

Comparing the strength of covalent bonds, intermolecular hydrogen bonds and other intermolecular interactions for organic molecules. X-ray diffraction data and quantum chemical calculations

Angelo Gavezzotti

Dipartimento di Chimica, Università di Milano, Milano, Italy

## SUPPLEMENTARY INFORMATION

### Table of contents

#### **SUPPLEM1 Intramolecular bond lengths**

SUPPLEM1.1 Retrieval of crystal structures from the CSD

SUPPLEM1.2 Average or Gaussian-fit derived best bond distances for intramolecular bonds:

complete schemes, Tables S1-S9 with DDF(R) distribution profiles with Gaussian fitting  
for 82 bond types

1.2.1 The carbonyl carbon

1.2.2 The *sp*-hybridized carbon

1.2.3 The *sp*<sup>2</sup>-hybridized carbon

1.2.4 The *sp*<sup>3</sup>-hybridized carbon

1.2.5 Nitrogen

1.2.6 Sulfur

1.2.7. C-H bond distances from neutron diffraction data

Table S10. Average C-C bond lengths in comparison with older surveys

Figure S1. The relationship between low-T and room-T bond lengths

SUPPLEM1.3 Calculations of bond stretching profiles by ab initio MP2/6-31G\*\*

Figures S2-S8 Bond stretching energy profiles

SUPPLEM1.4 Calculation of approximate vibrational frequencies and vibrational level spacings

from bond stretching force constants

#### **SUPPLEM2. The intermolecular hydrogen bond**

Figures S9-S11 DDF(R) plots for the various types of hydrogen bonding with Gaussian fit

Figures S12-S15. Calculated hydrogen-bond energy profiles for model compounds.

Table S11 Comparison of PIXEL and ab initio results for some hydrogen-bonded dimers

Table S12 Detail of average hydrogen bond lengths with number of data points

#### **SUPPLEM3 The weak bond**

Table S13 Equilibrium distances, binding energies and stretching force constants from cubic  
fitting for all calculated profiles of hydrogen bonds, halogen bonds and weak bonds

Figure S16. DDF(R) functions for the CH...O and CH...N contacts in crystals

Figure S17. Binding energy profiles for some weak-bond dimers

Units are Å for distances, kJ mol<sup>-1</sup> for energies and kJ mol<sup>-1</sup>Å<sup>-2</sup> for force constants.

## SUPPLEM1 Intramolecular bond lengths

### SUPPLEM1.1 Retrieval of crystal structures from the Cambridge Structural Database and preparation of the databases for the present work

A master dataset was created from the CSD under the following conditions: radiation source: X-rays; year: after 1980 (exclude visual/photographic methods); heaviest element Cl;  $Z' \leq 2$  (where  $Z'$  is the number of molecules in the asymmetric unit; this excludes suspicious materials); number of atoms in molecule  $\leq 35$ ; number of residues = 1 (exclude molecular complexes, solvates); average  $\sigma(\text{C-C bonds}) < 0.005$ ; conventional crystallographic R-factor  $\leq 5\%$ ; no disorder, no errors, no polymeric structures, no powder data, only organic compounds. Structures at non-ambient pressure were screened out by negating the text qualifier 'pressure'. This yielded 25423 hits. The CLP modules Retcif, Rtecor, Retcha and Clpcry (see <http://users.unimi.it/gavezzot>, the CLP program suite and its reference manual) check for correct interatomic connectivity (formal atomic valence), for cases with dubious electron counts, or for structures with suspiciously short intermolecular contacts, reducing the size of the dataset to 23248 crystal structures. For each group of duplicate determinations of the same crystal structure, only the one with the lowest R-factor was kept, leaving 21712 crystal structures irrespective of temperature of the X-ray determination (T-ALL data set), or 12532 entries with  $201 < T < 320$  K ("T-ROOM" data set), and 9679 entries with  $0 < T < 200$  K ("T-LOW" data set). The total room-T+low-T (22211) is greater than the number in the all-T database because many structures have been determined both at low and high temperature. With slightly more tolerant criteria, keeping only one best entry for duplicates but without temperature screening, datasets of 1017 bromine-containing structures and 411 iodine-containing structures were retrieved "BR" and "IOD" databases, respectively. Neutron diffraction crystal structure data were retrieved with the limits of  $R < 5\%$ , one residue (excluding salts, where H-atoms in charged species are often in special electronic environments, and hydrates), for 217 hits, reduced to 119 best unique, without temperature screening (the "NEU" database). Unfortunately the number of neutron determinations is very small. CSD Refcode names for all the structures in the databases above are available from the author upon request. Table S1 shows the types of atom considered with their labeling.

The database for hydrogen-bonding studies, the "HBDON" database, was prepared from the T-ALL dataset by requesting the presence of at least one of the hydrogen-bonding donor atoms (H5, H6, H7 or H8 in Table S1). This gave 11429 entries.

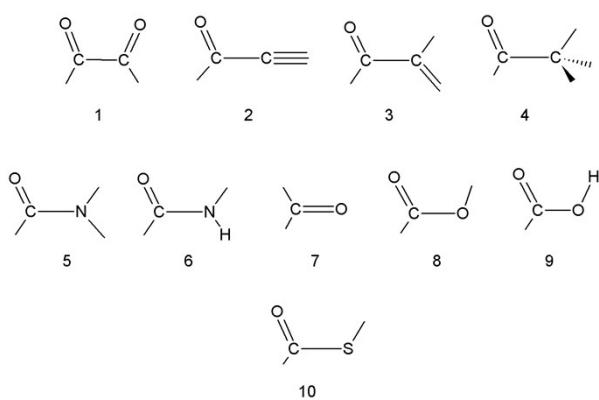
Table S1. Atomic species considered and their indicator label in the data treatment software

| hydrogen  | indicator |   |
|---|-----------|---|
| acetylene $\equiv\text{CH}$   | 1         | to sp carbon                                    |
| $=\text{CH}_2$ , $=\text{CHX}$ , aromatic CH                              | 2         | to $\text{sp}^2$ carbon                         |
| aliphatic CH, $\text{CH}_2$ , $\text{CH}_3$                               | 3         | to $\text{sp}^3$ carbon                         |
| R-OH, R-SH  | 5         | alcohol, thiol                                  |
| COO-H   | 6         | acid  |
| CON)-H  | 7         | amide   |
| $\text{R}_2\text{NH}$ , $\text{RNH}_2$ , $(\text{R}_3\text{N}^+)\text{H}$ | 8         | amine, ammonium                                 |
| carbon  |           |   |
| $\text{C}=\text{(O)}$   | 10        | carbonyl carbon (ketone, aldehyde, acid, amide) |
| $\equiv\text{C}-$   | 11        | sp carbon                                       |
| $\text{sp}^2\text{ C}$  | 12        | ethylenic, allene, aromatic                     |
| $\text{sp}^3\text{ C}$  | 13        |   |
| C(node)   | 14        | aromatic core, e.g. 9-10 in naphthalene         |
| nitrogen  |           |   |

|  |    |   |
|--|----|---|
| (R <sub>n</sub> H <sub>4-n</sub> )N <sup>+</sup> | 16 | ammonium  |
| (R <sub>n</sub> H <sub>3-n</sub> )N              | 17 | sp <sup>3</sup> nitrogen                        |
| arom.N, R=N(H)                                   | 18 | sp <sup>2</sup> nitrogen                        |
| -C≡N,-N=N  | 19 |   |
| nitro N  | 20 |   |
| amide N (CONH,CONH <sub>2</sub> )                | 21 |   |
| <b>oxygen</b>                                    |    |   |
| -O-  | 23 | ether   |
| C=O, COO <sup>-</sup>                            | 27 | carbonyl oxygen (ketone, aldehyde, acid, amide) |
| (C=O)-OH   | 28 | hydroxyl oxygen in acids                        |
| R-OH   | 29 | alcohol oxygen                                  |
| N=O  | 30 |   |
| S=O  | 31 |   |
| P=O  | 32 |   |
| <b>sulfur</b>                                    |    |   |
| -S-  | 34 | thioether                                       |
| (C)=S  | 35 | sulfide   |
| (O)=S  | 36 | sulfone, sulfoxide, sulfonate                   |
| R-S(H)   | 37 | thiol   |
| <b>heteroatoms</b>                               |    |   |
| F  | 41 |   |
| Cl   | 42 |   |
| Br   | 43 |   |
| I  | 44 |   |

**SUPPLEM1.2 Average or Gaussian-fit derived best bond distances for intramolecular bonds.** For each set of distances, a Scheme with indication of the bond under consideration (the central horizontal one), a Table of bond length values, and some examples of the N(R) profiles with Gaussian fit and difference between experitnal N(R) and Gaussian fit. T-ROOM database, 201<T<320, 12532 entries.

