Evaluation of several correlated electronic methods in the context of calculations common to drug design

Daniel L. Cheney

Computer Aided Drug Design
Bristol-Myers Squibb, P.O.Box 5400, Princeton, NJ 08543-5400

Q-Chem Workshop
March 2009

* *E-Mail cheneyd@bms.com*

INTRODUCTION

As the cost of computational infrastructure continues to decrease, quantum chemical methods are becoming more commonplace in computer-aided drug design

Roles of quantum chemistry in drug design at Bristol Myers – Squibb.

- Global conformational analysis (Jaguar, Q-Chem, AMPAC)
- Torsional scans (Jaguar, Q-Chem)
- Intermolecular interactions.
  - H-bond calculations (Jaguar, MolPro on 256 GB RAM servers)
  - Other polar and non-polar interactions (MolPro; DFT:M06-2X)
- PKa calculations (Jaguar)
- Electrostatic potentials (Spartan; Jaguar)
- Transition states and other specialized calculations (Gaussian)
- MM / QC (*QSite*, Schrodinger, Inc; Q-Chem / Charmm)

Semi-automated protocol for estimating hydrogen bonding strengths.

Calculation of binding energy for the complex A:B

Semi-automated

A

(1) X3LYP/631G** geometry opt

(2) LMP2/cc-pV**T**Z(-**f**) energy

(3) LMP2/cc-pV**Q**Z(-g) energy

B

A:B

(1) X3LYP/631G** geometry opt

(2) LMP2/cc-pV**T**Z(-**f**) energy

(3) LMP2/cc-pV**Q**Z(-g) energy

(4) LMP2/cc-pV**T**Z(-f) CP for **A**:B

(5) LMP2/cc- pV**Q**Z(-g) CP for **A**:B

(6) LMP2/cc-pV**T**Z(-f) CP for A**:B**

(7) LMP2/cc- pV**Q**Z(-g) CP for A**:B**

- *Optimize geometry*
- *Counterpoise corrections*
- *CBS extrapolation*
- *Apply CCSD(T) correction based on training set (13 pairs)*

Summary of comparison with previously published data:

24 strong, moderate and weakly bound systems spanning a broad range of chemical structures relevant to drug discovery for which CCSD(T) reference data is available.

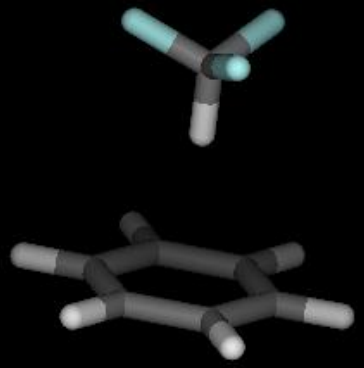
Mean absolute deviation:	0.77 kcal
Max absolute deviation:	1.6 kcal
Average percent error:	6.5 %
Max percent error:	21 %

Largest relative error:

$H\text{-}CF_3$ / Benzene

-3.3 (-4.2 kcal*)

Largest percent error (21%).



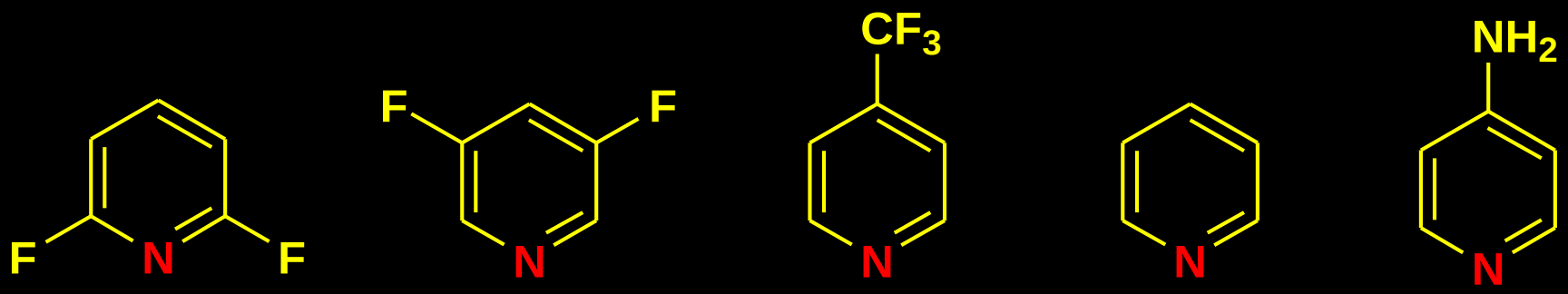
Hydrogen bond calculator: example

Pyridine analogues bound to water:

Strong inductive effect

Weak inductive effect

Strong resonance effect augments H bond strength



- 4.9 kcal

- 6.2 kcal

- 6.0 kcal

- 6.9 kcal

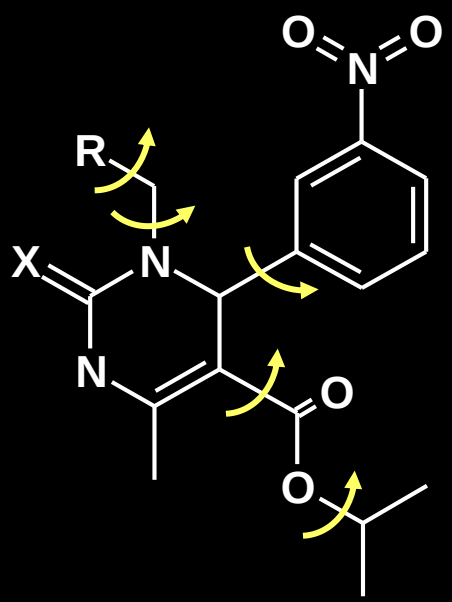
- 9.2 kcal

weaker

stronger

Example of a successful application of QC in T-Type Ca $^{2+}$ channel blockers

Generate hypothesis for bioactive conformation using crystallographic and quantum mechanical conformational analysis.



Monte-Carlo methods* to sample unusual ring systems

- **Minimize with OPLS-AA**
- **SGB solvent model**
- **< 5 kcal cutoff**

Remove redundant conformations

Minimize: DFT/HF/ PB **

Remove redundant conformations

(L)MP2/ cc-PVTZ(-f)++/ PB**

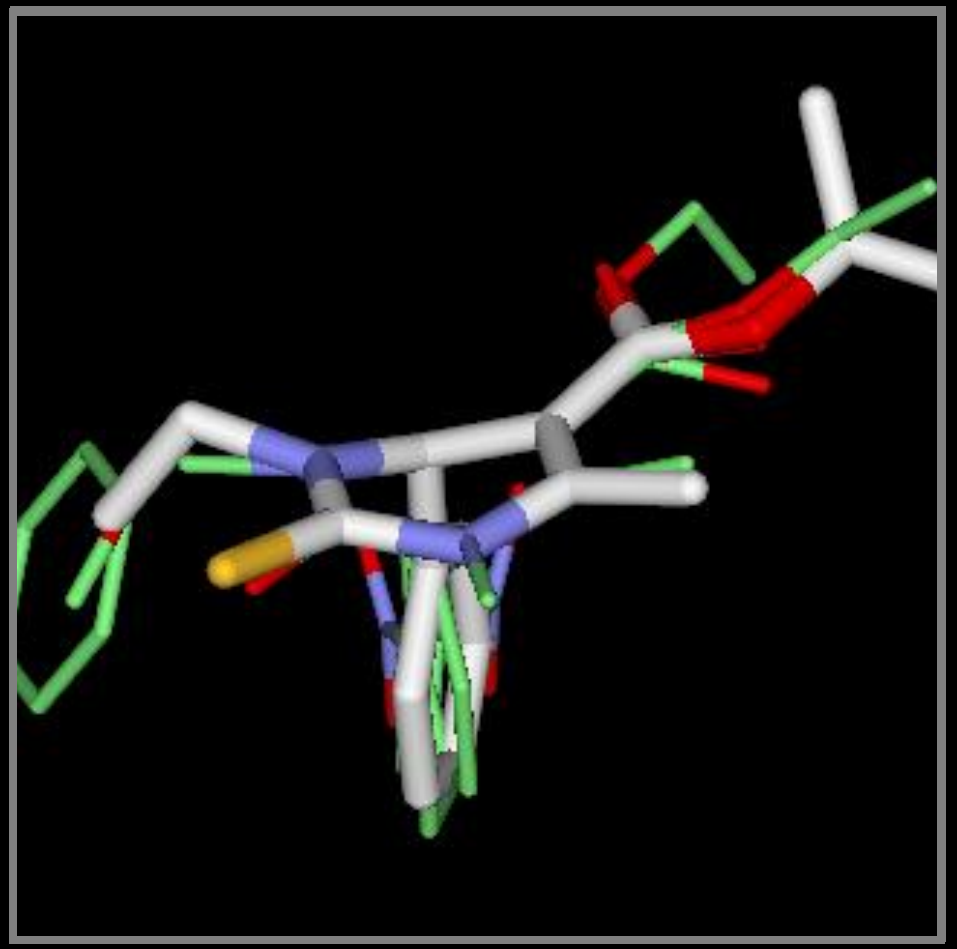
Hangeland, Jon J.; Cheney, Daniel L, et al Bioorganic & Medicinal Chemistry Letters (2008), 18(2), 474-478.

Successful application of QC in T-Type Ca $^{2+}$ channel blockers

Quantum chemistry and crystallographic data point to one preferred conformation.

