

# Principles of Helix-Helix Packing in Proteins: The Helical Lattice Superposition Model

Dirk Walther<sup>1\*</sup>, Frank Eisenhaber<sup>1,2</sup> and Patrick Argos<sup>1</sup>

<sup>1</sup>European Molecular Biology Laboratory, Meyerhofstraße 1 Postfach 10.2209, 69012 Heidelberg, Germany

<sup>2</sup>Biochemisches Institut der Charité, der Humboldt-Universität zu Berlin, Hessische Straße 3–4 10115 Berlin, Germany

The geometry of helix-helix packing in globular proteins is comprehensively analysed within the model of the superposition of two helix lattices which result from unrolling the helix cylinders onto a plane containing points representing each residue. The requirements for the helix geometry (the radius  $R$ , the twist angle  $\omega$  and the rise per residue  $\Delta$ ) under perfect match of the lattices are studied through a consistent mathematical model that allows consideration of all possible associations of all helix types ( $\alpha$ -,  $\pi$ - and  $3_{10}$ ). The corresponding equations have three well-separated solutions for the interhelical packing angle,  $\Omega$ , as a function of the helix geometric parameters allowing optimal packing. The resulting functional relations also show unexpected behaviour. For a typically observed  $\alpha$ -helix ( $\omega = 99.1^\circ$ ,  $\Delta = 1.45 \text{ \AA}$ ), the three optimal packing angles are  $\Omega_{a,b,c} = -37.1^\circ$ ,  $-97.4^\circ$  and  $+22.0^\circ$  with a periodicity of  $180^\circ$  and respective helix radii  $R_{a,b,c} = 3.0 \text{ \AA}$ ,  $3.5 \text{ \AA}$  and  $4.3 \text{ \AA}$ . However, the resulting radii are very sensitive to variations in the twist angle  $\omega$ . At  $\omega_{\text{triple}} = 96.9^\circ$ , all three solutions yield identical radii at  $\Delta = 1.45 \text{ \AA}$  where  $R_{\text{triple}} = 3.46 \text{ \AA}$ . This radius is close to that of a poly(Ala) helix, indicating a great packing flexibility when alanine is involved in the packing core, and  $\omega_{\text{triple}}$  is close to the mean observed twist angle. In contrast, the variety of possible theoretical solutions is limited for the other two helix types. Besides the perfect matches, novel suboptimal “knobs into holes” hydrophobic packing patterns as a function of the helix radius are described. Alternative “knobs onto knobs” and mixed models can be applied in cases where salt bridges, hydrogen bonds, disulphide bonds and tight hydrophobic head-to-head contacts are involved in helix-helix associations.

An analysis of the experimentally observed packings in proteins confirmed the conclusions of the theoretical model. Nonetheless, the observed  $\alpha$ -helix packings showed deviations from the  $180^\circ$  periodicity expected from the model. An investigation of the actual three-dimensional geometry of helix-helix packing revealed an explanation for the observed discrepancies where a decisive role was assigned to the defined orientation of the  $C^\alpha$ - $C^\beta$  vectors of the side-chains. As predicted from the model, helices with different radii (differently sized side-chains in the packing core) were observed to utilize different packing cells (packing patterns). In agreement with the coincidence between  $R_{\text{triple}}$  and the radius of a poly(Ala) helix, Ala was observed to show greatest propensity to build the packing core. The application of the helix lattice superposition model suggests that the packing of amino acid residues is best described by a “knobs into holes” scheme rather than “ridges into grooves”. The various specific packing modes made salient by the model should be useful in protein engineering and design.

© 1996 Academic Press Limited

**Keywords:** protein; helix; protein folding; helix packing; protein secondary structure

\*Corresponding author

## Introduction

The topic of helix-helix pairwise packing in proteins was addressed soon after helical structures

had been suggested. Several models were developed and were mostly devoted to surface complementarities upon packing. Crick's model (Crick, 1953), later referred to as “knobs into holes”,

introduced the unrolling of regular helices onto a plane and then finding the best fit of the resulting lattices (one point per residue). This was achieved by superposition in a face-to-face manner through rotation followed by translation such that residues of one helix (knobs) fit into cells formed by neighbouring residues in the other helix (holes). Assuming a helix radius  $R$  of 5.0 Å and a twist angle  $\omega$  of 100.0° between residues along the helix path, he found optimal packing at a dihedral packing angle  $\Omega$  between the helix axes at +20° (coiled-coil structures) and a suboptimal packing at  $\Omega = -70^\circ$ . Richmond & Richards (1978) also pursued the knobs into holes model and concluded further that the packing angle is inversely correlated to the helix radius. They suggested three possible classes of helix-helix packing and, for each class, listed possible amino acids central to the contact. These preferences were utilized to predict spatial helical arrangements from primary structural information (Richmond & Richards, 1978; Cohen *et al.*, 1979; Cohen & Kuntz, 1987).

Chothia *et al.* (1977, 1981) introduced another and now widely accepted interpretation of the superimposed "helical" lattices. Instead of "knobs into holes" packing, they coined "ridges into grooves". Here, the ridges formed by residues with sequential spacing  $i$  in the first helix fit into grooves formed by residues in the second helix with spacing  $j$ . By assuming mean observed helix geometries, they found three basic packing types by varying  $i$  and  $j$ ; namely,  $\Omega_{i=1,j=4} = -105^\circ$ ,  $\Omega_{i=4,j=4} = -52^\circ$  and  $\Omega_{i=3,j=4} = +23^\circ$ . In principle, yet other combinations of  $i$  and  $j$  were possible (e.g.  $\Omega_{i=3,j=3} = -109^\circ$ ); however, as they noted, these classes were barely distinguishable from the former because of their similar packing angle and pattern of amino acid contacts. They introduced yet another packing class ("crossed ridge" packing), where the ridges of two helices cross with expected packing angles at +55°, -15° and -105°. Chothia and his co-workers also argued that the observed preference for packing angles around  $\Omega = -52^\circ$  can be understood in that ridges, formed by contact residues spaced by  $i = 4$ , dominate the shape and surface of the helical face since they make the smallest angle to the helix axis.

Efimov (1979) attempted to relate the packing angle with preferred rotational states of the side-chains along the helix. He distinguished two types of packing; polar and apolar, each giving rise to different combinations of rotational isomeric states of the contacting amino acid residues. For a best fit, he proposed three discrete packing angles for the apolar case ( $\Omega \approx +30^\circ$ ,  $\approx -30^\circ$ ,  $\approx 90^\circ$ ) and a range of possible docking angles in the polar case ( $-30^\circ \leq \Omega \leq 30^\circ$ ). Reddy & Blundell (1993) correlated the distance of closest approach between two packed helices to the volume of the interface-forming amino acid residues and used the resulting linear dependency to predict the interhelical distance of structurally

equivalent helices in homologous proteins. Other efforts have focused on the energetic aspects of helix-helix packing where different interaction potentials ranging from burial of hydrophobic residues (Ptitsyn & Rashin, 1975) and other simplified interaction potentials (Solovyov & Kolchanov, 1984) to atomic energy minimization and Monte-Carlo sampling (Chou *et al.*, 1983, 1984; Tufféry & Lavery, 1993) have been applied. Murzin & Finkelstein (1988) attempted to predict the topology and orientation of certain helical assemblies by arranging them in polyhedral shells. Harris *et al.* (1994) have performed a careful study of the diversity in four-helix bundle proteins.

The work presented here was stimulated by the observation that observed helix-helix packing angles demonstrate a pronounced preference for  $\Omega \approx -50^\circ/130^\circ$ . It is difficult to imagine why this preference should be a result of the relative length of one ridge along one helical side, as argued by Chothia *et al.* (1981), or due to the less splayed character of residues in the  $i = 4$  ridge (Chothia *et al.*, 1981; Hutchinson *et al.*, 1994). The contact-forming residues in helix association need not belong to one and the same ridge. Maximizing the burial of hydrophobic surface upon contact (presumably favouring smaller packing angles) or an easier and fitter packing of amino acid side-chains at a certain packing angle would seem to provide more natural explanations. Thus, the model of unrolled helix lattices was further investigated and treated mathematically in a rigorous fashion. Which set of helix parameters (the radius of the helix  $R$ , twist angle  $\omega$  and the rise per residue  $\Delta$ ) guarantees an optimal match and association of two identical and ideal helical lattices in a face-to-face manner after translating one of them (homogeneous packing)? Can the ambiguities in the ridges-into-groove model be resolved by considering optimization of the packing density? To approach these questions, the conditions for optimal packing were mathematically formulated to allow careful consideration of all solutions. To check the theoretical model, a statistical analysis of experimentally determined helix-helix packings was effected. The latter showed that a 180° periodic selection in  $\Omega$  was not uniform. An explanation for this is provided here based on the tertiary structural configuration of helices, especially the  $C^\alpha$ - $C^\beta$  bond direction. To the authors' knowledge, the treatment here is mathematically rigorous in contrast to all the previous works where more visual approaches were adopted and various helical geometric parameters were held fixed. The non-uniform  $\Omega$  distribution has not been previously addressed. The various optimal and suboptimal packing modes made salient by the model should aid in protein engineering and design, especially in selection of residue types to achieve specific helical contact sites or axial orientations.

**Table 1.** Definition of symbols

Symbol	Definition
$R$	Radius of a helix
$\Delta$	Rise per residue along the helix axis
$\omega$	Angular twist per residue along the helix path
$\Omega$	Dihedral packing angle (sign conventions as in Chothia <i>et al.</i> , 1981)
$a, b, c$	Identifiers for the three optimal solutions of the functions describing the lattice superposition
$a_m, b_m, c_m$	Identifiers for the three optimal solutions of the model equations where the mean values of $\Delta$ and $\omega$ are taken from helices observed in protein tertiary structures
$\Omega_N$	Dihedral packing angle using the $180^\circ$ rotation symmetry of ideal helix-helix packing; i.e. $\Omega_N = \Omega + 180^\circ$ if $\Omega < 0^\circ$ ; otherwise $\Omega_N = \Omega$
$\tau$	Angle measuring the deviation of the helix axis from the contact plane of the helix pair; i.e. the plane normal to the line of closest approach. The angle is non-zero when, for straight or curved helix axes, the line of closest approach is not perpendicular to at least one of the respective helix axes and crosses the axes at the helical termini
$d$	Distance of closest approach between two fitted helix axes
$d_i$	Distance of closest approach between a local helix axis assigned to the residue $i$ of the first helix and the second fitted helix axis; i.e. the shortest distance between the position obtained by drawing a perpendicular from the geometric centre of the side-chain atoms of residue $i$ ( $C^\alpha$ for Gly) to its fitted helix axis to the second fitted helix axis
$\alpha$	Skew angle (Harris <i>et al.</i> , 1994) between a vector, obtained by drawing a perpendicular from the geometric centre of the side-chain ( $C^\alpha$ for Gly) to its fitted helix axis and the local line of closest approach between the interacting helices
$P_c$	Site of contact between a residue of one helix and a second helix; i.e. the geometric centre of the positions of two side-chain atoms of the residue of the first helix that are closest to the second helix axis ( $C^\alpha$ only for Gly and $C^\beta$ only for Ala)
$P_{tip}$	Position of the side-chain atom of a contacting helical residue that is furthest from the fitted helix axis ( $C^\alpha$ for Gly)
$R_a$	Apparent helix radius defined as the distance of the $P_{tip}$ -atom to the fitted helix axis; atomic radii were not considered and were assumed to be compensated by side-chain-side-chain interdigitation upon packing

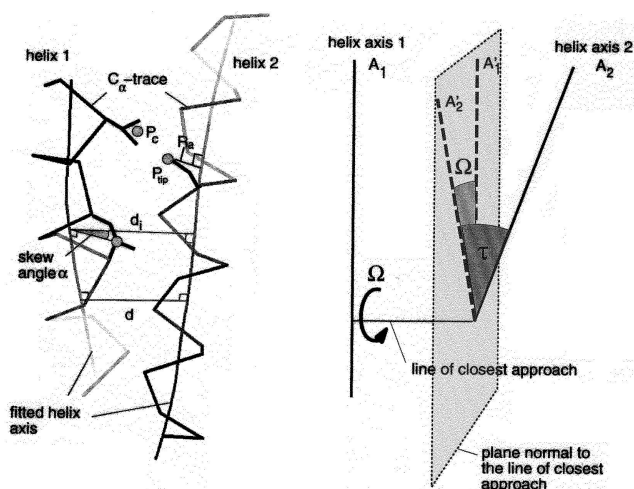
Several of these parameters are illustrated by Figure 1.