### SUPPLEM1.2.1 The carbonyl carbon



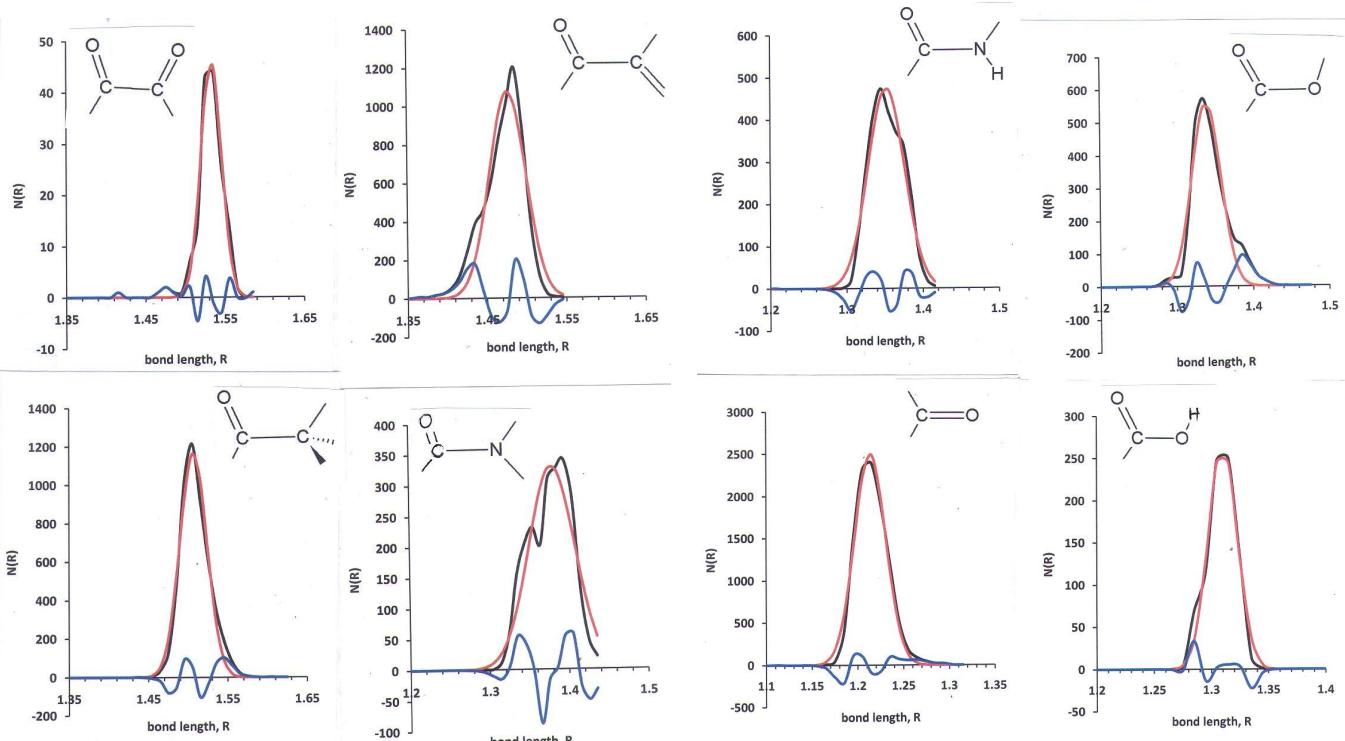
| Table S2      |            |                      |                           |                             |                 |
|---------------|------------|----------------------|---------------------------|-----------------------------|-----------------|
| bond scheme   | total hits | average <sup>a</sup> | gaussian fit <sup>b</sup> | linkage number <sup>c</sup> | species Table 1 |
| (O=)C – C(=O) | 157        | 1.531(19)            | 1.532(13)                 | 1                           | 10-10           |

|   |       |           |                          |    |       |
|---|-------|-----------|--------------------------|----|-------|
| (O=)C – C( <i>sp</i> <sup>2</sup> )         | 6423  | 1.470(26) | 1.447(20)sh<br>1.485(15) | 3  | 10-12 |
| (O=)C – C( <i>sp</i> <sup>3</sup> )         | 5075  | 1.509(18) | 1.507(17)                | 4  | 10-13 |
| (O=)C – N(X <sub>2</sub> )                  | 2358  | 1.378(28) | 1.349(16)sh<br>1.391(17) | 5  | 10-17 |
| (O=)C – N(HX),N(H <sub>2</sub> ), C-N amide | 2870  | 1.353(23) | 1.352(24)                | 6  | 10-21 |
| (O=)C – OX<br>C-O ester                     | 2836  | 1.345(23) | 1.337(16)<br>1.379(16)sh | 8  | 10-23 |
| C=O carbonyl (any)                          | 10677 | 1.216(18) | 1.214(17)                | 7  | 10-27 |
| (O=)C – OH,C-O carboxylic acid              | 831   | 1.308(12) | 1.310(13)                | 9  | 10-28 |
| (O=)C-SH, C-S thioester                     | 151   | 1.774(16) | 1.773(14)                | 10 | 10-34 |

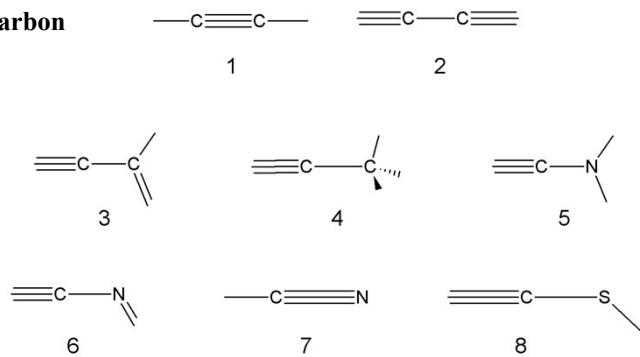
<sup>a</sup>Average of all data, irrespective of profile shape. X is any non-hydrogen atom, mostly carbon. <sup>b</sup>Profile fit by gaussian(s); more than one entry means multimodal profile. 'sh' indicates that the feature is a shoulder, not a separate peak. <sup>c</sup>Entries in parentheses are from scattered or too few data for significant statistics. <sup>d</sup>See Scheme.

Profiles of the N(R) functions in Table S2.

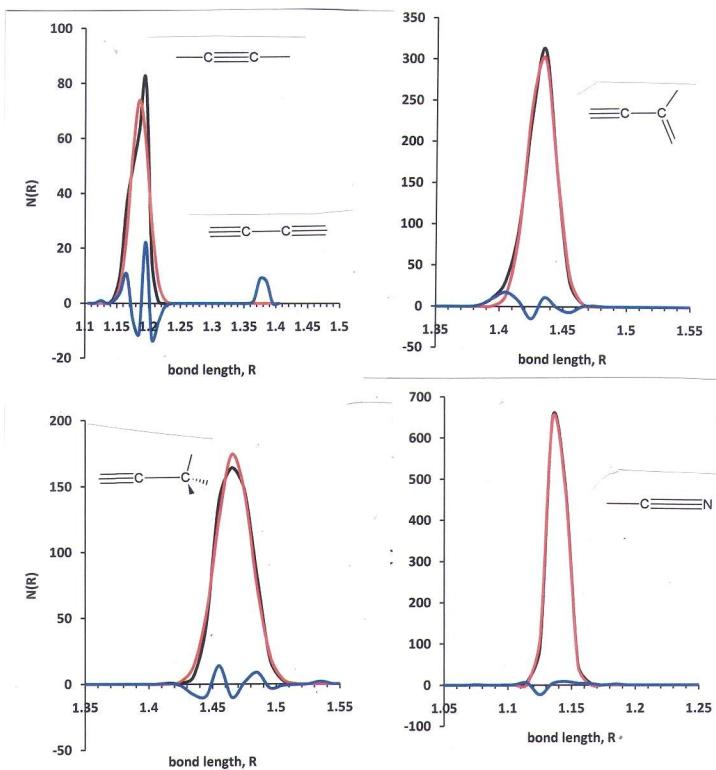
Black: experimental, red: Gaussian fitting, blue: difference exptl-fit.



**SUPPLEM1.2.2 The *sp*-hybridized carbon**



| Table S3                        | total hits | average        | gaussian fit | linkage number | species (Table 1) |
|---------------------------------|------------|----------------|--------------|----------------|-------------------|
| $\text{C}\equiv\text{C}$        | 260        | 1.183(13)      | -            | 1              | 11-11             |
| $\equiv\text{C---C}\equiv$      | 18         | [1.38]         |              | 2              |                   |
| $\text{C}(sp)---\text{C}(sp^2)$ | 877        | 1.431(12)      | 1.433(11)    | 3              | 11-12             |
| $\text{C}(sp)---\text{C}(sp^3)$ | 609        | 1.467 (14)     | 1.467(14)    | 4              | 11-13             |
| $\text{C}(sp)---\text{N}(sp^3)$ | 10         | [1.34]         |              | 5              | 11-17             |
| $\text{C}(sp)---\text{N}(sp^2)$ | 37         | [1.27 scatter] |              | 6              | 11-18             |
| $\text{C}\equiv\text{N}$        | 1287       | 1.139 (7)      | 1.139(7)     | 7              | 11-19             |
| $\text{C}(sp)---\text{S}(X)$    | 26         | [1.69]         |              | 8              | 11-34             |



### SUPPLEM1.2.3 The sp<sup>2</sup>-hybridized carbon

Number 4 is for the central bond in naphthalene rings or for the C(sp<sup>2</sup>)-benzene bond.

