DNA ligases. The TOPO vector is linearised, has a 5'–(C/T)CCTT-3' overhang and DNA topoisomerase I covalently linked to the vector. The DNA fragment must be adenylated to allow the DNA fragment to ligate to the 3' thymine overhand of the vector.

Adenylated PCR products were inserted into pCR2.1-TOPO vector. 0.5 µL pCR2.1-TOPO vector (Invitrogen), 500 ng insert DNA and 1 µL salt solution were mixed and incubated at 25 °C for 30 min. 20 µL of Mach1-T1 competent cells were added to the TOPO cloning reaction and the mixture was incubated

at 0 °C for 20 min. The mixture was heatshocked at 42 °C for 30 s then incubated at 0 °C for 2 min. 200 µL of pre-warmed (37 °C) SOC medium (2 % (w/v) bacto-tryptone, 0.5 % (w/v) bacto-yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$ and 20 mM glucose) was added and the mixture was shaken at 37 °C for 30 min. The mixture was plated onto pre-warmed (37 °C) LB agar plates containing 50 µg/mL kanamycin and 1 mM 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside and the plates were incubated at 30 °C overnight.

3.1.13 Transformation of DNA vector into competent cells

A DNA vector containing the desired cloned DNA fragment of a recombinant protein must be transformed into competent cells such as *E. coli* BL21 cells, to allow the expression of the recombinant protein through growth of the *E. coli* cells containing the transformed DNA.

DNA was added to competent cells and incubated for 20 min at 4 °C. The cells were heatshocked at 42 °C for 45 s then incubated for 2 min at 4 °C. 800 µL SOC prewarmed to 37 °C was added and the mixture was shaken for 30 min at 37 °C. The mixture was plated onto a 37 °C prewarmed LB agar plate containing appropriate antibiotics and the plates were grown at 30 °C for less than 14 h.

3.1.14 Preparation of *E. coli* glycerol stocks

To maintain *E. coli* cells transformed with the desired recombinant protein DNA, the cells must be stored at -140 °C in LB containing 40 % glycerol, which is a cryogenic.

E. coli cells were grown in 5 mL LB containing appropriate antibiotics at 30 °C overnight. The cells were harvested by centrifugation at 5,445 g for 5 min at 4 °C. The bacterial pellet was resuspended in 1 mL LB containing 40 % glycerol. The stock was placed in a Cryovial tube and stored at – 140 °C.

3.1.15 DNA sequencing

To identify that the correct DNA fragment has been cloned into the DNA vector or that the correct mutation has been inserted into the DNA fragment, the purified DNA is sequenced and compared to the theorectical DNA sequence.

Purified DNA (Section 3.1.1) was sequenced using 3.2 pmol forward and 3.2 pmol reverse primers at DBS Genomics (Durham University) using a 48-capillary 3730 DNA Analyzer.

3.2 Protein Biochemistry

3.2.1 Protein quantification by Bradford assay

To quantify the amount of recombinant protein that was purified the Bradford assay method, which is a spectroscopic, colorimetric protein assay was used. The Coomassie dye changes from an unstable red form to a stable blue form on binding with protein. The disadvantage of the Bradford assay is that it is only linear for small quantities of protein.

Protein was quantified using the bovine serum albumin (BSA)-calibrated Bradford assay method (Bradford, 1976). BSA concentrations of 0 - 20 µg/µL measured at 595 nm were used to generate a standard line and concentrations of sampled proteins were kept within this range (Figure 3.1).

3.2.2 Protein quantification by UV absorbance

Another method to quantify the amount of recombinant protein that was purified is to measure its UV absorbance at 280 nm. Amino acids with aromatic rings, such as tryptophan, tyrosine and phenylalanine, absorb UV at 280 nm. Therefore to calibrate the UV absorbance, a protein extinction coeffieient must be calculated.

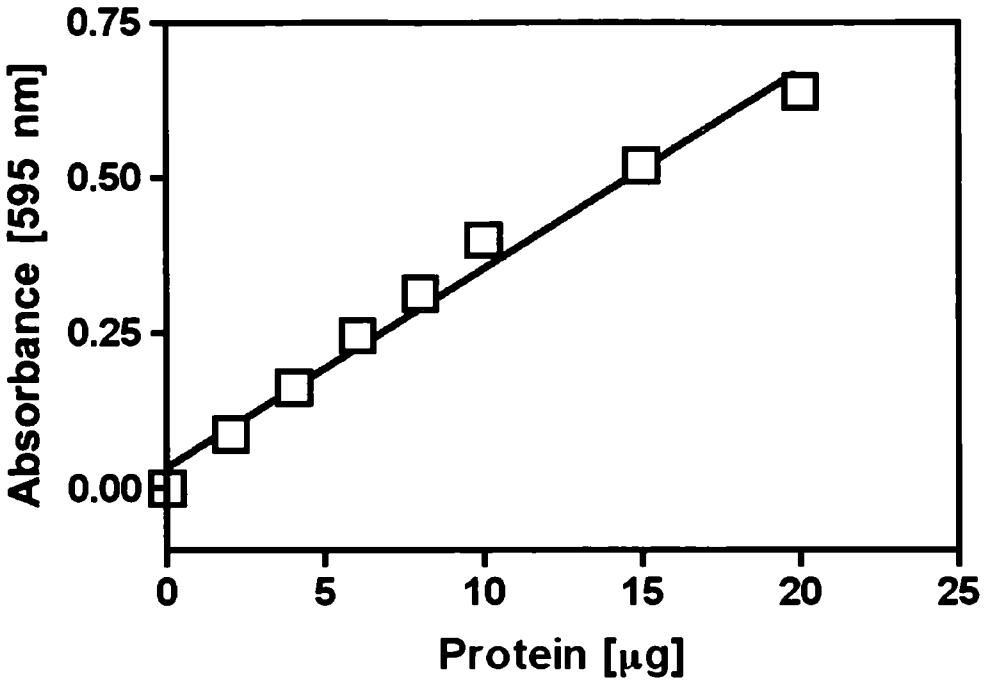
Protein was quantified at 280 nm using a quartz cuvette in a S2100 Diode Array Spectrophotometer and the concentration was calculated by dividing the absorbance by the protein extinction coefficient, obtained by entering the protein sequence into Protean (DNA Star). The spectrophotometer was blanked using appropriate dialysis buffer.

3.2.3 SDS-PAGE

SDS-PAGE is a technique that separates proteins by their electrophoretic mobility. SDS binds and unfolds protein, giving it a uniform negative charge. The polyacrylamide gel has a specific pore size, which separates the protein mixture by size and charge.

Protein molecular weight and purity were assessed using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (Davis, 1964; Ornstein, 1964). 2x loading buffer (100 mM Tris-HCl pH 6.8, 200 mM dithiothereitol, 4 % (w/v) SDS, 0.2 % (w/v) bromophenol blue, 20 % (v/v) glycerol) was added to protein, loaded onto a 18% (w/v) resolving gel with a 5% (w/v) stacking gel and electrophoresed at a voltage of 18 V/cm. The gel was stained with Coomassie Brilliant Blue G (3 mM Coomassie Brilliant Blue G, 12 M methanol and 2 M glacial acetic acid) then destained with destaining solution (12 M methanol and 2 M glacial acetic acid). The destained gel was photographed using a Biorad Gel Doc 2000.

Figure 3.1: Standard curve of BSA for the quantification of protein in the Bradford assay.


0.75
0.50
0.25
0.00
Absorbance [595 nm]
0
5
10
15
20
25
Protein [µg]

3.2.4 His-tag staining of proteins

To identify that the protein with the correct molecular weight is the desired expressed protein with a C-terminal His-tag, the SDS-PAGE gel was stained with a His-tag stain. This stain binds specifically to the oligohistidine domain of a His-tagged recombinant protein and can be visualised at 560 nm.

2 µg protein was run on a 18% (w/v) resolving gel with a 5% (w/v) stacking gel at a voltage of 18 V/cm. The gel was fixed with 200 mL Fixing Solution (9 M ethanol and 2 M glacial acetic acid) for 1 h at 25 °C. The Fixing Solution was decanted and the gel was washed twice with ultra pure H_2O. The gel was incubated with 25 mL InVision His (histidine)-tag In-gel Stain for 1 h at 25 °C. The stain was decanted and the gel was washed twice with 200 mL 20 mM phosphate buffer pH 7.8 for 10 min. The gel was then imaged by excitation at 560 nm using a Typhoon 9400 Variable Mode Imager (Amersham).

3.2.5 pH measurements

pH measurements and various pH values were measured at the temperature stated in the results using a pH meter (Hamilton Biotrode) connected to a computer with a PC card (Orion pH SensorLink) in a final volume of 100 µL. All pH measurements were accurate to ±0.02 pH units (manufacturers specifications). The Orion pH SensorLink program has a time-driven acquisition mode that can measure pH values every second. The pH meter was calibrated using standard pH buffered solutions at pH 4, 7 and 10 (SLS). NaCl/$NaHCO_3$/CO_2 was added to AC assays after the stated time and measurements taken for the time interval stated in the results.

3.2.6 Expression and purification of Rv1625c$_{204-443}$ wild type and mutant proteins

E. coli cells transformed with a pQE30 vector containing the desired recombinant protein gene are grown in LB broth and induced with the addition of IPTG, a lactose metabolite that triggers transcription of the lac operon and the recombinant protein gene. The expression of the recombinant protein also inserts an oligomeric His-tag at the C-terminal. The recombinant protein can then be purified on a His-tag affinity column, which specificially binds the His-tagged recombinant protein. The Recombinant protein is eluted from the column by the addition of imidazole and is dialysed with the appropriate buffer.

pQE30-Rv1625c$_{204-443}$ wt in *E. coli* BL21 [pREP4] and pQE30-Rv1625c$_{204-443}$ mutants in *E. coli* M15 [pREP4] obtained from glycerol stocks (Section 3.1.14) were grown in 100 mL LB broth containing 100 µg/mL ampicillin and 50 µg/mL kanamycin overnight at 30 °C. The overnight culture was added to 8 x 1 L LB containing 100 µg/mL ampicillin and 50 µg/mL kanamycin and grown at 30 °C to O.D.$_{600}$ = 0.6. Rv1625c$_{204-443}$ protein production was induced with 30 µM isopropyl-β-D-thiogalactopyranoside (IPTG) and incubated at 30 °C for 3 h. The culture was harvested by centrifugation at 12,000 g for 10 min at 4 °C, washed once with 20 ml per 1 L culture of bacterial wash buffer (50 mM Tris [tris{hydroxymethyl}aminomethane]-HCl pH 8.5, 1 mM EDTA), centrifuged at 5,445 g for 20 min at 4 °C and the bacterial pellet stored overnight at -80 °C. Cells were brought to 4 °C in a water bath then resuspended in 20 mL per original 1 L culture of cell lysis buffer (50 mM Tris-HCl pH 8.5, 300 mM NaCl, 10 mM 1-thioglycerol and 2 mM $MgCl_2$). The cells were broken by sonication at 74 W for 150 s at 4 °C. Cell debris was removed by centrifugation at 50,000 g for 30 min at 4 °C. 0.5 mL of nickel-nitrilotriacetic acid (Ni-NTA) slurry per original 1 L culture was added to the supernatant and the mixture was incubated for 2 h at 4 °C. The mixture was centrifuged at 405 g for 1 min and the slurry poured into a Polyprep column for washing. The columns were washed with 10x bed volume wash buffer A (50 mM Tris-HCl pH 8.5, 10 mM 1-thioglycerol, 2 mM $MgCl_2$, 400 mM NaCl and 5 mM imidazole), 10x bed

volume wash buffer B (wash buffer A + 15 mM imidazole), 10x bed volume wash buffer C (wash buffer B + 10 mM NaCl) and eluted with 5x bed volume elution wash (wash buffer C + 150 mM imidazole). The eluates were dialysed against 50 mM Tris-HCl pH 8.5, 10 mM NaCl, 2 mM $MgCl_2$, 2 mM 1-thioglycerol and 20% glycerol for 1 h at 4 °C and changed three times (Guo *et al.*, 2001).

3.2.7 Expression and purification of CyaB1$_{595-859}$ and Slr1991$_{120-337}$ wild type and mutant proteins

pQE30-CyaB1$_{595-859}$ wt in *E. coli* M15 [pREP4], pQE30-Slr1991$_{120-337}$ wt in *E. coli* M15 [pREP4] and pQE30-Slr1991$_{120-337}$ mutants in *E. coli* M15 [pREP4] obtained from glycerol stocks (Section 3.1.14) were grown in 100 mL LB broth containing 100 µg/mL ampicillin and 50 µg/mL kanamycin overnight at 30 °C. The overnight culture was added to 8 x 1 L LB containing 100 µg/mL ampicillin and 50 µg/mL kanamycin and grown at 30 °C to $O.D._{600}$ = 0.6. Protein production was induced with 300 µM IPTG and incubated at 30 °C for 3 h. The culture was harvested by centrifugation at 12,000 g for 10 min at 4 C, washed once with 20 ml per 1 L culture of bacterial wash buffer (50 mM Tris-HCl pH 8.5, 1 mM EDTA), centrifuged at 5,445 g for 20 min at 4 °C and the bacterial pellet stored overnight at -80 °C. Cells were purified following the same procedure as Rv1625c$_{204-443}$ (section 3.2.6).

3.2.8 Expression and His-tagged purification of Rv1625c$_{212-443}$ K296E/F363R/D365C

pRSET-Rv1625c$_{212-443}$ in *E. coli* BL21 (C43) [pREP4] obtained from a glycerol stock (Section 3.1.14) was grown in 100 mL LB broth containing 100 µg/mL ampicillin overnight at 30 °C. The overnight culture was added to 8 x 1 L LB containing 100 µg/mL ampicillin and grown at 30 °C to $O.D._{600}$ = 0.6. Protein production was induced with 500 µM IPTG and incubated at 37 °C for 3 h. The culture was harvested by centrifugation at 12,000 g for 10 min at 4 °C, washed once with 20 mL per 1 L culture of phosphate-buffered saline,

centrifuged at 5,445 g for 20 min at 4 °C and the bacterial pellet stored overnight at -80 °C. Cells were brought to 4 °C in a water bath then resuspended in 20 mL per original 1 L culture of cell lysis buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM 2-mercaptoethanol (2-ME) and 2 mM phenylmethylsulfonyl fluoride (PMSF)). The cells were broken by sonication at 74 W for 150 s at 4 °C. Cell debris was removed by centrifugation at 50,000 g for 30 min at 4 °C. 0.5 mL of Ni-NTA slurry per original 1 L culture and 250 mM NaCl was added to the supernatant and the mixture was incubated for 2 h at 4 °C. The mixture was centrifuged at 405 g for 1 min and the slurry poured into a Polyprep column for washing. The columns were washed with 10x bed volume wash buffer A (20 mM Tris-HCl pH 8.0, 5 mM 2-ME and 500 mM NaCl), 10x bed volume wash buffer B (20 mM Tris-HCl pH 8.0, 5 mM 2-ME, 100 mM NaCl and 10 mM imidazole) and eluted with 5x bed volume elution wash (20 mM Tris-HCl pH 8.0, 5 mM 2-ME, 100 mM NaCl, 150 mM imidazole and 10 % glycerol). The eluates were dialysed against 20 mM Tris-HCl pH 8.0, 5 mM 2-ME and 10% glycerol for 1 h at 4 °C and changed three times. The dialysed protein was concentrated to 1 mL and subjected to anion-exchange chromatography on a Biorad Biologic Duo Flow FPLC system using a Q-Sepharose column (25 x 1 cm). The column was equilibrated with dialysis buffer and the protein was injected into the column. The column was washed with 50 mL dialysis buffer at a flowrate of 3 mL/min and then a linear salt gradient of 0-500 mM NaCl was applied for a total volume of 175 mL. The gradient was increased to 1 M NaCl in a total volume of 50 mM and finally the column was washed with 25 mL dialysis buffer. 5 mL fractions were collected at the start of the linear salt gradient. The fractions containing eluted protein were pooled and dialysed into the Superdex 200 running buffer (20 mM HEPES (N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)-NaOH pH 7.5, 5 mM 2-ME and 10 % glycerol) and concentrated to 130 µL. The protein was injected into a Superdex 200 gel filtration column. 30 mL of running buffer at a flowrate of 0.5 mL/min was used to elute the protein. 0.25 mL fractions were collected. Fractions containing pure protein were pooled and concentrated to 14 mg/mL (Ketkar *et al.*, 2004).

3.2.9 Expression and purification of mammalian 7C1a

pProEx-HAH6-7C1a$_{263-476}$ in *E. coli* BL21 (DE3) obtained from glycerol stocks (Section 3.1.14) were grown in 100 mL LB broth containing 50 µg/mL ampicillin overnight at 25 °C. The overnight culture was added to 8 x 1 L LB containing 50 µg/mL ampicillin and grown at 25 °C to O.D.$_{600}$ = 0.4. AC protein production was induced with 30 µM IPTG and incubated at 25 °C for 19 h. The culture was harvested by centrifugation at 12,000 g for 10 min at 4 °C, washed once with 20 ml per 1 L culture of bacterial wash buffer (50 mM Tris-HCl pH 8.5, 1 mM EDTA), centrifuged at 5,445 g for 20 min at 4 °C and the bacterial pellet stored overnight at -80 °C. Cells were brought to 4 °C in a water bath then resuspended in 20 mL per original 1 L culture of Solution A (20 mM Tris-HCl pH 8.0, 5 mM 2-ME, 2 mM $MgCl_2$ and 1 mM PMSF). 0.1 mg/mL lysozyme was added and incubated for 30 min at 4 °C. The cells were broken by sonication at 74 W for 90 s at 4 °C. Cell debris was removed by centrifugation at 50,000 g for 30 min at 4 °C. 0.5 mL of Ni-NTA slurry per original 1 L culture and 100 mM NaCl was added to the supernatant and the mixture was incubated for 2 h at 4 °C. The mixture was centrifuged at 405 g for 1 min and the slurry poured into a Polyprep column for washing. The columns were washed with 10x bed volume wash buffer A (Solution A + 400 mM NaCl), 20x bed volume wash buffer B (Solution A + 100 mM NaCl + 20 mM imidazole), and eluted with 6x bed volume elution wash (Solution A + 100 mM NaCl + 150 mM imidazole). The eluates were dialysed against Solution B (50 mM Tris-HCl pH 8.5, 10 mM NaCl, 2 mM $MgCl_2$, 2 mM 1-thioglycerol) and 20% glycerol for 1 h at 4 °C and changed three times. The dialysed protein was concentrated to 1 mL and subjected to anion-exchange chromatography on a Biorad Biologic Duo Flow FPLC system using a Q-Sepharose column (25 x 1 cm). The column was equilibrated with Solution B and the protein was injected into the column. The column was washed with 50 mL Solution B at a flowrate of 2 mL/min and then a linear salt gradient of 100-500 mM NaCl was applied for a total volume of 175 mL. The gradient was increased to 1 M NaCl in a total volume of 50 mM and finally the column was washed with 25 mL Solution B. 5 mL fractions were

collected at the start of the linear salt gradient. Fractions containing pure protein were pooled and concentrated to 5 mg/mL (Yan and Tang, 2002).

3.2.10 Expression and purification of mammalian 2C2a

pQE60-2C2$_{821-1090}$ in *E. coli* M15 (DE3) [pREP4] obtained from glycerol stocks (Section 3.1.14) were grown in 100 mL LB broth containing 50 µg/mL ampicillin and 50 µg/mL kanamycin overnight at 30 °C. The overnight culture was added to 8 x 1 L LB containing 50 µg/mL ampicillin and 50 µg/mL kanamycin and grown at 30 °C to O.D.$_{600}$ = 0.4. AC protein production was induced with 30 µM IPTG and incubated at 25 °C for 15 h. The culture was harvested by centrifugation at 12,000 g for 10 min at 4 °C, washed once with 20 ml per 1 L culture of bacterial wash buffer (50 mM Tris-HCl pH 8.5, 1 mM EDTA), centrifuged at 5,445 g for 20 min at 4 °C and the bacterial pellet stored overnight at -80 °C. Cells were brought to 4 °C in a water bath then resuspended in 20 mL per original 1 L culture of lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM 2-ME, 50 mM NaCl and mixed protease inhibitors). 0.2 mg/mL lysozyme was added and incubated for 30 min at 4 °C. The cells were broken by sonication at 74 W for 90 s at 4 °C. Cell debris was removed by centrifugation at 50,000 g for 30 min at 4 C. 0.5 mL of Ni-NTA slurry per original 1 L culture and 100 mM NaCl was added to the supernatant and the mixture was incubated for 2 h at 4 °C. The mixture was centrifuged at 405 g for 1 min and the slurry poured into a Polyprep column for washing. The columns were washed with 15x bed volume wash 1 (lysis buffer + 2 mM $MgCl_2$ + 400 mM NaCl + 5 mM imidazole), 12x bed volume wash 2 (50 mM Tris-HCl pH 8.0, 10 mM 2-ME, 2 mM $MgCl_2$, and 15 mM imidazole), 8x bed volume wash 3 (50 mM Tris-HCl pH 8.0, 10 mM 2-ME, 10 mM NaCl and 15 mM imidazole) and eluted with 8x bed volume elution wash (50 mM Tris-HCl pH 8.0, 10 mM 2-ME, 2 mM $MgCl_2$, 10 mM NaCl and 150 mM imidazole). The eluates were dialysed against buffer A (50 mM Na-HEPES pH 8.0, 2 mM $MgCl_2$, 1 mM EDTA and 2 mM dithiothreitol) for 1 h at 4 °C and changed three times. The dialysed protein was concentrated to 1 mL and subjected to anion-exchange

chromatography on a Biorad Biologic Duo Flow FPLC system using a Q-Sepharose column (25 x 1 cm). The column was equilibrated with buffer A and the protein was injected into the column. The column was washed with 50 mL buffer A at a flowrate of 2 mL/min and then a linear salt gradient of 0-300 mM NaCl was applied for a total volume of 120 mL. The gradient was increased to 1 M NaCl in a total volume of 50 mM and finally the column was washed with 25 mL buffer A. 5 mL fractions were collected at the start of the linear salt gradient. Fractions containing pure protein were pooled and concentrated to 1 mg/mL (Whisnant *et al.*, 1996).

3.2.11 Adenylyl cyclase assays

AC protein was assayed for specific activity using radiolabelled [α-^{32}P] ATP and measuring the amount of [2,8-^{3}H] cAMP produced over a specific time and with a specific amount of protein. These assays allow for the measurement of specific activity with variation of pH value, temperature, protein concentration and time. To measure the amount of [2,8-^{3}H] cAMP produced, the [2,8-^{3}H] cAMP must be separated from unconverted [α-^{32}P] ATP and counted on a scintillation counter. Running the reaction mixture through a Dowex 50W-X4 resin column and an Aluminium oxide (type WN-3) column separates the [2,8-^{3}H] cAMP for counting (Section 4.4).

Proteins were assayed in a final volume of 100 µL. Reactions typically contained 22% glycerol, 50 mM Tris-HCl pH 7.5, with 2 mM Mn^{2+} as co-factor and 200 µM [α-^{32}P] ATP (25 kBq) and 2 mM [2,8-^{3}H] cAMP (150 Bq) unless otherwise stated. For the pH dependence assay for Rv1625c$_{204-443}$, the buffers used were 50 mM MES (2-morpholinoethanesufonic acid) (pH 6.5) Tris (pH 7.0 and 7.5) and CHES (2-(cyclohexyamino)ethanesulfonic acid) (pH 8.0 and 8.5). Further details are stated in the results. All assays were started by addition of substrate and stopped with stop buffer (1 M Tris-HCl pH 7.5, 3 mM ATP, 3 mM cAMP, 1.5 % (w/v) SDS) (Salomon *et al.*, 1974). Error bars are stated with figure. If an error bar is not seen, the data point is bigger than the error bar. All error bars are standard error of the mean. Assays were separated using a two column chromatography method. Reactions were decanted into columns (0.4 x

15 cm) containing 1 mL Dowex 50W-X4 resin pre-equilibrated with 5 mL H_2O and left to flow under gravity. Columns were washed with 5 mL H_2O and eluted with 2 mL H_2O into columns (0.4 x 15 cm) containing 1 mL Aluminium oxide (type WN-3) pre-equilibrated with 5 mL 20 mM Tris-HCl pH 7.5. Columns were eluted with 3 mL 20 mM Tris-HCl pH 7.5 into 6 mL scintillation vials. 3 mL Ecoscint XR scintillation fluid was added and vials were shaken well and counted with a Packard Liquid Scintillation Analyzer 1600 TR. Data were analysed using Excel and Prism.

3.2.12 Error bars of graphs

All error bars are standard error of the mean (SEM). SEM represents the precision of the mean. The number of replicates (*n*) is stated in the figure. The p-value is calculated by One-way ANOVA with Dunnett's post test, performed using GraphPad Prism version 5.02 for Windows, GraphPad Software, San Diego California USA, www.graphpad.com. If the mean of the data is significantly different compared to the other data ranges, the data are highlighted with (*$p<0.05$).

3.2.13 Preparation of 100 mM $NaHCO_3$ pH 6.5 for AC assays

To add CO_2 to AC assays, 100 mM $NaHCO_3$ was added at pH 6.5, where 38.7 % of the total Ci was CO_2.

100 mM $NaHCO_3$ was prepared to a final volume of 10 mL with the addition of 50 mM MES pH 6.5 at 37 °C. The solution was buffered to pH 6.5 using approximately 3 x 50 µL 1 M HCl and the pH measured (Section 3.2.5). Final pH adjustments were made using 1 M HCl or 1 M NaOH. The buffered solution was used immediately in AC assays.

3.2.14 Preparation of CO_2 saturated H_2O for Ci dis-equilibrium AC assays

Another method to prepare CO_2 for Ci dis-equilbrium AC assays where the added Ci is mostly CO_2 with a low concentration of $NaHCO_3$ in the solution.

50 mL of H_2O on ice was gently bubbled with 10 % (v/v) CO_2 gas for 1 h. The H_2O was saturated with CO_2 after this time and the concentration of CO_2 was approximately 30 mM based on titration against NaOH (data not shown).

3.2.15 β-galactosidase assay

To measure *in vivo* cAMP levels, a plasmid containing *LacZ* gene was transformed into *E. coli* cells containing the plasmid with $Rv1625_{204-443}$ recombinant gene. *In vivo* CO_2 stimulation of Rv1625c would increase cAMP levels. cAMP will bind to the lac operon of *LacZ* and up-regulate the transcription of β-galactosidase. On addition of the substrate ortho-nitrophenol-β-D-galactopyranoside (ONPG), β-galactosidase converts the substrate into ortho-nitrophenol (ONP), a yellow substance which can be quantified.

pCTXP1LacZ was transformed into *E. coli* M15 [pREP4] pQE30-$Rv1625c_{204-443}$ cells. *E. coli* M15 [pREP4] pQE30-$Rv1625c_{204-443}$, pCTXLacZ was grown in 200 mL LB with 100 µg/mL ampicillin, 50 µg/mL kanamycin and 5 µg/mL tetracycline at 30 °C until $O.D._{600}$ = 0.6. *E. coli* M15 [pREP4] pQE30-pCTXP1LacZ was grown in 200 mL LB with 5 µg/mL tetracycline at 30 °C until $O.D._{600}$ = 0.6. pQE30-$Rv1625c_{204-443}$, pCTXP1LacZ and pQE30, pCTXP1LacZ were induced with 30 µM IPTG for 3 h. Cells were centrifuged at 4,000 g for 10 min and resuspended into 192 mL LB Tris-HCl pH 7.1 and aliquoted into 50 mL tubes containing 12 mL LB. Tubes were bubbled with either 10 % (v/v) CO_2 in air or air for 30 min at 30 °C. Cell membranes were disrupted with the addition of 30 µM sodium deoxycholate and 120 µM toluene to 8.4 mL cell culture. Samples were shaken for 10 min at 30 °C. Enzyme assay was performed by adding 2 mL 0.1 M sodium phosphate (pH 7.0) and 0.2 mL 0.01 M ONPG to 2 mL lysed cell culture. Samples were incubated for 15 min at 30

°C. Reactions were stopped with the addition of 0.5 mL 2 M sodium carbonate and centrifuged at 15,700 g for 5 min. Absorbance was read at 420 nm. A standard curve was generated using 0-250 µM ONP and absorbance measured at 420 nm.

3.2.16 CO_2 Binding Assays

To identify CO_2 binding to ACs, radiolabelled $NaH^{14}CO_3$ was added to AC protein. The $^{14}CO_2$ bound protein was separated from unbound $^{14}CO_2$ and $NaH^{14}CO_3$ through gel filtration with G50 sephadex columns.

1 mL (50% w/v) of G50 sephadex was decanted into 0.8 mL Handee Centrifuge columns and centrifuged with cap on at 1,500 g for 30 s. The eluant was discarded. 450 µM protein in 50 mM MES pH 6.5 was incubated for 30 min with 30 mM $NaH^{14}CO_3$ (185 kBq) pH 6.5 in a total volume of 50 µL. The protein-CO_2 mixture was applied to the Handee Centrifuge column and spun at 1,500 g for 30 s into a microcentrifuge tube containing 50 µL 2 M NaOH. 1 mL H_2O and 2 mL scintillation fluid were added to the eluant and the mixture was counted on a scintillation counter (Lorimer, 1979).

3.2.17 Modification of K296C with 2-carboxyethyl methanethiosulfonate

To mimic CO_2 forming a carbamate with K296 of Rv1625c, the mutated K298C was modified with 2-carboxyethyl methanethiosulfonate (Section 6.3) which produced a mimic that would not be stimulated by CO_2.

12 mg $Rv1625c_{204-443}$ K296C (4 mg/mL) was dialysed into 70 mM CHES pH 9.5, 2 mM $MgCl_2$. 50 µL 0.2 M MTSCE was added to the protein and mixed at 25 °C for 2 h. The reaction was monitored by assaying 10 µL protein with 50 µL Ellman's reagent (4 mg/mL in 0.1 M Tris pH 8.0) and measuring the absorbance at 412 nm. A further additon of 50 µL 0.2 M MTSCE (Davis *et al.*, 1999) was added to the mixture to ensure complete modification. To remove excess MTSCE, the mixture was dialysed with dialysis buffer (Section 3.2.6a) at 4 °C for 1 h and changed three times.

3.3 Physical Chemistry

3.3.1 Mass spectroscopy

Mass spectroscopy is an analytical technique that determines the elemental composition of a molecule. It is a qualitative and quantitative technique that identifies the molecular weight of a protein and thus is used as confirmation that the expressed recombinant protein is the correct protein.

100 µg protein in 50 mM Tris-HCl pH 7.0 was analysed by Micromass LCT ES (electrospray)-TOF (time of flight) mass spectroscopy (MS). The sample was placed on HPLC (High Performance Liquid Chromatography) C4 column with a flow rate of 0.2 mL/min. The protein was washed with 0.5 % formic acid in H_2O and then washed with a linear gradient of CH_3CN with 0.5 % formic acid. The protein was eluted with 100 % CH_3CN and loaded onto the mass spectrometer.

3.3.2 CD spectroscopy

CD spectroscopy is a technique that can determine the structure of macromolecules, such as the secondary structure of proteins, through the differential absorption of left and right-handed circularly polarised light.

Protein was dialysed with AC assay dialysis buffer without $MgCl_2$ for 3 h with the buffer changed every hour. ATP was buffered using 50 mM MES pH 6.5. Protein samples were mixed in a microcentrifuge tube with conditions stated in the results to a final volume of 500 µL and pipetted into a 700 µL cuvette. 10 mM $MgCl_2$ was added to both the 20 µM protein sample and the 20 µM protein with 5 mM ATP sample and the final volumes were made up to 510 µL. Circular dichroism (CD) spectra were recorded using a JASCO J-810 Spectropolarimeter scanning at 50 nm s^{-1} with a band pass of 2 nm and response time of 4 s. Points were recorded at 1 nm intervals from 250-350 nm. N_2 gas was pumped through the machine for 1 h prior to use and remained running during experimentation. Spectrum Measurement for Windows was

used to analyse data. Spectra were normalised to 0 degree cm^{-2} dmol $residue^{-1}$ at 350 nm and the blank spectrum was subtracted from all spectra.

3.3.3 FTIR ATR spectroscopy

IR spectroscopy measures the amount of energy if IR light absorbed by a sample with a range of wavelengths. FTIR spectroscopy is a measurement technique that collects numerous IR spectra at once, by IR light passing through an interferometer, which performs a fourier transformation (mathematical transformation). The ATR adapter allows samples to be examined directly in the solid or liquid state without further preparation.

Protein was dialysed with dialysis buffer without $MgCl_2$ and glycerol for 3 h with the buffer changed every hour (Section 3.2.6). Protein samples typically contained as stated in results, 3 mM protein, 20 mM $MgCl_2$, 5 mM ATP pH 6.5 and 3 mM CO_2. The blank spectra samples typically contained as stated in results, 20 mM $MgCl_2$, 5 mM ATP pH 6.5 and 3 mM CO_2. 30 mM CO_2 solution was obtained by bubbling CO_2 gas from dry ice into H_2O for 60 min. Fourier transform infrared (FTIR) spectra were collected using a Perkin Elmer Spectrum 100 FTIR Spectrophotometer with a Miracle ATR (ZnSe) accessory (Pike Technologies Inc.) with 64 scans. Points were recorded at 1 cm^{-1} intervals from 1000-4000 cm^{-1}. Spectrum for Windows was used to analyse the data. Samples were mixed in a microcentrifuge tube to a total volume of 5 µL and pipetted directly onto the ATR crystal. H_2O and protein spectra were subtracted from sample spectra using Spectrum for Windows to obtain new spectra from 2320-2400 cm^{-1} and data were normalised to 100 % transmittance at 2500 cm^{-1}.

3.3.4 HCO_3^- ratiometric europium complex spectroscopy

Luminescence probe spectroscopy is a technique that measures a concentration change of a molecule. Light excitation of the probe bound to the molecule causes an electron to be raised to a higher energy state. A return of

the electron to the original energy state is accompanied by an emission of a photon, which can be measured and quantified.

HCO_3^- ratiometric luminescence spectra were recorded using a Fluorolog 3-11 Spectrophotometer and data were analysed using DataMax for Windows v2.1. Excitation and emission slits were 3 nm and 1.5 nm respectively and the bandpass was 1 nm. An excitation wavelength of 408 nm (sensitisation of the complex via acridone chromophore) was used to obtain the luminescence spectra. Points were recorded at 1 nm intervals from 570-720 nm with a 0.25 s integration time. EuL^4 probe (Bretonniere *et al.*, 2004) was kindly donated by Dr. Robert Pal and Prof. David Parker (Durham University) and was resuspended to 10 mM in H_2O. Samples were made to a final volume of 50 µL and pipetted into a Fluorimeter Sub Micro Cell. A standard curve was generated from 0.1-30 mM $NaHCO_3$ pH 6.5 using 100 µM EuL^4. Protein samples typically contained 500 µM protein, 50 mM MES pH 6.5, 100 µL EuL^4, 100 mM $MgCl_2$ and 10 mM ATP pH 6.5. Ratios of intensity calculated at wavelengths 616 nm and 592 nm and at wavelengths 616 nm and 712 nm were used to obtain the quantitative HCO_3^- ratios.

3.3.5 Crystallisation of Rv1625c$_{212-443}$ K296E/F363R/D365C protein

Protein X-ray crystallography is a method to determine the arrangement of a protein within the protein crystal. By measuring the 2D X-ray diffraction pattern of the protein crystal at a synchrotron, a 3D model of the electron density of the protein can be solved.

Crystallisation trials were set up using the hanging-drop vapour-diffusion method. Screens were performed using 24 well VDX plates with siliconised glass circle cover slides. The protein was diluted to 30 mg/mL in 20 mM HEPES)-NaOH pH 7.5 containing 5 mM 2-ME and 5 % (v/v) glycerol. The plate was set up using 5–10 % polyethylene glycol (PEG) 8000 in 0.1 M Tris pH 8.5 and 0–6 % glycerol in 500 µL reservoir solution. A second plate was set up using 5–10 % PEG 6000 in 0.1 M Tris pH 8.0 and 0–6 % (v/v) glycerol in 500 µL reservoir solution. The crystals were set up by mixing 1 µL protein with 1 µL reservoir solution and 2 µL protein with 1 µL reservoir solution (Ketkar *et*

al., 2004). Crystals appeared within 7 days at 20 °C and were left to grow for a further 1 month. Optimal sized crystals were cryo-protected by consecutively coating the crystal with reservoir solution containing 5–30 % glycerol. Cryo-protected crystals were immediately frozen in liquid nitrogen. The frozen crystals were stored -140 °C before transport to a synchrotron for X-ray diffraction.

3.3.6 Exposure of Rv1625c$_{212-443}$ K296E/F363R/D365C protein crystals to CO_2

Optimal sized crystals from the crystallisation trials where chosen to expose to CO_2. The crystals were cryo-protected by consecutively coating the crystal with reservoir solution containing 5–30 % glycerol 3-4 times. The cryo-protected crystal was immediately placed in a CO_2 chamber (Figure 3.2) and exposed to between 3–5 bar CO_2 for between 2–10 minutes. The crystal was then immediately removed from the CO_2 chamber and frozen in liquid nitrogen. The frozen CO_2 exposed crystals were stored at -140 °C before transport to a synchrotron for X-ray diffraction.

3.3.7 X-ray diffraction and data collection of Rv1625c$_{212-443}$ K296E/F363R/D365C protein crystals

X-ray diffraction data of CO_2 exposed protein crystals were collected at beam line X06DA equipped with a Mar225 CCD detector at the Swiss Light Source (Paul Scherrer Institut, Villigen, Switzerland). Data were processed and scaled with HKL2000 (Otwinowski and Minor, 1997). The structure was solved with molecular replacement using the CCP4 programme package (Collaborative Computational Project, 1994) and a 3D model was generated (Section 6.6).

Figure 3.2: CO_2 chamber used for exposing protein crystals to CO_2. Cyro-protected protein crystals were exposed to between 3–5 bar CO_2 for between 2–10 minutes prior to being frozen in liquid nitrogen and transported to a synchrotron.


1-20(µl Natural Pipette Tip

3.4 Table of DNA primers.

Protein	Primer	Sequence
Rv1625c$_{204-443}$		
R310A	MJC233	GTTGTCAGCGGCGTTCCGGCGCCCCGGCCTGACCATACG
	MJC234	CGTATGGTCAGGCCGGGGCGCCGGAACGCCGCTGACAAC
H315A	MJC231	CCGCGGCCCCGGCCTGACGCGACGCAAGCACTGGCGGAC
	MJC232	GTCCGCCAGTGCTTGCGTCGCGTCAGGCCGGGGCCGCGG
K334A	MJC223	AATGTCGCAGCGCAATTGGCGGATCCACGCGGCAACCCG
	MJC224	CGGGTTGCCGCGTGGATCCGCCAATTGCGCTGCGACATT
R337A	MJC235	GCGCAATTGAAGGATCCAGCGGGCAACCCGGTGCCGCTG
	MJC236	CAGCGGCACCGGGTTGCCCGCTGGATCCTTCAATTGCGC
R344A	MJC237	GGCAACCCGGTGCCGCTGGCGGTGGGCCTGGCCACCGGC
	MJC238	GCCGGTGGCCAGGCCCACCGCCAGCGGCACCGGGTTGCC
R360A	MJC239	GCGGGTGTGGTGGGTTCTGCGCGGTTCTTCTACGACGTG
	MJC240	CACGTCGTAGAAGAACCGCGCAGAACCCACCACACCCGC
R361A	MJC241	GGTGTGGTGGGTTCTCGAGCGTTCTTCTACGACGTGTGG
	MJC242	CCACACGTCGTAGAAGAACGCTCGAGAACCCACCACACC
V366A	MJC262	CGACGGTTCTTCTACGACGCGTGGGGCGACGCGGTCAAT
	MJC263	ATTGACCGCGTCGCCCCACGCGTCGTAGAAGAACCGTCG
R395A	MJC243	CCAGACGAGGTTTACGAGGCGCTCAAGGACGACTTCGTG
	MJC244	CACGAAGTCGTCCTTGAGCGCCTCGTAAACCTCGTCTGG
K397A	MJC225	GAGGTTTACGAGCGTCTCGCGGACGACTTCGTGTTGCGC
	MJC226	GCGCAACACGAAGTCGTCCGCGAGACGCTCGTAAACCTC
K405A	MJC245	GACTTCGTGTTGCGCGAGGCGGGCCACATAAACGTCAAG
	MJC246	CTTGACGTTTATGTGGCCCGCCTCGCGCAACACGAAGTC
K411A	MJC227	CGCGGCCACATAAACGTCGCGGGGAAAGGCGTAATGCGC
	MJC228	GCGCATTACGCCTTTCCCCGCGACGTTTATGTGGCCGCG
K413A	MJC229	CACATAAACGTCAAGGGGGCGGGCGTAATGCGCACGTGG
	MJC230	CCACGTGCGCATTACGCCCGCCCCCTTGACGTTTATGTG
Slr1991$_{120-337}$		
D136A	MJC208	ATTACGGTGTTGGTGGCGGCCATGCGTAACTTCACCGGC
	MJC209	GCCGGTGAAGTTACGCATGGCCGCCACCAACACCGTAAT
D181A	MJC210	GTGGATAGCTACATCGGTGCCGCGGTCATGGCCATCTGG
	MJC211	CCAGATGGCCATGACCGCGGCACCGATGTAGCTATCCAC
Vector		
pQE30	MJC104	GAGCGGATAACAATTTCAC
	MJC108	GTTCTGAGGTCATTACTGG

3.5 Sources of Materials.

All chemicals used were from Sigma Aldrich unless otherwise stated.

