Supporting Information

for

Molecular Crystals Sublimation: Structural Clusterization of the Database as an Approach for Prediction of the Thermodynamic Functions

German L. Perlovich

G.A. Krestov Institute of Solution Chemistry of the Russian Academy of Sciences, 153045, Ivanovo, Russia. E-mail: glp@isc-ras.ru

*To whom correspondence should be addressed: Telephone: +7-4932-533784; Fax: +7-4932-336237; E-mail glp@isc-ras.ru

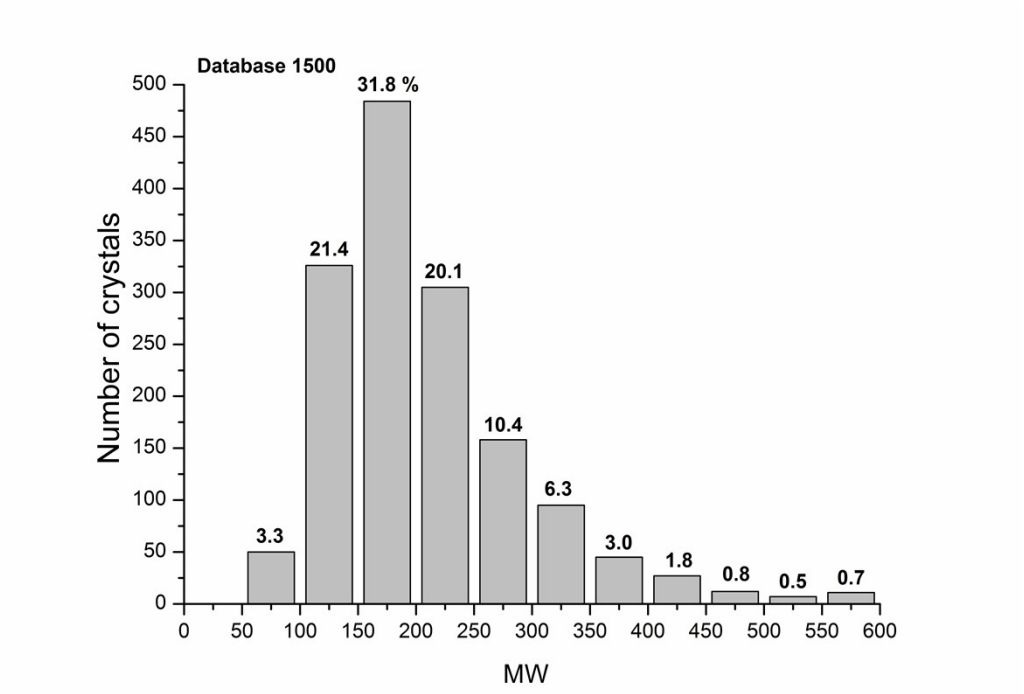

Figure S1. Function distribution crystals versus MW for Database with 1500 occurrences (number corresponds to percentage from the total number of the crystals)

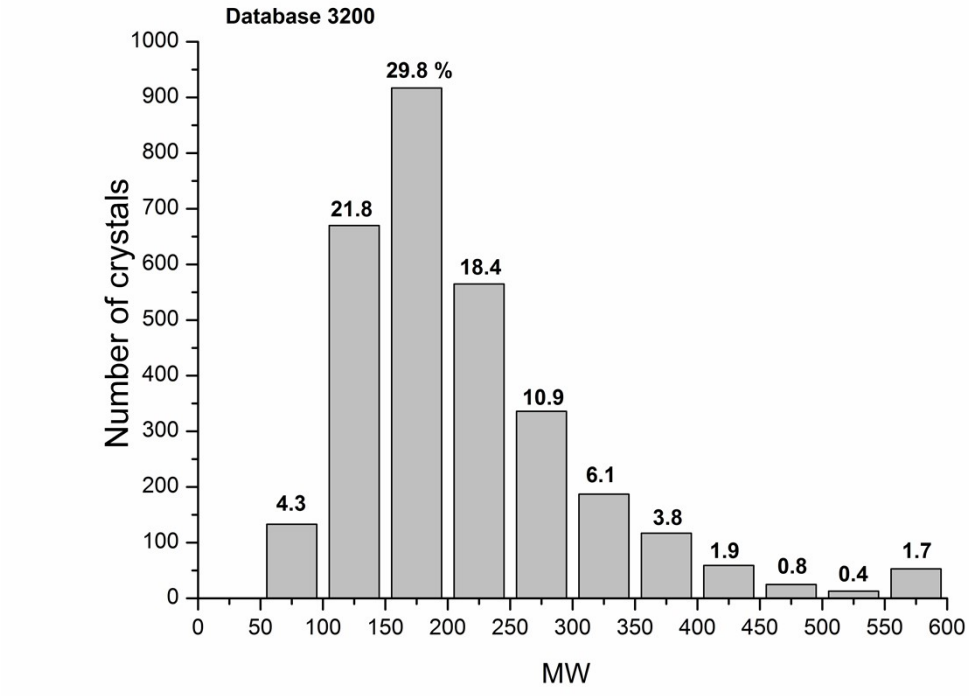

Figure S2. Function distribution crystals versus MW for Database with 3200 occurrences (number corresponds to percentage from the total number of the crystals)

