Response Equation Based Thermochemical Analysis of Singlet Bipolaron Structures in Oligo(3-Methyl-Thiophenes)

Eufrozina A. Hoffmann[1], Zoltan A. Fekete[2], Csaba Visy[1] and Tamás Körtvélyesi*,[1,2]

[1]*Department of Physical Chemistry, University of Szeged, H-6720 Szeged, Rerrich ter 1., Hungary*

[2]*HPC group, University of Szeged, H-6725 Szeged, Szikra u. 2., Hungary*

Abstract: A thermochemical investigation of singlet dication (bipolaron) structures in a set of 3-methyl-thiophene oligomers (up to the length of 24-mers) was carried out. Equlibrium structures and heats of formation were calculated with the semiempirical quantum chemical method RHF/PM3 for the doubly oxidized forms. All-*cisoid* planar structural variants of the form $A_{n1}Q_{n2}A_{n3}$ were considered, where A and Q are thiophene monomer units in aromatic and quinoidal chain segments, respectively. The types of segments were determined from their bond-length pattern. At this level of theory, diverse structures combining both types of segments were found stable.

The so-called "response equation based quantitative structure-property relationship" (REQ-QSPR) method, recently proposed by Fishtik *et al*. was used to perform and interpret regression data analysis. The REQ-based approach rationalizes mathematical treatment of stoichiometric equations. It provides a simple concept for breaking down a dataset for heat of formation into linear combination of terms that are easy to interpret in the framework of chemical groups. REQ-QSPR includes arbitrary structural descriptors rather than being restricted to using only the number of chemical constituents. Applied to our system, it is demonstrated that modelling nonlinear effects becomes possible as well. Supra-linear dependence of the energy with respect to the bipolaron localization length is indicated by REQ-QSPR, and detailed analysis suggests specific further calculations to carry out for improving the model.

Keywords: Bipolaron model, Response equation based quantitative structure-property relationship (REQ-QSPR).

1. INTRODUCTION

Electronically conducting polymers are a remarkable class of advanced materials, since under certain conditions they have a small electric resistance like metals, while can be transformed into an insulating state as well [1-8]. The changes in the electrical conductivity and in the coupled physical properties upon the redox transformation of the conjugated polymers have been generally interpreted on the basis of the polaron/bipolaron model [6, 9-25]. Despite the tremendous research done on the electronic properties of these molecules, however, less attention has been paid to the relative thermodynamic stability of their various possible structures [26, 27].

Polythiophenes, especially their 3-alkyl substituted derivatives (P3ATs) form a practically important family of these materials [28-32], because they can be synthesized and processed easily, and their doped states are stable in air [1, 4, 6, 31, 33-41].

We investigated a set of 3-methyl-thiophene oligomer dications (up to the length of 24-mers) (O3MT) as simple P3AT models. Their general molecular structure is a quinoidal region sandwiched between flanking aromatic parts on either side; Scheme **1** outlines the bonding pattern with a simple Kekule-like diagram. One should keep in mind that in reality the bond orders are both intermediate (between single

and double), and variable depending on the degree of aromatic/quinoidal nature. The specific choice for structures to be considered is motivated by our goal of uncovering stability trends related to quinoidal distortion upon bipolaron formation [26, 42, 43].
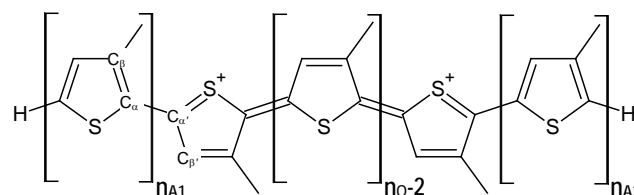
Scheme 1. Generic structure of $A_{nA1}Q_{nQ}A_{nA2}$ bipolaronic species considered.

Semiempirical quantum chemical PM3 calculations [44-46] were applied for this theoretical thermochemical analysis. The relatively low-level method [46-48] was chosen so that a large number of long-chain structures can be included in this exploratory study and its planned follow-ups; using the long-established PM3 Hamiltonian also facilitates comparison with earlier studies [1, 10, 47]. Our current report serves as a proof-of-concept for the response equation based quantitative structure-property relationship (REQ-QSPR) protocol presented, justifying the use of higher-level quantum chemical methods in refining this simple model developed. Besides this principal goal, the results also demonstrate that – at least at the RHF/PM3 level of theory - relatively short P3ATs already allow, by spontaneous change of their

*Address correspondence to this author at the Department of Physical Chemistry, University of Szeged, H-6720 Szeged, Rerrich ter 1., Hungary; E-mail: kortve@chem.u-szeged.hu

bond-length alteration pattern, to develop aromatic-quinoidal transitions along the chain [17-19]. Therefore, these oligomers are capable of supporting diverse stable structures with both aromatic and quinoidal sections.

2. METHODOLOGICAL AND COMPUTATIONAL ASPECTS

2.1. Theory of Response Equations

The so-called response reactions (RERs) were introduced by Fishtik *et al.* [49-52], forming the basis for an interpretative concept in thermochemistry. RERs were defined as uniquely determined chemical reactions based on the stoichiometric properties of systems with multiple equilibria. In these systems there may occur processes that are difficult to interpret without the aid of RERs [52-55]. They were also found useful to rationalize extracting thermochemical data from quantum chemical calculations [56-58].

Later on the RER concept has been extended to cover generalized equations, *i.e.* those analogous to chemical reactions but involving abstract species rather than actual chemical entities. Although the terminology of Fishtik [59, 60] has retained the name RER for these, we prefer to avoid confusion and will switch to the term "response equations" (REQ) from now on (unless the equations are actually ordinary chemical reactions in the non-generalized case). Full details of REQ-based methodologies for treating simple group-additivity problems [56-58, 60], as well as general quantitative structure-properties relationships [59] have already been published. Therefore only the principal ideas necessary to follow our analysis are summarized here for easy reference.