## Mathematical Description

### Homogeneous (hydrophobic) packing

#### Optimal (perfect) packing

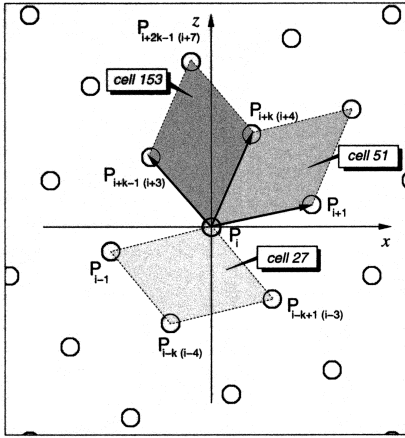
The model used here assumes regular and straight helices of radius  $R$ , twist angle  $\omega$  between successive residues along the helix path and a rise



**Figure 1.** Schematic drawing of the parameters used for the description of helix-helix packing geometries.  $A_1$  and  $A_2$  correspond to the helix axes projected onto the contact plane, which is normal to the line of closest approach. Definitions of the parameters and associated symbols are given in Table 1.

per residue  $\Delta$  along the helix axis. Various symbols utilized throughout the text are listed in Table 1 and their definitions are illustrated in Figure 1. Unrolling an ideal helix onto a plane towards the observer results in a regular lattice, as shown in Figure 2, where each point represents a residue. In associating  $\alpha$ -helices, each of the same geometry, one lattice must be rotated relative to the other about a lattice position such that the points of the two lattices overlap. Then an appropriately chosen translation of one lattice must be effected so that the knobs (points in one lattice) fall into the centre of parallelograms (holes) in the other helix (Figure 2). The parallelograms are formed by connecting four neighbouring points in one of the lattices. This packing optimization, where infinite lattices of unrolled helices overlap, is justified by the assumption that the global two-dimensional optimum coincides with the best possible local packing optimum in three dimensions.

This phenomenon can be mathematically formulated. Each lattice can be respectively described with two base vectors ( $\mathbf{v}_1$  and  $\mathbf{v}_2$ ;  $\mathbf{v}_1''$  and  $\mathbf{v}_2''$ ). In face-to-face packing of the helices, the base vectors are related through mirror symmetry:  $v_{x,1,2}' = -v_{x,1,2}$  and  $v_{z,1,2}' = v_{z,1,2}$ , where  $x$  and  $z$  represent respective vector components and the mirror plane contains the  $z$ -axis (Figure 2). Superposition requires rotation of one lattice such that  $\mathbf{v}_{1,2}'' = R_\Omega \mathbf{v}_{1,2}'$ , where  $R_\Omega$  is a rotation matrix corresponding to two helices with axial packing angle  $\Omega$ . The lattice point  $P_i$  is the centre of rotation (Figure 2). Under the condition of perfect



**Figure 2.** Detail of a helical lattice created by unrolling an infinite regular helix onto a plane towards the observer and specified by an  $xz$ -coordinate system. The origin is set onto one lattice point  $P_i$  representing residue  $i$ . The vectors correspond to possible base vectors for the lattice. The indices of the points refer to respective amino acid sequence positions along the helix relative to residue  $i$ , where  $k = [2\pi/\omega] + 1$ , where  $2\pi/\omega$  is made integral through truncation and  $\omega$  is in radians. The sequence separations for average  $\alpha$ -helices are given in parentheses. Two of any of the three base vectors shown can be chosen to specify the lattice (three possibilities). The grey coloured parallelograms correspond to the three possible topologically possible packing cells (holes) into which the lattice points (knobs) can be fit, resulting in helix-helix packing. The identifier of a given cell is calculated from indices of lattice points associated with the cell, which are summed after becoming powers to the base 2, and  $k$  is assumed as 4; for example, a cell bonded by  $i, i+3, i+4$  and  $i+7$  is identified by  $153 = 2^0 + 2^3 + 2^4 + 2^7$ ; similarly for cell 27 ( $2^0 + 2^1 + 2^3 + 2^4$ ); cell 51 and so forth.

overlap, the mirrored and rotated lattice can be described through a linear combination of the original base vectors ( $\mathbf{v}_{1,2}$ ) with four integer factors ( $n_1, n_2, n_3$  and  $n_4$ ) such that:

$$\mathbf{v}_1^i = n_1 \mathbf{v}_1 + n_2 \mathbf{v}_2 \quad (1)$$

$$\mathbf{v}_2^i = n_3 \mathbf{v}_1 + n_4 \mathbf{v}_2 \quad (2)$$

where  $n_{1..4} = 0, +1$  or  $-1$ .

Without loss of generality, two base vectors with respective components can be selected (Figure 2) such that:

$$\mathbf{v}_1 = \mathbf{P}_{i+1} - \mathbf{P}_i = \begin{pmatrix} AR \\ \Delta \end{pmatrix} \quad (3)$$

$$\mathbf{v}_2 = \mathbf{P}_{i+k} - \mathbf{P}_i = \begin{pmatrix} BR \\ k\Delta \end{pmatrix} \quad (4)$$

where  $A = \omega$ ,  $B = k\omega - 2\pi$  and  $k = (2\pi/\omega) + 1$ , where  $2\pi/\omega$  is truncated to an integral value and  $\omega$  is in radians. Since the helices are considered as cylinders, the  $x$ -co-ordinate corresponds to arcs on the cylinder. As shown in the Appendix, this system of equations yields three distinct solutions (Table 2) for the packing angle  $\Omega$ , corresponding to packing classes designated here as  $a, b$  and  $c$ . The solutions were found to possess a  $180^\circ$  periodicity as indicated by their signs. The functions for each class, corresponding to particular values of  $n_{1..4}$ , are subsequently given where the (b) relationships for  $R/\Delta$  are derived by squaring and summing the (a) relationships:

class  $a$ :

$$\left. \begin{aligned} \cos(\Omega) &= \pm \frac{(1-k)B + kA}{kA - B}, \\ \sin(\Omega) &= \pm \frac{R}{\Delta} \frac{2AB - B^2}{B - kA} \end{aligned} \right\} \quad (5a)$$

$$\left(\frac{R}{\Delta}\right)^2 = \frac{kA(2k-4) + kB(2-k)}{B^3 - 4AB^2 + 4BA^2} \quad (5b)$$

class  $b$ :

$$\cos(\Omega) = \pm \frac{A - kB}{kA - B}, \quad \sin(\Omega) = \pm \frac{R}{\Delta} \frac{A^2 - B^2}{B - kA} \quad (6a)$$

$$\left(\frac{R}{\Delta}\right)^2 = \frac{k^2 - 1}{A^2 - B^2} \quad (6b)$$

**Table 2.** The six solutions for optimal superposition of ideal helical lattices

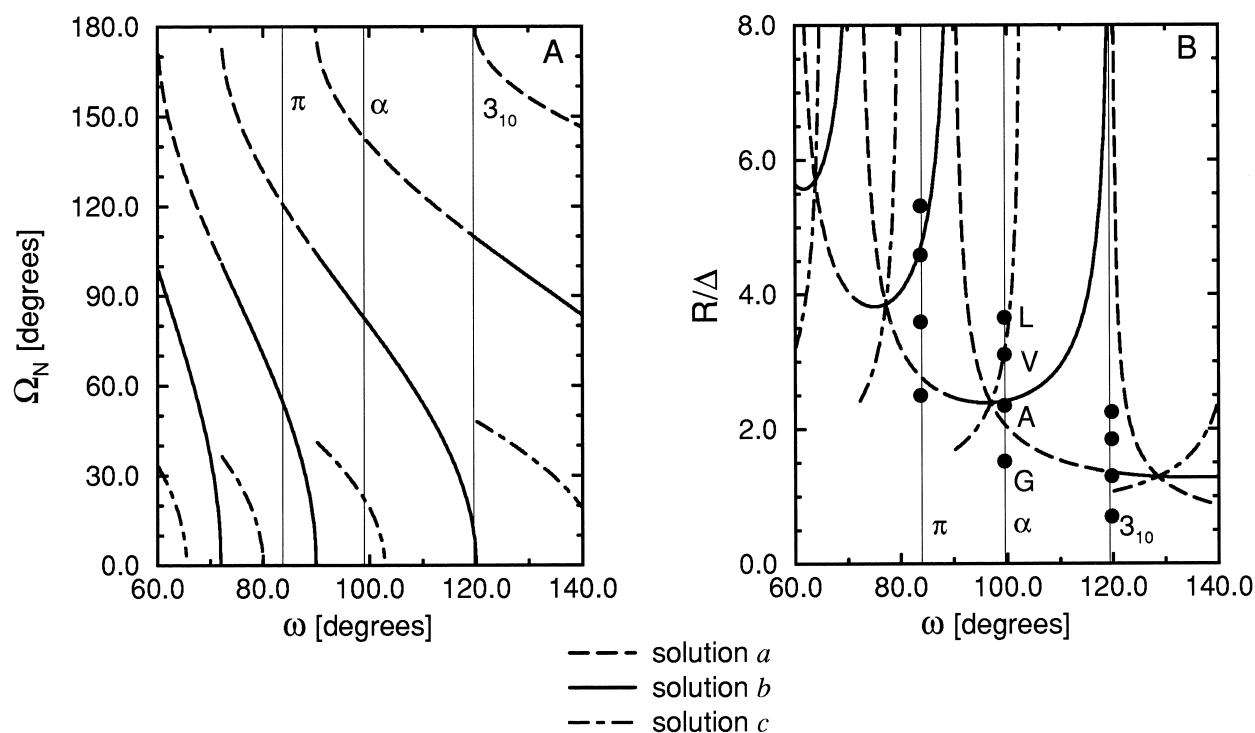
	$n_1$	$n_2$	$n_3$	$n_4$	Corresponding <sup>a</sup> solution equation:	$\Omega^b$ (deg.)	$R^c$ (Å)	Class <sup>d</sup> ident.
(1)	-1	1	0	1	(5a, b): + sign	-37.1	3.0	$a$
(2)	1	-1	0	-1	(5a, b): - sign	142.9	3.0	$a$
(3)	0	1	1	0	(6a, b): + sign	-97.4	3.5	$b$
(4)	0	-1	-1	0	(6a, b): - sign	82.6	3.5	$b$
(5)	-1	0	-1	1	(7a, b): + sign	22.0	4.3	$c$
(6)	1	0	1	-1	(7a, b): - sign	-158.0	4.3	$c$

<sup>a</sup> Solution corresponding to the given sign conditions in the specified equations.

<sup>b</sup> Values given for the packing angle assume a twist angle  $\omega = 99.1^\circ$ , which is the mean observed value in known protein structures.

<sup>c</sup> The helical radius given assumes  $\omega = 99.1^\circ$  and  $\Delta = 1.45$  Å, values most often observed in actual helices.

<sup>d</sup> Class identification (see the text).



