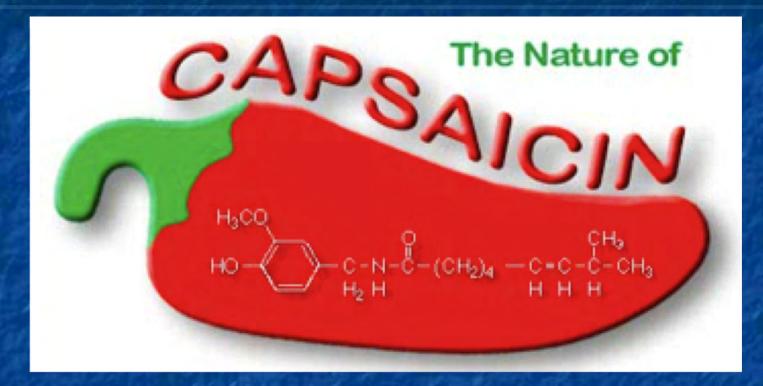
## MOLECULAR RECOGNITION: The Devil's in the details!



GARLAND R. MARSHALL, PH.D.
PROFESSOR OF BIOCHEMISTRY AND MOLECULAR
BIOPHYSICS AND OF BIOMEDICAL ENGINEERING
WASHINGTON UNIVERSITY, ST. LOUIS
USA

garland@biochem.wustl.edu

## ADD A LITTLE SPICE TO YOUR LIFE



In 1846, L.T. Thresh reported in the *Pharmacy Journal* that the pungent ingredient of peppers (capsicum) could be isolated in crystalline form and gave the name capsaicin.



### **CAPSAICIN**

- Although capsaicin has no odor or flavor, it is one of the most pungent compounds known, detectable to the palate in dilutions of one to seventeen million.
- Capsaicin was first synthesized in 1930 by E. Spath and F.S. Darling.

- Capsaicin prevents animals from eating chiles, so that they can be consumed by fruit-eating birds who specialize in red fruits with small seeds. Mammals perceive a burning sensation from capsaicin but birds do not. The seeds pass through the birds' digestive tract intact and encased in a perfect natural fertilizer.
- A person of average weight would have to consume nearly a half gallon of Tabasco® Sauce to overdose and become unconscious.



## **CAPSAICIN: A CHEMIST VIEW**

ani casa

- SUBNANOTECNOLOGY angstroms (Å), not nanometers (nm) - carboncarbon bond length = 1.54 Å or 0.154 nm
- Line-drawing structures curse of the Egyptians (papyrus and hieroglyphics); molecules are three-dimensional! And usually flexible, so shapes are not fixed.





	SOLUCIOSA	
Capsaicin		
HO	#	
Systematic name	(E)-N- (4-hydroxy- 3-methoxybenzyl) - 8-methylnon-6-enamide	
Molecular formula	C <sub>18</sub> H <sub>27</sub> NO <sub>3</sub>	
SMILES	CC(C)/C=C/CCCCC(NCC1=CC (OC)=C(O)C=C1)=O	
CAS number	[404-86-4]	
Molecular mass	305.41 g/mol	
Melting point	62 - 65 °C	

### **CAPSAICIN**

- Capsaicin is a hydrophobic molecule. Drinking water to reduce the burning caused by capsaicin is ineffective. Consuming foods high in fats and oils, such as milk or bread and butter, will help alleviate the burning. Alcoholic beverages also dissolve capsaicin due to the solvent characteristics of ethanol. Of course, over time the capsaicin will dissipate on its own. This illustrates that the physical properties of a molecule impact its behavior in biological systems.
- Biology often uses chemical messengers to signal between different tissues. Capsaicin interacts with a receptor on a neuron membrane. The burning and painful sensations of capsaicin result from binding to a receptor called the transient receptor potential vanilloid receptor subtype 1 (TRPV1) on sensory neurons. First cloned in 1997, TRPV1 is an ion channel-type receptor. TRPV1, which can also be stimulated with heat and physical abrasion, permits cations to pass through the cell membrane and into the cell from outside when activated. The resulting "depolarization" of the neuron stimulates it to signal the brain. Thus, capsaicin produces the same effect that excessive heat or abrasive damage cause, explaining why capsaicin is perceived as a burning sensation.
- Drugs often bind to receptors, either activating them as does capsaicin (agonists), or blocking activation (antagonists).



### **CAPSAICIN**

- Capsaicin is also the active ingredient in the riot-control agent pepper spray MACE®. When the spray comes in contact with skin, eyes or mucous membranes, it is very painful.
- Capsaicin was originally used in topical ointments to relieve the pain of peripheral neuropathy, for example post- herpetic neuralgia caused by shingles (*Herpes zoster*).
- With chronic exposure to capsaicin, neurons are depleted of neurotransmitters (specifically Substance P).
- Intranasal administration of capsaicin has shown some promise in treating certain kinds of headaches and chronic sinus infections. Is the cure worse than the disease?





## **TRPV1 Receptor**

- Activation -
  - Capsaicin
  - ■Resiniferatoxin (RTX) from the sap of resin spurge (Euphorbia resinifera), a cactus-like plant commonly found in Morocco
  - Temperature
  - Low pH (Acid)

#### Inhibitors.

Capsazepine

# Resiniferatoxis Capsaicin

Scutigeral

Capsazepine

#### Diverse structures act at same receptor.

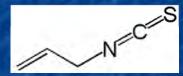
- Szallasi, A. and Blumberg, P.M. Resiniferatox in, a phorbol-related diterpene, acts as an ultrapotent analog of capsaicin, the irritant constituent in red pepper. Neuroscience 30: 515-520 (1989).
- Christopher S. J. Walpole, et al (1996). "Similarities and Differences in the Structure-Activity Relationships of Capsaicin and Resiniferatoxin Analogues". J Med Chem 39: 2939 -2952.