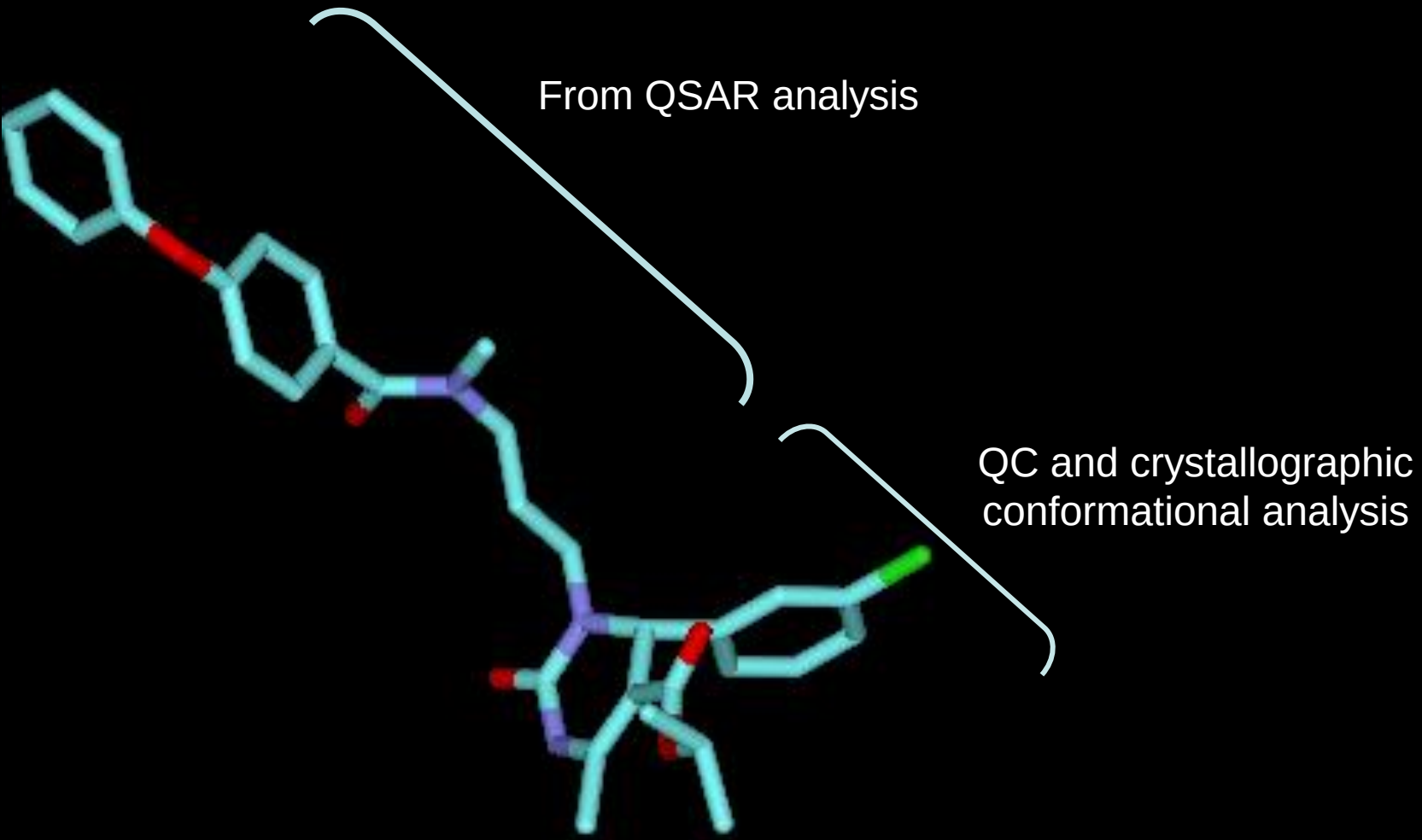
Preferred QC structure in white

Prevalent crystal conformation in green



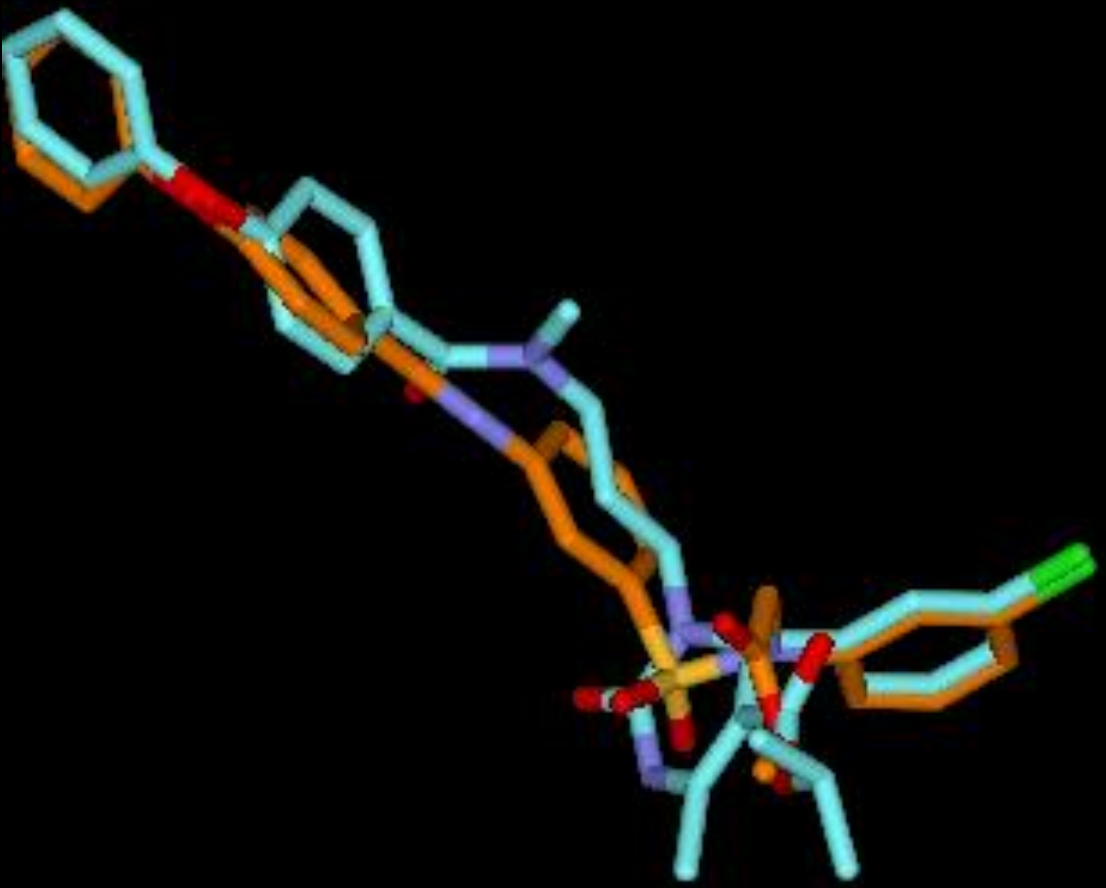
Hangeland, Jon J.; Cheney, Daniel L, et al Bioorganic & Medicinal Chemistry Letters (2008), 18(2), 474-478.

Successful application of QC in T-Type Ca $^{2+}$ channel blockers



Hangeland, Jon J.; Cheney, Daniel L, et al Bioorganic & Medicinal Chemistry Letters (2008), 18(2), 474-478.

Successful application of QC in T-Type Ca $^{2+}$ channel blockers



Hangeland, Jon J.; Cheney, Daniel L, et al Bioorganic & Medicinal Chemistry Letters (2008), 18(2), 474-478.

Successful application of QC in T-Type Ca $^{2+}$ channel blockers

- Ki = 70 nM
- Novel chemotype in heavily worked area
- Equipotent with lead series

Hangeland, Jon J.; Cheney, Daniel L, et al Bioorganic & Medicinal Chemistry Letters (2008), 18(2), 474-478.

METHODS

Methods which were compared:

• Canonical MP2 / cc-PVTZ (as reference)	Q-Chem v3.1
• Tri-atomics in Molecules (TRIM) / cc-PVTZ	Q-Chem v3.1
• Resolution of Identity (RI-MP2) / cc-PVTZ	Q-Chem v3.1
• local MP2 (LMP2) / pseudro-spectral cc-PVTZ	Jaguar v7.0
• DFT using X3LYP / 6-31G**	

Reference: "The X3LYP extended density functional for accurate descriptions of nonbond interactions, spin states, and thermo chemical properties." Xu, Xin and Goddard III, William A. PNAS March 2, 2004 vol. 101 no. 9 2673–2677

METHODS

Quantum Chemical Calculations

Equilibrium Geometries: All equilibrium geometries were obtained using X3LYP / 6-31G** (Jaguar v7.0).

- 24 torsional scans of drug like substructure were done over 360^{o} in 10^{o} increments. Internal coordinates were allowed to relax at each point except the varied torsion.
- 10 Tautomeric and 18 rotameric / configuration pairs were optimized at X3LYP / 6-31G** (Jaguar v7.0).

Single point energies: Q-Chem 3.1 or Jaguar 7.0 at the following levels of theory:

- TRIM-MP2 / cc-PVTZ (Q-Chem)
- RI-MP2 / cc-PVTZ (Q-Chem)
- LMP2 /ps-cc-PVTZ (Jaguar)
- X3LYP / 6-31G** (geometries)

Benchmark energies and potentials: All reference energies and torsion potentials were calculated at MP2 / cc-PVTZ (Q-Chem v3.1) using equilibrium geometries generated with X3LYP / 6-31G** (Jaguar v7.0).