**Figure 3.** Graphical representation for the three solutions (packing classes *a*, *b* and *c*) of optimally overlapped helical lattices. Calculated optimal packing angle  $\Omega_N$  and calculated ratio  $R/\Delta$  are shown as functions of the helix twist angle  $\omega$  for the given interval. The vertical lines correspond to the mean twist angles for the three helix types ( $\alpha - \langle\omega\rangle = 99.1 \pm 15.0^\circ$ ),  $3_{10}$  and  $\pi$ -helices). The filled circles in B correspond respectively to the ratios  $R/\Delta$  for different amino acids obtained by dividing the mean observed distances of the “tip” atoms ( $P_{\text{tip}}$ ) of the residues Gly, Ala, Val and Leu to the axis of  $\alpha$ -helices (i.e. the apparent helix radius) with the corresponding mean rise per residue of the respective helix type ( $\langle\Delta\rangle = 1.45 (\pm 1.1) \text{ \AA}$ ) for the  $\alpha$ -helix,  $2.0 \text{ \AA}$  ( $3_{10}$ -helix) and  $1.1 \text{ \AA}$  ( $\pi$ -helix)). The radii for the  $3_{10}$  and the  $\pi$ -helix were corrected for their different  $C^\alpha$  based radii ( $1.9$  and  $2.8 \text{ \AA}$ , respectively).  $\langle\omega\rangle_{\alpha\text{-helix}}$  was taken from the mean observed twist angle between the geometric centres of side-chain atoms for two consecutive helix residues. These mean values ( $\langle\omega\rangle$  and  $\langle\Delta\rangle$ ) did not include observations based on the four residues at either helix terminus where the helix axis is less accurately determined (see the text). Values for the  $3_{10}$  and  $\pi$ -helices were taken from Schulz & Schirmer (1979).

class *c*:

$$\left. \begin{aligned} \cos(\Omega) &= \pm \frac{B + (k-1)A}{kA - B}, \\ \sin(\Omega) &= \pm \frac{R}{\Delta} \frac{2AB - A^2}{B - kA} \end{aligned} \right\} \quad (7a)$$

$$\left(\frac{R}{\Delta}\right)^2 = \frac{A(2k-1) + B(2-4k)}{A^3 - 4A^2B + 4AB^2} \quad (7b)$$

It is clear that, for each class, the packing angle  $\Omega$  is a function of the twist angle  $\omega$ , as is the ratio  $R/\Delta$ .  $\Omega_N$  and  $R/\Delta$  can be plotted against  $\omega$  (Figure 3) for each class. As evident from the equations, each solution has a restricted definition space and there are singularities at different twist angles. For different  $k$  dependent upon  $\omega$ , the solutions repeat but have different periods. The solutions for all three classes recurrently cross at points (Figure 3) in the radius dependency (for the triple point closest to the mean observed twist angle for  $\alpha$ -helices  $\langle\omega\rangle_\alpha$ ,  $\omega_{\text{triple}} = 96.9^\circ$ ). The corresponding helical lattices are regular hexagons such that the three solutions for

the packing angle can accommodate the same helix radius. At  $\langle\omega\rangle_\alpha$ , the three packing classes display different radii. The smallest accommodated radius is found in packing class *a* ( $R = 3.0 \text{ \AA}$ ) and the largest in class *c* ( $R = 4.3 \text{ \AA}$ ) assuming the mean observed rise per residue  $\langle\Delta\rangle = 1.45 \text{ \AA}$ . Only for helices of the  $\alpha$ -type is the simultaneous occurrence of all three packing classes allowed and the mean observed value  $\langle\omega\rangle_\alpha$  is found closer to a triple point than the mean twist angle of the other helix types ( $\pi$  and  $3_{10}$ ). Thus, the helical lattice obtained from unrolling an average  $\alpha$ -helix more closely resembles a regular hexagonal lattice, which allows three packing angles simultaneously. The  $\pi$ -helix does not correspond to solution *c* and the mean twist angle for  $3_{10}$  helices ( $\langle\omega\rangle_{3_{10}} = 120.0^\circ$ ) coincides with an ambiguity where the solution switches from  $k = 4$  to  $k = 3$  with increasing  $\omega$ . Furthermore, the required helical radii corresponding to  $\omega$  intervals about the observed mean of  $\alpha$ -helices fall in a biologically reasonable range (Figure 3B), whereas for the other helix types, the required radius for one solution has to be either infinite (solution *c* for  $\pi$ -helices) or is ambiguously defined (solution *a* and *b* for  $3_{10}$

helices). Solution  $c$  for  $3_{10}$  helices is found just at the transition to an infinite radius.

### Suboptimal packing

Apart from analysing perfect lattice overlap, helix-helix axial association angles may be observed corresponding to local packing optima where, for example, base vectors are superimposed such that regularity is achieved in only one lattice direction. Since helices often pack over a few turns, the condition of infinite lattice superposition may be too strong and not fulfilled. Thus, only the nearest six neighbours around a central lattice point are considered to determine suboptimal packing. As before, the central points of two sublattices are brought into coincidence. Subsequently, one sublattice is rotated at the angle  $\Omega$  around the central point. A packing parameter  $S_P$  was constructed to measure the degree of overlap between the two sublattices:

$$S_P(\Omega_N, R) = \sum_m \min \left( \frac{|\mathbf{P}_m(R) - \mathbf{Q}_{n=1..6}(\Omega_N, R)|}{R} \right) \quad (8)$$

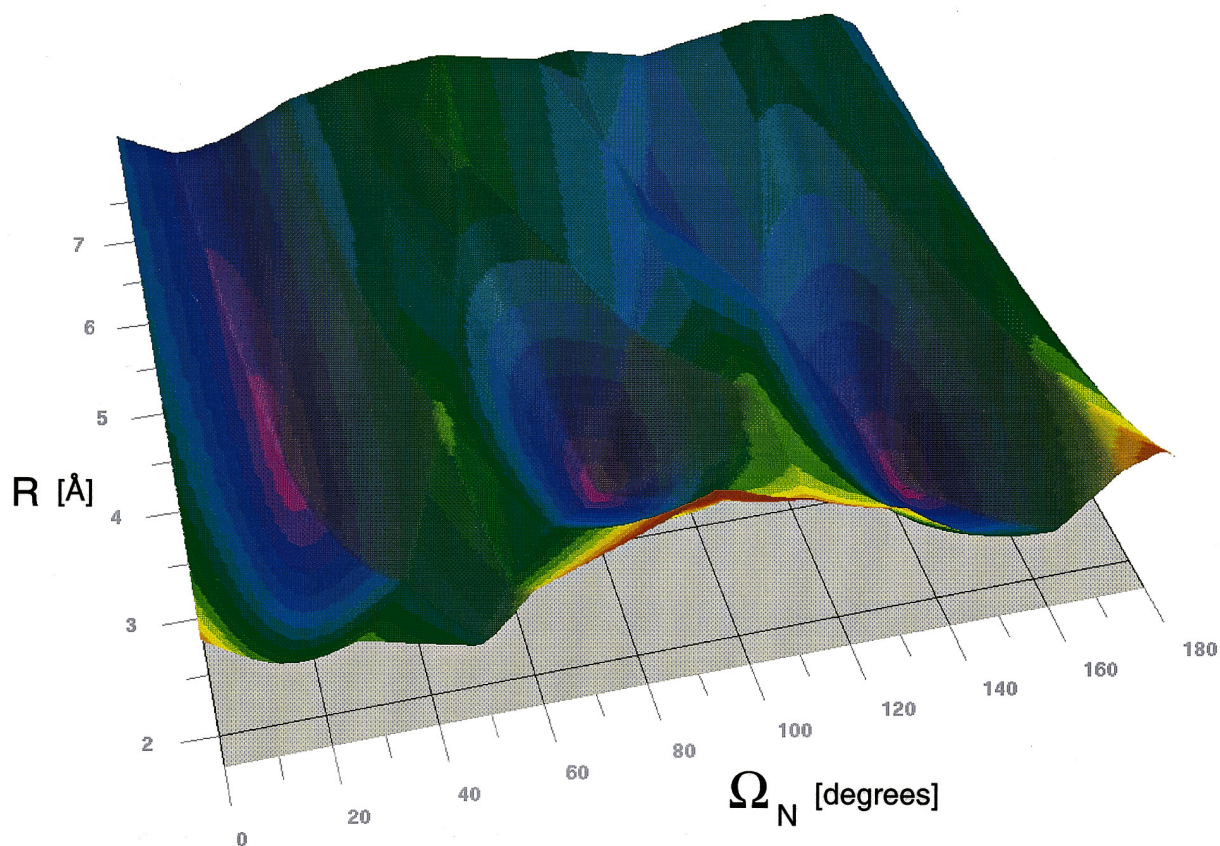
where  $m = 1..6$  corresponds to the six neighbouring lattice points of  $P_i$  (Figure 2), the lattice vectors to points  $P_m$  are held fixed and  $Q_n$  are the vector

positions in the mirrored and rotated lattice. The angle  $\Omega_N$  covers the range 0 to  $180^\circ$  and implies a second possible packing angle at  $\Omega = \Omega_N - 180^\circ$  due to the rotational symmetry of ideal helices. Since in the model an increased helix radius implies the same for the helix-forming side-chains, the distances between the lattice points were normalized to the helical radius ( $R$ ).

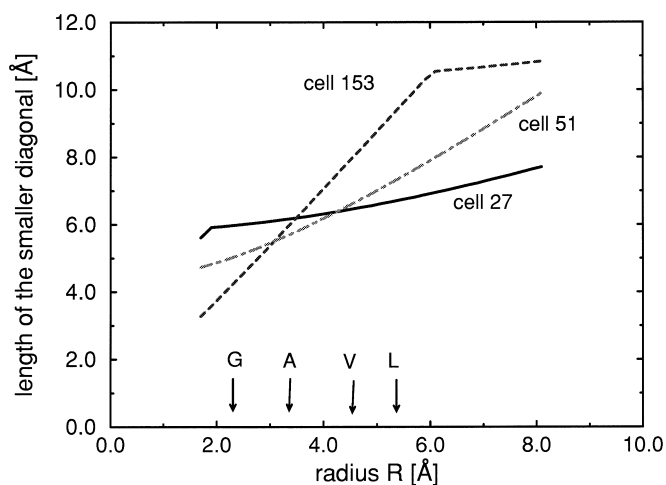
The behaviour of the function  $S_P(\Omega_N, R)$  is shown in Figure 4, where its value is plotted against  $\Omega_N$  with  $\Delta$  and  $\omega$  taken as the mean observed values for  $\alpha$ -helices. The three minima of the optimal packings ( $S_P(\Omega_N, R) = 0$ ) can be readily identified. For a given radius more than one packing angle meets the requirements of little steric clash, increasing the possible number of helix-helix packing angles. For example, at larger radii, in addition to the steep minimum at  $\Omega_N \approx 20^\circ$ , a broad but shallow minimum is found near  $\Omega_N = 120^\circ$ .

### Determination of the translation vector for the superimposed lattice

Only optimal superposition of lattice points through rotation has been thus far considered. For actual helix-helix packing, the translation vector by which one of the superimposed lattices is shifted to bring the knobs (lattice points) into holes (lattice



**Figure 4.** Three-dimensional plot of the function  $S_P = S_P(\Omega_N, R)$ . The colour spectrum corresponds to different isosurfaces where violet colors belong to minimal values of  $S_P$ . The twist angle and the rise per residue were respectively taken as the mean observed values,  $\omega = 99.1^\circ$  and  $\Delta = 1.45 \text{ \AA}$ . The grey shadows are shown merely for dimensional perspective and reflect the direction of the light source.



**Figure 5.** Size of the packing cells 27, 51 and 153 as a function of the helix radius. The size is defined by the length of the smaller diagonal associated with the respective packing cells. The twist angle was set to  $99.1^\circ$  and the rise per residue  $\Delta$  to  $1.45 \text{ \AA}$ , the respective average observed values. Arrows indicate observed radii of helices composed only of Gly (G), Ala (A), Val (V) and Leu (L).

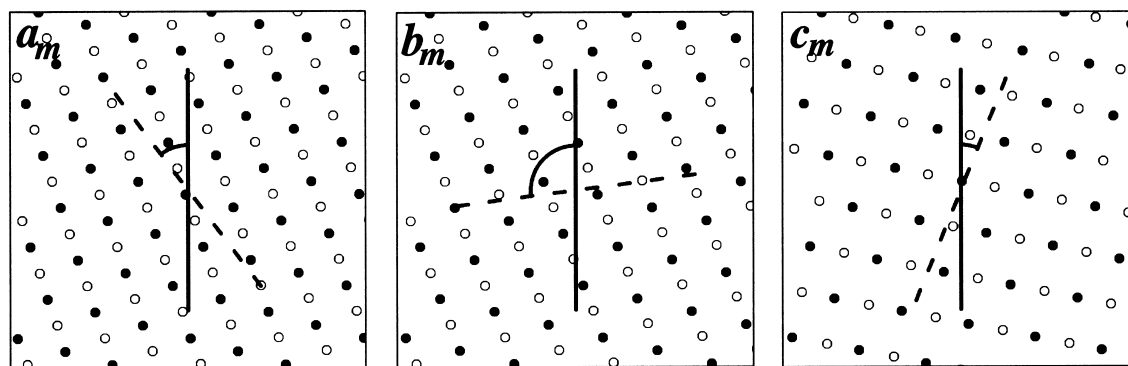
cells) must be applied. In accordance with the three topologically possible lattice cells (referred to as 153, 27 and 51), three translation vectors are possible where lattice points are shifted to their centres (Figure 2). The cellular designations are explained in the legend to Figure 2. To achieve the most homogeneous and dense packing, the largest possible cell must be selected for association with a side-chain of an interacting helix. The length of the smaller diagonal of each cell was chosen as a simple estimate of cellular size. (The area of an inscribed circle, an alternative, does not exist for parallelograms.) The plots of Figure 5 demonstrate that a cell's capacity depends on the helix radius and thus different cells are preferentially occupied at different helix radii. For helices with smaller radii, cell 27 should be favoured, while for helices with

bigger radii, side-chains should prefer to pack into cell 153. According to the size criterion adopted here, cell 51 is never the largest packing cell. For this reason, and since in cylindrically shaped helices cell 51 is mostly oriented away from the helix-helix interface, it will not be considered further.

