Supplementary information to

Z-Effect Reversal in Carboxylic Acid Associates

M. G. Medvedev, I. S. Bushmarinov, K. A. Lyssenko

Table of Contents

CSD search details

Data was extracted from the CSD with the CCDC ConQuest[1]. Unless otherwise specified, the search was limited by R value < 7.5%, and only *inter*-molecular hydrogen bonds were considered. Data visualization and analysis were performed using R[2] and CCDC Mercury[3].

Hydrogen bonds were defined during the searches as having O…X distance shorter than the sum of O and X van der Waals radii, and O-H…X angle larger than 150°. O=C-O-X torsions of 0–40° are were considered *cis*, and 140–180° were denoted as *trans*.

CSD statistics

Table S1. The CSD statistics (median and quartiles) for COOH…X fragments after exclusion of COOH dimers

X	O=C-O-H torsion	Count	%	X…O, Å		
				median	Q1	Q3
N	*cis*	3351	99.4	2.652	2.611	2.689
	trans	21	0.6	2.654	2.621	2.690
O	*cis*	10952	96.0	2.596	2.548	2.646
	trans	461	4.0	2.566	2.498	2.628
F	*cis*	65	100	2.581	2.476	2.628
S	*cis*	19	100	3.136	3.101	3.195
Cl	*cis*	387	97.5	3.006	2.975	3.047
	trans	10	2.5	2.976	2.959	3.036
Br	*cis*	71	98.6	3.141	3.115	3.202
	trans	1	1.4	3.183	3.183	3.183
I	*cis*	7	100	3.391	3.370	3.421

Table S2. Finite mixture modeling of $d_{O...O}$.

Conformation	Component	λ	μ	σ	Estimated count
trans	1	0.1847	2.4621	0.0195	85
	2	0.6473	2.5743	0.0593	298
	3	0.1680	2.7109	0.1232	77
cis	1	0.041	2.463	0.017	444
	2	0.913	2.598	0.067	9997
	3	0.047	2.722	0.110	512

Computational methods and additional tables

Geometries of all complexes included in this study were optimized at the uM11[4]/aug-cc-pV(T+d)Z[5] level of theory. M11 method has been chosen for its best performance for weak interactions among other popular DFT functionals and MP2 method[6]. All the same calculations with the uMP2 method and the same basis set were also carried out, providing quantitatively equal results (deviations in R^2 values for all discussed correlations did not exceed 0.02). Basis set aug-cc-pV(T+d)Z was used in spherical harmonics form, differing from aug-cc-pVTZ[7–9] by one extra tight d-function for elements Na-Ar, which is known to improve accuracy[5,10]. The calculations have been performed using the program package GAMESS-US[11,12] (version 20130501R1). For all complexes, except two with bromine, we have used an unpublished large grid by Curtis Janssen as implemented in GAMESS-US (JANS=2 keyword). It uses 155 radial points in the Euler-MacLaurin quadrature for all atoms, and prunes from a Lebedev grid whose largest size is 974, thus using about 71000 grid points/atom. For complexes with bromine we used even larger grid: 330 radial points for all atoms and 1730 angular points in the Lebedev grid. Four complexes ("non-alt" Z-conformers with formaldehyde and acetaldehyde and "non-alt" E-conformers with NO_3^- and BF_4^- anions) were optimized at C_s symmetry to avoid formation of additional bonds with acid molecule. All of them were then bent and reoptimized at C_1 symmetry to give multiple-bonded "alt" structures. All other calculations have been carried out at C_1 symmetry. The harmonic vibrational frequency calculations using the same method as for the geometry optimization were used to ascertain the presence of local minima. The "atoms-in-molecules" (AIM) analysis was performed at the same level of theory. The calculation of AIM properties and visualization of results was done using the AIMALL[13] program.

Table S3. Summary of changes in total energy, CH_3COOH QTAIM charges[a] and O…X distance upon transition from *cis* to *trans* conformation of hydrogen-bonded acetic acid with various acceptors (CH_3COOH…X), sorted by decrease of conformer energy difference, MP2/aug-cc-pV(T+d)Z calculations.

	ΔE_Z,	$\Delta q(\Omega) \cdot 10^2$, e					
H-bond acceptor	**kcal·mol^{-1}**	**H**	**O1**	**C**	**O2**	**CH_3**	**$\Delta d_{O\cdots X}$, Å**
H_2O	-7.70	1.7	-0.8	-0.9	-2.8	4.7	-0.086
$O{=}CH_2$[b]	-6.51	2.0	-1.2	-0.1	-3.1	2.9	0.002
OMe_2	-6.12	0.9	-1.5	-1.3	-2.6	4.1	-0.040
O=C(H)Me	-6.09	1.9	-1.2	-0.2	-2.9	2.6	-0.011
CF_4	-5.50	1.5	-1.7	-1.5	-1.9	4.0	0.014
triazole	-5.35	1.5	-1.2	0.0	-2.6	2.2	-0.047
none	-5.23	1.5	-1.9	-1.3	-2.3	4.1	0.000
$O{=}CMe_2$[b]	-4.37	1.6	-0.5	0.2	-2.6	2.1	-0.001
N≡CH	-3.70	1.2	-1.3	-0.9	-1.9	2.9	-0.021
$O{=}CH_2$	-2.85	1.1	-1.2	-0.8	-1.2	2.3	0.013
$O{=}CMe_2$	-2.01	1.2	-1.1	-0.4	-0.8	1.5	0.010
BF_4^-	2.56	1.5	-1.2	2.8	0.8	-4.1	-0.030
NO_3^-	3.16	0.8	-1.0	2.8	0.9	-4.4	-0.045
N_3^-	4.24	0.5	-0.7	3.2	1.2	-4.5	-0.040
NCO^-	4.24	0.4	-0.7	3.2	1.2	-4.7	-0.048
BF_4^-[c]	5.05	2.0	-0.8	4.4	1.2	-7.0	-0.038
NO_3^-[c]	5.06	1.0	-0.6	3.4	1.1	-5.4	-0.041
Br^-	5.25	1.1	-1.2	3.8	1.7	-4.4	-0.036
Cl^-	5.45	0.7	-0.9	3.8	1.7	-5.1	-0.042
N_3^-[c]	5.64	1.1	-1.1	4.2	1.2	-5.8	-0.056

[a] for O-H hydrogen (H), O-H oxygen (O1), C=O carbon (C), C=O oxygen (O2) and methyl group (CH_3)

[b] Alternative *cis* conformation

[c] Alternative *trans* conformation

Table S4. The energies of *cis* and *trans* conformations of selected free acids and their RCOOH…NCO⁻ complexes.

	Energies, hartree				ΔE, kcal/mol	
Acid	cis_{free}	$trans_{free}$	$cis_{\mathrm{NCO^-}}$	$trans_{\mathrm{NCO^-}}$	ΔE_{free}	$\Delta E_{\mathrm{NCO^-}}$
HCOOH	-189.7801	-189.7733	-357.9608	-357.9679	-4.29	4.45
CH_3COOH[a]	-	-	-	-	-5.09	4.73
EtCOOH	-268.3971	-268.3894	-436.5734	-436.5806	-4.81	4.56
PrCOOH	-307.6981	-307.6905	-475.8746	-475.8817	-4.77	4.45
H2C=CHCOOH	-267.1640	-267.1555	-435.3423	-435.3489	-5.33	4.09
H2NCH2COOH	-284.4453	-284.4376	-452.6260	-452.6319	-4.81	3.69

[a]see Table S5 for energies

Effect of secondary interactions

Some deviations from linear fit in the analyzed series are the consequence of weak secondary interactions. The total energy of such interactions can be successfully[14] quantified using Espinosa-Lecomte-Molins correlation[15–17] by summing their "contact energy"

$$E_{cont}(\text{hartree}) \approx -0.5 \times V(\mathbf{r}_{\mathrm{BCP}})$$

where $V(\mathbf{r})$ is local potential energy density and $\mathbf{r}_{\mathrm{BCP}}$ are the coordinates of the bond critical point corresponding to the weak interaction. Indeed, if we exclude the complexes with $\sum E_{cont} > 10$ kJ · mol^{-1}, the R^2 exceeds 0.97 for all fits to individual $\Delta q(\Omega)$. The fit to both $\Delta q(\Omega)_{CH_3}$ and $\Delta q(\Omega)_{O2}$, however, does not improve.

The dipole-dipole interaction, on the other hand, does not seem to play a major role since both "trans" and "alt-trans" structures of azide anion complex (Table 1 in the main text) have very similar $\Delta q(\Omega)$ and ΔE.

Source function analysis

The analysis of the source function proposed by Gatti et. al.[18] allows to quantify the influence of each atom in a system in determining the amount of electron density at a given point. In our system it provides insight into the stereolectronic contributions to the C=O bond critical point (BCP). On one hand, in the *cis* conformation the contribution of the O-H oxygen

remains nearly constant, allowing to associate it with the effect of the lone pair. On the other hand, in the *trans* conformation it strongly correlates with ΔE_Z. In other words, strong hydrogen bonds increase the contribution of O-H oxygen atomic source to the C=O bond critical point in *trans* conformation, making it similar to the contribution in *cis* conformer. These facts support the hypothesis that O-H group involved in a strong hydrogen bond can serve as a "surrogate" lone pair acting upon the C=O bond. With R^2 of 0.80, the relation is, however, only semi-quantitative.

ANOVA analysis

Summaries of linear fits

Below are provided the results of ANOVA analysis for all discussed relationships in the data, as implemented in R.[2] All relationships are statistically significant, but only for the fit against $\Delta q(\Omega)_{CH_3}$ together with $\Delta q(\Omega)_{O2}$ the residuals are close to normality. The data fields and variables are self-descriptive, delta_energy is ΔE_Z.

