Structure-function and transcription regulation studies of angiotensin-converting enzyme 2

Cheng-Fu Chang

Supervisor: Prof. E. D. Sturrock

Thesis presented for the Degree of

Doctor of Philosophy

In the Division of Medical Biochemistry

University of Cape Town

August 2009

Declaration

I, Cheng-Fu Chang, declare that this thesis is my own, unaided work (except where acknowledge indicate otherwise). Neither the whole work nor part thereof has been, is being, or is to be submitted for any degree or examination at any other university.

I empower the University of Cape Town to reproduce for the purpose of research either the whole or any part of the contents of this thesis, an any manner whatsoever.

Signed by Candidate Signature Removed

Signature of candidate: ____

Signed on the 26th day of August, 2009

Acknowledgement

Although by definition I am required to submit a thesis which encompasses my original and unaided work, but the materialization of this thesis would have been impossible without the many people who have supported this work each in their own ways. To these people, I thank you.

To my supervisor, Prof. E. D. Sturrock, I give you my deepest and most sincere thanks, for your untiring guidance, your timely encouragements when things get tough, your dedication and your brilliant insights. I shall never forget your fine teaching of thinking outside the box, and I pray I can nurture this learning into a great mind someday.

To Ms. Sylva Schwager, I give you my gratitude of equal magnitude. Your enthusiasm for science has been a great inspiration for me throughout the years. The expertise and support that you have offered is everything that any student could have dreamed for, and more.

To all other members of the Zinc Metalloprotease Group, current and past, I thank you for being exactly who you are. There are many research groups in the world, and there is the Zinc Metalloprotease Group at the University of Cape Town. I am greatly privileged to be amongst great minds, and I thank science (and again, Ed) for bringing us all together into the same group. I will not list your names here, as I am selfish and I will not share you amazing people with the rest of the world.

To our overseas collaborators, Dr. Nigel Hooper, Dr. Vincent Dive, Dr. Anthony Chubb, Dr. Amy La Corte, Dr. Joseph Penninger and Dr. Dan Lambert, my countless thanks to you. Without your support and generosity, this study would have not been possible.

To the Division of Medical Biochemistry at the University of Cape Town, thank you for providing such an energetic research environment by attracting wonderful researchers into one big family. My special thanks go to the members of the cancer biology group and the T-box factor research group for all the helpful assistance and friendliness I have received over the years.

To Dr. Virna Leaner, a great thank you for your expert knowledge and advice on the topic of transcription regulation. Also many thanks to you for all the technical assistance you have provided.

My gratitude also goes to the funders, the Wellcome Trust, the National Research Foundation and the University of Cape Town, for their financial support of this study.

Last but not least, my eternal gratitude to every smile, every belief and every thought bestowed upon me along this journey.

Table of contents

List of abbreviations

°C	degree Celsius
µl	micro-litre
µM	micro-molar
Å	angstrom
ACE	angiotensin-converting enzyme
AMPS	ammonium persulfate
Ala	alanine
Ang	angiotensin
ARDS	acute respiratory distress syndrome
Asn	asparagine
Asp	aspartic acid
ATP	adenosine triphosphate
bp	base pair
CpA	carboxypeptidase A
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
Dnp	4, 7-dinitrophenol
dNTP	deoxyribnucleotide triphosphate
DTT	dithiothreitol
EDTA	7-methxoycoumarin-4-acetic acid
FCS	foetal calf serum
g	gravity
Glu	glutamic acid
HBS	HEPES buffered saline
HEK293	human embryonic kidney 293 cell
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	histidine
HNF-1β	hepatocyte nuclear factor 1 beta
kb	kilo-base pair
kDa	kilo-dalton
K_i	inhibition constant
KKS	kinin-kallikrein system
K_M	Michaelis constant
mA	milli-ampere
MAX1	MYC associated factor X 1
Mca	7-methxoycoumarin-4-acetic acid
mg	milli-gram
min	minute
ml	milli-litre
mM	milli-molar
MOPS	3-(N-Morpholino) propanesulfonic acid
NEP	Neprilysin
ng	nano-gram
nM	nano-molar
PAGE	polyacrylamide gel electrophoresis

PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	RCSB Protein Data Bank
Phe	phenylalanine
pmol	pico-mole
PMSF	phenylmethylsulphonyl fluoride
PNK	poly-nucleotide kinase
Pro	Proline
RAS	renin-angiotensin system
REST	RE1-silencing transcription factor
RNA	ribonucleic acid
RREB-1	*Ras* repressive element binding protein 1
S.D.	standard deviation
SARS CoV	severe acute respiratory syndrome coronavirus
SDS	sodium dodecyl sulfate
sec	second
SHR	spontaneous-hypertensive rat
siRNA	small interference ribonucleic acid
TBE	tris boric acid EDTA buffer
TBS	tris buffered saline
TEMED	Tetramethylethylenediamine
Thr	throenine
Tris	tris-(hydroxymethyl)-aminomethane
Tyr	tyrosine
UTE	untranslated exon
V_{max}	maximum velocity
ZEB1	zinc finger E-box binding homeobox 1
ZNF202	zinc finger protein 202

Abstract

The angiotensin-converting enzyme 2 (ACE2) is a novel component of the renin-angiotensin system (RAS). Advances in ACE2 research have suggested a role of ACE2 as a major regulator of angiotensin II in the RAS, and ACE2 activities in various tissues has been reported to cause beneficial tissue protective effects. Despite the knowledge gathered to date for ACE2, the physical controls of ACE2 activity at the protein-structure level and the expression level remain poorly understood. To address this, the active site structure-function and potential transcription regulators of ACE2 were investigated in this thesis.

The inhibitory action of the novel ACE2 inhibitor 416F2 on soluble ACE2 was determined by two different inhibition kinetic analyses, namely the dose-dependent response plot and the direct linear plot. Utilizing synthetic fluorogenic substrates of ACE2 (Mca-APK-Dnp and Mca-YVADAPK-Dnp), inhibition constants of low nanomolar range were elucidated, indicative of tight binding of 416F2 to ACE2. Subsequently, the structure of 416F2 was used as the ligand in an *in silico* molecular docking simulation to identify ACE2 residues that interact with the P_2-leucyl side chain of 416F2. In the docked model, no interaction was observed between the ACE2 active site residues and the P_2-leucyl group of 416F2, and hence the S_2 sub-site was not identified. However, a hydrogen bond between the Tyr510 terminal hydroxyl group and the acetyl-carbonyl group of 416F2 was seen in the docked model. Investigation of the structure-function significance of Tyr510 showed that mutation of Tyr510 to a threonine residue abolished ACE2 activity, while ACE2 activity was retained in the Tyr510 to phenylalanine mutant. This finding suggested that Tyr510 is a critical residue for ACE2 activity, owing to its bulky hydrophobic component, and its terminal hydroxyl component exerts a more subtle effect on the enzymatic activity of ACE2.

The control of ACE2 activity through its transcription regulation was investigated by analysis of the region immediately upstream of the ACE2 translation start codon. Truncated mutants of the immediate upstream region of the ACE2 gene were

subjected to transcription activity assays, and the presence of negative transcription elements in the -742 to -704 region upstream of the ACE2 translation start codon was revealed. Bioinformatics search through the -742 to -704 region for putative transcription factor binding sites elucidated the presence of multiple sites. Results from transcription activity assessments of putative binding site mutants and *in vitro* binding experiments confirmed the binding of two transcription factors, name ZEB1 and ZNF202, to the -742 to -704 region upstream of the ACE2 translation start codon to repress transcription activity. It was concluded that ZEB1 and ZNF202 are potential transcription repressors of ACE2 and contribute towards the regulation of ACE2 transcription in development and vascular maintenance respectively.

Chapter 1
Introduction

1.1. The renin-angiotensin system

The renin-angiotensin system (RAS) plays a major role in blood pressure and fluid homeostasis. The metabolism of a series of angiotensin peptides by the various RAS enzymes results in effective blood pressure regulation (Fyhrquist and Saijonmaa, 2008; Acharya *et al.*, 2003). The enzyme renin, secreted by the juxtaglomerular apparatus of the kidney, is responsible for the conversion of the angiotensin peptide precursor, angiotensinogen, to the biologically inactive angiotensin I (AngI). A zinc dipeptidyl carboxypeptidase, angiotensin-converting enzyme (ACE), converts AngI to angiotensin II (AngII) through the removal of the C-terminal histidyl-leucine dipeptide. AngII is a potent vasopeptide which mediates its physiological function through the angiotensin II receptors, and AngII induced responses associated with blood pressure elevation are mediated through the angiotensin II type 1 (AT_1) receptor. The typical responses elicited by activated AT_1 receptors are aldosterone and vasopressin release, as well as vasoconstriction. Aldosterone and vasopressin release results in the retention of salt and water by the kidney, which in effect increases the volume of the blood and leads to increased blood pressure. In addition, the AT_1 receptor has also been shown to mediate cell proliferation responses (Hansen *et al.*, 2008). AngII has also been shown to stimulate the angiotensin type II type 2 receptor to elicit physiological effects such as vasodilation and inhibition of cell proliferation (Igarashi *et al.*, 2007; Hsieh *et al.*, 2005; Carey and Padia, 2008). Through the generation of AngII, ACE plays a critical role in the regulation of blood pressure and influences cell proliferation, and thus ACE is an obvious target for therapeutic intervention in the treatment for hypertension-related cardiovascular and cardiac remodelling diseases.

While inhibitors that block AngII production have been successfully designed and employed as treatment for various cardiovascular diseases, an AngII metabolizing enzyme within the RAS was unknown until the discovery of the ACE homologue

termed angiotensin-converting enzyme 2 (ACE2) (Tipnis *et al.* 2000; Donoghue *et al.* 2000). The discovery of ACE2 led to the elucidation of a new axis in the RAS that serves as a regulatory mechanism of ACE activity. ACE2 acts as a monopeptidyl carboxypeptidase on AngII to form the vasopeptide angiotensin 1-7 (Ang1-7) (Fig.1.1). Many recent studies have uncovered information showing that Ang1-7 mediates various responses such as vasodilation and anti-cell proliferation by binding to the Mas receptor (Santos *et al.*, 2003; Iusuf *et al.*, 2008). ACE2 also acts on AngI to form angiotensin 1-9 (Ang1-9) (Donoghue *et al.*, 2000; Vickers *et al.*, 2002). The function of Ang1-9 is to date still unknown. Ang1-7 is also generated from Ang1-9 through the dipeptidyl carboxypeptidase action of ACE. The ability of ACE2 to metabolize both AngI and AngII allows ACE2 to regulate the ACE-axis of RAS through two different avenues: ACE2 lowers the level of AngII by metabolising AngII to Ang1-7, and Ang1-7 subsequently mediates responses countering those of AngII; secondly, ACE2 directly competes with ACE for the substrate AngI.

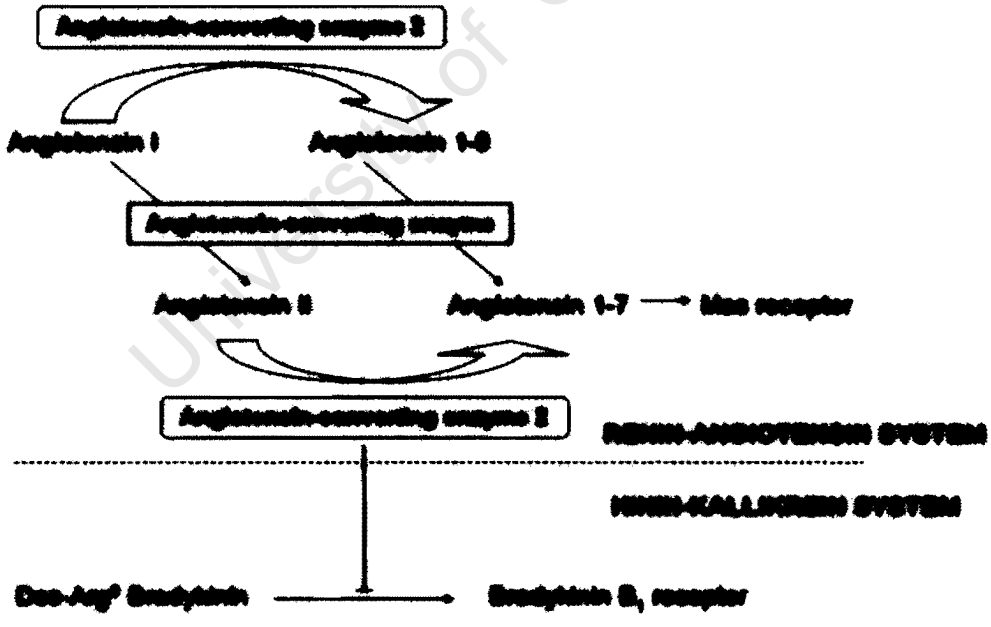

Fig.1.1. Schematic overview of the role of ACE2 in the processing of various vasopeptides of the RAS and KKS. In the RAS, ACE2 activity on AngI and AngII counter-regulates that of ACE. In the KKS, ACE2 degrades the vasoconstrictor des-Arg^9 bradykinin.

In addition to the RAS, it has also been documented that ACE2 is involved in the kinin-kallikrein system (KKS) (Fig.1.1). In the KKS, ACE2 deactivates des-Arg^9 bradykinin (Donoghue *et al.*, 2000; Vickers *et al.*, 2002). Des-Arg^9 bradykinin is a kinin peptide which elicits inflammation as well as a hypertensive response by binding to the bradykinin B_1 receptor (Ni *et al.*, 2003). Given the participation of ACE2 in the metabolism of numerous hormone peptides, combined with its identity as a novel homologue to the well studied ACE, ACE2 soon attracted immense research attention.

1.2. Physiological role of ACE2

1.2.1. ACE2 and cardiac function

As a novel component of the RAS and a homologue of ACE, ACE2 was immediately associated with blood pressure homeostasis when it was cloned in 2000 (Donoghue *et al.*, 2000; Tipnis *et al.*, 2000). The finding of highly localized expression of ACE2 in the heart and kidney provided support for such thinking (Tipnis *et al.*, 2000). The very first role of ACE2 suggested by animal knock-out (KO) studies was cardiac protection. In 2002, Crackower and colleagues reported that disruption of the ACE2 gene in mice resulted in the thinning of the left ventricular wall and reduced contractility of the left ventricle (Crackower *et al.*, 2002). The introduction of an *ace* deletion to an *ace2* KO background resulted in the rescue of the impaired heart phenotype, which suggested ACE is at least partially responsible for the dysfunctional cardiac phenotype that was prevented by ACE2. Crackower and colleagues also reported that the loss of ACE2 was not associated with an increase in blood pressure despite elevated levels of AngII were detected in the heart, kidney and plasma. Instead, a modest drop in blood pressure was detected in six-month *ace2*-null male mice and the authors reasoned that this was a result of weakened cardiac function (Crackower *et al.*, 2002).

Although the findings presented by Crackower *et al.* suggested ACE2 was critical for normal development of the heart in mice, later studies presented conflicting data. The *ace2*-null mice generated by Yamamoto and colleagues displayed nearly identical heart weight, cardiac function, basal blood pressure and AngII levels in the heart and kidney to that observed in wild-type mice (Yamamoto *et al.*, 2006). However, under pressure overload condition produced by transverse aortic constriction (TAC), cardiac hypertrophy was more severe in *ace2*-null mice when compared with wild-type mice, with the left ventricle of *ace2*-null mice exhibiting a significant dilatation and decreased contractility (Yamamoto *et al.*, 2006). Additionally, a greater increase in AngII levels in the heart and plasma after TAC was observed in *ace2*-null mutant mice than wild-type mice.

In a separate study by Gurley and colleagues, nearly identical cardiac physiology between *ace2*-null and wild-type mice was again reported (Gurley *et al.*, 2006).

Although the *ace2*-null mice displayed normal cardiac physiology, variable systolic blood pressures were measured for different cohorts of *ace2*-null mice. The mixed genetic makeup of the *ace2*-null mice was thought to be an attribute responsible for the variation in blood pressure (Gurley *et al.*, 2006). The findings reported by Gurley and colleagues suggested ACE2 is not directly involved in the maintenance of normal cardiac physiology or systolic blood pressure regulation, but exerts an effect on systolic blood pressure when associated with other genetic factors. In addition, *ace2*-null mice were more vulnerable to AngII induced hypertension as the elevation in blood pressure after AngII infusion was more exaggerated in *ace2*-null mice (Gurley *et al.*, 2006). The enhanced response to AngII infusion shown by *ace2*-null mice, and the occurrence of severe cardiac hypertrophy and cardiac fibrosis in *ace2*-null mice after TAC (Yamamoto *et al.*, 2006), collectively implied that ACE2 acts to preserve normal cardiac physiology under pathological conditions.

1.2.2. ACE2 and renal function

Although early *ace2* KO studies focused on the elucidation of the role of ACE2 in cardiac tissues, the first evidence of ACE2 function in the kidney came from a KO study where the loss of ACE2 resulted in progressive development of kidney injury (Oudit *et al.*, 2006). In young male mice, loss of ACE2 did not result in obvious abnormalities in kidney structure and function. However, deposition of fibrillar collagen, which is an indication of mesangial injury, was detected under electron microscopy. In male mice of twelve months, *ace2*-null mutants displayed diffused glomerulosclerosis and increased level of albumin in the urine which indicated kidney dysfunction (Oudit *et al.*, 2006). The observed kidney injury associated with loss of ACE2 was shown to be AngII-dependent, as treatment of the *ace2*-null mice with the AT_1 receptor blocker irbesartan prevented the development of glomerulosclerosis (Oudit *et al.*, 2006).

The protective role of ACE2 in the kidney was demonstrated by an investigation into the effect of ACE2 KO on diabetic kidney injury. A link between ACE2 and diabetic kidney pathology was established by the discovery of decreased ACE2 levels in kidneys of streptozocin-induced diabetic rats (Tikellis *et al.*, 2003). In a study by Wong *et al.*, deletion of *ace2* in mice carrying type1 diabetes mellitus ($ins^{WT/C96Y}$)

genotype potentiated the excretion rate of urinary albumin, which was indicative of worsened kidney injury, even though similar blood glucose levels were detected between *ace2*$^{-/y}$ *ins*$^{WT/C96Y}$ and *ace2*$^{+/y}$ *ins*$^{WT/C96Y}$ mice (Wong *et al.*, 2007). Further evidence of accelerated kidney injury in the diabetic mouse model came from the observation of increased immunostaining of fibronectin and α-smooth muscle actin in the glomeruli of *ace2*$^{-/y}$ *ins*$^{WT/C96}$ mice. Similar to the Oudit study, a dependency of accelerated kidney injury on AngII was implicated by the reduction in albumin excretion rate of the *ace2*$^{-/y}$ *ins*$^{WT/C96Y}$ mice to that of *ace2*$^{+/y}$ *ins*$^{WT/C96Y}$ mice after treatment with the AT_1 receptor blocker irbesartan (Wong *et al.*, 2007). Furthermore, inhibition of ACE2 with the inhibitor MLN4760 potentiated glomerular injury in streptozotocin-induced diabetic mice (Soler *et al.*, 2007). Although the exact mechanism implemented by ACE2 to achieve its protective role in kidney injury remains unclear, these ACE2 KO studies strongly suggest the involvement of RAS activity, particularly AngII metabolism.

1.2.3. ACE2 and pulmonary function

In 2003, Li *et al.* revealed a function for ACE2 as a potential receptor of the SARS coronavirus (CoV) in the lung (Li *et al.*, 2003). It was later confirmed that ACE2 is crucial for the development of SARS CoV-induced acute respiratory distress syndrome (ARDS) *in vivo*, as reduced levels of viral titre and SARS CoV Spike RNA were detected in the lungs of *ace2*-null mice (Kuba *et al.*, 2005). At the same time, ACE2 was shown to contribute towards protection against severe acute lung injury (Imai *et al.*, 2005). In mice, acute lung injury induced by acid aspiration was associated with increased lung elastance, lung oedema, infiltration of inflammatory cells and hyaline membrane formation. These conditions were found to be potentiated in mice lacking ACE2. ACE2 KO mice also suffered similar worsened loss of lung function, pulmonary oedema and leukocyte accumulation in the lung when subjected to sepsis-induced acute lung injury (Imai *et al.*, 2005). Interestingly, injection of recombinant human ACE2 into acid-treated mice was able to dampen the severity of lung injury in ACE2 KO mice, while wild-type mice injected with recombinant human ACE2 resulted in the reversal of lung dysfunction and pulmonary oedema (Imai *et al.*, 2005).

Since the protective effect was only observed in mice that were injected with catalytically active recombinant human ACE2, it was suggested that ACE2 activity contributes significantly towards the protection against acute lung injury. Furthermore, inactivation of the ACE gene improved the pathogenesis of acid aspiration induced lung injury in both wild-type and *ace2*-null mice (Imai *et al.*, 2005). Given the key roles of both ACE and ACE2 in the metabolism of AngII, it appears that the increased levels of AngII observed post acute lung injury in the absence of ACE2 are at least partly responsible for the aggravated disease in *ace2*-null mice (Imai *et al.*, 2005; Kuba *et al.*, 2005). In support of this view, mice treated with AT_1 receptor antagonist also showed reduced severity of lung injury (Imai *et al.*, 2005; Kuba *et al.*, 2005). Collectively these findings provided evidence for a protective role played by ACE2 in acute lung injury, and the protective function is mediated through the metabolism of AngII.

1.2.4. ACE2 and blood pressure

Although an abundance of evidence for ACE2 serving a tissue protective role has been accumulated thus far, there is a scarcity of definitive findings regarding the exact role played by ACE2 in blood pressure homeostasis. There is inconsistency amongst reports of basal blood pressures observed for *ace2*-null mice (Crackower *et al.*, 2002; Gurley *et al.*, 2006; Yamamoto *et al.*, 2006). The interplay between *ace2* KO mutation and other background genetic factors has been suggested as a contributing factor for the inconsistency in reported studies (Gurley *et al.*, 2006). In addition, ACE2 activity seems to contribute little towards the modification of basal blood pressure as similar basal plasma levels of AngII have been reported for wild-type and *ace2*-null mice (Gurley *et al.*, 2006; Yamamoto *et al.*, 2006). An examination of ACE2 expression in the kidney of spontaneous hypertensive rats (SHR) provided clues to the involvement of ACE2 in blood pressure regulation. Prior to the onset of hypertension, the ACE2 expression level in SHR was similar to that found in the normotensive Wistar Kyoto (WKY) rats (Tikellis *et al.*, 2006). In older animals, a reduction of ACE2 expression was correlated with the onset of hypertension. The authors concluded such developmental alteration in ACE2 expression is linked to the disruption of RAS and adult-onset hypertension, albeit further research is necessary for the full elucidation of the role played by ACE2 in blood pressure homeostasis.

1.3. ACE2 structure

1.3.1. Overall structure of ACE2

Thus far, ACE2 has been identified as the sole ACE homologue expressed in humans, while the gene for another ACE homologue, ACE3, has been described (Rella *et al.*, 2007). ACE2 shares a 42% amino acid sequence identity with its RAS counterpart ACE and there is a greater similarity between the N-domain of ACE and ACE2 (Donoghue *et al.*, 2000; Tipnis *et al.*, 2000). Like ACE, ACE2 is a type I transmembrane glycoprotein which consist of a N-terminal ectodomain exposed to the extracellular environment and anchored by a single transmembrane region that extends into the cytosolic C-terminal tail (Fig.1.2). The C-terminus of ACE2 is highly homologous to collectrin, a non-catalytic membrane glycoprotein which has been implicated in proper renal absorption of amino acids (Danilczyk *et al.*, 2006). Within the N-terminal ectodomain, the ACE-like catalytic domain is further divided into two sub-domains connected at the floor of the active site cavity (Towler *et al.*, 2004). The catalytic zinc and the N-terminus of the catalytic domain are located in sub-domain I whereas sub-domain II contains the C-terminus of the catalytic domain. The two sub-domains have been discovered to make a hinge-like movement upon ligand binding (Towler *et al.*, 2004) (Fig.1.3).

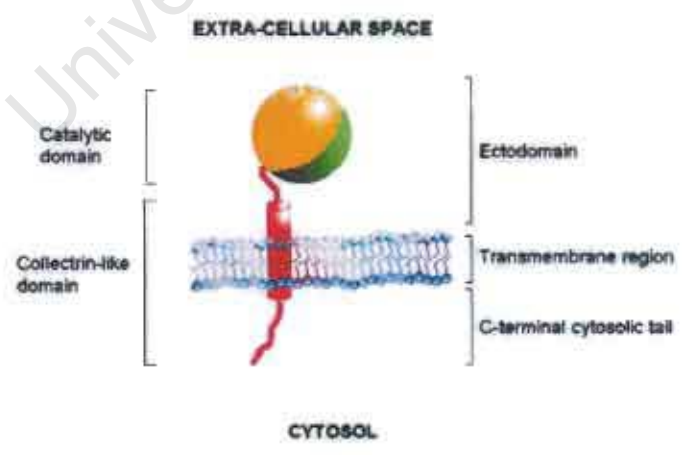

Fig.1.2. Schematic representation of the overall structure of ACE2. The ACE-like catalytic domain which forms part of the ectodomain is divided into sub-domain I (orange) and sub-domain II (green). The collectrin-like domain makes up the rest of the enzyme that is not catalytically active (red), which includes the transmembrane region and the C-terminal cytosolic tail.

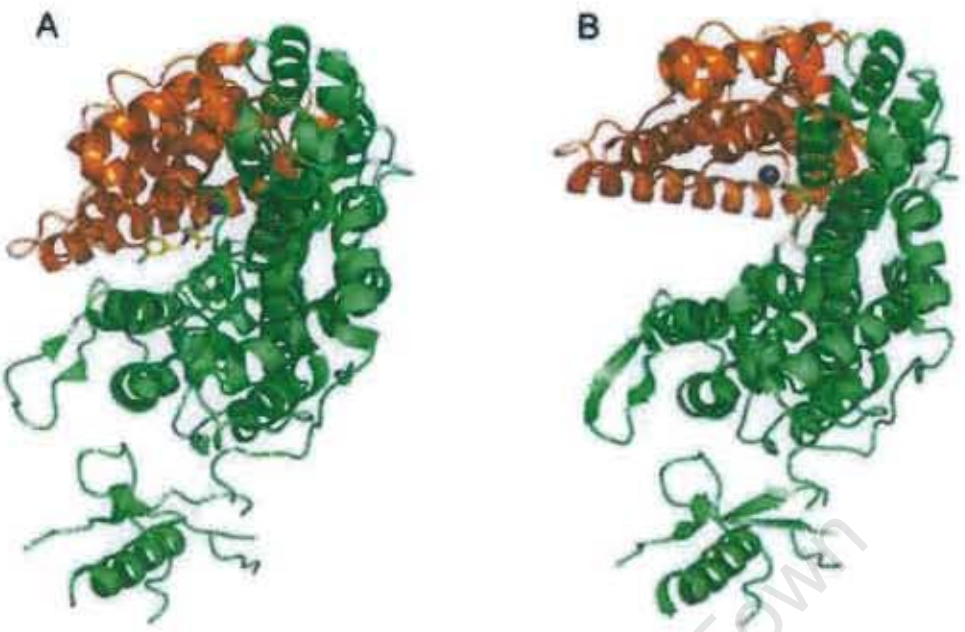

Fig.1.3. Ribbon representations of the x-ray crystal structure of ACE2 ectodomain, showing the hinge-like movement of the sub-domains upon ligand binding. **A.** Crystal structure of MLN4760-bound ACE2 (PDB code: 1R4L). **B.** Crystal structure of ligand-free ACE2 (PDB code: 1R42). Sub-domain I of the ACE2 catalytic domain is shown in orange, and sub-domain II is shown in green. The inhibitor MLN4760 is shown as stick model in yellow. The active site catalytic zinc is shown as sphere in dark gray. Images were generated using PyMol Molecular Viewer (DeLano Scientific LLC.).

1.3.2. Modelling of ACE2 active site

As a zinc metalloprotease, the consensus zinc binding sequence HEXXH is conserved in the amino acid sequence of ACE2. An additional glutamate residue (Glu402), which serves as the third zinc coordinating residue (corresponding to Glu389 in the N-domain and Glu987 in the C-domain of somatic ACE), is also conserved in ACE2 (Tipnis *et al.*, 2000). Several residue substitutions within the S_2' region account for the difference in the substrate selectivity between ACE and ACE2, as revealed by a comparative modelling study where a model of ACE2 active site was built using the crystal structure data of the testicular isoform of ACE (Guy *et al.*, 2003). The substitution of Gln281 in testis ACE with Arg273 in ACE2 prevents the accommodation of a P_2' substrate residue by ACE2. The arginine residue was modelled to interact with the substrate C-terminal carboxylate and thus limits the ability of the ACE2 active site to accommodate more than one residue on the primed

side (Fig.1.4). The ACE2 active site model also suggested that respective substitution of Lys511, Tyr520 and Thr282 in testis ACE with Leu503, Phe512 and Phe274 in ACE2 resulted in the occlusion of the S_2' sub-site (Fig.1.4). On the unprimed side of the active site, the modelled S_1 sub-site of ACE2 was expected to be smaller than that of testis ACE, as Val518 in testis ACE is substituted by a larger Tyr510 in ACE2 (Guy *et al.*, 2003).

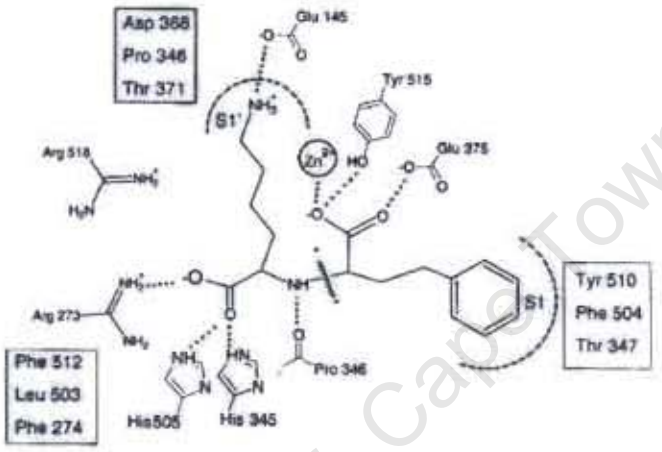

Fig.1.4. Schematic representation of the ACE2 active site model built using the active site structure of testis ACE. The modelled hydrogen bond interactions (dotted lines) between the ACE2 active site residues and the inhibitor lisinopril are shown. Residues located within the various sub-sites are listed in boxes. The P_2' prolyl group of lisinopril is not shown in this model as the S_2' binding site is absent in ACE2. Figure taken from (Guy *et al.*, 2003).

The modelled ACE2 active site not only provided an explanation for the inability of ACE inhibitors to act on ACE2, but also generated insights into the substrate preference of ACE2. The model provided support for the consensus substrate sequence [Pro-Xaa$^{(1-3)}$-Pro-hydrophobic residue] derived from a sequence comparison of various biological substrates of ACE2 (Vickers *et al.*, 2002). In the ACE2 active site model, the small S_1 sub-site of ACE2 formed by hydrophobic residues offered only sufficient space to accommodate small hydrophobic substrate residues and not bulky residues. This notion supported the observed conservation of a P_1 prolyl residue in biological peptides most effectively cleaved by ACE2 (Vickers *et al.*, 2002).

1.3.3. ACE2 active site crystal structure

Elucidation of the crystal structure of ACE2 (PDB code: 1R4L & 1R42) (Towler *et al.*, 2004) not only confirmed the involvement of the residues proposed by the previous ACE2 active site model in substrate binding (Guy *et al.*, 2003), but also shed light on the ACE2 active site structure-function relationships. In the ligand bound crystal structure of ACE2 (PDB code: 1R4L), the inhibitor MLN4760 was bound in the reverse orientation to what was anticipated from its design and the previous modelling study (Dales *et al.*, 2002; Guy *et al.*, 2003). Interestingly, the large 3, 5-dichlorobenzylimidazole did not occupy the S_1 and S_2 sub-sites of ACE2 as expected, but instead occupied the S_1' sub-site. There was a switch of carboxylates as the terminal carboxylate became the zinc coordinating group, and the original zinc coordinating carboxylate was found to interact with Arg273 as the terminal carboxylate (Towler *et al.*, 2004). In addition to Arg273, the terminal carboxylate of MLN4760 was also coordinated by His505 and His345 through hydrogen-bonding (Fig.1.5.A), as proposed by the modelled ACE2 active site (Guy *et al.*, 2003).

In the crystal structure, the S_1 sub-site of ACE2 is made up of Arg514, Thr347 and two bulky hydrophobic Phe504 and Tyr510 (Fig.1.5.B). The Tyr510 residue is positioned such that the hydrophobic ring forms the roof of the S_1 sub-site (Towler *et al.*, 2004). The substrate selectivity conferred by the small S_1 sub-site is evident from the inability of ACE2 to cleave biological substrates with a tyrosine or phenylalanine in the P_1 position (Vickers *et al.*, 2002).