Figure 6 shows the three possible perfectly superimposed helix lattices where the mean observed twist angle is taken. The selected packing cells are cell 27 for solutions  $a_m$  and  $b_m$  and cell 153 for  $c_m$ .

### Non-homogeneous packing

The knobs-into-holes model assumes that the amino acid side-chains pack isotropically into a hole formed by four side-chains of the second helix and with parallelogram cross-section. This might not be necessary if some other interaction joined the two helices, such as bonds formed between the associating amino acids, including disulphide bonds, salt bridges, hydrogen bonds or tight hydrophobic head-to-head contact (knobs onto knobs packing). If the packing site merely consisted of this type of contact alone, the preferred packing angles would remain the same as those derived from superposition of the helical lattices but no translation would be necessary. This situation is unlikely. Nonetheless, a mixture of knobs-into-holes and bonded contacts are yet possible. Chothia *et al.* (1981) have coined the term "crossed ridge helix packing" for these cases. The possible packing angles for this association type can be obtained by a consideration of suboptimal packing in the model developed here. The superposition of the two central lattice points can now be interpreted as a residue-residue bond of any type. Since the neighbouring residues should still obey the normal knobs-into-holes scheme but without shifting, the function  $S_p(\Omega_N, R)$  for the sublattice has now to be maximal instead of minimal. In agreement with the angles predicted by Chothia for the crossed ridge case, Figure 4 reveals three isolated maxima at



**Figure 6.** Helical lattices according to the theoretical solutions for packing classes  $a_m$ ,  $b_m$  and  $c_m$  at the mean observed parameters ( $\omega = 99.1$ ,  $\Delta = 1.45 \text{ \AA}$ ) and the corresponding packing angles  $\Omega(a_m, b_m, c_m) = -37.1^\circ, -97.4^\circ$  and  $+22.0^\circ$ , respectively. Starting from perfect superposition achieved by rotation of the mirrored lattice (open circles), one lattice was shifted to centre the lattice points in the appropriately chosen packing cells (see the text). The continuous (broken) line denotes the helix axis of the lattice with the filled (open) circles.

$\Omega_N \approx 55^\circ$ ,  $115^\circ$  and  $175^\circ$ . In addition, helices with large radii should also pack at  $\Omega_N \approx 75^\circ$ .

## Analysis of Observed Helix-Helix Associations

The theoretical model used here assumes that ideal helices pack. In the following section, the experimental verification of the conclusions drawn from the mathematical lattice superposition model is discussed.

### Data

A total of 220 protein tertiary structures, determined at 2.0 Å resolution or better and with mutual sequence similarity less than 35% as selected by the program OBSTRUCT (Heringa *et al.*, 1992, available via World-Wide-Web; URL:

[http://www.embl-heidelberg.de/obstruct/obstruct\\_info.html](http://www.embl-heidelberg.de/obstruct/obstruct_info.html)

were used for a statistical analysis of helix-helix association (the set is available upon request by e-mail to walther@embl-heidelberg.de). The assignments of the  $\alpha$ -helical stretches were taken from the program DSSP (Kabsch & Sander, 1983). The angle between two consecutive carbonyl bonds was not allowed to exceed  $65^\circ$ ; otherwise, the helix was divided into two at this residue. Two helices were defined to be in close contact if at least three residues of each helix had at least one interhelical atom-atom contact with maximal threshold distance of 4.5 Å between atom centres. The resulting dataset of proteins used in this study contained 687 closely packed pairs of helices. Membrane proteins were not included and only heavy atoms were considered.

### Definition of the helix axis

The definition of the helix axis from which many contact characteristics are measured bears critically on the results. Since helices can be bent, a procedure to fit a local helix axis,  $A_i$ , to every residue  $i$  along the helix was adopted. It takes advantage of a straightforward algorithm for the overall axial definition given by Chothia *et al.* (1981). The vector coincident with the local helix axis of residue  $i$ ,  $\mathbf{u}_i$ , can be determined from the cross product of the vectors  $\mathbf{B}_i$  and  $\mathbf{B}_{i+1}$  such that:

$$\mathbf{u}_i = \mathbf{B}_i \times \mathbf{B}_{i+1} \quad (9)$$

where:

$$\mathbf{B}_i = \mathbf{r}_i + \mathbf{r}_{i+2} - 2\mathbf{r}_{i+1} \quad (10)$$

and  $\mathbf{r}$  is the position vector of the  $C^\alpha$  atom in residue  $i$ . At the C terminus of the helix, where the residue indices would go beyond those in the helix, the local line vector is taken from the closest helical constituent residues. A point on the local axis  $A_i$  is assigned by calculating the geometric centre of the

closest four consecutive  $C^\alpha$  positions around the residue  $i$  (i.e.  $C_{j+i-1}^\alpha \dots C_{j+i+2}^\alpha$  where  $j=0$  for the inner helical residues and appropriately chosen at the helix termini and the points are correspondingly shifted along  $\mathbf{u}_i$ ). The length of the local axis is first set to 1.5 Å. The direction of the local axis  $A_i$  associated with residue  $i$  is then smoothed by taking the average direction of three consecutive local vectors centred at  $i$  (two at the helix ends). To achieve a continuous axial curve over the entire helix, the new starting and ending points of consecutive local lines are joined by calculating the middle point between the end point of the first local stretch and the starting point of the next local stretch. The new lengths and directions of the local axes are then recalculated. This smoothing procedure is repeated three times. Despite the simplicity of this algorithm, the improvement for the fit of the local axis is considerable. The standard deviation of the distances of each  $C^\alpha$  atom to the helix axis decreased from  $\sigma = 0.34$  Å for a globally defined axis (obtained by averaging the vectors  $\mathbf{u}_i$  over the whole helix and taking the geometric centre of every  $C^\alpha$  position) to  $\sigma = 0.14$  Å when using the local axes. When only the inner helical residues were considered (four residues subtracted at either helix termini), the accuracy was improved. The standard deviation decreased from  $\sigma = 0.37$  Å for the global axes to  $\sigma = 0.07$  Å for local axes.

The packing angles are positive if the background helix is rotated clockwise with respect to the frontal helix when facing them. The helices are parallel with respect to their sequence direction at  $\Omega = 0^\circ$ . The packing angles are sometimes normalized to the interval  $0^\circ < \Omega_N < 180^\circ$  because of the 2-fold rotation axis of ideal helices; i.e.  $\Omega_N = \Omega + 180^\circ$  if  $\Omega < 0^\circ$ ; otherwise  $\Omega_N = \Omega$ . The line of closest approach between two helices was calculated by determining the line of closest approach with minimum length amongst all local axes pairs. To ensure a face-to-face packing, distorted helix-helix associations where the line of closest approach intersected at the helix termini and  $\tau > 5^\circ$  (Table 1 and Figure 1) were omitted, leaving 449 helix pairs for analysis.

### Packing cell determination

The packing cell of a second helix utilized by a contacting residue in the first helix was determined by the sequence separation of four residues containing, respectively, one of the four closest atoms (one closest atom per residue) to the geometric centre of the two atoms in the contacting residue ( $C^\alpha$  for Gly and  $C^\beta$  for Ala) in the first helix that are closest to the axis of the second helix (position of contact,  $P_c$ ; see Figure 1 for illustration). To ensure a real packing conformation, the third closest residue of the second helix to the contacting residue in the first helix was required to be within a distance of 6.0 Å. Despite the imprecision of the cell-determining procedure, the observed ranking of cell usage is as predicted. The three topologically

possible cells (153, 27 and 51) were detected most often with respective counts 767, 647 and 228. Other determined cells such as 23 and 275 had frequencies of 48 and 42, and were followed by others.

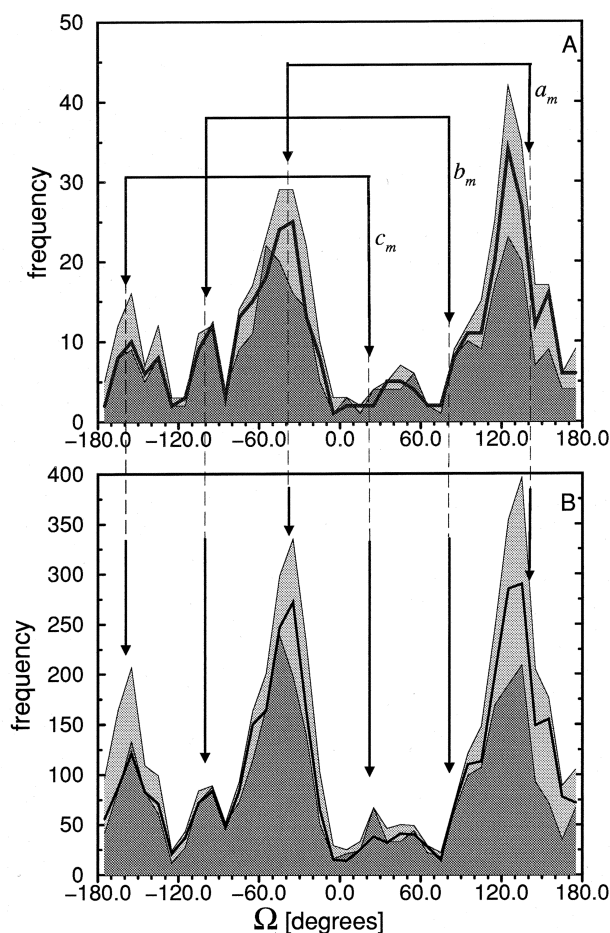
### Algorithm for interhelical "bond" determination

Interhelical bonds were determined on the basis of geometric pattern recognition. A bond was identified between two side-chains in different helices if their corresponding  $P_c$  sites were mutually the closest to each other. The  $P_c$  sites must be no more than 4 Å apart and the closest  $P_c$  site for other residues in the same helix must be 5 Å or greater. Furthermore, the angle between the local line of closest approach and the vector joining the two mutually closest positions  $P_c$  was required to be smaller than 45°. These conditions assured knobs-onto-knobs packing, and that identified residues literally faced each other and did not pack into a cell formed by the oppositely facing helix. For 95 helix-helix pairs, this definition was fulfilled; 80 such pairs had only one interhelical bond while 15 displayed two.

## Results

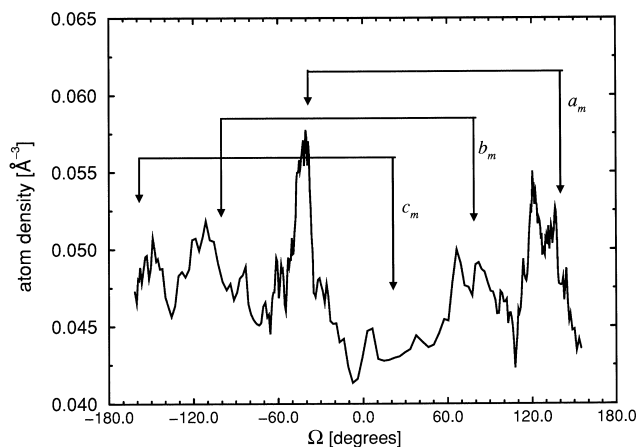
The distribution of the observed global (per helix-helix pair) and local (per amino acid residue along the two packed helices) dihedral packing angles in the selected set of proteins is shown in Figure 7. To a certain extent, the histogram of the local packing angles biases the observations to more parallel or antiparallel associations because of longer possible contact regions. Yet, it allows considerations at which angle packings are possible over a longer stretch where the lattice model is certainly more critical. To account for possible restrictions due to short loops connecting two successive helices along the chain, which disallows parallel packing, the condition of more than 20 intervening residues was applied in a second histogram. In a third histogram, all helix-helix pairs were used except those displaying interhelical bonds.

