

## Extended electron distributions applied to the molecular mechanics of some intermolecular interactions. II. Organic complexes

J.G. Vinter

*Cambridge Centre for Molecular Recognition, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.*

Received 22 April 1996

Accepted 11 June 1996

**Keywords:** Molecular mechanics; Atomic charge; Intermolecular interaction; Extended electron distribution; Atom-centred charges

---

### Summary

Extended electron distributions (XEDs) have been used to simulate the formation of complexes by intermolecular interaction via: (i) aromatic stacking; and (ii) hydrogen bonding. The results qualitatively reproduce experimental observations. In contrast, atom-centred partial charges fail to reproduce highly hydrogen-bonded systems, but make little difference in cases where interactions are driven largely by van der Waals forces. The dielectric constant used in the Coulombic term has been shown to be significant in defining the type and properties of these interactions when XEDs are employed. Some consideration has been given to solvation and entropy effects.

---

### Introduction

Atom-centred partial charges (ACCs) in the Coulombic term of molecular mechanics can be an approximation which sometimes leads to poor results. In the previous paper in this series, extended electron distributions (XEDs), originally stimulated by ideas from Hunter and Sanders [1], were incorporated into a larger molecular modelling framework [2]. The new XED force-fields were validated using a number of small-molecule complexes. The results suggested that the use of the ACCs can often lead to poor simulation, especially when intermolecular interactions are driven predominantly by Coulombic forces rather than dispersive enthalpy.

In this paper, further tests of the XED approach are reported on three larger organic nonbonded complexes. One of these is predominantly held together by van der Waals forces, whilst the other two examples are hydrogen-bonded.

### Methods

The methods are as described previously [2]. In addition, the rigid-docking routine was modified to allow the mobile species to begin its docking trajectory from a very small sphere around the centroid of the fixed molecule. In

this way, docking from inside a cavity could be undertaken (Fig. 1).

The docking technique (named XEDOCK) produces a map of most of the possible local minima associated with the interaction of one species (the bullet) with another (the target) from 'all' directions. Each dock is carried out separately. The final map is a conglomerate of the 259 independent docks from a new position on the docking sphere and gives a picture of the most favourable final local minimum between the target and the bullet from the particular starting point. This map therefore does not represent a picture of the build-up of a given ligand. The binary complex with the lowest energy of interaction is expected to be the experimentally favoured species. Many favoured orientations of the final binary are often created – all of which may be significant if their energies of interaction allow them to be observable at room temperature (3 kcal/mol or less above the global minimum).

In order to gain an insight into, for example, a solvent shell, the docking routine was adjusted so that after each run, the best-docked pair was used as the target for a further docking series. Thus, to create a water shell around a target, a water molecule was used as a bullet. The first docking run yielded 259 docked water molecules. The complex with the lowest energy was extracted and fed back to the docker as the target. This was repeated until

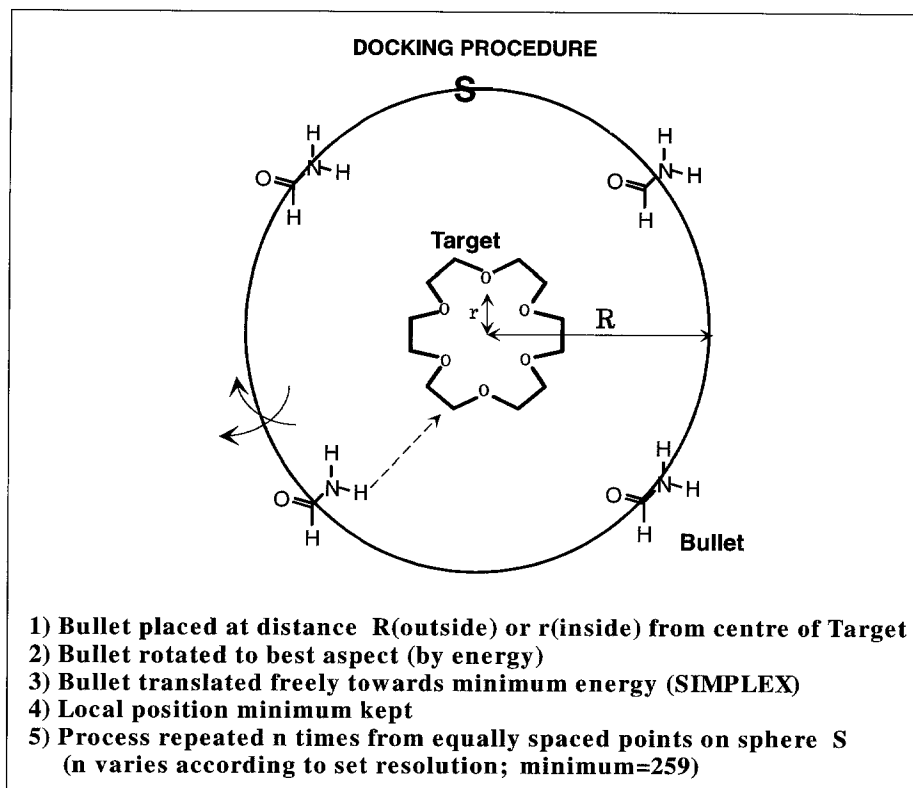


Fig. 1. Computational intermolecular docking procedure for two species.

the lowest energy of each new water complex had levelled off to a value corresponding with the average interaction energy with bulk water. The actual number of docked waters depended on the size of the solute, but 40 waters were usually sufficient to reach an average bulk-water energy of solvation of about 6 kcal/mol at a dielectric constant of unity (program named XEDSEQDOK).

All work was carried out on a Silicon Graphics Indy R4000 running IRIX v. 5.2 with 32 Mbytes of memory.

## Results and Discussion

### Stacking via van der Waals forces

The experimental results of Iverson and Lokey [3] were simulated using the docking technique. The two naphtha-

lene derivatives (Fig. 2) were cocrystallized by these workers to give a stable, stacked arrangement in water. The results of the simulation are outlined in Table 1. Four calculations at various dielectric values were carried out using ACCs and repeated using XEDs.

