Mechanism of Silencing the Catalytic Domain by the Regulatory Membrane Lipid Binding Domain of an Amphitropic Cytidylyltransferase

by
Harris Kuo-Hao Huang
B.Sc., University of British Columbia, 2004

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in the

Department of Molecular Biology and Biochemistry
Faculty of Science

© Harris Kuo-Hao Huang 2011

SIMON FRASER UNIVERSITY

Summer 2011

All rights reserved. However, in accordance with the *Copyright Act of Canada*, this work may be reproduced, without authorization, under the conditions for *Fair Dealing*. Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.

Approval

Name: **Harris Kuo-Hao Huang**

Degree: **Doctor of Philosophy**

Title of Thesis: **Mechanism of silencing the catalytic domain by the regulatory membrane lipid binding domain of an amphitropic cytidylyltransferase**

Examining Committee:

Chair: **Dr. Edgar C. Young**
Associate Professor, Department of Molecular Biology and Biochemistry

_____

Dr. Rosemary Cornell
Senior Supervisor
Professor, Department of Molecular Biology and Biochemistry

_____

Dr. George R. Agnes
Supervisor
Professor, Department of Chemistry

_____

Dr. David Vocadlo
Supervisor
Associate Professor, Department of Chemistry

_____

Dr. Dipankar Sen
Supervisor
Professor, Department of Molecular Biology and Biochemistry

_____

Dr. Lisa Craig
Internal Examiner
Associate Professor, Department of Molecular Biology and Biochemistry

Dr. Lawrence McIntosh
External Examiner
Professor, Department of Chemistry
University of British Columbia

Date Defended/Approved: June 1, 2011

SFU SIMON FRASER UNIVERSITY
LIBRARY

Declaration of
Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website <www.lib.sfu.ca> at: <http://ir.lib.sfu.ca/handle/1892/112>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

Abstract

CTP: phosphocholine cytidylyltransferase (CT) is an auto-regulated, homodimeric enzyme that catalyzes the rate-limiting step in the synthesis of phosphatidylcholine (PC). CT is activated by binding to PC-depleted membranes using a lipid-induced amphipathic helix (domain M). This same M domain functions to silence catalysis when the enzyme is not membrane-engaged. While the structure of CT's catalytic domain has recently been solved (PDB, 3HL4), if and how domain M makes contact with it to silence catalysis remains a mystery. To identify contact sites between domain M and other CT domains, I constructed single cysteine substitutions along domain M and conjugated each site to a sulfhydryl-reactive, biotinylated, benzophenone, BBP. After photo-cross-linking and trypsin digestion, the cross-linked peptides were affinity-purified and identified by mass spectrometry.

Each domain M-conjugated site forged cross-links to the same set of catalytic domain peptides, which flank one side of the active site, and these contacts were broken upon membrane insertion of domain M. I then targeted the most conserved region in domain M for mutagenesis. Mutation of F289 and F293 to aspartates obliterated cross-links between the catalytic domain and the BBP-labeled site at neighboring Cys-301. However, loss of this contact did not relieve inhibition of catalysis, suggesting that this site is but one of several cooperating inhibitory segments in the M domain.

Evidence for a partially disordered domain M which acquires order upon contact with the catalytic domain was obtained by fluorescence anisotropy, monitoring Oregon Green conjugated to sites in three distinct CT domains. All domain M sites had similar anisotropy values intermediate between the rigid catalytic domain and the very flexible C-terminal tail. Urea denaturation analysis suggested a weakly folded structure for domain M. Breaking a contact between M and C by the F → D mutation dramatically lowered the anisotropy of that sub-region and its resistance to unfolding by urea.

These results support an emerging model in which silencing by domain M is achieved through multiple transient, alternating, contacts with the catalytic domain. Contacts with the catalytic domain may introduce ordering into an otherwise flexible domain M. These inhibitory contacts are replaced with protein-lipid interaction upon activation.

Keywords: cytidylyltransferase; auto-inhibition; photo-cross-linking; mass spectrometry; fluorescence anisotropy

Acknowledgements

I would like to thank my supervisor, Rosemary Cornell, for her enthusiasm and continuing assistance. Completion of my thesis, and more importantly, development as a scientist would not have been possible without her unrelenting effort in mentorship. I would also like to thank my committee members, Drs. George Agnes, David Vocadlo, and Dipankar Sen for their time, helpful suggestions, and support throughout the years.

Thanks to all my labmates, past and present, for being a great group to work with. Thanks to Mike Bogan and Dr. Melanie O'Neill for their lessons in operating the mass spectrometer and fluorimeter. Thanks to Ziwei Ding for his assistance and mentorship in molecular coloning. Special thanks to Svetla Taneva for her mentorship, patience, help, and understanding. Thanks also to Edward Lau, Jane Huang, Diem Ly Van, Apollos Kim, and all of the many friends I made at SFU.

Thanks to all the MBB labs for sharing reagents and equipment. Thanks also to the department staff for their assistance. Thank the students and professors that I had the opportunity to meet in my TAships, for the fun and rewarding experience.

Finally, thanks to my family, my mom, dad, and brothers for their continuing support and belief in me. Special thanks to my son, Justin, for his innocence and unquestioned love that shade light into the darkest times of my life.

Table of Contents

List of Figures

List of Tables

Glossary

AA	acetamide
ACTH	Adrenocorticotropic hormone fragment 18-39
BBP	(2-[Nα-Benzoylbenzoicamido-N^6-(6-biotinamidocaproyl)-L-lysinylamido]ethyl methanethiosulfonate)
CD	Circular dichroism
β-ME	Beta-mercaptoethanol
CFTR	Cystic fibrosis transmembrane conductance regulator
CHCA	α-Cyano-4-hydroxycinnamic acid
CK	Choline kinase
CT	CTP:phosphocholine cytidylyltransferase
CT-cat	CT catalytic domain
CuPhe	Copper phenanthroline
DAG	Diacylglycerol
DMSO	Dimethyl sulphoxide
DOPG	dioleoylphosphatidylglycerol
DTT	Dithiothreitol
DTT-IAA	Derivatization method where sample is reduced with DTT first, then alkylated with IAA
EC_{50}	Concentration of urea resulting in 50% protein unfolding
ECT	Ethanolamine cytidylyltransferase
EDTA	Ethylenediaminetetraacetic acid
EPR	Electron paramagnetic resonance
ER	Endoplasmic reticulum
ETS	E-twenty six (a transcription factor unique to metazoans)
FAK	Focal Adhesion Kinase

FRET	Fluorescence resonance energy transfer
GCT	Glycerol-3-phosphate cytidylyltransferase
IAA	Iodoacetamide
ID	Intrinsically disordered
LR	Linear mode
m/z	Mass/charge
MALDI-MS	Matrix-assisted laser desorption/ionization mass spectrometry
MTAC (model)	Multiple, transient, alternating contacts (model)
NBD	Nucleotide binding domain
NMR	Nuclear magnetic resonance
OG	Oregon Green
PAGE	Polyacrylamide gel electrophoresis
PAX	Paired box transcription factor
PC	Phosphatidylcholine
PDB	Protein data bank
PDGF	Platelet derived growth factor
PE	Phosphatidylethanolamine
PH (domain)	Pleckstrin homology domain
PIP_2	Phosphatidylinositol 4,5-bisphosphate
PVDF	Poly(vinylidine difluoride
RC (model)	Rigid Contact (model)
SDS	Sodium dodecyl sulphate
Sin	Src interacting (Signal integrating) protein
SNARE	Soluble NSF Attachment Protein Receptor
SRC (kinase)	Sarcoma (tyrosine kinase)
SUV	Small unilamellar vesicle
SAV-HRP	Streptavidin Horse radish peroxidase (in western blot)
Tris	Tris(hydroxymethyl)methylamine

TX	Triton-X
WASP	Wiscott-Aldrich Syndrome Protein
WT	Wild-type

1: Introduction

1.1 Protein regulation by auto-inhibition

CTP:phosphocholine cytidylyltransferase (CT) is an auto-regulated enzyme responsible for the biosynthesis of phosphatidylcholine (PC). When I began my thesis research, the catalytic mechanism of CT and lipid requirement for activation had been studied in much detail. However, the auto-inhibition mechanism remained largely speculative. There are many other proteins that also contain auto-inhibitory properties. Detailed mechanisms of auto-inhibition have been established for some of these, which can provide a starting point for building hypotheses for CT auto-inhibition. Moreover, the general approaches undertaken with other enzymes to decipher their auto-inhibitory mechanisms can be considered for CT. Therefore, before specifically discussing CT auto-inhibition, I will review the types of regulatory mechanisms reported for some well-studied auto-inhibited proteins.

Auto-inhibition is a valuable feature that enables site-specific repression in proteins that require regulation. Mediation of inhibition comes from defined region(s) of the enzyme that possess controlling elements. It often involves specific intra-molecular interaction between the regulatory region and other functional domains, which are often linked together through disordered regions that exhibit the required flexibility for efficient transition from the inactive to active state of the functional domain[1]. The simplest way, and often the first step in identifying auto-inhibition, is to perform an activity analysis comparing fragments of the protein with the full length protein. A deleted region is auto-inhibitory if activity increases upon deletion. However, complexity arises in deciphering the details of how auto-inhibition is exerted by these regions. This often requires activity analyses of mutated variants in combination with several biophysical/biochemical techniques to probe structure, such as circular dichroism (CD) spectroscopy, protease accessibility mapping, fluorescence spectroscopy, nuclear magnetic resonance (NMR) spectroscopy, and X-ray crystallography.

Auto-inhibition can be simply achieved through direct interference with ligand/substrate binding sites in the functional domain[2-6]. For example, masking of the activation domain of the yeast transactivator, Leu3p, is achieved through a ligand-dependent, intra-molecular interaction[2]. In another example, the functional domain of TATA box binding protein negatively regulates itself inter-molecularly through dimerization that results in the burial of DNA binding sites[7].

Direct steric interference is also observed in some auto-inhibited *enzymes*. For example in PIP_2-regulated protein kinase B, the PH domain silences the kinase domain by direct interaction with the active site. In the silenced form, the PH domain inserts into a partially hydrophobic pocket in the kinase active site. Upon activation (by displacement of the PH-PIP_2 complex) this pocket becomes occupied by the phosphorylated activation loop of the kinase[8].

On the other hand, auto-inhibition can be achieved allosterically through indirect structural alteration of the functional domain and/or confinement of a nonproductive conformation in the functional domain, as exemplified by SRC kinases[9], SNARE proteins[10], and p53 proteins[11]. For example, in the protein kinase Src, the contact interface between the regulatory and catalytic domain is distant from the active site. In the inactive form of the kinase, tight contact is made between the phosphotyrosine binding pocket of the SH2 and phospho-Tyr527 at the tail of the kinase domain. Additionally a sticky SH2-SH3 linker peptide runs between the regulatory and kinase domains[12]. These rigid contacts sometimes referred to as *latches*, serve to generate a non-productive active site by altering the structure and orientation of the C-helix and activation loop (in the kinase domain) that are critical for substrate docking. Breaking of these contacts by dephosphorylation of phosphor-Tyr527 dissociates the regulatory and catalytic domains and allows reorganization of these active site motifs, so that the activation loop can become phosphorylated[12].

Relief of auto-inhibition by the direct interference mechanism is most commonly achieved by displacement of the auto-inhibitory domain by another molecule, resulting in an exchange of intra-molecular interactions (inhibitory) with inter-molecular interactions (activating)[9,13,14]. The other molecule could be a partnering protein or a small molecule ligand. For example, in SRC kinase, when the auto-inhibitory SH2 and SH3 domains

interact inter-molecularly with the phosphotyrosine or pro-rich motif in an activating protein partners such as FAK (focal adhesion kinase), Sin (Src interacting or Signal integrating protein), or the PDGF (Platelet derived growth factor) receptor, the intra-molecular auto-inhibitory contact is broken, allowing activation to occur[9]. Other strategies, that can either disrupt direct inhibitory interactions or work via allostery, include post-translational modification. For example, phospholipase C remains inactive through intramolecular interaction between its catalytic domain and inhibitory SH2 domain. Binding to and phosphorylation by receptor tyrosine kinase at specific sites in the SH2 domain induces structural rearrangements of the SH2 domain with respect to the catalytic domain, and lead to this enzyme's activation[15]. Though not as common, activation can also be achieved through site-specific proteolysis of the inhibitory domain, as examplified by activating proteolysis of Notch proteins upon ligand binding[16]. These regulatory mechanisms, except for proteolysis, are reversible and allow for efficient on/off switches and tight control of biological regulatory pathways.

Significant conformational changes may occur in either or both the regulatory and functional domains. For example, in the transcription factor Ets-1, three inhibitory α-helices, HI-1, HI-2, and HI-5 flanking the DNA-binding ETS domain cooperatively exert auto-inhibition on the ETS domain through intra-molecular contacts that stabilize a rigid fold[17,18]. When Ets-1 binds to protein partners such as PAX-5 [19], dramatic unfolding of the auto-inhibitory helix HI-1 occurs, as shown by an increase in proteolytic accessibility and reduced helical content by CD spectroscopy[17]. The ETS domain, freed of contact with the auto-inhibitory segment, becomes more malleable and this aids its interaction with the DNA helix[18,20].

Most intra-molecular interactions involved in auto-inhibition are highly specific in that the residues involved in the interaction sites are often conserved. Site-directed mutagenesis of key residues can disrupt the auto-inhibitory contact and relieve auto-inhibition[2,21]. This implies that these auto-inhibitory domains and their interacting partners are likely to be well folded structures with highly complementary surfaces for maximum contact. In support of this, addition of separately prepared inhibitory domains to the constitutively active functional domains of sigma factor[4] or WASP[22] led to inhibition *in trans*. Furthermore, high-resolution X-ray crystallographic studies show

rigid, well-folded structures for domains locked in auto-inhibited states for some of the well characterized examples[3,9,10].

On the other hand, evidence is accumulating that intrinsically disordered (ID) segments are abundant in multi-domain proteins[23] and that they can serve as regulatory switches[24-26]. This often involves ligand or protein partner-induced ordering of the ID segment, but alternatively, the intrinsic disorder is maintained during the regulatory interaction, with only localized transient ordering. For example, the disordered R region of the cystic fibrosis transmembrane conductance regulator (CFTR) Chloride conductance channel negatively modulates the channel conductance by multiple transient contacts with the nucleotide-binding domain of this channel, and the contact results in only transient ordering of small segments in R, as assessed by NMR analyses[27]. In another example, multiple sites on the disordered cyclin-dependent kinase inhibitor, Sic1, make inhibitory contacts with one core binding site on Sic1's receptor, Cdc4. These contacts, enhanced by phosphorylation, are only suboptimal and switch among the sites in a dynamic equilibrium[25].

How does CT fit into the realm of auto-inhibition among these proteins? Does auto-inhibition in CT involve specific contact between well-folded domains or is the CT regulatory module disordered? Does the regulatory domain of CT exhibit inhibition directly or allosterically? Is there dynamic restructuring during transition between the activated and auto-inhibited state? Before tackling these unknowns, I will first give a brief review on what is known about CT, in terms of its function, regulation and structure.

1.2 A brief history of the role of CT in Phosphatidylcholine metabolism

PC is the major component of eukaryotic membranes, and the most abundant class among all phospholipids[28-30]. PC contributes to the maintenance of the bilayer structure of membranes. It is a primary component in pulmonary surfactant[29-31], and is required for the secretion of very-low-density-lipoproteins[32]. In addition, PC acts as a source for several important signal molecules, such as diacylglycerol (DAG), phosphatidic acid, lysoPC, and arachidonic acid. These products are involved in a wide variety of cellular events, such as cell growth, proliferation, and apoptosis, DNA

synthesis, lipid synthesis, and regulation of ion channel activity[28,33-35].

Three distinct pathways for biosynthesis of PC had been identified. The first pathway involves series of three methylations catalyzed by phosphatidylethanolamine-*N*-methyltransferase that converts phosphatidylethanolamine (PE) to PC. It occurs in many organisms but mainly in the hepatic cells of animals[28,30-32]. The second pathway, found so far in some bacteria, converts choline and CDP-DAG into PC and CMP[30]. The third pathway, the CDP-choline pathway, also known as the Kennedy pathway, transforms dietary choline into PC in three steps. In the first step choline is phosphorylated, catalyzed by choline kinase (CK). Then CT catalyzes the transfer of CMP from CTP to phosphocholine to form CDP-choline. In the final step, CDP-choline condenses with DAG to form PC, catalyzed by CDP-choline:1,2-DAG cholinephosphotransferase[28,30-32].

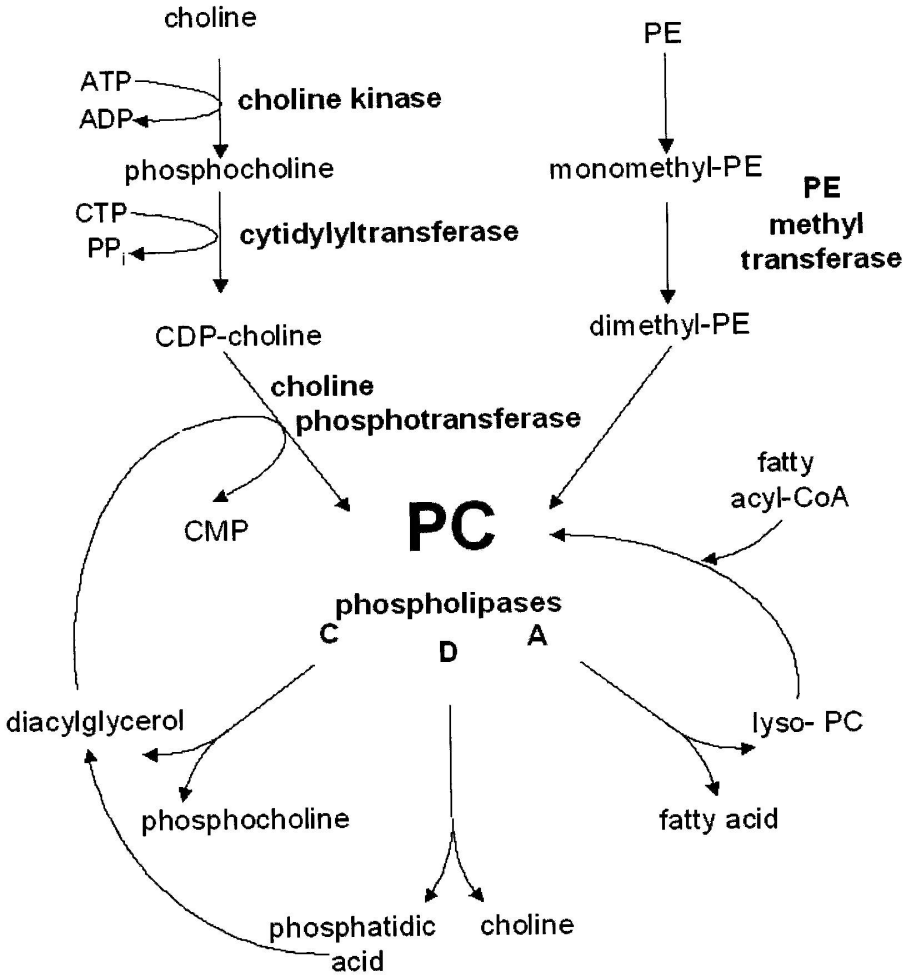


Figure 1.0 – PC metabolic pathways (MSc thesis, Jillian Smith, 2005)

Measurements of radiolabeled choline flux through the CDP-choline pathway strongly suggest that CT catalyzes the rate limiting and regulatory step in the biosynthesis of PC[36,37]. For example in several cultured cells, PC synthesis increased dramatically in response to phospolipase C, which catalyzes PC degredation[38,39]. This increase could not be attributed to changes in the activity of any enzymes in the CDP-choline pathways except for CT[38,39]. Many other agents that alter PC synthesis rates have been shown to work through modulation of CT, mainly at the post-transcriptional level[40].

There are typically two CT genes in animal cells, encoding proteins that are nearly identical in their catalytic domains, but with variable regulatory domains[41]. The importance of each CT isoform is demonstrated by the fact that in mice, disruption of the gene for CTα led to early embryonic lethality[42] and disruption of the tissue-specific CTβ gene led to gonadal dysfunction[43]. While the beta form is cytoplasmic/ER bound[44], the alpha form is found in the nucleus in many but not all cells[45-49].

1.3 Biochemical characterization of mammalian CTα Purification and cloning

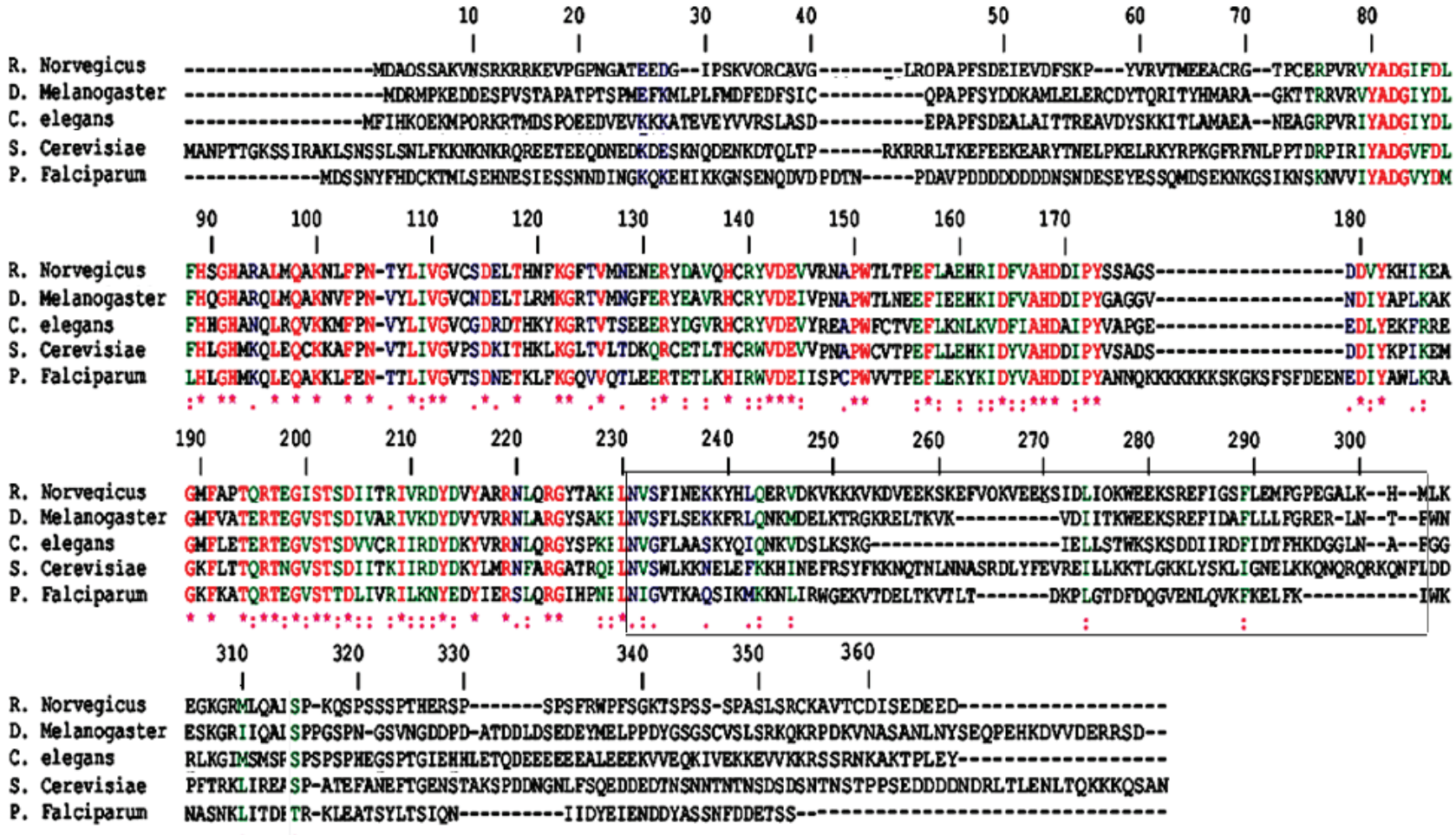
Kennedy and Weiss were the first to discover CT while uncovering the mechanism behind conversion of phosphocholine to PC in rat liver extracts. They also noted that CT can be found in both membrane-bound and soluble forms[50]. Despite attempts from many labs, CT's tendency to aggregate and lose activity resulted in the delay of its purification for 30 years. Finally, it was Weinhold and his colleagues who designed a successful purification protocol, in which detergents and lipids were added, essentially treating CT as if it were a membrane protein[51]. The purified CT had an apparent mass of 45 kDa on SDS-PAGE, and was shown to be a homodimer by sedimentation[52] and chemical cross-linking[53].

Cloning of CT cDNA was first achieved for the yeast enzyme[54], then for rat CTα[55]. A comparison of the two deduced protein sequences yielded a highly conserved region as a candidate for the catalytic domain. Adjacent to this, a lipid binding domain capable of forming amphipathic α-helices was proposed through secondary structural predictions[55]. Cloning of CTs from different organisms, including mammals[56,57]; invertebrates[58,59]; plants[60,61]; and even protist[62], as well as genomic sequencing of

organisms spanning most of the kingdoms of life have now given us a large number of CT sequences for comparative analysis. Sequence identities are 95-99% among mammalian CTs, and greater than 80% among the few vascular plants whose sequences are available. While little similarity was found in other domains, the proposed catalytic domain was highly conserved across numerous species, with a number of residues being absolutely conserved. The designation of catalytic domain was further strengthened by the discovery of a bacterial homolog, glycerol-3-phosphate cytidylyltransferase (GCT), which contains a sequence similar to and no larger than the designated mammalian catalytic domain[63]. Below is an alignment of lipid-regulated CTs using ClustalW.

Figure 1.1 – Sequence alignment of CTs from different species

CT from R. Norvegicus (CTα, Uniprot ID:P19836), D. Melanogaster (Q9W0D9), C. elegans (P49583), S. Cerevisiae (P13259), and P. Falciparum (P49587) were aligned using ClustalW. Residue numbers are the numbering for rat CTα. Boxed is the approximate region for domain M (residues 231 to 302). (*) indicates sequence identity. (:) indicates strong similarity. (.) indicate weak similarity.


R. Norvegicus
D. Melanogaster
C. elegans
S. Cerevisiae
P. Falciparum

1.4 CT catalysis

Due to a high degree of similarity in the primary structures between GCT and CT, GCT was used as a model to work out the mechanism of catalysis in this protein family. GCT exhibits a random order, rapid equilibrium reaction mechanism, with the dimeric enzyme essentially inactive until it is fully loaded with substrates[64]. This supports the concept that the enzymatic mechanism does not involve the enzyme as an intermediate nucleophile. Instead, direct attack of the phosphorylated alcohol substrate on CTP is involved (Figure 1.2). This random order of substrate binding was confirmed in purified recombinant rat CTα[65].

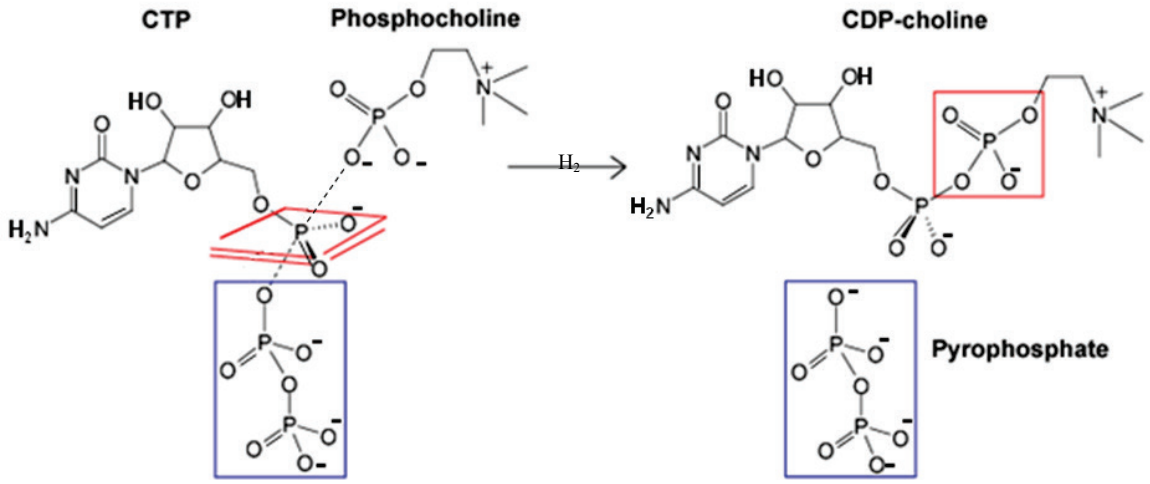


Figure 1.2 – CT catalyzed reaction

Formation of CDPcholine procedes via direct SN2 nucleophilic attack of the phosphate oxygen of phosphocholine on the alpha phosphorus of CTP, displacing diphosphate. The transition state (left) is a pentavalent trigonal bipyrimidyl geometry at the α-phosphate.

A few key sequence motifs critical for catalysis were identified for CT, based on mutagenesis and enzyme kinetic analysis. A defining element of the nucleotidyltransferase superfamily, HXGH, is found in both GCT and CT. Work with amino acyl tRNA synthetases, which are also members of the GCT family of nucleotidyl transferases, suggests that these His residues may be involved in the stabilization of a transition state[66-68]. Similarly, in GCT[69] and CT[70] these residues may also stabilize the pentavalent transition state formed by the phosphate oxygen of phosphocholine attacking the α phosphorous of CTP (Figure 1.2). In *B. subtilis* GCT and rat CT, mutations at

either His residues or glycine in the HXGH motif reduce k_{cat} and/or increase in K_m for CTP[65,69,70].

A second motif identified for the CT family is RTEGISTS, which forges extensive interactions with CTP in both GCT[71,72] and CT[73]. Mutation of R in this motif results in a large decrease in the specific activity[71,74]. Finally, very well conserved residues, Lys-44 and Lys-46 of GCT, and Lys-122 of CT, appeared to be involved in the binding/stabilization of phosphoglycerol in GCT[72], or phosphocholine in CT[73,74]. Mutations of these lysines result in a large increase in K_{cat} and K_m for the nucleophilic substrate, suggesting that these lysines may be responsible for aligning the substrates to attack the α phosphorous of CTP[69,71-73].

A recently solved structure of a mammalian CT (truncated version corresponding to the catalytic domain plus the N-terminal extension) revealed additional structural features in the specific recognition for phosphocholine (Figure 1.3) [73]. Hydrophobic side chains (Tyr-173, Val-181, and Tyr-182) from an extension in the L5 loop (absent in GCT) make van der Waals interactions with the choline methyl groups. Together with His-168, Tyr-173 (functions analogous to Lys-44 of GCT) coordinates the phosphate of CTP, and likely assists in the positioning for nucleophilic attack during catalysis. Confirmed by kinetic studies, His-168, buried deep in the active site cleft, is crucial for the binding of phosphocholine[73]. In essence, catalysis is accomplished by surrounding the negatively charged α-phosphate with electro-positive species contributed by several side chain and backbone nitrogens[73]. Based on structural changes of GCT crystallized with and without substrate[71,72,75], and the thermal B factors associated with elements of the solved CT structure[73], two major mobile elements were postulated. Specifically, the flap movements of two loops (L2 and L5) would function to allow access to the active site for the substrates (while excluding water), and efficient release of product during catalytic cycles[73].

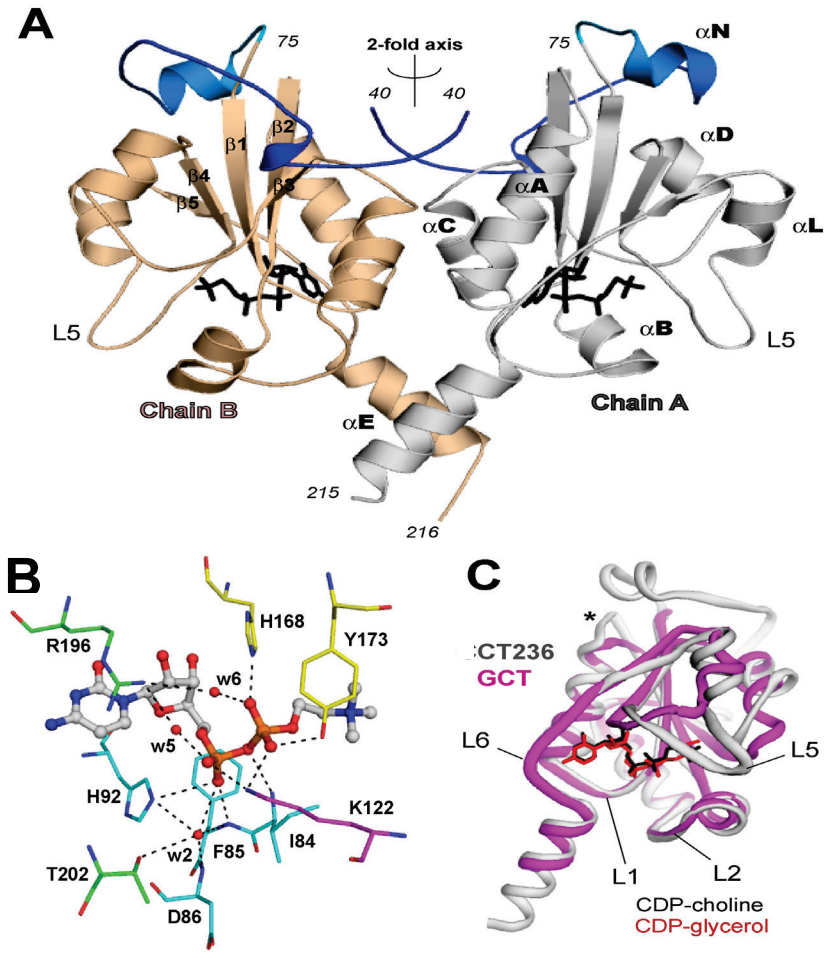


Figure 1.3 – The structure of the catalytic domain and active site of rat CTα
Images obtained from Lee et al (CTα PBD ID = 3HL4; GCT PBD ID = 1N1D)[73]. Panel A shows the homodimer, labeling the 5 helices flanking the β-sheet scaffold. CDP-choline is in stick representation (black). Panel B highlights some of the contacts between the α and β phosphates of CDP-choline and the CT active site: contacts from loop L-1, cyan; from L-2, purple; from L-5, yellow, and from L-6, green. In panel C, the loops are highlighted in an overlay of monomers of the solved structures of CT236 and GCT.

1.5 Regulation of CT by lipids and phosphorylation

Since CT is the key regulatory enzyme in the biosynthesis of PC, much effort has been focused on understanding the regulation of CT. From studies done on CT structure, sub-cellular location, and post-translational modification, several modes of regulation have been proposed.

1.5.1 Membrane lipid binding

CT's ability to bind to membrane lipids was first recognized when CT was consistently found in both the particulate and soluble fractions of disrupted cells, with the particulate fraction having much higher activity[76]. Activation through lipid binding was was proposed since the activity of the CT in the soluble fraction showed a significant increase upon addition of exogenous lipids. With these observations, an *in vivo* model, later termed amphitropism[77], emerged where the inactive soluble CT is activated upon translocation to the membrane in response to signals that connote a need for increased PC synthesis, such as a decrease in the ratio of PC to other phospholipids[78,79].

In support of this model, membrane translocation of CT accompanies treatment with exogenous fatty acids[80], phospholipase C (which generates DAG from PC catabolism), as well as addition of exogenous DAG[78]. As well, activation of DAG production via phorbol esters was correlated with an increase in membrane-associated CT[81]. Translocation of CT can also be caused by altered phospholipid compositions, as shown in cells starved for choline[82], which results in a decrease in the PC/PE ratio[78,79]. Therefore, products of PC catabolism (fatty acids and DAG), or a change in phosp3lipid head group composition can cause translocation of CT.

Lipid binding and activation has been characterized extensively in vitro with purified CT and small unilamellar vesicles[83-85]. Anionic lipids, such as phosphatidylglycerol, oleate, or phosphatidylinositol mixed with phosphatdylcholine are the most potent activators of CT. CT can also be effectively activated by mixed micelles of Triton-X-100 and anionic lipids[84]. These analyses demonstrated that CT binding and activation is dependent on the negative charge density of the micelles or vesicles[83,84]. In addition, DAG and other neutral lipids with small polar head groups, such as unsaturated phosphatidylethanolamine and oxidized PC, can activate CT, although to a lesser degree[84-87]. Lipids with small polar head groups tend to create surface defects and promote negative curvature strain in the lipid bilayer. CT binds to these defects in the membrane and relieves tension built up from the negative curvature[86]. The synergistic activation by anionic lipids and small head group lipids suggests CT is very sensitive to

the levels of these lipid signals in vivo[83]. As DAG and fatty acids are common lipid metabolites of PC in vivo, this synergy suggests a "feed-back" mechanism of PC synthesis, where the catabolites of PC could effectively recruit and activate CT in PC-deficient regions of the lipid membrane[40,88].

1.5.2 Phosphorylation

Post-translational modification is often utilized in the regulation of enzymes in cells, offering quick responses to keep up with the dynamic nature of cellular events. One of the most prevalent modifications is phosphorylation. In CTα, phosphorylation has been unambiguously localized to the 16 serines in the C-terminal tail[89]. Data from numerous studies has shown that a higher degree of phosphorylation is associated with the soluble, inactive form of CT, compared to the membrane bound form[48,82,90]. Moreover, phosphorylation decreases the membrane binding affinity of pure CT in vitro[89,91]. Phosphorylation, therefore, has been proposed as an antagonist of membrane binding and thus could play a role in the auto-inhibition and activation of CT.

Despite these convincing findings that correlate the state of phosphorylation with the partitioning of CT between soluble and membrane bound form, the physiological significance of this mode of regulation is not as clear. Wild-type, phopho-mimic, and mutant CT incapable of phosphorylation all displayed similar characteristics in terms of translocation induced by excess oleic acid, activation, and cell growth when either construct was expressed in a CHO58 cell line with a conditionally null endogenous CT[92]. Other studies demonstrated that the high density of negative charge from phosphorylation did not preclude membrane association[93], provided the lipid activator concentration was high. Thus the notion has emerged that phosphorylation serves to fine-tune the membrane affinity of CT, but the mechanism whereby it does so remains unsolved. Alternatively, phosphorylation may play a role in regulation of CT by other means. For example, posphorylation might be used to enhance the half-life of CT within the cell[92].

1.6 Domain Structure of CT

Four domains have been identified in CT, with approximate boundaries (Figure 1.4; see also Figure 2.2). **Domain N,** residues 1-74, houses the nuclear localization

signal at the N-terminus (residues 8-28)[49]. The solved structure of CTα, residues 1-236 (truncated form), revealed that the C-terminal half of domain N (residues 40 to 74) participates in dimer formation, and situates on top (opposite from active site) of the catalytic domain. The first 39 residues are not visible in the solved structure, presumably due to high flexibility[73]. Although this region may not be involved in stabilization of the dimer or CT catalysis, it may have additional role in membrane cross-bridging[94].

Figure 1.4 – Sub-division of CT sequence into four functional domains
The numbering is for rat CTα

Domain C (residues 74-230) is responsible for catalysis. It is highly conserved across species, and shares high sequence and structural similarities to other members of the nucleotidyltransferase family[71]. It contains an α/β nucleotide binding fold consisting of a twisted five-stranded parallel β-sheet flanked by five helices. The ligands (CTP and phosphocholine) bind to a pocket at the base of the β-sheet (see Figure 1.3)[71,73]. The active site cleft contains two arms: one arm, for binding CTP, extends from the top of helix E towards the base of the β-sheet; the other arm (for phosphocholine) extends along the base of the β sheet toward the L4 and L5 loops. The key components participating in catalysis locate to L1, L2, L5, and L6 and the N-terminus of Helix E[73].

Domain M (residues ~231-302) is the auto-regulatory domain. Deletion of this domain results in constitutive (lipid-independent) activity in rat CTα[95] and C. elegans CT[58]. It is worth noting that rat CTα lacking domain M and P shows a much higher K_m for the CTP substrate, as compared to the fully activated full length CT. This tentatively suggests that domain M might also assist in stabilization of the CTP arm of the active site in the active conformation.

In vertebrate CTs, domain M possesses an 11mer motif, repeated 4 times, with a consensus sequence of: **KSKEFVXKϕEE**. This motif uniquely resembles the 11mer motif (**KTKEGVϕXϕXX**) that is repeated 6 times in α-synuclein, the Parkinson's Disease associated protein (Lee, J., Semenec, L., Chen, J., Cornell, R. unpublished). This region in both CT[96] and α-synuclein[97] comprises a lipid-inducible amphipathic α-helix.

Compared to the catalytic domain, which shows > 60% similarity between human and yeast CTs, the sequence conservation in domain M is not high, and I will return to this issue in Chapter 3.

