Tautomeric polymorphism in omeprazole†

Prashant M. Bhatt and Gautam R. Desiraju*

Received (in Cambridge, UK) 12th January 2007, Accepted 15th March 2007
First published as an Advance Article on the web 11th April 2007
DOI: 10.1039/b700506g

Crystalline omeprazole exists as solid solutions of two tautomers in a continuous composition range, and this raises questions pertaining to the definition of the term *polymorph*.

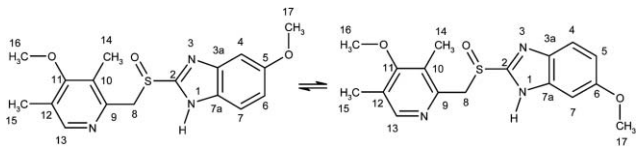
Omeprazole, 5(6)-methoxy-2-{(4-methoxy-3,5-dimethyl-2-pyridinyl) methylsulfinyl}-1*H*-benzimidazole **1** (**2**) is a blockbuster anti-ulcer drug. In recent times, polymorphism of drugs has emerged as a major topic of research because crystal forms with novel and interesting properties may qualify for independent patent protection, effectively extending the marketable life of the active pharmaceutical ingredient (API).[1,2] While there have been several definitions of the term *polymorph* in organic solid state chemistry and crystal engineering, the gist of these definitions is that they involve different arrangements of the same molecule in its solid forms. As long as the molecules concerned are rigid and there are no great ambiguities in their crystal structures, the meaning of the term polymorph is uncomplicated. With small molecule X-ray crystallography becoming a high throughput activity, the number and complexity of crystal structures now available has greatly increased. This leads to a reconsideration of the definition of the term polymorph. Notably, how similar should the same molecules be and how dissimilar should the different crystal structures be in order for them to qualify as polymorphs?[3] The present communication deals with the former question in the context of a little known variety of polymorphism, namely *tautomeric polymorphism*.

Taking the idea of molecular sameness further, it has been argued that if a pair of tautomers are in rapid equilibrium in solution or in the melt, the crystals formed by each of them are polymorphic.[4] In general, crystals of isomers that interconvert rapidly in solution would be classified as polymorphs, while those of slowly interconverting isomers would be classed as different compounds. There is an element of subjectivity in this definition because the rates of interconversion are generally temperature dependent. Accordingly, depending on the temperature of the experiment, a pair of crystal structures could be called polymorphs or different compounds! Still, given the fact that tautomerism is widely prevalent in solution, one might have expected there to be many well documented examples of tautomeric polymorphism.

In reality, this is not the case. There are, at best, two unambiguous reported instances of this phenomenon. The first report dates back to 1983 and is concerned with 2-amino-3-hydroxy-6-phenylazopyridine, which exists as hydroxyazo and quinonehydrazone crystals with different colours.[5] The second case has been reported in recent years and deals with sulfasalazine, which exists as an amide and an imide, and also as a hydrated imide and a DMF-solvated imide.[6,7] This phenomenon has also been referred to as *desmotropy*, with reference to 3-phenyl-1*H*-pyrazole and 5-phenyl-1*H*-pyrazole,[8] and to irbesartan, a tetrazole-containing pharmaceutical compound.[9] Other examples are more equivocal: in anthranilic acid, one polymorph contains neutral and zwitterionic molecules in the asymmetric unit, while the other contains only neutral molecules.[10] The dipeptide L-His–Gly crystallises as a hemihydrate, with both the more favourable Nε–H and the less favourable Nδ–H tautomers in the same crystal.[11] Similar situations, wherein two tautomers are present in the same crystal, prevail in *N*-(3-hydroxysalicylidene)-4-methoxyaniline,[12] 3(5)-phenyl-4-bromo-5(3)-methylpyrazole[13] and 4(5)-nitro-5(4)-methoxyimidazole.[14] Form **II** of ranitidine hydrochloride might also exist as a mixture of eneamine and nitronic acid tautomers.[15] In this communication, we present evidence that the crystal forms of omeprazole contain different tautomeric compositions, and that the phenomenon of tautomeric polymorphism in this system also leads to further questions regarding the definition of the term polymorph itself.

