

## Design of Diastereomeric Salt Resolution via Multicomponent System Characterization: A Case Study with Hydrate Formation

Miklós H. Bosits,<sup>a,b</sup> Laura Bereczki,<sup>c</sup> Petra Bombicz,<sup>c</sup> Zsófia Szalay,<sup>b</sup> Hajnalka Pataki\*<sup>a</sup> and Ádám Demeter<sup>b</sup>

### Supporting Information

#### S1. X-ray powder diffractograms and dynamic vapor sorption measurements of the relevant solid phases

Solid phases of each experiment were analyzed by X-ray powder diffraction (XRPD). Crystal structures were identified qualitatively based on their XRPD pattern (Figure S1). The measured diffractograms of (*R*)-pregabalin L-tartrate monohydrate and (*S*)-pregabalin D-tartrate monohydrate are the same as expected. The diffractograms of racemic pregabalin monohydrate, (*S*)-pregabalin L-tartrate monohydrate, and (*S*)-pregabalin D-tartrate monohydrate measured by variable-temperature X-ray powder diffraction (VT-XRPD) are presented on Figure S2-S4, respectively. The hydrate sorption and desorption isotherms of racemic pregabalin, (*S*)-pregabalin L-tartrate, and (*S*)-pregabalin D-tartrate measured by dynamic vapor sorption (DVS) are presented on Figure S5-S7, respectively.

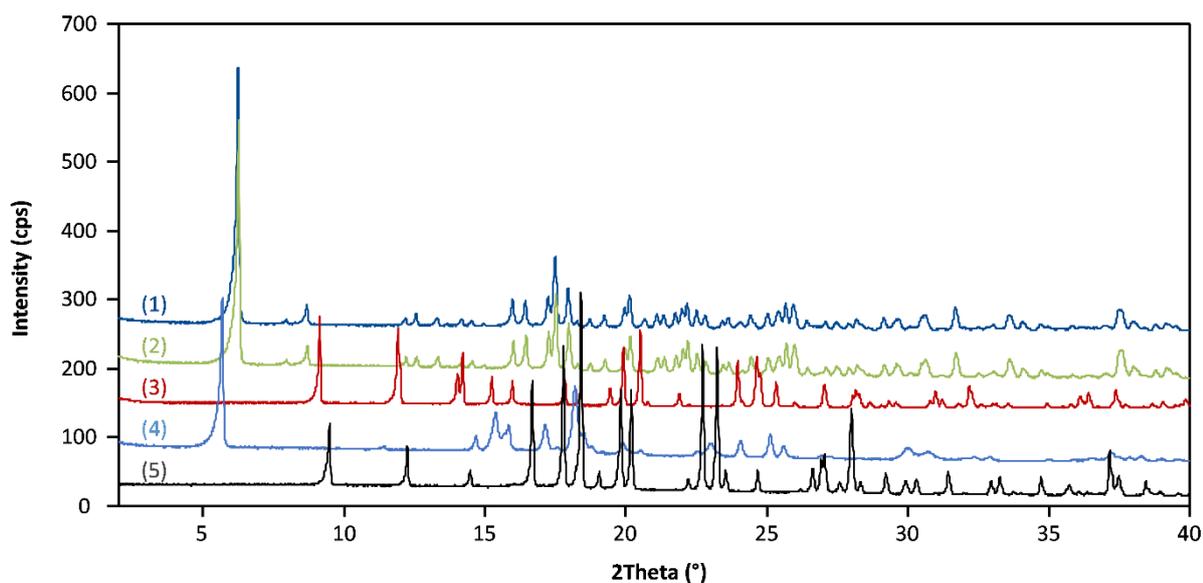


Figure S1. X-ray powder diffractograms of (1) (*R*)-pregabalin L-tartrate monohydrate, (2) (*S*)-pregabalin D-tartrate monohydrate, (3) (*S*)-pregabalin L-tartrate monohydrate, (4) racemic pregabalin monohydrate, and (5) (*S*)-pregabalin

<sup>a</sup> Department of Organic Chemistry and Technology, Faculty of Chemical Technology and Biotechnology, Budapest University of Technology and Economics, Műgyetem rkp. 3., H-1111 Budapest, Hungary.

<sup>b</sup> Gedeon Richter Plc., P.O. Box 27, H-1475 Budapest, Hungary.