All equipment used was from SLS unless otherwise stated.

Mini-prep kit was from Qiagen.

Restriction enzymes, restriction buffers, DNA ladders and protein ladders were from Fermentas

Ligation enzymes, ligation buffers, DNA polymerase, PCR buffers and trypsin were from Promega.

dNTPs and agarose were from Bioline.

TOPO cloning kit and His-tag In-gel Stain were from Invitrogen.

Cryovial tubes were from Simport.

Polyprep columns and Dowex resin and Prep-a-gene matrix were from Biorad

Radiochemicals were from Amersham and Perkin Elmer.

Scintillation vials were from Starstedt.

Scintillation fluid was from National Diagnostics.

Fluorimeter Sub Micro Cell was from Starna.

Handee Centrifuge columns were from Pierce.

Q-Sepharose and Superdex columns were from GE Healthcare.

MTSCE was from Toronto Research Chemicals.

VDX plates and cover slides were from Hampton Research.

4 Investigation into *in vitro* and *in vivo* inorganic carbon stimulation of $Rv1625c_{204-443}$

4.1 Introduction

Class IIIb ACs were previously thought to be the only sub-group of ACs responsive to Ci and the Ci form that caused this response was HCO_3^- (Cann *et al.*, 2003). Experiments conducted on Class IIIa ACs for Ci sensitivity showed a non-response to Ci (Chen *et al.*, 2000; Cann *et al.*, 2003). Further work has identified that the Ci species that caused the Ci response in Class IIIb ACs is CO_2 (Hammer *et al.*, 2006). As these two biologically active forms of Ci are in an equilibrium that is affected by pH, a further investigation into the response of Class IIIa ACs to CO_2 should be conducted at reduced pH, where the total percentage of CO_2 in Ci is greater. It is important to identify a potential Ci sensitivity in Class IIIa ACs as all mammalian tmACs and a large number of prokaryotic ACs are in this sub-group and a Ci sensitivity in Class IIIa ACs would suggest that these ACs are involved in CO_2 sensing in mammals and prokaryotes.

The catalytic domain of Rv1625c, a Class IIIa AC from *M. tuberculosis* has been previously expressed and characterised as a recombinant protein (Guo *et al.*, 2001) and was therefore investigated for Ci sensitivity. There is a link between elevated cAMP levels and virulence (Lowrie *et al.*, 1975) seen *in vivo*, which underpins the importance of identifying the role of ACs in pathogenesis of *M. tuberculosis*. Thus if Rv1625c is a CO_2 sensor, then it may be involved in *M. tuberculosis* pathogenesis as increased CO_2 concentrations are found at the site of *M. tuberculosis* infections.

4.2 Purification and analysis of $Rv1625c_{204-443}$

An alignment of the catalytic domain of Class III a-d ACs (Figure 4.1) shows the conserved residues required for catalysis. The T to D polymorphism of the Class IIIb ACs and Class IIIa ACs was thought to confer HCO_3^- sensitivity to the AC, hence Class IIIa ACs were thought to be non-responsive to HCO_3^- (Cann *et al.*, 2003; Linder and Schultz, 2003).

The catalytic domain of the Class IIIa AC from *M. tuberculosis* $Rv1625c_{204-443}$ was overexpressed in *E. coli* and purified by affinity chromatography to homogeneity to generate a recombinant protein. The protein was analysed for purity by MS and SDS-PAGE.

The SDS-PAGE gel of the expressed and purified $Rv1625c_{204-443}$ recombinant protein (Figure 4.2A(1)) shows a single protein band at approximately 30 kDa. The theoretical weight of $Rv1625c_{204-443}$ is 26790 Da. The His-tag stained gel (Figure 4.2A(2)) suggests that the protein band is the correctly expressed protein as there is a His-tag attached to the protein and it has not been proteolysed at the N-terminus.

The MS data (Figure 4.2B) confirms that the actual molecular weight of the purified protein matches the theoretical molecular weight of 27690 Da for $Rv1625c_{204-443}$ and proved that the purified protein is $Rv1625c_{204-443}$. The data also shows that there was no contaminating protein.

4.3 Control pH assays for AC assays

Class IIIb ACs were previously thought to be stimulated by HCO_3^-, however we have found that they are actually stimulated by CO_2. This suggested that Class IIIa ACs, previously thought to be non-responsive to HCO_3^-, may respond to CO_2 instead. To assess whether $Rv1625c_{204-443}$, a Class IIIa AC, might respond to CO_2, a pH dependence assay was conducted. The percentage of Ci in the form of CO_2 and HCO_3^- varies depending on the pH of the solution (Figure 4.3).

Figure 4.1: Alignment of the catalytic domains of Class III ACs and comparison of their adenine binding mechanisms. The Class IIIa AC is Rv1625c of *M. tuberculosis* H37Rv, the Class IIIb ACs are CyaB1 of *Anabaena* PCC 7120, Slr1991 of *Synechocystis* PCC 6803 and human sAC. The Class IIIc AC is Rv1624 of *M. tuberculosis* H37Rv. The Class IIId AC is CyaA of *Leishmania donovani*. Numbers denote amino acid sequence number. Arrows indicate conserved metal binding Asp residues. Triangle indicates conserved substrate binding Lys residue. Square indicates conserved transition state binding Arg residue. Circle indicates the polymorphic Asp/Thr of Class IIIa/b ACs.

```
                                      ↓                                                 ▼  ↓
a Mycobacteria  Rv1625c   252 VLFADIVGFTERASST........APADLVRFLDRLYSAFDELVDQHGLEKIKVSGDSYM
  Anabaena      CyaB1     602 VLFSDIRGYTTLTENL........GAAEVVSLLNQYFETMVEAVFNYEGTLDKFIGDALM
b Synechocystis Slr1991   132 VLVADMRNFTGMAQQV........EEELLSMLIGNWFRQAGHILREAGSWVDKYIGDAVM
  Human         sAC1       43 LMFVDISGFTAMTEKFSTAMYMDRGAEQLVEILNYYISAIVEKVLIFGGDILKFAGDALL
c Mycobacteria  Rv1264    218 VAFADLVGFTQLGEVV........SAEELGHLAGRLAGLARDLTAPPVWF.IKTIGDAVM
d Leishmania    CyaA      934 LLFTDIESSTALWAAL.........PQLMSDAIAAHHRVIRQLVKKYGCYEVKTIGDSFM


a Mycobacteria  Rv1625c   304 VV........................SGVPRPRPDHTQALADFALDMTNVAAQLKDPRGN
  Anabaena      CyaB1     654 AV..........................FGAPLPLTENHAWQAVQSALDMRQRLKEFNQR
b Synechocystis Slr1991   185 AI......................WFHGYNEATPAEIIQILHAVNRLQAMTAKLNQ....
  Human         sAC1      103 AL......................W........KVERKQLKNIITVVIKCSLEIHGLFEA
c Mycobacteria  Rv1264    269 LV........................CPDPAP.......LLDTVLKLVEVV....DTDNN
d Leishmania    CyaA      986 IACRSAHSAVSLACEIQTKLLKHDW..GTEALDRAYREFELARVDTLDDYEPPTARLSEE


                                                               ●               ■
a Mycobacteria  Rv1625c   340 .....PVPLRVGLATG...PVVAGVVGSRRFFYDVW.GDAVN...VASRMEST
  Anabaena      CyaB1     688 RIIQPQPQIKIGIGISSGEVVSGNIGSHKRMDYTV.IGDGVN...LSSRLETV
b Synechocystis Slr1991   219 KYELPFP.LRIGTGINTGYAMVGNTGSGDHPDYTA.IGDTVN...AAFRLESA
  Human         sAC1      133 KEVEEGLDIRVKIGLAAGHITMLVFGDETRNYFLV.IGQAVDD..V..RLAQN
c Mycobacteria  Rv1264    694 .....FPRLRAGVASG...MAVSRAGD........WFGSPVN...VASRVTGV
d Leishmania    CyaA     1044 EYAALWCGLRVRVGIHTGLTDIRYDEVTKGYDY...YGDTSN...MAARTEAV
```

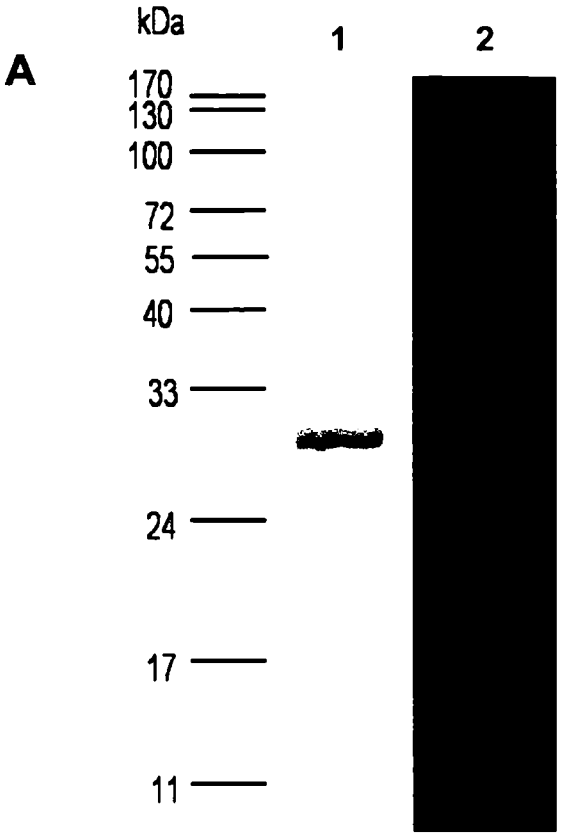
At pH 6.5, the percentage of Ci in the form of CO_2 is 38.7% and the percentage of Ci in the form of HCO_3^- is 61.3%. However raising the pH to 7.5 decreases the percentage of Ci in the form of CO_2 to 5.9% and increases the percentage of Ci in the form of HCO_3^- to 93.9%. Thus by lowering the pH value of Ci, the percentage of Ci in the form of CO_2 can be selectively increased in comparison to HCO_3^-.

Since pH played an important role in determining the percentage of CO_2 or HCO_3^-, it is important that during the AC assays the pH of the mixture is stable and accurate. To demonstrate that the pH at various pH values remained constant even with the addition of Ci, pH control assays were performed using a mock AC assay (without enzyme added) and measuring the pH of the solution every minute for 70 minutes.

The pH controls for the AC assays (Figure 4.4) demonstrates that at the five pH values, pH 6.5–8.5, the pH of the assay remained stable during the 30 minute pre-incubation period, the addition of substrate, and the 40 minute assay period. The pH value of the assay was also similar both in the presence of $NaHCO_3$ and NaCl. At all pH values, the pH value decreased during the first three minutes. This was because the assay mixture was kept at 0 °C until it was incubated in a 37 °C water bath. On increasing temperature, buffered solutions decrease in pH.

It can therefore be assumed that any changes in the specific activity of $Rv1625c_{204-443}$ in the assay will be due to the response of the enzyme to Ci and not a change in the pH value.

Figure 4.2: A) 1) SDS-PAGE gel of $Rv1625c_{204-443}$. 2 µg of protein was applied to each lane with molecular weight standards on the left of each gel and the gel was stained with Coomassie Brilliant Blue G. 2) His-tag stained SDS-PAGE gel of $Rv1625c_{204-443}$. B) MS data for $Rv1625c_{204-443}$.


A
kDa
1
2
170
130
100
72
55
40
33
24
17
11

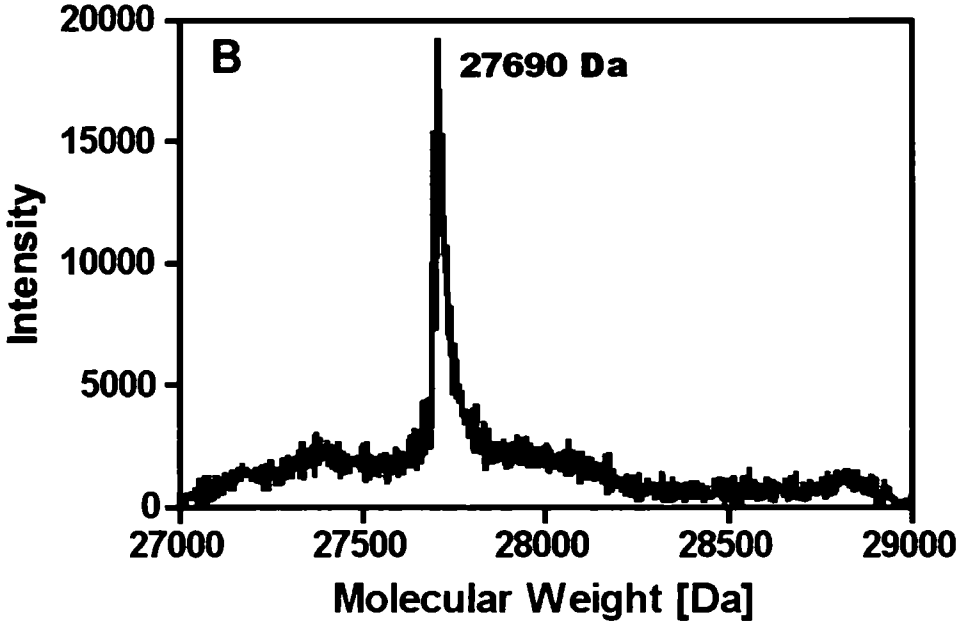
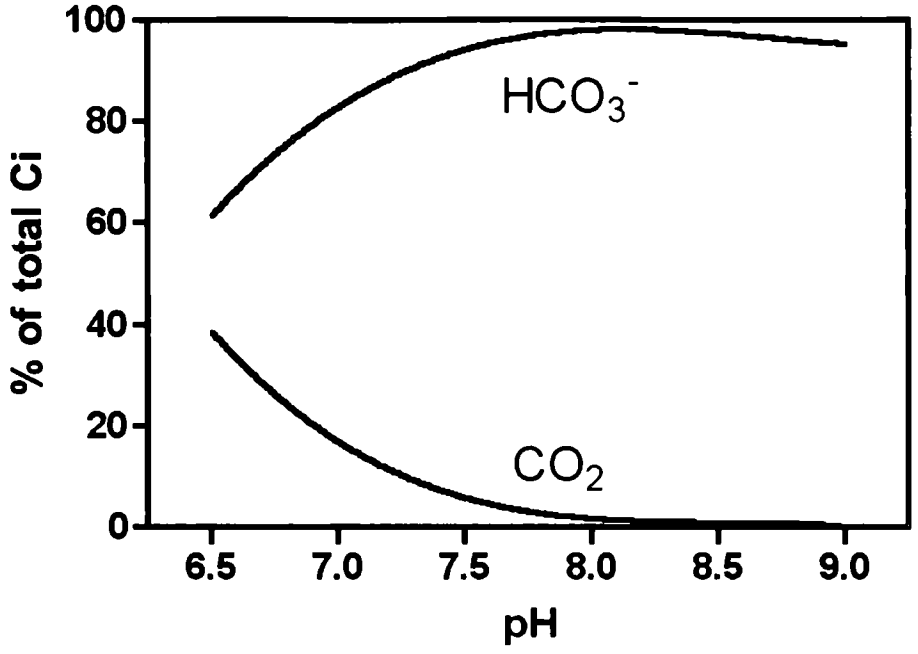

B
27690 Da
Intensity
20000
15000
10000
5000
0
27000
27500
28000
28500
29000
Molecular Weight [Da]

Figure 4.3: Percentage of total Ci in the form of CO_2 and HCO_3^- at different pH values.

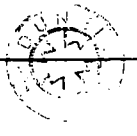

100
80
60
40
20
0
% of total Ci
HCO_3^-
CO_2
6.5
7.0
7.5
8.0
8.5
9.0
pH

4.4 AC assays

The AC assays were all conducted using similar methodologies with minor changes, depending on the assay, that are mentioned in the figure captions. The assay mixture typically contained 50 mM buffer, 2 mM metal (usually $MnCl_2$), salt (usually 20 mM NaCl/$NaHCO_3$) 22 % glycerol and protein. The concentration of protein added to the assay was adequate to maintain a substrate conversion of < 10 %, so as to reduce the inhibitory effects of the formed products on the AC. The assays were incubated for 30 minutes in a water bath (usually at 37 °C), then removed from the water bath and placed on ice. 200 µM ATP (spiked with 25kBq [α-^{32}P] ATP) was added to the assay mixture and immediately returned to the water bath for the assay time (usually 40 minutes). The assay was stopped with the addition of stop buffer that contained SDS to disrupt the protein and ATP to compete with [α-^{32}P] ATP for catalysis. A tracer amount of [2,8-^{3}H] cAMP (150 Bq) was added and used as a control for calculating recovery of [2,8-^{3}H] cAMP. The reactions were decanted into Dowex 50W-X4 columns which bound [α-^{32}P] ATP and not [2,8-^{3}H] cAMP. [2,8-^{3}H] cAMP was eluted off the Dowex columns into aluminium oxide columns which further purified the [2,8-^{3}H] cAMP from eluted [α-^{32}P] ATP. Ecoscint XR scintillation fluid was added to the eluted [2,8-^{3}H] cAMP and counted on a scintillation anaylzer.

4.5 pH dependence assay of Rv1625c$_{204-443}$ in the presence of Ci

Having completed the control pH assays for the AC assays and finalised the conditions for the AC assay, a series of pH dependence assay ranging from pH 6.5 to 8.5 was performed to ascertain the effect of Ci on the specific activity of Rv1625c$_{204-443}$ at various pH values. The graph (Figure 4.5) shows that in the presence of NaCl, the optimal pH value was approximately 8.0, however in the presence of $NaHCO_3$ the optimal pH value was 6.5. Interestingly, there was a 5-fold stimulation in specific activity due to the



presence of Ci that decreased with increasing pH until the stimulation disappeared at pH 7.5 and higher. This finding may explain why $Rv1625c_{204-443}$ was previously shown to be non-responsive to Ci as the assays were done at pH 7.5 where there was no observable Ci stimulation of the enzyme (Cann *et al.*, 2003).

There are three possible explanations for the Ci stimulation of $Rv1625c_{204-443}$ at pH 6.5 and pH 7.0. Firstly, the effect may be due to the cation, Na^+ stimulating the enzyme at lower pH values. This hypothesis can be tested by running a cation dependence assay with $Rv1625c_{204-443}$ and using a different cation, such as K^+ to observe its effect on $Rv1625c_{204-443}$ at low pH. Secondly, the effect may be due to the anion, HCO_3^- being better able to ionically interact at lower pH by stabilising the substrate or the active site of the enzyme and thus increasing specific activity. Thirdly, the stimulation may be due to the other biologically active form of Ci, CO_2 that forms 38.7 % (7.74 mM) of the total Ci at pH 6.5 and 16.6 % (3.32 mM) of the total Ci at pH 7.0. This hypothesis may be tested by performing an assay where predominately HCO_3^- or CO_2 was present.

4.6 Cation dependence assay of $Rv1625c_{204-443}$ at pH 6.5

To test the hypothesis that the cation may be stimulating $Rv1625c_{204-443}$, a salt dependence assay was conducted using sodium salts and potassium salts at pH 6.5. Figure 4.6 shows that the presence of NaCl or KCl did increase the specific activity of the enzyme, probably due to the stabilisation of the protein structure by interactions with the salt and thus increasing enzyme turnover. However, $NaHCO_3$ and $KHCO_3$ both stimulated the enzyme 4-fold as compared to their associated Cl^- salts, thus proving that the stimulation was independent of the cation used and that Ci dose can cause the stimulation. The specific activities of $Rv1625c_{204-443}$ in the presence of NaCl and $NaHCO_3$ were comparable with the values at pH 6.5 of the pH dependence assay (Figure 4.5). The first theory of cation dependence has been proved to cause a small fold stimulation compared to the stimulation observed with the addition of Ci, and it would be an interesting experiment to fully investigate this effect.

Figure 4.4: pH control figure for AC assays measured over 70 minutes. 200 µM ATP was added after 30 minutes (arrow). pH value was measured every minute. The assay mixture contained 50 mM buffer, 2 mM $MnCl_2$ and either 20 mM $NaHCO_3$ (square) or 20 mM NaCl (triangle). A) 50 mM MES pH 6.5 B) 50 mM Tris pH 7.0 C) 50 mM Tris pH 7.5 D) 50 mM Tris pH 8.0 E) 50 mM Tris pH 8.5.

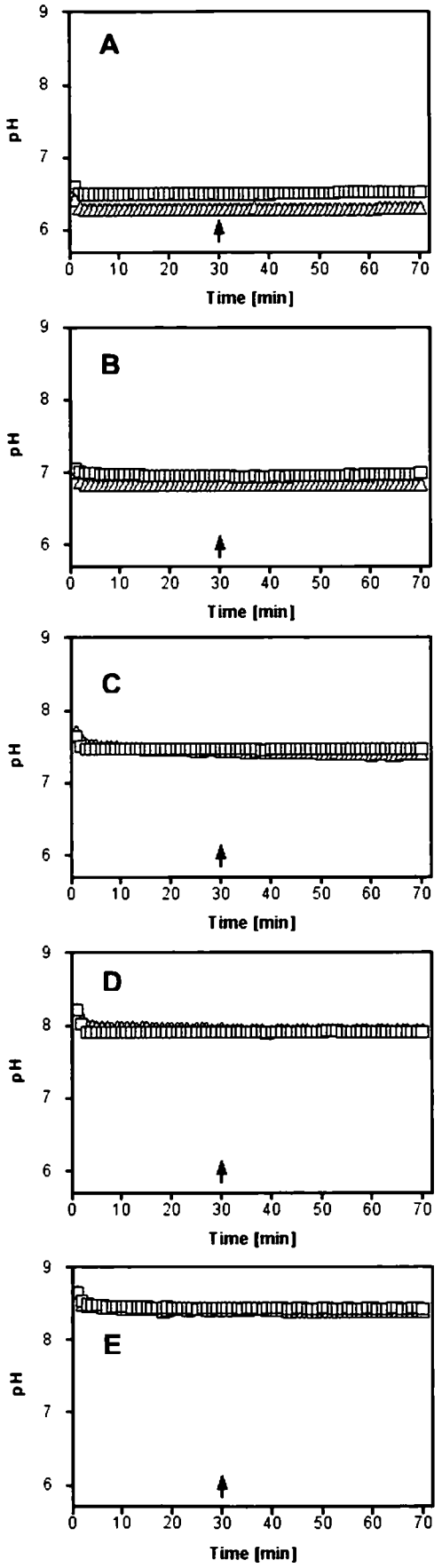

A
B
C
D
E
pH
9
8
7
6
0
10
20
30
40
50
60
70
Time [min]

Figure 4.5: pH dependence assay of $Rv1625c_{204-443}$ in the presence of Ci. 1.81 µM $Rv1625c_{204-443}$ was assayed at 37 °C at varying pH values in the presence of 200 µM ATP and 2 mM $MnCl_2$ for 40 minutes with a pre-incubation time of 30 minutes. The assay contained either 20 mM $NaHCO_3$ (square) or 20 mM NaCl (triangle). Error bars are standard error of the mean (*n*=8).

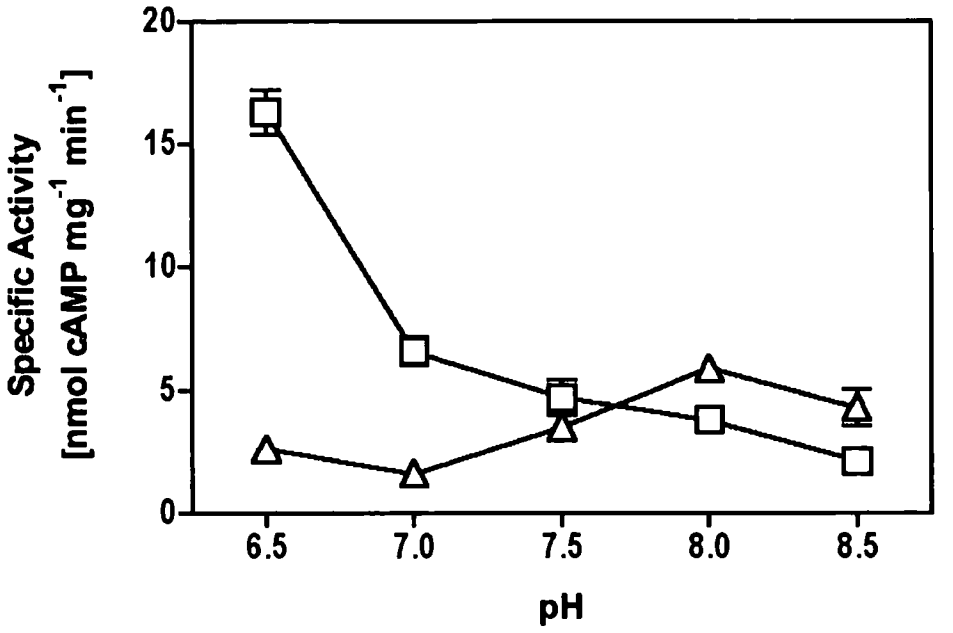
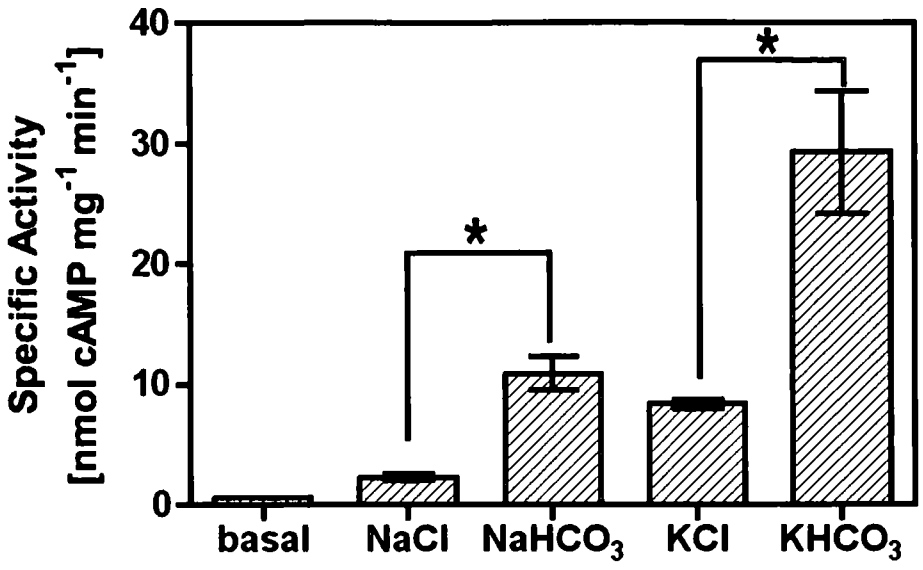

Specific Activity
[nmol cAMP mg-1 min-1]
20
15
10
5
0
6.5
7.0
7.5
8.0
8.5
pH

Figure 4.6: Cation dependence assay of $Rv1625c_{204-443}$ at pH 6.5. 433 nM $Rv1625c_{204-443}$ was assayed at 37 °C at pH 6.5 (50 mM MES) with 2 mM Mn^{2+} in the presence of 200 μM ATP and 20 mM salt for 40 minutes with a pre-incubation time of 30 minutes. Error bars are standard error of the mean (n=6) (*$p<0.05$).


Specific Activity
[nmol cAMP mg-1 min-1]
40
30
20
10
0
*
*
basal
NaCl
NaHCO3
KCl
KHCO3

However the explanation of Ci stimulation was investigated, with two theories left to test, the theory that HCO_3^- may stabilise the protein and cause the stimulation and the theory that CO_2 may be the Ci form causing the stimulation at pH 6.5 in Figure 4.5.

4.7 Ci dis-equilibrium assays

An assay with Ci in dis-equilibrium (Cooper and Filmer, 1969) would identify which form of Ci caused the stimulation of $Rv1625c_{304-443}$ as the added Ci would be predominately in the form of HCO_3^- or CO_2. Thus if HCO_3^- stimulated $Rv1625c_{204-443}$ then the second theory of stabilisation of the protein at lower pH is correct and if CO_2 stimulated $Rv1625c_{204-443}$ then the third theory is correct. The Ci dis-equilibrium was achieved by slowing the acquisition of the Ci equilibrium in the assay mixture so that $Rv1625c_{204-443}$ would be subjected to predominately HCO_3^- or CO_2, depending on which Ci form was added to the assay and there were two methods for accomplishing this. Firstly, lowering the temperature of the assay slowed the acquisition of the Ci equilibrium. This was done by running the assay at 0 °C compared with the normal assay temperature of 37 °C. Secondly, the assay time was reduced so that the Ci equilibrium would not be reached within the shortened assay time and $Rv1625c_{204-443}$ would be subjected to Ci that was predominately in the form added. To identify the length of the shortened time period for the assay, dis-equilibrium pH control assays at 0 °C were initially conducted. These dis-equilibrium pH control assays at 0 °C were conducted by adding either HCO_3^-, CO_2 or NaCl, to the assay mixture that contained only 5 mM buffer so that the acquisition of Ci equilibrium could be monitored with a pH probe every second for 80 seconds. Thus a time period whereby the Ci equilibrium had not been reached was chosen where the Ci was predominately in the form added. The Ci equilibrium was observed through a change in the pH of the weakly buffered solution as an increase in pH corresponded with the loss of H^+ from the solution due to the formation of CO_2 and a decrease in pH corresponded with the release of H^+ from the solution due to the formation of HCO_3^+. Carbonic anhydrase (CA), a zinc-metalloenzyme that speeds up the acquisition of the Ci

equilibrium was added to assays to demonstrate the complete acquisition of the Ci equilibrium, as observed by a steady pH value. CA has a rate constant for converting CO_2 to HCO_3^- of between 10^4 and 10^6 reactions per second (Chirica, 1997). Thus the limiting factor of the acquisition of the Ci equilibrium will be the diffusion rate of CO_2 to the active site of CA.

Figure 4.7A shows the Ci dis-equilibrium pH control assay with the addition of 20 mM $NaHCO_3$ after 20 seconds (arrow). In the presence of CA, the Ci equilibrium was observed to be reached 3 seconds after addition of $NaHCO_3$, as demonstrated by a stable pH value. In the absence of CA, the Ci equilibrium was observed to be almost reached 60 seconds after addition of $NaHCO_3$. The addition of $NaHCO_3$ increased the pH of the weakly buffered solution because CO_2 was formed from the added HCO_3^- to maintain the Ci equilibrium, which removed an H^+ ion from the buffer to form H_2O thus increasing pH.

Figure 4.7B shows the Ci dis-equilibrium pH control assay with the addition of 20 mM CO_2 after 20 seconds (arrow). In the presence of CA, the Ci equilibrium was observed to be reached 3 seconds after addition of CO_2, seen again by a stable pH value. In the absence of CA, the Ci equilibrium was observed to be reached 60 seconds after addition of CO_2. The addition of CO_2 decreased the pH of the weakly buffered solution because HCO_3^- was formed from the added CO_2 to maintain the Ci equilibrium, which added an H^+ ion to the buffer thus decreasing pH.

Figure 4.7C shows the Ci dis-equilibrium pH control assay with the addition of 20 mM NaCl after 20 seconds (arrow). Both in the presence and absence of CA there was no change in pH. This control assay showed that the changes in pH were not due to a dilution effect from the addition of a solution to the assay mixture.

From these Ci dis-equilibrium pH control assays a time period of 10 seconds was chosen whereby the Ci equilibrium was not reached in the assay mixture. Thus the Ci in the assay mixture was predominately in the form in which it was added and the Ci form that stimulated $Rv1625c_{204-443}$ should be identifiable.

Figure 4.7: Ci dis-equilibrium pH control assays. pH curves of (A) HCO_3^-, (B) CO_2 and (C) Cl^- in 5 mM MES pH 6.4 in the presence and absence of carbonic anhydrase (CA) at 0 °C. 20 mM $NaHCO_3/CO_2/NaCl$ was added after 20 seconds (arrow) and pH was measured every second for a further 60 seconds. The assay either had 132 U CA (square) or no CA (triangle). (D) pH curve of $Rv1625c_{204-443}$ in 100 mM MES pH 6.5 at 0 °C with the addition of either 20 mM $NaHCO_3$ (square), CO_2 (circle) or NaCl (triangle) after 5 seconds (arrow) and pH was measured every second for a further 10 second.

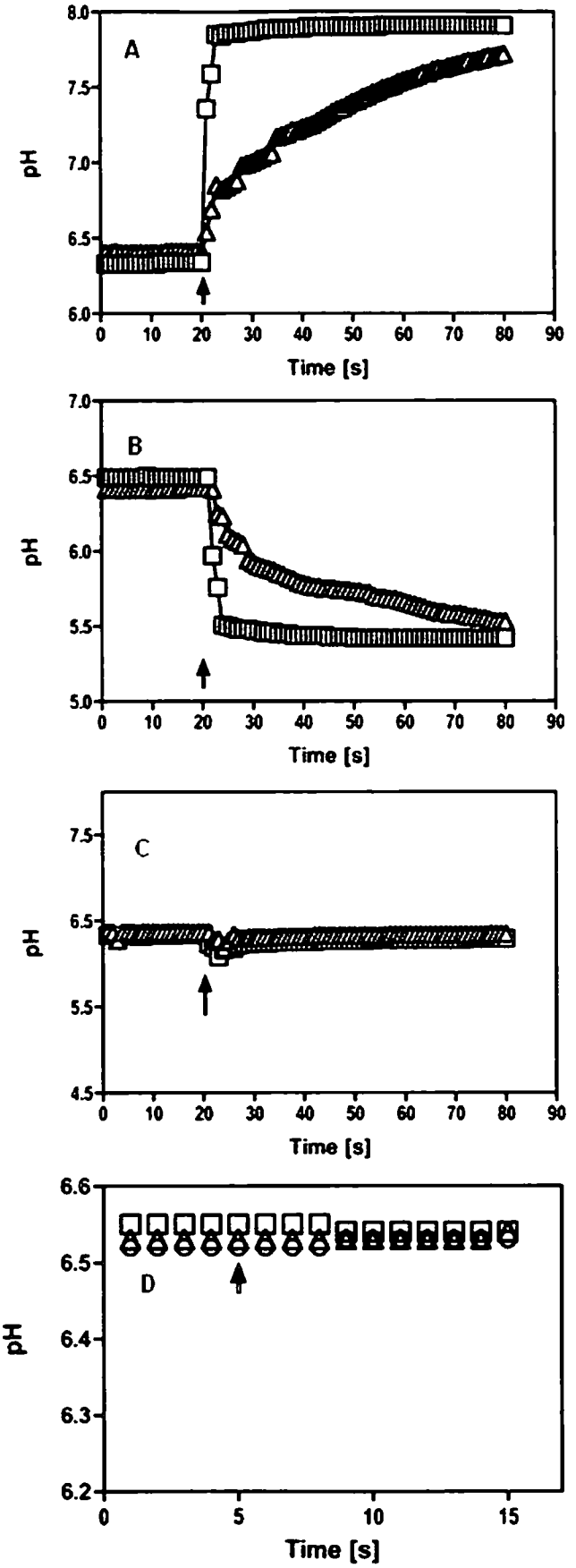
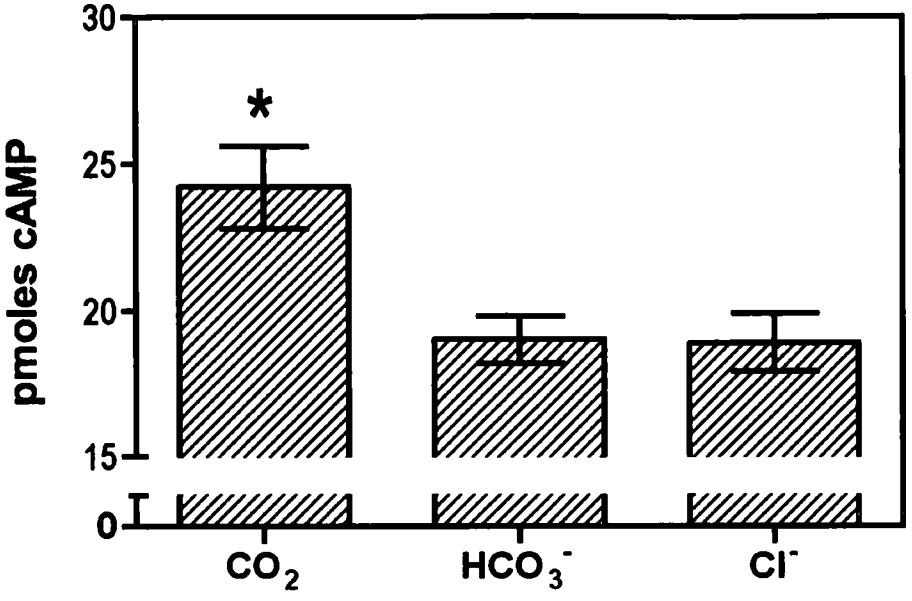

A
pH
8.0
7.5
7.0
6.5
6.0
0 10 20 30 40 50 60 70 80 90
Time [s]
B
pH
7.0
6.5
6.0
5.5
5.0
0 10 20 30 40 50 60 70 80 90
Time [s]
C
pH
7.5
6.5
5.5
4.5
0 10 20 30 40 50 60 70 80 90
Time [s]
D
pH
6.6
6.5
6.4
6.3
6.2
0 5 10 15
Time [s]

Figure 4.8: Ci dis-equilibrium assay of $Rv1625c_{204-443}$. 36 µM $Rv1625c_{204-443}$ was assayed at 0 °C at pH 6.0 (100 mM MES) with 2 mM Mn^{2+} in the presence of 100 mM NaCl, 200 µM ATP for 10 seconds and the addition of 20 mM $CO_2/HCO_3^-/Cl^-$. Error bars are standard error of the mean (n=20) (*$p<0.05$).


30
25
20
15
0
pmoles cAMP
*
CO_2
HCO_3^-
Cl^-

The final pH control assay to perform was to determine the buffer concentration required to maintain the assay pH at 6.5 with the addition of Ci.

This was done so that the stimulation observed in the Ci dis-equilibrium assay was due to the added Ci and not a change in pH. Figure 4.7D shows that with 100 mM MES the pH of the assay is stable for 10 seconds after the addition of Ci and NaCl. The final conditions that will be used for the 10 second Ci dis-equilibrium assay are much changed from the usual AC assay for $Rv1625c_{204-443}$, which demonstrates the difficulty in identifying the Ci form that is causing the stimulation. With the many alterations in conditions for the 10 second assay, the specific activity of $Rv1625c_{204-443}$ obtained from the 10 second assay may not be comparable to its specific activity in other assays.

Figure 4.8 shows the results of the 10 second Ci dis-equilibrium assay of $Rv1625c_{204-443}$. The figure shows that $Rv1625c_{204-443}$ was stimulated by the addition of CO_2 and not HCO_3^- as compared with the NaCl control. The units of the assay were not standard specific activity (nmol cAMP/mg/min) due to two reasons: a large concentration of protein was required for the assay to work and a short time period was used to maintain Ci dis-equilibrium. Therefore converting the raw data into specific activity would not produce a useful value. Instead, the amount of product formed was used for the figure. The stimulation of $Rv1625c_{204-443}$ was only 1.25-fold, much smaller than the 5-fold stimulation seen in the pH dependence assay (Figure 4.5) and in the salt dependence assay (Figure 4.6). This result may be due firstly, to the low assay temperature, low pH of the solution and high buffer concentration used in the assay that were outside of the enzyme's optimal conditions thus adversely affecting $Rv1625c_{204-443}$. Secondly, the result may be due to the high protein concentration and high salt concentration used in the assay to make it work. As the assay conditions were far from optimal, more protein was required to measure the activity of $Rv1625c_{204-443}$ and more salt was required to stabilise the large amount of protein. These high concentrations have unknown effects on the protein. Thirdly, the result may be due to the short assay period. The activity of $Rv1625c_{204-443}$ is unlikely to be in a linear relationship with time. In addition, CO_2 stimulation may be time dependent, explaining the low fold stimulation. The most probable answer is that the reason for the low fold

stimulation with the 10 second Ci dis-equilibrium assay is a combination of all the reasons stated.

