Acta Cryst. (1996). B52, 882–891

Symmetry-Modified Conformational Mapping and Classification of the Medium Rings from Crystallographic Data. IV. Cyclooctane and Related Eight-Membered Rings

FRANK H. ALLEN,[a]* JUDITH A. K. HOWARD[b]* AND NIGEL A. PITCHFORD[b]

[a]*Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, England, and* [b]*Department of Chemistry, University of Durham, South Road, Durham CH1 3LE, England*

(*Received 15 April 1996; accepted 5 June 1996*)

Abstract

Crystallographic observations of eight-membered ring conformations, retrieved from the Cambridge Structural Database, have been mapped and classified using symmetry-adapted deformation coordinates, principal component analysis and cluster analysis. Seven subsets of eight-membered rings, containing 11–32 conformational observations, have been analysed: cyclooctane (dataset 8C1), cyclooctene (8C2), cycloocta-1,3-diene (8C3), mono-*exo*-unsaturated carbocycles (8C4), monohetero (8A1), 1,5-dihetero (8A2) and 1,3,5,7-tetrahetero rings (8A3). The energetically preferred (by $\sim$7 kJ mol^{-1}) boat–chair form is adopted by 26 of the 32 examples of 8C1, although varying degrees of twist are induced by fusion to rings of sizes three, four and five. Crystallographic results for other subsets also populate the lower-energy areas of the appropriate potential energy hypersurface, but the analyses are complicated by the effects of ring fusion and by the small numbers of relevant crystal structures in some cases.

1. Introduction

The purpose of this short series of papers is to examine, map and classify the conformations of the medium rings (carbocyclic and heterocyclic) using experimental observations obtained by crystal structure analysis. Conformational analyses of this type are of fundamental chemical interest, since it is reasonable to assume that the crystallographic observations will be closely aligned with the low-energy features of the relevant potential energy hypersurface (see Allen, Harris & Taylor, 1996, and references therein). In earlier parts of this series we have used symmetry-adapted deformation coordinates, principal component analysis, dissimilarity calculations and cluster analysis to map and classify the observed conformations of cycloheptane derivatives (Allen, Howard & Pitchford, 1993), *exo*-unsaturated and heterocyclic seven-membered rings (Allen, Howard, Pitchford & Vinter, 1994) and *endo*-unsaturated seven-membered rings (Allen, Garner, Howard & Pitchford, 1994). In each case the distribution of experimentally observed conformations was correlated with either the original or published energy calculations. In this final paper of the series we apply these same techniques to the conformational analysis of cyclooctane and related unsaturated or heterocyclic eight-membered rings. The Cambridge Structural Database (CSD: Allen *et al.*, 1991) is used as the source of experimental crystallographic results.

2. Conformational space and energies for cyclooctane

A detailed discussion of conformational space for eight-membered rings was presented by Evans & Boeyens (1988) and only a brief summary need be given here. The conformational hypersurface for n = eight-membered rings is $n - 3$ = five-dimensional and may be described by the five symmetry-adapted puckering coordinates of Cremer & Pople (1975). These coordinates quantify the symmetrical deformations from a planar D_{8h} system as two amplitude–phase pairs (q_2,φ_2) and (q_3,φ_3), and a unique amplitude (q_4). These coordinates define two, q,φ-pseudorotational pathways and a single inversional subspace, corresponding here to the crown form of D_{4d} symmetry. Evans & Boeyens

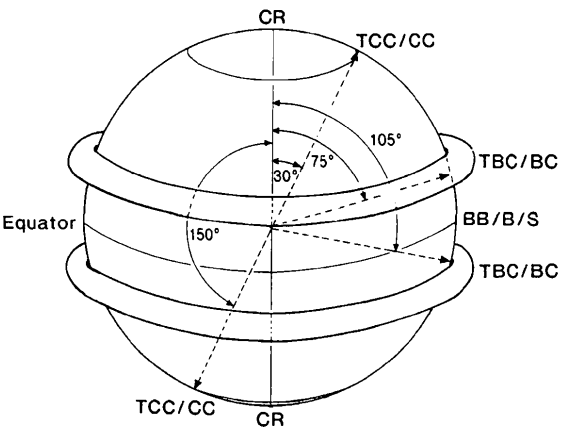

Fig. 1. Three-dimensional representation of the five-dimensional conformational space for cyclooctane (angles shown are of θ; see also Table 1).

Table 1. *The ten canonical conformations of cyclooctane and their characteristic torsion angles (τ_1–τ_8° in: Hendrickson, 1967)*

Puckering amplitudes (q_2–q_4 in Å), polar angles (θ, θ' in °) and index notation (i) are from Evans & Boyens (1988) and relative energies (ΔE in kJ mol^{-1}) from force-field calculations by Anet & Krane (1973). Full conformational descriptors are in Fig. 2.

Conformation	CR	BB	S	B	C	TC	TCC	CC	BC	TBC
τ_1	87.5	52.5	64.9	−73.5	119.9	37.3	56.2	66.0	65.0	88.0
τ_2	−87.5	52.5	37.6	0.0	−76.2	−109.3	−82.4	−105.2	44.7	−93.2
τ_3	87.5	−52.5	−64.9	73.5	0.0	109.3	114.6	105.2	−102.2	51.9
τ_4	−87.5	−52.5	−37.6	0.0	76.2	−37.3	−82.4	−66.0	65.0	44.8
τ_5	87.5	52.5	64.9	−73.5	−119.9	−37.3	56.2	66.0	−65.0	−115.6
τ_6	−87.5	52.5	37.6	0.0	76.2	109.3	−82.4	−105.5	102.2	44.8
τ_7	87.5	−52.5	−64.9	73.5	0.0	−109.3	114.6	105.2	44.7	51.9
τ_8	−87.5	−52.5	−37.6	0.0	−76.2	37.3	−82.4	−66.0	−65.0	−93.2
Symmetry	D_{4d}	D_{2d}	S_4	D_{2d}	C_{2h}	C_{2h}	D_2	C_{2v}	C_s	C_2
q_2	0.0	1.54	1.63	1.64	0.0	0.0	0.46	0.46	1.07	0.67
q_3	0.0	0.0	0.0	0.0	0.96	0.96	0.0	0.0	0.61	0.82
q_4	0.87	0.0	0.0	0.0	0.0	0.0	0.80	0.80	0.33	0.28
θ	0.0	90.0	90.0	90.0	90.0	90.0	30.0	30.0	75.0	75.0
θ'	180.0	90.0	90.0	90.0	90.0	90.0	150.0	150.0	105.0	105.0
i	l	h	h	h	k	k	hl	hl	hk	hk
ΔE	6.3	11.7	11.7	46.9	31.4	32.2	3.3	7.5	0.0	7.1

(1988) show that a three-dimensional representation of the five-dimensional hypersurface is best given by a set of tori on the surface of a sphere, as shown in Fig. 1. They identify ten canonical symmetric forms of cyclooctane and present idealized Cartesian coordinates from which the reference descriptors (torsion angles and puckering amplitudes) of Table 1 and the graphical representations of Fig. 2 are derived.