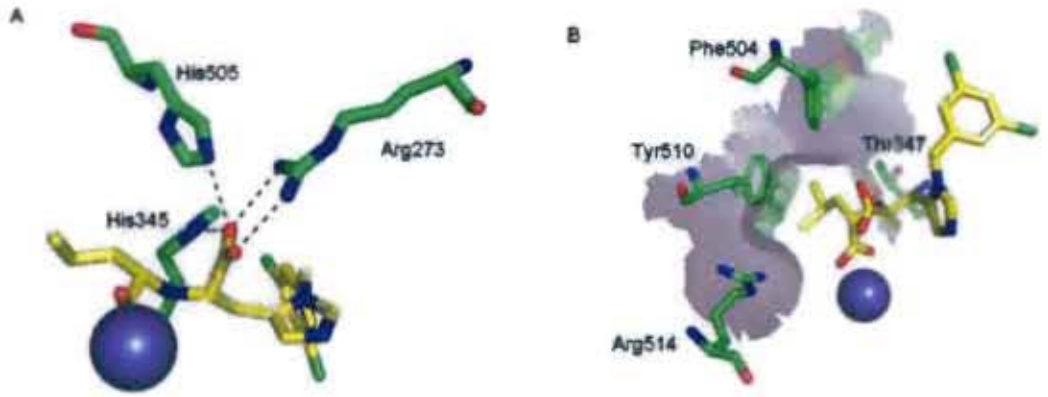

Fig.1.5. Schematic representations of the active site crystal structure of ACE2 (PDB code: 1R4L). The catalytic zinc is shown as grey sphere. **A.** The C-terminal carboxylate of the inhibitor MLN4760 (yellow) was coordinated by Arg273, His505 and His345 by hydrogen bond interactions (dashed lines). **B.** The residues making up the S_1 sub-site are Thr347, Phe504, Tyr510 and Arg514. The surface representation of the S_2 sub-site residues is also shown. Images were generated using PyMol Molecular Viewer (DeLano Scientific LLC.).

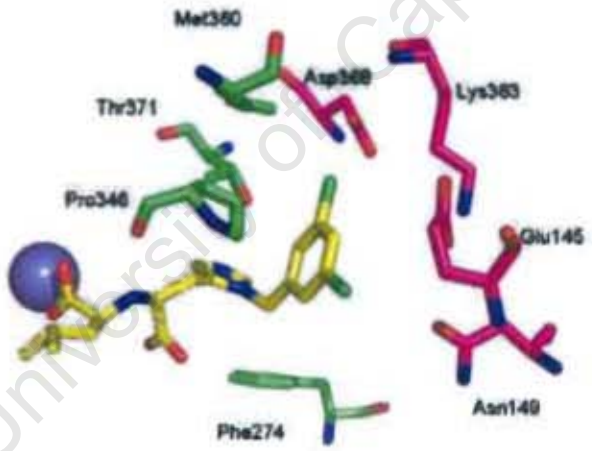

Fig.1.6. S_1' sub-site crystal structure of ACE2 (PDB code: 1R4L). The residues shown in green provide a hydrophobic environment which confers selectivity for hydrophobic substrate residues. The residues shown in magenta confer selectivity for basic substrate residues. The inhibitor MLN4760 is shown in yellow, and the catalytic zinc is shown as grey sphere. Image was generated using PyMol Molecular Viewer (DeLano Scientific LLC.).

While the S_1 sub-site of ACE2 active site is made up of residues grouped to sub-domain II of the ACE2 catalytic domain, the much larger S_1' sub-site is formed as part of the cleft between the two sub-domains. In the crystal structure, the large 3, 5-dichlorobenzylimidazole is well accommodated by the S_1' sub-site (Fig.1.6). The

hydrophobic environment created by Phe274, Pro346, Thr371, Met360 is particularly noteworthy, as it confers a substrate selectivity for a hydrophobic P_1' substrate residue (Vickers *et al.*, 2002). Additionally, the presence of Glu145, Asn149, Lys363 and Asp368 allows for the fitting of basic P_1' residues into the S_1' sub-site (Towler *et al.*, 2004) (Fig.1.6). Collectively these factors explain the substrate preference of ACE2 for large hydrophobic or basic residues at the P_1' position (Vickers *et al.*, 2002).

While it is clear that active site sub-sites determine the substrate selectivity of ACE2, structural attributes at the single residue level were also found to significantly affect the enzymatic activity of ACE2. In 2005, Guy *et al.* identified several active site residues to be essential for the catalytic activity of ACE2 (Guy *et al.*, 2005). Mutation of Arg273 to either a glutamine or a lysine led to complete abolition of ACE2 activity on a synthetic substrate, as did the replacement of His345 and His505 with an alanine or leucine (Guy *et al.*, 2005). In the case of Arg273, a positive charge at the correct distance was suggested to be crucial for proper binding of the substrate C-terminal carboxylate, as both the removal of the positive charge by a glutamine substitution and altered positioning of the positive charge by a lysine substitution resulted in the disruption of ACE2 activity (Guy *et al.*, 2005). In the MLN4760-bound ACE2 crystal structure, the C-terminal carboxylate of the inhibitor was coordinated by Arg273, His345 and His505 (Fig.1.5.A) (Towler *et al.*, 2004), and mutation of either histidine residues resulted in dramatically reduced activity (up to 300-fold) when compared to that of wild-type ACE2 (Guy *et al.*, 2005). The identification of these critical active site residues indicates the substrate selectivity of ACE2 is also controlled at the single residue level.

1.4. Chloride activation of ACE2

From the sequence alignment of ACE2 with its homologue ACE, it was noted that the residues responsible for chloride binding in ACE were conserved in ACE2 (Guy *et al.*, 2003). Although both chloride binding sites of ACE were conserved in ACE2, it was discovered later that only one of the conserved sites is bound with a chloride ion in the crystal structure (Towler *et al.*, 2004). The conservation of chloride ion binding sites strongly suggested the likelihood of ACE2 also experiencing chloride activation like its homologue ACE. The first report of the chloride sensitivity of ACE2 came from a study by Vickers and colleagues in 2003, where optimal ACE2 hydrolysis of the synthetic substrate Mca-APK-Dnp was detected at 1 M NaCl (Vickers *et al.*, 2002). Interestingly, the hydrolysis of AngII by ACE2 was shown to be inhibited by the presence of NaCl (approximately three-fold reduction in activity at 1 M NaCl) (Guy *et al.*, 2003). These findings suggested that chloride activation of ACE2 is a substrate-dependent event. While an optimal activity of ACE2 with the substrate Mca-APK-Dnp was reported at 1 M NaCl, pH 6.5 (Vickers *et al.*, 2002), optimal ACE2 activity has also been reported at 300 mM NaCl, pH 7.0, and at pH 8.0 in the absence of chloride (Guy *et al.*, 2003). Together these data indicated that chloride activation of ACE2 is a complex event which is both substrate and pH dependent.

In the ACE2 crystal structure (PDB code: 1R4L), a single chloride ion is coordinated by Arg169, Trp477, and Lys481 (Towler *et al.*, 2004). The Arg169 residue was confirmed to be a critical residue for the chloride sensitivity of ACE2, as a substitution of Arg169 with a glutamine abolished the chloride-activated activity of ACE2 with the physiological substrate AngII (Rushworth *et al.*, 2008). Similarly, Trp271 was also identified as a critical residue for the chloride activation of ACE2. Intriguingly, while chloride binding in the second conserved chloride binding site is abolished (Guy *et al.*, 2005; Towler *et al.*, 2004), mutation of the Arg514 residue in the second conserved chloride binding site resulted in a shift in substrate selectivity of ACE2 (Rushworth *et al.*, 2008). It was reasoned that since Arg514 also forms part of the S_1 sub-site of ACE2, the mutation of Arg514 potentially alters the environment of S_1 sub-site and affects substrate selectivity (Fig.1.5.B).

1.5. ACE2 modulating compounds

Soon after the discovery of ACE2, inhibitors of ACE2 were designed and synthesized with the intent to produce molecular tools for the elucidation of ACE2 physiological functions. While inhibitors of the well characterized homologue ACE had been developed long before the discovery of ACE2, design of ACE2 inhibitors had little to gain from the success of ACE inhibitors as these compounds fail to inhibit ACE2 (Tipnis *et al.*, 2000). To date, several potent ACE2 inhibitors have been successfully developed, and a brief summary of various ACE2 inhibitors reported in the literature are presented in Table 1.

1.5.1. Substrate-based ACE2 inhibitors

The designs of the first wave of ACE2 inhibitors was based on the natural substrates of ACE2 (Vickers *et al.*, 2002). Knowledge of the preference of ACE2 for its natural substrates was immediately utilized in the development of the first potent ACE2 inhibitor named MLN4760. The design of MLN4760 was based on the ligand P_1 prolyl group preference of ACE2 and its low efficiency at cleaving the C-terminal His-Leu bond of AngI (Vickers *et al.*, 2002). An imidazole scaffold with a zinc-coordinating carboxylate group was used in the synthesis of MLN4760, and the incorporation of a large 3, 5-dichlorobenzyl group at the N-3 position of the imidazole, together with a leucyl group as the C-terminal residue, led to a potent inhibitor of ACE2 with an IC_{50} value of 440 pmol (Dales *et al.*, 2002). MLN4760 was also found to be highly selective for ACE2, as indicated by its weak inhibition of human testis ACE (IC_{50} greater than 100 μM) and porcine carboxypeptidase A (IC_{50} value = 27 μM) (Dales *et al.*, 2002).

In addition to small molecule inhibitors of ACE2, short polypeptides with ACE2 inhibitory effect have also been reported. A strong peptide inhibitor of ACE2 named DX600 was discovered through phage display of constraint loop peptide libraries (Huang *et al.*, 2003). Huang and colleagues identified ten motifs from a cluster sequence analysis of various ACE2 binding peptides, and DX600 was derived from the C*X*P*X*R*XX*P*W**XX*C motif. This N-terminally acetylated, C-terminally amidated, twenty-six-residue polypeptide was found to inhibit ACE2 with an inhibition constant

(K_i) of 2.8 nM. The inhibition was also reported to be highly specific for ACE2 as DX600 displayed almost no inhibition of ACE even at a concentration of 100 μM (Huang *et al.*, 2003). The inhibitory action of DX600 has been related to the presence of an arginine residue between the two proline residues as well as the different spacing of the proline residues in the C*X***P***X***R***XX**P**W***XX*C motif. In addition, the constraint loop structure formed by the disulfide bridge has also been suggested as one of features preventing hydrolysis of DX600 by ACE2 (Huang *et al.*, 2003).

Table 1. Summary of ACE2 inhibitor reported to date. The inhibition constants for ACE2 and carboxypeptidase A (CpA) are shown.

Compound	**Design basis**	**K_i for ACE2 (nM)**	**K_i for CpA (nM)**
MLN4760	Substrate	0.44	27000
DX600	Substrate	2.8	N.D
NAAE	Structure	459000	N.D
4S-16659	Structure	62000[#]	N.D
1o (Deaton *et al.*, 2008a)	Structure	1.4	11000
1h (Deaton *et al.*, 2008b)	Structure	0.85	3300
41_{FII} (Mores *et al.*, 2008)	Substrate/Structure	0.4	1050

[#] IC_{50} value was reported
N.D. Not determined

1.5.2. ACE2 x-ray structure-based inhibitors

Elucidation of the ACE2 crystal structure has made possible the structure-based design of ACE2 inhibitors. Structural knowledge of ACE2 allowed for better prediction of the interaction between its active site residues and a ligand, and this was invaluable in the pursuit of inhibitors with even higher affinity and selectivity. The finding of MLN4760 binding in the reverse manner to its original design (Towler *et al.*, 2004) highlighted a weakness of substrate-based design of inhibitors, namely one

cannot fully deduce how a ligand will be accommodated by the active site of an enzyme solely from its substrate preference.

The first structure-based approach employed in the search for ACE2 inhibitors made use of the ACE2 crystal structure knowledge and *in silico* high-throughput screenings of small molecule databases. A small molecule with ACE2 inhibitory property yielded by this approach was N-(2-Aminoethly)-1 aziridine-ethanamine (NAAE) (Huentelman *et al.*, 2004). Although the specificity of NAAE towards ACE2 was not reported, and NAAE displayed a relatively high K_i of 459 μM for ACE2, this study presented a proof of concept for structure-based high-throughput screening of small molecules as potential ACE2 inhibitors. Similarly, a later study by Rella *et al.* employed the same strategy and compounds which inhibited ACE2 with IC_{50} values ranging from 62 to 179 μM were identified (Rella *et al.*, 2006).

In a recent study by Deaton *et al.*, the GlaxoSmithKline compound collection was screened for potential lead compounds to be further developed into ACE2 inhibitors, and a thio-acid compound was identified as a potent ACE2 inhibitor (K_i of 86 nM) (Deaton *et al.*, 2008a). However, this compound also displayed strong inhibitory activities towards ACE (30 nM) and NEP (1.1 nM). In an attempt to improve the potency and the selectivity of the inhibitor towards ACE2, structural analogs with various substitutions at the P_1 position were synthesized and characterized. The first finding presented by characterization of the P_1 analogs was that the (*R,R*) stereoisomers, in which the P_1 groups were positioned in a conformation not found in the biological substrates, failed to inhibit ACE2 (Deaton *et al.*, 2008a). It was also noted that analogs with linear alkanes as P_1 substituents displayed high potencies of inhibition, while analogs with P_1 substituents of a small α- branched nature (such as cyclobutane and cyclopentane) showed improved selectivity towards ACE2. Although higher ACE2 selectivity was achieved by analogs with larger cylcohexyl and phenyl P_1 substituents, this increased selectivity was associated with the loss of ACE2 inhibition (Deaton *et al.*, 2008a). In order to rationalize the potencies and specificities displayed by the different analogs, Deaton *et al* proceeded to perform molecular docking of the (R)-*sec*-butyl P_1 analog structure into the ACE2 crystal structure. In the docked model, the (R)-*sec*-butyl moiety was securely accommodated by Thr347,

Phe504, Tyr510 and Arg514 of the S_1 sub-site. The small S_1 sub-site also accommodated P_1 substituents such as cyclobutane and cyclopentane, but was less tolerant of cyclohexane (Deaton *et al.*, 2008a).

Subsequent to the investigation of the S_1 sub-site, Deaton *et al.* explored the S_1' sub-site using the same approach. A series of analogs with different P_1' substituents were synthesized and assessed for inhibitory activities towards ACE2, ACE, NEP and CpA (Deaton *et al.*, 2008b). It was discovered that analogs with a large P_1' group such as methylene *o*-phenoxyphenyl or methylene *p*-benzoxyphenyl moieties conferred high ACE2 inhibition (K_i values of the analogs ranged from 0.90 to 2.7 nM). Tethering of the methylene *p*-benzoxyphenyl group by addition of fluorine to the P_1'-end phenyl ring resulted in a further increase in inhibitory potency of the analog. The 3, 4-difluorobenzyl tyrosine P_1' analog was reported as the analog possessing the highest inhibitory activity and selectivity towards ACE2 (Deaton *et al.*, 2008b). While the increase in size of the P_1' substituent seemed to be associated with improved ACE2 inhibition and selectivity profile of the thio-acid analogs, the reduced potency of the analog with a larger trifluoromethyl P_1' substituent indicated such large P_1' moiety exceeds the space provided by the S_1' sub-site (Deaton *et al.*, 2008b). A comparison of the molecular models of different thio-acid P_1' analogs docked into the ACE2 crystal structure revealed that the ACE2 S_1' sub-site is able to bind the ligand P_1' moiety in different modes. In the S_1' sub-site, the P_1' 3, 4-difluorobenzyl tyrosine group of analog **1j** was observed to occupy a very different space to that occupied by the P_1' *o*-phenoxyphenyl group of analog **1r** (Fig.1.7) (Deaton *et al.*, 2008b). The overlay of the two docked models disclosed a S_1' sub-site shaped with two separate cavities. Such distinctly shaped S_1' sub-site provides a promising target for the development of future ACE2 inhibitors of not only even higher potency, but also a further increased selectivity towards ACE2.

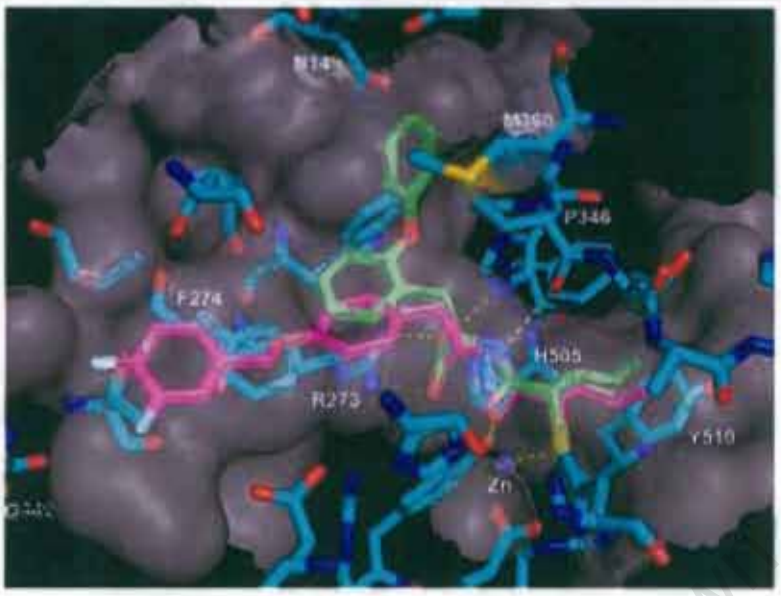

Fig.1.7. Docked models of thio-acid P_1' analogs **1j** (shown in green) and **1r** (shown in magenta) with the ACE2 crystal structure (PDB code: 1R4L). The different occupation of P_1' groups in the S_1' sub-site indicated that ACE2 binds its ligand with more than one binding mode at the S_1'/P_1' position. Figure taken from (Deaton *et al.*, 2008b).

The most recent ACE2 inhibitors reported in the literature are the phosphinic pseudo-peptides. The success in the development of potent, high domain-selective phosphinic ACE inhibitors (Dive *et al.*, 1999; Georgiadis *et al.*, 2003) has fueled the pursuit of phosphinic ACE2 inhibitors. A series of phosphinic ACE2 inhibitors were designed and synthesized with the goal of producing inhibitors with high bio-availability, as these inhibitors would provide invaluable aid to the full elucidation of ACE2 physiological functions (Mores *et al.*, 2008). The first group of compounds synthesized by Mores and colleagues were phosphinic dipeptides, and these compounds showed poor inhibition of ACE2 activity. However, strong inhibition of ACE2 was achieved by the phosphinic tripeptide inhibitors with P_2 groups. A variety of P_2 groups were accepted by the S_2 sub-site as similar inhibition constants were observed for compounds that differed only at the P_2 position. Interestingly, analogs not carrying a P_1 prolyl group showed decreased inhibition potencies. In the assessment of a docked model, Mores *et al.* noted the P_1 prolyl residue creates good van der Waals contacts with the Tyr510 and Phe504 residues, and such interaction may not exist for other P_1 residues placed in the same space (Mores *et al.*, 2008). Mores *et al.* also suggested the structural rigidity of the prolyl group could also contribute towards the difference in the observed inhibition potencies.

Although the presence of a P_2 group and a P_1 prolyl moiety appeared to be the major contributors towards the potency of the phosphinic ACE2 inhibitors, the selectivity of the inhibitors toward ACE2 was not mediated by either group. Several of the strong ACE2 phosphinic inhibitors were also associated with a strong inhibition of CpA (Mores *et al.*, 2008). In these inhibitors, increased ACE2 selectivity was gained through substitution of the P_1' phenyl group with a phenyl-isoxazole group. This substitution resulted in a more than 70-fold reduction in inhibition potency of the ACE2 phosphinic inhibitor towards CpA (Mores *et al.*, 2008). The increased ACE2 selectivity was explained by the better acceptance of large P_1' residues by the S_1' sub-site of ACE2 than by that of CpA. Given the fact that peptide sequences of these ACE2 phosphinic inhibitors mimicked those found in the biological substrates of ACE2, this study presented itself as an invaluable example of how both substrate-based and structure-based inhibitor designs were combined to produce potent ACE2 inhibitors of high selectivity.

1.5.3. ACE2 activator

Even though many potent ACE2 inhibitors have been successfully developed to date, apart from their use in investigating the structure-function of ACE2, the therapeutic application of ACE2 inhibitors is unclear. Based on the current *in vivo* physiological data, the activity of ACE2 is favourable in many pathophysiological conditions. Although the development of an ACE2 activator as potential therapeutic agent is desirable, the absence of a clear allosteric site on the enzyme presents a major obstacle. In the structure of ACE2, apart from the two conserved chloride binding sites, there is an absence of clear sites for binding of effector molecules. However, the ligand bound ACE2 assumed a very different conformation to that of the ligand-free ACE2 in the crystal structure (Towler *et al.*, 2004). In a search for potential ACE2 activators, Hernández Prada *et al.* took full advantage of this structural knowledge and identified conformation-specific sites, which are potentially accessible to small molecules, through a comparison of the ligand-bound and ligand-free ACE2 crystal structures (Hernandez Prada *et al.*, 2008). A total of three potential sites (Fig.1.8) were selected for the screening of a chemical compound library using molecular docking simulation. Only the search with the site at the hinge region of ligand-free ACE2 (Fig.1.8, site 1) yielded docked compounds with drug-like physiochemical

characteristics (Hernandez Prada *et al.*, 2008). As the hinge region provides a pivot point for large conformational movement of the sub-domains upon ligand binding, it was postulated that the binding of small effector molecules to the hinge region could induce the enzyme to assume a conformation that promotes binding of substrate or facilitated release of product.

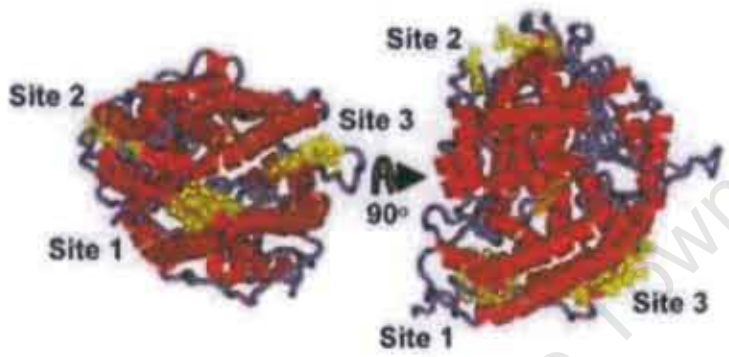

Fig.1.8. Relative positions of target sites (yellow) used in the search for small ACE2 effector molecules. The inhibitor-bound ACE2 (PDB code: 1R4L) is shown in ribbon representation (red). Site 1 was defined using the ligand-free conformation of ACE2 and superimposed to show its relative position to other sites. Figure take from (Hernandez Prada *et al.*, 2008).

Amongst the compounds which scored well in the screening process, two compounds, namely xanthenone (XNT) and resorcinolnaphthalein, were shown *in vitro* to specifically enhance ACE2 activity (Hernandez Prada *et al.*, 2008). Further *in vivo* testing of XNT demonstrated that the compound exerts antihypertensive effects such as lowering of blood pressure and potentiation of the hypotensive effect induced by bradykinin treatment. The enhanced ACE2 activity brought about by chronic XNT administration was also shown to reverse fibrosis of heart and kidney tissue of SHR (Hernandez Prada *et al.*, 2008). While ACE2 activating compounds such as XNT have opened a new avenue for development of novel treatment for hypertension, further research is essential to confirm the feasibility of such activators as therapeutic agents.

1.6. Expression regulation of ACE2

The flurry of studies carried out on ACE2 soon after its discovery owes much to the novelty of ACE2 as the latest addition to the RAS, and the role played by ACE2 as a functional receptor of the SARS CoV. While plenty of physiochemical and structural information about ACE2 has been gathered, the regulation of ACE2 expression remains poorly understood. The understanding of ACE2 expression regulation is likely to be beneficial in terms of the development of novel therapeutic intervention, as altered expression of ACE2 has been reported to be associated with different pathological conditions. During SARS-CoV infection, a down-regulation of ACE2 protein levels was observed in both *in vitro* and *in vivo* models upon the introduction of SARS-CoV Spike protein (Kuba *et al.*, 2005). Decrease in ACE2 mRNA and protein levels in the renal tubule has also been reported for a streptozocin-induced diabetic rat model (Tikellis *et al.*, 2003). Supporting evidence for this finding was provided by the detection of lowered ACE2 mRNA and protein levels in the kidney tissue of diabetic patients (Mizuiri *et al.*, 2008).

In 2005, Igase and colleagues reported that the blocking of AT_1 receptors with the antagonist olmesartan in 12-week old male SHR led to the detection of increased ACE2 mRNA levels in the thoracic aorta when compared to that of vehicle-treated animals (Igase *et al.*, 2005). The five-fold increase in ACE2 mRNA levels after treatment with olmesartan suggested that a regulatory effect on ACE2 expression is mediated through the AT_1 receptor. In a cell model of neonatal rat cardiomyocytes, aldosterone was found to reduce ACE2 mRNA expression and increase ACE mRNA, and this aldosterone induced altered expression of ACE2 and ACE was completely blocked by treatment with the aldosterone antagonist eplerenone (Yamamuro *et al.*, 2008). However, conflicting evidence for the role of aldosterone as a negative regulator of ACE2 expression has also been reported. In Dahl salt-sensitive hypertensive rats on a high sodium diet, treatment with AT_1 receptor blocker candesartan resulted in a significant increase of ACE2 mRNA in the heart, and this increase in ACE2 mRNA was not observed in animals treated with eplerenone (Takeda *et al.*, 2007).

Recently the mitogen activated protein (MAP) kinase/phosphatase pathway has been

shown as a mechanism through which AngII mediates its repressor effect on ACE2 expression, as treatment with the MAP kinase kinase inhibitor PD98059 attenuated the AngII mediated reduction of ACE2 gene expression in a rat vascular smooth muscle cell model (Gallagher *et al.*, 2008). While it appears that expression of ACE2 is modulated by complex interplay between various hormone peptides, little is known about the factors involved at the molecular level. In a search for gene target of the transcription factor hepatocyte nuclear factor 1 beta (HNF1β), Senkel and colleagues identified ACE2 as one of the direct target genes of HNF1β in HEK293 cells (Senkel *et al.*, 2005). A total of eight potential HNF1 binding sites were identified in the 1.1 kb region immediately upstream of the translation start site, and 5' deletion ACE2 promoter constructs with the removal of HNF1 site farthest from the translation start codon yielded a significantly reduced activation of promoter by HNF1β overexpression (Senkel *et al.*, 2005). Although the study did not pursue the implication of ACE2 gene activation by HNF1β, nevertheless it described one of the first transcription-modulating elements of the ACE2 gene.

The HNF1 binding sites reported by Senkel *et al.* were initially thought to be located in the promoter region of the ACE2 gene. However, a later study by Itoyama *et al.* revealed the presence of a new extended region of exon 1, as well as an untranslated exon (Itoyama *et al.*, 2005). The finding of a 5' untranslated exon placed the HNF1 binding sites within a region immediately upstream of the ACE2 translation start codon that does not form part of the promoter region. As precedents of *cis*-acting elements within intron 1 regions that regulate transcription activities have been reported for a variety of genes (LeBlanc *et al.*, 2006; Simon *et al.*, 2008; Girousse *et al.*, 2009), the findings documented by *Senkel et al.* have also suggested the involvement of intron 1 of the ACE2 gene in its transcription regulation.

1.7. Aim of study

As highlighted in the literature, ACE2 activity is beneficial in several pathological conditions. However, the understanding of the control of ACE2 activity at the protein-structure and expression level is incomplete. Further understanding of both the structure-function of ACE2 as well as its transcription regulation would prove immensely invaluable for the conceptualization of novel treatment for various pathological conditions. In order to gain further knowledge of the control of ACE2 activity, the two main aims devised for this study were:

- To investigate the interactions between ACE2 active site residues and the P_2 side chain of a bound ligand, and determine the structure-function relationship of these interactions.
- To elucidate potential transcription factors involved in the regulation of ACE2 expression.

To meet these aims, the objectives of this thesis were as follows:

- Characterization of the inhibitory parameters of a novel phosphinic ACE2 inhibitor (416F2).
- *In silico* molecular docking simulation of 416F2 into the ligand-bound crystal structure of ACE2.
- Identification of interactions between the P_2-leucyl group of 416F2 and ACE2 residues and assess the structure-function of identified interactions by site-directed mutagenesis.
- Screening of the region immediate upstream of the ACE2 translation start site for potential binding sites of putative transcription factors using transcription activity assays.
- Identification of putative transcription binding sites using a bioinformatics approach.
- Determination of the *in vitro* binding of putative transcription regulators of ACE2 by electro-mobility gel shift assay.

Chapter 2

Kinetic analyses of a novel ACE2 inhibitor

2.1. Introduction

As a recently discovered component of the renin-angiotensin system, the knowledge of ACE2 substrate selectivity remains incomplete. A previous biochemical investigation into the biological substrates of ACE2 has resulted in the postulation of a consensus sequence for ACE2 substrates (Vickers *et al.*, 2002). The consensus substrate sequence of ACE2 has helped the development of potent substrate-based inhibitors (Dales *et al.*, 2002). In a recent ACE2 inhibitor study, it was shown that tripeptide inhibitors with P_2 residues conferred greater inhibitory activities than did dipeptide inhibitors with only P_1 and P_1' groups (Mores *et al.*, 2008). These tripeptide inhibitors present themselves as potential probes to investigate the S_2 sub-site structure-function of ACE2.

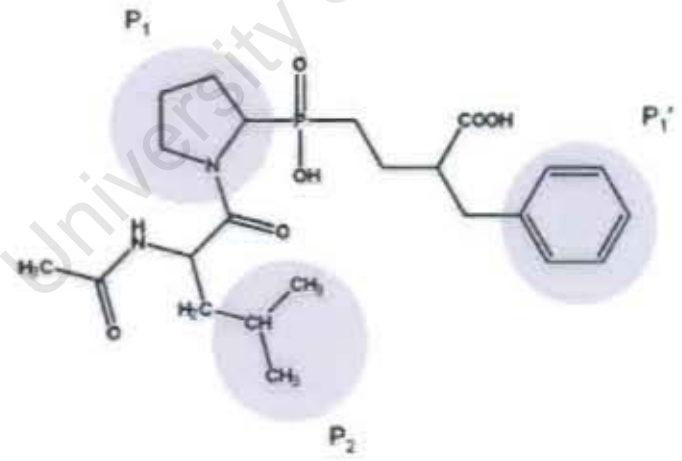

Fig.2.1. Chemical structure of the novel phosphinic tripeptide inhibitor 416F2. The tripeptide inhibitor mimics an amino acid sequence of Leu-Pro-Phe. The various functional groups are shown. The structure was generously provided by V. Dive, SIMOPRO, France.

In this study, the main objective was to assess the ACE2 inhibitory activity of the novel phosphinic tripeptide inhibitor 416F2 (Fig.2.1) using different enzyme kinetic

analyses. Many previous studies have successfully employed soluble ACE2 for enzyme activity analysis because of its retained functional integrity and ease to isolate (Guy *et al.*, 2005; Tipnis *et al.*, 2000; Vickers *et al.*, 2002). Similarly, soluble ACE2 was utilized in this study. ACE2 inhibitory activity of 416F2 was assessed using dose-dependent response curve and direct linear plot methods. The use of direct linear plot method also allowed for determination of the inhibition-type exhibited by 416F2 (Cornish-Bowden, 1995; Eisenthal and Cornish-Bowden, 1974)

2.2. Materials and methods

2.2.1. Propagation of pCI-NEO-sACE2FLAG

The DNA construct pCI-NEO-sACE2FLAG is an ACE2 cDNA fragment inserted into the mammalian expression vector plasmid pCI-NEO (Promega Corp.) at the *Xho*I and *Xba*I sites (a generous gift from D.W. Lambert and N.M. Hooper, University of Leeds). The C-terminal transmembrane and cytoplasmic regions of ACE2 were replaced by a FLAG tag in the sACE2FLAG construct (Guy *et al.*, 2005). In addition, the region coding for the N-terminal leader peptide sequence was also removed (Fig.2.2).

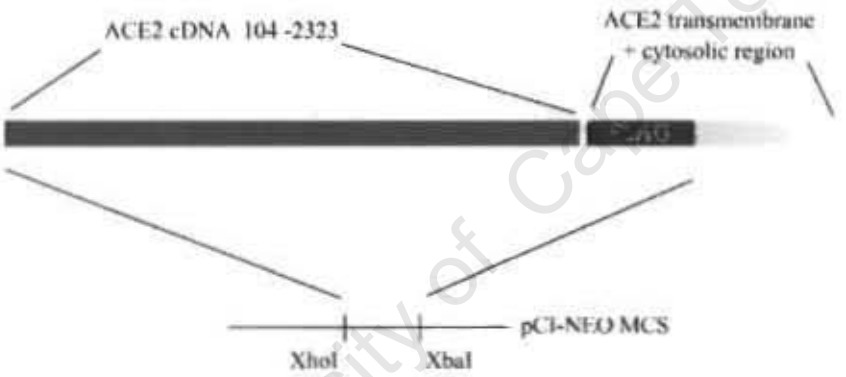

Fig.2.2. DNA construct of sACE2FLAG. The ACE2 transmembrane and C-terminal cytosolic regions of ACE were replaced by a FLAG epitope tag. The ACE2 cDNA sequence 104 to 2323 codes for the amino acid residues 1 to 740.

The plasmid pCI-NEO-sACE2FLAG was transformed into *E.coli DH5α* (see appendix A.1 for bacteria culturing method) using the standard heat shock transformation method and plated onto ampicillin-containing agar plates (see appendix A.4) (Ausubel *et al.*, 1992). Heat shock competent *E.coli DH5α* was prepared as described in appendix A.3. Ampicillin-resistant transformants were subjected to restriction enzyme analysis using the *Xho*I and *Xba*I enzymes, followed by electrophoresis on 1% (w/v) agarose gel in 1× Tris borate EDTA buffer (TBE) (see appendix A.5). Large-scale purification of pCI-NEO-sACE2FLAG plasmid DNA was carried out using the HiSpeed® Plasmid Kit (Qiagen), following the protocol provided by the manufacturer.

