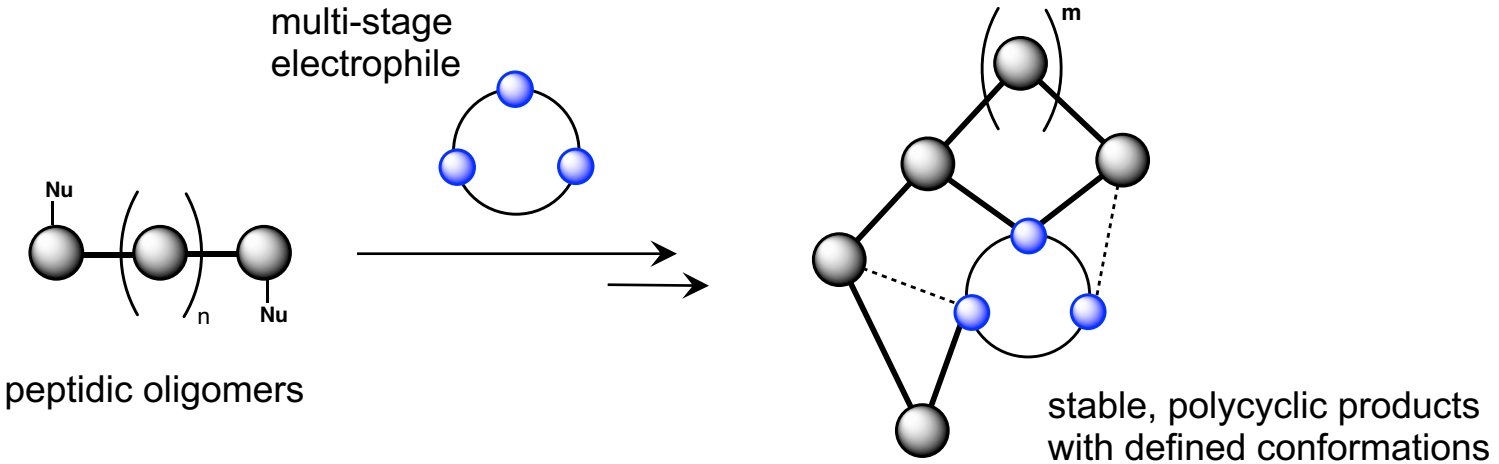
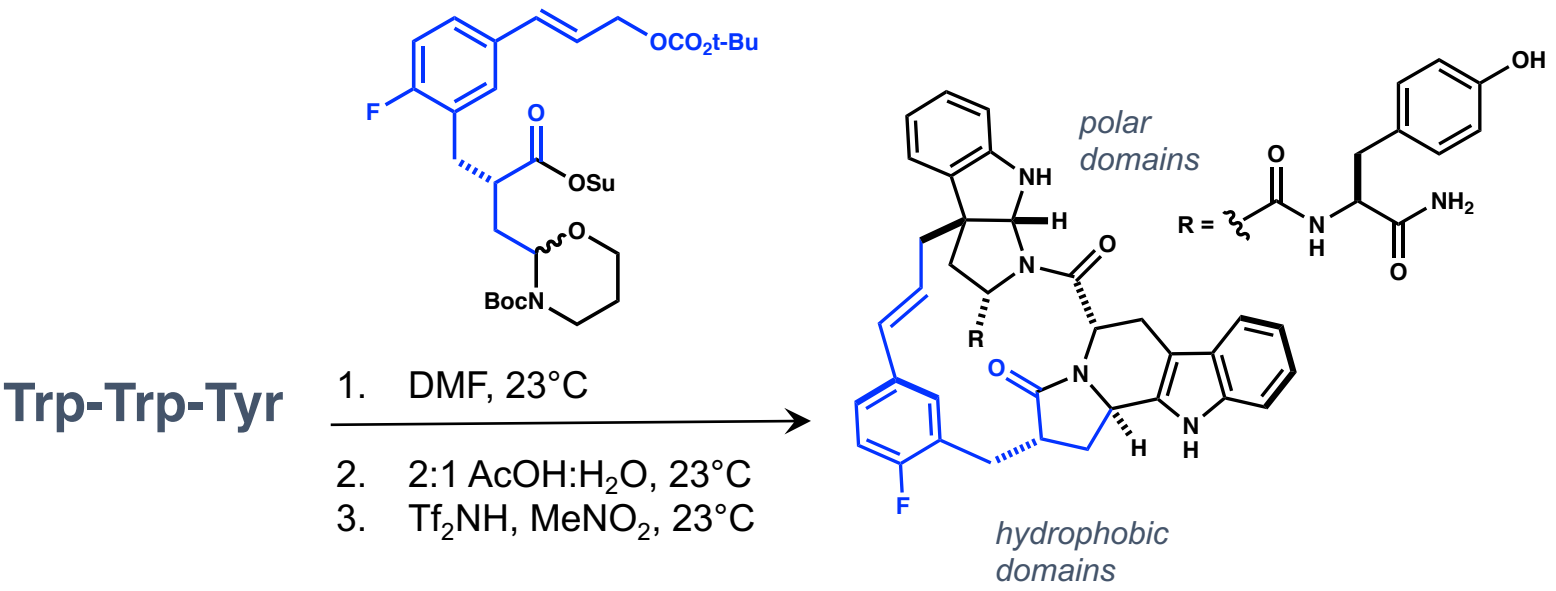
EXPLORING A HYPOTHETICAL GIGALIBRARY FOR NOVEL STRUCTURES AND FUNCTIONS.

Ishika Saha

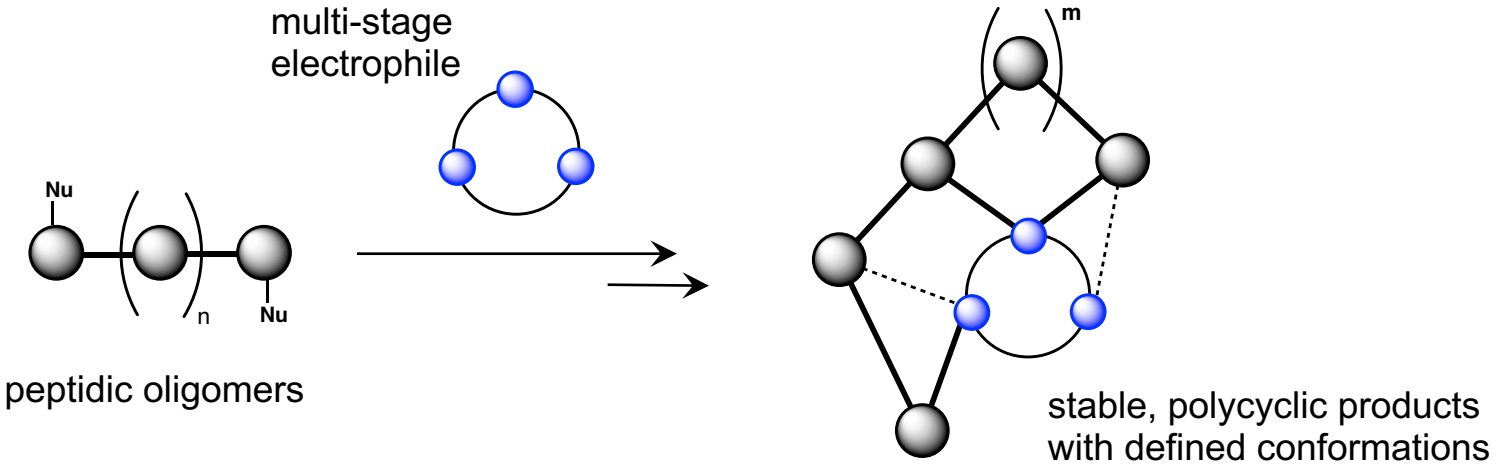
03/09/2022

Complex Molecules Made Simply

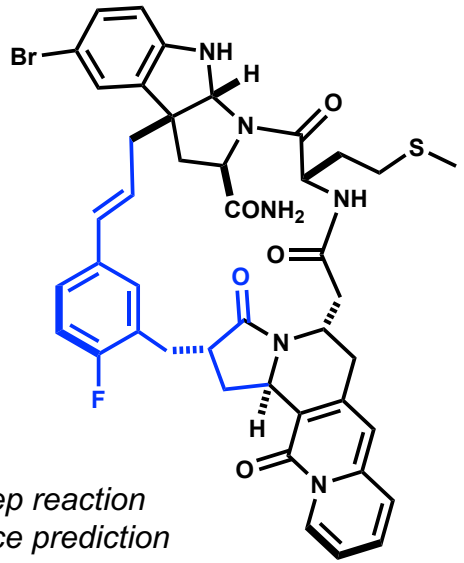




Complex Molecules Made Simply

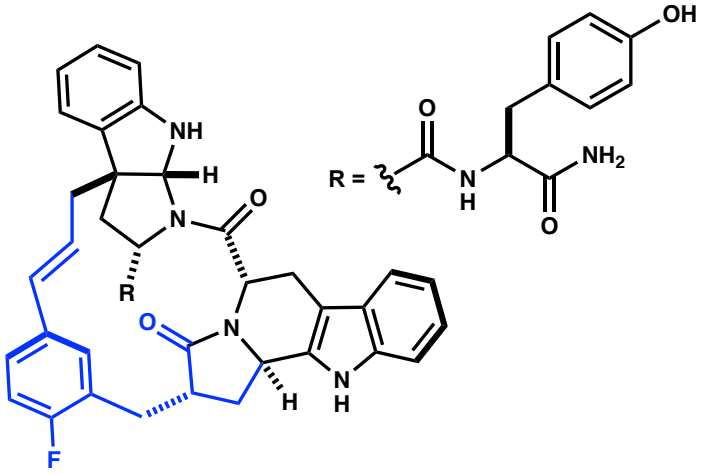


Computationally simulated structural variant



Multi-step reaction sequence prediction

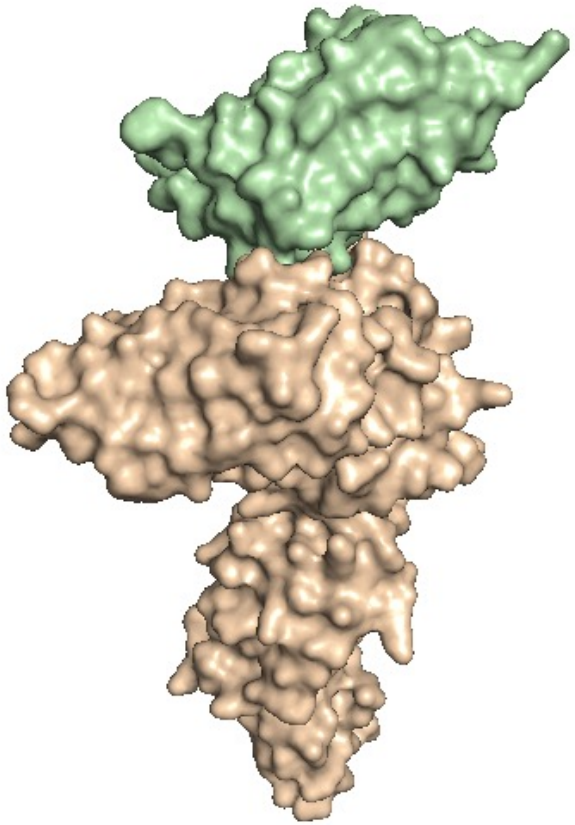
Experimentally synthesized and characterized