This set of results undoubtedly proved that CO_2 was the Ci form that stimulated $Rv1625c_{204-443}$ and demonstrated that $Rv1625c_{204-443}$ was the first Class IIIa AC to be identified as being stimulated by CO_2.

This finding compares similarly with Class IIIb ACs that are stimulated by CO_2 (Hammer *et al.*, 2006) and raises questions about whether other Class IIIa ACs are also stimulated by CO_2. Since Class IIIa and IIIb ACs are stimulated by CO_2, Class IIIc and IIId ACs may also be stimulated by CO_2. Characterisation of the CO_2 stimulation of $Rv1625c_{204-443}$ was conducted to ascertain any potential physiological relevance of the CO_2 response.

4.8 CO_2 dose dependence assays on $Rv1625c_{204-443}$

The CO_2 stimulation was characterised biochemically by identifying the $E.C._{.50}$ value for CO_2 stimulation of $Rv1625c_{204-443}$ and the kinetic parameters for CO_2 stimulation on $Rv1625c_{204-443}$. Standard AC assay conditions were used, with an assay time of 40 minutes.

Figure 4.9 shows the CO_2 dose dependence assay of $Rv1625c_{204-443}$ at pH 6.5 from 0 to 11.6 mM CO_2. This assay was performed in order to ascertain whether the CO_2 stimulation of $Rv1625c_{204-443}$ was a physiologically relevant activation by measuring whether the CO_2 stimulation of $Rv1625c_{204-443}$ was within the values of physiological CO_2 concentrations in prokaryotes. Problems with CO_2 de-gassing from the Ci stock solution limited the highest CO_2 concentration used in the assay, which explains why the curve did not form a sigmoidal shape and thus the calculated $E.C._{.50}$ value was only an apparent $E.C._{.50}$ value based on the available data. There was also an inhibition of $Rv1625c_{204-443}$ specific activity observed at higher Ci concentrations (data not shown) of unknown origin.

Figure 4.9: CO_2 dose dependence assay of $Rv1625c_{204-443}$ at pH 6.5. 433 nM $Rv1625c_{204-443}$ was assayed at 37 °C for Ci sensitivity at pH 6.5 (50 mM MES) with 2 mM Mn^{2+} in the presence of 200 µM ATP and a total salt concentration of 30 mM for 40 minutes with a pre-incubation time of 30 minutes. Error bars are standard error of the mean (n=12), (*$p<0.05$) compared to the 0 mM added CO_2 value (first point on figure).

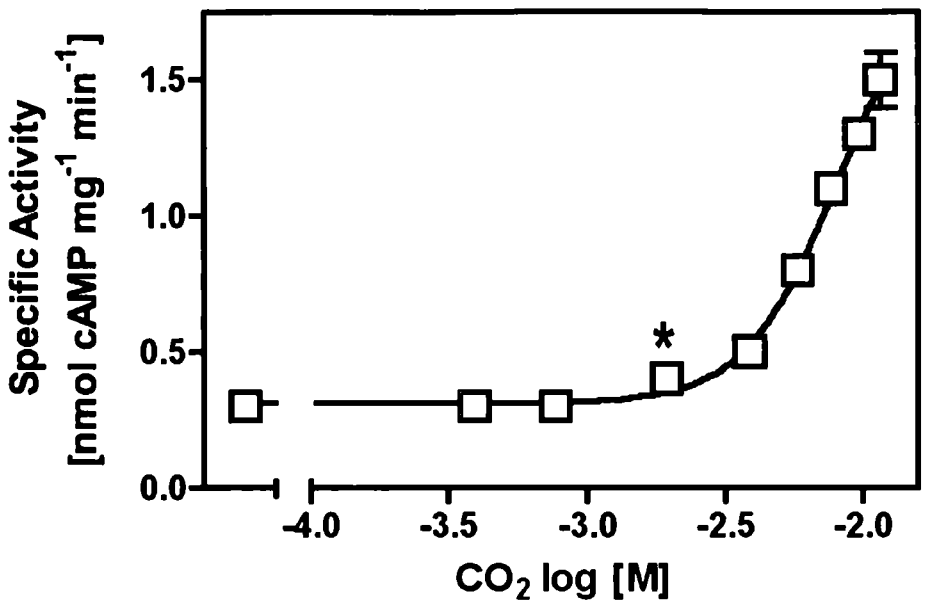
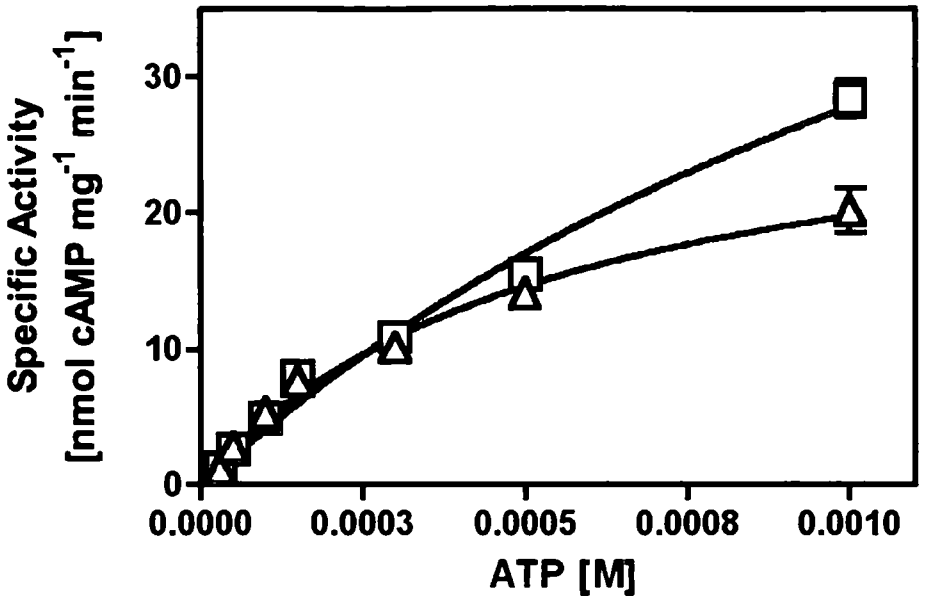

Specific Activity
[nmol cAMP mg-1 min-1]
1.5
1.0
0.5
0.0
*
-4.0
-3.5
-3.0
-2.5
-2.0
CO2 log [M]

Figure 4.10: Kinetic parameters of $Rv1625c_{204-443}$ at pH 6.5. 433 nM $Rv1625c_{204-443}$ was assayed at 37 °C to ascertain K_m and V_{max} at pH 6.5 (100 mM MES) with 2 mM Mn^{2+} and a total salt concentration of 20 mM for 40 minutes with 30 minutes pre-incubation. The assay contained either 20 mM $NaHCO_3$ (7.74 mM CO_2) (square) or 20 mM NaCl (triangle). Error bars are standard error of the mean (*n*=8).

Table 4.1: Kinetic parameters of $Rv1625c_{204-443}$ in the presence and absence of CO_2.



Kinetic parameter	Addition	
	Cl^-	CO_2
V_{max} [nmol cAMP mg^{-1} min^{-1}]	30.4 ± 0.8	76.0 ± 2.8
$K_{m\,[ATP]}$ [mM]	0.54 ± 0.02	1.72 ± 0.09
Hill Slope	0.91 ± 0.07	1.08 ± 0.06
k_{cat} [s^{-1}]	5.9	14.6
ε [L $mmol^{-1}$ s^{-1}]	10.9	8.5

The specific activities at 0 mM added CO_2 and 20 mM added CO_2 do not correspond with the values seen in the pH dependence assay (Figure 4.5) or the cation dependence assay (Figure 4.6) because the total salt concentration of the CO_2 dose dependence assay was 30 mM compared to 20 mM for the pH and cation dependence assays, however the CO_2 stimulation was similar (4-fold for CO_2 dose dependence assay compared with 5-fold for the pH and cation dependence assays).

The CO_2 E.C.$_{50}$ (app) was 13.2 ± 0.6 mM, which was a value that was higher than physiological levels of CO_2 and thus, the physiological relevance of the CO_2 E.C.$_{50}$ (app) value was difficult to identify from the limited data. The high E.C.$_{50}$ (app) value may be an artefact of the incomplete CO_2 dose dependence assay curve due to the problems of CO_2 degassing discussed before.

A different technique that measured the CO_2 K_d value was conducted using an equilibrium dialysis assay (data not shown), however, the experiment failed due to CO_2 degassing from the equipment at pH 6.5.

4.9 Kinetic parameters of Rv1625c$_{204-443}$ at pH 6.5

Assays to identify the effect of CO_2 on the affinity (via K_m) of Rv1625c$_{204-443}$ for ATP were conducted to compare the mode of CO_2 stimulation of Class IIIa ACs with Class IIIb ACs. Figure 4.10 shows the kinetics assay in the presence and absence of CO_2 with a substrate concentration ranging from 10–1000 µM. The maximum ATP concentration used was 1 mM as higher concentrations of substrate caused a decrease in the observed specific activity due to a smaller ATP:[α-^{32}P] ATP ratio, which led to less [2,8-^{3}H] cAMP being formed in relation to cAMP and thus a false decrease in specific activity.

Analysis using non-linear regression gave identical kinetic values (K_m and V_{max}) for Rv1625c$_{204-443}$ as stated in Figure 4.10. The results show that CO_2 increased K_m for ATP 3-fold. Thus CO_2 decreased the affinity of Rv1625c$_{204-443}$ for substrate compared to basal enzyme activity. However, CO_2 also increased V_{max} 2-fold, which increased the maximum amount of substrate that can be catalysed compared to basal enzyme activity. Thus the kinetic effect of CO_2

stimulation is increasing the turnover number (k_{cat}) of the enzyme 2-fold. The Hill Slope of $Rv1625c_{204-443}$ both in the presence and absence of CO_2 was approximately 1, suggesting that there is no increased or decreased cooperation between the two ATP binding sites of $Rv1625c_{204-443}$ and the efficiency of the enzyme (k_{cat}/K_m) decreased in the presence of CO_2 (Table 4.1). Therefore CO_2 stimulation of $Rv1625c_{204-443}$ decreases the efficiency of the enzyme in catalysing ATP to cAMP, however, it increases the turnover of the enzyme, thus making the enzyme work harder, but less efficiently compared to basal enzyme activity. This finding that CO_2 stimulates a Class IIIa AC through increasing k_{cat} was similar to the CO_2 stimulation seen in Class IIIb ACs and thus supports the idea that both Class IIIa and Class IIIb ACs have a similar CO_2 response mechanism (Hammer *et al.*, 2006).

4.10 CO_2 regulation of a mammalian Class IIIa AC

As a prokaryotic Class IIIa AC has been shown to be stimulated by CO_2 *in vitro*, other Class IIIa ACs may also be regulated by CO_2. A reconstituted mammalian Class IIIa AC comprising of a C1domain from AC7 and a C_2 domain from AC2 was investigated. The C_1 domain for AC7 and the C_2 domain from AC2 were overexpressed in *E. coli* and purified by affinity chromatography to homogeneity to generate recombinant proteins. The proteins were analysed for purity by SDS-PAGE.

The SDS-PAGE gel (Figure 4.11) of the expressed and purified $7C_1$ and $2C_2$ recombinant proteins shows a single protein band at approximately 24 kDa for $7C_1$ (Figure 4.11 (1)) and a single protein band at approximately 30 kDa for $2C_2$ (Figure 4.11 (2)). The theoretical weight of $7C_1$ is 23931 Da and of $2C_2$ is 30610 Da. The SDS-PAGE gel suggests that the protein bands are the correctly expressed proteins.

The reconstituted $7C_1.2C_2$ mammalian AC was assayed for CO_2 activity with a pH dependence assay in the presence and absence of Ci from pH 6.5 to 8.5. Figure 4.12 shows that the optimal pH value for basal activity of the reconstituted AC is pH 7.5. However, in the presence of Ci, there is a 2.5-fold Ci stimulation at pH 6.5 that decreases to a 1.8-fold Ci stimulation at pH 7.0.

Figure 4.11: SDS-PAGE gel of mammalian ACs. 2 µg of (1) $7C_{1\ 263-476}$ and (2) $2C_{2\ 821-1090}$ was applied to each lane with molecular weight standards on the left of each gel and the gel was stained with Coomassie Brilliant Blue G.

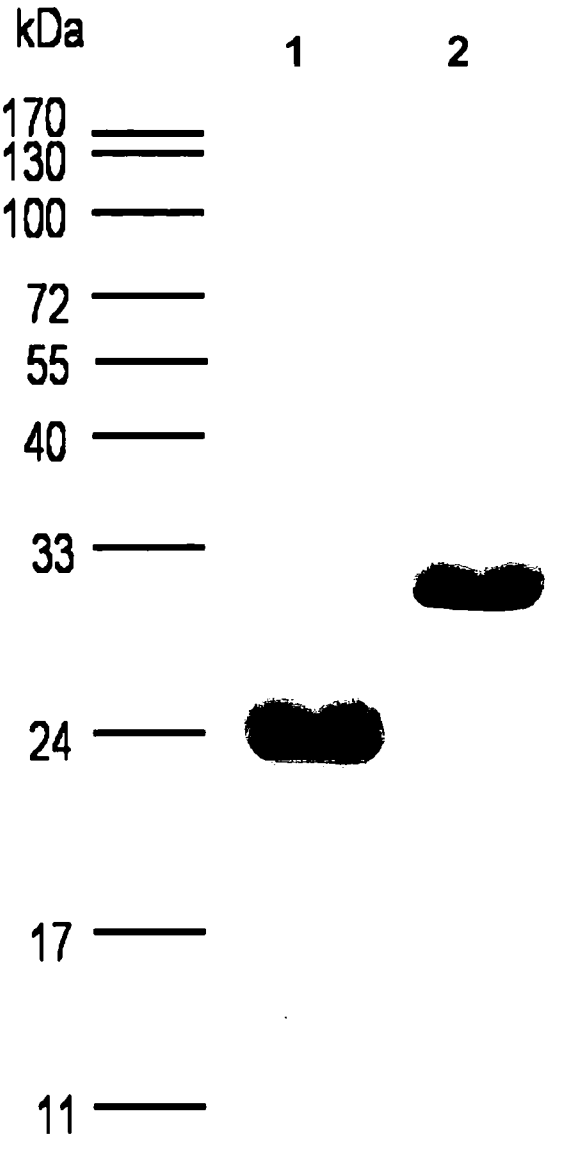
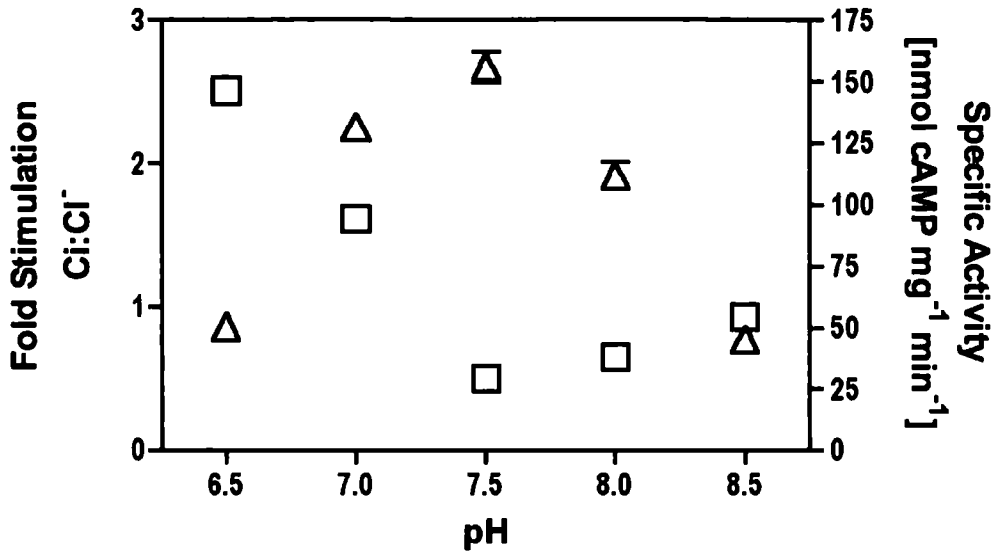

kDa
1
2
170
130
100
72
55
40
33
24
17
11

Figure 4.12: pH dependence assay of the reconstituted $7C_1.2C_2$ mammalian AC in the presence of Ci. 2 µM $7C_{1\ 263\text{-}476}$ and 350 nM $2C_{2\ 821\text{-}1090}$ were assayed at 30 °C at varying pH values in the presence of 1 mM ATP, 10 mM $MgCl_2$ and 10 µM forskolin for 20 minutes. The assay also contained either 20 mM $NaHCO_3$ (data not shown) or 20 mM NaCl (triangle). The squares represent Ci fold stimulation. Error bars are standard error of the mean (*n*=6).


Fold Stimulation
Ci:Cl⁻
Specific Activity
[nmol cAMP mg⁻¹ min⁻¹]
0
1
2
3
0
25
50
75
100
125
150
175
6.5
7.0
7.5
8.0
8.5
pH

At pH 7.5 and higher there is no Ci stimulation. This result is similar to the pH dependence assay of $Rv1625c_{204-443}$ where Ci stimulation was seen at pH 6.5 and 7.0 and suggests that CO_2 is the Ci species causing the stimulation. This is an interesting discovery, as mammalian tmACs were previously thought to be non-responsive to Ci owing to their insensitivity to HCO_3^-. If tmACs are sensitive to CO_2, they would be the first identified CO_2 sensors in mammals. However, further work into CO_2 regulation of tmACs would be required to confirm this proposal.

4.11 *In vivo* CO_2 stimulation of $Rv1625c_{204-443}$

If Rv1625c is a possible CO_2 sensor in *M. tuberculosis*, it is important to also demonstrate *in vivo* CO_2 stimulation of Rv1625c as well as *in vitro* CO_2 stimulation. It was difficult to observe the affect of CO_2 on Rv1625c in *M. tuberculosis* because the *M. tuberculosis* genome has at least 15 AC genes (Cole *et al.*, 1998). Work done previously on *M. tuberculosis* pathogenesis *in vivo* by infecting mice with a *Rv1625c* knockout strain demonstrated that Rv1625c alone was not a virulence factor, probably due to redundancy with the other 14 ACs (Guo *et al.*, 2005). Therefore, CO_2 responsiveness of Rv1625c activity *in vivo* was observed using *E. coli* as the host cell.

A plasmid containing the $Rv1625c_{204-443}$ gene was transformed into *E. coli*, since the *E. coli* genome has a Class I AC gene *cya* that is downregulated by cAMP and so any AC response will not be due to the endogenous Cya AC (Inada *et al.*, 1996). To measure any *in vivo* CO_2 stimulation of $Rv1625c_{204-443}$, a plasmid containing the *lacZ* gene behind a CRP promoter was also transformed into *E. coli*. (Figure 4.13) cAMP binds to the CRP and up-regulates the transcription of β-galactosidase. Adding ONPG, a substrate that β-galactosidase converts to galactose and ONP, which is a yellow coloured product that can be quantified. Thus, in this assay an increase in cAMP due to CO_2 stimulation of $Rv1625c_{204-443}$ in *E. coli* cells will produce an increase in ONP. An ONP standard curve was generated (Figure 4.14A), which was used to ascertain the concentration of ONP produced in the *in vivo* β-galactosidase

assay. The β-galactosidase assays for $Rv1625c_{204-443}$ and LacZ in *E. coli* and the control experiment of LacZ in *E. coli* were incubated at pH 7.1 in either 10 % (v/v) CO_2 in air or air for 30 minutes. The negative control was LacZ in *E.coli* that did not contain the recombinant AC gene. The amount of ONP produced was recorded and is shown in Figure 4.14B. The figure shows that in *E. coli* cells, $Rv1625c_{204-443}$ was stimulated by CO_2 as compared to $Rv1625c_{204-443}$ in *E. coli* cells incubated in air. The CO_2 stimulation of $Rv1625c_{204-443}$ in *E. coli* cells is also greater than the negative control experiments of LacZ vector in *E. coli* cells incubated both in CO_2 and air, thus demonstrating that the CRP in *E. coli* was not activated by CO_2. The CO_2 stimulation of $Rv1625c_{204-443}$ was much lower *in vivo* compared to *in vitro*, which may be due to three reasons. Firstly, there may be a lag phase *in vivo* between CO_2 sensing and transcription of *lacZ* that may account for the lower fold stimulation. Secondly, the growth media was buffered to pH 7.1, more alkaline than the *in vitro* assays at pH 6.5 and therefore there is less Ci in the form of CO_2. Finally, the amount of CO_2 that was sensed by $Rv1625c_{204-443}$ *in vivo* was unknown and may be a lower amount than in the *in vitro* assays, possibly due to the transfer of CO_2 into the medium being a limiting factor.

The *in vivo* β-galactosidase assay shows that $Rv1625c_{204-443}$ is stimulated by CO_2 both *in vitro* and *in vivo* and Rv1625c is the first characterised bacterial AC to be stimulated by CO_2 *in vivo*. Thus, Rv1625c may be a CO_2 sensor for *M. tuberculosis*, however, further work done *in vivo* in *M. tuberculosis* would need to be carried out to prove this theory.

Figure 4.13: Diagram of *in vivo* CO_2 stimulation of $Rv1625c_{204-443}$. CO_2 (green) enters *E. coli*, binds and stimulates $Rv1625c_{204-443}$ on the pQE30 vector. cAMP (blue) is produced and binds to the CRP promoter on the pCTXLacZ vector, which up-regulates the transcription of *lacZ*. β-galactosidase converts the added ONPG into galactose and ONP, a yellow compound that is quantified.

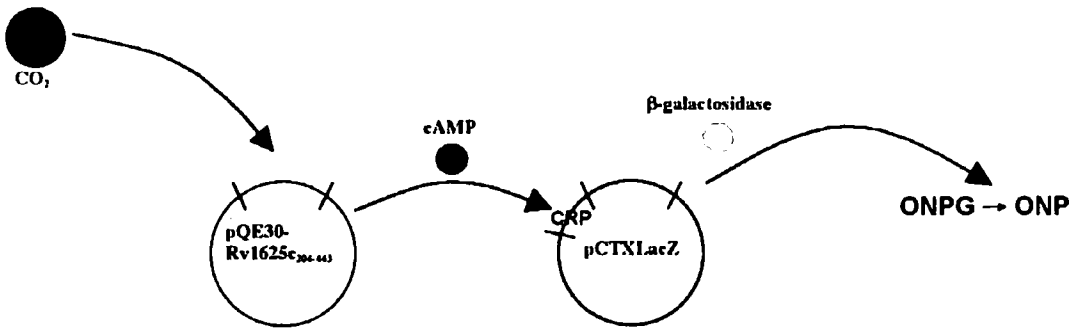
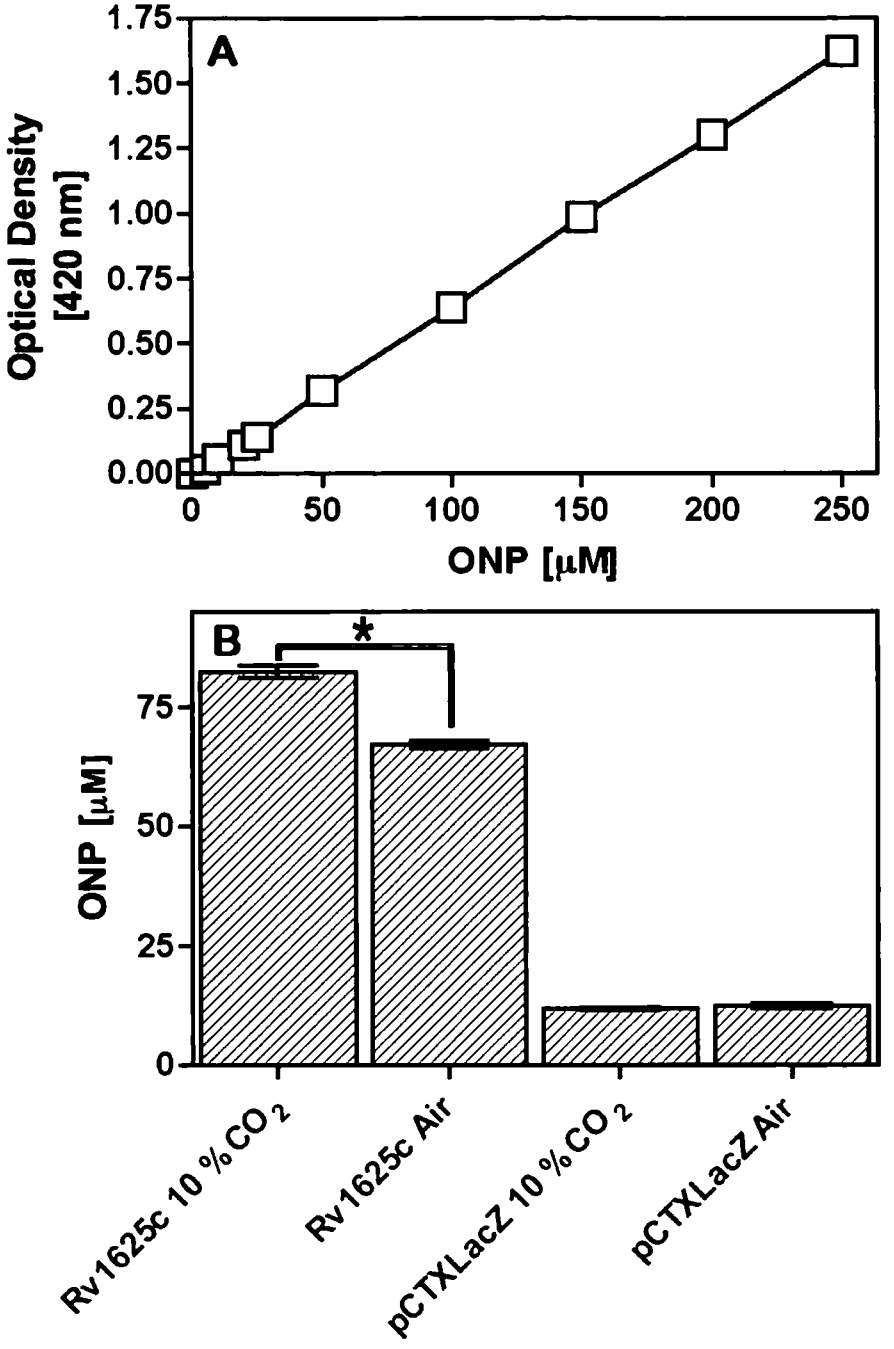

CO_2
cAMP
β-galactosidase
pQE30-
Rv1625c
CRP
pCTXLacZ
ONPG → ONP

Figure 4.14: (A) Standard curve of ONP to determine the ONP concentration for the β-galactosidase assay. (B) *In vivo* CO_2 stimulation of $Rv1625c_{204-443}$. Amount of ONP produced with either $Rv1625c_{204-443}$ with pCTXLacZ in *E. coli* or pCTXLacZ in *E. coli* incubated in either 10 % CO_2 or air. Error bars are standard error of the mean (n=9) (*$p<0.05$).


A
Optical Density [420 nm]
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
50
100
150
200
250
ONP [μM]
B
*
ONP [μM]
75
50
25
0
Rv1625c 10 % CO 2
Rv1625c Air
pCTXLacZ 10 % CO 2
pCTXLacZ Air

4.12 Conclusion

$Rv1625c_{204-443}$ was stimulated by CO_2 at pH 6.5 and is the first characterised Class IIIa AC shown to be stimulated by CO_2. $Rv1625c_{204-443}$ had an apparent $E.C._{50}$ for CO_2 of 13.2 ± 0.6 mM, which was greater than the value of physiological CO_2 in *M. tuberculosis*, but was an erroneous value due to an incomplete CO_2 dose dependence assay. CO_2 stimulated $Rv1625c_{204-443}$ by increasing V_{max} and consequently k_{cat}, but also decreased K_m compared to basal enzyme activity. The CO_2 stimulation was also observed *in vivo* by incubating $Rv1625c_{204-443}$ transformed into *E. coli* cells in 10 % CO_2 in air as compared to $Rv1625c_{204-443}$ transformed into *E. coli* cells incubated in air.

Identification of the conditions required for CO_2 to bind to $Rv1625c_{204-443}$ was investigated next as this may elucidate the mechanism of CO_2 stimulation of ACs, for example whether the metal co-factor or substrate are required for CO_2 binding. Once the CO_2 binding conditions have been identified, the actual CO_2 binding site on $Rv1625c_{204-443}$ will hopefully be identified.

5 Elucidation of AC CO_2 binding conditions

5.1 Introduction

Having demonstrated that CO_2 stimulates a model Class IIIa AC in addition to Class IIIb ACs, an investigation of the conditions for CO_2 binding is required. CO_2 may bind to apoprotein alone, or may have a requirement for metal co-factor, substrate or metal co-factor and substrate. Finding the CO_2 binding conditions would help to identify the CO_2 binding site on the Class IIIa ACs and shed significant light on the mechanism of activation.

Here, a number of spectroscopic techniques, such as CD spectroscopy, FTIR spectroscopy and HCO_3^- dependent luminescence probe spectroscopy, as well as a number of biochemical techniques, such as Ci stimulation of catalytic $Rv1625c_{204-443}$ single mutant proteins, Mn^{2+} dose dependence assay on $Rv1625c_{204-443}$ and $^{14}CO_2$ binding assays on ACs, have been used to try and identify the conditions required for CO_2 binding. Unlike AC assays, these techniques do not measure enzyme specific activity but other changes that occur on CO_2 binding to protein such as a conformational change or changes in infrared properties. These changes can be measured in the absence of metal co-factor or substrate or both, to identify the requirements for binding.

5.2 Ci stimulation of catalytic Rv1625c$_{204-443}$ single mutant proteins

In the literature there are two opposing hypotheses to explain Ci stimulation of Class IIIb ACs. The first hypothesis is that Ci binds to the substrate defining Lys in the active site (Cann *et al.*, 2003) and the second hypothesis is that Ci binds to the metal ion in the active site which causes the recruitment of a second metal ion (Steegborn *et al.*, 2005). These two hypotheses were tested with Rv1625c$_{204-443}$ to ascertain which hypothesis was correct for Class IIIa ACs.

According to the first hypothesis CO_2 binds to the substrate defining Lys, thus mutating the Lys residue to Ala should ablate the CO_2 stimulation. As there are 4 other conserved catalytic amino acids between the studied Class IIIa and Class IIIb ACs (see alignment in Figure 4.1, mammalian Class IIIa AC active site in Figure 2.5 and Rv1625c$_{204-443}$ active site in Figure 5.1), these residues may also represent potential CO_2 binding sites and so were mutated to create single mutant proteins that were also assayed for CO_2 dose dependency.

The catalytic domains of the single Ala mutant proteins of Rv1625c$_{204-443}$ were overexpressed in *E. coli* and purified by affinity chromatography to homogeneity to generate recombinant mutant proteins. The mutant proteins were analysed for purity by SDS-PAGE (Figure 5.2). The SDS-PAGE gel of the mutant proteins shows a single protein band at approximately 30 kDa, which is consistent with the theoretical weight of Rv1625c$_{204-443}$ of 26790 Da.

Figure 5.3 shows the CO_2 dose dependence curves for the Rv1625c$_{204-443}$ mutant proteins. As with the Rv1625c$_{204-443}$ wt protein in Section 4.5, problems with CO_2 de-gassing from the Ci stock limited the highest CO_2 concentration used in the assay, thus the CO_2 dose dependence curves did not form a sigmoidal shape that is expected due to a simple binding of CO_2 to the protein. The calculated E.C.$_{50}$ values were only apparent E.C.$_{50}$ values based on the available data. The concentration of mutant protein assayed differed in each assay so as to obtain a measurable amount of [2,8-^{3}H] cAMP whilst also keeping the conversion rate below 10 %. However, E.C.$_{50}$ value is

dependent on enzyme concentration, so to be able to compare the $E.C._{50}$(app) values obtained for the different mutant proteins, the $E.C._{50}$(app) value must be multiplied by enzyme concentration to produce an arbitrary value. This corrected value is stated with the $E.C._{50}$(app) value. Figure 5.3A shows the CO_2 dose dependence curve for the $Rv1625c_{204-443}$ D256A mutant protein, which is a mutation of the Asp that binds metal ion B in the active site. The mutant protein was active, with a low specific activity but was also stimulated by CO_2 and has an $E.C._{50}$(app) value of 4.2 ± 1.1 mM and a comparable value of 60.5 ± 15.8 mM μM protein. Figure 5.3B shows the CO_2 dose dependence curve for the $Rv1625c_{204-443}$ K296A mutant protein, which is a mutation of the Lys that specifies the substrate ATP. The mutant protein was active with an $E.C._{50}$(app) value of 28.4 ± 7.3 mM and a comparable value of 204.5 ± 52.6 mM μM protein. Figure 5.3C shows the CO_2 dose dependence curve for the $Rv1625c_{204-443}$ D300A mutant protein, which is a mutation of the Asp that binds metal ion A in the active site. The mutant protein was active, with a low specific activity and was stimulated by CO_2 with an $E.C._{50}$(app) value of 6.4 ± 1.1 mM and a comparable value of 46.1 ± 7.9 mM μM protein. Figure 5.3D shows the CO_2 dose dependence curve for the $Rv1625c_{204-443}$ D365A mutant protein, which is a mutation of the Asp that is required for substrate specificity of the AC for ATP. The mutant protein was still active and was stimulated by CO_2 with an $E.C._{50}$(app) value of 5.3 ± 1.1 mM and a comparable value of 76.3 ± 15.8 mM μM protein. Figure 5.3E shows the Ci dose dependence curve for the $Rv1625c_{204-443}$ R376A mutant protein, which is a mutation of the Arg required for transition state stabilisation. The mutant protein was still active and was stimulated by CO_2 with an $E.C._{50}$(app) value of 13.2 ± 1.4 mM and a comparable value of 7.3 ± 0.8 mM μM protein.

These CO_2 dose dependence assays show that all the mutant proteins assayed were stimulated by CO_2, suggesting that CO_2 did not bind exclusively to any of these residues. The corrected $E.C._{5-}$(app) values imply that $Rv1625c_{204-443}$ R376A binds CO_2 the strongest and $Rv1625c_{2-4-443}$ K296A binds CO_2 the weakest. However, as these comparable values were calculated from apparent $E.C._{50}$ values, the results are not conclusive and should not be analysed too carefully.

Figure 5.1: Active site structure of $Rv1625c_{204-443}$. Catalytic amino acids are shown with their labelling (blue) with ATP (black) and metal ions (pink) in the active site. Hydrogen bonds are shown (red).

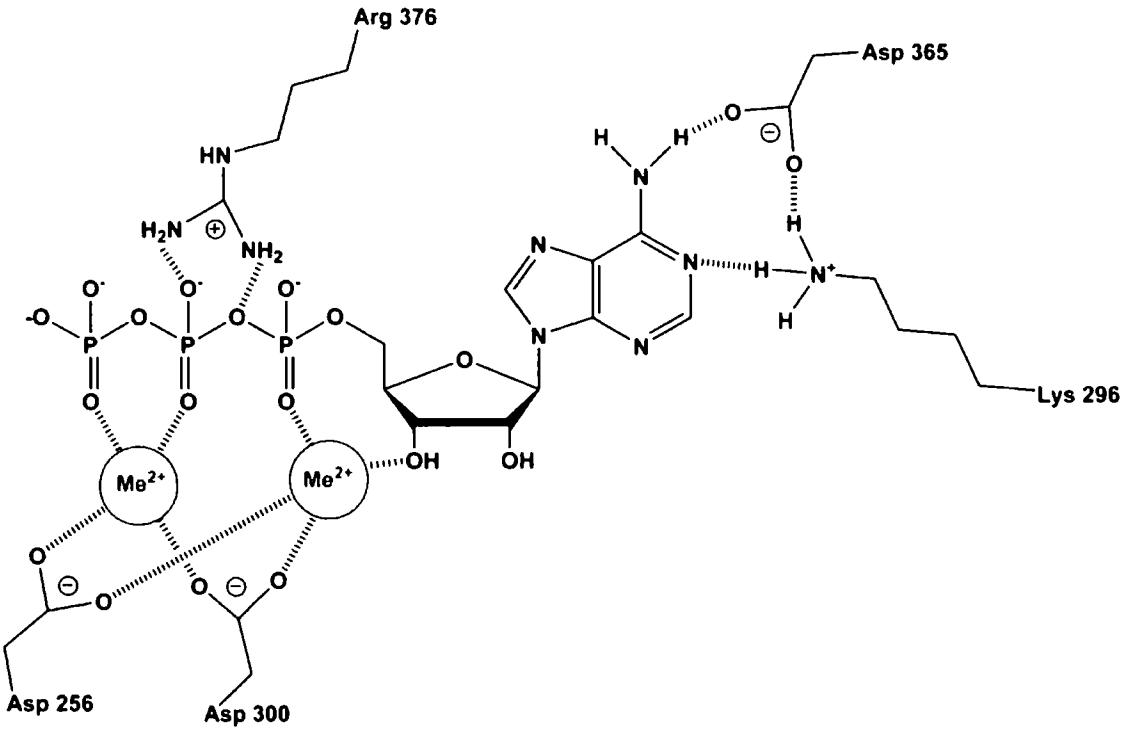

Arg 376
Asp 365
Lys 296
Asp 256
Asp 300
Me2+
Me2+
OH
OH

Figure 5.2: SDS-PAGE gel of Rv1625c$_{204-443}$ mutant proteins. 2 µg protein was applied with molecular weight standards on the left of the gel and stained with Coomassie Brilliant Blue G. 1) D256A 2) K296A 3) D300A 4) D365A 5) R376A.

