Tracking Dehydration Mechanisms in Crystalline Hydrates with Molecular **Dynamics Simulations**

Anders S. Larsen,^{†,§} Michael T. Ruggiero,^{‡,§} Kristoffer E. Johansson,[¶] J. Axel Zeitler,[‡] and Jukka Rantanen*,†

Department of Pharmacy, University of Copenhagen, Universitetsparken 2, 2100 København Ø, Denmark, Department of Chemical Engineering and Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge, CB3 0AS, United Kingdom, and Department of Biology, University of Copenhagen, Ole

Maaløes Vej 5, 2200 København N, Denmark

Received August 15, 2017; E-mail: jukka.rantanen@sund.ku.dk

Abstract:

Dehydration of crystalline solids is a widespread phenomenon, yet the fundamental mechanisms by which dehydration occurs is not properly understood. This arises due to technical limitations in studying such fast processes with sufficient sensitivity, nevertheless understanding dehydration pathways is critical for designing optimal properties for materials, particularly in the case of pharmaceutical solids. The computational methods presented here allow for accurate determination of the dehydrated species' crystal structure and to develop an understanding of the mechanism of dehydration at the molecular level. This work also highlights the critical role of explicitly taking into account the dynamical aspect of molecules using computational techniques, rather than relying on static energy minimization approaches. Specifically, the crystalline active pharmaceutical agent naproxen sodium, and its hydrates, is studied in silico using density functional theory and molecular dynamics, ultimately elucidating the face-specific dehydration mechanisms and revealing highly complex diffusion and nucleation behaviour. Additionally, the results indicate that the method is a viable way to explore dehydration pathways and predict new dehydrated crystal structures.

Hygroscopicity and water-solid interaction of crystalline solids represent a critical aspect of material engineering, given that the ability of a solid to absorb or release water dictate a vast amount of potential applications.¹ This is especially important for pharmaceutical crystals, where the formation of hydrates and solvates can directly impact bulk characteristics such as drug solubility and stability.²⁻⁴ Specifically, the dehydration process has attracted a great deal of attention since it is of key importance to storage and processability. While experiments have provided kinetic and thermodynamic insight into dehydration at the macroscopic level, a full understanding of dehydration mechanisms at the molecular level is still lacking. A common technique for investigating dehydration mechanisms is through the detailed comparison of the X-ray crystal structures of the hydrate and dehydrated species. Unfortunately, this often leads to inconclusive results, especially when the crystal forms have very different structures. Different mechanistic descriptions

of dehydration such as nucleation and growth, 1D/2D/3Ddiffusion has lead to the development of kinetic models, which are usually empirically fitted to experimental data such as thermal gravimetric analysis (TGA) data. Again, the results are often not suitable to discriminate between different dehydration models since many of them fit the data equally well. 5,6

Theoretical techniques can be used to obtain a great deal of insight into the energetics and related properties of crystalline materials.^{7–10} Common density functional theory (DFT) simulations are not the most suitable computational approach to effectively probe the mechanistic aspect of hydration, as they typically can only be used to determine the energies of the lowest energy structures. However, it has been shown that DFT can be used in certain cases where the structural changes associated with the dehydration are relatively small. In one such case it was shown that the dehydration process can be approximated by performing a simple energy minimization of the hydrate crystal structure with the water molecules deleted.¹¹ It is important to note that, whilst such a method can ultimately predict the correct structure of the dehydrated species, it provides no information related to the actual mechanism of dehydration. Moreover, since the process of dehydration is significantly dependent on thermal energy it is likely that the dynamics are critical, not just for forcing the water molecules out of the crystal but also for the structural change. Understanding the pathway between hydrates and dehydrated products is therefore critical to gain molecular level insight into dehydration.



Figure 1. The two experimental unit cells of sodium naproxen hydrates. Panel (a) shows the triclinic P1 DH-I structure, while (b) shows the monoclinic $P2_1$ structure.