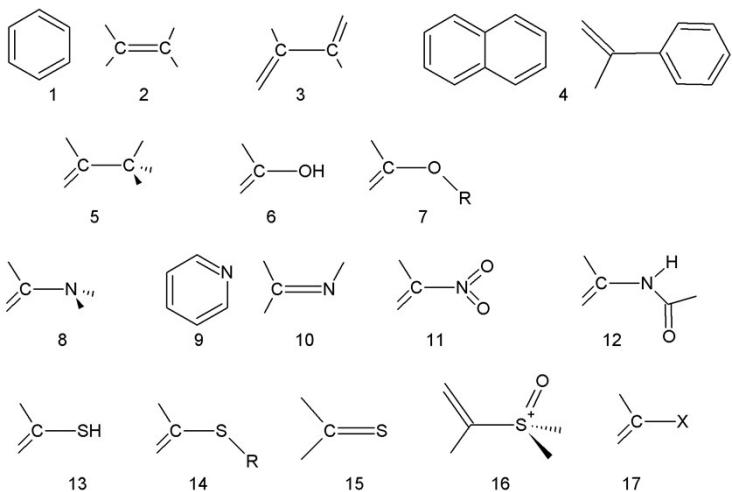
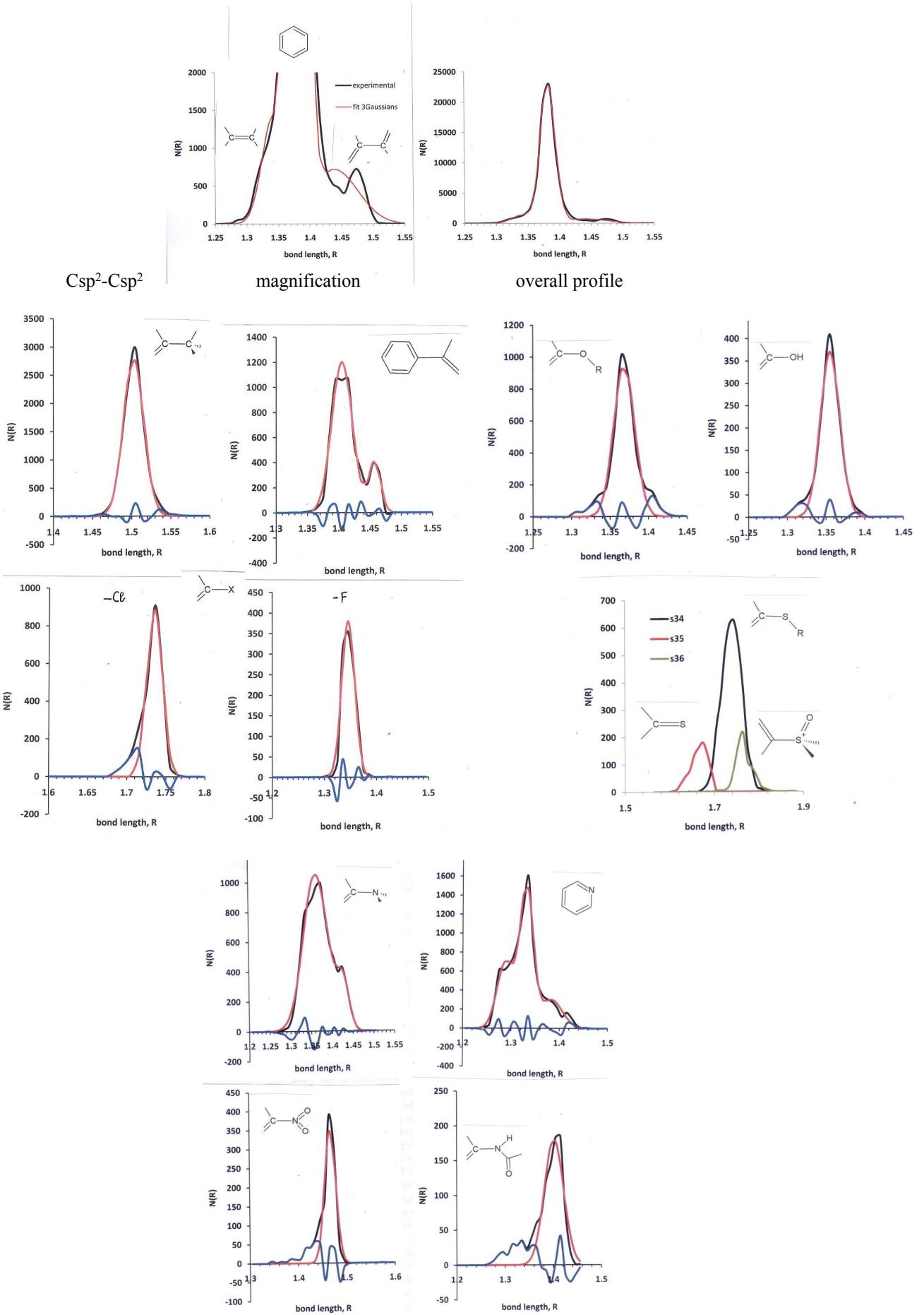
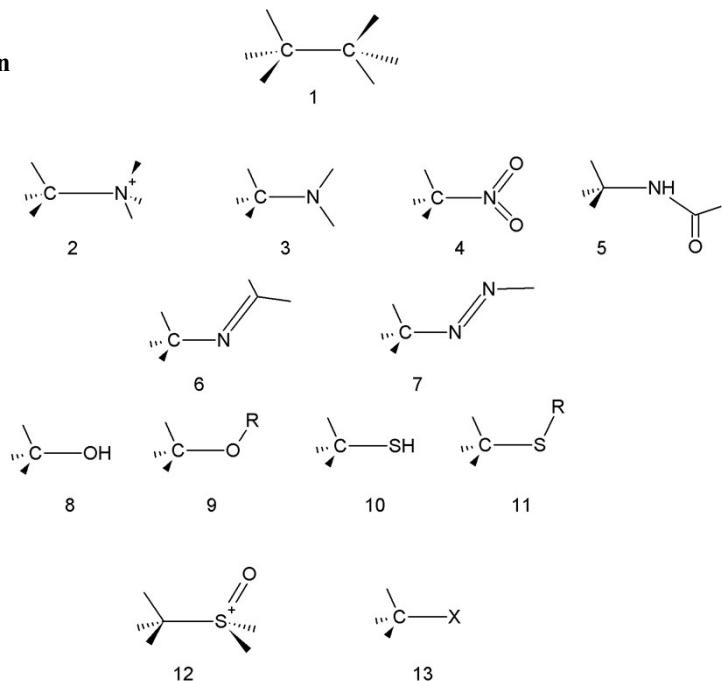


Table S4.

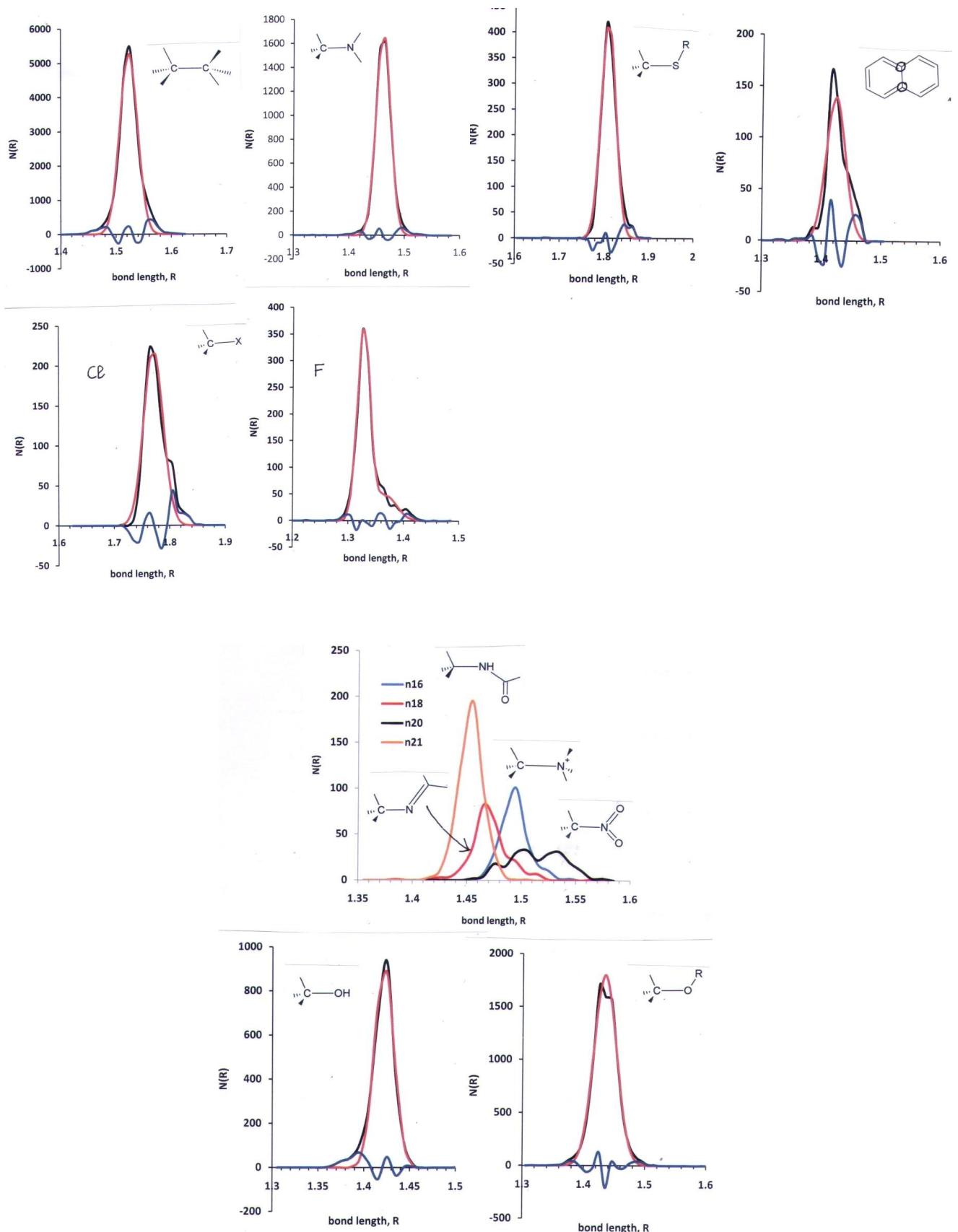
| bond type  | total hits | average    | gaussian fit                        | linkage number | species Table 1 |
|--|------------|------------|-------------------------------------|----------------|-----------------|
| C(sp <sup>2</sup> )-C(sp <sup>2</sup> )<br>>C=C<<br>--C---C-- aromatic<br>=C - C= single       | 79737      | -          | 1.342(18)<br>1.382(12)<br>1.439(38) | 2<br>1<br>3    | 12-12           |
| C(sp <sup>2</sup> )-C(sp <sup>2</sup> ,node)   | 5795       | 1.413 (24) | -                                   | 4              | 12-14           |
| C(sp <sup>2</sup> )-C(sp <sup>3</sup> )  | 9436       | 1.503(15)  | 1.503(13)                           | 5              | 12-13           |
| C(sp <sup>2</sup> )-N(sp <sup>3</sup> )<br>C-N(H <sub>2</sub> , RH, R <sub>2</sub> )<br>Ph - N | 8687       | 1.371(35)  | 1.363(30)<br>1.428(14)sh            | 8              | 12-17           |
| C(sp <sup>2</sup> )-N(sp <sup>2</sup> )<br>>C=N- double<br>aromatic -N---N-                    | 9813       | -          | 1.289(18)<br>1.331(14)              | 10<br>9        | 12-18           |
| C(sp <sup>2</sup> )-N(nitro)<br>mostly Ph-NO <sub>2</sub>                                      | 1227       | 1.457 (23) | 1.466(12)                           | 11             | 12-20           |
| C(sp <sup>2</sup> )-N(amide)   | 1078       | 1.385 (35) | 1.402(20)                           | 12             | 12-21           |
| C(sp <sup>2</sup> )-OX ether   | 3600       | 1.369 (19) | 1.369(14)                           | 7              | 12-23           |
| C(sp <sup>2</sup> )-OH<br>mostly phenols   | 1257       | 1.354 (15) | 1.356(13)                           | 6              | 12-29           |
| C(sp <sup>2</sup> )-SX thioether   | 3691       | 1.739 (23) | 1.740(22)                           | 14             | 12-34           |
| C(sp <sup>2</sup> )=S  | 859        | 1.665 (19) | 1.669(19)                           | 15             | 12-35           |
| C(sp <sup>2</sup> )-S(=O <sub>x</sub> )  | 840        | 1.765(22)  | 1.764(17)                           | 16             | 12-36           |
| C(sp <sup>2</sup> )-F  | 1148       | 1.347 (12) | 1.346(13)                           | 17             | 12-41           |
| C(sp <sup>2</sup> )-Cl   | 2465       | 1.731 (14) | 1.735(10)                           | 17             | 12-42           |
| C(sp <sup>2</sup> )-Br   | 1038       | 1.892(14)  | -                                   | 17             | 12-43           |
| C(sp <sup>2</sup> )-I  | 400        | 2.095(15)  | -                                   | 17             | 12-44           |
| C(node)-C(node)  | 601        | 1.425 (19) | 1.423(16)                           | 4              | 14-14           |



