Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The architecture of the GroEL–GroES–$(ADP)_7$ chaperonin complex. II. Heptagrammal characterization of the folding

A. Janner

Institute for Theoretical Physics, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands

Correspondence e-mail: alo@sci.kun.nl

Received 18 July 2002
Accepted 12 February 2003

The heptagrammal forms derived in part I [Janner (2003*a*). *Acta Cryst.* D**59**, 783–794] enclose chain segments of symmetry-related monomers in the GroEL–GroES–$(ADP)_7$ chaperonin complex. A chain reaching the boundary of a given form either ends, proceeds to a neighbouring form or has to fold. C^α atoms corresponding to these folding points are identified in each of the nine forms of the chaperonin and are approximated by ideal positions having integral coordinates (the indices) with respect to a symmetry-adapted basis. Mutual structural relations between the indexed positions are derived in terms of integral scale-rotations (similar to those that leave the form invariant). The magnesium ions at the binding sites of the nucleotides ADP and ATP are shown to be symmetry-related to these folding points. The change in folding (polymorphism) observed in the *cis* ring of GroEL arising from binding to GroES is discussed. In particular, the form segmentation is conserved in the polymorphic transition. The geometric and algebraic restrictions imposed on the indexed positions and on their structural relations by the integrality condition are presented in an appendix.

1. Prelude

Before discussing the heptagrammal characterization of the folding in the GroEL–GroES–$(ADP)_7$ chaperonin complex, it is opportune to consider the folding problem in the more general context of the geometry of axial symmetric proteins. The basic folding of a C^α chain occurs at the level of a monomer, which is eventually assembled into a multimer. In this process, the minor adjustment of the folded monomer can usually be neglected. Accordingly, the point-group symmetry appears to be irrelevant in protein folding.

This is certainly not the case for GroEL, where scaling symmetries of the multimer, which are normally not included in the point group, relate the external envelope to the central channel and imply the existence of geometric correlations between folding points of the monomer.

Moreover, the changes induced in GroEL on complexation with GroES are not negligible and substantially modify the folding of the monomers of the *cis* ring. In this case, the folding of the monomer and the geometric properties of the multimer are mutually dependent.

Investigation of 12 other axial symmetric proteins has shown that the interplay between multimeric symmetries and monomeric folding observed in the GroEL–GroES chaperonin has a generic character. This is quite natural, because the folding of the monomer and the geometry of multimer as an active unit are both essential for the biological processes and

have therefore been involved together in the evolutionary selection.

It can therefore be attempted to identify the relevant folding points of the monomer starting from the symmetry properties of the multimer or, conversely, to predict the symmetries of the multimer on the basis of the geometry of the folded monomer. In the following an attempt is made to present the typical steps that are possibly involved in these two complementary approaches.

1.1. Multimeric characterization of the monomeric folding

1.1.1. Molecular forms. The starting point is represented by the scale-rotational symmetry of the envelope and channel in terms of indexed molecular forms of the multimeric quaternary structure of an axial symmetric protein. These forms are delimited by planes perpendicular or parallel to the rotation axis and have vertices with integral coordinates (the indices) with respect to a symmetry-adapted basis. The following are typical.

(i) Scaling factors which depend on the rotational symmetry.

(ii) Characteristic fixed ratios between the inter-planar distance of delimiting planes parallel and perpendicular to the axis of rotation. Loosely speaking, these ratios imply a relation between the height and width of the molecular form.

For example, in GroEL the scaling factors are heptagrammal and are therefore connected with the sevenfold rotational symmetry; in the *trans* ring the height of the monomers is equal to the radius of the envelope of the multimer. For the symmetry-adapted basis (a, c) defined in the next section, one has c: $a = 1$.

1.1.2. C^{α} positions at form boundaries. In a second step, C^{α} positions are identified which have the following properties.

(i) A location at/near the boundaries of a molecular form. Typically, basal positions are found when the delimiting plane is perpendicular to the axis of rotation and lateral positions for delimiting planes parallel to the axis.

(ii) Integral coordinates with respect to the symmetry-adapted basis or, in other words, an ideal position with integral indices.

(iii) At least a pairwise relation by invertible integral transformations (typically representing scale-rotations).

1.1.3. Correlated folding points of the monomer. Making use of the rotational symmetry of the multimer in terms of monomers, the structural relations between C^{α} positions in the multimer are expressed as structural relations between C^{α} positions of a single monomer. The symmetry of the multimer is encoded in the geometry of these correlated folding points.

This is the approach adopted in the present work: molecular forms in part I and heptagrammal correlated folding points in part II.

1.2. Axial symmetric multimer from a folded monomer

It is possible in principle to derive the expected axial symmetry of the multimer assembled from a given folded monomer. The self-assembly process certainly involves more properties than the geometry of the monomer in a C^{α} backbone description, which is the only property considered here. However, the geometric properties substantially reduce the number of *a priori* possibilities. In the present approach, only a limited number of C^{α} positions are expected to be relevant: those denoted above as correlated folding points, which are those that become symmetry-related in the assembled conformation. They predispose the monomer to the symmetry adopted by the multimer. A number of these positions are characterized by some extreme property. Much more cannot be said at present, but it is clear that the path presented in the previous approach cannot simply be reversed. The insights gained so far suggest an algorithmic approach. At present, it is not more than a strategy, the validity of which has not yet been tested.

1.2.1. Axial direction. The first step of an algorithmic loop is to orient the monomer with respect to the direction of the axis of rotation expected for the multimer, taken to be the z axis of an orthogonal coordinate system. The fitting of the monomer into prismatic forms delimited by lateral planes parallel to the axial direction and by basal planes perpendicular to it can help to find the z direction. Other dominant secondary structures, such as α-helices or β-strands, may also suggest possible solutions. For a chosen orientation, the positions with extreme values of the z coordinate fix the height of the monomer or, to put it better, of its prismatic envelope, which is the same as the height of the axial-symmetric multimer.

1.2.2. Axis of rotation. A pair of C^{α} positions in the monomer that are related in the multimer by a planar scaling transformation have the same z coordinate and are aligned in a radial direction in the xy plane. Therefore, the intersection points of straight lines through C^{α} positions at (approximately) the same height are expected to show, in a projection along the z direction, a concentration at the axis of rotation. This should allow fixation of the location of the z axis, the cylindrical envelope of the multimer and its central channel.

1.2.3. Polygrammal scaling and rotations. Pairs of C^{α} positions that are

(i) at the same height z,

(ii) along the same radial direction in the xy plane and

(iii) at maximal and minimal radial distance from the rotational axis

are expected to be mutually related by a polygrammal scaling. If so, the scaling factor of their radial distances depends in a recognizable way on the order of the possible rotational symmetry. Both rotations and scalings are restricted by the crystallographic condition imposed by the existence of a common symmetry-adapted basis. This basis is indeed such that the elements of a scale-rotation point group are represented by integral invertible matrices. The axial symmetric multimer is then generated from the monomer by the subgroup of rotations.

1.2.4. Molecular forms of envelope and channel. A molecular form of the multimer obtained is constructed which has delimiting planes perpendicular and parallel to the rotation axis. If the vertices have integral indices with respect to a

symmetry-adapted basis, the loop of the algorithm is closed. Otherwise it is necessary to start again.

The validity of this algorithm is based on scaling properties which, together with rotations, eventually lead to a set of correlated folding points encoding the geometry of the multimer in the monomer. In the formulation given above, the algorithm only indicates a logical possibility. Its practical formulation will certainly require additional probabilistic elements. At present, it can be hoped to find cases that are simple enough to allow a decoding of the cryptogram represented by a folded monomer.

2. Introduction

In the first part of this work (Janner, 2003*a*), here denoted as part I, the architecture of the GroEL–GroES chaperonin complex has been analyzed in terms of molecular forms enclosing given segments of a set of polypeptide chains related by the sevenfold rotation R, a nearly exact symmetry of the chaperonin-ring structure with a central channel. This leads to forms consisting of two heptagonal prisms, one external and one internal, the two being related by a scaling transformation S. The basic property which justifies the present work is that these forms are not only heptagonal but also heptagrammal.

The heptagrammal property of a form requires sevenfold rotational symmetry and vertices with integral coordinates (the *indices* of the vertices of a heptagram) when expressed with respect to a symmetry-adapted basis set of vectors. In this basis, R and S are represented by invertible integral matrices. This property restricts the order of the rotation R and the scaling factor of S. Technically speaking, the basis vectors generate a three-dimensional free $\mathbb{Z}$-module M of rank 7, with R and S automorphisms of M generating the point group of the form.

A form enclosing a chain segment implies the existence of folding points at C^α atoms near enough to the form boundary. The aim of the present second part is to identify and to analyze these folding points or, in other words, to characterize that part of the molecular folding arising from the confinement by the given form. Actually, this perspective reverses cause and effect, because the form follows from particular folding points and not conversely. However, the present approach is purely phenomenological and inductive. It is intended to show that, with few exceptions, the folding points at the form boundaries have integral indices and that there are structural relations between these folding points expressible by integral invertible matrices, generalizing the few special cases presented in a further paper (Janner, 2003*b*). The existence of both properties explains the symmetry of the forms observed without implying an understanding of the origin of the symmetry.

As outlined in a previous paper (Janner, 2001), in a molecular crystallographic approach such as the present one, these structural relations are obtained from elements of a so-called *structural group* which has the symmetry group of the system as a subgroup. This symmetry group includes what in biochemistry are called non-crystallographic symmetries, which leave the molecule invariant, at least approximately, but not the embedding crystal. The elements of the structural group do not generally leave the molecule invariant nor the (idealized) backbone description, nor even the whole set of folding points involved in the form confinement.

Nevertheless, the knowledge of a structural group is essential for characterization of the architectural elements of the molecule. The present case can help to grasp how this is possible. Once folding points arising from a basal confinement (at C^α positions near the boundary faces perpendicular to the sevenfold axis) have been identified, it can be observed that they appear in a heptagrammal arrangement in the axial projection and are related by scale-rotations leaving the $\mathbb{Z}$-module M invariant. Conversely, by considering folding points at the lateral faces of a given form, it is found that their projection in a direction perpendicular to the axis fits with the underlying axial $\mathbb{Z}$-submodule.

In other words, integral indices can be assigned to a number of C^α positions which occur at boundary points and their mutual positions can be related by heptagrammal point-group transformations. Not all C^α positions near the boundaries possess this property, but there are enough folding points to justify the heptagrammal properties of the enclosing forms derived in part I. After some generalities given in the next section, in §4 this analysis is performed for the forms of the *trans* ring of GroEL and in §5 for the *cis* ring. The corresponding treatment of the co-chaperonin GroES follows in §6. In §7 the binding of the nucleotides ADP and ATP is considered in the complexed and in the uncomplexed states of the *cis* ring of GroEL. In both cases similar relations emerge between the positions of the Mg ions and form-folding points. The results are summarized in §8, where an attempt is made to gain a first insight in the polymorphism of the GroEL chains. In an appendix the geometry of the heptagrammal arrangements observed in the various axial projections of the folding points is first considered, followed by a characterization of heptagrammal point-group transformations. This appendix is intended to complement the appendix of part I.

