Sampling rifamycin conformational variety by cruising through crystal forms: implications for polymorph screening and for biological models†

Alessia Bacchi,* Mauro Carcelli and Giancarlo Pelizzi

Received (in Montpellier, France) 19th March 2008, Accepted 16th June 2008
First published as an Advance Article on the web 7th August 2008
DOI: 10.1039/b804746d

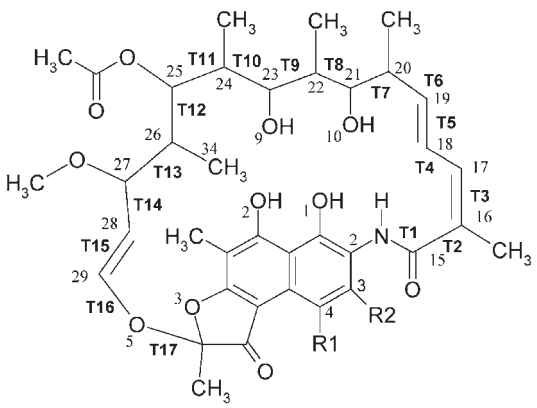
In this work, we report on a non-solvate and on a new solvate forms of rifamycin S, demonstrating that another stable conformation of the active molecule exists and it may be isolated by planning the crystallization conditions. In fact, in these structures rifamycin S has a previously uncharacterised, bent conformation: implications of these findings on the models used to explain rifamycin activity are discussed. Polymorphism has also been reported for the rifamycin-based antibiotic rifaximin, but without single-crystal X-ray diffraction characterization. Here the structure of a tetrahydrate form of rifaximin is discussed; confirming the flexibility of the drug, the two independent molecules in the unit cell are an example of conformational isomorphism.

Introduction

The rifamycins were first isolated by Sensi *et al*. at Lepetit SA in Milan in 1959 as a complex mixture of several congeners.[1] Rifamycins display a broad spectrum of antibiotic activity against Gram-positive and, to a lesser extent, Gram-negative bacteria and they have been used for the last 40 years as one of the mainstay agents in the treatment of tuberculosis, leprosy and AIDS-associated mycobacterial infections.[2] Rifamycins presently on the market are rifampicin, rifabutin, rifapentine and rifaximin.

Rifamycins belong to the family of naphthalenic ansamycins since the molecular structure (Scheme 1) comprises a naphthoquinonic chromophore and a 17-membered *ansa* chain. The overall molecular shape is reminiscent of a basket, since the *ansa* chain average plane is roughly perpendicular to the aromatic basal plane, to which it is connected by two hinges at the opposite side of the naphthoquinone, an amide and a ether junction. Several derivatives mainly differ for the type of substituent R inserted on the chromophore at C3/C4 (Scheme 1); rifamycins are intensely coloured from yellow to red, depending on R. Rifamycin polymorphism is currently a hot field of investigation and a issue of high interest by industries.[3–5] A vast number of rifamycin modifications have been isolated, described and characterized,[6] and the crystal structures of 18 different compounds are known, two of which have been isolated in two different solvate forms (Chart 1).

Rifamycins exert their antibiotic activity by acting on bacterial DNA-dependent RNA-polymerase (DDRP) and by inhibiting the first steps of chain elongation.[7,8] However, not all the derivatives show antibiotic properties, and the range of anti-bacterial activity of differently substituted derivatives spans several orders of magnitude.[9] The first attempts to explain the mechanism of action of rifamycins at a molecular and structural level date back to the exemplary work of Brufani *et al*.,[10] who evidenced that the drug–enzyme interaction had to be based on hydrogen bonds involving the four –OH groups O1, O2, O9 and O10 (Scheme 1). It was readily realized that the spatial arrangement of these four oxygen atoms was fundamental to ensure that the molecule could correctly interact with DDRP. In particular, a spatial pattern for active molecules was suggested, where the four oxygens stuck out of one molecular face, perpendicularly to the *ansa* chain and following some defined O···O interatomic distances[11] (Scheme 2(a)). However, the molecular structural parameters associated to the activity had not been clearly defined, mainly because the conformational variability of the *ansa* chain had been partially discussed in many solution studies, but it had never been systematically rationalized. In particular, solution NMR studies had shown that different conformers exist in water and in $CDCl_3$ and that they differ mainly for the positions of H18 and H27–H29, but this difference was attributed to the flexibility of the junctions, while it was stated that the reciprocal positions of the four



Scheme 1

Dipartimento di Chimica Generale ed Inorganica, Chimica Analitica, Chimica Fisica, Università di Parma, Viale G.P. Usberti, 17A – 43100 Parma, Italy. E-mail: alessia.bacchi@unipr.it

† CCDC reference numbers 681840–681843. For crystallographic data in CIF or other electronic format see DOI: 10.1039/b804746d

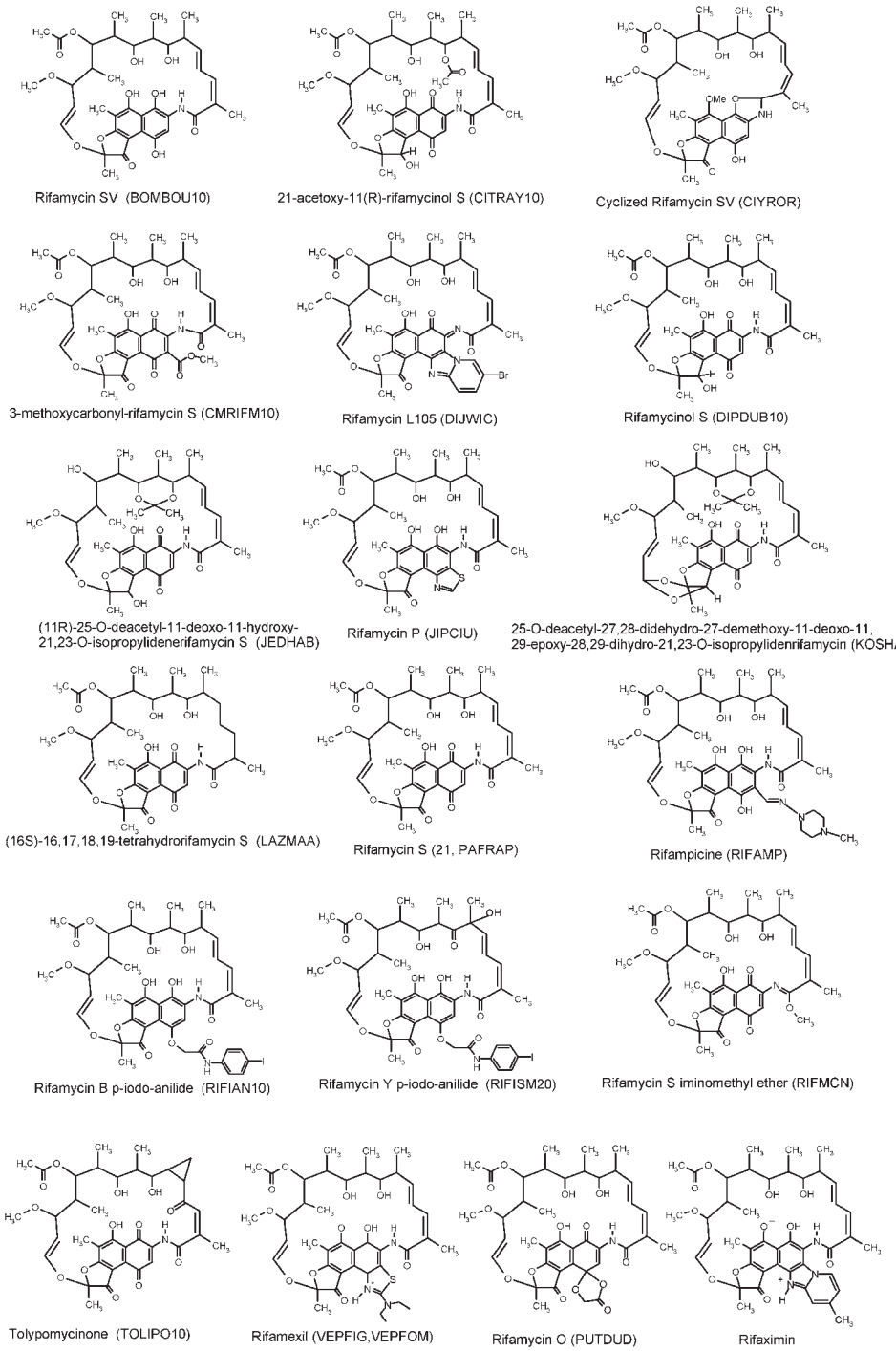
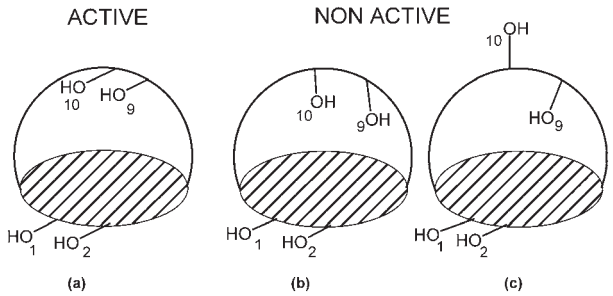