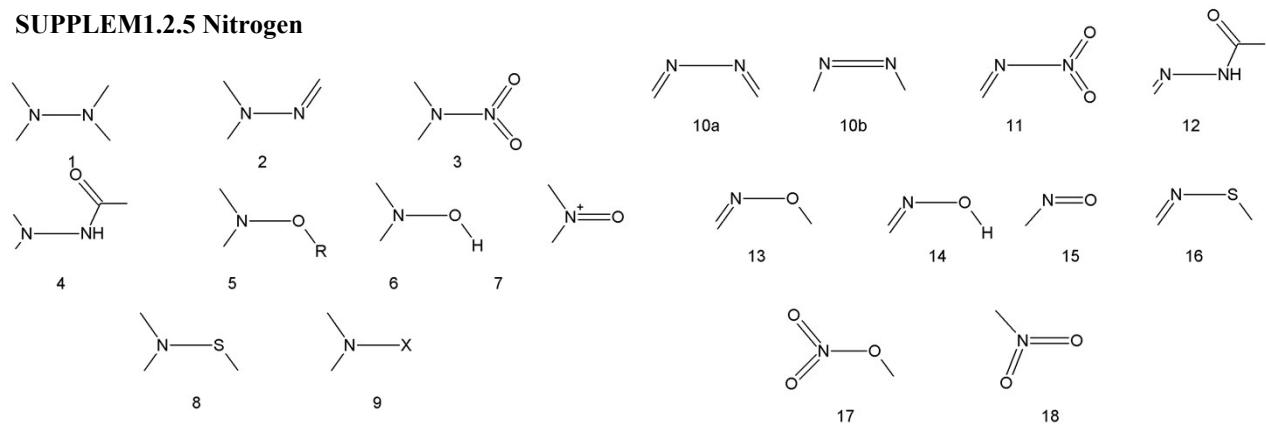
**SUPPLEM1.2.4 The  $sp^3$ -hybridized carbon**



| Table S5.  | total hits | average    | gaussian fit             | linkage number | species Table 1 |
|--|------------|------------|--------------------------|----------------|-----------------|
| C( $sp^3$ )– C( $sp^3$ )                           | 22828      | 1.523(21)  | 1.523(16)                | 1              | 13-13           |
| C( $sp^3$ )– N <sup>+</sup> ( $sp^3$ ) ammonium    | 289        | 1.494(13)  | 1.492(11)                | 2              | 13-16           |
| C( $sp^3$ )– N( $sp^3$ ) C-N( $X_2$ )              | 5744       | 1.462 (16) | 1.461(13)                | 3              | 13-17           |
| C( $sp^3$ )– N( $sp^2$ )                           | 275        | 1.471 (21) | 1.469(13)                | 6,7            | 13-18           |
| C( $sp^3$ )–N(nitro)                               | 218        | 1.514 (25) | 1.514(27)                | 4              | 13-20           |
| C( $sp^3$ )–N(amide)                               | 531        | 1.454(12)  | 1.454(11)                | 5              | 13-21           |
| C( $sp^3$ )– OX ether                              | 8375       | 1.435 (20) | 1.435(19)                | 9              | 13-23           |
| C( $sp^3$ )– OH alcohol                            | 2571       | 1.420 (14) | 1.422(11)                | 8              | 13-29           |
| C( $sp^3$ )– SX thioether                          | 1756       | 1.812 (19) | 1.809(17)                | 11             | 13-34           |
| C( $sp^3$ )– S(=O <sub>x</sub> ) sulfone,sulfoxide | 621        | 1.794(35)  | 1.749(6)<br>1.795(32)    | 12             | 13-36           |
| C( $sp^3$ )– F main peak F-cycloalkanes            | 1316       | 1.336 (26) | 1.328(12)<br>1.367(21)sh | 13             | 13-41           |
| C( $sp^3$ )– Cl                                    | 993        | 1.776 (20) | 1.771(18)                | 13             | 13-42           |
| C( $sp^3$ )– Br                                    | 304        | 1.960(20)  | -                        | 13             | 13-43           |
| C( $sp^3$ )– I                                     | 87         | 2.164(29)  | -                        | 13             | 13-44           |



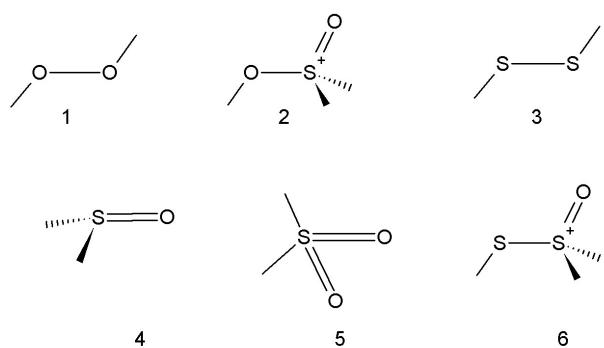
### SUPPLEM1.2.5 Nitrogen



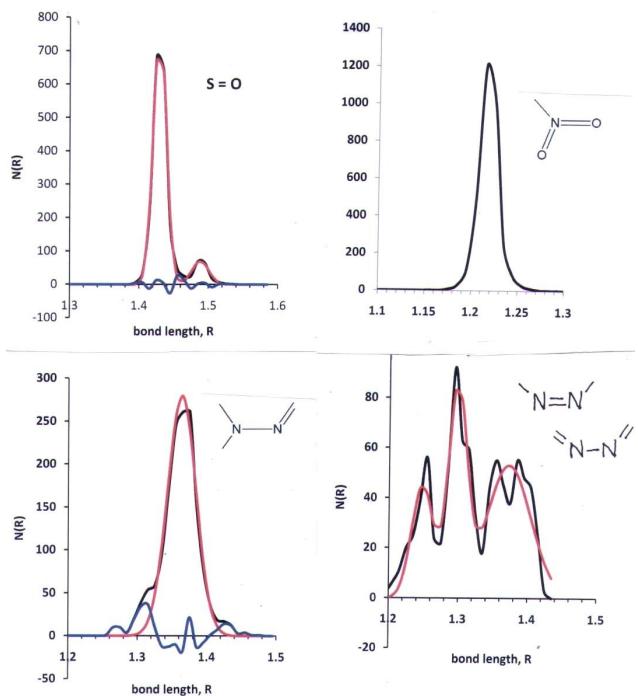
| Table S6.  | total hits | average    | Gaussian fit | linkage number | species Table 1 |
|--|------------|------------|--------------|----------------|-----------------|
|  |            |            |              |                |                 |
| $\text{N}(sp^3)-\text{N}(sp^3)$<br>$>\text{N}-\text{N}<$ | 223        | 1.400 (31) | 1.403(17)    | 1              | 17-17           |
| $\text{N}(sp^3)-\text{N}(sp^2)$<br>$>\text{N}-\text{N}=$ | 1557       | 1.361 (29) | 1.364(21)    | 2              | 17-18           |
| $\text{N}(sp^3)-\text{N}(nitro)$                         | 109        | 1.376 (30) | -            | 3              | 17-20           |
| $\text{N}(sp^3)-\text{N}(amide)$                         | 78         | 1.398 (28) | -            | 4              | 17-21           |
| $\text{N}(sp^3)-\text{O}X$                               | 129        | 1.427 (38) | -            | 5              | 17-23           |
| $\text{N}(sp^3)-\text{O}(\text{H})$                      | 68         | 1.399 (29) | -            | 6              | 17-29           |
| $\text{N}(sp^3)\rightarrow\text{O}$                      | 165        | 1.280 (34) | -            | 7              | 17-30           |

| Table S7   | total hits | average    | Gaussian fit                        | linkage number    | species Table 1 |
|--|------------|------------|-------------------------------------|-------------------|-----------------|
|  |            |            |                                     |                   |                 |
| $\text{N}(sp^2)-\text{N}(sp^2)$<br>$=\text{C}-\text{N}=\text{N}-\text{C}=$<br>$-\text{N}=\text{N}-$ double<br>$=\text{N}-\text{N}=$ single | 882        | -          | 1.249(16)<br>1.299(14)<br>1.373(32) | 10b<br>10b<br>10a | 18-18           |
| $\text{N}(sp^2)=\text{N}$ ( <i>terminal</i> )<br>$=\text{N}=\text{N}$  | 99         | 1.119 (12) | -                                   |                   | 18-19           |
| $\text{N}(sp^2)-\text{N}(amide)$   | 293        | 1.372 (29) | -                                   | 12                | 18-21           |
| $\text{N}(sp^2)-\text{O} =\text{N}-\text{OX}$  | 322        | 1.399 (25) | -                                   | 13                | 18-23           |
| $=\text{N}-\text{OH}$  | 162        | 1.395 (16) | -                                   | 14                | 18-29           |
| $-\text{N}=\text{O}$ nitro group   | 3204       | 1.218(12)  | 1.218(10)                           | 18                | 20-30           |
| $\text{N}(sp^2)-\text{SX}$   | 264        | 1.631(45)  | -                                   | 16                | 18-34           |

