

Organic Process Research & Development

Subscriber access provided by American Chemical Society

[View the Full Text HTML](#)



ACS Publications
High quality. High impact.

Organic Process Research & Development is published by the American Chemical Society, 1155 Sixteenth Street N.W., Washington, DC 20036

Structural and Thermodynamic Explanations of Ostwald's Rule

Terry Threlfall

Chemistry Department, Southampton University, Highfield, Southampton SO17 1BJ, UK

Abstract:

Ostwald proposed that the solid first formed on crystallisation of a melt or a solution would be the least stable polymorph. It is shown that this can be rationalized on the basis of irreversible thermodynamics, structural relationships, or a combined consideration of statistical thermodynamics and structural variation with temperature. It is concluded that the rule cannot be a universal law but is only in the nature of a possible preferred tendency.

Introduction

In 1897, Wilhelm Ostwald published a truly remarkable paper, entitled “Studien ueber die Bildung und Umwandlung fester Koerper” (Studies of the formation and transformation of solid substances).¹ In it the role of seeding of supercooled melts and supersaturated solutions was discussed, including the spatial and temporal role of seeds in the atmosphere in causing adventitious seeding. He was particularly concerned to establish just how small a quantity of seed crystal was needed to bring about crystallisation. This was attempted by using serial 10-fold dilutions “after the manner of the homeopaths” by trituration of the crystalline substance in an inert powder such as lactose or quartz. He discussed the physical state of the active constituent at dilutions at which it was no longer effective, the impossibility of removing the activity of hydrated crystal seeds by allowing them to effloresce (a consequence of reversible equilibrium), the dependence of the stability of a metastable polymorph on its method of preparation, the relationship between enantiotropy and monotropy, and the problem highlighted by the ability of isomorphous salts to seed as effectively as the actual salt, even though the solution is not saturated with respect to the isomorphous salt. He had earlier been responsible for the concept of metastable and labile zones in the crystallisation process and took the opportunity in this paper to explain and distinguish their properties, to counteract misconceptions which had arisen. He discussed also the analytical application of his results to detection limits down to 10^{-11} g, which equalled the best contemporary spectrochemical detection limits. All this was based on a few simple experiments carried out by visual examination on supercooled melts and solutions in test tubes and crystallisation dishes or by low-powered microscopic examination of drops on microscope slides. He also formulated from some of these observations what has subsequently become known as Ostwald's Law of Stages or Ostwald's Rule, namely that, in the case of a compound capable of crystallising in several

forms, it will be the least stable form which is first produced by spontaneous crystallisation, followed successively by forms of increasing stability. It is with this part of his paper that we shall be concerned here. Ostwald justified the rule on the basis of an argument of classical thermodynamics, by analogy with vapor–liquid equilibration near the critical temperature and pressure region. There have been previous discussions of Ostwald's work,² including attempts at thermodynamic justifications.^{3–6} We shall be revisiting Ostwald's analysis and in addition shall justify the rule on the basis of crystal structure concepts and of irreversible and statistical thermodynamics. In doing this it is hoped that the basis of the rule, its extent, and its limitations may all be clarified. The discussion here is centered on polymorphic forms. It is applicable, *mutus mutandi*, to hydrates and solvates, but the need for another component (water, solvent of solvation) leading to the necessity to involve the thermodynamic activity of the other component complicates such a discussion.

Ostwald's Rule

The most significant sentence in this paper is the one which has subsequently become commonly known in English as Ostwald's Rule or Ostwald's Law of Stages: “When leaving a given state and in transforming to another state, the state which is sought out is not the thermodynamically stable one, but the state nearest in stability to the original state”. This must be the next least stable state. For example, it implies that, in a crystallisation from the melt or from solution, the solid first formed will be that which is the least stable of the polymorphs, the one with the largest Gibbs free energy. The word which Ostwald uses in referring to this sentence is “Satz”. “Satz” in German can imply a vast range of meanings, from “statement” or “sentence” through “proposition” to “rule” or even “law”. The word itself therefore gives no indication of the strength which may be applied to the sentence, and so Ostwald's view of it must be sought within the paper. There are several examples which suggest that he might have regarded it as a universal law. First, he quotes how widely applicable the rule is and refers not only to different solid forms but also to the production of isomers. For example, the vapor of cyanic acid and that of cyanuric acid are identical, but from whatever source the vapor is produced, it deposits only cyanic acid, despite this being so

(2) Zott, R. *Angew. Chem., Int. Ed.* **2003**, 42, 3990.

(3) Cardew, P. T.; Davey, R. *Symposium Papers, Inst. Chem. Eng. North Western Branch* **1982**, 2, 1.

(4) Dufor, L.; Defay, R. *Thermodynamics of Clouds*; Academic Press: New York, 1963.

(5) Dunning, W. J. *Zettlemoyer* **1969**, 1.

(6) Nyvlt, J. *Cryst. Res. Technol.* **1995**, 30, 445.

(1) Ostwald, W. *Z. Phys. Chem.* **1897**, 22, 289.