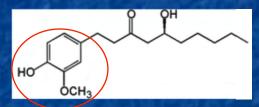
## **TRPA1** Receptor



- Activation Pungent natural products
  - Cinnamon oil (cinnamaldehyde)
  - Wintergreeen oil (methyl salicylate)
  - Clove oil (eugenol)
  - Mustard oil (allyl isothyocianate)
  - Ginger (gingerol)
  - Garlic (allicin)
  - Wasabi (allyl isothiocyanate)



ally I isothiocy anate



gingerol

alllicin

Things that are perceived very similarly (pungent, burning flavor) do not necessarily act by same mechanism. Receptors see compounds differently than structural diagrams.

## **TRPV1 Receptor**

- ABSTRACT Ion Channels: An Emerging Target (CHI, October 23-26, 2006, Boston, MA). TRPV1
   Antagonists as Anti-Hyperalgesics: Role of Differential Pharmacology
   . Narender Gavva, Ph.D., Senior Scientist, Department of Neuroscience, Amgen, Inc.
- TRPV1 is activated by chemical ligands (capsaicin, RTX, anandamide, and protons [pH < 5.7]), components of the inflammatory soup, and heat (>42° C), making it a molecular integrator of multiple noxious stimuli. TRPV1 antagonists representing different chemotypes that block all modes of activation produce anti-hyperalgesic effects in models of inflammatory, surgical-incision pain as well as analgesia in cancer pain models. Antagonists of TRPV1 that show differential pharmacology by their interaction through capsaicin-binding pocket were discovered by medicinal chemistry efforts. Mechanism(s) of differential pharmacology and the effects of TRPV1 antagonists that exhibit differential pharmacology on inflammation-induced hyperalgesia will be discussed.



November 4, 2008



Bio5476 - G.R.Marshall

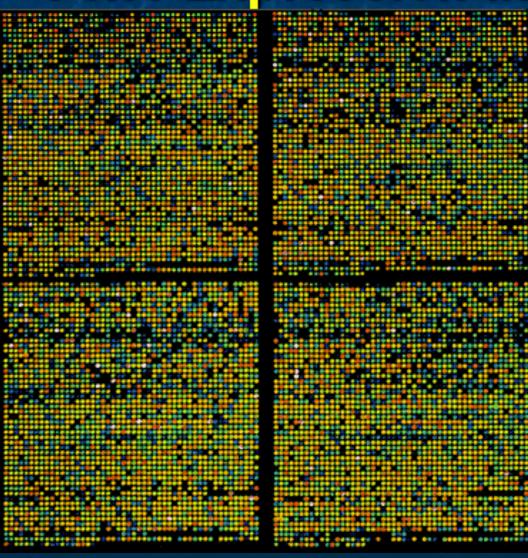




## **Medicinal Chemistry Efforts**

- Target Selection
- Computer-Aided Drug Design target structure known
- Pharmacophore concept and exploitation
- Lead Identification
  - Natural products
  - Combinatorial chemistry
- Optimization of lead structures
  - ADME (adsorption, distribution, metabolism, excretion)

## Gene Expression in Disease



Cyclooxygenase isozymes -

COX 1 - Normally expressed

**COX 2 - Expressed in inflammation** 

Specific (?) inhibitors of COX 2 - Celebrex, Vioxx

Non-specific inhibitors block prostaglandin in stomach and produce ulcers.



## Molecular Biology and AIDS (1983)

- COMPLETE SEQUENCE OF HIV GENOME (1985)
- IDENTIFICATION OF CODING REGIONS (GENES)
- DETECTION OF DTG MOTIF (ASPARTYL PROTEASES)
- KNOCKOUT MUTATION INHIBITED VIRUS (1988)
- CLONING AND EXPRESSION OF HIV PROTEASE
- USE WITH STRUCTURE-BASED DRUG DESIGN TO DEVELOP HIV PROTEASE INHIBITORS (1995)

## **HIV Life Cycle**

In targeting a pathogen such as HIV, some molecular aspect of the life cycle is selected to inhibit. Conventional wisdom suggests that a molecule/mechanism in the pathogen that is not present in the host is preferred to minimize side effects.

Life cycle of HIV each molecular event provides a potential target for therapeutic intervention. QuickTime™ and a Sorenson Video decompressor are needed to see this picture.

**Animation by Trimeris** 

November 4, 2008

Bio 5476 - Garland R. Marshall

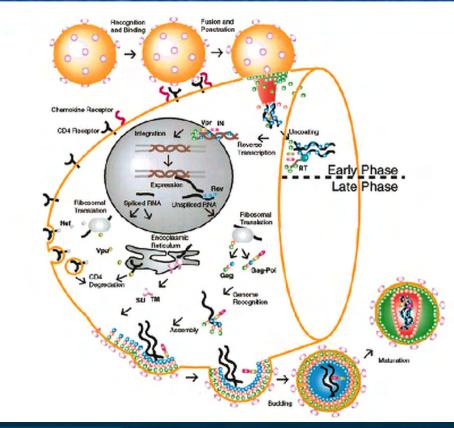
### Resistance

Targeting viral molecules quickly leads to resistance. Unless viral reproduction is inhibited completely, mutants in targeted proteins arise that

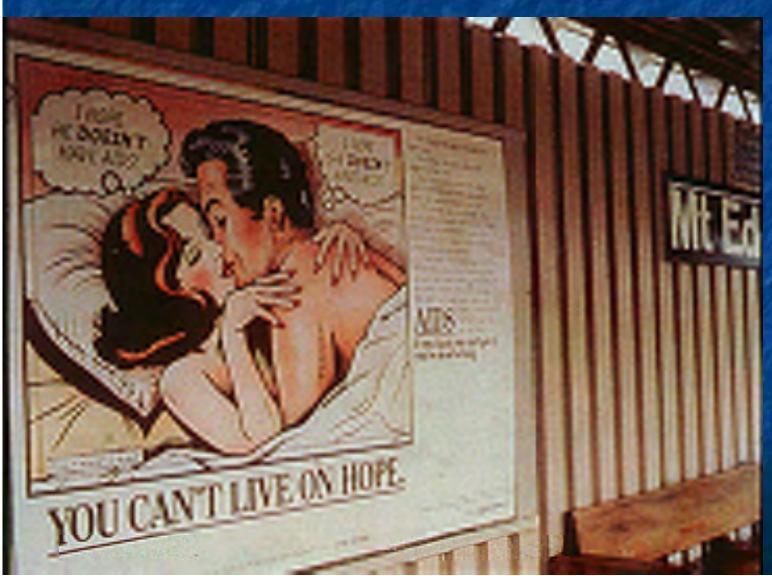
resist action of drug.

