A comparative review of computational methods as applied to gold(I) complexes and mechanisms

by

Jessica Reel

Department of Chemistry
Duke University

Date: ______________________

Approved:

_________________________
Ross A. Widenhoefer, Supervisor

_________________________
Richard A. MacPhail

_________________________
Jennifer L. Roizen

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Chemistry in the Graduate School of Duke University

2016
ABSTRACT

A comparative review of computational methods as applied to gold(I) complexes and mechanisms

by

Jessica Reel

Department of Chemistry
Duke University

Date:___________________

Approved:

___________________________
Ross A. Widenhoefer, Supervisor

___________________________
Richard A. MacPhail

___________________________
Jennifer L. Roizen

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Chemistry in the Graduate School of Duke University

2016
Abstract

In the last two decades, the field of homogeneous gold catalysis has been extremely active, growing at a rapid pace. Another rapidly-growing field—that of computational chemistry—has often been applied to the investigation of various gold-catalyzed reaction mechanisms. Unfortunately, a number of recent mechanistic studies have utilized computational methods that have been shown to be inappropriate and inaccurate in their description of gold chemistry. This work presents an overview of available computational methods with a focus on the approximations and limitations inherent in each, and offers a review of experimentally-characterized gold(I) complexes and proposed mechanisms as compared with their computationally-modeled counterparts. No aim is made to identify a “recommended” computational method for investigations of gold catalysis; rather, discrepancies between experimentally and computationally obtained values are highlighted, and the systematic errors between different computational methods are discussed.
Dedication

To the dear friends who helped me through it all—

Mom, Esther & Bud, Rachel, Sarah,

Christine, Caroline, Krista,

and of course my Walter—

Thank you.
Contents

Abstract .................................................................................................................................................. iv
List of Tables .......................................................................................................................................... vii
List of Figures ......................................................................................................................................... viii
Acknowledgements ............................................................................................................................. ix
1. Introduction and background ........................................................................................................... 1
2. A brief primer on computational chemistry .................................................................................... 4
   2.1 The evolution of computational methods ...................................................................................... 6
       2.1.1 Wavefunction-based methods ................................................................................................. 7
       2.1.2 Density functional theory ........................................................................................................ 12
       2.1.3 Basis sets .................................................................................................................................. 14
       2.1.4 Effective core potentials and relativistic effects ....................................................................... 16
   2.2 Selecting an appropriate computational method .......................................................................... 18
3. Computational modeling of gold(I) complexes ................................................................................ 24
   3.1 Gold(I)-alkene complexes ............................................................................................................. 26
   3.2 Gold(I)-alkyne complexes ............................................................................................................. 36
   3.3 Summary and conclusions ............................................................................................................. 41
4. Computationally-derived mechanisms of gold(I)-catalyzed reactions: the need to benchmark .......................................................... 43
5. Final thoughts and closing remarks ................................................................................................. 52
References ............................................................................................................................................... 54
List of Tables

Table 1: A “Jacob’s Ladder” of DFT functionals. ................................................................. 13

Table 2: Number of basis functions required for sample calculations. ............................ 16

Table 3: Computed bond lengths (Å) for the model complex [(ethylene)Au]+. ......................... 27

Table 4: Calculated§ and experimental† bond lengths (Å) for (alkene)AuCl complexes. ................................................................................................................................. 31

Table 5: Calculated§ and experimental† bond lengths (Å) and angles (°) for NHC- and phosphine-based gold(I)-alkene and -diene complexes. ............................................................................ 34

Table 6: Computed bond lengths (Å) for model complexes [(alkyne)Au]+. ............................. 37

Table 7: Calculated§ and experimental† bond lengths (Å) and angles (°) for (alkyne)AuCl complexes. ................................................................................................................................. 38

Table 8: Calculated§ and experimental† bond lengths (Å) and angles (°) for NHC- and phosphine-based L-Au-(alkyne) complexes. ......................................................................................... 40

Table 9: Energy deviations (kcal/mol) for stationary points in Figure 13, relative to CCSD/def2-TZVPP//CCSD/def2-SVP reference. Cells marked “dnc” indicate that the energy calculation failed to converge. Adapted from Faza et al., ref. [368]. ......................... 48
List of Figures

Figure 1: A simplified diagram of the self-consistent-field method. ................................. 8

Figure 2: Higher-order electron configurations in post-Hartree-Fock methods. ............... 10

Figure 3: Approaching an exact solution through a stepwise increase of basis set size and level of theory. .............................................................................................................. 19

Figure 4: Calculated ethylene C=C bond lengths, at different levels of theory across a range of basis sets, shown as difference from experimental value (1.339 Å).................. 20

Figure 5: Calculated acetylene C≡C bond lengths, at different levels of theory across a range of basis sets, shown as difference from experimental value (1.203 Å)........... 20

Figure 6: Gold-ethylene adducts synthesized by Dias and co-workers. Bond lengths in Å.............................................................................................................................................. 28

Figure 7: Computed bond lengths (Å) for [(ethylene)Au]⁺. Data points correspond to entries in Table 3.............................................................................................................................................. 29

Figure 8: Supporting ligands used in experimental and computational studies of gold(I) complexes........................................................................................................................................... 33

Figure 9: Alkenes and dienes employed in the experimentally- and computationally-characterized gold(I) complexes listed in Table 5................................................................. 34

Figure 10: Alkynes employed in experimentally- and computationally-characterized gold(I) complexes.......................................................................................................................... 38

Figure 11: Top: Key intermediates in the gold(I)-catalyzed reactions of propargylic esters. Bottom: Ligand effects on product ratios in the cycloisomerization of 6. See reference [366]. .............................................................................................................................................. 45

Figure 12: Calculated energies (relative to intermediate I) associated with the “golden carousel”. Round/square brackets correspond to L = IMe and PMe3, respectively; numbers adjacent to arrows represent transition states. Adapted from reference [365].. 46

Figure 13: The stationary points in Correa’s “golden carousel” that were used by Faza et al. to benchmark the performance of different computational methods. See refs. [96], [368] .............................................................................................................................................. 47
Acknowledgements

Thanks are due to my adviser, Dr. Ross Widenhoefer, for assistance preparing my NSF-GRFP proposal (even before I was officially a member of his group), as well as for his guidance the past two years, and particularly for allowing me to write the review that I really wanted to. Thanks also to Dr. Jennifer Roizen for assistance with the aforementioned NSF proposal, as well as the occasionally proffered word of advice or encouragement. I would additionally like to express my gratitude to several of my undergraduate professors—Dr. Monty Fetterolf, Dr. Chad Leverette, Dr. Gerard Rowe, Dr. Susan Glenn—for their congeniality and generosity throughout the years.

I am grateful to my current and former committee members—Dr. Jennifer Roizen, Dr. Richard MacPhail, Dr. Steven Malcolmson, and Dr. Weitao Yang—for their patience and understanding. Thank you also to Dr. Yang for valuable conversations in the area of computational chemistry, and to Dr. MacPhail for his willingness to step in as a last-minute substitute for Dr. Yang.

To my lab mates, past and present, you all have my appreciation. In particular, thank you to Dr. Robert Harris for his mentorship during my first year in the lab, and to Jacob Timmerman, Bob Carden, Nana Kim, and Nathan Lam for discussion, critiques, feedback, and proofreading.

Last but not least, I’d like to thank Dr. Jiyong Hong, Trish McMillan, Cecilia Eichenberger, and Caroline Morris for ensuring my awareness of requirements and
deadlines and for helping to guide me through the often-challenging logistics of committee changes and scheduling difficulties.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1106401. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.
1. Introduction and background

One of the most significant developments in homogeneous catalysis over the past decade has been the emergence of gold(I) complexes as catalysts for the functionalization of C–C multiple bonds, such as found in alkenes, alkynes, allenes, dienes, and enynes, among others. Beginning in 1998 with the work of Teles et al.\cite{1} on the activation of alkynes toward nucleophilic attack, the “gold rush” has exploded dramatically, with a huge number of gold(I)-catalyzed organic reactions being developed, investigated, optimized, and used in total synthesis applications. Gold-catalyzed reactions offer clear advantages over alternative approaches, as they typically involve mild reaction conditions, tolerate a wide range of functional groups, and are able to generate high levels of complexity with good atom efficiency under low catalyst loading. The interest and growth in the field of gold catalysis is demonstrated by the abundance of review publications, whether general\cite{2-14} or dealing with different reaction types,\cite{15-61} specific catalysts or substrates,\cite{62-74} key intermediates and mechanisms,\cite{75-102} or the potential use in total synthesis,\cite{103-127}

Despite the impressive synthetic and methodological accomplishments in this field, there is limited mechanistic understanding of gold(I) catalysis. The majority of proposed mechanisms are based, in part if not entirely, on computational modeling insofar as it agrees with experimental observation. Such computational models, unfortunately, are limited in scope at best; at worst, they utilize outdated methods that
may be inaccurate for modeling of transition metal complexes. Like homogeneous gold catalysis, the field of computational chemistry has been very active in the last two decades, with frequent publications relating improved algorithms and formulas. Regrettably, this rapid proliferation of computational approaches is difficult to remain abreast of, much less properly assess for applicability in the study of gold(I) catalysis.

Recent work by the Widenhoefer group has been directed toward experimental elucidation of the mechanisms\[128-137\] of gold(I)-catalyzed transformations,\[15, 138-159\] including the synthesis and study of relevant gold(I) complexes.\[90, 160-172\] The results of these investigations have highlighted significant inconsistencies between experimental observations and computationally-derived proposals for reaction intermediates. Such discrepancies are of great concern, as they call into question the validity of many of the proposed mechanisms upon which much of our understanding of gold(I) catalysis is based. Lamentably rare are examples of computationally-augmented mechanistic studies in gold catalysis that report on the validation or benchmarking of the computational methods used.

To that end, I submit this review and comparison of computational methods as applied to mechanistic studies in gold(I) catalysis. Chapter 2 begins this work with a very brief, non-technical overview of the various computational methods commonly applied to gold chemistry, with an emphasis on the advantages and limitations of each. Chapter 3 presents a comparative study of experimentally and computationally
determined geometries of several gold(I) complexes, and Chapter 4 examines proposed mechanisms for various gold-catalyzed reactions, and benchmarking studies that aim to prescribe a recommended computational methodology.
2. A brief primer on computational chemistry

“...the era of computing chemists, when hundreds if not thousands of chemists will go to the computing machine instead of the laboratory for increasingly many facets of chemical information, is already at hand. There is only one obstacle, namely, that someone must pay for the computing time.”

– Robert S. Mulliken, 1966 Nobel lecture

The IUPAC defines computational chemistry as “a discipline using mathematical methods for the calculation of molecular properties or for the simulation of molecular behavior”, and defines computer-assisted molecular modeling as “… the investigation of molecular structures and properties using computational chemistry and graphical visualization techniques.” Essentially, computational chemistry is the application of computer models and quantum mechanical calculations to solve chemical problems. These models and calculations can be used to great effect to study various atomic and molecular properties, including but not limited to: molecular structures, excited and transition states, reaction energies and pathways, magnetic properties, and spectroscopic frequencies.

