sciendo

Atomic partial charge model in chemistry: chemical accuracy of theoretical approaches for diatomic molecules

Matej Uhliar

Institute of Physical Chemistry and Chemical Physics, Slovak University of Technology in Bratislava, Radlinského 9, SK-812 37 Bratislava, Slovakia
matej.uhliar@stuba.sk

Abstract: Atomic partial charges cannot be physically measured but they play a significant role in many chemical theories and theoretical models. Therefore, they are, evaluated from experimentally acquired properties or calculated by quantum chemistry computational methods. This study is focused on determining chemical accuracy of various theoretical methods of computing atomic partial charges based on quantum chemistry. Values of gas-phase atomic partial charges were acquired by Mulliken (MUL) population analysis, natural bond analysis (NBO), Merz-Singh-Kollman (MSK) scheme, and atomic polar tensor (APT) charges computed considering Density Functional Theory and *ab initio* Møller-Plesset up to the second order levels. Correlations between the calculated values were determined by principal component analysis (PCA) and further confirmed by linear regression. The best agreement between experimentally evaluated atomic partial charges and theoretical values was obtained with the MSK scheme.

Keywords: electric dipole moment; atomic charges; Mulliken charges; principal component analysis

Introduction

Physical properties of compounds and their chemical reactivity are affected by their molecular as well as electronic structure. In case of aromatic organic molecules, their molecular structure can be described using various aromaticity descriptors, especially popular are HOMA-based Harmonic Oscillator Model of Electron Delocalization (HOMED) (Raczyńska et al., 2010) and Harmonic Oscillator Model of Aromaticity for Heterocyclic Electron Delocalization (HOMHED) (Frizzo et al., 2012) which are based on theoretical or experimental X-ray, bond lengths, respectively. These indices seem to be suitable for aromaticity description of linear oligomers and their ring-fused analogues (Cagardová et al, 2020; Cagardová et al., 2022) as well as for large two-dimensional molecules, e.g., sunflower molecules (Cagardová et al., 2019; Cagardová et al., 2022). Bond length of relevant substituent groups also correlates very well with antioxidant activity (Klein et al., 2006), chemical reactivity (Klein et al., 2007), or biological activity (Štellerová et al., 2022). Description of the electronic structure is based on the knowledge of electron density distribution in the molecule or its changes during optical transitions, which opens a possibility to investigate several properties, i.e., atomic partial charges, total charge density, spin density for radicals, energies of selected frontier molecular orbitals, and energies of electronic transitions. Although, partial electric charges of atoms cannot be measured directly, they have a significant role in theoretical chemistry, i.e., atomic partial charges serve to estimate the reactivity of molecules (Klein et al., 2006; Mehmood et al., 2017), can help to determine structural differences between molecules and their conformers (Rostkowski et al., 2007), and be implemented for the prediction of pKa (Gross et al., 2002; Svobodová Vařeková et al., 2013; Ugur et al., 2014). Additionally, there are reports on good correlation of selected partial charges with toxicity of para phenols (Michalík et al., 2018a) and 3-hydroxynaphthalene-2-carboxanilides (Michalík et al., 2018b).

The physical concept of partial atomic charge is not uniquely defined. Classical electrostatics defines the partial electric charge of a diatomic polar molecule using the electric dipole moment (p), which is a measure of the separation of positive and negative electrical charges within a system. Two-point charges located on the atoms, one with positive ($+q$) and the other one with negative ($-q$) values, are separated by distance d. Absolute value of these charges is identical. Since the electric dipole moment and the interatomic distance are experimentally obtainable quantities, absolute value of the partial charge on the atom is determined as

$$|q| = p/d \tag{1}$$

In quantum chemistry, partial atomic charge can be evaluated from the electron wave function or electron density. From the historical point of view, the oldest method for atomic partial charge

determination is the Mulliken population analysis based on the molecular orbital scheme (Mulliken, 1955) based on the Hartree-Fock (HF) approach. Obtained Mulliken (MUL) partial charges are, in principle, sensitive to the basis set choice. Hypothetically, a complete basis set for a molecule can be obtained by locating a large number of atomic orbitals on a single atom, i.e., all electrons in the system would be assigned to that single atom. And thus, the method has no complete basis set limit as the exact value of MUL partial charge depends on the way the limit is approached. Such dependency results in incorrectly defined calculated charges. Additionally, the basis set convergence of charges does not exist and different basis set families may yield significantly different results (Thakkar, 2021). These problems can be overcome by newer methods for net atomic charge computing, e.g., electrostatic potential Merz-Singh-Kollman analysis (MSK) (Breneman et al., 1990) or natural bond orbital (NBO) analysis (Reed et al., 1985). Both these methods can be performed at the HF level, but also in calculations where correlation energy is considered, i.e., post HF approach or DFT.

In the MSK scheme suggested by U.C. Singh and P.A. Kollman (Singh et al., 1984), partial atomic charges are fit to reproduce molecular electrostatic potential (MEP) at a number of points around the molecule. Thus, MEP is calculated at gridpoints located on several layers around the molecule. Layers are constructed as an overlay of van der Waals spheres around each atom discarding all points located inside the van der Waals volume. The best results are achieved by sampling points not too close to the van der Waals surface and the van der Waals radii are therefore modified using scaling factors.

NBO analysis is based on the Lewis's structure picture, where the input atomic orbital basis set is transformed via natural atomic orbitals (NAOs) and natural hybrid orbitals (NHOs) into natural bond orbitals (NBOs). The obtained NBOs correspond to the widely used Lewis's picture, in which two-centre bonds and lone pairs are localized.

Among the population analysis schemes available in literature, atomic partial charges evaluated from generalized atomic polar tensor (GAPT) can be completely determined from experimental data, infrared intensities and frequencies, equilibrium dipole moments, and molecular geometries (Biarge et al., 1961; Person et al., 1974).

Although many theoretical works deal with the comparison of partial atomic charges calculated using various treatments of atomic molecules, systematic analysis statistically comparing chemical accuracy of standardly used population analysis has not been performed yet. Therefore, a comparative study of theoretical ab initio and DFT calculations of diatomic polar molecules is presented in this paper. Partial aims of this study are to calculate optimal geometries and atomic charges for 17 representative molecules. Chemical accuracy of obtained results will be compared with experimental values deduced from experimental gas-phase dipole moments and distances. Finally, the Principal Component Analysis (PCA) (Jolliffe, 2002) will be used to determine possible correlations between the applied treatments.

Calculation details

Quantum chemical computations were carried out using the Gaussian 16 program package (Frisch et al., 2016) using 6-31G, 6-31G* and 6-311++G(d,p) basis set (Hariharan et al., 1973). The calculations were performed at the ab initio MP2 (Møller-Plesset up to the second order) level where all electrons were correlated. Next, the density functional theory with hybrid meta exchange-correlation functional M0-62X (Zhao et al., 2008) and 3-parameter hybrid B3LYP functional (Becke, 1993; Lee et al., 1988; Vosko et al., 1980; Stephens et al., 1994) was used. B3LYP and M0-62X functionals are the most commonly used in calculations of inorganic and organic molecules. Additionally, all electrons were corelated in case of the MP2 approach.

Results and Discussions

Experimental gas-phase values for interatomic distances and dipole moments of studied diatomic molecules (Irikura, 2007; Huber et al., 2013; Nelson et al., 1967; Johnson III, 2022) were used to calculate the estimated experimental absolute partial charge (Exp) of the molecule; minimal value of 0.020e was found for CO, while maximal value of 0.878e was found for the LiF molecule. Physical constant e stands for elementary charge in this notation. The MUL, APT, NBO and MSK values of atomic partial charge were evaluated at DFT (B3LYP and M06-2X functionals) and MP2 levels. MUL charges were calculated at HF/MP2/6-311++G(d, p) level of theory in MP2 calculations as single point calculation throughout MP2, B3LYP and M06-X, using the 6-31G* basis set. MUL* values were computed in the same way as MUL' but applying the 6-31G basis set, as it is implied that a smaller basis set is more suitable for the calculations of electric properties of molecules.

The highest Exp value was found for the NaF molecule to be 0.878, whereas the lowest Exp value was 0.020 for the CO molecule. Experimentally acquired interatomic distances of these molecules are 1.93 Å and 1.13 Å, respectively. In case of the

MP2 approach, the highest values of atomic partial charges calculated are mostly for the NaF molecule obtained with MUL (0.838), MUL' (0.703), NBO (0.988), and MSK (0.874). The highest value of atomic partial charges determined using MUL* was for the NaCl (0.710), and with APT for the NO (1.194) molecule. On the other hand, the lowest values were computed for other molecules than CO: NO with MUL (0.025), CS with MUL* (0.019), NBO (0.069), and APT (0.031), BrCl with MUL', whereas MSK values showed lowest value of 0.000 for the CO molecule.

In case of the DFT-B3LYP approach, the highest value was found for NaF with MUL (0.757), NBO (0.978), and MSK (0.874), NaCl with MUL' (0.564) and MUL* (0.622), and for LiF with APT (0.860). The lowest values were obtained for CO with Mul (0.010), Mul* (0.013), and MSK (0.000), for BF with MUL' (0.097), and for BrCl with NBO (0.102) and with APT (0.103).

The highest values of atomic partial charges computed using the DFT-M0-62X method were found for the same molecules as with the DFT-B3LYP approach: for NaF at MUL 0.780, at MUL' 0.610, at NBO 0.978, and at MSK 0.874. A similar value of 0.618 obtained with MUL* was computed for the NaCl molecule and the value of 0.860 for the LiF molecule was found with the APT approach. The lowest values of atomic partial charges were computed for the CO molecule with the MUL (0.018) and MSK (0.000) approach. Secondly, the lowest value of 0.018 was calculated with MUL* for the BF molecule, while the BrCl molecule has the lowest values of atomic partial charges acquired with MUL' (0.029), NBO (0.095), and APT (0.103).

Calculated interatomic distances show similar values with mostly insignificant differences compared to experimentally obtained values of interatomic distances. For example, experimentally determined interatomic distance of the smallest HF molecule is 0.92 Å, and calculated interatomic distances of this molecule are 0.916 Å, 0.922 Å, and 0.918 Å at MP2, DFT-B3LYP and DFT-M0-62X levels of theory, respectively. Experimental value of interatomic distance for the largest molecule NaBr is equal to 2.50 Å. Yet, calculations at all considered levels of theory provide values of 2.528 Å at MP2, 2.526 Å at DFT-B3LYP, and 2.511 Å at DFT-M0-62X levels of theory. The largest differences in interatomic distances in a molecule were observed for the BrCl molecule, where experimental value is 2.32 Å while calculated values of interatomic distances are 2.164 Å, 2.192 Å, and 2.158 Å at MP2, DFT-B3LYP and M0-62X levels of theory, respectively.

Mutual correlation between individual values of atomic partial charges at various levels of theory was determined via PCA using the first (PC1) and second (PC2) principal components of each analysis since their coverage over the whole dataset of relations is 95.89 % for the B3LYP, 95.98 % for the M06-2X, and 92.37 % for the MP2 dataset. Chemical accuracy is then further estimated using linear regression between Exp and individual values. Exp values, in atomic unit, were calculated according to:

$$Exp = \left(\frac{p}{d}\right) / e \qquad (2)$$

Where Exp is the calculated value of atomic partial charge determined from experimentally obtained dipole moment p converted from Debeye to C·m, and experimentally determined interatomic distance d, and then divided by elementary charge e.

Tab. 1. Experimental gas-phase electric dipole moments (p), interatomic distances (d) taken from (Nelson et al., 1967; Johnson III, 2022), evaluated experimental absolute values of partial charges (atomic units) and Mulliken charges obtained from experimental geometry calculations.

Molecule	*p*/*D*	*d*/Å	Exp
HF	1.83	0.92	0.416
HCl	1.09	1.27	0.178
HBr	0.83	1.41	0.122
ClF	0.88	1.63	0.113
BrF	1.42	1.76	0.168
BrCl	0.52	2.32	0.047
NaF	8.12	1.93	0.878
NaCl	8.97	2.36	0.791
NaBr	9.09	2.50	0.756
LiF	6.28	1.56	0.836
LiCl	7.09	2.02	0.730
LiBr	7.23	2.17	0.693
BF	0.50	1.27	0.082
CO	0.11	1.13	0.020
SO	1.55	1.50	0.215
CS	1.96	1.53	0.266
NO	0.16	1.15	0.029

Out of all studied molecules, the most polar one is NaBr with experimentally obtained value of induced dipole moment of 9.09 *D* and the least polar molecule is CO with the dipole moment of 0.11 *D* as shown in Table 1. These two molecules have interatomic distances of 2.50 Å and 1.13 Å, respectively. Experimentally obtained data result in partial atomic charges of 0.756 for NaBr and 0.020 for CO.

Theoretical gas-phase interatomic distances and absolute values of Mulliken partial atomic charges are shown in Table 2. As it was mentioned before, individual theoretical values of interatomic distances slightly deviate from those calculated from experimental data. However, when Mulliken values of atomic partial charges are compared to Exp values, rather high deviations are observed. Yet, values of partial atomic charges found from theoretically optimized geometries show only small deviations from values obtained from single point calculations in experimental geometries.

Therefore, values of atomic partial charges are suggested to be mostly influenced by theoretical means of calculation, while they show little to no dependency on the geometry used; values of atomic partial charges obtained by various methods on selected levels of theory are shown in Table 3.

In comparison to the NaBr molecule (experimentally most polar) with its Exp value of atomic partial charge of 0.756, the MP2 level of theory provides the closest value of 0.762 with the MSK approach, while the highest deviation from the Exp values was observed for the NBO (0.924), followed by the MUL* approach (0.579). In case of the CO molecule, the smallest deviation from the Exp value (0.020) was determined again with the MSK approach (0.000), even though this value suggests that the CO molecule is not polar. The highest deviations were observed for the NBO (0.594) followed by the MUL' approach (0.307). Complete results of atomic partial charges with bond lengths of individual diatomic molecules obtained at the MP2 level of theory are shown in Table 3a.

Individual values of atomic partial charges vary in contrast to experimentally acquirable data. Even their differences depend on the studied molecule. This behaviour can be observed also for the B3LYP and M0-62X levels of theory. Compared to MP2 and DFT levels, MSK values of atomic partial charges show the least change, to the point they remain the same in most of the studied molecules. Results of the DFT-B3LYP level of theory for the NaBr molecule show the same value of 0.762 of atomic partial charge for the MSK approach, even though other approaches provide different values than those obtained at the MP2 level of theory. The highest difference in value from Exp was acquired with the MUL* approach (0.469), followed by the MUL' value (0.488) and the MUL value (0.566). In comparison to MP2 results in general, individual values of various approaches, aside from MSK values, changed together with their deviations from each Exp value. The most notable difference was observed for the radical NO molecule for APT partial charge at the MP2 level of computation; this behavior was not observed at the DFT level of theory.

Values obtained at the DFT-M0-62X level of theory show similar trends as those obtained at the DFT-B3LYP level of theory. Thus, MSK values show the smallest deviation from Exp values in most cases, while APT values also show rather small diffe-

Tab. 2. Comparison of gas-phase Mulliken values of partial atomic charges (atomic units) calculated using experimental (Exp) and optimized geometries at various levels of theory.

Molecule	Exp		B3LYP		M0-62X		MP2	
	d/Å	MUL	*d*/Å	MUL	*d*/Å	MUL	*d*/Å	MUL
HF	0.92	0.290	0.922	0.291	0.918	0.296	0.916	0.302
HCl	1.27	0.133	1.287	0.136	1.282	0.145	1.273	0.132
HBr	1.41	0.107	1.427	0.110	1.420	0.123	1.410	0.099
ClF	1.63	0.141	1.678	0.154	1.646	0.150	1.672	0.171
BrF	1.76	0.221	1.811	0.234	1.782	0.214	1.807	0.261
BrCl	2.32	0.011	2.192	0.025	2.158	0.035	2.164	0.031
NaF	1.93	0.757	1.954	0.757	1.936	0.780	1.972	0.838
NaCl	2.36	0.537	2.383	0.545	2.366	0.563	2.376	0.631
NaBr	2.50	0.558	2.526	0.566	2.511	0.585	2.528	0.648
LiF	1.56	0.619	1.583	0.627	1.576	0.647	1.595	0.702
LiCl	2.02	0.400	2.025	0.402	2.021	0.415	2.016	0.463
LiBr	2.17	0.423	2.179	0.426	2.179	0.441	2.179	0.484
BF	1.27	0.233	1.271	0.232	1.268	0.237	1.272	0.150
CO	1.13	0.011	1.128	0.010	1.122	0.018	1.139	0.105
SO	1.50	0.220	1.510	0.222	1.490	0.222	1.537	0.256
CS	1.53	0.248	1.539	0.241	1.528	0.196	1.539	0.049
NO	1.15	0.090	1.148	0.091	1.138	0.097	1.134	0.025

Tab. 3a. Comparison of gas-phase absolute values of partial atomic charges (atomic units) calculated using selected approaches at MP2 level of theory.

Molecule	MUL	MUL'	MUL*	*NBO*	*APT*	*MSK*
MP2						
HF	0.302	0.517	0.370	0.555	0.423	0.474
HCl	0.132	0.243	0.259	0.243	0.212	0.258
HBr	0.099	0.256	0.213	0.168	0.142	0.195
ClF	0.171	0.359	0.351	0.375	0.279	0.150
BrF	0.261	0.393	0.471	0.461	0.333	0.192
BrCl	0.031	0.008	0.149	0.115	0.084	0.045
NaF	0.838	0.703	0.646	0.988	0.888	0.874
NaCl	0.631	0.667	0.710	0.958	0.844	0.791
NaBr	0.648	0.590	0.579	0.946	0.819	0.762
LiF	0.702	0.661	0.611	0.975	0.871	0.835
LiCl	0.463	0.499	0.495	0.939	0.796	0.734
LiBr	0.484	0.441	0.353	0.924	0.767	0.725
BF	0.150	0.233	0.153	0.623	0.276	0.086
CO	0.105	0.307	0.180	0.594	0.135	0.000
SO	0.256	0.456	0.447	0.640	0.245	0.243
CS	0.049	0.023	0.019	0.069	0.031	0.243
NO	0.025	0.195	0.200	0.500	1.194	0.012

Tab. 3b. Comparison of gas-phase absolute values of partial atomic charges (atomic units) calculated using selected approaches with B3LYP functional at DFT level.

Molecule	MUL	MUL'	MUL*	*NBO*	*APT*	*MSK*
B3LYP						
HF	0.291	0.465	0.317	0.548	0.417	0.452
HCl	0.136	0.226	0.244	0.253	0.211	0.250
HBr	0.110	0.231	0.177	0.184	0.131	0.184
ClF	0.154	0.282	0.232	0.335	0.279	0.150
BrF	0.234	0.298	0.324	0.405	0.355	0.192
BrCl	0.025	0.006	0.141	0.102	0.103	0.045
NaF	0.757	0.552	0.471	0.978	0.858	0.874
NaCl	0.545	0.564	0.622	0.928	0.811	0.791
NaBr	0.566	0.488	0.469	0.909	0.777	0.762
LiF	0.627	0.526	0.446	0.961	0.860	0.835
LiCl	0.402	0.405	0.399	0.925	0.787	0.734
LiBr	0.426	0.348	0.230	0.907	0.752	0.725
BF	0.232	0.097	0.037	0.550	0.271	0.111
CO	0.010	0.174	0.013	0.478	0.255	0.000
SO	0.222	0.364	0.275	0.601	0.320	0.243
CS	0.241	0.098	0.047	0.172	0.127	0.243
NO	0.091	0.112	0.067	0.500	0.186	0.012

rences from Exp. MUL, MUL', MUL* and NBO approaches provide results with rather high deviations from Exp.

Individual bond lengths are comparable at each level of calculation; however, MSK and APT charges show rather satisfactory results in comparison to the values calculated from experimental data. Each type of atomic partial charge shows only small changes in their values at various computational levels.

Assuming the basis set dependency of Mulliken values, MUL' values should be located between MUL

Tab. 3c. Comparison of absolute values of partial atomic charges (atomic units) calculated using selected approaches with M0-62X functional at DFT level.

Molecule / M0-62X	MUL	MUL'	MUL*	*NBO*	*APT*	*MSK*
HF	0.296	0.482	0.341	0.548	0.417	0.457
HCl	0.145	0.237	0.255	0.253	0.211	0.255
HBr	0.123	0.252	0.216	0.190	0.131	0.196
ClF	0.150	0.304	0.267	0.346	0.279	0.150
BrF	0.214	0.313	0.373	0.415	0.355	0.192
BrCl	0.035	0.029	0.114	0.095	0.103	0.045
NaF	0.780	0.610	0.541	0.978	0.858	0.874
NaCl	0.563	0.586	0.618	0.940	0.811	0.791
NaBr	0.585	0.516	0.521	0.925	0.777	0.762
LiF	0.647	0.577	0.511	0.964	0.860	0.835
LiCl	0.415	0.427	0.406	0.929	0.787	0.734
LiBr	0.441	0.377	0.295	0.912	0.752	0.725
BF	0.237	0.121	0.018	0.565	0.271	0.099
CO	0.018	0.192	0.052	0.487	0.255	0.000
SO	0.222	0.384	0.310	0.614	0.320	0.243
CS	0.196	0.092	0.044	0.172	0.127	0.243
NO	0.097	0.119	0.101	0.500	0.186	0.012

and MUL* values since they were calculated using medium basis set of 6-31G*. This is the case only with HCl, BrF, NaF, NaCl, NaBr, LiF, LiBr, BF, and CS molecules. In case of the BrCl molecule, its MUL' value is significantly lower than MUL and MUL* values, and for the rest of the studied molecules, MUL' values are higher than MUL and MUL* values. These inconsistencies suggest not only the basis set dependency of Mulliken values of atomic partial charges, but also recommend their calculation with several basis sets so the trend of Mulliken values can be estimated, more precisely.

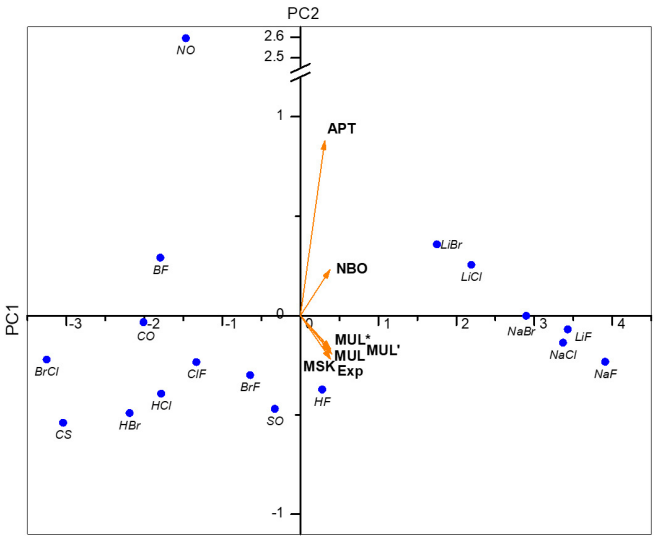
Additionally, most methods of atomic partial charges determination provide nearly identical results at the DFT level of theory regardless of the used functional. This fact implies independency of the studied population analyses from the functional used, or rather small influence of the functional on their results compared to the influence of the level of theory.

PCA was used to determine chemical accuracy of the used methods resulting in biplots (Fig. 1) and covariance matrices (Table 5). The main difference between PCA biplots was observed between MP2 and the remaining methods. Not only the relative ordering (top to bottom) and angles are different, but also the position of NO molecule is significantly different implying errors in the description of this radical molecule at the MP2 level of theory, which corresponds to data shown in Table 3a. Each dataset is divided into two groups: the first group consists of sodium and lithium halides and are located mostly in the first and fourth quadrant of a biplot; the second group consists of the other studied molecules, and it is located in the second and third quadrant with the exception of HF molecule.

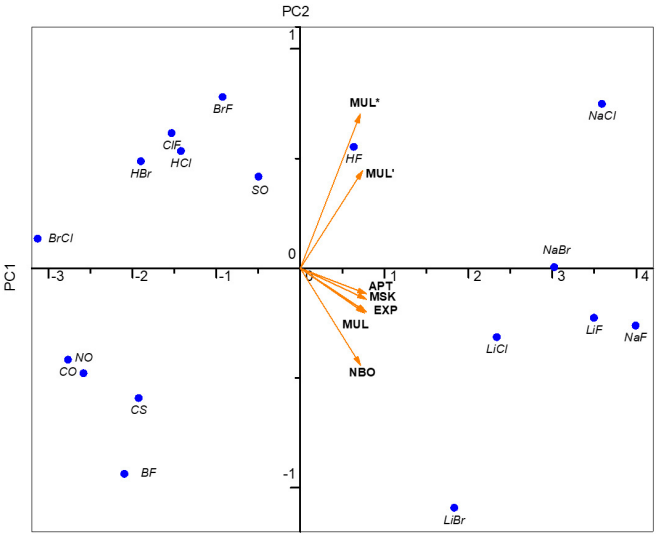
As mentioned before, differences observed between the MP2 and DFT biplots are mainly because of the APT charge value of NO, which is significantly displaced in comparison to other biplots. Such observation suggests that this radical molecule is not described well enough. It is also possible that APT method provides inaccurate values for radical molecules at the MP2 level of theory.

Biplots suggest that the most accurate method of partial atomic charges determination is MUL, followed by MSK. Mulliken method is expected to be rather accurate for MP2 geometries. However, by evaluation of covariance matrices (see Table 4), the most accurate method is MSK followed by APT at the DFT level of theory, probably due to an inconsistency in MUL data.

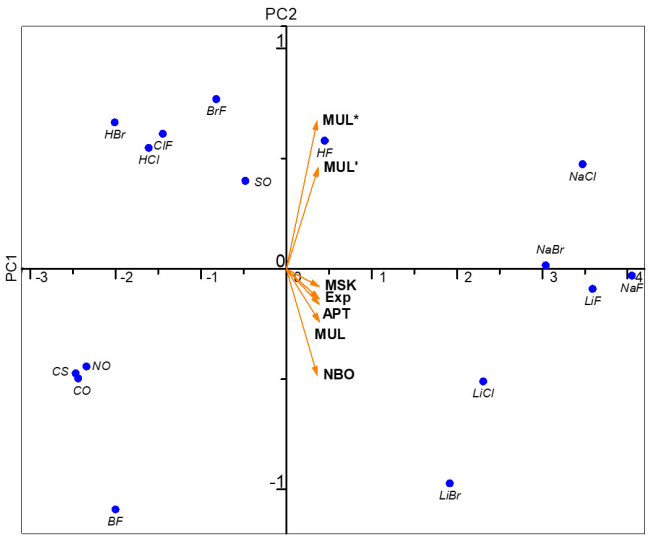
Table 5 shows the list of parameters for linear regression of every studied approach and Exp values. Linear regression was performed to assess the accuracy of individual approaches for theoretical determination of atomic partial charges. Chemical accuracy is determined from values of intercepts, slopes, and correlation coefficients. The closer



a)



b)



c)

Fig. 1. Primary component analysis for MP2 (a), B3LYP (b) and M06-2X (c) methods.

Tab. 4. Covariance table of PCA for individual methods of calculation provided by a computational tool (Jolliffe, 2002).

Group\Group	Exp	MUL	MUL'	MUL*	NBO	APT	MSK
				MP2			
Exp	1.000	0.950	0.818	0.804	0.816	0.659	0.995
MUL	0.950	1.000	0.913	0.900	0.873	0.656	0.946
MUL'	0.818	0.913	1.000	0.942	0.874	0.621	0.829
MUL*	0.804	0.900	0.942	1.000	0.803	0.624	0.811
NBO	0.816	0.873	0.874	0.803	1.000	0.756	0.798
APT	0.659	0.656	0.621	0.624	0.756	1.000	0.636
MSK	0.995	0.946	0.829	0.811	0.798	0.636	1.000
				B3LYP			
Exp	1.000	0.952	0.858	0.829	0.872	0.997	0.995
MUL	0.952	1.000	0.846	0.806	0.857	0.947	0.942
MUL'	0.858	0.846	1.000	0.911	0.809	0.872	0.877
MUL*	0.829	0.806	0.911	1.000	0.703	0.842	0.845
NBO	0.872	0.857	0.809	0.703	1.000	0.861	0.854
APT	0.997	0.947	0.872	0.842	0.861	1.000	1.000
MSK	0.995	0.942	0.877	0.845	0.854	1.000	1.000
				M06-2X			
Exp	1.000	0.956	0.865	0.842	0.868	0.960	0.996
MUL	0.956	1.000	0.869	0.836	0.869	0.934	0.950
MUL'	0.865	0.869	1.000	0.941	0.820	0.880	0.880
MUL*	0.842	0.836	0.941	1.000	0.733	0.846	0.858
NBO	0.868	0.869	0.820	0.733	1.000	0.951	0.853
APT	0.960	0.934	0.880	0.846	0.951	1.000	0.953
MSK	0.996	0.950	0.880	0.858	0.853	0.953	1.000

Tab. 5. List of intercepts (A), slopes (B), and determination coefficients R^2 of linear regression for correlations of individual theoretical absolute values of atomic partial charges and Exp values.

Data	B3LYP			M0-62X			MP2		
	A	B	R^2	A	B	R^2	A	B	R^2
MUL	–0.047	1.407	0.905	–0.043	1.369	0.913	0.001	1.184	0.903
MUL'	–0.118	1.594	0.736	–0.131	1.526	0.748	–0.114	1.264	0.670
MUL*	–0.035	1.537	0.687	–0.057	1.466	0.708	–0.103	1.305	0.646
NBO	–0.155	0.922	0.761	–0.154	0.911	0.753	–0.115	0.824	0.666
APT	–0.102	1.077	0.921	–0.102	1.077	0.921	0.081	0.596	0.434
MSK	–0.019	1.010	0.993	–0.020	1.009	0.992	–0.018	1.005	0.991

values of intercepts to 0 and slopes to 1 are, the more accurate the theoretical approach is, which is further supported by the correlation coefficient value.

Some theoretical approaches show higher accuracy than others, yet the worst correlation coefficient is observed for APT charge calculated at the MP2 level of theory. This is caused by inclusion of APT charge for the NO molecule. If data obtained for the NO molecule are disregarded in this case, the intercept value changes from 0.081 to –0.035 and the slope value changes from 0.596 to 0.962. The observed shift in values towards 0 for intercept and towards 1 for slope results in the determination coefficient R^2 changing to 0.913. Thus, APT values are reasonably accurate at the MP2 level of theory unless radicals are considered. With this observation in mind, all methods provide consistent results in terms of chemical accuracy. Thus, the most accurate results were obtained with the MSK method, followed by APT, MUL, NBO and least accurate results were observed with the MUL* method. Figure 2 shows

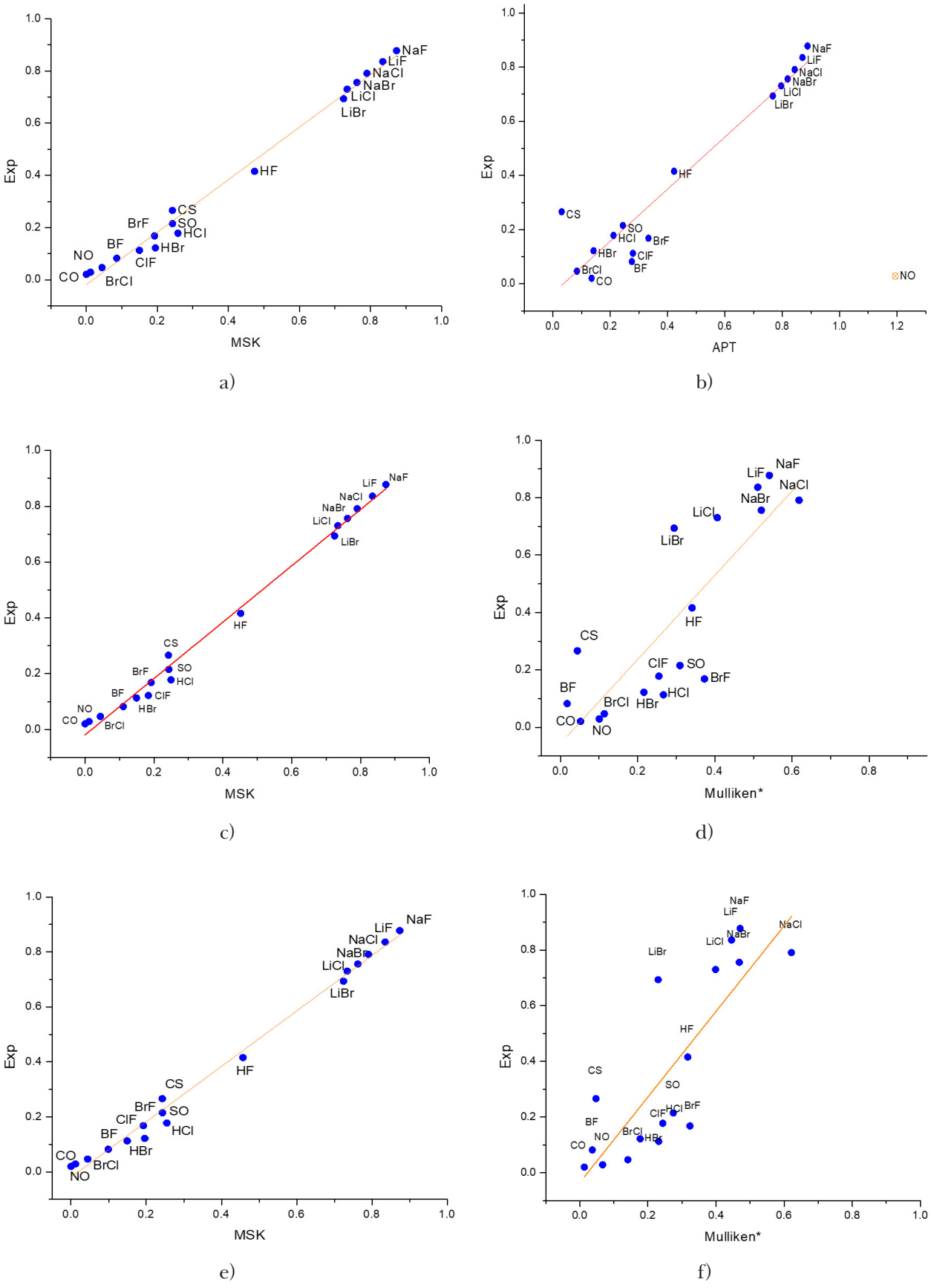


Fig. 2. Selected dependences of theoretical absolute values of partial atomic charges on experimental values. Selection is based on best and worst R^2 factor. Hollow symbols indicate data excluded from linear regression. for: a) MSK (MP2), b) APT(MP2), c) MSK(B3LYP), d) MUL*(B3LYP), e) MSK(M0-62X) and f) MUL*(M06-2X).

linear regression for the most and the least accurate results, whereas linear regression in Fig. 2b has a satisfactory course but it is acquired by exclusion of data for the NO molecule.

Conclusions

In this study, partial atomic charges were evaluated at different levels of theory and were further compared to their values obtained from experimentally acquired parameters.. To find the most accurate method of partial atomic charge determination, the principal component analysis was used. Furthermore, PCA determined the most accurate method tested.

MSK scheme provided results consistently and most accurately at the DFT and MP2 levels of calculation. MUL population analysis performed at the 6-311++G** basis set provided consistent but less accurate results than MSK, while NBO values are consistent and less accurate at the DFT and MP2 computational levels. Last evaluated method, APT, provides rather accurate results at the DFT level, but at the MP2 level of calculation, it showed most significant deviations from the obtained experimental values for radical molecules. The least accurate results were obtained with the Mulliken population analysis computed using the 6-31G* (MUL') and 6-31G basis set (MUL*) at the DFT level. The order of the least accurate methods at the MP2 level changes to MUL*, APT, and MUL', where MUL* is the least accurate method.

A slight dependency on the used functional has been found from further study of results of various methods of atomic partial charges determination. The difference between results is significant in the selection of the level of theory that calculations are to be performed at.

Each of the used methods showed the same trends but the calculated values should be checked, by another computational method, which is also recommended when using the Mulliken population analysis with smaller basis sets. Principal component analysis is rather suitable for the determination of trends and dependencies of chemical and physical parameters.

Correlations of individual theoretical approaches with the experiments using linear regression have confirmed MSK to be the most accurate approach at all levels of theory. It was also determined that the APT approach performed at the MP2 level of theory provides rather accurate results in case radical molecules are not included in the studied set.

Acknowledgment

This work was supported by the Slovak Research and Development Agency (APVV-19-0024) and VEGA 1/0461/21.

References

Becke AD (1993) J. Chem. Phys., 98: 5648—5652.

Besler BH, Merz KM, Kollman PA (1990) Journal of Computational Chemistry, 11(4): 431—439.

Biarge J, Herranz J, Morcillo J (1961) Annals of the Royal Spanish Society of Physics and Chemistry, 57, p 81.

Breneman MC, Wiberg BK (1990) Journal of Computational chemistry 11(3): 361.

Cagardová D, Matúška J, Poliak P, Lukeš V (2019) J. Phys. Chem. C.

Cagardová D, Matúška J, Michalík M, Poliak P, Lukeš V (2020) Synthetic Metals, 260: 116263.

Cagardová DM, Matúška J, Michalík M, Lukeš V (2022) Materials Today Communications, 33: 104688.

Frisch MJ, Head-Gordon M, Pople JA (1990) Chem. Phys. Lett., 166: 275—280.

Gaussian 16, Revision B.01, Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich AV, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark MJ, Heyd JJ, Brothers EN, Kudin KN, Staroverov VN, Keith TA, Kobayashi R, Normand J, Raghavachari K, Rendell AP, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski JW, Martin RL, Morokuma K, Farkas O, Foresman JB, Fox DJ. Gaussian, Inc., Wallingford CT, 2016.

Gross KC, Seybold PG, Hadad CM (2009) Int. J. Quantum Chem., 109: 3679—3684.

Hariharan PC, Pople JA (1973) Theor. Chim. Acta., 28: 213—222.

Huber KP, Herzberg G (2013) Molecular Spectra and Molecular Structure. Springer New York, NY.

Johnson III RD (2022, May 22), NIST Computational Chemistry Comparison and Benchmark Database, http://cccbdb.nist.gov/.

Jolliffe IT (2002) Principal Component Analysis in: Springer Series in Statistics, Springer New York, NY.

Irikura KK (2007) Journal of Physical and Chemical Reference Data, 36: 389—397.

Klein E, Lukeš V (2006) Chem. Phys., 330(3): 515—525.

Klein E, Lukeš V (2007) Journal of Molecular Structure: THEOCHEM, 805(1-3): 153—160.

Lee C, Yang W, Parr RG (1988) Phys. Rev. B, 37: 785—789.

Mehmood A, Janesko BG (2017) Angew. Chem. Int., 56: 6878—6881.

Michalik M, Poliak P, Klein E, Lukeš V (2018a) Chem. Phys. Let., 709: 71—76.

Michalik M, Poliak P, Lukeš V (2018b) Acta Chimica Slovenica, 65(1): 23—33.

Mulliken RS (1955) The Journal of Chemical Physics, 23(10): 1833—1840.

Nelson Jr RD, Lide Jr DR, Maryott AA (1967) National Standard Reference Data Series — National Bureau of Standards 10, U.S. Department of Commerce.

Person W, Newton J (1974) J. Chem. Phys. 61, 1040—1049.
Raczyńska ED, Hallman M, Kolczyńska K, Stepniewski (2010) Symmetry, 2: 1485—1509.
Reed AE, Weinstock RB, Weinhold F (1985), J. Chem. Phys., 83(2): 735—746.
Richter WE, Duarte LJ, Bruns RE (2021) J. Chem. Inf. Model., 61(8): 3881—3890.
Rostkowski M, Paneth P (2007) Pol. J. Chem., 81: 711—720.
Singh UC, Kollman PA (1984) J. Comp. Chem., 5: 129—145.
Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) J. Phys. Chem, 98: 11623—11627.
Štellerová D, Michalík M, Lukeš V (2022) Phytochemistry, 203: 113387.
Svobodová Vařeková R, Geidl S, Ionescu CM, Skřehota O, Kudera M, Sehnal D, Bouchal T, Abagyan R, Huber HJ, Koča J (2011) J. Chem, Inf. Model, 51(8): 1795—1806.
Thakkar AJ (2021) Quantum Chemistry (Third Edition): A concise introduction for students of physics, chemistry and materials science, Institute of Physics Publishing, Bristol, United Kingdom.
Ugur I, Marion A, Parant SP, Jensen JH, Monard G (2014) J. Chem. Inf. Model., 54: 2200—2213.
Vosko SH, Wilk L, Nusair M (1980) Can. J. Phys., 58: 1200—1211.
Williams DE, Yan JM (1988) Adv. Atomic Mol. Phys, 23: 87.

sciendo

Substitution effect of phenol derivatives on electrochemical oxidation potentials: Correlation of theoretical reaction Gibbs free energies

Andrea Kováčová

Institute of Physical Chemistry and Chemical Physics, Slovak University of Technology in Bratislava, Radlinského 9, SK-812 37 Bratislava, Slovakia
andrea.kovacova@stuba.sk

Abstract: Reaction Gibbs free energies for electron abstraction from phenol and its 74 derivatives were calculated using the composite *ab-initio* approach (G4) in combination with the implicit solvation model. Resulting values were correlated with 38 oxidation potentials obtained from cyclic voltammetry measurements. Substitution effect and the role of substituents in *ortho*-, *meta*-, and *para*-position were also quantified by Hammett constants. The evaluated linear dependences can be used to reliably estimate electrochemical potentials of substituted phenols solvated in water.

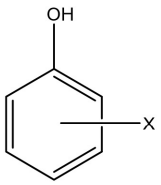
Keywords: Reaction Gibbs energies; Gn approach; Thermodynamics; Implicit solvent model

Introduction

Phenol and its derivatives are key moieties in the production of polycarbonates, detergents, herbicides, synthetic drugs, dyes, and antioxidants. Phenolic structural units are also widely distributed in plants and they play important roles in living systems, such as defence mechanisms against viruses, bacteria (Albuquerque et al., 2021) or reactive radicals (Michalík et al., 2019a). Phenolic compounds can also represent hazardous industrial pollutants (Guerra, 2001). On the other hand, chelating ability of polyphenols (Michalík et al., 2021; Šimunková et al., 2022) can potentially be used to eliminate solvated metal ions in wastewater. The main chemical transformation pathway of these molecules is often associated with their oxidation where electron loss can occur electrochemically (Pavitt et al., 2017) or in the presence of reducing metal cations (Ignaczak et al., 1987).

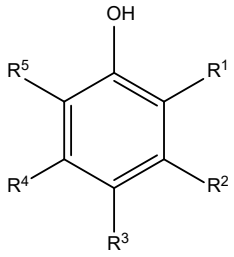
Theoretical calculations using quantum chemical methods play a very important role in this context; although they do not always provide accurate absolute values of quantities, they correctly reproduce experimental trends. Structural or energetic parameters of phenol-based molecules can also be correlated with biological activity (Štellerová et al., 2022). Currently, Density Functional Theory (DFT) calculations are most commonly used due to their lower computational cost. However, chemical accuracy of the results depends on the choice of DFT functional and the basis set. In case of thermochemical data calculations, the Gaussian-n (Gn) method is generally more accurate as it is based on a series of calculations using different methods and basis sets to obtain accurate energetics (Curtiss et al., 2007). In addition, various molecule-specific empirically independent parameters and higher-level correction terms are included to estimate the remaining deficiencies, assuming they are systematic (Curtiss et al., 2007). The use of geometries and zero-point energies from the DFT, application of CCSD(T) energies, inclusion of estimated *ab initio* Hartree-Fock limit energy, and addition of higher level correction parameters in the G4 method contribute to the most significant improvement of the original Gn theory (Curtiss et al., 2007). The combination of the implicit solvent model with G4 or DFT has been successfully used to predict the reaction enthalpies and Gibbs energies of phenol and its analogues (Cagardová et al., 2019; Michalík et al., 2019a; Biela et al., 2023). Interestingly, the use of pure *ab initio* MPn approach can fail in the theoretical prediction of bond dissociation enthalpies (Klein and Lukeš, 2006).

Description of solvent-solute effects can significantly affect the quality of the results obtained. Inclusion of explicit solvent molecules surrounding the molecule under investigation generates numerous geometry structures and consequently leads to rapid increase of computational time. The use of implicit solvent approximation assumes that solvent molecules can be replaced by a homogeneously polarisable and thermally averaged isotropic medium (Mennucci et al., 2002). Many theoretical physicochemical properties calculated for flavonols (Michalík et al., 2019b), phenols (Michalík et al., 2017), alcohols (Michalík and Lukeš, 2016), model amines (Rimarčík et al., 2010), and toluidine isomers (Lukeš and Hartmann, 2021) solvated in different polar solvents are reasonably well estimated using implicit solvent models combined with DFT.



X: H, NMe_2, F, Br, CF_3, NO, COH, Me, Et, OMe, NO_2, Ph, Cl, OH, CN, COMe, COOH, NH_2

a)



Name	Notation	Name	Notation
2,3-dimethylphenol	23Me	3,4-dimethoxyphenol	34OMe
2,4-dimethylphenol	24Me	3,5-dimethoxyphenol	35OMe
2,5-dimethylphenol	25Me	2,4-dinitrophenol	$24NO_2$
2,6-dimethylphenol	26Me	dopamine	dopamine
3,4-dimethylphenol	34Me	2,4,6-trimethylphenol	246Me
3,5-dimethylphenol	35Me	2,4,6-trichlorophenol	246Cl
2,3-dimethoxyphenol	23OMe	2-methoxy-4-ethylphenol	2OMe-4Et
2,4-dimethoxyphenol	24OMe	2-Hydroxy-4-methoxybenzaldehyde	2OMe-4COH
2,5-dimethoxyphenol	25OMe	2,4-dinitro-2-methylphenol	$2Me\text{-}24NO_2$
2,6-dimethoxyphenol	26OMe	2,6-dinitro-4-methylphenol	$4Me\text{-}26NO_2$

(b)

Fig. 1. Schematic structures and notations of studied molecules (a) mono- (b) polysubstitued.

Satisfactory predictions were also obtained for hydroxyl groups attached to non-aromatic rings (Žemlička et al., 2014) and alkyl chains (Štellerová and Lukeš, 2021).

One-electron oxidation reactions of phenols can affect various types of chemical processes. Therefore, knowledge of electrochemical properties of these compounds is crucial for the identification of these reactions. In this context, experimental study by Pavitt et al. (2017) is interesting, where half-wave oxidation potentials ($E_{1/2}$) were measured and determined electrochemically for a series of 38 phenol derivatives in water and isopropyl alcohol. These electrochemical potentials were correlated with calculated DFT potentials; solvent effects were included using the COSMO and SMD models. Although a relatively large number of substances have been studied experimentally, the role of substitution has not been sufficiently discussed. In particular, the electron-donating or electron-withdrawing effect of the substituent as well as the role of the substituent position have not been sufficiently addressed. Therefore, partial aims of this study include: (1) calculation of G4 reaction Gibbs free energies of the reaction between radical cations and neutral molecules; (2) analysis of the substituent effect using Hammett constants; (3) correlation of obtained data with available experimental $E_{1/2}$ values. In this study, 74 derivatives have been investigated (Fig. 1).