3. Generalities

The analysis of the various forms requires some preliminary general considerations. While in part I the data of three PDB files were considered together, 1grl for the unliganded GroEL (Braig *et al.*, 1994), 1der for GroEL complexed with ATPγS (Boisvert *et al.*, 1996) and 1aon for GroEL–GroES–$(ADP)_7$ complex (Xu *et al.*, 1997), here all the positions indicated are those of the residues found in 1aon. §§6 and 7 also require, of course, the GroEL *cis* ring from 1der.

The consistency with part I is ensured by adopting the same coordinate system and the $\mathbb{Z}$-module basis $\{a_1, \ldots, a_7\}$ as in equation (1) of part I,

$$a_k = a(\cos k\varphi, \sin k\varphi, 0), \quad \varphi = \frac{2\pi}{7}, \ k = 1, 2, \ldots, 6,$$
$$a_7 = c(0, 0, 1), \tag{1}$$

where

$$a_0 = -(a_1 + a_2, + \ldots + a_6) \quad (2)$$

is oriented along the x axis. This implies that before plotting, the data of 1aon are first translated to a new origin O by $\Delta x = -78.84$ Å, $\Delta y = 51.59$ Å and Δz with a value ranging from -4.0 to 0.0 Å and rotated around the z axis by an angle γ varying from 11 to 8°. The varying values of Δz and γ take into account possible local deformations with respect to ideal positions, as explained in part I. All C^α positions at form-boundary points are approximated by ideal positions with integral coordinates (rational indices) with respect to a single $\mathbb{Z}$-module M for the whole GroEL–GroES–$(ADP)_7$ complex. The fitting of the data to this $\mathbb{Z}$-module leads to the following values for the parameters a and c of (1),

$$a = 70.3 \text{ Å}, \quad c = a/12. \quad (3)$$

Rational indices have only a structural meaning for small integers (say, not much larger than ten in absolute value). For GroEL and GroES this is only the case if locally different submodules of M are adopted. In particular, as in part I, the submodules M_t, M_c and M_s are used for the GroEL *trans* ring, for the GroEL *cis* ring and for GroES, respectively. The corresponding $\mathbb{Z}$-module parameters are then

$$\begin{aligned} M_t: &\quad a_t = a, \quad c_t = 3c, \\ M_c: &\quad a_c = a, \quad c_c = 3\lambda_N c, \\ M_s: &\quad a_s = \mu_E a, \quad c_s = 2\lambda_N c. \end{aligned} \quad (4)$$

As $\lambda_N = 4\cos\varphi + 6\cos 2\varphi = 1.1588\ldots$ and $\mu_E = -1 + 2\cos\varphi - 2\cos 2\varphi = 0.6920\ldots$ are scaling factors of transformations that leave the $\mathbb{Z}$-module M invariant (see appendix), M_t, M_c and M_s are indeed submodules. In the figures, circles containing dots indicate the C^α positions at form-boundary points and neighbouring empty circles indicate the corresponding indexed ideal positions.

While in part I scaling transformations are characterized by their scaling factor labelled by a capital letter, here these scaling factors are indicated in their two-decimal approximation. For example, S_{μ_E} is indicated as $S_{0.69}$ and S_{λ_N} as $S_{1.15}$. The corresponding exact values can be found in the appendix. This characterization is only unambiguous for scaling matrices with small integral entries, as required by a meaningful indexing of related atomic positions. Following a suggestion by B. Souvignier, the entries of the scaling matrices considered (and of their inverse) are of absolute value not larger than 5.

The general procedure is similar to that adopted in part I for deriving the molecular forms of the GroEL–GroES chaperonin. It involves two steps leading to indexing and to structural relations. For determining the confined points these two steps are applied to the forms one after the other: three for the *trans* ring and three for the *cis* ring of GroEL and three for GroES.

3.1. Indexing

The first step consists of identifying C^α positions sufficiently well approximated by points with small integral indices at basal and/or lateral boundaries, which then define corresponding ideal C^α positions. The condition on the indices to be small is essential, but depends on the $\mathbb{Z}$-module basis adopted. Normally, this basis is the same as for the form considered. In some cases it is necessary to convert to an equivalent (local) basis. This set of ideal positions with integral coordinates has a heptagonal symmetry and is, by definition, a heptagram. Examples of heptagrams observed in GroEL–GroES are presented in the appendix.

3.2. Structural relations

In a second step $\mathbb{Z}$-module automorphisms are looked for (denoted as heptagrammal point-group transformations) which relate points of the given heptagram. Only those transformations are admitted which have small integral entries when expressed in the same basis as the points and an inverse with the same property. In the appendix, heptagrammal scaling transformations are given as parameterized integral matrices, together with a general expression for the corresponding scaling factors.

The C^α positions connected by scale-rotational point-group elements are interpreted as correlated folding points. For the remaining unrelated C^α positions, attempts are made to find alternative geometrical characterizations which still could be meaningful. In any case a question mark is then justified, either on the indexed idealization adopted or on having included these positions in the heptagram.

The figure of merit of an assignment depends on the following.

(i) The discrepancy between observed and ideal positions.
(ii) The simplicity of the heptagrammal structure.
(iii) The value of the rational indices.
(iv) The integral entries of the transformations.
(v) The number of non-trivial structural relations.

This figure of merit has not been cast into a formula. It is only intended as a basis for a qualitative evaluation of the results derived.

4. Heptagrammal folding in the *trans* ring of GroEL

4.1. The N-terminal form

Let us begin with the N-terminal form (see Fig. 3*a* of part I). The segments enclosed range from the residues Ala2 to Leu187 of the chains H to N of GroEL given in the PDB file 1aon. As shown in Fig. 1(*a*), four C^α positions have been selected: two basal, A109 and T181, and two lateral, K65 and D140. The corresponding heptagrammal arrangement has been marked by lines connecting the ideal positions with integral indices in the basis (a_t, c_t) indicated in (1) and (4), according to step 1. Given here are the indices of the positions labelled in Fig. 1(*a*). The other set of indices follows by heptagonal rotations.

$$\begin{aligned} \text{lateral}: \quad \text{K65} &= (1\,2\,0\,0\,2\,1, 1)_t, \\ \text{D140} &= (0\,0\,0\,\bar{1}\,0\,0, 2)_t, \\ \text{basal}: \quad \text{A109} &= (\bar{1}\,0\,\bar{1}\,\bar{1}\,0\,\bar{1}, 0)_t, \\ \text{T181} &= (1\,1\,2\,\bar{1}\,1\,2, 3)_t. \end{aligned} \quad (5)$$

In order to derive possible structural relations between these positions (step 2), it is immediately observed that D140 is rotationally equivalent to $B(2)$, where $B(n)$ denotes the point with indices $(1\,1\,1\,1\,1\,1, n)_t$. The parabolic automorphism P_t of M_t [already considered in equation (32) of part I and indicated again in the appendix] transforms $B(n)$ to $B(n+1)$ and the scaling $S_{-0.35}$ (specified in the same appendix) transforms $B(1)$ to K65. This implies that K65 and D140 are point-group-equivalent positions, with a structural relation given by

$$S_{-0.35}P_t^{-1}R^3(\mathrm{D140}) = \mathrm{K65}. \qquad (6)$$

In a similar way $B(2)$ is transformed to $B(0)$ by P_t^{-2} and then to A109 by the scaling $S_{-0.55}$, leading to the relation

$$S_{-0.55}P_t^{-2}R^3(\mathrm{D140}) = \mathrm{A109}. \qquad (7)$$

Therefore, A109 is also point-group equivalent to K65.

The remaining position T181 is not simply related to the others. Geometrically, it appears along an edge of the star heptagon {7/2} (shown in Fig. 8 of part I) scaled by a heptagrammal factor 0.24698... from one of the vertices. The indices indicated above have been obtained using this property. This scaling, however, is not centred at the origin, so that it does not imply the existence of a heptagrammal point-group relation with the other positions, even if it can be expressed in terms of elements belonging to the point group:

$$[S_{0.24}(R^2 - \mathbb{1}) + \mathbb{1}]R^{-3}P_tR^3(\mathrm{D140}) = \mathrm{T181}, \qquad (8)$$

where $\mathbb{1}$ denotes the unit matrix.

It is sometimes convenient to indicate the planar scaling only, considering the corresponding relations as given modulo the heptagonal rotations and the parabolic transformations of M_t. Doing so, the first two structural relations derived above are simply given by:

$$\mathrm{D140} \simeq S_{-0.35}^{-1}(\mathrm{K65}) \simeq S_{-0.55}^{-1}(\mathrm{A109}). \qquad (9)$$

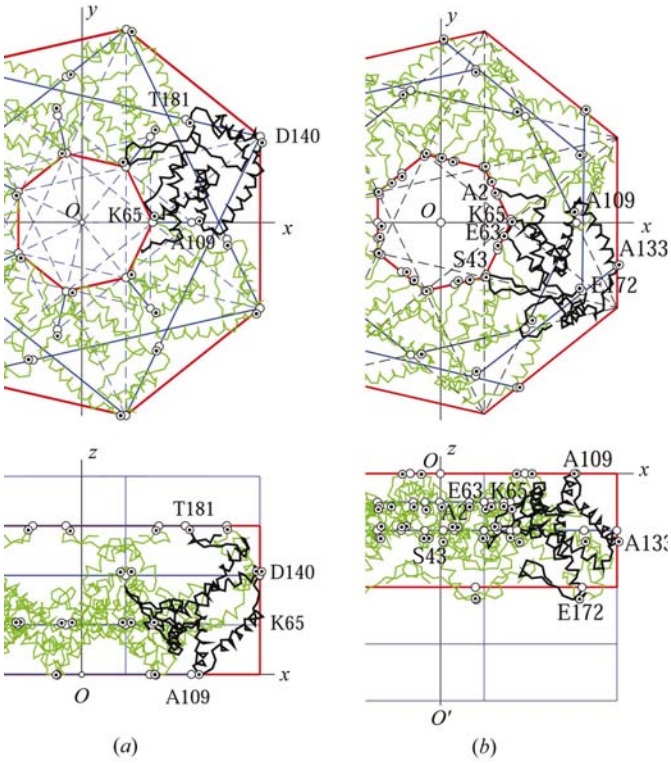

Figure 1
(*a*) Selected folding points arising from confinement in the N-terminal form of *trans* GroEL: T181 and A109 at the bases and K65 and D140 at the lateral faces of the form are indicated by circles containing dots in the axial and in a perpendicular projection. The corresponding ideal positions, indicated by open circles, belong to an indexed heptagram. The dashed lines indicate the existence of integral scaling relations discussed in the text. The same conventions are also used in the following figures. (*b*) N-terminal form of *cis* GroEL with folding points at form boundaries. The lateral folding points are related by planar integral scale-rotations: E63, K65 at $n_7 = -\frac{1}{2}$ and A2, S43, A133 at $n_7 = -1$. The two basal positions A109 and E172 at $n_7 = 0$ and $n_7 = -2$, respectively, are point-group equivalent. The planar indices are related by an integral scale-rotation and the axial indices by an integral parabolic transformation. This last type of relation is implicitly implied in the equivalences mentioned in other figures.

