Independent sets and the prediction of addition patterns for higher fullerenes

Patrick W. Fowler,[a] Kevin M. Rogers,[a] Ken R. Somers †[a] and Alessandro Troisi[b]

[a] School of Chemistry, University of Exeter, Stocker Road, Exeter, UK EX4 4QD
[b] Dipartimento di Chimica, 'G. Ciamician', Universitá di Bologna, via F. Selmi 2, 40126 Bologna, Italy

Under the assumptions that no two sp^3 carbon atoms are adjacent in the end product of bromination of a fullerene and that the residual π system is a closed shell, graph theory predicts maximum stoichiometries $C_{60}Br_{24}$, $C_{70}Br_{26}$, $C_{76}Br_{28}$, $C_{84}Br_{32}$ and rules out all but 58 of the $\sim 10^{23}$ addition patterns conceivable for these molecules.

Introduction

In the search for simple rationalisations of fullerene reactivity, several concepts taken from graph theory have proved useful. One is the *independence number*, which gives a mathematical embodiment to the chemical notion of steric strain. The independence number, $I(G)$, of a graph G is the maximum number of its vertices that can be marked simultaneously such that no two marked vertices are adjacent.

For icosahedral C_{60} the independence number is 24, and this fact has been used[1] to rationalise the structure of the end product of the bromination of [60]fullerene, *i.e.* $C_{60}Br_{24}$.[2] The argument goes as follows. Assume that the cage takes up the maximum number of *exo* addends compatible with two rules, one steric and one electronic. Rule (i): no two brominated carbon atoms should be adjacent. Rule (ii): the final addition pattern should contain only closed-shell π subsystems. In more formal mathematical language, rules (i) and (ii) are obeyed by those *maximum independent sets* (if any such exist) in which all *components* of the graph of unmarked vertices are *even* in size and have adjacency spectra split into equal halves across a *non-zero gap*.

In C_{60} the independence number $I(C_{60}) = 24$ predicts the stoichiometry, and of the $\sim 3 \times 10^{14}$ conceivable isomers of $C_{60}Br_{24}$ only 1085 are maximum independent sets. Of these, all but one are radicals. In fact, for C_{60}, a weaker version of rule (ii) would achieve the same reduction: we can simply forbid isolated bare carbon atoms (unmarked vertices). The sole survivor under the simultaneous action of rule (i) and of either version of rule (ii) is the experimental T_h isomer, in which the 36 bare carbon centres are arranged as 18 isolated double bonds.[1]

The present work is concerned with the extension of these very simple rules to the prediction of the so-far uncharacterised products of bromination of higher fullerenes. A new invariant related to the independence number is required.

Closed-shell independence number

The essential difference between C_{60} and the higher fullerenes investigated here is that for the latter, all maximum independent sets are radical in character. Achievement of a closed π shell therefore requires either termination of the process of addition at a smaller number of addends or introduction of adjacencies between addend sites. We investigate here the consequences of

---

rigid adherence to the non-adjacency rule: addition is supposed to take place under rules (i) and (ii), to achieve the *largest number* of non-adjacent addends for which there is at least one pattern where a closed shell is possible for every isolated π subsystem. Crane[3] has also considered the consequences of addition under these two constraints. This *closed-shell independence number*, $I^-(G)$, is clearly an integer, characteristic of the molecular graph, obeying eqn. (1).

$$I^-(G) \leq I(G) \qquad (1)$$

The numbers $I^-(G)$ and $I(G)$ can be equal (as for C_{20} and icosahedral C_{60}) but $I(G)$ can be odd whereas $I^-(G)$ cannot. A trivial upper bound on $I(G)$, following from the individual independence numbers of 2 for each of the 12 pentagons and 3 for the $n/2 - 10$ hexagons of a general n-atom fullerene, is given by eqn. (2).