Virus mutation rate is very high and reproductive cycle is very short; leads to resistance to drugs.





## You can't live on hope!



Despite
advances in
therapy,
AIDS is
currently an
incurable
disease!

HIV is arguably a much greater threat to the global economy than terrorism!



#### RATIONALIZATION FOR CATALYSIS

Linus Pauling (1945) — "enzyme catalysis is based on stabilization of the transition state"

Corollary: a compound mimicking the geometry of the transition state of the enzymatic reaction should show tighter binding to the active site than substrate.



Transition-state inhibitors of amide hydrolysis have stable tetrahedral geometry at usual position of carbonyl carbon of amide bond.

## HIV Protease Inhibitors with Different Chirality at Transition State

Stable tetrahedral intermediate



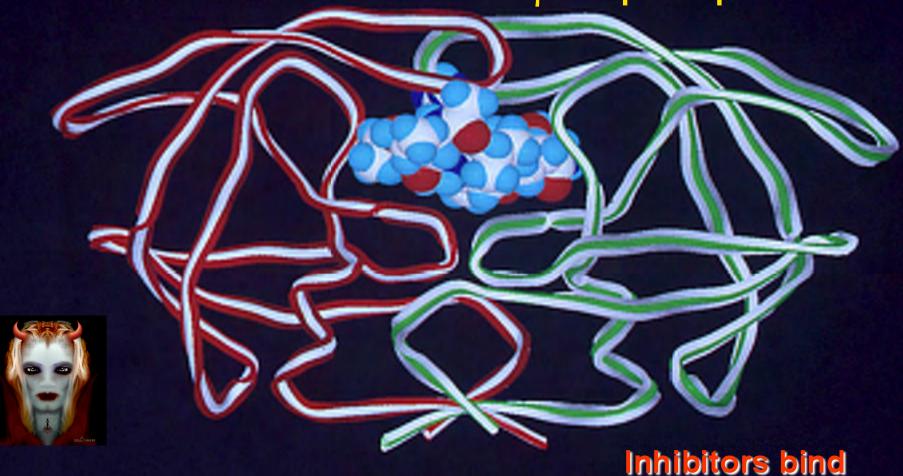
Prototypical HIV protease inhibitors with transition-state analogs

Active sites of enzymes are very stereoselective, aren't they?

Rich, D. H.; Sun, C. Q.; Vara Prasad, J. V.; Pathiasseril, A.; Toth, M. V.; Marshall, G. R.; Clare, M.; Mueller, R. A.; Houseman, K., Effect of hydroxyl group configuration in hydroxyethylamine dipeptide isosteres on HIV protease inhibition. Evidence for multiple binding modes. *J Med Chem* 1991, 34, (3), 1222-5.



## HIV Protease Complexed with Inhibitor MTV-101 β-hairpin flaps



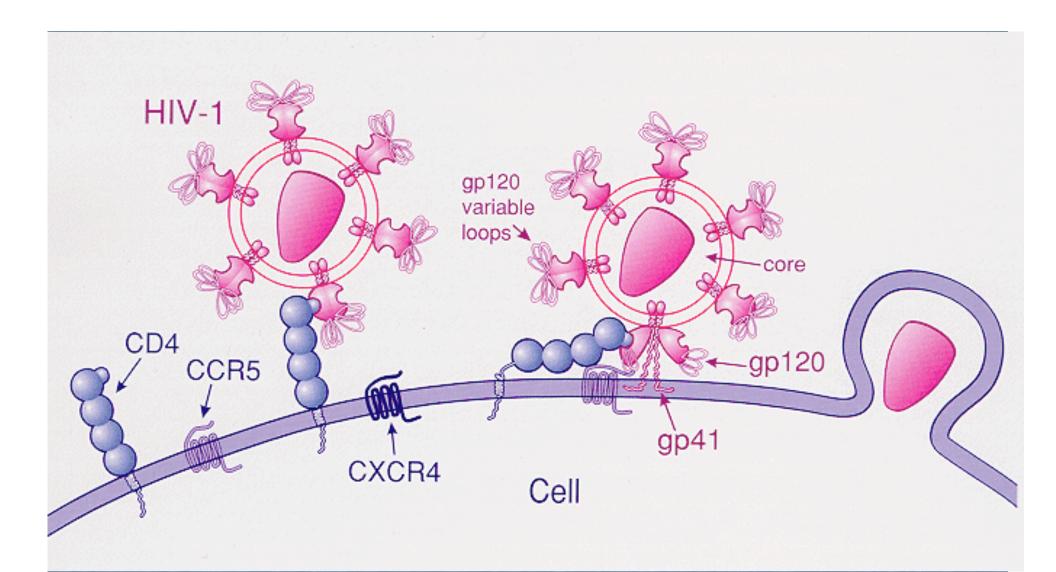
MVT-101 COMPLEX

with flaps down.

Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B.; Wlodawer, A., Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 A resolution. *Science* 1989, 246, (4934), 1149-52.

### Mammalian Proteins as Therapeutic Targets

- 1. CD4 protein on cell surface required, but not sufficient, for HIV infection
- 2. CCR5 and CXCR4 co-receptors G-coupled protein receptors that mediate chemokine signaling
- 3. Convertases proteolytic enzymes that process viral coat protein gp160 to gp120 and gp 41



Schematic diagram of viral entry into the cell. Either GPCR (CCR5 or CXCR4) can function as co-receptor depending on amino acid sequence of gp120 variable loops (from www.niaid.nih.gov/daids/dtpdb)

## ASKING 3D QUESTIONS OF RECEPTORS WHAT ARE THE RIGHT QUESTIONS?

Overall Objective – Determine recognition motif of receptor of unknown structure

Approach - Development of a variety of "conformational templates as molecular probes", i.e. relatively rigid scaffolds that satisfy at least three requirements:

- A. Possess limited 3D structures (a few well-determined 3D conformers);
- B. Be readily accessible synthetically, and
- C. Uniquely orient the peptide side chains (both  $\alpha$ - $\beta$  and  $\beta$ - $\gamma$  bonds) that are believed to transfer most of the information during peptide-receptor interactions.

