Racemic progesterone: predicted *in silico* and produced in the solid state†

Robert W. Lancaster,[a] Panagiotis G. Karamertzanis,[a] Ashley T. Hulme,[a] Derek A. Tocher,[a] Douglas F. Covey[b] and Sarah L. Price*[a]

Received (in Cambridge, UK) 10th August 2006, Accepted 27th September 2006
First published as an Advance Article on the web 16th October 2006
DOI: 10.1039/b611599c

A computational prediction that mixing the synthetic mirror image of progesterone with its natural form would produce a specific racemic crystal structure was validated.

Our computational studies of the potential crystal structures of the steroid hormone progesterone so clearly predicted the superior thermodynamic stability of a centrosymmetric structure over the known enantiomorphic polymorphs, that it inspired us to produce a fully racemic crystal of this biologically important steroid. This is a very early example of computational crystal structure prediction fulfilling its anticipated potential for the design of novel crystal structures[1–6] and chiral separation crystallization processes.[7,8]

Progesterone is used in menopausal hormone replacement therapies and as a key hormonal constituent of many oral contraceptives.[9] It has six chiral centres and, since it is synthesized from a natural product, it is only commercially available in the optically pure form, *nat*-progesterone. Two monotropically related polymorphs are known,[10,11] first characterized by hot stage microscopy,[12] and later determined by single crystal X-ray diffraction.[13,14] During our extensive polymorph screening studies, the metastable form 2 of *nat*-progesterone appeared to have 'disappeared'[15] until it was reproduced by templating.[16] In contrast, the mirror image (*ent*-progesterone) first crystallized as form 2, following its multi-stage total synthesis.[17] Thus, the factors determining the crystallization of the chirally pure forms are clearly complex, and so a computational study was undertaken to assess the possible packing modes of this steroid.

The computational prediction of the thermodynamically feasible crystal structures of progesterone had to consider its conformational flexibility (Fig. 1), which is restricted to the side chain at position 17. Quantum mechanical calculations (see ESI†) and consideration of crystallographic data for a large set of '20-oxo'-substituted pregnanes[18] led to the search for stable packing arrangements being performed with four rigid-body conformations with the keto group exposed in the vicinity of the global conformational minimum and a fifth conformation corresponding to a local minimum with the keto group less exposed. The molecular geometries were held rigid throughout the systematic generation of over 20 000 close packed crystal structures and their subsequent refinement by lattice energy minimization, using a detailed model of the *ab initio* molecular charge density for the electrostatic forces (see ESI†). The resulting crystal energy landscape is shown in Fig. 2.

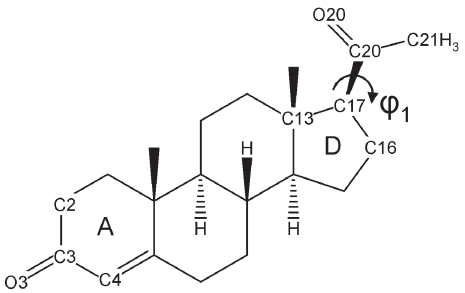

Fig. 1 Molecular structure of *nat*-progesterone. The effect of the packing forces on the rotation of the methyl-keto group φ_1 (C21–C20–C17–C13) was considered in searching for low energy crystal structures.

The two known polymorphs were correctly found as the two most stable structures within the enantiomorphic space groups (*i.e.* those that can be adopted by the optically pure molecule), consistent with the results of our polymorph screen. The computational results (Fig. 2) clearly demonstrate that there are several significantly more stable packing motifs that could only be

