Showcasing research from the Group of Prof. Ignacy Cukrowski at Pretoria University, South Africa

Reliability of HF/IQA, B3LYP/IQA, and MP2/IQA data in interpreting the nature and strength of interactions

Fundamental understanding of a chemical change, e.g., adduct, cluster and complex formation, a reaction mechanism, etc., through changes in the properties of and interactions between poly-atomic fragments is of special interest. This work shows that fragment attributed molecular system energy change (FAMSEC)-based B3LYP/IQA pictures are exact and, quantitatively, perfectly comparable with CCSD/BBC1/IQA data.

As featured in:

Reliability of the HF/IQA, B3LYP/IQA and MP2/IQA data in interpreting the nature and strength of interactions

Ignacy Cukrowski

*Department of Chemistry, Faculty of Natural and Agricultural Sciences, University of Pretoria, Lynnwood Road, Pretoria, 0002 South Africa*

*E-mail: ignacy.cukrowski@up.ac.za*

**Keywords**

IQA, HF, B3LYP, MP2, FAMSEC, AIMAll, interactions

**Abstract**

With main focus on the B3LYP level, reliability of level of theory (LoT) dependent IQA-defined energy terms (self- and additive atomic energies, interaction energy and its components) in interpreting interactions was investigated at three LoT using default settings in AIMAll software with explicitly implemented the actual B3LYP exchange-correlation functional. Reliability was ‘quantified’ using relative errors (REs) defined as, e.g. \( \text{RE} = \text{B3LYP/IQA(computed)} - \text{B3LYP/IQA(expected)} \), using reference CCSD/BBC1/IQA data to obtain the LoT/IQA(expected) terms. On average, B3LYP produced most accurate IQA energies among LoTs investigated with REs being an order of magnitude smaller than those at HF. B3LYP/IQA-description of the O4⋯H6 and O3⋯O4 interactions in glycol conformers compared well with the CCSD/BBC1/IQA-generated picture. Exceptionally reliable data were obtained at B3LYP for changes in the IQA energies computed for structural changes of glycol. The FAMSEC-based interpretation produced exact qualitative description and, quantitatively, perfectly comparable values with CCSD/BBC1/IQA data. It is documented that the \( \Delta E_{\text{IQA}} = \Delta E \) criterion (representing changes between the final and any suitably selected initial structure of a molecular system) ‘validates’ or predicts usefulness of changes in the LoT/IQA computed energy terms in terms of interpretability of interactions. This is supported by (1) smallest errors in the IQA energy changes at MP2/Müller even though \( E_{\text{IQA}} \) largely overestimated \( E \) (by \(-170\) kcal mol\(^{-1}\) due to large REs in self-atomic energies) and (2) reasonable HF-generated FAMSEC descriptors, regardless of largest REs in the HF/IQA data. Finally, adding Grimme’s D3 empirical dispersion correction had insignificant impact on REs.
1. Introduction

The concept and physical nature of chemical bonds is at the centre of Chemistry and it is being investigated for many decades. Arguably, the interacting quantum atoms (IQA) method\(^3\) is the most powerful tool (at least among methods under the Quantum Chemical Topology heading\(^4,5\)) in the study of chemical bonds as well as inter- and intramolecular interactions on a fundamental level. This is because the molecular energy is recovered by its partitioning to two main energy components, self-atomic energies, \(E_{\text{self}}^A\) and diatomic interaction energies between each unique pair of atoms A and B, \(E_{\text{int}}^{A,B}\). The latter term is particularly useful in characterizing the nature of an interaction as it is made of a classical term, \(V_{\text{el}}^{A,B}\), used to quantify Coulombic component of an interaction and the exchange-correlation (XC) term, \(V_{\text{XC}}^{A,B}\), describing the degree of covalency of the interaction (or a bond). The two terms, \(E_{\text{self}}^A\) and \(E_{\text{int}}^{A,B}\), and their components, are of a fundamental significance as they provide an invaluable direct theoretical insight on how properties of atoms as well as the strength and nature of interactions/bonds vary with a change in a chemical environment due to, \textit{e.g.} a conformational transformation, additional functionality in a molecule or along a reaction pathway.

\textit{Exact} IQA-defined energy components can be obtained at any level of theory (LoT) when the two-electron density matrix (2EDM) is used to compute \(V_{\text{ee}}\) energy contributions (intra- and inter-atomic, \(V_{\text{ee}}^A\) and \(V_{\text{ee}}^{A,B}\), respectively) as well as (de)localisation indices. Unfortunately, 2EDM is well defined only for variational CI-based approaches, such as CASSCF or CISD and, consequently, IQA is ‘restricted’ to a few-atom molecular systems due to exorbitant computational cost. For instance, importance of dynamic correlation and the explicit evaluation of the electron–electron contribution to the electronic energy as well as numerous IQA-defined energy terms was reported recently on very small (few atoms) molecular systems.\(^6-8\) Clearly, it would be advantageous to extend the use of IQA to real-life, hence much larger molecules
(among them of biological importance), by making use of wavefunctions generated at significantly lower LoT. To achieve that, appropriate approximations for 2EDM, based on one-electron density matrix (1EDM), must be implemented. To this effect, the Müller functional is implemented in most commonly used dedicated software for topological calculations, AIMAll. It approximates 2EDM in terms of natural orbitals of the one-electron density matrix and thus it allows calculating 2EDM-dependent properties. We have shown recently, that in the case of widely used perturbative methods (such as MP2) much more accurate IQA-defined energy terms can be obtained when proposed by Gritchenko et al. correction of the Müller functional is used, commonly known as BBC1. Among three levels of approximation studied, BBC1, BBC2 and Müller, the latter performed the worst when accuracy of the IQA data is considered.

Although the MP2/BBC1 data compared with those obtained at the CCSD/BBC1 level very well, it is still extremely expensive when larger molecular systems (10 < No. of atoms < 50; an arbitrarily selected range) are to be investigated by IQA. The B3LYP density functional is most commonly used in the computational studies of larger molecules (e.g. of biological relevance). Furthermore, the B3LYP and M06-2X DFT density functionals, due to computational cost (affordability) and their explicit implementation in AIMAll, became methods of choice in many applications (and references therein). Implementation (in AIMAll) of the actual exchange-correlation functional of the corresponding DFT model resulted in unambiguous IQA additive atomic energies that (i) when summed up, the electronic energy of a system (molecule) is fully recovered in the gas phase and (ii) showed a high correlation with the experimental rate constants and with the complexation energies.

Although obvious, it is important to stress that IQA descriptors computed from wavefunctions generated at different LoT differ in values. This is nicely illustrated in recent studies where a comparative analysis of numerous IQA descriptors coming from IQA decompositions performed at HF, MP2, CISD, CCSD, CCSD(T) and almost full configuration interaction (FCI) approximations. In this case the reference data were obtained from DMRG (density matrix
renormalization group) wavefunctions and their associated density matrices, approaching the FCI level with typical errors smaller than 1 kcal mol\(^{-1}\).  

Clearly, IQA data computed from DFT wavefunctions must differ from those obtained at much higher levels of theory that, in principle, are not feasible in daily use in the study of large molecular systems due to computational time required not only for electronic structure optimisations but also for IQA calculations. The most convenient approach would be to use, e.g., B3LYP for optimisation and subsequent IQA calculations provided that systems of interest are ‘suitable’ for the studies at the DFT level. Hence, with the main focus on B3LYP, the aim of this work is a thorough investigation of reliability of selected B3LYP/IQA descriptors that are very useful and most frequently used in interpreting and characterizing interactions, namely:

(a) Diatomic interaction energy, \(E_{\text{int}}^{A,B}\), and its components, \(V_{\text{el}}^{A,B}\) and \(V_{\text{XC}}^{A,B}\).

(b) Self-atomic and additive atomic energies, \(E_{\text{self}}^A\) and \(E_{\text{add}}^A\), respectively.

Assessing reliability of the above specified IQA-defined terms (computed from the B3LYP wavefunction) comes down to evaluating suitability of B3LYP for the study of interactions through the exact IQA energy partitioning scheme. To achieve that, a suitable set of IQA reference data is required. To this end, the B3LYP/IQA data will be compared against the CCSD/BBC1/TWOe/refined settings data that, as we have shown recently, is able to recover molecular energy well and, even more importantly, the computed \(V_{\text{ee}}\) terms reproduced the exact values to a small fraction of kcal mol\(^{-1}\). It is important to stress that molecules of interest in this work were, on purpose, fully energy optimised at each level of theory instead of performing single point calculations (SPC) on CCSD structures. As a consequence, atomic basins (their volumes and interatomic surfaces) and density distributions at each level of theory must be somewhat different. This is of no concern in this work as our goal is to assess how B3LYP/IQA data differ from CCSD/IQA data and how impactful the differences in the IQA descriptors are in interpreting interatomic interactions. Performing SPC would assure that structural integrity
obtained at higher level (3D placement of atoms would be preserved at DFT) but density distribution throughout a molecule would be that as calculated from a DFT wavefunction. One must remember that to obtain meaningful structures at DFT one might need to account for dispersion by, e.g. making use of Grimme’s empirical correction, as is the case also in this study.

Moreover, we expanded our study to data obtained from wavefunctions generated at the HF and MP2/Muller levels using default (auto) settings for all the IQA calculations in AIMAll software (there are many reasons behind this selection of LoTs and they will become obvious in sections that follow). One must note that by selecting full IQA calculations in AIMAll, the entire set of QTAIM\textsuperscript{21}-defined properties is computed as well.

Furthermore, it was also of importance to find out whether the selected IQA descriptors obtained at the three levels of theory under examination could be used for comparative analyses. To this effect, we will examine computed changes in selected energy terms obtained for an \( n \)-atom \((2 \leq n \leq 4)\) molecular fragment \( G \), among others:

(a) A self-fragment, \( \Delta E^G_{\text{self}} \), and fragment’s additive, \( \Delta E^G_{\text{add}} \), energy change,

(b) An intra-fragment interaction energy change, \( \Delta E^G_{\text{int}} \), and

(c) A change in the interaction energy between (i) an atom A of a fragment \( G \) with remaining atoms of a molecule, \( \Delta \sum_{X \neq A} E^X_{\text{int}} \), and (ii) all atoms of fragment \( G \) and remaining atoms of a molecular system, \( \text{i.e. fragment } \mathcal{H}, \Delta E^G_{\text{int} \mathcal{H}} \),

on a reference to final \((\text{ref } \rightarrow \text{fin})\) structural change of a molecular system. These energy terms are used in the fragment attributed molecular system energy change (FAMSEC) method.\textsuperscript{22} The FAMSEC method is making use of the IQA and IQF\textsuperscript{23} (interacting quantum fragments) concepts and its usefulness in the study of intramolecular interactions and they role in explaining relative stability of conformers has been demonstrated recently.\textsuperscript{11,24,25} A schematic representation of an
approach used to investigate LoT/AIMAll-generated IQA/IQF/FAMSEC data is presented in Scheme S1 in the ESI.

Three conformers of glycol (Fig. 1) are investigated in this work because we could make use of the CCSD/BBC1/TWOe data\textsuperscript{11} as a suitable reference for:

(1) A direct comparison (against reference data) of expected at lower LoTs IQA energy terms needed to evaluate relative errors (REs) in computed IQA energy terms for the same atoms, molecular fragments and interactions in different molecular environments (conformers) at lower LoTs, and

(2) Evaluating usefulness of FAMSEC-defined energy terms used to explain relative stability of these conformers. Hence, reliability in changes, relative to a reference structure (linear conformer, Lin) obtained on structural transformation to either the highest energy (eclipsed, Ecl) or lowest energy conformer, LEC, of glycol was investigated.

![Fig. 1](image1.png)

*Fig. 1* Ball-and-stick representation of the glycol conformers optimised at the CCSD/aug-cc-pVDZ level also showing atoms numbering and, relative to linear conformer, electronic energy differences (in kcal mol\textsuperscript{−1}) for the lowest energy (LEC) and eclipsed conformers of glycol.

