Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 18 November 2009
Accepted 13 January 2010

Form, symmetry and packing of biomacromolecules. II. Serotypes of human rhinovirus

A. Janner

Theoretical Physics, FNWI, Radboud University, Heyendaalseweg 135, NL-6525 AJ Nijmegen, The Netherlands. Correspondence e-mail: a.janner@science.ru.nl

The differentiation of the human rhinovirus into serotypes, all having very similar structures and the same architecture, is shown to be related to the packing of the viruses in the crystal and to its space-group symmetry. The molecular crystallographic properties (here described in terms of a molecular lattice Λ_M instead of the form lattice Λ_F considered in previous publications) appear to be compatible with the crystal structure and with the packing lattice Λ_P, introduced in Part I [Janner (2010). *Acta Cryst.* A**66**, 301–311]. On the basis of the enclosing forms of the capsid, a sphere packing is considered, where the spheres touch at kissing points. Residues of each of the four coat proteins (VP1, VP2, VP3, VP4), having a minimal distance from the kissing points, define a set of kissing point related (KPR) residues. In this set only four different residues occur, one for each coat protein, ordered into symmetric clusters {already classified in a previous publication [Janner (2006). *Acta Cryst.* A**62**, 270–286]} and indexed by neighbouring lattice points of Λ_P (or equivalently of Λ_M). The indexed KPR residues allow a fingerprint characterization of the five rhinovirus serotypes whose structures are known (HRV16, HRV14, HRV3, HRV2 and HRV1A). In the fingerprint they occur as internal (if inside the given capsid), as external (if belonging to the neighbouring viruses) or as a contact residue (if at a kissing point position). The same fingerprint, periodically extended, permits a coarse-grained reconstruction of the essential properties of the crystal packing, invariant with respect to the space group of the serotype.

1. Introduction

Molecular forms and the corresponding molecular lattices of axial-symmetric clusters of the coat proteins VP1, VP2, VP3 and VP4 (decamers, pentamers, hexamers, trimers and tetramers) of the human rhinovirus (HRV) have been analysed in a previous publication (Janner, 2006*a*). The striking result was obtained that all the different serotypes HRV16, 14, 3, 2 and 1A have correspondingly the same enclosing forms for all those clusters. In contrast, the space-group symmetry of their crystals is, in general, different: $P22_12_1$ for HRV16, $P2_13$ for HRV14, $P2_122_1$ for HRV3, $I222$ for HRV2 and $P6_322$ for HRV1A (for HRV16 and HRV3 only the setting is different). It is challenging to find out how the same molecular forms and architecture can give rise to different crystal symmetries.

Beside this specific question, it is of fundamental importance to learn more about the ability of viruses to undergo morphological variations, ensuring resistance to antiviral compounds and immune defence, while keeping their own functional infective properties. For the rhinovirus this variation has produced more than 100 different serotypes, with a high level of antigenic diversity and a high specificity to receptor molecules. Serotype is a typical property of the viral surface and is still not well understood (Verdaguer *et al.*, 2000).

The ability to crystallize by clustering through non-covalent interactions is also mainly a surface property. It is, therefore, not surprising that the space-group symmetry adopted by the virus depends on the serotype, despite the shared enclosing forms of the various clusters. The complexity of the system and the subtle properties involved are such that one cannot expect a direct connection between packing binding sites and antigenicity, certainly not within a purely geometric approach. Nevertheless, the present work reveals for the rhinovirus new, interesting features of structural properties which depend on the serotype.

In §2, the crystal packing is considered for two alternative enclosing forms of the capsid: the ico-dodecahedron, with icosahedral symmetry (Janner, 2006*a*,*b*), which in fact is the Keplerian triacontahedron (Janner, 2008), and the sphere inscribed in the cube which circumscribes the ico-dodecahedron. The *kissing points* of the sphere packing are where two spheres of the same size just touch, or *kiss* (Conway & Sloane, 1988). They are called contact points by Koch & Fischer (1999). In the biomacromolecular case kissing points

represent proximity between amino acids of different viruses in the cluster.

In §3, kissing point related (KPR) residues, defined as the nearest ones to a given kissing point, are identified in the virion of the five serotypes for each of the four coat proteins. In most of the cases these residues appear in pairs. Sometimes, however, more than two are at the same minimal distance, within the experimental precision.

In §4, these KPR residues are indexed by neighbouring points of the packing lattice. This allows recognition of the so-called *contact residues*, which are expected to be responsible for the clustering process. Together with the remaining KPR residues they represent, so to say, a *fingerprint* of the serotype. Moreover, this fingerprint encodes the clustering, as it allows one to reconstruct a packing, invariant with respect to the space group of the crystal, in a way specific for each serotype, as discussed in §5.

Finally, in §6, the implications of the results obtained are highlighted from a crystallographic and from a biochemical point of view.

2. Crystal packing

As explained in Part I of the present work (Janner, 2010) a *packing lattice* Λ_P has the crystal lattice Λ as a sublattice and it is possible to define a molecular enclosing form with vertices at the points of Λ_P. In the present case, this form cannot be one of the two forms mentioned in §1: neither the ico-dodecahedron, which requires a six-dimensional lattice because of the icosahedral symmetry, nor the sphere, which is non-indexable because not discrete. A possible alternative molecular enclosing form is the cube, easily indexable and already mentioned, with half-edge $a_c = \tau a_0$, where a_0 is the icosahedral lattice parameter and $\tau = (1 + 5^{1/2})/2$, the golden number. Then the sphere inscribed in the cube and with radius $R_0 = a_c$ is a natural packing unit, no longer requiring to be indexed. The symmetry axes of this cube, which circumscribes the ico-dodecahedron, are not always oriented as the crystal axes and possibly deviate by a small rotation around one axis. One can, however, neglect these deviations and visualize the crystal structure as a packing of ico-dodecahedra or of spheres with radius R_0 and centred at the virions. The metrical relations derived by these descriptions only approximate those of the real structure. In what follows the experimental values are indicated, in brackets, after the ideal ones. The cube itself is an enclosing form of a single capsid and has vertices at points of a *molecular lattice* Λ_M, which only differs slightly from Λ_P.

The lattice Λ_M plays a similar role as the *molecular form lattice* Λ_F considered in Part I (Janner, 2010) and in previous articles. The subtle difference between the two requires some explanation. Both lattices allow the vertices of molecular enclosing forms to be indexed, but in the case of Λ_M the forms considered are always three-dimensional and enclose the biomolecule treated as a packing unit of the crystal. The enclosing forms indexed by Λ_F have the point-group symmetry of the molecule, and this is not necessarily the case for the forms indexed by Λ_M. Moreover, the lattice Λ_M plays a similar role to the packing lattice Λ_P, but in the case of Λ_M the crystal lattice is not involved, as is the case for Λ_P. The corresponding forms are approximately equal, because the enclosed entity (the biomacromolecule) is the same, and this is the reason why the two lattices Λ_M and Λ_P can be considered as each other's approximation. An approximate description of the real structure is essential, but it requires a very careful definition of the concepts used for catching the underlying unity.

2.1. The serotype HRV14

The crystal of the human rhinovirus HRV14, with data available from the Brookhaven Protein Data Bank (PDB) (PDB code 4rhv) and the primary citation (Arnold & Rossmann, 1990), crystallizes according to the cubic space group

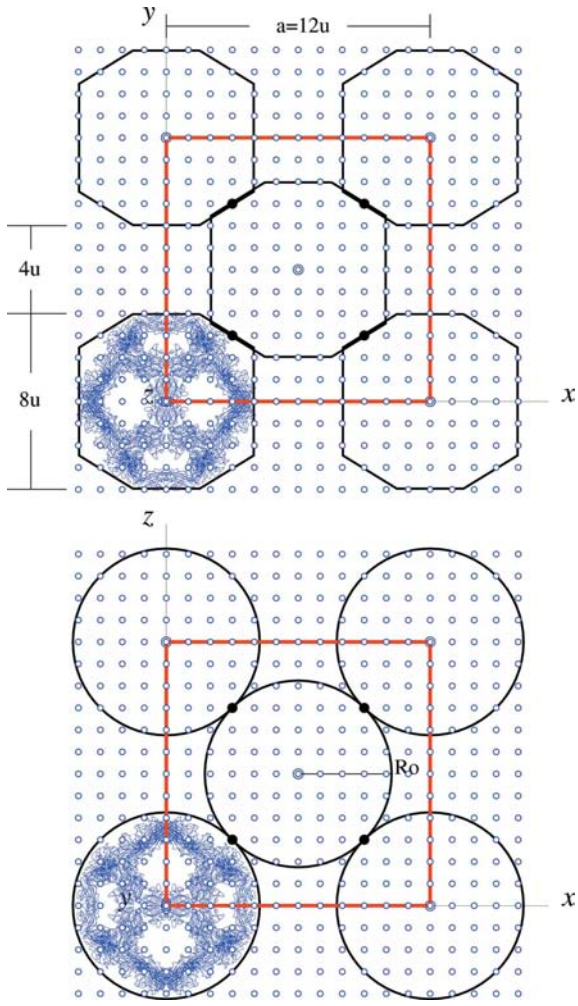

Figure 1
Crystal packing of HRV14 enclosed in the icosahedral ico-dodecahedra (upper view) and in spheres with radii R_0 (lower view) with kissing points (black filled dots) at points of a cubic packing lattice $\Lambda_P(u)$ (empty dots) fitting with the cube (not shown) with edge $8u$ circumscribing the ico-dodecahedron and with the crystal lattice $\Lambda(a)$ as a sublattice. The two lattices are related by $a = N_a u = 12u$. The coat protein VP2 is indicated to illustrate how the capsid fits with the forms. The centres of the spheres (double rings) define a face-centred-cubic (f.c.c.) lattice $23F$.

$P2_13$. With respect to the crystal axes, the capsid is slightly turned (by about 6°) around the threefold axis. It is convenient to characterize the crystal packing disregarding this small deviation. The result is shown in Fig. 1 in projections along the cubic z and y axes. The figure shows the packing lattice Λ_P related to the crystal unit cell, the enclosing spheres and the non-rotated samples of the coat protein chains together with the enclosing ico-dodecahedra, starting from a virus centred at the origin. The fitting is practically perfect.

Both the packing lattice Λ_P and the crystal lattice Λ are of the same Bravais type 23P. The lattice parameters are related by $u = a/N_a$, with $N_a = 12$, so that indeed $\Lambda(a) \subset \Lambda_P(u)$. Each sphere has 12 neighbours and the kissing number is $k = 12$. The sphere packing is a cubic closest packing (c.c.p.), with centres at a face-centred cubic (f.c.c.) lattice. The kissing points of the virus at the origin are at $[\pm\frac{1}{4}, \pm\frac{1}{4}, 0]$ and cyclic permutation. The half-edge a_c of the cubic enclosing form is approximately given by $a_c \simeq n_a u$, for the integral value $n_a = 4$.

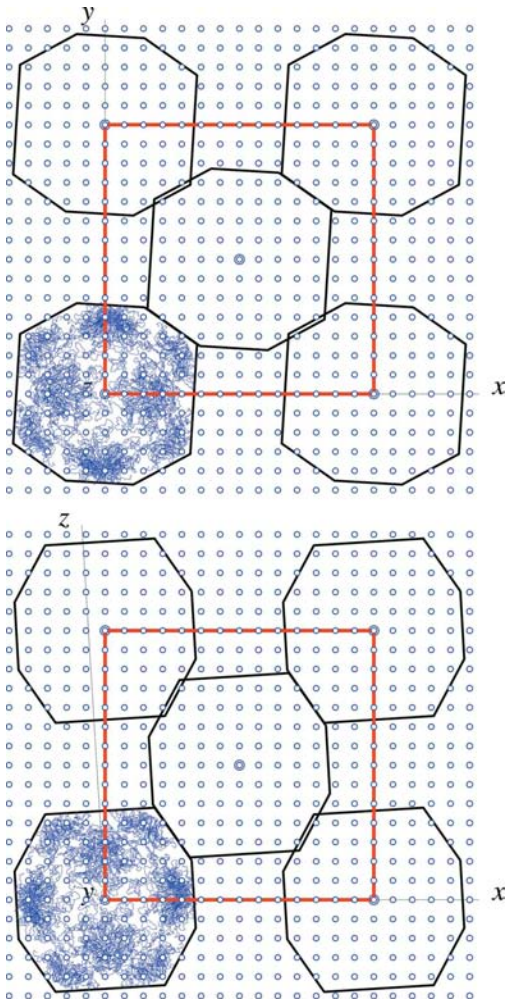

Figure 2
Effect of the different orientation of the capsid with respect to the crystal (corresponding to a 6° rotation around a threefold axis), neglected in Fig. 1. Only the packing of ico-dodecahedra is shown, because the sphere packing is the same for the two orientations. The coat protein VP1 is here used as a filling example.

