**5.08J Biological Chemistry II (Spring 2016)**
**ANSWERS to Problem Set 7 Module 5**

1. Biocatalysis has emerged as a general approach to asymmetric synthesis in organic chemistry. Recently the unique protonation machinery of a squalene /hopene cyclase (SHC, Fig. 1) was evolved and engineered to function as a Bronsted acid catalyst in water. The success of the approach was highlighted by the stereospecific synthesis of a variety of cyclohexanoids (Fig. 5). One of the active areas of research today is to harness nature's catalytic machineries to generate unique catalysts.

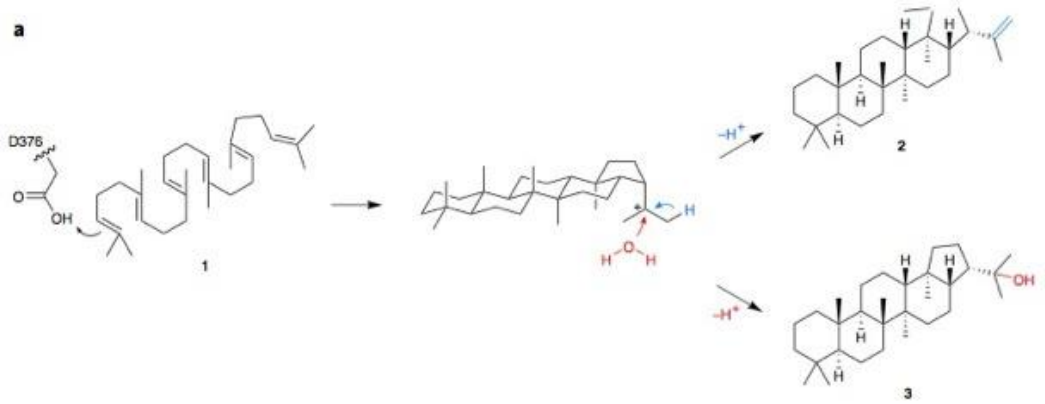
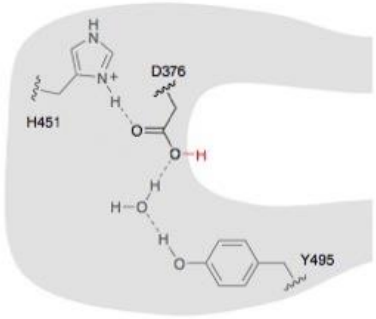



**Fig 1** SHC catalyzes the polyene cyclization of squalene (1) to pentacyclic products hopene (2) and hopanol (3). [Recall from Lecture, the formation of lanosterol from the epoxide of squalene.]

The catalytic machinery of SHC is shown in Fig.2 where D376 plays a key role in the cyclization process. Shown in Fig. 3 is the active site of SHC colored in gray. The Bronsted acid D376 is shown in gray sticks (red oxygens).



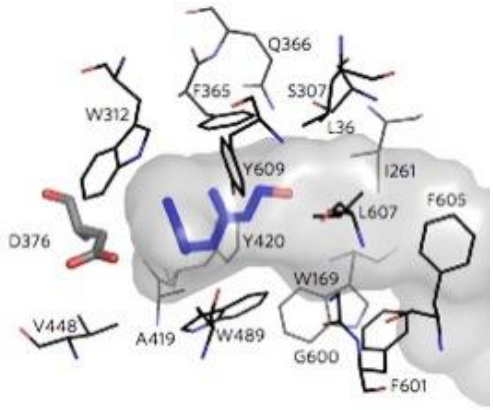




**Fig 2**. (left) Active site machinery of SHC and Bronsted acid D376

**Fig 3**. (right) The active site with its shape highlighted in gray. D376, the putative Bronsted acid, is shown in gray sticks with red oxygens and geraniol shown in blue sticks, is modeled into the active site.



A library of 34 single variants of SHC was generated by site-directed mutagenesis and screened using the monoterpenoid geraniol **4** or the (S)-6, 7-epoxygeraniol **7** revealing mutant proteins that had excellent catalytic activity and high stereoselectivity. In the case of the conversion of **4** to **6**, a single point mutation was sufficient to convert an almost inactive, wt-enzyme into a practical, useful catalyst.