The tautomers of omeprazole are the 5- and 6-methoxy derivatives (**1** and **2**, respectively, Scheme 1), and these have been previously detected in solution.[16,17] Solid forms of omeprazole have been investigated using two approaches—with PXRD, and with single crystal XRD and Raman analysis—but there is little correlation between them. Three forms, **A**, **B** and **C**, have been patented and are characterized by their PXRD traces. In patent application WO 99/08500, it is stated that form **A** is more stable than form **B**.[18] In patent application US 2004/0122056, it is stated that form **C** is easier to prepare than **A** or **B**.[19] A solitary crystal structure of what appears to be the 6-methoxy tautomer[20,21] (mistakenly called the 5-methoxy compound in this paper, corrected subsequently by Claramunt *et al.*[16]) is identified in the WO 99/08500 patent application as form **B**. Another patent, US 6,780,880, claims that omeprazole crystals contain mixtures of **1** and **2**, and states that single crystal X-ray methods may be used to estimate the relative amounts of the two tautomers, without providing any further information.[22] In summary, the literature is confusing.

We obtained single crystals of five different forms of omeprazole, with the **1** : **2** ratio varying from 0 : 100 to 15 : 85.



Scheme 1 Tautomeric interconversion in omeprazole.

School of Chemistry, University of Hyderabad, Hyderabad 500 046, India. E-mail: gautam_desiraju@yahoo.com

† Electronic supplementary information (ESI) available: ORTEP structures, PXRD patterns and packing diagrams. See DOI: 10.1039/b700506g

Crystals of the pure 6-methoxy tautomer **2** (**I**) were obtained from a 2% methanolic solution of NaOH by slow evaporation over 2 d. A stereoview of the crystal structure is shown in Fig. 1.‡ This structure is the prototype for the rest of the group. Crystals containing increasing amounts of **1** were obtained by following procedures in the US 6,780,880 patent. All of these other crystals are essentially isostructural to **I**, and the diffraction data were modelled by refinement of the site occupancy factors of the MeO-group between the 5- and 6- positions of the benzimidazole ring (Fig. 2). It may be noted, however, that the proportions of **1** in the crystals obtained were different from those reported in the literature. Crystals **II–V** were obtained as follows. **II**: 8% **1**, 92% **2**, from concentrated ammoniacal MeOH over 3 d at r.t.; **III**: 10% **1**, 90% **2**, from dilute ammoniacal MeOH over 3 d at r.t.; **IV**: 12% **1**, 88% **2**, from acetone or 70 : 30 MeOH–CCl_4 at 5 °C; **V**: 15% **1**, 85% **2**, from $CHCl_3$ over 2 d at r.t.

There is adequate evidence that **2** is more stable than **1** and correspondingly that crystals containing greater proportions of **2** are more stable than those containing less.[22] According to the literature, **2** is photostable while **1** is photosensitive. Crystals of **I** are white and do not change colour. We observed that as the proportion of **1** increases (**II** → **V**) the crystals darken with increasing ease upon standing. It is also possible that, at the crystal level, **2** packs slightly better than **1**. What is interesting is that the crystal packing is such that the MeO group may be situated either at the 5- or 6- position of the benzimidazole ring without affecting the gross mutual disposition of molecules. Accordingly, crystals **I–V** may be viewed as substitutional solid solutions of **1** in **2**. Because they are crystals with differing amounts of tautomeric structures, the term *tautomeric polymorphism* is justified.

Simulation of the PXRD patterns of crystals **I–V** showed that **III** corresponds to Form **A** of the WO 99/08500 patent application, **IV** corresponds to Form **C** and **V** is form **B**. It may be noted that the reported crystal structure of omeprazole is innocent of the possibility that the crystal contains both **1** and **2**.[20,21] This structure was refined as if it contained only **2**, but the simulated PXRD matches the crystal with 15% of tautomer **1**. It is interesting that forms **A**, **B** and **C** seem to be sufficiently distinctive in terms of their stabilities and other properties, and at least distinctive enough