$$I(G) \leq n/2 - 2 \qquad G = \text{fullerene} \qquad (2)$$

In the special case of isolated-pentagon fullerenes that are also *leapfrog* fullerenes C_n, where the set of faces derived from those of the parent $C_{n/3}$ cage (*i.e.* the Clar faces[4]) span the vertices of the larger cage, the bound is tighter [eqn. (3)]. Here we have used the observation that if the vertices of G are spanned by a set of disjoint subgraphs G_i, then $I(G)$ can be at most the sum of the individual independence numbers $I(G_i)$, a bound which may or may not be sharp, depending on the way the subgraphs fit together. Thus, $I(C_{60}) = 24$ follows from the fact that the truncated icosahedron consists of 12 disjoint pentagons, each of independence number 2. C_{20}, the smallest general fullerene, realises the bound of eqn. (2); C_{60}, the smallest leapfrog and smallest isolated-pentagon fullerene, realises the bound of eqn. (3). Non-leapfrog isolated-pentagon fullerenes with $I(G) > n/2 - 6$ are also known, *e.g.* I_h C_{80} with $I(G) = 36$.

$$I(G) \leq n/2 - 6 \qquad G = \text{leapfrog fullerene} \qquad (3)$$

These results also place bounds on the closed-shell independence number, by eqn. (1), but a more detailed argument gives a better bound on $I^-(G)$ for fullerenes (and other trivalent graphs). Let m sites of an n-atom fullerene be occupied by addends. In a pattern that obeys rules (i) and (ii), every empty site is neighbour to at most 2 addend sites. If a_i is the number of empty sites adjacent to exactly i addend sites, counting sites gives eqn. (4) and counting empty–occupied edges gives eqn. (5). Thus, in such a pattern, combination of eqns. (4) and (5)

$$a_0 + a_1 + a_2 = n - m \qquad (4)$$

$$a_1 + 2a_2 = 3m \qquad (5)$$

yields eqn. (6) and identification of m with $I^-(G)$ gives an upper

$$m = 2n/5 - a_1/5 - 2a_0/5 = a_1/3 + 2a_2/3 \qquad (6)$$

bound for the closed-shell independence number [eqn. (7)]

$$I^-(G) \leq 2n/5 \qquad G = \text{trivalent graph} \qquad (7)$$

which improves on eqns. (1) and (2) for $n > 20$ and on eqns. (1) and (3) for $n > 60$. The equality found by Crane [eqn. (3) of ref. 3] is effectively a statement that this bound would be sharp for cases where all sp^2 sites are in isolated double bonds.

Both C_{20} and C_{60} have a maximum independent pattern in which all bare carbon sites are paired in localised double bonds, and for these patterns $a_0 = a_1 = 0$, $a_2 = 3m/2$. In fact, for C_{20} the localised pattern is the only maximum independent set. For C_{60}, there is just one localised pattern out of 1085. *Maximum* independent sets always have $a_0 = 0$, but are not generally closed-shell and hence may have $a_3 \neq 0$, for which the appropriate versions of eqns. (4) and (5) give eqn. (8) and the bound

$$I(G) = n/4 + a_2/4 + a_3/2 =$$
$$2n/5 - a_1/5 + a_3/5 = n/2 - a_1/3 - a_2/6 \quad (8)$$

$n/4 \leq I(G) \leq n/2$ for any trivalent graph. The trivial upper bound of $n/2$ is obtained from partitioning the vertices into disjoint pairs along the $n/2$ double-bond edges of a Kekulé structure, of which any trivalent graph has at least three. It is easily improved for fullerenes, as shown by eqns. (2) and (3). In an addition pattern that violates rule (ii), the maximum possible number of odd components for a given number of empty sites would be when every one is an isolated vertex, and so a_3 is trivially bounded by $n - m$, but since any pentagon contains at least two adjacent unmarked sites, a better bound is $a_3 \leq n - m - 24$ for an isolated-pentagon fullerene.

An extreme example of radical character in a maximum independent set is given by the graphite sheet. As Fig. 1 shows, choice of the marked vertices in one Clar hexagon is enough to fix the unique pattern on the whole plane. The marked set includes half of the vertices and every excluded vertex is also isolated. A similar radical pattern occurs for Clar polyhedra composed of squares and hexagons. The graphite example suggests that for very large fullerenes, the difference between $I(G)$ and $I^-(G)$ will become a linear function of n so that closed-shell coverage will fall increasingly short of the maximum.

Given the definition, we are now ready to see how the closed-shell independence number performs for some specific cases of higher fullerenes. As there is already a closed-shell maximum independent set for C_{60}, the first case to explore is C_{70}, the second isolated-pentagon fullerene. The experimentally characterised isomers of C_{76} and C_{84} are also considered, with a view to predicting the maximum extent of bromination in each case.