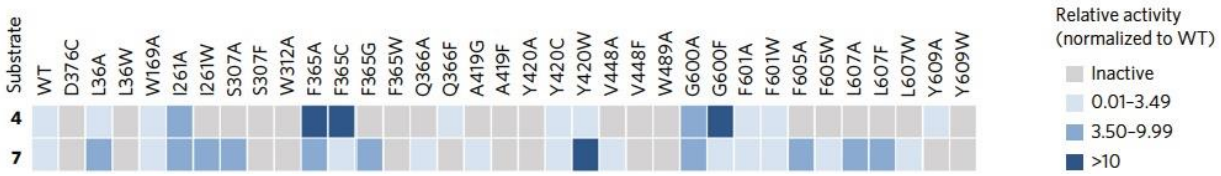



**Fig 4** The generation and screening of a variety of point mutations of SHC for activity with either **4** or **7**.

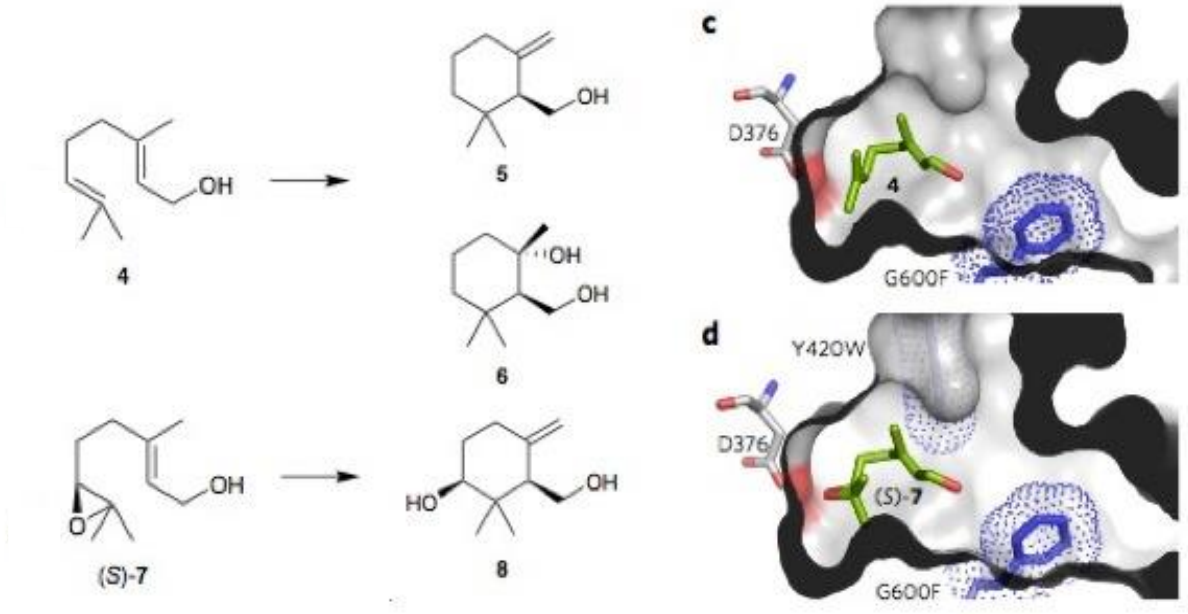



**Fig 5.** Cyclization of geraniol **4** and (S)-6, 7-epoxygeraniol **7**. The transformations of **4** to **5** and **6** and of **7** to **8** are shown above. The putative active conformation of 4 (c above) and 7 (d above) are modeled into the active site of SHC.



**Questions:**

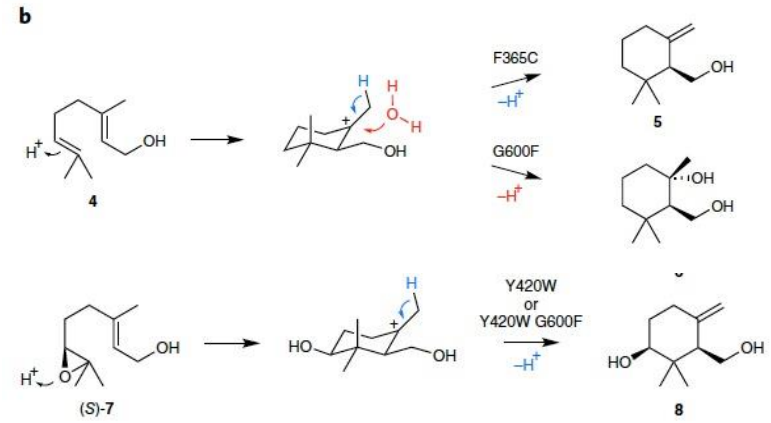
1. Look at the active site machinery in Figs 2 and 3. Think about the basic rules for the chemistry of terpene reactions discussed in Lecture.

a. What is unusual about D376 in Fig 2 and what might it tell you about its function?
In the cartoon in Figure 2, D376 is protonated. Since the pKas of D are around 4.8, this protonation state is unusual. However, a close look at the structure suggests D376 is in a hydrophobic environment (W312, V448, A 419) that could elevate its pKa so that it would remain protonated. In addition, the H bonding network shown in Figure 2, shows that the more acidic, anti-oriented proton of the carboxylate is directed toward the substrate (olefin or epoxide, right). Note D376C mutant is inactive, supporting the importance of this residue.