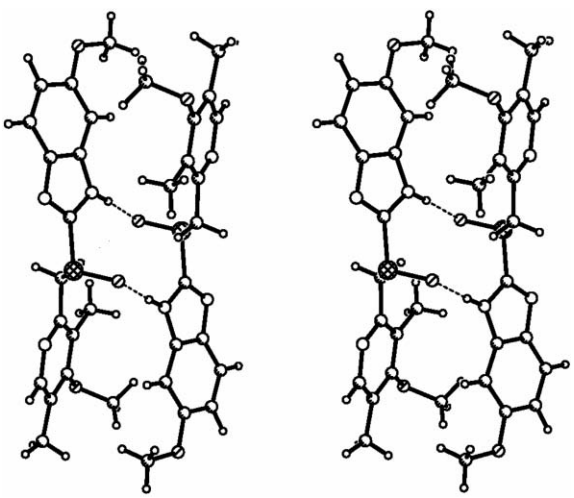


Fig. 1 Omeprazole as the 6-methoxy tautomer in the crystal structure of **I**. Note the N–H···O=S hydrogen bonds in the dimer.

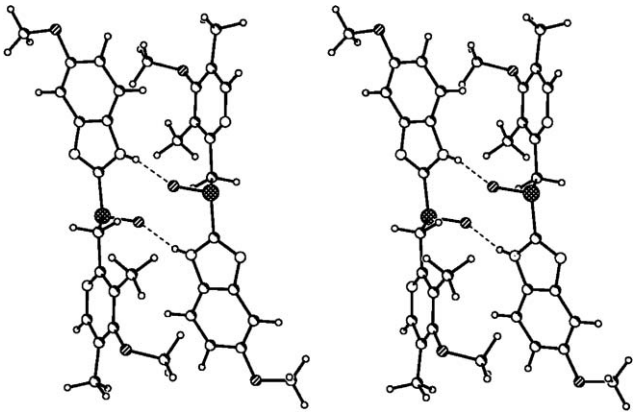


Fig. 2 Idealised view of the 5-methoxy tautomer of omeprazole in the crystal structure of **V**. The percentage of this tautomer is so low that only the methoxy group is "seen" as an entity distinct from the atoms of the 6-methoxy tautomer. Even then, refinement at variable occupancies provides inaccurate exocyclic angles for the 5-methoxy group. Accordingly, the coordinates used to generate this Figure were obtained by treating the 5-methoxy group as a rigid body with a fixed geometry with respect to the benzo ring.

that they enjoy independent patent protection. However, structurally speaking, they occur on a structural continuum that begins with a pure 6-OMe crystal and ends with a 85 : 15 mixture of 6- and 5-OMe tautomers.

This example highlights several interesting features of general significance. We have recently described structural modulation in crystalline aspirin, where the crystals are best described as intergrowths of two domains, each of which, if they existed independently, would constitute a pair of polymorphs.[23,24] Our description is at odds with the original report on this compound.[25] In omeprazole, the situation is different in that the modulation is at the molecular level, but there are still some points of comparison. Forms **I–V** contain different tautomeric compositions, therefore they are tautomeric polymorphs. But how many polymorphs of omeprazole really exist? Is it one or two or infinite? Would each **1** : **2** composition qualify for independent patent protection or would it be more meaningful to claim protection for compositional ranges? It is interesting to note that at present, the patented forms (**A**, **B** and **C**) are defined in terms of properties (stability, ease of preparation) rather than in terms of structure. The role of the PXRD trace is merely as a fingerprint of a form with a particular property rather than diagnostic of a particular structure type, because these PXRD traces (like the crystals they seek to characterise) constitute a structural continuum. The most significant aspect of this is that function seems to be a more meaningful criterion of polymorph patentability than structure. Indeed, this is clearly indicated in Buerger's definition of polymorphs as being different forms of the same chemical compound that have distinctive properties.[26] Therefore, if function is more significant than structure, this raises more provocative issues: (1) Should the definition of polymorphism rely so heavily on structural differences? (2) Are subtle structural differences really meaningful, especially in the context of the kind of modulation seen in omeprazole and aspirin? (3) Just as minor differences in crystal structure may be interpreted subjectively, may this also be said of molecular sameness? (4) Accordingly, how important is the

 This journal is © The Royal Society of Chemistry 2007

criterion that the molecular structure should be exactly the same if two crystals are to be called polymorphs? (5) In the context of polymorphism, is it more reasonable then to speak of a structural landscape that includes a number of solvated and unsolvated variations of the same molecular species, without insisting on a rigorous stoichiometric and chemical identity? In the end, one is reminded of Bernstein's realistic assessment that "an all-encompassing definition of polymorphism is elusive".[4] Perhaps such a definition is also unnecessary.