The impact that computational chemistry methods have had on the field of chemistry in general is demonstrated by the Nobel Prize in Chemistry having been awarded in 1998 to John Pople and Walter Kohn, and in 2013 to Martin Karplus,

a. “for his development of computational methods in quantum chemistry”
b. “for his development of the density-functional theory”
Michael Levitt, and Arieh Warshel. Today, computational chemistry is often faster, cheaper, and more environmentally friendly than purely experimental work, and indeed, there are certainly advantages in utilizing computational methods to interpret, optimize, analyze, or predict experimental results. At the same time, the user must bear in mind that *all computational models are approximations*, with more accurate approximations generally requiring more time and more computational resources (such as memory and processing power). Therefore, in choosing a computational method, the user must not only weigh the availability of resources against the desired accuracy, but also consider and understand the underlying theory—especially the approximations used—to ensure the method is appropriate for the problem. In that regard, this chapter will attempt to present a very brief, qualitative (i.e., non-mathematical) primer on the variety of computational methods available, with a particular emphasis on the strengths and limitations of each. Discussion will be limited to *ab initio* wavefunction-based and density functional theory (DFT) methods; molecular mechanics and semi-empirical methods are not generally applied to investigations of gold catalysis, and so are outside the scope of this work.

The primary references for this chapter include texts by Lewars,[175] Cramer,[176] Jensen,[177] Young,[178] and Bachrach,[179] and the interested reader is thus referred to them e. “for the development of multiscale models for complex chemical systems”

---
d. Latin, “from first principles” or “from the beginning”; that is, formulated without being fitted to experimental data.
2.1 The evolution of computational methods

“The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.”

– Paul A. M. Dirac[180]

The large variety of computational chemistry methods available represent the various approaches to solving the Schrödinger equation for a many-electron system. The reason that the Schrödinger equation is so difficult to solve for molecular systems is, in essence, the foundation of chemistry itself: particles interact. For non-interacting particles, the electric potential of the molecular system only influences each particle separately, so the total energy (as given by the Schrödinger equation) can be expressed as the sum of energies for each particle, and the total molecular wavefunction can be written as a product of the individual particle wavefunctions. These convenient decompositions are not possible, however, for interacting particles, and therefore the Schrödinger equation can be solved exactly only for the hydrogen atom and related one-electron systems. For $N$-electron systems where $N \geq 2$, the wavefunction is a function of the positions of all $N$

---

[180] True, at least, within the Born-Oppenheimer approximation (the assumption that the electronic motion and the nuclear motion in molecules can be separated).

[1] Indeed, an exercise in many undergraduate physical chemistry courses involves deriving this solution.
electrons together, and there is no closed-form solution.\textsuperscript{9} In addition, the computational resources required to arrive at even an approximated solution to the Schrödinger equation tend to grow exponentially as the model system (e.g., molecule) grows in size; the wavefunction of an $N$-body system is a function of $3N$ variables ($x, y, z$ for each particle). With more particles, the problem rapidly becomes intractable: every time an extra particle is added to the system, the computational resources must be doubled.\textsuperscript{181}

2.1.1 Wavefunction-based methods

Of the various \textit{ab initio} computational chemistry methods, the lowest-level approximation is the Hartree-Fock (HF) method,\textsuperscript{182-184} also known as the self-consistent field (SCF) method as it uses a \textit{variational} (i.e., iterative) approach to arrive at a numerical solution to the Schrödinger equation (\textbf{Figure 1}). The starting point for an HF calculation is a set of approximate one-electron wavefunctions known as \textit{basis functions}, typically constructed as a \textit{linear combination of atomic orbitals} (LCAO) represented mathematically as a Slater determinant.\textsuperscript{185-186} In each iteration, the weighted coefficients of the LCAO are varied until the resultant energy is minimized;\textsuperscript{h} the best possible solution is at the \textit{Hartree-Fock limit} as the set of basis functions (i.e., \textit{basis set}) approaches completeness.

The major assumption made by HF theory is that the many-electron molecular

\textsuperscript{g} The term “closed-form” describes a formula that can be evaluated in a finite number of standard mathematical operations; it can be considered to be an “analytical” or “mathematically exact” solution.

\textsuperscript{h} The \textit{variational theorem} states that the energy of any approximate wavefunction is always an upper bound to the exact energy, and therefore the best LCAO is the one that gives the lowest energy. A proof of the theorem is given by Prof. David Sherrill at http://vergil.chemistry.gatech.edu/notes/quantrev/node28.html
wavefunction can be represented as a product of individual, non-interacting one-electron wavefunctions. In essence, the HF method evaluates the energy of each electron as though it were moving in an averaged Coulombic field produced by all the other electrons, rather than taking into account the instantaneous electron-electron repulsions. This neglect of electron correlation,[187-190] due to the use of only a single Slater determinant to express the electronic configuration, can lead to large deviations from experimental results, particularly for energy calculations. Nonetheless, HF remains a reasonable starting point for computational work prior to the use of more expensive methods.

Figure 1: A simplified diagram of the self-consistent-field method.

A number of approaches, collectively referred to as post-Hartree-Fock methods, have been devised to better account for electron correlation and thus better approximate the true energy. Any \( N \)-electron wavefunction can, in theory, be expressed exactly as a linear combination of all possible \( N \)-electron Slater determinants formed from a “complete” (i.e., infinite) set of basis functions.[191-195] This ideal method is called full configuration-interaction, or full-CI, and constitutes the exact solution to the non-
relativistic Schrödinger equation within the basis set used. In reality, such computations are only practical with either minimal basis sets or very small molecules; the number of determinants required grows factorially with the number of electrons.\textsuperscript{[196]}

The use of a necessarily finite number of Slater determinants and basis functions yields an approximate solution; such methods are known as configuration interaction (CI) methods, where the prefix “full” has been dropped. CI methods express the wavefunction as a linear combination of multiple Slater determinants generated by substituting some number of occupied electronic orbitals in the HF “reference” determinant for previously vacant orbitals (i.e., “promoting” or “exciting” one or more electrons; see Figure 2 for examples), and use a variational method similar to HF to determine the lowest energy. Examples of CI methods include CISD, which includes all determinants containing either one or two excited electrons,\textsuperscript{i} and multi-reference CI (MRCI), which uses more than one reference determinant to generate the excitations. A problem of such truncated CI methods is that they are not size-consistent; for a system A+B, in which A and B are separated by a sufficiently large distance so there is essentially zero shared electron density, the energy given by a size-consistent method should equal the sum of the individual energies of A and B.\textsuperscript{[197]} The property of size consistency is of particular importance for calculations of dissociation reactions.

\textsuperscript{i} CISD stands for “configuration-interaction, singles & doubles”, where “singles” and “doubles” refer to the inclusion of all one- and two-electron excitations, respectively; the method CID, as an example, would include only doubles.
An alternate approach to including electron correlation is perturbation theory, which deconstructs the wavefunction into parts that are solvable (the “reference” wavefunction) and unsolvable (the “perturbation”). The reference wavefunction is solved first, and the resulting energy is then corrected to account for the perturbation. These corrections are expressed as an infinite series of terms, which become smaller and smaller for well-behaved systems. Møller–Plesset perturbation methods,\textsuperscript{[198-199]} which use HF as the reference wavefunction, are among the best-known; these methods are typically designated as MP\textit{n}, where \textit{n} represents the number of terms included from the infinite series: MP2 calculates through second-order terms, MP4 through fourth-order. Unlike
truncated CI methods, MP methods are size-consistent. The main limitation of perturbation methods is the assumption that the reference wavefunction is a reasonable approximation to the exact wavefunction, i.e., that the perturbation is sufficiently small. The more poorly the HF wavefunction describes the system, the larger the correction must be, and thus the more perturbation terms must be included to achieve a given level of accuracy. If the reference wavefunction is a poor description of the system, convergence may be so slow or erratic as to be useless; it has also been demonstrated that convergence may depend on the size of the basis set.

The coupled cluster (CC) method is related to both the CI approach and perturbation theory. Coupled-cluster theory expresses the correlation correction to the HF wavefunction as a Taylor series expansion with terms corresponding to excited determinants, analogous to those in CI; including singles and doubles gives CCSD. For many applications, CCSD, while relatively inexpensive computationally, does not provide sufficient accuracy except for the smallest systems and so at least an approximate treatment of triple excitations is needed. The most well-known coupled cluster method that provides an estimate of connected triples is CCSD(T), in which an estimate to the triples contribution is calculated using perturbation theory. Standard CC methods, like MP methods, are based on a single-determinant reference wavefunction, and like MP, they work best if that reference is sufficiently good. Due to the summation of terms through infinite order, however, CC is somewhat more tolerant of a poor
reference than MP methods. CCSD(T) calculations are, generally speaking, the current benchmark for practical calculations on molecules of up to moderate size,\textsuperscript{[208]} typically achieving accuracy within 2 kcal/mol or better, given a sufficiently large basis set.\textsuperscript{[209-213]}

### 2.1.2 Density functional theory

For density functional theory (DFT)\textsuperscript{[214-216]} the framework is established by the theorems of Hohenburg and Kohn\textsuperscript{[217]} and the equations of Kohn and Sham.\textsuperscript{[218]} In brief, DFT seeks to describe the electronic state of a molecule as an electron density function in 3 spatial coordinates, greatly reducing the computational expense as compared to wavefunction-based methods (which require $3N$ spatial coordinates for a system of $N$ electrons). Like the HF approach, DFT is a variational method, as it iteratively attempts to minimize the energy with respect to the electron density; this relationship of the energy to the density is the functional. The exact functional is as yet unknown,\textsuperscript{[219-220]} so a plethora of approximations have been developed; these approximations have been described as forming a “Jacob’s ladder” of increasing accuracy (Table 1).\textsuperscript{[221-223]}

The simplest approximation (i.e., the bottom rung on the ladder) is the local density approximation (LDA),\textsuperscript{[224]} in which the energy depends only on the local electron density where the functional is evaluated ($\rho(r)$). An improvement of this description is the generalized gradient approximation (GGA),\textsuperscript{[225]} which also considers the gradient of the electron density at the evaluation point ($\nabla \rho(r)$). Such approaches are quite rough for most chemical problems but have the advantage of scaling very well for larger systems;
meta-GGA functionals retain the convenient scaling properties but improve accuracy by including dependence on the kinetic energy density \((\tau(\rho))\) as well.