It may be noted from Table 1 that: (i) there is little change in energy contribution with change of dielectric constant; and (ii) no significant difference is observed between the results from ACCs and XEDs. The complexation is driven almost entirely by dispersive attraction which is why the products are water-stable. The Coulombic forces are minimal but are increased at low dielectric constants. In cases which may have more Coulombic influences contributing to their complexation (especially if, like porphyrins, they can contain metals), such second-

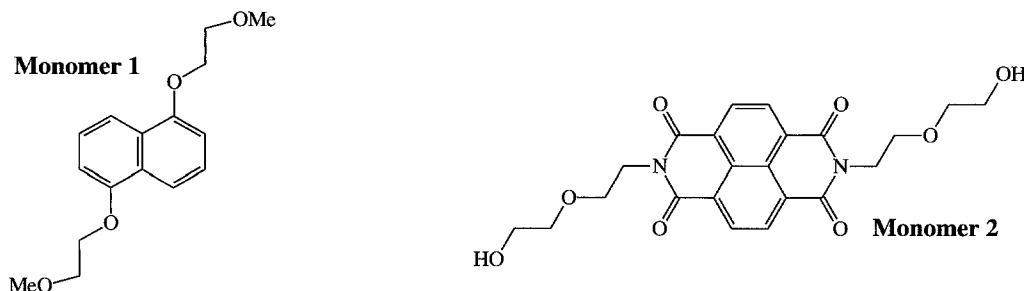


Fig. 2. The two monomers cocrystallized by Iverson and Lokey and used in this study to simulate the crystal structure of the resulting stacked complex.

TABLE 1  
RESULTS FROM DOCKING MONOMER 1 WITH MONOMER 2 AT VARIOUS DIELECTRIC VALUES

Dielectric constant (D)	Atom-centred partial charges				Extended electron distributions			
	Total energy (kcal/mol)	vdW energy (kcal/mol)	Coulomb energy (kcal/mol)	AA	Total energy (kcal/mol)	vdW energy (kcal/mol)	Coulomb energy (kcal/mol)	AA
1.0	-28.0	-23.5	-4.5	54	-27.9	-23.2	-4.7	30
4.0	-25.0	-24.6	-0.4	60	-25.0	-24.7	-0.3	59
8.0	-24.9	-24.7	-0.2	55	-24.8	-24.7	-0.1	39
20.0	-24.8	-24.7	-0.1	43	-24.8	-24.7	-0.1	18

The contributions to the total energy are the van der Waals interaction (vdW) and the electrostatic energy (Coulomb). AA is the number of arrangements accessible at room temperature. The docking procedure produces a map of local minimum orientations of the bullet on the target. The AA column records all species with greater than 0.1% mole fraction from a Boltzmann distribution at room temperature. (At room temperature, a large number indicates that many orientations can exist whilst a small number indicates a restricted set.)

ary forces are probably responsible for orientating the largely surface-bound vdW 'stickiness' which would be expected to have a low potential to sliding over each surface.

Lokey and Iverson reported a binding constant of  $130 \text{ M}^{-1}$  for sodium carboxylate derivatives of the monomers. It is reasonable to assume that a similar binding energy

will apply to the neutral monomers used in the X-ray study and the calculations reported here. Thus, a Gibbs free energy of binding of  $-2.9 \text{ kcal/mol}$  needs to be reconciled with the calculated enthalpy of binding of  $25 \text{ kcal/mol}$  (Table 1). Two processes are expected to accompany enthalpic complexation: (i) solvation; and (ii) entropy changes.

