

ASSESSING THE MEMBRANE TOPOLOGY OF THE YEAST RAS CONVERTING ENZYME (RCE1P)

by

RANJITH KUMAR KRISHNANKUTTY

(Under the Direction of Walter K. Schmidt)

ABSTRACT

The Ras converting enzyme (Rce1p) is an isoprenylation specific endoprotease involved in the post-translational processing of isoprenylated proteins, such as Ras and yeast mating pheromone **a**-factor. Inhibition of the enzyme leads to attenuation of transforming effects of mutant activated Ras and partial mislocalization of the protein, without affecting cell viability. Rce1p is an integral membrane protein localized on the endoplasmic reticulum membrane. There has not yet been any successful derivation of its protein structure using conventional techniques. This study has used the substituted cysteine accessibility method for determining the membrane topology of Rce1p. Two putative topology models have thus been constructed for Rce1p, bearing either five or seven transmembrane spans. Topology determination for Rce1p is a step forward in the determination of its structure and catalytic mechanism, which in turn has implications in the development of Rce1p inhibitors as novel anti-cancer drugs targeting the maturation of Ras protein.

INDEX WORDS: Cysteine scanning, Rce1p, PEG-Maleimide, **a**-factor, Membrane topology.

ASSESSING THE MEMBRANE TOPOLOGY OF THE YEAST RAS CONVERTING
ENZYME (RCE1P)

by

RANJITH KUMAR KRISHNANKUTTY

M.B.B.S, Medical College Thiruvananthapuram, India, 2006

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment
of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2009

© 2009

RANJITH KUMAR KRISHNANKUTTY

All Rights Reserved

ASSESSING THE MEMBRANE TOPOLOGY OF THE YEAST RAS CONVERTING
ENZYME (RCE1P)

by

RANJITH KUMAR KRISHNANKUTTY

Major Professor: Walter K. Schmidt

Committee: J. David Puett
Lance Wells

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2009

ACKNOWLEDGEMENTS

I am sincerely grateful to all those who have assisted me in this endeavor. I would especially like to thank my major advisor, Dr. Walter K. Schmidt, for his unending support and guidance and also for teaching me everything that I learnt during my time in the Schmidt lab. His optimism, patience in teaching and skill in molecular biology techniques will always be an inspiration for me.

I am also deeply indebted to Drs. Emily Hildebrandt and John Deaton for constructing many of the plasmids and strains used in this study, and also for giving me technical guidance to create the rest. I would also like to thank the members of my thesis committee, Drs. J. David Puett and Lance Wells for their valuable ideas and supervision over this project.

My experience in the Schmidt Lab has always been enjoyable, and for that I am grateful to all its members, not only for being there at all times to help with suggestions and ideas but also for creating an atmosphere filled with respect and amiability.

Last, but definitely not the least, I would also like to thank my friends here at the University of Georgia for their unconditional support, and for making the past three years an unforgettable part of my life.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
CHAPTER	
1 INTRODUCTION AND LITERATURE REVIEW	1
1.1 Ras Protein	1
1.2 Ras Signaling Pathway	3
1.3 Ras Protein Maturation	5
1.4 Ras Converting Enzyme (Rce1p)	7
1.5 Catalytic Activity of Rce1p	10
1.6 Techniques to Determine Rce1p Membrane Topology	12
1.7 Summary and Hypothesis	17
1.8 References	18
2 TOPOLOGY OF RAS CONVERTING ENZYME (RCE1P) AS DETERMINED BY SUBSTITUTED CYSTEINE ACCESSIBILITY METHOD	30
3 CONCLUSION AND DISCUSSION	70
APPENDIX	75
A PROTEOLYTIC PROCESSING OF CERTAIN CAAX MOTIFS CAN OCCUR IN THE ABSENCE OF THE RCE1P AND STE24P CAAX PROTEASES	76

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

The Ras converting enzyme is a protease involved in the maturation of a wide variety of isoprenylated proteins. The most notable biomedically relevant targets are the Ras GTPases. Since the Ras protein is significantly involved in tumorigenesis and cancer, the study of the Ras converting enzyme has useful applications in cancer pharmacotherapy.

1.1. Ras Protein

The Ras proteins are approximately 21,000 Da proteins with GTPase activity that are translated from various *RAS* genes. The oncogenic versions of the *RAS* genes were discovered first, when certain leukemia-causing retroviruses, on being transfected into newborn mice, were observed to induce sarcomas in their murine hosts (1). These oncogenic viral genes, named as *v-h-ras* and *v-k-ras* ('ras' from rat sarcoma), identified in the Harvey (H-MSV) and Kirsten (K-MSV) mouse sarcoma viral strains respectively, were similar to certain gene sequences found in the rat genome, and subsequently found to hybridize with DNA isolated from certain human bladder and lung cancer cell lines. The mammalian versions of the *v-h-ras* and *v-k-ras* genes were named as *HRAS* and *KRAS* (2). Subsequently, another member of the *RAS* oncogene family, *NRAS* was identified from human neuroblastoma cells (3). Later, as other proteins were identified which were biochemically similar to the protein products of the *RAS* genes, they were

classified into subfamilies within a large Ras superfamily of proteins. The Ras superfamily of proteins consist of over 150 small monomeric proteins with GTPase activity and are divided into the Ras, Rho, Rab, Ran, Arf and Kir subfamilies. Through their binding of guanine nucleotides and their GTPase activity, they serve as molecular switches in various intracellular signaling pathways (4). The Ras subfamily of proteins include more closely related proteins with the most well known being H-Ras, N-Ras, and the two K-Ras proteins formed from alternative slicing of the *KRAS* gene, K-Ras 4A and K-Ras 4B. All four of these proteins have a molecular size of approximately 21000 Da (therefore they are also known as p21 Ras), and have approximately 85% sequence similarity. The other members of the Ras subfamily include the M-Ras, R-Ras, Ral and Rap proteins, which share approximately 40-50% sequence similarity with the p21 Ras proteins (4).

Although the Ras protein has now been identified to be integral to a wide variety of cellular functions, it was its role in cell proliferation and cell survival that first brought attention to the involvement of Ras in cancer. After being discovered as oncogenes, the normal cellular versions of the *RAS* gene were soon cloned, and this led to the discovery that the oncogenic versions were mutated versions of the normal cellular *RAS* genes (5). Normal cellular *RAS* genes did not transform immortal cell lines like murine NIH-3T3 on transfection, but did so if viral oncogenic enhancers were present (6). Different chemical induced tumors in cell lines were also found to be caused by chemical induced mutations in the *RAS* genes (7). Further studies of the role of *RAS* genes in tumorigenesis were done by using inducible *RAS* transgenes in mice, which could be 'switched on' in specific tissues at specific times (8). This enabled the study of roles played by specific types of Ras proteins in cancers arising from different tissues. Such studies have shown that *RAS* gene mutations are associated with approximately 30% of all cancers, and

specifically with 90% of pancreatic adenocarcinomas, approximately 50% of colon cancers, 50% of thyroid carcinomas, 33% of lung adenocarcinomas and 35% of acute myeloid leukemias to name a few (9). K-Ras mutations are found to be mainly associated with pancreatic and colon cancers, while N-Ras mutations are seen more in hematological malignancies(9).

1.2. Ras Signaling Pathway

The Ras proteins play important roles in numerous complicated intracellular signaling pathways that modulate cellular growth, proliferation, differentiation and apoptosis (10). These intricate pathways are extremely interconnected, involve other members of the Ras superfamily, and are yet to be completely elucidated.

The Ras proteins are monomeric GTPase proteins that alternate between an activated guanosine triphosphate (GTP) bound form and a neutral or inactive guanosine diphosphate (GDP) bound form. They lie attached to the inner face of the plasma membrane in part through a lipid modification, which is described in detail in the next section (11). The Ras protein acts as a transducer protein for signals received by receptor tyrosine kinase receptors. They include the epidermal growth factor receptor, T cell receptors, B cell receptors, and IgE receptors (12-14). On binding to their specific ligands, these receptors dimerize, and their intracellular domains, which have tyrosine kinase activity, autophosphorylate themselves on Tyr residues. The phosphorylated Tyr residues are recognized by certain adaptor proteins with src homology 2 (SH2) domains, such as the growth factor receptor bound protein (Grb2) (15). These proteins also have an N-terminal SH3 domain through which they bind guanine nucleotide exchange factor (GEF) such as the Sos1/2, RasGRP or RasGRF1/2 proteins (16). The GEF proteins promote the loss of GDP bound to the inactive Ras-GDP protein. Once lost, the GDP is quickly

replaced by GTP, by virtue of it being present in abundant amounts in the cytosol. This replacement of GDP with GTP induces an allosteric change in the Switch 1 region of Ras, where the Ras effector proteins bind, and in Switch 2, and thus activates the protein by enabling its binding with its downstream effectors (17).

The effector proteins that Ras interacts with downstream depend on which cellular signaling pathway the signal is intended for. The major signaling pathways in which Ras is involved include the MAPK-ERK pathway that regulates gene expression and cell proliferation (18, 19), the Phosphoinositide 3'-OH kinase (PI3K)-Akt pathway that modulates cell survival and apoptosis (20), the PI3K-RacGEF/Rho pathway that regulates the intracellular actin cytoskeletal system (21), and the Tiam1-Rac (22) and RalGEF-Ral pathways (23), both of which modulate anchor-independent growth and metastasis of cancerous cells. Apart from these effectors, there are proteins which serve as negative regulators of Ras activity, known as GTPase activating proteins (GAPs). These proteins, such as p120RasGAP, NF1, GAP1m, and GAP1H accelerate the intrinsic GTPase activity of Ras by inducing an allosteric modification of the Ras active site thus leading to the hydrolysis of the GTP to GDP (24-27). This 'switching off' of Ras activity helps keep the signaling pathway in check, such that signals are transduced only when the upstream receptor is activated. In those Ras mutations that lead to cancer, such as those that occur at positions 12, 13 or 61, the Ras protein becomes insensitive to RasGAP activity, and so remains in the 'on' position even when there is no upstream signal. This then leads on to indiscriminate cell proliferation, inhibition of apoptosis and other changes that cause tumorigenesis (9).

1.3. Ras Protein Maturation

All proteins in the Ras subfamily are known to have a common motif of four amino acids in the C-terminus, known as the Ca_1a_2X motif. C indicates a cysteine residue present four amino acid residues away from the C-terminus; a_1 and a_2 indicate any aliphatic amino acid; and X stands for a number of different amino acids (28). This Ca_1a_2X motif directs post-translational modifications of the protein, without which the protein is not optimally functional. It is in this process that the Ras converting enzyme plays an important role.

The Ras proteins are translated as pro-p21 proteins in the cytoplasm, by free ribosomes (29). These become a substrate for the farnesyl transferase (FTase) enzyme, which attaches an isoprenoid C_{15} farnesyl group to the Cys present in the Ca_1a_2X motif. In yeast this FTase has been observed to be a heterodimer of Ram1p and Ram2p proteins. The presence of a Leu or Phe at the X position however, can make the protein a substrate of the geranylgeranyl transferase (GGTase) enzyme. This attaches a C_{20} geranylgeranyl group to the cysteine instead of a farnesyl group (30). This isoprenylation step targets the protein to the endoplasmic reticulum membrane, on the cytosolic face of which the subsequent steps take place. Following isoprenylation, a proteolytic cleavage takes place, which removes the last three a_1a_2X amino acids from the C-terminus. In mammalian and yeast cells, this is performed by the Ras converting enzyme (Rce1p). This is then followed by carboxymethylation of the now C-terminus cysteine, thus making it more hydrophobic (31).

In addition to these post-translational modifications, certain Ras proteins also undergo palmitoylation. This is the reversible attachment of a palmitoyl group by a protein acyl transferase enzyme to a cysteine residue that is located just proximal to the isoprenylated cysteine on the protein. H-Ras, N-Ras, K-Ras 4A and yeast Ras1p and Ras2p undergo

palmitoylation, and this modification, along with isoprenylation, help link these proteins to the inner face of the plasma membrane. K-Ras 4B on the other hand, does not have the palmitoylated cysteine, and instead has a polybasic chain of amino acids at the C-terminus. These help to link the C-terminus end to the negatively charged phospholipids in the plasma membrane (32). The nature of these additional post-translational modifications also decides the localization of the Ras proteins on the plasma membrane. Palmitoylated proteins such as H-Ras and N-Ras are seen to associate in lipid rafts, along with their upstream receptors, such as TCR and BCR, while K-Ras 4B is seen to associate in disordered plasma membrane regions (12, 13, 33).

In addition to Ras, there are numerous other essential proteins that carry the Ca_1a_2X motif and undergo the same post-translational processing described for Ras proteins. Some of them include Rac and Cdc42, proteins belonging to the Rho family of GTPases, that are important for intracellular signaling and cancer formation; the G_γ subunit of heterotrimeric G proteins, which is also involved in cell signaling; the nuclear lamins A and B, which are essential for formation of the nuclear envelope; centromeric proteins such as CENPE and CENPF, which are needed for mitotic control and transcriptional regulation; and the yeast mating pheromone **a**-factor (34).

The yeast **a**-factor warrants a closer look as one of the substrates of the Rce1p enzyme not only because of its use in this study, but also as a model to study post-translational events occurring in proteins with Ca_1a_2X motifs. It is a mating pheromone secreted by the haploid *MATa* strains of *Saccharomyces cerevisiae*, and its complete maturation and secretion are essential for it to mate with the haploid *MATa* cells. On secretion from the *MATa* cells, it is recognized by the Ste3 receptor on the surface of *MATa* cells, and this recruits the *MATa* cells to grow cellular projections towards the source of the pheromone, ultimately leading to mating. The

steps in biogenesis of **a**-factor are summarized in Figure 1.2. Its C-terminal Ca₁a₂X processing is identical to that of the Ras protein (35).

Isoprenylation as a post-translational modification has recently been discovered to be essential not just for linking the protein to hydrophobic membranes, but also for protein-protein interaction. In yeast, farnesylated Ras2p has greatly increased affinity to its effector protein adenylyl cyclase as compared to unfarnesylated Ras2p. Similarly, Sos has been shown to complex with farnesylated K-Ras but not with unprocessed K-Ras (36). In the case of Ras, the absence of isoprenylation or Ca₁a₂X proteolysis has been observed to reduce function of mutated, activated Ras protein (37). Thus, this opens a window of opportunity for an effective treatment of those cancers which involve hyperactive Ras protein activity, by the inhibition of these post-translational processing enzymes, especially the Ras converting enzyme.

1.4. Ras Converting Enzyme (Rce1p)

Pancreatic membrane extracts were found to have endoproteolytic activity that cleaved the last three amino acids from the C-terminus of prenylated K-Ras protein (38). Similar studies when conducted with yeast and rat liver membrane preparations showed activity that performed the same endoproteolytic cleavage on farnesylated **a**-factor *in vitro* (39). Biochemical efforts did not lead to identification of the gene product responsible for this activity. However, the gene for the first Ca₁a₂X protease, *STE24* (also known as *AFC1*) was identified via genetic approaches using *S. cerevisiae* as the model system. This coded for an integral membrane protein, Ste24p, with Zn-dependent metalloprotease activity. The presence of another Ca₁a₂X protease was suspected by the fact that membrane preparations from *STE24* mutants showed only a slight decrease in endoproteolytic activity, and this was of the same magnitude as that seen when wild-

type membrane extracts were treated with Zn chelators such as 1,10-phenanthroline. Screening for this gene by looking for mutations that caused complete loss of **a**-factor production among *STE24* mutants led to the discovery of the *RCE1* gene (37).

So far, orthologs of Rce1p have been discovered in all eukaryotes examined so far, including humans, plants (*Arabidopsis thaliana*), worm (*Caenorhabditis elegans*), fly (*Drosophila melanogaster*), protozoa (*Trypanosoma brucei*) and other yeast (*Schizosaccharomyces pombe*). While their sequence similarity and size vary largely, they are all functionally identical to the yeast enzyme with similar substrate specificity, and have been observed to be able to substitute for yeast Rce1p in yeast where this has been examined (40-44) (Mokry, D.Z. et al., unpublished observation).

The Rce1p enzyme is a polytopic integral membrane protein, 315 amino acids long, with a predicted size of 35 kDa in yeast. Its function in yeast was studied using *RCE1* mutants and double delete mutants where both the *RCE1* and *STE24* genes were disrupted. It was observed that the Rce1p enzyme is not essential for viability of yeast cells, as was with the double deletion of both Ca₁a₂X proteases. In terms of **a**-factor production, there was only a slight decrease in *RCE1* mutants, while the double mutants showed complete lack of **a**-factor production as was evident from their inability to mate with *MAT α* cells. There was also no effect seen on the mating ability of *MAT α* cells, implying the Rce1p enzyme has no role in the biogenesis of the other yeast mating pheromone, α -factor (37). The effect of Rce1p disruption was studied in the maturation of the yeast Ras proteins, Yeast Ras mutants show sensitivity to heat shock and starvation in yeast cells. It was observed that the increased heat shock sensitivity seen in yeast cells with mutant activated Ras proteins was attenuated when *RCE1* was disrupted. No effect was seen in *STE24* mutants, implying that Rce1p was specific for Ras maturation while **a**-factor

maturation had dual specificity for either enzyme (37). Studies using Ras-GFP proteins also showed that when *RCE1* was disrupted, the Ras proteins were partially mislocalized, and were found to lie on intracellular membranes as opposed to normally observed sites on the plasma membrane (37). Also, while *RAS* gene deletions were lethal in yeast cells, the cells grew normally when *RCE1* was disrupted. In mammalian cells however, the growth of immortalized cell lines were decreased by disruption of *RCE1*. The transforming ability of cells with mutant *RAS* genes were also decreased in *RCE1* disrupted cells (45).

Since the post-translational processing of the Ras protein is essential for its function, and the role of mutated Ras in cancer initiation and metastasis was well known, disrupting Ras protein maturation has always been an attractive means to treat cancer. One method involves the use of farnesyltransferase inhibitors which were found to be very effective in retarding cancer cell growth and caused no systemic toxicity in mice (46). However, their use was complicated by the fact that most solid tumors in humans are associated with mutations of the K-Ras and N-Ras proteins, which can alternatively be geranylgeranylated by the GGTase-1 enzyme in the presence of FTase inhibition (47). Use of GGTase inhibitors in combination with FTase inhibitors or alone, have also been discouraging due to the high levels of toxicity observed (48). In view of the problems with isoprenoid transferase inhibition, the inhibition of the $\text{Ca}_1\text{a}_2\text{X}$ proteolysis step, which follows both farnesylation and geranylgeranylation alike, is being investigated as a favorable option. Studies have shown that Rce1p inhibitors inhibit growth of K-Ras transformed mammalian cell lines without affecting viability of non-transformed cells, and also limit the anchorage independent growth of transformed cells in mice (49). These results encourage the need to study methods to inhibit the Rce1p enzyme, which by implication shows the need to study its mechanism of catalytic activity.

1.5. Catalytic Activity of Rce1p

The yeast Rce1p enzyme is ER membrane localized, and is predicted by hydropathy analysis to be a polytopic integral membrane protein with multiple membrane spans (50). As an integral membrane protein with proteolytic activity, it is unusual, and joins a select group of other proteases with polytopic membrane spans, such as the Ste24p enzyme; presenilin, which cleaves A β peptide; SPP which cleaves the signal peptide of translated proteins; Site 2 Protease, which participates in the processing of sterol response element binding protein; and Rhomboid, which is necessary for production of epidermal growth factor in *Drosophila* (51, 52). Numerous inhibitor profiles of Rce1p have been attempted in order to classify its proteolytic activity within the 5 known types of protease mechanisms, namely Cysteine proteases, Serine/Threonine proteases, Aspartate proteases, Metalloproteases, and Glutamic proteases, but have been unsuccessful. Rce1p also lacks any of the specific identifiable motifs present in each class of peptidases.

Rce1p has been found to be insensitive to known Serine/Threonine protease inhibitors such as antipain, leupeptin, chymostatin, PMSF; and to known Aspartate protease inhibitors such as pepstatin (53). It was once considered to be a Cysteine protease on the basis of inhibition of the enzyme by diazo-ketone Ac-(far)CDK, a farnesylated compound that mimics the transition state of cysteine proteases, and also by TPCK, a cysteine protease inhibitor (53, 54). The presence of a conserved cysteine at position 251 also seemed to suggest the sulfhydryl moiety of that residue to be the active site nucleophile. However, mutational analyses of the enzyme showed the non-essential nature of the Cys²⁵¹ residue, which was the only conserved Cys among the various orthologs studied. Instead, three other conserved amino acid residues, Glu¹⁵⁶, His¹⁹⁴ and His²⁴⁸ were found to be essential for the activity of the enzyme (51). The need for two His

and one Glu, which is characteristic of metalloproteases, now suggests that to be the catalytic class of Rce1p enzyme. Consistent with this classification, Rce1p is sensitive to the metalloprotease inhibitor, 1,10-phenanthroline and excess Zn. The previous TPCK inhibition of Rce1p could also be explained on the basis of the hydrophobic interactions that could occur between the enzyme, which is an integral membrane protein, and TPCK (51). Its lesser sensitivity to TLCK, a hydrophilic relative of TPCK, also favors this conclusion. However, the insensitivity of the enzyme to excess EDTA treatment, and the inability to restore catalytic activity with excess metal ions after treatment with 1,10-phenanthroline, which would be expected to occur if it was a metalloprotease, throws an element of doubt on this hypothesis. The sensitivity to 1,10-phenanthroline could readily be explained on the basis of hydrophobic interactions that could possibly occur between the hydrophobic compound and the enzyme.

The most conclusive evidence for classifying the enzyme as a metalloprotease would be to determine the crystal structure of the enzyme in complex with a metal ion. However, being the integral membrane protein that Rce1p is, there have been inherent difficulties in just purifying the enzyme. One step towards identifying the catalytic mechanism of the enzyme would be to determine the membrane topology of the enzyme. It is predicted that the catalytic activity of the enzyme must face the cytosolic side of the ER membrane (50), so by determining the membrane topology, the orientation of the three essential amino acid residues mentioned above, Glu¹⁵⁶, His¹⁹⁴ and His²⁴⁸, could be determined, and thus their involvement in the catalytic activity of the enzyme inferred.