**Table 1: A “Jacob’s Ladder” of DFT functionals.**

<table>
<thead>
<tr>
<th>Rung</th>
<th>Approximation</th>
<th>Evaluation</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>dispersion-corrected, meta-hybrids, double hybrids, etc.</td>
<td>(\rho, \nabla \rho, \tau, \text{HF})</td>
<td>M06, B2PLYP, wB97, cam-B3LYP, LC-wPBE, LC-BLYP, mPW2PLYP</td>
</tr>
<tr>
<td>4</td>
<td>hyper-GGA, hybrids</td>
<td>(\rho, \nabla \rho, \text{HF})</td>
<td>B3LYP, PBE0, TPSSh, B3PW91, wPBEh, mPW3PBE, HSE06</td>
</tr>
<tr>
<td>3</td>
<td>meta-GGA</td>
<td>(\rho, \nabla \rho, \tau)</td>
<td>M06L, TPSS, PBEKCI, BB95, VSXC, MN12L, tHCTH, PKZP, KCIS</td>
</tr>
<tr>
<td>2</td>
<td>GGA</td>
<td>(\rho, \nabla \rho)</td>
<td>PBE, BLYP, PW91, BP86, HCTH</td>
</tr>
<tr>
<td>1</td>
<td>LDA</td>
<td>(\rho)</td>
<td>SVWN, LSDA, VWN80, PW92, PZ81</td>
</tr>
</tbody>
</table>

Hybrid functionals compute some percentage of the energy using HF methods and so (due to cancellation of errors) generally perform better than GGAs; the most popular DFT method, B3LYP,[225-228] is a hybrid method. Incorporation of additional dispersion interactions and corrections generally shows an improvement over the standard hybrids; this category contains the highest-level functionals in routine use, with B2PLYP[229-230] in particular having shown comparable accuracy to very high-level coupled-cluster methods[231] at a fraction of the computational time.

Unfortunately, there is not yet a single, “universal” functional[220, 222] that has been demonstrated to be the best for a wide variety of chemical problems: for every property there seem to be one or two functionals that are superior, but only for that property.[232-234] In contrast to wavefunction-based methods, DFT methods typically do not offer a
straightforward means by which to systematically improve the accuracy of calculations (e.g., CCD → CCSD → CCSDT, etc.); the main benefit of DFT is that calculations tend to be much faster, scaling as $N^4$—compared with $N^5$ for MP2 or $N^8$ for CISD(T)—while generally offering an accuracy comparable to MP2 calculations. For insight into the current challenges and limitations of DFT—some of which will be discussed at the end of this chapter—the reader is referred to the 2012 review article by Yang and co-workers.\textsuperscript{[235]}

2.1.3 Basis sets

A basis set is a set of mathematical functions (basis functions), linear combinations of which yield molecular orbitals. These functions are usually centered on atomic nuclei, with several basis functions describing the electron distribution around an atom; combining atomic basis functions yields the electron distribution in the molecule as a whole. A particular type of basis function known as a Slater-type orbital (STO)\textsuperscript{[236]} is a good approximation to atomic orbital wavefunctions, but is difficult to evaluate; therefore, most basis sets use various combinations of Gaussian-type orbitals (GTOs) to approximate the STOs.\textsuperscript{[237]} A minimal basis set contains only the minimum number of basis functions required to represent all of the electrons on each atom (e.g., 1s for hydrogen and helium; 1s, 2s, 2p\textsubscript{x}, 2p\textsubscript{y}, 2p\textsubscript{z} for Li–Ne, etc.); the most common minimal basis set is STO-3G. Minimal basis sets are computationally less expensive than their larger counterparts, but typically provide far less accurate results. Split-valence basis sets
offer an improvement in accuracy by representing valence orbitals by more than one basis function; these are typically denoted as double-zeta, triple-zeta, etc. Some of the most well-known split-valence basis sets are those of the Pople group,[238] which include the 3-21G, 6-31G, and 6-311G basis sets. All of these sets can be optionally augmented with polarization and/or diffuse functions. Polarization functions add flexibility by including higher orbital types, such as p-orbitals on hydrogen or d-orbitals on carbon; these are denoted by one (*) or two (**) asterisks or a listing of the extra functions included (e.g. (d,p)), after the basis set name. Diffuse functions are very shallow Gaussian functions that more accurately represent electron distribution at a distance from the nucleus, which can be important when considering anions or “soft” atoms. The inclusion of diffuse functions is denoted by one (+) or two (++) plus signs, or by the prefix aug-. Another set of widely-used basis sets are the correlation-consistent basis sets developed by Dunning and coworkers,[239] since they are designed to systematically improve results with increasing basis set size and thus allow extrapolation to the complete basis set (CBS) limit.[240] These sets are named cc-pV\(n\)Z, where \(n\) = D, T, Q, 5, etc.; as an example, cc-pVTZ is a correlation-consistent polarized, valence triple-zeta basis set. The Pople 6-31G* and the Dunning cc-pVDZ are roughly comparable in size; correlation-consistent basis sets sometimes[241] give results superior to similarly-sized Pople sets, but not always.[242] It is useful to have a general idea of the number of basis functions in a given molecule, as this will determine the cost of the calculation (Table 2).
Table 2: Number of basis functions required for sample calculations.

<table>
<thead>
<tr>
<th>Basis set</th>
<th>Description</th>
<th>number of functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>STO-3G</td>
<td>minimal</td>
<td>H 1</td>
</tr>
<tr>
<td>3-21G</td>
<td>double split-valence</td>
<td>2</td>
</tr>
<tr>
<td>6-31G*</td>
<td>double split-valence with polarization</td>
<td>2</td>
</tr>
<tr>
<td>6-31G**</td>
<td>same, with p functions on H</td>
<td>5</td>
</tr>
<tr>
<td>6-311+G**</td>
<td>triple split-valence with polarization, p functions on H, and diffuse functions on heavy atoms</td>
<td>6</td>
</tr>
</tbody>
</table>

2.1.4 Effective core potentials and relativistic effects

At about the third row (K–Kr) of the periodic table, the large number of electrons in each atom begin to significantly slow down calculations due to the many two-electron repulsion integrals they engender. One method of overcoming this problem is to consider only the collective effect of the core electrons on the valence electrons (which are still considered explicitly). This average core effect is called an effective core potential (ECP) or a pseudopotential. The use of an ECP stands in contrast to using all-electron basis sets such as the Pople or Dunning sets mentioned above. In addition to speeding up calculations (with, of course, a tradeoff in accuracy), ECPs can also be used to account for relativistic effects in heavy atoms such as transition metals.\[243-249\]

As atomic nuclear charge (Z) increases, electrons that penetrate to the nucleus (the s electrons) increase their average velocity and as a consequence of relativity, their mass. This relativistic effect causes the s electrons (and to a lesser extent, the p electrons) to be in smaller orbitals than if this effect were absent. Therefore, in the heavier
elements, the $s$ electrons are more strongly bound and shield the nuclear charge from the other electrons (especially $d$ and $f$) more effectively; the $d$ and $f$ electrons are thus less bound and occupy larger orbitals. These effects scale roughly with $Z^2$ and become important for elements heavier than the lanthanides;\textsuperscript{[250-251]} for the elements Au–Bi the impact on energies is comparable with chemical bond energies.\textsuperscript{[252]} Indeed, gold, because it is also at the end of the 5$d$ orbital filling (third transition series contraction effect), following on the 4$f$ filling (lanthanide contraction), exhibits maximum impact from this relativistic effect.\textsuperscript{[251, 253-254]} For gold ($Z = 79$) the 1$s$ electrons will be going 58\% of the speed of light, increasing the relativistic mass by a factor of 1.22 and contracting the atomic radius by 22\%.\textsuperscript{[255]} The importance of the relativistic effect in gold chemistry was noted by Pitzer\textsuperscript{[252]} and Pyykkö and Desclaux\textsuperscript{[254]} nearly twenty years ago. More recently Kaltsoyannis\textsuperscript{[253]} has summarized the impact of the effect for inorganic and organometallic chemistry, and Pyykkö has provided a comprehensive series of reviews on relativistic effects,\textsuperscript{[251, 254, 256-263]} particularly as they relate to gold chemistry.\textsuperscript{[264-266]} Certainly, accurate calculations of complexes that contain the heavier elements should attempt to account for relativistic effects, either through use of a relativistic ECP (RECP) basis set (such as LANL2DZ\textsuperscript{[267-269]} or SDD\textsuperscript{[270-271]}) or via a relativistic all-electron approximation such as the Douglas-Kroll-Hess (DKH)\textsuperscript{[272-274]} or the zeroth-order regular approximation (ZORA)\textsuperscript{[275-277]} methods.
2.2 Selecting an appropriate computational method

There is something to be said for having such a wide variety of computational tools available; on the other hand, Dewar certainly had a point when, three decades ago, he described the situation as a “chaotic proliferation”. There are dozens of books of practical advice that aim to provide a feel for the appropriateness of various computational methods, and by simply reviewing the literature one may learn what approaches have been successfully applied to various problems, especially those related to one’s own research. That said, it would be inappropriate to assume that the method used in a published work is the one most suitable for the problem, especially if no indication of benchmarking or validation studies have been provided. Unfortunately, the “most appropriate” method and basis set for a given problem can rarely be picked from the start as a matter of routine; rather, it often must be found via a systematic series of trials, and comparison of computational results with experimental facts as far as possible. (Although, when egregious deviations from experiment are found at computational levels that experience—and experts—suggest should be reliable, one may be justified in questioning not only the computational methods, but perhaps the experimental methods as well.)

For novel systems for which there is no prior work to serve as a guide, a rational approach to selecting an appropriate computational method may be to conduct an initial exploration of the problem using either a semi-empirical or a Hartree-Fock method with...
a minimal basis set, and move up to larger basis sets with post-Hartree-Fock or DFT methods in a stepwise approach toward chemical accuracy (Figure 3).

![Diagram showing increasing level of theory vs. basis set size]

**Figure 3**: Approaching an exact solution through a stepwise increase of basis set size and level of theory.

Typically, a balanced approach between the size of the basis set and the level of theory is recommended, since the use of a very large basis set is pointless in many cases;\[281\] on the other hand, some problems yield only to very large basis set calculations (if at all). Note that it is possible for results to actually become *worse* with increasing basis set size;\[282-284\] due to chance cancellation of the larger errors found in lower-level methods; several papers emphasize the goal as one of finding “the right answer for the right reason”.\[285-288\] This fluctuation in accuracy with basis set size is demonstrated in
Figure 4 and Figure 5 for the difference from experimental values for calculated C-C bond distances in ethylene and acetylene, respectively.\cite{footnote}

Figure 4: Calculated ethylene C=C bond lengths, at different levels of theory across a range of basis sets, shown as difference from experimental value (1.339 Å).

Figure 5: Calculated acetylene C≡C bond lengths, at different levels of theory across a range of basis sets, shown as difference from experimental value (1.203 Å).
Interestingly, the DFT methods (LDA, BLYP, B3LYP, mPW1PW, M06-2X) show a tendency to overbind (i.e., the calculated bond lengths are shorter than the experimental value) in almost all cases for ethylene (Figure 4), but tend to move from underbinding to overbinding with increasing basis set size for acetylene (Figure 5). Given the widely-accepted “chemical accuracy” target goal for calculated bond lengths of within a few picometers (say, ±0.025 Å) the data summarized here are not terrible; however, considering that experimental X-ray crystallography (XRC) data can be accurate (for a crystal of sufficient purity and regularity) to within a few thousandths of an angstrom, and that a given bond length in two similar structures may indeed differ by less than 0.01 Å, one may be understandably wary at accepting conclusions based on even “chemically accurate” computational methods.\footnote{On the other hand, Dias and co-workers\cite{291} remind us that often, changes in the C–C bond length upon coordination of coinage metals to unsaturated substrates are harder to discern using routine XRC techniques, because relatively small lengthenings are often overshadowed by high estimated standard deviations.}

Another particularly important objective is to calculate thermochemical quantities to chemical accuracy, generally within 1 kcal/mol. Attempting to achieve reliable, consistent chemical accuracy for a general polyatomic molecule from a single calculation has proven to be an impossible task at this point in the evolution of computational chemistry techniques. Ideally, one would use a basis set near the CBS limit; the issue in performing such calculations is the computational scaling of the method with the size of the basis set. The lowest-level wavefunction-based method that
can be used to obtain near thermochemical accuracy is CCSD(T), which computationally scales as \( M^7 \), where \( M \) is the number of basis functions. To directly achieve thermochemical accuracy with CCSD(T) without extrapolation would require an all-electron treatment using a basis set of at least aug-cc-pCV6Z quality,\(^{292}\) which is currently feasible only for very small molecules.