ΔE_Z against total charge:

```
> summary(lm(delta_energy ~ XO_q_total, data = fullv))

Residuals:
    Min      1Q  Median      3Q     Max
-3.2981 -1.2409 -0.2667  1.0834  3.3353

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -4.8324     0.4963  -9.736 8.08e-09 ***
XO_q_total  -10.0746     0.7584 -13.285 4.56e-11 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.719 on 19 degrees of freedom
Multiple R-squared:  0.9028,      Adjusted R-squared:  0.8977
F-statistic: 176.5 on 1 and 19 DF,  p-value: 4.565e-11
```

ΔE_Z against $\Delta q(\Omega)_{CH_3}$:

```
> summary(lm(delta_energy ~ delta_q_CH3, data = fullv))

Call:
lm(formula = delta_energy ~ delta_q_CH3, data = fullv)

Residuals:
     Min       1Q   Median       3Q      Max
-2.26254 -0.64786  0.02242  1.13308  1.52799

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   -1.0881     0.2663  -4.086  0.00063 ***
delta_q_CH3 -125.3564     6.4837 -19.334 5.89e-14 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.213 on 19 degrees of freedom
Multiple R-squared:  0.9516,      Adjusted R-squared:  0.9491
F-statistic: 373.8 on 1 and 19 DF,  p-value: 5.889e-14
```

ΔE_Z against $\Delta q(\Omega)_{O2}$:

```
> summary(lm(delta_energy ~ delta_q_O2, data = fullv))

Call:
lm(formula = delta_energy ~ delta_q_O2, data = fullv)

Residuals:
    Min      1Q  Median      3Q     Max
-1.9519 -0.2657 -0.1209  0.8016  1.8028

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.9933     0.2502   7.967 1.78e-07 ***
delta_q_O2  301.8277    13.1749  22.909 2.66e-15 ***
---
```

```
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.031 on 19 degrees of freedom
Multiple R-squared:  0.9651,      Adjusted R-squared:  0.9632
F-statistic: 524.8 on 1 and 19 DF,  p-value: 2.661e-15
```

ΔE_Z against $\Delta q(\Omega)_C$:

```
> summary(lm(delta_energy ~ delta_q_C, data = fullv))

Call:
lm(formula = delta_energy ~ delta_q_C, data = fullv)

Residuals:
   Min     1Q Median     3Q    Max
-3.185 -0.512  0.107  1.157  2.373

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -2.9648     0.4256  -6.966 1.23e-06 ***
delta_q_C   228.7436    17.5581  13.028 6.39e-11 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.75 on 19 degrees of freedom
Multiple R-squared:  0.8993,      Adjusted R-squared:  0.894
F-statistic: 169.7 on 1 and 19 DF,  p-value: 6.388e-11
```

ΔE_Z against $\Delta q(\Omega)_{O2}$ together with $\Delta q(\Omega)_{CH_3}$:

```
> summary(lm(delta_energy ~ delta_q_CH3 + delta_q_O2, data = fullv))

Call:
lm(formula = delta_energy ~ delta_q_CH3 + delta_q_O2, data = fullv)

Residuals:
     Min       1Q   Median       3Q      Max
-1.40172 -0.31566  0.03239  0.30420  1.13288

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.6563     0.2666   2.462   0.0241 *
delta_q_CH3 -57.5170     9.5392  -6.030 1.06e-05 ***
delta_q_O2  172.5765    22.8076   7.567 5.37e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6094 on 18 degrees of freedom
Multiple R-squared:  0.9884,      Adjusted R-squared:  0.9871
F-statistic: 768.9 on 2 and 18 DF,  p-value: < 2.2e-16
```

ΔE_Z against $\Delta d_{C=O}$:

```
> summary(lm(delta_energy ~ delta_CO_len, data=fullv))

Call:
```

```
lm(formula = delta_energy ~ delta_CO_len, data = fullv)

Residuals:
    Min      1Q  Median      3Q     Max
-2.6587 -1.1334  0.1862  1.0334  2.7024

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)     4.0198     0.4826   8.329 9.16e-08 ***
delta_CO_len -903.1213    64.5085 -14.000 1.84e-11 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.639 on 19 degrees of freedom
Multiple R-squared:  0.9116,      Adjusted R-squared:  0.907
F-statistic:   196 on 1 and 19 DF,  p-value: 1.84e-11
```

Δd_{C-C} against $\Delta d_{C=O}$:

```
> summary(lm(delta_CC_len ~ delta_CO_len, data=fullv))

Call:
lm(formula = delta_CC_len ~ delta_CO_len, data = fullv)

Residuals:
      Min        1Q    Median        3Q       Max
-0.005648 -0.003476  0.001378  0.002337  0.003910

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.002445   0.001007   2.428   0.0253 *
delta_CO_len -1.186977   0.134611  -8.818 3.83e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.003421 on 19 degrees of freedom
Multiple R-squared:  0.8036,      Adjusted R-squared:  0.7933
F-statistic: 77.75 on 1 and 19 DF,  p-value: 3.834e-08
```

Diagnostic plots for the best fit

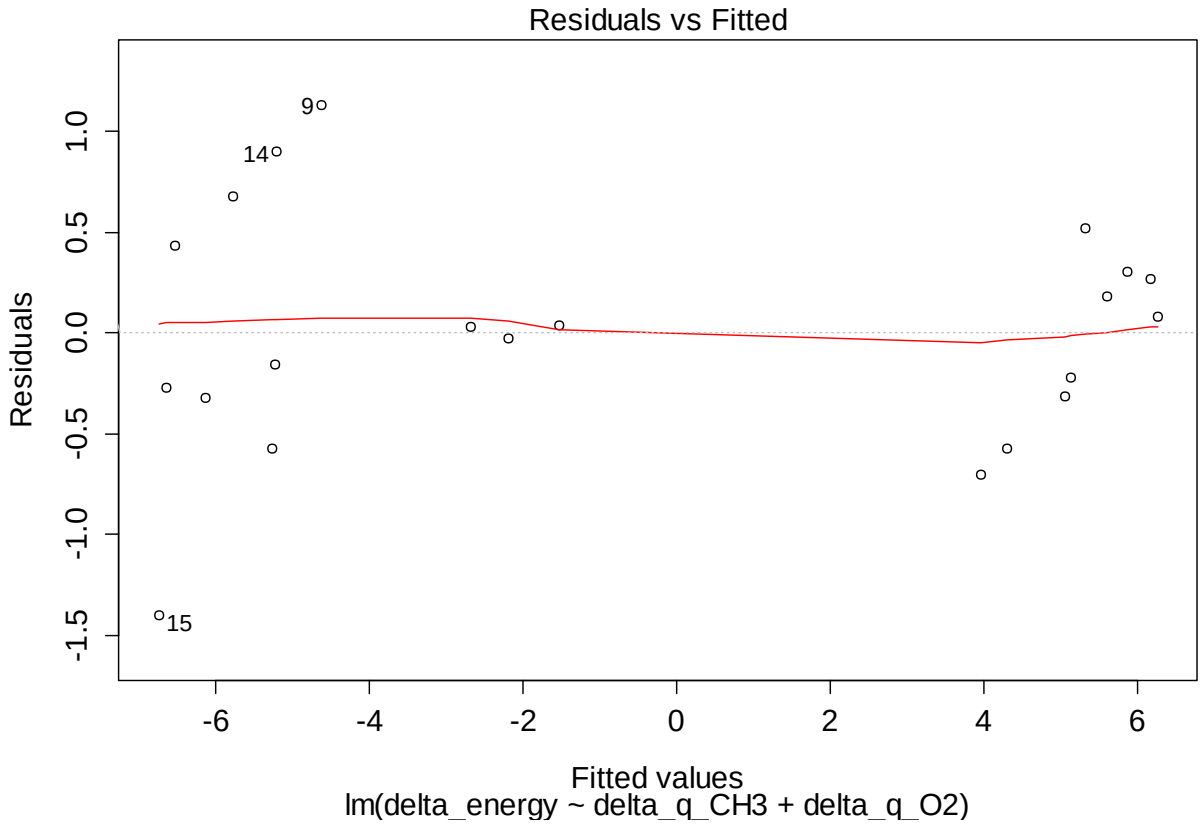


Figure S1. Residuals plot for the relationship $\Delta E_Z\ (\mathbf{kcal \cdot mol^{-1}}) \approx \mathbf{0.7} - \mathbf{58}\Delta q(\Omega)_{CH_3} + \mathbf{173}\Delta q(\Omega)_{O2}$. Numbers on plot denote three furthest outlying data points and correspond to row numbers in Table .