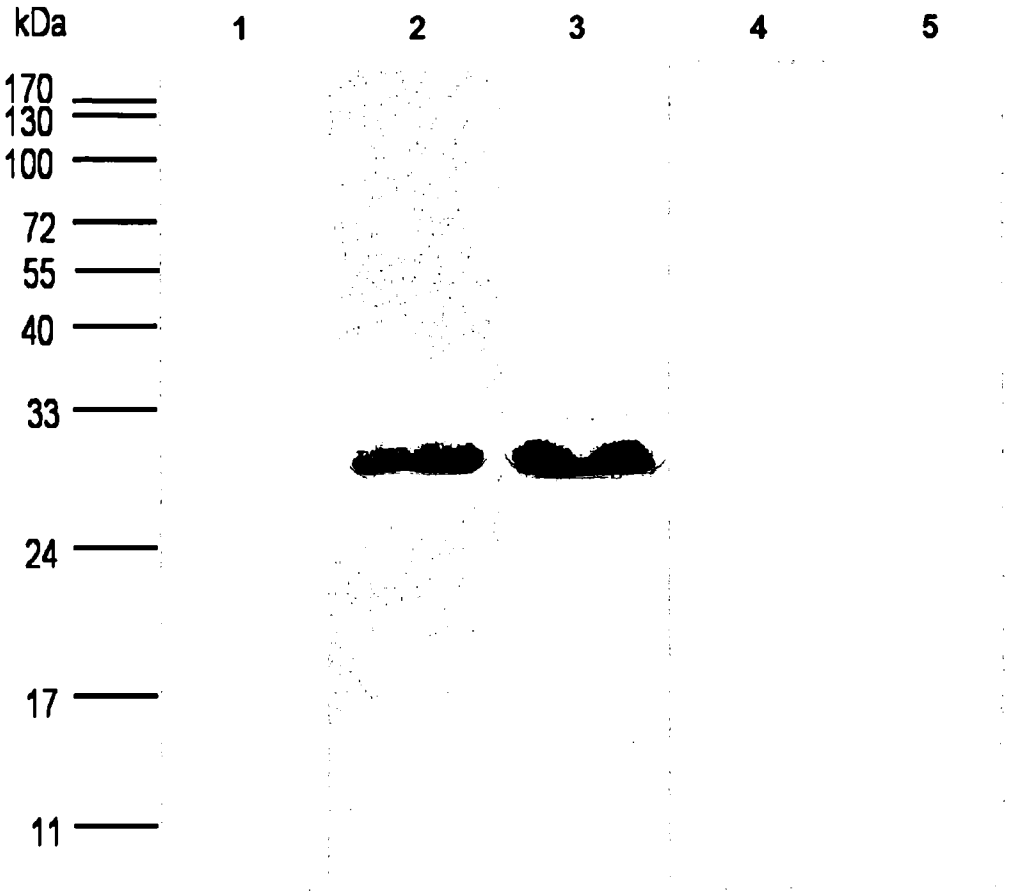
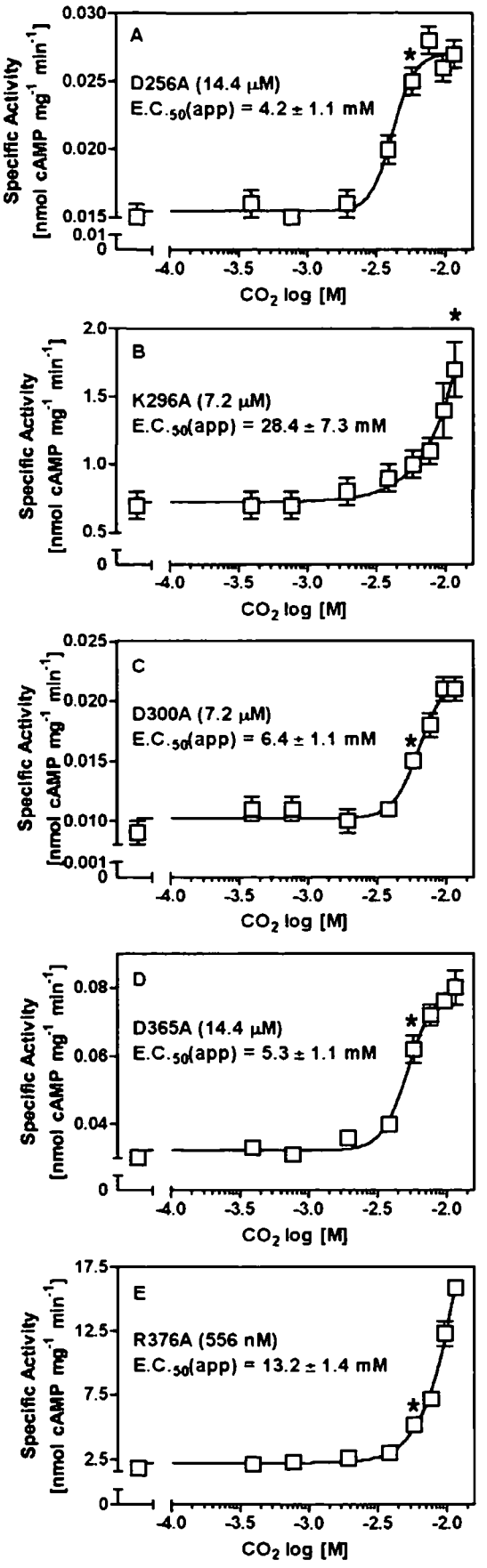

kDa
1
2
3
4
5
170
130
100
72
55
40
33
24
17
11

Figure 5.3: CO_2 dose dependence assay of $Rv1625c_{204-443}$ mutant proteins at pH 6.5. $Rv1625c_{204-443}$ mutant proteins were assayed at 37 °C at pH 6.5 (50 mM MES) with 2 mM Mn^{2+} in the presence of 200 µM ATP and a total salt concentration of 30 mM for 40 minutes with a pre-incubation time of 30 minutes. $E.C._{50}$(app) values are included in figures. Error bars are standard error of the mean (n=6), (*$p<0.05$) compared to 0 mM added CO_2. A) 14.4 µM D256A, B) 7.2 µM K296A, C) 7.2 µM D300A, D) 14.4 µM D365A, E) 556 nM R376A.


A
D256A (14.4 μM)
E.C.50(app) = 4.2 ± 1.1 mM
B
K296A (7.2 μM)
E.C.50(app) = 28.4 ± 7.3 mM
C
D300A (7.2 μM)
E.C.50(app) = 6.4 ± 1.1 mM
D
D365A (14.4 μM)
E.C.50(app) = 5.3 ± 1.1 mM
E
R376A (556 nM)
E.C.50(app) = 13.2 ± 1.4 mM
Specific Activity [nmol cAMP mg-1 min-1]
CO2 log [M]

Thus the first hypothesis that the substrate defining Lys exclusively binds CO_2 is not correct for Rv1625c$_{204-443}$. The second hypothesis of CO_2 binding to metal ions in the active site was therefore tested.

5.3 Mn^{2+} dose dependence assay of Rv1625c$_{204-443}$

To observe the effect of CO_2 on Mn^{2+} affinity on Rv1625c$_{204-443}$ and ascertain the effect CO_2 has on metal recruitment into the active site a Mn^{2+} dose dependence assay in the presence and absence of CO_2 was performed. Figure 5.4 shows the affect of CO_2 on the Mn^{2+} dose dependence of Rv1625c$_{204-443}$. In the presence of 20 mM $NaHCO_3$ pH 6.5 the CO_2 stimulation increased up to 0.5 mM Mn^{2+}, where specific activity levelled then started to decrease at 2 mM Mn^{2+}. There was an increase in the Hill slope of the Mn^{2+} dose response to 6.6. In the absence of CO_2, specific activity initially increased with increasing concentration of Mn^{2+} but then decreased at 5 mM Mn^{2+} with a Hill slope of 3.0. The doubling of the Hill slope in the presence of CO_2 suggests an increase in the co-operativity of the two binding sites. This result agrees with the hypothesis of Steegborn *et al.* that Ci recruits a second metal ion into the active site of the AC but it does not prove that CO_2 is binding to that metal.

5.4 CD spectroscopy

To confirm whether a metal ion is required for CO_2 binding to protein three spectroscopic techniques were used to observe CO_2 binding to protein with the addition of a divalent metal ion, substrate, or metal ion and substrate. The first spectroscopic technique used to identify CO_2 binding was CD spectroscopy. The CD spectra in the near UV (250–350 nm) describes the secondary structure of a protein through the differential absorbance of left and right handed circularly polarised light (Mathews, 2000).

Rv1625c$_{204-443}$ was analysed in the presence and absence of Ci under conditions that differed from conditions required for catalysis. A conformational

change in the secondary structure of the protein was hypothesised to occur when CO_2 bound to the protein, which should be seen by CD spectroscopy, thus identifying the conditions that are required for CO_2 binding to $Rv1625c_{204-443}$. Figure 5.5A shows the CD spectra of $Rv1625c_{204-443}$ in the presence or absence of CO_2, and with the addition of $MgCl_2$ or ATP, or $MgCl_2$ and ATP. Figure 5.5A demonstrates a conformational change due to the binding of ATP into the active site both in the absence and presence of metal ions. However with the addition of CO_2, there was no observable conformational change that would suggest the binding of CO_2 to the protein. A conformational change due to CO_2 binding may be too small to be detected by CD spectroscopy. Alternatively CO_2 binding does not cause a conformational change in the protein. Figure 5.5B shows the activity of $Rv1625c_{204-443}$ using the conditions required for obtaining the CD spectra and showed that the enzyme was active in the presence of metal ions and was not active in the absence of metal ions. This result demonstrates that ATP can bind to the active site of the protein without metal present and thus without catalysis occurring but was uninformative with respect to requirements for CO_2 binding.

5.5 Expression and purification of $Slr1991_{120-337}$ and $CyaB1_{595-859}$ for FTIR spectroscopy

As CD spectroscopy did not help identify the requirements for CO_2 binding, other spectroscopic techniques, such as FTIR spectroscopy was used to try and identify the requirements for CO_2 binding. As CO_2 is IR active, FTIR spectroscopy was used to measure a decrease in the CO_2 signal due to CO_2 binding to the protein or the appearance of a new carbonyl peak formed by the bound CO_2 on the protein, for example from the binding of CO_2 to Lys to form a carbamate. Another spectroscopic technique used was the addition of a HCO_3^- dependent luminescent probe that would measure the amount of HCO_3^- in solution. Thus a quenching of the probe signal should be observed due to CO_2 binding the protein and decreasing the amount of HCO_3^- in solution.

Figure 5.4: Mn^{2+} dose dependence assay of $Rv1625c_{204-443}$ at pH 6.5. 925 nM $Rv1625c_{204-443}$ was assayed at 37 °C at pH 6.5 (50 mM MES) in the presence of 200 µM ATP and salt concentration of 20 mM for 40 minutes with a pre-incubation time of 30 minutes. The assay contained either 20 mM $NaHCO_3$ pH 6.5 (7.74 mM CO_2) (square) or 20 mM NaCl pH 6.5 (triangle). Error bars are standard error of the mean (n=6) (*$p<0.05$).

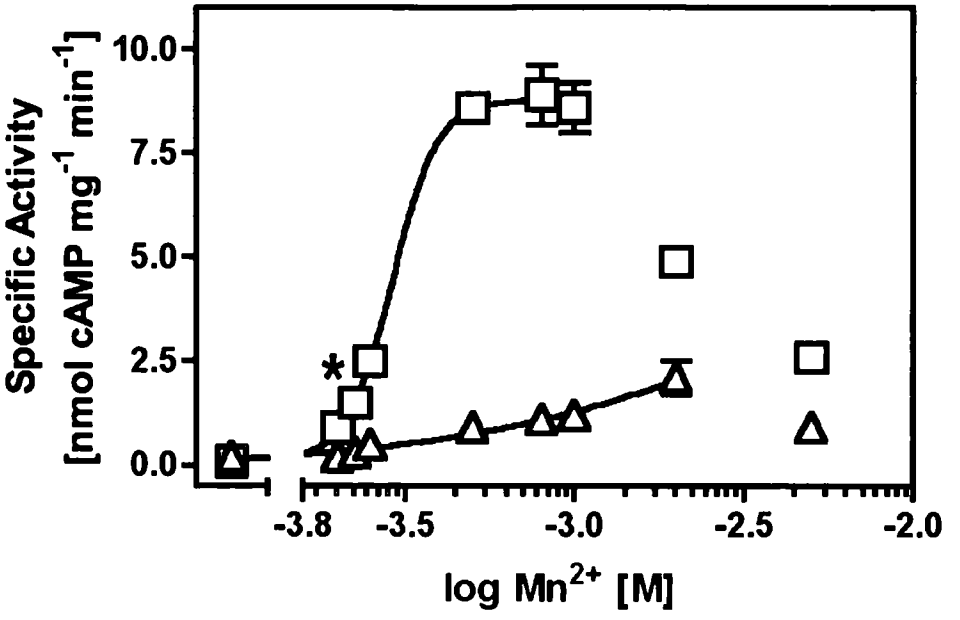
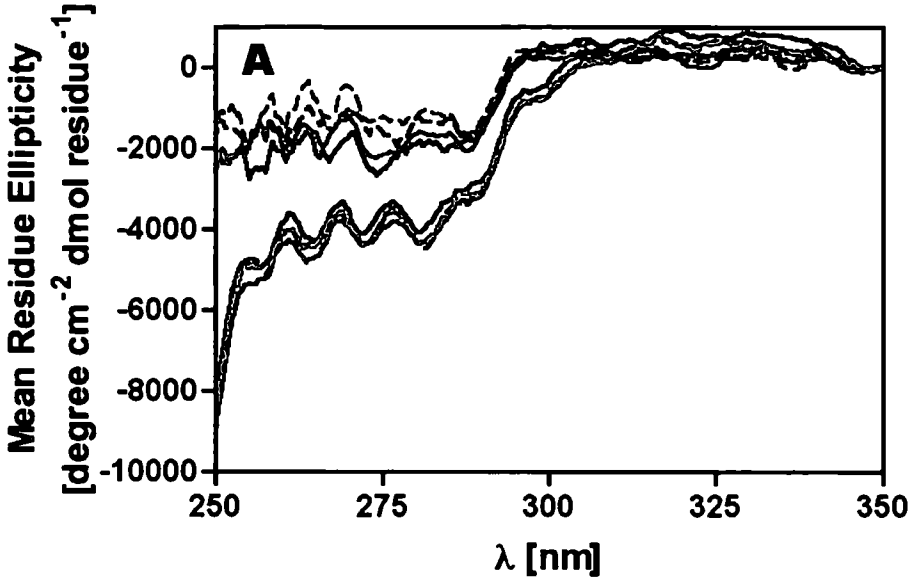

Specific Activity
[nmol cAMP mg-1 min-1]
10.0
7.5
5.0
2.5
0.0
-3.8
-3.5
-3.0
-2.5
-2.0
log Mn2+ [M]

Figure 5.5: A) Circular dichroism spectra of 20 μM $Rv1625c_{204-443}$ at pH 6.5 in the presence of the following conditions: nothing (solid black line), CO_2 (dotted black line), 10 mM $MgCl_2$ (solid red line), 10 $MgCl_2$ and 7.74 mM CO_2 (dotted red line), 5 mM ATP (solid blue line), 5 mM ATP and 7.74 mM CO_2 (dotted blue line), 5 mM ATP and 10 mM $MgCl_2$ (solid green line) and 5 mM ATP, 10 mM $MgCl_2$ and 7.74 mM CO_2 (dotted green line). B) AC assay of $Rv1625c_{204-443}$ in the absence and presence of 10 mM $MgCl_2$. Error bars are standard error of the mean (n=4) (*$p<0.05$).


A
Mean Residue Ellipticity
[degree cm-2 dmol residue-1]
0
-2000
-4000
-6000
-8000
-10000
250
275
300
325
350
λ [nm]

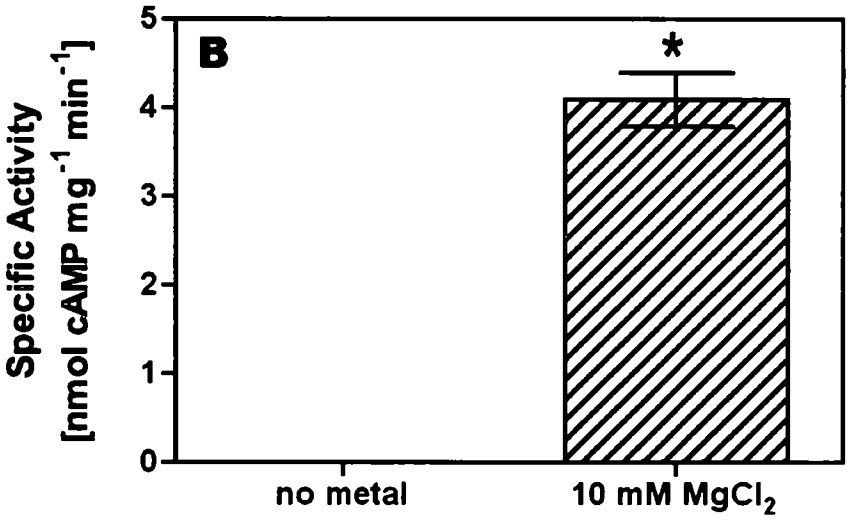
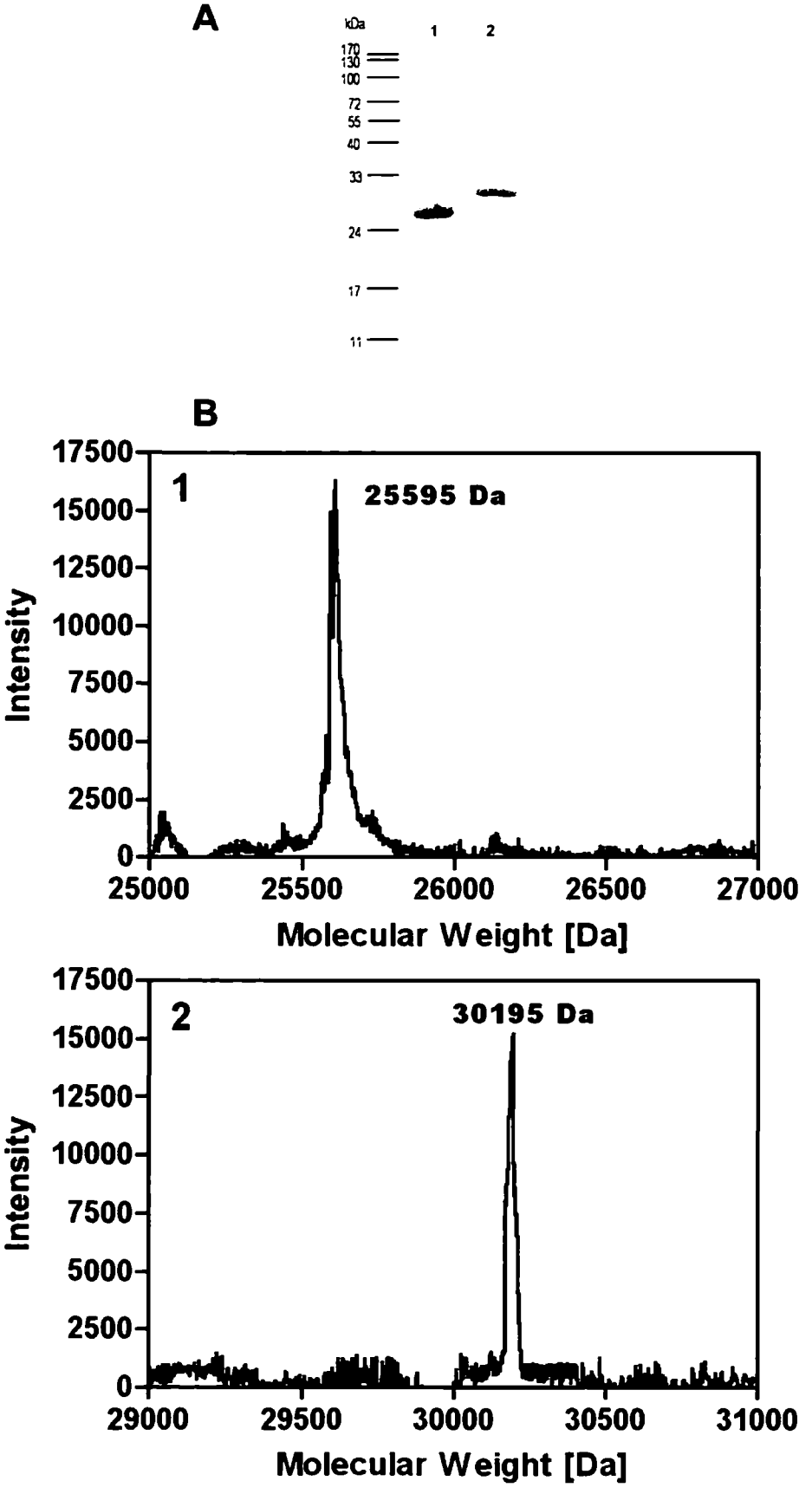

B
*
Specific Activity
[nmol cAMP mg-1 min-1]
5
4
3
2
1
0
no metal
10 mM MgCl2

Figure 5.6: A) SDS-PAGE gel of 1) $Slr1991_{120-337}$ and 2) $CyaB1_{595-859}$. 2 µg protein was applied with molecular weight standards on the left of each gel and stained with Coomassie Brilliant Blue G. B) MS data of 1) $Slr1991_{120-337}$ and 2) $CyaB1_{595-859}$.


A
kDa
1
2
170
130
100
72
55
40
33
24
17
11
B
1
25595 Da
Intensity
Molecular Weight [Da]
2
30195 Da
Intensity
Molecular Weight [Da]

These two techniques required large amounts of protein, such as 0.5 mg protein per spectrum for FTIR spectroscopy and 6.4 mg protein per assay for the HCO_3^- dependent luminescent probe. As the yield of $Rv1625c_{204-443}$ is typically 8 mg per 8 L culture, this protein was not suited to these experiments. Other ACs that were CO_2 stimulated and overexpressed in larger quantities were used to verify the techniques before investing time in $Rv1625c_{204-443}$ expression.

Two Class IIIb ACs that have been previously demonstrated to be stimulated by CO_2, $Slr1991_{120-337}$ from *Synechocystis* and $CyaB1_{595-859}$ from *Anabaena* (Hammer *et al.*, 2006) produced better yields from overexpression (approximately 50 mg per 8 L culture) and were therefore used for these spectroscopic techniques.

The catalytic domains of the Class IIIb ACs $Slr1991_{1230-337}$ from *Synechocystis* and $CyaB1_{595-859}$ from *Anabaena* were overexpressed in *E. coli* and purified by affinity chromatography to homogeneity. The proteins were analysed for purity by MS and SDS-PAGE.

The SDS-PAGE gels of the expressed and purified $Slr1991_{120-337}$ and $CyaB1_{595-859}$ recombinant proteins (Figure 5.6A) show single protein bands at approximately 25 kDa and 30 kDa respectively. The theoretical weight of $Slr1991_{120-337}$ is 25595 Da and $CyaB1_{595-859}$ is 30195 Da.

The MS data (Figure 5.6B) show that the actual molecular weight of the purified proteins match the theoretical molecular weights of 25595 Da for $Slr1991_{120-337}$ and 30195 Da for $CyaB1_{595-859}$ and proved that the purified proteins are $Slr1991_{120-337}$ and $CyaB1_{595-859}$ The data also show that there were no contaminating protein bands.

5.6 FTIR spectroscopy

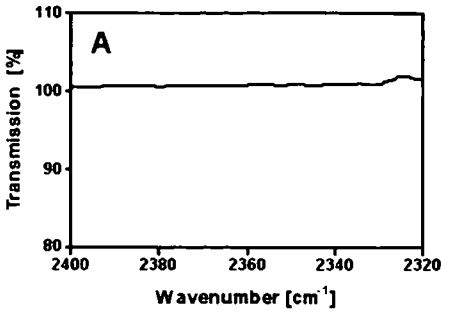
CO_2 is IR active, therefore FTIR spectroscopy was used to identify the conditions for CO_2 binding to AC. CO_2 is a linear molecule that has three characteristic stretching modes. CO_2 can stretch symmetrically, with both oxygen atoms stretching in opposite directions, however this mode is not IR active. Another CO_2 stretching mode is asymmetrical stretching, with the

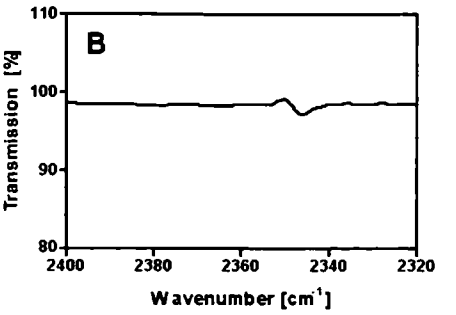
oxygen atoms stretching in the same directions. This mode can be observed at about 2345 cm^{-1}. The final mode of stretching of CO_2 is a vertical bend of the molecule, which is predicted at approximately 546 cm^{-1} (Schultz *et al.*, 1996). This mode lies in the 400–1500 cm^{-1} region of the spectrum which is called the fingerprint region owing to the unique signals formed by intra-molecular bonds of the molecule, which overlap each other, thus making identification of the vertical bend of CO_2 difficult (Schultz *et al.*, 1996). FTIR spectroscopy can also be used to determine the secondary structure of proteins by looking at vibrations of the peptide backbone (Krimm and Bandekar, 1986). These vibrations form bands termed Amide I band and Amide II band. The Amide I band is the C=O stretching mode of the peptide and is usually in the region 1600–1700 cm^{-1} (Krimm and Bandekar, 1986). The Amide II band is the N-H bend and C-N stretch in the peptide backbone and is usually seen at 1550 cm^{-1} (Krimm and Bandekar, 1986). However, both these regions are masked by the H_2O bending vibrations at 1650 cm^{-1}, thus to observe the Amide bands samples need to be dissolved in D_2O. The spectra of proteins in D_2O were measured (data not shown), however the O-D stretching vibrations mask the asymmetric CO_2 signal at 2345 cm^{-1} and as observed previously in the CD spectroscopy, there was no change in the structure of the protein on addition of CO_2 under any conditions (data not shown). Therefore, the protein spectra were measured in H_2O and the asymmetric CO_2 signal at 2345 cm^{-1} alone was observed. When CO_2 is added to a sample containing AC in the required conditions for CO_2 binding, it is hypothesised that a CO_2 signal should be visible at around 2345 cm^{-1} with another signal at a lower wavenumber visible due to CO_2 bound to the protein. Therefore, the region around the asymmetric CO_2 stretch was the main area of observation.

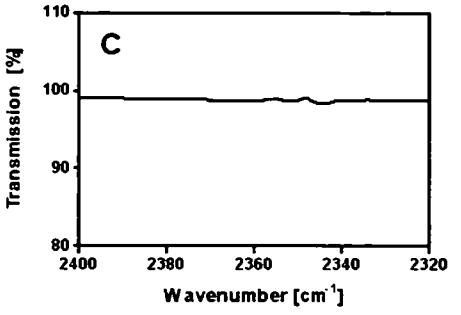
Control FTIR spectra were obtained with no protein present and with the H_2O spectrum subtracted to give the spectrum of the added factors only. As the CO_2 signal is the target of the experiment, all spectra are focused into the region 2400 – 2320 cm^{-1} where CO_2 peaks were observed.

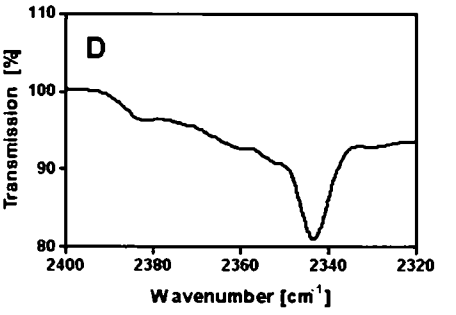
Figure 5.7A shows the spectrum of buffer only. As expected in the region 2400–2320 cm^{-1} there were no signals. Figure 5.7B shows the spectrum for buffer and $MgCl_2$ and demonstrated that there were no additional signals in the region of interest.

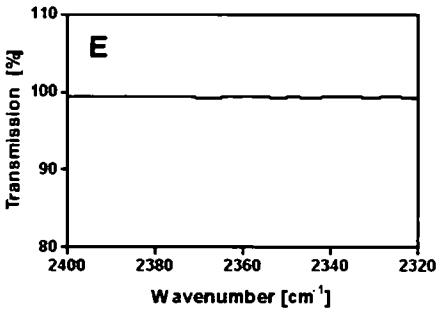
Figure 5.7: FTIR spectra with the subtraction of the H_2O spectrum of the control conditions with the following additions, A) 50 mM MES, B) 50 mM MES and 20 mM $MgCl_2$, C) 50 mM MES and 5 mM ATP, D) 50 mM MES and 3 mM CO_2, E) 50 mM MES, 20 mM $MgCl_2$ and 5 mM ATP, F) 50 mM MES, 20 mM $MgCl_2$ and 3 mM CO_2, G) 50 mM MES, 5 mM ATP and 3 mM CO_2, H) 50 mM MES, 20 mM $MgCl_2$, 5 mM ATP and 3 mM CO_2.

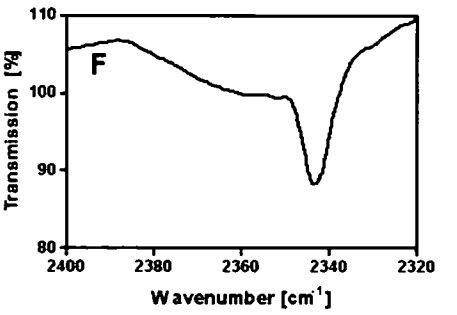

A
Transmission [%]
110
100
90
80
2400
2380
2360
2340
2320
Wavenumber [cm-1]

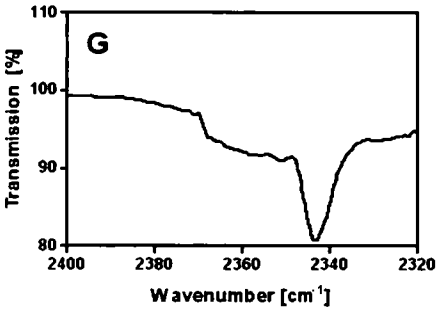

B
Transmission [%]
110
100
90
80
2400
2380
2360
2340
2320
Wavenumber [cm-1]


C
Transmission [%]
110
100
90
80
2400
2380
2360
2340
2320
Wavenumber [cm-1]


D
Transmission [%]
110
100
90
80
2400
2380
2360
2340
2320
Wavenumber [cm-1]


E
Transmission [%]
110
100
90
80
2400
2380
2360
2340
2320
Wavenumber [cm-1]


F
Transmission [%]
110
100
90
80
2400
2380
2360
2340
2320
Wavenumber [cm-1]


G
Transmission [%]
110
100
90
80
2400
2380
2360
2340
2320
Wavenumber [cm-1]

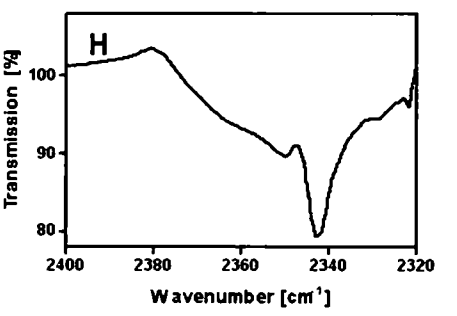
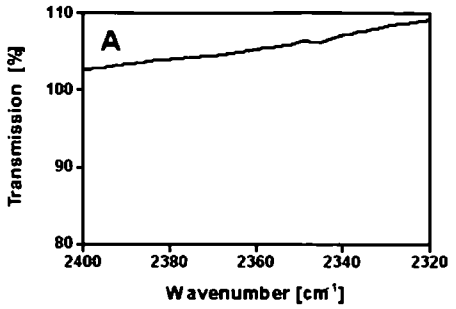
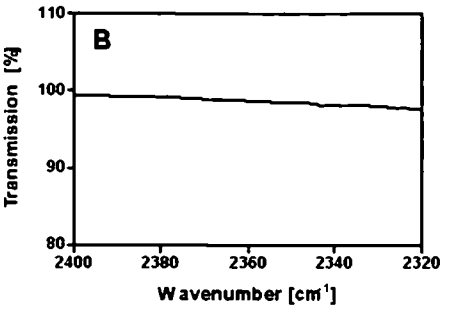
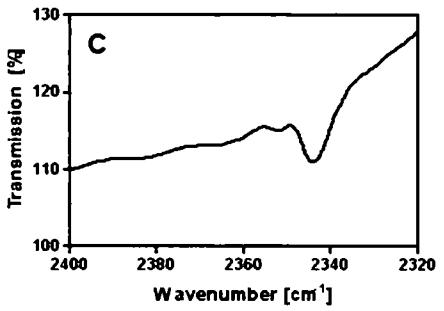
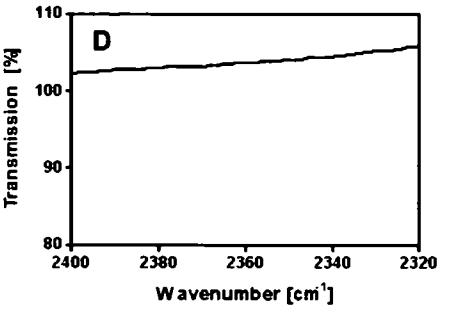

H
Transmission [%]
100
90
80
2400
2380
2360
2340
2320
Wavenumber [cm-1]

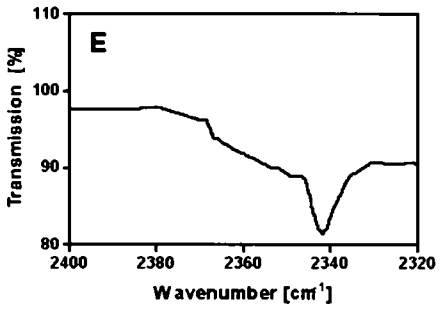
Figure 5.8: FTIR spectra with the subtraction of the protein and H_2O spectrum of BSA with the following additions, A) 3 mM BSA and 20 mM $MgCl_2$, B) 3 mM BSA and 5 mM ATP, C) 3 mM BSA and 3 mM CO_2, D) 3 mM BSA, 20 mM $MgCl_2$ and 5 mM ATP, E) 3 mM BSA, 20 mM $MgCl_2$ and 3 mM CO_2, F) 3 mM BSA, 5 mM ATP and 3 mM CO_2, G) 3 mM BSA, 20 mM $MgCl_2$, 5 mM ATP and 3 mM CO_2.

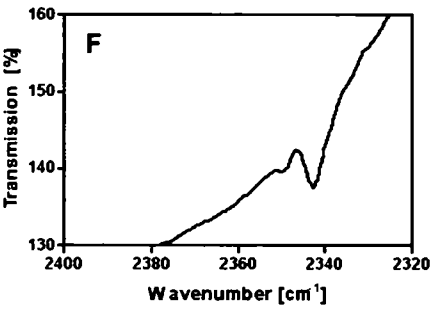

A
Transmission [%]
110
100
90
80
2400
2380
2360
2340
2320
Wavenumber [cm-1]


B
Transmission [%]
110
100
90
80
2400
2380
2360
2340
2320
Wavenumber [cm-1]


C
Transmission [%]
130
120
110
100
2400
2380
2360
2340
2320
Wavenumber [cm-1]


D
Transmission [%]
110
100
90
80
2400
2380
2360
2340
2320
Wavenumber [cm-1]


E
Transmission [%]
110
100
90
80
2400
2380
2360
2340
2320
Wavenumber [cm-1]


F
Transmission [%]
160
150
140
130
2400
2380
2360
2340
2320
Wavenumber [cm-1]

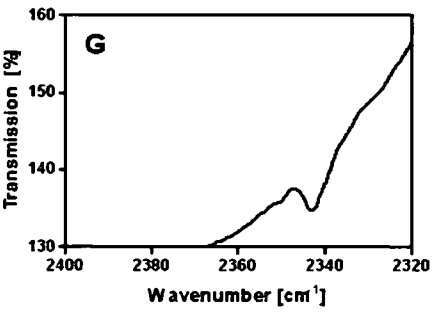

G
Transmission [%]
160
150
140
130
2400
2380
2360
2340
2320
Wavenumber [cm-1]

Figure 5.9: FTIR spectra with the subtraction of the protein and H_2O spectrum of $CyaB1_{595-859}$ with the following additions, A) 3 mM $CyaB1_{595-859}$ and 20 mM $MgCl_2$, B) 3 mM $CyaB1_{595-859}$ and 3 mM CO_2, C) 3 mM $CyaB1_{595-859}$, 20 mM $MgCl_2$ and 3 mM CO_2, D) 3 mM $CyaB1_{595-859}$, 20 mM $MgCl_2$, 5 mM ATP and 3 mM CO_2.

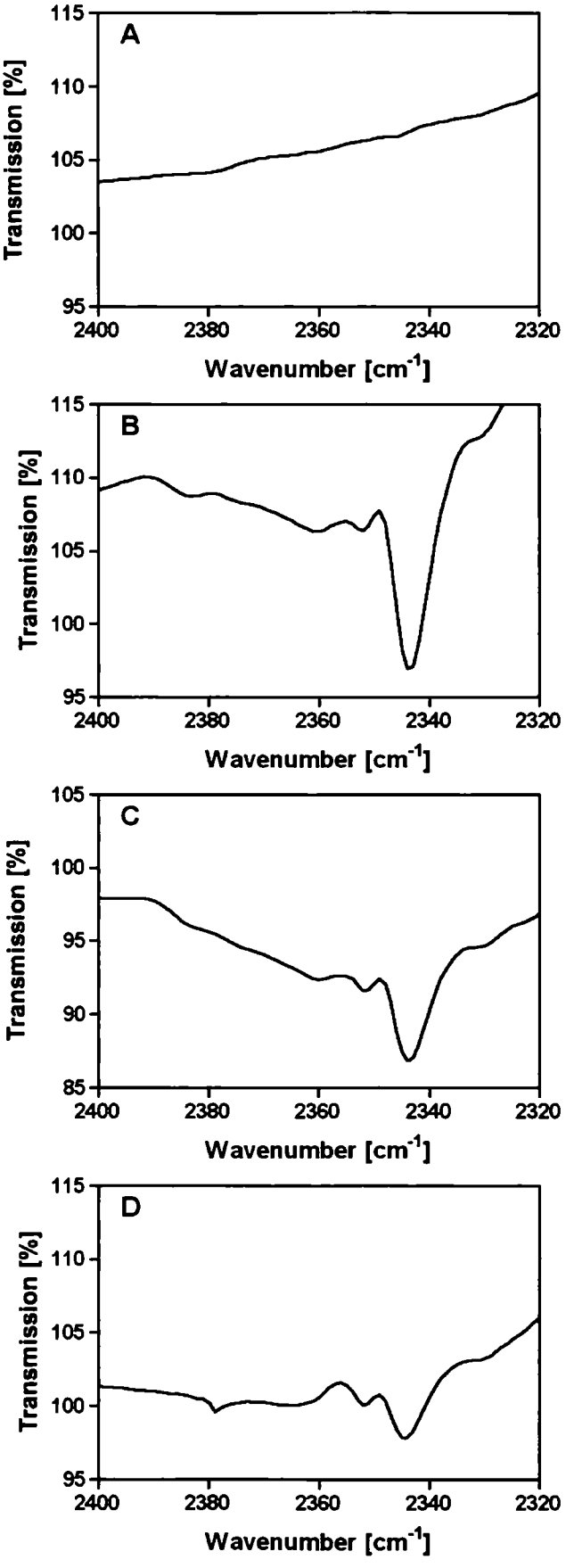

A
B
C
D
Transmission [%]
Wavenumber [cm-1]
115
110
105
100
95
90
85
2400
2380
2360
2340
2320

Figure 5.7C shows the spectrum for buffer and ATP and again demonstrated that there were no signals in the region of interest. There was a faint signal at about 2345 cm^{-1} which represented atmospheric CO_2. Figure 5.7D shows the spectrum of buffer and CO_2. There was a signal at 2345 cm^{-1} corresponding to the asymmetric stretch of CO_2. The spectrum was not flat owing to discrepancies in the subtraction of H_2O and buffer from the spectrum.

Figure 5.7E shows the spectrum of buffer, $MgCl_2$ and ATP and as expected, there were no signals in the region of interest. Figure 5.7F shows the spectrum of $MgCl_2$ and CO_2. Again, the asymmetric CO_2 stretch was seen and there was also a small signal adjacent to that signal at 2355 cm^{-1}, which was initially thought to be $^{13}CO_2$, however, $^{13}CO_2$ accounts for 1.1 % of atmospheric CO_2 and thus would not be visible on the spectra. Therefore the signal at 2355 cm^{-1} must be an artefact of the subtraction of H_2O from the spectrum. Figure 5.7G shows the spectrum of buffer, ATP and CO_2 and the asymmetric CO_2 stretch was again seen at 2345 cm^{-1}. The artefact at 2355 cm^{-1} due to H_2O subtraction was also observed. Figure 5.7H shows the spectrum of $MgCl_2$, ATP and CO_2. The asymmetric CO_2 stretch and the H_2O subtraction artefact at 2355 cm^{-1} were observed, but there were no other discernible signals seen.

These control spectra demonstrated that in the region 2400 – 2320 cm^{-1} the only signals visible were CO_2 signals. The concentration of CO_2 added was the same as the protein concentration, therefore, if CO_2 bound to the AC a depletion of the CO_2 signal would be visible.

Before running FTIR spectra on AC protein, further controls using BSA were done to ascertain whether there were any signals in the 2400 - 2320 cm^{-1} region due to non-specific protein interactions with CO_2.