1.6. Techniques to Determine Rce1p Membrane Topology

The most useful step towards identifying an enzyme's catalytic mechanism would be to determine its three dimensional structure. However, in the case of Rce1p, since it is an integral membrane protein, the conventional techniques that are available to solve three dimensional structures of soluble proteins are not easily applicable, mostly due to the inability to purify the enzyme till date. Integral membrane proteins rely extensively on their surrounding environment for both folding into their final structure and also for maintaining that structure. Any change in their natural hydrophobic environment, including the hydrophobic core of the membrane and the membrane-water interface region, would cause the protein to alter its structure, thereby affecting its functionality.

The conventional techniques used for determining three dimensional structures of soluble proteins include X-Ray Crystallography and Nuclear Magnetic Resonance (NMR) spectroscopy. In order to perform X-Ray Crystallography, three dimensional crystals of the protein need to be obtained first. This in turn requires purification of the protein. Integral membrane proteins do not readily form three dimensional crystals as indicated by the relatively few membrane protein structures resolved till date. The overexpression of membrane proteins is likely to cause misfolding of the protein due to overloading of the membrane protein assembly infrastructure within the cell, such as the Sec61 translocon (55). NMR, on the other hand places a size limit on the protein of interest and is extremely difficult to perform on large molecules such as integral membrane proteins (56). The conventional techniques also require solubilization of the protein, which in the case of the membrane proteins involves the use of detergents to disrupt the membrane. However, this alteration of the protein's natural environment will lead to an alteration in the native state of the protein, and thus the structure determined using these

techniques will not necessarily be accurate. Newer approaches have therefore arisen for preferential use with membrane proteins. They include techniques such as Atomic force microscopy, Cryo-EM and Electron crystallography.

It is also useful to determine the membrane topology of the protein, which would not only give information regarding the number and orientation of the TM segments of the protein, but also the location of the N and C-termini of the protein with respect to the membrane. Some of the techniques used to determine membrane topology are listed below.

1.6.1. *Use of reporter proteins:*

The main principle in the use of reporter proteins to analyze membrane topology lies in the placement of the reporter at specific locations on the polypeptide chain, and then using the properties of the reporter protein to determine whether that site on the peptide chain lies in the cytosol or in other locations (e.g. the ER lumen). The reporter proteins may be tagged at the end of the protein, after truncating the protein to different lengths; or may be incorporated within the sequence of the protein. The former method has the obvious disadvantage of altering the length of the protein, which in turn may alter its structure from its native state. Incorporating the reporter in between would ensure the peptide chain remains full length but may also disrupt structure if too large. Examples of reporter proteins for eukaryotic systems include the alkaline phosphatase protein (PhoA), Suc2 invertase enzyme, and the His4 reporter among others.

The Suc2 reporter takes advantage of the numerous N-X-S/T N-linked glycosylation sites it has (57). If located within the ER lumen, the Suc2 reporter or a portion containing just the glycosylation domain would get extensively glycosylated, and this in turn produces a noticeable increase in size of the protein, which can be detected by Western. Since glycosylation occurs

only within the ER lumen, this can rule out a cytosolic location for the reporter protein. However, another possibility in such a case is the protein being misfolded and this affecting glycosylation. Hence a better approach is to use two reporters at once, with one specific for the cytosolic compartment and the other for the ER lumen. In yeast, a commonly used reporter specific for the cytosolic compartment is the His4 protein, a histidinol dehydrogenase enzyme. Presence of the His4 reporter in the cytosol helps convert histidinol into histidine, and thus enables the cell to selectively grow in a media containing histidinol as the sole source of histidine (57). A luminal location of the His4 reporter would lead to no growth in this selective media. It must also be mentioned that the His4 reporter can only be used in a His4 yeast knockout background.

The disadvantage of using reporters in determining membrane topology is the physical change that is produced by the incorporation of the reporter within the protein. In the case of Rce1p, which is 315 amino acids long, the change in length produced by an addition of Suc2 or the 53 amino acid size Suc2 glycosylation motif is considerable. Therefore, any topology information gleaned would have to be considered with these caveats. Similar caveats also apply to His4, which is much larger than Suc2. Furthermore, it should be remembered that in the case of N-Linked glycosylation, the Asn residue must lie on a loop that is larger than 25 amino acids long, and anything smaller than that length implies a decrease in glycosylation frequency (58). Therefore any lack of glycosylation observed using Suc2 reporters could be despite the reporter still being present on an intraluminal loop except that the loop is too short for glycosylation to occur.

1.6.2. *Use of computational algorithms:*

Computational algorithms that empirically predict a topology model for membrane proteins are in common use. The models predicted by these programs are not taken to be final; rather, these models are used as an aid upon which information gathered by other techniques are built on. They therefore give useful information such as the likely number of TM segments and their location on the protein sequence, the orientation of the N and C-termini, and at the fundamental level help to initially differentiate between a membrane protein and a soluble protein from the gene sequence. This information can then be used to select appropriate locations within the peptide sequence for placing reporter proteins or for inducing selective mutations for use in other techniques.

Computational algorithms usually predict models on the basis of hydrophobicity of amino acids and the expected free energy levels when these amino acids are placed within the hydrophobic core of the protein. It is based on the assumption that the native state of the protein is the state with the lowest free energy. Other factors, such as the ‘positive inside rule’ which is a statistical finding that shows the higher prevalence of positively charged amino acids in the cytosolic loops than in the luminal loops, are also taken into consideration (59). Newer computational algorithms are trained on a database of known sequences with experimentally derived topologies, and the salient features learnt from these are applied to the protein under study. Some programs like SCAMPI (60) avoid machine learning methods such as above and instead predict models using physical principles, such as estimating each amino acid’s contribution towards free energy on being inserted into the membrane. Other programs, like ConPred II (61) look for a consensus model from a list of models predicted by other algorithms such as TMHMM, MEMSAT, SOSUI, TMPred and HMMTOP. Table 2.1 shows the TM

segments predicted for Rce1p by some of the more commonly used topology prediction programs. Figure 2.3 also depicts the model predicted by TMHMM 2.0 in a cartoon form. Another computational technique that can be used to determine the structural model of a membrane protein is by using homology modeling, where homologs of the protein of interest are studied for their structural characteristics and this information applied to predict a model for this protein. In the case of Rce1p however, there are no homologs for which structures have been determined. This is a common problem. For example, four different topology models have been predicted for presenilin (62).

1.6.3. *Substituted cysteine accessibility method:*

This technique (63) operates on the same principle as reporter proteins, but avoids the disadvantages mentioned above in the use of reporters. Site directed mutagenesis techniques are used to systematically introduce cysteine residues into the protein sequence at specific points from one end of the protein to the other. Pre-existing cysteines are left alone, or substituted with other amino acids, preferable Ser, Thr or Ala to create numerous cysteine mutants having only one cysteine in their sequence. The membrane protein of interest is isolated in intact membrane preparations, enriched for the protein of choice. These membranes are then labeled with a reagent called Maleimide that covalently links to the thiol group on the cysteine residues. The Maleimide molecules are linked to long PEG polymer chains, and are thus membrane impermeable. Whenever a cysteine is labeled by a PEG-Maleimide, there is a detectable increase in molecular mass of the protein on account of the PEG group attached. By looking for this increase in molecular mass, following treatment with PEG-Maleimide both in the absence and presence of detergent, it can be determined if the particular site on the protein sequence that

bears the cysteine is on the exposed or luminal side of the ER membrane. Those cysteine residues that lie on the luminal side of the membrane get labeled by PEG-Maleimide only if the membranes are disrupted by detergent, thus making them accessible to the membrane impermeable PEG-Maleimide. The sites where the cysteines are introduced on the protein are usually chosen with guidance from the topology models predicted by computer algorithms mentioned above. This is to maximize the chances of selecting those sites which lie in accessible loops of the protein.

The main advantage of the substituted cysteine accessibility method is that introduction of single cysteines within the protein minimizes significant changes to the native state of the protein (63). Therefore the topology determined is more likely to be accurate. One of the handicaps of this technique however is its dependency on intact membrane in the membrane preparations. Any compromise in the integrity of the membrane can lead to inaccurate interpretations of the results. This can be overcome by use of a control to demonstrate the intact nature of the membranes. Another potential drawback that must be taken into consideration is the use of this technique in proteins that have cysteine pair disulfide bonds essential for maintaining the native structure of the protein. Creation of Cys-less mutants of these proteins would alter the native structure, and so any topology predicted will be inaccurate. This however can be ruled out by testing the functionality of all the cysteine mutants created, since an intact function implies an intact native structure of the protein (63).

1.7. Summary and Hypothesis

Mutations of the Ras protein leading to a constantly activated state are associated with over 30% of all cancers. The Ras converting enzyme (Rce1p) is an endoprotease that catalyses

an important step in Ras protein maturation. Inhibition of Rce1p attenuates Ras overactivity, and reduction in transforming ability of such cells, without affecting cell viability. Therefore Rce1p is being considered as a target for developing anticancer pharmacotherapy. However, neither its structure nor catalytic mechanism has been determined. Conventional methods to determine structure or membrane topology have not been effective. In this study, we hypothesize that through substituted cysteine scanning accessibility method, the membrane topology of the enzyme can be determined, and provide a step towards determining the catalytic activity of the enzyme. Such findings should enable further studies on Rce1p and pave the way for newer anticancer treatment modalities.

1.8. References

1. Harvey JJ: An Unidentified Virus Which Causes the Rapid Production of Tumours in Mice. *Nature* 204: 1104-5, 1964.
2. Der CJ, Krontiris TG and Cooper GM: Transforming genes of human bladder and lung carcinoma cell lines are homologous to the ras genes of Harvey and Kirsten sarcoma viruses. *Proc Natl Acad Sci U S A* 79: 3637-40, 1982.
3. Shimizu K, Goldfarb M, Perucho M and Wigler M: Isolation and preliminary characterization of the transforming gene of a human neuroblastoma cell line. *Proc Natl Acad Sci U S A* 80: 383-7, 1983.
4. Ehrhardt A, Ehrhardt GR, Guo X and Schrader JW: Ras and relatives--job sharing and networking keep an old family together. *Exp Hematol* 30: 1089-106, 2002.

5. Reddy EP, Reynolds RK, Santos E and Barbacid M: A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. *Nature* 300: 149-52, 1982.
6. Malumbres M and Barbacid M: RAS oncogenes: the first 30 years. *Nat Rev Cancer* 3: 459-65, 2003.
7. Zarbl H, Sukumar S, Arthur AV, Martin-Zanca D and Barbacid M: Direct mutagenesis of Ha-ras-1 oncogenes by N-nitroso-N-methylurea during initiation of mammary carcinogenesis in rats. *Nature* 315: 382-5, 1985.
8. Sinn E, Muller W, Pattengale P, Tepler I, Wallace R and Leder P: Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: synergistic action of oncogenes in vivo. *Cell* 49: 465-75, 1987.
9. Bos JL: Ras oncogenes in human cancer: A review *Cancer Research* 49: 4682-4689, 1989.
10. Lowy DR and Willumsen BM: Function and regulation of ras. *Annual Review of Biochemistry* 62: 851-91, 1993.
11. Willingham MC, Pastan I, Shih TY and Scolnick EM: Localization of the src gene product of the Harvey strain of MSV to plasma membrane of transformed cells by electron microscopic immunocytochemistry. *Cell* 19: 1005-14, 1980.
12. Cheng PC, Brown BK, Song W and Pierce SK: Translocation of the B cell antigen receptor into lipid rafts reveals a novel step in signaling. *J Immunol* 166: 3693-701, 2001.
13. Xavier R, Brennan T, Li Q, McCormack C and Seed B: Membrane compartmentation is required for efficient T cell activation. *Immunity* 8: 723-32, 1998.

14. Sheets ED, Holowka D and Baird B: Critical role for cholesterol in Lyn-mediated tyrosine phosphorylation of FcεRI and their association with detergent-resistant membranes. *J Cell Biol* 145: 877-87, 1999.
15. Lowenstein EJ, Daly RJ, Batzer AG, Li W, Margolis B, Lammers R, Ullrich A, Skolnik EY, Bar-Sagi D and Schlessinger J: The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell* 70: 431-42, 1992.
16. Quilliam LA, Rebhun JF and Castro AF: A growing family of guanine nucleotide exchange factors is responsible for activation of Ras-family GTPases. *Prog Nucleic Acid Res Mol Biol* 71: 391-444, 2002.
17. Boriack-Sjodin PA, Margarit SM, Bar-Sagi D and Kuriyan J: The structural basis of the activation of Ras by Sos. *Nature* 394: 337-43, 1998.
18. Yordy JS and Muise-Helmericks RC: Signal transduction and the Ets family of transcription factors. *Oncogene* 19: 6503-13, 2000.
19. Seth A, Gonzalez FA, Gupta S, Raden DL and Davis RJ: Signal transduction within the nucleus by mitogen-activated protein kinase. *J Biol Chem* 267: 24796-804, 1992.
20. Rodriguez-Viciana P, Warne PH, Dhand R, Vanhaesebroeck B, Gout I, Fry MJ, Waterfield MD and Downward J: Phosphatidylinositol-3-OH kinase as a direct target of Ras. *Nature* 370: 527-32, 1994.
21. Rodriguez-Viciana P, Warne PH, Khwaja A, Marte BM, Pappin D, Das P, Waterfield MD, Ridley A and Downward J: Role of phosphoinositide 3-OH kinase in cell transformation and control of the actin cytoskeleton by Ras. *Cell* 89: 457-67, 1997.
22. Michiels F, Habets GG, Stam JC, van der Kammen RA and Collard JG: A role for Rac in Tiam1-induced membrane ruffling and invasion. *Nature* 375: 338-40, 1995.

23. Feig LA: Ral-GTPases: approaching their 15 minutes of fame. *Trends Cell Biol* 13: 419-25, 2003.
24. Ahmadian MR, Stege P, Scheffzek K and Wittinghofer A: Confirmation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras. *Nat Struct Biol* 4: 686-9, 1997.
25. Adari H, Lowy DR, Willumsen BM, Der CJ and McCormick F: Guanosine triphosphatase activating protein (GAP) interacts with the p21 ras effector binding domain. *Science* 240: 518-21, 1988.
26. Xu GF, O'Connell P, Viskochil D, Cawthon R, Robertson M, Culver M, Dunn D, Stevens J, Gesteland R, White R and et al.: The neurofibromatosis type 1 gene encodes a protein related to GAP. *Cell* 62: 599-608, 1990.
27. Maekawa M, Li S, Iwamatsu A, Morishita T, Yokota K, Imai Y, Kohsaka S, Nakamura S and Hattori S: A novel mammalian Ras GTPase-activating protein which has phospholipid-binding and Btk homology regions. *Molecular and Cellular Biology* 14: 6879-85, 1994.
28. Cox AD and Der CJ: Ras family signaling: therapeutic targeting. *Cancer Biol Ther* 1: 599-606, 2002.
29. Ulsh LS and Shih TY: Metabolic turnover of human c-rasH p21 protein of EJ bladder carcinoma and its normal cellular and viral homologs. *Molecular and Cellular Biology* 4: 1647-52, 1984.
30. Moores SL, Schaber MD, Mosser SD, Rands E, O'Hara MB, Garsky VM, Marshall MS, Pompliano DL and Gibbs JB: Sequence dependence of protein isoprenylation. *J Biol Chem* 266: 14603-10, 1991.
31. Casey: Protein Lipidation in Cell Signaling. *Science* 268: 221-225, 1995.

32. Hancock JF, Paterson H and Marshall CJ: A polybasic domain or palmitoylation is required in addition to the CAAX motif to localize p21ras to the plasma membrane. *Cell* 63: 133-9, 1990.
33. Roy S, Luetterforst R, Harding A, Apolloni A, Etheridge M, Stang E, Rolls B, Hancock JF and Parton RG: Dominant-negative caveolin inhibits H-Ras function by disrupting cholesterol-rich plasma membrane domains. *Nat Cell Biol* 1: 98-105, 1999.
34. Winter-Vann AM and Casey PJ: Post-prenylation-processing enzymes as new targets in oncogenesis. *Nat Rev Cancer* 5: 405-12, 2005.
35. Tam A, Nouvet F, Fujimura-Kamada K, Slunt H, Sisodia SS and Michaelis S: Dual roles for Ste24p in yeast **a**-factor maturation: NH₂-terminal proteolysis and COOH-terminal CAAX processing. *J. Cell Biol.* 142: 635-49, 1998.
36. Zhang FL and Casey PJ: Protein prenylation: molecular mechanisms and functional consequences. *Ann Rev Biochem* 65: 241-69, 1996.
37. Boyartchuk VL, Ashby MN and Rine J: Modulation of Ras and **a**-factor function by carboxyl-terminal proteolysis. *Science* 275: 1796-800, 1997.
38. Hancock J, Cadwallader K and Marshall C: Methylation and proteolysis are essential for efficient membrane binding of prenylated p21K-ras(B). *EMBO J.* 10: 641-646, 1991.
39. Ashby M, King D and Rine J: Endoproteolytic processing of a farnesylated peptide *in vitro*. *Proc. Natl. Acad. Sci. USA* 89: 4613-17, 1992.
40. Freije JM, Blay P, Pendas AM, Cadinanos J, Crespo P and Lopez-Otin C: Identification and chromosomal location of two human genes encoding enzymes potentially involved in proteolytic maturation of farnesylated proteins. *Genomics* 58: 270-80, 1999.

41. Otto JC, Kim E, Young SG and Casey PJ: Cloning and characterization of a mammalian prenyl protein-specific protease. *J. Biol. Chem.* 274: 8379-82, 1999.
42. Cadiñanos J, Schmidt WK, Fueyo A, Varela I, Lopez-Otin C and Freije JM: Identification, functional expression and enzymic analysis of two distinct CaaX proteases from *Caenorhabditis elegans*. *Biochem J* 370: 1047-54, 2003.
43. Cadiñanos J, Varela I, Mandel D, Schmidt WK, Díaz-Perales A, López-Otín C and JMP F: AtFACE-2, a prenylated-protein protease from *Arabidopsis thaliana* related to Ras converting enzymes. *Journal of Biological Chemistry* 278: 42091-7, 2003.
44. Gillespie JR, Yokoyama K, Lu K, Eastman RT, Bollinger JG, Van Voorhis WC, Gelb MH and Buckner FS: C-terminal proteolysis of prenylated proteins in trypanosomatids and RNA interference of enzymes required for the post-translational processing pathway of farnesylated proteins. *Mol Biochem Parasitol* in press: 2007.
45. Bergo MO, Ambroziak P, Gregory C, George A, Otto JC, Kim E, Nagase H, Casey PJ, Balmain A and Young SG: Absence of the CAAX endoprotease Rce1: effects on cell growth and transformation. *Mol Cell Biol* 22: 171-81, 2002.
46. Kohl NE, Omer CA, Conner MW, Anthony NJ, Davide JP, deSolms SJ, Giuliani EA, Gomez RP, Graham SL, Hamilton K and et al.: Inhibition of farnesyltransferase induces regression of mammary and salivary carcinomas in ras transgenic mice. *Nat Med* 1: 792-7, 1995.
47. Whyte DB, Kirschmeier P, Hockenberry TN, Nunez-Oliva I, James L, Catino JJ, Bishop WR and Pai JK: K- and N-Ras are geranylgeranylated in cells treated with farnesyl protein transferase inhibitors. *J Biol Chem* 272: 14459-64, 1997.
48. Lobell RB, Omer CA, Abrams MT, Bhimnathwala HG, Brucker MJ, Buser CA, Davide JP, deSolms SJ, Dinsmore CJ, Ellis-Hutchings MS, Kral AM, Liu D, Lumma WC, Machotka

- SV, Rands E, Williams TM, Graham SL, Hartman GD, Oliff AI, Heimbrook DC and Kohl NE: Evaluation of farnesyl:protein transferase and geranylgeranyl:protein transferase inhibitor combinations in preclinical models. *Cancer Res* 61: 8758-68, 2001.
49. Chen Y: Selective inhibition of Ras-transformed cell growth by a novel fatty acid-based chloromethyl ketone designed to target Ras endoprotease. *Annals of the New York Academy of Sciences* 886: 103, 1999.
50. Schmidt WK, Tam A, Fujimura-Kamada K and Michaelis S: Endoplasmic reticulum membrane localization of Rce1p and Ste24p, yeast proteases involved in carboxyl-terminal CAAX protein processing and amino-terminal α -factor cleavage. *Proc Natl Acad Sci USA* 95: 11175-80, 1998.
51. Plummer LJ, Hildebrandt ER, Porter SB, Rogers VA, McCracken J and Schmidt WK: Mutational analysis of the ras converting enzyme reveals a requirement for glutamate and histidine residues. *J Biol Chem* 281: 4596-605, 2006.
52. Tam A, Schmidt WK and Michaelis S: The multispanning membrane protein Ste24p catalyzes CAAX proteolysis and NH₂-terminal processing of the yeast α -factor precursor. *J Biol Chem* 276: 46798-806, 2001.
53. Chen Y, Ma YT and Rando RR: Solubilization, partial purification, and affinity labeling of the membrane-bound isoprenylated protein endoprotease. *Biochemistry* 35: 3227-37, 1996.
54. Dolence JM, Steward LE, Dolence EK, Wong DH and Poulter CD: Studies with recombinant *Saccharomyces cerevisiae* CaaX prenyl protease Rce1p. *Biochemistry* 39: 4096-104, 2000.
55. Drew D, Froderberg L, Baars L and de Gier JW: Assembly and overexpression of membrane proteins in *Escherichia coli*. *Biochim Biophys Acta* 1610: 3-10, 2003.

56. Kainosho M, Torizawa T, Iwashita Y, Terauchi T, Mei Ono A and Guntert P: Optimal isotope labelling for NMR protein structure determinations. *Nature* 440: 52-7, 2006.
57. Deak PM and Wolf DH: Membrane topology and function of Der3/Hrd1p as a ubiquitin-protein ligase (E3) involved in endoplasmic reticulum degradation. *J Biol Chem* 276: 10663-9, 2001.
58. Nilsson IM and von Heijne G: Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J Biol Chem* 268: 5798-801, 1993.
59. Nilsson J, Persson B and von Heijne G: Comparative analysis of amino acid distributions in integral membrane proteins from 107 genomes. *Proteins* 60: 606-16, 2005.
60. Bernsel A, Viklund H, Falk J, Lindahl E, von Heijne G and Elofsson A: Prediction of membrane-protein topology from first principles. *Proc Natl Acad Sci U S A* 105: 7177-81, 2008.
61. Arai M, Mitsuke H, Ikeda M, Xia JX, Kikuchi T, Satake M and Shimizu T: ConPred II: a consensus prediction method for obtaining transmembrane topology models with high reliability. *Nucleic Acids Res* 32: W390-3, 2004.
62. Kim J and Schekman R: The ins and outs of presenilin 1 membrane topology. *Proc Natl Acad Sci U S A* 101: 905-6, 2004.
63. Bogdanov M, Zhang W, Xie J and Dowhan W: Transmembrane protein topology mapping by the substituted cysteine accessibility method (SCAM(TM)): application to lipid-specific membrane protein topogenesis. *Methods* 36: 148-71, 2005.