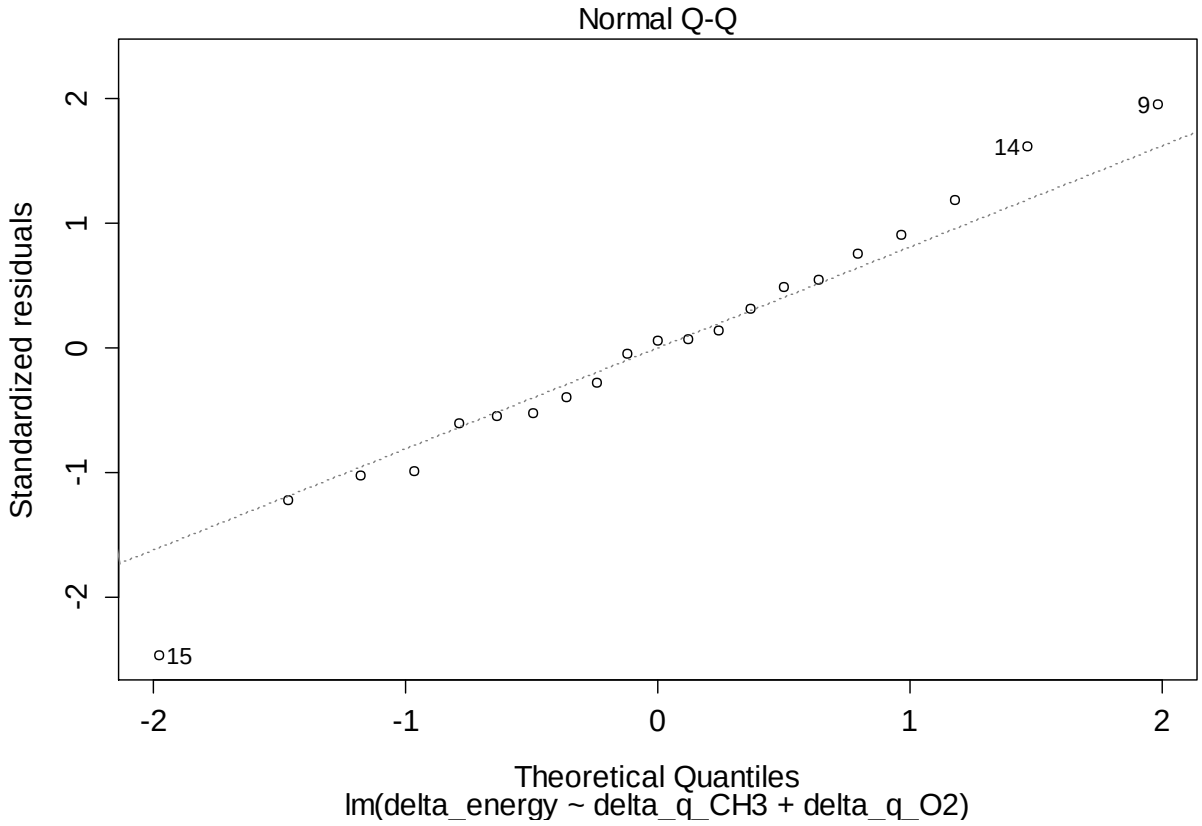


Figure S2. Q-Q plot for the relationship $\Delta E_Z\ (\mathbf{kcal \cdot mol^{-1}}) \approx \mathbf{0.7 - 58}\Delta q(\Omega)_{CH_3} + \mathbf{173}\Delta q(\Omega)_{O2}$. Numbers on plot denote three furthest outlying data points and correspond to row numbers in Table .

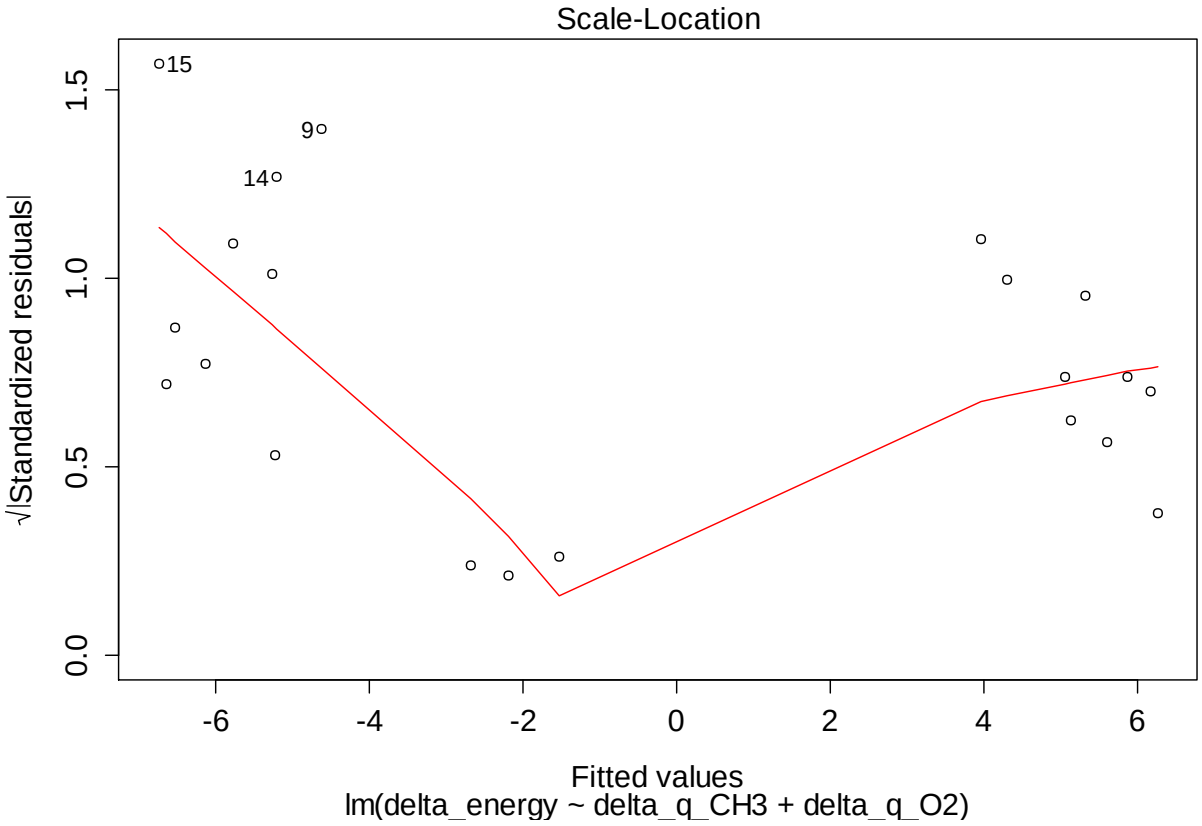


Figure S3. Scale-location plot for the relationship $\Delta E_Z\ (\mathbf{kcal \cdot mol^{-1}}) \approx \mathbf{0.7} - \mathbf{58}\Delta q(\Omega)_{CH_3} + \mathbf{173}\Delta q(\Omega)_{O2}$. Numbers on plot denote three furthest outlying data points and correspond to row numbers in Table .

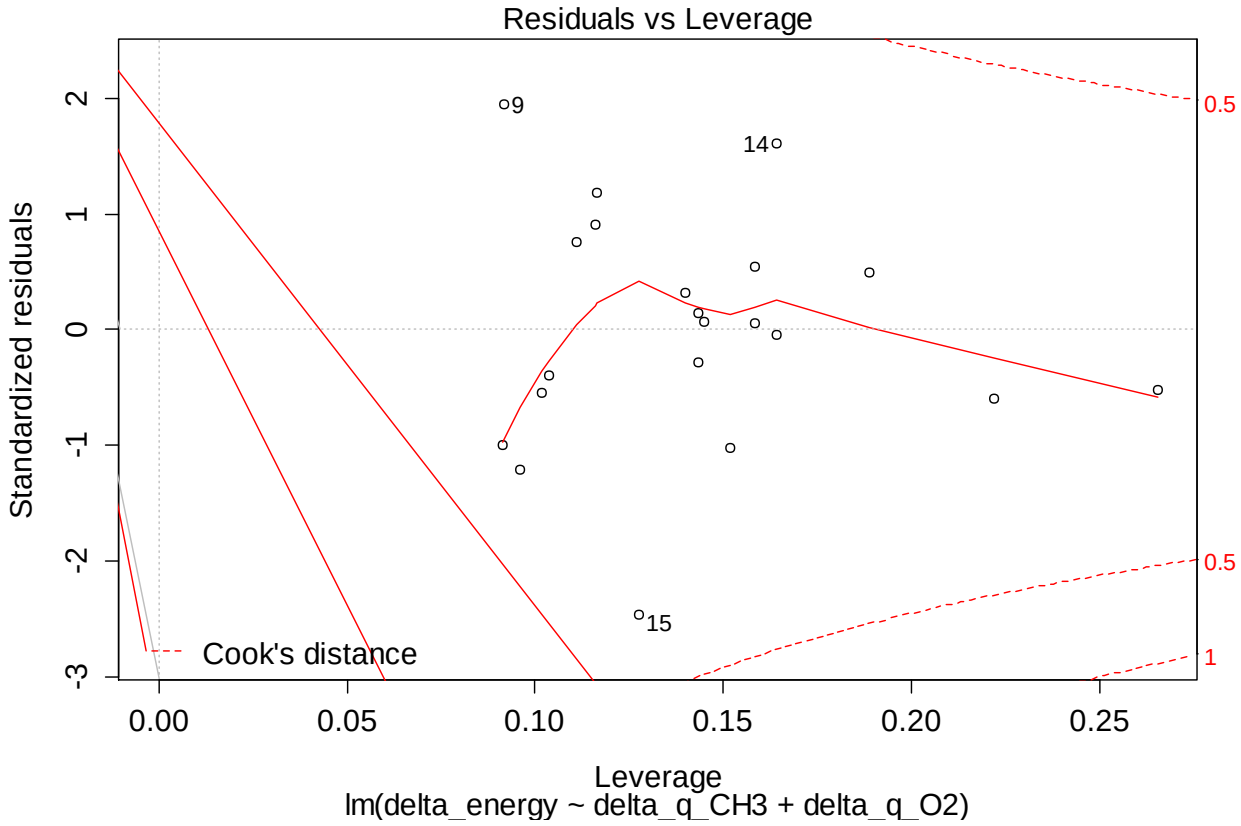


Figure S4. Residuals vs leverage plot for the relationship $\Delta E_Z\ (\mathbf{kcal \cdot mol^{-1}}) \approx \mathbf{0.7} - \mathbf{58}\Delta q(\Omega)_{CH_3} + \mathbf{173}\Delta q(\Omega)_{O2}$. Numbers on plot denote three furthest outlying data points and correspond to row numbers in Table .