Identifying Surrogates of Regulatory Proteins

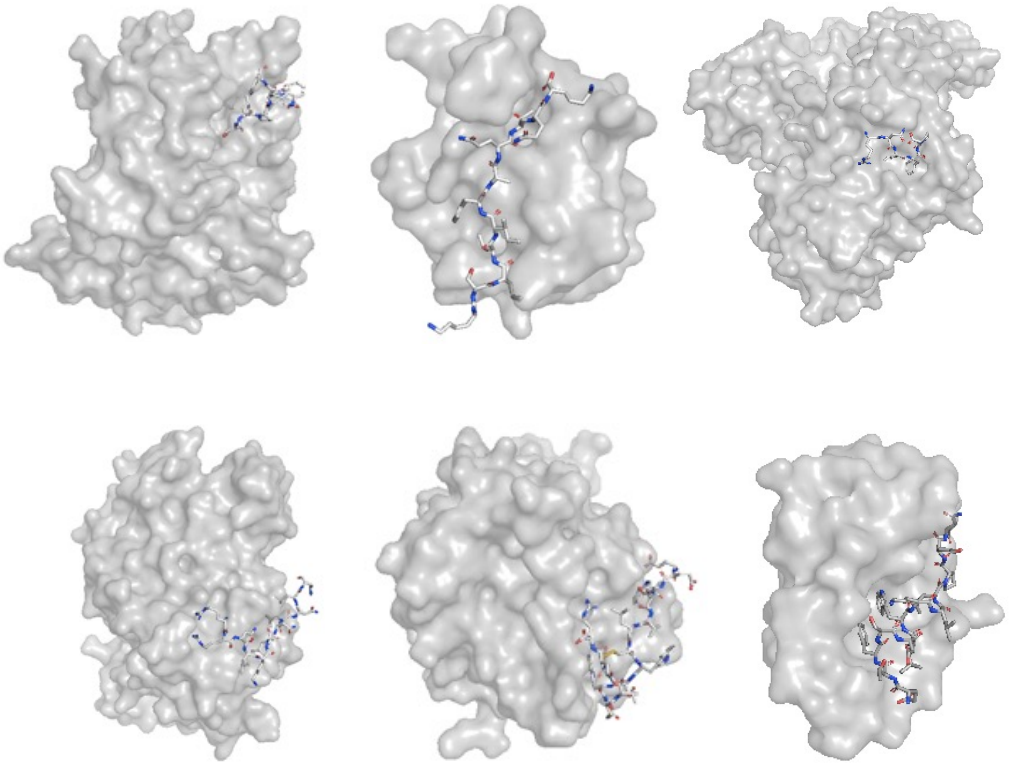
UCLA Chemistry & Biochemistry

Protein – Protein Interactions



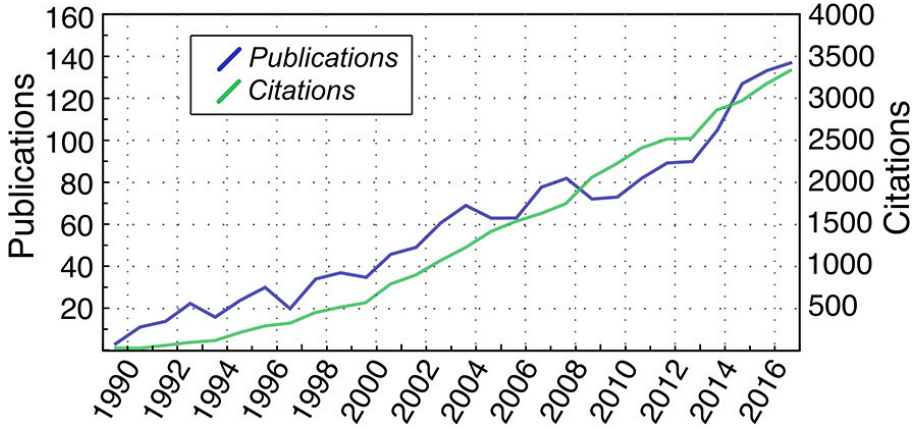
Proc. Natl. Acad. Sci. USA (2008) 105 (8), 3011-3016

Protein – Peptide Interactions

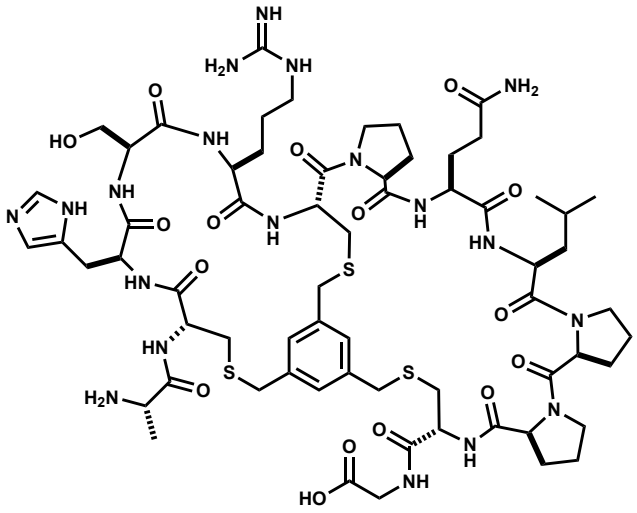


Curr Opin Struct Biol. (1995), 1, 103-13
Structure (2010) 18, 188–199

Macrocyclic Peptides in Drug Discovery

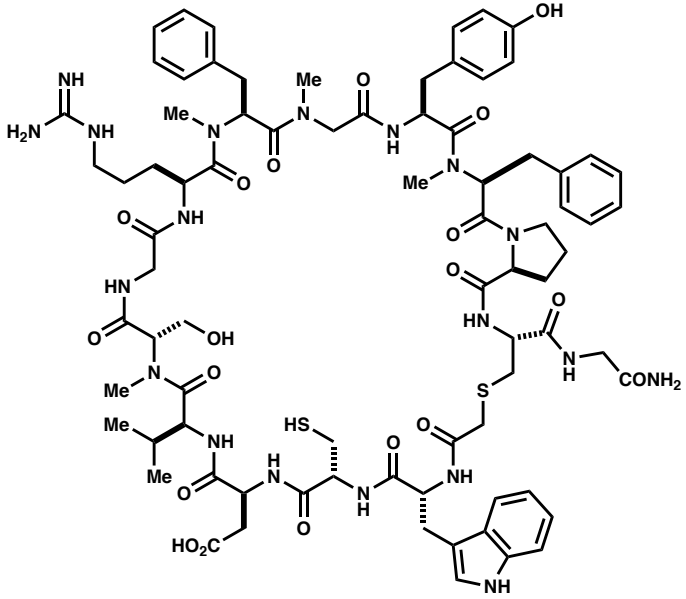


J. Am. Chem. Soc. (2019) 141, 10, 4167–4181



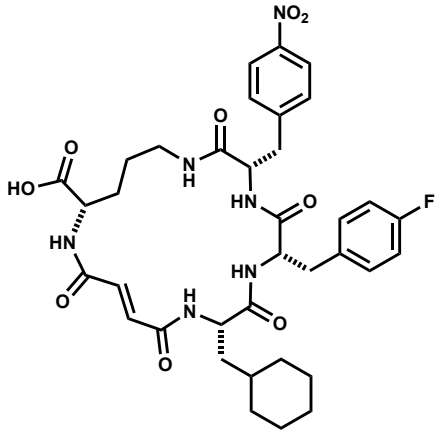
Sortase A inhibitor, phage display screen

ACS Med Chem Lett. (2016) 7(6), 606–611



E6AP Ub ligase inhibitor, RaPID screen

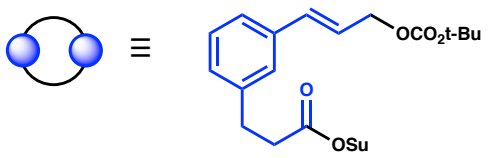
Chem. Biol. (2011) *18* (12), 1562–1570



SRC inhibitor, DNA-templated screen

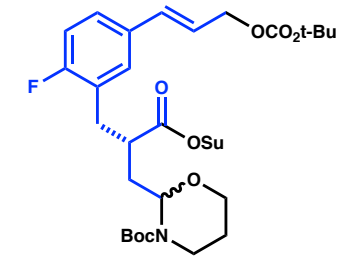
Nat Chem Biol. (2012) 8(4), 366–374

Alternative Template Based Method



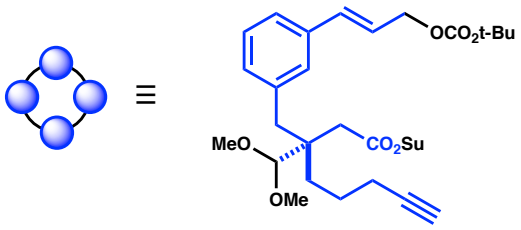
G1

Proc. Natl. Acad. Sci. USA (2013) E3753
Tetrahedron (2013) 7683
Chem. Sci. (2016) 4158



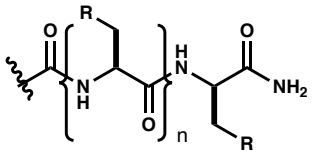
G2

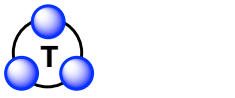
Chem. Sci. (2015) 2219

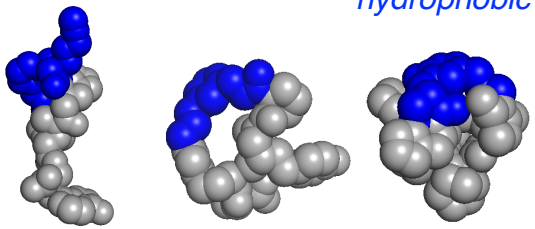


G3

J. Org. Chem. (2018) 3090







rod disk sphere

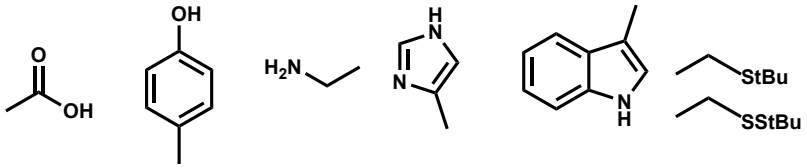
hydrophobic domains residual from template **T**

stable, polycyclic products with defined conformations

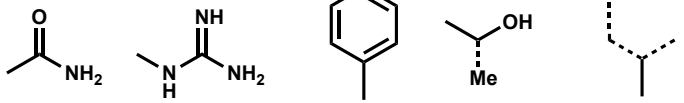
polar domains residual from oligomer

where R =

Participants



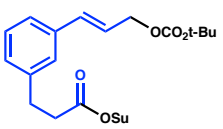
Spectators



Representative Outcomes

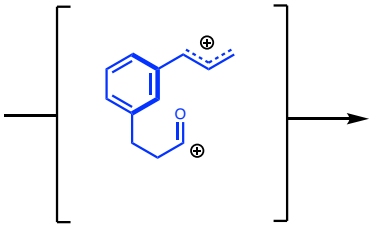
UCLA Chemistry & Biochemistry

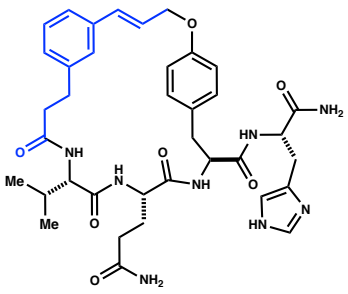
TEMPLATE | **INTERMEDIATE REACTIVITY** | **REPRESENTATIVE PRODUCTS**



G1

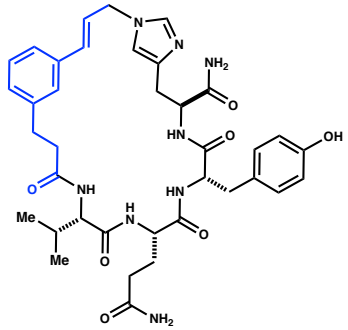
Proc. Natl. Acad. Sci. USA (2013) E3753
Tetrahedron (2013) 7683
Chem. Sci. (2016) 4158