2.2.2. sACE2FLAG transfection and isolation

2.2.2.1. Cell culture transfection

Human embryonic kidney epithelial cells (HEK293; a generous gift from the UCT/MRC Research Group for Receptor Biology, University of Cape Town) were cultured in Dulbecco's Modified Eagle Medium (DMEM; Highveld Biological) containing 10% (v/v) foetal calf serum (FCS; Gibco®, Invitrogen) and a final concentration of 50 mM HEPES, pH 7.5. Prior to transfection, cell cultures were grown until approximately 60% confluent in 6-well tissue culture plates (Nunc Corp.). An amount of 1.5 µg of plasmid pCI-NEO-sACE2FLAG was transfected into HEK293 cell culture using the Profection Transfection Kit (Promega Corp.) following the protocol provided by the manufacturer. Briefly, fresh growth medium was placed onto cell cultures 3 hours prior to the transfection, and the calcium chloride-DNA complex was prepared by mixing plasmid DNA with 18 µl of 2 M calcium chloride (Profection Transfection Kit; Promega Corp.) and sterile nuclease-free water to a final volume of 150 µl. The plasmid DNA mixture was bubbled though equal volumes of 2 × HEPES buffered saline (HBS) (Profection Transfection Kit; Promega Corp.). This final mixture was incubated at room temperature for 30 min and added directly to cell cultures in a drop-wise fashion with gentle swirling of culturing plates. After 4 hours of incubation at 37°C, a glycerol shock procedure was performed on the transfected cells. Transfected cells were rinsed once with 1 × phosphate buffered saline (PBS) (137 mM NaCl, pH 7.4; 2.7 mM KCl; 20 mM Na_2HPO_4; 2 mM KH_2PO_4) followed by the application of 500 µl of sterile 15% (v/v) glycerol in PBS. The cell cultures were exposed to 15% glycerol solution for a period no longer than 2 min and cell cultures were rinsed twice with 1 × PBS solution to remove all traces of glycerol. The cell cultures were allowed to recover for 24 hours in fresh culture medium after the glycerol shock procedure. Selection for stable sACE2FLAG expressing cells was achieved by culturing the transfected cells in growth medium containing 0.8 mg/ml of the antibiotic G418 (Sigma-Aldrich, Inc.).

2.2.2.2. FLAG-tagged protein purification

Stable pCI-NEO-sACE2FLAG transfected HEK293 cells were grown until approximately 80% confluent in two T150 cell culture flasks (Nunc Corp.) followed by incubation in 25 ml of DMEM growth medium containing 2% (v/v) FCS. The culture medium containing secreted sACE2FLAG protein was collected after an incubation period of approximately 48 hours at 37°C. The harvest process was repeated twice.

The collected culture medium was centrifuged at 1800 × *g* for 5 min at 4°C to remove the cells from the media. Soluble ACE2FLAG was purified by affinity column chromatography using the ANTI-FLAG M2® resin (Sigma-Aldrich, Inc.). Binding of sACE2FLAG to the column was achieved by passing the collected media through the column twice at 4°C. The column was then washed with 20 column volumes of ice-cold Tris buffered saline (TBS) (50 mM Tris, pH 7.5; 150 mM NaCl). The sACE2FLAG was eluted with ten 1-ml aliquots of elution buffer (0.1 M glycine HCl, pH3.5). The purification process was monitored by assaying eluted fractions for ACE2 activity as described in 2.2.3.1 using the FRET-based synthetic substrate Mca-Ala-Pro-Lys-Dnp-OH (Mca-APK-Dnp; AnaSpec, Inc.). The collected fractions containing active sACE2FLAG were pooled and concentrated using Amicon® Ultra centrifugal 30 kDa cut-off filter devices (Millipore Corp.). The amount of sACE2FLAG isolated was determined by the standard Bradford assay (Bradford, 1976) using Bio-Rad protein assay reagent (Bio-Rad Laboratories). The purity of the isolated sACE2FLAG was monitored by SDS PAGE (see appendix A.8) using a 10% gel (Ausubel *et al.*, 1992). The purified sACE2FLAG was stored at -20°C.

2.2.3. ACE2 activity determination

2.2.3.1. ACE2 activity fluorogenic assay

The activities of ACE2 proteins were determined using synthetic fluorogenic ACE2 substrates Mca-APK-Dnp and Mca-Tyr-Val-Ala-Asp-Ala-Pro-Lys-Dnp-OH (Mca-YVADAPK-Dnp; R&D Systems, Inc.). Since the substrates were dissolved in DMSO,

the ACE2 activity assay was carried out such that less than 0.7% (v/v) DMSO is present in the final reaction mixture (Vickers *et al.*, 2002). Activity assays were carried out using U96- PP plates (Nunc Corp.) in a total reaction volume of 100 µl at 25°C. The reaction mixture contained final concentrations of 100 mM Tris, pH 7.0 and 300 mM NaCl. The reaction product formation was measured by the increase in fluorescence detected at excitation and emission wavelengths of 320 nm and 430 nm respectively, using a Cary Eclipse fluorescence spectrophotometer (Varian, Inc.). For continuous assays, the increase in fluorescence signal was monitored over a time period of 60 min.

2.2.3.2. Construction of fluorogenic substrate standard curves

The amount of reaction product formed was determined using a standard curve generated with 7-methxoycoumarin-4-acetic acid (Mca; Sigma-Aldrich, Inc.). Standard curves were generated by plotting the fluorescence signals measured against varied amounts of Mca (0 nmol to 15 nmol). In addition, standard curves were also generated by measuring the fluorescence released by completely hydrolyzed fluorogenic substrate Mca-APK-Dnp (AnaSpec, Inc.) of amounts ranging from 0 to 15 nmol. The complete hydrolysis of Mca-APK-Dnp was achieved by incubation of the substrate with 0.35 µg purified sACE2FLAG at 25°C. Complete hydrolysis of the substrate was assessed by comparison of the fluorescence intensities (F.i) recorded after 2 hours and 3 hours of incubation. The fluorescence intensities of the standards were measured as described in 2.2.3.1 in a total volume of 100 µl containing 100 mM Tris, pH 7.0 and 300 mM NaCl. Similarly, a standard curve of completely hydrolyzed Mca-YVADAPK-Dnp was also generated.

2.2.4. Kinetic characterization of the ACE2 inhibitor 416F2

While the isolated sACE2FLAG protein was utilized in the standard curve construction (see 2.2.3.2), another soluble ACE2 (a generous gift from J. Penninger, Apeiron Biologics; hereafter abbreviated as ACE2ab) was used to facilitate subsequent enzyme kinetic experiments, as the isolated sACE2FLAG stock was lost due to storage problems. The ACE2ab protein is identical to sACE2FLAG but lacking the FLAG tag. Enzyme inhibitor complexes were formed by incubation of ACE2ab, at

an amount of 5.27 pmol (final concentration of 65.11 nM), in ACE2 activity assay buffer (100 mM Tris, pH 7.0; 300 mM NaCl) with various concentrations of 416F2 (a generous gift from V. Dive, SIMOPRO, France) at room temperature for 1 hour. A volume of 3 µl of this mixture was used for the ACE2 fluorogenic assay (described in 2.2.3.1) containing various amounts of synthetic ACE2 substrates. The Mca-APK-Dnp substrate was used in inhibition assays analyzed by the dose-dependent response curve method, and a less expensive substrate Mca-YVADAPK-Dnp was used for the inhibition assays analyzed by the direct linear plot method.

In order to establish the initial inhibitor concentration range required for determination of the inhibition constant (K_i), inhibition assays were first performed using 0 to 1 mM inhibitor concentration with a Mca-APK-Dnp substrate concentration of 20 µM. An inhibitor concentration range of 10 to 100 nM was used for accurate determination of the inhibition constant. Mca-APK-Dnp substrate concentrations of 20 µM and 5 µM were chosen for the initial determination of IC_{50} values. IC_{50}'s were calculated from a plot of initial velocities (v) against a log scale of inhibitor concentrations using the dose-dependent response curve function of the software Prism4 (GraphPad Software). v was defined as the pmol of Mca-AP formed per min with less than 15% complete hydrolysis. The inhibition constant was determined from the IC_{50} value using equation 2.1 (Cheng and Prusoff, 1973) with the reported K_M value of 147±0.7 µM for Mca-APK-Dnp (Vickers *et al.*, 2002).

$$K_i = \frac{IC_{50}}{1 + \frac{[\text{substrate}]}{K_M}}$$

equation 2.1.

The IC_{50} value was confirmed by repeating the inhibition assay using the same inhibitor concentration range with Mca-APK-Dnp substrate concentrations of 5, 8, and 10 µM and K_i values were calculated using equation 2.1 as before.

The K_i value determined using the dose-dependent response curve was confirmed using the direct linear plot analysis (Cornish-Bowden, 1995; Eisenthal and Cornish-Bowden, 1974). Inhibition assays were carried out using various concentrations of 416F2 (concentration ranging from 0 to 40 nM) and Mca-YVADAPK-Dnp substrate (concentration ranging from 5 to 20 μM). The apparent Michaelis constant (K_M^{app}) and apparent maximum velocity (V_{max}^{app}) values of ACE2 for Mca-YVADAPK-Dnp were estimated by the direct linear plot method in the presence of 0, 10, 20 and 40 nM 416F2 (Cornish-Bowden, 1995; Cornish-Bowden and Eisenthal, 1978; Eisenthal and Cornish-Bowden, 1974). Briefly, estimation of K_M^{app} and V_{max}^{app} values by the direct linear plot method involved plotting of the v (v defined as pmol of Mca-YVADAP formed per min with less than 15% complete hydrolysis) as the Y-intercept and the negative substrate concentration value as the X-intercept. A series of such intercepts representing different initial velocities in the presence of different substrate concentrations were plotted and lines were fitted to the intercepts. The intersections of the various lines were noted, and the co-ordinate of median was taken as the estimated K_M^{app} and V_{max}^{app} values (Fig.2.3).

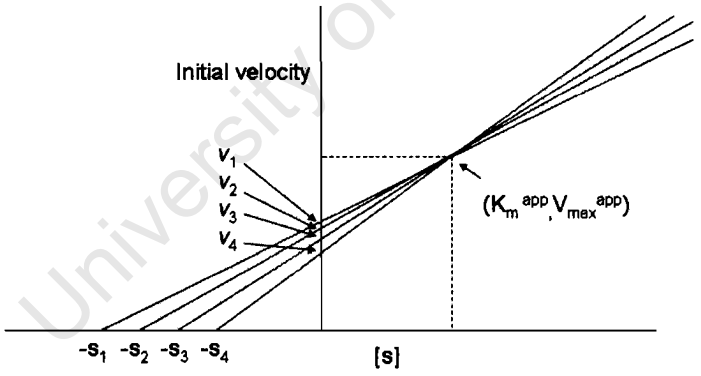

Fig.2.3. The direct linear plot. Each observation of initial velocity at a substrate concentration is plotted as Y- and X- intercepts respectively. Given the K_M and V_{max} values are constants for any specific enzyme hydrolyzing a particular substrate, all extrapolations of observed initial velocities at specific substrate concentrations would pass through the point representing the estimated apparent K_M^{app} and V_{max}^{app}.

The type of inhibition displayed by 416F2 was determined by comparing K_M^{app} and V_{max}^{app} values in the presence of inhibitor to those determined in the absence of inhibitor. Following the determination of the inhibition type, the K_i value of 416F2 for ACE2 was calculated using the following equations:

$$V_{max}^{app} = \frac{V_{max}}{1 + [I]/K_{iu}} \qquad V_{max}^{app}/K_M^{app} = \frac{V_{max}/K_M}{1 + [I]/K_{ic}}$$

$\therefore 1/V_{max}^{ap} (V_{max} = 1/K_{iu}([I]) + 1$ $\therefore (K_M^{app}/V_{max}^{app})(V_{max}/K_M) = 1/K_{ic}([I]) + 1$

if $1/V_{max}^{app} = 0$ if $K_m^{app}/V_{,max}^{app} = 0$

$-1 = 1/K_{iu}([I])$ $-1 = 1/K_{ic}([I])$

$-[I] = K_{iu}$ $-[I] = K_{ic}$

equation 2.2 equation 2.3

The inhibition constant estimated by the direct linear plot consisted of two components (K_{iu} and K_{ic}). K_{iu} and K_{ic} values were estimated by plotting $1/ V_{max}^{app}$ and K_M^{app}/ V_{max}^{app} against inhibitor concentrations [I] respectively. The K_{ic} value predicts binding of inhibitor to hinder the substrate binding process, and the K_{iu} value predicts binding of inhibitor to inhibit catalysis (Cornish-Bowden, 1995). The K_{ic} value calculated was compared to the K_i value calculated from the IC_{50} value determined by the dose-dependent response curve.

2.3. Results

2.3.1. Propagation of pCI-NEO-sACE2FLAG

The sACE2FLAG plasmid was propagated in *E.coli DH5α*. The positive transformants were screened by restriction enzyme digestion using enzymes *Xho*I and *Xba*I (Fig.2.4). The DNA fragment that migrated at a molecular weight between 3530 and 2027 bp corresponds to the 2205 bp of sACE2FLAG construct. Clone number 1 was used for large-scale DNA isolation and subsequent HEK293 cell transfections.

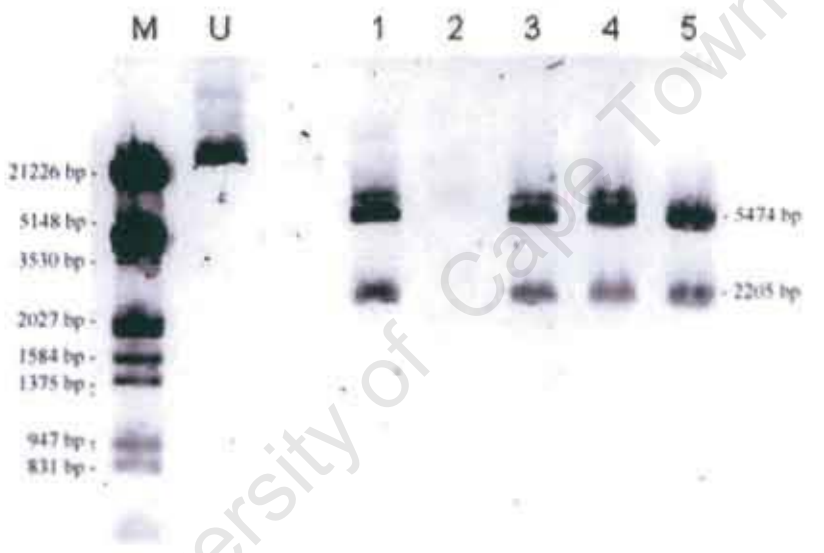

Fig.2.4. Plasmid sACE2FLAG transformed into *E.coli DH5α*. **Lane M:** DNA marker *EcoRI/Hind*III digested λ DNA. **Lane U:** undigested plasmid sACE2FLAG. **Lane 1, 3, 4 & 5:** positive clones carrying the correct sACE2FLAG construct. The expected size of the sACE2FLAG was 2205 bp.

2.3.2. Expression and purification of sACE2FLAG

Successful isolation of soluble ACE2 from transfected HEK293 has been described previously (Guy *et al.*, 2005; Guy *et al.*, 2005; Lambert *et al.*, 2005; Lambert *et al.*, 2008). Taken together with the consideration of proper processing of the enzyme by human cell lines, the expression of sACE2FLAG was carried out in the same protein expression system.

In order to confirm active sACE2FLAG protein was isolated by affinity anti-FLAG-tag chromatography, the presence of ACE2 in each eluted fractions was assessed using the ACE2 activity fluorogenic assay described in section 2.2.3.1 (Fig.2.5.A). The activity profile of the purification process indicated that not all sACE2FLAG expressed into the culture medium was isolated by the chromatography, as residual ACE2 activity was detected in the flow through fraction (Fig.2.5.A). The bulk of the eluted sACE2FLAG was found in fractions 6 to 9, as indicated by the ACE2 activity profile. These fractions were pooled and concentrated to a concentration of 68 µg/ml.

The purity of the isolated sACE2FLAG was assessed by SDS-PAGE (Fig.2.5.B). The single band travelling at a size corresponding to approximately 100 kDa represents the purified sACE2FLAG protein. The observed molecular weight of approximately 100 kDa is greater than that of the expected 85 kDa (predicted from the amino acid sequence of sACE2FLAG using PAWS, ProteoMetrics Software), indicating glycosylation of the purified sACE2FLAG protein. The presence of this single band was indicative of the isolated sACE2FLAG being free of contaminants and degradation (Fig.2.5.B).

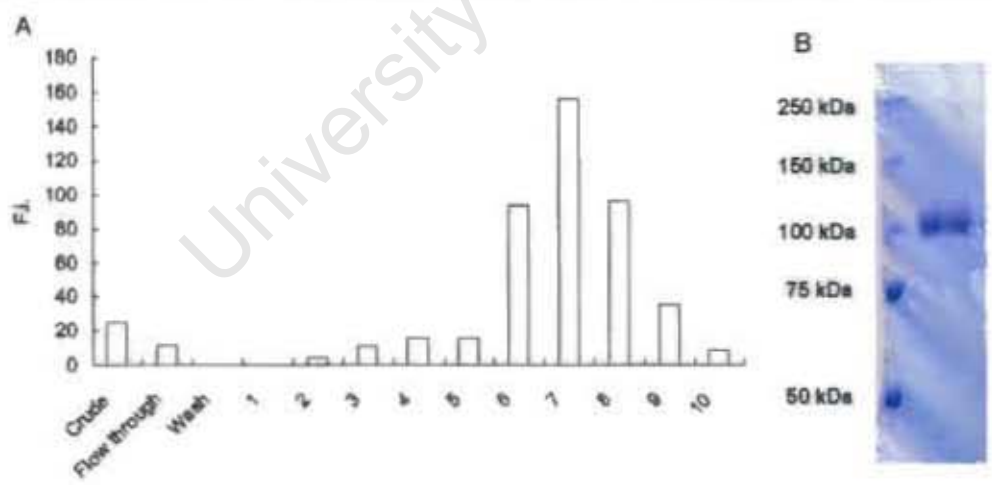

Fig.2.5. Purification of sACE2FLAG using anti-FLAG affinity chromatography. **A.** The ACE2 activity profile determined with 10 µl of each of the culture medium loaded onto column (Crude), the void volume (Flow through), the TBS wash buffer past through column (Wash) and the eluted fractions (1-10). ACE2 activity was recorded as increase in F.i. **B.** Purity of the isolated sACE2FLAG was assessed by fractionation of concentrated pooled fraction (30 × the concentration of eluted fractions) by SDS PAGE using a 10% gel.

2.3.3. Establishment of standard curves for the ACE2 activity assay

In order to quantify product formation in the ACE2 activity assay, standard curves were constructed with the fluorophores Mca and Mca-AP. The Mca group attached to the synthetic ACE2 substrates is the fluorophore responsible for the fluorescence at 430 nm when excited with light of wavelength 320 nm, thus a standard curve was generated using pure Mca (Fig.2.6.A). In addition, standard curves were generated by using completely hydrolyzed Mca-APK-Dnp substrate (Fig.2.6.B). The complete hydrolysis of Mca-APK-Dnp was achieved after 2 hours of incubation as further incubation of the substrate with purified sACE2FLAG did not result in further increase in fluorescence signal (data not shown).

A substantial difference between the standard curves was observed (Fig.2.6.A and B). At lower substrate amounts, the fluorescence signals released by completely hydrolyzed Mca-APK-Dnp increased in a linear fashion and the fluorescence intensities detected were comparable to those of the Mca standards (0 to 1.2 nmol). At higher substrate amounts (5 to 10 nmol), the standard curve generated with completely hydrolyzed Mca-APK-Dnp failed to increase further and the linearity of the standard curve was lost (Fig.2.6.B). Although peptide cleavage by ACE2 frees the fluorophore Mca from the 4,7-dinitrophenol (Dnp) fluorescence quencher, inter-molecular quenching of the fluorophore occurs at high concentrations of freed Dnp quencher (Phu *et al.*, 2005). This inter-molecular quenching explains the observed distortion of the completely hydrolyzed Mca-APK-Dnp standard curve. To circumvent this inter-molecular fluorescence quenching, all subsequent fluorogenic assays were conducted using substrate amounts of less than 2 nmol (equivalent to a concentration of 20 µM in a 100 µl total reaction volume). Substrate concentrations for subsequent inhibition assays were adjusted such that the amount of product formed at 15% complete hydrolysis does not exceed 2 nmol.

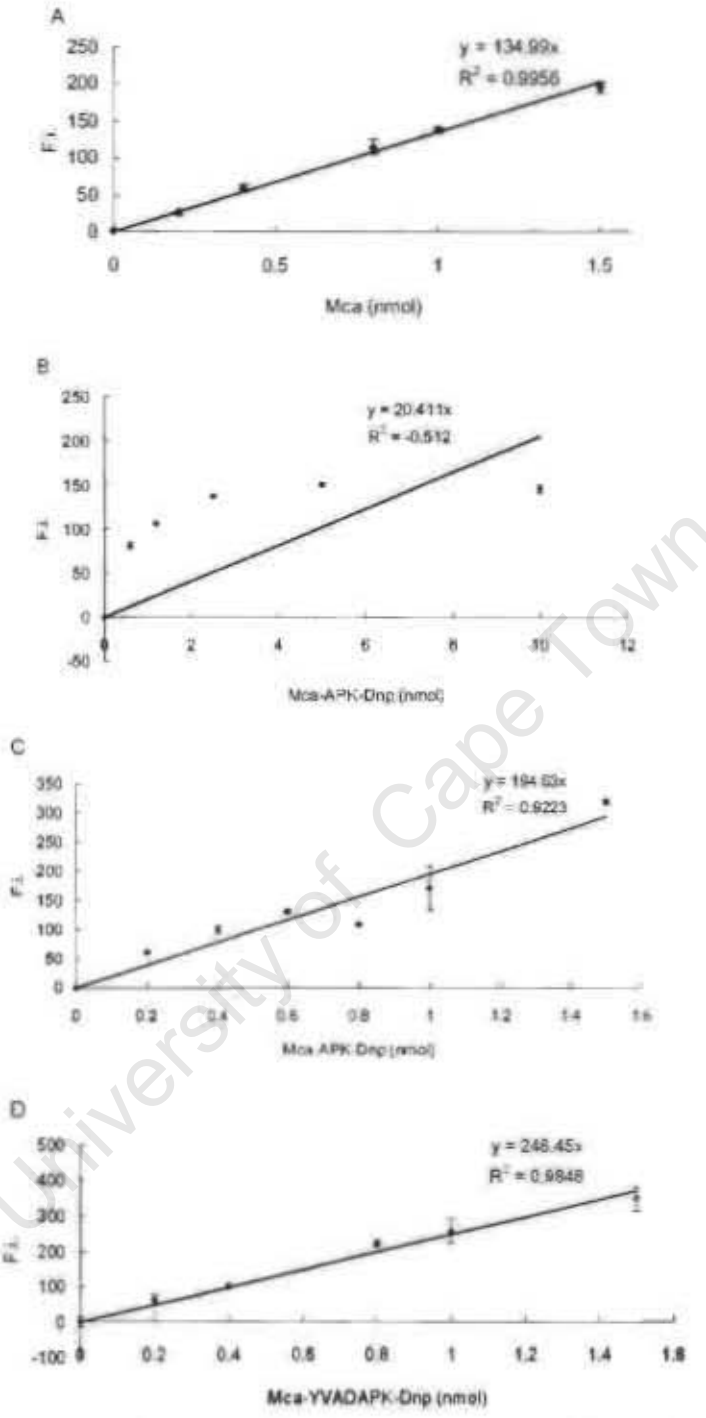

Fig.2.6. Standard curves for the quantification of Mca-product formation. **A.** Standard curve of pure Mca and **B.** Standard curve of completely hydrolyzed Mca-APK-Dnp substrate at amounts ranging from 0 to 10 nmol. **C.** Standard curve of completely hydrolyzed Mca-APK-Dnp substrate and **D.** Standard curve of completely hydrolyzed Mca-YVADAPK-Dnp substrate at amounts ranging from 0 to 1.5 nmol. A higher sensitivity setting of the spectrophotometer was selected in measuring f.i. used in the construction of standard curves shown in C and D. All readings were normalised using fluorescence signals detected for 0 nmol standard. Each data point represents a mean ± S.D. of F.i. determined in triplicate.

The standard curve of completely hydrolyzed Mca-APK-Dnp was subsequently generated with Mca-APK-Dnp quantities below 2 nmol (Fig.2.6.C) as intermolecular quenching occurs at higher values (Fig.2.6.B). In order to accurately detect the fluorescence intensities from such low amounts of Mca-APK-Dnp, a high photomultiplier tube (PMT) voltage detection sensitivity setting of the spectrophotometer was selected in measuring these standards. Hence, higher fluorescence intensity signals were recorded than those used in the construction of the previous Mca-APK-Dnp standard curve at corresponding substrate amounts (Fig.2.6.B and C). Similarly, a completely hydrolyzed Mca-YVADAPK-Dnp standard curve was also constructed using the same quantity of substrates (Fig.2.6.D). At the high PMT voltage sensitivity, the pure Mca standard curve had a detection range of 0 to 2 nmol (data not shown). The Mca standard curve was not used because its slope was 5-fold greater than those of completely hydrolyzed Mca-APK-Dnp and Mca-YVADAPK-Dnp. The construction of the various standard curves allowed for the quantification of product formed and the determination of initial reaction velocities for the subsequent inhibition assays.

2.3.4. Kinetic characterization of ACE2 inhibition by 416F2

The inhibition of ACE2 by 416F2 was assessed using two enzyme kinetic analyses. A broad inhibition range was first determined with the synthetic substrate Mca-APK-Dnp, and the inhibition profile of 416F2 at a concentration range of 0 to 1000 µM is shown in Fig.2.7. It was clear that 416F2 achieved near complete inhibition at a concentration of 100 nM, thus the range between 10 and 100 nM (10, 20, 30, 40, 50, 70 and 100 nM) was used for the inhibition constant determination.

An estimation of the IC_{50} value of 416F2 on ACE2 activity was determined using two different substrate concentrations (20 and 5 µM). The amount of product formed per min by the incubation of Mca-APK-Dnp with ACE2ab (final concentration of approximately 2 nM) in the presence of 416F2 (at concentrations of 10, 20, 30, 40, 50, 70 and 100 nM) was fitted to a non-linear dose-dependent response curve, and the IC_{50} was calculated (Fig.2.8). While an IC_{50} value was determined using inhibition assays carried out with 20 µM of Mca-APK-Dnp, the dose-dependent response curve of inhibition assays carried out with 5 µM substrate concentration failed to converge

and therefore an IC_{50} could not be determined (Fig.2.8). The estimated K_i of 416F2 calculated from the IC_{50} value determined for inhibition assays carried out with 20 μM substrate is shown in Table 2.1.

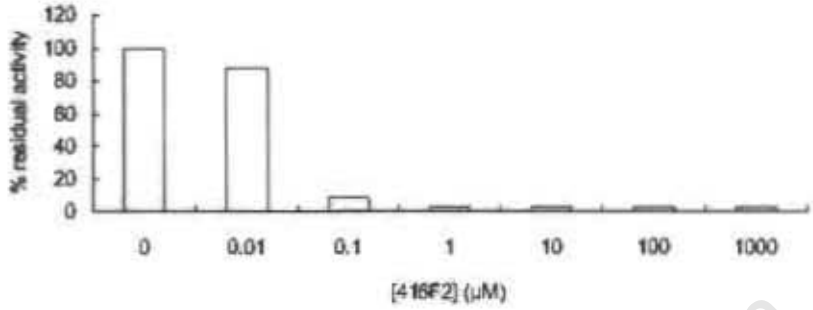

Fig.2.7. Broad concentration range inhibition profile of phosphinic inhibitor 416F2 on ACE2 activity. ACE2 activity was assessed as the increase in F.I. measured after 60 min incubation with the substrate Mca-APK-Dnp. Near complete inhibition of ACE2 was achieved at an inhibitor concentration of 100 nM.

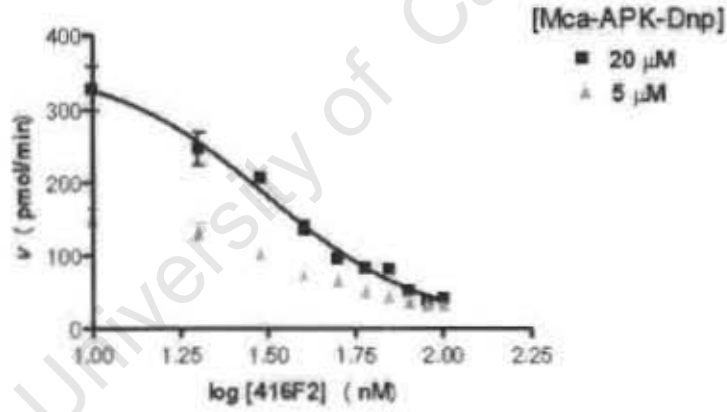

Fig.2.8. Non-linear dose-dependent sigmoidal curve of 416F2 inhibition on ACE2 activity with Mca-APK-Dnp. The inhibition profile was generated with substrate concentrations of 20 and 5 μM. Dose-dependent response curve of the inhibition profile generated with 5 μM failed to converge. Each data point represents a mean ± S.D. of inhibition assays performed in triplicate.

As IC_{50} value was only successfully determined for inhibition assays performed with 20 μM Mca-APK-Dnp, the estimated K_i value was further verified by repeating the dose-dependent response analysis using substrate concentrations of 5, 8 and 10 μM with the same inhibitor concentrations (Fig.2.9.). The K_i values calculated using IC_{50} values determined in the presence of the different substrate concentrations were in

agreement with each other and an average Ki value of 27.91±5.59 was determined (Table 2.1).

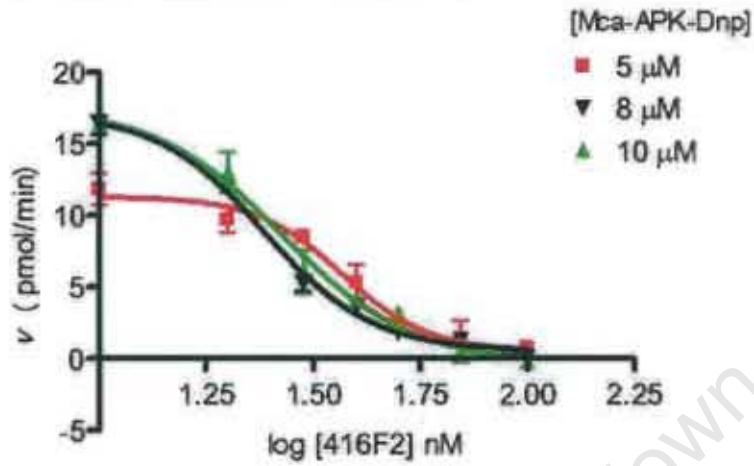

Fig.2.9. Non-linear dose-dependent sigmoidal curve of 416F2 inhibition on ACE2 activity with Mca-APK-Dnp. The inhibition profile was generated with substrate concentrations of 5, 8 and 10 µM. Each data point represents a mean ± S.D. of inhibition assays performed in triplicate.

Table 2.1. Inhibition constants of 416F2 inhibition of ACE2 activity with Mca-APK-Dnp, calculated from IC_{50} values determined at a range of Mca-APK-Dnp concentrations.

Substrate conc. (µM)	**IC_{50} (nM)**	**K_i (nM)**
5	37.03	35.82
8	24.32	23.06
10	26.86	25.15
20	31.38	27.62
	Avg. K_i	**27.91±5.59**

In addition to the dose-dependent response curve method, K_i value of 416F2 inhibition was also determined using the direct linear plot with the synthetic substrate Mca-YVADAPK-Dnp (Cornish-Bowden, 1995; Eisenthal and Cornish-Bowden, 1974). The direct linear plot is a simple method to estimate enzyme kinetic parameters. The construction of direct linear plots requires no mathematical transformation of the raw data, and is hence free of any bias that incurs through data transformation. In addition to kinetic parameter determinations, the direct linear plot analysis is also able to distinguish between different types of inhibition displayed by an inhibitor. A

representative direct linear plot is shown in Fig.2.10.

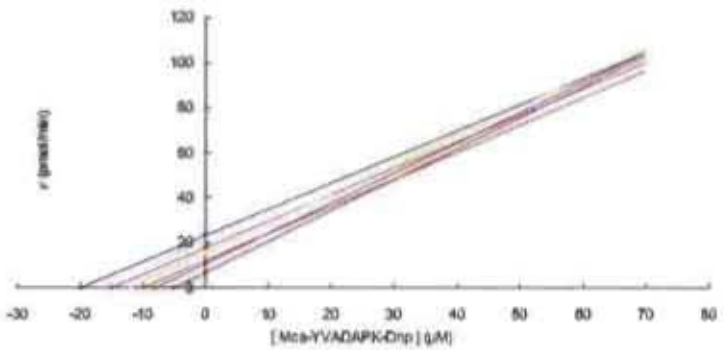

Fig.2.10. A representative direct linear plot of v (pmol/min) versus substrate concentration (μM). The linear lines were generated by utilizing corresponding v values and negative substrate concentrations (20, 15, 12, 10, 8 and 5 μM) as y- and x-intercepts respectively. Intersections of the straight lines were determined using equation of the extrapolated lines. Each intersection in the positive quadrant was considered as an estimate of observed V_{max} and K_M. The median value of the set of intersection was taken as the best estimates of observed V_{max}^{app} and K_M^{app}.

The K_i value is determined as the factor by which the K_M^{app} and V_{max}^{app} values observed in the absence of inhibition shift to new values under inhibitory conditions. Depending on the type of inhibition, observed K_M^{app} and V_{max}^{app} values are influenced differently, therefore this study first determined the type of inhibition displayed by 416F2. The plot of the observed K_M^{app} and V_{max}^{app} values on a set of K_M and V_{max} axis showed a shift of both K_M^{app} and V_{max}^{app} values without a constant V_{max}^{app}/ K_M^{app} value (see Table 2.2 and Fig.2.11), which was indicative of a mixed type of inhibition exhibited by 416F2 (Cornish-Bowden, 1995; Eisenthal and Cornish-Bowden, 1974).