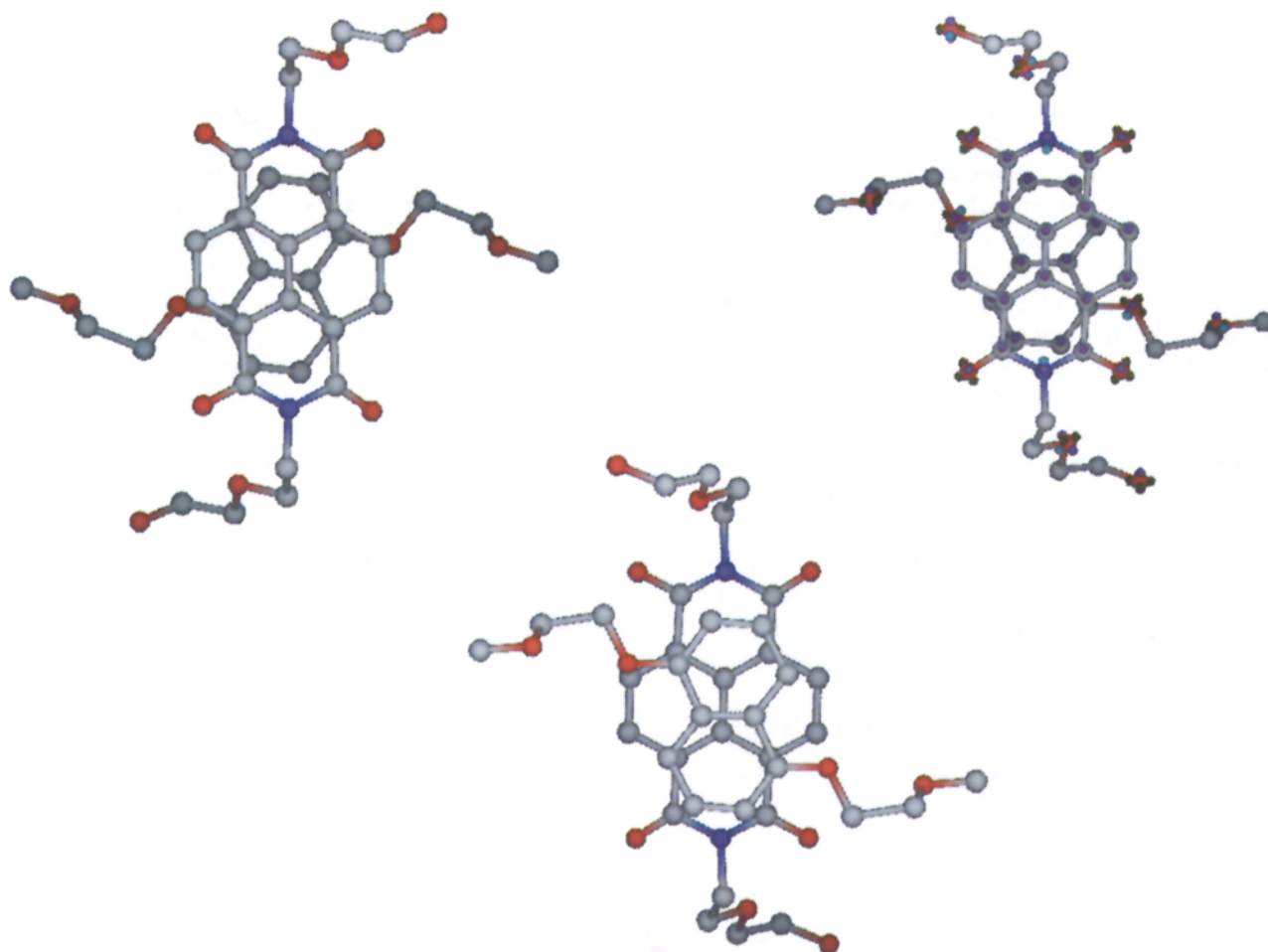


Fig. 3. The energetically most favourable complexes from the docking of Monomer 1 with Monomer 2 at a dielectric of unity. Top left: energies calculated using ACCs; top right: energies calculated using XEDs; bottom: the experimental complex from X-ray analysis.

TABLE 2  
CALCULATED ENERGIES OF SEQUENTIAL DOCKING OF WATER WITH ITSELF AND THE SPECIES DESCRIBED BY IVERSON AND LOKEY

Docking number	Water	Monomer 1	Monomer 2	Stacked complex of 1 and 2
1	-3.68	-6.51	-8.72	-9.66
2	-5.35	-6.44	-8.70	-9.30
3	-5.80	-6.22	-6.76	-7.88
4	-5.83	-7.04	-7.06	-7.73
5	-5.97	-6.14	-7.00	-7.28
6	-4.62	-7.02	-8.24	-7.03
7	-5.64	-5.84	-6.81	-6.99
8	-5.41	-6.75	-7.04	-6.83
9	-7.27	-6.71	-6.99	-6.92
10	-6.44	-7.37	-8.23	-6.72
11	-5.83	-5.96	-6.71	-7.20
12	-6.31	-7.51	-7.21	-8.46
Average	-6.22	-6.62	-7.46	-7.67

Only the first 12 out of 40 docks of water on itself are presented.

### Solvation

Sequential docking of water as described in the Methods section was used to investigate solvation patterns and any apparent adjustments to solvation during complexation. XEDs were added to all species for sequential solvation calculations.

The first stage was to characterise the behaviour of bulk water. Sequential docking of 40 water molecules on water yielded an average enthalpy of  $-6.2$  kcal/mol for the inclusion of a water molecule into bulk water. Twelve water molecules were similarly docked onto Monomer 1 and twelve onto Monomer 2. Finally, the binary complex was solvated with 12 waters (Table 2).

Computer graphical examination of the positions of each water on the monomers and the complex revealed that all the solvent waters with energies above that of bulk water ( $> 1$  kcal/mol difference) do not associate with the aromatic stackers, but instead solvate the chains. Furthermore, the average energies of solvation are close to that for bulk water. There appears to be no evidence for solvation rearrangement in thermodynamic terms on complexation of 1 and 2 to form the 1/2 complex. Water may therefore not be an important participant in the Iverson–Lokey stacking process.

### Entropy

Entropy changes on binary complexation can be approximated using the ideas of Williams et al. [4]. Table 3 outlines the entropy contributions from each species.

It is not obvious how to assess the number of rotatable bonds that are frozen out during complexation. The change in translational and rotational entropy can be calculated from a difference in complexed and monomer values:  $T\Delta S_{\text{complex}} = TS_{\text{tr1/2}} - (TS_{\text{tr1}} + TS_{\text{tr2}}) = 18.5 - 34.3 = -15.8$  kcal/mol at 298 K. The enthalpy of formation of the complex is  $-27.9$  kcal/mol. If  $\Delta G = -2.9$  kcal/mol, then  $-2.9 + 27.9 - 15.8 = TS_{\text{bond freezing}} = +9.2$  kcal/mol.

If the freezing of a rotatable bond is worth 1.3 kcal/mol [4], the equivalent of approximately 7 bonds are frozen out of a total of 22 on complexation of Monomer 1 and 2. This would seem reasonable from the examination of the complex and leads to the possible conclusion that entropy changes, and not solvation, result in the overall reconciliation of experimental free energy and calculated enthalpy.

XEDs represent no advantage to the study of the Iverson–Lokey complex. In both cases, the crystal structure was reproduced reasonably well by the lowest energy complex ('global minimum') from each docking experi-

TABLE 3  
ENTROPY ESTIMATIONS FOR THE TWO MONOMERS DESCRIBED BY IVERSON AND LOKEY AND THE COMPLEX FORMED FROM THEM

Structure	MW	T (K)	Rotatable bonds	$TS_{\text{bond rot}}$ (kcal/mol)	$TS_{\text{trans}}$ (kcal/mol)	$TS_{\text{rot}}$ (kcal/mol)
Monomer 1	276.3	298	10	13.0	6.4	10.3
Monomer 2	442.4	298	12	15.6	6.5	11.1
1/2 Stacked complex	718.7	298	22	28.6	6.7	11.8

Loss of entropy for bond rotation is for all rotatable bonds. Entropy loss for translation and molecular rotation are for rigid bodies. (The freezing of a rotatable bond is worth 1.3 kcal/mol. Translational entropy is a function of molecular weight and is calculated from the Sackur–Tetrode equation. Rotational entropy is related to the moments of inertia [4].)

