Tunable recognition of the steroid α-face by adjacent π-electron density

T. Friščić, R. W. Lancaster, L. Fábián, P. G. Karamertzanis

Table of contents

The electronic supplementary information provided herein contains only limited powder X-ray diffraction and reflectance FTIR characterisation data. Further data is available as Supplementary sections S8 and S9 at:

http://www.chem.ucl.ac.uk/basictechorg/publications/si_steroid_alpha_face_pnas.pdf

S1. Experimental

All reagents and solvents were purchased from Sigma-Aldrich Chemical Co. and were used without further purification.

S1.1. Mechanochemical screening

Mechanochemical screening for the formation of cocrystals was performed using a Retsch MM2200 grinder mill operating at 30 Hz. Each experiment was performed by placing 100 mg of a mixture of the steroid and a cocrystal former in a 1:1 stoichiometric ratio into a 10 mL volume stainless steel jar, along with 30 μL of nitromethane and two stainless steel balls of 7 mm diameter. The mixture was then ground for 20 minutes, left to stand open in air for several minutes and examined using PXRD/FTIR-ATR. During grinding, the jars were exposed to a strong stream of air, preventing significant increase in the average temperature of the reaction mixture. Changing the grinding liquid did not affect the outcome of grinding, as verified in case of progesterone and pyrene, where ethanol, methanol and acetonitrile were also used as grinding liquids.

S1.2. Cocrystal formation and determination of cocrystal stoichiometry

The formation of a cocrystal upon liquid-assisted grinding was recognised through the appearance of new reflections in the PXRD pattern, which could not be assigned to any known solid forms (e.g. polymorphs or solvates or hydrates) of the reactants. In some cases, the only visible change to the PXRD pattern of the reactant mixture was the disappearance of reflections for the cocrystal former. This was interpreted through the amorphisation of the cocrystal former and was, consequently, interpreted as a lack of cocrystal formation.

In cases where the cocrystal formed, the stoichiometry of the product was established by varying the ratio of the cocrystal former and the steroid in the reaction mixture, until the PXRD pattern exhibited only reflections belonging to the cocrystal. This is illustrated on Figure S1 for (**pro**)·(**13**), a cocrystal composed of the steroid and the cocrystal former in a 1:1 stoichiometric ratio.

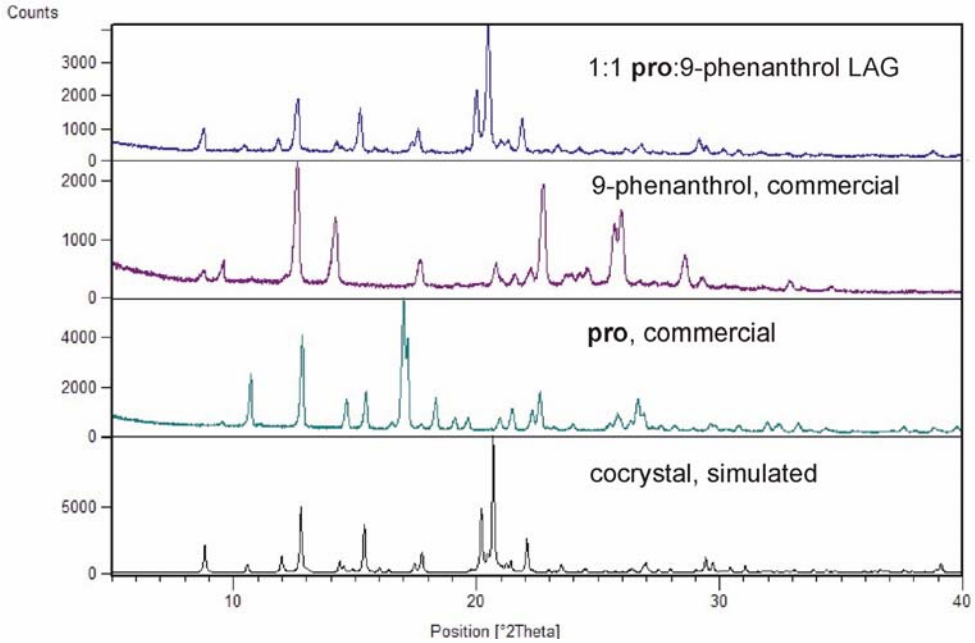

Figure S1 Overlay of PXRD patterns (bottom to top): simulated for the (**pro**)·(**13**) cocrystal; commercial **pro**; commercial **13** and the product of a LAG reaction of **pro** and **13** in 1:1 stoichiometric ratio. Nitromethane was used as the LAG liquid.

The formation of a 2:1 steroid:cocrystal former product is illustrated in Figure S2 for the cocrystal (**pro**)$_2$(**22**). The product for the 1:1 reaction clearly shows reflections belonging to excess pyrene (**22**, indicated by the vertical dotted lines). However, grinding of **pro** and **22** in a 2:1 ratio provided the pure cocrystal, with a PXRD pattern that completely matched the one calculated from the experimentally determined crystal structure.

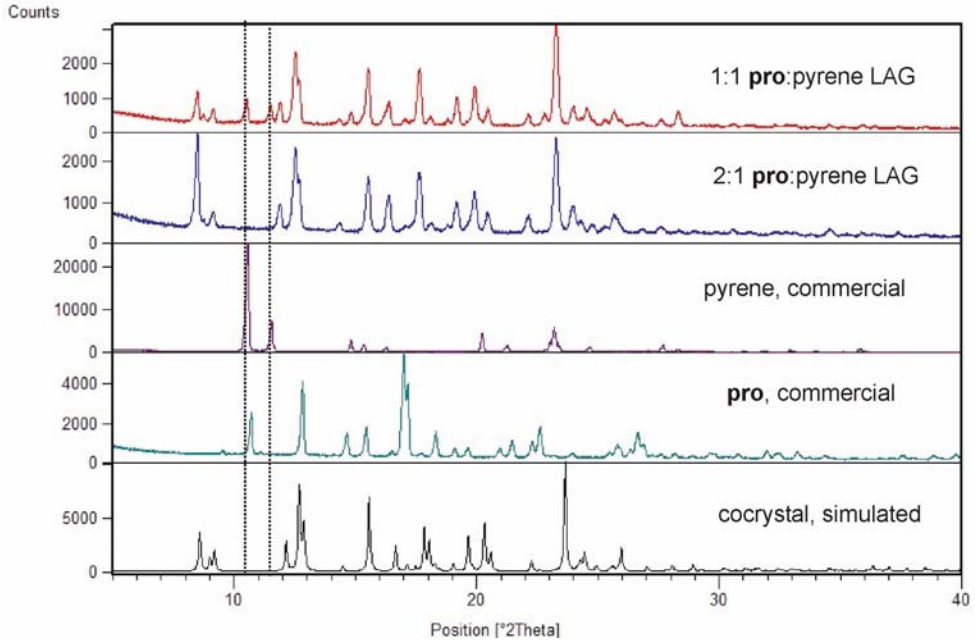

Figure S2. Overlay of PXRD patterns (bottom to top): simulated for the (**pro**)$_2$·(**22**) cocrystal; commercial **pro**; commercial **22;** the product of a LAG reaction of **pro** and **22** in 2:1 stoichiometric ratio and the product of a LAG reaction of **pro** and **22** in 1:1 stoichiometric ratio. Nitromethane was used as the LAG liquid. The vertical dotted lines indicate the positions of two characteristic reflections of **22**.

Figure S3 illustrates the formation of a (**pro**)$_3$·(**21**) cocrystal composed of the steroid and the cocrystal former in a 3:1 stoichiometric ratio. The product from the 1:1 reaction clearly shows

reflections belonging to excess phenanthrene (**21**, indicated by the vertical dotted lines). However, grinding of **pro** and **21** in a 3:1 ratio provided the pure cocrystal, with a PXRD pattern that entirely matches the one calculated from the experimentally determined crystal structure.

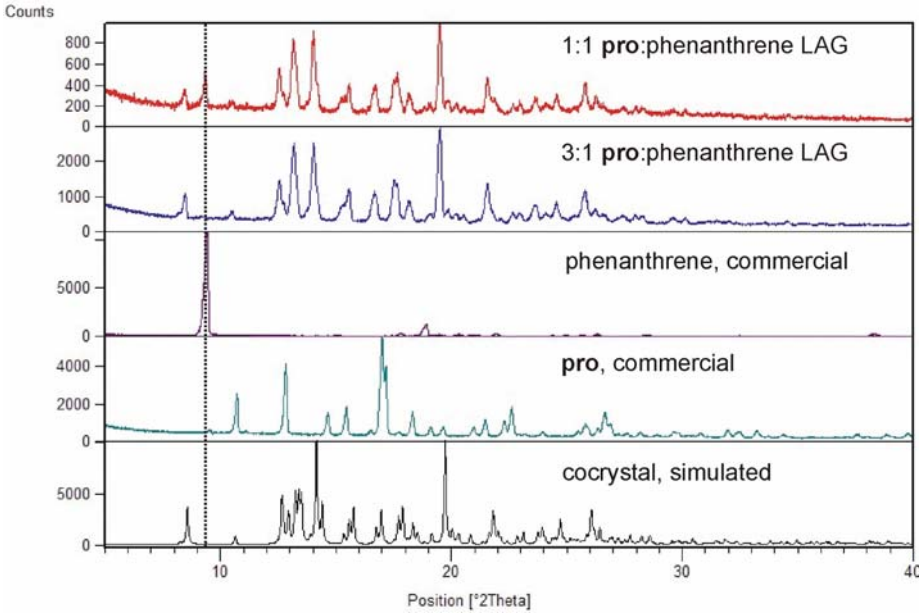

Figure S3. Overlay of PXRD patterns (bottom to top): simulated for the (**pro**)$_3$·(**21**) cocrystal; commercial **pro**; commercial **21;** the product of a LAG reaction of **pro** and **21** in 3:1 stoichiometric ratio and the product of a LAG reaction of **pro** and **21** in 1:1 stoichiometric ratio. Nitromethane was used as the LAG liquid. The vertical dotted lines indicate the positions of two characteristic reflections of **21**.

S1.3. Single crystal growth

Single crystals have been obtained by mixing stoichiometric quantities of the steroid and the cocrystal former (at the 0.1 mmol level) in nitromethane. Brief sonication (30 sec) resulted in the formation of fine crystals of the cocrystal, that were dissolved by heating the suspension to boiling and cooling it down to room temperature. Single crystals were obtained by prolonged standing (ca. 1 day) of the cooled solution.

S1.4. Slurry screening for polymorphs of 9-phenanthrol (13)

The screening for polymorphs of **13** was performed by placing 50 mg of the commercial solid into 1 mL of the following solvents: chloroform, ethyl acetate, nitromethane, di(*n*-butyl)ether and 1,1,1-trifluoroethanol. The suspension was then slurried for 24 hours at ambient temperature and subsequently filtered and analysed using PXRD. In cases where **13** was readily soluble in the slurrying solvent, a saturated solution was first prepared by adding 50 mg portions of **13** to the solvent until no more dissolved.

S1.5. Analytical methods

Powder X-ray diffraction (PXRD). PXRD data was collected on a laboratory Philips X'Pert Pro diffractometer, equipped with an X'celerator RTMS detector, using Ni-filtered CuKα radiation, using a flat plate configuration. Data were typically collected in the 2θ range 5-40^o within a time period of 2 mins 40 sec.

Single crystal X-ray diffraction. Single crystal X-ray diffraction data were collected on a Nonius Kappa CCD diffractometer equipped with a graphite monochromator and an Oxford cryostream, using Mo*K*α radiation. Structure solution and refinement was performed using SHELX available with the WinGX package of crystallographic tools, running on a Pentium-based PC under MS Windows XP.

Differential scanning calorimetry (DSC). DSC thermograms for the prepared cocrystals were recorded on a Mettler DSC822^e calorimeter with typical sample weights of 7-8 mg. Each sample was placed in pierced lid sealed aluminium pan of 40 μL volume. Measurements were performed in a dynamic atmosphere of nitrogen with a flow of 80 $cm^{-3}min^{-1}$.

S1.6. Rietveld refinement and PXRD crystal structure solution for (pre)·(2)

The PXRD pattern of the (**pre**)·(**2**) cocrystal sample suggested isostructurality with the known (**pre**)·(4-iodophenol) cocrystal.[1] The experimental pattern could be readily indexed to give similar unit cell

parameters to those of (**pre**)·(4-iodophenol), and the systematic absences were consistent with the expected space group ($P2_12_12_1$). Pawley refinement in DASH yielded a satisfactory fit (profile χ^2 = 3.60), except for the presence of small extra peaks at 16.1 and 16.3 degrees. Structure solution with DASH confirmed the isostructurality of the 4-iodophenol and 4-bromophenol cocrystals of **pre**. Rietveld refinement was carried out using GSAS and EXPGUI. A single phase refinement did not give satisfactory results, and the presence of extra peaks suggested that unreacted **pre** was present in the sample. Therefore, the powder sample was described as a mixture of two phases: the structure determined for (**pre**)·(**2**) and the already known **pre** structure (CCDC code: PREGOL).[2] Atomic coordinates in the **pre** phase were kept fixed during the refinement. Bond lengths and angles of **pre** in the cocrystal phase were restrained to values taken from the (**pre**)·(4-iodophenol) structure, while the bond lengths and angles of **2** were restrained to values from the crystal structure of **2**.[3] A planar group restraint was applied for the aromatic ring. A common isotropic displacement parameter was used for all non-hydrogen atoms in each phase, while the displacement parameters of all hydrogen atoms were constrained to be 1.2 times the common displacement parameter value. Peak profiles were modelled by an asymmetry corrected pseudo-Voigt function,[4] and an eighteen parameter Chebyshev polynomial of the 1st kind was used to describe the background. The Rietveld refinement converged to yield acceptable figures of merit (R_p = 0.0386, R_{wp} = 0.0508). The refinement indicates that the sample contained 4.37(9)% m/m **pre**.

S1.7. Crystallographic data

Table S1. General and crystallographic parameters for cocrystals (**pro**)·(**2**), (**pro**)$_2$·(**12**), (**pro**)·(**13**), (**pro**)$_2$·(**14**) and (**pro**)·(**15**)

	(**pro**)·(**2**)	(**pro**)$_2$·(**12**)	(**pro**)·(**13**)	(**pro**)$_2$·(**14**)	(**pro**)·(**15**)
Formula	$C_{27}H_{35}BrO_3$	$C_{58}H_{70}O_5$	$C_{35}H_{40}O_3$	$C_{52}H_{68}O_6$	$C_{28}H_{36}O_6$
M_r	487.5	847.2	508.7	789.1	468.6
Crystal	monoclinic	orthorhombic	orthorhombic	triclinic	orthorhombic
Space group	$P2_1$	$P2_12_12_1$	$P2_12_12_1$	P1	$P2_12_12_1$
a (Å)	7.1675(2)	10.3853(1)	9.2036(2)	7.4957(1)	7.2877(1)
b (Å)	17.1472(5)	10.8906(1)	14.7632(3)	11.2843(2)	14.0524(2)
c (Å)	19.5212(8)	41.1705(5)	20.0513(5)	13.4237(3)	23.3852(4)
α (°)	90	90	90	78.283(2)	90
β (°)	90.382(1)	90	90	83.673(2)	90
γ (°)	90	90	90	74.829(1)	90
V (Å^3)	2399.2(1)	4656.47(1)	2724.46(1)	1071.06(3)	2394.87(6)
ρ (g cm^{-3})	1.35	1.21	1.24	1.22	1.30
λ (Å)	0.71073	0.71073	0.71073	0.71073	0.71073
μ (mm^{-1})	1.739	0.075	0.077	0.078	0.090
R (or Rp)	0.059	0.069	0.058	0.057	0.048
wR_2 (or Rwp)	0.121	0.199	0.133	0.166	0.121
S	1.125	1.035	1.071	1.068	1.036
$\rho_{min,\ max}$ (e Å^{-3})	-0.55, 0.40	-0.32, 0.79	-0.25, 0.50	-0.45, 0.21	-0.19, 0.23

Table S2. General and crystallographic parameters for cocrystals (**pro**)$_3$·(**21**), (**pro**)$_2$·(**22**), (**pre**)·(**2**), (**pre**)·(**5**) and (**bes**)·(**22**)

	(**pro**)$_3$·(**21**)	(**pro**)$_2$·(**22**)	(**pre**)·(**2**)	(**pre**)·(**5**)	(**bes**)·(**22**)
Formula	$C_{77}H_{100}O_6$	$C_{58}H_{70}O_4$	$C_{27}H_{37}BrO_3$	$C_{31}H_{40}O_3$	$C_{34}H_{34}O_2$
M_r	1121.6	831.1	489.5	460.6	474.6
Crystal	monoclinic	triclinic	orthorhombic	orthorhombic	orthorhombic
Space group	$P2_1$	$P1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
a (Å)	7.4342(1)	9.8401(2)	6.2615(1)	6.2027(1)	7.5090(1)
b (Å)	39.6492(3)	11.3738(2)	15.4833(2)	15.7511(3)	17.9935(4)
c (Å)	11.1961(1)	11.4241(3)	25.4155(5)	26.0999(6)	19.0703(5)
α (^{o})	90	115.037(1)	90	90	90
β (^{o})	107.318(1)	91.454(1)	90	90	90
γ (^{o})	90	91.791(1)	90	90	90
V (Å^3)	3150.56(6)	1156.82(4)	2463.98(7)	2549.94(9)	2576.65(9)
ρ (g cm^{-3})	1.18	1.19	1.32	1.20	1.22
λ (Å)	0.71703	0.71703	1.5418	0.71703	0.71703
μ (mm^{-1})	0.073	0.073	2.462	0.075	0.074
R (or *R*p)	0.048	0.058	0.058 (*R*p)	0.037	0.062
wR_2 (or *R*wp)	0.112	0.137	0.051 (*R*wp)	0.092	0.154
S	1.017	1.039	1.73	1.084	1.059
$\rho_{min, max}$ (e Å$^-$	-0.19, 0.19	-0.18, 0.22	-0.32, 0.32	-0.16, 0.16	-0.27, 0.39

Table S3. General and crystallographic parameters for (androsterone)·(**5**) cocrystal and 9-phenanthrol (**13**).

