

## SARACEN – molecular structures from theory and experiment: the best of both worlds

Norbert W. Mitzel<sup>a</sup> and David W. H. Rankin<sup>b</sup>

<sup>a</sup> *Institut für Anorganische und Analytische Chemie, Westfälische Wilhelms-Universität Münster, Wilhelm-Klemm-Strasse 8, 48149 Münster, Germany*

<sup>b</sup> *School of Chemistry, University of Edinburgh, West Mains Road, Edinburgh, UK EH9 3JJ. E-mail: d.w.h.rankin@ed.ac.uk*

Received 19th June 2003, Accepted 19th August 2003  
First published as an Advance Article on the web 1st September 2003

Structures of molecules in the gas phase, determined experimentally, provide definitive information about their identity, reactivity and other properties, free from intermolecular interactions. Available methods have not been applicable to large and asymmetric molecules. Now the SARACEN (Structure Analysis Restrained by *Ab initio* Calculations for Electron diffraction) method, using data from computational methods to complement experimental data, has opened the door to full structure determination for all sufficiently volatile molecules.

### Why determine structures?

Structure is so fundamental a part of our chemical thinking that it is easy to forget how important it is. Some 80% of papers in the top inorganic journals report new geometrical information (*i.e.* bond lengths and angles) and about half of the remainder make specific use of such data. In organic journals the proportions are lower, but structural information in the broader sense of connectivity is ubiquitous, often accompanied by two-dimensional representations of three-dimensional structures which conceal almost as much as they reveal.

Why is so much effort still expended in an area where so much is already known? First, for identification. We make a compound, and need to know what it is. A crystal structure determination can now provide a rapid and effective answer and often, particularly for transition metal compounds, it may be the only appropriate method. This limited objective of

identification does not of itself require great precision in the structural measurements, but this is essential if the structural data are to be used to understand the properties of a compound. This can include interpretation of spectroscopic data, rationalising and predicting chemical behaviour and, less directly, explaining physical properties.

However, the whole body of structural data is much more valuable than the sum of the individual structures. At the most basic level, the essence of the way in which we visualise molecules and reactions is founded on the vast accumulation of structural data. Molecular modelling programs, which can quantify our ideas, use parameters derived from these hundreds of thousands of structures, and particularly from those for simple, representative compounds, usually studied in the gas phase. Smaller numbers of structures, for series of related compounds, can provide pictures of reaction pathways. For example, the activation of C–H bonds and the transfer of a hydrogen atom from carbon to a transition metal atom can be represented by the structures of compounds that happen to lie at different points along the path.<sup>1</sup> Each was studied individually as an end in itself; but together they provide an insight that was not available to any of the original researchers.

### Accurate structures of gas-phase molecules

Comparisons of structures of two or more similar compounds usually involve differences between related parameters; for such differences to be significant, accuracy is of paramount

*Norbert W. Mitzel was born in Augsburg, Germany, in 1966. He studied chemistry at the Technische Universität München, Germany, where he graduated in 1991 and received his Dr. rer. nat. degree in 1993 with Professor Hubert Schmidbaur. He was a Marie–Curie fellow of the European Union with Professor David W. H. Rankin at the University of Edinburgh from 1994 to 1996, before working on his habilitation at the Technische Universität München, which he finished in 2000. After one semester at the Technische Universität Berlin he moved to the Westfälische Wilhelms-Universität in Münster, Germany, where he has been Professor of Inorganic Chemistry since 2002. His research interests presently include preparative chemistry of group 13, 14 and 15 elements, weak secondary interactions, structural chemistry in different phases and volatile molecular material precursors. Characteristic of this research has been the interplay of synthetic studies of small molecules with a minimum of organic substitution with instrumental techniques including gas-phase electron diffraction, in-situ crystal growth for X-ray diffraction and multinuclear NMR spectroscopy.*

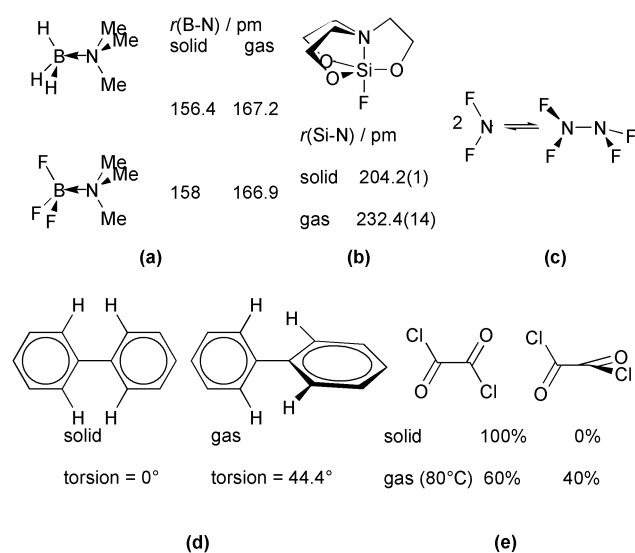


Norbert W. Mitzel and David W. H. Rankin

*David Rankin is Professor of Structural Chemistry at the University of Edinburgh, where he has held a succession of positions, starting as an ICI Research Fellow in 1969. His first degree, in Chemistry (with Rowing) was gained in Cambridge, followed by research under the supervision first of Evelyn Ebsworth and then George Sheldrick. Synthesis of germyl compounds paved the way to spectroscopic and structural studies, and an introduction to gas electron diffraction. After moving to Edinburgh the focus of the synthetic work shifted to fluorophosphine derivatives, but the structural emphasis remained, with the eventual establishment of an electron diffraction laboratory. Other sources of structural information also became important, first experimental, including liquid crystal NMR spectroscopy, and more recently computational methods. He has received the Royal Society of Chemistry Award for Structural Chemistry (1990) and was awarded a Tilden Lectureship (1996), and elected Fellow of the Royal Society of Edinburgh in 1991. Latterly he has also studied relationships between metals in plants and soils, hunting rhododendrons on Chinese limestone mountains.*

importance. Moreover, crystal packing effects can cause sufficient perturbation to mask the underlying structural changes. Even the overused phrase 'no significant inter-molecular contacts' (which really means that there are no inter-molecular distances that are much shorter than the appropriate sums of van der Waals radii) can hide extremely significant structural changes caused by those contacts.

The simple Lewis-acid-base adducts  $\text{H}_3\text{B-NMe}_3$  and  $\text{F}_3\text{BNMe}_3$ , for example, have dative B-N bonds, which can be considered as partial bonds, being driven towards completion if the environment of the molecules is changed from vacuum to a polar crystal.<sup>2</sup> This phase change results in a shrinkage of the B-N bond of about 10 pm in each case [Scheme 1(a)].<sup>3</sup> Similar intramolecular changes of bond lengths were observed in silatranes [e.g. fluorosilatrane, Scheme 1(b)], the Si-N bonds of which are larger in the gas phase and shorter in the solid state.<sup>4</sup> Homolytic bond dissociation and its structural consequences may also be observed in the gas phase. For example,  $\text{N}_2\text{F}_4$  has a long N-N bond in its vapour,<sup>5</sup> where it exists in equilibrium with its dissociation product  $\cdot\text{NF}_2$  [Scheme 1(c)], which was also structurally characterised by GED.<sup>6</sup>



Scheme 1

Drastic conformational differences between molecules in gaseous and solid states can be observed, e.g. for biphenyl<sup>7</sup> [Scheme 1(d)] and for oxalyl chloride [Scheme 1(e)].<sup>8</sup> In the crystal the two phenyl groups in biphenyl are coplanar, with quite close repulsive contact between the four *ortho* hydrogen atoms, whereas in the gas phase the torsional angle between the two phenyl rings is 44.4(12)°, thus avoiding repulsive forces between the hydrogen atoms but losing  $\pi$  conjugation between the phenyl rings. Oxalyl chloride, on the other hand, is present as the planar *anti* conformer in the solid state, but in the gas phase there is a significant amount of a second conformer (~40% at 80 °C). This *gauche* conformer has reduced conjugation of the C-C single bond, and is predicted to be a stable conformation only by a minority of theoretical methods.

It is clear that accurate gas-phase data are vital, and ideally all molecular structures should be studied in both gaseous and crystalline phases, although of course such an ideal is attainable in only a minority of cases. Unfortunately, the techniques available for determining structures of molecules in the gas phase are few, and provide limited information. Electron diffraction is the most widely used method. However, it provides only one-dimensional information about three-dimensional structures, and the difficulty or impossibility of resolving similar distances (represented by overlapping peaks in the radial distribution curve) means that rarely can more than 10–15 independent

geometrical parameters be refined. Scattering of electrons also depends on the atomic numbers of the atoms involved, so location of light atoms in the presence of heavy ones is inaccurate or impossible. The technique, therefore, is applicable mainly to small and/or highly symmetrical molecules.

Rotation constants, obtained from pure rotational or high-resolution vibrational spectra, also provide geometrical information, but no more than three constants can be measured for any one molecule. Isotopic substitution can help but, unless the required isotopomers occur naturally in sufficient abundance, the preparative work and the expense involved can be formidable obstacles. Rotational spectroscopy, while extremely valuable for some types of small molecule, is extremely limited in its range of applicability.

Combining the merits of electron diffraction and rotational spectroscopy, by analysing diffraction data and rotation constants simultaneously, is useful,<sup>9</sup> as the types of information provided by the two methods can be complementary. However, the total range of gas-phase molecules that can be studied is not greatly extended by this type of combined analysis. We have also pioneered the supplementation of gas-phase data by dipolar coupling constants, measured by NMR spectroscopy of solutions in liquid crystal solvents.<sup>10</sup> This method can provide extremely accurate structures, but great care is needed to ensure that the liquid crystal solvent does not distort the molecules from their gas-phase geometries, and even then the joint analysis technique is restricted to small molecules.

At this point we run out of widely applicable methods for studying structures of 'free' molecules, but we have by no means run out of volatile compounds for which structural information would be extremely valuable. We have therefore turned to theoretical methods, which have developed and continue to advance in concert with the developments in computing power, to fill the gaps that are left in experimental structural studies of complex molecules. In this we do not merely *compare* theory and experiment; we *combine* them, to give structures which make use of the best of both worlds.

## The way forward: the SARACEN method – Structure Analysis Restrained by *Ab initio* Calculations for Electron diffraction

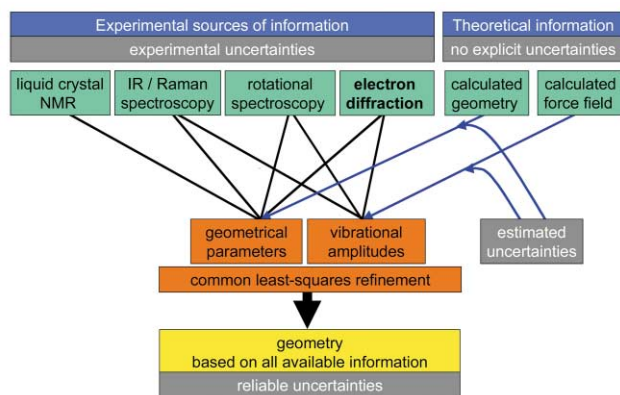
The SARACEN method, first described in two papers in 1996,<sup>11,12</sup> is not in fact limited to electron diffraction data. It can make use of measurements from any relevant source, experimental or theoretical; in effect it treats computed parameters in the same way as experimental observations, and then refines to a structure based on all the available observations. It is built on two important foundations, differing from earlier work chiefly in the *extent* to which theoretical data are used as *restraints*, and in the emphasis on refining *all* significant parameters, no matter how complex the molecule under investigation.