The two largest peaks occur in the intervals  $-70^\circ \leq \Omega \leq -20^\circ$  and  $+110^\circ \leq \Omega \leq +140^\circ$ . The medium peaks are found at  $-170^\circ \leq \Omega \leq -150^\circ$  and  $-110^\circ \leq \Omega \leq -90^\circ$ . Fewer helical pairs pack at  $+10^\circ \leq \Omega \leq +60^\circ$  and  $+160^\circ \leq \Omega \leq +180^\circ$ . As indicated by arrows in Figure 7, in the negative angular range the optimal solutions ( $a_m$ ,  $b_m$  and  $c_m$ ) of the helical lattice superposition model match the observed peaks well. In the positive range, the class  $a$  peak at  $\Omega \approx +142^\circ$  misses the observed peak by  $20^\circ$ , which rather corresponds to the angle of the predicted suboptimal solution for larger helix radii. The positive class  $b$  peak falls at a peak shoulder. The expected peak at  $\Omega \approx +22^\circ$  is little observed in the histogram of the global packing angles.



**Figure 7.** Frequency histogram for the observed dihedral helix-helix packing angle  $\Omega$  (bin width  $10^\circ$ ). A, Packing angle about the global line of closest approach. B, Histogram of local packing angles; i.e. the packing angle about the local line of closest approach defined for each contacting amino acid in the helix-helix pair (Figure 1). The light grey filled histogram corresponds to data based on all observed helix-helix pairs while the dark grey filled histogram was determined from those with more than 20 intervening residues between the end of one helix in a contacting pair and the beginning of the other helix. The third histogram (thick line) corresponds to all helix-helix pairs with no detected interhelical bond. Arrows show the predicted packing angles for the three optimal solutions ( $a_m$ ,  $b_m$  and  $c_m$ ) according to the theoretical model developed here; the mean observed twist angle was taken as  $99.1^\circ$ .

The correlation between frequencies of packing angles in the negative range to its periodic angle in the positive range; i.e.  $r_{f(\Omega < 0), f(\Omega + 180^\circ)}$ , was 0.74. The absence of a pronounced peak at  $\Omega \approx +22^\circ$  is perhaps explained by the repulsion of the dipole moments in nearly parallel helices. However, the importance of this effect is doubted by several authors, and theoretical and experimental results indicate that the effect may be limited (Chou & Zheng, 1992; Robinson & Sligar, 1993). Furthermore, in the histogram of the local packing angles (Figure 7B), the expected peak at  $\Omega \approx +22^\circ$  is better observed, indicating that the packings at larger

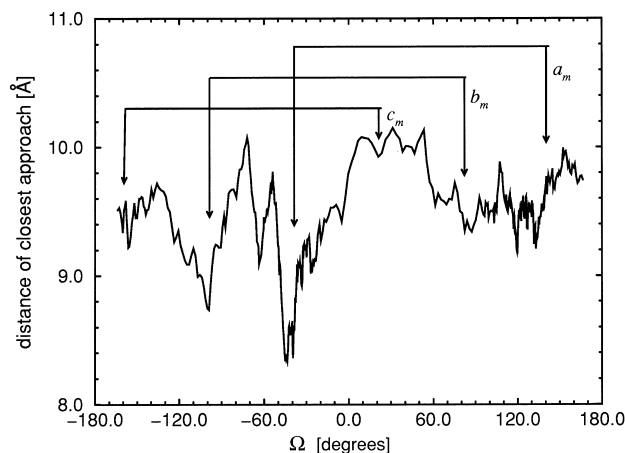


**Figure 8.** Observed density of interface atoms at different packing angles. The core region is defined by a sphere centred at the middle of the line of closest approach (length  $d$ ) between the two helices and with radius  $r = d/2$ . Only atoms belonging to contact-forming amino acids have been considered; i.e. at least one interhelical atom-atom contact was within 4.5 Å. Furthermore, it was required that all helix termini ( $C^\alpha$  positions) be outside the defined sphere. Arrows correspond to the optimal packing angles for classes  $a_m$ ,  $b_m$  and  $c_m$ . Results were obtained by a running average over 20 successive point clusters of the ordered helix data pairs with respect to  $\Omega_N$  with a mean standard deviation  $\langle\sigma\rangle = 0.013 \text{ \AA}^{-3}$ .

positive angles are accommodated by shorter contact regions for which the lattice approach is naturally less stringent. An examination of the packing parameter  $S_P$  (Figure 4) at  $\Omega_N \approx +60^\circ$  and  $\Omega \approx 0^\circ$  indicates steric unacceptability at all helical radii, which explains the paucity of observations near these packing angles where only interhelical bond packing is possible. Nonetheless, from the theoretical model described here, the observed distribution is likely to consist of a mixture of optimal and suboptimal solutions that are a function of the helical radii as well as knobs-onto-knobs packing cases. A detailed structural investigation will be subsequently described.

### Goodness of packing at different packing angles

Helix-helix interfaces should display a high atom density at their association site (core). In Figure 8 this property is shown as a function of packing angle. The packing core was defined by a sphere centred at the middle position of the line of closest approach and with a radius of half the length of the line ( $r = d/2$ ). The running average of the densities has pronounced maxima. Except packings at  $\Omega \approx 22^\circ$  (very few data), high densities are generally achieved in agreement with the predicted optimal packing classes  $a_m$ ,  $b_m$  and  $c_m$  (i.e. highest possible homogeneous packing density), which are derived from the theoretical model by using  $\omega = 99.1^\circ$ , the mean observed value. The peak at  $\Omega \approx 140^\circ$  is broad and extends to  $\Omega = +120^\circ$ , a suboptimal packing



**Figure 9.** Radius dependency as a function of the dihedral packing angle as expressed through a 30-point running average of the distances of closest approach ordered with ascending  $\Omega$  values. Arrows correspond to the optimal packing angles for classes  $a_m$ ,  $b_m$  and  $c_m$ . The mean standard distribution for the 30-point clusters was 1.2 Å.

angle of helices with large radii (*vide infra*). Interestingly, the highest core densities are found at the most frequently observed packing angles ( $\Omega \approx -40^\circ/+130^\circ$ , Figure 7). The plot shows deviations from  $180^\circ$  periodicity.

### Radius dependency of helix-helix packing angles

In the present model, it is assumed that the associating helices have the same radius. A check was made regarding this assumption. The distance of the side-chain atom that belongs to a contacting residue in one helix and that is the furthest from its own helix axis was projected onto the local line of closest approach between the two helices and then normalized by division with the closest approach distance. Residues were considered only if the angle between the perpendicular drawn from the geometric centre of the side-chain to its helix axis and the local line of closest approach was smaller than  $25^\circ$  to ensure that the residue points almost directly to the second helix along the local line of closest approach (“radius” at the interaction site). It was further required that identified contacting residues be surrounded by at least three side-chains of the second helix with closest side-chains atoms within 6 Å to the  $P_c$  site of the contacting residue. Good knobs-into-holes packing was thus guaranteed. The normalized distances for residues fulfilling these conditions have a mean of 0.44 and a standard deviation  $\sigma$  of 0.08 over 491 residues. The relatively low value of  $\sigma$  confirms that the assumption of similar radii is reasonable for the many contacting regions.

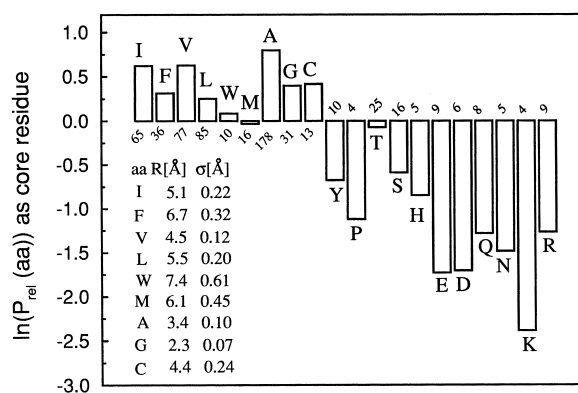
It is clear from Figure 3B and the behaviour of the function  $S_P(\Omega_N, R)$  (Figure 4) that for the optimal solutions the radii of the helices and thus the distance of closest approach between packed

helices should increase in the order  $R(a_m) < R(b_m) < R(c_m)$ . As shown in Figure 9 (mean closest approach distances versus the packing angle), these predictions are confirmed by the observed data in the negative angular range. Shortest distances of closest approach are found in agreement with the optimal knobs-into-holes packing classes, which allow interdigitation of the side-chains. As predicted, the distances increase from  $a_m$  to  $b_m$  and  $c_m$ . However, the distribution is, like the histogram of the packing angles (Figure 7), not periodic. In the positive range, very close distances are not observed. Only a shallow minimum (knobs-into-holes interdigitation) is observed in an interval where a suboptimal packing for larger helix radii is predicted. This is consistent with the observation that the main peak in the histogram of packing angles in the positive range (Figure 7) actually agrees with the suboptimal packing arrangement of helices with larger helix radii (Figure 4). It seems as if the required small helix radii for the optimal solutions  $a_m$  and  $b_m$  (short distances of closest approach), practically corresponding to packings of Gly or Ala in the packing core, are less compatible with positive packing angles. This recurrent and consistent deviation from periodicity can be explained by the actual spatial packing geometries of the two periodic solutions that are naturally excluded in the two-dimensional lattice model (see Discussion).

From Figure 4, it is evident that helices with radii between 3.0 Å and 4.5 Å fall simultaneously into the three minima of the function  $S_P(\Omega_N, R)$ . Thus, residue types that impart such helix radii should provide the largest packing flexibility and therefore be advantageous over others and correspondingly be often observed as radius-determining residues in the packing core; i.e. they should display a small skew angle  $\alpha$  (see Table 1 and Figure 1 for definition and illustration) to the other helix axis. At larger  $\alpha$  angles, the size of the residue becomes significant, such that extended residues can compensate for the cylindrical shape of helices and are observed more often because of the closer distance to the second helix. It can be shown that the mean skew angle increases from Ala to Val to Leu ( $\langle\alpha\rangle_{\text{Ala}} = 41.1^\circ$ ,  $\langle\alpha\rangle_{\text{Val}} = 43.2^\circ$  and  $\langle\alpha\rangle_{\text{Leu}} = 47.7^\circ$ ). The logarithmic relative frequencies of the 20 amino acids as packing core residues normalized to the mean relative frequency of the residue within helices generally is shown in Figure 10. The relative probability is highest for Ala followed by Val and Ile. The radius of a poly(Ala) helix is near 3.4 Å and 4.5 Å for poly(Val) and 5.2 Å for poly(Ile). The preference of Ala and Val is thus explained by their ability to allow flexible packing arrangements to achieve an optimal protein fold.

### Radius dependency of packing cell (holes) occupancy

Larger residues of one helix should preferably pack into the largest cell formed by the residues  $i$ ,

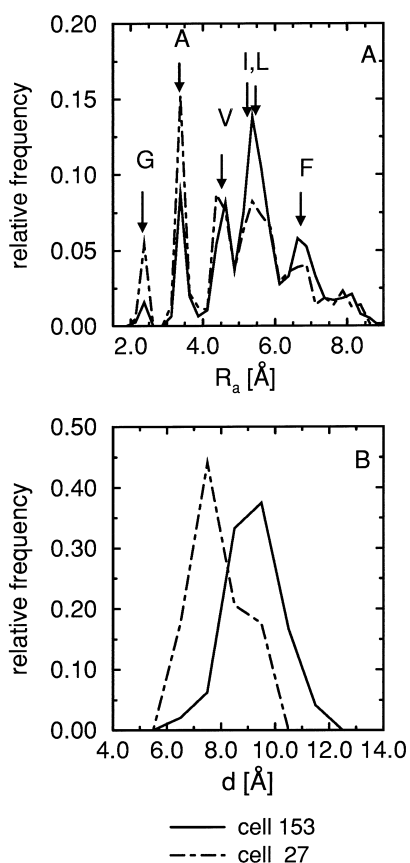


**Figure 10.** Logarithmic relative probability for the occurrence of individual amino acids in the packing core. Only side-chains with a corresponding skew angle  $\alpha$  smaller than  $25^\circ$  were taken under the condition that  $d_i < d + 1.0$  Å (Figure 1, Table 1). The obtained relative frequencies were normalized by the relative frequencies of the appropriate amino acid as found in all helices of the dataset. Listed in the plot are the specific helical radii corresponding to helices composed entirely of a given amino acid type and the respective observed standard deviations. Numbers near the bars correspond to the number of observations.