We thank the DST for financial support. P. M. B. thanks the CSIR for the award of a SRF.

Notes and references

‡ *Crystal data*: X-Ray data were collected on a Bruker SMART 4K-CCD area detector.

Omeprazole (**I**): $C_{17}H_{19}N_3O_3$, M = 345.42, triclinic, a = 9.6421(9), b = 10.3865 (10), c = 10.1539 (10) Å, α = 89.929(2), β = 110.939(2), γ = 116.937(2)°, V = 830.78(14) Å^3, T = 100(2) K, space group P$\bar{1}$, Z = 2, μ(Mo-K$_\alpha$) = 0.216 mm^{-1}, size 0.21 × 0.19 × 0.09 mm. 9197 total reflections, of which 3262 were independent, 2723 observed [$I > 2\sigma(I)$]. (R_{int} = 0.0337). Refinement against F^2 with 225 parameters, R_1 [$I > 2\sigma(I)$] = 0.0567, wR_2 [$I > 2\sigma(I)$] = 0.1209. CCDC 633382.

Omeprazole (**II**): $C_{17}H_{19}N_3O_3$, M = 345.42, triclinic, a = 9.6674(71), b = 10.3370(76), c = 10.2918(75) Å, α = 90.044(11), β = 111.552 (12), γ = 116.451 (12)°, V = 839.6(11) Å^3, T = 100(2) K, space group P$\bar{1}$, Z = 2, μ(Mo-K$_\alpha$) = 0.213 mm^{-1}, size 0.20 × 0.17 × 0.12 mm. 7586 total reflections, of which 3248 were independent, 1950 observed [$I > 2\sigma(I)$]. (R_{int} = 0.0563). Refinement against F^2 with 249 parameters, R_1 [$I > 2\sigma(I)$] = 0.0533, wR_2 [$I > 2\sigma(I)$] = 0.1196. CCDC 633383.

Omeprazole (**III**): $C_{17}H_{19}N_3O_3$, M = 345.42, triclinic, a = 9.6380(50), b = 10.2645(54), c = 10.3238(54) Å, α = 90.085(9), β = 111.732(8), γ = 116.288(8)°, V = 833.7(8) Å^3, T = 100(2) K, space group P$\bar{1}$, Z = 2, μ(Mo-K$_\alpha$) = 0.215 mm^{-1}, size 0.25 × 0.20 × 0.08 mm. 6800 total reflections, of which 3281 were independent, 1824 observed [$I > 2\sigma(I)$]. (R_{int} = 0.0563). Refinement against F^2 with 249 parameters, R_1 [$I > 2\sigma(I)$] = 0.0620, wR_2 [$I > 2\sigma(I)$] = 0.1449. CCDC 633384.

Omeprazole (**IV**): $C_{17}H_{19}N_3O_3$, M = 345.42, triclinic, a = 9.6439(16), b = 10.2621(17), c = 10.3322(17) Å, α = 90.216(2), β = 111.762(2), γ = 116.113(2)°, V = 835.5(2) Å^3, T = 100(2) K, space group P$\bar{1}$, Z = 2, μ(Mo-K$_\alpha$) = 0.214 mm^{-1}, size 0.23 × 0.17 × 0.07 mm. 8968 total reflections, of which 3280 were independent, 2917 observed [$I > 2\sigma(I)$]. (R_{int} = 0.0219). Refinement against F^2 with 249 parameters, R_1 [$I > 2\sigma(I)$] = 0.0504, wR_2 [$I > 2\sigma(I)$] = 0.1288. CCDC 633385.

Omeprazole (**V**): $C_{17}H_{19}N_3O_3$, M = 345.42, triclinic, a = 9.7014(26), b = 10.2585(28), c = 10.6942(29) Å, α = 91.720(4), β = 112.117(4), γ = 115.642(4)°, V = 864.8(4) Å^3, T = 298(2) K, space group P$\bar{1}$, Z = 2, μ(Mo-K$_\alpha$) = 0.207 mm^{-1}, size 0.26 × 0.21 × 0.06 mm. 8775 total reflections, of which 3373 were independent, 1736 observed [$I > 2\sigma(I)$]. (R_{int} = 0.0520). Refinement against F^2 with 249 parameters, R_1 [$I > 2\sigma(I)$] = 0.0611, wR_2 [$I > 2\sigma(I)$] = 0.1410. CCDC 633386.