Table 2.2. K_M^{app} and V_{max}^{app} values estimated for ACE2 activity with Mca-YVADAPK-Dnp in the presence of 416F2 concentrations ranging from 0 to 40 nM by direct linear plot

[416F2] (nM)	K_M^{app} (μM)	V_{max}^{app} (pmol/min)	V_{max}^{app}/ K_M^{app}
0	79.25±13.37	418.86±88.88	5.28
10	56.62±39.52	176.83±97.46	3.12
20	35.61±10.88	81.06±18.03	2.28
30	59.02±3.28	90.65±2.60	1.53
40	17.27±10.02	38.01±9.38	2.20

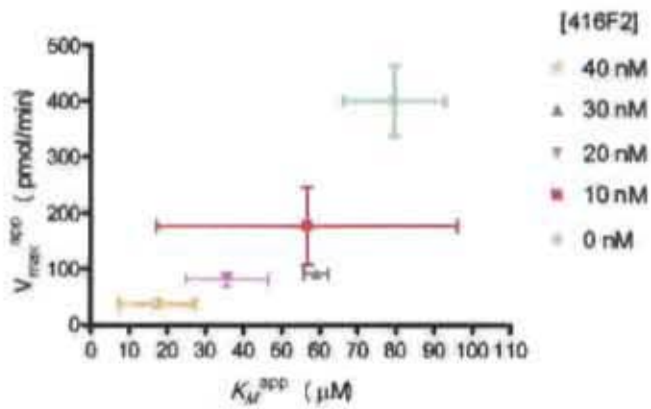

Fig.2.11. Plot of observed K_M^{app} and V_{max}^{app} values for ACE2 activity with Mca-YVADAPK-Dnp in the presence of various 416F2 concentrations. The shifting of K_M^{app} and V_{max}^{app} values without constant V_{max}^{app}/K_M^{app} value indicated a mixed type of inhibition exhibited by 416F2. Experiments were carried out in duplicate and the data represented by the mean with S.D.

Identification of a mixed type of inhibition displayed by 416F2 led to the choice of equations 2.2 and 2.3 for K_i calculation. Plots of $1/V_{max}^{app}$ and K_M^{app}/V_{max}^{app} against versus various inhibitor concentrations revealed a K_{iu} value of 2.436 nM and a K_{ic} value of 31.10 nM respectively (Fig.2.12). The K_{ic} value (31.1 nM), which signifies a competitive inhibition of ACE2 substrate binding by 416F2, was in agreement with the average K_i value (27.91±5.59 nM) calculated from the IC_{50} values determined by the dose-dependent response curve analysis (Table 2.3). However a significantly lower K_{iu} value (2.436 nM) was determined for 416F2 inhibition of ACE2. Taken together, these K_i values indicate 416F2 is a strong inhibitor of ACE2.

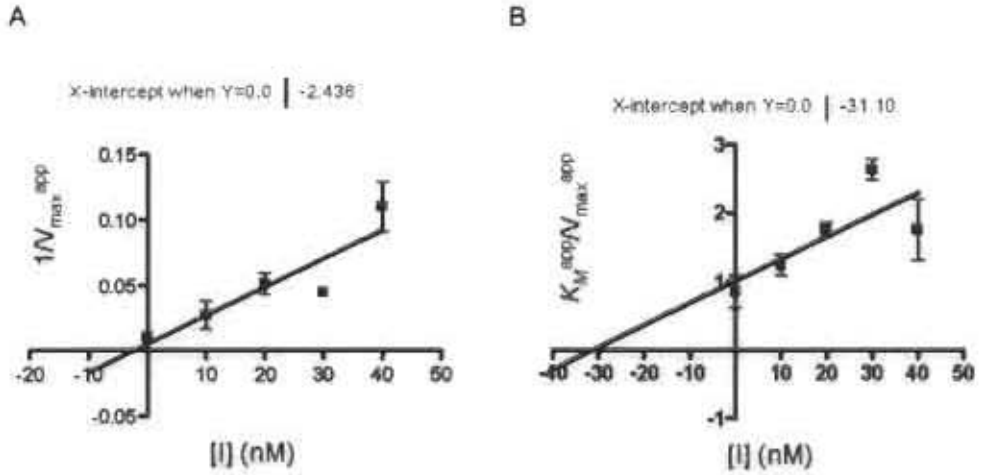

Fig.2.12. K_i value determination by direct linear plot (Cornish-Bowden, 1995; Eisenthal and Cornish-Bowden, 1974). **A.** K_{iu} value of 2.436 nM was determined from the x-intercept of a plot of $1/V_{max}^{app}$ versus inhibitor concentrations 0, 10, 20, 30 and 40 nM. **B.** K_{ic} value of 34.12 nM was determined from the x-intercept of a plot of K_M^{app}/V_{max}^{app} against inhibitor concentrations 0, 10, 20, 30 and 40 nM. Experiments were carried out in duplicate and the data represented by the mean with S.D.

Table 2.3. Summary of K_i values for 416F2 inhibition of ACE2 activity with substrates Mca-APK-Dnp and Mca-YVADAPK-Dnp, calculated using different enzyme kinetic methods.

Enzyme kinetics method	**Substrate**	**K_i (nM)**
Dose-dependent response curve	Mca-APK-Dnp	27.91±5.59 (Avg.)
Direct linear plot (K_{ic})	Mca-YVADAPK-Dnp	31.10
Direct linear plot (K_{iu})	Mca-YVADAPK-Dnp	2.44

2.4. Discussion

The inhibitory effect of the novel ACE2 inhibitor 416F2 was examined by two different enzyme kinetic analyses using a FRET-based fluorogenic assay with ACE2 synthetic substrates Mca-APK-Dnp and Mca-YVADAPK-Dnp. An average K_i value of 27.91±5.59 nM was determined for 416F2 inhibition of ACE2 activity with Mca-APK-Dnp using the dose-dependent response curve method, while the use of the direct linear plot method with substrate Mca-YVADAPK-Dnp gave a K_{ic} value of 31.10 nM and a K_{iu} value of 2.436 nM (Table 2.4). The direct linear plot analysis also led to the elucidation of a mixed type of inhibition displayed by 416F2 (Fig.2.11). This finding was unexpected, because the inhibitor design was based on the consensus sequence of ACE2 biological substrates (Fig.2.1), and was thus expected to inhibit ACE2 in a simple competitive fashion. Instead, the determination of the greater K_{ic} value than the K_{iu} value implies a mixed inhibition of predominantly uncompetitive inhibitory nature was displayed by 416F2. It was noted the large standard deviations associated with the estimated K_M^{app} and V_{max}^{app} values (Table 2.3) might have distorted the inhibition type analysis, and this could be resolved by verification with other inhibition type distinguishing enzyme kinetic analyses, such as the one proposed by Brandt *et al.* (Brandt *et al.*, 1987). However, consideration has also been made for the possibility of the small tripeptide-mimicking structure of 416F2 to bind elsewhere in the ACE2 structure to cause a conformational change that would render sub-optimal hydrolysis of bound substrate. This thinking is supported by the report which described the binding of small effector molecules to specific sites on the surface of ACE2 promotes ACE2 activity (Hernandez Prada *et al.*, 2008). Although it is unlikely for 416F2 to bind to the surface of ACE2, however it is possible for 416F2 to bind active site sub-sites distal to the catalytic zinc and promote a conformational change in the active site. Nevertheless, the K_{ic} estimated by the direct linear plot method was in agreement with the K_i value calculated from IC_{50}'s determined by dose-dependent response curves (Table 2.3), and the low nanomolar K_i values were indicative of strong binding of 416F2 to free ACE2.

Interestingly, an ACE2 inhibition constant of 0.13 nM reported for 416F2 by a separate study (Mores *et al.*, 2008) was significantly lower than the K_i values determined in this study. The discrepancy could be attributed to the different buffer

systems used for the inhibition assays. In this study, the ACE2 inhibition assays were carried out in final concentrations of 100 mM Tris, 300 mM NaCl, pH 7.0 while the ACE2 inhibitory assays reported by Mores and colleagues were conducted in final concentrations of 50 mM HEPES, 1 M NaCl, 10 μM $ZnCl_2$, pH 6.8 (Mores *et al.*, 2008). Particularly, the difference in chloride concentrations of the assay buffers would lead to very different observed inhibition of ACE2 by the same compound. ACE2 inhibition has been reported to be a chloride sensitive event where increased inhibitor potency was found at higher chloride concentrations (Rushworth *et al.*, 2008). While it is unlikely that a 3-fold increase in chloride concentration would translate into a 100-fold increase of inhibitor binding to ACE2, other factors such as the difference in pH of the assay buffers and the use of different enzyme kinetic analysis methods (Mores *et al.* performed an enzyme kinetic analysis which considers the occupancy of total enzyme by different concentrations of inhibitor for K_i determination (Horovitz and Levitzki, 1987)) could contribute towards the difference in the observed K_i values.

The low nanomolar inhibition displayed by 416F2 is similar to the other successful phosphinic inhibitors reported in the literature. As an example, RXP407 is a phosphinic inhibitor that selectively inhibits the N domain of the ACE2 homologue ACE with a K_i value of 12 nM (Dive *et al.*, 1999). RXPA380 is another phosphinic inhibitor that selectively inhibits the C domain of ACE with a K_i of 3 nM (Georgiadis *et al.*, 2003). RXPA380 has been successfully co-crystallized with testis ACE (an ACE isoform identical to the C domain of ACE), and the crystal structure has led to identification of interactions between the inhibitor and the enzyme active site residues (Corradi *et al.*, 2007). The success of the studies on phosphinic inhibitors supports the choice of 416F2 as a suitable probe for the investigation of interactions between a ligand and ACE2 active site residues.

Chapter 3
Molecular docking simulation of 416F2 into the ACE2 crystal structure

3.1. Introduction

The development of novel ACE2 inhibitors, in combination with molecular docking simulations, has to date elucidated several structure-function relationships of ACE2 (Deaton *et al.*, 2008a; Deaton *et al.*, 2008b; Mores *et al.*, 2008). The findings of the *in silico* studies were in agreement with the insights from the crystal structure, which included the revelation of a large S_1' sub-site of ACE2. This sub-site has been reported by the various molecular docking simulations and the crystal structure to readily accommodate both large basic or hydrophobic substrate residues (Deaton *et al.*, 2008b; Mores *et al.*, 2008; Towler *et al.*, 2004). On the unprimed side of the scissile bond, the small S_1 sub-site is highly selective for P_1 substrate residues. While previous studies have revealed much information about the two sub-sites adjacent to the catalytic zinc, the contribution of sub-sites distal to the catalytic cleavage site remain unclear.

In this study, a 3D structure of a novel potent phosphinic inhibitor of ACE, 416F2, was docked by molecular simulation into the active site of ACE2 crystal structure (PDB code: 1R4L). The docked model was examined for potential interactions formed between active site residues and the extended P_2 residue of 416F2. Subsequently the significance of S_1 residue Tyr510 was evaluated by mutagenesis.

3.2. Materials & Methods

3.2.1. Molecular docking of 416F2 into the ACE2 crystal structure

The ACE2 structure used as the receptor in the docking experiment was prepared from the crystal structure of the inhibitor MLN4760-bound ACE2 (PDB code: 1R4L). The PDB file for the inhibitor-bound ACE2 crystal structure was obtained from the RCSB Protein Data Bank (http://www.rcsb.org). The preparation of molecular structures for the docking experiment was carried out on the Discovery Studio v.1.7 platform (Accelrys Software Inc.). The structure of the inhibitor MLN4760 was removed from the active site of the ACE2 crystal structure and the integrity of the ACE2 structure was examined for any missing residues. The incorrect valence of the catalytic zinc from the removal of the co-crystallized inhibitor was corrected by the assignment of a +2 charge to the zinc atom. The missing hydrogen atoms were also added to the structure.

The Dundee PRODRG2 Server online application (Schuttelkopf and van Aalten, 2004) was used to generate PDB files containing the co-ordinates of the 3D structure of 416F2. The chemical structure of 416F2 was used as input into the PRODRG2 server and the generated structure of 416F2 was energy minimized by GROMOS87 force field (Schuttelkopf and van Aalten, 2004). The C-terminal carboxylate and the phosphinate group of 416F2 were adjusted to a partial double bond to simulate the dissociated state of the acidic groups.

The docking experiment was performed using the Dock Ligands (LigandFit) protocol of Discovery Studio v.1.7. The docking protocol was carried out using the CFF force field, and a variable number of Monte Carlo Steps defined by Discovery Studio v.1.7 was chosen as the conformation search number of Monte Carlo trials. The docking protocol allowed for the flexibility of the ligand (416F2) while the receptor (ACE2) structure was defined as a rigid structure.

The binding site for the docking experiment was defined by two different methods. The volume of the MLN4760 inhibitor co-crystallized with ACE2 was taken as the ligand binding site for the re-docking of the MLN4760 structure (Fig.3.1.1). The

binding site for the docking of 416F2 was manually defined by selecting the volume of interior space of ACE2 and this volume was trimmed to include only the S_1 and S_1' sub-site space (Fig.3.1.2).

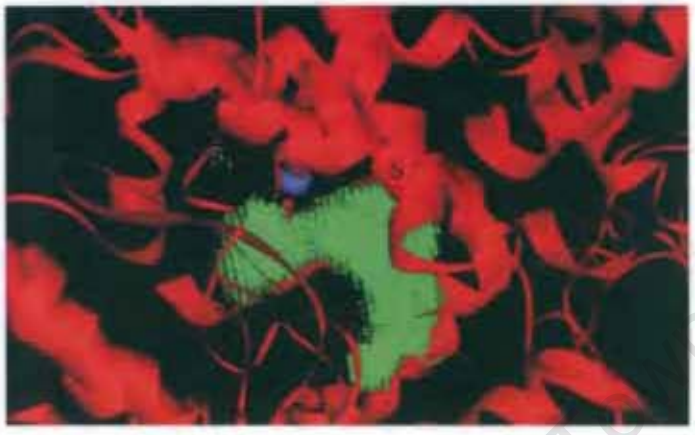

Fig.3.1.1. Ligand binding site defined for the re-docking of MLN4760. The volume of the MLN4760 (green matrix) co-crystallized with ACE2 (red ribbon representation) was taken directly as the ligand binding site. The catalytic zinc is shown as grey sphere. The relative position of the S_1 and S_1' sub-sites are also shown. The image was generated using Discovery Studio v.1.7 (Accelrys Software Inc.).

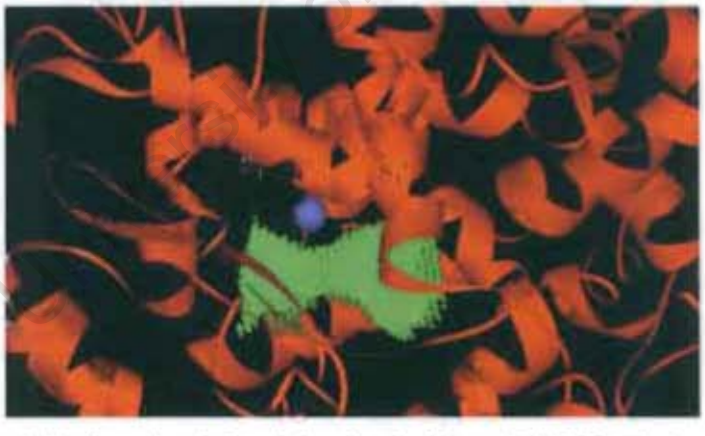

Fig.3.1.2. Ligand binding site defined for the docking of 416F2. The interior space of ACE2 was manually trimmed and the volume of the S_1 and S_1' sub-site of ACE2 was retained. The volume defined as the binding site is shown as the green matrix and the ACE2 structure is shown as the orange ribbon model. The catalytic zinc is shown as grey sphere. The relative position of the S_1 and S_1' sub-sites are also shown. The image was generated using Discovery Studio v.1.7 (Accelrys Software Inc.).

The poses from each docking experiment were subjected to a post docking energy minimization procedure. The energy minimization step was achieved by using the Ligand Minimization protocol of Discovery Studio v.1.7. The CFF force field was

chosen for the minimization protocol. The quality of the docking experiment protocol was evaluated by comparing the poses of the ACE2-MLN4760 docked model to the crystal structure. The positioning of the zinc coordinating group and the interaction between the ligand C-terminal carboxylate and residue Arg273 were chosen as the key criteria in the assessment of the docked models. All poses that failed to retain the key interactions were discarded. In addition, the root mean square deviations (RMSD) of the retrieved docked MLN4760 poses were determined using Discovery Studio Visualizer ver.1.7 (Accelrys Software Inc.) with reference to the crystal structure.

In the docking experiment of 416F2 into ACE2 crystal structure, the overall protein structure of the docked model was superimposed onto the crystal structure and the RMSD determined for the C-α backbone. All poses of the docked model were visually assessed for proper zinc coordination by the phosphinate and C-terminal carboxylate coordination with Arg273. Docked poses which do not contain these interactions were discarded.

3.2.2. Mutagenesis of Tyr510 residue

The Tyr510 codon of the sACE2FLAG construct was mutated to phenylalanine and threonine by site directed mutagenesis. The nucleotide sequence of the mutagenesis primers are shown in Table 3.1. A silent mutation was incorporated into the mutagenesis primers to introduce an *Xmn*I restriction site as a screening marker. The mutagenesis reaction mixtures of 50 µl final volume contained final concentrations of 0.6 µM of each forward and reverse mutagenesis primers, 0.4 mM dNTP mix (Sigma-Aldrich, Inc.), 1 x Hi-Fi® buffer (Kapa Biosystems) and 1 unit of Hi-Fi® polymerase (Kapa Biosystems). The mutagenesis reaction was carried out using a thermal cycling condition of an initial incubation at 95°C for 5 min, followed by 16 cycles of incubation for 30 sec at 95°C, 15 sec at 67°C and 4 min at 72°C. The choice of an elongation period of 4 min at 72°C was based on recommendation made by manufacturer of the Hi-Fi® polymerase (Kapa Biosystems). The melting temperature of the mutagenesis primer, provided by the manufacturer (Inqaba Biotech), was chosen as the annealing temperature for the mutagenesis reaction. After thermal cycling, a volume of 20 µl of mutagenesis reaction was treated with 10 units of *Dpn*I

and incubated at 37°C for at least 2 hours and 5 µl of the digestion mixture was transformed into *E.coli DH5α* following a heat shock transformation procedure (see appendix A.4). Positive bacteria colonies carrying plasmids with the desired mutation were identified by *Xmn*I restriction enzyme digestion (see appendix A.2 and A.5). Plasmid DNA samples that screened positive were subjected to DNA oligonucleotide sequencing analysis (see appendix A.6) to confirm proper introduction of the desired mutation, as well as to identify any spurious mutation within the constructs. Following confirmation of desired mutations in the mutant sACE2FLAG constructs, the construct fragments were sub-cloned back into expression plasmid pCI-NEO (Promega Corp.) at the *Xho*I and *Xba*I sites to eliminate any spurious mutation introduced into the vector by the mutagenesis reaction. The final sACE2FLAG-Y510F and sACE2FLAG-Y510T plasmids were prepared in bulk with Plasmid Midi Kit (Qiagen) following instruction provided by the manufacturer.

Table 3.1. Oligonucleotide sequences used in site directed mutagenesis of Tyr510 residue of ACE2. Letters in lower case indicate mutated nucleotides. The underlined sequence indicate screening marker *Xmn*I restriction site introduced by a silent mutation. The *Xmn*I site is disrupted in the sACE2FLAG-Y510T mutagenesis primers (*Xmn*I recognition site: GAA NNNN TTC).

Primer name	Nucleotide sequence
sACE2FLAG-Y510F Fwd	5'-CTCTGTTCCATGTTTCgAATGATTtCTCATTCAT TCGATATTAC-3'
sACE2FLAG-Y510F Rev	5'-GTAATATCGAATGAATGAGaAATCATTcGAAA CATGGAACAGAG-3'
sACE2FLAG-Y510T Fwd	5'-CTCTGTTCCATGTTTCgAATGATacCTCATTCAT TCGATATTAC-3'
sACE2FLAG-Y510T Rev	5'-GTAATATCGAATGAATGAGgtATCATTcGAAAC ATGGAACAGAG-3'

3.2.3. Activity assay of sACE2FLAG Y510 mutants

The sACE2FLAG Y510F and sACE2FLAG Y510T plasmids were transfected into HEK293 cells using the Profection Transfection Kit (Promega) as described in 2.2.2.1. Briefly, 1.5 µg of plasmid was transfected into cells in a 6-well tissue culture plate (Nunc Inc.) cultured to 60% confluent following the procedure described by the manufacturer (also see 2.2.2.1). Selection for positively transfected clones was

achieved by incubation of transfected cells in culture medium containing 0.8 mg/ml mammalian antibiotic G418. The transfected cell cultures were plated into 6-well tissue culture plates and grown to confluence. The confluent cell cultures were rinsed with 1 × PBS and the culture medium was replaced with 1 ml of serum-free OptiMEM culture medium (Gibco®, Invitrogen). The culture medium was collected after 24 hours and 10 µl of the medium was subjected to a FRET-based ACE2 fluorogenic assay using a concentration of 10 µM of the synthetic substrate Mca-YVADAPK-Dnp (as described in 2.2.3.1). The activity assays were performed in triplicate on each transfection. Upon failure to detect ACE2 activity in culture medium exposed to transfected cells for 24 hours, expression of the transfected soluble enzymes was assessed by western blot analysis performed with 5 µg of total protein from the collected serum-free culture medium. The cell lysates were also examined for any incorrectly processed ACE2 protein. Cell lysates were harvested from confluent transfected cell culture in 6-well plates using 0.5 ml of Triton lysis buffer (50 mM HEPES, pH 7.5; 500 mM NaCl; 1% (v/v) Triton X-100; 1 mM PMSF) and an amount of 5 µg total protein was subjected to western blot analysis.

3.2.4. Western blot analysis of ACE2 protein

Total protein amount of 10 µg was fractionated by SDS-PAGE using an 8% gel (see appendix A.8). After electrophoresis, the gel was soaked in blotting buffer (25 mM Tris, pH 8.2; 200 mM glycine; 20% (v/v) methanol) and the Mini Trans-Blot transfer apparatus (Bio-Rad Laboratories) was assembled with the gel placed on the cathode end. The transfer process was carried out at 300 mA for 60 min. After the transfer process, the nitrocellulose membrane was blocked by an overnight incubation in 5% (w/v) skim milk in Tris buffered saline Tween (TBS-T) (50 mM Tris, pH 7.4; 200 mM NaCl; 0.1 % (v/v) Tween) at 4°C overnight. Following the blocking procedure, the nitrocellulose membrane was incubated on a shaker at room temperature with monoclonal anti-human ACE2 ectodomain antibody (R&D Systems, Inc.) diluted 1:1000 in 5% (w/v) skim milk in TBS-T for at least 90 min. After the antibody incubation, the membrane was rinsed and washed in 5% (w/v) skim milk in TBS-T for 15 min. The wash step was repeated another three times (5 min per wash). The washed membrane was incubated with 1:2000 diluted secondary anti-mouse antibody

conjugated with horse radish peroxidase (Amersham, GE Healthcare) in 5% (w/v) skim milk in TBS-T at room temperature on shaker for 60 min. After the secondary antibody incubation, the membrane was rinsed and washed as before in TBS-T. Detection of blotted target protein was carrying out in the dark using ECL Plus™ Western Blotting Detection kit (Amersham, GE Healthcare). The membrane was exposed to the detection solution for 5 min and the chemi-luminescence released was captured using Hyperfilm™ ECL photographic films (Amersham, GE Healthcare).

3.3. Results

3.3.1. Accuracy assessment of the docking protocol

The inhibitor bound ACE2 crystal structure was used to verify the accuracy of the molecular docking. The structure of MLN4760 taken from the crystal structure served as a control for the accuracy of the docking protocol, as the MLN4760 crystal structure was expected to dock into ACE2 active site in the same conformation. As the docking experiment involved the fitting of a relatively small molecule into the large interior cavity of ACE2, a careful choice of the ligand binding site was necessary to ensure proper docking of the ligand to the expected site of interaction. Thus in the assessment for the accuracy of the docking protocol, the space occupied by MLN4760 in the crystal structure was chosen as the volume of the binding site.

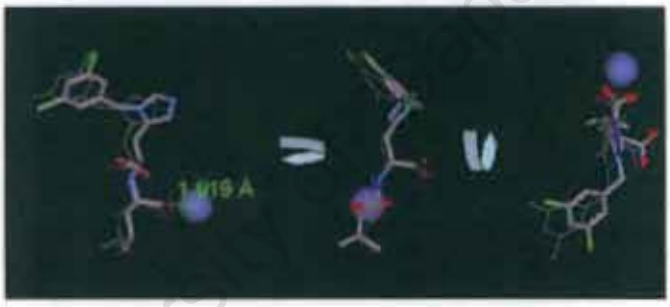

Fig.3.2.1. Superimposition of docked MLN4760 model on MLN4760 structure observed in the crystal structure. Docked model is represented in stick model, and MLN4760 structure in the crystal structure is shown as line model. The image was generated using Discovery Studio v.1.7 (Accelrys Software Inc.).

In the model of MLN4760 docked into the ACE2 crystal structure, the MLN4760 backbone in the docked model assumed a position very similar to that observed in the crystal structure (Fig.3.2.1). The docked MLN4760 was identically orientated and its zinc coordinating carboxylate was 1.9 Å from the active site zinc. The interaction between the C-terminal carboxylate of MLN4760 and Arg273 was preserved in the docked model, however the hydrogen bonds between the C-terminal carboxylate and His345 and His505 were lost as the C-terminal carboxylate was pulled closer towards Arg273 (Fig.3.2.2). The 3, 5-dichlorobenzylimidazole side chain moiety of the docked MLN4760 assumed a highly similar conformation to that observed in the crystal structure (Fig.3.2.1). In the docked model, the MLN4760 P_1 isobutyl side

chain extended towards Tyr510 like that found in the crystal structure. The interaction between the imidazole of MLN4760 and Thr371 found in the crystal structure was also conserved in the docked model (Fig.3.2.2).

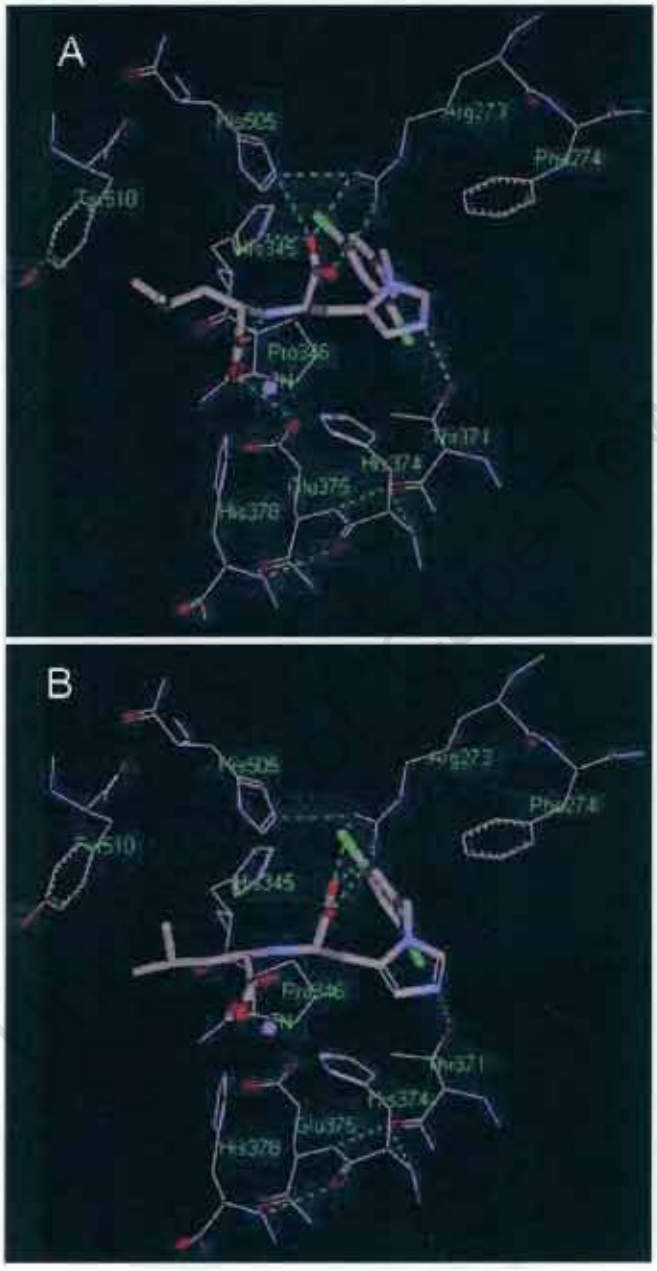

Fig.3.2.2. Comparison of the interactions between bound ligand and ACE2 active residues found in crystal structure and the docked model. The structure of MLN4760 is shown as stick model and the ACE2 residues are shown as line model. The green dash lines represent hydrogen bonds. **A.** The interactions between MLN4760 and ACE2 observed in the crystal structure are shown. **B.** Docked model of MLN4760 into the active site of ACE2. The interaction between Arg273 and the C-terminal carboxylate of MLN4760, and the interaction between Thr371 with the imidazole ring was conserved in the docked model. The image was generated using Discovery Studio v.1.7 (Accelrys Software Inc.).

An additional assessment of the docking protocol accuracy was the determination of RMSD of the docked MLN4760 poses in reference to the pose of MLN4760 found in the crystal structure. RMSD values were determined for all docked poses where interactions of the MLN4760 carboxylates with the active site zinc and Arg273 were retained (Table 3.2). An averaged RMSD of 2.05 Å was obtained for the docking of MLN4760 back into the ACE2 crystal structure. Considering these findings, the same docking protocol was applied to the subsequent molecular docking of 416F2 into the ACE2 structure.

Table 3.2. RMSD of docked MLN4760 poses. The pose of MLN4760 in the crystal structure was used as the reference ligand.

Pose no.	**RMSD (Å)**
Reference molecule	0.0
1	1.21
2	1.24
3	2.44
4	2.44
5	2.37
6	2.59
Average	2.05

3.3.2. Model of 416F2 docked into ACE2 crystal structure

The overall structure of the ACE2 was retained in the 416F2 docked model when compared to the crystal structure, as indicated by the C-α backbone RMSD value of 0.02 Å determined for the superimposition of 416F2 docked model to the ACE2 crystal structure. Furthermore, both the docked model and the crystal structure yielded identical Ramachandran plots by the ProCheck software (Laskowski *et al.*, 1993) (data not shown). These findings were expected as the docking experiments were carried out using the rigid protein receptor approach. The low RMSD and identical Ramachandran plots indicated the overall integrity of the ACE2 protein was preserved during the docking process.

The poses retrieved for the docking of 416F2 are shown in Fig.3.3. The zinc-coordinating phosphinate and the P_1 prolyl groups were similarly positioned in all poses, indicating highly constrained fitting of these moieties (Fig.3.3). In the docked model, the phosphinate group was positioned close to the active site zinc atom, and an ionic interaction between the C-terminal carboxylate of 416F2 and Arg273 was observed (Fig.3.4). However, the C-terminal carboxylate coordinating hydrogen bond provided by His345 and His505 was not observed. The hydrophobic P_1' benzyl group of 416F2 extended towards Phe274 in the large S_1' sub-site (Fig.3.4). The P_1 prolyl group of 416F2 was found to be accommodated by the ACE2 S_1 sub-site (Fig.3.6.1). The P_1 prolyl groups in different poses were near identically positioned (Fig.3.3), indicating the P_1 prolyl group is highly restrained in structure. A high degree of movement was observed for the N-terminal acetyl-leucyl moiety of 416F2 (Fig.3.3 and Fig.3.5). A closer examination of the ACE2 active site revealed a large cavity which occurs beyond the S_1 sub-site (Fig.3.6.1 and Fig.3.6.2), and this large space allowed rotation of the P_2 leucyl side chain of 416F2. A hydrogen bond between the terminal hydroxyl of Tyr510 and the carbonyl of the acetyl group of 416F2 distanced at 3.05 Å apart was observed (Fig.3.4 and Fig.3.5), while another was seen between the carbonyl group of the 416F2 P_2-leucyl group and the amine group of Ala348 distanced at 2.99 Å apart (Fig.3.5).

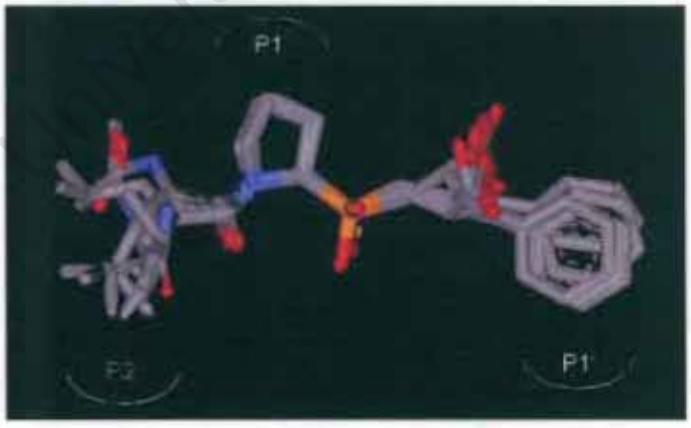

Fig.3.3. Docked poses retrieved for the docking of 416F2 into ACE2 crystal structure. The P_1 and phosphinate moieties assumed similar conformations in all docked poses, indicating constrained fitting of these groups. The P_2 acetyl-leucyl groups of the different poses were positioned in highly dissimilar manner.