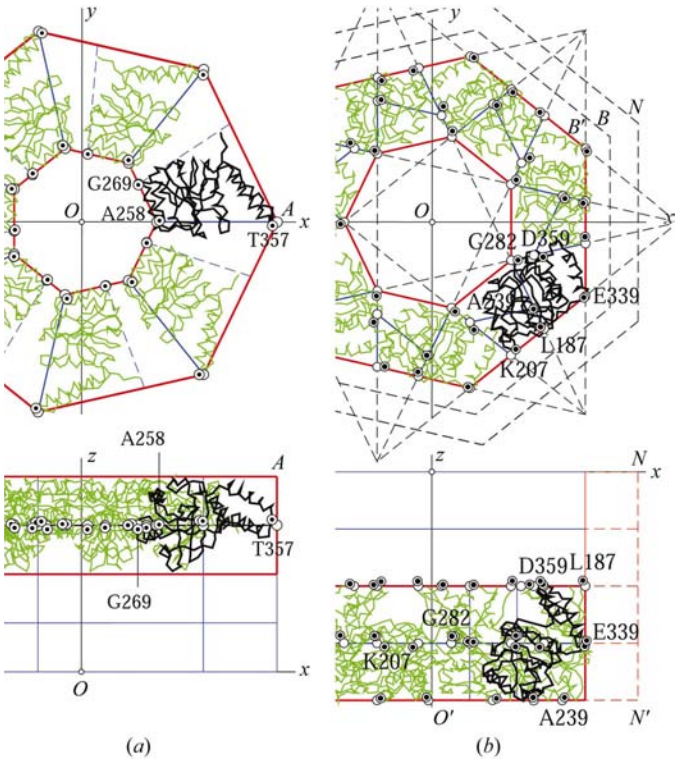

Figure 2
(*a*) Folding points at lateral form-boundaries belonging to the intermediate form of *trans* GroEL together with the corresponding ideal heptagram: all these positions share the same axial coordinate $n_7 = 3$. (*b*) The heptagrammal structure occurring in the intermediate form of *cis* GroEL consists of basal folding points (D359, L187 and A239) and of lateral-confined folding points (K207, G282, E339 and again L187, even if the latter represents the transition to the N-terminal form rather than a folding point). The corresponding ideal positions are all point-group equivalent. In the axial projection they appear at the vertices of the extended heptagram discussed in the appendix and shown in Fig. 9.

4.2. The intermediate form

The intermediate form of the *trans* ring corresponds to the apical domain (Xu *et al.*, 1997) and ranges from Leu187 to Val376. Three lateral C^{α} positions have been selected (A258, G269 and T357 indicated in Fig. 2*a*) and four basal ones (E209, K311, E315 and A340 shown in Fig. 3). All the lateral C^{α} positions have approximately the same height, fitting with the axial index $n_7 = 3$. Two are at the central hole and one at the vertex of the external heptagonal boundary (see Fig. 2*a*). Of the four basal C^{α} positions, three are at the top of the *trans* ring ($n_7 = 4$) and one at the lower basis, mid-height of the ring ($n_7 = 2$).

Starting with the lateral positions, one can immediately assign to T357 the ideal position $A(3)$, where $A(n)$ has the indices $(\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}, n)_t$. The ideal position of A258 then follows by the scaling transformation $S_{0.38}$. The linear scaling $R^{-3}Y_{-0.28}R^3$ transforms A258 into G269, where Y_{μ} scales the y coordinate by a factor μ and leaves the x coordinate invariant (see the appendix),

$$S_{0.38}(\mathrm{T357}) = \mathrm{A258}, \quad R^{-3}Y_{-0.28}R^3(\mathrm{A258}) = \mathrm{G269}. \qquad (10)$$

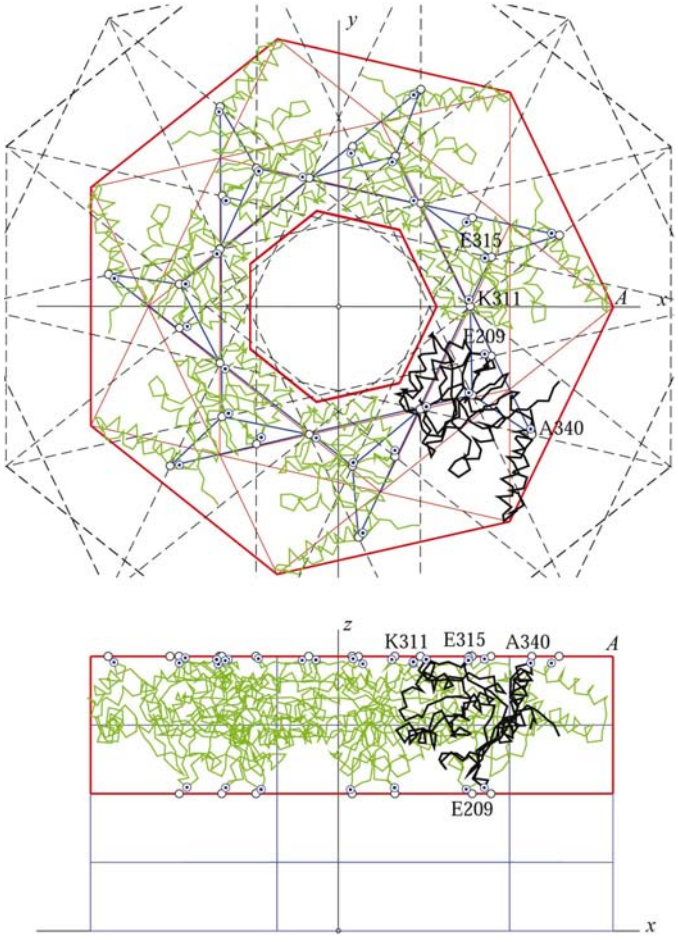

Figure 3
Folding points at the two basis faces with $n_7 = 4$ and $n_7 = 2$ of the same intermediate form of *trans* GroEL as in the previous Fig. 2(*a*). In the axial projection they show a heptagrammal arrangement with points related by integral scale-rotations.

These relations allow assignment of the following indices to the corresponding lateral-confined folding points indicated in Fig. 2(*a*),

$$\begin{aligned} \mathrm{T357} &= (\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}, 3)_t, \\ \mathrm{A258} &= (5\,1\,3\,3\,1\,5, 3)_t, \\ \mathrm{G269} &= (1\,0\,1\,1\,\bar{1}\,2, 3)_t. \end{aligned} \qquad (11)$$

Moving to the basal positions E209, K311, E315 and A340 shown in Fig. 3, it can be seen in the view along the heptagonal axis that these positions are arranged according to a highly symmetric pattern and that K311 follows from the position $A(4)$ by repeating twice a {7/2} star-heptagon construction

$$\mathrm{K311} = S^2_{0.69}A(4) = (\bar{9}\,\bar{2}\,\bar{6}\,\bar{6}\,\bar{2}\,\bar{9}, 4)_t. \qquad (12)$$

These indices are fairly high, but still acceptable. Not acceptable, however, are those which follow for the other basal points, namely

$$\begin{aligned} \mathrm{E315} &= (81\,\bar{13}\,75\;\;10\,39\,52, 4)_t, \\ \mathrm{A340} &= (\bar{20}\,3\,\bar{19}\,\bar{3}\,\bar{10}\,\bar{13}, 4)_t, \\ \mathrm{E209} &= (52\,39\,10\,75\,\bar{13}\,81, 2)_t. \end{aligned} \qquad (13)$$

These high indices have been computed using the symmetry of the arrangement. For example, K311 and A340 are in the same mutual position as the points labelled C_0 and $-V_3$ in the reverse heptagram of Fig. 9 (compare with the indices of Table 1 in the appendix). This observation suggests changing the basis (a_t, c_t) to an equivalent basis $(a_{t'}, c_{t'})$ defined by

$$(a_{t'}, c_{t'}) = (\lambda_T a_t, c_t), \quad \lambda_T = 1.3418\ldots \qquad (14)$$

obtained by the scaling transformation $T = S^2_{0.69}S^{-1}_{0.35}$. The indices obtained with respect to this local basis of the $\mathbb{Z}$-module M_t are quite reasonable,

$$\begin{aligned} \mathrm{E209} &= (2\,1\,0\,3\,\bar{1}\,3, 2)_{t'}, \\ \mathrm{K311} &= (1\,2\,0\,0\,2\,1, 4)_{t'}, \\ \mathrm{E315} &= (3\,\bar{1}\,3\,0\,1\,2, 4)_{t'}, \\ \mathrm{A340} &= (1\,1\,1\,0\,1\,2, 4)_{t'}. \end{aligned} \qquad (15)$$

All these positions are point-group equivalent. Their mutual relations are independent of the basis chosen for the indices. It follows that

$$\begin{aligned} \mathrm{A340} &= R^{-3}Y^{-1}_{0.28}R^3(\mathrm{K311}), \\ \mathrm{E315} &= S_{0.69}R(\mathrm{A340}), \\ \mathrm{E209} &= m_y(\mathrm{E315}), \end{aligned} \qquad (16)$$

where $Y_{0.28}$ denotes a linear scaling in the y direction by a factor $0.28\ldots$ and m_y the reflection which changes the sign of the y coordinate.