We used Lin as a reference state of molecular system because it does not possess any specific intramolecular H-bonds or steric clashes. Classically, relative to Lin, the decrease in energy, by about −2.3 kcal mol\textsuperscript{−1} at CCSD, computed for the LEC could be attributed to the presence of intramolecular H-bond between O4 and H6. On the other hand, the increase of energy, by +9.2
kcal mol\(^{-1}\), found for \(E_{cl}\) at CCSD would be attributed to the steric clash between O3 and O4 atoms. It was then of great interest to find out how interpretation of relative stability of glycol conformers obtained from the LoT/AIMAll/auto-settings combination compares with that at CCSD/BBC1/TWOe with refined-settings.

Finally, we will also examine (i) how important a recovery of molecular energy from IQA calculations is; note that due to Müller approximation, \(E_{IQA}\) cannot recover \(E\) at the MP2 level, and (ii) an impact of Grimme’s empirical dispersion correction (GD3) on IQA descriptors as this correction is not accounted for in IQA calculations performed in AIMAll.

2. Computational methods

Optimisations, with a keyword ‘opt=verytight’, at the HF, B3LYP and MP2 levels were performed using Gaussian 09, revision D,\(^{26}\) which was also used to generate relevant wavefunctions. In case of B3LYP, optimisations were performed either without or with Grimme’s empirical dispersion correction with a keyword ‘empiricaldispersion=GD3’. The Dunning double zeta basis set, aug-cc-pVDZ, which is augmented by diffuse functions, was used throughout. One imaginary frequency has been present at all levels of theory only in the case of eclipsed form of glycol that was optimised with a dihedral angle \(\text{DA(O3,C2,C1,O4)}\) constrained to \(0^\circ\). Coordinates of optimised structures at all levels of theory are included in the ESI.

Topological and IQA analyses (HF, B3LYP, MP2) were conducted using AIMAll\(^{10}\) with default (auto) settings. This means that the Müller approximation was used in the case of MP2. The TWOe program, version 14, developed by Polestshuk\(^{27,28}\) was used to compute \(V_{ee}^{AA}\) throughout.

Finally, the IQA data used as a reference were obtained by optimizing glycol conformers at CCSD/aug-cc-pVDZ level with frozen core electrons and generating electron densities in the PRIRODA14 program of Laikov and Ustynyuk;\(^{29}\) relevant wavefunctions were submitted to the TWOe program, version 14. The BBC1 functional was used because it generates significantly
better quality IQA data when compared with the BBC2 and Müller approximations.\(^{11}\) To further improve the IQA data computed by use of the CCSD/BBC1 combination, the increased Lebedev grids were selected for angular integrations of oxygen atoms and the highest Lebedev and medium sized Gauss-Legendre angular grids have been chosen for C-atoms integrations. As a result, all atomic Laplacian \(|L(\Omega)|\) values have not exceeded \(5 \times 10^{-5}\) au.\(^{11}\) For comparison, the \(|L(\Omega)|\) values (at B3LYP) have not exceeded \(3 \times 10^{-5}\), \(4 \times 10^{-4}\) and \(5 \times 10^{-5}\) au for H, C and O atoms when auto setting were used in AIMAll.

3. Theoretical background

3.1. Interacting quantum atoms (IQA)

IQA is scientifically rigorous and exact an energy partitioning scheme (typically applied on a molecule) that is simple to grasp. A molecule is being considered as made of atoms that fill in entire space occupied by a molecule. This means that there are no voids in 3D molecular space or regions of overlapping atoms; hence, each atom has well-defined interatomic boundaries. The concept of IQA atom comes directly from Bader’s QTAIM.\(^{21}\) The IQA procedure makes use of integration techniques (algorithms to compute two-electron integrals over irregularly shaped regions of space) that allow ‘the explicit evaluation of the electron–electron contribution to the electronic energy in a manner independent on the fulfilment of the virial theorem’.\(^{6}\) Concerning electronic (or \textit{ab initio}) molecular energy \(E\), it is partitioned into IQA-defined additive (or total) atomic energies such that, in an ideal case, \(E = E_{\text{IQA}}\), where

\[
E_{\text{IQA}} = \sum_{A} E_{\text{add}}^{A}
\]

(1)

In accord with a chemical intuition, all IQA atoms of a molecule are involved in interactions (either bonded or not). Furthermore, all atoms are treated on equal footing regardless if we see them covalently or otherwise bonded. From this follows that the total atomic energy must consist of two major components, namely the energy of an atom itself (self-atomic energy) and the sum of interaction energies atom \(A\) is experiencing with each other atom \(X\) of a molecule. Note that
to make atomic energy additive, the total interaction energy is halved and it involves all possible unique atom pairs $A,X$, hence also atoms that are distanced from each other (no chemical bond is present)

$$E_{\text{add}}^A = E_{\text{self}}^A + 0.5 \sum_{X \neq A} E_{\text{int}}^{AX}$$  \hspace{1cm} (2)

There are many interaction energy components. Those that are most useful in describing (quantifying and understanding the nature of a chemical bond) are related to electrostatic component of an interaction (called a classical term) and electron sharing (due to quantum-mechanic effects that also incorporate the Pauli Exclusion Principle) known as the exchange-correlation (XC) term that is useful in interpreting the degree of bond covalency,

$$E_{\text{int}}^{AB} = V_{\text{cl}}^{AB} + V_{\text{XC}}^{AB}$$  \hspace{1cm} (3)

It is important to stress that both terms, classical and XC, are always present. This implies that interplay always occurs between them in different kinds of chemical bonds/interactions.

To fully understand challenges the IQA scheme is faced with, it is important to explain how the main components of molecular energy are computed. It is well-known fact that a molecule consists of interacting nuclei and electrons and this is rigorously accounted for in the IQA scheme. To this effect, the self-atomic energy is partitioned into three components,

$$E_{\text{self}}^A = T^A + V_{\text{ee}}^{AA} + V_{\text{en}}^{AA}$$  \hspace{1cm} (4)

that account for kinetic energy of electrons of atom A ($T^A$), the electron–electron potential energy (always repulsive $V_{\text{ee}}^{AA}$ term) and the electron–nucleus potential energy (always attractive $V_{\text{en}}^{AA}$ term) where AA notation is used to indicate that these terms are computed within atomic basin of atom A, $\Omega_A$; hence, all these terms describe intra-atomic energy. Although these terms are seldom used in the study of bonding, they provide an important insight on atomic energy distribution throughout a molecule. From the fact that electrons are predominantly located close
to the atom’s core (nucleus) it follows that $E_{\text{self}}^A >> E_{\text{int}}^{AB}$; hence, the sum of all self-atomic energies

$$E_{\text{self}}^{\text{Tot}} = \sum_A E_{\text{self}}^A,$$

(5)

contributes most to $E$, typically about 98%. Although the total interaction energy (between all atoms in a molecule)

$$E_{\text{int}}^{\text{Tot}} = 0.5\sum_A \sum_{B \neq A} E_{\text{int}}^{AB},$$

(6)

accounts only for about 2% of $E$, ($E = E_{\text{self}}^{\text{Tot}} + E_{\text{int}}^{\text{Tot}}$) to accurately account for this, each diatomic interaction energy $E_{\text{int}}^{AB}$ must be partitioned to smaller and physically meaningful components

$$E_{\text{int}}^{AB} = V_{nn}^{AB} + V_{ne}^{AB} + V_{en}^{AB} + V_{ee}^{AB}$$

(7)

where consecutive terms on the right-hand side account for nucleus-nucleus, nucleus-electron, electron-nucleus, and electron-electron components, respectively, of interacting atoms A and B. The energy terms in Eq. 7 provide an invaluable and on a fundamental level insight on their interplay (or significance) in different chemical bonding; in most cases, however, a classical chemist would make use of just two terms in Eq. 3. Regardless, to generate exact IQA data, IQA-defined $V_{ee}$ energy contributions, i.e., the intra-atomic $V_{ee}^{AA}$ term in Eq. 4 and inter-atomic $V_{ee}^{AB}$ term in Eq. 7 must be computed. Unfortunately, as mentioned above, this is computationally demanding due to the need of the exact form of 2EDM.

There are three possible approaches that one can consider to overcome this hurdle, but each is at the cost of accuracy in the LoT-dependent IQA-defined energy terms. With the main aim of this work in mind, one can make use of:

1. An MP2 level with a specific approximation; this is least desirable in terms of computational time needed in the analysis of large molecules. Moreover, we have shown
recently\textsuperscript{11} that BBC1 performed by far best among Müller, BBC1 and BBC2 approximations. Recall that the Müller approximation is only implemented in AIMAll at present.

(2) The affordable B3LYP level and due to the demand for IQA-based insight (not only on chemical bonding) a big effort was made by Keith (a sole developer of AIMAll software) to incorporate the explicit implementation of the B3LYP functional; details related to this implementation are well-described elsewhere.\textsuperscript{16}

(3) A highly time-efficient HF level for which the exact 2EDM is known; one must realize that dynamical correlation is not included at HF.

To make an educated choice in selecting LoT, it is necessary to gain some insight on significance of differences between LoT/IQA data and the high-level reference data; this is one of the major goals of the present work.

3.2. Fragment attributed molecular system energy change (FAMSEC)

The concept of FAMSEC stems from the need to understand changes taking place throughout a molecule when it finds itself in new environment and explain these changes in terms of classical thinking. A change of environment can be seen as a broad spectrum of chemically relevant events, such as:

(a) Conformational change (to understand relative stability of molecules and role played by intramolecular interactions as well steric clashes).

(b) Formation of adducts and clusters (what drives them to form).

(c) Reaction pathway from reactants through a transitional state to products (this covers inter- and intramolecular interactions, bond breaking and new bond formation; all needed to explain reaction mechanism and preferential substitution sites).

(d) Formation of metal complexes and their relative stability (\textit{e.g.} in terms of formation of 5- and 6-membered coordination rings and their influence on strength of coordination bonds), and many more.
To address the above, one must compare the properties of atoms and chemically meaningful molecular fragments between two structures of a molecule (or molecular system), i.e., when it changed form a particular initial structure (it is used as a reference, ref) to a molecular final structure that is of interest (fin structure). Two major approaches are used in FAMSEC\textsuperscript{20} that focus on:

(1) Changes in properties confined to a 3D space occupied by specific $n$-atom fragment $G$ of a molecule and related to it energetic effect when the $\text{ref} \rightarrow \text{fin}$ structural transformation takes place; this can be seen as a localised to within a fragment $G$ event and loc-FAMSEC energy term applies.

(2) How changes in properties of $G$ and remaining atoms of a molecule (typically treated as another molecular fragment $H$) impact on entire molecule when $\text{ref} \rightarrow \text{fin}$ occurs and what are energetic consequences in terms of stability of a molecule; this is seen as a global, on a molecular scale, event and mol-FAMSEC energy term applies.

The $\text{loc}$-FAMSEC term is defined as

$$E_{\text{attr-loc}}^G = \Delta E_{\text{self}}^G + \Delta E_{\text{int}}^G$$ \hspace{1cm} (8)

where $\Delta E_{\text{self}}^G$ accounts for self-fragment energy change (sum of self-atomic energy changes of atoms constituting a molecular fragment $G$) and $\Delta E_{\text{int}}^G$ quantifies the intra-fragment interaction energy change; in case of a fragment made of two atoms it simply means a diatomic interaction energy change. This is a useful term that quantifies and describes the nature of fragment’s energy change, stabilizing or otherwise. Hence, $E_{\text{attr-loc}}^G$ is suitable to identify strained fragments in the $\text{fin}$ state of a molecule.\textsuperscript{11,22}

The $\text{mol}$-FAMSEC is defined as

$$E_{\text{attr-mol}}^G = E_{\text{attr-loc}}^G + \Delta E_{\text{int}}^{G,H}$$ \hspace{1cm} (9)
where interactions between atoms of $G$ with the remaining atoms of a molecule, treated as molecular fragment $H$, are accounted for through the $\Delta E_{\text{int}}^{G,H}$ energy term (the inter-fragment interaction energy change). Moreover, when $\Delta E_{\text{int}}^{G,H} < 0$, then it implies that $G$ found itself, relative to ref structure, in more attractive (stabilizing) molecular environment when in the fin structure. The interplay between the two components, $E_{\text{attr-loc}}^G$ and $\Delta E_{\text{int}}^{G,H}$, decides whether the molecular fragment $G$ has added to stability of the fin structure of a molecular system ($E_{\text{attr-mol}}^G < 0$) or contributed in a destabilizing manner.