	(androsterone)·(**5**)	9-phenanthrol (**13**)
Formula	$C_{29}H_{38}O_3$	$C_{14}H_{10}O$
M_r	434.6	194.2
Crystal system	orthorhombic	orthorhombic
Space group	$P2_12_12_1$	$P2_12_12_1$
a (Å)	7.2711(1)	4.6819(1)
b (Å)	17.7776(2)	12.3504(6)
c (Å)	18.4071(3)	16.766(1)
α (°)	90	90
β (°)	90	90
γ (°)	90	90
V (Å^3)	2379.35(6)	969.5(1)
ρ (g cm^{-3})	1.21	1.33
λ (Å)	0.71703 (MoKα)	1.5418 (CuKα)
μ (mm^{-1})	0.076	0.647
R (or *R*p)	0.039	0.059 (*R*p)
wR_2 (or *R*wp)	0.107	0.040 (*R*wp)
S	1.044	0.45
$\rho_{min,\ max}$ (e Å^{-3})	-0.19, 0.27	-0.16, 0.16

S1.8 Thermal ellipsoid plots

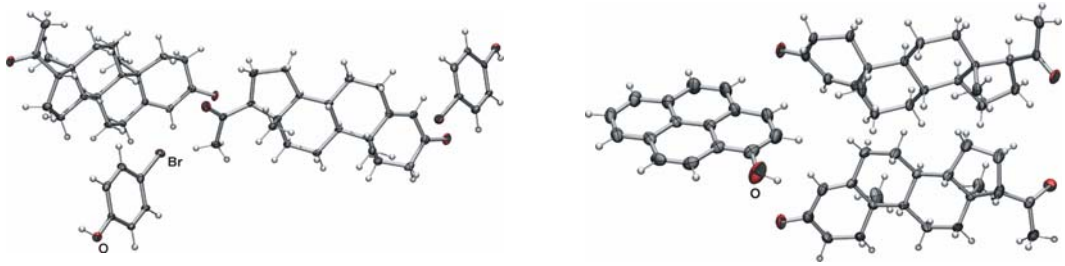

Figure S4. Asymmetric unit of: (a) (**pro**)·(**2**) and (b) (**pro**)$_2$·(**12**) cocrystal. Thermal ellispoids are shown at the 30% probability level.

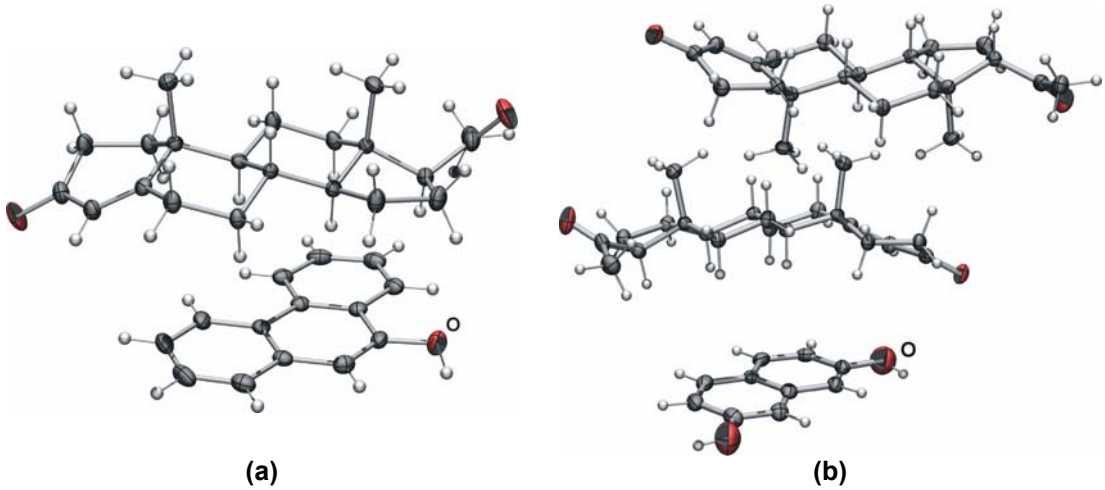

Figure S5. Asymmetric unit of: (a) (**pro**)·(**13**) cocrystal and (b) (**pro**)$_2$·(**14**) cocrystal. Thermal ellispoids are shown at the 30% probability level.

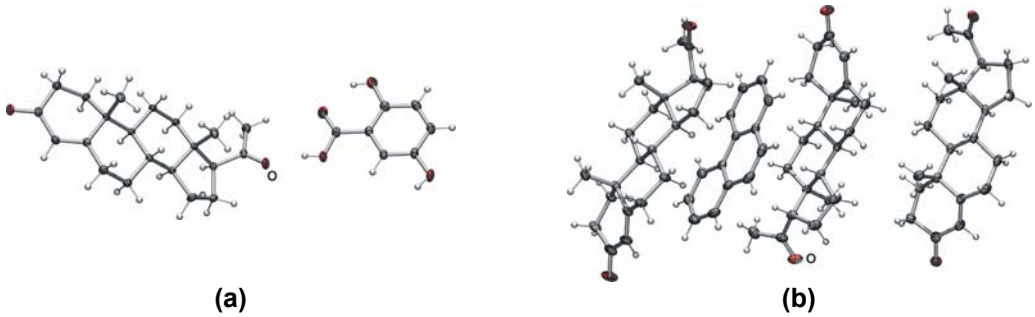

Figure S6. Asymmetric unit of: (a) (**pro**)·(**15**) cocrystal and (b) (**pro**)$_3$·(**21**) cocrystal. Thermal ellispoids are shown at the 30% probability level.

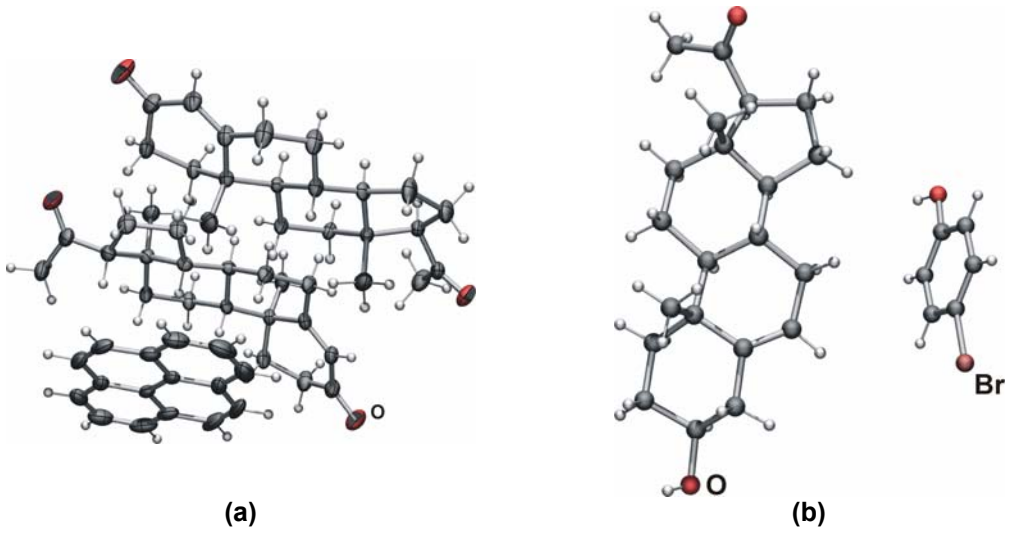

Figure S7. Asymmetric unit of: (a) (**pro**)$_2$·(**22**) cocrystal and (b) (**pre**)·(**2**) cocrystal. Thermal ellispoids are shown at the 30% probability level.

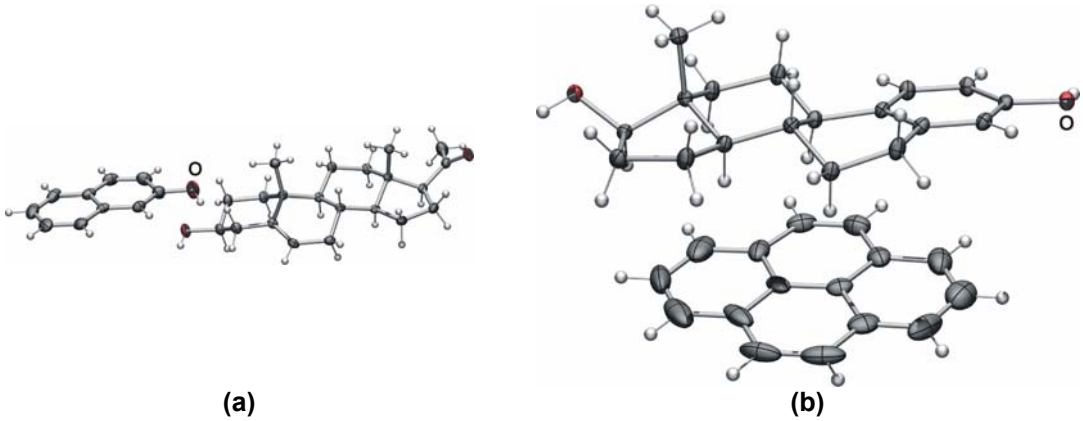

Figure S8. Asymmetric unit of: (a) (**pre**)·(**5**) cocrystal and (b) (**bes**)·(**22**) cocrystal. Thermal ellispoids are shown at the 30% probability level.

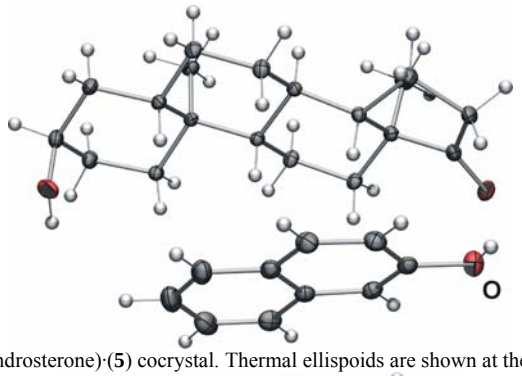

Figure S9. Asymmetric unit of (androsterone)·(**5**) cocrystal. Thermal ellispoids are shown at the 30% probability level.

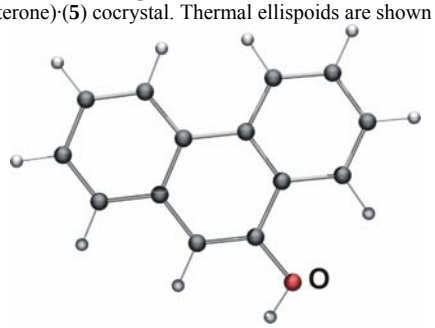

Figure S10. Asymmetric unit of **13**. Thermal ellispoids are shown at the 30% probability level.

S2. PXRD analysis of LAG screening results

2.1. Progesterone experiments

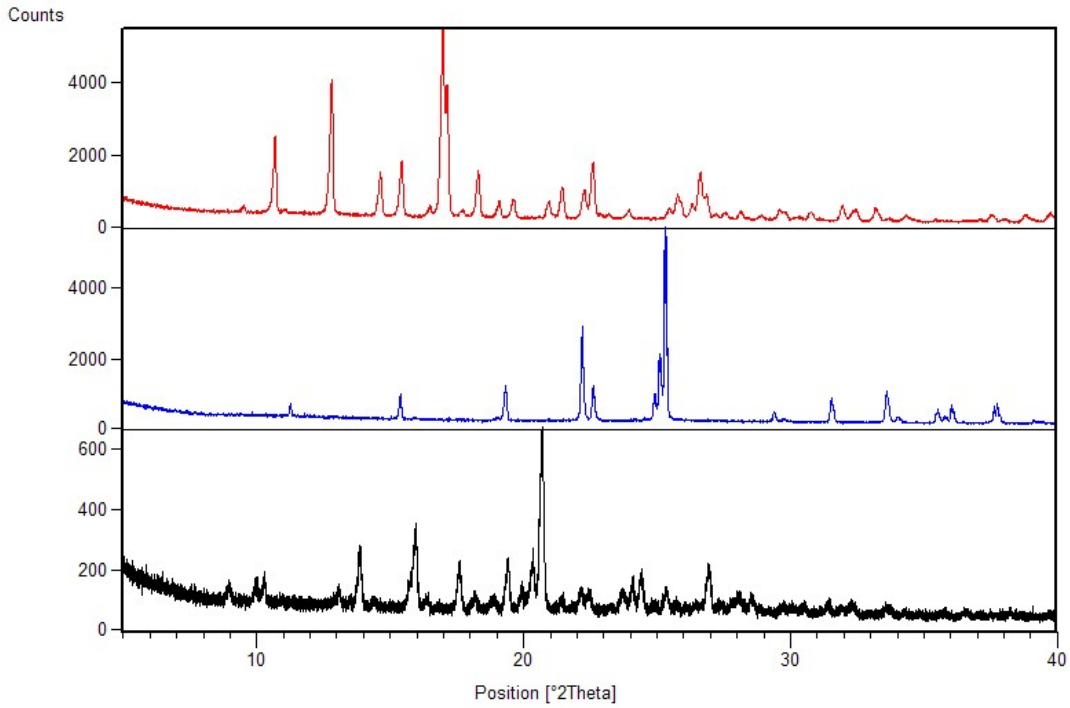

Figure S11. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) 4-bromophenol and (c) product of LAG reaction of **pro** and 4-bromophenol in 1:1 stoichiometric ratio. Cocrystal forms.

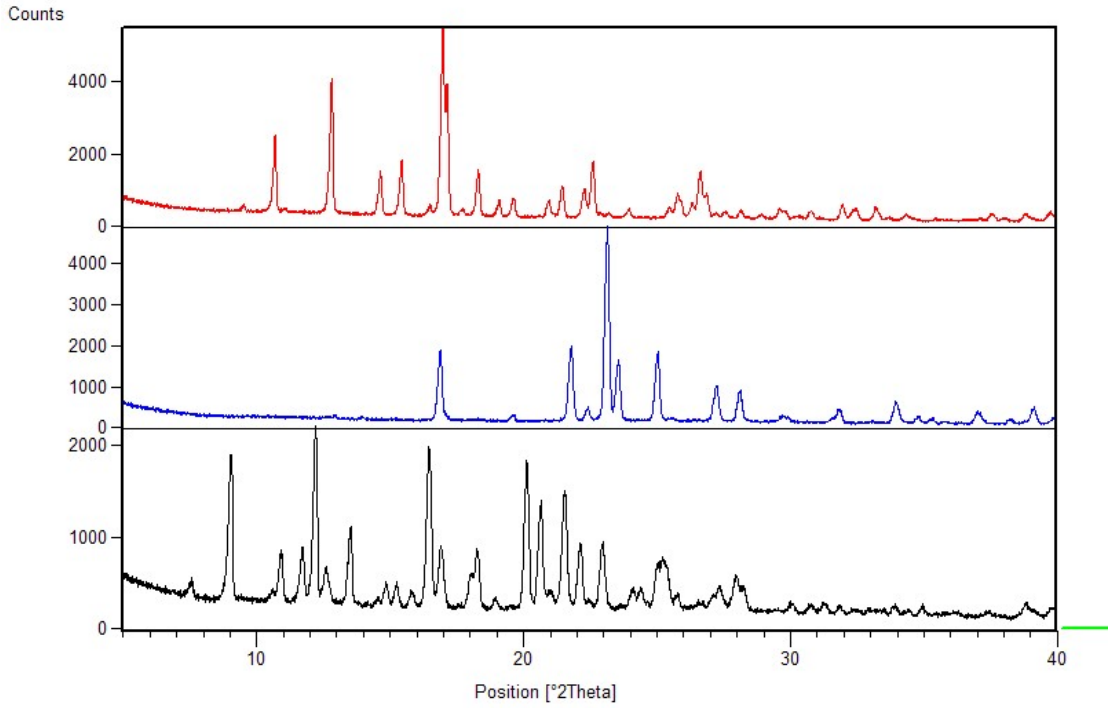

Figure S12. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) phloroglucinol and (c) product of LAG reaction of **pro** and phloroglucinol in 1:1 stoichiometric ratio. Cocrystal forms.