LRSE in human insulin

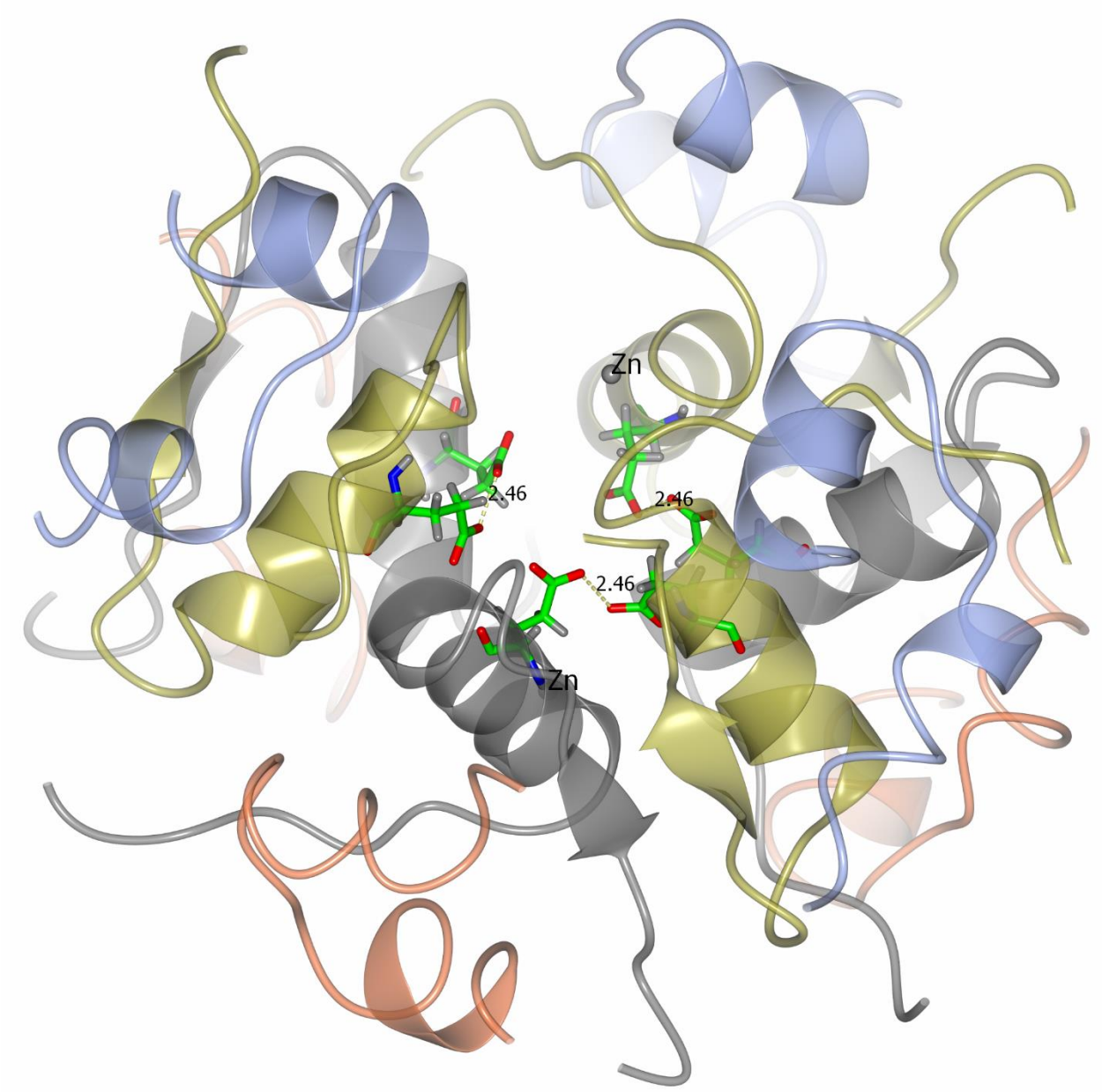
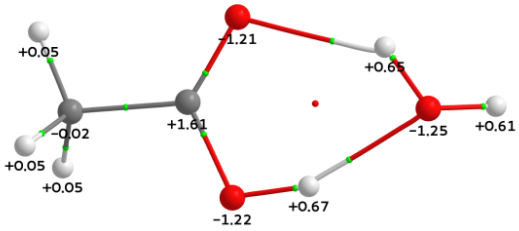
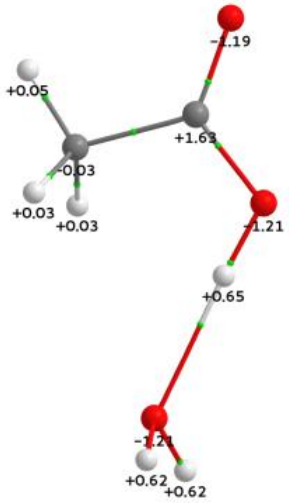
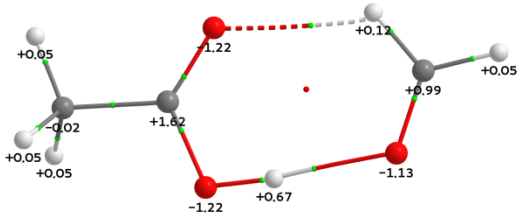
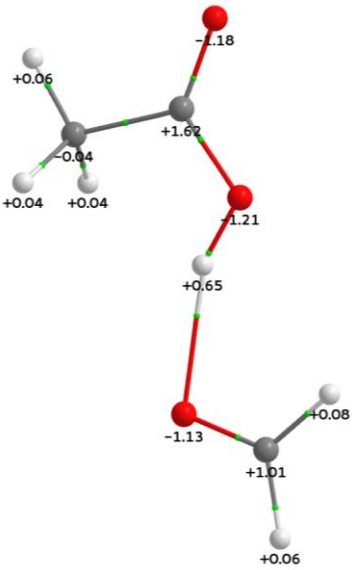


Figure S5. Contacting GluB13 residues in the insulin T_6 hexamer. $O^{\varepsilon 2}$…$O^{\varepsilon 2}$ distances are marked on the image. Protein chains are colored by symmetry equivalence.

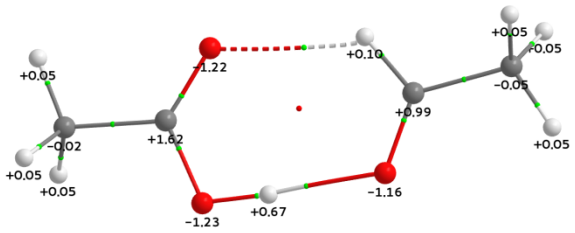
OPLS-aa force field evaluation for LRSE

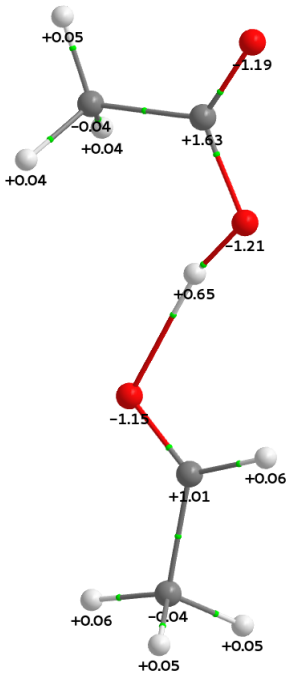
Two butyric acid - butyrate associates, modeling the interaction of Glu13B side chains, were optimized with uM11/aug-cc-pV(T+d)Z method (GAMESS-US) and OPLS-aa force field (TINKER[20]): trans-trans (as observed in T_6 insulin) and cis-cis geometries. In the case of trans-trans conformer in OPLS-aa, the dihedral C…O(H)…O…C and valence O(H)…O…C angles were restrained to 180 and 120 degrees respectively to preserve the correct dimer geometry, which was significantly altered otherwise. While the M11 calculation predicts trans-trans form to be 4.5 kcal/mol more stable than cis, OPLS-aa predicts them to have identical energies.

Molecular graphs

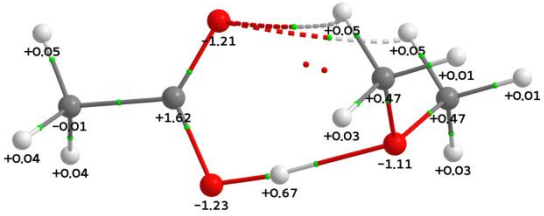
	cis	trans
H_2O		
$O{=}CH_2$[a]		

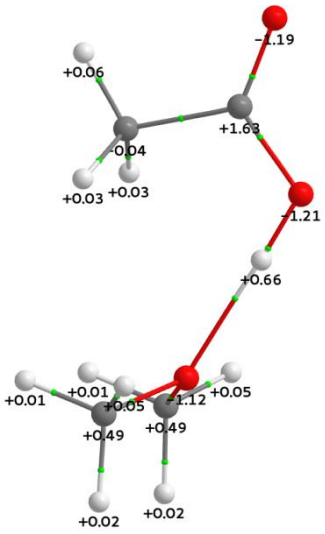
O=C(H)Me
[a]


+0.05
+0.05
−1.22
+0.10
+0.05
+0.05
−0.05
+0.99
+0.05
−0.02
+1.62
+0.05
+0.05
−1.23
+0.67
−1.16

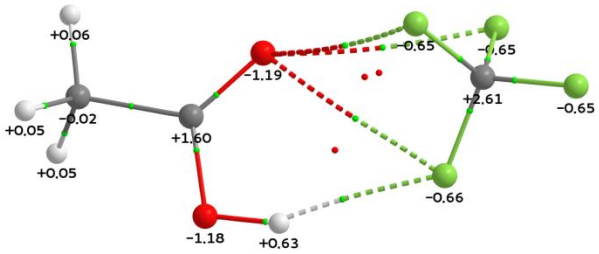

+0.05
−1.19
−0.04
+0.04
+1.63
+0.04
−1.21
+0.65
−1.15
+0.06
+1.01
+0.06
−0.04
+0.05
+0.05

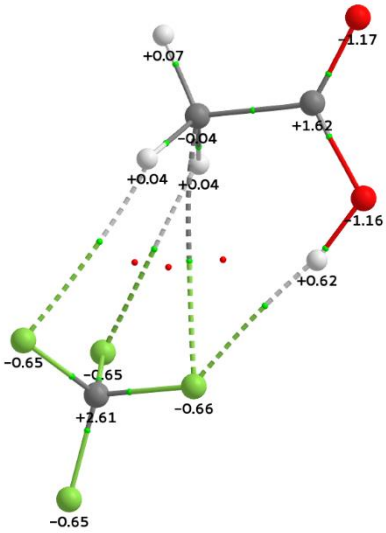
OMe_2


+0.05
−1.21
+0.05
+0.05
+0.01
+0.01
+0.47
−0.01
+1.62
+0.47
+0.03
+0.04
+0.04
−1.11
+0.03
−1.23
+0.67