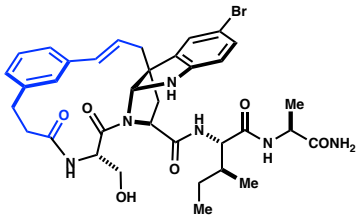
C-O bonded, switchable selectivity

1



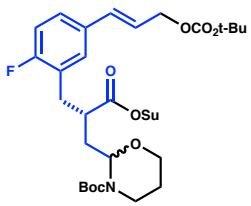
C-N bonded

2



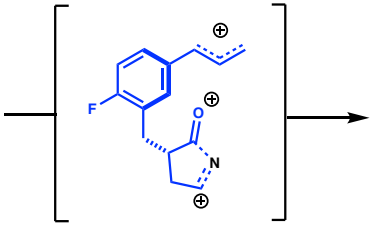
C-C bonded, ansa-bridged

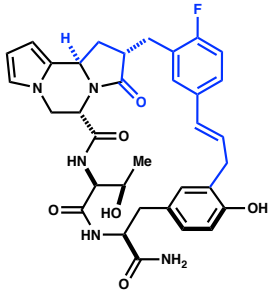
3



G2

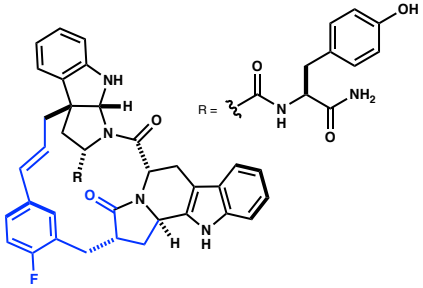
Chem. Sci. (2015) 2219





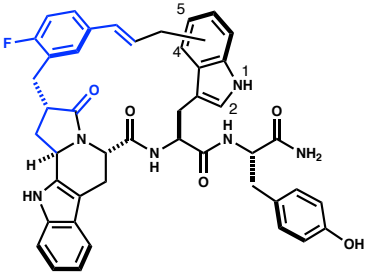
polycyclic, C-C bonded

4



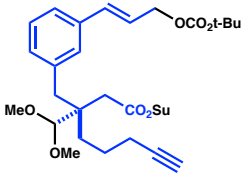
polycyclic, ansa-bridged

5



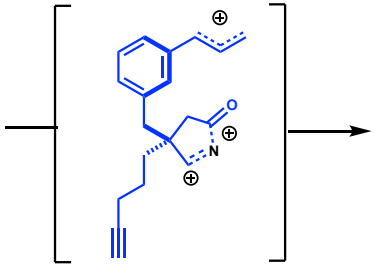
variable macrocycle size

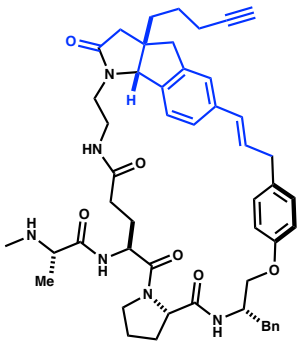
6

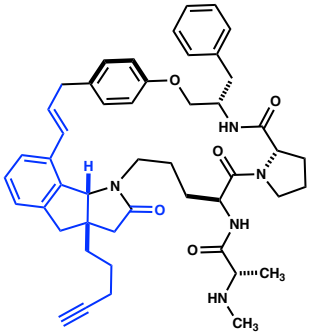


G3

J. Org. Chem. (2018) 3090

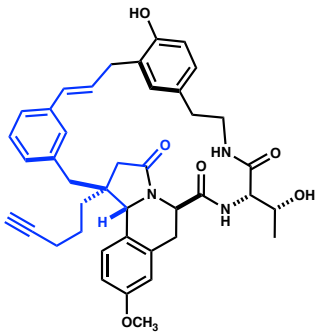






further diversification through template engagement

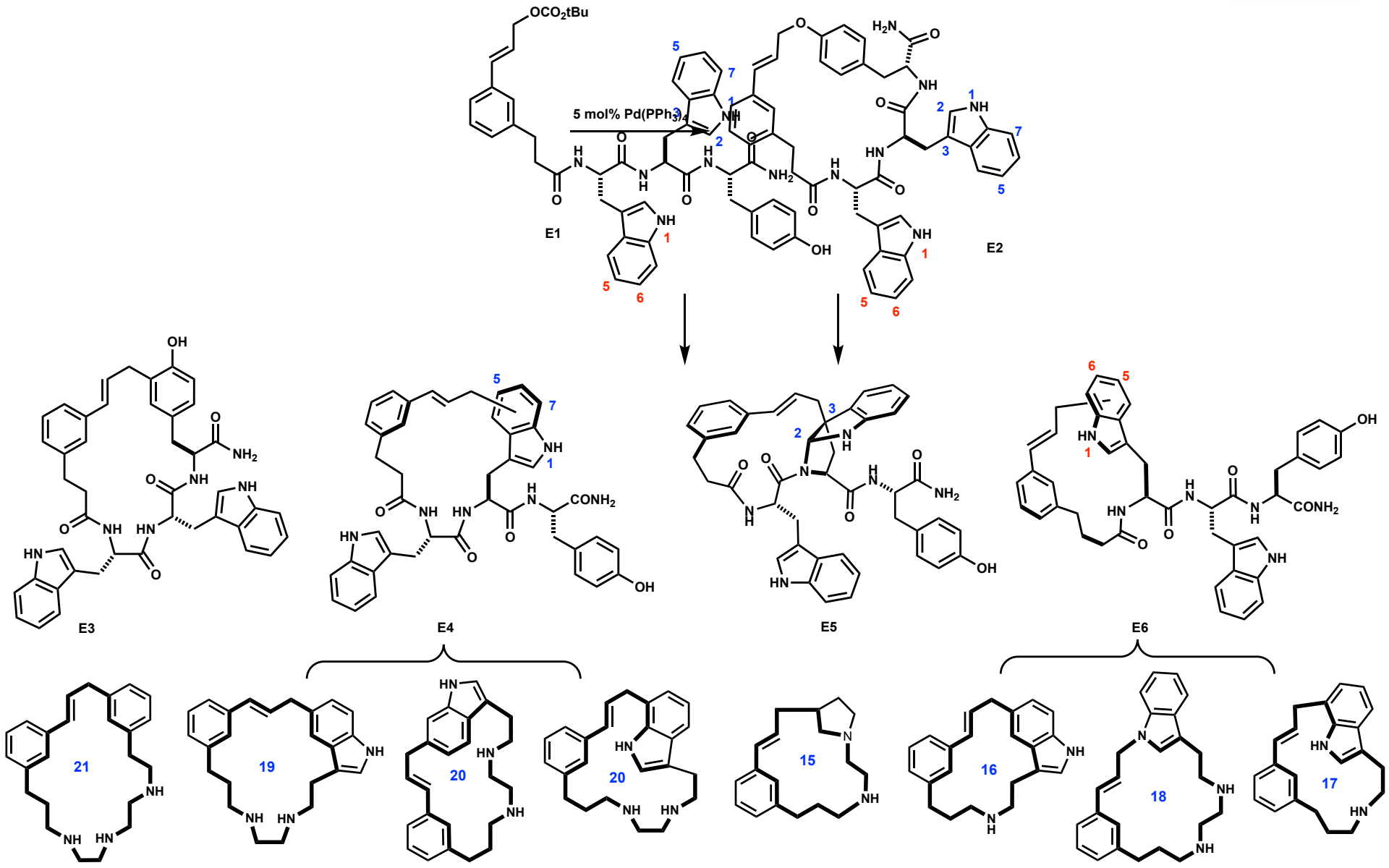
7 **8**



C-C bonded

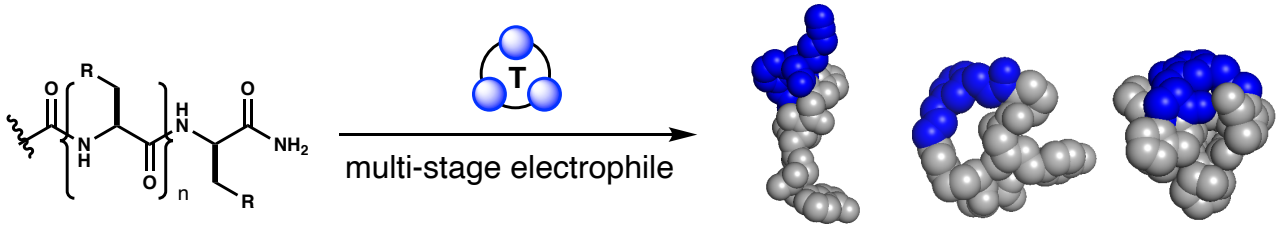
9

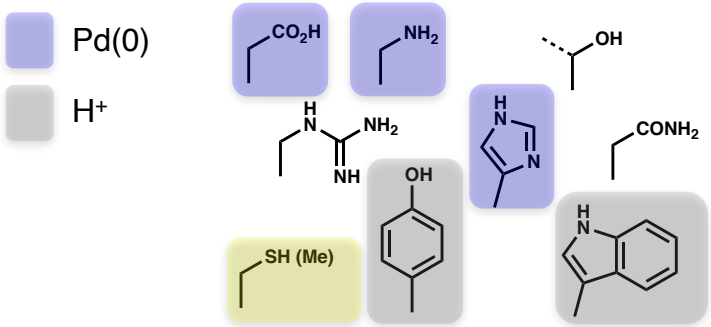
Representative Outcomes

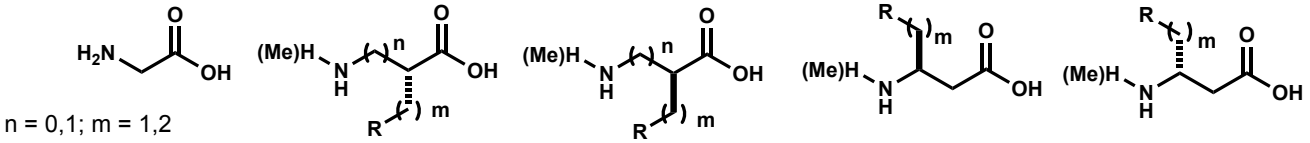


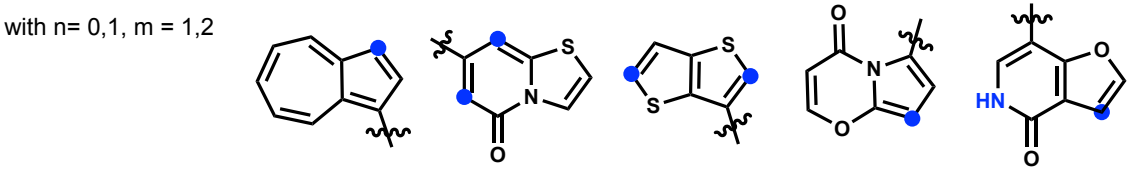
Tetrahedron (2013) 69, 7683-7691

Established Scope Anticipates a Broad Platform









CPMG – Composite Peptide Macrocycle Generator

UCLA Chemistry & Biochemistry

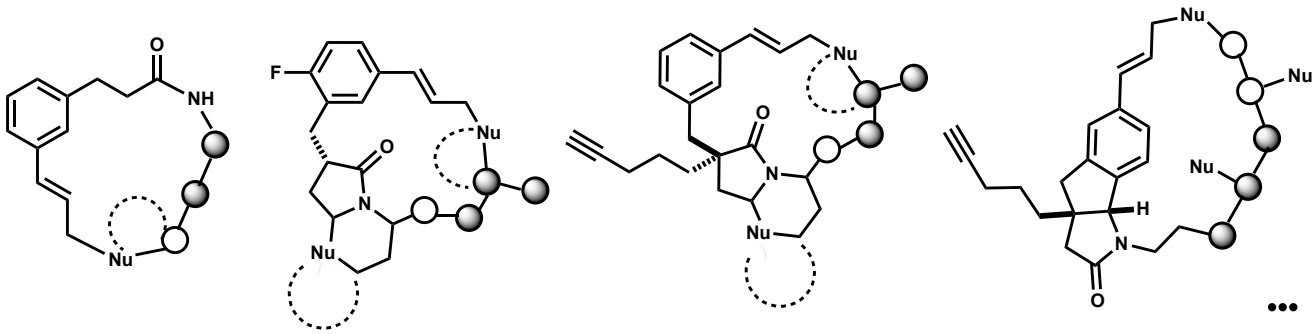
Computational generation of an annotated gigalibrary of synthesizable, composite peptidic macrocycles