[†]University of Copenhagen, Department of Pharmacy

[‡]University of Cambridge, Department of Chemical Engineering and Biotechnology

[¶]University of Copenhagen, Department of Biology [§]Contributed equally to this work

In this work, we investigate the dehydration of naproxen sodium, a common nonsteroidal anti-inflammatory drug that has been extensively studied using a large number of experimental and theoretical techniques (Figure 1). $^{12-14}$ The many detailed studies on the static structures of naproxen sodium hydrates, coupled with the lack of an appropriate molecular mechanism of dehydration, make this a good model system for understanding such processes in hydrates of salt compounds.^{15–17} Three different computational approaches are used, the first is a straightforward minimization of the structure using DFT, while the second and third use molecular dynamics $(MD)^{18-21}$ to elucidate the importance of dynamics to the dehydration mechanism. The crystal structures of the dihydrate form I (DH-I) and the monohydrate (MH) form of sodium naproxen have been solved previously^{22,23} and are presented in Figure 1. Experimental dehydration of DH-I occurs at temperatures above 298 K and results in the MH structure. At higher temperatures, DH-I dehydrates directly to the anhydrate, but this processes is believed to progress via formation of the monohydrate, as evidenced by a step-wise water loss in the TGA data.²⁴

Initially the naproxen crystal structures were optimised via fully-periodic DFT simulations with the CRYSTAL14 quantum mechanical software package²⁵ coupled with the triple-zeta 6-311G(2d,2p) basis set²⁶ and the D3-corrected^{27,28} Perdew-Burke-Ernzerhof²⁹ (PBE) functional. The initial atomic coordinates were taken from the experimental structure. The energy minimization resulted in average errors on the unit cell of 1.32 % for the DH-I and 1.43 % for the MH compared to the experimental structures, more details about the energy minimization is available in the SI.

Six models of a dehydrated unit cell were prepared by deleting two out of four water molecules in the energy minimised DH-I unit cell. The six structures were fully optimized without any symmetry constraints. The resulting dehydrated unit cells all yielded distorted unit cells compared with the experimental MH structure (see Table 1), with none of the dehydration models transforming to a $P2_1$ monoclinic system. In none of the six models did the remaining DH-I hydrogen bonds change to having hydrogen bonds between carboxylate groups and water molecules along the *b*-axis like in the MH. All the six dehydration models have energies substantially higher than the optimised MH structure and fall in the range of 23.10 to $50.93 \text{ kJ mol}^{-1}$ relative to the MH energy. The unit cell of model 3 looks deceptively close the MH structure but the water molecules have not moved significantly resulting in a hole in the structure and thus the high energy.

Table 1. DFT optimized lattice parameters of DH-I, MH, and the six dehydrated DH-I structures. The relative energies of the monohydrate structures (in relation to the experimentally observed structure) are provided.

Structure	A	B	C	α	β	γ	ΔE
	(A)	(A)	(A)	(deg)	(deg)	(deg)	$(kJ mol^{-1})$
DH-I exp	22.81	5.81	5.43	89.53	85.53	92.61	
DH-I	21.90	5.75	5.36	88.71	86.38	93.26	
MH exp	21.78	5.79	5.44	90.00	91.41	90.00	
MH	20.96	5.66	5.38	90.00	91.10	90.00	0.0
Model 1	23.22	5.64	5.01	84.90	85.87	93.20	50.93
Model 2	21.68	5.55	5.25	95.81	89.10	88.03	23.10
Model 3	22.27	5.70	4.95	90.82	91.20	90.61	39.12
Model 4	22.67	5.74	4.86	88.84	89.73	90.65	46.47
Model 5	22.27	5.45	5.30	88.51	87.09	93.19	42.77
Model 6	21.84	5.64	5.10	89.51	90.19	89.61	31.83