<sup>c</sup> Centre for Structural Sciences, Research Centre for Natural Sciences, Magyar tudósok körútja 2., H-1117 Budapest, Hungary.

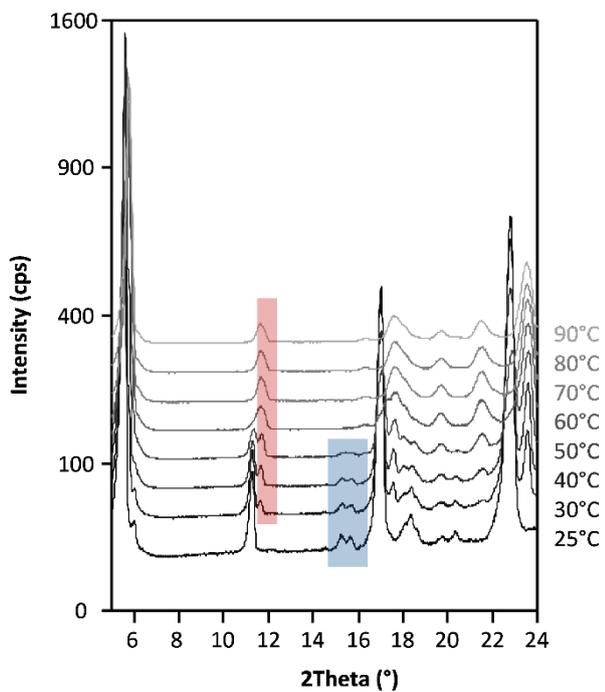


Figure S2. Diffractograms of racemic pregabalin at different temperatures. Blue and red colours correspond to the characteristic peaks of monohydrate and dehydrated forms, respectively.

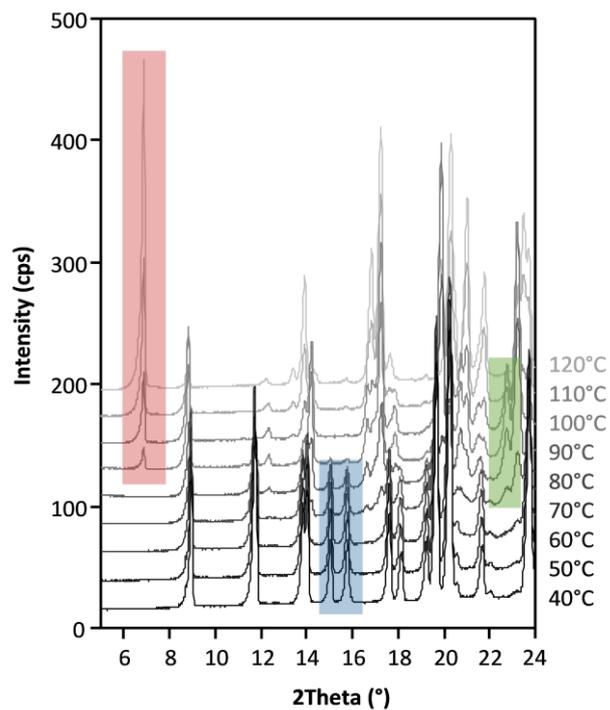


Figure S4. Diffractograms of (*S*)-pregabalin L-tartrate at different temperatures. Blue, green, and red colours correspond to the characteristic peaks of monohydrate, dehydrated, and anhydrate forms, respectively.

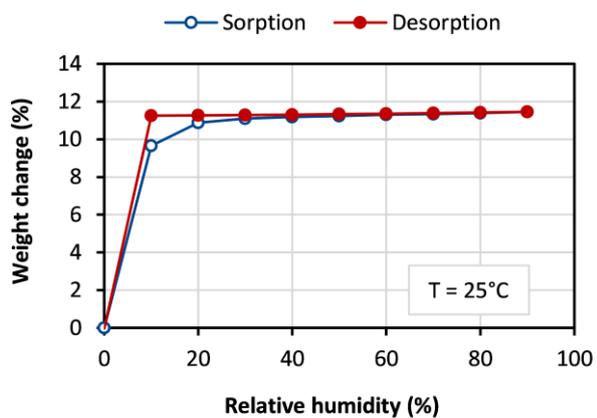


Figure S3. Sorption and desorption isotherms of racemic pregabalin measured by DVS.