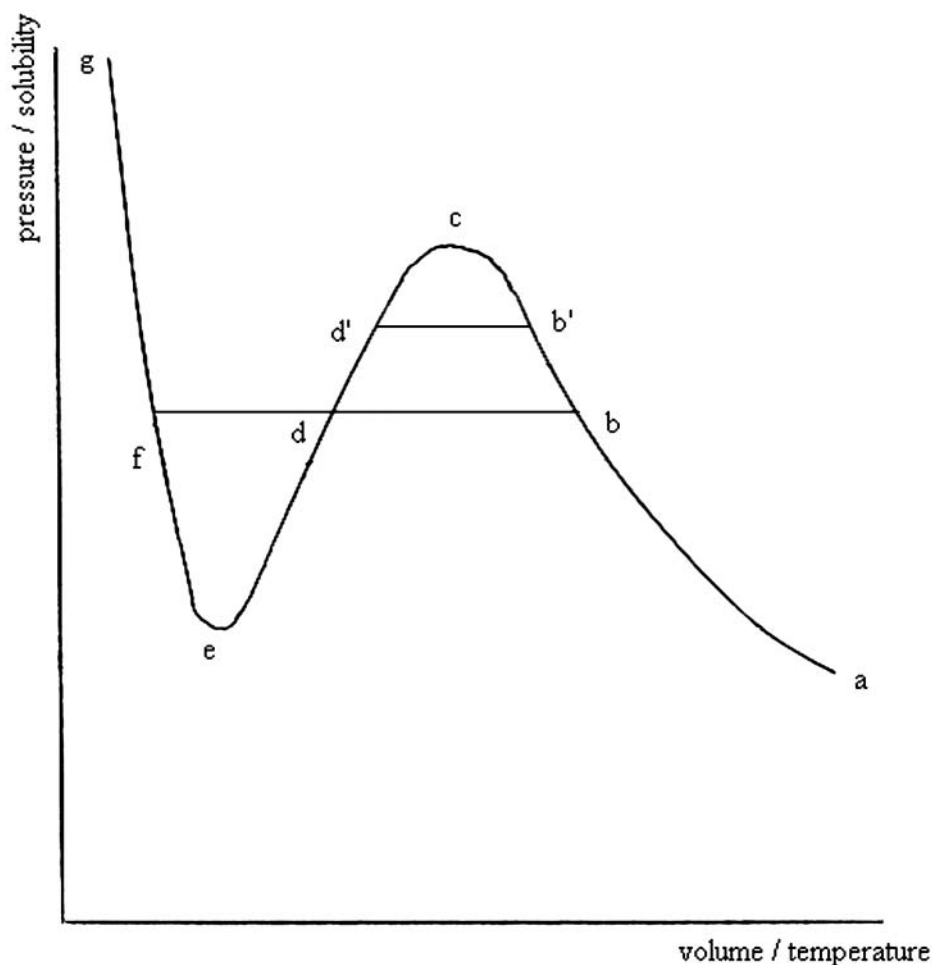


Figure 1. Curve *ab*, followed by the straight line through to *f*, then up to *g*, is the normal pressure–volume representation of gas–liquid equilibrium. J. J. Thompson is quoted by Ostwald¹ as first showing that it was possible to bypass the condensation of liquid during the compression of gas to produce the third-order curve, *abcdefg*. In Ostwald’s analogy *ab* is the undersaturated solution or the melt above the melting point, *bc* is the metastable region, and *cd* or *ce* is the labile region. If the diagram is generalised by the addition of a third orthogonal axis of temperature, then *c* can be recognised as a point on the catastrophe surface dividing the “stable” (metastable) region from the “catastrophic” (labile) region. In Figure 5.5 of Saunders, R. T. *An Introduction to Catastrophe Theory*; Cambridge University Press, 1980, the catastrophe–recovery cycle is closely analogous to the crystallisation/dissolution or crystallisation/melting processes under discussion.

unstable as to polymerise explosively on gentle warming to cyanmelide, rather than depositing the stable cyanuric acid. He goes on to state that he knows of no example where two compounds give identical vapors, where it is not the least stable compound which is condensed. He even invokes chemical reactions, for example, potassium hydroxide and chlorine yield potassium chloride and potassium hypochlorite rather than potassium chloride plus hydrogen or potassium chloride plus potassium chlorate. Second, in referring to crystallisation from solution, he points out that it is the liquid rather than the solid which is the next most stable form and quotes the example of acidification of alkaline benzoic acid solutions which produces benzoic acid initially as droplets. In condensing from the vapor, sulfur first appears even at temperatures below the melting point as a liquid, which then solidifies. He goes on to state that even in those cases where a liquid is not first observed this may be due to a lack of a means of observing it before it transforms further to a solid. The proposition that “if you cannot observe it, it must nevertheless be there” is a curiously unscientific statement from so eminent a scientist, unless he is using it in the context

of a universal law and merely justifying why some cases may not be observable. Although he expresses reservations about this philosophy at his paragraph 37, these and other examples suggest that he regarded the rule as a thermodynamic inevitability and that, where the sequence was not observed, this was due only to the rapid kinetics of further transformation to a more stable form. Ostwald’s authority probably explains why in earlier times it was referred to as Ostwald’s Law. It has become clear over the century since Ostwald’s paper, that, although it is a useful indicator of a possible sequence of production of crystalline forms, it is not as universal and reliable⁷ (as it is suggested here) that Ostwald had supposed. This is reflected in the increasing modern usage of the term “Ostwald’s rule” or “Ostwald’s principle” rather than “Ostwald’s law”.

Ostwald’s Thermodynamic Explanation of the Rule

The thermodynamic justification which Ostwald proposed is an ingenious one based on an analogy with gas–liquid

(7) Bernstein, J. *Polymorphism in Molecular Crystals*; Clarendon Press: Oxford, 2002.

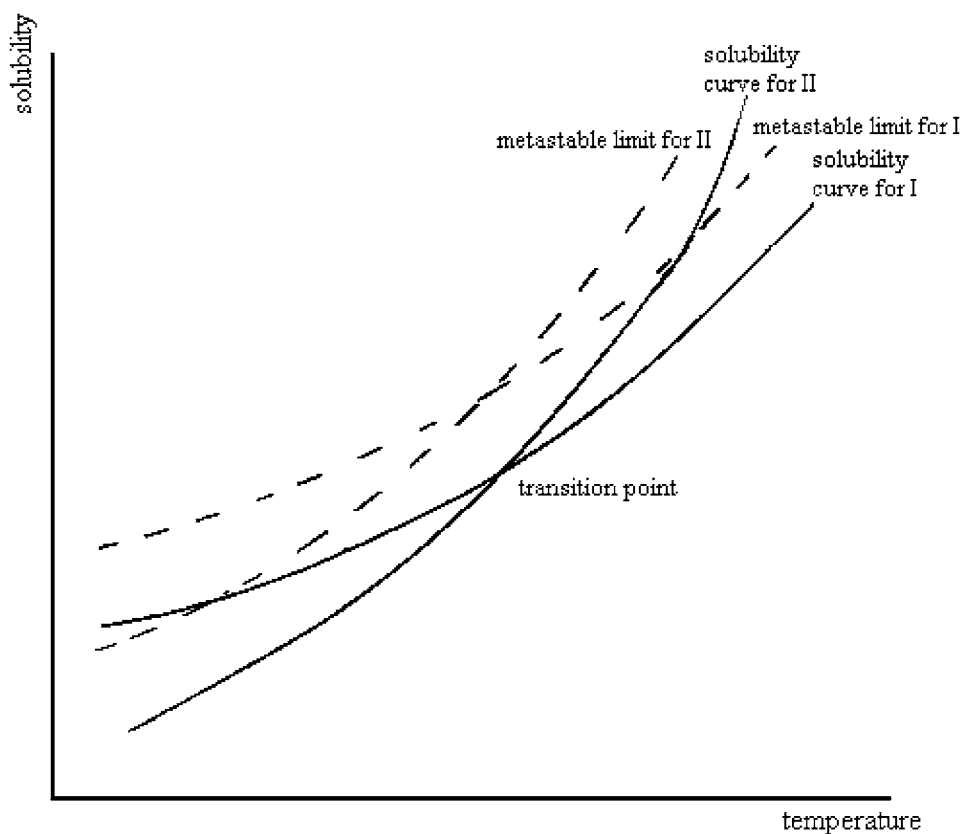


Figure 2. Usual representation of the solubility curves and metastable zone limits of a dimorphic system. The solubility curves and particularly the metastable zone limits are difficult to establish experimentally for metastable forms, but measured metastable zone widths⁸ are nearly always of the order of 10–50 °C.