Ishika Saha[a,1], Eric K. Dang[b,1], Dennis Svatunek[a], Kendall N. Houk[a,2], and Patrick G. Harran[a,2]

[a]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095; and [b]Department of Computer Science, University of California, Los Angeles, CA 90095

PNAS | October 6, 2020 | vol. 117 | no. 40 | 24679–24690

Heterocycles | Monomers | Oligomers | Reactions

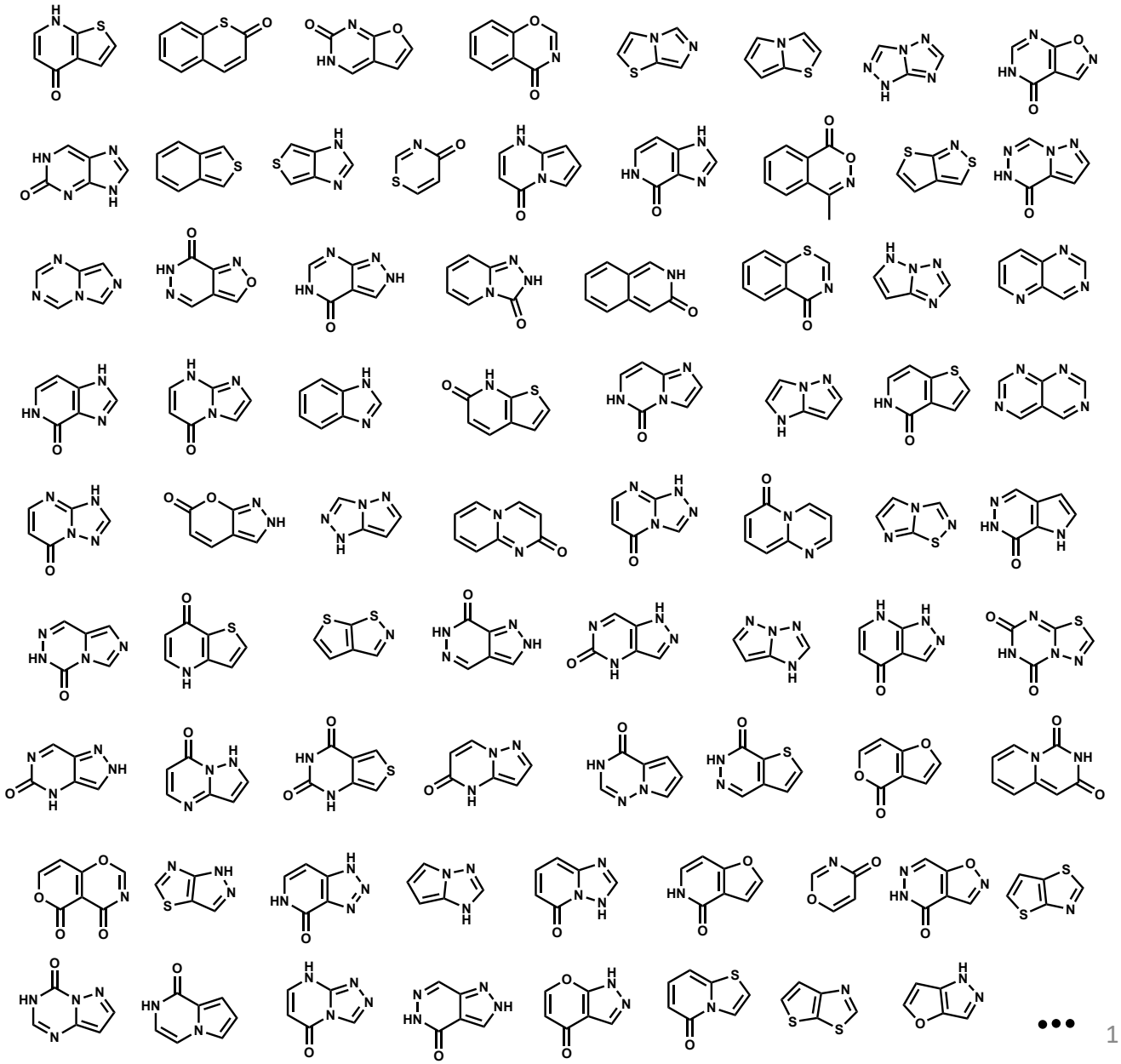


2×10^9 macrocycles

Identifying Heterocycles for CPMG

UCLA Chemistry & Biochemistry

VEHICLE database



J. Med. Chem. (2009) 52, 9, 2952–2963

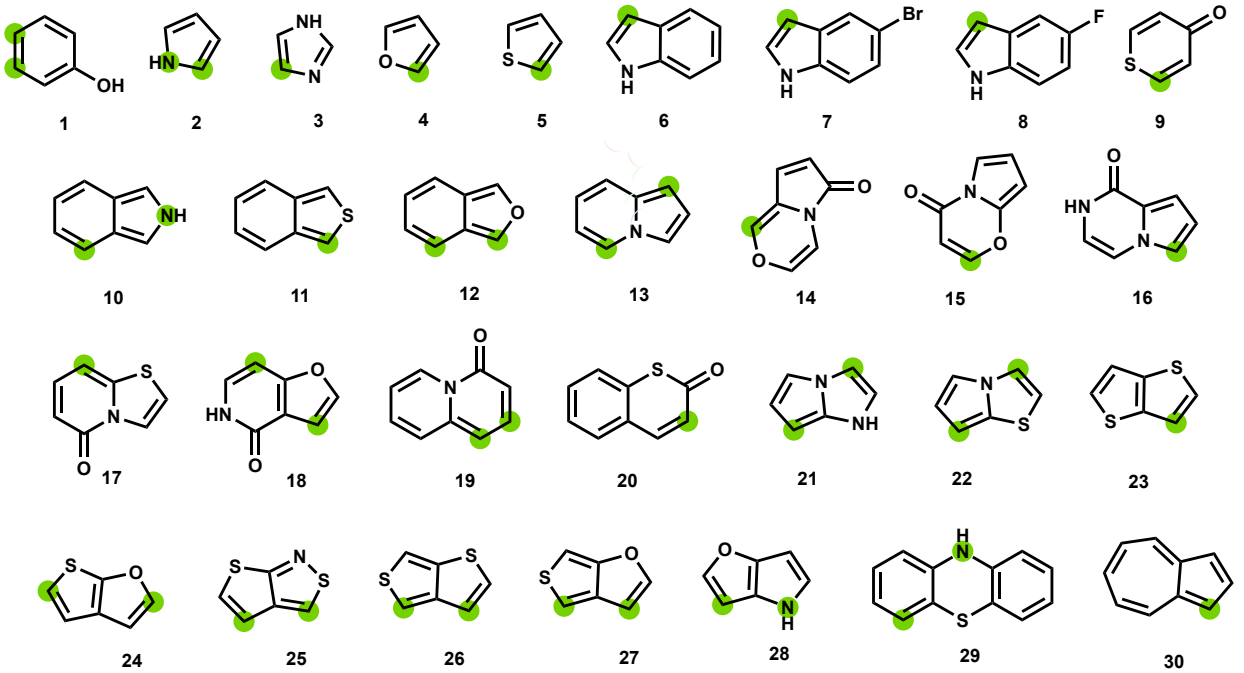
Identifying Heterocycles for CPMG

UCLA Chemistry & Biochemistry

VEHICLE database

RegioSQM

Jaguar



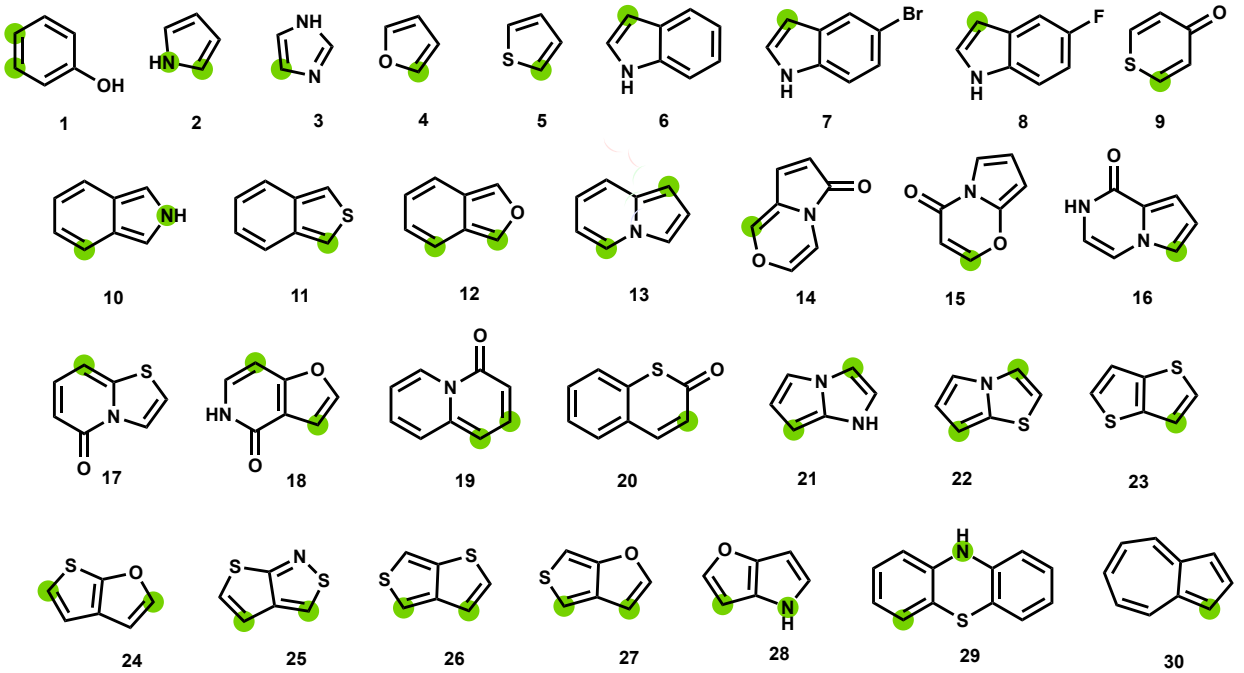
Chem. Sci., (2018) 9, 660-665
Int. J. Quantum Chem. (2013) 113, 2110– 2142
J. Med. Chem. (2009) 52, 9, 2952–2963