The radius of the sphere, Q (Fig. 1), is given by*

$$Q = (q_2^2 + q_3^2 + q_4^2)^{1/2} \quad (1)$$

and the spherical polar angle θ in the range $0 \leq \theta \leq \pi$ is

$$\theta = \cos^{-1}(q_4/Q). \quad (2)$$

The unique D_{4d}-crown form (Table 1) occurs at $\theta = 0°$ with its inverse at $\theta = 180°$, a situation which is analogous to that of the unique D_{3d}-chair form of cyclohexane (Picket & Strauss, 1970; Cremer & Pople, 1975). The tori describing canonical forms occur on the sphere at specific θ values and have major and minor radii q_2 and q_3, respectively. The positions of individual conformers and of their permutational and inversional variants occur on these tori at specific values of φ_2 and φ_3. Thus (Fig. 1, Table 1), the TCC/CC and BB/B/S pseudo-rotational pathways occur at $\theta \simeq 30$ and $\theta \simeq 90°$, respectively, with inverse conformers having $\theta \simeq 150$ or 90°. Since $q_3 = 0$ for all these conformers, the two tori contract to latitudinal circles, with permutational isomers defined by specific values of φ_2. The C and TC forms have $q_2 = 0$ and their pseudo-rotational pathway, therefore, follows a meridian, with symmetry variants at specific values of φ_3. Finally, the energetically preferred BC conformers have q_2 and $q_3 \neq 0$, so that its permutational variants lie on helical tracks on a torus at $\theta \simeq 75°$ (inverse conformers have $\theta \simeq 105°$). The TBC conformers also lie on these pseudo-rotational helical tracks and at φ_2, φ_3 positions that are midway between the BC variants. This situation is directly analogous to the description of the C/TC pseudo-rotation itinerary of seven-membered rings (see Allen, Howard & Pitchford, 1993, and references therein).

Evans & Boyens (1988) define a nomenclature for eight-membered ring conformers that specifies the angular (θ,φ_2,φ_3) positions of canonical forms and their symmetry variants. The nomenclature involves three indices h, k, l *via* angular definitions

* The equations of Evans & Boyens (1988) imply that Q^2 is given by summation of q_m^2 for $m = 2, 3 \ldots (n/2 - 1)$ for an n-atom ring, *i.e.* $Q^2 = q_2^2 + q_3^2$ here. Their equation should indicate a summation of q_i^2 for $i = 2$ to $n/2$, yielding Q^2 as indicated in (1).

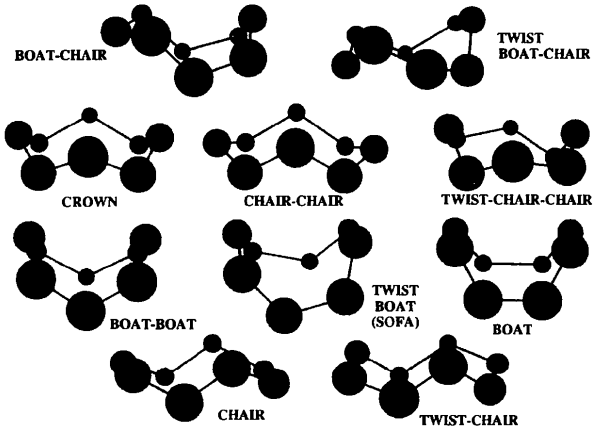

Fig. 2. The ten canonical forms of cyclooctane.

of the form $\varphi_2 = h\pi/16$, $\varphi_3 = k\pi/16$, $\theta = l\pi/16$, where $-16 \le h, k \le 16$ $(h, k = 32 - h, k)$ and $0 \le l \le 16$. The canonical forms are distinguished by only one or two integers (Table 1) and inverse conformers occur at $\varphi_2' = \varphi_2 + \pi = \varphi_2(h + 16)$; $\varphi_3' = \varphi_3(k + 16)$, $\theta' = \pi - \theta = \theta(16 - l)$. Fig. 3 shows the φ_2, φ_3 mapping of the helical BC/TBC pseudorotational pathway showing the 16 permutations/inversions of the BC form and their corresponding h, k indices. TBC conformers lie midway between the BC variants and Fig. 3 here compares directly with Fig. 2 of Allen, Howard & Pitchford (1993) for seven-membered rings.

Energy-minimization calculations for ten conformers (Table 1) of eight-membered rings were first performed by Hendrickson (1967). He found that the BC form was preferred and that the TBC form was some 8.4 kJ mol^{-1} higher in energy. In Table 1 we quote energy differences computed using later force-field methods by Anet & Krane (1973), where the TBC–BC difference is now 7.1 kJ mol^{-1}, with the TCC, crown and the CC forms all within 7.5 kJ mol^{-1} of the preferred BC form.

3. Methodology

3.1. *Data retrieval*

The April 1994 release of the CSD (version 5.07) was used for substructure searches and data analyses using the programs *QUEST3D*, *GSTAT* and *VISTA* (Cambridge Structural Database, 1994, 1995). Chemical search fragments are depicted in Fig. 4 and fall into three main groups: (i) parent cyclooctane (dataset 8C1), (ii) carbocycles with endocyclic or exocyclic double bonds, including fusion with aromatic ring systems (datasets 8C2–8C4), and (iii) heterocycles (datasets 8A1–8A3). All primary substructure searches were qualified by a number of secondary search criteria, applied using the CSD bit-screen mechanisms; (*a*) error-free atomic coordinates were available, (*b*) the structure contained no reported disorder, (*c*) the structure was 'organic' according to CSD definitions, (*d*) the crystallographic $R \le 0.12$ and (*e*) the ring was not bridged, nor formed part of a highly complex ring assembly for which automated ring analysis (Allen *et al.*, 1991) was incomplete (<3% of all CSD entries). Within these criteria, the numbers of CSD entries (N_e) and the numbers of independent ring fragments (N_f) for each dataset are indicated in Fig. 4. The CSD reference codes for structures used in this study are given in Table 2. No other statistically useful subsets of eight-membered rings could be located in the CSD.

3.2. *Data analysis*

Data analysis for cyclooctane fragments (dataset 8C1) was based on the five Cremer–Pople (Cremer &

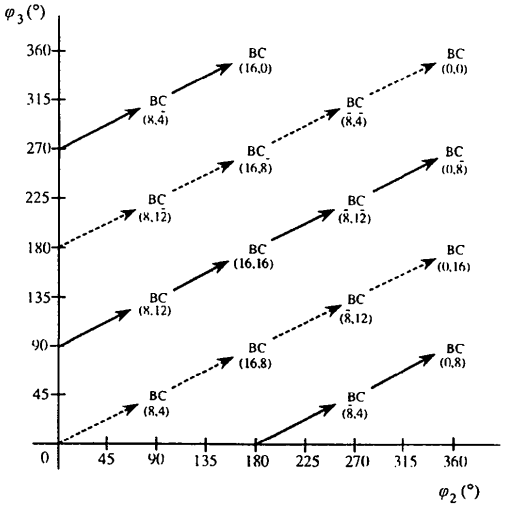

Fig. 3. Orthogonal φ_2, φ_3-mapping of the helical BC/TBC pseudorotation pathways showing the ideal positions of the 16 permutations/inversion of the BC forms (described by their respective h, k values). Conformers on the solid lines are on the torus at $\theta = 75°$, those on the dotted lines are on the torus at $\theta = 105°$.