4.3. The C-terminal form

The C-terminal form of *trans* GroEL has segments ranging from Val376 to Pro525 (or to Lys526 in the uncomplexed structure determination reported in 1der). The mobile part of the remaining 23 residues of each chain is missing (Xu *et al.*, 1997). One basal position (E434) and three lateral positions

(K425, N487 and P525) have been identified and indexed (see Fig. 4*a*). An additional position (A384) has been included because it could possibly be point-group equivalent with K425, as their axial indices differ by two. One finds

$$\begin{aligned} \text{lateral}: \quad \text{A384} &= (1\,1\,0\,0\,1\,1, 5/2)_t, \\ \text{K425} &= (\bar{1}\,2\,\bar{3}\,1\,0\,0, 1/2)_t, \\ \text{N487} &= (0\,0\,\bar{1}\,0\,0\,0, 1)_t, \\ \text{P525} &= (\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,0\,\bar{2}, 1)_t, \qquad (17) \\ \text{basal}: \quad \text{E434} &= (0\,\bar{1}\,\bar{1}\,0\,0\,\bar{1}, 0)_t. \qquad (18) \end{aligned}$$

As already mentioned, P525 is not the C-terminus of the chain, but only the last C^{α} whose position has been determined. It can safely be assumed that it forms the pivot of the more mobile tail and can therefore be considered as a lateral confined folding point. It is related to N487, which is also lateral at the same height, and to the basal E434 at the interface with the *cis* ring,

$$\text{P525} = S_{0.24}S_{0.80}^{-1}R^{-2}P_tR^2(\text{E434}) = S_{-0.24}R^{-3}(\text{N487}). \qquad (19)$$

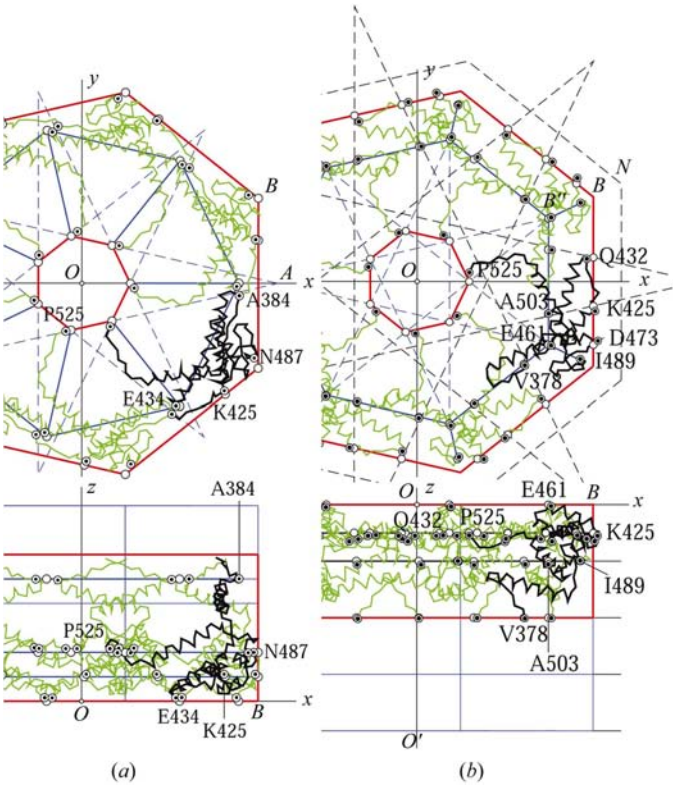

Figure 4
(*a*) C-terminal form of *trans* GroEL. The two pairs of folding points at the lateral faces are in integral scale-rotation relations and differ by an integer in their n_7 coordinate: A384 and K425 by $\delta z = 2$ and P525 and N487 by $\delta z = 0$. The basal folding point E434 at $z = 0$ is in a integral scale-rotation relation with P525 and N487. (*b*) C-terminal form of *cis* GroEL. The folding points at form boundaries occur in two pairs of scale-rotational related positions of star heptagons (indicated by dashed lines). There is a planar set with $n_7 = -\frac{1}{2}$ with D473, Q432, P525 and K425 and a planar set having axial coordinates which differ by an integer: E461 with $n_7 = 0$, I489, A503 with $n_7 = -1$ and V378 at the form boundary $n_7 = -2$.

The planar indices of K425 follow from the observation that its position is scaled from a vertex along the edge of the external heptagon (Fig. 4*a*), in a similar way as found for T181 along the edge of a star heptagon (see Fig. 1*a*). A384 has the same planar indices as E434 but a $\mathbb{Z}$-module inequivalent height.

5. Heptagrammal folding in the *cis* ring of GroEL

5.1. The N-terminal form

As in the *trans* ring, the N-terminal form of the *cis* ring ranges from the residues Ala2 to Leu187, now of the chains *A* to *G* of GroEL. The folding is strongly influenced by the presence of the co-chaperonin GroES and the confinement of the seven chains in a $\mathbb{Z}$-module indexed form is only very approximate. Difficulty in indexing point-group correlated folding positions would therefore be expected. In reality this indexing is straightforward, even if the discrepancy between observed and ideal positions is larger than that found so far. As in the previous case, the two basal positions in a chain appear to be correlated with the five lateral positions and are shown together in Fig. 1(*b*). The residue A2 has been included (even if strictly speaking it is not a folding point) because it fits nicely at the heptagonal boundary of the central channel. The indices given are the coordinates with respect to the basis (a_c, c_c) of the local $\mathbb{Z}$-module M_c of (4). The parabolic transformation connecting different axial indices n_7 is P_c. With respect to the new basis, this transformation has the same matrix representation as P_t with respect to the (a_t, c_t) basis [see equation (33) of part I]. We still write $P_cA(n) = A(n-1)$ for $A(n)$ with indices $(\bar{1}, \ldots, \bar{1}, n)_c$ and in a similar way $P_cB(n) = B(n+1)$ for $B(n) = (1, \ldots, 1, n)_c$.

The positions labelled in the view along the axis of Fig. 1(*b*) belong to the chain *D*. The corresponding indices are

$$\begin{aligned} \text{basal}: \quad \text{A109} &= (2\,0\,1\,1\,0\,2, 0)_c, \\ \text{E172} &= (\bar{1}\,\bar{1}\,0\,\bar{1}\,\bar{1}\,0, \bar{2})_c, \qquad (20) \\ \text{lateral}: \quad \text{A2} &= (0\,1\,\tfrac{1}{2}\,\tfrac{\bar{1}}{2}\,\tfrac{1}{2}\,1, \bar{1})_c, \\ \text{S43} &= (1\,\bar{1}\,\bar{1}\,1\,0\,\bar{1}, \bar{1})_c, \\ \text{K65} &= (1\,2\,0\,0\,2\,1, \tfrac{\bar{1}}{2})_c, \\ \text{E63} &= (0\,6\,1\,\bar{2}\,5\,3, \tfrac{\bar{1}}{2})_c, \\ \text{A133} &= (0\,1\,\bar{3}\,2\,\bar{1}\,0, \bar{1})_c. \qquad (21) \end{aligned}$$

If one forgets about A2, a better fit with these indices would have been obtained by adopting a slightly greater value for the basis parameter c_c, but we have no justification for doing so. The same non-integral value of the axial index in K65 and E63 implies non-equivalence with the positions having integral values and suggests a possible scaling relation between the two. Both lateral positions E63 and A133 are in a off-centre scaling relation by a factor 0.2469 ... along a heptagonal edge. The remaining structural relations are between point-group-equivalent positions, easily expressed with respect to $B(0)$,

$$\begin{aligned} \text{A109} &= S_{-0.69}B(0), \\ \text{S43} &= R^{-1}S_{-0.35}P_c^{-1}B(0), \\ \text{E172} &= S_{0.80}R^3P_c^{-2}B(0). \end{aligned} \tag{22}$$

5.2. The intermediate form

In this case the segments of the chains *A* to *G* range from Leu187 to Val378. Despite the different folding, the correlated folding points of the *trans* and *cis* intermediate forms have similar constraints requiring an equivalent local basis $(a_{c'}, c_{c'})$ to ensure low indices. In the present case,

$$(a_{c'}, c_{c'}) = (\mu_N a_c, c_c), \quad \mu_N = 0.86294\ldots \tag{23}$$

The indexed position $(1\ 1\ \ldots\ 1, n)_{c'}$ is indicated by $B'(n)$, so that $B'(n) = S_{0.86}B(n)$. The parabolic element $P_{c'}$ transforms $B'(n)$ into $B'(n+1) = P_{c'}B(n')$.

The lateral and the basal folding points fit into the ideal positions of a heptagram obtained by extending that given by the intersection points of the union of two star heptagons {7/2} ∪ {7/3} (see Fig. 9). Fig. 2(*b*) shows how well the two perpendicular views of the three basal and the three lateral folding points fit with the underlying $\mathbb{Z}$-module structure. All these positions are point-group equivalent with $B = B(0)$. The following scaling relations are found:

$$\begin{aligned} \text{E339} &= R^3B'(\bar{3}), \\ \text{G282} &= S_{0.69}^{-1}S_{0.35}(\text{E339}), \\ \text{L187} &= R^{-1}Y_{0.28}R^{-3}B'(\bar{2}), \\ \text{K207} &= R^{-1}Y_{-0.28}R(\text{E339}), \\ \text{A239} &= R^3Y_{0.38}R^3Y_{0.28}R^3B'(\bar{4}), \\ \text{D359} &= R^3Y_{0.11}RB'(\bar{2}). \end{aligned} \tag{24}$$

The indices of the selected C^α positions belonging to the chain *D* are

$$\begin{aligned} \text{basal}: \quad \text{L187} &= (3\,3\,0\,1\,4\,2, \bar{2})_{c'}, \\ \text{D359} &= (\bar{1}\,\bar{2}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{2}, \bar{2})_{c'}, \\ \text{A239} &= (\bar{3}\,\bar{1}\,0\,\bar{2}\,\bar{2}\,0, \bar{4})_{c'}, \end{aligned} \tag{25}$$

$$\begin{aligned} \text{lateral}: \quad \text{K207} &= (\bar{3}\,\bar{4}\,\bar{1}\,\bar{1}\,\bar{4}\,\bar{2}, \bar{3})_{c'}, \\ \text{G282} &= (\bar{5}\,\bar{3}\,1\,\bar{3}\,\bar{5}\,0, \bar{3})_{c'}, \\ \text{E339} &= (0\,0\,\bar{1}\,0\,0\,0, \bar{3})_{c'}. \end{aligned} \tag{26}$$

L187 is not actually a folding point and is simply a transition point to the neighbouring N-terminal form.