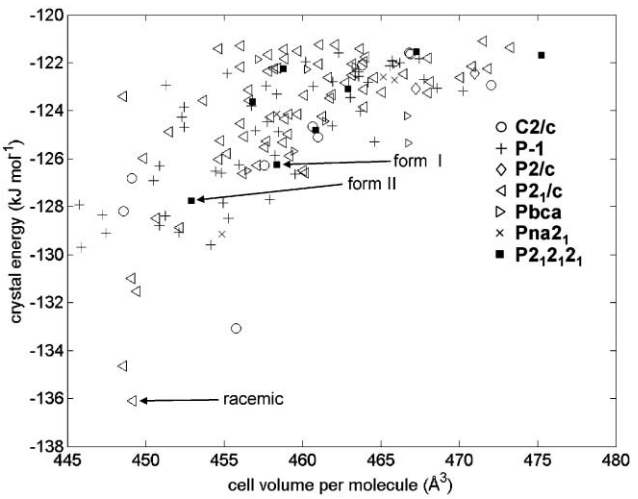

Fig. 2 Crystal energy *vs.* cell volume for the computed crystal structures of progesterone within 15 kJ mol^{-1} of the global minimum. Open and solid symbols denote crystal energy minima in racemic and enantiomorphic space groups respectively.

[a]*Department of Chemistry, University College London, 20 Gordon Street, London, UK WC1H 0AJ. E-mail: s.l.price@ucl.ac.uk; Fax: +44 (0)20 7679 7463; Tel: +44 (0)20 7679 4622*
[b]*Department of Molecular Biology and Pharmacology, Washington University School of Medicine, St. Louis, MO 63110, USA*
† Electronic supplementary information (ESI) available: Computational methodology; discussion of carbonyl···carbonyl interactions; experimental information. See DOI: 10.1039/b611599c

generated by the racemate. The predominant feature of these structures is the antiparallel arrangement of the carbonyl groups of two inversion- or glide-related molecules. In the global minimum energy structure this motif adopts a perfectly planar, almost rectangular, geometry with the two '20-keto' groups related *via* an inversion centre. Quantum mechanical calculations on propanone dimers[19] showed that this carbonyl arrangement corresponds to a particularly stable configuration with an interaction energy comparable to a medium strength hydrogen bond. Such carbonyl–carbonyl interactions are frequently observed in statistical surveys of organic crystal structures[19,20] and have been linked to the stabilization of secondary protein structure motifs.[21] The structure is further stabilized by three weak C=O···H–C interactions involving the less hindered and more basic conjugated C3=O3 carbonyl. In contrast, in both enantiomerically pure polymorphs the progesterone molecules are aligned in a head-to-tail arrangement interacting with weak C=O···H–C bonds, involving both carbonyl groups and methylene hydrogen atoms at positions 2 and 16, and a sub-optimal relative orientation of the carbonyl groups (see ESI†). Thus, the analysis of the predicted structure in terms of its functional group interactions confirmed that it was highly likely that this predicted structure would be observed if natural progesterone was crystallized in the presence of its enantiomer.

The cell constants of the expected racemic structure were deposited in St. Louis, for later validation of this prediction, prior to obtaining a sample (2.5 mg) of recently synthesized *ent*-progesterone.[17] Equal quantities of the two enantiomers were mixed, initially producing an oily residue, before crystals suitable for X-ray diffraction were ultimately crystallized from dichloromethane (see ESI†). The experimental determination revealed a centrosymmetric structure,‡ which corresponded to the predicted lowest energy racemic structure. The visual overlay (Fig. 3) confirms the excellent match. The onset of melting for the crystalline sample was determined by DSC to be 175 °C, and when compared to the published melting points for forms 1 and 2 (129 and 122 °C respectively), unambiguously confirms its superior stability.

Why was this structure predicted so successfully? Whilst it has long been recognized[22,23] that the inversion operator usually leads to more stable packing arrangements, this is usually expected to be a steric effect. However, analysis of the known and computed low energy structures (ESI†, Table S2) shows little variation in the density and the dominant repulsion–dispersion contribution to the lattice between chiral and racemic structures. It is the weak electrostatic interactions, which we have modelled particularly realistically, that predominantly determine the relative stability of the crystal structures. The racemate can optimize both the packing and the carbonyl–carbonyl interaction more effectively than *nat*-progesterone.