Fig. 1.1. *Ras signaling pathways.*

Growth factor receptors (e.g. RTKs) on the plasma membrane dimerize and autophosphorylate themselves on binding to their ligand. SH2 domain bearing modulator proteins such as Grb2 recognize the phosphorylated Tyr residues and bind to these RTKs. Ras GEFs, such as Sos1, are recruited by the Grb2 protein, which in turn promote GDP to GTP exchange and activate Ras protein that is attached to the membrane via its lipid. Activated Ras then transduces the signal further downstream via a wide variety of effectors, viz Raf, PI3K, RalGEF and Tiam1, among others, to modulate the effects of their respective cellular pathways. Ras GAP proteins ‘switch off’ the signal by promoting intrinsic GTPase activity of Ras, thus taking it back to the inactive GDP bound state.

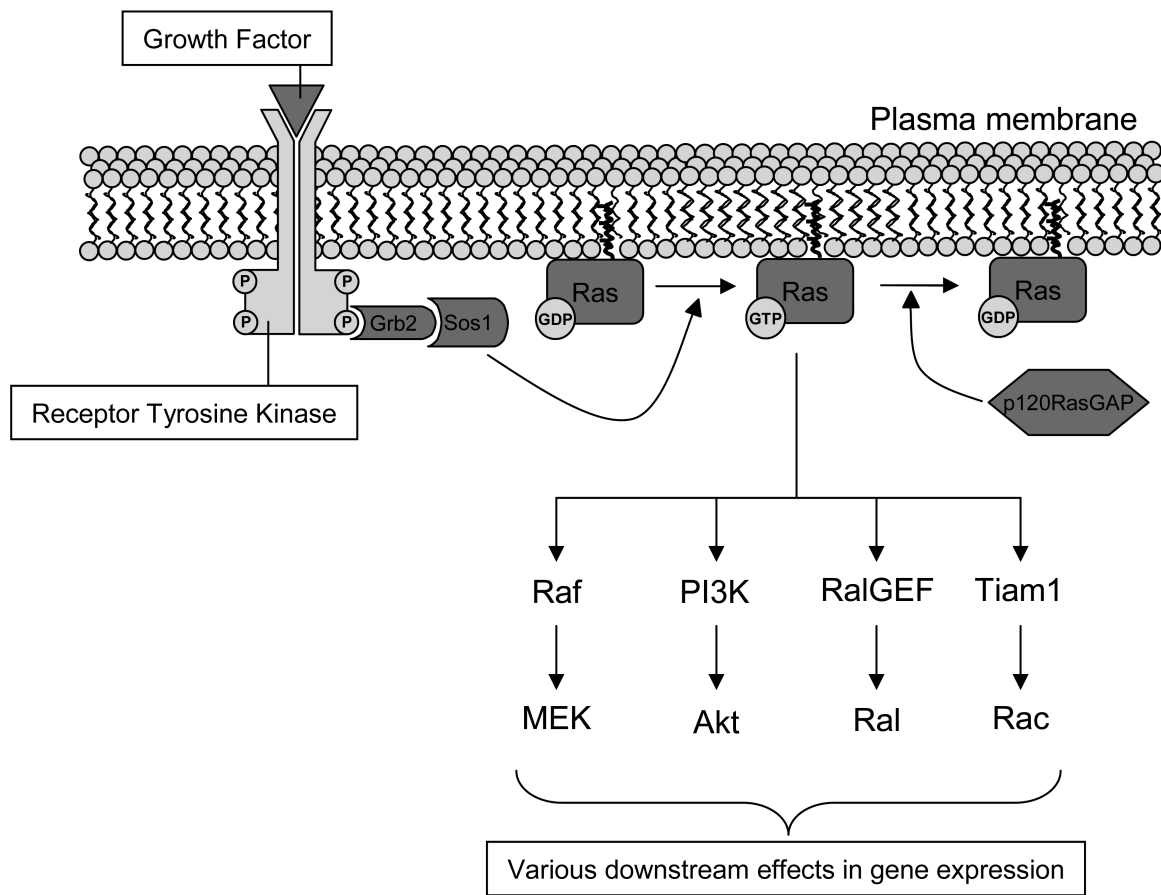
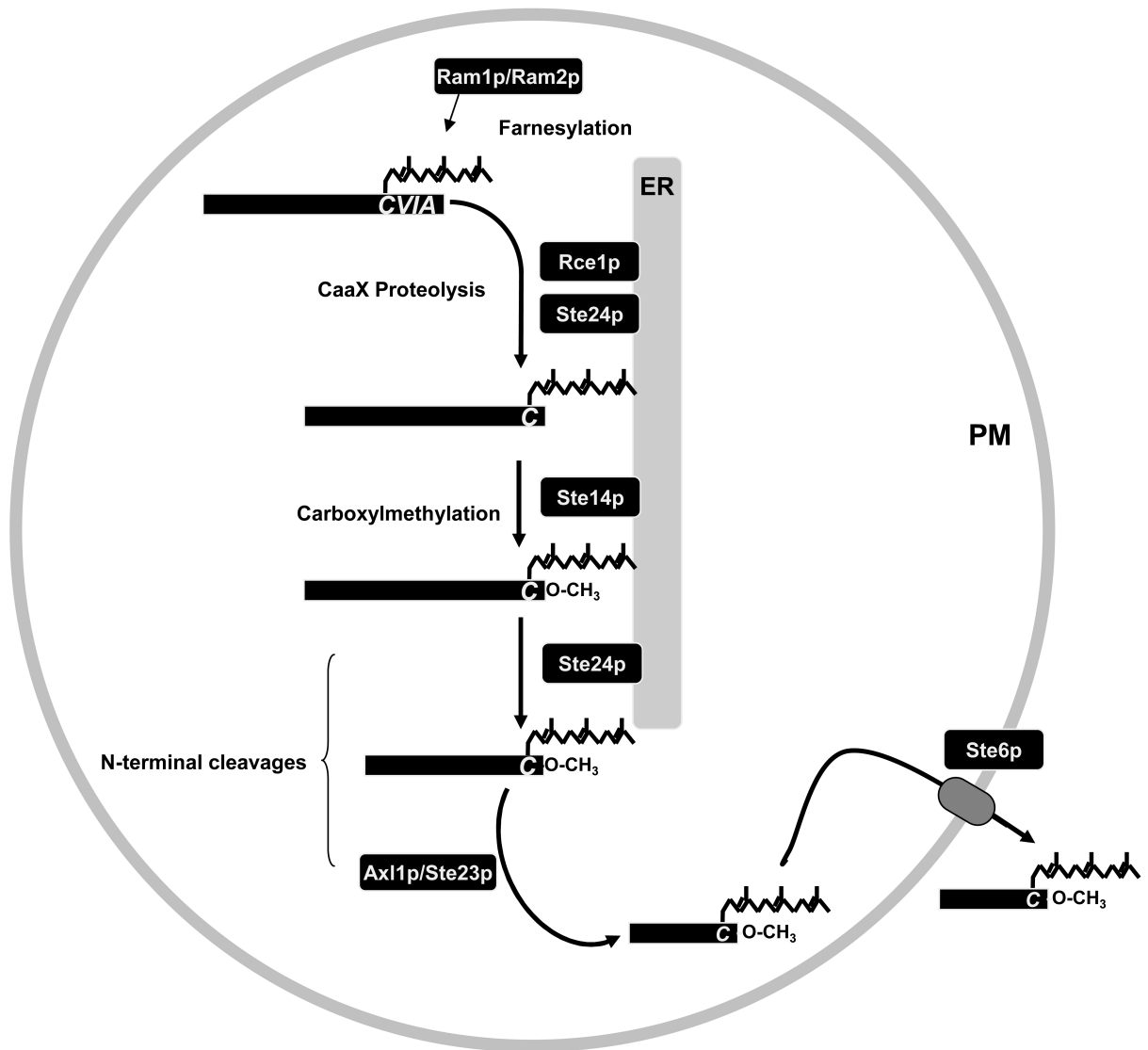


Fig. 1.2. *a*-factor biogenesis pathway.

Cartoon depicting biogenesis of *S. cerevisiae* **a**-factor. Following translation of mRNAs derived from redundant *MFA1/MFA2* genes, the precursor **a**-factor undergoes farnesylation by the Ram1p/Ram2p farnesyltransferase. This is followed by CaaX proteolysis by the ER-localized Rce1p or Ste24p enzymes, which cleave the VIA tripeptide from the CVIA motif. Either Rce1p or Ste24p can mediate this step, but the enzymes are otherwise unrelated. Ste14p completes the modification pathway by transferring a methyl group to the C-terminus containing the farnesylated cysteine. The above constitute the CaaX processing steps common to all CaaX proteins. In the case of **a**-factor, two successive N-terminal proteolytic events take place, first by Ste24p and then by Axl1p or Ste23p, which trim the **a**-factor precursor down to 12 amino acids length. The mature form of **a**-factor is ultimately secreted out of the cell via the Ste6p ABC transporter.



CHAPTER 2

TOPOLOGY OF RAS CONVERTING ENZYME (RCE1P) AS DETERMINED BY SUBSTITUTED CYSTEINE ACCESSIBILITY METHOD¹

¹Ranjith K. Krishnankutty, Emily Hildebrandt, John Deaton, and Walter K. Schmidt
To be submitted to *The Journal Of Biological Chemistry*

INTRODUCTION

The Ras converting enzyme (Rce1p) is a proteolytic enzyme involved in the processing of the carboxy terminal 'CaaX' motif found in prenylated proteins. Isoprenylation is a common post-translational modification seen in membrane-associated proteins. The tetrapeptide CaaX motif, where 'C' represents cysteine, 'a' indicates any aliphatic amino acid and 'X' stands for one of several amino acids, is a common site for attachment of an isoprene derived group. Once the isoprenoid group is covalently linked to the cysteine of the CaaX motif by the enzyme farnesyltransferase, the 'aaX' tripeptide is cleaved off by one of the two known CaaX proteases, Rce1p or Ste24p. Proteolysis is then followed by carboxylmethylation by isoprenylcysteine carboxyl methyltransferase, which completes the modification pathway of most CaaX proteins (1, 2). Isoprenylated proteins are seen in all eukaryotic organisms and play vital roles in cellular functions. Examples of isoprenylproteins include signaling molecules (*i.e.*, Ras and RhoB), nuclear proteins (*i.e.*, CENP-E, CENP-F, and nuclear lamins), and fungal mating pheromones (*e.g.*, *S. cerevisiae* a-factor). The Ras protein is a G-protein that plays a significant role in signal transduction and development of cancer. Ras mutations are associated with over 30% of all cancers (3).

The role played by Rce1p in Ras protein maturation is the basis for investigating the enzyme as a means to develop anti-cancer pharmacotherapy. Although farnesyltransferase inhibitors (FTI) are currently being studied in clinical trials for efficacy, data has not conclusively shown any advantages for using FTIs as anti-cancer drugs, for reasons outlined in the introduction to this thesis. (4). Inhibition of Rce1p is also being considered as an alternative therapeutic approach. However, the lack of definitive structural information regarding Rce1p has

hampered this approach. Hence, determining the structure of Rce1p is expected to help in finding more efficient ways to inhibit the enzyme.

The difficulty in resolving the structure of Rce1p lies in the fact that it is an integral membrane protein, located in the ER membrane (5). Conventional methods for structural determination such as X-Ray crystallography and Nuclear Magnetic Resonance spectroscopy while theoretically feasible, are technically difficult due to the need for purified enzyme. Moreover, purification disrupts the enzyme's native environment, which in turn may affect the physiological function of the enzyme. To date, Rce1p has not been purified in useful biochemical quantities. Therefore, alternative techniques that retain the enzyme's native environment and thus its function, are preferred for assessing protein structure. One such method is the substituted cysteine accessibility method (SCAM). The basis of SCAM is the strong, covalent linking of a sulfhydryl specific agent, such as maleimide, to the thiol group on specific cysteine residues on the protein of interest (Figure 2.1), and then deducing the topology of the protein from the accessibility of that particular cysteine residue to the sulfhydryl specific agent (6). Here, we have used a membrane impermeable form of maleimide, PEG-Maleimide with yeast microsomal membranes containing Rce1p to study the orientation of the transmembrane loops of Rce1p. Accessibility of the cysteine to the PEG-Maleimide can be deduced by an apparent increase in the molecular mass of the enzyme, as seen on a Western blot, owing to the 5 KDa size of the PEG group in PEG-Maleimide (Figure 2.2). In such instances, a reactive cysteine would be considered as cytosolically disposed.

Various computer transmembrane (TM) topology prediction models have suggested Rce1p to have from six to eight TM loops (See Table 2.1) (7). In order to determine the membrane topology of Rce1p using the SCAM method, we have used a triple HA tagged

Rce1p protein that can be detected on Western blot using the anti-HA antibody. Figure 2.3 depicts a putative model topology for Rce1p as predicted by the TMHMM 2.0 program. The eight naturally occurring cysteine residues in this protein were first substituted out to create a cysteine-less mutant. Subsequently, cysteine residues were substituted into specific sites on the primary sequence of the protein, those which were predicted by this model to lie within extra-membrane loops of the protein. By using SCAM, we were able to determine the orientation of these substituted cysteine sites with respect to the membrane, and thus deduce a proposed topology of the Rce1p.

EXPERIMENTAL PROCEDURES:

1. Strains: All strains used in this study are listed in Table 2.2. Strains yWS1114, yWSS787 - yWS794, yWS977 - yWS981, and yWS1113 were created by transforming strain SM3614 with the plasmids pWS781, pWS787 - pWS794, pWS864 - pWS868 and pWS911, respectively, using yeast transformation methods previously described (8). Yeast strains yWS1019 - yWS1032 and yWS1117 were created by transforming strain yWS198 with the plasmids pWS822 - pWS829, pWS875 - pWS879 and pWS933 respectively. All strains were routinely cultured on appropriate synthetic complete dropout (SC-) media at 30 °C.

2. Plasmids: All plasmids used in this study are listed in Table 2.3. Plasmids pWS781, pWS787 - pWS794 were created by Dr. John Deaton, while the plasmids pWS822 - pWS829, pWS864 - pWS868, and pWS875 - pWS878 were created by Dr. Emily Hildebrandt. The plasmids pWS781, pWS787 - pWS794, pWS864 - pWS868 and pWS911 were all created using the Quickchange site-directed cysteine mutagenesis protocol. The Cys-less mutant pWS787, lacking

the naturally occurring eight Cys codons in RCE1::HAc were created by sequentially replacing Cys with Ser using this method. Plasmid pWS781 was created when seven of the naturally present eight Cys codons in the RCE1::HAc ORF were substituted, barring the Cys codon located at position 354. pWS787 was subsequently used as a template to insert Cys codons at specific sites on the primary sequence using Quickchange. Cys residues were introduced to replace Ser residues at sites 5, 30, 71, 96, 145, 175, 187, 206, 260 on the primary sequence, yielding pWS788, pWS789, pWS790, pWS791, pWS865, pWS866 and pWS868 respectively. Cys residues were introduced to replace Thr residues at sites 155 and 314 on the primary sequence to create pWS792 and pWS911 respectively. A Phe residue at site 234 was replaced with Cys to create pWS794, while an Ala residue at site 249 was replaced with Cys to create pWS867.

Plasmids pWS822 - pWS830, pWS875 - pWS879 and pWS933 are high copy 2μ versions of plasmids pWS781, pWS787 - pWS794, pWS864 - pWS868 and pWS911 respectively. They were created by subcloning the Not1-Xho1 fragment from the appropriate low copy *CEN* plasmids into pRS426.

3. *Yeast mating assay*: *MATa* strains expressing Rce1p cysteine mutants and *MAT α* strains were grown in appropriate liquid selection media to mid-log phase, and cell density matched to A_{600} of 1.00 by dilution with fresh media. The *MAT α* cells were dispensed in a 96 well microtiter plate with 90 μ l in each well for 5 rows. To the first row of wells, 10 μ l of the appropriate *MATa* strain was added. After mixing with a pipettor, 10 μ l was then taken from the 1st row of wells and added to the 2nd row. After mixing, 10 μ l was transferred from the 2nd row and added to the 3rd and so on, till the last row of wells was mixed. This method created a 10x dilution gradient of *MATa* cells. 5 μ l of cells from each well was then spotted on a minimal media plate and colony

growth observed after 48 hours. 5 μ l of cells from each well were also loaded on a SC-Lys plate to verify the 10x dilution gradient of *MATa* cells in each column of wells.

4. Protein extract preparation: The Rce1p protein was isolated in yeast microsome preparations. Yeast cells were grown to mid-log phase in appropriate liquid selection media, harvested by centrifugation at 3000g for 5 min, washed with cold 10 mM sodium azide, and reisolated by centrifugation. The cell pellet obtained was resuspended in lysis buffer (800 mM Mannitol, 100 mM KCL, 50 mM Tris at pH 7.5, 1 mM EGTA) and transferred to a microfuge tube, containing silica beads. The mixture was mixed on a vortexer for 4 min, chilled on ice for 2 min, and the cycle repeated until 16 min of mixing was achieved. The lysed cells were clarified of cell debris by centrifugation, and the supernatant centrifuged at 16000g for 15 min to obtain the membrane pellet. The pellet was resuspended in lysis buffer by gentle, but repeated pipetting. The protein concentration in the membrane extracts was determined using the Biorad Protein Assay and were then adjusted to 1 mg/ml by dilution with lysis buffer.

5. PEG-Maleimide treatment and detection: Membrane extracts (50 μ g of protein) were incubated in two sets on ice, with lysis buffer and 1 mM methoxypolyethylene glycol 5,000 maleimide (PEG-Maleimide) (Sigma-Aldrich). In one set of reaction mixtures, Triton X-100 (TX-100) was added at 0.25% concentration (final), and in the other, equivalent amounts of lysis buffer added. After incubating the mixtures for 1 hr on ice, the reactions were quenched by adding 1 μ l of 1 mM β -mercaptoethanol and further incubated on ice for 10 min. 2X Sample buffer (6% w/v SDS, 24% w/v Sucrose, 75 mM Tris at pH 6.8, 3% β -mercaptoethanol, 0.01 % Bromophenol blue) was added, and the samples were run on 12.5% SDS-PAGE gels, transferred

to nitrocellulose membranes, and blots probed by using anti-HA antibodies (1:10,000) for Rce1p and anti-Kar2p antibodies (1:10,000) for Kar2p.

RESULTS

Functionality of Rce1p cysteine mutants:

As the main premise of using SCAM to determine topology of Rce1p is that the protein's native environment and functionality will be retained, it had to be ensured that all cysteine mutants constructed were actually functional. This was verified by performing a yeast mating test. As mentioned above, yeast **a**-factor mating pheromone is an isoprenylated protein, and its maturation and CaaX motif processing mimics that of Ras proteins. Production of normal amounts of mature **a**-factor is essential for haploid *MATa* yeast to mate with a *MATα* strain to form a diploid. This requires that the CaaX proteolytic activity must be functional in a *MATa* strain. Therefore, in order to verify functionality of the various cysteine mutants of Rce1p enzyme, a *MATa* yeast strain disrupted for genomic *RCE1* and *STE24* genes was transformed with plasmids encoding mutant *RCE1* genes, and the resultant strains evaluated for mating with a *MATα* strain.

The Rce1p mutants created for the SCAM experiment include S5C, S30C, S71C, S96S, S145C, T155C, S175C, S187C, S206C, F234C, A249C, S260C, T314C and 354C, in an otherwise cysteine-free background. In addition, a Cys-less Rce1p mutant was created in order to compare the Rce1p functionality of the Cys-less Rce1p protein with that of WT and the other cysteine mutants via its mating properties. Figure 2.4 shows the results of the dilution mating test. Panel A shows mating on synthetic dextrose (SD) solid media, which selects for diploid colony growth. All the cysteine mutants, except for F234C show functional Rce1p activity, as

seen by diploid colony formation. By comparing the dilution series as a whole, differences in mating efficiency can be identified. For example, S71C, S145C, S187C and S260C showed lower mating activity than the other mutants, implying that these Rce1p mutants were less functional than the others, which in turn implies that SCAM results these mutants has to be taken with that much more circumspection. Incidentally, it was also considered at one point of time that the Rce1p enzyme was a cysteine protease (9). This hypothesis is clearly disproved by the strong mating activity that is observed for the Cys-less mutant, which is comparable to that of the wild-type. Panel B, which shows haploid *MATa* cells growing on SC-Lysine solid media, serves as a control for the 10 fold dilutions of *MATa* cells prepared in Panel A.

Optimization of Mannitol concentration in Lysis buffer using Kar2p control:

The results from the SCAM experiment are interpreted on the basis of whether or not each cysteine residue introduced into the Rce1p protein sequence is accessible to PEG-Maleimide in the absence and presence of detergent. The detergent serves to disrupt the membrane, thus making previously inaccessible cysteine residues accessible. Therefore, in order to accurately interpret the results from the SCAM experiment, the membrane integrity of Rce1p-containing microsomes in the absence of detergent is essential. Hence, as a control to monitor the intact nature of these ER microsomes, the intralumenal ER protein Kar2p was chosen. Kar2p is a 74.5 KDa chaperone protein that bears one cysteine residue in its sequence. Since it is an intralumenal protein and separated from the cytosol by the ER membrane, its cysteine residue should not be accessible to PEG-Maleimide molecules and thus would not be modified by the reagent in the absence of detergent. Only in the presence of detergent would it be accessible to

the membrane impermeable PEG-Maleimide and thus be modified, which would be observed as an increase in molecular mass on Western detection.