The starting point is expressing conventional chemical reactions in terms of linear algebra. They simply are those linear combinations of the B_i chemical species which preserve mass-balance; that is, they conserve the type and number of components (which in the chemical practice are either atoms or simple groups that constitute the species in question). It is well-known that such stoichiometric equations can only be unequivocal (*i.e.* not expressible as linear combinations of simpler ones) if they contain no more than $q + 1$ species, in a system with q components. RERs by definition are reactions with at most $q + 1$ species, and are mathematically the most elementary building blocks of a given stoichiometric system [60, 61].

Fishtik has pointed out that handling a system of QSPR equations can be carried out in complete analogy with the stoichiometric systems [59]. Here, instead of dealing with q components in the chemical sense, one has certain values for a set of q descriptors assigned to each species. Introducing a "descriptor-balance", *i.e.* requiring that linear combinations of the B_i chemical species conserve the type and number of descriptors, leads to the definition of *isostructural equations* – natural generalization of the stoichiometric equations. With these one can then speak of *isostructural REQs* (ISREQ): those with at most $q + 2$ species; this number is higher than $q + 1$ by one, because typical QSPR regressions involve an extra degree of freedom (the intercept), to which the unit vector belongs as an extra "descriptor", besides those counted by q. Except this formal difference (which is merely a matter of marking the degrees of freedom with a different label, $(q - 1)$ *vs.* q), ISREQs for systems of QSPR equations are perfect analogues of RERs for stoichiometric systems.

Similarly to what had been demonstrated for the application of RERs, partitioning gross quantities into their elementary contributions provides deeper insight into the underlying data when analyzing with ISREQs. Due to the mathematical foundation of defining them, they provide a detailed picture for the distribution of errors in QSPR fitting. This is a crucial aid in detecting outliers, which is a non-trivial task considering that mere evaluation of the residuals is often not sufficient for identifying them [58, 59].

2.2. Quantum Chemical Calculations

Semiempirical quantum mechanical calculations (PM3 [46]) were carried out with the PC-GAMESS program [44, 45]. Manipulation of the input and output data was facilitated by custom made scripts, which we release to the public[†].

The aim of this study was not to decide the multiplicity (*i.e.* determine relative stability of singlet bipolaron *vs.* triplet two-polaron structures), rather to analyze the thermochemistry of possible singlet structures. Therefore the doubly-charged molecular ions were forced to have singlet multiplicity, and spin-restricted Hartree-Fock calculations (RHF) were done.

The starting point was a fully optimized regioregular 24-mer chain, with head-to-tail tacticity. An all-*cisoid* planar conformation was enforced by fixing inter-ring linking dihedral angles ($C_\beta C_\alpha C_{\alpha'} C_{\beta'}$, see Scheme **1**) at 180 degrees. In some cases the location of the quinoidal segment was directed to select positions along the chain, by setting the length of an appropriately located inter-ring bond, C_α-$C_{\alpha'}$, to the optimized value obtained for the purely aromatic-aromatic intrachain linkage (144.06 pm).

Fixing one such bond at both sides of the segment made it possible to model quinoidal regions [26, 42, 43] with various artificially shortened lengths. (We note that our preliminary results – not yet included in this report here - also indicate to feasibility of modeling lengthened regions as well, by fixing at appropriately spaced positions the delimiting bonds with a value around 141 pm, corresponding to the intermediate aromatic-quinoidal linkages). Calculations of shorter chains proceeded from the geometry obtained with truncating the 24-mer and reoptimizing.

The bulk of our calculations was carried out on detrended data: $\Delta_f^{dt} H_i = \Delta_f H_i - (m_{dt} n_{\text{Chain}} + b_{dt})$, where m_{dt} and b_{dt} are the slope and intercept of linear regression of $\Delta_f H_i$ versus n_{Chain} (calculated before the rest of the calculations for all $\Delta_f H_i$, not excluding those points dropped from the later fits). The reason for this transformation is purely practical: it was found to eliminate floating-point error accumulation problems that were encountered in our scripts when working on large datasets without the detrending.

2.3. Data Processing with REQ-QSPR

First the outputs for the optimized geometries were scanned to extract their C_α-$C_{\alpha'}$ inter-ring linking bond lengths. See Fig. (**1**) for an illustrative example of the linkage length variation along the chains. Also, at this stage we found it
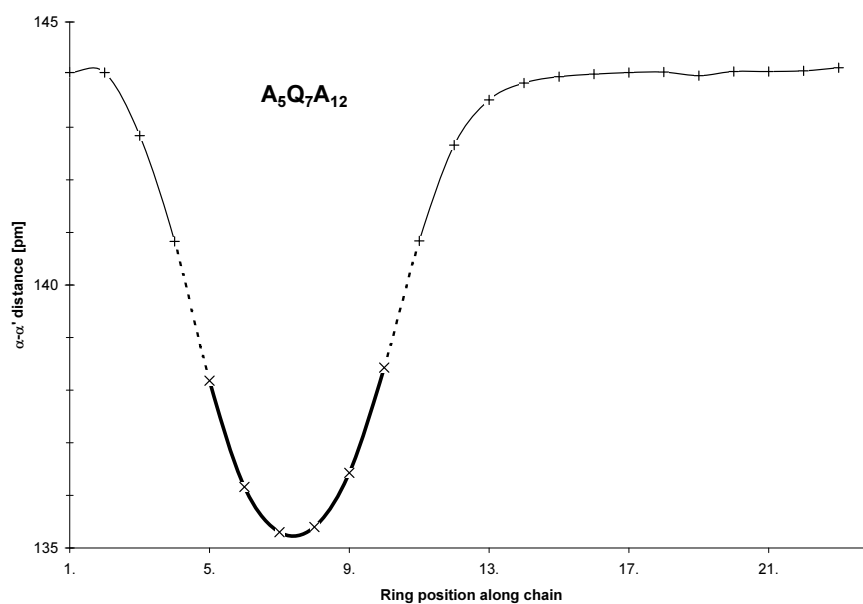
Fig. (1). Example of α-α' linking bond length variation: $A_5Q_7A_{12}$ species; quinoidal segment: heavy line with slanted crosses, aromatic segments: light line with straight crosses, intermediate linkage: dotted line (see text for details).

very useful to monitor the progress of calculations *via* quick visual display of the optimized structures. For this purpose our scripts produce PDB molecular coordinate files (released as Supporting Material), which have embedded color-coding to emphasize variation in these distances. The rings were then classified as either aromatic or quinoidal (considering 141 pm to be the characteristic distance separating aromatic and quinoidal segments).