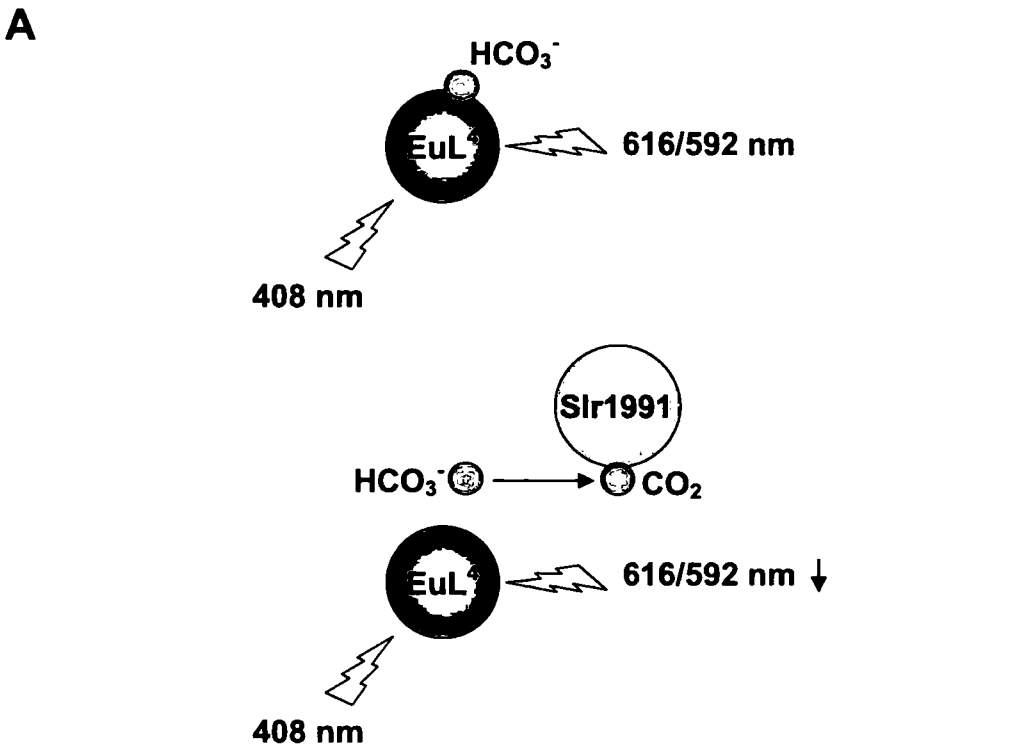
Figure 5.8 shows the region 2400 – 2320 cm^{-1} of the FTIR spectra of BSA with the protein and H_2O spectrum subtracted to give the spectrum of the factors added. Figure 5.8A shows the spectrum of BSA and $MgCl_2$. As with the control spectrum of buffer and $MgCl_2$, there were no signals in the region 2400 2320 cm^{-1}, although atmospheric CO_2 was visible at 2345 cm^{-1}. Figure 5.8B shows the spectrum of BSA and ATP and again, no signals were visible in this region. Figure 5.8C shows the spectrum of BSA and CO_2 and there was an asymmetric CO_2 stretch seen at 2345 cm^{-1}. Figure 5.8D shows the spectrum of

BSA, $MgCl_2$ and ATP and there were no signals visible in this region. Figure 5.8E shows the spectrum of BSA, $MgCl_2$ and CO_2 and there was an asymmetric CO_2 stretch seen at 2345 cm^{-1} and also the H_2O subtraction signal at 2355 cm^{-1} was observed. Figure 5.8F shows the spectrum of BSA and ATP and CO_2 with the asymmetric CO_2 stretch at 2345 cm^{-1} and the H_2O subtraction signal at 2355 cm^{-1}. Figure 5.8G shows the spectrum of BSA, $MgCl_2$, ATP and CO_2 and the asymmetric CO_2 stretch at 2345 cm^{-1} and the H_2O subtraction signal at 2355 cm^{-1} were observed.

These spectra demonstrate that there were no new CO_2 signals formed in the region 2400–2320 cm^{-1} due to non-specific protein interactions to CO_2 and no depletion of the CO_2 signal was observed, thus any new signals or depleted CO_2 signals seen in the spectra with AC would be due to specific CO_2 binding to the AC. Figure 5.9 shows the region 2400–2320 cm^{-1} of the FTIR spectra of $CyaB1_{595\text{-}859}$ with the protein and H_2O spectrum subtracted to give the spectrum of the factors added. Figure 5.9A shows the spectrum with the addition of $MgCl_2$ to the protein and as expected, there were no new signals in the spectrum. Figure 5.9B shows the addition of CO_2 to the protein and the asymmetric stretch signal was visible at 2345 cm^{-1}. There was a small signal adjacent to the CO_2 peak at 2355 cm^{-1} which was the H_2O subtraction signal. There were no other signals observed in the region. Figure 5.9C shows the addition of $MgCl_2$ and CO_2 and again, the asymmetric CO_2 signal and the adjacent H_2O subtraction signal at 2355 cm^{-1} were observed but no other signals were seen.

Figure 5.9D shows the addition of $MgCl_2$ ATP and CO_2 and again, the asymmetric CO_2 signal at 2345 cm^{-1} and the adjacent H_2O subtraction signal at 2355 cm^{-1} were observed. There was also another signal at 2380 cm^{-1}, not seen in the other spectra that could have been CO_2 bound to $CyaB1_{595\text{-}859}$. However, repeated FTIR analysis of these conditions with $CyaB1_{595\text{-}859}$ did not reproduce this signal, suggesting that the signal was formed during the subtraction of protein and H_2O from the spectrum.

Figure 5.10: A) Diagram of HCO_3^- binding to EuL^4, excitation at 408 nm and emission measured at 616/592 nm. HCO_3^- and CO_2 are in equilibrium in solution. CO_2 binding to $Slr1991_{120-337}$ decreases binding of HCO_3^- to EuL^4, decreasing luminescence. B) Structure of HCO_3^- dependent luminescence probe, EuL^4.


A
HCO_3^-
EuL
616/592 nm
408 nm
Slr1991
HCO_3^-
CO_2
EuL
616/592 nm
408 nm

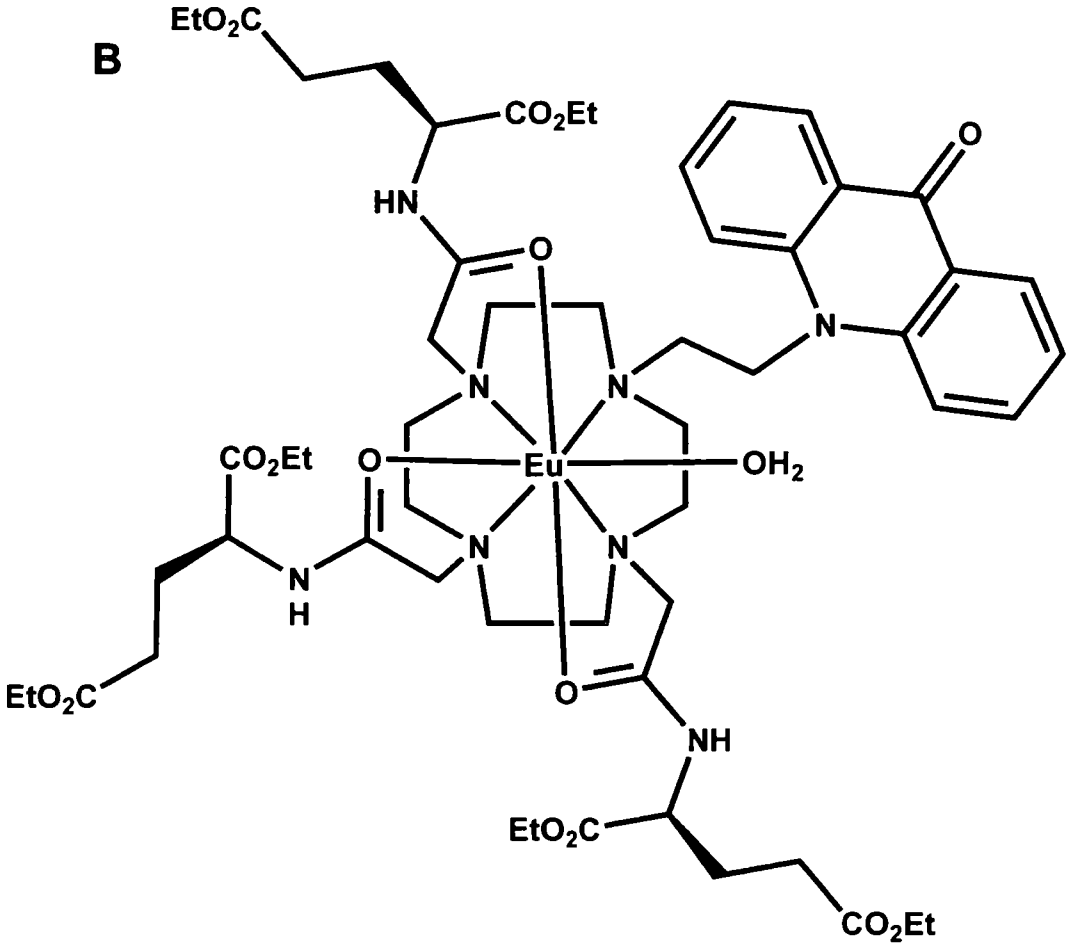

B
EtO_2C
CO_2Et
HN
O
N
N
N
O
Eu
OH_2
CO_2Et
O
N
N
N
H
EtO_2C
O
NH
EtO_2C
CO_2Et

Figure 5.11: HCO_3^- dependent luminescence probe spectroscopy. A) Ci dose dependence curve of the EuL^4 complex as a function of the emission intensity ratio (616/592 nm). B-E) Intensity of EuL^4 complex spectra in the presence of $Slr1991_{120-337}$ with (red line) or without (black line) 5 mM $NaHCO_3$ pH 6.5 (1.94 mM CO_2) and the addition of (B) no additional factors, (C) 100 mM $MgCl_2$, (D) 10 mM ATP and (E) 100 mM $MgCl_2$ and 10 mM ATP. The emission spectra were obtained by exciting EuL^4 at 408 nm.

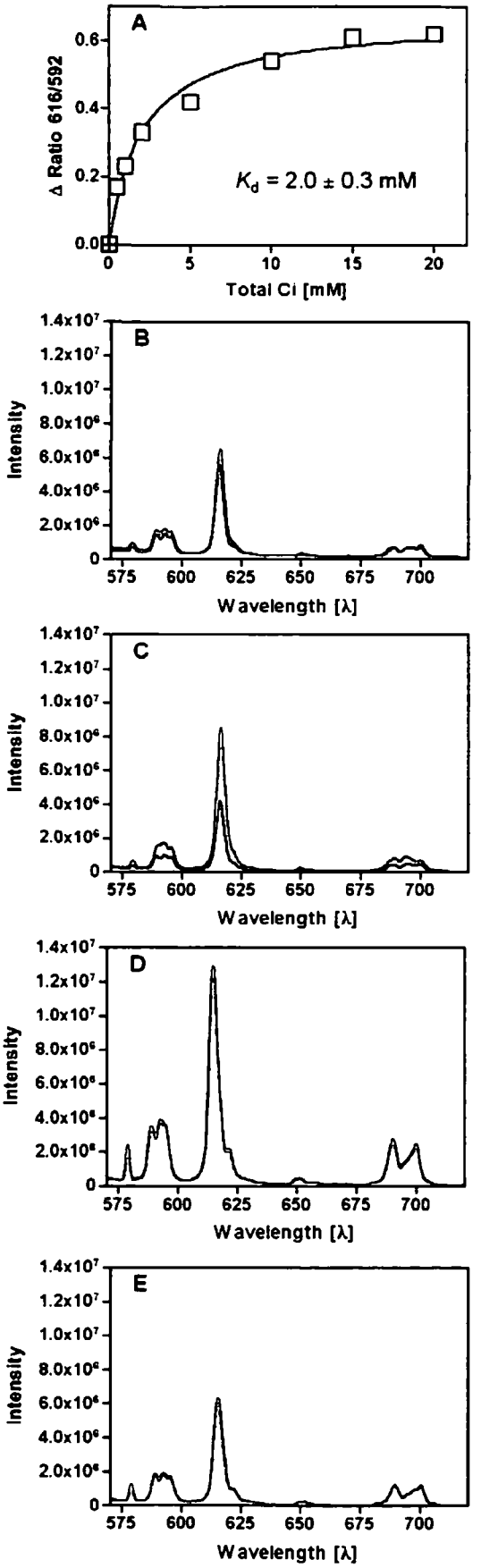

A
Δ Ratio 616/592
0.0
0.2
0.4
0.6
K_d = 2.0 ± 0.3 mM
0
5
10
15
20
Total Ci [mM]
B
C
D
E
Intensity
1.4x10⁷
1.2x10⁷
1.0x10⁷
8.0x10⁶
6.0x10⁶
4.0x10⁶
2.0x10⁶
0
575
600
625
650
675
700
Wavelength [λ]

Table 5.1: HCO_3^- dependent luminescence probe as a function of the emission intensity ratios of A) 616/592 nm or B) 616/700 nm in the presence and absence of 5 mM $NaHCO_3$ pH 6.5 (1.94 mM CO_2) with the following conditions, i) basal, ii) 10 mM $MgCl_2$ iii) 5 mM ATP and iv) 5 mM ATP and 10 mM $MgCl_2$ in the presence of the following proteins, 1) no protein and 2) 500 µM $Slr1991_{120-337}$.

A				
616/592 nm	**1) No protein**		**2) Slr1991**	
i) basal	2.66	Ratio HCO_3^-:basal	3.69	Ratio HCO_3^-:basal
+ HCO_3^-	3.09	1.15	3.94	1.07
ii) $MgCl_2$	2.43	Ratio HCO_3^-:$MgCl_2$	3.89	Ratio HCO_3^-:$MgCl_2$
+ HCO_3^-	3.41	1.40	4.15	1.07
iii) ATP	3.00		2.95	
+ HCO_3^-	2.97		2.96	
iv) ATP+$MgCl_2$	3.15		3.24	
+ HCO_3^-	3.14		3.27	
B				
616/700 nm	**1) No protein**		**2) Slr1991**	
i) basal	4.93	Ratio HCO_3^-:basal	6.48	Ratio HCO_3^-:basal
+ HCO_3^-	6.61	1.34	8.39	1.29
ii) $MgCl_2$	5.28	Ratio HCO_3^-:$MgCl_2$	7.67	Ratio HCO_3^-:$MgCl_2$
+ HCO_3^-	10.02	1.90	10.10	1.32
iii) ATP	4.67		4.77	
+ HCO_3^-	4.76		4.65	
iv) ATP+$MgCl_2$	5.05		5.26	
+ HCO_3^-	5.11		5.41	

FTIR spectroscopy was unable to aid in the identification of the CO_2 binding conditions or the CO_2 binding site on the AC due to issues with the substraction of the H_2O and buffer spectrum which produced artefacts in the spectrum that made it difficult to identify potential new CO_2 peaks or a depletion of the original CO_2 peak.

5.7 HCO_3^- dependent luminescence probe spectroscopy

As CO_2 binding may not cause a measurable conformational change by CD spectroscopy or a change in IR properties as measured by FTIR, a HCO_3^- dependent luminescent probe was used (Bretonniere *et al.*, 2004) to measure the amount of HCO_3^- in solution. Quenching of the probe signal due to CO_2 binding to the protein should occur as CO_2 binds to the protein, reducing the HCO_3^- in solution and therefore luminescence, thus allowing for identification of the conditions required for CO_2 binding to protein (Figure 5.10A).

The probe used for the HCO_3^- dependent luminescence spectroscopy functions by observing europium emissions on exciting the europium complex of the probe at specific wavelengths (Figure 5.10B). The complex, EuL^4 has an acridone chromophore attached, which improves the absorption of light and the movement of energy to the europium ion, thus intensifying the signal (Pal and Parker, 2007). As the excited europium ion relaxes back to the ground state energy level, energy is released as emitted light that gives the complex its characteristic emission spectra.

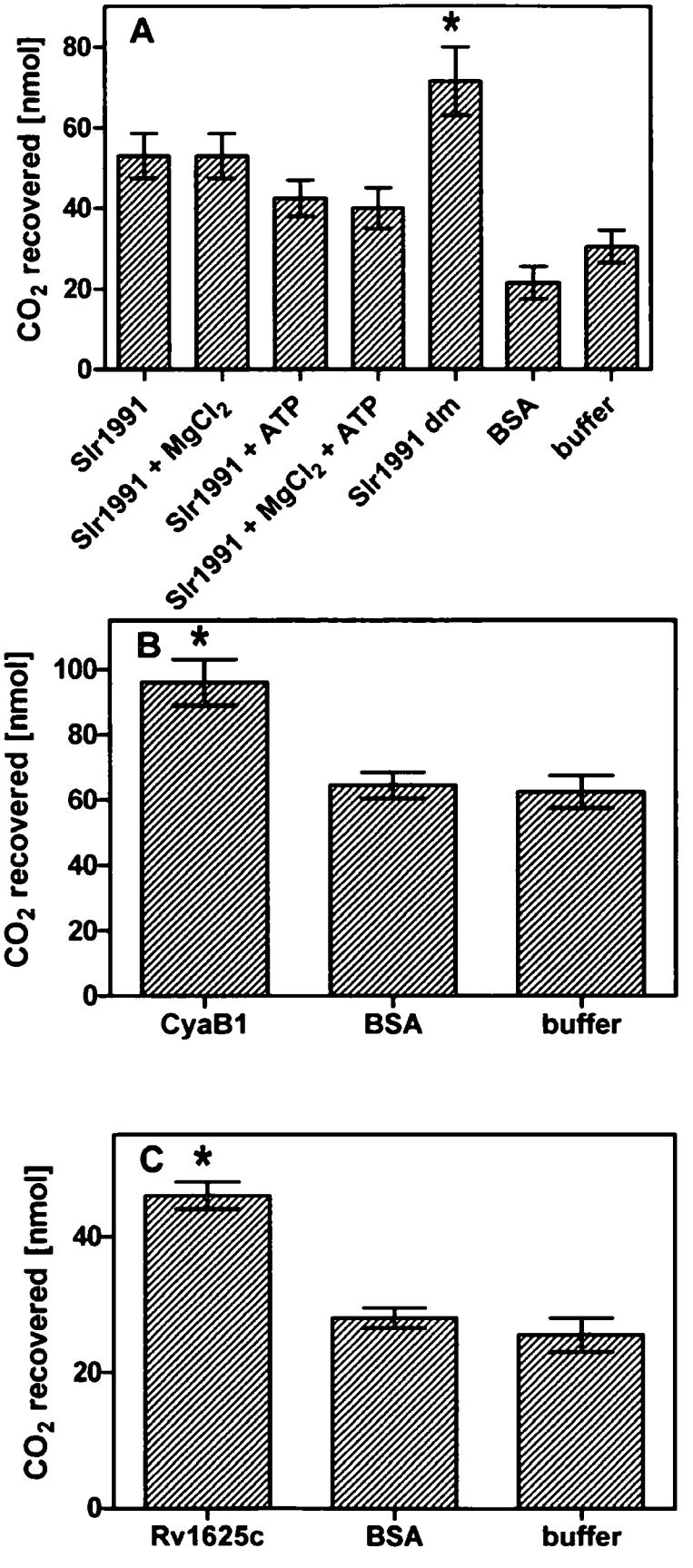
The ratio of the $\Delta J=1$ transition (around 592 nm) or the $\Delta J=4$ transition (around 700 nm) with the $\Delta J=2$ (around 616 nm) transition gives a quantifiable measure of the concentration of HCO_3^- in solution (Bretonniere *et al.*, 2004).

Figure 5.11A shows the HCO_3^- dose dependence curve of the complex EuL^4 and gave a K_d value of 2.0 ± 0.2 mM. This means that the probe will be able to measure small changes in HCO_3^- concentration through a drop in luminescence, which will occur due to CO_2 binding to protein. The small change in HCO_3^- concentration can only be observed if an appropriate concentration of protein is added. If the protein concentration is too low, the change in HCO_3^- concentration due to CO_2 binding will be too small to

observe. The concentration of protein used in these experiments was the maximum concentration obtainable. Figure 5.11B shows the EuL^4 emission spectrum in the presence of 500 µM $Slr1991_{120-337}$ with or without Ci. The 616/592 nm ratio increased in the presence of Ci from 3.69 to 3.94, confirming that HCO_3^- was present (Table 5.1A). Figure 5.11C shows the EuL^4 emission spectrum in the presence of $Slr1991_{120-337}$ and 100 mM $MgCl_2$ with and without Ci. Again the 616/592 nm ratio confirms that HCO_3^- was present on addition of Ci with an increase in ratio from 3.89 to 4.15 (Table 5.1A). Figure 5.11D shows the EuL^4 emission spectrum in the presence of $Slr1991_{120-337}$ and 10 mM ATP with and without Ci. This spectrum demonstrates a magnification of the ratio, where the spectrum overlapped both in the presence and absence of Ci giving a ratio of 2.95 and 2.96 (Table 5.1A). The 616/592 nm ratios derived in the presence of ATP erroneously suggests that HCO_3^- was at the same concentration both in the presence and absence of Ci. The explanation for this result was that the probe also binds to phosphate groups and so was binding to the added ATP, which caused a large excitation and masked any HCO_3^- binding to the probe. Figure 5.11E shows the EuL^4 emission spectrum in the presence of $Slr1991_{120-337}$, 100 mM $MgCl_2$ and 10 mM ATP with and without Ci. This spectrum also shows a magnification of the signal and overlapping of the spectrum with its 616/592 nm ratio (3.24 and 3.27 (Table 5.1A)) and thus also suggests that HCO_3^- was at the same concentration both in the presence and absence of Ci. Thus the results from the EuL^4 emission spectra in the presence of ATP and $MgCl_2$ with ATP are not useful. The same results were seen when calculated with the 616/700 nm ratios (Table 5.1).

The 616/592 nm and 616/700 nm ratios were calculated with the control spectra of no protein present, and the 616/700 nm ratio was also calculated for 0.5 mM $Slr1991_{120-337}$ in the presence of no added co-factor, $MgCl_2$, ATP or $MgCl_2$ and ATP (Table 5.1). These results agree with the initial findings of the 616/592 nm ratios with $Slr1991_{120-337}$ that in the presence of ATP the spectra were distorted and that no difference can be seen between the ratios either in the absence or presence of Ci when ATP was present.

Figure 5.12: A) CO_2 binding assays of Class IIIa and Class IIIb ACs. 30 mM $NaH^{14}CO_3$ at pH 6.5 (11.6 mM $^{14}CO_2$) was assayed with the following additions: 450 µM $Slr1991_{120-337}$, 450 µM $Slr1991_{120-337}$ with 2 mM $MgCl_2$, 450 µM $Slr1991_{120-337}$ with 200 µM ATP, 450 µM $Slr1991_{120-337}$ with 2 mM $MgCl_2$ and 200 µM ATP, 450 µM $Slr1991_{120-337}$ dm (D256A, D300A), 450 µM BSA and 50 µM buffer. B) CO_2 binding assays with 30 mM $NaH^{14}CO_3$ at pH 6.5 (11.6 mM $^{14}CO_2$) with 450 µM $CyaB1_{595-859}$, 450 µM BSA or 50 mM buffer. C) CO_2 binding assays with 30 mM $NaH^{14}CO_3$ at pH 6.5 (11.6 mM $^{14}CO_2$) with 450 µM $Rv1625c_{204-443}$, 450 µM BSA or 50 mM buffer. Error bars are standard error of the mean (n=12) (*$p<0.05$).


A
*
CO_2 recovered [nmol]
0
20
40
60
80
Slr1991
Slr1991 + $MgCl_2$
Slr1991 + ATP
Slr1991 + $MgCl_2$ + ATP
Slr1991 dm
BSA
buffer
B
*
CO_2 recovered [nmol]
0
20
40
60
80
100
CyaB1
BSA
buffer
C
*
CO_2 recovered [nmol]
0
20
40
Rv1625c
BSA
buffer

Further analysis of the ratios of Table 5.1 showed a quenching of the probe signal in the presence of $Slr1991_{120-337}$. By calculating the HCO_3^- present:HCO_3^- absent ratio (at 616/592 nm and 616/700 nm) of both the basal and $MgCl_2$ conditions, with and without $Slr1991_{120-337}$, a decrease in the new calculated ratios in the presence of $Slr1991_{120-337}$ was seen (Table 5.1) (compare column 4 with column 2). These calculated ratios suggest that the CO_2 binding conditions to $Slr1991_{120-337}$ apoprotein, are either in the presence of metal ion or in the absence of metal ion. However, there is no data for CO_2 binding in the presence of ATP or $MgCl_2$ and ATP (due to the distortion in the spectra caused by ATP) that can support either hypothesis or rule out the requirement of ATP for CO_2 binding. To further investigate whether CO_2 is binding to apoprotein alone or binding in the presence of metal, another technique, $^{14}CO_2$ binding assays was used.

5.8 $^{14}CO_2$ binding assays

$^{14}CO_2$ binding assays have been used to identify CO_2 binding to Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase) (Lorimer, 1979). This technique separates protein bound $^{14}CO_2$ from unbound $^{14}CO_2$ using a spin gel filtration column and thus, identifies how much CO_2 is recovered with the protein.

A large amount of protein is required for this assay to be able to recover enough $^{14}CO_2$ from the gel filtration column to count on the scintillation counter. Due to the requirement of large amounts of protein, $Slr1991_{120-337}$ was initially used to observe CO_2 binding with the addition of different conditions as compared to buffer and BSA. Measuring the amount of recovered $^{14}CO_2$ in the buffer containing BSA (Figure 5.12A) shows that the same amount of $^{14}CO_2$ was recovered in both and that $^{14}CO_2$ was not binding non-specifically to protein. Adding $Slr1991_{120-337}$ showed a significant increase in the amount of $^{14}CO_2$ recovered compared to the controls, proving that $Slr1991_{120-337}$ binds CO_2. Repeating this assay in the presence of other conditions such as metal co-factor, substrate, and both metal co-factor and substrate showed no significant change in the amount of $^{14}CO_2$ recovered.

Indeed, the $Slr1991_{120-337}$ D256A, D300A metal binding mutant protein also recovered $^{14}CO_2$ further suggesting that metal is not required for CO_2 binding to the AC. $Slr1991_{120-337}$ therefore, bound CO_2 solely to the apoprotein and no additional cofactors or ions were required. This result disproves the hypothesis of Steegborn *et al.* that CO_2 binds to a metal ion in the active site of the protein.

The $^{14}CO_2$ binding assays were repeated using the other Class IIIb AC, $CyaB1_{595-859}$ (Figure 5.12B) and the Class IIIa AC $Rv1625c_{204-443}$ (Figure 5.12C). In both cases there was a significant increase in the amount of $^{14}CO_2$ recovered with the AC proteins as compared to buffer and BSA. These $^{14}CO_2$ binding assays prove that $Slr1991_{120-337}$, $CyaB1_{595-859}$ and $Rv1625c_{204-443}$ all bind CO_2 to the apoprotein and that CO_2 binding to ACs does not require any other cofactors or ions.

5.9 Conclusion

CO_2 binds to all Class IIIa and Class IIIb AC proteins studied so far, however, the actual CO_2 binding site has not been identified. CO_2 was demonstrated not to bind to the catalytic amino acid residues in the active site of $Rv1625c_{204-443}$, particularly the substrate defining Lys, but was shown to increase the affinity of $Rv1625c_{204-443}$ to Mn^{2+} and increase the co-operativity of the two binding sites in recruiting metal ions. Using spectroscopic techniques such as CD spectroscopy and FTIR spectroscopy failed to identify whether metal ions or any other cofactors were required for CO_2 binding. HCO_3^- dependent luminescence probe spectroscopy suggested that CO_2 bound the apoprotein and that metal may also be required. $^{14}CO_2$ binding assays proved that CO_2 binds to the apoprotein and that CO_2 binding to AC permits metal recruitment without CO_2 binding to the metal.

Elucidation of the CO_2 binding site was the next aim. The most obvious CO_2 binding site is likely to be an amine or amide sidechain group of an amino acid like Lys, Arg or His that form a carbamate with CO_2, as was observed in Rubisco. Analysis of the possible amino acid residues in the active site of $Rv1625c_{204-443}$ was conducted.

6 Investigation into potential CO_2 binding sites

6.1 Introduction

CO_2 has been identified as binding to apoprotein and that no cofactor or substrate is required (see Chapter 5). In this chapter two hypotheses were investigated to explain CO_2 binding to apoprotein. The first hypothesis is that CO_2 is bound covalently to a single amino acid in the active site of $Rv1625c_{204-443}$. The second hypothesis is that CO_2 is bound non-covalently in a pocket defined by multiple amino acids. If CO_2 is bound to a single amino acid in the active site of $Rv1625c_{204-443}$, it is most likely to form a covalent bond with an amino acid that has an amine or amide side group, such as Arg, Lys and His, to produce a carbamate. To test this hypothesis, 12 amino acids were tested by creating single mutant $Rv1625c_{204-443}$ proteins that are conserved or near-conserved residues, as compared with other CO_2 activated Class III ACs. Another potential CO_2 binding site was mutated ($Rv1625c_{204-443}$ K296C) and modified with MTSCE to mimic the carbamate bond hypothesised to form through CO_2 binding. The second hypothesis was investigated by analysing double and triple mutant $Rv1625c_{204-443}$ proteins, identified by aligning the HCO_3^- binding residues of sAC (which also binds CO_2) with $Rv1625c_{204-443}$. Finally Rv1625c protein crystals were grown and exposed to CO_2 in an attempt to observe CO_2 bound to the protein through X-ray diffraction.

6.2 Analysis of potential CO_2 binding residues in $Rv1625c_{204-443}$

CO_2 is most likely to form a covalent bond with the amine or amide side group of an amino acid to form a carbamate. The amino acids that fall into this category are Arg, Lys or His. From the sequence alignment of CO_2 activated Class IIIb ACs (Figure 6.1), conserved or near-conserved Arg, Lys and His residues of $Rv1625c_{204-443}$ were investigated.

Figure 6.1: Alignment of the catalytic domains of prokaryotic Class III ACs and the identification of potential CO_2 binding amino acid side chains. The sequences of Class IIIa AC Rv1625c of *M. tuberculosis* H37Rv, the Class IIIb AC Slr1991 of *Synechocystis* PCC 6803 and the Class IIIb AC CyaB1 from *Anabaena* PCC 7120 are aligned for comparison. Numbers denote amino acid sequence number. Arrows indicate conserved metal binding Asp residues. Triangle indicates conserved substrate binding Lys residue. Circle indicates the polymorphic Asp/Thr of Class IIIa/b ACs. Highlight indicates residues mutated to Ala to form single mutant proteins for CO_2 dependence assays.

```
                                                                  ↓
Rv1625c 204 DTARAEAVMEAEHDRSEALLANMLPASIAERLKEPERNIIADKYDEASVLFADIVGFTERAS
Slr1991 120 -TQT------------------NVLH---------ERRLIT--------VLVADMRNFTGMAQ
CyaB1   595 ----------------------------------------GERKEVTVLFSDIRGYTTLTE

                                          ▼  ↓
Rv1625c 266 STAPADLVRFLDRLYSAFDELVDQHGLEKIKVSGDSYMVV--SGVPSPR-PDHTQALADFAL
Slr1991 147 QVEEELLSMLIGNWFRQAGHILREAGSWVDKYIGDAVMAIWFHGYNEATPAEIIQILHAVNR
CyaB1   616 NLGAAEVVSLLNQYFETMVEAVFNYEGTLDKFIGDALMAV--FGAPLPLTENHAWQAVQSAL

                                                   ●
Rv1625c 325 DMTNVAAQLKDPR----GNPVPLRVG--LATGPVVAGVVGS-RRFFYDVWGDAVNVASRMES
Slr1991 209 LQAMTAKLNQKYE----LPF-PLRIGTGINTGYAMVGNTGSGDHPDYTAIGDTVNAAFRLES
CyaB1   676 DMRQRLKEFNQRRIIQPQPQIKIGIG--ISSGEVVSGNIGSHKRMDYTVIGDGVNLSSRLET

Rv1625c 380 -TDSVGQIQ--VPDEVYERLKD--DFVLREKGH-INVKGKGVMRTWY---------LIGRKV
Slr1991 267 ATKQAHFDLAMSEKTFSYLQDLPQWSATVQQHTIELKGYTNPITIYGLPFATLGTLLDHTSL
CyaB1   736 -VTKEYGCDIILSEFTYQLCSDRIWVRQLDKIRVKGKHQAVNIYELISDRSTPLDDNTQEFL

Rv1625c 427 AADPGEVRGAEPRTAGV
Slr1991 328 EDTTPASEGP
CyaB1   797 FHYHNGRTAYLVRDFTQAIACFNSAKHIRPTDQAVNIHLERAYNYQQTPPPPQWDGVWTIFTK
```

Figure 6.2: SDS-PAGE gel of $Rv1625c_{204-443}$ mutant proteins. 2 µg of each protein was applied to each lane with molecular weight standards on the left of each gel. Gels were stained with Coomassie Brilliant Blue G. Gel A): Lane 1) R310A 2) H315A 3) K334A 4) R337A 5) R344A 6) R360A. Gel B): Lane 7) R381A 8) R395A 9) K397A 10) R405A 11) K411A 12) K413A.

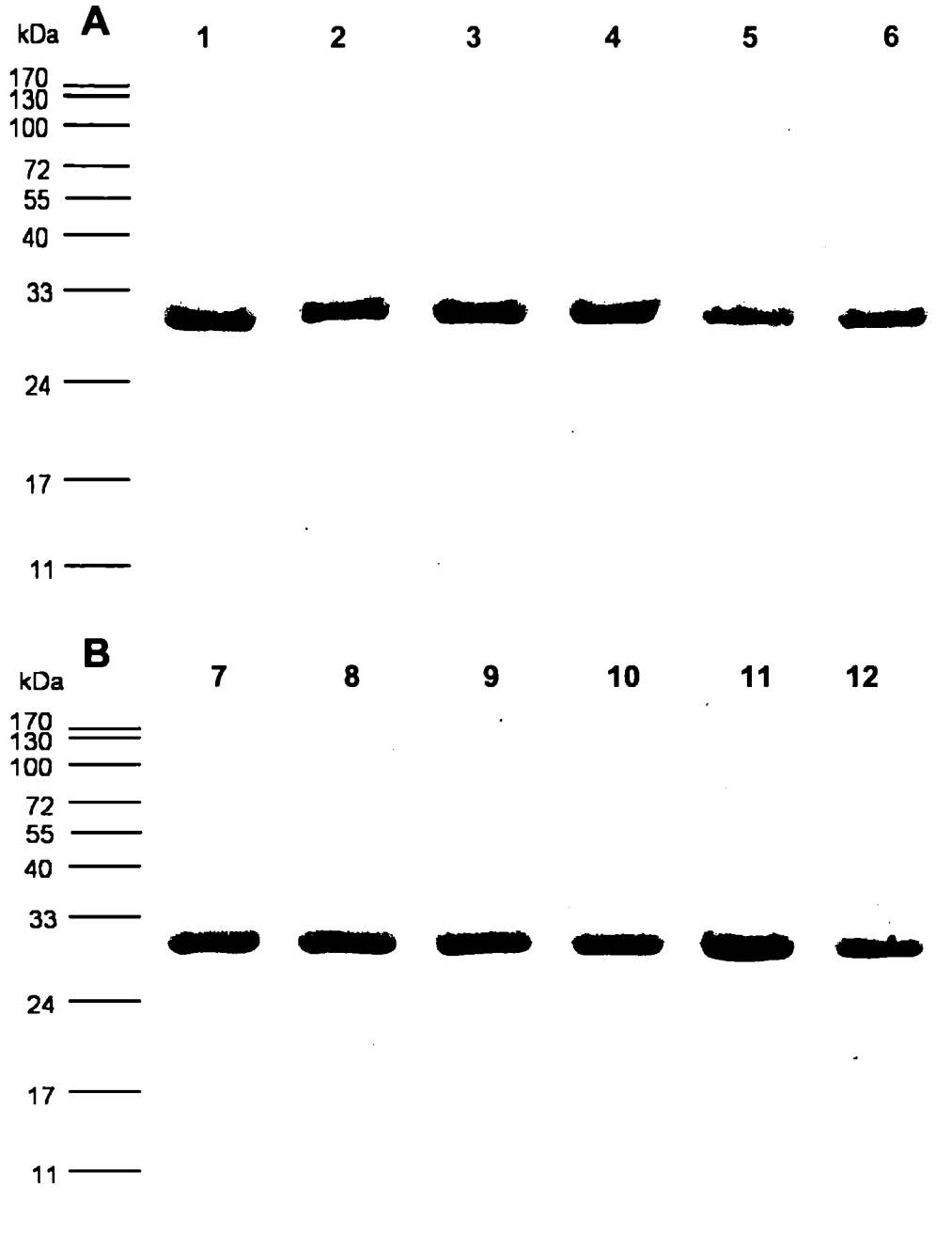
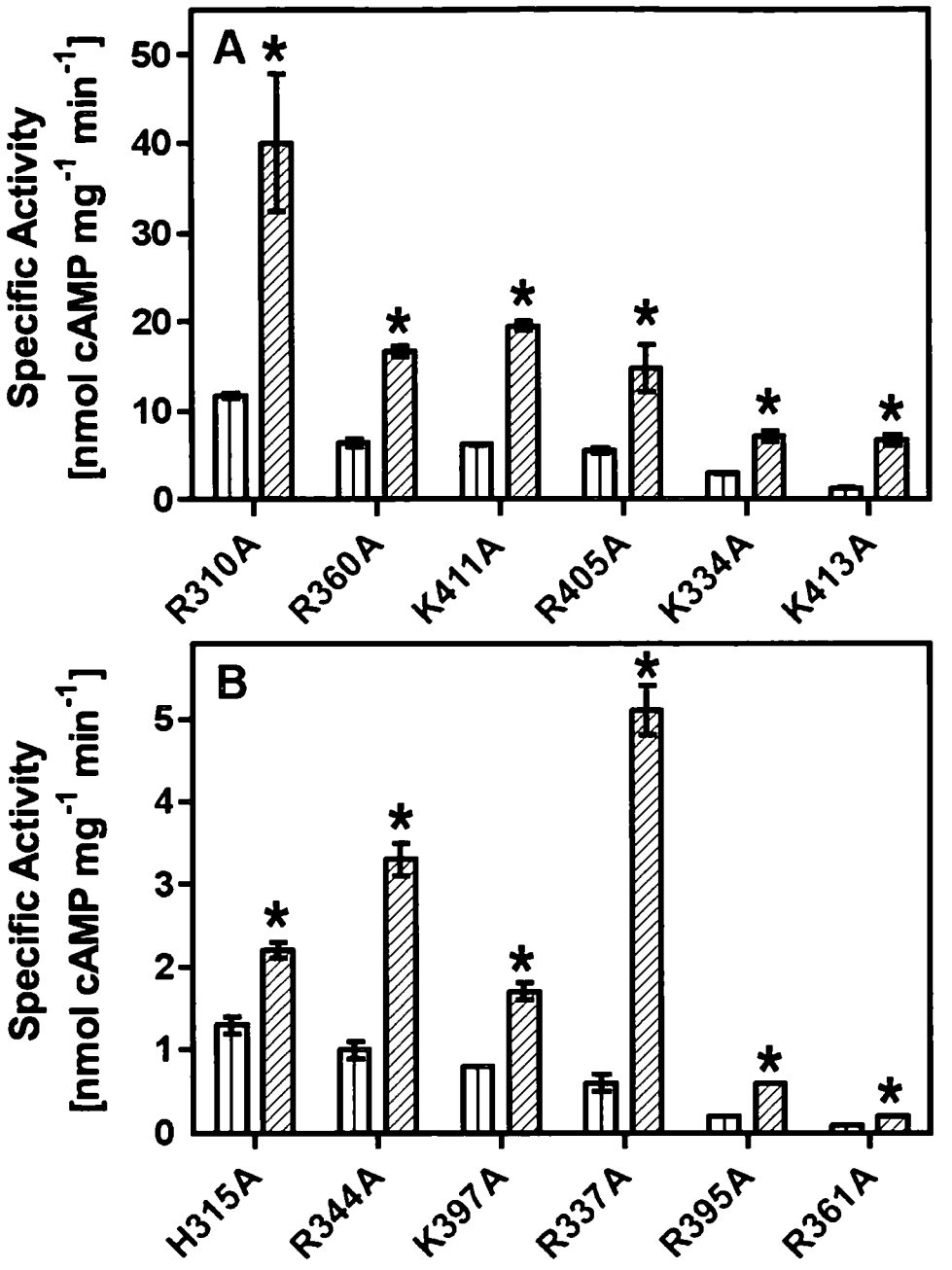

A
kDa
1
2
3
4
5
6
170
130
100
72
55
40
33
24
17
11
B
kDa
7
8
9
10
11
12
170
130
100
72
55
40
33
24
17
11

Figure 6.3: CO_2 dependence assays for $Rv1625c_{204-443}$ mutant proteins. Assays were performed at 37 °C at pH 6.5 (50 mM MES) with 2 mM Mn^{2+} in the presence of 200 µM ATP and 20 mM NaCl or Ci for 40 minutes with a pre-incubation time of 30 minutes. The bars on the left with vertical stripes represent the addition of 20 mM NaCl and the bars on the right with diagonal stripes represent the addition of 20 mM Ci. Error bars are standard error of the mean (n=6) (*$p<0.05$) compared to corresponding Cl^- value. A) 1.8 µM R310A, 1.8 µM R360A, 1.8 µM K411A, 1.8 µM R405A, 1.8 µM K334A, 1.8 µM K413A. B) 1.8 µM H315A, 1.8 µM R344A, 1.8 µM K397A, 14.5 µM R395A, 3.6 µM R361A.