The effect of the disregarded rotation around the threefold axis is shown in Fig. 2. The sphere packing is exactly the same, that of the rotated ico-dodecahedra deviates only a little, as expected.

2.2. The serotype HRV2

The space-group symmetry of the serotype HRV2 is $I222$ (Verdaguer *et al.*, 2000) (PDB code 1fpn). The relation between the crystal unit cell, the packing lattice $\Lambda_P(u, v, w)$ and the vertices of the viral ico-dodecahedron, projected along twofold axes, is shown in Fig. 3. One sees that the

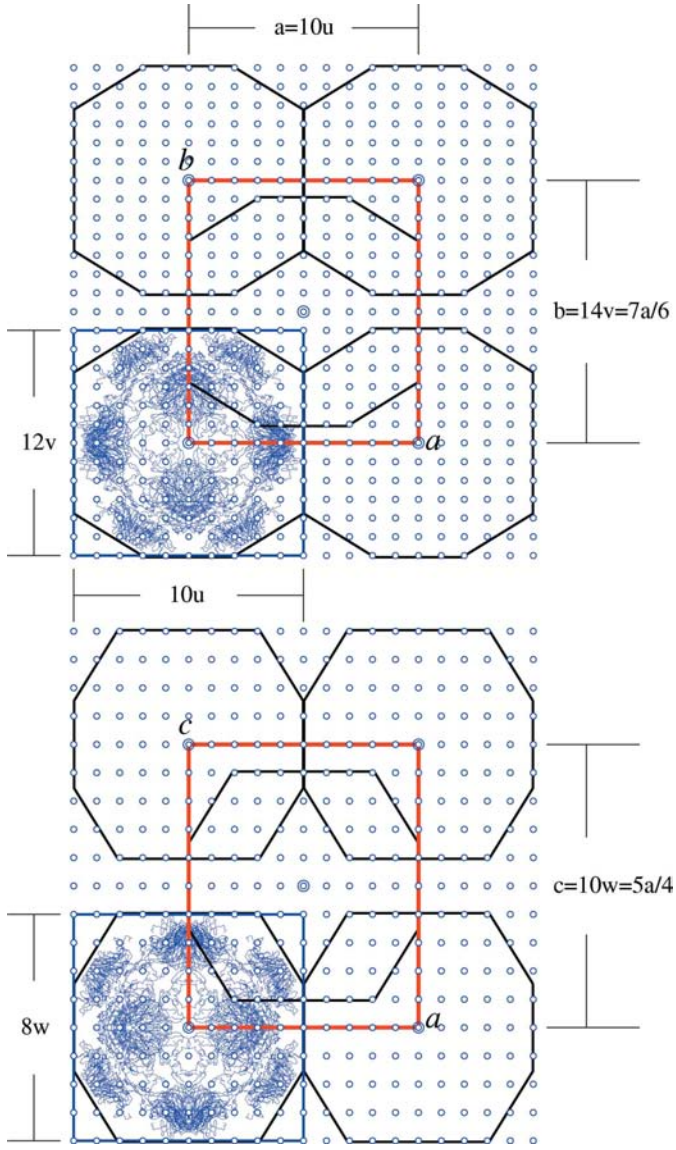

Figure 3
The orthorhombic packing lattice $\Lambda_P(u, v, w)$ of HRV2 is connected with the crystal lattice $\Lambda(a, b, c)$ by the relations $a = N_a u$, $b = N_b v$, $c = N_c w$ with $N_a = N_c = 10$, $N_b = 14$ and fits with the ico-dodecahedra enclosing the capsids. Indeed, the circumscribing cubes have edges given (approximately) by $10u = 12v = 8w$. This leads to the axial ratios $a/b = 6/7$, $c/a = 5/4$ and $c/b = 15/14$. The filling example given is the coat protein VP1.

correspondence with the crystal lattice parameters a, b, c is given by

$$u = \frac{a}{N_a},\ v = \frac{b}{N_b},\ w = \frac{c}{N_c}, \quad N_a = N_c = 10,\ N_b = 14. \quad (1)$$

The cubic molecular envelope with half-edge a_c is expressible in terms of the integers n_a, n_b, n_c and the molecular lattice parameters r, s, t of Λ_M:

$$a_c = n_a r = n_b s = n_c t, \quad n_a = 5, n_b = 6, n_c = 4, \quad (2)$$

approximately related to the lattice parameters of $\Lambda_P(u, v, w)$:

$$a_c \simeq n_a u = n_b v = n_c w. \quad (3)$$

These relations lead to rational axial ratios for the orthorhombic crystal lattice $\Lambda(a, b, c)$:

$$\begin{aligned} \frac{b}{a} &= \frac{N_b n_a}{N_a n_b} = \frac{7}{6} = 1.167\ (1.143), \\ \frac{c}{a} &= \frac{N_c n_a}{N_a n_c} = \frac{5}{4} = 1.250\ (1.233), \\ \frac{c}{b} &= \frac{N_c n_b}{N_b n_c} = \frac{15}{14} = 1.071\ (1.077), \end{aligned} \quad (4)$$

where the experimental values are indicated in brackets. An alternative characterization of the crystal as a sphere packing (Fig. 4) shows that the kissing number is $k = 10$, with eight contact points at mid-distance between sphere centres along the body diagonal of the unit cell ($\pm[\frac{1}{4}\frac{1}{4}\frac{1}{4}]$, $\pm[\bar{\frac{1}{4}}\frac{1}{4}\frac{1}{4}]$ and cyclic), and two more contact points between spheres along the x direction $\pm[\frac{1}{2}00]$. The rational indices indicated are the fractional coordinates with respect to the virion centred at the origin of the crystal lattice (Fig. 3). These ideal kissing positions only approximate the real ones, as one sees from the numerical values obtained for the expressions R_0, R_1, R_2 of the radii of alternative spherical enclosing forms:

$$\begin{aligned} R_0 &= a_0\tau = 95\tau = 153.71, \\ R_1 &= (1/4)(a^2 + b^2 + c^2)^{1/2} = 150.96, \\ R_2 &= a/2 = 154.34. \end{aligned} \quad (5)$$

R_0 is the radius of the sphere inscribed in the cubic enclosing form of the icosahedric capsid. R_1 is that of the spheres touching along the body diagonal of the orthorhombic crystal unit cell, whereas R_2 is the radius of the spheres touching at the mid-edge $a/2$ of the same unit cell. The numerical value 95 has already been obtained in Janner (2006*a*) from a graphical fitting. It is convenient to adopt the finer packing lattice defined by the the values $N_a = N_c = 20$, $N_b = 28$, instead of those of Figs. 3 and 4, in order to ensure that the kissing points also belong to Λ_P.

2.3. The serotype HRV3

The packing of this serotype is similar to that of the previous HRV2, despite the different space-group symmetry, which for HRV3 is $P2_122_1$ (PDB code 1rhi) (Zhao *et al.*, 1996). In the case of HRV3 the orientation of the corresponding twofold axes of the crystal and of the viral capsid coincide. As is standard for this space group, the centre of the virus has been chosen at $[0\bar{\frac{1}{4}}0]$ and not at the origin of the crystal lattice. The kissing number of the sphere packing is again $k = 10$ (Fig. 5). Two kissing points are at $\pm[00\frac{1}{2}]$ and eight at $\pm[\frac{1}{4}0\frac{1}{4}]$, $\pm[\frac{1}{4}0\bar{\frac{1}{4}}]$, $\pm[\frac{1}{4}\frac{1}{2}\frac{1}{4}]$ and $\pm[\frac{1}{4}\frac{1}{2}\bar{\frac{1}{4}}]$. Taking into account the shift in origin and the setting of the space group, these are correspondingly as in HRV2. As in equation (5), the radius of the packed sphere can be expressed in different ways:

$$\begin{aligned} R_0 &= a_0\tau = 95\tau = 153.71, \\ R_1 &= (1/4)(a^2 + b^2 + c^2)^{1/2} = 151.28, \\ R_2 &= c/2 = 150.85. \end{aligned} \quad (6)$$

The parameters of the orthorhombic packing lattice $\Lambda_P(u, v, w)$ are related to those of the crystal lattice $\Lambda(a, b, c)$ (see Fig. 5) by the integers $N_a = N_b = 16$ and $N_c = 12$:

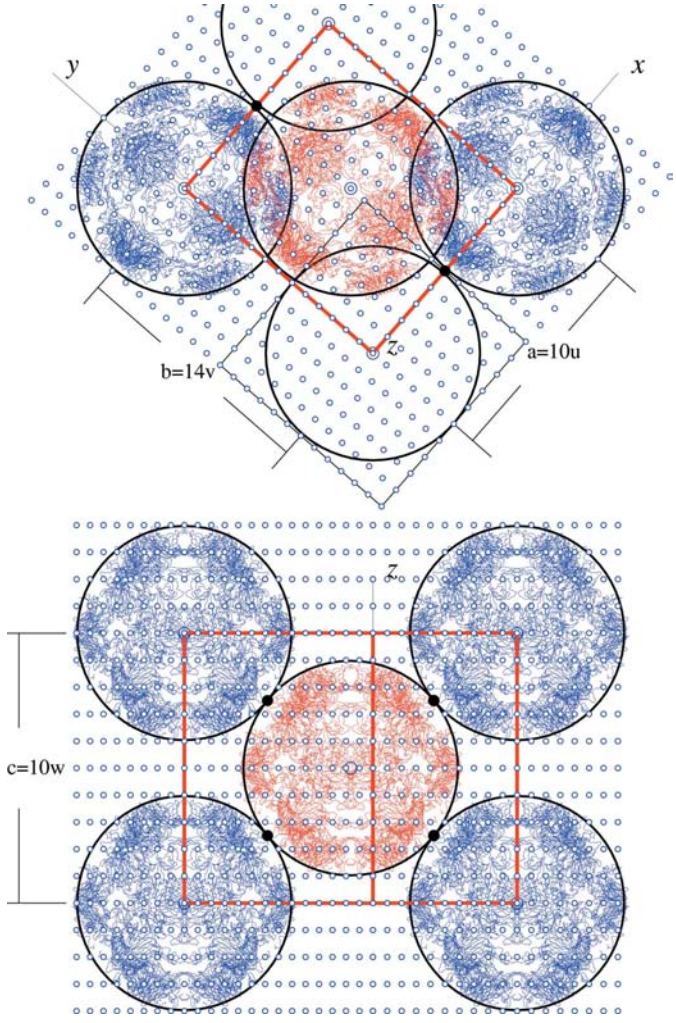

Figure 4
The sphere packing of HRV2, shown together with the same packing lattice as in Fig. 3, has centres forming a 222*I* sublattice. There are eight kissing points along the body diagonals and two more along the x axis in the a direction (black filled dots). These two play the same crystallographic role as other lattice points delimiting the molecular enclosing forms along the b and c directions, and can be disregarded as kissing points. In order to ensure that the packing lattice also includes the remaining kissing points, the finer lattice with $N_a = N_c = 20$, $N_b = 28$ has been adopted as packing lattice (not shown here).

$$u = a/16,\ \ v = b/16,\ \ w = c/12, \tag{7}$$

and to the integers $n_a = n_c = 6$, $n_b = 7$ for the molecular lattice Λ_M indexing the cubic enclosing form. The axial ratios of the crystal lattice Λ follow from these values

$$\begin{aligned}\frac{a}{b} &= \frac{n_b}{n_a}\frac{N_a}{N_b} = \frac{7}{6} = 1.167\ (1.164),\\ \frac{a}{c} &= \frac{n_c}{n_a}\frac{N_a}{N_c} = \frac{4}{3} = 1.333\ (1.319),\\ \frac{b}{c} &= \frac{n_c}{n_b}\frac{N_b}{N_c} = \frac{8}{7} = 1.143\ (1.133),\end{aligned} \tag{8}$$

with the experimental values in brackets.

2.4. The serotype HRV16

The space-group symmetry of the serotype HRV16 (PDB code 1aym) (Hadfield *et al.*, 1997) is the same as for HRV3. The setting $P22_12_1$ is, however, different and in both cases not the standard one, $P2_12_12$, of *International Tables for Crystallography*. In a recently revised version of the data file with PDB code 1aym, the centre of the capsid is chosen at $[\frac{1}{4}\,0\,0]$ and the molecular twofold axes turned around x by $R_x[(\pi/2) - \delta]$, with $\delta = 3.74^\circ$. The effect of the δ deviation for the packing lattice is small, as one can see from Fig. 6, and is disregarded in what follows. The remaining rotation by 90° leads to a different orientation for the projected views adopted in Fig. 6 with respect to previous ones.