The MOCED (Molecular Orbital Constrained Electron Diffraction) technique was described by Schäfer *et al.* in 1982.<sup>13</sup> Parameters that could not be refined in analyses of electron diffraction data, particularly differences between similar distances or related angles, were fixed at their computed values. This valuable method allowed otherwise inaccessible structures to be determined and continues to be used by many groups, including, until very recently, ourselves. But it is less than perfect, because the theoretical information is applied as *rigid* constraints; each such constrained parameter is implicitly assumed to be *exactly* correct. Consequently, if one or more refining parameters are correlated with constrained parameters (and that is the usual reason for having to impose the constraint), the standard deviations may be underestimated, by a large but unknown amount.

The second foundation on which the SARACEN method is built is the technique of using predicate observations,

introduced by Bartell *et al.* in 1975.<sup>14</sup> This method involves the application of restraints (we use this word to indicate a flexibly imposed restriction, rather than a constraint, which is presumed to be rigidly fixed), based on any knowledge that one might have about the restrained parameter, usually no more than experience of similar bond lengths or angles in related compounds. In turn, this method was derived from the diagnostic least-squares procedure.<sup>15</sup> This similarly used flexible restraints based on previous general knowledge, but was specifically applied to structures for which some rotation constants were available, but not enough to allow full structure determination. In this case the procedure was designed to fit the experimental rotation constants exactly.

In the SARACEN method, experimental and theoretical data are analysed together (Scheme 2), in effect treating the theoretical method as if it were another experimental technique.<sup>11,12</sup> Computed parameters are then introduced into the joint refinement just like experimental data, weighted according to their uncertainties, and so they act as flexible restraints, with the degree of flexibility determined by the assigned uncertainties. Weights assigned to additional data are inversely proportional to the squares of their uncertainties, and are scaled to the electron diffraction data, for which uncertainties are taken to be the r.m.s. deviation between observed and calculated values. The uncertainties assigned to computed parameters are necessarily to some extent subjective, but the aim is to make them as objective as possible. In practice, calculations are done using a series of basis sets and at several different levels of theory. Parameters will vary according to the level of treatment, and by considering these variations, which decrease in magnitude as the treatment gets more and more sophisticated, it is possible to estimate the uncertainty for each parameter in the value obtained at the highest level of calculation attainable, *i.e.* the one that gives the lowest total molecular energy. One can also be guided by experience of calculations of the same types for similar molecules, and it is better to be conservative in estimating the uncertainty of a restraint than to define it too closely. Of course, differences between two chemically related distances or angles are likely to be more accurately computed than absolute values, so the choice of parameter to be restrained is critical.



Scheme 2

However, care must be taken to ensure that restraints are not derived from converged but systematically erroneous calculations. This point is critical in only a minority of cases. Usually restraints are applied to parameters describing the peripheral substituents, whose geometrical properties are reliably calculated, rather than to the molecular core, which is often the central point of interest of the structure determination.

The second essential feature of the SARACEN method is that *all* significant parameters are refined. If, as is often the case in normal analyses of electron diffraction data, some parameters are not included in the refinement, not only are they

themselves not investigated, but their possible influence on the values and standard deviations of refining parameters is neglected. Some researchers make allowance for this by arbitrarily increasing the standard deviations of the parameters that they report. But in some cases such increase may not be sufficient, while in others hard-earned precision may be unnecessarily discarded. The outcome of a SARACEN refinement is therefore a full set of structural parameters, with realistic estimates of the uncertainty for each of them. This is feasible even for very complex structures, which would otherwise be impossible to solve without making many assumptions. (Note that the improved fit to the experimental data is not simply a consequence of increasing the number of refining parameters. This increase is matched by a corresponding increase in the number of 'experimental' data, as restraints are added.) Thus in the end the structure is the best fit to all available structural data, theoretical and experimental, and should represent the best estimate that is currently attainable, with meaningful estimated standard deviations.

### An example of the SARACEN method, tin(II) acetate

The way in which computed restraints can supplement GED data is illustrated by a study of tin(II) acetate.<sup>16</sup> At least four possible starting structures have to be considered, involving asymmetrically and symmetrically bonded acetate groups with and without a stereochemically active lone pair at tin. It is the least symmetrical of these, with asymmetric ligands and an active lone pair, and thus  $C_2$  symmetry (Fig. 1), that proved to be correct. In this structure there are two similar but different Sn–O distances, two O–C distances and so on, and it cannot be assumed that the heavy atoms in each ligand are coplanar. As resolution of similar distances is not feasible using GED data alone, more information is required. Computed restraints were therefore applied to some parameters defining differences between distances, in particular differences between the two Sn–O distances and between the two C–O distances, as well as to small deviations from idealised local symmetry. Full details are given in the original paper.<sup>16</sup> The whole structure could then be determined reliably. Selected details are included in Fig. 1.

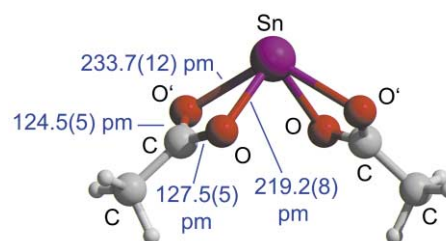


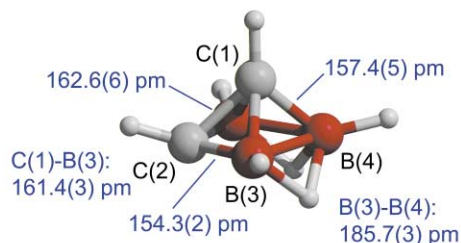
Fig. 1 Gas-phase structure of tin(II) acetate, with asymmetrically coordinated acetate ligands. The acetate groups are planar, but twisted by  $16.8(11)^\circ$  from a  $C_{2v}$  structure.

### The SARACEN method with rotation constants

The method is not limited merely to analysis of GED data with computed restraints. *Any* experimental data relating to structures can be incorporated. In a study of the carbaborane *nido*-1,2- $C_2B_3H_7$  (Fig. 2)<sup>17</sup> the very small amount of material available precluded the recording of a complete data set and as a compromise data were obtained from only one camera distance in the ED experiment. (Normally two or three data sets with different ranges of the scattering angle would be recorded.) However, there were nine rotation constants available from different isotopomers of *nido*-1,2- $C_2B_3H_7$ , which were included in a combined refinement of ED and MW data and *ab initio* restraints. The result was a complete molecular structure, consistent with measured and calculated  $^{11}B$  NMR chemical shifts and theoretical predictions of the geometry. In the first refined structure, derived from microwave data alone, the C–C

**Table 1** PFH<sub>2</sub> parameters and restraints

Parameter	Refined value	Uncertainty of restraint	Refined standard deviation
$r(\text{P-F})/\text{pm}$	160.4	2.0	0.1
$r(\text{P-H})/\text{pm}$	142.3	1.0	0.1
$\angle(\text{FPH})^\circ$	98.1	1.0	0.5
$\angle(\text{HPH})^\circ$	92.1	1.0	0.3

**Fig. 2** Nido-1,2-C<sub>2</sub>B<sub>3</sub>H<sub>7</sub>, showing similar B–B, B–C and C–C distances, all of which have been refined using the SARACEN method.

distance was reported to be 1.453 Å, *i.e.* 9 pm shorter than the sum of the covalent radii, implying some double bond character to be retained in the cage. The new value of 1.626(6) Å indicates less than single-bond order for this C–C unit.

The SARACEN procedure is not restricted to GED data analysis, and is in principle applicable to all kinds of data analysis where insufficient experimental data can be augmented by reliable calculated parameters. One example is the structure determination of the monohalogenophosphines PH<sub>2</sub>X.<sup>18</sup> For a single isotopic species of, say, PFH<sub>2</sub>, three rotation constants can be measured. However, four parameters, two bond lengths and two angles, are required to define the structure. Without further information, none of these parameters can be determined individually. Use of data for a second isotopic species may solve the problem, but replacement of H by D changes three of the four parameters significantly, so we are no further forward, and P, F and I only have one stable isotope. But application of even weak restraints to the four parameters allows all of them to be refined – and they all return *small* standard deviations (Table 1). Accurate information is present in the three rotation constants, but it is locked up until released by the application of quite gentle restraints.

### Scraping the barrel? Small molecules for which there is insufficient experimental information

For the purpose of simplicity one is often tempted to assume that a molecule as a whole has some particular symmetry or that some part of it has local symmetry. However, surprisingly often our chemists' intuition turns out to be misleading. This has led to oversimplified or wrong molecular structures being published. Reducing molecular symmetry increases the number of parameters and makes it necessary to adopt sensible assumptions if it is not possible to refine all of these parameters independently. The SARACEN procedure is here the best way to extract the maximum amount of information from the experimental data, but careful selection of restraints and assignment of weights to them are necessary in order not to suppress the experimental information. Often the structure of the important molecular backbone or core can be determined from the experimental data alone, and theory is only needed to adjust the geometry of the outer ligand sphere and thus to reduce correlation with the core parameters.

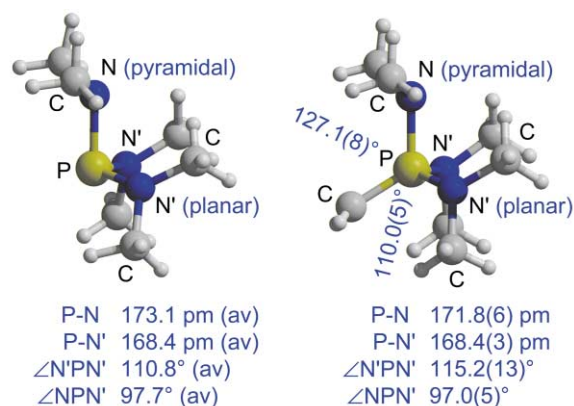
Sometimes the geometry of even a very simple molecule can be underdetermined even though a huge amount of experimental information is available. Usually the problem arises from the presence of similar internuclear distances (overlapping Gaussian peaks in the GED radial distribution curve cannot be resolved) or from the small contribution of weakly scattering

atoms (particularly hydrogen) to the experimental information if strong scatterers are also present. Here we show how information from the data is released by refining parameters that have been restrained, directly or indirectly, rather than fixing them at certain values. The application of restraints is not limited to geometrical parameters, but can also be used to support the refinement of amplitudes of vibration and other parameters, such as energy differences between the constituents of conformational mixtures.