equilibria, as shown in Figure 1. On increasing the pressure on a gas it decreases in volume according to the hyperbola *ab* until the point *b* at which condensation can take place. If it does so, the system follows the straight line *bdf* representing the heterogeneous system of gas plus liquid with varying proportions of the two phases. At point *f* the system is totally liquid, and its low compressibility results in the very steep line *fg*. If, on the other hand, no condensation of liquid takes place at *b*, then the compressed gas follows the curve *bc* up to the point *c* at which the system becomes unstable. In a similar way, it is asserted that the liquid can be expanded, with some difficulty, before vapor is generated, resulting in the curve *fe*. In the region of the diagram *cde* the slope of the curve is reversed; decreasing pressure is associated with decreasing volume, so this is truly an unstable situation in which spontaneous change occurs. In the analogy drawn by Ostwald, *bc* is the supersaturated or supercooled metastable zone for the crystallising liquid case, and *cd* is the supersaturated or supercooled labile region. Approached from the other direction, *fe* represents superheated solid. Although Ostwald recognised that the equilibrium (lowest free energy) situation at *b'* can be by-passed, and indeed must be by-passed in order to obtain a supercooled melt, he does not comment that the same can happen at *d'* so that the least stable form may not be produced. Presumably, this is because he was emphasising the distinction between the metastable region *bc* and the labile region *cd*. However it is a matter of common knowledge that a solution can deposit one or more crystals, yet take a long time, even months, before crystal-

lising to the point of exhausting the saturation of the solution. Clearly, kinetics is the dominant feature in the production of crystals even in the monomorphic case. Whether a given polymorph will or will not be produced in competition with the stable form, either from a labile or a metastable solution, must be a kinetic competition reliant on the rate at which the molecules can assemble in the different packings associated with the different forms. One encounters also an analogy in the common situation in which a polymorph stable at high temperature is cooled below a transition point without transition occurring. In this case the reduced mobility of the molecules in the crystal must have a significant effect on the kinetics, but clearly there is no difference in principle. The very existence of different forms at a given temperature is proof of the triumph of kinetics over thermodynamics.

A further conclusion which Ostwald drew from analysis of this diagram is that the metastable zone of a metastable form must always lie totally within the metastable zone of the stable form, i.e., *b'* and *d'* must always be within the limits of the line *bd*. He was supported in this belief by observations on the crystallisation of sodium sulfate. The first product of crystallisation of highly supersaturated aqueous sodium sulfate solutions in sealed tubes is the heptahydrate rather than Glauber's salt, the decahydrate (paragraph 32 of Ostwald's paper). The spontaneous crystallisation in a sealed system indicated that the heptahydrate must be in its labile region even though it must be nearer to its solubility curve than the less-soluble decahydrate is to its curve. This comment is extended to the solidification of

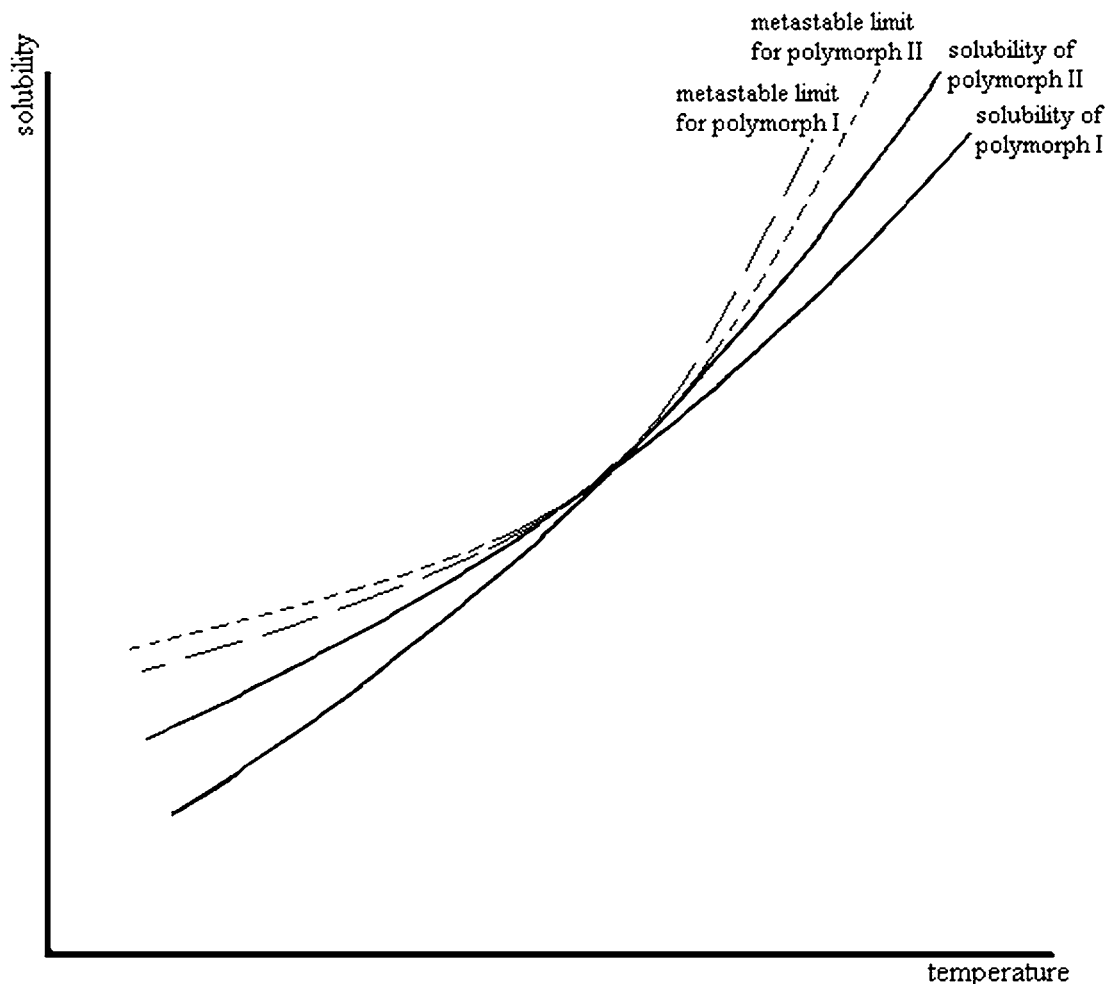


Figure 3. Consequence of the thermodynamic analogy presented by Ostwald is that the region b'd' of Figure 1 must always lie totally within the region bd. This restricts the metastable region of the metastable forms to the area totally within the metastable zone of the stable form. The highly improbable consequence of this is shown in the diagram.

melts and vapours and even homogeneous chemical reactions. Although he recognised some differences between the vapor–liquid and the solution–crystal cases (such as the difficulty in exceeding the melting point without causing melting, in contrast to the ease of superheating a liquid), he ascribed this difference to surface energies, and concluded that the analogy appeared reliable.

Despite the evidence just quoted and the apparent theoretical basis, this cannot be correct. Drawing the metastable zone in the usual manner for a dimorphic system, as in Figure 2, shows that the metastable zone for a less stable form must lie in some regions of the graph outside that of a more stable form. The attempt to draw the metastable zones for the metastable forms entirely within the metastable zones of the stable forms inevitably leads to a physically unrealistic kink in the metastable zone lines, illustrated in Figure 3. The existence of a potential equilibrium transition to another polymorph cannot have any effect on the smooth variation with temperature of the properties of the polymorph in question, such as the metastable zone width. Indeed if it did so, it would be an absolutely reliable indicator of the existence, or otherwise, of as-yet undiscovered polymorphs. Apart from common experience that the metastable zone width does not become negligible in the area

of a transition point, literature traces confirm that this is not the case. For example, Beckmann⁸ draws a metastable curve for polymorph B of Abecarnil which shows a typical metastable zone width but which must deny Ostwald's belief both for Polymorph A and for Polymorph B above and below 80 °C, respectively. The universality of the “least stable first” principle would produce a curiosity in the case of phenylacetic acid, for example, in that when precipitated above 76 °C, the solid would be first formed and then melt, whereas when precipitated below 76° the liquid would first form, then solidify. It is also incompatible with the concept of nucleation in which nuclei are regarded as small crystals which have only just reached the point of thermodynamic stability, although it is compatible with Gavezzotti and Filippini's proposition of nucleation as a critical mass effect.⁹ The discussion has centered here on the formation of solid from liquid, but it applies equally to solid–solid transformation. Until Weygand's work on the chalcones,¹⁰ there was no system known of a higher complexity than trimorphic. Consequently there would not have been much opportunity

(8) Beckmann, W.; Nickisch, K.; Budde, U. *Org. Process Res. Dev.* **1999**, *2*, 296.