Computational Details

Quantum chemical calculations in water were performed at Gaussian-4 level (G4) of theory (Curtiss et al., 2007), which selects a suitable basis set, e.g. in this work, the GTBas3 basis set, i.e. 6-31G(2df,p) for molecules containing H, C, N, F, Cl, and 6-31G(2fg) for molecules possessing Br atom. Contributions of the solvent effect were approximated by the SMD model (Solvation Model based on the quantum mechanical charge Density of a solute molecule interacting with a continuum) (Marenich et al., 2009), which is based on generalised Born equation represented by an approximation of the Poisson equation suitable for arbitrary cavity shapes. Gibbs free energies were evaluated for the temperature of 298.15 K and the pressure of 101.325 kPa. Calculations were performed using the Gaussian 09 software (Frisch et al., 2009). Vibrational analysis included in the G4 calculation process showed no imaginary frequencies, confirming the geometry corresponding to the lowest energy minimum. From the optimal geometries obtained for all studied molecules, optimal geometries of the *ortho*- and polysubstituted phenol derivatives were visualised using the Jmol viewer (Fig. 2) (Jmol development team, 2016). In this work, the type of reaction studied is presented in Eq. (1)

$$X—Ar—OH \rightarrow (X—Ar—OH)^{\bullet+} + e^- \qquad (1)$$

where X represents the substituents used. Spin state of neutral compounds is a singlet and corresponding cation radicals consist of one unpaired electron, i.e., their spin state is doublet.

Reaction Gibbs free energies expressing electron abstraction ($\Delta_r G^e$) are

$$\Delta_r G^e = G((X—Ar—OH)^{\bullet+}) - G(X—Ar—OH) + G(e^-) = \\ = \Delta G + G(e^-) \qquad (2)$$

where $G((X—Ar—OH)^{\bullet+})$ and $G(X—Ar—OH)$ denote the Gibbs free energies of the cation radical and neutral compound, respectively, and $G(e^-)$ is the Gibbs free energy of the solvated electron. The value of $G(e^-)$ is constant for all reactions and it is deduced from linear dependences (see the intercept of Eq. 6). For this reason, all energy values in this manuscript are given without the value of $G(e^-)$. According to the thermodynamic theory (Atkins and Paula, 2006), the absolute electrochemical potential can be calculated from the reaction Gibbs energies using the equation

$$E = -\Delta_r G^e / nF \quad (3)$$

where E denotes the electrochemical potential, $\Delta_r G^e$ is the reaction Gibbs free energy, n represents the number of exchange electrons in the reaction, and F is the Faraday constant (96 485 $C \cdot mol^{-1}$). Assuming that the absolute electrochemical potential of the Standard Hydrogen Electrode (SHE) is 4.28 V (E^{SHE}), according to (Ho et al., 2016), the value of the electrochemical potential is calculated as follows

$$E^{cal} = E - E^{SHE} \quad (4)$$

In our study, the dependence between the available Hammett constants (Tab. 1) and the experimental values of the electrochemical potentials for *ortho*, *meta* and *para* derivatives has been examined (Fig. 3).

Tab. 1. Takahata-Chong Hammett type constants for *ortho* isomers (σ_o) (Takahata and Chong, 2005) and Hammett constants for *meta* (σ_m) and *para* (σ_{p^+}) isomers (Hansch et al., 1991).

Substituent	σ_o	σ_m	σ_{p^+}
H	0.00	0.00	0.00
Me	–0.17	–0.07	–0.31
Et	–0.31	–0.07	–0.30
OMe	–0.39	0.12	–0.78
NO_2	0.80	0.71	0.79
Ph		0.06	–0.18
Cl	0.20	0.37	0.11
OH	–0.05	0.12	–0.92
CN	0.71	0.56	0.66
COMe		0.38	
COOH		0.37	0.42
NH_2	–0.41	–0.16	–1.30
COH		0.35	0.73
F	0.24	0.34	–0.07
Br		0.39	0.15
CF_3		0.43	0.61
NO		0.62	
NMe_2	–0.73	–0.16	–1.70

Results and discussion

Optimised geometries of selected species from the studied molecules are shown in Fig. 2. The phenolic group and the substituents are in one plane with the benzene ring. Exceptions are sterically bulky substituents such as **Ph, bisphenol A** and **NMe_2**. In case of *ortho*-substituents containing oxygen or halogen atoms, an intramolecular hydrogen bond is formed with the **OH** group.

Calculated ΔG values for aqueous environment are listed in Tab. 3 and Tab. 4. The reaction Gibbs free energies range from 473 $kJ \cdot mol^{-1}$ for **$4NMe_2$** to 672 $kJ \cdot mol^{-1}$ for **$24NO_2$** and their values indicate that the reactions are exothermic in the reduction direction.

Results obtained from the theoretical G4 reaction Gibbs energies are summarised in Tab. 3 and Tab. 4; a comparison of the theoretical values with the available experimental values indicates large deviations. The values range from 0.33 for **3Cl** to 1.19 for **$4NH_2$**. For example, for **$4NH_2$**, the experimental value is 0.425 V vs. SHE and the theoretical value is –0.77 V. A possible explanation is related to the limitations of both the implicit model and the DFT approach.

In physical organic chemistry, Hammett constants are used to estimate the substitution effect of aromatic molecules. This approach provides a simple way of determining the electron-donating (negative values) and electron-withdrawing (positive values) groups, respectively. They are different for *meta* (σ_m) and *para* (σ_p) substituents, i.e., a substituent can be electron-withdrawing in the *meta* position and electron-donating in the *para* position, e.g., **Ph**, **F**, **OCH_3**. For *ortho* substituents, there are several *Hammett-type substituent constants* (σ_o), e.g., the **Takahata constants** (Takahata & Chong, 2005). For *para* substituents, in addition to σ_p, there are also the Brown-Okamoto constants (σ_{p^+} and σ_{p^-}) (Brown and Okamoto, 1957), the use of which depends on the reaction centre of the molecule. Therefore, Hammett constants have long been used for different types of dependences where it is necessary to find a relationship between different types of substituents and a given quantity (e.g., BDE, ionisation potentials, pK_a) (Bordwell et al., 1994; Klein et al., 2009; Lukeš et al., 2017). The *ortho* Hammett constants proposed by Takahata are not available for all substituents (Takahata and Chong, 2005).

The dependence of Hammett constants (σ) on the experimental electrochemical potentials (E) is linear

$$E/V = 0.377 \times \sigma_{o,m,p^+} - 1.016 \quad (5)$$

with the coefficient of determination (R^2) of 0.918 (Fig. 3). Errors of the regression parameters

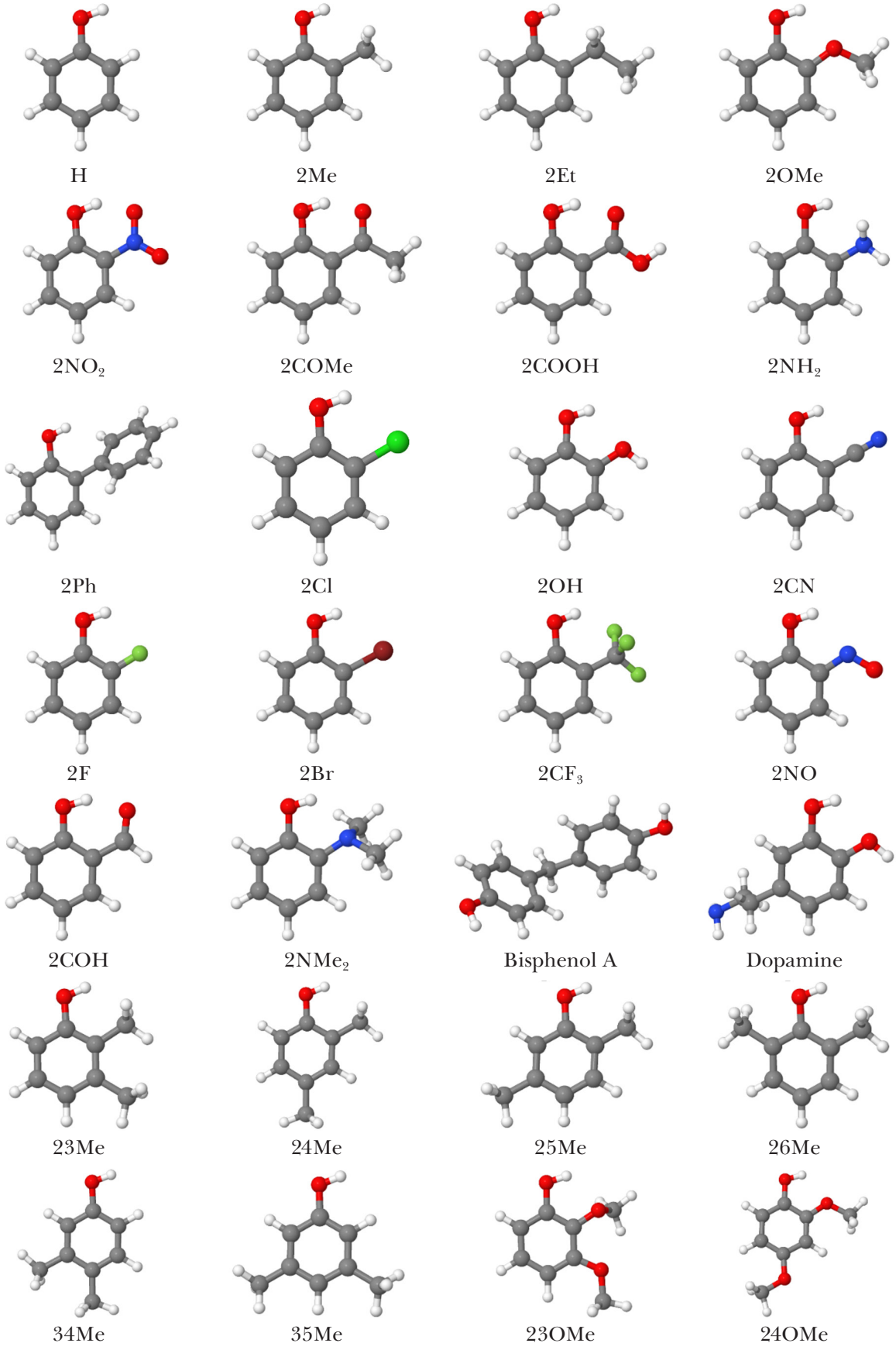


Fig. 2a. Optimal geometries of neutral forms of studied *ortho*- and poly-substituted species in aqueous solution.

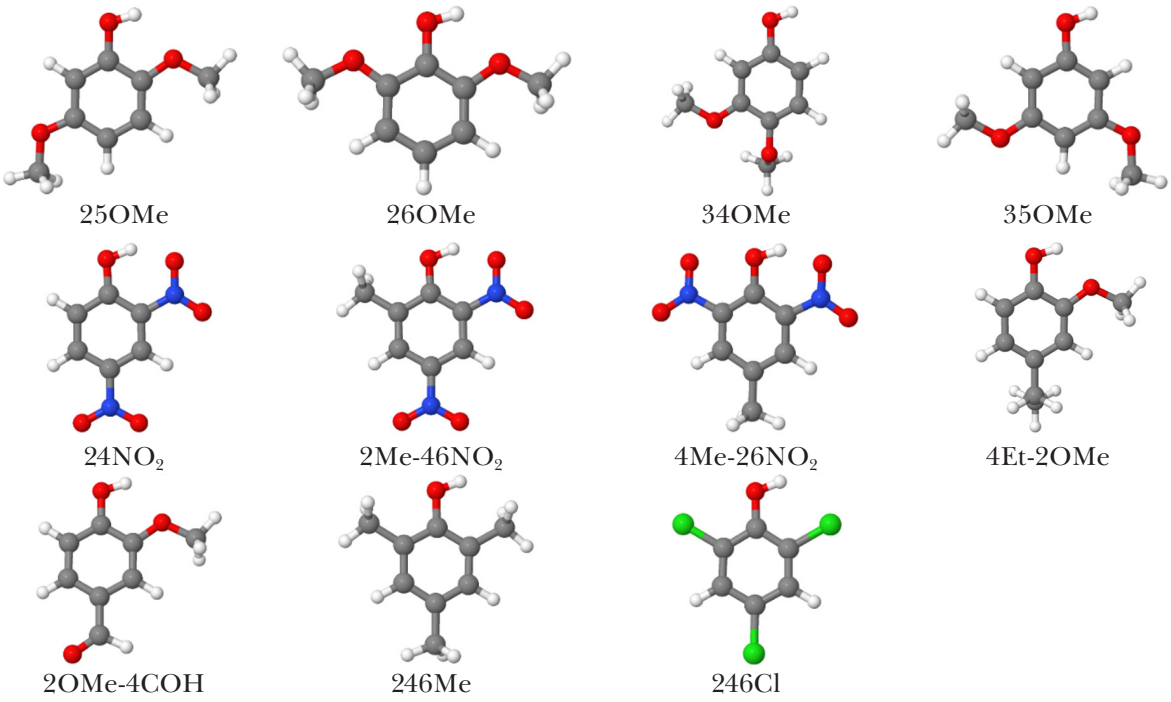


Fig. 2b. Optimal geometries of neutral forms of studied *ortho*- and poly-substituted species in aqueous solution.

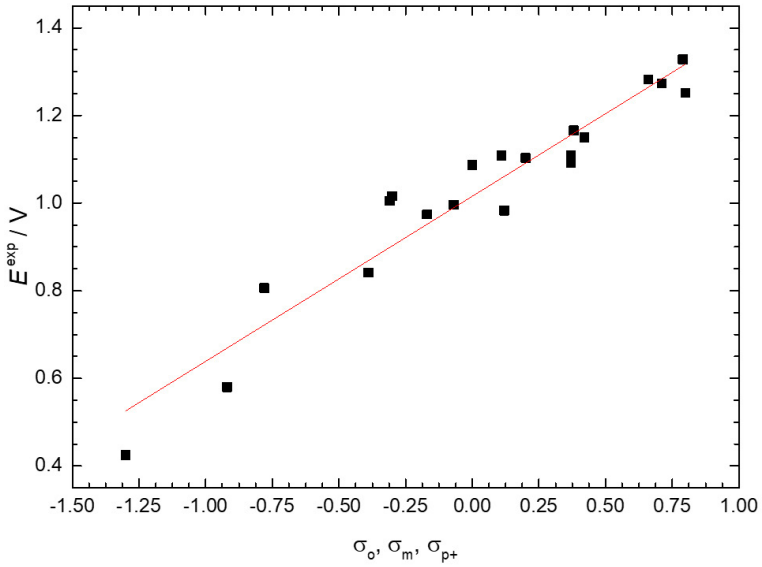


Fig. 3. Correlation between Hammett (Hansch et al., 1991; Takahata and Chong, 2005) and available experimental electrochemical potential values (Pavitt et al., 2017) for studied species in water.

are 0.015 V for intercept and 0.026 V for slope, respectively. In case of *para* substituents, σ_{p+} constants were used in the linear dependence since the dependence employing σ_p and σ_{p-} constants, respectively, showed worse correlation. For compounds with more substituents or where Hammett constants are not available, the correlation between Hammett constants and experimental electrochemical potential cannot be reliably evaluated.

Shortcomings of the theoretical model used can be overcome by direct correlation of the available electrochemical oxidation potentials $E_{1/2}$ with the calculated ΔG. As it is shown in Fig. 4, this correlation is linear and can be described by the following equation

$$E/\mathrm{V} = 0.005 \times \Delta G/\mathrm{kJ \cdot mol^{-1}} - 2.174 \qquad (6)$$

with a satisfactory value of coefficient of determi-

Tab. 2. Parameters a, b with errors and coefficient of determination (R^2) for linear correlations described by the equation $E/V = a \times \Delta G/kJ \cdot mol^{-1} - b$ for *ortho-* with polysubstitueted phenols (E_o), *meta-* (E_m) and *para-* (E_p) substituted phenols.

	$a\ 10^{-3}/mol \cdot V \cdot kJ^{-1}$	b/V	$\sigma_a\ 10^{-4}/mol \cdot V \cdot kJ^{-1}$	σ_b/V	R^2
E_o	5.5	–2.18	2.4	0.15	0.977
E_m	4.2	–1.41	9.1	0.54	0.773
E_p	5.8	–2.36	3.9	0.23	0.958

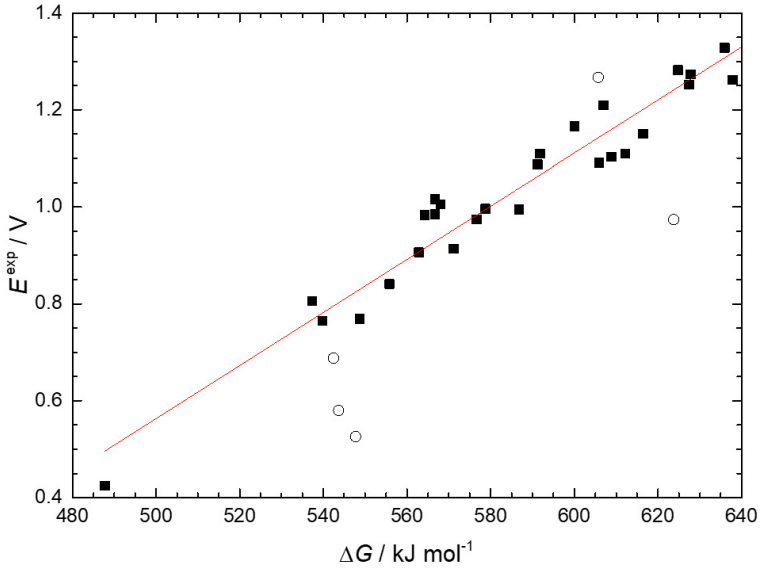


Fig. 4. Correlation between the difference in Gibbs energies of the cation radical forms and neutral molecules and the available experimental electrochemical potential values (Pavitt et al., 2017) for studied species in water. Open circles represent groups excluded from the linear regression.

nation, $R^2 = 0.953$, for studied derivatives. Errors of the regression parameters are 0.14 V for intercept and 2.3 10^{-4} $mol \cdot V \cdot kJ^{-1}$ for slope, respectively. In addition to this correlation, correlations were also performed separately for *ortho-* with polysubstituted phenols (E_o), *meta-* (E_m) and *para-* (E_p) substituted phenols separately. Regression parameters *a*, *b* with the regression errors and the coefficient of determination (R^2) obtained from these linear correlations are summarised in Tab. 2. In these regressions, the same points were excluded from the joint correlation. As the regressions contain different numbers of points, our calculations are based on Eq. 6.

The intercept includes $G(e^-)$, which has a value of $-149\ kJ \cdot mol^{-1}$ (Zhan and Dixon, 2003; Fifen, 2013), while the slope covers the imperfections of the implicit model for the specified type of substituted molecules. Based on the correlation coefficient, experimental values of electrochemical potentials are in good agreement with our calculated Gibbs energies. All 36 experimental values of the electrochemical potential were available for the studied compounds (6 for *ortho-* and *meta-*substituted, 11 for *para-*substituted, and 9 for polysubstituted derivatives). Some substituents had to be excluded from the correlation (Fig. 4) because they did not follow the general trend and were significantly scattered from the regression line. **2COOH** shows a slight deviation from the linear dependence due to steric effects occurring in phenolic derivatives with substituents in the *ortho* position between the substituent and the OH group, but also due to expected dissociation of the **COOH** group. Thus, it is possible that the available experimental value of the electrochemical potential for compounds containing the **COOH** group belongs to its carboxylate anion. Deviations from the linear dependence were also observed for **4OH**, **26OMe**, **246Cl,** and **dopamine** (Fig. 4).

Estimated electrochemical potentials (E) corresponding to the oxidation of all phenolic derivatives studied based on these equations are summarised in Tab. 3 and Tab. 4. The Gibbs energies and associated electrochemical potentials range from 509 $kJ \cdot mol^{-1}$ (0.61 V) for **2NMe$_2$** to 627 $kJ \cdot mol^{-1}$ (1.26 V) for **2NO$_2$**, from 496 $kJ \cdot mol^{-1}$ (0.54 V) for **3NMe$_2$** to 628 $kJ \cdot mol^{-1}$ (1.26 V) for **3NO$_2$**, from 473 $kJ \cdot mol^{-1}$ (0.41 V) for **4NMe$_2$** to 636 $kJ \cdot mol^{-1}$

Tab. 3. Calculated G4(SMD=water) Gibbs free energies (ΔG) in kJ · mol^{-1} for electron abstraction, electrochemical potential values calculated from Eq. 4 vs. SHE, scaled electrochemical potential values obtained from Eqs. 5—6 in V, and experimental electrochemical potential E vs. SHE (Pavitt et al., 2017). Values in parentheses stand for the Gibbs reaction energies for electron abstraction evaluated from electrochemical potentials (Eq. 3).

Substituent	*ortho*					*meta*					*para*				
	ΔG	E^{cal}	$E^{\Delta G}$	E^{σ}	E^{exp}	ΔG	E^{cal}	$E^{\Delta G}$	E^{σ}	E^{exp}	ΔG	E^{cal}	$E^{\Delta G}$	E^{σ}	E^{exp}
H	591	0.30	1.06 (102)	1.02	1.09	591	0.30	1.06 (102)	1.02	1.09	591	0.30	1.06 (102)	1.02	1.09
Me	577	0.15	0.98 (95)	0.95	0.97	579	0.17	0.99 (96)	0.99	1.00	568	0.06	0.93 (90)	0.90	1.01
Et	580	0.19	1.00 (97)	0.90		576	0.15	0.98 (94)	0.99		567	0.05	0.93 (89)	0.90	1.02
OMe	556	–0.06	0.87 (84)	0.87	0.84	564	0.02	0.91 (88)	1.06	0.98	537	–0.25	0.77 (74)	0.72	0.81
NO_2	627	0.68	1.26 (121)	1.32	1.25	628	0.68	1.26 (122)	1.28	1.27	636	0.77	1.30 (126)	1.31	1.33
Ph	587	0.26	1.04 (100)	1.02	1.00	583	0.22	1.02 (98)	1.04		562	0.00	0.90 (87)	0.95	
Cl	609	0.49	1.16 (112)	1.09	1.10	606	0.46	1.14 (110)	1.16	1.09	592	0.31	1.06 (103)	1.06	1.11
OH	563	0.01	0.91 (87)	1.00		574	0.12	0.96 (93)	1.06		544	–0.19	0.80 (77)	0.67	0.58
CN	626	0.67	1.25 (121)	1.28		622	0.62	1.23 (119)	1.23		625	0.65	1.24 (120)	1.26	1.28
COMe	598	0.38	1.10 (106)	1.02		600	0.39	1.11 (107)	1.16	1.17	607	0.47	1.15 (111)	1.02	1.21
COOH	606	0.45	1.14 (110)	1.02	1.27	612	0.52	1.17 (113)	1.16	1.11	616	0.56	1.20 (116)	1.17	1.15
Bisphenol A											571	0.10	0.95 (92)	1.02	0.91
NH_2	515	–0.49	0.64 (62)	0.86		517	–0.46	0.65 (63)	0.96		488	–0.77	0.49 (48)	0.53	0.43
Br	616	0.56	1.20 (115)	1.02		610	0.50	1.17 (112)	1.16		597	0.37	1.09 (106)	1.07	
CF_3	620	0.60	1.22 (118)	1.02		610	0.49	1.16 (112)	1.18		626	0.66	1.25 (121)	1.25	
F	602	0.42	1.12 (108)	1.11		604	0.44	1.13 (109)	1.14		587	0.26	1.04 (100)	0.99	
NMe_2	509	–0.55	0.61 (59)	1.02		496	–0.68	0.54 (52)	0.96		473	–0.92	0.41 (40)	0.37	
NO	592	0.31	1.07 (103)	1.02		590	0.29	1.05 (102)	1.25		576	0.14	0.98 (94)	1.02	
CHO	607	0.46	1.14 (110)	1.02		610	0.50	1.16 (112)	1.15		614	0.54	1.18 (114)	1.29	

(1.30 V) for **4NO$_2$**, and from 521 kJ · mol^{-1} (0.68 V) for **24OMe** to 672 kJ · mol^{-1} (1.14 V) for **24NO$_2$**. The above-mentioned values show that the **NMe$_2$** group has the lowest and **NO$_2$** the highest electrochemical potential for each position in agreement with the corresponding Hammett constants.

As it can be seen from the comparison of the reaction Gibbs energies for cation formation in aqueous

Tab. 4. Calculated G4(SMD=water) Gibbs free energies (ΔG) in kJ · mol^{-1} for electron abstraction, electrochemical potential values calculated from Eq. 4 vs. SHE, scaled electrochemical potential values obtained from Eqs. 5—6 in V, and experimental electrochemical potential E vs. SHE (Pavitt et al., 2017). Values in parentheses stand for the Gibbs reaction energies for electron abstraction evaluated from electrochemical potentials (Eq. 3).

Substituent	ΔG	E^{cal}	$E^{\Delta G}$	E^{exp}	Substituent	ΔG	E^{cal}	$E^{\Delta G}$	E^{exp}
24NO$_2$	672	1.14	1.50 (145)	1.50	**23OMe**	557	–0.05	0.87 (84)	
2Me-46NO$_2$	652	0.93	1.39 (134)	1.40	**24OMe**	521	–0.43	0.68 (65)	
4Me-26NO$_2$	638	0.79	1.32 (127)	1.26	**25OMe**	523	–0.40	0.69 (66)	
Dopamine	548	–0.15	0.82 (79)	0.53	**26OMe**	542	–0.20	0.79 (77)	0.69
23Me	573	0.11	0.96 (93)		**34OMe**	528	–0.35	0.71 (69)	
24Me	556	–0.07	0.86 (83)		**35OMe**	568	0.06	0.93 (90)	
25Me	563	–0.01	0.90 (87)	0.91	**2OMe-4Et**	540	–0.23	0.78 (75)	0.77
26Me	567	0.06	0.93 (90)		**2OMe-4CHO**	567	0.05	0.93 (89)	0.99
34Me	551	–0.12	0.84 (81)		**246Me**	549	–0.14	0.83 (80)	0.77
35Me	568	0.07	0.93 (90)		**246Cl**	624	0.64	1.24 (119)	0.97

Tab. 5. Calculated G4(SMD=water) deprotonation Gibbs free energies ($\Delta_r G$) in kJ · mol^{-1} of proton abstraction for selected molecules calculated according to $\Delta_r G = G$ (deprotonated) — G (protonated) + + G (H^+). Asterisk symbol denotes the molecule with intramolecular hydrogen bond.

Substituent	*ortho*		*meta*		*para*		*poly*	
	$\Delta_r G_1$	$\Delta_r G_2$	$\Delta_r G_1$	$\Delta_r G_1$	$\Delta_r G_1$	$\Delta_r G_2$	$\Delta_r G_1$	$\Delta_r G_2$
OH	87*	102	98	98	106	106		
COOH	34*	99	47	94	51	82		
Bisphenol A					99	99		
Dopamine							90*	102

Tab. 6. Calculated G4(SMD=water) Gibbs free energies (ΔG) in kJ · mol^{-1} for electron abstraction from deprotonated species and scaled electrochemical potential ($E^{\Delta G}$) values obtained from Eq. 6 in V.

Substituent	*ortho*				*meta*				*para*				*poly*			
	ΔG_1	$E^{\Delta G}$	ΔG_2	$E^{\Delta G}$	ΔG_1	$E^{\Delta G}$	ΔG_2	$E^{\Delta G}$	ΔG_1	$E^{\Delta G}$	ΔG_2	$E^{\Delta G}$	ΔG_1	$E^{\Delta G}$	ΔG_2	$E^{\Delta G}$
OH	444	0.25	438	0.22	454	0.31	453	0.30	424	0.14	424	0.14				
COOH	632	1.28			584	1.02			579	0.99						
Bisphenol A									434	0.20	434	0.20				
Dopamine													426	0.16	429	0.17

media, the values range from 40 kJ · mol^{-1} (**4NMe$_2$**) to 145 kJ · mol^{-1} (**24NO$_2$**). In case of molecules with a carboxyl group or an additional hydroxyl group, the hypothetical deprotonated form may also undergo ionisation in aqueous media. This may affect the measurement of the electrochemical potential under investigation. Considering the experimental Gibbs free energy for the proton (–1121 kJ · mol^{-1}) (Tawa et al. 1998; Ishikawa and Nakai 2016), the corresponding reaction Gibbs energies of deprotonation of carboxyl groups are 34 kJ · mol^{-1} (**2COOH**), 47 kJ · mol^{-1} (**3COOH**), and 51 kJ · mol^{-1} (**4COOH**). These values are lower than those of the OH group deprotonation, i.e., 99 kJ · mol^{-1} (**2COOH**), 94 kJ · mol^{-1} (**3COOH**), and 82 kJ · mol^{-1} (**4COOH**). Energetics are influenced by intramolecular hydrogen bonding for the *ortho* position.

In case of molecules containing more than one OH group, the intramolecular hydrogen bonding also affects the reaction energies. Calculated Gibbs deprotonation energies (see Tab. 5) indicate that the energy interval overlaps with the energy interval for the heterolytic cleavage of the O—H bond in hydroxybenzoic acid.

Deprotonation from the hydroxyl group of the **2OH** molecule is energetically more favourable due to stabilisation by an intramolecular hydrogen bond. By abstracting an electron from the anion of the corresponding hydroxybenzoic acid, the ΔG_1 values

change to 632 kJ·mol^{-1} (**2COOH**), 584 kJ·mol^{-1} (**3COOH**), and 579 kJ·mol^{-1} (**4COOH**). Compared to the formation of a cation radical from a neutral molecule, these values are higher for **2COOH** and lower for **3COOH** and **4COOH**. Due to the higher deprotonation energy of the OH group of hydroxybenzoic acid, in comparison with the COOH group, the corresponding electrochemical potentials of the anionic form were not calculated (see Tab. 6).

The corresponding scaled electrochemical potentials for electron abstraction from selected deprotonated species obtained from Eq. 6 are 1.28 V (**2COOH**), 1.02 V (**3COOH**), 0.99 V (**4COOH**) (see Tab. 6). This supports the hypothesis that the experimental potential for **2COOH** can be combined from the values of the two deprotonation processes from the COOH and the OH group. **4OH** and **dopamine** show a lower value of the scaled electrochemical potential when the presence of an anion is assumed. The scaled electrochemical potentials obtained from the deprotonated form in case of **2COOH**, **4OH**, and **dopamine** follow a linear trend compared to the electrochemical potentials obtained from the neutral forms.

Conclusion

In this study, thermodynamics of electron abstraction of phenol and its derivatives has been theoretically investigated. Linear correlation between Hammett constants and calculated Gibbs free energies of the reaction with published experimental electrochemical potentials are presented. Obtained values of electrochemical potentials have an increasing character from electron-donating substituents to electron-accepting substituents, which is in agreement with the corresponding Hammett constants. As the Hammett constants for *ortho* substituents are not available in literature, this approach is only suitable for mono *meta* and *para* substituted derivatives. On the other hand, the correlation between calculated ΔG and experimental E is also possible for *ortho* derivatives. Furthermore, these correlations provided better coefficients of determination than the linear correlations based on the Hammett constants. It can be concluded that the knowledge of calculated ΔG value for polysubstituted phenol in combination with the linear Eq. 6 allows a reliable determination of the scaled electrochemical potential.

Acknowledgement

The work has been supported by the Ministry of Education, Science, Research and Sport of the Slovak Republic, project VEGA 1/0461/21.

References

Albuquerque BR, Heleno SA, Oliveira B, Barros L, Ferreira ICFR (2021) Food Funct. 12: 14–29.

Atkins PW, De Paula J (2006) Atkins' Physical Chemistry. 8th Edition, Oxford University Press, Oxford, New York.

Biela M, Kleinová A, Uhliar M, Klein E (2023) J. Mol. Graph. Model. 122: 1–6.

Bordwell FG, Zhang XM, Satish AV, Cheng JP (1994) J. Am. Chem. Soc. 116: 6605–6610.

Brown HC, Okamoto Y (1957) J. Am. Chem. Soc. 79: 1913–1917.

Cagardová D, Michalík M, Klein E, Lukeš V, Marković Z (2019) Acta Chim. Slovaca 12: 225–240.

Curtiss LA, Redfern PC, Raghavachari K (2007) J. Chem. Phys. 126: 084108.

Fifen JJ (2013) J. Chem. Theory Comput. 9: 3165–3169.

Frisch MJ, Trucks GW, Schlegel HB, GE Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr., Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Keith T, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS,Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J and Fox DJ (2009) Gaussian 09, Revision D.01, Gaussian, Inc. Wallingford CT.

Guerra R (2001) Chemosphere 44: 1737–1747.

Hansch C, Leo A, Taft R (1991) Chem. Rev. 91: 165–195.

Ho J, Coote LM, Cramer CJ, Truhlar DG (2016) Theoretical Calculation of Reduction Potentials. In: Organic Electrochemistry, 5th edn. CRC Press, pp 229–259.

Ignaczak M, Dziegiec' J, Prawicki K (1987) Kinet. Catal. (Engl Transl); (United States) 27(6): 1305–1309.

Ishikawa A, Nakai H (2016) Chem. Phys. Lett. 650: 159–164.

Jmol development team (2016) Jmol: an open-source Java viewer for chemical structures in 3D (14.6). http://jmol.sourceforge.net/.

Klein E, Lukeš V (2006) J. Phys. Chem. A 110: 12312–12320.

Klein E, Rimarčík J, Lukeš V (2009) Acta Chim. Slovaca 2: 37–51.

Lukeš V, Hartmann H (2021) Color. Technol. 137: 389–398.

Lukeš V, Škorňa P, Michalík M, Klein E (2017) Chem. Phys. Lett. 687: 66–72.

Marenich AV, Cramer CJ, Truhlar DG (2009) J. Phys. Chem. B 113: 6378–6396.

Mennucci B, Tomasi J, Cammi R, Cheeseman JR, Frisch MJ, Devlin FJ, Gabriel S, Stephens PJ (2002) J. Phys. Chem. A 106: 6102–6113.

Michalík M, Biela M, Cagardová D, Lukeš V (2021) Chem. Phys. Lett. 762: 138142.

Michalík M, Lukeš V (2016) Acta Chim. Slovaca 9: 89–94.

Michalík M, Poliak P, Lukeš V, Klein E (2019a) Phytochemistry 166: 112077.
Michalík M, Rimarčík J, Lukeš V, Klein E (2019b) Acta Chim. Slovaca 12: 108—118.
Michalík M, Sádecká L, Lukeš V (2017) Acta Chim. Slovaca 10: 91—95.
Pavitt AS, Bylaska EJ, Tratnyek PG (2017) Environ. Sci.: Process Impacts 19: 339—349.
Rimarčík J, Lukeš V, Klein E, Ilčin M (2010) J. Mol. Struct. theochem. 952: 25—30.
Šimunková M, Biela M, Štekláč M, Hlinčík A, Klein E, Malček M (2022) J. Mol. Liq. 359: 119230
Štellerová D, Lukeš V (2021) Acta Chim. Slovaca 14: 32—37.
Štellerová D, Michalík M, Lukeš V (2022) Phytochemistry 203: 113387.
Takahata Y, Chong DP (2005) Int. J. Quantum Chem. 103: 509—515.
Tawa GJ, Topol IA, Burt SK, Caldwell RA, Rashin AA (1998) J. Chem. Phys. 109: 4852—4863.
Žemlička L, Fodran P, Lukeš V, Vagánek A, Slováková M, Staško A, Dubaj T, Liptaj T, Karabín M, Birošová L, Rapta P (2014) Monatsh. Chem. 145: 1307—1318.
Zhan CG, Dixon DA (2003) J. Phys. Chem. B 107: 4403—4417.

sciendo

Solvent retention capacity as a useful tool for quality evaluation of flours produced from Slovak bred wheat varieties

Tatiana Holkovičová, Zlatica Kohajdová, Michaela Lauková, Lucia Minarovičová, Mária Babulicová

Department of Food Technology, Institute of Food Science and Nutrition, Faculty of Chemical and Food Technology, Slovak University of Technology, Radlinského 9, 812 37 Bratislava, Slovak Republic
tatiana.holkovicova@stuba.sk, zlatica.kohajdova@stuba.sk

Abstract: Wheat varieties (IS Danubius and MS Luneta) bred in Slovakia were assessed for their bakery potential. It was found that flour produced from IS Danubius was characterised by significantly higher level of wet and dry gluten content (30.91 and 20.53 %, respectively) and exhibited higher gluten swelling capacity (29.3 %) in comparison to commercial wheat flour that is usually available in Slovak markets. Solvent retention capacity (solvation in specific solvents) of the tested flours was also determined. The measurements showed that these parameters significantly differ from those determined for commercially available flour, whereas the flour produced from IS Danubius wheat variety had the highest lactic acid retention capacity (124.15 %). Correlation analysis indicated strong positive correlations between solvent retention capacity values and gluten characteristics. Furthermore, it was found that baked rolls prepared from IS Danubius flour showed significantly higher loaf volume compared to commercial wheat flour. This study proves lactic acid retention capacity and sucrose retention capacity as parameters enabling the prediction of gluten network quality in dough and of qualitative parameters of baked goods. From the sensory evaluation resulted that the highest score for overall acceptance was observed in IS Danubius baked rolls, which significantly differs from both MS Luneta and commercial wheat flour.

Keywords: baked rolls, gluten, sensory evaluation, solvent retention capacity, wheat varieties

Introduction

Wheat is the most widely grown crop in different environments and climate conditions because of its great nutritional and economic potential. It is cultivated in two forms, spring, and winter (Dizlek and Awika, 2023). Worldwide, wheat is harvested on more than 220 million hectares while spring forms of wheat still predominate globally (Tadesse et al., 2019). This cereal grain currently provides approximately 20 percent of the total caloric intake of humans. It contains essential amino acids, minerals, and vitamins, coupled with nutritionally beneficial secondary metabolites and dietary fibres (Hellemans et al., 2018).

Enhancing crop yield is important to keep pace with population growth and increasing demand. However, persistent climate trends pose a threat to wheat yields, potentially limiting growth in various regions. Spring wheat is currently cultivated in tropical and subtropical environments often nearing or surpassing optimal temperatures especially during the later stages of grain filling. In addition to conventional agronomic adjustments, a frequently suggested adaptation approach involves the development of new wheat varieties. These varieties aim to integrate enhanced heat tolerance with other coveted traits, including disease resistance and the potential for high yields (Hellemans et al., 2018; Kiszonas and Morris, 2018; Dizlek and Awika, 2023). In the past decades, breeding has improved wheat quality, including flour quality, dough rheological properties, and final product quality (Kiszonas and Morris, 2018).

As wheat flour constitutes the main component of baked products, evaluating its quality contributes to final product quality (Kweon et al., 2011; López and Simsek, 2021). The baking performance of wheat flour is determined by its components, mainly by the quality and concentration of proteins. The most important contribution to baking performance is attributed to the major endosperm protein, gluten, which represents an essential food component because of its rheological and functional properties (Laidig et al., 2017). Analysing the specific functional contributions of each component in flour leads to a more accurate prediction of overall flour functionality. This understanding of the mechanisms behind dough mixing and baking allows reaching optimized product quality. In this regard, solvent retention capacity (SRC) method is a valuable tool to assess flour functionality for soft wheat applications (López and Simsek, 2021; Gong et al., 2023). The SRC test functions as a

solvation assay for flours, relying on the enhanced swelling behaviour of individual polymer networks in specific diagnostic solvents – water, lactic acid, sodium carbonate, and sucrose. These solvents are strategically chosen to predict the functional contribution of each flour component. Widely adopted by wheat breeders, millers, bakers, and cereal research scientists, the SRC method has found application in evaluating relationships between flour SRC profiles and final product quality (Švec a Hrušková, 2017; Gong et al., 2023).

The aim of this study was to evaluate the potential of two Slovak registered wheat varieties (IS Danubius and MS Luneta) for bakery purposes. Gluten characteristics, proximate composition, hydrating properties, and SRC of flours were also determined.

Material and Methods

Raw materials

Wheat flours produced from IS Danubius and MS Luneta varieties were assessed in this study. These samples were obtained from the Research Institute of Plant Production, Piešťany, Slovakia. IS Danubius represents a wheat variety bred by crossbreeding the varieties IS Escoria and Bohemia. Commercially produced fine wheat flour (Mlyn Kolárovo a.s., Slovak Republic, particle size 160–250 μm) (used as a control sample) and other ingredients (vegetable oil, sugar, salt, yeasts) used for the preparation of wheat rolls were purchased in a Slovak local market.

Evaluation of flour samples

The following gluten characteristics were determined in wheat flour samples: wet gluten content, gluten ductility (cm), swelling capacity (cm^3) (STN ISO 46 1011-9), dry gluten content (Kohajdová et al., 2013). Gravimetric method was used to determine moisture in the samples (ICC Standard No. 110/1). Flour acidity was measured according to Kohajdová et al. (2013). Retention capacity of wheat flours and gluten performance index were estimated according to a method adapted from Dvořáček et al. (2012). Starch content was determined by the polarimetric method of Ewers (Alkurdi and Supuka, 2015). Hydrating properties, water holding capacity (WHC), and water retention capacity (WRC) of flour samples were measured according to the method described by Raghavendra et al. (2004).

Rolls preparation

Wheat rolls were prepared according to Kohajdová et al. (2013). The original recipe includes 300 g of wheat flour, 12.06 g of yeasts, 7.5 g of vegetable oil, 5.63 g of salt, 3.22 g of sugar and water to achieve farinographic consistency of 400 BU (Brabender units). The procedure is presented in Fig. 1 (Kohajdová et al., 2013).

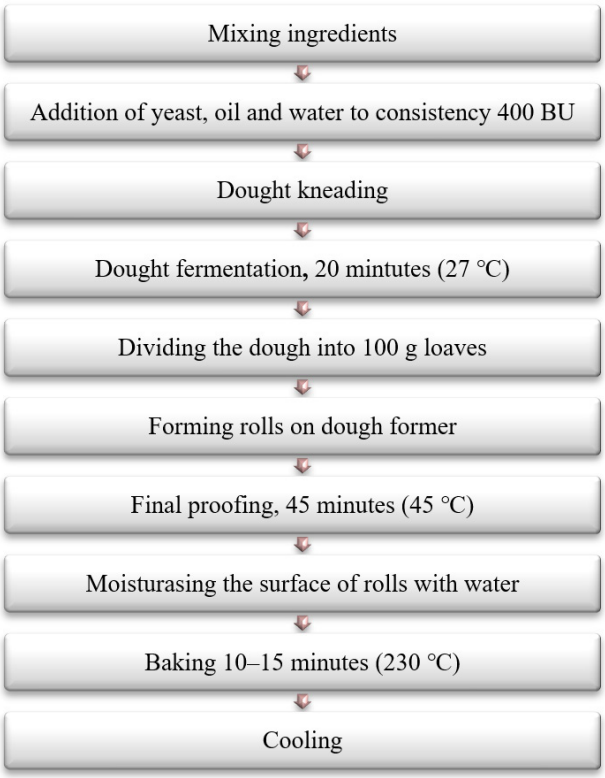


Fig. 1. Flow chart of wheat roll preparation (Kohajdová et al., 2013).

Evaluation of baked rolls

Qualitative parameters of baked rolls were determined two hours after baking. Baking losses (%) represent the weight loss of the product after baking. This parameter was computed as a difference between product weight before and after baking (Kohajdová et al., 2013). Loaf volume (cm^3) of baked rolls was estimated by the rapeseed displacement method (Lauková et al., 2016). Loaf specific volume (cm^3/100 g) was calculated according to Kohajdová et al. (2013). Cambering was calculated as a ratio of rolls height and width (Lauková et al., 2016).

Sensory evaluation

Products were evaluated by seven trained judges using a five-point hedonic scale according to the method described by Kohajdová et al. (2013). The panel consisted of staff and students of the Faculty of Chemical and Food Technology, Slovak University of Technology, Bratislava, Slovakia. The assessed parameters were: shape of products, colour and hardness of crust and crumb, flavour and taste of baked rolls, adhesiveness to palate, springiness, and porosity. Overall acceptability of baked rolls was determined using a 100 mm graphic non-struc-

tured line segment with the description of extremes (100 % — fully acceptable and 0 % — nonacceptable) (Lauková et al., 2016).

Statistical analysis

All analyses were performed in triplicate. The results were expressed as the mean value ± standard deviation. Analysis of variance was conducted to establish differences ($p < 0.05$). Fisher's least significant differences (LSD) test was used to describe the significance of the differences between the Slovak commercial flour sample (control) and flours produced from IS Danubius and MS Luneta wheat varieties. Pearson's correlation test was applied to determine the relationship between SRC values, gluten characteristics, hydrating properties, and starch content using the statistical software Statgraphic version 19 (Statsoft-Inc., Virginia, USA).

Results and Discussion

Flour evaluation

Wheat flour quality is mainly affected by the amount and characteristics of flour proteins (Dizlek and Awika, 2023). Wheat contains a complex mixture of proteins characterized by their distinct ability to create a viscoelastic dough when mixed with water. Gluten characteristics of tested flours are listed in Tab. 1, with gluten, protein constituent of flour responsible for imparting elasticity and strength to the dough, described as the resilient mass remaining after washing wheat dough to eliminate starch granules and water-soluble components (Balamurugan et al., 2018; STN 46 1011-9). Wet gluten yield of tested wheat flours represented 28.24, 28.51 and 30.91 % in dry matter, for control flour, and that produced from MS Luneta and IS Danubius wheat variety, respectively. According to Popovska (2023), optimal wet gluten content for high-quality bread flour is within the range of 30—36 % in dry matter. Dry gluten is the final product after gluten extraction and drying (Kaushik et al., 2015). The percentage of dry gluten in wheat samples was 18.20, 19.32, and 20.53 % in dry matter. According to Kohajdová et al. (2013), dry gluten content represents approximately 1/3 of the wet gluten amount. Thus, flour produced from IS Danubius wheat variety has the best quality from the tested samples. Rheological and functional properties of gluten depend on the ratio of gliadins and glutenins. Each constituent has a distinct function important for the determination of viscoelastic properties and final product quality. Wheat prolamins (gliadins) contribute to gluten ductility, the ability to maintain cohesion under tension (Biesiekierski, 2017). Based on the results, gluten ductility of the studied samples fell within the range considered medium ductile gluten, which is the most suitable for baked goods production (Kohajdová et al., 2013). Gluten swelling capacity is the ability to increase volume in weak acid conditions while maintaining its consistency; this parameter defines flour quality and suitability for bakery purposes (Kohajdová et al., 2013). According to Zálešáková et al. (2004), swelling capacity of gluten washed from high-quality wheat flour is higher than 13 cm^3. All tested samples met this criterion, with IS Danubius and MS Luneta exhibiting significantly higher values compared to the control sample.

Statistical analysis showed that flour produced from the IS Danubius wheat variety had significantly higher wet and dry gluten content than other samples. Statistically significant differences were also found in gluten characteristics of flours produced from the studied wheat (results are not shown).

Chemical composition and hydrating properties of wheat flours are summarised in Tab. 2. Moisture content of powdered materials is a critical parameter influencing their cohesiveness and affecting product stability during storage (Fitzpatrick, 2013). It was determined that flours varied significantly in moisture content, but these values were in agreement with legislative requirements defined in the Food Code of the Slovak Republic (15 %).

Flour acidity serves as an indicator of its freshness and the way in which it is stored. During storage, titratable acidity of flour gradually increases, while the rate of acidity increase in flour is directly proportional to the storage temperature and ash content (Rosentrater and Evers, 2018). The acidity of control sample was significantly higher compared with other samples, although they were

Tab. 1. Gluten characteristics of tested flours.

	Wet gluten content (% in dry matter)	Dry gluten content (% in dry matter)	Gluten ductility (cm)	Gluten swelling capacity (cm^3)
Control	28.24 ±0.31	18.20 ±0.24	11.90 ±0.14	22.67 ±0.47
IS Danubius	30.91 ±0.37*	20.53 ±0.47*	12.07 ±0.17	29.33 ±0.47*
MS Luneta	28.51 ±0.02	19.32 ±0.42	11.67 ±0.12	25.00 ±0.00*

*denotes a statistically significant difference at $p < 0.05$ level.