### P(NMe<sub>2</sub>)<sub>3</sub> and CH<sub>2</sub>=P(NMe<sub>2</sub>)<sub>3</sub>, not at all threefold propellers<sup>19</sup>

We became interested in the structure of one of the simplest triaminophosphines, P(NMe<sub>2</sub>)<sub>3</sub>, when we realised that in many crystal structures with P(NMe<sub>2</sub>)<sub>3</sub> fragments (mostly as complex ligands), pronounced deviation from local C<sub>3</sub> symmetry was found and often geometries close to local C<sub>s</sub> symmetry were adopted. There was also discussion in the literature<sup>20</sup> for many years about the structure of tris(dimethylamino)phosphine, P(NMe<sub>2</sub>)<sub>3</sub>, with arguments for and against C<sub>s</sub> and C<sub>3</sub> symmetry, based on the results of photoelectron spectra. In the 1970s the structure of P(NMe<sub>2</sub>)<sub>3</sub> was determined by GED, but analysed under the assumption of C<sub>3</sub> symmetry. However, we eventually showed that it adopts C<sub>s</sub> symmetry in the crystal structure, having two planar NMe<sub>2</sub> groups with short P–N bonds, as well as one pyramidal NMe<sub>2</sub> group with a long P–N bond. We therefore undertook a new refinement using the GED data and a SARACEN-like method, and confirmed that this C<sub>s</sub> symmetric structure describes all the molecules present in the gas phase. *Ab initio* calculations supported these findings.

In CH<sub>2</sub>=P(NMe<sub>2</sub>)<sub>3</sub> there seems to be a better reason for the whole molecule to adopt C<sub>s</sub> symmetry, as the ylidic methylene unit does not allow exact C<sub>3</sub> symmetry. In fact this molecule does have C<sub>s</sub> symmetry, with two different NMe<sub>2</sub> group geometries, just as in P(NMe<sub>2</sub>)<sub>3</sub>. This is obviously not due to the presence of the CH<sub>2</sub> unit, but an inherent phenomenon of the PN<sub>3</sub> unit. Most intriguing in CH<sub>2</sub>=P(NMe<sub>2</sub>)<sub>3</sub> are the grossly different C=P–N angles: 127.1(8)° to the pyramidal and 110.0(5)° to the planar NMe<sub>2</sub> groups (Fig. 3). Without the application of suitable restraints describing the differences in the bonding parameters of the non-equivalent NMe<sub>2</sub> groups (the

**Fig. 3** Crystal structure of P(NMe<sub>2</sub>)<sub>3</sub> and gas-phase structure of CH<sub>2</sub>=P(NMe<sub>2</sub>)<sub>3</sub>, including selected parameters. Average values (av.) are provided for the crystal structure, which did not show molecular symmetry. The methylene hydrogen atoms of CH<sub>2</sub>=P(NMe<sub>2</sub>)<sub>3</sub> are behind one another.

SARACEN method) a reliable refinement could not have been achieved either for  $\text{P}(\text{NMe}_2)_3$  or for  $\text{CH}_2=\text{P}(\text{NMe}_2)_3$ .

#### Ring systems without symmetry, $\text{Me}_2\text{NN}(\text{SiH}_2\text{CH}_2)_2$ and $\text{Me}_2\text{HCN}(\text{SiH}_2\text{CH}_2)_2$ <sup>21</sup>

These molecules comprise a highly flexible five-membered ring fused to either an  $\text{NMe}_2$  group or an isoelectronic  $\text{CHMe}_2$  group. In the first case a high barrier to rotation about the N–N bond is to be expected because of N–N lone pair repulsion, but in the second case the barrier should be relatively low. Early attempts to refine these structures assuming local  $C_{2v}$  symmetry at the ring nitrogen atoms and  $C_2$  symmetry for the rings failed completely. It turned out that the ring geometries deviated from these idealised structures too much to allow a good fit to the experimental data. Furthermore, a secondary interaction between one silicon atom and the nitrogen atom of the  $\text{NMe}_2$  group in  $\text{Me}_2\text{NN}(\text{SiH}_2\text{CH}_2)_2$  leads to a drastic deviation from equal Si–N–N angles. The many parameters necessary to describe the structures completely in  $C_1$  symmetry made it impossible to refine all of them freely. Constraining most of the differences between similar parameters (*i.e.* application of the MOCED method) led to an improvement in the fit to the experimental data, but did not allow retrieval of all the important structural information. By changing these constraints into restraints it was possible to reach a good fit to the experimental data in a stable and convergent least-squares refinement. The resulting two structures are shown in Fig. 4.

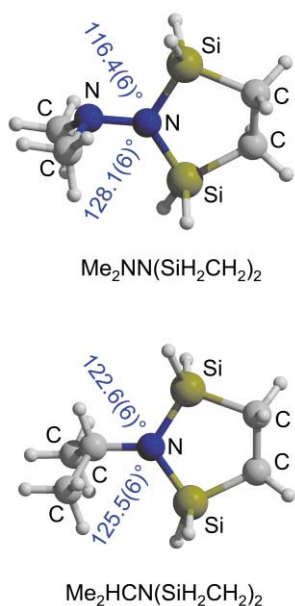


Fig. 4 Structures of  $\text{Me}_2\text{NN}(\text{SiH}_2\text{CH}_2)_2$  and  $\text{Me}_2\text{HCN}(\text{SiH}_2\text{CH}_2)_2$ .

#### Other molecules with low symmetry

One of the major advances that has been made possible by the introduction of the SARACEN method is the application of GED to molecules with low symmetry, so there are many examples that could have been chosen here. Vanadyl nitrate,  $\text{VO}(\text{NO}_3)_3$ , could in principle have threefold symmetry, being  $C_3$  or even  $C_{3v}$ . In fact it has  $C_s$  symmetry, with all three nitrate groups being planar, one lying on the mirror plane and the other two related to one another by the mirror symmetry (Fig. 5).<sup>22</sup> As each of the groups is bidentate, but with one nitrogen much closer to the vanadium atom than the other, there are four different V–O distances in the molecule, as well as the V=O bond distance. Similarly, there are altogether six different N–O distances, and groups of related angles. It is quite impossible to refine all of these independently, as they are of course very strongly correlated. Adding restraints to differences between related parameters removes most of this correlation,

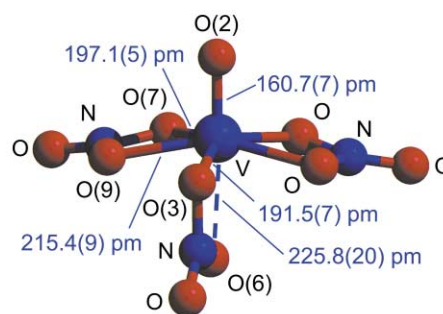


Fig. 5 The structure of  $\text{VO}(\text{NO}_3)_3$ , showing key inter-atomic distances.

allowing the whole structure, best described as being derived from a pentagonal bipyramid, to be determined.

Other examples include  $\text{Bu}^t\text{NSNBu}^t$ , which exists effectively entirely as the *E,Z* conformer (Fig. 6).<sup>23</sup> This structure is therefore more complex than that of its larger relative  $\text{S}(\text{NBu}^t)_3$ , because it has substantially different parameters for the two *N*-butyl groups, including SNC angles of 117 and 126°. It is essential to allow for such effects in the refinement, and this is best done by adding restraints on the differences, but then refining both of the parameters. Note that computed absolute values for the S=N distances are unreliable, even with the most modern DFT functionals, so it is very important only to apply restraints to difference parameters.

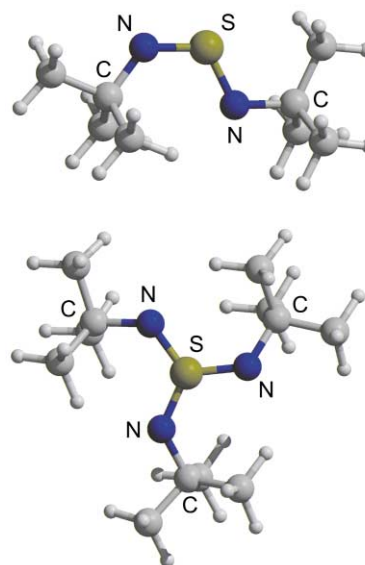
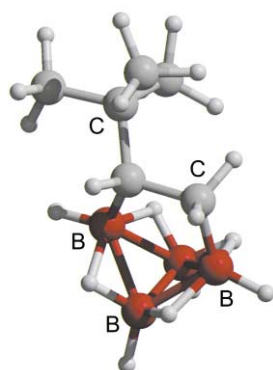


Fig. 6 Structures of  $\text{S}(\text{NBu})_2$  and  $\text{S}(\text{NBu})_3$ .

Our studies of boranes and their derivatives have also been greatly enhanced by the ability to deal properly with molecules with low symmetry. The structures of the unusual ‘basket’ compounds that are derived from  $\text{B}_4\text{H}_{10}$  by having two of the ‘wingtip’ hydrogen atoms replaced by a  $-\text{CR}_2\text{CR}_2-$  linking moiety (Fig. 7) can easily be determined, even when there is a single *t*-butyl substituent on the linking group, removing all symmetry.<sup>24</sup>

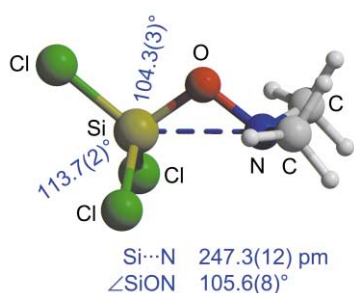
#### Structures dominated by a group of strong scatterers

In GED studies it may be difficult or impossible to determine all the geometrical parameters if there are a few intensely scattering atoms present in a group while other less strong scatterers comprise an important part of the structure. The silicon and three chlorine atoms in  $\text{Cl}_3\text{SiONMe}_2$ <sup>25</sup> form such a unit, which dominates the whole scattering information, even in the non-bonding region of the radial distribution curve. Problems arose



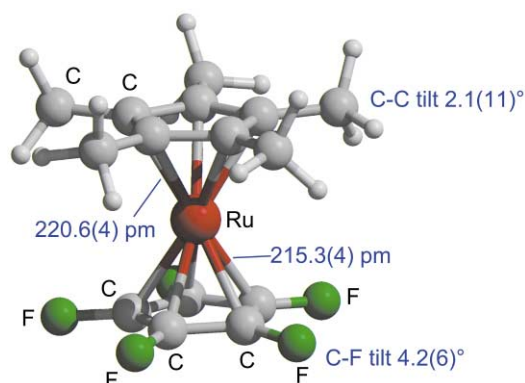
**Fig. 7** The structure of  $B_4H_8CH_2CHBu^t$  is asymmetric, but all heavy-atom parameters were determined.

in the refinement of the geometry of the  $ONMe_2$  group without restraints. Both the  $ONC$  and  $CNC$  angles adopted values that seemed far too large in comparison to those in other O-silyl hydroxylamines, and were also inconsistent with those found in the crystal structure of this compound. Simply restraining the sum of the angles at this nitrogen atom to the computed value, with an uncertainty estimated to be  $2^\circ$  (based on the deviations between experimental and calculated parameters for many other  $R_3SiONMe_2$  compounds), led to a stable refinement, only marginally higher in  $R$ -factor. Other parameter values were almost unchanged from those obtained in the non-restrained refinement (Fig. 8).