Due to their much lower computational cost compared with higher-level wavefunction-based methods, DFT methods—particularly B3LYP—have remained a popular choice for computational investigations,\(^k\) in spite of identified limitations of DFT in general\(^{219,294}\) and B3LYP in particular.\(^{294-297}\)\(^{298}\) DFT calculations are known to underestimate transition state energies and reaction barriers, and fail to accurately describe near-degenerate states, such as arise in transition metal systems or the breaking of chemical bonds.\(^{219,294}\) An issue known as “overdelocalization”, in which structures with more-delocalized electron densities are preferentially stabilized over more-localized alternatives, leads to erroneous predictions of higher-symmetry structures being preferred over lower-symmetry ones.\(^{235}\)\(^{299}\) Consequences of overdelocalization include poor predictions for such investigations as nucleophilic substitution, competing cycloaddition pathways, and rotations around single bonds in conjugated systems. Note that HF theory tends to be inaccurate for such situations in the opposite direction from

\(^k\)Sousa et al. report\(^{232}\) that in the period 2002–2006, each year B3LYP has accounted for ~80% of functional names mentioned in the literature; indeed, the original publications by Becke\(^{228}\) and Lee, Yang, and Parr\(^{226}\) are the most-cited articles in the entire field of chemistry\(^{293}\).
DFT, and so again, hybrid functionals tend to show improved performance by this offsetting of errors. Even so, the inaccuracy of hybrid DFT methods (such as B3LYP) for describing transition metal chemistry has been particularly well-noted,[297, 300-302] yet there remains a tendency to use B3LYP “in order to maintain comparability with previous studies” even as its inadequacies are noted by those using it.[303]

Many recently developed DFT methods include various corrections to account for such known issues.[304-319] It remains to be seen whether such methods—assuming they are as generally applicable—will replace the ever-popular B3LYP as the method of choice for computational investigations.
3. Computational modeling of gold(I) complexes

“It doesn’t matter how beautiful your theory is, it doesn’t matter how smart you are. If it doesn’t agree with experiment, it’s wrong.”

--Richard P. Feynman

From a practical standpoint, realistic modeling of a typical gold(I)-catalyzed organic transformation—involving dozens of atoms and multiple reaction steps—using very high-level computational methods can be prohibitively expensive in terms of computer time and memory. Although CCSD(T) may be considered the “gold standard” in computational chemistry, reliably providing a high level of accuracy at the lowest possible cost, its $N^7$ scaling with the number of basis functions effectively limits its general application to the smallest of systems. For mechanistic studies in homogeneous gold catalysis, that require simulation of the substrate, metal center, and ligands (and ideally, counterions and explicit solvent molecules as well), the complexity of the model system oftentimes consigns us to using less-computationally-expensive methods such as DFT to achieve “reasonably” accurate results.

Benchmarking studies using highly accurate methods, such as CCSD(T) and beyond, have necessarily focused on very small gold compounds,[320-325] with the usual bulky phosphine or N-heterocyclic carbene (NHC) supporting ligand often being severely truncated. Unfortunately, such artificially constructed systems are of course too unrealistic to be synthesized and experimentally studied—and the effect of the ligands, counterions, and solvent molecules involved in a given reaction cannot be neglected as
these have all been shown to affect the catalytic pathways and products.\cite{69, 326-330} On the other hand, computational studies of such “artificial” systems might still provide valuable information regarding electron distribution in a ligand-metal-substrate system, and aid prediction and interpretation of coordination geometries, barriers for ligand exchange, etc. When the factors contributing to each aspect of bonding, geometry, or structure are dissected and identified, they can sometimes prove a good reference point for the development of catalytic systems with desired properties.\cite{79, 91, 327, 331}

Two recent reviews by Faza and co-workers\cite{83, 96} discussed the lack of comprehensive benchmarking studies seeking to validate the adequacy of the most popular computational methods being used for the investigation of gold-catalyzed organic reactions. The authors report that most studies used DFT, with B3LYP being the most popular functional, and 6-31G* with LANL2DZ (as an effective core potential for gold) the most popular basis set. Although this combination has historically provided consistently good results at low computational cost for organic molecules, it had never been explicitly tested for gold complexes. Even as larger basis sets and newer, improved functionals started to be more commonly applied to the investigation of gold catalysis, this lack of validation persisted. Faza and López state:

\begin{quote}
Among the mechanistic studies which cared about justifying the choice of functional or basis set, the most extended argument in support of a given combination was the citation of some other previous mechanistic studies which had used it, or an a posteriori agreement of the calculated data with the experimental observations.\cite{96}
\end{quote}
Of the handful of benchmarking studies that have sought to make a recommendation of computational methods to use in modeling homogeneous gold catalysis, most were purely theoretical in nature; that is, the accuracy of a tested method was determined as compared to calculations done at the highest level of theory, rather than as compared to experimental data.[332] Given the fact that all calculations—even those at the highest level—are still approximations, this chapter seeks to validate the conclusions drawn from such high-level calculations by offering a comparison of experimentally- and computationally-determined geometries in various gold(I)-alkene and -alkyne complexes.

Brooner and Widenhoefer provided a critical review in 2013 covering cationic two-coordinate gold π complexes,[90] following a more general 2010 review of gold η2-coordination to unsaturated substrates by Schmidbaur and Schier.[80] Of the previously reviewed complexes, the present work is concerned primarily with those for which experimental bond distances have been determined; these are summarized in table form in the following pages. As well, many of the earlier examples included in the Schmidbauer review are unrepresentative of compounds typically involved in a gold-catalyzed reaction scheme, and therefore will generally not be repeated here.

### 3.1 Gold(I)-alkene complexes

Contemporary with the first experimental reports of gold π complexes, theoretical approaches to the interaction between ethylene and cationic gold using semi-empirical computational methods were published.[333-334] Ziegler et al. in 1979 published
results of Hartree-Fock calculations for the model complex \((\text{ethylene})\text{Au}^+\) that indicated a symmetrical \(\eta^2\)-bonding motif, but also exhibited a too-long Au–C distance of 2.47 Å (significantly higher than all later values).\(^{335}\) Well-defined crystal structure information for a gold(I)-alkene complex was finally obtained in 1987 by Strähle and co-workers for \((\text{cis-cyclooctene})\text{AuCl}\).\(^{336}\) In the 2000’s, many more examples of gold(I)-alkene complexes were characterized both experimentally and computationally, but to date (as far as this author can tell) no serious analysis comparing the results of the two approaches has been presented.

A 1996 study by Hertwig et al.\(^{320}\) was among the first “benchmarking” studies of computational methods as applied to gold(I) complexes. The theoretical “naked” cationic gold complex \((\text{ethylene})\text{Au}^+\) was optimized at three different levels of theory.

<table>
<thead>
<tr>
<th>#</th>
<th>method</th>
<th>C=C</th>
<th>Au–C</th>
<th>ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B3LYP/6-31G**/LANL2DZ</td>
<td>1.400</td>
<td>2.243</td>
<td>[21]</td>
</tr>
<tr>
<td>2</td>
<td>B3LYP/TZ2P</td>
<td>1.390</td>
<td>2.231</td>
<td>[320]</td>
</tr>
<tr>
<td>3</td>
<td>B3LYP/pVDZ/SDD</td>
<td>1.397</td>
<td>2.253</td>
<td>[337]</td>
</tr>
<tr>
<td>4</td>
<td>B3LYP/6-31+G*/SDD</td>
<td>1.399</td>
<td>2.249</td>
<td>[337]</td>
</tr>
<tr>
<td>5</td>
<td>B3LYP/pVDZ/LANL2DZ</td>
<td>1.400</td>
<td>2.242</td>
<td>[337]</td>
</tr>
<tr>
<td>6</td>
<td>B3LYP/6-31+G*/LANL2DZ</td>
<td>1.403</td>
<td>2.239</td>
<td>[337]</td>
</tr>
<tr>
<td>7</td>
<td>B3LYP/6-31+G*/LANL2DZ</td>
<td>1.404</td>
<td>2.226</td>
<td>[338]</td>
</tr>
<tr>
<td>8</td>
<td>BP86/TZP/ZORA</td>
<td>1.414</td>
<td>2.165</td>
<td>[321]</td>
</tr>
<tr>
<td>9</td>
<td>LDA/TZ2P</td>
<td>1.371</td>
<td>2.180</td>
<td>[320]</td>
</tr>
<tr>
<td>10</td>
<td>MP2/TZ2P</td>
<td>1.403</td>
<td>2.099</td>
<td>[320]</td>
</tr>
<tr>
<td>11</td>
<td>mPW1PW/pVDZ/SDD</td>
<td>1.394</td>
<td>2.213</td>
<td>[337]</td>
</tr>
<tr>
<td>12</td>
<td>mPW1PW/6-31+G*/SDD</td>
<td>1.396</td>
<td>2.209</td>
<td>[337]</td>
</tr>
<tr>
<td>13</td>
<td>mPW1PW/pVDZ/LANL2DZ</td>
<td>1.397</td>
<td>2.201</td>
<td>[337]</td>
</tr>
<tr>
<td>14</td>
<td>mPW1PW/6-31+G*/LANL2DZ</td>
<td>1.400</td>
<td>2.199</td>
<td>[337]</td>
</tr>
</tbody>
</table>
(MP2, LDA, and B3LYP) using the same basis set for all. A similar benchmarking study on the same theoretical model was carried out in 2006 by Cinellu and co-workers.[337] The computed bond distances for C=C and Au–C from these two studies, along with those from various other calculations on the same complex, are summarized in Table 3 (vide supra) and presented graphically in Figure 7 (vide infra). For experimental studies of gold-ethylene complexes we are unfortunately limited to exotic structures bearing multidentate ligands that provide improved stability; such adducts have been synthesized and characterized by Dias and co-workers,[339-340] and the experimentally determined bond lengths are presented in Figure 6.

![Complex Structures](image)

Figure 6: Gold-ethylene adducts synthesized by Dias and co-workers. Bond lengths in Å.
Figure 7: Computed bond lengths (Å) for [(ethylene)Au]+. Data points correspond to entries in Table 3.