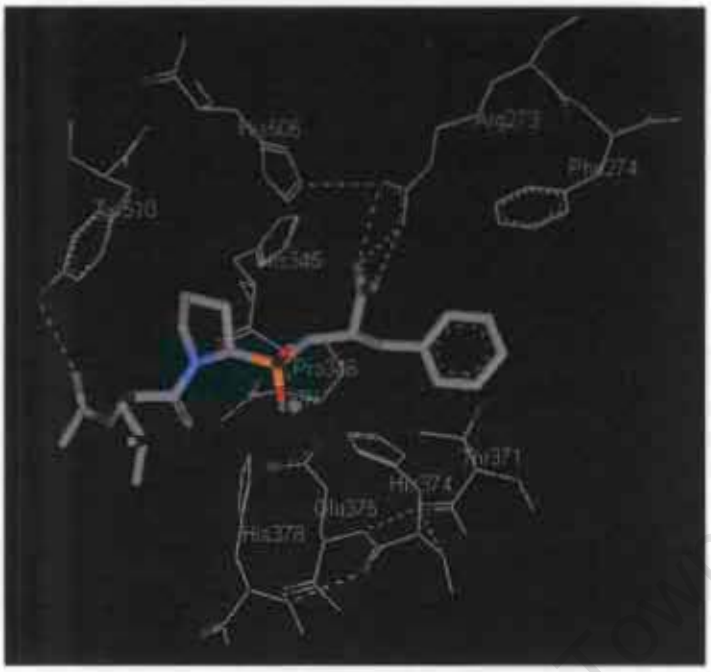

Fig.3.4. Interactions between 416F2 (shown as stick model) and ACE2 active site residues (shown as line model) in docked model. There interaction between the terminal carboxylate of 416F2 and Arg273 was preserved. The zinc coordinating phosphinate was positioned accordingly close to the active site zinc atom (represented as grey sphere). The image was generated using Discovery Studio v.1.7 (Accelrys Software Inc.).

3.3.3. Investigation of the ACE2 S_2 sub-site

From the comparison of the various poses of 416F2 in the docked model it was observed that the acetylated P_2-leucyl residue of 416F2 displayed a greater degree of freedom than the P_1 and P_1' groups (Fig.3.3 and Fig.3.5). A closer examination of the docked model revealed a lack of interactions between the P_2-leucyl group of 416F2 and the ACE2 active site residues. In the docked model, the presence of a clear S_2 sub-site in the ACE2 active site was not observed as the ACE2 active site opens into a large space beyond the S_1 sub-site (Fig.3.6.1 and Fig.3.6.2). The active site cavity narrows as it approached the S_1 sub-site where the P_1 prolyl group of 416F2 was tightly accommodated. An examination of the active site cavity surface identified Tyr510 as a residue that protrudes into the cavity and its terminal hydroxyl group provides a point of interaction with the ligand (Fig.3.7). Based on this it was decided for this study to further investigate the role of Tyr510 in the enzymatic activity of ACE2.

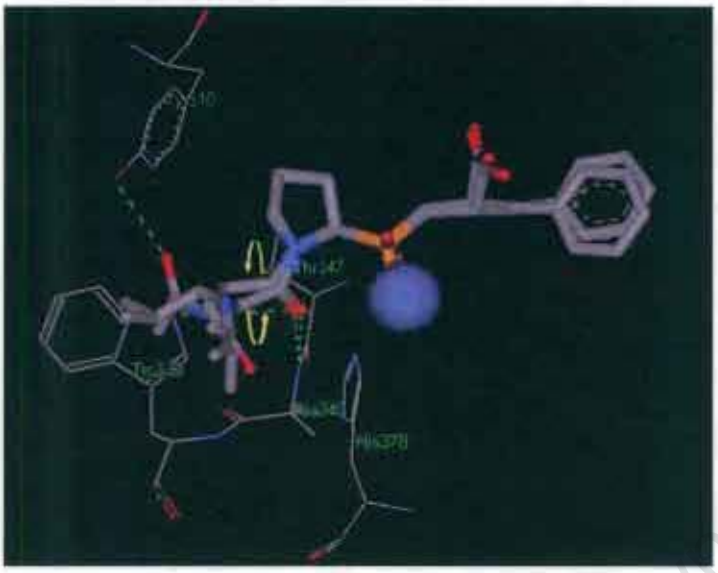

Fig.3.5. Overlay of selected poses from the molecular docking of 416F2 in ACE2. Good correlations between the different poses at the P_1' and P_1 positions were observed. The position of the zinc coordinating phosphinate and the C-terminal carboxylate were highly similar between the poses. The greatest difference between the poses was at the P_2 position where the leucyl group was able to rotate around the bond between its α-carbon and carbonyl group (indicated by yellow arrows). The structure of 416F2 is shown as stick model and the ACE2 residues are shown as line model. The image was generated using Discovery Studio v.1.7 (Accelrys Software Inc.).

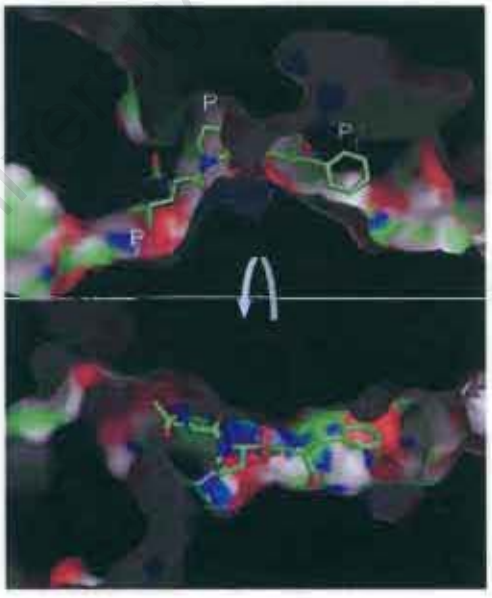

Fig.3.6.1. Model of 416F2 docked into ACE2. Most of the 416F2 molecule was tightly bound to the ACE2 active site. The P_1 prolyl side chain is well accommodated by the S_1 sub-site, while the P_1'-benzyl and the P_2-leucyl groups extended into less confined spaces. Images were generated using PyMol Molecular Viewer (DeLano Scientific LLC.).

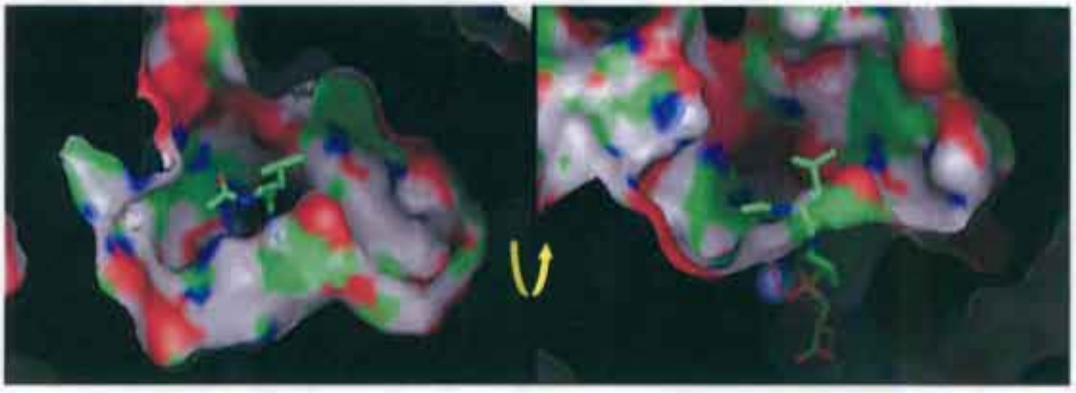

Fig.3.6.2. Docked model of 416F2 (shown as stick model) into the ACE2 active site showed a large cavity beyond the S_1 sub-site. The P_2 acetyl-leucyl moiety of 416F2 extended freely into the large cavity and a clear S_2 sub-site was no observed. Images were generated using PyMol Molecular Viewer (DeLano Scientific LLC.).

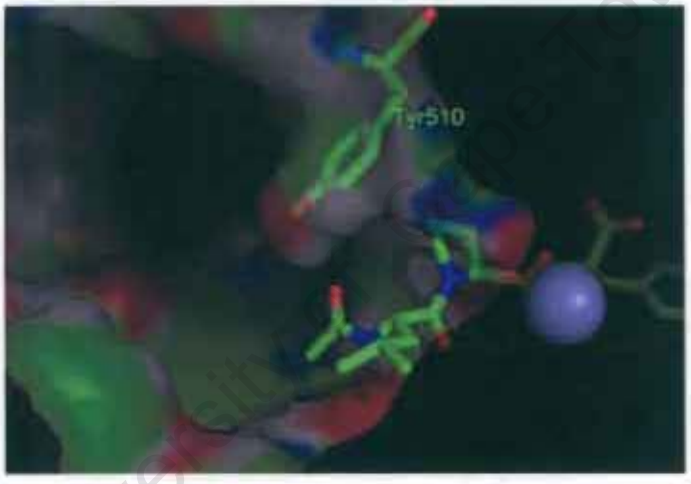

Fig.3.7. Tyr510 that forms part of the S_1 sub-site protrudes into the cavity of the ACE2 active site beyond the S_1 sub-site. The extension of the terminal hydroxyl of Tyr510 into the cavity allowed Tyr510 to interact with the carbonyl of the acetyl moiety of 416F2.

3.3.4. The role of Tyr510 in the enzymatic activity of ACE2

3.3.4.1. Mutagenesis of sACE2FLAG

In order to assess the contribution of Tyr510 in the ligand binding of ACE2, Tyr510 was mutated into a phenylalanine residue to determine the involvement of the terminal hydroxyl moiety. A separate mutation of Tyr510 to a threonine residue was made to elucidate the significance of the tyrosine hydrophobic ring in ligand binding. The Tyr510 mutations were introduced into the pCI-NEO-sACE2FLAG construct and

the proposed Y510F mutation was properly introduced into the sACE2FLAG construct, as indicated by *Xmn*I restriction enzyme digestion (Fig.3.8.1). Restriction enzyme digestion of the sACE2FLAG-Y510T mutant construct with *Xmn*I revealed digestion patterns identical to that of the unmutated sACE2FLAG (Fig.3.8.1). The sACE2FLAG-Y510T mutagenesis primer pair contained a disrupted *Xmn*I site (see Table 3.1), and thus *Xmn*I digestion failed to confirm the introduction of desired tyrosine to threonine mutation (Fig.3.8.1). Proper tyrosine to threonine mutation was verified by nucleotide sequencing analysis, and partial sequence alignment of sACE2 Y510T to ACE2 cDNA indicated proper tyrosine to threonine mutation (Fig.3.8.2).

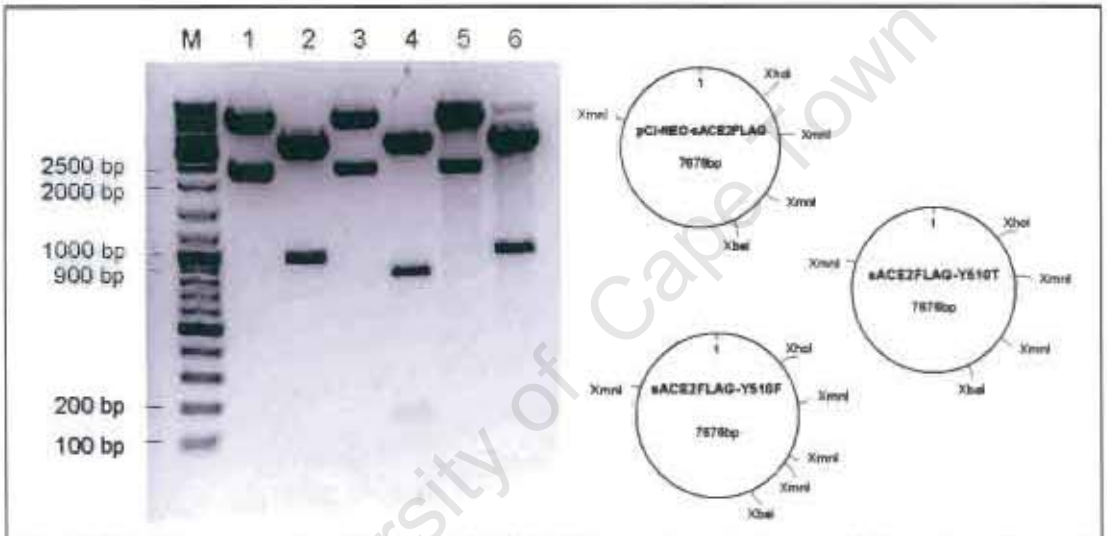

Fig.3.8.1. Mutagenesis of the sACE2FLAG construct. Successful re-insertion of mutated sACE2FLAG constructs into expression vector pCI-NEO was indicated by the excision of a 2205 bp fragment corresponding to the size of the sACE2FLAG construct with restriction enzyme *Xho*I and *Xba*I (**Lane 1:** sACE2FLAG. **Lane 3:** sACE2FLAG-Y510F. **Lane 5:** sACE2FLAG-Y510T). The new *Xmn*I restriction site introduced as screening marker led to the cleavage of the 944 bp fragment to an 805 bp and a 135 bp fragment, as observed for the Y510F mutation (**Lane 4:** *Xmn*I digested sACE2FLAG-Y510F). The *Xmn*I screening marker was disrupted by the nucleotide changes introduced by the ACE2 Y510T mutagenesis primers, therefore identical restriction enzyme digestion pattern was observed for sACE2FLAG and sACE2FLAG-Y510T(**Lane 2:** *Xmn*I digested sACE2FLAG. **Land 6:** *Xmn*I digested sACE2FLAG-Y510T). **Lane M:** O'GeneRuler™ DNA Ladder Mix (Fermentas Life Sciences)

Fig.3.8.2. Partial nucleotide sequence alignment of sACE2FLAG-Y510T mutant construct to ACE2 mRNA (GenBank Accession number: NM_021804). The desired mutation and the screening marker silent mutation were properly introduced.

3.3.4.2. Activity of sACE2 Tyr510 mutants

The mutated sACE2FLAG constructs were transfected into HEK293 cells and protein expressed into culture medium was assessed by an ACE2 activity assay. ACE2 activity was detected in medium harvested from sACE2FLAG and sACE2FLAG-Y510F transfected cells but not in that of sACE2FLAG-Y510T transfected cells even after prolonged period of incubation with substrate (Fig.3.9).

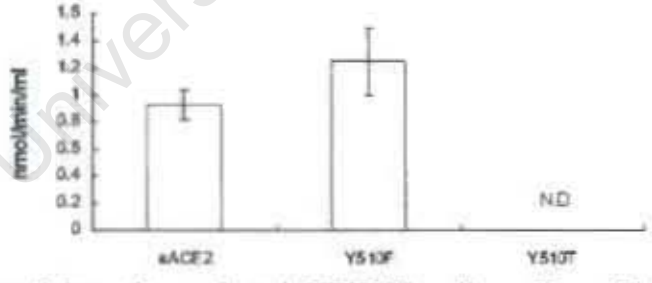

Fig.3.9. ACE2 activity of transfected HEK293 cell medium. Culture medium of HEK293 cells transfected with sACE2FLAG, sACE2FLAG-Y510F or sACE2FLAG-Y510T constructs was harvested after 24 hours of incubation, and ACE2 activity in the medium was assayed using the synthetic substrate Mca-YVADAPK-Dnp. ACE2 activity is expressed as nmol of Mca-YVADAP formed per min per ml of medium. Experiments were carried out in triplicate and the data represented by the mean with S.D. N.D: not detected.

In order to determine whether the sACE2FLAG-Y510T mutant was properly expressed but not active, western blot analysis was performed on the culture medium

and transfected cells. The strong detection of a protein by anti-ACE2 antibody in the cell culture medium and not in the cell lysate confirmed the expression and secretion of soluble sACE2FLAG-Y510T (Fig.3.10.). These results clearly show that the sACE2FLAG-Y510T mutant is not catalytically active despite its expression into the culture medium.

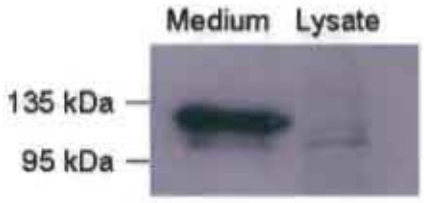

Fig.3.10. Western blot of the transfected cells and cell medium of sACE2FLAG-Y150T. The strong presence of ACE2 detected in the culture medium was indicative of expression of soluble sACE2FLAG-Y510T into the culture medium.

3.4. Discussion

An investigation of the ACE2 S_2 sub-site was performed by *in silico* molecular docking in which the structure of a potent phosphinic inhibitor, 416F2, was docked into the closed conformation of the ACE2 crystal structure. The accuracy of the docking protocol of choice was assessed by docking of MLN4760 from the ACE2-MLN4760 crystal structure into the ACE2 molecule. The MLN4760 docked into a very similar position to what was found in the crystal structure (Fig.3.2.1), and key interactions of the MLN4760 carboxylates with the active site zinc and Arg273 were conserved in the docked model (Fig.3.2.2). Although a relatively high average RMSD (2 Å) was determined for the docked MLN4760 poses (Table 3.2), the RMSD values for the first two poses (1.21 and 1.24 Å for pose 1 and 2 respectively) were below the threshold RMSD of 2 Å commonly considered as the cut-off for successful docking experiments (Rarey *et al.*, 1996; Verdonk *et al.*, 2003).

In the docked model of 416F2 in ACE2, the ligand P_1' benzyl group of 416F2 extends into the large S_1' sub-site without fully occupying the S_1' sub-site, unlike the relatively larger P_1' 3, 5-dichlorobenzylimidazole of MLN4760 in the crystal structure. The hydrogen bond found between Thr371 and the P_1' imidazole of MLN4760 in the crystal structure was lost in the docked model due to the hydrophobic nature of the P_1' benzyl group of 416F2. Hydrogen bonding of Arg273 to the ligand C-terminal carboxylate was preserved in the docked model, however the other C-terminal carboxylate coordinating hydrogen bonds provided by His345 and His505 were lost. In the docked model 416F2 into ACE2, the ligand C-terminal carboxylate was extended closer towards Arg273 and away from His505 and His345 (compare Fig.3.2.2 and Fig.3.4). At the S_1/P_1 position, the P_1 prolyl group of 416F2 was well fitted into the S_1 sub-site, similar to the P_1 isobutyl group of MLN4760 found in the crystal structure. The 416F2 P_2-leucyl moiety assumed a variety of positions due to rotation around the bond between its α-carbon and carbonyl group (Fig.3.5).

The docked model did not reveal any interaction between the P_2-leucyl side chain of 416F2 and the ACE2 active site residues. Considering the substantial increase in active site cavity beyond the S_1 sub-site (Fig.3.6.1 and Fig.3.6.2), the possibility of some conformational change in the S_2 sub-site in the presence of a ligand P_2 group

should not be overlooked. However, this was averted in the molecular docking by the adoption of a rigid protein receptor approach. The active site residues were prevented from moving and this approach excluded the possibility of an induced-fit ligand binding mechanism from taking place. In retrospect, the choice of a flexible-receptor docking strategy might facilitate the elucidation of the S_2 sub-site. The flexible-receptor docking strategy was not adopted in this study due to consideration of the molecular dynamics of ACE2-ligand binding. The ACE2 active site consisted of residues located on different sub-domains of the ACE2 catalytic domain and the sub-domains undergo significant movement upon ligand binding (Towler *et al.*, 2004). The utilization of a flexible-receptor docking strategy is likely to disrupt the specific conformation of ligand-bound ACE2 and alter the active site structure.

Considering the mixed type inhibition displayed by 416F2 (see chapter 2), the uncompetitive nature of its ACE2 inhibitory activity could not be easily explained by the docked structure reported here. The possibility of the relatively small 416F2 binding elsewhere within the ACE2 active site cavity should not be overlooked. Similar to the flexible-receptor docking experiment, a major challenge faced by a search for alternative small molecule binding sites within the ACE2 active site is the consideration of its structural dynamics. The binding of 416F2 to a secondary site to exert uncompetitive inhibition is an event likely mediated through structural change of the ACE2 active site after 416F2 binding, and flexible-receptor molecular modelling could be used to explore this alternative mechanism. Moreover, x-ray crystallography could be employed to investigate the precise binding of 416F2 to the ACE2 active site and its interactions with the active site residues.

While the S_2 sub-site was not defined in the 416F2 docked model, a hydrogen bond between the terminal hydroxyl of Tyr510 and the acetyl carbonyl group of 416F2 was revealed in the docked model (Fig.3.4 and Fig.3.5). The bulky Tyr510 protrudes into the active site cavity beyond the S_1 sub-site (Fig.3.7), and thus it was decided to investigate the significance of Tyr510 in ACE2 activity. The mutation of Tyr510 into a threonine residue resulted in the complete loss of ACE2 activity (Fig.3.9) despite normal expression of the soluble protein into the culture medium (Fig.3.10). This finding suggested Tyr510 as an essential residue for ACE2 activity. Structurally, Tyr510 is well defined as a S_1 residue and the bulky hydrophobic ring of Tyr510

In summary, molecular docking of the potent ACE2 inhibitor 416F2 into the closed conformation of the ACE2 crystal structure revealed a weak hydrogen bond between the terminal carboxylate of Tyr510 and the carbonyl of the acetyl group of 416F2. Removal of the Tyr510 terminal hydroxyl by mutagenesis did not result in the abolition of ACE2 activity for the fluorogenic substrate Mca-YVADAPK-Dnp, indicating the hydrogen bond exerts a subtle effect in the enzymatic activity of ACE2. However, the mutation of Tyr510 to a threonine completely abolished ACE2 activity, indicating the hydrophobic environment and the steric hindrance provided by Tyr510 is essential for ACE2 activity.

Chapter 4

Elucidation of potential ACE2 transcription regulators

4.1. Introduction

Many physiological studies using animal models have elucidated the involvement of ACE2 in the protection of renal, cardio and respiratory tissues against injuries (Imai *et al.*, 2005; Oudit *et al.*, 2006; Yamamoto *et al.*, 2006). While intense research efforts have been invested in investigations of the functional role played by ACE2 in various body tissues and the RAS, the transcription regulation of ACE2 is still poorly understood. The findings of altered ACE2 expression under various pathological conditions, such as diabetes (Mizuiri *et al.*, 2008) and SARS CoV infection (Kuba *et al.*, 2005), highlight the involvement of ACE2 expression in the progression of these pathological conditions. Further understanding of the factors involved in ACE2 expression is likely to reveal plausible mechanisms in the pathogenesis of these events.

In an attempt to identify potential targets of hepatocyte nuclear factor 1 beta (HNF-1β) in HEK293 cells, Senkel *et al.* identified the ACE2 gene as a direct target gene of HNF-1β (Senkel *et al.*, 2005). In HEK293 cells, co-transfection of HNF-1β expression vector and ACE2 promoter constructs resulted in the activation of ACE2 promoter constructs as assessed by luciferase assays. However, the implication of the association between HNF-1β over-expression and ACE2 promoter activity is unclear.

The potential HNF-1β binding sites identified by Senkel *et al.* were originally thought to be located within the ACE2 promoter sequence (Senkel *et al.*, 2005). However, with the later identification of a 5′ untranslated exon (5′UTE) and an extended untranslated region of exon 1 (Itoyama *et al.*, 2005), the potential HNF-1β binding sites were in fact located within the 5′UTE and intron 1 regions of the ACE2 gene, and not the promoter. The findings by Senkel *et al.* demonstrated the involvement of the 5′UTE and intron 1 regions in the transcription regulation of ACE2, and additional transcription factors could confer ACE2 transcription regulatory effects by binding to these regions.

In this study, the 5′UTE and intron 1 regions of the ACE2 gene were collectively denoted as the immediately upstream region of the ACE2 translation start codon (ACE2 IUR). The ACE2 IUR was investigated for presence of potential transcription regulatory elements by promoter activity assays. The transcription activities of various truncated ACE2 IUR constructs were determined *in vitro*. The regions of ACE2 IUR found to contain potential transcription regulatory elements by promoter activity assays were subjected to bioinformatics searches to identify potential transcription factor binding sites within the ACE2 IUR. The binding of potential transcription factors from the bioinformatics search was evaluated by an *in vitro* binding assay.

4.2. Materials and methods

4.2.1. Human genomic DNA extraction

Human embryonic kidney (HEK293) cells were cultured in a 6-well plate until confluent and genomic DNA was extracted from cells according to a previously described procedure (Laird *et al.*, 1991). Briefly, the cultured cells were washed twice with 1 ml of PBS to remove residual culture medium. Cell lysis was achieved by direct addition of 0.5 ml lysis buffer (100 mM Tris, pH 8.5; 5 mM EDTA; 0.2% (w/v) SDS; 200 mM NaCl; 100 µg/ml Proteinase K) to the cells followed by incubation at 37°C for 30 min. Genomic DNA was precipitated from the cell lysate by addition of an equal volume of isopropanol. The precipitated genomic DNA was transferred into a 1.5 ml Eppendorf tube and washed with ice cold 70% (v/v) ethanol. Genomic DNA was then pelleted by centrifugation at 16000 × *g* using a micro-centrifuge. The supernatant was decanted and the DNA pellet was dried using a Savant Speedi Vac System (GMI, Inc.). The dried DNA was resuspended in nuclease-free water and the DNA concentration was determined using the absorbance reading at 260 nm measured with an Anthelie spectrophotometer (Secomam).

4.2.2. Generation of ACE2 IUR mutants from human genomic DNA

DNA fragments representing various lengths of the ACE2 IUR were amplified by the polymerase chain reaction (PCR) method. The nucleotide sequences of the primers used are shown in appendix Table A.2. Genomic DNA (100 ng) was used as template in all reactions. The ACE2 C2, C3 and C4 fwd primers were paired with the ACE2 Con Rev primer for the amplification of truncated ACE2 IUE mutant fragments (Fig.4.1.1). Nested deletion mutant fragments (Fig.4.1.2) were amplified from the C2 Forward construct using the various forward primers (-1131, -1077, -1036, -963, -903, -837, -784, -742, -704) and the ACE2 Pro Con Rev primer (see appendix Table A.2). The reaction mixtures contained 0.6 µM of each of the forward and the reverse primers, 0.4 mM dNTP mix (Sigma-Aldrich, Inc.), 1 × *Pfu* DNA polymerase buffer (Promega Corp.) and 5 units of *Pfu* DNA polymerase (Promega Corp.). Nuclease-free water was used to make up a total reaction volume of 50 µl in all reactions. The reaction mixtures were subjected to thermal cycling incubation of 95°C for 5 min,

followed by 25 cycles of thermal cycling at 95°C for 30 sec, 50°C for 30 sec and 72°C for varied elongation periods. Incubations at 72°C were timed such that it allowed the extension of 1 kb of DNA per minute. At the end of the 25 cycles a final incubation at 72°C for 5 min was included to allow strand completion. Finally the reaction was cooled to 4°C after thermal cycling. Thermal cycling was carried out using a Hybaid PCR Sprint thermal cycler (Mandel Scientific Inc.). All PCR products were analyzed by electrophoresis using 1% (w/v) agarose gel (see appendix A.5).

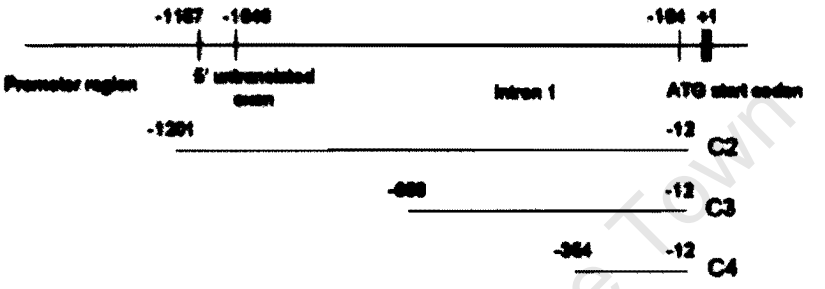

Fig.4.1.1. ACE2 IUR deletion mutants amplified from human genomic DNA. The deletion mutants were screened by luciferase assays to determine their activities in driving gene expression.

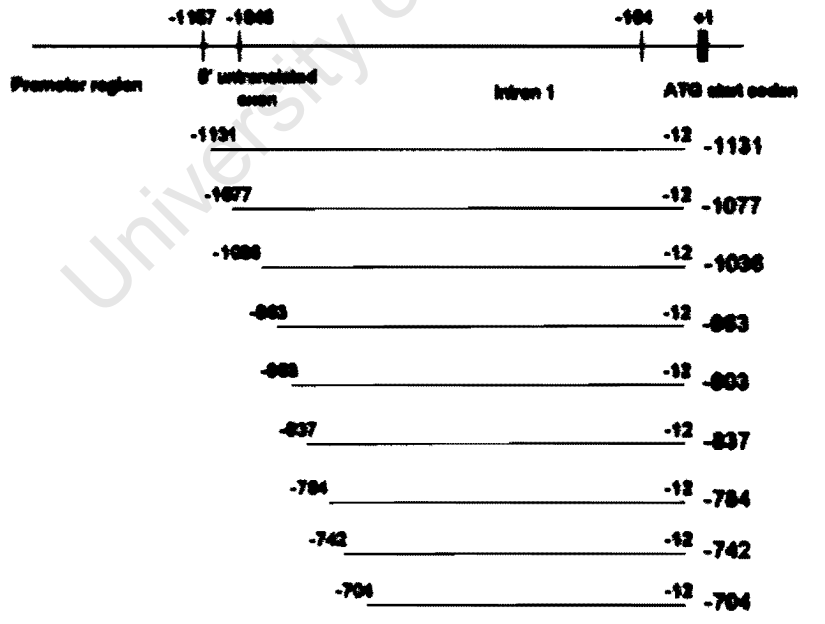

Fig.4.1.2. ACE2 IUR nested deletion mutant constructs amplified from the C2 construct. These mutants represent the segmentation of the -1201 to -658 region of ACE2 IUR and their activities in driving gene expression were assessed by dual luciferase assays.

4.2.3. Sub-cloning of ACE2 IUR mutants

4.2.3.1. Sub-cloning of ACE2 IUR mutant constructs C2, C3 and C4

The ACE2 IUR mutants generated by PCR were analyzed by electrophoresis using 1% (w/v) agarose gel (see appendix A.5) and purified using the Wizard SV Gel and PCR clean-up system (Promega Corp.). The mutant fragments were inserted into the cloning vector pBluescript at the *EcoR*V site. In order to facilitate ligation of the mutant fragments with pBluescript, the mutant fragments were treated with polynucleotide kinase (PNK) (Roche Ltd.) in 1 × T4 ligase buffer (Promega Corp.) at 37 °C for 1 hour. Restriction enzyme *EcoR*V digested pBluescript vector was treated with shrimp alkaline phosphatase (Fermentas Life Sciences) in 1 × reaction buffer at 37 °C for 1 hour. After PNK and shrimp alkaline phosphatase treatment, heat inactivation of the enzymes were achieved by incubation at 65°C for 20 min. Ligation was achieved through overnight incubation of linear pBluescript vector and ACE2 IUR mutant fragments (at a molar ratio of 1:3) at 4°C in a total volume of 10 µl, containing 1 × ligase buffer (Promega Corp.) and 3 units of T4 DNA ligase (Promega Corp.). Ligation products were transformed into *E.coli DH5*α (see appendix A.4) and positive clones were screened using restriction enzyme digestion with *BamH*I and *Hind*III. The plasmid DNA used in all restriction enzyme digestion screens were prepared using a small scale plasmid preparation (see appendix A.2). As the ACE2 IUR mutant fragments were inserted into vector pBluescript by blunt-end ligation, an additional screening step using restriction enzyme digestion with *Hind*III and *Age*I was performed to determine the orientation of inserted constructs. Positive screens of pBluescript containing inserted ACE2 IUR mutant fragments were prepared using Zippy® Mini-Prep Kit (Zymo Research) and subjected to partial DNA sequencing (see appendix A.6). DNA sequencing was performed to confirm the orientation of the inserted fragments, as well as the nucleotide sequence of the fragments. Constructs of ACE2 IUR mutant fragments reversely inserted into pBluescript were retained as a control for subsequent transfections and dual luciferase assays. Constructs with the correct sequence were sub-cloned from pBluescript into the reporter vector pGL3-BASIC (Promega Corp.) at the *Sma*I and *Hind*III sites. The final pGL3-BASIC

constructs were propagated in *E.coli DH5α* and prepared using the Zippy® Mini-Prep Kit (Zymo Research).

4.2.3.2. Sub-cloning of ACE2 IUR mutant constructs -1131, -1077, -1036, -963, -903, -837, -784, -742 and -704

The nested deletion ACE2 IUR mutant fragments were amplified by PCR using ACE2 IUR mutant C2 construct as the template, together with the respective forward primers and the ACE2 Pro Con Rev (see appendix Table A.2.). The thermal cycling profile described in 4.2.3.1 was followed. The PCR products were analyzed by electrophoresis using 1% (w/v) agarose gel and purified by using the Wizard® SV Gel and PCR clean-up system (Promega Corp.). Restriction enzyme digestion of the PCR products was carried out with *Hind*III (see appendix A.5). The digested PCR products were purified by gel extraction method using the Wizard® SV Gel and PCR clean-up system. Purified restriction enzyme digestion products were incubated with 10 units of polynucleotide kinase (Roche Ltd.) in 1 × T4 DNA ligase buffer (Promega Corp.) at 37°C for 1 hour, followed by heat inactivation of the polynucleotide kinase through incubation at 65°C for 20 min. The vector pBluescript was prepared by double restriction enzyme digestion using *Sma*I and *Hind*III carried out in the presence of 10 units of shrimp alkaline phosphatase (Fermentas Life Sciences). Ligations of the nested deletion mutant fragments with pBluescript were carried out as described in 4.2.3.1. The screening of correctly inserted fragments was achieved by restriction enzyme digestion using *Xho*I and *Xba*I. The orientation of the nested deletion mutant fragments ligated into pBluescript was confirmed by restriction enzyme digestion using *Age*I and *Xho*I and DNA nucleotide sequencing. The nested deletion mutant fragments were sub-cloned into the reporter plasmid pGL3-BASIC using restriction enzymes *Sac*I and *Xho*I. The final pGL3-BASIC constructs were propagated in *E.coli DH5α* and prepared using the Zippy® Mini-Prep Kit (Zymo Research).