$i + 3$ ,  $i + 4$  and  $i + 7$  (cell 153) and smaller residues would prefer the smaller cell formed by residues  $i$ ,  $i + 1$ ,  $i + 3$  and  $i + 4$  (cell 27; Figures 2 and 5). Corresponding preferences for these cells are confirmed qualitatively by the experimentally observed data (Figure 11A). The small amino acids Gly and Ala clearly prefer to occupy cell 27. At an apparent helix radius of 4.5 Å (distance of the tip atom of the residue occupying the cell to its helix axis), the occupancy is balanced between cell 27 and cell 153. Larger residues (Ile, Leu and Phe) prefer cell 153 as predicted. The assignment of the packing cell is based on the position of contact ( $P_c$ ) and not the tip atom ( $P_{\text{tip}}$ ). For bigger side-chains these two positions may differ, likely explaining the weak difference for the two cells in occupancies at  $R > 7.0$  Å. It was observed that cell 51 also prefers larger amino acids; since it is somewhat to the side of the helix-helix interface, extended residues must reach like arms to fill it. The discrete nature of residue sizes used in packing is also evident from the clear peaks in the plot of Figure 11A. Furthermore, the more direct approach relating the distance of closest approach to the predominantly used packing cell type at a given helix-helix interface also confirms the theoretical conclusions (Figure 11B). More closely packed helices predominantly utilize cell 27, while cell 153 dominates for helices further apart at the association site.

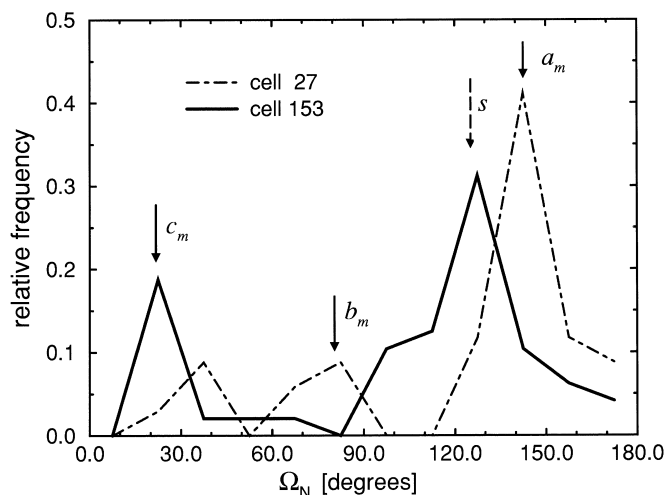
### Correlation of packing angle and preferred packing cell

Since the helix radius is related to the packing angle and to the packing cell predominantly occupied, a well-defined correlation should exist



**Figure 11.** Correlation between the helix radius and the occupied packing cell. A, Normalized histogram of occupancies of a specific packing cell (hole) are plotted as a function of the length (size) of the occupying side-chain defined by the distance of its tip atom to its helix axis (apparent helix radius  $R_a$ , Figure 1). Packing cell 27 is indicated by a broken line and cell 153 by a continuous line. For comparison, the mean distances for selected amino acids as found in all helices of the protein dataset are indicated by the arrows. The bin width was taken as 0.25 Å. B, Normalized histogram of observed distances of closest approach for helices with a predominantly packed cell 27 (broken line, 34 examples) and cell 153 (continuous line, 48 examples). The respective difference in the number of occupied packing cells of the two types for a given helix-helix contact region was larger than 2 to ensure cell-type dominance. Conditions that deem a cell occupied are discussed in the text. The bin width was taken as 1 Å.

between the packing angle and the preferred packing cell. At packing angles preferred by larger (smaller) helices, the packing cell 153 (27) should be mainly occupied. By assigning each pair of packed helices to one cell class determined by the prevailing cell type used, the resulting distribution is in good agreement with the predictions of the model (Figure 12). Because of the sparseness of data, the packings were normalized to the range  $0^\circ \leq \Omega_N \leq 180^\circ$ . Cell 153 is preferably occupied over two packing angle intervals. It is the dominating cell at  $\Omega_N \approx 25^\circ$  and occurs also at  $\Omega_N \approx 130^\circ$ , a suboptimal packing angle for larger radii. The peaks



**Figure 12.** Relative occupancies of packing cells (holes) as a function of packing angle  $\Omega_N$ . A helical pair was assigned to only one cell packing type according to the most frequently occupied cell along the contact. Counts are registered only if the number of occupied cells of type 153 (continuous line) is larger by at least +3 than the number of occupied cell types 27 for a single pair of packed helices (48 examples) and *vice versa* for cell 27 (broken lines, 34 examples).

are clearly distinct for cell 27 at  $\Omega_N \approx 150^\circ$  where the model predicts association of helices with smaller radii. Peaks for cell 27 are found also at  $\Omega_N \approx 80^\circ$  and  $\Omega_N \approx 40^\circ$ . The former angle corresponds to an optimal solution (class *b*) and the latter can be identified as suboptimal for helices with smaller radii. The distribution of  $S_p(\Omega_N, R)$  for helices with large radii has two minima, in contrast to that of helices with smaller radii, which exhibits three minima (Figure 4). This is confirmed by the data shown in Figure 12, which reveals that the broad peak at  $\Omega_N \approx 130^\circ$  actually comprises two different packing modes, optimal packings (cell 27 peak) and suboptimal packings (cell 153 peak).

### Non-homogeneous associations

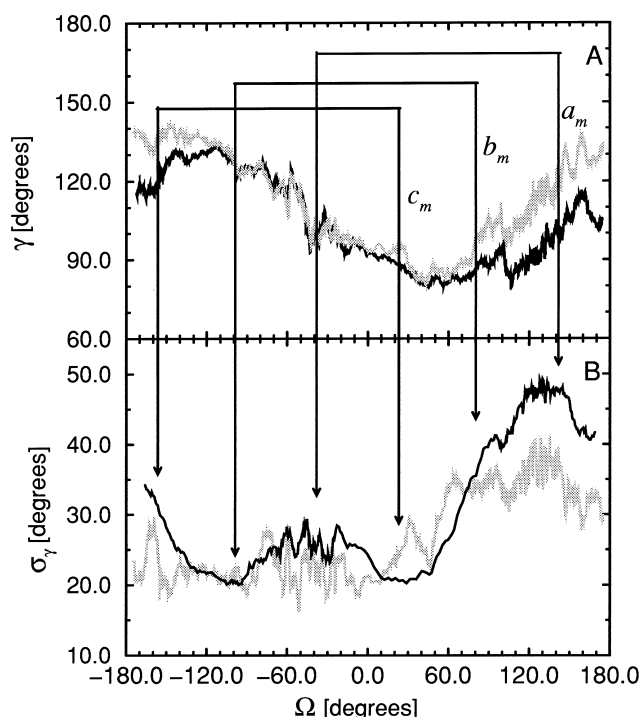
Besides side-chain interdigitation facilitated by van der Waals contacts of apolar atoms (knobs into holes), interhelical salt bridges, disulphide bonds, hydrogen bonds and tight head-to-head van der Waals contacts can constitute interhelical contacts, referred to here as interhelical bond interactions (knobs-onto-knobs). Indeed, cysteine and charged residues, and the polar asparagine were found to show the highest propensity of forming such bonds. Helix-helix associations with only one such interhelical bond were found more often at the expected packing angles (*vide supra*), provided that at least one helix of the pair had less than 12 residues (37 examples, data not shown). In longer helices with larger contact regions, the packing angles behaved according to knobs-into-holes where hydrophobic contacts dominate.

## Discussion

### Deviation from the 180° periodicity, limitations of the two-dimensional approach

A model for helix-helix packing based on superposition of two planar lattices yields 180° periodic solutions in the packing angle  $\Omega$ . However, the observed properties show deviations from periodicity. In particular, the predicted optimal solutions  $a_m$  and  $b_m$  are not convincingly represented by the experimental data in the positive angular range, neither the packing angles (Figure 7) nor at the expected smaller radii (Figure 9). What causes this discrepancy? Why is packing with short distances of closest approach (small helix radii) disfavoured in the positive  $\Omega$  range? Three main features of the real spatial structure of  $\alpha$ -helices are not described by a two-dimensional model: (1) the cylindrical shape; (2) the radii along the helix are discrete rather than continuous, as are the side-chain orientations (rotamers); and (3) the non-orthogonal extensions of side-chains; i.e. the  $C^\alpha$ - $C^\beta$  vectors leave the helix backbone under a defined angle (extension angle) and are not, as assumed by the model, straight extensions of the perpendicular drawn from the  $C^\alpha$ -positions to the helix axis. This latter property has been shown important in causing different oligomerization states of coiled coils (Harbury *et al.*, 1993).

Principally, for a real three-dimensional but regular helix, the lattice obtained by unrolling such a helix onto a plane coincides with that used in the model. Despite an apparently smaller helix radius for the same helix-building amino acids caused by the extension angle, the solutions for the packing angles would still be 180°-periodic but differ only in a translation of one lattice. However, in three dimensions, the extension angle entails different alignments of the  $C^\alpha$ - $C^\beta$  vectors of side-chains performing interhelical contacts (angle  $\gamma$ ) and thus different mutual orientations for the contacting residues; i.e. between the knob and corresponding hole residues. The angle  $\gamma$  between the  $C^\alpha$ - $C^\beta$  vectors of two contacting side-chains (at least one inter-helix atom-atom contact shorter than 4.5 Å) correlates at 71% with the angle between the corresponding  $C^\alpha$  geometric-centre-of-side-chain vectors. Obviously, the alignment angle  $\gamma$  depends on the packing angle  $\Omega$ , as demonstrated by Figure 13. The sinusoidal shape of the observed mean reflects the full-circle rotation in  $\Omega$ . Further, not only does  $\gamma_{\text{mean}}$  vary with the packing angle but also the observed standard deviations  $\sigma_\gamma$ . High  $\sigma_\gamma$  values reflect side-chain-side-chain contacts of residues with alternately nearly parallel (small  $\gamma$ ) and antiparallel  $C^\alpha$ - $C^\beta$  (large  $\gamma$ ) vector pairs, whereas smaller deviations point to more regular packing with the corresponding mean  $\gamma$  in the 90 to 120° range. In this respect, the optimal solutions  $a_m$  and  $b_m$  show more regularity in the negative packing angle range than in the positive range. For solution



**Figure 13.** Observed angle ( $\gamma$ ) between the  $C^\alpha$ - $C^\beta$  vectors of two interhelically contacting side-chains. Shown are the mean value (A) and the standard deviations (B) as a function of the packing angle of the corresponding helix-helix pair obtained by a 100-points (black lines) and 50-points (grey lines) running average of the  $\Omega$ -ordered data points. The black line corresponds to all observed  $\gamma$ -angles (4122 events) in A. In B, The raw data for the black curve were the standard deviations of  $\gamma$  per helix-helix pair. The 100-point clusters of the running average had a mean standard deviation of 10.3°. The grey lines were obtained for observations where at least one  $C^\alpha$ - $C^\beta$  vector of the contacting residue-residue pair made an angle to the global line of closest approach to the adjacent helix smaller than 45° (2747 events); i.e. residues centrally involved in the helix-helix packing. Arrows correspond to the three periodic optimal solutions of the helical lattice superposition model.