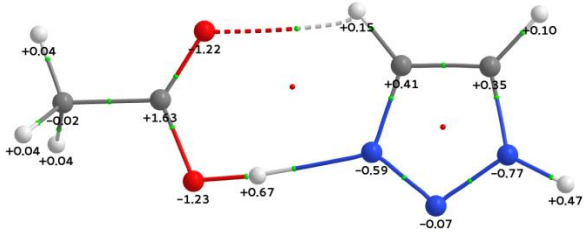

−1.19
+0.06
+1.63
−0.04
+0.03
+0.03
−1.21
+0.66
+0.01
+0.01
+0.05
−1.12
+0.05
+0.49
+0.49
+0.02
+0.02

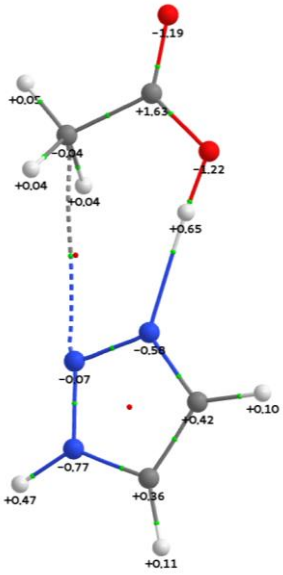
CF_4



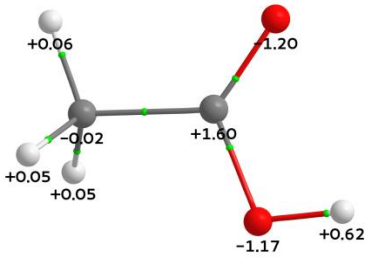


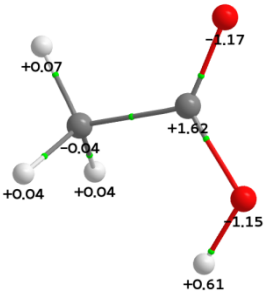
triazole



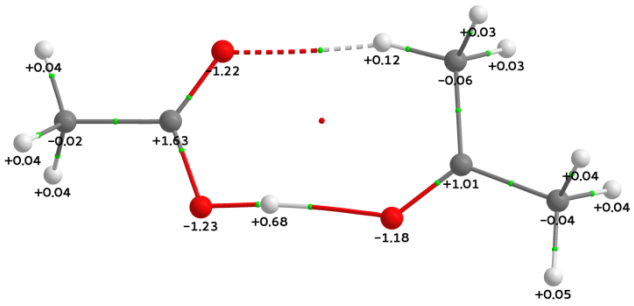


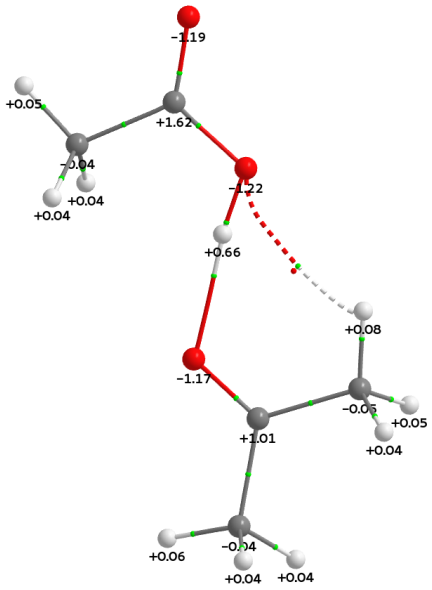
none



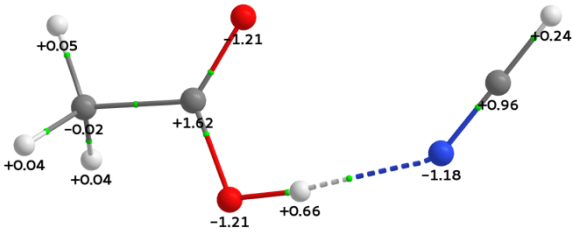


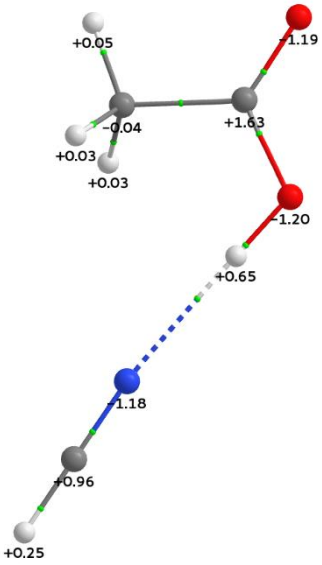
O=CMe$_2$[a]


+0.04
+0.04
+0.04
−0.02
+1.63
−1.22
+0.12
+0.03
+0.03
−0.06
+1.01
+0.04
+0.04
−0.04
+0.05
−1.23
+0.68
−1.18

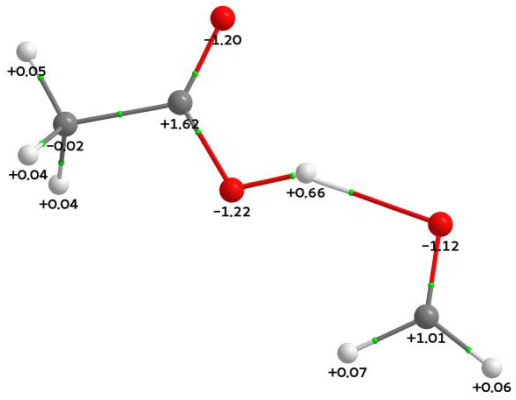

−1.19
+0.05
+1.62
−0.04
+0.04
+0.04
−1.22
+0.66
+0.08
−1.17
−0.05
+0.05
+0.04
+1.01
+0.06
−0.04
+0.04
+0.04

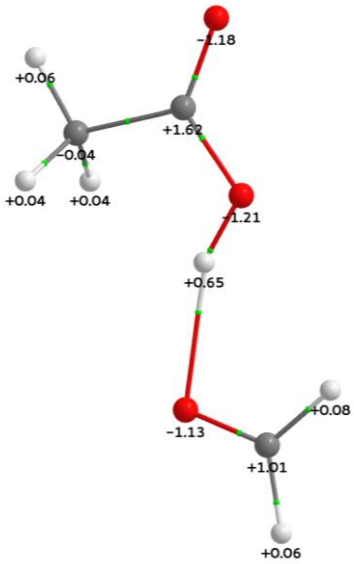
N≡CH


+0.05
−1.21
+0.24
+0.96
−0.02
+1.62
+0.04
+0.04
−1.18
−1.21
+0.66

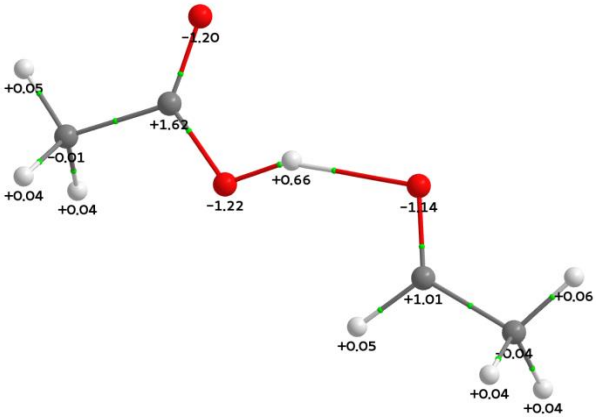

+0.05
−1.19
−0.04
+1.63
+0.03
+0.03
−1.20
+0.65
−1.18
+0.96
+0.25

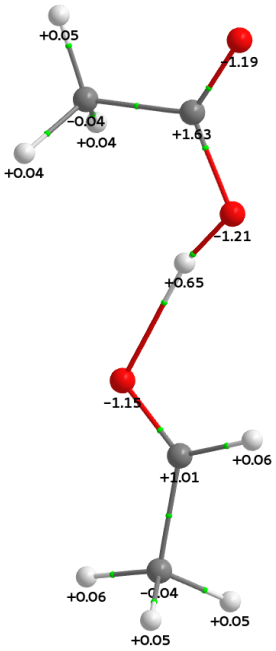
O=CH_2


−1.20
+0.05
+1.62
−0.02
+0.04
+0.04
−1.22
+0.66
−1.12
+1.01
+0.07
+0.06

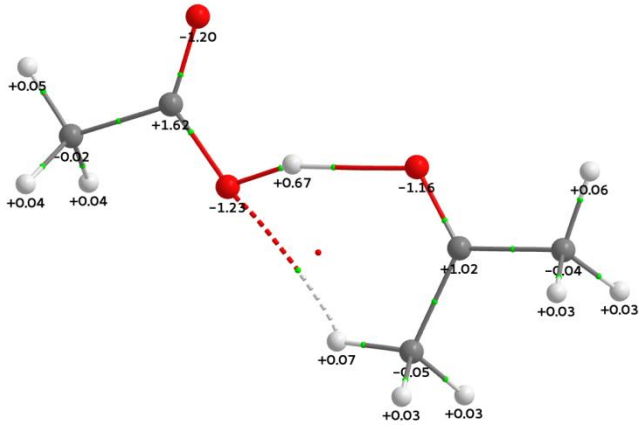

−1.18
+0.06
+1.62
−0.04
+0.04
+0.04
−1.21
+0.65
−1.13
+0.08
+1.01
+0.06

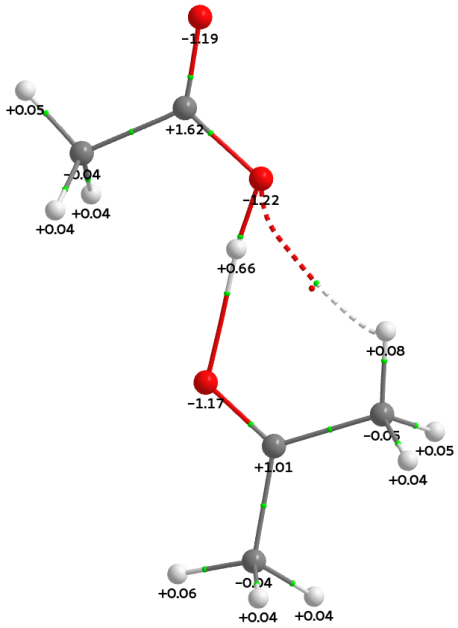
O=C(H)Me


−1.20
+0.05
+1.62
−0.01
+0.04
+0.04
−1.22
+0.66
−1.14
+1.01
+0.06
+0.05
−0.04
+0.04
+0.04