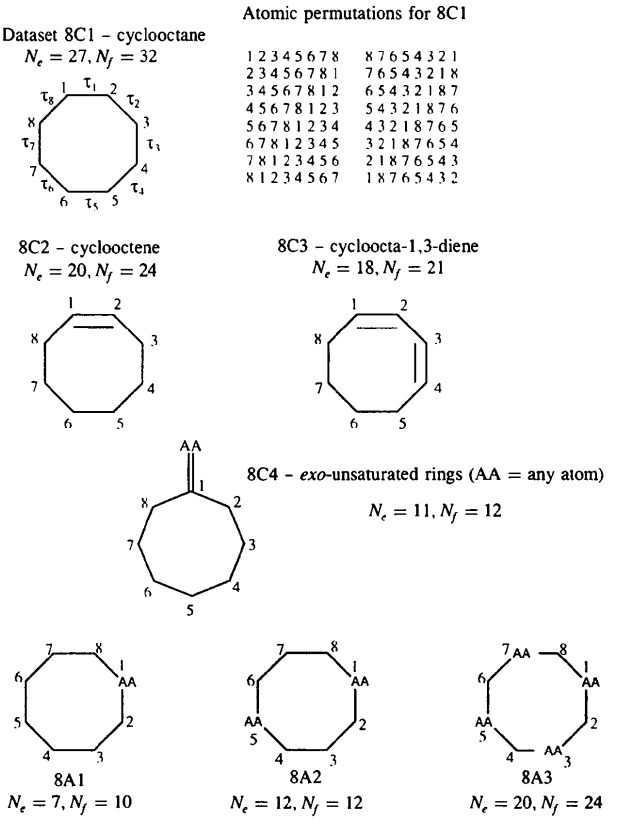

Fig. 4. Chemical rings for which conformational analysis has been performed. N_e is the number of CSD entries retrieved and N_f is the number of occurences of each ring that are included in the analyses.

Pople, 1975) parameters (hereafter referred to as CP parameters), *viz.* q_2, φ_2, q_3, φ_3 and q_4, the polar angle θ calculated using equation (1) and the eight intra-annular torsion angles τ_1–τ_8, calculated in cyclic order with τ_1 referring to the torsion angle about $C_1—C_2$. The torsion angles formed the basis for principal component (PC) analysis (see, *e.g.* Taylor & Allen, 1994) and for symmetry-modified cluster analysis using the Jarvis–Patrick (Jarvis & Patrick, 1973) algorithm, whose application to conformational analysis from crystallographic data is fully described elsewhere (Allen, Doyle & Taylor, 1991; Allen, Howard & Pitchford, 1993). For datasets 8C2–8A3 we employed only the cluster analysis routines, together with visual inspection of these small datasets.

In all cases, however, it was necessary to expand the original data for each individual fragment so that data for all possible atomic permutations and their coordinate inversions were also included in the analyses. The substructure search procedure in *QUEST3D* will place each fragment in some arbitrary asymmetric unit of conformational space and, here, we must apply the D_{8h} topological symmetry of the planar search fragment 8C1, together with the inversion operation, to fill the 32 asymmetric units of the full conformational space. Topological symmetry for fragments 8C2–8A3 is lower, but must be treated similarly. The rationale for generating both atomic and torsional permutational operators for (cyclically enumerated) ring systems is described by Allen, Doyle & Auf der Heyde (1991) and Taylor & Allen (1994). Thus, the 32-fold expansion of the 32 rings of dataset 8C1 yields 1024 symmetry equivalents for data analysis.

4. Conformational analysis of cyclooctanes (8C1)

4.1. *Polar and phase angle plots*

The histogram (Fig. 5*a*) of the polar angle θ (equation 1) shows that conformations on the BC/TBC pseudo-rotation itinerary with $\theta \simeq 75$, 105° (Table 1) dominate the expanded dataset, with 26 of the 32 examples (81%) in this category. Four of the remaining fragments have $\theta \simeq 45°$, presumably distorted variants of CC/TCC conformers (Table 1, Fig. 1), while only two examples exhibit the crown conformation at $\theta \simeq 0$, 180°. The scatterplot of the CP phase angles φ_2, φ_3 (Fig. 5*b*) is directly comparable to the Fig. 3 mapping and confirms that the energetically preferred BC conformer is dominant on the BC/TBC pathway with, perhaps, some evidence for deformations in the TBC direction. Pure TBC conformers, which should occur on the helical pathway (which appears as the diagonal tracks in the