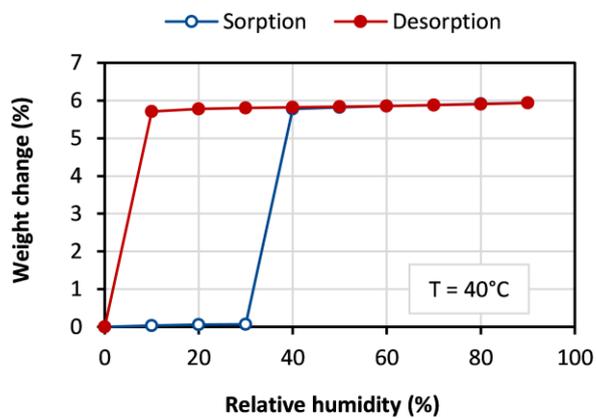


Figure S5. Sorption and desorption isotherms of (*S*)-pregabalin L-tartrate measured by DVS.

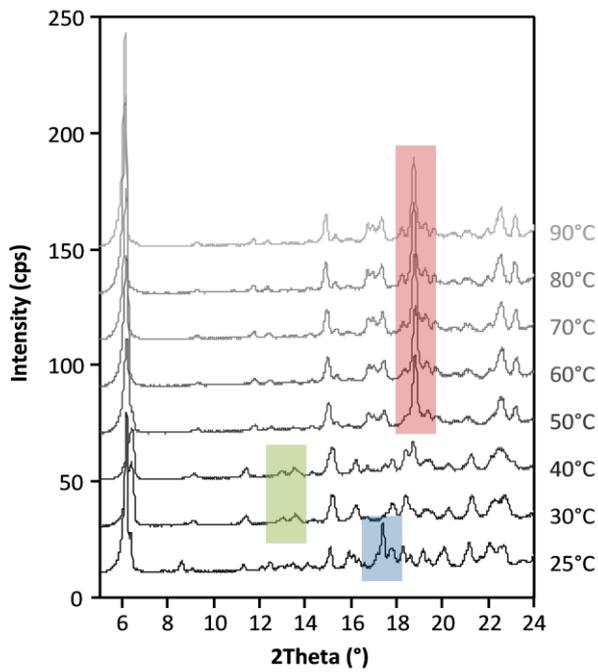


Figure S6. Diffractograms of (*S*)-pregabalin D-tartrate at different temperatures. Blue, green, and red colours correspond to the characteristic peaks of monohydrate, hemihydrate, and dehydrated forms, respectively.

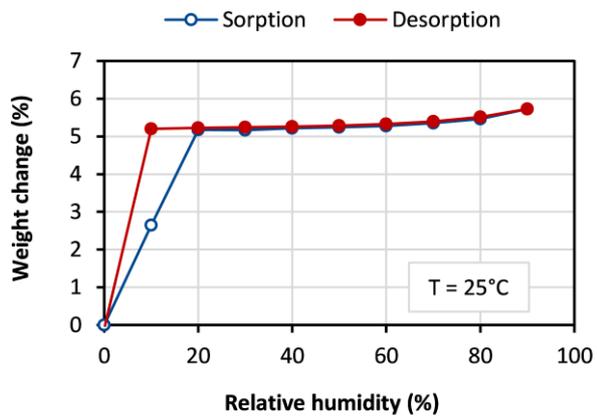


Figure S7. Sorption and desorption isotherms of (*S*)-pregabalin D-tartrate measured by DVS.

## S2. Crystal structure of racemic pregabalin monohydrate

The colourless, platelet crystals of racemic pregabalin hydrate are very soft because the layers easily slip. The structure explains this macroscopical property. The sliding layers can be recognised in the polarized light under the microscope (Figure S8a). After testing several crystals the highlighted section (Figure S8b) separated from the rest of the agglomerate was the subject of single crystal data collection. The diffraction pattern verified the poor crystal quality having prolonged spots.

