J Comput Aided Mol Des (2010) 24:473–474
DOI 10.1007/s10822-010-9357-1

Overview of the perspectives devoted to tautomerism in molecular design

Yvonne Connolly Martin

Received: 9 April 2010 / Accepted: 13 April 2010 / Published online: 29 April 2010

Abstract This communication summarizes the important points made in each contribution. They show that there remain issues in the cheminformatic handling of tautomers and predicting the relative energies of various tautomers.

Keywords Tautomers · Quantum chemistry · Cheminformatics · Structure searching · Crystal structures · Drug design

This Perspectives issue of Journal of Computer-Aided Molecular Design presents various invited viewpoints on tautomerism as it relates to molecular design. It is not meant to be complete, but rather to provide a glimpse into what some experts view as the current state tautomerism as it pertains to biological issues, particularly of importance to drug discovery. Two who have thought especially long and hard about this issue are not represented: Peter Taylor and Peter Kenny. This issue is dedicated to them in respect to their important contributions to this field.

In the lead article of this issue, Katritzky et al. review the important types of annular tautomerism—this provides a context for the following articles. The concepts are illustrated with several frequently prescribed drugs. They also remind us that there are some drugs that theoretically tautomerize, but that at the pH's of biological interest they ionize with the result that tautomerism is not an issue. The relationship between protonation and deprotonation and tautomerism is also discussed by Greenwood et al. and Sayle.

Y. C. Martin (✉)
Martin Consulting, Waukegan, IL, USA
e-mail: yvonnecmartin@comcast.net

Roger Sayle discusses tautomerism in the context of mesomerism, resonance, protonation state, and aromaticity. His contribution reflects his extensive experience with the issues of tautomerism in cheminformatics, docking, and property calculations and includes many interesting examples of the difficulities of representing particular molecules as a Lewis structure. Although most software enumerates tautomers using rules, he describes a more holistic approach that places the specific number of protons on a topological scaffold of the heavy atoms.

Wendy Warr reviews how many cheminformatics software vendors handle tautomerism. She notes that there is no consensus on such issues as: How are structures represented in a database—is it as a unique tautomer, the entered tautomer, or all tautomers? How does tautomerism affect substructure searches? Which structure should be shown if a tautomer of the registered molecule matches the query? Which tautomers are considered.?

Sitzmann et al. also provide an extensive investigation on the cheminformatic aspects of tautomerism and show that because different software systems register tautomers differently, approximately 10% of compounds registered in one database are registered as a different tautomer in another database. They found that using their rules more than 2/3 of the molecules their database of existing compounds can tautomerize.

Is tautomerism an important issue in drug design? Several authors present different evidence. Cruz-Cabeza et al. show that all 136 PDB complexes of indazoles and all 157 complexes of pyrimidones bind as the low-energy tautomer. In contrast, Bill Porter discusses the example of warfarin, which can exist in 40 topologically different structures. In particular, the often overloopked ring-chain tautomerism (and not addressed by any other author) alone produces eight different ring stereoisomers. The various

warfarin tautomers are recognized by different enzymes and are calculated to have pK_as that vary by ten orders of magnitude and log D values that vary by 3.4! Rodger Henry provides examples of small molecule crystal structures of very similar drug-like molecules that crystallize as different tautomers. Several of the authors (Martin et al. and Clark) remind us that binding a rare tautomer incurs a penalty in affinity.

The Cambridge Structural Database (CSD) provides experimental information on the frequency of different tautomers in different environments. Rodger Henry provides two examples of such switches. Cruz-Cabeza show that the CSD contains no examples of crystal structures of 37 of their 84 tautomerizable heterocyclic ring systems. Of those found, the lowest energy tautomer was observed unless the relative energy difference between tautomers was small.

Tautomer ratios are related to the differences in pK_as of the corresponding structures. Because of the success in predicting pK_as, this suggests that empirical methods can be used to predict tautomer ratios. For example, Greenwood et al. present their strategy of using Hammett-Taft equations to estimate tautomer energy. However, Tim Clark takes a different view and argues that a tautomer-aware force-field would be more appropriate. Both groups use quantum chemical methods as well as observed pK_as to derive the parameters for their approach. This Perspective issue is not comprehensive, in that others calculate tautomer ratios using approaches not presented in this set of papers [1–4].

Because of the relationship between pK_a and tautomer ratio, I present simple model-based equations for QSAR of a series of related molecules using Hammett σ to account for tautomerisation.

Dick Cramer presents the argument that because any difference in tautomer ratio of related compounds is related to fundamental differences in the properties of the molecules (for example, electronic structure), topomer CoMFA automatically corrects for tautomer differences. It does so because it follows strict rules for structure consistency of the common parts of the analogues that in turn emphasize the difference between them, which in turn might be the tendency to tautomerize.

Several authors discuss using quantum chemical methods to calculate tautomer ratios. Andreas Klampt reviews the results of using COSMO RS, which is based on BP/TZVP-DFT calculations, for the SAMPL blind prediction of tautomer energies. Applying the MP2 correction reduced the average prediction error to 1.2 kcal/mol. However, an empirical correction applicable to molecules that can undergo keto-enol tautomerism reduced the RMSE to 0.61 kcal/mol. Cruz-Cabeza et al. compare the energy differences calculated with MP2/6-311++G** using a polarisable continuum model with the relative abundance of the tautomers in the Cambridge Structural Database. They find that in seven of eight cases the lower energy tautomer was also the most abundant in that database. The exception is 2H-1,2,3-triazole, which is calculated to be 4.0 kJ/mol (0.96 kcal/mol) more stable than the 1H-1,2,3-triazole tautomer, but is found 11 times in the CSD compared to 27 for its tautomer. On the other hand, Jeremy Greenwood et al. provide two examples to support their conclusion that DFT M06-2X/aug-cc-pVTZ9(-f) with PBSCRF correctly predicts the relative distribution of tautomers. In contrast, Tim Clark provides no opinion as to whether he would use DFT or ab initio methods to calculate tautomer ratios. Haranczyk studies the relative stability of tautomers anionic nucleic acid bases, species are important in radiation-induced mutagenesis. After generating all possible tautomers, the most stable are identified by energy-based screening at the DFT/B3LYP/631++G** level followed by energy refinement at the MP2/AVDZ level, all gas-phase calculations. He also shows how cheminformatic approaches decrease the number of structures that need to be optimized. In summary, although quantum chemical calculations provide much insight into the relative energies of tautomers, there appears to be no consensus on the optimal method. Additionally is 1 kcal/mol as close as we can get? Is this close enough?

The articles illustrate that the problem of tautomers in drug design is no means settled. Convergence of cheminformatic issues would ensure that databases from different sources are compatible and that structure searching would identify the appropriate compounds. The accurate prediction of tautomer ratio should improve the ability to forecast biological potency—whether such an advance will come from quantum chemistry or more empirical methods is as yet unknown.

References

1. Milletti F, Storchi L, Sforna G, Cross S, Cruciani G (2009) J Chem Inf Model 49:68–75
2. ten Brink T, Exner TE (2009) J Chem Inf Model 49:1535–1546
3. http://www.chemaxon.com/marvin/help/chemicalterms/EvaluatorFunctions.html#isomers_functions
4. http://www.molecular-networks.com/files/docs/tautomer/