Figure 2.5 shows a Western blot displaying modification of Kar2p in the presence and absence of detergent. Since membrane integrity depends on osmolarity of the Lysis buffer used, various concentrations of Mannitol were evaluated for the Lysis buffer. As seen on the left panel, at lower mannitol concentrations, the microsomes were not intact, as seen by the PEG-Maleimide modification of Kar2p in the absence of detergent. With increasing concentrations, it was established that microsomal extracts prepared with Lysis buffer containing 0.8 M Mannitol contained intact microsomes, as seen by the absence of modification of Kar2p in the absence of detergent. The right panel displays the modification observed when microsomes are treated with PEG-Maleimide in the presence of detergent.

PEG-Maleimide treatment of Rce1p cysteine mutants in the absence of detergent:

As a first step towards determining the membrane topology of the Rce1p enzyme, the compartmental localization of cysteine residues introduced in the Rce1p cysteine mutants was assessed in the absence of detergent. A total of fourteen different single cysteine mutants which were created, were evaluated. The wild-type enzyme was used as a positive control, while a Cys-less mutant was used as a negative control. Following treatment with PEG-Maleimide, the reaction was quenched in excess β -mercaptoethanol, and samples evaluated by Western blot to detect the Rce1p enzyme and the control Kar2p protein.

PEG-Maleimide is a membrane impermeable compound due to its 5 KDa PEG moiety. Thus any Rce1p cysteine mutant protein that is modified in the absence of detergent must have a Cys residue exposed to the reaction buffer containing PEG-Maleimide. The bead beating method

employed to isolate ER microsomes in this study produces right side out microsomes. In other words, the intraluminal compartment within these microsomes is the same as the intraluminal compartment within the endoplasmic reticulum of cells. The extraluminal compartment in these membrane preparations is continuous with the cytosolic compartment of the cell. Therefore, this implies that any Rce1p cysteine mutant that is modified by PEG-Maleimide must have its cysteine lying in the extraluminal compartment (*i.e.*, on the cytoplasmic face of the protein). Also, the cysteinyl thiol group, to which the Maleimide binds to, lies at a distance of 8-10 Å away from the peptide backbone, so this implies that the PEG-Maleimide will have no access to any cysteine residues deeper than 8-10 Å from the membrane surface, (*i.e.*, a position on the TM segment that is 5-6 amino acids distant from the cytosolic membrane surface). If the cysteine is not modified however, there could be three explanations. The simplest explanation is that the cysteine lies on the intraluminal face. Another explanation could be that the cysteine lies within a TM segment traversing the membrane, deep enough to be inaccessible to the PEG-Maleimide. Lastly, it is possible that the cysteine is on the cytosolic face, but is in such a position that is tightly packed against another portion of Rce1p and thus protected against modification.

Figure 2.6 shows the fate of Rce1p cysteine mutants following the PEG-Maleimide treatment in the absence of detergent. The wild-type protein (lane 1) is modified, as is expected from the high probability that at least one of the eight naturally occurring Cys residues present in the Rce1::HAc ORF has a cytoplasmic orientation. Modified Rce1p (lane 1) shows an increase in molecular size, owing to the 5 KDa PEG group linked to the cysteine residue on the protein via the Maleimide compound. The protein is not completely modified, as indicated by the presence of unmodified protein on the Western blot. The reasons for incomplete modification are

unknown. The Rce1p Cys-less mutant is not modified, as expected. Lane 17, which is an untreated sample is also not modified and serves as an additional negative control.

Many cysteine mutants appear to be modified by the PEG-Maleimide with the exception of S5C (lane 3) and A249C (lane 13). The strongest shifted bands were observed with S30C, S71C, S145C, T155C, S175C, S206C, T314C and C354. The mutants S96C, S187C, F234C and S260C are modified relatively weakly, as is evident from the lesser intensity of the shifted band in the lanes 6, 10, 12 and 14 respectively. Kar2p, which was used to demonstrate intact membranes in the absence of detergent, is not modified by PEG-Maleimide in any of the sixteen lanes, thus implying that the ER microsomes were intact in all samples used during the experiment.

PEG-Maleimide treatment of Rce1p cysteine mutants in the presence of detergent:

The next step in the application of the substituted cysteine accessibility method (SCAM) to determine membrane topology of Rce1p was to confirm that the various Rce1p cysteine mutants were modifiable in the presence of detergent. The technique followed was the same as previously described in the absence of detergent, except that 0.25% Triton X-100 (final concentration) was added to the reaction mix. Since detergents disrupt membranes, addition of Triton X-100 to the reaction mix leads to disruption of the ER microsomes and makes all intraluminal contents accessible to solution reagents (*i.e.*, PEG-Maleimide).

Figure 2.7 shows the effect of PEG-Maleimide modification of Kar2p and the various Rce1p cysteine mutants in the presence of Triton X-100. As expected, the negative controls in lanes 2 and 17 representing the Rce1p Cys-less mutant and the untreated sample respectively, shows no modification. All other samples, including all the Rce1p cysteine mutants and the wild-

type Rce1p are labeled by the PEG-Maleimide. The control protein Kar2p however, does not show complete modification in all lanes. While it is distinctly clear in lanes 1, 5-11, 14 and 15, it is only faintly visible for lanes 4 and 16. Lanes 2, 3, 12 and 13 however, are not modified by PEG-Maleimide.

DISCUSSION

The Ras converting enzyme is an important member of the Ras protein maturation pathway. As a membrane protein localized to the ER membrane, it is one of the few proteases that are also integral membrane proteins with multiple membrane spans. Such proteins are a challenge to purify for studying their enzymatic activity and biochemical properties. Conventional techniques used to study enzymes and enzymatic activity, are not readily applicable to the membrane embedded proteins. For example, X-Ray crystallography cannot usually be applied to investigate structures of activation states and complex formation with substrates.

The inability to purify Rce1p has further hampered biochemical investigations of Rce1p. Since the three dimensional structure of Rce1p cannot be determined at this stage, a step forward to identifying the catalytic site of the enzyme would be to determine its membrane topology. Since three amino acids essential for enzyme activity have already been identified (Glu¹⁵⁶, His¹⁹⁴ and His²⁴⁸ (7)), determining the topology will provide a better understanding of their orientation across the membrane, and thus potentially valuable insight regarding the disposition of the active site.

We have used the substituted cysteine accessibility method (SCAM) to determine the topology of Rce1p. This method has the advantage of causing minimal change in the protein's

native state and structure, and does not disrupt the native membrane environment of Rce1p, both of which are essential for its function. In this endeavor, we created one cysteine-less mutant and fourteen cysteine mutants, each having a single cysteine at a specific location that ranges from position 5 to position 354 on the C-terminal HA tag linked to protein for use in detection. The cysteine residues were introduced into the Cys-less mutant of the protein mostly at locations that were predicted by topology modeling programs to be exposed elements of the protein (*i.e.*, not transmembrane segments).

All the cysteine mutants created were tested for functionality. This was done to ensure that there were no essential cysteine residues or structurally important disulfide bonds. All except one of the cysteine mutants created were functional. The lack of mating observed for F234C is not due to lack of function, since the expression of Rce1p was seen to be normal (Fig. 2.6 and Fig. 2.7). Previous mutational studies on the Rce1p enzyme (7) did not evaluate the F234 position on the yeast enzyme. It is however a highly conserved position, with the identical F residue present in four out of five Rce1p homologs studied, the residue being a Gln in the outlier. This study also demonstrates the near wild-type level functional activity of the Cys-less mutant of Rce1p. This conclusively rules out the initially considered hypothesis that Rce1p is a cysteine protease (9).

All cysteine mutants were evaluated for reactivity with PEG-Maleimide in the absence and presence of detergent. Position 5 was shown to not lie within the intracytoplasmic compartment as it did not show labeling with PEG-Maleimide in the absence of detergent, even as the Kar2p control showed no leakage from the microsomes. This implies that position 5 lies in the intraluminal face of the protein, or in a TM segment within the hydrophobic core of the membrane, where PEG-Maleimide is unable to access it. Also, since PEG-Maleimide is expected

to have access up to 5-6 amino acids from the cytosolic surface, this suggests that the N-terminus of the protein lies either in the intraluminal compartment or is within the hydrophobic core of the membrane. This is however contrary to the orientation proposed by topology models predicted by computational algorithms such as TMHMM 2.0 (Figure 2.3). A better idea of the location of position 5 can be obtained by using a Maleimide reagent such as N-ethylmaleimide (NEM), which is specific for covalent linkage to thiolate anions (10), in conjunction with PEG-Maleimide. NEM labels only those cysteines that lie on the extra-membrane regions of the peptide chain, as compared to those cysteines that lie within the hydrophobic core of the membrane. Since the modification of Cys residues by NEM is not detectable by Western, owing to the small size of the NEM molecule, it can be used as a blocking agent prior to the application of PEG-Maleimide. The readily membrane permeable NEM molecule, therefore modifies all cysteine residues lying outside the hydrophobic membrane, both on the cytoplasmic and intraluminal face of the protein. If this NEM treatment is followed by treatment with PEG-Maleimide in the absence of detergent, none of the normally modified cytoplasmic Cys residues will be modified by PEG-Maleimide. Instead, if it is followed by treatment with PEG-Maleimide in the presence of detergent, only those cysteines that lie within the hydrophobic core of the membrane will be visibly modified by PEG-Maleimide. None of the intraluminal Cys residues would be modified because of the prior labeling by NEM.

Positions 30 and 71, on the other hand, are readily labeled by PEG-Maleimide, which implies that they are likely to lie on the cytosolic face of the protein. In the case of position 96, the Rce1p cysteine mutant appears to be lightly modified in the absence of detergent, indicating a position partially accessible from the cytoplasmic face of the membrane, such as within a TM segment, but at a distance less than five amino acid residues from the membrane surface.

Positions 145, 155, 175 and 206 show results similar to that of positions 30 and 71, and are therefore likely to be located on the cytosolic face of the protein.

In the case of positions 187, 234 and 260, these cysteine mutants are lightly modified in the absence of detergent, without showing any evidence of leaky membranes. They are therefore likely to be in locations similar to that of position 96, which are only partially accessible to the PEG-Maleimide, such as a site on a TM segment traversing the membrane, but within the 5-6 amino acids cut-off distance from the cytosolic surface. It must be remembered however, that the F234C Rce1p mutant was nonfunctional, which suggests an altered structure for that mutant. Hence any inference of the orientation of position 234 should be made with caution. Position 249 on the other hand, shows no modification at all by PEG-Maleimide, with the Kar2p control showing no leakage. This implies that this position, similar to that of position 5, lies either in the intraluminal loop of the protein, or lies within the hydrophobic core of the membrane, at a site inaccessible to PEG-Maleimide. Positions 314 and 354 appear to be strongly modified, which indicates that, similar to positions 30 and 71, they also lie within the intracytoplasmic face of Rce1p.

One caveat to be observed is the absence of modification of the Kar2p control in the presence of detergent for lanes 3, 12 and 13 (Figure 2.7). There is no clear explanation for this observation, but this could raise a question with regards to the validity of the use of Kar2p control to demonstrate the intact nature of membranes. However, the presence of Kar2p modification seen in all the other lanes in the presence of detergent (Figure 2.7), along with the uniformly observed Rce1p modification in all these lanes, shows that the lack of Kar2p labeling in lanes 3, 12 and 13 is likely to be due to conditions restricted to these samples, and do not

affect the validity of the use of Kar2p control in general. Furthermore, previous observations for these lanes show similar Rce1p modification results, but with modified Kar2p controls.

The above findings have been used to construct a putative topology model for Rce1p (Figure 2.8A). When this model is compared to that predicted by the TMHMM 2.0 algorithm for the enzyme, there are a few differences to be noted. As mentioned above, the N-terminus is suggested to lie in the cytosol by the TMHMM 2.0 model (Figure 2.3), while our results indicate that it is likely to lie on the intraluminal face of the enzyme or within the hydrophobic core of the membrane. Also, the second and third TM segments predicted by the TMHMM 2.0 program, extending from residues 44-61 and 76-93, are shown to be absent in the SCAM model. This is due to the fact that all three positions, 30, 71 and 96 have been observed to lie on the cytoplasmic face of the membrane. It should be noted that there is a clear consensus among all the five computational algorithms considered in this study (Table 2.1) with regards to the presence of the second TM segment. Also the third TM segment has been predicted by two of the five programs, TMHMM 2.0 and ConPred II. In addition to the N-terminus, there is also a discrepancy with respect to the orientation of the C-terminus of the enzyme. According to the TMHMM 2.0 model, the C-terminus of the enzyme, along with the HA tag, is predicted to lie within the intraluminal compartment. The SCAM model however implies that the C-terminus lies within the intracytoplasmic compartment. This is also in contrast to observations from previous studies that used a HA/Suc2p/His4c dual activity topology reporter tagged to the C-terminal end of the protein (Bhatt, S. et al, unpublished observation), which suggested an intraluminal orientation for the C-terminus. Also, since both positions, 260 and 314, lie on the cytoplasmic face of the enzyme, another possibility is the presence of two TM spans between these two positions (Figure 2.8B). This alternate model will bring the number of TM segments to seven, from the five

predicted by the model shown in Figure 2.8A. The similarities observed between both SCAM models and the TMHMM 2.0 model are the orientation of the positions 145, 155, 175, 187, 206, 249 and 260.

It has previously been considered that the active site of the Rce1p enzyme is most likely to lie on the cytoplasmic face of the protein or partially buried within the hydrophobic core of the membrane, with the isoprenoid group linked to the CaaX motif on the substrate serving to present the peptide bond to the active site of the enzyme from the cytoplasmic face of the membrane (5, 11). Two of the three residues found to be essential for Rce1p function, Glu¹⁵⁶ and His¹⁹⁴, are both predicted to lie on an intracytoplasmic loop of the protein according to the TMHMM 2.0 model. This prediction is supported by the SCAM model, which also shows that these residues are likely to lie on the cytoplasmic face of the enzyme. Therefore, one possibility from this result is that these residues play a role in constituting the active site of the enzyme, with the His residue serving as one of the coordinating amino acids of the metal ion. Alternately, these two residues may be essential for the maintenance of the native structure of the enzyme, and thus be essential for functional activity. The third essential amino acid, His²⁴⁸, lies adjacent to position 249, whose location was determined by SCAM to lie either within the intraluminal compartment or within the hydrophobic core of the membrane (Figure 2.8A). A TM segment location within the hydrophobic core of the membrane for position 249 will correlate with the location predicted by the TMHMM 2.0 model. The His²⁴⁸ residue may therefore either play a role in maintaining the native structure of the enzyme, with that particular TM segment being an essential segment, or it could even be constituting the active site of the enzyme, with the active site partially buried within the hydrophobic part of the membrane. Phe²³⁴, which was found to be

essential for enzyme activity in this study, may also be playing a role in the maintenance of the native structure of the enzyme.

Although the SCAM results have enabled the construction of two topology models for the enzyme, further refinement is needed. This can be achieved by using the same SCAM technique, however, for newer cysteine mutants of the Rce1p enzyme, that would help in determining the orientation of those sites on the Rce1p primary sequence that are yet to be localized. These newer Rce1p cysteine mutants include those that have Cys residues introduced at the positions marked in Figure 2.8B. They include positions such as a site on the intralumenal loop between the second and third TM segment, a site on the intracytoplasmic loop between the fifth and sixth TM spans, and a site on the intralumenal loop between the sixth and seventh TM spans. In addition, a site adjacent to Phe²³⁴ on the intralumenal loop between the fourth and fifth TM spans can also be investigated to determine the essential nature of that loop for the function of the enzyme. Also, the orientation of the N-terminus can be confirmed by using N-terminal tagged reporter proteins.

In summary, the substituted cysteine accessibility method used to determine membrane topology of the Rce1p enzyme shows evidence that the N-terminus of the protein lies intralumenally, with position 249 placed either intralumenally or within the hydrophobic core of the membrane. It also suggests that positions 145, 155, 175, 187, 206, 314 and 354, all lie within the cytosolic side, and the two essential residues, Glu¹⁵⁶ and His¹⁹⁴, by virtue of lying on the side that the active site is considered to face, may play a role in constituting the active site of the enzyme. These results have been used to construct two possible putative topology models for Rce1p, with one bearing five TM spans and the other bearing seven TM spans, with both having

the N-terminus oriented towards the intraluminal compartment and the C-terminus towards the intracytoplasmic compartment.

REFERENCES

1. Chen P, Sapperstein S, Choi JD and Michaelis S: Biogenesis of the *Saccharomyces cerevisiae* mating pheromone **a**-factor. J. Cell. Biol. 136: 251-69, 1997.
2. Zhang FL and Casey PJ: Protein prenylation: molecular mechanisms and functional consequences. Ann Rev Biochem 65: 241-69, 1996.
3. Schafer WR and Rine J: Protein prenylation: genes, enzymes, targets, and functions. Annual Review of Genetics 26: 209-37, 1992.
4. Downward J: targeting Ras signaling pathways in cancer therapy. Nature 3: 11-22, 2002.
5. Schmidt WK, Tam A, Fujimura-Kamada K and Michaelis S: Endoplasmic reticulum membrane localization of Rce1p and Ste24p, yeast proteases involved in carboxyl-terminal CAAX protein processing and amino-terminal **a**-factor cleavage. Proc Natl Acad Sci USA 95: 11175-80, 1998.
6. Bogdanov M, Zhang W, Xie J and Dowhan W: Transmembrane protein topology mapping by the substituted cysteine accessibility method (SCAM(TM)): application to lipid-specific membrane protein topogenesis. Methods 36: 148-71, 2005.
7. Plummer LJ, Hildebrandt ER, Porter SB, Rogers VA, McCracken J and Schmidt WK: Mutational analysis of the ras converting enzyme reveals a requirement for glutamate and histidine residues. J Biol Chem 281: 4596-605, 2006.
8. Elble R: A simple and efficient procedure for transformation of yeasts. BioTechniques 13: 18-20, 1992.

9. Dolence JM, Steward LE, Dolence EK, Wong DH and Poulter CD: Studies with recombinant *Saccharomyces cerevisiae* CaaX prenyl protease Rce1p. *Biochemistry* 39: 4096-104, 2000.
10. Lu J and Deutsch C: Pegylation: a method for assessing topological accessibilities in Kv1.3. *Biochemistry* 40: 13288-301, 2001.
11. Trueblood CE, Boyartchuk VL, Picologlou EA, Rozema D, Poulter CD and Rine J: The CaaX proteases, Afc1p and Rce1p, have overlapping but distinct substrate specificities. *Mol. Cell. Biol.* 20: 4381-92, 2000.
12. Tam A, Nouvet F, Fujimura-Kamada K, Slunt H, Sisodia SS and Michaelis S: Dual roles for Ste24p in yeast **a**-factor maturation: NH₂-terminal proteolysis and COOH-terminal CAAX processing. *J. Cell Biol.* 142: 635-49, 1998.
13. Cadiñanos J, Varela I, Mandel D, Schmidt WK, Díaz-Perales A, López-Otín C and JMP F: AtFACE-2, a prenylated-protein protease from *Arabidopsis thaliana* related to Ras converting enzymes. *Journal of Biological Chemistry* 278: 42091-7, 2003.
14. Michaelis S and Herskowitz I: The **a**-factor pheromone of *Saccharomyces cerevisiae* is essential for mating. *Mol. Cell. Biol.* 8: 1309-18, 1988.
15. Sikorski RS and Hieter P: A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* 122: 19-27, 1989.