Comparison of Fig. 6 with Figs. 3 and 5 makes clear that the kissing points at $\pm[\frac{1}{2}\,0\,0]$ for HRV2 and at $\pm[0\,0\,\frac{1}{2}]$ for HRV3 are no longer there, so that the kissing number is reduced to $k = 8$. All the remaining discussion is similar. The packing and the crystal lattices are related, as above, by $N_a = N_b = 12$, $N_c = 16$ and the molecular form by $n_a = n_b = 5$, $n_c = 7$, leading to the axial ratios

$$\begin{aligned}\frac{a}{b} &= 1.0\ (1.045),\\ \frac{a}{c} &= \frac{21}{20} = 1.05\ (1.083),\\ \frac{b}{c} &= \frac{n_c}{n_b}\frac{N_b}{N_c} = \frac{21}{20} = 1.05\ (1.036).\end{aligned} \tag{9}$$

2.5. The serotype HRV1A

In this case a threefold axis of the capsid is oriented along the hexagonal axis of the crystal, which has the space-group symmetry $P6_322$ (PDB code 1r1a) (Kim *et al.*, 1989). In the projection along the trigonal z axis of the capsid, the ico-dodecahedra appear as regular hexagons with radius R_h arranged around an equilateral triangle centred at the hexagonal screw axis of the crystal, forming in this way a quasi-regular tessellation of the type $\{^3_6\}$ (Coxeter, 1963). The corresponding vertices, projection of the various ico-dodecahedra, are well approximated by points of the packing lattice $\Lambda_P(u, v, w)$, defined with respect to the crystal lattice $\Lambda(a, b, c)$ by the set of integers $N_a = N_b = N_c = 12$ (Fig. 7). The centres of the viral capsids, at the Wyckoff position 2(*d*) $[\frac{2}{3}\,\frac{1}{3}\,\frac{1}{4}]$, $[\frac{1}{3}\,\frac{2}{3}\,\frac{3}{4}]$ of $P6_322$, are at points of the packing lattice. The kissing points of the sphere packing are at $[\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2}]$, $[\frac{1}{2}\,\frac{1}{2}\,0]$, $[0\,\frac{1}{2}\,\frac{1}{2}]$, $[0\,\frac{1}{2}\,0]$, $[\frac{1}{2}\,0\,\frac{1}{2}]$, $[\frac{1}{2}\,0\,0]$, as indicated in Fig. 7. They correspond to the Wyckoff positions 6(*g*) of $P6_322$ for $x = \frac{1}{2}$. The kissing number of the packing is thus $k = 6$. The sphere with radius R_0 inscribed in the cubic envelope with half-edge a_c is also inscribed in the hexagonal projected boundary with radius R_h of the ico-dodecahedron, whereas its height H_h (in the direction of the same threefold axis) is a bit larger than the diameter of the sphere. All are expressible in terms of the icosahedral lattice parameter a_0:

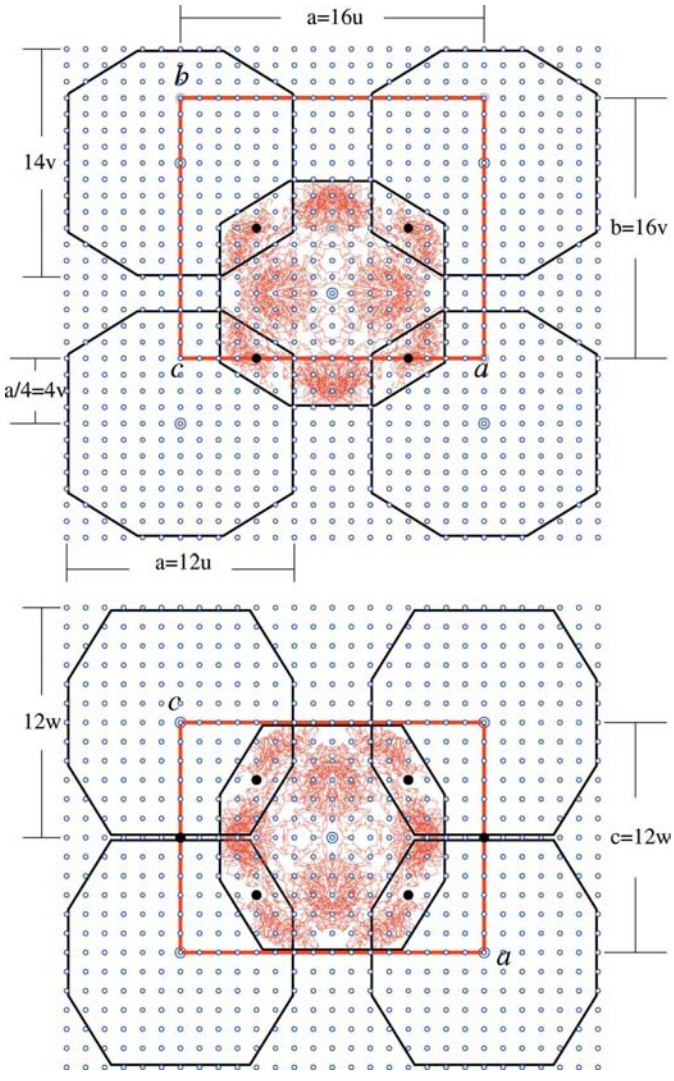

Figure 5
The compatibility of the molecular enclosing forms (defined in terms of ico-dodecahedra and of spheres) and the crystal packing for HRV3 is revealed by the orthorhombic packing lattice $\Lambda_P(u, v, w)$, which has the crystal lattice $\Lambda(a, b, c)$ as a sublattice. One has $N_a = N_b = 16u = 16v$, $N_c = 12w$ for the lattices and $6u = 7v = 6w$ for the half-edges of the cube defined by the ico-dodecahedron of the capsid. The centres of the molecular forms define a 222*I* lattice. Each packing spheres has ten kissing points. As in the previous case, only the eight along the body diagonals of the packing are taken into account. Note that the centre of the starting capsid is at $-4v$ distance from the origin of Λ.

$$R_0 = a_c = \tau a_0, \quad R_h = \frac{2\tau}{3^{1/2}} a_0, \quad H_h = 23^{1/2} a_0, \tag{10}$$

and indeed $2\tau = 3.236$ is less than $2(3)^{1/2} = 3.464$. The enveloping sphere has a boundary passing through points of the packing lattice spanned by the vectors $\mathbf{u} = \mathbf{a}/N_a$, $\mathbf{v} = \mathbf{b}/N_b$, $\mathbf{w} = \mathbf{c}/N_c$. It allows one to derive the axial ratio of the crystal lattice from different expressions for the radius R_0 of the spheres:

$$R_0 = |\mathbf{u} - \mathbf{v} + 3\mathbf{w}| = |\mathbf{u} - 2\mathbf{v}|. \tag{11}$$

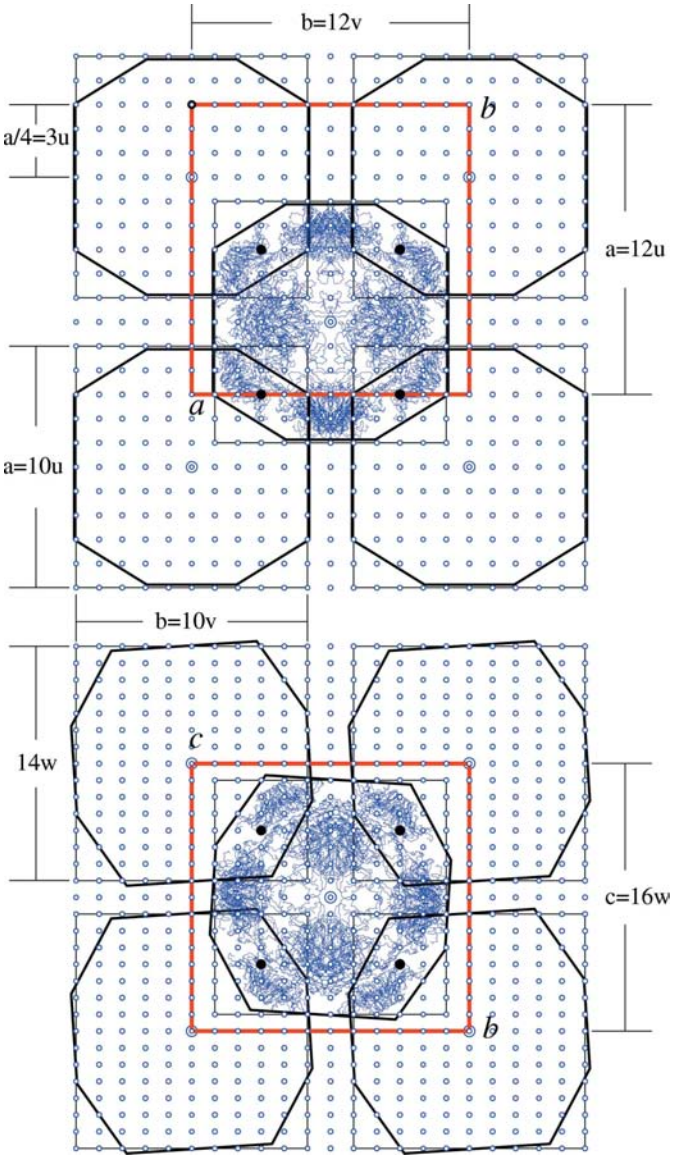

Figure 6
Sphere packing 222*I* for HRV16, with eight kissing points along the body diagonals. The orthorhombic packing lattice $\Lambda_P(u, v, w)$ is related to the crystal lattice $\Lambda(a, b, c)$ by $a = N_a u$, $b = N_b v$, $c = N_c w$ with $N_a = N_b = 12$, $N_c = 16$. Note that the orientation of the twofold axes is different from that of the previous figures and that the centres of the ico-dodecahedra are shifted with respect to the origin of Λ (at the upper left side) by $3u = a/4$ along the a direction. The radius of the spheres (not shown) is given by $R_0 = 5u = 5v = 7w$.

This relation implies the value $\gamma = \frac{4}{3} = 1.333$ (1.365) for the axial ratio c/a, with the experimental value in brackets, as in previous expressions.

Tables 1 and 2 summarize the results obtained so far for the five serotypes discussed above. It is indeed possible to verify that the lattice parameters of Λ_M approximate well those of Λ_P.

The origin of Λ and of Λ_P is the standard one for the given space group, whereas the origin of Λ_M is at the centre C_0 of a chosen virus. These origins are, in general, different and this implies a corresponding change in the indices of the same position. In Table 3 the indices of a kissing point exemplify this dependency for the five serotypes of the rhinovirus.

3. Residues related to kissing points

Kissing points represent proximity between amino acids in different viruses in the cluster. This interaction is not of the covalent type, so that the mutual distance between residues of interacting amino acids can vary considerably. In the present purely geometric approach, the selection criterion chosen for

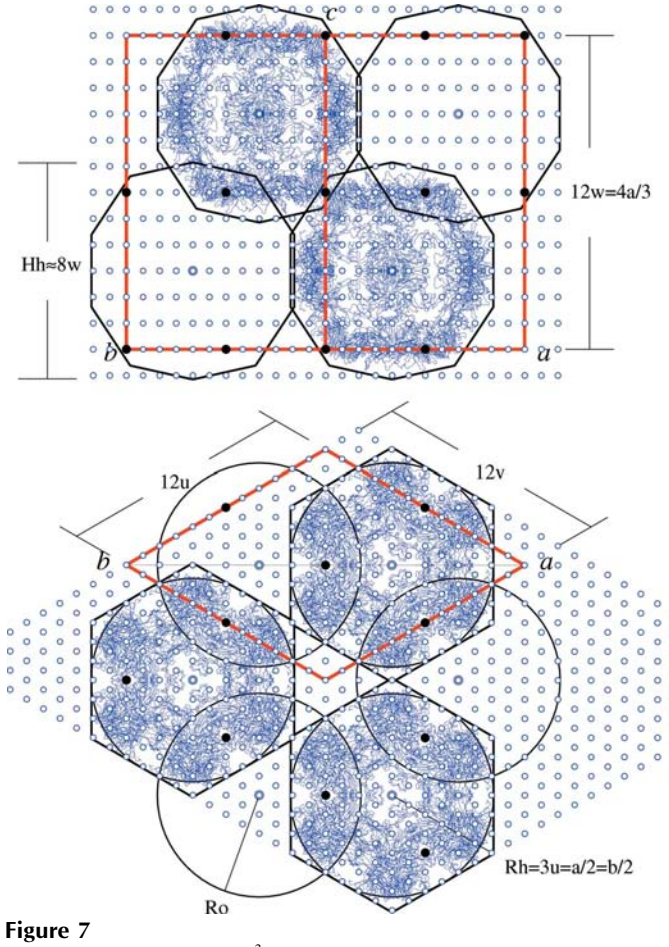

Figure 7
Quasi-regular tessellation $\{^3_6\}$ of the ico-dodecahedra, enclosing HRV1A, projected along the 6_3 screw axis, leading to six kissing points for the corresponding sphere packing. The hexagonal packing lattice $\Lambda_P(u, v, w)$ is related to the crystal lattice $\Lambda(a, b, c)$ by $a = b = N_a u = N_b v$, $c = N_c w$ with $N_a = N_b = N_c = 12$. The sphere packing leads to the axial ratio $c/a = 4/3$.