**Fig. 8** The structure of  $Cl_3SiONMe_2$ , showing key parameters.

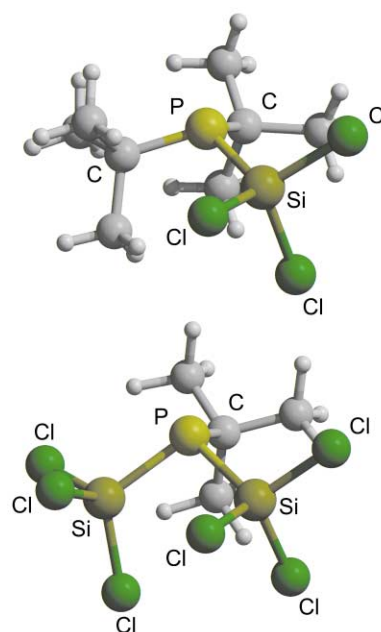
A study of  $Ru(C_5Me_5)(C_5F_5)$ , which is the first structure including a  $C_5F_5$  group to be determined in the gas phase, had problems due to a single dominant scatterer and due to overlapping peaks for similar distances.<sup>26</sup> The two different  $Ru-C$  distances are similar, as are the  $Ru \cdots C(Me)$  and  $Ru \cdots F$  distances, so resolution of them is difficult, and the contributions from the remaining atoms pairs are all relatively weak, although there are many different distances. Moreover, it is not easy to do reliable calculations of the structure of this molecule. Very high levels of theory (unrealistically expensive) are needed for *ab initio* methods. Density functional theory is the only practical way to compute data, and it is essential to calibrate the calculations against known structures (such as ferrocene). Restraints on differences between equivalent parameters for the two rings were used, and it was then possible to refine both the  $Ru-C$  distances, which were 215.3(4) pm for the  $C_5F_5$  ring and 220.6(4) pm for the  $C_5Me_5$  ring. Both of these are shorter than in ruthenocene itself. Although the difference was restrained, the esd for the refined value was smaller than the uncertainty of the restraint, so we can tell that there is some information about the difference in the experimental data. It was also possible to determine the extent to which the  $C-F$  and  $C-C(Me)$  bonds are bent out of their respective ring planes, 4.2(6) and 2.1(11) $^\circ$ , respectively (Fig. 9). Both these values are much larger than in ruthenocene itself, suggesting that the displacement is steric in origin, rather than associated with the electronegativity of the substituents.



**Fig. 9** The structure of  $Ru(C_5Me_5)(C_5F_5)$ , showing  $Ru-C$  distances and  $C-F/Me$  tilt angles (away from the ruthenium atom).

### When local symmetry is important

We have seen that assumptions about molecular symmetry, made to simplify the analysis of GED data, may sometimes be inaccurate. Similarly, it is normal to make assumptions about the local symmetry of fragments of a molecule, even though these are in principle inconsistent with the molecular symmetry. For example, one would almost always assume that a methyl group has  $C_{3v}$  local symmetry, although that is rarely strictly true. That is hardly going to matter, but an  $SiCl_3$  group in a crowded molecule may be substantially distorted, and as the chlorines contribute significantly to the scattering, neglect of the distortion may seriously affect refined parameters. The SARACEN method makes it possible to apply restraints to parameters such as differences between the various  $ClSiCl$  angles, and consequently the large number of geometrical parameters needed to describe such an asymmetric group can all be refined. This has been done in studies of the crowded molecules  $PBu^t_2(SiCl_3)$ <sup>27</sup> and  $PBu^t(SiCl_3)_2$  (Fig. 10).<sup>28</sup> In these compounds the *tert*-butyl and trichlorosilyl groups are all twisted, in the same direction, away from the fully staggered positions, to reduce 1,3-interactions between methyl groups and chlorine atoms. The molecules therefore have  $C_1$  symmetry, and altogether 27 parameters are required to define the positions of carbon and chlorine atoms. Nevertheless, all of these could be refined, revealing big differences within the sets of  $PCC$  and  $PSiCl$  angles, but rather little distortion within the groups,



**Fig. 10** The structures of  $PBu^t_2(SiCl_3)$  and  $PBu^t(SiCl_3)_2$ . The butyl and trichlorosilyl groups are tilted, predominantly towards the lone pair of electrons on phosphorus.

*i.e.* of CCC and ClSiCl angles. For the latter groups of parameters, once it was confirmed both that the angle differences were small and that they were not significantly correlated with other refining parameters, it was possible to fix the differences either to zero, or to the small computed values. Attention was then focussed on the tilting of the groups, which involved large components towards the phosphorus lone pair, and smaller components around the equatorial belt.

Allowance for possible reduction from the idealised maximum possible local symmetry for substituent groups is now routine in our work. Other examples of asymmetric groups include the *tert*-butyl groups in  $\text{Bu}^t_2\text{SiHSiH}_2\text{Bu}^t$ ,<sup>29</sup> the nitrate groups in  $\text{VO}(\text{NO}_3)_3$ <sup>22</sup> (see the section on molecules with low symmetry, above), and the bis(trimethylsilyl)methyl groups in the phosphorus and arsenic radicals described in the section on gas–solid differences, below.<sup>30</sup>

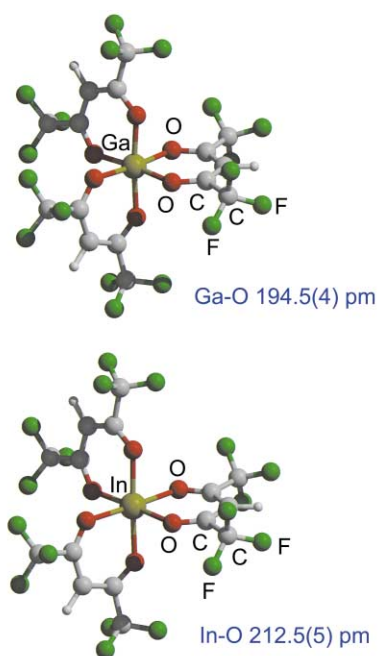
### Bigger and worse? Large and complex molecules

The sheer size of a molecule can be a problem for structural investigation in the gas phase, as usually many similar inter-nuclear distances occur in such systems. Of course, a large number of atoms does not of itself make a structure difficult to solve. The structures of  $(\text{Bu}^t\text{GaS})_4$  and its selenium analogue, based on a central  $\text{Ga}_4\text{S}_4$  cube, have only nine independent parameters, and can be determined without need for the SARACEN method,<sup>31</sup> while  $\text{C}_{60}$  only requires two geometric parameters to be completely determined.<sup>32</sup> Although large molecules tend to be less volatile than smaller ones, developments have been made to allow GED data collection at lower vapour pressures, and so there are now very many compounds for which gas-phase structures are potentially accessible. These large molecules also present major problems for high-level computational methods. Nevertheless, molecules of comparatively large size have been studied using the SARACEN method.

### Forty and more non-hydrogen atoms

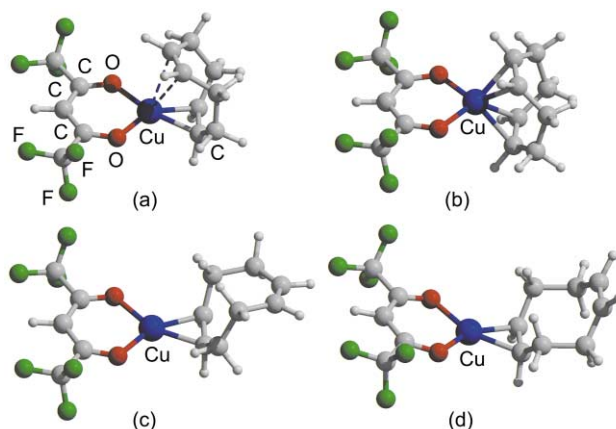
An example is provided by the two homologous compounds  $\text{Ga}(\text{hfac})_3$  and  $\text{In}(\text{hfac})_3$  (hfac = 1,1,1,5,5,5-hexafluoropentane-2,4-dionato),<sup>33</sup> both containing 40 non-hydrogen atoms. These compounds are useful chemical vapour deposition precursors for materials containing Ga and In and therefore their structures in the gas phase are of interest. Despite having  $D_3$  symmetry, the number of parameters needed to describe the molecular geometry is large and there are many important amplitudes of vibration to refine. In such cases SARACEN turns out to be extremely useful for refining the complete molecular structure. For example, in this case the C–C<sub>ring</sub> and the C–C<sub>F</sub> bond lengths were each restrained. The uncertainties were derived from the differences between values obtained by SCF and DFT calculations on these systems, bearing in mind the fact that SCF calculations often underestimate these values, while they are overestimated by DFT calculations. Further restraints were also applied to the CCF angles and CH bond lengths, as well as to a number of amplitudes or groups of them. The final structures of the compounds shown in Fig. 11 deviate in some torsional angles OCCF from the theoretical predictions. This reflects the high flexibility of these molecular sub-units, which make it difficult to get very accurate torsional parameters by computational methods. Even larger deviations of this type are observed in the crystal structure of  $\text{Ga}(\text{hfac})_3$ , which is also of lower molecular symmetry ( $C_1$ ) than the free molecule ( $D_3$ ), showing the ease with which the conformation is distorted.

The hfac ligand also features in an interesting copper complex, with fewer atoms than  $\text{Ga}(\text{hfac})_3$ , but lower symmetry. The complex  $\text{Cu}(1,5\text{-cod})(\text{hfac})$  is useful as a precursor for the deposition of a conducting layer of copper from the gas phase,



**Fig. 11** The refined structures of  $\text{In}(\text{hfac})_3$  and  $\text{Ga}(\text{hfac})_3$  are almost identical, the main difference being the gallium–oxygen distance of Ga–O 194.5(4) pm, compared with 212.5(5) pm for indium–oxygen.

decomposing by disproportionation to give the volatile copper(II) complex  $\text{Cu}(\text{hfac})_2$  and free cyclooctadiene. Knowledge of the gas-phase structure is a prerequisite for understanding the process of deposition on a hot surface, but there are many possibilities. The 1,5-cyclooctadiene ligand can be singly or doubly coordinated to the copper atom, or possibly with one strong link and one weak, and if singly coordinated, the second double bond can be *endo* or *exo*. Performing calculations at a high enough level to be reliable is a challenge, but eventually it could be shown that a structure with an asymmetrically chelating ligand and  $C_1$  symmetry (Fig. 12) had the lowest energy, and the GED data were consistent with this.<sup>34</sup> Without computed information it would be impossible to be confident in the results of an experimental determination for a structure of this complexity. On the other hand, using some of the computed parameters as restraints, one can use the experimental data to give values for those parameters that are least well determined theoretically, particularly those relating to the copper atom and its immediate coordination environment. One does indeed get the best of both worlds, experimental and theoretical.



**Fig. 12** Four possible structures of  $\text{Cu}(1,5\text{-cod})(\text{hfac})$ . Structure (a) was found to be predominant in the gas phase. One olefinic group of the cyclooctadiene ligand is coordinated to the square-planar copper atom [refined Cu–C distances 194.0(13) and 194.4(9) pm]. The second C=C double bond is weakly associated with the copper atom [Cu  $\cdots$  C distances 267.2(23) and 276.9(25) pm].

A structure with an even larger number of atoms is the fullerene  $C_{70}$ , which was studied by Hedberg *et al.*<sup>35</sup> These authors did not use the SARACEN method explicitly, but applied five restraints to support the refinement with assigned uncertainties drawn from experience, perhaps rather smaller than one would derive from a graded series of computed values.