$c_m$ , the standard deviations are slightly smaller in the positive range but the orientations of the  $C^\alpha$ - $C^\beta$  vectors have less impact on the packing because of the larger required helix radii. Figure 14 reveals the consequences of the systematically different  $\gamma$ -angles on the helix-helix packing. In the case of alternating parallel and antiparallel  $C^\alpha$ - $C^\beta$  vectors (henceforth called alternating packing), where the optimal solutions  $a_m$  and  $b_m$  are in the positive packing angle range (Figure 13), the three-dimensional packing differs from the more regular ( $\gamma$ -angles) packings (henceforth called regular packing). Figure 14 illustrates this for solution  $a_m$ . Solutions  $a_m$  and  $b_m$  require small helix radii (Figure 3) and, consequently, short distances of closest approach. This is achieved by small residues in the packing core (preferentially Gly, Ala or Pro). In three dimensions, the planar lattice approach

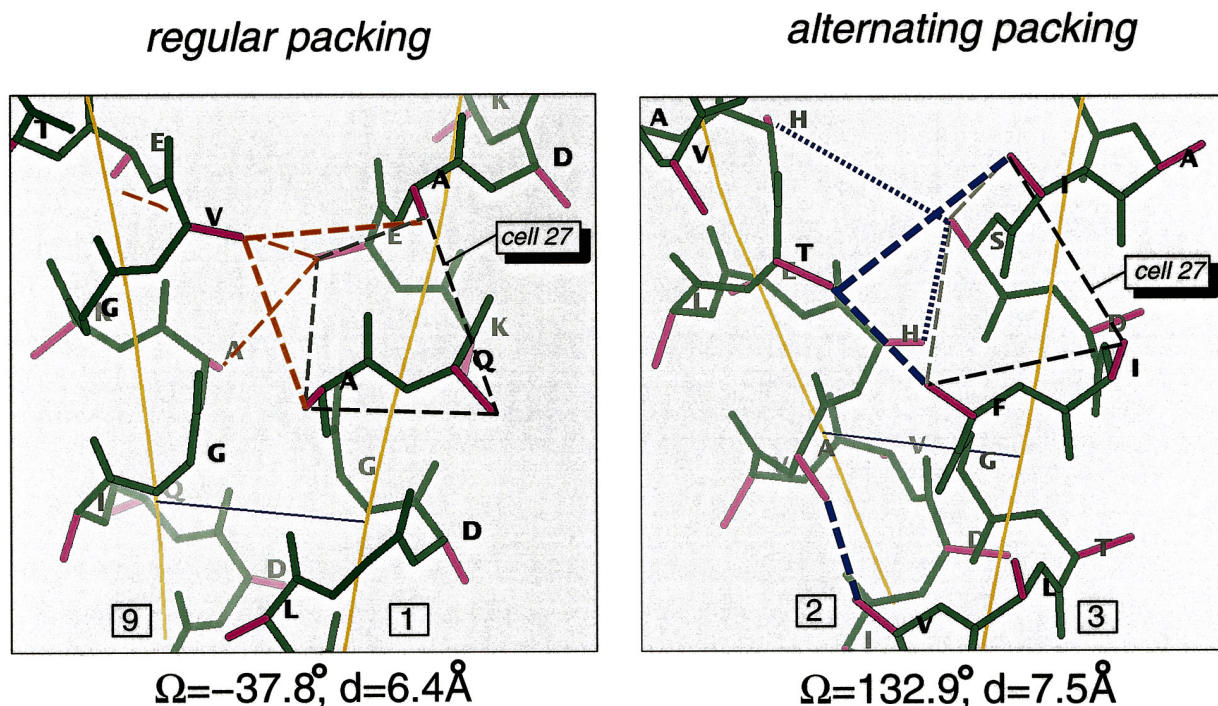
may be understood as packings of helices “unrolling” their side-chains onto the surface of the other. Thus, side-chains outside the packing core may be larger, thereby fulfilling the planar packing conditions and filling the crevice that would be opened up by the packing of ideal cylinders. This is supported by experimental observations such as the increasing mean skew angle from Ala to Val to Leu (*vide supra*). This mutual (“gearwheel”) unrolling is different for regular and alternating packings. In the alternating packing case, knob-residues repeatedly pack with hole-residues from the other helix with nearly antiparallel  $C^\alpha-C^\beta$  vectors (Figure 14). Obviously, given a corresponding packing angle  $\Omega$ , alterations of the side-chain sizes are less tolerable in this case where steric clash of the respective side-groups from the two helices can easily result because of the parallel or facing  $C^\alpha-C^\beta$  vectors. In the regular case, steric hindrance is less likely because the hole-residues point away from the interface and may even be extended. Consequently, regular packings may have short closest approach distances and more sequences (greater tolerance to different side-chain sizes) fulfil the requirement for small helix radii for solutions  $a_m$  and  $b_m$ . Alternating packing generally has larger distances of closest approach and thus, instead of utilizing the optimal

solution  $a_m$ , they go to the next accessible solution for helices with larger radii; i.e.  $\Omega \approx 120$  to  $130^\circ$  (Figure 4). The same principle considerations hold for the  $b_m$  periodic solutions. The next accessible packing mode for solution  $b_m$  is also the suboptimal with  $\Omega \approx 120$  to  $130^\circ$ , resulting in frequent observations for this packing angle range (Figure 7).

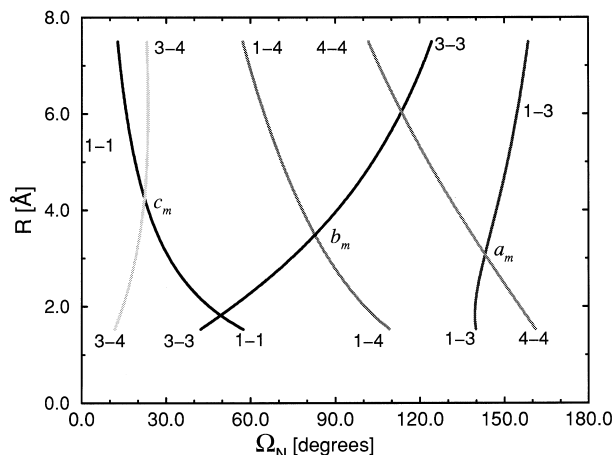
#### Ridges into grooves: a model lacking structural details

In the work presented here, helix-helix packing was studied theoretically from the perspective of the helical lattice superposition concept, which allowed all possible associations to be systematically considered from a purely mathematical perspective and is not found in previous work (Crick, 1953; Efimof, 1979; Chothia *et al.*, 1981). Thus, a more complete understanding of packing options, both optimal and suboptimal, has been achieved.

The lattice superposition model treats the packing problem on the basis of individual side-chains as the smallest packing unit, while higher-order structures are assumed by the ridges into groove (r/g) model where the dominating shape feature of helices are considered smooth, and continuous



**Figure 14.** Differences in the packing between the two  $180^\circ$ -periodic solutions of class  $a_m$ ; illustration of the regular and alternating packing mode. The pictures show real examples of packed helices with corresponding packing angles and distances of closest approach illustrating the differences in the  $C^\alpha-C^\beta$  vector alignments: regular packing (PDB entry codes and sequence numbers) 1dbp, helix 1, 43 to 53, helix 9, 237 to 253; alternating packing (right graph) 1thl, helix 2, 137 to 151, helix 3, 159 to 179. The  $C^\alpha-C^\beta$  bonds are drawn in magenta. Interhelical contacts between residues with nearly perpendicular  $C^\alpha-C^\beta$  vectors are denoted by broken red lines.  $C^\alpha-C^\beta$  vectors with antiparallel orientation are indicated by broken blue lines and the ones with nearly parallel orientation are shown with dotted blue lines. The yellow curved lines are the helix axes, the thin blue continuous lines are the lines of closest approach. The broken dark grey lines connecting the  $C^\beta$  positions denote the packing cell (cell 27). The sequences of the helices are given in the one-letter code.



**Figure 15.** Expected packing angles  $\Omega_N$  for the ‘ridges into groove’ model as a function of the helix radius. The numbers correspond to the combinations of the ridges and grooves; i.e. in the terminology of the model presented here, the oriented angles are given for the six possible combinations of the three base vectors (Figure 2) of one helical lattice with the corresponding three base vectors of the other; i.e. mirrored but not rotated lattice ( $\Delta = 1.45 \text{ \AA}$ ,  $\omega = 99.1^\circ$ ).

ridges and grooves are formed by residues at regular sequence separation. These ridges and grooves correspond to base vectors in the model presented here, where helix-helix packing involves their alignment in the respective lattices such that the condition  $\mathbf{v}_i = \lambda R_\Omega \mathbf{v}'_j$  is fulfilled. The term  $\lambda$  is a scalar value,  $R_\Omega$  is a rotation matrix with the corresponding packing angle  $\Omega$ ,  $\mathbf{v}'_j$  and  $\mathbf{v}_i$  are vectors joining lattice points with sequence spacing  $i$  and  $j$  (e.g.  $i = j = 4$  for class 4-4), and the prime denotes the applied mirror operation corresponding to face-to-face packing. The resulting packing angles are plotted in Figure 15 as a function of the helix radius. In the helix lattice superposition model, not only is the direction of a pair of base vectors considered but also their length and the packing properties of their neighbours. In most cases, this coincides with a ‘knobs into holes’ (k/h) packing scheme. The equivalent k/h graph is given in Figure 4. The three k/h optimal solutions ( $a_m$ ,  $b_m$  and  $c_m$ ) are found at the intersection points in the r/g model where three different base vectors are involved (Figure 15). The k/h treatment deletes some of the possible solutions of the r/g model due to steric clashes at other lattice points; for example, the 1-4 and 1-3 r/g classes at larger helical radii or the smaller radial segments of the 3-3 class. In the k/h approach, the optimal solutions delineate the preferred packing angles. For different classes of the r/g model, packing angles are not as distinguishable. Nonetheless, both approaches rely on the direction of base vectors and thus some packing solutions are commonly predicted.

It is obvious that the ridges and grooves are ‘bumpy’ and that protruding side-chains and local depressions are more appropriate helical surface descriptions. A smooth sliding of a ridge into a

groove is therefore unlikely. There exists a register allowing only discrete translations where the side-chains of one helix can click into the local depressions of the other helix (knobs into holes). Only through consideration of these key features of helices can successful prediction of the radius dependencies and occupancies of packing cells (holes) according to packing angle be achieved. Though two continuous ridges can certainly be aligned, others must inevitably cross. This conflict can be resolved only by assuming a discrete nature for ridges and grooves. Furthermore, only 27.8% of the helical residues make intra-helical side-chain-side-chain contacts (atom-atom distances smaller than  $4.0 \text{ \AA}$ ). Thus, smooth ridges hardly predominate.

Through a consistent mathematical treatment, three and only three solutions for the perfect superposition of  $\alpha$ -helical lattices have been demonstrated. Not only is suboptimal packing evident in the model but also the relationship between occupancy of packing cells and the helix radius. Further, it is shown that within the preferred packing angle range  $120^\circ < \Omega_N < 160^\circ$ , there are two topologically different packing arrangements (small helix radii/cell 27 occupancy and large helix radii/cell 153 occupancy). This result cannot be inferred from the r/g model where packing cells are not considered.

### Regularity of helices

The helix superposition model assumes regularity and that packing of two helices is strainless. The helix pairs must also display similar radii, possess relatively straight helical axes and constant twist angles and rises per residue. Significant violation lessens the applicability of the model. Despite the large variations in the observed twist angles  $\omega_{i,i+1}$  between consecutive centres of side-chains (standard deviation  $\sigma$  of  $15^\circ$ ), the side-chains are covalently bound to the  $C^\alpha$  backbone atoms which are very regular in their  $\omega_{i,i+1}$  ( $\sigma = 3.7^\circ$  for inner helical residues).

### Helix radius dependency of packing

By analysing the helix radius dependency of the packing angle, it was possible to reshape the suggestions of Richmond & Richards (1978), who inferred that the radius is inversely correlated with the packing angles defined as the smaller of the two complementary angles with  $180^\circ$ . The model used here shows that the dependency is not a monotonous function, as observed also by Reddy & Blundell (1993), albeit without the detailed explanations provided here.

The helix geometric parameters that allow optimized packing were examined. It is noteworthy that the structural characteristics of  $\alpha$ -helices designed by nature best and most consistently satisfy the requirements in the helix parameters

(Figure 3). This would allow considerable and advantageous flexibility in achieving the protein fold. Apart from internal structural strains, the clear disadvantage of other helical types in packing flexibility ( $\pi$ -helix and  $3_{10}$ -helix) in viable folded proteins is evident.

It has been shown that alanine as a helical constituent provides the largest flexibility in possible packing angles, since the radius of a poly(Ala)  $\alpha$ -helix is closest to that associated with the  $\omega$  triple point. Alanine has accordingly been observed to be very often involved in helix-helix contacts as a central, radius-determining amino acid (Figure 10). This alanine preference in helices is thus explained not only by its compatibility with helical structure as such but also by the attendant variety allowed for packing arrangements.

The observed relationship between packing angle and helix radius is likely to be of use in the engineering of protein structure. If, for instance, the designing task required helices packed with  $20^\circ$  (or  $-160^\circ$ ), leucine would be the ideal candidate for hydrophobic associations. If a packing angle of about  $-40^\circ$  is desired, glycine would be the better choice. This is supported by the work of Chou *et al.* (1984) who, in their energetic analysis of helix-helix packing, found the lowest interaction energies at  $-154^\circ$  ( $\Omega_N = 26^\circ$ ) for packing of poly(Leu) helices and at  $144^\circ$  ( $\Omega_N = 144^\circ$ ) for poly(Ala) helices.

The model in this work also explains the observed increased occurrence of leucine and the decreased frequency of glycine and proline in four-helix bundle proteins, where helices pack at about  $\Omega_N \approx 20^\circ$  (Paliakasis & Kokkinidis, 1992). Alanine was also often involved, which supports the model in that alanine was shown to possess greatest packing flexibility. The significance of packing cell type and helix radius, and the corresponding need for a good residue fit into a specific cell should further aid in associating helices. Minimally, the number of possible interaction sites can be reduced for any two specific helices. Attempts in this prediction direction are in progress.