The inability of static DFT energy minimizations to accurately predict the correct monohydrate structure, even though the structural differences between DH-I and MH are on a similar magnitude to those of the morphinanes studied by Braun et al.,¹¹ highlights the importance of the dynamical nature of solid state structures in this system. Two approaches to explore the dehydration process based on MD simulations were therefore investigated. The first method involved the random removal of water molecules at regular intervals of 200 ps from a fully periodic supercell of the DH-I structure until half of the water molecules were fully removed. The second method consisted of modelling a periodic slab, with a corresponding vacuum region above the surface, which allows for the physical diffusion of the water molecules. The vacuum region was added to three faces of the crystal, face A with miller index (± 100) , face B with Miller index (0 ± 10) and face C with Miller index (00 ± 1) , which were judged to be the most important to dehydration based on a Bravais, Friedel, Donnay and Harker³⁰ morphology calculation. The MD simulations were performed using the GROMACS 5.1.1 software package³¹ together with the CHARMM General Force Field (CGenFF).^{32,33} More information about the computational setup is described in the SI.



Figure 2. (A) Change in supercell angles during dehydration using the random removal of water from a 3D-periodic model. (B) water loss during dehydration within the slab/vacuum model.

The first MD dehydration method resulted in a structure in which the DH-I supercell angles resembled those of the MH at 30 ns simulation time, Figure 2a. For the first 5 ns the removal of water molecules did not significantly disrupt the integrity of the hydrogen bond network, but between 10-20 ns the hydrogen bond network breaks down and the wa-

ter/sodium region becomes severely disordered. After 20 ns, features of the MH structure can be observed and around 30 ns the system has fully relaxed into the MH structure. During the simulation the α and γ angles both transformed until they reached 90° while the β angle changed to an average value of 92.2° . The smooth change of the super cell angles suggest that no stable intermediate structure exists between the DH-I and the MH. By energy minimising the dehydrated system with the force field, the unit cell and space group $P2_1$ was detected, which indicates a triclinic (P1) to monoclinic $(P2_1)$ transformation. The ability of the force field to model the DH-I and MH was checked by comparing the simulated unit cells with the experimental cell parameters. The DH-I has a average error on the unit cell of 0.83% and the MH errors of 1.25%, which is in excellent agreement with both the experiment and the static-DFT geometry minimizations.

While the periodic removal of water molecules from a supercell did result in the correct MH structure, such simulations are a poor representation of the real physical mechanism given that water molecules will naturally diffuse through the crystal to the external environment rather than disappear in discrete steps. In order to get a more physical picture, the second method of MD dehydration was performed, which involved the simulation of a solid separated from an adjacent vacuum region. Dehydration was observed from the face B and C crystal surface during the 400 ns of simulation time (see Figure 2b), while face A did not experience any water loss, see SI Figure S2. For face C, the loss of water occurred at a constant rate until around 160 ns, at which point the partially dehydrated surface layers became significantly disordered while the DH-I structure remained intact deeper inside the crystal (Figure 3a). At 160 ns the hydrogen bonding network of the bulk crystal broke down and an increase in the rate of water loss was observed. The structure exhibited highly disordered surface layers with trapped water molecules in the bulk, yet the MH structure clearly emerges, (see Figure 3b). Trapped water molecules could indicate that the dehydration stops from face C after the top layers transform to the MH structure. Like face C, the dehydration from face B initially commenced at a constant rate with water molecules originating from the surface layers emerging from the crystal face until 50 ns (Figure 2b). At this point the bulk of the crystal started to dehydrate. Unlike in the case of face C, the water/sodium layers of face B did not dehydrate at the same time and this resulted in the variability in the loss of water. Since the disordered surface layers were anhydrous the ratio of water to naproxen decreased below 1.0. The dehydrated structure was characterised by severely disordered surface regions in addition to 7-8 well formed layers of monohydrate (Figure 3a). The full visualisation including the disordered surface regions is available in the SI (Figure S3).