5.3. The C-terminal form

This form ranges from Val378 to Pro525. It has a terminal loop similar to the one occurring in the *trans* ring. It is now V378 that connects the intermediate and the C-terminal forms. It is included here together with A503 because both are mirror images of each other in projection and are in an off-centre scaling relation by a factor 0.246... with respect to the basal folding point E461. As can be seen from Fig. 4(*b*), in the axial projection E461 is in a scaling relation with the position I489. Moreover, the *z* coordinates of these four positions have integral values and could therefore be connected by the parabolic transformation P_c, once their equivalence to positions of the type $B(n) = (1\ \ldots\ 1, n)_c$ has been expressed. This setup leads to fairly high indices, which can be avoided by converting to an equivalent local basis, as performed in a previous case. We now define the new $\mathbb{Z}$-module basis of M_c by

$$(a_{c''}, c_{c''}) = (\mu_M a_c, c_c), \quad \mu_M = 0.643\ldots/0.862\ldots = 0.745\ldots \tag{27}$$

With respect to the new basis, the position $B''(n)$ has the same indices as $B(n)$ in the original basis (a_c, c_c), and the parabolic transformation $P_{c''}$ has the same matrix representation as P_c,

$$\begin{aligned} B''(n) = (1\,1\,1\,1\,1\,1, n)_{c''} &= S_{0.86}^{-1}S_{0.64}(1\,1\,1\,1\,1\,1, n)_c \\ P_{c''}(a_{c''}, c_{c''}) &= P_c(a_c, c_c). \end{aligned}$$

We now have the structural relations

$$\text{E461} = R^3B''(0), \quad \text{I489} = R^3S_{0.80}^{-1}B''(\bar{1}). \tag{29}$$

Despite the mutual relations indicated at the beginning, the point-group equivalence of V378 and A503 with E461 and I489 could not be proved.

The situation of another set of form-limited folding positions (P525, D473, Q432, K425) is simpler because these are co-planar. They all have the same (ideal) *z* coordinate $n_7 = -\frac{1}{2}$ (see Fig. 4*b*). P525 is at a vertex of the heptagonal channel and is scaled by $S_{0.35}$ from $B''(\bar{\frac{1}{2}})$ as indicated in Fig. 4(*b*) by the dashed lines of the smaller {7/3} star polygon. The other positions Q432, K425, D437 are along the lateral external boundary and follow from $B(\bar{\frac{1}{2}})$ by off-centre scalings with factors $-0.35\ldots, 0.35\ldots$ and $(0.35\ldots)^2$, respectively. In fact, these off-centre scalings correspond to linear scalings. Indeed, in particular

$$\begin{aligned} \text{K425} &= Y_{0.28}R^3B(\bar{\tfrac{1}{2}}), \\ \text{Q432} &= m_y(\text{K425}), \\ \text{D473} &= Y_{0.38}^{-1}(\text{K425}). \end{aligned} \tag{30}$$

Knowing these relations, the indices of the height positions labelled in Fig. 4(*b*) can be derived,

$$\begin{aligned} \text{basal}: \quad \text{V378} &= (\bar{1}\,2\,\bar{3}\,1\,0\,0, \bar{2})_{c''}, \\ \text{E461} &= (0\,0\,\bar{1}\,0\,0\,0, 0)_{c''}, \end{aligned} \tag{31}$$

$$\begin{aligned} \text{lateral}: \quad \text{P525} &= (1\,2\,0\,0\,2\,1, \bar{\tfrac{1}{2}})_{c''}, \\ \text{D473} &= (\bar{3}\,\bar{1}\,\bar{3}\,\bar{2}\,\bar{1}\,\bar{3}, \bar{\tfrac{1}{2}})_{c''}, \\ \text{Q432} &= (\bar{4}\,1\,\bar{5}\,0\,\bar{3}\,\bar{2}, \bar{\tfrac{1}{2}})_{c''}, \\ \text{K425} &= (\bar{2}\,\bar{3}\,0\,\bar{5}\,1\,\bar{4}, \bar{\tfrac{1}{2}})_{c''}, \end{aligned} \tag{32}$$

$$\begin{aligned} \text{additional}: \quad \text{I489} &= (0\,\bar{1}\,0\,\bar{1}\,0\,0, \bar{1})_{c''}, \\ \text{A503} &= (0\,1\,\bar{3}\,2\,\bar{1}\,0, \bar{1})_{c''}. \end{aligned} \tag{33}$$

6. The co-chaperonin GroES

The role of the extremal lateral point $A(n)$ in GroEL is taken over in GroES by $E(n)$, which in the basis (a_s, c_s) of the $\mathbb{Z}$-module M_s [see (4)] has the same corresponding indices: $E(n) = (\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}, n)_s$. The structural compatibility between GroEL and GroES appears in a clear way when expressing the position $E(0)$ at the interface between the two co-chaperonins with respect to their bases (a_c, c_c) and (a_s, c_s), respectively,

$$\begin{aligned} E(0) &= (\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}, 0)_s = (2\,0\,1\,1\,0\,2, 0)_c, \\ B(\bar{4}) &= (1\,1\,1\,1\,1\,1, \bar{4})_c = (1\,2\,1\,1\,2\,1, \bar{6})_s, \\ E(\bar{6}) &= S_{-0.69}B(\bar{4}). \end{aligned} \tag{34}$$

6.1. The N/I-combined form

Three forms have been derived in part I for GroES: an N-terminal form, an intermediate form and a C-terminal form. From the point of view of folding, however, the first two must be considered together, in an N/I-combined form, as can be seen by considering the three basal positions A31, G52, E53, the two lateral positions I25, E50 and the two positions V6 and G62 which are basal and lateral for the N-terminal form and the intermediate form, respectively (see Fig. 5). Looking at their z coordinates, it can be deduced that I25 and E50 belong together as do the remaining five positions. For the first two positions, the following structural relation is found:

$$\mathrm{E50} = R^2 S_{0.69} S_{0.24} P_s^{-2} R^{-3}(\mathrm{I25}). \tag{35}$$

The other five are also at point-group equivalent positions:

$$\begin{aligned} \mathrm{G52} &= R^2 S_{0.24} E(\bar{3}), \quad \mathrm{E53} = R^2 S_{0.30} E(\bar{3}), \\ \mathrm{G62} &= R^{-1} S_{-0.69} E(\bar{2}), \quad \mathrm{V6} = R^3 S_{-0.69}(\mathrm{G62}), \\ \mathrm{A31} &= R Y_{0.38}^{-1} R^2 S_{0.24} E(0), \end{aligned} \tag{36}$$

where $Y_{0.38}$ is a linear scaling of the y coordinate by the scaling factor indicated. The same transformation can also be expressed as an off-centre scaling at E along the edge of the star heptagon {7/2} by a factor 0.4450... The indices of these folding points are

$$\begin{aligned} \mathrm{I25} &= (0\,0\,1\,0\,0\,0, \tfrac{\bar{1}}{2})_s, \quad \mathrm{E50} = (\bar{5}\,1\,\bar{5}\,0\,\bar{3}\,\bar{3}, \tfrac{\bar{5}}{2})_s, \\ \mathrm{A31} &= (\bar{1}\,\bar{1}\,\bar{1}\,0\,\bar{1}\,\bar{2}, 0)_s, \quad \mathrm{G52} = (1\,\bar{1}\,1\,0\,0\,0, \bar{3})_s, \\ \mathrm{E53} &= (\bar{2}\,1\,\bar{2}\,0\,\bar{1}\,\bar{1}, \bar{3})_s, \quad \mathrm{G62} = (2\,1\,1\,2\,0\,2, \bar{2})_s, \\ \mathrm{V6} &= (\bar{7}\,2\,\bar{7}\,0\,\bar{4}\,\bar{4}, \bar{2})_s. \end{aligned} \tag{37}$$

6.2. The C-terminal form

This form is centred in height with respect to the previous combined forms and has the normal heptagonal prism boundaries. Despite the good fitting of the chain segments into the indexed form, the deviations between observed and ideal C^α positions (see Fig. 6) are larger than in most of the other cases, in particular for the positions D79 and S89. If this was the rule, this approach would not appear to be meaningful. In the present case, however, the ideal positions of the vertices of the form (implying that they are at the same time basal and lateral) could explain the corresponding sharp folding of the chain. The indices of these form-confined folding points are

$$\begin{aligned} \mathrm{Y71} &= (1\,2\,1\,0\,2\,2, \bar{1})_s, \quad \mathrm{D79} \simeq (\bar{1}\,\bar{1}\,0\,0\,\bar{1}\,0, \bar{1})_s, \\ \mathrm{G62} &= (\bar{2}\,0\,\bar{2}\,0\,\bar{1}\,\bar{1}, \bar{2})_s, \quad \mathrm{S89} \simeq (1\,2\,1\,0\,2\,2, \bar{2})_s, \\ \mathrm{A97} &= (0\,\bar{1}\,\bar{1}\,0\,0\,\bar{1}, \bar{2})_s. \end{aligned} \tag{38}$$

These positions are point-group equivalent as can be seen from the following structural relations, all expressed with respect to $E(\bar{2})$, so that the connection with the folding points of the previous forms becomes evident [see (36)],

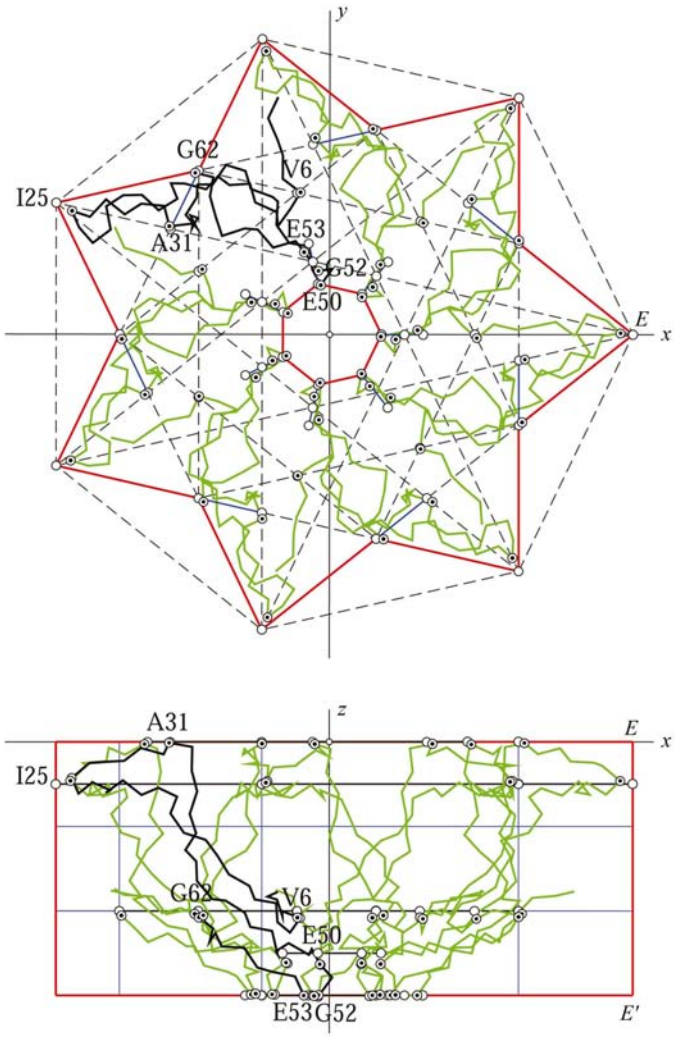

Figure 5
Correlated folding points of the N/I-combined form of GroES (where I indicates intermediate). The lateral folding points, sharing the planar heptagram of the form (I25 and E50), differ in their axial coordinates by an integer $\delta z = 2$. The basal folding points (E53, G52 with $n_7 = -4$ and A31 at $z = 0$) have planar indices related by integral scale-rotations. The two other positions G62 and V6, which are in the respective forms both planar and lateral, are in a {7/2} star-polygon relation and are point-group equivalent with the other folding points. G62 is a limit point shared by the intermediate form and by the C-terminal form (see Fig. 6).

$$\begin{aligned} \text{Y71} &= R^2 S_{-0.35} E(\bar{1}), \quad & \text{S89} &\simeq R^2 S_{-0.35} E(\bar{2}), \\ \text{A97} &= R^{-1} S_{0.80} E(\bar{2}), \quad & \text{D79} &\simeq R^{-2} S_{0.80} E(\bar{1}), \\ \text{G62} &= R^2 S_{-0.69} E(\bar{2}). \end{aligned} \tag{39}$$

Note that G62, which is a position shared by the two forms of GroES, is not presented in (37) and (38) with the same indices, because in Figs. 5 and 6 different chains have been selected for graphical reasons.