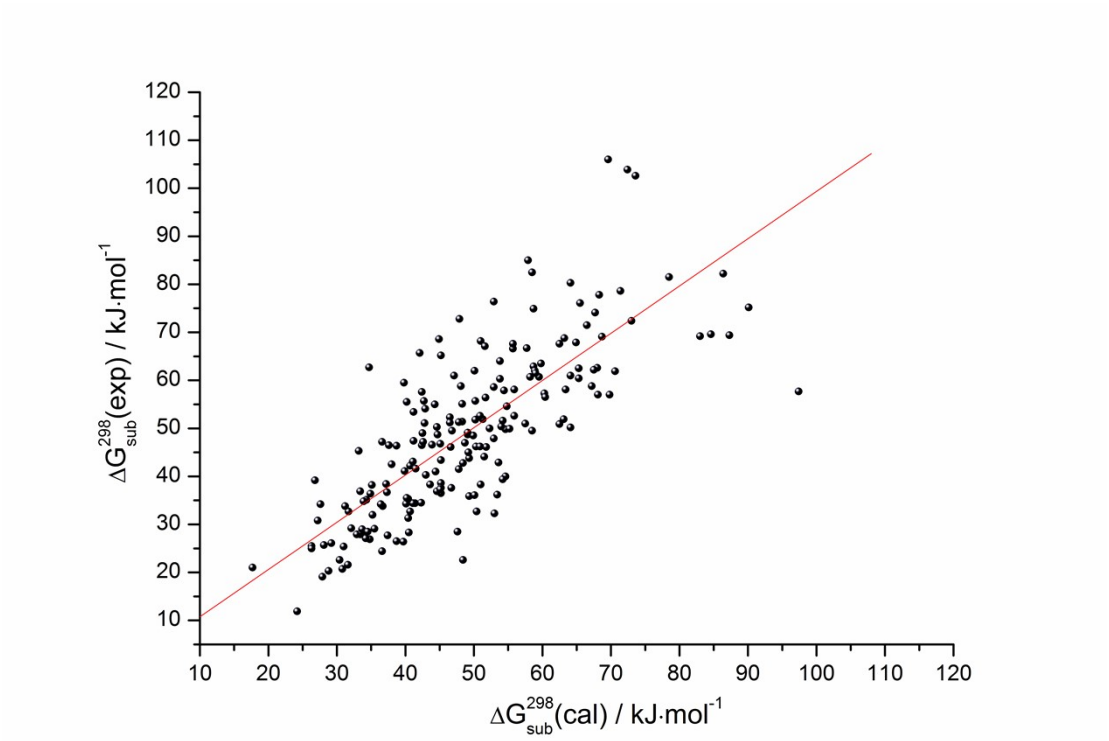

Figure S3. Dependence $\Delta G^{298}_{sub}(\exp)$ versus $\Delta G^{298}_{sub}(cal)$ (calculated by the correlation model using HYBOT descriptors without clusterization)

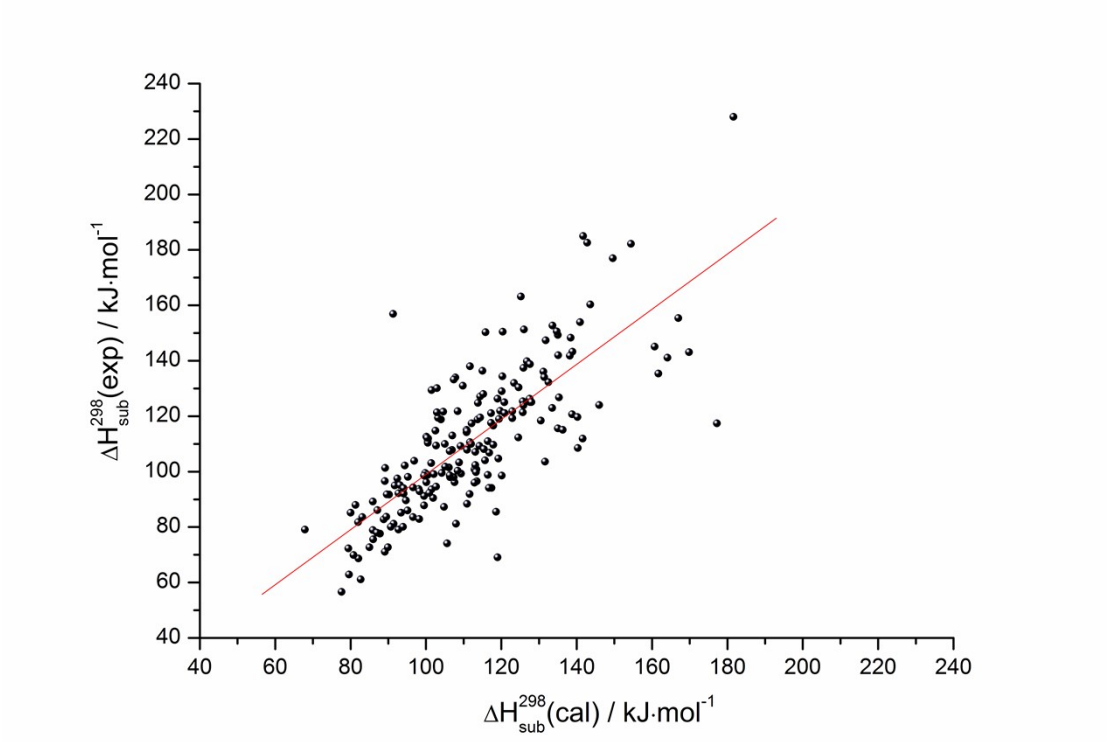

Figure S4. Dependence $\Delta H^{298}_{sub}(\exp)$ versus $\Delta H^{298}_{sub}(cal)$ (calculated by the correlation model using HYBOT descriptors without clusterization)