**SUPPLEM1.2.6 Sulfur**



| Table S8                                     | total hits | average        | Gaussian fit | linkage number | species Table 1 |
|--|------------|----------------|--------------|----------------|-----------------|
| -O-O- peroxide                               | 24         | [1.48]         | -            | 1              | 23-23           |
| -O-S(O <sub>2</sub> ), -O-S=O sulfonate      | 125        | [1.60 scatter] | -            | 2              | 23-36           |
| >S-O   | 2015       | 1.436(21)      |              |                | 31-36           |
| >S=(O <sub>2</sub> )                         |            |                | 1.429(10)    | 5              |                 |
| >S=O   |            |                | 1.489(12)    | 4              |                 |
| -S-S-  | 220        | 2.054 (29)     | -            | 3              |                 |
| -S-S(O <sub>2</sub> ), -S-S(O <sub>3</sub> ) | 21         | [2.10]         | -            | 6              |                 |



**SUPPLEM1.2.7. Average bond distances involving hydrogen, from neutron diffraction data.**

| Table S9                            | total hits | average   |
|-------------------------------------|------------|-----------|
| C(sp <sub>2</sub> )–H               | 307        | 1.077(31) |
| C(sp <sub>3</sub> )–H               | 670        | 1.085(26) |
| R–O–H                               | 93         | 0.975(27) |
| (CO)–O–H                            | 11         | 1.015     |
| (CO)–N–H                            | 37         | 1.018     |
| N(sp <sub>3</sub> ) <sup>+</sup> –H | 29         | 1.034(14) |
| N(sp <sub>3</sub> )–H               | 40         | 1.018(14) |

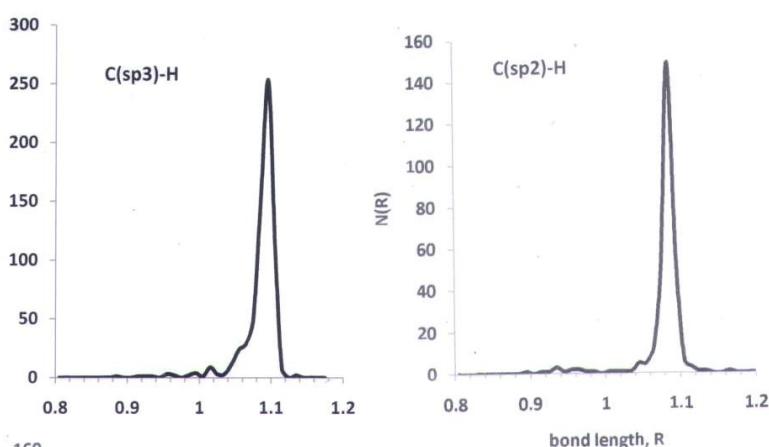


Table S10. Average C–C bond lengths in comparison with older surveys. A.G.Orpen, L.Brammer, F.H.Allen, O.Kennard, D.G.Watson and R.Taylor, in Structure Correlation, H.-B.Buerger and J.D.Dunitz, Eds., VCH, Weinheim 1994, vol. 2, Appendix A.

| experimental X-ray, present            |                                   |                                   | 1994 survey <sup>a</sup> |
|--|-----------------------------------|-----------------------------------|--------------------------|
| bond type                              | R <sub>G</sub> or R <sub>Av</sub> | σ <sub>G</sub> or σ <sub>Av</sub> |                          |
| -C≡C-                                  | 1.183                             | 13                                | 1.183                    |
| >C=C<                                  | 1.342                             | 18                                | 1.331                    |
| ≡C – C≡                                | 1.378                             | 6                                 | 1.377                    |
| C <sub>arom</sub> ---C <sub>arom</sub> | 1.382                             | 12                                | 1.380                    |
| ≡C – Csp <sup>2</sup>                  | 1.433                             | 11                                | 1.431                    |
| Csp <sup>2</sup> – Csp <sup>2</sup>    | 1.439                             | 38                                | 1.460                    |
| ≡C – Csp <sup>3</sup>                  | 1.467                             | 14                                | 1.466                    |
| Csp <sup>2</sup> – Csp <sup>3</sup>    | 1.503                             | 13                                | 1.507                    |
| Csp <sup>3</sup> – Csp <sup>3</sup>    | 1.523                             | 16                                | 1.530                    |

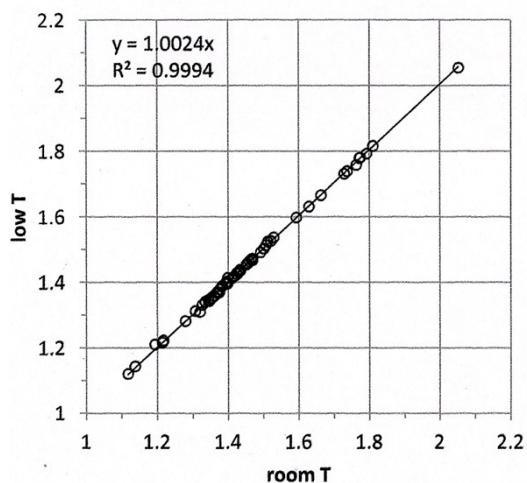


Figure S1. The relationship between low-T and room-T experimental determination. Low-T bond lengths are slightly longer due to a systematic error in the treatment of X-ray data.

### SUPPLEM1.3 Calculations of bond stretching profiles by ab initio MP2/6-31G\*\*

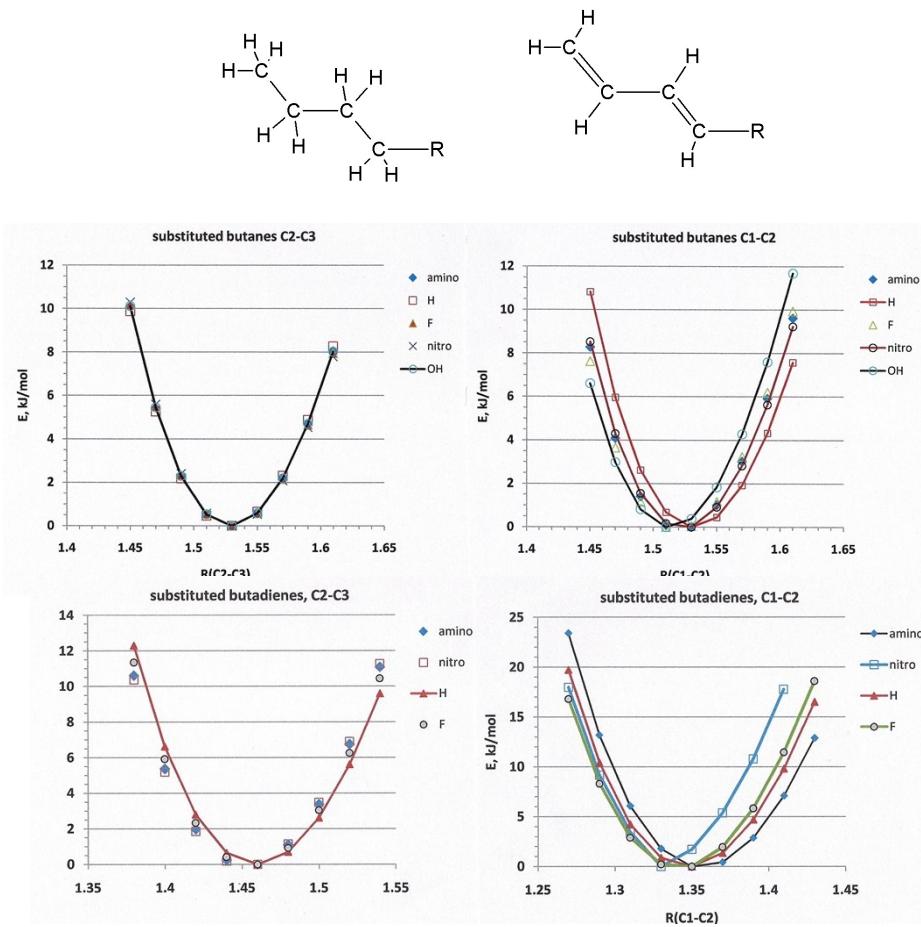


Figure S2. Bond stretching profiles for carbon-carbon bonds in 1-substituted butanes and butadienes. The C2-C3 central bond is in both cases quite insensitive to substitution. The C1-C2 bond, close to the substitution point, is marginally sensitive to substitution, but without a clear trend as a function of the electronic character of the substituent. All curves are strictly harmonic within 0.1 Å from the minimum. Note the different energy scale for the C1-C2 (double) bond in butadiene.

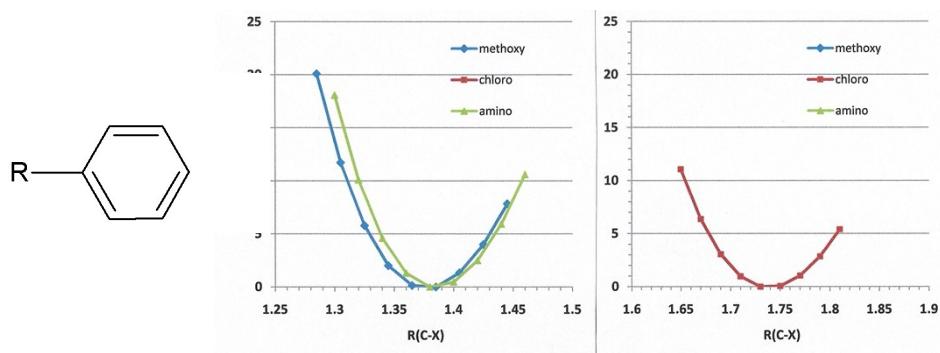


Figure S3. Bond stretching profiles for the C-substituent bond in substituted benzenes. All curves are strictly harmonic within 0.1 Å from the minimum.