7. GroEL polymorphism: ATP/ADP-binding sites and folding points

There are two main consequences for GroEL of the formation of a complex with GroES.

(i) A change in the folding of the chains *A* to *G* in the *cis* ring, with a corresponding breaking of the dyadic symmetry between the *cis* and the *trans* rings of GroEL.

(ii) The elimination of the original ADP/ATP-binding sites from the *trans* ring in the asymmetric (complexed) state.

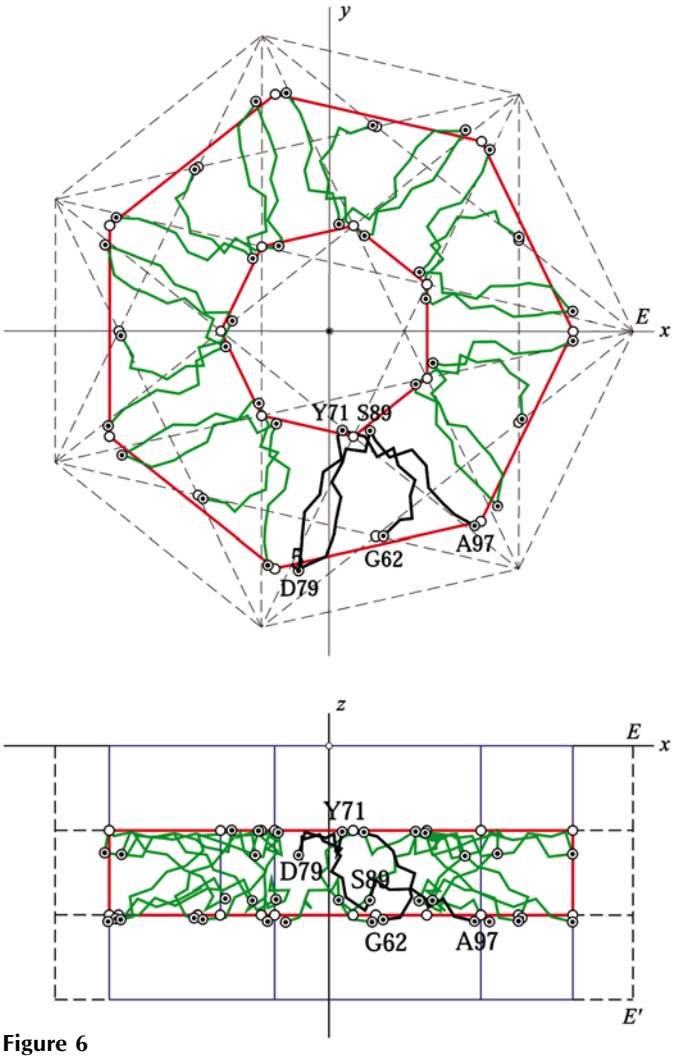

Figure 6
C-terminal form of GroES. The correlated folding points deviate from ideal indexed positions more than in previous cases, but have the remarkable property of being (approximately) at the vertices of the form itself and are both basally and laterally confined.

These facts imply a strong correlation between molecular forms, nucleotide binding and the alternative chain folding known as *chain polymorphism*. The existence of non-trivial relations between chain polymorphism and symmetry is suggested by the many form properties shared by the *trans* and the *cis* rings in the asymmetric state pointed out and summarized in part I. At the atomic level the following two questions naturally arise.

(i) What is the relation between the ATP/ADP-binding sites and the symmetry of the form where the nucleotide is embedded? In particular: what is the relation between the ATP/ADP positions and the point-group-correlated folding points?

(ii) What happens to form-confined folding points by the transition to an alternative form of the chain?

Within the limits of the present paper it is not possible to give a full answer to these questions. In the two following subsections it is only intended to suggest that the present approach represents an appropriate method for discussing these questions.

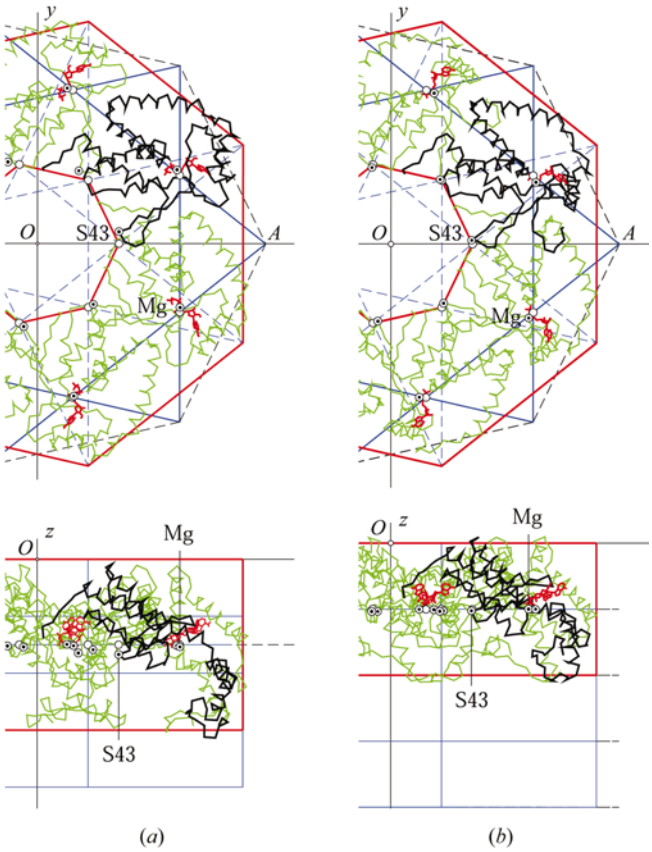

Figure 7
The binding sites of the nucleotides ADP and ATP are shown in the *cis* ring of GroEL–GroES–$(ADP)_7$ (PDB code 1aon) on the left and of GroEL–ATPγS (PDB code 1der) on the right. Despite the different folding, in both cases the bound magnesium ion, indicated by Mg, is related to the position *A* labelling an external boundary of GroEL (see in particular Figs. 1 and 2 of part I) by a star heptagon {7/2} with the same mid-height *z* coordinate of the corresponding form as the folding point S43. The latter is at the boundary of the same channel which is related by the star heptagon {7/3} to the position *B*, opposite *A*. Accordingly, Mg and S43 conserve the same integral scale-rotation relation during the polymorphic transformation.

7.1. Nucleotide-binding sites

In both GroEL states, the symmetric uncomplexed and the asymmetric complexed state, the nucleotide-binding site is in close contact with the equatorial subunits of the N- and C-termini of a chain (Xu *et al.*, 1997). This implies that the corresponding terminal forms should be considered in the complexed and the uncomplexed cases (Figs. 1*a* and 1*b*, and 4*a* and 4*b*). Here, only the N-terminal forms are compared (Fig. 7).

The first observation is that both forms have the same lateral confinement with a central heptagonal channel scaled by $S_{-0.35}$ with respect to the external enclosing heptagon (see Fig. 1*a* and 1*b*). Both have S43 at a vertex position of the channel at a mid-height of the corresponding form: $n_7 = -1$ if complexed and $n_7 = -\frac{3}{2}$ if uncomplexed.

The second observation is that the bound Mg^{2+} ion plays a similar role for the nucleotides as the C^{α} for the amino acids and occurs in the same equatorial plane as S43 (Janner, 2003*b*).

Finally, in both states the Mg^{2+} ions are related to the *A* heptagon as in a star polygon {7/2}, the reverse of the *B* heptagon of the external boundary.

All these properties are the consequence of the structural relation

$$\mathrm{Mg} = S_{-0.69} S_{0.35}^{-1}(\mathrm{S43}) \qquad (40)$$

between a folding point at a form boundary and a nucleotide-binding position at the magnesium ion position Mg in both the complexed and the uncomplexed states. The indices corresponding to the positions labelled in Fig. 7 are given by

$$\begin{aligned} \text{uncomplexed}: \quad & \mathrm{S43} = (1\,2\,0\,0\,2\,1, \bar{1})_t, \quad \mathrm{Mg} = (1\,\bar{1}\,1\,\bar{1}\,1\,0, \bar{1})_t \\ \text{complexed}: \quad & \mathrm{S43} = (1\,2\,0\,0\,2\,1, \bar{\tfrac{3}{2}})_c, \quad \mathrm{Mg} = (1\,\bar{1}\,1\,\bar{1}\,1\,0, \bar{\tfrac{3}{2}})_c. \end{aligned} \qquad (41)$$

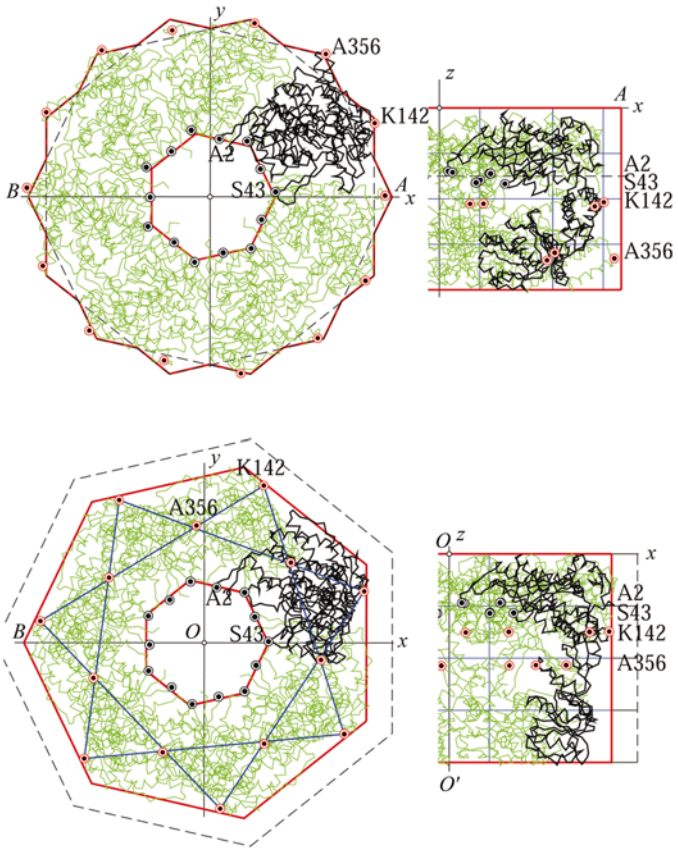

Figure 8
Folding points which are lateral form boundaries and point-group equivalent in the *cis* ring of the uncomplexed GroEL (PDB code 1der, shown in the upper part) are still point-group equivalent in the complexed state (PDB code 1aon, shown in the lower part), despite the change in folding evident from the projections perpendicular to the sevenfold axis. A2 and S43 remain at the boundary of the central channel (with same radial size but different heights). A356 and K142, which in the uncomplexed state have a radial extremal position (corresponding to *A* and *B*, respectively) in the complexed state appear in projection at the vertices of a common star heptagon {7/2}. Note that positions having integral axial coordinates have half-integral values and *vice versa*. This is possibly purely accidental.