Numerous studies, each with a different approach, have shown that this domain is the principle membrane binding domain[58,96,98-102]. It responds to changes in membrane lipid composition *indirectly*, by sensing ensuing changes in membrane physical properties[77,86,103]. Upon encountering a suitable membrane surface, domain M is attracted by electrostatics and docks onto that surface. There it transitions from a mixture of conformers into a long amphipathic α-helix and inserts half way into one monolayer[96,101,104]. By an as yet unknown mechanism this conformational change in domain M communicates to the catalytic domain to relieve auto-inhibition. CD studies with domain M peptides, up to 62 amino acids in length, showed that they lack a stable structure in solution. In the presence of activating lipids, however, the peptides gain significant α-helical content[104]. Homonuclear 2D NMR analyses of two overlapping domain M peptides in the presence of the membrane mimetic, SDS, provided strong support that membrane-bound domain M forms a long continuous amphiphatic α-helix (Figure 1.5)[96]. In water, the peptide spectra did not provide evidence for structure. CD deconvolution analysis of full-length CT with and without activating lipids also indicated a change in helical content equivalent to conversion of ~60 residues (likely domain M) into a helical conformation[104]. None of these analyses provide detailed insight into the structure of domain M in its non-membrane bound form. This problem is crucial to understanding the mechanism of auto-inhibition by M. To date crystallization trials of CT constructs containing domain M have not been successful.

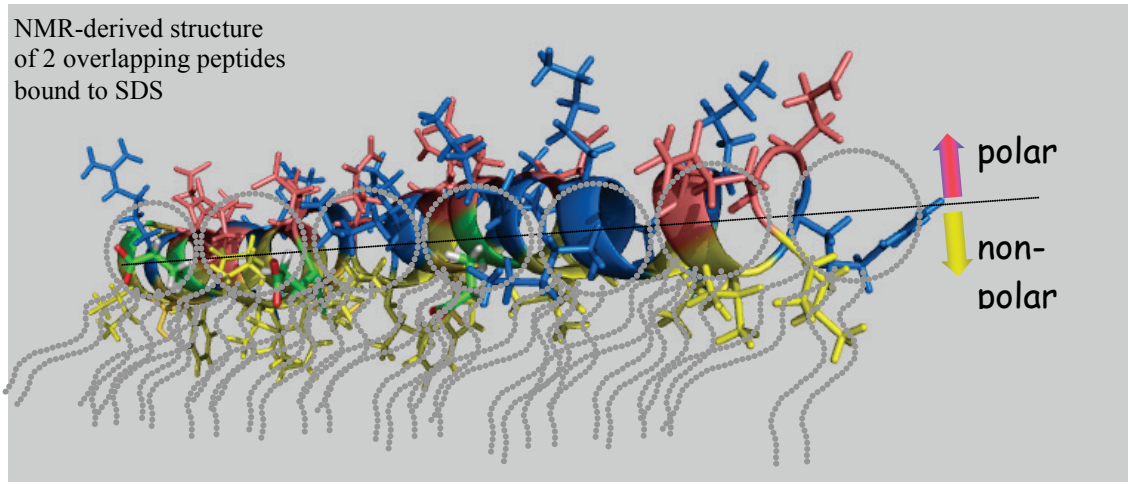


Figure 1.5 – NMR structure of SDS-bound (lipid-mimic) domain M peptide
Stucture of domain M peptides bound to SDS (peptide: SDS = 1: 40 M/M) solved as described[96]. The lipid monolayer is etched in by hand. yellow residues = hydrophobic; blue = basic; salmon = acidic or polar/nonionic. Three glutamates that become protonated at the surface of anionic membranes are shown as green with oxygens in red[103].

The15-aa linker sequence between domain M and domain P, which is not conserved, is of unknown structure and function. It may in fact be a portion of domain M, and participate in membrane binding and auto-inhibition.

Domain P (residues 315 to 367) contains 16 serines in CTα with the potential to be phosphorylated[89]. This region has very poor sequence conservation across orthologs, but tends to be around 50 residues in length, rich in proline, and is otherwise lacking in hydrophobic character. As discussed previously, binding analyses with phosphorylated and dephosphorylated forms implicate a role for modulation of lipid binding through phosphorylation. Comparative CD-analysis and disorder predictions indicate this domain is mostly unstructured and disordered[93,104]. This is supported by its high accessibility to protease digestion[100].

1.7 Proposed mechanisms for auto-inhibition of CT

As there is no solved structure for the soluble form of a CT containing domain M, how domain M silences the catalytic domain remains to be explored. The simplest mechanism to propose would be a direct contact between the catalytic domain and domain M, both possessing rigid, stable structures with high complementary surfaces for strong auto-inhibitory interaction (Figure 1.6A), as exemplified by other well-

characterized auto-inhibitory proteins discussed earlier. Some evidence for this rigid-contact inhibition model emerged from a recent study on C. elegans CT, where lipid independent activity increased significantly when specific hydrophobic residues in domain M were non-conservatively mutated. Individual substitutions of L-246, W-249, I-256, or I-257 to serine had no significant effect on lipid-independent activity, but the combination of all four substitutions to serine resulted in a > 10 fold increase[105]. Mutation of a highly conserved F-260 to another hydrophobic residue (I, V, or L) was without effect, but changing this one F to aspartate resulted in nearly complete lipid-independent activity. This data is compatible with a hydrophobic surface mediating inhibitory contact between domain M and C.

However, other data are unsupportive of stable domain M-domain C contacts: (1) Limited protease digestion analysis: If domain M docks onto domain C one should see an imprint of contact sites on domain C that would be at least partly protected from proteases. However, the proteolytic pattern of peptides from the catalytic domain was the same comparing constructs with or without domain M, as well as full-length CT +/- activating lipid vesicles[100]. (2) Although auto-inhibition by domain M is a regulatory feature of CTs from plasmodium[62] to mammals[84], domain M shows poor conservation in terms of sequence, length, and presence of the 11mer motif. Since the catalytic domain *is* highly similar, a lock and key silencing mechanism is not consistent with the divergent M domain.

An alternative mechanism of auto-inhibition (Figure 1.6B) was recently suggested that does not rely on M/C direct contact. Rather, a silencing mechanism was proposed that involves a highly disordered domain M in its lipid-free state. The disorder would be propagated through the M/C linker, through helix E to generate local disordering of the residues at the N-terminus of Helix E that participate in catalysis[73]. This high entropy status prevents their proper alignment for contact with CTP. According to this model, lipid binding of domain M would propagate order from the membrane-bound α-helical M domain all the way into helix E. This disorder-inhibition model is based on analysis of crystal structures of catalytic domains of ethanolamine cytidylyltransferase (ECT), GCT, and CT, showing disorder to order transitions for helix E[71,72,75] (which is immediately adjacent to active site residues) upon binding substrate or product. Although this model

escapes the requirement for specific auto-inhibitory contact, there has not been any direct experimental evidence to support it. Emerging from my thesis work is a third silencing mechanism that incorporates both direct contact and disorder in the silencing domain.

Details of auto-inhibitory mechanisms for some of the auto-inhibitory proteins have been elucidated successfully with specific structural data from NMR experiments[6,8] and crystallography[9,10]. Unfortunately, it has been very difficult to obtain proper crystal for structural determination of the full length CT with X-ray crystallography. As well, since CT is a stable homodimer in solution, with a combined molecular weight of 84 kDa, it exceeds the upper limit for structural determination by NMR. Therefore, other structural approaches must be taken for deciphering mechanism of CT auto-inhibition.

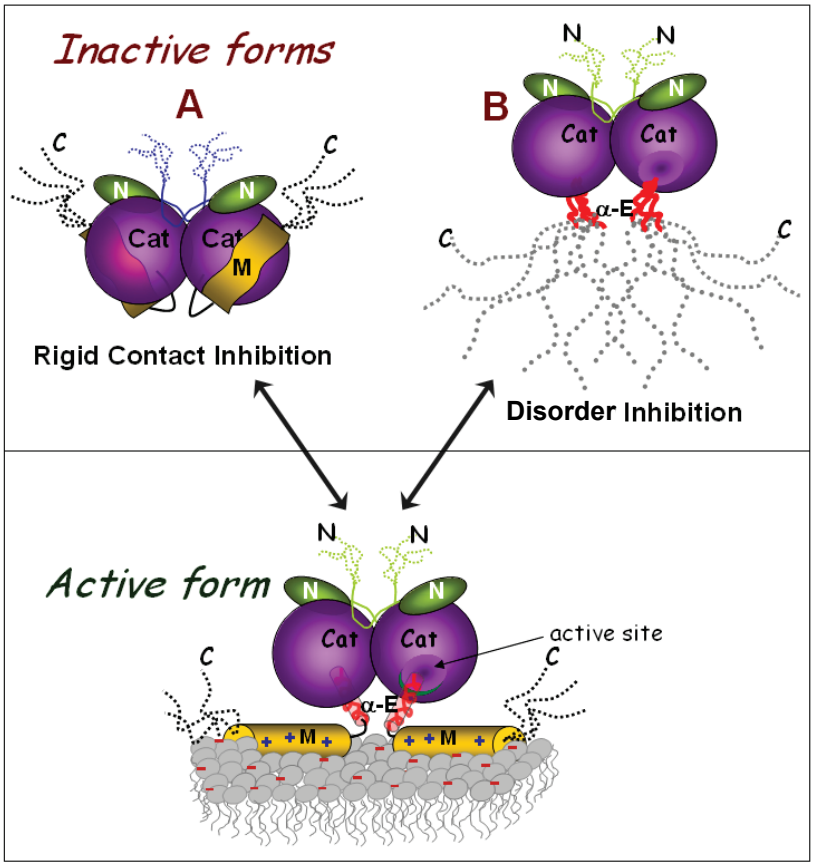


Figure 1.6 – Possible auto-inhibitory mechanisms proposed for CT

In my thesis work I have used two approaches: photo-cross-linking to map the inter-domain interactions of domain M, and fluorescence anisotropy to probe the structural dynamics of domain M in comparison to folded domain C and the floppy C-terminus of the protein. As proposed in the Disorder-inhibition model, the flexibility of domain M may be important in auto-inhibition. Anisotropy analysis can provide a lower resolution alternative to NMR for exploring structural dynamics[106]. Details of the theory and application for these two techniques are introduced in the next two sections.

1.8 Photo-cross-linking

Photo-cross-linking is a useful tool for analyzing molecular interactions, including protein-protein, protein-ligand, and domain-domain interactions[107-110]. It is used in combination with other analytical tools such as proteolysis and mass spectrometry.

To be useful for structural applications, photo-cross-linkers must meet the following requirements[111]: 1) attachment of the photo-cross-linker must not disrupt the folding of protein; 2) the optimal conditions for photo-cross-linker activation must not damage the protein's structure and function; 3) the lifetime of the excited state of the photo-cross-linker should be shorter than the lifetime of specific interaction being probed (eg. interaction between auto-inhibitory domain and functional domain); 4) the photo-cross-linker should react non-specifically with any neighboring group so that interaction sites are not biased; 5) the probe must be available commercially or convenient to synthesize[111]. Compounds that generate carbenes and nitrenes on irradiation, benzophenone derivatives that produce reactive radicals, and aryldiazonium generating cations best meet most of these requirements[112-114]. Of these photophores, a benzophenone derivative, BBP (Figure 1.7), was chosen for my project, since it was commercially available at the beginning of this project, and had been used successfully to map protein-protein interactions[115]. BBP also has several other useful features, described below.

A simplified version of the benzophenone reaction is outlined in Figure 1.7. One major advantage benzophenone has over other photo-cross-linkers is its unique, *reversible* photo-activation at non-damaging wavelength (~350 nm). The benzophenone

will relax electronically if it fails to form a covalent bond with its target within the lifetime of its excited state (which can persist up to ~100 μsec) [114], but may still maintain its binding and photo-activation potential. As a result, it can undergo many excitation-relaxation cycles until a favorable geometry for covalent modification is achieved. The probability of finding a favorable geometry for the duration of photo-irradiation may be increased further with the use of a flexible linker to the site of protein attachment[114].

A common approach in the investigation of inter-molecular or inter-domain interaction with a photo-cross-linker involves attaching the photo-reactive cross-linker onto a specific site in the molecule of interest, irradiating to initiate photo-cross-linking, and analyzing photo-cross-linked species via gel electrophoresis and/or mass spectrometry (MS)[107,110,116,117]. The quality of MS spectra can be improved by purifying photo-labeled molecules or parts of the molecules. BBP is suitable for these tasks, as it contains sulhydryl-reactive methanethiosulfonate for site-specific attachment to engineered cysteines, a flexible linker for efficient photo-cross-linking, and a biotin group for purification via avidin affinity chromatography.

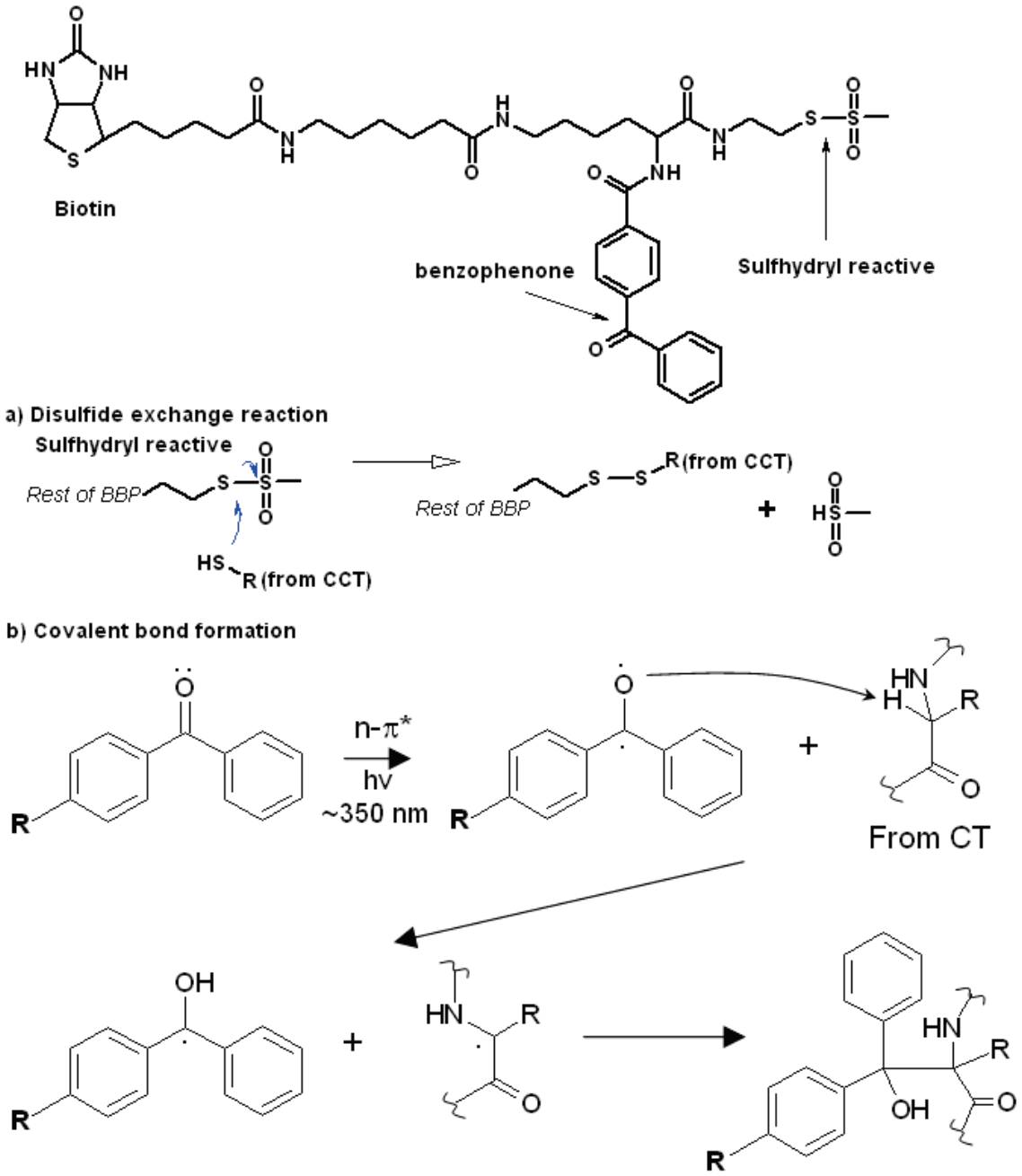


Figure 1.7 – Reaction of a biotinylated sulfhydryl reactive photo-cross-linking reagent, BBP (2-[Nα-Benzoylbenzoicamido-N^6-(6-biotinamidocaproyl)-L-lysinylamido]ethyl methanethiosulfonate)
a) The sulfhydryl reactive group on BBP reacts with the SH of a single cysteine in CT to form a stable disulfide. b) UV irradiation at ~350 nm causes promotion of one electron from a nonbonding sp2-like n-orbital on oxygen to an antibonding π*orbital of the carbonyl group. The electron-deficient oxygen then abstracts H from weak C-H bonds, giving rise to the ketyl and alkyl radicals that quickly recombine to generate new C-C bond.

1.9 Fluorescence Anisotropy

Fluorescence is the emission of a photon upon the return of an electron from a higher energy orbital (excited state) to a lower orbital (ground state). Since there is a significant time delay between absorption of a photon (~ 10 fsec)[118] and fluorescence emission (~ 10 nsec)[119], a fluorophore can undergo significant positional displacement between absorption and emission.

A light source usually emits electromagnetic waves as a collection of all orientations (unpolarized). It can be preferentially aligned in one plane (polarized). The probability of absorption is highest when the electronic transition moment of the fluorophore is parallel or at a very small angle to the plane of electric oscillations of the exciting light beam[119]. In an isotropic solution of randomly orientated fluorophores (Figure 1.8A), the correctly orientated fluorophores will be preferentially excited (Figure 1.8B). In a hypothetical rigid medium, rotational motion of excited fluorophores will be very low (Figure 1.8C), resulting in retention of polarization of the emitted light. On the other hand, in a very fluid medium an unconstrained fluorophore will rotate in the interval between absorption and emission; thus the emission will be depolarized. (Figure 1.8D)[118]. The degree of depolarization of the emitted light can be monitored by detectors oriented to capture light parallel or perpendicular to the plane of the incident light.

There are two types of fluorescence anisotropy measurements. A steady-state measurement involves continuous excitation and is an average of the time-resolved phenomena over the course of irradiation. It is simple to perform, and information derived from this type of measurement can be straightforward. On the other hand, a time-resolved decay measurement involves a sub-nsec excitation pulse, followed by subsequent monitoring of the decay of polarization of the emission. Time-resolved anisotropy decay monitors dynamics of the sample on a much finer scale, and can offer a wealth of information not obtainable by steady-state measurements. However, this type of measurement requires a fluorimeter capable of measuring intensities on the nsec time scale[119], which was unfortunately not available for this project. Therefore, I will focus on factors involved in steady-state measurements.

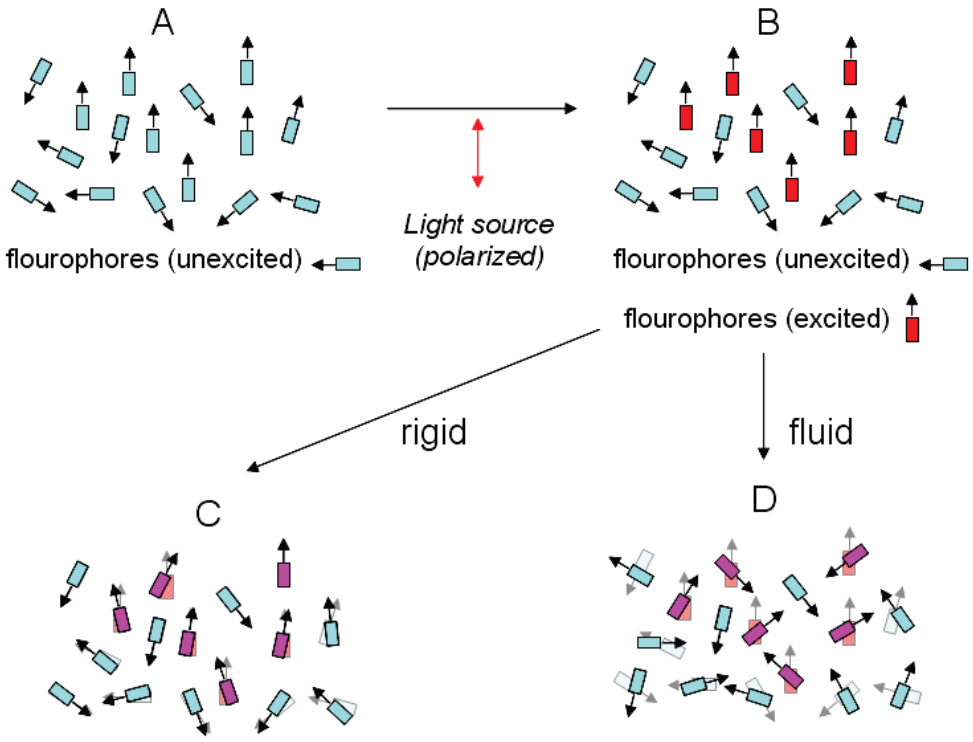


Figure 1.8 - Fluorophore excitation and emission has orientational dependence
Emission becomes depolarized with fluorophore rotation

The equation used to define steady state anisotropy (r) is $r = (I_{II} - I_{\perp}) / (I_{II} + 2I_{\perp})$, where I_{II} is the fluorescence intensity from vertically polarized emission, and $I_{\perp}$ is from horizontally polarized emission. The theoretical maxima for anisotropy is between -0.2 (isotropic) to 0.4 (frozen)[119]. For comparative purposes, it is more informative to express freedom of rotation of the fluorophore with a rotational correlation time. Rotational correlation time (θ) is most accurately obtained from a time-resolved anisotropy decay, but can be approximated using the Perrin equation: $r = r_o/(1+\tau/\theta)$, where r_o is the limiting (i.e. maximal) anisotropy, and τ is fluorophore life-time (average time the molecule spends in the excited state prior to return to the ground state). Most unbound fluorophores rotate at 50 – 100 psec scale, where their life-times are in the 1-10 nsec range, resulting in extremely low anisotropy. Conjugation of a fluorophore to a protein/macromolecule greatly increases anisotropy by slowing down the probe's rate of rotational motion. As rotational correlation times for most proteins are comparable to typical fluorescence life-times (eg. 25 kDa protein has ~ 10 nsec rotational

time),measurement of anisotropy will be sensitive to any factor that affects the rate of rotational diffusion. As defined by the equation: $\theta = (\eta V)/(RT)$, these factors are viscosity of the solvent (η), volume of the rotating molecule (V), and temperature (T) the measurement is taken.

Other non-trivial factors unrelated to rotational freedom may affect the anisotropy values, and should be minimized if not avoidable. First, light scattering from molecules in the analyzed sample can greatly interfere with the measurement of anisotropy. For example, turbidity from aqueous suspensions of membranes in biological samples can cause scattering of the incident polarized light, resulting in excitation of fluorophores in many different orientations. Scattering can also originate from the emitted photons. Both of these scattering sources can under-estimate the actual anisotropy value. Fortunately, scattering can be detected by sample dilution. If dilution does not change the anisotropy value, then significant depolarization from scattering is unlikely. The other side-cause of depolarization is radiative transfer, which is the reabsorption of emitted photons. This process may increase apparent life-time of excited state as emitted photons can be reabsorbed and re-emitted several times before leaving the sample[120]. This is often the cause for low anisotropy values for the fluorophores that have large spectral overlap between excitation and emission wavelengths, such as fluoresceins and its derivatives. In fact, a single radiative transfer step can reduce the anisotropy to 28% of its initial value[120]. As this process is not easily avoided as with scattering, care must be taken in the selection of fluorophore, and in maintaining a dilute fluorophore concentration.

Anisotropy can also be influenced by resonance energy transfer (RET) between fluorophores, which can result in additional angular displacement of emission. In fact, a single RET step can reduce anisotropy to as much as 4% of the initial value[120]. This factor is addressed further in Chapter 3.

Despite potential complications from scattering, radiative transfers, and RET, many studies have successfully utilized anisotropy measurements to assess mobility and complex formation of macromolecules. For example, mobility and stabilization (in response to ligands) of an important helix in nuclear receptors was compared between wild-type and defective mutants with a fluorescein derivative, using both steady-state and

time-resolved anisotropy[121]. Formation of a ternary SNARE complex upon ligand binding was followed by observing the increase in fluorescence anisotropy of a fluorophore tag[122]. Through proper controls and careful experimental design in the steady-state anisotropy measurements, I have gathered useful data on the dynamics of CT domains that may be involved in the auto-regulation of CT.

1.10 Objectives

My major goal has been to elucidate the mechanism of CT auto-regulation by domain M. There is an on-going effort in the Cornell lab to solve the structure of CT in its active and inactive forms, and elucidate dynamic changes at the secondary, tertiary, and quarternary levels.

Some specific questions I have asked are:

1) With what sites in CT does domain M make contact in its inactive, soluble and active, membrane-bound forms?
2) What are the structural elements in domain M that enable it to efficiently silence the catalytic domain? How flexible is domain M?
3) What are the structural transformations required for relief of auto-inhibition during membrane binding and CT activation?

I have addressed these questions with the research presented in this thesis. In particular, I used photo-cross-linking, coupled with MALDI-MS, in an attempt to capture lipid-sensitive, auto-inhibitory interactions between the regulatory domain M and the catalytic domain. In addition, I investigated the domain dynamics of the regulatory domain M with the use of steady-state anisotropy, and gained insights into the inhibitory behaviour and dynamics of domain M in the context of full length CT.

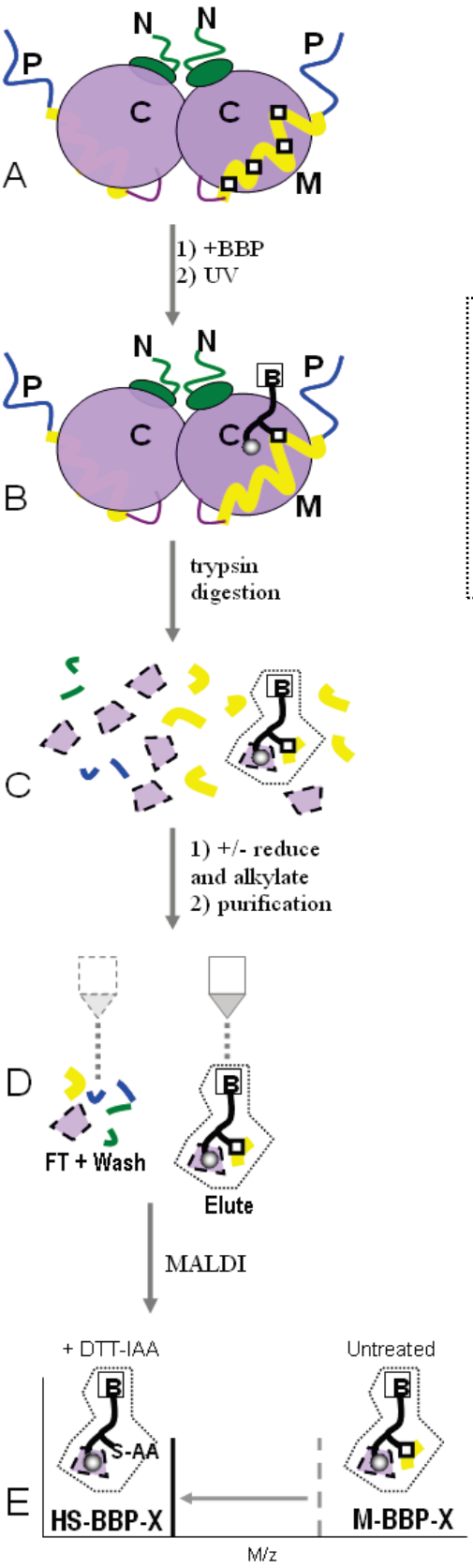
2: Identification of lipid-sensitive interactions between domain M and other CT domains via photo-cross-linking/mass spectrometry

2.1 Introduction and summary of Chapter 2

As discussed in Chapter 1, the mechanism by which the M domain suppresses catalytic activity is unknown. The prevailing model at the time I began this work was that the M domain makes contact with the catalytic domain. Examination of the proteolysis pattern for full-length *vs* CT truncated before domain M (CT236) had suggested a possible interaction between domains M and N rather than M and C. Specifically there was no apparent foot print of M (protection from cleavage) on the catalytic domain when comparing the soluble and membrane bound forms[100]. To identify contacts between domain M and any other portions of CT, I engineered single cysteines at 7 sites along the length of domain M using a cysteine-free rat CTα. After purification of each single cysteine variant, I worked out methods for cysteine-specific labeling with BBP, benzophenone-mediated cross-linking, and detection of these modifications. This method is an unbiased search for any contacts within any region of CT with the engineered cysteine site that are sufficiently stable over the ~100 μsec lifetime of the activated benzophenone.

In this chapter I present the development of these protocols. I showed specificity of BBP labeling (by blocking with sulfhydryl-targeted alkylation), efficient removal of unconjugated BBP from labeled samples, and sufficient photo-cross-linking (formation of reductant-resistant covalent linkage via benzophenone). After complete proteolysis of the CTs and isolation of biotinylated peptides by avidin affinity purification, the peptides cross-linked to the domain M peptide containing the BBP adduct were identified by MALDI-MS. (See Figure 2.1 for a flow chart of this protocol). I present my method for authenticating these peptides as contact points with domain M, which involved parallel

analysis of peptides generated with 4 distinct conditions (BBP-labeled, BBP-photo-labeled; each +/- reduction and acetamidylation of the sulfydryl). I found that each of four sites in domain M generated a similar photo-cross-linking pattern, involving contacts throughout other portions of domain M, P and C. The contact sites in domain C were mapped to peptides flanking the active site of the enzyme, and were absent when CT was bound to lipid vesicles, in keeping with a function in silencing the enzyme. Mutations in a conserved motif near in the C-terminus of domain M eliminated contacts to domain C but did not relieve auto-inhibition of CT, suggesting that other sites participate in the silencing function. These results tentatively support a model where flexible M domains make multiple alternating contacts with a defined region on the catalytic domain. I hypothesize that these contacts mediate auto-inhibition.



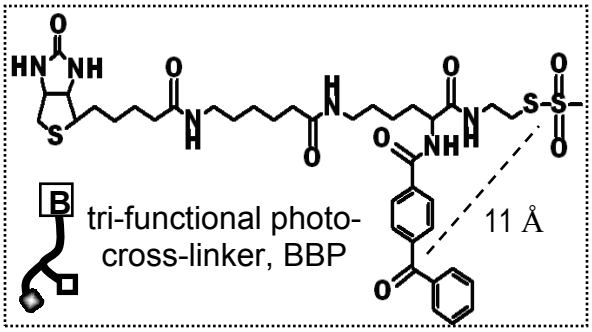


Figure 2.1 - Outline of protocol identifying inter-domain contacts with domain M

A. Single cysteine variants at various sites within domain M (▣) were purified and labeled with BBP selectively at the engineered cysteine by disulfide exchange. **B.** BBP-labeled CT was irradiated (366 nm) for 1 h at 4°C. **C.** Photo-cross-linked samples were digested with trypsin. **D and E.** Biotin-tagged peptides from half of the digested sample were purified by avidin affinity chromatography without pre-treatment, and the other half of the digested sample was reduced and alkylated (DTT-IAA) prior to avidin chromatography. Peptides were identified by MALDI-MS. The mass shift in the spectra shown in **E** is equal to the mass of the domain M peptide (with modified cysteine) minus acetamide (56 Da). **Inset** displays the tri-functional photo-cross-linker, BBP, which has a ~ 11Å distance separation between the sulfhydryl-reactive methanethiosulphonate and the photo-reactive benzophenone moiety.

2.2 Results

2.2.1 All single cysteine CT variants retained functional activity

Wild-type rat CTα contains seven endogenous cysteines. These were mutated to serines by Mingtang Xie with only a ~ 2-fold reduction in enzyme activity[123]. Beginning with this cysteineless CT, I engineered eight single cysteine CT variants along the length of domain M (Figure 2.2).

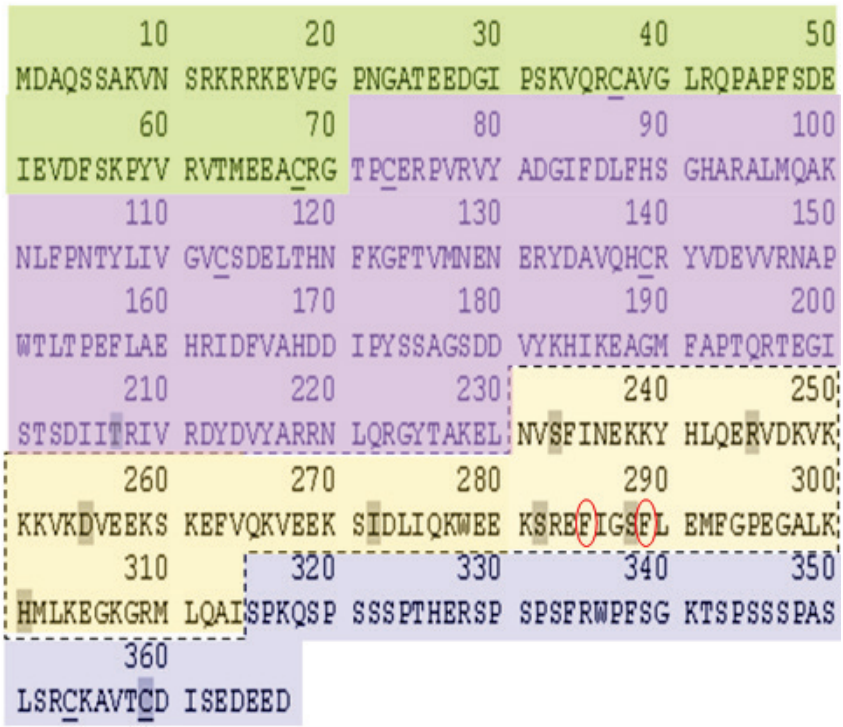


Figure 2.2 – Primary sequence of CT with sites of engineered single cysteines

Single cysteine CT variants were generated by first mutating all cysteines (underlined) to serines, followed by site-directed mutagenesis to replace selected residues to cysteines (gray high-lighted). The domains of CT are shaded as follows: domain N, green; domain C, violet; domain M, yellow; domain P, blue. The exact domain boundaries for each domain are unknown. The 289F and 293F that were mutated to aspartates in section 2.2.7 are in red ovals.

The cysteine substitutions were aimed to be non-perturbing, thus I substituted at serine residues or at sites that are poorly conserved (see Figure 1.1). Although all CT variants were slightly less active than the wild-type CT, their activities were slightly

above or in the range of the cysteine-free CT. Notably, all remained highly lipid-regulated (Table 2.1). This indicates that the overall protein structure and inter-domain interaction responsible for lipid regulation and enzyme activity was not affected by the mutations made.

Table 2.1 - Specific activities of CT constructs
Wild-type (WT) and single cysteine variants of CT were assayed for activity as described in Methods, with 0.1 mM PC/PG sonicated vesicles. All values are means ± SEM (N=3), except for WT, which was an average of duplicates.

	WT	**S233C**	**R245C**	**D255C**	**I272C**	**S282C**	**S288C**	**H301C**	**C359**
PC/PG	9650 ± 260	5910 ± 210	7320 ± 250	4840 ± 130	6810 ± 330	2690 ± 45	6920 ± 300	7610 ± 100	7170 ± 360
No lipid	188 ± 23	242 ± 13	401 ± 83	127 ± 27	73 ± 29	235 ± 17	189 ± 34	186 ± 27	178 ± 58

2.2.2 M domains can make contact with each other in solution

Previous analysis of the dimer interface of CT had indicated inter-subunit contacts involving domain C and its N-terminal extension[73,123]. Interestingly, all single cysteine domain M variants can form disulfide bridged dimers upon oxidation (Figure 2.3A). To test whether the disulfide bridges were due to random collisions between dimers I prepared a heterodimer of CT236/CT-S288C which has only one domain M. This heterodimer did not undergo disulfide formation upon oxidation (Figure 2.3B), suggesting that the disulfide bridges between M domains were formed within the same homodimer. This result reveals that either the residues in the M domains are close to each other in a stable complex, or that there is sufficient flexibility along the length of domain M to allow brief inter-chain contact that could be trapped by disulfide bonding.

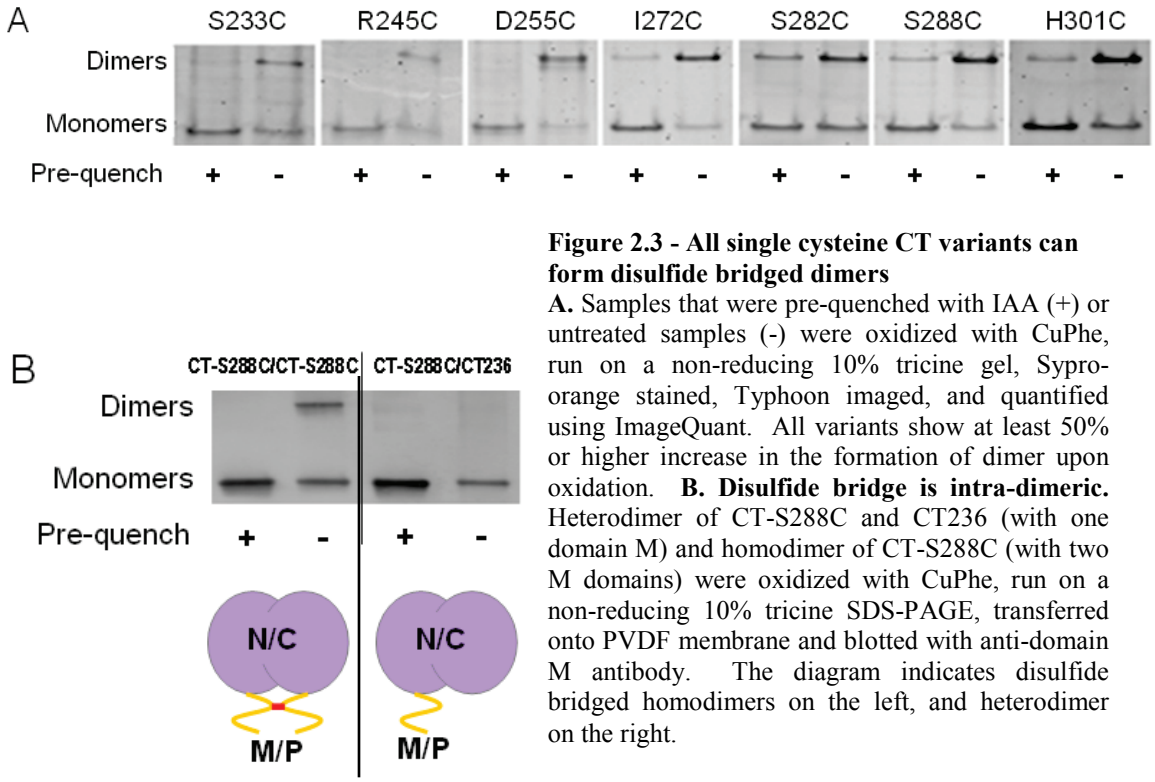


Figure 2.3 - All single cysteine CT variants can form disulfide bridged dimers
A. Samples that were pre-quenched with IAA (+) or untreated samples (-) were oxidized with CuPhe, run on a non-reducing 10% tricine gel, Sypro-orange stained, Typhoon imaged, and quantified using ImageQuant. All variants show at least 50% or higher increase in the formation of dimer upon oxidation. **B. Disulfide bridge is intra-dimeric.** Heterodimer of CT-S288C and CT236 (with one domain M) and homodimer of CT-S288C (with two M domains) were oxidized with CuPhe, run on a non-reducing 10% tricine SDS-PAGE, transferred onto PVDF membrane and blotted with anti-domain M antibody. The diagram indicates disulfide bridged homodimers on the left, and heterodimer on the right.

2.2.3 Assessment of the specificity and efficiency of labeling and photo-labeling with BBP

In developing my photo-labeling protocol it was important to demonstrate that the BBP was faithfully conjugated to the specific cysteine in domain M, that excess unconjugated BBP was removed prior to photolysis, and that the BBP linkage after photo-irradiation became resistant to disulfide reduction (indicating coupling via the benzophenone moiety). Taking advantage of the N-terminal His-tag on each single cysteine CT variant, I first immobilized the CT on Ni-agarose beads, exposed the CT to 100 molar excess BBP in DMSO for < 5 min, immediately removed excess non-conjugated BBP by multiple washes with DMSO and with buffer, and then recovered labeled CT from the beads with imidazole washes.