Identifying Heterocycles for CPMG

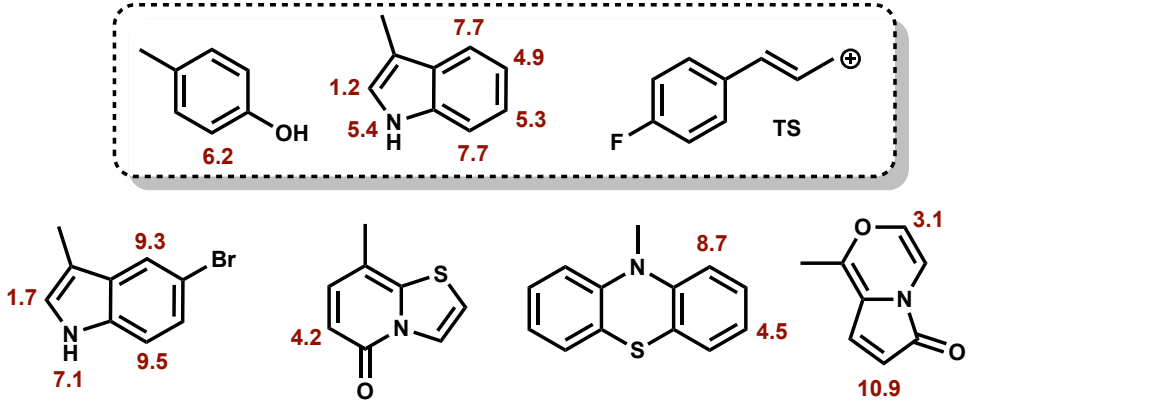
UCLA Chemistry & Biochemistry

VEHICLE database

RegioSQM
Jaguar



(ωB97X-D-SMD(methanol)/6–31G(d)) in Gaussian16 RevA.03



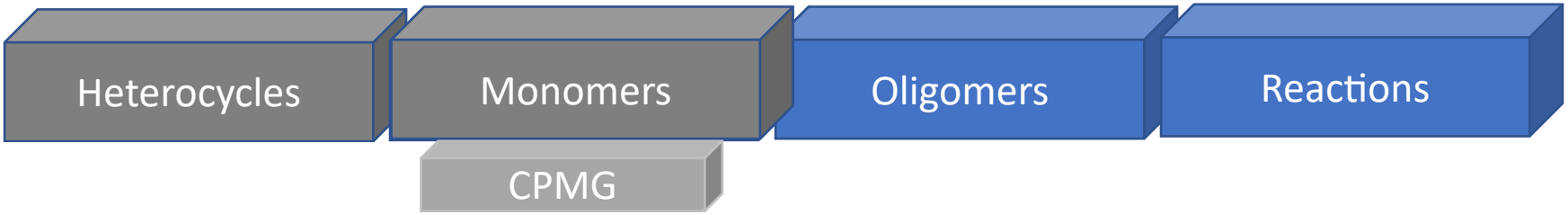
Chem. Sci., (2018) 9, 660-665
Int. J. Quantum Chem. (2013) 113, 2110– 2142
Phys. Chem. Chem. Phys. (2008) 10, 6615–6620

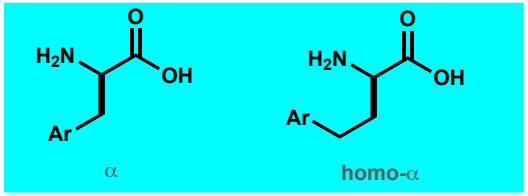
Computational generation of an annotated gigalibrary of synthesizable, composite peptidic macrocycles

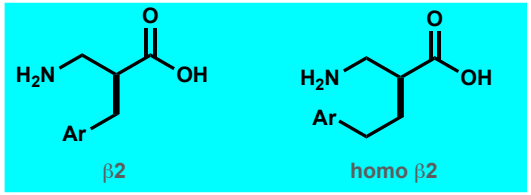
Ishika Saha[a,1], Eric K. Dang[b,1], Dennis Svatunek[a], Kendall N. Houk[a,2], and Patrick G. Harran[a,2]

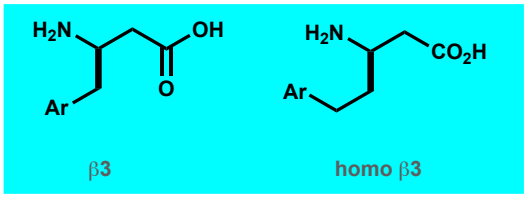
[a]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095; and [b]Department of Computer Science, University of California, Los Angeles, CA 90095

PNAS | October 6, 2020 | vol. 117 | no. 40 | 24679–24690







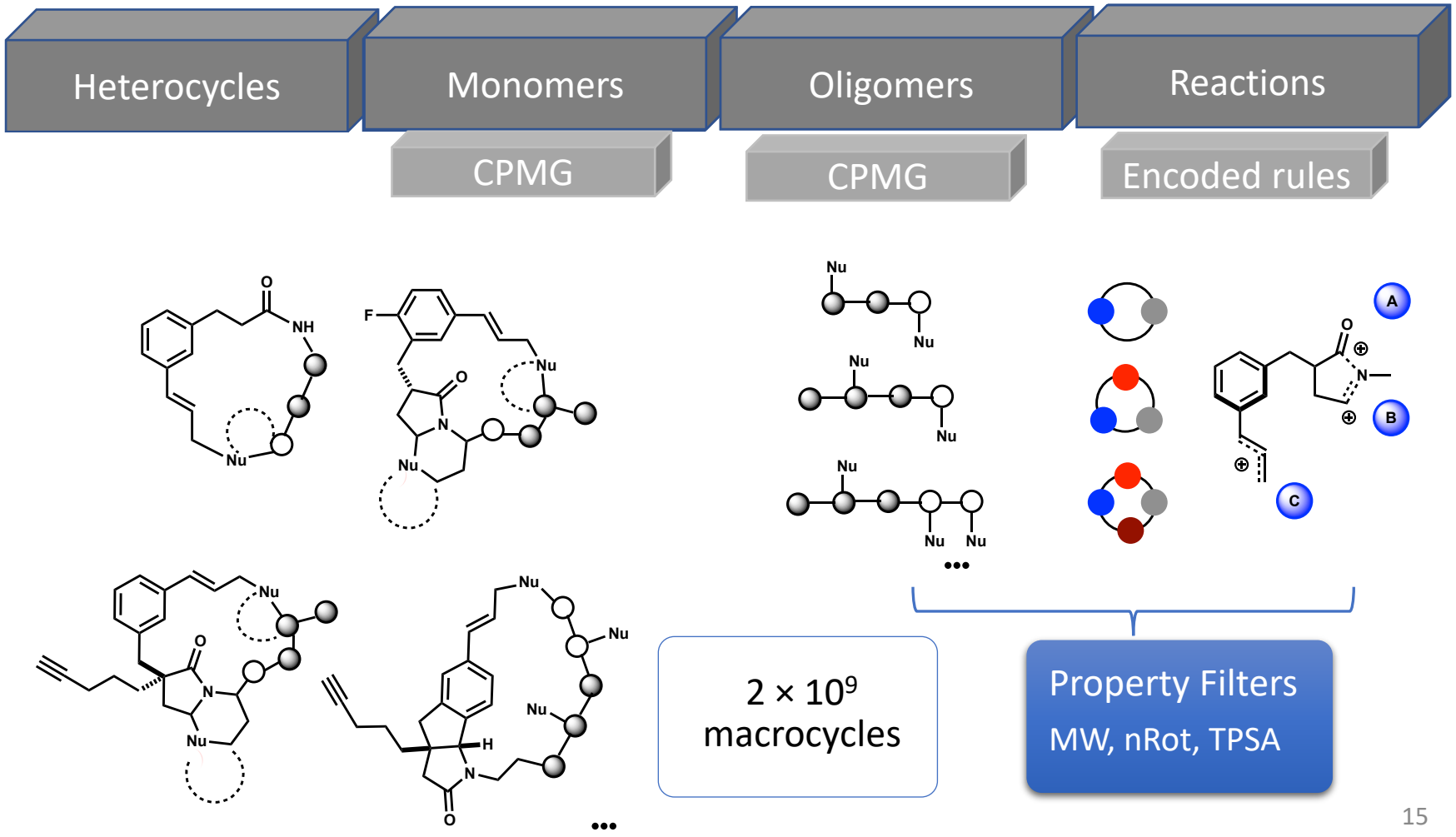


Computational generation of an annotated gigalibrary of synthesizable, composite peptidic macrocycles

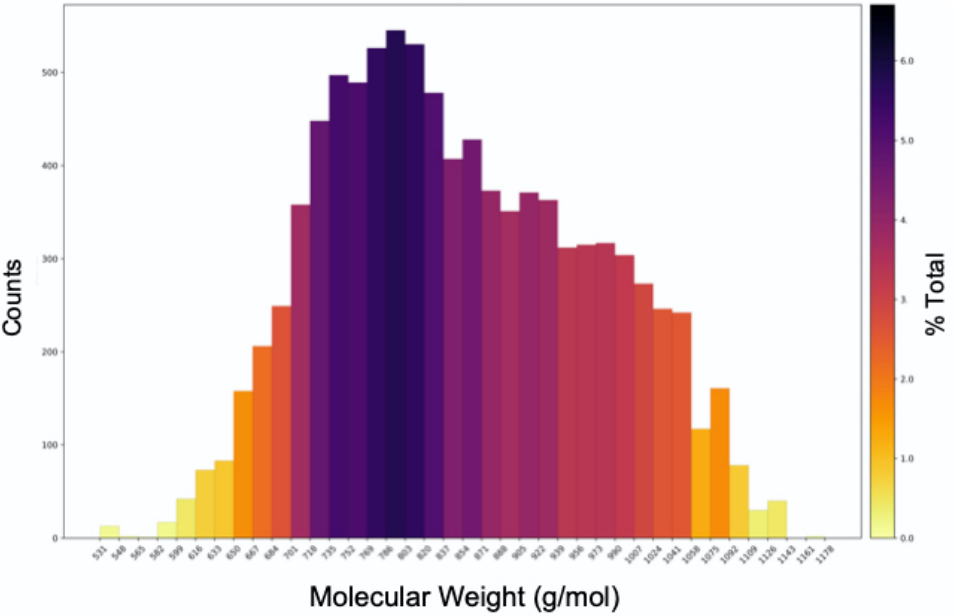
Ishika Saha[a,1], Eric K. Dang[b,1], Dennis Svatunek[a], Kendall N. Houk[a,2], and Patrick G. Harran[a,2]

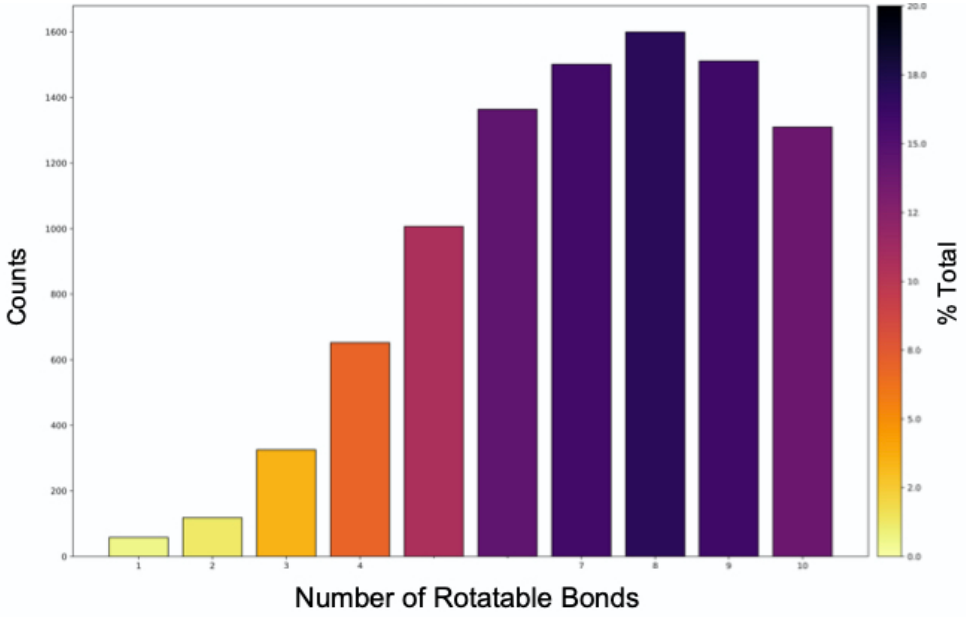
[a]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095; and [b]Department of Computer Science, University of California, Los Angeles, CA 90095