### **Reverse-Turn Mimetics**

Cyclization strategy:

 $\begin{array}{c|c}
R_2 & H \\
\hline
CO & CO \\
HN & R_4
\end{array}$ 

Cyclic pentapeptides, pentaazacrowns (metal complexes)

The Marshall Lab (2001, 2003)

Introducing exocyclic rings (Pro):

**D-Pro-NMe-AA** 

$$\begin{array}{c|c}
Me & R_3 \\
\hline
-CO & CO \\
\hline
N & R_1
\end{array}$$

Pro-NMe-D-AA

The Marshall Lab (1995, 1998)

Reverse-turn populations

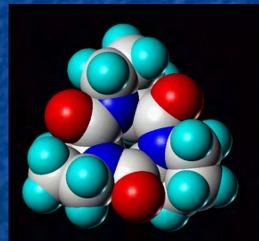
**D-Pro-Pro** > 60 %

**Pro-***D***-Pro** > 60 %

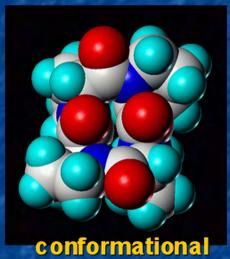
### **Explore Conformations of Cyclic Templates**

#### **Privileged Templates**

- minimize *intra*-molecular interaction
- maximize *inter*-molecular interaction



conformational fluctuation



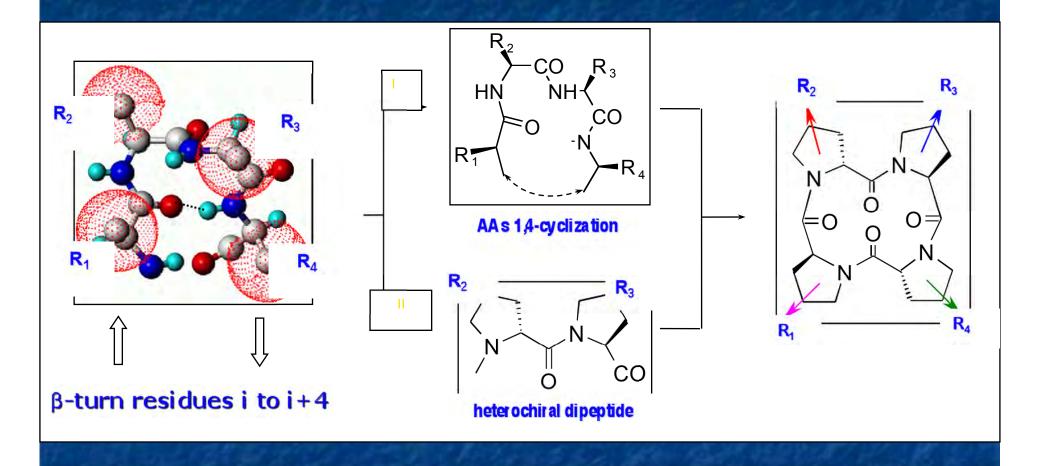
conformation change

Which cyclic peptides are suitable for templates?



#### **EVOLUTION OF CYCLIC TETRAPEPTIDE TEMPLATES**

(Che and Marshall, J. Med. Chem., 49:111-124 2006)



#### Cyclic Tetrapeptides Containing Chimeric Amino Acids

A cyclic tetrapeptide c(Pro-pro-Pro-pro) has been shown to adopt two conformers in solution, one with the amide bonds preceding the L-prolines in the cis-configuration and the other with the amides preceding the D-prolines having cis-amides. All seven protons (one a-proton, six side-chain protons) on the proline ring can be substituted with side-chain functionality to generate chimeric amino acids. We are exploring the conformational impact of hydrogen substitution and change of  $\alpha$ -carbon to nitrogen (azPro) as well as mapping the potential surface to determine the relative stabilities of conformers.

$$\frac{3}{4} \frac{1}{\sqrt{\frac{4}{N}}} \frac{1}{\sqrt{\frac{2}{N}}} \frac{1}{\sqrt{\frac{4}{N}}} \frac{1}{\sqrt{\frac{4}{N}}}} \frac{1}{\sqrt{\frac{4}{N}}} \frac{1}{\sqrt{\frac{4}{N}}} \frac{1}{\sqrt{\frac{4}{N}}} \frac{1}{\sqrt{\frac{4}{$$

The ctct and tctc conformers of cyclo-(DLDL)-Pro4. The D-prolines are in blue, and the L-prolines are in red. The cyclic peptides are numbered, shown in the center of each pyrrolidine ring. The two conformers isomerize at 45° C in DMF and 80° C in H<sub>2</sub>O, with  $\Delta H^{\sharp} = 44$  and 28 kcal mol<sup>-1</sup> and  $\Delta S^{\sharp} = 59$  and 22 cal K<sup>-1</sup> mol<sup>-1</sup>, respectively (Mastle et al. 1991).