b. What is unusual about the active site cavity of the enzyme shown in Fig. 3 that will be important for catalysis and specificity? (Think about discussions in Lecture) The authors suggest that the active site is unusual because of the apparent separation of D376, the putative general acid catalyst, from the substrate binding site. The amino acids involved in the protonation domain are mainly part of a rigid, well ordered network, while the squalene binding site (gray shape in Fig. 3) is more malleable, shape selective, and filled with aromatics which turnout to be on flexible loops (not clear from the modd structures you were given). The authors goal is to

convert SHC to new catalysts that can catalyze specific chemistry on small analogs of the larger substrate. These analogs bind because they small, however, the initial binding is likely non-productive. What is important for optimizing SHC to catalyze these reactions (Figure 5) is that the wt enzyme has very low activity with these small analogs. Herschlag made a proposal that has withstood the test of time. He proposed that evolving better protein catalysts requires a low, initial level of activity. Terpene cyclases in general and SHC in this problem, are promiscuous. Thus the authors thought that with appropriate mutations of the wt SHC, better catalysts could results from rational mutations. Their approach was to make a library of mutants and examine their activity (Figure 4). Several mutants (G600F and Y420W, see data in Figure 4) likely work by stabilizing a reactive, folded conformation of the substrate analog. Finally, we discussed in lecture that aromatics are often found in the substrate binding site of terpene cyclases to help fold them and then potentially to stabilize proposed carbocation intermediates by pi-cation interactions. (F, Y, W stacking on the carbocation). Since carbocation intermediates have remained elusive experimentally to date, this hypothesis remains to be established. However in SHC the aromatic residues are apparent.

2. The reactions that the investigators were interested in catalyzing are shown in Fig 5. Propose a mechanism for the role of D376 in each of these reactions. Draw a mechanistic scheme (show the intermediates) for the conversion of **4** to **5** and **6** and for the conversion of **7** to **8**.
This problem was taken from *Nat Chem Biol* 11. 121 (2015). SHCs are protonases for stereoselective Bronsted acid catalysis (Figure 2). The two cases shown below are examples of two general paradigms discussed in class: protonation of an olefin or an epoxide to initiate cyclization. Note that protonation of the epoxide proceeds at the oxygen lone pair, whereas alkene protonation takes place at the pi system. Thus the protonation would depend on slightly different orientations of the active chair conformer and different mutations might effect each reaction in a different fashion.






**Summary**: The distinct properties of enzymes involved in terpene biosynthesis, particularly their promiscuity, offer many opportunities for their use as versatile catalysts taking advantage of our understanding of these types of reactions and our ability to make libraries of mutants that can be screened.

2. "The volumetric heating values of today's biofuels are too low to power energy-intensive aircraft, rockets and missiles." Recently dimers of the natural product pinene (Fig 6) has been

shown to have a volumetric heating value similar to the tactical fuel JP-10. Several labs have thus focused their efforts on engineering *E. coli* to provide a sustainable source of pinene.