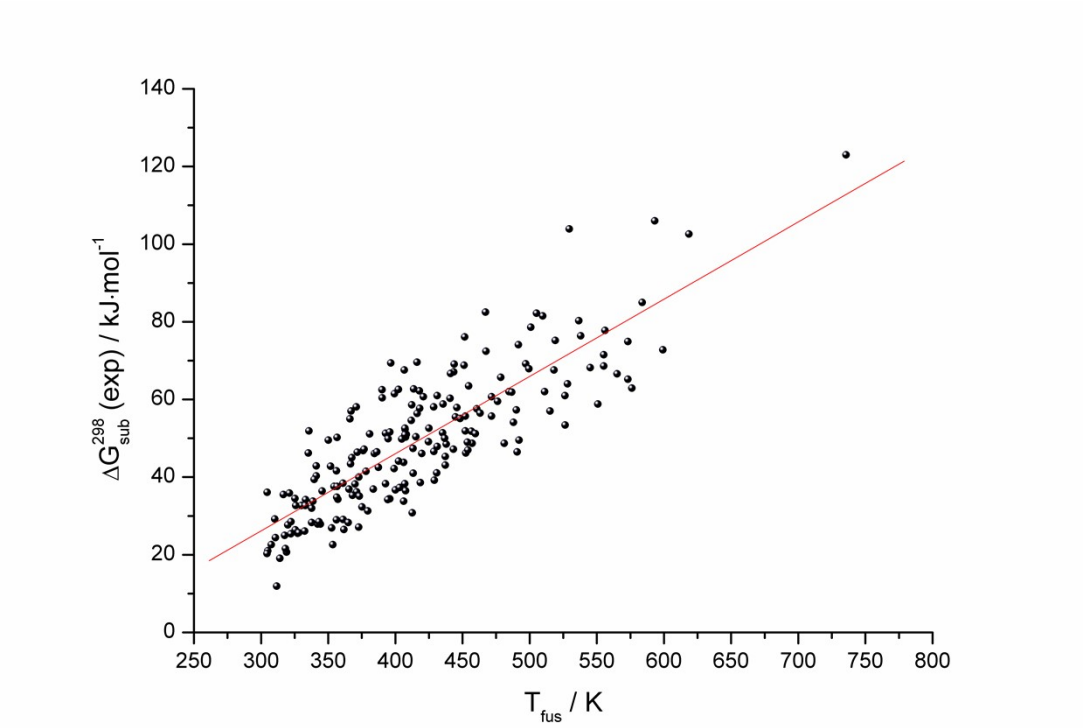

Figure S5. Dependence $\Delta G^{298}_{sub}(\exp)$ versus T_{fus}

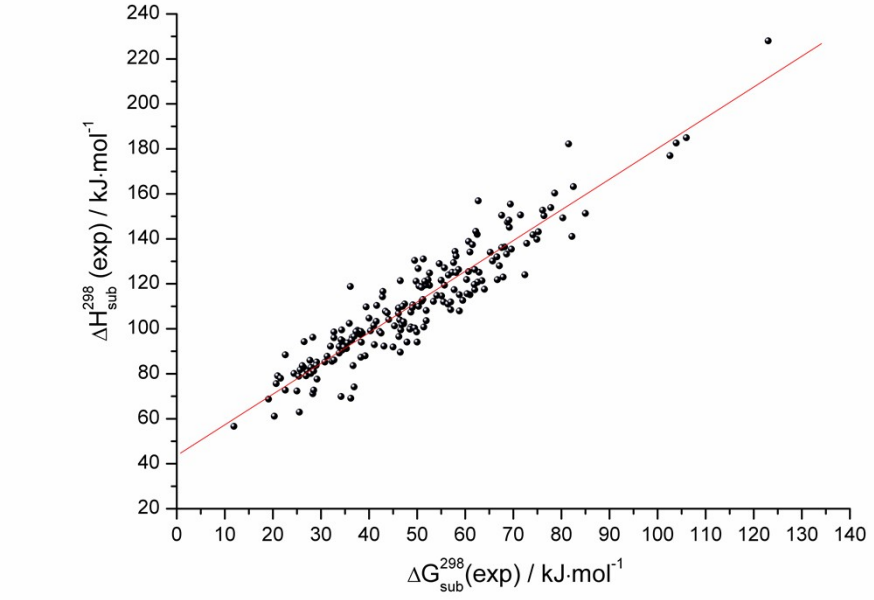

Figure S6. Dependence $\Delta H^{298}_{sub}(\exp)$ versus $\Delta G^{298}_{sub}(\exp)$

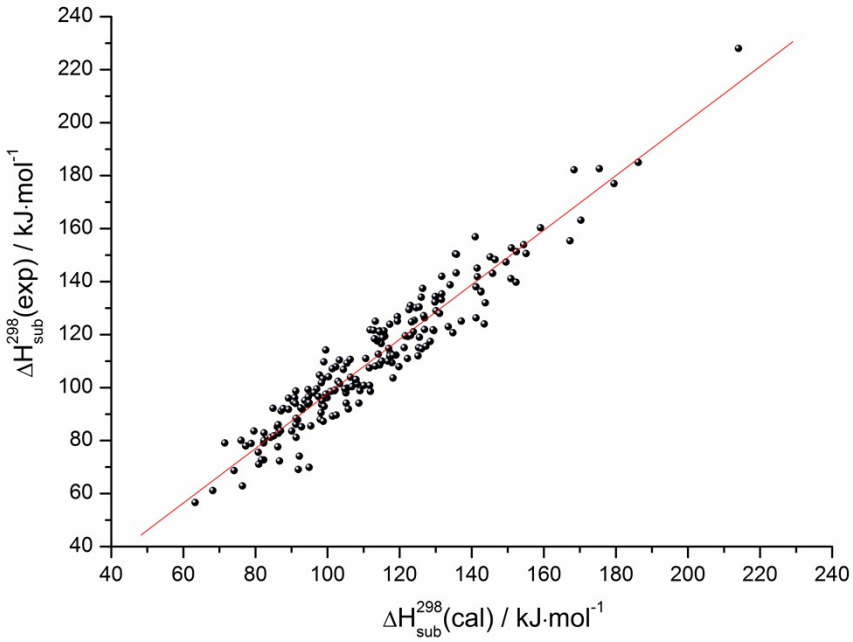

Figure S7. Dependence $\Delta H_{sub}^{298}(\exp)$ versus $\Delta H_{sub}^{298}(cal)$ (calculated by the correlation model $\Delta G_{sub}^{298}(\exp) = A + B \cdot T_{fus}$ with clusterization)