Chart 1

oxygens O1, O2, O9 and O10 were not affected by the solvent.[12]

By contrast, in successive works based on multivariate statistical analysis of structural data, molecular mechanics and molecular dynamics calculations, we evidenced that the active pattern was the result of the favourable combination of two principal conformational degrees of freedom of the *ansa* chain.[13,14] The first one involves the bending around the T6 and T13 bonds of the central part (C18–C28) of the *ansa* chain, which remains rigid. The effect of a deformation at this level is that O8, O9 and O11 lean towards the chromophore and that the C–O vectors point perpendicularly to the aromatic rings, while the methyl group C34 is pulled out of the *ansa* cavity (Scheme 2(b)). This kind of arrangement is found for example in the presence of bulky substituents at C4, as in the crystal structure of rifamycin O.[13] Similar results had been evidenced in a NMR solution study of the effects of the substituents on the *ansa* chain conformation of different rifamycin derivatives.[15] The second way of destroying the pharmacophore active pattern is the alteration of the torsion



Scheme 2

angles around T7 and T8 by local rearrangement of the reciprocal orientation of the C–O9 and C–O11 vectors. This occurs in the presence of chemical modifications in the C20–C22 sector of the *ansa* chain, as in rifamycin Y (Scheme 2(c)).[10] Another point previously highlighted was that the junctions between the *ansa* chain and the chromophore may easily rotate without affecting the overall arrangement of the pharmacophore.[12,13,16]

The breakthrough in the explanation of rifamycin activity has been the determination of the crystal structure of the rifamycin–DDRP adduct, showing that the drug molecule inhibits the initial chain propagation steps simply 'being there', acting as a physical block to the elongation of the polymer; the correct arrangement of the four oxygens is needed to place the rifamycin in the right position.[7] However, this model alone cannot completely account for the effects of C3/C4 substituents on the resistance to DDRP mutations and an allosteric effect involving Mg^{2+} ions has recently been invoked.[8] In these studies the rifamycin molecular model was based on the 'active' (Scheme 2(a)) conformation observed in most of the crystal structures determined so far.

Our principal aim in searching for new crystal forms of rifamycins was to induce conformational polymorphism[17,18] in order to explore the torsional flexibility of the *ansa* chain as a response to varied intermolecular environments, which could be comparable to the situation experienced by the drug approaching DDRP.[19]

In this work we report on the first non-solvate (RS I) and one new solvate (RS II) forms of rifamycin S. These structures contain rifamycin in a bent conformation, showing for the first time that these molecules are distributed over a population of conformers according to the above mentioned principal degrees of freedom of the *ansa* chain.

Rifaximin (Chart 1) is a semisynthetic, rifamycin-based non-systemic antibiotic; it is licensed in particular for use to treat travellers' diarrhoea caused by *Escherichia coli*, but new uses are emerging.[20] In a recent patent application[5] two different polymorphs (rifaximin α and β, depending on the water content) are synthesised and characterised and it was claimed that, on the basis of the two polymorphic forms of rifaximin, it is possible to modulate its level of systemic adsorption. As a contribution to the study of the polymorphism of this important drug, we tried to obtain its X-ray structure: here we report on the crystal structure of a tetrahydrate form of rifaximin (RX4), obtained by slow evaporation of a water–ethanol solution at room temperature.

Results and discussion

Preparation of new crystal forms of rifamycin S

Molecular mechanics[13] and molecular dynamics[14] studies of ryfamycin S have shown that methyl C34 could oscillate back and forth over the chromophore by means of the rocking of the central part of the *ansa*-chain around the hinges at T4/T6 and T13, with a modest energy expense of less than 5 kcal mol^{-1}, if no steric hindrance is present. The two limiting conformers correspond to an open (T4 = −160°, T6 = −60°, T13 = 180°) (Fig. 1, left) and to a closed conformation (T4 = 40°, T6 = 100°, T13 = 80°) (Fig. 1, middle). In the former the oxygens of the pharmacophore are involved in the intramolecular hydrogen bonds O10–H···O9–H···O8, that lock the conformation of the central part of the *ansa* chain, and O2–H···O1 on the chromophore; O8, O9 and O10 do not interact with O1 and O2 because they are too far apart. This arrangement exposes polar groups at the molecular surface and should be favoured in polar solvents. All the rifamycins reported up to now, and in particular the two solvates of rifamycin S (orthorhombic[21] and monoclinic[11]), are crystallized from polar media (water, ethanol, methanol, isopropyl alcohol, ethyl acetate). Effectively, the open conformation is observed in all these crystal structures (Fig. 1, right), and torsion angles similar to the proposed model are adopted (Table 1). On the other end of the oscillation, methyl C34 is exposed on the molecular surface, and O9 and O10 are hidden

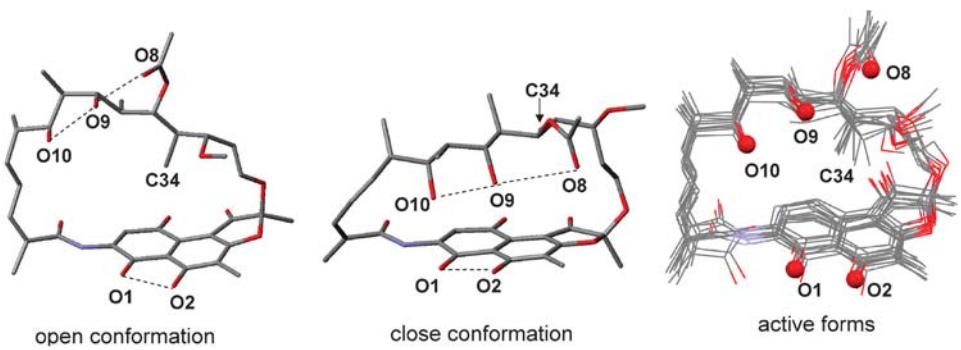


Fig. 1 Comparison of the open (left) and closed (center) conformations of rifamycin S evidenced by molecular modelling studies,[14] intramolecular hydrogen bonds are dashed. Right: superimposition of the crystal structures of all the active forms reported in literature shows that the open conformation is adopted in all cases.