It is important to stress that the loc- and mol-FAMSEC terms can be computed for all unique, 2-, 3-, … n-atom, fragments. From that one can establish which fragments were most locally (de)stabilized and which ones (de)stabilized a molecule the most, etc. This is very useful information in interpreting many chemical phenomena and also puts the energies attributed to a selected fragment in molecular-scale perspective.

Computing loc-FAMSEC requires inputs only from few atoms (typically between 2 and 4) but mol-FAMSEC requires computing on all unique atom pairs and this is very expensive. However, when data related just to a specific fragment is required, one can make use of different IQA-defined energy terms to compute mol-FAMSEC. It has been shown that two expressions can be used and, e.g., for $G = \{A,B\}$ (a two-atom fragment) they are

$$E_{\text{attr-mol}}^G = 2 \Delta E_{\text{add}}^G - E_{\text{attr-loc}}^G$$  \hspace{1cm} (10)

and

$$E_{\text{attr-mol}}^G = \Delta E_{\text{self}}^G - \Delta E_{\text{int}}^G + \Delta \sum_{X \neq A} E_{\text{int}}^{A,X} + \Delta \sum_{X \neq B} E_{\text{int}}^{B,X}$$  \hspace{1cm} (11)

where $\Delta E_{\text{add}}^G$ accounts for fragment’s additive energy change (simply a sum of atomic additive energy changes) and the last two terms in Eq. 11 describe how the sum of interactions between an atom A (or B) belonging to $G$ and each other atom X in a molecule changed on ref $\rightarrow$ fin
transformation. From that one can gain an insight on how (un)favourably the molecular environment changed in relation to each atom of $G$ fragment. Furthermore, one can also compute the $\Delta E_{\text{int}}^{G,H}$ term from

$$\Delta E_{\text{int}}^{G,H} = \Delta \sum_{X=A} E_{\text{int}}^{X,A} + \Delta \sum_{X=B} E_{\text{int}}^{B,X} - 2 \Delta E_{\text{int}}^G$$

(12)

to assess and quantify how '(un)friendly’ new environment in fin structure is towards the entire molecular fragment $G$. The advantage of Eqs. 10-12 is in that one has to run the IQA calculation only for the selected few atoms constituting $G$ because all terms in these equations are always computed by AIMAll. Note that the same applies also to the energy terms in Eqs. 2-4. Hence it is clear that, in principle, one could investigate with confidence and in a very reasonable time inter- and intramolecular interactions in large(er) molecular system provided that reliability of IQA energy terms computed from the B3LYP wavefunction (i.e. B3LYP/IQA data) was known.

3.3. Evaluation of uncertainty in the LoT/IQA energy terms

It is impossible to directly assess reliability of LoT/IQA data when the exact form of 2EDM is not well defined. To this effect, consider a general expression for the additive atomic energy computed at a specific LoT for which the exact form of 2EDM is not known,

$$\text{Comput} E_{\text{add}}^{A} = \left( E_{\text{add}}^{A} \right)^w + \text{err}_{\text{add}}^A$$

(13)

where $\left( E_{\text{add}}^{A} \right)^w$ stands for the exact value of the additive atomic energy; by exact we mean the true value and only one true value exists for, e.g. additive atomic energy of atom A in a specific molecular environment. The $\text{err}_{\text{add}}^A$ term is the error in the computed additive atomic energy (note that $\text{err}_{\text{add}}^A$ as well as the $\text{Comput} E_{\text{add}}^{A}$ are LoT dependent). It is obvious that neither $\left( E_{\text{add}}^{A} \right)^w$ nor $\text{err}_{\text{add}}^A$ values are easily achievable for real life molecules. Hence, the values computed at CCSD/BBC1 have been used as best available approximation of the $\left( E_{\text{add}}^{A} \right)^w$ values.$^{11}$
One must realize, however, that neither molecular energies nor IQA energy components are directly comparable when computed at different LoTs. To illustrate this point, see the data obtained for the LEC (Table 1) showing that there is no obvious (if at all) correlation between differences, relative to the reference data used, in molecular electronic ($\Delta E, E_{\text{CCSD}} = -229.71346$ au), total self ($\Delta E_{\text{self}}^\text{Tot}, (E_{\text{self}}^\text{Tot})_{\text{CCSD/BBC1}} = -226.52801$ au) and total interaction ($\Delta E_{\text{int}}^\text{Tot}, (E_{\text{int}}^\text{Tot})_{\text{CCSD/BBC1}} = -3.18485$ au) energies computed at the specified LoT.

**Table 1** Electronic energies (in au) and differences (in kcal mol$^{-1}$), relative to the CCSD/BBC1 data, in the indicated energies of the LEC for the specified levels of theory

<table>
<thead>
<tr>
<th></th>
<th>HF</th>
<th>B3LYP</th>
<th>B3LYP-GD3</th>
<th>MP2/Müller</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electronic energies, $E$</td>
<td>-228.97714</td>
<td>-230.29514</td>
<td>-230.30011</td>
<td>-229.67804</td>
</tr>
<tr>
<td>$\Delta E = E_{\text{LoT}} - E_{\text{CCSD}}$</td>
<td>462.1</td>
<td>-365.0</td>
<td>-368.1</td>
<td>22.2</td>
</tr>
<tr>
<td>$\Delta E_{\text{self}}^\text{Tot} = E_{\text{self}}^\text{Tot}(\text{LoT}) - E_{\text{self}}^\text{Tot}(\text{CCSD/BBC1})$</td>
<td>820.3</td>
<td>-307.3</td>
<td>-309.4</td>
<td>-158.0</td>
</tr>
<tr>
<td>$\Delta E_{\text{int}}^\text{Tot} = E_{\text{int}}^\text{Tot}(\text{LoT}) - E_{\text{int}}^\text{Tot}(\text{CCSD/BBC1})$</td>
<td>-358.3</td>
<td>-58.8</td>
<td>-57.0</td>
<td>9.2</td>
</tr>
</tbody>
</table>

Let us recall that $E_{\text{IQA}}$ is made of a sum of additive atomic energies (Eq. 1) at any level of theory. Providing that $E = E_{\text{IQA}}$, it is reasonable to assume that each atom in a molecule contributes a specific %-fraction to $E_{\text{IQA}}$, hence $E$, regardless of LoT. To support this supposition, one can expand the $(E_{\text{add}}^A/E)_{\text{LoT}}$ ratio in the series

\[
\frac{(E_{\text{add}}^A)_{\text{LoT}}}{E_{\text{LoT}}} = \frac{(E_{\text{add}}^A)_{\text{CCSD}} + \Delta E_{\text{add}}^A}{E_{\text{CCSD}}} \approx \frac{(E_{\text{add}}^A)_{\text{CCSD}}}{E_{\text{CCSD}}} + \frac{\Delta E_{\text{add}}^A}{E_{\text{CCSD}}} + \frac{\Delta E}{E_{\text{CCSD}}} \approx \frac{(E_{\text{add}}^A)_{\text{CCSD}}}{E_{\text{CCSD}}} \tag{14}
\]

where $\Delta E_{\text{add}}^A$ stands for the error (the exact value, $err_{\text{add}}^A$, is unknown) and $\Delta E$ is the difference in molecular energies obtained at a specific level of theory and CCSD. Eq. 14 holds because $\Delta E \ll E_{\text{CCSD}}$ and it is equally applicable to all IQA energy components of interest in this work, such as additive and self-atomic energies, or interaction energy and its components.
The above paves the way for computing the expected (nearly error-free) IQA energy terms at a selected LoT by use of CCSD/BBC1 reference data. For instance, one can compute the \( R_{\text{self}}^A = \left( \frac{E_{\text{self}}^A}{E} \right)_{\text{CCSD/BBC1}} \) ratio for self-atomic energy of an atom A and use it to compute the expected value at a selected LoT, e.g., B3LYP, as \( \left( \frac{\text{Expect} E_{\text{self}}^A}{E} \right)_{\text{B3LYP}} = R_{\text{self}}^A \times E_{\text{B3LYP}} \). Finally, by comparing the expected and computed values at a specific LoT one obtains the relative error, RE (relative, through the \( R_{\text{self}}^A \) term, to the best available reference data), \( \text{RE} = \Delta_{\text{self}}^A = \left( \frac{\text{Comput} E_{\text{self}}^A}{E} \right)_{\text{B3LYP}} - \left( \frac{\text{Expect} E_{\text{self}}^A}{E} \right)_{\text{B3LYP}} \), that serves as an excellent estimate of reliability of B3LYP/IQA descriptors in sound interpreting of, e.g. intermolecular interactions. This protocol was employed here to evaluate IQA and FAMSEC energy terms computed at all levels of theory investigated in this work.

4. Results and discussion

4.1. Molecular energy change

Although the \( E_{\text{LoT}} \) values are not directly comparable, changes in \( E \), such as \( \Delta E_{\text{LoT}} = E_{\text{fin}} - E_{\text{ref}} \) computed at the specific LoT, are expected to be highly comparable between levels of theory. As shown in Fig. 2, the \( \Delta E_{\text{LoT}} \) values computed for the \textbf{Lin} \( \rightarrow \) \textbf{LEC} and \textbf{Lin} \( \rightarrow \) \textbf{Ecl} structural changes at all levels of theory reproduced those obtained at CCSD, \( \Delta E_{\text{CCSD}} = -2.34 \) and +9.17 kcal mol\(^{-1}\), respectively, quite well.

![Fig. 2 Changes in molecular electronic energies computed at each level of theory for the Lin to LEC (part a) and Lin to Ecl (part b) structural transformations of glycol. B3-GD3 = B3LYP-GD3.](image-url)
On average, (i) the smallest errors are observed at B3LYP (with and without GD3) for which \( \Delta E_{\text{CCSD}} \) values were reproduced to within a small fraction of kcal mol\(^{-1}\) and (ii) the largest errors (up to about 1 kcal mol\(^{-1}\)) were obtained at HF.

Having established that \( \Delta E_{\text{LoT}} \approx \Delta E_{\text{CCSD}} \) (this can be seen as a prerequisite for any IQA-based analyses that follow) it was of critical importance to test:

(1) The impact of inequality \( E_{\text{IQA}} \neq E \) at MP2 level on REs in the computed MP2/IQA descriptors,

(2) The impact of GD3 on REs in the computed B3LYP-GD3/IQA descriptors, and

(3) How well \( \Delta E_{\text{IQA}} \) reproduced \( \Delta E \) at each LoT for both structural changes; this is a new criterion proposed by us recently to assess quality of FAMSEC data generated at post HF levels of theory with different approximations.\(^{11}\) We test validity of the \( \Delta E_{\text{IQA}} = \Delta E \) criterion by expanding our studies to (i) lower levels of theory and (ii) changes in the LoT/IQA descriptors commonly used in describing interactions.

4.2. \( E_{\text{IQA}} \) vs. electronic molecular energy \( E \)

Importance of \( \Delta E = E_{\text{IQA}} - E \approx 0 \) has been strongly advocated,\(^{16}\) hence relevant data is included in Table 2. It is seen that, within typical in IQA calculations integration errors, this equality has been met at the HF and B3LYP levels perfectly but failed dismally at MP2/Müller.

<table>
<thead>
<tr>
<th>Level of theory</th>
<th>LEC</th>
<th>Ecl</th>
<th>Lin</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF</td>
<td>0.01</td>
<td>–0.01</td>
<td>–0.01</td>
</tr>
<tr>
<td>B3LYP</td>
<td>0.00</td>
<td>0.04</td>
<td>0.08</td>
</tr>
<tr>
<td>B3LYP-GD3</td>
<td>3.07</td>
<td>2.61</td>
<td>3.11</td>
</tr>
<tr>
<td>MP2/Müller</td>
<td>−170.8</td>
<td>−171.8</td>
<td>−170.4</td>
</tr>
</tbody>
</table>

In case of optimised structures at the B3LYP-GD3 level, \( \Delta E \) of about 3 kcal mol\(^{-1}\) is observed and this is due to the Grimme’s empirical dispersion correction. Large \( \Delta E \) values (on average for
three conformers of about $-171$ kcal mol$^{-1}$) were obtained for the MP2/Müller level; $E_{\text{IQA}}$ highly overestimated electronic energy.