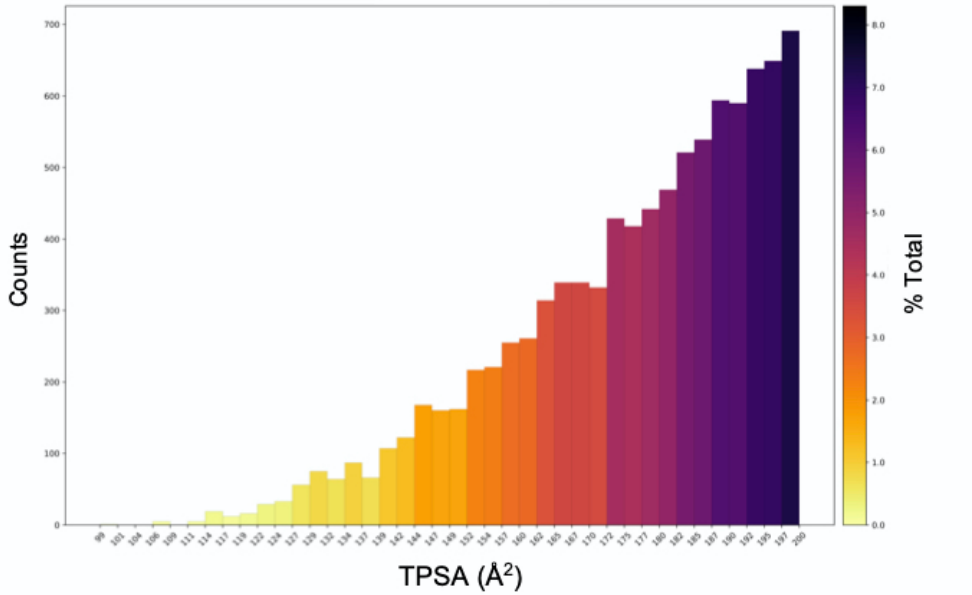
PNAS | October 6, 2020 | vol. 117 | no. 40 | 24679–24690



Convex Hull Sampling of 10^4 Structures



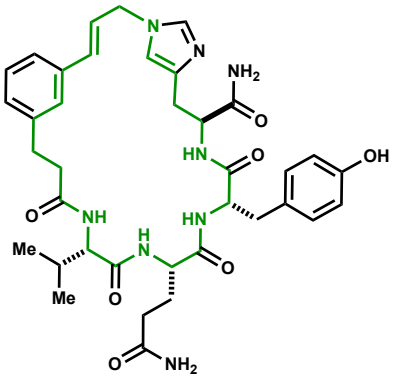




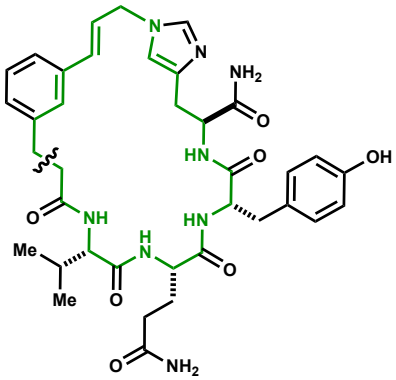
Conformational Analysis – ConfBuster++

UCLA Chemistry & Biochemistry

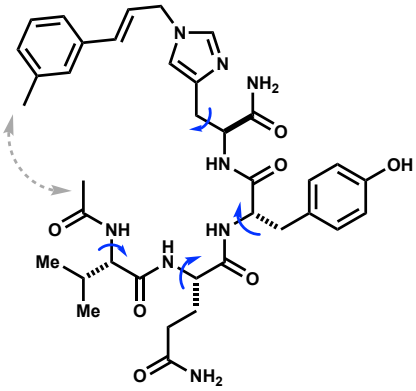
1. Embed molecule, identify backbone and 'cleavable' bonds



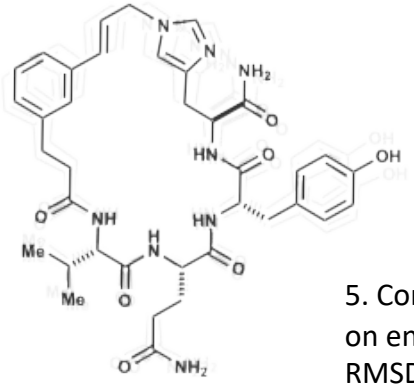
2. Cleave bonds to generate *n* conformers per linearized form



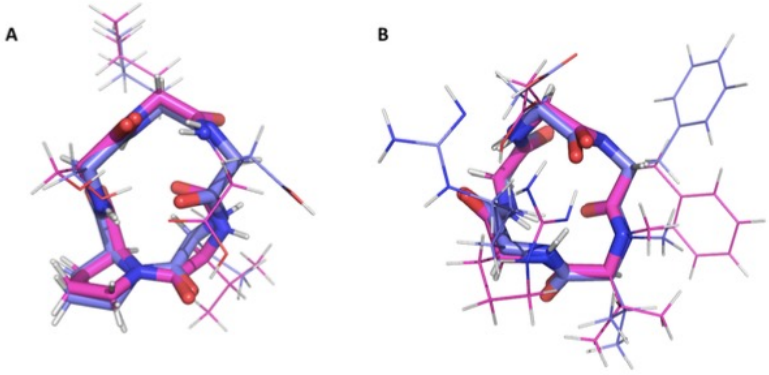
3. Rotate dihedrals to generate *N* new conformers



4. Genetic algorithm to retain *M* conformers



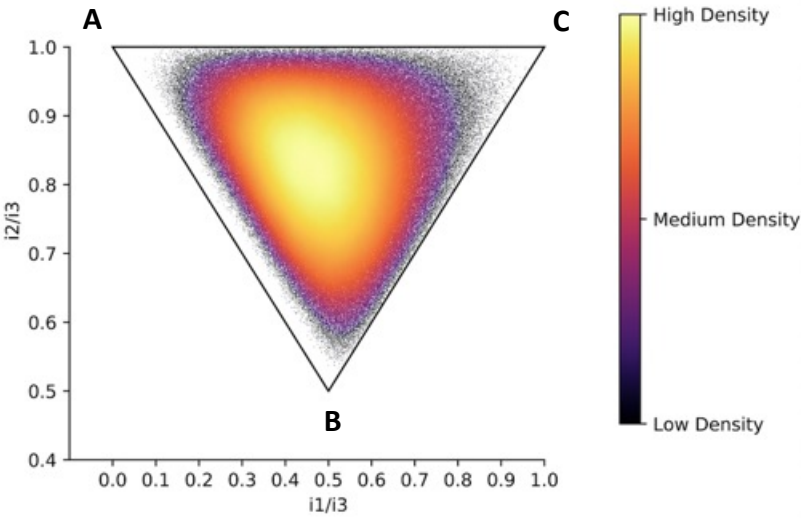
5. Conformers filtered on energetic and RMSD thresholds

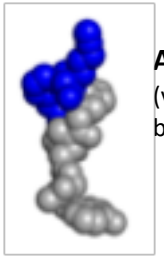


Shape Diversity of CPMG Library

UCLA Chemistry & Biochemistry

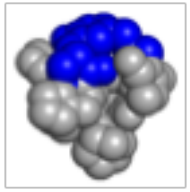
Principal Moment of Inertia Plot



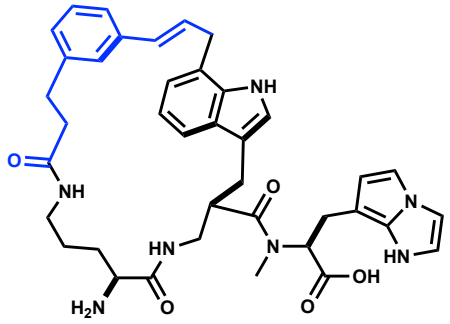


A. Rod shaped
(vertex A exemplified by diacetylene)

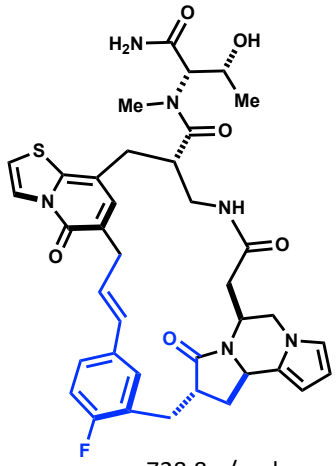
B. Disk shaped
(vertex B exemplified by benzene)



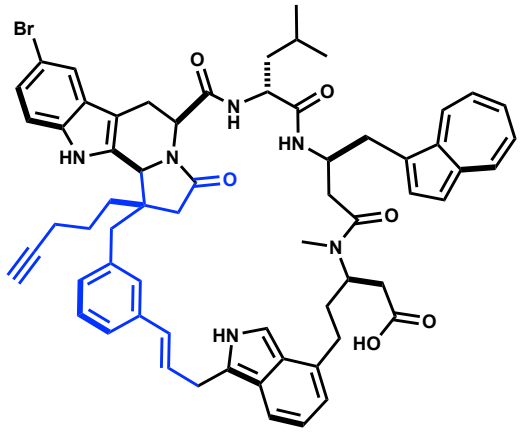
C. Sphere shaped
(vertex C exemplified by adamantane)