Although a direct comparison of the model complexes [(ethylene)Au]+ with the complexes 1a–1c reported by Dias et al. would be misleading due to the characteristics
of the ligands in the latter case, a few observations merit discussion. Firstly, other than a few outliers (particularly the LDA calculation,\textsuperscript{a} data point #9), the computed C=C bond lengths (Figure 7, dark circles) fall within a fairly narrow range spanning only 0.024 Å. This is good, since different methods should yield very similar results for the same complex, but not surprising, considering that most of the popular DFT methods have been optimized to describe organic compounds. Cinellu and co-authors\textsuperscript{[337]} note the similar C=C bond lengths (as also compared with those obtained by Hertwig\textsuperscript{[320]} and Lee et al.\textsuperscript{[338]}), as well as the good agreement of their calculations on free ethylene (1.330–1.335 Å) versus the experimental value (1.339 Å).\textsuperscript{[341]} On the other hand, the C=C bond distances of gold-ethylene adducts 1a–1c, as measured by X-ray crystallography (XRC), fall within the same narrow range of 0.024 Å (Figure 6), highlighting an unfortunate situation where the uncertainty in computational measurements on the same compound is of the same order of magnitude as the ligand-influenced bond length difference.

With respect to the Au–C bond distances, the experimentally-determined measurements for 1a–1c (Figure 6) are fairly consistent (within a range of 0.019 Å) and quite a bit lower ($d_{max} = 2.108$ Å) than almost all\textsuperscript{b} of those calculated for the model [(ethylene)Au]+ complexes (Figure 7, light squares). The calculated Au–C bond lengths for the model ethylene-gold cations also cover a wider range (0.088 Å) than the corresponding C=C bond lengths. This is also observed by Cinellu and co-workers,

\textsuperscript{a} Like Hartree-Fock, LDA is known to consistently overbind, resulting in bonds that are artificially short.
\textsuperscript{b} Considering the systematically too-long Au–C bonds produced by DFT methods (as noted in the following pages), the MP2 calculation—though an outlier here—could potentially be the most accurate in this dataset.
particularly the underbinding (overestimation of bond lengths) exhibited by the B3LYP functional as compared with the mPW1PW functional; additional DFT calculations by the same authors also showed that mPW1PW more closely reproduces experimental results than does B3LYP. Notably, within each series, the SDD RECP gives the poorest estimates of the bond dissociation energy (BDE),\textsuperscript{c} corresponding to the longest Au–C and shortest C=C bond distances.

Unfortunately, severely truncated model systems such as [(ethylene)Au]\textsuperscript{+}, while possibly providing useful benchmarking data, are not necessarily accurate for describing realistic gold(I) complexes used in catalysis. A step above such models would be to include the chlorine atom as a simple supporting ligand, especially given the fact that AuCl has been shown to be catalytically active.\textsuperscript{[344-346]} A search of the literature for experimentally-characterized chlorogold(I)-alkenes, however, returned few examples, which are summarized in Table 4.

<table>
<thead>
<tr>
<th>alkene</th>
<th>C=C (Å)</th>
<th>Au–C1 (Å)</th>
<th>Au–C2 (Å)</th>
<th>Cl–Au (Å)</th>
<th>ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ethylene\textsuperscript{§}</td>
<td>1.382</td>
<td>2.248</td>
<td>2.248</td>
<td>**</td>
<td>[21]</td>
</tr>
<tr>
<td>propene\textsuperscript{§}</td>
<td>1.386</td>
<td>2.230</td>
<td>2.290</td>
<td>**</td>
<td>[21]</td>
</tr>
<tr>
<td>cis-cyclooctene\textsuperscript{†}</td>
<td>1.38</td>
<td>2.15</td>
<td>2.21</td>
<td>2.266</td>
<td>[336]</td>
</tr>
<tr>
<td>dicyclopentadiene\textsuperscript{‡}</td>
<td>1.38</td>
<td>2.20</td>
<td>2.16</td>
<td>2.276</td>
<td>[347]</td>
</tr>
</tbody>
</table>

\* not reported.

With such a limited data set, not many comparisons can be drawn, especially

\textsuperscript{c} Cinellu et al. remark\textsuperscript{[337]} that the BDE for the [(η\textsubscript{2}-ethylene)M]\textsuperscript{+} moiety has been utilized as a criterion for the validation of theoretical results (see also [320], [338], [342-343]). In the case of M = Au, a lower limit of 59 kcal/mol for the BDE was established experimentally,\textsuperscript{[342]} while values in the range 62–71 kcal/mol have been calculated at various levels of theory.\textsuperscript{[330, 341]}
given the cyclic versus acyclic nature of the alkenes in the experimental versus computational complexes, respectively.d Interestingly, while the C=C bond lengths appear to be fairly consistent between the computational and experimental complexes, the Au–C bond distances are significantly longer for the computational models. Given the expectation for strained cyclic alkenes to bind more strongly to transition metals (and thus to exhibit shorter Au–C bonds), the trend seems reasonable at first glance; however, note that this trend mirrors that seen in the [(ethylene)Au]⁺ complexes above, indicating a systematic Au–C underbinding across all functionals used.

The most catalytically relevant gold-alkene complexes that have been structurally characterized include contributions by Widenhoefer and co-workers,¹⁶⁰-¹⁶¹, ¹⁶³ and the group of Green and Russell,³⁴⁸ utilizing supporting ligands that have become widely popular in gold-catalyzed reactions (Figure 8). In 2009, Brown, Dickens, and Widenhoefer characterized a series of IPr-Au-alkene complexes,¹⁶⁰ followed by a set of (o-biphenyl)di-tert-butylphosphine (JohnPhos, P₁) supported gold-alkene complexes,¹⁶¹ Green & Russell also in 2009 reported a group of tri-tert-butylphosphine (P₂) gold alkene complexes,³⁴⁹ and a series of P₁-supported gold dienes was published in 2011 by Brooner and Widenhoefer.¹¹⁶³

---

d. Strained cyclic substrates are expected to bind more strongly to metals, since donation of electron density rehybridizes the substrate thereby relieving some ring strain.
Figure 8: Supporting ligands used in experimental and computational studies of gold(I) complexes.

Computational models of non-truncated, catalytically-relevant gold-alkene complexes are seemingly rare, but a 2010 paper by Macchioni and co-workers reported calculations on a series of IPr-Au-alkenes,[33] which are included in Table 5 alongside the experimentally-characterized complexes. The entire set of alkenes and dienes employed in the above studies is shown in Figure 9.
Figure 9: Alkenes and dienes employed in the experimentally- and computationally-characterized gold(I) complexes listed in Table 5.

Table 5: Calculated§ and experimental† bond lengths (Å) and angles (°) for NHC- and phosphine-based gold(I)-alkene and -diene complexes.

<table>
<thead>
<tr>
<th>alkene</th>
<th>L</th>
<th>free</th>
<th>bound</th>
<th>Au–C1</th>
<th>Au–C2</th>
<th>L–Au</th>
<th>L–Au–∥</th>
<th>ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2a§</td>
<td>IPr</td>
<td>1.339</td>
<td>1.380</td>
<td>2.272</td>
<td>2.401</td>
<td>2.054</td>
<td>**</td>
<td>[331]</td>
</tr>
<tr>
<td>2a§</td>
<td>MeIPr</td>
<td>1.339</td>
<td>1.379</td>
<td>2.280</td>
<td>2.399</td>
<td>2.057</td>
<td>**</td>
<td>[331]</td>
</tr>
<tr>
<td>2b§</td>
<td>IPr</td>
<td>1.356</td>
<td>1.402</td>
<td>2.346</td>
<td>2.345</td>
<td>2.053</td>
<td>**</td>
<td>[331]</td>
</tr>
<tr>
<td>2c§</td>
<td>IPr</td>
<td>1.34</td>
<td>1.39</td>
<td>2.25</td>
<td>2.42</td>
<td>2.05</td>
<td>**</td>
<td>[331]</td>
</tr>
<tr>
<td>2b†</td>
<td>IPr</td>
<td>**</td>
<td>1.346</td>
<td>2.239</td>
<td>2.23</td>
<td>1.998</td>
<td>177</td>
<td>[160]</td>
</tr>
<tr>
<td>2d†</td>
<td>IPr</td>
<td>**</td>
<td>1.331</td>
<td>2.199</td>
<td>2.285</td>
<td>2.006</td>
<td>172</td>
<td>[160]</td>
</tr>
<tr>
<td>2e†</td>
<td>IPr</td>
<td>**</td>
<td>1.374</td>
<td>2.224</td>
<td>2.248</td>
<td>1.996</td>
<td>175</td>
<td>[160]</td>
</tr>
<tr>
<td>2b†</td>
<td>P1</td>
<td>**</td>
<td>1.325</td>
<td>2.292</td>
<td>2.293</td>
<td>2.287</td>
<td>163</td>
<td>[161]</td>
</tr>
<tr>
<td>2c†</td>
<td>P1</td>
<td>**</td>
<td>1.319</td>
<td>2.199</td>
<td>2.308</td>
<td>2.286</td>
<td>162</td>
<td>[161]</td>
</tr>
<tr>
<td>2f†</td>
<td>P1</td>
<td>**</td>
<td>1.369</td>
<td>2.210</td>
<td>2.365</td>
<td>2.299</td>
<td>165</td>
<td>[161]</td>
</tr>
<tr>
<td>2d†</td>
<td>P2</td>
<td>**</td>
<td>1.349</td>
<td>2.224</td>
<td>2.350</td>
<td>2.293</td>
<td>**</td>
<td>[349]</td>
</tr>
<tr>
<td>2e†</td>
<td>P2</td>
<td>1.334</td>
<td>1.366</td>
<td>2.281</td>
<td>2.299</td>
<td>2.298</td>
<td>175</td>
<td>[349]</td>
</tr>
<tr>
<td>3a†</td>
<td>P2</td>
<td>1.337</td>
<td>1.370</td>
<td>2.273</td>
<td>2.296</td>
<td>2.294</td>
<td>**</td>
<td>[349]</td>
</tr>
<tr>
<td>3b†</td>
<td>P1</td>
<td>**</td>
<td>1.363</td>
<td>2.226</td>
<td>2.332</td>
<td>**</td>
<td>163</td>
<td>[163]</td>
</tr>
<tr>
<td>3c†</td>
<td>P1</td>
<td>**</td>
<td>b</td>
<td>2.245</td>
<td>2.320*</td>
<td>**</td>
<td>166*</td>
<td>[163]</td>
</tr>
<tr>
<td>3d†</td>
<td>P1</td>
<td>**</td>
<td>1.367*</td>
<td>2.214*</td>
<td>2.308*</td>
<td>**</td>
<td>167*</td>
<td>[163]</td>
</tr>
<tr>
<td>3e†</td>
<td>P1</td>
<td>**</td>
<td>1.352</td>
<td>2.264</td>
<td>2.295</td>
<td>**</td>
<td>172</td>
<td>[163]</td>
</tr>
</tbody>
</table>

a. angle between the gold-coordinated atom of the supporting ligand, the gold atom, and the centroid of the alkene C=C bond. b. meaningful analysis of bond distance was precluded by orientational disorder of the diene. * weighted average of two conformers. ** not reported.
Among the collected measurements, the only complex (ligand-Au-alkene) that is repeated is IPr-Au-2,3-dimethyl-2-butene (IPr-Au-2b), which was characterized both computationally and experimentally. In the computational study, the investigators used the BLYP GGA functional with the TZ2P polarized triple-zeta basis set for all atoms, ZORA for relativistic effects, and implicit solvent modeling via COSMO using the dielectric of CH₂Cl₂.