(9) Gavezzotti, A.; Filippini, G. *Chem. Commun.* **1998**, 28.

(10) Weygand, C.; Baumgaertel, H. *Liebigs Ann.* **1929**, 469, 225.

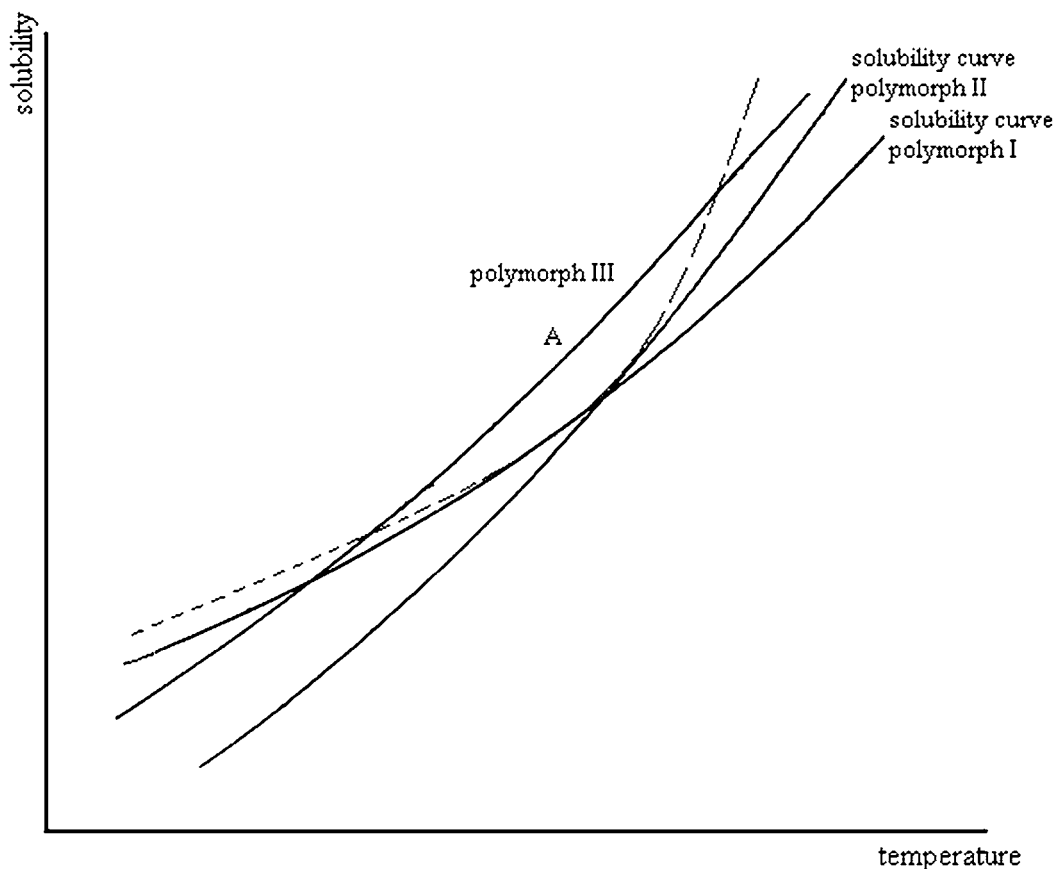


Figure 4. Trimorphic systems are impossible within the restraints of Ostwald's thermodynamic analogy. Both the solubility curve and the metastable limit would need to lie totally within the region abcde, necessitating all solid–solid transition temperatures to coincide exactly.

for Ostwald to have investigated the sequences of solid–solid transformations. There are now many papers in which one or more forms are reported as only obtainable by a specific operation such as desolvation of a certain solvate.¹¹ This shows that there is no one-to-one mapping of instability to first appearance of a polymorph. Furthermore, the existence of the state of affairs contemplated by Ostwald would make a trimorphic system impossible as shown in Figure 4. Point A of the third polymorph would have to coincide with the transition between polymorphs I and II. Indeed, all transition points of a polymorphic compound would have to be identical, which is contrary to experience.

It must be concluded that the analogy relied upon by Ostwald between gas–liquid and solution–solid behaviour is not totally valid, although it is certain that there must be an underlying thermodynamic reason for the ready formation of the less stable forms.

An Alternative Thermodynamic Approach

Ostwald's analysis shows why the formation of metastable polymorphs might occur, but perhaps less so why they do, and certainly cannot show such formation to be a universal law. An alternative thermodynamic approach, which was not available to Ostwald, because it had not then been invented, is that of irreversible thermodynamics. I am grateful to M.

Descamps of the University of Lille, whose lecture at the 6th PhANTa conference in Ascona¹² first made me aware of the importance of irreversible thermodynamics in relation to phase transitions. M. Descamps has conducted experiments which provide the first convincing explanation of a widespread but little known phenomenon, namely that it is possible to convert a stable polymorph to a less stable one by grinding. There is a smattering of examples in the literature^{13–15} of the conversion of stable polymorphs to other crystal forms, although amorphization by grinding is better known.¹⁶ It is clear that many practitioners have encountered sporadic examples but considered them to be anomalous observations and not publishable.¹⁷ What M. Descamp's experiments have made clear is that it is a general occurrence, determined by the competition between two opposed driving forces. The minimisation of ΔG is the classical thermodynamic driver, always quoted, leading to the stable form,

(11) Bergren, M. S.; Chao, R. S.; Meulman, P. A.; Sarver, R. W.; Lyster, M. A.; Havens, J. L.; Hawly, M. *J. Pharm. Sci.* **1996**, *88*, 834.

(12) 6th International conference on Pharmacy and Applied Physical Chemistry, Ascona, Switzerland, May 26–30, 2002.

(13) Beckmann, W.; Otto, W.; Budde, U. *Org. Process Res. Dev.* **2001**, *5*, 387.

(14) Shachtshneider, T. P.; Boldryev, V. V. *Drug Dev. Int. Pharm.* **1993**, *19*, 2055.

(15) Otsuka, M.; Otsuka, K.; Nobuyoshi, K. *Drug Dev. Int. Pharm.* **1994**, *20*, 1649.

(16) Crowley, K. J.; Zografi, G. *J. Pharm. Sci.* **2002**, *91*, 492.

(17) There do not appear to be any suggestions in the literature. A discussion at the conference, *100 Years of Ostwald's Rule*, UMIST, Manchester, UK, November 10–11, 1999, produced several ideas including effects of heating, intermediate amorphization, and pressure but no consensus as to the extent and true reason for the examples which had been encountered.