Strictly speaking, the fact that $\Delta E = E_{\text{IQA}} - E = 0$ was obtained might simply imply that the sum of:

(a) Self-atomic and all unique diatomic interaction energies, or

(b) Additive atomic energies

reproduced $E$ very well. However, HF/IQA and B3LYP/IQA data must have some REs due to these theories intrinsic limitations. Furthermore, it was also of importance and interest to find out about RE distribution among the IQA descriptors at all LoTs under investigation.

To address this, let us start with analysis of relative errors, $\Delta_{\text{add}}$, in individual additive atomic energies, $E_{\text{add}}^A$, obtained at each LoT. To fully appreciate the importance of the approach taken here in comparing data computed at different LoT, let us analyse CCSD/BBC1 and, as an example, HF data obtained for the LEC (Table 3). We note large differences between computed values at both LoTs, $\Delta E_{\text{add}}^A = (E_{\text{add}}^A)_{\text{HF}} - (E_{\text{add}}^A)_{\text{CCSD/BBC1}}$, of about 135 and 80 kcal mol$^{-1}$ for O- and C-atoms, respectively. This shows that, e.g. the computed $(E_{\text{add}}^O)_{\text{HF}}$ values appear to be significantly smaller (less negative, hence underestimated by $\sim 135$ kcal mol$^{-1}$) when directly compared with atomic additive energies computed at the CCSD/BBC1 level. However, for the error-free data, it is expected that the IQA additive energy of, say O3-atom in the LEC, should contribute the same fraction, $R_{\text{LoT}}^{O3} = R_{\text{CCSD}}^{O3} = (E_{\text{add}}^{O3})_{\text{LoT}} / E_{\text{LoT}} = (E_{\text{add}}^{O3})_{\text{CCSD}} / E_{\text{CCSD}} = 0.32762$, to $E$ of LEC at any level. When the nearly error-free or expected at HF additive atomic energy of the O3-atom was computed, $E_{\text{add}}^{O3} = 0.32762 \times E_{\text{HF}} = -75.01825$ au, it revealed that it is smaller (less negative) than the computed value, $E_{\text{add}}^{O3} = -75.04475$ au. This clearly shows that the computed additive energy of O3 at HF carries an error and, more specifically, it overestimated (being too negative) the expected value.
Table 3  Analysis of additive atomic energies of the LEC obtained at CCSD/BBC1 and HF levels.\(^a\)

<table>
<thead>
<tr>
<th>Atom A</th>
<th>(E_{\text{add}}^A)</th>
<th>(E_{\text{add}}^A / E)</th>
<th>Comput(E_{\text{add}}^A)</th>
<th>Expect(E_{\text{add}}^A)</th>
<th>(\Delta E_{\text{add}}^A)</th>
<th>(\Delta^A_{\text{add}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>-37.94313</td>
<td>0.16518</td>
<td>-37.81718</td>
<td>-37.82151</td>
<td>79.0</td>
<td>2.7</td>
</tr>
<tr>
<td>C2</td>
<td>-37.94389</td>
<td>0.16518</td>
<td>-37.81888</td>
<td>-37.82227</td>
<td>78.4</td>
<td>2.1</td>
</tr>
<tr>
<td>O3</td>
<td>-75.25949</td>
<td>0.32762</td>
<td>-75.04475</td>
<td>-75.01825</td>
<td>134.7</td>
<td>-16.6</td>
</tr>
<tr>
<td>O4</td>
<td>-75.27158</td>
<td>0.32768</td>
<td>-75.05523</td>
<td>-75.03030</td>
<td>135.8</td>
<td>-15.6</td>
</tr>
<tr>
<td>H5</td>
<td>-0.48480</td>
<td>0.00211</td>
<td>-0.48129</td>
<td>-0.48325</td>
<td>2.2</td>
<td>1.2</td>
</tr>
<tr>
<td>H6</td>
<td>-0.48892</td>
<td>0.00213</td>
<td>-0.48673</td>
<td>-0.48735</td>
<td>1.4</td>
<td>0.4</td>
</tr>
<tr>
<td>H7</td>
<td>-0.58017</td>
<td>0.00253</td>
<td>-0.56819</td>
<td>-0.57831</td>
<td>7.5</td>
<td>6.4</td>
</tr>
<tr>
<td>H8</td>
<td>-0.58022</td>
<td>0.00253</td>
<td>-0.56759</td>
<td>-0.57836</td>
<td>7.9</td>
<td>6.8</td>
</tr>
<tr>
<td>H9</td>
<td>-0.58095</td>
<td>0.00253</td>
<td>-0.56962</td>
<td>-0.57909</td>
<td>7.1</td>
<td>5.9</td>
</tr>
<tr>
<td>H10</td>
<td>-0.57971</td>
<td>0.00252</td>
<td>-0.56765</td>
<td>-0.57785</td>
<td>7.6</td>
<td>6.4</td>
</tr>
</tbody>
</table>

\(^a\) Additive atomic energies are in au. Differences \(\Delta E_{\text{add}}^A = (E_{\text{add}}^A)_{\text{HF}} - (E_{\text{add}}^A)_{\text{CCSD/BBC1}}\) and relative errors \(\Delta^A_{\text{add}} = \text{ Comput}E_{\text{add}}^A - \text{ Expect}E_{\text{add}}^A\) are in kcal mol\(^{-1}\).  

Hence, relative to the reference used our best estimate of the RE in the additive energy, \(\Delta^O_{\text{add}}\), at HF is \(\text{ Comput}E_{\text{add}}^O - \text{ Expect}E_{\text{add}}^O = -16.6\) kcal mol\(^{-1}\) which is an order of magnitude smaller (in absolute value) than the direct difference between computed values, \(\Delta E_{\text{add}}^O = (\Delta E_{\text{add}}^O)_{\text{HF}} - (\Delta E_{\text{add}}^O)_{\text{CCSD/BBC1}} = 134.7\) kcal mol\(^{-1}\). This example nicely illustrates how misleading direct comparison between data generated at different levels of theory can be (in other words closeness between the HF and CCSD data), not only in the difference in values (large or small) but also in directionality of a RE in terms of under- or overestimation. Relevant data obtained at all levels for three conformers of glycol is presented in Tables S1-S11 in the ESI.

Relative error analysis for individual atoms in the LEC (Fig. 3(a)) and Ecl and Lin conformers (Fig. S1 in the ESI) shows that relative to the CCSD/BBC1 data:

(a) Individual additive atomic energies computed at all LoTs do carry significant relative errors; the largest, of up to nearly \(-50\) kcal mol\(^{-1}\), are observed for O-atoms at MP2/Müller. Unidirectional REs are observed at MP2/Müller (all additive atomic energies were overestimated, they are too negative). This contributed to a large departure of \(E_{IQA}\) at
MP2/Müller from $E_{\text{MP2}}$ and more specifically, $E_{\text{IQA}(\text{MP2}/\text{Müller})} \ll E_{\text{MP2}}$. It is then clear that in case of $E_{\text{IQA}(\text{LoT})}$ being largely different then $E_{\text{LoT}}$, one must expect large uncertainties in additive atomic energies as well as self-atomic energies as they contribute most to $E_{\text{add}}^A$.

(b) Smaller and bidirectional REs are observed at HF and B3LYP levels. At both levels, largest REs are observed for O-atoms (as was the case at MP2/Müller) but surprisingly disproportionally large REs were obtained for H-atoms at HF.

c) Importantly, smallest REs in additive atomic energies were obtained at B3LYP level with a standard deviation of about ±5 kcal mol$^{-1}$ (±9 kcal mol$^{-1}$ was obtained for the HF data).

Furthermore, regardless of the degree and directionality of REs in $E_{\text{add}}^A$, the $\Delta E_{\text{IQA}}$ values mimic those obtained for the $\Delta E$ values (Fig. 2a) for the Lin $\rightarrow$ LEC structural change – see data in Fig. 3b – even in the case of worst performing MP2/Müller/IQA (relevant data obtained for the Lin$\rightarrow$ Ecl structural change is shown in Fig. S2 in the ESI). The data shown in Fig. 3 clearly demonstrates that (i) significant REs in the computed $E_{\text{add}}^A$ values are present (part a) and significance and directionality of relative errors is LoT- and atom-dependent, (ii) these REs are

![Figure 3](image_url)

**Fig. 3** Using the CCSD/BBC1 data as reference, relative error (RE) at the indicated levels of theory in individual additive atomic energies obtained for the lowest energy conformer (LEC) of glycol – part (a). Change in $E_{\text{IQA}}$ for the Lin $\rightarrow$ LEC structural transformation – part (b), where dashed line indicates the $\Delta E$ value obtained at the CCSD/BBC1 level; B3-GD3 stands for the B3LYP-GD3 level and MP2/M for the MP2/Müller level.
of systematic nature as their perfect cancelation took place when changes $\Delta E_{\text{IQA}}$ were computed (part b).

The equality $\Delta E_{\text{IQA}} = \Delta E$ found at each LoT is promising as it suggests that differences in the computed energy terms (for a variety of descriptors) computed on a structural change of glycol might be suitable in describing chemical bonds and interactions. To this effect, see data in Fig. 4, where remarkable agreement in the computed individual $\Delta E_{\text{add}}^A$ values at all LoTs is observed for the $\text{Lin} \rightarrow \text{Ecl}$ structural change.

![Fig. 4](image)

Fig. 4 Change in individual additive atomic energies computed at the indicated levels of theory for the $\text{Lin} \rightarrow \text{Ecl}$ structural transformation of glycol.

Regardless of the LoT, data in Fig. 4 shows that, as one would expect, due to steric clash O-atoms became destabilised the most in the $\text{Ecl}$ conformer (by +6.6 kcal mol$^{-1}$ at CCSD/BBC1). For the remaining levels, additive energy of O-atoms increased, on average, by +7.82±0.05 kcal mol$^{-1}$. Interestingly and most likely not easily predictable, the H-atoms bonded to O-atoms became stabilised by −2.0 kcal mol$^{-1}$ at CCSD/BBC1; this has been reproduced very well at remaining LoTs, on average by −1.9±0.1 kcal mol$^{-1}$. For data obtained on the $\text{Lin} \rightarrow \text{LEC}$ structural transformation see Fig. S3 in the ESI. This is an important finding because it shows that one can obtain a reliable and invaluable insight on most (de)stabilized molecular fragments for the $\text{ref} \rightarrow \text{fin}$ structural change from any level of theory tested here. This, in turn, can be used.
not only to explain the change in the molecular energy but also to point at specific molecular region the energy of which has decreased or increased most.

**Fig. 5** Topology of $\Delta E^A_{\text{add}}$ in the Ecl conformer (after Lin $\rightarrow$ Ecl transformation) showing an increase (red solid circle) and decrease (blue dashed circles) in additive atomic energies; relative thickness of lines is nearly proportional to the energy change, e.g. for O-atoms, $\Delta E^O_{\text{add}} = +6.6$ kcal mol$^{-1}$ and H5,H6-atoms $\Delta E^H_{\text{add}} = -2.0$ kcal mol$^{-1}$, both values at the CCSD/BBC1 level.

This can be seen as a kind of topology of $\Delta E^A_{\text{add}}$ distribution throughout a molecule in the fin structure of a molecule and, as an example, a pictorial presentation of $\Delta E^A_{\text{add}}$ in the Ecl conformer (after Lin $\rightarrow$ Ecl transformation) is shown in Fig. 5; importantly, this picture holds for all LoTs and the same approach can be taken to draw topologies for changes in other energy terms. Note that relative thickness of rings in Fig. 5 is proportional to computed $\Delta E^A_{\text{add}}$ values.