Table 1 Comparison of the torsion angles (°) along the *ansa* chain (T1–T17, Scheme 1) and of the interatomic distances (Å) between the oxygens important for activity in the compounds studied in this work (RS I, two molecules; RS II; RX4, two molecules, e.s.d.'s reported), in the two other solvate forms of rifamycin S (monoclinic methanol solvate and orthorhombic water solvate[11,21]), in rifamycin O[13] and the average values for known active forms in the open conformation.[14] Main differences are evidenced in bold; compounds with lipoid conformations are indicated by asterisks

	RS I a*	RS I b*	RS II*	O*	S. H_2O	S. MeOH	Average open	RX4 a	RX4 b
T1	−173.3(5)	−173.6(5)	−163.1(8)	−176.4	−163.4	−177.3	177.6	168.5(5)	163.7(5)
T2	63.0(7)	64.7(7)	48(1)	63.6	132.9	92.8	74.0	−134.2(6)	56.9(8)
T3	0.6(6)	−4(1)	2(1)	1.1	−3.5	−3.0	3.1	3.6(9)	1(1)
T4	32(1)	37(1)	−175(1)	36.5	−173.5	178.8	176.0	−178.9(7)	−169.4(7)
T5	−174.1(6)	−174.2(6)	−177.4(9)	−178.6	176.2	−175.3	−177.8	−172.0(6)	177.3(6)
T6	103.9(7)	100.5(7)	−13(1)	117.4	−60.7	−45.6	−45.5	−40.4(9)	−66.6(8)
T7	−172.8(5)	−177.4(5)	−173.3(8)	−172.5	−177.3	175.6	179.3	170.2(5)	174.0(5)
T8	−165.5(4)	−165.9(4)	−177.4(7)	−172.1	−170.3	−174.8	−175.9	−171.3(5)	−173.4(5)
T9	76.6(6)	90.8(6)	62.5(9)	63.2	64.9	60.4	58.3	58.5(7)	53.2(7)
T10	−176.4(4)	−173.0(4)	175.3(7)	177.9	−168.4	−171.1	−177.6	−176.4(5)	172.6(5)
T11	−173.1(4)	−175.4(4)	176.1(7)	−177.7	172.1	173.4	167.7	173.6(5)	170.3(5)
T12	−178.3(4)	177.3(4)	−179.5(7)	−174.6	−164.6	−174.1	179.9	173.4(5)	175.2(5)
T13	**61.5(6)**	**60.0(6)**	**59(1)**	**56.0**	**−165.0**	**−172.3**	**−170.9**	**−167.6(5)**	**−172.7(6)**
T14	124.9(6)	125.0(5)	121(1)	113.8	−113.7	−116.5	−119.5	107.9(8)	111(1)
T15	−176.6(6)	−179.0(5)	174.6(9)	175.5	−179.6	−177.8	−176.8	−173.8(7)	−173(1)
T16	35.2(9)	35.6(8)	70(1)	65.9	−134.8	−117.2	−121.1	37(1)	51(2)
T17	−76.9(6)	−78.6(5)	−78(1)	−81.7	−65.0	−61.6	−63.8	−77.7(7)	−76.5(7)
O1–O2	2.549(5)	2.549(5)	2.552(8)	2.54	2.56	2.57	2.50	2.437(6)	2.461(7)
O1–O8	**5.896(7)**	**5.346(7)**	**7.157(9)**	**7.11**	**9.71**	**8.78**	**9.08**	**9.895(7)**	**9.088(8)**
O1–O9	**3.851(5)**	**3.672(6)**	**4.130(8)**	**4.30**	**8.38**	**7.25**	**7.11**	**8.809(6)**	**7.380(7)**
O1–O10	**2.955(5)**	**3.036(6)**	**2.977(9)**	**2.91**	**7.55**	**6.21**	**6.03**	**8.054(6)**	**6.417(9)**
O2–O8	**4.435(7)**	**3.976(7)**	**5.401(9)**	**5.45**	**9.85**	**8.98**	**9.04**	**9.851(7)**	**9.261(8)**
O2–O9	**3.822(5)**	**3.892(6)**	**3.262(9)**	**3.61**	**9.21**	**8.17**	**7.84**	**9.609(6)**	**8.300(7)**
O2–O10	**4.559(6)**	**4.748(6)**	**3.844(9)**	**3.98**	**9.00**	**7.84**	**7.54**	**9.468(6)**	**8.052(9)**
O8–O10	5.659(8)	5.442(7)	6.228(9)	6.10	5.44	5.49	6.04	6.115(8)	6.26(1)
O8–O9	3.015(7)	2.884(7)	3.620(9)	3.43	2.83	2.81	3.39	3.492(7)	3.447(9)
O9–O10	2.682(6)	2.651(6)	2.627(8)	2.70	2.63	2.69	2.72	2.695(7)	2.878(8)

inside the cavity of the *ansa*-chain, linked to O1 and O2 by intramolecular hydrogen bonds. This closed form should be favoured in non-polar media and it is exemplified by rifamycin O (T4 = 36°, T6 = 117°, T13 = 56°).

Therefore, crystallization of rifamycin S from apolar/aprotic solvents was attempted in order to stabilize the less polar conformation. Solvents used include: *n*-hexane, chloroform, toluene, cyclohexane, acetone, THF and their mixtures. Single crystals of rifamycin S were first obtained in a sealed vessel by careful stratification of *n*-hexane as anti-solvent on a concentrated chloroform solution of rifamycin S seeded with a small amount of powder of rifamycin S 'as received'. Large crystals of a new monoclinic non-solvate rifamycin S form grew overnight (RS I); the same form was obtained also by chlorobenzene–hexane (0.5/1 v/v). The X-ray powder diffraction pattern of RS I is identical to that of the 'as received' compound, and it is different from those of the previously known solvate forms of rifamycin S. Crystallization was repeated several times in similar conditions, and a second form (RS II) was obtained, as a chloroform–water solvate.

Crystal structure of RS I

RS I contains two almost identical independent molecules (superimposable within rms = 0.02 $Å^2$), related by a pseudo 2_1 axis along ***c***. The molecules present the lipoid closed conformation (Fig. 2), where O8, O9 and O10 are locked together by intramolecular hydrogen bonds that fix the

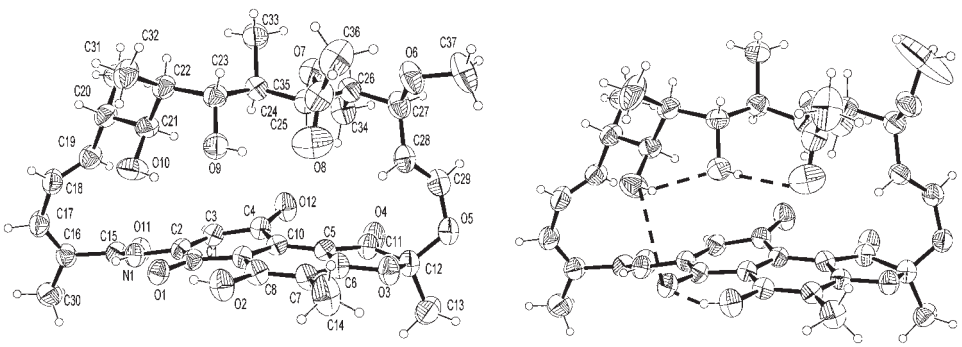


Fig. 2 Perspective view and labelling of the molecular structures of the two independent molecules of RS I, showing for the first time that the closed conformation of rifamycin S may be stabilized in the solid state. Thermal ellipsoids at 30% probability level. Intramolecular hydrogen bonds common to both molecules are indicated on the right.

Table 2 Hydrogen bonds in RS I, RS II and RX4. Atoms of molecules B in RS I and in RX4 are primed. Water molecules are labeled with a W

	D···A/Å	D–H···A/°		D···A/Å	D–H···A/°
RS I					
O2···O1	2.549(5)	146.0	O2′···O1′	2.549(5)	147.3
O9···O8	3.015(7)	162.3	O9′···O8′	2.884(7)	158.5
O10···O1	2.955(5)	107.8	O10′···O1′	3.036(6)	102.9
O10···O9	2.682(6)	137.1	O10′···O9′	2.651(6)	133.8
RS II					
O2···O1	2.552(8)	126(11)	O1W···O8	2.85(1)	165(11)
O9···O2W	2.67(1)	163(16)	O1W···O12(i)	2.85(1)	178(9)
O10···O1	2.977(9)	94(8)	O2W···O13	2.71(1)	144(7)
O10···O9	2.628(8)	104(9)	O2W···O11(i)	2.88(1)	162(14)
RX4					
N1···O11′	2.853(6)	128.0	O1W···O10′	3.43(1)	147.6
N3···O4	2.791(6)	141.8	O1W···O1(v)	2.778(7)	104.7
O1···O2	2.437(6)	141.2	O2W···O2(ii)	2.869(8)	132.3
O9···O8	3.492(7)	159.2	O2W···O4′(ii)	2.975(8)	156.0
O10···O9	2.695(7)	143.6	O3W···O4′(ii)	2.87(1)	143.4
N1′···O1W	2.848(7)	145.3	O3W···O7W	2.81(1)	169.8
N3′···O4′	2.743(6)	142.4	O4W···O2W	2.83(1)	175.7
O1′···O2′	2.461(7)	150(8)	O4W···O8(iii)	2.86(1)	134.5
O9′···O8′	3.447(9)	148.6	O5W···O7W	3.27(2)	134.0
O9′···O10(v)	2.752(6)	115.6	O5W···O8W	2.59(2)	113.9
O10′···O9′	2.878(8)	124.0	O6W···O4W	2.85(1)	146.1
O10′···O11(v)	2.711(9)	134.9	O6W···O4(iv)	2.88(1)	128.9
			O7W···O2′(iv)	2.80(2)	163.8
			O8W···O8′	2.92(1)	117.9