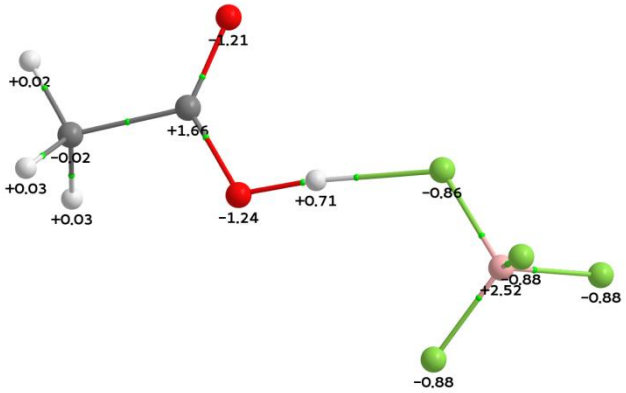

+0.05
−1.19
−0.04
+0.04
+1.63
+0.04
−1.21
+0.65
−1.15
+0.06
+1.01
+0.06
−0.04
+0.05
+0.05

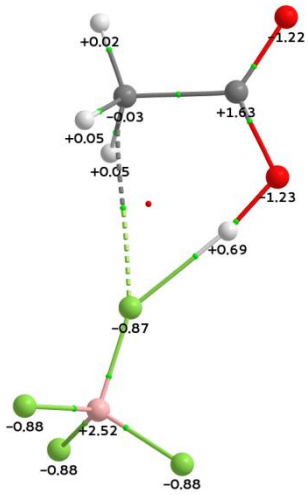
$O{=}CMe_2$



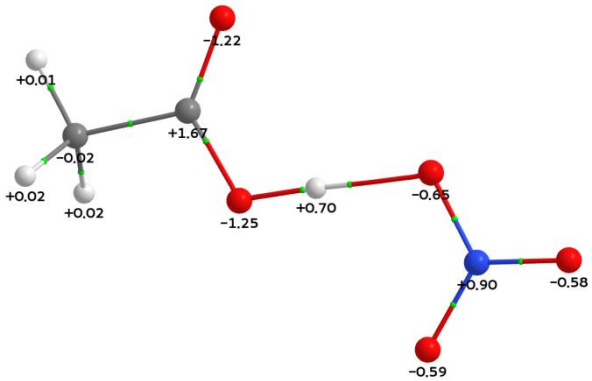


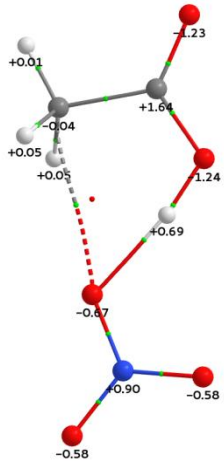
BF_4^-



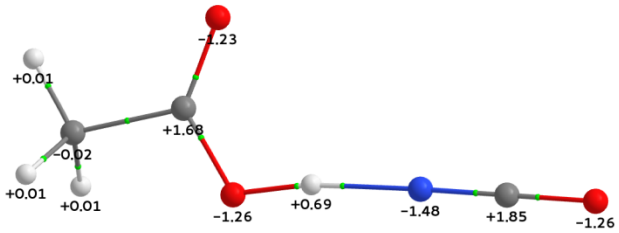


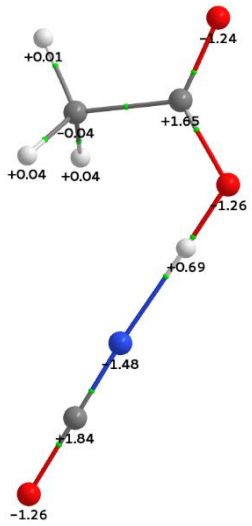
NO_3^-



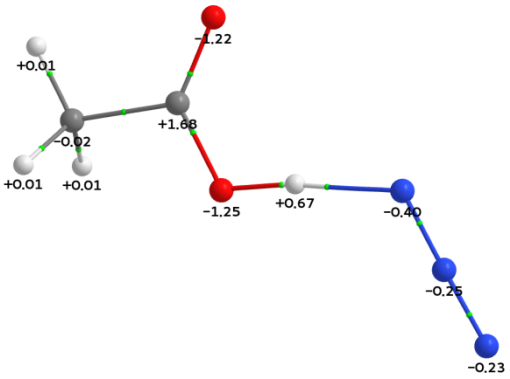


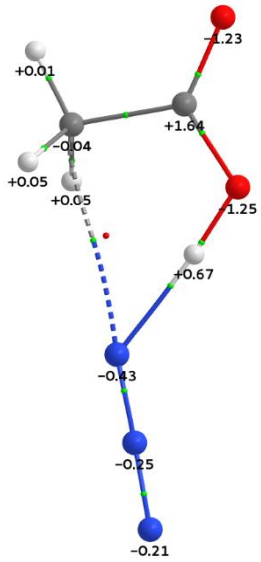
NCO^-


+0.01
−1.23
+1.68
−0.02
+0.01
+0.01
−1.26
+0.69
−1.48
+1.85
−1.26

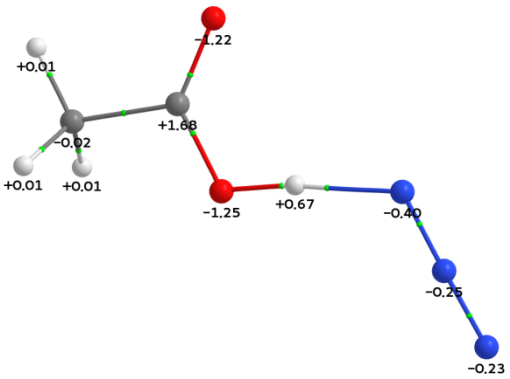

+0.01
−1.24
+1.65
−0.04
+0.04
+0.04
−1.26
+0.69
−1.48
+1.84
−1.26

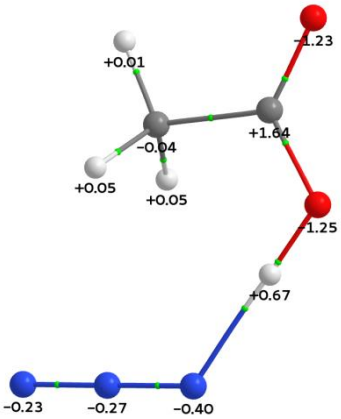
N_3^-


+0.01
−1.22
+1.68
−0.02
+0.01
+0.01
−1.25
+0.67
−0.40
−0.26
−0.23

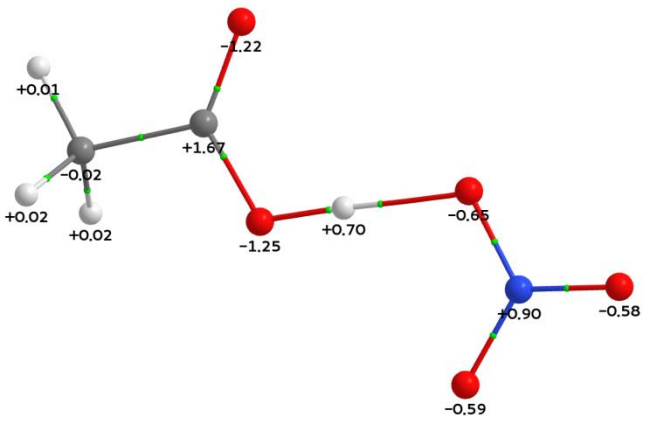

+0.01
−1.23
+1.64
−0.04
+0.05
+0.05
−1.25
+0.67
−0.43
−0.25
−0.21

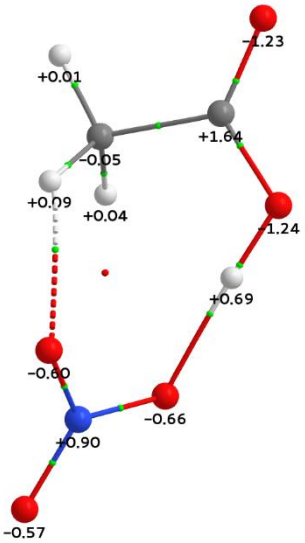
$N_3^{-[b]}$


+0.01
−1.22
+1.68
−0.02
+0.01
+0.01
−1.25
+0.67
−0.40
−0.26
−0.23

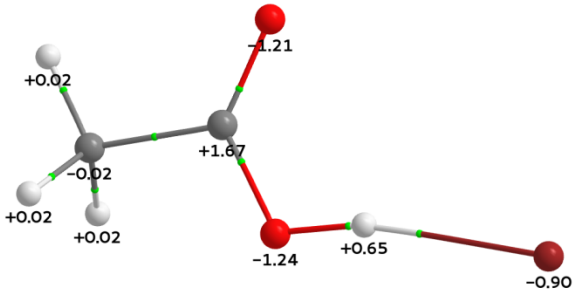

+0.01
−1.23
+1.64
−0.04
+0.05
+0.05
−1.25
+0.67
−0.23
−0.27
−0.40

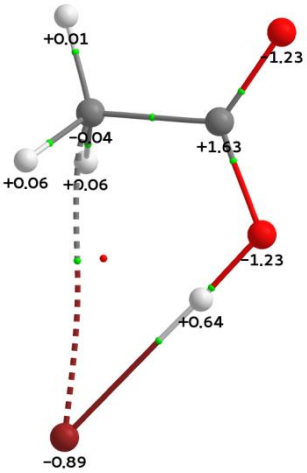
NO_3^{-}[b]