## cyclo-(DLDL)-Pro4



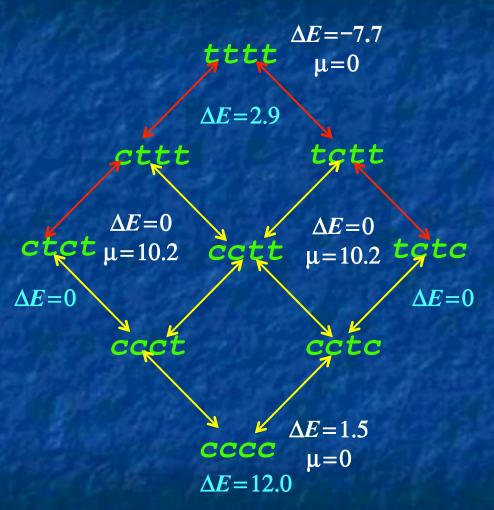


all trans-amide conformer

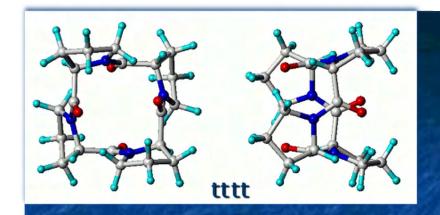


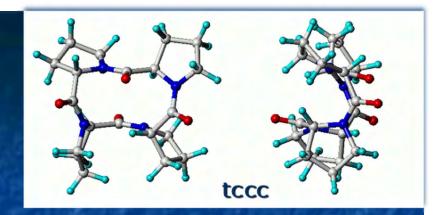


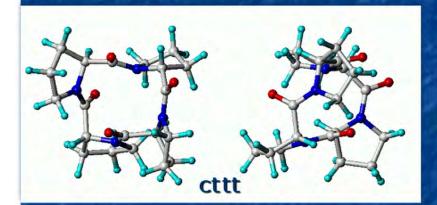
all as-amide conformer

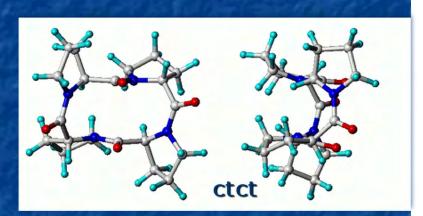


DFT calculations (B3LYP/6-31G\*):  $\Delta E$  (kcal/mol) (white = in vacuo, blue = water) and  $\mu$  (debye)

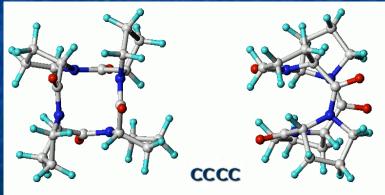








#### Amide-bond Conformers Of c[Pro-pro-Pro-pro]

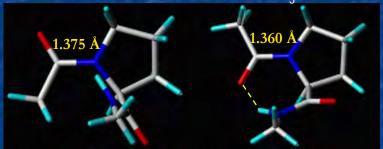


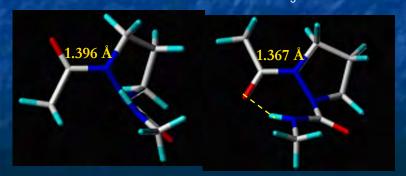
(Ye and Marshall, Engineering cyclic tetrapeptides containing chimeric amino acids as preferred reverse-turn scaffolds. J. Med. Chem 49:111-124 2006)

### The Impact of azPro on *cis⇔trans* Isomerism

CH<sub>3</sub> IN Ac-Pro-NHMe

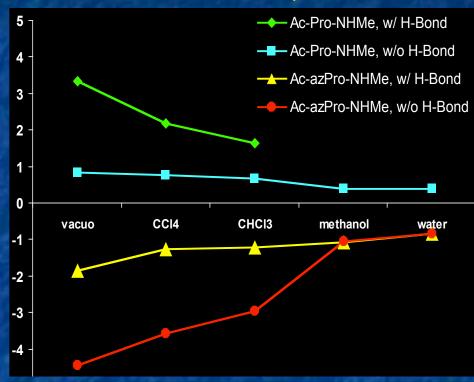
H<sub>3</sub>C-N H CH<sub>3</sub>





Che and Marshall, Impact of azaproline on peptide conformation. J Org Chem 2004, 69, (26), 9030-42.

∆G (cis - trans) kcal/mol

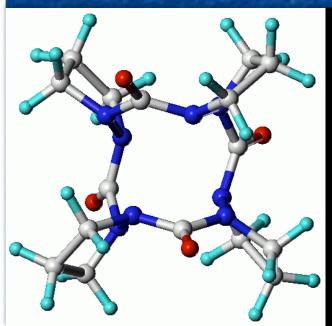


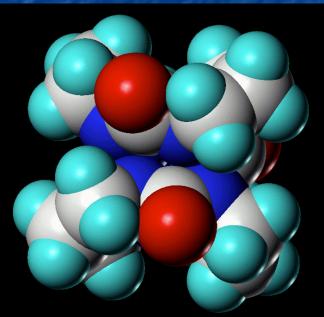
DFT calculations (B3LYP/6-31+G")

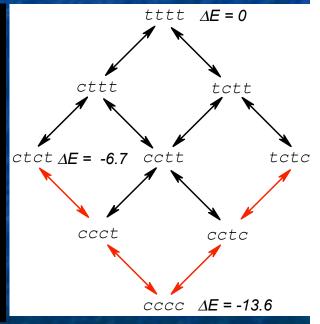
If carboxymide group of azPro has hydrogen, internal hydrogen bonding in *trans* dominates.

Bio5476 - G.R.Marshall

The DFT potential surface of cyclo-azPro4 with structure of the most stable ccc-amide conformer. The probable transition path (red) between ctct and tctc, which is through the ccc-amide isomer, is "upside down" compared to that of cyclo-(D-pro-L-Pro-D-pro-L-Pro).







# Cyclic Tetraproline as a Privileged Reverse-Turn Scaffold (side chains replacing any of seven hydrogens on proline, or pipecolic acid, generate chimeric amino acids)

## Constraints on side-chain torsion angles in proline and pipecolic chimeric amino acid analogs

$$\chi^{1}$$
 $\chi^{1}$ 
 $\chi^{1$ 

## How to Stabilize Different Conformers of Cyclic Tetraproline Analogs?

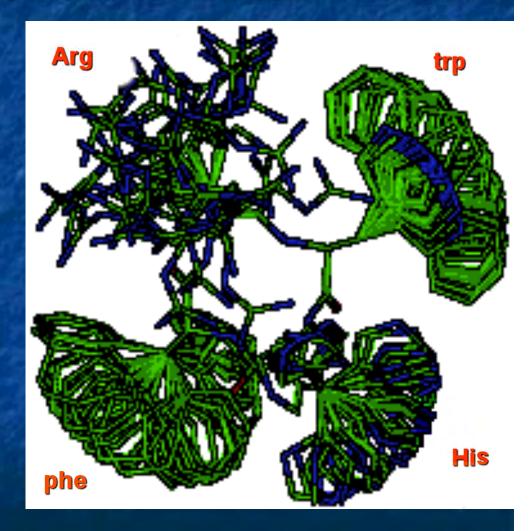
- 1. Pattern of substituents test with single methyl substitutions All 6 positions of each D or L proline in the three different conformers (tttt, ctct and tctc) 36 compounds were minimized and relative energies compared; significant impact on relative stabilities detected.
- 2. Change hybridization of backbone replace  $\alpha$ -carbon with nitrogen (aza-amino acids). The potential surface of c(azPro)<sub>4</sub> was examined by DFT calculations (ccc conformer stabilized).
- 4. Replace cyclic amino acids such as Pro and Pip (6-membered ring) with N-methyl amino acids Pro replacements stabilize tttt conformers.