In all of these cases, disorder was modelled with the constraints DELU and SIMU in the refinement (SHELX-97). For crystallographic data in CIF or other electronic format see DOI: 10.1039/b700506g

1 *Polymorphism in Pharmaceutical Solids, Drugs and the Pharmaceutical Sciences*, ed. H. G. Brittain, Marcel Dekker, New York, 1999, vol. **95**.
2 *Solid State Characterization of Pharmaceuticals*, ed. A. Zakrzewski and M. Zakrzewski, Assa International Inc., Danbury, Connecticut, USA, 2006.
3 T. L. Threlfall, *Analyst*, 1995, **120**, 2435–2460.
4 J. Bernstein, *Polymorphism in Molecular Crystals*, Clarendon, Oxford, 2002, pp. 2–4, 240–255 and 297–307.
5 G. R. Desiraju, *J. Chem. Soc., Perkin Trans. 2*, 1983, 1025–1030.
6 L. A. Filip, M. R. Caira, S. I. Fărcaş and M. T. Bojiţă, *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.*, 2001, **57**, 435–436.
7 A. J. Blake, X. Lin, M. Schröder, C. Wilson and R. X. Yuan, *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.*, 2004, **60**, o226–o228.
8 M. A. Garcia, C. Lopez, R. M. Claramunt, A. Kenz, M. Pierrot and J. Elguero, *Helv. Chim. Acta*, 2002, **85**, 2763–2776.
9 M. Bauer, R. K. Harris, R. C. Rao, D. C. Apperley and C. A. Rodger, *J. Chem. Soc., Perkin Trans. 2*, 1998, 475–481.
10 W. H. Ojala and M. C. Etter, *J. Am. Chem. Soc.*, 1992, **114**, 10288–10297.
11 T. Steiner and G. Koellner, *Chem. Commun.*, 1997, 1207–1208.
12 H. Pizzala, M. Carles, W. E. E. Stone and A. Thevand, *J. Mol. Struct.*, 2000, **526**, 261–268.
13 A. L. Llamas-Saiz, C. Foces-Foces, C. Fontenas, N. Jagerovic and J. Elguero, *J. Mol. Struct.*, 1999, **484**, 197–205.
14 M. Kubicki, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2004, **60**, 191–196.
15 M. Mirmehrabi, S. Rohani, K. S. K. Murthy and B. Radatus, *J. Cryst. Growth*, 2004, **260**, 517–526.
16 R. M. Claramunt, C. López, I. Alkorta, J. Elguero, R. Yang and S. Schulman, *Magn. Reson. Chem.*, 2004, **42**, 712–714.
17 R. M. Claramunt, C. López and J. Elguero, *ARKIVOC (Gainesville, FL, USA*, 2006, **5**, 5–11.
18 L. Karin, N. David, S. Gunnel and Y. Ingvar, *Pat. Appl.*, WO 99/08500, 1999.
19 N. H. Milac and A. Copar, *US Pat.*, 2004/0122056 A1, 2004.
20 H. Ohishi, Y. In, T. Ishida, M. Inoue, F. Sato, M. Okitsu and T. Ohno, *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.*, 1989, **45**, 1921–1923.
21 J. Deng, Y. Chi, F. Fu, X. Cui, K. Yu, J. Zhu and Y. Jiang, *Tetrahedron: Asymmetry*, 2000, **11**, 1729–1732.
22 R. R. Whittle, F. D. Sancilio and G. W. Stowell, *US Pat.*, US 6,780,880 B1, 2004.
23 A. D. Bond, R. Boese and G. R. Desiraju, *Angew. Chem., Int. Ed.*, 2006, **46**, 615–617.
24 A. D. Bond, R. Boese and G. R. Desiraju, *Angew. Chem., Int. Ed.*, 2006, **46**, 618–622.
25 P. Vishweshwar, J. A. McMahon, M. Oliveira, M. L. Peterson and M. J. Zaworotko, *J. Am. Chem. Soc.*, 2005, **127**, 16802–16803.
26 M. J. Buerger and M. C. Bloom, *Z. Kristallogr.*, 1937, **96**, 182–200.