Automated Labeling of *cis/trans* Isomers

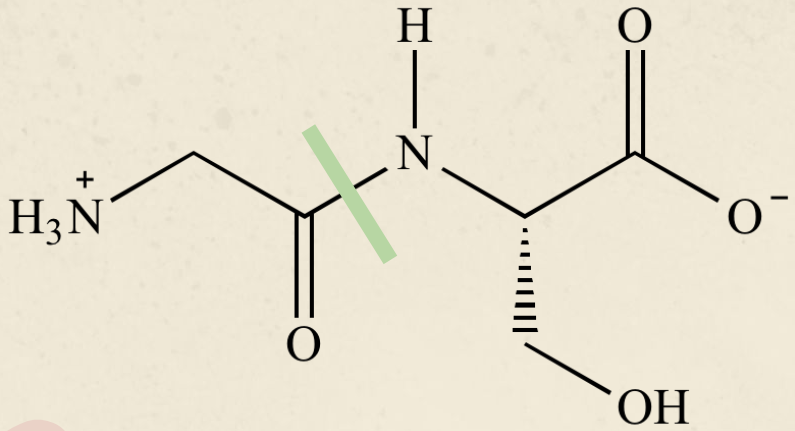
AKA: classifying ProXxx dipeptide conformers that were theoretically generated by the Poutsma Group

A Presentation By: Miko Miwa

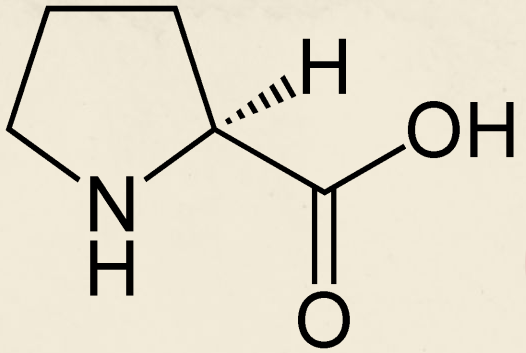
Introduction

What are dipeptides?

- Chain of two (2) amino acids
- In the Poutsma Group, we are specifically interested in dipeptides containing the amino acid proline



Gly Ser



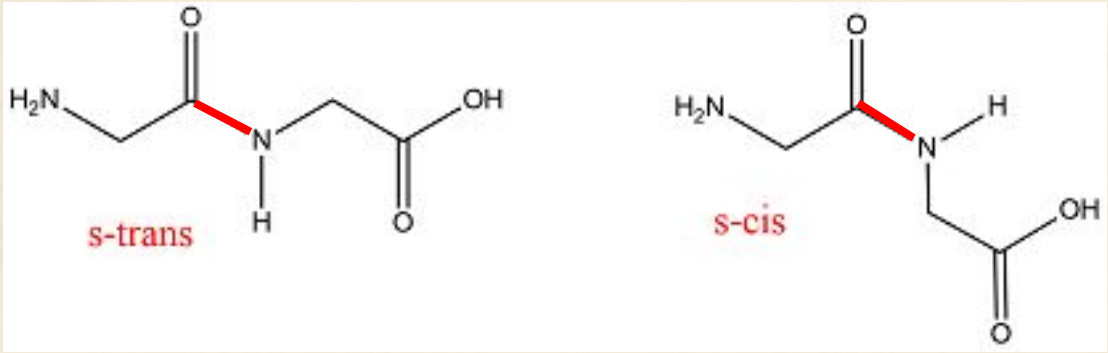
Pro

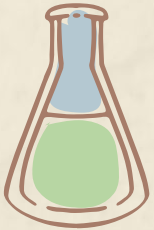
What are conformational isomers?

- Molecules consisting of the same atoms and connections, distinguished only by the rotation of one or more single bonds

What is *cis-trans* isomerism?