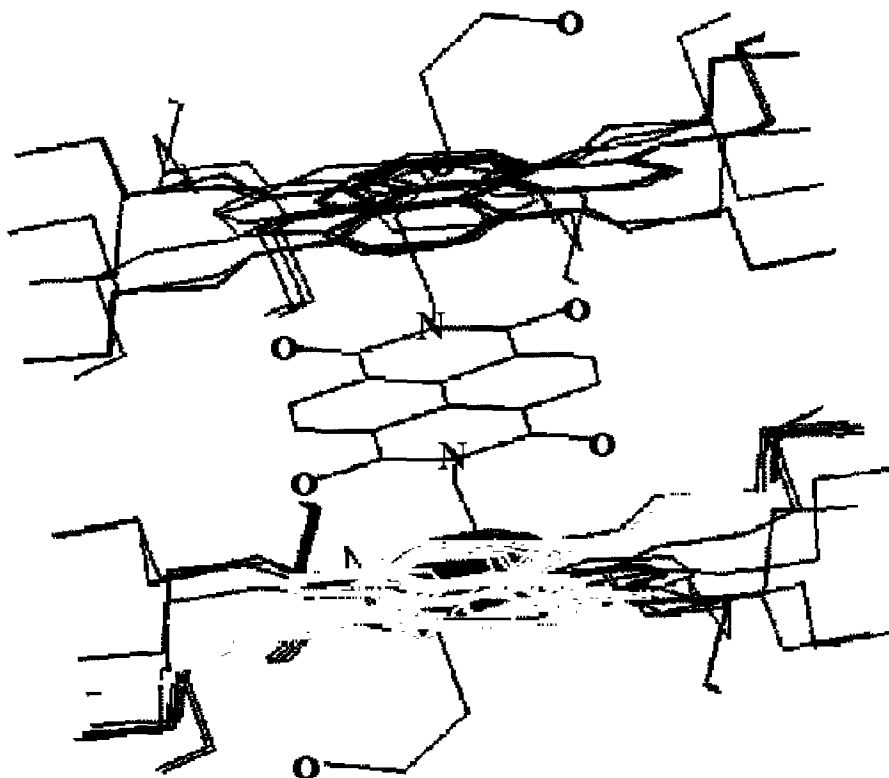


Fig. 4. A 'screen dump' of 31 orientations of the docking of Monomer 1 onto Monomer 2 all of which are within 3 kcal/mol of the global minimum and therefore accessible at room temperature.

ment. The distance between the stacked pair was 3.33 Å for the X-ray structure and 3.29 Å for the best dock. The alignment (Fig. 3) of the monomers in the crystal was slightly skewed, presumably due to crystal-packing forces,

whilst the best simulation was essentially symmetric. Over a 3-kcal/mol range from this global minimum, a variation of positions was observed covering 31 docked structures (Fig. 4).

TABLE 4  
CALCULATED ENERGIES OF SEQUENTIAL DOCKING OF THE MONOMER LEADING TO THE 'TENNIS BALL' AND 'SUPER TENNIS BALL'

Simulation	Dielectric constant (D)	Atom-centred partial charges				Extended electron distributions			
		Total energy (kcal/mol)	vdW energy (kcal/mol)	Coulomb energy (kcal/mol)	AA	Total energy (kcal/mol)	vdW energy (kcal/mol)	Coulomb energy (kcal/mol)	AA
'Tennis ball'	1.0	-31.7	-16.9	-14.8	91 <sup>a</sup>	-51.6	-2.8	-48.8	6 <sup>b</sup>
	2.0					-30.7	-16.9	-13.8	48 <sup>a</sup>
	3.0					-26.1	-17.1	-9.0	55 <sup>a</sup>
	4.0					-23.9	-17.1	-6.7	38 <sup>a</sup>
	8.0					-20.5	-17.2	-3.3	45 <sup>a</sup>
	20.0					-19.7	-19.6	-0.1	40 <sup>a</sup>
Fully relaxed minimisation	1.0					-72.3	-2.5	-69.9	
'Super tennis ball'	1.0	-38.5	-29.7	-8.8	4 <sup>a</sup>	-45.3	-4.0	-41.3	17 <sup>b</sup>
	0.5					-98.0	14.1	-112.1	2 <sup>b</sup>
Fully relaxed minimisation	1.0					-100.0	-3.5	-96.4	

For the structures, see Fig. 5. The contributions to the total energy are the van der Waals interaction (vdW) and the electrostatic energy (Coulomb). AA is the number of arrangements accessible at room temperature. The docking procedure produces a map of local minimum orientations of the bullet on the target. The AA column records all species with greater than 0.1% mole fractions from a Boltzmann distribution at room temperature. (At room temperature, a large number indicates that many orientations can exist whilst a small number indicates a restricted set.)

<sup>a</sup> No 'tennis ball' complex formed.

<sup>b</sup> 'Tennis ball' complex formed.