I addressed the efficiency and specificity of BBP-labeling and removal of non-conjugated BBP via electrophoretic analysis (Figure 2.4). Conjugation of BBP onto cysteines in the CT sample greatly reduced the availability of these cysteines to form

disulfide bridges under non-reducing conditions (resulting in a decreased proportion of disulfide-bridged dimer as shown in Figure 2.4A lanes 1 and 2). This indicates that I had labeled the CT with high yield. The specificity of BBP labeling at cysteine was demonstrated by the ability to block BBP-conjugation through alkylation of cysteines with iodoacetamide (IAA) (Figure 2.4B). Furthermore, I removed most, if not all, of non-conjugated BBP prior to photo-irradiation (Figure 2.4C). Importantly, I qualitatively confirmed that photo-cross-linking was efficient, as most of the BBP on photo-labeled CT had formed β-mercaptoethanol (β-ME) resistant, covalent bonds with the protein via the benzophenone moiety (Figure 2.4A lane 8 and 11). The analysis presented here showing results with CT-H301C has been performed for all CT variants subjected to photo-cross-linking analysis in this study, and yielded similar results.

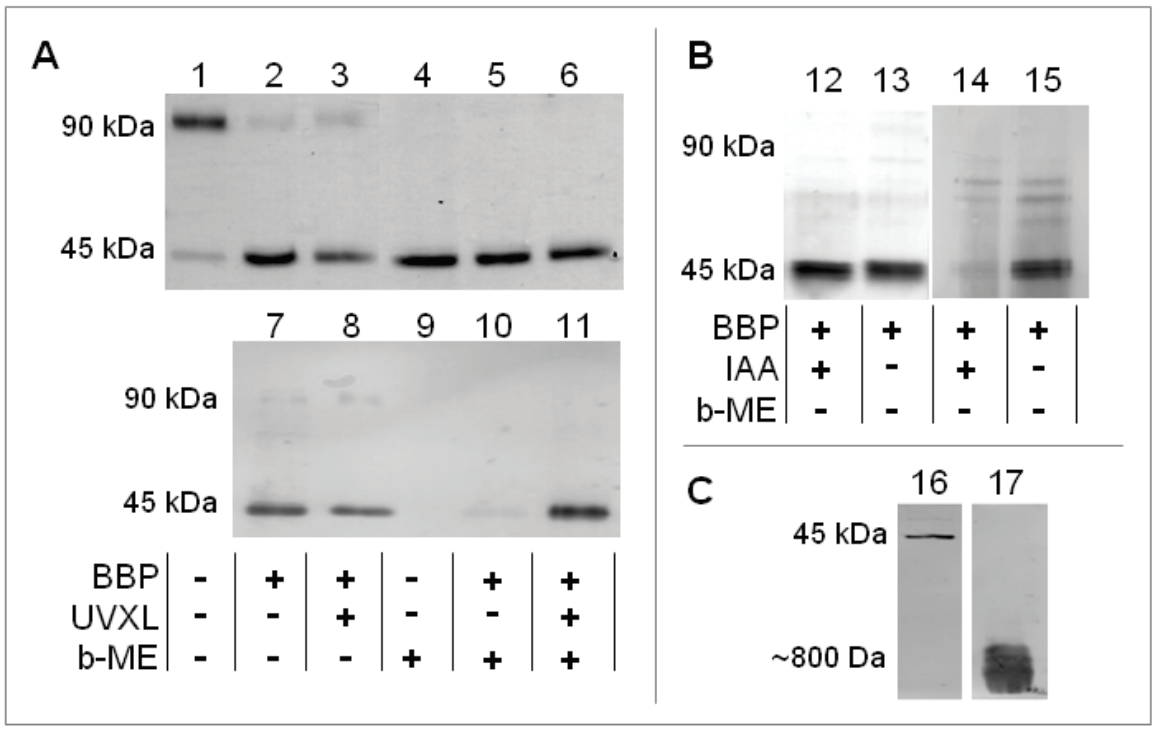


Figure 2.4 - Specific labeling of CT variants exemplified with CT-H301C

A. Assessment of labeling and photo-labeling efficiency. Top panel: Sypro-Orange stained 10% tricine-SDS-PAGE of unlabeled, BBP-labeled, and photo-labeled CT-H301C. **Bottom panel:** The same samples as shown in the top gel were processed for detection of biotin-containing BBP with Streptavidin HRP. **B. BBP labeling is disulfide dependent**. Silver-stained SDS-PAGE (lane 12 and 13) and Streptavidin HRP (lane 14 and 15) for pre-quenched (with IAA) versus unquenched, BBP-labeled CT-H301C. **C. Washing protocol removes excess free BBP**. Streptavidin HRP blot of BBP-labeled CT-H301C after labeling, washing 6 times, and elution from Nickel-agarose (lane 16). The unmodified BBP (input) is shown in lane 17. Samples were run on 16.5% tricine SDS-PAGE.

To assess whether this photo-labeling protocol affected the native structure of CT, I evaluated CT's tertiary fold by limited chymotrypsin digestion, and I measured its activity in the absence and presence of anionic lipid vesicles. With CT-H301C as an example (Figure 2.5), the proteolysis pattern was identical before and after the labeling procedure, and the specific activity was largely unaffected. This analysis assured us that the cross-links to be identified would reflect interactions of the native enzyme.

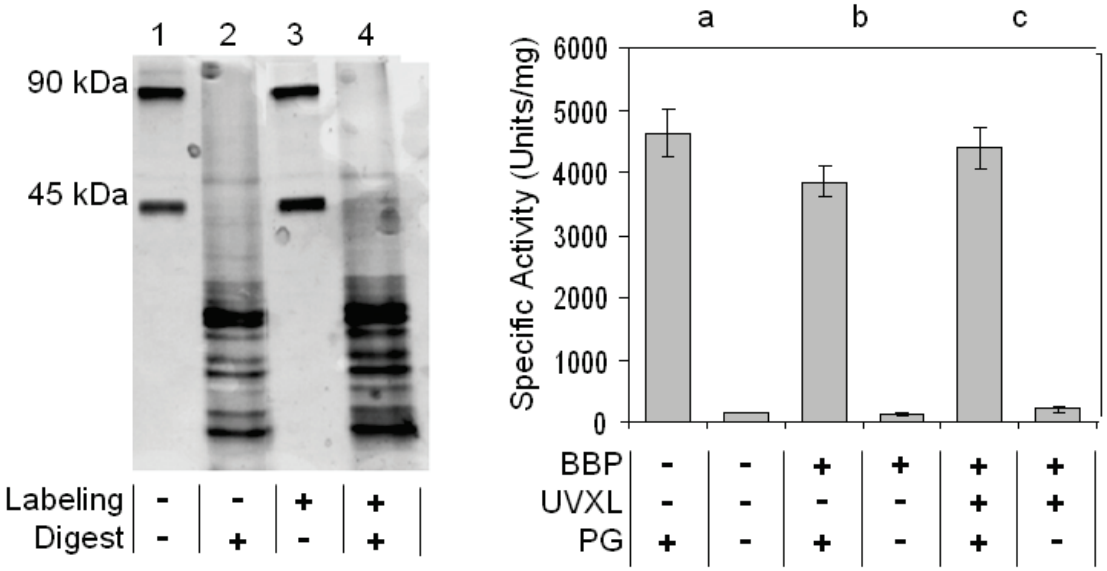


Figure 2.5 - CT structure and activity was preserved after the process of BBP-labeling and photo-cross-linking

Left: Silver-stained, 12% non-reducing SDS-PAGE showing the limited chymotrypsin digestion patterns of untreated CT-H301C (lane 1 and 2) and CT-H301C that had gone through the BBP-labeling procedure (lane 3 and 4). For digestion, CT-H301C was incubated with 1:200 (w/w) ratio of chymotrypsin:CT for 5 min at 37^{o}C with shaking.
Right: Activity analysis of untreated (a), BBP-labeled (b), and photo-labeled (c) CT-H301C. Activity assays contain 10 mM DTT. Specific activity (nmol CDPcholine formed /min/mg protein) were means +/- range of two independent determinations for a, or means +/- S.E.M of 4 independent determinations for b and c.

2.2.4 Photocross-linking analysis of four single cysteine sites indicates that all sub-regions of domain M make similar contacts with multiple sites in domains N, C, M, and P.

Although we generated seven single cysteine variants for photo-cross-linking analysis, three variants were not analyzed by photo-cross-linking. CT-S282C did not incorporate BBP label. The reason for this is not clear, since there is no implication from structural predictions that this region is particularly buried, and a protease accessibility

analysis showed relatively high accessibility[100]. CT-S233C and CT-D255C were unusually insoluble when expressed, and contained much higher impurities relative to other variants upon purification. Thus, my photo-cross-linking analysis focused on identifying photo-cross-links from cysteine sites at position 245, 272, 288, and 301, which span most of domain M.

2.2.4.1 Identification of photo-cross-links from MS spectra

To identify photo-cross-linked sites, I reduced and alkylated half of the trypsin digested, photo-labeled sample prior to affinity purification on avidin beads so that I could compare the mass spectrum of unreduced, cross-linked species with that of reduced, alkylated species derived from the parent photo-cross-linked peptide. This treatment produces identifiable mass shifts only for the peptides that contained a BBP adduct through photo-cross-linking. To be certified as a peptide that was cross-linked during photo-labeling, a peak matched to that peptide must satisfy the following criteria: 1) be unique to the spectrum for the photo-cross-linked sample (ie. if that peak were also found in spectra of non-UV irradiated sample, it would be discarded); 2) show a mass shift between reduced and non-reduced spectra equal to the mass of the domain M peptide with modified cysteine minus acetamide; 3) match its theoretical mass within 1 Da.

The certification process is illustrated in the set of spectra from an analysis of CT-H301C, shown in Figure 2.6. The peak in panel C at 3116.8 Da (yellow), unique to the photo-labeled sample, matches the mass of a peptide containing residues 301-304 connected to residues 284-300 through BBP ($M_{[301-304]}$-BBP - $M_{[284-300]}$). After reduction of the disulfide and alkylation this species generates two peptides: a 2683.2 Da peak (shown in the bottom panel, yellow) corresponding to residues 284-300 with covalently attached, acetamide-modified BBP (AA-BBP-$M_{[284-300]}$) and acetamidated 301-304, which disappeared as it no longer carried BBP and had been removed during avidin-affinity purification. Also illustrated in this figure are two species unique to the photo-labeled samples (panel C) representing cross-links between $M_{[301-304]}$ and domain C peptides, $C_{[79-94]}$ or $C_{[187-196]}$ (purple). The species at 3050 and 2353 Da in the 3rd panel represent the cross-linked species, which upon reduction and alkylation, produce the

BBP-conjugated and acetamidated domain C peptides at 2616 and 1918 Da, respectively, shown in the bottom panel. Some peaks in the spectra for photolysed samples are discounted, since they are also observed in the non-photolysed sample (blue dashed lines in Figure 2.6)

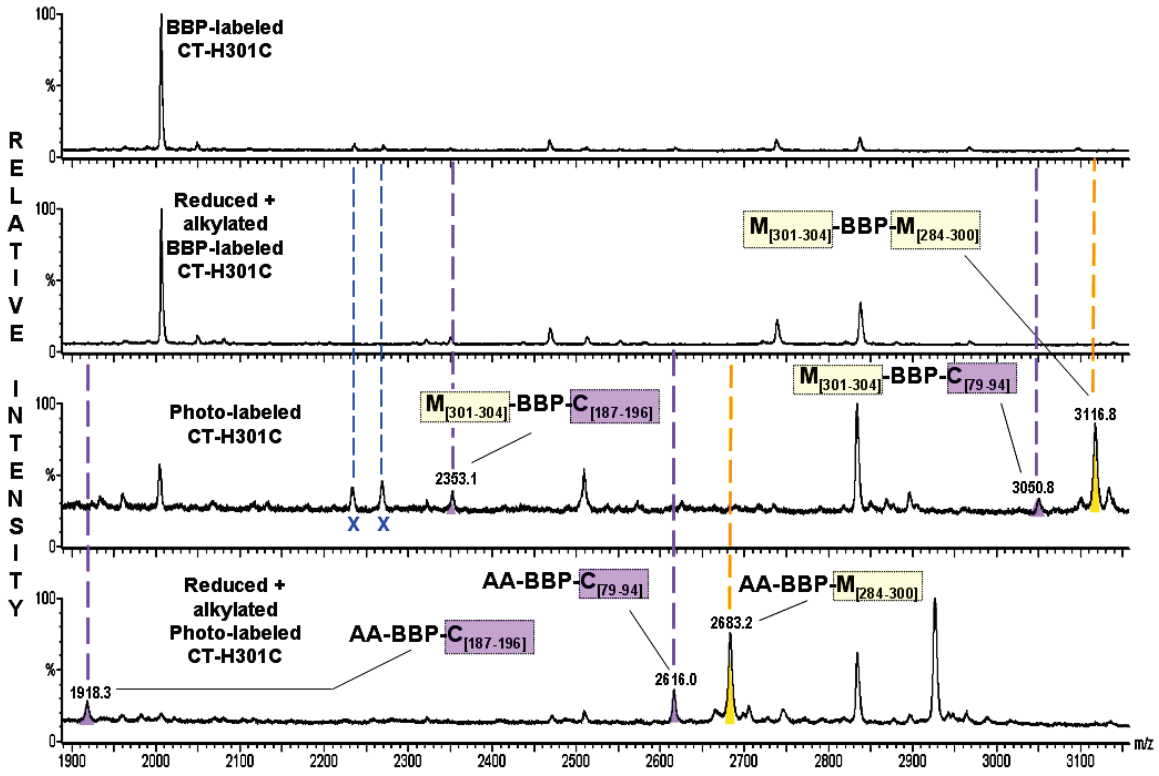


Figure 2.6 - Method for assigning bona-fide cross-linked peptides, illustrated with a set of spectra of peptides from CT-H301C

CT-H301C was BBP-labeled. Half of the sample was UV irradiated as described in Methods. All samples were digested with trypsin. Half of both non-UV irradiated and UV irradiated samples were reduced and alkylated. All samples were purified with monomeric avidin beads, and analyzed on MALDI-MS. Blue dashed lines represent examples of peaks that were not unique to photo-labeled sample spectra, even though they showed mass shift comparing spectra for untreated and reduced samples. These were not included in the final selection of peaks for identification. Violet and gold dashed lines indicate peaks that were specific to photo-irradiated samples, and that shifted mass upon reduction and alkylation, revealing the species cross-linked to the domain M C301 site.

Figure 2.7 shows the spectrum of photo-cross-linked CT-H301C derived peptides analyzed before reduction/alkylation. The most intense peaks matched to photo-cross-links between the modified cysteine and residues of immediate proximity (denoted self-cross-links), such as $M_{[301-304]}$-BBP-$M_{[284-300]}$ and AA-BBP-$M_{[284-300]}$. The high probability of these self-cross-links is expected since this is a nearest neighbour analysis. Nevertheless, photo-cross-links to other parts of CT, such as domain P ($M_{[301-307]}$-BBP-

$P_{[356-367]}$) and C (AA-BBP-$C_{[79-94]}$), were identified. Although they were minor compared to the intensity of the peaks corresponding to self-cross-links, they were very consistent across repeated trials. The same trend (strong self-cross-links, and weak cross-links to other domains) was seen with all single cysteine variants analyzed (Appendix I). A complete list of photo-cross-links identified for each CT can be found in Table 2.2. The percent of total MALDI peaks unique to the photo-cross-linked spectra that were assigned to CT tryptic peptides is also listed in Table 2.2, and was > 80%. Some spectra generated a few peaks that matched within 1 Da to two or more CT tryptic peptides. These species are listed in Appendix II.

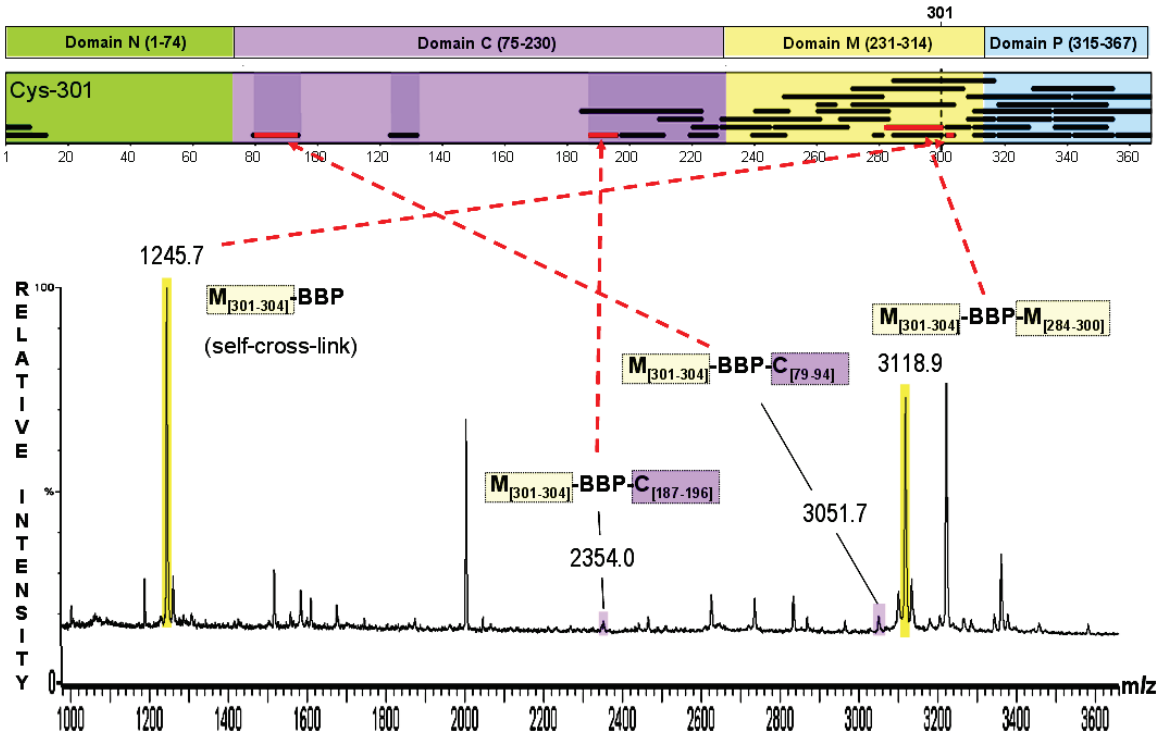


Figure 2.7 - Certified photo-cross-linked peptides were used to generate a photo-cross-link coverage map

Using the spectrum for CT-H301C as an example, photo-cross-linked peptides, identified by the method shown in Figure 2.6, were plotted as bars representing their position and length in CT. The identities of four cross-linked peaks are shown. The peak at 1245.7 Da corresponding to $M_{[301-304]}$-BBP matches to both self-cross-link and non-cross-linked species. Some if not all of the species desorbing at this m/z are likely self-cross-links because reduction and alkylation showed loss of this peak from the spectrum, and conversion to species with two AA moities (AA-$M_{[301-304]}$-BBP-AA).

Table 2.2 – Identified photo-cross-linked peptides from MALDI-MS for each single cysteine CT variant

CT samples were labeled with BBP, photo-cross-linked, digested, affinity purified, and analyzed on MALDI-MS as described in Methods. Peaks that match to more than one CT tryptic peptide were not listed here. Some species were represented by multiple peaks in the mass spectra where the sequence was identical but varied in modification such as phosphorylation states and oxidation. These species were listed only once. The list contains BBP-conjugated peptides from spectra of both non-reduced/alkylated and reduced/alkylated samples. Each table is sum of at least two repeated trials of the analyzed variant. For each analyzed spectra, the percent of total peaks that were matched to CT was 80% or higher. Peptides are color-coded based on their domain location. Yellow = domain M; Purple = Domain C; Green = Domain N; Blue = Domain P. A BBP dimer species that was routinely observed is in orange. Each peptide's residues are indicated in brackets. AA = Acetamide. **DOPG** = sonicated lipid vesicles composed of dioleoyl phosphatidylglycerol

CT-R245C no lipid		
Identity	Average Observed Mass (Da)	Sequence of peptide from contact
AA-BBP	810.3	
AA-BBP-BBP-AA	1620.0	
M[240-248]-BBP	1886.3	YHLQECVDK
AA-BBP-C[187-196]	1918.3	EAGMFAPTQR
BBP-C[219-228]	1960.4	RNLQRGYTAK
M[239-248]-BBP	2014.4	YHLQECVDK
M[240-250]-BBP	2113.6	YHLQECVDK
M[239-250]-BBP	2241.7	KYHLQECVDKVK
M[239-248]-BBP-N[His-tag-His-tag]	2395.9	MAK
M[240-248]-BBP-M[224-228]	2441.9	GYTAK
AA-BBP-P[342-353]	2547.2	TSPSSSPASLSR
AA-BBP-C[79-94]	2615.0	VYADGIFDLFHSGHAR
AA-BBP-M[252-266]	2631.2	KVKDVEEKSKEFVQK
AA-BBP-P[342-355]	2681.4	TSPSSSPASLSRSK
AA-BBP-M[284-300]	2698.2	EFIGSFLEMFGPEGALK
M[240-250]-BBP-M[278-281]	2705.2	WEEK
M[240-248]-BBP-M[253-259]	2749.2	VKDVEEK
M[240-250]-BBP-M[262-266]	2780.3	EFVQK
BBP-N[His-tag-16]	2790.3	SAMDAQSSAKVNSRKRRK
M[240-248]-BBP-N[9-15]	2801.8	VNSRKRR
BBP-C[219-238]	3134.7	RNLQRGYTAKELNVSFINEK
M[239-248]-BBP-M[229-238]	3207.8	ELNVSFINEK
M[239-248]-BBP-C[123-132]	3211.0	GFTVMNENER
AA-BBP-P[305-328]	3435.0	EGKGRMLQAISPKQSPSSSPTHER
AA-BBP-M[282-304]	3451.2	SREFIGSFLEMFGPEGALKHMLK
BBP-C[184-208]	3530.1	HIKEAGMFAPTQRTEGISTSDIITR
M[240-248]-BBP-C[197-211]	3547.2	TEGISTSDIITRIVR
AA-BBP-C[187-211]	3560.2	EAGMFAPTQRTEGISTSDIITRIVR
AA-BBP-P[329-355]	3823.2	SPSPSFRWPFSGKTSPSSSPASLSRSK
M[240-250]-BBP-N[His-tag-14]	3866.5	SAMDAQSSAKVNSRKR
AA-BBP-N[His-tag-13]	4072.7	MAKHHHHHHIEGRSAMDAQSSAKVNSRK
M[239-248]-BBP-M[282-300]	4129.5	SREFIGSFLEMFGPEGALK
BBP-M[278-309]	4495.3	WEEKSREFIGSFLEMFGPEGALKHMLKEGKGR
AA-BBP-P[336-367]	4675.5	WPFSGKTSPSSSPASLSRSKAVTSDISEDEED
AA-BBP-P[308-341]	4886.2	GRMLQAISPKQSPSSSPTHERSPSPSFRWPFSGK
BBP-M[278-317]	5380.4	WEEKSREFIGSFLEMFGPEGALKHMLKEGKGRMLQAISPK

CT-I272C no lipid		
Identity	Average Observed Mass (Da)	Sequence of peptide from contact
AA-BBP	810.0	
AA-BBP-BBP-AA	1620.0	
AA-BBP-M[246-248]	1172.3	VDK
AA-BBP-M[301-304]	1338.6	HMLK
M[271-277]-BBP	1556.7	SCDLIQK
AA-BBP-P[310-317]	1698.2	MLQAISPK
AA-BBP-N[His-tag-8]	1805.6	SAMDAQSSAK
AA-BBP-C[212-219]	1867.9	DYDVYARR
BBP-M[262-270]	1889	EFVQKVEEK
AA-BBP-C[187-196]	1916.8	EAGMFAPTQR
BBP-C[219-228]	1959.5	RNLQRGYTAK
AA-BBP-C[209-218]	2080.1	IVRDYDVYAR
AA-BBP-M[267-277]	2101.9	VEEKSCDLIQK
M[271-281]-BBP	2129.4	SCDLIQKWEEK
BBP-P[356-367]	2223.4	AVTSDISEDEED
AA-BBP-C[209-219]	2237	IVRDYDVYARR
BBP-C[212-223]	2323	DYDVYARRNLQR
BBP-N[His-tag-13]	2333.9	SAMDAQSSAKVNSRK
M[271-283]-BBP	2373.3	SCDLIQKWEEKSR
AA-BBP-C[197-211]	2470.7	TEGISTSDIITRIVR
AA-BBP-M[252-266]	2631.3	KVKDVEEKSKEFVQK
AA-BBP-M[284-300]	2698	EFIGSFLEMFGPEGALK
M[271-277]-BBP-M[229-238]	2751.4	ELNVSFINEK
BBP-N[His-tag-16]	2790.5	SAMDAQSSAKVNSRKRRK
AA-BBP-M[282-300]	2926.2	SREFIGSFLEMFGPEGALK
AA-BBP-M[253-270]	2989.7	VKDVEEKSKEFVQKVEEK
M[271-277]-BBP-P[305-317]	3004.7	EGKGRMLQAISPK
M[271-277]-BBP-P[318-328]	3090.6	QSPSSSPTHER
M[271-277]-BBP-C[79-94]	3362.9	VYADGIFDLFHSGHAR
BBP-P[305-328]	3363.6	EGKGRMLQAISPKQSPSSSPTHER
M[271-277]-BBP-P[342-355]	3428.6	TSPSSSPASLSRSK
BBP-M[284-304]	3450.6	EFIGSFLEMFGPEGALKHMLKEGK
AA-BBP-M[260-281]	3521.6	SKEFVQKVEEKSCDLIQKWEEK
BBP-P[329-355]	3687.5	SPSPSFRWPFSGKTSPSSSPASLSRSK
AA-BBP-M[282-307]	3749.3	SREFIGSFLEMFGPEGALKHMLKEGK
M[271-281]-BBP-P[310-328]	4244.1	MLQAISPKQSPSSSPTHER
M[271-281]-BBP-M[249-266]	4322.4	VKKKVKDVEEKSKEFVQK
AA-BBP-C[101-132]	4497	NLFPNTYLIVGVSSDELTHNFKGFTVMNENER

AA-BBP-M[284-317]	4604.7	EFIGSFLEMFGPEGALKHMLKEGKGRMLQAISPK
AA-BBP-P[308-341]	4887.2	GRMLQAISPKQSPSSSPTHERSPSPSFRWPFSGK
BBP-P[318-353]	4904.2	QSPSSSPTHERSPSPSFRWPFSGKTSPSSSPASLSR
BBP-M[278-317]	5379.9	WEEKSREFIGSFLEMFGPEGALKHMLKEGKGRMLQAISPK
BBP-C[184-223]	5449.4	HIKEAGMFAPTQRTEGISTSDIITRIVRDYDVYARRNLQR

CT-I272C + DOPG		
Identity	Average Observed Mass (Da)	Sequence of peptide from contact
AA-BBP	810.0	
M[271-277]-BBP	1557.9	SCDLIQK
M[271-281]-BBP	2130	SCDLIQKWEEK
AA-BBP-P[308-335]	4222.9	GRMLQAISPKQSPSSSPTHERSPSPSFR

CT-S288C no lipid		
Identity	Average Observed Mass (Da)	Sequence of peptide from contact
AA-BBP	810.0	
AA-BBP-BBP-AA	1620.0	
BBP-M[246-251]	1470.3	VDKVKK
AA-BBP-N[9-16]	1854	VNSRKRRK
AA-BBP-C[187-196]	1918.2	EAGMFAPTQR
BBP-C[219-228]	1961.1	RNLQRGYTAK
BBP-P[342-355]	2224.5	TSPSSSPASLSRSK
AA-BBP-C[212-223]	2379.2	DYDVYARRNLQR
AA-BBP-C[79-94]	2615.7	VYADGIFDLFHSGHAR
M[284-300]-BBP	2640.2	EFIGCFLEMFGPEGALK
AA-BBP-P[336-353]	2771.5	WPFSGKTSPSSSPASLSR
AA-BBP-M[284-300]-AA	2756.3	EFIGCFLEMFGPEGALK
BBP-M[239-254]	2779.9	KYHLQERVDKVKKKVK
AA-BBP-M[282-300]-AA	2999.8	SREFIGCFLEMFGPEGALK
M[284-304]-BBP	3149.2	EFIGCFLEMFGPEGALKHMLK
M[284-300]-BBP-M[249-252]	3158.4	VKKK
AA-BBP-P[308-328]	3266.7	GRMLQAISPKQSPSSSPTHER
M[282-304]-BBP	3392.9	SREFIGCFLEMFGPEGALKHMLK
M[282-300]-BBP-M[301-304]	3412.2	HMLK
M[282-300]-BBP-C[224-228]	3423.7	GYTAK
AA-BBP-C[197-219]	3510.9	TEGISTSDIITRIVRDYDVYARR
AA-BBP-M[284-307]	3524	EFIGCFLEMFGPEGALKHMLKEGK
BBP-P[310-355]	3594.3	MLQAISPKQSPSSSPTHERSPSPSFR
M[284-300]-BBP-N[His-tag-8]	3637.4	SAMDAQSSAK
M[284-300]-BBP-M[301-309]	3696.8	HMLKEGKGR
AA-BBP-M[224-248]	3821.9	GYTAKELNVSFINEKKYHLQERVDK
M[284-300]-BBP-C[123-132]	3837.0	GFTVMNENER
M[282-304]-BBP-C[220-223]	3923.7	NLQR
M[284-300]-BBP-M[260-270]	3990.8	SKEFVQKVEEK
M[282-300]-BBP-M[262-270]	4037.8	EFVQKVEEK
M[284-300]-BBP-N[His-tag-13]	4252.3	SAMDAQSSAKVNSRK
AA-BBP-P[310-335]	4372.7	TSPSSSPASLSRSKAVTSDISEDEED
BBP-M[219-250]	4659.7	RNLQRGYTAKELNVSFINEKKYHLQERVDKVK
BBP-P[318-353]	4665.1	QSPSSSPTHERSPSPSFRWPFSGKTSPSSSPASLSR
AA-BBP-P[336-367]	4676.9	WPFSGKTSPSSSPASLSRSKAVTSDISEDEED
M[282-300]-BBP-N[His-tag-15]	4793.1	SAMDAQSSAKVNSRKRR
M[271-304]-BBP	4794.0	SIDLIQKWEEKSREFIGCFLEMFGPEGALKHMLK
M[284-300]-BBP-P[310-328]	4898.4	MLQAISPKQSPSSSPTHER

M[284-300]-BBP-P[336-355]	4991.6	WPFSGKTSPSSSPASLSRSK
M[284-300]-BBP-P[329-353]	5278.2	SPSPSFRWPFSGKTSPSSSPASLSR
AA-BBP-M[278-317]	5453.5	WEEKSREFIGCFLEMFGPEGALKHMLKEGKGRMLQAISPK
M[284-300]-BBP-P[308-335]	6109.2	GRMLQAISPKQSPSSSPTHERSPSPSFR
M[282-300]-BBP-P[305-335]	6284.6	EGKGRMLQAISPKQSPSSSPTHERSPSPSFR

CT-S288C + DOPG		
Identity	Average Observed Mass (Da)	Sequence of peptide from contact
AA-BBP	810.0	
M[284-300]-BBP	2640.2	EFIGCFLEMFGPEGALK
AA-BBP-P[336-353]	2769.0	WPFSGKTSPSSSPASLSR
M[282-300]-BBP	2883.3	SREFIGCFLEMFGPEGALK
M[284-304]-BBP	3166.4	EFIGCFLEMFGPEGALKHMLK
BBP-P[308-328]	3207.8	GRMLQAISPKQSPSSSPTHER
BBP-P[318-341]	3507.6	QSPSSSPTHERSPSPSFRWPFSGK

CT-H301C no lipid		
Identity	Average Observed Mass (Da)	Sequence of peptide from contact
AA-BBP	810.0	
AA-BBP-BBP-AA	1620.0	
M[301-304]-BBP	1245.9	CMLK
BBP-M[278-281]	1345.1	WEEK
AA-BBP-M[301-304]-AA	1361.5	CMLK
BBP-M[301-309]	1776.3	CMLKEGKGR
AA-BBP-P[310-317]	1777.4	MLQAISPK
BBP-P[308-317]	1871.1	GRMLQAISPK
AA-BBP-C[187-196]	1918.4	EAGMFAPTQR
BBP-C[219-228]	1960	RNLQRGYTAK
AA-BBP-P[342-355]	2281.8	TSPSSSPASLSRSK
BBP-N[His-tag-13]	2333.4	SAMDAQSSAKVNSRK
M[301-304]-BBP-C[123-132]	2443.8	GFTVMNENER
AA-BBP-C[197-211]	2472.5	TEGISTSDIITRIVR
AA-BBP-C[79-94]	2615.2	VYADGIFDLFHSGHAR
M[301-307]-BBP-C[220-228]	2628.1	NLQRGYTAK
BBP-P[336-353]	2633.2	WPFSGKTSPSSSPASLSR
AA-BBP-M[284-300]	2699.1	EFIGSFLEMFGPEGALK
AA-BBP-C[209-223]	2748.5	IVRDYDVYARRNLQR
M[301-307]-BBP-P[356-367]	2870.4	AVTSDISEDEED
AA-BBP-M[282-300]	2926.4	SREFIGSFLEMFGPEGALK
AA-BBP-M[267-283]	2943.5	VEEKSIDLIQKWEEKSR
M[301-304]-BBP-C[79-94]	3050.2	VYADGIFDLFHSGHAR
M[301-307]-BBP-M[239-250]	3102.7	KYHLQERVDKVK
AA-BBP-P[336-355]	3305.2	WPFSGKTSPSSSPASLSRSK
M[301-304]-BBP-P[310-328]	3345	MLQAISPKQSPSSSPTHER
BBP-N[His-tag-8]	3373.7	MAKHHHHHHIEGRSAMDAQSSAK
M[301-304]-BBP-M[229-245]	3410.1	ELNVSFINEKKYHLQER
BBP-N[9-33]	3476.2	VNSRKRRKEVPGPNGATEEDGIPSK
BBP-P[310-335]	3594.9	MLQAISPKQSPSSSPTHERSPSPSFR
AA-BBP-M[260-283]	3775.9	SKEFVQKVEEKSIDLIQKWEEKSR
M[301-304]-BBP-P[318-341]	4241.1	QSPSSSPTHERSPSPSFRWPFSGK
M[301-307]-BBP-M[246-270]	4564.8	VDKVKKKVKDVEEKSKEFVQKVEEK
BBP-P[318-353]	4665.2	QSPSSSPTHERSPSPSFRWPFSGKTSPSSSPASLSR
M[271-304]-BBP	4714.1	SIDLIQKWEEKSREFIGSFLEMFGPEGALKCMLK
AA-BBP-M[229-261]	4828.5	ELNVSFINEKKYHLQERVDKVKKKVKDVEEKSK
AA-BBP-M[249-281]	4843.5	VKKKVKDVEEKSKEFVQKVEEKSIDLIQKWEEK
AA-BBP-P[308-341]	4887.4	GRMLQAISPKQSPSSSPTHERSPSPSFRWPFSGK

M[301-304]-BBP-P[329-355]	4916.5	SPSPSFRWPFSGKTSPSSSPASLSRSK
M[301-304]-BBP-M[284-317]	4990.6	EFIGSFLEMFGPEGALKCMLKEGKGRMLQAISPK
M[271-307]-BBP	5043.1	SIDLIQKWEEKSREFIGSFLEMFGPEGALKCMLKEGK
AA-BBP-C[95-132]	5140.9	ALMQAKNLFPNTYLIVGVSSDELTHNFKGFTVMNENER
AA-BBP-P[336-367]	5155.8	WPFSGKTSPSSSPASLSRSKAVTSDISEDEED
BBP-C[184-223]	5449.8	HIKEAGMFAPTQRTEGISTSDIITRIVRDYDVYARRNLQR

CT-H301C + DOPG		
Identity	Average Observed Mass (Da)	Sequence of peptide from contact
AA-BBP	810.0	
M[301-304]-BBP	1278.7	CMLK
AA-BBP-P[310-317]	1714.3	MLQAISPK
AA-BBP-M[284-300]	2683.9	EFIGSFLEMFGPEGALK
AA-BBP-M[282-300]	2928.8	SREFIGSFLEMFGPEGALK
AA-BBP-P[342-367]	3492.6	TSPSSSPASLSRSKAVTSDISEDEED
M[301-304]-BBP-M[278-300]	3936.4	WEEKSREFIGSFLEMFGPEGALK
M[271-304]-BBP	4730.6	SIDLIQKWEEKSREFIGSFLEMFGPEGALKCMLK
M[271-307]-BBP	5043.0	SIDLIQKWEEKSREFIGSFLEMFGPEGALKCMLKEGK
BBP-P[329-367]	5137.8	SPSPSFRWPFSGKTSPSSSPASLSRSKAVTSDISEDEED

2.2.4.2 Assembly of photo-cross-link coverage maps.

To facilitate a comparison of the cross-linking patterns from different BBP sites, we generated photo-cross-link coverage maps, as shown in Figure 2.8, from photo-cross-linked peptides identified by MS. The peaks assigned to the peptides in this figure matched to only one CT tryptic peptide. A major finding is that the BBP-labeled CTs share similar photo-cross-linking patterns, regardless of the site of conjugation within domain M. First, photo-cross-links to the N-terminus of domain N were identified. Considering that in the solved CT structure (PDB: 3HL4) the base of domain N (residue 40-73) is situated on the side of the catalytic domain opposite that of the linker to domain M, a domain M – N cross-link implies that either domain M or the N-terminus of domain N, or both, are very flexible to make the contact possible. Secondly, relatively unfocussed photo-cross-links to domain P and within domain M were found, likely due to some degree of disorder in the structure of these domains. Finally, and most

significantly, I found photo-cross-linking to selective peptides within domain C, which I discuss at length in the next section.

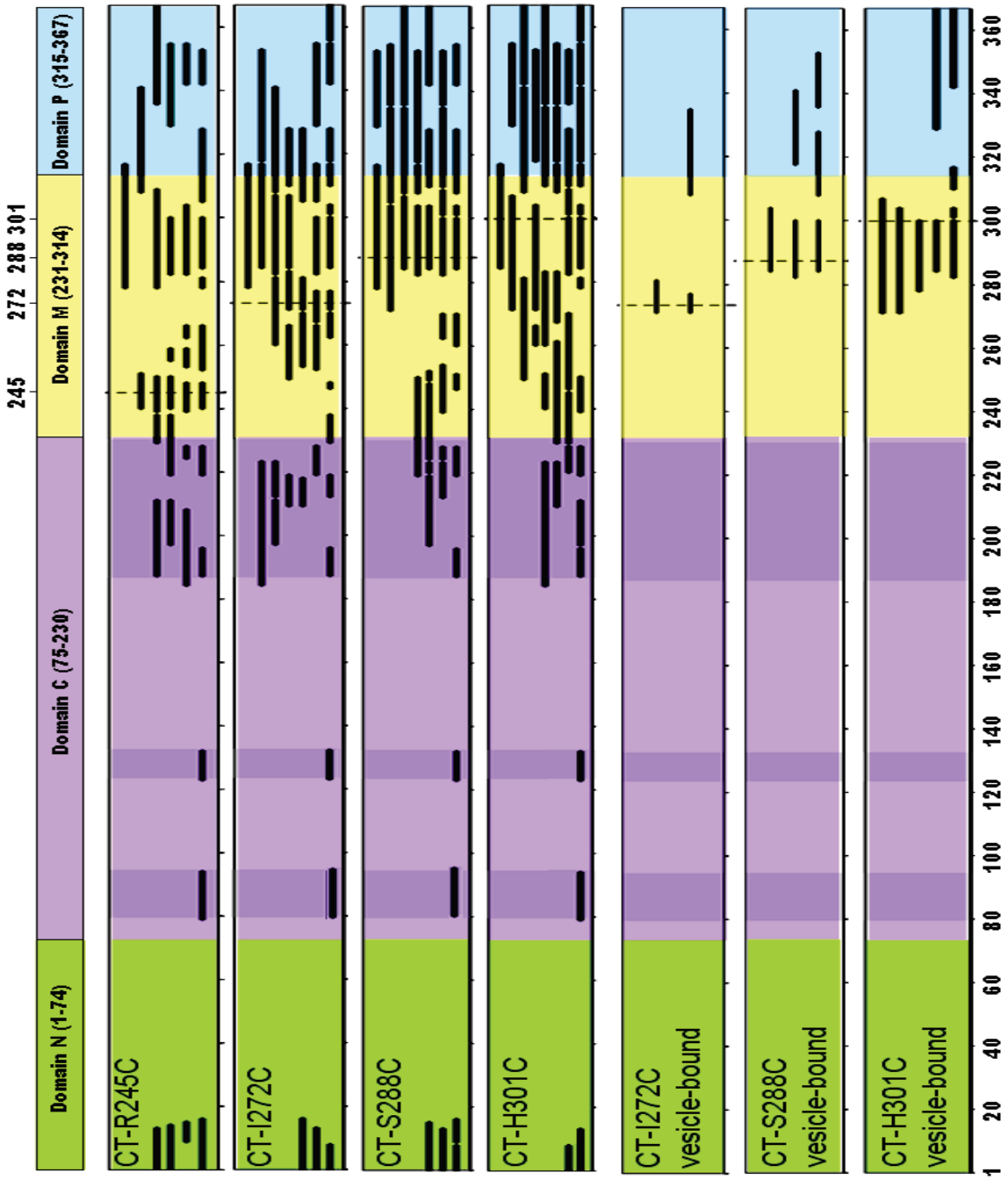


Figure 2.8 - Photo-cross-link coverage maps of different CT variants

Each bar maps photo-cross-linked peptides identified via MALDI-MS spectra. Some of the peptides contain missed cleavage sites. The dashed line represents the BBP-modified cysteine residue in each single cysteine CT variant. The three darkened areas in domain C identify the specific location of focused photo-cross-links within domain C. Each map is sum of at least two repeated trials of each analyzed variant. Regions that do not contain any photo-cross-links, for all CT variants, include residues 20 to 78, 95-122, and 133-183.