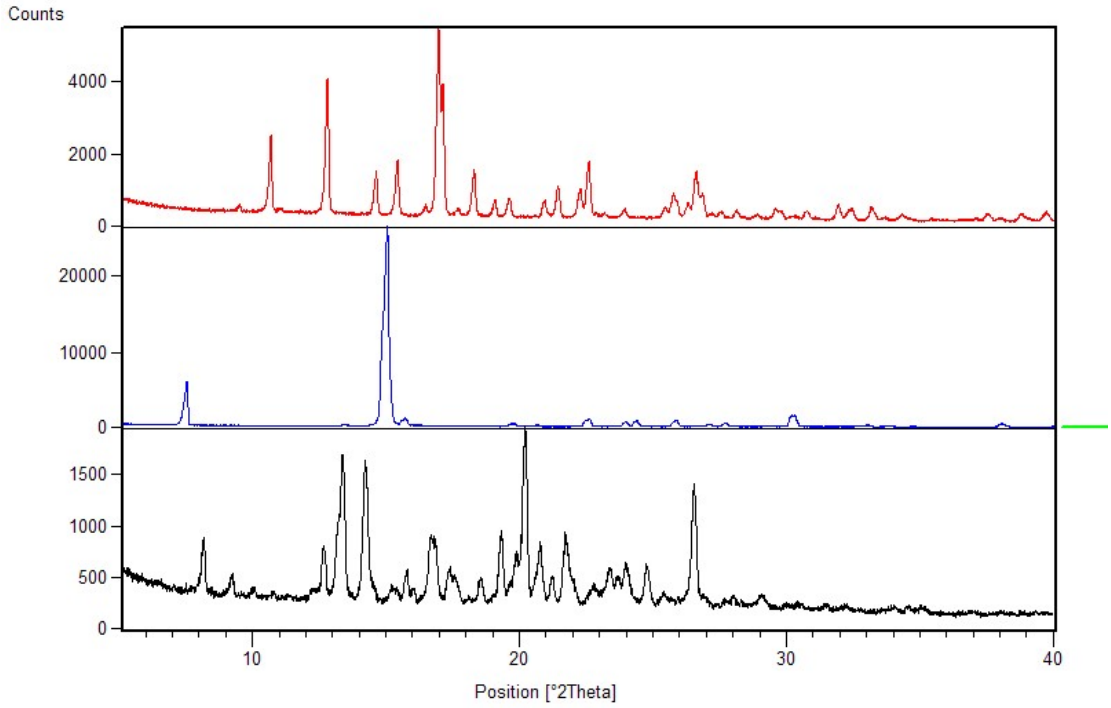

Figure S13. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) 1-naphthol and (c) product of LAG reaction of **pro** and 1-naphthol in 1:1 stoichiometric ratio. Cocrystal forms.

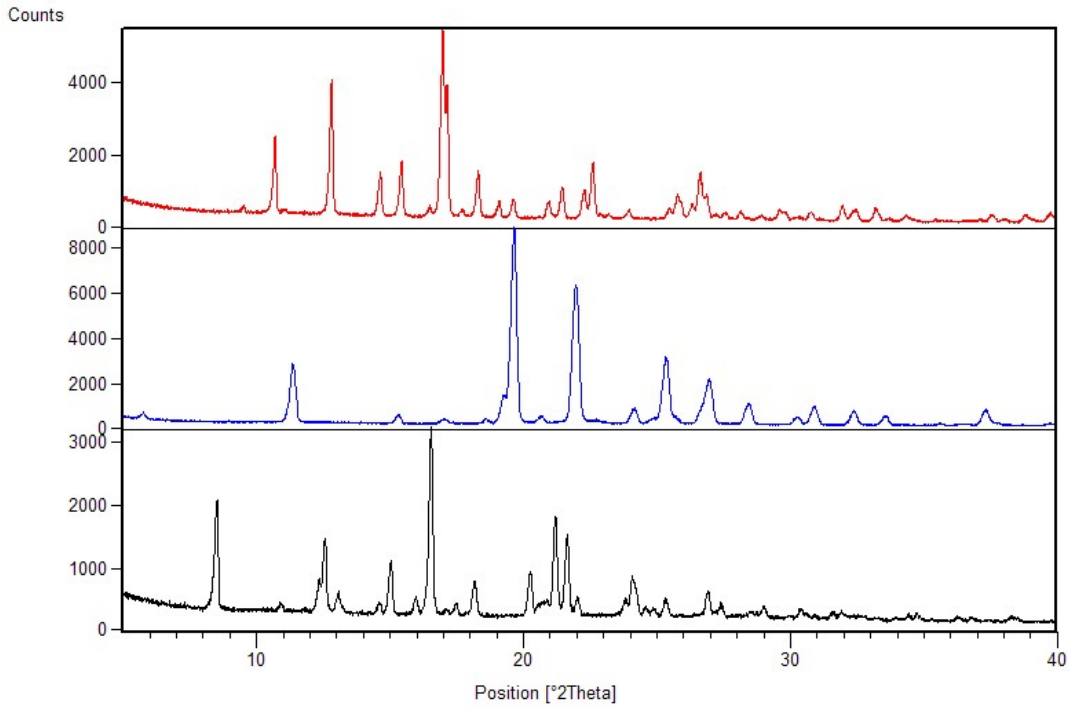

Figure S14. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) 2-naphthol and (c) product of LAG reaction of **pro** and 2-naphthol in 1:1 stoichiometric ratio. Cocrystal forms.

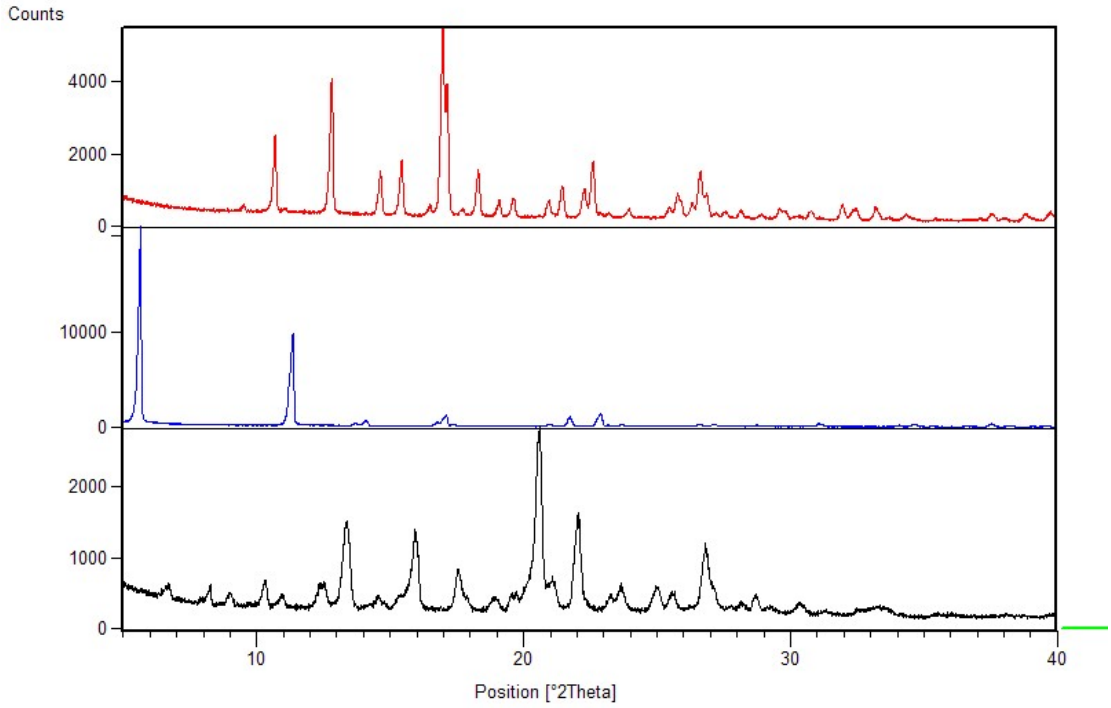

Figure S15. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) 1-naphthoic acid and (c) product of LAG reaction of **pro** and 1-naphthoic acid in 1:1 stoichiometric ratio. Cocrystal forms.

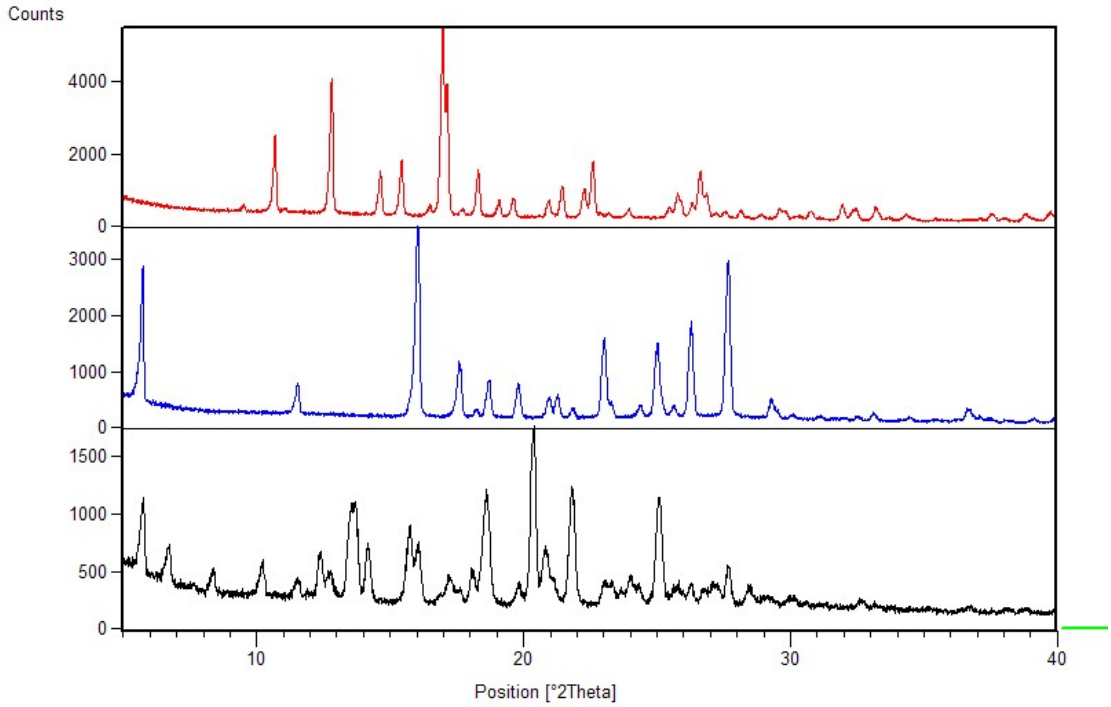

Figure S16. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) 2-naphthoic acid and (c) product of LAG reaction of **pro** and 2-naphthoic acid in 1:1 stoichiometric ratio. Cocrystal forms.

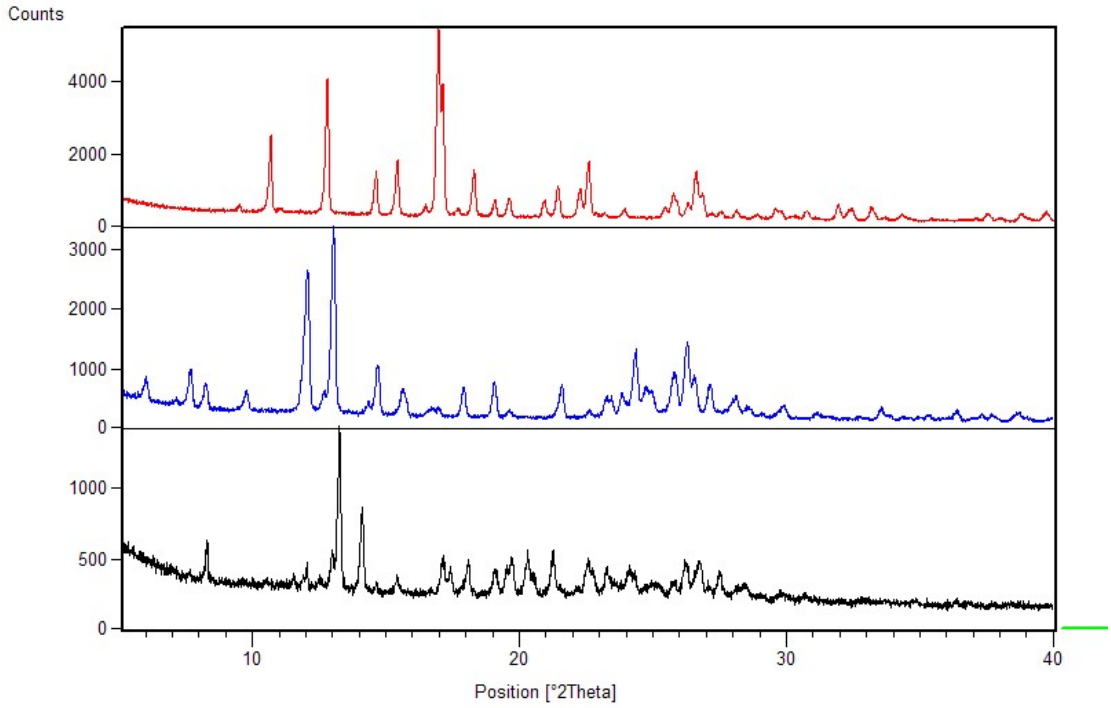

Figure S17. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) 2-hydroxy-1-naphthoic acid and (c) product of LAG reaction of **pro** and 2-hydroxy-1-naphthoic acid in 1:1 stoichiometric ratio. Cocrystal forms.

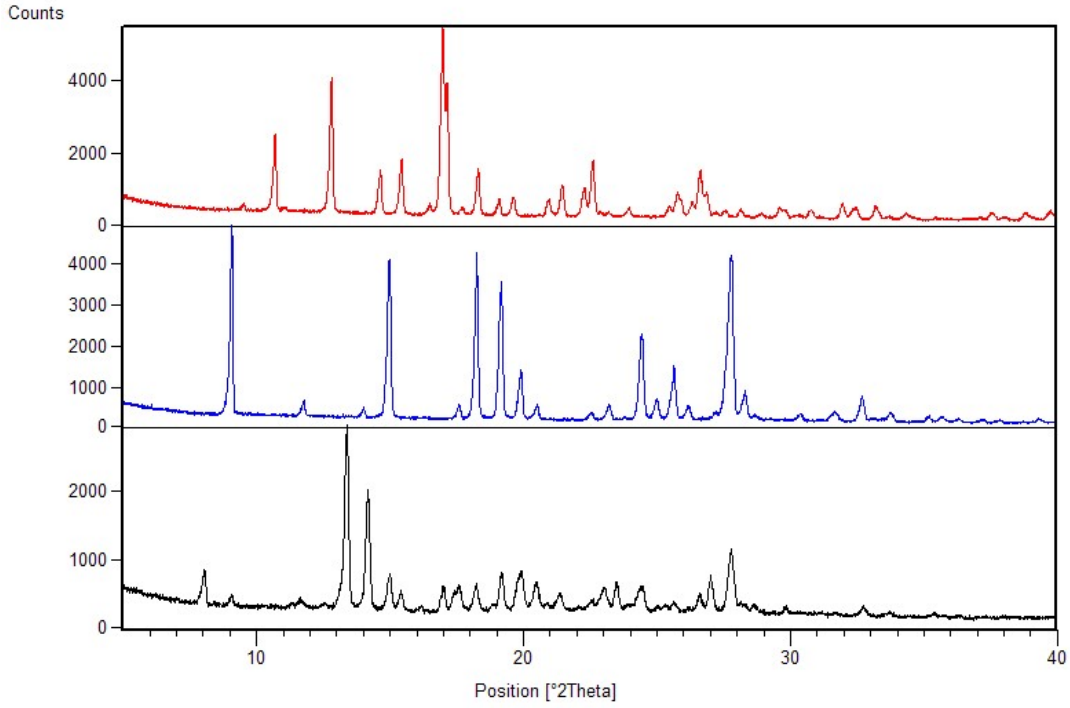

Figure S18. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) 2-hydroxy-3-naphthoic acid and (c) product of LAG reaction of **pro** and 2-hydroxy-3-naphthoic acid in 1:1 stoichiometric ratio. Cocrystal forms.