A
Specific Activity
[nmol cAMP mg-1 min-1]
50
40
30
20
10
0
R310A
R360A
K411A
R405A
K334A
K413A
B
Specific Activity
[nmol cAMP mg-1 min-1]
5
4
3
2
1
0
H315A
R344A
K397A
R337A
R395A
R361A

A sequence alignment of the catalytic domains of three prokaryotic Class IIIa and Class IIIb ACs that are stimulated by CO_2, $Rv1625c_{204-443}$, $Slr1991_{120-337}$ and $CyaB1_{595-859}$, was used to identify the Arg, Lys or His residues that were conserved or that were near-conserved (Figure 6.8) and that may be potential CO_2 binding residues.

In total, 12 residues in $Rv1625c_{204-443}$ were identified that were potential CO_2 binding sites due to conservation in the Class IIIb ACs Slr1991 from *Synechocystis* PCC 6803 and CyaB1 from *Anabaena* PCC 7120. Single Ala mutant proteins of each residue were created by site-directed mutagenesis. All the mutant proteins were overexpressed in *E. coli* and purified by affinity chromatography to homogeneity. The mutant proteins were analysed for purity by SDS-PAGE.

The SDS-PAGE gel of the purified mutant proteins (Figure 6.2) shows a protein band at approximately 27 kDa for all mutant proteins, similar to the molecular weight of the $Rv1625c_{204-443}$ wt protein. All the mutant proteins appeared pure except R395A and K397A, which had contaminating bands at approximately 25, 72 and 100 kDa for R395A and 25 kDa for K397A, but both mutant proteins were estimated to be 95 % pure by comparison of the intensity of the gel bands.

CO_2 dependence assays were performed on the 12 mutant proteins to ascertain if there was any ablation of CO_2 stimulation. Ablation of activity would give direct evidence that the mutated residue was the CO_2 binding site for $Rv1625c_{204-443}$.

Figure 6.3 shows firstly that all the mutant proteins were active, with varying basal specific activities and secondly, that all the mutant proteins were stimulated by CO_2 with fold stimulation varying from 2-fold to 8-fold. Interestingly, a few of the mutant proteins (R310A, R360A, K411A, R405A and K334A) had higher basal specific activities than wt protein, suggesting that the mutations of these residues increase catalysis, most likely through the faster release of cAMP from the active site. Further work on the kinetics of these mutant proteins would need to be conducted to identify their mechanism of stimulation. It is also interesting to note the large increase in CO_2 fold stimulation of certain mutant proteins (R310A and R337A) suggests that these residues may constrain the response of $Rv1625c_{204-443}$ to CO_2. Again further

work on the kinetics of these mutant proteins would need to be conducted to identify their function. Another interesting note is the large difference in basal specific activity of R360A and R361A, residues that are adjacent to each other, which suggest that R361 plays an important unknown role in catalysis, possibly in stabilising the enzyme structure.

The CO_2 dependence assays on the 12 selected Arg, Lys and His mutant proteins did not identify the CO_2 binding site. This finding questions the theory that CO_2 is binding to a single amino acid sidechain group to form a carbamate and it is unlikely that CO_2 binds to a carboxylate group. CO_2 has also previously been demonstrated to not require metal or substrate for stimulation, suggesting that CO_2 does not bind to metal or substrate. It is possible that more than one residue is involved in CO_2 binding, thus creating a single mutant protein will not help to identify the CO_2 binding site as the other CO_2 binding residue will still facilitate the CO_2 stimulation of the enzyme.

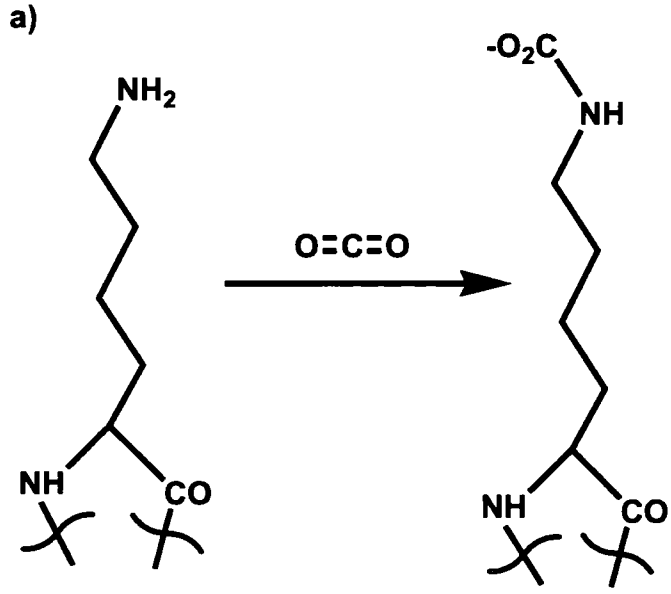
6.3 Modification of Rv1625c$_{204-443}$ K296C to form a carbamate

As the 12 potential CO_2 binding amino acids were all shown to not affect CO_2 activation, the substrate defining Lys of Rv1625c$_{204-443}$. K296 was re-investigated. This was because the corresponding Lys in CyaB1 of *Anabaena* PCC 7120 has previously been demonstrated to be required for Ci activity by the ablation of CO_2 stimulation through a K646A point mutant protein (Cann *et al.*, 2003). The Rv1625c$_{204-443}$ K296A mutant (Section 5.2) was previously shown to be stimulated by CO_2 and thus was not solely responsible for CO_2 binding. However K296A may be indirectly involved in CO_2 stimulation of Rv1625c$_{204-443}$, for example through stabilisation of the active site during catalysis or through being part of a CO_2 binding pocket.

Chemical modification of a catalytic amino acid was used to investigate a possible role for K296 as a redundant part of a CO_2 binding pocket in Rv1625c$_{204-443}$. If CO_2 forms a carbamate with the amine sidechain group of K296 in conjunction with other binding amino acid partners, it is reasonable to presume that the resulting carbamate will be important in enzyme activation.

Figure 6.4: Schematic diagram to demonstrate the isosteric resemblance of a) the hypothesised carbamate formation on Lys with CO_2, and b) the modification of Cys mutant with MTSCE.

a)



b)

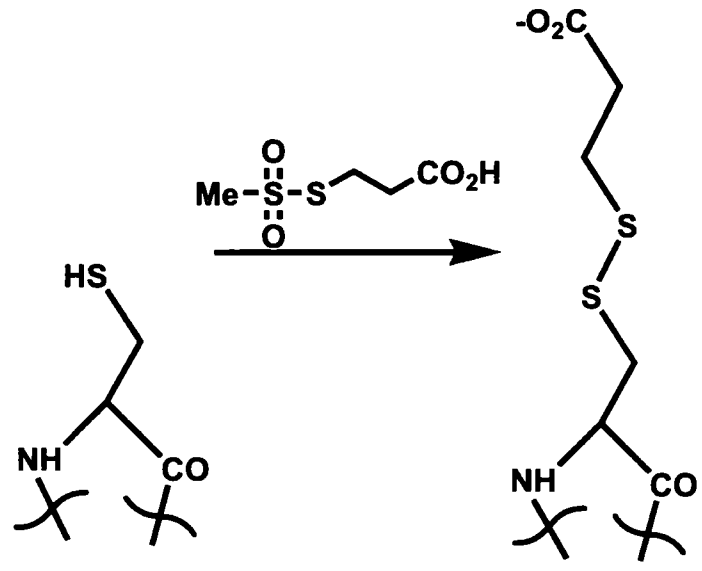
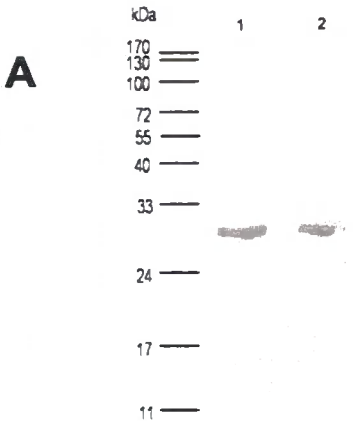
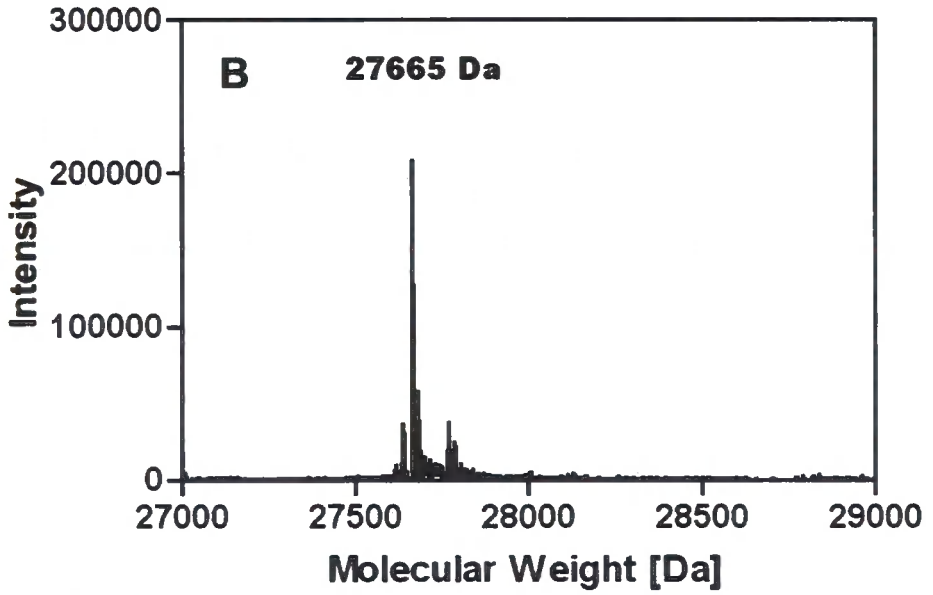


Figure 6.5: A) SDS-PAGE gel of Rv1625c$_{204-443}$ mutant proteins. 2 µg of 1) K296C and 2) K296C modified with MTSCE was applied to each lane with molecular weight standards on the left of each gel and the gel was stained with Coomassie Brilliant Blue G. B) MS data for Rv1625c$_{204-443}$ K296C. C) MS data for Rv1625c$_{204-443}$ K296C modified with MTSCE.


A
kDa
1
2
170
130
100
72
55
40
33
24
17
11


B
27665 Da
300000
200000
100000
0
Intensity
27000
27500
28000
28500
29000
Molecular Weight [Da]

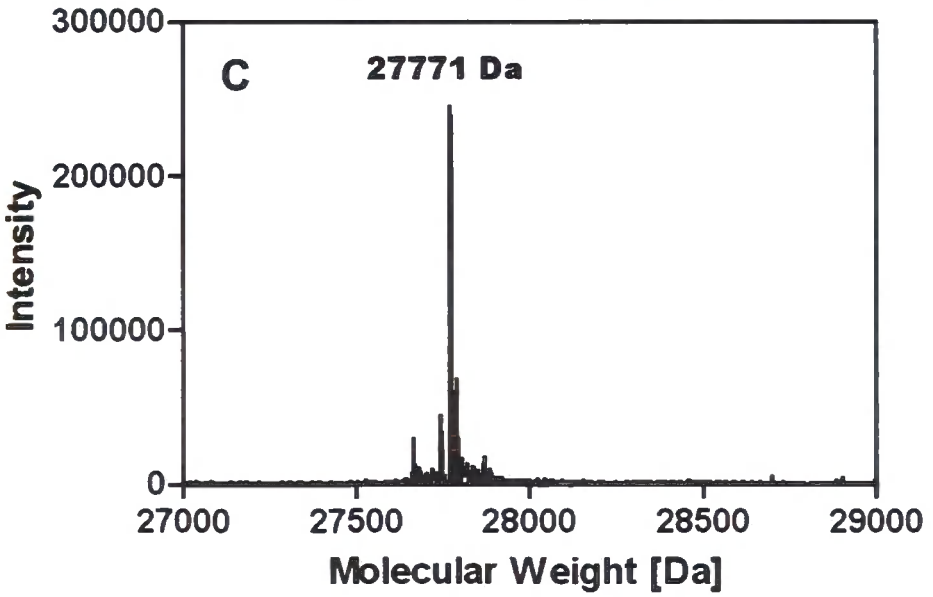
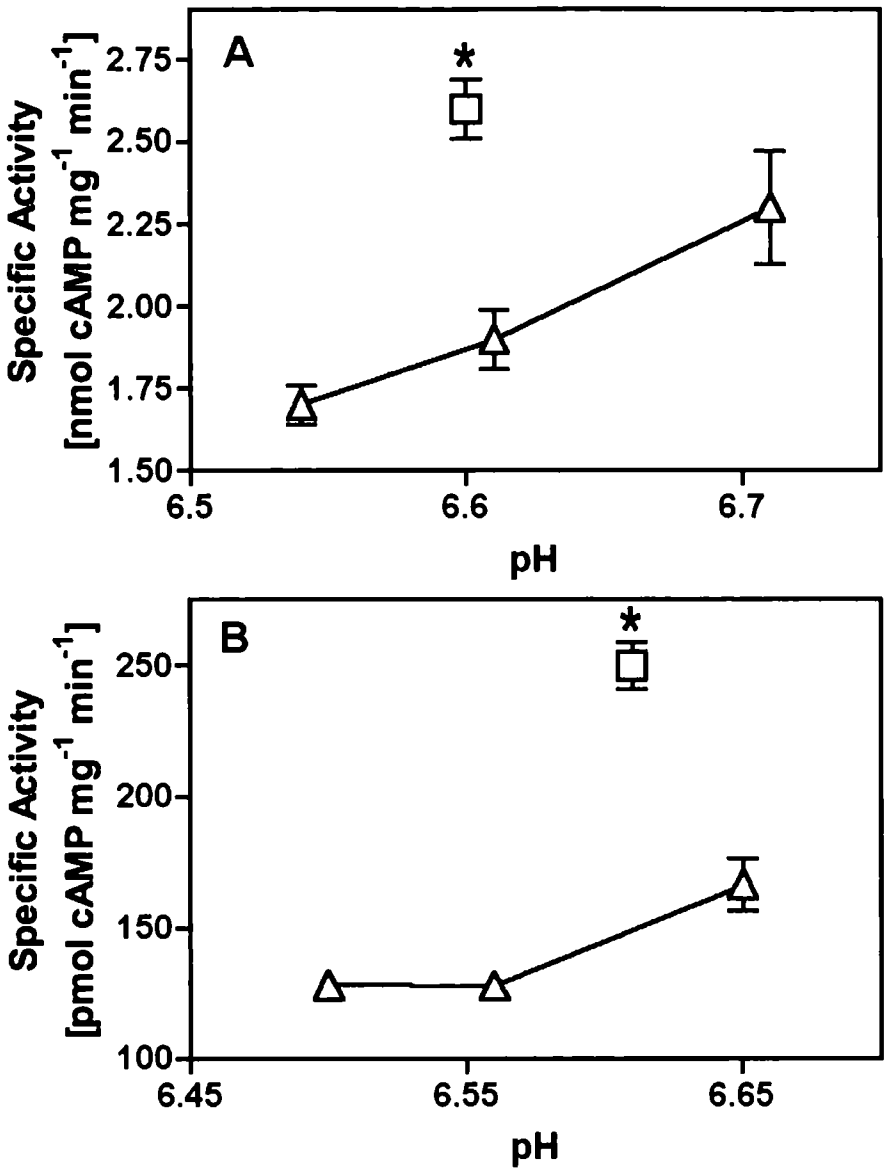

C
27771 Da
300000
200000
100000
0
Intensity
27000
27500
28000
28500
29000
Molecular Weight [Da]

Figure 6.6: CO_2 dependence assay for $Rv1625c_{204-443}$ K296C mutant proteins. K296C mutant proteins were assayed at 37 °C for CO_2 sensitivity at pH 6.5 (50 mM MES) with 2 mM Mn^{2+} in the presence of 200 µM ATP and 20 mM $NaHCO_3$ or NaCl for 40 minutes with a pre-incubation time of 30 minutes. Error bars are standard error of the mean (n=6) (*$p<0.05$) compared to Cl^- value at corresponding pH value. A) 7.2 µM K296C. B) 7.2 µM K296C modified with MTSCE.


A
*
Specific Activity
[nmol cAMP mg-1 min-1]
2.75
2.50
2.25
2.00
1.75
1.50
6.5
6.6
6.7
pH
B
*
Specific Activity
[pmol cAMP mg-1 min-1]
250
200
150
100
6.45
6.55
6.65
pH

The formation of a stable carbamate at K296, as a chemical mimic for CO_2 activation of the enzyme, was therefore investigated.

The formation of a steric analogue of the carbamate at K296 may permanently activate Rv1625c$_{204-443}$ and therefore the enzyme will no longer be responsive to CO_2. K296 was mutated to Cys, which allowed for the modification of the thiol group of Cys with MTSCE (see Section 3.2.17) to form a carboxylate group that mimics the hypothesised carbamate binding of CO_2 to K296 (Figure 6.4). Therefore the modified Rv1625c$_{204-443}$ K296C should be CO_2 insensitive as the modification mimics the CO_2 bound form of Rv1625c$_{204-443}$, however, specific activity should be comparable to CO_2 activated Rv1625c$_{204-443}$ wt.

K296C was overexpressed in *E. coli* and purified by affinity chromatography to homogeneity to generate a recombinant protein. The protein was analysed for purity by SDS-PAGE and MS. The SDS-PAGE gel of K296C (Figure 6.5A1) shows a single protein band at approximately 27 kDa. There were no contaminating protein bands. The MS data (Figure 6.5B) for the K296C mutant protein confirmed that the actual molecular weight of the purified protein matched the theoretical molecular weight of 27665 Da for Rv1625c$_{204-443}$ K296C mutant protein

500 µM of the K296C mutant protein was modified with MTSCE to add the carboxylate group (Figure 6.4). The modified protein was analysed for purity by SDS-PAGE and MS. The SDS-PAGE gel of K296C (Figure 6.5A2) shows a single protein band at approximately 27 kDa. There were no contaminating protein bands. The MS data (Figure 6.5C) for the modified K296C mutant protein confirmed that the actual molecular weight of the purified protein matched the theoretical molecular weight of 27771 Da for the modified Rv1625c$_{204-443}$ K296C mutant protein.

The K296C and modified K296C mutant proteins were analysed for CO_2 dependence by performing a CO_2 dependence assay against a Cl^- control curve. A small shift in the pH value of the assay containing CO_2 was noticed that could be suggested as the cause of the increase in specific activity, so a Cl^- control curve at three different pH values was conducted to differentiate between small changes in assay pH and reduced responses to CO_2. The result of the CO_2 dependence assay on K296C (Figure 6.6A) shows that

K296C was stimulated by CO_2 with a 1.4 fold stimulation. The result of the CO_2 dependence assay on the modified K296C (Figure 6.6B) shows that the modified K296C was also stimulated by CO_2 with a CO_2 fold stimulation of 1.6. Noticeably, there is a 1000-fold drop in basal specific activity of the modified K296C, most likely due to the modification of the Cys at K296 in the active site, which placed a larger carbamate group in the active site, thus making the binding of ATP more difficult and decreasing specific activity. A kinetics assay would need to be performed to determine this hypothesis precisely. The Cl^- assays demonstrated that the increase in specific activity was due to CO_2 stimulation and not a change in pH. As $Rv1625c_{204-443}$ K296C mutant protein was stimulated by CO_2, the substrate defining Lys could not be the sole CO_2 binding site in $Rv1625c_{204-443}$. As the modified protein is also stimulated, K296 would appear to not be part of a CO_2 binding pocket despite the contrary findings in CyaB1 of *Anabaena* PCC 7120 (Cann *et al.*, 2003). The identification of the substrate defining Lys as being involved in Ci stimulation was discovered in the Class IIIb AC $CyaB1_{595-859}$, thus it may be possible that Class IIIa ACs and Class IIIb ACs have different mechanisms of CO_2 stimulation.

6.4 sAC structure

During the third year of research on this thesis, the mammalian sAC protein was crystallised with HCO_3^- bound and a diffraction set collected by another research group (Saalau-Bethell *et al.*, 2007). Analysis of the diffraction data of the sAC crystal with HCO_3^- bound showed that the three residues that are required for HCO_3^- binding are the substrate defining Lys K95, the transition state stabilising Arg, R176 and a valine (Val) of unknown activity, V167 (Figure 6.7). The diffraction data was entered into Pymol, a program that generates protein structures and the HCO_3^- was identified by the diffraction data to sit in a binding pocket. HCO_3^- binds to the side chains of K95 and R176 and also binds to the 'backbone' main chain of V167. sAC is activated by CO_2 and HCO_3^- but whether the binding site is the same for CO_2 and HCO_3^- or whether there are two binding sites is unknown.

Figure 6.7: Pymol generated image of HCO_3^- binding to sAC from crystallisation data (Saalau-Bethell *et al.*, 2007). a) Whole catalytic domain of sAC with HCO_3^- (red molecule) in binding pocket. b) HCO_3^- (red molecule) binding to residues K95 (blue residue on left), V167 (grey residue on right) and R176 (blue residue on right).

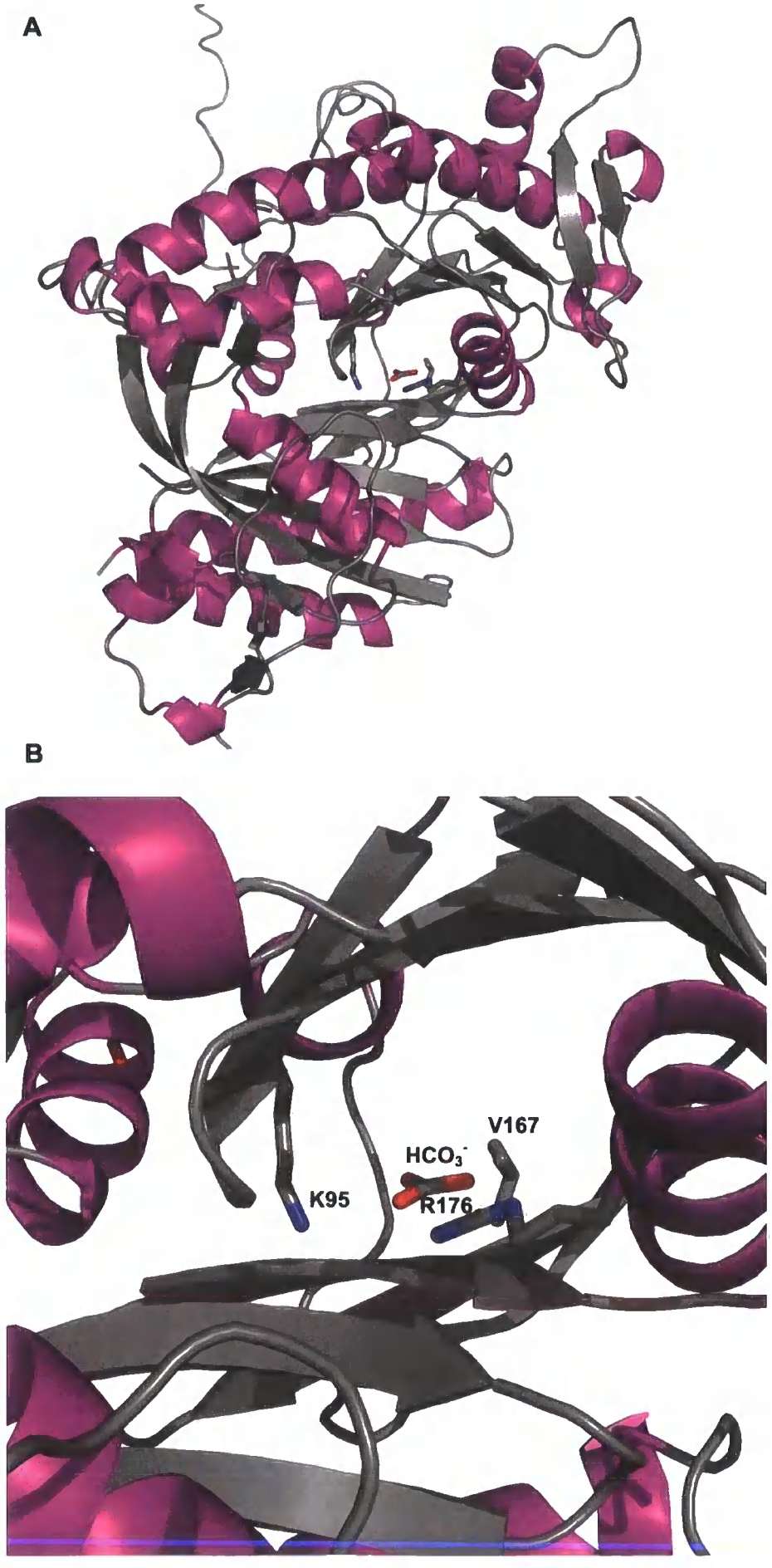

A
B
V167
HCO_3^-
K95
R176

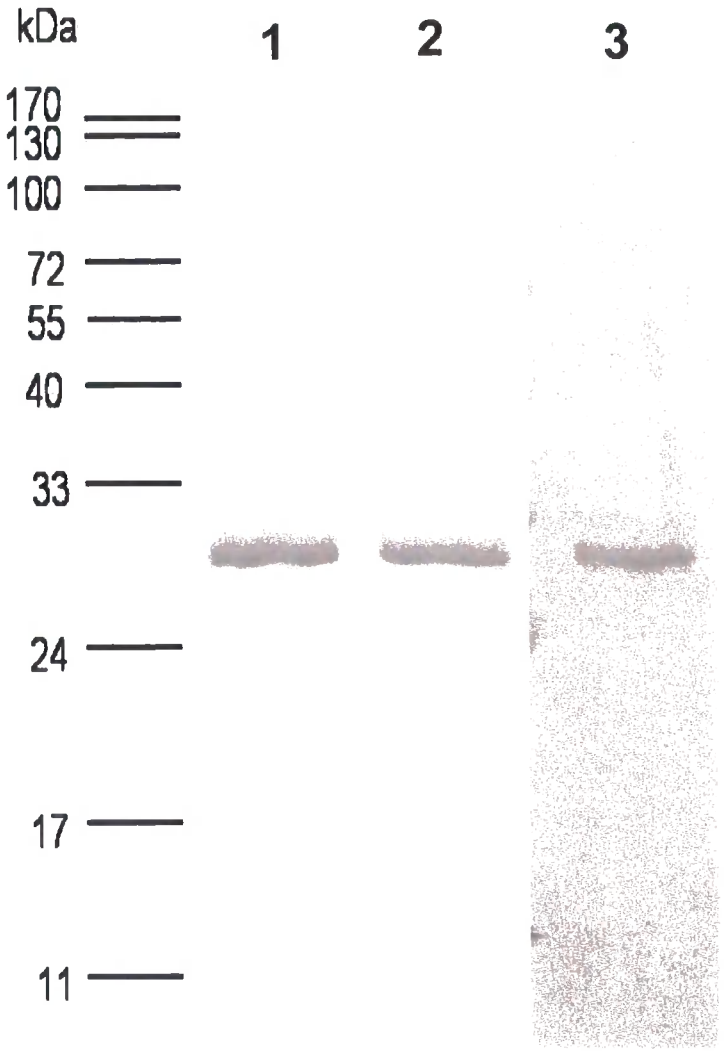
Figure 6.8: Alignment of the catalytic domain of Rv1625c from *M. tuberculosis* and sAC from *Homo sapiens*. K95, V167 and R176, the residues of sAC that bound Ci (arrows) were aligned with the catalytic domain of Rv1625c to obtain the corresponding residues in Rv1625c.

```
                                                                                  ↓
Mycobacteria  Rv1625c  252 VLFADIVGFTERASST........APADLVRFLDRLYSAFDELVDQHGLEKIKVSGDSYM
Human         sAC1      43 LMFVDISGFTAMTEKFSTAMYMDRGAEQLVEILNYYISAIVEKVLIFGGDILKFAGDALL

Mycobacteria  Rv1625c  304 VV.......................SGVPRPRPDHTQALADFALDMTNVAAQLKDPRGN
Human         sAC1     103 AL.....................W........KVERKQLKNIITVVIKCSLEIHGLFEA

                                                          ↓             ↓
Mycobacteria  Rv1625c  340 .....PVPLRVGLATG...PVVAGVVGSRRFFYDVW.GDAVN...VASRMEST
Human         sAC1     133 KEVEEGLDIRVKIGLAAGHITMLVFGDETRNYFLV.IGQAVDD..V..RLAQN
```

Figure 6.9: SDS-PAGE gel of $Rv1625c_{204-443}$ mutant proteins. 2 µg of 1) V366A 2) K296A, R376A and 3) K296A, V366A, R376A was applied to each lane with molecular weight standards on the left of each gel and the gel was stained with Coomassie Brilliant Blue G.


kDa
1
2
3
170
130
100
72
55
40
33
24
17
11

6.5 Analysis of double and triple mutant proteins of $Rv1625c_{204-443}$

As the single K296 mutant protein of $Rv1625c_{204-443}$ and the twelve identified single amine group residue mutant $Rv1625c_{204-443}$ proteins have all been demonstrated not to be the sole CO_2 binding site, the possibility that there was more than one residue involved in CO_2 binding was considered. Thus CO_2 may be forming a carbamate with the amino side group of a residue, but may also be stabilised with other residues in the active site. No ablation of CO_2 activity was observed in the single mutant proteins as the other residues in the binding site were able to 'bind' CO_2 and stimulate $Rv1625c_{204-443}$.

A crystal structure of mammalian sAC with bound HCO_3^- showed that the residues K95, V167 and R176 all bound HCO_3^-. By creating an alignment of mammalian sAC and Rv1625c, the corresponding residues in Rv1625c can be identified and mutated. As Rv1625c corresponds to half of a mammalian AC, it is likely that the same binding pocket will exist in Rv1625c as in mammalian sAC (Figure 6.7). A logical experiment would be to create the corresponding single, double and triple mutant proteins in $Rv1625c_{204-443}$ to ascertain by AC assay, whether these residues are required for CO_2 binding in a Class IIIa AC.

The corresponding three residues that bind HCO_3^- in sAC, the substrate defining Lys, the transition state stabilising Arg and the backbone residue Val, were identified in Rv1625c by an alignment of the catalytic domains of sAC and Rv1625c (Figure 6.8). The Rv1625c residues are K296, V366 and R376. Single mutant proteins of each of these residues were constructed, overexpressed and assayed for CO_2 dependency as was the double substrate defining mutant protein, K296A, R376A and the triple mutant protein, K296A, V366A, R376A. As K296A and R376A single mutant proteins have previously been overexpressed and purified (See Section 5.2), V366A single mutant protein was also overexpressed and purified. The K296A, R376A double mutant protein and K296A, V366A, R376A triple mutant protein were formed through restriction enzyme digestion and ligation of the appropriate single

mutant DNA, double checked through forward and reverse DNA sequencing of the DNA and finally expressed and purified.

The $Rv1625c_{204-443}$ V366A single mutant protein, K296A, R376A double mutant protein and K296A, V366A, R376A triple mutant protein were overexpressed in *E. coli* and purified by affinity chromatography to homogeneity to generate three recombinant proteins. The mutant proteins were analysed for purity by SDS-PAGE.

The SDS-PAGE gel of the expressed and purified $Rv1625c_{204-443}$ mutant proteins (Figure 6.9) shows a single protein band at approximately 30 kDa for the V366A single mutant protein (Figure 6.9(1)). The K296A, R376A double mutant protein also shows a band at approximately 30 kDa (Figure 6.9(2)) and the K296A, V366A, R376A triple mutant protein also shows a band at approximately 30 kDa (Figure 6.9(3)). The theoretical weight of $Rv1625c_{204-443}$ is 26790 Da and so the expressed $Rv1625c_{204-443}$ mutant proteins match the theoretical weight of the $Rv1625c_{204-443}$ mutant and the mutant proteins are pure as there are no contaminating protein bands seen on the SDS-PAGE gel. The three single mutant proteins, K296A, V366A and R376A, the double mutant protein, K296A, R376A and the triple mutant protein, K296A, V366A, R376A were analysed for CO_2 responsiveness by performing CO_2 dependence assays against Cl^- control curves. The Cl^- control curves were included to investigate the theory that a small change in pH of the assay due to the addition and degassing of CO_2 may be the cause of the fold stimulation of $Rv1625c_{204-443}$ and not that $Rv1625c_{204-443}$ is CO_2 sensitive. The pH values of the Cl^- control curves were between pH 6.45 and 6.6 and the pH value of the assay containing CO_2 was measured and kept within the Cl^- curve. Figure 6.10A shows the CO_2 dependence assay for the K296A single mutant protein. K296A is stimulated by CO_2 with a 1.5-fold CO_2 stimulation. The specific activities of the Cl^- curve and the CO_2 stimulated assay correspond with the results previously observed with K296A (see Section 5.2). Figure 6.10B shows that the V366A single mutant protein is stimulated by CO_2 with a 1.5-fold CO_2 stimulation. The specific activities of the Cl^- curve correspond to the basal specific activity of $Rv1625c_{204-443}$ wt. Figure 6.10C shows that the R376A single mutant protein is stimulated by CO_2 with a 2-fold CO_2 stimulation.

Figure 6.10: CO_2 dependence assay for $Rv1625c_{204-443}$ K269A, V366A, R376A mutant proteins. Mutant proteins were assayed at 37 °C for CO_2 sensitivity at pH 6.5 (50 mM MES) with 2 mM Mn^{2+} in the presence of 200 µM ATP and (20 mM Ci) for 40 minutes with a pre-incubation time of 30 minutes. Error bars are standard error of the mean (n=6) (*$p<0.05$) compared to all Cl^- values observed. A) 7.2 µM K296A, B) 1.2 µM V366A, C) 556 nM R376A, D) 18 µM K296A, R376A, E) 70 µM K296A, V366A, R376A.

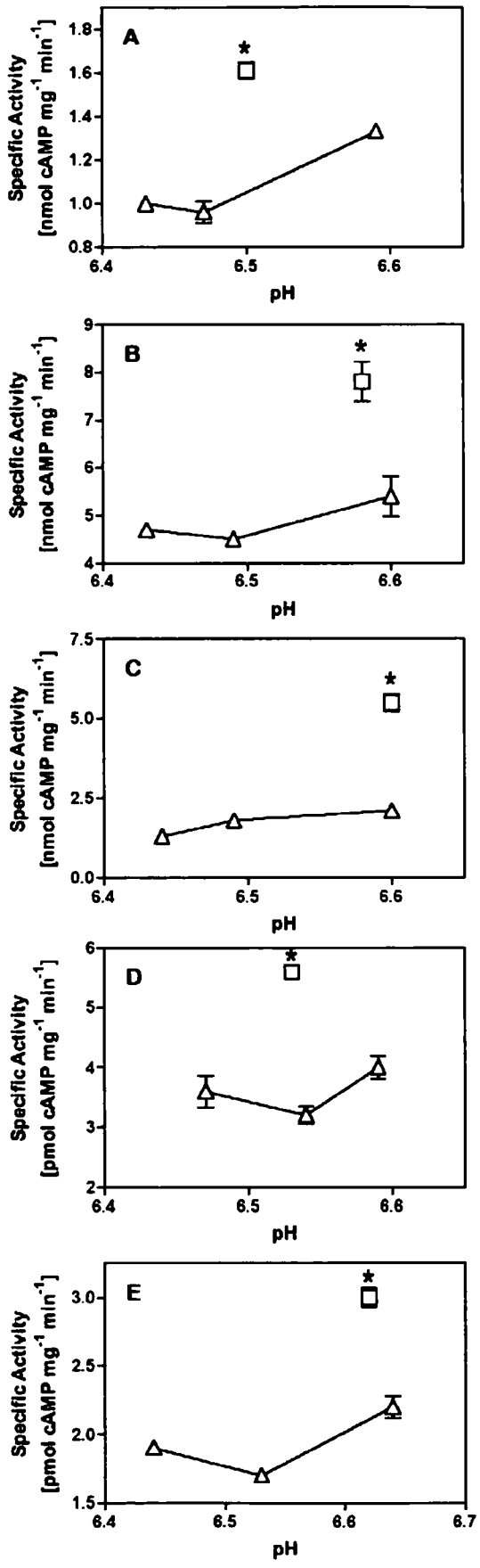
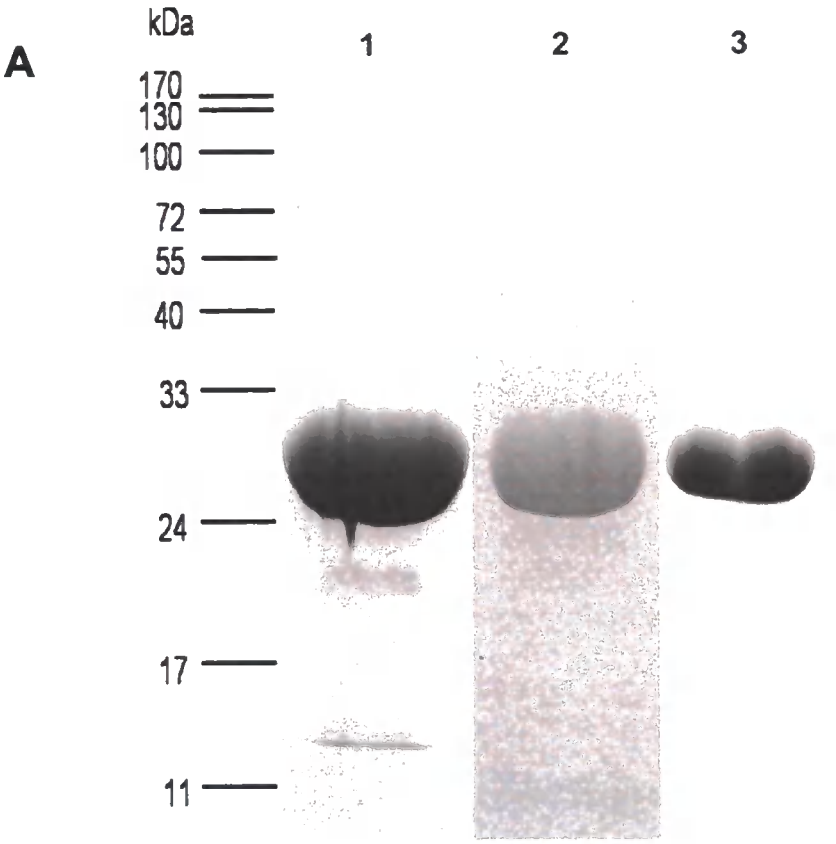
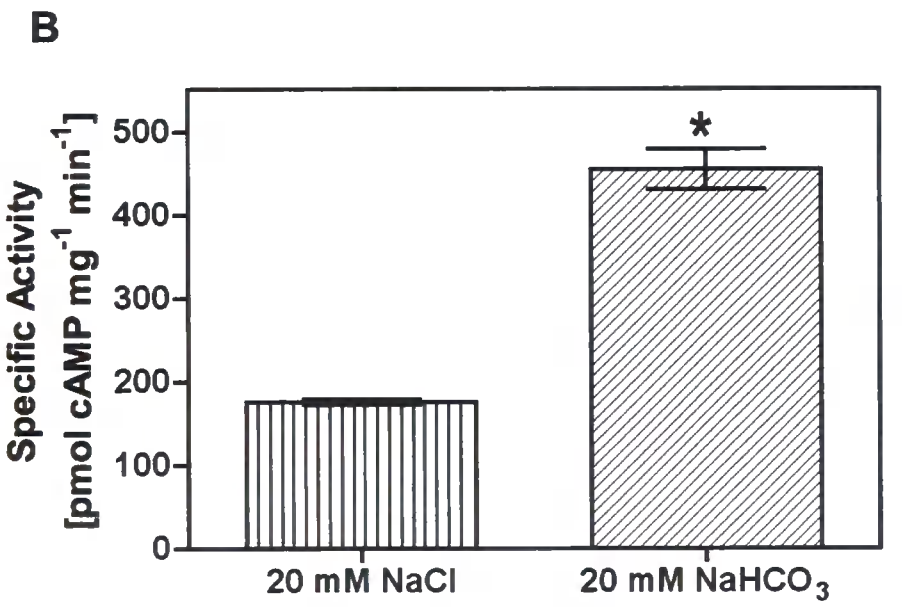

A
B
C
D
E
Specific Activity
[nmol cAMP mg-1 min-1]
[pmol cAMP mg-1 min-1]
pH
6.4
6.5
6.6
6.7

Figure 6.11: A) SDS-PAGE gel of $Rv1625c_{212-443}$ K296E, F363R, D365C mutant protein. 10 µL of 1) Ni-NTA purified, 2) Q-sepharose purified and 3) Superdex 200 purified mutant protein was applied to each lane with molecular weight standards on the left of each gel and the gel was stained with Coomassie Brilliant Blue G. B) CO_2 dependence assay for $Rv1625c_{204-443}$ K296E, F363R, D365C mutant protein. Mutant protein was assayed at 37 °C for CO_2 sensitivity at pH 6.5 (50 mM MES) with 2 mM Mn^{2+} in the presence of 200 µM ATP and (20 mM Ci) for 40 minutes with a pre-incubation time of 30 minutes. Error bars are standard error of the mean (n=6) (*$p<0.05$).