Table 2.1. Topology Models

Algorithm	1	2	3	4	5	6	7	8	9
TMHMM 2.0	4-23	44-61	76-93	105-127			208-230	237-259	279-298
ConPred II	4-24	43-63	72-92	105-125			208-228	238-258	279-299
SCAMPI- seq	2-22	44-64		106-126			206-226	234-254	282-302
OCTOPUS	3-23	42-62		101-121	149-169	178-198	209-229	235-255	280-300
PRODIV	5-25	46-66				172-192	213-233	238-258	279-299

Table 2.2. Yeast Strains

Strains	Genotype	Reference
SM3614	<i>MATa trp1 leu2 ura3 his4 can1 ste24::LEU2 rce1::TRP1</i>	(12)
yWS 122	<i>MATa trp1 leu2 ura3 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc]</i>	(13)
SM1068	<i>MATa lys1</i>	ATCC#204279
yWS164	<i>MATa trp1 leu2 ura3 his4 mfa1Δ mfa2Δ rce1::TRP1 ste24::KAN^R</i>	(13)
yWS198	<i>MATa trp1 leu2 ura3 his4 can1</i>	(14)
yWS787	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less]</i>	This study
yWS788	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (S5C)]</i>	This study
yWS789	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (S30C)]</i>	This study
yWS790	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (S71C)]</i>	This study
yWS977	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (S96C)]</i>	This study
yWS791	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (S145C)]</i>	This study
yWS792	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (T155C)]</i>	This study
yWS978	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (S175C)]</i>	This study
yWS979	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (S187C)]</i>	This study
yWS793	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (S206C)]</i>	This study
yWS794	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (F234C)]</i>	This study
yWS980	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (A249C)]</i>	This study
yWS986	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (S260C)]</i>	This study
yWS1113	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (T314C)]</i>	This study
yWS1114	<i>MATa trp1 leu2 his4 can1 ste24::LEU2 rce1::TRP1 [CEN URA3 RCE1::HAc cys-less (C354)]</i>	This study
yWS1020	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc cys-less]</i>	This study
yWS1021	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc cys-less (S5C)]</i>	This study
yWS1022	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc</i>	This study

yWS1023	<i>cys-less (S30C)]</i> <i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc</i> <i>cys-less (S71C)]</i>	This study
yWS1028	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc</i> <i>cys-less (S96C)]</i>	This study
yWS1024	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc</i> <i>cys-less (S145C)]</i>	This study
yWS1025	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc</i> <i>cys-less (T155C)]</i>	This study
yWS1029	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc</i> <i>cys-less (S175C)]</i>	This study
yWS1030	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc</i> <i>cys-less (S187C)]</i>	This study
yWS1026	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc</i> <i>cys-less (S206C)]</i>	This study
yWS1031	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc</i> <i>cys-less (A249C)]</i>	This study
yWS1032	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc</i> <i>cys-less (S260C)]</i>	This study
yWS1019	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc</i> <i>cys-less (C354)]</i>	This study
yWS1117	<i>MATa trp1 leu2 his4 can1 rce1::TRP1 [2μ URA3 RCE1::HAc</i> <i>cys-less (T314C)]</i>	This study

Table 2.3. Plasmids

Plasmid	Genotype	Reference
pRS426	<i>2μ URA3</i>	(15)
pSM1314	<i>CEN URA3 RCE1::HAc</i>	(5)
pWS787	<i>CEN URA3 RCE1::HAc cys-less</i>	This study
pWS788	<i>CEN URA3 RCE1::HAc cys-less (S5C)</i>	This study
pWS789	<i>CEN URA3 RCE1::HAc cys-less (S30C)</i>	This study
pWS790	<i>CEN URA3 RCE1::HAc cys-less (S71C)</i>	This study
pWS864	<i>CEN URA3 RCE1::HAc cys-less (S96C)</i>	This study
pWS791	<i>CEN URA3 RCE1::HAc cys-less (S145C)</i>	This study
pWS792	<i>CEN URA3 RCE1::HAc cys-less (T155C)</i>	This study
pWS865	<i>CEN URA3 RCE1::HAc cys-less (S175C)</i>	This study
pWS866	<i>CEN URA3 RCE1::HAc cys-less (S187C)</i>	This study
pWS793	<i>CEN URA3 RCE1::HAc cys-less (S206C)</i>	This study
pWS794	<i>CEN URA3 RCE1::HAc cys-less (F234C)</i>	This study
pWS867	<i>CEN URA3 RCE1::HAc cys-less (A249C)</i>	This study
pWS868	<i>CEN URA3 RCE1::HAc cys-less (S260C)</i>	This study
pWS911	<i>CEN URA3 RCE1::HAc cys-less (T314C)</i>	This study
pWS781	<i>CEN URA3 RCE1::HAc cys-less (C354)</i>	This study
pWS823	<i>2μ URA3 RCE1::HAc cys-less</i>	This study
pWS824	<i>2μ URA3 RCE1::HAc cys-less (S5C)</i>	This study
pWS825	<i>2μ URA3 RCE1::HAc cys-less (S30C)</i>	This study
pWS826	<i>2μ URA3 RCE1::HAc cys-less (S71C)</i>	This study
pWS875	<i>2μ URA3 RCE1::HAc cys-less (S96C)</i>	This study
pWS827	<i>2μ URA3 RCE1::HAc cys-less (S145C)</i>	This study
pWS828	<i>2μ URA3 RCE1::HAc cys-less (T155C)</i>	This study
pWS876	<i>2μ URA3 RCE1::HAc cys-less (S175C)</i>	This study
pWS877	<i>2μ URA3 RCE1::HAc cys-less (S187C)</i>	This study
pWS829	<i>2μ URA3 RCE1::HAc cys-less (S206C)</i>	This study
pWS878	<i>2μ URA3 RCE1::HAc cys-less (A249C)</i>	This study
pWS879	<i>2μ URA3 RCE1::HAc cys-less (S260C)</i>	This study
pWS933	<i>2μ URA3 RCE1::HAc cys-less (T314C)</i>	This study
pWS822	<i>2μ URA3 RCE1::HAc cys-less (C354)</i>	This study

Fig. 2.1. *Maleimide modification of cysteine residue.*

Reaction showing covalent binding of Maleimide with thiol group on Cys residues. The PEG group linked to the Maleimide, represented by 'R', is thus attached to the protein via the Cys residue.

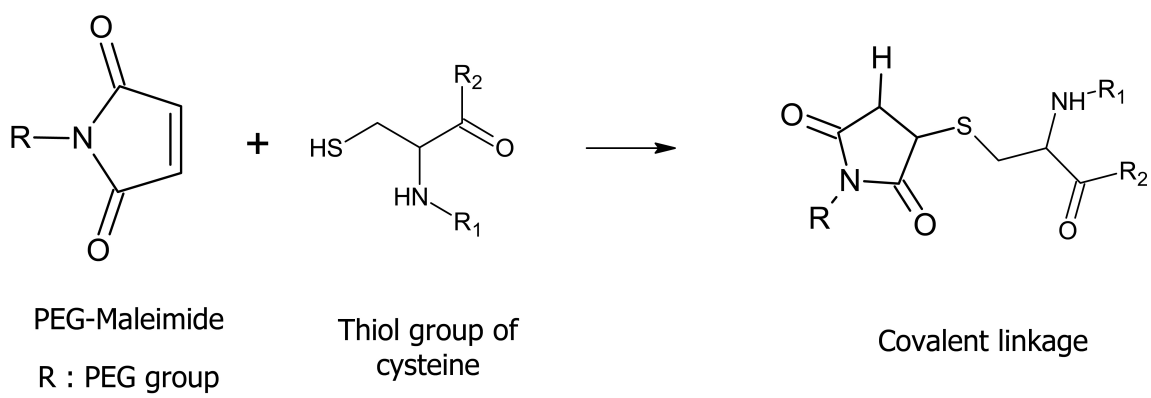
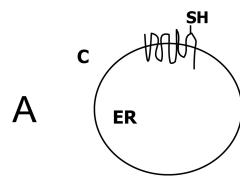
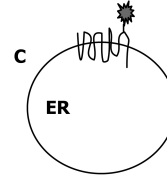


Fig. 2.2. *PEG-Maleimide treatment of yeast microsomes.*

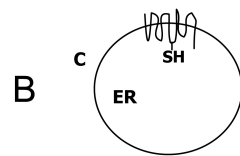
Yeast microsomes carrying the Rce1p protein with one Cys residue are treated with PEG-Maleimide (PEGMAL) in the absence of detergent (Fig. 2.3A and Fig. 2.3B) and in the presence of detergent, Triton X-100 (Fig. 2.3C). Only those cysteine mutants with the Cys residue on the cytosolic side of the membrane are modified by PEG-Maleimide, leading to a 5 KDa increase in apparent molecular mass on Western detection (Fig. 2.3A), while those mutants of the protein with Cys residues on the luminal side are not, showing no change in molecular mass of treated protein on Western (Fig. 2.3B). However the same cysteine mutants used in Fig. 2.3B, when treated with PEG-Maleimide in presence of Triton X-100, show modification of Cys residue, as noted by an apparent increase in molecular mass on Western blots (Fig. 2.3C). ‘C’ indicates cytosol, while ‘ER’ indicates the endoplasmic reticulum microsomal intraluminal compartment.



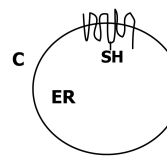
Add PEG-Maleimide
→



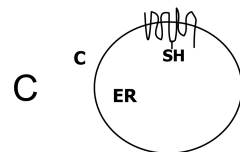
PEGMAL	+	-
40KD	■	
35KD		■



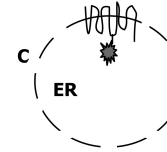
Add PEG-Maleimide
→



PEGMAL	+	-
40KD		
35KD	■	■



Add PEG-Maleimide
+ Triton X-100
→



PEGMAL	+	-
40KD	■	
35KD		■

Fig. 2.3. *Topology model for Rce1p::HA as predicted by TMHMM 2.0.*

Cartoon representation of membrane topology of Rce1p::HA drawn based on TM segments predicted by TMHMM 2.0 program. Amino acid residues indicated by hexagons show triple HA tag on protein. Naturally occurring Cys residues are indicated by bold 'C'. Residues marked by bold circles and with bold lettering represent sites where Cys residues were inserted separately by site directed mutagenesis. Residues marked by filled circles represent the 3 amino acid residues essential for enzyme activity (7).

Fig. 2.4. *Functionality of Rce1p cysteine mutants.*

Dilution mating test performed between *MATa* cells carrying various Rce1p cysteine mutants and *MATα* cells. There is a 10-fold dilution of *MATa* cells loaded from uppermost to lowermost row. Panel A shows SD minimal media plates with diploid colonies formed as a result of mating. Panel B shows SC-Lysine plates showing *MATa* cells loading control. Only Rce1p F234C mutant shows no mating, while S260C and S187C show reduced mating. All other mutants show satisfactory mating, with WT Rce1p showing highest mating efficiency.

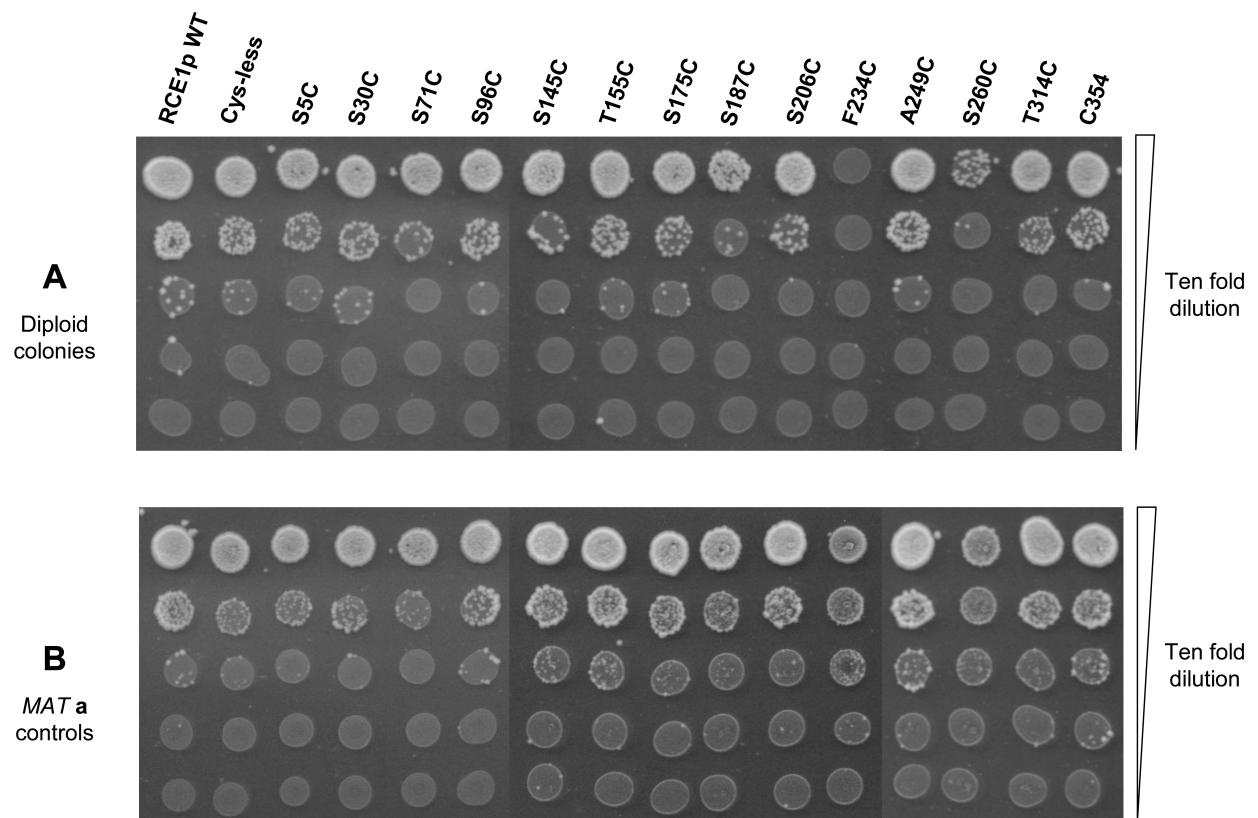


Fig. 2.5. *Optimization of mannitol concentration in lysis buffer using Kar2p control.*

Western blot showing the effect of increasing mannitol concentration in Lysis buffer on yeast microsome integrity. Yeast microsomes prepared using Lysis buffer containing increasing concentrations of Mannitol were treated with PEG-Maleimide in the absence of 0.25% Triton X-100 (left panel) and in the presence of 0.25% Triton X-100 (right panel). Following quenching with β -mercaptoethanol, samples were run on SDS-PAGE and Western blot performed for Kar2p. Only microsomes prepared with 0.8 M Mannitol Lysis buffer shows no modification of Kar2p by PEG-Maleimide in the absence of detergent, indicating no leakage from microsomes.

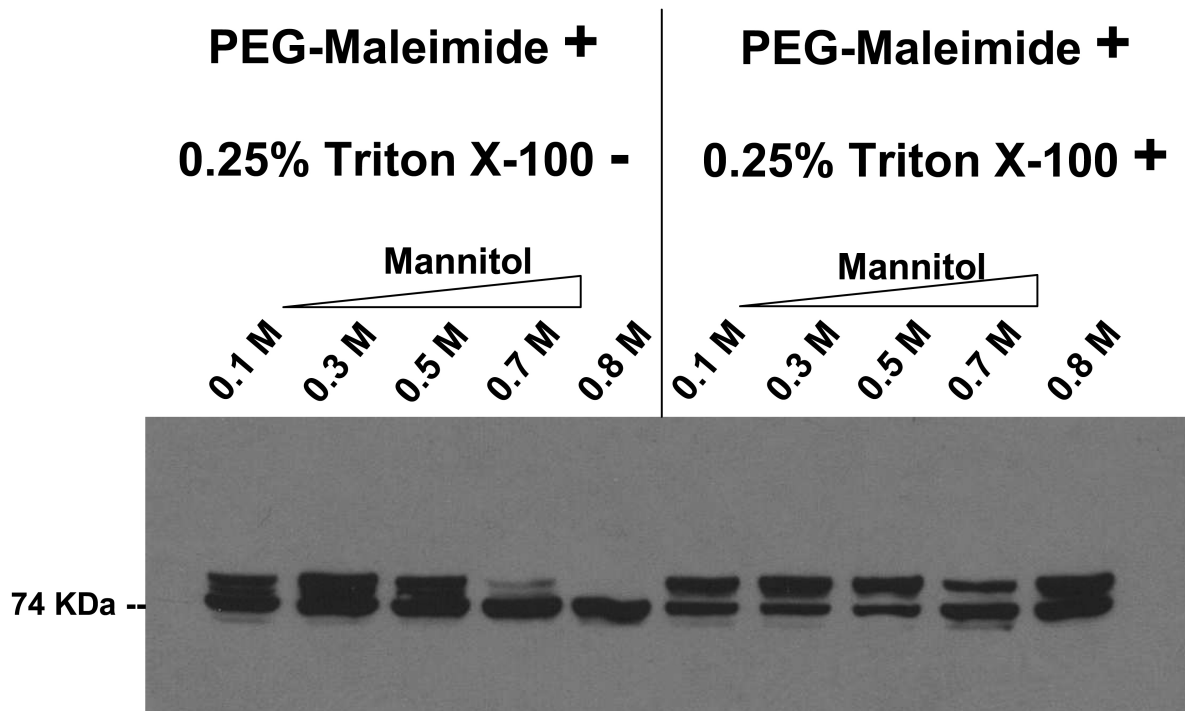


Fig. 2.6. *Orientation of cysteine residues on Rce1p cysteine mutants with respect to membrane.*

Western blot depicting the PEG-Maleimide modification of the various Rce1p cysteine mutants in the absence of detergent Triton X-100. ER microsome preparations were treated with 1 mM PEG-Maleimide in lysis buffer and incubated for one hour on ice, following which the reaction was quenched with excess β -mercaptoethanol. Samples were run on SDS-PAGE gels following addition of 2X sample buffer, and Western detection performed with anti-HA antibodies for detecting Rce1p (35KDa) and anti-Kar2p antibodies for detecting Kar2p (70KDa). PEG-Maleimide modified cysteine mutants show an increase in molecular mass. Absence of modification of Kar2p indicates intact microsomes without leakage. Cysteine mutants S5C and A249C show no PEG-Maleimide modification, indicating that positions 5 and 249 were inaccessible to PEG-Maleimide. All other mutated cysteine sites were accessible to membrane impermeable PEG-Maleimide. The Kar2p control protein is also modified by PEG-Maleimide in the lanes 1-9, implying leakage from microsomes in those samples.

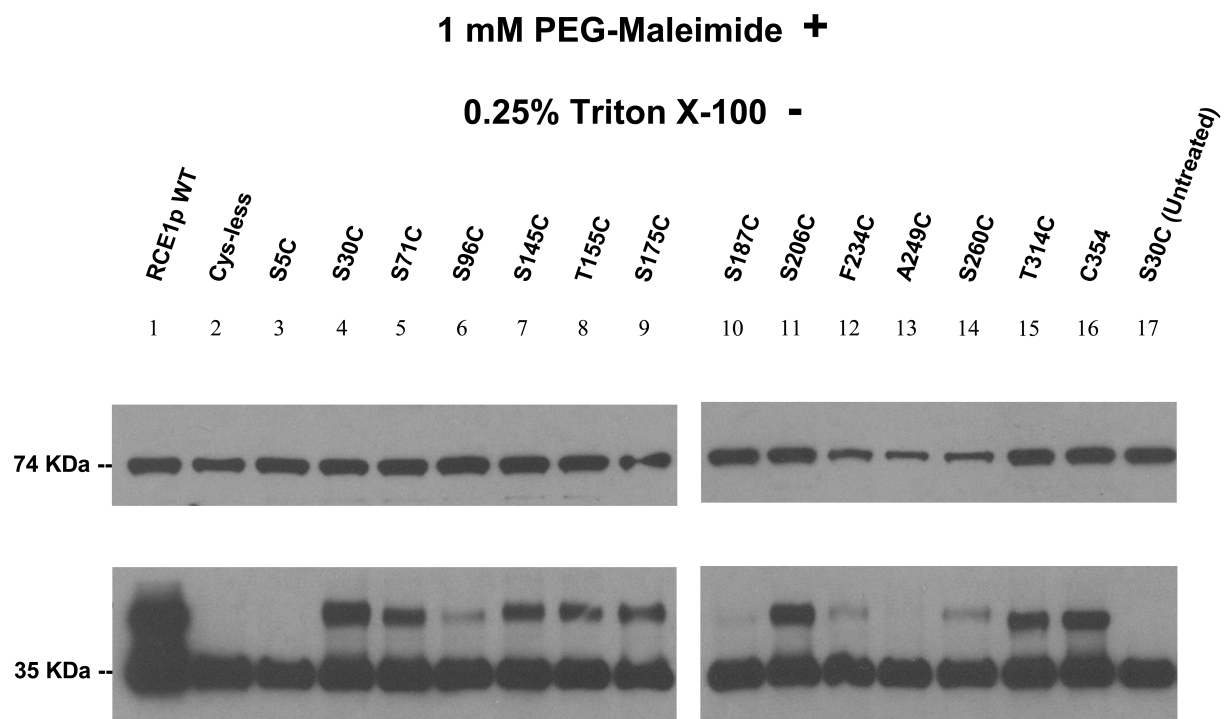


Fig. 2.7. *Accessibility of Rce1p cysteine mutation sites in the presence of detergent.*

Western blot depicting the PEG-Maleimide modification of the various Rce1p cysteine mutants in the presence of detergent Triton X-100. The experimental steps were similar to the ones used in the previous experiment, except that detergent was added to all samples. ER microsomal preparations were treated with 1 mM PEG-Maleimide and 0.25% Triton X-100 in lysis buffer and incubated for one hour on ice, following which the reaction was quenched with excess β -mercaptoethanol. Samples were run on SDS-PAGE gels following addition of 2X sample buffer, and Western detection performed with anti-HA and anti Kar2p antibodies as in previous experiment. All cysteine mutants showed PEG-Maleimide modification in the presence of detergent. PEG-Maleimide modification of Kar2p modification by PEG-Maleimide is observed only in lanes 1-12 and 14. Lanes 13, 15 and 16 show no modification by PEG-Maleimide.