(Che, Y.; Marshall, G.R., Engineering cyclic tetrapeptides containing chimeric amino acids as preferred reverse-turn scaffolds. J Med Chem 2006, 49, (1), 111-24.)

Solution NMR structure of c[His-phe-Arg-trp] in DMSO. The cyclic tetrapeptide containing alternating D-and L-amino acids has four trans-amide bonds (as defined by coupling constants and the absence of CaH-CaH cross peaks). There are 50 structures represented in the figure and all have an identical backbone conformation. The structure found emphasizes a very rigid backbone with side chains oriented in unique relative positions for molecular recognition. Note limited options for side-chain interactions with a receptor.

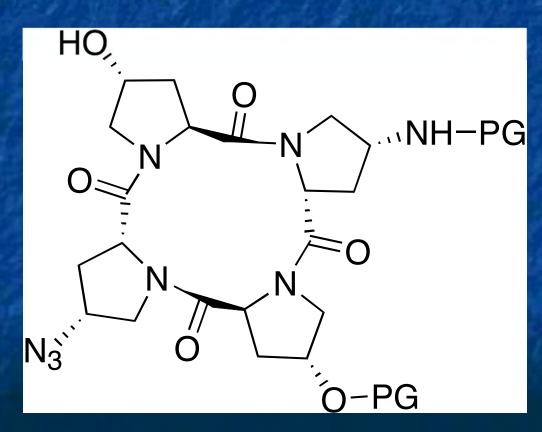
By mapping the potential energy surface and the transition paths between conformers, we can readily estimate the number of conformers, their distributions and half-lives that should be observed experimentally. This allows a rational basis for selection of substitution patterns on templates to stabilize, or destabilize, particular conformers.

Experimental Structure of cyclic tetrapeptide in *tttt* conformation (Smythe et al., unpublished)



### **C**imeric Tetraproline Libraries

- Provide unique side-chain orientations on rigid scaffold
- Synthetically accessible privileged scaffold
- Select chimeric proline analogs so that each R position can be selectively modified



- Alkylations
- Coupling
- Staudinger Ligation

Easy to design; often hard to make!

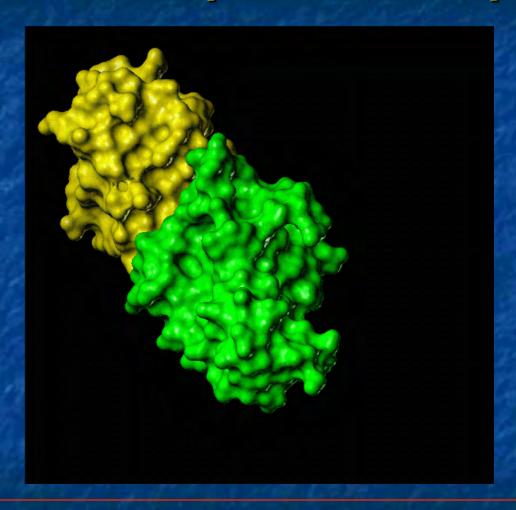


## Synthesis of Chimeric *Trans* and *Cis* L-Proline Analogs

a . Tet. Lett., 2001, 2459-2460

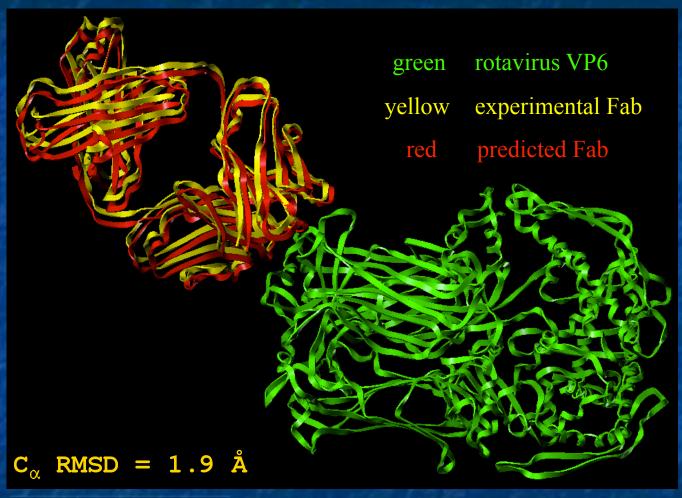
<sup>&</sup>lt;sup>b</sup> J. Peptide. Res. 2005, 65, 298–310.

## **Predict Protein-protein Complexes**



Which one is the native complex?