691.8 g/mol

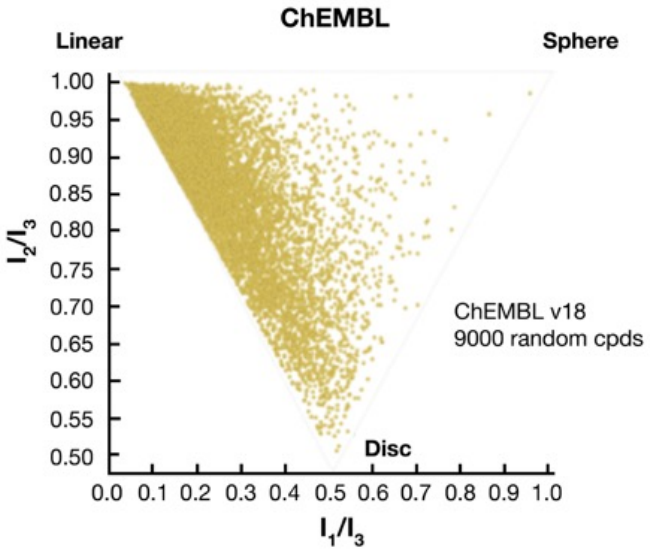


728.8 g/mol

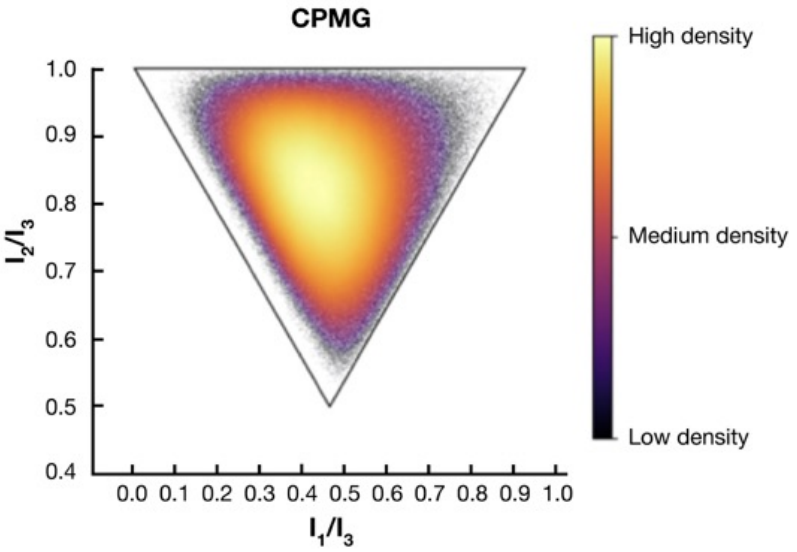


1096.2 g/mol

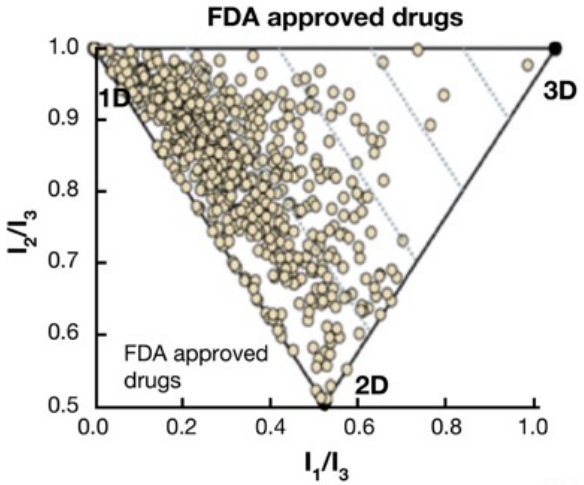
Shape Diversity of CPMG Library

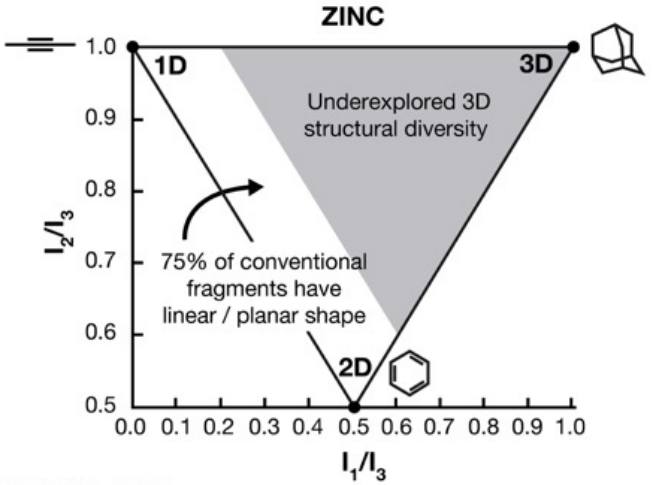


J. Med. Chem. 2016, 59, 10, 4443-4458



PNAS, 2020, 117 (40), 24679-24690.

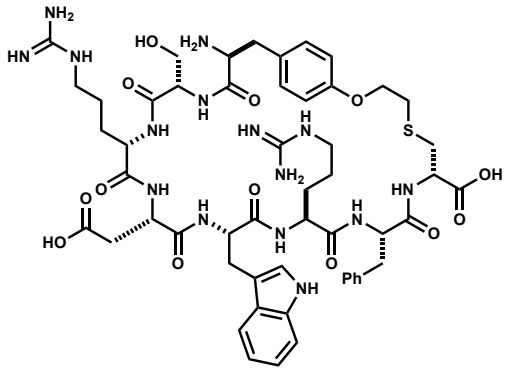
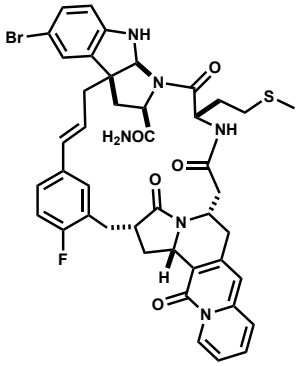
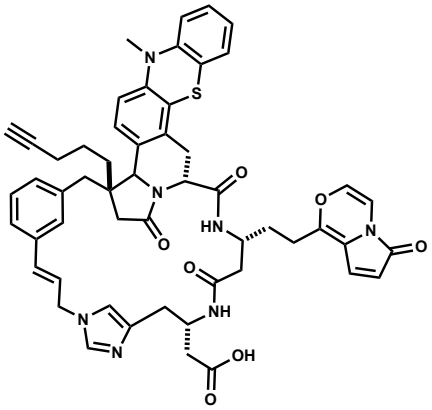


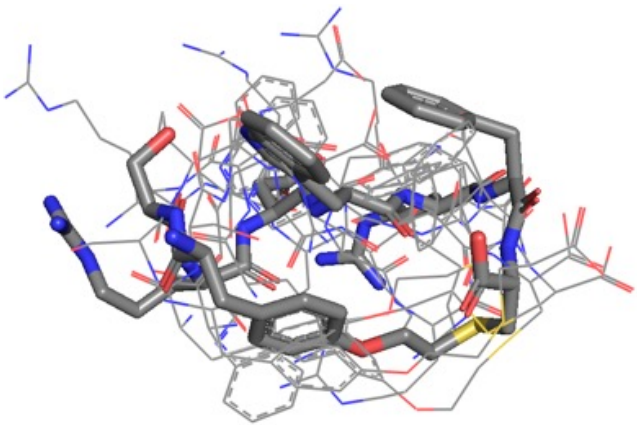


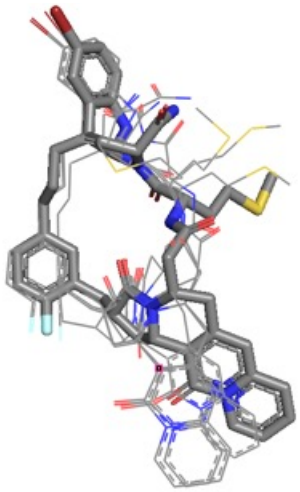
Chem. Sci., 2020,11 1216-1225

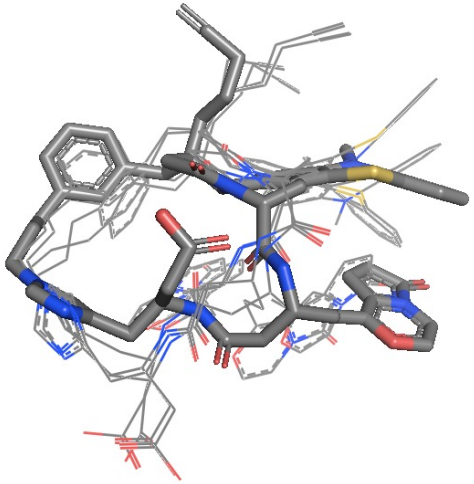
Advantages versus State of the Art

UCLA Chemistry & Biochemistry

	Phage Display Platform	CPMG	CPMG
			
	Single constraint / native backbone	*Multiple constraints / abiotic functionality / variable stereochemistry*	
MW (g/mol)	1157.5	853.8	944.1
Rotatable bonds	16	4	7
TPSA (Å²)	498.7	171.6	197.3
Average ring RMSD of 5 LECS (Å)	2.76	1.10	1.24



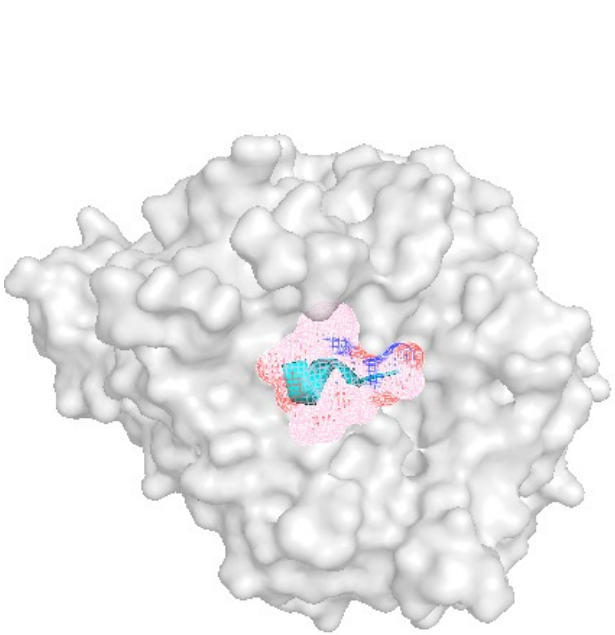




ACS Cent Sci. (2020) 6(3), 368-381

Computationally Guided Application

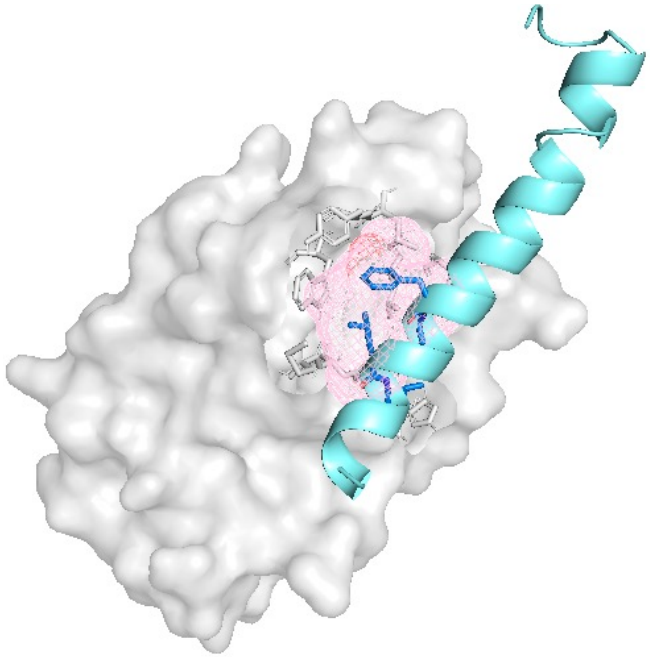
UCLA Chemistry & Biochemistry



Clathrin protein

Nat Struct Mol Biol (2004) 11, 242-248

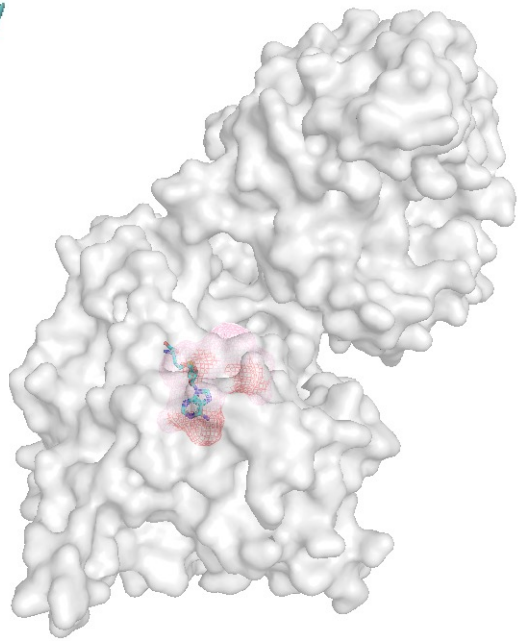
known consensus binding sequence
Thr-Leu-Pro-Trp-Asp-Leu-Trp-Thr-Thr