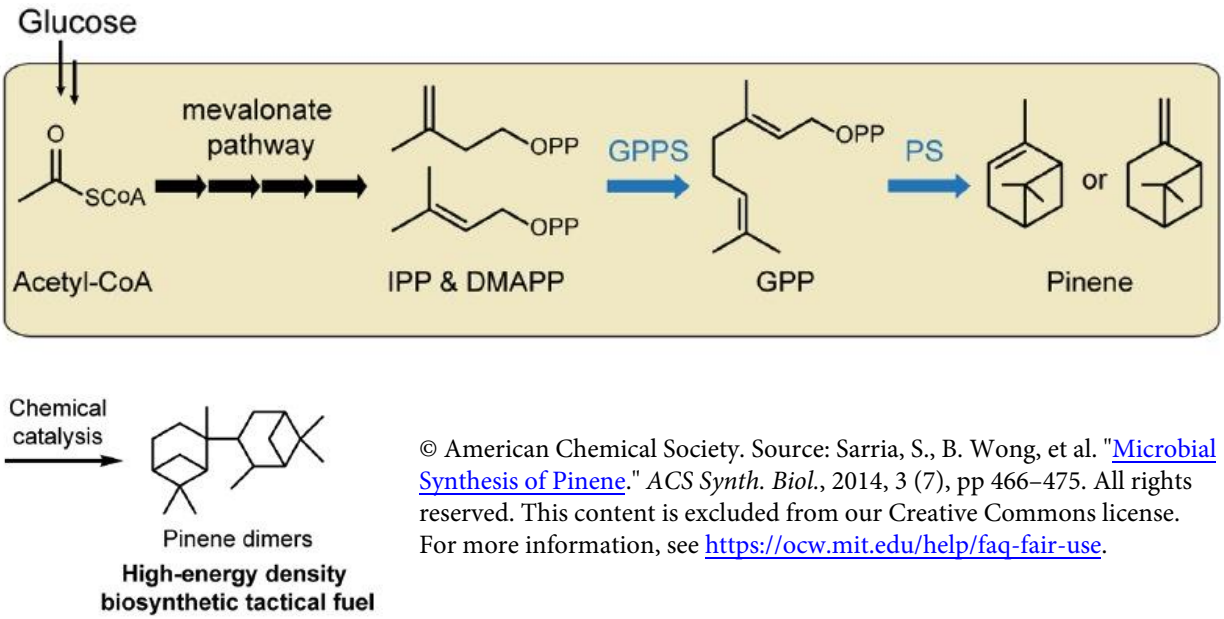


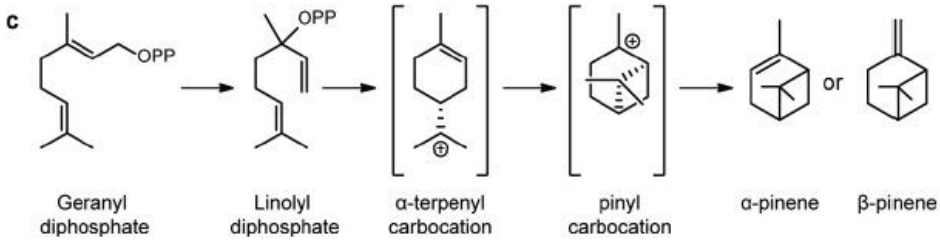



**Fig. 6** Microbial synthesis of pinene using the mevalonate pathway to produce IPP and DMAPP and geranyl diphosphate synthase (GPPS) and pinene synthase (PS) to produce pinene.

Extensive experimentation was required to reduce toxicity associated with the high concentrations of GPP and the pinenes. GPPS and PS from many different organisms were examined and many problems associated with levels of protein expression were encountered. They found that the *A. grandis* PS resulted in the highest production of pinene, even though in *E. coli* it was one of the more poorly expressed PSs. Furthermore, the ratio of the two pinene isomers varied with GPPS in a way that is not understand. Work is progressing to try to optimize production of pinene and understand the basis for and fix the problems encountered.

**Questions**:

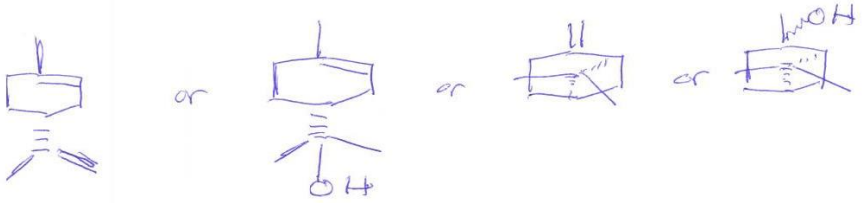
1. Propose a mechanism by which GPP might be converted to the two isomers of pinene. Show the intermediates that might be involved. This problem was taken from ACS *Synthetic Biology* 3, 466 -475 (2014).






2. How might you design a rapid chemical quench experiment to detect your intermediates? Show design and the structures of the expected trapped species. Carbocations in the above mechanism might be trapped by deprotonation to form olefins, addition of $H_2O$ to

form alcohols, and linolyl diphosphate may be isolable if it is stable. An experiment would involve mixing the enzyme in one syringe, with geranylPP in the second syringe under conditions where all the substrate is bound. If there is any equilibration between the intermediates, one might have a chance of trapping them with an acid, $H^+$, quench. While linolyl diphosphate may be stable enough to isolate, all of the proposed cation intermediates are unstable and thus only trapped products would be observed.




3. Brown, Goldstein and coworkers have studied the Scap protein in an effort to understand its function as a cholesterol sensor. The amino acid sequence and topology of the membrane domain of Scap is shown in Fig 7.

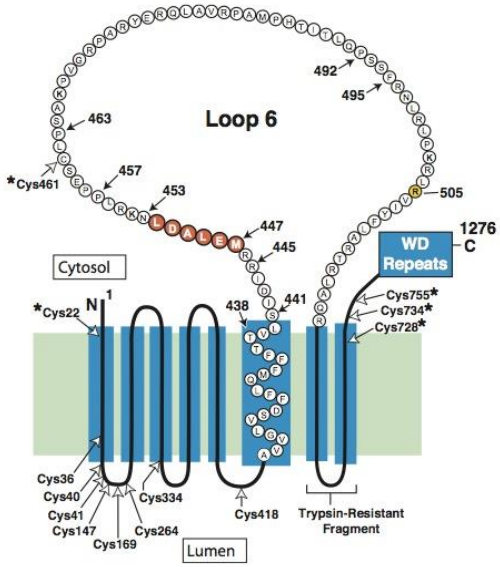



**Fig 7** A cartoon representation of Scap. Loop 1 (between transmembrane helices 1 and 2 has many cysteines on the lumen side).