METHODS

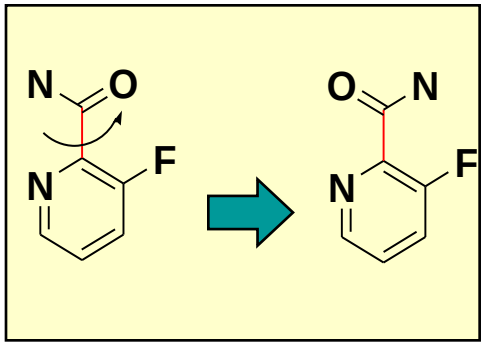
Selection of molecules:

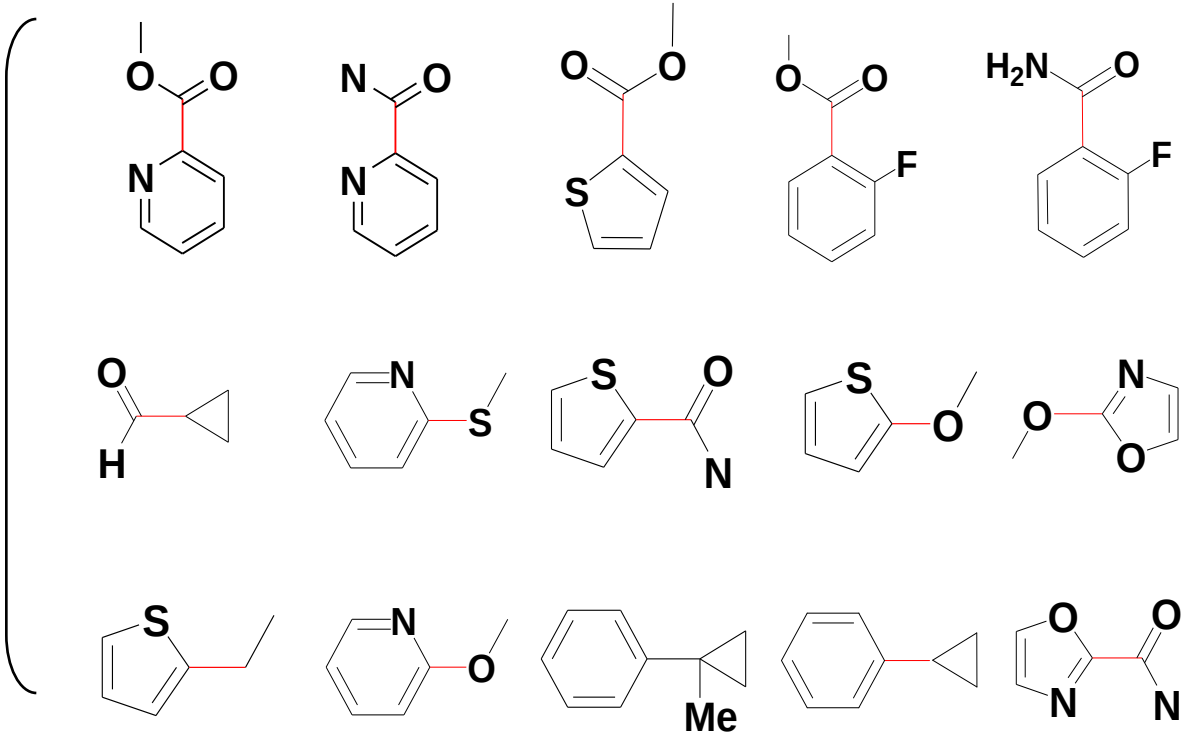
Structures were chosen which are

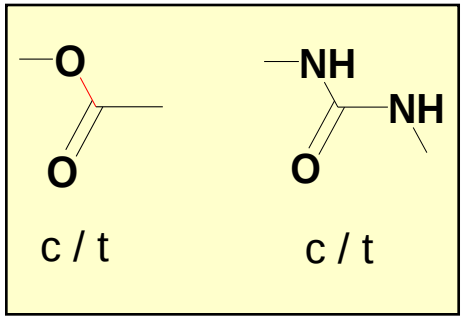
- **commonly encountered in drug design scenarios, and**
- **offer a reasonable level of chemical diversity**

A. Relative energies of 18 rotameric and configurational pairs

Data set







Range of relative energies of conformer pairs:

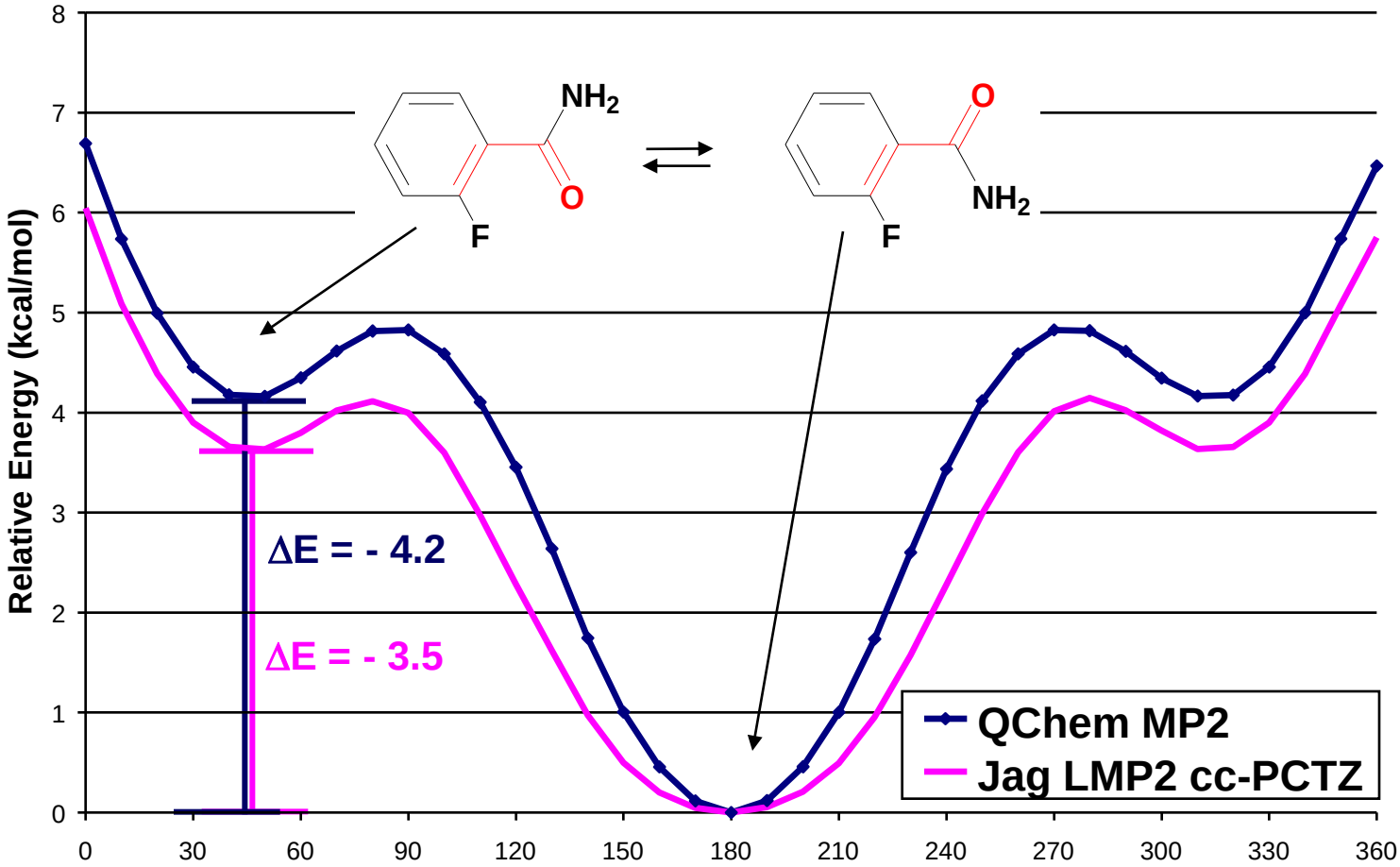
Mean absolute ΔE = 2.7 kcal

Median ΔE = 1.8

Max ΔE = 9.4

A. Relative energies of 18 rotameric and configurational pairs

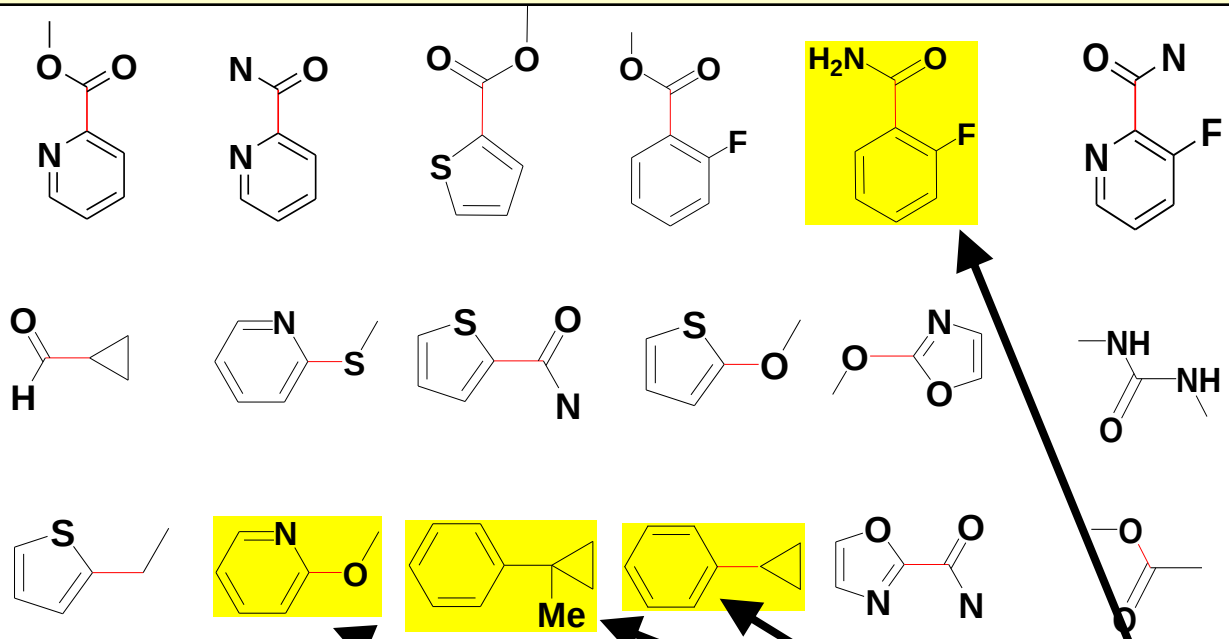
Example



	50º	180º	Abs. Deviation from MP2 reference (kcal)
MP2/ cc-PVTZ	0.0	- 4.2	
LMP2/ cc-PVTZ	0.0	- 3.5	0.7