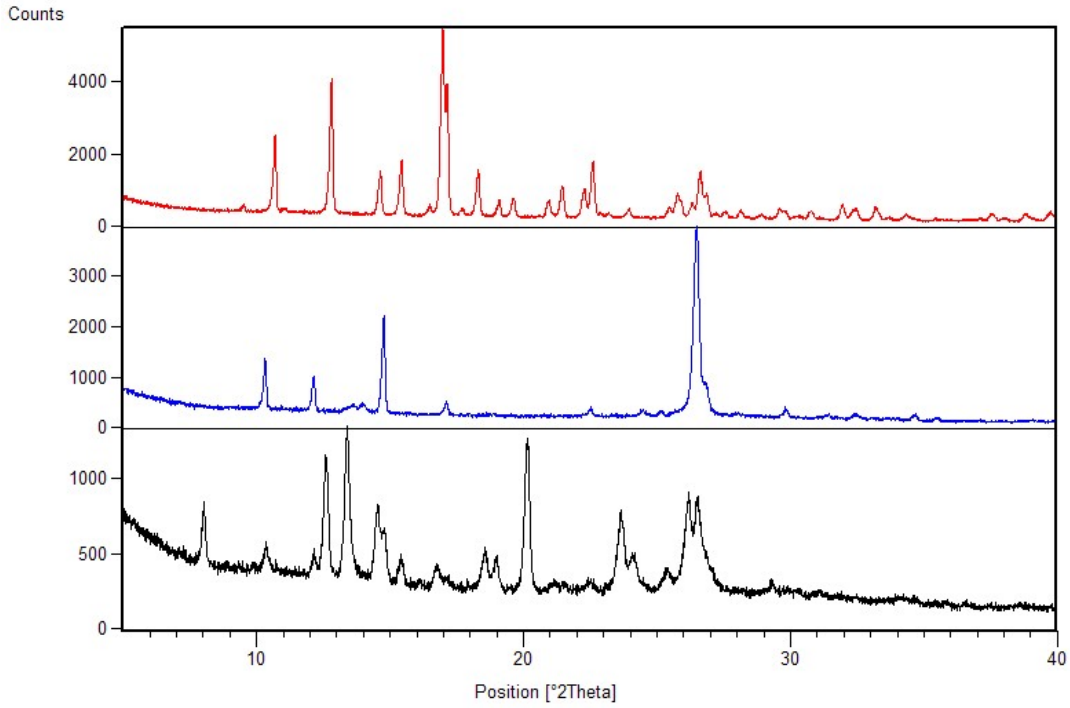

Figure S19. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) 1-hydroxy-2-naphthoic acid and (c) product of LAG reaction of **pro** and 1-hydroxy-2-naphthoic acid in 1:1 stoichiometric ratio. Cocrystal forms.

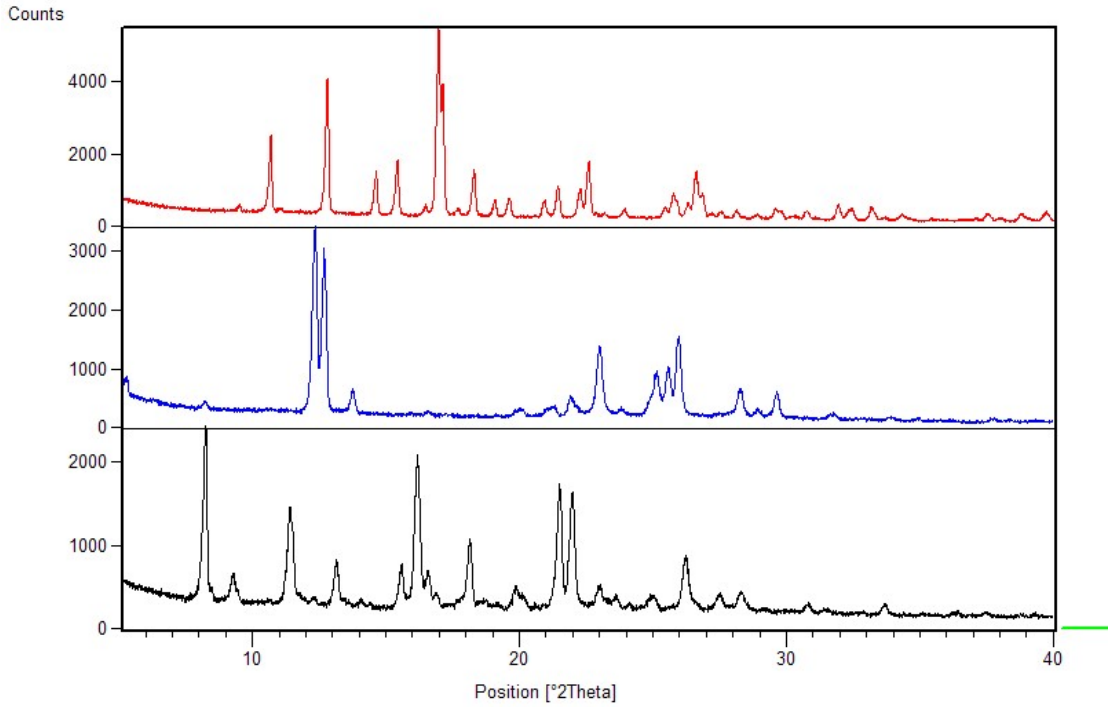

Figure S20. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) pyrenol and (c) product of LAG reaction of **pro** and pyrenol in 1:1 stoichiometric ratio. Cocrystal forms.

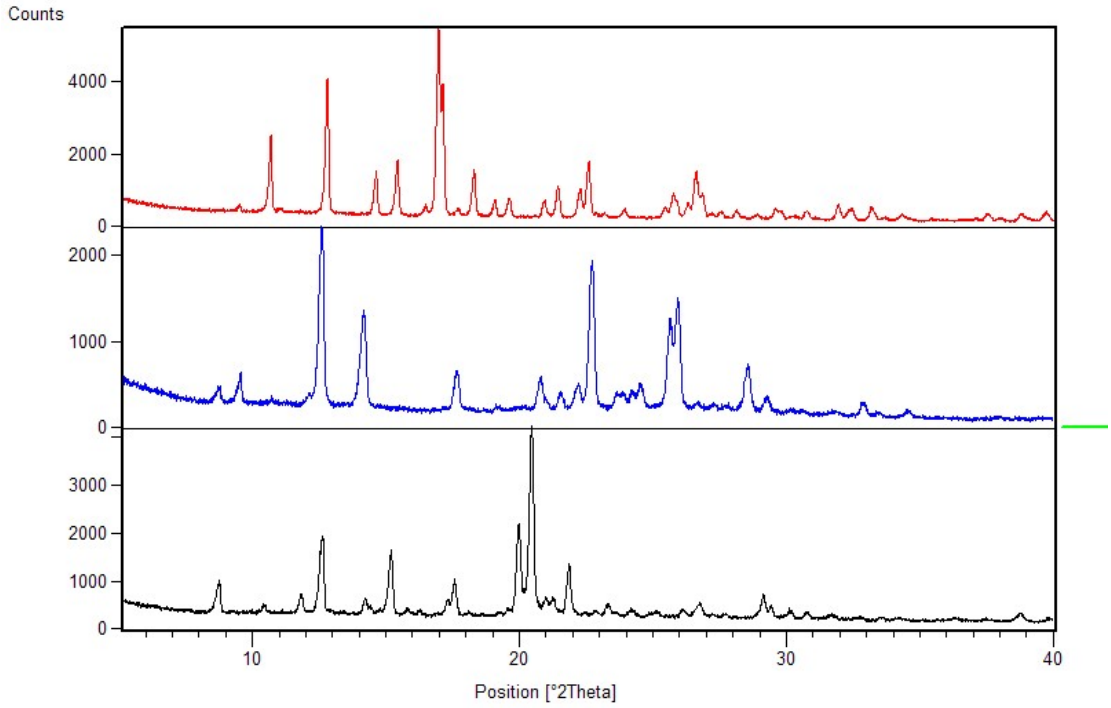

Figure S21. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) 9-phenanthrol and (c) product of LAG reaction of **pro** and 9-phenanthrol in 1:1 stoichiometric ratio. Cocrystal forms.

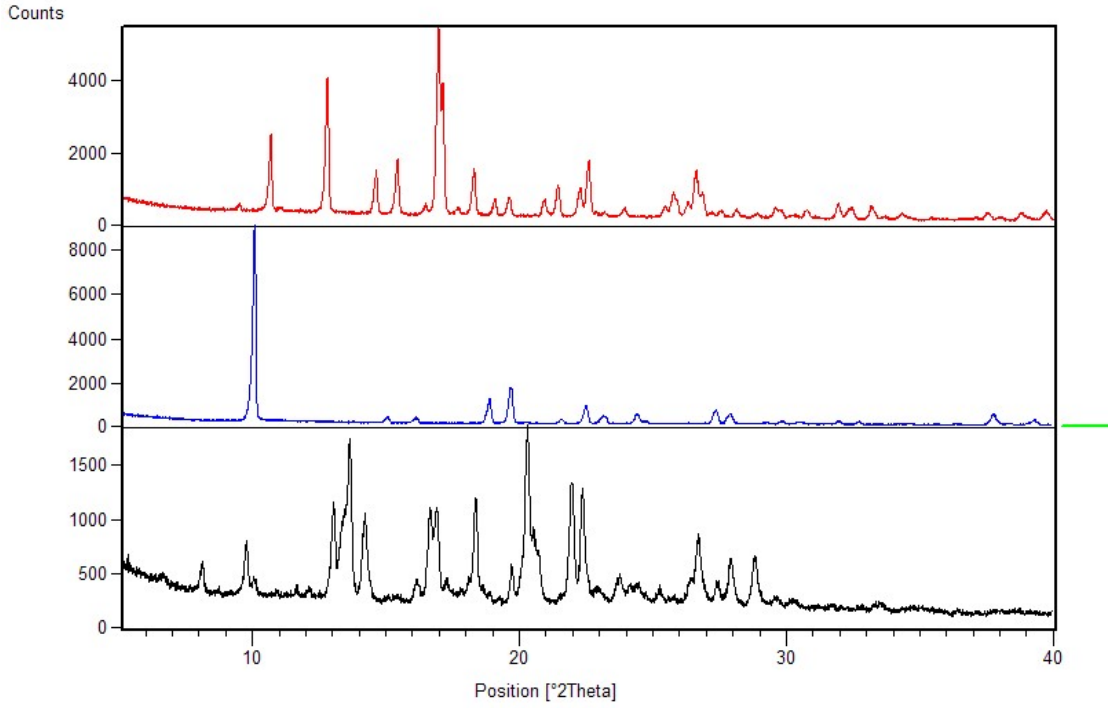

Figure S22. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) 2,7-dihydroxynaphthalene and (c) product of LAG reaction of **pro** and 2,7-dihydroxynaphthalene in 1:1 stoichiometric ratio. Cocrystal forms.

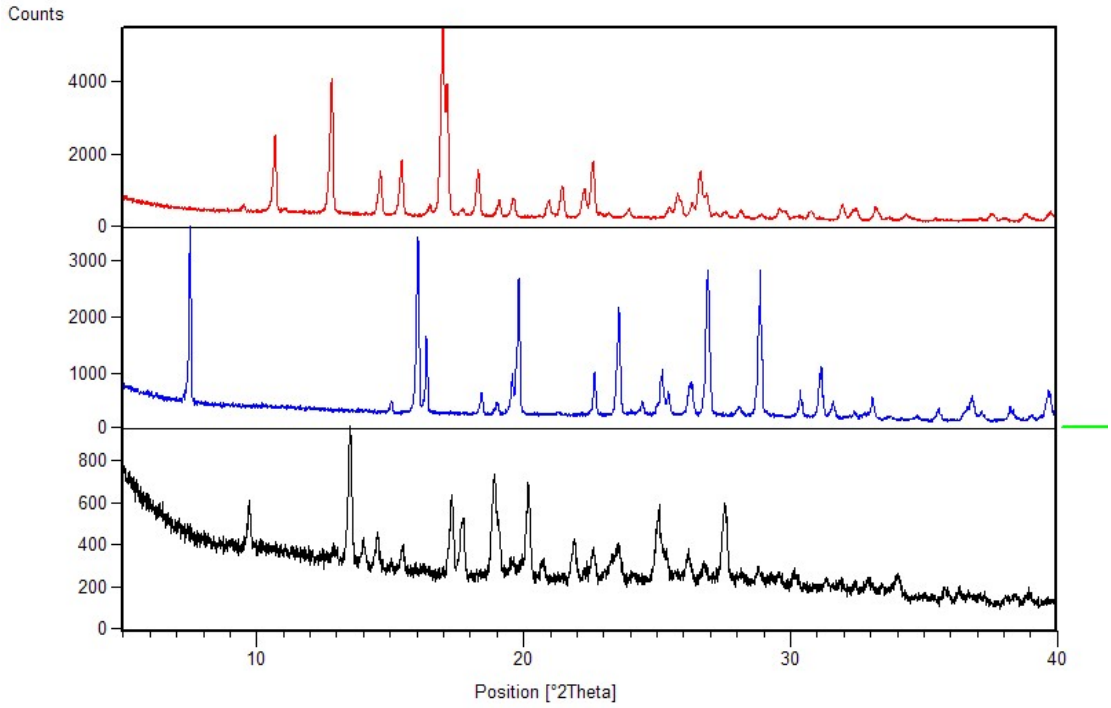

Figure S23. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) gentisic acid and (c) product of LAG reaction of **pro** and gentisic acid in 1:1 stoichiometric ratio. Cocrystal forms.

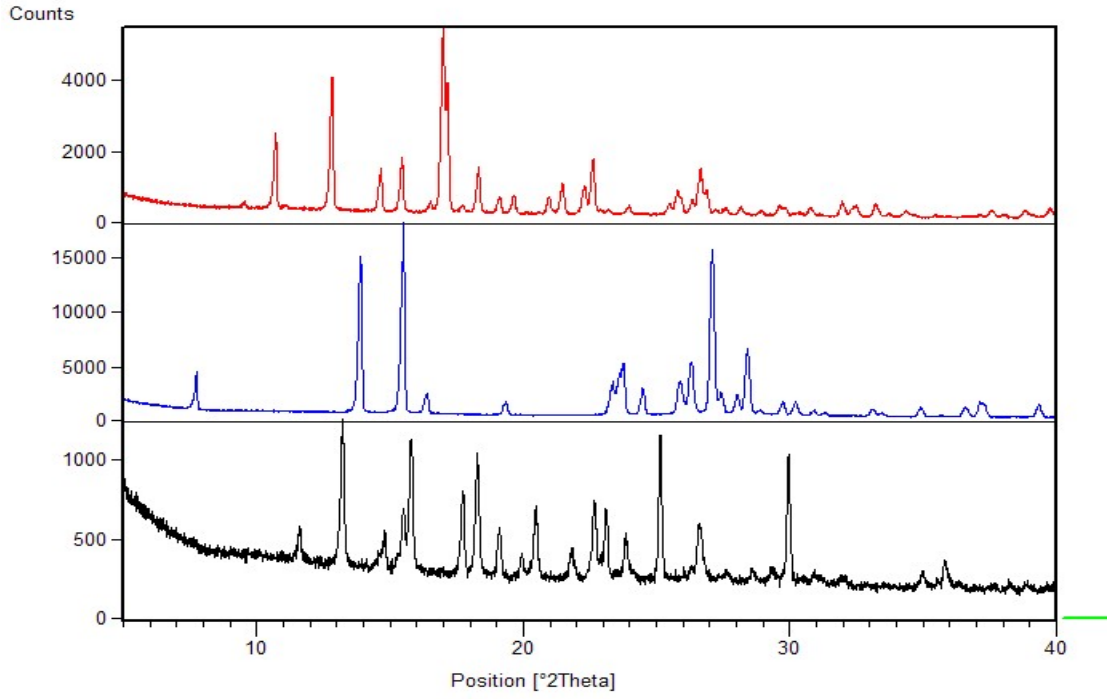

Figure S24. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) phthalimide and (c) product of LAG reaction of **pro** and pthalimide in 1:1 stoichiometric ratio. Cocrystal forms.