A
kDa
1
2
3
170
130
100
72
55
40
33
24
17
11


B
Specific Activity
[pmol cAMP mg-1 min-1]
500
400
300
200
100
0
*
20 mM NaCl
20 mM NaHCO3

The specific activities of the Cl^- curve and the CO_2 stimulated assay correspond with the results previously observed with R376A (see Section 5.2). Figure 6.10D shows that the double mutant protein, K296A, R376A is stimulated by CO_2 with a 2-fold CO_2 stimulation.

Figure 6.10E shows that the triple mutant protein, K296A, V366A, R376A is stimulated by CO_2 with a 1.5-fold CO_2 stimulation. In all the Cl^- control assays, the specific activity increased with an increase of pH from 6.5 to 6.6, an expected result as the optimal pH for $Rv1625c_{204-443}$ is pH 7.5 (Guo *et al.*, 2001). However, this pH-dependent increase in specific activity was less than half the increase in specific activity due to CO_2 sensitivity. The CO_2 stimulation is not a pH effect as the addition of CO_2 to the assay would decrease the pH value as CO_2 equilibrates to form a small amount of HCO_3^- and H^+. Therefore basal enzyme activity does increase with an increase in pH towards the optimal pH value, however the fold stimulation caused by the addition of CO_2 to the assay has a larger effect.

All the single, double and triple $Rv1625c_{204-443}$ mutant proteins were stimulated by CO_2 with CO_2 fold stimulations of between 1.5–2. This result demonstrates that none of these residues by themselves is the CO_2 binding site or that CO_2 is binding to more than one of these residues to stimulate $Rv1625c_{204-443}$. This result suggests that CO_2 and HCO_3^- have different binding sites on ACs, a logical presumption as the two Cis are different shapes and have different ionic charges. As sAC is stimulated by both CO_2 and HCO_3^-, sAC must have two different binding sites for the two Cis.

6.6 Crystallographic studies on $Rv1625c_{212-443}$

As all attempts to identify the CO_2 binding site have been unsuccessful, X-ray crystallography was employed to directly observe CO_2 bound to the active site of Rv1625c protein crystals. Rv1625c has previously been crystallised (Ketkar *et al.*, 2004), however due to stabilisation problems with the wild type protein, attributed to a mixture of monomeric and dimeric catalytic domains, the K296E, F363R, D365C triple mutant protein was the only stable mutant protein of Rv1625c that formed crystals as the triple mutant existed

predominately as a monomer (Shenoy *et al.*, 2003). In Rv1625c, K296 and D365 are the substrate defining residues and F363 is required for efficient AC activity. The triple mutant was initially generated to convert Rv1625c from an AC to a GC, however the mutations failed to convert Rv1625c to a GC and reduced its AC activity and altered its oligomeric state (Shenoy *et al.*, 2003).

$Rv1625c_{212-443}$ K296E, F363R, D365C was overexpressed in *E. coli* and purified by affinity chromatography, anion exchange chromatography and gel filtration chromatography to generate a recombinant protein. After each chromatographic step the protein was analysed for purity by SDS-PAGE.

The SDS-PAGE gel of the expressed and purified $Rv1625c_{212-443}$ K296E, F363R, D365C (Figure 6.11A) shows the protein bands from the three chromatographic steps. After the Ni-NTA column, the eluted protein (Figure 6.11A (1)) gives a large band at approximately 30 kDa of approximately 15 µg protein. The theoretical weight of $Rv1625c_{212-443}$ K296E, F363R, D365C is 26874 Da. The gel also shows that there are contaminating protein bands at approximately 20 kDa, 130 kDa and 170 kDa. The eluted protein after the Q-Sepharose column (Figure 6.11A (2)) shows a large band at approximately 30 kDa of approximately 10 µg protein. The gel shows that there are no other contaminating protein bands. The eluted protein after the final column, the Sephadex 100 column (Figure 6.11A (3)) shows a large band at approximately 30 kDa of approximately 8 µg protein. Again the gel shows that there are no other contaminating protein bands. Thus the mutant protein is > 99 % pure and is pure enough for crystallisation trials. The SDS-PAGE gels (Figure 6.11) also show a decrease in the quantity of mutant protein recovered from the different chromatographic steps. This loss of protein is due to non-specific interactions of the protein to the columns and also incomplete recovery of the eluted proteins from the eluant. The SDS-PAGE gels also show that the mutant protein was pure after the Q-Sepharose column, however the mutant protein was purified on the final column so as to replicate the published conditions and to change the protein buffer to 20 mM HEPES-NaOH, a buffer known to aid crystallisation of $Rv1625c_{212-443}$ and also to desalt the buffer from the high salt concentration used in the anion exchange chromatography (Ketkar *et al.*, 2004).

Figure 6.12: Crystals of Rv1625c$_{212-443}$ K296E, F363R, D365C mutant protein grown by the hanging-drop method. Crystal size is 150 x 30 x 30 Å.

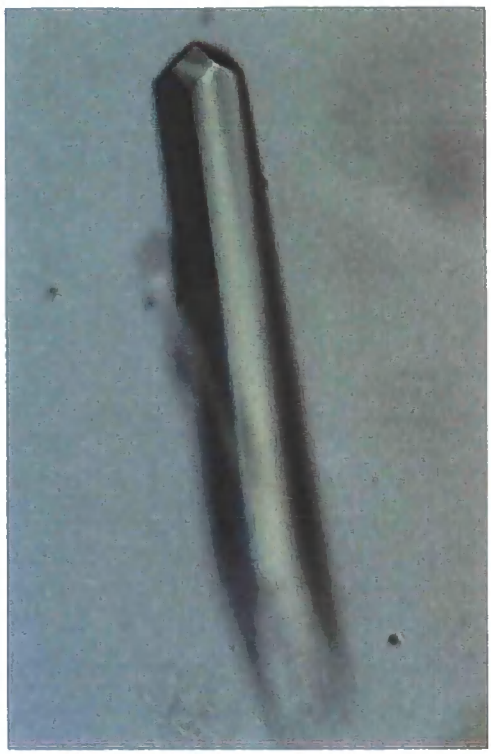
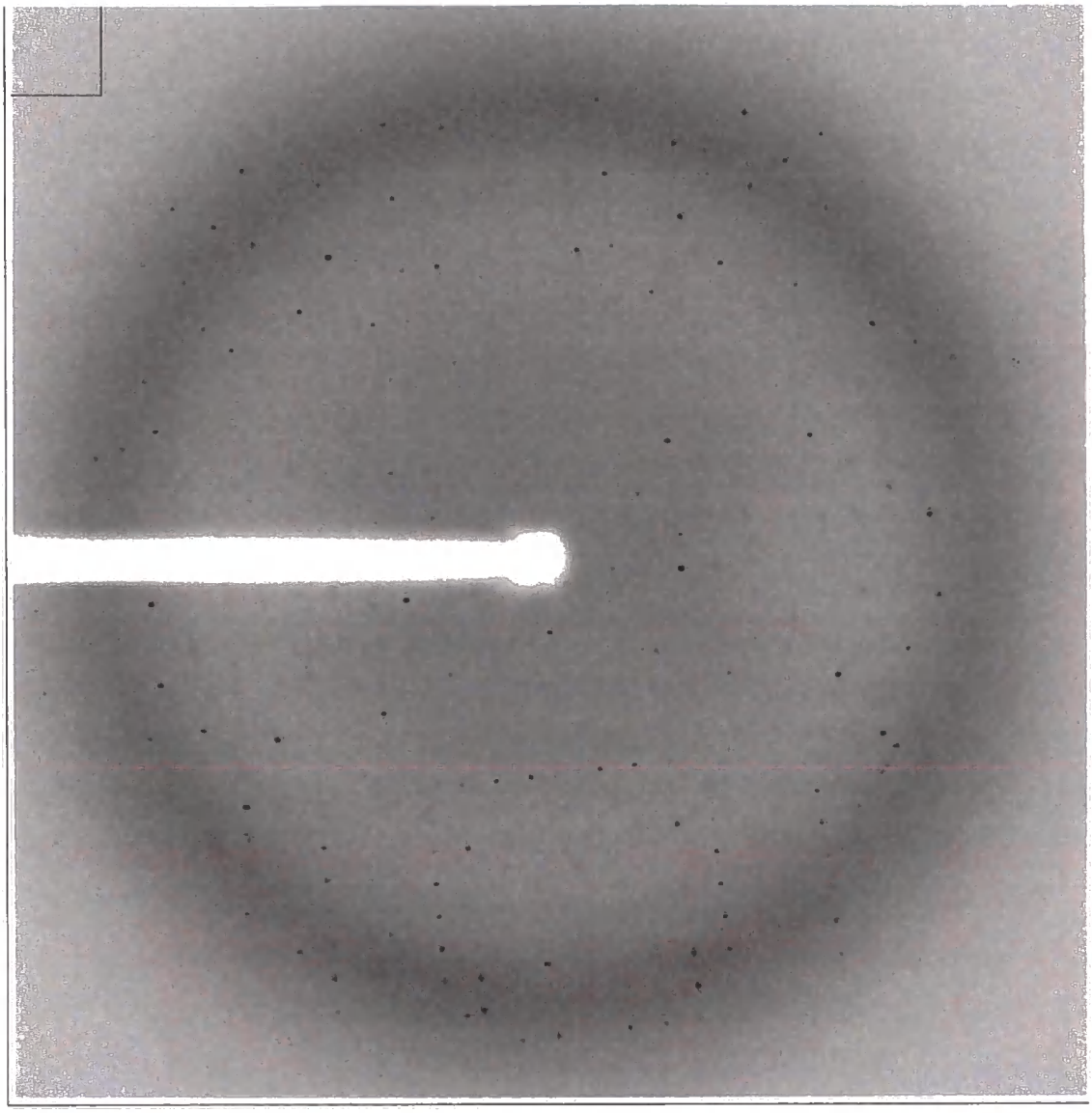


Figure 6.13: Diffraction pattern of $Rv1625c_{212-443}$ K296E, F363R, D365C mutant protein crystal exposed to CO_2, collected at the Swiss Light Source (Paul Scherrer Institut, Villigen, Switzerland).



For the crystallisation of Rv1625c$_{212-443}$ K296E, F363R, D365C in the presence of CO_2 the mutant protein was first shown to be activated by CO_2. If the mutant protein was not stimulated by CO_2, it may have been due to the CO_2 binding residue(s) being one of the mutated residues and thus the crystal structure would not show CO_2 bound to the protein. A CO_2 dependence assay on Rv1625c$_{212-443}$ K296E, F363R, D365C was performed (Figure 6.11B).

Figure 6.11B shows that Rv1625c$_{212-443}$ K296E, F363R, D365C was stimulated by CO_2 with a 2.5 fold stimulation. Therefore the mutant protein does bind CO_2 and thus the protein crystals incubated with CO_2 will hopefully produce a diffraction data set that when solved will display CO_2 bound to residue(s) of Rv1625c$_{212-443}$ K296E, F363R, D365C.

With the help of Dr K. Betz (Department of Chemistry, Durham University), crystallisation trials were set up with Rv1625c$_{212-443}$ K296E, F363R, D365C as described in Section 3.3.5 and crystals with a typical length of 20-30 µm grown under the conditions of 7-8 % PEG 8000 and 6 % glycerol were selected for CO_2 exposure (Figure 6.12) as the largest crystals grew under these conditions. Optimal sized crystals were removed from the crystallisation trials, exposed to CO_2 in CO_2 chamber (see Section 3.3.6) and cryogenically frozen for transportation to a synchrotron.

Both cryogenically frozen native Rv1625c$_{212-443}$ K296E, F363R, D365C protein crystals and CO_2 exposed protein crystals were subjected to X-ray diffraction studies and data sets were collected (Figure 6.13). The Rv1625c$_{212-443}$ K296E, F363R, D365C protein crystallised in space group $P4_1$, with unit-cell parameters $a = b = 77.3$, $c = 110.3$ Å, which are similar parameters to the published data (Ketkar *et al.*, 2004) but with a resolution of 2.8 Å, which is better than the published data of 3.4 Å. However, the structure of the CO_2 bound protein had a high R-factor, which measures the agreement between the crystallographic model and the experimental X-ray diffraction data (Morris *et al.*, 1992), of >35 %, a more acceptable value is <25 %.

The high R-factor causes the phase error to be high and thus the electron density is poor. This result was observed for a second native protein crystal so is a real effect caused by the crystals. With a resolution of 2.8 Å, CO_2 would not be seen in the structure of the CO_2 exposed protein crystals.

6.7 Conclusion

In this chapter numerous experiments were conducted to identify whether CO_2 bound to $Rv1625c_{204-443}$ through a single amino acid, with the most likely residue being K296, or via other residues that contain an amine or amide sidechain group, such as Lys, His and Arg. However, CO_2 dose dependence assays of the single mutants showed that all the single mutant proteins assayed were stimulated by CO_2 and were therefore not the CO_2 binding site. Thus experiments into multiple amino acids being the potential CO_2 binding site were investigated. The multiple residues were identified through aligning the HCO_3^- binding residues of sAC with $Rv1625c_{204-443}$, however, CO_2 dependence assays on the multiple residue mutant proteins showed that they were all stimulated by CO_2. Finally Rv1625c was investigated by growing protein crystals, exposing the crystals to CO_2 and collecting a diffraction data set, however, the resolution of the diffracted crystal was too low to observe CO_2 in the structure.

Unfortunately the CO_2 binding site of Rv1625c was not identified, however, this work has identified a Class IIIa AC as being activated by CO_2, identified that CO_2 binds solely to the apoprotein and has narrowed down the potential binding sites of CO_2 to $Rv1625c_{204-443}$. A range of techniques for determining the mode of CO_2 binding to ACs were identified but were not effective. It can only be a matter of time before the CO_2 binding site on Rv1625c is discovered.

7 Discussion

7.1 Final discussion

Rv1625c is an excellent model to use for studying AC biochemistry for a number of reasons.

- the Rv1625c protein consists of six putative transmembrane helices and a single Class IIIa AC catalytic domain (Guo *et al.*, 2001). This corresponds to exactly one half of a mammalian G protein-regulated AC. In addition, the six key catalytic residues distributed between the two catalytic domains of mammalian ACs are also found in the active site of Rv1625c (Guo *et al.*, 2001). Rv1625c is the closest bacterial progenitor of mammalian ACs identified to date, therefore, it is likely that Rv1625c responds to Ci stimulation in a similar fashion to mammalian ACs.

- Rv1625c can be expressed using a bacterial expression system *in vivo* as the protein does not require post-translational modification for activity. Using a bacterial expression system, compared to a mammalian expression system, enables large amounts of protein to be expressed and also simplifies subsequent purification steps. A greater, purer yield is therefore, obtained with fewer purification steps. With protein expressed quickly, experiments can be conducted faster and results obtained in a shorter time than if a mammalian AC was used. Expressing mammalian ACs using a bacterial expression system does not post-translationally modify the protein and yield is lower. Using a mammalian expression system to express mammalian ACs would be preferable, however, the maintenance of mammalian cells and the subsequent purification of the protein is more problematic.

- The conditions required for an AC assay using Rv1625c are simpler than a mammalian AC assay. Rv1625c does not require any activators for basal AC activity, whereas mammalian ACs require the plant diterpene forskolin and/or the monomeric $G_s\alpha$ subunit to detect significant basal AC activity (data not shown). Thus, $G_s\alpha$ subunits have to be

expressed, purified, and activated. A $G_s\alpha$ dose dependence curve generated, a forskolin dose dependence curve generated and/or a synergistic forskolin/$G_s\alpha$ dose dependence curve must be generated before commencing on mammalian AC assays. As mammalian ACs are heterodimers, for AC activity two subunits are required for the formation of a catalytic site, whereas Rv1625c is a homodimer and thus dimerises with two single subunits to form two catalytic sites.

Although Rv1625c is a useful model for G protein-regulated mammalian ACs, there is, however, a limitation to studying CO_2 stimulation of Rv1625c solely as a model for CO_2 stimulation in mammalian ACs. CO_2 stimulation of Rv1625c only suggests that G protein-regulated mammalian ACs are also stimulated by CO_2 since Rv1625c is a bacterial AC and mammalian ACs may respond differently to CO_2 since there are numerous activators specific to mammalian ACs that do not activate Rv1625c. To prove that mammalian ACs are CO_2 responsive, assays using mammalian ACs must be conducted. Work done by Townsend *et al.* has demonstrated that G protein-regulated mammalian ACs are stimulated by CO_2 (Townsend *et al.*, 2009), thus, supporting the use of Rv1625c as a model for CO_2 stimulation in G protein-regulated mammalian ACs.

The *in vitro* study of CO_2 stimulation of Rv1625c has numerous advantages over *in vivo* experiments.

- *In vitro* AC assays allow for closely controlled experimental conditions on an isolated catalytic domain of the enzyme. Therefore, any changes in AC activity can be attributed to the change in experimental conditions. Kinetic parameters of AC stimulation can be calculated to understand the mechanism of the stimulation.

- Expression and purification of recombinant AC protein is straightforward and yield is high, allowing for numerous *in vitro* AC assays to be conducted. By contrast, *M. tuberculosis* grows slowly and

as its genome contains 15 putative AC domains, thus, it is difficult to associate any change in cAMP concentration with Rv1625c.

The disadvantage of *in vitro* assays is that they do not explain how the enzyme functions *in vivo*. AC proteins possess regulatory domains, such as the HAMP domain of Rv1319c of *M. tuberculosis* that decreases specific activity by 70 % (Linder 2004), that can modulate the activity of the AC domain and must therefore, be considered. There are also other activators and inhibitors in the *in vivo* signalling cascade, such as PDEs that degrade cAMP, that also affect AC activity and are unaccounted for with *in vitro* assays. In *M. tuberculosis*, there is the added problem of 14 other ACs that may interact and affect the response of CO_2 activated Rv1625c. This interaction of other *M. tuberculosis* ACs in the absence of Rv1625c is demonstrated clearly by work done to identify Rv1625c *in vivo* pathogenecity by infecting a mouse with a *rv1625c* knockout. The *rv1625c* knockout still caused pathogenesis, likely due to redundancy of the 14 other ACs (Guo *et al.*, 2005).

There are numerous difficulties with experiments that require CO_2.

- CO_2 saturation of water causes the acidification of the water due to formation of carbonic acid. Therefore the addition of CO_2 saturated water to the AC assays requires a high concentration of buffer and close regulation of the pH of the buffer for the duration of the experiment. Control assays must have an identical pH value because at low pH value, a pH change of as little as 0.1 could mask an AC stimulation due to AC activity increasing as pH increases. At pH values approaching 7.5, significant amounts of CO_2 in the assay solution are converted to HCO_3^- (so as to maintain the Ci equilibrium) and thus, there is an overall lower concentration of CO_2 in the assay than in the stock CO_2 saturated water. Thus there is a lower concentration of CO_2 in the assay and a diminished CO_2 stimulation is observed at pH 7.5.

- CO_2 degasses readily from a saturated CO_2 solution due to a steep concentration gradient between the solution and the atmosphere.

- CO_2 saturated water gives a 30 mM CO_2 solution, which limits the addition of CO_2 to the assay.

AC assays were an integral part of the investigation into CO_2 stimulation of Rv1625c. The use of [α-^{32}P] ATP as a substrate allowed for the measurement, via a scintillation counter, of the amount of [2,8-^{3}H] cAMP formed during the assay and taking into account the concentration of enzyme used, the AC activity was calculated. There are a few considerations made for this AC activity calculation.

- Individual Dowex and aluminium oxide columns, which separate the [2,8-^{3}H] cAMP from [α-^{32}P] ATP, are used for each replicate sample, thus there may be discrepancies between replicates. This discrepancy can be reduced by increasing the number of replicates per assay and repeating assays while randomising the columns used.

- There is a maximum amount of ATP that can be added to the assays. If the concentration of ATP used is higher than 100 mM, the ratio of ATP:[α-^{32}P] ATP is too great and the amount of [2,8-^{3}H] cAMP produced decreases as more cAMP is produced, which gives a decrease in the sensitivity of the AC assay and erroneously suggests that there is a decrease in AC activity.

- The kinetic calculations of the AC assay assume that all measurements of specific activity are of the initial rate of reaction. To maintain this assumption, the AC assay conditions must be controlled so that the conversion of substrate is kept below 10 %, otherwise there will be a significant drop in substrate concentration and specific activity will no longer be a measure of the initial rate of reaction. Another reason for maintaining the substrate conversion at below 10 % is that high PPi concentrations can inhibit AC activity.

The 'Investigation into *in vitro* and *in vivo* Ci stimulation of $Rv1625c_{204-443}$' chapter showed that $Rv1625c_{204-443}$ was stimulated by CO_2 at pH 6.5 and that CO_2 stimulated $Rv1625c_{204-443}$ through increasing k_{cat}. $Rv1625c_{204-443}$ was also stimulated by CO_2 *in vivo*. $Rv1625c_{204-443}$ is the first Class IIIa AC shown to be stimulated by CO_2, although numerous other prokaryotic ACs have been previously shown to be CO_2 responsive. The Class IIIb ACs Slr1991 from *Synechocystis* PCC 6803 and CyaB1 from *Anabaena* PCC 7120 were both proven to be regulated by CO_2 through an increase in k_{cat} (Hammer *et al.*, 2006). As ACs from both Class IIIa and Class IIIb ACs are CO_2 responsive, it is possible that all Class III ACs may be responsive to CO_2. In the literature, Class IIIb ACs are defined as being HCO_3^- responsive (Linder and Schultz, 2003), however, Hammer *et al* suggest that the wrong Ci form has been previously identified as the activator, as HCO_3^- is in equilibrium with CO_2. This anomaly is likely due to performing ACs assays at pH 7.5, where the Ci is 93.9 % in the form HCO_3^- and 5.9 % in the form CO_2. Indeed, experiments conducted on the Class IIIb ACs mammalian sAC and CyaC from *S. platensis* suggested that HCO_3^- stimulated these enzymes (Chen *et al.*, 2000), however, work done by Townsend *et al* demonstrated that mammalian sAC is regulated not only by HCO_3^- but by CO_2 as well (Townsend *et al.*, 2009). The Class IIIa G-protein regulated mammalian AC, $7C_1.2C_2$ was also demonstrated to be CO_2 activated (Townsend *et al.*, 2009). It is likely that the CO_2 binding mechanism is similar to the CO_2 binding mechanism of Rv1625c. Work done on the crystallisation of CyaC in the presence of HCO_3^- did not produce crystals and soaking CyaC crystals in HCO_3^- dissolved the crystals (Steegborn *et al.*, 2005). Flash soaking and freezing of the crystal demonstrated a conformational change, but as HCO_3^- was not identified in the structure, the conformational change may be induced by CO_2 instead.

In support of the *in vivo* CO_2 stimulation of Rv1625c, another Class IIIa AC, mammalian $7C_1.2C_2$ has been demonstrated to be regulated *in vivo* (Townsend *et al.*, 2009). $7C_1.2C_2$ in HEK 293T cells were activated by 5 % CO_2 (compared to atmospheric CO_2) in the presence of isoproterenol, a β-adrenergic receptor agonist.

The 'Elucidation of AC CO_2 binding conditions' chapter showed that CO_2 increased the affinity of $Rv1625c_{204-443}$ for Mn^{2+} and increased the co-

operativity of the two binding sites for recruiting metal ions. This finding agreed with Steegborn *et al.*, who suggested that Ci stimulates AC by recruiting a second metal ion to the active site (Steegborn *et al.*, 2005). The authors also suggested that Ci stimulation of CyaC causes a marked active site closure, however, CD spectroscopy of $Rv1625c_{204-443}$ (Figure 5.5) did not identify any change in secondary structure due to the addition of CO_2. In the 'Elucidation of AC CO_2 binding conditions' chapter, spectroscopic techniques such as FTIR spectroscopy and HCO_3^- dependent luminescence probe spectroscopy were explored, but were unable to identify the conditions required for CO_2 binding. Binding assays with $Rv1625c_{204-443}$ using $^{14}CO_2$ (which had previously been used to identify CO_2 binding to Rubisco (Lorimer, 1979)) identified that CO_2 bound to apoprotein and no other co-factors, such as Mg^{2+} or Mn^{2+}, were required. AC assays can not be used to identify whether metal ions are required for CO_2 binding as if the metal ions are removed from the assay, no catalysis can occur. As CO_2 is not binding directly with the metal ion in the active site but aiding metal recruitment, CO_2 is likely bound to a single conserved residue or a conserved site in the active site of Rv1625c. A likely candidate for a CO_2 binding site is the conserved substrate defining Lys (K296 in Rv1625c). Cann *et al.* demonstrated that a point mutation of the substrate defining Lys of CyaB1 from *Anabaena* PCC 7120 (K646) ablated Ci activation (Cann *et al.*, 2003). The authors suggest that Ci stimulation (in the form of HCO_3^-) occurs through the formation of a hydrogen bond with K646. However, further work from the same lab has identified CO_2 and not HCO_3^- as the activator of CyaB1 (Hammer *et al.*, 2006) and thus the mechanism of activation is likely to be CO_2 forming a carbamate with the amino group of the substrate defining Lys, as observed with Rubisco (Lorimer *et al.*, 1976).

The 'Investigation into potential CO_2 binding sites' chapter showed that CO_2 was not binding solely to the substrate defining Lys in Rv1625c, K296, nor was it solely binding to 12 other amino group containing residues in the active site. The formation of a carbamate is the most likely interaction that CO_2 would make to an amino acid, as observed with Rubisco (Lorimer *et al.*, 1976). However, CO_2 may be forming a carbamate with a single residue but also require stabilisation via other residues in a binding pocket. This would explain why no ablation of CO_2 activity was seen in the CO_2 dose dependence assays

of single mutant proteins, as other CO_2 binding residues in the binding pocket are able to 'bind' CO_2 and cause an activation. With the difficulty in measuring $E.C._{50}$ values of the mutant proteins, changes in the affinity of CO_2 to the mutant proteins can not be calculated.

sAC can bind both HCO_3^- and CO_2, however, it is unknown whether the Ci species bind at the same binding site or whether there are two distinct sites. Creating the corresponding multiple residue binding sites in Rv1625c of the HCO_3^- binding residues of mammalian sAC, which were identified from the crystal structure of sAC and HCO_3^- (Saalau-Bethell *et al.*, 2007), did not ablate CO_2 sensitivity and so are unlikely to be the CO_2 binding site. This could be due to two main reasons. First, the identified coordinates of HCO_3^- from the sAC crystal structure may not fit into the binding pocket of sAC correctly. This was demonstrated by the need to rotate the coordination of the sidechain of R176 in the sAC crystal structure to represent its normal orientation and create a binding pocket in the structure. If the coordinates of the binding residues from the sAC crystal structure were incorrect, the corresponding residues in Rv1625c would also be incorrect and the CO_2 binding site would not be identified. Second, as sAC can bind both HCO_3^- and CO_2, perhaps there are two different binding sites for the two Ci species for sAC and as Rv1625c does not bind HCO_3^-, mutation of the residues that correspond to the HCO_3^- site would have no affect on CO_2 stimulation of Rv1625c.

Crystals of Rv1625c were grown and exposed to CO_2. The diffraction data that were collected gave a higher resolution model than previously published, but had a high R-factor, which increased the phase error and made the electron model density poor. However, even if the Rv1625c structure was solved the CO_2 binding site would not be identified since CO_2 would not be visible in the structure. This is because CO_2, with an approximate length of 234 pm (2.34 Å), is smaller than the resolution obtained from the Rv1625c crystal (2.8 Å) and therefore would not be resolved. As expressing, purifying and growing Rv1625c crystals took over 1 month and the diffraction data could only be collected at a synchrotron, there was only time for one attempt at solving the crystal structure of Rv1625c with CO_2 bound.

7.2 Future Work

The initial future work is to answer the question, what are the CO_2 binding residues of Rv1625c? The crystal structure of Rv1625c exposed to CO_2 is required and so expression, purification and growth of more $Rv1625c_{212-443}$ K296E, F363R, D365C mutant protein are required. The exposure of the Rv1625c crystal to CO_2 should be repeated and another data set collected at a synchrotron facility. To improve the resolution of the diffraction data collected, larger crystals are required. The concentration of protein solution added to the crystallisation screens may be decreased, as the concentrations used (30 mg/mL) formed numerous small crystals, whereas a lower concentration of protein solution (e.g. 15 mg/mL) may form fewer crystals, but, incubated for a longer time period, the crystals may be larger. An ideal time period to initially test is 2 months of growing. The exposure of CO_2 can also be modified to find the optimal time period for exposure, too short a time and CO_2 would not be able to bind, too long a time and the crystal would dry out and not diffract correctly. Protein crystals of other CO_2 activated AC proteins should also be grown, such as Slr1991 from *Synechocystis* PCC 6803 and CyaB1 from *Anabaena* PCC 7120 to confirm that the CO_2 binding site is conserved between ACs. As the growing conditions of the new ACs are not known, it may take time to identify the correct conditions for optimal crystal growth. Mutant proteins may also need to be considered as they were with Rv1625c, because native AC protein may dimerise and inhibit crystallisation. If mutant proteins are required due to problems with crystallisation of the wt proteins, the Slr1991 and CyaB1 corresponding triple mutants of $Rv1625c_{212-443}$ K296E, F363R, D365C would be obvious initial tests, however, other mutations may need to be investigated in order to inhibit dimerisation. sAC would also be an interesting protein to crystallise and expose to CO_2, since the HCO_3^- binding site has been previously identified by crystallisation (Saalau-Bethell *et al.*, 2007), it would be interesting to observe the CO_2 binding site.

To back up the theory that all Class III ACs are CO_2 responsive, Class IIIc and Class IIId ACs should be assayed for CO_2 responsiveness. Rv1264 (Linder *et al.*, 2002) and Rv1900c (Sinha *et al.*, 2005) are both Class IIIc ACs from *M. tuberculosis* that have previously been expressed as recombinant

proteins and assayed for AC activity. The AC from *Candida albicans* is a Class IIId AC that has previously been expressed as a recombinant protein and assayed for AC activity (Klengel *et al.*, 2005). The protein constructs can be obtained and the AC proteins expressed and purified. Conducting CO_2 dose dependence assays on the ACs will determine whether they are responsive or non-responsive to CO_2.

Once all the subdivisions of Class III ACs have been tested for CO_2 responsive, the next step would be to examine the CO_2 responsiveness of the other AC classes. Are all AC classes CO_2 responsive? As the other classes of ACs were formed by convergent evolution, they all have unique catalytic domains and it is therefore, unlikely that CO_2 responsiveness is conserved within the other classes of ACs. The Class I AC Cya from *E. coli* (Linder, 2008), the Class II AC EF from *B. anthracis* (Guo *et al.*, 2004) and the Class IV AC AC2 from *A. hydrophilia* (Sismeiro *et al.*, 1998) have previously been expressed as a recombinant proteins, purified and assayed for AC activity. The Class V AC from *P. ruminicola* (Cotta *et al.*, 1998) and the Class VI AC from *R. etli* (Tellez-Sosa *et al.*, 2002) have not been expressed as recombinant proteins and thus molecular cloning to insert the catalytic domains of the AC genes into plasmids would be required before assaying for CO_2 responsiveness.

With Rv1625c responsiveness to CO_2 characterised *in vitro*, another area of future work would be to study *in vivo* CO_2 responsiveness of *M. tuberculosis* and any potential links with virulence via the halting of phagosomal acidification (Lowrie *et al.*, 1979), and Rv1625c. Is Rv1625c the CO_2 sensor for *M. tuberculosis*? As the *M. tuberculosis* genome encodes for 15 putative AC genes, it may be difficult to assign any change in cAMP production to Rv1625c and thus, prove that Rv1625c is the CO_2 sensor. A *rv1625c* knockout infected in mice was still pathogenic (Guo *et al.*, 2005) and the result is likely due to the redundancy provided by the other 14 ACs. *M. tuberculosis* could be grown both in ambient and high CO_2 levels and the concentration of cAMP formed measured by ELISA. It would be hypothesised that as long as high CO_2 levels do not inhibit *M. tuberculosis* growth, cAMP levels will be higher in the presence of high CO_2 than compared to ambient CO_2 levels. This result could be compared to a *rv1625c* knockout strain grown

both in ambient and high CO_2 levels. However there may be an identical increase in cAMP due to CO_2 stimulation of the other ACs. Therefore, as CO_2 may play an important role in pathogenesis, the CO_2 sensor in *M. tuberculosis* may not be a single AC, but may be a combination of all ACs.

In summary, a Class IIIa AC from *M. tuberculosis*, Rv1625c, has been demonstrated to be stimulated by CO_2 both *in vitro* and *in vivo*. The conditions for CO_2 binding to Rv1625c were investigated using numerous techniques and it was identified that CO_2 bound solely to the Rv1625c apoprotein, without the requirement of any cofactors or metal ions. The identification of the CO_2 binding site of Rv1625c was attempted using numerous techniques, however the binding site was not elucidated. Further work is required to identify the CO_2 binding site.

8 References

Reference

Abdel Motaal, A., I. Tews, J. E. Schultz and J. U. Linder (2006). "Fatty acid regulation of adenylyl cyclase Rv2212 from Mycobacterium tuberculosis H37Rv." Febs J **273**(18): 4219-4228.

Anes, E., M. P. Kuhnel, E. Bos, J. Moniz-Pereira, A. Habermann and G. Griffiths (2003). "Selected lipids activate phagosome actin assembly and maturation resulting in killing of pathogenic mycobacteria." Nat Cell Biol **5**(9): 793-802.

Aravind, L. and C. P. Ponting (1999). "The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins." FEMS Microbiol Lett **176**(1): 111-116.

Bahn, Y. S. and F. A. Muhlschlegel (2006). "CO_2 sensing in fungi and beyond." Curr Opin Microbiol **9**(6): 572-578.

Baillie, L. and T. D. Read (2001). "*Bacillus anthracis*, a bug with attitude!" Curr Opin Microbiol **4**(1): 78-81.

Bakalyar, H. A. and R. R. Reed (1990). "Identification of a specialized adenylyl cyclase that may mediate odorant detection." Science **250**(4986): 1403-1406.

Barzu, O. and A. Danchin (1994). "Adenylyl cyclases: a heterogeneous class of ATP-utilizing enzymes." Prog Nucleic Acid Res Mol Biol **49**: 241-283.

Bertini, I., C. Del Bianco, I. Gelis, N. Katsaros, C. Luchinat, G. Parigi, M. Peana, A. Provenzani and M. A. Zoroddu (2004). "Experimentally exploring the conformational space sampled by domain reorientation in calmodulin." Proc Natl Acad Sci U S A **101**(18): 6841-6846.

Bhat, S. V., B. S. Bajwa, H. Dornauer, N. J. D. Souza and H. W. Fehlhaber (1977). "Structures and Stereochemistry of New Labdane Diterpenoids from *Coleus Forskohlii Briq*." Tetrahedron Letters(19): 1669-1672.

Bhaya, D., N. R. Bianco, D. Bryant and A. Grossman (2000). "Type IV pilus biogenesis and motility in the cyanobacterium *Synechocystis* sp. PCC6803." Mol Microbiol **37**(4): 941-951.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Anal Biochem **72**: 248-254.

Braun, T. (1990). "Inhibition of the soluble form of testis adenylate cyclase by catechol estrogens and other catechols." Proc Soc Exp Biol Med **194**(1): 58-63.

Braun, T. and R. F. Dods (1975). "Development of a Mn^{2+}-sensitive, "soluble" adenylate cyclase in rat testis." Proc Natl Acad Sci U S A **72**(3): 1097-1101.

Braun, T., H. Frank, R. Dods and S. Sepsenwol (1977). "Mn^{2+}-sensitive, soluble adenylate cyclase in rat testis. Differentiation from other testicular nucleotide cyclases." Biochim Biophys Acta **481**(1): 227-235.

Bretonniere, Y., M. J. Cann, D. Parker and R. Slater (2004). "Design, synthesis and evaluation of ratiometric probes for hydrogencarbonate based on europium emission." Org Biomol Chem **2**(11): 1624-1632.

Bretscher, A. J., K. E. Busch and M. de Bono (2008). "A carbon dioxide avoidance behavior is integrated with responses to ambient oxygen and

food in *Caenorhabditis elegans*." Proc Natl Acad Sci U S A **105**(23): 8044-8049.

Buck, J., M. L. Sinclair, L. Schapal, M. J. Cann and L. R. Levin (1999). "Cytosolic adenylyl cyclase defines a unique signaling molecule in mammals." Proc Natl Acad Sci U S A **96**(1): 79-84.

Cali, J. J., J. C. Zwaagstra, N. Mons, D. M. Cooper and J. Krupinski (1994). "Type VIII adenylyl cyclase. A Ca^{2+}/calmodulin-stimulated enzyme expressed in discrete regions of rat brain." J Biol Chem **269**(16): 12190-12195.

Cann, M. J. (2007). "Sodium regulation of GAF domain function." Biochem Soc Trans **35**(Pt 5): 1032-1034.

Cann, M. J., A. Hammer, J. Zhou and T. Kanacher (2003). "A defined subset of adenylyl cyclases is regulated by bicarbonate ion." J Biol Chem **278**(37): 35033-35038.

Carter, P. B. and F. M. Collins (1974). "Growth of typhoid and paratyphoid bacilli in intravenously infected mice." Infect Immun **10**(4): 816-822.

Castro, L. I., C. Hermsen, J. E. Schultz and J. U. Linder (2005). "Adenylyl cyclase Rv0386 from *Mycobacterium tuberculosis* H37Rv uses a novel mode for substrate selection." Febs J **272**(12): 3085-3092.

Chabardes, D., D. Firsov, L. Aarab, A. Clabecq, A. C. Bellanger, S. Siaume-Perez and J. M. Elalouf (1996). "Localization of mRNAs encoding Ca^{2+}-inhibitable adenylyl cyclases along the renal tubule. Functional consequences for regulation of the cAMP content." J Biol Chem **271**(32): 19264-19271.

Chaloupka, J. A., S. A. Bullock, V. Iourgenko, L. R. Levin and J. Buck (2006). "Autoinhibitory regulation of soluble adenylyl cyclase." Mol Reprod Dev **73**(3): 361-368.

Chen, Y., M. J. Cann, T. N. Litvin, V. Iourgenko, M. L. Sinclair, L. R. Levin and J. Buck (2000). "Soluble adenylyl cyclase as an evolutionarily conserved bicarbonate sensor." Science **289**(5479): 625-628.

Chen, Y., A. Harry, J. Li, M. J. Smit, X. Bai, R. Magnusson, J. P. Pieroni, G. Weng and R. Iyengar (1997). "Adenylyl cyclase 6 is selectively regulated by protein kinase A phosphorylation in a region involved in Galphas stimulation." Proc Natl Acad Sci U S A **94**(25): 14100-14104.

Choi, E. J., Z. Xia and D. R. Storm (1992). "Stimulation of the type III olfactory adenylyl cyclase by calcium and calmodulin." Biochemistry **31**(28): 6492-6498.

Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead and B. G. Barrell (1998). "Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence." Nature **393**(6685): 537-544.

Collaborative Computational Project, N. (1994). "The CCP4 suite: Programs for Protein Crystallography." Acta Crystallog. sect. D **50**: 760-763.

Cooper, T. G. and D. Filmer (1969). "The active species of "CO_2" utilized by ribulose diphosphate carboxylase." J Biol Chem **244**(3): 1081-1083.

Cornelis, G. R., A. Boland, A. P. Boyd, C. Geuijen, M. Iriarte, C. Neyt, M. P. Sory and I. Stainier (1998). "The virulence plasmid of *Yersinia*, an antihost genome." Microbiol Mol Biol Rev **62**(4): 1315-1352.

Cotta, M. A., T. R. Whitehead and M. B. Wheeler (1998). "Identification of a novel adenylate cyclase in the ruminal anaerobe, *Prevotella ruminicola* D31d." FEMS Microbiol Lett **164**(2): 257-260.

Curtiss, R., 3rd and S. M. Kelly (1987). "*Salmonella typhimurium* deletion mutants lacking adenylate cyclase and cyclic AMP receptor protein are avirulent and immunogenic." Infect Immun **55**(12): 3035-3043.

Davis, B. G., X. Shang, G. DeSantis, R. R. Bott and J. B. Jones (1999). "The controlled introduction of multiple negative charge at single amino acid sites in subtilisin *Bacillus lentus*." Bioorg Med Chem **7**(11): 2293-2301.