Computational strategy

Three general strategies for obtaining non-radical isomers when $I(G) \neq I^-(G)$ can be envisaged: (a) add further Br atoms to saturate the radical patches; (b) stop the addition at a number short of $I(G)$ that allows all subsystems to have a closed shell; (c) move Br atoms around within the set so as to make all patches even. Strategy (a) would introduce at least 24 adjacencies of addend sites for C_{70} (see below). Strategy (c) generally also produces a high number of adjacencies, and by itself it cannot help when $I(G)$ is odd, as in C_{70} (see below). Though a small number of adjacencies may be tolerated when coverage is
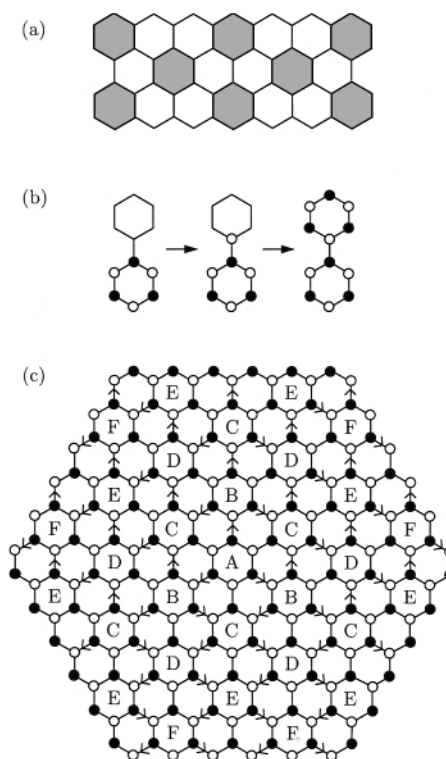

Fig. 1 The unique maximum independent set of graphite is constructed from one of the three Clar coverings of the sheet (a). Assignment of three marked vertices to each Clar hexagon gives the maximum independent set and forces the pattern in Clar neighbours (b). Thus the disposition of marked vertices in hexagon A determines their positions in B, which fix those in C, and so on as shown by the arrows. The final pattern (c) is unique to translation and rotation and has a radical site on every other position.

small, as in $C_{60}Br_6$, or when the atoms are rigorously pyramidal as in the claimed $C_{20}Br_{20}$ molecule[5] (but see ref. 6), adjacency will become increasingly disfavoured for larger fullerenes and higher coverage. It is found that $C_{60}Br_{60}$, for example, is not even a local minimum on the potential surface at some levels of theory;[6] the single adjacency in $C_{60}Br_6$ disappears on addition of further bromine[7] to give $C_{60}Br_8$ and of course $C_{60}Br_{24}$.

Only option (b) of backtracking to reach $I^-(G)$ addends is compatible with rigorous exclusion of adjacencies. A computer program was written to generate independent sets of even size $I(G)$, $I(G) - 1$, $I(G) - 2$ ... and to filter out those with odd components or with even components with open π shells in Hückel theory, thus successively testing the possible values for $I^-(G)$. The technique adopted for C_{70} and larger cases was to split the molecule into hemispherical halves, build independent sets on each and combine them with elimination of any non-independent patterns thereby generated. Cases with internal odd components could be removed before combination, reducing the computational effort. The test for radical character was performed in two stages: (i) addition patterns with at least one odd component were eliminated; (ii) all non-trivial, all-even components were tested for zero HOMO–LUMO gap.