**4.3. Relative errors in the total self-molecular and total interaction energy terms**

The main component of $E_{\text{IQA}}$ is always, and by far, the total self-molecular energy, $E^{\text{Tot}}_{\text{self}}$; for instance, at the CCSD/BBC1 level it constitutes 98.6% of the LEC electronic energy. Hence, the contribution made by all diatomic interaction energies, $E^{\text{Tot}}_{\text{int}}$, is just 1.4% of $E_{\text{CCSD}}$. The relative errors in the total (i) self-molecular energy ($\Delta^{\text{Tot}}_{\text{self}} = \frac{\text{Comput} E^{\text{Tot}}_{\text{self}} - \text{Expect} E^{\text{Tot}}_{\text{self}}}{\text{Comput} E^{\text{Tot}}_{\text{self}}}$) and (ii) interaction energy ($\Delta^{\text{Tot}}_{\text{int}} = \frac{\text{Comput} E^{\text{Tot}}_{\text{int}} - \text{Expect} E^{\text{Tot}}_{\text{int}}}{\text{Comput} E^{\text{Tot}}_{\text{int}}}$) are shown for the LEC in Fig. 6 (relevant data for the Lin and Ecl conformers are shown in Fig. S4 in the ESI).
From the analysis of REs, $\Delta_{\text{self}}^{\text{Tot}}$ and $\Delta_{\text{int}}^{\text{Tot}}$, we note that:

(a) The same directionality in REs is observed at HF, B3LYP and B3LYP-GD3 for which $E_{\text{self}}^{\text{Tot}}$ and $E_{\text{int}}^{\text{Tot}}$ were under- and overestimated, respectively, relative to CCSD/BBC1. Recall that both these energy terms are always negative irrespective of the level of theory used.

(b) By far the largest REs, $\Delta_{\text{self}}^{\text{Tot}}$ and $\Delta_{\text{int}}^{\text{Tot}}$, are observed for HF. The self-molecular energy was underestimated by staggering 364.6 kcal mol$^{-1}$ whereas the total interaction energy, $E_{\text{int}}^{\text{Tot}}$, was underestimated by the same value in accord with $E_{\text{IQA}} = E$ at HF.

(c) Almost an order of magnitude smaller (relative to HF) REs in $E_{\text{self}}^{\text{Tot}}$ and $E_{\text{int}}^{\text{Tot}}$ are found for B3LYP. Here again, summing up the $\Delta_{\text{self}}^{\text{Tot}}$ and $\Delta_{\text{int}}^{\text{Tot}}$ relative errors results in $\Delta E = E_{\text{IQA}} - E = 0$.

(d) Very different picture emerged in the case of MP2/Müller. The total self-molecular energy was largely overestimated, by $-180$ kcal mol$^{-1}$, but interestingly, the $E_{\text{int}}^{\text{Tot}}$ term is underestimated just by few kcal mol$^{-1}$ resulting in the $\Delta_{\text{int}}^{\text{Tot}}$ RE being smallest among all LoTs. This fully explains the origin of $\Delta E = E_{\text{IQA}} - E < 0$ obtained from the MP2/Müller data (as shown in Table 2).
In terms of absolute values, we found B3LYP ~ B3LYP-GD3 << MP2/Müller << HF trend in Δ_{self}^{Tot} relative errors and MP2/Müller << B3LYP ~ B3LYP-GD3 << HF trend in Δ_{int}^{Tot} REs. For a full set of data obtained for the expected and computed E_{self}^{Tot} and E_{int}^{Tot} values for all conformers of glycol as well as associated REs see Table S12 of the ESI.

Although different degrees of and directionality in the REs errors in computed E_{self}^{Tot} and E_{int}^{Tot} values for each conformer were found, remarkably the changes in the total self-molecular and interaction energy terms, ΔE_{self}^{Tot} and ΔE_{int}^{Tot}, computed at the CCSD/BBC1 level for the Lin → LEC structural change were reproduced quite well at all LoTs – Fig. 7a. The ΔE_{self}^{Tot} value of +2.7 kcal mol\(^{-1}\) at the CCSD/BBC1 level (relative to Lin, self-molecular energy of LEC increased, became less negative) was consistently slightly overestimated (on average by about 3 kcal mol\(^{-1}\)) at other LoTs, ΔE_{self}^{Tot} = 5.6±1.6 kcal mol\(^{-1}\). On the other hand, ΔE_{int}^{Tot} of −11.5 kcal mol\(^{-1}\) at CCSD/BBC1 (a contribution of the interaction energy term to E strengthened; it became more negative when in the LEC) was slightly underestimated (on average by about 3 kcal mol\(^{-1}\)) at other LoTs, ΔE_{int}^{Tot} = −8.4±1.2 kcal mol\(^{-1}\). From the combined data shown in Fig. 7 it follows that:

Fig. 7 Changes in the self-molecular and total interaction energies at the indicated levels of theory obtained for Lin to LEC (part a) and Lin to Ecl (part b) structural transformations of glycol. B3-GD3 = B3LYP-GD3, MP2/M = MP2/Müller.
(a) HF appears least reliable; on average absolute changes in the $E_{\text{self}}^{\text{Tot}}$ and $E_{\text{int}}^{\text{Tot}}$ terms at HF are significantly larger when compared with B3LYP and MP2/Müller.

(b) Relative to CCSD/BBC1, trends in the $\Delta E_{\text{self}}^{\text{Tot}}$ and $\Delta E_{\text{int}}^{\text{Tot}}$ values are reproduced for both structural changes only at B3LYP, B3LYP-GD3 and MP2/Müller. This suggests that HF data is more sensitive to molecular system complexity, interactions involved, etc.

(c) Grimme’s empirical dispersion correction (GD3) plays small and indecisive impact on the computed $\Delta E_{\text{self}}^{\text{Tot}}$ and $\Delta E_{\text{int}}^{\text{Tot}}$ values as: (i) GD3 has minimised $\Delta E_{\text{self}}^{\text{Tot}}$ for both structural changes but (ii) inconsistency of the GD3 impact is observed in the case of the $\Delta E_{\text{int}}^{\text{Tot}}$ value as they slightly increased/decreased for Lin $\rightarrow$ LEC and Lin $\rightarrow$ Ecl, respectively.

4.4. Relative errors in computed individual self-atomic energies

Full analyses of relative errors in self-atomic energies and differences $\Delta E_{\text{self}}^{A} = (E_{\text{self}}^{A})_{\text{LoT}} - (E_{\text{self}}^{A})_{\text{CCSD/BBC1}}$ computed for each conformer at all levels of theory are presented in commented Tables S13-S24 in the ESI. In general, these data nicely illustrates a significance of using expected, relative to the CCSD/BBC1 data, values in calculating REs, $\Delta_{\text{self}}^{A}$, rather than misleading differences (closeness) in computed self-atomic energies at two levels of theory, $\Delta E_{\text{self}}^{A}$. Moreover, the consistency in the trends identified (Tables S13-S24 in the ESI) strengthens the earlier conclusion that significance of REs (large/small) and their direction (negative/positive) is (i) atom-dependent and (ii) predominantly and highly dependent on (hence characteristic for) a LoT used (including the level of approximation, if applicable). This can be strengthened even further by noting that $\Delta_{\text{self}}^{C} > \Delta_{\text{self}}^{O}$ at HF and both B3LYP methods, whereas the opposite trend in REs, $\Delta_{\text{self}}^{C} < \Delta_{\text{self}}^{O}$, is observed for the MP2/Müller level even though the same basis set and auto settings in AIMAll were used throughout. This reflects the impact made by the Müller approximation for which main origin of errors is rooted in the improper evaluation
of the $V_{ee}^A$ term;\textsuperscript{11} thus, the larger number of electrons populating an atomic basin (here O-atom) the larger error in the computed self-atomic energy is observed.

In search for the origin of the REs in the total self-molecular energy, $\Delta_{\text{self}}^{\text{Tot}}$, we performed analysis of individual atoms’ self-energies. The REs, $\Delta_{\text{self}}^A = \text{Comput}_{\text{self}}^A - \text{Expect}_{\text{self}}^A$, obtained for the LEC are pictorially presented in Fig. 8. Relevant data obtained for the Ecl and Lin conformers is shown in Fig. S5 in the ESI. Remarkable consistency is observed in the REs computed for the same atoms in all conformers at a specific LoT; they are nearly the same.

![Graph showing relative errors in self-atomic energies for LEC of glycol](image)

**Fig. 8** Using CCSD/BBC1 data as a reference, relative errors (RE) in computed self-atomic energies in the LEC of glycol at the indicated levels of theory.

Analysis of relative errors reveals that:

(a) The largest REs are generated at HF and MP2/Müller; the $E_{\text{self}}^A$ values were underestimated (too positive) and overestimated (too negative) by up to $+80$ and $-50$ kcal mol\textsuperscript{-1} at HF and MP2/Müller levels, respectively. In both cases a unidirectional REs are observed. Possibly, a very poor performance of HF can be attributed to the absence of the dynamic electron-electron correlation (recall that no approximation is used in this case and for HF single-determinant wavefunction the expression used for the 2EDM is exact). Rather poor performance observed at MP2/Müller must be linked, as mentioned above, with intrinsic shortcomings of the Müller approximation.
(b) Nearly one order of magnitude smaller REs were obtained at B3LYP and B3LYP-GD3 levels. Clearly, at least partly due to the explicit implementation of B3LYP density functional in AIMAll, the B3LYP/IQA data has produced outstanding results relative to HF/IQA and MP2/Müller/IQA.

(c) Consistently, the smallest REs were obtained for H-atoms at all LoTs but they are disproportionately large when compared with the $\Delta_{\text{self}}^O$ and $\Delta_{\text{self}}^C$ values and particularly so at HF.

Typically, small changes in self-atomic energies are expected on a structural change and this is indeed observed at all LoTs with HF producing, relative to CCSD/BBC1, too large values in most cases - see data obtained for Lin $\rightarrow$ Ecl and Lin $\rightarrow$ LEC in Fig. 9. Possibly most unexpectedly, MP2/Müller outperformed all other LoTs even though large REs for individual self-atomic energies and the largest departure from the $E_{\text{IQA}} = E$ equality was obtained at this level. We found the following trend for averaged $|\Delta E_{\text{self}}^A|$ values computed for all atoms on both structural changes: $0.4 \pm 0.5$ (MP2/Müller) $< 0.6 \pm 0.6$ (B3LYP) $= 0.6 \pm 0.6$ (B3LYP-GD3) $< 1.3 \pm 1.5$ (HF) – all values in kcal mol$^{-1}$. Furthermore, the trend in $\Delta E_{\text{self}}^A$ seen at CSD/BBC1 was reproduced exactly only at MP2/Müller.

![Fig. 9](image-url) Changes in individual self-atomic energies at the indicated levels of theory for the Lin $\rightarrow$ LEC (part a) and Lin $\rightarrow$ Ecl structural transformation of glycol.

Focusing on most interesting/important observations related to the data shown in Fig. 9 we note that:
(a) Considering the LEC, there are 2-atom fragments that experienced most significant change in $E_{\text{self}}^A$, namely (i) a backbone (C1–C2) with opposite trends for each C-atom and (ii) a proton donor O3–H6 group; this shows that X–H···Y–Z representation recommended by IUPAC$^{28}$ does not fully recover variations in $E_{\text{self}}^A$; in this case X–H···Y–Z–W would be more appropriate.

(b) All atoms in the Ecl conformer experienced significant changes in $E_{\text{self}}^A$ with the C1–C2 backbone being influenced most. Only atoms of two OH functional groups decreased their self-atomic energy when in Ecl; an opposite trend to that observed for atoms involved in a bonding interaction in the LEC.

Finally, the $\Delta E_{\text{self}}^A$ values obtained at other LoTs compare well with those obtained at CCSD/BBC1 except HF that generated overestimated values in many instances.

4.5. Relative errors in computed diatomic interaction energy and its components

Arguably, the diatomic interaction energy, $E_{\text{int}}^{A,B}$, and its components, the classical term $V_{\text{cl}}^{A,B}$ and XC-term $V_{\text{XC}}^{A,B}$, are most useful, hence most frequently used, in describing/characterising chemical bonds and interactions from the IQA perspective. We decided to take advantage of full IQA calculations on entire molecules (the three glycol conformers) and performed a comprehensive analysis of all possible interactions. To facilitate interpretation of data obtained at all LoTs, a full set of CCSD/BBC1 data obtained is presented in Tables S25-S30 in the ESI for all 45 unique atom-pairs in three conformers of glycol as well as the Lin $\rightarrow$ LEC and Lin $\rightarrow$ Ecl structural changes.