A. Relative energies of 18 rotameric and configurational pairs

Results

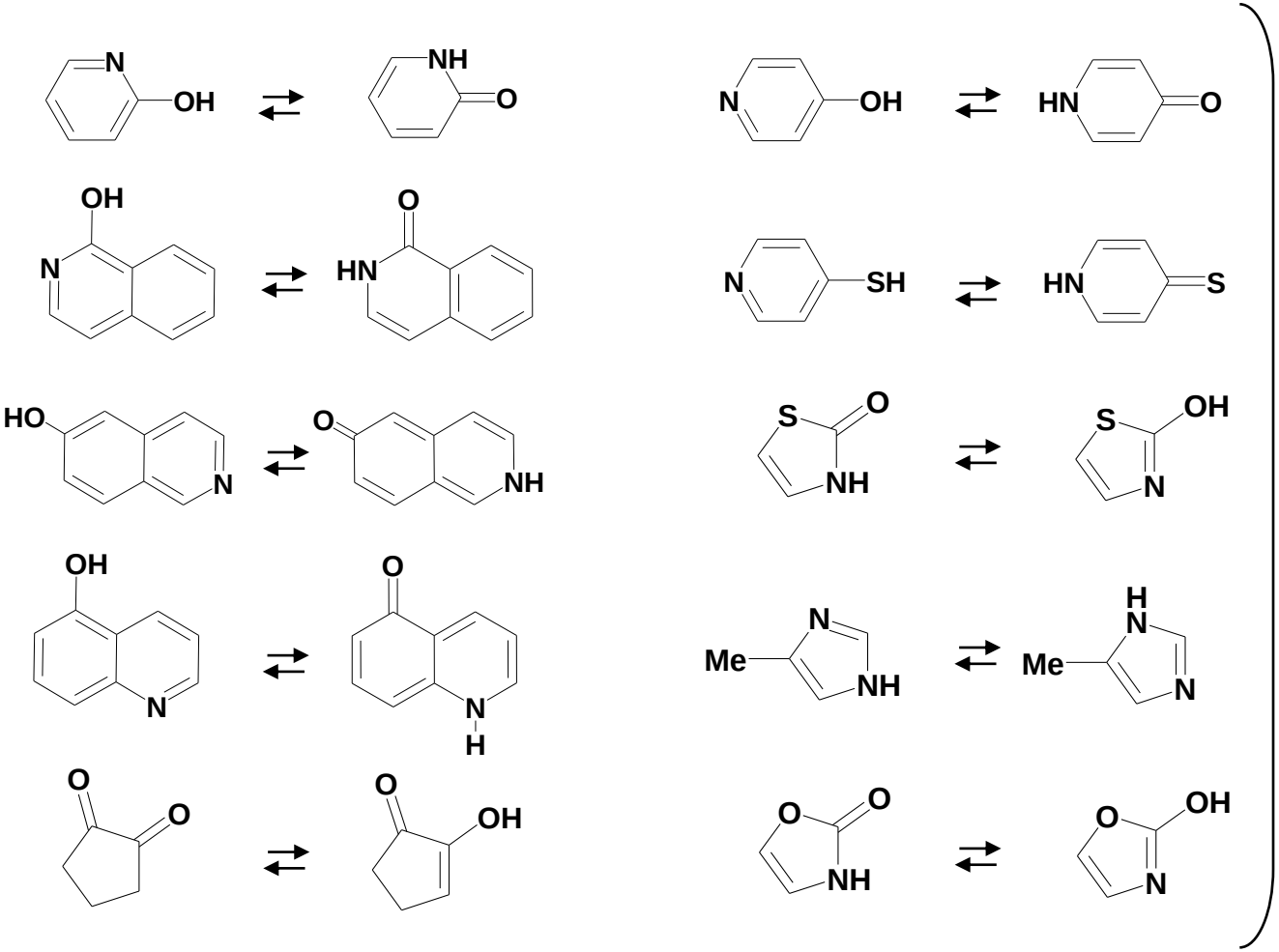


Software	Method	Mean Abs. Deviation from MP2 benchmarks (kcal)	Error*	
			Median	Maximum
Q-Chem	TRIM-MP2/ cc-PVTZ	0.085	0.02	0.8
Q-Chem	RI-MP2 / cc-PVTZ	0.044	0.0014	0.7
Jaguar	LMP2 /ps-cc-PVTZ	0.32	0.27	0.7
--	X3LYP / 6-31G**	0.36	0.23	1.6

*Based on absolute values of errors

B. Relative energies of 10 tautomeric pairs.

Data set



- **Chemically diverse**
- **Wide energy range:**

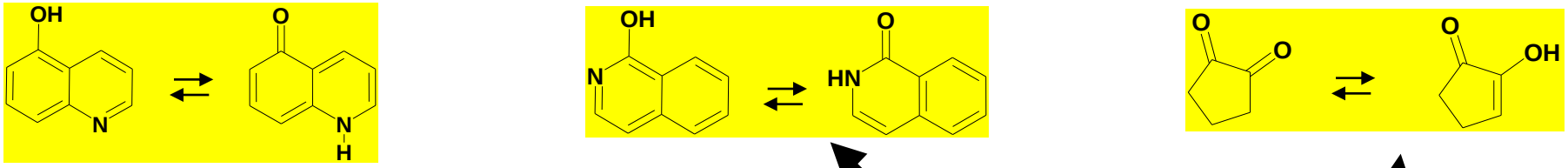
  Mean ΔE = 8.5 kcal

  Median ΔE = 5.8

  Max ΔE = 18.3

B. Relative energies of 10 tautomeric pairs.

Results



Software	Method	Ave. Abs. Dev	ΔΔE (error; kcal)*	
			Median	Max
Q-Chem	TRIM-MP2 / cc-PVTZ	0.085	0.12	0.44
Q-Chem	RI-MP2 / cc-PVTZ	0.041	0.004	0.22
--	X3LYP / 6-31G**	3.2	2.64	7.2
Jaguar	LMP2 / ps-cc-PVTZ	NA**		

**Jaguar LMP2 tautomeric energies were not included since delocalization of non-aromatic tautomers must be done manually, and so cannot be easily compared to automatic delocalization methods such as TRIM- and RI-MP2

Momentary Digression …

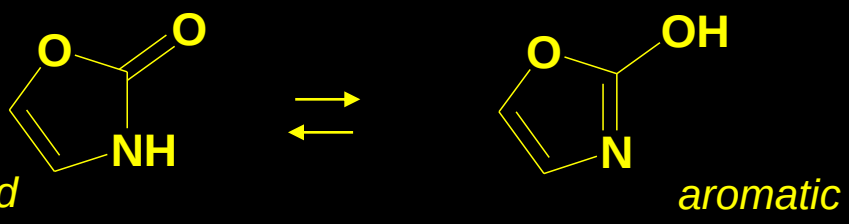
Jaguar LMP2

Jaguar LMP2 requires user-specified delocalization.

- *difficult for casual users*
- *how does one know a priori how to define delocalization?*

➢ *It was decided not to include Jaguar tautomeric energies.*

Example:

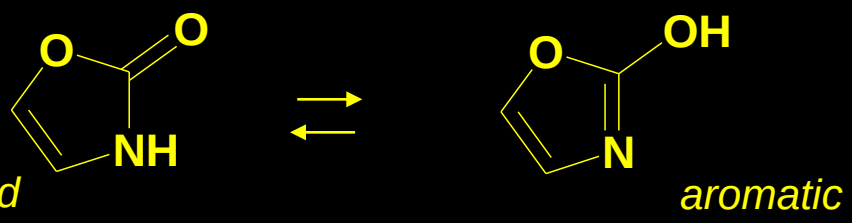


delocalized *aromatic*

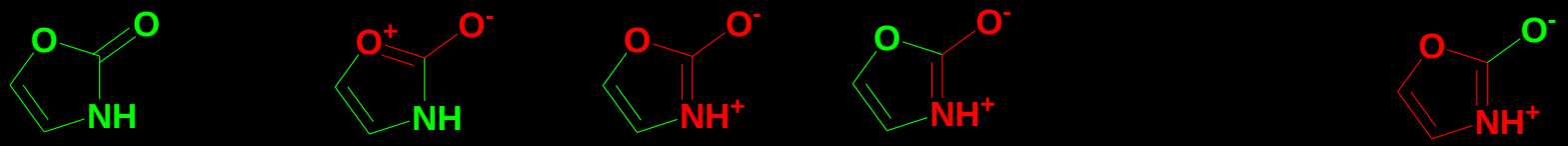
Jaguar LMP2 requires user-specified delocalization.