Equivalent positions: (i) $x-1, y, z$; (ii) $1-x, y-1/2, 1-z$; (iii) $1-x, y-1/2, 2-z$; (iv) $x, y, z-1$; (v) $2-x, y-1/2, 2-z$.

geometry of the central section of the *ansa* chain; the *ansa* chain is rotated around T6 and T13 to approach O10 to O2, forming the O10–H···O2 contact (Table 2). Methyl C34 is pushed outside the *ansa* nest (C34···C10 = 4.5 Å, compared to the open form C34···C10 = 3.5 Å). The differences between this closed geometry and the open conformation observed in the previous solvate forms and in all active rifamycins are highlighted by comparing the values of the seventeen torsion angles along the *ansa* chain (Table 1). RS I clusters together with rifamycin O and defines the characteristics of the lipoid conformation in terms of torsion angles T4/T6 and T13 (40, 100 and 60°, respectively). The pattern of interatomic distances between the five oxygens O1, O2, O8, O9 and O10 belonging to the pharmacophore differs in the two forms (Table 1). The mobility of the junctions between the *ansa* and the chromophore is evidenced by the high spread of T2 and T16 (Table 1). In the closed lipoid conformation of RS I, all the –OH functional groups are involved in intramolecular hydrogen bonds (Table 2), while the amidic –NH group does not participate to any intra- or intermolecular contact. The different properties of the closed and open forms of rifamycin S are reflected by the different values of their hydrophobic and hydrophilic surface areas (SA): the hydrophobic SA of the lipoid form RS I is larger than that calculated for the open rifamycin S methanol solvate[11] (673 and 637 Å^2, respectively), while the hydrophilic SA is smaller (126 and 215 Å^2, respectively).

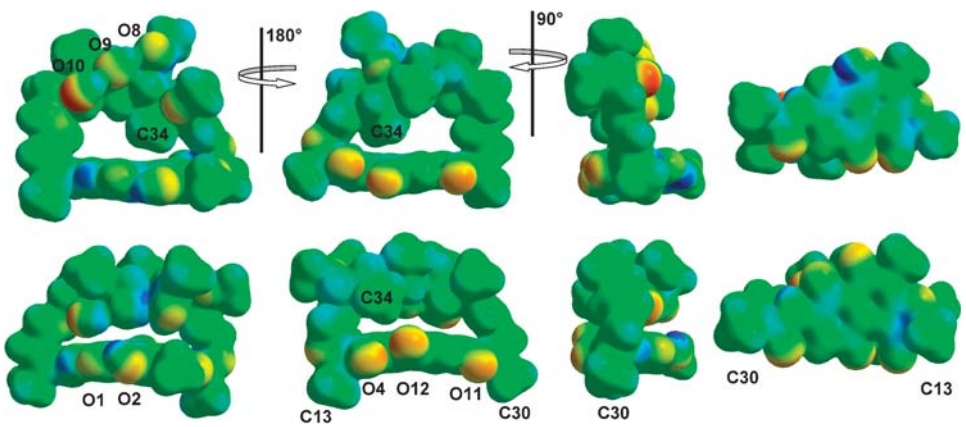


Fig. 3 Molecular electrostatic potential (MEP) at the van der Waals surface (0.008 a.u.) for rifamycin S in the open form (top, PAFRAP) and in the closed form (bottom, RS I). MEP ranges between −0.13 (red) and +0.13 (blue). Different molecular projections are shown, from left to right: front side, containing the oxygen atoms O1, O2, O8, O9 and O10 which define the pharmacophore; back side, containing O4, O12 and exposing methyl C34 in the closed conformation; edge-on view from the amidic junction side; bottom side of the chromophore.

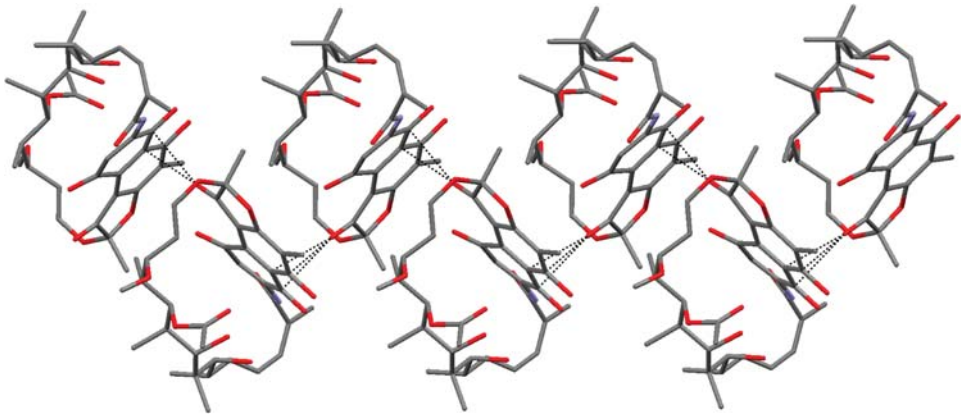


Fig. 4 Crystal packing of RS I. The most significant intermolecular contacts between O4 and the bottom side of the chromophore are shown as dotted lines.

The crystal packing of RS I and the different supramolecular properties of the open and of the closed conformations of rifamycin S may be discussed by considering the electrostatic potential distribution (Fig. 3, first line): on the front side the open conformation (rifamycin S, methanol solvate[11]) exposes an array of negatively charged groups carried by the *ansa* chain, accompanied by the alternation of negative and positive groups on the chromophore which create a polar nest on the side with the pharmacophore; the opposite side of the molecule has three strongly negative points in correspondence of the chromophoric oxygens. There is an accumulation of positive potential on the bottom face of the naphthoquinonic system. Conversely, in the closed conformation (RS I, Fig. 3, second line) the molecular front face containing the pharmacophore groups is folded inside and the polar groups are much less exposed. The rear side maintains negative polar groups and the bottom aromatic side is positively charged. Consequently the crystal packing of RS I is dominated by contacts between the furanic C═O on the molecular rear and the bottom side of the aromatic rings (O4···(C1A, C8A, C9A) ranging from 2.89 to 3.04 Å; O4A···(C1, C8, C9) from 2.80 to 3.29 Å). This results in arrays alternating the two independent molecules oriented such that the aromatic planes are reciprocally perpendicular and the rear side of one molecule perfectly adheres to the bottom side of the following one. A pseudo 2_1 axis along ***c*** relates pairs of independent molecules along the array (Fig. 4).

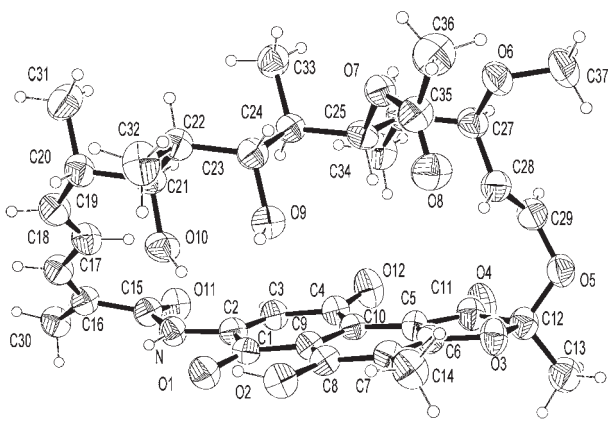


Fig. 5 Perspective view and labelling of rifamycin S in RS II. Solvation chloroform and water not shown for clarity. Thermal ellipsoids at the 30% probability level.