**Summary of what you have learned in class and from required reading**: Scap is as a cholesterol (Ch) sensor that regulates transport of SREBP from the ER membrane to the Golgi where it gets processed to become a transcription factor (TF). When Ch levels are low in the ER membrane, SREBP binds to the WD domain of Scap and Scap is in a conformation where the MELADL sequence in loop 6 (red, Fig 7) recruits proteins associated with COPII coated vesicle formation. This process facilitates budding of the ER membrane and both Scap and SREBP move to the Golgi. Once in the Golgi, the SREBP is processed by two proteases that generate the soluble transcription factor 75,000 Da from the full length SREBP which is 115,000 kDa. The TF then moves from the cytosol to the nucleus and activates genes involved in fatty acid, cholesterol, and triglyceride production.

When Ch levels are high on the other hand, Scap binds to an ER integral membrane protein Insig, the MELADL sequence becomes sequestered and Scap and SREBP remain in

the ER membrane. **The experiments and data presented below were designed to understand the mechanism behind Scap sensing of Ch, more specifically, the role of Loop1 in this process**.

**Experimental information is given in Figures 8 through 12.** Loop 1 contains 245 amino acids (40-284) and faces the ER lumen. Several constructs of Loop 1 only have been designed and expressed. In addition, [$^3$H] Ch is available which can be analyzed by scintillation counting. In some cases the results with the different Loop 1 constructs are compared with the results from the full length (FL) Scap and FL-Scap-mutants. The experimental design is described in each of the Figure legends.

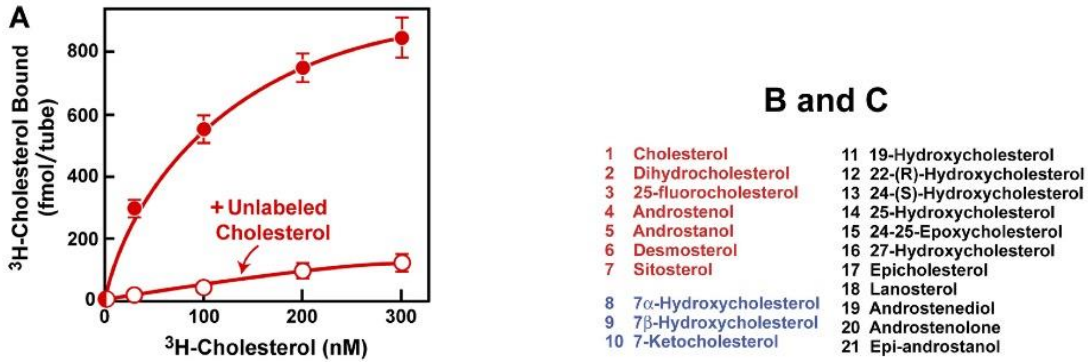



**Fig 8 A.** [$^3$H] Sterol binding to His$_6$-Scap (Loop1) where Loop 1 is residues 46-269. The control is identical to the experiment except that unlabeled cholesterol is added. **Fig 8B and C** are a list of the unlabeled sterols used in Figure **9B** and **9C** below (note there is no 9A). These sterols are used to displace [$^3$H]-cholesterol bound from the different Scap constructs.

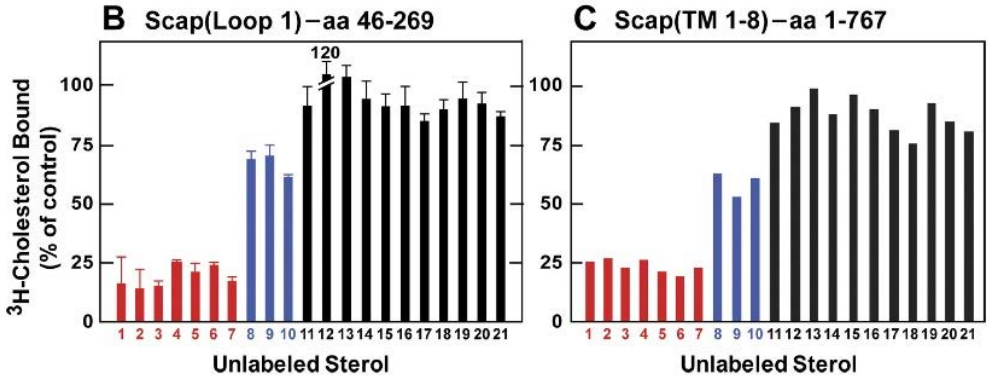



**Fig 9**. Comparison of the sterol specificity of [$^3$H]cholesterol binding to Scap(Loop 1, **B**) and Scap (TM1-8, where TM1-8 indicates full length Scap with 8 transmembrane helices, **C**). The assay uses scintillation counting to determine the amount of [$^3$H]cholesterol that remains bound to the small domain (**left panel, Scap(Loop 1)-aa 46-269**) of Scap and **right panel**, **full length** Scap (**Scap(TM 1-8), aa 1-767**).