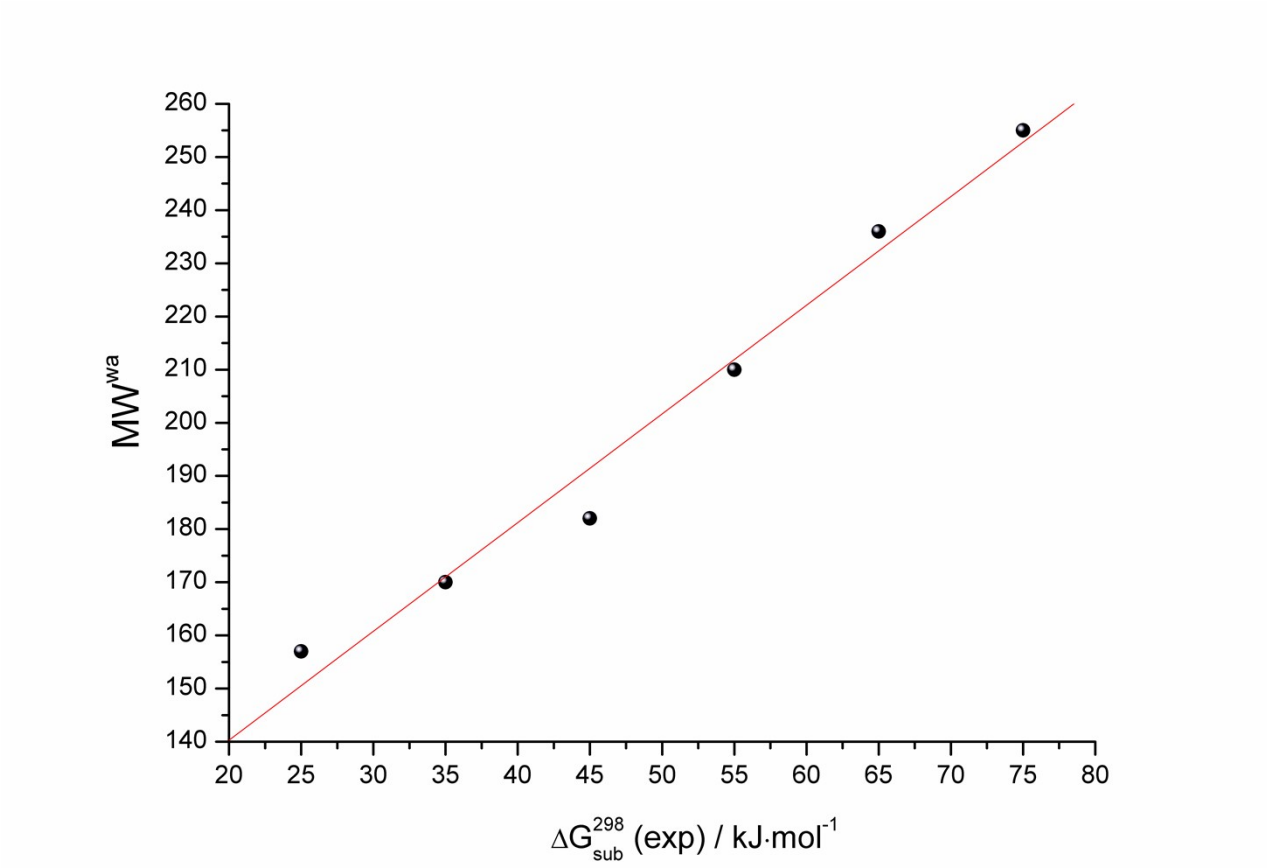

Figure S8. Dependence MW^{wa} versus $\Delta G_{sub}^{298}(\exp)$

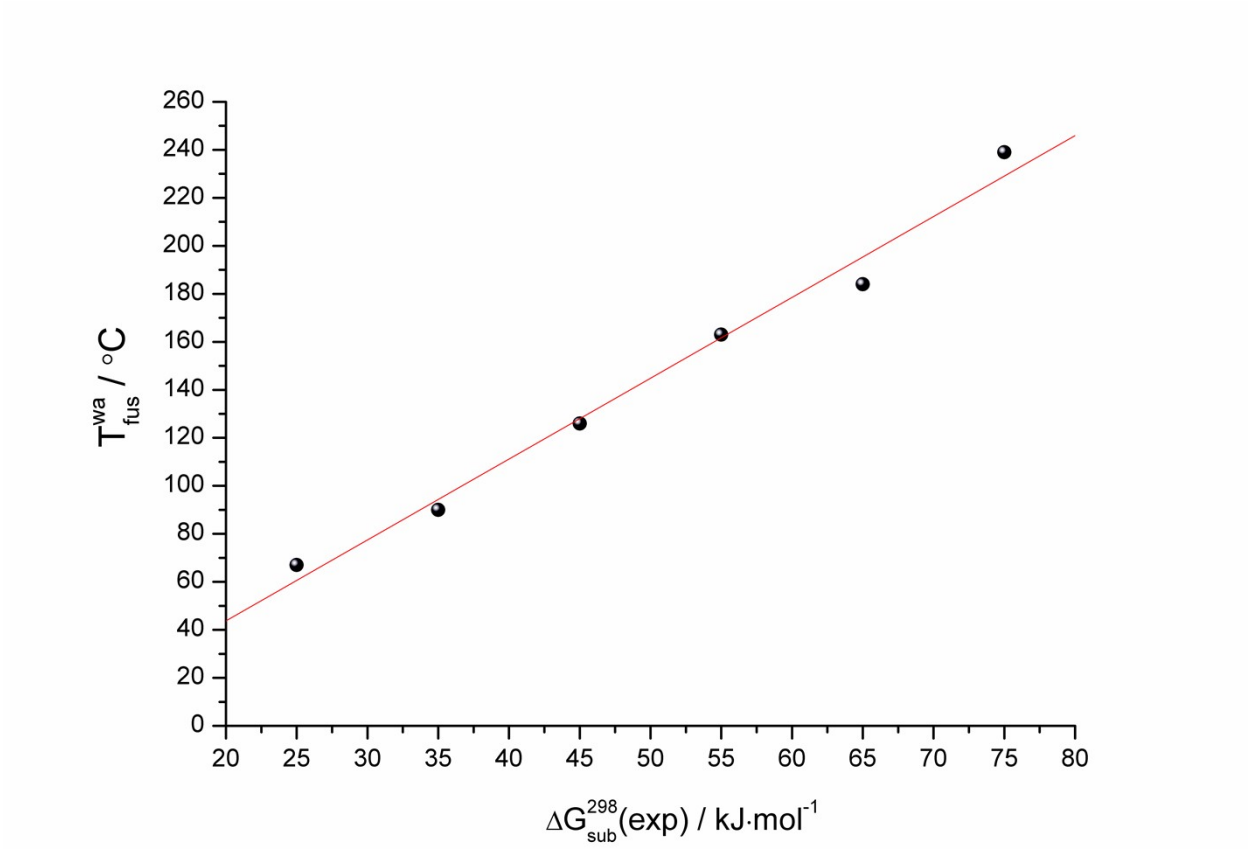

Figure S9. Dependence T_{fus}^{wa} versus $\Delta G_{sub}^{298}(\text{exp})$

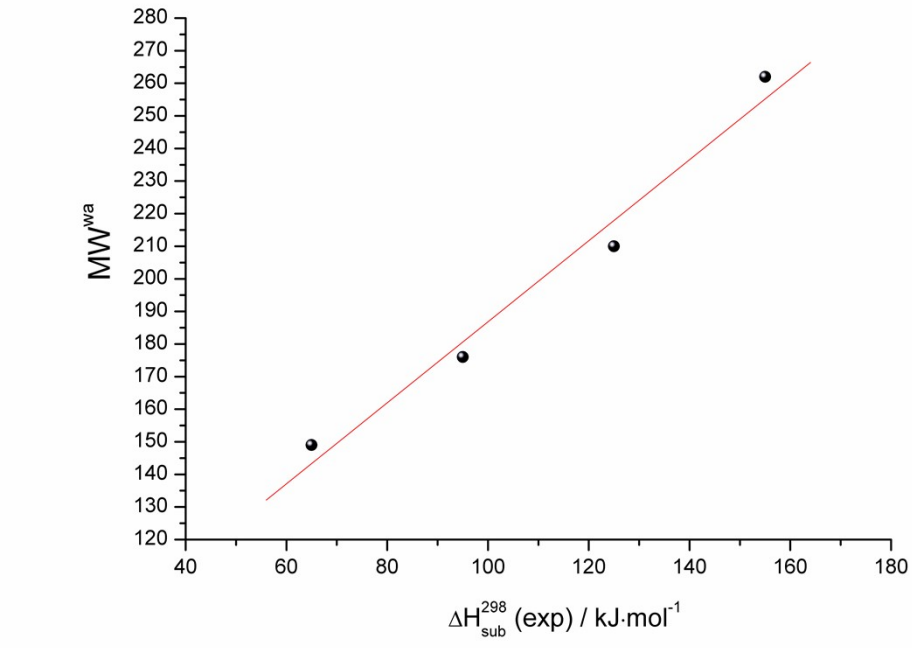