Crystal structure of RS II

RS II crystallizes with a chloroform and two water molecules per unit cell. Here, rifamycin S adopts a lipoid conformation (Fig. 5), differing from RS I mainly in the sector N1–C21 of the *ansa* chain, described by torsion angles T1–T6 (Table 1). In particular, the diene fragment has a cisoid conformation in RS I and a transoid one in RS II. Also in this case the *ansa* chain is folded inside, and the methyl C34 is pushed away from the chromophore (C34···C10 = 3.94 Å). The lipoid conformation observed in RS II is very similar to the one found for rifamycin

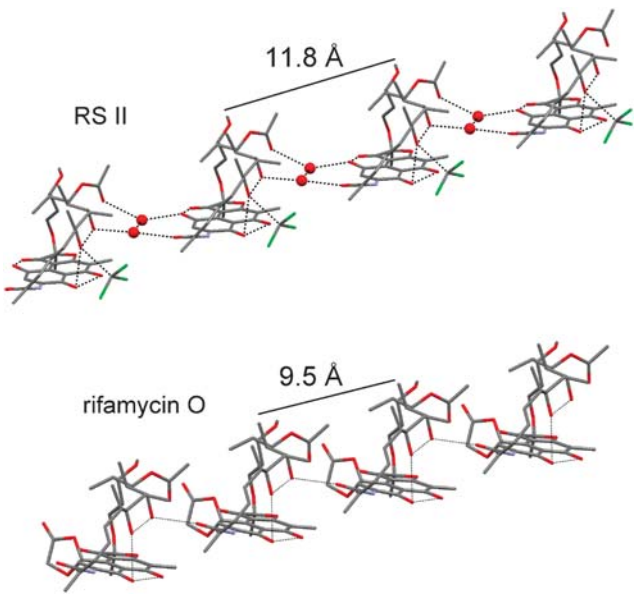


Fig. 6 Comparison of the molecular arrays in the crystal packing of rifamycin S chloroform dihydrate (RS II, top) and of rifamycin O[13] (bottom). In RS II two water molecules (red balls) bridge between the chromophore edge and the oxygens of the *ansa* chain; chloroform molecules decorate the array. In rifamycin O a similar translational array is obtained by direct interactions between rifamycin molecules, without water intercalation.

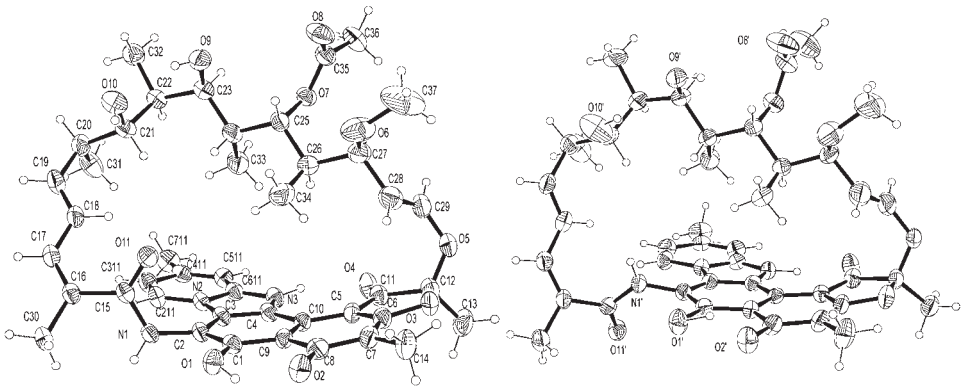


Fig. 7 Perspective view and labelling of the two independent molecules in the crystal structure of rifaximin tetrahydrate, showing different conformations at the amidic junction between the *ansa* chain and the chromophore (N1–C15–O11). Water molecules not shown. Thermal ellipsoids drawn at the 30% probability level.

O (Scheme 2b),[13] where the bending of the *ansa* chain is forced by steric hindrance at C4. As seen from the pattern of interatomic O···O distances within the pharmacophore (Table 1), the presence of the crystallization solvent in RS II does not alter significantly the spatial distribution of interatomic contacts with respect to non-solvate RS I, since O10 is still hydrogen bonded to O1 so that the central part of the *ansa* chain is bent in the lipoid conformation. However, the local arrangement of the central part of the *ansa* chain, carrying O8, O9 and O10, is assisted in RS II by two water molecules O1W and O2W, which link indirectly O9 to O8, while in RS I these were hydrogen bonded together (Table 2). The packing is arranged so that the two water molecules also bridge O8 and O9 with O11 and O12 on adjacent molecules, creating hydrogen bonded translational arrays (Fig. 6); the arrays are decorated by chloroform molecules anchored by CH···O and OH···Cl contacts to O10 and O2 placed in a convenient arrangement in the lipoid conformation (C–H···O: C38···O10, 3.11(1) Å, C–H···O 161.4(6)°; O–H···Cl: O2···Cl1, 3.435(9) Å, O–H···Cl 128(9)°). These translational arrays are also observed in the crystal structure of rifamycin O, where they are organized by direct intermolecular hydrogen bonds without intercalation of water molecules (O9···O11(x, y, z + 1) 2.834 Å, Fig. 6). By comparing the packing of RS I, RS II and rifamycin O it could be argued that, respectively, the inclusion of water (RS II) and the bulky substituent (rifamycin O) hinder the optimization of electrostatic contacts between the negatively charged O4–O12 and the positively charged bottom side of the aromatic rings, as observed in RS I. In fact, in RS II O4 and O12 accept C–H···O contacts from C33 and C21 on the *ansa* chain of an adjacent molecule.

Crystal structure of rifaximin·4H_2O (RX4)

Rifaximin crystallizes with two independent molecules and eight waters, for an overall tetrahydrate stoichiometry. Its characterisation allows to conclude that this is the so-called β rifaximin of the literature.[5] The two independent molecules are related by a pseudo two-fold rotation around ***a***, and both present the open conformation typical of active compounds crystallized from polar protic solvents (Fig. 7), as evidenced by the distribution of the torsion angles along the *ansa* chain, and by the pattern of interatomic distances within the pharmacophore (Table 1). The two independent molecules differ only for the local conformation at the amidic junction, expressed by T2: it is an example of conformational isomorphism.[22] It is interesting to note that these two conformations of the amide plane with respect to the chromophore ring were proposed to explain some NMR data.[12] It is clear that the junctions between the *ansa* chain and the chromophore are quite flexible and, at the same time, that their arrangement does not affect the disposition of the five oxygens important for activity. Compared to the lipoid arrangement of RS I and RS II, in the open conformation of RX4 the intramolecular hydrogen bonds O1–H···O2 and O9–H···O10 are conserved, while the others are replaced by a net of intermolecular contacts involving other rifaximin molecules and waters (Table 2). In particular, the amidic carbonyl O11 of molecule A is twisted towards the middle of the *ansa* chain and points in the direction of O9 and O10; oxygens O9′ and O10′ of molecule B fit in the cavity of the *ansa* chain of molecule A and bridge O9–O10 with O11 by forming the interactions O10′–H···O9′–H···O10(v)–H and O10′–H···O11(v). Conversely, the NH amidic group of molecule B is directed towards

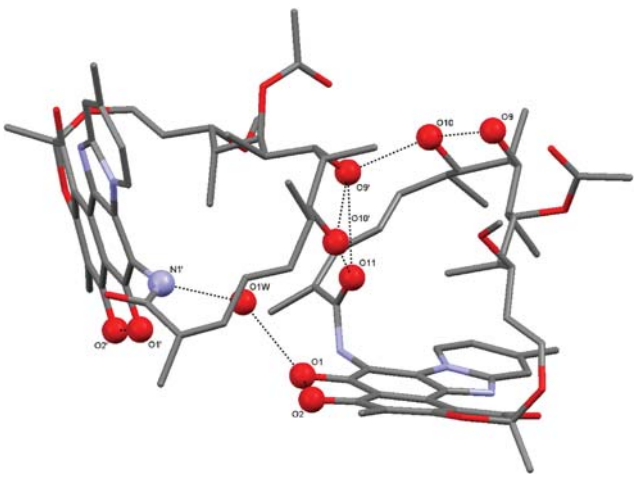


Fig. 8 Direct intermolecular hydrogen bonds between oxygens O9, O10 and O11 on the *ansa* chain of the two independent rifaximin molecules in RX4. Water molecule O1W bridges between the amidic N1′ on one molecule and the chromophore O1 on the other.