2.2.5 Photo-labeled domain C peptides map to the active site and linker between domain C and M

All analyzed CTs contained photo-cross-links to specific regions in domain C (Figure 2.8 darkened areas). We mapped these peptides onto the solved structure of CT236 (PDB: 3HL4)[73]. These peptides form part of the active site (the portion proximal to the dimer interface), helix E, and the linker between domains C and M (Figure 2.9). We identified residues 201-205 as the converging contact region, i.e. the site within 11 Å of all common photo-cross-linked domain C peptides. Briefly, using SwissDeepView, residues within 11 Å of the most commonly photo-cross-linked domain C peptides (187-196, 123-132, 79-94, and helix E) were selected. The overlapping region from these residues was determined to be the converging site (Figure 2.10). If a BBP modified region of domain M were to dock onto this site, BBP could potentially react with all the photo-cross-linked domain C peptides. This region, highlighted in red in Figure 2.9, could serve as a site for the docking of domain M and for its silencing action.

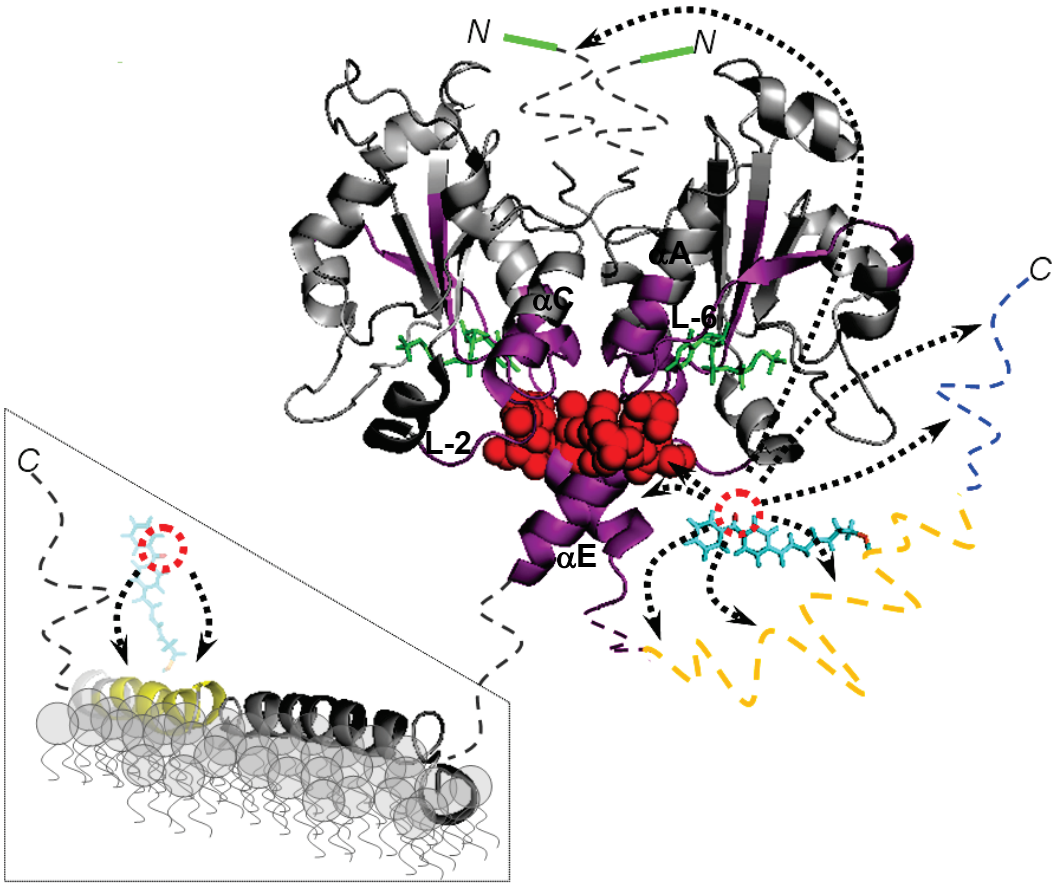


Figure 2.9 - BBP forges photo-cross-links to all regions within CT, including a region flanking the active site

CT structure composite: The catalytic domain of rat CTα (PDB ID = 3HL4) is shown in Pymol cartoon mode. CDP-choline (green sticks) is bound to the active site. Identified photo-cross-links from domain M to domain C are purple, with the proposed docking site (residues 201-205) for domain M shown as red balls. The dashed lines represent unsolved regions of CT (domain M in yellow, domain P in blue, and domain N in green). Part of the BBP molecule (from benzophenone to sulfhydryl-reactive center) is represented as cyan sticks, with the benzophenone included in a red dashed circle. **Inset**. In the membrane-bound form, photo-cross-links were limited to sites of immediate proximity to the modified cysteine. Domain M is modeled from the solved NMR structure, PDB IDs = 1PEH and 1PEI.

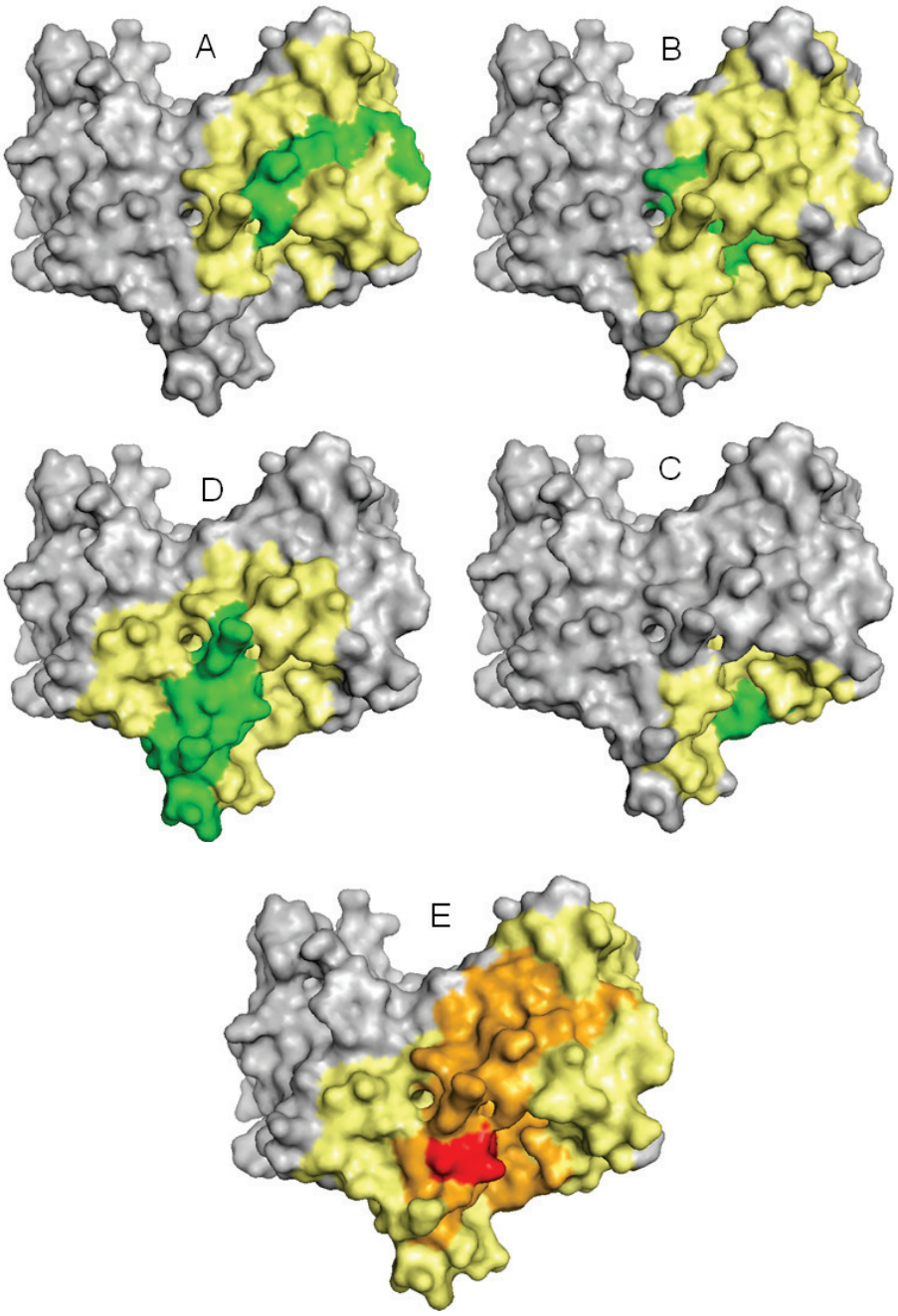


Figure 2.10 - Calculated converging point from photo-cross-linked peptides identified in MALDI-MS could potentially serve as docking site for domain M

SwissDeepView was used to determine residues within 11 Å (yellow in **A, B, C,** and **D**) of the most common photo-cross-linked domain C peptides (green in **A** [187-196], **B** [79-94], **C** [123-132], and **D** [197-208]). **E** shows regions of overlap from these calculated residues. Specifically, yellow represents no overlaps, orange represents overlap of calculated residues from two or more different photo-cross-linked domain C peptides, and red represents overlap of all four domain C peptides. At the proposed docking site (red, aa 201-205) BBP attached to the docked domain M, can reach all of the common photo-cross-linked domain C peptides found in the MS analysis. Images were created using surface topology depiction in Pymol.

2.2.5.1 Nearly all peptides derived from CT can be detected in the crude digest by MALDI-MS

The data I have described so far suggest that the domain M contact with domain C is restricted to a location flanking the active site. However selective cross-linking could reflect artifacts of analysis. To examine whether peptides in the missing regions of domain C on the photo-cross-link coverage map were absent due to difficulty in detection on MALDI-MS, I analyzed digests of photo-labeled CT-H301C before and after avidin purification. As shown in Figure 2.11 and listed in Table 2.3, I readily identified in the crude digests peptides such as $C_{[163-183]}$, $C_{[148-162]}$ and $C_{[101-122]}$ that were absent in the photo-cross-link coverage map. Sequence coverage for this set of spectra was 99% (missing a peptide encompassing residues 184-186). Additionally, several peptides that represented very minor peaks in the crude digest were augmented in the spectra following avidin affinity purification; These peptides were mapped to species containing BBP. Therefore, the absence of photo-cross-links to peptides derived from large stretches within the catalytic domain is not due to failure to generate peptides within these regions during trypsin digestion, nor inability to detect these species on MALDI-MS. Rather, it is due to lack of reaction with the BBP-conjugated residue within domain M.

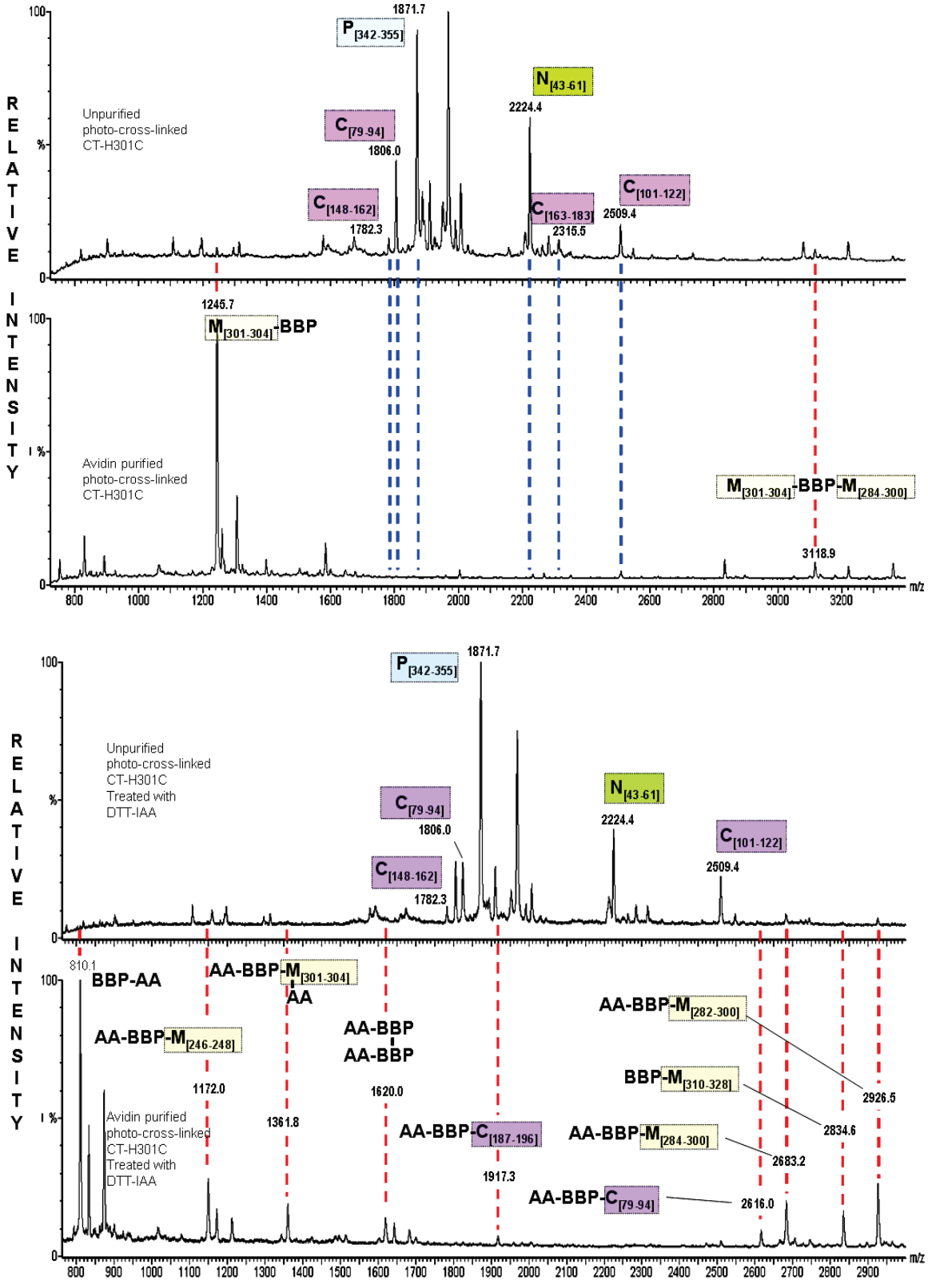

RELATIVE INTENSITY
Unpurified photo-cross-linked CT-H301C
P[342-355]
1871.7
N[43-61]
2224.4
C[79-94]
1806.0
C[148-162]
1782.3
C[163-183]
2316.5
C[101-122]
2509.4
Avidin purified photo-cross-linked CT-H301C
1245.7
M[301-304]-BBP
M[301-304]-BBP-M[284-300]
3118.9
m/z
Unpurified photo-cross-linked CT-H301C Treated with DTT-IAA
Avidin purified photo-cross-linked CT-H301C Treated with DTT-IAA
810.1
BBP-AA
AA-BBP-M[301-304]
AA
AA-BBP-M[246-248]
AA-BBP
AA-BBP
1172.0
1620.0
1361.8
AA-BBP-M[282-300]
BBP-M[310-328]
2926.5
2834.6
AA-BBP-C[187-196]
AA-BBP-M[284-300]
2683.2
1917.3
AA-BBP-C[79-94]
2616.0

Figure 2.11 – Purification of CT tryptic peptides with monomeric avidin beads selectively enhances relative intensity (and quality) of peaks corresponding to peptides with BBP adducts

CT-H301C was labeled with BBP, photo-cross-linked for an hour at 4^{o}C, and digested with trypsin as described in Method. The samples in the first and third spectra were analyzed without purification, while the samples in second and fourth spectra were purified with monomeric avidin beads. Samples in the first and second spectra were not reduced and alkylated, while third and fourth were. Blue dashed lines highlight some of the peaks that match to peptides not carrying BBP, which were removed by avidin purification. Red dashed lines highlight peaks that match to peptides (containing BBP adducts,) that were enhanced by avidin purification.

Table 2.3 – Tryptic peptides present in a digest of photo-cross-linked CT-H301C before avidin purification

CT-H301C was labeled with BBP, photo-cross-linked, digested, and analyzed on MALDI-MS without purification on monomeric avidin beads. Most of the peptides lack BBP label and are not color-coded. The BBP-labeled peptides are color coded as in Table 2.2. Only peaks that match within 1 Da of theoretical mass were considered. Peaks that match to more than one CT tryptic peptide were not listed. Peaks that correspond to an identified peptide of same sequence but with different modifications (such as phosphorylation or oxidation) were listed only once. This is the output of one trial, which was repeated with very similar results.

Identity	Average Observed Mass (Da)	Sequence of peptide/peptide from contact
His-tag-8	2605.7	MAKHHHHHHIEGRSAMDAQSSAK
His-tag-14	4172.4	MAKHHHHHHIEGRSAMDAQSSAKVNSRKR
His-tag-16	3645.9	MAKHHHHHHIEGRSAMDAQSSAKVNSRKRRK
His-tag-33	5351.6	MAKHHHHHHIEGRSAMDAQSSAKVNSRKRRKEVPGPNGATEEDGIPSK
His-tag-His-tag	1297.1	HHHHHHIEGR
His-tag-8	995	SAMDAQSSAK
His-tag-15	1894.7	SAMDAQSSAKVNSRKRR
9-14	759.4	VNSRKR
AA-BBP-N[13-36]	3459.4	KRRKEVPGPNGATEEDGIPSKVQR
16-33	1825	KEVPGPNGATEEDGIPSK
16-36	2209.6	KEVPGPNGATEEDGIPSKVQR
34-42	985.3	VQRSAVGLR
43-61	2224.4	QPAPFSDEIEVDFSKPYVR
62-69	924	VTMEEASR
70-100	3428	GTPSERPVRVYADGIFDLFHSGHARALMQAK
79-94	1805.8	VYADGIFDLFHSGHAR
M[301-304]-BBP-C[79-94]	3050.2	VYADGIFDLFHSGHAR
79-100	2464.4	VYADGIFDLFHSGHARALMQAK
101-122	2509.4	NLFPNTYLIVGVSSDELTHNFK
101-140	4662.3	NLFPNTYLIVGVSSDELTHNFKGFTVMNENERYDAVQHSR
123-132	1197.4	GFTVMNENER
AA-BBP-C[123-140]	2963.8	GFTVMNENERYDAVQHSR
141-147	880.9	YVDEVVR
148-162	1782.8	NAPWTLTPEFLAEHR
163-183	2314.9	IDFVAHDDIPYSSAGSDDVYK
187-196	1108.3	EAGMFAPTQR
197-211	1659.9	TEGISTSDIITRIVR
209-228	2458.5	IVRDYDVYARRNLQRGYTAK
212-218	901.7	DYDVYAR
M[301-304]-BBP-M[219-245]	4581.8	RNLQRGYTAKELNVSFINEKKYHLQER
220-238	2225.2	NLQRGYTAKELNVSFINEK
220-239	2353.3	NLQRGYTAKELNVSFINEKK
224-254	3721	GYTAKELNVSFINEKKYHLQERVDKVKKKVK
229-238	1193.1	ELNVSFINEK
229-248	2490.6	ELNVSFINEKKYHLQERVDK
AA-BBP-M[229-251]	3655.8	ELNVSFINEKKYHLQERVDKVKK
229-266	4648.4	ELNVSFINEKKYHLQERVDKVKKKVKDVEEKSKEFVQK
229-281	6506	ELNVSFINEKKYHLQERVDKVKKKVKDVEEKSKEFVQKVEEKSIDLIQKWEEK
246-261	1887.7	VDKVKKKVKDVEEKSK

M[301-304]-M[246-261]	3132.2	VDKVKKKVKDVEEKSK
246-283	4618.3	VDKVKKKVKDVEEKSKEFVQKVEEKSIDLIQKWEEKSR
249-254	729.9	VKKKVK
249-300	6331.3	VKKKVKDVEEKSKEFVQKVEEKSIDLIQKWEEKSREFIGSFLEMFGPEGALK
251-277	3233.4	KKVKDVEEKSKEFVQKVEEKSIDLIQK
252-261	1191.9	KVKDVEEKSK
253-266	1692.9	VKDVEEKSKEFVQK
255-270	1951.9	DVEEKSKEFVQKVEEK
260-270	1352.2	SKEFVQKVEEK
271-309	4490	SIDLIQKWEEKSREFIGSFLEMFGPEGALKCMLKEGKGR
M[271-317]-BBP	6109.2	SIDLIQKWEEKSREFIGSFLEMFGPEGALKCMLKEGKGRMLQAISPK
278-304	3195.2	WEEKSREFIGSFLEMFGPEGALKCMLK
282-300	2116.6	SREFIGSFLEMFGPEGALK
AA-BBP-M[282-300]	2925.6	SREFIGSFLEMFGPEGALK
284-300	1872.7	EFIGSFLEMFGPEGALK
AA-BBP-M[284-300]	2681.6	EFIGSFLEMFGPEGALK
M[301-304]-BBP-M[284-300]	3117.2	EFIGSFLEMFGPEGALK
M[284-317]-BBP	4513.7	EFIGSFLEMFGPEGALKCMLKEGKGRMLQAISPK
M[301-304]-BBP	1244.6	CMLK
AA-BBP-M[301-304]-AA	1361.3	CMLK
301-307	810.6	CMLKEGK
301-309	1021.7	CMLKEGKGR
BBP-M[301-328]	3869.5	CMLKEGKGRMLQAISPKQSPSSSPTHER
AA-BBP-M[301-328]	3911.5	CMLKEGKGRMLQAISPKQSPSSSPTHER
AA-BBP-M[301-335]	4685.5	CMLKEGKGRMLQAISPKQSPSSSPTHERSPSPSFR
308-317	1181.7	GRMLQAISPK
AA-BBP-M[79-94]	2615.6	VYADGIFDLFHSGHAR
AA-BBP-P[308-328]	3186.2	GRMLQAISPKQSPSSSPTHER
308-341	4476.6	GRMLQAISPKQSPSSSPTHERSPSPSFRWPFSGK
AA-BBP-P[310-335]	3668.1	MLQAISPKQSPSSSPTHERSPSPSFR
310-335	4210.8	MLQAISPKQSPSSSPTHERSPSPSFR
BBP-P[310-341]	4700.3	MLQAISPKQSPSSSPTHERSPSPSFRWPFSGK
318-328	1213.1	QSPSSSPTHER
318-335	2212.2	QSPSSSPTHERSPSPSFR
318-341	3155	QSPSSSPTHERSPSPSFRWPFSGK
318-355	4529.7	QSPSSSPTHERSPSPSFRWPFSGKTSPSSSPASLSRSK
329-341	1481.8	SPSPSFRWPFSGK
329-353	3359.2	SPSPSFRWPFSGKTSPSSSPASLSR
329-367	5185.6	SPSPSFRWPFSGKTSPSSSPASLSRSKAVTSDISEDEED
336-341	721.7	WPFSGK
336-355	2734.6	WPFSGKTSPSSSPASLSRSK

343-353	1337.1	TSPSSSPASLSR
342-355	1953.2	TSPSSSPASLSRSK
354-367	1844.3	SKAVTSDISEDEED
M[301-304]-BBP-P[354-367]	3010.5	SKAVTSDISEDEED

2.2.6 Lipid binding breaks contact with domains N and C, and orders domain M via transition into an amphiphatic α-helix

If the contacts between domain M and the catalytic domain function in silencing CT's activity, then they should be broken upon membrane binding. Thus I repeated the photo-cross-linking analysis of CT-I272C, CT-S288C, and CT-H301C in the presence of ~4000 molar excess dioleoyl-phosphatidylglycerol (DOPG) vesicles. Figure 2.12 shows representative spectral regions from 3 CT variants analyzed for photo-cross-links of the soluble and membrane-bound forms. Peptides mapping to the three sub-regions of domain C in the soluble form of CT were absent from the spectra of membrane-bound CT. The similar intensities of peaks flanking the domain C peaks in these spectra verify the change specific to the domain C sites. The results of this analysis are summarized in the bottom three panels of Figure 2.8, and the peptides cross-linked to BBP-labeled domain M sites are listed in Table 2.2. Complete elimination of contacts with the catalytic domain, when photo-cross-linking was performed in the presence of activating lipids, supports the hypothesis that these contacts are auto-inhibitory. Photo-cross-links within domain M for the lipid-bound samples were restricted to sites of immediate proximity (Figure 2.8). Changes observed in the photo-cross-linking pattern within domain M upon lipid binding reflects a structural transition from a relatively unstructured segment (resulting in unfocused photo-cross-link) to an ordered, α-helical domain M (resulting in limitation to self-cross-links).

Although I expected to find photo-cross-links to lipids for samples irradiated in the presence of DOPG, I could not identify any BBP-modified lipid molecule on MS. This could be due to difficulty in detection of DOPG on MALDI-MS. It is also likely that structure formed in domain M restricts the proper orientation of the photo-cross-linkers required to form covalent bond with the lipids.

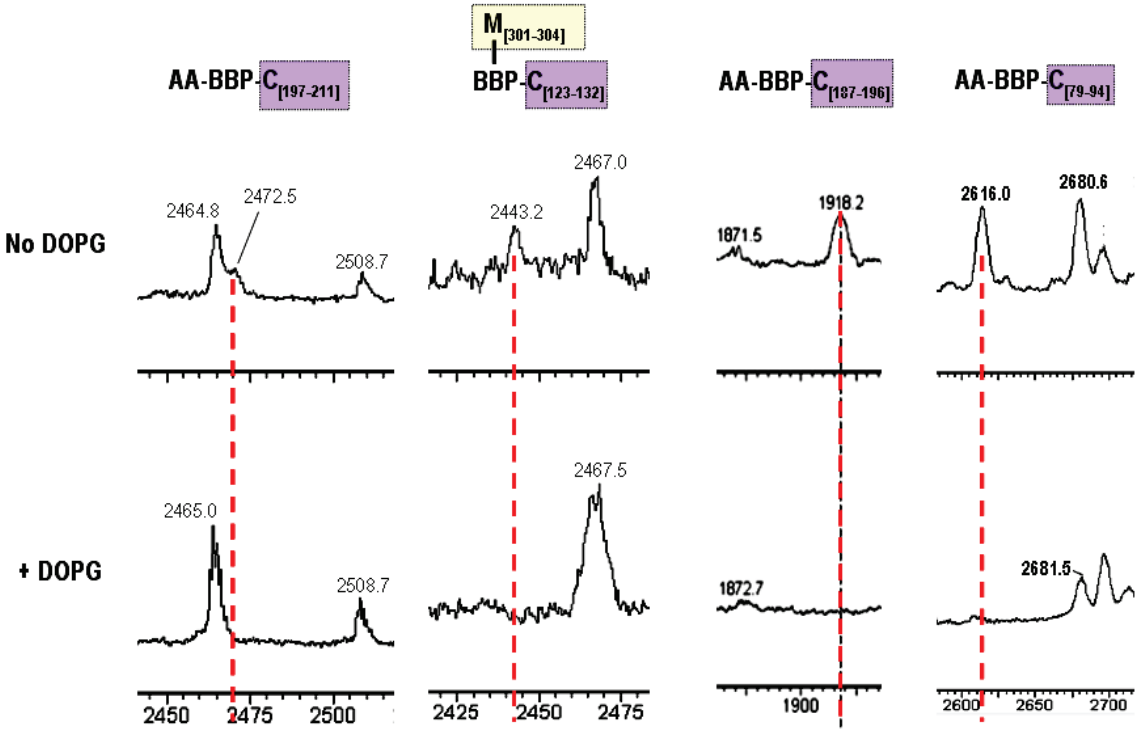


Figure 2.12 – Photo-cross-links to the catalytic domain do not occur in the presence of excess activating lipids

BBP-labeled CT-H301C was photo-cross-linked at 4°C in the absence of DOPG (**Top spectra)** or in the presence of 4000 molar excess DOPG (**Bottom spectra**), and digested with trypsin. BBP-containing peptides were purified with monomeric avidin beads, and analyzed on MALDI-MS as described in Methods.

2.2.7 Mutation at a conserved region in domain M breaks contacts with domain N and C, but does not obliterate the inhibitory action of domain M

To search for a conserved sequence that may contribute to auto-inhibition of CT, I aligned sequences of different CTs from evolutionarily divergent organisms (Figure 2.13). I identified a highly conserved sequence motif, $F^{289}LEMF^{293}$, at the C-terminus of domain M. I hypothesized that non-conservative mutations made in this motif would disrupt contacts with the catalytic domain, and subsequently eliminate auto-inhibition from domain M. I mutated the two phenylalanines in the FLEMF motif into aspartates in CT-H301C to produce CT-2D301, purified the protein and conducted BBP photo-cross-linking analysis. This domain M variant (CT-2D301) did not generate any photo-cross-links to domains N or C (Figure 2.14A), suggesting that the mutation had disrupted contacts in the region of C-301. However, the auto-regulation remained similar to the

parent, CT-H301C (Figure 2.14B). Both CTs were highly auto-inhibited in the absence of activating lipids, and were similarly activated by PG vesicles. This somewhat surprising result suggests that the FLEMF contact with the catalytic domain is not sufficient for inhibition of catalysis. Other sites must contribute, and in so doing can compensate for the loss of contact from FLEMF.

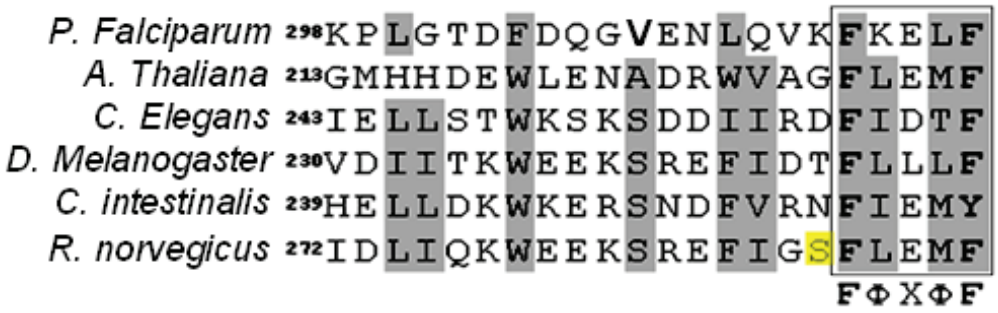


Figure 2.13 – Primordial 22 residue segment in domain M

Sequence alignment of CTs from different organisms identified a highly conserved FLEMF motif (solid box) at the C-terminus of domain M. Highly similar or identical residues are shaded in gray. This motif resides beside one of the residues that had been shown to contact helix E and the active site, 288 (yellow). ϕ = hydrophobic residue.

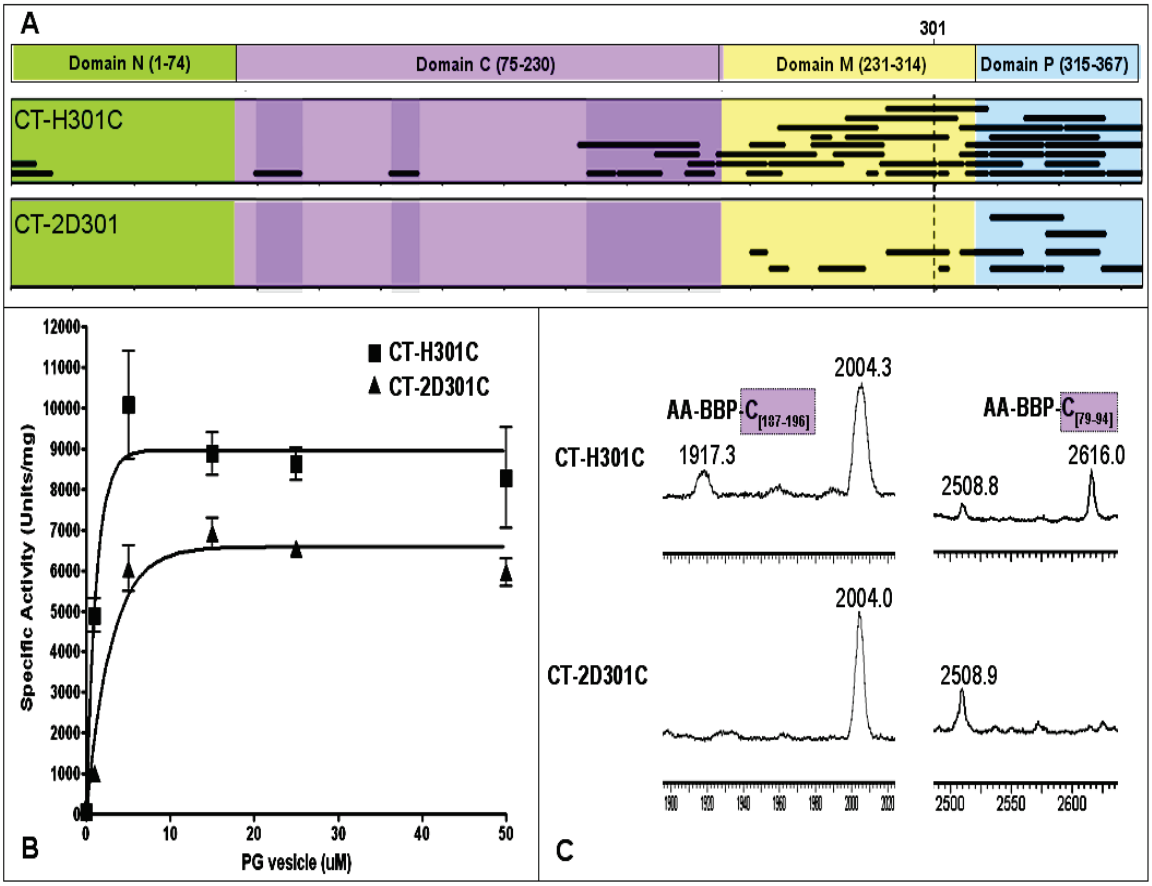


Figure 2.14 – Mutation of conserved phenylalanines in the FLEMF site inhibit contact of the region near residue 301 with the active site, but does not relieve overall auto-inhibition

A. Photo-cross-link coverage map showing elimination of all photo-cross-links to domain C and N due to the F2D mutation. **B.** Lipid regulation of CT activity is not affected by mutation of the conserved FLEMF motif. The specific activity of CT-H301C and the same construct with additional mutation at F^{289} and F^{293} to aspartate (CT-2D301) was determined in the absence or presence of the indicated concentrations of egg PC/PG (1/1) sonicated vesicles. Data are averages +/- S.D. of four determinations. **C.** Examples of MS peaks corresponding to photo-cross-linked domain C peptides that are not observed in CT-2D301. Spectra are from BBP-photo-labeled, reduced and alkylated samples.

2.2.8 Conjugated BBP at each site in domain M interferes minimally with auto-inhibition of CT activity.

Based on the results of Figure 2.4, it appears that the BBP conjugation reactions were very efficient, resulting in the majority of CT molecules incorporating a BBP adduct at the specific cysteine. If the site of conjugation participates in contacting the catalytic domain or in any way contributes to enzyme silencing, BBP might interfere with that process. I assayed the activity of each of the CT single cysteine variant with and without

the BBP adduct, as well as with and without activating lipids. The results (Figure 2.15) showed that auto-inhibition of CT was affected only marginally by the BBP addition. Because the enzyme activity is challenging to detect in the absence of lipid, 2-fold effects are not significant. The tolerance to the BBP adduct at any of several sites in domain M suggests that the domain M – C contacts does not rely on a stringent, structurally rigid and highly complementary interaction surface. Rather, this result supports a loose M-C interaction, and/or participation of several domain M sites. For example, if a BBP adduct at Cys-245 interfered with contacts at that site, if other sites participate in the silencing process, they might be unimpeded by the adduct at Cys-245. This model invokes a requirement for a flexible domain M.

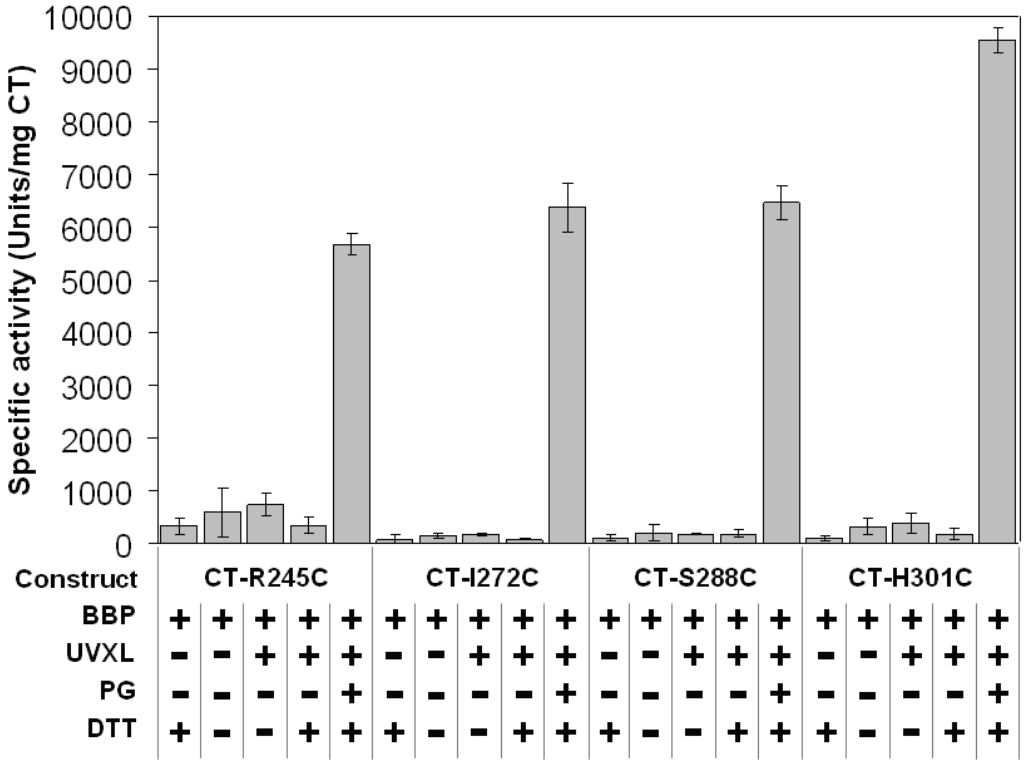


Figure 2.15 - Activity analysis on BBP-labeled and photo-labeled CT variants

CT variants were labeled with BBP, photo-labeled as described in Methods, and assayed for enzyme activity under different conditions. DTT treatment releases the sulfhydryl-mediated BBP link to the engineered cysteine. Each data set was the average ± S.D. of 3 independent trials.

2.2.9 Photo-cross-linking with CT-T207C was not successful

Residues 201-205 in helix E of the catalytic domain were proposed to be the docking site for domain M (Figure 2.10). To provide further support for this claim, I made a single cysteine variant (CT-T207C) at a position close to this region (residue 207). This residue is not as well conserved compared to the rest of helix E (Figure 1.1), and is exposed to the environment as shown in the solved structure of CT236[73]. The purified CT-T207C had a specific activity of 6990 $\pm$ 274 Units/mg with PC/PG lipid vesicle and 151 $\pm$ 69 Units/mg without lipids (N=4). I labeled CT-T207C with BBP successfully, but failed to obtain any photolysis reaction as detected by SAV-HRP blot (eg. BBP did not form a β-ME resistant covalent bond). Failure to photo-cross-link cannot be attributed to the lack of specific contact from domain M to this region, as even cross-links to the immediate proximity on helix E (self-cross-links) were not observed. Instead, it is possible that this region lacks the flexibility that domain M possesses to allow BBP to sample as many orientations as needed for successful photolysis. It is also possible that BBP might be somehow trapped within small pockets on the catalytic domain near helix E, or the dimer interface region in helix E, restricting proper positioning of BBP for photolysis.

2.3 Discussion

2.3.1 Domain M silencing does not involve a tight complex with the active site

The original model for inhibition by domain M envisioned domain M as a rigid clamp making long-lived contact with the catalytic domain. I will refer to this as the Rigid Contact (RC) model. However, there are many observations incompatible with this model.