It should be noted that the number of candidate patterns increases rapidly with the distance from $I(G)$, as not only all ancestors of the maximum independent sets are under consideration, but also many sets that are independent on some smaller number of vertices but would meet a dead end before the full number of $I(G)$ addends. For ease of coding, projection of symmetry-distinct solutions was carried out only after the generation and filtering steps, as the total numbers of patterns involved were still manageable. More economical procedures could be designed. The final set of isomers is lexicographically ordered according to the locations of addends in the appropriate IUPAC numbering scheme.[8]
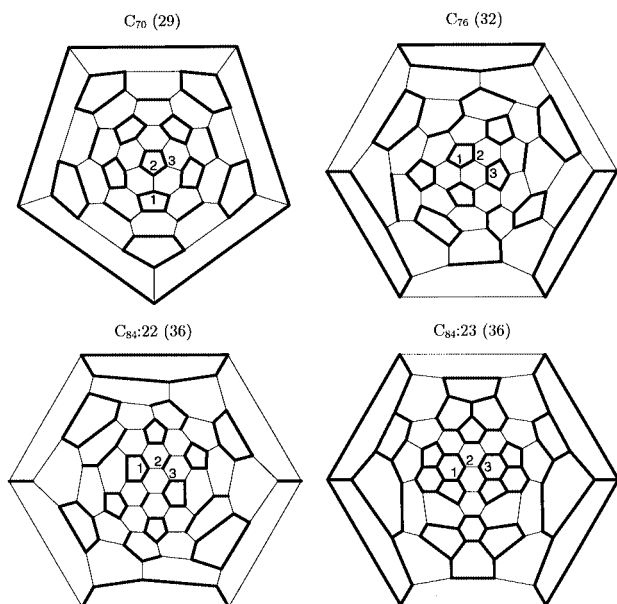
Fig. 2 Independence numbers for four higher fullerenes. The illustrated decompositions into components (bold lines) are spanning sets, *i.e.* they cover each vertex exactly once. For C_{70}, C_{76} and C_{84}:22, the components are twelve pentagons ($P_5 = 12$) plus p_r paths of length r. C_{84}:23 is composed of six pyracylene units ($P_y = 6$). Each set generates a bound on the independence number $I(G) = 2P_5 + 3p_6 + 2p_4 + p_2 + 6P_y$ that turns out to be sharp in all four cases. The numbers 1, 2, 3 indicate the start of the vertex spiral that is the basis of the IUPAC numbering scheme for fullerene derivatives.[8]

Results

(i) $C_{70}Br_x$

The independence number of the D_{5h} experimental isomer of [70]fullerene is $I(C_{70}) = 29$. This follows from the fact that all vertices of the graph are covered exactly once by a disjoint set of twelve pentagons and five equatorial pairs (see Fig. 2); $I(C_{70})$ is hence at most $(12 \times 2) + (5 \times 1) = 29$, and as patterns with 29 non-adjacent vertices are readily found, $I(C_{70})$ is exactly 29. Hence, any maximum independent set of C_{70} contains odd components and is therefore radical. Computer enumeration and construction of the maximum independent sets yields 2972 symmetry-distinct isomers of $C_{70}Br_{29}$ (118 C_s, all others of C_1 symmetry). These are disqualified under rule (ii) by the presence of 11 to 17 odd patches, and every one of the 2972 patterns includes at least eight isolated empty sites.

It follows from eqn. (7) that $I^-(C_{70}) \leq 28$. By the exhaustive approach outlined above it was established that all independent sets of size 28 on C_{70} are also radical, but that ten distinct closed-shell patterns (7 C_1, 2 C_2 and 1 C_s) can be constructed with 26 independent marked vertices. Thus $I^-(C_{70}) = 26$. The ten are illustrated as Schlegel diagrams in Fig. 3. The same ten structures are identified by Crane in his computer search for maximally non-adjacent, closed-shell isomers of $C_{70}X_{26}$.[3] They all have a similar make-up, with localised double-bond pairs constituting the bulk of the bare-carbon patches. No remaining π subsystems contain more than six atoms.