Table 1
Serotypes of the human rhinovirus.

				Packing lattice			Molecular lattice				Axial ratios		
Serotype	PDB code	Space group	Kissing points	N_a	N_b	N_c	a_0	n_a	n_b	n_c	a/b	c/a	c/b
HRV16	1aym	$P22_12_1$	8	12	12	16	91	5	5	7	1	20/21	20/21
HRV14	4rhv	$P2_13$	12	12	12	12	91	4	4	4	1	1	1
HRV3	1rhi	$P2_122_1$	8 (10)	16	16	12	95	6	7	6	7/6	3/4	7/8
HRV2	1fpn	$I222$	8 (10)	20 (10)	28 (14)	20 (10)	95	10	12	8	6/7	5/4	15/14
HRV1A	1r1a	$P6_322$	6	12	12	2	93	6	6	8	1	4/3	4/3

Table 2
Lattice parameters.

		Crystal lattice Λ			Packing lattice Λ_P			Molecular lattice Λ_M		
Serotype	Bravais	a	b	c	u	v	w	r	s	t
HRV16	$222P$	362.60	347.10	334.90	30.21	28.82	20.93	29.45	29.45	21.03
HRV14	$23P$	445.10	445.10	445.10	37.09	37.09	37.09	36.81	36.81	36.81
HRV3	$222P$	397.90	341.80	301.70	24.87	21.36	25.14	25.62	21.96	25.62
HRV2	$222I$	308.68	352.98	380.48	15.33	12.60	19.02	15.37	12.81	19.21
HRV1A	$3P$	341.30	341.30	465.90	28.44	28.44	37.99	28.96	28.96	40.27

Table 3
Indices of a chosen kissing point (K_1).

Lattice-related indices: $[m_1m_2m_3]_\Lambda = [n_1, n_2, n_3]_{\Lambda_P} = [p_1\ p_2\ p_3]_{\Lambda_M}$.

		Kissing point K_1		
Serotype	Viral centre C_0	$\Lambda(a, b, c)$	$\Lambda_P(u, v, w)$	$\Lambda_M(r, s, t)$
HRV16	$[\frac{1}{4}\,0\,0] = [3, 0, 0]$	$[\frac{1}{2}\ \frac{1}{4}\ \frac{1}{4}]$	[6, 3, 4]	[3 3 4]
HRV14	$[0\,0\,0] = [0, 0, 0]$	$[\frac{1}{4}\ \frac{1}{4}\ 0]$	[3, 3, 0]	[3 3 0]
HRV3	$[0\,\frac{\overline{1}}{4}\,0] = [0, \overline{4}, 0]$	$[\frac{1}{4}\ 0\ \frac{1}{4}]$	[4, 0, 3]	[4 4 3]
HRV2	$[0\,0\,0] = [0, 0, 0]$	$[\frac{1}{4}\ \frac{1}{4}\ \frac{1}{4}]$	[5, 7, 5]	[5 7 5]
HRV1A	$[\frac{2}{3}\ \frac{1}{3}\ \frac{1}{4}] = [8, 4, 3]$	$[\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}]$	[6, 6, 6]	$[\overline{2}\ 2\ 3]$

each of the coat proteins is the minimal distance of residues from a given kissing point of the packed structure, independently of whether the corresponding amino acids interact or not.

Starting from the kissing points of the virus with the centre at the origin C_0 of Λ_M, the set of KPR residues has been determined for the five serotypes considered above, together with the chains to which they belong. In a previous publication (Janner, 2006*a*), the 60 identical copies of each coat protein chains of these rhinoviruses, whose coordinates are given in the PDB files indicated above, have been numbered in a standard way by adopting a common ordering of the elements of the icosahedral group, instead of the order adopted by the various authors. It is, therefore, not surprising to find KPR residues sharing the same chain number and grouped according to the clusters derived in Janner (2006*a*) for the given point-group symmetry. The KPR residues belong either to the central capsid, and are then called *internal*, or to the neighbouring viruses, and are considered to be *external*. In order to emphasize the difference between the new concepts introduced so far, let us recapitulate. *Kissing sites* are at points of the packing lattice where two neighbouring enclosing forms (or packing forms, respectively) meet. *KPR sites* are at the C_α position of the residues of a given chain, which have a minimal distance from a kissing site. Normally, residues are inside the form, so that kissing and KPR sites are different. It is only in the coarse-grained (packing) lattice approximation that KPR sites can coincide with a kissing point. In this case the corresponding residues are *contact residues*.

The result is summarized in Table 4. The layout, which allows a very compact characterization of the KPR residues, requires some explanation. One sees that in all serotypes and for each of the four coat proteins, one residue only occurs in the KPR set. The number indicated is that of the C_α in the polypeptide chain. The distance of these residues from the kissing point (given in Å) depends on the kissing point and on whether the residue is internal, *i.e.* it belongs to the central capsid at C_0, or is external. Moreover, these residues and their polypeptide chain are arranged in clusters of a given point-group symmetry, as analysed and classified in a previous publication (Janner, 2006*a*). The indices of the kissing points (expressed with respect to the conventional unit cell of the crystal) are indicated in the same order of succession as the chain numbers in the cluster.

For example, in HRV3 the residue Gly203 of the coat protein VP3 is at 8.8 Å from the kissing point $[\frac{1}{4}\,0\,\frac{1}{4}]$ if it belongs to the internal chain 24, or as an external residue it has a distance of 3.0 Å from the same kissing point and belongs to the chain 59.

Fig. 8 allows one to visualize differences between serotypes, despite the same architecture of their capsid. Compared are the coat proteins VP2 enclosed in their ico-dodecahedron, for the two serotypes HRV3 (upper part) and HRV14 (lower part) in projected views along a twofold icosahedral axis. The similarity in the chain patterns is apparent, even without going into the details of the indexed enclosing forms involved. In contrast, the differences already appear in number and location of the kissing points, which occur at different points of molecular lattices having different lattice parameters ($2n_a = 2n_c = 12$, $2n_b = 14$ for HRV3 and $2n_a = 2n_b = 2n_c = 8$ for HRV14). The corresponding KPR residues (internal and

Table 4
Kissing point related (KPR) residues.

Residues and distances					Kissing points	
VP1	VP2	VP3	VP4	222 cluster chains	of internal residues	of external residues
(*a*) Orthorhombic HRV2						
Ile280	Ser234	Pro203	Leu43			
24.6	5.5	5.5	57.4	{2, 34, 39, 48}	$[\bar{\frac{1}{4}}\,\bar{\frac{1}{4}}\,\frac{1}{4}]$, $[\bar{\frac{1}{4}}\,\frac{1}{4}\,\bar{\frac{1}{4}}]$, $[\frac{1}{4}\,\bar{\frac{1}{4}}\,\bar{\frac{1}{4}}]$, $[\frac{1}{4}\,\frac{1}{4}\,\frac{1}{4}]$	$[\frac{1}{4}\,\frac{1}{4}\,\bar{\frac{1}{4}}]$, $[\frac{1}{4}\,\bar{\frac{1}{4}}\,\frac{1}{4}]$, $[\bar{\frac{1}{4}}\,\frac{1}{4}\,\frac{1}{4}]$, $[\bar{\frac{1}{4}}\,\bar{\frac{1}{4}}\,\bar{\frac{1}{4}}]$
24.6	5.5	5.5	57.4	{4, 31, 36, 45}	$[\frac{1}{4}\,\bar{\frac{1}{4}}\,\frac{1}{4}]$, $[\frac{1}{4}\,\frac{1}{4}\,\bar{\frac{1}{4}}]$, $[\bar{\frac{1}{4}}\,\bar{\frac{1}{4}}\,\bar{\frac{1}{4}}]$, $[\bar{\frac{1}{4}}\,\frac{1}{4}\,\frac{1}{4}]$	$[\frac{1}{4}\,\bar{\frac{1}{4}}\,\frac{1}{4}]$, $[\frac{1}{4}\,\frac{1}{4}\,\bar{\frac{1}{4}}]$, $[\bar{\frac{1}{4}}\,\bar{\frac{1}{4}}\,\bar{\frac{1}{4}}]$, $[\bar{\frac{1}{4}}\,\frac{1}{4}\,\frac{1}{4}]$
(*b*) Orthorhombic HRV3						
Ile285	Asn229	Gly203	Leu67			
23.5	6.2	8.8	30.2	{5, 24, 51, 56}	$[\frac{1}{4}\,\frac{1}{2}\,\frac{3}{4}]$, $[\frac{1}{4}\,0\,\frac{1}{4}]$, $[\frac{1}{4}\,\frac{1}{2}\,\frac{1}{4}]$, $[\frac{3}{4}\,0\,\frac{3}{4}]$	$[\frac{3}{4}\,\frac{1}{2}\,\frac{3}{4}]$, $[\frac{3}{4}\,0\,\frac{1}{4}]$, $[\frac{1}{4}\,\frac{1}{2}\,\frac{1}{4}]$, $[\frac{1}{4}\,0\,\frac{3}{4}]$
18.6	9.6	3.0	30.5	{8, 22, 54, 59}	$[\frac{1}{4}\,\frac{1}{2}\,\frac{1}{4}]$, $[\frac{1}{4}\,0\,\frac{3}{4}]$, $[\frac{3}{4}\,\frac{1}{2}\,\frac{3}{4}]$, $[\frac{3}{4}\,0\,\frac{1}{4}]$	$[\frac{3}{4}\,\frac{1}{2}\,\frac{3}{4}]$, $[\frac{3}{4}\,0\,\frac{3}{4}]$, $[\frac{1}{4}\,\frac{1}{2}\,\frac{3}{4}]$, $[\frac{1}{4}\,0\,\frac{1}{4}]$
(*c*) Orthorhombic HRV16						
Glu38	Asn233	Pro201	Asp44			
29.2	6.3		61.3	{2, 34, 39, 48}	$[\frac{1}{2}\,\frac{1}{4}\,\frac{1}{4}]$, $[1\,\frac{1}{4}\,\frac{3}{4}]$, $[\frac{1}{2}\,\frac{3}{4}\,\frac{3}{4}]$, $[1\,\frac{1}{4}\,\frac{3}{4}]$	$[\frac{1}{2}\,\frac{3}{4}\,\frac{1}{4}]$, $[1\,\frac{1}{4}\,\frac{1}{4}]$, $[\frac{1}{2}\,\frac{1}{4}\,\frac{3}{4}]$, $[1\,\frac{3}{4}\,\frac{3}{4}]$
29.6	6.7		61.7	{5, 24, 51, 56}	$[\frac{1}{2}\,\frac{3}{4}\,\frac{1}{4}]$, $[\frac{1}{2}\,\frac{1}{4}\,\frac{3}{4}]$, $[1\,\frac{1}{4}\,\frac{1}{4}]$, $[1\,\frac{3}{4}\,\frac{3}{4}]$	$[\frac{1}{2}\,\frac{1}{4}\,\frac{1}{4}]$, $[\frac{1}{2}\,\frac{3}{4}\,\frac{3}{4}]$, $[1\,\frac{3}{4}\,\frac{1}{4}]$ $[1\,\frac{1}{4}\,\frac{3}{4}]$
		4.2		{8, 22, 54, 59}	$[\frac{1}{2}\,\frac{1}{4}\,\frac{1}{4}]$, $[\frac{1}{2}\,\frac{3}{4}\,\frac{3}{4}]$, $[1\,\frac{3}{4}\,\frac{1}{4}]$, $[1\,\frac{1}{4}\,\frac{3}{4}]$	$[\frac{1}{2}\,\frac{3}{4}\,\frac{1}{4}]$, $[\frac{1}{2}\,\frac{1}{4}\,\frac{3}{4}]$, $[1\,\frac{1}{4}\,\frac{1}{4}]$, $[1\,\frac{3}{4}\,\frac{3}{4}]$
		4.2	61.7	{4, 31, 36, 45}	$[\frac{1}{2}\,\frac{1}{4}\,\frac{3}{4}]$, $[1\,\frac{3}{4}\,\frac{3}{4}]$, $[\frac{1}{2}\,\frac{3}{4}\,\frac{1}{4}]$, $[1\,\frac{1}{4}\,\frac{1}{4}]$	$[\frac{1}{2}\,\frac{3}{4}\,\frac{3}{4}]$, $[1\,\frac{1}{4}\,\frac{3}{4}]$, $[\frac{1}{2}\,\frac{1}{4}\,\frac{1}{4}]$, $[1\,\frac{3}{4}\,\frac{1}{4}]$
(*d*) Cubic HRV14						
Asp91	Gly138	Thr231	Ser40			
	20.2	9.8	40.2	{9, 23, 50, 55}	$[\frac{1}{4}\,\bar{\frac{1}{4}}\,0]$, $[\frac{1}{4}\,\frac{1}{4}\,0]$, $[\bar{\frac{1}{4}}\,\bar{\frac{1}{4}}\,0]$, $[\bar{\frac{1}{4}}\,\frac{1}{4}\,0]$	$[\bar{\frac{1}{4}}\,\bar{\frac{1}{4}}\,0]$, $[\bar{\frac{1}{4}}\,\frac{1}{4}\,0]$, $[\frac{1}{4}\,\bar{\frac{1}{4}}\,0]$, $[\frac{1}{4}\,\frac{1}{4}\,0]$
9.4				{8, 22, 54, 59}		
	20.2	9.8	40.2	{10, 15, 29, 43}	$[0\,\bar{\frac{1}{4}}\,\frac{1}{4}]$, $[0\,\bar{\frac{1}{4}}\,\bar{\frac{1}{4}}]$, $[0\,\frac{1}{4}\,\bar{\frac{1}{4}}]$, $[0\,\frac{1}{4}\,\frac{1}{4}]$	$[0\,\frac{1}{4}\,\bar{\frac{1}{4}}]$, $[0\,\frac{1}{4}\,\frac{1}{4}]$, $[0\,\bar{\frac{1}{4}}\,\bar{\frac{1}{4}}]$, $[0\,\bar{\frac{1}{4}}\,\frac{1}{4}]$
9.4				{14, 19, 28, 42}		
	20.2	9.8	40.2	{3,30,35,49}	$[\frac{1}{4}\,0\,\frac{1}{4}]$, $[\bar{\frac{1}{4}}\,0\,\bar{\frac{1}{4}}]$, $[\frac{1}{4}\,0\,\bar{\frac{1}{4}}]$, $[\bar{\frac{1}{4}}\,0\,\frac{1}{4}]$	$[\frac{1}{4}\,0\,\bar{\frac{1}{4}}]$, $[\bar{\frac{1}{4}}\,0\,\frac{1}{4}]$, $[\frac{1}{4}\,0\,\frac{1}{4}]$, $[\bar{\frac{1}{4}}\,0\,\bar{\frac{1}{4}}]$
9.4				{2, 34, 39, 48}		
(*e*) Hexagonal HRV1A						
Val279	Leu179	Thr92	Ala36			
11.0	23.1	15.2		{2, 6, 24; 34, 52, 56}	$[\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2}]$, $[\frac{1}{2}\,0\,\frac{1}{2}]$, $[1\,\frac{1}{2}\,\frac{1}{2}]$; $[\frac{1}{2}\,0\,0]$, $[\frac{1}{2}\,\frac{1}{2}\,0]$, $[1\,\frac{1}{2}\,0]$	
10.7	22.9	15.0		{2, 6, 24; 34, 52, 56}		$[\frac{1}{2}\,\frac{1}{2}\,0]$, $[\frac{1}{2}\,0\,0]$, $[1\,\frac{1}{2}\,0]$; $[\frac{1}{2}\,0\,\frac{1}{2}]$, $[\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2}]$, $[1\,\frac{1}{2}\,\frac{1}{2}]$
			41.0	{1, 5, 23; 33, 51, 55}	$[\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2}]$, $[\frac{1}{2}\,0\,\frac{1}{2}]$, $[1\,\frac{1}{2}\,\frac{1}{2}]$; $[\frac{1}{2}\,0\,0]$, $[\frac{1}{2}\,\frac{1}{2}\,0]$, $[1\,\frac{1}{2}\,0]$	
			41.1	{1, 5, 23; 33, 51, 55}		$[\frac{1}{2}\,\frac{1}{2}\,0]$, $[\frac{1}{2}\,0\,0]$, $[1\,\frac{1}{2}\,0]$; $[\frac{1}{2}\,0\,\frac{1}{2}]$, $[\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2}]$, $[1\,\frac{1}{2}\,\frac{1}{2}]$