Figure S10. Dependence MW^{wa} versus $\Delta H_{sub}^{298}(\text{exp})$

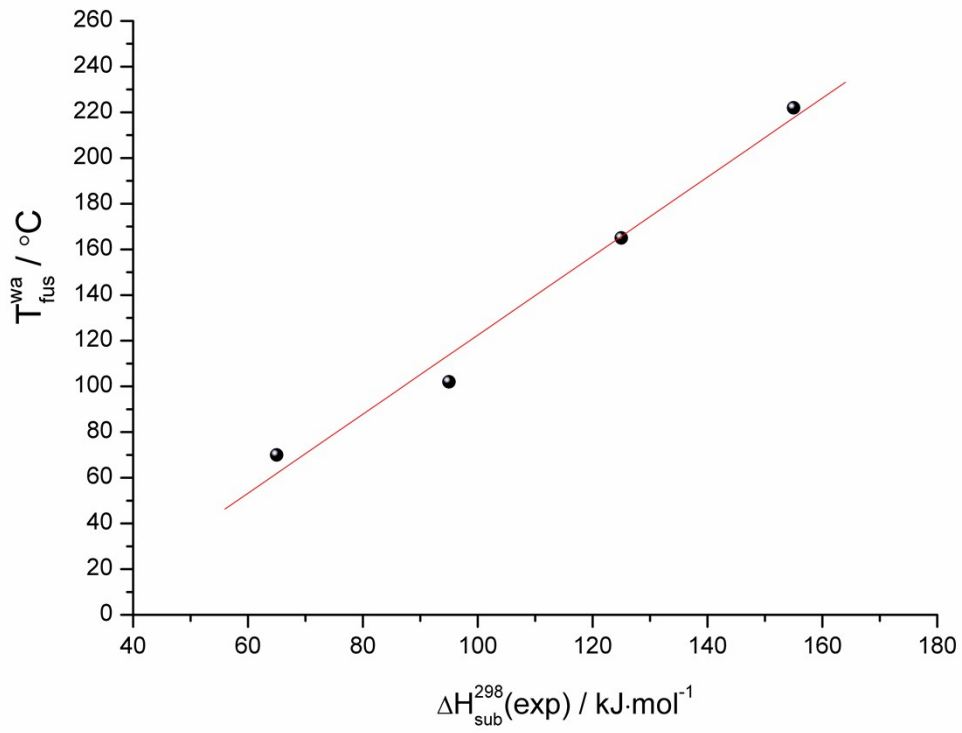

Figure S11. Dependence T_{fus}^{wa} versus ΔH_{sub}^{298}(exp)

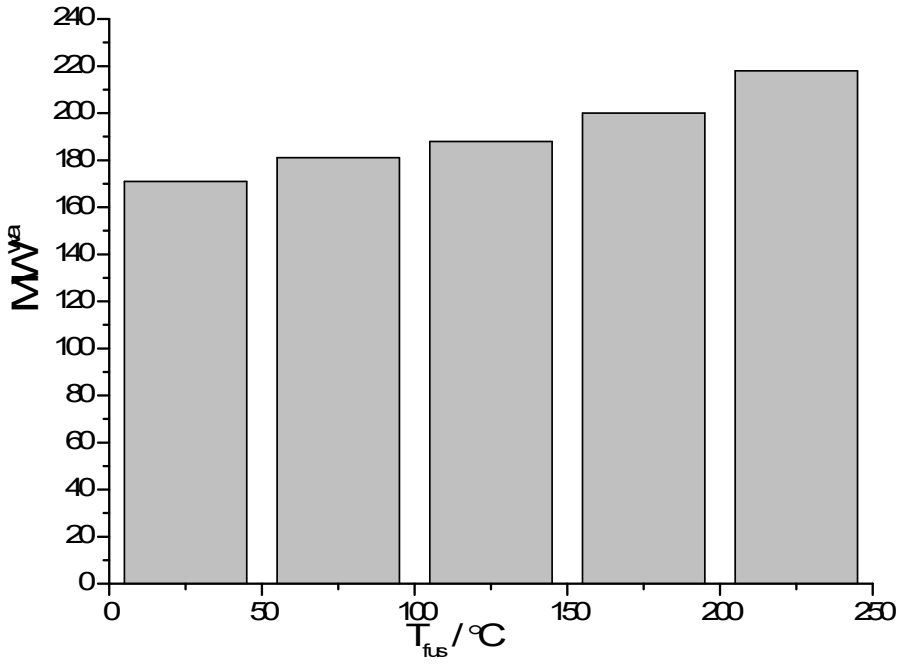

Figure S12. Function distribution MW^{wa} for the groups of molecular crystals with different values of T_{fus}