Following the protocol used for the additive and self-atomic energies, full analyses of REs, $\Delta_{\text{int}}^{A,B}$ and differences $\Delta E_{\text{int}}^{A,B}$ computed for covalently bonded atom-pairs in each conformer at all levels of theory (incorporating two intramolecular interactions of interest, O4···H6 in the LEC and O3···O4 in the Ecl conformer, as well as both components of diatomic interactions) is
included in Tables S31-S66 in the ESI; pictorial presentation of the $\Delta_{\text{int}}^{A,B}$, $\Delta_{\text{cl}}^{A,B}$ and $\Delta_{\text{XC}}^{A,B}$ REs obtained for the three conformers is shown in Figs. S6-S8 in the ESI.

To gain quantitative evaluation of typical for each level of theory REs, $\Delta_{\text{int}}^{A,B}$, $\Delta_{\text{cl}}^{A,B}$ and $\Delta_{\text{XC}}^{A,B}$, combined analyses for covalently bonded atoms, i.e. C–C, C–O, C–H, and O–H, in glycol conformer are presented in Tables S67-S69 in the ESI. For convenience, the data in these tables was collated in Table 4 that, together with data shown in Figs. S6-S8 in the ESI, reveals several important trends:

(1) It is instantly apparent that HF data carries largest and systematic REs, $\Delta_{\text{int}}^{A,B}$; in many instances they are an order of magnitude larger than those at remaining LoTs. Furthermore, except the C–C backbone, all diatomic interactions were largely overestimated at HF. For instance, in the case of the C–O covalent bonds the relative error is about $-115$ kcal mol$^{-1}$ that is nearly 24% more than the expected value.

(2) The main source of the $\Delta_{\text{int}}^{A,B}$ REs computed at HF for the atom-pairs forming C–O, C–C and O–H covalent bonds in all conformers is nearly entirely due to errors in the classical component, $\Delta_{\text{cl}}^{A,B}$. For instance, we found for the LEC $\Delta_{\text{int}}^{\text{C1,O4}}/\Delta_{\text{cl}}^{\text{C1,O4}} = -109.1/-110.4$ kcal mol$^{-1}$ or $\Delta_{\text{int}}^{\text{O3,H6}}/\Delta_{\text{cl}}^{\text{O3,H6}} = -49.1/-46.3$ kcal mol$^{-1}$. Considering the C–H covalent bonds, REs in the XC-term, $\Delta_{\text{XC}}^{A,B}$, contributed most to inaccuracy of interaction energy, e.g., $\Delta_{\text{int}}^{\text{C1,H7}}/\Delta_{\text{XC}}^{\text{C1,H7}} = -28.5/-18.5$ kcal mol$^{-1}$ was obtained for the LEC.

(3) Regardless of the level of theory, in most cases relative errors in the classical term are significantly larger than that in the XC-term. In some instances they can be an order of magnitude larger and particularly so at HF, e.g., REs of $-204\pm133\%$ and $10.5\pm0.1\%$ were found for the classical and the XC-terms, respectively, at HF in the case of the C–H bonds.
Table 4 Collated data from analyses of relative errors in the computed diatomic interaction energy and its components obtained for the covalently bonded atoms in three conformers of glycol at the indicated levels of theory: interaction energy – part A; classical component – part B; exchange-correlation component – part C. All values are in kcal mol\(^{-1}\).

<table>
<thead>
<tr>
<th>Part A</th>
<th>HF</th>
<th>B3LYP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bond</td>
<td>(E_{\text{int}}^{\text{A,B}})</td>
<td>(\Delta_{\text{int}}^{\text{A,B}})</td>
</tr>
<tr>
<td>C–C</td>
<td>-59.3±7.1</td>
<td>27.8±1.6</td>
</tr>
<tr>
<td>C–O</td>
<td>-484.4±7.9</td>
<td>-114.4±4.0</td>
</tr>
<tr>
<td>C–H</td>
<td>-167.7±3.2</td>
<td>-29.4±1.2</td>
</tr>
<tr>
<td>O–H</td>
<td>-361.6±3.1</td>
<td>-48.6±0.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part B</th>
<th>HF</th>
<th>B3LYP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bond</td>
<td>(V_{\text{el}}^{\text{A,B}})</td>
<td>(\Delta_{\text{el}}^{\text{A,B}})</td>
</tr>
<tr>
<td>C–C</td>
<td>120.7±3.7</td>
<td>41.5±1.8</td>
</tr>
<tr>
<td>C–O</td>
<td>-330.2±6.4</td>
<td>-115.7±3.8</td>
</tr>
<tr>
<td>C–H</td>
<td>7.2±3.1</td>
<td>-11.1±1.2</td>
</tr>
<tr>
<td>O–H</td>
<td>-241.3±5.0</td>
<td>-45.7±0.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part C</th>
<th>HF</th>
<th>B3LYP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bond</td>
<td>(V_{\text{xc}}^{\text{A,B}})</td>
<td>(\Delta_{\text{xc}}^{\text{A,B}})</td>
</tr>
<tr>
<td>C–C</td>
<td>-180.0±3.4</td>
<td>-13.7±0.3</td>
</tr>
<tr>
<td>C–O</td>
<td>-154.2±1.8</td>
<td>1.3±0.4</td>
</tr>
<tr>
<td>C–H</td>
<td>-175.0±0.2</td>
<td>-18.4±0.2</td>
</tr>
<tr>
<td>O–H</td>
<td>-120.3±1.9</td>
<td>-2.9±0.4</td>
</tr>
</tbody>
</table>

(4) Not only smallest but typically well below 10 and 3% REs are observed at MP2/Müller for the classical and XC-term, respectively. On the other hand, REs obtained at B3LYP are about twice as large as those at MP2/Müller and several times smaller when compared with HF.
(5) It is clear that REs (their values and directionality) depend on the kind of the covalent bond and the LoT. For instance: the %-errors in the $E_{\text{int}}^{\text{Comput} A,B}$ term obtained for (i) the C–C bonds are largest among all covalent bonds at all LoTs and the computed interaction energy was largely underestimated at HF (by $-47.5\pm8.9\%$) and overestimated at other levels, by $22.5\pm0.9\%$ at B3LYP and $7.9\pm0.2\%$ at MP2/Müller and (ii) the C–O covalent bonds show just opposite trend, namely the $E_{\text{int}}^{\text{Comput} A,B}$ term was largely overestimated by $-114.4\pm4.0$ kcal mol$^{-1}$ resulting in $23.6\pm0.5$ %-error at HF whereas it was underestimated at remaining levels with %-errors of about $3\%$.

(6) Interestingly, trends in REs are nearly identical at B3LYP and MP2/Müller, e.g., the interaction energy was slightly underestimated only in the case of the C–O bonds or the XC-term was overestimated for all covalent bonds at these levels. (Note that results obtained at B3LYP and B3LYP-GD3 are nearly identical.)

(7) In general, diatomic interaction energies carry the largest at HF and smallest at MP2/Müller REs. Recall that the origin of large departure of $E_{\text{IQA}}$ from $E$ at MP2/Müller is, due to the Müller approximation for 2EDM, predominantly linked with self-atomic energies.

We note with interest that REs obtained at B3LYP for the intramolecular interactions of particular interest in this work, O4⋯H6 in the LEC and O3⋯O4 in the Ecl conformer, are rather small. As a matter of fact they can be seen as perfectly suited for a general description from which a nature of each interaction and relative contributions made by two energy terms (pointing at classical term as a major contributor and reflecting non-negligible fraction of the XC-term) are fully comparable with the picture generated at CCSD/BBC1 (Figs. S6-S7 in the ESI). In support of the above we found for the computed energy terms at B3LYP:

(i) $E_{\text{int}}^{\text{H6,O4}}$ of $-86.1$ and $+121.4$ kcal mol$^{-1}$ were underestimated by 5 and 0.1 kcal mol$^{-1}$, respectively (Tables S32 and S40 in the ESI).
(ii) $V_{\text{cl}}^{O4\cdots H6}$ and $V_{\text{cl}}^{O3\cdots O4}$ of $-83.8$ and $+129.9$ kcal mol$^{-1}$ were under- and overestimated by 4.3 and 6.9 kcal mol$^{-1}$, respectively (Tables S44 and S52 in the ESI).

(iii) $V_{\text{cl}}^{O4\cdots H6}$ and $V_{\text{cl}}^{O3\cdots O4}$ of $-2.3$ and $-8.4$ kcal mol$^{-1}$ were underestimated by 0.8 and 6.9 kcal mol$^{-1}$, respectively (Tables S56 and S64 in the ESI).

Availability of a single set of data for just two different intramolecular interactions does not allow for any statistical treatment. Hence, we decided to analyse the O4···H6 and O3···O4 interactions in all conformers, regardless of the interatomic distance (Table S70 in the ESI). Data obtained (i) confirmed the general trend in LoT-dependent REs, HF > B3LYP = B3LYP-GD3 > MP2/Müller, (ii) REs in the computed interaction energy and its classical component are comparable and typically within 5% of the expected values at B3LYP, (iii) REs in the computed XC-term are an order of magnitude larger, between 40-55% of the expected values at B3LYP, and (iv) B3LYP-generated data systematically under-estimates the expected values of the interaction energy and its components.

This is very encouraging when a study of week interactions goes. Hence, we decided to perform an extensive analyses of all distant C···H, O···H and H···H interactions in all conformers and levels of theory as (i) this should provide an additional insight on how LoT investigated might perform in terms of evaluating common intramolecular interactions and, in addition, (ii) such information might be of importance for researchers who are involved, e.g. in the design of force field(s) needed in modelling of large(r) molecules of biological importance. For those interested in exact values, relevant computed data is included in Tables S71-S106 in the ESI. Based on these data, the averaged absolute REs in the interaction energy and its components computed for the distant C···H, O···H and H···H intramolecular interactions are presented in Table 5. In general, trends discussed for the covalent bonds and intramolecular interactions are also observed in this case: (i) REs are smallest at MP2/Müller for all three energy terms shown in
Table 5, (ii) much larger REs were obtained at HF whereas (ii) both modes of B3LYP generated quite trustworthy data with REs typically smaller than twice the values obtained at MP2/Müller.

Table 5  Averaged, for the combined data presented in Tables S71-S106 in the ESI, absolute relative errors in the interaction energy and its components computed for the distant C--H, O--H and H--H intramolecular interactions in all conformers of glycol at the indicated levels of theory.\(^a\)

<table>
<thead>
<tr>
<th></th>
<th>HF</th>
<th>B3</th>
<th>GD3</th>
<th>MP2</th>
<th>HF</th>
<th>B3</th>
<th>GD3</th>
<th>MP2</th>
<th>HF</th>
<th>B3</th>
<th>GD3</th>
<th>MP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avr:</td>
<td>13.61</td>
<td>3.76</td>
<td>3.85</td>
<td>1.89</td>
<td>13.47</td>
<td>3.73</td>
<td>3.82</td>
<td>1.68</td>
<td>0.24</td>
<td>0.18</td>
<td>0.18</td>
<td>0.47</td>
</tr>
<tr>
<td>StDev</td>
<td>9.42</td>
<td>2.07</td>
<td>2.14</td>
<td>0.48</td>
<td>9.18</td>
<td>2.23</td>
<td>2.30</td>
<td>0.51</td>
<td>0.16</td>
<td>0.10</td>
<td>0.10</td>
<td>0.22</td>
</tr>
<tr>
<td>Avr:</td>
<td>8.49</td>
<td>3.08</td>
<td>3.12</td>
<td>1.59</td>
<td>9.09</td>
<td>2.89</td>
<td>2.94</td>
<td>1.98</td>
<td>1.03</td>
<td>0.55</td>
<td>0.53</td>
<td>0.39</td>
</tr>
<tr>
<td>StDev</td>
<td>1.80</td>
<td>1.11</td>
<td>1.11</td>
<td>0.65</td>
<td>2.27</td>
<td>0.72</td>
<td>0.74</td>
<td>0.57</td>
<td>0.61</td>
<td>0.17</td>
<td>0.16</td>
<td>0.12</td>
</tr>
<tr>
<td>Avr:</td>
<td>2.42</td>
<td>0.86</td>
<td>0.87</td>
<td>0.73</td>
<td>2.40</td>
<td>0.84</td>
<td>0.85</td>
<td>0.64</td>
<td>0.19</td>
<td>0.10</td>
<td>0.10</td>
<td>0.12</td>
</tr>
<tr>
<td>StDev</td>
<td>1.62</td>
<td>0.70</td>
<td>0.71</td>
<td>0.47</td>
<td>1.65</td>
<td>0.70</td>
<td>0.71</td>
<td>0.52</td>
<td>0.23</td>
<td>0.11</td>
<td>0.11</td>
<td>0.12</td>
</tr>
</tbody>
</table>

\(^a\) All values are in kcal mol\(^{-1}\). B3 = B3LYP. GD3 stands for B3LYP-GD3 and MP2 = MP2/Müller.