It is evident from the data that this particular computational model is inadequate for accurate modeling of this gold(I)-alkene complex, yielding bond distances that are systematically longer than those found in the XRC structure (Δd(C=C) 0.056 Å, Δd(Au–C, average) 0.111 Å). The 4–5% relative error in bond length is well outside the accepted “chemical accuracy” goal of within a few picometers. This large discrepancy was acknowledged by the authors of the computational investigation—who then go on to remark that the Au–C bond distance is a crucial factor for understanding the nature of the gold(I)-alkene complex—but the results were nevertheless published. Is the computational model, then, considered to describe the gold(I)-alkene bonding “well enough” to pass the peer-review process?

By comparing the experimentally-characterized complexes IPr-Au-2b and P1-Au-2b (Δd(C=C) 0.021 Å, Δd(Au–C, average) 0.058 Å; 2–3% relative error), we may obtain a sense of the differences in C=C and Au–C bond lengths that result from a change in the

---

*In the authors’ own words: “It is important to note that the Au–C distance is a crucial factor for the modulation of the nature of the bond in this type of gold(I)-alkene (alkyne) complexes. In particular, the metal-to-substrate π-back-donation component of the Dewar-Chatt-Duncanson model is most sensitive to the substrate distortion and its distance from gold(I).”*
supporting ligand. Such analysis also confirms just how poorly the computational model simulates these gold(I)-alkene complexes: the difference between the computational and experimental bond distances for the exact same complex is twice as large as the experimentally-measured ligand-influenced bond length difference for two similar complexes. A comparison of the XRC bond distances for other “ligand-swapped” experimental pairs (IPr-Au-2d vs. P2-Au-2d; IPr-Au-2e vs. P2-Au-2e) reveals similarly minor differences in bond lengths (1–3% relative error). How can we reliably utilize these kinds of computational methods to investigate such ligand effects, when the margin of error is so much larger than the anticipated difference?

The short answer is: until and unless the aforementioned “chemical accuracy” target can be demonstrably attained in the application of one or more computational methods to gold(I) chemistry, we should exercise a modicum of caution when interpreting such computational results.

3.2 Gold(I)-alkyne complexes

In contrast to the early work on gold(I)-alkene complexes, it was not until the late 1990s that the structural chemistry of corresponding alkyne complexes began to be developed. Preliminary accounts in this area include the oligomeric gold(I) alkynyl (RC≡CAu)n catenes by Mingos et al.,[351] the interaction of chelating diacetylenetitanium “pincers” with alkylgold by Lang and co-workers,[352-353] and the complexation of thiacycloalkynes to AuCl by Schulte and Behrens.[354] Besides bearing fairly little
resemblance to catalytically important gold(I)-alkyne complexes, these have all been well-covered in the review by Schmidbauer\cite{80} and so will not be discussed further here.

The first structurally-characterized gold(I) complexes of simple alkynes were not reported until 2009, independently by the groups of Dias\cite{290, 355} and Fürstner.\cite{356} Slightly earlier, in 2008, computational studies from the Echavarren group\cite{21} included the theoretical “naked” cationic gold species [(acetylene)Au]$^+$ and [(propyne)Au]$^+$, which are collected with similar examples in Table 6.

<table>
<thead>
<tr>
<th>complex</th>
<th>method</th>
<th>free C≡C</th>
<th>bound C≡C</th>
<th>Au–C1</th>
<th>Au–C2</th>
<th>ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(acetylene)Au$^+$</td>
<td>BP86/TZP/ZORA</td>
<td>1.205</td>
<td>1.25</td>
<td>2.135</td>
<td>2.135</td>
<td>[321]</td>
</tr>
<tr>
<td>(acetylene)Au$^+$</td>
<td>B3LYP/6-31G**/LANL2DZ</td>
<td>1.205</td>
<td>1.24</td>
<td>2.221</td>
<td>2.223</td>
<td>[21]</td>
</tr>
<tr>
<td>(propyne)Au$^+$</td>
<td>B3LYP/6-31G**/LANL2DZ</td>
<td>1.207</td>
<td>1.253</td>
<td>2.073</td>
<td>2.551</td>
<td>[21]</td>
</tr>
<tr>
<td>(cyclooctyne)Au$^+$</td>
<td>BP86/TZVPP</td>
<td>**</td>
<td>1.270</td>
<td>2.128</td>
<td>2.128</td>
<td>[357]</td>
</tr>
</tbody>
</table>

Although there are far fewer examples than for the [(ethylene)Au]$^+$ complexes discussed in the last section, similar trends are evident: the C≡C bond distance appears to be fairly consistent across compounds and computational methods, while the Au–C bond lengths exhibit more variation. Of particular note are the first two entries in Table 6, as they both describe the same model complex [(acetylene)Au]$^+$, yet differ by 0.094 Å in the Au–C bond length. This is similar to the range noted earlier for the Au–C bond length in [(ethylene)Au]$^+$ complexes, with the lowest and highest values again being produced by the BP86 and B3LYP functionals, respectively.

More data, both experimental and computational in origin, is available for
complexes (alkyne)-AuCl, as characterized by the groups of Dias,[290, 357] Fürstner,[356] Echavarren,[21] Nava,[89] and Belpassi & Tarantelli,[323, 332] the alkynes are depicted in Figure 10 and the measured bond distances and angles are summarized in Table 7.

![Figure 10: Alkynes employed in experimentally- and computationally-characterized gold(I) complexes.](image)

**Table 7: Calculated\(^8\) and experimental\(^†\) bond lengths (Å) and angles (°) for (alkyne)AuCl complexes.**

<table>
<thead>
<tr>
<th>alkyne</th>
<th>C≡C</th>
<th>Au–C1</th>
<th>Au–C2</th>
<th>Cl–Au</th>
<th>R–C≡C(^a)</th>
<th>ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>acetylene(^b)</td>
<td>1.24</td>
<td>2.09 (Au-C≡C midpoint)</td>
<td>**</td>
<td></td>
<td>163</td>
<td>[323]</td>
</tr>
<tr>
<td>acetylene(^c)</td>
<td>1.24</td>
<td>2.13</td>
<td>2.13</td>
<td>2.24</td>
<td>**</td>
<td>[89]</td>
</tr>
<tr>
<td>acetylene(^d)</td>
<td>1.24</td>
<td>2.14</td>
<td>2.14</td>
<td>2.25</td>
<td>164.6</td>
<td>[332]</td>
</tr>
<tr>
<td>acetylene(^e)</td>
<td>1.235</td>
<td>2.221</td>
<td>2.221</td>
<td>**</td>
<td>**</td>
<td>[21]</td>
</tr>
<tr>
<td>propyne(^e)</td>
<td>1.238</td>
<td>2.184</td>
<td>2.296</td>
<td>**</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>4a(^f)</td>
<td>1.224</td>
<td>2.152</td>
<td>2.172</td>
<td>2.270</td>
<td>166.9</td>
<td>[290]</td>
</tr>
<tr>
<td>4a(^g)</td>
<td>1.247</td>
<td>2.206</td>
<td>2.231</td>
<td>2.304</td>
<td>165.7</td>
<td>[290]</td>
</tr>
<tr>
<td>4d(^g)</td>
<td>1.256</td>
<td>2.130</td>
<td>2.130</td>
<td>2.262</td>
<td>**</td>
<td>[357]</td>
</tr>
<tr>
<td>4e(^f)</td>
<td>1.224</td>
<td>2.064 (Au-C≡C midpoint)</td>
<td>**</td>
<td></td>
<td>165</td>
<td>[356]</td>
</tr>
<tr>
<td>4e(^g)</td>
<td>1.256</td>
<td>2.154</td>
<td>2.157</td>
<td>2.263</td>
<td>163.1</td>
<td>[356]</td>
</tr>
</tbody>
</table>

\(^a\) angle between the alkynyl carbons and the connected (R = C or H) atoms; i.e., deviation from linearity.
\(^b\) BLYP/TZ2P/ZORA  
\(^c\) CCSD(T)/def2-QZVPP  
\(^d\) DF-LCCSD(T)/def2-TZVPP  
\(^e\) B3LYP/6-31G**/LANL2DZ  
\(^f\) B3LYP/ccVDZ (C, H)/aug-cc-pVDZ (Cl)/LANL2DZ (Au)  
\(^g\) BP86/def2-TZVPP  
\(^\text{** not reported.}\)

The same trends observed thus far continue to hold; very little variation is exhibited in the C≡C distances, while the Au–C bond lengths cover a range of 0.091 Å
(higher if the large Au–C2 distance for the propyne complex is included). The calculated C≡C bond lengths in the complexes 4d-AuCl and 4e-AuCl merit mention here, since they are quite a bit longer than the corresponding distances in both the computational (acetylene)AuCl and the experimental 4e-AuCl complexes. Of note is that the calculations for (4d, 4e)-AuCl utilized the BP86 functional, which was shown previously to yield Au–C distances that are too short; indeed, these complexes also follow this trend. Considering that the Dewar-Chatt-Duncanson model predicts just such an inverse relationship between the C-C and M-C bond lengths for metal-coordinated unsaturated species, this computational trend is at least somewhat reassuring.

Examples of crystallographically characterized cationic gold-alkyne complexes utilizing NHC or phosphine supporting ligands are less common than their alkene counterparts, but have an advantage in that most of the investigators in this area have also conducted DFT calculations on the same complexes, allowing for a more direct comparison. The available published data are summarized below in Table 8.
Table 8: Calculated§ and experimental† bond lengths (Å) and angles (°) for NHC- and phosphine-based L-Au(alkyne) complexes.

<table>
<thead>
<tr>
<th>alkyne</th>
<th>L</th>
<th>C≡C</th>
<th>Au–Cl</th>
<th>Au–C2</th>
<th>L–Au</th>
<th>R–C≡C</th>
<th>ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>4b§,b</td>
<td>IPr</td>
<td>1.24</td>
<td>2.27</td>
<td>2.29</td>
<td>2.04</td>
<td>166</td>
<td>[358]</td>
</tr>
<tr>
<td>4b§,b</td>
<td>P4</td>
<td>1.24</td>
<td>2.31</td>
<td>2.29</td>
<td>2.32</td>
<td>167</td>
<td>[358]</td>
</tr>
<tr>
<td>4c†</td>
<td>P2</td>
<td>1.220</td>
<td>2.238</td>
<td>2.239</td>
<td>**</td>
<td>166, 168</td>
<td>[359]</td>
</tr>
<tr>
<td>4d†</td>
<td>P3</td>
<td>1.213*</td>
<td>2.207*</td>
<td>2.193*</td>
<td>2.315*</td>
<td>157*</td>
<td>[357]</td>
</tr>
<tr>
<td>4d§,c</td>
<td>P3</td>
<td>1.245</td>
<td>2.222</td>
<td>2.222</td>
<td>2.379</td>
<td>**</td>
<td>[357]</td>
</tr>
<tr>
<td>4d†</td>
<td>6-DIPP</td>
<td>1.213</td>
<td>2.199</td>
<td>2.201</td>
<td>2.022</td>
<td>157</td>
<td>[357]</td>
</tr>
<tr>
<td>4d§,c</td>
<td>6-DIPP</td>
<td>1.244</td>
<td>2.202</td>
<td>2.202</td>
<td>2.043</td>
<td>**</td>
<td>[357]</td>
</tr>
<tr>
<td>4e†</td>
<td>IPr</td>
<td>1.206</td>
<td>2.142 (Au–C≡C midpoint)</td>
<td>**</td>
<td>168</td>
<td>[356]</td>
<td></td>
</tr>
<tr>
<td>4e§,c</td>
<td>IPr</td>
<td>1.246</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>163</td>
<td>[356]</td>
</tr>
<tr>
<td>4e†</td>
<td>6-DIPP</td>
<td>1.218</td>
<td>2.14 (Au–C≡C midpoint)</td>
<td>**</td>
<td>160</td>
<td>[356]</td>
<td></td>
</tr>
<tr>
<td>4e§,c</td>
<td>6-DIPP</td>
<td>1.247</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>166</td>
<td>[356]</td>
</tr>
</tbody>
</table>

a. angle between the alkynyl carbons and the connected (R = C or H) atoms; i.e., deviation from linearity.
b. BLYP/TZ2P/ZORA  c. BP86/def2-TZVPP  * average of two conformers.  ** not reported.