- *While not user-friendly, Jaguar LMP2 offers certain strengths*
- *Local – local MP2 offers insights into stabilizations resulting from specific de-localizations*

Example:



delocalized *aromatic*

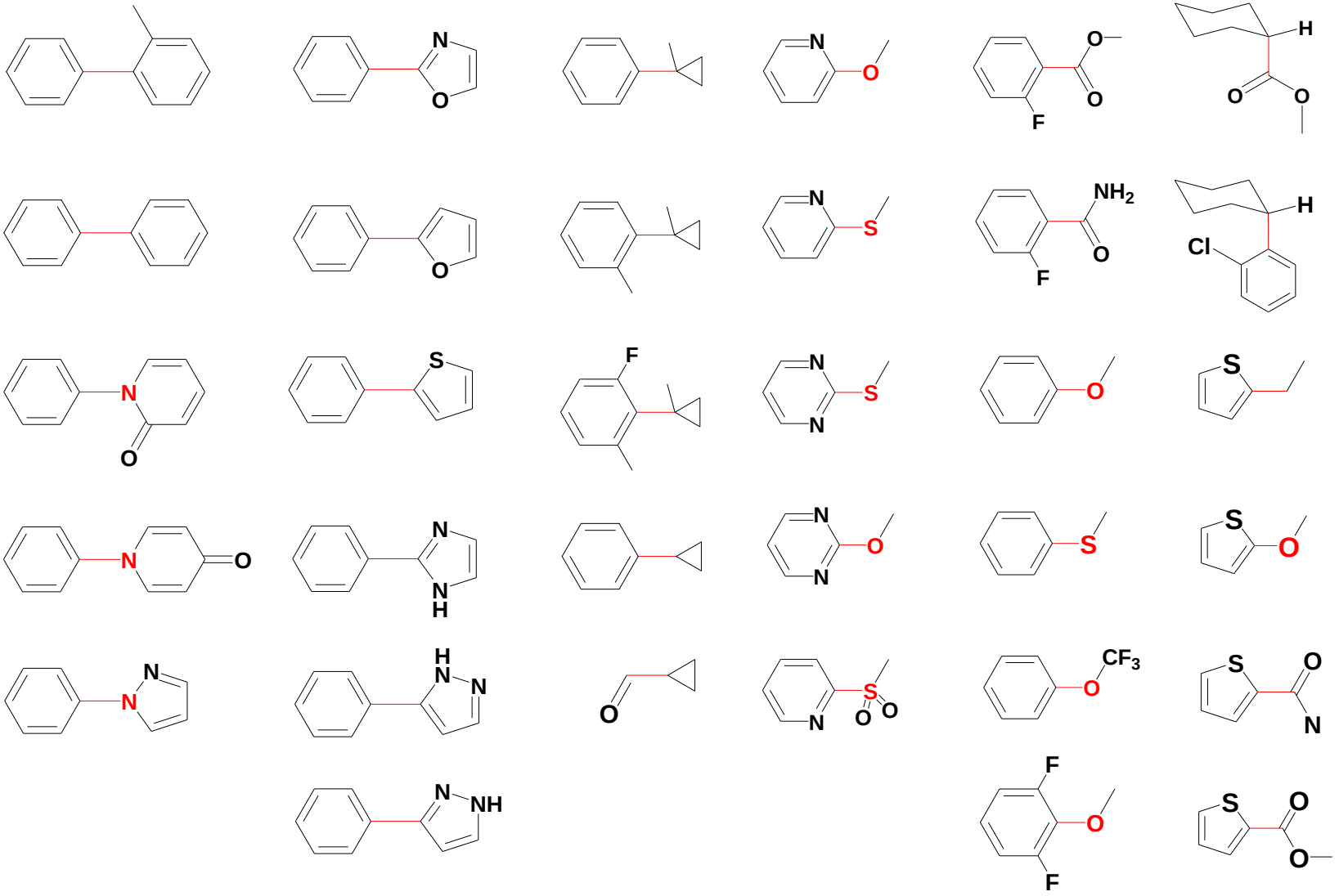


0.0 kcal *- 1.4* *- 1.6* *- 1.9* *- 3.1*

No delocalization *Full delocalization*

33 Torsional Scans (360° by 10° increments)

Data set: Chemically diverse, pharmaceutically relevant.



33 Torsional Scans (360^o by 10^o increments)

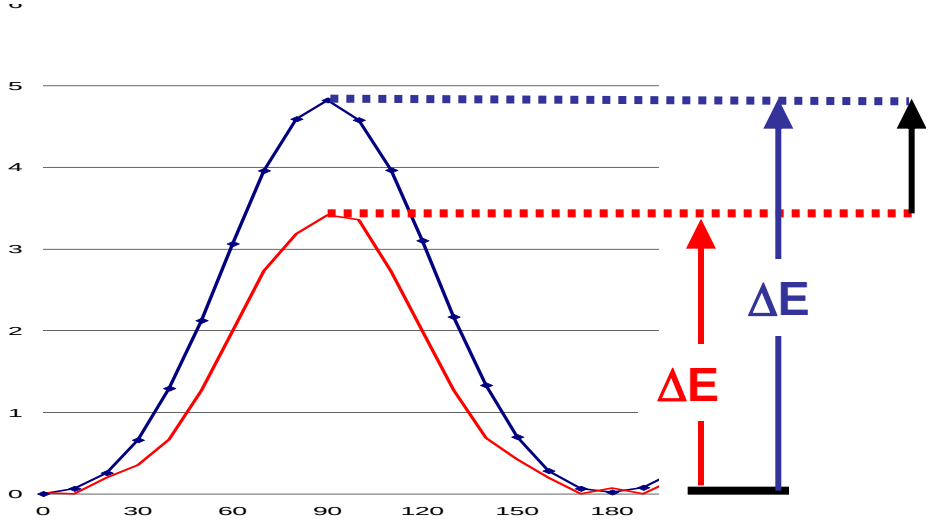
Results: Overall Quality of Torsional Potentials

Distorted Potential: Shape of potential differed mildly in some aspect, e.g., involving additional maxima (rare), or a reversal of local and global minima relative to benchmark potentials calculated at MP2/cc-PVTZ//X3LYP/631G**

Software	Method	# Distorted Potentials Relative to MP2
Q-Chem	TRIM-MP2 / cc-PVTZ	4 / 33
Q-Chem	RI-MP2 / cc-PVTZ	1 / 33
Jaguar	LMP2 / ps-cc-PVTZ	4 / 33
--	X3LYP / 6-31G**	2 / 33

RESULTS

Barrier Heights



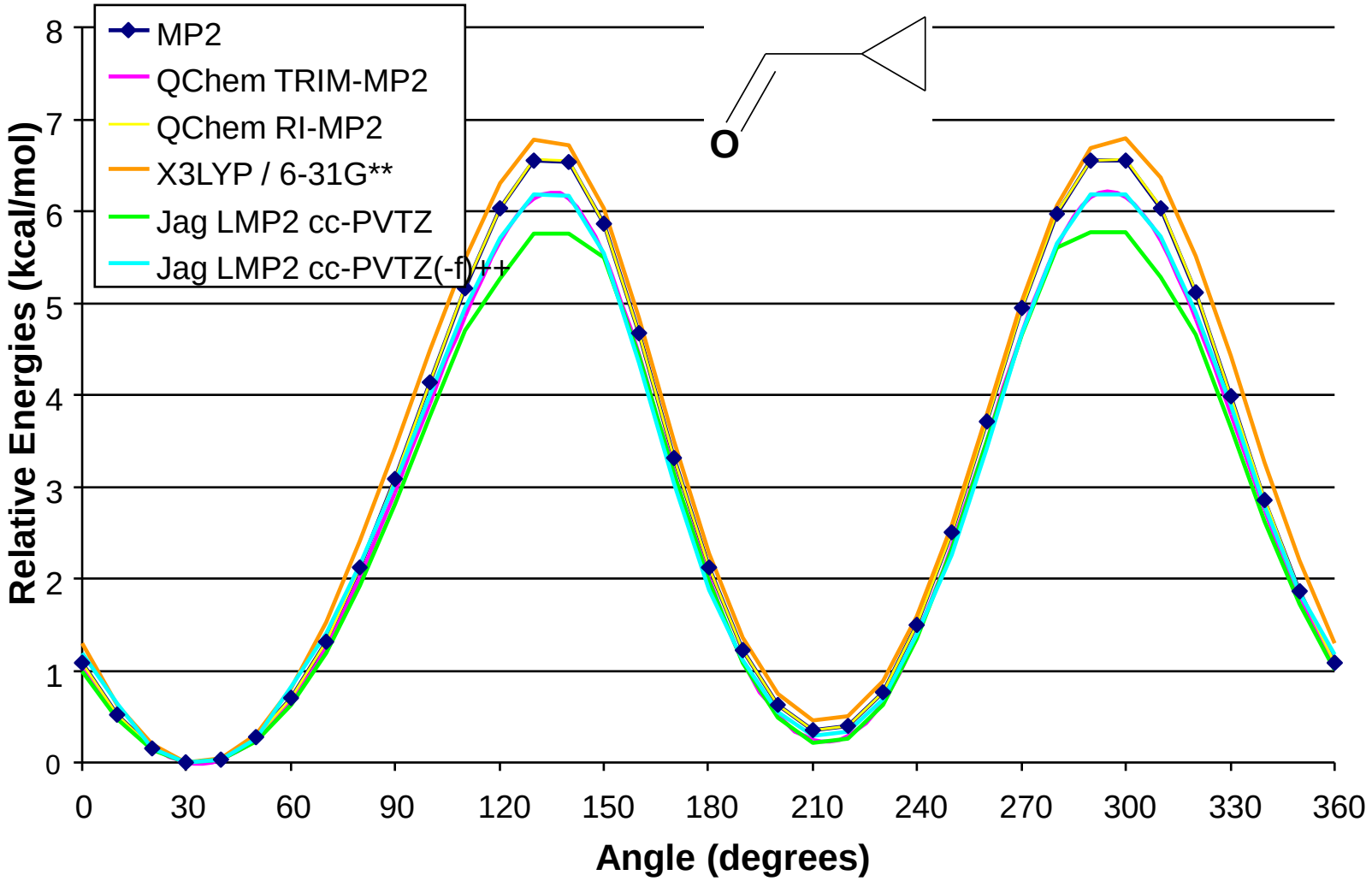
ΔΔE = difference in barrier heights

		ΔΔE (error; kcal)*		
Software	**Method**	**Median***	**Ave abs ΔΔE**	**Max. err.**
Q-Chem	TRIM-MP2/ cc-PVTZ	0.093	0.23	1.5
Q-Chem	RI-MP2 / cc-PVTZ	0.001	0.01	0.26
--	X3LYP / 6-31G**	0.38	0.58	4.6
Jaguar	LMP2 / ps-cc-PVTZ	0.48	0.50	1.3