(1) Domain M sequence is poorly conserved. The RC model would necessitate sequence conservation in both domain C and M at the sites of contact. However, CT M domains from diverse species that are lipid-regulated differ markedly in sequence, length, and the presence of 11mer motifs (Figure 1.1)

(2) Domain M and C do not interact "in trans", unlike other auto-inhibitory proteins[4,27,124,125]. An attempt to capture a tight interaction between domain M

and catalytic domain failed when His-tag CT236 immobilized on Ni^{+} beads was used as bait to trap domain M peptides or the complete CT tail domain, residues 237-367 (Cornell, unpublished data). In addition, the specific activity of truncated CT lacking domain M was not inhibited by addition of domain M peptide into the reaction mixture. Therefore, if there is a direct interaction between domain M and the catalytic domain, it may be very weak and require these two domains to be linked via the polypeptide chain.

(3) There is no footprint of domain M on the catalytic domain based on proteolytic pattern of CT +/- membrane. If domain M makes long-lived auto-inhibitory contacts, photo-cross-linking should have mapped out a region of contact at the catalytic domain characterized by major peaks in the mass spectra. Additionally, each BBP site should have generated at least partially differing patterns of cross-linked species. Although I did identify lipid sensitive photo-cross-links to the catalytic domain, these cross-links were typically minor components in the mass spectra. Importantly, each BBP site forged contacts to an identical set of peptides, mapping to the C-terminal underside of domain C, flanking the portion of the active site bordering the dimer interface, as well as helix E and the linker between M and C (Figure 2.9, 2.10). The absence of a tight M-C complex has also been supported by deuterium exchange/Mass spec analyses of CT236 (lacking domain M) and CT312 (containing domain M). The only portion of the catalytic domain that shows reduced exchange associated with the presence of domain M was the helix E region. (Schriemer, Lee, de Silva, Cornell, unpublished).

2.3.2 Domain M silences the active site by multiple, transient, alternating contacts (MTAC)

An alternative model for silencing involves multiple transient alternating contacts (MTAC) that exert inhibition in a collective or cooperative manner. Many observations from my data support this:

(1) Contacts from **multiple** parts of domain M to the catalytic domain were identified via photo-cross-linking analysis. Specifically, residues from each analyzed CT variant (distributed from N to C terminus of domain M) focussed their contacts to

the catalytic domain near the active sites and helix E, perhaps converging at residues 201-205. Although I lack definitive proof, these contacts may be important for auto-inhibition since they are lipid- sensitive. The contact between domain M and the catalytic domain seemed to be driven by hydrophobic interactions, as mutation from phenylalanines to aspartates in CT-2D301 completely eliminated these inter-domain contacts. The importance of multiple hydrophobic residues participating in auto-inhibition was also demonstrated in C. elegans CT. The substitution of conserved hydrophobic residues one at a time had little effect on activity, but when four of these were substituted *en* bloc to serine, a large increase in lipid- independent activity ensued[105].

(2) The **transient** nature of domain M contacts is revealed in the low relative intensity of photo-cross-links to the catalytic domain compared to the intra-domain cross-links, reflecting a low probability of contact with domain C. Activity analysis on photo-cross-linked samples agrees with this. Photo-cross-links to the catalytic domain are expected to irreversibly inhibit CT. Since lipid-activation of CT remained mostly unchanged for photo-cross-linked samples (Figure 2.15), they are likely to be a minority species in the photo-cross-linked population.

(3) Activity analysis on BBP-labeled CT variants and CT-2D301 support the **alternating** feature of MTAC. A potentially disruptive modification (BBP is much bulkier than cysteine) at various domain M sites did not completely activate CT (Figure 2.15). In fact, BBP modification at each of 4 sites resulted in only ~2 to 3-fold increase in lipid-independent activity. Photo-cross-linking, involving predominantly cross-links within domain M near the modified cysteine residues, resulted in ~5-fold increase in lipid-independent activity. These effects are small compared to activation by lipids (20 -90 fold). Similar compensatory effect was observed with the activity analysis of CT-2D301. Although contacts to the catalytic domain were completely eliminated, auto-inhibition remained largely unchanged (Figure 2.14). In addition to these analyses, my proposed docking site (residues 201-205) would require multiple contacts from domain M (ranging from 245 to 301; $\geq$ 60 residues) to be made in alternating fashion.

(4) **Cooperativity** is suggested in my finding that elimination of the contact in the region of the FLEMF motif was insufficient to inhibit silencing by domain M. It is also strongly suggested by the findings of the Friesen group working with CT from C. elegans, where simultaneous mutation of four distantly distributed hydrophobic residues into serines was required to overcome auto-inhibition by domain M. in C. elegans CT[105]. Curiously, in this analysis, a single non-conservative mutation at one of the phenylalanines in the FLEMF caused a dramatic increase in lipid-independent activity.

(5) Cooperative interactions of domain M with the active site could also involve alternation between each M domain of the dimer converging on the same spot near the helix E dimer interface. Disulfide bridging analysis of domain M showed M domains contact each other in the soluble form, hinting that they may work **cooperatively** as an unit to bring about auto-inhibition. In support of this hypothesis, Taneva *et al* showed that a CT heterodimer with only one domain M can efficiently turn the enzyme on and off[94]. Since the domain M docking sites (residues 201-205) of the α-E helices are part of the dimer interface, it may be possible to inhibit both active sites of the CT homodimer with contact from only one domain M. In a cooperative, and alternating manner, one domain M might search the lipid membrane for a suitable surface to bind while the other domain M keeps the CT active site silenced. Close communication between M domains allows for the lipid binding of one domain M to relieve auto-inhibition from the other M domain as well. This mechanism is highly advantageous as it allows for efficient silencing and activation of CT.

2.3.3 Domain M Flexibility requirement for the MTAC model.

Flexibility is required for the transient nature of contacts from domain M to domain C, and alternating between multiple contact sites. Flexibility was suggested by the unfocussed photo-cross-linking pattern within domain M in the soluble form and also by the facility with which disulfide bridging occurs between adjacent M domain cysteines. If domain M were a rigid, compact structure that makes specific contact with a small region in the catalytic domain, then a drastic mutation at any place in domain M

might affect the binding strength of the entire domain and relieve at least some of the auto-inhibition. This was not observed.

Nevertheless, results from disulfide bridge analysis and photo-cross-linking and are merely suggestive. Ambiguity exists, at least for the interpretation of the unfocussed photo-cross-linking pattern. The BBP molecule contains an 11 Å flexible linker (it contains at least two freely rotatable single bonds), and this will allow a sizeable search radius. Since domain M flexibility is important for the MTAC model, it was necessary to confirm the flexibility of domain M by probing its structural dynamics with a more direct approach, such as fluorescence anisotropy. This is the subject of Chapter 3.

2.4 Methods

2.4.1 Construction of single cysteine variants of rat CTα in pVL1392

A cysteineless full length CT, with the seven native cysteines substituted to serine, was prepared by Mingtang Xie. This construct in pBluescript was used as a template for QuickChange site-directed mutagenesis to introduce cysteines into the various sites indicated in Figure 2.2 [123]. For all CT variants used in this project, with the exception of CT-T207C, the baculovirus shuttle vectors were constructed as follows: The copied mutant DNA strand was cut using EcoRI producing a 1075 bp fragment containing codons for residues 29 to 367. This fragment was purified and ligated with the 10 kb vector fragment of EcoRI-cut pVL1392-His-tagged 236 vector (having codons for His-tag and first 29 CT residues that do not contain any cysteines). This strategy replaced all the CT sequence that contained cysteines. All constructs contained His tags followed by a linker peptide which contains a Factor Xa cleavage site. The N-terminal sequence of all CTs is MAKHHHHHHIEGRSA- CTstart.

2.4.2 Expression and purification of CT constructs in T. ni cells

Single cysteine variants in pVL1392 were co-transfected into *Trichoplusia ni* cells with Baculovirus as described[104]; while CT-T207C was co-transfected into COS cells as described[103,123]. All CT variants, except CT-T207C, were expressed and purified under

native condition using His-tag column chromatography as described[123]. 1 mM DTT was included in all steps of purification. As these full length CTs tend to fall out of solution, dialysis to remove imidazole was performed with the addition of 0.25 mM TX-100 for 1.5 h at 4°C. CT-T207C expressed in COS cells were a gift from Ziwei Ding. It was not soluble upon expression and needed to be purified from pellet under denatured condition as described[126,127]. The imidazole was removed as above prior to storage. Purity of all CT variants was estimated by SDS PAGE and Coomassie staining, and can be assessed from the gels shown in Figure 2.3 - 2.5. The concentration of purified stocks was determined by the method of Bradford[128]. All purified proteins were stored at -80C in small aliquots. The CT samples that were utilized for all analyses in this thesis contained the His tag and linker (see 2.4.1). Taneva *et al* have shown previously that the His tag on rat CTα does not interfere with its folding or activity[94].

2.4.3 BBP labeling

20 μg (0.5 nmol) of His-tagged single cysteine CT variant, in dialysis buffer (20 mM NaH_2PO_4, 150 mM NaCl, 1 mM DTT, 0.25 mM TX-100, pH 7.4), was bound to 100 μL of prepared Ni-beads for 20 min at room temperature, with rotating wheel spinning at 5 rpm. After removing supernatant, Ni-beads were re-suspended in 100 μl of 240 μM BBP in DMSO, vortexed briefly, and centrifuged to pellet beads. Supernatant containing excess BBP was removed. Beads were washed once with pure DMSO, four times with equilibration buffer (50 mM NaH_2PO_4, 500 mM NaCl, 0.25 mM TX-100, pH 8.0), and once with wash buffer (10 mM Tris, 150 mM NaCl, 0.25 mM TX-100, pH 7.4). All washes were 100 μl at RT. BBP-labeled CT variants were then recovered from Ni-beads with 4 x 100 μl elution buffer (350 mM imidazole, 10 mM Tris, 150 mM NaCl, 0.25 mM TX-100, pH 7.4).

2.4.4 Photo-irradiation of BBP-labeled samples.

Sample recovered from the Ni-beads was further diluted with 400 μl wash buffer. Diluted BBP-labeled CT was UV irradiated on ice in the cold room (4°C) for 1 h, using a hand held UV GL25 mineralight UV lamp set at 365 nm, as the light source[129-131].

Distance between surface of sample solution (in eppendorf tube or tube lid) and light source was about 1 cm. In my initial trial I varied the time of irradiation between 5 and 60 min. The extent of photo-crosslinking, as assessed by reducing SDS-PAGE and reactivity with streptavidin-HRP, showed that 1 h provided maximum benzophenone mediated conjugation.

For photolysis in the presence of lipid vesicles, 6 mM DOPG sonicated unilamellar vesicles (SUVs) were prepared in Buffer A.15 (10 mM Tris , 150 mM NaCl, 1 mM EDTA; pH 7.4), following established protocol[83]. In place of the 400 µl wash buffer added prior to UV irradiation described above, 400 µl of the DOPG SUVs were added to the BBP-labeled samples. The DOPG to CT molar ratio was about 4000:1. Diluted samples, either with wash buffer (samples without lipids) or with DOPG SUVs (samples with lipids) were UV irradiated as described above.

2.4.5 Delipidation

As lipids and detergents can interfere with desorption and detection of analytes on MALDI-MS, it was necessary to remove these before MS analysis. For CT-I272C and CT-H301C delipidation, pure TX-100 was added to UV irradiated samples to the final concentration of 1%. After complete digestion with trypsin, excess DOPG and TX-100 were removed with extensive washes during purification of BBP-labeled peptides, as described below. For CT-S288C delipidation, a solvent extraction procedure was followed[100]. Solvent extracted samples were reconstituted in dilution buffer (10 mM Tris, 150 mM NaCl,,175 mM imidazole, 0.25 mM TX-100, pH 7.4) prior to digestion described below.

2.4.6 Trypsin digestion and purification of BBP-labeled peptides

800 µl BBP-labeled and photo-cross-linked samples were completely digested with trypsin at w/w ratio of 1/5 (trypsin/CT) for 1 h at 37^{o}C. The digestion was quenched with 2 mM PMSF. Half of each sample was reduced with 10 mM DTT for 15 min at 37^{o}C, and alkylated with 500 mM IAA for 30 min at RT in the dark. Monomeric avidin beads were prepared following manufacturer's protocol from Pierce. DTT-reduced and

IAA alkylated, as well as untreated samples were bound to 5 µl monomeric avidin beads for 1 h at RT. Beads were washed extensively with 4 x 50 µl wash buffer (100 mM NaH_2PO_4, 150 mM NaCl, 0.25 mM TX-100, pH 7) and 2 x 100 µl ddH_2O. BBP-labeled peptides were then recovered with 4 x 5 µl acidic elution buffer (70% ACN + 30% TFA).

2.4.7 MS analysis

Purified BBP-labeled peptides were mixed 1:1 with 10 mg/ml CHCA in acidic elution buffer, and co-crystalized on sample wells of a MALDI target plate (12 x 8 + 6 x 4.wells). Standard mix, containing angiotensin II (1047 Da), P14R (1534 Da), ACTH fragment 18-39 (2466 Da), insulin oxidized B chain (3496 Da), and insulin (5734 Da), was also mixed 1:1 v/v with 10 mg/ml CHCA in acidic elution buffer, and co-crystalized on a MALDI target plate. Samples were allowed to dry at room temperature in dark. Mass spectra of the samples were collected on a MALDI-time of flight-MS equipped with delayed ion extraction and operated in linear mode (MALDI-LR; Waters Technologies). For all samples, the source voltage was 15 kV, the detector voltage was 1.8 kV, and the delay time was 500 ns. The mass range was set at 700 Da to 10,000 Da. Each mass spectrum was the sum of at least 200 laser shots. MS-Digest program was used to calculate the theoretical mass of tryptic digests of different CT constructs[132]. MassLynx (from Waters) was used to visualize MS data. All peak matching and identification were done manually, using the criteria described in Chapter 2 **(2.2.4).** The parameters I used to evaluate peptide masses include the following potential modifications: oxidation of methionine, acetylation of N-terminus of protein, phosphorylation of serines in domain P, alkylation from IAA, and reduction without subsequent alkylation.

2.4.8 Preparation of CT236-C288 heterodimer

This protocol was adapted from Taneva *et al* [94]. His-tagged CT-S288C and Strept-tagged CT236 were first denatured separately in the dialysis buffer with 1 mM SDS at 37^{o}C for 15 min. These two proteins were then mixed together at 1 to 1 molar ratio, and incubated at 37^{o}C for 5 min to form heterodimers. SDS in the buffer was diluted to 0.5

mM, and dialyzed at 4°C over-night to remove SDS. Dialzyed sample, containing a total of 1.5 nmol of Strept-tagged proteins (includes heterodimer and CT236 homodimer) were incubated with 100 µl Streptactin beads prepared in dialysis buffer, for 30 min at 4°C on a rotating wheel. The sample was spun briefly to pellet the beads and the supernatant was collected. The beads were washed with 6 x 100 µl wash buffer (100 mM Tris, 150 mM NaCl, 1 mM EDTA, 0.25 mM TX100, 2 mM DTT; pH 8.0). Strept-tagged proteins were recovered with 2 aliquots of 100 µl elution buffer (100 mM Tris, 150 mM NaCl, 1 mM DTT, 2 mM DTT, 5 mM desthiobiotin, pH 8.0). While this sample contains a mixture of Strep-tagged CT236 homodimers and strep-tagged CT236/His full length CT-S288C heterodimers, I probed the oxidative disulfide bridge formation of only the heterodimer by detecting only those species reactive to anti-domain M antibody. It was not necessary to separate the heterodimer from Strept-tagged CT236 homodimer as Strept-tagged CT236 do not react with anti-domain M antibody.

2.4.9 Oxidation with copper phenanthroline

5 mM $CuSO_4$ and 15 mM 1,10-phenanthroline stocks were prepared in ddH_2O. $CuSO_4$ and 1,10-phenanthroline were mixed to form CuPhe immediately prior to use following the protocol of Xie *et al* [123]. For pre-quenched samples, 24 pmol of CT was first incubated with 0.6 µmol of IAA, prepared fresh in phosphate buffer (20 mM NaH_2PO_4, 150 mM NaCl, 0.25 mM TX-100, 1 mM DTT, pH 7.4), for 30 min at room temperature in the dark. Both untreated and pre-quenched samples (0.8 uM in 30 µl reaction volume) were reacted with 0.2 mM CuPhe for 10 min at 37°C with shaking. Oxidation was quenched with 20 mM IAA.

2.4.10 Determination of converging point of BBP-mediated cross-links on the catalytic domain

The photo-crosslinked domain C peptides identified by mass spec (aa 187-196, 79-94, 123-132, 197-208) were mapped onto the solved structure of CT236 (3HL4)[73] using SwissDeepView. Backbone residue within 11 Å (linker length of BBP) of each peptide were identified using the group selection tool of this program. For the segment 187-196 the neighboring residues within 11 Å are: 134, 59-68, 75, 77-86, 89-103, 106,

108-110, 122, 141, 151-174, and 178-205; For the peptide 79-94 the residues within 11 Å are: 45-48, 59, 61-64, 66-68, 71, 95-171, 173, 175, 180-183, 185-186, 189-210, 200-210 of other subunit, and 212-213 of other subunit; For peptide 123-132 the residues within 11 Å are: 44-47, 81-89, 92, 96, 110-139, 142, 144-151, 173-176, 201-207, 209-210, and 197-214 of other subunit. For peptide 197-208 the residues within 11 Å are: 83-99, 118, 122-128, 132, 135-136, 141, 168-171, 173, 193-216, 85-89 of other subunit, 115 of other subunit, 119 of other subunit, 125-136 of other subunit, 138 of other subunit, and 201-216 of other subunit. These were colored coded in Pymol for visualization (Figure 2.10). Overlapping residues in common to all peptides were then evaluated manually. All photo-cross-linked domain C peptides had overlaps at region 201-205. If a domain M site (with the BBP attached) were to dock onto this region, BBP could react with all of the photo-cross-linked domain C peptides identified.

3: Investigation of structural dynamics of domain M by fluorescence anisotropy

3.1 Introduction

Structural flexibility within domain M was tentatively suggested by both photo-cross-linking analysis and oxidation experiments, as described in Chapter 2. As this property may be a key factor in the auto-regulation of CT, it was important to confirm the dynamic nature of domain M by another method. Several methods are available for assessing protein dynamics, such as NMR, EPR, and fluorescence spectroscopies.

NMR spectroscopy is perhaps the gold standard in the investigation of structural dynamics. Many parameters in NMR, such as chemical shifts, line widths, relaxation rates, and residual dipolar couplings can provide detailed information about local conformational dynamics that polypeptides adopt at atomic level of resolution and internal mobility/dynamics on a time-scale ranging from picoseconds to seconds[133-135]. However, CT is a 84 kDa homodimeric protein, and would exceed the upper molecular weight limit (~50 kDa) for obtaining high resolution structure using NMR spectroscopy.

Electron paramagnetic resonance (EPR) spectroscopy is a technique that specifically detects unpaired electrons. To monitor protein dynamics a reporter label possessing a paramagnetic species is introduced at a specific site in a protein in a process called site directed spin labeling. Structural changes and interactions in proteins often affect the mobility of the label and its EPR spectrum, and are readily detected as line-shape changes with this method. EPR spectroscopy has emerged as a valuable tool for assessing dynamics on the nsec-μsec timescale and mapping elements of secondary structure in a variety of proteins, including those (e.g. membrane proteins)[97] that could not be studied using NMR or X-ray crystallography[136,137]. However, procedures and data analysis involved in EPR experiments are often complex. As it was inaccessible to me without collaboration, EPR was not chosen for this project.

Fluorescence spectroscopy has been used successfully in many studies involving structural dynamics of proteins. Several parameters of fluorescence spectroscopy, such as fluorescence quantum yield (proportional to fluorescence intensity), fluorescence life-time, and fluorescence anisotropy can be carefully examined to yield useful structural information concerning local environment within a protein, structural movement of particular region of a protein, and interactions between domains of the same protein or between proteins and ligands/other proteins. A distinct advantage of this technique over the others is its availability as fluorometers are much cheaper (especially for steady-state anisotropy). Measurements are simple, and data interpretation is often straightforward. In some cases, use of an intrinsic fluorophore (tryptophan) offers much less perturbation of protein native structure. Alternatively, commercially available fluorophores with reactive groups for cysteines or lysines can be conjugated onto sites in the protein. As a fluorometer was immediately available, and the single cysteine variants were on hand and the method for conjugation had been developed, I chose this method to explore domain M dynamics in the context of the whole enzyme.

A review of steady-state anisotropy theory can be found in Chapter 1. Briefly, in steady-state anisotropy measurement, the rotational freedom of a fluorophore attached to a specific site in a protein is evaluated. As outlined in Figure 3.1, an excited fluorophore attached to a rigid region of a protein will have greater restriction on its rotation/freedom of movement. Consequently, emitted photons from the fluorophore will remain polarized, yielding a high anisotropy value. On the other hand, more freedom in fluorophore rotation attached to a highly mobile region will give rise to a lower anisotropy value.

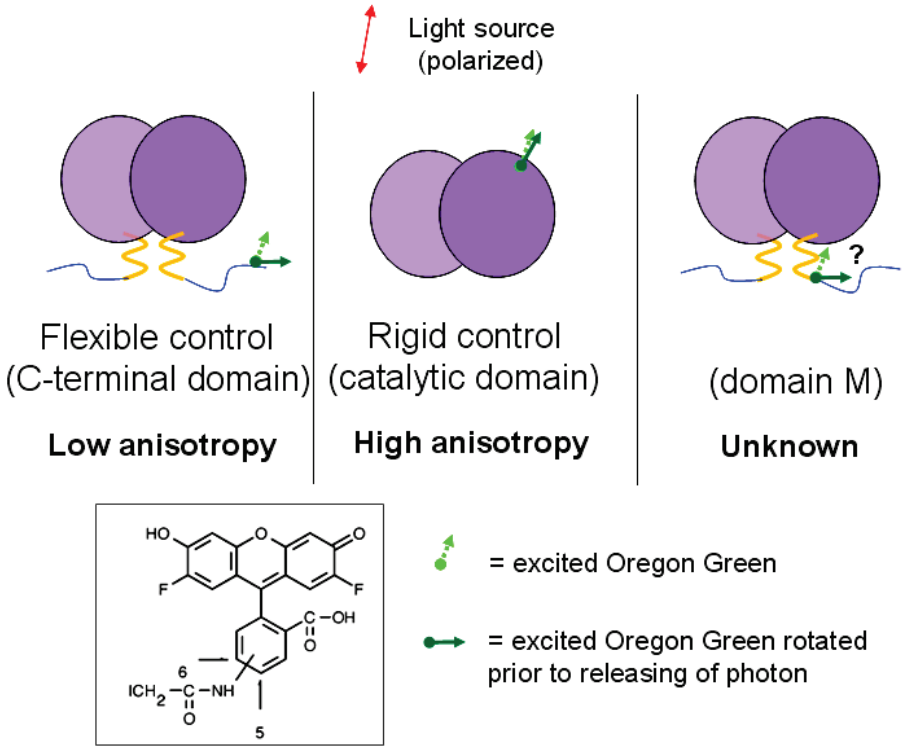


Figure 3.1 – Comparative anisotropy analysis of different constructs would reveal relative degree of order in different CT domains

CT constructs were labeled with OG (structure shown in box) as described in Methods. C-terminal domain (left, a control for flexible domain) is expected to contain low anisotropy value, while the catalytic domain (middle, a control for rigid control) is expected to contain high anisotropy value. Relative degree of order of domain M (right) can be obtained through comparison

Selecting an appropriate fluorophore is the first step in successful anisotropy experiments. Of several commercially available fluorophores (such as eosin, stilbene, or naphthalenes), Oregon Green-488 iodoacetamide (OG, a derivative of fluorescein) was chosen for its short linker length, high extinction coefficient (68,000 L $mol^{-1}cm^{-1}$), solubility in water, and sulfhydryl-reactivity[138,139].

Every site within domain M that I analyzed (residues 233, 245, 272, 288, 301) had an anisotropy value intermediate between several sites in the rigid catalytic domain and flexible control (C-terminal domain). Importantly, the 2D mutation, which had lost its contact with the catalytic domain (see Chapter 2), showed significant increase in disorder. This suggests that contacts with the catalytic domain may induce structural order and

stability into otherwise flexible domain M. The molten globular-like nature of domain M was further confirmed with urea denaturation analysis.

3.2 Results

3.2.1 Fluorophore conjugation: selection of sites and effect on CT activity

The goal was to compare the steady state anisotropy of sites within domain M with sites within a rigid CT domain and within a very disordered CT domain. For a rigid domain, I labeled the ordered cysteines in the catalytic domain of CT236 with OG. In this chapter I refer to this control construct as CT-cat. It consists of residues 1-236, thus is missing all of domain M and P. It was expected that this control would provide a high anisotropy value, because each of the 5 natural cysteines (residues 37, 68, 72, 113, 139) is known from the solved structure to reside in an ordered secondary structural element[73] (except for residue 37). CT-C359 was used to represent a disordered region. Residue 359 resides at the extreme C-terminus, and is known from limited proteolysis to be the most exposed, flexible region of the protein[100]. The single cysteine CT variants (CT-S233C, CT-R245C, CT-I272C, CT-S288C, CT-H301C, and CT-2D301) described in Chapter 2 were used to assess the anisotropy of domain M.

As the method for efficient, specific labeling of cysteines in these CTs had been worked out with BBP, the same protocol (with slight modification on washing procedure) was used for labeling with OG. Briefly, CT was immobilized on Ni^{2+} agarose beads, exposed to OG, washed extensively with buffer to remove unbound OG, and the OG-labeled CTs were recovered with imidazole. The OG/CT protein molar ratio was determined using the extinction coefficient for OG and a Bradford assay for purified CT. Similar to unmodified CT and BBP-labeled CT, OG-labeled CT remained inactive in the absence of lipids, and stimulated >20-fold by anionic lipid vesicles (Figure 3.2). Thus we could be assured that OG was reporting on the native conformation.

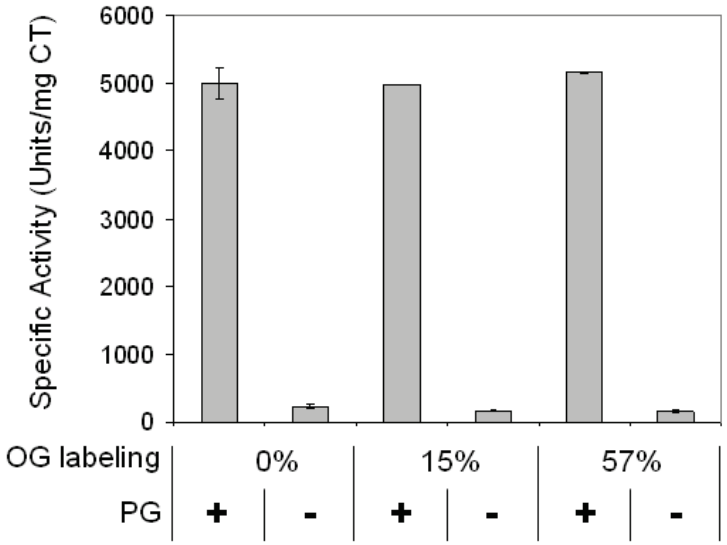


Figure 3.2 - CT-H301C lipid response is not affected by OG-labeling

CT-H301C was labeled with varying concentrations of OG by varying the OG/CT molar ratio during labeling as described in the Methods. OG labeling is the total % labeling, calculated as ([OG] / [protein])*100%.

3.2.1.1 OG fluorescence is insensitive to its specific location on CT

Having shown that OG labeling does not interfere with CT activation by lipids, I considered other potential caveats of anisotropy measurements. Since environmental influences on fluorescence lifetime can impact on anisotropy values[119], it was wise to consider whether the different locations of the OG in CT affected its lifetime. Increase in fluorescence lifetime is directly related with a decrease in anisotropy, as the fluorophore would experience a longer time for rotation prior to emitting its photon for detection. Unfortunately I did not have access to an instrumental set-up for lifetime measurements. However, since steady-state intensity is related to lifetime by $I_{(t)} = I_o\, e^{t/\tau}$; where t is time, $I_{(t)}$ is intensity at t, I_o is intensity at t = 0, τ = life-time of fluorophore, I measured the fluorescence intensities for OG at all cys positions on samples where the OG concentrations were the same. The intensities (in cps) were very similar at six distinct engineered cys sites as well as the 5 natural cys sites in the catalytic domain (Figure 3.3). The similar intensities at 11 distinct sites argue that differences observed in fluorescence anisotropies would reflect differences in rotational freedom at each position, instead of differences in fluorescence lifetimes.

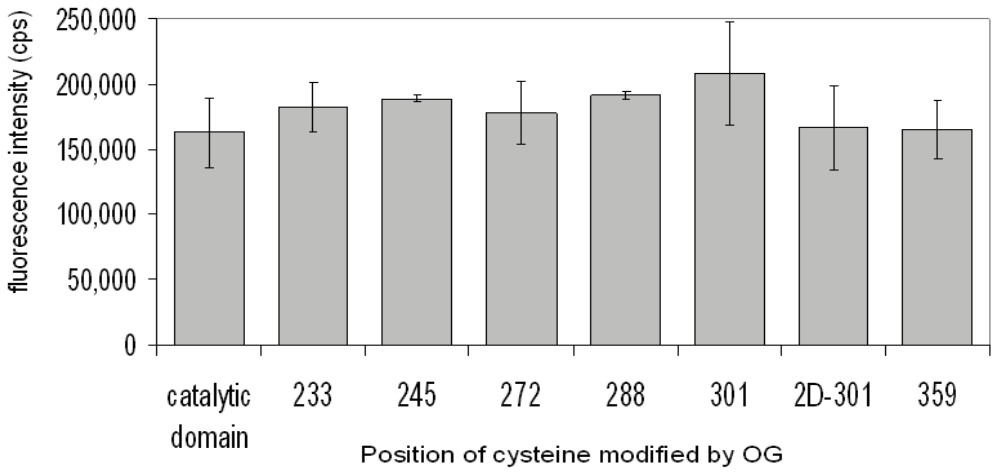


Figure 3.3 – OG fluorescence intensity is insensitive to the site of conjugation sites in CT

Fluoresence intensity was measured with excitation wavelength set at 491 nm, and emission wavelength scanned from 510 to 650 nm. [OG] = 100 nM in all samples. Intensity at maximum emission wavelength was picked (typically at 522 nm). Each data set is an average of at least two independent measurements.

3.2.2 The anisotropy measurement of OG-labeled CT is FRET-sensitive

While I was determining optimal conditions for labeling CTs with OG I found that anisotropy values decreased for the same labeled position with increasing amount of conjugated OG (illustrated for two OG-labeled constructs in Fig 3.4). This change in anisotropy values is likely due to FRET.

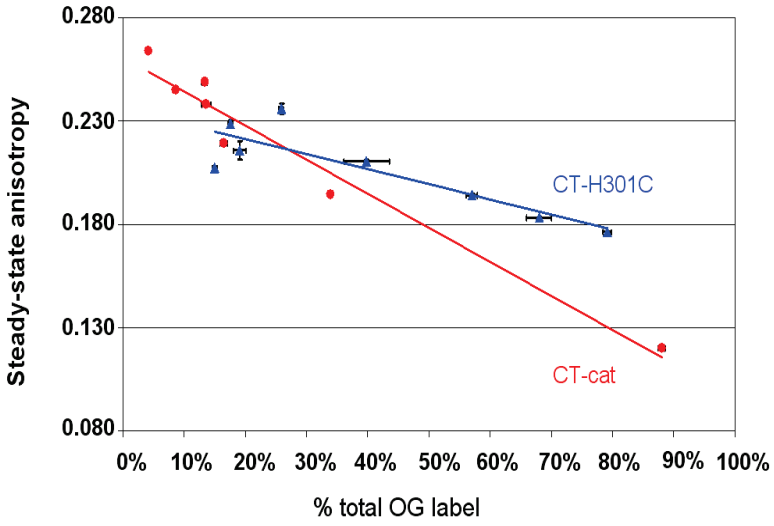


Figure 3.4 - Steady-state anisotropy of OG-labeled dimeric CTs decreases with increasing percent label

CT-H301C (blue) and CT-cat (red) were labeled with different amount of OG as described in Methods. Each data point is an average of 2 aliquots, each with 30 measurements

FRET is one of many possible paths of relaxation for an excited fluorophore, and involves transferring of energy to an acceptor fluorophore through non-radiative dipole-dipole coupling. It occurs when there is spectral overlap between the emission spectrum of the donor and the absorption of the acceptor. Since CT is a homodimer and since OG has a narrow Stokes shift (distance between peaks of fluorescence absorption and emission; ie, more spectral overlap), homo-FRET could occur for samples that are doubly labeled. The probability of FRET would increase with percent labeling. I have shown that the engineered cysteines in domain M come within contact distance for disulfide bond formation (Figure 2.3), thus FRET between two fluorophores, one on each subunit, is highly likely. FRET can result in additional angular displacement of the emission oscillator and hence lower anisotropy values (Figure 3.5)[119].

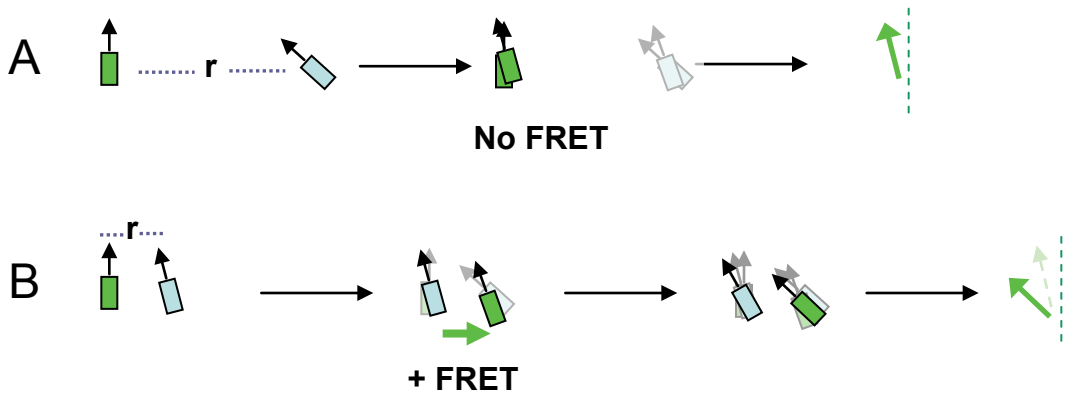
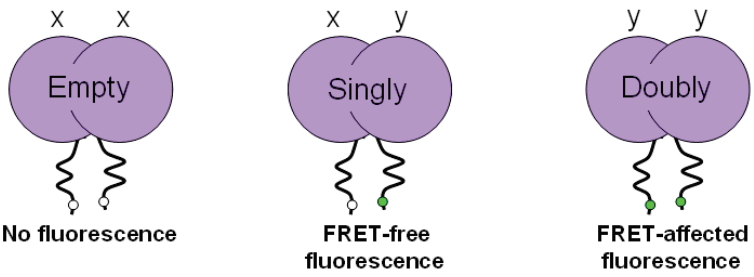


Figure 3.5 - Possible mechanism for lowering of steady-state anisotropy by homo-FRET

A. The emission from a fluorophore with rigid orientation will remain polarized in the absence of FRET (either there is no spectral overlap between emission and excitation wavelength, or distance between two fluorophore (r) is much greater than Foster distance). **B.** In homo-FRET, the dipole of the excited OG interacts with another OG dipole, and energy is transferred as shown by the transfer of green color. The second excited OG may rotate further before emitting a photon for detection.

Since it was difficult to prepare samples with an exactly pre-defined and low percent OG conjugation from one labeling reaction to another, it was desirable to obtain FRET-free anisotropy in order to compare between different OG-labeled positions. This could be done by examining anisotropy at various concentrations of OG (i.e different % total labeling), and extrapolating to [OG] = 0. For each total percent labeling, the probability distribution of unlabeled, singly labeled, and doubly labeled pairs distribute as

a simple polynomial, as shown in Figure 3.6. The probability of both FRET-affected and FRET-free species is thus easy to calculate.



Total % label (can be measured)	Calculated probability of each species		
100%	0%	0%	100%
50%	25%	50%	25%
15%	72%	25.5%	2.3%
10%	81%	18%	1%

x = Unmodified cysteine; y = cysteine modified with OG

Figure 3.6 - Probability of obtaining FRET-free and FRET-affected species can be calculated from total % label

Total % label for each OG-labeled samples can be measured ([OG] from absorbance reading at 491 nm and [protein] from Bradford assay). If x = the probability of a CT monomer with an unmodified cysteine, and y = the probability of a monomer modified by OG, the distribution of the three dimeric species shown above can be calculated as $x^2 + 2xy + y^2$.

To confirm that the decreased anisotropy as a function of doubly labeled species is caused by homo-FRET, I explored *how in theory* the FRET-free fluorescence would vary as a function of doubly labeled dimer, using the polynomial expression above. Would the relationship be linear, parabolic, or of another mathematical function? Figure 3.7 shows a shallow parabolic relationship between the percent of the monomer population experiencing no FRET *vs* the percent dimer experiencing FRET. Notably, this second-order polynomial relationship between anisotropy and FRET was observed in a study of Fluorescein-tagged mellitin that forms a complex promoted by high salt condition[140].

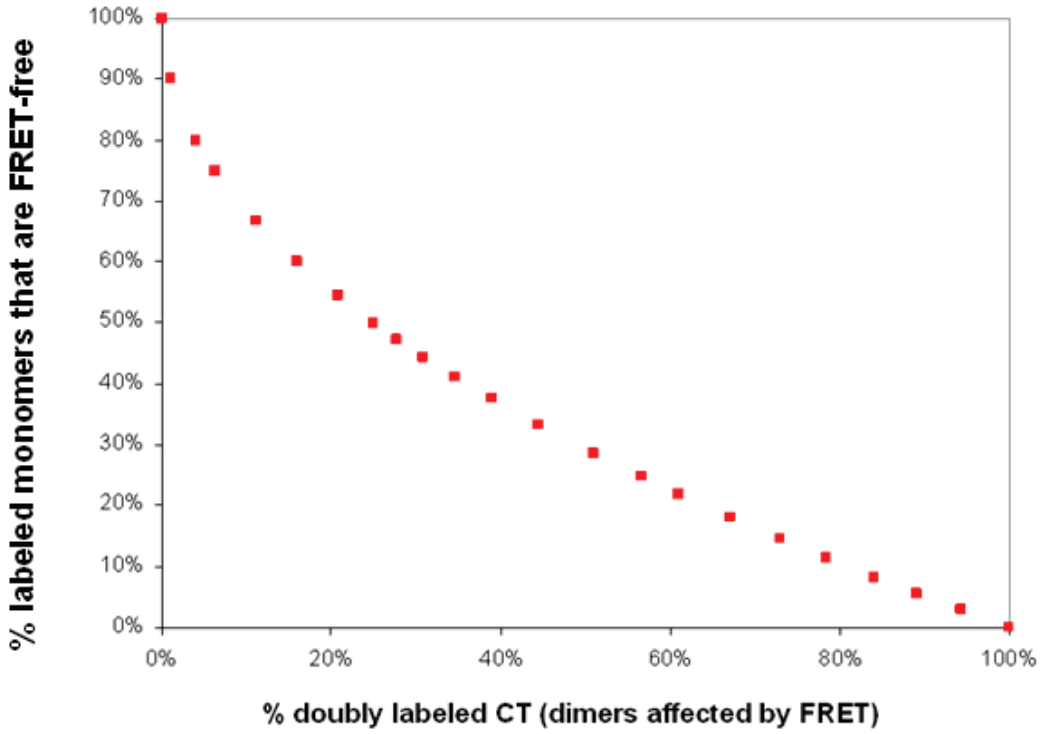


Figure 3.7 – Relationship between FRET-free anisotropy and % doubly labeled CT dimer is second-order polynomial

% doubly labeled CT is calculated from the probability distribution equation as described in Figure 3.6. % of labelled CT monomers not experiencing FRET is calculated as follows: (singly labeled dimer)/((0.5 x singly labeled dimer) + doubly labeled dimer). For example, for total OG label of 50%, % doubly labeled dimer = 25% (X-axis), while %FRET-free monomer = 50% (Y-axis).

3.2.3 Anisotropy values for all sub-regions of domain M are intermediate between that of domain C and the extreme C-terminus

I then attempted to extrapolate FRET-free anisotropy values for all analyzed samples from a curve of fluorescence anisotropy vs % doubly labelled CT, (as shown in Figure 3.8), by extrapolation to the Y-intercept. Using Prism, the model that generated the best fit with the data was a second-order polynomial, as was obtained in the theoretical FRET analysis shown in Figure 3.7. Thus it is likely that the decrease in anisotropy with increased doubly labeled CT dimers is due to FRET. Extrapolation to zero percent double labeling with Prism obtained FRET- free anisotropy values for each CT (Table 3.1). This analysis showed that CT-cat had the highest anisotropy (0.267$\pm$ 0.011) and CT359 had the lowest (0.115$\pm$ 0.004). The anisotropy values of the OG at the sites in domain M fell between this range.