Tab. 2. Proximate composition and hydrating properties of tested flours.

	Moisture (%)	Starch (%)	WHC ($g \cdot g^{-1}$)	WRC ($g \cdot g^{-1}$)	Flour acidity (mmol/kg)
Control	7.85 ±0.03	76.35 ±0.27	1.95 ±0.22	1.06 ±0.01	32.65 ±0.55
IS Danubius	8.89 ±0.01*	75.64 ±0.13	1.69 ±0.02	1.01 ±0.00*	29.11 ±1.12*
MS Luneta	8.81 ±0.04*	77.29 ±0.36	1.76 ±0.01	0.98 ±0.00*	27.96 ±0.00*

Note: WHC — water holding capacity, WRC — water retention capacity.
*denotes a statistically significant difference at $p < 0.05$ level.

all in acceptable ranges recommended for wheat (20—40 mmol/kg) (Kohajdová et al., 2013).

Starch is the major component of wheat flour and its content and characteristics directly influence rheological properties of dough and quality of final products (Kohajdová et al., 2013). No significant differences in this parameter were found in the control sample and other flours. On the other hand, flours produced from IS Danubius and MS Luneta significantly differ in starch content. However, all flour samples confirm the claim that starch makes up about 80.0 % of wheat flour (Eriksson et al., 2014).

Functional properties (including hydrating properties WHC and WRC) of wheat flour determine its application and use for various food products based on the desirable texture (Lauková et al., 2016). WHC is connected to the porous matrix structure created by polysaccharide chains, which have the capability to retain substantial amounts of water through hydrogen bonding (Zhu et al., 2015). WRC in wheat flour refers to the ability of the flour to hold and retain water after the application of an external force (Liu et al., 2018). No significant differences in WHC were found in the evaluated samples. However, WRC values of flours produced from IS Danubius and MS Luneta significantly differed from those of control flour, which is probably due to their higher content of hydrophilic components, i.e., starch and dietary fibre (Garcia-Valle et al., 2019).

Solvent retention capacity

SRC is a solvation test predicting flour functionality and swelling ability of flour polymeric constituents. For the SRC method, four standard solvents are used to evaluate which specific flour component is responsible for a given variation in flour-swelling behaviour (Kweon et al., 2011; Wessels et al., 2020). Diagnostic solvents included: 5 % lactic acid associated with glutenin characteristics and gluten strength, 5 % sodium carbonate with damaged starch content, 50 % sucrose with pentosans (arabinoxylans) content, and distilled water which is influenced by all flour components (Li et al., 2014). SRC values of studied flours are presented in Tab. 3. Concerning the standard solvents, the trend of SRC profile of all samples is water retention capacity (WSRC) < sodium carbonate retention capacity (SCSRC) < sucrose retention capacity (SUSRC) < lactic acid retention capacity (LASRC), which is in accordance with the trend obtained by López and Simsek (2021), where the supplemental solvents and gluten quality were studied. According to the obtained results, SCSRC of studied samples showed the higher variation in values (73.62—89.85 %), while the lower variation was found for WSRC (68.03—77.75 %).

LASRC value of control flour significantly differs from other flours. Previously, Li et al. (2014) described that this fact is related to protein content and gluten functionality. However, Lindgren and Simsek (2016) and López and Simsek (2021) reported that LASRC is specifically connected with glutenin content instead of total protein content since glutenins are soluble in acid solutions. In general, higher LASRC values indicate bakery products with higher loaf volume, dough strength, and better gluten quality (Kweon et al., 2011).

Alkaline pH of the sodium carbonate solution is essential for interactions with damaged starch. Damaged starch represents small starch particles torn away from bigger starch granules during wheat milling which have high water holding capacity (Li et al., 2014; Lindgren and Simsek, 2016). Control flour sample showed the highest SCSRC, and its value differs significantly from the tested flours. Higher SCSRC values were already reported to indicate higher starch damage. Damaged smaller starch granules have better hydration and swelling capacity during dough preparation. A certain degree of starch damage is desirable to achieve high baking quality (Dvořáček et al., 2012).

Considering the polysaccharide component of grain, non-starch polysaccharide content, namely soluble and insoluble pentosans (arabinoxylans), is mentioned as a significant quality component. These substances play a role in the structural function of gluten complex and, at the same time, in water holding capacity (Dvořáček et al., 2012). Neutral pH of the sucrose solvent facilitates the swelling

of arabinoxylans and interacts specifically with the xylan backbone (Kweon et al., 2011; Lindgren and Simsek, 2016). According to Dvořáček et al. (2012), the interval range of SUSRC for high baking quality of flour is 104.4—113.0 %. It was found that only sample IS Danubius meets the set criteria for high baking quality. Wessels et al. (2020) came to the same conclusion when evaluating the SUSRC of wheat flour in different varieties cultivated in South Africa.

Water represents a reference SRC solvent as it is associated with all functional flour components (Kweon et al., 2011). Among the three samples, IS Danubius flour (72.74 %) showed higher WSRC values. Also, it was determined that WSRC of the control sample was significantly different compared to the other flours studied. The WSRC value is specifically connected with water holding capacity of flour, a crucial factor for the processing and final quality of baked goods. However, the other three SRC solvents exhibit better compatibility with the individual polymer components of flour (Lindgren and Simsek, 2016).

Gluten performance index (GPI) is a derived value based on the lactic acid, sodium carbonate, and sucrose SRC values. It is a predictive parameter which better indicates overall gluten performance in the presence of other modulation networks as the LASRC value alone (Li et al., 2014). GPI can be used as a credible parameter for bread-baking performance prediction (Jeon et al., 2019). GPI of the control sample was significantly lower than those calculated for other tested flours. Jeon et al. (2019) documented that baked goods produced from flours with higher GPI values (0.60—0.70) are likely to exhibit a firmer and more viscoelastic texture after baking. Consistent with these results, IS Danubius and MS Luneta showed significantly higher GPI values than the control sample, suggesting better functionality and quality during the baking process.

Correlation analysis

In cereal research, correlation analysis is used to measure the strength of the linear relationship between two variables and to compute their association. This statistical method is useful in revealing the magnitude and direction of the relationship between different parameters contributing to the quality of cereal products (Chong et al., 2014).

Correlation analysis was applied to estimate the relationships between the SRC values, gluten characteristics, functional properties, and starch content (Tab. 4). High correlations were observed among different SRCs and GPI. SCSRC showed strong positive correlation with WSRC (r = 0.968) and SUSRC (r = 0.967) at the 0.05 % level of significance. High correlation between SCSRC, WSCR, and starch content indicate that starch damage is the main indicator for the determination of water absorption in flours (Ram et al., 2005). Similar results were also reported by Siddiqi et al. (2022) in their study of different wheat cultivars. GPI showed high negative correlations with WSRC (r = –0.925), SUSRC (r = –0.925), SCSRC (r = –0,798), and significant correlation with dry gluten content (r = 0.789, $p < 0.05$). Positive correlation between GPI and dry gluten can be explained by the fact that GPI is a predictor of gluten quality (Siddiqi et al., 2022).

LASRC values positively correlate with wet gluten content (r = 0.945, $p < 0.05$), dry gluten content (r = 0,737, $p < 0.05$), and gluten swelling capacity (r = 0.994, $p < 0.05$) since lactic acid SRC reflects gluten characteristics and functionality (Lindgren and Simsek, 216). Likewise, according to the findings of Wessels et al. (2020) and Gong et al. (2023), positive correlations in LASRC and gluten characteristics were found. Strong negative correlation between LASRC and starch content was also found (r = –0.935) as higher levels of damaged starch led to an increase in lactic acid retention (Dvořáček et al., 2012). Also, WRC was found to be in positive relation with WSRC (r = 0.992, $p < 0.05$), SUSRC (r = 0.992, $p < 0.05$), SCSRC (r = 0.928, $p < 0.05$), and WHC (r = 0.800, $p < 0.05$). WHC showed positive correlation with SUSRC (r = 0.720, $p < 0.05$) and WRSC (r = 0.719), and non-significant negative correlations with wet gluten (r = –0.770), dry gluten (r = –0.960), and GPI (r = –0,929). This can be explained due to the ability of water to

Tab. 3. Retention capacity of wheat flours.

(%)	WSRC	SUSRC	LASRC	SCSRC	GPI
Control	71.75 ±0.76	106.12 ±0.43	117.06 ±0.13	89.85 ±0.59	0.60 ±0.00
IS Danubius	72.74 ±0.26*	100.04 ±1.75*	124.15 ±0.74*	85.14 ±0.58*	0.67 ±0.01*
MS Luneta	68.03 ±0.11*	94.36 ±0.57*	114.71 ±0.37*	73.62 ±0.32*	0.68 ±0.00*

Note: WSRC — water retention capacity, SUSRC — sucrose retention capacity, LASRC — lactic acid retention capacity, SCSRC — sodium carbonate retention capacity, GPI — gluten performance index

*denotes a statistically significant difference at $p < 0.05$ level.

Tab. 4. Correlation analysis.

Parameters	WGC (%)	DGC (%)	GSC (cm^3)	WSRC (%)	SUSRC (%)	LASRC (%)	SCSRC (%)	GPI (%)	SC (%)	WHC ($g \cdot g^{-1}$)	WRC ($g \cdot g^{-1}$)
WGC (%)	1										
DGC (%)	0,917*	1									
GSC (cm^3)	0,903*	0,658*	1								
WSRC (%)	–0,110	–0,496	0,328	1							
SUSRC (%)	–0,111	–0,498	0,326	0,999*	1						
LASRC (%)	0,945*	0,737*	0,994*	0,222	0,220	1					
SCSRC (%)	0,145	–0,261	0,556*	0,968*	0,967*	0,461	1				
GPI (%)	0,480	0,789*	0,057*	–0,925	–0,925	0,166	–0,798	1			
SC (%)	–0,767	–0,449	–0,968	–0,553	–0,552	–0,935	–0,746	0,194	1		
WHC ($g \cdot g^{-1}$)	–0,770	–0,960	–0,421	0,719	0,720*	–0,519	0,520*	–0,929	0,182	1	
WRC ($g \cdot g^{-1}$)	–0,233	–0,601	0,207	0,992*	0,992*	0,098	0,928*	–0,965	–0,444	0,800*	1

Note: WGC — wet gluten content, DGC — dry gluten content, GSC — Gluten swelling capacity, WSRC — water retention capacity, SUSRC — sucrose retention capacity, LASRC — lactic acid retention capacity, SCSRC — sodium carbonate retention capacity, GPI — gluten performance index, SC — starch content, WHC — water holding capacity, WRC — water retention capacity

*indicates a statistically significant difference at $p < 0.05$ level.

swell up and hydrate all polymeric components of flour (Nogueira et al., 2021). Wet gluten was significantly positively correlated with dry gluten content ($r = 0.917$, $p < 0.05$) and swelling capacity ($r = 0.903$, $p < 0.05$).

Qualitative parameters of baked rolls

Qualitative parameters of baked goods refer to the characteristics and attributes that define the sensory and physical qualities of final products (Nashat and Abdullah, 2016). Qualitative parameters of baked rolls are presented in Tab. 5 and their photo documentation is showed in Fig. 2. Loaf volume and specific volume represent important criteria in the assessment of the quality of fresh bread in both industrial quality control and consumer evaluation (Rossmann et al., 2020). While it serves as an indicator of gluten content in bread, it is essential to note that other components, including starch and fibre, also play significant role in determining the volume and specific volume of bread. Therefore, specific volume is a comprehensive measure reflecting various constituents that collectively contribute to the overall quality of bread (Laidig et al., 2022). Significantly higher loaf volume and specific volume was observed in baked rolls produced from the IS Danubius wheat variety compared to other samples; this can be ascribed to higher gluten content in this flour sample (Jankielsohn and Miles, 2017). These findings are in agreement with the study of Schopf and Scherf (2021) who concluded that variations in loaf volume are most affected by the amount of total gluten proteins.

Cambering, expressed as the ratio of loaf height to width, indicates the proportion of these dimensions in loaves. A higher cambering value is preferred and indicates a product with more appealing shape. Ideally, the loaf should exhibit an arched form, smoothly transitioning from the bottom to the upper crust, with higher width-to-height ratio (Bojňanská and Mocko, 2014). Cambering values falling within the range of 0.60 to 0.70 are regarded

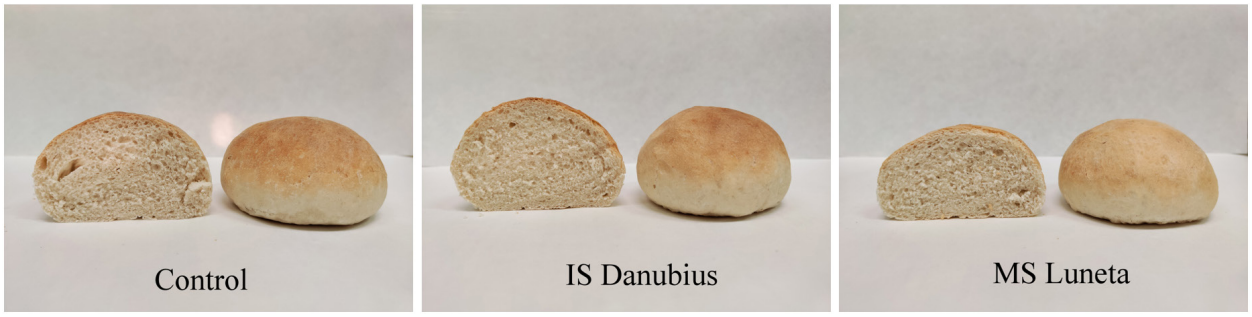


Fig. 2. Baked rolls.

Tab. 5. Qualitative parameters of baked rolls.

	Baking loss (%)	Loaf volume (cm^3)	Loaf specific volume (cm^3/100g)	Cambering
Control	10.80 ±0.44	261.25 ±4.15	287.18 ±3.66	0.62 ±0.02
IS Danubius	8.49 ±0.30*	266.25 ±4.15*	292.14 ±2.98	0.64 ±0.04
MS Luneta	11.54 ±0.61	240.00 ±3.54*	267.12 ±3.25*	0.61 ±0.03

*denotes a statistically significant difference at $p < 0.05$ level.

as favourable, while values below 0.50 are considered inadequate (Kohajdová et al., 2013). From the results it can be concluded that all evaluated samples met the criteria and had good cambering (above 0.60). Best cambering value was observed in the IS Danubius flour sample (0.64), which showed about 3.13 % higher values than the control sample (0.62).

Baking losses are defined as the reduction in the weight of baked goods after the baking process. These losses primarily occur due to the evaporation of moisture during baking. Factors influencing baking losses include the weight of the bakery product, its shape, and moisture content. Typical baking losses for baked goods generally fall within the range from 10 to 15 % (Bojňanská and Mocko, 2014). Baking losses were appropriate in all assessed samples, however, IS Danubius sample showed significantly lower value (8.49 %) than MS Luneta and control samples.

Sensory evaluation

Results of sensory evaluation of individual samples are presented in Fig. 3. Colour of crust and crumb of the baked rolls had the following decreasing trend: IS Danubius > Control > MS Luneta. The colour of crust can be influenced mainly by the flour type, baking temperature and time, presence of enzymes such as amylases and proteases, initial and final humidity (Castro et al., 2017). Baked rolls made from the IS Danubius flour sample showed significantly higher ($p < 0.05$) values of hardness of crust and crumb than baked rolls made from MS Luneta, which can be influenced by higher gluten content in the raw material. The protein content in flour, primarily gluten, plays a vital role in the texture of the baked product. Higher protein content often results in a firmer structure. Moreover, the way the dough is handled, including kneading and shaping, influences gluten development, overall structure, and hardness (Puerta et al., 2021). Baked goods made from IS Danubius wheat variety and control flour had a more acceptable flavour compared to those made from MS Luneta flour. Adhesiveness describes the tendency of the food to stick or adhere to surfaces, such as teeth or the mouth, and reflects the texture and mouthfeel experienced during eating (Durán-Aranguren et al., 2023). Adhesiveness of IS Danubius and MS Luneta baked rolls exhibited significantly higher scores compared to the control flour sample. Dough with good adhesiveness displays a well-developed gluten network (Yovchev et al., 2015). Springiness denotes the speed at which a sample returns to its original shape after deformation (Guiné et al., 2022). The highest value of this textural parameter was found in IS Danubius baked rolls. High springiness values of the baked goods are connected with the freshness

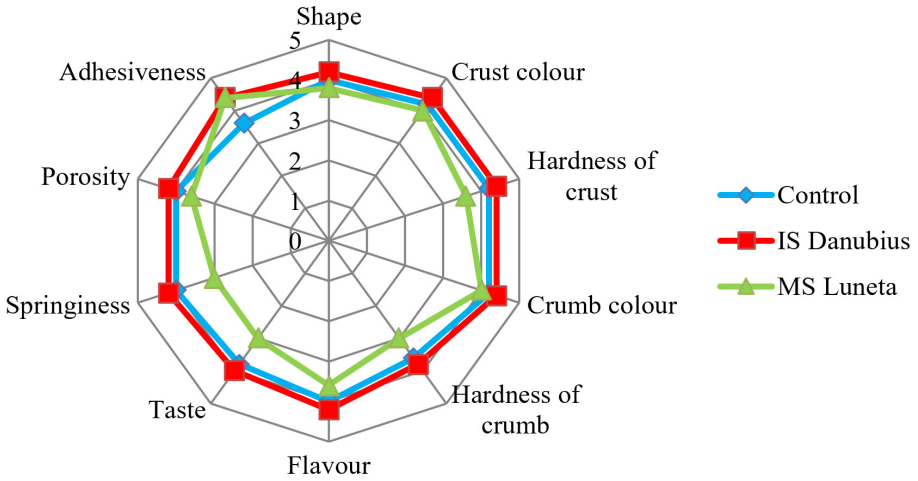


Fig. 3. Sensory evaluation of baked rolls.

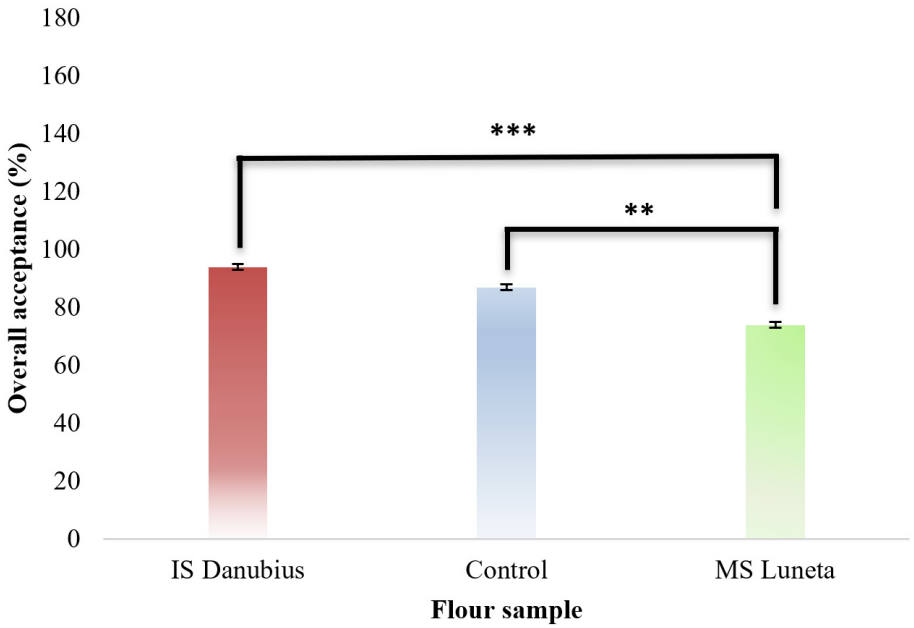


Fig. 4. Overall acceptance of baked rolls.
Asterisk brackets denote a statistically significant difference between samples at $p < 0.01$ (**) and $p < 0.001$ (***) level, respectively.

and elasticity, and thus are preferred (Král et al., 2018). Statistical analysis showed that MS Luneta baked rolls had significantly lower springiness than the control sample. Statistically significant differences in this parameter were also found between IS Danubius and MS Luneta baked rolls. The results show that the porosity of IS Danubius baked rolls significantly differs from that of MS Luneta baked rolls. According to Rathnayake et al. (2018), flour with better protein quality and quantity results in baked goods with improved porous structure characterized by uniformly sized gas cells. IS Danubius and control samples obtained the highest score for the shape of baked rolls. The higher content of gluten and the development of dough structure help in maintaining the shape of the baked product while suitable gluten development contributes to their elasticity and structure (Dufour et al., 2024).
Regarding the overall acceptance (Fig. 4), all samples were evaluated as pleasant. The scores for overall acceptance showed significant differences ($p < 0.01$) between control sample and MS Luneta baked rolls. Moreover, significant differences ($p < 0.001$) were also found between IS Danubius and MS Luneta baked rolls. Baked rolls produced from the IS Danubius wheat variety were rated as the most acceptable by assessors (94 %).

Conclusion

Bakery potential of two wheat varieties (IS Danubius and MS Luneta) cultivated in Slovakia was evaluated in this study. Results have distinctly indicated that there is notable ($p < 0.05$) variability in the physicochemical, functional, and sensory characteristics of the introduced wheat cultivars. Statistical analysis showed that significant differences in gluten content, gluten swelling capacity, and in retention capacity parameters (WSRC, SUSRC, LASRC, SCSRC) were found not only between the control sample and flour prepared from the studied wheat varieties but also between IS Danubius and MS Luneta flours. Baked rolls prepared from IS Danubius wheat variety showed about 20 % higher sensory acceptability than those from MS Luneta flour. In conclusion, the winter wheat variety IS Danubius showed good suitability for baked goods production and has thus been recommended for cultivation and subsequent use in the baking industry. Further studies will be performed to characterise rheological and textural properties of baked rolls to provide data on the elastic and viscous behaviour of dough and the characteristics of baked rolls.

Acknowledgement
This publication was supported by grant VEGA No. 1/0583/20 and by Young Research Support Program/ Excellent Team of STU, Slovak Republic, grant No. 1352.

References

Alkurdi MI, Supuka J (2015) Acta Scient. Polon. 14(2): 3–16.
Balamurugan VS, Sivabalan S, Thamizhselvan I (2018) Int. J. Res. Anal. Rev. 5(3): 426–431.
Biesiekierski JR (2017) J. Gastroen. Hepatol. 32: 78–81.

Bojňanská T, Mocko K, (2014) J. Microb. Biotech. Food Sci. 3(3): 190—194.
Castro W, Oblitas J, Chuquizuta T, Avila-George H (2017) J. Cereal Sci. 74: 194—199.
Dizlek H, Awika JM (2023) Food Chem. 404: 134648.
Dufour M, Chaunier L, Lourdin D, Réguerre AL, Hugon F, Dugué A, Della Valle G (2024) Food Hydrocol. 146: 109214.
Durán-Aranguren DD, Muñoz-Daza LF, Castillo-Hurtado LJ, Posada JA, Mussatto SI, Sierra R, Hernández-Carrión M (2023) LWT — Food Sci. Tech. 185: 1—11, 115174.
Dvořáček V, Papoušková L, Hermuth J, Dotlačil L, Faměra O (2012) Využití metody retenční kapacity mouky pro predikci technologické kvality pšenice v ČR. Výzkumný ústav rostlinné výroby Praha.
Eriksson E, Koch K, Tortoe C, Akonor PT, Baidoo EA (2014) British J. Appl. Sci. Tech. 4 (11): 1609—1621.
Fitzpatrick J (2013) In: Bhandari B, Bansal N, Zhang M, Schuck P. (Ed) Handbook of Food Powders, Vol. 1 (pp 285—308). Woodhead Publishing, Amsterdam.
Garcia-Valle DE, Bello-Perez LA, Flores-Silva PC, Agama-Acevedo E, Tovar J (2019) Frontiers Nutri. 6(2): 1—9.
Gong W, Wang X, Wang F, Wang J (2023) Foods 12(9): 1879.
Guiné RP (2022) Appl. Sci. 12(17): 1—26.
Hellemans T, Landschoot S, Dewitte K, Van Bockstaele F, Vermeir P, Eeckhout M, Haesaert, G (2018) J. Agri. Food Chem. 66(11): 2491—2509.
Chong PH, Yusof YA, Aziz MG, Chin NL, Syed Muhammad SK (2014) Inter. Food Resear. J. 21(4): 1—5.
ICC Standard No. 110/1 (1976) Determine. Moisture Cont. Cereals Cereal Products (Practical method)
Jankielsohn A, Miles C (2017) J. Hortic. Sci. Resear. 1: 42—47.
Jeon S, Baik BK, Kweon M (2019) Cereal Chem. 96(3): 497—507.
Kaushik R, Kumar N, Sihag MK, Ray A (2015) J. Food Sci. Tech. 52: 5930—5937.
Kiszonas AM, Morris CF (2018) Cereal Chem. 95(1): 17—34.
Kohajdová Z, Karovičová J, Jurasová M, Magala M (2013) Cereálie, Cukor, Cukrovinky I — Návody na laboratórne cvičenia FCHPT STU Bratislava.
Král M, Pokorná J, Tremlová B, Ošťádalová M, Trojan V, Vyhnánek T, Walczycka M (2018) Acta Univer. Agricul. Silvicul. Mendelian. Brunen. 66(3): 685—690.
Kweon M, Slade L, Levine H (2011) Cereal Chem. 88(6): 537—552.
Laidig F, Hüsken A, Rentel D, Piepho HP (2022) Theor. Appl. Genetics 135(4): 1331—1343.
Laidig F, Piepho HP, Rentel D, Drobek T, Meyer U, Huesken A (2017) Theor. Appl. Genetics 130: 223—245.
Lauková M, Kohajdová Z, Karovičová J (2016) Acta Chim. Slov. 9(1): 14—18.
Li J, Hou GG, Chen Z, Chung AL, Gehring K (2014) LWT — Food Sci. Tech. 55(1): 43—50.
Lindgren A, Simsek S (2016) J. Food Process. Preserv. 40(2): 131—139.
Liu Y, Xu M, Wu H, Jing L, Gong B, Gou M, Li W (2018) J. Food Sci. Tech. 55: 5142—5152.
López AMM, Simsek S (2021) J. Cereal Sci. 102(1): 103339.
Nashat S, Abdullah MZ (2016) In: Da-Wen Sun (Ed) Computer Vision Technology for Food Quality Evaluation (pp 525—589). Academic Press, London.
Nogueira ADC, Aguiar EVD, Capriles VD, Steel CJ (2021) Cereal Chem. 98(3): 716—728.
Popovska O (2023) Eu. J. Agri. Food Sci. 5(4): 8—12.
Puerta P, Garzón R, Rosell CM, Fiszman S, Laguna L, Tárrega A (2021) LWT — Food Sci. Tech. 140: 1—8, 110718.
Raghavendra SN, Rastogi NK, Raghavarao KSMS, Tharanathan RN (2004) Eu. Food Resear. Tech. 218: 563—567.
Ram S, Dawar V, Singh RP, Shoran J (2005) J. Cereal Sci. 42(2): 261—266.
Rathnayake HA, Navaratne SB, Navaratne CM (2018) Internat. J. of Food Sci. 2018: 1—15 , 8187318.
Rosentrater KA, Evers AD (2018) Kent's Technology of Cereals, Vol. 5. Woodhead Publishing, Cambridge.
Rossmann A, Scherf KA, Rühl G, Greef JM, Mühling KH (2020) J. Cereal Sci. 93: 1—6, 102944.
Schopf M, Scherf KA (2021) Foods 10(2): 228.
Siddiqi RA, Singh TP, Rani M, Sogi DS, Bhat MA (2020) Front. Nutri. 7: 141.
STN 46 1011-9: 1988/1. Testing of cereals, legumes and oil-bearing crops. Cereal testing. Gluten test. Determination of gluten ductility and swelling capacity. The Slovak Office of Standards, Metrology and Testing (UNMS SR).
Švec I, Hrušková M (2017) Chemi. Listy: 111(1), 41—46.
Tadesse W, Sanchez-Garcia M, Assefa SG, Amri A, Bishaw Z, Ogbonnaya FC, Baum M (2019) Crop Breed. Genetics Genomi. 1(1): 1—28.
Food Code of the Slovak Republic. (Potravinový kódex Slovenskej republiky). 2017. Part Three — Commadity Heads. (Tretia časť — komoditné hlavy). Available on: http://www.svps.sk/legislativa/legislativa_kodex.asp.
Wessels R, Wentzel B, Labuschagne MT (2020) J. Cereal Sci. 93: 1—8, 102983.
Yovchev A, Scanlon MG, Nickerson MT (2015) Canadian Food Insights, 31—33.
Zálešáková A, Bielková S, Gregová E, Kraic J (2004) Nova Biotechnologica 235—245.
Zhu F, Du B, Xu B (2015) J. Cereal Sci. 65: 43—47.

sciendo

Prediction of selected properties of aflatoxin molecules by the QSAR method

Alexandra Štecová, Dagmar Štellerová

Institute of Physical Chemistry and Chemical Physics, Slovak University of Technology in Bratislava, Radlinského 9, SK-812 37 Bratislava, Slovakia
dagmar.stellerova@stuba.sk

Abstract: Aflatoxins are naturally occurring compounds produced by fungi, mainly of *Aspergillus* species. All aflatoxins are proved to cause acute toxicity to human health, some even causing chronic diseases such as cancer. These molecules clearly proved that even natural molecules can be directly related to cancer and the formation of tumours. In general, aflatoxins can be characterised as organic compounds; B1, G1, B2, and G2 are produced directly by fungi and these four are subsequently metabolised in biological systems of microbes, animals or humans into other forms, such as M1, EB1, AFL etc. This manuscript provides a brief overview of 14 aflatoxins, their molecular structure and its possible relationship to aflatoxins biological activity. This information, in combination with additional calculations, offers the possibility to investigate the mentioned compounds and their properties using the QSAR approach**.**

Keywords: aflatoxins, biological activity, QSAR analysis, principal component analysis

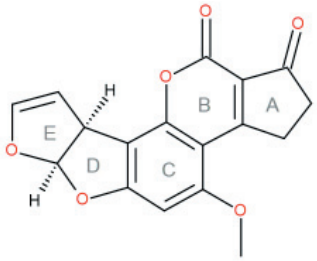
Introduction

Aflatoxins, mycotoxins produced exclusively by fungi, have been studied extensively since the 1960's. These compounds have proved that molecules of purely natural origin can cause not only acute toxicity, but also chronic diseases including immune suppression and malignant cell growth (Hsieh, 1988; Pickova et al., 2021). Liver and kidney cancer diseases are typically associated with aflatoxins and their intermediate metabolites (Benkerroum, 2020a). However, the risk to human body is via the ingestion of contaminated food crops, as well as the consumption of milk, eggs, meat and products made of them in case of livestock and poultry intoxication during feeding. Intensive monitoring of these substances showed their frequent occurrence in agricultural crops. Public concern about safety and quality crop products has increased extensively, forcing regulatory authorities and related stakeholders to develop robust safety evaluation and control policies. (Bennett and Klich, 2003; Elgioushy et al., 2020; World Health Organization, 2018; European Commission, 2006a; European Commission, 2006b).

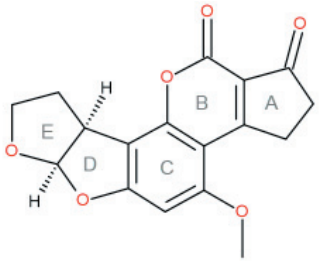
Chemical structure of a molecule has a direct impact on its properties such as bioactivity at the cellular level. Increased probability of toxicity is observed for less polar and more lipophilic compounds (Hughes et al., 2008). Molecular lipophilicity can be expressed as the base ten logarithm of the octanol-water partition coefficient (Log P); lipophilic substances are more soluble in octanol than in water and the value of this coefficient is above zero. On the other hand, negative Log P values indicate hydrophilic substances. The Log P quantity can also be evaluated theoretically using the density functional theory combined with an implicit solvent model (Michalík and Lukeš, 2016). Calculated Log P values correlate well with biological concentration of phenols and their possible quinone metabolites (Michalík et al., 2018a), 3-hydroxynaphthalene-2-carboxanilides (Michalík et al., 2018b), 1-hydroxynaphthalene-2-carboxanilides, and 2-hydroxynaphthalene-1-carboxanilides (Škorňa et al., 2017). Empirical Log P values can be determined using the approaches of chemical informatics, namely the Quantitative Structure-Activity Relationship (QSAR) approach (Muratov et al., 2020). Generally, the QSAR method allows generating relationships between structures and selected chemical and biological properties of molecules.

Based on the severity of the effects aflatoxins have on human health, there are publications devoted to systematic analysis of their chemical structure and its relationship to their biological activity or reactivity (Goldblatt, 1969; Heathcote and Hibbert, 1978; Betina, 1990). Most of available studies are found in the field of toxicology and agricultural sciences, where experimentally selected properties regarding aflatoxins mutagenic or carcinogenic potential are observed for only a few molecules (Wong and Hsieh, 1976). It is also possible to find various studies in the field of chemistry, but these mostly analyse the reactivity of the molecules (Valerio et al., 2007). Regarding experimental data, information on relative mutagenicity is available through the Ames test (Tejs, 2008), which evaluation includes the number of revertants, i.e. the number of *Salmonella* bacteria cells that restored a mutation in their histidine operon. A systematic view of the chemical structure and an

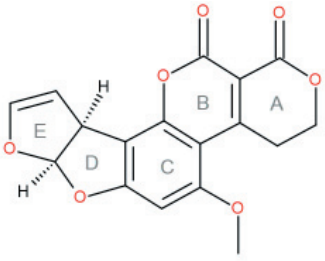
estimate of its possible influence on properties influencing biological effectiveness are missing. For this reason, the study of aflatoxin molecules at the theoretical level was chosen.



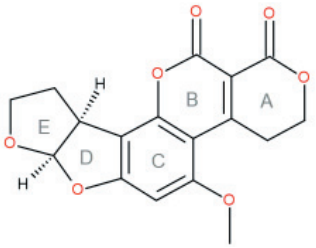
Aflatoxin B1 (**B1**)



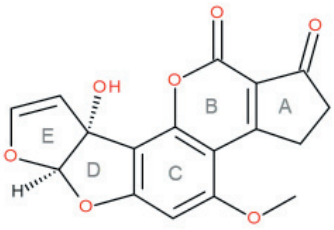
Aflatoxin B2 (**B2**)



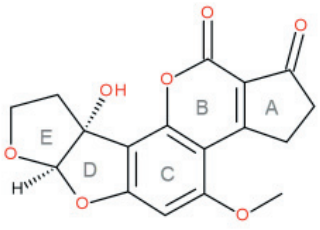
Aflatoxin G1 (**G1**)



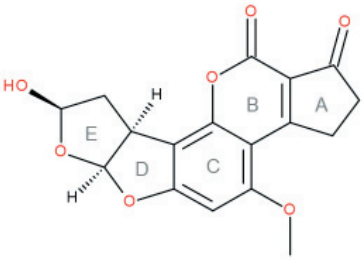
Aflatoxin G2 (**G2**)



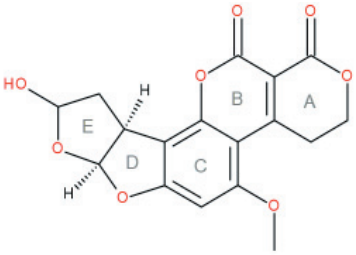
Aflatoxin M1 (**M1**)



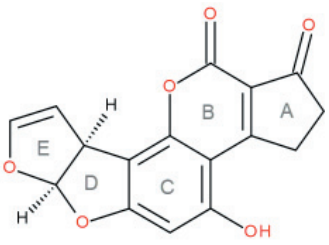
Aflatoxin M2 (**M2**)



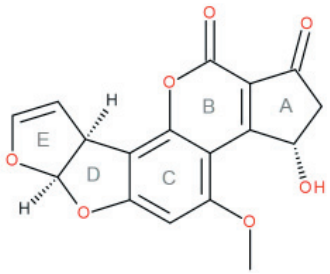
Aflatoxin B2a (**B2a**)



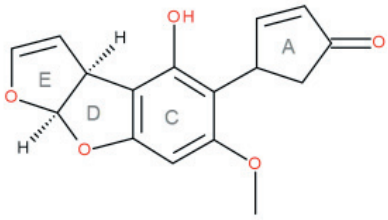
Aflatoxin G2a (**G2a**)



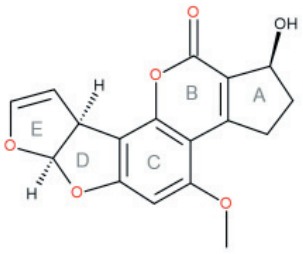
Aflatoxin P1 (**P1**)



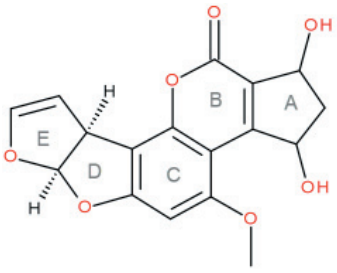
Aflatoxin Q1 (**Q1**)



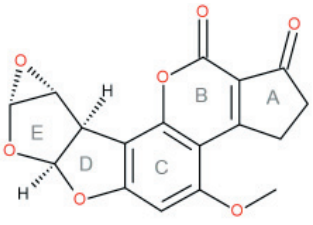
Aflatoxin D1 (**D1**)



Aflatoxicol (**AFL**)



Aflatoxicol H1 (**H1**)



2,3-epoxyaflatoxin B1 (**EB1**)

Fig. 1. Structural formulas and ring notations of studied molecules.

The main aims of the study are: (1) discuss the chemical structure of selected aflatoxins (see Fig. 1), (2) determine chemical and biological properties of the studied molecules using the QSAR method, (3) correlate calculated parameters using statistical principal component analysis (PCA), and (4) determine possible corelation between the evaluated QSAR values with available experimental data for relative mutagenicity of the studied molecules.

Methodology

QSAR is a method based on known experimental data and a statistical model, which enables to determine mathematical quantitative relationship between the structure and activity of the molecule. Properties such as topological polar surface area (TPSA), lipophilicity (miLogP), G-protein-coupled receptors (GPCR), kinase inhibitor (KI), nuclear receptor ligand (NRL), enzyme inhibitor (EI), and protease inhibitor (PI) were monitored. Molinspiration supports internet chemistry community by offering free online services for calculation of the above mentioned important molecular properties (Molinspiration Cheminformatics, 2023). The TPSA descriptor expresses the sum of polar atoms surface area of a molecule; this relevantly characterises the transport of a molecule including its absorption and bioavailability (Kulkarni et al., 2021). According to the Veber's rule (Veber et al., 2002), molecules with TPSA value of 140 $Å^2$ or lower are likely to have good intestinal absorption (Viswanadhan et al., 1989). The miLogP coefficient represents lipophilicity, indicating the degree of molecular hydrophobicity, which directly affects the ability of the molecule to penetrate the cell membrane, its metabolism, and quantifies the hydrophobic interactions at the molecule-receptor level (Soliman et al., 2021).

GPCR, KI, NRL, EI and PI parameters are used to predict biological activity. Biological activity score for organic molecules can be interpreted as active when its value is higher than 0, moderately active when it is between –5.0 and 0.0, and inactive when it is lower than –5.0 (Flores-Holguín et al., 2019).

GPCR parameter describes the biological activity of individual molecules in terms of whether the given molecule can bind to G-protein-coupled receptors. The receptor thus activates intermediate products called G-proteins (Srivastava, 2021), which subsequently activate enzymes, open ion channels, and initiate intracellular signalizing cascades.

Kinase inhibitor is a substance that blocks a type of enzyme called kinase. This feature is characterised by the KI parameter. Human cells have many different kinases which help control important functions such as cell signalizing involved in many diseases including cancer, their metabolism, division, and survival (Kulkarni et al., 2021).

Ligands that show the ability to bind to nuclear receptors and thereby activate them are called ligands for nuclear receptors (NRL). Typically, ligand-biomolecule complexes are formed based on non-covalent interactions (Burdge and Overby, 2020). They include lipophilic substances such as endogenous hormones, vitamins A and D, and xenobiotic hormones. Since the expression of a large number of genes is regulated by nuclear receptors, ligands that activate these receptors can have a significant effect on the organism (Sharma and Raghav, 2021).

Protease inhibitors (PI) are compounds that act by interfering with enzymes that cleave proteins, e.g. inhibitors preventing viral replication by selectively binding to viral proteases and blocking proteolytic cleavage of protein precursors (Shamsi et al., 2021).

Enzyme inhibitors (EI) are substances that block the action of an enzyme. Enzymes help speed up chemical reactions in the body and are involved in many cell functions, including cell signalizing, growth, and division. In cancer treatment, enzyme inhibitors can be used to block certain enzymes that cancer cells need to grow (Kulkarni et al., 2021).

Principal Component Analysis (PCA) can be used for larger sets of data, where possible intercorrelations are expected. This statistical approach is based on orthogonal data transformation allowing to create new element sets. By visualizing them, correlations between groups of evaluated properties can be identified (Sheoran et al., 2020). This statistical approach can show the existence or non-existence of a correlation between the selected properties.

Results

The subject of research are 14 selected aflatoxins. Molecules of aflatoxin **B1** and **G1**, together with their hydrogenated derivates **B2** and **G2** are those best described in literature. Furthermore, there are aflatoxins **M1** and **M2**, which are the 4-hydroxy and 4-dihydroxy derivates of **B1** and **B2,** respectively. Other investigated molecules are their known derivates or metabolites of animal metabolism. An overview of investigated molecules is provided in Figure 1.

From the point of view of organic chemistry, chemical structure of the studied molecules can be limited to three basic fragments formed by five-membered or six-membered rings. The first fragment is formed by only one cyclic ring A, which ensures the greatest variability in structures. In aflatoxins **B1**, **B2**, **B2a**, **M1**, **M2**, **P1**, **Q1**, **AFL** and **H1**,

ring A represents cyclopentene. For aflatoxins **G1**, **G2**, **G2a**, cycle A is a 6-membered lactone ring. The second fragment represents coumarin containing rings B and C, where ring C may contain a methoxy or hydroxyl group. Analogous to hydroxyl group, the methoxy group can also represent important reaction site in metabolic processes and chemical reactivity. For example, the position of methoxy groups has direct influence on anti-cancer chemopreventive effects as inhibitors of P-glycoprotein (Štellerová et al., 2022). In case of guaiacol oxidation, methoxy group demethylation significantly affects thermochemistry of homolytic abstraction of the hydrogen atom from the neighbouring hydroxyl group (Biela et al., 2022). This abstraction leads to cyclization or quinone molecule formation. The hydroxyl group itself represents a potential site of metal ion chelation (Michalík et al., 2021). The third fragment consists of two difuran nuclei D and E, where a double or single bond can alternate on ring E. At the same time, one hydroxyl group can appear on this fragment. In case of the connecting bond between rings E and D, there are two alternative options for the orientation of the hydrogens, resp. hydrogen and a substituent. In this work, only one type of isomer was studied, where the hydrogen atoms or the hydrogen-substituent pair are oriented in the same direction. In this case, the third fragment is the least deformed and therefore more energetically advantageous. Based on this analysis, the selected group of aflatoxins can be divided into two main groups. The first group represents difurocumarocyclopentenones, which typically consist of a series of aflatoxins **B** and their derivates. The second group represents difurocumarolactones with the main representatives of the aflatoxin **G** series, including **G1**, **G2**, and **G2a**. It can be noted that aflatoxins with a saturated (**G2** and **M2**) or hydrated (**B2a**, **G2a**) terminal furan ring show the lowest toxicity. The most widespread and most dangerous aflatoxin **B1** has a special double bond in the cyclic ring E, which is also observed in aflatoxin **G1** and **M1**. Within this framework, several experiments, both *in vivo* and *in vitro,* showed that the potency of aflatoxin toxicity decreases in order **B1** > **G1** > > **B2** > **G2** (Zuckerman et al., 1967; Wogan et al., 1971, Veselý et al., 1983). This suggests the key role of the double bond in this furan group (Benkerroum, 2020b). Structures with a double bond on ring E have poor solubility in water and are prone to undergo an epoxidation reaction. In this regard, one of the direct metabolites of aflatoxin **B1**, molecule **EB1**, is considered as the principle responsible for genotoxicity. As an example of such **EB**1 action, its active isomer, *exo*-8,9-epoxide, interacts with DNA through S_N2 reaction mechanism creating a covalent bond with its guanyl N^7 atom (Johnson and Guengerich, 1997). Molecule **D1** is structurally different as ring B is not formed.

As for the QSAR analysis, based on known experimental data and a statistical model, a mathematical quantitative relationship between the structure and activity of a molecule can be determined. Table 1 summarises, in addition to the parameters mentioned above, basic properties such as molar mass (M), molecular volume (V), and number of atoms (N_{at}), given without hydrogen atoms.

The smallest volume is shown by molecules **P1**, **D1** and **B1**, and on the contrary, molecules **G2a** and **G**2 have the largest volume. All G-molecules contain a 6-membered ring (ring A), which causes a considerably larger bulk compared to the others. The highest value of TPSA is shown by aflatoxin **G2a**, which is closely related to the number of atoms and molar mass, as these values are the highest in **G2a**. It follows that the molecule with the lowest number of atoms and the smallest molar mass shows the lowest TPSA value. Specifically, it is a **D1** molecule that does not contain a B ring. As for miLogP, the highest value is recorded in case of **EB1**, which means that it has the highest ability to penetrate the cell membrane. On the other hand, compound **H1** showed the lowest ability to penetrate the cell membrane. By comparing the GPCR values, it can be concluded that all molecules show moderate biological activity in this property. The exception is the **D1** molecule, which is biologically active. Based on the values listed in Tab. 1, it can be concluded that all aflatoxins investigated show moderate biological activity and thus may have moderate effect on blocking kinases. By monitoring the NRL parameter, it was found that **B1** shows a significantly different value from the rest of the molecules as this value is negative, while for others it is positive. This means that **B1** has lower ability to bind to nuclear receptors than other molecules. Furthermore, it is evident that the PI values are negative and similar for all aflatoxins, therefore, all of them have moderate effect on blocking protein cleavage. In case of the EI parameter, **B1** and **G1** show only negative values, thus showing moderate biological activity. Other aflatoxins are biologically active and are therefore able to block the action of enzymes. It should be noted that this prediction is based on the molecular structure which affects the spatial conditions of chemical reactions. The common structural feature of compounds **B1** and **G1** is that they have a methoxy group on ring C, one double bond on ring E and no hydroxyl group is present. Nevertheless, the enzymatic reactions are a more complex problem as they are affected

also by thermodynamical processes and chemical kinetics. Unfortunately, to the best our knowledge, no experimental values are available in literature, therefore, correct estimation of molecular structure influence on the evaluated parameters is not possible.

Table 1 also shows experimental data for relative mutagenicity and the amount of revertants for the studied compounds (Wong and Hsieh, 1976). Mutagenicity represents a very complex quantity which is dependent on many factors end molecular properties. To better understand the generated QSAR data, statistical principal component analysis (Jolliffe and Cadima, 2016) has been done by linear transformation of the original data into a new coordinate system where the variation in the data can be described within a lower number of dimensions. The linear reliance between each pair of descriptors, Pearson's *r* correlation coefficients, is shown in Table 2. The best positively correlated descriptors are GPCR/KI with the magnitude of 0.852 and EI/PI with the magnitude of 0.796. Relationship between particular inhibitor activities may be related to similarities in QSAR design of enzyme receptor active sides as well as in relation of aflatoxin structures. As for the best negative correlation, TPSA/miLogP surface descriptors showed the value of –0.676. This relationship comes from the structural definition of the descriptors, where miLogP represents lipophilicity of studied aflatoxins and TPSA the proportion of polar groups in their structures.