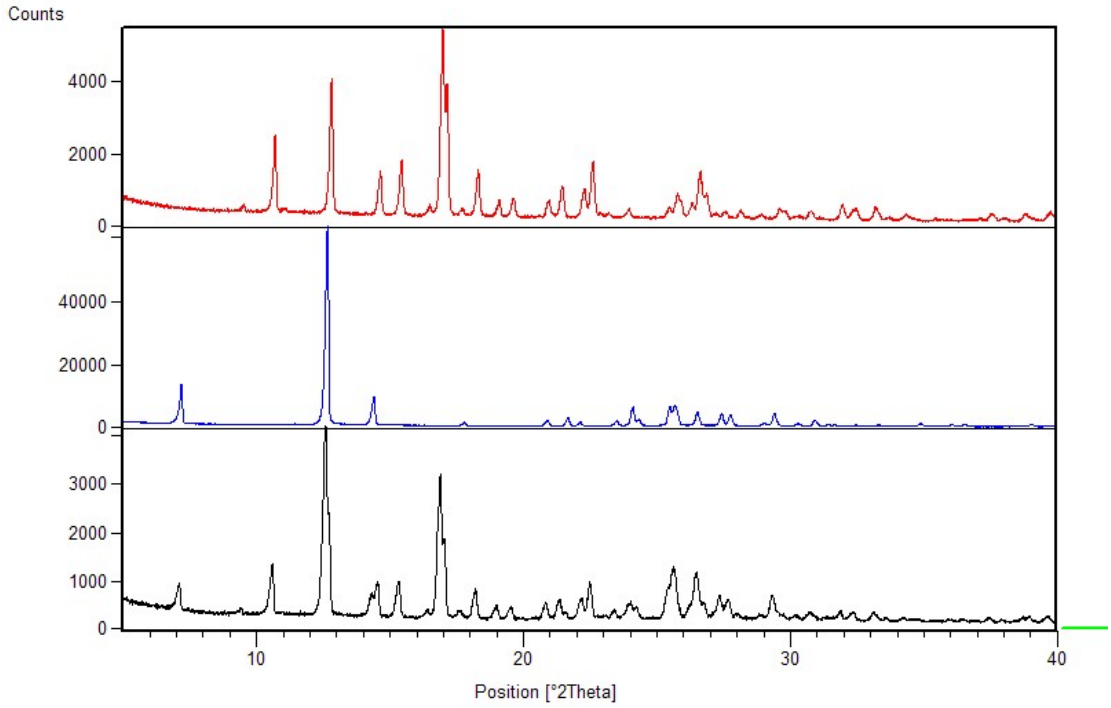

Figure S25. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) theophylline and (c) product of LAG reaction of **pro** and theophylline in 1:1 stoichiometric ratio. Cocrystal does not form.

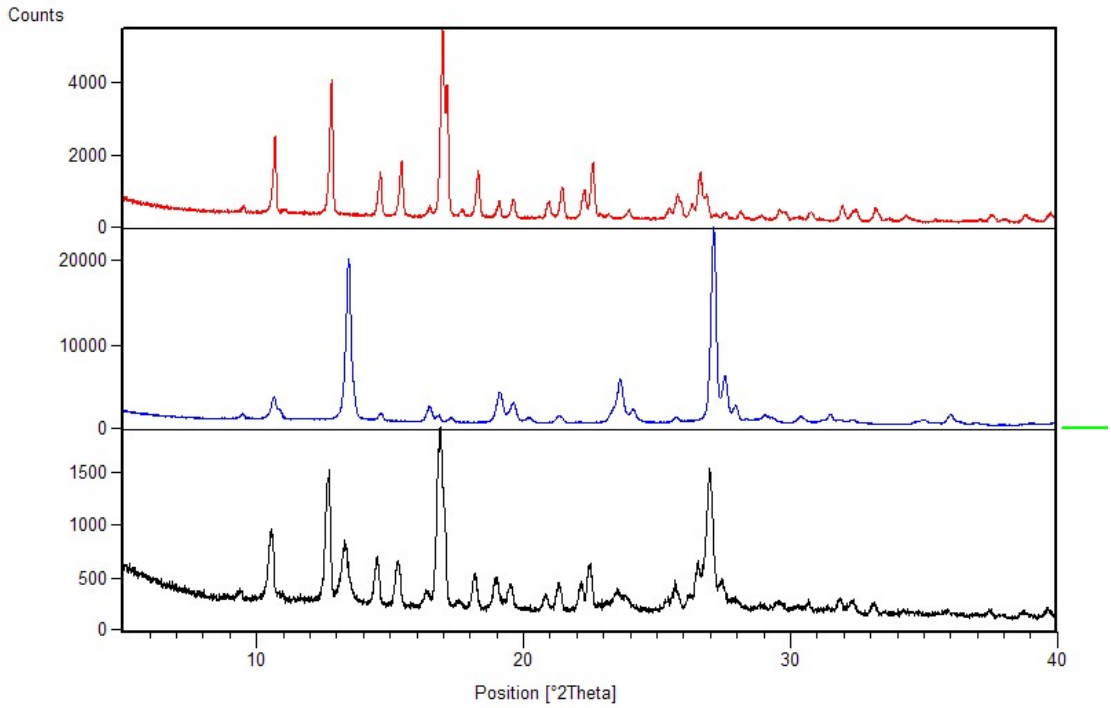

Figure S26. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) theobromine and (c) product of LAG reaction of **pro** and theobromine in 1:1 stoichiometric ratio. Cocrystal does not form.

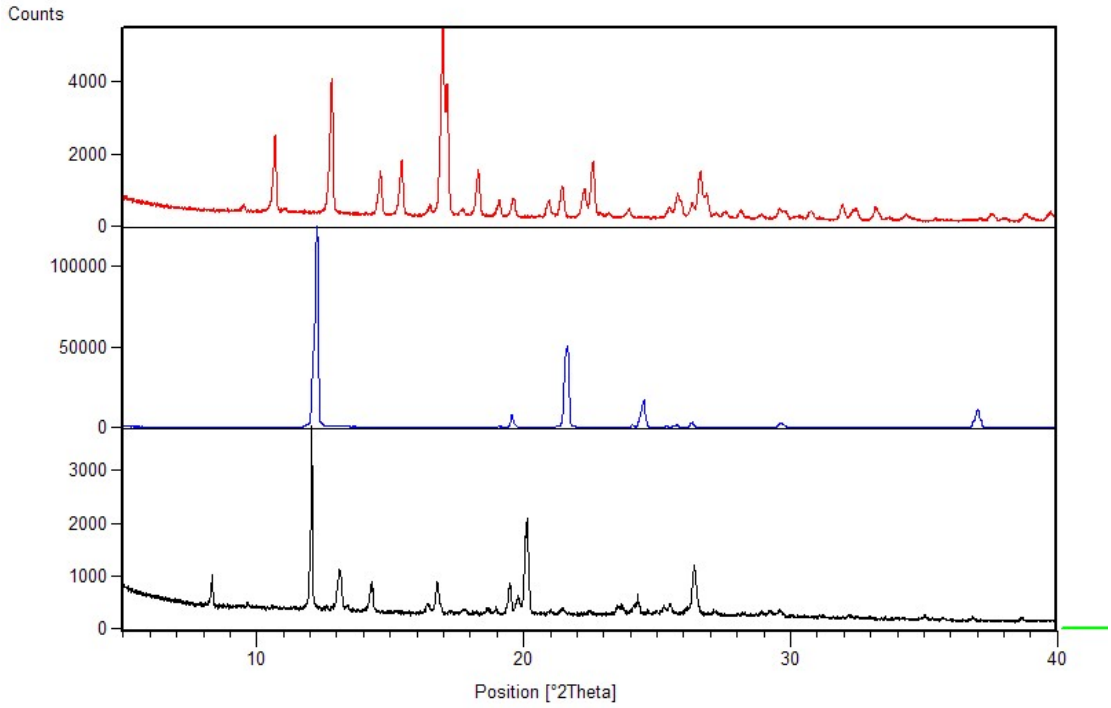

Figure S27. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) naphthalene and (c) product of LAG reaction of **pro** and naphthalene in 1:1 stoichiometric ratio. Cocrystal forms.

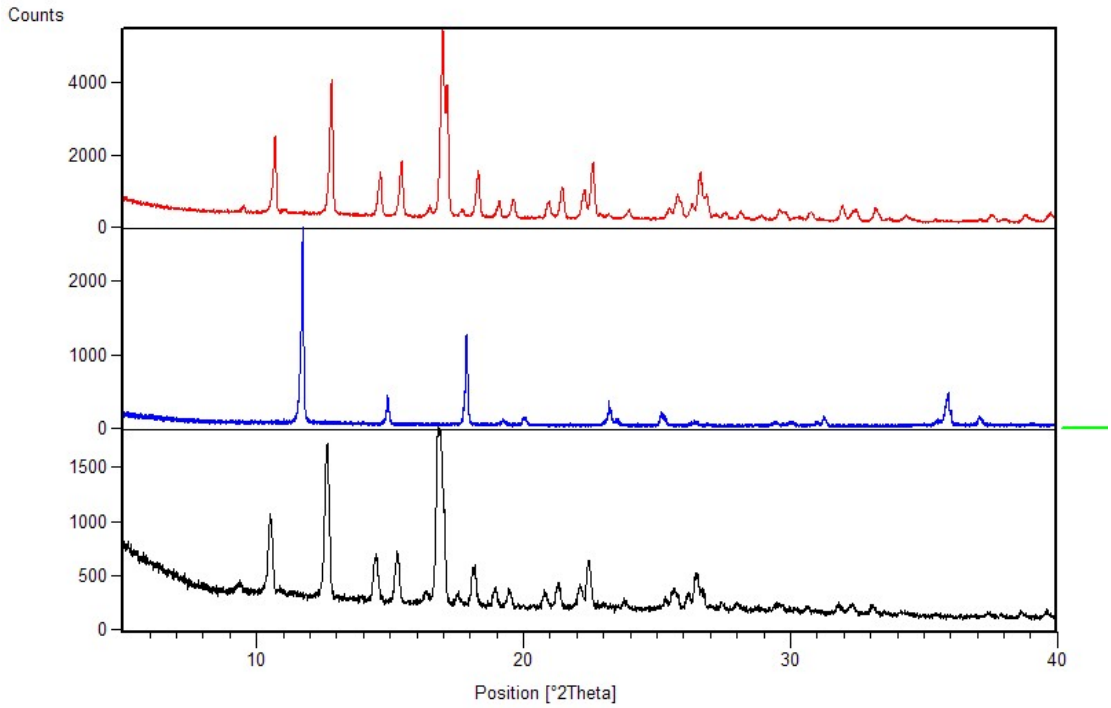

Figure S28. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) octafluoronaphthalene and (c) product of LAG reaction of **pro** and octafluoronaphthalene in 1:1 stoichiometric ratio. Cocrystal does not form, and the disappearance of octafluoronaphthalene reflections is explained by amorphisation.

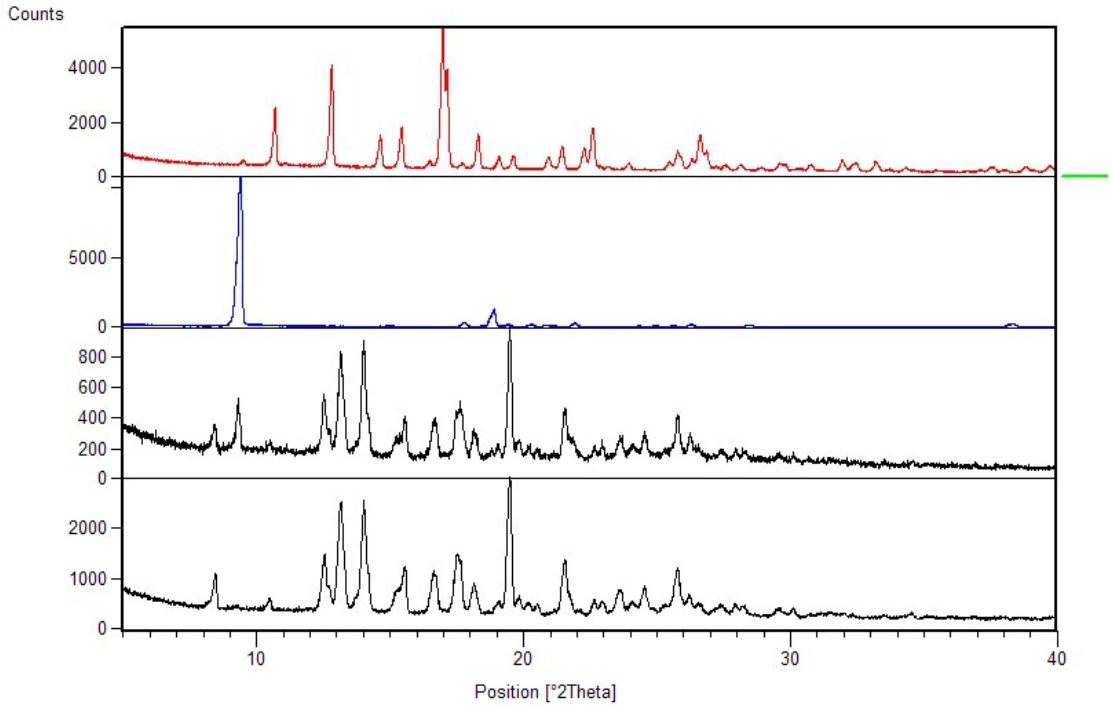

Figure S29. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) phenanthrene; (c) product of LAG reaction of **pro** and phenanthrene in 1:1 stoichiometric ratio and (d) (c) product of LAG reaction of **pro** and phenanthrene in 3:1 stoichiometric ratio. Cocrystal of composition (**pro**)$_3$·(phenanthrene) forms.

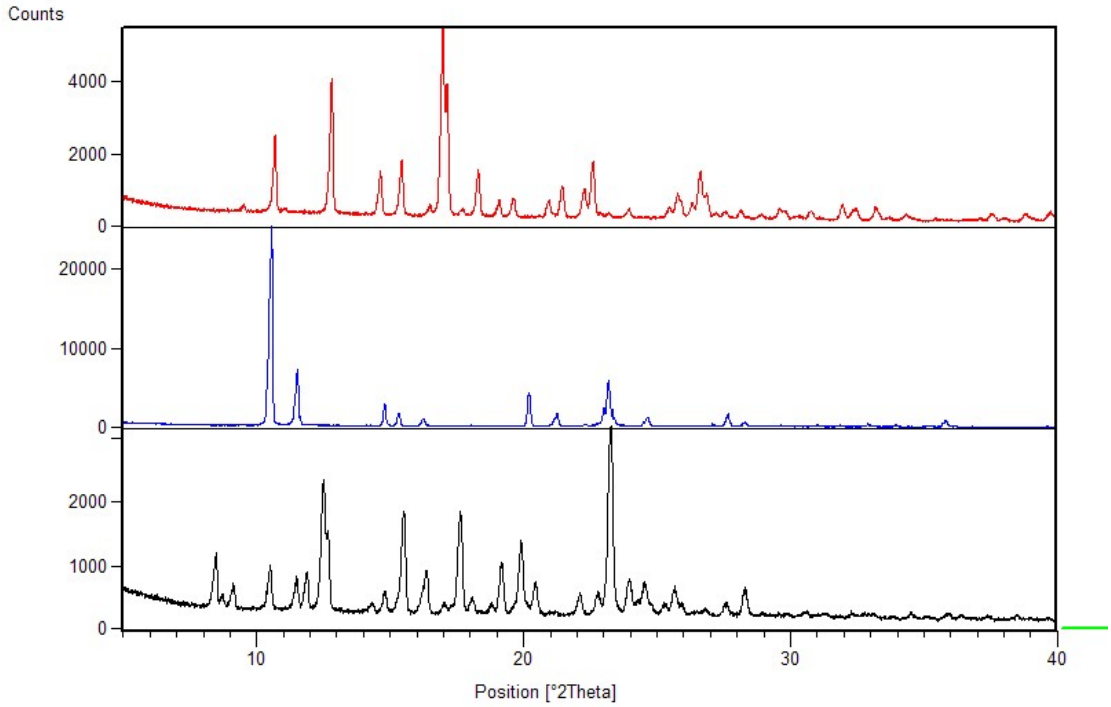

Figure S30. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) pyrene and (c) product of LAG reaction of **pro** and pyrene in 1:1 stoichiometric ratio. Cocrystal forms.