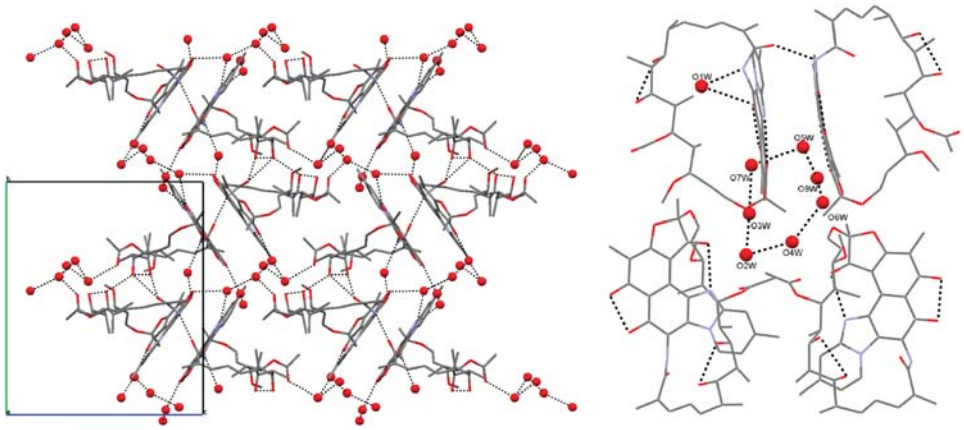


Fig. 9 Left: crystal packing of RX4, showing arrays built by π–π stacking and by the inter-rifamycin contacts shown in Fig. 8. Water molecules O2W–O8W are involved in an extended network of hydrogen bonds and form heptameric clusters (right).

the middle of the *ansa* chain, while the carbonyl is tilted away, and is replaced inside the *ansa* chain by a water molecule, O1W (Fig. 7 and Fig. 8). The numerous intermolecular hydrogen bonds contribute to change the aspect of the IR spectrum of RX4 compared to rifamycin S. In particular, a broad band centered at 3463 cm^{-1} replaces the two distinct bands at 3450 (ν(OH)) and 3365 cm^{-1} (ν(NH)); the furanoic ν(CO) shifts from 1740 cm^{-1} in rifamycin S to 1717 cm^{-1} in RX4. The packing is based on pairing of independent molecules facing each other by π–π association of the bottom side of the chromophores, assisted by a intermolecular hydrogen bond between the amidic NH group of one molecule and the CO group of the other one (Fig. 8, Table 2). π–π stacking of chromophores is not uncommon in the packing of rifamycins, having been observed in other six cases (CSD refcodes, Chart 1: VEPFOM, VEPFIG, DIPDUB10, JEHDAB, JIPCIU, KOSHAB); in one of these (VEPFOM) the stacking is also reinforced by a hydrogen bond between amidic groups, as in RX4. Rifaximin dimers are linked together by the hydrogen bonds involving the OH groups on the *ansa* chain described above (Table 2), and the resulting pattern is an array of directly interacting molecules. The eight water molecules per asymmetric unit that establish a tight network of hydrogen bonds in the crystal packing may be differentiated according to their role: while O1W occupies the cavity inside the *ansa* chain of molecule B, the remaining seven water molecules form a cyclic pattern (Fig. 9).

Relation between packing and docking

The distribution of torsion angles along the *ansa* chain and the pattern of interatomic distances between oxygens belonging to the pharmacophore (Table 1) have shown that in general rifamycin molecules can crystallize in a closed conformation, where the *ansa* chain is folded inwards and the oxygens make intramolecular hydrogen bonds, and in an open conformation, where the oxygens point outwards and are hydrogen bonded with the surrounding molecules in the lattice, that are often represented by solvent of crystallization. It could be argued that the open conformation would be observed in solvate structures, whereas for non-solvate compounds the molecule should prefer the closed folding. In fact we have shown that the closed conformation is favoured by crystallization from apolar aprotic solvents, even in the presence of solvation molecules, as in RS II. From the results shown here it seems that in the crystal the molecules retain the lipoid conformation

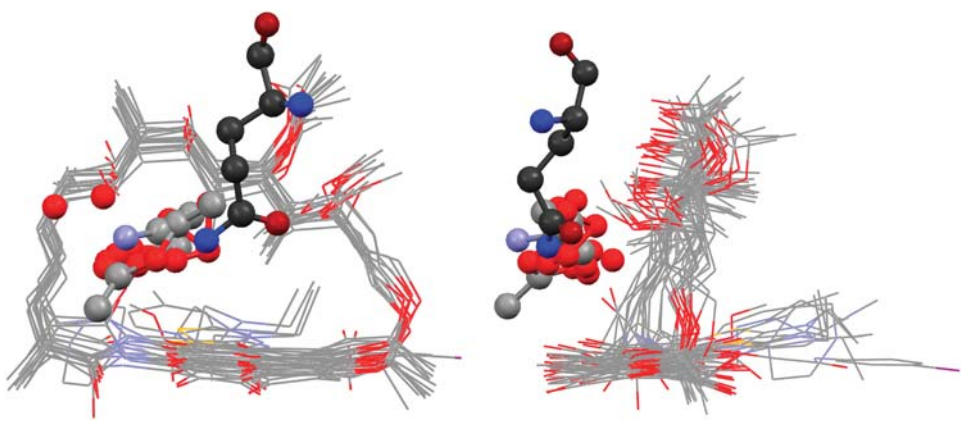


Fig. 10 Conservation of a common supramolecular contact site in the packing of all the crystal structures of rifamycins open forms, here superimposed (front side, left; side view, right). The position of the contacts in the different structures is shown as balls. This site recalls the location of residue C:GLN393 (darker ball and stick) in the crystal structure of the Taq RNA polymerase–rifampicin complex (PDB code: 1YNN).

favoured in a lipophilic solvent.[23] Since the crystal is a self-organized system where the contacts of the molecules with the surroundings are optimized in a subtle interplay between conformation and symmetry,[18,24] it might be assumed that the molecular environment in the crystal packing of active rifamycins could suggest how the macromolecular target should be preorganized to optimize the docking of the drug. Hence we investigated the environment of the open conformations observed in the packing of all the active rifamycins, to verify whether some common patterns recur. We find that all the open conformations pack with a common intermolecular contact with a group in the nest created by the *ansa* chain (Fig. 10). Since this intermolecular contact is conserved in the crystal packing of all active rifamycins in the open conformation, probably it is a very stabilizing interaction; this suggests that an environment able to place a polar group inside the *ansa* chain could favour the locking of the drug molecule in the right geometry. Effectively, this contact position is very close to the location of residue C:GLN393 when the molecule docks on DDRP, as shown by the comparison with the crystal structure of the Taq RNA polymerase–rifampicin complex[4] (PDB code 1YNN) (Fig. 10).

Thermal analysis

The thermal analysis of RX4 does not require any particular comments. In the TGA before 100 °C there is a loss of weight corresponding to the solvent (8% calculated and expected); the DSC trace presents two endothermic peaks related to desolvation and melting (223 °C), then followed by decomposition. More interesting is the thermal analysis of RS I and RS II, in particular if compared with the data of rifampicin,[25] rifamexil[16] and rifamycin O.[13] In fact, besides several solvates, for both rifampicin and rifamexil a non-solvated form, stable at high temperature, has been described, form I. Form I is obtained from the commercially available compound (form II) by recrystallization at high temperature of the melt, by grinding form II in the presence of lipophilic media, or by recrystallization of the amorphous form in lipophilic environment, and for this reason it was designated as 'lipoid'. The crystal structure of form I has never been determined up to now, although some hypotheses were presented based on molecular modelling and spectroscopic data.[3] As shown by molecular modelling and by X-ray analysis, rifamycin O is sterically constrained in the closed, lipoid conformation; its DSC diagram is related to the thermal behaviour of form I:[13] the transition form II→ form I is not present, and rifamycin O does not melt but rather decomposes. In the same way, the DSC of RS I (Fig. 11), crystallized from $CHCl_3$–*n*-hexane in the closed, lipoid conformation, corresponds to those of rifampicin and rifamexil forms I and of rifamycin O. It is reasonable to conclude that form I corresponds to the polymorph containing the closed, lipoid conformer, also for rifampicin and rifamexil.