Davis, B. J. (1964). "Disc Electrophoresis. Ii. Method and Application to Human Serum Proteins." Ann N Y Acad Sci **121**: 404-427.

de Rooij, J., F. J. Zwartkruis, M. H. Verheijen, R. H. Cool, S. M. Nijman, A. Wittinghofer and J. L. Bos (1998). "Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP." Nature **396**(6710): 474-477.

Dessauer, C. W. and A. G. Gilman (1997). "The catalytic mechanism of mammalian adenylyl cyclase. Equilibrium binding and kinetic analysis of P-site inhibition." J Biol Chem **272**(44): 27787-27795.

Dessauer, C. W., J. J. Tesmer, S. R. Sprang and A. G. Gilman (1998). "Identification of a Gialpha binding site on type V adenylyl cyclase." J Biol Chem **273**(40): 25831-25839.

Dessauer, C. W., J. J. Tesmer, S. R. Sprang and A. G. Gilman (1999). "The interactions of adenylate cyclases with P-site inhibitors." Trends Pharmacol Sci **20**(5): 205-210.

Dittrich, D., C. Keller, S. Ehlers, J. E. Schultz and P. Sander (2006). "Characterization of a *Mycobacterium tuberculosis* mutant deficient in pH-sensing adenylate cyclase Rv1264." Int J Med Microbiol **296**(8): 563-566.

Edelhoff, S., E. C. Villacres, D. R. Storm and C. M. Disteche (1995). "Mapping of adenylyl cyclase genes type I, II, III, IV, V, and VI in mouse." Mamm Genome **6**(2): 111-113.

Elliott, J. L., J. Mogridge and R. J. Collier (2000). "A quantitative study of the interactions of *Bacillus anthracis* edema factor and lethal factor with activated protective antigen." Biochemistry **39**(22): 6706-6713.

Emmer, M., B. deCrombrugghe, I. Pastan and R. Perlman (1970). "Cyclic AMP receptor protein of *E. coli:* its role in the synthesis of inducible enzymes." Proc Natl Acad Sci U S A **66**(2): 480-487.

Esposito, G., B. S. Jaiswal, F. Xie, M. A. Krajnc-Franken, T. J. Robben, A. M. Strik, C. Kuil, R. L. Philipsen, M. van Duin, M. Conti and J. A. Gossen (2004). "Mice deficient for soluble adenylyl cyclase are infertile because of a severe sperm-motility defect." Proc Natl Acad Sci U S A **101**(9): 2993-2998.

Fagan, K. A., R. Mahey and D. M. Cooper (1996). "Functional co-localization of transfected Ca($^{2+}$)-stimulable adenylyl cyclases with capacitative Ca^{2+} entry sites." J Biol Chem **271**(21): 12438-12444.

Falkowski, P. G. (1997). "Photosynthesis: the paradox of carbon dioxide efflux." Curr Biol **7**(10): R637-639.

Farrell, J., L. Ramos, M. Tresguerres, M. Kamenetsky, L. R. Levin and J. Buck (2008). "Somatic 'soluble' adenylyl cyclase isoforms are unaffected in Sacy tm1Lex/Sacy tm1Lex 'knockout' mice." PLoS ONE **3**(9): e3251.

Feinstein, P. G., K. A. Schrader, H. A. Bakalyar, W. J. Tang, J. Krupinski, A. G. Gilman and R. R. Reed (1991). "Molecular cloning and characterization of a Ca^{2+}/calmodulin-insensitive adenylyl cyclase from rat brain." Proc Natl Acad Sci U S A **88**(22): 10173-10177.

Findeisen, F., J. U. Linder, A. Schultz, J. E. Schultz, B. Brugger, F. Wieland, I. Sinning and I. Tews (2007). "The Structure of the Regulatory Domain of the Adenylyl Cyclase Rv1264 from *Mycobacterium tuberculosis* with Bound Oleic Acid." J Mol Biol **369**(5): 1282-1295.

Florio, V. A. and E. M. Ross (1983). "Regulation of the catalytic component of adenylate cyclase. Potentiative interaction of stimulatory ligands and 2',5'-dideoxyadenosine." Mol Pharmacol **24**(2): 195-202.

Forte, L. R. (1983). "Activation of renal adenylate cyclase by forskolin: assessment of enzymatic activity in animal models of the secondary hyperparathyroid state." Arch Biochem Biophys **225**(2): 898-905.

Gao, B. N. and A. G. Gilman (1991). "Cloning and expression of a widely distributed (type IV) adenylyl cyclase." Proc Natl Acad Sci U S A **88**(22): 10178-10182.

Garau, J. and L. Gomez (2003). "*Pseudomonas aeruginosa* pneumonia." Curr Opin Infect Dis **16**(2): 135-143.

Garbers, D. L., D. J. Tubb and R. V. Hyne (1982). "A requirement of bicarbonate for Ca^{2+}-induced elevations of cyclic AMP in guinea pig spermatozoa." J Biol Chem **257**(15): 8980-8984.

Gille, A., G. H. Lushington, T. C. Mou, M. B. Doughty, R. A. Johnson and R. Seifert (2004). "Differential inhibition of adenylyl cyclase isoforms and soluble guanylyl cyclase by purine and pyrimidine nucleotides." J Biol Chem **279**(19): 19955-19969.

Goulding, E. H., J. Ngai, R. H. Kramer, S. Colicos, R. Axel, S. A. Siegelbaum and A. Chess (1992). "Molecular cloning and single-channel properties of the cyclic nucleotide-gated channel from catfish olfactory neurons." Neuron **8**(1): 45-58.

Gross-Langenhoff, M., K. Hofbauer, J. Weber, A. Schultz and J. E. Schultz (2006). "cAMP is a ligand for the tandem GAF domain of human phosphodiesterase 10 and cGMP for the tandem GAF domain of phosphodiesterase 11." J Biol Chem **281**(5): 2841-2846.

Guillou, J. L., H. Nakata and D. M. Cooper (1999). "Inhibition by calcium of mammalian adenylyl cyclases." J Biol Chem **274**(50): 35539-35545.

Guo, Q., Y. Shen, N. L. Zhukovskaya, J. Florian and W. J. Tang (2004). "Structural and kinetic analyses of the interaction of anthrax adenylyl cyclase toxin with reaction products cAMP and pyrophosphate." J Biol Chem **279**(28): 29427-29435.

Guo, Y. L., U. Kurz, A. Schultz, J. U. Linder, D. Dittrich, C. Keller, S. Ehlers, P. Sander and J. E. Schultz (2005). "Interaction of Rv1625c, a mycobacterial class IIIa adenylyl cyclase, with a mammalian congener." Mol Microbiol **57**(3): 667-677.

Guo, Y. L., H. Mayer, W. Vollmer, D. Dittrich, P. Sander, A. Schultz and J. E. Schultz (2009). "Polyphosphates from *Mycobacterium bovis*--potent inhibitors of class III adenylate cyclases." FEBS J **276**(4): 1094-1103.

Guo, Y. L., T. Seebacher, U. Kurz, J. U. Linder and J. E. Schultz (2001). "Adenylyl cyclase Rv1625c of *Mycobacterium tuberculosis*: a progenitor of mammalian adenylyl cyclases." Embo J **20**(14): 3667-3675.

Haber, N., D. Stengel, N. Defer, N. Roeckel, M. G. Mattei and J. Hanoune (1994). "Chromosomal mapping of human adenylyl cyclase genes type III, type V and type VI." Hum Genet **94**(1): 69-73.

Hacker, B. M., J. E. Tomlinson, G. A. Wayman, R. Sultana, G. Chan, E. Villacres, C. Disteche and D. R. Storm (1998). "Cloning, chromosomal mapping, and regulatory properties of the human type 9 adenylyl cyclase (ADCY9)." Genomics **50**(1): 97-104.

Hallem, E. A. and P. W. Sternberg (2008). "Acute carbon dioxide avoidance in *Caenorhabditis elegans*." Proc Natl Acad Sci U S A **105**(23): 8038-8043.

Haller, J. C., S. Carlson, K. J. Pederson and D. E. Pierson (2000). "A chromosomally encoded type III secretion pathway in *Yersinia enterocolitica* is important in virulence." Mol Microbiol **36**(6): 1436-1446.

Hammer, A., D. R. Hodgson and M. J. Cann (2006). "Regulation of prokaryotic adenylyl cyclases by CO_2." Biochem J **396**(2): 215-218.

Heidemann, S. R., H. C. Joshi, A. Schechter, J. R. Fletcher and M. Bothwell (1985). "Synergistic effects of cyclic AMP and nerve growth factor on neurite outgrowth and microtubule stability of PC12 cells." J Cell Biol **100**(3): 916-927.

Hellevuo, K., R. Berry, J. M. Sikela and B. Tabakoff (1995). "Localization of the gene for a novel human adenylyl cyclase (ADCY7) to chromosome 16." Hum Genet **95**(2): 197-200.

Hess, K. C., B. H. Jones, B. Marquez, Y. Chen, T. S. Ord, M. Kamenetsky, C. Miyamoto, J. H. Zippin, G. S. Kopf, S. S. Suarez, L. R. Levin, C. J. Williams, J. Buck and S. B. Moss (2005). "The "soluble" adenylyl cyclase in sperm mediates multiple signaling events required for fertilization." Dev Cell **9**(2): 249-259.

Hopker, V. H., D. Shewan, M. Tessier-Lavigne, M. Poo and C. Holt (1999). "Growth-cone attraction to netrin-1 is converted to repulsion by laminin-1." Nature **401**(6748): 69-73.

Horinouchi, S., M. Kito, M. Nishiyama, K. Furuya, S. K. Hong, K. Miyake and T. Beppu (1990). "Primary structure of AfsR, a global regulatory protein for secondary metabolite formation in *Streptomyces coelicolor* A3(2)." Gene **95**(1): 49-56.

Hu, J., C. Zhong, C. Ding, Q. Chi, A. Walz, P. Mombaerts, H. Matsunami and M. Luo (2007). "Detection of near-atmospheric concentrations of CO_2 by an olfactory subsystem in the mouse." Science **317**(5840): 953-957.

Hueck, C. J., M. J. Hantman, V. Bajaj, C. Johnston, C. A. Lee and S. I. Miller (1995). "*Salmonella typhimurium* secreted invasion determinants are homologous to Shigella Ipa proteins." Mol Microbiol **18**(3): 479-490.

Hyne, R. V. and D. L. Garbers (1979). "Calcium-dependent increase in adenosine 3',5'-monophosphate and induction of the acrosome reaction in guinea pig spermatozoa." Proc Natl Acad Sci U S A **76**(11): 5699-5703.

Ide, M. (1969). "Adenyl cyclase of *Escherichia coli*." Biochem Biophys Res Commun **36**(1): 42-46.

Inada, T., H. Takahashi, T. Mizuno and H. Aiba (1996). "Down regulation of cAMP production by cAMP receptor protein in *Escherichia coli*: an assessment of the contributions of transcriptional and posttranscriptional control of adenylate cyclase." Mol Gen Genet **253**(1-2): 198-204.

Ishikawa, Y., S. Katsushika, L. Chen, N. J. Halnon, J. Kawabe and C. J. Homcy (1992). "Isolation and characterization of a novel cardiac adenylyl cyclase cDNA." J Biol Chem **267**(19): 13553-13557.

Jaiswal, B. S. and M. Conti (2001). "Identification and functional analysis of splice variants of the germ cell soluble adenylyl cyclase." J Biol Chem **276**(34): 31698-31708.

Jameson, D. M. and J. F. Eccleston (1997). "Fluorescent nucleotide analogs: synthesis and applications." Methods Enzymol **278**: 363-390.

Janda, J. M. (1991). "Recent advances in the study of the taxonomy, pathogenicity, and infectious syndromes associated with the genus *Aeromonas*." Clin Microbiol Rev **4**(4): 397-410.

Johnson, R. A., W. Saur and K. H. Jakobs (1979). "Effects of prostaglandin E1 and adenosine on metal and metal-ATP kinetics of platelet adenylate cyclase." J Biol Chem **254**(4): 1094-1101.

Johnson, R. A. and I. Shoshani (1990). "Inhibition of *Bordetella pertussis* and *Bacillus anthracis* adenylyl cyclases by polyadenylate and "P"-site agonists." J Biol Chem **265**(31): 19035-19039.

Jones, W. D., P. Cayirlioglu, I. G. Kadow and L. B. Vosshall (2007). "Two chemosensory receptors together mediate carbon dioxide detection in *Drosophila*." Nature **445**(7123): 86-90.

Kalamidas, S. A., M. P. Kuehnel, P. Peyron, V. Rybin, S. Rauch, O. B. Kotoulas, M. Houslay, B. A. Hemmings, M. G. Gutierrez, E. Anes and G. Griffiths (2006). "cAMP synthesis and degradation by phagosomes regulate actin assembly and fusion events: consequences for mycobacteria." J Cell Sci **119**(Pt 17): 3686-3694.

Kamenetsky, M., S. Middelhaufe, E. M. Bank, L. R. Levin, J. Buck and C. Steegborn (2006). "Molecular details of cAMP generation in mammalian cells: a tale of two systems." J Mol Biol **362**(4): 623-639.

Kanacher, T., A. Schultz, J. U. Linder and J. E. Schultz (2002). "A GAF-domain-regulated adenylyl cyclase from *Anabaena* is a self-activating cAMP switch." Embo J **21**(14): 3672-3680.

Kasahara, M., K. Yashiro, T. Sakamoto and M. Ohmori (1997). "The *Spirulina platensis* adenylate cyclase gene, cyaC, encodes a novel signal transduction protein." Plant Cell Physiol **38**(7): 828-836.

Katayama, M. and M. Ohmori (1997). "Isolation and characterization of multiple adenylate cyclase genes from the cyanobacterium *Anabaena* sp. strain PCC 7120." J Bacteriol **179**(11): 3588-3593.

Katsushika, S., L. Chen, J. Kawabe, R. Nilakantan, N. J. Halnon, C. J. Homcy and Y. Ishikawa (1992). "Cloning and characterization of a sixth adenylyl cyclase isoform: types V and VI constitute a subgroup within the mammalian adenylyl cyclase family." Proc Natl Acad Sci U S A **89**(18): 8774-8778.

Kaupp, U. B., T. Niidome, T. Tanabe, S. Terada, W. Bonigk, W. Stuhmer, N. J. Cook, K. Kangawa, H. Matsuo, T. Hirose and et al. (1989). "Primary

structure and functional expression from complementary DNA of the rod photoreceptor cyclic GMP-gated channel." Nature **342**(6251): 762-766.

Kawasaki, H., G. M. Springett, N. Mochizuki, S. Toki, M. Nakaya, M. Matsuda, D. E. Housman and A. M. Graybiel (1998). "A family of cAMP-binding proteins that directly activate Rap1." Science **282**(5397): 2275-2279.

Ketkar, A. D., A. R. Shenoy, M. M. Kesavulu, S. S. Visweswariah and K. Suguna (2004). "Purification, crystallization and preliminary X-ray diffraction analysis of the catalytic domain of adenylyl cyclase Rv1625c from *Mycobacterium tuberculosis*." Acta Crystallogr D Biol Crystallogr **60**(Pt 2): 371-373.

Khelef, N., A. Zychlinsky and N. Guiso (1993). "*Bordetella pertussis* induces apoptosis in macrophages: role of adenylate cyclase-hemolysin." Infect Immun **61**(10): 4064-4071.

Kim, Y. R., S. Y. Kim, C. M. Kim, S. E. Lee and J. H. Rhee (2005). "Essential role of an adenylate cyclase in regulating *Vibrio vulnificus* virulence." FEMS Microbiol Lett **243**(2): 497-503.

Klein, C., R. K. Sunahara, T. Y. Hudson, T. Heyduk and A. C. Howlett (2002). "Zinc inhibition of cAMP signaling." J Biol Chem **277**(14): 11859-11865.

Klengel, T., W. J. Liang, J. Chaloupka, C. Ruoff, K. Schroppel, J. R. Naglik, S. E. Eckert, E. G. Mogensen, K. Haynes, M. F. Tuite, L. R. Levin, J. Buck and F. A. Muhlschlegel (2005). "Fungal adenylyl cyclase integrates CO_2 sensing with cAMP signaling and virulence." Curr Biol **15**(22): 2021-2026.

Krimm, S. and J. Bandekar (1986). "Vibrational Spectroscopy and Conformation of Peptides, Polypeptides, and Proteins." Advances in Protein Chemistry **38**: 181-364.

Krupinski, J., F. Coussen, H. A. Bakalyar, W. J. Tang, P. G. Feinstein, K. Orth, C. Slaughter, R. R. Reed and A. G. Gilman (1989). "Adenylyl cyclase amino acid sequence: possible channel- or transporter-like structure." Science **244**(4912): 1558-1564.

Kwon, J. Y., A. Dahanukar, L. A. Weiss and J. R. Carlson (2007). "The molecular basis of CO_2 reception in *Drosophila*." Proc Natl Acad Sci U S A **104**(9): 3574-3578.

Ladant, D. and A. Ullmann (1999). "*Bordatella pertussis* adenylate cyclase: a toxin with multiple talents." Trends Microbiol **7**(4): 172-176.

Lehotay, D. C. and B. E. Murphy (1974). "Changes in spleen weight, cyclic-AMP content, and serum-corticosterone levels of rats in response to CO_2." Can J Physiol Pharmacol **52**(6): 1074-1079.

Leppla, S. H. (1982). "Anthrax toxin edema factor: a bacterial adenylate cyclase that increases cyclic AMP concentrations of eukaryotic cells." Proc Natl Acad Sci U S A **79**(10): 3162-3166.

Linder, J. U. (2006). "Class III adenylyl cyclases: molecular mechanisms of catalysis and regulation." Cell Mol Life Sci **63**(15): 1736-1751.

Linder, J. U. (2008). "Structure-function relationships in *Escherichia coli* adenylyl cyclase." Biochem J.

Linder, J. U., L. I. Castro, Y. L. Guo and J. E. Schultz (2004). "Functional chimeras between the catalytic domains of the mycobacterial adenylyl cyclase Rv1625c and a Paramecium guanylyl cyclase." FEBS Lett **568**(1-3): 151-154.

Linder, J. U., A. Hammer and J. E. Schultz (2004). "The effect of HAMP domains on class IIIb adenylyl cyclases from *Mycobacterium tuberculosis*." Eur J Biochem **271**(12): 2446-2451.

Linder, J. U., A. Schultz and J. E. Schultz (2002). "Adenylyl cyclase Rv1264 from *Mycobacterium tuberculosis* has an autoinhibitory N-terminal domain." J Biol Chem **277**(18): 15271-15276.

Linder, J. U. and J. E. Schultz (2003). "The class III adenylyl cyclases: multi-purpose signalling modules." Cell Signal **15**(12): 1081-1089.

Lorimer, G. H. (1979). "Evidence for the existence of discrete activator and substrate sites for CO_2 on ribulose-1,5-bisphosphate carboxylase." J Biol Chem **254**(13): 5599-5601.

Lorimer, G. H., M. R. Badger and T. J. Andrews (1976). "The activation of ribulose-1,5-bisphosphate carboxylase by carbon dioxide and magnesium ions. Equilibria, kinetics, a suggested mechanism, and physiological implications." Biochemistry **15**(3): 529-536.

Lowrie, D. B., V. R. Aber and P. S. Jackett (1979). "Phagosome-lysosome fusion and cyclic adenosine 3':5'-monophosphate in macrophages infected with *Mycobacterium microti*, *Mycobacterium bovis* BCG or *Mycobacterium lepraemurium*." J Gen Microbiol **110**(2): 431-441.

Lowrie, D. B., P. S. Jackett and N. A. Ratcliffe (1975). "*Mycobacterium microti* may protect itself from intracellular destruction by releasing cyclic AMP into phagosomes." Nature **254**(5501): 600-602.

Makman, R. S. and E. W. Sutherland (1965). "Adenosine 3',5'-Phosphate in *Escherichia Coli*." J Biol Chem **240**: 1309-1314.

Masuda, S. and T. A. Ono (2004). "Biochemical characterization of the major adenylyl cyclase, Cya1, in the cyanobacterium *Synechocystis* sp. PCC 6803." FEBS Lett **577**(1-2): 255-258.

Masuda, S. and T. A. Ono (2005). "Adenylyl cyclase activity of Cya1 from the cyanobacterium *Synechocystis* sp. strain PCC 6803 is inhibited by bicarbonate." J Bacteriol **187**(14): 5032-5035.

Mathews, C. K., van Holde, K. E., Ahern, K. G. (2000). Biochemistry.

Matsuoka, I., G. Giuili, M. Poyard, D. Stengel, J. Parma, G. Guellaen and J. Hanoune (1992). "Localization of adenylyl and guanylyl cyclase in rat brain by in situ hybridization: comparison with calmodulin mRNA distribution." J Neurosci **12**(9): 3350-3360.

McCue, L. A., K. A. McDonough and C. E. Lawrence (2000). "Functional classification of cNMP-binding proteins and nucleotide cyclases with implications for novel regulatory pathways in *Mycobacterium tuberculosis*." Genome Res **10**(2): 204-219.

Mekalanos, J. J., D. J. Swartz, G. D. Pearson, N. Harford, F. Groyne and M. de Wilde (1983). "Cholera toxin genes: nucleotide sequence, deletion analysis and vaccine development." Nature **306**(5943): 551-557.

Mock, M. and A. Ullmann (1993). "Calmodulin-activated bacterial adenylate cyclases as virulence factors." Trends Microbiol **1**(5): 187-192.

Morris, A. L., M. W. MacArthur, E. G. Hutchinson and J. M. Thornton (1992). "Stereochemical quality of protein structure coordinates." Proteins **12**(4): 345-364.

Ohmori, K., M. Hirose and M. Ohmori (1992). "Function of cAMP as a mat-forming factor in the cyanobacterium *Spirulina platensis*." Plant Cell Physiol **33**: 21-25.

Ornstein, L. (1964). "Disc Electrophoresis. I. Background and Theory." Ann N Y Acad Sci **121**: 321-349.

Otwinowski, Z. and W. Minor (1997). "Processing of X-ray diffraction data collected in oscillation mode." Methods Enzymol **276**(4986): 307-326.

Pal, R. and D. Parker (2007). "A single component ratiometric pH probe with long wavelength excitation of europium emission." Chem Commun (Camb)(5): 474-476.

Pastan, I. and R. L. Perlman (1969). "Stimulation of tryptophanase synthesis in *Escherichia coli* by cyclic 3',5'-adenosine monophosphate." J Biol Chem **244**(8): 2226-2232.

Pastor-Soler, N., V. Beaulieu, T. N. Litvin, N. Da Silva, Y. Chen, D. Brown, J. Buck, L. R. Levin and S. Breton (2003). "Bicarbonate-regulated adenylyl cyclase (sAC) is a sensor that regulates pH-dependent V-ATPase recycling." J Biol Chem **278**(49): 49523-49529.

Perez-Garcia, M. T., L. Almaraz and C. Gonzalez (1990). "Effects of different types of stimulation on cyclic AMP content in the rabbit carotid body: functional significance." J Neurochem **55**(4): 1287-1293.

Perlman, R. L. and I. Pastan (1968). "Regulation of beta-galactosidase synthesis in *Escherichia coli* by cyclic adenosine 3',5'-monophosphate." J Biol Chem **243**(20): 5420-5427.

Petersen, S. and G. M. Young (2002). "Essential role for cyclic AMP and its receptor protein in *Yersinia enterocolitica* virulence." Infect Immun **70**(7): 3665-3672.

Peterson, K. M. and J. J. Mekalanos (1988). "Characterization of the *Vibrio cholerae* ToxR regulon: identification of novel genes involved in intestinal colonization." Infect Immun **56**(11): 2822-2829.

Premont, R. T., O. Jacobowitz and R. Iyengar (1992). "Lowered responsiveness of the catalyst of adenylyl cyclase to stimulation by GS in heterologous desensitization: a role for adenosine 3',5'-monophosphate-dependent phosphorylation." Endocrinology **131**(6): 2774-2784.

Premont, R. T., I. Matsuoka, M. G. Mattei, Y. Pouille, N. Defer and J. Hanoune (1996). "Identification and characterization of a widely expressed form of adenylyl cyclase." J Biol Chem **271**(23): 13900-13907.

Rall, T. W. and E. W. Sutherland (1958). "Formation of a cyclic adenine ribonucleotide by tissue particles." J Biol Chem **232**(2): 1065-1076.

Riggs, A. D., G. Reiness and G. Zubay (1971). "Purification and DNA-binding properties of the catabolite gene activator protein." Proc Natl Acad Sci U S A **68**(6): 1222-1225.

Rink, J., E. Ghigo, Y. Kalaidzidis and M. Zerial (2005). "Rab conversion as a mechanism of progression from early to late endosomes." Cell **122**(5): 735-749.

Roos, A. and W. F. Boron (1981). "Intracellular pH." Physiol Rev **61**(2): 296-434.

Russell, D. G. (2003). "Phagosomes, fatty acids and tuberculosis." Nat Cell Biol **5**(9): 776-778.

Saalau-Bethell, S. M., A. Cleassby, T. A. Sambrook, J. Coyle and M. Vinkovic (2007). Crystal structure of human soluble adenylate cyclase. http://www.wipo.int/pctdb/en/wo.jsp?wo=2007010285&IA=GB2006002750&DISPLAY=STATUS. W. I. P. Organization. **WO/2007/010285**.

Sakamoto, T., N. Murata and M. Ohmori (1991). "The concentration of cyclic AMP and adenylate cyclase activity in cyanobacteria." Plant Cell Physiol **32**: 581-857.

Salomon, Y., C. Londos and M. Rodbell (1974). "A highly sensitive adenylate cyclase assay." Anal Biochem **58**(2): 541-548.

Sarfati, R. S., V. K. Kansal, H. Munier, P. Glaser, A. M. Gilles, E. Labruyere, M. Mock, A. Danchin and O. Barzu (1990). "Binding of 3'-anthraniloyl-2'-deoxy-ATP to calmodulin-activated adenylate cyclase from *Bordetella pertussis* and *Bacillus anthracis*." J Biol Chem **265**(31): 18902-18906.

Schmid, A., Z. Sutto, M. C. Nlend, G. Horvath, N. Schmid, J. Buck, L. R. Levin, G. E. Conner, N. Fregien and M. Salathe (2007). "Soluble adenylyl cyclase is localized to cilia and contributes to ciliary beat frequency regulation via production of cAMP." J Gen Physiol **130**(1): 99-109.

Schultz, C. P., H. H. Eysel, H. H. Mantsch and M. Jackson (1996). "Carbon dioxide in tissues, cells and biological fluids detected by FTIR spectroscopy." J Phys Chem **100**: 6845-6848.

Seamon, K. B., W. Padgett and J. W. Daly (1981). "Forskolin - Unique Diterpene Activator of Adenylate-Cyclase in Membranes and in Intact-Cells." Proc Natl Acad Sci U S A **78**(6): 3363-3367.

Shen, T., Y. Suzuki, M. Poyard, N. Miyamoto, N. Defer and J. Hanoune (1997). "Expression of adenylyl cyclase mRNAs in the adult, in developing, and in the Brattleboro rat kidney." Am J Physiol **273**(1 Pt 1): C323-330.

Shen, Y., Y. S. Lee, S. Soelaiman, P. Bergson, D. Lu, A. Chen, K. Beckingham, Z. Grabarek, M. Mrksich and W. J. Tang (2002). "Physiological calcium concentrations regulate calmodulin binding and catalysis of adenylyl cyclase exotoxins." EMBO J **21**(24): 6721-6732.

Shenoy, A. R., N. P. Sreenath, M. Mahalingam and S. S. Visweswariah (2005). "Characterization of phylogenetically distant members of the adenylate cyclase family from mycobacteria: Rv1647 from *Mycobacterium tuberculosis* and its orthologue ML1399 from *M. leprae*." Biochem J **387**(Pt 2): 541-551.

Shenoy, A. R., N. Srinivasan, M. Subramaniam and S. S. Visweswariah (2003). "Mutational analysis of the *Mycobacterium tuberculosis* Rv1625c adenylyl cyclase: residues that confer nucleotide specificity contribute to dimerization." FEBS Lett **545**(2-3): 253-259.

Sinha, S. C., M. Wetterer, S. R. Sprang, J. E. Schultz and J. U. Linder (2005). "Origin of asymmetry in adenylyl cyclases: structures of *Mycobacterium tuberculosis* Rv1900c." Embo J **24**(4): 663-673.

Sismeiro, O., P. Trotot, F. Biville, C. Vivares and A. Danchin (1998). "*Aeromonas hydrophila* adenylyl cyclase 2: a new class of adenylyl cyclases with thermophilic properties and sequence similarities to proteins from hyperthermophilic archaebacteria." J Bacteriol **180**(13): 3339-3344.

Skorupski, K. and R. K. Taylor (1997). "Cyclic AMP and its receptor protein negatively regulate the coordinate expression of cholera toxin and toxin-coregulated pilus in Vibrio cholerae." Proc Natl Acad Sci U S A **94**(1): 265-270.

Smith, K. S. and J. G. Ferry (2000). "Prokaryotic carbonic anhydrases." FEMS Microbiol Rev **24**(4): 335-366.

Smith, R. S., M. C. Wolfgang and S. Lory (2004). "An adenylate cyclase-controlled signaling network regulates *Pseudomonas aeruginosa* virulence in a mouse model of acute pneumonia." Infect Immun **72**(3): 1677-1684.

Soderling, S. H. and J. A. Beavo (2000). "Regulation of cAMP and cGMP signaling: new phosphodiesterases and new functions." Curr Opin Cell Biol **12**(2): 174-179.

Sprang, S. R. (1997). "G protein mechanisms: insights from structural analysis." Annu Rev Biochem **66**: 639-678.

Steegborn, C., T. N. Litvin, K. C. Hess, A. B. Capper, R. Taussig, J. Buck, L. R. Levin and H. Wu (2005). "A novel mechanism for adenylyl cyclase inhibition from the crystal structure of its complex with catechol estrogen." J Biol Chem **280**(36): 31754-31759.

Steegborn, C., T. N. Litvin, L. R. Levin, J. Buck and H. Wu (2005). "Bicarbonate activation of adenylyl cyclase via promotion of catalytic active site closure and metal recruitment." Nat Struct Mol Biol **12**(1): 32-37.

Stengel, D., J. Parma, M. H. Gannage, N. Roeckel, M. G. Mattei, R. Barouki and J. Hanoune (1992). "Different chromosomal localization of two adenylyl cyclase genes expressed in human brain." Hum Genet **90**(1-2): 126-130.

Stessin, A. M., J. H. Zippin, M. Kamenetsky, K. C. Hess, J. Buck and L. R. Levin (2006). "Soluble adenylyl cyclase mediates nerve growth factor-induced activation of Rap1." J Biol Chem **281**(25): 17253-17258.

Stewart, D., T. Kamiyama, D. T. Mason, R. Miller and J. Wikman-Coffelt (1978). "Alteration in canine cardiac basal levels of cAMP and cGMP, and elevated tissue PCO_2 levels, induced by coronary ligation." J Mol Cell Cardiol **10**(2): 125-136.

Sturgill-Koszycki, S., P. H. Schlesinger, P. Chakraborty, P. L. Haddix, H. L. Collins, A. K. Fok, R. D. Allen, S. L. Gluck, J. Heuser and D. G. Russell (1994). "Lack of acidification in *Mycobacterium* phagosomes produced by exclusion of the vesicular proton-ATPase." Science **263**(5147): 678-681.

Sunahara, R. K., A. Beuve, J. J. Tesmer, S. R. Sprang, D. L. Garbers and A. G. Gilman (1998). "Exchange of substrate and inhibitor specificities between adenylyl and guanylyl cyclases." J Biol Chem **273**(26): 16332-16338.

Sunahara, R. K., C. W. Dessauer and A. G. Gilman (1996). "Complexity and diversity of mammalian adenylyl cyclases." Annu Rev Pharmacol Toxicol **36**: 461-480.

Sunahara, R. K., C. W. Dessauer, R. E. Whisnant, C. Kleuss and A. G. Gilman (1997). "Interaction of Gsalpha with the cytosolic domains of mammalian adenylyl cyclase." J Biol Chem **272**(35): 22265-22271.

Susstrunk, U., J. Pidoux, S. Taubert, A. Ullmann and C. J. Thompson (1998). "Pleiotropic effects of cAMP on germination, antibiotic biosynthesis and morphological development in Streptomyces coelicolor." Mol Microbiol **30**(1): 33-46.

Sutherland, E. W. (1972). "Studies on the mechanism of hormone action." Science **177**(47): 401-408.

Sutherland, E. W. and T. W. Rall (1958). "Fractionation and characterization of a cyclic adenine ribonucleotide formed by tissue particles." J Biol Chem **232**(2): 1077-1091.

Sutherland, E. W., T. W. Rall and T. Menon (1962). "Adenyl cylase. I. Distribution, preparation, and properties." J Biol Chem **237**: 1220-1227.

Suzuki, Y., T. Shen, M. Poyard, M. Best-Belpomme, J. Hanoune and N. Defer (1998). "Expression of adenylyl cyclase mRNAs in the denervated and in the developing mouse skeletal muscle." Am J Physiol **274**(6 Pt 1): C1674-1685.

Tacket, C. O., F. Brenner and P. A. Blake (1984). "Clinical features and an epidemiological study of Vibrio vulnificus infections." J Infect Dis **149**(4): 558-561.

Takai, K., Y. Kurashina, C. Suzuki-Hori, H. Okamoto and O. Hayaishi (1974). "Adenylate cyclase from *Brevibacterium liquefaciens*. I. Purification, crystallization, and some properties." J Biol Chem **249**(6): 1965-1972.

Tao, M. and F. Lipmann (1969). "Isolation of adenyl cyclase from *Escherichia coli*." Proc Natl Acad Sci U S A **63**(1): 86-92.

Taylor, R. K., V. L. Miller, D. B. Furlong and J. J. Mekalanos (1987). "Use of phoA gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin." Proc Natl Acad Sci U S A **84**(9): 2833-2837.

Tellez-Sosa, J., N. Soberon, A. Vega-Segura, M. E. Torres-Marquez and M. A. Cevallos (2002). "The *Rhizobium etli* cyaC product: characterization of a novel adenylate cyclase class." J Bacteriol **184**(13): 3560-3568.

Terauchi, K. and M. Ohmori (1999). "An adenylate cyclase, Cya1, regulates cell motility in the cyanobacterium *Synechocystis* sp. PCC 6803." Plant Cell Physiol **40**(2): 248-251.

Terauchi, K. and M. Ohmori (2004). "Blue light stimulates cyanobacterial motility via a cAMP signal transduction system." Mol Microbiol **52**(1): 303-309.

Tesmer, J. J., R. K. Sunahara, A. G. Gilman and S. R. Sprang (1997). "Crystal structure of the catalytic domains of adenylyl cyclase in a complex with Gsalpha.GTPgammaS." Science **278**(5345): 1907-1916.

Tesmer, J. J., R. K. Sunahara, R. A. Johnson, G. Gosselin, A. G. Gilman and S. R. Sprang (1999). "Two-metal-Ion catalysis in adenylyl cyclase." Science **285**(5428): 756-760.

Tews, I., F. Findeisen, I. Sinning, A. Schultz, J. E. Schultz and J. U. Linder (2005). "The structure of a pH-sensing mycobacterial adenylyl cyclase holoenzyme." Science **308**(5724): 1020-1023.

Townsend, P. D., P. M. Holliday, S. Fenyk, K. C. Hess, M. A. Gray, D. R. Hodgson and M. J. Cann (2009). "Stimulation of mammalian G protein-responsive adenylyl cyclases by carbon dioxide." J Biol Chem **284**(9): 784-791.

Via, L. E., D. Deretic, R. J. Ulmer, N. S. Hibler, L. A. Huber and V. Deretic (1997). "Arrest of mycobacterial phagosome maturation is caused by a block in vesicle fusion between stages controlled by rab5 and rab7." J Biol Chem **272**(20): 13326-13331.

Walsh, D. A., J. P. Perkins and E. G. Krebs (1968). "An adenosine 3',5'-monophosphate-dependant protein kinase from rabbit skeletal muscle." J Biol Chem **243**(13): 3763-3765.

Wang, Y., C. S. Lam, F. Wu, W. Wang, Y. Duan and P. Huang (2005). "Regulation of CFTR channels by $HCO(_3)^-$-sensitive soluble adenylyl cyclase in human airway epithelial cells." Am J Physiol Cell Physiol **289**(5): C1145-1151.

Watson, P. A., J. Krupinski, A. M. Kempinski and C. D. Frankenfield (1994). "Molecular cloning and characterization of the type VII isoform of mammalian adenylyl cyclase expressed widely in mouse tissues and in S49 mouse lymphoma cells." J Biol Chem **269**(46): 28893-28898.

Welton, A. F. and B. A. Simko (1980). "Metal ion and guanine nucleotide regulation of the inhibition of lung adenylate cyclase by adenosine analogs." Biochim Biophys Acta **615**(1): 252-261.

Whisnant, R. E., A. G. Gilman and C. W. Dessauer (1996). "Interaction of the two cytosolic domains of mammalian adenylyl cyclase." Proc Natl Acad Sci U S A **93**(13): 6621-6625.

Wolfgang, M. C., V. T. Lee, M. E. Gilmore and S. Lory (2003). "Coordinate regulation of bacterial virulence genes by a novel adenylate cyclase-dependent signaling pathway." Dev Cell **4**(2): 253-263.

Wu, K. Y., J. H. Zippin, D. R. Huron, M. Kamenetsky, U. Hengst, J. Buck, L. R. Levin and S. R. Jaffrey (2006). "Soluble adenylyl cyclase is required for netrin-1 signaling in nerve growth cones." Nat Neurosci **9**(10): 1257-1264.

Xia, Z. G., C. D. Refsdal, K. M. Merchant, D. M. Dorsa and D. R. Storm (1991). "Distribution of mRNA for the calmodulin-sensitive adenylate cyclase in rat brain: expression in areas associated with learning and memory." Neuron **6**(3): 431-443.

Yahr, T. L., A. J. Vallis, M. K. Hancock, J. T. Barbieri and D. W. Frank (1998). "ExoY, an adenylate cyclase secreted by the *Pseudomonas aeruginosa* type III system." Proc Natl Acad Sci U S A **95**(23): 13899-13904.

Yan, S. Z. and W. J. Tang (2002). "Construction of soluble adenylyl cyclase from human membrane-bound type 7 adenylyl cyclase." Methods Enzymol **345**: 231-241.

Yashiro, K., T. Sakamoto and M. Ohmori (1996). "Molecular characterization of an adenylate cyclase gene of the cyanobacterium Spirulina platensis." Plant Mol Biol **31**(1): 175-181.

Yoshimura, H., S. Yoshihara, S. Okamoto, M. Ikeuchi and M. Ohmori (2002). "A cAMP receptor protein, SYCRP1, is responsible for the cell motility of *Synechocystis* sp. PCC 6803." Plant Cell Physiol **43**(4): 460-463.

Yoshimura, M. and D. M. Cooper (1992). "Cloning and expression of a $Ca(^{2+})$-inhibitable adenylyl cyclase from NCB-20 cells." Proc Natl Acad Sci U S A **89**(15): 6716-6720.

Young, G. M., D. H. Schmiel and V. L. Miller (1999). "A new pathway for the secretion of virulence factors by bacteria: the flagellar export apparatus functions as a protein-secretion system." Proc Natl Acad Sci U S A **96**(11): 6456-6461.

Yu, H. J., J. R. Unnerstall and R. D. Green (1995). "Determination and cellular localization of adenylyl cyclase isozymes expressed in embryonic chick heart." FEBS Lett **374**(1): 89-94.

Zhang, G., Y. Liu, A. E. Ruoho and J. H. Hurley (1997). "Structure of the adenylyl cyclase catalytic core." Nature **386**(6622): 247-253.

Zimmermann, G., D. Zhou and R. Taussig (1998). "Mutations uncover a role for two magnesium ions in the catalytic mechanism of adenylyl cyclase." J Biol Chem **273**(31): 19650-19655.

Zippin, J. H., Y. Chen, P. Nahirney, M. Kamenetsky, M. S. Wuttke, D. A. Fischman, L. R. Levin and J. Buck (2003). "Compartmentalization of bicarbonate-sensitive adenylyl cyclase in distinct signaling microdomains." Faseb J **17**(1): 82-84.

Zubay, G., D. Schwartz and J. Beckwith (1970). "Mechanism of activation of catabolite-sensitive genes: a positive control system." Proc Natl Acad Sci U S A **66**(1): 104-110.

LDS