### Conformational nightmares

Conformational analysis is one of the most important applications of gas-phase electron diffraction, which still is most impressively demonstrated by the classic work of Nobel laureate Odd Hassel on the conformations of cyclohexane and decalin.<sup>36</sup> Although GED generally does not give relative amounts of several components of a gaseous sample with high accuracy, the proportions may be very sensitive to the energy differences, and so experiments are often more accurate than computational methods for the determination of relative energies of conformers.

Since the time of Hassel many conformational mixtures have been examined, including complex cases with many more than two conformers present simultaneously in the gas phase. The limited amount of experimental information has to be used to determine the structures of two or more conformers. In the past one had to assume that, apart from the conformation-defining torsional angles, the geometrical parameters are the same in the different conformers. This might be true enough in some cases, but it is certainly not the case in others. Such simplifying assumptions can be extremely misleading. The SARACEN method can provide the missing information in such cases, by introducing restraints on the differences between related parameters in the various conformers. The composition of the gas can then be determined, as well as the parameters for all components, and so important experimental information is released, both thermodynamic and geometric.

The principles of the procedure are illustrated by a study of the structures of 1,4-disilabutane and 1,5-disilapentane.<sup>37</sup> The disilabutane exists predominantly as the *anti* conformer in the gas phase [experiment 76(2)%, theory (levels up to CCSD(T) and basis sets up to 6-311G\*\*) *ca.* 83%], the remainder being the *gauche* form, with the SiCCSi angle close to  $80^\circ$  (Fig. 13). Most parameters for the two conformers are calculated to be very similar, except for the SiCC angles, which are calculated to be  $3.7^\circ$  wider in the *gauche* form. This value was used as a restraint, with an uncertainty derived from the extent of its variation in the different types of calculation, and then the SiCC angles in both conformers were refined. Similarly, the small differences between Si–C distances, and between C–C distances, in the two forms were also refined, subject to restraints. In this case, however, the esds for the refined differences were very close to the uncertainties applied to the corresponding restraints. This tells us that there is effectively no information about these differences in the experimental data. If the least-squares correlation matrix tells us that no other refining param-

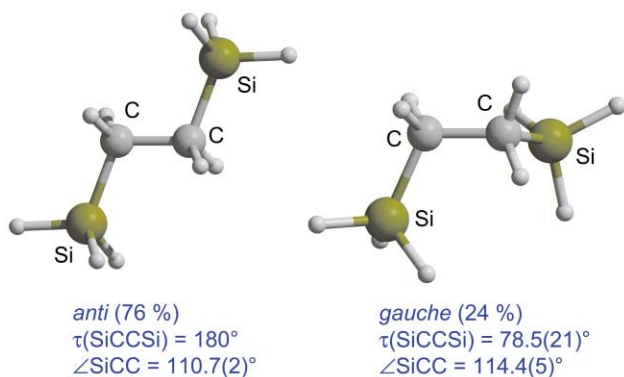


Fig. 13 The two conformers of 1,4-disilabutane.

eters are correlated with the restrained differences, it is possible to replace the restrained parameter by one fixed at the calculated value. In these circumstances, the refinement is exactly equivalent to a MOCED refinement, but with the important distinction that one can be sure that the estimated standard deviations are realistic, and not unrealistically reduced by ignoring some correlation to a fixed parameter.

Another example of the use of the SARACEN method in the study of two conformers is given in the section on secondary bonding, below. In other cases there have been more than two conformers, and in such situations electron diffraction is not good at giving relative proportions. If in addition the species present have different values for major geometrical parameters, then GED may have little to offer at all, and there is no other gas-phase structural technique to bring to bear on the problem. But if differences in parameters can be restrained at computed values, and relative populations are fixed as derived from calculated energies, the primary structural parameters can be obtained experimentally. This has been done for some compounds containing three  $\text{SiMe}_2\text{H}$  groups bound to carbon, which can have no less than eleven conformations. In  $\text{CH}(\text{SiMe}_2\text{Br})_3$  just three of these are significantly populated at the temperature of the GED experiment, and with the application of suitable restraints on difference parameters relating to the positions of the heavier atoms, the structures could be refined.<sup>38</sup> However, although the calculated abundances of the three conformers were 76, 15 and 9%, the data could only be fitted satisfactorily with approximately equal amounts of all three. If the bromine atoms are replaced by hydrogen, nine of the eleven conformers become significantly populated at the temperature of the experiment,<sup>38</sup> and in  $\text{C}[\text{SiMe}_2\text{H}]_3\text{SiH}_3$  all eleven (Fig. 14) fall within a total energy range of just  $3 \text{ kJ mol}^{-1}$ , and have abundances of between 3 and 16%.<sup>39</sup> With such an amazingly narrow energy range, one has to include all the conformers in the analysis, and if they were treated separately there would be a huge number of refining parameters. So a simplifying strategy was adopted, with the  $\text{SiMe}_2\text{H}$  ligands grouped according to their local environment. Thus the 27 different dihedral angles were reduced to seven, and only three  $\text{Si}(\text{H}_3)\text{CSi}(\text{Me}_2\text{H})$  angles were needed, a much more manageable proposition. Thus once again, even in this extreme situation, the experimental GED data could provide the key parameters, without having to make unwarranted assumptions.

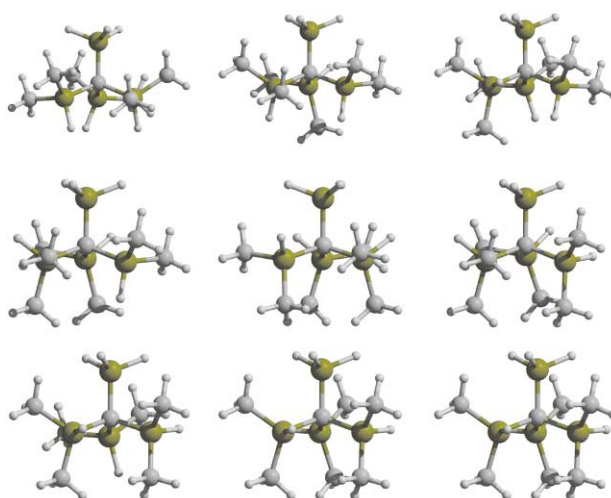


Fig. 14 The nine conformers of  $\text{C}(\text{SiMe}_2\text{H})_3\text{SiH}_3$ , out of eleven possible, that are most significantly populated.

### Other structural changes associated with large-amplitude torsional motions

The very fast electron diffraction experiment averages scattering information over all conformations present in the gas phase,



and if the torsional potential function is relatively flat the significantly populated distribution of instantaneous conformations covers a wide range. This continuous distribution can be modelled with a few discrete conformations, weighted according to their relative energies. If there are parameters that vary significantly as the torsion angle changes a problem arises, because of the huge number of parameters necessary to describe all these different conformers. The problem can be solved by applying restraints to difference parameters, in much the same way as in a study of several distinct conformations.

Nitrosomethane and its derivatives exhibit very large-amplitude torsional motions, due to the interactions between the three substituents on carbon at one end of the C–N bond and the nitrogen lone pair of electrons and the oxygen atom at the other end. This leads in principle to a six-fold barrier to rotation, with low barrier heights and therefore rapid interconversion of the different conformers. In a study of chlorodifluoronitrosomethane,  $\text{CClF}_2\text{NO}$ ,<sup>40</sup> *ab initio* calculations indicated that it consists of two conformers, in which the N–O bond approximately eclipses the C–Cl ( $C_s$  symmetry) and C–F bonds ( $C_1$ ). These two forms differ in energy by only  $1.1 \text{ kJ mol}^{-1}$  [MP2/TZ2P + ZPE(MP2/DZP)] and are separated by a barrier of about  $1 \text{ kJ mol}^{-1}$ . There is therefore barely restricted torsional motion over an extremely wide range, the only obstruction to free internal rotation being a barrier of  $5\text{--}10 \text{ kJ mol}^{-1}$  between the two equivalent  $C_1$  conformations. Associated with the torsional motion there were huge variations of angles at carbon and nitrogen, up to  $9^\circ$  for NCCl and  $5^\circ$  in the case of the CNO angle (Fig. 15). Eleven static conformations were used to represent the dynamic situation, and differences between the highly variable angles in these conformers were initially restrained. Once it became clear that these difference parameters were not correlated with other refining parameters, it was possible to fix them, thus simplifying the refinement process considerably. Overall, a highly complex dynamic situation has been modelled accurately, and one can be sure that there are no correlations that could, by being hidden, lead to unrealistically small standard deviations in refined parameters.

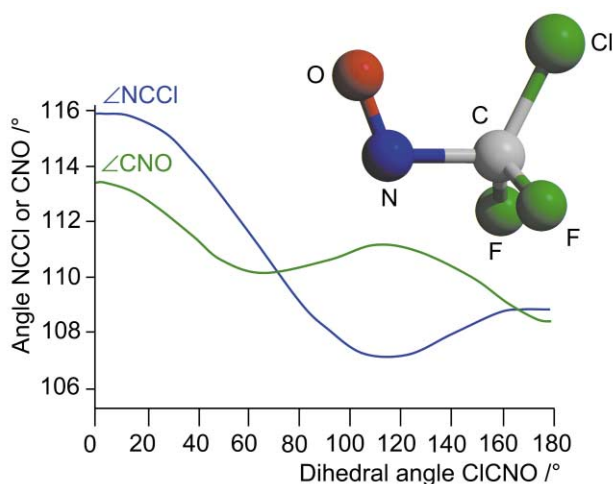


Fig. 15 Variations in the CNO and NCCl angles as functions of the ClCNO dihedral angle in  $\text{CCl}_2\text{NO}$ .

## Gases and solids

Although one often reads that in a crystal structure there are ‘no significant intermolecular interactions’, the very fact that the compound has crystallised tells us that this statement is not true. The interactions may only involve van der Waals forces, but even these can have a profound effect on the structure of the molecules. In other cases, specific interactions are very likely to

cause significant differences between structures in gaseous and crystalline phases. This is one reason why it is so important to determine structures in the gas phase.

An extreme example of structural change with phase is provided by the bis[bis(trimethylsilyl)methyl]phosphido radical,  $\text{P}[\text{CH}(\text{SiMe}_3)_2]_2$ , and its dimer,  $\text{P}_2[\text{CH}(\text{SiMe}_3)_2]_4$  (Fig. 16).<sup>30</sup> The diphosphine, studied by X-ray crystallography, is under severe steric strain, which is shown particularly in distortion of inter-bond angles from their normal values. For example, there is a difference of as much as  $12^\circ$  between the two PCSi angles at a single carbon atom. There is thus a lot of strain energy stored in the molecules, but *ab initio* calculations show that this is not enough for the molecules simply to dissociate, and so they are perfectly stable in the crystal. However, dissociation can occur in the vapour phase, because one of the  $\text{CH}(\text{SiMe}_3)_2$  groups on one phosphido radical can rotate about its P–C link, reducing the strength of the interactions between the two groups in the one radical. The relaxation of the steric strain is thus the main contribution of energy necessary to break the P–P bond, which was a normal, strong bond, and not one of exceptional length and on the verge of breaking. The diphosphine only holds together in the crystalline phase because of the packing forces. The dissociation energy was thus stored in the many distorted angles and lengthened bonds in the diphosphine – a so-called ‘Jack-in-the-box’ molecule.