Since the quinoidal segment is located in the middle of the chain, this assignment provides 3 principal descriptors for each structure investigated: n_Q, n_{Ashort} and n_{Along}, the numbers of rings in the quinoidal, shorter and longer aromatic segments respectively. The structural designations of the 30 species included in this study are shown in Table **1**, along with their total chain length (in units of thiophene rings) and PM3 heats of formation data. A 2-D display of the distribution of n_Q and n_{Ashort} values is presented in Fig. (**2**).

Table 1.   Bipolaronic Oligo(3-methyl-thiophene) Structures, their Calculated (PM3) Heats of Formation Data, and Descriptors to be Used for REQ-QSPR

#	Species	$\Delta_f H$ kcal/mol	n_{Chain}	$\Delta_f{}^{dt}H^1$ kcal/mol	n_{Ashort}	n_Q	n_{Along}	n_{Atotal}	n_Q^2	2^{n_Q-3}
1.	$A_6Q_3A_6$	797.77	15	3.42	6	3	6	12	9	1
2.	$A_5Q_4A_6$	796.01	15	1.66	5	4	6	11	16	2
3.	$A_1Q_5A_9$	797.40	15	3.05	1	5	9	10	25	4
4.	$A_4Q_7A_4$	792.69	15	-1.66	4	7	4	8	49	16
5.	$A_4Q_7A_5$	821.16	16	-1.59	4	7	5	9	49	16
6.	$A_4Q_7A_6$	849.66	17	-1.49	4	7	6	10	49	16
7.	$A_3Q_6A_9$	879.81	18	0.26	3	6	9	12	36	8
8.	$A_5Q_7A_6$	878.34	18	-1.21	5	7	6	11	49	16
9.	$A_4Q_7A_7$	878.45	18	-1.10	4	7	7	11	49	16
10.	$A_3Q_7A_8$	878.60	18	-0.95	3	7	8	11	49	16

(Table 1). Contd…..

#	Species	$\Delta_f H$ kcal/mol	n_{Chain}	$\Delta_f^{dt}H^1$ kcal/mol	n_{Ashort}	n_Q	n_{Along}	n_{Atotal}	n_Q^2	2^{n_Q-3}
11.	$A_4Q_8A_6$	878.20	18	-1.35	4	8	6	10	64	32
12.	$A_4Q_7A_8$	906.72	19	-1.22	4	7	8	12	49	16
13.	$A_4Q_7A_9$	935.27	20	-1.07	4	7	9	13	49	16
14.	$A_8Q_4A_9$	967.30	21	2.56	8	4	9	17	16	2
15.	$A_7Q_6A_8$	964.97	21	0.23	7	6	8	15	36	8
16.	$A_7Q_7A_7$	964.18	21	-0.56	7	7	7	14	49	16
17.	$A_4Q_7A_{10}$	963.81	21	-0.93	4	7	10	14	49	16
18.	$A_4Q_7A_{11}$	992.36	22	-0.77	4	7	11	15	49	16
19.	$A_4Q_8A_{11}$	1020.92	23	-0.61	4	8	11	15	64	32
20.	$A_{10}Q_4A_{10}$	1052.98	24	3.05	10	4	10	20	16	2
21.	$A_9Q_5A_{10}$	1051.60	24	1.67	9	5	10	19	25	4
22.	$A_9Q_6A_9$	1050.65	24	0.72	9	6	9	18	36	8
23.	$A_8Q_7A_9$	1049.74	24	-0.19	8	7	9	17	49	16
24.	$A_7Q_7A_{10}$	1049.67	24	-0.26	7	7	10	17	49	16
25.	$A_6Q_7A_{11}$	1049.57	24	-0.36	6	7	11	17	49	16
26.	$A_5Q_7A_{12}$	1049.51	24	-0.42	5	7	12	17	49	16
27.	$A_4Q_7A_{13}$	1049.53	24	-0.40	4	7	13	17	49	16
28.	$A_3Q_7A_{14}$	1049.91	24	-0.02	3	7	14	17	49	16
29.	$A_8Q_8A_8$	1049.86	24	-0.07	8	8	8	16	64	32
30.	$A_3Q_8A_{13}$	1049.64	24	-0.29	3	8	13	16	64	32

[1]Detrended data: $\Delta_f^{dt}H_i = \Delta_f H_i - (m_{dt}n_{Chain} + b_{dt})$, m_{dt}=28.397, b_{dt}=368.40.

Next the descriptors used for the regression model were selected. So far we have investigated combinations of the following variables: n_Q, n_{Ashort}, n_{Along}, n_Q^2, 2^{n_Q-3}, $n_{Atotal} = n_{Ashort} + n_{Along}$, and $n_{Chain} = n_Q + n_{Atotal}$, but arbitrary others can be included if so desired. Once the variables are selected and the corresponding stoichiometric matrix is set up as input, our RER software[‡] (an earlier version of which has already been published) generates a listing of all the ISREQs [59].