−1.22
+0.01
+1.67
−0.02
+0.02
+0.02
−1.25
+0.70
−0.65
+0.90
−0.58
−0.59

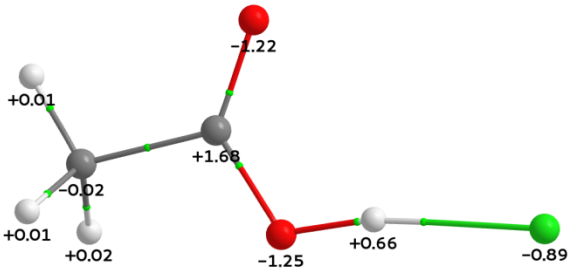

−1.23
+0.01
+1.64
−0.05
+0.09
+0.04
−1.24
+0.69
−0.60
−0.66
+0.90
−0.57

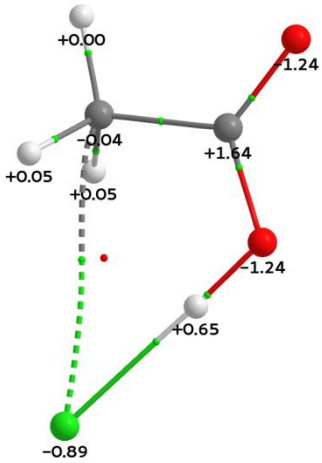
Br^{-}


−1.21
+0.02
+1.67
−0.02
+0.02
+0.02
−1.24
+0.65
−0.90


+0.01
−1.23
−0.04
+1.63
+0.06
+0.06
−1.23
+0.64
−0.89

Cl^{-}


−1.22
+0.01
+1.68
−0.02
+0.01
+0.02
−1.25
+0.66
−0.89


+0.00
1.24
−0.04
+1.64
+0.05
+0.05
−1.24
+0.65
−0.89

$BF_4^{-[b]}$

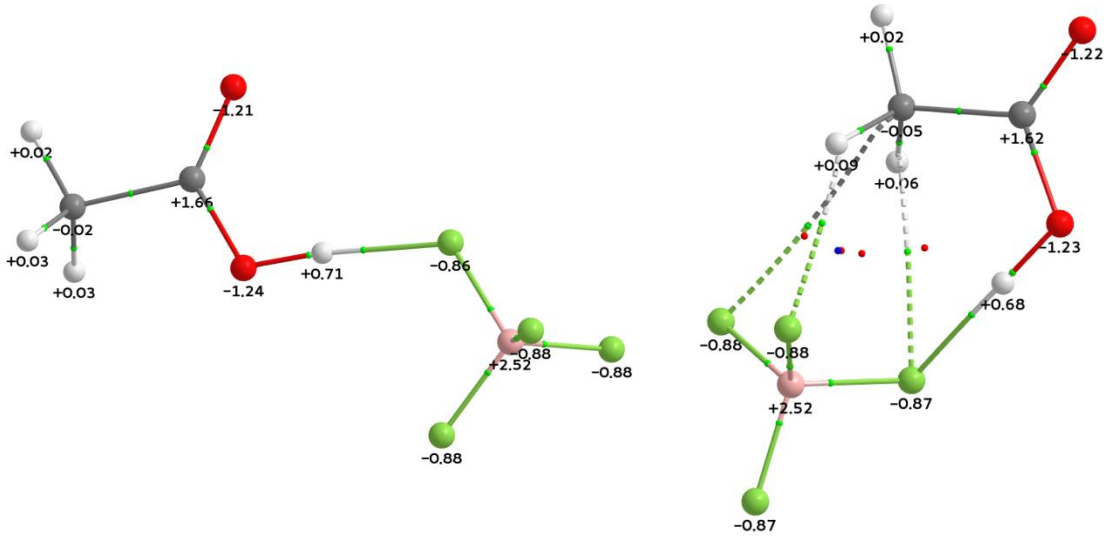


Figure S5. Optimized geometries at the M11/aug-cc-pV(T+d)Z level of theory in CPK colors with AIM atomic charges and AIM distribution of critical points (green spheres – BCP, red spheres – RCP, blue spheres – CCP). Adduct pairs are sorted by decrease of conformer energy difference (the same order with the Table 1). [a] Denotes alternative cis conformation, while [b] denotes alternative trans conformation.

References

1 F. H. Allen, *Acta Crystallogr. Sect. B-Struct. Sci.*, 2002, **58**, 380–388.
2 R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2015.
3 C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek and P. A. Wood, *J. Appl. Crystallogr.*, 2008, **41**, 466–470.
4 R. Peverati and D. G. Truhlar, *J. Phys. Chem. Lett.*, 2011, **2**, 2810–2817.
5 T. H. Dunning Jr., K. A. Peterson and A. K. Wilson, *J. Chem. Phys.*, 2001, **114**, 9244–9253.
6 R. Peverati and D. G. Truhlar, *Philos. Trans. R. Soc. Lond. Math. Phys. Eng. Sci.*, 2014, **372**, 20120476.
7 R. A. Kendall, T. H. D. Jr and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796–6806.
8 D. E. Woon and T. H. D. Jr, *J. Chem. Phys.*, 1993, **98**, 1358–1371.
9 A. K. Wilson, D. E. Woon, K. A. Peterson and T. H. D. Jr, *J. Chem. Phys.*, 1999, **110**, 7667–7676.
10 S. Yockel and A. K. Wilson, *Theor. Chem. Acc.*, 2007, **120**, 119–131.
11 M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis and J. A. Montgomery, *J. Comput. Chem.*, 1993, **14**, 1347–1363.
12 M. S. Gordon and M. W. Schmidt, in *Theory and Applications of Computational Chemistry*, eds. G. E. Scuseria, C. E. Dykstra, G. Frenking and K. S. Kim, Elsevier, Amsterdam, 2005, pp. 1167–1189.
13 T. A. Keith, *AIMAll (Version 15.05.18, aim.tkgristmill.com)*, TK Gristmill Software, Overland Park KS, USA, 2015.
14 K. A. Lyssenko, *Mendeleev Commun.*, 2012, **22**, 1–7.
15 E. Espinosa, E. Molins and C. Lecomte, *Chem. Phys. Lett.*, 1998, **285**, 170–173.
16 E. Espinosa, C. Lecomte and E. Molins, *Chem. Phys. Lett.*, 1999, **300**, 745–748.
17 E. Espinosa and E. Molins, *J. Chem. Phys.*, 2000, **113**, 5686–5694.
18 R. F. W. Bader and C. Gatti, *Chem. Phys. Lett.*, 1998, **287**, 233–238.
19 E. Espinosa and E. Molins, *J. Chem. Phys.*, 2000, **113**, 5686–5694.

Data tables

Table S5. Full energy, relevant atomic charges (q), lengths (Å) and QTAIM properties of COOH…X bond (E_{cont} in kcal·mol^{-1} calculated according to the Espinosa-Molins-Lecomte correlation[19]), sum of $\rho(\mathrm{r})$ and E_{cont} for all intermolecular bonding (suffix "total") lengths (Å) and ellipticities of relevant bonds (CC denotes C-CH_3, CO denotes C=O) and atomic source contributions to the C=O (3,-1) critical point for the M11 calculation results. Prefix "0" denotes values for the *cis* conformer, prefix "180" denotes values for *trans*. Row number denotes the specific cis-trans pair; all data with the same row number should be combined into one row to reproduce the calculations discussed in the manuscript.

	Name	0 energy	0 q H	0 q O1	0 q C	0 q O2	0 q CH3	0 q total	0 OH len	0 CO len	0 CC len
1	AcOH	-229.0946	0.6242	-1.1693	1.5997	-1.1953	0.1400	-0.0007	0.9651	1.1996	1.5005
2	AcOH+BF4	-653.7689	0.7065	-1.2390	1.6595	-1.2138	0.0520	-0.9998	0.9930	1.2076	1.5137
3	AcOH+BF4	-653.7689	0.7065	-1.2390	1.6595	-1.2138	0.0520	-0.9998	0.9930	1.2076	1.5137
4	AcOH+Br	-2803.1164	0.6539	-1.2411	1.6686	-1.2144	0.0333	-0.9999	1.0107	1.2077	1.5175
5	AcOH+CF4	-666.6953	0.6343	-1.1762	1.6029	-1.1938	0.1408	0.0010	0.9657	1.2013	1.5000
6	AcOH+Cl	-689.4038	0.6584	-1.2477	1.6760	-1.2201	0.0188	-1.0004	1.0322	1.2102	1.5204
7	AcOH+N3	-393.3629	0.6688	-1.2520	1.6771	-1.2242	0.0178	-0.9997	1.0542	1.2123	1.5205
8	AcOH+N3	-393.3629	0.6688	-1.2520	1.6771	-1.2242	0.0178	-0.9997	1.0542	1.2123	1.5205
9	AcOH+NCH	-322.5217	0.6571	-1.2069	1.6172	-1.2058	0.1168	-0.0001	0.9733	1.2036	1.5035
10	AcOH+NCO	-397.2709	0.6859	-1.2635	1.6821	-1.2280	0.0086	-0.9998	1.0660	1.2140	1.5221
11	AcOH+NO3	-509.5303	0.6982	-1.2537	1.6713	-1.2224	0.0276	-0.9993	1.0295	1.2116	1.5183
12	AcOH+NO3	-509.5303	0.6982	-1.2537	1.6713	-1.2224	0.0276	-0.9993	1.0295	1.2116	1.5183
13	AcOH+OCMe2	-422.2384	0.6713	-1.2331	1.6216	-1.1977	0.1062	-0.0008	0.9817	1.2016	1.5052
14	AcOH+OCMe2	-422.2429	0.6766	-1.2265	1.6269	-1.2167	0.1131	-0.0005	0.9834	1.2081	1.5041
15	AcOH+OH2	-305.5540	0.6719	-1.2164	1.6144	-1.2133	0.1355	0.0001	0.9822	1.2123	1.5005
16	AcOH+OHCH	-343.6058	0.6598	-1.2186	1.6146	-1.1954	0.1179	0.0005	0.9762	1.2004	1.5035