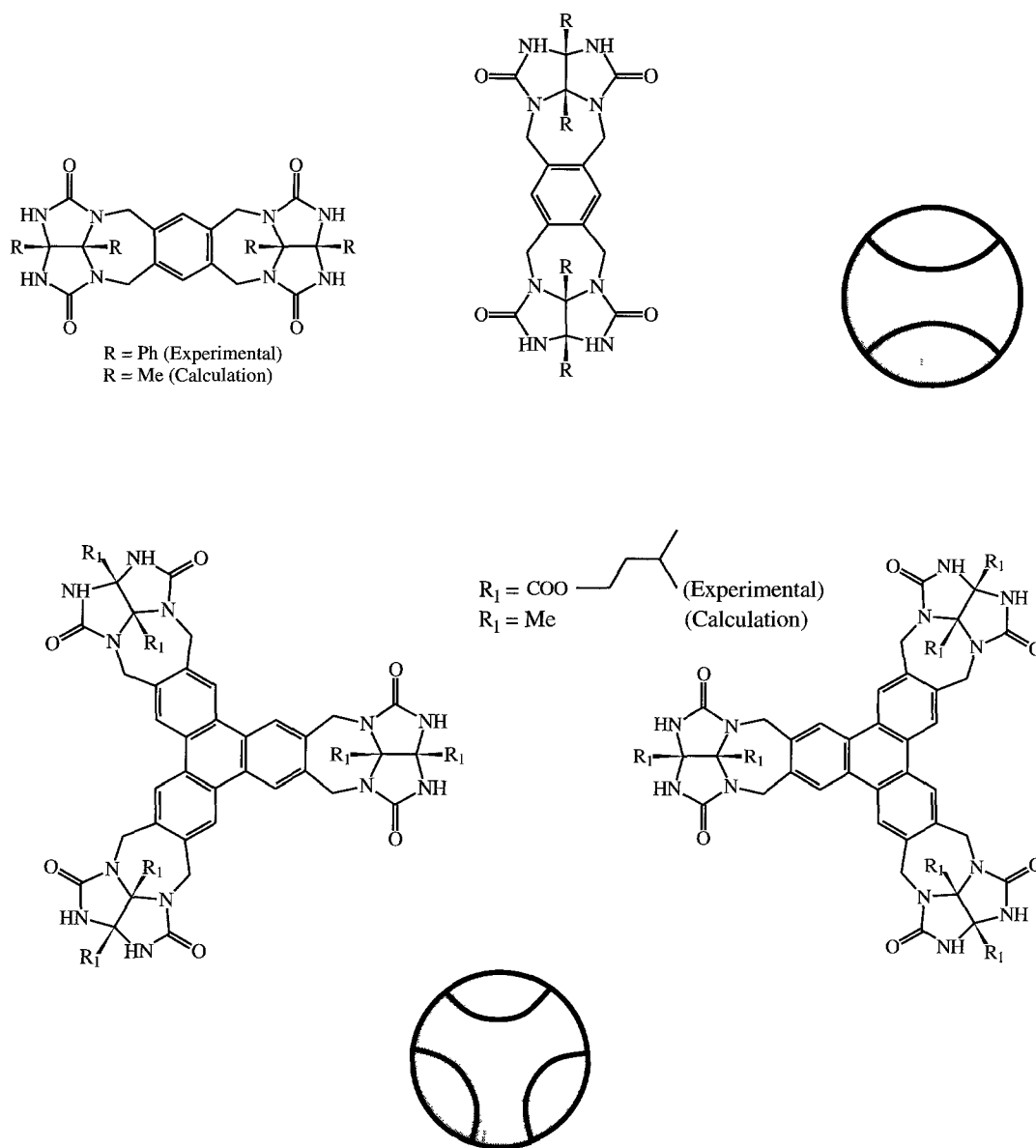


Fig. 5. Scheme showing the di-glycoluril and tri-glycoluril species used to create the 'tennis ball' (top) and 'super tennis ball' (below) [6] (see Table 4).

It would appear that a variety of linkers could be used to produce Iverson's 'aedamers' without loss of binding between the stacking pairs. By adding polar atoms into the aromatic stack, it may be possible to increase flexibility of orientation allowing new arrangements to appear and widening further the spectrum of linkers which could be used.

#### Hydrogen-bond formation

The glycoluril monomers reported by Rebek et al. [5,6] and shown in Fig. 5, represent a suitable example of complex formation through H-bonding. The dimeric complexes are formed in chloroform and retain their integrity up to 20% DMSO. The first of these, the 'tennis ball' [5], is held together by eight hydrogen bonds. The 'super tennis ball' [6] uses 12 H-bonds for stability. Mod-

ified monomers with methyl replacements (Fig. 5) were docked as before and the results are recorded in Table 4.

Even at a dielectric constant of unity, no complex from either monomer was formed by the molecular mechanics docking routines using ACCs. Any further docking at high dielectric constants would not give a different result, as the Coulombic contribution is attenuated with increased dielectric constant (Table 4). Breakdown of the symmetric complex formed with XEDs, occurred between a dielectric constant of 1 and 2 for both 'tennis balls'. A contrived dielectric constant of 0.5 enhanced the effects of hydrogen bonding and produced a high degree of  $D_{3d}$  complex symmetry – as can be seen from the last column of Table 4, which records that only two arrangements were collected within the accessible Boltzmann population.

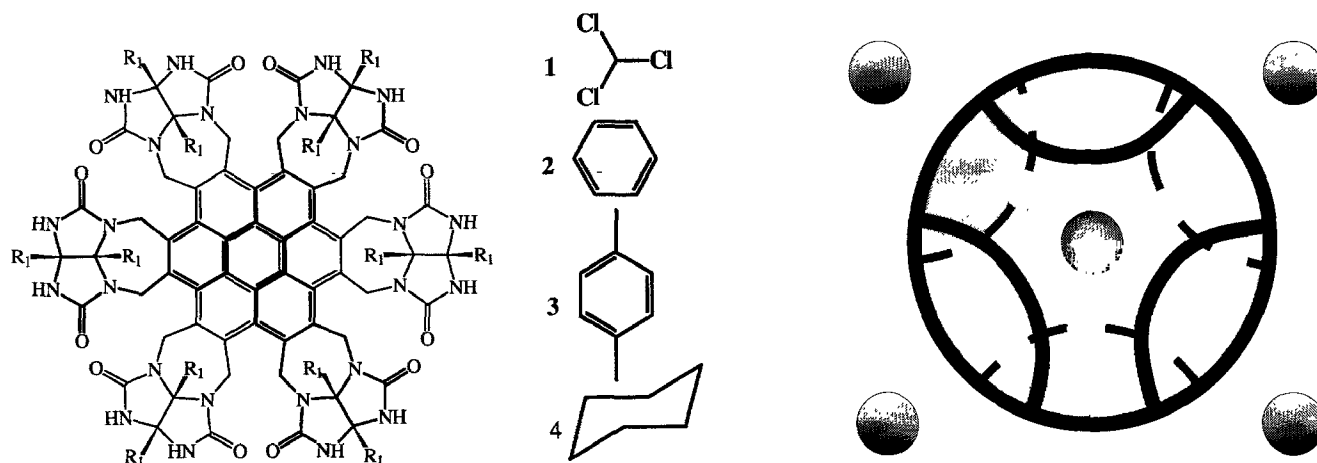


Fig. 6. Each solvent species was docked with the 'super tennis ball' as described by Rebek (at its fully relaxed minimum) from both outside and inside the ball (see Methods). Two separate docking patterns were generated, each emphasising the binding importance of the inner and outer surfaces of the ball (see Fig. 7).