In Figure 2, the correlation biplot shows the relationship between every QSAR parameter. Each feature is plotted as a vector with origin at point [0; 0]. Important information is the angle of the vectors.

Tab. 1. Properties generated by the QSAR approach (N_{at} is the number of atoms, *M* is molar mass, *V* is volume of the molecule, TPSA is surface area of the molecule, miLogP is lipophilicity, GPCR is a ligand for receptors coupled to G-proteins, KI is kinase inhibitor, NRL is a ligand for nuclear receptors and EI is enzyme inhibitor, PI is protease inhibitor). Experimental relative mutagenicity (RM) and the number of revertants (Rev) have been taken from ref. Wong and Hsieh, 1976.

Molecule	N_{at}	*M* g·mol^{-1}	*V* Ĺ^3	TPSA Ĺ^2	miLogP	GPCR	KI	NRL	EI	PI	RM	Rev
B1	23	312.28	253.24	74.98	1.48	–0.26	–0.43	–0.09	–0.09	–0.36	100	8527
G1	24	328.28	262.22	84.22	1.52	–0.24	–0.38	0.01	–0.01	–0.38	3.3	285
M1	24	328.28	260.93	95.21	0.90	–0.16	–0.42	0.21	0.18	–0.29	3.2	275
EB1	24	324.29	259.20	76.92	1.96	–0.25	–0.57	0.04	0.01	–0.40	–	–
B2a	24	330.29	267.47	95.21	0.91	–0.14	–0.41	0.15	0.25	–0.18	0.0	2
Q1	24	328.28	261.28	95.21	0.49	–0.07	–0.34	0.21	0.17	–0.22	1.1	99
P1	22	298.25	235.71	85.98	1.20	–0.23	–0.41	0.23	0.22	–0.33	0.1	10
G2a	25	346.29	276.45	104.45	0.95	–0.13	–0.36	0.07	0.15	–0.21	0.0	–
D1	21	286.28	244.13	65.00	1.25	0.23	–0.15	0.16	0.39	–0.14	–	–
B2	23	314.29	259.42	74.98	1.57	–0.22	–0.53	0.10	0.23	–0.19	0.2	18
G2	24	330.29	268.41	84.22	1.61	–0.19	–0.47	0.03	0.14	–0.20	0.1	9
M2	24	330.29	267.12	95.21	1.00	–0.13	–0.47	0.26	0.29	–0.16	–	–
AFL	23	314.29	259.10	78.14	1.14	–0.02	–0.28	0.34	0.20	–0.27	22.8	1940
H1	24	330.29	267.14	98.37	0.15	–0.05	–0.26	0.22	0.14	–0.20	2.0	170

Tab. 2. Pearson's *r* correlation coefficients for each pair of QSAR descriptors

Pearson's r	V	TPSA	miLogP	GPCR	KI	NRL	EI	PI
V	–	0.627	–0.244	–0.147	–0.191	–0.110	–0.154	0.290
TPSA	0.627	–	–0.676	–0.180	–0.045	0.262	0.022	0.210
miLogP	–0.244	–0.676	–	–0.409	–0.550	–0.580	–0.327	–0.469
GPCR	–0.147	–0.180	–0.409	–	0.852	0.483	0.679	0.621
KI	–0.191	–0.045	–0.550	0.852	–	0.389	0.392	0.367
NRL	–0.110	0.262	–0.580	0.483	0.389	–	0.683	0.400
EI	–0.154	0.022	–0.327	0.679	0.392	0.683	–	0.796
PI	0.290	0.210	–0.469	0.621	0.367	0.400	0.796	–

The cosine of the angle between a pair of vectors quantifies the correlation between the corresponding variables, cos 90 ° = 0, cos 180 ° = –1, cos 0 ° = 1. Therefore, vectors of uncorrelated properties are perpendicular to each other. The smaller this angle, the greater the correlation they describe. The sign of the correlation is then plus. When both angles are greater than 90 degrees, the correlation is also greater and the sign of the correlation is minus (Kohler and Luniak, 2005).

As illustrated in Figure 2, the vector of TPSA forms an almost right angle with the GPCR vector and a slightly smaller angle with the KI and EI vectors, which indicates the smallest correlation with TPSA among the selected properties. A greater correlation can be seen between the miLogP and NRL, PI vectors which make an angle of approximately 180 degrees, so the cosine of the angle is –1 and the correlation coefficients are negative. There is a fairly strong dependency between TPSA, *V* and GPCR, KI, EI vectors. The correlation coefficients are positive in both cases. Points that are close to each other represent objects with similar values of analysed properties. Points located in the lower part of the plot show similar behaviour for miLogP, GPCR, KI and EI properties; these are points **AFL**, **P1** and **B2**. Points **Q1**, **M2** and **B2a** are the closest to each other, and thus there is the greatest similarity between them. The same trend is shown for point **B1** and its metabolite **EB1**, which are placed distantly from the other points. An object at a large distance from the origin has a large interaction effect with at least one variable, aflatoxin **B1** and 2,3-epoxyaflatoxin **EB1** show this effect with miLogP. Point **G2a** is located approximately on the extended *V* vector due to the largest volume of its molecule among the monitored molecules. Points for compounds **H1** and **D1** stand alone, with no company of other molecules around; thus different behaviour of these molecules can be assumed. It is important to note that the lengths of the vector arrows indicate the characters with the greatest influence on the layout of the graph. However, most vectors are of similar length. It follows that these vectors have a similar effect, except for KI and *V*, with significantly shorter vectors. Thus, these QSAR descriptors have the lowest effect on the arrangement of the graph (Fig. 2).

In case of published experimental data, only four molecules can be selected with the best ratio between the measured value and the estimated measurement error. These include molecules **B1**, **AFL**, **Q1** and **H1**, where there is a double bond enabling the formation of an epoxy bond on the E ring. According to Wong and Hsieh (1976), B1 molecule exhibits the highest mutagenicity (Tab. 1). For **Q1** and **H1** molecules, possible conversion of hydroxyl groups into keto groups can play a specific role as the mechanism of such transformation is connected with proton and electron transfer processes (Lukeš et al., 2022). Despite the relative mutagenicity being mainly connected to aflatoxin epoxy metabolites; we tried to find a correlation between QSAR descriptors and available experi-

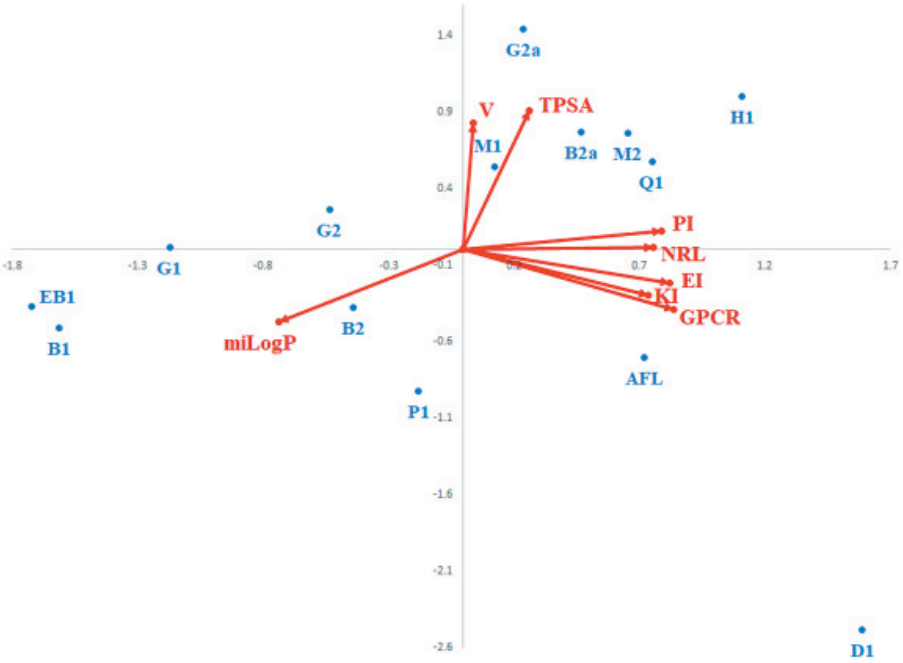
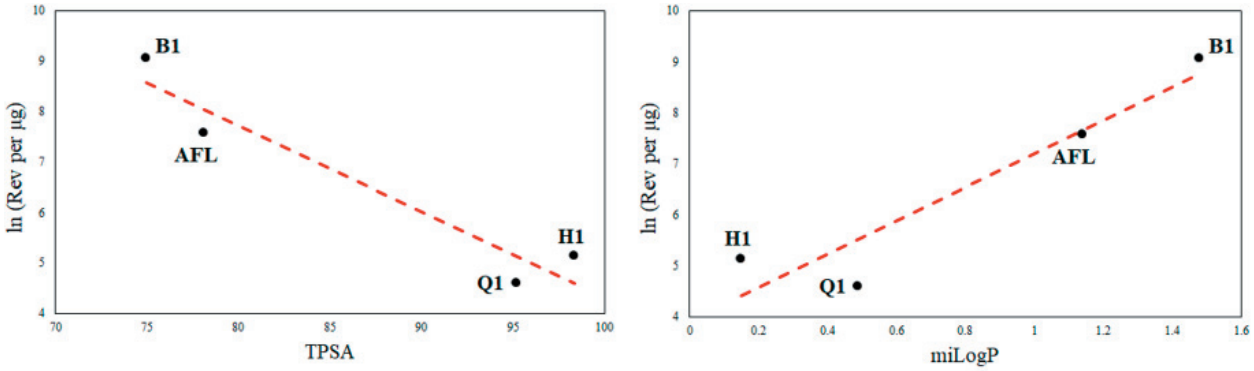


Fig. 2. Correlation between, TPSA, miLogP, GPCR, V, KI, NRL, EI and PI. Red arrows represent correlation vectors.



a) b)

Fig. 3. Dependence of the natural logarithm of the number of revertants on TPSA (a) and miLogP (b).

mental data. There appears to be an interesting correlation with TPSA and miLogP values (curve in Fig. 3), these exponential dependencies are shown in a linearised form for better visualisation. The calculated equations are

$$\ln (\text{Rev}) = -0.1698 \times TPSA + 21.305 \quad (1)$$

$$\ln (\text{Rev}) = 3.2587 \times miLogP + 3.9322 \quad (2)$$

with determination coefficient of 0.9218 for TPSA and 0.8874 for miLogP. Considering linear dependencies for the four parent molecules with experimental data, it is plausible to expect that chemical structures of their mutagenic metabolites are fairly comparable. Unfortunately, the number of points is not sufficient to verify whether this assumption is correct.

Conclusion

To summarise, a systematic QSAR analysis has been carried out during the research. However, to fully confirm the stated results, it is advisable to expand the database of the studied molecules. With more data, the analysis of the quantitative relationship between structure and activity would be more informative. Next, correlations between individual calculated parameters have been determined with the best PCA positive correlation for QSAR inhibition descriptors. On the other hand, miLogP and TPSA parameters correlate negatively. The correlation of experimental data with the TPSA parameter and miLogP are interesting, especially in case of compounds **B1**, **AFL**, **Q1** and **H1**, all of which contain a double bond that may be responsible for the carcinogenic and toxic properties. Future research should concentrate on the characterisation of the chemical structure of aflatoxins and their epoxy metabolites quantum chemically as well as on the reactivity of the mentioned molecules.

Acknowledgements

The work has been supported by the Slovak Research and Development Agency (APVV-20-0213) and VEGA 1/0461/21.

References

Benkerroum N (2020a) Int. J. of Environ. Res. Public Health 17(2): 423.

Benkerroum N (2020b) Int. J. of Environ. Res. Public Health 17(4): 1215.

Bennett JW, Klich M (2003) Clin. Microbiol. Rev. 16(3): 497–516.

Betina, V (1990) Mykotoxíny: Chémia-biológia-ekológia. Alfa, Bratislava.

Biela M, Kleinová A, Klein E (2022) Free Radic. Res. 56(11–12): 730–739.

Burdge J, Overby J (2020) Chemistry: Atoms first (4th Ed). McGraw Hill, New York.

Elgioushy MM, Elgaml SA, El-Adl MM, Hegazy AM, Hashish EA (2020) Environ. Sci. Pollut. Res. Int. 27(28): 35526–35534.

EC-European Commission. (2006a) Commission Regulation (EC) No 1881/2006 of 19 December 2006 setting maximum levels for certain contaminants in foodstuffs. OJEU 70, 12–34.

EC-European Commission. (2006b) Commission Regulation (EC) No 401/2006 of 23 February 2006 laying down the methods of sampling and analysis for the official control of the levels of mycotoxins in foodstuffs. OJEU 364, 5–24.

Flores-Holguín N, Frau J, Glossman-Mitnik D (2019) In: Stefaniu A, Rasul A, Hussain G (Ed) Cheminformatics and its Applications, (pp 78–79). IntechOpen, London.

Goldblatt LA (1969) Aflatoxin: Scientific Background, Control, and Implications. Academic Press, London.

Heathcote JG, Hibbert JR (1978) Aflatoxins: chemical and biological aspects. Elsevier Scientific Publishing Co., Amsterdam.

Hsieh D (1988) In: Natori S, Hashimoto K, Ueno Y (Ed) Mycotoxins and phytotoxins. Third Joint Food and Agriculture Organization (pp 69–80). International Conference of Mycotoxins, Elsevier, Amsterdam.

Hughes JD, Blagg J, Price DA, Bailey S, DeCrescenzo GA, Devraj RV, Ellsworth E, Fobian YM, Gibbs ME, Gilles RW, Greene N, Huang E, Krieger-Burke T, Loesel J, Wager T, Whiteley L, Zhang Y (2008) Bioorg. Med. Chem. Lett. 18(17): 4872—4875.
Johnson WW, Guengerich FP (1997) PNAS 94(12): 6121—6125.
Jolliffe IT, Cadima J (2016) Philos. Trans. A Math. Phys. Eng. Sci. 374(2065): 20150202.
Kohler U, Luniak M (2005) Stata J. 5(2): 208—223.
Kulkarni PS, Walunj YS, Dongare ND (2021) The Chemist 92(1): 53—59.
Lukeš V, Kováčová A, Hartmann H (2022) J. Mol. Liq. 360: 119356.
Michalík M, Lukeš V (2016) Acta Chimica Slovaca 9(2): 89—94.
Michalík M, Poliak P, Klein E, Lukeš V (2018a) Chem. Phys. Lett. 709: 71—76.
Michalík M, Poliak P, Lukeš V (2018b) Acta Chim. Slov. 65(1): 23—33.
Michalík M, Biela M, Cagardová D, Lukeš V (2021) Chem. Phys. Lett. 762: 138142.
Molinspiration Cheminformatics. calculation of Molecular Properties and Bioactivity Score. Computer software (2023), https://www.molinspiration.com/cgi-bin/properties.
Muratov EN, Bajorath J, Sheridan RP, Tetko IV, Filimonov D, Poroikov V, Tudor I, Oprea TI, Baskin II, Varnek A, Roitberg A, Isayev O, Curtalolo S, Fourches D, Cohen Y, Aspuru-Guzik A, Winkler DA, Agrafiotis D, Cherkasov A, Tropsha A (2020) Chem. Soc. Rev. 49(11): 3525—3564.
Pickova D, Ostry V, Toman J, Malir F (2021) Toxins 13(6): 399.
Shamsi TN, Parveen R, Fatima S (2016) Int. J. Biol. Macromol. 91: 1120—1133.
Sharma K, Raghav N (2021) Comput. Toxicol. 19: 100174.
Sheoran OP, Kumar V, Poonia H, Malik K (2020) Int. J. Eng. Adv. Technol. 9(3): 3050—3054.
Škorňa P, Michalík M, Lukeš V, Klein E (2017) J. Mol. Struct. 1144: 473—481.
Soliman K, Grimm F, Wurm CA, Egner A (2021) Sci. Rep. 11(1): 6991.
Srivastava R (2021) ACS Omega 6(38): 24891—24901.
Štellerová D, Michalík M, Lukeš V (2022) Phytochemistry 203: 113387.
Tejs S (2008) Environ. Biotechnol. 4(1): 7—14.
Valerio LG, Arvidson KB, Chanderbhan RF, Contrera JF (2007) Toxicol. Appl. Pharmacol. 222(1): 1—16.
Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD (2002) J. Med. Chem. 45(12): 2615—2623.
Veselý D, Veselá D, Jelínek R (1983) Toxicol. Lett. 15: 297—300.
Viswanadhan VN, Ghose AK, Revankar GR, Robins RK (1989) J. Chem. Inf. Comput. Sci. 29(3): 163—172.
World Health Organization (2018) Aflatoxins Food Safety Digest. World Health Organization, Geneva, Switzerland.
Wong JJ, Hsieh DPH (1976) Proc. Natl. Acad. Sci. USA 73(7): 2241—2244.
Wogan GN, Edwards GS, Newberne PM (1971) Cancer Res. 31: 1936—1942.
Zuckerman AJ, Rees KR, Inman D, Petts V (1967) Nature 214: 814—815.

sciendo

Novel synthesis of known carcinogen precursors

Milata Viktor, Branislav Pavilek, Dušan Bortňák, Zuzana Svítková, Michaela Halinkovičová

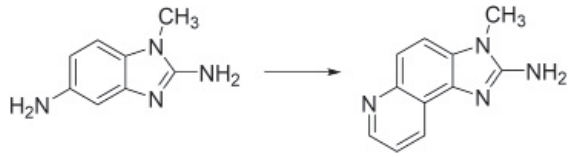
Department of Organic Chemistry, Institute of Organic Chemistry, Catalysis and Petrochemistry, Faculty of Chemical and Food Technology, Slovak University of Technology, Radlinského street 9, SK-812 37 Bratislava, Slovakia
viktor.milata@stuba.sk

Abstract: 5-Amino-1-methylbenzimidazole (**2**) prepared from 2,4-dinitrochlorobenzene in four steps is a suitable reagent for nucleophilic vinylic substitution with various enolethers (**3**) thus affording *N*-substituted enaminoesters — aminoethylene derivatives (**4**). Their thermal cyclocondensation reaction under the Gould-Jacobs protocol gives regioselectively angularly annelated 8-substituted imidazo[4,5-*f*]quinolones (**5**), whose mechanism of origin has been attempted. The obtained results give suitable structures for the synthesis of food-borne carcinogen IQ (**1**).

Keywords: enaminoesters, food-borne carcinogens, Gould-Jacobs reaction, imidazoquinoline, IQ, ^{1}H and ^{13}C NMR

Introduction

Food-borne carcinogen 3-methyl-3*H*-imidazo[4,5-*f*] quinolin-2-amine (IQ, **1**) isolated from thermally treated meat has been prepared by derivatisation or transformation of the imidazole ring (Milata, 2001; Lakshmi, 2004). Our research group simplified this method (Bella, 2012) by developing a convenient and direct IQ synthesis, which involves the incorporation of a pyridine ring in the last reaction step. This synthesis starts from 2,5-diamino-1-methylbenzimidazole (Fig. 1):



1, IQ

Fig. 1. Preparation of food-borne carcinogen (**1**).

If amination of the imidazole ring in position 2 is considered as the final transformation, then various 3-methylimidazo[4,5-*f*]quinolines (**5**) should be suitable precursors for the synthesis of carcinogen IQ (**1**). This synthesis involves the transformation of fused quinolone ring to quinoline (Milata, 1988) in the penultimate reaction step.

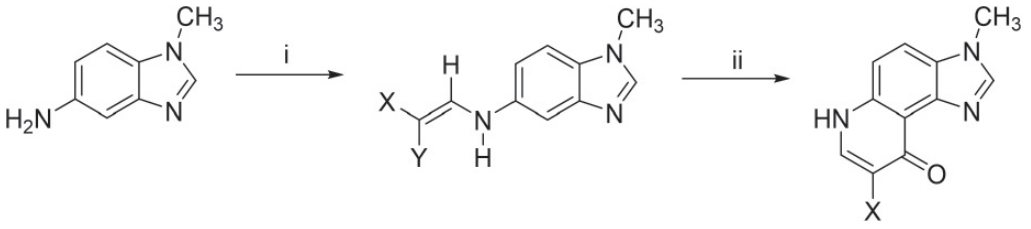
Material and Methods

Methods

All NMR spectra were obtained using an INOVA NMR 300 MHz spectrometer (operating frequencies: 300 MHz (^{1}H) and 75 MHz (^{13}C), equipped with an inverse triple resonance probe and a standard tuneable X/H probe. Tetramethylsilane was used to calculate the ^{1}H and ^{13}C chemical shift scales and for correct reference using the (residual) solvent signals (2.50 and 39.52 ppm for DMSO-d_6, and 7.26 and 77.00 ppm for $CDCl_3$).

Chemicals

Solvents and Dowtherm® were distilled before use. Enolethers (**3**) (Fig. 2) were commercially accessible



2 **4** **5**

Fig. 2. Reaction scheme of the preparation of compounds **4** and **5**.
i: RO—CH═CXY; **3**, **4**, **5a**: X—Y = $(COO)_2CMe_2$ (in this case for **5**: X = H), b: X = Y = COOMe, c: X = Y = COOEt, d: X = COMe, Y = COOMe, e: X = CN, Y = COOEt, ii: 250 °C, Dowtherm.

(**3a**, **3d**) or prepared according to known procedures (Milata, 1989; 2005) using tri(m)ethyl orthoformate and corresponding active methylene compounds (methyl acetoacetate (**3b**), ethyl cyanoacetate (**3c**) and Meldrum's acid (**3e**), respectively). Stannous chloride, hydrochloric acid, and sodium hydroxide were obtained from local vendors.

*5-Amino-1-methylbenzimidazole (**2**)*

Stannous dichloride dihydrate (21 g, 0,093 mol) dissolved in concentrated hydrochloric acid (36 %, 90 mL) was added in parts to stirred 5-nitro-1-methylbenzimidazole (5 g, 0,028 mol). The reaction mixture was refluxed for 5 h at approx. 150 °C in an oil bath. Subsequently, the reaction mixture was allowed to cool to laboratory temperature and precipitated 5-amino-1-methylbenzimidazole hydrochloride was filtered off, washed with ethanol, and dried with the yield of about 4,7 g (90 %). This hydrochloride was neutralised with sodium hydroxide solution (10 %, 1 g/5 ml) and the obtained solution was evaporated to dryness on a rotary vacuum evaporator (RVO) while solid residue was extracted using dichloromethane (4 × 30 ml). Collected extracts were evaporated to dryness on RVO. Yield 3,7 g (89 %), m.p. 165 °C (ethanol) (156—157 °C Ellis, 1974 (benzene)).

^{1}H NMR (DMSO — d_6): 3.71 (3H, s, CH_3); 4.74 (2H, s, NH_2); 6.61 (1H, dd, $^3J_{6,7}$ = 8.5 Hz, $^4J_{6,4}$ = 2.0 Hz, H-6); 6.77 (1H, d, $^3J_{4,6}$ = 2.0 Hz, H-4); 7.19 (1H, d, $^3J_{7,6}$ = 8.5 Hz, H-7); 7.89 (1H, s, H-2)

*General procedure to prepare 5-aminoethylene derivatives (**4**)*

To a suspension of 2 (1 g, 0,006 mol) in methanol or ethanol (60 mL), according to the alkoxygroup in **3,** enolether (**3**) was added (0,007 mol: **3a** — 1,40 g; **3b** — 1,38 g; **3c** — 1,04 g; **3d** — 1,15 g; **3e** — 1,54 g). The reaction mixture was refluxed for 0,5 — 2 hours (**3a** — 0,5 h, **3b** — 0,5 h, **3c** — 2 h, **3d** — 0,5 h, **3e —** 2 h) and monitored using TLC (eluent chloroform : methanol 9 : 1). After cooling of the reaction mixture, the separated crystals were filtered off and washed with cold methanol or ethanol, respectively, and dried. Yields and melting points are presented in Tab. 1:

*General procedure to synthesize 8-substituted imidazo[4,5-f]quinolines (**5**)*

A mixture of derivative 4 and Dowtherm® (15—50 ml) was refluxed at 250—260 °C for 1—9 hours (see Table 2). Reaction progress was monitored by TLC (chloroform : methanol 5 : 1). After cooling of the reaction mixture, the precipitated solid was filtered off, washed by hexane, and dried in vacuum at 80 °C for 6 h to remove traces of Dowtherm. The yields and melting points of the products (**5**) are provided in Tab. 2.

Results and discussion

The starting 5-amino-1-methylbenzimidazole (**2**) has been prepared according to a known procedure (Ellis, 1974) using cheap and commercially available 2,4-dinitrochlorobenzene treated with aqueous methylamine, the nitro group in position 2 was partially reduced with sodium hydrogen sulphide (Zinnin reduction) and the obtained 2-amino-4-nitro-*N*-methylaniline was cyclised with formic acid in total yield of 75—85 %. Initially, 5-nitro-1-methylbenzimidazole was reduced catalytically with hydrogen using Raney-Ni as catalyst at slightly elevated pressure (about 20 kPa). However, the results were not satisfactory, most likely due to

Tab. 1. Yields and melting points of 5-aminoethylene derivatives (**4**).

Compound	3a	3b	3c*	3d	3e**
Yield (%)	70	93	66	82	60
M.p. (°C)	170—172	224—227	180—183	209—212	127—130

See also: *Milata, 1989 ; **Ilavský, 1985.

Tab. 2. 8-Substituted imidazo[4,5-*f*]quinolines (**5**).

Starting derivatives 4	Dowtherm (ml)	Reflux (h)	Yield (%)	M.p. (°C)
4a	15	2	60	245—250
4b	15	0,5	60	255—257
4c	30	3	47	248—252*
4d	50	8,5	40	> 360
4e	20	1,5	50	245—248**

See also: *Milata, 1994; **Ilavský, 1985.

the presence of traces of sulphur or some sulphur-containing compounds poisoning the catalyst. Therefore, stannous dichloride in hydrochloric acid was used to obtain the starting raw material (**2**) more efficiently and in greater amounts.

A set of 8-substituted-3-methylimidazo[4,5-*f*]quinolines (**5**) was prepared, with 5-aminoethylene derivatives of 1-methylbenzimidazole (**4**) as appropriate building blocks for this synthesis. Derivatives **4** were obtained by the reaction of amine **2** and enolethers **3** (Fig. 2).

Due to the possible transesterification, enolethers with the same alkoxygroup attached to a tertiary carbon atom as the alkoxygroups on the ester moiety (**3a—d**) were used. Even though this group is replaced in a nucleophilic vinylic substitution in the first step of the reaction to avoid any possible transesterification, alcohol with the same alkoxygroup was used as a reaction solvent. Nucleophilic vinylic substitution (S_NV) of unequally substituted enol ethers (**3b**, **3c**) takes place with retention of the configuration (Saloň, 2005). These aminoethylenes (**4b**, **4c**) display geometrical isomerism of the double bond (Couchouron, 1983; Milata, 1989; Saloň, 2005) and antiperiplanar configuration of the aminomethylene fragment —NH—CH═ with $^3J_{HH}$ between 10.85—12.84 Hz stabilised with an intramolecular hydrogen bond between proton NH and carbonyl oxygen of the acetyl/alkoxycarbonyl group. In case of derivative **4b**, intramolecular bond with the acetyl group is preferred over the one with the methoxycarbonyl group (Couchouron, 1983; Milata, 1989) — Tabs. 3, 4 Figs. 3, 4:

Thermal cyclisation of **4** in inert media of Dowtherm under Gould-Jacobs reaction conditions afforded only angularly annelated imidazo[4,5-*f*]quinolines. This fact is confirmed by coupling constant $^3J_{HH}$ of the protons of the benzene ring (H_4, H_5 of about

Tab. 3: ^{1}H NMR spectra of compounds **4**.

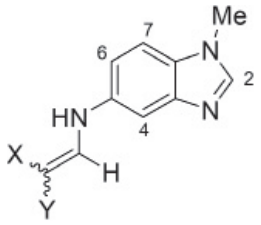


Fig. 3: Numbering of atoms of compounds **4**.

Compound	Solvent	E/Z (%)	H-2	H-4	H-6	H-7	—NH—	—CH═	X, Y	—CH_3
4a	DMSO	–	8.20 s	7.82 d ^4J = 2.1	7.44 dd $^3J_{6,7}$ = 8.7 $^4J_{6,4}$ = 2.1	7.61 d ^3J = 8.7	11.34 d ^3J = 14.5	8.58 d ^3J = 14.5	1.69 s	3.85 s
4b	$CDCl_3$	–	7.89 s	7.58 d ^4J = 2.0	7.13 dd $^3J_{6,7}$ = 8.6 $^4J_{6,4}$ = 2.0	7.37 d ^3J = 8.6	11.15 d ^3J = 13.8	8.61 s	3.87 s 3.85 s	3.78 s
4c*	DMSO	–	7.64 s	a	7.29 d $^3J_{6,7}$ = 8.6	7.57 d ^3J = 8.6	10.85 d ^3J = 14.0	8.44 d ^3J = 14.0	1.25t, 4.10q 1.26t, 4.20q ^3J = 7.02	3.82 s
4d	$CDCl_3$	E:(62) Z:(38)	7.92 s	7.56 d ^3J = 1.9	7.12 dd $^3J_{6,7}$ = 8.6 $^4J_{6,4}$ = 1.9	7.36 d ^3J = 8.6	12.84 d ^3J = 12.8 11.16 d ^3J = 12.8	8.51 s 8.66 d ^3J = 12.8	1.29t, 4.20q ^3J = 7.1 2.50 s 1.34t, 4.28q ^3J = 7.1 2.44 s	3.84 s
4e**	DMSO	Z:(53) E:(47)	7.92 s	7.64 d ^3J = 2.0 7.79 d ^3J = 1.9	7.39 dd $^3J_{6,7}$ = 8.7 $^4J_{6,4}$ = 2.0 7.12 dd $^3J_{6,7}$ = 8.7 $^4J_{6,4}$ = 1.9	7.57 d ^3J = 8.7	a 11.16 d ^3J = 14.0	8.50 d ^3J = 9.7 8.63 d ^3J = 14.0	1.29t, 4.20q ^3J = 7.12 1.30t, 4.28q ^3J = 7.12	3.82 s

[a] signal not observed; See also: *Milata, 1994; **Ilavský, 1985.

Tab. 4. ^{13}C NMR spectra of compounds **4**.

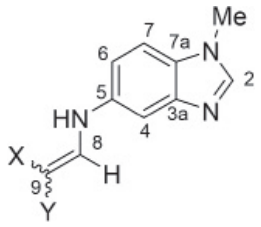


Fig. 4. Numbering of atoms of compounds **4**.

Compound	Solvent	C-2	C-3a	C-4	C-5	C-6	C-7	C-7a	C-8	C-9	Me	X, Y
4a	DMSO	143.8	141.8	104.0	138.9	120.1	114.6	133.1	153.3	98.9	30.8	26.5, 133.5, 162.7, 162.4
4b	$CDCl_3$	144.6	140.0	108.5	142.5	117.3	114.8	132.0	145.1	92.6	31.2	51.3, 166.1, 51.4, 169.5
4c*	DMSO	141.7	139.5	113.2	139.6	119.1	116.2	124.1	149.0	97.2	36.0	14.2, 62.7, 164.9, 64.4, 170.2
4d	$CDCl_3$	143.4	140.1	107.8	134.7	118.3	114.7	132.4	152.5	102.1	30.9	13.9, 31.5, 59.7, 166.6, 200.0
4e**	DMSO	143.7	141.1	107.9	141.7	118.6	115.3	132.5	153.3	73.5	29.9	14.3, 59.4, 115.3, 164.0

See also: *Milata, 1994; **Ilavský, 1985.

9 Hz, Tab. 5, Fig. 6). Similar NMR attributes were found in many previous experiments (Milata, 1987; 1989; 1994; 2000; etc.). The direction of the ring closure occurs preferentially at the C_4-position of the benzimidazole ring which is explained by higher thermodynamically stabilised -complex of two resonance hybrids of the angularly annelated product against the less stabilised linear one (Ishiwata, 1969; Fig. 5):

High temperature of the cyclisation reaction is necessary to eliminate alcohol from the alkoxycarbonyl group of the enaminoester fragment thus producing iminoketene which attacks electron-rich carbon C_4 of the enamine system —N—C_5=C_4. Hydrogen for alcohol production comes from the —NH— group and therefore cyclisation in inert media is possible only if the nitrogen atom is unsubstituted (Ishiwata, 1969; Milata, 2000). For iminoketene formation, reaction temperature close to 250 °C and dilution are very important. Very reactive cyanoiminoketene **4c** (X = CN) requires dilution of 1 g of the substrate/50 ml of Dowtherm, less reactive acetyl analog **4b** (X = $COCH_3$): 1 g/30 ml and ester and their equivalent (**4a**, **4d**, **4e**) 1 g/15–20 ml. Higher dilution complicates product isolation. Upscaling of the reaction requires lower dilutions.

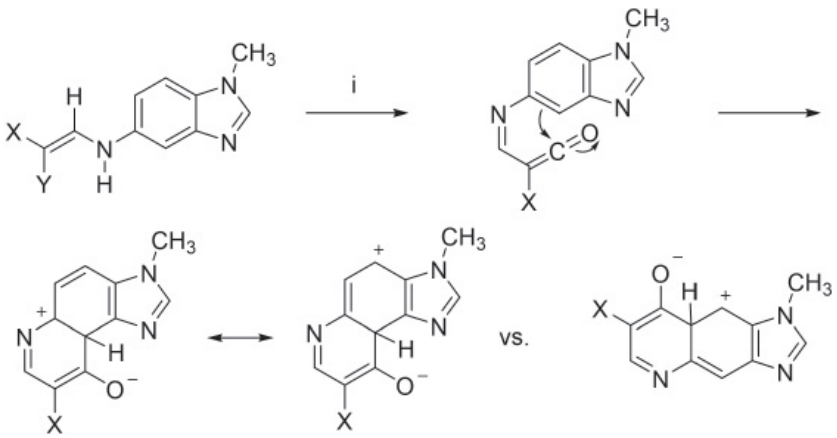


Fig. 5. Proposed mechanism explaining regioselectivity of the cyclisation reaction of compounds **4**. i: –250 °C: —ROH.

Tab. 5: ^{1}H NMR spectra of compounds **5**.

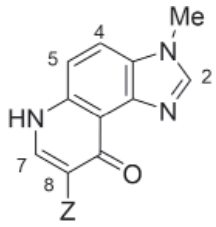


Fig. 6: Numbering of atoms of compounds **5**.

Compound	Solvent	H-2	H-4	H-5	H-7	H-8	Me	—NH—	Z
5a	DMSO	9.04 s	8.06 d ^{3}J = 9.4	8.57 d ^{3}J = 6.7	8.17 d ^{3}J = 9.4	7.37 d ^{3}J = 6.7	4.05 s	11.34 s	–
5b	TFA	9.34 s	8.33 d ^{3}J = 9.5	8.47 d ^{3}J = 9.5	9.32 s	–	4.04 s	11.34 s	4.23 s
5c*	TFA	9.25 s	8.22 d ^{3}J = 9.6	8.35 d ^{3}J = 9.6	9.23 s	–	4.13 s	11.34 s	1.24t, 4.42q ^{3}J = 7.3
5d	TFA	9.29 s	8.15 d ^{3}J = 9.3	8.28 d ^{3}J = 9.3	9.14 s	–	4.04 s	11.30 s	2.56 s
5e**	DMSO	8.54 s	7.52 d ^{3}J = 8.8	7.98 d ^{3}J = 8.8	8.32 s	–	3.93 s	10.70 s	–

See also: *Milata, 1994; **Ilavský, 1985.

Tab. 6: ^{13}C NMR spectra of compounds **5**.

Compound	Solvent	C-2	C-3a	C-4	C-5	C-5a	C-7	C-8	C-9	C-9a	C-9b	Me	Z
5a	DMSO	143.7	123.2	122.5	118.9	141.5	133.1	110.9	171.4	111.4	127.3	36.1	–
5d	DMSO	144.6	125.1	122.1	118.2	141.5	133.4	112.1	174.5	110.8	127.4	35.7	27.0, 205.0, 206.1

Compounds **5b**, **5c**, **5e** are insoluble in DMSO, $CDCl_3$ or TFA.

Conclusions

The Gould-Jacobs reaction has been proven to be an appropriate pathway to fuse a 3-substituted 4-pyridine nucleus to the starting amine, in our case 5-amino-1-methylbenzidazole (**2**). Therefore, 5-amino-1-methylbenzimidazole with a suitable type of enolether can be used to prepare intermediates for food-borne carginogen IQ (**1**). These intermediates were subsequently subjected to thermal cyclocondensation in inert media. A cyclisation mechanism similar to other types of pyridine nucleus fusion (like Skraup, Combes, Conrad-Limpach, Knorr, Meth-Cohn, Doebner reactions, etc. Li (2005)) has been proposed. At the same time, regioselectivity of the thermal cyclocondensation reaction has been explained and all conditions important for successful nuclei transformations studied have been described.

In particular, the potential of nucleus transformation into unsubstituted parent heterocycle was applied to our previous methodology (Milata, 1988) and subsequent amination of the 1-methylimidazole ring showed new perspective for the preparation of food-borne carcinogen IQ (**1**) (Waterhouse, 1985), thus opening a new synthetic route, an alternative to our synthesis (Bella, 2012) or classical methods based on cyclisation of 5,6-diaminopyridines (Milata, 2001). The prepared precursors and IQ (**1**) may have application in characterising chemical carcinogens in cancer hazard identification studies or in describing mechanisms of chemical carcinogenesis (Guyton, 2018).

Acknowledgment

This contribution has been created thanks to the support of the Operational Program Integrated Infrastructure for the project: "Strategic research in the field of SMART monitoring, treatment, and preventive protection against the coronavirus (SARS-CoV-2)", Project no. 313011ASS8, co-financed by the European Regional Development and the Fund of the Agency for Scientific and Technical Assistance based on contract no. APVV-17-0513 and APVV-20-0213.

References

Bella M, Milata V, Larina L (2012) J. Heterocycl. Chem. 49(2): 293—296.

Couchouron B, Le Saint J, Courtot P (1983) Bull. Soc. Chim. Fr. 66—72.

Ellis GP, Jones RT (1974) J. Chem. Soc. Perkin Trans 1, 903—909.

Guyton KZ, Rusyn I, Chiu WA, Corpet DE, van den Berg M, Ross MK, Christiani DC, Beland FA, Smith MT (2018) Carcinogenesis 39: 614—622.

Ilavský D, Bobošík V, Martvoň A (1985) Chem. Pap. 39(4): 527—536.

Ishiwata S, Shiokawa Y (1969) Chem. Pharm. Bull. 17(12): 2455—2460.

Lakshmi VM, Hsu FF, Zenser TV (2004) Chem. Res. Toxicol. 17(5): 709—716.

Li J-J (2005) Name Reactions in Heterocyclic Chemistry, John Wiley and Sons, New Jersey, ISBN 0-47130215-5.

Milata V, Ilavský D (1987) Coll. Czechoslov. Chem. Commun. 52: 2918—2925.

Milata V, Ilavský D, Leško J (1988) Coll. Czechoslov. Chem. Commun. 53: 1068—1077.

Milata V, Ilavský D, Leško J (1998) Coll. Czechoslov. Chem. Commun. 53: 1068—1077.

Milata V, Ilavský D, Goljer I, Leško J, Chahinian M, Henry-Basch E (1989) Coll. Czechoslov. Chem. Commun. 55: 1038—1048.

Milata V, Ilavsky D, Goljer I, Zalibera Ľ (1994) Coll. Czechoslov. Chem. Commun. 59: 1145—1152.

Milata V, Claramunt RM, Elguero J, Zálupský P (2000) Targets in Heterocyclic Systems: 4: 167—203.

Milata V (2001) Tricyclic Azoloquinolines in: Advances in Heterocyclic Chemistry, ed. A.R. Katritzky, Academic Press, San Diego 78: 189—268.

Saloň J, Milata V, Gatial A, Prónayová N, Leško J, Černuchová P, Vo-Thanh G, Loupy A (2005) Eur. J. Org. Chem. 22: 4870—4878.

Waterhouse AL, Rapoport H (1985) J. Labelled Compds. Radiopharm. 22: 201—216.

sciendo

Advancements and challenges in extraction and determination of pesticide residues in soya products

Agneša Szarka, Barbora Vidová, Nicolette Viktoryová, Svetlana Hrouzková

Slovak University of Technology in Bratislava, Faculty of Chemical and Food Technology, Institute of Analytical Chemistry, Radlinského 9, 812 37 Bratislava, Slovak republic
svetlana.hrouzkova@stuba.sk

Abstract: Soya beans, a vital source of protein, oil, and dietary fibre, are cultivated extensively, with a growing demand since 1950. However, the extensive use of pesticides in soya bean cultivation poses potential risks to human health and the environment, especially given the lack of strict regulations for maximum residue levels in soya-based products within the European Union. This study provides a comprehensive overview of isolation and determination methods of pesticide residues from soya products, emphasising the challenges posed by the high-fat content of soya beans. Various chromatographic techniques coupled with mass spectrometry are discussed for their efficacy in pesticide residues analysis. The manuscript also reviews sample preparation techniques, with a focus on the QuEChERS (Quick, Easy, Cheap, Effective, Rugged, and Safe) method, and its modifications to enhance extraction and purification efficiency. Additionally, the study explores the application of liquid-liquid extraction and solid-phase extraction in the analysis of soya oil and beverages and innovative methods, such as dispersive liquid-liquid microextraction and matrix solid-phase dispersion, for comprehensive pesticide residue analysis. This review highlights the importance of multiresidue analytical methods in ensuring the safety of soya products thereby supporting regulatory compliance and public health protection.

Keywords: soya products, pesticides, sample preparation, real sample analysis

Introduction

The vegetarian diet is popular among the population. Vegetarians supplement their diet with legumes (e.g., soy, lentils, peas), nuts (e.g., cashews, macadamia, almonds, pecans, walnuts) and seeds (e.g., chia, flax seeds) or pseudocereals (quinoa, buckwheat). The range of products includes vegan foods such as dairy alternatives (e.g., oat, almond, soy drinks) and cheeses or meat substitutes (e.g., soy-based). A significant part of vegetarian diet is soya beans and various soya products. Soya beans (*Glycine max* (L.) Merr.) have been grown in the Far East since early times and have become of supreme importance as a source of oil and protein throughout the world during the 20th century. Soya bean is considered one of the most widely consumed legume crops in the world. Usually, soya seeds consist of about 36.5 % protein, 19.9 % lipids, 30 % carbohydrates, and 9.3 % dietary fibre. According to information from the World Wildlife Organization, the demand for soya beans has increased up to 15 times since 1950. However, increased use of pesticides in soya bean cultivation is also related to this increased demand. The use of pesticides presents a danger of residues in the final products, posing a risk to the environment and human health. To guarantee the safety of food and feed and to regulate international trade, the European Union (EU) has set maximum residue levels (MRLs) to minimise the presence of pesticide residues in various foods. Agricultural practice in soya bean production requires large amounts of pesticides and there is no current legislation establishing MRLs for pesticide residues in soya-based products, which makes it difficult to control the indiscriminate the use of these compounds and to preserve health of the consumers.

Determination of pesticides in high-fat samples such as soya beans is very complicated, therefore, prior to the actual sample analysis, it is very important to treat the sample with a suitable extraction technique. Chromatographic techniques in conjunction with mass spectrometry (MS) are commonly used to analyse extracts as they provide satisfactory selectivity and sensitivity.

The aim of the study is to provide an overview of methods for the detection and determination of pesticides in soya bean products and of sample preparation techniques used for sample analysis to detect and determine pesticides at low concentration levels. The paper critically evaluates the up-to-date achievements and the authors set challenges for the near future in this area.

Soya and soya products

Soya is an annual legume belonging to the *Fabacean* family with edible seeds. Soya beans are among

the six most useful agricultural crops in the world because they are used as food for humans and animals, but also as a raw material for biodiesel production (Augustyn, 2019).

Soya beans were first cultivated in America in 1804 and remained a rare garden plant for nearly 100 years. They can be grown in almost all types of soil, but produce best in warm, fertile, well-drained sandy loam away from the danger of frost. Like other legumes, soya beans enrich the soil with nitrogen (Rigby, 2001).

Soya products and soya bean oil became very scarce food during World War II when the demand for edible oil increased. Nowadays, soya bean oil is also used in industry, mainly in the production of paints, varnishes, and other drying oil products (Shurtleff, 2009).

Soya bean oil and soya bean meal are a source of fuels and bioproducts. Consumer interests and political interests are focused on achieving high content of biological materials in consumer goods to meet the growing expectations regarding sustainability and the use of renewable resources. Along with these consumer and legislative trends, stricter environmental standards, increased ability to provide improved performance characteristics, and more cost-effective chemical conversion processes lead to increased use of soya products as raw materials and materials for industrial products. These factors provide motivation to develop new materials from soya beans, which may create new markets (Schmitz et al., 2008).

Generally, soya bean proteins represent 30—60 % on dry weight basis; e. g., soy milk contains around 3—4 % of protein (Olías et al., 2023) while storage proteins are the most prevalent. β-Conglycinin (7S) and glycinin (11S) globulins are two storage proteins that make up 80 % of the total protein content of soya. In addition, there are other less abundant storage proteins such as 2S, 9S, and 15S globulins. In addition to these proteins, soya also contains enzymes, protease inhibitors, lectin, and others (García et al., 1997).

Soya beans are abundant in lipids, representing up to approximately 18—24 % of their total dry weight. Storage lipids are deposited mainly in the form of triacylglycerols in oil bodies. A triacylglycerol molecule is composed of three fatty acids esterified to a glycerol backbone. The majority of fatty acids (FAs) found in mature soya beans are unsaturated fatty acids, accounting for approximately 85 % of the total lipid content. Among them, linoleic acid (C18:2n C6) is the most abundant, making up to 50 % of the total fatty acid content (Olías et al., 2023), followed by oleic (18:1), palmitic (16:0), linolenic (18:3), and stearic (18:0) acids (Medic et al., 2014).

Soya beans contain a significant amount of isoflavones, with the main isoflavone compounds, genistein, daidzein, and glycitein, occur in four chemically different forms as aglycone, β-glucoside, acetyl-β-glucoside, and malonyl-β-glucoside. Soya isoflavones have the ability to alleviate symptoms associated with menopause and prevent age-related dysfunctions such as cardiovascular disease and osteoporosis in postmenopausal women, among others, by exhibiting antioxidant and anticarcinogenic effects. On the other hand, there is growing concern about excessive human exposure to isoflavones due to their estrogen-like properties. The presence of isoflavones has also been reported in soya bean leaves, branches, pods, and roots obtained from living plants (Carneiro et al., 2020).

Various diseases and plant pests can have negative impact on the production of soya products. Losses caused by insect damage reduce grain yield and quality and thus also seed quality. Today, grain quality is not an important factor in the sale of soya beans, but this may change in the future. There has been limited research on the effects of disease and insects on these traits, which will change as markets for these traits develop and rapid and accurate assessment of these traits is required. Most of these diseases and insects directly affect soya beans (Rupe, 2008).

Nowadays, new soya bean varieties are developed with the help of hybridisation and genetic modification. Hybridisation enables the isolation of types that provide good yields, are resistant to diseases, and are adaptable in different maturity requirements as well as in different environments. Genetically modified soya beans are engineered, e.g., to be resistant to glyphosate and are among the most widely used genetically modified organisms (GMOs). To ensure soya bean production and prevent pests, a large number of pesticides are used, which can then also be present in form of residues in the final soya bean products, thus endangering consumers (Rigby, 2001).