17	AcOH+OHCH	-343.6126	0.6692	-1.2197	1.6203	-1.2167	0.1227	0.0000	0.9828	1.2090	1.5025
18	AcOH+OHCMe	-382.9234	0.6645	-1.2247	1.6184	-1.1971	0.1117	-0.0002	0.9791	1.2011	1.5044
19	AcOH+OHCMe	-382.9301	0.6730	-1.2250	1.6237	-1.2180	0.1174	0.0005	0.9864	1.2095	1.5032
20	AcOH+OMe2	-384.1205	0.6722	-1.2282	1.6199	-1.2139	0.1149	0.0003	0.9903	1.2086	1.5037
21	AcOH+trz	-471.3237	0.6695	-1.2331	1.6283	-1.2199	0.1071	-0.0005	0.9935	1.2100	1.5048

	0 OH ellip	**0 CO ellip**	**0 CC ellip**	**0 hbond Rho**	**0 hbond Econt**	**0 hbond AB length**	**0 Econt total**	**0 Rho total**	**0 sour O5C7**	**0 sour O6**
1	0.0118	0.0808	1.5005	0.0000	0.0000	0.0000	0.0000	0.0000	90.8306	5.2182
2	0.0089	0.0747	1.5137	0.0476	-16.3854	2.6049	-16.3854	0.0476	90.2329	5.7639
3	0.0089	0.0747	1.5137	0.0476	-16.3854	2.6049	-16.3854	0.0476	90.2329	5.7639
4	0.0084	0.0758	1.5175	0.0371	-8.3259	3.1339	-8.3259	0.0371	90.2135	5.7776
5	0.0117	0.0786	1.5000	0.0100	-2.5360	3.1283	-2.5360	0.0100	90.6079	5.2796
6	0.0082	0.0737	1.5204	0.0530	-13.9922	2.9149	-13.9922	0.0530	90.0589	5.9150
7	0.0083	0.0721	1.5205	0.0762	-23.4724	2.5863	-23.4724	0.0762	89.9600	6.0117
8	0.0083	0.0721	1.5205	0.0762	-23.4724	2.5863	-23.4724	0.0762	89.9600	6.0117
9	0.0104	0.0765	1.5035	0.0246	-6.4882	2.9408	-6.4882	0.0246	90.5470	5.4267
10	0.0084	0.0710	1.5221	0.0843	-27.0067	2.5431	-27.0067	0.0843	89.8509	6.1071
11	0.0090	0.0721	1.5183	0.0708	-23.6019	2.5690	-23.6019	0.0708	89.9846	5.9677
12	0.0090	0.0721	1.5183	0.0708	-23.6019	2.5690	-23.6019	0.0708	89.9846	5.9677
13	0.0108	0.0795	1.5052	0.0352	-10.1988	2.7557	-11.7940	0.0425	90.1953	5.6755
14	0.0100	0.0701	1.5041	0.0375	-11.3426	2.7427	-14.2303	0.0509	90.6669	5.3598
15	0.0092	0.0654	1.5005	0.0363	-10.8229	2.7243	-17.8447	0.0633	89.9699	5.7203
16	0.0113	0.0806	1.5035	0.0287	-7.7914	2.7674	-7.7914	0.0287	90.1404	5.6849
17	0.0101	0.0684	1.5025	0.0378	-11.1993	2.7629	-14.1382	0.0518	90.7380	5.2835
18	0.0108	0.0798	1.5044	0.0326	-9.1879	2.7456	-9.1879	0.0326	90.1199	5.7100
19	0.0100	0.0677	1.5032	0.0414	-12.4838	2.7329	-15.3681	0.0552	90.6901	5.3356
20	0.0106	0.0681	1.5037	0.0474	-14.6823	2.6905	-18.2945	0.0651	90.1450	5.6515
21	0.0094	0.0681	1.5048	0.0443	-12.2314	2.7729	-14.6696	0.0562	90.0965	5.7444

	180 energy	180 q H	180 q O1	180 q C	180 q O2	180 q CH3	180 q total	180 OH len	180 CO len	180 CC len	180 O H ellip
1	-229.0865	0.6084	-1.1507	1.6150	-1.1717	0.0990	0.0000	0.9604	1.1924	1.5110	0.0126
2	-653.7791	0.6844	-1.2269	1.6187	-1.2240	0.1168	-1.0016	0.9804	1.2073	1.5087	0.0093
3	-653.7741	0.6930	-1.2297	1.6301	-1.2186	0.0951	-0.9996	0.9822	1.2065	1.5122	0.0097
4	-2803.1262	0.6428	-1.2284	1.6306	-1.2292	0.0797	-0.9984	0.9985	1.2104	1.5115	0.0092
5	-666.6860	0.6175	-1.1579	1.6192	-1.1732	0.0997	-0.0003	0.9607	1.1928	1.5117	0.0124
6	-689.4139	0.6525	-1.2379	1.6366	-1.2352	0.0709	-1.0000	1.0129	1.2129	1.5132	0.0089
7	-393.3707	0.6700	-1.2464	1.6426	-1.2343	0.0655	-1.0001	1.0239	1.2129	1.5154	0.0090
8	-393.3721	0.6683	-1.2456	1.6393	-1.2337	0.0753	-0.9990	1.0205	1.2127	1.5153	0.0091
9	-322.5162	0.6453	-1.1953	1.6276	-1.1857	0.0855	0.0002	0.9677	1.1966	1.5129	0.0110
10	-397.2784	0.6867	-1.2580	1.6486	-1.2375	0.0563	-1.0001	1.0313	1.2143	1.5170	0.0089
11	-509.5362	0.6892	-1.2418	1.6423	-1.2290	0.0712	-1.0000	1.0072	1.2112	1.5157	0.0097
12	-509.5396	0.6874	-1.2450	1.6379	-1.2318	0.0803	-0.9990	1.0048	1.2108	1.5129	0.0088
13	-422.2361	0.6579	-1.2211	1.6248	-1.1899	0.0918	-0.0014	0.9774	1.1975	1.5128	0.0116
14	-422.2361	0.6579	-1.2211	1.6248	-1.1899	0.0918	-0.0014	0.9774	1.1975	1.5128	0.0116
15	-305.5410	0.6533	-1.2065	1.6267	-1.1863	0.0883	0.0010	0.9709	1.1969	1.5129	0.0113
16	-343.6016	0.6473	-1.2078	1.6236	-1.1839	0.0943	0.0002	0.9723	1.1957	1.5121	0.0116
17	-343.6016	0.6473	-1.2078	1.6236	-1.1839	0.0943	0.0002	0.9723	1.1957	1.5121	0.0116
18	-382.9199	0.6518	-1.2125	1.6254	-1.1874	0.0914	0.0002	0.9750	1.1968	1.5127	0.0111
19	-382.9199	0.6518	-1.2125	1.6254	-1.1874	0.0914	0.0002	0.9750	1.1968	1.5127	0.0111
20	-384.1109	0.6603	-1.2115	1.6342	-1.1865	0.0727	-0.0012	0.9756	1.1976	1.5157	0.0115
21	-471.3152	0.6522	-1.2172	1.6286	-1.1935	0.0843	-0.0010	0.9781	1.1990	1.5133	0.0103

	180 CO ellip	180 CC ellip	180 hbond Rho	180 hbond Econt	180 hbond AB.length	180 Econt total	180 Rho total	180 sour O5C7	180 sour O6	Remark
1	0.0982	1.5110	0.0000	0.0000	0.0000	0.0000	0.0000	91.1764	4.9599	
2	0.0763	1.5087	0.0400	-13.3278	2.6616	-19.4175	0.0663	90.4628	5.5760	
3	0.0775	1.5122	0.0426	-14.7564	2.6259	-17.2360	0.0523	90.4197	5.6232	180alt
4	0.0740	1.5115	0.0343	-7.6731	3.1830	-9.4083	0.0431	90.1990	5.7600	
5	0.0980	1.5117	0.0091	-2.3714	3.1847	-5.6152	0.0244	91.1467	4.9912	
6	0.0713	1.5132	0.0472	-12.2992	2.9653	-14.3053	0.0567	90.0747	5.8939	
7	0.0718	1.5154	0.0636	-19.1478	2.6356	-21.4164	0.0733	90.0865	5.9067	
8	0.0719	1.5153	0.0614	-18.3237	2.6496	-18.3237	0.0614	90.0549	5.9323	180alt
9	0.0933	1.5129	0.0236	-6.1924	2.9710	-6.1924	0.0236	90.9146	5.2054	
10	0.0709	1.5170	0.0698	-21.6150	2.5979	-21.6150	0.0698	89.9561	6.0255	
11	0.0732	1.5157	0.0581	-18.6235	2.6207	-21.2639	0.0682	90.1845	5.8063	
12	0.0732	1.5129	0.0577	-18.5595	2.6215	-21.2168	0.0704	90.1416	5.8632	180alt
13	0.0917	1.5128	0.0390	-11.7531	2.7459	-13.5113	0.0466	90.8940	5.2140	
14	0.0917	1.5128	0.0390	-11.7531	2.7459	-13.5113	0.0466	90.8940	5.2140	0alt
15	0.0930	1.5129	0.0312	-9.1488	2.8211	-9.1488	0.0312	90.8988	5.2149	
16	0.0945	1.5121	0.0336	-9.7793	2.7549	-9.7793	0.0336	90.9890	5.1254	
17	0.0945	1.5121	0.0336	-9.7793	2.7549	-9.7793	0.0336	90.9890	5.1254	0alt
18	0.0930	1.5127	0.0374	-11.1873	2.7455	-11.1873	0.0374	90.9208	5.1911	
19	0.0930	1.5127	0.0374	-11.1873	2.7455	-11.1873	0.0374	90.9208	5.1911	0alt
20	0.0931	1.5157	0.0378	-11.5016	2.7515	-11.5016	0.0378	90.8299	5.2863	
21	0.0903	1.5133	0.0356	-9.5022	2.8520	-10.4695	0.0404	90.8037	5.3113	