No experimental binding data for the formations of these hydrogen-bonded complexes have been reported. The enthalpy contributions from solvation could be calculated on the basis that most of the solvent from the inner faces of the two reacting monomers would need to be swept out on complexation. These calculations, which would follow the same regime as for water solvation on the Iverson–Lokey complex, have not been done because solvent binding enthalpies are small and occlusion can take place (see below). Translational and rotational entropy contributions can be assessed as above assuming no entropy loss due to bond freezing; the formation of the 'tennis ball' losing 16.2 kcal/mol of entropy and the 'super tennis ball' losing 17.9 kcal/mol at 298 K. When subtracted from the fully relaxed enthalpies of formation and assuming small solvent rearrangement contributions,

the average binding enthalpy for each hydrogen bond on the 'tennis ball' is 7.0 kcal/mol and for the 'super tennis ball' 6.8 kcal/mol. These are consistent and reasonable values.

When Rebek [6] observed that titration of the chloroform solution of the 'super tennis ball' with benzene generated a new set of NMR peaks, he concluded that chloroform, occluded during complex formation, was being exchanged for benzene. Further experiments showed that *p*-xylene could be displaced by cyclohexane but over a much longer period of time.

It would be a further validation of the method described here, to attempt to simulate these observations. This process involves a sequential two-stage dock; firstly, to form the ball and take the best-docked structure. Secondly, to dock each solvent with the ball (in the same way as

TABLE 5  
RESULTS FROM THE DOCKING OF VARIOUS SOLVENTS WITH THE 'SUPER TENNIS BALL'

Docking start position	Solvent	Extended electron distributions <sup>a</sup>				# Molecules passing through ball from a total of 259 start positions
		Total energy (kcal/mol)	vdW energy (kcal/mol)	Coulomb energy (kcal/mol)	AA	
Outside	Chloroform	−10.8				4 entries
	Benzene	−12.4	−5.1	−7.3	28	no entries
	<i>p</i> -Xylene	−13.3				1 entry
	Cyclohexane	−9.2	−9.5	0.3	134	no entries
Inside	Chloroform	−17.1	−15.5	−1.6	151	0 outside
	Benzene	−21.7	−20.4	−1.3	41	215 outside
	<i>p</i> -Xylene	−26.8	−26.5	−0.3	3	256 outside
	Cyclohexane	−24.5	−24.7	0.2	31	221 outside

For the structures, see Fig. 6. Energies are those of the best binding position. The contributions to the total energy are the van der Waals interaction (vdW) and the electrostatic energy (Coulomb). AA is the number of arrangements accessible at room temperature. The docking procedure produces a map of local minimum orientations of the bullet on the target. The AA column records all species with greater than 0.1% mole fraction from a Boltzmann distribution at room temperature. (At room temperature, a large number indicates that many orientations can exist whilst a small number indicates a restricted set.) The energies recorded in the 1st and 3rd rows are for the best binding to the outside of the 'ball', but are not global minima because both chloroform and *p*-xylene entered the 'ball' from the outside.

<sup>a</sup> XEDs for D = 1.

the Iverson–Lokey complex was water-solvated). Since there is no provision to open and reform the hydrogen bonds of the ball in the present regime, docking small molecules from the outside of the ball would not guarantee entry during the computer simulation. Two experiments on each solvent were therefore done from outside, and from inside, using a starting sphere radius of 0.5 Å. Figure 6 and Table 5 summarise the procedure and results.

It may be inferred from Table 5 that the solvents in the list are held predominantly by van der Waals forces (except benzene). The global minimum energies generated from within the ball were used to compare with experimental and simulated events. From the calculated enthalpies, benzene would be expected to replace chloroform, but *p*-xylene retains a slight advantage over cyclohexane and would not exchange on enthalpic grounds.

Figure 7 shows the docking patterns of the various solvents from outside and inside the 'super tennis ball'. Under certain conditions, all the solvents studied here could pass through the intact cage wall of the 'super tennis ball'. The calculation allowed both chloroform and *p*-xylene to enter the ball from the outside. No chloroform molecules escaped from the capsule when docked from the inside and all 259 local minima (C1, Fig. 7) were within 2.8 kcal/mol of each other. These effects are most likely to be manifestations of the computational procedure but may reflect some chemistry. Chloroform may be small enough to exchange through the cage wall unhindered. The fact that *p*-xylene could also enter the unbroken ball whereas benzene did not, may be due to a degree of directionality bestowed on *p*-xylene by its methyl groups. They may hydrophobically help to lower the energy barrier to entry below the final binding energy (as must