external ones), in Fig. 8 all belonging to VP2 chains, do not have the same residue and chain numbers and are at different distances from their kissing point (see Table 4). There is one residue only for each of the coat proteins: for VP2, Asn229 in HRV3 and Gly138 in HRV14 at the chains 222 clusters {5, 24, 51, 56}, {8, 22, 54, 59} for HRV3 and {9, 23, 50, 55}, {10, 15, 29, 43}, {3, 30, 35, 49} for HRV14. The conclusion is that KPR residues are good candidates for a fingerprint characterization of the serotype, despite the impressive reduction in the number of residues (four instead of about 800) and of the chains involved (in Fig. 8, eight or 12 instead of the 60 implied by the icosahedral symmetry).

4. Fingerprint of the serotype

In order to extract from the KPR set the type of information graphically obtained in §2, the KPR residues are indexed by the integral approximation of the coordinates of the C_α atoms in the {**u, v, w**} basis of the packing lattice Λ_P. The same procedure applied to the molecular lattice Λ_M instead of Λ_P yields correspondingly equivalent sets of indices. Possible deviations are accidental. These integral indices then become rational, when expressed as crystallographic ones (with respect to the crystal lattice Λ), and are at the basis of a connection between the crystallographic properties of a single virus and the crystal structure.

One considers again the kissing points of the capsid with the centre at the origin C_0 of Λ_M. The residues having the same indices as a kissing point are called contact residues. Contact residues always combine internal with external residues at the boundary of the capsid. They normally occur in pairs: one internal and one external in such a close contact with each other that their coordinates have correspondingly the same integral approximation, and thus the same set of indices (up to a possible change of origin). This justifies the name. The remaining KPR residues also have a structural meaning related to the space group of the crystal, as discussed further on. Not all kissing points derived in §2 from the crystal packing have associated contact residues. Those without contact residues are, from the point of view of clustering, more or less

accidental and can, therefore, be neglected. For this reason the number of kissing points considered for HRV2 and for HRV3 is reduced from ten to eight (as indicated in Table 1).

The result of indexing the KPR residues is summarized in Table 5. There the indices of kissing point representatives are given with respect to the crystal lattice Λ and to the packing lattice Λ_P (which have the same origin). In order to avoid confusion as much as possible and without overloading the notation, the set of (rational) indices with respect to Λ are written as $[m_1 m_2 m_3]$, whereas the integral ones, with respect to Λ_P, are indicated as $[n_1, n_2, n_3]$, where one has, in particular, $n_1 = N_a m_1$, $n_2 = N_b m_2$, $n_3 = N_c m_3$ (the indices of a point belonging to the molecular lattice Λ_M are given as $[z_1\ z_2\ z_3]$).

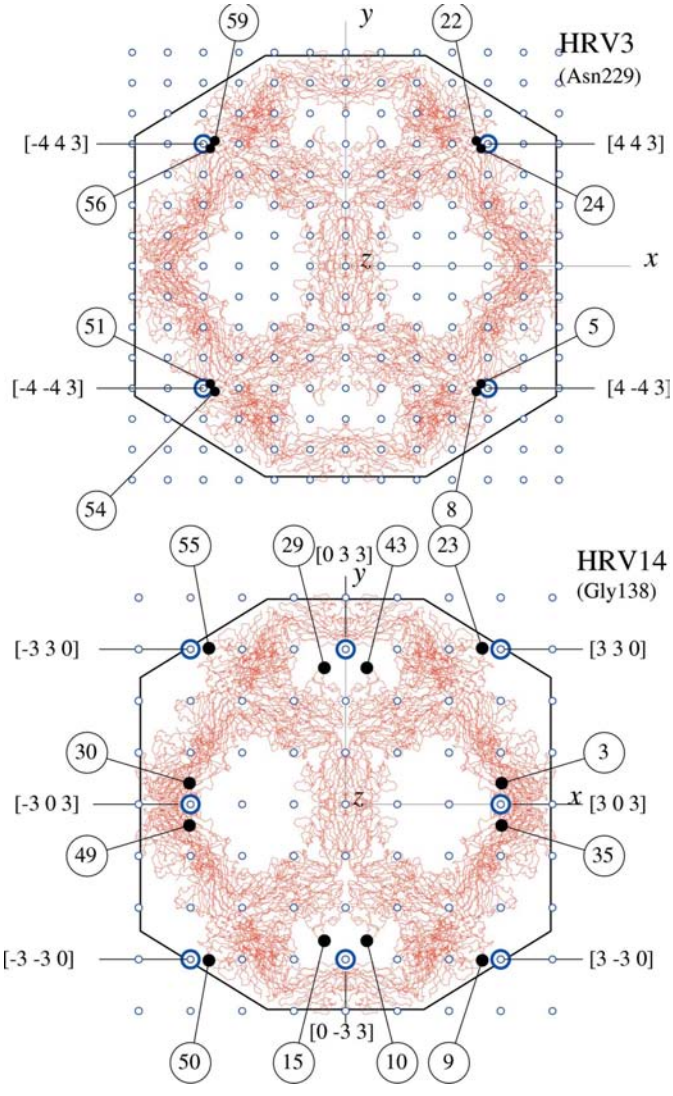

Figure 8
Comparison of the indexed positions of kissing points (double circles) belonging to the packing lattice Λ_P of HRV3 (upper part) and of HRV14 (lower part), and the positions of residues (black dots) belonging in both cases to the coat protein VP2, which are at minimal distance from the kissing points. These residues are Asn229 for HRV3 and Gly138 for HRV14, occurring in 222 clusters of VP2 chains (two clusters for HRV3 and three for HRV14). The whole set of the VP2 coat proteins (in red) shows that the architecture is the same for both serotypes.

The indices of the other kissing points are obtained by applying to the representative the transformations of the appropriate point group: the orthorhombic 222 for HRV2, HRV3 and HRV16, the cubic 23 for HRV14 and the trigonal 32 for HRV1A. The corresponding chain clusters are indicated together with the representative where an indexed KPR residue occurs. The indices of the latter are reported for each of the coat proteins, together with the C_α number of the residue in question. For contact residues, the name of their coat protein is printed in bold.

For example in HRV16, according to the file with PDB code 1aym, the C_α of the residue Asn233 of the coat protein VP2 is N4128. It occurs with the same indices [6, 3, 4] as the kissing point $[\frac{1}{2}\frac{1}{4}\frac{1}{4}]$, in chain 2 as internal and in chain 5 as external. It is thus a contact residue. In a similar way, one finds another contact residue: again Asn233, now associated with the kissing point $[\frac{1}{2}\frac{3}{4}\frac{1}{4}]$ in the internal chain 5 and in the external chain 2, for the virus centred at C_0. For the non-contact residue Glu38 (E282) of VP1, the one associated with $[\frac{1}{2}\frac{1}{4}\frac{1}{4}]$ and belonging to chain 2 has indices [6, 3, 3] and is internal, whereas the one belonging to chain 5 is external and has indices [6, 4, 5]. The additional information given here (and not indicated in Table 5) can be deduced by comparison with the data reported in Table 4.

Neglecting the small metrical differences between the bases of the packing lattice and of the molecular lattice, the indices of a given position can be translated in the way indicated for a kissing point in Table 3. Then the information contained in Table 5 can be represented graphically in projected views of the ico-dodecahedron of the central capsid, together with the points of the corresponding molecular lattice (or of the nearly identical packing lattice). This has been done in a diagrammatic way in Figs. 9, 10, 11, 12 and 13 for the serotypes HRV14, HRV2, HRV16, HRV3 and HRV1A, respectively. In these diagrams the set of indexed KPR residues represents a fingerprint of the virus and its serotype. The indices of these residues, expressed with respect to the molecular lattice Λ_M centred at the capsid and specified by their coat protein, are given in Table 6 for inequivalent positions of the point groups involved (indicated in brackets).

5. From fingerprint to crystal structure

In this section we explore the possibility of a lattice-periodic extension of the KPR residues of the fingerprint diagram, which eventually leads to their indexed positions in a model for the crystal of the given human rhinovirus serotype, and from this to the full crystal structure for the atomic positions of the PDB data file.

From the fingerprint diagrams, or alternatively from Table 6, one sees that contact residues appear in pairs at opposite positions, say K_1 and K_{-1}, with respect to the origin C_0 of the molecular lattice Λ_M. This is also the case for the two spheres which kiss at those points and have correspondingly opposite centres, say at C_1 and C_{-1}. Accordingly, kissing spheres and their contact residues are aligned along the radial directions

Table 5
Indices of KPR residues.

(*a*) HRV2. 222 clusters and chain representatives 48 ∈ {2, 34, 39, 48}, 36 ∈ {4, 31, 36, 45}.