4.2.4. Reverse transcription of the ACE2 transcript

HEK293 and COS-1 cells were cultured until confluent in 10 mm culture dish (Nunc Corp.) and total RNA was extracted using QIAzol reagent (Qiagen). Briefly, QIAzol

reagent (1 ml) was directly added to cell cultures and the cells were scraped off the culture plate and collected into 1.5 ml Eppendorf tubes. A volume of 0.2 ml of chloroform was added to the collected cells and immediately mixed by gentle inversion of the tube for 15 sec. The mixture was incubated on ice for 10 min followed by centrifugation at 12000 × *g* for 15 min at 4°C. The aqueous phase was collected and mixed with 0.5 ml of isopropanol. The mixture was incubated at -20°C for at least 2 hours to allow precipitation of RNA. The precipitated RNA was collected by centrifugation at 12000 × *g* for 30 min at 4°C. The RNA pellet was washed using 1 ml of ice cold 70% (v/v) ethanol and re-collected by centrifugation at 12000 × *g* for 30 min at 4°C. The supernatant was decanted and the RNA pellet was allowed to air-dry. The purified total RNA was resuspended in 30 µl of DEPC treated sterile water, and the concentration was determined using absorbance reading at 260 nm measured with an Anthelie spectrophotometer (Secomam). The RNA concentration was calculated with a conversion factor of 40 ng per unit absorbance reading at 260 nm. The quality of the isolated total RNA was assessed by gel electrophoresis using 1% (w/v) agarose denaturing gel containing 7% (v/v) formaldehyde and 1 × MOPS buffer (0.1 M MOPS; 5 mM sodium acetate; 0.5 mM EDTA, pH 7.0). The fractionated total RNA was stained with ethidium bromide (0.1 µg/ml) for 10 min followed by visualization under UV light.

Reverse transcription of the ACE2 transcript was carried out with the isolated total RNA using the OneStep RT-PCR Kit (Qiagen). The reaction mixture included a final concentration of 1 × Qiagen OneStep RT-PCR buffer (Qiagen), 0.6 µM of the ACE2 RT Fwd and the ACE2 RT Rev primers (see appendix Table A.2), and 400 µM of the dNTP mix (Sigma-Aldrich, Inc.). A volume of 2 µl of Qiagen OneStep RT-PCR Enzyme Mix (Qiagen) was added to the reaction mixture followed by the addition of 2 µg, 200 ng or 20 ng total RNA. A total reaction volume of 50 µl was achieved by the addition of ribonuclease-free water. The reverse transcription reactions were performed using a Hybaid PCR Sprint thermal cycler (Mandel Scientific Inc.). The reaction mixtures were subjected to initial incubation at 50°C for 30 min followed by 95°C for 15 min. Amplification of the reverse transcription products was achieved by thermal cycling at 94°C for 1 min, 55°C for 30 sec and 72°C for 30 sec, for 25 cycles. A final incubation step at 72°C for 5 min was included to allow final extension of the

RT-PCR products. The RT-PCR products were analyzed by gel electrophoresis using 1% (w/v) agarose gel.

4.2.5. Dual luciferase assay

The pGL3-ACE2 IUR mutant constructs and the pRL-CMV vector (Promega Corp.) were transiently co-transfected into COS-1 cell cultures. All transfection reactions were carried out using the Fugene® reagent (Promega Corp.). The COS-1 cells were cultured in 12-well plates (Nunc Corp.) until 50% confluent prior to transfection. A volume of 1.5 µl of the Fugene® reagent was added to 38.5 µl of OptiMEM serum-free culture medium (Gibco®, Invitrogen), and the mixture was allowed to stand at room temperature for 5 min. Plasmid DNA amounted at 550 ng (containing pGL3-ACE2 IUR constructs and the pRL-CMV construct at a ratio of 10: 1) was added to the mixture. The mixture was incubated at room temperature for 15 min and applied directly to cell culture in 12-well plates. The culture medium was changed after 4 hours of incubation and the cells were allowed to recover for 26 hours. Prior to cell lysate collection, the transfected cell culture was rinsed with 0.5 ml PBS. A volume of 175 µl of 1 × passive lysis buffer (Dual Luciferase Kit, Promega Corp.) was added to the cells and cell lysis was carried out by shaking the culture plate at room temperature for 15 min. The cell lysate was collected and stored at -80°C. Transfections were repeated three times and assayed twice by the dual luciferase assay.

At the start of the dual luciferase assay, a volume of 50 µl of LARII substrate (Dual Luciferase Kit, Promega Corp.) was pre-dispensed in an U96-PP plate (Nunc Corp.) and 10 µl of cell lysate was added. The luminescence signal was immediately measured using a Glomax 96-well plate luminometer (Promega Corp.) and the luminescence signal was recorded as the firefly luciferase activity. Post recording of the firefly luciferase luminescence signal, a volume of 50 µl of Stop&Glo reagent (Dual Luciferase Kit, Promega Corp.) was added and the luminescence reading was measured. This second luminescence signal was recorded as the *Renilla* luciferase activity. The *Renilla* luciferase activity served as an internal control for the measured firefly luciferase activity. All assays were performed in duplicates. The transcription

activity of the ACE2 IUR mutant constructs were determined as the percentage of luminescence signal produced by the reporter firefly luciferase over that of the *Renilla* luciferase.

4.2.6. Scanning for putative transcription factor binding sites

The DNA sequence of the -742 to -704 region of the ACE2 IUR was scanned for putative transcription factor binding sites using two different bioinformatics web applications. The querying of two different bioinformatics databases was performed to identify distinct transcription factor binding sites. The sequence was queried against the vertebrate group of the MatInspector library (Matrix Family Library Version 7.0) by MatInspector (Cartharius *et al.*, 2005), while the TRANSFACT matrices, TRANSFACT factors and the JASPER matrices model libraries were queried using ChIP Mapper (Marinescu *et al.*, 2005b). The search results were compared and the sequence alignments of the putative binding sites to the consensus sequence were evaluated.

4.2.7. Mutagenesis and sub-cloning of C2F ΔRREB-1, C2F ΔZEB1 and C2F ΔRREB-1/ΔZEB1 constructs

The mutated C2F fragments C2F ΔRREB-1 and C2F ΔZEB1 were generated by PCR based mutagenesis reaction using Hi-Fi® high fidelity polymerase (Kapa Biosystems). The ΔRREB-1 and ΔZEB1 mutagenesis reactions were performed using the C2F-pGL3 construct as template, along with the ΔRREB-1 (fwd/rev) and ΔZEB1 (fwd/rev) primer sets (see appendix Table A.2). The ΔRREB-1/ΔZEB1 double mutation was introduced into the C2F ΔZEB1 construct using the ACE2 Pro d_mut primers (see appendix Table A.2). The mutagenesis reaction mixtures were established in a final volume of 50 µl containing final concentrations of 0.6 µM of each forward and reverse mutagenesis primers, 0.4 mM dNTP mix, 1 x Hi-Fi® buffer (Kapa Biosystems) and 1 unit of Hi-Fi® polymerase. Thermal cycling was carried out using a Hybaid PCR Sprint thermal cycler (Mandel Scientific Inc.). The thermal cycling profile consisted of an initial melt (95°C for 30 sec), followed by 20 cycles of melting at 95°C for 15 sec, annealing at 76°C for 15 sec and elongation at 72°C for 3.5 min. The annealing temperature and elongation time were selected on the basis as previously

described (See 3.2.2). A final elongation at 72°C for 6 min was included to allow strand completion. The mutagenesis reactions were analyzed by electrophoresis using 1% (w/v) agarose gel. A volume of 20 µl of successfully amplified mutagenesis product was directly incubated with 10 units of *Dpn*I restriction enzyme for 2 hours. After *Dpn*I restriction enzyme digestion, 3 µl of the digestion mixture was directly transformed into *E.coli DH5α* following a heat shock bacterial transformation protocol (see appendix A.4). Positive colonies carrying the mutated C2F construct were identified by restriction enzyme digestion using *EcoR*I and *Xba*I. Positive colonies carrying plasmid constructs with the ΔRREB-1/ΔZEB1 mutation were screened by restriction enzyme digestion using *Xba*I. The mutated C2F constructs were sub-cloned into the *Sac*I and *Hind*III sites of the vector pBluescript and subjected to DNA nucleotide sequencing (see appendix A.6) to confirm the absence of spurious mutations. The confirmed constructs were sub-cloned into the reporter vector pGL3-BASIC at the *Sac*I and *Hind*III sites. The final pGL3-BASIC constructs were propagated in *E.coli DH5α* and prepared using the Zippy® Mini-Prep Kit (Zymo Research).

4.2.8. *In vitro* binding assay

4.2.8.1. Nuclear protein extraction

Cultured HEK293 cells from four confluent T175 culture flasks (Nunc Corp.) were washed twice with 10 ml of ice cold PBS and harvested using a cell scraper. Cells were pelleted by centrifugation (3000 *g*) for 5 min at 4°C and the cell pellet was resuspended in 1 packed cell volume of buffer A (10 mM HEPES, pH 8.0; 1.5 mM $MgCl_2$; 10 mM KCl; 1 mM DTT; 1 x protease inhibitor cocktail). The protease inhibitor cocktail was the EDTA-free inhibitor cocktail set III from CalBiochem® (Merck (Pty.) Ltd.).The resuspended cells were incubated on ice for 15 min to allow swelling of cells. The swollen cells were lysed by mechanical disruption through 5 to 10 cycles of slow aspiration and rapid ejection of the cell suspension through a 21-gauge syringe needle using a 1 ml syringe. Cell lysis was monitored under light microscope. The cell homogenate was centrifuged at 4°C for 5 min using a bench-top microcentrifuge (12000 *g*) and the nuclear pellet was resuspended in 2/3 original

packed cell volume of buffer C (20 mM HEPES, pH 8.0; 1.5 mM $MgCl_2$; 25% (v/v) glycerol; 420 mM NaCl; 0.2 mM EDTA; 1 mM DTT; 0.5 mM PMSF; 1 x Protease inhibitor cocktail). The nuclear suspension was incubated at 4°C for 30 min with shaking to achieve nuclear lysis. Nuclear debris was cleared by centrifugation in a bench-top microcentrifuge for 5 min at 4°C. The supernatant containing the nuclear proteins was collected and dialysed against buffer D (20 mM HEPES, pH 8.0; 20% (v/v) glycerol; 100 mM KCl; 0.2 mM EDTA; 1 mM DTT; 0.5 mM PMSF; 1 x protease inhibitor cocktail) at a 1:100 (v/v) ratio for 2 hours. The protein concentration of the nuclear extract was determined by the Bradford assay (Bradford, 1976) using Bio-Rad Protein Assay reagent (Bio-Rad Laboratories, Inc.). The isolated nuclear extract was flash frozen with liquid nitrogen and stored at -80°C.

4.2.8.2. Probe oligonucleotide labelling

Oligonucleotides (40 µg) representing the forward and reverse complementary sequences of the -742 to -704 region of the ACE2 IUR ($^{5'}$ TTA GCT GGG CGT GGT GGT GGG CAC CTG TAG TCC CAG CT $^{3'}$ / $^{5'}$ AGC TGG GAC TAC AGG TGC CCA CCA CCA CGC CCA GCT AA $^{3'}$) were annealed by step-wise down-ramping of temperature in a Hybaid PCR Sprint thermal cycler (Mandel Scientific Inc.). The temperature down-ramping involved 15 min incubations of the oligonucleotides at 95°C, 75°C, 55°C, 35°C, 4°C. An identical annealing procedure was also carried out using the ACE2 Pro d_mut primer pair (served as mutated sequence control in electro-mobility shift assay). The annealed oligonucleotides were analysed by PAGE using a 16% gel. After visualization by ethidium bromide staining, the desired double stranded oligonucleotides were excised. The desired annealing products were eluted from polyacrylamide gel fragments by overnight incubation at 37°C with shaking in 500 µl elution buffer (0.5 M ammonium acetate; 10 mM magnesium acetate; 10 mM EDTA, pH 8.0). The eluted oligonucleotides were collected in the supernatant after centrifugation (12000 *g*) in a bench-top microcentrifuge for 1 min. The collected supernatant was subjected to sodium acetate DNA precipitation (see appendix A.7). The purified oligonucleotides were resuspended in 30 µl of nuclease-free water and the concentration determined using a NanoDrop spectrophotometer (Thermo Fisher Scientific Inc.).

The annealed oligonucleotides representing the sequence of the -742 to -704 region of the ACE2 IUR were labelled with radioactive P^{32}-ATP using T4 polynucleotide kinase (Roche Ltd.). The labelling reaction was carried out in a total reaction volume of 25 µl containing 100 ng of double stranded oligonucleotides, 2.5 µl of P^{32}-ATP (AEC-Amersham), 2.5 µl of 10 × T4 PNK forward buffer (Roche Ltd.), 18 µl of sterile nuclease-free water, and 10 units of T4 polynucleotide kinase (Roche Ltd.). The labelling reaction was allowed to proceed for 30 min at 37°C. After the labelling reaction, the T4 polynucleotide kinase was heat inactivated by incubation at 65°C for 10 min. Unincorporated P^{32}-ATP was removed by passing the labelling reaction mixture through a sephadex G25 spin column (GE Healthcare), and labelled oligonucleotides were collected in the flow through. The radioactivity of the labelled oligonucleotides was counted using a Tri-Carb liquid scintillation counter (Packard®, GMI Inc.).

4.2.8.3. Electro-mobility shift assay

The labelled oligonucleotides were incubated with varied amounts of HEK293 nuclear extract in order to determine the appropriate amount of nuclear protein needed for the binding assay. The binding reaction mixture contained 4 µl of 5 × incubation buffer (100 mM HEPES, pH 7.9; 250 mM KCl; 2.5 mM DTT; 10 mM EDTA; 5.0 mM $MgCl_2$; 20% (w/v) Ficol 400), 2 µl of the non-specific competitor poly dI/C (1 µg/µl) and varied amounts (0, 4 and 8 µg) of nuclear extract. The reaction volume was made up to 19 µl using sterile nuclease-free water. The reaction mixture was subjected to a brief centrifugation and the mixture was incubated at room temperature for 15 min. A volume of 1 µl of the labelled oligonucleotides (20000 cpm/µl) was added to the reaction mixture followed by incubation at 4°C for 30 min. A volume of 2 µl of 0.25% (w/v) bromophenol blue was added to the binding reaction mixture prior to the loading of the binding reactions onto 0.5 × TBE 5% (w/v) polyacrylamide gel. The polyacrylamide gel was equilibrated by electrophoresis at 140 V for 30 min prior to the loading of the binding reactions. Electrophoresis was carried out at 140 V for 60 min after loading of the binding reactions. At the end of the electrophoresis, the polyacrylamide gel was dried at 65 °C for 90 min using a slab gel dryer (Hoefer Inc.).

Detection was achieved by exposure of Hyperfilm™ ECL photographic film (Amersham, GE Healthcare) to the dried gel overnight at -80°C.

In order to determine the specificity of the binding reactions, the electro-mobility gel shift assay was repeated with the inclusion of unlabelled oligonucleotides and mutant oligonucleotides (ACE2 Pro d_mut primer pair). Prior to the addition of the labelled probe, unlabelled oligonucleotides at 100- or 500-fold excess to the labelled probe were added to the binding reaction followed by incubation at room temperature for 20 min. The competition assay using mutant oligonucleotides was performed using only 100-fold excess of unlabelled mutant oligonucleotides. In order to identify the factors involved in the formation of the observed DNA-protein complex, binding reaction was repeated with the incorporation of 1 µg of antibodies specific for transcription factors MAX1 (Abcam Plc.), REST1 (Abcam Plc.), RREB-1 (Rockland Inc.), ZEB1 (Santa Cruz Biotechnology, Inc.) and ZNF202 (Abcam Plc.) prior to the addition of labelled probe oligonucleotides. The identical procedure was followed for all binding reactions after the addition of 40000 cpm of labelled probe.

4.2.9. Statistical analysis

For all data sets the means and standard deviations were calculated, and where appropriate, statistical analysis was carried out using student t-test assuming equal variance.

4.3. Results

4.3.1. Construction of ACE2 IUR mutant constructs

The various ACE2 IUR mutant constructs were generated using genomic DNA extracted from HEK293 cells. The amplification of the C2, C3 and C4 fragments yielded DNA fragments corresponding to the expected size of 1189, 646 and 342 bp respectively (Fig.4.2). All constructs extended from -12 position of the ACE2 IUR to positions -1201, -658, -354 for the C2, C3 and C4 ACE2 IUR mutant fragments respectively (the translation start site of ACE2 was denoted as +1).

Fig.4.2. PCR amplification of ACE2 IUR mutant fragments from human genomic DNA. Successfully amplified C2 (1189 bp; **lane 1**), C3 (646 bp; **lane 2**) and C4 (342 bp; **lane 3**) fragments were resolved on 1% (w/v) agarose gel. **Lane M:** *Hind*III/*Eco*RI digested λ DNA marker.

The insertion of the ACE2 IUR mutant fragments into the vector pBluescript was achieved through blunt-end ligation at the *Eco*RV site and it was necessary to determine the orientations of inserted constructs by double restriction enzyme digestion using *Hind*III and *Age*I. All mutant IUR constructs contained a single *Age*I restriction site at position -292 upstream from the 3′ end (Fig.4.3). Thus, all

pBluescript vectors with an ACE2 IUR mutant fragment ligated in the forward direction would yield a 292 bp fragment when digested with *Hind*III and *Age*I.

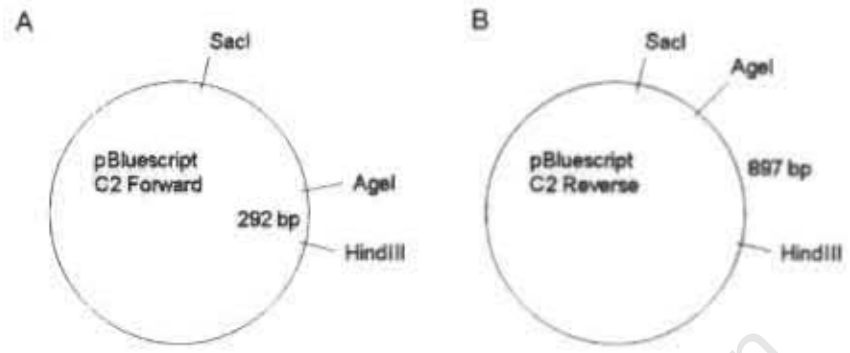

Fig.4.3. Position of the *Age*I restriction site within the ACE2 IUR mutant fragment C2 inserted into pBluescript. The inserted fragment was flanked by the *Sac*I and *Hind*III restriction sites. A. Constructs inserted in the forward orientation would yield a 292 bp fragment by double restriction enzyme digestion with *Age*I and *Hind*III. B. Constructs inserted in the reverse orientation would not yield a 292 bp fragment by double restriction enzyme digestion with *Age*I and *Hind*III

The DNA nucleotide sequencing of the pBluescript-ACE2 IUR mutant constructs revealed no spurious mutations in the inserted fragments (data not shown). The C2, C3 and C4 fragments of both orientations were successfully sub-cloned into the reporter vector pGL3-BASIC at the *Sma*I and *Hind*III sites, as indicated by the *Sac*I/*Hind*III restriction enzyme digestion (The *Sma*I site was destroyed by the blunt-end ligation) (Fig.4.4). The orientation of the inserted constructs into pGL3-BASIC was monitored using *Hind*III/*Age*I restriction enzyme digestion (Fig.4.5).

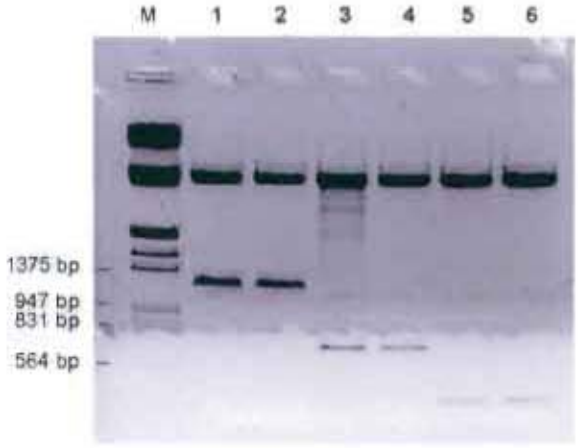

Fig.4.4. Sub-cloning of the ACE2 IUR mutant fragments into the reporter vector pGL3-BASIC. Successful sub-cloning was indicated by the excision of the insert fragment using *Sac*I and *Hind*III. Electrophoresis of the restriction enzyme digestion products on 1% (w/v) agarose gel revealed excised fragments corresponding to the expected sizes of the C2 (1189 bp; **lane 1 & 2**), C3 (646 bp; **lane 3 & 4**) and C4 (342 bp; **lane 5 & 6**) fragments. **Lane M**: *Hind*III/*EcoR*I digested λ DNA marker.

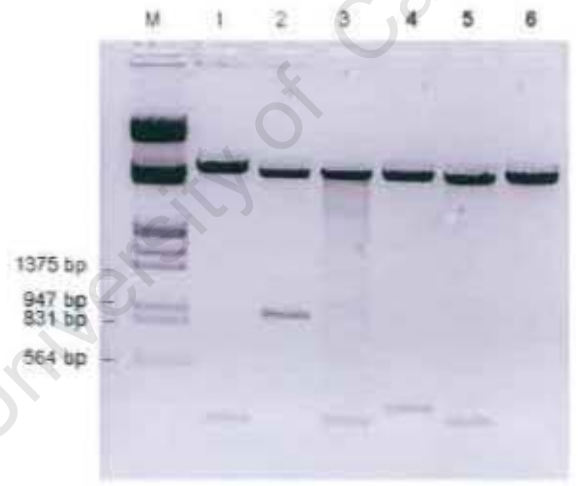

Fig.4.5. Determination of the orientation of ACE2 IUR mutant fragments inserted into the reporter vector pGL3-BASIC. The excision of a 292 bp fragment by restriction enzyme digestion using *Hind*III and *Age*I indicated the fragments were inserted in a forward orientation. (**Lane 1, 3 and 5**: C2 Forward, C3 Forward, C4 Forward constructs respectively). The *Hind*III/*Age*I restriction enzyme digestion of constructs with reversely inserted fragments yielded DNA fragments of various sizes, such as an 897 bp fragment for the C2 Reverse construct (**Lane 2**) and a 354 bp fragment for the C3 Reverse construct (**Lane 4**). The expected 50 bp fragment from the *Hind*III/*Age*I restriction enzyme digestion of the C4 Reverse construct was too small to be resolved by electrophoresis using 1% (w/v) agarose gel (**Lane 6**). **Lane M:** *Hind*III/*EcoR*I digested λ DNA marker.

The nested deletion ACE2 IUR mutant fragments (-1131; -1077; -1036; -963; -903; -837; -784; -742; -704; See Fig.4.1.2) covering the region of difference between the C2 and C3 ACE2 IUR mutant fragments were amplified from the ACE2 IUR mutant construct C2F. The amplifications yielded various nested deletion fragments (Fig.4.6). The amplified fragments were cloned into the vector pBluescript at the *Sma*I and *Hind*III sites (data not shown). Sequencing analysis of all the pBluescript-constructs revealed no spurious mutations in the inserted fragments (data not shown).

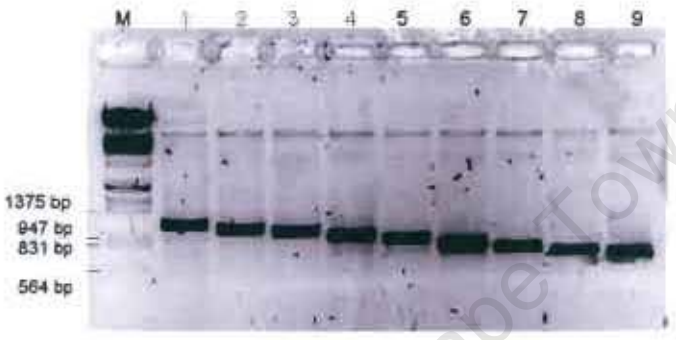

Fig.4.6. PCR amplification of ACE2 IUR nested deletion mutant fragments. The fragments amplified from the C2F construct were analyzed by electrophoresis on 1% (w/v) agarose gel. The sizes of the amplified fragments were 1119 bp (**Lane 1**), 1065 bp (**Lane 2**), 1011 bp (**Lane 3**), 951 bp (**Lane 4**), 891 bp (**Lane 5**), 825 bp (**Lane 6**), 772 bp (**Lane 7**), 730 bp (**Lane 8**) and 692 bp (**Lane 9**). **Lane M:** *Hind*III/*Eco*RI digested λ DNA marker.

The nested deletion mutant fragments were sub-cloned into the reporter vector pGL3-BASIC at the *Sac*I and *Xho*I sites (Fig.4.7). Successful sub-cloning of the nested deletion ACE2 IUR mutant fragments into the reporter vector pGL3-BASIC was demonstrated by double restriction enzyme digestion using *Sac*I and *Xho*I (Fig.4.7). The 292 bp fragment excised from all constructs by double restriction enzyme digestion using *Age*I and *Xho*I provided confirmation of the forward orientation of the inserted constructs (Fig.4.8).

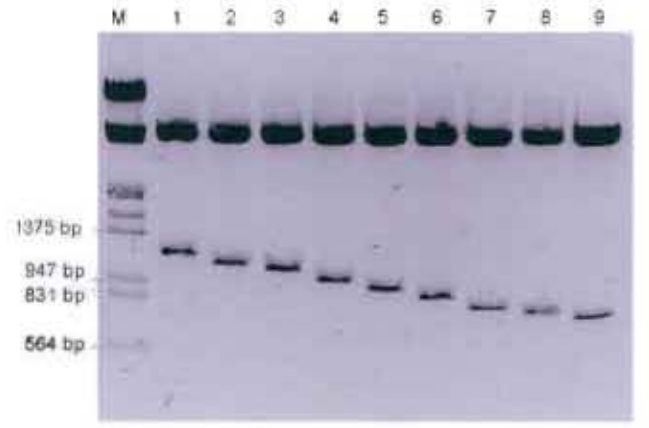

Fig.4.7. Sub-cloning of nested deletion ACE2 IUR mutant fragments into the reporter vector pGL3-BASIC. Double restriction enzyme digestion using *Sac*I and *Xho*I resulted in the excision of fragments corresponding to the expected sizes of the inserted fragments (**Lane 1 to 9**). **Lane M:** *Hind*III/*Eco*RI digested λ DNA marker.

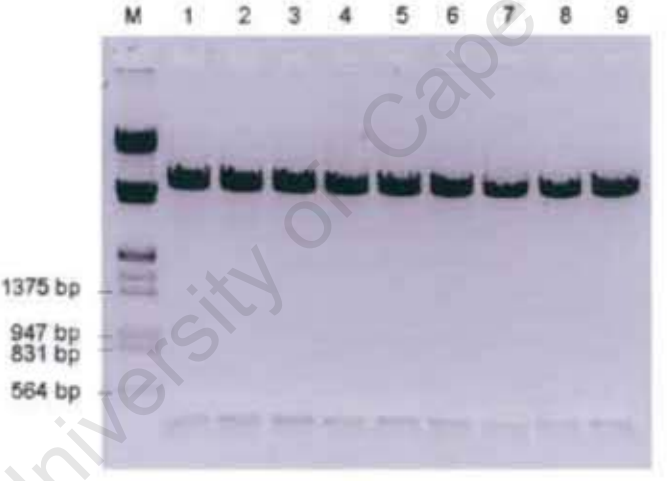

Fig.4.8. Orientation determination of nested deletion mutant ACE2 IUR constructs sub-cloned into pGL3-BASIC. The forward orientation of all inserted fragments was confirmed by the excision of a 292 bp fragment using *Age*I/*Xho*I restriction enzyme digestion (**Lane 1 to 9**). **Lane M:** *Hind*III/*Eco*RI digested λ DNA marker.

4.3.2. Reverse transcription of ACE2 mRNA

An important factor in promoter activity studies is the choice of cell line, as a low endogenous promoter drive is likely to lead to the production of low signals that are difficult to distinguish from the background. The *C.aethiops* kidney epithelial cell line Vero E6 is a cell line model chosen in previous studies to investigate the interactions between ACE2 and the SARS spike protein due to its high endogenous ACE2

expression (Li *et al.*, 2007; Zhao *et al.*, 2007). In this study, the *C.aethiops* kidney fibroblast cell line COS-1 was chosen.

In order to verify the level of ACE2 expression in COS-1 cells, reverse transcription of the ACE2 transcript was carried out on total RNA isolated from HEK293 and COS-1 cells and compared. HEK293 cells have been shown previously to possess a low level of expression of ACE2 (Warner *et al.*, 2005). As expected, a relatively high level of ACE2 transcript was detected in COS-1 total RNA while a low level of ACE2 transcript was present in HEK293 total RNA (Fig.4.9). ACE2 transcript was successfully detected in 200 ng of total COS-1 RNA while ACE2 transcript was barely detectable in 2 µg of HEK293 total RNA.

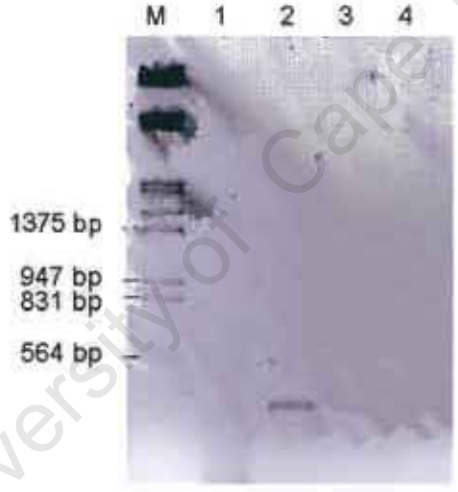

Fig.4.9. Reverse transcription PCR of ACE2 transcripts in different amounts of HEK293 and COS-1 total RNA. The primers target the region from position +129 to +544 of the ACE2 mRNA. Analysis of reverse transcription PCR product by electrophoresis on 1% (w/v) agarose gel revealed the presence of a 415 bp fragment, indicating successful detection of ACE2 mRNA. **Lane 1:** HEK293 total RNA (2 µg). **Lane 2:** COS-1 total RNA (2 µg). **Lane 3:** COS-1 total RNA (200 ng). **Lane 4:** COS-1 total RNA (20 ng). **Lane M:** *Hind*III/*Eco*RI digested λ DNA marker.

4.3.3. Dual luciferase assay

The transcription activity of the ACE2 IUR mutant constructs was determined in COS-1 cells using the dual luciferase assay. After normalization with the pRL-CMV internal control, the C2 construct displayed a lower activity than the C3 construct, while the C4 construct displayed an extremely low activity (Fig.4.10). The reversely inserted ACE2 IUR mutant constructs were found to possess low activities similar to that of pGL3 (data not shown). The 4-fold difference in the transcription activity of the C2F and the C3F constructs suggested that the presence of negative transcription regulatory elements in the -1201 to -658 region of the ACE2 IUR.

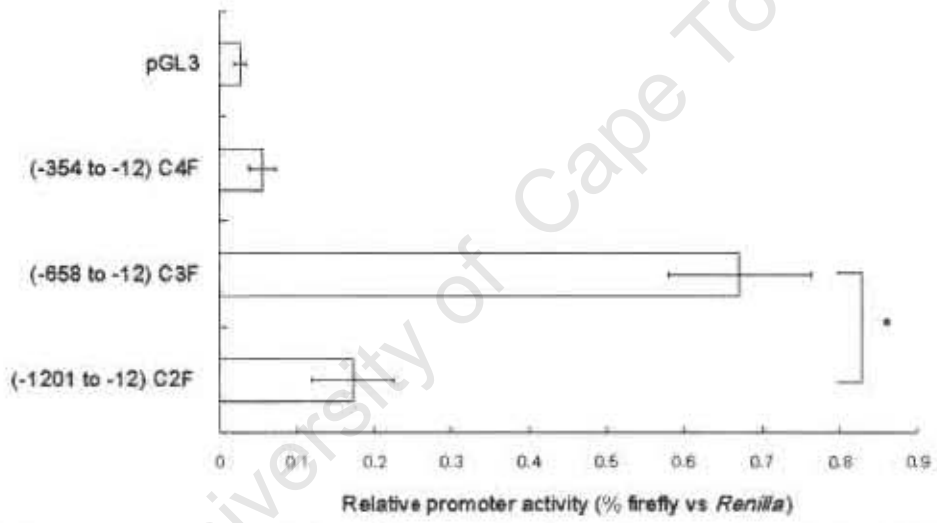

Fig.4.10. Transcription activity of the ACE2 IUR mutant constructs. The ACE2 IUR mutant construct activities were expressed as a percentage of luminescence signals measured for the firefly luciferase over the internal control *Renilla* luciferase. Numbers in bracket indicate the region of the ACE2 IUR represented by the constructs (also see Fig.4.1.1). Experiments were carried out in triplicate and the data represented by the mean with S.D. *p<0.001

In order to elucidate the negative transcription elements responsible for the observed transcription inhibition, the 543 bp region (-1201 to -658) was further analyzed through generation of nested deletion mutants of the C2F construct and their transcriptional activities assessed (Fig.4.11). The -1311 construct, which most resembled the C2F construct in size, displayed an activity almost identical to that of the C2F construct, and the -704 construct, which most resembled the C3F construct, displayed an activity that was similar to that measured for the C3F construct

(Fig.4.11). The most significant increase in transcription activity over that of the C2F construct was achieved by the -704 construct ($p<0.001$). The removal of a 38 bp sequence from position -704 to -742 resulted in a statistically significant two-fold increase in transcription activity ($p<0.005$). This two-fold increase in transcription activity was the greatest increase achieved amongst all the nested deletion ACE2 IUR mutant constructs (Fig.4.11). This observation indicated that the region from position -742 to -704 contributed most significantly to the repression of the transcription activity of the C2F construct.