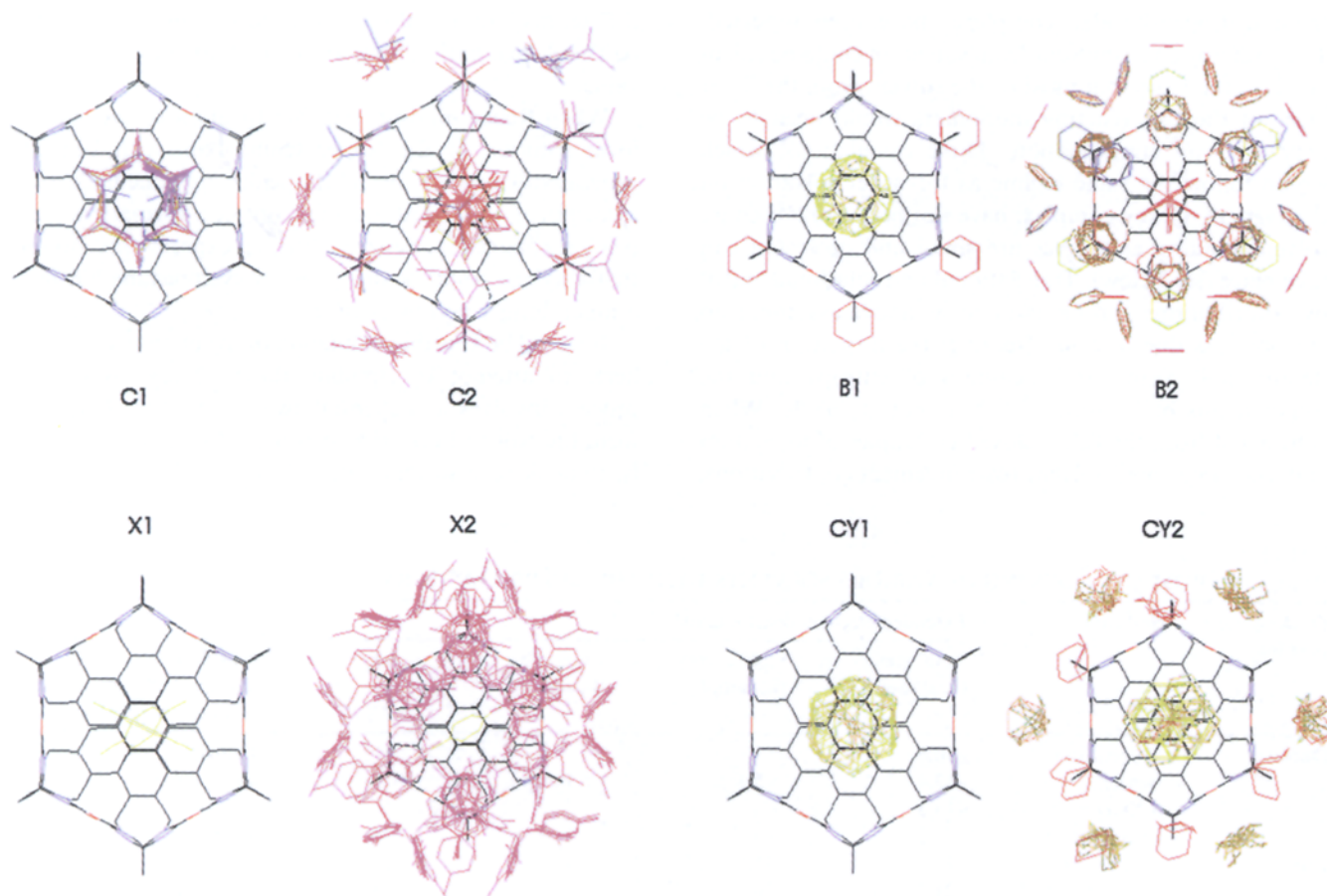


Fig. 7. Docking patterns of the various solvents onto the 'super tennis ball' from inside and outside. The colour coding progresses from yellow, through green to red and finally blue as the energy of docking increases. (High-energy local minima (blue) are not included, unless within 3 kcal/mol of the global minimum.) C1: chloroform from inside and C2: from outside, over a 3-kcal/mol range from the global minimum. B1: benzene from inside. All the yellow/green structures are inside the 'Rebek' complex and are within 3 kcal/mol of each other. Those that have escaped from the complex and have found local minima on the outer surface (red) are >6 kcal/mol above the global minimum. B2: benzene docked from outside (note that there are no benzene molecules within the cage). X1: *p*-xylene from inside showing a 3-kcal/mol energy range from the global minimum. X2: from outside. One low-energy structure can be seen inside the cage and is >13 kcal/mol below any others on the outside. CY1: cyclohexane from inside within 3 kcal/mol of the global minimum. CY2: from outside. None has entered the cage and all are >9 kcal/mol above the global minimum.