The output obtained is processed further by our scripts implementing the REQ-QSPR equations [59] as follows. Formulas are presented below for a general dependent variable y_i, keeping in mind that in the present case $y_i \equiv \Delta_f H_i$, the heats of formation data for species indexed by i. The residuals are, by definition: $e_i = y_{-i} \hat{y}_i$, where $\hat{y}_i$ are the fitted values from the QSPR model. It should be emphasized, however, that the main result of the REQ-QSPR treatment is that given by Eq. 2: the residuals are calculated directly, without actually calculating the fitted values $\hat{y}_i$! Rather they

[‡]<http:// www.staff.u-szeged.hu/~fekete/req-polythiophene/>, and also as Supplementary Material at the publisher.
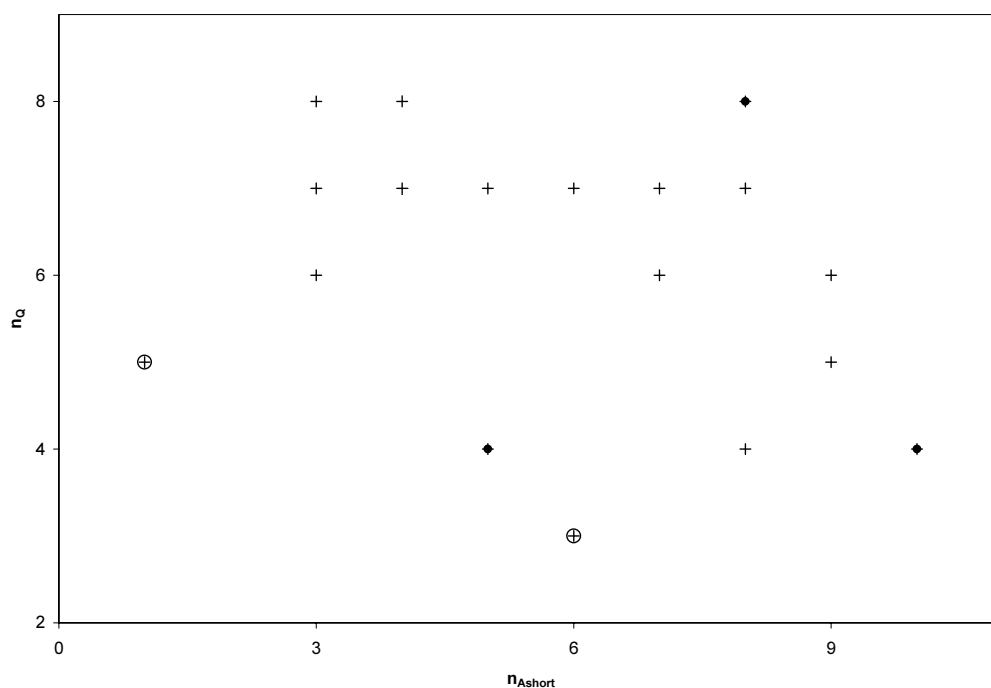
Fig. (2). Two-dimensional display of values for descriptor variables n_Q vs. n_{Ashort}; datapoints eliminated in model reduction are marked with large open circles (first stage), and small filled ones (second stage); see text for details.

are derived from the formal solution given in Eq. 36 of Ref. [59], from which follow the Eqs. 1-5 shown below.

For a model with n species and q descriptors (plus the constant term) the number of degrees of freedom is $m = n - q - 1$. Each ISREQ (indexed by k) has a change of property denoted by $\delta y(k)$ as:

$$\delta y(k) = \sum_j v_j(k) y_j \tag{1}$$

From these quantities and the stoichiometric coefficients in the ISREQs the residuals are calculated according to Eqs. 2-4.

$$e_i = \frac{1}{F} \sum_k v_i(k) \delta y(k); \qquad i = 1,\ldots,n \tag{2}$$

$$F = \frac{1}{m} \sum_k g(k) \tag{3}$$

$$g(k) = \sum_i v_i^2(k) \tag{4}$$

Moreover, an expression given in Eq.5 for individual ISREQ error $s(k)$ is evaluated:

$$s^2(k) = \frac{1}{mF^2} \sum_i (v_i(k) \delta y(k))^2 \tag{5}$$

In our experience this statistics has proven useful not only for characterizing individual ISREQs, but more importantly for suggesting possible outliers from among the data points. The following procedure was used: the list of ISREQs is first pruned to contain only those with $s(k)$ above some threshold value (chosen typically 2-3 orders of magnitude below the maximum); then from this list we collect the following partial error sum for each species index i:

$$M_i^s = \sum_{v_i(k) \neq 0} s(k) \tag{6}$$

That is, large values of M_i^s indicate those species whose overall contributions to the errors of ISREQs are large.

3. RESULTS AND DISCUSSION

As mentioned in Section 2.3, there are 3 principal descriptors for each structure investigated: n_Q, n_{Ashort} and n_{Along}. In addition to these, we generated the following trial variables: n_{Atotal}, the overall number of aromatic rings; n_Q^2, to capture quadratic dependence on the length of quinoidal segment; and the exponentially changing $2^{n_Q - 3}$, in order to include a variable to cover a steeper variation with this length. REQ-QSPR models with several combinations from this set of descriptors were tested. A summary of all runs from the first stage is presented in Table **2**, and an example of the evaluated residuals lists is shown in Table **3** for the model with n_Q.

It was noted that the consensus from the best fitting equations had indicated two species as most likely outliers: $A_6Q_3A_6$ and $A_1Q_5A_9$. (This can be justified on structural grounds: the former structure has an extremely short quinoidal section, while the latter a singular aromatic ring termination; therefore their features could only be adequately covered in a study with many more points.) The next round of analysis was then performed with omitting these two data points. Perusal of results from this stage (not shown) suggests that further removal of species indexed 2, 20 and 29 (*i.e.* the structures $A_5Q_4A_6$, $A_{10}Q_4A_{10}$ and $A_8Q_8A_8$) may significantly improve the fit, too. Therefore our final dataset is reduced to 25 data points, with REQ-QSPR models characterized by results listed in Table **4**.