1 mM PEG-Maleimide +

0.25% Triton X-100 +

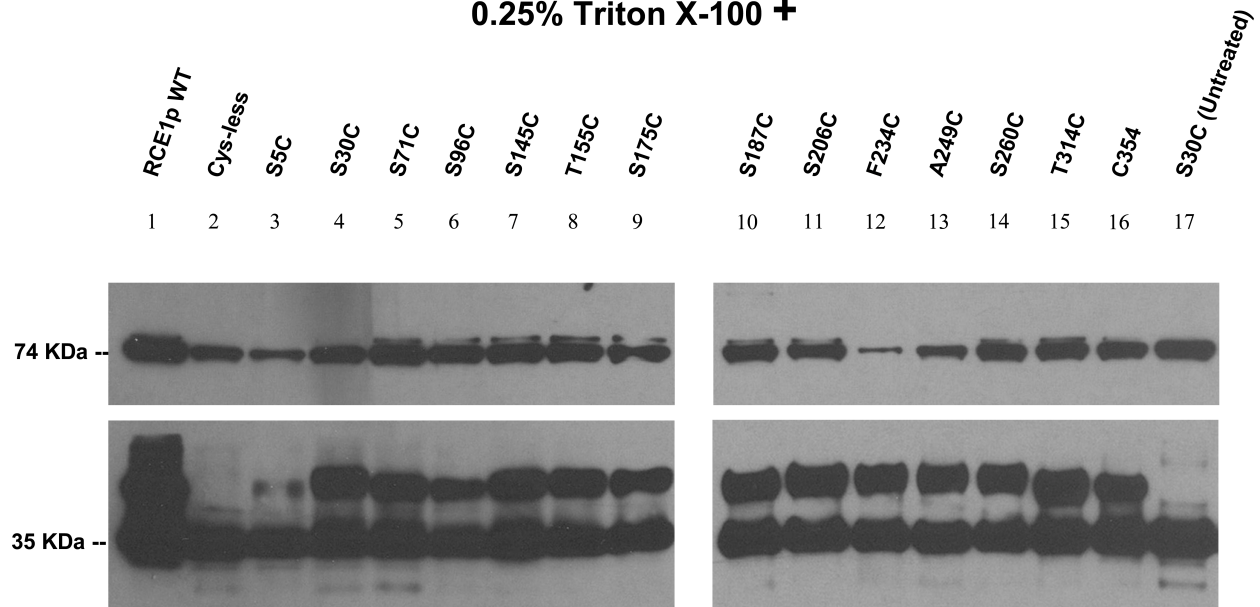
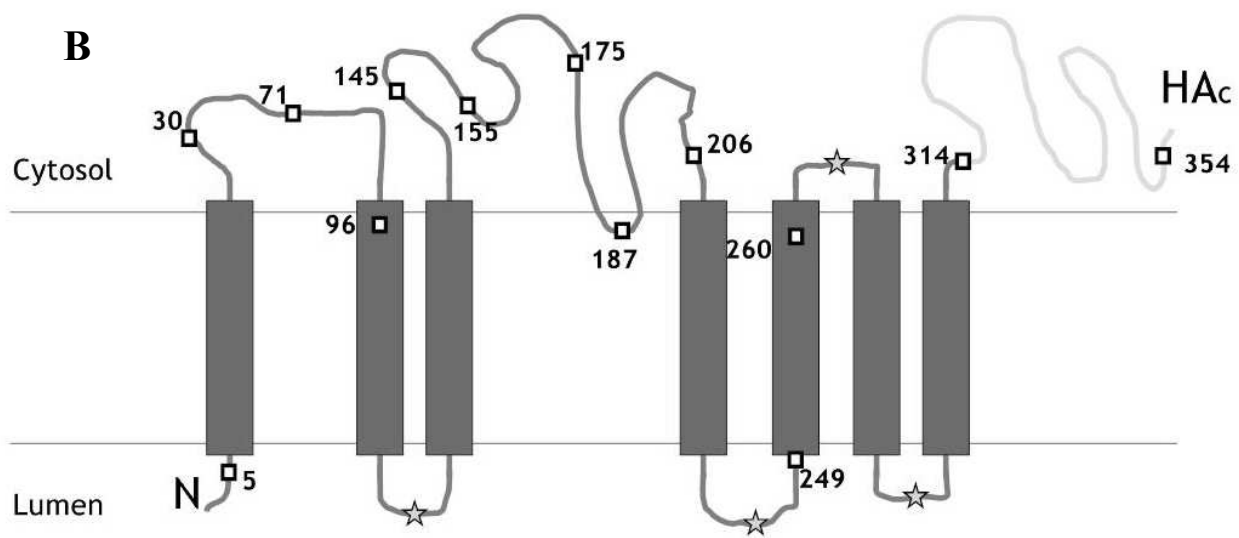
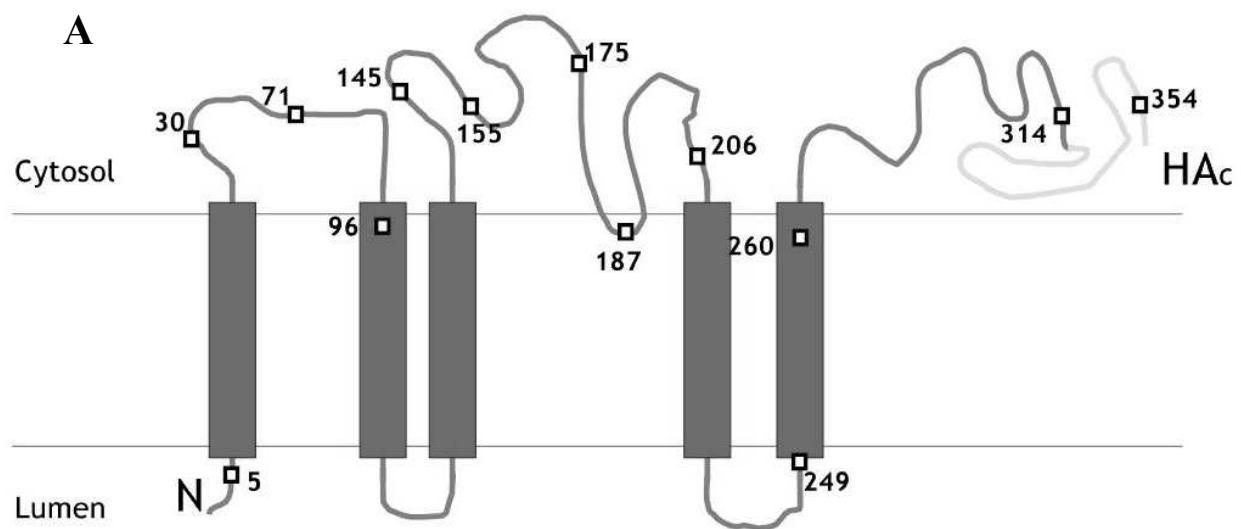


Fig. 2.8. *Putative topology models for Rce1p as constructed from SCAM results.*

Two topology models can be constructed for Rce1p on the basis of SCAM results. Figure 2.8A shows a model with five TM spans, and Figure 2.8B shows an alternate model with seven TM spans. The positions of the Cys residues in each Rce1p cysteine mutant, as determined by SCAM, have been shown in both models by empty squares. The sites at which Cys residues need to be introduced, in order to create newer Rce1p cysteine mutants for use in SCAM technique, for the purpose of refining these putative topology models, have been shown as grey stars in Figure 2.8B. The N and C-termini have been shown in both models, with the HA tag located at the C-terminus delineated in a lighter color.



CHAPTER 3

CONCLUSION AND DISCUSSION

In this study, we have used the substituted cysteine scanning method (SCAM) to investigate the membrane topology of the Ras converting enzyme (Rce1p). It is one of the few known proteases that is also an integral membrane protein having multiple membrane spans. In addition, the Rce1p enzyme can be considered as a model towards investigating other proteases that are also integral membrane proteins, such as Presenilin, Ste24p, SPP, and S2P (1, 2). Not knowing the topology or the three dimensional structure of Rce1p has been an impediment towards a comprehensive study of the catalytic activity of the enzyme. Biochemical analysis of the enzyme is needed in part due to the important role Rce1p plays in Ras protein maturation, which in turn is involved in cancer. For example, determination of the catalytic activity of Rce1p enzyme will help in discovering newer approaches to inhibit the enzyme. This, as mentioned previously will serve an important role in developing newer anticancer pharmacotherapy.

The SCAM technique has been used in this study of the membrane topology of Rce1p. It is an alternative to the use of reporters in the study of membrane topology of enzymes. It bypasses the disadvantages of traditional reporters that often alter the native structure of the protein being studied. SCAM works on the principle of labeling a cysteine thiol group on the protein with maleimide, which forms a covalent bond with the thiol group. The maleimide is in turn linked to another moiety, such as a PEG group that alters mobility of the modified protein in such a way that it can easily be detected by Western blot. The PEG-Maleimide molecule is also

impermeable across the membrane, and hence will label only those cysteines that are oriented to the same side of the membrane to which PEG-Maleimide molecules have access to. Therefore, the introduction of cysteine residues into the primary sequence of the protein at salient positions extending from the N-terminus to the C-terminus allows for a determination of the orientation of all these positions, and thus yielding a topology model for the entire protein (3).

The substituted cysteine accessibility method has been used so far to construct the membrane topology of many proteins, such as that of the yeast vacuolar ATPase subunit, Vph1p (4), and that of the voltage-sensing domain of the potassium channel KvAP in *E. coli* (5). However there have not yet been any reports published that use this technique to study membrane topology of yeast ER-localized membrane proteins. Furthermore there have not been any similar studies which have attempted to use a control to demonstrate integrity of membranes. This study has aimed to be the first to show the use of Kar2p as a control to demonstrate intact membranes while applying SCAM to study the membrane topology of an ER-localized integral membrane protein. Thus using the SCAM technique we have been able to construct two putative topology models for the Rce1p enzyme (Figure 2.8). However, in order to refine the accuracy of the model, further studies need to be conducted. These include the construction of newer Rce1p cysteine mutants that have Cys residues incorporated at those sites that have not been covered by this study. Use of the SCAM technique on those mutants will enable us to identify the orientation of those positions in the membrane topology of the protein. The new positions on the primary structure where Cys residues need to be introduced can be chosen following a computational analysis of the current SCAM model, and thus select more accurately sites that are likely to lie within an intracytoplasmic or intraluminal loop on the protein. Another useful addition to the technique would be the use of other maleimide compounds such as N-ethylmaleimide and biotin

maleimide, with different membrane permeability properties that can be used to differentially localize the various cysteine residues introduced into the protein. The use of NEM has been detailed in the previous chapter. Biotin maleimide on the other hand, is similar to PEG-Maleimide in terms of membrane permeability and does not cross the membrane. However, being smaller in size as compared to PEG-Maleimide, biotin maleimide has an extended reach compared to the bigger PEG group, and hence will label even those Cys residues that are only weakly modified by PEG-Maleimide, thus validating existing PEG-Maleimide hits. The labeling with biotin can also be detected using horse-radish peroxidase linked avidin on Western blot instead of using antibodies against the protein.

The significance of this SCAM study on Rce1p membrane topology lies in its applications in the discovery of newer anti-cancer drugs. FTase inhibitors have not been shown to be conclusively effective, either due to intolerable toxicity or due to bypassing of the FTase enzyme in the maturation of certain Ras proteins like K-Ras and N-Ras. Therefore, the idea of Rce1p inhibitors in the treatment of cancer has gained traction recently. However, while investigating the application of Rce1p inhibitors in the treatment of cancer, it is also essential to study the effect of Rce1p inhibition in mammalian tissues, and to identify, if any, deleterious effects of this inhibition of Rce1p. Rce1p inhibitors have been shown to partially mislocalize the Ras protein from the plasma membrane to intracellular membranes (6), and have also shown to reduce the tumorigenic capacity of mutated Ras proteins in cells (7). This is in contrast to the observation that Rce1p knockout mice die on embryonic day 15, although no specific cause of death has been identified (6). Use of the Cre-LoxP system has enabled the study of Rce1p defects in post-natal mice, and in specific tissues. This has shown that *RCE1* disruption can lead to severe heart defects such as dilatory cardiomyopathy, causing mice to die within 3 to 5 months

(8). No such deleterious effect has however been observed in other tissues. Another study that aimed to create an *in vivo* model of cancer treated with Rce1p inhibition, studied the effect of Rce1p inactivation in mice where the oncogenic allele of K-Ras was activated (9). This showed that Rce1p deficiency in fact worsened the development of malignant disease caused by the activated K-Ras. The myeloproliferative disease thus caused, resulted in the mice dying more rapidly. Therefore, although Rce1p inhibition does hold promise in the treatment of cancer, extensive studies regarding the effects of its inhibition in mammalian tissues, both normal and cancerous, need to be conducted. The absence so far of any structural information regarding the enzyme has hampered the investigation of Rce1p inhibition as a therapeutic target. This study aims to bridge this information gap by determining the membrane topology of the enzyme, using a less invasive and thus more accurate approach such as SCAM. The membrane topology of Rce1p will serve as a welcome first step towards determining the catalytic mechanism of the enzyme, and thus lead to better approaches to inhibit the enzyme.

REFERENCES

1. Plummer LJ, Hildebrandt ER, Porter SB, Rogers VA, McCracken J and Schmidt WK: Mutational analysis of the ras converting enzyme reveals a requirement for glutamate and histidine residues. J Biol Chem 281: 4596-605, 2006.
2. Tam A, Schmidt WK and Michaelis S: The multispanning membrane protein Ste24p catalyzes CAAX proteolysis and NH₂-terminal processing of the yeast a-factor precursor. J Biol Chem 276: 46798-806, 2001.

3. Bogdanov M, Zhang W, Xie J and Dowhan W: Transmembrane protein topology mapping by the substituted cysteine accessibility method (SCAM(TM)): application to lipid-specific membrane protein topogenesis. *Methods* 36: 148-71, 2005.
4. Wang Y, Toei M and Forgac M: Analysis of the membrane topology of transmembrane segments in the C-terminal hydrophobic domain of the yeast vacuolar ATPase subunit a (Vph1p) by chemical modification. *J Biol Chem* 283: 20696-702, 2008.
5. Neale EJ, Rong H, Cockcroft CJ and Sivaprasadarao A: Mapping the membrane-aqueous border for the voltage-sensing domain of a potassium channel. *J Biol Chem* 282: 37597-604, 2007.
6. Kim E, Ambroziak P, Otto JC, Taylor B, Ashby M, Shannon K, Casey PJ and Young SG: Disruption of the mouse Rce1 gene results in defective Ras processing and mislocalization of Ras within cells. *J. Biol. Chem.* 274: 8383-8390, 1999.
7. Bergo MO, Ambroziak P, Gregory C, George A, Otto JC, Kim E, Nagase H, Casey PJ, Balmain A and Young SG: Absence of the CAAX endoprotease Rce1: effects on cell growth and transformation. *Mol Cell Biol* 22: 171-81, 2002.
8. Bergo MO, Lieu HD, Gavino BJ, Ambroziak P, Otto JC, Casey PJ, Walker QM and Young SG: On the physiological importance of endoproteolysis of CAAX proteins: heart-specific RCE1 knockout mice develop a lethal cardiomyopathy. *J Biol Chem* 279: 4729-36, 2004.
9. Wahlstrom AM, Cutts BA, Karlsson C, Andersson KM, Liu M, Sjogren AK, Swolin B, Young SG and Bergo MO: Rce1 deficiency accelerates the development of K-RAS-induced myeloproliferative disease. *Blood* 109: 763-8, 2007.

APPENDIX

The following is a manuscript co-authored by me, and submitted to *Yeast* in November 2008. My contributions to the manuscript are listed below:

1. All experiments that led to construction of Figure 4.2.
2. All experiments that led to construction of Figure 4.3.
3. Editing of the manuscript
4. Creation of plasmid pWS912

**PROTEOLYTIC PROCESSING OF CERTAIN CAAX MOTIFS CAN OCCUR IN THE
ABSENCE OF THE RCE1P AND STE24P CAAX PROTEASES¹**

¹ Ranjith K. Krishnankutty*, Sayali S. Kukday*, Amanda J. Castleberry, Sarah R. Breevoort, and
Walter K. Schmidt

Submitted to *Yeast*, November 2008

* These individuals contributed equally to this work

ABSTRACT

The CaaX motif directs C-terminal protein modifications that include isoprenylation, proteolysis, and carboxylmethylation. Proteolysis is generally believed to require either Rce1p or Ste24p. While investigating the substrate specificity of these proteases using the yeast *a* factor mating pheromone as a reporter, we observed Rce1p and Ste24p-independent mating (RSM) when the CKQQ CaaX motif was used in lieu of the natural *a*-factor CVIA motif. Uncharged or negatively charged amino acid substitutions at the $\alpha 1$ position of the CKQQ motif prevented RSM. Alanine substitutions at the $\alpha 2$ and X positions enhanced RSM. Random mutagenesis of the CaaX motif provided evidence that RSM occurs with approximately 1% of all possible CaaX motif permutations. Combined mutational and genetic data indicates that RSM-promoting motifs have a positively charged amino acid at the $\alpha 1$ position. Two of nine naturally occurring yeast CaaX motifs conforming to this pattern promoted RSM. The activity of the isoprenylcysteine carboxyl methyltransferase Ste14p was required for RSM, indicating that RSM-promoting CaaX motifs are indeed proteolyzed. RSM was enhanced by the over-expression of Axl1p or Ste23p, suggesting a role for these M16A subfamily metalloproteases in this process. We have also determined that an N-terminal extension of the *a*-factor precursor, which is typically removed by the yeast M16A enzymes, is required for optimal RSM. These observations suggest a model that involves targeting of the *a*-factor precursor to the peptidosome cavity of M16A enzymes where subsequent interactions between RSM-promoting CaaX motifs and the active site of the M16A enzyme lead to proteolytic cleavage.

INTRODUCTION

The CaaX motif is a C-terminal tetrapeptide sequence generally described as having an invariant cysteine (C), two aliphatic amino acids (a_1 and a_2), and one of several amino acids in the terminal position (X). Eukaryotic proteins having a CaaX motif (CaaX proteins) typically undergo three ordered post-translational modifications (reviewed in (1, 2)) (Figure 3.1). The first is isoprenylation of the cysteine by either the C15 farnesyl transferase (FTase) or the C20 geranylgeranyl transferase I (GGTase I). The context of the CaaX motif can dictate which isoprenoid is attached, with geranylgeranylated proteins often having Leu and Phe at the X position. Isoprenylation is followed by an endoproteolytic cleavage that removes the last three amino acids of the motif (*i.e.*, aaX). Two proteases, Rce1p and Ste24p, have been identified that can perform CaaX proteolysis (3, 4). CaaX proteolysis is followed by carboxylmethylation of the farnesylated cysteine by an isoprenylcysteine carboxyl methyltransferase (ICMT). Collectively, these modifications modulate the activity, membrane partitioning, subcellular localization, stability, and/or protein-protein interaction properties of the modified protein (3, 5-11).

CaaX proteins have diverse, biologically important functions. Pertinent examples of CaaX proteins include signaling molecules (*i.e.*, Ras and RhoB), nuclear proteins (*i.e.*, CENP-E, CENP-F, and nuclear lamins), Hsp40 chaperones (*i.e.*, Ydj1p and Dnj3), and fungal mating pheromones (*e.g.*, *Saccharomyces cerevisiae* α -factor). Because of the prominence of CaaX proteins in association with disease (*e.g.* Ras and cancer), it is generally hypothesized that interfering with CaaX modifications could be incorporated into disease intervention strategies. This hypothesis has led to the development of FTase inhibitors (FTIs) that are currently being investigated for the treatment of cancer and progeroid syndromes (12, 13). Inhibitors of the CaaX proteases and ICMT hold similar therapeutic potential and are being investigated (14-17). A

problematic issue in this research area is the ability of CaaX proteins to be processed by partially redundant activities. For example, several proteins are known to be isoprenylated by GGTase I in the presence of FTIs. Likewise, it is possible that targeted inhibition of Rce1p can lead to alternative processing by Ste24p, and *vice versa*. This issue is less of a concern for targeted inhibition of the ICMT because there appears to be no alternative enzyme that can perform the carboxyl methylation of CaaX proteins.

The two CaaX proteases are both ER-localized membrane proteins, but are otherwise unrelated by primary sequence (18). Ste24p is a zinc-dependent metalloprotease that has been purified and demonstrated to possess *in vitro* CaaX proteolytic activity (4). The mechanism of Rce1p remains undefined. Several lines of evidence support the function of Rce1p as a CaaX protease, including genetic and over-expression studies (3, 4, 10, 19). Bioinformatic and inhibitor profiles suggest that it is a metalloprotease (15, 20).

Rce1p and Ste24p have partially overlapping substrate specificity, meaning that each enzyme has specific substrates and also shared ones. For example, Rce1p specifically modifies Ras GTPases, Ste24p specifically modifies prelamin A, and both enzymes modify the yeast **a**-factor precursor (3, 21, 22). Yeast **a**-factor has been a convenient reporter for investigating CaaX modifications because defects in any of the three post-translational events results in a sterile mating phenotype and because it can be used to readily monitor either Rce1p or Ste24p activity. The yeast system is also useful for the evaluation of CaaX proteases from other eukaryotic species because they all have the ability to recognize yeast **a**-factor precursor as a substrate (23-26).

While investigating the substrate specificities of the yeast CaaX proteases using **a**-factor as a reporter, we observed the ability of certain CaaX motifs to promote yeast mating in the

absence of Rce1p and Ste24p. This study compares Rce1p and Ste24p-independent mating (RSM) with mating promoted by the established CaaX proteases and provides evidence that a substantial number of CaaX motifs, including naturally occurring yeast motifs, can promote RSM. Moreover, we provide evidence that the yeast M16A metalloproteases Axl1p and Ste23p, which normally cleave an N-terminal extension found on the **a**-factor precursor, can enhance RSM, suggesting that these enzymes are responsible for this activity.

MATERIALS AND METHODS

Yeast strains – The yeast strains used in this study are listed in Table 4.1. yWS829 was created by disrupting the *STE14* gene in yWS164 using the BamHI-ClaI fragment from pSM284 (7). The disruption was confirmed by PCR methods. Yeast strains were routinely grown at 30 °C on rich media (YEPD) or appropriate synthetic dropout media (SC-) when propagating plasmid-transformed strains (27). Yeast DNA transformations were carried out according to published methods (28).

Yeast plasmids – The yeast plasmids used in this study are listed in Table 4.2. pWS610 and pWS612 were constructed by subcloning the appropriate NotI-XhoI fragment encoding **a**-factor from pSM1605 and pWS196, respectively, into pRS415. pWS817 was created similarly but with pWS654 and pRS315. pWS196 and all other **a**-factor encoding plasmids bearing altered CaaX motifs were created by PCR-directed plasmid-based recombination, which involves repair of a plasmid having a double stranded break by a PCR product having homology to the plasmid on both sides of the break point (29). The parent plasmid (*i.e.*, pSM1605, pWS438 or pWS610) was treated with MluI, which cuts in the 3' untranslated region (UTR) of the *MFAl* gene very near the 3' end of the open reading frame (ORF). In most instances, plasmids were additionally treated

with SphI, which also cuts in the 3' UTR but more distal to the ORF. The digested plasmid was co-transformed into yeast with a PCR product having sequence homology to the plasmid in regions flanking the restriction site(s) to allow for gap repair. The PCR product was generated using a mutagenic forward oligo that contained 39 bases of homology to the parent plasmid, nine bases encoding the desired aaX sequence, and an 18-21 base extension for annealing to a template (*i.e.*, pSM1605, pWS438 or pWS610). The reverse oligo was complementary to the vector outside the polylinker into which the *MFAl* encoding fragment was subcloned; its use generates homology to the plasmid on the SphI side of the digested plasmid. Following co-transformation of the digested plasmid and PCR product, individual yeast colonies surviving appropriate selection (SC-Ura or SC-Leu) were screened for those containing a plasmid encoding the altered *MFAl* gene, as determined by restriction enzyme mapping and subsequent sequencing of isolated plasmids; a silent site (e.g. SphI or PstI) was typically incorporated along with the desired mutation. All plasmids derived from pWS438 were converted to low-copy plasmids by subcloning the NotI-XhoI *MFAl*-encoding fragment into pRS315 at the same sites. pWS196 was the only plasmid derived from pSM1605, and it was manipulated as described above. pWS601 and pWS602 were also created by PCR-directed plasmid-based recombination. These plasmids encode *AXL1* and *STE23*, respectively, behind the constitutive phosphoglycerate kinase (*PGK*) promoter. pSM703 was the recipient vector used in the construction of these plasmids, which was gapped within its polylinker prior to use.

Constructs encoding ubiquitin fusions were created by PCR-directed plasmid-based recombination essentially as described above for the creation of **a**-factor CaaX motif mutants. A PCR fragment encoding the CKQQ motif was derived from pWS718 and recombined into MluI linearized pSM1368 and pSM1369 to create pWS892 and pWS893, respectively. To create

pWS894, a PCR fragment also derived from pWS718 was produced that would incorporate the DNA sequence encoding mature **a**-factor upon recombination into MluI linearized pWS892.