Coat proteins	Residues	$[\frac{1}{4}\,\frac{1}{4}\,\frac{1}{4}] = [5, 7, 5]$		$[\frac{\bar{1}}{4}\,\frac{\bar{1}}{4}\,\frac{\bar{1}}{4}] = [\bar{5}, \bar{7}, \bar{5}]$	
		48	36	36	48
VP1	Ile280 = I2129	[4, 6, 6]	[7, 6, 5]	$[\bar{3}, \bar{8}, \bar{5}]$	$[\bar{6}, \bar{8}, \bar{4}]$
VP2	Ser234 = S3913	[5, 7, 5]	[5, 7, 5]	$[\bar{5}, \bar{7}, \bar{5}]$	$[\bar{5}, \bar{7}, \bar{5}]$
VP3	Pro203 = P5687	[5, 7, 5]	[5, 7, 5]	$[\bar{5}, \bar{7}, \bar{5}]$	$[\bar{5}, \bar{7}, \bar{5}]$
VP4	Leu43 = L6140	[2, 5, 5]	[8, 9, 5]	$[\bar{2}, \bar{5}, \bar{5}]$	$[\bar{8}, \bar{9}, \bar{5}]$

(*b*) HRV3. 222 clusters and chain representatives 5, 24 ∈ {5, 24, 51, 56}, 54, 59 ∈ {8, 22, 54, 59}.

Coat proteins	Residues	$[\frac{1}{4}\ 0\ \frac{1}{4}] = [4, 0, 3]$		$[\frac{\bar{1}}{4}\,\frac{\bar{1}}{2}\,\frac{\bar{1}}{4}] = [\bar{4}, \bar{8}, \bar{3}]$	
		24	59	54	5
VP1	Ile285 = I2128	[4, 0, 2]	[4, 1, 3]	$[\bar{4}, \bar{7}, \bar{3}]$	$[\bar{4}, \bar{8}, \bar{4}]$
VP2	Asn229 = N3884	[4, 0, 3]	[4, 0, 3]	$[\bar{4}, \bar{8}, \bar{3}]$	$[\bar{4}, \bar{8}, \bar{3}]$
VP3	Gly203 = G5698	[4, 0, 3]	[4, 0, 3]	$[\bar{4}, \bar{8}, \bar{3}]$	$[\bar{4}, \bar{8}, \bar{3}]$
VP4	Leu67 = L6275	$[3, \bar{1}, 2]$	[5, 1, 4]	$[\bar{3}, \bar{7}, \bar{2}]$	$[\bar{5}, \bar{9}, \bar{4}]$

(*c*) HRV16. 222 clusters and chain representatives 2 ∈ {2, 34, 39, 45}, 8 ∈ {8, 22, 54, 59}, 5 ∈ {5, 24, 51, 56}, 36 ∈ {4, 31, 36, 45}.

Coat proteins	Residues	$[\frac{1}{2}\,\frac{1}{4}\,\frac{1}{4}] = [6, 3, 4]$				$[\frac{1}{2}\,\frac{3}{4}\,\frac{1}{4}] = [6, 9, 4]$			
		2	5	8	36	5	2	36	8
VP1	Glu38 = E282	[6, 3, 3]	[6, 4, 5]			$[6, \bar{2}, 3]$	[6, 9, 5]		
VP2	Asn233 = N4128	[6, 3, 4]	[6, 3, 4]			[6, 9, 4]	[6, 9, 4]		
VP3	Pro201 = P5900			[6, 9, 4]	[6, 9, 4]			[6, 3, 4]	[6, 3, 4]
VP4	Asp44 = D6409	[5, 3, 1]	[6, 5, 5]		[7, 3, 7]	[6, 11, 3]	[7, 9, 7]	[5, 9, 1]	

(*d*) HRV14. 23 clusters and chain representatives 22, 59 ∈ {8, 22, 54, 59}, 23, 55 ∈ {9, 23, 50, 55}.

Coat proteins	Residues	$[\frac{1}{4}\,\frac{1}{4}\,0] = [3, 3, 0]$			
		22	59	23	55
VP1	Asp91 = D625	[3, 3, 0]	[3, 3, 0]		
VP2	Gly138 = G3356			[3, 3, 0]	[3, 3, 0]
VP3	Thr231 = T6103			[3, 3, 0]	[3, 3, 0]
VP4	Ser40 = S6440			[2, 2, 0]	[4, 4, 0]

(*e*) HRV1A. 32 clusters and chain representatives 2, 52 ∈ {2, 6, 24; 34, 52, 56}, 1, 51 ∈ {1, 5, 23; 33, 51, 55}.

Coat proteins	Residues	$[\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2}] = [6, 6, 6]$			
		2	52	1	51
VP1	Val279 = V2193	[6, 6, 6]	[6, 6, 6]		
VP2	Leu179 = L3587	[6, 5, 6]	[5, 6, 6]		
VP3	Thr92 = T4933	[6, 6, 6]	[6, 6, 6]		
VP4	Ala36 = A6169			[7, 6, 5]	[6, 7, 7]

from C_0 towards the kissing points, in an alternating succession of sphere centres and contact residues as

$$\ldots, C_{-1}, K_{-1}, C_0, K_1, C_1, \ldots. \qquad (12)$$

This property allows a lattice-periodic extension of the original KPR residues of the fingerprint. One may describe the approach by the following conceptual steps.

(1) Model of the capsid. An extremely reduced and coarse-grained model of the capsid is obtained from the four KPR residues (one for each coat protein) indicated in Table 5, and their indexed positions at points of the lattice Λ_M according to the fingerprint diagram and to Table 6.

(2) Empty sphere packing. The alternating sequences of centres and kissing points, obtained for all the contact residues of the fingerprint, define the sphere packing of the crystal.

(3) Decorated sphere packing. The central sphere at C_0 is decorated by the internal and the contact residues given in the fingerprint. The neighbouring sphere, say kissing at K_1 and with centre C_1, is decorated by the external residues at K_1, together with the translated image of those at the opposite kissing point K_{-1} and belonging thus to the sphere at C_{-1}. The following kissing sphere in the same direction is again decorated by the (translated) internal residues of the sphere at the origin. Repeating this step for all the kissing directions, one obtains a periodic packing of spheres, decorated by internal and by external KPR residues, sharing contact residues. This

extended set of KPR residues is a kind of simplified and coarse-grained model of the crystal, which reflects the simple model of the capsid indicated above.

(4) Crystal space group. The space group of the decorated sphere packing is expected to be the same as that of the serotype crystal, but with lattice parameters a', b', c' multiples of the r, s, t parameters of the molecular lattice Λ_M, which are slightly different from the a, b, c lattice parameters of the crystal structure ($a' \simeq a, b' \simeq b, c' \simeq c$).

(5) Verification. Instead of deriving the space-group symmetry of the decorated sphere packing, one simply verifies that the indexed positions of the crystal model are left invariant by the space group of the crystal indicated in the PDB file. This should be done after shifting the origin of Λ_M to that of the packing lattice Λ_P and after translating accordingly the positions from points of Λ_M to the corresponding neighbouring ones of Λ_P. In other words, the data of Table 5 are used instead of those of Table 6. The space group of this configuration should then be that of the real crystal. The above verification is needed in order to ensure that the integral approximation of the real positions leading to indexed residues (as explained in §4) and the simplification represented by the crystal model obtained from the fingerprint did not change the symmetry of the serotype.

(6) Crystal structure. Eventually the full crystal structure is obtained by applying the verified space group to the atomic coordinates indicated in the PDB file.

In the orthorhombic case, the eight kissing spheres around the central one imply a $222I$ lattice for the empty sphere packing, with (conventional) cell parameters given by $a = 20u$, $b = 28v$, $c = 20w$ for HRV2, $a = 16u$, $b = 16v$,

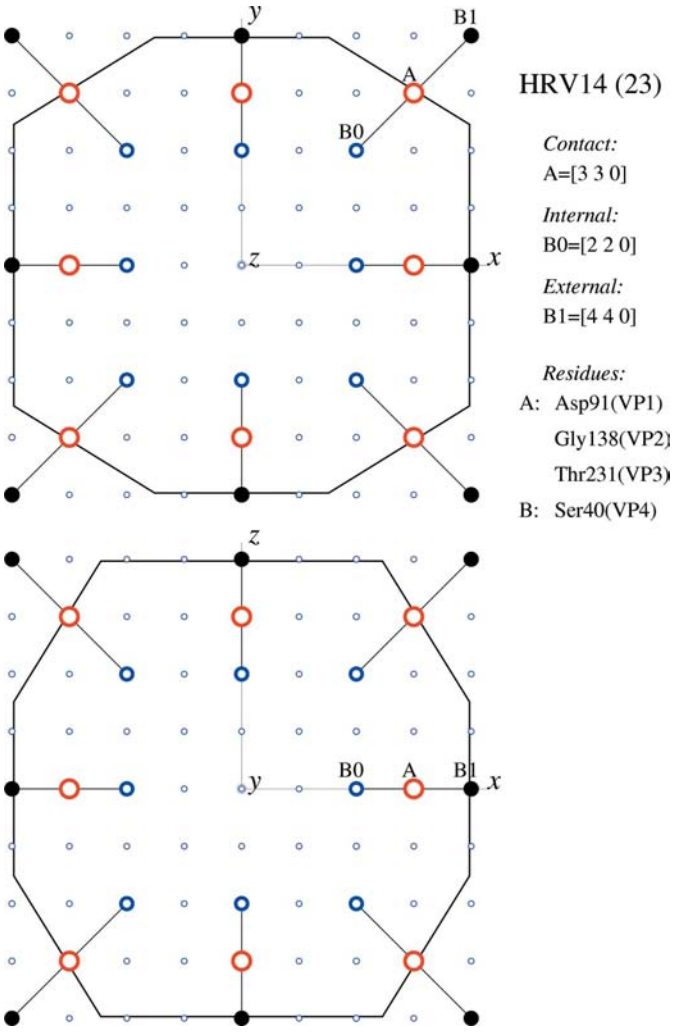

Figure 9
Projected views of the fingerprint diagram for the cubic HRV14. Shown are the indexed KPR residues at points of the molecular lattice Λ_M. This lattice could be replaced by the packing lattice Λ_P with only a small variation in the metrical features as required for a fitting with the crystal lattice. The internal residues (small blue circles) are inside this capsid. The indexed contact residues (large red circles) coincide with the kissing points and are thus at the boundary, whereas the external KPR residues (black filled circles) belong to the neighbouring capsids. The full set of KPR residues can be obtained from the indexed positions indicated, by applying the transformations of the point group indicated in brackets (see Tables 5 and 6). The lines connecting KPR positions play a graphical role only.

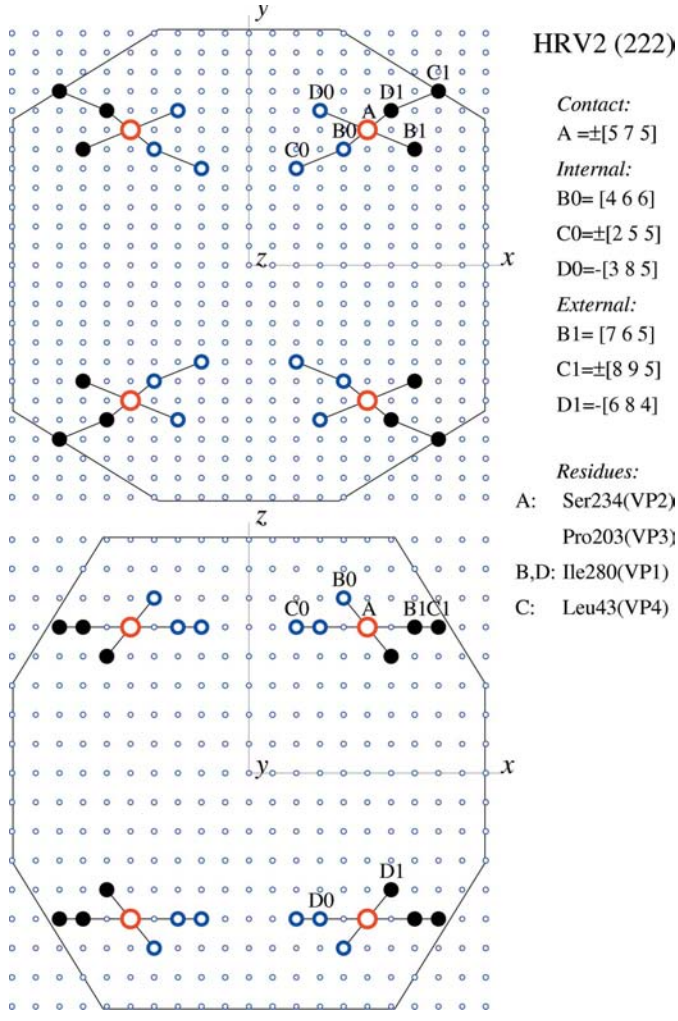

Figure 10
Fingerprint diagram for the orthorhombic HRV2 (with the same conventions as in Fig. 9).