(ii) $C_{76}Br_x$

The experimental D_2 isolated-pentagon isomer of [76]fullerene has $I(C_{76}) = 32$, as demonstrated by its decomposition into twelve pentagons, two paths of length six and two isolated pairs (Fig. 2). Computer search yields 8677 symmetry-distinct maximum independent sets (8651 C_1 and 26 C_2), all of radical character, with 12 to 20 odd components, and each with at least ten isolated bare carbons. Although the $2n/5$ bound gives $I^-(C_{76}) \leq 30$, the computer check again finds the true value to be smaller by two, at $I^-(C_{76}) = 28$. A total of 36 patterns obey-
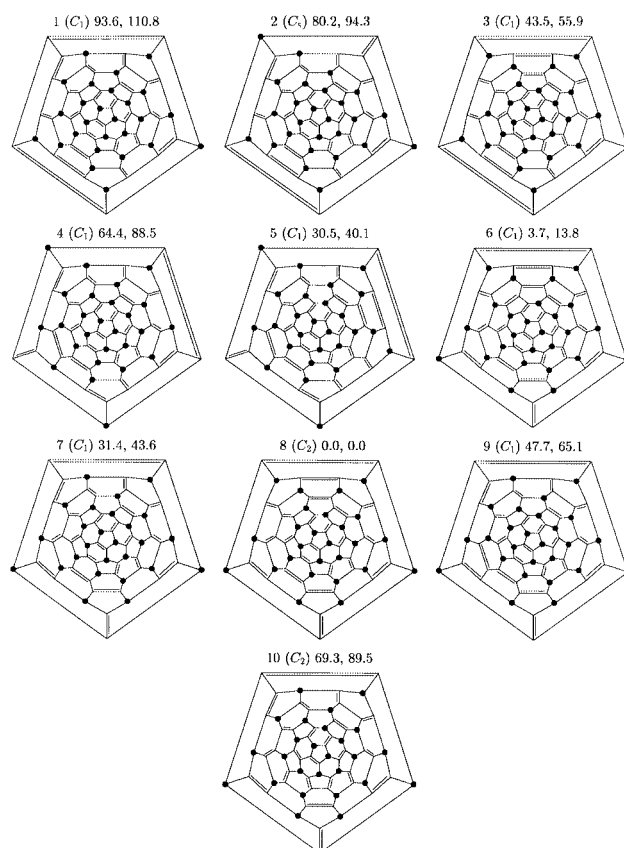
ing rule (ii) are found (33 C_1, 2 C_2, 1 D_2), again with the bare carbons arranged predominantly as localised pairs (see Fig. 4).

(iii) $C_{84}Br_x$

The main C_{84} fraction extracted from Krätschmer–Huffman soot is a 2:1 mixture of isomers, which from analysis of ^{13}C NMR spectra and total-energy calculations have been assigned[9] as 84:22 (D_2) and 84:23 (D_{2d}) in the spiral sequence[10] of 24 isolated-pentagon fullerenes. Both have $I(G) = 36$, as shown by their decompositions (Fig. 2) into pentagons and paths of lengths two or four (D_2) or six pyracylene fragments (D_{2d}). The maximum independent sets for both isomers are, however, all radical in character. 5150 sets (5052 C_1, 94 C_2 and 4 D_2) found for the D_2 isomer have 16 to 24 odd components, at least 14 of which are isolated bare carbon atoms in any one case. The D_{2d} isomer yields 2553 maximum independent sets (2528 C_1, 21 C_2, 1 D_2 and 3 S_4) with the same range of odd components and the same minimum of 14 isolated bare carbons. Eqn. (7) gives a bound of 32 on $I^-(C_{84})$ which is in fact realised for both fullerenes: $I^-(84:22) = I^-(84:23) = 32$. The eleven 32-vertex independent sets (seven for 84:22 and four for 84:23) that are compatible with the closed-shell rule (ii) are illustrated in Fig. 5; all have 18 non-conjugated double bonds and four butadiene chains.

Discussion

Table 1 summarises the results of this purely graph-theoretical survey of the maximum extent of addition of bulky ligands. C_{60} turns out to be a special case in which the unmodified independence number is achievable as a closed π shell. In the higher fullerenes, addition is limited, not by steric crowding, but by the impossibility of achieving a closed shell at the highest

Fig. 3 The complete set of isomers of $C_{70}Br_{26}$ obeying no-adjacency and closed-shell rules. Isomers are lexicographically ordered according to the locations of addends in the IUPAC numbering scheme[8] (see also Fig. 2) and are labelled by symmetry, MNDO and AM1 energies (in kJ mol^{-1}) relative to the isomer of lowest energy (no. 8).