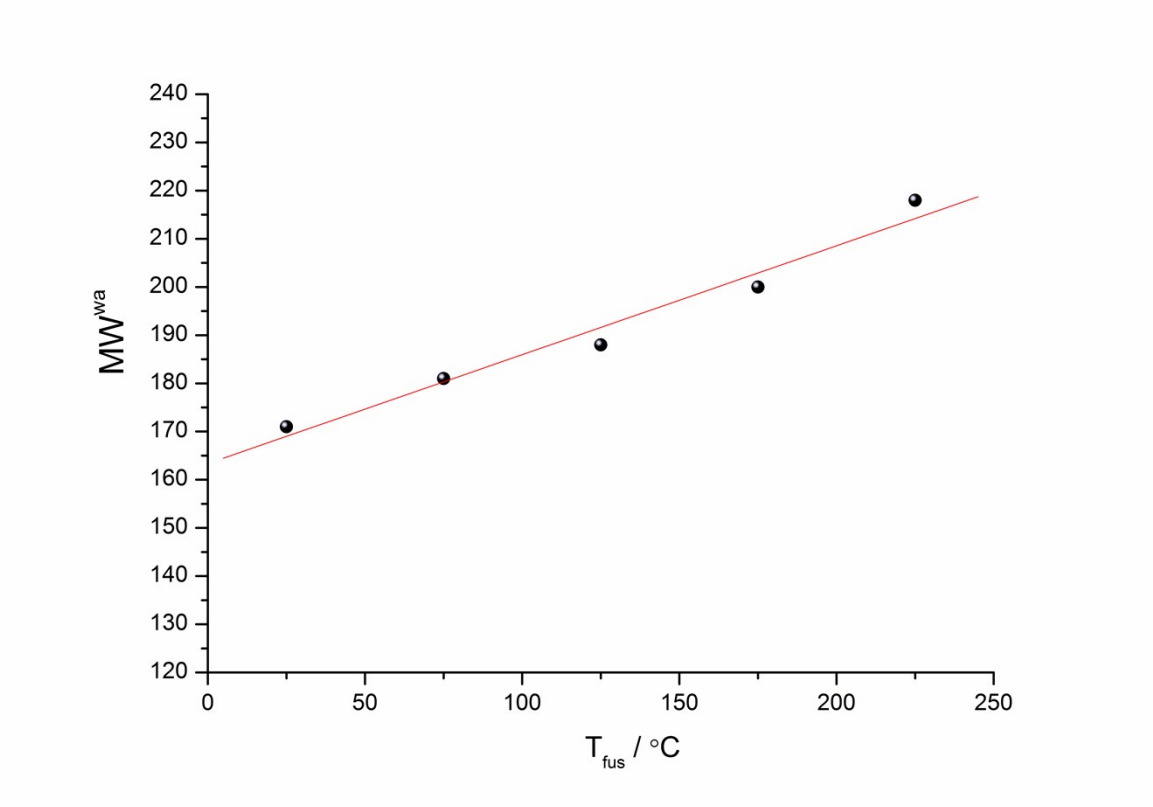

Figure S13. Dependence MW^{wa} versus T_{fus}

Table S1. Characteristics of correlation equation $\Delta G_{sub}^{298}(\text{exp}) = A + B \cdot \Delta H_{sub}^{298}(\text{exp})$ for groups of molecular crystals with different ς_{TS} values