$c = 12u$ for HRV3, and $a = 12u$, $b = 12v$, $c = 16w$ for HRV16, with u, v, w the parameters of the (primitive) packing lattice Λ_P (see Table 5 and compare with Table 1). The decoration of the spheres by internal/external KPR residues reduces the periodicity to the sublattice 222*P* for HRV16 and HRV3, and keeps 222*I* for HRV2. In a similar way, the centred 23*F* lattice periodicity of the empty sphere packing of HRV14, implied by the 12 kissing points of the central sphere and with lattice parameter $a = 12u$, is reduced by the decoration to the sublattice 23*P*. Finally, for HRV1A, the lattice 32*P* with parameters $a = b = 12u = 12v$, $c = 12w$ is conserved by the decoration. Only the point-group symmetry is reduced from 622 to 32, with the twofold axes in perpendicular orientation with respect to hexagonal edges of the enclosing boundary (see Fig. 13).

The space-group invariance of the set of indices reported in Table 5, extended by the lattice periodicity given above, has been verified for the five serotypes: HRV16 ($P22_12_1$), HRV14 ($P2_13$), HRV3 ($P2_122_1$), HRV2 ($I222$) and HRV1A ($P6_322$). This verification leads to the set of space-group-inequivalent positions occupied by the indexed KPR residues reported in Table 7, where the Wyckoff positions and the residues involved are also indicated.

6. Conclusion

The general aim of this series of papers is to explore the connection between the molecular crystallographic properties observed in many biomacromolecules and their crystal structures. The present work confirms for the human rhinovirus the compatibility between the two types of crystallographic order, already illustrated by tutorial examples in Part I (Janner, 2010). The more specific aim of gaining an insight into the structural peculiarities of different serotypes has been partially realized. More work is to be done. The variation in space-group symmetry could be expressed in terms of very simple models, one for the viral capsid and one for the crystal, both involving, for each of the five serotypes with known structure, four residues only (one for each coat protein) at minimal distance from the kissing point in their indexed

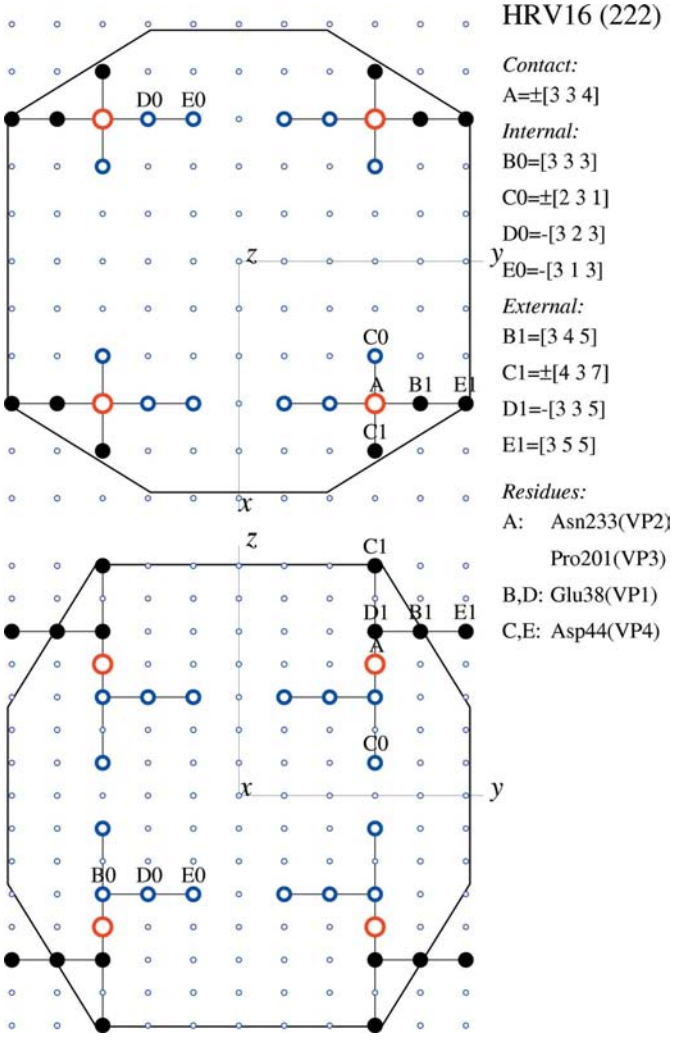

Figure 11
Fingerprint diagram for the orthorhombic HRV16 (with the same conventions as in Fig. 9).

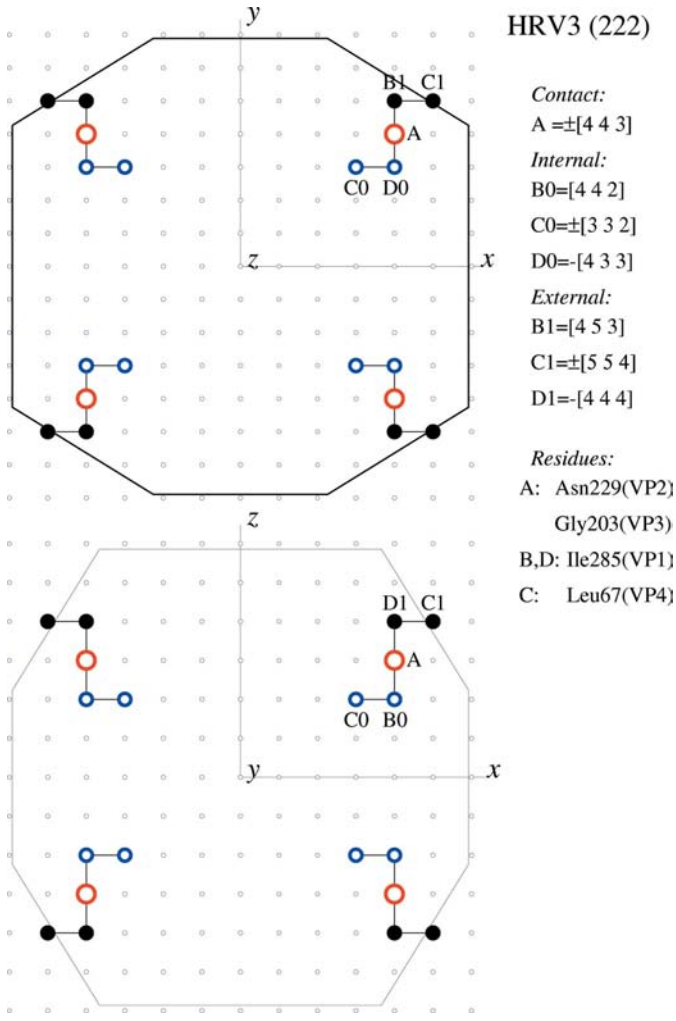

Figure 12
Fingerprint diagram for the orthorhombic HRV3 (with the same conventions as in Fig. 9).

position and a reduced number of the 60 chains implied by the icosahedral symmetry. These indexed positions allow one to define a fingerprint of the serotype. From this approach, however, no information could be obtained on the antiviral binding site, on which the classification of serotypes is based (see, for example, Kim *et al.*, 1993). It would be great to establish generic properties valid for other viruses as well. This is a first step towards a still missing general underlying theory. Here, new crystallographic and biochemical insights have been obtained, commented on from these two points of view.

From the crystallographic point of view, one should be aware that no higher-dimensional description has been used, not even for the ico-dodecahedron which is the enclosing form with icosahedral symmetry of the rhinovirus capsid (Janner, 2006*a*). This is based on the fact that the crystals involved have a three-dimensional space-group symmetry and are not icosahedral quasi-crystals. Moreover, the sphere, a natural molecular enclosing form for this virus, is here considered together with the cube (circumscribing the ico-dodecahedron), the latter being indexable in three dimensions. This leads to a

Table 6
Fingerprint residues indexed by the molecular lattice Λ_M.

C, I, E indicate contact, internal and external KPR residues, respectively.

Coat proteins	Type	HRV2 (222)		HRV3 (222)		HRV16 (222)		
VP2, VP3	C	[5 7 5]	[$\bar{5}$ $\bar{7}$ $\bar{5}$]	[4 4 3]	[$\bar{4}$ $\bar{4}$ $\bar{3}$]	[3 3 4]	[$\bar{3}$ $\bar{3}$ $\bar{4}$]	
VP1	I	[4 6 6]	[$\bar{3}$ $\bar{8}$ $\bar{5}$]	[4 4 2]	[$\bar{4}$ $\bar{3}$ $\bar{3}$]	[3 3 3]	[$\bar{3}$ $\bar{2}$ $\bar{3}$]	
	E	[7 6 5]	[$\bar{6}$ $\bar{8}$ $\bar{4}$]	[4 5 3]	[$\bar{4}$ $\bar{4}$ $\bar{4}$]	[3 4 5]	[$\bar{3}$ $\bar{3}$ $\bar{5}$]	
VP4	I	[2 5 5]	[$\bar{2}$ $\bar{5}$ $\bar{5}$]	[3 3 2]	[$\bar{3}$ $\bar{3}$ $\bar{2}$]	[2 3 1]	[$\bar{2}$ $\bar{3}$ $\bar{1}$]	[$\bar{3}$ $\bar{1}$ $\bar{3}$]
	E	[8 9 5]	[$\bar{8}$ $\bar{9}$ $\bar{5}$]	[5 5 4]	[$\bar{5}$ $\bar{5}$ $\bar{4}$]	[4 3 7]	[$\bar{4}$ $\bar{3}$ $\bar{7}$]	[3 5 5]

Coat proteins	Type	HRV1A (32)
VP1, VP3	C	[4 2 3]
VP2	I	[3 2 3]
	E	[5 3 3]
VP4	I	[3 1 2]
	E	[5 2 4]

Coat proteins	Type	HRV14 (23)
VP1, VP2, VP3	C	[3 3 0]
VP4	I	[2 2 0]
	E	[4 4 0]

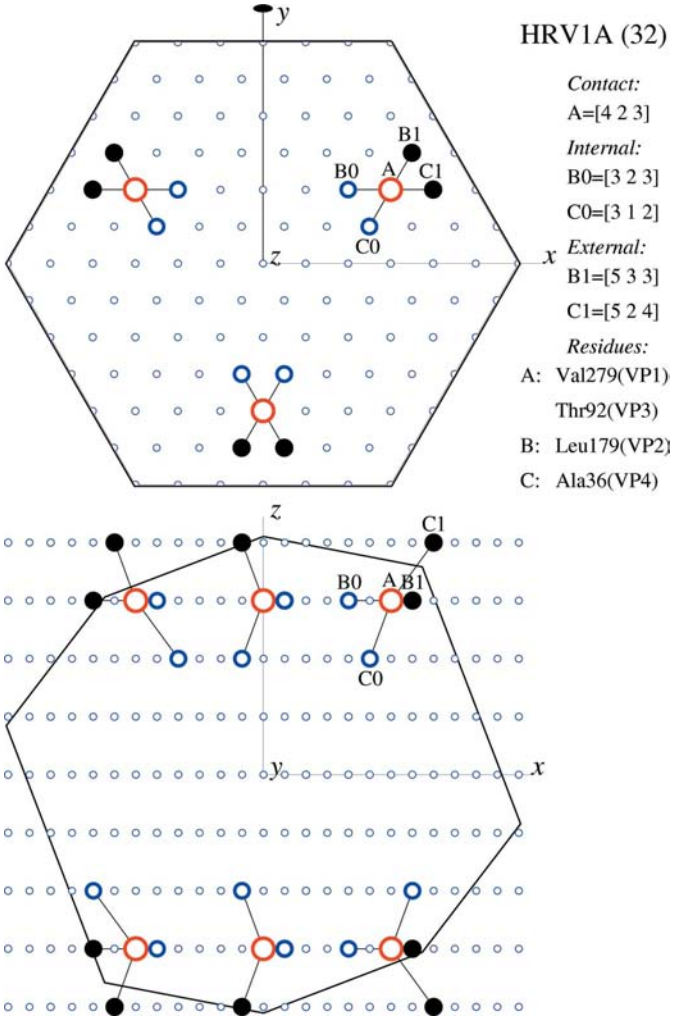

Figure 13
Fingerprint diagram for the hexagonal HRV1A (with the same conventions as in Fig. 9).

sphere packing, to kissing points and to packing lattices. As explained in Part I (Janner, 2010), a packing lattice has the crystal lattice as a sublattice and fits (more or less well) to the vertices of molecular enclosing forms mentioned. The molecular crystallographic properties, instead of being based on a form lattice Λ_F as in previous publications on biomacromolecules, are here expressed in terms of a three-dimensional molecular lattice Λ_M equal to the packing lattice Λ_P in a coarse-grained approximation only. In a similar coarse-grained way the position of a residue can be replaced by that of the nearest lattice point of Λ_P (or of Λ_M) and indexed accordingly. In this lattice-point approximation, one considers as contact residues those coinciding with kissing points, and *vice versa*.