Alanine scanning site-directed mutagenesis, SDM, is a method used to replace the natural amino acid with Ala at various positions within the protein of interest, in this case Scap Loop 1 (aa 47-248) and to examine the consequences of this substitution on the processing of SREBP into the TF. These Ala scanning mutants within the Loop 1 (47-248) region of Scap were each transfected into Chinese hamster ovary cells (CHO) cells that lack Scap.

The cells contained a plasmid (pSRE-firefly luciferase) with a steroid responsive element (SRE) placed in front of the firefly luciferase gene. When the gene is transcribed and translated into luciferase, the luciferase activity can be easily measured even at very low amounts of protein by light emission in the presence of ATP. Thus, the researchers could assay in a sensitive fashion, the transcription factor activity of SREBP. The data presented in Fig 10 are a subset of mutants examined for luciferase activity in the presence and absence of cholesterol.

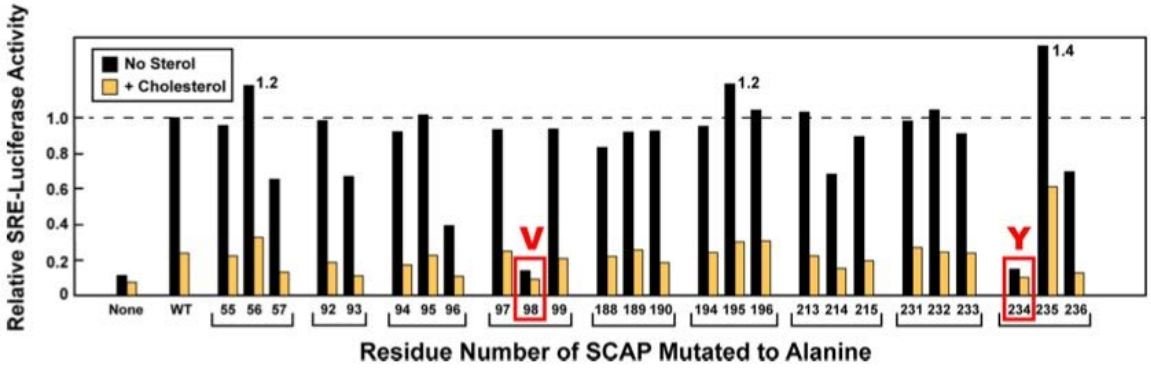



**Fig 10.** Alanine scanning in which site-directed mutagenesis, SDM, was used to replace the natural amino acid with Ala at various positions within Scap loop 1 (aa 47-248) and to examine the consequences of this substitution on the processing of SREBP into the TF. The SRE has been engineered in front of the gene that codes for luciferase. Thus, liberation of the TF activates the transcription of the luciferase gene and ultimately the production of protein that generates light that can be detected in a continuous assay in the presence of ATP.



To determine if the results observed with the Scap Loop1 construct are similar to the full length wt-Scap, the full length Y234A mutant was made and studied as described in Figure legend 11.