Presence of pesticide residues in soya and soya products

Pesticide presence in soya and soya products is a significant concern due to the extensive use of chemicals such as glyphosate in soya farming, especially with genetically modified (GM) soya beans (Rigby, 2001). Pesticides can enter soya products through several pathways during cultivation, harvest, and processing. Direct application of pesticides to soya crops, absorption from contaminated soil and water, and air drift are primary sources during cultivation. Post-harvest handling and processing can also introduce pesticide residues through

contaminated equipment and cross-contamination (Páleníková et al., 2015). Environmental factors, such as contaminated soil and water, as well as transport and storage practices, further contribute to the presence of pesticides. Regulatory bodies set MRLs to ensure safety, and most soya products meet these standards. Regular testing is conducted to monitor pesticide residues using methods such as gas chromatography and mass spectrometry.

To minimise contamination, integrated pest management (IPM), organic farming, thorough washing, and proper processing are essential, alongside regular testing and monitoring to ensure that safety standards are met.

Organic soya beans, grown without synthetic pesticides, have lower residues but are not entirely free from contamination. Health risks associated with pesticide residues include endocrine disruption and reproductive issues leading to increased consumer demand for organic and non-GMO soya products.

Sample preparation of soya-based samples

Samples of soya and soya-based products can occur in different forms, with different content of potential interferences (from the point of analysis) in the matrix, therefore analytical methods are here divided into subsections according to the type of matrix. An overview of the analytical methods is summarised in Table 1, where, in addition to samples and types of analytes, extraction techniques, cleaning techniques and instrumental techniques for separation and detection of analytes are listed. The limits of detection and detection of pesticides in real samples are also provided.

Soya beans

Most studies in the field of soya bean product analysis have focused on the analysis of raw soya beans, while up to 40 % of literature reviewed pesticides determination in these products (Table 1). The determination of pesticide residues in soya beans is crucial for ensuring food safety and compliance with regulatory standards. Soya beans, being a major agricultural product, are subject to pesticide application, making it essential to monitor residue levels effectively. Advanced analytical techniques have been developed to detect and quantify these residues with high precision and sensitivity.

For soya bean samples, the most frequently used sample preparation method is the quick, easy, cheap, effective, rugged, and safe (QuEChERS) sample preparation process, which was originally developed by Anastassiades et al. in 2003 (Anastassiades et al., 2003). It has been used for multiresidue pesticide analysis of a wide variety of plant and animal-based matrices. QuEChERS involves a salt-out acetonitrile extraction/partitioning step, followed by a clean-up step using dispersive solid phase extraction (dSPE) with primary secondary amine (PSA) prior to the analysis by gas chromatography-mass spectrometry (GC-MS) or liquid chromatography-tandem mass spectrometry (LC-MS/MS). In the traditional QuEChERS method, acetonitrile is used as the extraction solvent, which is also the preferred option for soya bean samples analysis, however, it was mixed with water in a 1:1 ratio in some cases. Soya bean samples represent a very complex matrix for GC-MS, LC-MS analysis and therefore, the cleaning step (dSPE) is very important. Different types of sorbents have been applied to eliminate interferences from the acetonitrile extract using mostly PSA in combination with other sorbents. The combination of PSA and Florisil was compared with the combination of PSA, Florisil, and C18 (Huertas Perez at al., 2015). The mixture of PSA and C18 sorbents was found to be insufficient for the elimination of matrix effects, where three pesticides, tebuconazole, diazinon, and chlorpyrifos, showed matrix suppression. On the other hand, with the combination of the three sorbents, the recoveries of problematic pesticides ranged between 60 and 75 %, while the recovery of imidacloprid increased up to 90 %. Other compounds were extracted with recoveries ranging from 80 to 90 % (Huertas Pérez et al., 2015).

The QuEChERS procedure without a purification step was used to isolate a mixture of pesticides from soya beans. Acetonitrile with the addition of 1 % acetic acid along with $MgSO_4$ and NaOAc was used for extraction. Satisfactory separation of pesticides (thiabendazole, aminocarb, imazil, methoxuron, carbofuran, imazapyr, and metosulam) was obtained after the addition of 100 mM acetic acid (pH 2.4). Recoveries of the pesticides ranged between 85—120 % with a relative standard deviation (RSD) of less than 6.1 % in all cases. The limits of detection (LODs) were lower than 0.1 $\mu g \cdot kg$ for all pesticides tested, even several times lower than the established MRLs for pesticides in soya beans (Daniel, 2018).

Herrera López et al. (2019) compared different types of sorbents for the purification of soya bean methanol extract. Several sorbents, EMR-Lipid, CAPTIVA, weak cation exchange sorbents (WCX), hydrophilic-lipophilic balanced sorbent (HLB), and a reversed-phase sorbent operating with strong cation exchange (MCX) were tested in the purification process to evaluate their concentration factors and matrix effects. None of the sorbents alone proved suitable for the isolation of all pesticides, and a combination of the two was used for final pu-

rification providing good separation and improving the shape of the chromatographic peaks. Reliable recoveries in the range of 70—120 % were obtained by combining the sorbents. The extracts were analysed by hydrophilic interaction chromatography in combination with MS. An analysis of real samples showed that this method is reliable and can be used in practice.

A new and environmentally friendly method determining pesticide residues in soya beans has been developed by Petrarca et al. (2024), offering a reliable alternative to conventional extraction techniques. This innovative approach integrates the classical QuEChERS extraction method with a dispersive liquid-liquid microextraction (DLLME) step, utilising a deep eutectic solvent (DES) made of camphor and hexanoic acid in a 1:1 molar ratio. This DES has not been previously used in pesticide analysis of complex matrices such as soya beans through gas chromatography-mass spectrometry (Petrarca et al., 2024).

A method based on the determination of pesticides by gas chromatography (GC) combined with isotope dilution (MS) was used for the detection of diazinon, malathion, chlorpyrifos, captan, alpha-endosphan, pendosulphan, iprodione, and cypermethrin in soya beans. Ethyl acetate was used as the extraction solvent, then the extract was dried with $MgSO_4$ and injected into the chromatographic system. The method was not suitable for the determination of tebuconazole as the presence of both the quantifier and qualifier ion was interfered by impurities, but it was possible to detect all the other pesticides investigated. The recoveries ranged from 80 to115 % (Huertas-Pérez et al., 2019).

Maldaner et al. (2008) used the matrix solid-phase dispersion (MSPD) technique, which involves dispersing a sample over a solid support and eluting the compounds with a small volume of solvent. The major advantage of the method is that extraction and purification can be performed in a single step, reducing the overall analysis time. Silica was added to the soya bean sample, the mixture was mixed and injected into propylene syringes containing the co-adsorbent. The syringes were connected to vacuum and pesticides were eluted with ethyl acetate and methanol. Subsequently, the extract was dried with nitrogen and dissolved in acetonitrile. The final extract was then analysed by high performance liquid chromatography (HPLC) (Maldaner, 2008).

Hong et al. (1993) chose to determine pesticides in soya beans by the GC-MS method. A sample of 25 g was ground and extracted with a 2:1 mixture of solvent, acetone and methanol. After filtration through a Buchner funnel, the filtrate was evaporated on a rotary evaporator and water and saturated sodium chloride solution were added. After cleaning the sample with SX-3 sorbent, the sample was dried under a stream of nitrogen and dissolved in 2 ml of hexane. This method of analysis gave reliable recoveries with reasonable precision, and it appears to be suitable for the determination of a mixture of several pesticides. Each pesticide was identified by retention time and characteristic fragment ions using GC-MS (Hong et al., 1993).

From this chapter it is evident that various modifications of the QuEChERS method have been commonly used to analyse soya beans for pesticide extraction. It has been shown that using a mixture of different sorbents, the best purification of the extracts and removal of fat particles from the matrix can be achieved. Different chromatographic techniques, mainly GC and LC coupled with MS, are used for extract analysis. Table 1 summarises a review of articles on the determination of pesticides in soya beans.

Soya oil

Pesticide residues determination in soya oil is a critical aspect of food safety, ensuring that the oil is free from harmful levels of pesticide residues. The complexity of soya oil due to its high fat content and the potential for various types of pesticides presence requires sophisticated analytical techniques. The goal is to achieve accurate, sensitive, and reliable detection and quantification of pesticides. The extraction of pesticides from soya oil often involves liquid-liquid extraction (LLE) or solid-phase extraction (SPE) (Table 1). These techniques aim to separate pesticides from the oil matrix. Common solvents used include acetonitrile, hexane, and ethyl acetate. Post-extraction, a clean-up step with different sorbents is necessary to remove interfering substances.

Dias et al. (2016) conducted a comparative study on different sorbents, evaluating their efficiency in eliminating fats from soya bean oil samples. Utilising the QuEChERS method to extract 165 pesticides, they compared PSA, Z-Sep, and a specialised sorbent, EMR-Lipid, in the purification step. Subsequent analysis by UHPLC-MS/MS revealed that EMR-Lipid offered significant advantages over other sorbents, including more reliable pesticide recoveries and higher repeatability, with recoveries ranging from 70—120 % and RSD <20 % for most pesticides (Dias et al., 2016).

Further research explored a combination of matrix-assisted solid-phase dispersion (MSPD) with mini-column size exclusion chromatography (M-CS-EC) for pyrethroid pesticides separation from soya bean oil. Using an Extrelut-3 cartridge packed with diatomaceous earth and acetonitrile as the extraction

solvent, this method allowed for efficient fat removal and reduced solvent consumption. Di Muccio et al. (1999) reported recoveries of 80—111 % for nine out of 14 pesticides investigated (Di Muccio et al., 1999).

In conclusion, the determination of pesticides in soya oil is essential for ensuring food safety, given the complexity of soya oil fat content and the variety of pesticides used. Accurate, sensitive, and reliable detection methods are critical, often involving LLE or SPE to separate pesticides from the oil matrix. Advanced techniques such as UHPLC-MS/MS and various clean-up procedures, including the use of specialised sorbents like EMR-Lipid, have shown effectiveness in achieving high recoveries and repeatability. Comparative studies highlight the advantages of different extraction and purification methods, such as SMDP combined with mini-column size exclusion chromatography for specific pesticides such as pyrethroids.

Soya milk and other soya-based beverages

Soya milk is a plant-based beverage which became popular as the alternative to cow's milk and it is gaining popularity among people with lactose intolerance and milk protein intolerance. Soya milk is rich in proteins, vitamins, minerals, and the content of lipids is relatively low compared to the original soya beans as lipid content of soya milk is about 2 %. The increased consumption of soya beverages in Brazil and associated pesticide concerns prompted the use of the UHPLC-MS method for pesticide content determination. The QuEChERS method, using 10 ml of acetonitrile, was optimised for purification with C18 and PSA sorbents. May et al. (2017) observed that C18 provided recoveries of pesticides closer to 100 %, with LOD and LOQ ranging from 3—8 $\mu g \cdot L^{-1}$ and 10—25 $\mu g \cdot L^{-1}$, respectively. The chosen combination of QuEChERS and UHPLC-MS methods was found to be suitable for the determination of larger amounts of pesticides in soya beverages in routine laboratory practice (May et al., 2017).

Campillo et al. (2015) investigated the determination of seven strobilurin fungicides in soya beverages using DLLME. Precipitation of proteins in an acidic medium was necessary to prevent analytical interference. After extraction and clean-up, LC with a diode array detector (DAD) was used for the analysis, although none of the fungicides were detected in real samples (Campillo et al., 2015).

The effect of NaOH addition on the pesticide extraction was studied in the determination of pesticides in soya milk by gas chromatography coupled with a flame ionisation detector (GC-FID). For a suitable extraction solvent, the following properties were emphasised: miscibility with aqueous solution, immiscibility with n-hexane, and high extraction efficiency in the presence of aqueous solution and n-hexane. Therefore, four different solvents, i.e. acetonitrile, acetone, methanol, and isopropyl alcohol, were investigated, demonstrating that only ACN formed a three-phase system in the presence of n-hexane after the addition of the Na_2SO_4 reagent. During the pesticides' extraction, n-hexane was shown to be partially dissolved in acetonitrile and the organic phase was collected at the top of the aqueous solution due to its lower density. Removal of the organic phase with the extracted analytes was problematic, and thus, NaOH was added in different volumes and concentrations. Best recoveries were achieved at the NaOH concentrations of 0.1 $mol \cdot L^{-1}$ and volume of 500 µL (Abbaspour et al., 2019a).

Hernández-Borges et al. (2005) determined herbicides in soya milk using capillary electrophoresis (CE) combined with MS. Following dilution, acidification, and SPE extraction, herbicides were eluted with acetonitrile and analysed, achieving LODs at ppb levels with RSD between 3.8 and 8.1 % (Hernández-Borges et al., 2005).

Rejczak et al. (2016) used HPLC-DAD for sulphonylurea herbicide determination in soya milk. Following acetonitrile extraction and Z-Sep purification, recoveries ranged from 61—108 % with RSD <15 %. Chlorosulphuron was detected in real samples at 14.2 $ng \cdot mL^{-1}$ (Rejczak, 2016).

Gionfriddo et al. (2020) introduced a novel DI-SPME (direct immersion solid-phase microextraction) method for soya milk analysis, a representative matrix for food commodities high in proteins and lipids. The method employed a matrix-compatible SPME fibre (PDMS/DVB/PDMS), which outperformed PDMS/DVB fibres in coating robustness and repeatability for extracting pesticides used on soya plants. Key to the method's development was optimising the SPME fibre cleaning protocol, involving a 10-second pre-desorption rinse in a 9:1 water-acetone solution and a 1-minute post-desorption wash in a 1:1 water-acetone solution. This cleaning ensured optimal extraction for up to 120 consecutive extractions in soya milk samples diluted in the ration of 1:1 with ultrapure water. To improve analyte recovery, various organic modifiers were tested and evaluated (Gionfriddo et al., 2020).

The increasing consumption of soya beverages has driven the optimisation of methods for pesticide determination, with approaches such as QuEChERS combined with UHPLC-MS proving suitable for routine laboratory analysis. Additional techniques, such as DLLME and GC with various detectors, have also been explored, emphasising

the importance of solvent selection and clean-up steps to avoid analytical interference. Innovative methods, such as DI-SPME with matrix-compatible fibres, show advancement in pesticide analysis for complex matrices such as soya milk.

Flour

Soya flour is a product made by grinding roasted or non-roasted soybeans into a fine powder. It is a versatile ingredient that can be used in a variety of culinary applications. Soya flour is rich in protein, dietary fibre, and other nutrients. Owing to the dry nature of this sample, special approach to sample extraction must be assumed.

For the detection of pesticides in soya bean flour, a solid-liquid extraction (SLE) method was employed, followed by HPLC with a diode array detector (HPLC-DAD). The extraction process relied on the differential affinities of sample components across three phases: aqueous solution, acetonitrile, and n-hexane. Initially, acetonitrile was added to the solid sample, facilitating the extraction of analytes along with other compounds, such as lipids. To purify the sample, an aqueous phase (NaOH solution), a cleaning solvent (n-hexane), and a phase-separation agent (NaCl) were introduced to the extract. The addition of NaCl resulted in a three-phase system where hydrophobic compounds were extracted into n-hexane, polar compounds into the aqueous phase, and the analytes remained in the acetonitrile phase. The extraction yielded efficiencies between 55 % and 90 % (Abbaspour et al., 2019b).

Pizzutti et al. (2007) developed a method using liquid chromatography-mass spectrometry (LC-MS) for pesticide determination in soya bean flour. Flour samples underwent acetone-based extraction without subsequent purification by sorbents. The method demonstrated excellent linearity for calibration curves in the range of 0.1 or 0.25 to 10.0 $ng \cdot mL^{-1}$, with LODs generally spanning from 0.1 to 0.25 $ng \cdot mL^{-1}$. Recovery rates ranged from 70 % to 120 %. This acetone-based extraction method proved reliable for quantifying significant pesticide residues in the complex matrix of soya bean rice, thus proving suitable for routine surveillance and research application (Pizzutti et al., 2007).

Soya bean rice is similar to soya flour concerning the composition and nature. For both soya bean rice and flour, LLE and acetone-based extraction methods were utilised. Post-extraction, liquid chromatography (LC) was employed to separate the analytes. The overall recovery rates for pesticides ranged from 55 % to 120 % across both methods, demonstrating the effectiveness of these extraction and analysis techniques in dealing with complex food matrices.

The use of LLE followed by HPLC-DAD and LC-MS methods has proved to be effective for the determination of pesticides in soya bean flour and rice. The extraction processes, optimised with different solvents and phase-separation techniques, ensure high recovery rates and accurate detection, making these methods suitable for both routine and detailed pesticide analysis in complex matrices.

Dietary supplements

Soya-based dietary supplements cover a wide range of products intended for direct consumption improving the health of the population. They may contain soya powder, rich in proteins, or they are based on extracts with a variable composition. Research on the determination of pesticides in these supplements primarily focuses on developing and validating analytical methods to ensure the safety and quality of these products. Pesticides, used in agriculture to protect crops from pests, can remain as residues in food products including soya-based supplements. Given the potential health risks associated with pesticide exposure, it is crucial to monitor and regulate these residues.

Páleníková et al. (2015) investigated the optimisation of pesticide extraction methods from soya dietary supplement samples using a modified QuEChERS approach. Their study evaluated various extraction solvents and purification procedures. The most effective conditions were identified using a combination of PSA, C18, GBC, and Zr-Sep+ sorbents with ethyl acetate as the solvent. This method consistently yielded reliable recoveries between 70 % and 120 % for 92 % of the pesticides studied. In contrast, the use of acetonitrile as the solvent resulted in reliable recoveries for only 28 % of the pesticides, which is likely connected to analyte loss during evaporation (Páleníková et al., 2015).

Alves et al. (2016) employed a modified rapid extraction method, quick polar pesticides method (QuPPe), coupled with liquid chromatography and mass spectrometry to determine polar pesticides in soya dietary supplements. Their study explored various experimental conditions, including solvent selection, oxygenation, extraction time, and cleaning sorbents. Acetonitrile acidified with formic acid was used as the extraction solvent, leading to a viscous mixture unsuitable for injection into the liquid chromatograph. Optimal separation was achieved using a hypercarb-porous graphitic carbon stationary phase. Pesticide recoveries ranged from 46 % to 119 % (Alves et al., 2016).

Soya-based infant foods were analysed by LC/ESI-MS/MS. The method includes liquid–liquid partitioning and solid-phase extraction (SPE) with Oasis HLB cartridges to remove soya proteins and

Tab. 1. Overview of sample analysis and preparation details of published analytical methods applied for soya products analysis.

Sample	Pesticides	Sample preparation	Clean up	Instrumentation	LOD (µg·kg^{-1} or µg·L^{-1})	Real sample findings	Ref.
soya bean	10 pesticides	QuEChERS	C18, PSA, Florisil	LC-MS	1—5	ND	Huerta-Peréz, 2015
soya bean	Glyphosate and 13 polar pesticides	SLE (methanol)	EMR-Lipid, CAPTIVA, WCX, HLB, MCX	HILIC LC-MS	0.02—0.05	ND	Herrera, 2019
soya bean	8 pesticides	SLE	PSA, C18	GC-IDMS	10—25	ND	Huerta-Pérez, 2019
soya bean	7 pesticides	QuEChERS	$MgSO_4$, PSA	CE-MS	less than 0.1	ND	Daniel, 2018
soya bean	25 pesticides	SLE	Bio-Beads S-X3	GC-MS	2—50	ND	Hong, 1993
soya bean	6 pesticides	MSPD	InertSep C8	HPLC	40—80 µg · g^{-1}	ND	Maldaner, 2008
soya bean	30 OPP	QuEChERS	PSA, DSC-18	GC-MS	0.2—3	Chlorpyriphos, Pirimiphos-methyl	Marchis, 2012
soya bean	257 pesticides and micotoxins	QuEChERS	Florisil	LC-MS	5—10	ND	Martínez-Domínquez, 2016
soya bean	203 pesticides	SLE (hexane, acetonitrile)	SPE Florisil	GC-MS	less than 10	ND	Shin, 2018
soya bean	10 pesticides	QuEChERS	DLLME (200 µL DES)	GC-MS	5—26	ND	Petrarca, 2024
soya bean oil	169 pesticides mixture	QuEChERS	PSA, Z-Sep, EMR-Lipid	UHPLC-QqQ-MS	0.01—0.05	Azoxystrobin <LOQ	Dias, 2016
soya bean oil, and other oils	group of 14 pyrethroid pesticides	SMDP	Florisil, Alumina	M-CS-EC	13—53	ND	Di Muccio, 1999
soya bean oil, nuts, sesame oil	28 pesticides	dSPE (PSA, C18)	Low temperature cleanup	GC-FPD, GC-MS	20—255	ND	Li 2007
soya bean oil	126 pesticides	LLE (hexane, acetonitrile)	SPE (Captiva EMR-Lipid)	GC-MS/MS	5—50	ND	Wang, 2021
soya milk	7 pesticides	LLE DLLME	PSA	DSPE	0.035—0.12	Chlopyriphos 19 µg · L^{-1}	Abbaspour, 2019
soya milk	group of triazolopyrimidine sulphonanilide pesticides	SPE	NaOH	CE-MS	less than 74	ND	Hernández, 2005
soya milk	8 sulphonylurea herbicides	QuEChERS	Z-Sep	HPLC-DAD, LC-MS	10—100	Chlorsulphuron 14.2 ng · mL^{-1}	Rejczak, 2016
soya milk	10 pesticides	DI-SPME (PDMS/DVB/PDMS)	–	GC-MS	1—2.5	Trifluralin Dimethoate Malathion Chlorpyrifos Thiabendazole λ-Cyhalothrin β-Cyfluthrin Esfenvalerate 0.05—8 µg · kg^{-1}	Gionfriddo 2020

Tab. 1. (continue) Overview of sample analysis and preparation details of published analytical methods applied for soya products analysis.

Sample	Pesticides	Sample preparation	Clean up	Instrumentation	LOD (μg·kg^{-1} or μg·L^{-1})	Real sample findings	Ref.
soya drink	group of 39 pesticides	QuEChERS d-SPE	C18, PSA	UHPLC-MS	10—25	ND	May, 2017
soya drink	strobilurin fungicides	DLLME	ND	LC-DAD-MS	0.8—4.5	ND	Campillo, 2015
soya rice	group of 169 pesticides	LLE	ND	LC-MS	0.001—0.025	ND	Pizzutti, 2007
soya flour	group of 9 pesticides	LLE LLME	C18, Florisil	HPLC-DAD	0.016—0.06	ND	Abbaspour, 2019
soya-based infant formula	13 pesticides	SLE (acetonitrile, hexane)	SPE (Oasis HLB Plus)	LC-MS/MS	0.1—0.6	ND	Wang, 2006
soya-based infant formula	glyphosate and AMPA	SLE (chloroform)	SPE	HPLC-FD	20	Glyphosate 0.03—1.08 mg · kg^{-1} AMPA 0.02—0.17 mg · kg^{-1}	Rodrigues, 2018
soya nutritional supplements	4 polar pesticides (Chlorate, Fosetyl-Al, Maleic Hydrazide, Perchlorate)	QuPPe	Hypercarb	LC-MS	4—100	Chlorate 63 to 1642 μg · kg^{-1}	Alves, 2016
soya nutritional supplements	group of 177 pesticides	QuEChERS	PSA, C18, GBC, Z-Sep	GC-QqQ-MS	0.1—10	Malathion 11.1 μg · kg^{-1}, Pyriproxyfen 1.5 μg · kg^{-1}	Páleníková, 2015

other interferences and isolate pesticides from the samples. Isoprocarb was used as the internal standard (IS), which helps compensate for pesticide losses and variations during extraction, mitigating matrix effects and ensuring accurate quantitative results. However, due to matrix effects and poor accuracy, only semiquantitative analyses should be done for propoxur and aldicarb (Wang et al., 2006). Overall, these studies highlight the importance of method optimisation and selection of appropriate extraction solvents and purification procedures for accurate pesticide analysis in soya dietary supplements to ensure consumer safety and regulatory compliance.

Real-life findings

The analysis of pesticide residues in soya and soya products is crucial for ensuring food safety and public health protection. Pesticide residues in soya products can pose health risks to consumers if present at levels exceeding regulatory limits or if they include pesticides known to be harmful to human health. Analysing real samples allows for the assessment of actual exposure risks and ensures that products on the market are safe for consumption.

Studies have identified various pesticides, including glyphosate and organo-chlorine pesticides. Glyphosate, in particular, is frequently detected due to its widespread use in soya bean cultivation. Although glyphosate is a universally used pesticide, monitoring has been hampered by analytical difficulties due to its high polarity. Kolakowski et al. (2020) examined 204 samples of soya-based products, including canned/frozen soya beans, soya beverages, and meat alternatives. The prevalence of glyphosate residues varied by product type, with no residues detected in canned soya beans and tofu, and up to 37 % in the analysed soya flour. The maximum glyphosate levels ranged from 0.013 ppm in frozen soya beans to 6.0 ppm in soya flour. These levels were significantly below the maximum residue limits. Soya flour consistently showed higher contaminant levels compared to other soya products (Kolakowski et al., 2020). Glyphosate and amino methyl phosphonic acid (AMPA) were detected in baby food formula by Rodrigues and Souza (2018), marking the first study to report glyphosate in such products. Their find-

ings showed a detection rate of 50.5 % for glyphosate and 19 % for AMPA, with mean concentrations of 0.19 $mg \cdot kg^{-1}$ and 0.05 $mg \cdot kg^{-1}$, respectively (Rodrigues et al., 2018).

Detection of polar pesticides like chlorate and perchlorate has also been investigated, as it requires specialised analytical methods due to the chemical properties of these compounds. Although chlorate has been banned in the EU since 2010, it was determined in five soya-based dietary supplements out of the fourteen analysed samples at high concentrations ranging from 63 to 1642 $\mu g \cdot kg^{-1}$ (Alves et al., 2016).

A high number of pesticides were detected in commercial samples of soya milk from two different brands (Gionfriddo et al., 2020). The concentration of dimethoate in one of the samples was higher than the MRL set for soya products (118.9 $\mu g \cdot kg^{-1}$). Overall, analysing real soya samples for pesticide findings is essential for safeguarding public health, ensuring regulatory compliance, maintaining product quality, and minimising environmental impact associated with pesticide use in soya bean production.

Future trends, challenges and conclusions

This manuscript highlights the growing significance of soya beans in vegetarian diet and their extensive use in various industries while emphasising the critical need to monitor pesticide residues in soya products for food safety. The increased demand for soya beans has led to greater pesticide use, posing risks to human health and the environment. Various analytical methods, particularly modifications of the QuEChERS method, have been developed and optimised to effectively extract pesticide residues in complex soya matrices. The review covers sample preparation and analysis of soya beans, soya oil, soya beverages, flour, and dietary supplements, demonstrating the efficacy of advanced chromatographic and mass spectrometric techniques. Ensuring the safety of soya products requires continual improvement and validation of these analytical methods to meet regulatory standards and protect consumers. The findings underscore the necessity for stringent monitoring and innovative approaches to managing pesticide residues in the face of increased soya bean consumption. Future trend is the development of analytical methods applicable to the detection and quantification of multiple groups of analytes in a single analysis with increased degree of sensitivity selectivity, as well as of faster, easier and solventless or solvent-minimised methods. Progress in procedures meeting the green analytical sample preparation and green analytical chemistry requirements is expected.

Acknowledgment

This work was supported by the Slovak Research and Development Agency under Contract No. APVV-19-0149, the Scientific Grant Agency of the Ministry of Education of the Slovak Republic (VEGA project no. 1/0332/24), as well as by the STU Young Researchers Support Program.

References

Abbaspour M (2019a) J. Sci. Food Agri. 99(8): 4094—4104.

Abbaspour M, Farajzadeh MA, Sorouraddin SM, Mohebbi A (2019b) Anal. Methods 11(31): 4022—4033.

Alves DR, Romero-González R, López-Ruiz R, Jimenéz-Medina ML, Garrido Frenich A (2016) Anal. Bioanal. Chem. 408(28): 8089—8098.

Anastassiades M, Lehotay SJ, Stajnbaher D, Schenck FJ (2003) J. AOAC Int. 86 (2): 412—431.

Augustyn A (2019) Soybean (Glycine max). Retrieved from https://www.britannica.com 2023-12-13.

Campillo N, Iniesta MJ, Vinas P, Hernández-Córdoba M (2015) Food Add. Cont. 32(12): 2039—2047.

Carneiro AM, Moreira AE, Bragagnolo FS, Borges SM, Pilon AC, Rinaldo D, Funari CS (2020) Food Res. Intern. 130: 108949.

Daniel D (2018) Applications Agilent Technologies 2018: 1—6.

Dias JV, Cutillas V, Lozano A, Pizzutti IR, Fernández-Alba AR (2016) J. Chromatogr. A 1462: 8—18.

Di Muccio A, Pelosi P, Barbini DA, Generali T, Girolimetti S, Stefanelli P, Leonelli A, Amendola G, Vergori L, Fresquet EV (1999) J. Chromatogr. A 833(1): 19—34.

Gionfriddo E, Gruszecka D, Li X, Pawliszyn J (2020) Talanta 211: 120746.

García MC, Torre M, Marina ML, Laborda F (1997) Crit. Rev. Food Sci. Nut. 37(4): 361—391.

Hernández-Borges J, Rodriguez-Delgado MA, García-Montelongo FJ, Cifuentes A (2005) J Sep Sci 28(9—10): 948—956.

Herrera López S (2019) J. Chromatogr. A 1594: 93—104.

Hong J, Fo Y, Rhee J, Kim T, Kim K (1993) J. Chromatogr. A 639(2): 261—271. doi: 10.1016/0021-9673(93)80262-7.

Huertas Pérez JF, Olsen SB, Alba ARF, Schimmel H, Dabrio M (2015) Talanta 137: 120—129.

Huertas-Pérez JF, Olsen BS, Gokcen T, Sandor F, Schimmel H, Dabrio M (2019) Food Add. Cont. 36(1): 96—108.

Li L, Xu Y, Pan C, Zhou Z, Jianc S, Liu F (2007) J. AOAC Int. 90 (5): 1387—1394.

Maldaner (2008) J. Liq. Chromatogr. R. T. 31(7): 972—983.

Marchis D, Ferro GL, Brizio P, Squadrone S, Abete MC (2012) Food Control 25 (1): 270—273.

Martínez-Dominguez G, Romero-González R, Arrebola FJ, Frenich AG (2016) Food Control 59: 218—224.

May MM, Ferronato G, Bandeira N, Prestes O, Zanella R, Adaime MB (2017) Food Anal. Methods 10(2): 369—378.

Medic J, Atkinson C, Hurburgh CR (2014) J. Am. Oil Chem. Soc. 91: 363—384.

Olías R, Delgado-Andrade C, Padial M, Marín-Manzano MC, Clemente A (2023) Foods 12: 2665.
Páleníková A, Martínez-Domínguez G, Arrebola FJ, Romero-González R, Hrouzková S, Frenich AG (2015) Food Chem. 173: 796—807.
Petrarca MH, Cunha SC, Fernandes JO (2024) J. Chromatogr. A 1727: 464999.
Pizzutti IR, Kok A, Zanella R, Adaime MB, Hiemstra M, Wickert C, Prestes OD (2007) J. Chromatogr. A 1142(2): 123—136.
Rejczak T (2016) J. Chromatogr. A 1473: 56—65.
Rigby D (2001) Calif. Manag. Rev. 43: 139—160.
Rodrigues NR, de Souza APF (2018) Food Add. Cont. 35: 724—731.
Rupe J, Luttrell RG (2008) Soybeans — Chemistry, Production, Processing, and Utilization, ISBN 978-1-893997-64-6, 93—116.
Schurtleff W, Aoyagi A (2009) History of soybeans and soyfoods in South America, ISBN 978-1928914-23-5.
Schmitz JF, Erhan SZ, Sharma BK, Johnson LA, Myers DJ (2008) Soybeans — Chemistry, Production, Processing, and Utilization, ISBN 978-1-893997-64-6, 539—612.
Shin Y, Lee J, Kim JH (2018) App. Bio. Chem. 61: 543—548
Wang J, Wang C (2006) J. AOAC Int. 89: 214—224.
Wang C., Zhang Y, Wang D Shi P, Yang B (2021) Chin. J. Pest. Sci. 23 (2) 405—413.

sciendo

Reducing the impact of artificial blue light on the skin: A spectroscopic study

Silvia Martiniaková[1], Jarmila Hojerová[1], Zuzana Turányiová[1], Dana Dvoranová[2], Miriama Malček Šimunková[2]

[1]Institute of Food Science and Nutrition and
[2]Institute of Physical Chemistry and Chemical Physics, Faculty of Chemical and Food Technology, Slovak University of Technology in Bratislava, Radlinského 9, SK-812 37 Bratislava, Slovak Republic
silvia.martiniakova@stuba.sk, jarmila.hojerova@stuba.sk

Abstract: As people spend many hours looking at digital screens, the negative effects of artificial blue light are becoming more apparent. While most research has focused on its effects on eyes, less is known about the effects of blue light on the skin, where similar photoreceptors are located. Unlike the sunscreens against UVB and UVA radiation, there is no standard method for determining skin protection against blue light. The lipophilic complex Carotolino, a system combining carrot root extract, carrot seed oil, and β-carotene, was chosen as a model substance for this research. Spectrophotometric investigation demonstrated the ability of Carotolino to absorb radiation in the blue light region (400—500 nm). After a 60-minute LED@450 nm exposure, corresponding to the maximum wavelength of radiation from the displays of common smartphones, only small changes (1.4 %) in the optical spectra were observed. The spectra showed sufficient photostability of Carotolino and its stabilizing effect on the photolabile Ubiquinone. In the 415—455 nm wavelength range associated with oxidative stress, Carotolino (0.4 % wt.) reduced blue light by ~97.0 %. EPR spin trapping showed that blue light with a maximum wavelength of 450 nm causes significant formation of reactive free radicals, which can be partially eliminated by the application of Carotolino. The results confirmed the suitability of both methods to evaluate the effectiveness of substances to reduce physical impact of incident light on the skin. Further methods are needed to investigate biological protection of skin against blue light by promising substances.

Key words: blue light; carotenoids; EPR spin trapping; free radicals; light emitting diodes; skin care

Abbreviations: BL: blue light; DMPO: 5,5-dimethyl-1-pyrroline N-oxide; EPR: electron paramagnetic resonance; INCI: International Nomenclature of Cosmetic Ingredients; LED: light-emitting diode; ROS: reactive oxygen species; UV: ultraviolet; VIS: visible

Introduction

Since the creation of human life on Earth, humans have been exposed to varying levels of visible and invisible radiation from the sun. In general, blue light (BL) represents about 25 % of visible light (~380—780 nm), and its actual intensity depends on location, time of day, season, and various weather conditions and atmospheric pollution. Blue light is known as high-energy visible (HEV) light in the wavelength range of ~400—500 nm with maximum emission from 440 to 460 nm (Coats et al., 2021; Pourang et al., 2022). Over millions of years, specialized cells in the human eye have evolved to respond to natural BL with wavelengths between 470 and 490 nm, thus keeping track of time and regulating the body's biological functions. Adequate BL exposure is important for synchronizing the natural circadian rhythm, which can affect many processes including sleep quality, metabolism, immune function, and even mood. Sufficient BL, especially in the morning, induces wakefulness, while its deficiency in the evening triggers the production of the sleep hormone melatonin (Heiting, 2019; Françon, 2024).

Recently, due to modern life consequences, humans are not exposed to adequate levels of natural BL during the day but are overexposed to artificial BL outside daylight hours. Such light is emitted by the screens of digital devices such as smartphones, laptops, tablets, computers, and televisions illuminated by light emitting diodes (LEDs) as well as LED street, household, architectural, and security lighting. Unlike the continuous and nearly uniform visible sunlight, cold-white LEDs emit spectrally unbalanced light characterized by a high-energetic blue peak at 450 nm (~47.0 %), a low-energetic green-yellow peak, and a weak emission in the red region of the spectrum (Zhao et al., 2018; Heiting, 2019; Françon, 2024). Especially smartphones have become an integral part of our daily life. According to statistics, the average person spends 4 hours and 37 minutes on their smartphone every day (Howarth, 2024). Teran et al. (2018) evaluated the light emission spectrum of 36 smartphone displays from various producers using the integrating sphere method. When in the regular mode, the power spectra of the smartphones resembled a white LED source,

showing three characteristic peaks at about 450, 545, and 603 nm. For almost all of them, the blue peak (~450 nm) had twice as much power as the yellow (~545 nm) and red (~603 nm) ones. Ultraviolet (UV) and infrared regions were practically null compared to the three emission peaks in the visible range (Teran et al., 2018). Therefore, it is BL that is being investigated for its potentially harmful effects on humans.

Although the intensity of artificial BL is low compared to BL from the sun, its cumulative effect on the body represents serious health risk for people who are in front of LED devices for too long during the day and sometimes even at night, and the screen is too close to their eyes and skin. They are exposed to a relatively large intensity of blue wavelengths outside of normal daytime hours, so the evening and night-time production of the hormone melatonin in their bodies is suppressed. The consequence can be not only disruption of the circadian rhythm and sleep disorders, but also other problems, for example metabolic dysregulation (Zerbini et al., 2020; Haghani et al., 2024). In addition, particularly short-wave BL with wavelengths between 415 nm and 455 nm is closely related to oxidative and phototoxic damage to photoreceptors in human eyes, especially the retina (Marie et al., 2018; Zhao et al., 2018; Wang et al., 2024; Haghani et al., 2024).

While most of the research to date has focused on the effects of BL on eyes, less is known about its impact on skin, where similar photoreceptors are found. BL has been shown to cause a wide range of beneficial as well as harmful effects depending on the exposure dose, applied wavelength, and skin condition. Therapeutic effects of blue LED light with specific wavelengths are used in the treatment of dermatological diseases such as psoriasis, eczema, acne vulgaris, actinic keratosis, and skin malignancies, as well as in cosmetic treatments. Controlled short-term exposure to blue LED light is used in phototherapy in hyperbilirubinemic newborns (Sadowska et al., 2021; Olagbenro, 2022; Lebleu et al., 2023). LED emission doses from these sources are significantly lower than the minimal doses reported to induce skin damage within practical exposure times (Brown et al., 2024).

However, the results of *in vitro, in vivo,* and clinical studies show that repeated long-term exposure to BL causes direct and indirect harmful effects to the skin. Blue wavelengths, especially in the wavelength range of 415—455 nm, exert their direct oxidizing effect on the skin by interacting with chromophores (flavins, porphyrins, nitrosated proteins, and opsins) in which they trigger a change from ground state to activated state. Activation of these molecules leads to the release of reactive forms of nitrogen, i.e., nitric oxide (NO), and to the overproduction of reactive oxygen species (ROS), e.g., superoxide radical anion, while UVA light induces predominant formation of singlet oxygen (Sadowska et al., 2021; Suitthimeathegorn et al., 2022; Kumari et al., 2023). In an *in vitro study* on human dermal fibroblasts, Mann et al. (2020) found that the rate of ROS generation was highest between the wavelengths of 400—450 nm and then decreased continuously, with only minimal effects observed at wavelengths above 500 nm. Similarly, Mamalis et al. (2016) exposed dermal fibroblasts to various effects of blue LED light (415 nm; irradiance 5—80 $J \cdot cm^{-2}$) and demonstrated a dose-dependent increase in ROS. As BL penetrates the skin deeper than UV light (Suitthimeathegorn et al., 2022), the reactive free radicals can cause DNA damage of epidermal keratinocytes and dermal fibroblasts, reduction of normal cell proliferation and uneven melanogenesis leading to hyperpigmentation (Dong et al., 2019; Campiche et al., 2020; Pourang et al., 2022; Lebleu et al., 2023; Lago, 2024).

The skin, as the surface of the human body, is immediately exposed to radiation and chemical pollutants from external environment, including free radicals causing oxidative stress outside and inside the organism (Ácsová et al., 2019; Martiniaková et al., 2022). Therefore, consequences of free radicals from BL on the skin will be manifested not only by premature external photoaging (promotion of wrinkles), but also by accelerated internal aging of the organism with all the health consequences (Arjmandi et al., 2018). In addition, as excessive exposure to BL can disrupt the circadian rhythm, it can negatively affect skin renewal processes that take place during the night (Suitthimeathegorn et al., 2022).

From this point of view, measures should be taken to protect human skin and eyes during frequent and long-term exposure to blue LED light (Coats et al., 2021; Kumar et al., 2023). One of the possible ways to reduce the adverse light effects of blue LED light on the skin is the application of cosmetic products containing BL protective ingredients. However, unlike the evaluation of the effectiveness of sunscreen products against UVB and UVA radiation, there is currently no standardized method for determining skin protection against BL. Therefore, this research is oriented on verification of the spectrophotometric method for evaluating the ability of a model substance to reduce physical effect of BL in the 415—455 nm wavelength range, which is closely related to oxidative stress. The indirect technique of electron

paramagnetic resonance (EPR) spectroscopy, spin trapping, was applied to determine transient paramagnetic intermediates content formed upon BL exposure.

Materials and Methods

Chemicals and reagents

Based on our previous research (unpublished results), the lipophilic substance Carotolino (reddish-orange liquid) from Lipoid Kozmetik (Steinhausen, Switzerland) was selected for this research from several cosmetic substances currently presented as effective against the impact of blue light on the skin. This natural complex consists of *Canola Oil, Daucus Carota Sativa Seed Oil, Daucus Carota Sativa Root Extract, Helianthus Annuus Seed Oil, Tocopheryl Acetate, Beta-Carotene* (names according to the International Nomenclature of Cosmetic Ingredients, INCI). Ubiquinone (INCI: *Ubiquinone*), the oxidized form of coenzyme Q10 (orange powder), was obtained from Kaneka Corporation (Osaka, Japan) and was used as a model lipophilic excipient. Canola oil (INCI: *Canola Oil*) from Palma (Bratislava, Slovakia) was used as a lipophilic solvent and a reference sample. The 5,5-dimethyl-1-pyrroline N-oxide (DMPO) spin trapping agent was obtained from Sigma-Aldrich (Merck) and distilled prior to the application.

Spectrophotometric experiments

Electronic absorption spectra within the 400–550 nm wavelength range were recorded using a UV-1800 Shimadzu spectrophotometer (Kyoto, Japan) in a quartz cell (optical path length of 10 mm). Canola oil served as the reference. The reaction mixtures were irradiated by LED@450 nm in a Keva3 photochemical reactor (Brno, Czech Republic) (Figure 1) for defined time and the electronic absorption spectra were recorded. The wavelength was chosen considering the findings of Teran et al. (2018) regarding the power spectra emitted by common smartphone displays. To demonstrate the effect of Carotolino, a lipophilic substance that absorbs radiation in the blue light region but exhibits photolability was sought. The Ubiquinone substance (oxidized form of coenzyme Q10) in Canola oil (5 % wt.) was chosen. The investigated mixture (Carotolino (0.4 % wt.) in Canola oil or Carotolino (0.4 % wt.) plus Ubiquinone (5 % wt.) in Canola oil) were prepared and optical spectra before and after a 60-minute LED@450 nm exposure were recorded.

EPR spin trapping experiments

EPR experiments were carried out at room temperature using an EPR spectrometer EMX*Plus* (Bruker, Mannheim, Germany) operating at 100 kHz field modulation using a high sensitivity

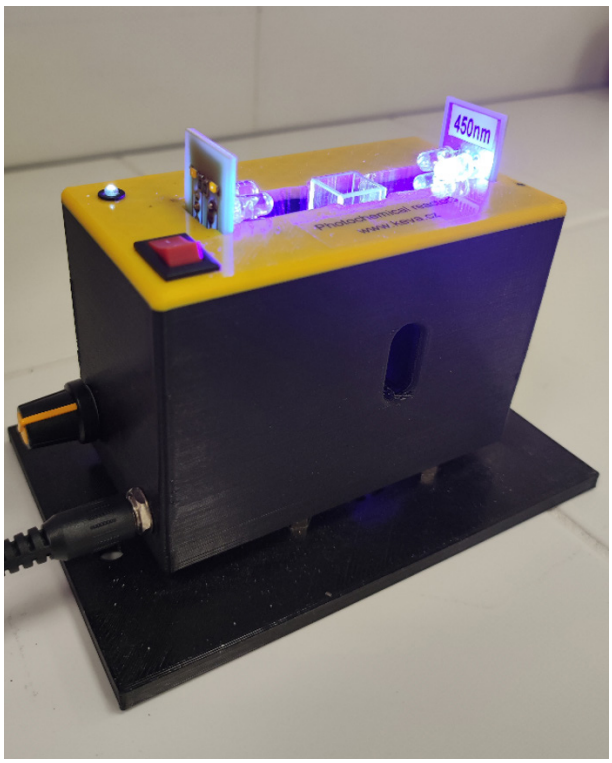


Fig. 1. Reaction system in Keva 3 photochemical reactor upon LED@450 nm exposure.

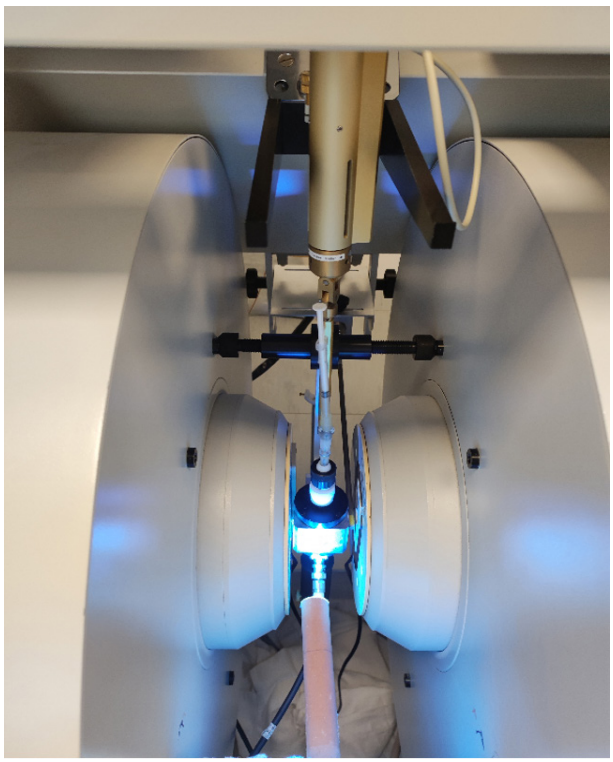


Fig. 2. Reaction system during blue light exposure in the cavity of EPR spectrometer.

probe-head in a small quartz flat cell (optical cell length of 0.45 mm, WG 808-Q, Wilmad-LabGlass, Vineland, USA). Cold-light source KL 1600 LED (T = 5600 K; Schott, Germany) was applied to provide the VIS-light irradiation directly in the EPR cavity, characterized by the illumination intensity of 160 klx measured by a digital lux meter (Metra Blansko, Czech Republic), and a blue filter was set to select blue light (λ_{max} = 450 nm). The 5,5-dimethyl-1-pyrriline N-oxide (DMPO) spin trapping agent was used to determine the content of transient paramagnetic species.

The reference system was prepared directly before the measurement by mixing 200 µL of Canola oil (reference sample) with 50 µL of DMPO (210 mM). The system was immediately transferred to a flat cell under air and inserted into the cavity of the EPR spectrometer (Figure 2). Acquisition of EPR spectra started 2 minutes after mixing the reaction system. The system was first analyzed before irradiation (10 minutes in the dark). Subsequently, it was irradiated with BL for 10 minutes, while EPR spectra were measured *in situ* during exposure (Figure 2). The reaction system prepared directly before the measurement consisted of 50 µL of a solution of Carotolino in Canola oil (1:1 vol.), 150 µL of Canola oil and 50 µL of DMPO. Thus, Carotolino (10 % wt.) solution in Canola oil was analyzed before and after the irradiation using the same procedure as for the reference sample. The experimental EPR spectra were analyzed using the WinEPR software (Bruker).