7.2. Correlated folding points and chain polymorphism

The previous subsection gave an example of a correlated folding point (S43) which conserves its character during the polymorphic transition (Figs. 7 and 8). This is, however, not the general case. In this subsection, another type of conservation is exemplified.

Considered are the positions A2, S43, K142, and A356. In the uncomplexed state all are lateral folding points with a heptagrammal ordering with integral and half-integral axial indices n_7. A similar type of heptagrammal ordering is observed for these positions in the complexed state. Only A2 and S43 remain at the conserved boundary of the central channel, whereas K142 and A356, instead of being at the vertices of heptagons in a reverse orientation, in the complexed state are at the vertices of a star heptagon {7/2} of different size and different orientation. Moreover, in the polymorphic transition the *z* coordinates with an integral value change into a half-integral one and *vice versa*.

It is not expected that this behaviour reflects a general situation. Typically in Figs. 1(*a*) and 2(*a*) the neighbouring positions D140 and T357 have been selected and not K142 and A356, as here. This underlines the complexity of the order of the GroEL–GroES chaperonin, where certainly many more structural relations exist than those analyzed in this paper.

8. Concluding remarks

This paper does not pretend to solve the general problem of protein folding. The present approach is interpretative, phenomenological and non-predictive. Furthermore, only a given type of folding points has been analyzed: those at the boundaries of prismatic molecular forms having vertices with integral coordinates related by scale-rotation transformations represented by integral invertible matrices.

The atomic order is very complex, as expected. Part of the structure reflects an abstract higher dimensional crystallographic order and this is unexpected. The higher dimensional

lattice structure appears in projection in the three-dimensional physical space as a $\mathbb{Z}$-module, as in quasi-crystals, but in the molecular case without translations (Janner, 2001). What is observed in the GroEL–GroES chaperonin is only a very incomplete image of a higher dimensional symmetric object. This image is much more incomplete than that obtained from cut-projection in the case of quasi-crystals. This situation requires extension of the familiar concept of *symmetry* to that of *structural relations*, which imposes less rigid constraints. Nevertheless, the ideal underlying crystallographic order can be identified, both at the geometric and at the algebraic levels presented in the appendix. Conceptually, the problem is the same as for crystals: from the geometry of forms with rational indices, find the atomic positions related by invertible integral matrix transformations (Janner, 2001). In crystals the atomic structure is eventually arrived at *via* X-ray diffraction. For biomacromolecules, the sequence in time is reversed: starting from an atomic structure known from X-ray diffraction, attempts are made to characterize the architecture in terms of a possible (higher-dimensional) crystallographic order. This is, of course, not the only reasonable way of obtaining an insight into the biomacromolecular building principles.

APPENDIX *A*

A1. Algebraic integers and their geometry in the heptagonal case

The analysis of the architecture of GroEL–GroES occurs at two levels.

(i) The geometric level, with the identification of a number of C^{α} positions in a heptagrammal arrangement. This allows assignment of integral indices $(n_1, \ldots, n_7)$ to these special positions expressed with respect to a $\mathbb{Z}$-module basis of the type (a, c) as given in (1).

(ii) The algebraic level, in terms of invertible integral matrices relating heptagrammal positions. Related atomic positions are said to be point-group equivalent.

Equivalent heptagrammal positions have z coordinates which differ by an integral value δz. Two cases are considered:

(i) Co-planar equivalent positions: $\delta z = 0$.

(ii) Positions at $z_0 + \delta z$ equivalent to $A(z_0) = (\bar{1}, \ldots, \bar{1}, z_0)$: δz = integer. This case is treated as $(\bar{1}, \ldots, \bar{1}, 0)$ with respect to a z_0-shifted basis (a, c).

In the second case, the parabolic transformation P with matrix representation

$$P(a,c) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \qquad (42)$$

allows the z coordinates occurring in a heptagram which differ by integral values to be related, so that the problem is reduced to planar equivalence. The general case has not been investigated so far.

Table 1
Indexed intersection points of the heptagram in the upper part of Fig. 9.

By changing all the signs, the corresponding indices of the reverse heptagram are obtained.

k	A_k	$-C_k$	D_k	$-E_k$	U_k	V_k
0	$\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}$	$\bar{1}\,\bar{2}\,0\,0\,\bar{2}\,\bar{1}$	$2\,1\,1\,1\,1\,2$	$\bar{2}\,0\,\bar{1}\,\bar{1}\,0\,\bar{2}$	$0\,0\,1\,\bar{1}\,0\,1$	$1\,0\,\bar{1}\,1\,0\,0$
1	$1\,0\,0\,0\,0\,0$	$\bar{1}\,\bar{2}\,0\,0\,\bar{2}\,\bar{1}$	$\bar{2}\,0\,\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}$	$2\,0\,2\,1\,1\,2$	$\bar{1}\,\bar{1}\,\bar{1}\,0\,\bar{2}\,\bar{1}$	$0\,1\,0\,\bar{1}\,1\,0$
2	$0\,1\,0\,0\,0\,0$	$1\,0\,\bar{1}\,1\,1\,\bar{1}$	$1\,\bar{1}\,1\,0\,0\,0$	$\bar{2}\,0\,\bar{2}\,0\,\bar{1}\,\bar{1}$	$1\,0\,0\,0\,1\,\bar{1}$	$0\,0\,1\,0\,\bar{1}\,1$
3	$0\,0\,1\,0\,0\,0$	$\bar{2}\,\bar{1}\,0\,\bar{1}\,\bar{2}\,0$	$0\,1\,\bar{1}\,1\,0\,0$	$1\,\bar{1}\,1\,\bar{1}\,1\,0$	$1\,2\,1\,1\,1\,2$	$\bar{1}\,\bar{1}\,\bar{1}\,0\,\bar{1}\,\bar{2}$
4	$0\,0\,0\,1\,0\,0$	$0\,\bar{2}\,\bar{1}\,0\,\bar{1}\,\bar{2}$	$0\,0\,1\,\bar{1}\,1\,0$	$0\,1\,\bar{1}\,1\,\bar{1}\,1$	$\bar{2}\,\bar{1}\,0\,\bar{1}\,\bar{1}\,\bar{1}$	$2\,1\,1\,1\,2\,1$
5	$0\,0\,0\,0\,1\,0$	$2\,2\,0\,1\,2\,1$	$0\,0\,0\,1\,\bar{1}\,1$	$\bar{1}\,\bar{1}\,0\,\bar{2}\,0\,\bar{2}$	$1\,\bar{1}\,0\,1\,0\,0$	$\bar{1}\,1\,0\,0\,0\,1$
6	$0\,0\,0\,0\,0\,1$	$\bar{1}\,1\,1\,\bar{1}\,0\,1$	$\bar{1}\,\bar{1}\,\bar{1}\,\bar{1}\,0\,\bar{2}$	$2\,1\,1\,2\,0\,2$	$0\,1\,\bar{1}\,0\,1\,0$	$\bar{1}\,\bar{2}\,0\,\bar{1}\,\bar{1}\,\bar{1}$

Accordingly, in the following the discussion is restricted to a two-dimensional rank six heptagonal $\mathbb{Z}$-module and to the corresponding planar indices $(n_1, \ldots, n_6)$.

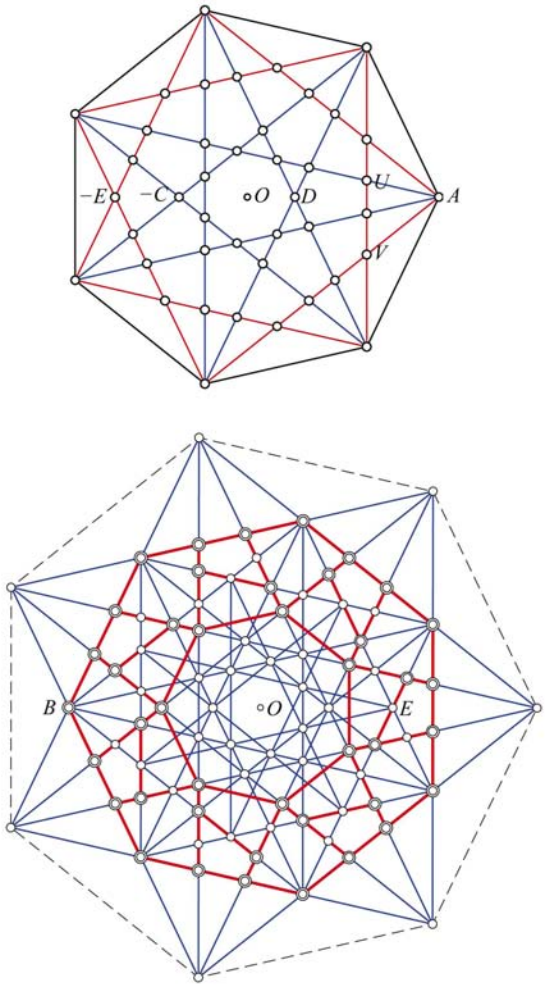

Figure 9
(*a*) Heptagram obtained from the intersection of the two star heptagons {7/2} and {7/3}. The indices of the vertices labelled as in other figures are indicated in Table 1. (*b*) Extended version of the previous heptagram (in a reverse orientation) occurring at heptagrammal folding points confined by the intermediate form of *cis* GroEL (see Fig. 2*b*). The correspondence is outlined and the ideal folding positions are indicated by double circles. These are point-group equivalent. The other encircled vertices of this heptagram have integral indices, but their point-group equivalence has not been proved.