Table 2.   **Summary of all REQ-QSPR Runs Performed at the First Stage**

Model label	s_y	$\max|e_i|$	$i:\max|e_i|$
A_t, A_s, Q_1, Q_2, Q_x	0.43	1.34	3
A_t, Q_1, Q_2, Q_x	0.46	1.70	3
A_t, A_s, Q_1, Q_x	0.48	1.79	3
A_t, Q_2, Q_x	0.47	1.97	3
A_t, Q_1, Q_x	0.51	2.17	3
A_t, A_s, Q_1, Q_2	0.53	1.92	3
A_t, Q_1, Q_2	0.55	2.28	3
A_s, Q_1, Q_2, Q_x	0.56	1.32	3
Q_1, Q_2, Q_x	0.56	1.16	3
A_t, A_s, Q_1	0.58	1.69	3
Q_2, Q_x	0.58	1.52	3
A_t, Q_1	0.60	2.04	3
A_s, Q_1, Q_x	0.63	1.92	3
Q_1, Q_x	0.62	1.69	3
A_t, A_s, Q_2	0.66	1.65	3
A_s, Q_1, Q_2	0.67	2.04	3
A_t, Q_2	0.67	1.95	3
Q_1, Q_2	0.67	1.76	3
A_s, Q_1	0.70	1.83	3
Q_1	0.69	1.55	3
A_s, Q_2	0.76	1.78	3
Q_2	0.76	1.57	29
AQ, x	1.04	2.41	3
Q_x	1.05	1.83	29
A_t, A_s	1.35	3.97	3
A_s	1.34	3.98	3
A_t	1.36	3.60	1
A_1	1.44	3.48	1

Table 3.   **Example of Evaluated Residuals for REQ-QSPR with Variable n_Q (Stage 1)**

#	Token	e_i	M_i^s
1.	$A_6Q_3A_6$	0.02	0.60
3.	$A_1Q_5A_9$	1.55	0.49
29.	$A_8Q_8A_8$	1.38	0.49
2.	$A_5Q_4A_6$	-0.73	0.44
30.	$A_3Q_8A_{13}$	1.16	0.44
20.	$A_{10}Q_4A_{10}$	0.61	0.42
19.	$A_4Q_8A_{11}$	0.85	0.37
14.	$A_8Q_4A_9$	0.14	0.36
4.	$A_4Q_7A_4$	-1.11	0.31
5.	$A_4Q_7A_5$	-1.04	0.29

#	Token	e_i	M_i^s
6.	$A_4Q_7A_6$	-0.94	0.27
11.	$A_4Q_8A_6$	0.14	0.27
8.	$A_5Q_7A_6$	-0.66	0.21
12.	$A_4Q_7A_8$	-0.68	0.21
21.	$A_9Q_5A_{10}$	0.23	0.21
9.	$A_4Q_7A_7$	-0.56	0.19
13.	$A_4Q_7A_9$	-0.53	0.19
28.	$A_3Q_7A_{14}$	0.48	0.19
10.	$A_3Q_7A_8$	-0.41	0.18
17.	$A_4Q_7A_{10}$	-0.40	0.17
23.	$A_8Q_7A_9$	0.32	0.17
16.	$A_7Q_7A_7$	-0.04	0.16
18.	$A_4Q_7A_{11}$	-0.25	0.16
24.	$A_7Q_7A_{10}$	0.25	0.16
25.	$A_6Q_7A_{11}$	0.15	0.16
26.	$A_5Q_7A_{12}$	0.10	0.16
27.	$A_4Q_7A_{13}$	0.12	0.16
7.	$A_3Q_6A_9$	-0.19	0.14
15.	$A_7Q_6A_8$	-0.22	0.14
22.	$A_9Q_6A_9$	0.25	0.14

Table 4.   Summary of all REQ-QSPR Runs Performed at the Third Stage

Model label	s_y	max$|e_i|$	i:max$|ei|$
A_t, Q_2, Q_x	0.18	0.47	7
A_t, Q_1, Q_x	0.19	0.54	7
A_t, Q_1, Q_2, Q_x	0.18	0.39	7
A_t, A_s, Q_1, Q_x	0.20	0.50	7
A_t, A_s, Q_1, Q_2, Q_x	0.18	0.31	7
A_t, Q_1, Q_2	0.23	0.63	7
A_t, A_s, Q_1, Q_2	0.23	0.59	7
A_t, Q_1	0.33	0.68	30
A_t, A_s, Q_1	0.33	0.64	30
A_t, Q_2	0.38	0.77	14
A_t, A_s, Q_2	0.39	0.76	14
A_s, Q_1, Q_2, Q_x	0.45	0.95	28
A_s, Q_1, Q_x	0.46	0.94	28
A_s, Q_1, Q_2	0.48	0.91	28
Q_2, Q_x	0.47	0.83	4
Q_1, Q_2, Q_x	0.47	0.79	4
Q_1, Q_x	0.49	0.85	4

(Table 4). Contd…..