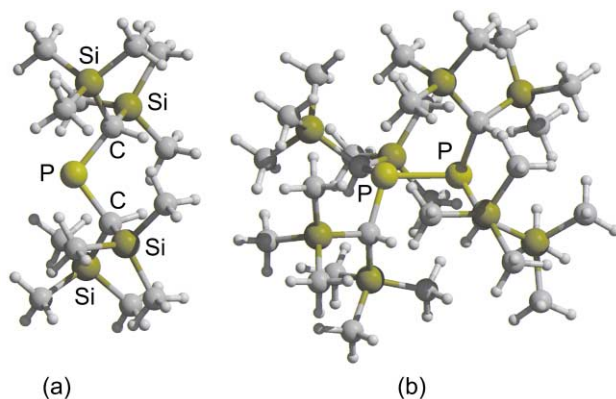
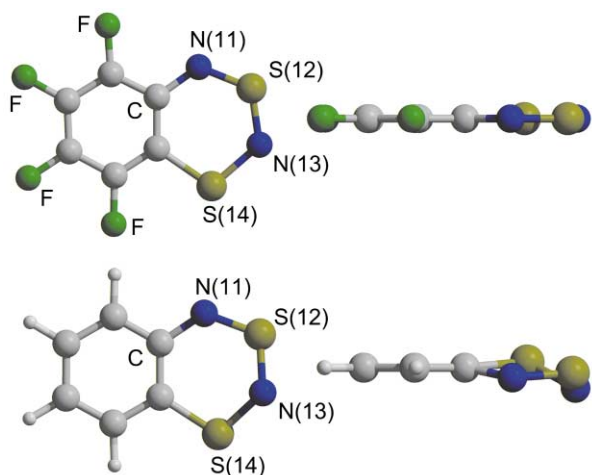


Fig. 16 (a) The gas-phase structure of the radical  $\text{P}[\text{CH}(\text{SiMe}_3)_2]_2$  and (b) the structure of its dimer in the crystalline phase.

The structure of the radical is a challenge for GED – even though it is only ‘half a molecule’! The  $\text{CH}(\text{SiMe}_3)_2$  groups have no local symmetry, and there are several different P–C and Si–C distances that are very similar, and give rise to overlapping peaks in the radial distribution curve, as do many of the non-bonded distances. Until quite recently the size of the diphosphine and the radical (for which unrestricted calculations are needed) made high-level *ab initio* calculations excessively expensive, while parameterisation for simpler calculations, such as molecular mechanics, was not available. So the ability to perform calculations at a high enough level to derive reliable restraints, coupled with the SARACEN method for interpretation of the experimental diffraction data, has at last allowed the gas-phase structure to be determined. That has uncovered a most unusual dissociation behaviour, which has in turn led to a new approach to the thermodynamics of bond dissociation.<sup>41</sup>

The continuing importance of experimental information for gas-phase structures is illustrated by the molecules  $1,3\lambda^4\delta^2,2,4$ -benzodithiadiazine,  $\text{C}_6\text{H}_4\text{NSNS}$  and its 5,6,7,8-tetrafluoro derivative. The parent compound is planar in the crystalline phase,<sup>42</sup> whereas when it is fluorinated the ring containing the sulfur and nitrogen atoms is bent along the N11–S14 line.<sup>43</sup> (See Fig. 17 for the atom numbering.) However, Density Functional Theory calculations at the B3LYP/6-31G\* level suggest that for the free molecules the situation is reversed (the parent is



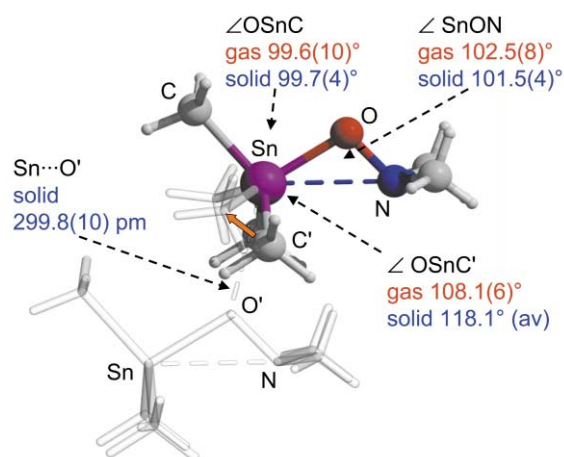
**Fig. 17** Top and side views of  $C_6F_4NSNS$  (planar) and  $C_6H_4NSNS$  (non-planar).

non-planar and the fluoro derivative is planar), while *ab initio* calculations at the MP2/6-31G\* level suggest that both molecules are non-planar. For compounds of this type, with multiple bonds between second-row and first-row elements, DFT methods, even with the latest functionals, are notoriously unreliable. This is clearly a case where experimental gas-phase data are needed, but the molecules must be modelled in  $C_1$  symmetry, and there are many sets of similar distances, which cannot be resolved by GED.

The SARACEN method yielded a solution to this problem, but great care was needed to ensure that computed restraints were only used for parameters that could safely be assumed to be reliable.<sup>44</sup> The B3LYP and MP2 calculations were repeated, using the 6-311+G\* basis set for nitrogen, fluorine and sulfur. In the SARACEN refinement parameters relating to the benzo rings of the molecules were fairly tightly restrained, because the values were consistent between the different calculations. For the heterocyclic ring the important distances and angles, and the all-important parameters defining possible deviation from planarity, were not restrained. The outcome was an unequivocal demonstration that the parent compound,  $C_6H_4NSNS$ , is non-planar in the gas phase, whereas  $C_6F_4NSNS$  is planar. In this case, DFT calculations are correct in predicting the conformation, whereas MP2 calculations are incorrect, although the DFT distances can not be relied on. But the most important conclusion is that the conformations in the gas phase are the exact opposite of those for the crystalline phases. This demonstrates that data for both phases are needed; one cannot assume that the structure is the same in both phases.

An example which shows how much the molecular structure of organometallic compounds is dependent on even very weak intermolecular contacts is the structural analysis of  $Me_3SnONMe_2$ .<sup>45</sup> Fig. 18 shows a superposition of the gas phase structure and the crystal structure of this molecule. An intermolecular  $Sn \cdots O$  contact of almost 300 pm length leads to a drastic distortion of the coordination sphere at the tin atom, away from a tetrahedral coordination geometry with a fifth intramolecular contact to the N atom (4 + 1 coordination) to a coordination geometry of a trigonal pyramid with two further contacts, intramolecular to an N and intermolecular to an O atom (4 + 2 coordination).

There are, of course, many other examples of major differences between structures in solid and gaseous phases. Bis(*tert*-butyl)ferrocene<sup>46</sup> is another organometallic one, where there is also a very complicated conformational problem, caused by rotation of the rings relative to one another, restricted only when the butyl groups are nearly eclipsed. Other examples are given in the next section of this article.



**Fig. 18** Superposition of the gas-phase (bold) and crystal (feint) structures of  $Me_3SnONMe_2$  showing the change of structure upon phase change (red arrow).

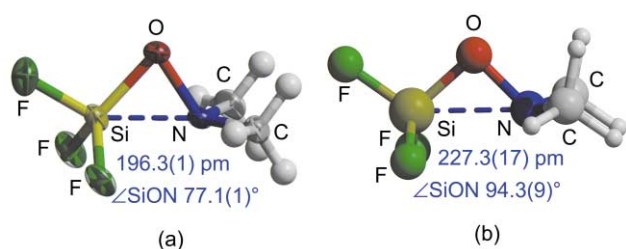
## Secondary interactions

In this final section describing applications of the SARACEN method, we look at a family of compounds in which there are  $\beta$ -donor interactions with silicon atoms. In general, silyl-substituted amines and ethers have wide angles at nitrogen and oxygen. Trisilylamines and disilylamines are almost invariably planar, and in silyl ethers angles at oxygen are typically  $125^\circ$  if there is a single substituent and  $140$ – $180^\circ$  when there are two. It was rather a surprise, therefore, to find angles much less than tetrahedral in silyl-oxygen and silyl-nitrogen compounds. Moreover, the angles proved to be extremely variable, and in studying structures of these compounds in the gas phase, examples that could fall into all of the categories covered in this article have been encountered. There have been large differences between gas-phase and crystalline phase structures; major discrepancies between theoretical and experimental results; compounds with multiple conformers; compounds with huge internal structural changes associated with conformational changes; and over and over again, molecules with low symmetry. Taken together, they illustrate the potential of the SARACEN method for elucidating a fascinating area of inorganic chemistry, which repeatedly challenged our preconceptions.

The story begins with silanes in which one, two, three or all four hydrogen atoms were replaced by *N,N*-dimethylaminoxy groups, studied by X-ray diffraction of single crystals. The SiON angles in  $Si(ONMe_2)_4$  were  $107$ – $110^\circ$ , approximately tetrahedral, but nevertheless  $16^\circ$  smaller than in the isoelectronic  $Si(OCHMe_2)_4$ .<sup>47</sup> The angle compression was attributed to  $\beta$ -donor interactions from nitrogen atoms to silicon. The corresponding angles in  $HSi(ONMe_2)_3$  and  $H_3SiONMe_2$  were  $102$ – $105^\circ$  and  $102.6^\circ$ , respectively, but in  $H_2Si(ONMe_2)_2$  the SiON angles decreased dramatically, to  $94.2$  and  $96.2^\circ$ .<sup>47</sup> In this last compound, the two  $N \cdots Si$  linkages were directly opposite one another, with the  $N \cdots Si \cdots N$  angle  $179.7(4)^\circ$ . All these structural parameters were reproduced well by calculations at MP2/6-31G\* or MP2/6-31++G\*\* levels.

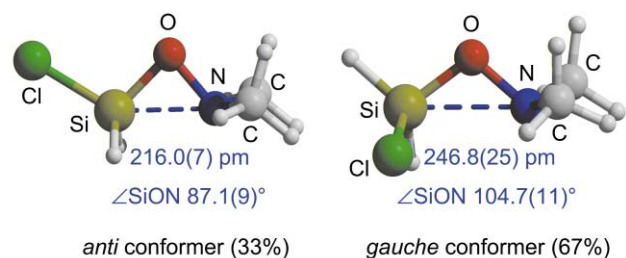
There is much less good agreement of the computed and experimental gas-phase SiON angles in  $Cl_3SiONMe_2$ ,<sup>25</sup> which was discussed earlier in the context of domination of the scattering by the heavier atoms. The GED refinement gave a value of  $105.6(8)^\circ$ , but calculations at the MP2/6-31G\* level yielded a much smaller angle of  $100.4^\circ$ . The discrepancy is even greater for the germanium analogue,  $Cl_3GeONMe_2$ , with GED and MP2/6-31G\* values of  $104.0(11)$  and  $95.6^\circ$  respectively. It is clear that care must be taken with computed structures for compounds of this kind, because the  $\beta$ -donor interactions are not easily modelled. If the chlorine substituents are replaced by

fluorine, to give  $F_3SiONMe_2$ , the interaction becomes much stronger, and the  $SiON$  angle reduces to  $94.3(9)^\circ$  in the gas phase (Fig. 19).<sup>48</sup> Calculations using the MP2 method with the fairly large 6-311G(d,p) basis set reproduce this very well, giving  $94.1^\circ$ , but the B3LYP method with the even larger 6-311++G(d,p) basis set misses by a mile, giving  $105.5^\circ$ . When the angle in the crystalline phase was found to be an astonishingly small  $77.1(1)^\circ$ , with the non-bonded  $Si \cdots N$  distances only a little larger than the sum of covalent radii, at  $196.3(1)$  pm, some explanation was needed. Optimising the structure at the MP2/6-31G\* level of theory while changing the conformation of the  $SiF_3$  group stepwise showed that this internal rotation was associated with a  $26^\circ$  variation in  $SiON$  angle, from  $86$  to  $112^\circ$ , while the barrier to rotation was only  $23$  kJ mol<sup>-1</sup>. In a situation as fluid as this, the importance of experimental structural data is self-evident.