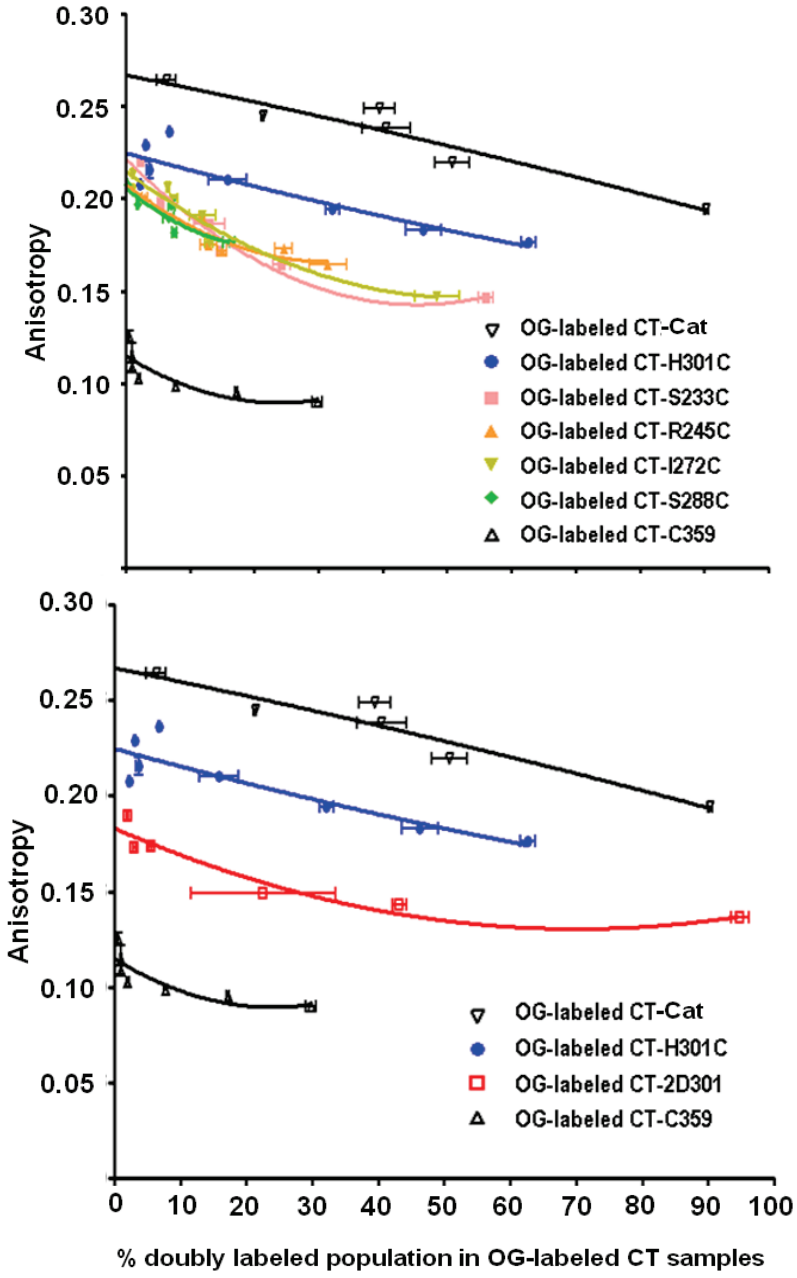


Figure 3.8 - Extrapolated FRET-free anisotropy values indicate domain M contains structural order intermediate between flexible and rigid controls

Each CT construct was OG-labeled to various percent of total as described in Methods. % doubly labeled population is calculated from the measured total OG label with the formulas: X+Y = 100% and $X^2 + 2XY + Y^2 = 100\%$; where X represents unlabeled CT and Y represents OG-labeled CT monomer (See Figure 3.6). Each data point consists of two independent sets of anisotropy measurements (30 readings each aliquot). Prism was used for finding best line fit for each data set. Source of error for the Y-axis is the minor fluctuations within the flourimeter during measurement. Source of error for X-axis is in the protein and OG quantitative analysis.

Top graph: data for rigid control (CT-cat), flexible control (CT-C359), and domain M variants

Bottom graph: data for rigid control (CT-cat), flexible control (CT-C359), CT-H301C (no mutation in FLEMF), and CT-2D301 (contains non-conservative mutation: FLEMF→ DLEMD).

The anisotropy values were converted into rotational correlation times (θ, using the Perrin equation: $r = r_o/(1+\tau/\theta)$, where r_o = the limiting anisotropy and τ = fluorescence lifetime of OG, 4.1 ns [138]. The limiting anisotropy (r_o) value for OG was determined by analysis in glycerol at -10°C. The value obtained was 0.381 ± 0.002. This is in agreement with published values for OG[139]. The rotational correlation time for the folded CT-cat construct was 9.6 ± 1.3 ns, in comparison to only 1.8 ± 0.1 ns for the flexible tail, CT-C359 (Table 3.1). A rough estimate of the rotational correlation time for the entire CT dimer can be made based on its mass (~90 kDa)[119]. Although, we do not know the hydration value for the CT dimer, the lower limit for θ should be ~ 40 nsec. Thus the θ values for the CT head and C-terminal tail reflect segmental motion at the OG-conjugated site, rather than tumbling of the whole molecule. For domain M sites, CT-S233C, CT-R245C, CT-I272C, CT-S288C, and CT-H301C all gave similar rotational correlation times averaged around 5 ns. This suggests that these parts of domain M are more flexible than the catalytic domain, but more ordered than the C-terminus of CT.

Table 3.1 – Analysis of steady-state anisotropy for OG-labeled CT constructs
FRET-free steady-state anisotropy values were obtained by extrapolation from the plot generated by Prism as shown in Figure 3.8. Rotational correlation times were calculated using the following formula: $r = r_o/(1+\tau/\theta)$, where τ = 4.1 ns and r_o = 0.381 ± 0.002.

Domain	Position of OG	Anisotropy	Rotational Correlation time (θ in ns)
C	Various (rigid)	0.267 ± 0.011	9.6 ± 1.3
P	359 (flexible)	0.115 ± 0.004	1.8 ± 0.1
M	233	0.222 ± 0.007	5.7 ± 0.4
M	245	0.208 ± 0.010	4.9 ± 0.5
M	272	0.215 ± 0.006	5.3 ± 0.3
M	288	0.206 ± 0.006	4. 8 ± 0.3
M	301 (without 2D mutation)	0.225 ± 0.007	5.9 ± 0.4
M	301 (with 2D mutation)	0.183 ± 0.005	3.8 ± 0.2

3.2.4 Anisotropy at most of the OG-labeled positions in domain M did not increase in the presence of activating lipids.

The effect of binding to vesicles on the anisotropy of each OG-labeled CT was investigated by mixing the CTs with a 400 to 1200-fold excess of PG/lyso PC (1/4) micelles. CT can bind to this micelles system efficiently. This micelles system had been used previously in CD spectroscopy of CT to avoid light scattering interference[104]. Surprisingly I found that, with the exception of CT-245C, lipid interaction did not increase anisotropy (Table 3.2). Since the lipid concentration was in excess of that needed to fully activate CT[104], the failure to order the M domain cannot be ascribed to lack of CT binding to the vesicles. The reasons for these results are explained in the Discussion.

Table 3.2 – Lipid binding did not increase order in domain M variants, except for CT-R245C
Steady-state anisotropy was measured for each OG-labeled CT in the presence or absence of various fold excess of PG/lysoPC micelles.

Position of OG	% total label	Lipid:CT	Steady-state Anisotropy
245	45%	0	0.177 ± 0.001
245	45%	891	0.221 ± 0.001
245	32%	0	0.176 ± 0.003
245	32%	408	0.230 ± 0.002
245	32%	1223	0.236 ± 0.002
272	11%	0	0.213 ± 0.001
272	11%	800	0.192 ± 0.001
288	27%	0	0.187 ± 0.003
288	27%	745	0.184 ± 0.003
301	40%	0	0.214 ± 0.001
301	40%	748	0.198 ± 0.001

3.2.5 Mutation of the conserved FLEMF motif causes a large decrease in anisotropy near that site.

In Chapter 2 I showed that mutation of ^{289}FLEMF^{293} to ^{289}DLEMD^{293} resulted in disruption of BBP mediated contacts to the active site (Figure 2.13). If that contact helps to order domain M, disruption should decrease the anisotropy. Figure 3.8 (bottom) shows that this is precisely what I observed. The anisotropy value was reduced from 0.225 to

0.183, and ϕ was reduced from 5.9 ± 0.4 ns (CT-H301C) to 3.8 ± 0.2 ns (CT-2D301); (Table 3.1).

3.2.6 Urea denaturation reveals a loosely folded structure for domain M.

The anisotropy results suggested dynamics for domain M sites that are intermediate between the well-folded catalytic domain and the unstructured CT tail. To get an additional handle on the relative foldedness of domain M, I measured the effect of increasing urea concentration on the anisotropy of all OG-labeled CTs. As a folded structure is denatured, its anisotropy should decrease. Strongly folded units require a higher urea concentration than weakly folded units for the unfolding transition. The cooperativity of the unfolding transition can also provide hints as to the degree of tertiary structure[141].

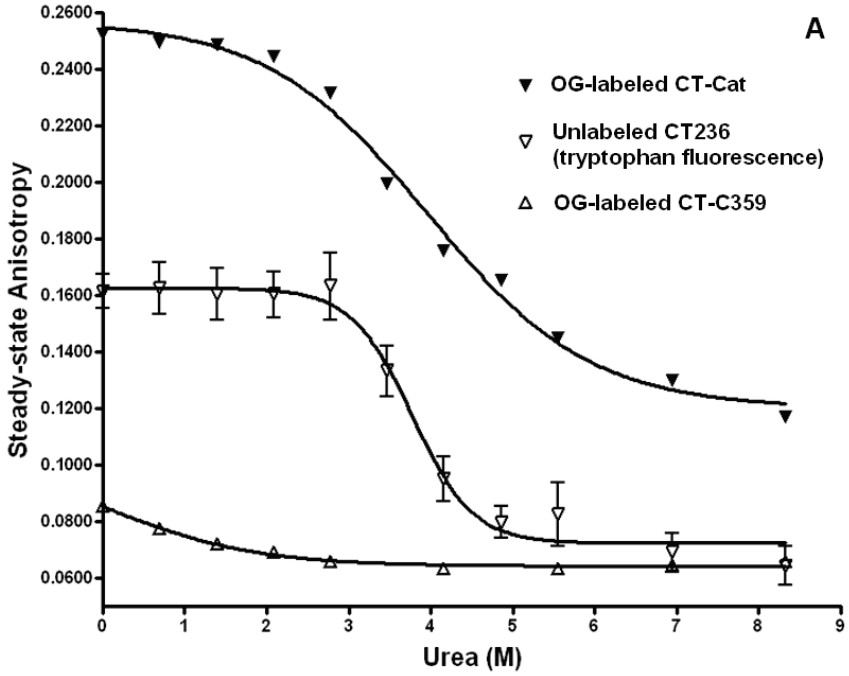
As shown in Figure 3.9, panel A, OG-labeled CT-cat with its five cysteines showed a relatively high resistance to urea denaturation (EC_{50} = 4.0 ± 0.1 M). EC_{50} refers to the concentration of urea resulting in 50% unfolding. However its transition of unfolding was not highly cooperative (slope = -0.4 ± 0.1 and complete unfolding required 5 molar units of urea), likely due to different unfolding properties of the individual sites. To clarify the characteristics of the folded CT catalytic domain, I measured the urea-induced unfolding of non-OG-labeled CT-cat by monitoring intrinsic tryptophan fluorescence instead of the OG fluorescence. CT-cat has a single tryptophan residue at position 151. This tryptophan is known to contribute to the active site and resides at a site which in the solved structure with a low thermal factor[73]. In contrast to OG-labeled CT-cat, the anisotropy for Trp151 showed a relatively sharp transition of unfolding (slope = -1.2 ± 0.2; unfolds over 2 molar units of urea), expected from the cooperative unfolding of a rigid structure. Importantly, the denaturation transition showed an EC_{50} identical to that obtained with the OG-labeled CT-cat. (EC_{50} = 3.8 ± 0.1 M). In comparison, the flexible control (CT-C359) showed no resistance to urea denaturation, as the anisotropy decreased gradually with increasing urea concentration, without a visible unfolding transition (Figure 3.9, panel A).

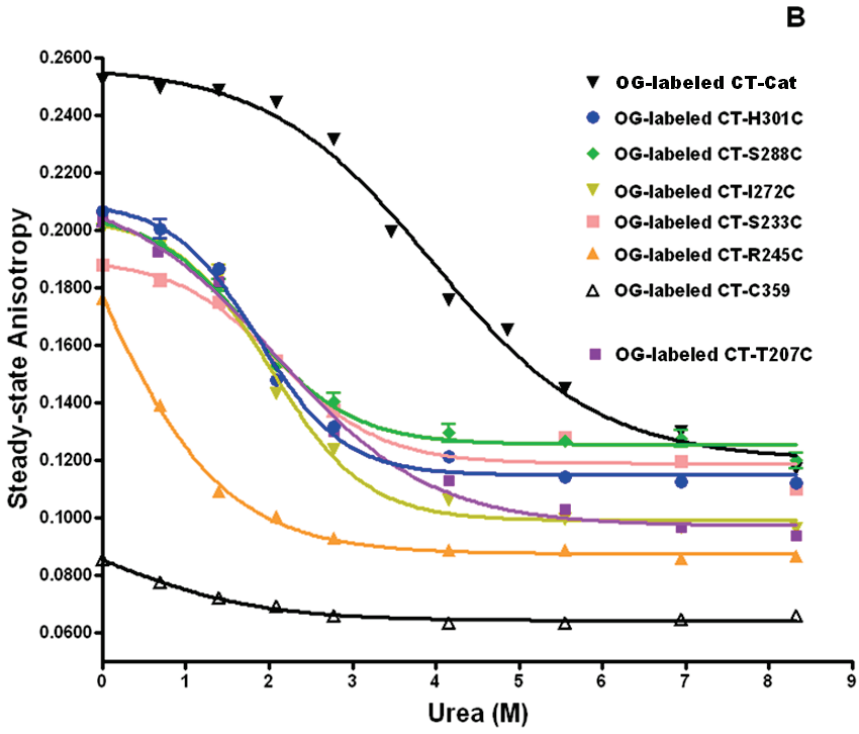
All OG-labeled domain M sites showed denaturation responses that were intermediate between CT-cat and CT-C359 (Figure 3.9, panel B). Unlike the denaturation of CT-C359, clear unfolding transitions were observed. With the exception of CT-R245C, the EC_{50} values were near 2 M, significantly lower than that of the rigid controls. However, their transitions of unfolding were relatively sharp compared to the multiply laeled CT-cat (average slope = -0.7; unfolding over 2 molar units of urea), but less cooperative than the unfolding reflected by Trp 151. CT-R245C unfolding transition appeared at < 1 M urea, indicative of more disorder at this portion of domain M compared to the other sites probed.

CT-T207C (cys engineered into Helix α-E at the C-terminus of the catalytic domain) was labeled with OG. Interestingly, residue 207 displayed a denaturation response similar to that of other domain M sites (except for 245). It had the same EC_{50} (2.1 + 0.3 M), though it has a slightly softer transition of unfolding (slope = -0.4; unfolds over 2 molar units of urea; Figure 3.9).

3.2.7 Mutation of the conserved FLEMF motif disorders local structure

The 2D mutation at the FLEMF motif had dramatic disordering effects on the adjacent C301 site (Figure 3.9, panel C). Not only was the starting anisotropy lowered by this mutation, but the unfolding transition was eliminated. The gradual reduction in anisotropy with increasing urea concentration resembled the behaviour at the disordered CT tail (CT-C359). This finding supports the hypothesis that contact with domain C induces order in this segment of domain M. An alternative explanation involves disorder as a result of loss of structure prior to losing contact. It is possible that domain M naturally possess stable structure required to make auto-inhibitory contact with the catalytic domain. Mutation caused significant melt-down of the structure, and disrupted contact with the catalytic domain.


A
OG-labeled CT-Cat
Unlabeled CT236
(tryptophan fluorescence)
OG-labeled CT-C359
Steady-state Anisotropy
Urea (M)


B
OG-labeled CT-Cat
OG-labeled CT-H301C
OG-labeled CT-S288C
OG-labeled CT-I272C
OG-labeled CT-S233C
OG-labeled CT-R245C
OG-labeled CT-C359
OG-labeled CT-T207C
Steady-state Anisotropy
Urea (M)

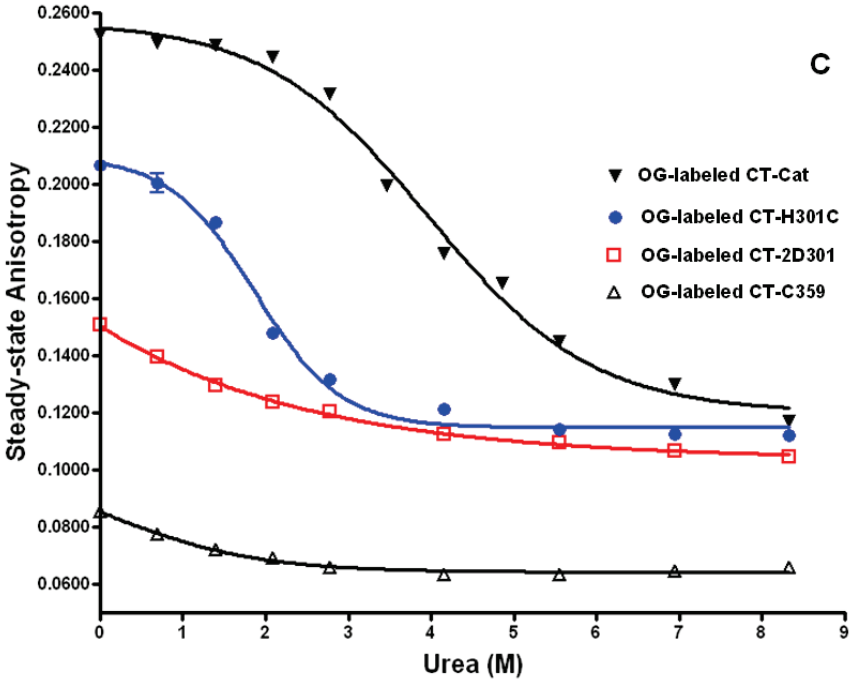


Figure 3.9 – Urea denaturation analysis revealed different structural order among CT domains
OG-labeled CT samples were treated with the indicated concentrations of urea for 5 min and their steady-state anisotropies were analyzed at 20°C, as described in Methods.

3.2.8 Domain M is predicted to contain both ordered and disordered sub-regions

In another approach to investigate the structure of domain M, I investigated the disorder propensity of domain M using four different prediction programs (DISORPRED2, Poodle, DISOclust, and Disprot), with the full length rat CTα as the query (Figure 3.10). These programs were selected out of many others[142] for their accuracy in assigning the highest order to the regions within the CT catalytic domain with the lowest crystallographic B-factors[73]. Unlike the catalytic domain, none of the programs predict a rigid globular structure for the entire domain M. DISOclust and Disprot predicted most of domain M to be disordered, while Poodle and DISORPRED2 predicted that the first half of domain M is disordered. Importantly, all four predicted the region around 280-290 to contain higher order than the rest of domain M. This analysis agrees with the steady-state anisotropy results showing that OG at site C301 has slightly higher anisotropy than other positions in domain M, but not as high as the catalytic domain.

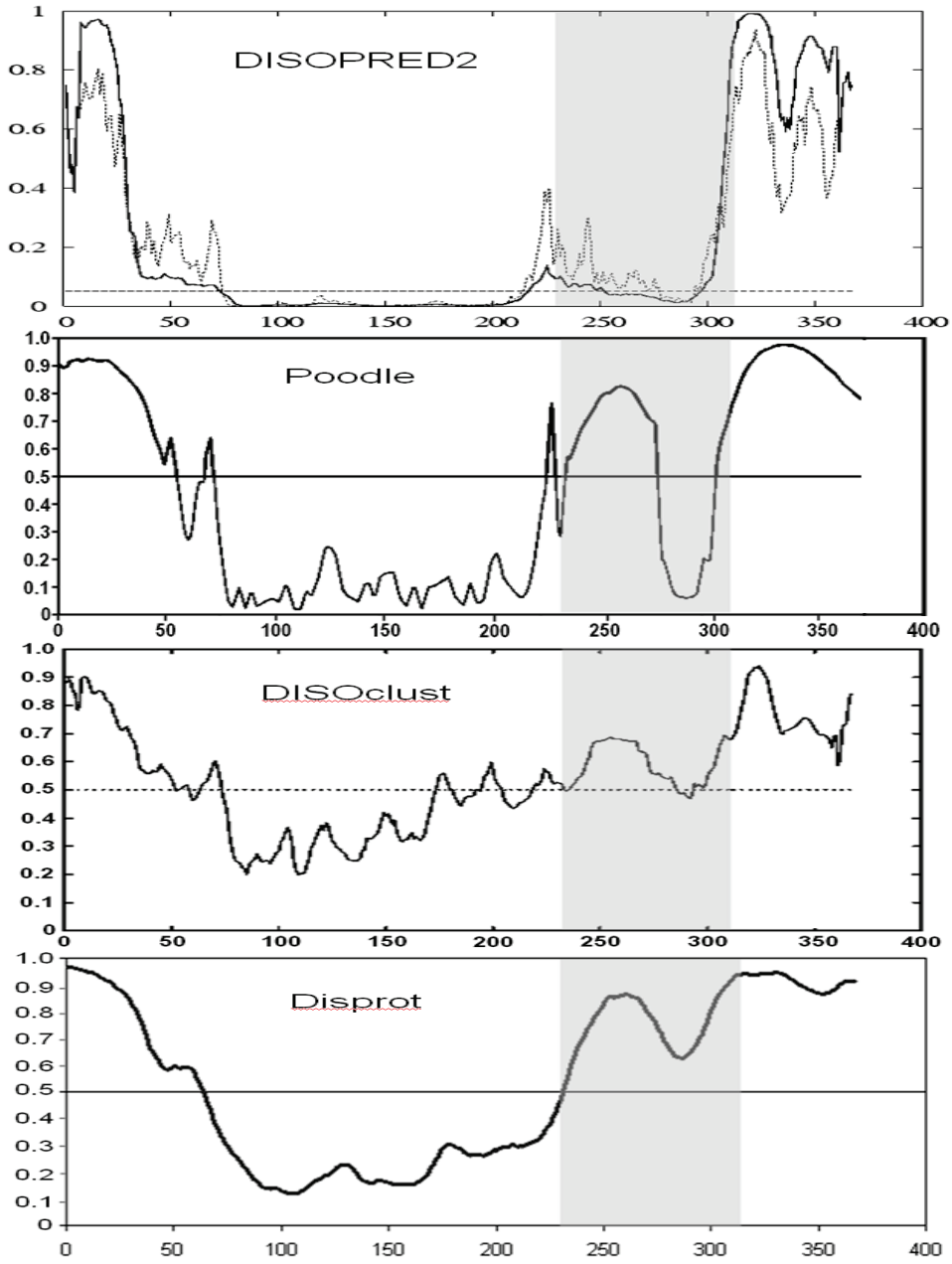


Figure 3.10 - CTα Domain M is predicted to contain both disordered and ordered regions

Full length CT was submitted to four different prediction server (DISOPRED2, Poodle, DISOclust, and Disprot) for order-disorder prediction. For each plot, the solid or solid horizontal line is the threshold set by the program, where values above the threshold are considered disordered. Shaded area shows residues in domain M. On each graph, Y-axis is disorder probability, and X-axis is residue numbers in CT

Next I examined the structural predictions for CTs from divergent species that have a conserved catalytic domain and a conserved mechanism for regulation of domain M by lipids[58,83,106,143,144]. I aligned CTs from different organisms with Clustal W (Figure

3.11), and found that relative to the highly conserved catalytic domain (residues 75 to 230), there is very little sequence conservation across species for domain M (residues 231 to 314). Since it is possible that two proteins have different sequences and yet share a highly similar structural fold and function, could the structural preference of CT M domains be conserved across different species? As a solved structure for soluble form of domain M is not available for any of these CT, I examined this possibility by analyzing the disorder propensity for the M domains of each CT with DISOClust (Figure 3.12). For each CT Disoclust predicts the first half of domain M to be disordered. This is followed by a distinct increase in order in all CTs. For rat CT, Drosophila, and C. elegans this occurs at around the FLEMF region identified earlier (285-295 for rat; 276-286 for Drosophila; 271-281 for C. elegans). For both yeast and P. Falciparum, this occurs near the FLEMF region (FLEMF region is 312-322 for yeast; 311-321 for P. Falciparum). Finally, the disorder increases sharply at end of domain M, toward the beginning of domain P. These similar structural tendencies suggest that the structural malleability of domain M may have been advantageous for lipid regulation of CTs, and thus remained conserved through evolution.

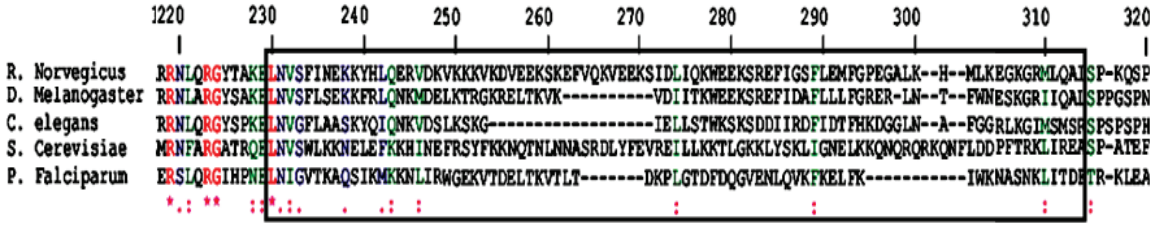


Figure 3.11 – Sequence alignment of domain M among CTs from different species

CT from R. Norvegicus (CTa, Uniprot ID:P19836), D. Melanogaster (Q9W0D9), C. Elegans (P49583), S. Cerevisiae (P13259), and P. Falciparum (P49587) were aligned using ClustalW. Residue numbers are the numbering for rat CTα. Boxed is the approximate region for domain M: R. Norvegicus = 230-314; D. Melanogaster = 233-305; C. Elegans = 234-300; S. Cerevisiae = 258-344; P. Falciparum = 263-333).

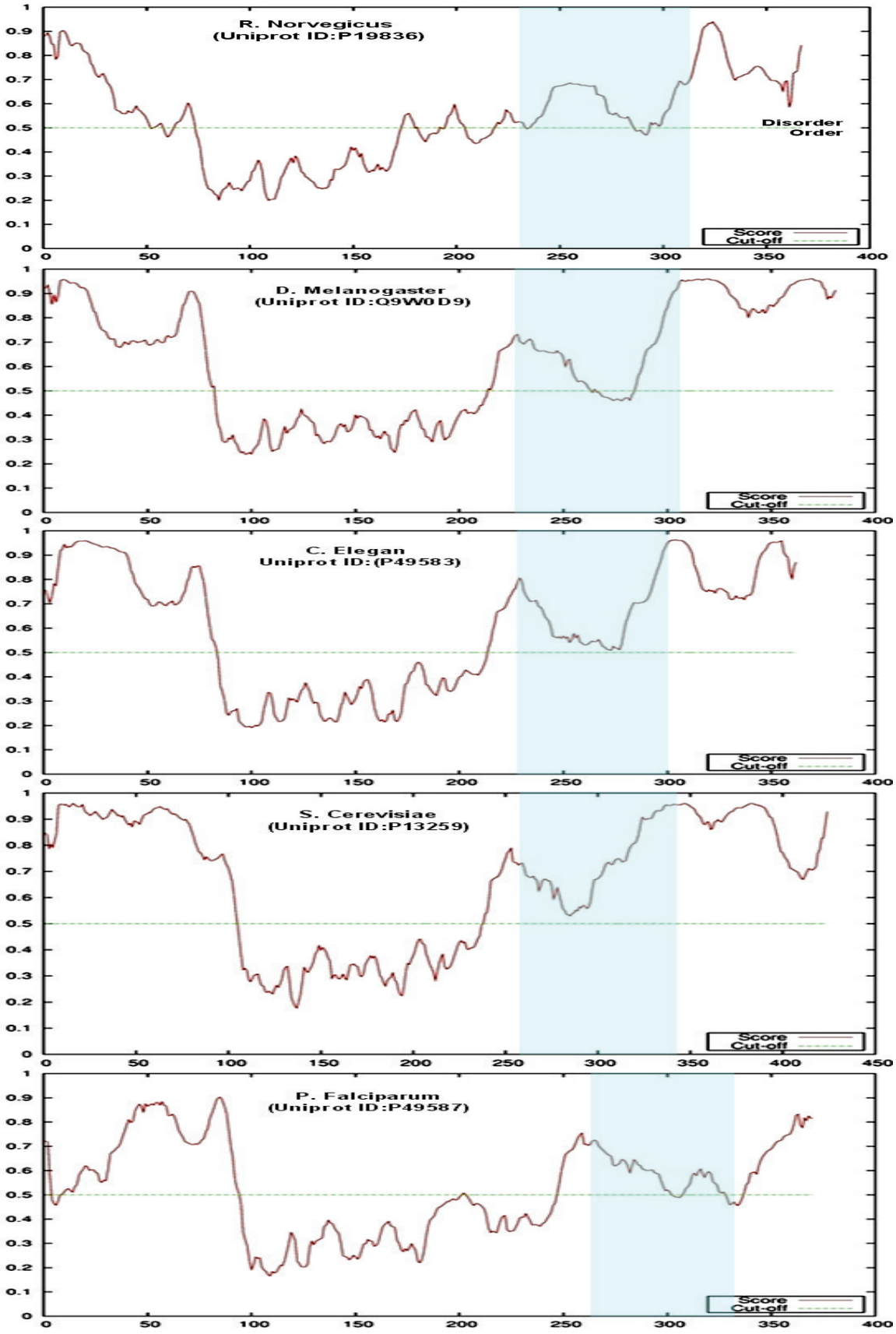

R. Norvegicus
(Uniprot ID:P19836)
Disorder
Order
Score
Cut-off
D. Melanogaster
(Uniprot ID:Q9W0D9)
Score
Cut-off
C. Elegan
Uniprot ID:(P49583)
Score
Cut-off
S. Cerevisiae
(Uniprot ID:P13259)
Score
Cut-off
P. Falciparum
(Uniprot ID:P49587)
Score
Cut-off

Figure 3.12 – Structural disorder prediction for CTs of divergent species

CT from different organisms were analyzed using DISOClust disorder prediction program. For each plot, the dashed line is the threshold set by the program, where values above the threshold are considered disordered. On each graph Y-axis is disorder probability, and X-axis is residue numbers in CT. Approximate regions for M domains (based on ClustalW alignment) are shaded in blue (R. Norvegicus: 230-314; D. Melanogaster: 233-305; C. elegans: 234-300; S. Cerevisiae: 258-344; P. Falciparum: 263-333).

3.3 Discussion

The fluorescence anisotropy and denaturation analyses have provided strong evidence that domain M is neither a well-folded globular domain nor a region lacking in structure. The anisotropy values at 5 different sites were similar, and corresponded to rotational frequencies that were only 2-fold faster than sites in the folded catalytic domain, but 2.5-fold slower than the site at the C-terminal tail. The urea denaturation analysis also supported a partially-folded structure for domain M, in that the unfolding was cooperative. The structure of domain M was dependent on its contact with domain C, as the segment in the vicinity of a mutation that prevents contact near the C-terminius of domain M (^{289}DLEMD^{293}) became very disordered and non-cooperative in its unfolding.

Similarity in the anisotropy values and response to urea for all sites in domain M (except 245) argue that the malleability of domain M is spread out over the entire region, with the highest tendency for order at site 301. Since order appeared to depend on interaction with the catalytic domain at least around the FLEMF motif, similar behaviour is expected for other sites that show inter-domain contacts. There may be multiple structurally inducible sites in domain M that contact the catalytic domain.

3.3.1 Implications for silencing models:

The ability of domain M to form disulfide bridges and an unfocussed intra-domain photo-cross-linking pattern in the soluble form (see Chapter 2) had suggested that domain M possesses some degree of structural flexibility. While disproving the Rigid Contact Model, these results could be used to support the Disorder-inhibition Model discussed in the Introduction, Chapter 1. In the Disorder-inhibition Model, silencing of the catalytic domain arises from disordering of the active site by the instability in helix E,

which is propagated from a very disordered and flexible domain M. According to this model the observed weak contacts between M and C would result from random collisions of highly flexible M domains located near the base of the active sites of the CT dimer. The lack of an effect on activity upon disruption of contact from any one site in domain M would also be consistent with the Disorder-inhibition model, because specific interactions are not required.

What is not compatible with the Disorder-inhibition model is the anisotropy/denaturation findings, showing that the entire domain M, although more disordered than the rigid catalytic domain, is much more folded than the C-terminal tail. In fact, the anisotropy and urea sensitivity of most sites in domain M resemble that of C207 – in the middle of helix E, the terminal helix in the catalytic domain. In support of this, limited protease digestion pattern showed resistance to digestion (i.e. inaccessibility) in the region between residues 286 and 300, likely due to the foldedness of this sub-region. In fact, comparative CD analysis tentatively suggests that domain M possesses some β-strand components in the soluble form, which are transformed, along with random coils, into α-helical components upon membrane binding[104].

Anisotropy analysis supported the MTAC model in which alternating local contacts of M with C suppress catalysis. There are two types of MTAC models for silencing of the catalytic domain by domain M (Figure 3.13). Model A describes a domain M that contains a firm structure that can slide about to create alternative auto-inhibitory contacts between various parts of domain M and the catalytic domain. This sliding mechanism is made possible with a flexible region at the N-terminus of domain M or at the hinge between helix E and domain M. In this model, the aspartate mutations at FLEMF cause a dislocation and barring of that region from interaction with the catalytic domain, leading to an apparent increase in disorder due to lower restriction on the rotational freedom of the fluorescence probe. In Model B, domain M does not possess a firm, fixed structure by itself, but portions of M gain temporary structure/order upon transient contacts with the catalytic domain. Unlike the sliding model, the aspartate mutations at FLEMF disrupt contact with the catalytic domain, resulting in the disruption of induced ordering at that region.

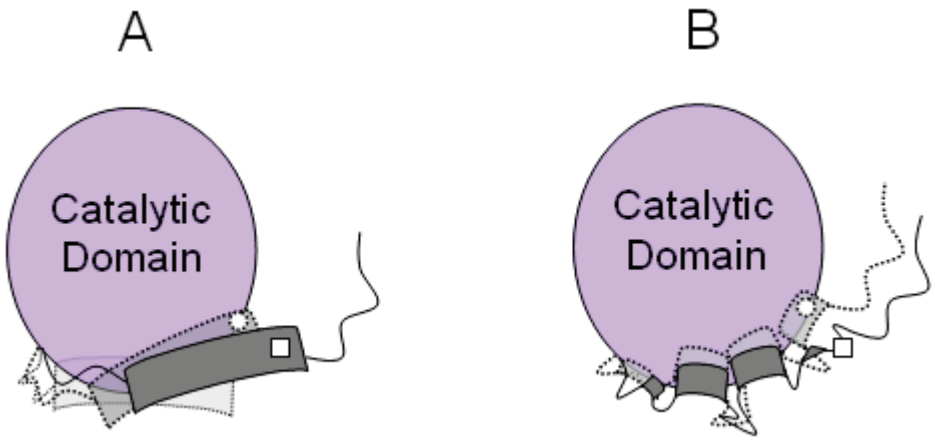


Figure 3.13 - Possible modes of contact for domain M

In scenario A, domain M (grey) is a stable structure with its N-terminus and/or the linker connecting domain M and the catalytic domain being highly flexible. Different parts of domain M, such as the phenylalanines in FLEMF (o), take turns making contact with the catalytic domain (purple) in a sliding motion. With the mutation to aspartate (□), repulsion dissociates it from the catalytic domain, reducing the restriction on fluorophore movement.

In scenario B, only certain sub-regions of domain M (grey) form contact with the catalytic domain, and obtain transient structure upon contact. The FLEMF segment, when mutated to DLEMD, becomes disordered since contact is disrupted.

The results of urea denaturation analysis favour MTAC Model B. The disruption of contact by the mutation at the FLEMF region leads to a dramatic melt down of structure in that region, as indicated by complete loss of resistance to unfolding induced by urea. The requirement for contact with domain C to create local structure in domain M may be explained by an induced fit mechanism. In this mechanism, domain M is devoid of structure by itself, but its folding is enabled (catalyzed) by the binding surface on domain C. An alternative explanation is that the binding surface in domain C selects for a pre-formed conformation from the population of disordered and partly folded domain M conformers[145]. In fact, comparative CD-analysis has revealed that domain M is 37.2 % disordered, in the soluble form[104]. CD analysis of the tail domain (M+P) in the absence of lipid or the catalytic domain reveal a value of 35% disordered (MSc thesis, Jospeh Lee, 2011).

The urea denaturation profiles suggest an unfolding behaviour for domain M reminiscent of proteins with malleable molten globular structure. Molten globular structures unfold at lower concentration of denaturant, but retain some cooperativity of the unfolding process[26]. Having a malleable property is advantageous as it allows domain

M to bind to different partners of completely different shape, size and physical-chemical properties. In the soluble form, domain M can form-fit with a site in the catalytic domain. However, a distinct highly alpha-helical conformation is induced by domain M's other contacting partner, the membrane surface. Flexibility and malleability are critical properties which enable domain M to sweep out the membrane surface searching for a suitable site that will facilitate its folding into an α-helix, and its lipid insertion that leads to CT activation.

Urea denaturation analysis also revealed subtle differences between sub-regions of domain M. In particular, the region around residue 245 appeared to be less ordered, as it did not possess a well-defined transition of unfolding as seen for all other sub-regions of domain M analyzed. Instead, it behaved similarly to the highly disordered C-terminal tail (Figure 3.9). Agreeing with this observation, this region had a higher disorder prediction compared to the other sub-regions of domain M (Figure 3.10). The region around residue 245 contains the highest concentration of charged residues in domain M (Figure 3.14), which are more commonly found in disordered regions of proteins. In comparison, the C-terminal portion of domain M contains mini-clusters of hydrophobic residues that in an aqueous environment would experience pressure to become buried within a folded structure. Disorder from the N-terminal sub-region might be one of the main contributors in the flexibility of domain M. In addition, this sub-region near residue 245 may facilitate the transient nature of the interaction between the catalytic domain and domain M, for the benefit of quicker activation. Similar to the fly-casting model[146], lose structures and transient contacts can both greatly enhance the range and efficiency of searching for a suitable site for folding and insertion on the membrane surface.

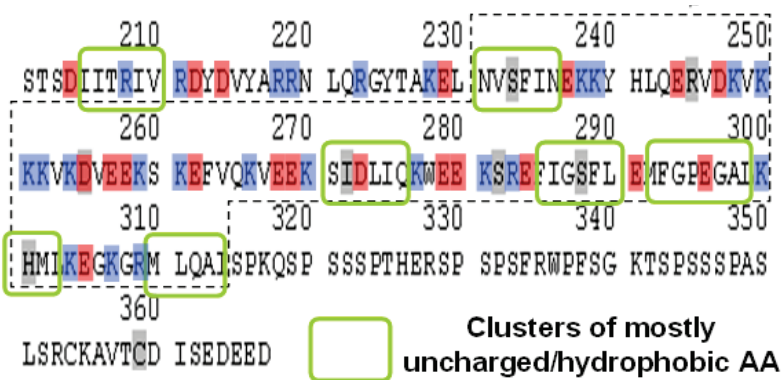


Figure 3.14 – Differences in disorder propensity within domain M may be influenced by clusters of residues of certain polarities

Domain M (residues ~231-314) is included in dashed box. Charged residues from residues 201 to 314 are highlighted in red (for acidic residues) and blue (for basic residues). Gray highlights residues that have been mutated to cysteines in single cysteine CT variants.

With the many advantages offered by the MTAC model for auto-regulation, this mechanism may have emerged early in the evolution of CTs. Although conservation does not appear to be accomplished through primary sequence, most CTs analyzed here have been reported to be lipid responsive[58,83,106,143,144]. It is likely the auto-silencing mechanism is shared as well. In support of this, the disorder propensity appeared to be similar (high disorder in the N-terminal region of domain M followed by low disorder near the C-terminus of domain M; Figure 3.12) in CTs spanning single-celled protists to vascular plants to mammals. Recently proof of a conserved mechanism for auto-regulation emerged with the finding that a chimeric enzyme composed of the catalytic domain of rat CTα fused to the C-terminal regulatory tail of C. elegans or D. melanogaster behaved nearly identically to the wild-type rat enzyme, i.e. was silenced in the absence of lipids and stimulated to similar extents by addition of anionic lipid vesicles (Ding and Cornell, unpublished).