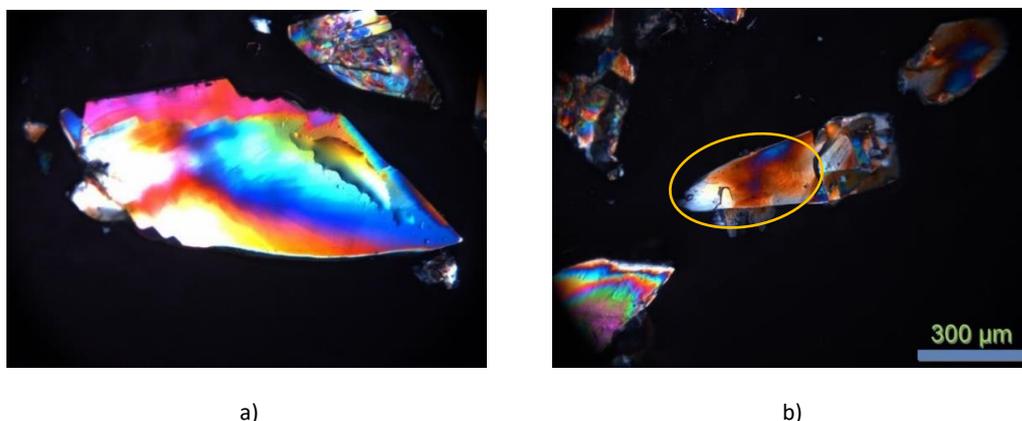


Figure S8. The platelet-type racemic crystals of pregabalin under polarized light. a) The soft, sliding layers diffract the polarized light differently. b) The highlighted single crystal was used in the SXR D experiment.

Crystal data and details of the structure determination and refinement are listed in Table S1. Racemic pregabalin hydrate crystallizes in the triclinic crystal system, in the centrosymmetric space group  $P-1$  (#2). There is one molecule in the asymmetric unit (Figure S9).

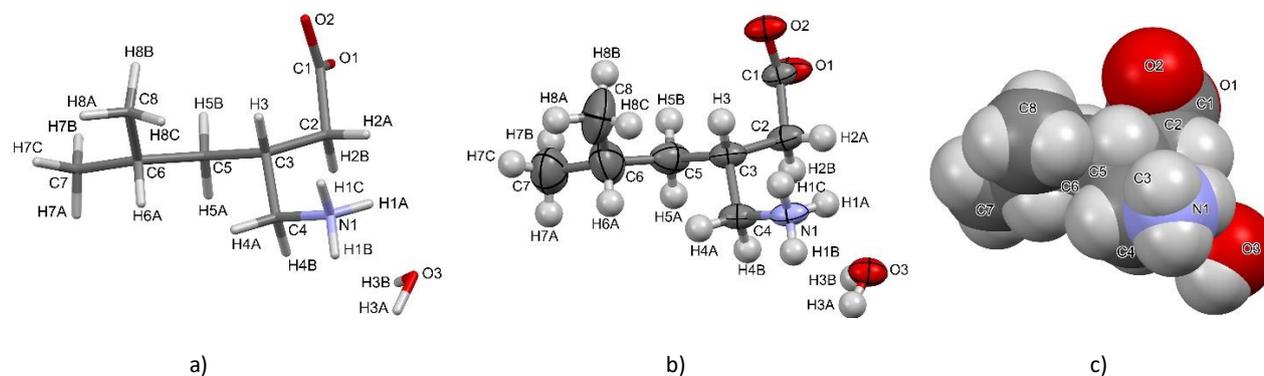


Figure S9. a) Caped stick presentation of (*RS*)-PG with atom labelling. The chiral centre in the molecule is C3. b) Molecular structure ORTEP presentation showing the main occupancies of the disordered positions of C7 and C8. The displacement ellipsoids are drawn at the 50% probability level. c) Space-filling model of (*RS*)-PG.

The C7 and C8 atoms of the branching alkyl termini of the molecule are disordered having available space in the absence of significant intermolecular interactions in the crystal lattice (Figure S10). The two disordered positions occur with almost equal opportunity, the refined value is 53% for C7 and C8, while 47% for C7' and C8'. Two hydrogen positions could be identified and geometrically constrained on the connecting C6. There is no residual solvent accessible void in the crystal lattice. The Kitaigorodskii packing coefficient<sup>1</sup> can not be calculated due to the disordered C7 and C8 atoms. The calculated crystal density is 1.054 g/cm<sup>3</sup>.

<sup>1</sup>A.L. Spek, *J. Appl. Cryst.*, 2003, **36**, 7-13.