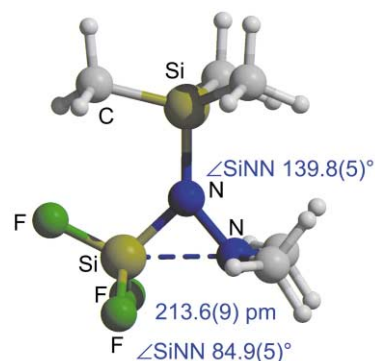
**Fig. 19** The structures of  $F_3SiONMe_2$  (a) in the crystalline phase and (b) in the gas phase.

If the silyl group is asymmetric, as in  $ClH_2SiONMe_2$ , it is possible to have two conformers. Both are present in significant amounts in the gas phase,<sup>49</sup> and although their energies are almost equal they are very different in their geometries. (Fig. 20) The interpretation of the GED data is therefore much more complex, because it is impossible to assume similar (or equal) geometrical parameters for the two conformers. Moreover, its geometry in the solid state is different again, so it is even more difficult to say to what extent calculations are reliable. As with the related compounds, the occurrence of a  $\beta$ -donor secondary bond made these calculations of the structure difficult, as one could not be sure about the quality by comparison with related cases. As with  $F_3SiONMe_2$ , there was an enormous variation in the  $SiON$  angle, this time over  $30^\circ$ , as the silyl group was rotated. The angles for the two minima on the potential energy curve were  $91.6^\circ$  when the  $Si-Cl$  bond was *anti* to the  $O-N$  bond, and  $104.5^\circ$  for the *gauche* conformer (MP2/6-311G\*\*), in reasonable agreement with the experimental GED values of  $87.1(9)$  and  $104.7(11)^\circ$ . The angle in the crystal was again much smaller, at  $79.7(1)^\circ$ . The SARACEN refinement was carried out in a way that left the important  $Si-O-N$  and  $Cl-Si-O$  angles for both conformers free to refine and not subject to restraints. Only the less important parameters were restrained as necessary, as were amplitudes of vibration. In the end 100 parameters were refined. It turned out that the refined geometries did not exactly match the calculated structures, but they were close enough to justify the use of the restraints and their assigned uncertainties. This was the crucial experiment for getting more detailed insight into this type of weak interaction over two classical bonds.



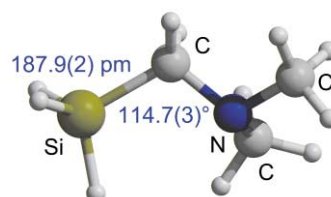
**Fig. 20** The structures of the two conformers of  $ClH_2SiONMe_2$ .

Does the type of secondary interaction seen in these *O*-silyl hydroxylamines depend on the presence of oxygen as the linking atom? No. We have already seen the example of  $Me_2NN(SiH_2CH_2)_2$ ,<sup>21</sup> in which there is a close link between the nitrogen atom of the dimethylamino group and one of the silicon atoms, with the consequence that the two  $SiNN$  angles differ by as much as  $12^\circ$ . In general, in hydrazines with one silyl substituent the  $SiNN$  angle is around  $105^\circ$ , whereas  $120^\circ$  would be expected in the absence of any specific  $Si \cdots N$  interaction. If there are two silyl groups on one of the nitrogen atoms, only one of them can interact with the  $\beta$  nitrogen, so two very different  $SiNN$  angles are found. In  $(Me_3Si)(F_3Si)NNMe_2$  (Fig. 21) the difference is amazingly more than  $50^\circ$ .<sup>50</sup> Such molecules are thus forced to have low symmetry, and so the SARACEN method is essential.



**Fig. 21** The structure of  $(Me_3Si)(F_3Si)NNMe_2$  in the gas phase.

In contrast, when the linking atom between  $Si$  and  $N$  is changed from oxygen or nitrogen to carbon, the  $\beta$  interaction is not observed. Indeed, in  $H_3SiCH_2NMe_2$  the  $SiCN$  angle is widened with respect to the tetrahedral reference value.<sup>51</sup> This compound was studied as the simplest compound containing an  $SiCN$  unit, because interaction between nitrogen and silicon had been postulated to account for its unusual reactivity and reduced nitrogen basicity.<sup>52</sup> The refinement procedure involved the use of restraints for differences between similar distances, while allowing the core angles to refine freely. However, in this case the structure obtained in the SARACEN refinement was somewhat different from the geometry predicted *ab initio*. (Fig. 22) The SARACEN values for the torsion angles  $SiCNC$  are  $62.4(12)$  and  $-173.7(14)^\circ$ , in much closer accord with the ideal staggered geometry ( $60$  and  $-180^\circ$ ) that chemical intuition would lead one to expect than the *ab initio* values of  $82.7$  and  $-156.5^\circ$ . An explanation can be found in the shallow potential for torsional movement, and so the GED data represent the thermally averaged structure, whereas the calculated data correspond to the hypothetical equilibrium geometry without any real physical counterpart.



**Fig. 22** The structure of  $H_3SiCH_2NMe_2$  in the gas phase.

## Conclusions

### Theory or experiment – or both?

After having heard so much about using theoretical information to support experiments one might be tempted to ask whether it is necessary to continue doing GED at all. Indeed

there are some molecules that are not worth studying experimentally, because that would not give a better answer than calculations, but would be far more expensive to do. Simple hydrocarbons are nowadays predicted so accurately, even from computationally cheap force-field programs, that any experimental effort to determine the gas-phase structure seems unjustified. However, in quite a few cases the data show that experiments give results that differ from theoretical predictions to some extent, sometimes very significantly. This can be true even when part of the information necessary to determine the complete structure stems from theory, such as when valuable structural information is released from the experimental data only after restraining parts of the geometry to be refined. It is important to continue to use experimental data to define the parts of structures that are not reliably computed, and to apply restraints to more predictable parts: to use each method for the purposes for which it is best suited.

As we have seen in the examples discussed in this perspective, computational methods are not consistently reliable for several types of compound. Amongst these are transition metal complexes. *Ab initio* calculations for ferrocenes have to be taken to an extraordinarily high (and expensive) level to reproduce iron-carbon distances accurately.<sup>53</sup> Some Density Functional Theory functionals do the job much more efficiently, but one has to choose the functional carefully, and what is best in one situation is not necessarily the best in another. The methods have to be calibrated – and that requires accurate experiments. Another example of the inadequacies of both *ab initio* and DFT methods is in the calculation of the V=O distance in VO(NO<sub>3</sub>)<sub>3</sub>.<sup>22</sup> Similarly, distances between atoms of second-row and first-row elements are poorly reproduced, even with the latest functionals. Indeed, there are so many parameters used in some of these functionals, that the computational method could be described as semi-empirical, which seems to lose the point of the purely theoretical nature of calculations. The work on the sulfur imides,<sup>23</sup> where experimental S=N distances differ substantially from calculated values, illustrates this point. Other problems with computational methods arise with relatively weak intramolecular interactions, such as the β-donor interactions in silyl hydroxylamines discussed above, and in predicting relative energies (and therefore populations) of conformers.

We have also seen that even weak intermolecular contacts or dipole fields can lead to large distortions of the geometry of a free molecule when it is incorporated into the solid phase. As such effects are hard to predict and even more difficult to treat correctly by means of calculations, single crystal X-ray crystallography cannot be used for the purpose of comparison or calibration in these cases. Gas-phase methods are the only experimental way to guide the development of theory in this respect.

Above all, theory can never supplant experiment. For how can the theory be validated? There is, and always will be, a need for accurate structural data, for the widest range of structural types. The only method applicable to large and complex gaseous molecules is electron diffraction. By itself, it cannot cope with such molecules, but when supported by computed restraints it can be used to study almost any sufficiently volatile compound.

#### The SARACEN method: a great leap sideways – or forwards?

What has the introduction of the SARACEN method achieved?

- Structures are more reliable. Structural analyses are much less dependent on assumptions, such as overall or local symmetry. It is much easier to explore a wide range of possible structures, or to identify possible conformers, and include several of them in the analysis. However, great care is needed; it is too easy to presume that the theoretical part of the structure is absolutely correct, and blindly follow its suggestions. It is

only when theory and experiment are broadly in accordance at that they can support each other, and both must be treated as objectively as possible.

- Structures are more accurate. Because there need be no unsupported assumptions about symmetry, and because all parameters can be allowed to refine, the possibility of introducing systematic errors can almost be eliminated.

- Estimated errors are realistic. If a refining parameter is correlated with one that is fixed, its estimated standard deviation must necessarily be too small. There is no way of knowing by how much it is too small. Some researchers multiply their standard deviations by two or three to allow for such correlations, as well as for other systematic errors. These factors may be too small, so that the quoted uncertainties are misleadingly optimistic; but they may also be too large, thus wasting valuable precision. With the SARACEN method, one can quote the uncertainty simply as the standard deviation, confident that it is meaningful.

- Most molecules can be studied, regardless of complexity. The limitations of electron diffraction can always be overcome by the use of sufficient restraints, whether they are based on other experimental observations or on computed parameters. Large molecules and those with low symmetry no longer present insuperable problems. All that is needed is that the compound is sufficiently volatile. Applications in sophisticated experiments, such as those involving formation of weakly associated species in expanded jets, study of short-lived species, or investigation of reaction dynamics,<sup>54</sup> are equally possible.

Electron diffraction can thus now be applied to a much wider range of compounds than ever before. By bringing together theoretical and experimental data in a single combined analysis, one does indeed get the best of both worlds.