From comparative CD-analysis[104], NMR spectroscopy[96], and photo-cross-linking, it is well-established that domain M transitions into a long α-helix in the presence of activating lipids. Although this transition was expected to yield an increase in anisotropy (ie. structure becomes more ordered) with lipid binding, this was not observed (Table 3.2). One explanation for this observation is that in the soluble form, the fluorophores are restricted in rotational freedom due to entrapment from the interaction between domain M and the active site. Upon membrane binding, the hydrophilic fluorophore would

extend into the aqueous phase, where it would experience more freedom of rotation, as it is no longer interacting with the polypeptide. On the other hand, residue 245 is in a very polar section of domain M that may remain more superficially bound to the membrane surface than the C-terminus of domain M. Upon membrane binding the fluorophore attached to residue 245 may maintain substantial contact with the polypeptide. With formation of an α-helix, site 245 exerts further restriction onto the attached fluorophore and thus leading to the observed increase in anisotropy.

3.4 Methods

3.4.1 OG labeling of CT variants

His-tagged single cysteine variants (1.2 nmol to 5 nmol) in dialysis buffer (20 mM NaH_2PO_4, 150 mM NaCl, 1 mM DTT, 0.25 mM TX-100, pH 7.4), were bound to 100 µl of prepared Ni-beads for 20 min at room temperature. After removing supernatant, Ni-beads were re-suspended in 100 µl of OG-IAA (various concentrations, ranging from 0.24 nmol to 0.05 nmol) in equilibration buffer (50 mM NaH_2PO_4, 500 mM NaCl, 0.25 mM TX-100, pH 8.0), vortexed briefly, and centrifuged to pellet beads. Supernatant containing excess OG was removed. Beads were washed ten times with equilibration buffer, and once with wash buffer (10 mM Tris, 150 mM NaCl, 0.25 mM TX-100, pH 7.4). All washes were of 100 µl volume. OG-labeled CT variants were then recovered from Ni-beads with 4 x 100 µl elution buffer (350 mM imidazole, 10 mM Tris, 150 mM NaCl, 0.25 mM TX-100, pH 7.4). Recovered sample was further diluted with 400 µl of the wash buffer, and aliquoted for different analyses.

3.4.2 Determination of total OG percent labelling

OG concentration was obtained from absorbance wavelength scan from 400 to 550 nm. A background scan was performed on sample buffer (10 mM Tris, 150 mM NaCl, 0.25 mM TX-100, 175 mM imidazole; pH 7.4). Background-corrected peak absorbance at around 491 nm was used to calculate OG concentration using Beer's law. Protein concentration was determined by Bradford assays. The percent of CT labeled by OG was calculated as: 100% x (OG concentration)/(protein concentration).

3.4.3 Steady-state anisotropy measurements

OG-labeled CT sample (110 μl) in a quartz curvette was equilibrated to 20^{o}C in the thermal-controlled sample chamber for 5 min. Anisotropy was measured with the spectroflourimeter using steady-state mode, with either of the two settings. Setting A was used for OG fluorescence: 491 nm for excitation wavelength; 522 nm for emission wavelength; 8 nm slit widths for excitation and emission wavelengths; 1 s integration time. Setting B was used for tryptophan fluorescence: 295 nm for excitation wavelength; 348 nm for emission wavelength; 1 nm slit width for excitation wavelength; 12 nm slit width for emission wavelength; 1 s integration time). 30 readings were recorded per sample aliquot. Two aliquots were analyzed for each sample.

3.4.4 Urea denaturation analysis

Protein sample (110 μl) was transferred to a quartz curvette and equilibrated to 20^{o}C in the thermal-controlled sample chamber for 5 min. Anisotropy was measured as described above for undenatured sample. Sample was denatured sequentially with solid urea, vortexed, incubated in the sample chamber for 5 minute, and analyzed at each urea concentration. Volume increase was miniscule with each addition of solid urea. The estimated error for the weighing the urea (ultra-pure grade) was around 5 $\pm$ 0.2 mg.

4: Impact of my thesis

My work has provided valuable information concerning the inter-domain contact sites that may be crucial in the auto-inhibition of CT. In particular, I generated a low-resolution map of possible docking sites on the catalytic domain for all sub-regions of domain M. The approach I used, photo-cross-linking combined with MALDI-MS enabled identification of weak interactions that have escaped detection by other techniques, such as inhibition of catalysis by the regulatory domain supplied *in trans*, pull-down assays, and limited protease digestion pattern analysis.

While shortcomings such as selection of false positives, interference from photo-cross-linkers, and the low resolution inherent in the photo-cross-linking approach are not to be minimized, evidence for domain M – domain C contact obtained using this method was validated by steady-state anisotropy experiments. Specifically, when the contact with domain C was disrupted by mutation in one sub-region of domain M (FLEMF), the local structure at that site in M dissolved.

More detailed information about domain M dynamics was obtained using urea denaturation. While urea denaturation in combination with fluorescence anisotropy has been used widely in studies involving protein folding[147-151], I extended this strategy using fluorescence spectroscopy to characterize differently folded regions of the same protein. While urea-induced unfolding of separate domains within a protein has been analyzed by NMR methods[152], the use of fluorescence anisotropy to accomplish this goal is novel. Importantly, this simple technique is capable of differentiating foldedness between regions that are merely 12 residues apart in the same protein domain. Because of the relatively fast dynamics of domain M that have emerged from both fluorescence and cross-linking, a high-resolution structure of a CT containing domain M via crystallography may be unachievable. The methods utilized in this study proved to be a workable alternative.

A new working model, MTAC (Figure 4.1), has emerged from our experimental data. In this model, auto-inhibition by domain M is achieved through contacts to the catalytic domain, either directly to the active site, and/or to helix E that situates right below the active site, resulting in subtle active site restructuring that reduces catalytic efficiency. Contacts from domain M are likely to be transient, and driven mainly by hydrophobic interactions, since it is this pattern of hydrophobic residues in domain M that is conserved across diverse species that show domain M mediated silencing. The inhibitory contacts could alternate between multiple parts of domain M, and between the two M domains in a cooperative manner. Immediately upstream of helix E lie key points of contact with the substrate CTP at residues 202, 200, 197, and 196. At this time we do not know how the contact with domain M segments alters the positioning of these residues, but it remains our working hypothesis that these positions in helix E are affected to slow catalysis. Helix E is also a key participant in forming the dimer interface. Taneva et *al* have provided evidence for cooperativity between the active sites of the CT dimer[94]. Hence silencing by MTAC with helix E may impact on that cooperativity. Upon lipid binding, domain M is released from interacting with helix E and the active site, allowing catalysis to proceed.

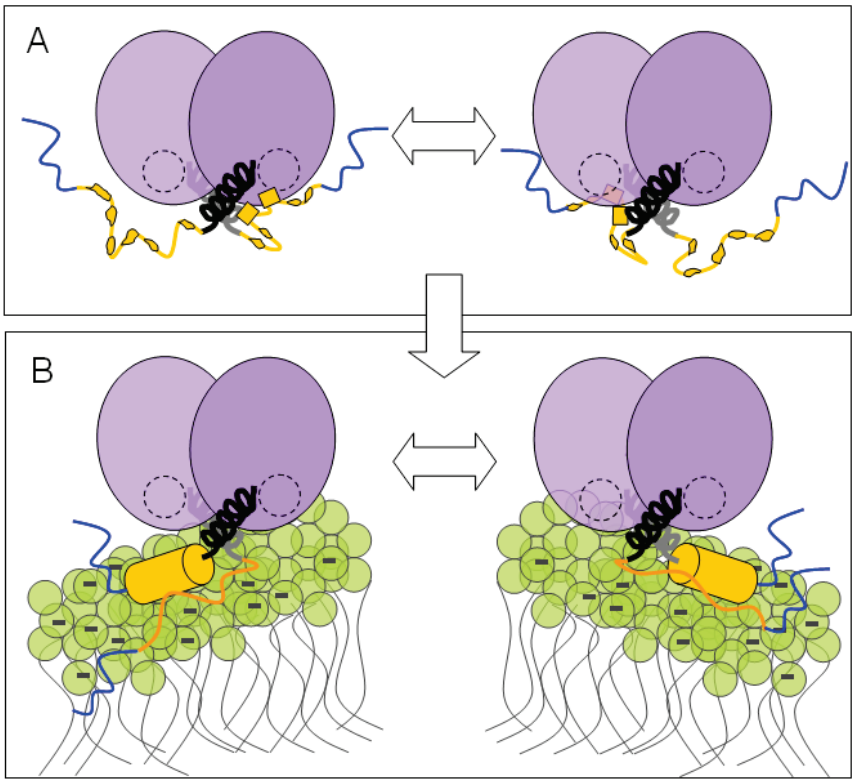


Figure 4.1 - Multiple Transient Alternating Cooperative Contact model for CT auto-inhibition

A. Auto-inhibition in the soluble form. Parts of the M domains (unstructured shapes in yellow) in CT may form transient contact with active site (dashed circle in purple) and helix E (solid black), which induces local order in M (yellow solid box). The remainder of domain M is disordered (yellow solid line). This type of contact is only meta-stable and alternates between different sub-regions of domain M. **B.** In the presence of activating lipids, domain M inserts itself into the lipid bilayer, making it impossible to form any inhibitory contacts to the catalytic domain. The model shown here incorporates the idea presented in Taneva *et al* that the M helices bind membranes one at a time.

The disorder predictions for domain M from diverse species, and the mutational analysis of multiple hydrophobic residues in domain M from C. elegans CT (refer to Chapter 2 discussion) tentatively suggest that the MTAC model may be a conserved mechanism of auto-inhibition, providing a facile mechanism for both silencing and membrane binding, and the interconversion between these two functions. While this type of mechanism involving structural malleability differs from the more common auto-inhibition involving rigid, auto-inhibited complexes solvable by crystallagraphy (eg. SRC kinase, SNARE, Ets-1)[9,10,19], would MTAC be a novel mechanism unique to lipid-regulated cytidylyltransferases? Or can this mechanism be applied to other proteins whose auto-inhibitory mechanisms have yet to be uncovered?

A similar mechanism involving a disordered but structure-inducible auto-regulatory domain has been reported for the CFTR[27]. This study utilized NMR exclusively to examine the auto-inhibitory mechanism as X-ray crystallographic structures are more difficult to obtain, and would provide only a static picture of one conformation in a dynamic ensemble. The dynamic engagement of the regulatory region (R) and nucleotide-binding domains (NBDs) of CFTR was examined. Similar to CT's domain M, this intrinsically disordered R region appears to sample multiple, heterogenous, and rapidly-inter-converting conformations with varying degrees of compactness and secondary structural features. Multiple sites from the R region bind transiently as α-helices to NBDs to exhibit inhibition. In the activated state, phosphorylation at non-specific sites in the R region lowers the amount of helical structure and reduces interaction with the NBD. Unlike the approach I took with CT domain M, this study was performed with two domains that had been artificially separated. The observed R-NBD interaction was examined *in trans*, whereas a C-M interactions in CT was not observable *in trans* (from activity analysis and pull-down assay). Thus, the silencing interaction in CT is much weaker and more transient. Interestingly, NBDs act as a heterodimers, and the dimer interface has been proposed as one of the possible docking sites for the R region. However, unlike CT where the dimers remain mostly intact through activation-inactivation cycles, NBD is proposed to be inhibited by the R region that prevents proper dimer formation for full channel activity.

Features of the MTAC model have also been demonstrated in the classic Ball-and-Chain model for the inactivation domains of voltage gated potassium channels[124,153,154]. These integral membrane proteins cycle among three states: closed, open, and inactive. The inactive state is achieved through a highly flexible "chain" that carries out a random search until the "ball" plugs the open channel, effectively blocking the channel pore. Both hydrophilic and highly conserved non-polar residues in the "ball" region are important in auto-inhibitory interaction[153,155], and the length of the "chain" is important in controlling rate of inactivation. As a tetrameric structure is required to form a single pore, four inactivation domains are available for inactivation[154,156,157]. However, analysis of a tetramer with only one inactivating plug showed that binding from only one of the four identical inactivating domains is sufficient to block the pore[158,159]. Thus the

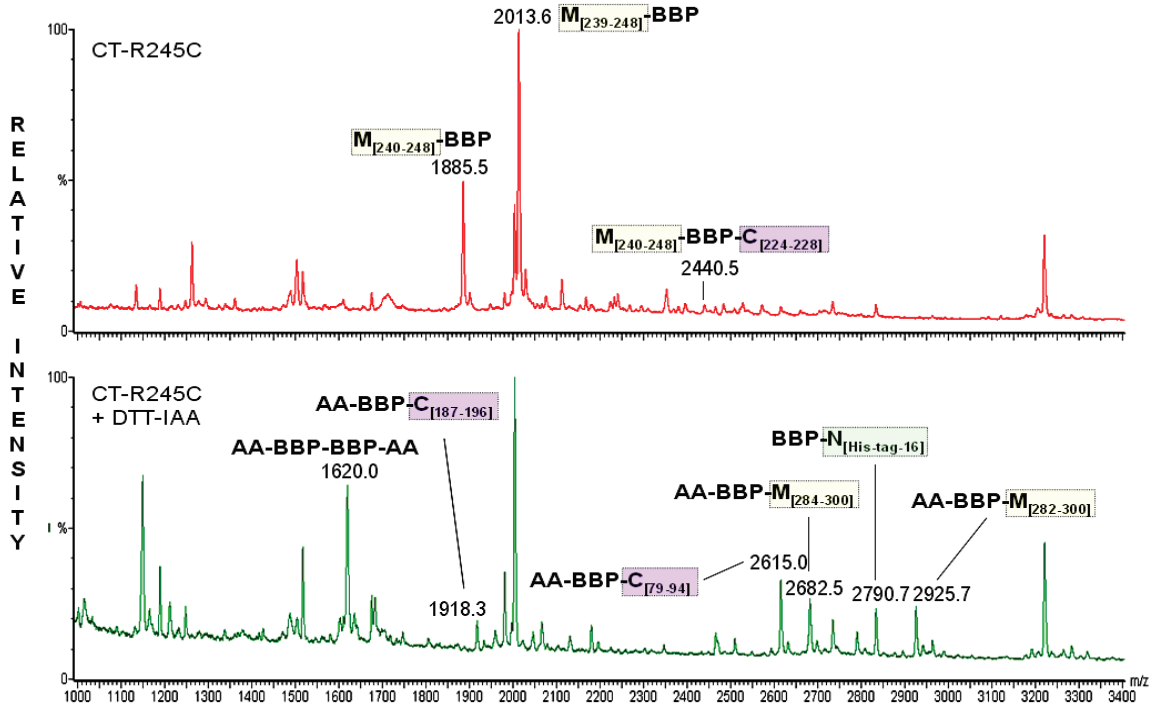
binding of each inactivating segment occurs in a mutually exclusive manner, and the inactivation segments appear to be working in an alternating fashion. CT auto-inhibition works similarly with sub-region(s) possessing higher probability for disorder (eg. region around residue 245) that may work as the small, flexible "chain" to swing the structurally-inducible inactivating regions (eg. hydrophobic clusters identified in Figure 2.14) for proper contact with docking sites in the catalytic domain. The inhibitory sub-regions from both M domains can have equal potential to form contact with the same sites on the catalytic domain, and thus work in an alternating fashion.

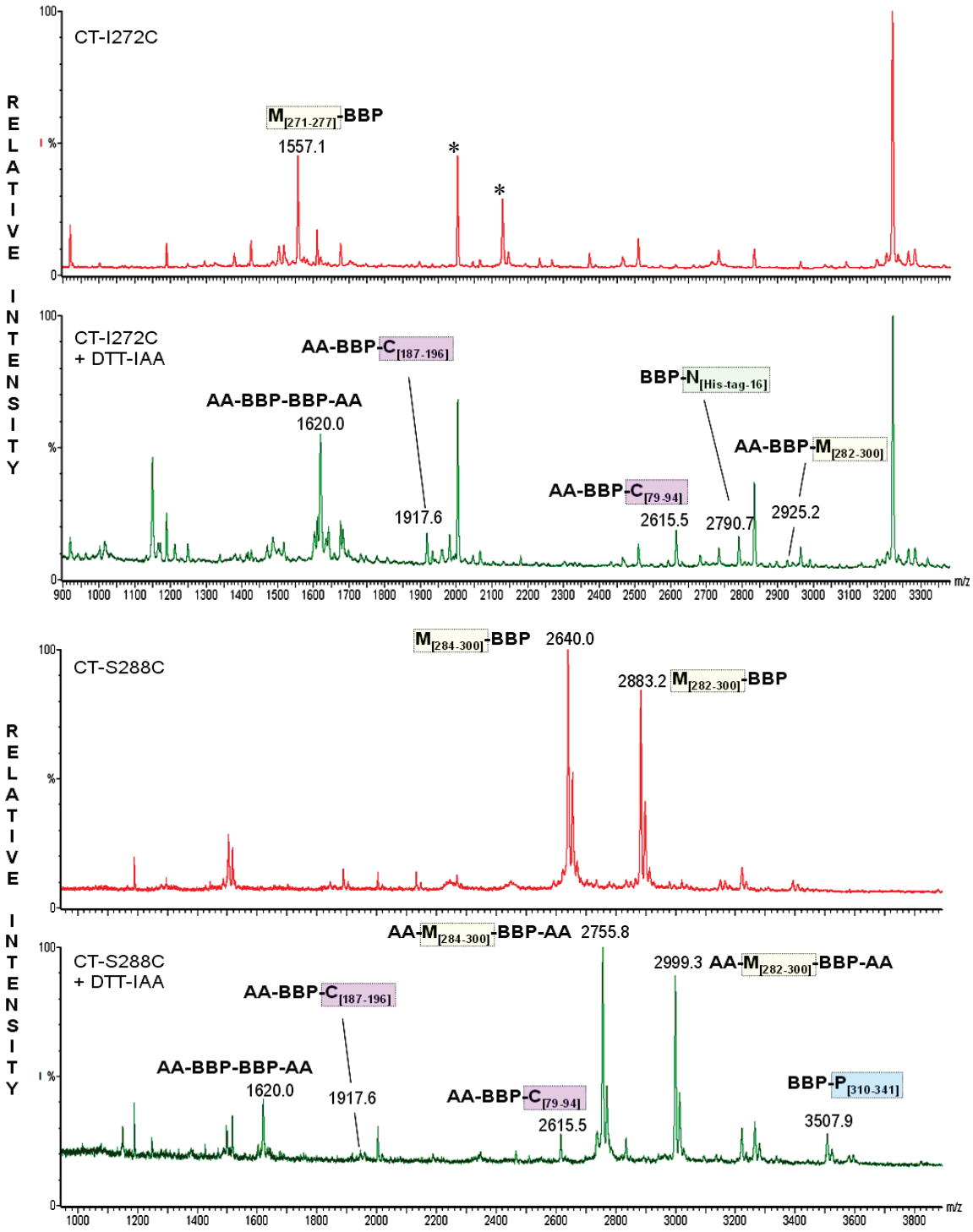
When compared to better characterized auto-inhibitory mechanisms, the MTAC model for auto-inhibitory mechanism appeared to be novel at the first glance. However, similarities shared with the auto-inhibition mechanisms for CFTR and potassium ion channel suggest that at least some of the key elements of MTAC model can be applied generally to other regulated proteins. Analysis performed in this thesis can offer helpful hints for deciphering the mechanisms of inhibition for other proteins, especially the ones with structures that tend to escape current methods of determination.

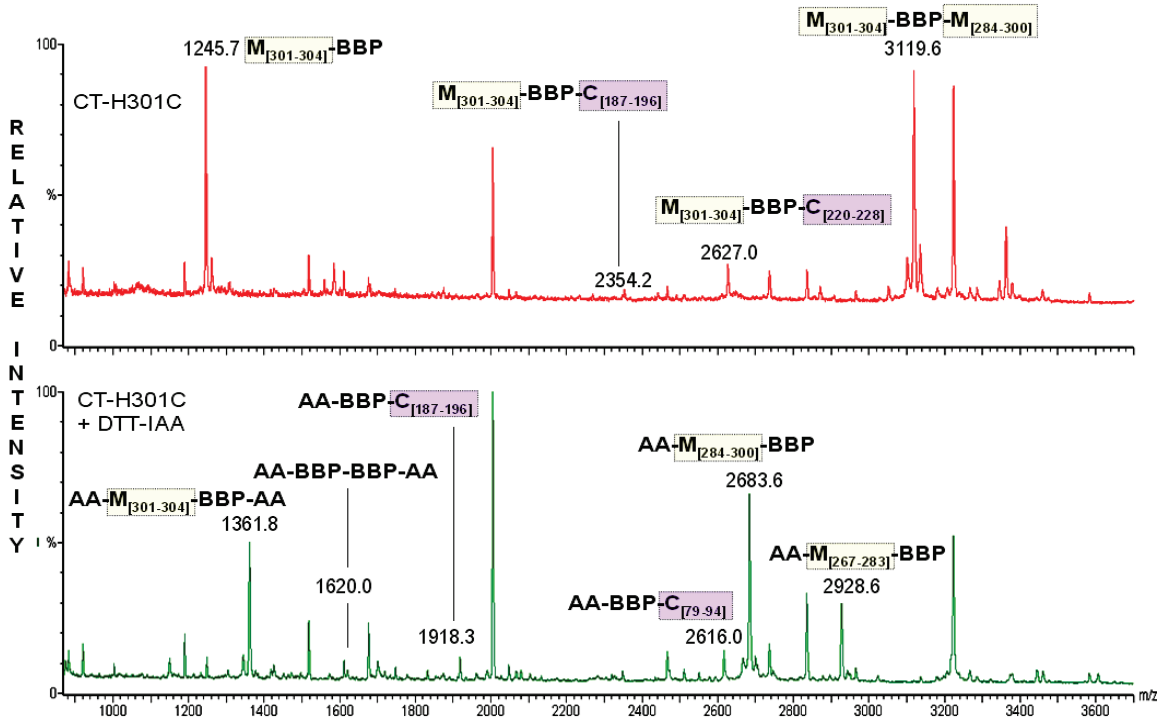
Appendices

Appendix I - Representative spectra sets from photo-cross-linking analysis of CT variants

All mutants were BBP-labeled, photo-cross-linked, tryptic digested, avidin purified, and analyzed on MALDI-MS as described in Methods in Chapter 2. Best representative spectra set for each mutant is shown, with a few of the photo-cross-links identified. Relative intensities of some junk peaks that are not unique to photo-cross-linked samples (eg. appeared in non-irradiated sample spectra) can be quite high. Two examples (*) can be found in the spectra set for CT-I272C.




RELATIVE INTENSITY
CT-I272C
M[271-277]-BBP
1557.1
*
*
CT-I272C
+ DTT-IAA
AA-BBP-C[187-196]
AA-BBP-BBP-AA
1620.0
BBP-N[His-tag-16]
AA-BBP-M[282-300]
AA-BBP-C[79-94]
1917.6
2615.5
2790.7
2925.2
m/z
M[284-300]-BBP 2640.0
CT-S288C
2883.2 M[282-300]-BBP
AA-M[284-300]-BBP-AA 2755.8
2999.3 AA-M[282-300]-BBP-AA
CT-S288C
+ DTT-IAA
AA-BBP-C[187-196]
AA-BBP-BBP-AA
1620.0
1917.6
AA-BBP-C[79-94]
2615.5
BBP-P[310-341]
3507.9
m/z


RELATIVE INTENSITY
CT-H301C
1245.7 M[301-304]-BBP
M[301-304]-BBP-C[187-196]
2354.2
M[301-304]-BBP-C[220-228]
2627.0
M[301-304]-BBP-M[284-300]
3119.6
CT-H301C
+ DTT-IAA
AA-BBP-C[187-196]
1918.3
AA-BBP-BBP-AA
1620.0
AA-M[301-304]-BBP-AA
1361.8
AA-M[284-300]-BBP
2683.6
AA-M[267-283]-BBP
2928.6
AA-BBP-C[79-94]
2616.0
m/z

Appendix II: Ambiguously identified photo-cross-linked peptides from MALDI-MS for each CT

CT-R245C no lipid		
Identity	Average Observed Mass (Da)	Sequence of peptide from contact
M[240-248]-BBP-C[219-223]	2572.3	RNLQR
M[240-250]-BBP-N[13-15]	2572.6	KRR
M[240-250]-BBP-N[14-16]	2572.6	RRK
M[239-250]-BBP -N[14-15]	2572.3	RR
M[240-251]-BBP-N[14-15]	2572.3	RNLQR
M[240-248]-BBP-C[219-223]	2572.3	RNLQR
AA-BBP-P[342-355]	2682.4	TSPSSSPASLSRSK
AA-BBP-M[284-300]	2683.2	EFIGSFLEMFGPEGALK
M[240-248]-BBP-M[253-259]	2748.5	VKDVEEK
M[240-250]-BBP-M[255-259]	2748.7	DVEEK
M[240-248]-BBP-C[187-196]	2993.8	EAGMFAPTQR
M[240-250]-BBP-C[141-147]	2993.0	YVDEVVR
M[239-248]-BBP-C[187-196]	3122.1	EAGMFAPTQR
M[239-250]-BBP-C[141-147]	3120.9	YVDEVVR
M[240-251]-BBP-C[141-147]	3120.9	YVDEVVR
AA-BBP-C[219-238]	3191.8	RNLQRGYTAKELNVSFINEK
AA-BBP-C[187-208]	3192.7	EAGMFAPTQRTEGISTSDIITR
AA-BBP-M[284-304]	3192.9	EFIGSFLEMFGPEGALKHMLK
AA-BBP-C[101-122]	3319.9	NLFPNTYLIVGVSSDELTHNFK
AA-BBP-M[219-239]	3319.9	RNLQRGYTAKELNVSFINEKK
M[239-248]-BBP-C[79-94]	3691.5	VYADGIFDLFHSGHAR
M[240-250]-BBP-P[342-355]	3690.2	TSPSSSPASLSR
M[240-248]-BBP-P[342-355]	3757.9	TSPSSSPASLSRSK
M[240-250]-BBP-N[14-20]	3756.8	MAKHHHHHHIEGR
M[240-248]-BBP-M[247-262]	3819.8	EFVQKVEEKSIDLIQK
M[239-248]-BBP-C[79-94]	3819.8	VYADGIFDLFHSGHAR
M[239-250]-BBP-P[342-353]	3818.1	TSPSSSPASLSR
M[240-251]-BBP-P[342-353]	3818.1	TSPSSSPASLSR
M[239-248]-BBP-P[342-355]	3886.2	TSPSSSPASLSRSK
M[239-250]-BBP-N[His-tag-His-tag]	3884.7	MAKHHHHHHIEGR
M[240-251]-BBP-N[His-tag-His-tag]	3884.7	MAKHHHHHHIEGR
M[240-250]-BBP-M[240-248]	3999.3	(K)YHLQECVDK(V)
M[240-248]-BBP-P[310-328]	3999.9	MLQAISPKQSPSSSPTHER
M[240-250]-BBP-P[336-353]	3993.1	WPFSGKTSPSSSPASLSR
M[239-250]-BBP-N[His-tag-14]	3994.7	SAMDAQSSAKVNSRKR
M[240-251]-BBP-N[His-tag-14]	3994.7	SAMDAQSSAKVNSRKR
M[240-250]-BBP-P[342-355]	3985.4	TSPSSSPASLSRSK
M[240-248]-BBP-P[310-328]	3983.9	MLQAISPKQSPSSSPTHER

M[240-248]-BBP-P[310-328]	4127.8	MLQAISPKQSPSSSPTHER
M[239-248]-BBP-P[310-328]	4128.2	MLQAISPKQSPSSSPTHER
M[240-248]-BBP-M[239-250]	4127.2	KYHLQECVDKVK
M[240-248]-BBP-M[240-251]	4127.2	YHLQECVDKVKK
M[240-248]-BBP-M[262-281]	4392.4	EFVQKVEEKSIDLIQKWEEK
M[240-248]-BBP-M[260-277]	4390.4	SKEFVQKVEEKSIDLIQK
M[239-248]-BBP-M[260-277]	4390.4	SKEFVQKVEEKSIDLIQK
M[240-250]-BBP-C[79-100]	4561.8	VYADGIFDLFHSGHARALMQAK
M[239-248]-BBP-P[318-335]	4561.8	QSPSSSPTHERSPSPSFR

CT-I272C no lipid		
Identity	Average Observed Mass (Da)	Sequence of peptide from contact
AA-BBP-M[267-277]	2102.5	VEEKSCDLIQK
AA-BBP-P[318-328]	2103.3	QSPSSSPTHER
AA-BBP-C[197-208]	2103.5	TEGISTSDIITR
AA-BBP-P[342-355]	2682.4	(K)TSPSSSPASLSRSK(A)
AA-BBP-M[284-300]	2683.2	(R)EFIGSFLEMFGPEGALK(H)
AA-BBP-N[43-61]	3035.5	(R)QPAPFSDEIEVDFSKPYVR(V)
AA-BBP-C[220-238]	3035.6	(R)NLQRGYTAKELNVSFINEK(K)
AA-BBP-C[101-122]	3319.9	(K)NLFPNTYLIVGVSSDELTHNFK(G)
AA-BBP-C[219-239]	3319.9	(R)RNLQRGYTAKELNVSFINEKK(Y)
BBP-P[310-355]	3594.2	(R)MLQAISPKQSPSSSPTHERSPSPSFR(W)
AA-BBP-C[70-94]	3596.1	(R)GTPSERPVRVYADGIFDLFHSGHAR(A)
M[271-281]-BBP-P[318-335]	4261.8	(K)QSPSSSPTHERSPSPSFR(W)
M[271-283]-BBP-M[246-261]	4261.6	(R)VDKVKKKVKDVEEKSK(E)
M[271-283]-BBP-P[342-355]	4261.7	(K)TSPSSSPASLSRSK(A)
M[271-277]-BBP-P[329-355]	4587.3	(R)SPSPSFRWPFSGKTSPSSSPASLSRSK(A)
M[271-281]-BBP-P[336-353]	4585.6	(R)WPFSGKTSPSSSPASLSR(S)

CT-S288C no lipid		
Identity	Average Observed Mass (Da)	Sequence of peptide from contact
AA-BBP-M[262-270]	1946.3	EFVQKVEEK
BBP-C[229-238]	1946.3	ELNVSFINEK
AA-BBP-P[310-328]	3212.3	MLQAISPKQSPSSSPTHER
AA-BBP-C[209-228]	3211.8	IVRDYDVYARRNLQRGYTAK
M[284-300]-BBP-C[95-100]	3302.0	ALMQAK
M[282-300]-BBP-N[34-36]	3303.0	VQR
AA-BBP-C[101-122]	3319.9	NLFPNTYLIVGVSSDELTHNFK
AA-BBP-C[219-239]	3319.9	RNLQRGYTAKELNVSFINEKK
AA-BBP-M[246-266]	3330.0	VDKVKKKVKDVEEKSKEFVQK
AA-BBP-P[336-353]	3330.0	WPFSGKTSPSSSPASLSR
M[284-300]-BBP-C[123-132]	3837.5	GFTVMNENER
M[284-300]-BBP-P[308-317]	3837.5	GRMLQAISPK
M[282-300]-BBP-P[329-335]	3838.3	SPSPSFR
M[284-304]-BBP-C[219-223]	3837.7	RNLQR
M[282-300]-BBP-N[His-tag-8]	3880.6	SAMDAQSSAK
M[284-304]-BBP-M[249-254]	3880.9	VKKKVK
M[284-300]-BBP-C[209-219]	4082.8	IVRDYDVYARR
M[282-300]-BBP-C[123-132]	4081.8	GFTVMNENER
M[282-300]-BBP-P[308-317]	4081.8	GRMLQAISPK
M[284-300]-BBP-C[79-94]	4446.2	VYADGIFDLFHSGHAR
M[282-300]-BBP-P[329-341]	4445.2	SPSPSFRWPFSGK
M[282-300]-BBP-M[249-261]	4446.4	VKKKVKDVEEKSK
M[282-304]-BBP-C[220-228]	4444.8	NLQRGYTAK
M[284-300]-BBP-C[123-140]	4794.5	GFTVMNENERYDAVQHSR
M[284-300]-BBP-N[14-33]	4794.6	RRKEVPGPNGATEEDGIPSK
M[282-300]-BBP-N[His-tag-15]	4793.6	SAMDAQSSAKVNSRKRR
M[284-304]-BBP-N[His-tag-His-tag]	4794.7	MAKHHHHHHIEGR
M[284-300]-BBP-C[123-140]	4794.5	GFTVMNENERYDAVQHSR

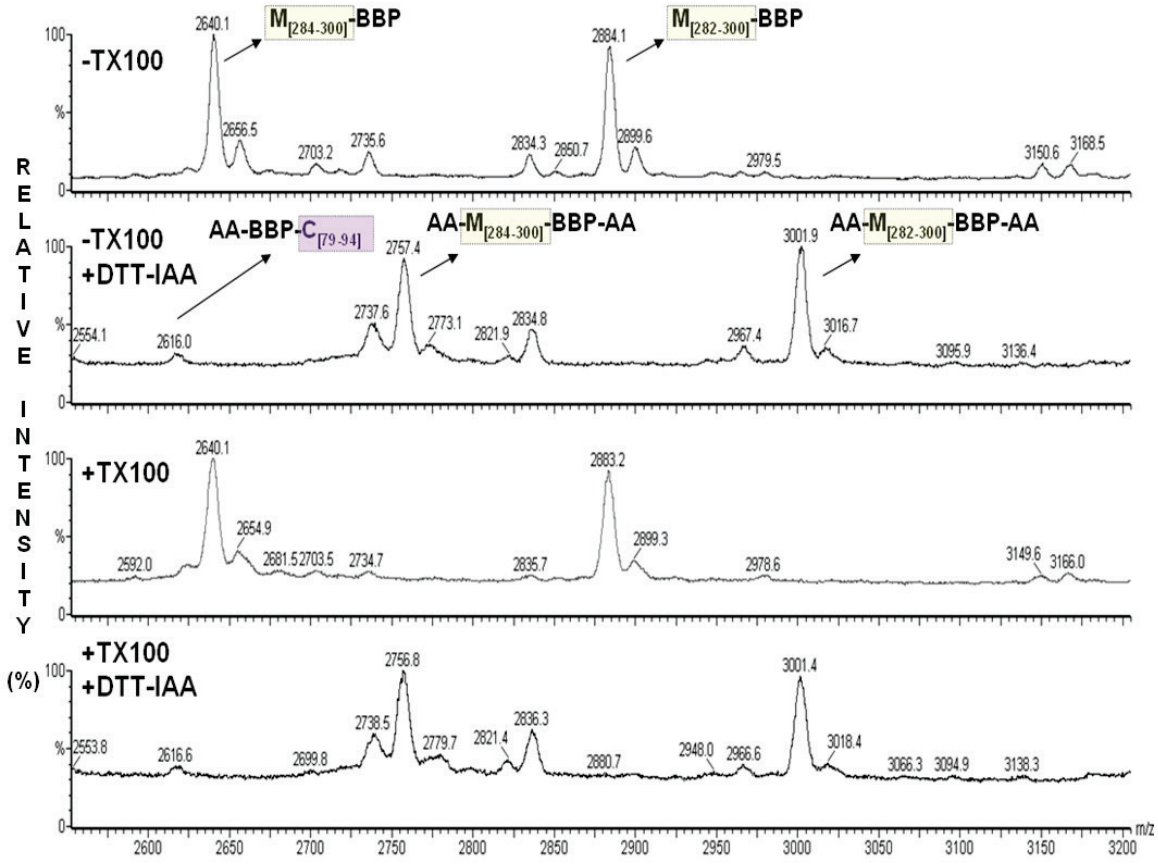
CT-H301C no lipid		
Identity	Average Observed Mass (Da)	Sequence of peptide from contact
BBP-M[260-270]	2104.5	SKEFVQKVEEK
AA-BBP-C[197-208]	2103.5	TEGISTSDIITR
AA-BBP-P[342-356]	2682.4	TSPSSSPASLSRSK
AA-BBP-M[284-300]	2683.2	EFIGSFLEMFGPEGALK
M[301-304]-BBP-C[79-94]	3067.9	VYADGIFDLFHSGHAR
M[301-304]-BBP-M[252-266]	3068.0	KVKDVEEKSKEFVQK
AA-BBP-C[101-122]	3319.9	NLFPNTYLIVGVSSDELTHNFK
AA-BBP-C[219-239]	3319.9	RNLQRGYTAKELNVSFINEKK
M[301-304]-BBP-N[17-36]	3343.2	EVPGPNGATEEDGIPSKVQR
BBP-M[282-304]	3343.1	SREFIGSFLEMFGPEGALKCMLK
M[301-304]-BBP-M[282-300]	3378.4	SREFIGSFLEMFGPEGALK
M[301-304]-BBP-P[318-335]	3378.0	QSPSSSPTHERSPSPSFR
M[301-304]-BBP-M[282-300]	3394.4	SREFIGSFLEMFGPEGALK
M[301-304]-BBP-P[318-335]	3394.4	QSPSSSPTHERSPSPSFR
M[301-304]-BBP-M[229-245]	3394.4	ELNVSFINEKKYHLQER
AA-BBP-N[13-36]	3460.0	KRRKEVPGPNGATEEDGIPSKVQR
AA-BBP-M[282-304]-AA	3459.4	SREFIGSFLEMFGPEGALKCMLK
AA-BBP-N[17-42]	3475.0	EVPGPNGATEEDGIPSKVQRSAVGLR
BBP-M[260-281]	3475.1	SKEFVQKVEEKSIDLIQKWEEK
BBP-N[9-33]	3476.0	VNSRKRRKEVPGPNGATEEDGIPSK
AA-BBP-N[His-tag-8]	3489.0	MAKHHHHHHIEGRSAMDAQSSAK
AA-BBP-M[284-307]	3489.2	EFIGSFLEMFGPEGALKCMLKEGK
BBP-P[336-355]	3488.2	WPFSGKTSPSSSPASLSRSK
M[301-304]-BBP-C[101-122]	3755.7	NLFPNTYLIVGVSSDELTHNFK
M[301-304]-BBP-M[219-239]	3755.8	RNLQRGYTAKELNVSFINEKK
M[301-304]-BBP-N[His-tag-8]	3924.8	MAKHHHHHHIEGRSAMDAQSSAK
M[301-304]-BBP-M[284-307]	3925.1	EFIGSFLEMFGPEGALKCMLKEGK
AA-BBP-M[271-300]	4297.1	SIDLIQKWEEKSREFIGSFLEMFGPEGALK
BBP-P[310-341]	4297.0	MLQAISPKQSPSSSPTHERSPSPSFRWPFSGK
BBP-N[His-tag-15]	4297.0	MAKHHHHHHIEGRSAMDAQSSAKVNSRKRR
AA-BBP-N[His-tag-15]	4312.0	MAKHHHHHHIEGRSAMDAQSSAKVNSRKRR
BBP-P[310-341]	4313.0	MLQAISPKQSPSSSPTHERSPSPSFRWPFSGK
BBP-N[His-tag-15]	4313.0	MAKHHHHHHIEGRSAMDAQSSAKVNSRKRR
M[271-304]-BBP	4745.7	SIDLIQKWEEKSREFIGSFLEMFGPEGALKCMLK
M[301-304]-BBP-P[342-367]	4745.5	TSPSSSPASLSRSKAVTSDISEDEED
AA-BBP-M[251-283]	4859.7	KKVKDVEEKSKEFVQKVEEKSIDLIQKWEEKSR
BBP-P[336-367]	4859.4	WPFSGKTSPSSSPASLSRSKAVTSDISEDEED
M[301-304]-BBP-P[310-341]	4949.9	MLQAISPKQSPSSSPTHERSPSPSFRWPFSGK

M[301-309]-BBP-P[354-367]	4948.8	SKAVTSDISEDEED
M[301-304]-BBP-C[101-132]	4950.0	NLFPNTYLIVGVSSDELTHNFKGFTVMNENER
AA-BBP-M[271-304]	5119.1	SIDLIQKWEEKSREFIGSFLEMFGPEGALKCMLKEGK
BBP-C[62-100]	5117.9	VTMEEASRGTPSERPVRVYADGIFDLFHSGHARALMQAK

CT-H301C + lipid		
Identity	Average Observed Mass (Da)	Sequence of peptide from contact
M[301-304]-BBP-P[305-309]	1793.1	EGKGR
M[301-304]-BBP-M[220-223]	1793.1	NLQR
M[301-304]-BBP-M[282-300]	3378.1	SREFIGSFLEMFGPEGALK
M[301-304]-BBP-P[318-335]	3378.6	QSPSSSPTHERSPSPSFR
M[301-304]-BBP-M[282-300]	3394.1	SREFIGSFLEMFGPEGALK
M[301-304]-BBP-P[318-335]	3394.6	QSPSSSPTHERSPSPSFR
M[301-304]-BBP-M[229-245]	3394.9	ELNVSFINEKKYHLQER
M[301-304]-BBP-M[282-300]	3410.1	SREFIGSFLEMFGPEGALK
M[301-304]-BBP-M[229-245]	3410.9	ELNVSFINEKKYHLQER
AA-BBP-N[17-42]	3475.0	EVPGPNGATEEDGIPSKVQRSAVGLR
BBP-M[260-281]	3475.1	SKEFVQKVEEKSIDLIQKWEEK
BBP-N[9-33]	3476.0	VNSRKRRKEVPGPNGATEEDGIPSK
AA-BBP-M[271-300]	4297.1	SIDLIQKWEEKSREFIGSFLEMFGPEGALK
BBP-P[336-367]	4299.5	WPFSGKTSPSSSPASLSRSKAVTSDISEDEED
AA-BBP-M[271-300]	4314.8	SIDLIQKWEEKSREFIGSFLEMFGPEGALK
BBP-P[342-367]	4316.5	TSPSSSPASLSRSKAVTSDISEDEED
M[301-304]-BBP-N[His-tag-15]	4780.5	MAKHHHHHHIEGRSAMDAQSSAKVNSRKRR
M[301-304]-BBP-P[342-367]	4778.2	TSPSSSPASLSRSKAVTSDISEDEED
AA-BBP-M[271-304]-AA	4846.7	VKKKVKDVEEKSKEFVQKVEEKSIDLIQKWEEK
AA-BBP-M[282-317]	4846.9	SREFIGSFLEMFGPEGALKCMLKEGKGRMLQAISPK
M[271-304]-BBP	4778.6	SIDLIQKWEEKSREFIGSFLEMFGPEGALKCMLK

Appendix III: MALDI-MS spectra for photo-cross-linked samples prepared in the presence or absence of 0.25 mM TX-100 remained similar

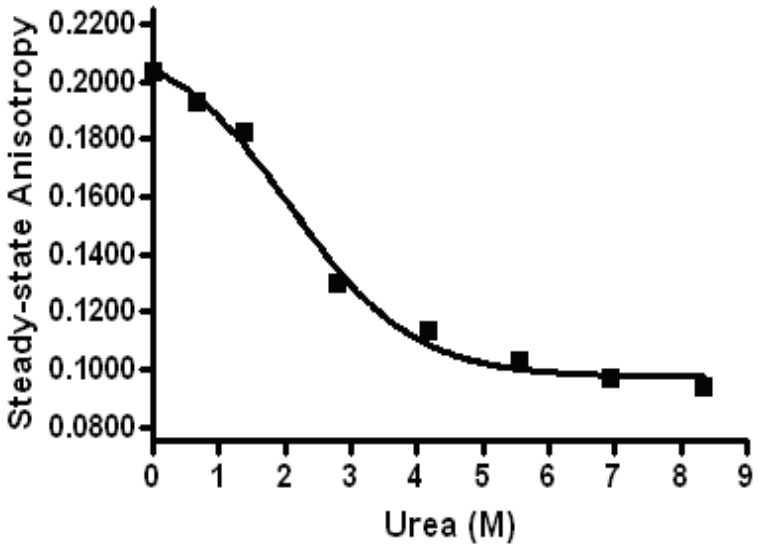
CT-S288C was BBP-labeled, photo-cross-linked, tryptic digested, avidin purified, and analyzed on MALDI-MS as described in Methods in Chapter 2. A focused view of the spectra (scanned from 700 to 10000 Da), with a few identified peaks, is shown below. Other parts of spectra not shown here were also similar to each other. For samples without TX-100, elution buffer used to recover BBP-labeled samples from Ni beads did not include TX-100. To compensate for significant loss of sample eluted in the absence of TX-100 (~50%), input for these trials was doubled. DTT-IAA = samples that have been reduced and alkylated with IAA.