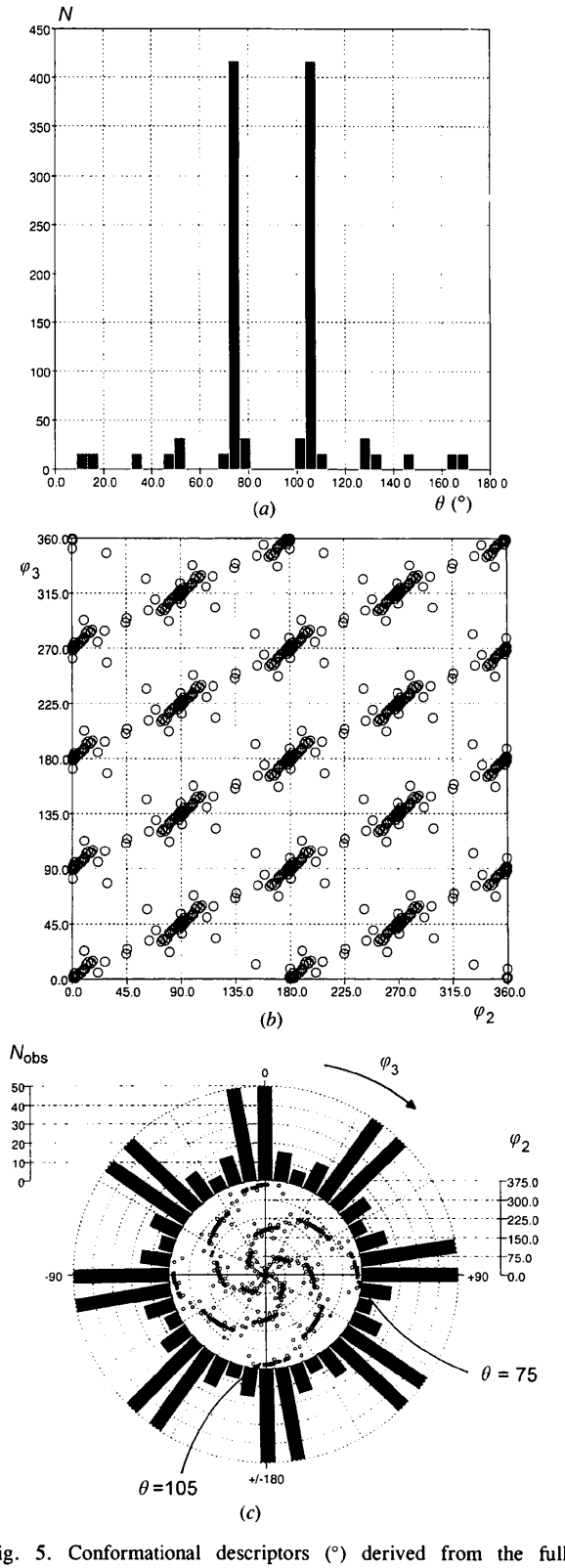

Fig. 5. Conformational descriptors (°) derived from the fully symmetrized dataset of crystallographic observations of cyclooctane (8C1): (*a*) histogram of θ from equation (2), (*b*) orthogonal φ_2,φ_3-mapping and (*c*) polar φ_2, φ_3-scatterplot.

Table 2. *CSD reference codes for the datasets of eight-membered rings of Fig. 4*

Dataset 8C1			
AMCOCA	BAGPII	BCOCTB	CLCOCT
COCOAC	COCOXA10	COVLUU	CURBIA
CUVZEY	CYOCDL	DEZPUT	ECOTDA
EOCNON10	GATRAU	GIVBAO	HOXTHD
HUMULB10	KESVIN	MSOCYO10	OCSHYD
PCDODO	SATKIH	SEJFIW	SPOCTC10
SPTZBN	VALGOE	VASWOB	
Dataset 8C2			
BMEOPH10	BUNPRA	BUTGOM	CACKAS
CBCOEK	CEPHAL10	COCNBZ01	COCNBZ10
CRPLST10	CTSCOC	DEMRES	DINSIC10
DIWZOY	ECOTEC	FEKRUI	FEMGEJ
FETKEU	IPRINC	KAKHEJ	TCHPDO
Dataset 8C3			
BRISTG	BUBXUR	CETZUW	CICHOL
CICHOL01	CIFNAG	CILSOF	COLBOU
COLBUA	DENWOI	FETRAX	GEDVAM
HISPIC	ISTEGO	JECKIL	JECKOR
KAKHAF	KEHTOG		
Dataset 8C4			
AOCINB	BOVMUU	BURJED	CAWLAN
CYCOXA	CYCOXB	DEJJUX	DEMRIW
POCARA	POCARD	POITDL	
Dataset 8A1			
FINZAD	HEPAMC	KETVUA	KEVBES
KEVBIW	LAURCN10	VEZJUF	
Dataset 8A2			
ACTHCO10	BODVAR10	BPHZOC10	CARYAV
CARYEZ	CATJIQ	CATJIQ01	FAVMOE
MSAZOP01	MSAZOP10	THETAC	TSHOCZ
Dataset 8A3			
ANTZCO	BALRIP	BEJVIV	BUKDAM
COWBEV	DEDBUJ	KAKLIR	METALD
MOAOSP10	MTCHLR	OCHTET	OCHTET01
OCHTET03	OCHTET04	OCHTET12	OMSIOC
TACTAZ10	TETIOC10	TOXOCN	TSELOC

orthogonal φ_2, φ_3 plot of Fig. 5*b*) between the BC peaks, are almost completely absent.

The polar scatterplot of Fig. 5(*c*) provides a more appropriate visualization of the helical pathways in φ_2,φ_3-space. The helix on the torus at $\theta = 75°$ (solid lines in the orthogonal plot of Fig. 3) begins at $\varphi_2 = 360$, $\varphi_3 = 90°$ in Fig. 5(*c*), passes through the origin and terminates at $\varphi_2 = 360$, $\varphi_3 = -90\,(270)°$. The $\theta = 105°$ helix (dotted lines in Fig. 3) begins at $\varphi_2 = 360$, $\varphi_3 = 0°$ and proceeds through the origin to $\varphi_2 = 360$, $\varphi_3 = 180°$. The eight peaks in the circular distribution occur around the eight φ_3 values (see Fig. 5*b*) that characterize the BC conformers, *viz.* 0, 45, 90, 135, 180, $-135\,(225)$, $-90\,(270)$ and $-45\,(315)°$. Each peak subsumes two symmetry variants of the BC conformation, one from each of the $\theta = 75$, 105° tori. The slight asymmetry of the peaks with respect to their ideal positions reflects the slight twist distortions of the BC form that occur in the available crystallographic observations.

4.2. *Cremer–Pople (CP) coordinate plots*

The five CP puckering parameters q_2,φ_2; q_3,φ_3; q_4 can be re-expressed in Cartesian form (see Fig. 1): $CP1 = q_2 \cos\varphi_2$, $CP2 = q_2 \sin\varphi_2$ (mapping the BB/B/S equatorial conformers at $\theta = 90°$), $CP3 = q_3 \cos\varphi_3$, $CP4 = q_3 \sin\varphi_3$ (mapping the meridional CC/TCC conformers) and $CP5 = q_4$ (mapping the polar crown conformers at $\theta = 0,180°$). These CP axes are mutually orthogonal in the five-dimensional conformational space and the ten possible two-dimensional scatterplots for the expanded 8C1 dataset are shown in Fig. 6.

A number of these plots can be immediately correlated with different two-dimensional projections of the five-dimensional hypersurface representation of Fig. 1. The plots of CP5 *versus* CP1 (or CP2), shown in Fig. 6(*c*), are orthogonal views parallel to the $\theta = 90°$ equator, which is itself unrepresented due to the lack of BB/B/S conformers. Above and below the equator are the densely populated BC/TBC tori at $\theta \simeq 75,105°$. Symmetry variants of the crown conformers appear at the poles, with symmetry variants of the CC/TCC forms spread out along the obvious meridional lines that define the projection of the sphere. The CP5 *versus* CP3 (or CP4) maps (Fig. 6*f*) are similarly interpreted, except that the meridional conformers now appear as groups linking the poles with the BC/TBC tori. The CP3 *versus* CP4 mapping (Fig. 6*e*) represents a polar projection of Fig. 1, with the inner (lower θ) CC/TCC pseudo-rotation circles and the outer (higher θ) BC/TBC tori clearly defined.

Other plots in Fig. 6 are, perhaps, less intuitive. The CP1 *versus* CP2 plot views the BC/TBC tori and the CC/TCC circles along a polar axis. Fig. 3 shows that the dominant BC conformers exist at only four specific values of φ_2, *i.e.* at 0, 90, 180, 270°, hence the outer high density in the CP1,CP2 plot represents a superimposition of these φ_2 symmetry variants of BC conformers. The origin of the plot reflects crown conformer density, while density linking this origin to the BC peaks is from the CC/TCC variants. The CP1 *versus* CP3 (or CP4) plots (Fig. 6*b*) represent orthogonal projections (side-on views) of the BC/TBC helical pathway on the $\theta = 75,105°$ tori. The CP2 *versus* CP3 (or CP4) plots of Fig. 6(*d*) are similar to Fig. 6(*b*), but are rotated by 90°.