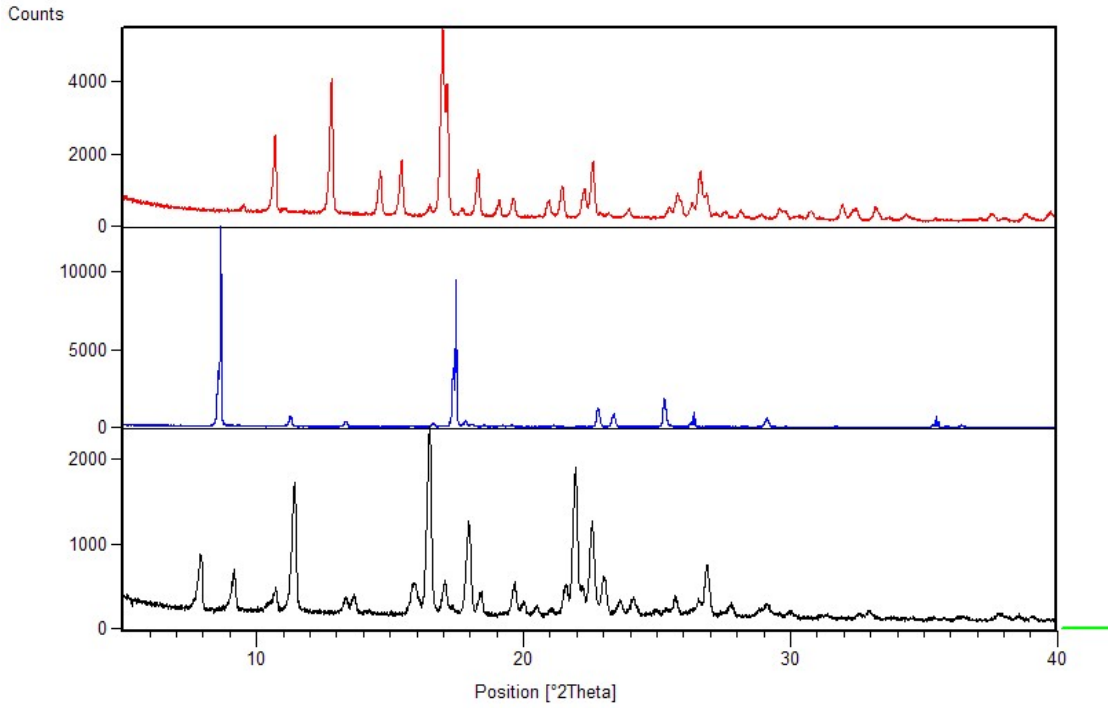

Figure S31. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) perylene and (c) product of LAG reaction of **pro** and perylene in 1:1 stoichiometric ratio. Cocrystal forms.

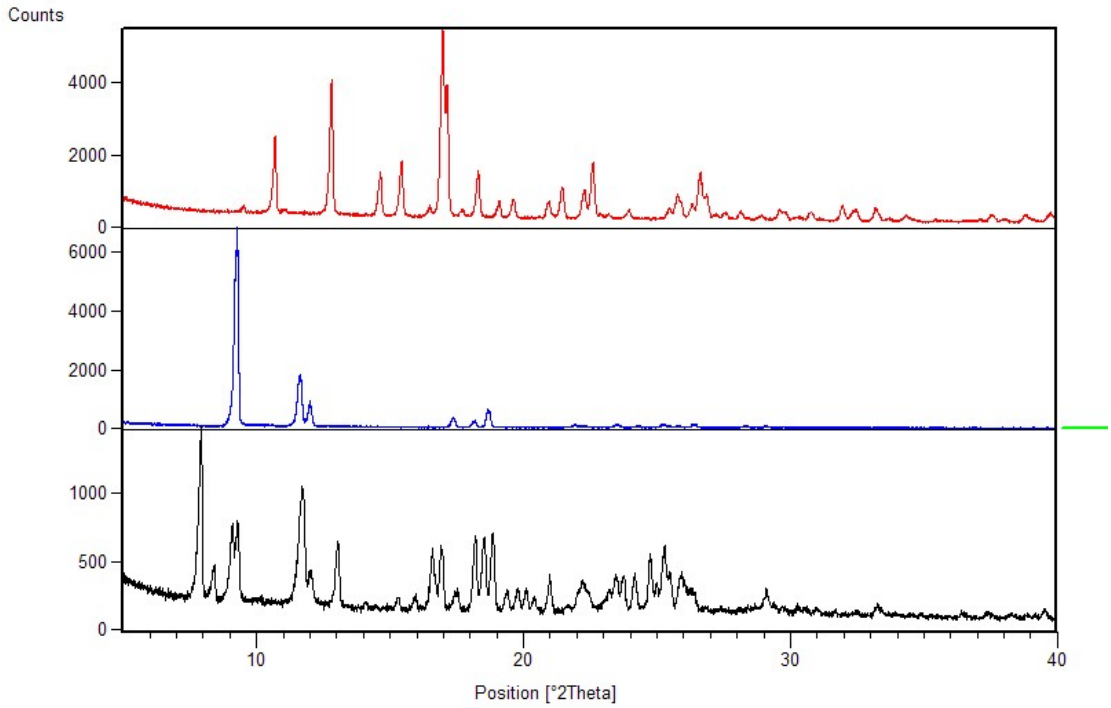

Figure S32. Overlay of PXRD patterns (top to bottom) for: (a) **pro**; (b) coronene and (c) product of LAG reaction of **pro** and coronene in 1:1 stoichiometric ratio. Cocrystal forms.

S3. FTIR-ATR solid-state spectroscopy

Infrared spectra were recorded at 4 cm^{-1} resolution on a Perkin Elmer Spectrum One FTIR spectrometer equipped with a diamond attenuated total reflectance (ATR) accessory. Potential cocrystals were compared with co-added spectra of the steroid and cococrystal former. Examples of positively confirmed cocrystals are depicted below along with several unsuccessful cocrystallisation attempts.

In cases where there is potential for hydrogen bonding, one would generally expect changes in the position of carbonyl or hydroxyl bands. When cocrystal formation arises due to α···π stacking, the region of interest for differentiating a cocrystal from a mixture of pure-component crystals is the in- and out of plane deformation (fingerprint) region.

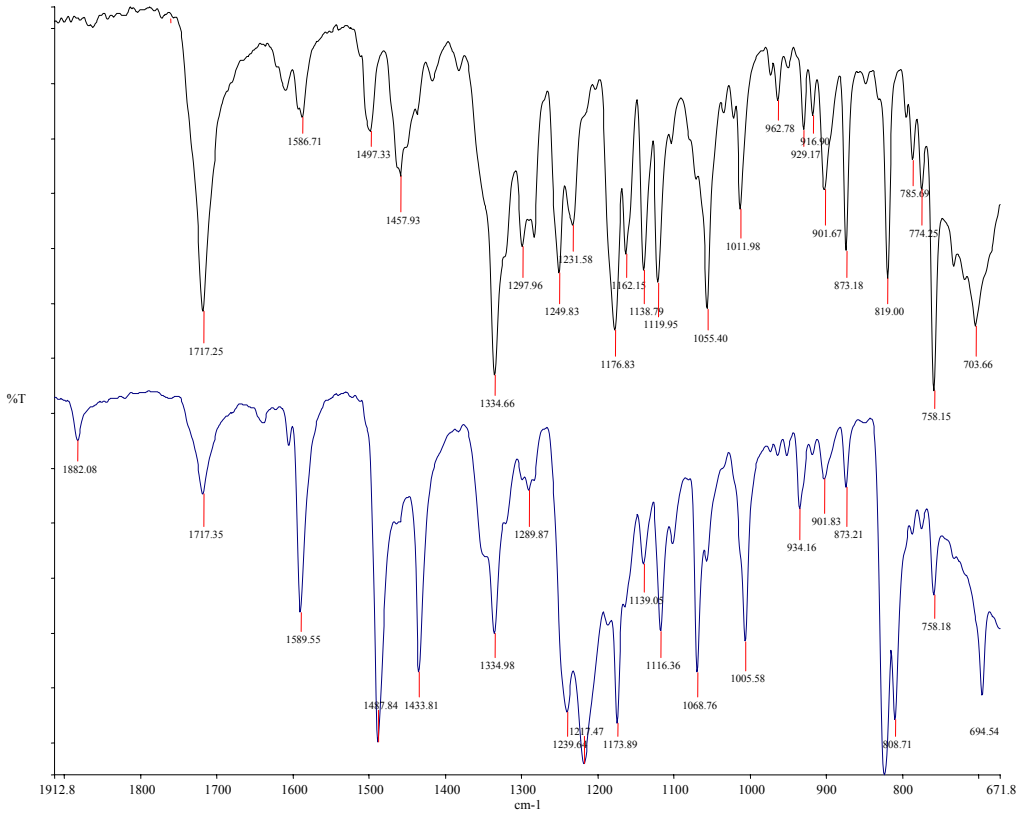

Figure S33. Progesterone: 4-bromophenol cocrystal (bottom) vs spectral addition (top). Cocrystal formed.

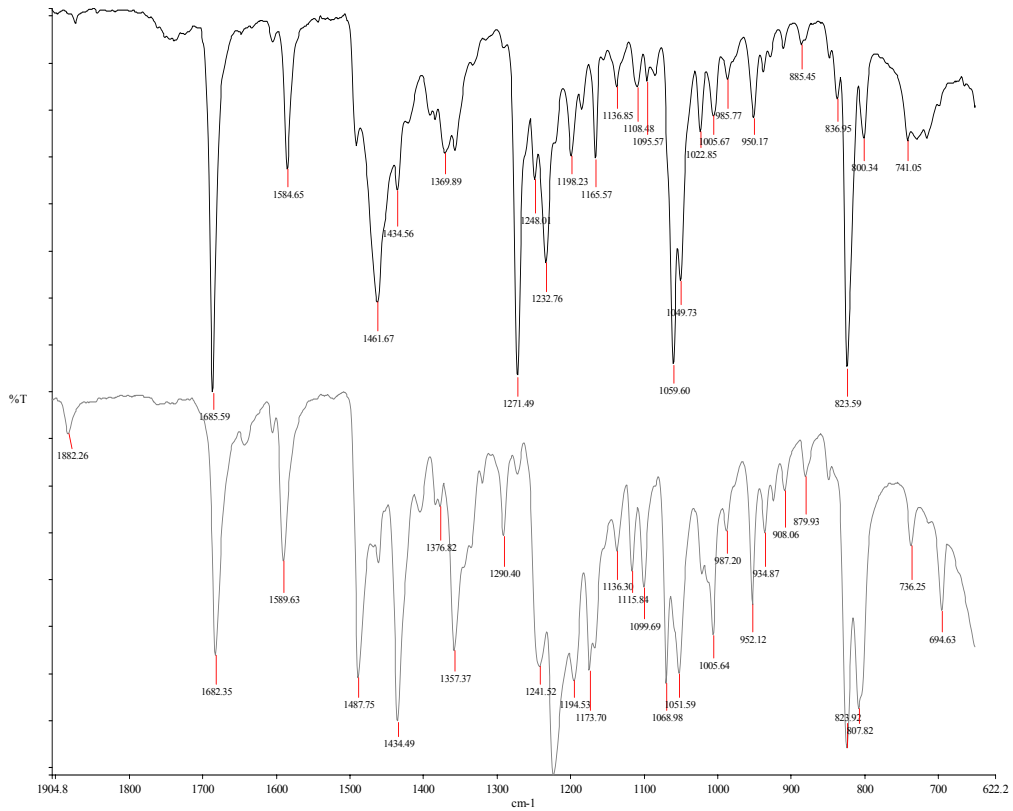

Figure S34. Pregnenolone: 4-bromophenol cocrystal (top) vs spectral addition (bottom). Cocrystal formed.

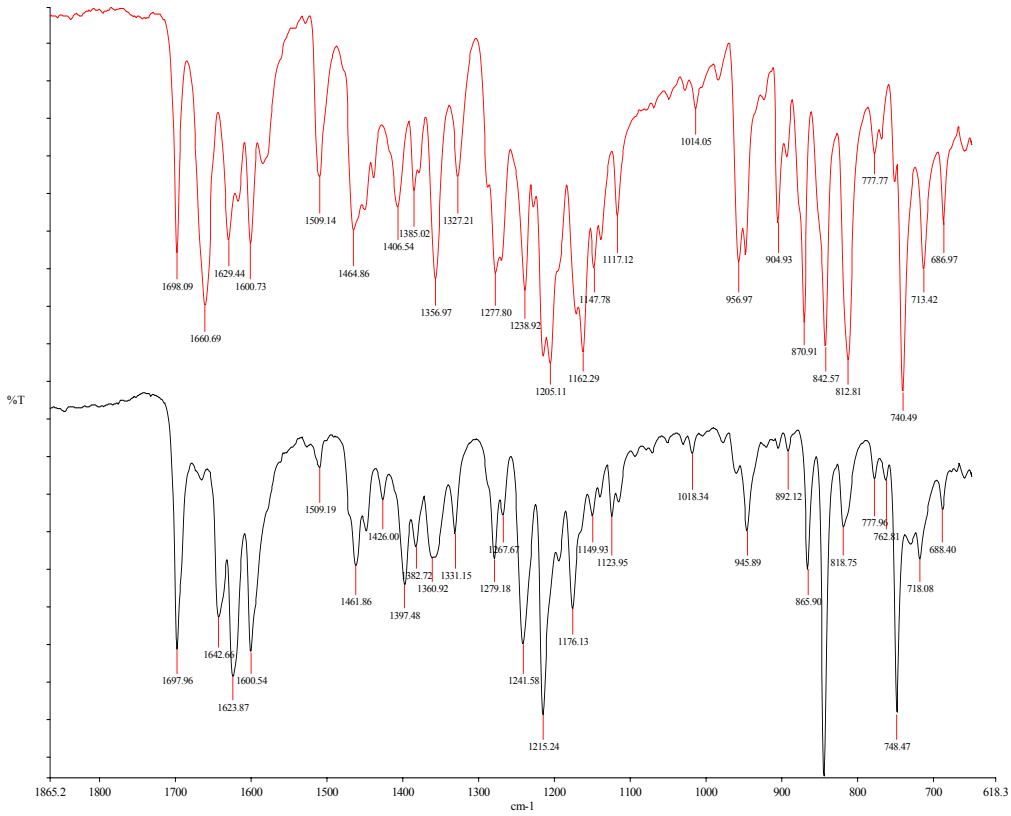

Figure S35. Progesterone:2-naphthol cocrystal (bottom) vs spectral addition (top). Cocrystal formed.

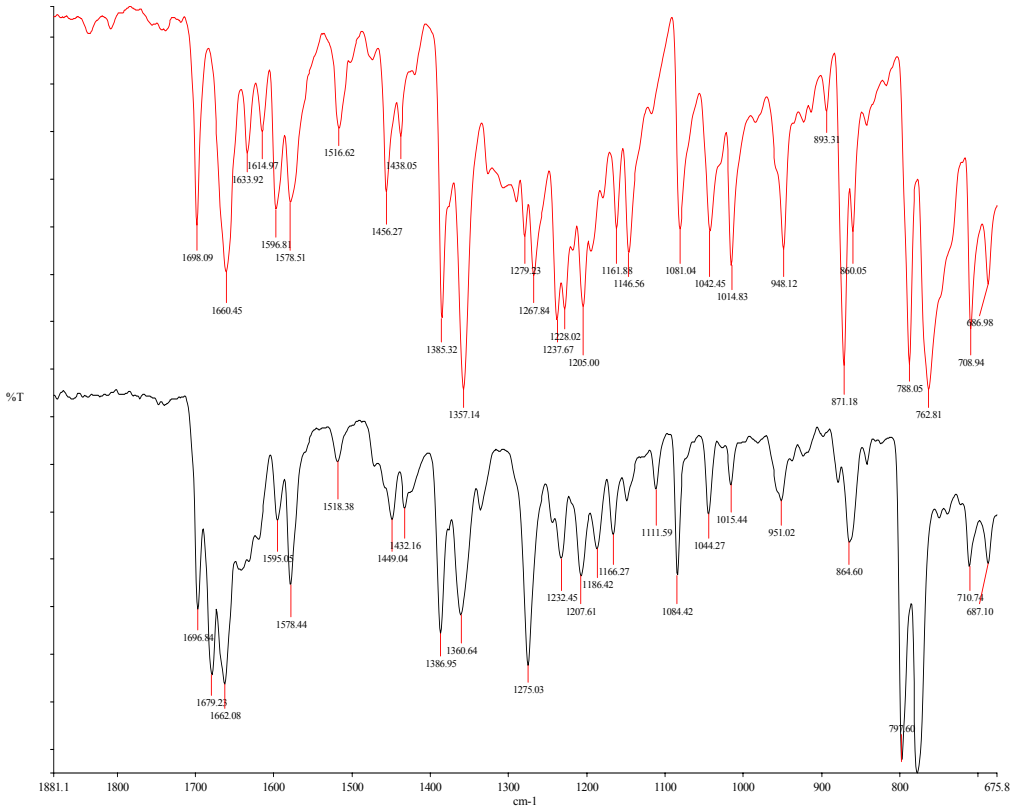

Figure S36. Progesterone: 1-naphthol cocrystal (bottom) vs spectral addition (top). Cocrystal formed.