Appendix IV: Urea denaturation analysis of Helix E, and the effect of CTP and lipids on anisotropy of Helix E

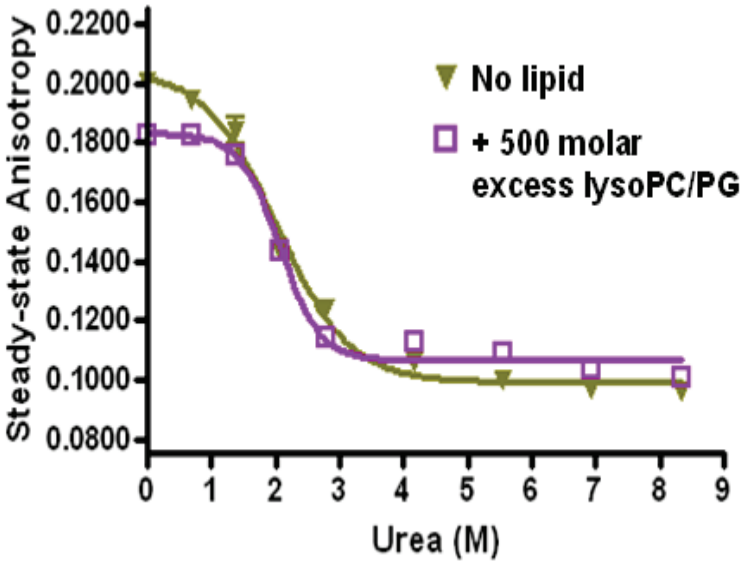
Steady-state anisotropy values for CT-T207C (near N-terminus of helix E) with and without CTP or lipids were determined. In all samples, total OG labelling efficiency was 17% (1137 nM CT; 189 nM OG). For urea denaturation analysis shown below the table, OG-labeled CT-T207C was treated with the indicated concentrations of urea for 5 min and its steady-state anisotropies were analyzed at 20°C, as described in Chapter 3 Methods. Two different % OG labeled samples were analyzed with similar result. EC_{50} is 2.1 $\pm$ 0.1 M urea; transition of unfolding is - 0.5 + 0.1.

Lipid:CT (M/M)	CTP:CT (M/M)	Steady-state anisotropy
0	0	0.180 + 0.001
0	7300	0.181 + 0.001
530	0	0.181 + 0.002



Appendix V: Urea denaturation analysis of lipid-bound CT-I272C

Urea denaturation analysis was performed on OG-labeled CT-I272C with and without 500 molar excess lysoPC/PG lipid mixture. Samples were treated with the indicated concentrations of urea for 5 min and its steady-state anisotropies were analyzed at 20°C, as described in Chapter 3 Methods.



Reference List

1. Pufall, M.A. & Graves, B.J. Autoinhibitory domains: modular effectors of cellular regulation. *Annu Rev Cell Dev Biol* **18**, 421-62 (2002).
2. Wang, D., Zheng, F., Holmberg, S. & Kohlhaw, G.B. Yeast transcriptional regulator Leu3p. Self-masking, specificity of masking, and evidence for regulation by the intracellular level of Leu3p. *J Biol Chem* **274**, 19017-24 (1999).
3. Ziegler, W.H., Liddington, R.C. & Critchley, D.R. The structure and regulation of vinculin. *Trends Cell Biol* **16**, 453-60 (2006).
4. Dombroski, A.J., Walter, W.A. & Gross, C.A. Amino-terminal amino acids modulate sigma-factor DNA-binding activity. *Genes Dev* **7**, 2446-55 (1993).
5. Prehoda, K.E., Scott, J.A., Mullins, R.D. & Lim, W.A. Integration of multiple signals through cooperative regulation of the N-WASP-Arp2/3 complex. *Science* **290**, 801-6 (2000).
6. Aghazadeh, B., Lowry, W.E., Huang, X.Y. & Rosen, M.K. Structural basis for relief of autoinhibition of the Dbl homology domain of proto-oncogene Vav by tyrosine phosphorylation. *Cell* **102**, 625-33 (2000).
7. Jackson-Fisher, A.J., Chitikila, C., Mitra, M. & Pugh, B.F. A role for TBP dimerization in preventing unregulated gene expression. *Mol Cell* **3**, 717-27 (1999).
8. Calleja, V., Laguerre, M. & Larijani, B. 3-D structure and dynamics of protein kinase B-new mechanism for the allosteric regulation of an AGC kinase. *J Chem Biol* **2**, 11-25 (2009).
9. Xu, W., Doshi, A., Lei, M., Eck, M.J. & Harrison, S.C. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol Cell* **3**, 629-38 (1999).
10. Munson, M., Chen, X., Cocina, A.E., Schultz, S.M. & Hughson, F.M. Interactions within the yeast t-SNARE Sso1p that control SNARE complex assembly. *Nat Struct Biol* **7**, 894-902 (2000).
11. Hupp, T.R., Sparks, A. & Lane, D.P. Small peptides activate the latent sequence-specific DNA binding function of p53. *Cell* **83**, 237-45 (1995).
12. Cowan-Jacob, S.W. et al. The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* **13**, 861-71 (2005).
13. Goetz, T.L., Gu, T.L., Speck, N.A. & Graves, B.J. Auto-inhibition of Ets-1 is counteracted by DNA binding cooperativity with core-binding factor alpha2. *Mol Cell Biol* **20**, 81-90 (2000).

14. Buck, M., Xu, W. & Rosen, M.K. Global disruption of the WASP autoinhibited structure on Cdc42 binding. Ligand displacement as a novel method for monitoring amide hydrogen exchange. *Biochemistry* **40**, 14115-22 (2001).

15. Gresset, A., Hicks, S.N., Harden, T.K. & Sondek, J. Mechanism of phosphorylation-induced activation of phospholipase C-gamma isozymes. *J Biol Chem* **285**, 35836-47 (2010).

16. Tiyanont, K. et al. Evidence for Increased Exposure of the Notch1 Metalloprotease Cleavage Site upon Conversion to an Activated Conformation. *Structure* **19**, 546-54 (2011).

17. Petersen, J.M. et al. Modulation of transcription factor Ets-1 DNA binding: DNA-induced unfolding of an alpha helix. *Science* **269**, 1866-9 (1995).

18. Lee, G.M. et al. The affinity of Ets-1 for DNA is modulated by phosphorylation through transient interactions of an unstructured region. *J Mol Biol* **382**, 1014-30 (2008).

19. Garvie, C.W., Hagman, J. & Wolberger, C. Structural studies of Ets-1/Pax5 complex formation on DNA. *Mol Cell* **8**, 1267-76 (2001).

20. Lee, G.M. et al. The structural and dynamic basis of Ets-1 DNA binding autoinhibition. *J Biol Chem* **280**, 7088-99 (2005).

21. Lin, R., Mamane, Y. & Hiscott, J. Structural and functional analysis of interferon regulatory factor 3: localization of the transactivation and autoinhibitory domains. *Mol Cell Biol* **19**, 2465-74 (1999).

22. Rohatgi, R., Ho, H.Y. & Kirschner, M.W. Mechanism of N-WASP activation by CDC42 and phosphatidylinositol 4, 5-bisphosphate. *J Cell Biol* **150**, 1299-310 (2000).

23. Dunker, A.K., Silman, I., Uversky, V.N. & Sussman, J.L. Function and structure of inherently disordered proteins. *Curr Opin Struct Biol* **18**, 756-64 (2008).

24. Smock, R.G. & Gierasch, L.M. Sending signals dynamically. *Science* **324**, 198-203 (2009).

25. Mittag, T. et al. Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proc Natl Acad Sci U S A* **105**, 17772-7 (2008).

26. Uversky, V.N. & Dunker, A.K. Understanding Protein Unfolding. *Biochim. Biophys. Acta* **1804**, 1231-1264 (2010).

27. Baker, J.M. et al. CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. *Nat Struct Mol Biol* **14**, 738-45 (2007).

28. Cui, Z. & Houweling, M. Phosphatidylcholine and cell death. *Biochim Biophys Acta* **1585**, 87-96 (2002).

29. Clement, J.M. & Kent, C. CTP:phosphocholine cytidylyltransferase: insights into regulatory mechanisms and novel functions. *Biochem Biophys Res Commun* **257**, 643-50 (1999).

30. Kent, C. Regulatory enzymes of phosphatidylcholine biosynthesis: a personal perspective. *Biochim Biophys Acta* **1733**, 53-66 (2005).

31. Vance, D.E. Boehringer Mannheim Award lecture. Phosphatidylcholine metabolism: masochistic enzymology, metabolic regulation, and lipoprotein assembly. *Biochem Cell Biol* **68**, 1151-65 (1990).

32. Li, Z. & Vance, D.E. Phosphatidylcholine and choline homeostasis. *J Lipid Res* **49**, 1187-94 (2008).

33. Exton, J.H. Signaling through phosphatidylcholine breakdown. *J Biol Chem* **265**, 1-4 (1990).

34. Exton, J.H. Phosphatidylcholine breakdown and signal transduction. *Biochim Biophys Acta* **1212**, 26-42 (1994).

35. Billah, M.M. & Anthes, J.C. The regulation and cellular functions of phosphatidylcholine hydrolysis. *Biochem J* **269**, 281-91 (1990).

36. Choy, P.C., Paddon, H.B. & Vance, D.E. An increase in cytoplasmic CTP accelerates the reaction catalyzed by CTP:phosphocholine cytidylyltransferase in poliovirus-infected HeLa cells. *J Biol Chem* **255**, 1070-3 (1980).

37. Cornell, R.B. & Goldfine, H. The coordination of sterol and phospholipid synthesis in cultured myogenic cells. Effect of cholesterol synthesis inhibition on the synthesis of phosphatidylcholine. *Biochim Biophys Acta* **750**, 504-20 (1983).

38. Sleight, R. & Kent, C. Regulation of phosphatidylcholine biosynthesis in cultured chick embryonic muscle treated with phospholipase C. *J Biol Chem* **255**, 10644-50 (1980).

39. Wright, P.S., Morand, J.N. & Kent, C. Regulation of phosphatidylcholine biosynthesis in Chinese hamster ovary cells by reversible membrane association of CTP: phosphocholine cytidylyltransferase. *J Biol Chem* **260**, 7919-26 (1985).

40. Cornell, R.B. Regulation of mammalian CTP:phosphocholine cytidylyltransferase *in Gross, RW., ed.. Advances in Lipobiology: Volume 1*, 1-38 (1996).

41. Jackowski, S. & Fagone, P. CTP: Phosphocholine cytidylyltransferase: paving the way from gene to membrane. *J Biol Chem* **280**, 853-6 (2005).

42. Wang, L., Magdaleno, S., Tabas, I. & Jackowski, S. Early Embryonic Lethality in Mice with Targeted Deletion of the CTP:Phosphocholine Cytidylyltransferase α Gene (Pcyt1a). *Mol Cell Biol* **25**, 3357-63 (2005).

43. Jackowski, S. et al. Disruption of CCTβ2 Expression Leads to Gonadal Dysfunction. *Mol Cell Biol* **24**, 4720-33 (2004).

44. Lykidis, A., Murti, K.G. & Jackowski, S. Cloning and characterization of a second human CTP:phosphocholine cytidylyltransferase. *J Biol Chem* **273**, 14022-9 (1998).

45. Morand, J.N. & Kent, C. Localization of the membrane-associated CTP:phosphocholine cytidylyltransferase in Chinese hamster ovary cells with an altered membrane composition. *J Biol Chem* **264**, 13785-92 (1989).

46. Watkins, J.D. & Kent, C. Immunolocalization of membrane-associated CTP:phosphocholine cytidylyltransferase in phosphatidylcholine-deficient Chinese hamster ovary cells. *J Biol Chem* **267**, 5686-92 (1992).

47. Wang, Y., Sweitzer, T.D., Weinhold, P.A. & Kent, C. Nuclear localization of soluble CTP:phosphocholine cytidylyltransferase. *J Biol Chem* **268**, 5899-904 (1993).

48. Wang, Y., MacDonald, J.I. & Kent, C. Regulation of CTP:phosphocholine cytidylyltransferase in HeLa cells. Effect of oleate on phosphorylation and intracellular localization. *J Biol Chem* **268**, 5512-8 (1993).

49. Wang, Y., MacDonald, J.I. & Kent, C. Identification of the nuclear localization signal of rat liver CTP:phosphocholine cytidylyltransferase. *J Biol Chem* **270**, 354-60 (1995).

50. Kennedy, E.P. & Weiss, S.B. The function of cytidine coenzymes in the biosynthesis of phospholipides. *J Biol Chem* **222**, 193-214 (1956).

51. Weinhold, P.A., Rounsifer, M.E. & Feldman, D.A. The purification and characterization of CTP:phosphorylcholine cytidylyltransferase from rat liver. *J Biol Chem* **261**, 5104-10 (1986).

52. Weinhold, P.A., Rounsifer, M.E., Charles, L. & Feldman, D.A. Characterization of cytosolic forms of CTP: choline-phosphate cytidylyltransferase in lung, isolated alveolar type II cells, A549 cell and Hep G2 cells. *Biochim Biophys Acta* **1006**, 299-310 (1989).

53. Cornell, R. Chemical cross-linking reveals a dimeric structure for CTP:phosphocholine cytidylyltransferase. *J Biol Chem* **264**, 9077-82 (1989).

54. Tsukagoshi, Y., Nikawa, J. & Yamashita, S. Molecular cloning and characterization of the gene encoding cholinephosphate cytidylyltransferase in Saccharomyces cerevisiae. *Eur J Biochem* **169**, 477-86 (1987).

55. Kalmar, G.B., Kay, R.J., Lachance, A., Aebersold, R. & Cornell, R.B. Cloning and expression of rat liver CTP: phosphocholine cytidylyltransferase: an amphipathic protein that controls phosphatidylcholine synthesis. *Proc Natl Acad Sci U S A* **87**, 6029-33 (1990).

56. Kalmar, G.B., Kay, R.J., LaChance, A.C. & Cornell, R.B. Primary structure and expression of a human CTP:phosphocholine cytidylyltransferase. *Biochim Biophys Acta* **1219**, 328-34 (1994).

57. Rutherford, M.S. et al. The gene for murine CTP:phosphocholine cytidylyltransferase (Ctpct) is located on mouse chromosome 16. *Genomics* **18**, 698-701 (1993).

58. Friesen, J.A., Liu, M.F. & Kent, C. Cloning and characterization of a lipid-activated CTP:phosphocholine cytidylyltransferase from Caenorhabditis elegans: identification of a 21-residue segment critical for lipid activation. *Biochim Biophys Acta* **1533**, 86-98 (2001).

59. Helmink, B.A. & Friesen, J.A. Characterization of a lipid activated CTP:phosphocholine cytidylyltransferase from Drosophila melanogaster. *Biochim Biophys Acta* **1683**, 78-88 (2004).

60. Nishida, I., Swinhoe, R., Slabas, A.R. & Murata, N. Cloning of Brassica napus CTP: phosphocholine cytidylyltransferase cDNAs by complementation in a yeast cct mutant. *Plant Mol Biol* **31**, 205-11 (1996).

61. Jones, P.L., Willey, D.L., Gacesa, P. & Harwood, J.L. Isolation, characterisation and expression of a cDNA for pea cholinephosphate cytidylyltransferase. *Plant Mol Biol* **37**, 179-85 (1998).

62. Yeo, H.J., Larvor, M.P., Ancelin, M.L. & Vial, H.J. Plasmodium falciparum CTP:phosphocholine cytidylyltransferase expressed in Escherichia coli: purification, characterization and lipid regulation. *Biochem J* **324 (Pt 3)**, 903-10 (1997).

63. Mauel, C., Young, M. & Karamata, D. Genes concerned with synthesis of poly(glycerol phosphate), the essential teichoic acid in Bacillus subtilis strain 168, are organized in two divergent transcription units. *J Gen Microbiol* **137**, 929-41 (1991).

64. Sanker, S., Campbell, H.A. & Kent, C. Negative cooperativity of substrate binding but not enzyme activity in wild-type and mutant forms of CTP:glycerol-3-phosphate cytidylyltransferase. *J Biol Chem* **276**, 37922-8 (2001).

65. Veitch, D.P., Gilham, D. & Cornell, R.B. The role of histidine residues in the HXGH site of CTP:phosphocholine cytidylyltransferase in CTP binding and catalysis. *Eur J Biochem* **255**, 227-34 (1998).

66. Jones, M.D., Lowe, D.M., Borgford, T. & Fersht, A.R. Natural variation of tyrosyl-tRNA synthetase and comparison with engineered mutants. *Biochemistry* **25**, 1887-91 (1986).

67. Leatherbarrow, R.J., Fersht, A.R. & Winter, G. Transition-state stabilization in the mechanism of tyrosyl-tRNA synthetase revealed by protein engineering. *Proc Natl Acad Sci U S A* **82**, 7840-4 (1985).

68. Perona, J.J., Rould, M.A. & Steitz, T.A. Structural basis for transfer RNA aminoacylation by Escherichia coli glutaminyl-tRNA synthetase. *Biochemistry* **32**, 8758-71 (1993).

69. Park, Y.S. et al. Identification of functional conserved residues of CTP:glycerol-3-phosphate cytidylyltransferase. Role of histidines in the conserved HXGH in catalysis. *J Biol Chem* **272**, 15161-6 (1997).

70. Veitch, D.P. & Cornell, R.B. Substitution of serine for glycine-91 in the HXGH motif of CTP:phosphocholine cytidylyltransferase implicates this motif in CTP binding. *Biochemistry* **35**, 10743-50 (1996).

71. Weber, C.H., Park, Y.S., Sanker, S., Kent, C. & Ludwig, M.L. A prototypical cytidylyltransferase: CTP:glycerol-3-phosphate cytidylyltransferase from bacillus subtilis. *Structure* **7**, 1113-24 (1999).

72. Pattridge, K.A. et al. Glycerol-3-phosphate cytidylyltransferase. Structural changes induced by binding of CDP-glycerol and the role of lysine residues in catalysis. *J Biol Chem* **278**, 51863-71 (2003).

73. Lee, J., Johnson, J., Ding, Z., Paetzel, M. & Cornell, R.B. Crystal structure of a mammalian CTP: phosphocholine cytidylyltransferase catalytic domain reveals novel active site residues within a highly conserved nucleotidyltransferase fold. *J Biol Chem* **284**, 33535-48 (2009).

74. Helmink, B.A., Braker, J.D., Kent, C. & Friesen, J.A. Identification of lysine 122 and arginine 196 as important functional residues of rat CTP:phosphocholine cytidylyltransferase alpha. *Biochemistry* **42**, 5043-51 (2003).

75. Fong, D.H., Yim, V.C., D'Elia, M.A., Brown, E.D. & Berghuis, A.M. Crystal structure of CTP:glycerol-3-phosphate cytidylyltransferase from Staphylococcus aureus: examination of structural basis for kinetic mechanism. *Biochim Biophys Acta* **1764**, 63-9 (2006).

76. Tronchere, H., Record, M., Terce, F. & Chap, H. Phosphatidylcholine cycle and regulation of phosphatidylcholine biosynthesis by enzyme translocation. *Biochim Biophys Acta* **1212**, 137-51 (1994).

77. Cornell, R.B. & Northwood, I.C. Regulation of CTP:phosphocholine cytidylyltransferase by amphitropism and relocalization. *Trends Biochem Sci* **25**, 441-7 (2000).

78. Sleight, R. & Kent, C. Regulation of phosphatidylcholine biosynthesis in mammalian cells. III. Effects of alterations in the phospholipid compositions of Chinese hamster ovary and LM cells on the activity and distribution of CTP:phosphocholine cytidylyltransferase. *J Biol Chem* **258**, 836-9 (1983).

79. Jamil, H., Hatch, G.M. & Vance, D.E. Evidence that binding of CTP:phosphocholine cytidylyltransferase to membranes in rat hepatocytes is modulated by the ratio of bilayer- to non-bilayer-forming lipids. *Biochem J* **291**, 419-27 (1993).

80. Pelech, S.L. & Vance, D.E. Regulation of rat liver cytosolic CTP: phosphocholine cytidylyltransferase by phosphorylation and dephosphorylation. *J. Biol. Chem.* **257**, 14198-14202 (1982).

81. Kolesnick, R.N. & Hemer, M.R. Physiologic 1,2-diacylglycerol levels induce protein kinase C-independent translocation of a regulatory enzyme. *J Biol Chem* **265**, 10900-4 (1990).

82. Watkins, J.D. & Kent, C. Regulation of CTP:phosphocholine cytidylyltransferase activity and subcellular location by phosphorylation in Chinese hamster ovary cells. The effect of phospholipase C treatment. *J Biol Chem* **266**, 21113-7 (1991).

83. Arnold, R.S. & Cornell, R.B. Lipid regulation of CTP: phosphocholine cytidylyltransferase: electrostatic, hydrophobic, and synergistic interactions of anionic phospholipids and diacylglycerol. *Biochemistry* **35**, 9917-24 (1996).

84. Cornell, R.B. Regulation of CTP:phosphocholine cytidylyltransferase by lipids. 2. Surface curvature, acyl chain length, and lipid-phase dependence for activation. *Biochemistry* **30**, 5881-8 (1991).

85. Johnson, J.E. et al. Comparison of the lipid regulation of yeast and rat CTP: phosphocholine cytidylyltransferase expressed in COS cells. *Biochem J* **285 (Pt 3)**, 815-20 (1992).

86. Davies, S.M., Epand, R.M., Kraayenhof, R. & Cornell, R.B. Regulation of CTP: phosphocholine cytidylyltransferase activity by the physical properties of lipid membranes: an important role for stored curvature strain energy. *Biochemistry* **40**, 10522-31 (2001).

87. Drobnies, A.E., van Der Ende, B., Thewalt, J.L. & Cornell, R.B. CTP:phosphocholine cytidylyltransferase activation by oxidized phosphatidylcholines correlates with a decrease in lipid order: a 2H NMR analysis. *Biochemistry* **38**, 15606-14 (1999).

88. Pelech, S.L. & Vance, D.E. Signal Transduction Via Phosphatidylcholine Cycles. *Trends in Biochemical Sciences* **14**, 28-30 (1989).

89. MacDonald, J.I. & Kent, C. Identification of phosphorylation sites in rat liver CTP: phosphocholine cytidylyltransferase. *J Biol Chem* **269**, 10529-37 (1994).

90. Wieprecht, M., Wieder, T., Paul, C., Geilen, C.C. & Orfanos, C.E. Evidence for phosphorylation of CTP:phosphocholine cytidylyltransferase by multiple proline-directed protein kinases. *J Biol Chem* **271**, 9955-61 (1996).

91. Weinhold, P.A., Charles, L. & Feldman, D.A. Regulation of CTP: phosphocholine cytidylyltransferase in HepG2 cells: effect of choline depletion on phosphorylation, translocation and phosphatidylcholine levels. *Biochim Biophys Acta* **1210**, 335-47 (1994).

92. Arnold, R.S., DePaoli-Roach, A.A. & Cornell, R.B. Binding of CTP:phosphocholine cytidylyltransferase to lipid vesicles: diacylglycerol and enzyme dephosphorylation increase the affinity for negatively charged membranes. *Biochemistry* **36**, 6149-56 (1997).

93. Dennis, M.K., Taneva, S.G. & Cornell, R.B. The intrinsically disordered nuclear localization signal and phosphorylation segments distinguish the membrane affinity of two cytidylyltransferase isoforms. *J Biol Chem* **286**, 12349-60 (2011).

94. Taneva, S., Dennis, M.K., Ding, Z., Smith, J.L. & Cornell, R.B. Contribution of each membrane binding domain of the CTP:phosphocholine cytidylyltransferase-alpha dimer to its activation, membrane binding, and membrane cross-bridging. *J Biol Chem* **283**, 28137-48 (2008).

95. Friesen, J.A., Campbell, H.A. & Kent, C. Enzymatic and cellular characterization of a catalytic fragment of CTP:phosphocholine cytidylyltransferase alpha. *J Biol Chem* **274**, 13384-9 (1999).

96. Dunne, S.J., Cornell, R.B., Johnson, J.E., Glover, N.R. & Tracey, A.S. Structure of the membrane binding domain of CTP:phosphocholine cytidylyltransferase. *Biochemistry* **35**, 11975-84 (1996).

97. Georgieva, E.R., Ramlall, T.F., Borbat, P.P., Freed, J.H. & Eliezer, D. Membrane-bound alpha-synuclein forms an extended helix: long-distance pulsed ESR measurements using vesicles, bicelles, and rodlike micelles. *J Am Chem Soc* **130**, 12856-7 (2008).

98. Wang, Y. & Kent, C. Identification of an inhibitory domain of CTP:phosphocholine cytidylyltransferase. *J Biol Chem* **270**, 18948-52 (1995).

99. Craig, L., Johnson, J.E. & Cornell, R.B. Identification of the membrane-binding domain of rat liver CTP:phosphocholine cytidylyltransferase using chymotrypsin proteolysis. *J Biol Chem* **269**, 3311-7 (1994).

100. Bogan, M.J., Agnes, G.R., Pio, F. & Cornell, R.B. Interdomain and membrane interactions of CTP:phosphocholine cytidylyltransferase revealed via limited proteolysis and mass spectrometry. *J Biol Chem* **280**, 19613-24 (2005).

101. Johnson, J.E., Aebersold, R. & Cornell, R.B. An amphipathic alpha-helix is the principle membrane-embedded region of CTP:phosphocholine cytidylyltransferase. Identification of the 3-(trifluoromethyl)-3-(m-[125I]iodophenyl) diazirine photolabeled domain. *Biochim Biophys Acta* **1324**, 273-84 (1997).

102. Johnson, J.E. & Cornell, R.B. Membrane-binding amphipathic alpha-helical peptide derived from CTP:phosphocholine cytidylyltransferase. *Biochemistry* **33**, 4327-35 (1994).

103. Johnson, J.E., Xie, M., Singh, L.M., Edge, R. & Cornell, R.B. Both acidic and basic amino acids in an amphitropic enzyme, CTP:phosphocholine cytidylyltransferase, dictate its selectivity for anionic membranes. *J Biol Chem* **278**, 514-22 (2003).

104. Taneva, S., Johnson, J.E. & Cornell, R.B. Lipid-induced conformational switch in the membrane binding domain of CTP:phosphocholine cytidylyltransferase: a circular dichroism study. *Biochemistry* **42**, 11768-76 (2003).

105. Braker, J.D., Hodel, K.J., Mullins, D.R. & Friesen, J.A. Identification of hydrophobic amino acids required for lipid activation of C. elegans CTP:phosphocholine cytidylyltransferase. *Arch Biochem Biophys* **492**, 10-6 (2009).

106. Larvor, M.P. et al. Characterization of the lipid-binding domain of the Plasmodium falciparum CTP:phosphocholine cytidylyltransferase through synthetic-peptide studies. *Biochem J* **375**, 653-61 (2003).

107. Grant, J.E. et al. The N terminus of GTP gamma S-activated transducin alpha-subunit interacts with the C terminus of the cGMP phosphodiesterase gamma-subunit. *J Biol Chem* **281**, 6194-202 (2006).

108. Luo, Y. et al. Troponin-I interacts with the Met47 region of skeletal muscle actin. Implications for the mechanism of thin filament regulation by calcium. *J Mol Biol* **316**, 429-34 (2002).

109. Cai, K., Itoh, Y. & Khorana, H.G. Mapping of contact sites in complex formation between transducin and light-activated rhodopsin by covalent crosslinking: use of a photoactivatable reagent. *Proc Natl Acad Sci U S A* **98**, 4877-82 (2001).

110. Jahn, O., Eckart, K., Tezval, H. & Spiess, J. Characterization of peptide-protein interactions using photoaffinity labeling and LC/MS. *Anal Bioanal Chem* **378**, 1031-6 (2004).

111. Vodovozova, E.L. Photoaffinity labeling and its application in structural biology. *Biochemistry (Mosc)* **72**, 1-20 (2007).

112. Borden, W.T. et al. The interplay of theory and experiment in the study of phenylnitrene. *Acc Chem Res* **33**, 765-71 (2000).

113. Brunner, J., Senn, H. & Richards, F.M. 3-Trifluoromethyl-3-phenyldiazirine. A new carbene generating group for photolabeling reagents. *J Biol Chem* **255**, 3313-8 (1980).

114. Dorman, G. & Prestwich, G.D. Benzophenone photophores in biochemistry. *Biochemistry* **33**, 5661-73 (1994).

115. Ahrends, R. et al. Identifying an interaction site between MutH and the C-terminal domain of MutL by crosslinking, affinity purification, chemical coding and mass spectrometry. *Nucleic Acids Res* **34**, 3169-80 (2006).

116. Guo, L.W. et al. Sulfhydryl-reactive, cleavable, and radioiodinatable benzophenone photoprobes for study of protein-protein interaction. *Bioconjug Chem* **16**, 685-93 (2005).

117. Davidson, W. et al. Characterization of the binding site for inhibitors of the HPV11 E1-E2 protein interaction on the E2 transactivation domain by photoaffinity labeling and mass spectrometry. *Anal Chem* **76**, 2095-102 (2004).

118. Mann, T.L. & Krull, U.J. Fluorescence polarization spectroscopy in protein analysis. *Analyst* **128**, 313-7 (2003).

119. Lakowicz, J.R. *Principles of fluorescence spectroscopy*, Chapter 10 (pp 355-382). (Springer, New York, 2006).

120. Berberan-Santos, M.N., Nunes Pereira, E.J., Martinho, J.M.G. Stochastic theory of molecular radiative transport. *Journal of Chemical Physics* **103**, 3022-3028 (1995).

121. Kallenberger, B.C., Love, J.D., Chatterjee, V.K. & Schwabe, J.W. A dynamic mechanism of nuclear receptor activation and its perturbation in a human disease. *Nat Struct Biol* **10**, 136-40 (2003).

122. Burkhardt, P., Hattendorf, D.A., Weis, W.I. & Fasshauer, D. Munc18a controls SNARE assembly through its interaction with the syntaxin N-peptide. *EMBO J* **27**, 923-33 (2008).

123. Xie, M., Smith, J.L., Ding, Z., Zhang, D. & Cornell, R.B. Membrane binding modulates the quaternary structure of CTP:phosphocholine cytidylyltransferase. *J Biol Chem* **279**, 28817-25 (2004).

124. Zagotta, W.N., Hoshi, T. & Aldrich, R.W. Restoration of inactivation in mutants of Shaker potassium channels by a peptide derived from ShB. *Science* **250**, 568-71 (1990).

125. Calakos, N., Bennett, M.K., Peterson, K.E. & Scheller, R.H. Protein-protein interactions contributing to the specificity of intracellular vesicular trafficking. *Science* **263**, 1146-9 (1994).

126. Holzinger, A., Phillips, K.S. & Weaver, T.E. Single-step purification/solubilization of recombinant proteins: application to surfactant protein B. *Biotechniques* **20**, 804-6, 808 (1996).

127. Qiagen. *QiaExpressionist. A handbook for high-level expression and purification of 6xHis-tagged proteins.*, (2001).

128. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-54 (1976).

129. Inoue, T. & Forgac, M. Cysteine-mediated cross-linking indicates that subunit C of the V-ATPase is in close proximity to subunits E and G of the V1 domain and subunit a of the V0 domain. *J Biol Chem* **280**, 27896-903 (2005).

130. Ward, D.G., Brewer, S.M., Cornes, M.P. & Trayer, I.P. A cross-linking study of the N-terminal extension of human cardiac troponin I. *Biochemistry* **42**, 10324-32 (2003).

131. Chin, J.W., Martin, A.B., King, D.S., Wang, L. & Schultz, P.G. Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. *Proc Natl Acad Sci U S A* **99**, 11020-4 (2002).

132. http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msdigest.

133. Cavanagh, J., Fairbrother, W.J., Palmer III, A.G., Rance, M., Skelton, N.J. *Protein NMR Spectroscopy: Principles and Practice*, (2007).

134. Bernado, P., Bertoncini, C.W., Griesinger, C., Zweckstetter, M. & Blackledge, M. Defining long-range order and local disorder in native alpha-synuclein using residual dipolar couplings. *J Am Chem Soc* **127**, 17968-9 (2005).

135. Dyson, H.J. & Wright, P.E. Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. *Adv Protein Chem* **62**, 311-40 (2002).

136. Columbus, L. & Hubbell, W.L. A new spin on protein dynamics. *Trends Biochem Sci* **27**, 288-95 (2002).

137. Hubbell, W.L., Cafiso, D.S. & Altenbach, C. Identifying conformational changes with site-directed spin labeling. *Nat Struct Biol* **7**, 735-9 (2000).

138. http://www.invitrogen.com/site/us/en/home/References/Molecular-Probes-The-Handbook.html.

139. Rusinova, E., Tretyachenko-Ladokhina, V., Vele, O.E., Senear, D.F. & Alexander Ross, J.B. Alexa and Oregon Green dyes as fluorescence anisotropy probes for measuring protein-protein and protein-nucleic acid interactions. *Anal Biochem* **308**, 18-25 (2002).

140. Runnels, L.W. & Scarlata, S.F. Theory and application of fluorescence homotransfer to melittin oligomerization. *Biophys J* **69**, 1569-83 (1995).

141. Uversky, V.N. & Ptitsyn, O.B. All-or-none solvent-induced transitions between native, molten globule and unfolded states in globular proteins. *Fold Des* **1**, 117-22 (1996).

142. http://rosettadesigngroup.com/blog/521/17-protein-disorder-prediction-servers/.

143. Friesen, J.A., Park, Y.S. & Kent, C. Purification and kinetic characterization of CTP:phosphocholine cytidylyltransferase from Saccharomyces cerevisiae. *Protein Expr Purif* **21**, 141-8 (2001).

144. Tilley, D.M., Evans, C.R., Larson, T.M., Edwards, K.A. & Friesen, J.A. Identification and characterization of the nuclear isoform of Drosophila melanogaster CTP:phosphocholine cytidylyltransferase. *Biochemistry* **47**, 11838-46 (2008).

145. Wright, P.E. & Dyson, H.J. Linking folding and binding *Current Opinion in Structural Biology* **19**, 31-38 (2009).

146. Shoemaker, B.A., Portman, J.J. & Wolynes , P.G. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc Natl Acad Sci U S A* **97**, 8868-8873 (2000).

147. Morrow, J.A. et al. Apolipoprotein E4 forms a molten globule. A potential basis for its association with disease. *J Biol Chem* **277**, 50380-5 (2002).

148. Ghosh, S., Jana, S., Nath, D. & Guchhait, N. Fluorescent probing of protein bovine serum albumin stability and denaturation using polarity sensitive spectral response of a charge transfer probe. *J Fluoresc* **21**, 365-74 (2011).

149. Forsyth, W.R., Bilsel, O., Gu, Z. & Matthews, C.R. Topology and sequence in the folding of a TIM barrel protein: global analysis highlights partitioning between transient off-pathway and stable on-pathway folding intermediates in the complex folding mechanism of a (betaalpha)8 barrel of unknown function from B. subtilis. *J Mol Biol* **372**, 236-53 (2007).

150. Chakrabarty, A., Mallick, A., Haldar, B., Das, P. & Chattopadhyay, N. Binding interaction of a biological photosensitizer with serum albumins: a biophysical study. *Biomacromolecules* **8**, 920-7 (2007).

151. Sigman, J.A. et al. Flexibility in substrate recognition by thimet oligopeptidase as revealed by denaturation studies. *Biochem J* **388**, 255-61 (2005).

152. Schulman, B.A., Kim, P.S., Dobson, C.M. & Redfield, C. A residue-specific NMR view of the non-cooperative unfolding of a molten globule. *Nat Struct Biol* **4**, 630-4 (1997).

153. Antz, C. et al. NMR structure of inactivation gates from mammalian voltage-dependent potassium channels. *Nature* **385**, 272-5 (1997).

154. Hoshi, T., Zagotta, W.N. & Aldrich, R.W. Biophysical and molecular mechanisms of Shaker potassium channel inactivation. *Science* **250**, 533-8 (1990).

155. Prince-Carter, A. & Pfaffinger, P.J. Multiple intermediate states precede pore block during N-type inactivation of a voltage-gated potassium channel. *J Gen Physiol* **134**, 15-34 (2009).

156. Liebovitch, L.S., Selector, L.Y. & Kline, R.P. Statistical properties predicted by the ball and chain model of channel inactivation. *Biophys J* **63**, 1579-85 (1992).

157. Podlaha, O. & Zhang, J. Positive selection on protein-length in the evolution of a primate sperm ion channel. *Proc Natl Acad Sci U S A* **100**, 12241-6 (2003).

158. MacKinnon, R., Aldrich, R.W. & Lee, A.W. Functional stoichiometry of Shaker potassium channel inactivation. *Science* **262**, 757-9 (1993).

159. Gomez-Lagunas, F. & Armstrong, C.M. Inactivation in ShakerB K+ channels: a test for the number of inactivating particles on each channel. *Biophys J* **68**, 89-95 (1995).