In conclusion, the observed preference for packing angles near  $-40^\circ$  and  $+130^\circ$  may not be explained by a better packing of side-chains alone. The presented study revealed that there are three optimal periodic solutions for the packing angle. Furthermore, the preferred angle in the positive range is not the calculated optimal solution and therefore corresponds to a suboptimal solution (*vide supra*), hence, other determinants like entropic effects or surface burial differences might be important.

## References

- Chothia, C., Levitt, M. & Richardson, D. (1977). Structure of proteins: packing of  $\alpha$ -helices and pleated sheets. *Proc. Natl Acad. Sci. USA*, **74**, 4130–4134.
- Chothia, C., Levitt, M. & Richardson, D. (1981). Helix to helix packing in proteins. *J. Mol. Biol.* **145**, 215–250.
- Chou, K. C. & Zheng, C. (1992). Strong electrostatic loop-helix interactions in bundle motif protein structures. *Biophys. J.* **63**, 682–688.
- Chou, K. C., Némethy, G. & Scheraga, H. A. (1983). Energetic approach to the packing of  $\alpha$ -helices. 1. Equivalent helices. *J. Phys. Chem.* **87**, 2869–2881.
- Chou, K. C., Némethy, G. & Scheraga, H. A. (1984). Energetic approach to the packing of  $\alpha$ -helices. 2. General treatment of nonequivalent and nonregular helices. *J. Am. Chem. Soc.* **106**, 3161–3170.
- Cohen, F. E. & Kuntz, I. D. (1987). Prediction of the three-dimensional structure of human growth hormone. *Proteins: Struct. Funct. Genet.* **2**, 162–166.
- Cohen, F. E., Richmond, T. J. & Richards, F. M. (1979). Protein folding: evaluation of some simple rules for the assembly of helices into tertiary structures with myoglobin as an example. *J. Mol. Biol.* **132**, 275–288.
- Crick, F. H. C. (1953). The packing of  $\alpha$ -helices: simple coiled coils. *Acta Crystallog.* **6**, 689–697.
- Efimov, A. V. (1979). Packing of  $\alpha$ -helices in globular proteins. Layer-structure of globin hydrophobic cores. *J. Mol. Biol.* **134**, 23–40.
- Harbury, P. B., Zhang, T., Kim, P. S. & Alber, T. (1993). A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science*, **262**, 1401–1407.
- Harris, N. L., Presell, S. R. & Cohen F. E. (1994). Four helix bundle diversity in globular proteins. *J. Mol. Biol.* **236**, 1356–1368.
- Heringa, J., Sommerfeldt, H., Higgins, D. & Argos, P. (1992). OBSTRUCT: a program to obtain largest cliques from a protein sequence set according to structural resolution and sequence similarity. *CABIOS*, **8**, 599–600.
- Hutchinson, E. G., Morris, A. L. & Thornton, J. M. (1994). Structural patterns in globular proteins. In *Structure Correlation* (Burgi, H. B. & Dunitz, J. D., eds) Verlag Chemie, Weinheim, vol. 2, pp. 643–650.
- Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **2**, 2577–2637.
- Murzin, A. G. & Finkelstein, A. V. (1988). General architecture of the  $\alpha$ -helical globule. *J. Mol. Biol.* **204**, 749–769.
- Paliakasis, C. D. & Kokkinidis, M. (1992). Relationships between sequence and structure for the four- $\alpha$ -helix bundle tertiary motif in proteins. *Protein Eng.* **5**, 739–748.
- Ptitsyn, O. B. & Rashin A. A. (1975). A model of myoglobin self-organisation. *Biophys. Chem.* **3**, 1–20.
- Reddy, B. V. B. & Blundell, T. L. (1993). Packing of secondary structural elements in proteins. Analysis and prediction of inter-helix distance. *J. Mol. Biol.* **233**, 464–479.
- Richmond, T. J. & Richards, F. M. (1978). Packing of  $\alpha$ -helices: geometric constraints and contact area. *J. Mol. Biol.* **119**, 537–555.
- Robinson, C. R. & Sligar, S. G. (1993). Electrostatic stabilization in four-helix bundle proteins. *Protein Sci.* **2**, 826–837.
- Schulz, G. E. & Schirmer, R. H. (1979). *Principles of Protein Structure*. Springer-Verlag, Berlin.
- Solov'yov, V. V. & Kolchanov, N. A. (1984). A simple method for the calculation of low energy packings of  $\alpha$ -helices—a threshold approximation. I. The use of the method to estimate the effects of amino acid substitutions, deletions and insertions in globins. *J. Theoret. Biol.* **110**, 67–91.

Tufféry, P. & Lavery, R. (1993). Packing and recognition of protein structural elements: a new approach applied to the 4-helix bundle of myohemerythrin. *Proteins: Struct. Funct. Genet.* **15**, 413–425.

## Appendix

Under the condition of perfect overlap of the two helical lattices, equations (1) and (2) must be satisfied. The base vectors  $\mathbf{v}_1''$  and  $\mathbf{v}_2''$  describing the second lattice result from  $\mathbf{v}_{1,2}'' = R_\Omega \mathbf{v}_{1,2}'$  where the vectors  $\mathbf{v}_{1,2}'$  are mirrors of the base vectors of the first lattice  $\mathbf{v}_{1,2}$  and  $R_\Omega$  is a rotation matrix (see Mathematical Description). By substituting the selected vectors of equations (3) and (4) into equations (1) and (2) the following system of equations results:

$$-AR \cos(\Omega) - \Delta \sin(\Omega) = n_1 AR + n_2 BR \quad (\text{A1})$$

$$-AR \sin(\Omega) + \Delta \cos(\Omega) = n_1 \Delta + n_2 k \Delta \quad (\text{A2})$$

$$-BR \cos(\Omega) - k \Delta \sin(\Omega) = n_3 AR + n_4 BR \quad (\text{A3})$$

$$-BR \sin(\Omega) + k \Delta \cos(\Omega) = n_3 \Delta + n_4 k \Delta \quad (\text{A4})$$

Given specific helix geometric parameters ( $R$ ,  $\omega$ ,  $\Delta$ ), this system of equations would contain five unknowns ( $\Omega$ ,  $n_1$ ,  $n_2$ ,  $n_3$  and  $n_4$ ). However, for the integer variables  $n_{1..4}$ , several boundary conditions apply:

$$n_{1,2,3,4} \in (-1, 0, 1) \quad (\text{A5})$$

$$n_1^2 + n_2^2 \neq 0 \quad \text{and} \quad n_3^2 + n_4^2 \neq 0 \quad (\text{A6})$$

$$|n_1 + n_2| < 2 \quad \text{and} \quad |n_3 + n_4| < 2 \quad (\text{A7})$$

$$n_2 n_3 \neq n_1 n_4 \quad (\text{A8})$$

These conditions reflect restrictions in the length and orientation of the base vectors. Obviously,  $n_1$  and  $n_2$  may not be simultaneously zero; the same holds for  $n_3$  and  $n_4$  (equation (A6)). The base vectors may not exceed in magnitude the distance of the closest hexagonal lattice points around the point  $P_i$  (Figure 2; equations (A5) and (A7)) and they may not be linearly dependent (equation (A8)). These boundary conditions reduce the number of possible combinations of values for  $n_1$  to  $n_4$  from 81 to 24. Further restrictions can be elicited by reformulating equations (A1) to (A4). (1) Multiplying equation (A1) and equation (A2) by  $B$  and (A3) and (A4) by  $A$  and subsequently subtracting equation (A1) from (A3) and (A2) from (A4) yields:

$$\begin{aligned} \Delta(B - kA)\sin(\Omega) &= (AB(n_4 - n_1) \\ &+ n_3 A^2 - n_2 B^2)R \end{aligned} \quad (\text{A9})$$

$$\begin{aligned} \Delta(kA - B)\cos(\Omega) \\ = ((n_3 + kn_4)A - (n_1 + n_2 k \Delta)B)\Delta \end{aligned} \quad (\text{A10})$$

Dividing equation (A9) by equation (A10), the 180° periodicity of the solutions become obvious; namely:

$$\tan(\Omega) = \frac{R}{\Delta} f(\omega) \quad (\text{A11})$$

where  $f$  indicates a function.

(2) Separating  $R$  and  $\Delta$  in equations (A1) to (A4) and equating one side of equation (A1) with (A2) and one side of equation (A3) with (A4) yields:

$$n_2(kA - B)\cos(\Omega) = A - (n_1 + kn_2)(n_1 A + n_2 B) \quad (\text{A12})$$

$$n_3(kA - B)\cos(\Omega) = (n_3 + kn_4)(n_3 A + n_4 B) - kB \quad (\text{A13})$$

(3) By multiplying equation (A1) and (A3) by  $\Delta$  and equation (A2) and (A4) by  $AR$  and subsequently subtracting equation (A1) from (A2) and (A3) from (A4) and multiplying equation (A1) and (A3) by  $k\Delta$  and (A2) and (A4) by  $BR$  and subsequently subtracting (A1) from (A2) and (A3) from (A4), it can be shown that:

$$\begin{aligned} (k\Delta^2 - ABR^2)\sin(\Omega) + \Delta R(kA + B)\cos(\Omega) \\ = n_1 \Delta R(B - kA) \end{aligned} \quad (\text{A14})$$

$$\begin{aligned} (\Delta^2 - A^2 R^2)\sin(\Omega) + 2AR\Delta\cos(\Omega) \\ = n_2 \Delta R(kA - B) \end{aligned} \quad (\text{A15})$$

$$\begin{aligned} (k^2 \Delta^2 - B^2 R^2)\sin(\Omega) + 2kBR\cos(\Omega) \\ = n_3 \Delta R(B - kA) \end{aligned} \quad (\text{A16})$$

$$\begin{aligned} (k\Delta^2 - ABR^2)\sin(\Omega) + \Delta R(kA + B)\cos(\Omega) \\ = -n_4 \Delta R(B - kA) \end{aligned} \quad (\text{A17})$$

These latter transformations restrict the possible combinations of  $n_{1,2,3,4}$ . Comparing equations (A14) and (A17), it directly follows that  $n_1 = -n_4$ , given that  $B - kA \neq 0$ . The remaining cases of possible combinations must be investigated separately. If  $n_2 = n_3 = 0$  and  $n_1 = -n_4 = \pm 1$ , then it follows from equation (A10) that:

$$(kA - B)\cos(\Omega) = \pm(kA - B) \quad (\text{A18})$$

Since  $kA - B \neq 0$ , then  $\cos(\Omega) = \pm 1$  and thus  $\sin(\Omega) = 0$ . Under these conditions equations (A15) and (A16) yield  $\pm 2AR\Delta = 0$  and  $2kBR = 0$ . Since  $A = \omega$  and for any real helix  $\omega$  cannot be zero, then the combinations of  $n$ -values above are dismissable. If  $n_1 = n_4 = 0$  and  $n_{2,3} = \pm 1$ , then it follows from equations (A12) and (A13) that:

$$n_2(kA - B)\cos(\Omega) = A - kB \quad (\text{A19})$$

$$n_3(kA - B)\cos(\Omega) = A - kB \quad (\text{A20})$$

Since  $kA - B \neq 0$ ,  $n_2 = n_3$  providing  $A - kB \neq 0$ . If  $A - kB = 0$ , the twist angle  $\omega$  must be  $2k\pi/(k^2 - 1)$  and from equation (A14),  $\Delta/R = 2\pi/(k^2 - 1)$ . A detailed examination of equation (A9) taken with these values for  $\omega$  and  $\Delta/R$  shows that  $n_2 = n_3$ .

Consequently, six possible combinations of  $n_{1,2,3,4}$  remain. The solutions for the packing angle  $\Omega$  can now be obtained directly by using these possible sets in equations (A9) and (A10). The allowed sets are given in Table 2 of the main text and correspond to packing classes designated here as  $a$ ,  $b$  and  $c$  (each

with  $180^\circ$  periodicity). Note that the equations are solved under the conditions of lattice superposition for two associating helices.  $\Omega$  represents the rotation angle required for one lattice to achieve the overlap. Actual packing is, of course, a result of lattice translation as well.

***Edited by B. Honig***

*(Received 4 July 1995; accepted 9 October 1995)*