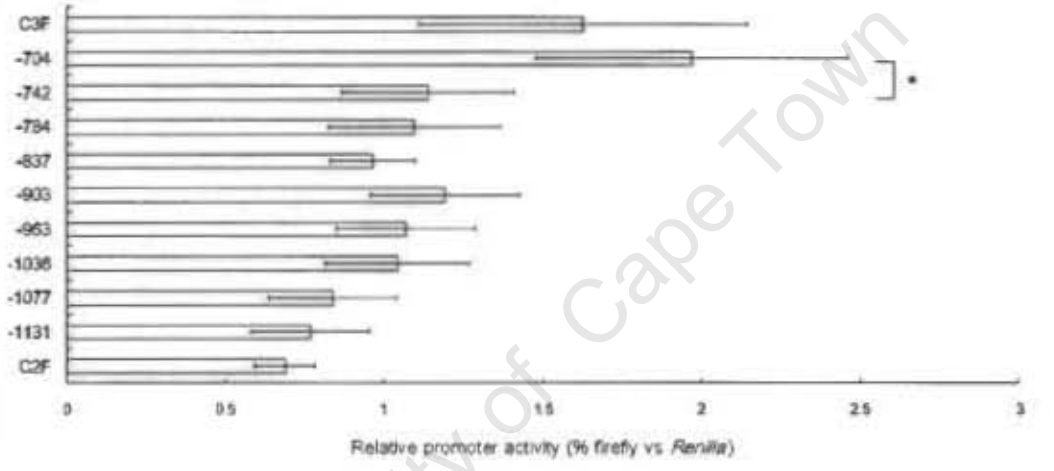

Fig.4.11. Transcription activity of the nested deletion ACE2 IUR mutant constructs. The relative sizes of the various ACE2 mutant IUR constructs are shown in Fig.4.1.2. The activities of the ACE2 IUR mutant constructs were expressed as a percentage of fluorescence signal measured for the firefly luciferase over the internal control *Renilla* luciferase. The largest shift in activity compared to the C2F construct was observed for the -704 construct. The removal of the sequence from position -742 to -704 resulted in the most significant increase in transcription activity. Experiments were carried out in triplicate and the data represented by the mean with S.D. *$p<0.005$

4.3.4. Scanning of putative transcription factor binding sites

The region from position -742 to -704 of the ACE2 IUR was scanned for putative transcription binding sites using bioinformatics web applications MatInspector and ChIP Mapper. A number of four potential transcription binding sites were identified by MatInspector (Cartharius *et al.*, 2005) (Table 4.1) while ten entries were retrieved by ChIP Mapper (Marinescu *et al.*, 2005a) (Table 4.2).

Table 4.1. Putative transcription binding sites within the -742 to -704 region of the ACE2 IUR returned by the MatInspector web application (Cartharius *et al.*, 2005).

Family	Family Information	Matrix	Position (from–to)	Str.	Core sim.	Matrix sim.	Sequence (capitals: core sequence)
EGRF	EGR/nerve growth factor induced protein C & related factors	EGR3.01	5 - 21	(+)	1.000	0.782	ctggGCGTggtggtggg
ZBPF	Zinc binding protein factors	ZNF202.01	5 - 27	(-)	1.000	0.753	caggtgCCCAccaccac gcccag
MYOD	Myoblast determining factors	MYOGENIN.01	17 - 33	(-)	1.000	0.924	ggaetACAGgtgcccac
ZFHX	Two-handed zinc finger homeodomain transcription factors	ZEB1.03	18 - 30	(+)	1.000	0.968	tgggCACCtgtag

Table 4.2. Putative transcription binding sites within the -742 to -704 region of the ACE2 IUR returned by the ChIP Mapper web application (Marinescu *et al.*, 2005a).

Factor	Strand	Start	End	Score	E- value
NRSF form 2	-	1	21	2.1	16
RREB-1	-	2	21	1.5	11
Sp1	+	6	15	2.7	14
SP1	+	7	15	2.4	2.5
Ahr-ARNT	+	9	13	1.6	15
ZEB1	+	19	30	4.9	8.8
Lmo2 complex	-	19	27	4.4	18
Sn	-	19	29	4.8	13
Max1	+	22	27	1.4	14
Snail	-	22	27	1.9	18

The transcription factor ZEB1 was retrieved by both bioinformatics tools and the ZEB1 binding sequence had the highest matrix similarity and the best score. Overlap of the putative transcription binding sites in the -742 to -704 region was noted in both

sets of the search results. In the result set retrieved using the MatInspector, the putative transcription factor binding sites splits the -742 to -704 region into two sections. While many entries were retrieved using the ChIP Mapper, the putative binding sites also focused on two sections of the -742 to -704 region. Sequence alignment of several identified putative transcription factor binding sites revealed the focus of putative binding sites in the -742 to -722 and -721 to -716 regions (Fig.4.12).

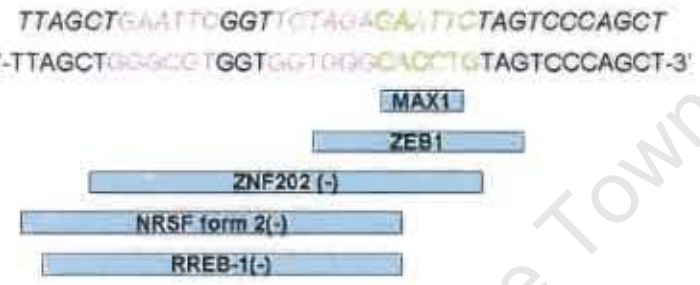

Fig.4.12. Sequence of the -742 to -704 region of the ACE2 IUR identified to contain putative transcription repressors. Blue bars indicate the positions of the putative binding sites of the various transcription factors. The (-) sign denotes the occurrence of the putative binding site on the reverse DNA strand. Sequence on top shown in italic indicates the mutations introduced to disrupt the various putative binding sites. Nucleotide changes introduced into the mutant construct ΔRREB-1 are shown in pink. Nucleotide changes introduced into the mutant construct ΔZEB1 are shown in green.

4.3.5. Mutagenesis of ACE2 IUR -742 to -704 region

The bioinformatics analysis of the ACE2 IUR -742 to -704 region yielded several putative transcription factor binding sites located within the sequence, as listed in Table 4.2 and Table 4.3. A closer examination of the putative binding site positions revealed considerable overlapping of the binding sites (Fig.4.12.). Large portions of ZNF202, REST/NRSF form 2 and RREB-1 binding sites overlap with each other while the MAX1 binding site is incorporated within the ZEB1 binding site. Mutations in the overlapping nucleotide sequences (highlighted sequence in Fig.4.12) to common restriction sites (*Eco*RI and *Xba*I sites) were introduced into the C2F construct to assess the contribution of the various putative binding sites to the observed transcription repression of the C2F construct.

Successful mutagenesis of the C2F-pGL3 construct was confirmed by restriction enzyme digestion. Since the putative transcription factor binding sites were disrupted with the introduction of *EcoR*I and *Xba*I restriction sites, the restriction digest patterns of the ΔRREB-1 and ΔZEB1 mutants were expected to be different to that of the unmutated C2F-pGL3 construct. The ΔRREB-1 mutagenesis primers introduced new *EcoR*I and *Xba*I restriction sites and digestion with restriction enzyme *Xba*I resulted in the observation of the anticipated 338 bp and 3604 bp fragments for the C2F ΔRREB-1 mutant construct (Fig.4.13). The *EcoR*I restriction enzyme digestion of C2F-pGL3 resulted in the appearance of 5390 bp and 611 bp fragments, and the 611 bp fragment was reduced to a 447 bp and a 164 bp fragment with the introduction of an extra *EcoR*I site in the ΔZEB1 mutant (Fig.4.13). The *Xba*I restriction enzyme digestion patterns of the C2F-pGL3 and C2F ΔZEB1 constructs were identical (Fig.4.13), indicating the correct mutation had been introduced in the C2FΔZEB1 construct.

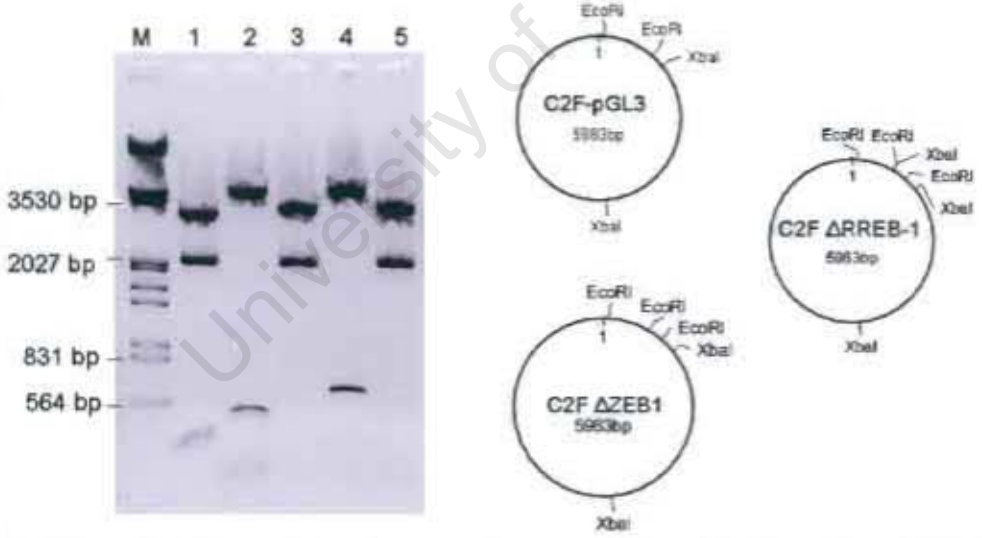

Fig.4.13. Mutagenesis of the C2F-pGL3 construct to introduce the ΔRREB-1 and ΔZEB1 mutations. Correct introduction of desired mutations were monitored by restriction enzyme digestion followed by electrophoresis on 1% (w/v) agarose gel. **Lane 1:** C2F ΔRREB-1 construct digested with restriction enzyme *Xba*I yielded 3604 bp, 2059 bp and 338 bp fragments. **Lane 2:** *EcoR*I restriction enzyme digestion pattern of C2F ΔZEB1 included 5390 bp, 447 bp and 164 bp fragments. **Lane 3:** Restriction enzyme digestion of C2F ΔZEB1 with *Xba*I restriction enzyme produced 3942 bp and 2059 bp fragments. **Lane 4:** Unmutated C2F-pGL3 construct digested with *EcoR*I restriction enzyme yielded 5390 bp and 611 bp fragments. **Lane 5:** Unmutated C2F-pGL3 construct digested with *Xba*I restriction enzyme yielded 3942 bp and 2059 bp fragments. **Lane M:** *Hind*III/*EcoR*I digested λ DNA marker.

The disruption of all putative transcription factor binding sites in the -742 to -704 region of the ACE2 IUR was achieved by construction of the C2F ΔRREB-1/ΔZEB1 mutant using the C2F ΔZEB1 construct as the template. The C2F ΔRREB-1/ΔZEB1 mutant fragment was sub-cloned into pBluescript and pGL3-BASIC at the *Sac*I and *Xba*I sites. In both vectors, the *Sac*I restriction site is located 5′ of the inserted fragment, thus an 815 bp fragment is excised from unmutated C2F-pGL3 construct by *Sac*I/*Xba*I double restriction enzyme digestion (data not shown). The additional *Xba*I restriction site introduced by the mutagenesis primers led to the production of 477 bp and 338 bp fragments by *Sac*I/*Xba*I double restriction enzyme digestion (Fig.4.14). The absence of the 815 bp fragment indicated correct introduction of the ΔRREB-1/ΔZEB1 mutations.

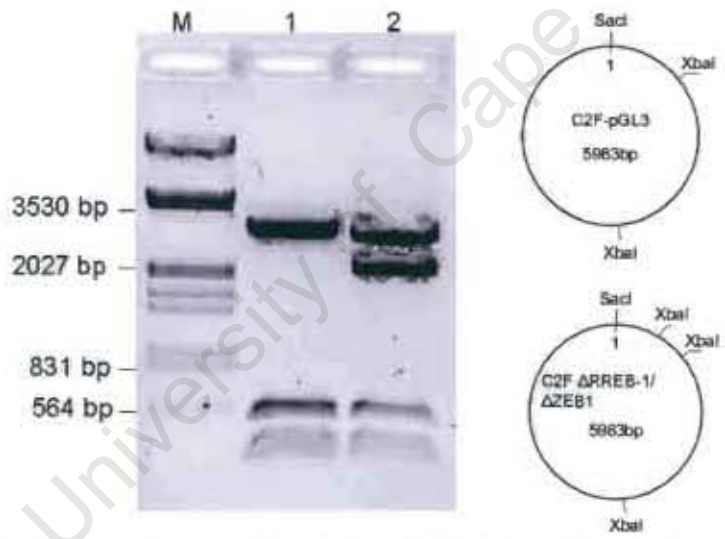

Fig.4.14. Sub-cloning of the C2F ΔRREB-1/ΔZEB1 mutant fragment. Restriction enzyme *Sac*I/*Xba*I digestion of the unmutated C2F construct would produce an 815 bp fragment (top plasmid map on right). Introduction of an extra *Xba*I restriction site by the ACE2 Pro d_mut mutagenesis primers resulted in the reduction of the 815 bp fragment to 477 bp and 338 bp fragments after *Sac*I/*Xba*I double restriction enzyme digestion. **Lane 1:** C2F ΔRREB-1/ΔZEB1 inserted into pBluescript. **Lane 2:** C2F ΔRREB-1/ΔZEB1 inserted into pGL3-BASIC. **Lane M:** *Hind*III/*Eco*RI digested λ DNA marker.

In the assessment of transcription activity of the mutated C2F constructs, the disruption of the putative binding sites in the C2F ΔRREB-1 and the C2F ΔZEB1 constructs resulted in transcription activities of approximately 1.5-fold higher than that of the C2F-pGL3 construct (Fig.4.15.1). Upon disruption of all putative

transcription factor binding sites in the ACE2 IUR -742 to -704 region (represented by the ΔRREB-1/ΔZEB1 mutant construct), a 1.63-fold increase in transcription activity was observed (Fig.4.15.2). Together, these results suggest the involvement of the transcription factor binding sites within the -742 to -704 region in transcriptional repression. In order to identify the transcription factors binding to the -742 to -704 region, subsequent *in vitro* DNA binding assays were conducted.

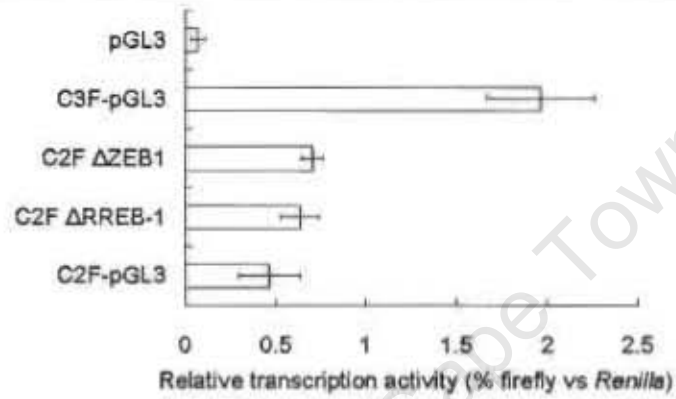

Fig.4.15.1. Transcription activity of mutated C2F constructs with disrupted putative transcription factor binding sites in the ACE2 IUR -742 to -704 region. Both the C2F ΔRREB-1 and C2F ΔZEB1 mutant constructs displayed slight but statistically insignificant increases in transcription activity compared to that of the C2F-pGL3 construct. Experiments were carried out in triplicate and the data represented by the mean with S.D.

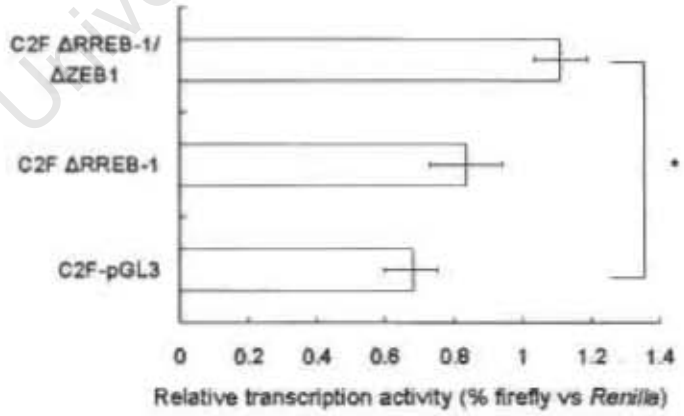

Fig.4.15.2. Transcription activity of mutated C2F constructs carrying disrupted putative transcription factor binding sites in the ACE2 IUR -742 to -704 region. The disruption of all putative transcription binding sites in C2F ΔRREB-1/ΔZEB1 mutant construct resulted in a statistically significant increase in the transcription activity of the construct. Experiments were carried out in triplicate and the data represented by the mean with S.D. *$p < 0.005$.

4.3.7. *In vitro* binding assay

An *in vitro* binding assay such as the electro-mobility shift assay is a simple, effective way to determine the binding of nuclear factors to a DNA sequence of interest. In this study, the double stranded oligonucleotide representing the -742 to -704 region of the ACE2 IUR was obtained by the annealing of synthesized oligonucleotides. The change in mobility of the annealed oligonucleotides (Lane 3 and 6, Fig.4.16.1) compared to that of the single stranded oligonucleotides (Lane 1, 2, 4 and 5, Fig.4.16.1) on a 16% (w/v) polyacrylamide gel allowed for identification of double stranded oligonucleotides to be excised from the gel and purified. Annealed wild-type and mutant oligonucleotides of high purity (Fig.4.16.2) were obtained and stored at a concentration of 100 ng/µl.

Fig.4.16.1. Annealed oligonucleotides for *in vitro* binding assay separated on 16% (w/v) polyacrylamide gel. The mobility of double stranded annealed oligonucleotides was distinct to that of the single stranded oligonucleotides. **Lane 1 & 2:** Forward and reverse ACE2 Pro d_mut mutagenesis primers. **Lane 3:** Annealed ACE2 Pro d_mut mutagenesis primers. **Lane 4 & 5:** Forward and reverse strand of the ACE2 IUR -742 to -704 sequence. **Lane 6:** Annealed ACE2 IUR -742 to -704 oligonucleotides.

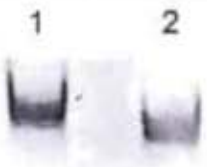

Fig.4.16.2. Purified annealed oligonucleotides for *in vitro* binding assay. **Lane 1:** ACE2 Pro d_mut mutagenesis primers. **Lane 2:** The ACE2 IUR -742 to -704 oligonucleotides.

The annealed ACE2 IUR -742 to -704 oligonucleotides (100 ng) were utilized in the labelling reaction. The radioactivity of the labelled oligonucleotides was determined as 1.7×10^6 counts per min (cpm) per µl (specific activity of 4.2×10^8 cpm per µg oligonucleotides) by scintillation counting. The incubation of the labelled DNA probe with HEK293 nuclear extract resulted in the formation of DNA protein complexes (Fig.4.17).

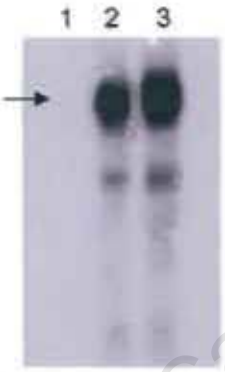

Fig.4.17. Titration of nuclear extract required for electro-mobility shift assay. In the absence of nuclear extract, no DNA protein complex was observed (**Lane 1**). The amount of DNA protein complex formed was increased with increased amounts of protein extract used in the binding assay. The focus of the study was placed on the larger DNA protein complex (indicated by arrow). **Lane 1:** No nuclear extract control. **Lane 2:** HEK293 nuclear extract (4 µg). **Lane 3:** HEK293 nuclear extract (8 µg).

Pre-inclusion of unlabelled oligonucleotides at 100-fold excess in the binding assay resulted in the reduced formation of the observed DNA protein complex (lane 2, Fig.4.18), and the pre-inclusion of unlabelled oligonucleotides at 500-fold excess completely prevented the formation of the DNA protein complex (lane 3, Fig.4.18). The specificity of the binding reaction was further confirmed by the inability of the unlabelled mutant oligonucleotides at 100-fold excess to compete with the binding of nuclear factors to the labelled probe (lane 4, Fig.4.18). Binding assays were also carried out with the inclusion of different antibodies to identify the factors bound to the labelled probe. Should a transcription factor interact with the labelled probe, inclusion of antibodies in the binding assay would result in either a super-shift (formation of a larger complex) or a loss of binding. The inclusion of the MAX1, REST and RREB-1 antibodies in the binding assay did not influence the formation of the observed DNA protein complex, indicating these factors do not bind the labelled

probe. The inclusion of the ZEB1 antibody resulted in the appearance of a larger super-shifted complex, indicating binding of the ZEB1 factor to the labelled probe (lane 8, Fig.4.18). Disappearance of the original DNA protein complex was observed for the binding reaction in which the ZNF202 antibody was included (lane 9, Fig.4.18). Competition of the ZNF202 antibody with the labelled probe to prevent DNA protein complex formation indicated an interaction between ZNF202 and the labelled probe. Moreover, the ZNF202 antibody prevents the formation of the original DNA protein complex by blocking the DNA binding site of ZNF202, as the antigen used to raise the antibody was a part of the ZNF202 DNA binding zinc finger domain (Abcam Plc., cat# ab58179). These findings suggest ZEB1 and ZNF202 bind to the -742 to -704 region of the ACE2 gene to potentially repress ACE2 transcription.

Fig.4.18. Electro-mobility shift assay performed to confirm the specificity of observed DNA protein bindings, and the identification of factors involved in the DNA protein complex formation. **Lane 1:** HEK293 nuclear extract (4 µg). **Lane 2:** Unlabelled oligonucleotides competition (100 × excess). **Lane 3:** Unlabelled oligonucleotides competition (500 × excess). **Lane 4:** Unlabelled mutant oligonucleotides competition (100 × excess). **Lane 5:** MAX1 antibody inclusion. **Lane 6:** REST antibody inclusion. **Lane 7:** RREB-1 antibody inclusion. **Lane 8:** ZEB1 antibody inclusion. **Lane 9:** ZNF202 antibody inclusion.

4.4. Discussion

In order to gain further insight into the transcription control of ACE2 , a search for potential transcription regulatory elements of ACE2 was conducted on the region immediately upstream of the translation start codon of the ACE2 gene. This included the untranslated 5′ exon and intron 1 of the ACE2 gene (Itoyama *et al.*, 2005). The initial assessment of the transcription activity of the ACE2 IUR mutant constructs by dual luciferase assays indicated the presence of negative regulatory elements binding to the region -1201 to -658 upstream of the ACE2 translation start codon (Fig.4.10). Further transcription activity analysis of the nested deletion ACE2 IUR mutant constructs representing the -1201 to -658 region (Fig.4.1.2) showed the most significant increase in transcription activity amongst the nested deletion mutants was achieved by deletion of the region from -742 to -704 (Fig.4.11). The bioinformatics screening for putative transcription factor binding sites in the -742 to -704 region revealed the presence of several putative binding sites, and the disruption of these putative binding sites led to a statistical significant increase in the transcription activity of the ACE2 IUR mutant construct C2F (Fig.4.15.2). The resultant increase in transcription activity from the disruption of a combination of putative transcription factor binding sites (Fig.4.15.2) suggested the binding of more than one repressor element to the -742 to -704 region. The formation of DNA-protein complexes after co-incubation of labelled ACE2 IUR -742 to -704 oligonucleotides with HEK293 nuclear factors (Fig.4.17) confirmed the binding of transcription factors *in vitro* to the ACE2 IUR -742 to -704 region. Two of such transcription factors were identified as ZNF202 and ZEB1 (Fig.4.18).

The factor ZNF202 is a zinc finger protein that has eight C2H2 zinc finger domains and 25 potential phosphorylation sites (Monaco *et al.*, 1998). The involvement of ZNF202 in the expression regulation of genes related to lipid metabolism and vascular maintenance has been established *in vitro* (Wagner *et al.*, 2000). In HepG2 human hepatoma cells, the over-expression of ZNF202 resulted in a 5- to10-fold reduction in the expression of reporter constructs driven by the promoters of *apoAIV* and *apoE* (*apoAIV* and *apoE* codes for apolipoproteins AIV and E respectively) (Wagner *et al.*, 2000). In the same study, ZNF202 was also reported to bind to the promoter of the vascular endothelial growth factor (VEGF) gene, albeit the transcription regulatory

effect of ZNF202 on VEGF expression was not determined. Studies on VEGF have indicated that VEGF enhances atherosclerotic plaque progression in cholesterol-fed apolipoprotein E/apolipoprotein B100 KO mice (Celletti *et al.*, 2001). In this regard, single nucleotide polymorphisms in the ZNF202 promoter have been proposed to predict severe atherosclerosis, as altered transcription of ZNF202 brought about by different ZNF202 promoter variants was speculated to alter the expression of VEGF (Stene *et al.*, 2008). Interestingly, overexpression of ACE2 in atherosclerosis-prone *apoE* KO mice resulted in less oil red-O-positive stained areas and less macrophage infiltration in the aorta, indicating limited endothelial dysfunction and reduced atherosclerosis (Lovren *et al.*, 2008). Given that ACE2 promotes proper endothelial function, it is speculated that ZNF202 also regulates ACE2 expression in a similar context as it does for other genes involved in vascular maintenance. While the *in vitro* binding of ZNF202 to many gene promoters has been reported (Wagner *et al.*, 2000), this study is a first to suggest the binding of ZNF202 to the ACE2 promoter.

The other transcription factor found to bind the -742 to -704 region upstream of the ACE2 translation start site is ZEB1. ZEB1 is a zinc finger E-box binding transcription repressor which has been reported to promote the epithelial to mesenchymal transition during development and cancer metastasis (Aigner *et al.*, 2007; Spaderna *et al.*, 2008). Strong repression of tumour suppressors and genes involved in the maintenance of polarized epithelial phenotype by ZEB1 in various cancer cells has been documented (Drake *et al.*, 2009; Adachi *et al.*, 2009; Clarhaut *et al.*, 2009; Spaderna *et al.*, 2008). Given the prominent involvement of ZEB1 in the regulation of genes associated with developmental events such as epithelial to mesenchymal transition, it is postulated that ZEB1 also regulates the developmental expression of ACE2, as observed in different animal models (Tikellis *et al.*, 2006; Wiener *et al.*, 2007).

While ZEB1 and ZNF202 bind *in vitro* to the -742 to -704 region upstream of the ACE2 translation start site, consensus binding sequence of the two transcription factors could occur elsewhere in the ACE2 IUR. A search through the rest of the ACE2 IUR sequence revealed no other occurrences of the consensus binding sites for ZEB1 and ZNF202 (data not shown). It is also noteworthy to mention that the nested deletion mutant constructs were only generated for the -1201 to -658 region (Fig.4.1.2), even though the removal of the -658 to -354 region in the ACE2 IUR

mutant construct C4 also led to a change in reporter gene expression (Fig.4.10). This observation suggested the presence of positive transcription regulators within the -658 to -354 region. The presence of both negative and positive transcription elements in the immediate upstream region of the ACE2 gene implied that the 5′ UTE and intron 1 of ACE2 contribute significantly towards the transcription regulation of ACE2.

Chapter 5
Conclusions and Future directions

Much research attention has been focused on ACE2 since it was first cloned in 2000 (Donoghue *et al.*, 2000; Tipnis *et al.*, 2000), and different aspects of ACE2 ranging from its physiological functions to its biochemical characteristics have been successfully elucidated. The activity of ACE2 has been reported to provide protection against various tissue injuries in many animal models (Imai *et al.*, 2005; Oudit *et al.*, 2006; Tikellis *et al.*, 2003; Wong *et al.*, 2007; Yamamoto *et al.*, 2006), however there is plenty that remains unclear in respect of the control of ACE2 activity. While the structure-function of the ACE2 S_1 and S_1' sub-sites has been extensively investigated (Deaton *et al.*, 2008a; Deaton *et al.*, 2008b; Mores *et al.*, 2008; Towler *et al.*, 2004), the relevance of the ACE2 S_2 sub-site remains elusive.

The structure-function of the S_2 sub-site of ACE2 was investigated in this study using the novel phosphinic ACE2 inhibitor 416F2. Inhibition of ACE2 by 416F2 was characterized using the dose-dependent response plot and the direct linear plot methods with synthetic ACE2 substrates Mca-APK-Dnp and Mca-YVADAPK-Dnp, and the estimated K_i values of low nanomolar range (average K_i of 27.91 nM and K_{ic} of 31.10 nM) were indicative of potent inhibition. The strong ACE2 inhibitory action of 416F2 is supported by a separate study where 416F2 was reported to inhibit ACE2 with a sub-nanomolar inhibition constant of 0.13 nM (Mores *et al.*, 2008).

In order to elucidate the S_2 sub-site, a molecular model of 416F2 structure docked into the ACE2 crystal structure (PDB code: 1R4L) was generated by *in silico* molecular dynamic simulation. In the docked model, the P_2-leucyl group of 416F2 was observed to extend into a large cavity beyond the S_1 sub-site and it assumed a variety of poses. Despite the strong inhibition of ACE2 displayed by 416F2, no interactions between the ACE2 active site residues and the P_2-leucyl group of 416F2 were observed in the docked model, and hence the S_2 sub-site was not defined. The non-bulky, hydrophobic nature of the P_2-leucyl group of 416F2 provides limited opportunities for extensive interactions with ACE2 active site residues. In future work, molecular

docking of tripeptide inhibitors with bulky or charged P_2 side chains into the ACE2 structure could reveal interactions between P_2 side chains and active site residues that could not be formed by a P_2-leucyl group. The possibility of induced-fit ligand binding could also be considered in future molecular docking experiments. However, the approach of flexible-protein receptor docking requires a consideration of the ligand binding conformation dynamics of ACE2, and appropriate control needs to be devised to ensure the integrity of the ACE2 active site. As an alternative approach to molecular docking, crystallographic studies of ACE2 co-crystallized with tripeptide inhibitors could also be conducted to elucidate the ligand P_2 side chain interactions with the ACE2 active site residues. Subsequent to the elucidation of P_2/S_2 residue interactions, mutagenesis experiments could be carried out to assess the significance of these interactions in the enzymatic activity of ACE2.

In the 416F2-ACE2 docked model in this study, it was noted that the terminal hydroxyl group of Tyr510 extended beyond the S_1 sub-site space and formed a hydrogen bond with the acetyl carbonyl group of 416F2. While Tyr510 has been defined by many studies as a component of the S_1 sub-site which contributes towards the S_1 selectivity for small hydrophobic P_1 residues (Deaton *et al.*, 2008a; Mores *et al.*, 2008; Towler *et al.*, 2004), its significance in ACE2 activity has not been characterized. The mutagenesis of Tyr510 carried out in this study has shown that the hydrophobic surface and the steric hindrance provided by Tyr510 is essential for ACE2 activity, as the substitution of Tyr510 with threonine resulted in the abolishment of enzymatic activity. It is postulated that the bulky hydrophobic ring of Tyr510 not only confers the substrate selectivity for small hydrophobic P_1 residues, but also contributes to the stabilization required for the formation of catalytic intermediates during the hydrolysis process by limiting the movement of ligand P_1 residues. To date, no residues of the S_1 sub-site have been reported to be essential for ACE2 activity (Guy *et al.*, 2005), and Tyr510 is the first S_1 sub-site residue described as a critical residue. In contrast, the terminal hydroxyl of Tyr510 was found to be non-essential for ACE2 activity, because enzymatic activity was retained by an ACE2 mutant with the substitution of Tyr510 with phenylalanine. The hydrogen-bonding capability of the Tyr510 terminal hydroxyl has been suggested to confer a subtle substrate selectivity by another study (Mores *et al.*, 2008). The hydrogen bond observed between the 416F2 acetyl carbonyl group and the hydroxyl group of Tyr510

found in the ACE2-416F2 docked model is suggestive of the involvement of the Tyr510 hydroxyl group in the coordination of ligand backbone. Nevertheless, the specific structure-function of the Tyr510 terminal hydroxyl group warrants further investigation through a detailed enzyme kinetic analysis.

The regulation of ACE2 activity through its expression was also investigated in this thesis. To elucidate the factors involved in the regulation of ACE2 expression, a search for potential transcription regulators binding to the region immediately upstream of the ACE2 translation start codon was conducted. Analyses of nested deletion mutant constructs and *in vitro* binding experiments showed that two transcription factors, namely ZEB1 and ZNF202, bind to the region -742 to -704 upstream of the ACE2 translation start codon. The binding of these two transcription factors was associated with repressed transcription activity as disruption of the putative binding sites in the ACE2 IUR construct C2F resulted in elevated reporter luciferase expression. The established role of ZNF202 as a transcription repressor that binds to the promoter of genes involved in vascular maintenance, such as VEGF (Wagner *et al.*, 2000), suggests that ACE2 expression is regulated by ZNF202 owing to its epithelial protective role (Lovren *et al.*, 2008; Rentzsch *et al.*, 2008). On the other hand, it is postulated that ZEB1 binds to the ACE2 promoter to repress ACE2 expression during development. In animal models, expression of ACE2 has been shown to be developmentally regulated in kidney and lung tissues (Tikellis *et al.*, 2006; Wiener *et al.*, 2007). The involvement of ZEB1 in the regulation of developmental events such as epithelial-mesenchymal transition poises ZEB1 to repress ACE2 expression during development.

In order to fully establish the roles of ZNF202 and ZEB1 as transcription repressors of ACE2, further investigations need to be conducted. The low expression level of ACE2 in HEK293 makes the cell line ideal for the investigation into the *in vivo* binding of ZEB1 and ZNF202 by chromatin immunoprecipitation. Future *in vivo* binding data would greatly complement the *in vitro* binding data presented in this study. In addition, the functional significance of ZNF202 and ZEB1 as regulators of ACE2 expression could be assessed using siRNA knock-down experiments. Subsequently, the significance of ZNF202 regulated ACE2 expression in pathological condition such

as atherosclerosis could be assessed to determine any relationship between ZNF202 and ACE2 expressions in atherosclerosis.