Finally, let us briefly discuss changes in the diatomic interaction energies obtained at the CCSD/BBC1 level and compare them with relevant data obtained at lower levels investigated in this work. Data obtained for the classical and the XC components are presented in extensively commented Figs. S9-S10 in the ESI. Focusing on Lin → LEC (Fig. 10a), due to the formation of intramolecular interaction, the \(\Delta E_{\text{int}}^{\text{O4}, \text{H6}}\) term dominates. At CCSD/BBC1, \(E_{\text{int}}^{\text{O4}, \text{H6}}\) has changed from −50.4 (in Lin) to −90.9 (in LEC) kcal mol\(^{-1}\); hence, it became more attractive by −40.5 kcal mol\(^{-1}\) and has contributed most (among all unique 45 diatomic interactions) to LEC stability. Furthermore, the interactions between covalently bonded C- and O- atoms have also changed significantly; \(E_{\text{int}}^{\text{C1}, \text{O4}}\) strengthened (became more negative) by −7.5 kcal mol\(^{-1}\) (note that O4 is a proton accepter in the O3–H6···O4–C1 interaction) whereas an opposite is observed for \(E_{\text{int}}^{\text{C2}, \text{O3}}\) that weakened by +8.0 kcal mol\(^{-1}\) to facilitate charge transfer to interatomic region between O4 and H6. It is then obvious that, in this particular case, formation of the intramolecular interaction
has significantly affected 5-atom molecular region, C2–O3–H6–O4–C1. Interestingly, similar phenomenon was observed on the formation of a ‘proper’, red-shifted and QTAIM-bonded intramolecular H-bond in β-alanine. This indicates that the same physical process(es) might take place on the formation of QTAIM-(non)bonded intramolecular H-bonds.

Besides most prominent and of stabilizing nature \( \Delta E_{\text{int}}^{O4,H6} \), there are additional nine diatomic (non)bonded interactions that strengthened in the LEC, hence added to this conformer stability (see Table S25 in the ESI). Their \( \Delta E_{\text{int}}^{A,B} \) values vary between \(-2\) and \(-13\) kcal mol\(^{-1}\) and these interactions mainly involve atoms of the H5O4C1–C2O3H5 fragment of glycol. Among them are the non-bonded and highly attractive interactions between the C3,H5 and C2,O4 atom pairs (with \( E_{\text{int}}^{C3,H5} \) and \( E_{\text{int}}^{C2,O4} \) of \(-50.4\) and \(-91.8\) kcal mol\(^{-1}\), respectively, in Lin) that became even stronger in the LEC, respectively, by \(-13.1\) and \(-10.6\) kcal mol\(^{-1}\) (the second and third most prominent change among all attractive in nature \( \Delta E_{\text{int}}^{A,B} \) values).

To complete the picture, there are also (non)bonded interactions that have changed in unfavourable manner on Lin \( \rightarrow \) LEC and for seven of them \(+2 < \Delta E_{\text{int}}^{A,B} < +40\) kcal mol\(^{-1}\) – Table S25 in the ESI. Among them is the O3,O4 atom pair that experienced the largest increase in the repulsive interaction, by +33.9 kcal mol\(^{-1}\), that is nearly twice as much as found for the
Lin → Ecl transformation where these two atoms are directly involved in the steric clash – Fig. 10b. One wonders how many chemists at large would predict that correctly? Considering $\Delta E_{\text{int}}^{01,04} = +18.0$ kcal mol$^{-1}$ obtained on Lin → Ecl – Fig. 10b – (from +103.2 (in Lin) to +121.2 (in Ecl) kcal mol$^{-1}$) one must stress that it is most significant among all 45 atom pairs – see Table S28 in the ESI. The second most unfavourable change took place between covalently bonded C-atoms. This C–C bond became weaker (one might say strained in Ecl) as $E_{\text{int}}^{\text{C1,C2}}$ became less negative, by +8.9 kcal mol$^{-1}$; note that opposite took place on Lin → LEC where the C1–C2 covalent bond became stronger, by −1.3 kcal mol$^{-1}$, in the LEC (Table S25 in the ESI).

All the above nicely illustrates that on a structural rearrangement (it results just in different placement of atoms relative to each other in 3D space) many and very significant changes can and indeed take place. Hence, focusing on just one obvious, from visual inspection, diatomic interaction does not provide a full picture and may lead to wrong conclusions, such as, e.g. the case of the QTAIM-bonded H⋯H interaction in cis-2-butene,$^{31}$ as demonstrated from molecular-wide analyses of this interaction.$^{24,32}

Most of these changes are rather impossible to predict just from the eye inspection of the structures. At the same time, they are so interesting and provide a wealth of invaluable information generated from the use of the IQA data obtained at the CCSD/BBC1 level. The question one then can ask is whether the same picture could be recovered at lower LoT investigated in this work. An inspection of data shown in Fig. 10 provides an affirmative answer; all major changes discussed above were reproduced in terms of trends exactly and, importantly, highly comparable (against CCSD/BBC1) value-wise data were obtained at remaining (lower) LoT. Relative to CCSD/BBC1 data, a full set of errors in changes in the interaction energy terms ($\Delta E_{\text{int}}^{AB}$, $\Delta V_{\text{el}}^{AB}$ and $\Delta V_{\text{xc}}^{AB}$) on Lin → LEC and Lin → Ecl is shown in Figs. S11-S12 in the ESI and the analyses of errors are included in Tables S107-S112 in the ESI.
4.6. FAMSEC-based analysis of intramolecular interactions

Despite large (at HF) and significant (at B3LYP and MP2/Müller) REs in IQA-defined energy components observed for numerous atoms, the computed at CCSD/BBC1 changes in these energies on structural transformations of glycol were in most cases reproduced (i) exactly considering trends and (ii) the CCSD/BBC1 values, typically, below 10%-errors (except HF data). It is then not entirely surprising that, regardless of a LoT, most commonly used FAMSEC descriptors (namely: $\Delta E_{\text{sel}}^G$, $\Delta E_{\text{int}}^G$, $E_{\text{attr-loc}}^G$, $\Delta E_{\text{int-attr}}^{G-H}$, and $E_{\text{attr-mol}}^G$) that, for simplicity, will be marked in relevant Figs. as 1, 2, 3, 4, and 5, respectively) show very much the same trends and highly comparable values (Fig. 11).

Hence, FAMSEC-based description and interpretation of, e.g. the O4⋯H6 interaction in the LEC, should be fundamentally sound regardless of a level of theory used; values obtained at B3LYP will be used for illustration purposes. Considering the two-atom $G = \{O4,H6\}$ fragment (it is a part of commonly used X⋯H⋯Y$^2$ or D⋯H⋯A$^{17}$ notations, where D and A stand for a proton donor and acceptor, respectively) it is seen that its self-energy, $\Delta E_{\text{sel}}^G$ (1 in Fig. 11a) has slightly increased (by 5.1 kcal mol$^{-1}$) and this is typically observed for atoms directly involved in intramolecular H-bonding.$^{22,25,33}$ The O4 and H6 atoms are involved in highly attractive
interaction in both conformers (−48.4 and −86.1 kcal mol⁻¹ in Lin and LEC, respectively). Their intra-fragment interaction energy, ΔEᵦₒₑᵦintros (2 in Fig. 11a) has changed in a stabilizing manner when in the LEC (by −37.7 kcal mol⁻¹) that compensated over the ΔEᵦₒₑᵦ.self term. As a result, this fragment became stabilised in LEC due to Eattr-loc being highly negative, −32.6 kcal mol⁻¹, (3 in Fig. 11a). However, largely less favourable interactions between atoms of G and atoms of fragment ℨ (all atoms of glycol but H6 and O4) are at all levels predicted to be over +55 kcal mol⁻¹ (less attractive) relative to the same interactions in Lin (ΔEᵦₒₑᵦˌ, 4 in Fig. 11a). For instance, at B3LYP the inter-fragment interaction energy, Eᵦₒₑᵦˌ, computed for Lin and LEC was −867.0 and −810.7 kcal mol⁻¹, respectively; hence, ΔEᵦₒₑᵦˌ of +56.3 kcal mol⁻¹ was obtained on the Lin → LEC structural change (+56.1 kcal mol⁻¹ was obtained at CCSD/BBC1). As a consequence, this fragment has contributed to molecular energy of LEC in destabilizing manner with Eattr-mol of +23.6 kcal mol⁻¹ (5 in Fig. 11a).

Relative to the interpretation of the {H6⋯O4} fragment, four FAMSEC descriptors (2 to 5) have opposite trends at all LoTs when a 4-atom O3–H6⋯O4–C1 fragment was selected according to the IUPAC recommendation, X–H⋯Y–Z₃₀ — Fig. 11b. The intrafragment interactions changed somewhat unfavourably, from −674.8 in Lin to −670.1 kcal mol⁻¹ in the LEC. Hence, ΔEᵦₒₑᵦˌ = +4.7 kcal mol⁻¹ (2 in Fig. 11b) and, in combination with ΔEᵦₒₑᵦ.self > 0 (1 in Fig. 11b), made this fragment slightly destabilized in the LEC on the intramolecular H-bonding formation as Eattr-loc > 0 (3 in Fig. 11b). Importantly, the inter-fragment interactions (Eᵦₒₑᵦˌ of −1092.9 and −1113.9 kcal mol⁻¹ at B3LYP in Lin and LEC, respectively) became stronger by −21.0 kcal mol⁻¹ (ΔEᵦₒₑᵦˌ, 4 in Fig. 11b) that is more significant than the Eattr-loc term of +7.7 kcal mol⁻¹. As a result, the UIPAC representation of the intramolecular H-bonding fully recovered a
general notion of stabilizing energy contribution made to a molecule as reflected by $E_{\text{attr-mol}}^G < 0$ (5 in Fig. 11b).

Data obtained at the B3LYP and B3LYP-GD3 levels are of special interest due to relatively low cost of the IQA calculations. Averaged data obtained from CCSD/BBC1, B3LYP and B3LYP-GD3 for both interactions, O4···H6 and O3···O4, using 2- and 4-atom fragments is included in Table 6. It shows that B3LYP-generated data differ from CCSD/BBC1 well within ±3 kcal mol⁻¹.