*Based on absolute values of errors

33 Torsional Scans (360° by 10° increments)

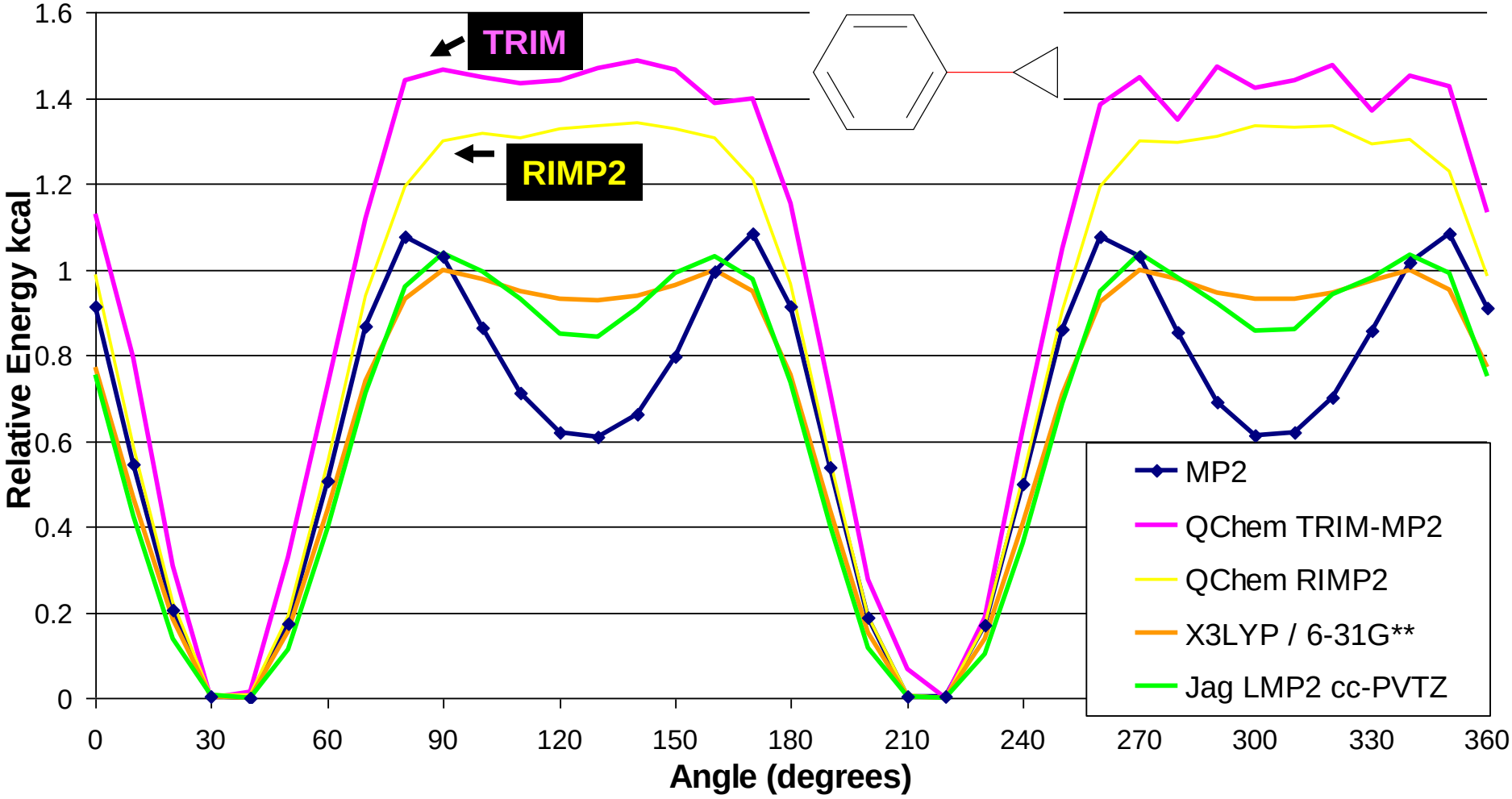
Results: Overall Quality of Torsional Potentials



- In some cases, all methods yield potentials with small deviations from MP2 benchmarks (dark blue with markers)

33 Torsional Scans (360° by 10° increments)

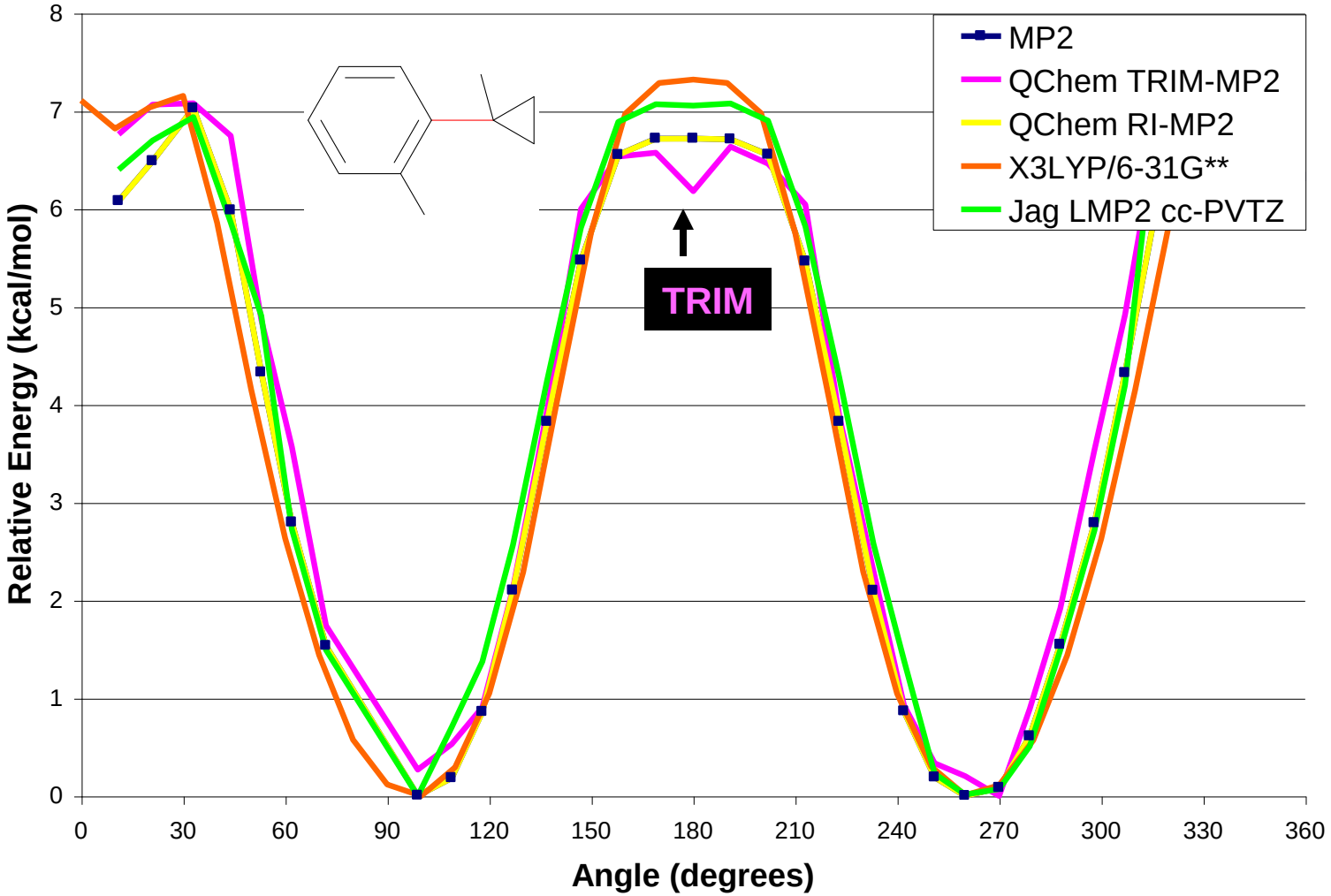
Results: Q-Chem RI-MP2 / TRIM-MP2 Distorted Potentials



- In this example, both Q-Chem TRIM and RI-MP2 have distorted potentials, missing the local MP2 minima

33 Torsional Scans (360º by 10º increments)

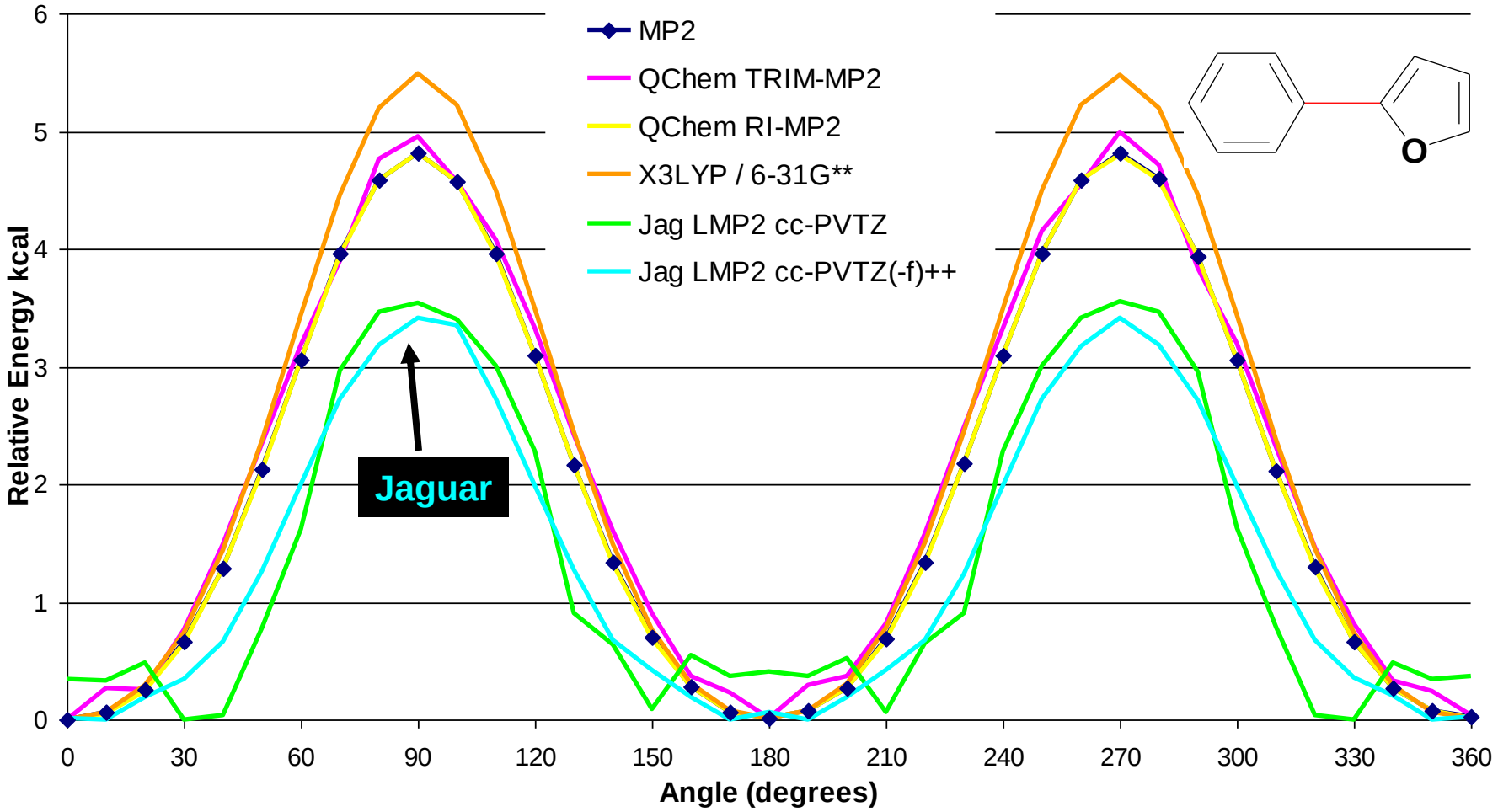
Q-Chem TRIM-MP2 results – distorted potential