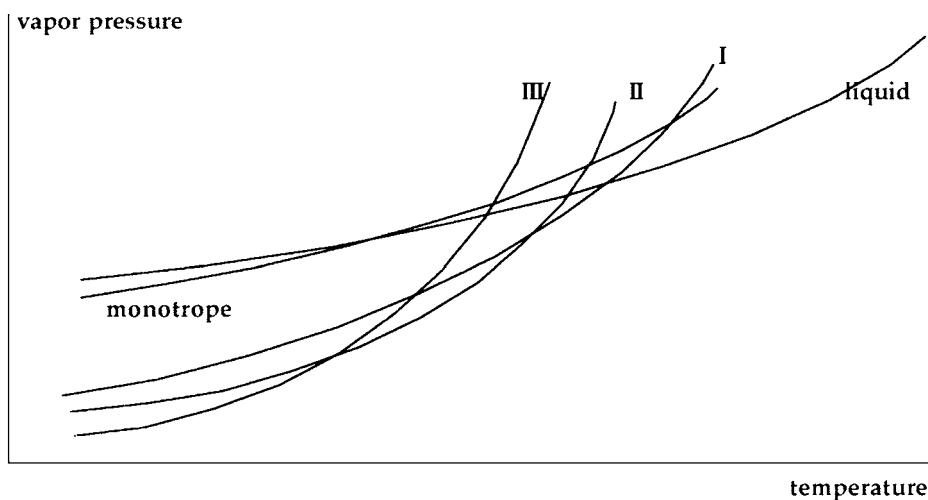


Figure 5. Phase diagram for a polymorphic system with several enantiomorphically related forms and a monomorphic form. The monomorphic form would be the stable form at high temperature if it did not melt before reaching the transition point. It is the least stable of the forms at room temperature.

whilst the maximization of the rate of entropy production is the driver in irreversible thermodynamics.¹⁸ This will lead to the less stable form. The equilibrium composition of the polymorphic mixture will depend on the rate at which excess energy is applied to the system. That which is applicable to recrystallisation must equally apply to initial crystallisation. In the grinding case the source of the excess energy is that of the impact forces. The thermodynamics then reflect the need to redistribute this efficiently. Insofar as the excess energy is not externalised as heat, it is most efficiently dissipated by conversion into free energy in the product, i.e. by generating the least stable product. The greater the impact forces, the larger the proportion of metastable polymorph formed. In the crystallisation case, the driving force is probably the need to remove the heat of crystallisation at the point of nucleation or crystallisation. Dependent on the enthalpy of crystallisation, the concentration of the solution, and the rapidity of nucleation or crystal growth, unstable forms may or may not be favored.

Structural Basis of Ostwald's Rule

As well as having a thermodynamic basis, polymorph formation has a structural basis, as the preferential routes of desolvation mentioned earlier illustrate. The polymorph which will crystallise preferentially from a melt or solution will be the one easiest to form (i.e. the one with the smallest energy barrier expressed in kinetic/thermodynamic terms) or the one whose structural organisation is most readily derived from the arrangement in the melt or solution. There is insufficient present knowledge of the structure of liquids or solutions (especially those on the point of nucleation) or of intermediates to be able to address the problem in general in these terms. However, it might be supposed that structures which are not easy to relate in terms of packing and bonding would be in general less easily converted than closely related structures would be. The few available comparative spectra of a series of polymorphs and melts, for example as shown

in Figure 1 of ref 19 suggest that the highest-melting form shows the greatest spectral similarity to the melt. Portions of the spectra of this diagram, on this occasion recorded as diamond anvil ATR spectra, are presented in Figure 6. From this one deduces that Polymorph I probably has the greatest structural similarity to that of the melt, borne out by detailed interpretation of the sulfathiazole polymorph and solvate spectra.²⁰ Some caution is needed in interpreting the infrared only in superficial terms. Instrumental factors and thermal disequilibrium both contribute to band broadening at elevated temperatures, which will mimic the broadening seen in melts as a result of the multiplicity of potential interactions in such nonperiodic systems. Band shifts of intermolecular interactions will be more significant, but even here care must be taken to recognise that what may be being observed is a general lengthening of the molecule–molecule distances and their associated hydrogen bonds rather than specific structural features. NMR spectra are more easily interpreted in terms of detailed structure. The number of examples of solid-state NMR spectra of polymorphs in the literature is rapidly growing, but rarely are associated melts examined, and in only a few cases is any comparison made of spectra of the polymorphs in relation to the structure of solutions or melts.²¹

In a series of enantiotropically related polymorphs such as sulfathiazole, the most stable polymorph at the highest temperature will be the least stable at the lowest temperatures. In the case of such enantiotropically related series, the structure closest to that of the melt will be the least stable structure at room temperature, and thus conformity with Ostwald's rule is facilitated. For the case of monotropically related polymorphs, the case may not appear at first glance to be so clear-cut. However, examination of phase diagrams of monotropically related forms, such as that shown in Figure 5, will reveal that a monotrope behaves as a high-temperature form, whose potential stability range has been cut off by

(19) Threlfall, T. L. *Analyst* **1995**, *195*, 2435.

(20) Hursthouse, M. B.; Threlfall, T.L.; Ward, S. J. Unpublished observations

(21) Stephenson, G. A.; Groleau, E. G.; Kleeman, R. L.; Xu, W.; Rigsbee, D. *R. J. Pharm. Sci.* **1998**, *87*, 536.