4.3. *Symmetry-modified principal component analysis*

The PC analysis was performed on τ_1–τ_8 for all 1024 fragments of the symmetry-expanded 8C1 dataset. The percentage variance accounted for by each PC, together

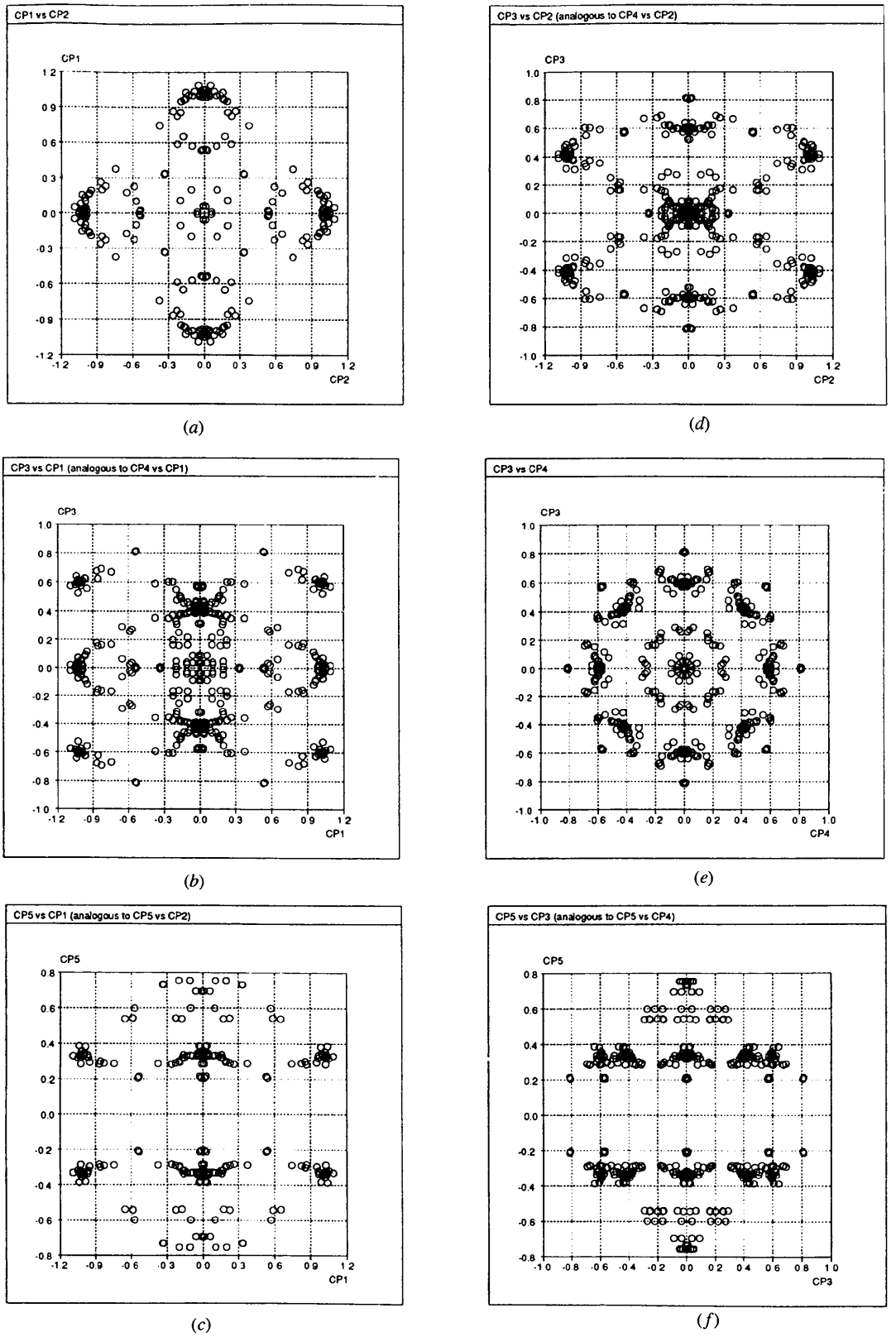

Fig. 6. Scatterplots of Cremer–Pople (Cremer & Pople, 1975) puckering parameters for the fully symmetrized dataset 8C1 (see text for discussion of individual plots).

Table 3. *Principal component analysis of the symmetry-expanded torsional dataset 8C1*

Var is the percentage of total variance accounted for by each PC, $L(\tau_n)$ is the PC loading for each torsion angle, τ_n (°).

	PC1	PC2	PC3	PC4	PC5
Var	37.27	21.45	21.45	9.89	9.89
$L(\tau_1)$	44.87	27.61	39.44	−31.46	−8.87
$L(\tau_2)$	−44.87	8.36	−47.41	8.87	−31.46
$L(\tau_3)$	44.87	−39.44	27.61	31.46	8.87
$L(\tau_4)$	−44.87	47.41	8.36	−8.87	31.46
$L(\tau_5)$	4.87	−27.61	−39.44	−31.46	−8.87
$L(\tau_6)$	−44.87	−8.36	47.41	8.87	−31.46
$L(\tau_7)$	44.87	39.44	−27.61	31.46	8.87
$L(\tau_8)$	−44.87	−47.40	−8.36	−8.87	31.46

with their loadings, are given in Table 3. As expected, five PC's account for ~100% of torsional variance in the dataset. They occur as a unique PC1, together with two degenerate pairs PC2,PC3, and PC4,PC5. The unique coordinate has the C_4 symmetry of the crown form, while the PC2,PC3 pair map chair-like (*m* symmetry) distortions and PC4,PC5 map boat-like (C_2 symmetry) distortions. As in earlier examples (Allen, Doyle & Auf der Heyde, 1991; Allen, Howard & Pitchford, 1993), the scatterplots based on PC scores are clearly identical to those based on CP puckering coordinates, except for mutual axial rotations.

4.4. *Symmetry-modified cluster analysis*

The symmetry-modified Jarvis–Patrick (JP) clustering algorithm (Allen, Doyle & Taylor, 1991; Allen, Howard & Pitchford, 1993) uses the torsion angles τ_1–τ_8 as a basis. Cluster formation is governed by four user-defined parameters: (*a*) the power factor, n, used in the calculation of torsional dissimilarities, D^n, where $n = 1$ represents the city-block metric, $n = 2$ represents the Euclidean metric, (*b*) K_{NN}: the number of nearest neighbours, based on D^n values, for each observed conformation, i, that are to be considered in the JP clustering procedure, (*c*) K_{JP}: the minimum number of nearest neighbours that must be common to the nearest neighbour tables of observations i_1, i_2 in order for this pair of observations to be assigned to the same cluster ($K_{JP} \leq K_{NN}$) and (*d*) D_{max}: a maximum dissimilarity (range $1 \rightarrow 0$) above which two observations are not considered as nearest neighbours, even though K_{NN} may not be satisfied.