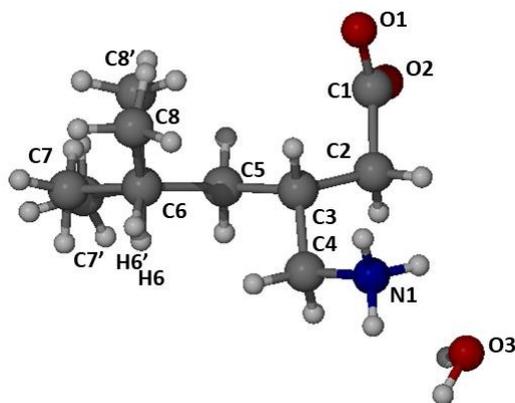


Figure S10. The terminal methyl C7 and C8 atoms are disordered in the crystal lattice along with the connected hydrogen atoms.

Two pregabalin molecules in the unit cell are connected by strong N-H...O type hydrogen bonds forming a dimer via two water molecules (Figure S11) arranged by a symmetry centre, the graph set descriptor<sup>2</sup> is  $R^2_4(6)$ .

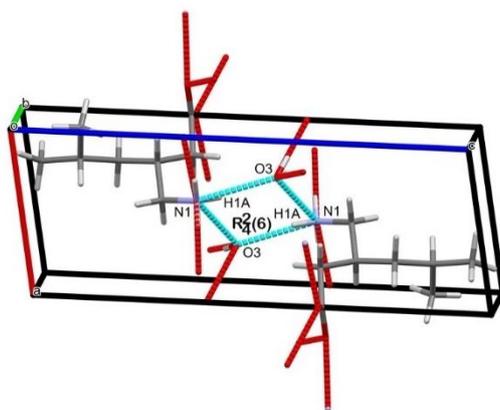


Figure S11. The unit cell ( $Z=2$ ,  $Z'=1$ ) of the racemic crystal of pregabalin showing two hydrate water molecules connected to two pregabalin molecules forming a ring of intermolecular interactions around a symmetry centre described by the graph set descriptor  $R^2_4(6)$ .

The zwitterionic region of the racemic pregabalin together with the water molecule are able to form several strong O-H...O and N-H...O type intermolecular interactions (Figure S12a, Table S2) creating a 2D network which is perpendicular to the  $c$  crystal axis (Figure S12b). The opposite end of the molecule is fully apolar, there are very weak van der Waals interactions among the terminal iso-butyl groups only, they can not participate even in weak C-H type hydrogen bonds as no acceptor atoms remain. The anisotropic distribution of intermolecular interactions, i.e. the strong contrast between the ionic and the apolar regions of the crystal lattice, results in a layered structure, which explains well the soft, sliding macroscopic property of the crystal. The packing arrangement is presented in Figure S13 enhancing the molecular layer structure. The calculated BFDH morphology shows why the parallel layers in the platelet-like crystal are able to split (Figure S14) making the crystals so fragile. The lack of strong intermolecular interactions explains the platelet crystal habit and the low crystal growth rate in the (001) direction.

<sup>2</sup>J. Grell, J. Bernstein and G. Tinhofer, *Acta Crystallogr.*, 1999, **B55**, 1030-1043.