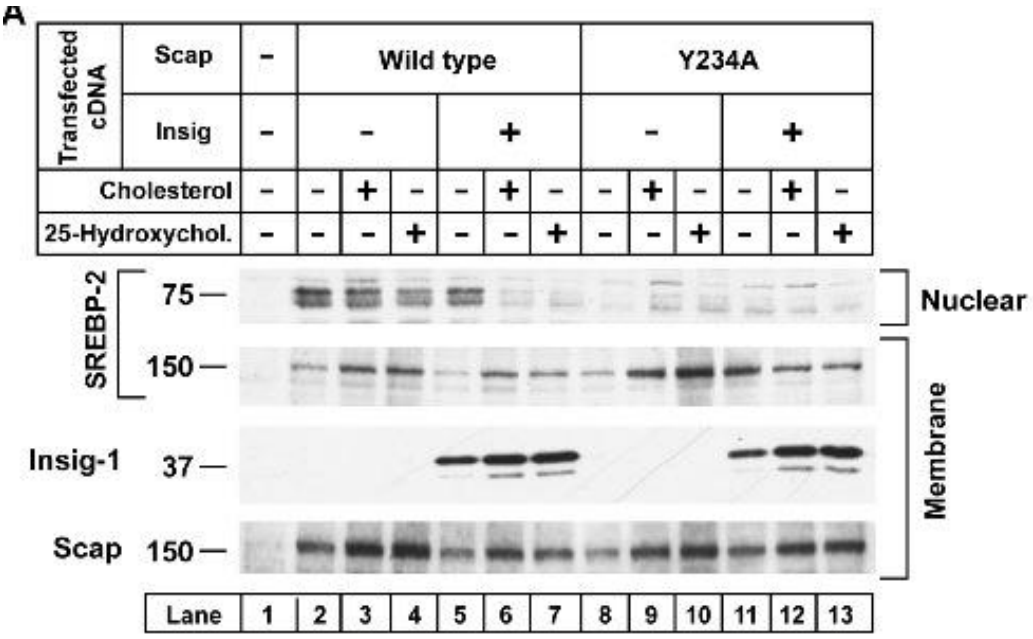



**Fig 11:** Biochemical characterization of the full length wt vs full length Y234A Scap (see Fig 10). Immunoblot of SREBP-2 in Scap-deficient cells transfected with WT or Y234A Scap (**full length versions**) in the absence or presence of transfected Insig-1 and in the absence or presence of cholesterol or 25-hydroxycholesterol. Nuclear or membrane indicate that separation of these cellular components has been achieved by centrifugation and then analyzed by SDS PAGE and western blotting. The TF part of SREBP moves to the nucleus and the rest remains in the ER membrane.

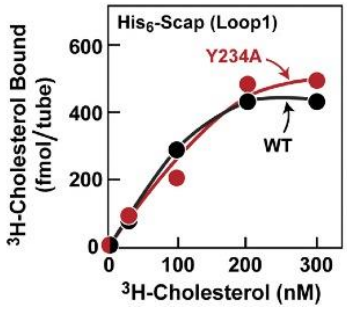



**Fig 12.** $(His)_6$-WT Scap (loop 1) and Y234A $(His)_6$-Scap (loop 1) were expressed and purified and cholesterol binding studied with the results shown.



**Questions**: **The experiments were designed to identify the mechanism of Scap sensing of cholesterol.** Note this data was **taken from *J. Biol Chem* 286, 18002 2011. The experiments and the conclusions by the authors are described in detail in this paper.**

1. A Loop 1 construct of Scap was prepared (46-269 residues) and contained a signal recognition sequence and a $(His)_6$-tag. Isolation of this construct required detergent. The experiments in Figures 2 and 3 were carried out to compare its properties with full length (TM1-8) Scaps. Describe what the results in these figures tell you.

Recombinant Scap Loop 1 binds [$^3$H]-cholesterol with saturation kinetics and a calculated $K_d$ of 67 μM. These results are similar to experiments carried out with binding studies on the entire membrane attached domain (TM domains 1-8). In Fig 9, a comparison of sterol binding (sterols analyzed are listed in Fig 8B-C) between these two Scap constructs is examined in an effort to tell how important Loop 1 is in sterol binding. In fact what they are looking at in this experiment is the ability of a variety of unlabeled Ch analogs at a fixed concentration to displace the bound [$^3$H] cholesterol. The results between full length and mutant Scaps are very similar with the Ch analogs indicated as red bars being effective competitors, blue bars being partial competitors, and black bars not very effective competitors. Since the Scap Loop 1 construct seems to behave similarly to the TM(1-8) Scap, the authors decided to continue work with this construct. They carried out scanning Ala mutagenesis studies to try to identify which groups might be important in binding of Ch and in the conformation of Scap that allows it to reorganize in the presence of Ch, to mask the MELADL domain in Loop 6, which binds to the machinery to transport Scap and SRE-BP to the Golgi through the COPII vesicles.

2. The requirement for detergent solubilization potentially raises concerns. What would these concerns be? How might one further characterize this new construct?

The small domain has no catalytic "activity" that can measured, only Ch binding assays are used as a method of characterization. The issue is whether this domain of Scap folds in detergent in the same way as it does in the membrane? Does this construct aggregate? Does the structure change with different detergents? In the paper they show similar Ch binding and saturation behavior, which seems unusual given that most of the transmembrane region of Scap is not present in the Loop 1 construct.

Further characterization of the new construct: In addition to binding cholesterol, Scap also interacts with Insig in the presence of cholesterol, its WD repeat domain interacts with the regulatory domain of SRE-BP and its (Scaps zip-code sequence (MELADL) in Loop 6 interacts with Sec23. In addition, depending on its conformation, Scap can be cleaved by different proteases (trypsin is used to cleave the RR sequence adjacent to the zip code domain for example in Loop 6 and a cleavage site also appears and disappears in Loop 7 depending on the conformation that is governed by sterol. (Note: you were not expected to know this, but proteases can be useful tools to help detect conformational changes.) Thus a comparison of similar experiments done on the detergent solubilized vs the ER Scap, could be informative. Mutants of this protein at these interaction sites and the consequences of the mutations in detergent solubilized and ER Scap would also support similar conformations. Membrane protein conformations (and quaternary structure) is a challenging problem.