Our earlier work had consistently indicated that values of $n = 2$, $K_{NN} = 7$–10, $K_{JP} \sim K_{NN}/2$ and $D_{max} = 0.05$–0.10 gave clusters that were numerically coherent and chemically meaningful. For the cyclooctane dataset (8C1), the optimum clustering structure was judged to be that with $n = 2$, $K_{NN} = 8$, $K_{JP} = 5$ and $D_{max} = 0.10$. This resulted in 22 of the 32 conformations being assigned to three clusters with populations, N_p, of 12, 6 and 4, respectively. Mean torsion angles for these clusters are given in Table 4(*a*). The largest cluster ($N_p = 12$) corresponds to pure BC conformers and the centroid of this cluster is close to a special position in conformational space such that cluster coalescence occurs (see Allen & Taylor, 1991). Cluster 2 ($N_p = 6$) is a BC/TBC intermediate that was recognized and described by Anet & Krane (1973) and which shows flattening of the angles τ_1 and τ_2. Further flattening in this area leads to cluster 3 ($N_p = 4$), another BC/TBC intermediate that has $\tau_2 \simeq 0°$. The remaining ten conformers in dataset 8C1 occur as: (*a*) one doublet ($N_p = 2$) corresponding to the two crown conformers identified in Fig. 5(*a*), (*b*) the singleton TCC conformer of Fig. 5(*a*) and (*c*) seven distorted BC/TCC conformers that differ significantly from one another and from clusters 1, 2 and 3 of Table 4(*a*). Torsion angles for the crown and TCC conformers are also exemplified in Table 4(*a*) for an individual CSD refcode. For clusters 1, 2 and 3, the refcode cited is for the experimental observation that is closest to the cluster centroid. As with other systems, the cluster analysis has been able to identify the major conformational subgroups present in dataset 8C1 in a rapid and effective manner, and from a torsional dataset that contains examples drawn from many different asymmetric units of conformational space.

4.5. *Chemical environments of preferred conformers*

All 12 members of the BC cluster 1 (Table 4*a*) are free (non-fused) cyclooctane rings and, in view of the energies cited in Table 1, the BC preference is to be expected. Cluster 2 contains six higher energy BC/TBC conformers and, in five of these cases, the cyclooctane ring is fused to four- or five-membered ring. This constraint appears sufficient to move the eight-ring conformation away from its preferred (isolated) BC form along the BC/TBC pseudo-rotation pathway. This process is continued in cluster 3, where fusion to three-membered rings or a number of quaternary C-ring atoms leads to further distortions of the preferred BC form of the eight-membered ring. Three-membered ring fusion also occurs in the crown and TCC examples in dataset 8C1.

5. Unsaturated eight-membered rings (8C2–8C4)

5.1. *Cyclooctene (8C2)*

Cluster analysis results for the 24 conformations in dataset 8C2 are given in Table 4(*b*). The *cis* and *trans* conformers have been well studied, both experimentally and computationally (see, *e.g.* Allinger & Sprague, 1972; Ermer & Lifson, 1973; Traetteberg, 1975; Favini, Buemi & Raimondi, 1968). Here, cluster 1 ($N_p = 9$) corresponds to *cis* conformers and cluster 2 ($N_p = 6$) corresponds to *trans* conformers, which are

Table 4. *Symmetry-modified Jarvis–Patrick (JP) clustering results for the eight-membered ring datasets of Fig. 4*

The JP control parameters K_{NN}, K_{JP}, D_{max} and n are described in the text, N_f is the number of crystallographic observations and N_p is the population of cluster number N_c. Where appropriate the CSD reference codes cited represent those rings which are closest to the centroid of the cluster described by the mean torsion angles (°) with e.s.d.'s in parentheses.* Other refcodes refer to individual rings for which τ_1–τ_8 are cited without e.s.d.'s to indicate typical values. Conformational identifiers are from Table 1 and Fig. 2.

N_c	N_p	Refcode	Conformational identifiers	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	τ_7	τ_8
(*a*) 8C1: $N_f = 32, K_{NN} = 8, K_{JP} = 5, D_{max} = 0.10, n = 2$											
1	12	BAGPII	BC	−100.9 (10)	43.2 (11)	65.0 (8)	−65.0 (8)	−43.2 (11)	100.9 (10)	−66.8 (10)	66.8 (10)
2	6	COVLUU	BC/TBC	−91.0 (9)	24.3 (18)	77.2 (10)	−57.3 (10)	−53.4 (7)	103.1 (13)	−65.3 (12)	68.9 (8)
3	4	SPTZBN	BC/TBC	−79.5 (4)	0.5 (21)	89.6 (24)	−52.4 (7)	−56.4 (13)	102.6 (32)	−71.2 (15)	76.3 (36)
		DEZPUT	Crown	70.4	−83.2	92.3	−73.3	63.8	−82.3	96.5	−82.0
		EOCNON10	TCC	47.7	−84.7	134.4	−85.3	48.7	−82.4	124.9	−80.7
(*b*) 8C2: $N_f = 24, K_{NN} = 8, K_{JP} = 4, D_{max} = 0.08, n = 2$											
1	9	IPRINC	BC/TBC	−3.5 (9)	81.6 (6)	−73.0 (11)	71.1 (8)	−100.0 (24)	50.4 (16)	53.7 (17)	−90.7 (10)
2	6	COCNBZ01	TCC	−139.7 (15)	88.5 (19)	−50.9 (18)	80.1 (20)	−114.0 (20)	80.1 (20)	−50.9 (18)	88.5 (19)
3	2	CEPHAL10	BC/TBC	5.4 (40)	81.9 (24)	−82.3 (96)	72.4 (46)	−99.8 (74)	71.4 (41)	28.4 (9)	−81.6 (39)
(*c*) 8C3: $N_f = 21, K_{NN} = 8, K_{JP} = 4, D_{max} = 0.10, n = 2$											
1	16	JECKIL	C_2-TBC	0.8 (15)	−65.5 (17)	0.8 (15)	90.5 (17)	−78.2 (33)	46.9 (54)	−78.2 (33)	90.5 (17)
(*d*) 8C4: $N_f = 12, K_{NN} = 6, K_{JP} = 3, D_{max} = 0.10, n = 1$											
1	8	CYCOXA	BC	98.9 (30)	−63.7 (14)	65.7 (11)	−104.6 (18)	48.3 (30)	60.7 (21)	−68.3 (36)	−39.7 (56)
(*e*) 8A1: $N_f = 10, K_{NN} = 5, K_{JP} = 3, D_{max} = 0.10, n = 1$											
1	5	KETVUA	CC	100.0 (15)	−56.0 (68)	78.8 (25)	−108.2 (47)	83.7 (15)	−60.5 (46)	74.4 (17)	−125.9 (55)
(*f*) 8A2: $N_f = 12, K_{NN} = 4, K_{JP} = 2, D_{max} = 0.12, n = 2$											
1	6	CARYAU	BC	103.8 (16)	−69.0 (6)	65.9 (7)	−110.6 (7)	54.7 (7)	61.9 (17)	−63.7 (10)	−45.7 (14)
(*g*) 8A1: $N_f = 10, K_{NN} = 5, K_{JP} = 3, D_{max} = 0.10, n = 1$											
1	12	COWBEV	Crown	97.1 (22)	−76.6 (30)	76.6 (30)	−97.1 (22)	97.1 (22)	−76.6 (30)	76.6 (30)	−97.1 (22)
2	5	ANTZCO	TC	103.8 (11)	−21.1 (24)	−42.5 (11)	116.6 (8)	−103.8 (11)	21.1 (24)	42.5 (11)	−116.6 (8)

* The e.s.d.'s of the mean torsion angles (σ_m) are given, from which $\sigma_{sample} = N_p$.

easily distinguished by τ_1 (C—C=C—C) $\simeq 0$ (*cis*) or 140° (*trans*).