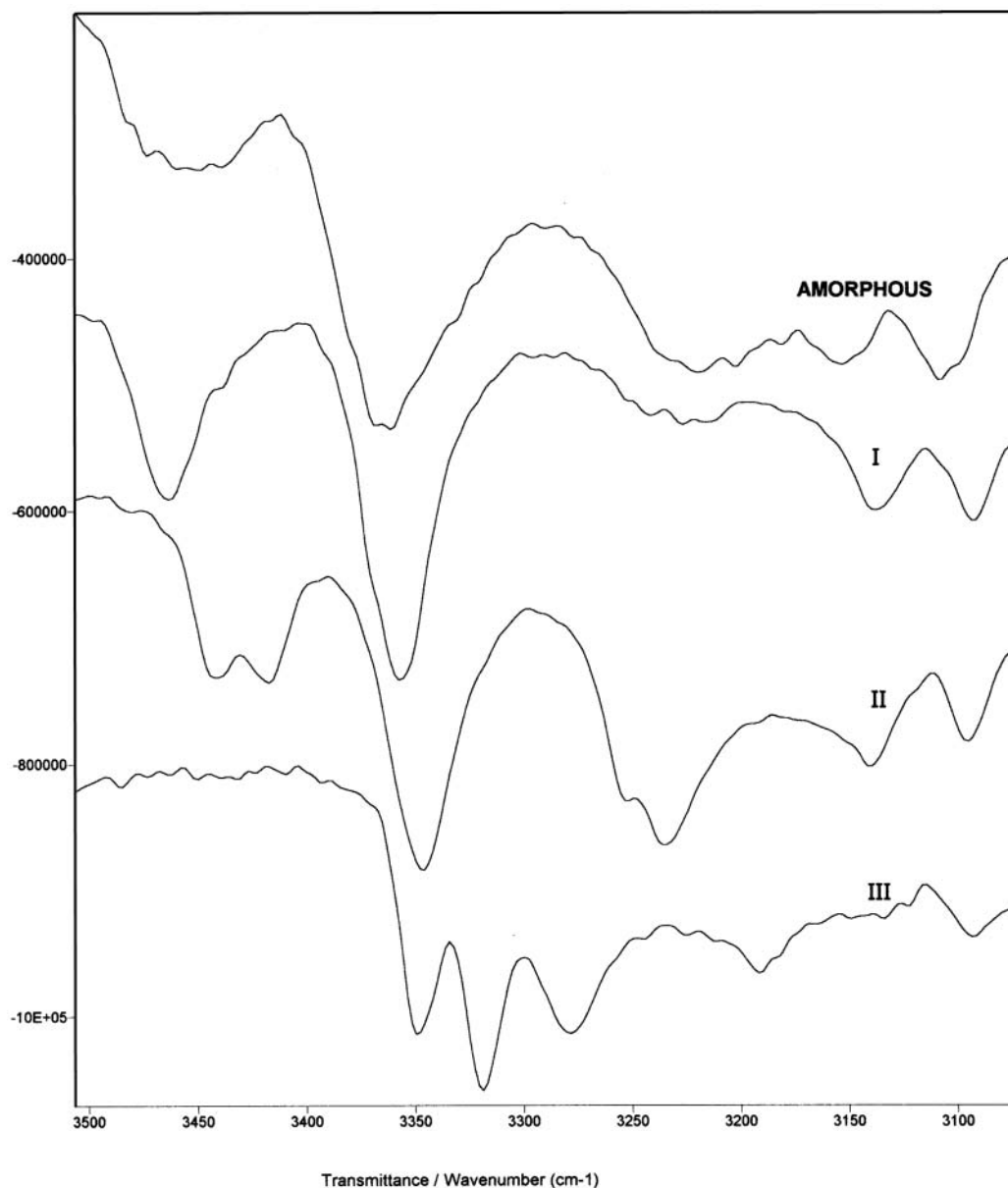


Figure 6. Infrared spectra of sulfathiazole polymorphs I (mp 203 °C), II (mp 197 °C), III (mp 175 °C), and the amorphous form. Polymorphs III, IV, and V (and III' and IV', unpublished) share very similar spectra, melting points, and structures, being members of a polytypic series. The numbering of the polymorphs is in accord with pharmaceutical usage. Polymorph II is Polymorph V of the Cambridge Crystallographic Data Base. The melt has a similar spectrum to that of the amorphous form except for the expected shifts due to the temperature difference.¹⁹ It can be seen that the spectrum of Polymorph I is closer to that of the amorphous form than to those of the other polymorphs. This is because the structures making up the lower-melting polymorphs are no longer present in the melt and amorphous form. In particular, the peculiar dimers forming the layers of the polytypic polymorphs and giving rise to the multiple bands between 3350 and 3270 cm^{-1} must be absent. Similarly the tetramers characteristic of Polymorph II³⁵ and giving rise to the doublet at 3440/3420 cm^{-1} must also be absent. The amorphous form is distinguished from Polymorph I, first because it is an isotropic glass and second by its strong band at 1518 cm^{-1} , a wavenumber at which none of the known sulphathiazole polymorphs absorb. Sulfathiazole is strikingly polyamorphous, and isotropic glasses with strong spectral resemblances to the other polymorphs can be produced, but these are metastable at elevated temperatures with respect to the form shown above.

the interposing of the liquid state. It is likely, therefore, to be even more akin to the liquid than any enantiotope but, of course, even less stable at room temperature.

A much clearer example of the production of the most structurally similar polymorph from the melt is that of sulfapyridine. When sulfapyridine melts are quenched to an amorphous form and subsequently gently reheated, they first crystallise around 100 °C, as shown by Kuhnert-Brandstaetter,²² to a form which subsequently changes on further

heating through a whole series of transformations. The DSC of quenched sulfapyridine melts is much illustrated in the literature of manufacturers of thermal analysis literature because of its variety. The spectral and other characteristics of this initial polymorph were recorded by Burger²³ as sulfapyridine polymorph III. This can also be obtained by pouring molten sulfapyridine at 200 °C under nitrogen into

(22) Kuhnert-Brandstaetter, M.; Wunsch, S. *Mikrochim. Acta* **1967**, 6, 1297.

(23) Burger, A.; Schulte, K.; Ramberger, R. *J. Therm. Anal.* **1980**, 19, 475.

boiling toluene, a procedure best carried out on a small scale for safety reasons. We have recently determined the crystal structure of this form.²⁴ The structural unit of this crystal consists of one molecule of the amido tautomer plus one molecule of the imido tautomer. The balance between amido and imido in the antibacterial sulfonamides is a delicate one, dependent mainly on the basicity of the heterocyclic ring. Thus, until our recent reinvestigations of the sulfadrag polymorphs,²⁴ all sulfathiazole, methyl sulfathiazole, and sulfapyridine polymorphs had been shown to possess the imido structures, whilst others had exclusively the amino structure. Amongst some 300 unpublished crystal structures which have been acquired at the University of Southampton of related sulfonamides in pursuit of structural systematics of polymorphs, only a handful of structures have been seen to possess a mixed amido–imido structure. Included amongst these are 5-chlorosulfapyridine and 5-methylsulfapyridine. However, when the exquisite restraints ensuring molecular uniformity in the crystal structure are removed, amino–imido mixtures are more readily formed. It was shown long ago by ultraviolet spectroscopic comparison of solutions of sulfapyridine and sulfathiazole with their methylated derivatives that these solutions consist of mixtures of amido and imido, the proportions of which are solvent dependent.²⁵ Melts of sulfapyridine and of sulfathiazole are similarly mixtures of amido and imido tautomers.

Thus, when molten sulfapyridine is quenched, it is unsurprising that the polymorph first formed is the one containing both imido and amido tautomers, in preference to one of the forms containing only imido tautomers.

Support for the concept of the structure of the melts being closer to that of the metastable forms than to the form stable at room temperature or below also comes from the melt crystallisation process itself. It is at first sight a curiosity that there is a general belief supported by example that more polymorphs are produced from the melt than from solution, despite the greater number of variable parameters available in the latter case. In addition the viscosity of the quenched melt compared with that of the solution would seem to restrict the formation of less easily (entropically) formed polymorphs. Examples of the formation of manifold polymorphs from the melt are the chalcones investigated by Weygand²⁶ by melt recrystallisation and benzophenone by hot-stage microscopy, summarised by Deffet.²⁷ Solution recrystallisation fails to produce as many polymorphs in either case. Further examples are given by Kuhnert-Brandstaetter.²⁸ The most probable explanation is that there are already clusters of the appropriate orientation and conformation in the melt capable of nucleating metastable forms which are then preserved by the low mobility of the molecules in the cooled melt.²⁹ By contrast, solvation in solution might

interfere with the ready formation of compact clusters and interfere with the putative metastable polymorph formation by transformation of early-formed nuclei to the stabler form. (A corollary of this discussion would be the implication that concentrated solutions are more likely to produce a variety of polymorphs than dilute solutions, because of the increased concentration of structural clusters.)