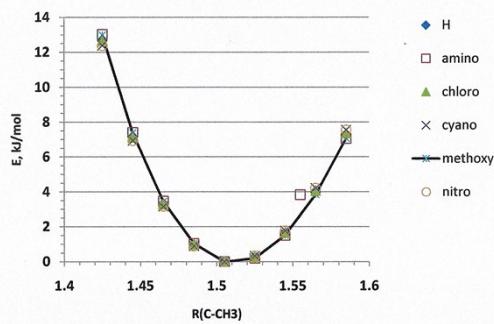
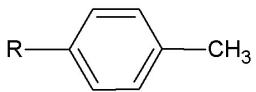


Figure S4. Bond stretching profiles for the C-CH<sub>3</sub> bond in p-substituted toluenes. Bond length and force constant are insensitive to substitution. All curves are strictly harmonic within 0.1 Å from the minimum.

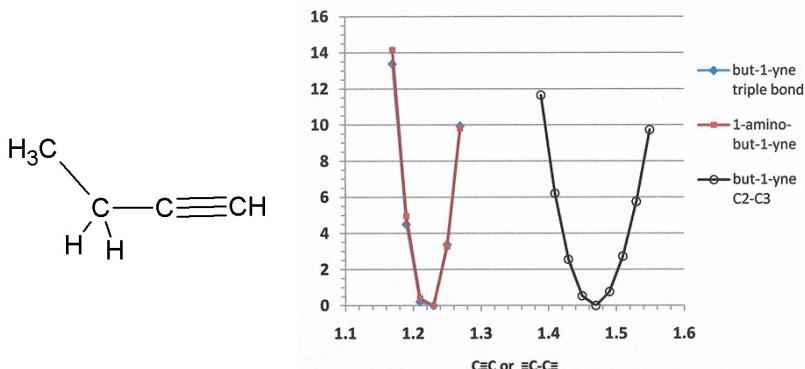


Figure S5. Bond stretching profiles in acetylenic compounds. The profile for triple bond is extremely narrow as expected, and is insensitive to substitution.

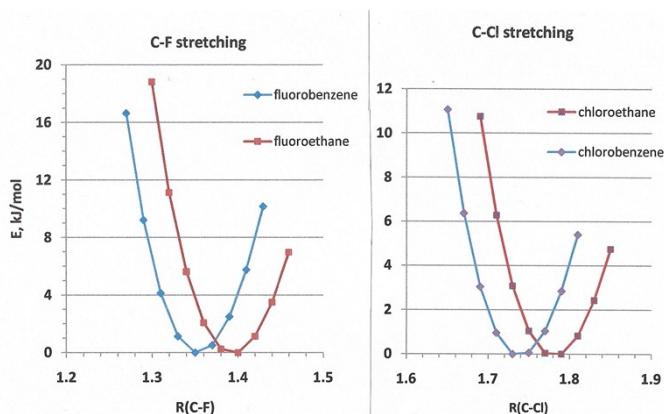


Figure S6. Bond stretching profiles for carbon-halogen bonds.

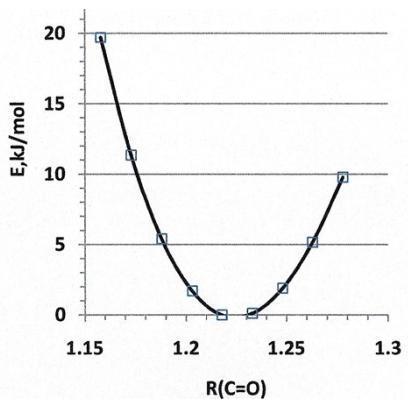
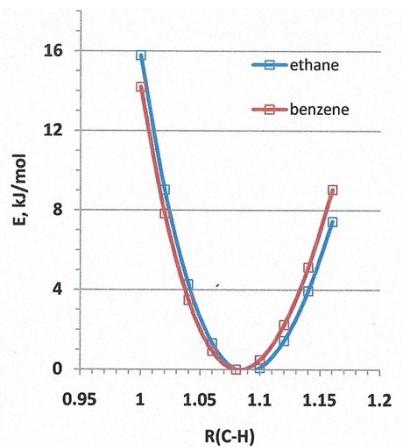


Figure S7. Bond stretching profile for the C=O bond in acetone.

Figure S8. Bond stretching profiles for C-H bonds.



#### SUPPLEM1.4 Calculation of approximate vibrational frequencies and vibrational level spacings from bond stretching force constants

In a first approximation the bond stretching can be assimilated to that occurring in a diatomic molecule:

$$\nu(\text{Hz}) = 1/(2\pi) [k \cdot 10^{23}/\mu]^{1/2} ; \quad \Delta E = h\nu$$

with  $k$  in  $\text{kJ}/(\text{mol } \text{\AA}^2)$  and the reduced mass  $\mu$  in kg. Assuming as a reference the ethane molecule as a diatomic molecule formed by two methyl groups,  $m = 0.0075$  kg and for  $k = 3000$  one gets  $\nu = 3 \cdot 10^{13}$  Hz, or approximately a wavenumber of  $1000 \text{ cm}^{-1}$ , and  $\Delta E = 12 \text{ kJ/mol}$ .

#### SUPPLEM2. The intermolecular hydrogen bond

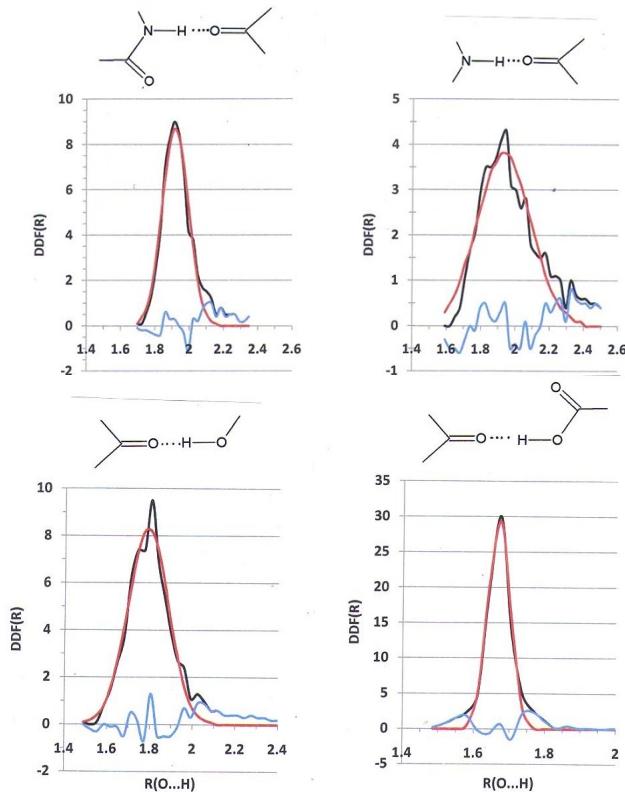
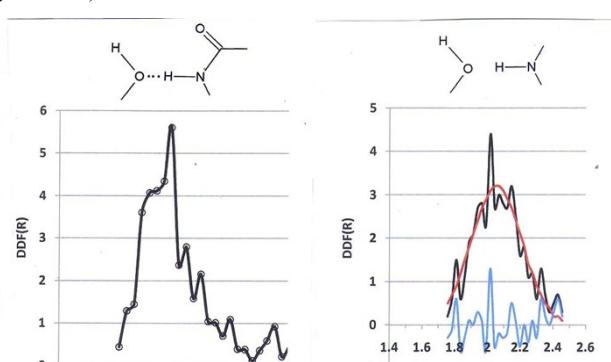


Figure S9. Distance density function (DDF) plots, with Gaussian fitting, for the distribution of hydrogen bonding distances of the carbonyl group with several donors. Black, experimental DDF; red: Gaussian fitting, blue: difference. Dangling bonds are to non-H atoms (mostly carbon).



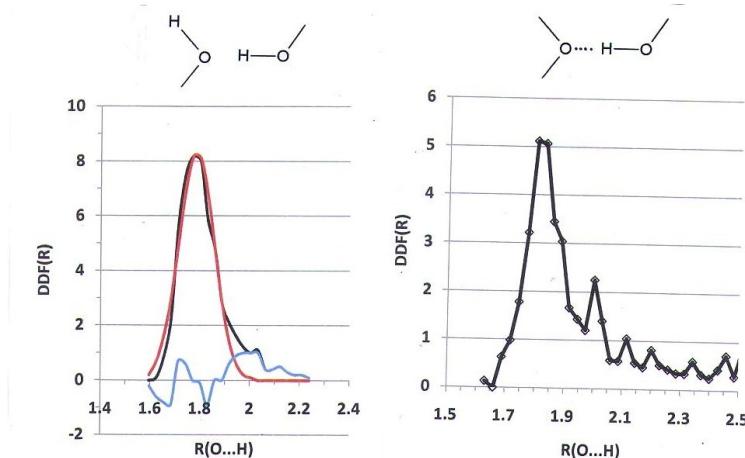


Figure S10. Distance density plots for the distribution of hydrogen bonding distances of the alcohol or ether group with several donors. Black, experimental DDF, red: Gaussian fitting, blue: difference. No Gaussian fitting when the distributions are multimodal and/or too noisy because of the small number of data. Dangling bonds are to non-H atoms (mostly carbon).