Results and Discussion

Figure 3 shows electronic absorption spectra of studied systems before and after LED@405 nm exposure. The absorption spectrum of Carotolino in Canola oil (red solid line) with absorption maxima in the blue region at 449 and 475 nm is characteristic for carotenoids with a π-bond conjugated system (Domenici at al., 2014). The absorbance value of Carotolino (0.4 % wt.) in Canola oil was $A_{450\ nm/0\ h}$ = 1.89 (Figure 3, solid red line) before exposure, and after a 60-minute irradiation, minor changes in optical spectra were observed with only negligible decrease (Figure 3, dashed red line), i.e., decrease of only 2.1 % was measured. The integrated area under the absorption curve over the entire BL wavelength range (400—500 nm) before and after irradiation was calculated. Decrease in the absorption capacity of the Carotolino solution was only 1.4 %, which means that Carotolino in Canola oil has sufficient photostability.

Due to known photodegradation of both forms of coenzyme Q10 (Ubiquinol and Ubiquinone) under visible light (Hojerová, 2012) and the absorption of radiation in the BL region, Ubiquinone was chosen as a model photolabile substance, as documented in Figure 3 (green lines). The absorbance of Ubiquinone (5 % wt.) in Canola oil dropped from $A_{450\ nm/0\ h}$ = 0.523 before irradiation (Figure 3, solid green line) to $A_{450\ nm/1\ h}$ = 0.201 after 60 minutes of irradiation (Figure 3, dashed green line), which represents a decrease of up to 61.7 %. The integrated area under the absorption curve in the 400—500 nm wavelength range before and after irradiation decreased by 57.0 %. The Ubiquinone solution can therefore be considered photounstable in the BL region and suitable for verifying the possible stabilizing effect of the Carotolino substance in this work.

After the addition of Carotolino (0.4 % wt.) to the solution of Ubiquinone (5 % wt.) in Canola oil, the absorbance values increased since the components absorb in the same region, but the shape of absorption curves before and after irradiation (Figure 3, solid and dashed black line) changed only slightly. In the Japanese patent JP 5344768B2 (JP, 2010), β-carotene in combination with ascorbic acid is used as a photodegradation inhibitor of coenzyme Q10. We assume that the substance Carotolino, composed of predominantly carotenoids, especially β-carotene, forms an active complex with Ubiquinone, which is resistant to blue light. However, these considerations require confirmation by analysis of the composition of the resulting complex. It can be concluded that Carotolino demonstrated not only stable absorption of BL before and after irradiation under the experimental conditions, but also photostabilizing effect on the photolabile substance Ubiquinone.

Furthermore, for a more comprehensible expression of the degree of BL protection by the evaluated substances, the BASF method (BASF, 2024) was applied. This method is currently used by BASF to evaluate the effectiveness of cosmetic UV filters against BL and is expressed as a percentage reduction in transmittance within the 400—450 nm wavelength range. However, the transmittance values in our work were calculated from the absorbance values in the wavelength range of 415—455 nm, which was reported to induce a considerable amount of reactive free radicals and oxidative stress (Marie et al., 2018; Zhao, 2018; Wang et al., 2023; Heiting, 2019; Pourang et al., 2022). The transmission dependencies are shown in Figure 4. The larger the area above the transmission curve, the more BL the substance blocks, i.e. the less BL is transmitted. Reciprocally, the larger the area under the transmission curve, the more BL is transmitted and the less is blocked. Finally, the percentage of BL reduction in the 415—455 nm wavelength range

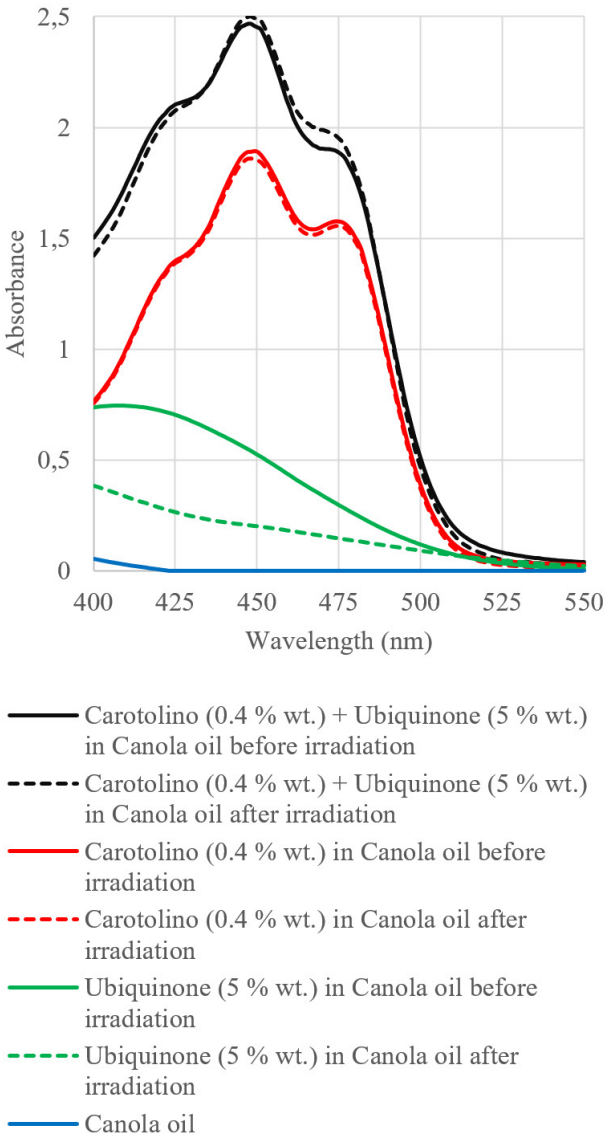


Fig. 3. Absorption spectra of studied substances in the region of 400–550 nm before and after a 60-minute exposure (LED@450 nm) in Keva 3 photochemical reactor.

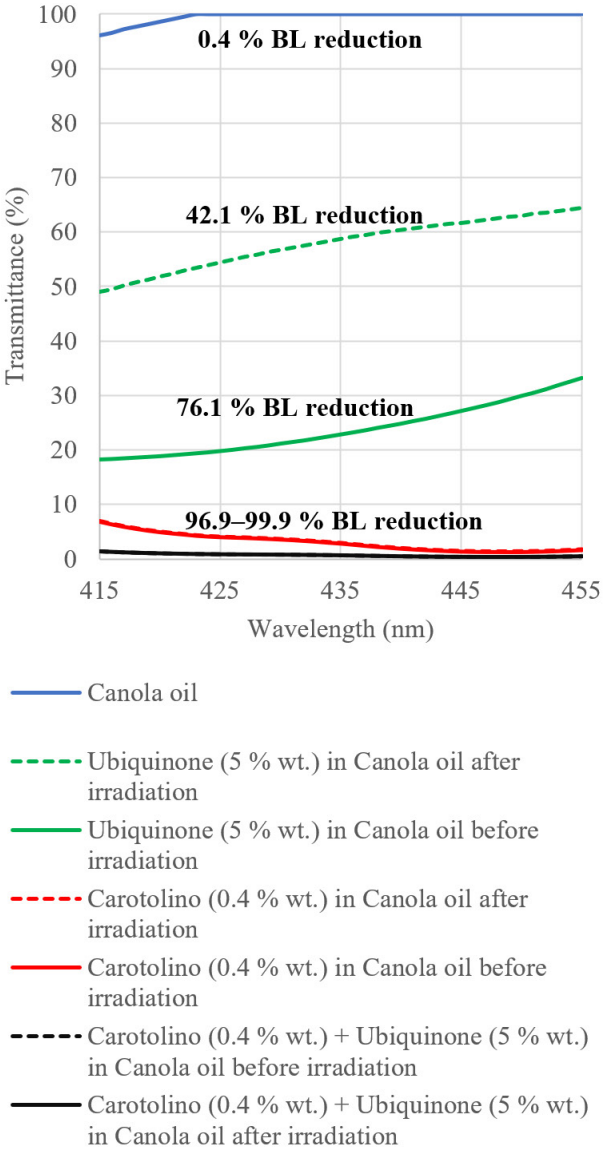


Fig. 4. Calculated transmission spectra of studied substances in the region of 415–455 nm before and after a 60-minute exposure (LED@450 nm) in Keva 3 photochemical reactor.

Tab. 1. Blue light reduction by the evaluated substances in the 415–455 nm wavelength range.

Substance	**Blue light reduction (%) range 415–455 nm**
Canola oil before irradiation	0.4
Ubiquinone (5 % wt.) in Canola oil before irradiation	76.1
Ubiquinone (5 % wt.) in Canola oil after irradiation	42.1
Carotolino (0.4 % wt.) in Canola oil before irradiation	97.0
Carotolino (0.4 % wt.) in Canola oil after irradiation	96.9
Carotolino (0.4 % wt.) + Ubiquinone (5 % wt.) in Canola oil before irradiation	98.6
Carotolino (0.4 % wt.) + Ubiquinone (5 % wt.) in Canola oil after irradiation	99.3

was calculated from the integrated area above the transmission curve using equation (1):

$$\text{Blue light reduction (\%)} = 100 \times \sum_{415\text{ nm}}^{455\text{ nm}} (1 - T) / 41 \quad (1),$$

where 41 is the number of wavelength steps in the 415—455 nm range with a step of 1 nm and T means transmittance.

As shown in Table 1, Carotolino (0.4 % wt.) in Canola oil reduces the potential exposure of skin to BL in the wavelength range of 415—455 nm (associated with oxidative stress) up to 97.0 %. Moreover, after a LED@450 nm exposure for 60 minutes, the Carotolino solution is still able to block up to 96.9 % of blue light at 415—455 nm.

The formation of reactive free radicals is one of the most harmful side effects of exposure to blue light. Free radicals generated by light exposure often possess a short lifetime and therefore cannot be detected by cw-EPR spectrometers, which requires the use of indirect techniques, such spin trapping, which utilize diamagnetic compounds to trap free radicals while forming a more stable spin-adduct (Dvoranová et al., 2014; Jomova et al., 2022). The amount of spin-adducts reflects the amount of trapped free radicals, which can be obtained by quantitative analysis of the EPR spectra. The determination of overall radical content in the studied system was performed in two parallels and the average of the double integrals values of EPR signal intensities was calculated. Double integral of the EPR signal intensity in non-irradiated Canola oil showed only insignificant amount of spin-adducts present in the reference system (Figure 5), thus indicating negligible amount of reactive free radicals. However, following a 10-minutes blue LED light exposure (λ_{max} = 450 nm) of Canola oil, the total amount of generated reactive free radicals, detected as spin-adducts, notably increased; expressed by the double integral of the EPR signal intensity of 8.03 × 10^9 (Figure 5, blue column). This observation indicates that the active substance Carotolino is able to eliminate some of the reactive free radicals generated upon blue LED light exposure due to its radical scavenging components, such as carotenoids. Another possible explanation is that the lower amount of spin-adducts in sample with Carotolino is due to its ability to absorb BL, which leads to less energy being absorbed by the components of Canola oil which are prone to form radicals.

Conclusions

As was mentioned, blue light occurs naturally as part of sunlight. Like all living things on Earth, humans have evolved to respond to the daily cycle of light and dark. In recent decades, because of modern lifestyle, people are not exposed to adequate levels of natural blue light during the day but are overexposed to relatively high levels of blue LED light at night. There is growing concern that the increased exposure to artificial sources of blue light from lighting and digital screens influences our health.

Although the full extent and exact nature of the biological effects of blue light on the skin are not yet fully understood, certain beneficial as well as harmful effects have already been confirmed. As the skin's

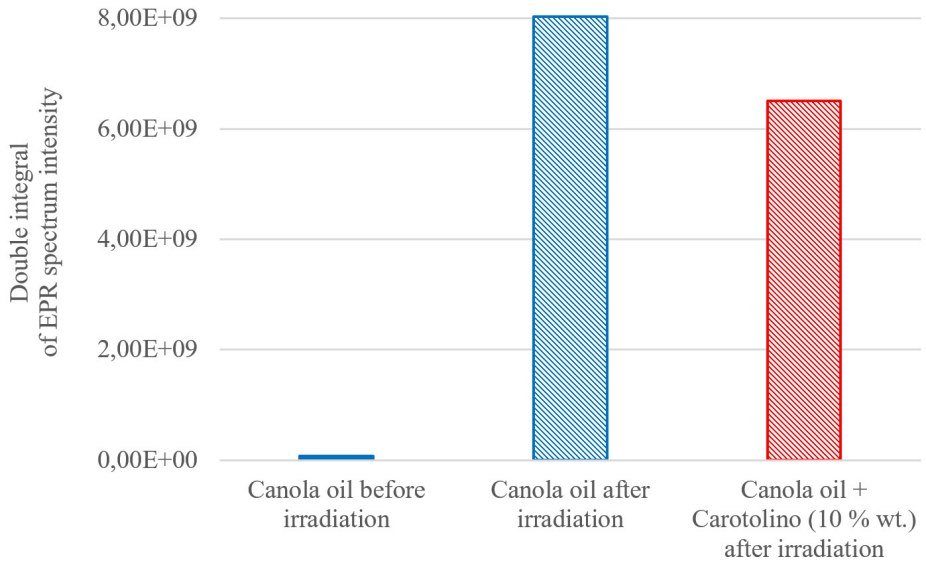


Fig. 5. Double integrated EPR spectra of studied systems before and after 10-minute blue LED light (λ_{max} = 450 nm) exposure in the cavity of EPR spectrometer.

exposure to blue light from artificial sources gradually increases, research into its protection is needed. When considering cosmetic interventions to protect the skin from blue light exposure, a combination of daytime protection and nighttime active repair strategy may be optimal. Ideal daily protection regimen includes reduced exposure to blue light by all available means and prevention or minimization of the formation of free radicals with skin care containing effective anti-blue light ingredients.

One such substance appears to be Carotolino, a stabilized complex of carotenoids and other components in Canola oil. In this research, the ability of Carotolino to *i)* absorb light in the wavelength range of 400—450 nm corresponding to blue light, *ii)* reduce the transmission of light at wavelengths from 415 to 455 nm, which generate reactive free radicals, and *iii)* maintain photostability in the entire evaluated range of wavelengths under irradiation conditions, was proven by absorption spectrophotometry.

Furthermore, an EPR spin trapping experiment showed that blue light with a maximum wavelength of 450 nm, which corresponds to the maximum wavelength of radiation from the displays of common smartphones, causes considerable formation of reactive free radicals which can be partially eliminated by applying the model cosmetic ingredient Carotolino.

Considering the results obtained, both methods seem to be suitable for the evaluation of the physical effectiveness of substances promising skin protection against blue light. It is highly possible that by appropriate *in vitro* or *in vivo* biological assays, for example by dermal fibroblast or epidermal keratinocyte cells viability evaluation, Carotolino substance will also demonstrate direct biological effectiveness in reducing skin damage caused by blue light. In our opinion, the future of topical products providing skin protection against free radical damage from digital screens is a combination of photostable blue light-reducing substances and photostable UVA filters effective even above the wavelength of 360 nm.

Acknowledgement

This publication was supported by the Operational Program Integrated Infrastructure within the project: Demand-driven Research for the Sustainable and Innovative Food, Drive4SIFood 313011V336, cofinanced by the European Regional Development Fund.

References

Ácsová A, Martiniaková S, Hojerová J (2019) Acta Chimica Slovaca 12 (2): 200—211.

Arjmandi N, Mortazavi G, Zarei S, Faraz M, Mortazavi SAR (2018) Journal of Biomedical Physics Engineering 8 (4): 447—452.

BASF (2024) The BASF Personal Care Sunscreen Simulator 4.0. Available on: https://sunscreensimulator.basf.com.

Brown A, Trullas C, Jourdan E (2024) Journal of Photochemistry and Photobiology 19: 100216.

Campiche R, Curpen SJ, Lutchmanen-Kolanthan V, Gougeon S, Cherel M, Laurent G, Gempeler M, Schuetz R (2020) International Journal of Cosmetic Sci 42: 399—406.

Coats JG, Maktabi B, Abou-Dahech MS, Baki G (2021) Journal of Cosmetic Dermatology 20: 714—717.

Cotter EJ, Cotter LM, Riley CN, Dixon J, VanDerwerker N, Ufot AI, Godfrey J, Gold D, Hetzel SJ, Safdar N, Grogan BF (2024) JSES International 8 (2): 328—334.

Dong K, Goyarts EC, Pelle E, Trivero J, Pernodet N (2019) International Journal of Cosmetic Sci 41 (6): 558—562.

Domenici V, Ancora D, Cifelli M, Serani A, Veracini CA, Zandomeneghi M (2014) Journal of Agricultural and Food Chemistry 62 (38): 9317.

Dvoranová D, Barbieriková Z, Brezová V (2014) Molecules 19: 17279.

Françon A, Behar-Cohen F, Torriglia A (2024) Environment International 184: 108471.

Haghani M, Abbasi S, Abdoli L, Shams SF, Baha'addini Baigy Zarandi BF, Shokrpour N, Jahromizadeh A, Mortazavi SA, Mortazavi SMJ (2024) Journal of Biomedical Physics and Engineering 14 (3): 213—228.

Heiting G (2019) How to Save Your Eyes in the Digital Age. Book. Available on: https://eyesafe.com/handbook.

Hojerová J (2012) Skin Health Benefits of Coenzyme Q_{10}. Chapter in book. Watson RR: Bioactive Dietary Factors and Plant Extracts in Dermatology. New York: Springer, 197—213.

Howarth J (2024) Time Spent Using Smartphones (2024 Statistics). Exploding Topics June 4. https://explodingtopics.com.

Jomova K, Hudecova L, Lauro P, Simunková M, Barbierikova Z, Malcek M, Alwasel SH, Alhazza IM, Rhodes CJ, Valko M (2022) Journal of Inorganic Biochemistry 266: 111635.

JP (2010) The Japanese Patent JP5344768B2. Coenzyme Q10 Photodegradation Inhibitor and Oral Composition. Available on: https://patents.google.com/patent/JP5344768B2/en.

Kumari J, Das K, Babaei M, Rokni GR, Goldust M (2023) Journal of Cosmetic Dermatology 22 (4): 1185—1190.

Lago JC, Ganzerla MD, Dias ALA, Savietto JP (2024) JID Innovations 4: 100252.

Lebleu A, Le Mestr A, Plaza C, Serre C, Botto J, Capallere C, Imbert I (2023) Journal of Investigative Dermatology 143 (5): S197.

Mamalis A, Koo E, Jagdeo J (2016) Dermatological Surgery 42 (6): 727—732.

Mann T, Eggers K, Rippke F (2020) Photodermatology Photoimmunology and Photomedicine 36 (2): 135—144.

Marie M, Bigot K, Angebault C, Barrau C, Gondouin P, Pagan D (2018) Cell Death & Disease 9 (3): 287.

Martiniaková S, Ácsová A, Hojerová J, Krepsová Z, Kreps F (2022) Acta Chimica Slovaca 15 (1): 1—11.

Pourang A, Tisack A, Ezekwe N, Torres AE, Kohli I, Hamzavi IH, Lim HW (2022) Photodermatology Photoimmunology and Photomedicine 38 (3): 191—196.
Sadowska M, Narbutt J, Lesiak A (2021) Life 11 (7): 670.
Suitthimeathegorn O, Yang Ch, Ma Y, Liu W (2022) Skin Pharmacology and Physiology 35 (6): 305—318.
Teran E, De Gracia P, Yee-Rendon C-M, Ortega-Salazar J, Molina-Reyes J, García ER (2018) Investigative Ophthalmology & Visual Science 59 (9): 4040.
Wang L, Yu X, Zhang D, Wen Y, Zhang L, Xia Y, Chen J, Xie Ch, Zhu H, Tong J, Shen Y (2023) Journal of Photochemistry and Photobiology B 240 (3): 112654.
Yang Q, Xia Y, Chen K, Wang Y, Song D, Zhu J, Tong, J, Shen Y (2024) Journal of Photochemistry and Photobiology B 255 (6): 112908.
Zerbini G, Kantermann T, Merrow M (2020) European Journal of Neuroscience 51 (12): 2355—2366.
Zhang Ch-X, Fan B, Chi J, Li Y-L, Jiao Q, Zhang Z-Y, Li G-Y (2024) Experimental Eye Research 244 (7): 109946.
Zhao Z-Ch, Zhou Y, Tan G, Li J (2018) International Journal of Ophthalmology 11 (12): 1999—2003.

sciendo

Quantum-chemical studies of infrared spectra of ^{15}N labeled diazene isomers

Martin Breza, Anton Gatial

Faculty of Chemical and Food Technology STU, Radlinskeho 9, SK-81237 Bratislava, Slovakia
martin.breza@stuba.sk

Abstract: The geometry of *trans*-HN═NH, *cis*-HN═NH and N═NH_2 containing no, one or two ^{15}N labeled atoms was optimized. The corresponding infrared vibrations were evaluated using a linear scaling factor. For each of these compounds at least one vibration can be found, which enables to distinguish between heteroisotopic $^{14}N═^{15}N$ and homoisotopic $^{14}N═^{14}N$ or $^{15}N═^{15}N$ species. Independent of the ^{15}N labeling, only *trans*-conformation should be found in the reaction mixture under equilibrium conditions.

Keywords: Coupled Cluster, isotopic labeling, geometry optimization, vibration transitions

Introduction

The reaction mechanism for homogeneous oxidation of hydrazine N_2H_4 by various oxidizing agents in aqueous solutions is more complex than in case of its heterogeneous oxidation on solid electrode surfaces (Higginson and Sutton, 1953; Cahn and Powell, 1954; Petek and Bruckenstein, 1973). Homogenous oxidation of a mixture of non-labeled ($^{14}N_2H_4$) and ^{15}N labeled ($^{15}N_2H_4$) hydrazine molecules yields $^{14}N_2$, $^{15}N^{14}N$, and $^{15}N_2$ molecules, whereas heterogeneous oxidation produces $^{14}N_2$ and $^{15}N_2$ molecules only. The formation of $^{14}N^{15}N$ molecules cannot be explained by simple removal of hydrogens from N_2H_4 molecules, the intermediate formation of N_4H_6 or N_4H_4 dimers is supposed. Quantum-chemical studies of N_4H_6 isomers in aqueous solutions at the CCSD/cc-pVTZ level of theory (Breza and Manova, 2023) explained their decomposition prevailingly into $^{15}N^{14}N$ molecules through $^{14}NH_3$...$^{14}N^{15}N$...$^{15}NH_3$ intermediates. The HN═NH and N═NH_2 species can also be formed. Due to the low stability of these species, their $^{15}N/^{14}N$ isotopic composition must be experimentally determined using sufficiently fast methods, such as infrared (IR) spectroscopy.

Early IR investigations of N_2H_2 were done by matrix isolation methods (Van Thiel and Pimentel, 1960; Milligan and Jacox, 1964; Blau and Hochheimer, 1964; Rosengreen and Pimentel, 1965; Bondybey and Nibler, 1973; Nibler and Bondybey, 1974) but the spectra were very complex due to impurities and extensive hydrogen bonding. Rather pure diazene (diimide, diamine) and significantly simplified IR spectra were obtained later (Minkwitz, 1975; Demaison et al., 1997; Craig et al., 2015). However, the correct band assignments were based on a theoretical study of Craig and Levin (Craig and Levin, 1979) (see Tab. 1). Based on high resolution IR spectroscopy of gas phase diazene, rotational constants were obtained and the structure of *trans*-diazene was confirmed (Trombetti, 1968; Carlotti et al., 1974; Hegelund et al., 1994, 1996, 1997) (see Tab. 2).

Isodiazene N═NH_2 (isodiimide, isodiimine) was obtained by photolysis of carbamoyl azide [$H_2N(CO)N_3$] and its IR spectra were also measured (Sylwester and Dervan, 1984; Teles et al., 1989) (see Tab. 1).

Quantum-chemical studies at CCSD(T) and similar levels of theory explained the energy relations among the three N_2H_2 isomers (*trans*-N_2H_2 being the most stable) and resolved the last problems of the assignment of diazene bands (Kobayashi et al., 1993; Martin and Taylor, 1999; Sekušak and Frenking, 2001; Matus et al., 2006; Spada et al., 2013) (see Tab. 2 for vibration transitions and Tab. 3 for optimal geometries and energies).

Only one old study dealing with the effect of ^{15}N labeled diazene in its experimental IR spectra was found (Rosengren and Pimentel, 1965). Triple splitting of the 1286 cm^{-1} band was observed with 1.3 and 2.6 cm^{-1} shifts for $H^{14}N═^{15}NH$ and $H^{15}N═^{15}NH$, respectively. Even larger shifts (up to 7 cm^{-1}) were predicted for another band (Craig and Levin, 1979).

The isotopic ^{15}N labeling of terminal nitrogen in $^{15}N^{14}NH_2$ lowers the 1574 cm^{-1} band to 1548 cm^{-1} (Sylwester and Dervan, 1984). Analogously, in deuterated isodiazene N═ND_2, the 1571 and 1552 cm^{-1} bands are assigned to $^{14}N═^{14}N$ and $^{14}N═^{15}N$ stretching, respectively.

The aim of our recent study was to investigate the effect of single and double ^{15}N labeling on the IR spectra of the three N_2H_2 isomers. This should enable confirmation of the above-mentioned N_2H_2 reaction intermediates (Breza and Manova, 2023) of homogenous hydrazine oxidation through dimerization because their instability can be problematic.

Tab. 1. Observed IR fundamentals (in cm^{-1}) for N_2H_2 and NNH_2 by various authors in matrices and in gas phase.

N_2H_2	N_2H_2	N_2H_2	N_2H_2	N_2H_2	N_2H_2	N_2H_2
Nibler and Bondybey, 1974	Minkwitz, 1975	Minkwitz, 1975	Demaison et al., 1997	Demaison et al., 1997	Craig et al., 2015	Craig et al., 2015
N_2 matrix IR, Raman	N_2 matrix IR	Ar matrix IR	N_2 matrix IR [b)]	Ar matrix IR [b)]	N_2 matrix IR	Ar matrix IR
1286	(1288) [a)]	(1283) [a)]	972 w	976 w	–	–
–	(1322) [a)]	(1313) [a)]	1288 vs	1283 vs	1284	1284
1529	–	–	1322 vs	1322 vs	1327	1312
1583	–	–	–	–	–	–
3128	–	3118	–	–	3131	3124
3131	3137	–	3137 w	3137 w	–	–

N_2H_2	N_2H_2	N_2H_2	NNH_2	NNH_2
Trombetti, 1971	Hallin et al., 1981	Hegelund et al., 1994	Sylwester and Dervan, 1984	Teles et al., 1989
Gas IR	Gas IR	Gas IR	Ar matrix IR	Ar matrix IR [c)]
–	–	–	1003	1002.7(1.00)
1359	1288.6386(16)	1288.64786(4)	1574	1287.5(0.008)
1404	1316.4129(15)	1316.41214(4)	1863	1574.2(0.31)
–	–	–	2141	1644.7(0.021)
3095	3120.2788(8)	3120.28676(8)	2808	2804.6(0.33)
–	–	–	2865	2862.0(0.75)

[a)]Values in parentheses are reassignments (Craig and Levin, 1979).

[b)]w – weak, vs-very strong.

[c)]relative intensities in parentheses.

Tab. 2. Calculated vibration transitions (in cm^{-1}) for N_2H_2 isomers by various authors at various levels of theory in gas phase.

Trans-HN=NH	*Trans*-HN=NH	*Cis*-HN=NH	N=NH_2	*Trans*-HN=NH	*Trans*-HN=NH
Martin and Taylor, 1997	Martin and Taylor, 1999	Martin and Taylor, 1999	Martin and Taylor, 1999	Spada et al., 2013	Spada et al., 2013
CCSD(T)/ cc-pVTZ	CCSD(T)/ cc-pVQZ	CCSD(T)/ cc-pVQZ	CCSD(T)/ cc-pVQZ	CCSD(T)/ aug-cc-pVTZ	CASSCF/ aug-cc-pVTZ
1294.2	1294.2	1231.9	991.0	1317	1321
1317.5	1317.4	1334.7	1292.7	1344	1354
1519.3	1528.2	1520.7	1560.4	1552	1527
1519.4	1578.5	1548.4	1665.3	1611	1617
3033.3	3051.0	2986.5	2769.7	3265	3124
3125.0	3133.3	3002.5	2866.5	3297	3158

Cis-diazene has not been detected experimentally. Isodiazene was detected shortly after formation (Sylwester and Dervan, 1984; Teles et al., 1989) probably due to its lower conversion rate.

Method

Using Gaussian16 software (Frisch et al., 2016), geometry optimization of *trans*-HN=NH, *cis*-HN=NH, and N=NH_2 with no, one or two ^{15}N labeled atoms was performed at the CCSD (Coupled Cluster using Single and Double substitutions from the Hartree-Fock determinant) level of theory (Scuseria et al., 1988) with cc-pVTZ basis sets for all atoms (Dunning, 1989). The optimized structures were tested for the absence of imaginary vibrations by vibrational analysis. Wavenumbers of vibrational transitions were multiplied by a scaling factor of 0.941 obtained from

Tab. 3. Calculated N—N, d(N—N), and N—H, d(N—H), bond lengths, N—N—H bond angles, α(N—N—H), relative energy differences, ΔE, for N_2H_2 isomers obtained at various levels of theory by various authors in gas phase.

Isomer	d(N—N) [Å]	d(N—H) [Å]	α(N—N—H) [deg]	ΔE [kJ/mol]	Method	Reference
Trans-HN=NH	1.2468	1.0281	106.17	–	CCSD(T)/cc-pVQZ	Kobayashi et al., 1993
Trans-HN=NH	1.2468	1.0281	106.17	0.00	CCSD(T)/cc-pVQZ	Martin and Taylor, 1999
Cis-HN=NH	1.2456	1.0331	111.88	21.78	CCSD(T)/cc-pVQZ	Martin and Taylor, 1999
N=NH_2	1.2172	1.0342	123.49	100.82	CCSD(T)/cc-pVQZ	Martin and Taylor, 1999
Trans-HN=NH	1.265	1.040	105.7	–	CCSD(T)/aug-cc-pVDZ	Sekušak and Frenking, 2001
Trans-HN=NH	1.2659	1.0415	105.75	–	CCSD(T)/aVDZ	Matus et al., 2006
Cis-HN=NH	1.2652	1.0464	111.57	–	CCSD(T)/aVDZ	Matus et al., 2006
Trans-HN=NH	1.2528	1.0318	106.13	0.0	CCSD(T)/aVTZ	Matus et al., 2006
Cis-HN=NH	1.2526	1.0365	111.77	21.3[a)]	CCSD(T)/aVTZ	Matus et al., 2006
Trans-HN=NH	1.247(1)	1.030(1)	106.3(1)	–	Exp.	Demaison et al., 1997

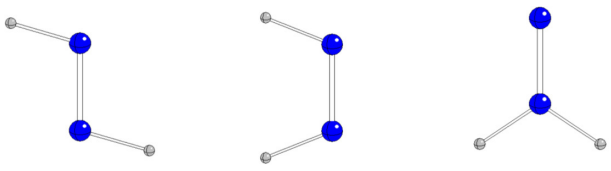
[a)]Enthalpy at 298 K

the National Institute of Standards and Technology (NIST, 2022) for CCSD/cc-pVTZ calculations.

Results and Discussion

Geometry optimization of *trans*-HN=NH, *cis*-HN=NH and N=NH_2 isomers gives stable structures of the C_{2h}, C_{2v}, and C_{2v} symmetry groups, respectively, with bond lengths and angles (see Fig. 1 and Tab. 4) very similar to previous CCSD(T)/cc-pVQZ studies (Kobayashi et al., 1993; Martin and Taylor, 1999). Single and double ^{15}N labeling causes only very small changes in structure parameters, which are lower than standard numerical accuracy. Gibbs energy data indicate the highest stability of *trans*-HN=NH, independent of ^{15}N labeling (Tab. 5), in agreement with analogous energy data from (Martin and Taylor, 1999; Matus et al., 2006), suggesting that only *trans*-HN=NH can be found in the reaction system under equilibrium conditions.

IR vibrational transitions calculated in previous studies (Martin and Taylor, 1997, 1999; Spada et al., 2013) do not include scaling factors. Their values are usually overestimated, although sometimes very exact. To improve the results, linear scaling factors for the CCSD method and cc-pVTZ basis sets (NIST, 2022) were applied to multiply the calculated wavenumbers of vibrational transitions.



Trans-diazene *Cis*-diazene Isodiazene

Fig. 1. CCSD/cc-pVTZ optimized structures of studied compounds (N — blue, H — gray).

Trans-diazene (C_{2h} symmetry group) has three IR active vibrations (one of a_u and two of b_u symmetries), the remaining three vibrations of a_g symmetry are active only in Raman spectroscopy (Tab. 6). Unlike other studies (Martin and Taylor, 1997, 1999; Spada et al., 2013), our data for the non-labeled compound are underestimated, and their deviations from the most exact experimental data (Hallin et al., 1981; Hegelund et al., 1994) are higher. The two most intense (the two lowest) transitions exhibit up to 2 cm^{-1} shifts of *trans*-H^{15}N=^{14}NH related to those of other isotopic composition. Its highest b_u vibration, despite having lower intensity, exhibits analogous shifts of –6 and +4 cm^{-1} related to homoisotopic compounds.

Cis-diazene (C_{2v} symmetry) has three a_1, one a_2, and two b_1 vibrations which can be IR active (Tab. 7). There are no experimental data for this isomer. The most intense b_1 vibration of *cis*-H^{15}N=^{14}NH exhibits –8 and +5 cm^{-1} shifts related to its homoisotopic counterparts. The other two intense vibrations also exhibit lower but measurable shifts.

Isodiazene (C_{2v} symmetry) has three a_1, two b_1, and one b_2 IR active vibrations (Tab. 8). Its three most intense transitions exhibit only negligible differences between their homoisotopic and heteroisotopic counterparts. Only the a_1 vibration of medium intensity is split four times. It contains two heteroisotopic bands separated by 6 cm^{-1} and two homoisotopic bands shifted from their center by +20 and –20 cm^{-1}.

Tab. 4. Bond lengths and angles of optimized geometries of compounds under study.

	Trans-HN═NH	*Cis*-HN═NH	N═NH_2
Bond lengths [Å]			
N—N	1.243	1.242	1.217
N—H	1.028	1.033	1.032
Bond angles [deg]			
N—N—H	106.1	111.9	123.6
H—N—H	–	–	112.9

Tab. 5. Absolute, G_{298}, and relative, ΔG_{298}, Gibbs energies at 298 K of studied systems (related to *trans*-isomer of the same isotopic composition).

	G_{298} [Hartree]	ΔG_{298} [kJ/mol]
Trans-H^{14}N═^{14}NH	–110.45183	0.00
Cis-H^{14}N═^{14}NH	–110.44397	20.64
^{14}N═$^{14}NH_2$	–110.41327	101.24
Trans-H^{15}N═^{15}NH	–110.45218	0.00
Cis-H^{15}N═^{15}NH	–110.44366	22.37
^{15}N═$^{15}NH_2$	–110.41361	101.27
Trans-H^{15}N═^{14}NH	–110.45201	0.00
Cis-H^{15}N═^{14}NH	–110.44349	22.37
^{15}N═$^{14}NH_2$	–110.41342	101.32
^{14}N═$^{15}NH_2$	–110.41346	101.21

Tab. 6. Symmetries, Γ (within the C_{2h} group), calculated wavenumbers, ν, and IR intensities, I, of vibration transitions of *trans*-diazenes of various ^{14}N/^{15}N isotopic composition. Most suitable transition for *trans*-H^{15}N═^{14}NH determination is highlighted in bold.

Γ	*Trans*-H^{14}N═^{14}NH		*Trans*-H^{15}N═^{15}NH		*Trans*-H^{15}N═^{14}NH	
	ν [cm^{-1}]	I [km/mol]	ν [cm^{-1}]	I [km/mol]	ν [cm^{-1}]	I [km/mol]
a_u	1273.1	94.5	1270.9	94.0	1272.3	94.3
b_u	1286.0	93.2	1283.8	92.7	1285.1	92.9
a_g	1526.4	0.0	1484.6	0.0	1508.3	0.0
a_g	1551.3	0.0	1531.9	0.0	1539.4	0.0
a_g	3118.4	0.0	3108.7	0.0	3111.6	0.3
b_u	3147.2	18.5	3137.1	18.6	**3141.1**	18.3

Tab. 7. Symmetries, Γ (within the C_{2v} group), calculated wavenumbers, ν, and IR intensities, I, of vibration transitions of *cis*-diazenes of various ^{14}N/^{15}N isotopic composition. Most suitable transition for *cis*-H^{15}N═^{14}NH determination is highlighted in bold.

Γ	*Cis*-H^{14}N═^{14}NH		*Cis*-H^{15}N═^{15}NH		*Cis*-H^{15}N═^{14}NH	
	ν [cm^{-1}]	I [km/mol]	ν [cm^{-1}]	I [km/mol]	ν [cm^{-1}]	I [km/mol]
a_2	1209.9	0.0	1203.0	0.0	1206.2	0.0
a_1	1312.2	0.6	1309.5	0.7	1310.5	0.7
b_1	1491.9	66.9	1478.9	68.0	**1484.4**	66.0
a_1	1549.3	2.9	1498.4	2.6	1524.5	4.2
b_1	3005.5	53.5	3000.8	53.2	3004.0	53.3
a_1	3087.5	30.6	3082.3	30.2	3085.9	30.4

Tab. 8. Symmetries, Γ (within C_{2v} group), calculated wavenumbers, ν, and IR intensities, I, of vibration transitions of isodiazenes of various $^{14}N/^{15}N$ isotopic composition. Most suitable transitions for $^{15}N{=}^{14}NH_2$ determination are highlighted in bold.

Γ	$^{14}N{=}^{14}NH_2$		$^{15}N{=}^{15}NH_2$		$^{15}N{=}^{14}NH_2$		$^{14}N{=}^{15}NH_2$	
	ν [cm^{-1}]	I [km/mol]	ν [cm^{-1}]	I [km/mol]	ν [cm^{-1}]	I [km/mol]	ν [cm^{-1}]	I [km/mol]
b_2	946.5	109.5	938.4	111.0	945.8	110.3	939.2	110.2
b_1	1264.0	2.7	1253.8	3.2	1260.8	2.9	1257.0	2.9
a_1	1527.5	13.3	1479.3	12.5	**1500.4**	12.8	**1506.9**	13.1
a_1	1647.5	0.0	1641.2	0.0	1647.4	0.0	1641.2	0.0
a_1	2995.2	64.5	2993.1	64.1	2995.0	64.7	2993.4	64.0
b_1	3009.1	56.4	2999.2	55.3	3009.1	56.4	2999.2	55.3

Finally, it can be concluded that heteroisotopic diazenes (*trans*-$H^{15}N{=}^{14}NH$ and *cis*-$H^{15}N{=}^{14}NH$) and isodiazenes ($^{15}N{=}^{14}NH_2$ and $^{14}N{=}^{15}NH_2$) can be detected by IR measurements. Another problem is finding suitable scaling factors for (iso)diazenes. It seems that linear scaling factors are insufficient for accurate calculation of IR active vibrations. More complex solutions to this problem should be proposed in future research.

Acknowledgement
The authors thank the HPC center at the Slovak University of Technology in Bratislava, which is a part of the Slovak Infrastructure of High-Performance Computing (SIVVP project ITMS 26230120002, funded by European Region Development Funds) for the computational time and resources made available.

References

Blau EJ, Hochheimer BF (1964) J. Chem. Phys. 41: 1174–1182.
Bondybey VE, Nibler JW (1973) J. Chem. Phys. 58: 2125–2134.
Breza M, Manova A (2023) Inorganics 11: 00413.
Cahn JW, Powell RE (1954) J. Am. Chem. Soc, 76: 2568–2572.
Carlotti M, Johns JWC, Trombetti A (1974) 52: 340–344.
Craig NC, Appiah KJ, Miller ChE, Seiden MV, Varley JE (2015) J. Mol. Spectr. 310: 3–7.
Craig NC, Kliewer MA, Shih NC (1979) J. Am. Chem. Soc. 101: 2480–2482.
Craig NC, Levin IW (1979) J. Chem. Phys. 71: 400–407.
Demaison J, Hegelund F, Burger H (1997) J. Mol. Struct. 413–414: 447–456.
Dunning Jr TH (1989) J. Chem. Phys. 90: 1007–1023.
Frisch GW, Trucks MJ, Schlegel B, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X et al. (2016) Gaussian 16, Revision B.01; Gaussian, Inc.: Wallingford, CT, USA.
Hallin K-EJ, Johns JWC, Trombetti A (1981) Can. J. Phys. 59: 663–672.
Hegelund F, Burger H, Polanz O (1994) J. Mol. Spectr. 167: 1–10.
Hegelund F, Burger H, Polanz O (1996) J. Mol. Spectr. 179: 142–149.
Hegelund F, Burger H, Polanz O (1997) J. Mol. Spectr. 181: 151–159.
Higginson WCE, Sutton D (1953) J. Chem. Soc. 1402–1406.
Kobayashi R, Bludsky O, Koch H, Jorgensen P (1993) Chem. Phys. Lett. 215: 576–581.
Martin JML, Taylor PR (1997) Spectrochim. Acta A53: 1039–1050.
Martin JML, Taylor PR (1999) Mol. Phys. 96(4): 681–692.
Matus MH, Arduengo III AJ, Dixon DA (2006) J. Phys. Chem. A 110: 10116–10121.
Milligan DE, Jacox ME (1964) J. Chem. Phys. 41: 2838–2841.
Minkwitz R (1975) Z. Anorg. Allgem. Chem. 411: 1–14.
Nibler JW, Bondybey VE (1974) J. Chem. Phys. 60: 1307–1312.
NIST Standard Reference Database. Release 22 (2022) Computational Chemistry Comparison and Benchmark DataBase. Available at https://cccbdb.nist.gov/vibscalejust.asp (accessed at July 15, 2024).
Petek M, Bruckenstein S (1973) Electroanal. Chem. Interrac. Electrochem. 1973, 47, 329–333.
Rosengren K, Pimentel GC (1965) J. Chem. Phys. 43: 507–516.
Scuseria GE, Janssen CL, Schaefer III HF (1988) J. Chem. Phys. 89: 7382–7387.
Sekušak S, Frenking G (2001) J. Mol. Struct. (Theochem) 541: 17–29.
Spada RFK, Ferrao LFA, Cardoso DVV, Roberto-Neto O, Machado FBC (2013) Chem. Phys. Lett. 557: 37–42.
Szopa K, Musial M, Kucharski SA (2008) Int. J. Quantum Chem. 108: 2108–2116.
Sylwester AP, Dervan PB (1984) J. Am. Chem. Soc. 106: 4648–4650.
Teles JH, Maier G (1989) Chem. Ber. 122: 293–323.
Teles JH, Maier G, Hess Jr. BA, Schaad LJ (1989) Chem. Ber. 122: 749–752.
Van Thiel M, Pimentel GC (1960) J. Chem. Phys. 32: 133–140.
Trombetti A (1968) Can. J. Phys. 46: 1005–1011.

sciendo

On the nature of copper binding to benzene

Jaroslav Granatier, Andrea Kováčová

Institute of Physical Chemistry and Chemical Physics, Slovak University of Technology in Bratislava, Radlinského 9, SK-812 37 Bratislava, Slovakia
jaroslav.granatier@stuba.sk

Abstract: Adsorption of copper atom on benzene surface has been studied at the *ab initio* MP2 and CCSD(T) theory levels. CCSD(T)/CBS binding energies of the copper atom adsorbed on hollow, top, and bridge positions are 16.77, 19.27 and 21.08 kJ/mol, respectively. Adsorption at the bridge position represents the most stable structure of the Cu-benzene complex with charge transfer from benzene to the copper atom playing a key role.

Keywords: Cu-benzene complex, correlation method, charge-transfer, benchmark CCSD(T)/CBS

Introduction

Weak van der Waals interaction between metal atoms and aromatic molecules is interesting from several points of view. One of the aspects is the absorption of aromatic molecules on metal surface. This is very important for understanding processes in surface science, heterogenous catalysis (Ago et al., 2012; Nieminen et al., 2008; Syomin et al., 2001), or molecular electronics (Cho et al., 2011; Cho, Choi and Kim, 2011; Heath and Ratner, 2003). On the other hand, interaction of graphene with metal atoms leads to formation of hybrid systems with application in sensing, catalysis, nanoelectronics, or hydrogen storage (Baby et al., 2010; Hong et al., 2010; Li et al., 2010a, 2010b; Scheuermann et al., 2009; Shan et al., 2010; Xiong et al., 2010, Singla et al., 2021). Hydrogen storage utilizes two different mechanisms (Malček and Bučinský, 2020): dissociation of molecular hydrogen by the host material and chemisorption of atomic hydrogen; and physisorption of molecular hydrogen on host material by weak van der Waals forces.

Theoretical calculations can provide essential information on the interaction between various metals and aromatic molecules or graphene surface. The description of this interaction is complicated due to the electronic structure of metals, which is affected by relativistic effects (Bučinský et al., 2011) and static and dynamic correlation requiring computational and time demanding methods to compute even simple systems.

Our previous study focused on interactions in Ag-benzene and Au-benzene complexes (Granatier et al., 2011). Charge transfer from benzene to the gold atom plays an important role in Au-benzene complexes. On the other hand, silver interacts with benzene by dispersion force. Thus, the silver atom preferred adsorption site is the center of the benzene ring, while the most stable structure of the Au-benzene complex is achieved by adsorption of the gold atom on the direct carbon atom or C—C bond.

In previous similar theoretical works (Bilić et al., 2006; Caputo et al., 2007; Chan, Neaton and Cohen, 2008; Giovannetti et al., 2008; Khomyakov et al., 2009; Manadé et. al., 2015), the interaction of paramagnetic copper with the carbon surface was calculated mostly at the DFT level. As copper is expected to be adsorbed on the carbon surface by a weak van der Waals bond, the DFT method may not provide correct description of this interaction. Therefore, high accuracy wave function theory (WFT) methods such as correlation MP2 and CCSD(T) were used in this study. The high computational demand of these methods allowed only simple systems to be calculated. As adsorption of the copper atom on a benzene molecule represents a very simple model for the description of copper atom adsorption on graphene surface, MP2 and CCSD(T) methods can be applied to obtain benchmark data and compare them with results calculated at the DFT theory level. The Cu-benzene complex was studied to provide better understanding of the nature of the bond between copper and the carbon surface.

Computational methods

Benchmark calculations of the Cu-benzene complex were performed using the spin-adapted CCSD(T) method with restricted closed-/open-shell reference function (Heckert et al., 2006; Neogrády and Urban, 1995; Urban, Neogrády and Hubač, 1997; Watts, Gauss and Bartlett, 1993) and the less demanding MP2 method (Møller and Plesset, 1934). The $3p^6 3d^{10} 4s^1$ shell of the copper atom and all electrons in benzene except for $1s^2$ electrons of the carbon atoms were explicitly correlated in correlation MP2 and CCSD(T) calculations.

Relativistic effects have been included using the scalar one-component Douglas-Kroll-Hess approximation (Douglas and Kroll, 1974; Hess and Chandra, 1987). All relativistic MP2 and CCSD(T) calculations were performed with two various types of relativistic all-electron basis sets. Both used basis sets provide high accuracy results for molecular properties such as ionization potential, electron affinity, electrical moments, or polarizabilities and are useful in calculations of binding energies of noncovalent interacting Cu-benzene complexes. The aug-cc-pVXZ-DK basis sets (Balabanov and Peterson, 2005, 2006; Dunning, 1989; Kendall, Dunning and Harrison, 1992) represent the first basis set type and are useful for benchmark calculations. The ANO-RCC basis set (Roos et al., 2004, 2005) was used to understand the interaction between the copper atom and the benzene molecule and for comparison with previous results (Granatier et al., 2011). The advantage of this basis set is that it is available for various degrees of contraction. CCSD(T) calculations were performed using ANO-RCC-VDZP and ANO-RCC-VTZP basis sets, and the MP2 binding energies were also computed using the ANO-RCC-VQZP basis set.