**Table 6** Averaged and combined from the CCSD/BBC1, B3LYP and B3LYP-GD3 levels FAMSEC descriptors (in kcal mol⁻¹) used to characterise the O4···H6 (in LEC) and O3···O4 (in Ecl) interactions

<table>
<thead>
<tr>
<th>$\Delta E_{\text{self}}^G$</th>
<th>$\Delta E_{\text{int}}^G$</th>
<th>$E_{\text{attr-loc}}^G$</th>
<th>$\Delta E_{\text{int}}^{G,H}$</th>
<th>$E_{\text{attr-mol}}^G$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G = {O4,H6}$ in LEC</td>
<td>4.8±0.3</td>
<td>-38.6±1.7</td>
<td>-33.7±2.0</td>
<td>56.2±0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G = {C1,O3,O4,H6}$ in LEC</td>
<td>2.2±1.0</td>
<td>4.3±1.4</td>
<td>6.5±2.3</td>
<td>-21.6±1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-15.1±3.6</td>
</tr>
<tr>
<td>$G = {O3,O4}$ in Ecl</td>
<td>1.0±2.3</td>
<td>18.6±0.5</td>
<td>19.6±2.8</td>
<td>-9.3±2.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10.3±0.9</td>
</tr>
<tr>
<td>$G = {C1,O3,O4,H6}$ in Ecl</td>
<td>2.0±2.8</td>
<td>8.5±1.1</td>
<td>10.5±1.7</td>
<td>5.6±0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16.1±1.5</td>
</tr>
<tr>
<td>$G = {O3,O4,H5,H6}$ in Ecl</td>
<td>-1.8±2.7</td>
<td>11.7±0.1</td>
<td>9.9±2.6</td>
<td>2.6±1.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12.5±1.1</td>
</tr>
</tbody>
</table>

Since IQA calculations were performed on entire molecules, we decided to expand FAMSEC-based analysis in order to gain a deeper insight on interpretability of data generated at different LoTs. A full set of results obtained at all levels of theory on $\text{Lin} \rightarrow \text{LEC}$ and $\text{Lin} \rightarrow \text{Ecl}$ structural changes for selected fragments containing meaningful functional groups, namely (i) 2-atom fragments: $\{O3,O4\}$, $\{O4,H5\}$, $\{O3,H6\}$, $\{O4,H6\}$, and $\{C1,C2\}$, (ii) 3-atom fragments: $\{O3,O4,H6\}$, $\{C1,H7,H9\}$ and $\{C2,H8,H10\}$ and (iii) 4-atom fragments: $\{O3,O4,H5,H6\}$ and $\{C1,O3,O4,H6\}$ are graphically presented in Figs. S13-S18 in the ESI. Moreover, relevant computed values of all FAMSEC descriptors are included in Tables S113-S118 in the ESI.
The extended analysis of error (relative to the CCSD/BBC1 data) in the FAMSEC-defined energy terms that includes all ten indicated above meaningful 2-, 3- and 4-atom functional groups of glycol is presented in Table S119 in the ESI. It shows that FAMSEC descriptors computed at the B3LYP and B3LYP-GD3 levels for the Lin → LEC and Lin → Ecl structural changes differ from values obtained at the CCSD/BBC1 level, on average, by 2.0±2.1 kcal mol⁻¹. Somewhat smaller an overall averaged error of 1.5±1.5 kcal mol⁻¹ was obtained at the MP2/Müller combination and poorer performance of HF is reflected by significantly larger errors and uncertainty, 4.8±4.8 kcal mol⁻¹. It is then clear that quite outstanding performance of B3LYP functional must be attributed to its explicit implementation in AIMAll in combination with cancelling of systematic REs in IQA-defined energy terms when a comparative analysis (such as FAMSEC) is performed.

5. Conclusions

There is a growing interest in using most popular B3LYP functional in computing interacting quantum atoms (IQA) data due to their usefulness in describing fundamental properties of atoms as well as the strength and nature of interactions/bonds and how they vary with a change in a chemical environment. However, reliability of, e.g. B3LYP/IQA energy terms (generated from the DFT wavefunction) in describing interactions is unknown.

To address this critical shortcoming, a thorough and extensive investigation of selected IQA- and FAMSEC-defined descriptors (they are very useful in interpreting and characterizing inter- and intramolecular interactions) was performed at HF, B3LYP, B3LYP with Grimme’s empirical D3 dispersion correction and MP2/Müller levels of theory (LoT) using default (auto) settings for the IQA calculations in most popular dedicated AIMAll software. Our focus was on (i) diatomic interaction energy, $E_{\text{int}}^{A,B}$, and its components, $V_{\text{el}}^{A,B}$ and $V_{\text{XC}}^{A,B}$, as well as (ii) self- and additive atomic energies, $E_{\text{self}}^{A}$ and $E_{\text{add}}^{A}$, respectively.
As a case study, three conformers of glycol were investigated because we could make use of data obtained from the CCSD/BBC1/TWOe combination\textsuperscript{11} as a suitable reference for:

(1) A \textit{direct} comparison (against reference data) of \textit{expected} at lower LoTs IQA energy terms needed to evaluate relative errors (REs) in computed IQA energy terms for the same atoms, molecular fragments and interactions in different molecular environments (conformers) at lower LoTs, and

(2) Evaluating usefulness of FAMSEC-defined energy terms (FAMSEC = fragment attributed molecular system energy change\textsuperscript{22}) used to explain and pointing at the origin of relative stability of these conformers. Hence, reliability in \textit{changes}, relative to a reference structure, \textit{ref} (linear conformer, \textbf{Lin}) obtained on structural transformation to a final molecular structure (\textit{fin}) either the highest energy (eclipsed, \textbf{Ecl}) or lowest energy conformer, \textbf{LEC}, of glycol was investigated.

The REs (they quantify inaccuracy of the LoT/IQA energy terms) revealed that:

(a) The self-molecular energy $E_{\text{self}}^{\text{Tot}}$ was overestimated at HF by staggering $-364.6$ kcal mol$^{-1}$ whereas the total interaction energy, $E_{\text{int}}^{\text{Tot}}$, was underestimated by the same value. By contrast, nearly an order of magnitude smaller REs in $E_{\text{self}}^{\text{Tot}}$ and $E_{\text{int}}^{\text{Tot}}$ ($\Delta_{\text{self}}^{\text{Tot}}$ and $\Delta_{\text{int}}^{\text{Tot}}$, respectively) were found at B3LYP; $E_{\text{self}}^{\text{Tot}}$ and $E_{\text{int}}^{\text{Tot}}$ were under- and overestimated, respectively, by the same value. Considering MP2/Müller, $E_{\text{self}}^{\text{Tot}}$ was overestimated (too negative) by $-170$ kcal mol$^{-1}$.

(b) Individual self-atomic energies, the $E_{\text{self}}^{A}$ values, were underestimated (too positive) and overestimated by up to $+80$ and $-50$ kcal mol$^{-1}$ at HF and MP2/Müller; B3LYP reproduced the expected values with REs of up to $+15$ kcal mol$^{-1}$. In all cases unidirectional REs in the $E_{\text{self}}^{A}$ values for all atoms are observed.

(c) Except the C–C backbone, all diatomic interactions, $E_{\text{int}}^{A,B}$, computed for covalently bonded atoms were largely overestimated at HF. For instance, in the case of the C–O bond the
averaged RE in the $E_{\text{int}}^{A,B}$ values for three conformers was about $-114$ kcal mol$^{-1}$ that is nearly 24% more than the expected value. An order of magnitude smaller RE was obtained at B3LYP. Combining all possible distant C···H, O···H and H···H intramolecular interactions in three conformers of glycol, the averaged absolute REs of $13.6\pm9.4$, $3.8\pm2.1$ and $1.9\pm0.5$ kcal mol$^{-1}$ in $E_{\text{int}}^{A,B}$ and $13.5\pm9.2$, $3.7\pm2.2$ and $1.7\pm0.5$ kcal mol$^{-1}$ in the classical term, $V_{\text{cl}}^{A,B}$, were obtained at HF, B3LYP and MP2/Müller, respectively.

It is important to stress that those REs (their values and directionality) in interaction energy terms depend on atoms involved as well as the LoT. Furthermore, in most instances, the main contributor to uncertainty in the computed interaction energies were REs in the classical term, $V_{\text{cl}}^{A,B}$.

Interestingly, data obtained at B3LYP for the intramolecular interactions of particular interest in this work, O4···H6 in the LEC and O3···O4 in the Ecl conformer, can be seen as perfectly suited for a general description as they are fully comparable with the picture generated at CCSD/BBC1. For instance, $E_{\text{int}}^{O4,H6}$ and $E_{\text{int}}^{O3,O4}$ of $-86.1$ and $+121.4$ kcal mol$^{-1}$ were underestimated by 5 and 0.1 kcal mol$^{-1}$, respectively. Furthermore, $E_{\text{int}}^{O4,H6}$ and $E_{\text{int}}^{O3,O4}$, when computed in all conformers, regardless of the interatomic distance between atoms of interest, carried averaged $-4.8$ and $-1.3$ %-errors, respectively, at B3LYP. It is important to note that, relative to CCSD/BBC1, B3LYP-generated data systematically underestimates the expected values of the intramolecular interaction energy and its components.

Exceptionally reliable data were obtained at B3LYP when changes in atomic energies (self and additive) as well interaction energy terms were computed on two structural changes, Lin $\rightarrow$ LEC and Lin $\rightarrow$ HEC. This is exemplified by:

(a) Using $\Delta E_{\text{add}}^{A}$, a reliable and invaluable insight on most (de)stabilized molecular fragments was obtained (this applies to all LoT tested). This, in turn, can be used not only to explain the
change in the molecular energy but also to point at specific molecular region the energy of which has decreased or increased most on the ref → fin structural transformation.

(b) The FAMSEC-based interpretations of the O4⋯H6 and O3⋯O4 interactions and their role in relative stability of the glycol conformers; the qualitative description was exact and quantitative data perfectly comparable with those obtained at the CCSD/BBC1 level.

In order to gain a deeper insight on quality of B3LYP/IQA data, FAMSEC analyses were performed on five 2-atom, three 3-atom and two 4-atom fragments containing meaningful functional groups of glycol (including the O4⋯H6 and O3⋯O4 interactions) for the Lin → LEC and Lin → Ecl structural changes. It has been found that averaged B3LYP/IQA and B3LYP-GD3/IQA data differ from CCSD/BBC1/IQA values by 2.0±2.1 kcal mol⁻¹ (1.5±1.5 and 4.8±4.8 kcal mol⁻¹ was found for MP2/Müller/IQA and HF/IQA, respectively). To further illustrate reliability of the B3LYP data, averaged values from CCSD/BBC1/IQA, B3LYP/IQA and B3LYP-GD3/IQA for \( \Delta E_{\text{self}}^G \), \( \Delta E_{\text{int}}^G \), \( E_{\text{attr}}^G \), \( \Delta E_{\text{int}}^{G,\text{attr}} \), and \( E_{\text{attr}}^G \) (where \( G = \{O4,H6\} \) in LEC) were found to be 4.8±0.3, −38.6±1.7, −33.7±2.0, +56.2±0.1, and +22.5±2.1 kcal mol⁻¹, respectively. Clearly, this quite outstanding performance of B3LYP functional must be attributed, at least partly, to its explicit implementation in AIMAll in combination with cancelling of systematic relative errors when a comparative analysis (such as FAMSEC) is performed.

Finally, it is documented that if recommended \( E = E_{\text{IQA}} \) equality is not met it does not necessarily disqualifies LoT/IQA data entirely. To support this, best performance was obtained at MP2/Müller for interaction energies even though \( E_{\text{IQA}} \) largely overestimated \( E \). This study has fully explained this by identifying the origin of \( E_{\text{IQA}} \neq E \) inequality; it resides nearly entirely in the self-atomic energies due to the Müller approximation.

Importantly, however, this work has demonstrated that meeting the \( \Delta E_{\text{IQA}} = \Delta E \) criterion proposed by us recently¹¹ (it is representing changes between the final and any suitably selected
initial structure of a molecular system) ‘validates’ reliability of changes in the above specified IQA-defined energy terms. Hence, one can perform a comparative analysis, such as FAMSEC, with confidence; B3LYP/aug-cc-pVDZ/IQA-based interpretation of many interactions produced exact qualitative explanation and, quantitatively, perfectly comparable with CCSD/BBC1/IQA data.

To conclude, it is our hope that this work will encourage chemists at large to confidently make use of IQA data obtained at B3LYP to gain a fundamental insight on how properties of atoms as well as the strength and nature of interactions/bonds vary with a change in a chemical environment due to, e.g. a conformational transformation, additional functionality in a molecule or along a reaction pathway. It is important to stress that the solvent effects and an empirical dispersion corrections, such as that by Grimme (e.g., GD3) are not accounted for in the IQA calculations (as far as we know not only in AIMAll). However, results of this work suggest that inclusion of GD3 has insignificant impact on overall quality or reliability of the B3LYP/IQA energy terms. Clearly, more work has to be done to explore suitability of B3LYP/IQA data further.

Acknowledgements
This work is based on the research supported in part by the National Research Foundation of South Africa (Grant Number 105855). The authors gratefully acknowledge the Centre for High Performance Computing (CHPC), South Africa, for providing computational resources to this research project. Constructive and informative comments made by Dr Keith are greatly acknowledged.

Statement:
There is no conflict of interest
References