Serial dilution mating assay – The ability of the various CaaX motifs to promote **a**-factor maturation was judged using a genetic assay that scores diploid formation resulting from the mating of haploid mating partners. The *MATa* haploid strain used (yWS164) lacks both CaaX protease-encoding genes and both **a**-factor-encoding genes (25). Mating competence was restored in this strain by co-transformation with plasmids encoding an **a**-factor species and a CaaX protease. Transformation with the latter was not necessary in the case of certain **a**-factor CaaX variants. In brief, the serial dilution mating assay involves the mixing of *MATa* and *MATα* cell suspensions on medium selective for diploid growth (26). The cultures are prepared by first growing the *MATa* yeast in selective media and the *MATα* yeast in non-selective YEPD for 24 hrs, then normalizing the cultures to a cell density of $A_{600} 1.00 \pm 0.05$ with appropriate sterile media. A portion of each normalized *MATa* culture was diluted ten-fold with a normalized *MATα* culture such that the final volume of the mating mixture was 100 μ l. This primary mixture was subjected to several additional ten-fold dilutions using normalized *MATα* cells as the diluent until a set of 5 samples was prepared. A portion of each serially diluted mixture (5 μ l) was spotted onto solid SD medium. Growth of diploid cells on SD medium was scored after 3-4 days growth at 30 °C. The results of the mating test were digitally recorded by scanning the plates using a standard flat bed scanner.

Genetic screen to identify CaaX motifs that permit RSM – A library of plasmids encoding all possible permutations of the CaaX motif appended to yeast **a**-factor was created in yWS164. The individual plasmid-bearing colonies were assessed by replica methods for the ability to produce

a-factor. Both mating and halo assays were used (30). In brief, the population of transformants was replica plated onto separate lawns of IH1793 and RC757. The lawns were prepared by scraping freshly grown strains from a YEPD plate (*i.e.*, 48 hrs growth at 30 °C), diluting the cells into liquid YEPD, adjusting the density to $A_{600} 1.00 \pm 0.05$, pouring the cell suspension onto a plate of SD (IH1793) or YEPD (RC757) (~3-5 ml/plate), immediately decanting the majority of the surface liquid, and allowing the residual liquid to absorb for 30 min at room temperature. The replica printed plates were incubated at 30 °C for 120 hours (IH1793 lawn) or 16 hrs (RC757 lawn) to allow for growth of diploids and formation of halos, respectively. Plasmids were isolated from colonies exhibiting mating competence and the ability to growth arrest RC757 cells (31). The plasmids were retransformed into yWS164, phenotypes reconfirmed, and plasmids sequenced.

The plasmid library was created by plasmid-based PCR-directed recombination. pWS654 was gapped with PstI, which cuts within the sequence encoding the aaX portion of the *MFAL* gene, and MluI, which cuts 3' of the *MFAL* ORF. The forward oligo used to generate the PCR fragment had 39 bases of homology to the *MFAL* gene (5' to PstI cut site), a nine base sequence that was randomized for every possible nucleotide combination (*i.e.*, the randomized aaX sequence), and 24 bases for annealing of the primer to the pWS438 plasmid used for target amplification; the first codon of the 24 base sequence encoded a stop codon. The reverse primer was homologous to DNA just outside the polylinker of pWS438 into which the *MFAL* gene was subcloned. This sequence is also present on pWS654. The plasmid-derived DNA fragments and the PCR-generated DNA fragments were co-transformed into yWS164 to facilitate recombination events that formed plasmids allowing for selective growth of yeast on SC-Leu medium.

RESULTS

1. **a-factor-CKQQ promotes Rce1p and Ste24p-independent mating (RSM):**

The specificities of the yeast Rce1p and Ste24p CaaX proteases can be monitored using the yeast **a-factor** mating pheromone as a reporter molecule. During such an investigation, we observed mating by a strain expressing the **a-factor-CKQQ** variant in the absence of the established CaaX proteases (Figure 4.2A). The CKQQ motif was derived from Pex19p and is also present on the mammalian Ser/Thr kinase Lkb1, a known tumor suppressor. Both proteins are known to be isoprenylated and thus substrates for CaaX proteolysis (32, 33). RSM was not observed when either wildtype **a-factor** (CVIA) or a variant known to be Ste24p-specific (CASQ) was expressed.

The **a-factor** CKQQ variant itself appears to undergo normal post-translational modification since it promotes relatively efficient mating in the presence of CaaX proteases. In particular, mating appears to be more efficient in the presence of the Ste24p by comparison to Rce1p (Figure 4.2B). Rce1p does not appear to recognize the CKQQ motif since mating observed for **a-factor-CKQQ** in the presence of Rce1p is no different than that observed in its absence. Based on these observations, we conclude that the CKQQ motif is normally recognized by Ste24p, much like the reportedly Ste24p-specific CASQ motif. Unlike the CASQ motif, however, the CKQQ motif can yield bioactive **a-factor** in the absence of both Rce1p and Ste24p.

2. **Other CaaX motifs also support RSM:**

To investigate the extent of motifs that support RSM, the a1, a2 and X positions of the CKQQ motif were independently altered to Ala. This analysis revealed that Lys at the a1 position was a critical determinant for RSM (Figure 4.3A). Alterations at the a2 and X position

did not abolish RSM. Both CKAQ and CKQA supported more efficient mating than the CKQQ motif. The possibility of a charge requirement at the a1 position was investigated in more detail by substituting various polar amino acids (Figure 4.3B). Of the motifs evaluated, CRQQ and CHQQ promoted RSM while CDQQ and CEQQ did not. The CHQQ motif supported less efficient mating than either the CKQQ or CRQQ motifs.

Given our findings, we hypothesized that the C(K/R/H)aX motif might be a good predictor of RSM substrates. To test this hypothesis, we examined additional natural yeast CaaX motifs and several synthetic sequences (*i.e.*, not occurring in yeast) corresponding to this consensus. Only a subset of these motifs promoted mating when appended to **a**-factor (Figure 4.3C and Table 4.3). These observations indicate that C(K/R/H)aX can be used to identify candidate substrates for RSM, but that this consensus sequence is not an absolute predictor of RSM substrates. Our result was somewhat expected since the consensus-matching CKIA motif has been previously identified as not promoting mating activity (34).

3. A relatively large number of CaaX motifs can support RSM:

To broadly investigate the propensity of CaaX motifs to promote mating in the absence of Rce1p and Ste24p, we set up a genetic screen to identify motifs capable of producing biologically active **a**-factor in an *rce1 ste24* null background. For the screen, a degenerate PCR oligonucleotide was used to create a population of plasmids encoding yeast **a**-factor with randomly appended aaX sequences. Theoretically, 8000 aaX permutations were possible. The plasmid library was created in yeast through recombination-mediated methods. Evaluation of over 3000 yeast colonies by replica-based mating tests revealed a substantial number having the ability to mate and induce growth arrest of MAT $\square$ *sst2-1* yeast. RSM was observed at a rate of

0.93% $\pm$ 0.61% suggesting that approximately 75 CaaX motifs can promote RSM. Six plasmids capable of promoting RSM were recovered and sequenced. This analysis revealed sequences having either Lys or Arg at the a1 position but no consistent pattern at the a2 and X positions (Table 4.3).

4. RSM is dependent on Ste14p:

Two hypotheses were developed to explain our observations for RSM. The most straightforward was that a third CaaX proteolytic activity is responsible for RSM (Figure 4.4A). Alternatively, it was possible that certain RSM-promoting motifs were uncleaved, and that the uncleaved motifs somehow mimicked the biophysical properties of a carboxymethylated C-terminus such that cellular export and receptor binding by the pheromone were now possible. To distinguish between these possibilities, we predicted that a proteolytic-dependent mechanism would require the isoprenylcysteine carboxyl methyltransferase (ICMT) for activation of the biological activity of **a**-factor, whereas a carboxymethyl mimic would not. We thus evaluated the dependence of RSM on the Ste14p ICMT. Using **a**-factor-CKAQ as a reporter, we observed that RSM was indeed dependent on Ste14p (Figure 4.4B). This observation strongly implicates involvement of a proteolytic activity in promoting RSM.

5. RSM is enhanced by the yeast M16A proteases Axl1p and Ste23p:

To further advance the hypothesis that RSM is promoted by a proteolytic activity, we sought to identify protease gene(s) involved. Using a candidate approach, we first examined other proteases associated with **a**-factor maturation, specifically the M16A subfamily proteases Axl1p and Ste23p. These proteases independently cleave an N-terminal extension found on the

a-factor precursor during **a**-factor biogenesis, with Ax11p being responsible for the majority of this activity (35). When over-expressed, each protease was capable of enhancing RSM associated with **a**-factor-CKQQ (Figure 4.5A). Protease over-expression did not promote RSM in the presence of wildtype **a**-factor or a charge switch mutant (CVIA and CDQQ, respectively) (Figure 4.5B). This observation suggests that the effect of protease over-expression is linked to recognition of a specific subset of CaaX motifs. Our results are consistent with Ax11p and Ste23p contributing to the proteolytic activity that promotes RSM but do not exclude the possibility of the M16A proteases activating a secondary protease having this role.

6. The N-terminal extension of a-factor is important for RSM:

Given the possible and likely involvement of M16A enzymes in cleaving RSM-promoting CaaX motifs, we hypothesized that the N-terminal extension of **a**-factor would somehow be involved in regulating RSM. The first third of this N-terminal extension is removed by Ste24p to yield a partial extension, which is subsequently fully removed by the activity of a yeast M16A enzyme (36). By analogy to substrates of other M16A enzymes, the partial N-terminal extension presumably binds an exosite on the M16A enzyme (37).

First, we examined whether the N-terminal extension shields the natural **a**-factor CaaX motif but not RSM-promoting CaaX motifs from M16A recognition. To test this possibility, **a**-factor was expressed with and without its N-terminal extension using a ubiquitin fusion approach that allows expression of very short peptides (36, 38). The fusions incorporated the full-length **a**-factor precursor sequence (Ubi-P1), a truncated sequence reflecting partial loss of the N-terminal extension (Ubi-P2), or the mature sequence of **a**-factor (Ubi-M) (Figure 4.6A). The fusions had an associated CaaX motif, either CVIA or CKQQ. None of the fusions containing the wildtype

CaaX motif CVIA were capable of promoting RSM despite encoding functional **a**-factor products (Figure 4.6B), indicating that the N-terminal extension does not simply shield CaaX motifs from M16A recognition. Of note, Ubi-M appeared less effective at promoting mating in the presence of CaaX proteases relative to its longer counterparts, which is consistent with its rapid turnover in cells (36).

We next examined whether the N-terminal extension is needed to recruit the **a**-factor precursor to the ‘peptidosome’ cavity of the M16A enzyme where CaaX motifs with appropriate biophysical properties (*i.e.*, RSM promoting motifs) would be cleaved. To test this possibility, ubiquitin-**a**-factor fusions having a CKQQ motif were evaluated (Figure 4.6C). In the presence of CaaX proteolytic activity, all of the ubiquitin fusions had reduced ability to promote mating relative to their CVIA counterparts, with Ubi-M(CKQQ) being completely incapable of promoting mating. RSM was observed with Ubi-P1(CKQQ) and Ubi-P2(CKQQ) but not with Ubi-M(CKQQ). The fact that Ubi-P2(CKQQ) promotes mating somewhat better than Ubi-P1(CKQQ) is consistent with our assertion that the N-terminal extension of **a**-factor, specifically the partial extension present on the P2 intermediate, is important for RSM. However, the fact that Ubi-M(CKQQ) cannot promote mating in the presence of CaaX proteases, precludes our ability to conclusively demonstrate an essential recruitment role for the N-terminal extension in RSM.

DISCUSSION

We have identified multiple CaaX motifs that, when used in lieu of the natural **a**-factor CVIA motif, can promote yeast mating in the absence of the established CaaX proteases Rce1p and Ste24p. Rce1p and Ste24p independent mating (RSM) can be promoted by several CaaX motifs naturally present in the yeast genome as well as multiple synthetic sequences (Table 4.3).

Our genetic analysis allows us to project that approximately 75 motifs can promote RSM. These motifs represent approximately 1% of all possible CaaX permutations. The fact that RSM-promoting motifs have been previously overlooked is probably not surprising given that only about 1% of CaaX motifs have been previously evaluated in the context of the **a**-factor reporter. (3, 26, 34). Moreover, with the exception of CKIA, none of the previously evaluated motifs were matches for the consensus sequence C(K/R/H)aX, which we have derived as a good but not absolute predictor of an RSM-promoting motif. Sequencing of all the RSM-promoting motifs identified by our unbiased genetic screen may provide additional insight into whether there is an RSM consensus motif. Evaluation of existing motifs suggests that a charged residue at the X position may not be compatible with RSM.

Our observations are consistent with RSM involving proteolysis of the susceptible motifs. Highly likely candidate proteases are the yeast M16A metalloproteases Axl1p and Ste23p that are normally involved in the proteolytic trimming of the N-terminal extension of the yeast **a**-factor precursor. Whether these proteases will recognize an RSM-promoting motif *in vitro* remains to be determined and will require purification of the yeast M16A enzymes and the synthesis of a compatible substrate. The hypothesis that yeast M16A enzymes can cleave CaaX motifs is further supported by additional observations. Three yeast activities, one membrane-associated and two soluble, have been previously identified that are able to cleave the **a**-factor CaaX motif *in vitro* (39, 40). The membrane-associated activity is likely a combination of Rce1p and Ste24p activities, and neither can be responsible for RSM due to their absence in our test strain. One of the soluble activities is a *PEP4*-dependent carboxypeptidase, most likely carboxypeptidase Y. This enzyme is a compartmentalized vacuolar protease and is not expected to come in contact with **a**-factor intermediates, which are hypothesized to be modified by

enzymes having cytosol-oriented active sites; mature **a**-factor is exported directly from the cytosol and across the plasma membrane by the Ste6p ABC-type transporter (41). The second soluble *in vitro* activity is associated with an undefined 110 kDa enzyme and is phenanthroline-sensitive. Axl1p and Ste23p are approximately this size and are both predicted to be phenanthroline-sensitive based on their functional homology to other M16A proteases, such as the human insulin-degrading enzyme (IDE) (35). Thus, they are likely responsible for the *in vitro* activity reported. However, one major inconsistency remains between our *in vivo* observations and the reported 100 kDa *in vitro* activity that cleaves CaaX motifs. We do not observe cleavage of the CVIA CaaX motif *in vivo* whereas this is observed *in vitro*. A major difference between the two types of experiments is, respectively, the use of a full-length biologically active reporter vs. a shorter peptide-based biologically inactive reporter, and this could underlie the specificity differences observed.

Another issue that remains to be resolved is the physiological impact of M16A enzymes on the maturation of CaaX proteins having RSM-promoting motifs. We believe that M16A enzymes cleave RSM-promoting motifs only in the specific context of the yeast **a**-factor reporter. This conclusion is based on the observation that the **a**-factor CKQQ variant, when produced without its N-terminal extension, is an unsuitable substrate for RSM. The N-terminal extension, by analogy to M16A other substrates, is likely involved in binding to the exosite of M16A enzymes. This extension helps anchor the substrate in the so-called “peptidosome” cavity of the M16A enzyme. We thus hypothesize that certain CaaX motifs, once drawn into the cavity of an M16A enzyme, have the propensity to compete against the ideal M16A cleavage sequence for binding to the active site, perhaps by their ability to form extended beta sheet interactions with the M16A enzyme. The ability of other substrates to behave similarly would be limited to those

having an exosite binding sequence and being small enough to fit within the M16A enzyme cavity, which is predicted to hold proteins less than 50 amino acids in size (37). None of the naturally occurring yeast proteins having RSM promoting motifs are small enough to fit in an M16A cavity. Hence, we believe that these proteins are not modified to any appreciable extent by M16A enzymes. However, we cannot discount the possibility that small CaaX proteins with RSM promoting motifs exist in other systems that would be suitable M16A substrates in those systems. If such candidates exist in humans, IDE might be able to help mature these proteins under conditions where CaaX protease inhibition is a therapeutic strategy.

ACKNOWLEDGEMENTS

We are grateful to members of the Schmidt lab for technical advice and critical discussions. This work was supported in part through a Georgia Cancer Coalition Distinguished Cancer Clinician/Scientist Scholar award (to WKS) and a grant from the National Institutes of Health (to WKS; R01 GM67092).

REFERENCES

1. Spence R and Casey PJ: Mechanism of catalysis by protein farnesyltransferase. In: Tamanoi F and Sigman D (eds). *The Enzymes*. Academic Press, New York, NY, 2001, pp 1-18.
2. Young SG, Ambroziak P, Kim E and Clarke S: Postisoprenylation protein processing: CXXX (CaaX) endoproteases and isoprenylcysteine carboxyl methyltransferase. In: Tamanoi F and Sigman DS (eds). *The Enzymes*. Academic Press, New York, NY, 2001, pp 155-213.
3. Boyartchuk VL, Ashby MN and Rine J: Modulation of Ras and **a**-factor function by carboxyl-terminal proteolysis. *Science* 275: 1796-800, 1997.

4. Tam A, Schmidt WK and Michaelis S: The multispanning membrane protein Ste24p catalyzes CAAX proteolysis and NH₂-terminal processing of the yeast **a**-factor precursor. J Biol Chem 276: 46798-806, 2001.
5. Hrycyna CA, Sapperstein SK, Clarke S and Michaelis S: The *Saccharomyces cerevisiae* *STE14* gene encodes a methyltransferase that mediates C-terminal methylation of **a**-factor and RAS proteins. EMBO J. 10: 1699-1709, 1991.
6. Marcus S, Caldwell GA, Miller D, Xue C-B, Naider F and Becker JM: Significance of C-terminal cysteine modifications to the biological activity of the *Saccharomyces cerevisiae* **a**-factor mating pheromone. Mol. Cell. Biol. 11: 3603-12, 1991.
7. Sapperstein S, Berkower C and Michaelis S: Nucleotide sequence of the yeast *STE14* gene, which encodes farnesylcysteine carboxyl methyltransferase, and demonstration of its essential role in **a**-factor export. Mol. Cell. Biol. 14: 1438-1449, 1994.
8. Kim E, Ambroziak P, Otto JC, Taylor B, Ashby M, Shannon K, Casey PJ and Young SG: Disruption of the mouse Rce1 gene results in defective Ras processing and mislocalization of Ras within cells. J. Biol. Chem. 274: 8383-8390, 1999.
9. Bergo MO, Leung GK, Ambroziak P, Otto JC, Casey PJ and Young SG: Targeted inactivation of the isoprenylcysteine carboxyl methyltransferase gene causes mislocalization of K-Ras in mammalian cells. J. Biol. Chem. 275: 17605-10, 2000.
10. Bergo MO, Ambroziak P, Gregory C, George A, Otto JC, Kim E, Nagase H, Casey PJ, Balmain A and Young SG: Absence of the CAAX endoprotease Rce1: effects on cell growth and transformation. Mol Cell Biol 22: 171-81, 2002.
11. Manandhar SP, Jacobsen WH and Schmidt WK: Chemical inhibition of CaaX protease activity disrupts yeast Ras localization. Eukaryot Cell: submitted.

12. Young SG, Fong LG and Michaelis S: Prelamin A, Zmpste24, misshapen cell nuclei, and progeria--new evidence suggesting that protein farnesylation could be important for disease pathogenesis. *J Lipid Res* 46: 2531-58, 2005.
13. Basso AD, Kirschmeier P and Bishop WR: Lipid posttranslational modifications. Farnesyl transferase inhibitors. *J Lipid Res* 47: 15-31, 2006.
14. Winter-Vann AM and Casey PJ: Post-prenylation-processing enzymes as new targets in oncogenesis. *Nat Rev Cancer* 5: 405-12, 2005.
15. Manandhar SP, Hildebrandt ER and Schmidt WK: Small molecule inhibitors of the Rce1p CaaX protease. *Journal of Biomolecular Screening* 12: 883-893, 2007.
16. Wang M, Tan W, Zhou J, Leow J, Go M, Lee HS and Casey PJ: A small molecule inhibitor of isoprenylcysteine carboxymethyltransferase induces autophagic cell death in PC3 prostate cancer cells. *J Biol Chem* 283: 18678-84, 2008.
17. Blum R, Cox AD and Kloog Y: Inhibitors of chronically active ras: potential for treatment of human malignancies. *Recent Pat Anticancer Drug Discov* 3: 31-47, 2008.
18. Schmidt WK, Tam A, Fujimura-Kamada K and Michaelis S: Endoplasmic reticulum membrane localization of Rce1p and Ste24p, yeast proteases involved in carboxyl-terminal CAAX protein processing and amino-terminal **a**-factor cleavage. *Proc Natl Acad Sci USA* 95: 11175-80, 1998.
19. Otto JC, Kim E, Young SG and Casey PJ: Cloning and characterization of a mammalian prenyl protein-specific protease. *J. Biol. Chem.* 274: 8379-82, 1999.
20. Pei J and Grishin NV: Type II CAAX prenyl endopeptidases belong to a novel superfamily of putative membrane-bound metalloproteases. *Trends Biochem Sci* 26: 275-7, 2001.

21. Bergo MO, Gavino B, Ross J, Schmidt WK, Hong C, Kendall LV, Mohr A, Meta M, Genant H, Jiang Y, Wisner ER, Van Bruggen N, Carano RA, Michaelis S, Griffey SM and Young SG: Zmpste24 deficiency in mice causes spontaneous bone fractures, muscle weakness, and a prelamin A processing defect. *Proc Natl Acad Sci U S A* 99: 13049-54, 2002.
22. Pendas AM, Zhou Z, Cadinanos J, Freije JM, Wang J, Hultenby K, Astudillo A, Wernerson A, Rodriguez F, Tryggvason K and Lopez-Otin C: Defective prelamin A processing and muscular and adipocyte alterations in Zmpste24 metalloproteinase-deficient mice. *Nat Genet* 31: 94-9, 2002.
23. Bracha K, Lavy M and Yalovsky S: The Arabidopsis AtSTE24 is a CAAX protease with broad substrate specificity. *J Biol Chem* 277: 29856-64, 2002.
24. Cadiñanos J, Schmidt WK, Fueyo A, Varela I, Lopez-Otin C and Freije JM: Identification, functional expression and enzymic analysis of two distinct CaaX proteases from *Caenorhabditis elegans*. *Biochem J* 370: 1047-54, 2003.
25. Cadiñanos J, Varela I, Mandel D, Schmidt WK, Díaz-Perales A, López-Otín C and JMP F: AtFACE-2, a prenylated-protein protease from *Arabidopsis thaliana* related to Ras converting enzymes. *Journal of Biological Chemistry* 278: 42091-7, 2003.
26. Plummer LJ, Hildebrandt ER, Porter SB, Rogers VA, McCracken J and Schmidt WK: Mutational analysis of the ras converting enzyme reveals a requirement for glutamate and histidine residues. *J Biol Chem* 281: 4596-605, 2006.
27. Michaelis S and Herskowitz I: The **a**-factor pheromone of *Saccharomyces cerevisiae* is essential for mating. *Mol. Cell. Biol.* 8: 1309-18, 1988.
28. Elble R: A simple and efficient procedure for transformation of yeasts. *BioTechniques* 13: 18-20, 1992.