Benchmark binding energies were evaluated at the approximate CCSD(T)/CBS theory level, where the difference between CCSD(T) and MP2 binding energies calculated in the small basis set is added to the MP2/CBS biding energy (Jurečka and Hobza, 2002). MP2 binding energy in the complete basis sets was obtained using the Halkier's extrapolation scheme (Halkier et al., 1998) from MP2 calculations in finite basis sets. While the less demanding and well-parallelizable MP2 calculations are performed in large basis sets, the highly demanding CCSD(T) calculations are computed in small basis sets which allows for faster completion of the basis set for the difference between the correlation CCSD(T) and MP2 energies convergences compared with the MP2 and CCSD(T) correlation energies themselves. CCSD(T)/CBS binding energies and equilibrium distance between the copper atom and benzene surface were obtained employing the aug-cc-pVXZ-DK and ANO-RCC-VXZP basis sets. Although the ANO-RCC basis set does not meet the requirements for extrapolation to an infinite basis set, it provides results comparable to standard extrapolation techniques (Granatier, 2017). On the other hand, the aug-cc-pVXZ-DK basis sets are standardly used to evaluate energies in the complete basis set.

All calculated WFT binding energies were treated for the basis set superposition error (BSSE) using the Boys-Bernardi counterpoise correction (Boys and Bernardi, 1970). MP2 as well as CCSD(T) binding energies were calculated using the OpenMolcas program package (Fdez Galván et al., 2019).

DFT calculations were performed with an empirical D3 correction term describing the dispersion energy (Grimme et al., 2010). According to the suggestions of the authors, TPSS (Tao et al., 2003) and PBE (Perdew, Burke and Ernzerhof, 1996, 1997) functionals with the def2-QZVP basis set (Weigend and Ahlrichs, 2005) were used. DFT-D3 calculations were done using the ORCA program package (Neese, 2012).

The structure of benzene was optimized at the MP2/cc-pVTZ level and was taken from our previous papers with C—C and C—H bond lengths equal to 1.394 Å and 1.082 Å, respectively (Granatier et al., 2011, 2013; Lazar et al., 2014). The geometry of benzene was assumed to be frozen in all subsequent WFT calculations. The copper metal atom was modeled as being adsorbed at one of three different positions: (h) the 'hollow' site above the center of the aromatic ring, (t) a 'top' site directly above a C atom, and (b) a 'bridge' site above the midpoint of a C—C bond. Each bonding position leads to a different nature of the bond. Bond lengths of the studied complexes correspond to the distance between the metal atom and the plane of the benzene ring.

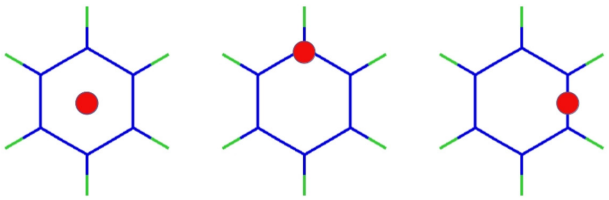


Fig. 1. Benzene molecule; three potential sites for metal atom adsorption: a) hollow, b) top, and c) bridge.

Results

All relativistic MP2 and CCSD(T) binding energies and the corresponding equilibrium distances calculated with ANO-RCC-VXZP (X = T, Q) and aug-cc-pVXZ-DK (X = D, T) basis sets are collected in Tab. 1. The relativistic BSSE corrected MP2/ANO-RCC-VTZP and CCSD(T)/ANO-RCC-VTZP potential energy curves for Cu-benzene with copper adsorbed at the hollow, top, and bridge positions are collected in Fig. 2.

Correct description of the Cu-benzene complex requires the use of correlation methods. The ROHF binding energy curves are repulsive for all three structures, but the inclusion of electron correlation leads to the appearance of an energy minimum. However, the correlation MP2 and CCSD(T) methods provide a different picture of the inter-

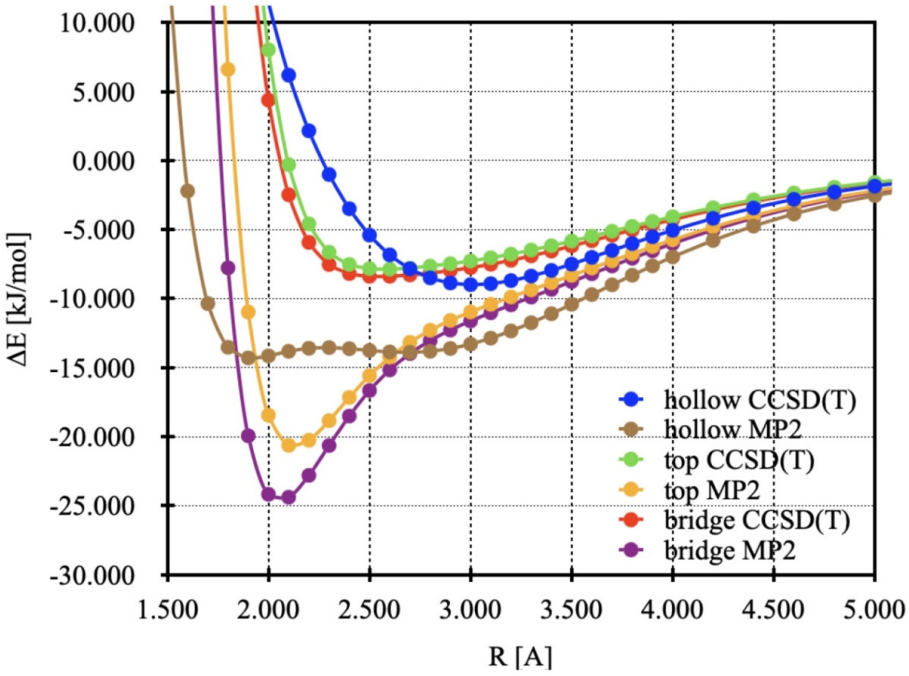


Fig. 2. DKH relativistic BSSE corrected CCSD(T)/ANO-RCC-VTZP and MP2/ANO-RCC-VQZP potential energy curves for Cu-benzene complex with copper adsorbed at the hollow, top, and bridge positions. The MP2 curve shows a double minimum. This minimum has no physical meaning and is the result of BSSE correction together with the properties of the ANO-type basis set.

Tab. 1. Extrapolated binding energies, Δ*E* [kJ/mol], and optimal bond lengths, R (in terms of the shortest distance between the copper atom and the benzene surface) [Å], for Cu-benzene complex calculated at the DKH-MP2 and DKH-CCSD(T) theory level. All binding energies are BSSE corrected.

		hollow	top	bridge
MP2/ANO-RCC-VTZP	ΔE [kJ/mol]	10.26	12.75	15.36
	R [Å]	3.00	2.20	2.12
CCSD(T)/ANO-RCC-VTZP	ΔE [kJ/mol]	8.99	7.87	8.40
	R [Å]	3.01	2.57	2.54
MP2/ANO-RCC-VQZP	ΔE [kJ/mol]	14.31	20.71	24.62
	R [Å]	1.92	2.13	2.05
MP2/aug-cc-pVDZ-DK	ΔE [kJ/mol]	12.32	13.34	15.26
	R [Å]	2.97	2.27	2.19
CCSD(T)/aug-cc-pVDZ-DK	ΔE [kJ/mol]	11.70	10.57	11.34
	R [Å]	2.93	2.56	2.55
MP2/aug-cc-pVTZ-DK	ΔE [kJ/mol]	14.13	20.52	24.17
	R [Å]	2.73	2.14	2.07

action between the copper atom and the benzene molecule. The Cu-benzene complex with copper atom adsorbed at the bridge position represents the most stable structure. The MP2/ANO-RCC-VTZP and MP2/ANO-RCC-VQZP binding energies are 15.36 and 24.62 kJ/mol, respectively. The binding energy of copper positioned above the carbon atom (top position) is by about 16 % lower. The hollow position of the benzene molecule represents the least stable site determined at the MP2 theory level. Binding energy calculated with ANO-RCC-VTZP and ANO-RCC-VQZP basis set is lower by about 32 % and 42 %, respectively. Distances between the copper atom and benzene plane estimated by the MP2/ANO-RCC-VTZP method are 3.00 Å, 2.20 Å, and 2.12 Å for hollow, top, and bridge adsorption sites, respectively. The larger ANO-RCC-VQZP basis set provides similar results for the top and bridge positions, but the bond distance is by about 1.1 Å shorter above the benzene ring.

The CCSD(T) approach provides a different order of preferred adsorption sites, but the calculated CCSD(T)/ANO-RCC-VTZP binding energies for all three Cu adsorption positions are similar, i.e. 8.99, 7.87, and 8.40 kJ/mol for the hollow, top, and bridge positions, respectively. The bond distance between the copper atom and benzene plane is 2.93 Å for the most stable hollow position; equilibrium distances for the top and bridge positions are by about 0.4 Å shorter.

The relativistic aug-cc-pVXZ-DK basis set provided similar results. The CCSD(T) method also prefers the hollow position. Compared to the CCSD(T)/ANO-RCC-VTZP, the CCSD(T)/aug-cc-pVDZ-DK binding energies are by about 3 kJ/mol higher. For the hollow, top, and bridge positions, the binding energies correspond to 11.70, 10.75, and 11.34 kJ/mol, respectively. While the CCSD(T) approach prefers the hollow position, the MP2 method indicates the bridge position as the most stable Cu-benzene complex while adsorption above the center of the benzene ring corresponds to the least stable configuration.

Determination of the most stable position is very important, however, both correlation methods calculated in the finite basis set give different results. Therefore, the calculated results were extrapolated to the complete basis set (CBS).

The evaluated MP2/CBS and CCSD(T)/CBS binding energies and equilibrium distances are collected in Tab. 2. The DFT-D3 approach was fitted to the results in the complete basis set. Therefore, Tab. 2. also contains the DFT-D3 data.

Benchmark CCSD(T)/CBS results (labeled as CCSD(T)/[avTQ-avT]) were obtained by adding the correction term (difference of CCSD(T) and MP2 binding energies) calculated with the aug-cc-pVTZ-DK basis set to the MP2/CBS binding energy, evaluated by the Halkier's method of binding energies extrapolation at the MP2/aug-cc-pVTZ-DK and MP2/aug-cc-pVQZ-DK theory level (labeled as MP2/[avTQ]). Benchmark CCSD(T)/[avTQ-avT] binding energies for the hollow, top, and bridge positions are 16.77, 19.27, and 21.08 kJ/mol, respectively. The corresponding equilibrium distances are 2.57, 2.28, and 2.25 Å for (h), (t), and (b) positions, respectively. Computing of the correction term in the smaller aug-cc-pVDZ-DK basis set led to a decrease in the computing demand without accuracy loss. The same order of energies was obtained by direct extrapolation of CCSD(T)/aug-cc-pVDZ-DK and CCSD(T)/aug-cc-pVTZ-DK binding energies (labeled as CCSD(T)/[avDT]) to complete basis set. Binding energies of the Cu-benzene complex are 14.56, 16.99, and 18.49 kJ/mol for the (h), (t), and (b) structure, respectively.

The MP2/CBS overestimates the binding energies and decreases the bond distance between the cop-

Tab. 2. Extrapolated DKH-MP2/CBS, DKH-CCSD(T)/CBS and DFT-D3 binding energies, ΔE [kJ/mol], and optimal bond lengths, R (in terms of the shortest distance between the copper atom and the benzene surface) [Å], for Cu-benzene complex. Results were evaluated from aug-cc-PVXZ-DK (X = D, T, Q) or ANO-RCC-VXZP (X = D, T, Q) data. Basis sets used in the Halkier's extrapolation scheme are shown in square parentheses. Indication after the dash labels the basis for calculating the correction term. Binding energies calculated by the WFT methods were BSSE corrected.

		hollow	top	bridge
MP2/[avTQ]	ΔE [kJ/mol]	23.74	26.70	31.49
	R_e [Å]	1.82	2.09	2.02
CCSD(T)/[avDT]	ΔE [kJ/mol]	14.56	16.99	18.49
	R_e [Å]	2.66	2.31	2.27
CCSD(T)/[avTQ-T]	ΔE [kJ/mol]	16.77	19.27	21.08
	R_e [Å]	2.57	2.28	2.25
MP2/[TQ]	ΔE [kJ/mol]	22.77	25.94	30.61
	R_e [Å]	1.83	2.10	2.03
CCSD(T)/[DT]	ΔE [kJ/mol]	12.26	12.89	13.81
	R_e [Å]	2.81	2.38	2.34
CCSD(T)/[TQ-T]	ΔE [kJ/mol]	15.59	17.45	18.98
	R_e [Å]	2.64	2.30	2.25
DFT-D3/TPSS	ΔE [kJ/mol]	27.75	34.32	35.39
	R_e [Å]	2.45	2.18	2.15
DFT-D3/PBE	ΔE [kJ/mol]	23.61	34.46	34.74
	R_e [Å]	2.49	2.18	2.16

per atom and benzene plane, especially for the (t) position. For example, while the energy difference between top and hollow positions is overestimated by about 20 % in comparison with the reference CCSD(T)/[avTQ-avT] results, the energy gap between the preferred bridge and top position is larger by about 100 %. The bond distance is shorter by about 0.2 Å for the top and bridge position, but the equilibrium distance in the hollow structure is shorter by about 0.7 Å.

The MP2/CBS results obtained using the ANO-RCC-VTZP and ANO-RCC-VQZP basis sets correspond to the MP2/[avTQ] results. The energy differences between MP2/[avTQ] and MP2/[TQ] are below 1 kJ/mol for all three positions.

Including the correction term when computing with the ANO-RCC-VTZP basis set leads to an increase in error; differences between CCSD(T)/[TQ-T] and benchmark CCSD(T)/[avTQ-avT]) binding energies are above 1 kJ/mol. However, it should be mentioned that compared to the aug-cc-pVTZ-DK basis set, the ANO-RCC-VTZP basis set is smaller and the calculations are less demanding.

Direct extrapolation of CCSD(T)/ANO-RCC-VDZP and CCSD(T)/ANO-RCC-VTZP binding energies to CBS led to a poorer description of the interactions in the Cu-benzene complex. This lower accuracy is caused by the low quality of the basis sets, especially the ANO-RCC-VDZP basis set which contains only polarization functions.

The DFT-D3/TPSS and DFT-D3/PBE approaches overestimate the binding energies. Compared to the reference CCSD(T)/[avTQ-avT] results, the DFT-D3 binding energies are larger by about 40—80 %. All these methods provide an order of the stability of the Cu-benzene complex corresponding to the (b), (t), and (h) positions. This is contrary to the results obtained using finite basis sets, where the adsorption of the copper atom at the hollow position was preferred.

Low binding energy for copper is indicative of non-covalent binding. Compared to our previous study (Granatier et al., 2011), the CCSD(T)/ANO-RCC-VTZP results indicate that the Cu-benzene and Ag-benzene complexes are characterized by a similar interaction model. Both, copper and silver atoms prefer the hollow position and the binding energies are very close, indicating that the copper atom is bound to benzene by a dispersion interaction. However, a more detailed analysis provides different conclusion. The ROHF and MP2 approaches suggest charge transfer from the benzene molecule to the 4s orbital of the copper atom. Compared to the silver and gold atoms, the magnitude of charge transfer in the Cu-benzene complex corresponds to the value of magnitude of charge transfer in the Au-benzene complex. Moreover, the charge transfer in the Cu-benzene complex strongly depends on the position of the adsorbed metal atom. Mulliken population analysis showed that the magnitude of charge transfer in top and bridge positions is more than twice that in hollow position. Similar dependence was observed in the Au-benzene complex. On the other hand, the Ag-benzene complex is characterized by a similar charge for all three positions.

Adsorption of metal atoms on graphene surface has been studied previously (Manadé et al., 2015), showing the most stable position of the copper atom determined at the DFT/PBE and DFT-D2/PBE theory level to correspond to the adsorption of the metal atom directly on the carbon atom (t). Preferring the (t) or (b) position indicates that charge transfer plays an important role in binding copper on the graphene surface. The binding energy of this structure is 22.19 and 51.14 kcal/mol at the DFT and DFT-D2 theory level, respectively. However, the bond distance between copper and carbon surface is 2.21 Å, which is in good agreement with our results. Compared to our DFT-D3 results, the DFT/PBE and DFT-D2/PBE strongly overestimate the bond between the copper atom and the carbon surface.

Conclusion

Decoration of graphene with metal atoms represents a way of properties modification and new materials development. This requires understanding the nature of the bond between the metal and graphene. The Cu-benzene complex is a very simple model to be studied to understand the interactions between the copper atom and the graphene surface.

Therefore, this study was focused on the interaction of copper atom with benzene molecule using accurate correlation methods. Correct description of the interactions in the Cu-benzene complex requires the use of high accuracy correlation methods and high-quality basis sets and their extrapolation to complete basis set.

The DFT-D3 method is characterized by lower demandingness and the ability to describe the van der Waals complexes. Compared to high accuracy correlation methods, the DFT-D3 method provides the correct order of stability but strongly overestimates the binding energies for all three positions.

The copper atom prefers the adsorption on the bridge position. This structure of the Cu-benzene complex is characterized by strong charge transfer from the benzene molecule to the copper atom. Analogous to the Au-benzene complex, the charge transfer plays an important role in the understanding of the interactions in the Cu-benzene complex.

Thus, the significant difference between binding energy in the Cu-benzene and Au-benzene complexes will be the subject of further study.

Acknowledgement

This paper was supported by the Slovak Grant Agency VEGA (1/0461/21). We are grateful to the HPC center at the Slovak University of Technology in Bratislava, which is a part of the Slovak Infrastructure of High-Performance Computing (SIVVP project, ITMS code 26230120002, funded by the European region development funds, ERDF) for the computational time and resources made available.

References

Ago H, Ogawa Y, Tsuji M, Mizuno S, Hibino H (2012) J. Phys. Chem. Lett. 3(16): 2228—2236.

Baby TT, Aravind SSJ, Arockiadoss T, Rakhi RB, Ramaprabhu S (2010) Sensors Actuators, B Chem. 145(1): 71—77.

Balabanov NB, Peterson KA (2005) J. Chem. Phys. 123(6): 1—15.

Balabanov NB, Peterson KA (2006) J. Chem. Phys. 125(7): 1—11.

Bilić A, Reimers JR, Hush NS, Hoft RC, Ford MJ (2006) J. Chem. Theory Comput. 2(4): 1093—1105.

Boys SF, Bernardi F (1970) Mol. Phys. 19(4): 553—566.

Bučinský L, Biskupič S, Ilčin M, Lukeš V, Laurinc V (2011) J. Comput. Chem. 32: 356—367.

Caputo R, Prascher BP, Staemmler V, Bagus PS, Wöll C (2007) J. Phys. Chem. A. 111(49): 12778—12784.

Chan KT, Neaton JB, Cohen ML (2008) Phys. Rev. B — Condens. Matter Mater. Phys. 77(23): 1—12.

Cho WJ, Cho Y, Min SK, Kim WY, Kim KS (2011) J. Am. Chem. Soc. 133(24): 9364—9369.

Cho Y, Choi YC, Kim KS (2011) J. Phys. Chem. C. 115(13): 6019—6023.

Douglas M, Kroll NM (1974) Ann. Phys. (N. Y.). 82(1): 89—155.

Dunning TH (1989) J. Chem. Phys. 90(2): 1007—1023.

Fdez Galván I, Vacher M, Alavi A, Angeli C, Aquilante F, Autschbach J, Bao JJ, Bokarev SI, Bogdanov NA, Carlson RK, Chibotaru LF, Creutzberg J, Dattani N, Delcey MG, Dong SS, Dreuw A, Freitag L, Frutos LM, Gagliardi L, Gendron F, Giussani A, González L, Grell G, Guo M, Hoyer CE, Johansson M, Keller S, Knecht S, Kovačević G, Källman E, Manni GLi, Lundberg M, Ma Y, Mai S, Malhado JP, Malmqvist PÅ, Marquetand P, Mewes SA, Norell J, Olivucci M, Oppel M, Phung QM, Pierloot K, Plasser F, Reiher M, Sand AM, Schapiro I, Sharma P, Stein CJ, Sørensen LK, Truhlar DG, Ugandi M, Ungur L, Valentini A, Vancoillie S, Veryazov V, Weser O, Wesołowski TA, Widmark P-O, Wouters S, Zech A, Zobel JP and Lindh R (2019) Openmolcas: From source code to insight, J. Chem. Theory Comput. 15(11): 5925—5964.

Giovannetti G, Khomyakov PA, Brocks G, Karpan VM, Van Den Brink J, Kelly PJ (2008) Phys. Rev. Lett. 101(2): 4—7.

Granatier J (2017) Acta Chim. Slovaca. 10(2): 159—164.

Granatier J, Dubecký M, Lazar P, Otyepka M, Hobza P (2013) J. Chem. Theory Comput. 9(3): 1461—1468.

Granatier J, Lazar P, Otyepka M, Hobza P (2011) J. Chem. Theory Comput. 7(11): 3743—3755.

Grimme S, Antony J, Ehrlich S, Krieg H (2010) J. Chem. Phys. 132: 154104.

Halkier A, Helgaker T, Jørgensen P, Klopper W, Koch H, et al. (1998) Chem. Phys. Lett. 286(3—4): 243—252.

Heath JR, Ratner MA (2003) Phys. Today. 56: 43—49.

Heckert M, Heun O, Gauss J, Szalay PG (2006) J. Chem. Phys. 124(12): 124105.

Hess B, Chandra P (1987) Phys. Scr. 36: 412.

Hong W, Bai H, Xu Y, Yao Z, Gu Z, Shi G (2010) J. Phys. Chem. C. 114(4): 1822—1826.

Jurečka P, Hobza P (2002) Chem. Phys. Lett. 365(1—2): 89—94.

Kendall RA, Dunning TH, Harrison RJ (1992) J. Chem. Phys. 96(9): 6796—6806.

Khomyakov PA, Giovannetti G, Rusu PC, Brocks G, Van Den Brink J, Kelly PJ (2009) Phys. Rev. B — Condens. Matter Mater. Phys. 79(19): 1—12.

Lazar P, Granatier J, Klimesd J, Hobzab P, Otyepka M (2014) Phys. Chem. Chem. Phys. 16: 20818—20827.

Li Y, Fan X, Qi J, Ji J, Wang S, et al. (2010a) Mater. Res. Bull. 45(10): 1413—1418.

Li Y, Fan X, Qi J, Ji J, Wang S, et al. (2010b) Nano Res. 3(6): 429—437.

Malček M, Bučinský L (2020) Theor. Chem. Acc. 139(11): 167.

Manadé M, Viñes F, Illas F (2015) Carbon N. Y. 95: 525—534.

Møller C, Plesset MS (1934) Phys. Rev. 46(7): 618—622.

Neese F (2012) Wiley Interdiscip. Rev. Comput. Mol. Sci. 2(1): 73—78.

Neogrády P, Urban M (1995) Int. J. Quantum Chem. 55(2): 187—203.

Nieminen V, Honkala K, Taskinen A, Yu Murzin D (2008) J. Phys. Chem. C. 112(17): 6822—6831.

Perdew JP, Burke K, Ernzerhof M (1996) Phys. Rev. Lett. 77(18): 3865—3868.

Perdew JP, Burke K, Ernzerhof M (1997) Phys. Rev. Lett. 78(7): 1396—1396.

Roos BO, Lindh R, Malmqvist PÅ, Veryazov V, Widmark PO (2004) J. Phys. Chem. A. 108(15): 2851—2858.

Roos BO, Lindh R, Malmqvist PÅ, Veryazov V, Widmark PO (2005) J. Phys. Chem. A. 109(29): 6575—6579.

Singla M, Sharma D, Jaggi N (2021) Int. J. Hydrogen Energy. 46(29): 16188—16201.

Scheuermann GM, Rumi L, Steurer P, Bannwarth W, Mülhaupt R (2009) J. Am. Chem. Soc. 131(23): 8262—8270.

Shan C, Yang H, Han D, Zhang Q, Ivaska A, Niu L (2010) Biosens. Bioelectron. 25(5): 1070—1074.

Syomin D, Kim J, Koel BE, Elison GB (2001) J. Phys. Chem. B. 105(35): 8387—8394.

Tao J, Perdew JP, Staroverov VN, Scuseria GE (2003) Phys. Rev. Lett. 91(14): 3—6.

Urban M, Neogrády P, Hubač I (1997) In Recent Advances in Coupled-Cluster Methods, pp. 275—306. Singapore. World Scie ed.

Watts JD, Gauss J, Bartlett RJ (1993) J. Chem. Phys. 98(11): 8718—8733.

Xiong Z, Zhang LL, Ma J, Zhao XS (2010) Chem. Commun. 46(33): 6099—6101.

Weigend F, Ahlrichs R (2005) Phys. Chem. Chem. Phys. 7(18): 3297—3305.

sciendo

Theoretical investigation of aniline derivatives: Correlation of theoretical reaction Gibbs free energies with experimental oxidation potentials

Andrea Kováčová, Martin Michalík

Institute of Physical Chemistry and Chemical Physics, Slovak University of Technology in Bratislava, Radlinského 9, SK–812 37 Bratislava, Slovakia
andrea.kovacova@stuba.sk

Abstract: Anilines and their derivatives are used in industrial production of dyes, pharmaceuticals, plastics, and synthetic antioxidants. Abstraction of electron and formation of a cation radical represent a significant step in the reactivity of this group of substances. The aim of this study was to theoretically investigate the effect of a substituent on the oxidative electrochemical potential of 64 species of aniline derivatives. Quantum-chemical calculations were performed using the composite G4 method. The obtained linear dependences for reaction Gibbs free energies were correlated with the Hammett constants and available experimental values.

Keywords: anilines, electrochemical potential, reaction Gibbs energies, Gn approach, thermodynamics, implicit solvent model

Introduction

In organic materials, a number of chemical processes are initiated by homolytic cleavage of chemical bonds. Reactivity and stability widely vary for such generated radicals which is the subject of many studies. In case of polymers, formation of radicals can lead either to the formation of new materials with interesting properties, i.e. cross-linked materials (Skákalová, Lukeš and Breza, 1997), or to simple degradation. Antioxidants which can trap these radicals are mostly phenol- (Kleinová et al., 2024), or aniline-based (Kováčová et al., 2023; Lukeš and Hartmann, 2021) compounds. Derivatives of the latter have a wide spectrum of applications in various industries, e.g., polymers, dyes, rubber, medicine, pharmacology, and electronics (Anjalin, Kanagathara and Baby Suganthi, 2020; Lukeš et al., 2011; Rappoport, 2007). *Para*-aniline derivatives also include hydroquinones (Rani and Shanker, 2017). Lately, Pavitt et. al published electrochemical potentials of aniline derivatives measured in water by the SCV method (Pavitt, Bylaska and Tratnyek, 2017). Their study inspired us to providing theoretical background while testing the current state-of-art calculation approach, G4.

The composite G4 method is often used as a more precise method for reference theoretical calculations, e.g. bond dissociation enthalpy (Biela et al., 2023). Needless to say, this method is much more computationally expensive as it uses a complex multi-step procedure. One of the steps involves single point calculation at the MP4 level of theory, for which 128 GB of RAM is necessary even for medium sized molecules. Still, a more conventional QSPR approach in the first step is adequate in many cases to obtain approximate classification of derivatives.

Despite being almost a century old, the Hammett constant still allows for reasonable quantification of electron-donating (negative values) or electron-withdrawing (positive values) properties of substituents. There are substituents with different effect in meta and in para position, it is also possible to use Hammett type constants (for example Brown-Okamoto (Hansch, Leo and Taft, 1991)) plus (σ^+) or minus (σ^-) for substituents in the para position. Once linear dependences are found and enumerated for a given dataset of derivatives with known properties, prediction for other compounds is possible using QSPR, e.g. BDE (Vagánek et al., 2013).

Considering the above points, the main aims of this paper are: (1) to find optimal geometries of studied molecules using steps within the G4 method; (2) to calculate reaction Gibbs free energies related to the reaction in Eq. 1; (3) to evaluate electrochemical potentials and to compare obtained values with available experimental data, and (4) to assess the substitution effect using the QSPR approach. This work is a continuation of a similar study carried out on phenolic derivatives (Kováčová, 2024).

Calculation details

Quantum-chemical calculations

Quantum-chemical calculations were performed by composite Gaussian-4 level (G4) of theory (Curtiss, Redfern and Raghavachari, 2007) as implemented in the Gaussian 09 program package (Frisch et al., 2009). For comparison, computa-

Acta Chimica Slovaca, Vol. 17, No. 1, 2024, pp. 74—82, DOI: 10.2478/acs-2024-0010

tional analyses were performed using the density functional theory at the B97D3/def2-TZVPP level (Grimme, Ehrlich and Goerigk, 2011; Weigend and Ahlrichs, 2005). The solvent effect contribution of water was simulated using SMD implicit solvent model (solvation model based on the quantum mechanical charge density of a solute molecule interacting with a continuum) (Marenich, Cramer and Truhlar, 2009). This model is based on the generalised Born equation, which is represented by an approximation of the Poisson equation suitable for arbitrary cavity shapes. Gibbs free energies were evaluated for the temperature of 298.15 K and the pressure of 101.325 kPa. Vibrational analysis included in the G4 calculation process showed no imaginary frequencies, which confirms geometry corresponding to the lowest energy minimum. Optimal geometries were visualised for the *ortho*- and *poly*-substituted aniline derivatives using the open-source Jmol viewer (Fig. 2) (Jmol development team, 2016). Pre-optimisation of some molecules was performed using the XTB program package (Bannwarth et al., 2021; Bannwarth, Ehlert and Grimme, 2019; Pracht, Bohle and Grimme, 2020). Structural schemes with notation of studied species are shown in Fig. 1.

Results and Discussion

To investigate the thermodynamics of one-electron transfer reaction according to Eq. 1

$$X—Ar—NH_2 \rightarrow (X—Ar—NH_2)^{\bullet+} + e^- \quad (1)$$

where X represents the substituents used, it was necessary to calculate the geometry of neutral molecule and the corresponding cation radical. Optimal geometries for selected *ortho*- and *poly*-substituted anilines in neutral states are depicted in Fig. 2. Reaction Gibbs free energies expressing electron abstraction ($\Delta_r G$) follow Eq. 2

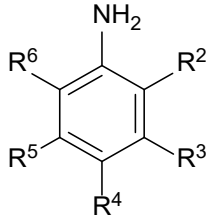
$$\Delta_r G = G((X—Ar—NH_2)^{\bullet+}) - G(X—Ar—NH_2) + G(e^-) = \\ = \Delta G + G(e^-) \quad (2)$$

where $G((X—Ar—NH_2)^{\bullet+})$ and $G(X—Ar—NH_2)$ denote the Gibbs free energies of the cation radical and neutral compound, respectively, and $G(e^-)$ is the Gibbs free energy of the solvated electron. Calculated $\Delta_r G$ values for all mono-substituted anilines can be found in Tab. 1. The value of $G(e^-)$ in Eq. 2 is constant (–149 kJ · mol^{-1}) for all reactions under study. This value was calculated by adding $G(e^-)$ for vacuum (–4 kJ · mol^{-1}, (Fifen, 2013)) with a solvation correction (water) –145 kJ · mol^{-1} (Ishikawa and Nakai, 2016).

According to the thermodynamic theory and after correction to standard hydrogen electrode (SHE) (Ho et al., 2016), the absolute electrochemical potential can be calculated from the reaction Gibbs energies using Eq. 3

$$E^{cal} = -(\Delta_r G / nF) - E^{SHE} \quad (3)$$

where E^{cal} denotes electrochemical potential, $\Delta_r G$ is the reaction Gibbs free energy, n represents the number of exchange electrons in the reaction, F is the Faraday constant (96 485 C · mol^{-1}), and E^{SHE} is the absolute electrochemical potential of SHE (4.28 V). For twelve aniline derivatives, experimental values of electrochemical potential, E, are



Name	Notation	Name	Notation
aniline	H	2,6-dimethylaniline	26Me
x-acetylaniline	xAc	3,4-dimethylaniline	34Me
x-bromoaniline	xBr	3,5-dimethylaniline	35Me
x-trifluoromethylaniline	xCF3	2,3-dimethoxyaniline	23OMe
x-chloroaniline	xCl	2,4-dimethoxyaniline	24OMe
x-cyanoaniline	xCN	2,5-dimethoxyaniline	25OMe
x-aminobenzaldehyde	xCOH	2,6-dimethoxyaniline	26OMe
x-aminobenzoic acid	xCOOH	3,4-dimethoxyaniline	34OMe
x-ethylaniline	xEt	3,5-dimethoxyaniline	35OMe
x-fluoroaniline	xF	2,4-dinitroaniline	24NO_2
x-methylaniline	xMe	2,4,6-trimethylaniline	246Me
x-aminoaniline	xNH2	2,4,6-trichloroaniline	246Cl
x-nitroaniline	xNO2	2,6-dinitro-4-methylaniline	4Me-26NO_2
x-nitrosoaniline	xNO	2-methoxy-4-ethylaniline	2OMe-4Et
x-methoxyaniline	xOMe	2-Hydroxy-4-methoxybenzaldehyde	2OMe-4COH
2,3-dimethylaniline	23Me	2-methyl-5-nitroaniline	2Me-5NO_2
2,4-dimethylaniline	24Me	2-methoxynitroaniline	2OMe-NO_2
2,5-dimethylaniline	25Me	4-methyl-3-nitroaniline	4Me-3NO_2

Fig. 1. Schematic structures and notations of studied molecules. X = 2 for *ortho*- derivatives, X = 3 for *meta*- derivatives and X = 4 for *para*- derivatives.

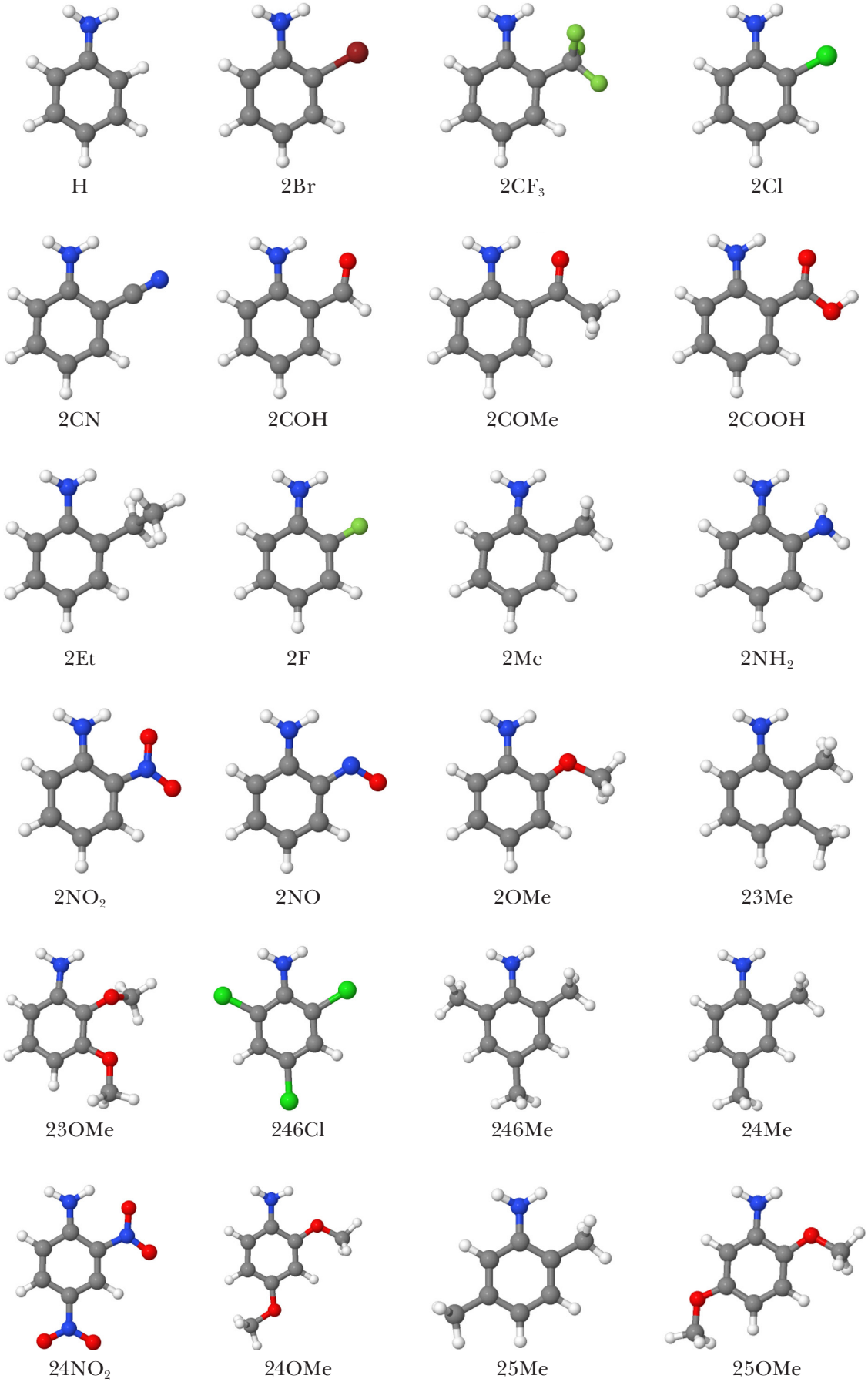


Fig. 2. G4 optimal geometries of *ortho*- and *poly*-substituted aniline derivatives.

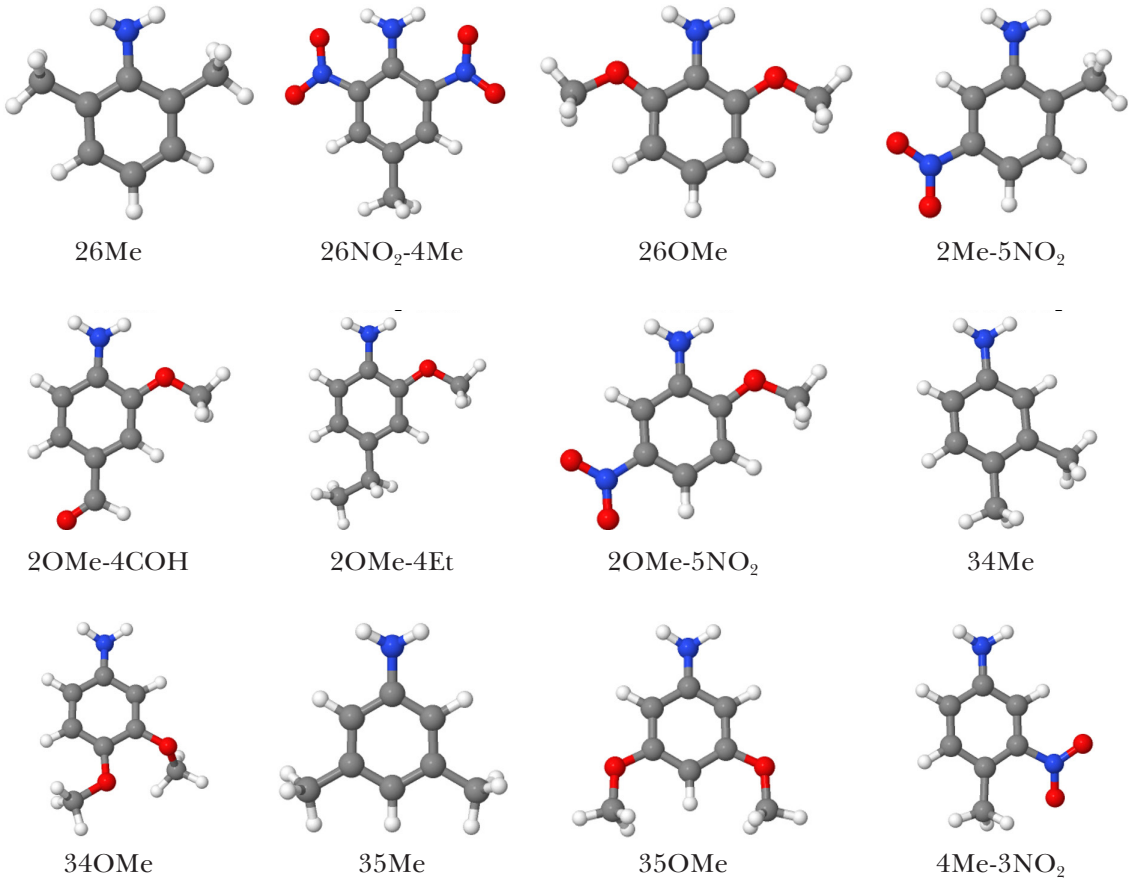


Fig. 2 continue. G4 optimal geometries of *ortho*- and *poly*-substituted aniline derivatives.

Tab. 1. Takahata-Chong Hammett type constants for *ortho* isomers (σ_o) (Takahata and Chong, 2005), and Hammett constants for *meta* (σ_m) and *para* (σ_{p+}) isomers (Hansch, Leo and Taft, 1991). Calculated G4(water) reaction Gibbs free energies for electron abstraction, experimental electrochemical potential E^{exp} vs. SHE (Pavitt, Bylaska and Tratnyek, 2017), and bond length (d) for C—N in monosubstituted anilines.

Substituent	Hammett constants			$\Delta_r G$/kJ·mol^{-1}			E^{exp}/V			d/Å		
	σ_O	σ_m	σ_{p+}	*ortho*	*meta*	*para*	*ortho*	*meta*	*para*	*ortho*	*meta*	*para*
H	0.00	0.00	0.00	369	369	369	1.07	1.07	1.07	1.394	1.394	1.394
Br		0.39	0.15	395	387	376				1.385	1.387	1.389
CF3		0.43	0.61	395	385	391				1.385	1.388	1.381
Cl	0.20	0.37	0.11	389	385	373	1.204	1.167	1.058	1.385	1.387	1.390
CN	0.71	0.56	0.66	406	397	398				1.372	1.385	1.369
COH		0.35	0.73	393	387	392				1.354	1.388	1.363
COMe		0.38		382	380	390				1.358	1.391	1.369
COOH		0.37	0.42	389	385	392		1.074	1.122	1.363	1.389	1.370
Et	–0.31	–0.07	–0.30	357	368	357				1.398	1.395	1.397
F	0.24	0.34	–0.07	383	382	367				1.392	1.388	1.398
Me	–0.17	–0.07	–0.31	363	363	352	1.017	1.029	0.927	1.396	1.395	1.398
NH_2	–0.41	–0.16	–1.30	325	342	296				1.405	1.396	1.408
NO_2	0.80	0.71	0.79	415	401	409	1.372	1.269	1.35	1.345	1.383	1.350
NO		0.62		386	401	389				1.343	1.385	1.344
OMe	–0.39	0.12	–0.78	349	369	325	0.894	1.023	0.766	1.397	1.393	1.403

available, ranging from 0.63 V for $4NH_2$ derivate (electron-donor) to 1.26 V for 4COH and 2CN derivatives (electron-acceptor), see Tab. 1.

In Fig. 3, the dependence between Takahata-Chong or Hammett constants and E^{exp} experimental values for *ortho*-, *meta*-, and *para*- derivatives, combined is shown. Coefficient of determination (R^2) of 0.915 was found from linear regression. Line parameters *a*, *b* with errors are summarised in Tab. 2.

The substituent constants however work best only for mono-substituted derivatives. Another descriptor, which does not have this limitation might be the C—N bond length, where N is amino nitrogen of neutral derivatives. While the changes are quite subtle, good correlation was found with E^{exp} and R^2 of 0.927 for 14 points (Fig. 5). The shorter the bond, the more is the lone pair of nitrogen involved in benzene ring conjugation. The trend does not follow substituents with the strongest effect on the ring, i.e., $2NO_2$, $4NO_2$, and 4COOH derivatives. On the other hand, if also *poly*-substituted anilines are included in the theoretical $\Delta_r G$ values, even better correlation can be found. Depicted in Fig. 4 (a) for 18 points, the value of R^2 is 0.955. Interestingly, for 5 *ortho* derivatives, the coefficient of determination even reached 0.993, which is not the case for *meta* and *para* derivatives.

A comparison between the linear correlations based on the B97D3/def2-TZVPP level of theory data presented in Fig. 4(b) and the original data set (see Tab. 2) revealed similar line parameters. Nevertheless, the coefficients of determination are slightly lower (R^2 of 0.939). However, it should be noted, that DFT takes only ca. 10 % of the overall computational time of the G4 method.

This relationship naturally arises from Eq. 3, but the intercepts found are far from the value of SHE electrochemical potential. In fact, electrochemical potentials, E^{cal}, for all anilines under study can be directly calculated based on Eq. 3 and compared with available data. A comparison of the three methods of theoretical prediction of electrochemical potential is shown in Fig. 6. While individual values obtained from correlation methods ($E^{\Delta G}$, E^{σ}) are relatively close to each other, E^{cal} values show large discrepancies. This might be attributed to the varying experimental conditions e.g., overpotential of the electrode, which is included in the regression models $E^{\Delta G}$ and E^{σ} trained on experimental data. The shift from experimental values is largely consistent when experimental conditions remain constant. Consequently, data from diverse sources with varying conditions may not be compatible. It can be argued that values scaling is a common practice in theoretical modelling, as it facilitates better correlation with experiments often correcting shortcomings of the implicit solvation model. Thermodynamic quantities, such as pK_a (Busch et al., 2022), bond dissociation energies (Bakalbassis, Lithoxoidou and Vafiadis, 2003), and ionisation potential (Fu et al., 2005), are particularly characteristic of this phenomenon, whereby computed theoretical values are mapped onto experimental ones. Some improvement is expected if explicit solvent molecules are added, but calculations including advanced explicit solvation models are beyond the scope of this paper.

Conclusion

In this study, the effect of a substituent on the oxidative electrochemical potential of 64 species of aniline derivatives has been theoretically investigated. The results presented here were obtained

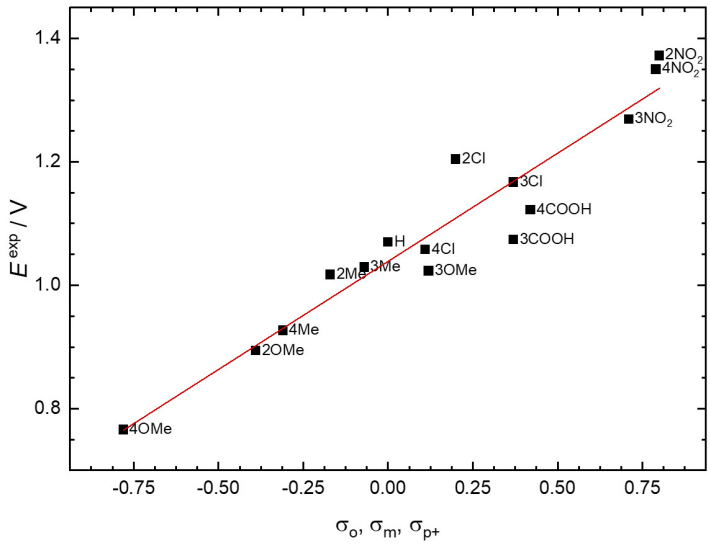


Fig. 3. Dependence of electrochemical potentials on Takahata-Chong Hammett type constants (*ortho*- derivatives) and Hammett constants (*meta*, *para*) for aniline derivatives.