Q-Chem LMP2 has dip at 180 degree maxima, this is seen in another very similar molecule.

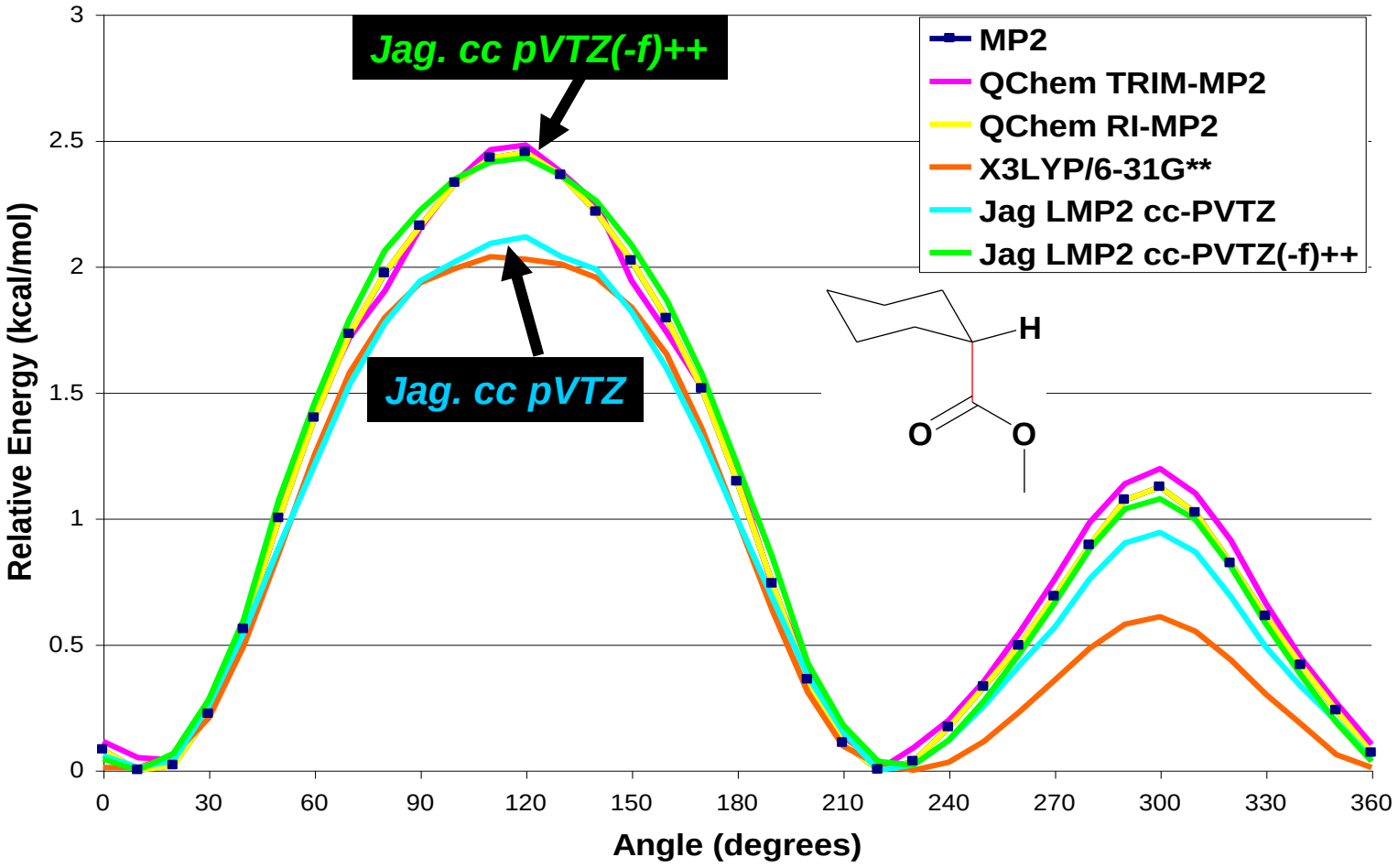
33 Torsional Scans (360º by 10º increments)

Results: Quality of torsional potentials



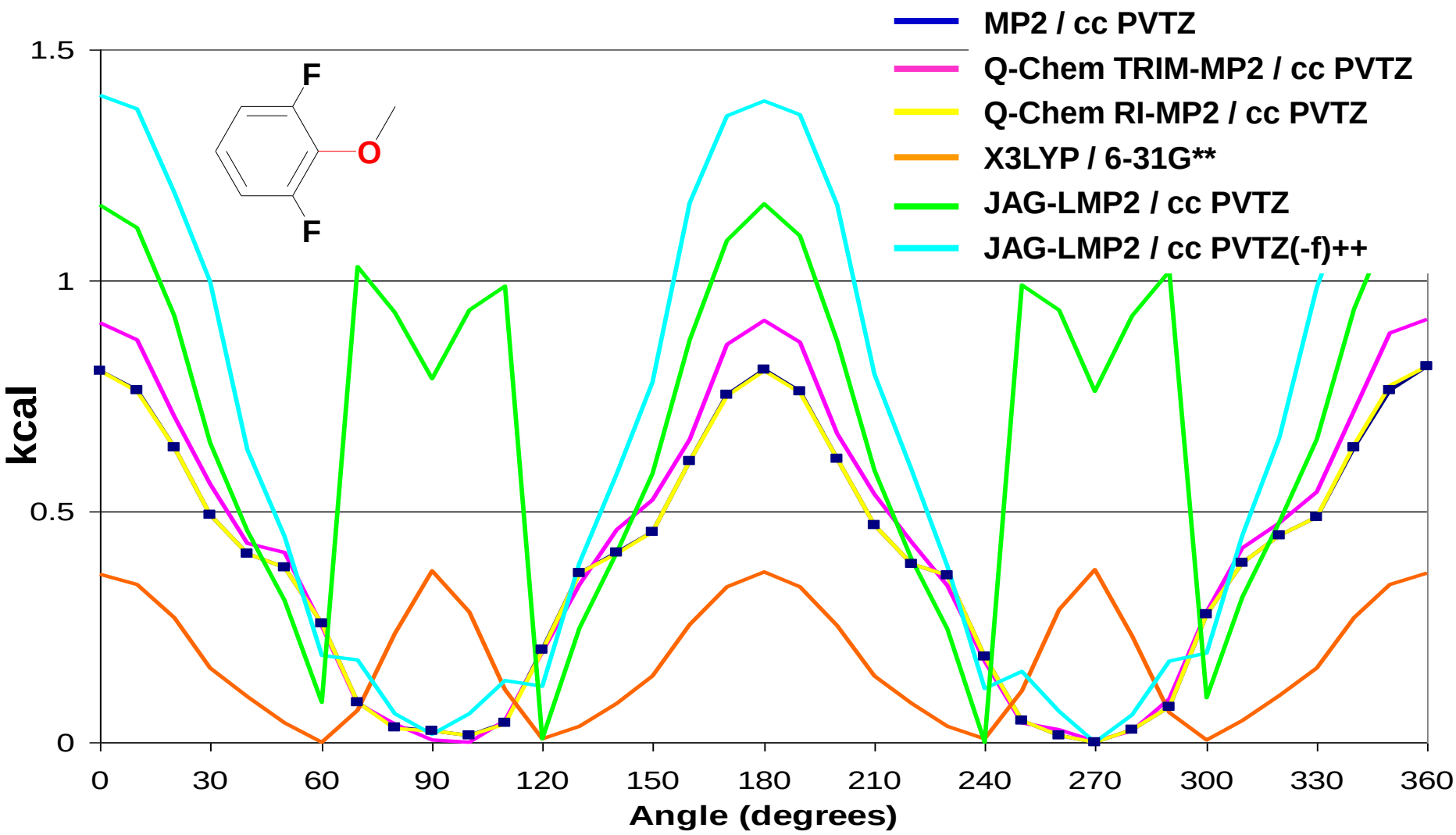
- Jaguar LMP2 / cc-PVTZ does not delineate the global minima at 0º and 180º. Instead angles +/- 30 degrees are determined to be the global minima.
- Using diffuse functions (cc-PVTZ(-f)++) fixes this problem.