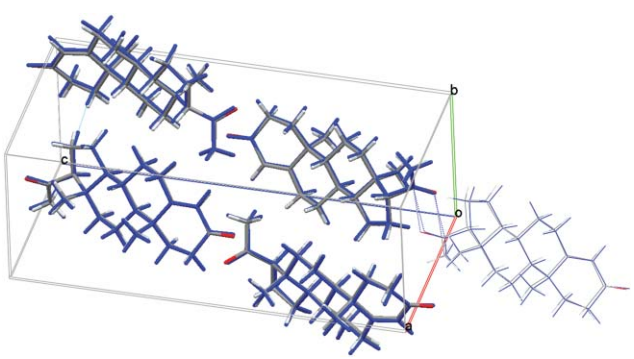

Fig. 3 Unit cell overlay of the racemic progesterone crystal structure (coloured by element) with the lowest energy structure from the search (coloured blue). A translationally related molecule is also shown (wireframe) to show the antiparallel carbonyl–carbonyl interaction (dotted lines).

The factors that determine whether a racemic mixture can be resolved by crystallization, as first observed by Pasteur for tartrates,[24] are clearly very subtle for steroids and will depend on the specific polar substituents and conformations. However, this work demonstrates the capability of crystal structure predictions for rationalizing the formation of homochiral and racemic crystals.

Understanding all the factors which determine the crystal structures of organic molecules is a major challenge[25–27] when there are many equally energetically feasible structures, and hence kinetic factors determine which are observed polymorphs. Experimentally establishing the exact crystallization conditions that can reproducibly produce the same polymorph can be challenging,[28,29] as we will report for *nat*-progesterone.[16] New experimental investigations on various molecular systems[30–32] are leading to the discovery of new polymorphs, which correspond to previously computed low energy structures,[33,34] but the existence of these specific polymorphs could not be predicted with confidence. However, when only a few structures allow optimal intermolecular interactions, as is clearly the case for progesterone, then we had sufficient confidence in our prediction to pursue experimentally a racemic structure. The success demonstrates that computational predictions can lead to novel solid forms of pharmaceuticals and so can assist pharmaceutical companies in satisfying the due diligence process in solid form selection.

This work was supported by the Basic Technology Program of the Research Councils UK www.cposs.org.uk (Grant Number GR/S24114/01 to S. L. P.) and the National Institute of Health (Grant Number GM 47969 to D. F. C.). Thermal analysis was carried out by Mrs Hilary F. Cannon (Pharmorphix, Cambridge, UK).

Notes and references

‡ *Crystal data for racemic progesterone*: $C_{21}H_{30}O_2$, M = 314.45, colourless plate, space group $P2_1/c$ (no. 14), a = 10.9219(17), b = 7.4400(12), c = 21.034(3) Å, β = 93.362(3)°, V = 1706.2(5) Å^3, Z = 4, ρ(calc.) = 1.224 g cm^{-3}, T = 150(2) K, 0.26 × 0.24 × 0.05 mm, 13 774 reflections measured, 3984 independent reflections (R_{int} = 0.063), 328 parameters refined, $R(F)$ = 0.0886 (2687 data with $I > 2\sigma(I)$), $wR(F^2)$ = 0.1750 (all data). Residual electron density extremes from the difference map: −0.27 and +0.29 e Å^{-3}. CCDC 606954. For crystallographic data in CIF or other electronic format see DOI: 10.1039/b611599c.

1 G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerrigter, R. G. Della Valle, E. Venuti, A. Dzyabchenko, J. D. Dunitz, B. Schweizer, B. P. van Eijck, P. Erk, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, F. J. J. Leusen, C. Liang, C. C. Pantelides, P. G. Karamertzanis, S. L. Price, T. C. Lewis, H. Nowell, A. Torrisi, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt and P. Verwer, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2005, **61**(5), 511–527.
2 S. L. Price, *Adv. Drug Delivery Rev.*, 2004, **56**(3), 301–319.
3 J. D. Dunitz, *Chem. Commun.*, 2003, 545–548.
4 A. Gavezzotti, *Model. Simul. Mater. Sci. Eng.*, 2002, **10**(3), R1–R29.
5 G. R. Desiraju, *Nature*, 2001, **6845**(412), 397–400.