Model label	s_y	max|e_i|	i:max|ei|
A_t, A_s	0.62	1.90	14
A_t	0.67	2.23	14
Q_1, Q_2	0.51	0.90	4
A_s, Q_1	0.55	1.26	30
Q_1	0.58	1.21	30
A_s, Q_2	0.59	1.34	30
Q_2	0.63	1.31	30
AQ, $_x$	0.71	1.60	14
A_s	0.74	1.93	14
A_1	0.90	2.87	14
Q_x	0.80	1.82	14
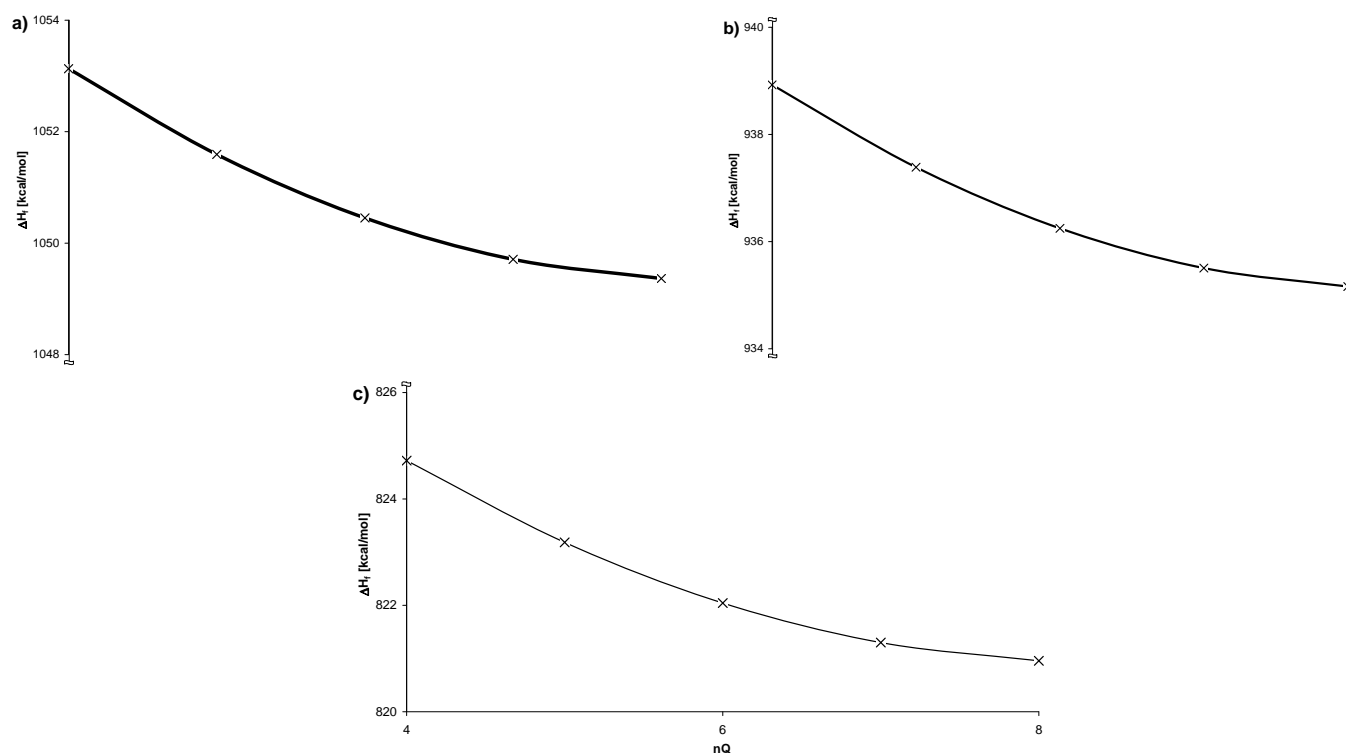
Fig. (3). Plot of $\Delta_f H$ vs. n_Q; **a**): n_{Total}=24, **b**): n_{Total}=20, **c**): n_{Total}=16.

One of the best models ($s_y = 0.23$) available at this stage is that with the set of 3 descriptors: n_{Atotal}, n_Q and n_Q^2; this happens to have a ready chemical interpretation, namely the first two terms correspond to linear group additivity values for the aromatic and quinoidal rings, respectively, and the quadratic coefficient is related to non-linearity. This fit, Eq.7 is our final result:

$$\frac{\Delta_f H_i}{[kcal / mol]} = (28.55 \pm 0.02) \cdot n_{Atotal} + (25.22 \pm 0.49) \cdot n_Q$$
$$+ (0.199 \pm 0.040) \cdot n_Q^2 + (378.0 \pm 1.6) \qquad (7)$$

With a mere 0.2 kcal/mol standard error of fitting, this expression describes well the trend of stability versus the descriptors, in the range covered: $8 \le n_{Atotal} \le 19$, $4 \le n_Q \le 8$ Fig. (**2**) depicts the distribution of all points considered). We note that analyzing the quadratic form of Eq. 7 for any fixed chain length, *i.e.* setting a constraint $n_{Atotal} + n_Q \equiv n_{Total}$, suggests a minimum for the heat of formation occurring slightly above $n_Q = 8$. Eq. 7 should not be used for extrapolating outside its interval of applicability, $4 \le n_Q \le 8$! The function of $\Delta_f H$ vs. n_Q is plotted on Fig. (**3**), for three different values of n_{Total}: 24, 20 and 16.

Naturally, we can improve the model by further calculations. First, considering the process of data point removal (see the omitted points marked on Fig. **2**), it is obvious that a

better description would be obtained if more points are available in areas insufficiently covered by this set of data. In particular, filling in the lower left corner of the field shown on Fig. (**2**) would give a firmer base for the regression; expanding the field over the top right corner would provide coverage for larger structures. Secondly, higher-level quantum chemical values are needed for more accurate heats of formation; careful selection of structures – which can be aided by the protocol exemplified here - is important in this respect for keeping the necessary computing resources under control.

For reason of brevity only a condensed summary of the evaluated results was presented above, more details of the model fitting are provided as Supplementary Material.[†] .In conclusion, it has been demonstrated that the REQ-QSPR treatment is a valuable tool for interpreting regression analysis. The model derived describes the enthalpies of formation for polythiophene oligomers (15-24 rings total chain length, and 4-8 rings quinoidal mid-section width) as simple functional dependence on the number of aromatic and quinoidal units.