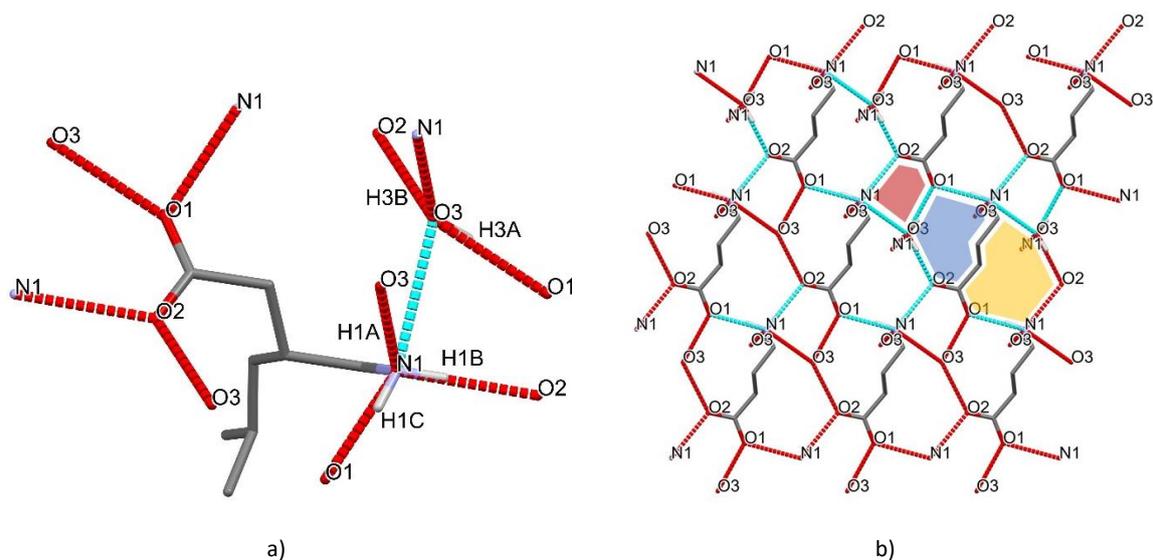


Figure S12. a) The strong O-H...O and N-H...O type intermolecular interactions that the pregabalin molecule is able to form in its racemic crystal. b) The 2D network created by the strong supramolecular interactions in the ab crystallographic sheet. The C5, C6, C7 and C8, as well as the carbon hydrogens are omitted for clarity. Highlighted rings formed by intermolecular interactions: red is  $R^3_3(8)$ , blue is  $R^2_3(11)$ , yellow is  $R^3_4(13)$ .

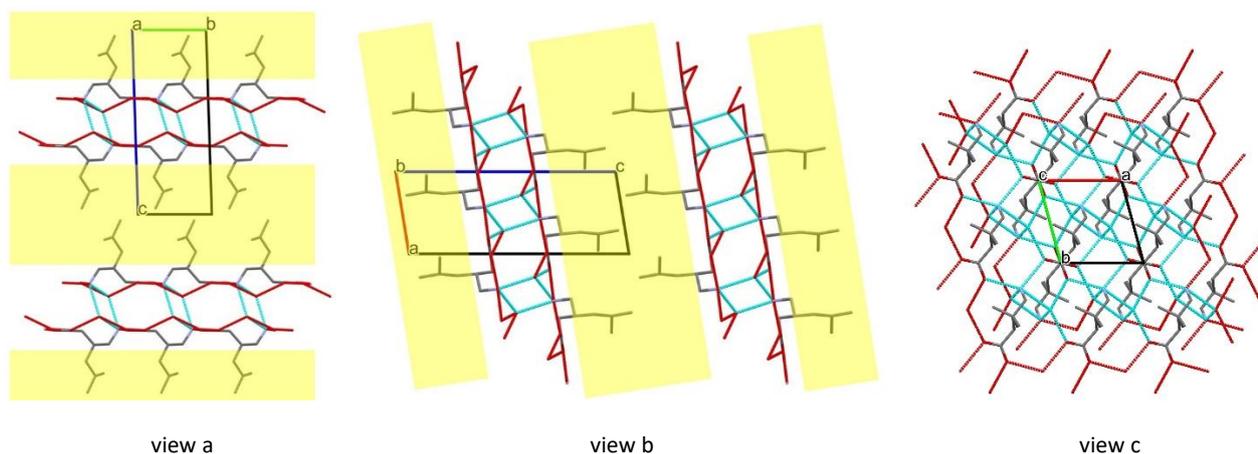


Figure S13. Crystal packing arrangement viewing from the a, b and c crystallographic directions, respectively. Intermolecular interactions are shown by cyan and red dotted lines. Weakly connected apolar layers are highlighted by yellow.

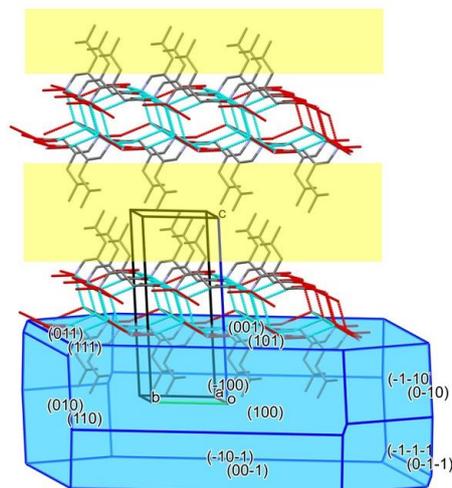


Figure S14. The calculated BFDH morphology with the assumption that the growing rate of the different crystal faces are equal. It can be well seen that the sliding sheets are parallel to the main plate (001) of the platelet-type crystal.