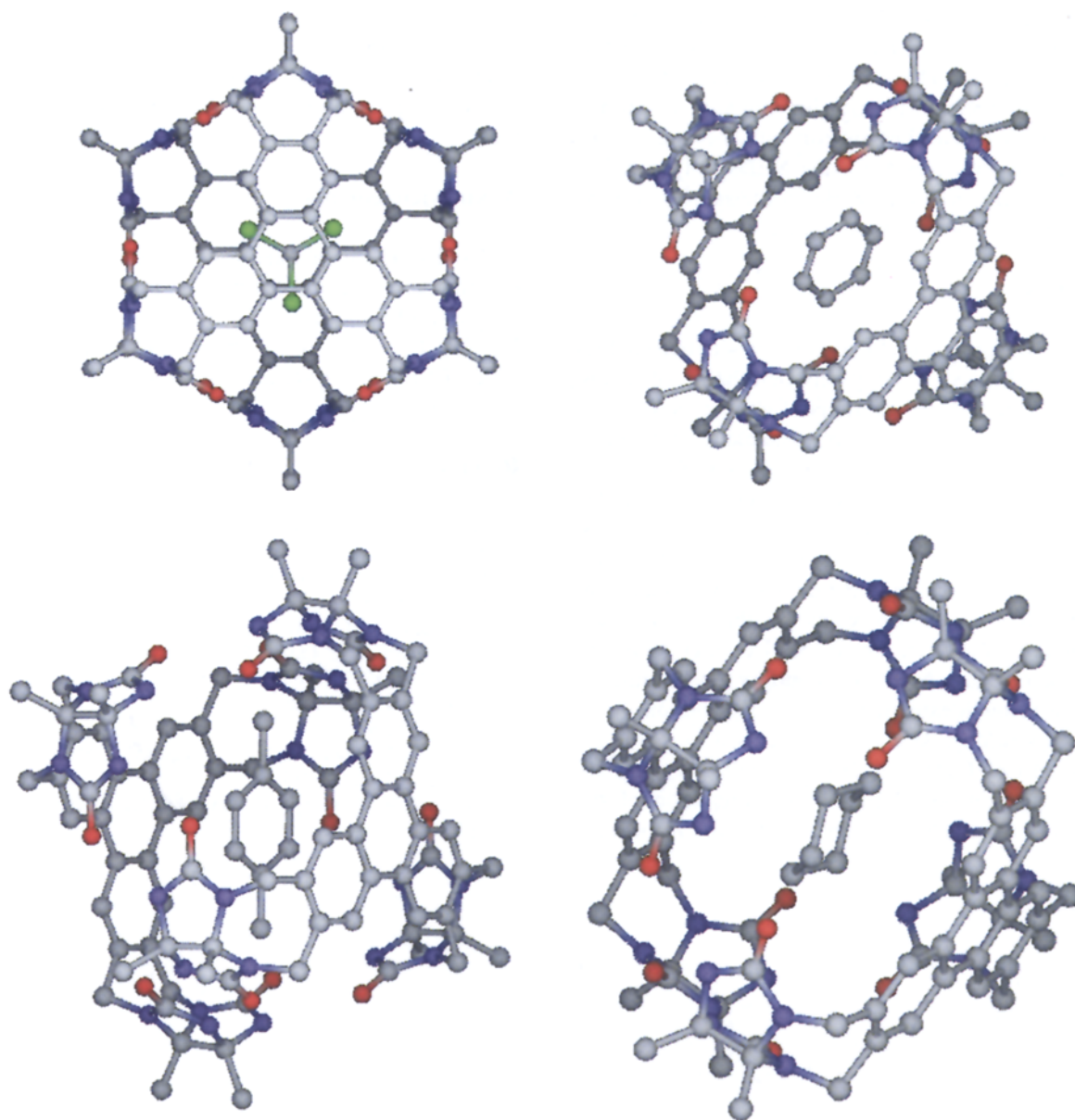


Fig. 8. The global minimum positions of each solvent with the 'super tennis ball' cage. Top left: chloroform; top right: benzene; bottom left: *p*-xylene; bottom right: cyclohexane.

happen in the simulation). Benzene may be unable to surmount this barrier because of its increased polar nature. From the experimental results [6], benzene would be expected to enter and leave the cage as readily as chloroform but the simulation allows it only to leave, not enter (B1 and B2 in Fig. 7). Benzene must therefore be very close to being able to enter, but some cage flexing may be necessary to facilitate this – a process not allowed by the simulation techniques. It is also of interest that *p*-xylene exists in only three limited and highly directional orientations (two shown in X1, Fig. 7) within the ball. Its replacement by cyclohexane under experimental conditions could be due to the entropy constraints of this limitation, particularly since the enthalpy of binding of cyclohexane

is slightly less favourable than *p*-xylene. Furthermore, the number of orientations that cyclohexane can adopt within the ball (CY1 in Fig. 7) far exceeds that of *p*-xylene and would add further to the increase in entropy of cyclohexane complexation. Figure 8 shows the global minimum positions of each solvent with the 'super tennis ball' cage.

Without XEDs, the simulation procedure cannot reproduce the experimentally observed formation of these 'tennis ball' complexes. The driving force for complex formation is the production of hydrogen bonds for which correct Coulombic interactions are necessary. We have demonstrated that atom-centred charges are inadequate for this job. The problem could be overcome by equipping the molecular mechanics armoury with some special

additional potentials which specifically force the hydrogen-bonding interaction. However, this is unnecessary if XEDs are employed. They have the advantage of a closer relationship to the nature and directionality of atomic orbitals. The same can be said for 'out-of-plane' potentials which are not necessary using XEDs;  $\pi$ -type XEDs are forced to align where possible, in the same way as  $\pi$ -orbitals align in quantum mechanics.

## Conclusions

Although the general usefulness of atom-centred charges remains unchallenged, it is apparent that the use of XEDs, or similarly extended descriptions of charge, may be necessary to better describe electrostatically dominated processes. We have demonstrated that using XEDs does not prejudice the results obtained in cases where complex formation is essentially driven by van der Waals forces. Thus, similar results are obtained using both XEDs and ACCs to simulate Iverson and Lokey's experimental observations. However, it also seems possible that  $\pi$ -stacking, although dominated by dispersive interactions at close quarters, may derive a secondary fine tuning to orientation through extended electron distribution.

Most importantly, ACCs cannot be used to reproduce the formation of Rebek's hydrogen-bonded 'tennis ball' complexes. In contrast, XEDs allow simulation of the complex formation and also account for the incorpora-

tion and exchange of solvent into the molecular complex. These conclusions are of importance in view of the common use of molecular mechanics calculations to model ligand-receptor interactions that occur in biological systems.

## Acknowledgements

The EPSRC is thanked for its financial support. Special thanks are due to Prof. B.L. Iverson, Dr. R. Scott Lokey and Prof. J. Rebek for providing unpublished data. Dr. T.I. Oprea (LANL) and Dr. C.M.R. Low (James Black Foundation) are also thanked for encouragement and scientific input.

## References

- 1 Hunter, C.A. and Sanders, J.K.M., *J. Am. Chem. Soc.*, 112 (1990) 5525.
- 2 Vinter, J.G., *J. Comput.-Aided Mol. Design*, 8 (1994) 653.
- 3 Lokey, R.S. and Iverson, B.L., *Nature*, 375 (1995) 303.
- 4 a. Williams, D.H., Cox, J.P.L., Doig, A.J., Gardner, M., Gerhard, U., Kaye, P.T., Lal, A.R., Nicholls, I.A., Salter, C.J. and Mitchell, R.C., *J. Am. Chem. Soc.*, 113 (1991) 7020.  
b. Williams, D.H. and Searle, M.S., *J. Am. Chem. Soc.*, 114 (1992) 10690.
- 5 Wyler, R., De Mendoza, J. and Rebek Jr., J., *Angew. Chem. Int. Ed. Engl.*, 32 (1993) 1699.
- 6 Grotzfeld, R.M., Branda, N. and Rebek Jr., J., *Science*, 271 (1996) 487.