The results of the simulations clearly show that water molecules readily leave the crystal through faces B and C, while no water molecules can be observed crossing the ordered hydrophobic layers of naproxen molecules of face A. The water molecules undergo 2D diffusion in the water/sodium channels of the ordered hydrate, but this view is further complicated by the rate of water loss from the two surfaces differing significantly. The fact that a stable MH structure is found means that the simulation predicts the MH to be more resistant to thermal dehydration compared with the DH-I, which is in good agreement with previous



Figure 3. An interior volume of naproxen sodium before and after dehydration with corresponding unit cells superimposed. (A) Left: simulation of face B at 0 ns, Right: simulation at 400 ns. (B) Left: simulation of face C at 0 ns, Right: simulation at 400 ns.

experimental work. $^{\rm 24}$

The four different water molecules of the DH-I unit cell were tracked during the dehydration simulation (Figure 2b). Neither crystal face showed significant differences in dehydration rates between the four respective water molecules, except for the beginning of the simulation for face B where water type 3/4 are more likely to leave the crystal suggesting that 3/4 dehydrate easier from the surface layers than type 1/2. Figure 4 shows the displacement of water molecules at 400 ns from their initial position. It can be concluded that almost all of the water molecules experience diffusion during the dehydration and that the final water molecules in the MH unit cell is rarely from the corresponding DH-I unit cell.



Figure 4. Magnitude of total displacement of water molecules at t = 400 ns of the dehydration simulation relative to their initial position at t = 0.

The analysis of the dehydration process based on using DFT energy minimizations alone failed to produce the desired MH unit cell. We therefore assume that in the case of sodium naproxen the transformation between the DH-I and MH forms involves a far more rugged energy landscape with moderate barriers and local minima. In such a scenario it would not be surprising that simple energy minimization is an insufficient technique to derive the structure of the molecular crystal following dehydration, even for a simple system.

By explicitly including the dynamics and temperature by means of MD simulations it is possible to overcome the limitations of the DFT approach and successfully cross the energy barriers between the DH-I and MH phases. The system undergoes a transient highly disordered state during the dehydration before it eventually relaxes into the MH form. Interestingly, the disorder is initially spatially confined to a random region of the supercell and quickly spreads to the whole system, which suggests that the process is a first order phase transition taking place by nucleation. The dynamics of the disordered state help the water molecules, ions and naproxen molecules to navigate the energy landscape connecting the DH-I and MH forms. In the pure DH-I movement of the molecules is restricted by the tight packing of the crystal while during dehydration the disorder allows for a wider sampling of angles, torsions and conformations.

By simulating the system using a surface and a vacuum region it was possible to resolve the complex dehydration behaviour in far greater detail. Our results can be interpreted that, at the molecular level, the dehydration of sodium naproxen DH-I does not follow any of the simple models of dehydration but can be considered a mix of diffusion and nucleation events.

The MD simulations indicate that the experimentally observed macroscopic water loss profile constitutes the average of many different microscopic dehydration rates that originate from different regions of the crystal. Additionally, the rate of dehydration from each crystal face varies with time, which means that obtaining mechanistic explanations from experimental data will be difficult. It is hence not surprising that simple fitting of the experimental macroscopic dehydration data, which is typically measured from polycrystalline samples spanning a range in particle size, against any of the available kinetic models leads to inconclusive results. Instead it would be more sensible to measure the face-specific dehydration from each face of a single crystal before fitting the data to any model, yet this is technically too challenging to achieve at present to the best of our knowledge.

Another interesting result is that using the MD simulation it was possible to successfully predict the MH structure from the DH-I structure alone. This result also highlights the potential for using MD simulations in the context of crystal structure prediction of hydrate forms. Clearly further work will be required to establish the general applicability and limitations of the method to explore new dehydration pathways.

By comparing DFT energy minimizations and MD simulations we were able to highlight the important role of the molecular dynamics and temperature in dehydration processes for organic molecular crystals. Whilst it was previously demonstrated how the minimization method can be used for such applications our results highlight some of the limitations of this approach in that the numerical method can get stuck in local energy minima. In contrast, we found that using the MD method the experimentally observed MH structure was correctly predicted.