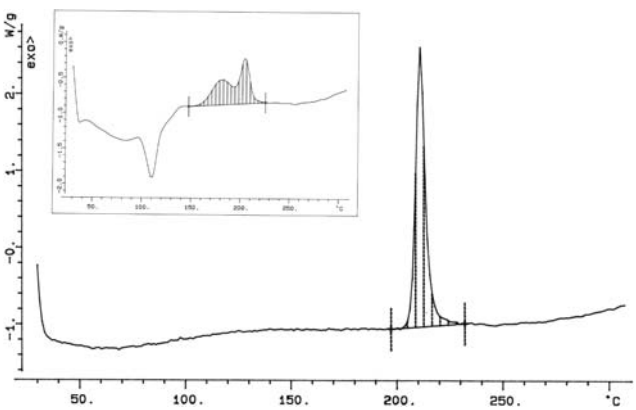


Fig. 11 DSC curve for RS I. RS I in the closed lipoid conformation melts directly at 211 °C, without phase transitions, similarly to form I of rifampicin[25] and rifamexil,[16] and the sterically constrained rifamycin O.[13] Inset: DSC trace of RS II: desolvation occurs below 150 °C, followed by a phase transition to RS I, which then melts.

It is also worth noting that, even if the high melting point of forms I has been attributed to a strong packing related to a favourable conformation of the acetyl group,[25] we have shown here that the packing advantage of the lipoid conformation derives essentially from a more compact molecular shape, proceeding from the optimization of intramolecular polar interactions, while the crystal cohesion is mainly based on several electrostatic dispersive intermolecular contacts involving lipophilic groups and aromatic rings.

Conclusions

Polymorphism is very important in the field of pharmaceutical products and the interest of industry and academy in the solid state of drugs has grown tremendously in recent decades.[26] Even though two crystalline forms (α and β) of rifaximin have been recently described,[5] no single-crystal X-ray characterization was reported. We succeeded in crystallizing a tetrahydrate rifaximin RX4, that is an interesting example of conformational isomorphism. In fact, the two independent molecules in the unit cell differ mainly in the conformation around the T2 bond: C═O is *syn* or *anti* to the methyl C30 group. The flexibility of the *ansa* chain is a relevant topic in the discussion of the structure and activity of rifamycins in general. Conformational flexibility of the *ansa* chain has often been neglected when considering the mode of action of the drug towards its target, DDRP, and rifamycins are generally defined as rigid molecules, where the spatial arrangement of the main four oxygens (O1, O2, O9, O10) cannot be altered. Nevertheless, a judicious consideration of precedent chemometric,[13] molecular modelling[14] and spectroscopic[15] evidence suggests that the *ansa* chain may attain at least two conformations which, in particular, differ significantly by the disposition of the four oxygens. Only one of these conformations is invariably observed in the crystal structures of unhindered rifamycins crystallized from polar solvents. The analysis of the molecular electrostatic potential indicates that this conformation is more hydrophilic than the other one, where the polar groups are hidden inside and the lipophilic groups are exposed at the surface. Also the thermal data indicated that rifamycins often occur in two polymorphic forms, I and II, and that form I may be obtained by grinding II in presence of lipophilic media.[16,25] Therefore, we designed a polymorph screening focused on the isolation of the less polar conformer from apolar/aprotic crystallization solvents. Two new crystal forms are obtained, RS I and RS II, both displaying the expected lipoid conformation. RS I in particular clarifies the structure of so-called

Table 3 Crystal data and structure refinement for RS I, RS II, and RX4

Compound	RS I	RS I low T (100 K)	RS II	RX4
Empirical formula	$C_{37}H_{45}NO_{12}$	$C_{37}H_{45}NO_{12}$	$C_{38}H_{50}Cl_3NO_{14}$	$C_{43}H_{59}N_3O_{15}$
M_r	695.74	695.74	851.14	857.93
Crystal system	Monoclinic	Monoclinic	Monoclinic	Monoclinic
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
a/Å	24.605(6)	24.618(5)	11.764(5)	13.8420(8)
b/Å	14.569(4)	14.588(5)	15.876(5)	19.863(1)
c/Å	10.318(3)	10.278(5)	12.222(5)	16.721(1)
β/°	91.82(2)	92.48(5)	112.34(5)	91.709(1)
V/Å^3	3697(2)	3688(2)	2111(1)	4595.3(4)
Z	4	4	2	4
D_c/Mg m^{-3}	1.250	1.253	1.339	1.240
μ/mm^{-1}	0.777	0.778	2.520	0.094
$F(000)$	1480	1480	896	1832
θ range for data collection (°)	3.53–70.04	3.52–70.47	3.91–71.22	1.22–22.47
Reflections collected	14 061	7198	4364	36 988
Independent reflections (R_{int})	13 496 (0.0190)	7057 (0.0632)	4348 (0.0324)	11 931 (0.0408)
Data/restraints/parameters	13 496/1/926	7057/1/925	4348/9/543	11 931/3/1138
Goodness-of-fit on F^2	1.063	0.873	1.229	0.992
Final R indices [$I > 2\sigma(I)$]	$R1 = 0.0590$	$R1 = 0.0579$	$R1 = 0.0676$	$R1 = 0.0545$
	w$R2 = 0.1557$	w$R2 = 0.1190$	w$R2 = 0.1627$	w$R2 = 0.1384$
R indices (all data)	$R1 = 0.1050$	$R1 = 0.1339$	$R1 = 0.1231$	$R1 = 0.0727$
	w$R2 = 0.2013$	w$R2 = 0.1494$	w$R2 = 0.2044$	w$R2 = 0.1483$
Largest ΔF max./min./e Å^{-3}	0.354/−0.283	0.440/−0.303	0.311/−0.471	0.736/−0.390

form I; it is of note that in the structure only intramolecular and no intermolecular hydrogen bonds are present. On the other hand, the comparison of all the intermolecular hydrogen bonds in the crystal packing of rifamycins crystallized from polar media show that the open form always requires the presence of a hydrogen bond acceptor in the nest created by the open form of the *ansa* chain; the hydrogen bond acceptor usually bridges the hydroxyl groups on the *ansa* chain to the oxygens on the chromophore. It seems from these results that a lipophilic environment in solution favours the less polar folding of the *ansa*-chain, inducing the optimization of intramolecular hydrogen bonds. The same folding is forced in presence of bulky substituents at C3/C4.[13,15] A polar solvent, by contrast, may unlock the net of intramolecular hydrogen bonds and place itself between the *ansa* chain and the chromophore. Self-inclusion is also possible, when the open conformation induced by the polar solvent allows to place a group belonging to a neighbouring rifamycin inside the *ansa* chain.

These results may contribute to explain the mechanism of action of rifamycins, which has been recently a matter of debate especially centered on the effects on drug activity of C3/C4 substituents, apparently lying far away the oxygen atoms involved in the critical contacts.[7,8] The model considered so far, in fact, did not account for a possible influence of protein–rifamycin contacts on the conformation of the *ansa* chain. The evidence of lipoid folding in RS I and RS II demonstrates that protein residues placed at the back side of rifamycin might push on the substituents at C3/C4 and cause the bending of the *ansa* chain, altering the pattern of oxygens at the rifamycin front side.

Experimental

Single crystals of RS I grew overnight from a concentrated chloroform solution of rifamycin S, seeded with a small amount of rifamycin S 'as received', by stratification of *n*-hexane; RS I was obtained also from chlorobenzene–hexane (0.5 : 1 v/v). RS II was obtained by slow evaporation from chloroform solution at room temperature. Crystals of RX4 were obtained by slow evaporation of a water–ethanol solution of commercial rifaximin at room temperature.