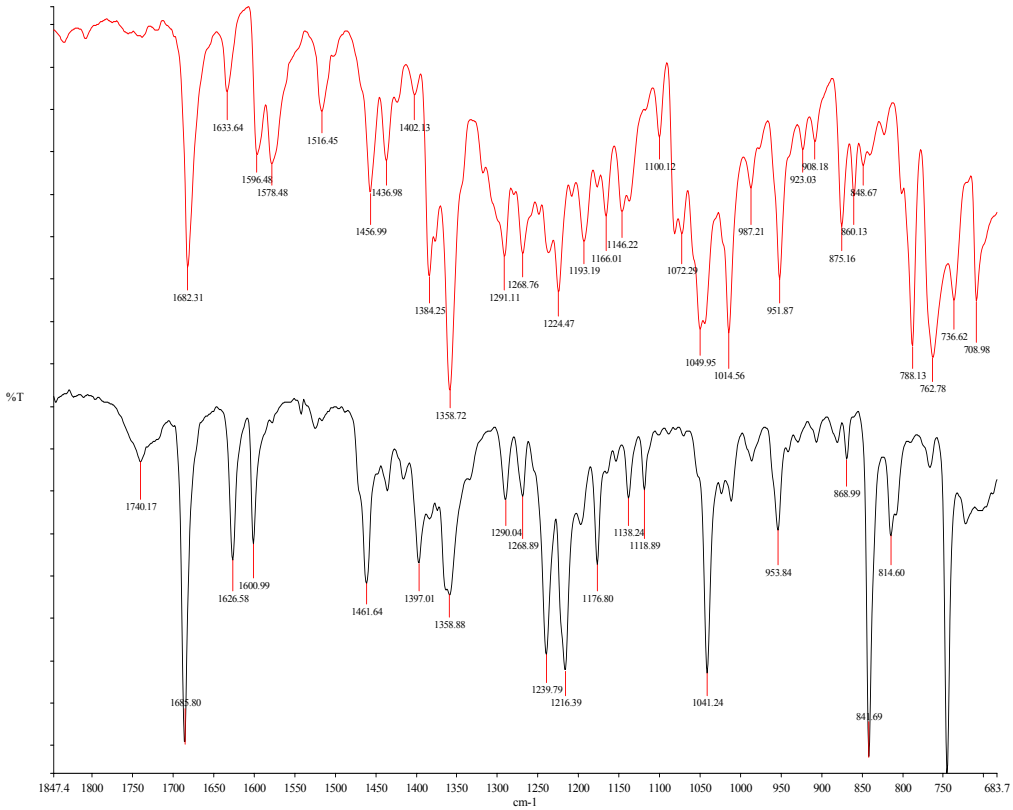

Figure S37. Pregnenolone: 1-naphthol cocrystal (bottom) vs co-addition (top). Cocrystal formed.

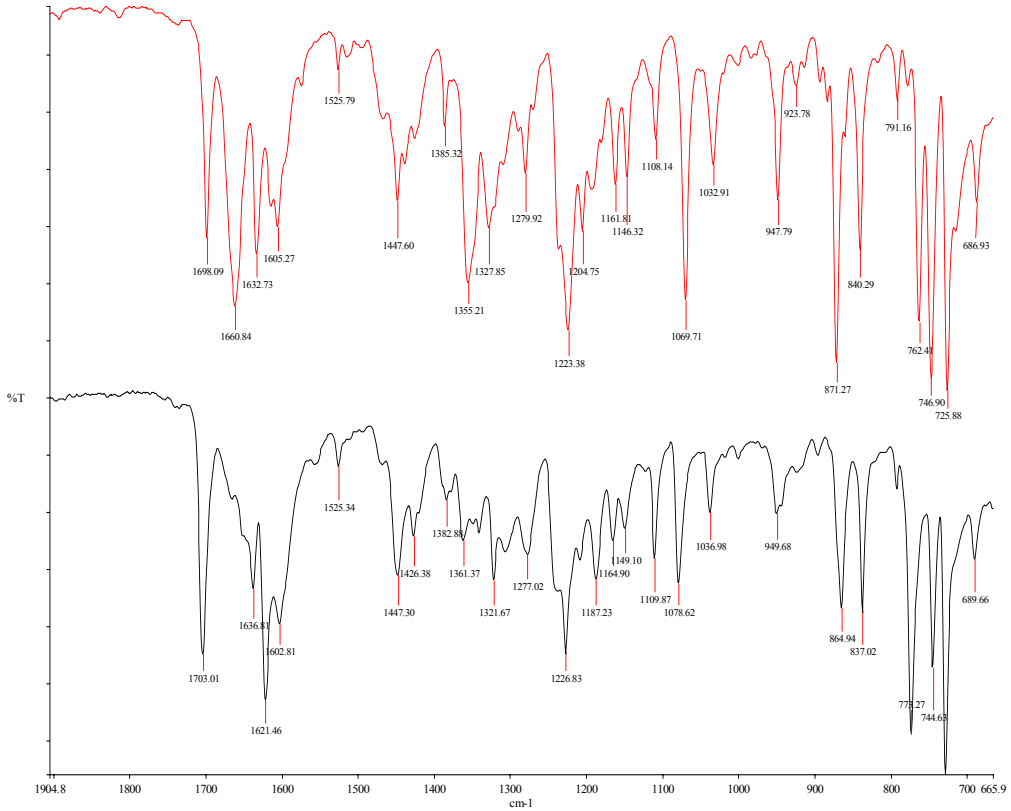

Figure S38. Progesterone: 9-phenanthrol cocrystal (bottom) vs co-addition (top). Cocrystal formed.

S4. Electrostatic surface potential calculations

In this section we present a comparative analysis of the electrostatic surface potentials for local conformational minima of **pro**, **pre**, **bes** and **est** not presented in the manuscript. We also report the ESP for androsterone and androsta-14-diene-3,17-dione. The potentials were modelled from the distributed multipole expansion[5] up to hexadecapole of the B3LYP/6-31G(d,p) charge density in isolation projected on the molecular surface defined by twice the van der Waals radii[6] (2×vdW surface). For the steroids with conformationally labile hydroxyl groups, the ESP was calculated for selected, low-energy gas-phase optimised conformational minima. The maximum and minimum of the electrostatic potential on this surface are indicated under each plot.

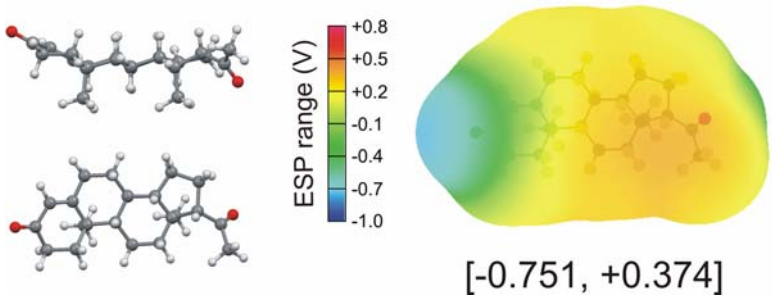

Figure S39. Side and face-on views and calculated ESP for the gas-phase optimised **pro** molecule.

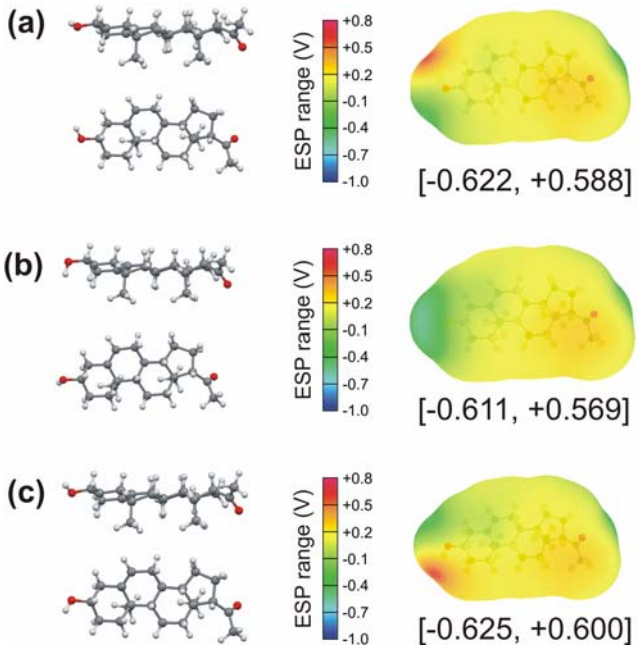

Figure S40. Side and face-on views and calculated ESP for three gas-phase optimised conformations of **pre**, differing in the orientation of the -OH group. Relative conformational energies: (a) 0 kJ mol^{-1}; (b) 0.29 kJ mol^{-1} and (c) 0.63 kJ mol^{-1}. In Figure 2 of the manuscript we show conformation (c) as representative.

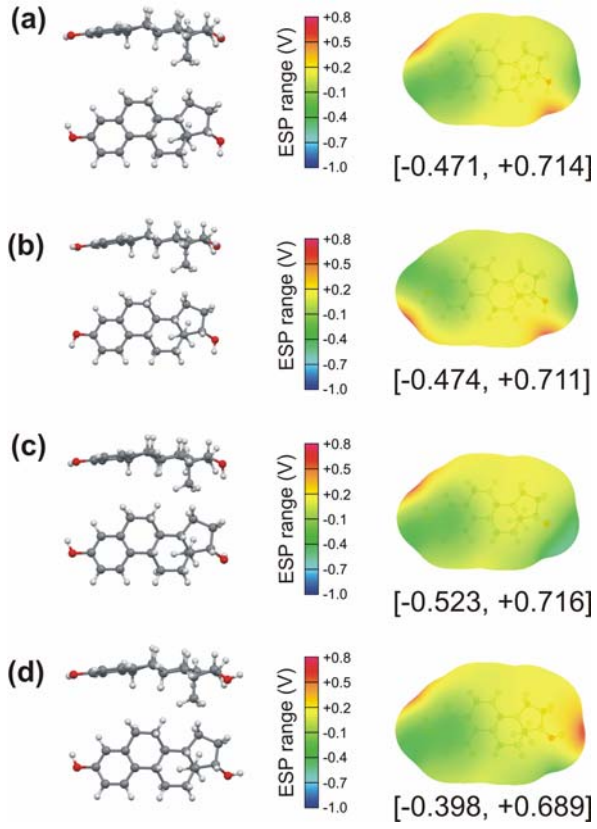

Figure S41. Side and face-on views and calculated ESP for four gas-phase optimised conformations of **bes** differing in the orientation of the hydroxyl groups. Conformations (a) and (b) have the hydroxyl group at position 17 in the energetically lower configuration, whilst conformations (c) and (d) contain the hydroxyl group at position 3 in the energetically lower configuration. Relative conformational energies: (a) 0 kJ mol^{-1}; (b) 0.06 kJ mol^{-1}; (c) 0.06 kJ mol^{-1} and (d) 1.68 kJ mol^{-1}. In Figure 2 of the manuscript we show conformation (b) as representative.

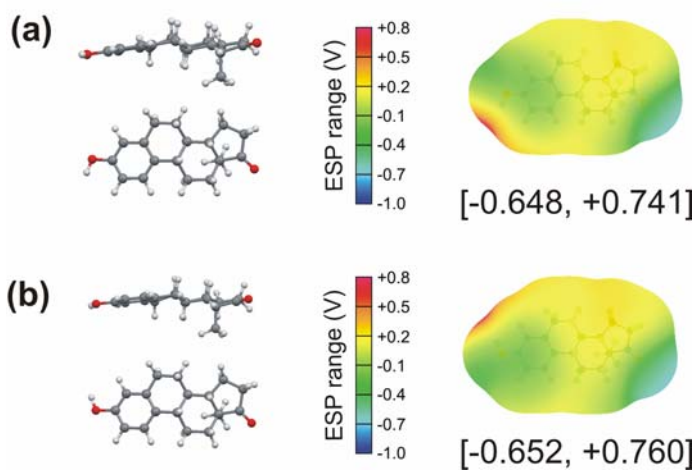

Figure S42. Side and face-on views and calculated ESP for two gas-phase optimised conformations of **est**, differing in the orientation of the hydroxyl group. Relative conformational energies: (a) 0 kJ mol^{-1} and (b) 0.25 kJ mol^{-1}. In Figure 2 of the manuscript we show conformation (a) as representative.

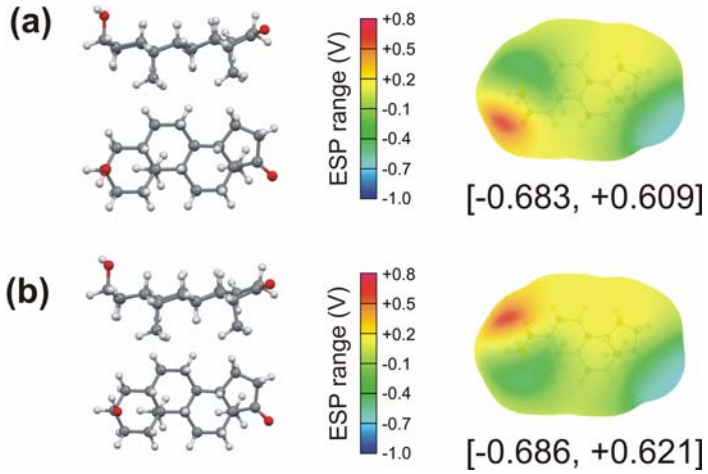

Figure S43. Side and face-on views and calculated ESP for two gas-phase optimised conformational minima of androsterone, differing in the orientation of the hydroxyl group. The difference in conformational energies is within the optimisation tolerance.

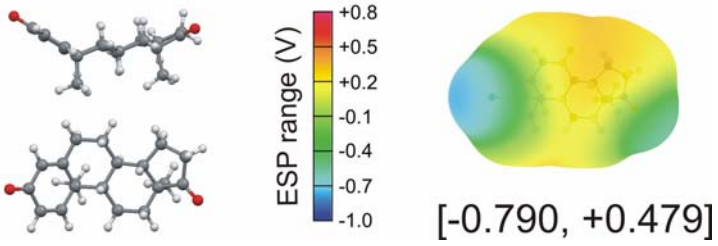

Figure S44. Side and face-on views and calculated ESP for the gas-phase optimised conformation of androsta-14-diene-3-17dione.

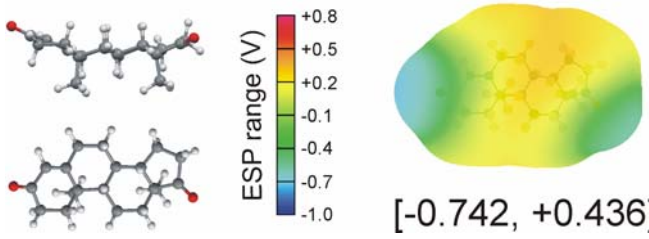

Figure S45. Side and face-on views and calculated ESP for the gas-phase optimised conformation of androstenedione.

S5. Crystal structure prediction

For validation purposes, the computational procedure used for cocrystal structure prediction was also applied for the prediction of crystal structures of the pure components **pro**, **13** and **22**. The sum of the lattice energies of the pure-component crystals were

contrasted with the lattice energies of the cocrystals to elucidate the thermodynamic grounds for cocrystallisation[7] and attempt to predict the stoichiometry of the (**pro**)·(**22**) asymmetric unit.[8,9] The lattice energy vs. density plots (lattice energy landscapes) for all predicted structures are summarised in Figure S46.