ς_{TS} %	*A*	*B*	R	σ	n
32-33	0.00 ± 0.46	0.517 ± 0.004	0.9975	0.84	110
33-34	-0.49 ± 0.43	0.500 ± 0.003	0.9973	0.85	121
34-35	-0.26 ± 0.40	0.475 ± 0.003	0.9965	0.84	159
35-36	-0.62 ± 0.35	0.454 ± 0.003	0.9959	0.78	194
36-37	-0.82 ± 0.38	0.433 ± 0.003	0.9939	0.78	197
37-38	-0.81 ± 0.34	0.408 ± 0.003	0.9942	0.77	190
38-39	-0.73 ± 0.39	0.382 ± 0.004	0.9928	0.80	149
39-40	-0.74 ± 0.47	0.356 ± 0.005	0.9881	0.74	123

Table S2. Distribution functions of molecular crystals by molecular weight for different groups of compounds selected by the ς_{TS} values

ς_{TS}	32-33	33-34	34-35	35-36	36-37	37-38	38-39	39-40
MW	%	%	%	%	%	%	%	%
50-100	6	4	3	6	2	5	5	6
100-150	14	21	27	27	36	52	43	29
150-200	**22**[b]	19	**40**	**54**	**80**	**75**	**70**	**53**
200-250	14	**34**	37	**54**	48	39	22	21
250-300	21	13	18	32	17	16	6	6
300-350	15	10	23	13	11	1	3	2
350-400	9	10	11	8	3	2	0	6
400-450	5	6	0	0	0	0	0	0
450-500	4	4	0	0	0	0	0	0
MW^{wa} [a]	239	224	213	209	188	175	169	175

[a] weighted average value;
[b] Maximal value of the distribution function (bold font);

Table S3. Distribution functions of molecular crystal by fusion temperatures for different groups of compounds selected according to ς_{TS} values.

ς_{TS}	32-33	33-34	34-35	35-36	36-37	37-38	38-39	39-40
T_{fus} /°C	%	%	%	%	%	%	%	%
0-50	0	0	1	0	0	0	10	17
50-100	9	10	12	19	32	68	**80**	**84**
100-150	16	21	33	59	**108**	**93**	53	19
150-200	25	**34**	**58**	**87**	50	25	6	3
200-250	25	32	43	27	7	4	0	0
250-300	**26**[b]	21	12	1	0	0	0	0
300-350	9	3	0	1	0	0	0	0
T_{fus}^{wa} [a]	210	193	179	161	131	115	90	76

[a] weighted average value;
[b] Maximal value of the distribution function (bold font);

Table S4. Distribution functions of molecular crystals by molecular weight for different groups of compounds selected for $\Delta G_{sub}^{298}(\text{exp})$ (in kJ·mol^{-1}).

$\Delta G_{sub}^{298}(\text{exp})$ MW	20-30	30-40	40-50	50-60	60-70	70-80
50-100	20	16	5	6	2	2
100-150	62	87	75	36	26	17
150-200	**73**[b]	**120**	**151**	**89**	26	15
200-250	23	50	77	79	**52**	11
250-300	3	17	33	40	33	**19**
300-350	3	3	7	27	25	16
350-400	0	2	5	9	13	7
400-450	0	0	2	3	5	4
450-500	0	7	1	5	0	3
MW[wa a]	157	170	182	210	236	255

[a] weighted average value;
[b] Maximal value of the distribution function (bold font);

Table S5. Distribution functions of molecular crystals by fusion temperatures for different groups of compounds selected for $\Delta G_{sub}^{298}(\text{exp})$ (in kJ·mol^{-1}).

$\Delta G_{sub}^{298}(\text{exp})$ T_{fus} /°C	20-30	30-40	40-50	50-60	60-70	70-80
0-50	58	26	3	0	0	0
50-100	**101**[b]	**157**	97	19	3	0
100-150	23	95	**148**	**102**	47	4
150-200	2	23	90	**102**	**60**	25
200-250	0	1	17	57	46	23
250-300	0	0	1	14	20	**31**
300-350	0	0	0	0	6	11
T_{fus}^{wa} [a]	67	90	126	163	184	239

[a] weighted average value;
[b] Maximal value of the distribution function (bold font);

Table S6. Distribution functions of molecular crystals by molecular weight for different groups of compounds selected for ΔH_{sub}^{298}(exp) (in kJ·mol^{-1}).

ΔH_{sub}^{298}(exp) MW	50-80	80-110	110-140	140-170
50-100	16	25	8	0
100-150	**46**[b]	179	76	22
150-200	42	**266**	**152**	20
200-250	11	126	134	30
250-300	3	48	74	30
300-350	3	9	49	**32**
350-400	1	6	21	14
400-450	0	8	5	6
450-500	0	14	5	4
MW[wa a]	149	176	210	262

[a] weighted average value;
[b] Maximal value of the distribution function (bold font);

Table S7. Distribution functions of molecular crystals by fusion temperatures for different groups of compounds selected for ΔH_{sub}^{298}(exp) (in kJ·mol^{-1}).