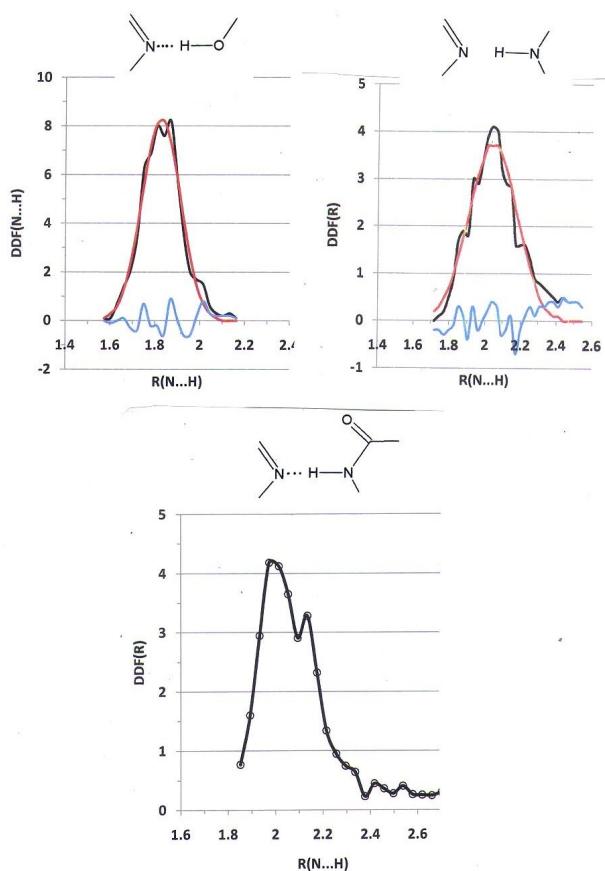
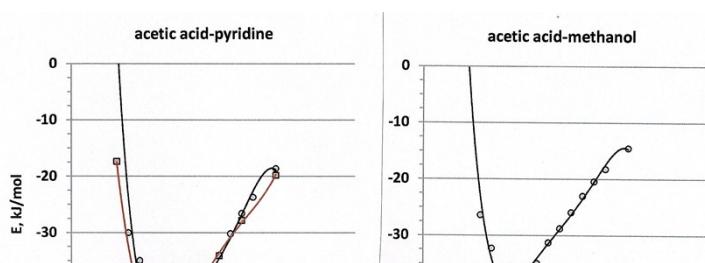
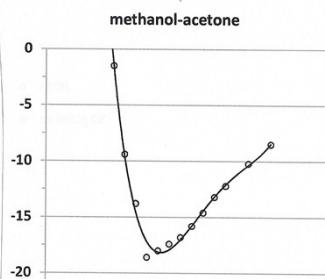
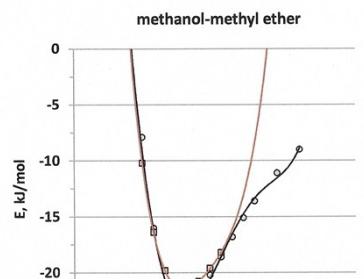
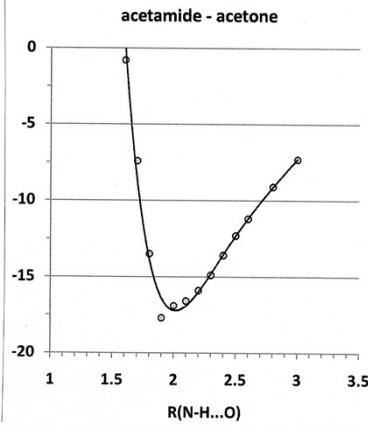
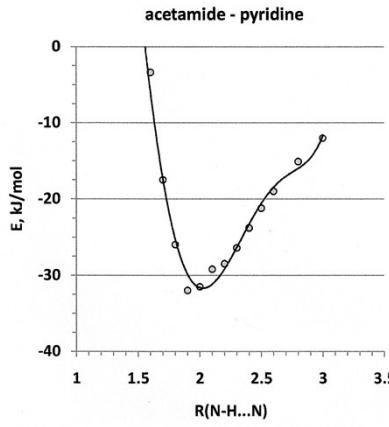
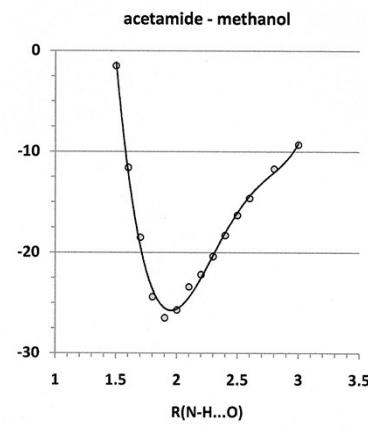
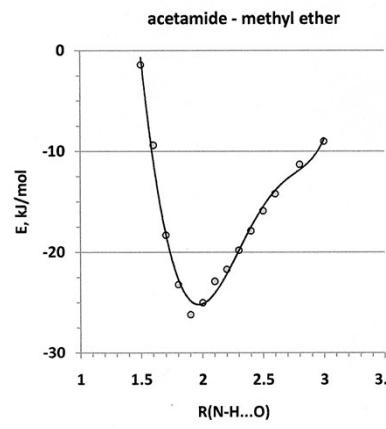
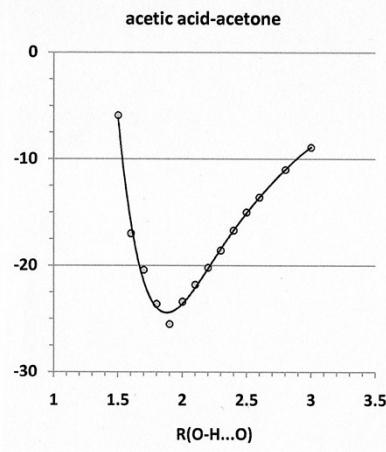
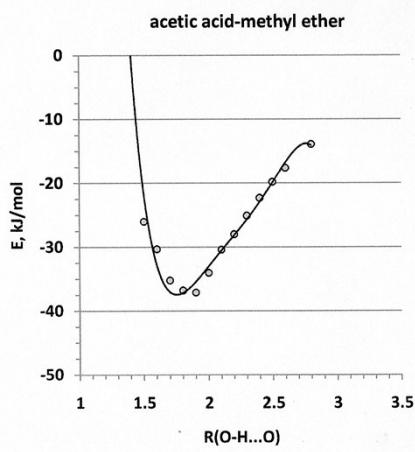


Figure S11. Distance density plots for the distribution of some N-H...O or N-H...N hydrogen bonding distances. Black, experimental DDF, red: Gaussian fitting, blue: difference. No Gaussian fitting when the distributions are multimodal and/or too noisy because of the small number of data. Dangling bonds are to non-H atoms (mostly carbon).

Figures S12-S15. Calculated hydrogen-bond energy profiles for model compounds.





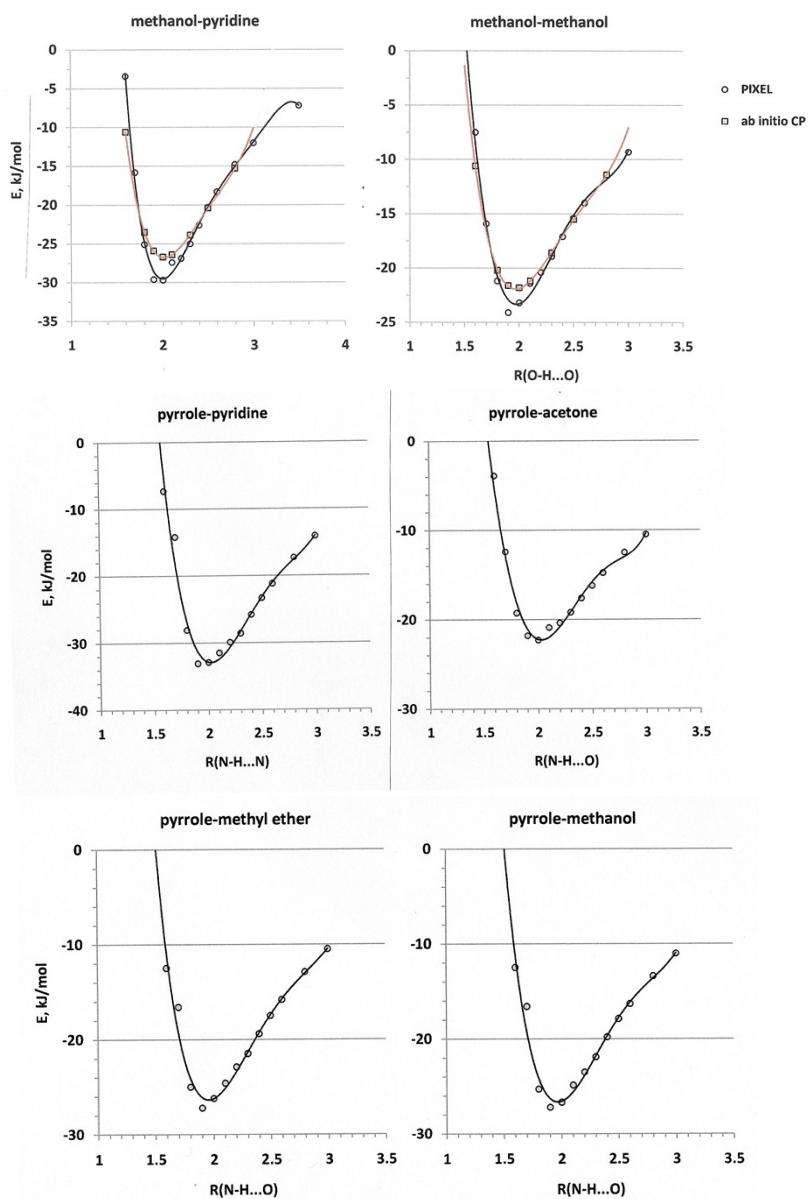


Table S11. First line: experimental top-frequency bond distance for common types of hydrogen bonding; 11429 crystal structures in the "HB-donor" data set. Second line: equilibrium distance, cohesive energy in  $\text{kJ mol}^{-1}$  and stretching force constant calculated by PIXEL for pairing of the indicated model molecules. Third line: same by ab initio MP2/6-31G\*\*

|                                      | O-H alcohol<br>exptl.<br>(methanol)     | COO-H acid<br>exptl.<br>(acetic acid) |
|--------------------------------------|---|---------------------------------------|
| acceptor (exptl)<br>(model molecule) |   |                                       |
| =N-or --N--arom<br>(pyridine)        | 1.826(86)<br>2.00 30 214<br>2.02 27 125 | -<br>1.85 49 342<br>1.82 47 227       |
| - O - ether<br>(dimethyl ether)      | 1.838(79)<br>1.97 23 146<br>1.95 22 125 | -<br>1.82 39 283                      |
| (R)O(H) alcohol<br>(methanol)        | 1.784(71)<br>1.92 25 241<br>1.95 23 152 | -<br>1.82 40 286                      |