Table 2
Parameter values and corresponding scaling factors for integral planar scalings $S_{\mu(t_0,t_1,t_2)}$ and linear scalings $X_{\mu(t_0,t_1,t_2)}$, $Y_{\mu(t_0,t_1,t_2)}$, with entries in S_μ and S_μ^{-1} of absolute value not larger than 5.

The linear scaling possibility is marked by *. For each scaling factor the number of times its absolute value occurs in the structural relations derived so far is also given for the GroEL *trans* and *cis* rings, for GroES and the bound Mg ion, even if these values give a global impression only.

t_0	t_1	t_2	$\mu(t_0, t_1, t_2)$	Heptagram	*trans*	*cis*	GroES	Mg	Total
1	0	2	0.1099...*			1			1
−2	1	−2	0.1370...						
−1	2	3	0.1588...						
1	−1	−1	0.1980...						
−1	1	0	0.2469...	{7/3}			3		3
1	−2	−4	0.2862...*	{7/2} ∪ {7/3}	2	3			5
2	−1	1	0.3079...				1		1
0	1	2	0.3568...	{7/3}	4	2	2		8
−3	2	−2	0.3840...*	{7/2} ∪ {7/3}	1	2	1		4
0	0	−1	0.4450...						
1	0	1	0.5549...		1				1
1	−1	−2	0.6431...			1			1
−1	1	−1	0.6920...	7/2	2	2	3	1	8
0	1	1	0.8019...		1	1	2		4
3	−1	2	0.8629...			4			4

A1.1. Geometry of heptagrams. The simplest heptagrams are the regular heptagon {7/1} and the star heptagons {7/2} and {7/3} found in Coxeter (1961) and in Fig. 8 of part I. The intersection points of these three simple cases are shown in Fig. 9 in the direct orientation. All have integral indices, as indicated in Table 1, with the corresponding labels A_k to E_k, U_k and V_k, where the subscript k indicates a point rotated counter-clockwise by an angle $2\pi k/7$ (normally, the subscript $k = 0$ is omitted). The indices of the points of the heptagram in the reverse orientation are obtained by inverting all the signs.

Most heptagrams of the forms found in GroEL–GroES are obtained by adding simple heptagonal patterns to these two basic heptagrams. The more complex case of the intermediate form in the *cis* ring of GroEL (see Fig. 2*b*) can be analyzed by combining an extended version of the {7/2} and {7/3} star heptagons, as shown in the lower part of Fig. 9. All the encircled intersection points of this extended heptagram have integral indices.

Folding points at lateral form boundaries frequently lead to heptagrams with points on heptagonal (or star-heptagonal) edges. Their indexing is not easy. Only few of the edge-points can be obtained by linear scaling, as exemplified in Fig. 2(*b*) and correspondingly in Fig. 9. Other low-index edge positions follow from off-centre scaling centred at a vertex. Several examples of this last case have been obtained from off-centre scaling with scaling factor 0.246... (see Figs. 1*a* and 1*b*, and 4*a* and 4*b*). Their structural meaning is still obscure, despite the fact that the same factor also occurs for one of the two planar scalings of the star heptagon {7/3} and that, in general, it ensures low indices.

A1.2. Point-group transformations. Discussed here are the automorphisms of the two-dimensional $\mathbb{Z}$-module M which play a role in the structural relations derived in this paper for C^α positions in GroEL–GroES at the boundaries of molecular forms. Some of these occur in the point groups of the enclosing forms and have been given in the appendix of part I. In particular, this is the case for the rotation R by an angle of $2\pi/7$ around the z axis and the reflection m that leaves the x axis (and in three dimensions also the z axis) invariant. Scaling transformations occurring in the star heptagons {7/2} and {7/3} have also been indicated, together with a couple of other cases.

Here, the general case of a two-dimensional planar scaling with scaling factor μ leaving the heptagonal $\mathbb{Z}$-module invariant and represented by a six-dimensional integral matrix S_μ is parameterized by three integers t_0, t_1, t_2 according to

$$S_{\mu(t_0,t_1,t_2)} = \begin{pmatrix} t_0 - t_1 & t_1 - t_2 & t_2 & 0 & -t_2 & -t_1 + t_2 \\ 0 & t_0 - t_2 & t_1 & t_2 & -t_2 & -t_1 \\ -t_1 + t_2 & t_1 - t_2 & t_0 & t_1 & 0 & -t_1 \\ -t_1 & 0 & t_1 & t_0 & t_1 - t_2 & -t_1 + t_2 \\ -t_1 & -t_2 & t_2 & t_1 & t_0 - t_2 & 0 \\ -t_1 + t_2 & -t_2 & 0 & t_2 & t_1 - t_2 & t_0 - t_1 \end{pmatrix}, \quad (43)$$

where the scaling factor is given by

$$\mu(t_0, t_1, t_2) = t_0 + 2\,t_1 \cos\frac{2\pi}{7} + 2\,t_2 \cos\frac{4\pi}{7}. \quad (44)$$

The condition $\det S_{\mu(t_0,t_1,t_2)} = 1$ ensures that the inverse scaling matrix also has integral entries. The indices of the transformed vertices $S_\mu A_k$ of the heptagon appear as the kth column vector of the matrix S_μ. A general point with Cartesian coordinates (x, y) is transformed according to

$$S_\mu(x, y) = (\mu x, \mu y). \quad (45)$$

For $|t_i| < 4$ and $0 < \mu < 1$, 21 different solutions are found. If it is furthermore required that the scaling matrices and their inverse have entries with maximal absolute value 5 (as suggested by B. Souvignier), the number of solutions is reduced to 12, as indicated in Table 2.

There is a direct relation between these automorphisms and the group of units of the cyclotomic field $\mathcal{F}(\zeta)$ with $\zeta = \exp(2\pi/7)$ (see, for example, Cohn, 1962), but this aspect is not discussed here. For self-similar quasi-crystals, this group of units must be enlarged by linear scaling (Janner, 1997). Moreover, scaling factors of both signs occur, despite the fact that the Euclidean point-group symmetry of the regular heptagon does not contain the total inversion.

The linear scalings considered here are automorphisms of M which only scale one of the Cartesian coordinates of a point (x, y) in the plane

$$X_\mu(x, y) = (\mu x, y), \quad Y_\mu(x, y) = (x, \mu y). \quad (46)$$

Their product is the planar scaling $S_\mu = X_\mu Y_\mu$. As previously discussed Janner (2002) and in the appendix of part I, linear scalings demonstrate that the positions labelled U and V in the heptagram {7/2} ∪ {7/3} of Fig. 9 are point-group equivalent with all the other intersection points of this heptagram.

From the parametric representation of $S_{\mu(t_0,t_1,t_2)}$ given above one derives by factorization a corresponding parameterization of $X_{\mu(t_0,t_1,t_2)}$ and $Y_{\mu(t_0,t_1,t_2)}$:

$$X_{\mu(t_0,t_1,t_2)} = \frac{1}{2}\begin{pmatrix} 1+t_0-2t_1+t_2 & t_1-2t_2 & t_2 & t_2 & t_1-2t_2 & 1+t_0-2t_1+t_2 \\ -t_1 & 1+t_0-2t_2 & t_1+t_2 & t_1+t_2 & -1+t_0-2t_2 & -t_1 \\ -2t_1+t_2 & t_1-t_2 & 1+t_0+t_1 & -1+t_0+t_1 & t_1-t_2 & -2t_1+t_2 \\ -2t_1+t_2 & t_1-t_2 & -1+t_0+t_1 & 1+t_0+t_1 & t_1-t_2 & 2t_1+t_2 \\ -t_1 & -1+t_0-2t_2 & t_1+t_2 & t_1+t_2 & 1+t_0-2t_2 & -t_1 \\ -1+t_0-2t_1+t_2 & t_1-2t_2 & t_2 & t_2 & t_1-2t_2 & 1+t_0-2t_1+t_2 \end{pmatrix} \quad (47)$$

and

$$Y_{\mu(t_0,t_1,t_2)} = \frac{1}{2}\begin{pmatrix} 1+t_0-t_2 & t_1 & t_2 & -t_2 & -t_1 & 1-t_0+t_2 \\ t_1 & 1+t_0 & t_1-t_2 & -t_1+t_2 & 1-t_0 & -t_1 \\ t_2 & t_1-t_2 & 1+t_0-t_1 & 1-t_0+t_1 & -t_1+t_2 & -t_2 \\ -t_2 & -t_1+t_2 & 1-t_0+t_1 & 1+t_0-t_1 & t_1-t_2 & t_2 \\ -t_1 & 1-t_0 & -t_1+t_2 & t_1-t_2 & 1+t_0 & t_1 \\ 1-t_0+t_2 & -t_1 & -t_2 & t_2 & t_1 & 1+t_0-t_2 \end{pmatrix}. \quad (48)$$

The possible values of the parameters are restricted by the conditions:

$$\det X_{\mu(t_0,t_1,t_2)} = 1, \quad \det Y_{\mu(t_0,t_1,t_2)} = 1, \quad t_0 \text{ odd}, \quad t_1,\ t_2 \text{ even}. \quad (49)$$

Only three of the 21 possible parameter values for $|t_i| < 4$ satisfy these conditions. In Table 2 these cases are marked with an asterisk.

The careful reading of the first manuscript by B. Souvignier and his pertinent suggestions have greatly helped the author to improve the presentation. Thanks are expressed to R. de Gelder for valuable remarks and to Annalisa Fasolino for stimulating discussions.

References

Boisvert, D. C., Wang, J., Otwinowski, Z., Horwich, A. L. & Sigler, P. B. (1996). *Nature Struct. Biol.* **3**, 170–177.
Braig, K., Otwinowski, Z., Hedge, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L. & Sigler, P. B. (1994). *Nature (London)*, **371**, 578–586.
Cohn, H. (1962). *A Second Course in Number Theory*. New York: John Wiley.
Coxeter, H. S. M. (1961). *Introduction to Geometry*. New York: John Wiley.
Janner, A. (1997). *GROUP21 Physical Applications and Mathematical Aspects of Geometry, Groups and Algebra*, Vol. 2, edited by H.-D. Doebner, W. Scherer & C. Schulte, pp. 949–953. Singapore: World Scientific.
Janner, A. (2001). *Acta Cryst.* A**57**, 378–388.
Janner, A. (2002). *Struct. Chem.* **13**, 279–289.
Janner, A. (2003*a*). *Acta Cryst.* D**59**, 783–794.
Janner, A. (2003*b*). *Proteins*, **51**, 126–136.
Xu, Z., Horwich, A. L. & Sigler, P. B. (1997). *Nature (London)*, **388**, 741–750.