X-Ray crystallography

Mo-Kα radiation ($\lambda = 0.71073$ Å) on a SMART AXS 1000 CCD diffractometer for RX4, Cu-Kα radiation ($\lambda = 1.54178$ Å) on a Enraf-Nonius CAD4 diffractometer for RS I and RS II. All data were collected at room temperature (293 K), RS I was also recollected at $T = 100$ K (see ESI†). Cooling of RS I to 100 K did not afford any significant structural change apart from a slight cell contraction as shown in Table 3. Lorentz, polarization and absorption corrections were applied.[27] Structures were solved by direct methods using SIR97[28] and refined by full-matrix least-squares on all F^2 using SHELXL97[29] implemented in the WingX package.[30] Hydrogen atoms were partly located on Fourier difference maps and refined isotropically and partly introduced in calculated positions. Anisotropic displacement parameters refined for all non-hydrogen atoms. Hydrogen bonds were analyzed with SHELXL97[29] and PARST97,[31] and extensive use was made of the Cambridge Crystallographic Data Centre packages[32] for the analysis of crystal packing. Table 3 summarizes crystal data and structure determination results. Electron density isosurfaces (mapped at $\rho = 0.008$ a.u.) coloured according to the value of the molecular electrostatic potential (MEP, ranging between −0.13 and +0.13 a.u.) were calculated and drawn by using Crystal Explorer.[33] The characteristics of the contacts between rifamycin and DDRP were evaluated by using the web resources of the Protein Data Bank site.[34] Surface area properties were calculated by using QikProp 3.0.[35]

CCDC 681840, 681841, 681842 and 681843 contain the supplementary crystallographic data for RS I, RS I at low temperature, RS II and RX4, respectively.

Thermal analyses

Thermal analyses were carried out with a Du Pont TA 2000 instrument equipped with a DSC cell and TG module. The working conditions were as follows: gas flow, N_2 at 25 ml min^{-1}; heating rate 10 °C min^{-1}; open sample aluminium pan; sample weight (mg): 4.383 (RS I), 21.062 (RS II), 4.559 (RX4).

Acknowledgements

The authors thank the "Centro Interfacoltà Misure Giuseppe Casnati" and the "Laboratorio di Strutturistica Mario Nardelli" of the University of Parma; they also express their gratitude to Dr Marino Nebuloni and Dr Alessio Lodola for their suggestions.

References

1 P. Sensi, P. Margalith and M. T. Timbal, *Il Farmaco*, 1959, **14**, 146.
2 G. Lancini and W. Zanichelli, Structure–Activity Relationships in Rifamycins, , in *Structure–Activity Relationships Among the Semisynthetic Antibiotics*, ed. D. Perlman, Academic Press, New York, 1977.
3 S. Agrawal, Y. Ashokraj, P. V. Bharatam, O. Pillai and R. Panchagnula, *Eur. J. Pharm. Sci.*, 2004, **22**, 127.
4 S. Q. Henwood, W. Liebenberg, L. R. Tiedt, A. P. Lotter and M. M. De Villiers, *Drug Dev. Ind. Pharm.*, 2001, **27**, 1017.
5 (*a*) G. C. Viscomi, M. Campana, D. Confortini, M. M. Barbanti and D. Braga, *Eur. Pat. Appl.*, EP 2005-4695 20050303, 2006; (*b*) G. C. Viscomi, M. Campana, M. Barbanti, F. Grepioni, M. Polito, D. Confortini, G. Rosini, P. Righi, V. Cannata and D. Braga, *CrystEngComm*, 2008, **10**, 1074.
6 H. G. Floss and T.-W. Yu, *Chem. Rev.*, 2005, **105**, 621.
7 E. A. Campbell, N. Korzheva, A. Mustaev, K. Murakami, S. Nair, A. Goldfarb and S. A. Darst, *Cell*, 2001, **104**, 901.
8 I. Artsimovitch, M. N. Vassylyeva, D. Svetlov, V. Svetlov, A. Perederina, N. Igarashi, N. Matsugaki, S. Wakatsuki, T. H. Tahirov and D. G. Vassylyev, *Cell*, 2005, **122**, 351.
9 N. Maggi, R. Pallanza and P. Sensi, *Antimicrob. Agents Chemother.*, 1965, **5**, 765.
10 M. Brufani, S. Cerrini, W. Fideli and A. Vaciago, *J. Mol. Biol.*, 1974, **87**, 409.
11 S. K. Arora and P. Arjuan, *J. Antibiot.*, 1992, **45**, 428.
12 L. Cellai, S. Cerrini and A. Segre, *J. Org. Chem.*, 1982, **47**, 2652.
13 A. Bacchi, G. Pelizzi, M. Nebuloni and P. Ferrari, *J. Med. Chem.*, 1998, **41**, 2319.
14 A. Bacchi and G. Pelizzi, *J. Comput. Aided Mol. Des.*, 1999, **13**, 385.
15 M. F. Dampier, C.-W. Chen and H. W. Whitlock, Jr, *J. Am. Chem. Soc.*, 1976, **98**, 7064.
16 A. Bacchi, G. Mori, G. Pelizzi, G. Pelosi, M. Nebuloni and G. B. Panzone, *Mol. Pharmacol.*, 1995, **47**, 611.
17 J. Bernstein and A. T. Hagler, *J. Am. Chem. Soc.*, 1978, **100**, 673.
18 A. Nangia, Molecular Conformation and Crystal Lattice Energy Factors in Conformational Polymorphism, , in *Models, Mysteries and Magic of Molecules*, ed. J. C. A. Boeyens and J. F. Ogilvie, Springer, Dordrecht, 2008.
19 C. Pascard, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1995, **51**, 407.
20 M. Majewski and R. W. McCallum, *Adv. Med. Sci.*, 2007, **52**, 139.
21 S. K. Arora, *J. Med. Chem.*, 1985, **28**, 1099.
22 R. A. Esteves de Castro, J. Canotilho, R. M. Barbosa, M. Ramos Silva, A. Matos Beja, J. A. Paixao and J. Simoes Redinha, *Cryst. Growth Des.*, 2007, **7**, 496, and references therein.
23 R. Yamasaki, A. Tanatani, I. Azumaya, H. Masu, K. Yamaguchi and H. Kagechika, *Cryst. Growth Des.*, 2006, **6**, 2007.
24 A. Gavezzotti and G. Filippini, *J. Am. Chem. Soc.*, 1995, **117**, 12299.
25 G. Pelizza, M. Nebuloni, P. Ferrari and G. G. Gallo, *Il Farmaco*, 1977, **32**, 471.
26 *Polymorphism in the Pharmaceutical Industry*, ed. R. Hilfiker, Wiley-VCH, Weinheim, Germany, 2006.
27 (*a*) *SAINT: SAX Area Detector Integration*, Siemens Analytical Instruments, Inc., Madison, WI, USA; (*b*) G. Sheldrick, *SADABS: Siemens Area Detector Absorption Correction Software*, University of Göttingen, Germany, 1996.
28 A. Altomare, M. C. Burla, M. Cavalli, G. Cascarano, C. Giacovazzo, A. Gagliardi, A. G. Moliterni, G. Polidori and R. Spagna, *SIR97: A New Program For Solving and Refining Crystal Structures*, Istituto di Ricerca per lo Sviluppo di Metodologie Cristallografiche CNR, Bari, 1997.
29 G. Sheldrick, *SHELXL97. Program for Structure Refinement*, University of Göttingen, Germany, 1997.
30 L. J. Farrugia, *J. Appl. Crystallogr.*, 1999, **32**, 837.
31 M. Nardelli, *J. Appl. Crystallogr.*, 1995, **28**, 659.
32 (*a*) F. H. Allen, O. Kennard and R. Taylor, *Acc. Chem. Res.*, 1983, **16**, 146; (*b*) I. J. Bruno, J. C. Cole, P. R. Edgington, M. Kessler, C. F. Macrae, P. McCabe, J. Pearson and R. Taylor, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2002, **58**, 389.
33 M. A. Spackman, J. J. McKinnon and D. Jayatilaka, *CrystEngComm*, 2008, **10**, 377.
34 H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J. D. Westbrook and C. Zardecki, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2002, **58**, 899.
35 *QikProp 3.0 Schrodinger*, LLC, New York, NY, 2007.