Generation of Heptagon-Containing Fullerene Structures by Computational Methods

Xiaoyang Liu

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Chemistry

Harry C. Dorn, Chair
Louis A. Madsen
Edward F. Valeev

November 18, 2016
Blacksburg, VA

Keywords: Fullerene, TNT-EMFs, Spiral Algorithm, Geometry, Structure Generation

Generation of Heptagon-Containing Fullerene Structures by Computational Methods

Xiaoyang Liu

ABSTRACT

Since the discovery three decades ago, fullerenes as well as metallofullerenes have been extensively investigated. However, almost all known fullerenes follow the classical definition, that is, classic fullerenes are comprised of only pentagons and hexagons. Nowadays, more and more evidence, from both theoretical and experimental studies, suggests that non-classical fullerenes, especially heptagon-containing fullerenes, are important as intermediates in fullerene formation mechanisms. To obtain fundamental understandings of fullerenes and their formation mechanisms, new systematic studies should be undertaken. Although necessary tools, such as isomer generating programs, have been developed for classical fullerenes, none of them are able to solve problems related to non-classical fullerenes. In this thesis, existing theories and algorithms of classical fullerenes are generalized to accommodate non-classical fullerenes. A new program based on these generalized principles is provided for generating non-classical isomers. Along with this program, other tools are also attached for accelerating future investigations of non-classical fullerenes. In addition, research to date is also reviewed.

Generation of Heptagon-Containing Fullerene Structures by Computational Methods

Xiaoyang Liu

GENERAL AUDIENCE ABSTRACT

The thesis describes the generation of heptagon-containing fullerene structures based on generalized spiral algorithm. Details of algorithms and designs are discussed. The thesis is helpful for people who are working on fullerene geometry studies. The algorithms and programs provided by this thesis could be a good start for other investigations.

For those who are not familiar with programming, the program provided by this thesis can also be used to generate and analyze heptagon-containing structures. In addition, this thesis can also be helpful for people working on areas that geometry analyses of molecules are necessary.

Table of Contents

1. Background: 1
   1.1 Current Evidence of Heptagons in Fullerenes, Metallofullerenes, and Nanotubes 1
   1.2 $LaSc_2N@Cs(hept)-C_{80}$: First Discovered Heptagon-Containing Metallofullerene 5
   1.3 The Role of Heptagon-Containing Fullerenes in Fullerene Formation 9
      1.3.1 Heptagon-containing Fullerenes in Cage Shrinkage 9
      1.3.2 Heptagon-containing Fullerenes in Cage Growth 16
      1.3.3 Conclusion 18
2 Fullerene Topology 19
   2.1 Fullerene Dual 19
   2.2 Point Groups of Fullerenes 21
3 Generalized Spiral Program: 22
   3.1 Euler's Polyhedra Formula 23
   3.2 Schlegel Projection and Spiral Graph 25
   3.3 Spiral Algorithm 28
   3.4 Generalized Fullerene Formation Program Based on Spiral Algorithm 31
   3.5 Research for Single-Heptagon-Containing Fullerenes in the Range C_{30}-C_{38} Structures and Symmetries 35
   3.6 Research for Two-Heptagon-Containing Fullerenes with High Symmetries 39
4 Conclusions 40
5 References: 42
6 Appendix: 44
   A. Methods for constructing fullerene cages from the output of Gspiral program 44
   B. Atlas of single- and two- heptagon-containing fullerenes 46
   C. Generalized Spiral Program(Gspiral.cc) 46

Table of Figures

Figure 1 The regular fullerene motif(left) and heptagon-containing motif (right). **1**

Figure 2 The chemical perforation reaction of C_{60}. **2**

Figure 3 Two views of the structure of LaSc2N@Cs(hept)-C80 as determined by a single-crystal X-ray diffraction. **6**

Figure 4 The structures of $LaSc_2N$@Cs(hept)-C_{80} and $LaSc_2N$@Ih-C_{80}. **8**