GIPr

Proc. Natl. Acad. Sci. USA (2007) 104 (35) 13942-13947

42 residue binding sequence



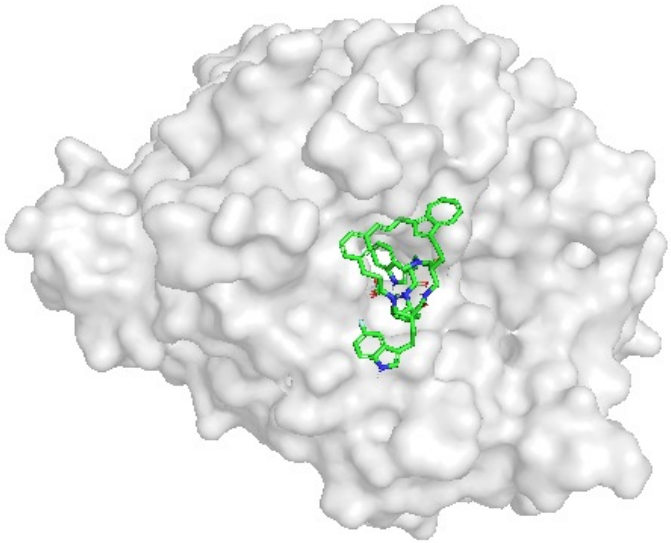
MTHFR

Biochimie (2021) 183, 100-107

S-adenosylhomocysteine analogue binders

Computationally Guided Application

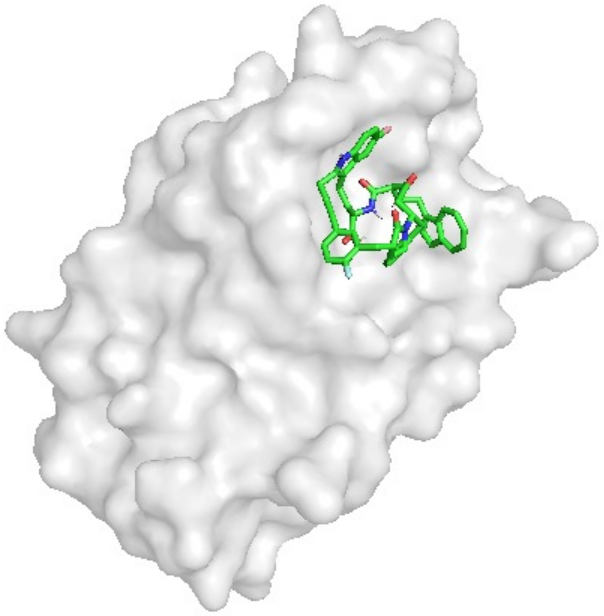
UCLA Chemistry & Biochemistry



Clathrin protein

Nat Struct Mol Biol (2004) 11, 242-248

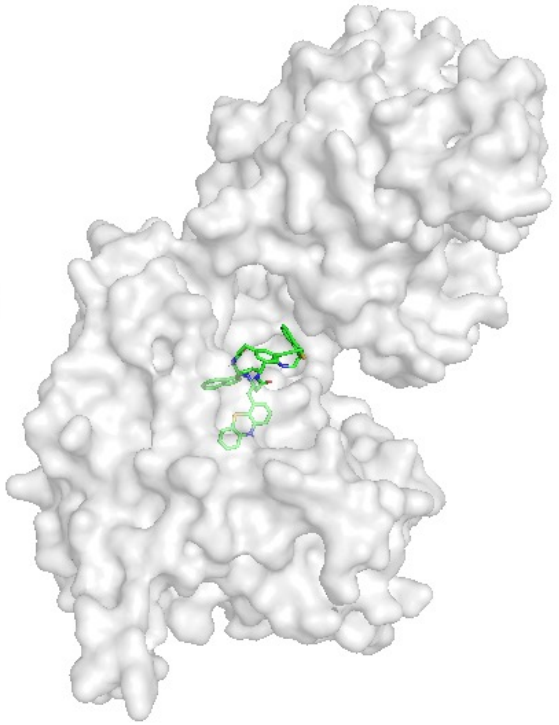
known consensus binding sequence
Thr-Leu-Pro-Trp-Asp-Leu-Trp-Thr-Thr



GIPr

Proc. Natl. Acad. Sci. USA (2007) 104 (35) 13942-13947

42 residue binding sequence



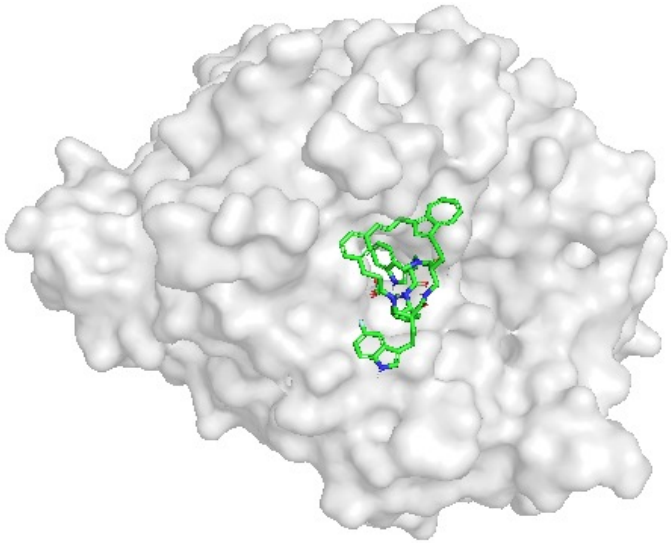
MTHFR

Biochimie (2021) 183, 100-107

S-adenosylhomocysteine analogue binders

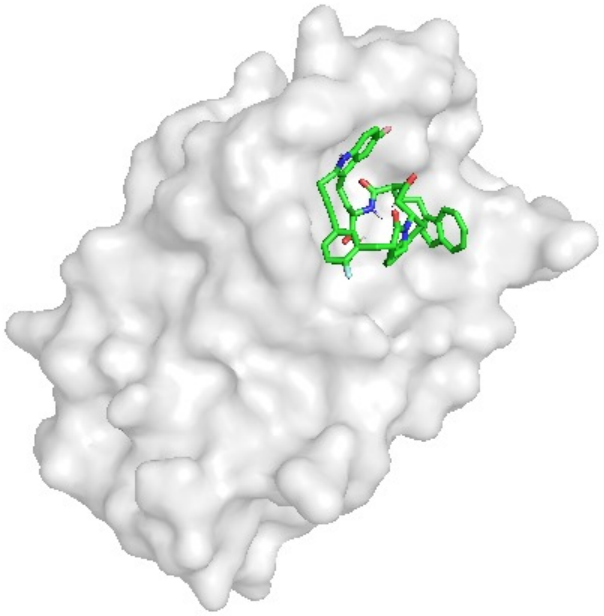
Computationally Guided Application

UCLA Chemistry & Biochemistry



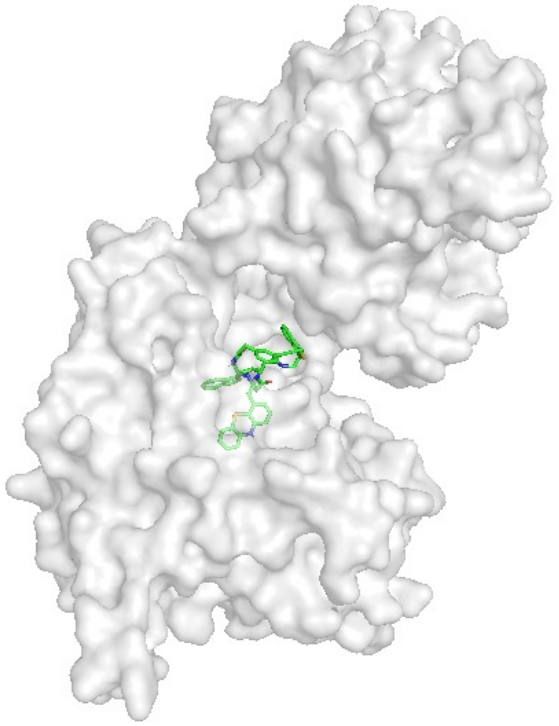
Clathrin protein

Nat Struct Mol Biol (2004) 11, 242-248



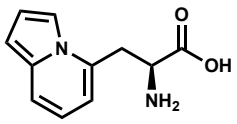
GIPr

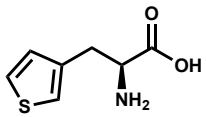
Proc. Natl. Acad. Sci. USA (2007) 104 (35) 13942-13947

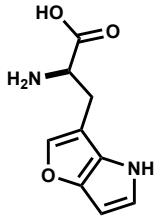


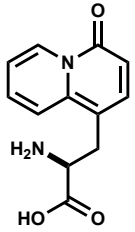
MTHFR

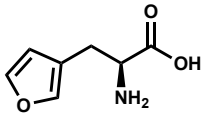
Biochimie (2021) 183, 100-107

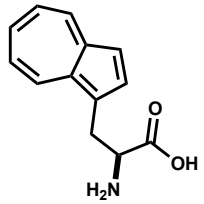






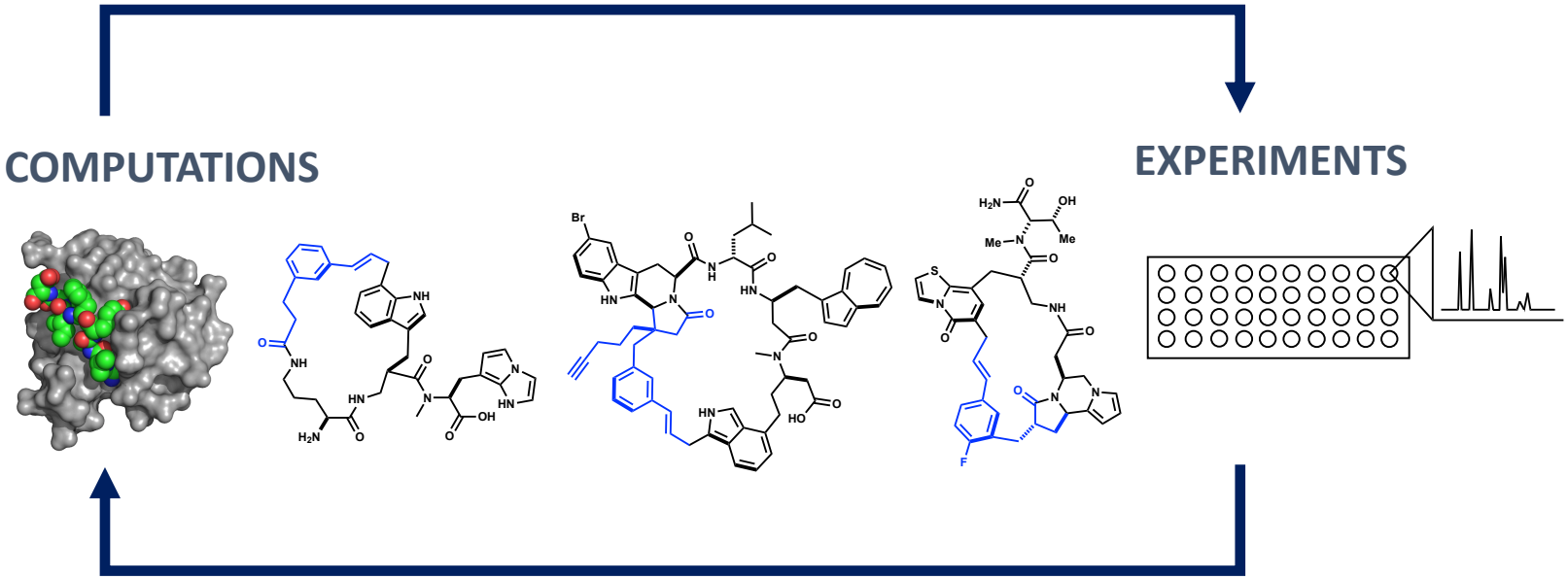






Efforts in Collaboration with Industry

UCLA Chemistry & Biochemistry



CONFORMATIONAL ANALYSES

PROTEIN MOLECULAR DYNAMICS

HIGH SPEED DOCKING

MONOMER ACQUISITION

OLIGOMER ASSEMBLY AND PROCESSING

HIGH THROUGHPUT PURIFICATION