Results: Dependency of Jaguar LMP2 on basis set



- Jaguar LMP2 calculations show different barrier heights when using diffuse functions
- X3LYP potential surface with maxima and minima identified correctly, however the magnitude of the barrier heights is incorrect.

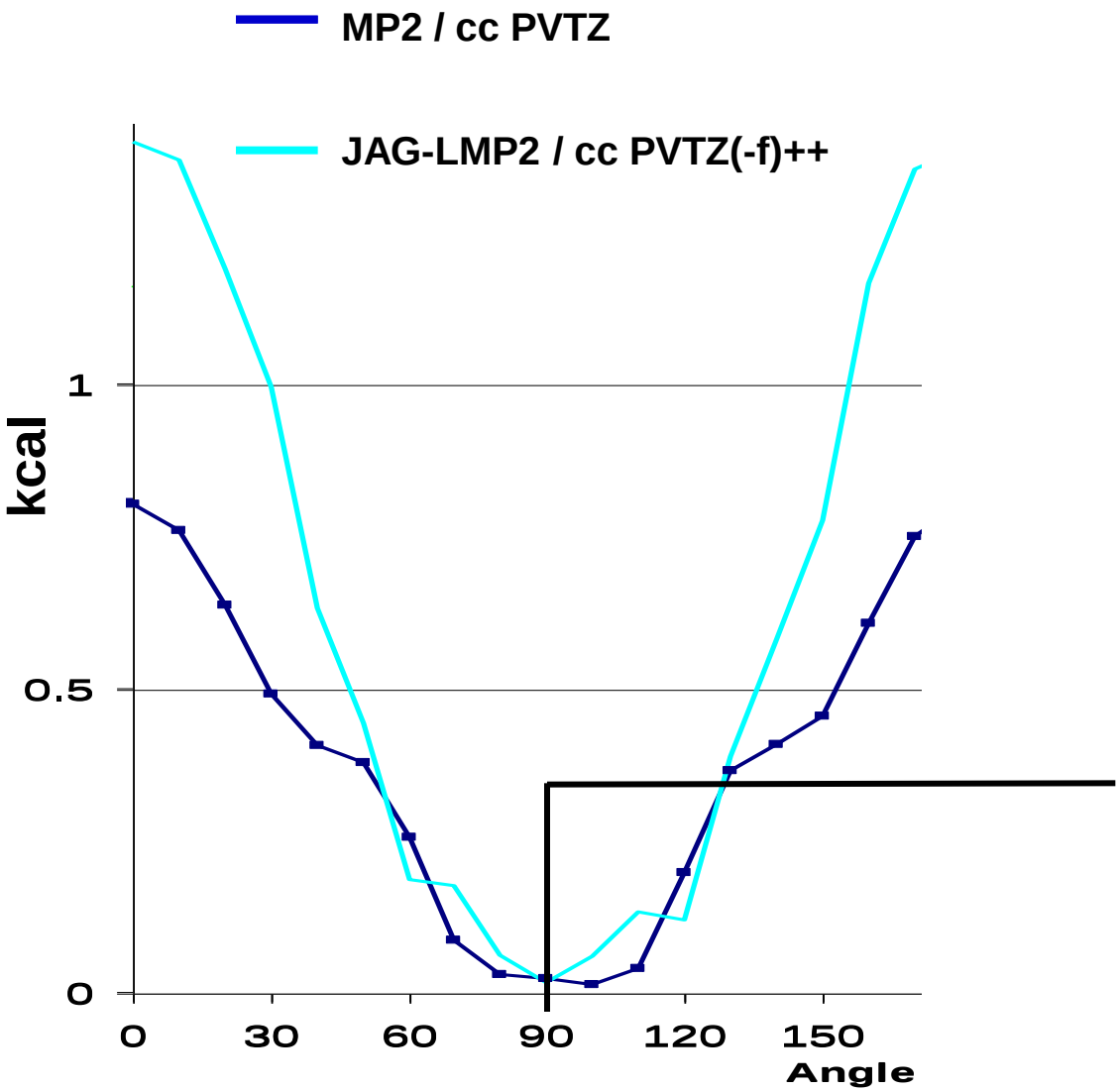
Results: Dependency of Jaguar LMP2 on basis set

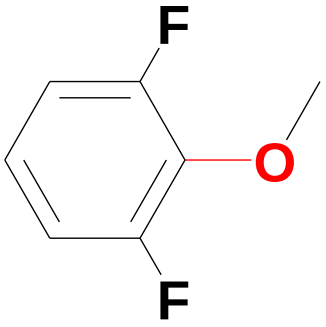


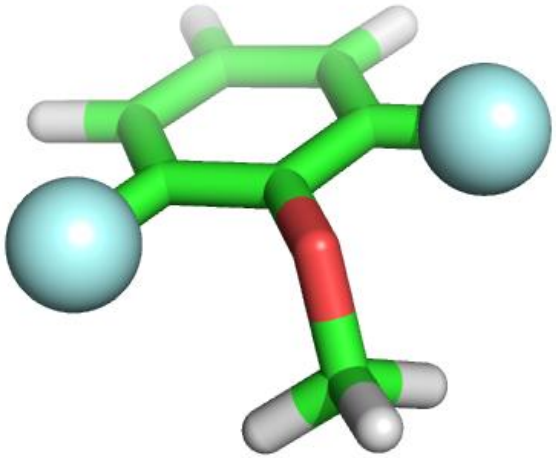
- Example of a distorted potential for DFT X3LYP showing global minima incorrectly identified as maxima
- Additionally, Jaguar LMP2 has the correct shape of the energy landscape, but overestimates the barrier heights

Results: Dependency of Jaguar LMP2 on basis set

Torsion is in an electrostatically complex environment







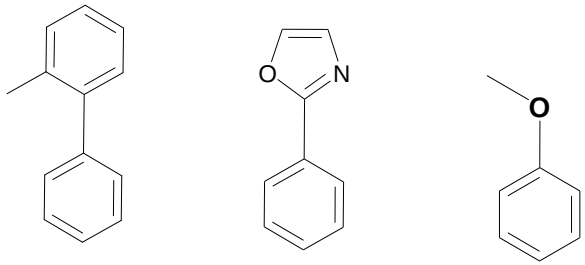
RESULTS

There appears to be some dependency of at least one implicit correlated method with basis set.

- *Perhaps a more fair comparison of methods may involve calculations with multiple basis sets (in lieu of a complete basis set)*

RESULTS

Timings (Min/CPU)



Q-Chem (“cpu” *)	MP2 / cc-PVTZ	496	125	33
Q-Chem (“cpu” *)	TRIM / cc-PVTZ	133	64	32
Q-Chem (“cpu” *)	**RI-MP2 / cc-PVTZ**	**42**	**23**	**13**
Jaguar (“user+sys” *)	LMP2 /ps-cc-PVTZ	82	39	15

- **Calculations for timings were performed on:** Beowulf Linux Cluster, 64 x AMD duo-core Opteron 246 chips.

***** Name given to timing value in output file.

SUMMARY

- RI-MP2 and TRIM-MP2 (Q-Chem) produced relative conformational energies which were nearly identical to MP2 benchmarks, with RI-MP2 having an average error of only 0.04 kcal.
- LMP2 (Jaguar) relative conformational energies had larger errors (average 0.32 kcal), while DFT using X3LYP/ 6-31G**, performed well yielding an average error of 0.35 kcal for relative conformational energies.
- RI-MP2 potentials, with 1 exception were very nearly identical to MP2 potentials
- TRIM-MP2 (Q-Chem) and LMP2 (Jaguar) torsion potentials each deviated in 4 of 33 cases. Potentials generated with X3LYP/ 6-31G** (Jaguar) were generally similar to benchmark potentials, although the barrier heights sometimes deviated.
- RI-MP2 - generated relative tautomeric energies which were nearly identical to MP2 benchmarks, while TRIM-MP2 (Q-Chem) reproduced, on average, relative tautomeric energies to within 0.1 kcal. DFT X3LYP produced significant errors (average = 3.2 kcal, max = 7 kcal).
- Jaguar LMP2 tautomeric energies were not included in this report since delocalization of non-aromatic tautomers must be done manually, and so cannot be easily compared to automatic delocalization methods such as TRIM- and RI-MP2.
- Timings of all the implicit MP2 methods were impressive. RI-MP2 was the fastest, which is remarkable given that it reproduced accurately nearly all MP2 benchmarks.

CONCLUSIONS

- ***All implicit MP2 methods performed reasonably well. However, in this study RI-MP2 (Q-Chem) demonstrated superior efficiency in terms of speed and accuracy, yielding results which were arguably indistinguishable from canonical MP2.***

- ***Basis set dependencies need to be taken into account when comparing implicit MP2 methods***

FUTURE WORK

- Extend protocol to account for basis set dependencies of correlated methods, i.e., repeat all calculations at cc-PVTZ(-f)++.
- Extend the current test set to include non-bonded pairs of molecules
- Extend the current test set to include small conformational sets of large, drug sized molecules.
- Other correlated techniques may be examined as well, most notably one or more of the recent Truhlar DFT functionals.

ACKNOWLEDGEMENTS

The authors wish to thank:

Malcolm Davis, Brian Claus, and Jano Jusuf for their support in maintaining the BMS-CADD computing infrastructure, which made this work possible.

Role of QC in Drug Discovery: calculation of hydrogen bond strengths

Geometry optimizations **X3LYP / 631G****

Train

$E_{bind} = [aE_{bind}(\text{cc - pVQZ}) - bE_{bind}(\text{cc - pVTZ})] / (a - b)$

counterpoise – corrected binding energies

12 dimers

Define a and b parameters by fitting (L)MP2 energies to energies from **CCSD(T) / cc-pVTZ**

Apply

$E_{bind} = [aE_{bind}(\text{cc - pVQZ}) - bE_{bind}(\text{cc - pVTZ})] / (a - b)$

counterpoise – corrected binding energies

Jaguar PS-LMP2

Compute binding energies using fixed a and b parameters

Estimated binding energies