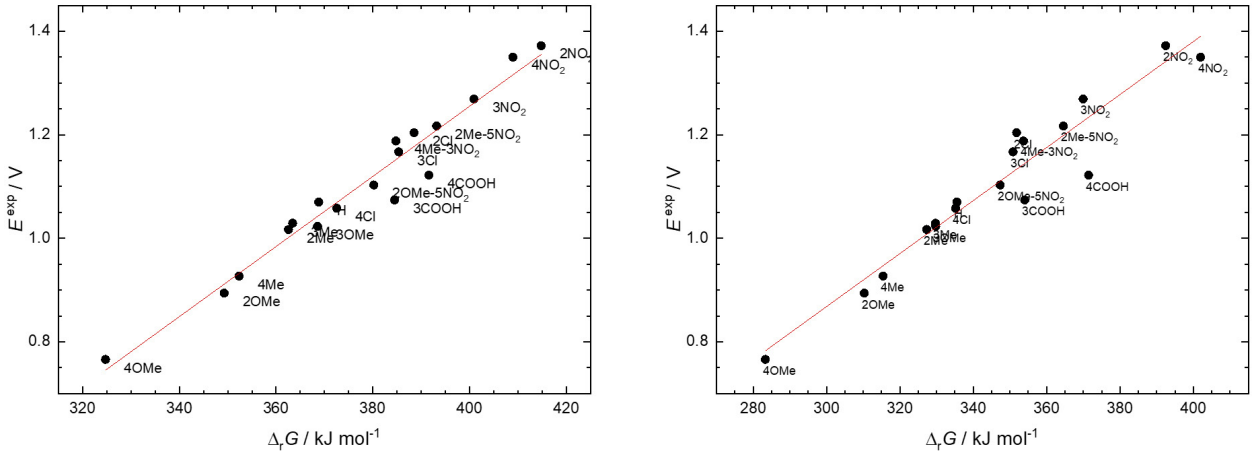


Fig. 4. Dependence of electrochemical potentials on reaction Gibbs free energies by (a) G4 method and (b) DFT-D3 level of theory.

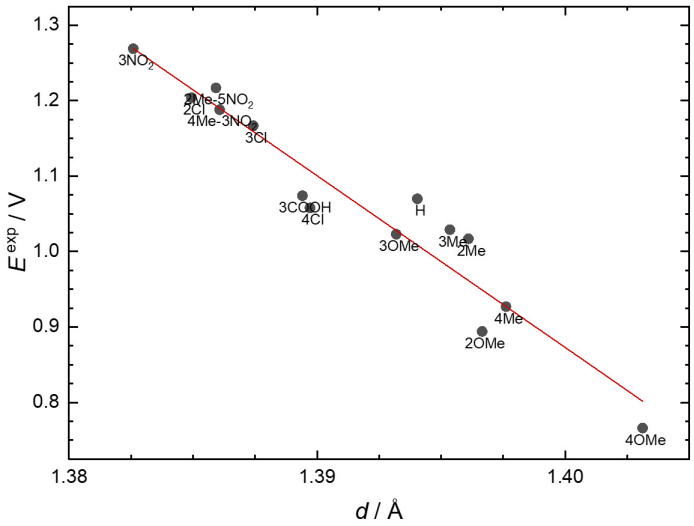


Fig. 5. Dependence of available experimental values of electrochemical potentials on C—N bond length (G4 method).

Tab. 2. a, b parameters with errors and coefficient of determination (R^2) for linear correlations described by equations E/V = a + b × ΔG/kJ·mol^{-1}, E/V = a + b × σ or E/V = a + b × d/Å, respectively, for *ortho*- (E_o), *meta*- (E_m), *para*- (E_p), and *poly*-substituted (E_{poly}) anilines. B97D3 denotes DFT data, otherwise, data from the G4 method are employed.

	a	***b* 10^{-3}**	**n**	**R^2**
$E_{o,\Delta G}$	–1.58 ±0.13	7.15 ±0.34	5	0.993
$E_{m,\Delta G}$	–1.24 ±0.50	6.2 ±1.3	6	0.846
$E_{p,\Delta G}$	–1.35 ±0.25	6.49 ±0.67	6	0.959
$E_{poly,\Delta G}$	–1.31 ±0.62	6.4 ±1.6	4	0.885
$E_{\Delta G}$	–1.45 ±0.14	6.77 ±0.37	18	0.955
$E_{\Delta G,B97D3}$	–0.71 ±0.12	5.24 ±0.33	18	0.939
E_{σ}	1.039 ±0.014	350 ±30	15	0.915
E_{d}	32.8 ±2.6	–22800 ±1800	14	0.927
$E_{d,B97D3}$	29.4 ±3.3	-20400 ±2400	14	0.859

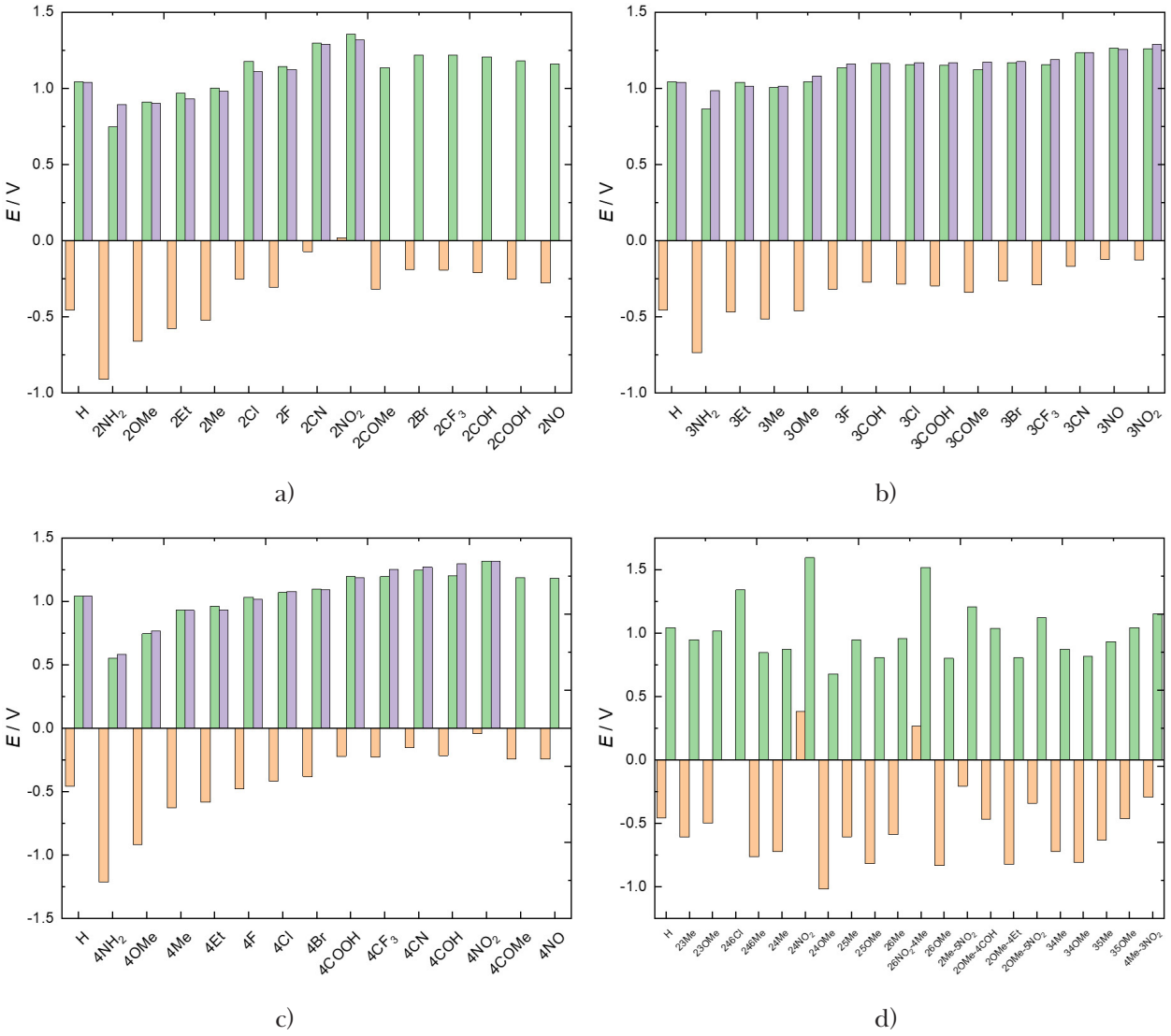


Fig. 6. Comparison of G4 calculated electrochemical potentials E^{cal} (orange), $E^{\Delta G}$ (green), E^{σ} (purple) for selected substituents. Values of E^{σ} were calculated for substituents with available Hammett constants only.

Tab. 3. Calculated G4(water) reaction Gibbs free energies ($\Delta_r G$) for electron abstraction, experimental electrochemical potential (E^{exp}) vs. SHE (Pavitt, Bylaska and Tratnyek, 2017) and bond length (d) for C—N of *poly*-substituted aniline derivatives.

Substituent	$\Delta_r G$/kJ·mol^{-1}	E^{exp}/V	d/Å	Substituent	$\Delta_r G$/kJ·mol^{-1}	d/Å
24N$_O$2	450		1.333	**23OMe**	365	1.392
2OMe-5NO$_2$	380	1.103	1.386	**24OMe**	315	1.405
4Me-26NO$_2$	439		1.331	**25OMe**	334	1.394
2Me-5NO$_2$	393	1.217	1.386	**26OMe**	333	1.400
23Me	354		1.398	**34OMe**	335	1.401
24Me	343		1.399	**35OMe**	368	1.393
25Me	355		1.397	**2OMe-4Et**	334	1.399
26Me	356		1.397	**2OMe-4CHO**	368	1.367
34Me	343		1.399	**246Me**	339	1.401
35Me	352		1.397	**246Cl**	413	1.373
4Me-3NO$_2$	385	1.188	1.386			

Tab. 4. Calculated B97D3/def2-TZVPP reaction Gibbs free energies for electron abstraction and bond length (*d*) for C—N of monosubstituted anilines.

Substituent	$\Delta_r G_{B97D3}$/kJ·mol^{-1}			*d*/Å		
	ortho	*meta*	*para*	*ortho*	*meta*	*para*
H	336	336	336	1.393	1.393	1.393
Br	359	353	338	1.383	1.387	1.389
CF3	366	356	364	1.381	1.387	1.378
Cl	352	351	335	1.384	1.387	1.390
CN	375	365	371	1.369	1.384	1.367
COH	373	359	375	1.351	1.386	1.358
COMe	359	352	364	1.355	1.390	1.365
COOH	362	354	371	1.360	1.388	1.366
Et	326	338	320	1.396	1.395	1.398
F	346	348	330	1.389	1.387	1.397
Me	327	330	315	1.395	1.395	1.398
NH_2	285	298	254	1.404	1.396	1.409
NO_2	392	370	402	1.345	1.381	1.348
NO	345	365	333	1.342	1.383	1.343
OMe	310	330	283	1.395	1.394	1.402

Tab. 5. Calculated B97D3/def2-TZVPP reaction Gibbs free energies for electron abstraction and bond length (*d*) for C—N *of poly*-substituted aniline derivatives.

Substituent	$\Delta_r G_{B97D3}$/kJ·mol^{-1}	*d*/Å	Substituent	$\Delta_r G_{B97D3}$/kJ·mol^{-1}	*d*/Å
24NO_2	449	1.334	**23OMe**	319	1.391
2OMe-5NO_2	347	1.383	**24OMe**	277	1.405
4Me-26NO_2	418	1.332	**25OMe**	295	1.394
2Me-5NO_2	365	1.383	**26OMe**	294	1.398
23Me	321	1.397	**34OMe**	291	1.402
24Me	309	1.399	**35OMe**	327	1.394
25Me	319	1.396	**2OMe-4Et**	293	1.399
26Me	322	1.395	**2OMe-4CHO**	343	1.362
34Me	310	1.399	**246Me**	303	1.400
35Me	326	1.396	**246Cl**	370	1.374
4Me-3NO_2	354	1.386			

by means of the expensive composite *ab initio* G4 method. It is preferable to predict electrochemical values with the use of a known line trend derived from a correlation with experimental values, as opposed to calculating them from a principal equation. As it has been demonstrated in context of mono-substituted derivatives, empirical substituent constants can be employed to a certain extent to determine the relationship with the quantity of interest. In certain cases, the requisite linear trend may not be observed within the available quantities. Moreover, Hammett constants and available experimental values were used to evaluate the trends. Consequently, more sophisticated calculation methods are required. However, it should be noted that the C—N bond length used in the correlation were calculated in the very first step of G4, at simple B3LYP/6-31G(2df,p) level of theory. While the G4 method is expected to yield highly accurate energies, concerns may arise regarding the precision of the solvation model employed. Additional DFT-D3 calculations revealed a similar discrepancy in comparison with experimental data, thereby confirming the limitations of the implicit model. It can be concluded that the knowledge of calculated ΔG value for all studied aniline derivatives in combination with the linear Eq. 6 allows for reasonable determination of electrochemical potential in water.

Acknowledgement
This paper has been supported by the Ministry of Education, Science, Research and Sport of the Slovak

Republic, project VEGA 1/0461/21. We are grateful to the HPC centre at the Slovak University of Technology in Bratislava, which is a part of the Slovak Infrastructure of High-Performance Computing (SIVVP project, ITMS code 26230120002, funded by the European region development funds, ERDF) for the computational time and resources made available.

References

Anjalin M, Kanagathara N, Baby Suganthi AR (2020) Mater. Today Proc. 33: 4751–4755.

Bakalbassis EG, Lithoxoidou AT, Vafiadis AP (2003) J. Phys. Chem. A. 107(41): 8594–8606.

Bannwarth C, Caldeweyher E, Ehlert S, Hansen A, Pracht P, et al. (2021) Wiley Interdiscip. Rev. Comput. Mol. Sci. 11(2): 1–49.

Bannwarth C, Ehlert S, Grimme S (2019) J. Chem. Theory Comput. 15(3): 1652–1671.

Biela M, Kleinová A, Uhliar M, Klein E (2023) J. Mol. Graph. Model. 122: 1–6.

Busch M, Ahlberg E, Ahlberg E, Laasonen K (2022) ACS Omega. 7(20): 17369–17383.

Curtiss LA, Redfern PC, Raghavachari K (2007) J. Chem. Phys. 127(12): 1–8.

Fifen JJ (2013) J. Chem. Theory Comput. 9(7): 3165–3169.

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr., Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Keith T, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS,Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J and Fox DJ (2009) Gaussian 09, Revision D.01, Gaussian, Inc. Wallingford CT.

Fu Y, Liu L, Yu HZ, Wang YM, Guo QX (2005) J. Am. Chem. Soc. 127(19): 7227–7234.

Grimme S, Ehrlich S, Goerigk L (2011) J. Comput. Chem. 32(7): 1456–1465.

Hansch C, Leo A, Taft R (1991) Chem. Rev. 91: 165–195.

Ho J, Coote LM, Cramer CJ, Truhlar DG (2016) In Organic Electrochemistry, pp. 229–259. CRC Press. 5th ed.

Ishikawa A, Nakai H (2016) Chem. Phys. Lett. 650: 159–164.

Jmol development team. (2016) Jmol: an open-source Java viewer for chemical structures in 3D (14.6). http://jmol.sourceforge.net/.

Kleinová A, Biela M, Lukeš V, Klein E (2024) J. Mol. Struct. 1303: 137646.

Kováčová A (2024) Acta Chim. Slovaca. 17(1): 12–21.

Kováčová A, Michalík M, Hartmann H, Lukeš V (2023) J. Mol. Liq. 389: 122811.

Lukeš V, Hartmann H (2021) Color. Technol. 137(4): 389–398.

Lukeš V, Rapta P, Idzik KR, Beckert R, Dunsch L (2011) J. Phys. Chem. B. 115(13): 3344–3353.

Marenich AV, Cramer CJ, Truhlar DG (2009) J. Phys. Chem. B. 113(18): 6378–6396.

Pavitt AS, Bylaska EJ, Tratnyek PG (2017) Environ. Sci. Process. Impacts. 19(3): 339–349.

Pracht P, Bohle F, Grimme S (2020) Phys. Chem. Chem. Phys. 22(14): 7169–7192.

Rani M, Shanker U (2017) J. Environ. Chem. Eng. 5(6): 5298–5311.

Rappoport Z (2007) The chemistry of Anilines, Part 1, Vol. 11 New York: Wiley. 1163.

Skákalová V, Lukeš V, Breza M (1997) Macromol. Chem. Phys. 198(10): 3161–3172.

Takahata Y, Chong DP (2005) Int. J. Quantum Chem. 103(5): 509–515.

Vagánek A, Rimarčík J, Ilčin M, Škorňa P, Lukeš V, Klein E (2013) Comput. Theor. Chem. 1014: 60–67.

Weigend F, Ahlrichs R (2005) Phys. Chem. Chem. Phys. 7(18): 3297–3305.

sciendo

Theoretical and experimental study of monoethanolamides and diethanolamides as potential additives to fossil diesel fuel

Teodora Kocsisová, Andrea Kováčová

Institute of Physical Chemistry and Chemical Physics, Slovak University of Technology in Bratislava, Radlinského 9, SK-812 37 Bratislava, Slovakia
andrea.kovacova@stuba.sk

Abstract: In this work, theoretical and experimental study of monoethanolamides (MEAD) and diethanolamides (DEAD) is presented. The aim was to investigate the thermodynamics of addition reactions used in the synthesis of MEAD and DEAD. The reactions have exothermic character and alkyl elongation affects the reaction enthalpy of DEAD production reactions. In the experimental part, the prepared MEAD and DEAD samples were preliminary tested as a potential additive in diesel fuel. The research focused on measuring three key properties: i.e., foaming, lubricity, and cetane number. The results suggest that MEAD is a suitable additive for diesel fuel.

Keywords: green chemistry, DFT, thermodynamics, eco-friendly additive, foaming, corrosivity

Introduction

Detergents are a group of surfactant molecules with an amphibole structure, where the long tail has a non-polar part and the other side has a short polar part. Alkanolamides may be considered such substances, exhibiting various interesting properties, which consist of a long non-polar alkyl chain and a polar amide moiety. While most surfactants increase the degree of corrosion, the diethanolamide solution protects both steel and iron from corrosion (Sanders, 1958). These substances also have anticorrosive properties, antifoaming properties, and they increase lubricity and cetane number of fuels (EP 1 227 143). One way to improve fuel economy of an engine and reduce fuel consumption is to reduce engine friction. Currently, fatty acid esters, fatty acid amides, and fatty acid esteramides are applied as environmentally friendly friction modifiers (Hui, 1998).

As it is reported in literature (Gaurav et al., 2017; Liu et al., 2021), alkanolamides can be prepared from various raw materials, e.g., FFA (free fatty acids), TAG (triacylglycerols), FAME (methyl ester of fatty acids), alkanolamines, i.e., monoethanolamine (MEA) or diethanolamine (DEA). Despite the wide variety of alkanolamides, monoethanolamides (MEAD) and diethanolamides (DEAD) are of particular interest in technical practice (Shahidi, 2005). An additive on this basis can potentially be added to any type of organic fuel (gasoline, diesel, FAME, ethanol) or to mixtures thereof. In general, optimum concentration of an additive is reported to be 1 to 5 % v/v relative to pure fuel (Kleinová, Cvengrošová and Cvengroš, 2013; Mikulec et al., 2010). In case of solutions with anhydrous ethanol or even with water, DEAD and MEAD have an emulsifying effect as they form a stable microemulsion. The presence of a nitrogen component in diesel, for example added urea, leads to in better combustion with lower CO, Nox, and solid particle emissions. This also results in lower fuel consumption. Therefore, ethanolamine or diethanolamine, from the production of MEAD or DEAD, need not be removed as an unreacted reactant from the mixture after amidation.

On the other hand, optimal adjustment of the concentration ratios of additives and fossil fuel with respect to different physicochemical properties is investigated. Considering the work published so far, the objective of this study is to investigate the thermodynamics of reactions occurring in the synthesis of MEAD and DEAD from fatty acids. In the experimental part, MEAD and DEAD samples were prepared by internal synthetical treatment and purification technics according to Pokorný and Dubská (1986). Consequently, these samples were tested as additives in diesel fuel. This part of research focused on three key properties i.e., foaming, lubricity and cetane number.

Materials and Methods

Samples preparation

Condensation reaction of MEA (Lachema, Brno) with technical oleic acid (Lachema, Brno) was used to prepare the MEAD sample. Acid number of oleic acid was 196.0 mg (KOH)/g and the declared acid abundances (acyl profile) were as follows: C16:0-9.8; C18:0-2.9; C18:1-61.2; C18:2-23.0; C18:3-2.2;

C20:0-1.0 (area percentages by GLC chromatogram).

To limit the formation of ester to a minimum, a slight molar excess of alkanolamine to fatty acids is usually used (1.02:1 to 1.10:1) at the working temperature for MEA, which is around its boiling point of 170 °C. Both reaction components are mixed at 80 °C to 120 °C. The reaction mixture is heated up to 160—170 °C. During the reaction, water is removed and the acid number decreases. Continuous analysis of the reaction mixtures was not performed, only the final products of the reaction were determined.

The reaction can be considered complete when the acid number drops to 5 mg KOH/g or lower. The excess MEA is distilled off under reduced pressure, at the yield of above 90 %. The reaction time of 4 h was determined according to Pokorný and Dubská (1986). The ratio of reactants was 1.1:1 (alkanolamine:FFA). As suggested in Scheme 1, up to three product species (**3/MEAD**, **4** and **5**) with supposed yields are considered to be formed in this type of reaction (Pokorný and Dubská, 1986). The reaction was carried out without a catalyst under constant stirring at significantly reduced pressure of 2 kPa at 160 °C. During the reaction, the decrease in the acid number was monitored, when the acid number did not decrease significantly any further, the reaction was considered complete. The final acid number was 0.7 mg KOH/g. The sample was finally purified by distillation on a molecular evaporator at 220 °C and 30 Pa. The resulting product was a distillate with the acid number of 0.9 mg KOH/g. The obtained product was solid at room temperature and insoluble in water.

Diethanolamides were prepared in a similar manner as MEAD according to Scheme 1 (see product **8/DEAD**). DEAD are produced by reacting FFA or FAME with DEA in a molar ratio of 1:2 at 150—170 °C. During the reaction, water or MeOH is removed and the acid number decreases. When the acid number drops to 5 mg KOH/g or lower, the reaction is considered to be complete (Hui, 1998).

The molar ratio of reactants was 2:1 (alkanolamine:FFA) and the reaction temperature was 150 °C at 2 kPa. The whole synthesis needed 150 minutes to complete. The sample was finally processed on a molecular evaporator at 180 °C and 20 Pa so that the front part was evaporated and the distillation residue was the product. The obtained DEAD had an acid number of 0.3 mg KOH/g and it was obtained as viscous liquid soluble in water.

To test the prepared compounds as additives, neat diesel fuel obtained from VÚRUP, Bratislava was used; the corresponding basic physicochemical properties are summarised in Tab. 1. This diesel fuel was mixed with synthesised MEAD and DEAD samples at 0.2, 0.5, 0.8, 1.4 and 2.0 wt. %, respectively. Homogenisation of the mixtures was carried out under constant stirring and at elevated temperature (40 °C).

Physicochemical measurement

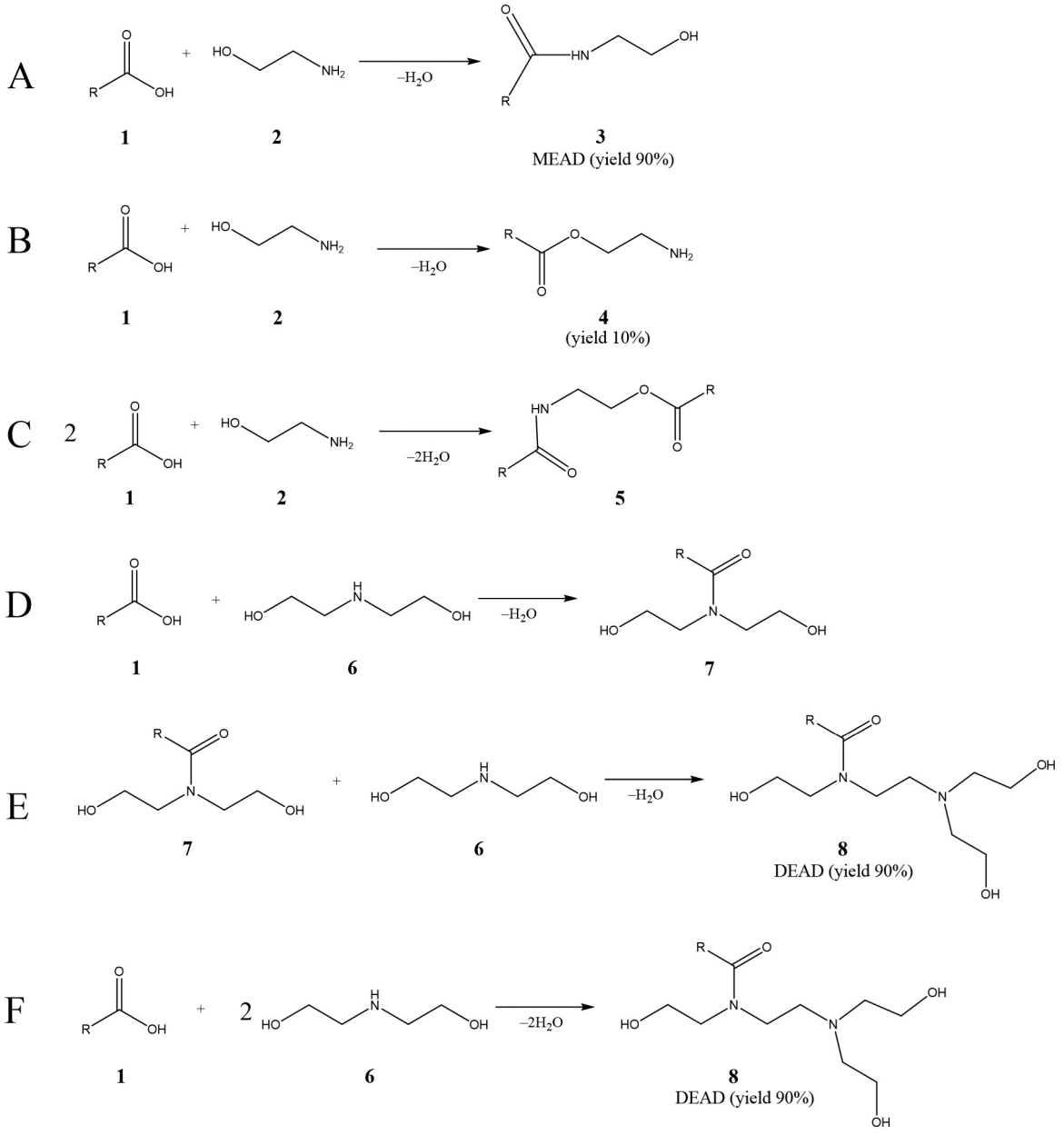
The prepared samples of MEAD and DEAD were tested as foam reducing additives, lubricity additives, and cetane number enhancers. The physicochemical

Tab. 2. Test methods and standards used to measure MEAD and DEAD samples.

Property	Applied test method
Foaming of diesel fuel	AFNOR NF MO-075
Lubricity	STN EN ISO 12156-1
Cetane number	STN EN ISO 5165
Acid number	STN EN ISO 660

Tab. 1. Physicochemical properties of nonadditive diesel fuel used for testing.

Property			
Density at 15 °C, kg/m^3	840.5	Distillation temperatures, °C	
Viscosity at 40 °C, mm^2/s	2.681	start of distillation	198.8
Sulphur content, mg/kg	6.2	10 vol. %	222.6
Cetane index (STN ISO 4264)	51.16	30 vol. %	242.5
Flash Point, °C	75	50 vol. %	266.3
Water content, mg/kg	36.7	70 vol. %	293.0
Ash content, wt. %	< 0.001	90 vol. %	329.4
Carbon residue, wt. %	< 0.01	95 vol. %	346.1
Total contamination, mg/kg	0.84	end of distillation	358.8
Copper strip corrosion (3 hr at 50 °C)	1a	Aromatic content, wt. %	31.50
Oxidation stability, g/m^3	0.0	— monoaromatics, wt. %	24.90
CLOUD point, °C	–8	— diaromatics, wt. %	5.90
Filterability, °C	–8	— triaromatics and higher, wt. %	0.7
Lubricity (wsd1.4), without additive	596	Polycyclic aromatic hydrocarbons, wt. %	6.6
Cetane number	49.9		



Scheme 1. Condensation reactions to synthesise MEAD and DEAD samples. R symbolised methyl, ethyl, and propyl group. Estimated yields are in parentheses (according to Pokorný and Dubská (1986)).

methods used are summarised in Tab. 2, where the technical standards and standardised procedures used are indicated (NACE TM0172-2015; Slovak Office of Standards, Metrology and Testing).

Quantum-chemical calculations

All calculations were performed using the Gaussian 16 program package (Frisch et al., 2016). The density functional theory with hybrid meta exchange-correlation functional M06-2X (Zhao and Truhlar, 2008) combined with the 6-311++G(d,p) basis set (Hariharan and Pople, 1973) was employed. The calculations performed at this theoretical level offer reliable description of thermodynamics of homolytic or heterolytic decay of N—H (Kováčová et al., 2023), O—H (Michalík et al., 2019), S—H (Rimarčík et al., 2011), C—CH (Lukeš and Hartmann, 2021), or C—CH_3 (Kleinová et al., 2024) bonds.

Optimisation of geometric structures was performed applying tight optimisation criteria and ultrafine DFT integration grid. True minima on the potential energy surface were confirmed through vibrational analysis (ensuring no imaginary frequencies are present). Thermodynamics of individual reactions were evaluated using reaction enthalpies for 298.15 K. Therefore, equilibria of studied reactions were not be discussed; reaction Gibbs free energy is beyond of the scope this publication (Atkins and Paula, 2006). Optimal geometries were visualised using the open-source Jmol viewer (Fig. 2) (Jmol

development team, 2016). Statical calculations were performed using exponential fit in OriginPro 2016 (OriginPro, 2016).

Results and Discussion

Theoretical calculations

Quantum-chemical calculations of the studied molecules were performed for gas-phase. Methyl- (Me), ethyl- (Et), and propyl- (Pr) groups (Scheme 1) were used as model alkyl chain (R) with expected induction effect along simple C—C bonds (Clayden, Greeves and Warren, 2012). Reactions A, B, C and F were mentioned according to (Pokorný and Dubská, 1986). Reactions D and E were suggested by authors for better understanding of the thermodynamics of investigated reaction schemes.

Since the studied molecules can have a larger number of conformations, only conformations with the occurrence of maximal number of intramolecular hydrogen bonds were considered. From a thermodynamic point of view, these conforma-

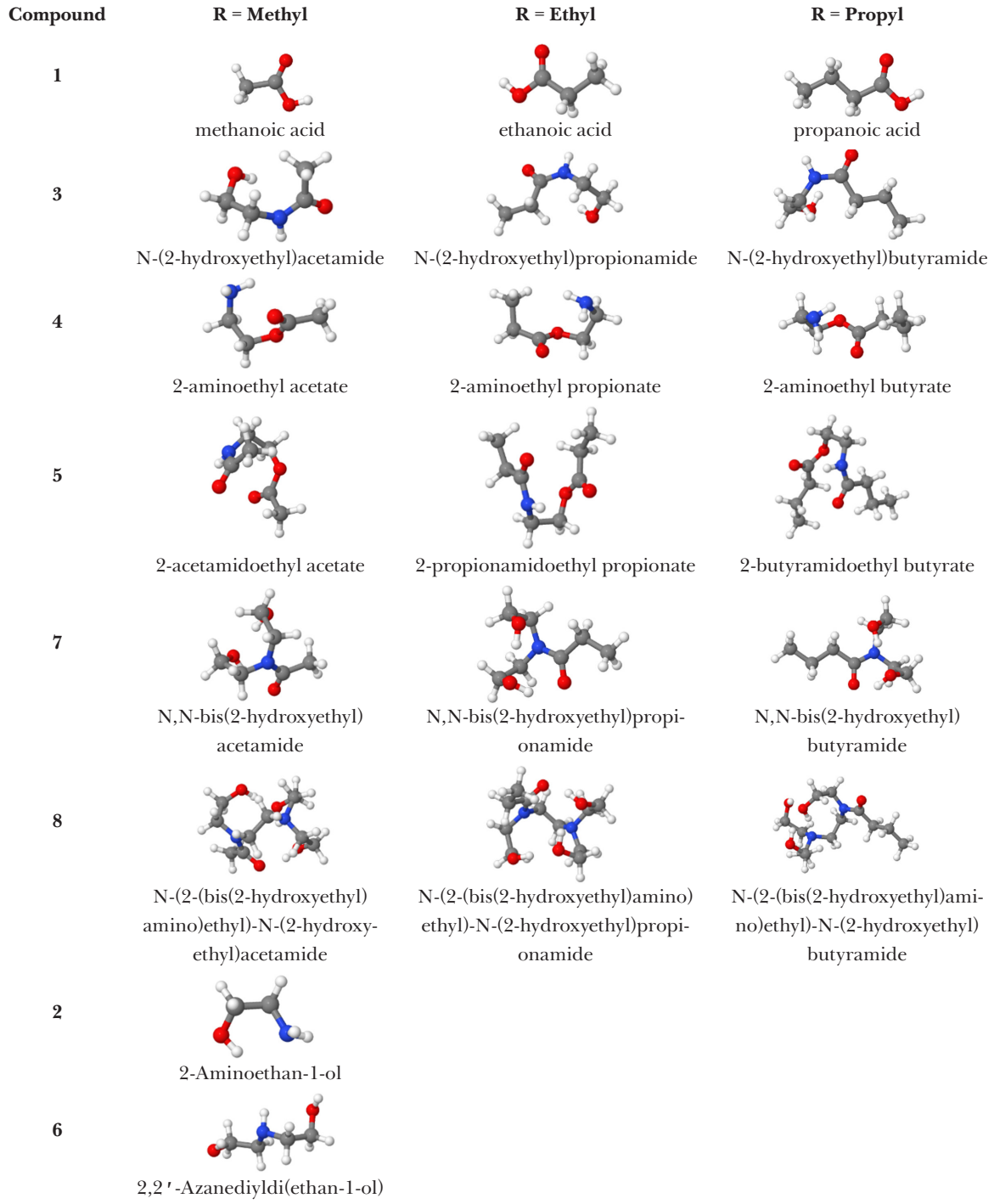


Fig. 1. Optimal geometries of studied molecules.

tions are the most energetically preferred and the most stable. This approach has been successfully used by Štellerová et al. (2022). Similarly to cyclic non-aromatic compounds, the presence of hydroxyl groups sterically modulates the geometry of organic molecule conformations (Rottmannová et al., 2013) and thus, *all*-trans conformer was considered for the propyl chain. Optimised geometries of studied organic molecules for individual substitutions of the alkyl group are shown in Fig. 1. Bond breakage in a molecule is caused by the formation of an intramolecular hydrogen bond between N—H and C—H. This phenomenon can be observed, for example, in case of MEAD, DEAD. It is typically the case if one intramolecular hydrogen bond exists within the molecules under study; however, in instances such as the product of reaction E (propyl group), multiple intramolecular hydrogen bonds are present.

Reaction enthalpies ($\Delta_r H$) were calculated as subtraction between calculated enthalpies of products and reactants with the corresponding reaction scheme (Scheme 1) in mind. As it is shown in Fig. 1, A and C reactions are considered as exothermic processes from the energetical point of view. Reaction enthalpies calculated for reaction A are –6 kJ · mol^{-1} for Me group and –4 kJ · mol^{-1} for Et and Pr groups. Similarly, substituent energies for Et and Pr groups for reaction C are comparable, i.e., prolongation of alkyl chain has only an insignificant effect on reaction thermodynamics. Reaction enthalpy for the largest model molecule with the Pr group was –19 kJ · mol^{-1} for reaction C. Reaction enthalpy of B reaction, 1 kJ · mol^{-1}, identifies weak endothermic process for the Et substituent group.

Exothermic character of chemical reactions is also predicted when reaction enthalpies for DEAD formation are considered. Higher reaction enthalpies were found for reactions A through C. It is worth to note that alkyl chain has more significant effect in E and F reactions. Elongation of the alkyl chain decreases the reaction enthalpy values by 2 kJ · mol^{-1} for hypothetical D reaction, 12 kJ · mol^{-1} for E reaction, and 13 kJ · mol^{-1} for F reaction compared to the Me group. The F reaction is the most preferred one for direct DEAD formation. Additional branching of the chain will possibly be energetically preferred as well. Energetical difference between E and F reactions is approximately 40 kJ · mol^{-1}.

Physicochemical measurements

The use of diesel fuel with very low sulphur content is associated with the need to adjust lubricity to avoid damage to selected aggregates and engine parts. Lubricity itself was expressed in this measurement as *WS1.4* in μm.

Lubricity of diesel fuel and MEAD and DEAD samples in diesel fuel was determined according to the High-Frequency Reciprocating Rig (HFRR) on a PCS Instruments HFR2 Specimens device according to STN EN ISO 12 156-1. The test principle is as follows: the fuel sample to be tested is placed in a small test vessel tempered at 60 °C. A fixed steel ball is clamped in a vertically mounted fixture and is pressed against a horizontally mounted static plate (plate) by means of an applied load. The test ball oscillates at a constant frequency at a constant stroke and the contact surface of the ball and the plate is completely immersed in a small vessel of the test device filled with the test sample. Metallurgical quality of the plate and ball as well as the test con-

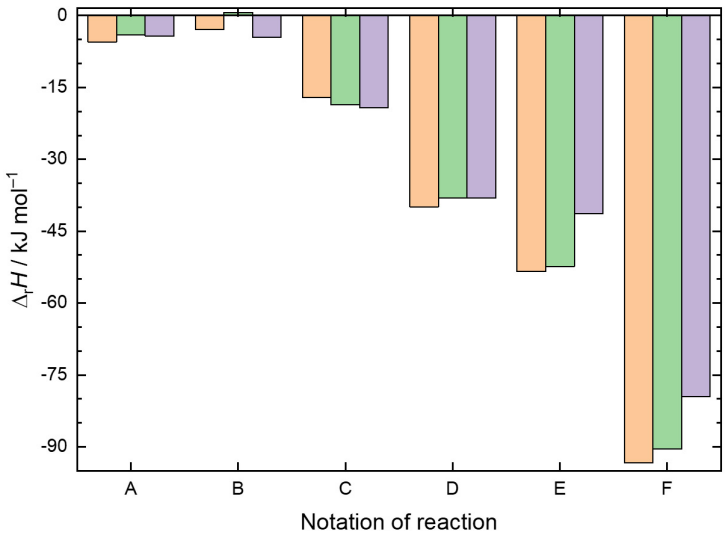


Fig. 2. Comparison of reaction enthalpies for studied types of reactions in dependence on alkyl chain length. Notation: methyl- (orange), ethyl- (green) and propyl- group (purple).

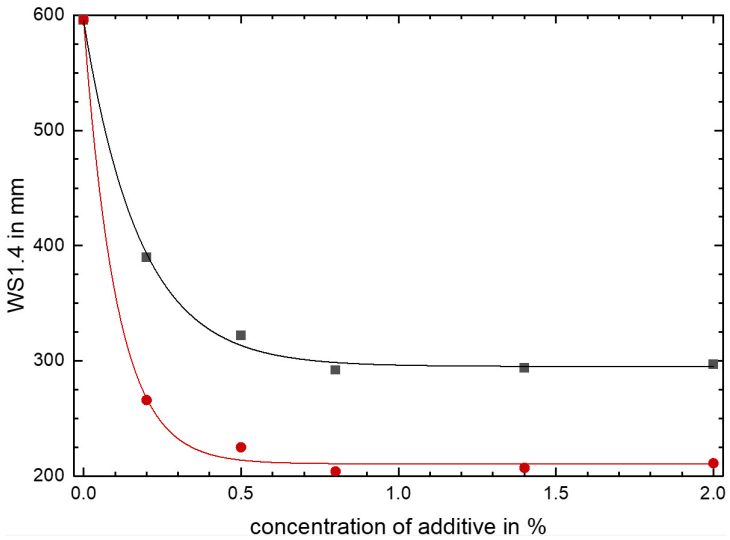


Fig. 3. Dependence of lubricity on the percentage concentration (c) of additive MEAD and DEAD.

ditions (fuel temperature, ball weight, frequency, stroke size, and test time) are specified in the standard. Diameter of the abrasion mark (MWSD value) after the test is read using a microscope and it is corrected depending on the temperature and humidity at the beginning and end of the test. The corrected diameter (WS value of 1.4) is a measure of the lubricity of the tested sample and the estimated uncertainty for this experiment is 2 %.

The tests showed that the synthesised MEAD and DEAD affect the lubricity; the dependences of lubricity on concentration (*c*) are shown in Fig. 3.

The obtained dependences satisfy the form of an exponential function with parameters of the linearised function:

$$WS1.4/\mu m = y_0 + A \times \exp(B \times c) \quad (1)$$

where c denotes the percentage concentration of the additive. Regression parameters y_0, A, B, and coefficient of determination (R^2) obtained from regressions are summarised in Tab. 3. Therefore, if R^2 is above 0.9980, the function correctly describes the concentration dependence of lubricity.

Tab. 3. Regression parameter of Eq. 1 and coefficient of determination for MEAD and DEAD.

	MEAD	**DEAD**
y_0	295.2 ±3.8	210.6 ±4.1
A	300.3 ±7.5	385.3 ±8.9
B	–5.60 ±0.38	–9.56 ±0.78
R^2	0.9981	0.9984

Comparison of the curves shows that the measured value of WS1.4 decreased significantly below the limit required by the Slovak Technical Standard STN EN 590, this condition was met already at the lowest investigated additive concentration. DEAD alone has a 10 % higher efficiency than MEAD. The estimated limit value is 210.6 μm.

Cetane number (CN) of diesel fuel and MEAD and DEAD samples in diesel fuel were determined according to STN EN ISO 5165 on a Waukesha test engine (STN EN ISO 5165). Results of the measurements are presented in Tab. 4, showing no positive effect of amides on the CN increase and the CN

Tab. 4. Effect of MEAD and DEAD addition (in wt. %) on the antifoaming properties of diesel fuel and on the cetane number of diesel fuel.

Percentage	0.0	0.2	0.5	0.8	2.0
MEAD					
cetane number	49.9	49.5	48.7	48.4	48.8
Foam break-up time/s	28.5	2.87	4.4	4.3	9.5
Foam height/mL	86	16	25	30	35
DEAD					
cetane number	49.9	49.9	49.9	49.5	48.3
Foam break-up time/s	28.5	43	19	22	31
Foam height/mL	86	86	75	77	83

even decreases with the increasing concentration. This fact was verified at a concentration of 20 wt. % of MEAD when the cetane number decreased to 47.3. Therefore, the estimated uncertainty for measured cetane number is ±0.2, the presented experiment indicates minimal effect of the studied additives on CN.

Antifoaming properties of the synthesised MEAD and DEAD additive samples were evaluated according to the French standard AFNOR NF M 07-075 (*Détermination de la tendance au moussage de gazoles*). In this test, a quantity of diesel fuel is injected at constant pressure into a calibrated glass measuring cup. The volume of foam produced, and the time taken for the foam to break up are then measured. Results of the effect of MEAD and DEAD addition on the antifoaming properties are summarised in Tab. 4. The MEAD sample was found to have very good antifoaming properties. The DEAD sample failed in this test because it had high foaming characteristics. The values for DEAD ranged from 75 to 86, with similar values achieved for 0.2 % and 2.0 %. MEAD addition showed better antifoaming properties compared to DEAD. The estimated uncertainty for measured antifoaming properties is 3 %.

Conclusions

This work was devoted to theoretical and experimental study of MEAD and DEAD samples. Thermodynamics of the relevant reactions leading to the production of the studied substances and by-products indicate exothermic nature of the reactions. The influence of the increase in the alkyl chain was demonstrated in case of DEAD formation. The syntheses of MEAD and DEAD were done employing the internal synthetical treatment. The final product can be separated on a molecular evaporator by distillation. Preliminary experimental characteristics of diesel fuel with MEAD or DEAD addition showed only a negligible positive effect on the cetane number. Lubricity properties of the samples were observed even at low MEAD and DEAD concentrations. While the MEAD sample showed good antifoaming properties, the DEAD sample did not have a significant effect on the antifoam properties.

It can be concluded that the addition of MEAD (0.2 wt. %) to diesel fuel is perspective for the preparation of novel types of fuels. Nevertheless, complex experimental characterisation of prepared samples should be presented in a systematic form.

Acknowledgement

This paper was supported by the Ministry of Education, Science, Research and Sport of the Slovak Republic, project VEGA 1/0461/21. We are grateful to the HPC centre at the Slovak University of Technology in Bratislava, which is a part of the Slovak Infrastructure of High-Performance Computing (SIVVP project, ITMS code 26230120002, funded by the European region development funds, ERDF) for the computational time and resources made available.

References

Atkins P, Paula J de (2006) Physical Chemistry 1072.

Clayden J, Greeves N, Warren S (2012) Org. Chem.

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr., Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Keith T, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS,Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J and Fox DJ (2009) Gaussian 09, Revision D.01, Gaussian, Inc. Wallingford CT.

Gaurav N, Sivasankari S, Kiran GS, Ninawe A, Selvin J (2017) Renew. Sustain. Energy Rev. 73 (September 2016): 205–214.

Hariharan PC, Pople JA (1973) Theor. Chim. Acta. 28(3): 213–222.

Hui YH (1998) Fatty Acids and Derivatives from Coconut Oil. In: Bailey's Industrial Oil and Fat Products. 5th Edn. Vol. 5. Ed. Y.H. Hui, J. Wiley & Sons, New York (USA), pp. 74–78.

Jmol development team. (2016) Jmol: an open-source Java viewer for chemical structures in 3D (14.6). http://jmol.sourceforge.net/.

Kleinová A, Biela M, Lukeš V, Klein E (2024) J. Mol. Struct. 1303: 137646.

Kleinová A, Cvengrošová Z, Cvengroš J (2013) Fuel. 106: 749–756.

Kováčová A, Michalík M, Hartmann H, Lukeš V (2023) J. Mol. Liq. 389: 122811.

Liu Y, Cruz-Morales P, Zargar A, Belcher MS, Pang B, et al. (2021) Cell. 184(6): 1636–1647.

Lukeš V, Hartmann H (2021) Color. Technol. 137(4): 389–398.

Michalík M, Poliak P, Lukeš V, Klein E (2019) Phytochemistry. 166(July): 112077.

Mikulec J, Cvengroš J, Joríková Ľ, Banič M, Kleinová A 2010 J. Clean. Prod. 18(9): 917–926.

NACE TM0172-2015. Determining Corrosive Properties of Insoluble Petroleum Product Pipeline Cargoes. 2015.

OriginPro, Version 2016. OriginLab Corporation, Northampton, MA, USA.

Pokorný J, Dubská L (1986) Technologie tuků Praha: Státní nakladatelství technické literatury. 450.
Rimarčík J, Lukeš V, Klein E, Rottmannová L (2011) Comput. Theor. Chem. 967(2—3): 273—283.
Roeder J et al. (2002) EP 1 227 143 Fuel additives.
Rottmannová L, Lukeš V, Ilčin M, Fodran P, Herich P, et al. (2013) J. Mol. Struct. 1049: 494—501.
Sanders HL (1958) J. Am. Oil Chem. Soc. 35(10): 548—551.
Shahidi F (2005) Bailey's industrial oil and fat products volume 5 sixth Wiley&sons.
Slovak Office of Standards, Metrology and Testing https://www.normoff.gov.sk/.
Štellerová D, Michalík M, Lukeš V (2022) Phytochemistry. 203: 113387.
STN EN ISO 5165. Petroleum products — Determination of the ignition quality of diesel fuels — Cetane engine method.
Zhao Y, Truhlar DG (2008) Theor. Chem. Acc. 120(1—3): 215—241.