Cyclooctene is the smallest cycloalkene for which a stable *trans* form is possible. The strained ring is energetically capable of adopting the CC or TCC conformation and θ values for the crystallographic examples lie in the expected 30–40° range. The C_2-TCC form has been calculated as the minimum energy form (Traetteberg, 1975), some 13.1 kJ mol^{-1} below the CC form. All six examples of dataset 8C2 show the expected C_2-TCC symmetry, with one of the torsion angles (τ_3) of the cyclooctane archetype of Table 1 showing greatly increased puckering to permit the *trans* arrangement about the C1=C2 double bond (τ_1 in dataset 8C2). The θ values for the *cis* conformers lie in the range 70–80°, which characterizes the BC/TBC pathway for the parent cyclooctane. The nine examples of cluster 1 for dataset 8C2 all exhibit an asymmetric BC/TBC conformational variant that is closely similar to the minimum-energy form obtained by force-field calculations (see, *e.g.* Favini, Buemi & Raimondi, 1968; Ermer & Lifson, 1973). Indeed, this BC/TBC conformer is also closely similar to that of cyclooctane (8C1) cluster 3 where, as already noted, the deformation away from the preferred BC form is due to fusion to three-membered rings – a common conformational mimic for an ethylenic linkage. Other cyclooctene conformations (nine examples in dataset 8C2) are distorted variants of the *cis* conformation induced, primarily, by fusion to other small rings. The doublet of cluster 3 (8C2) is typical and has additional flattening in the τ_7 area by comparison with the true *cis* conformers of cluster 1.

5.2. *Cycloocta-1,3-diene (8C3)*

Cluster analysis (Table 4*c*) has clearly isolated and symmetrized (Allen & Taylor, 1991) a C_2-symmetric conformer that is adopted by 16 of the 21 observations of dataset 8C3. This conformer is denoted as TBC by Anet & Yavari (1977), who calculate its energy to be only 2.3 kJ mol^{-1} below that of an asymmetric C_1-TB conformer, in close agreement with the 1.8 kJ mol^{-1} energy difference calculated by Allinger, Viskocil,

Burkert & Yuh, (1976). Both sets of authors agree that these are the only two low-energy conformers of cycloocta-1,3-diene.

The 1,3-diene system in the C_2-TBC conformation is non-planar with τ_2(C1—C2—C3—C4) = 65.5 (1.7) in Table 4(*c*) by comparison with τ_2 values of 54 and 55.7° calculated by Anet & Yavari (1977) and by Allinger, Viskocil, Burkert & Yuh (1976), respectively. However, we note that in all except one of the observations that comprise cluster 1 for dataset 8C3 (Table 4*c*), both 'double bonds' arise from dibenzo fusion that, almost certainly, leads to enhanced puckering in the C1–C4 area of the eight-membered ring. This supposition is enhanced by the crystallographic data for the remaining five conformations of dataset 8C3. All of these rings adopt distorted forms of the C_1-TB conformation of Anet & Yavari (1977). However, by comparison with their calculated τ_2 of 42°, four of the crystallographic examples (BUBXUR, COLBOU and two independent rings in DENWOI) have τ_2 in the range 64–75° and are dibenzo-cyclooctadienes. In only one example (KAKHAF) τ_2 is 46° and here the '1,3-diene' comprises a single benzo-fusion bond and a simple ethylenic linkage. It is the C_1-TB conformation that was identified by Traetteberg (1970) in an electron diffraction study of the parent cycloocta-1,3-diene. Obviously, crystal structures of simple derivatives of this parent ring will significantly enhance the crystallographic knowledge base in this area.

5.3. *Exo-unsaturated cyclooctane rings (8C4)*

There are 12 examples of cyclooctane rings that contain a single Csp^2 atom that forms an exocyclic double bond to C, N or O substituents. Eleven of these rings adopt the minimum-energy BC conformation of Table 1 and eight of these form the coherent cluster for which mean torsion angles are presented in Table 4(*d*). The only true outlier (POCARA) has the eight-membered ring fused to both three- and four-membered rings and exhibits a TCC conformation.

6. Heterocyclic eight-membered rings (8A1–8A3)

6.1. *Monohetero-cyclooctanes (8A1)*

Seven CSD entries yield ten monohetero-cyclooctanes, nine of which have hetero-O and one has hetero-N. Force-field calculations (Burkert, 1980) indicate that the BC form is preferred as in cyclooctane. By contrast, the crystallographic data yield a coherent cluster of five fragments (Table 4*e*) which adopt a slightly distorted CC conformation. However, all these rings form part of the extended fused systems. The remaining five rings all fall on the BC/TBC pseudo-rotation itinerary and are either isolated rings or fused to only one other ring. The free rings adopt almost perfect BC conformations, as predicted by computational methods.

6.2. *1,5-Dihetero-cyclooctanes (8A2)*

The 12 rings (all isolated) of dataset 8A2 comprise six S,S rings, four N,N rings and two S,N rings. The only coherent cluster (Table 4*f*) contains six BC conformers (four S,S rings, two S,N rings), the preferred form of the parent carbocycle. Three rings (CATJIQ and CATJIQ01, both N,N, and the S,S THETAC) all adopt the unusual TC form (Table 1) and all are disubstituted at the 3,7-positions. Other conformations represented are TCC/CC intermediates (two examples) and a single example of the crown form.