All of the above examples are cases in which a metastable or high-temperature form or most-easily crystallised form appears to have a structure which most closely resembles that of the melt or solution. The small extent to which such relationships have been addressed ensures that only a minute proportion of polymorph systems has been so investigated. Even a concentrated research programme would not generate a substantial proportion in relation to the number of polymorphic systems. The question must then arise as to whether these examples are representative of the normal behavior or are merely isolated exceptions. There is fortunately a theoretical basis for supposing that this may represent a general relationship, discussed in the immediately following paragraphs.

Relating Structural Drivers of Polymorphic Change to Thermodynamic Drivers. The thermodynamic factors and the structural factors are of course not completely independent phenomena, independent consideration of which can lead to contrary conclusions in respect to phase transitions.

At absolute zero, the most stable crystal form of a non-hydrogen-bonded solid will be the one with the highest density, because this maximizes the van der Waals interactions.

As the temperature rises, the increasing thermal vibrations initially produce lattice expansion, but when this becomes insufficient to absorb the vibrational requirements, that form becomes destabilised, and transition to a less dense (more open) form occurs. The matter may not appear to be so clear-cut in hydrogen-bonded solids, where there is competition between the dense packing that relies on many, weak, nondirectional, short-range van der Waals forces and the more open packing that is partly dependent (the van der Waals forces are still present) on a few, strong, directional, long-range hydrogen bonds to achieve overall energy minimisation. However, our work on anisotropic lattice expansion²⁴ has shown that the strongest hydrogen bonds tend to be those that are first affected by increase of temperature. This result at first appears surprising: it might have been thought that the weakest hydrogen bonds would be most easily altered and therefore most liable to alter. The strongest hydrogen bonds are, however, the shortest so that a greater relief of vibrational de-energisation is effected in the strongest bonds. Furthermore, since bond stretching is known from infrared spectra to be about twice the frequency of bond bending³⁰ that is in turn at much higher frequency than bond torsion, bond lengthening will take place preferentially via

(24) Hughes, D. S.; Hursthouse, M. B.; Threlfall, T.; Tavener, S. *Acta Crystallogr.* **1999**, C55, 1831.

(25) Shepherd, R. G.; Bratton, A. C.; Blanchard, K. C. *J. Am. Chem. Soc.* **1942**, 64, 2532.

(26) Weygand, C.; Mensdorf, L. *Ber. Dtsch. Chem. Ges.* **1935**, 68, 1825.

(27) Deffet, L. *Repertoire des Composés organiques polymorphes*; Desoer: Brussels 1942.

(28) Kuhnert-Brandstaetter, M. *Thermomicroscopy in the Analysis of Pharmaceuticals*; Pergamon: Oxford, 1952.

(29) A reviewer has suggested that the reason more forms may be obtainable from melts is that the induction time for all forms is much shorter from the melt than from solution.

(30) Colthup, N. B.; Daly, L. H.; Wibberley, S. E. *Introduction to Infrared and Raman Spectroscopy*, 3rd ed.; Academic: New York, 1990.

torsion or possibly bending of bonds, rather than directly by bond stretching. The lowest-frequency hydrogen bond stretchings are associated with the strongest hydrogen bonds but with the highest frequencies of other X–H modes.³⁰ However, this in turn implies that the still weaker H...acceptor modes will have the lowest deformation and torsion frequencies for the strongest hydrogen bonds. Thus, it is possible that the shortest hydrogen bonds are indeed the easiest to unravel. Even in hydrogen-bonded solids the same principles are then at work to drive phase transitions, namely the need to eliminate at higher temperatures those intermolecular interactions which interfere most with the intramolecular vibrational modes. The less-dense forms and the forms with the weaker hydrogen bonds are those with the lowest-frequency lattice vibrations and torsional modes. Those will therefore be the forms with the lowest stability at low temperatures and the shallowest slopes in the usual phase representations such as Figures 2 and 3. The melt is the most open of the condensed phases and the least stable one at low temperature. In particular, its molecules will be juxtaposed in arrangements in which those intermolecular interactions, which were so beneficial at low temperature, (close hump-to-hollow matching and short hydrogen bonds) have been eliminated. Consequently, the geometry of dimers and clusters in the melt is unlikely to mimic those strong interactions which lead to the ordering of molecules characteristic of the low-temperature forms.

The same result is arrived at via statistical thermodynamics. The slope of the lines in Figures 2 and 3 is determined by the heat capacity, ΔC_p which is itself determined by the equation³¹

$$C_p \approx k/n \sum_{\nu} (h\nu/kT)^2 \exp(h\nu/kt) / [\exp(h\nu/kt) - 1]^2$$

where n = number of moles in the sample, h = Planck's constant, k = Boltzmann's constant, ν = frequency.

This equation, because of the squared exponential on the bottom line, implies a reciprocal relationship between vibrational frequencies and heat capacity. [It is stated³¹ that this is ascertainable by differentiation. In my hands the result was not obvious. I am grateful to Dr. R. Moss of the Chemistry Department of the University of Southampton for a numerical differentiation using the mathematical program MAPLE, which indeed confirmed Burger and Ramberger's conclusion.] Thus, it is the lowest-frequency modes in the infrared spectrum, especially the lattice vibrations related to the intermolecular distances and therefore to the van der Waals interactions, and the low-frequency torsions such as the hydrogen-bond twisting modes, which drive C_p . To relate the mathematical and the structural models of lattice expansion and polymorphic transition discussed here, C_p may be regarded as the energy required to overcome the molecular friction. As concluded earlier, it is therefore the densest forms which ultimately become least stable at higher temperatures, culminating in the loss of the closest packing options, and thus, ultimately, it is the low-frequency vibrations associated

with the intermolecular interactions which bring about the possibility of transition.

Psychological Aspects of Ostwald's Rule

Ostwald nowhere writes of the possibility of subsequent crystallisation giving rise to a different result. Indeed, if the analysis of his beliefs detailed earlier is correct, it would be contrary to his thesis. He is concerned only with the order of appearance of polymorphs in a single experiment: the formation of the least stable form followed by the appearance of more stable forms as a result of transition. The extension of Ostwald's rule to the possibility of subsequent experiments giving rise to a stable form with consequent exclusion of the original form is a modern concept. A scientist encountering crystallisation in a polymorphic system for the first time would suppose, as did Ostwald's contemporaries and predecessors, not unreasonably, that the most stable form would also be the easiest to produce and therefore most likely to be the first generated. Subsequent formation of the less stable form as a result of deliberate experimentation would not be commented upon. However, the opposite case would be considered strange at first encounter and an explanation sought. The consequence might be a subsequent mention in the literature of the crystallisation obeying Ostwald's rule. Even the more experienced practitioners seem to comment on Ostwald's rule when it is followed but remain silent or express surprise when it is not followed. Presumably, for this reason, Ostwald's rule seems to be overemphasized in relation to the observed behavior of crystallisations.