#### References

- 1 D. Braga, F. Grepioni, K. Biradha and G. R. Desiraju, *J. Chem. Soc., Dalton Trans.*, 1996, 3925.
- 2 K. R. Leopold, M. Canagaratna and J. A. Phillips, *Acc. Chem. Res.*, 1997, **30**, 57.
- 3 (a) H<sub>3</sub>BNMe<sub>3</sub>: E. W. Hughes, *J. Am. Chem. Soc.*, 1956, **78**, 502; M. Bühl, T. Steinke, P. v. R. Schleyer and R. Boese, *Angew. Chem.*, 1991, **103**, 1179; (b) F<sub>3</sub>BNMe<sub>3</sub>, solid state: S. Geller and J. L. Hoard, *Acta Crystallogr.*, 1951, **4**, 399; (c) F<sub>3</sub>BNMe<sub>3</sub>, gas phase: M. Hargittai and I. Hargittai, *J. Mol. Struct.*, 1977, **39**, 79.
- 4 (a) Solid state: L. Parkanyi, P. Hencsei, L. Bihatsi and T. Müller, *J. Organomet. Chem.*, 1984, **269**, 1; (b) gas phase: G. Forgacs, M. Kolonits and I. Hargittai, *Struct. Chem.*, 1990, **1**, 245; (c) earlier structure determination on methyl silatrane: Q. Shen and R. L. Hilderbrandt, *J. Mol. Struct.*, 1980, **64**, 257.
- 5 M. M. Gilbert, G. Gundersen and K. Hedberg, *J. Chem. Phys.*, 1972, **56**, 1691.
- 6 R. K. Bohn and S. H. Bauer, *Inorg. Chem.*, 1967, **6**, 304.
- 7 (a) Gas phase: A. Almenningen, O. Bastiansen, L. Fernholt, B. N. Cyvin, S. J. Cyvin and S. Samdal, *J. Mol. Struct.*, 1985, **128**, 59; (b) solid state: A. Hargreaves, S. Rizvi and J. Trotter, *Proc. Chem. Soc.*, 1961, 122.
- 8 (a) D. D. Danielson, L. Hedberg, K. Hedberg, K. Hagen and M. Traetteberg, *J. Phys. Chem.*, 1995, **99**, 9374; (b) K. Hagen and K. Hedberg, *J. Am. Chem. Soc.*, 1973, **95**, 1003.
- 9 K. Kuchitsu, M. Nakata and S. Yamamoto, in *Stereochemical Applications of Gas-Phase Electron Diffraction*, ed. I. Hargittai and M. Hargittai, VCH Publishers Inc., New York, 1988, vol. 1, pp. 227–264.
- 10 D. W. H. Rankin, in *Stereochemical Applications of Gas-Phase Electron Diffraction*, ed. I. Hargittai and M. Hargittai, VCH Publishers Inc., New York, 1988, vol. 1, pp. 451–482.
- 11 A. J. Blake, P. T. Brain, H. McNab, J. Miller, C. A. Morrison, S. Parsons, D. W. H. Rankin, H. E. Robertson and B. A. Smart, *J. Phys. Chem.*, 1996, **100**, 12280.
- 12 P. T. Brain, C. A. Morrison, S. Parsons and D. W. H. Rankin, *J. Chem. Soc., Dalton Trans.*, 1996, 4589.
- 13 V. J. Klimkowski, J. D. Ewbank, C. van Alsenoy and L. Schäfer, *J. Am. Chem. Soc.*, 1982, **104**, 1476.
- 14 L. S. Bartell, D. J. Romenenko and T. C. Wong, in *Molecular Structure by Diffraction Methods*, Specialist Periodical Reports, The Chemical Society, 1975, vol. 3, p. 72.

- 15 R. F. Curl, *J. Comput. Phys.*, 1970, **6**, 367.
- 16 B. A. Smart, L. E. Griffiths, C. R. Pulham, H. E. Robertson, N. W. Mitzel and D. W. H. Rankin, *J. Chem. Soc., Dalton Trans.*, 1997, 1565.
- 17 M. A. Fox, R. Greatrex, A. Nikrahi, P. T. Brain, M. J. Picton, D. W. H. Rankin, H. E. Robertson, M. Bühl, L. Li and R. A. Beaudet, *Inorg. Chem.*, 1998, **37**, 2166.
- 18 B. A. Smart and D. W. H. Rankin, *Chem. Commun.*, 1997, 231.
- 19 N. W. Mitzel, B. A. Smart, K.-H. Dreihäupl, D. W. H. Rankin and H. Schmidbaur, *J. Am. Chem. Soc.*, 1996, **118**, 12673.
- 20 A. H. Cowley, M. J. S. Dewar, D. W. Goodman and J. R. Schweiger, *J. Am. Chem. Soc.*, 1973, **95**, 65; M. F. Lappert, J. B. Pedley, B. T. Wilkins, O. Stelzer and E. Unger, *J. Chem. Soc., Dalton Trans.*, 1975, 1207; J. H. Hargis and S. D. Worley, *Inorg. Chem.*, 1977, **16**, 1686.
- 21 N. W. Mitzel, H. Schmidbaur, D. W. H. Rankin, B. A. Smart, M. Hofmann and P. v. R. Schleyer, *Inorg. Chem.*, 1997, **36**, 4360.
- 22 B. A. Smart, H. E. Robertson, D. W. H. Rankin, E. G. Hope and C. J. Marsden, *J. Chem. Soc., Dalton Trans.*, 1999, 473.
- 23 S. L. Hinchley, P. Trickey, H. E. Robertson, B. A. Smart, D. W. H. Rankin, D. Leusser, B. Walford, D. Stalke, M. Bühl and S. J. Obrey, *J. Chem. Soc., Dalton Trans.*, 2002, 4607.
- 24 P. T. Brain, M. Bühl, M. A. Fox, R. Greatrex, D. Hnyk, A. Nikrahi, D. W. H. Rankin and H. E. Robertson, *J. Mol. Struct.*, 1998, **445**, 319.
- 25 U. Losehand, N. W. Mitzel and D. W. H. Rankin, *J. Chem. Soc., Dalton Trans.*, 1999, 4291.
- 26 B. F. Johnston, D. W. H. Rankin, H. E. Robertson, R. P. Hughes and J. R. Lompfrey, *Organometallics*, 2002, **21**, 4840.
- 27 W.-W. du Mont, L. Müller, R. Martens, P. M. Papathomas, B. A. Smart, H. E. Robertson and D. W. H. Rankin, *Eur. J. Inorg. Chem.*, 1999, 1381.
- 28 S. L. Hinchley, H. E. Robertson, D. W. H. Rankin and W.-W. du Mont, *J. Chem. Soc., Dalton Trans.*, 2002, 3787.
- 29 S. L. Hinchley, B. A. Smart, C. A. Morrison, H. E. Robertson, D. W. H. Rankin, R. Zink and K. Hassler, *J. Chem. Soc., Dalton Trans.*, 1999, 2303.
- 30 S. L. Hinchley, C. A. Morrison, D. W. H. Rankin, C. L. B. Macdonald, R. J. Wiacek, A. H. Cowley, M. F. Lappert, G. Gundersen, J. A. C. Clyburne and P. P. Power, *Chem. Commun.*, 2000, 2045; S. L. Hinchley, C. A. Morrison, D. W. H. Rankin, C. L. B. Macdonald, R. J. Wiacek, A. Voigt, A. H. Cowley, M. F. Lappert, G. Gundersen, J. A. C. Clyburne and P. P. Power, *J. Am. Chem. Soc.*, 2001, **123**, 9045.
- 31 W. M. Cleaver, M. Späth, D. Hnyk, G. McMurdo, M. B. Power, M. Stuke, D. W. H. Rankin and A. R. Barron, *Organometallics*, 1995, **14**, 690; M. B. Power, A. R. Barron, D. Hnyk, H. E. Robertson and D. W. H. Rankin, *Adv. Mater. Opt. Electron.*, 1995, **5**, 177.
- 32 K. Hedberg, L. Hedberg, D. S. Bethune, C. A. Brown, M. S. de Vries and R. D. Johnson, *Science*, 1991, **254**, 410.
- 33 P. T. Brain, M. Bühl, H. E. Robertson, A. D. Jackson, P. D. Lickiss, D. MacKerracher, D. W. H. Rankin, D. Shah and W. Thiel, *J. Chem. Soc., Dalton Trans.*, 1998, 545.
- 34 D. Hnyk, M. Bühl, P. T. Brain, H. E. Robertson and D. W. H. Rankin, *J. Am. Chem. Soc.*, 2002, **124**, 8078.
- 35 K. Hedberg, L. Hedberg, M. Bühl, D. S. Bethune, C. A. Brown and R. D. Johnson, *J. Am. Chem. Soc.*, 1997, **119**, 5314.
- 36 O. Hassel, *Quart. Rev.*, 1953, **7**, 221.
- 37 N. W. Mitzel, B. A. Smart, A. J. Blake, H. E. Robertson and D. W. H. Rankin, *J. Phys. Chem.*, 1996, **100**, 9339.
- 38 C. A. Morrison, D. W. H. Rankin, H. E. Robertson, C. Eaborn, A. Farook, P. B. Hitchcock and J. D. Smith, *J. Chem. Soc., Dalton Trans.*, 2000, 4312.
- 39 C. A. Morrison, D. W. H. Rankin, H. E. Robertson, P. D. Lickiss and P. C. Masangane, *J. Chem. Soc., Dalton Trans.*, 1999, 2293.
- 40 B. A. Smart, P. T. Brain, H. E. Robertson and D. W. H. Rankin, *Inorg. Chem.*, 1998, **37**, 2687.
- 41 K. B. Borisenko and D. W. H. Rankin, *J. Chem. Soc., Dalton Trans.*, 2002, 3135.
- 42 A. W. Cordes, M. Hojo, H. Koenig, M. C. Noble, R. T. Oakley and W. T. Pennington, *Inorg. Chem.*, 1986, **25**, 1137.
- 43 A. V. Zibarev, Y. V. Gatilov and A. O. Miller, *Polyhedron*, 1992, **11**, 1137.
- 44 F. Blockhuys, S. L. Hinchley, A. Yu. Marakov, Yu. V. Gatilov, A. V. Zibarev, J. D. Woollins and D. W. H. Rankin, *Chem. Eur. J.*, 2001, **7**, 3592.
- 45 N. W. Mitzel, U. Losehand and A. Richardson, *Organometallics*, 1999, **18**, 2610.
- 46 C. A. Morrison, S. F. Bone, D. W. H. Rankin, H. E. Robertson, S. Parsons, R. A. Coxall, S. Fraser, J. A. S. Howell, P. C. Yates and N. Fey, *Organometallics*, 2001, **20**, 2309.
- 47 N. W. Mitzel, A. J. Blake and D. W. H. Rankin, *J. Am. Chem. Soc.*, 1997, **119**, 4143.
- 48 N. W. Mitzel, U. Losehand, A. Wu, D. Cremer and D. W. H. Rankin, *J. Am. Chem. Soc.*, 2000, **112**, 4471.
- 49 N. W. Mitzel and U. Losehand, *J. Am. Chem. Soc.*, 1998, **120**, 7320.
- 50 L. J. McLachlan, K. Vojinovic, N. W. Mitzel and D. W. H. Rankin, unpublished observations.
- 51 N. W. Mitzel, C. Kiener and D. W. H. Rankin, *Organometallics*, 1999, **18**, 3437.
- 52 (a) M. G. Vorkonkov, V. P. Feshin, V. F. Mironov, S. A. Mikahilyants and T. K. Gar, *Zh. Obshch. Khim.*, 1971, **41**, 2211; (b) M. G. Voronkov, T. V. Kashik, E. Ya. Lukevits, E. S. Deriglazova, A. E. Pestunovich and R. Y. Moskovich, *Zh. Obshch. Khim.*, 1974, **44**, 778.
- 53 H. Koch and P. Jorgenson, *J. Chem. Phys.*, 1996, **104**, 9528.
- 54 J. C. Williamson, J. Cao, H. Ihee, H. Frey and A. H. Zewail, *Nature*, 1997, **386**, 159.