ACKNOWLEDGEMENT

REFERENCES

[1]   Alhalasah, W.; Holze, R. Electrochemical materials science: tailoring intrinsically conducting polymers. The example: substituted thiophenes. *J. Solid State Electrochem.,* **2005**, *9*(12), 836-844.
[2]   Heeger, A. J. Nobel lecture: semiconducting and metallic polymers: the fourth generation of polymeric materials. *Rev. Mod. Phys.,* **2001**, *73*(3), 681-700.
[3]   MacDiarmid, A. G. Nobel lecture: synthetic metals: A novel role for organic polymers. *Rev. Mod. Phys.,* **2001**, *73*(3), 701-712.
[4]   Pinter, E.; Fekete, Z. A.; Berkesi, O.; Makra, P.; Patzko, A.; Visy, C. Characterization of Poly(3-octylthiophene)/Silver nanocomposites prepared by solution doping. *J. Phys. Chem. C,* **2007**, *111*(32), 11872-11878.
[5]   Winokur, M. J. Structural studies of conducting polymers. In: *handbook conducting polymers*, Skotheim, T. A.; Elsenbaumer, R. L.; Reynolds, J. R., Eds. CRC Press, **1998**, Vol. 2, p. 707.
[6]   Winokur, M. J.; Chunwachirasiri, W. Nanoscale structure-property relationships in conjugated polymers: implications for present and future device applications. *J. Polymer Sci. Part B-Polymer Phys.,* **2003**, *41*(21), 2630-2648.
[7]   Reddinger, J. L.; Reynolds, J. R. Molecular engineering of pi-conjugated polymers. *Radic. Polymer Polyeletrolytes,* **1999**, *145*, 57-122.
[8]   McQuade, D. T.; Pullen, A. E.; Swager, T. M. Conjugated polymer-based chemical sensors. *Chem. Rev.,* **2000**, *100*(7), 2537-2574.
[9]   Bredas, J. L.; Themans, B.; Fripiat, J. G.; Andre, J. M.; Chance, R. R. Highly conducting polyparaphenylene, polypyrrole, and polythiophene chains - an abinitio study of the geometry and electronic-structure modifications upon doping. *Phys. Rev. B,* **1984**, *29*(12), 6761-6773.
[10] Geskin, V. M.; Bredas, J. L. Polaron pair versus bipolaron on oligothiophene chains: a theoretical study of the singlet and triplet states. *Chemphyschem,* **2003**, *4*(5), 498-505.
[11] Geskin, V. M.; Dkhissi, A.; Bredas, J. L. Oligothiophene radical cations: polaron structure in hybrid DFT and MP2 calculations. *Int. J. Quantum Chem.,* **2003**, *91*(3), 350-354.
[12] van Haare, J.; Havinga, E. E.; van Dongen, J. L. J.; Janssen, R. A. J.; Cornil, J.; Bredas, J. L. Redox states of long oligothiophenes: two polarons on a single chain. *Chemistry-a Eur. J.,* **1998**, *4*(8), 1509-1522.
[13] Chen, J.; Heeger, A. J.; Wudl, F. Confined soliton pairs (bipolarons) in polythiophene - insitu magnetic-resonance measurements. *Solid State Commun.,* **1986**, *58*(4), 251-257.
[14] Heeger, A. J.; Kivelson, S.; Schrieffer, J. R.; Su, W. P. Solitons in conducting polymers. *Rev. Mod. Phys.,* **1988**, *60*(3), 781-850.
[15] Patil, A. O.; Heeger, A. J.; Wudl, F. Optical-properties of conducting polymers. *Chem. Rev.,* **1988**, *88*(1), 183-200.
[16] Kurti, J.; Surjan, P. R.; Kertesz, M.; Frapper, G. Design of small-gap conjugated polymers. *Synth. Met.,* **1993**, *57*(2-3), 4338-4343.
[17] Kurti, J.; Surjan, P. R. Quinoid vs aromatic structure of polyisothianaphthene. *J. Chem. Phys.,* **1990**, *92*(5), 3247-3248.
[18] Cuff, L.; Kertesz, M. Evidence of quinonoid structures in the vibrational spectra of thiophene based conducting polymers: Poly(thiophene), poly(thieno[3,4-b]benzene), and poly(thieno[3,4-b]pyrazine). *J. Chem. Phys.,* **1997**, *106*(13), 5541-5553.
[19] Kertesz, M.; Choi, C. H.; Yang, S. J. Conjugated polymers and aromaticity. *Chem. Rev.,* **2005**, *105*(10), 3448-3481.
[20] Gao, Y.; Liu, C. G.; Jiang, Y. S. Electronic structure of thiophene oligomer dications: an alternative interpretation from the spin-unrestricted DFT study. *J. Phys. Chem. A,* **2002**, *106*(21), 5380-5384.
[21] Yurtsever, E. Multiple bipolaron formation in oligothiophenes. **1999**, *105*(3), 179-183.
[22] Casanovas, J.; Aleman, C. Comparative theoretical study of heterocyclic conducting oligomers: neutral and oxidized forms. **2007**, *111*(12), 4823-4830.
[23] Niehaus, T. A.; Di Carlo, A.; Frauenheim, T. Effect of self-consistency and electron correlation on the spatial extension of bipolaronic defects. **2004**, *5*(4), 167-174.
[24] Hutchison, G. R.; Zhao, Y. J.; Delley, B.; Freeman, A. J.; Ratner, M. A.; Marks, T. J. Electronic structure of conducting polymers: limitations of oligomer extrapolation approximations and effects of heteroatoms. *Phys. Rev. B,* **2003**, *68*(3), 035204.1-035204.13.
[25] Hutchison, G. R.; Ratner, M. A.; Marks, T. J. Electronic structure and band gaps in cationic heterocyclic oligomers. Multidimensional analysis of the interplay of heteroatoms, substituents, molecular length, and charge on redox and transparency characteristics. *J. Phys. Chem. B,* **2005**, *109*(8), 3126-3138.
[26] Moro, G.; Scalmani, G.; Cosentino, U.; Pitea, D. On the structure of polaronic defects in thiophene oligomers: a combined hartree-fock and density functional theory study. **2000**, *108*(2), 165-172.
[27] Chaudhury, P.; Bhattacharyya, S. P. Evolution of defect states in doped polythiophene: a study based on the method of simulated annealing. *Int. J. Quantum Chem.,* **2003**, *91*(6), 663-674.
[28] Roncali, J. Conjugated poly(thiophenes) - synthesis, functionalization, and applications. *Chem. Rev.,* **1992**, *92*(4), 711-738.
[29] Roncali, J. Linearly extended pi-donors: when tetrathiafulvalene meets conjugated oligomers and polymers. *J. Mater. Chem.,* **1997**, *7*(12), 2307-2321.
[30] Roncali, J. Synthetic principles for bandgap control in linear pi-conjugated systems. *Chem. Rev.,* **1997**, *97*(1), 173-205.
[31] McCullough, R. D. The chemistry of conducting polythiophenes. *Adv. Mater.,* **1998**, *10*(2), 93.
[32] Kline, R. J.; McGehee, M. D. Morphology and charge transport in conjugated polymer. *Polym. Rev.,* **2006**, *46*(1), 27-45.
[33] Chung, T. C.; Kaufman, J. H.; Heeger, A. J.; Wudl, F. Charge storage in doped poly(thiophene) - optical and electrochemical studies. *Phys. Rev. B,* **1984**, *30*(2), 702-710.
[34] Kaufman, J. H.; Chung, T. C.; Heeger, A. J.; Wudl, F. Poly(thiophene) - a stable polymer cathode material. *J. Electrochem. Soc.,* **1984**, *131*(9), 2092-2093.
[35] Lavarda, F. C.; Dossantos, M. C.; Galvao, D. S.; Laks, B. Insulator-to-metal transition in polythiophene. *Phys. Rev. B,* **1994**, *49*(2), 979-983.
[36] Alhalasah, W.; Holze, R. Electrochemical materials science: calculation vs. experiment as predictive tools in tailoring intrinsically conducting polythiophenes. *Microchim. Acta,* **2006**, *156*(1-2), 133-139.
[37] Hohnholz, D.; MacDiarmid, A. G., Line patterning of conducting polymers: new horizons for inexpensive, disposable electronic devices. *Synth. Met.,* **2001**, *121*(1-3), 1327-1328.