Increasing complexity of the new molecular entities of the pharmaceutical industry, coupled with increasing attention to the discovery of polymorphs, has led to the unveiling of systems of up to 17 polymorphs.³² From a purely statistical view the chances against encountering the stable polymorph of such a system at first crystallisation is minimal, although there are, of course, other factors which make the precise numerical application of such an approach indefensible. There are also many other circumstances in which the rule is immaterial. For example, there are substances which crystallise exclusively and reliably in one form above a given temperature, but in another form below that temperature. There are thousands of examples in the literature where it is claimed that one polymorph crystallises from one solvent but a different one from a second solvent. Although it might have been more accurate in many cases to report that the observed result was from the one crystallisation attempted from that particular solvent, there are undoubtedly many well-investigated examples which are reliably and extremely solvent dependent. In such cases the order of appearance of polymorphs is a matter of chance as to which solvent was first tried. The formation of one hydrate as opposed to another is often dependent on the water concentration of the crystallising solvent. Rapid crystallisation and harvesting from a concentrated solution may give rise, especially during chemical development, to a metastable form as opposed to a stable form from long crystallisation and drying, a consequence of subsequent transition. It is true that this is

(31) Burger, A.; Ramberger, R. *Mikrochim. Acta* **1979**, *II*, 259.

(32) Pesti, J. A.; Chorvat, R. A.; Huhn, G. F. *Chem. Innovation* **2002**, *Oct.*, 28.

an example of Ostwald's rule, but in the general case the subsequent discovery of a stable form on a bench scale may be only a reflection of the particular order in which a set of crystallising conditions were chosen. Exceptions to Ostwald's rule occur also in crystallisations from the melt and from vapor, as has been emphasised.⁷

It is extraordinarily difficult to eliminate beyond doubt the possibility of chance seeding¹ whether from retained nuclei in solution or in the melt, as in the case of benzophenone,⁷ from extraneous seeding from the nuclei atmosphere or within vessels, from sensitivity of the crystallising phase to dust or other foreign substances, including the materials of construction of the crystallisers and impellers, perhaps even to vibration or cosmic rays. For these reasons the discussion in this contribution has concentrated on underpinning the experimental evidence of non-Ostwald rule behavior by a theoretical analysis. It must be concluded that whether polymorphic systems follow Ostwald's rule is a matter of both chance and the competing forces in the crystallisation. Whilst there are many examples of the preferential formation of a metastable polymorph, there are also many examples of contrary behavior.⁷

Comments on the Application of Ostwald's Rule in Some Practical Situations

The deposition of a liquid rather than crystals from solution (oiling) can be a serious problem particularly in the manufacture of low-melting solids. The comments in the preliminary discussion of Ostwald's rule suggest that he regarded oiling out as a manifestation of the rule. The mechanistic explanations of this phenomenon would include a kinetic one and one involving solvent incorporation. The latter is likely to apply in the case of anti-solvent addition or salting out where the solution breaks into two phases: solidification follows the removal of the solvent from the oil. The more general explanation must be that the dissolved phase separates before the molecules have time to assemble into a lattice. This reasoning lends credence to Gavezotti and Filipino's model of nucleation.⁹ If it could be shown generally that the more stable the polymorph, the slower the kinetics of assembly, then this would provide a total explanation of Ostwald's rule. I think such a relationship unlikely. The reason the problem of oiling is more acute with lower-melting products may simply relate to the closeness of the precipitation temperature to the melting point. Another reason may lie in the mechanism of prenucleation organisation within solutions. Higher-melting solids are more likely to contain hydrogen-bonded networks. Hydrogen-bond propensities will surely result in the formation of hydrogen-bonded clusters in solution. It is therefore possible that solutions of hydrogen-bond solids have a predisposition to immediate crystallisation which those of some van der Waals solids may lack.

It is becoming increasingly accepted by process development chemists that the appearance of a new polymorph during chemical development is linked to the removal of key impurities, and sometimes to the presence of a new impurity. The mechanistic details have been insufficiently investigated, or perhaps insufficiently reported in industrial cases to allow

the development of a general understanding. However, there are laboratory-scale investigations, particularly by Davey's group,³³ which enable one to believe that the formation of a particular polymorph can be dramatically slowed even by the incorporation of single-impurity molecules. This is incompatible with a universal law describing the order of appearance of polymorphs.

Ostwald's rule is often quoted in relation to disappearing polymorphs.⁷ This phenomenon is only a reflection of the range of possibilities for the polymorphic relationships. If the new polymorph is a metastable form, then the original form is unlikely to disappear. There is always a range of conditions under which it is stable and under which, in principle, it can be produced by conversion, even if its area of polymorph formation space appears to have vanished. If the new polymorph is a stable form and the relationship is an enantiotropic one, then there will be conditions under which the original form will be the thermodynamically stable one. Even if the relationship is one of monotropy, there may still be convenient or at least readily attainable conditions under which formation of the metastable polymorph dominates. In the latter cases, whether there will be a disappearing polymorph will depend on the barrier to conversion. As stated earlier it is a state of affairs not countenanced by Ostwald. It is possible that Ostwald was correct in his view, that the different outcome is a result of the use of a crystallising solution, which is different, perhaps only marginally different, from the original. Given the subtlety of crystallisation and our limited understanding of it, we should not be surprised if it is not always possible to track down the source of the difference.

At early stages of development, solutions which deposit mixtures of polymorphs ("concomitant polymorphs"), may be encountered. The hope would be to avoid such a situation later. Any solutions yielding stable mixtures i.e., where interconversion is slow, probably represent exceptions to Ostwald's rule, as the polymorphs are unlikely to have formed sequentially or by transformation. One sees in this case why purity may have such a decisive effect on polymorphic outcome: even trace impurity could alter dramatically the balance of relative rates of growth of the competing polymorphs. Bernstein⁷ discusses several examples of concomitant polymorphs in relation to Ostwald's rule. Many of these involve conformational polymorphism, although there does not seem to be any reason conformational polymorphs should show distinct behavior. There will be an entropic contribution, and the kinetics of crystallisation may thereby be slowed, but this will be only one of many effects leading to competing kinetics between polymorphs.

The recent report of Peterson et al.³⁴ gives considerable support in a negative way to the propositions advanced in this contribution as explanations of the behavior described by Ostwald's rule. It is commented there that the dehydration of acetaminophen trihydrate gives the stable polymorph

(33) Blagden, N.; Davey, R.; Rowe, R.; Roberts, R. *Int. J. Pharm.* **1998**, *172*, 167.

(34) Peterson, M. L.; McIlroy, D.; Shaw, P.; Mustonen, J. P.; Oliviera, M.; Almarsson, O. *Cryst. Growth Des.* **2003**, *3*, 761.

(35) Blagden, N.; Davey, R. *J. Cryst. Growth Des.* **2003**. In press.

rather than the metastable one predicted by Ostwald's rule. There are two reasons why Ostwald's rule may not be relevant in this case, related respectively to the thermodynamic and structural explanations presented earlier. First, the dehydration is likely to occur under near-equilibrium conditions—even if the relative humidity is such as to imply a divergence from equilibrium, the slow escape of the water molecules from the crystal maze will mimic equilibrium. Thus, classical thermodynamics determines the production of the stable form. Second, there is no reason the structure of the hydrate should resemble the metastable form rather than the stable form. The expectation would generally be that, in the case of dehydration, all anhydrates would be statistically equally favored.

Conclusions

The favored formation of a less stable form from a crystallisation process can be explained either on a structural basis or on the basis of irreversible thermodynamics. Although the early formation of metastable forms often occurs, there are many exceptions. The exceptions as well as the rule can be understood within the theoretical framework discussed. Thus, for both experimental and theoretical reasons, the term "Ostwald's law" is a less desirable one than "Ostwald's rule".

Received for review June 12, 2003.

OP030026L