### **CAPRI Target 2: Bovine Rotavirus VP6/Fab**



Correct interface Fab residues VP6

Fab 27/27 VP6 26/27

November 4, 2008

Bio5476 - G.R.Marshall

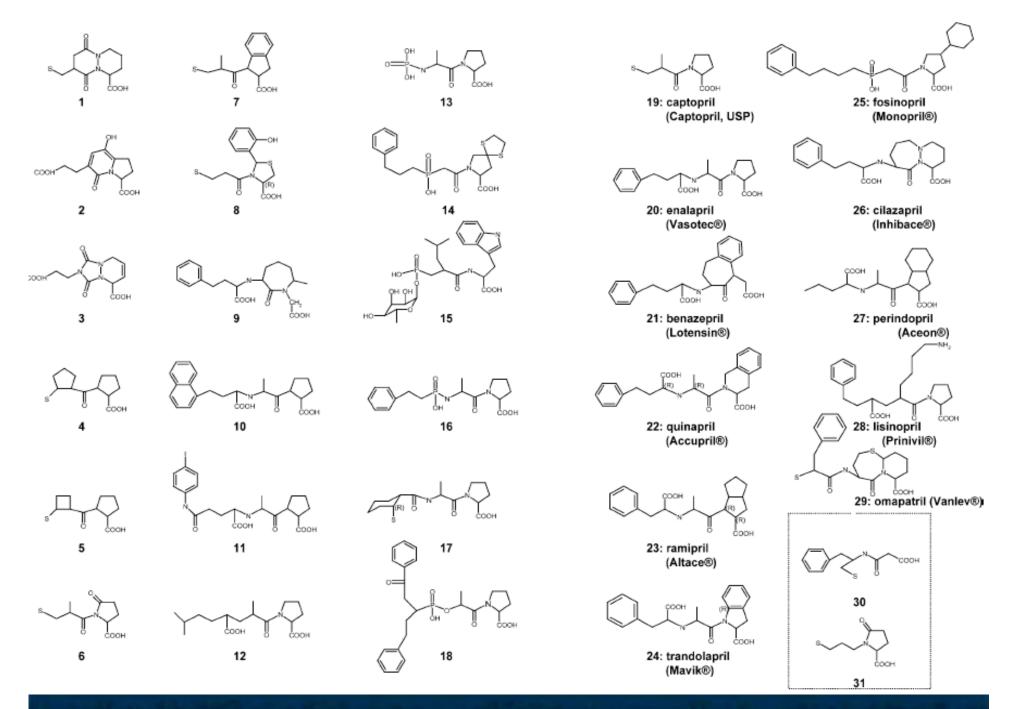
## Conformational Changes Limit the Accuracy of Current Approaches



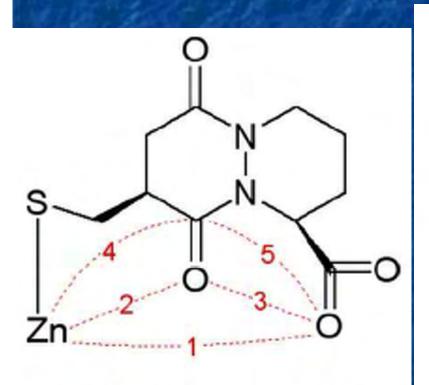
superimposed
crystal
structures of
CheY from 10
different
complexes
Proteins are dynamic
structures and undergo
changes when they
interact with other
molecules (induced fit/
allosteric changes)

## Pharmacophore versus Active-Site Modeling

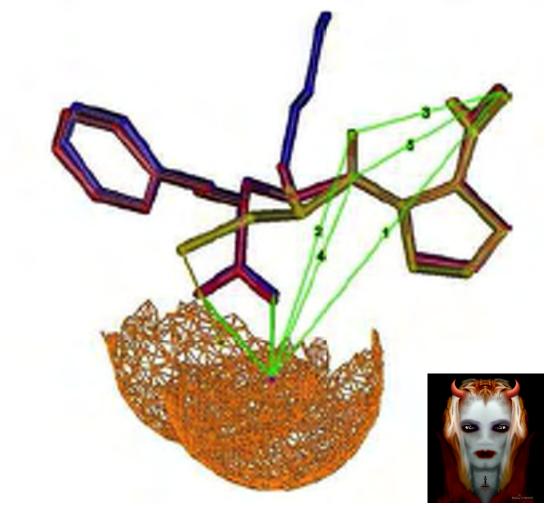
- A. Pharmacophore assumes physical overlap of similar groups from different ligands when bound.
- B. Active site is fixed and common to all bound ligands and optimal geometry of interaction is assumed.



## Kuster and Marshall. Validated Ligand Mapping of ACE Active Site. J Comp-Aided Drug Des 2005, 19, 609-615.



1:	Predicted model 7,188 - 7,812	Crystal structures 8,485 - 8,673	Deviation ~1.1
2	4.812 - 5.312	5.718 - 5.981	~0.6
3:	3,562	3.532 - 3.628	~0.0
1:	4.812 - 5.062	4.876 - 5.181	~0.1
5:	3.938	3.989 - 4.047	~0.0



Active site is not fixed in ACE; zinc moves to accommodate chemistry

## Volume Mapping

- Receptor must have space to bind active ligands.
- Inactivity can be due to steric repulsion even if correct pharmacophore can be presented.

Volume Union of active compounds



Sufrin, J. R., Dunn, D. A. & Marshall, G. R. (1981). Steric mapping of the L-methionine binding site of ATP: L-methionine S-adenosyltransferase. Mol. Pharmacol. 19, 307-313.

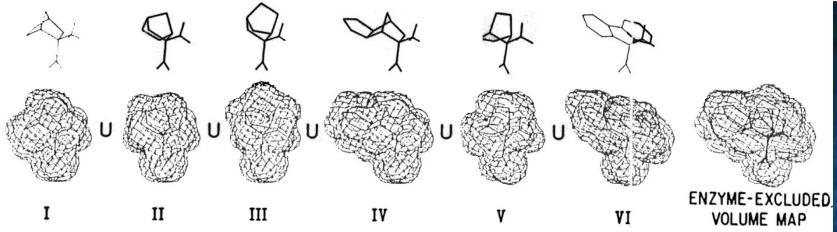


Fig. 2. Enzyme-excluded volume map

The structures of the six active analogues are shown on top; below each structure is the electron density map for that molecule. The union of these six individual electron density maps gives the enzyme-excluded volume map and defines that region of the methionine binding site available for binding by substrate, substrate analogues, or nonsubstrate analogues.