Table 1 Reduction of the isomer problem by application of no-adjacency and closed-shell rules[a]

n	60	70	76	84:22	84:23
$I(G)$	24	29	32	36	36
N_I	$\sim 3.0 \times 10^{14}$	$\sim 2.0 \times 10^{18}$	$\sim 6.7 \times 10^{20}$	$\sim 1.8 \times 10^{23}$	$\sim 9.0 \times 10^{22}$
$N_{I(G)}$	1085	2972	8677	5150	2553
$I^-(G)$	24	26	28	32	32
N_{I^-}	$\sim 3.0 \times 10^{14}$	$\sim 5.6 \times 10^{16}$	$\sim 1.2 \times 10^{20}$	$\sim 3.9 \times 10^{22}$	$\sim 2.0 \times 10^{22}$
$N_{I^-(G)}$	1	10	36	7	4

[a] n is the number of atoms in the fullerene cage, $I(G)$ is the independence number of the graph, N_I the *total* number of distinct addition patterns of stoichiometry $C_nBr_{I(G)}$ and $N_{I(G)}$ is the number of these that are maximum independent sets. $I^-(G)$ is the closed-shell independence number, N_{I^-} the *total* number of isomers with stoichiometry $C_nBr_{I^-}$ and $N_{I^-(G)}$ is the final size of the subset that remains after application of rules (i) and (ii). Exact totals for N_I and N_{I^-}, broken down by point group, are given in refs. 11–13.
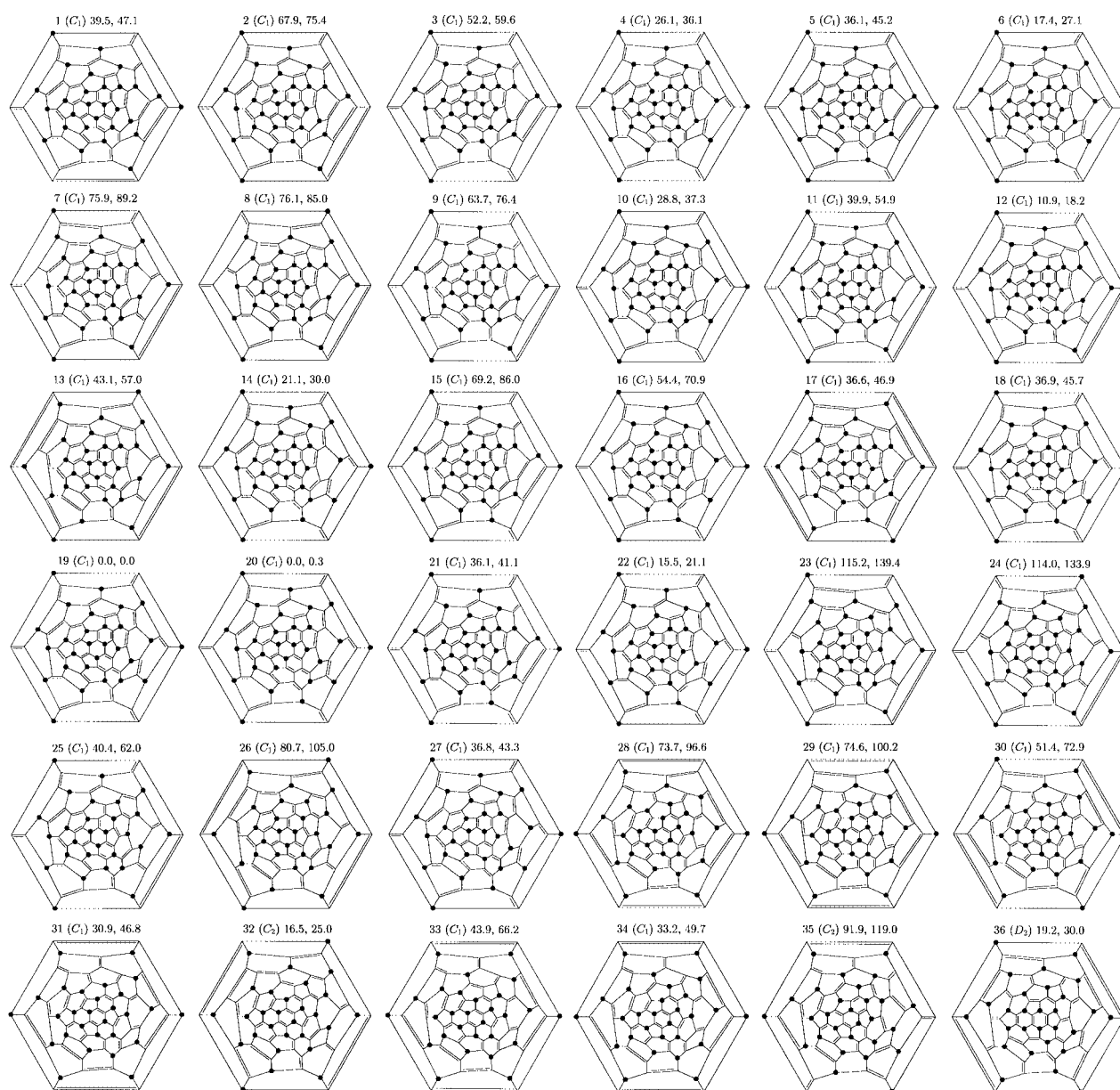