Table S1. Crystal data and details of the structure determination of racemic pregabalin hydrate.

Crystal data			Data collection	
Formula	$C_8H_{17}NO_2 \cdot H_2O$		Temperature (K)	143
Formula weight (g/mol)	177.24		Radiation ( $\text{\AA}$ )	0.71073
Crystal system	Triclinic		Theta min., max. ( $^\circ$ )	3.4, 25.3
Space group	$P-1$	(No. 2)	Dataset	-7 : 7; -7 : 7; -18 : 18
$a, b, c$ ( $\text{\AA}$ )	6.100(2)	6.259(2) 15.420(2)	Tot., Uniq. Data, $R(\text{int})$	10890, 2043, 0.105
$\alpha, \beta, \gamma$ ( $^\circ$ )	85.660(2)	80.031(2) 74.466(2)	Observed data, $I > 2\sigma(I)$	1395
$V$ ( $\text{\AA}^3$ )	558.4(3)		Refinement	
$Z, Z'$	2, 1		$N_{\text{ref}}, N_{\text{par}}$	2043, 140
$D_{\text{calc}}$ ( $\text{g/cm}^3$ )	1.054		$R, wR2, S$	0.0921, 0.2678, 1.16
$\mu - \text{MoK}\alpha$ (1/mm)	0.079		$w = 1/[\sigma^2(F_o^2) + (0.1029P)^2 + 0.4361P]$ , where $P = (F_o^2 + 2F_c^2)/3$	
$F(000)$	196		Max. and Av. shift/error	0.13, 0.01
Crystal size (mm)	0.30 x 0.90 x 1.70		Min. and Max. resd. dens. ( $\text{e}/\text{\AA}^3$ )	-0.24, 0.32

Table S2. Hydrogen bonds (Angstrom, Deg) of racemic pregabalin hydrate.

D-H...A	D-H ( $\text{\AA}$ )	H...A ( $\text{\AA}$ )	D...A ( $\text{\AA}$ )	D-H...A ( $^\circ$ )	symmetry operation
N1-H1A...O3	0.91	2.47	3.038(5)	120	within asymmetric unit
N1-H1A...O3	0.91	2.03	2.823(4)	145	1-x, 1-y, 1-z
N1-H1B...O2	0.91	1.85	2.756(4)	178	1+x, -1+y, z
N1-H1C...O1	0.91	1.87	2.766(4)	169	x, -1+y, z
O3-H3A...O1	0.85	1.92	2.753(4)	166	1+x, -1+y, z
O3-H3B...O2	0.96	1.82	2.767(4)	169	1+x, y, z

### S3. Prediction of resolution attributes in diastereomeric salt crystallization

The solubility product constants ( $K_{sp}$ ) are calculated from the measured pure component solubilities ( $x_{eq}$ ), where  $x_{S/R^+} = x_{L^-}$ :

$$K_{sp} = x_{S/R^+} * x_{L^-} = x_{S/R^+}^2 = x_{eq}^2$$

The multi-component system with the crystallized diastereomeric salts ( $n_{SLH}$ ,  $n_{RLH}$ ) was described in Table S3.

Table S3. The amount (mol) of the components in equilibrium: (*S*)-pregabalin, (*R*)-pregabalin, L-tartrate, and water.

	$n_{S^+}$	$n_{R^+}$	$n_{L^-}$	$n_{H_2O}$
Dissolved at the upper temperature	x	x	2x	1 - 4x
Crystallized at the lower temperature	$n_{SLH}$	$n_{RLH}$	$n_{SLH} + n_{RLH}$	$n_{SLH} + n_{RLH}$
Dissolved at the lower temperature	$x - n_{SLH}$	$x - n_{RLH}$	$2x - n_{SLH} - n_{RLH}$	$1 - 4x - n_{SLH} - n_{RLH}$

Solution at the upper temperature was assumed to be saturated:

$$K_{sp,SL}^{T1} = x_{S^+} * x_{L^-} = x * 2x = 2x^2$$

$$x = \sqrt{K_{sp,SL}^{T1}/2}$$

Total amount (mol) of dissolved components in equilibrium at the lower temperature:

$$(x - n_{SLH}) + (x - n_{RLH}) + (2x - n_{SLH} - n_{RLH}) + (1 - 4x - n_{SLH} - n_{RLH}) = 1 - 3(n_{SLH} + n_{RLH})$$

The mole fractions (listed in Table S4) were calculated by dividing specific amounts by the total amount.