6.3. *1,3,5,7-Tetrahetero-cyclooctanes (8A3)*

The 24 rings of dataset 8A3 are all isolated rings with the following hetero-atom constitutions: tetra-N: 12 examples, tetra-O: 3, tetra-S: 3, tetra-Se: 1, tetra-Si: 1, 1,5-N, 3,7-P: 3 and 1,5-N, 3,7-O:1. Cluster analysis shows a slightly distorted crown form to be dominant in these crystallographic examples (cluster 1, Table 4*g*, 12 observations). Membership of this group comprises six tetra-N rings, all three tetra-O rings and all of the mixed N,P and N,O rings with the exception of a single N,P example. Burkert & Allinger (1982) report that a BC form is again energetically preferred (at least for 1,3,5,7-tetraoxacyclooctane), but present arguments why the D_{4d} crown form is more commonly observed in solution and in crystal structures. The subsidiary cluster 2 of Table 4(*g*) contains five tetra-N rings that all adopt a distorted TC conformation, almost identical to that noted above for three of the 1,5-dihetero systems. The tetra-S and tetra-Se rings do adopt the BC form (as does one of the 1,5-N,3,7-P rings). The tetra-Si ring adopts a BB conformation, one of the lower energy forms identified by force-field methods for the tetra-O heterocycle (Burkert & Allinger, 1982).

7. Conclusions

This short series of papers, including the current report, has highlighted both the successes and the pitfalls of conformational analyses from crystallographic data. The work confirms that multivariate analyses using principal components and clustering techniques are of significant value in obtaining mappings and dissections of torsional datasets. For ring systems, the Cremer & Pople (1975) analysis is invaluable. We note that proper attention must be paid to the topological symmetry of substructures under study, in order that full advantage can be taken of these methods of data analysis. We also note that while *all* cluster analysis results must be scrutinized carefully for chemical sensibility, the current work indicates that extreme care must be taken in assessing these results for very small datasets, *i.e.* those that contain less than 20–30 conformational observations. Evidence presented elsewhere (Conklin, Fortier, Glasgow & Allen, 1996) would indicate that

conceptual clustering may have advantages over the numerical techniques used here.

The major pitfalls of interpretation are related to three interlinked factors: (*a*) the precision with which the substructure of interest is defined, (*b*) the chemical homogeneity of the substructural environment and (*c*) the number of observations retrieved for analysis. Obviously factors (*a*) and (*b*) are linked, although it is always possible to define a substructure with great precision and still retrieve a sample that is heavily biased towards *e.g.* a particular substitution or ring-fusion pattern. If the number of observations (*c*) is low, there is a tendency to relax substructural precision to enlarge the dataset. In this work, for example, we have used both aromatic and ethylenic bonds in definitions of intra-annular unsaturation. These imprecisions are acceptable and often engender a more interesting analysis, but their use must not be forgotten in presenting the observed conformational preferences and in comparing these results with other experimental or computational data. Thus, chemical bias is an important issue and should be carefully assessed, especially for smaller datasets: The analysis in this paper and in earlier parts of the series have shown, particularly, the effects of the fusion of target rings to small (three- and four-membered rings) or aromatic systems.

Despite these caveats, our conformational analyses of medium rings have shown a clear qualitative correlation between conformational preferences exhibited in experimental studies of a condensed phase, the crystal, and energy minima computed from gas-phase force-field parameters. A similar correlation between crystallographic conformational preferences, determined from large and chemically homogeneous datasets of acyclic substructures, and high-level *ab initio* (*in vacuo*) calculations has recently been presented (Allen, Harris & Taylor, 1996). Thus, crystallographic observations provide good indications of conformational preferences for use in molecular modelling and related applications.

Finally, we note that database analyses of this type can often provide indications of chemical structure types for which crystallographic data is either very sparse or not available at all. These considerations prompted our recent structure determinations of a non-fused derivative of cyclohepta-1,3-diene (Allen, Howard, Pitchford & Steel, 1995). Other chemical structures in this category are obvious from the analyses presented in the current paper.

References

Allen, F. H. & Taylor, R. (1991). *Acta Cryst.* B**47**, 404–412.

Allen, F. H., Davies, J. E., Galloy, J. J, Johnson, O., Kennard, O., Macrae, C. F., Mitchell, G. F., Smith, M. J. & Watson, D. G. (1991). *J. Chem. Inf. Comput. Sci.* **31**, 187–204.

Allen, F. H., Doyle, M. J. & Auf Der Heyde, T. P. E. (1991). *Acta Cryst.* B**47**, 412–424.

Allen, F. H., Doyle, M. J. & Taylor, R. (1991). *Acta Cryst.* B**47**, 41–49.

Allen, F. H., Garner, S. E., Howard, J. A. K. & Pitchford, N. A. (1994). *Acta Cryst.* B**50**, 395–404.

Allen, F. H., Harris, S. E. & Taylor, R. (1996). *J. Comput.-Aided Mol. Des.* **10**, 247–254.

Allen, F. H., Howard, J. A. K. & Pitchford, N. A. (1993). *Acta Cryst.* B**49**, 910–928.

Allen, F. H., Howard, J. A. K., Pitchford, N. A. & Steel, P. G. (1995). *Acta Cryst.* C**51**, 1946–1949.

Allen, F. H., Howard, J. A. K., Pitchford, N. A. & Vinter, J. G. (1994). *Acta Cryst.* B**50**, 382–395.

Allinger, N. L. & Sprague, J. T. (1972). *J. Am. Chem. Soc.* **94**, 5734–5742.

Allinger, N. L., Viskocil, J. F. Jr, Burkert, U. & Yuh, Y. (1976). *Tetrahedron*, **32**, 33–38.

Anet, F. A. L. & Krane, J. (1973). *Tetrahedron Lett.* **50**, 5029–5032.

Anet, F. A. L. & Yavari, I. (1977). *J. Am. Chem. Soc.* **99**, 6986–6991.

Burkert, U. (1980). *Z. Naturforsch. Teil B*, **35**, 1479–1485.

Burkert, U. & Allinger, N. L. (1982). *Molecular Mechanics.* ACS Monograph, No. 148. Washington, DC: American Chemical Society.

Cambridge Structural Database (1994). *User's Manual.* Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, England.

Cambridge Structural Database (1995). *User's Manual.* Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, England.

Conklin, D., Fortier, S., Glasgow, J. I. & Allen, F. H. (1996). *Acta Cryst.* B**52**, 535–549.

Cremer, D. & Pople, J. A. (1975). *J. Am. Chem. Soc.* **97**, 1354–1358.

Ermer, O. & Lifson, S. (1973). *J. Am. Chem. Soc.* **95**, 4121–4131.

Evans, D. G. & Boeyens, J. C. A. (1988). *Acta Cryst.* B**44**, 663–671.

Favini, G., Buemi, G. & Raimondi, M. (1968). *J. Mol. Struct.* **2**, 137–152.

Hendrickson, J. B. (1967). *J. Am. Chem. Soc.* **89**, 7047–7054.

Jarvis, R. A. & Patrick, E. A. (1973). *IEEE Trans. Comput.* **22**, 1025–1034.

Pickett, H. M. & Strauss, H. L. (1970). *J. Am Chem. Soc.* **92**, 7281–7288.

Taylor, R. & Allen, F. H. (1994). *Structure Correlation*, edited by H.-B. Buergi & J. D. Dunitz, pp. 111–162. Weinheim, Germany: VCH Publishers.

Traetteberg, M. (1970). *Acta Chem. Scand.* **24**, 2285–2290.

Traetteberg, M. (1975). *Acta Chem. Scand. Sect. B*, **29**, 29–40.