ΔH_{sub}^{298}(exp) T_{fus} /°C	50-80	80-110	110-140	140-170
0-50	38	57	4	0
50-100	**57**[b]	**273**	55	2
100-150	22	220	156	23
150-200	5	101	**161**	41
200-250	0	27	97	30
250-300	0	3	41	**47**
300-350	0	0	10	13
350-400	0	0	0	2
T_{fus}^{wa} [a]	70	102	165	222

[a] weighted average value;
[b] Maximal value of the distribution function (bold font);

Table S8. Distribution functions of molecular crystals by MW for different groups of compounds selected for T_{fus} (in °C)

T_{fus} / MW	0-50	50-100	100-150	150-200	200-250
50-100	9	12	15	7	4
100-150	23	90	99	51	22
150-200	**46**[b]	**148**	**127**	**96**	**42**
200-250	18	83	78	68	38
250-300	1	31	53	38	19
300-350	1	13	24	29	21
350-400	0	1	16	11	11
400-450	2	2	3	6	5
450-500	2	10	6	3	2
T_{fus}^{wa} [a]	171	181	188	200	218

[a] weighted average value;
[b] Maximal value of the distribution function (bold font);

Each group of crystals with $\Delta G_{sub}^{298}(\text{exp})$ values within 10 kJ·mol^{-1} corresponds to the specific value of T_{fus}^{wa} and MW^{wa}. At that, the weighted average values also increase with $\Delta G_{sub}^{298}(\text{exp})$ growth and can be described by the following correlation equations:

$$MW^{wa} = (99.4 \pm 7.8) + (2.05 \pm 0.15) \cdot \Delta G_{sub}^{298} / kJ \cdot mol^{-1} \quad \text{(S1)}$$

R = 0.9897; σ = 6.2; n = 6

$$T_{fus}^{wa} / °C = -(23.6 \pm 10.8) + (3.37 \pm 0.20) \cdot \Delta G_{sub}^{298} / kJ \cdot mol^{-1} \quad \text{(S2)}$$

R = 0.9927; σ = 8.6; n = 6

These equations give an indication of how sensitive the values of MW^{wa} and T_{fus}^{wa} to the changes in the values of $\Delta G_{sub}^{298}(\text{exp})$ are. A similar analysis can be done by rewriting these equations in the form:

$$\Delta G_{sub}^{298} / kJ \cdot mol^{-1} = -(46.6 \pm 7.1) + (0.48 \pm 0.03) \cdot MW^{wa} \quad \text{(S3)}$$

R = 0.9897; σ = 3.0; n = 6

$$\Delta G_{sub}^{298} / kJ \cdot mol^{-1} = (7.6 \pm 2.8) + (0.29 \pm 0.02) \cdot T_{fus}^{wa} / °C \quad \text{(S4)}$$

R = 0.9927; σ = 2.5; n = 6

The value $\Delta G_{sub}^{298}(\text{exp})$ is more sensitive to the changes in the molecular weight of a compound as compared to the variations in the melting point of a crystal.

Each group of crystals with $\Delta H_{sub}^{298}(\text{exp})$ values within 30 kJ·mol^{-1} corresponds to the specific value of T_{fus}^{wa} and MW^{wa}. At that, the weighted average values also increase with $\Delta H_{sub}^{298}(\text{exp})$ growth and can be described by the following correlation equations:

$$MW^{wa} = (62.5 \pm 15.4) + (1.24 \pm 0.13) \cdot \Delta H_{sub}^{298} / kJ \cdot mol^{-1} \quad \text{(S5)}$$

R = 0.9885; σ = 9.0; n = 4

$$T_{fus}^{wa} / °C = -(50.6 \pm 18.2) + (1.73 \pm 0.16) \cdot \Delta H_{sub}^{298} / kJ \cdot mol^{-1} \quad \text{(S6)}$$

R = 0.9918; σ = 10.6; n = 4

These equations give an indication of how sensitive the values of MW^{wa} and T_{fus}^{wa} to the changes in the values of $\Delta H_{sub}^{298}(\text{exp})$ are. A similar analysis can be done by rewriting these equations in the form:

$$\Delta H_{sub}^{298} / kJ \cdot mol^{-1} = -(46.6 \pm 17.3) + (0.79 \pm 0.08) \cdot MW^{wa} \quad \text{(S7)}$$

R = 0.9885; σ = 7.2; n = 4

$$\Delta H_{sub}^{298} / kJ \cdot mol^{-1} = (30.5 \pm 7.9) + (0.57 \pm 0.05) \cdot T_{fus}^{wa} / °C \quad \text{(S8)}$$

R = 0.9918; σ = 6.1; n = 4

Since for each interval of T_{fus} (each group) function variation there is a specific distribution function by the MW values, we analyzed this function similarly to the previous cases (Tables S8). The dependences of the weighted average value of molecular weight (MW^{wa}) on fusion temperature (T_{fus}^{wa}) (for each group) are shown in Figures S10, Figures S11. Thus, a specific value of MW^{wa} corresponds to each group of crystals with T_{fus}^{wa} values within 50 °C. At that, the weighted average values also increase with T_{fus}^{wa} growth and can be described by the following correlation equations:

$$MW^{wa} = (163.4 \pm 3.3) + (0.23 \pm 0.02) \cdot T_{fus}^{wa} / °C \quad \text{(S9)}$$

R = 0.9846; σ = 3.7; n = 5

$$T_{fus}^{wa} / °C = -(696.8 \pm 84.6) + (4.29 \pm 0.43) \cdot MW^{wa} \quad \text{(S10)}$$

R = 0.9846; σ = 16.0; n = 5