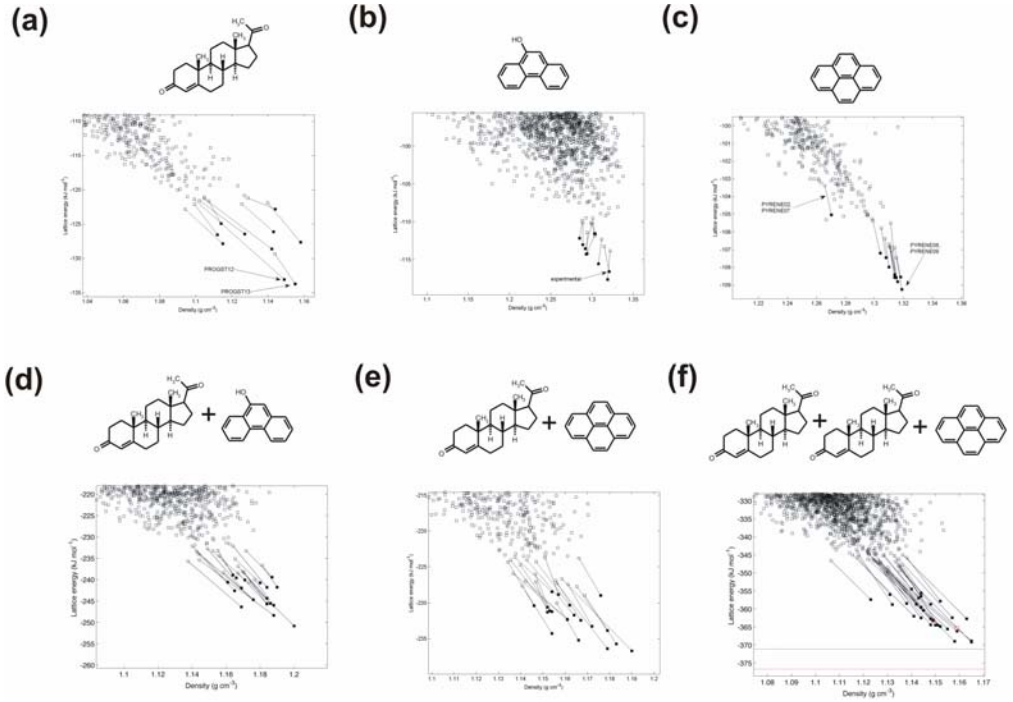

Figure S46. Predicted lattice energy vs. density landscapes for: (a) **pro**; (b) **13**; (c) **22**; (d) (**pro**)(**13**); (e) hypothetical (**pro**)(**22**) and (f) (**pro**)$_2$(**22**). For the cocrystal predictions (d-f), the horizontal black and red lines denote the sum of the lattice energies of the least and most stable polymorphs of the component molecules, respectively, minimised with the same flexible-molecule computational model, that was also used to minimise the experimental crystal structure (shown with red square). For (**pro**)$_2$(**22**) the search structure (also shown with a red square) that resembled most closely the minimised experimental cocrystal differs by 0.49 Å in the 20-molecule coordination sphere and is 2 kJ mol^{-1} less stable.

S5.1. Crystal structure prediction of pure components

Our calculations confirmed our earlier results[10] that the two known chiral **pro** crystal structures differ in energy by less than 1 kJ mol^{-1} and correspond to the two most stable packing arrangements. The two ambient pressure **22** polymorphs[11,12] interconvert at 110 K. Upon minimisation, these similar[13] structures coalesce into the same lattice energy minimum that is 4 kJ mol^{-1} less stable than the global minimum, which corresponds to the high-pressure polymorph when lattice-energy minimised under ambient pressure.[14]

A crystal structure of **13** has not been reported and our efforts to grow single crystals for X-ray diffraction were unsuccessful. Crystal structure solution from PXRD of a commercial sample failed at the indexing stage. However, the PXRD pattern was in good agreement with the simulated pattern for a ca. 4:1 mixture of two low-energy, predicted crystal structures (predicted polymorph A: -116.6 kJ mol^{-1} compared with -111.6 kJ mol^{-1} for predicted polymorph B) that are both within approximately 5 kJ mol^{-1} from the global minimum. Hot-stage microscopy and DSC experiments with a commercial sample of **13** revealed two endothermic events, consistent with melting of a two-component polymorph mixture.

The commercial sample was slurried in 1,1,1-trifluoroethanol to facilitate conversion to the most stable predicted polymorph (A). The PXRD pattern of the resulting solid showed a close resemblance to the simulated powder pattern of the predicted A crystal structure and a Pawley refinement starting from the predicted cell parameters and space group ($P2_12_12_1$) yielded a satisfactory fit (profile χ^2 = 3.40) in DASH.[15] Full structure solution starting from the measured PXRD pattern of the slurry product using DASH resulted in virtually the same structure as the predicted one. Rietveld refinement of this structure was carried out using GSAS,[16] with the EXPGUI interface.[17] Bond lengths and angles were restrained during the refinement to values taken from the predicted structure, and planar group restraints were applied for the aromatic rings. A common isotropic displacement parameter was used for all non-hydrogen atoms, while the displacement parameters of all hydrogen atoms were constrained to be twice the common displacement parameter value. Peak profiles were modelled by an asymmetry corrected pseudo-Voigt function,[4] and a fourteen parameter Chebyshev polynomial of the 1st kind was used to describe the background. The Rietveld refinement converged readily to yield acceptable figures

of merit (R_p = 0.0313, R_{wp} = 0.0399) and a reasonable structure. The torsion angle of the hydroxyl group cannot be determined reliably from the PXRD data, so the H atom of this group was restrained to be coplanar with the aromatic system.

The predicted less stable polymorph B was also obtained by thermal cycling in DSC. However, partial thermal decomposition prevented the collection of a PXRD pattern suitable for Rietveld refinement. Nevertheless, the measured PXRD pattern was in qualitative agreement with the simulated PXRD pattern of the predicted structure (Figure S47). A detailed, joint theoretical-experimental investigation of **13** will be reported in a forthcoming publication.

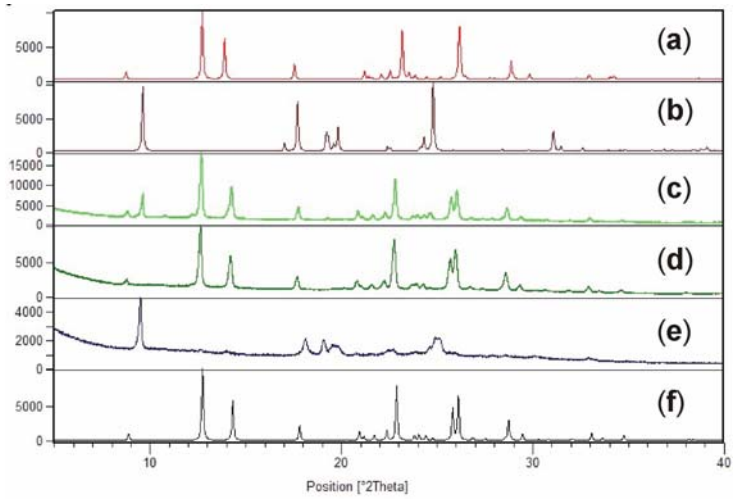

Figure S47. Overlay of PXRD patterns (from top to bottom): (a) simulated for the predicted lower-energy polymorph A of **13**; (b) simulated for the predicted higher-energy polymorph B; (c) commercial sample of **13**; (d) **13** after 1,1,1-trifluoroethanol slurry; (e) **13** after thermal DSC cycling and (f) simulated for the crystal structure obtained by Rietveld refinement of the predicted low-energy polymorph A against the PXRD pattern obtained after overnight slurrying of the commercial sample.

S5.2 Cocrystal structure prediction

The (**pro**)·(**13**) lattice energy landscape showed a variety of packing motifs, with the hydroxyl donor bonded to either **pro** carbonyl with equal frequency. The three most stable predicted cocrystal structures are additionally stabilised by α···π stacking and the experimental structure corresponds to the densest and most stable predicted cocrystal. The experimental structure is also the only that has a small thermodynamic advantage over the most stable polymorphs of **pro** and **13** crystallising independently. The most stable predicted cocrystal structure that lacks α···π stacking is ca. 5 kJ mol^{-1} less stable than the global minimum. The predicted lattice energy landscape clearly shows that α···π stacking not only does not disrupt close packing and hydrogen bonding, but also provides the extra stabilisation necessary for cocrystallisation.

To the best of our knowledge, the generated lattice energy landscape for (**pro**)$_2$·(**22**) is the most demanding crystal structure prediction reported to date, both in terms of molecular size and asymmetric unit complexity. Whilst 1:1 cocrystal structure prediction is now computationally tractable, a crystal structure that resembled the minimised experimental (**pro**)$_2$·(**22**) cocrystal was only found once after approximately one million lattice energy minimisations. Such calculations push the boundaries of what is possible in cocrystal structure prediction and, despite the search lacking completeness, the predictions are informative in analysing packing motifs. The lattice energy landscape shows limited packing diversity with most predicted structures exhibiting clear α···π stacking on one (1:1 stoichiometry) or both sides (2:1 stoichiometry) of **22**. However, there are no predicted structures that are energetically competitive with the pure component crystals and no obvious thermodynamic benefit was observed between the hypothetic 1:1 (Figure S46e) and the experimentally observed 1:2 cocrystal stoichiometry. The lattice energy differences involved are small and likely to change sign depending on the model for the intermolecular forces and entropy effects.[18] Many of the predicted (**pro**)$_2$·(**22**) sandwich structures differ only in the rotation of **22** on its aromatic plane (Figure S48). Such rotations with respect to the encasing **pro** molecules are likely to be labile and to correspond to low-frequency librations contributing to entropic stabilisation. This view is supported by the thermal ellipsoids of **22** in the experimental cocrystal, which become laterally elongated towards the periphery of the molecule (Figure S49). Such elongation is particularly notable in comparison with the analogous structure of (**pro**)$_2$·(**12**), where the pyrene system is anchored by O-H···O hydrogen bonding.

Hence, our static lattice energy results are informative in providing all thermodynamically plausible molecular arrangements in cocrystal structures and establishing the dominance of α···π

stacking in cocrystals of **pro** with both **13** and **22.** Despite this success, the confident prediction of cocrystallisation remains elusive for current methodologies as it would require computation of Gibbs free energies to accuracy of the order of 1 kJ mol^{-1} and consideration of crystallisation kinetics.

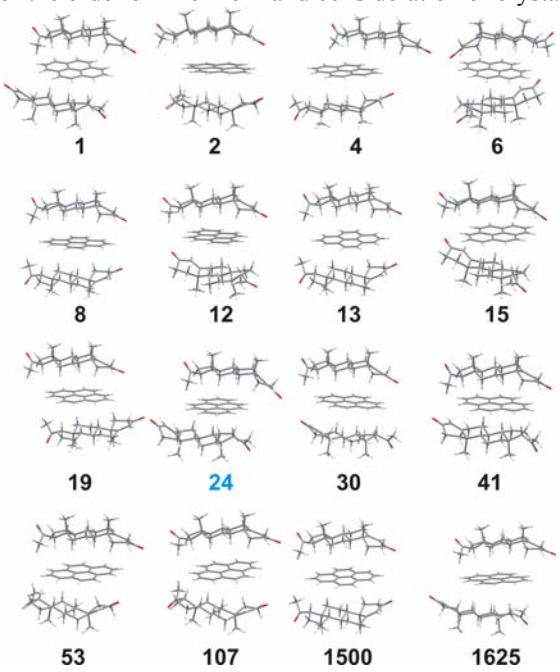

Figure S48. The "sandwich" trimer structures found among the 29 lowest energy predicted structures for the cocrystal (**pro**)$_2$·(**22**) that span a lattice energy range of 15 kJ mol^{-1} from the global minimum. The predicted structure that corresponds to the observed structure is highlighted in blue and is approximately 6 kJ mol^{-1} less stable than the global minimum (structure 6 in Figure).

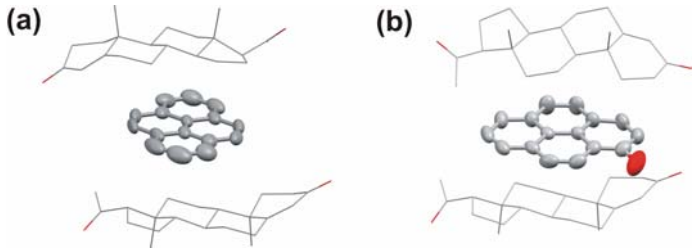

Figure S49. The trimer "sandwich" motifs in cocrystals (a) (**pro**)$_2$·(**22**) and (**pro**)$_2$·(**12**). For clarity, the pro molecules are shown in wireframe representation and the arenes are shown using 50% probability ellipsoids.

S6. Aromatic groups in the binding sites of progesterone and estrogen receptors

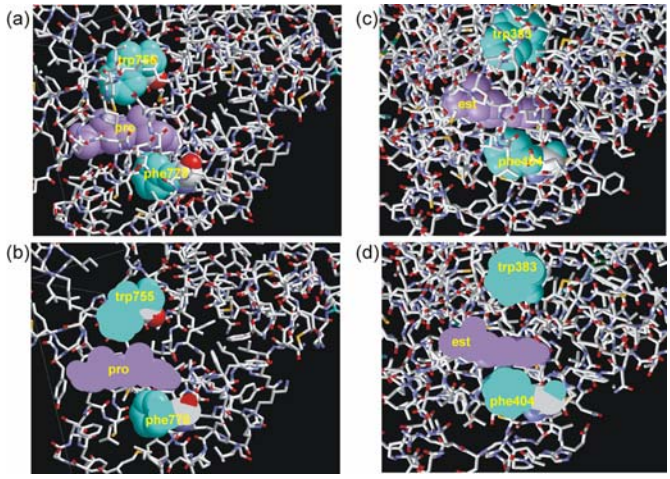

Figure S50. Fragment of crystal structures of (a) and (b) progesterone receptor with the **pro** molecule (purple) in the binding site;[19] (c) and (d) estrogen receptor with the **est** molecule (purple) in the active site.[20] For clarity, the aromatic Phe and Trp residues in the vicinity of the ligand are coloured blue and are shown in space-filling mode.

S7. References

1. Bhatt PM, Desiraju GR (2008) Co-crystal formation and the determination of absolute configuration *CrystEngComm* 10:1747-1749.
2. Lancaster RW, Karamertzanis PG, Hulme AT, Tocher DA, Lewis TC, Price SL (2007) The polymorphism of progesterone: Stabilization of a 'disappearing' polymorph by co-crystallization *J. Pharm. Sci.* 96:3419-3431.
3. Perrin M, Rantsordas S, Thozet A (1978) 4-Bromophenol C6H5BrO *Cryst. Struct. Commun.* **7**:59-62.
4. Finger LW, Cox DE, Jephcoat AP (1994) A correction for powder diffraction peak asymmetry due to axial divergence *J. Appl. Cryst.* 27:892-900.
5. Stone AJ (2005) Distributed multipole analysis: stability for large basis sets *J. Chem. Theory Comp.* 1:1128-1132.
6. Bondi A (1964) Van der Waals Volumes and Radii *J. Phys. Chem.* 68:441-451.
7. Karamertzanis PG, Kazantsev AV, Issa N, Welch GWA, Adijman CS, Pantelides CC, Price SL (2009) Can the formation of pharmaceutical cocrystals be computationally predicted? 2. Crystal structure prediction *J. Chem. Theory Comp.* 5:1432-1448.
8. Issa N, Karamertzanis PG, Welch GWA, Price SL (2009) Can the formation of pharmaceutical cocrystals be computationally predicted? I. Comparison of like energies *Cryst. Growth Des.* 9:442-453.
9. Cruz-Cabeza AJ, Day GM, Jones W (2008) Towards prediction of stoichiometry in crystalline multicomponent complexes *Chem. Eur. J.* 14:8830-8836.
10. Lancaster RW, Karamertzanis PG, Hulme AT, Tocher DA, Covey DF, Price SL (2006) Racemic progesterone: predicted in silico and produced in the solid state *Chem. Commun.* 4921-4923.
11. Kai Y, Hama F, Yasuoka N, Kasai N (1978) Structural chemistry of layered cyclophanes. III. Molecular structures of [2.2](2,7)pyrenophane-1,1'-diene and pyrene (redetermined) at -160 ^{o}C *Acta Cryst.* B34:1263-1270.
12. Frampton CS, Knight KS, Shankland N, Shankland K (2000) Single-crystal X-ray diffraction analysis of pyrene II at 93 K *J. Mol. Struc.* 520:29-32.
13. Jones W, Ramdas S, Thomas JM (1978) Novel approach to the determination of the crystal structures of organic molecular crystals: low temperature form of pyrene *Chem. Phys. Lett.* 54:490-493.
14. Fabbiani FPA, Allan DR, Parson S, Pulham CR (2006) Exploration of the high-pressure behaviour of polycyclic aromatic hydrocarbons: naphthalene, phenanthrene and pyrene *Acta Cryst.* B62:826-842.
15. David WIF, Shankland K, van de Streek J, Pidcock E, Motherwell WDS, Cole JC (2006) DASH: a program for crystal structure determination from powder diffraction data *J. Appl. Cryst.* 39:910-915.
16. Larson AC, Von Dreele RB (2004) *General Structure Analysis System (GSAS)*, Los Alamos National Laboratory Report LAUR 86-748.
17. Toby BH (2001) EXPGUI, a graphical user interface for GSAS *J. Appl. Cryst.* 34:210-221.
18. Price SL (2009) Computed crystal energy landscapes for understanding and predicting organic crystal structures and polymorphism *Acc. Chem. Res.* 42:117-126.
19. Williams SP, Sigler PB (1998) Atomic structure of progesterone complexed with its receptor *Nature*, 393:392-396.
20. Brzozowski AM, Pike ACW, Dauter Z, Hubbard RE, Bonn T, Engström O, Öhman L, Greene GL, Gustafsson J-Å, Carlquist M (1997) Molecular basis of agonism and antagonism in the oestrogen receptor *Nature* 389:753-758.