The need for geometric rules not taking into account physico-chemical interactions is the reason why, in a similar way as for the contact residues, additional residues at minimal distance from the kissing points are considered, the so-called KPR residues. This set for each of the four coat proteins, including VP4, which is buried inside the capsid, has amino acids at fairly large distances (30 Å or more) from the kissing points and, therefore, is not directly involved in the lattice/clustering formation. As a result, the set of KPR residues, alternatively indexed by both the packing and the molecular lattices, defines at the same time a simple model of the capsid, and one of the crystal.

A structural interpretation could be given for the axial ratios of the crystal lattice parameters, a property which, in general, is still unexplained from the crystallographic point of view (Janner, 2004; de Gelder & Janner, 2005*a*,*b*).

More could be said on all these mutual relations, involving the icosahedral symmetry of the virus and no less than four different lattices (Λ, Λ_P, Λ_M, not forgetting Λ_F). It is typical that four Tables (Tables 4, 5, 6 and 7) were needed for specifying complementary aspects of the structure, even if

Table 7
Space-group-inequivalent positions occupied by KPR indexed residues.

(*a*) HRV2 (1aym), space group *I*222, Wyckoff 8(*k*)

Ile280 = I2129 (VP1)	Ser234 = S3913 (VP2), Pro203 = P5687 (VP3)	Leu43 = L6140 (VP4)
$[4, 6, 6] = [\frac{1}{5}\frac{3}{14}\frac{3}{10}]$, $[7, 6, 5] = [\frac{7}{20}\frac{3}{14}\frac{1}{4}]$	$[5, 7, 5] = [\frac{1}{4}\frac{1}{4}\frac{1}{4}]$	$[8, 9, 5] = [\frac{2}{5}\frac{9}{28}\frac{1}{4}]$, $[2, 5, 5] = [\frac{1}{10}\frac{5}{28}\frac{1}{4}]$

(*b*) HRV3 (1rhi), space group $P2_122_1$, Wyckoff 4(*c*)

Ile285 = I2128 (VP1)	Asn229 = N3884 (VP2), Glu203 = G5698 (VP3)	Leu67 = L6265 (VP4)
$[4, 0, 2] = [\frac{1}{4}0\frac{1}{6}]$, $[4, 1, 3] = [\frac{1}{4}\frac{1}{16}\frac{1}{4}]$	$[4, 8, 3] = [\frac{1}{4}\frac{1}{2}\frac{1}{4}]$, $[4, 0, 3] = [\frac{1}{4}0\frac{1}{4}]$	$[5, 7, 4] = [\frac{5}{16}\frac{7}{16}\frac{1}{3}]$, $[5, 1, 8] = [\frac{5}{16}\frac{1}{16}\frac{2}{3}]$
$[4, 8, 4] = [\frac{1}{4}\frac{1}{2}\frac{1}{3}]$, $[4, 9, 3] = [\frac{1}{4}\frac{9}{16}\frac{1}{4}]$		$[5, 1, 4] = [\frac{5}{16}\frac{1}{16}\frac{1}{3}]$, $[3, 9, 2] = [\frac{3}{16}\frac{9}{16}\frac{1}{6}]$

(*c*) HRV16 (1aym), space group $P22_12_1$, Wyckoff 4(*c*)

Glu38 = E282 (VP1)	Asn233 = N4128 (VP2), Pro201 = P5900 (VP3)	Asp44 = D6409 (VP4)
$[6, 4, 5] = [\frac{1}{2}\frac{1}{3}\frac{5}{16}]$, $[6, 3, 3] = [\frac{1}{2}\frac{1}{4}\frac{3}{16}]$	$[6, 3, 4] = [\frac{1}{2}\frac{1}{4}\frac{1}{4}]$, $[0, 3, 4] = [0\frac{1}{4}\frac{1}{4}]$	$[7, 3, 7] = [\frac{7}{12}\frac{1}{4}\frac{7}{16}]$, $[6, 5, 5] = [\frac{1}{2}\frac{5}{12}\frac{5}{16}]$
$[0, 3, 5] = [0\frac{1}{4}\frac{5}{16}]$, $[0, 2, 3] = [0\frac{1}{6}\frac{3}{16}]$		$[5, 3, 1] = [\frac{5}{12}\frac{1}{4}\frac{1}{16}]$, $[1, 9, 1] = [\frac{1}{12}\frac{3}{4}\frac{1}{16}]$
		$[1, 3, 1] = [\frac{1}{12}\frac{1}{4}\frac{1}{16}]$, $[0, 1, 3] = [0\frac{1}{12}\frac{3}{16}]$

(*d*) HRV14 (1rhv), space group $P2_13$, Wyckoff 12(*b*)

Asp91 = D625 (VP1), Gly138 = G3356 (VP2), Thr231 = T6103 (VP3)	Ser40 = S6440 (VP4)
$[3, 3, 6] = [\frac{1}{4}\frac{1}{4}\frac{1}{2}]$, $[3, 3, 0] = [\frac{1}{4}\frac{1}{4}0]$	$[2, 2, 0] = [\frac{1}{6}\frac{1}{6}0]$, $[4, 4, 0] = [\frac{1}{3}\frac{1}{3}0]$
	$[2, 4, 6] = [\frac{1}{6}\frac{1}{3}\frac{1}{2}]$, $[4, 2, 6] = [\frac{1}{3}\frac{1}{6}\frac{1}{2}]$

(*e*) HRV1A, Space group $P6_322$

Wyckoff 6(*g*)	Wyckoff 12(*i*)	
Val279 = V2193 (VP1), Thr92 = T4933 (VP3)	Leu179 = L3587 (VP2)	Ala36 = A6169 (VP4)
$[6, 0, 0] = [\frac{1}{2}00]$	$[6, 1, 0] = [\frac{1}{2}\frac{1}{12}0]$	$[6, 1, 5] = [\frac{1}{2}\frac{1}{12}\frac{5}{12}]$

their content is essentially the same. The data of Table 4 refer to the real C_α position of the residues, those of Table 5 to coarse-grained positions indexed according to the packing lattice Λ_P, those of Table 6 are indexed with respect to the molecular lattice Λ_M and those of Table 7 indicate representatives of the occupied Wyckoff positions of the space group leaving the crystal model invariant. In each of these cases the presentation takes the corresponding symmetries into account. As a result, the mutual translation of the content of these tables, while in principle straightforward, requires a fair amount of work. Non-trivial crystallographic relations could then be obtained, like the molecular interpretation of the axial ratios for the crystal lattices of the various serotypes and the point-group symmetry of the fingerprint diagrams.

From the biochemical point of view, what is most remarkable is the identification of a set of four residues (one for each coat protein at a reduced number of chains among the 60 required by the icosahedral symmetry) allowing characterization of not only the serotype, but also the KPR residues, the indexed capsid and crystal models and, last but not least, the residues responsible for the assembling process from the viral capsid to the packed crystal. In particular, one learns that in all cases but one, two coat proteins only are involved in the cluster formation through their contact residues: VP2 and VP3 in HRV2, HRV3 and HRV16; VP1 and VP3 in HRV1A; whereas all three surface coat proteins interact in HRV14. This last case is a bit suspicious, because the pair of neighbouring residues, both Asp91 of VP1, are of the charged type, which seems not favourable for binding. In all the other cases there are four interacting residues at the given kissing point location: two of them are polar amino acids (Ser, Asn or Thr) and two are hydrophobic (Pro, Gly or Val) (see Table 5).

As pointed out by one of the referees, the assembly of the viruses into a lattice-periodic structure is not a self-assembling process, because it requires an experimental intervention. The fingerprint diagram, however, reveals that the crystal structure is, so to say, encoded in each individual virus, even if in a coarse-grained lattice approximation only. The question arises whether an appropriate modification which reduces the mismatch between the viral structure and its fingerprint diagram would enhance the crystallization process towards a self-assembling one.

A more direct analysis of the relation between serotype variation and conserved architecture, as discussed for the enclosing forms of the various axial-symmetric chain clusters, would require a reformulation in terms of (selected) residues responsible for the conserved architectural elements.

From the observed close connection between serotype and crystal structure arises the conjecture of a common evolution. From the crystallographic point of view, this seems to be a daring idea. It becomes more natural if one considers the existence of viral polyhedra. These are 'crystals of the viral polyhedrin protein which form a thought matrix protecting virus particles'. The quoted text has been taken from the abstract P04.13.311 of a poster presented at the IUCr 2008 Assembly in Osaka (Coulibaly *et al.*, 2008) based on a previous publication (Coulibaly *et al.*, 2007). There the baculovirus and the cypovirus trigger the crystallization of viral polyhedra. In this case, one can speak of a self-assembling process leading to a co-crystal of virus and viral polyhedra, because it occurs without experimental intervention. Quoting again these authors: 'This evolutionary convergence to very similar crystalline architectures from different building blocks is reminiscent of the wide use of the icosahedral symmetry in virus particles'.

An older publication on virus polyhedra deserves to be mentioned (Carstens *et al.*, 1986) because of the identification of a specific amino acid, essential for the polyhedrin crystallization into cubic virus polyhedra. The views one finds in these works are similar to those presented here.

Among the many possible investigations on the relation between forms, symmetry and packing, the next step, important in order to recognize the generic properties related to variability and conservation of viral structural properties, is the extension of the present approach to the family of picorna viruses, to which the human rhinovirus belongs. The application of the present approach to the co-crystals of viral polyhedrin and baculovirus (or cypovirus, respectively) should allow us to gain an understanding of the possible relation between the fingerprint encoding and a self-assembling crystallization.

Thanks are expressed to my daughter Anna-Maria for critical remarks and constructive suggestions.

References

Arnold, E. & Rossmann, M. G. (1990). *J. Mol. Biol.* **211**, 763–801.
Carstens, E. B., Krebs, A. & Gallerneault, C. E. (1986). *J. Virol.* **58**, 684–688.
Conway, J. H. & Sloane, N. J. A. (1988). *Sphere Packings, Lattices and Groups*, ch. 1. Berlin: Springer.
Coulibaly, F., Chiu, E., Ikeda, K., Gutmann, S., Haebel, P. W., Schulze-Briese, C., Mori, H. & Metcalf, P. (2007). *Nature* (*London*), **446**, 97–101.
Coulibaly, F., Chiu, E., Ikeda, K., Gutmann, S., Haebel, P. W., Schulze-Briese, C., Mori, H. & Metcalf, P. (2008). *Acta Cryst.* A**64**, C328.
Coxeter, H. S. M. (1963). *Regular Polytopes*, p. 60. New York: Macmillan.
Gelder, R. de & Janner, A. (2005*a*). *Acta Cryst.* B**61**, 287–295.
Gelder, R. de & Janner, A. (2005*b*). *Acta Cryst.* B**61**, 296–303.
Hadfield, A. T., Lee, W.-m., Zhao, R., Oliveira, M. A., Minor, I., Rueckert, R. R. & Rossmann, M. G. (1997). *Structure*, **5**, 427–441.
Janner, A. (2004). *Acta Cryst.* A**60**, 198–200.
Janner, A. (2006*a*). *Acta Cryst.* A**62**, 270–286.
Janner, A. (2006*b*). *Acta Cryst.* A**62**, 319–330.
Janner, A. (2008). *Comput. Math. Methods Med.* **9**, 167–173.
Janner, A. (2010). *Acta Cryst.* A**66**, 301–311.
Kim, S., Smith, Th. J., Chapman, M. S., Rossmann, M. G., Pevear, D. C., Dutko, F. J., Felock, P. J., Diana, G. D. & McKinlay, M. A. (1989). *J. Mol. Biol.* **210**, 91–111.
Kim, K. H., Willingmann, P., Gong, Z. X., Kremer, M. J., Chapman, M. S., Minor, I., Oliveira, M. A., Rossmann, M. G., Andries, K., Diana, G. D., Dutko, F. J., McKinlay, M. A. & Pevear, D. C. (1993). *J. Mol. Biol.* **230**, 206–227.
Koch, E. & Fischer, W. (1999). *International Tables for Crystallography*, Vol. C, Section 9.1, edited by A. J. C. Wilson & E. Prince, pp. 738–743. Dordrecht: Kluwer Academic Publishers.
Verdaguer, N., Blaas, D. & Fita, I. (2000). *J. Mol. Biol.* **300**, 1179–1194.
Zhao, R., Pevear, D. C., Kremer, M. J., Giranda, V. L., Kofron, J. A., Kuhn, R. J. & Rossmann, M. G. (1996). *Structure*, **4**, 1205–1220.