Fig. 4 The complete set of isomers of $C_{76}Br_{28}$ obeying no-adjacency and closed-shell rules. Isomers are ordered and labelled as in Fig. 3. MNDO and AM1 energies (in kJ mol^{-1}) are relative to the isomer of lowest energy (no. 19). Isomers 19 and 20 are isoenergetic within the precision of the methods.

coverage. In all cases, the non-adjacency rule is a powerful constraint on possible addition patterns, reducing the set of isomers under consideration by many orders of magnitude from the huge number that are theoretically possible;[11–13] further insistence on a closed-shell configuration is an increasingly severe additional constraint, cutting the number of candidates down to tens in the cases we have examined.

It remains to check whether the isomers predicted by these rules for maximum coverage are in fact of low energy, and so likely to be thermodynamically favoured. Calculations on $C_{70}Br_{26}$, $C_{76}Br_{28}$ and $C_{84}Br_{32}$ with all three semi-empirical MOPAC methods[14] give relatively narrow spreads of energy. Whilst it is dangerous to rely too heavily on the exact ordering of isomers from such methods, it is perhaps worth noting that
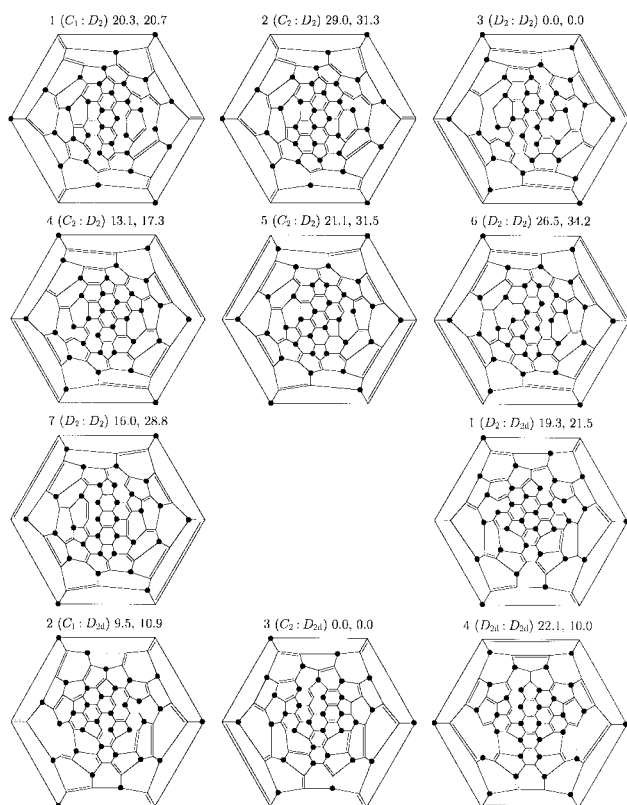
Fig. 5 The complete set of isomers of $C_{84}Br_{32}$ based on isomers 84:22 (D_2) and 84:23 (D_{2d}) and obeying no-adjacency and closed-shell rules. Isomers are ordered and labelled (as in Fig. 3) in two sequences, with symmetries ($g:h$) of both the addition pattern (g) and parent fullerene (h) given in each case. MNDO and AM1 energies (in kJ mol^{-1}) are relative to the isomer of lowest energy in each sequence.

the AM1 and MNDO methods agree on the identity of the most stable isomer in all four cases. The PM3 Hamiltonian has a documented problem of spurious attraction between non-bonded bromine atoms,[15–17] but still gives approximately the same range and rough order of isomers. The overall spread of energies within the sets favoured by the two rules is small (~100 kJ mol^{-1} for $C_{76}Br_{26}$, for example), and sample calculations on molecules with radical patterns or bromine adjacencies indicate large energy penalties for stepping outside the set.