Including a vacuum region in the MD simulation makes it possible to simulate a physically plausible diffusion process associated with the dehydration process. Whilst we fully acknowledge that, in the absence of unequivocal experimental observations, this molecular mechanism may not be the definitive mass transport process it nonetheless seems to provide a plausible mechanistic route. Here, the water molecules of the DH-I form were observed to diffuse along the water channels of the crystal hydrate into the vacuum region. The mass transport was followed by a crystal transform to the experimental MH structure. While it would be natural to extend the MD simulations by dehydrating the MH to the anhydrate, this mechanism involves the global rearrangement of naproxen molecules, requiring significantly longer time scales and possibly the explicit inclusion of nucleation sites, which would require additional method development. Additionally, this technique can be further developed to elucidate the role of the humidity in dehydration processes, by explicitly including water molecules in the vacuum region and varying the temperature, also yielding information related to the hydration mechanisms in molecular crystals.

We conclude that the dehydration of DH-I is a complicated mix of 2D diffusion and nucleation events which results in complicated dehydration profiles for single crystals of naproxen sodium. This means that the experimental data must be regarded as the average of many different dehydration rates that occur concomitantly over many surfaces. Given the polycrystalline nature of naproxen sodium DH-I in any technical applications, which additionally is complicated by various length scales of porosity when analysing a bulk sample, it is clear that there is little resemblance between the dehydration rates predicted at the molecular level when compared to the macroscopic measurement. In turn this means that the mechanistic analysis of dehydration based on such macroscopic observations is intrinsically limited. It is of critical importance to be mindful of this fundamental limitation when analysing experimental data with a view to exploring the molecular mechanism.

Acknowledgement A.S.L, K.E.J and J.R. gratefully acknowledges the Villum Foundation (Denmark) for financial support (project No. VKR023111). M.T.R. and J.A.Z. would like to acknowledge funding from the U.K. Engineering and Physical Sciences Research Council (EP/N022769/1).

References

- (1)Khankari, R. K.; Grant, D. J. Thermochimica Acta 1995, 248, 61 - 79
- (2)Galwey, A. K. Thermochimica Acta 2000, 355, 181 - 238.
- Giron, D.; Goldbronn, C.; Mutz, M.; Pfeffer, S.; Piechon, P.; Schwab, P. Journal of Thermal Analysis and Calorimetry **2002**, 68, 453-465.
- Petit, S.; Coquerel, G. Chem. Mater 1996, 8, 2247 2258. (4)Okoth, M. O.; Vrcelj, R. M.; Sheen, D. B.; Sherwood, J. N. CrystEngComm **2013**, 15, 8202–8213. (5)
- Zhou, D.; Schmitt, E. A.; Zhang, G. G.; Law, D.; Wight, C. A.; Vyazovkin, S.; Grant, D. J. Journal of Pharmaceutical Sciences
- **2003**, 92, 1367 1376. Reilly, A. M. et al. Acta Crystallographica Section B 2016, 72, (7)439 - 459.
- Pulido, A. et al. Nature 2017, 543, 657-664.
- van de Streek, J.; Neumann, M. A. Acta Crystallographica Section B 2010, 66, 544-558.
- Moellmann, J.; Grimme, S. The Journal of Physical Chemistry (10)C 2014, 118, 7615–7621.
- Braun, D. E.; Gelbrich, T.; Kahlenberg, V.; Griesser, U. J. (11)Molecular Pharmaceutics 2014, 11, 3145-3163, PMID: 25036525
- Mason, L.; Edwards, J.; Moore, R.; McQuay, H. Cochrane (12)Database Syst Rev 2004, 4.
- Malaj, L.; Čensi, R.; Martino, P. D. Crystal Growth & Design (13)**2009**, *9*, 2128–2136. (14)
- Kim, Y.; Paskow, H. C.; Rousseau, R. W. Crystal Growth & Ann, T.; Faskow, H. C.; Nousseau, R. W. Crystal Growth & Design 2005, 5, 1623–1632.
 Chen, L. R.; Jr., V. G. Y.; Lechuga-Ballesteros, D.; Grant, D. J. Journal of Pharmaceutical Sciences 1999, 88, 1191 – 1200.
 Khankari, R.; Chen, L.; Grant, D. J. W. Journal of Pharma-tic Logic Logic Logic Logic Logic Logic Logic Logic Logic (15)
- (16)
- ceutical Sciences 1998, 87, 1052-1061.
- Kons, A.; Rutkovska, L.; Berzins, A.; Bobrovs, R.; Actins, A. (17)CrystEngComm 2015, 17, 3627-3635