It is rather unfortunate for the present review that in both cases where the investigators augmented their experimental studies with computational analysis, the same DFT functional and basis set was used (BP86/def2-TZVPP); more unfortunate is the fact that this particular method has already been shown to overbind Au-(alkene) and Au-(alkyne) complexes, resulting in shortened Au–C and lengthened C–C bonds. This trend, however, is only evident here in the discrepancies for the C≡C distances; the calculated Au–C bond lengths manage to align exceptionally well with the XRC measurements. This is fairly discouraging: if a computational method is consistently high or low in its predictions, at least it can be corrected for; such does not seem to be the case for BP86 as applied to gold(I)-alkyne complexes.
The computationally-modeled complexes IPr-Au-4b and P4-Au-4b can be contrasted with the experimentally (and computationally) characterized complex 4a-AuCl (4a = 3-hexyne; 4b = 2-hexyne). The C≡C bond distances are comparable: 1.224 Å (experimental) versus 1.24–1.247 Å (calculated). The Au–C measurements are longer for the calculated IPr-Au-4b and P4-Au-4b (2.29 Å average) than for the calculated 4a-AuCl (2.22 Å average). Interestingly, the calculations on the 4b complexes predict a nearly symmetrical binding of 2-hexyne to gold, while those for the 4a complex indicated a more asymmetrical binding (Au–C1: 2.206, Au–C2: 2.231 Å) of 3-hexyne to gold. This goes somewhat against the intuition that the more symmetrical substrate (in this case, the 3-hexyne) should bind more symmetrically, especially if there is no steric influence from the supporting ligand (Cl in this case). However, XRC measurements of 4a-AuCl indeed reveal an asymmetrical binding (Au–C1: 2.152, Au–C2: 2.172 Å), although less pronounced than that predicted. Interestingly and in contrast, the experimentally characterized complex P2-Au-4c (4c = 4,4-dimethyl-2-pentyne) displayed symmetrical coordination to gold, with no significant lengthening of the C≡C bond relative to free 4c. Whether this unexpected symmetrical coordination is due to electronic effects from the alkyne itself, or a result of steric influence by the bulky tri-tert-butylphosphine ligand, is unclear; in the absence of experimental data for similar compounds, a much more rigorous computational study could possibly offer more insight into this result.

3.3 Summary and conclusions

Although the various computational models examined here (as well as others not
mentioned) may prove to be valuable tools for investigating gold(I) complexes, care
must be taken in both the choice and application of computational methods as well as
the interpretation of computational results. Other authors have bemoaned the lack of
validation of computational methods being applied to studies of gold chemistry and we
are in agreement, inasmuch as our comparison of experimentally versus
computationally determined bond distances in gold(I) π-complexes reveals some
disconcerting trends.

We have identified some fairly large discrepancies in calculated bond lengths,
both as compared to experimental measurements and as compared to results from other
computational methods. Even considering the extremely simplistic model complex
[(ethylene)Au]^+ , bond length estimates yielded by different computational methods
range from 1.371–1.414 Å for C=C (a range of 0.043 Å) and from 2.099–2.253 Å for Au–C
(a range of 0.154 Å). Among the two most commonly utilized DFT functionals, B3LYP
and BP86, the data suggest that B3LYP tends to overestimate Au–C bond lengths, while
BP86 tends to underestimate; however, it was also seen that BP86 does not seem to
produce results which are systematically high or low, thus precluding the possibility of
applying corrections. More troubling yet are the BLYP results for IPr-Au-2b, which
exhibit a 4–5% relative error as compared with XRC measurements for the same
complex. It is hoped that these observations will encourage additional validation and
thoughtful selection of computational methods as applied to gold chemistry.
4. Computationally-derived mechanisms of gold(I)-catalyzed reactions: the need to benchmark

Computationally calculated energies, in general, tend to be more erroneous than calculated geometries, especially when attempting to model transition states. Unfortunately, accurate calculations of such energies are crucial for evaluating prospective mechanisms, particularly in situations common to transition-metal catalyzed reactions, where there is a bifurcationa along the potential energy surface (PES). Such scenarios have been described for organic reactions[360] and have recently been invoked in descriptions of gold-catalyzed processes as well.[361-364] Given that proposed mechanistic pathways and transition states are entirely dependent on the shape of the calculated PES, a verification within “chemical accuracy”—that is, absolute error in calculated energies of less than 1 kcal/mol—is absolutely crucial to ensure valid conclusions.

In the field of homogeneous gold(I) catalysis, the enyne and propargylic ester rearrangement manifolds are two of the best-studied. Gold complexation to the alkyne fragment can lead to a series of intermediates that open the way to very rich reactivity, often being the starting point of a reaction cascade that easily generates complexity, with selectivity able to be tuned via supporting ligand, substrate substitution, counterions, and solvent. Clearly, with so many variables in play, mechanistic knowledge of these

---

a. Bifurcations arise when there are sequential transition states with no intervening energy minimum. For such systems, the shape of the potential energy surface and dynamic effects, rather than transition-state energetics, control selectivity.[360]
systems is invaluable, not only for permitting ease-of-use in routine synthetic work, but also for engineering new gold-catalyzed reactions. However, the short lifespan of many of the intermediates, along with the complex reaction pathways, make the experimental study of such mechanisms a difficult task, and ideally, computational chemistry methods would allow for characterization of intermediates and differentiation between competitive reaction mechanisms. Alas, small variations between methods in calculated reaction energies and barriers can sometimes lead to a dramatic reversal of product distributions, highlighting the need for highly-accurate computational methods.

One of the earliest computational studies on a full gold(I) catalytic cycle was that by Cavallo and co-workers on the gold-catalyzed reactions of propargylic esters,[965] specifically the mechanism by which the known 1,2- and 1,3-acyl migrations occur (Figure 11). The authors proposed, on the basis of DFT calculations, that the three Au-activated species (I, II, III) can interconvert easily, forming what they term “a golden carousel”. Truncated ligands PMe₃ and 2,3-dimethylimidazol-2-ylidene (IMe) were utilized in the computations. The BP86 functional was used with a TZVP basis set and SDD effective core potential, and solvent effects were accounted for via single-point calculations on the gas-phase optimized geometries, using the IEF-PCM solvation model with the CH₂Cl₂ dielectric constant.

The results of the above computational study indicate that this reaction is characterized by numerous highly competitive pathways, but also that the high reactivity of starting structure I (Figure 11) is at the origin of this diversity.
Figure 11. Top: Key intermediates in the gold(I)-catalyzed reactions of propargylic esters. Bottom: Ligand effects on product ratios in the cycloisomerization of 6. See reference [366].
The gold-coordinated alkyne species can undergo both 1,2 and 1,3-shift of the ester moiety, leading to the Au-carbene species II and to the Au-allene species III, respectively. All species are connected in the “golden carousel” catalytic cycle, shown in Figure 12 below.

Figure 12: Calculated energies (relative to intermediate I) associated with the “golden carousel”. Round/square brackets correspond to L = IMe and PMe, respectively; numbers adjacent to arrows represent transition states. Adapted from reference [365].

The main weakness of this study, other than the use of a DFT functional that is potentially unsuitable for modeling gold chemistry, is the omission of possible transition
state structures. Both of these issues were subsequently tackled in a thorough
benchmarking study by Faza and co-workers,[368] which used the “golden carousel” as a
reference set but also explicitly included transition structures (Figure 13). The reference
energies and geometries were obtained at the CCSD/def2-TZVPP//CCSD/def2-SVP level,
and deviations from this reference are highlighted in Table 9.

![Diagram of chemical structures](image)

Figure 13: The stationary points in Correa’s “golden carousel” that were used by Faza et al. to
benchmark the performance of different computational methods. See refs. [96], [368]

From the results of this benchmarking study, several conclusions were drawn
that led the authors to propose a general and routine method for modeling gold-
catalyzed reactions. DFT methods clearly outperform MP2; the small gain in accuracy
that sometimes obtained with MP2 is completely offset by the much larger
computational cost. In terms of geometries, the different generations of functionals in
Table 9: Energy deviations (kcal/mol) for stationary points in Figure 13, relative to CCSD/def2-TZVPP//CCSD/def2-SVP reference. Cells marked “dnc” indicate that the energy calculation failed to converge. Adapted from Faza et al., ref. [368]