3. An alanine scanning experiment is shown above (Figure 10 and Figure 11). This method was developed by the Wells group (UCSF) in an effort to define potential sites critical for ligand binding and function in the EGFR (epidermal growth factor receptor). It involves making many mutants where each residue or a cluster of residues is replaced with an ala. The mutants are then incorporated back into the cell from which the gene for Scap has been deleted and an assay is carried out to determine the consequences of the mutations.

The assay developed in this case takes advantage of the transcription factor (TF) properties of SRE-BP and its binding to SRE motifs which leads to transcriptional activation. Given the model described in lecture for regulation of cholesterol levels, explain the observed results with the WT Scap in the presence and absence of cholesterol (Figure 10). The two aberrant mutants are boxed in red (Figure 10). Provide an explanation for these results and for why Brown and Goldstein decided to study the Y234A-Scap in detail.

Figure 10 shows that w/t Scap leads to high levels of luciferase activity in the absence of sterol and low levels (20% wt) in the presence of sterol. This implies that at low Ch the SRE-BP got transported to the Golgi, got proteolytically processed to the transcription factor form, bound to the SRE of the luciferase gene in the nucleus and led to production of protein, which in the presence of its substrate gives off light which they can measure. Two of the mutants (red boxes) show no luciferase activity in the absence of Ch, while most of the other mutants behaved similar to the WT Scap. The hypothesis was that these mutants impaired the ability of Scap to move with SRE-BP to the Golgi complex by the CopII vesicles. They decided to focus on one of these mutants: Y234A.

4. To test their model the experiments described in Figure 11 were carried out. To think about this data, look at the model for regulation of TF release and the role of cholesterol and then compare the results with the Y234A-Scap FL(full length) mutant. You are looking at western blots of different players in this process under different conditions. Explain the data in each of the lanes of Figure 11 and how they relate to the WT model and proposed role for Loop 1 from all of the data.

The assay involves nuclear extracts and membrane pellets under each defined set of conditions that were analyzed by Western Blots in the presence or absence of Insig and in the presence or absence of sterol. [think about recitation this week) The important bands to look at are those of processed SREBP (75 kDa, nuclear, why there are two bands?, I do not know) vs full-length SREBP (150 kDa, membrane).
Lane 2: the SRE-BP in wt-cells in the absence of Insig and absence of Ch, moves to the nucleus as expected
Lane3/4: if Ch (or 25-OH-Ch) is present one sees some retention of SRE-BP in the ER in the absence of Insig.
Lane 5: in the presence of Insig and absence of Ch, most of the SRE-BP is found in the nucleus
Lane 6,7: when Insig and Ch (or 25-OH-Ch) are present, most of SRE-BP stays in the ER.

Thus the wt cells appear to behave as expected. Now one can compare these results with the Y234A mutant. A key result is that in the presence of Insig, regardless of whether Ch is present or not (lane 11 compared to lane 12/13), the SRE-BP remains in the ER (150 Kda). Thus this single mutation in loop 1 is altering Scap's ability to undergo the appropriate conformational change to bind the COPII machinery. Insig no longer has an effect on SRE-BP movement to the nucleus (compare results from lanes 8 and 9 with those in 11 and 12). Note the Western Blots with Insig1 antibody confirmed its presence or absence in these experiments (and same for Scap).

**The major conclusion from this study is that loop1 of Scap is a crucial component of the Ch-sensing mechanism**.

5. You are given that antibodies to all the players in cholesterol homeostasis are available. Given the experiments above and using Abs to Insig-1 for immunoprecipitation (IP) and SDS-PAGE/western analysis of the IP, what proteins would you expect to see in the presence and absence of cholesterol? Explain each prediction.

This experiment would allow you to determine if Y234A vs wt Scap binds to Insig in sterol depleted and sterol replete cells. The authors performed these experiments. They prepared detergent extracts of cells with Y234A or wt scap, performed the immunoprecipitation of Insig, and analyzed it by SDS-PAGE and Westerns. In sterol-depleted cells, the wt Scap did NOT co-immunoprecipitate with Insig. The addition of Ch, caused them to co-immunoprecipitate, indicating that wt Scap does not interact with Insig in the absence of Ch, but interacts in the presence of Ch, as expected. In contrast, Y234A Scap did co-immunoprecipitate with Insig in either the absence and the presence of Ch. These results again indicate that the loop 1 mutation alters the conformation of Scap even in the absence of Ch, such that it can interact with Insig regardless of Ch and also prevents interaction with the CopII targeting proteins.

# 5.08J Biological Chemistry II

Spring 2016