The region immediately upstream of the ACE2 translation start codon investigated in this study did not include the ACE2 promoter region. The ACE2 promoter region -1203 to -3380 upstream of the translation start codon was investigated by a collaborator of the laboratory. The nested deletion promoter assays performed by our collaborator revealed no significant finding as similar transcription activities were determined for the nested deletion constructs and a gradual loss of transcription activities was observed with the two shortest constructs (Amy Cilia La Corte, personal communication). It was noted that given the precedent of intronic and promoter transcription elements interacting with each other to form complex regulatory mechanisms (Girousse *et al.*, 2009; Jones *et al.*, 2007), it is likely that ZEB1 and ZNF202 interact with other transcription elements in the promoter region to fine tune ACE2 expression. Investigation of the promoter, 5′ UTE and intron 1 regions of the ACE2 gene together as a whole would facilitate the elucidation of the functional significance of ZEB1 and ZNF202 in the transcription regulation of ACE2. Furthermore, such investigation could also identify other novel transcription regulatory mechanisms controlling ACE2 expression.

Overall, the findings of this study contribute towards a better understanding of the control of ACE2 activity at both the structure-function and the transcription level. Knowledge of the structure-function relationship of ACE2 could be exploited in the design of more potent and selective novel ACE2 inhibitors, despite the fact that the therapeutic benefits of ACE2 inhibition are currently unclear. At the transcription level, a better understanding of the ACE2 transcription regulation mechanism could reveal opportunities where ACE2 transcription control could be capitalized on to achieve desirable benefits associated with ACE2 activity. However, in order to validate ACE2 as a meaningful therapeutic target, further elucidation of various aspects of ACE2 is still necessary.

Appendix

A.1. Bacterial culturing

The bacteria *Escherichia coli DH5*, strain (F^-, φ80d*lacZ*ΔM15, (*lacZYA-argF*)U169, *deoR*, *recA*1, *endA*1, *hsdR*17(rk^-, mk^+), *phoA*, *supE*44, λ^-, *thi*-1, *gyrA*96, *relA*1) was used for the propagation of DNA plasmids. All bacteria cultures were grown in Luria Broth culture medium (LB; 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl) medium at 37°C with orbital shaking (150 RPM).

A.2. Crude small scale plasmid preparation

A volume of 1 ml of overnight *E.coli* culture was collected in Eppendorf tube by centrifugation at maximum speed for 2 min using a bench-top microcentrifuge. The supernatant was removed and the collected bacterial cell culture was resuspended in 70 µl of cell lysis buffer (10 mM Tris, pH 8.0; 1 mM EDTA; 15% (w/v) sucrose; 100 µg/ml BSA; 20 µg/ml ribonuclease A; 2 mg/ml lysozyme). The resuspended culture was incubated at 37°C for 30 min, followed by boiling for 1 min. The boiled culture resuspension was incubated on ice for 10 min. The cellular debris was cleared by centrifugation at maximum speed for 10 min using a bench-top microcentrifuge.

A.3. Preparation of heat shock competent E.coli DH5α

An over night culture of *E.coli DH5*α was inoculated into 50 ml of LB at 1: 50 dilution. The culture was incubated at 37°C until an OD_{595} of 0.4 was reached. The culture was chilled on ice for 15 min before the cells were collected by centrifugation at 1000 × *g* for 5 min. The cell pellet was resuspended in 1/10 original volume of pre-chilled TFBI (30 mM KOAc; 100 mM RbCl; 10 mM $CaCl_2$; 50 mM $MnCl_2$; 15% (v/v) glycerol; adjusted to pH 5.8 with glacial acetic acid). The resuspended cells were incubated on ice for 15 min before subjected to centrifugation at 1000 × *g* for 5 min. The collected cells were resuspended in 1/25 original volume of pre-chilled TFBII (10 mM MOPS; 10 mM $RbCl_2$; 75 mM $CaCl_2$; 15% (v/v) glycerol; adjusted to pH 6.5 with glacial acetic acid). The final resuspension was divided into 100 µl fractions

followed by flash freeze using liquid nitrogen. The competent cells were stored at -80°C until used.

A.4. E.coli transformation (Ausubel *et al.*, 1992)

Aliquots of prepared competent cells were thawed on ice. An amount of 50 to 100 ng plasmid DNA was added to 100 µl of competent cells and the mixture was incubated on ice for at least 25 min before heat shock treatment. The heat shock process was carried out at 42°C for 45 sec. The heat shocked cells were immediately placed on ice and a volume of 900 µl of LB was added to the cells after 1 min incubation on ice. The cells were allowed to recover at 37°C for 60 min with shake, and 100 µl of the transformed cells were plated on LB plates containing 50 µg/ml ampicillin.

A.5. Restriction enzyme digestion and DNA electrophoresis

All restriction enzyme digestions were carried out in a total volume of 20 µl containing 1 × restriction enzyme buffer recommended by the manufacturer of the restriction enzyme. The amount of restriction enzyme used in each digestion ranged from 1 to 10 units. Restriction enzyme digestions were allowed to occur by incubation at 37°C for at least 60 min (up to 16 hours).

Restriction enzyme digestion and PCR products were fractionate by electrophoresis using 1% (w/v) agarose gel containing ethidium bromide (0.2 mg/ml) in 1 × TBE buffer (90 mM Tris; 90 mM boric acid; 2 mM EDTA). Electrophoresis was carried out at 70 volt for time period ranging from 30 min to 60 min.

A.6. DNA nucleotide sequencing

All DNA nucleotide sequencing reactions were conducted by an available DNA sequencing service (Dept. of Molecular and Cell Biology, University of Cape Town). For DNA fragments sub-cloned into the vector pBluescript, sequencing of the inserted fragments were performed using standard T7 and T3 sequencing primers. For the ACE2 mutants sub-cloned into the vector pCI-NEO, sequencing of the inserted ACE2

mutant fragments were performed using oligonucleotides listed in Table A.1, in addition to the standard T7 and T3 sequencing primers.

A.7. Sodium acetate DNA precipitation

Ice-cold (-20°C) absolute ethanol was added to sample at two times the sample volume together with a tenth final volume of 3M sodium acetate. The mixture was incubated at -70°C for at least 30min, followed by centrifugation at 4°C for 10 min in a bench-top microcentrifuge at full speed. The supernatant was decanted and the pellet was washed in 70% (v/v) ice-cold (-20°C) ethanol. The mixture was incubated at -20°C for 10 min and the DNA pellet collected by centrifugation at 4°C for 10 min in a bench-top microcentrifuge at full speed. The supernatant was decanted and the DNA pellet was dried using a Savant Speedi Vac System (GMI, Inc.).

A.8. SDS polyacrylamide gel electrophoresis

SDS-PAGE gel was casted using Mini-PROTEAN Tetra Electrophoresis System (Bio-Rad Laboratories). In the casting of a 10% (w/v) polyacrylamide running gel, 2.5 ml of 40% (w/v) acrylamide/bis-acrylamide (19:1) solution (Sigma-Aldrich, Inc.) was added to 3.3 ml of running gel buffer (1.1 M Tris, pH 8.8; 0.3% (w/v) SDS). The volume was made up to 9.9 ml with distilled water before the addition of 0.1 ml 10% (w/v) AMPS and 10 µl TEMED. The mixture was poured into casting plates and a volume of 100 µl of isopropanol was added to the top of gel. After polymerization of the running gel, the isopropanol was poured off and stacking gel solution was poured on top of the running gel. The 3% (w/v) stacking gel solution was prepared with 1.1 ml of 40% (w/v) acrylamide/bis-acrylamide (19:1) solution (Sigma-Aldrich, Inc.), 3.3 ml of stacking gel buffer (180 mM Tris, pH 6.8; 0.15% (w/v) SDS) and the volume made up to 9.7 ml with distilled water. After addition of 0.3 ml 10% (w/v) AMPS and 20 µl TEMED, the stacking gel solution was well mixed and immediately poured into gel casting plates with combs inserted on top. The Mini-PROTEAN Tetra Electrophoresis System (Bio-Rad Laboratories) was assembled with casted gel according to directions provided by the manufacturer.

Protein sample volumes exceeding 20 µl were adjusted to 20 µl by concentration using a Savant Speedi Vac System (GMI, Inc.) prior to the addition of 6 × sample loading buffer. The protein samples were reduced by boiling for 5 min in sample loading buffer (62.5 mM Tris, pH 6.8; 2% (w/v) SDS; 10% (v/v) glycerol; 5% (v/v) β-mercaptoethanol; 0.001% (w/v) bromophenol blue) and loaded directly onto SDS-PAGE gel. Electrophoresis was carried out in 1 × running buffer (25 mM Tris, pH 8.3; 190 mM glycine; 1% (w/v) SDS) at a constant current of 50 mA for 60 min.

After electrophoresis, the Mini-PROTEAN Tetra Electrophoresis System (Bio-Rad Laboratories) was disassembled and the running gel was subjected to staining in Coomassie staining solution (0.2% (w/v) Coomassie Blue; 7% (v/v) glacial acetic acid; 50% (v/v) ethanol) for at least 2 hours. De-staining of running gel was achieved by incubation of running gel in de-stain solution (7% (v/v) glacial acetic acid; 25% (v/v) ethanol) for 60 min with at least one change of de-stain solution.

Table A.1. Sequence of oligonucleotides used as primers in the DNA nucleotide sequencing of ACE2 constructs

Primer name	**Sequence**
sACE2 pCI NEO out	5′-CCCAAGCCCAGAGCCTC-3′
sACE2 mRNA200	5′-CCACGAAGCCGAAGACC-3′
sACE2 mRNA700	5′-GGGGATTATTGGAGAGG-3′
sACE2 mRNA1200	5′-GGACGACTTCCTGACAGC-3′
sACE2 mRNA1705	5′-CATGAAGGCCCTCTGCAC-3′
sACE2 mRNA2200	5′-GGCCATCAGGATGTCCC-3′

Table A.2. Sequences of oligonucleotides used in construction of ACE2 promoter mutants and chromatin immunoprecipitation experiments. Underlined sequences indicate the restriction sites introduced

Primer name	**Nucleotide sequence**
ACE2 C2 Fwd	5′-CTGAAATTTGACAAGATAAACCAC-3′
ACE2 C3 Fwd	5′-GAGCTTGCAGTGAGCCGAG-3′
ACE2 C4 Fwd	5′-TTTTCCGTGTATCTTTAACAGC-3′
ACE2 Con Rev	5′-GCCAAGATCACATCCACTGA-3′
ACE2 -1131 Fwd	5′-GATCCCATGGCTACAGAGG-3′
ACE2 -1077 Fwd	5′-CTTCCACGTGACCTTGAC-3′
ACE2 -1036 Fwd	5′-GGAGAAGGAGGCACTCAA-3′
ACE2 -963 Fwd	5′-GCAAAGGCAGATCAGGAG-3′
ACE2 -903 Fwd	5′-GGCCATGGAAATTAAAAC-3′
ACE2 -837 Fwd	5′-TTTGGGAGGCCGAGTT-3′
ACE2 -784 Fwd	5′-ACTAACACGGTGAAACCC-3′
ACE2 -742 Fwd	5′-TTAGCTGGGCGTGGT-3′
ACE2 -704 Fwd	5′-ACTCGGGAGGCTGAGG-3′
ACE2 Pro Con Rev	5′-GCCAAGATCACATCCACTGAAAGCTTGGG-3′
ACE2 RT Fwd	5′-CAAAGGCTGATAAGAGAGAAAAT-3′
ACE2 RT Rev	5′-CTGAGAGCACTGAAGACCCAT-3′
ΔRREB-1 Fwd	5′-AATACAAAAACTTAGCTGAATTCGGTTCTAGACACCTGTAGT CCCAGCTAC-3′
ΔRREB-1 Rev	5′-GTAGCTGGGACTACAGGTGTCTAGAACCGAATTCAGCTAAGT TTTTGTATT-3′
ΔZEB1 Fwd	5′- TTAGCTGGGCGTGGTGGTGGGGAATTCTAGTCCCAGCTACTC GGG- 3′
ΔZEB1 Rev	5′- CCCGAGTAGCTGGGACTAGAATTCCCCACCACCACGCCCAGC TAA- 3′
ACE2 Pro d_mut Fwd	5′- TTAGCTGAATTCGGTTCTAGAGAATTCTAGTCCCAGCTACTC GGG- 3′
ACE2 Pro d_mut Rev	5′- CCCGAGTAGCTGGGACTAGAATTCTCTAGAACCGAATTCAGC TAA- 3′

References

Acharya,K.R., Sturrock,E.D., Riordan,J.F., and Ehlers,M.R. (2003). Ace revisited: a new target for structure-based drug design. Nat. Rev. Drug Discov. *2*, 891-902.

Adachi,Y., Takeuchi,T., Nagayama,T., Ohtsuki,Y., and Furihata,M. (2009). Zeb1-mediated T-cadherin repression increases the invasive potential of gallbladder cancer. FEBS Lett. *583*, 430-436.

Aigner,K., Dampier,B., Descovich,L., Mikula,M., Sultan,A., Schreiber,M., Mikulits,W., Brabletz,T., Strand,D., Obrist,P., Sommergruber,W., Schweifer,N., Wernitznig,A., Beug,H., Foisner,R., and Eger,A. (2007). The transcription factor ZEB1 (deltaEF1) promotes tumour cell dedifferentiation by repressing master regulators of epithelial polarity. Oncogene *26*, 6979-6988.

Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A., and Struhl,K. (1992). Short Protocols in Molecular Biology. (New York: John Wiley and Sons).

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.

Brandt,R.B., Laux,J.E., and Yates,S.W. (1987). Calculation of inhibitor Ki and inhibitor type from the concentration of inhibitor for 50% inhibition for Michaelis-Menten enzymes. Biochem. Med. Metab Biol. *37*, 344-349.

Carey,R.M. and Padia,S.H. (2008). Angiotensin AT2 receptors: control of renal sodium excretion and blood pressure. Trends Endocrinol. Metab *19*, 84-87.

Cartharius,K., Frech,K., Grote,K., Klocke,B., Haltmeier,M., Klingenhoff,A., Frisch,M., Bayerlein,M., and Werner,T. (2005). MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics. *21*, 2933-2942.

Celletti,F.L., Waugh,J.M., Amabile,P.G., Brendolan,A., Hilfiker,P.R., and Dake,M.D. (2001). Vascular endothelial growth factor enhances atherosclerotic plaque progression. Nat. Med. *7*, 425-429.

Cheng,Y. and Prusoff,W.H. (1973). Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. Biochem. Pharmacol. *22*, 3099-3108.

Clarhaut,J., Gemmill,R.M., Potiron,V.A., Ait-Si-Ali,S., Imbert,J., Drabkin,H.A., and Roche,J. (2009). ZEB-1, a repressor of the semaphorin 3F tumor suppressor gene in lung cancer cells. Neoplasia. *11*, 157-166.

Cornish-Bowden,A. (1995). Fundamentals of enzyme kinetics. (London: Portland Press).

Cornish-Bowden,A. and Eisenthal,R. (1978). Estimation of Michaelis constant and maximum velocity from the direct linear plot. Biochim. Biophys. Acta *523*, 268-272.

Corradi,H.R., Chitapi,I., Sewell,B.T., Georgiadis,D., Dive,V., Sturrock,E.D., and Acharya,K.R. (2007). The structure of testis angiotensin-converting enzyme in complex with the C domain-specific inhibitor RXPA380. Biochemistry *46*, 5473-5478.

Crackower,M.A., Sarao,R., Oudit,G.Y., Yagil,C., Kozieradzki,I., Scanga,S.E., Oliveira-dos-Santos,A.J., da Costa,J., Zhang,L., Pei,Y., Scholey,J., Ferrario,C.M., Manoukian,A.S., Chappell,M.C., Backx,P.H., Yagil,Y., and Penninger,J.M. (2002). Angiotensin-converting enzyme 2 is an essential regulator of heart function. Nature *417*, 822-828.

Dales,N.A., Gould,A.E., Brown,J.A., Calderwood,E.F., Guan,B., Minor,C.A., Gavin,J.M., Hales,P., Kaushik,V.K., Stewart,M., Tummino,P.J., Vickers,C.S., Ocain,T.D., and Patane,M.A. (2002). Substrate-based design of the first class of angiotensin-converting enzyme-related carboxypeptidase (ACE2) inhibitors. J. Am. Chem. Soc. *124*, 11852-11853.

Danilczyk,U., Sarao,R., Remy,C., Benabbas,C., Stange,G., Richter,A., Arya,S., Pospisilik,J.A., Singer,D., Camargo,S.M., Makrides,V., Ramadan,T., Verrey,F., Wagner,C.A., and Penninger,J.M. (2006). Essential role for collectrin in renal amino acid transport. Nature *444*, 1088-1091.

Deaton,D.N., Gao,E.N., Graham,K.P., Gross,J.W., Miller,A.B., and Strelow,J.M. (2008a). Thiol-based angiotensin-converting enzyme 2 inhibitors: P1 modifications for the exploration of the S1 subsite. Bioorg. Med. Chem. Lett. *18*, 732-737.

Deaton,D.N., Graham,K.P., Gross,J.W., and Miller,A.B. (2008b). Thiol-based angiotensin-converting enzyme 2 inhibitors: P1' modifications for the exploration of the S1' subsite. Bioorg. Med. Chem. Lett. *18*, 1681-1687.

Dive,V., Cotton,J., Yiotakis,A., Michaud,A., Vassiliou,S., Jiracek,J., Vazeux,G., Chauvet,M.T., Cuniasse,P., and Corvol,P. (1999). RXP 407, a phosphinic peptide, is a potent inhibitor of angiotensin I converting enzyme able to differentiate between its two active sites. Proc. Natl. Acad. Sci. U. S. A *96*, 4330-4335.

Donoghue,M., Hsieh,F., Baronas,E., Godbout,K., Gosselin,M., Stagliano,N., Donovan,M., Woolf,B., Robison,K., Jeyaseelan,R., Breitbart,R.E., and Acton,S. (2000). A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. Circ. Res. *87*, E1-E9.

Drake,J.M., Strohbehn,G., Bair,T.B., Moreland,J.G., and Henry,M.D. (2009). ZEB1 enhances transendothelial migration and represses the epithelial phenotype of prostate cancer cells. Mol. Biol. Cell *20*, 2207-2217.

Eisenthal,R. and Cornish-Bowden,A. (1974). The direct linear plot. A new graphical procedure for estimating enzyme kinetic parameters. Biochem. J. *139*, 715-720.

Fyhrquist,F. and Saijonmaa,O. (2008). Renin-angiotensin system revisited. J. Intern. Med. *264*, 224-236.

Gallagher,P.E., Ferrario,C.M., and Tallant,E.A. (2008). MAP Kinase/Phosphatase Pathway Mediates the Regulation of ACE2 by Angiotensin Peptides. Am. J. Physiol Cell Physiol.

Georgiadis,D., Beau,F., Czarny,B., Cotton,J., Yiotakis,A., and Dive,V. (2003). Roles of the two active sites of somatic angiotensin-converting enzyme in the cleavage of angiotensin I and bradykinin: insights from selective inhibitors. Circ. Res. *93*, 148-154.

Girousse,A., Tavernier,G., Tiraby,C., Lichtenstein,L., Iacovoni,J.S., Mairal,A., Villarroya,F., and Langin,D. (2009). Transcription of the human uncoupling protein 3 gene is governed by a complex interplay between the promoter and intronic sequences. Diabetologia.

Gurley,S.B., Allred,A., Le,T.H., Griffiths,R., Mao,L., Philip,N., Haystead,T.A., Donoghue,M., Breitbart,R.E., Acton,S.L., Rockman,H.A., and Coffman,T.M. (2006). Altered blood pressure responses and normal cardiac phenotype in ACE2-null mice. J. Clin. Invest *116*, 2218-2225.

Guy,J.L., Jackson,R.M., Acharya,K.R., Sturrock,E.D., Hooper,N.M., and Turner,A.J. (2003). Angiotensin-converting enzyme-2 (ACE2): comparative modeling of the active site, specificity requirements, and chloride dependence. Biochemistry *42*, 13185-13192.

Guy,J.L., Jackson,R.M., Jensen,H.A., Hooper,N.M., and Turner,A.J. (2005). Identification of critical active-site residues in angiotensin-converting enzyme-2 (ACE2) by site-directed mutagenesis. FEBS J. *272*, 3512-3520.

Hansen,J.L., Aplin,M., Hansen,J.T., Christensen,G.L., Bonde,M.M., Schneider,M., Haunso,S., Schiffer,H.H., Burstein,E.S., Weiner,D.M., and Sheikh,S.P. (2008). The human angiotensin AT(1) receptor supports G protein-independent extracellular signal-regulated kinase 1/2 activation and cellular proliferation. Eur. J. Pharmacol. *590*, 255-263.

Hernandez Prada,J.A., Ferreira,A.J., Katovich,M.J., Shenoy,V., Qi,Y., Santos,R.A., Castellano,R.K., Lampkins,A.J., Gubala,V., Ostrov,D.A., and Raizada,M.K. (2008). Structure-based identification of small-molecule angiotensin-converting enzyme 2 activators as novel antihypertensive agents. Hypertension *51*, 1312-1317.

Horovitz,A. and Levitzki,A. (1987). An accurate method for determination of receptor-ligand and enzyme-inhibitor dissociation constants from displacement curves. Proc. Natl. Acad. Sci. U. S. A *84*, 6654-6658.

Hsieh,P.S., Tai,Y.H., Loh,C.H., Shih,K.C., Cheng,W.T., and Chu,C.H. (2005). Functional interaction of AT1 and AT2 receptors in fructose-induced insulin resistance and hypertension in rats. Metabolism *54*, 157-164.

Huang,L., Sexton,D.J., Skogerson,K., Devlin,M., Smith,R., Sanyal,I., Parry,T., Kent,R., Enright,J., Wu,Q.L., Conley,G., DeOliveira,D., Morganelli,L., Ducar,M., Wescott,C.R., and Ladner,R.C. (2003). Novel peptide inhibitors of angiotensin-converting enzyme 2. J. Biol. Chem. *278*, 15532-15540.

Huentelman,M.J., Zubcevic,J., Hernandez Prada,J.A., Xiao,X., Dimitrov,D.S., Raizada,M.K., and Ostrov,D.A. (2004). Structure-based discovery of a novel angiotensin-converting enzyme 2 inhibitor. Hypertension *44*, 903-906.

Igarashi,M., Hirata,A., Nozaki,H., Kadomoto-Antsuki,Y., and Tominaga,M. (2007). Role of angiotensin II type-1 and type-2 receptors on vascular smooth muscle cell growth and glucose metabolism in diabetic rats. Diabetes Res. Clin. Pract. *75*, 267-277.

Igase,M., Strawn,W.B., Gallagher,P.E., Geary,R.L., and Ferrario,C.M. (2005). Angiotensin II AT1 receptors regulate ACE2 and angiotensin-(1-7) expression in the aorta of spontaneously hypertensive rats. Am. J. Physiol Heart Circ. Physiol *289*, H1013-H1019.

Imai,Y., Kuba,K., Rao,S., Huan,Y., Guo,F., Guan,B., Yang,P., Sarao,R., Wada,T., Leong-Poi,H., Crackower,M.A., Fukamizu,A., Hui,C.C., Hein,L., Uhlig,S., Slutsky,A.S., Jiang,C., and Penninger,J.M. (2005). Angiotensin-converting enzyme 2 protects from severe acute lung failure. Nature *436*, 112-116.

Itoyama,S., Keicho,N., Hijikata,M., Quy,T., Phi,N.C., Long,H.T., Ha,l.D., Ban,V.V., Matsushita,I., Yanai,H., Kirikae,F., Kirikae,T., Kuratsuji,T., and Sasazuki,T. (2005). Identification of an alternative 5'-untranslated exon and new polymorphisms of angiotensin-converting enzyme 2 gene: lack of association with SARS in the Vietnamese population. Am. J. Med. Genet. A *136*, 52-57.

Iusuf,D., Henning,R.H., van Gilst,W.H., and Roks,A.J. (2008). Angiotensin-(1-7): pharmacological properties and pharmacotherapeutic perspectives. Eur. J. Pharmacol. *585*, 303-312.

Jones,E.A., Jang,S.W., Mager,G.M., Chang,L.W., Srinivasan,R., Gokey,N.G., Ward,R.M., Nagarajan,R., and Svaren,J. (2007). Interactions of Sox10 and Egr2 in myelin gene regulation. Neuron Glia Biol. *3*, 377-387.

Kuba,K., Imai,Y., Rao,S., Gao,H., Guo,F., Guan,B., Huan,Y., Yang,P., Zhang,Y., Deng,W., Bao,L., Zhang,B., Liu,G., Wang,Z., Chappell,M., Liu,Y., Zheng,D., Leibbrandt,A., Wada,T., Slutsky,A.S., Liu,D., Qin,C., Jiang,C., and Penninger,J.M. (2005). A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. Nat. Med. *11*, 875-879.

Laird,P.W., Zijderveld,A., Linders,K., Rudnicki,M.A., Jaenisch,R., and Berns,A. (1991). Simplified mammalian DNA isolation procedure. Nucleic Acids Res. *19*, 4293.

Lambert,D.W., Clarke,N.E., Hooper,N.M., and Turner,A.J. (2008). Calmodulin interacts with angiotensin-converting enzyme-2 (ACE2) and inhibits shedding of its ectodomain. FEBS Lett. *582*, 385-390.

Lambert,D.W., Yarski,M., Warner,F.J., Thornhill,P., Parkin,E.T., Smith,A.I., Hooper,N.M., and Turner,A.J. (2005). Tumor necrosis factor-alpha convertase (ADAM17) mediates regulated ectodomain shedding of the severe-acute respiratory

syndrome-coronavirus (SARS-CoV) receptor, angiotensin-converting enzyme-2 (ACE2). J. Biol. Chem. *280*, 30113-30119.

Laskowski,R.A., MacArthur,M.W., Moss,D.S., and Thornton,J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. 283-291.

LeBlanc,S.E., Jang,S.W., Ward,R.M., Wrabetz,L., and Svaren,J. (2006). Direct regulation of myelin protein zero expression by the Egr2 transactivator. J. Biol. Chem. *281*, 5453-5460.

Li,G.M., Li,Y.G., Yamate,M., Li,S.M., and Ikuta,K. (2007). Lipid rafts play an important role in the early stage of severe acute respiratory syndrome-coronavirus life cycle. Microbes. Infect. *9*, 96-102.

Li,W., Moore,M.J., Vasilieva,N., Sui,J., Wong,S.K., Berne,M.A., Somasundaran,M., Sullivan,J.L., Luzuriaga,K., Greenough,T.C., Choe,H., and Farzan,M. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature *426*, 450-454.

Lovren,F., Pan,Y., Quan,A., Teoh,H., Wang,G., Shukla,P.C., Levitt,K.S., Oudit,G.Y., Al Omran,M., Stewart,D.J., Slutsky,A.S., Peterson,M.D., Backx,P.H., Penninger,J.M., and Verma,S. (2008). Angiotensin converting enzyme-2 confers endothelial protection and attenuates atherosclerosis. Am. J. Physiol Heart Circ. Physiol *295*, H1377-H1384.

Marinescu,V.D., Kohane,I.S., and Riva,A. (2005a). MAPPER: a search engine for the computational identification of putative transcription factor binding sites in multiple genomes. BMC. Bioinformatics. *6*, 79.

Marinescu,V.D., Kohane,I.S., and Riva,A. (2005b). The MAPPER database: a multi-genome catalog of putative transcription factor binding sites. Nucleic Acids Res. *33*, D91-D97.

Mizuiri,S., Hemmi,H., Arita,M., Ohashi,Y., Tanaka,Y., Miyagi,M., Sakai,K., Ishikawa,Y., Shibuya,K., Hase,H., and Aikawa,A. (2008). Expression of ACE and ACE2 in individuals with diabetic kidney disease and healthy controls. Am. J. Kidney Dis. *51*, 613-623.

Monaco,C., Helmer,C.M., Caprini,E., Vorechovsky,I., Russo,G., Croce,C.M., Barbanti-Brodano,G., and Negrini,M. (1998). Molecular cloning and characterization of ZNF202: a new gene at 11q23.3 encoding testis-specific zinc finger proteins. Genomics *52*, 358-362.

Mores,A., Matziari,M., Beau,F., Cuniasse,P., Yiotakis,A., and Dive,V. (2008). Development of potent and selective phosphinic peptide inhibitors of angiotensin-converting enzyme 2. J. Med. Chem. *51*, 2216-2226.

Ni,A., Yin,H., Agata,J., Yang,Z., Chao,L., and Chao,J. (2003). Overexpression of kinin B1 receptors induces hypertensive response to des-Arg9-bradykinin and susceptibility to inflammation. J. Biol. Chem. *278*, 219-225.

Spaderna,S., Schmalhofer,O., Wahlbuhl,M., Dimmler,A., Bauer,K., Sultan,A., Hlubek,F., Jung,A., Strand,D., Eger,A., Kirchner,T., Behrens,J., and Brabletz,T. (2008). The transcriptional repressor ZEB1 promotes metastasis and loss of cell polarity in cancer. Cancer Res. *68*, 537-544.

Stene,M.C., Frikke-Schmidt,R., Nordestgaard,B.G., Grande,P., Schnohr,P., and Tybjaerg-Hansen,A. (2008). Functional promoter variant in zinc finger protein 202 predicts severe atherosclerosis and ischemic heart disease. J. Am. Coll. Cardiol. *52*, 369-377.

Takeda,Y., Zhu,A., Yoneda,T., Usukura,M., Takata,H., and Yamagishi,M. (2007). Effects of aldosterone and angiotensin II receptor blockade on cardiac angiotensinogen and angiotensin-converting enzyme 2 expression in Dahl salt-sensitive hypertensive rats. Am. J. Hypertens. *20*, 1119-1124.

Tikellis,C., Cooper,M.E., Bialkowski,K., Johnston,C.I., Burns,W.C., Lew,R.A., Smith,A.I., and Thomas,M.C. (2006). Developmental expression of ACE2 in the SHR kidney: a role in hypertension? Kidney Int. *70*, 34-41.

Tikellis,C., Johnston,C.I., Forbes,J.M., Burns,W.C., Burrell,L.M., Risvanis,J., and Cooper,M.E. (2003). Characterization of renal angiotensin-converting enzyme 2 in diabetic nephropathy. Hypertension *41*, 392-397.

Tipnis,S.R., Hooper,N.M., Hyde,R., Karran,E., Christie,G., and Turner,A.J. (2000). A human homolog of angiotensin-converting enzyme. Cloning and functional expression as a captopril-insensitive carboxypeptidase. J. Biol. Chem. *275*, 33238-33243.

Towler,P., Staker,B., Prasad,S.G., Menon,S., Tang,J., Parsons,T., Ryan,D., Fisher,M., Williams,D., Dales,N.A., Patane,M.A., and Pantoliano,M.W. (2004). ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. J. Biol. Chem. *279*, 17996-18007.

Verdonk,M.L., Cole,J.C., Hartshorn,M.J., Murray,C.W., and Taylor,R.D. (2003). Improved protein-ligand docking using GOLD. Proteins *52*, 609-623.

Vickers,C., Hales,P., Kaushik,V., Dick,L., Gavin,J., Tang,J., Godbout,K., Parsons,T., Baronas,E., Hsieh,F., Acton,S., Patane,M., Nichols,A., and Tummino,P. (2002). Hydrolysis of biological peptides by human angiotensin-converting enzyme-related carboxypeptidase. J. Biol. Chem. *277*, 14838-14843.

Wagner,S., Hess,M.A., Ormonde-Hanson,P., Malandro,J., Hu,H., Chen,M., Kehrer,R., Frodsham,M., Schumacher,C., Beluch,M., Honer,C., Skolnick,M., Ballinger,D., and Bowen,B.R. (2000). A broad role for the zinc finger protein ZNF202 in human lipid metabolism. J. Biol. Chem. *275*, 15685-15690.

Warner,F.J., Lew,R.A., Smith,A.I., Lambert,D.W., Hooper,N.M., and Turner,A.J. (2005). Angiotensin-converting enzyme 2 (ACE2), but not ACE, is preferentially localized to the apical surface of polarized kidney cells. J. Biol. Chem. *280*, 39353-39362.

Wiener,R.S., Cao,Y.X., Hinds,A., Ramirez,M.I., and Williams,M.C. (2007). Angiotensin converting enzyme 2 is primarily epithelial and is developmentally regulated in the mouse lung. J. Cell Biochem. *101*, 1278-1291.

Wong,D.W., Oudit,G.Y., Reich,H., Kassiri,Z., Zhou,J., Liu,Q.C., Backx,P.H., Penninger,J.M., Herzenberg,A.M., and Scholey,J.W. (2007). Loss of angiotensin-converting enzyme-2 (Ace2) accelerates diabetic kidney injury. Am. J. Pathol. *171*, 438-451.

Yamamoto,K., Ohishi,M., Katsuya,T., Ito,N., Ikushima,M., Kaibe,M., Tatara,Y., Shiota,A., Sugano,S., Takeda,S., Rakugi,H., and Ogihara,T. (2006). Deletion of angiotensin-converting enzyme 2 accelerates pressure overload-induced cardiac dysfunction by increasing local angiotensin II. Hypertension *47*, 718-726.

Yamamuro,M., Yoshimura,M., Nakayama,M., Abe,K., Sumida,H., Sugiyama,S., Saito,Y., Nakao,K., Yasue,H., and Ogawa,H. (2008). Aldosterone, but not angiotensin II, reduces angiotensin converting enzyme 2 gene expression levels in cultured neonatal rat cardiomyocytes. Circ. J. *72*, 1346-1350.

Zhao,J., Wang,W., Yuan,Z., Jia,R., Zhao,Z., Xu,X., Lv,P., Zhang,Y., Jiang,C., and Gao,X.M. (2007). A study on antigenicity and receptor-binding ability of fragment 450-650 of the spike protein of SARS coronavirus. Virology *359*, 362-370.