- De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Journal of Medicinal Chemistry 2016, 59, 4035–4061, PMID: 26807648.
 Nemkevich, A.; Burgi, H.-B.; Spackman, M. A.; Corry, B. Phys.
- Chem. Chem. Phys. 2010, 12, 14916-14929.
- (20) Larsen, A. S.; Rantanen, J.; Johansson, K. E. Journal of Pharmaceutical Sciences 2017, 106, 348–355.
 (21) Hamad, S.; Moon, C.; Catlow, C. R. A.; Hulme, A. T.; Price, S. L. The Journal of Physical Chemistry B 2006, 110, 2222, 2200, DMID. 16404246 3323–3329, PMID: 16494346.
- (22) Bond, A. D.; Cornett, C.; Larsen, F. H.; Qu, H.; Rantanen, J.
- (22) Bond, N. D., Ornett, C., Larsen, F. H., Qu, H., Randanen, S. IUCrJ 2014, 1, 328–337.
 (23) Kim, Y. B.; Park, I. Y.; Lah, W. R. Archives of Pharmacal Research 1990, 13, 166–173.
 (24) Raijada, D.; Bond, A. D.; Larsen, F. H.; Cornett, C.; Qu, H.;
- Rantanen, J. Pharmaceutical Research 2013, 30, 280-289.
- Dovesi, R.; Orlando, R.; Erba, A.; Zicovich-Wilson, C. M.; Civalleri, B.; Casassa, S.; Maschio, L.; Ferrabone, M.; De La Pierre, M.; D'Arco, P.; Noël, Y.; Causà, M.; Rérat, M.; Kirtman, B. International Journal of Quantum Chemistry **2014**, *114*, 1287–1317. (25)
- (26) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. The Journal of Chemical Physics 1980, 72.
- (27) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. The Journal of Chemical Physics 2010, 132, 154104.
 (28) Grimme, S.; Ehrlich, S.; Goerigk, L. Journal of Computational Chemistry 2011, 32, 1456–1465.
- (29)Perdew, J.; Burke, K.; Ernzerhof, M. Physical Review Letters
- **1996**, *77*, 3865–3868. (30) Donnay, J. D. H.; Harker, D. American Mineralogist 1937, 22,
- 446 467Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; (31)
- Hess, B.; Lindahl, E. SoftwareX 2015, 1-2, 19-25.
- Hess, B.; Lindahl, E. SoftwareX 2015, 1-2, 19–25.
 (32) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Macherell JR, A. D. Journal of computational chemistry 2009, 31, 671–690.
 (33) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Journal of Chemical Theory and Computation 2012, 8, 3257–3273, PMID: 23341755.

For Table of Contents Use Only

Tracking Dehydration Mechanisms in Crystalline Hydrates with Molecular Dynamics Simulations
Anders S. Larsen^{1,†}, Michael T. Ruggiero^{2,†}, Kristoffer E. Johansson³, J. Axel Zeitler², Jukka Rantanen^{*,1}
[†]Contributed equally to this work
¹ University of Copenhagen, Department of Pharmacy, University of Copenhagen, Department of Pharmacy, University of Cohemical Engineering and Biotechnology,
University of Cambridge, Philippa Fawcett Drive, Cambridge, CB3 0AS, United Kingdom
³ University of Copenhagen, Department of Biology, Ole Maaløes Vej 5, 2200 København N, Denmark



Synopsis

The dehydration of naproxen sodium dihydrate was investigated at the molecular level using the molecular dynamics based simulation. The partial dehydration of this dihydrate into the correct, experimentally reported, monohydrate was observed and the results indicate that the energy landscape crossed during this dehydration process is highly rugged. It was further observed that the computational dehydration rates at different faces of the crystal were not identical and indicated the complexity of the involved diffusion and nucleation processes.