<table>
<thead>
<tr>
<th></th>
<th>HF</th>
<th>MP2</th>
<th>LDA</th>
<th>BLYP</th>
<th>B3LYP</th>
<th>mPW1PW</th>
<th>B2PLYP</th>
<th>ωB97</th>
<th>M06</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.35</td>
<td>2.80</td>
<td>10.58</td>
<td>8.24</td>
<td>5.81</td>
<td>5.45</td>
<td>3.00</td>
<td>0.71</td>
<td>2.31</td>
</tr>
<tr>
<td>3 anti</td>
<td>3.12</td>
<td>1.36</td>
<td>0.17</td>
<td>3.48</td>
<td>1.57</td>
<td>1.09</td>
<td>1.14</td>
<td>1.41</td>
<td>2.12</td>
</tr>
<tr>
<td>3 syn</td>
<td>2.75</td>
<td>0.80</td>
<td>0.07</td>
<td>3.55</td>
<td>1.53</td>
<td>1.17</td>
<td>0.94</td>
<td>1.70</td>
<td>1.96</td>
</tr>
<tr>
<td>4 anti</td>
<td>5.98</td>
<td>5.37</td>
<td>dnc</td>
<td>2.89</td>
<td>1.95</td>
<td>0.18</td>
<td>2.91</td>
<td>0.42</td>
<td>1.11</td>
</tr>
<tr>
<td>4 syn</td>
<td>5.84</td>
<td>5.20</td>
<td>dnc</td>
<td>2.97</td>
<td>2.06</td>
<td>0.20</td>
<td>3.00</td>
<td>0.43</td>
<td>1.03</td>
</tr>
<tr>
<td>5</td>
<td>1.45</td>
<td>0.47</td>
<td>0.27</td>
<td>4.55</td>
<td>1.88</td>
<td>0.66</td>
<td>0.48</td>
<td>1.93</td>
<td>2.70</td>
</tr>
<tr>
<td>6</td>
<td>0.32</td>
<td>4.26</td>
<td>4.69</td>
<td>1.29</td>
<td>1.14</td>
<td>0.79</td>
<td>1.55</td>
<td>0.48</td>
<td>0.08</td>
</tr>
<tr>
<td>7</td>
<td>4.13</td>
<td>2.50</td>
<td>0.82</td>
<td>3.26</td>
<td>1.44</td>
<td>1.15</td>
<td>1.43</td>
<td>1.31</td>
<td>2.05</td>
</tr>
<tr>
<td>8</td>
<td>5.42</td>
<td>4.45</td>
<td>dnc</td>
<td>1.51</td>
<td>0.81</td>
<td>0.74</td>
<td>1.99</td>
<td>0.64</td>
<td>0.58</td>
</tr>
<tr>
<td>ts23 anti</td>
<td>2.56</td>
<td>3.10</td>
<td>7.62</td>
<td>1.26</td>
<td>0.84</td>
<td>1.32</td>
<td>1.24</td>
<td>1.26</td>
<td>0.05</td>
</tr>
<tr>
<td>ts23 syn</td>
<td>1.18</td>
<td>0.71</td>
<td>4.01</td>
<td>0.92</td>
<td>0.36</td>
<td>0.02</td>
<td>0.28</td>
<td>0.62</td>
<td>1.67</td>
</tr>
<tr>
<td>ts25</td>
<td>0.16</td>
<td>0.01</td>
<td>8.17</td>
<td>2.44</td>
<td>1.88</td>
<td>2.21</td>
<td>0.97</td>
<td>0.64</td>
<td>0.46</td>
</tr>
<tr>
<td>ts34 anti</td>
<td>2.05</td>
<td>0.42</td>
<td>dnc</td>
<td>5.70</td>
<td>3.63</td>
<td>2.67</td>
<td>0.95</td>
<td>0.21</td>
<td>2.33</td>
</tr>
<tr>
<td>ts34 syn</td>
<td>2.04</td>
<td>0.47</td>
<td>dnc</td>
<td>5.74</td>
<td>3.60</td>
<td>2.66</td>
<td>0.87</td>
<td>0.09</td>
<td>2.19</td>
</tr>
<tr>
<td>ts56</td>
<td>1.08</td>
<td>1.07</td>
<td>1.79</td>
<td>2.29</td>
<td>1.97</td>
<td>1.83</td>
<td>1.44</td>
<td>0.86</td>
<td>1.86</td>
</tr>
<tr>
<td>ts67</td>
<td>5.57</td>
<td>5.08</td>
<td>0.74</td>
<td>2.88</td>
<td>1.29</td>
<td>1.14</td>
<td>1.42</td>
<td>0.21</td>
<td>1.43</td>
</tr>
<tr>
<td>ts78</td>
<td>0.58</td>
<td>0.56</td>
<td>dnc</td>
<td>3.12</td>
<td>1.40</td>
<td>0.40</td>
<td>0.37</td>
<td>0.47</td>
<td>0.92</td>
</tr>
<tr>
<td>ts84</td>
<td>4.45</td>
<td>4.38</td>
<td>12.21</td>
<td>0.87</td>
<td>0.98</td>
<td>2.68</td>
<td>0.90</td>
<td>2.87</td>
<td>1.67</td>
</tr>
</tbody>
</table>

"Jacob’s ladder" showed systematic improvement, with LDA offering very poor results, and GGA and above significantly improving them. HF and MP2 proved to be more accurate than most GGA and meta-GGA functionals, and so it is not surprising that the largest improvement in DFT performance comes by way of the hybrid functionals as a result of the introduction of HF energy. The inclusion of dispersion corrections further reduces the error in both geometries and energies; the best overall results were obtained with the B2PLYP double hybrid functional, with the significantly less costly ωB97 and M06 functionals close behind. For energies, these three functionals displayed maximum deviations smaller than 3 kcal/mol with respect to the reference values. It should be
noted that the deviations in geometries and energies are not systematic along the entire reaction path; problematic species typically correspond to the initial $\eta^2$-coordinated gold-alkyne species, and structures with a large electron delocalization.

With the best methods for geometry optimizations established, the performance of different ECPs were assessed, with LANL2DZ and SDD yielding larger errors than the fully relativistic ecp-60-mwb, which offered results comparable to those obtained with the DKH method. Computational models for solvation effects were also investigated, with the Polarizable Continuum Model (PCM)\(^b\) introducing energy deviations of the same order or less than that resulting from variations in the effective core potential (ECP) or basis set. This indicates that careful selection of basis set and/or ECP may be more critical to high accuracy than (approximate) inclusion of solvent effects—at least for the reaction studied.

The use of CCSD as the reference for energies in the above benchmarking study was criticized by Nava et al.,\(^{[89]}\) since the systematic inclusion of triple excitations is deemed important to account for long-range dispersion effects. The original benchmarking study was later revisited by the original authors at a higher level of theory,\(^{[96]}\) with the observation that the relative energies of the transition states and intermediates barely changed. Their conclusion was that density functionals that perform well with respect to CCSD/def2-TZVPP data also do well when higher-level DLPNO-CCSD(T)/def2-QZVPP results are used as the standard. (The authors agreed,\(^{[89]}\)

\(^b\) PCM models solvation effects as a polarizable continuum, rather than as individual solvent molecules.
however, that the inclusion of triple excitations is highly recommended.) Similar results were obtained by Belpassi et al.[332] in a benchmarking study on the gold-catalyzed hydroamination of alkynes.

One question that has not been adequately addressed in such benchmarking studies is how well the reference data, calculated at high levels of theory, models reality. The need to resort to computational modeling is, of course, due to the fact that many of the intermediates and transition states considered in the benchmarking sets cannot be experimentally isolated and/or characterized. On the other hand, many recent benchmarking studies (such as the one by Belpassi and coworkers[332]) use severely truncated complexes in the models, which are certainly unrepresentative of species encountered along a given gold-catalyzed reaction pathway. Given the continual advances in both computational methods and computational power, one would hope to see more realistic models being increasingly applied in such benchmarking studies.

Some general conclusions about gold-catalyzed mechanisms can be extracted from computational studies. The first is that, for a given set of reactants and catalysts, there are usually several relatively low-energy reaction pathways available. It is not uncommon for such pathways to involve many different, often interconnected, steps where gold changes its coordination; hence, the availability and utilization of superbly accurate computational methods is crucial for unambiguous interpretation of calculated energies and mechanistic pathways.

The second conclusion, associated with the first, is the fact that homogeneous
gold catalysis is easily influenced by additional factors including solvent, counterion, and the structure of the catalyst and the substrate. Counterions and solvent in particular are often involved in proton relays, which can dramatically alter the preference for one reaction path over another, making the explicit consideration of solvent and counterions—often neglected in these kinds of studies—a necessity. Additionally, given the known effects that may be engendered by differences in steric and electronic effects imparted by both the supporting ligand and substrate, it would seem essential to employ exact, non-truncated models in computational investigations of any gold(I)-catalyzed mechanism.
5. Final thoughts and closing remarks

Early computational studies on gold-catalyzed reactions utilized DFT functionals (such as B3LYP and BP86) and basis sets that were known at the time to provide consistently good results at low cost for organic reactivity. As time advanced larger basis sets and more modern functionals became more common, but the main problem behind the choice of computational method remained—they had not been explicitly tested against organometallic gold systems, although in some papers a small-scale benchmarking study or comparison with the results of higher-level ab initio calculations (most commonly just MP2) was carried out. In this context, the 2009 finding by Benitez et al.\cite{Benitez2009} that the M06 functional correctly reproduced the influence of different ligands on the experimental bond rotation barriers of gold carbenoid species, while B3LYP and BP86 failed at this task, should have provided a warning sign that the choice of computational methodology should be approached with care—especially when applying methods that have not been developed or tested with the target system in mind. Nevertheless, B3LYP and BP86 continue to be utilized in mechanistic studies of gold catalysis as late as 2015.\cite{Roesch2014, Roesch2015}

A clear example of what can go wrong in a computational mechanistic study when the choice of method has not been carefully considered is recounted by Faza and López.\cite{Faza2014} A 2014 study by Rösch et al.\cite{Roesch2014} sought to calculate the reaction profiles corresponding to the four mechanisms proposed for a rhodium-catalyzed hydrosilylation of ethylene—a fairly complex system, since the four mechanisms are
connected at some intermediates and the potential energy profiles of some of the alternative pathways are relatively flat overall. The authors used a stepping approach, first optimizing all four profiles with B3LYP and BP86, then further optimizing two to three relevant barriers in each mechanism with an additional set of functionals, in an attempt to identify the rate-limiting step in each, which would then be evaluated with seven more functionals. The most relevant result from this study is that the barrier heights and step at which they occurred changed depending on the functional, with enthalpy differences of up to 10 kcal/mol. Although the reaction involved is not catalyzed by gold, the warning remains valid, especially given the similarly facile interconversion of intermediates in most gold-catalyzed reaction manifolds.

Given the difficulty in isolating and characterizing intermediates experimentally, computational chemistry remains an necessary tool for exploring the plethora of mechanistic pathways available for a given chemical transformation. Because of the activation power of gold, the energy barriers of most, if not all, of the steps involved in the diverse catalytic pathways are often very low, allowing for this reactivity to occur under very mild conditions. However, this double-edged sword of low energy barriers and facile interconversions, and the flat energy profile that they commonly exhibit, requires the cautious and prudent application and interpretation of extremely accurate and robust computational methodology in the search for mechanistic insight.
References


Z. Li, C. Brouwer, C. He, "Gold-catalyzed organic transformations". *Chem. Rev.* **2008**, *108*(8), 3239-3265.


63


R. S. Mulliken, Spectroscopy, molecular orbitals, and chemical bonding, Laboratory of Molecular Structure and Spectra, University of Chicago, 1967.


P. a. M. Dirac, "Quantum mechanics of many-electron systems". Proceedings of the Royal Society of London Series a-Containing Papers of a Mathematical and Physical Character 1929, 123(792), 714-733.


[196] Jensen, section 4.2.2, p141; Cramer, section 7.2.3, p211.

[197] Lewars, section 5.4.3.1, p276; Jensen, section 4.5, p153; Young, chapter 26, p223; Bachrach, section 1.2.2, p16.


[200] Cramer, section 7.4, p221; Jensen, section 4.8, p163; Bachrach, section 1.2.3, p17.


[208] Lewars, p275; Cramer, p226;


R. A. Wheeler, *Annual reports in computational chemistry*, Elsevier, **2012**.


Johnson 2015, p10-18

Cramer, sect 8.5.5


S. Guo, J. C. Bernhammer, H. Han Vinh, "1,2,4-triazole-derived carbene complexes of gold: Characterization, solid-state aggregation and ligand disproportionation". Dalton Trans. 2015, 44(34), 15157-15165.