6 G. R. Desiraju, *Science*, 1997, **5337**(278), 404–405.
7 F. J. J. Leusen, *Cryst. Growth Des.*, 2003, **3**(2), 189–192.
8 P. G. Karamertzanis and S. L. Price, *J. Phys. Chem. B*, 2005, **109**(36), 17 134–17 150.
9 E. Hirvonen, *Maturitas*, 1996, **23**, S13–S18.
10 A. Burger and R. Ramberger, *Mikrochim. Acta*, 1979, **11**, 259–271.
11 M. Muramatsu, M. Iwahashi and U. Takeuchi, *J. Pharm. Sci.*, 1979, **68**(2), 175–177.
12 M. Brandstätter-Kuhnert and A. Kofler, *Mikrochim. Acta*, 1959, 847–853.
13 E. Foresti Serantoni, A. Krajewski, R. Mongiorgi, L. Riva de Sanseverino and R. Cameroni, *Cryst. Struct. Commun.*, 1975, **4**, 189–192.
14 H. Campsteyn, L. Dupont and O. Dideberg, *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.*, 1972, **28**(10), 3032–3042.
15 J. D. Duntiz and J. Bernstein, *Acc. Chem. Res.*, 1995, **28**, 193–200.
16 R. W. Lancaster, P. G. Karamertzanis, A. T. Hulme, D. A. Tocher, D. F. Covey and S. L. Price, *J. Pharm. Sci.*, 2006, manuscript in preparation.
17 R. J. Auchus, A. S. Kumar, C. A. Boswell, M. K. Gupta, K. Bruce, N. P. Rath and D. F. Covey, *Arch. Biochem. Biophys.*, 2003, **409**(1), 134–144.
18 W. L. Duax, J. F. Griffin and D. C. Rohrer, *J. Am. Chem. Soc.*, 1981, **103**(22), 6705–6712.
19 F. H. Allen, C. A. Baalham, J. P. M. Lommerse and P. R. Raithby, *Acta Crystallogr., Sect. B: Struct. Sci.*, 1998, **54**, 320–329.
20 S. Lee, A. B. Mallik and D. C. Fredrickson, *Cryst. Growth Des.*, 2004, **4**(2), 279–290.
21 R. Paulini, K. Muller and F. Diederich, *Angew. Chem., Int. Ed.*, 2005, **44**(12), 1788–1805.
22 C. P. Brock, W. B. Schweizer and J. D. Dunitz, *J. Am. Chem. Soc.*, 1991, **113**(26), 9811.
23 O. Wallach, *Justus Liebigs Ann. Chem.*, 1895, **286**, 90.
24 M. L. Pasteur, *Ann. Chim. Phys., S3*, 1850, **28**, 56–117.
25 J. D. Dunitz and H. A. Scheraga, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**(40), 14 309–14 311.
26 R. J. Davey, K. Allen, N. Blagden, W. I. Cross, H. F. Lieberman, M. J. Quayle, S. Righini, L. Seton and G. J. T. Tiddy, *CrystEngComm*, 2002, **4**, 257–264.
27 A. Gavezzotti, *CrystEngComm*, 2002, **4**, 343–347.
28 T. Threlfall, *Org. Process Res. Dev.*, 2000, **4**(5), 384–390.
29 A. J. Florence, A. Johnston, A. L. Price, H. Nowell, A. R. Kennedy and N. Shankland, *J. Pharm. Sci.*, 2006, **95**(9), 1918–1930.
30 F. P. A. Fabbiani, D. R. Allan, S. Parsons and C. R. Pulham, *CrystEngComm*, 2005, **7**(29), 179–186.
31 P. Vishweshwar, J. A. McMahon, M. Oliveira, M. L. Peterson and M. J. Zaworotko, *J. Am. Chem. Soc.*, 2005, **127**(48), 16 802–16 803.
32 G. M. Day, A. V. Trask, W. D. S. Motherwell and W. Jones, *Chem. Commun.*, 2006, 54–56.
33 H. Nowell and S. L. Price, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2005, **61**, 558–568.
34 C. Ouvrard and S. L. Price, *Cryst. Growth Des.*, 2004, **4**(6), 1119–1127.