Table S12. Experimental top-frequency bond distance for common types of hydrogen bonding; 11429 crystal structures in the "HB-donor" data set, with the number of hits for each type.

|                 | donor                                   |                |                         |                    |
|-----------------|---|----------------|-------------------------|--------------------|
|                 | O-H alcohol                             | COO-H acid     | CON-H amide             | R <sub>2</sub> N-H |
| acceptor        |   |                |                         |                    |
| =N-or --N--arom | 1.826(86) <sup>a</sup> 893 <sup>c</sup> | -              | [1.98] <sup>b</sup> 394 | 2.041(134) 2150    |
| - O - ether     | 1.838(79) 751                           | -              | -                       | -                  |
| (R)O(H) alcohol | 1.784(71) 2987                          | -              | [1.97] 326              | 2.055(155) 584     |
| (C)=O carbonyl  | 1.790(95) 2046                          | 1.669(31) 1171 | 1.918(77) 3143          | 1.933(157) 2150    |

<sup>a</sup> $R_G$  and  $\sigma_G$  from a Gaussian fit of the DDF. <sup>b</sup>From visual estimation from the graph (too few data for Gaussian fitting). <sup>c</sup> Number of data points.

### SUPPLEM3 The weak bond

Table S13 Equilibrium distances, binding energies and stretching force constants from cubic fitting for all calculated profiles of hydrogen bonds, halogen bonds and weak bonds. Å, kJ/mol and kJ/(mol Å<sup>2</sup>) units.

| pair                                      | $R^\circ$ | $E^\circ$ | k     |
|---|-----------|-----------|-------|
| <b>hydrogen bonds</b>                     |           |           |       |
| methanol don.-pyridine accept. pix        | 2.002     | -30.0     | 214.2 |
| methanol don.-pyridine accept. ab in      | 2.020     | -27.0     | 124.6 |
| methanol don.-methyl ether accept. pix    | 1.966     | -23.2     | 146.0 |
| methanol don.-methyl ether accept. ab in  | 1.948     | -21.7     | 124.6 |
| methanol don.-methanol accept. pix        | 1.923     | -25.0     | 241.0 |
| methanol don.-methanol accept. ab in      | 1.949     | -22.6     | 151.6 |
| methanol don.-acetone accept. pix         | 2.018     | -18.4     | 125.4 |
| acetic acid don.-pyridine accept. pix     | 1.850     | -49.3     | 342.8 |
| acetic acid don. pyridine accept. ab in   | 1.824     | -46.8     | 227.4 |
| acetic acid don.-methyl ether accept. pix | 1.823     | -38.6     | 283.0 |
| acetic acid don.-methanol accept. pix     | 1.821     | -39.8     | 285.5 |
| acetic acid don.-acetone accept. pix      | 1.901     | -24.7     | 155.2 |
| acetamide don.-pyridine accept. pix       | 2.003     | -32.0     | 220.1 |

|   |       |       |       |
|---|-------|-------|-------|
| acetamide don.-methyl ether accept. pix | 1.952 | -25.7 | 180.9 |
| acetamide don.-methanol accept. pix     | 1.961 | -25.9 | 159.4 |
| acetamide don.-acetone accept. pix      | 2.019 | -17.7 | 134.8 |
| pyrrole don.-pyridine accept. pix       | 2.027 | -32.9 | 197.6 |
| pyrrole don.-methyl ether accept. pix   | 1.969 | -26.4 | 154.0 |
| pyrrole don.-methanol accept. pix       | 1.975 | -26.7 | 152.2 |
| pyrrole don.-acetone accept. pix        | 2.005 | -22.3 | 144.8 |
| acetic acid-acetic acid pix             | 1.78  | -32.  | 250.  |
| acetamide-acetamide pix                 | 1.96  | -28.  | 135.  |
| benzamide-benzamide pix                 | 1.95  | -27.  | 140.  |
| pyrazole-pyrazole pix                   | 1.95  | -39.  | 200.  |
| methanol-methanol pix                   | 1.95  | -22.  | 130.  |
| phenol-phenol pix                       | 1.90  | -25.  | 130.  |

#### halogen bonds

|                                     |       |       |      |
|-------------------------------------|-------|-------|------|
| Bromobenzene pyridine all ab initio | 3.197 | -7.2  | 59.0 |
| p-ohBromobenzene pyridine           | 3.229 | -6.7  | 51.8 |
| no2Brbenzene pyridine ab initio     | 3.140 | -11.4 | 64.9 |
| nh2Brbenzene pyridine all ab initio | 3.252 | -5.4  | 53.2 |
| cnBrbenzene pyridine                | 3.117 | -10.4 | 62.5 |
| coohBrbenzene pyridine              | 3.166 | -8.8  | 62.1 |
| c6h5cl benzoquinone abini           | 3.340 | -1.6  | 16.6 |
| c6f5cl furan ab ini                 | 3.163 | -5.2  | 37.8 |
| c6f5I furan ab ini                  | 3.182 | -11.3 | 66.6 |
| c6f5I benzonitrile                  | 3.175 | -16.7 | 74.5 |
| c6h5i-pyridine                      | 3.192 | -13.4 | 69.1 |
| c6h5br benzonitrile                 | 3.302 | -4.0  | 49.5 |
| c6f5br pyridine                     | 3.010 | -15.7 | 86.0 |
| c6f5br benzoquinone                 | 3.118 | -9.0  | 54.9 |
| c6h5br furan ab ini                 | 3.326 | -4.2  | 28.5 |

#### weak bonds

|                                  |       |      |      |
|----------------------------------|-------|------|------|
| acetone benzene ab initio        | 2.533 | -5.2 | 32.3 |
| formaldehyde ethylene ab initio  | 2.602 | -3.1 | 27.9 |
| urea benzene ab initio           | 2.385 | -7.6 | 54.7 |
| chlorobenzene-chlorobenzene pix  | 3.501 | -2.6 | 33.2 |
| hexaFclbenz-chlorobenzene pix    | 3.451 | -3.3 | 38.0 |
| benzene-chlorobenzene H...Cl pix | 2.535 | -4.9 | 40.6 |
| benzene acetonitrile pix         | 2.548 | -6.6 | 36.5 |

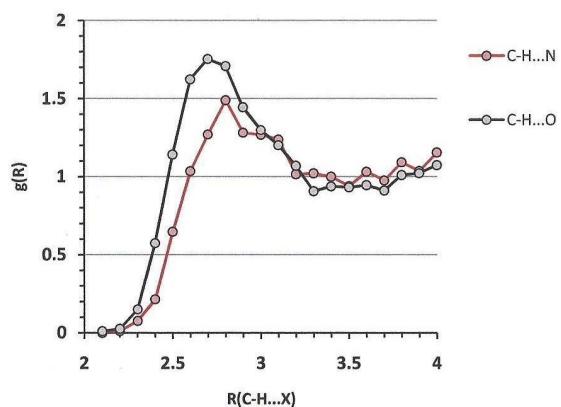


Figure S16. DDF(R) functions for the CH...O and CH...N contacts in crystals

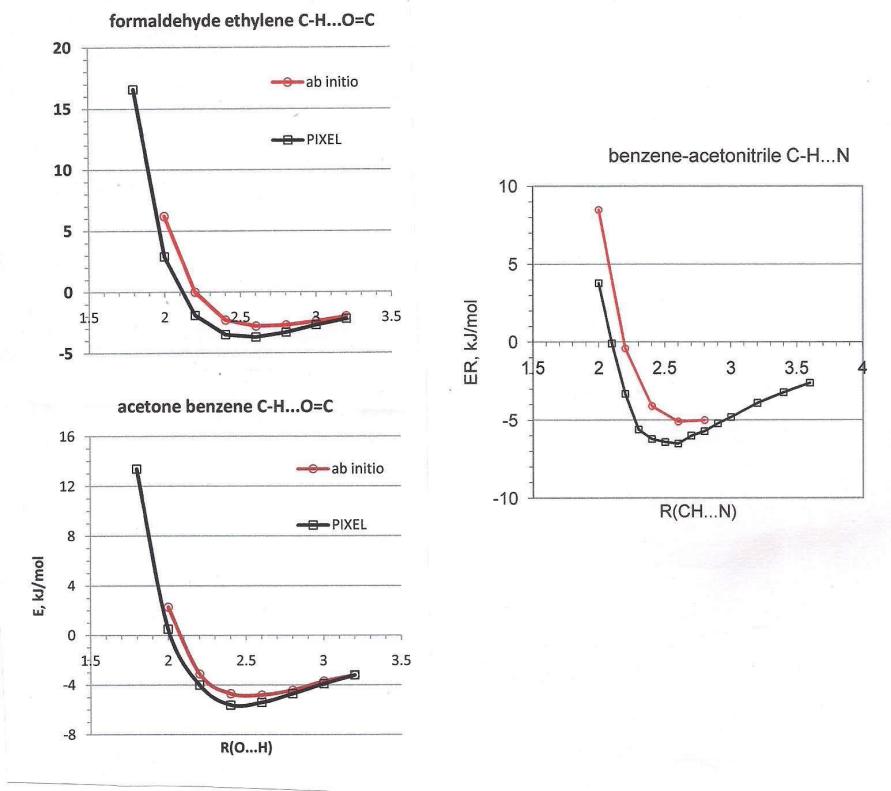


Figure S17. Binding energy profiles for some weak-bond dimers