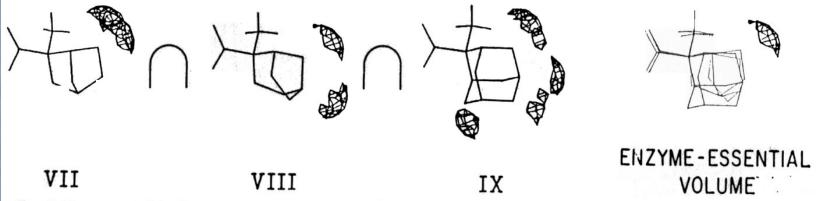


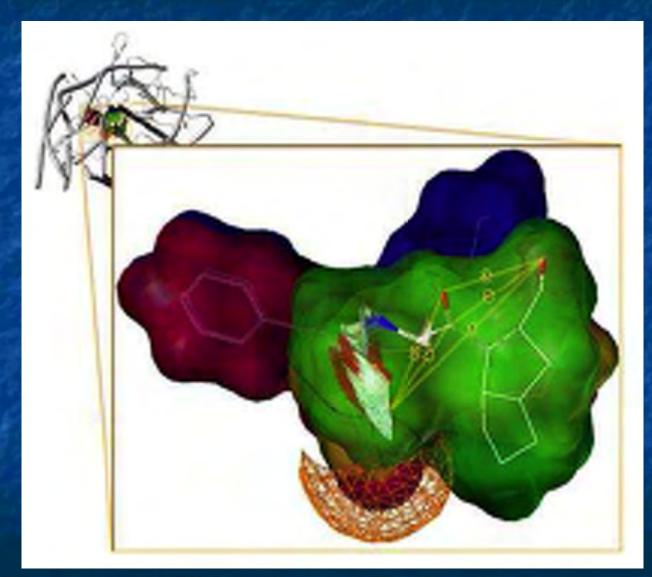
Fig. 3. Enzyme-essential volume map

Each inactive molecule is shown with its unique volume segments that are not part of the enzyme-excluded volume map. Intersection of the unique volume segments of each inactive analogue gives one region of unique volume overlap for all three molecules which defines the enzyme essential volume, i.e., a region occupied by the enzyme and therefore not available for occupancy by other molecules.

Sufrin, J. R., Dunn, D. A. & Marshall, G. R. (1981). Steric mapping of the L-methionine binding site of ATP: L-methionine S-adenosyltransferase. Mol. Pharmacol. 19, 307-313.

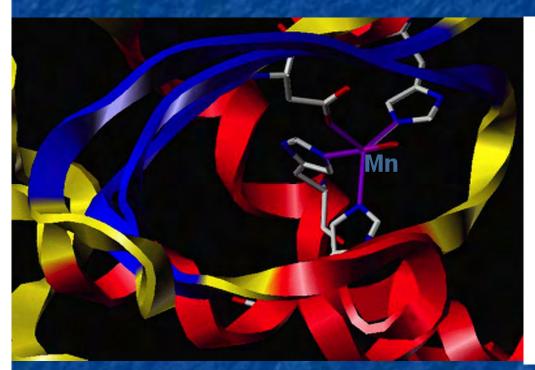
## Kuster and Marshall. Validated Ligand Mapping of ACE Active Site. J Comp-Aided Drug Des 2005, 19, 609-615.

Conformations of Ramipril that fit ACE active-site model. Volume requirements for Ramipril fit within known combined volume map for other ACE inhibitors.

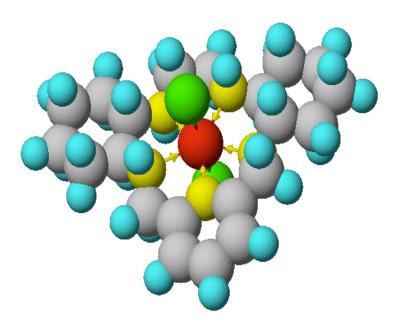


## M40403 IS NOT A COPY OF ACTIVE SITE OF MN SOD, BUT A MIMIC OF CATALYTIC MECHANISM

(designed by Dennis P. Riley, Ph.D., Metaphore Pharmaceuticals)



Active Site of human SOD – 3 His nitrogens and 1 Asp oxygen coordinate with Mn



M40403 has 5 nitrogen ligands to Mn from its pentaazacrown scaffold

#### Stereochemical Effects



#### **KEY INSIGHT PROVIDED BY DIFFERENCES IN ACTIVITY**

Small differences in stereochemistry can eliminate SOD activity of synzyme.

$$k_{cat}$$
 (pH=7.4)= 1.2 x 10<sup>+8</sup> M<sup>-1</sup> s<sup>-1</sup>

k<sub>cat</sub> (pH=7.4): Inactive

R,R,S,S,-isomer is rigidly planar; i.e., will not fold, thus no catalytic activity

Acknowledgments

Graduate Students - Rieko Arimoto (BME, Vertex), Lori Anderson (Bioorganic), Eric Welsh (Biophysics), Matt Anderson (Bioorganic), Ye Che (Biophysics, NIH), Sage Berg-Cross (Duke, Biochemistry), Dan Kuster (Iowa State, BME), J.W. Feng (Ohio State, CompBio), Yat Tang (UCLA, CompBio), Rob Yang (Alberta, CompBio) Xiaoming Zhang (Alaska, CompBio), Lee Tessler (Brown, CompBio), Aggie Bielska (Delaware, MSTP), Aaron Spivak (Davidson, CompBio), Rich Seeger (Georgia Tech, BME), Kenneth Borrelli (Duke, BME)

Collaborators - Prof. Gregory Nikiforovich (WUMS) - GPCR/peptide Modeling

Dr. Greg Bourne (Cambridge, WUMS) - Combinatorial Chemistry

Dr. Abby Fisher (Minnesota, WUMS) - Chimeric Proline Tetrapeptides

Dr. Yaniv Barda (Hebrew U, WUMS) - Targeted Cyclic Azacrowns

Dr. Christy Taylor (MIT, WUMS) - Rhodopsin/Tranducin Interface

Profs. David Cistola (WUMS) & Oleg G. Kisselev (SLU) - TrNOE NMR

Prof. Tom Baranski (WUMS) - Molecular Biology of GPCRs

Prof. Steve O. Smith (SUNY Stony Brook) - MAS NMR Studies

Prof. Wayne L. Hubbell and Dr. Ned Van Eps (UCLA) - ESR Studies

Prof. Victor Guallar (WUMS) - QM/MM

Many others in the past!