The main result of the present investigation is that the limit of coverage for bromination has been established. If more than $I^-(G)$ addends are attached to a fullerene, then the product must have sp^3 adjacencies, or an open π shell, or both. Patterns with fewer than $I^-(G)$ added bromine atoms may be kinetically or even thermodynamically favoured, but $I^-(G)$ is a limit to the addition process under the stated assumptions. The predictions

do not depend on the precise chemical identity of the addend, only on the existence of a significant steric requirement; for example, $C_{60}Cl_6$ and $C_{60}Br_6$ are isostructural[18] and chlorination of C_{60} terminates at 24 addends[19] (but see ref. 20) yielding a product apparently isostructural[19] with $C_{60}Br_{24}$. A crucial test of the relevance of the present purely mathematical model and its restrictions will come with the experimental characterisation of addition products for the higher fullerenes.

References

1 P. W. Fowler, P. Hansen, K. M. Rogers and S. Fajtlowicz, *J. Chem. Soc., Perkin Trans. 2*, 1998, 1531.
2 F. N. Tebbe, R. L. Harlow, D. B. Chase, D. L. Thorn, G. C. Campbell, J. C. Calabrese, N. Herron, R. J. Young and E. Wasserman, *Science*, 1992, **256**, 822.
3 J. D. Crane, *Fullerene Sci. Technol.*, 1999, **7**, 427.
4 P. W. Fowler and T. Pisanski, *J. Chem. Soc., Faraday Trans.*, 1994, **90**, 2865.
5 F. Wahl, J. Wörth and H. Prinzbach, *Angew. Chem., Int. Ed. Engl.*, 1993, **32**, 1722.
6 J. Cioslowski, L. Edgington and B. B. Stefanov, *J. Am. Chem. Soc.*, 1995, **117**, 10381.
7 P. R. Birkett, P. B. Hitchcock, H. W. Kroto, R. Taylor and D. R. M. Walton, *Nature*, 1992, **357**, 479.
8 E. W. Godley and R. Taylor, *Pure Appl. Chem.*, 1997, **69**, 1411.
9 D. E. Manolopoulos, P. W. Fowler, R. Taylor, H. W. Kroto and D. R. M. Walton, *J. Chem. Soc., Faraday Trans.*, 1992, **88**, 3117.
10 P. W. Fowler and D. E. Manolopoulos, *An atlas of fullerenes*, Oxford University Press, Oxford, 1995.
11 S. Fujita, *Bull. Chem. Soc. Jpn.*, 1991, **64**, 3215.
12 P. W. Fowler, D. B. Redmond and J. P. B. Sandall, *J. Chem. Soc., Faraday Trans.*, 1998, **94**, 2883.
13 Y. Shao, J. Wu and Y. Jiang, *J. Phys. Chem.*, 1996, **100**, 15064.
14 Calculated using MOPAC 6.00: J. J. P. Stewart, Quantum Chemistry Program Exchange, Department of Chemistry, Indiana University, Bloomington, Indiana, 47405, USA.
15 D. A. Dixon, N. Matsuzawa, T. Fukunaga and F. N. Tebbe, *J. Phys. Chem.*, 1992, **96**, 6107.
16 N. Matsuzawa, T. Fukunaga and D. A. Dixon, *J. Phys. Chem.*, 1992, **96**, 10747.
17 P. W. Fowler and J. P. B. Sandall, *J. Chem. Soc., Perkin Trans. 2*, 1995, 1247.
18 P. R. Birkett, A. G. Avent, A. D. Darwish, H. W. Kroto, R. Taylor and D. R. M. Walton, *J. Chem. Soc., Chem. Commun.*, 1993, 1230.
19 A. J. Adamson, J. H. Holloway, E. G. Hope and R. Taylor, *Fullerene Sci. Technol.*, 1997, **5**, 629.
20 F. Cataldo, *Carbon*, 1994, **32**, 437.

Paper 9/05405G