[†]<http://www.staff.u-szeged.hu/~fekete/req-polythiophene/>, and also as Supplementary Material at the publisher.

[38] Hohnholz, D.; Okuzaki, H.; MacDiarmid, A. G. Plastic electronic devices through line patterning of conducting polymers. *Adv. Funct. Mater.,* **2005**, *15*(1), 51-56.

[39] Dhoot, A. S.; Wang, G. M.; Moses, D.; Heeger, A. J. Voltage-induced metal-insulator transition in polythiophene field-effect transistors. *Phys. Rev. Lett.,* **2006**, *96*(24).

[40] Surin, M.; Leclere, P.; Lazzaroni, R.; Yuen, J. D.; Wang, G.; Moses, D.; Heeger, A. J.; Cho, S.; Lee, K. Relationship between the microscopic morphology and the charge transport properties in poly(3-hexylthiophene) field-effect transistors. *J. App. Phys.,* **2006**, *100*(3).

[41] Yang, C. Y.; Soci, C.; Moses, D.; Heeger, A. J. Aligned rrP3HT film: structural order and transport properties. *Synth. Met.,* **2005**, *155*(3), 639-642.

[42] Kozaki, M.; Yonezawa, Y.; Okada, K. Cyclopentadithiophene (CPDT) dimer dication: bipolaron model for quaterthiophenes. **2002**, *4*(25), 4535-4538.

[43] Casado, J.; Miller, L. L.; Mann, K. R.; Pappenfus, T. M.; Higuchi, H.; Orti, E.; Milian, B.; Pou-Amerigo, R.; Hernandez, V.; Navar-rete, J. T. L. Quinonoid oligothiophenes as electron-donor and electron-acceptor materials. a spectroelectrochemical and theoretical study. **2002**, *124*(41), 12380-12388.

[44] Granovsky, A. A. *PC GAMESS*, 6.4; **2004**.

[45] Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General atomic and molecular electronic-structure system. *J. Comput. Chem.,* **1993**, *14*(11), 1347-1363.

[46] Stewart, J. J. P. Optimization of parameters for semiempirical methods .1. method. *J. Comput. Chem.,* **1989**, *10*(2), 209-220.

[47] Davila, L. Y. A.; Caldas, M. J. Applicability of MNDO techniques AM1 and PM3 to ring-structured polymers. *J. Comput. Chem.,* **2002**, *23*(12), 1135-1142.

[48] Alves-Santos, M.; Davila, L. Y. A.; Petrilli, H. M.; Capaz, R. B.; Caldas, M. J. Application of standard DFT theory for nonbonded interactions in soft matter: Prototype study of poly-para-phenylene. *J. Comput. Chem.,* **2006**, *27*(2), 217-227.

[49] Fishtik, I.; Nagypál, I.; Gutman, I. *J. Chem. Soc. Faraday Trans.,* **1994**, *90*, 3245-3252.

[50] Fishtik, I.; Nagypal, I.; Gutman, I. New approach to sensitivity analysis of multiple equilibria in solutions. *J. Chem. Soc. Faraday Trans.,* **1994**, *90*(21), 3245-3252.

[51] Fishtik, I.; Nagypál, I.; Gutman, I. *J. Chem. Soc. Faraday Trans.,* **1995**, *91*, 259-267.

[52] Nagypal, I.; Hoffmann, E.; Gutman, I.; Fishtik, I. Response reactions: an interpretive concept in chemical thermodynamics. *Pure Appl. Chem.,* **1998**, *70*(3), 583-590.

[53] Galan, M.; Angulo, G. Some remarks on the application of relaxation techniques to chemical equilibria. *Chem. Phys.,* **2000**, *254*(2-3), 329-336.

[54] Hoffmann, E. A.; Nagypal, I. Response reactions: equilibrium coupling. *J. Phys. Chem. B,* **2006**, *110*(21)*,* 10581-10584.

[55] Fishtik, I.; Povar, I. Buffer capacity in multiple chemical reaction systems involving solid phases. *Can. J. Chem.-Revue Can. De Chimie,* **2006**, *84*(8)*,* 1036-1044.

[56] Fishtik, I.; Datta, R. Group additivity vs ab initio. *J. Phys. Chem. A,* **2003**, *107*(34), 6698-6707.

[57] Fishtik, F.; Urban, D.; Wilcox, J. The effect of stoichiometry on ab initio-based thermochemistry predictions. *Chem. Phys. Let.,* **2006**, *417*(1-3), 185-189.

[58] Fishtik, I. Group additivity methods without group values. *J. Phys. Chem. A,* **2006**, *110*(49), 13264-13269.

[59] Fishtik, I.; Datta, R. Stoichiometric approach to quantitative structure-property relationships (QSPR). *J. Chem. Info. Comp. Sci.,* **2003**, *43*(4), 1259-1268.

[60] Fishtik, I.; Datta, R.; Liebman, J. F. Response reactions: A mathematical well-defined way to obtain accurate thermochemistry from ab initio calculations. *J. Phys. Chem. A,* **2003**, *107*(5), 695-705.

[61] Hoffmann, E. A.; Nagypál, I. *Phys. Chem. Chem. Phys.* **2001**, *3*, 3107-3113.