29. Oldenburg KR, Vo KT, Michaelis S and Paddon C: Recombination-mediated PCR-directed plasmid construction *in vivo* in yeast. *Nucleic Acids Res* 25: 451-452, 1997.
30. Nijbroek GL and Michaelis S: Functional assays for the analysis of yeast *ste6* mutants. *Methods Enzymol* 292: 193-212, 1998.
31. Robzyk K and Kassir Y: A simple and highly efficient procedure for rescuing autonomous plasmids from yeast. *Nucleic Acids Research* 20: 3790, 1992.
32. Gotte K, Girzalsky W, Linkert M, Baumgart E, Kammerer S, Kunau WH and Erdmann R: Pex19p, a farnesylated protein essential for peroxisome biogenesis. *Mol Cell Biol* 18: 616-28, 1998.
33. Collins SP, Reoma JL, Gamm DM and Uhler MD: LKB1, a novel serine/threonine protein kinase and potential tumour suppressor, is phosphorylated by cAMP-dependent protein kinase (PKA) and prenylated *in vivo*. *Biochem J* 345 Pt 3: 673-80, 2000.
34. Trueblood CE, Boyartchuk VL, Picologlou EA, Rozema D, Poulter CD and Rine J: The CaaX proteases, Afc1p and Rce1p, have overlapping but distinct substrate specificities. *Mol. Cell. Biol.* 20: 4381-92, 2000.
35. Kim S, Lapham A, Freedman C, Reed T and Schmidt W: Yeast as a tractable genetic system for functional studies of the insulin-degrading enzyme. *J Biol Chem.* 280: 27481-27490, 2005.
36. Tam A, Nouvet F, Fujimura-Kamada K, Slunt H, Sisodia SS and Michaelis S: Dual roles for Ste24p in yeast **a**-factor maturation: NH₂-terminal proteolysis and COOH-terminal CAAX processing. *J. Cell Biol.* 142: 635-49, 1998.
37. Shen Y, Joachimiak A, Rosner MR and Tang WJ: Structures of human insulin-degrading enzyme reveal a new substrate recognition mechanism. *Nature* 443: 870-4, 2006.

38. Boyartchuk VL and Rine J: Roles of prenyl protein proteases in maturation of *Saccharomyces cerevisiae* **a**-factor. *Genetics* 150: 95-101, 1998.
39. Ashby M, King D and Rine J: Endoproteolytic processing of a farnesylated peptide *in vitro*. *Proc. Natl. Acad. Sci. USA* 89: 4613-17, 1992.
40. Hrycyna CA and Clarke S: Maturation of isoprenylated proteins in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 267: 10457-64, 1992.
41. Kuchler K, Sterne RE and Thorner J: *Saccharomyces cerevisiae* *STE6* gene product: a novel pathway for protein export in eukaryotic cells. *EMBO J.* 8: 3973-3984, 1989.
42. Powers S, Michaelis S, Broek D, Santa Anna S, Field J, Herskowitz I and Wigler M: *RAM*, a gene of yeast required for a functional modification of RAS proteins and for production of mating pheromone **a**-factor. *Cell* 47: 413-22, 1986.
43. Chen P, Sapperstein S, Choi JD and Michaelis S: Biogenesis of the *Saccharomyces cerevisiae* mating pheromone **a**-factor. *J. Cell. Biol.* 136: 251-69, 1997.
44. Adames N, Blundell K, Ashby MN and Boone C: Role of yeast insulin-degrading enzyme homologs in propheromone processing and bud site selection. *Science* 270: 464-67, 1995.
45. Sikorski RS and Hieter P: A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* 122: 19-27, 1989.
46. Zhang Y, Nijbroek G, Sullivan ML, McCracken AA, Watkins SC, Michaelis S and Brodsky JL: Hsp70 molecular chaperone facilitates endoplasmic reticulum-associated protein degradation of Cystic Fibrosis Transmembrane Conductance Regulator in yeast. *Mol Biol Cell* 12: 1303-14, 2001.

47. Fujimura-Kamada K, Nouvet FJ and Michaelis S: A novel membrane-associated metalloprotease, Ste24p, is required for the first step of NH₂-terminal processing of the yeast **a**-factor precursor. *J. Cell Biol.* 136: 271-85, 1997.
48. Schmidt WK, Tam A and Michaelis S: Reconstitution of the Ste24p-dependent N-terminal proteolytic step in yeast **a**-factor biogenesis. *J. Biol. Chem.* 275: 6227-33, 2000.

Table 4.1. Strains used in this study.

STRAIN	GENOTYPE	REFERENCE
IH1783	<i>MATa trp1 leu2 ura3 his4</i>	ATCC#204278
IH1793	<i>MATα lys1</i>	ATCC#204279
RC757	<i>MATα sst2-1 rme his6 met1 can1 cyh2</i>	(42)
SM2331	<i>MATa trp1 leu2 ura3 his4 mfa1Δ mfa2Δ</i>	(43)
yWS164	<i>MATa trp1 leu2 ura3 his4 mfa1Δ mfa2Δ rce1::TRP1 ste24::KAN^R</i>	(25)
yWS829	<i>MATa trp1 leu2 ura3 his4 mfa1Δ mfa2Δ rce1::TRP1 ste24::KAN^R ste14::URA3</i>	This study

Table 4.2. Plasmids used in this study.

PLASMID	GENOTYPE	REFERENCE
P80	<i>CEN URA3 AXL1</i>	(44)
pRS315	<i>CEN LEU2</i>	(45)
pRS316	<i>CEN URA3</i>	(45)
pRS415	<i>CEN LEU2</i>	(45)
pRS424	<i>2μ TRP1</i>	(45)
pSM284	<i>integrating ste14:URA3</i>	(7)
pSM703	<i>2μ URA3 P_{PGK}</i>	(46)
pSM1107	<i>CEN URA3 HA::STE24</i>	(47)
pSM1153	<i>CEN TRP1 AXL1</i>	(18)
pSM1366	<i>CEN URA3 UBI-MFA1 M</i>	(36)
pSM1368	<i>CEN URA3 UBI-MFA1 P1</i>	(36)
pSM1369	<i>CEN URA3 UBI-MFA1 P2</i>	(36)
pSM1605	<i>2μ URA3 MFA1</i>	(48)
pSM1314	<i>CEN URA3 RCE1::HAc</i>	(18)
pWS196	<i>2μ URA3 MFA1-CASQ</i>	This study
pWS438	<i>2μ LEU2 MFA1</i>	(25)
pWS601	<i>2μ URA3 P_{PGK}-AXL1</i>	This study
pWS602	<i>2μ URA3 P_{PGK}-STE23-BglII</i>	This study
pWS610	<i>CEN LEU2 MFA1</i>	This study
pWS612	<i>CEN LEU2 MFA1-CASQ</i>	This study
pWS654	<i>2μ LEU2 MFA1-CALQ</i>	(26)
pWS718	<i>2μ LEU MFA1-CKQQ</i>	This study
pWS727	<i>CEN LEU2 MFA1-CKQQ</i>	This study
pWS737	<i>CEN LEU2 MFA1-CAQQ</i>	This study
pWS738	<i>CEN LEU2 MFA1-CKAQ</i>	This study
pWS739	<i>CEN LEU2 MFA1-CKQA</i>	This study
pWS817	<i>CEN LEU2 MFA1-CALQ</i>	This study
pWS844	<i>CEN LEU2 MFA1-CRQQ</i>	This study
pWS845	<i>CEN LEU2 MFA1-CKIS</i>	This study
pWS846	<i>CEN LEU2 MFA1-CKQS</i>	This study
pWS847	<i>CEN LEU2 MFA1-CKGE</i>	This study
pWS848	<i>CEN LEU2 MFA1-CKCI</i>	This study
pWS849	<i>CEN LEU2 MFA1-CKYI</i>	This study
pWS850	<i>CEN LEU2 MFA1-CKCT</i>	This study
pWS851	<i>CEN LEU2 MFA1-CEQQ</i>	This study
pWS852	<i>CEN LEU2 MFA1-CHQQ</i>	This study
pWS853	<i>CEN LEU2 MFA1-CDQQ</i>	This study
pWS854	<i>CEN LEU2 MFA1-CRVK</i>	This study
pWS855	<i>CEN LEU2 MFA1-CRNR</i>	This study
pWS856	<i>CEN LEU2 MFA1-CRMV</i>	This study
pWS883	<i>2μ LEU2 MFA1-CKVA</i>	This study
pWS884	<i>2μ LEU2 MFA1-CRVA</i>	This study
pWS885	<i>2μ LEU2 MFA1-CRMS</i>	This study

pWS886	<i>2μ LEU2 MFA1-CRVN</i>	This study
pWS887	<i>2μ LEU2 MFA1-CKMT</i>	This study
pWS891	<i>2μ LEU2 MFA1-CKIT</i>	This study
pWS892	<i>CEN URA3 UBI-MFA1 P1 (CKQQ)</i>	This study
pWS893	<i>CEN URA3 UBI- MFA1 P2 (CKQQ)</i>	This study
pWS894	<i>CEN URA3 UBI- MFA1 M (CKQQ)</i>	This study
pWS912	<i>CEN LEU2 MFA1-CRME</i>	This study

Table 3. Summary of ability of CaaX motifs to promote Rce1p and Ste24p-independent mating (RSM).

MOTIF	RSM OBSERVED	SOURCE ^A
CALM	No ^b	synthetic
CALQ	No ^b	synthetic
CAMQ	No ^b	synthetic
CAQQ	No	synthetic
CASQ	No	<i>YDJI</i>
CDQQ	No	synthetic
CEQQ	No	synthetic
CHQQ	Yes	synthetic
CKAQ	Yes	synthetic
CKCI	No	<i>POP8</i>
CKCT	No	<i>YBL049w/MOHI</i>
CKGE	No	<i>YMR197c/VTII</i>
CKIS	Yes	<i>YPR092w^c</i>
CKIT	Yes	RSM screen
CKMT	Yes	RSM screen
CKQA	Yes	synthetic
CKQQ	Yes	<i>PEX19, LKB1, NAP1</i>
CKQS	Yes	<i>NAP1</i>
CKVA	Yes	RSM screen
CKYI	No	<i>YMR060c/SAM37/TOM37</i>
CRME	No	<i>YJL059W/YHC3</i>
CRMS	Yes	RSM screen
CRMV	Yes ^b	synthetic
CRNR	No	<i>YML041C/VPS71</i>
CRQQ	Yes	synthetic
CRVA	Yes	RSM screen
CRVK	No	<i>YMR158W/MRPS8</i>
CRVN	Yes	RSM screen
CVIA	No	<i>MFA1, MFA2</i>

^a All of the indicated genes are those of *S. cerevisiae* except for LKB1 and NAP1 which are human.

^b Krishnankutty and Schmidt, unpublished observation; see also (26)

^c Reported as a dubious ORF in the *Saccharomyces* Genome database (www.yeastgenome.org).

Figure 4.1. *CaaX proteins are extensively modified post-translationally.*

The C-terminal tetrapeptide CaaX motif directs three ordered post-translational modifications, including isoprenylation, proteolysis and carboxymethylation. Interfering with these steps can disrupt the activity and/or localization of the protein being modified.

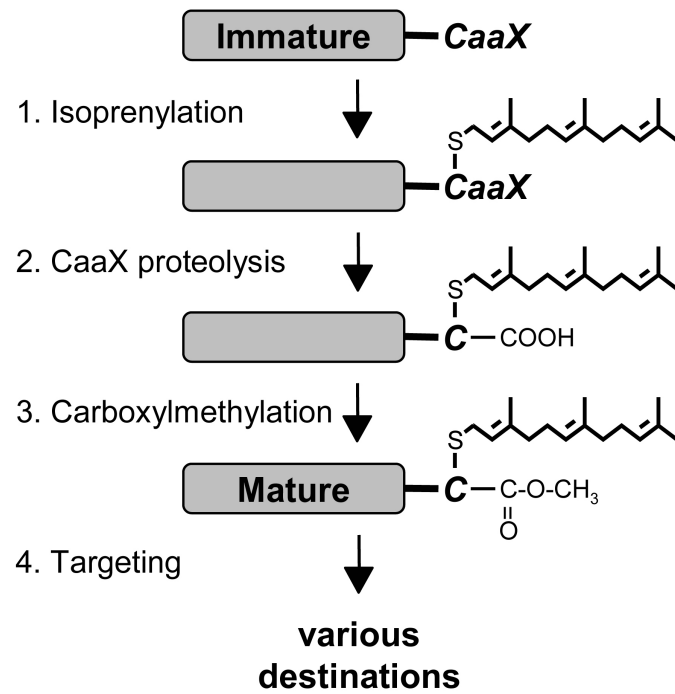


Figure 4.2. *Evidence for Rce1p and Ste24p-independent mating (RSM).*

A) The **a**-factor-CKQQ variant promotes mating in a yeast strain lacking endogenous copies of the CaaX proteases and **a**-factor genes (yWS164). This phenotype is not associated when **a**-factor is appended with its natural CaaX motif (CVIA) or one that is Ste24p-specific (CASQ). When co-expressed with Ste24p, but not Rce1p, the CKQQ variant promotes more efficient mating. The plasmids used were pRS316, pSM1107, pSM1314, pWS610, pWS612, and pWS727. **B)** Serial dilution mating tests were conducted to quantifiably compare the amount of mating promoted by **a**-factor having its natural CaaX motif (CVIA) or a variant (CKQQ) in the presence and absence of CaaX proteases. In the presence of Ste24p, the CKQQ variant mates as effectively as wildtype **a**-factor. The plasmids used were pWS610 (CVIA) and pWS727 (CKQQ).

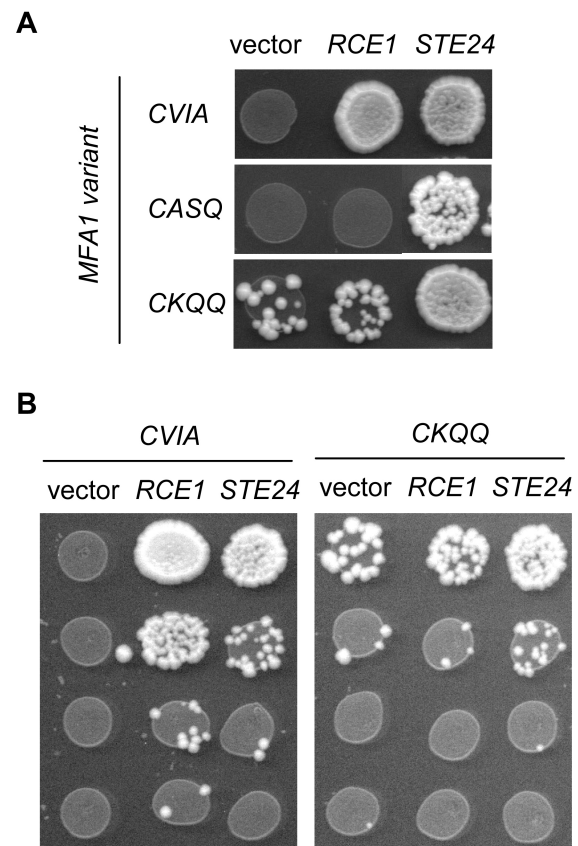


Figure 4.3. *Multiple CaaX motifs promote RSM.*

Serial dilution mating tests were conducted as described in Figure 2 using yWS164 and plasmids encoding the indicated a-factor CaaX variants. Only the first dilution spot is shown for each mating test. **A)** Ala substitutions at the a1, a2 and X positions of the CKQQ motif. The plasmids used were pWS610, pWS727, pWS737, pWS738, and pWS739. **B)** Charged amino acid substitutions at the a1 position. The plasmids used were pWS610, pWS727, pWS844, pWS851, pWS852, and pWS853. **C)** Naturally occurring motifs that correspond to the consensus C(K/R/H)aX. See Table 3 for the source gene of the natural CaaX motifs. The plasmids used were pWS845, pWS846, pWS847, pWS848, pWS849, pWS850, pWS854, pWS855, and pWS912.

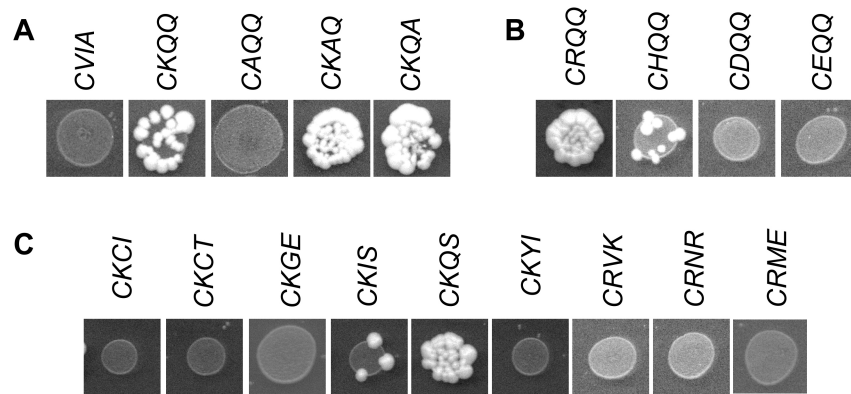


Figure 4.4. *RSM requires carboxymethylation.*

A) Models for RSM. In the absence of Rce1p and Ste24p, CaaX proteolysis of **a**-factor-CKQQ is mediated by either an ER (1) or non-ER-localized (2) protease (Rsm1p). If the latter, the dependence of RSM on Ste14p indicates that a trafficking step is required to return proteolyzed **a**-factor to the Ste14p ICMT that resides at the ER. Alternatively, non-proteolyzed **a**-factor could promote mating (3). In this scenario, RSM would be independent of Ste14p. ER – endoplasmic reticulum; PM – plasma membrane. **B)** Loss of Ste14p in a CaaX protease deficient strain prevents RSM. The plasmids used were pWS610 and pWS738, which were separately transformed into yWS164 and yWS829. yWS164-derived strains were additionally transformed with pRS316 to provide a comparable set of auxotrophic markers to that of the yWS829-derived strains.

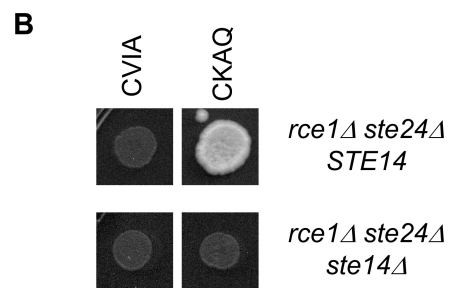
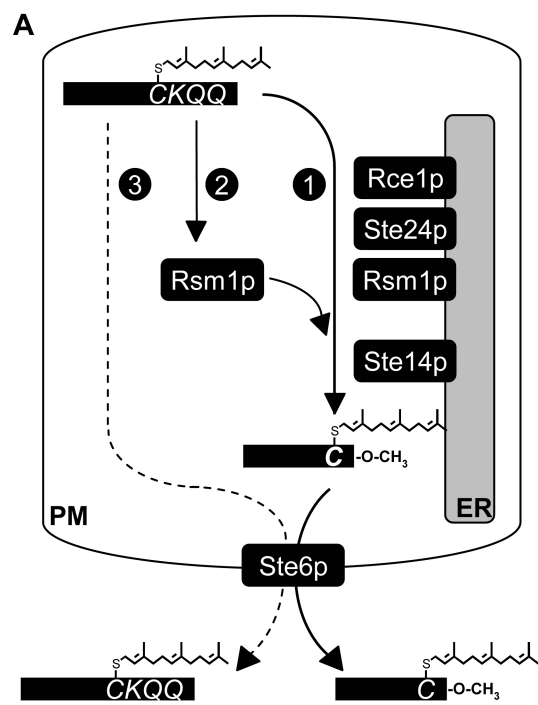


Figure 4.5. *RSM is enhanced by M16A proteases.*

A) Axl1p and Ste23p were independently over-expressed in yWS164 in the presence of the **a**-factor CKQQ variant, and the strains were subjected to a serial dilution mating test as described in Figure 2. The first two dilution spots are shown. The plasmids used were pRS316, pWS601, pWS602, and pWS727. **B)** The ability of over-expressed Ste23p to enhance RSM was evaluated for CaaX motifs that were previously identified as either not promoting (e.g. CVIA and CDQQ) or promoting RSM (e.g. CKQQ). The plasmids used were pWS602, pWS610, pWS727, and pWS853.

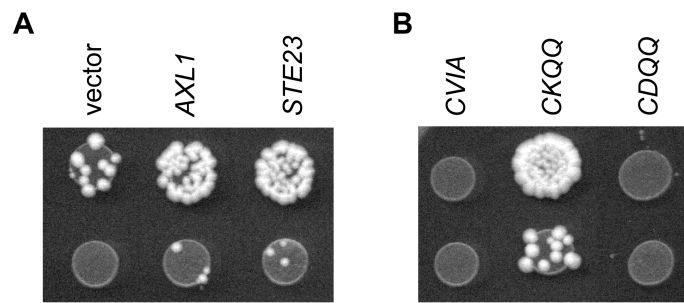


Figure 4.6. *The N-terminal extension of α -factor modulates RSM.*

A) Cartoon depicting the proteolytic cleavage sites in the α -factor precursor and ubiquitin fusions created to bypass certain proteolytic steps associated with α -factor biogenesis. See text for additional details on the ubiquitin fusions. **B)** Ubiquitin fusions to various lengths of the α -factor precursor were expressed in either SM2331 or yWS164, and the strains subjected to a serial dilution mating test as described in Figure 2. The plasmids used were pRS316, pSM1366, pSM1368, and pSM1369. **C)** A mating test was performed as described in panel B using ubiquitin- α -factor fusions appended with the CKQQ motif. The plasmids used were pRS316, pWS892, pWS893, and pWS894.