- When isomers differ based on the placement of functional groups with respect to each other on a single plane
    - *cis* - "this side of" ; *trans* - "other side of"

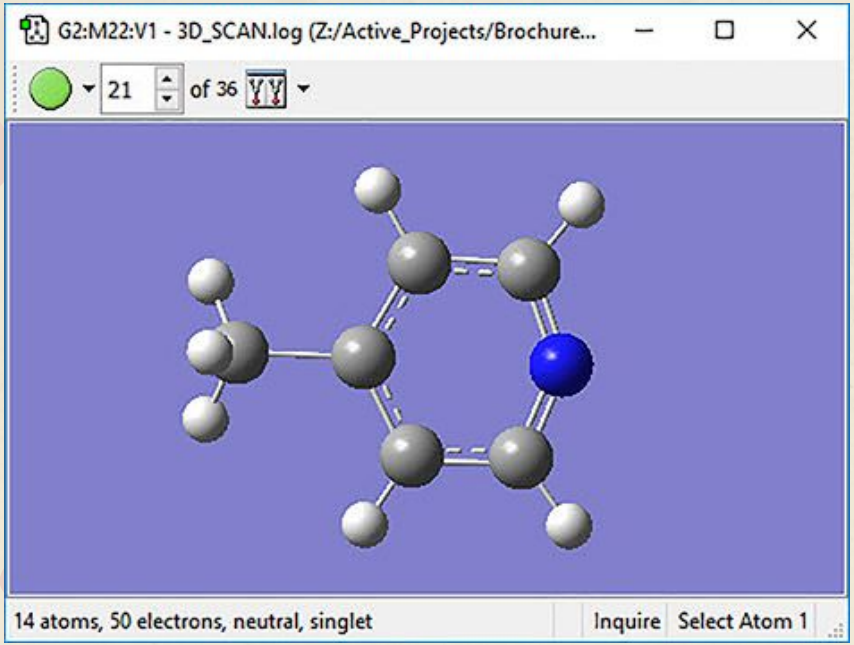


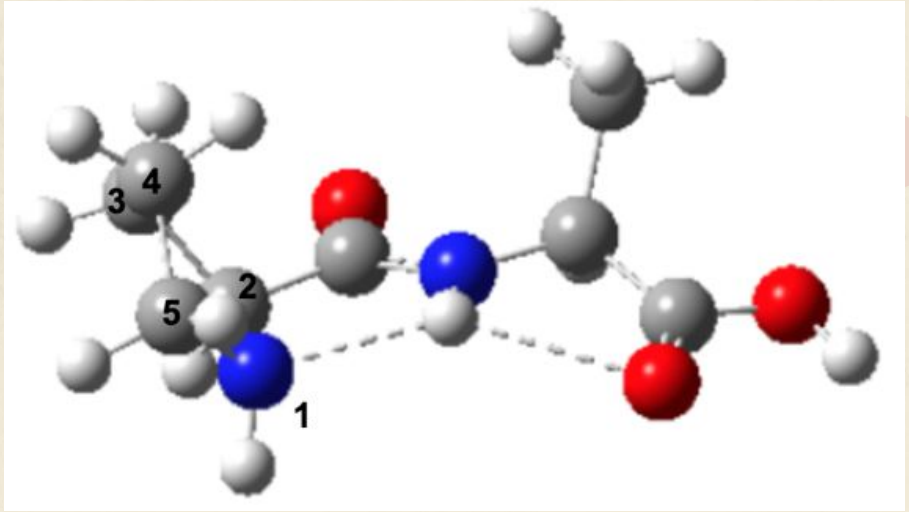


The Goals

Short-term:

1. Create a system for reading and processing raw Gaussian output files into something that is both usable and shareable
2. Create an algorithm that can correctly label ProXxx conformers as either *cis* or *trans*


G2:M22:V1 - 3D_SCAN.log (Z:/Active_Projects/Brochure...
21
of 36
14 atoms, 50 electrons, neutral, singlet
Inquire
Select Atom 1


1
2
3
4
5

The Process - Hidden and Visible

My project has two components - the hidden and the visible

Hidden stuff is on my personal device and includes the raw data generated by my lab (this information should remain private)	Visible stuff is on github and cleaned up to such a point that it is ok to be shared publicly