Table S4. The mole fractions (-) of the dissolved components in equilibrium: (*S*)-pregabalin, (*R*)-pregabalin, and L-tartrate.

	$x_{S^+}$	$x_{R^+}$	$x_{L^-}$
Dissolved at the lower temperature	$\frac{x - n_{SLH}}{1 - 3(n_{SLH} + n_{RLH})}$	$\frac{x - n_{RLH}}{1 - 3(n_{SLH} + n_{RLH})}$	$\frac{2x - n_{SLH} - n_{RLH}}{1 - 3(n_{SLH} + n_{RLH})}$

When only (*S*)-PG-L-TA diastereomeric salt is crystallized ( $n_{RLH} = 0$ ):

$$K_{sp,SL}^{T2} = x_{S^+} * x_{L^-} = \frac{x - n_{SLH}}{1 - 3n_{SLH}} * \frac{2x - n_{SLH}}{1 - 3n_{SLH}}$$

Solution of the quadratic equation:

$$n_{SLH} = \frac{3x - 6K_{sp,SL}^{T2} - \sqrt{(6K_{sp,SL}^{T2} - 3x)^2 - (4 - 36K_{sp,SL}^{T2})(2x^2 - K_{sp,SL}^{T2})}}{2 - 18K_{sp,SL}^{T2}}$$

The criterium for initial assumption:

$$K_{sp,RL}^{T2} \geq x_{R^+} * x_{L^-} = \frac{x}{1 - 3n_{SLH}} * \frac{2x - n_{SLH}}{1 - 3n_{SLH}}$$

$$K_{sp,RL}^{T2} - \frac{x * (2x - n_{SLH})}{(1 - 3n_{SLH})^2} \geq 0$$

When the expression above is false, both the diastereomeric salts are crystallized:

$$K_{sp,SL}^{T2} = x_{S^+} * x_{L^-} = \left( \frac{x - n_{SLH}}{1 - 3(n_{SLH} + n_{RLH})} \right) * \left( \frac{2x - n_{SLH} - n_{RLH}}{1 - 3(n_{SLH} + n_{RLH})} \right)$$

$$K_{sp,RL}^{T2} = x_{R^+} * x_{L^-} = \left( \frac{x - n_{RLH}}{1 - 3(n_{SLH} + n_{RLH})} \right) * \left( \frac{2x - n_{SLH} - n_{RLH}}{1 - 3(n_{SLH} + n_{RLH})} \right)$$

Solutions to the equations:

$$n_{SLH} = \frac{K_{sp,SL}^{T2} + 3(K_{sp,RL}^{T2} - K_{sp,SL}^{T2})x - x \sqrt{K_{sp,SL}^{T2} + K_{sp,RL}^{T2}}}{3(K_{sp,SL}^{T2} + K_{sp,RL}^{T2}) - \sqrt{K_{sp,SL}^{T2} + K_{sp,RL}^{T2}}}$$

$$n_{RLH} = \frac{K_{sp,RL}^{T2} + 3(K_{sp,SL}^{T2} - K_{sp,RL}^{T2})x - x \sqrt{K_{sp,SL}^{T2} + K_{sp,RL}^{T2}}}{3(K_{sp,SL}^{T2} + K_{sp,RL}^{T2}) - \sqrt{K_{sp,SL}^{T2} + K_{sp,RL}^{T2}}}$$

Diastereomeric excess (%):

$$de = \frac{n_{SLH} - n_{RLH}}{n_{SLH} + n_{RLH}} * 100\%$$

Yield (%):

$$Y = \frac{n_{SLH} + n_{RLH}}{x} * 100\%$$

Selectivity (-):

$$S = \frac{de}{100\%} * \frac{Y}{100\%}$$

Productivity (mg/g water):

$$P = S * \frac{x}{1 - 4x} * \frac{M_{SLH/RLH}}{M_{water}}$$

where  $M_{SLH/RLH}$  and  $M_{water}$  are the molar masses of the diastereomeric salt hydrates (309.3 g/mol) and water (18.0 g/mol).