Figure 5 The mechanism of the loss of C_2 from C_{60}. **10**

Figure 6 The representation of the pathway for the loss of C_2 carbide from Cs-$C_{86}Cl_{28}$ cage. **12**

Figure 7 The transformation from the asymmetric fullerene cage C_1(51383)-C_{84} to symmetric cages. **14**

Figure 8 The top-down mechanism for the fullerene formation involving low-symmetric fullerenes. **15**

Figure 9 The possible mechanisms for the C_2 carbide incorporation into a fullerene. **17**

Figure 10 The mechanism of the fullerene growth from C_{68} to C_{70} involving the C_2 incorporations and the Stone-Wales rearrangements. **18**

Figure 11 Planar and 3D embedding of two fullerene isomers, I_h-C_{20} and I_h-C_{60}. **19**

Figure 12 The decision tree for figuring out the point group for any fullerene isomer. **22**

Figure 13 The structures of I_h-C_{60} (left) and C_1-C_{36} (right) with one heptagonal face. **24**

Figure 14 Schlegel projection of I_h-C_{60}. **26**

Figure 15 The three valid sequences and corresponding spiral graphs. **30**

Figure 16 The main structure and the accessory components of the generalized spiral program **35**

Figure 17 The motifs of heptagon containing fullerenes that determine the relative energies. **38**

Figure 18 The structures of a series of two heptagon containing fullerenes with D_{7h} and D_{7d} symmetries alternately. **40**

1. Background:

1.1 Current Evidence of Heptagons in Fullerenes, Metallofullerenes, and Nanotubes

This chapter describes the investigations related to heptagonal faces. They are the foundations for my research on heptagon-containing fullerenes. These experimental data reveal the fact that heptagons exist in fullerenes and metallofullerenes, and affect a lot for the properties of the molecules. My research is focused on the studies of heptagon-containing fullerenes and metallofullerenes.

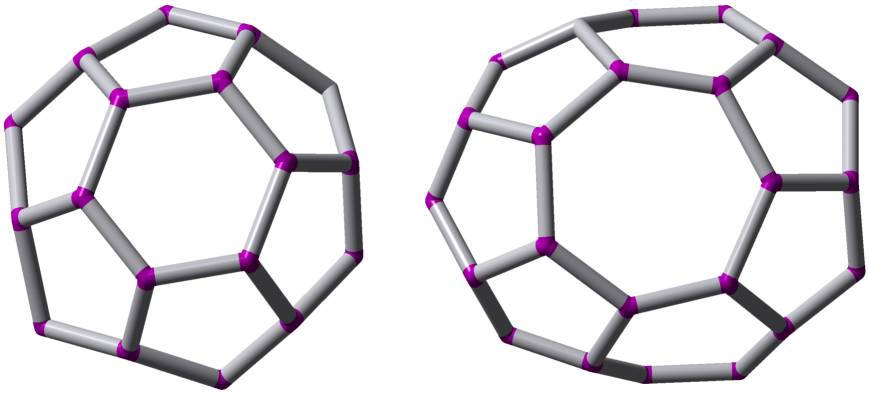

Figure 1 The regular fullerene motif(left) and heptagon-containing motif (right).

The fact that all known classical fullerenes are comprised of pentagons and hexagons[1] attracts close attentions from science communities. In 1991, Smeal et al. discovered from the TEM image that the terminations of nanotubes,[2] polyhedral caps in most cases, might be firstly transformed into conical shapes before the closure. Iijima et al.[3] reported that the cone angle was around 20°, and

concluded that the conversion from the cone-like growth to the cylindrical growth was caused by defects, commonly thought as the incorporation of heptagons and pentagons, in the hexagonal[1] networks. The discovery of heptagons in nanotubes strongly implies that heptagons may also exist in fullerene cages, due to the similarities of properties between these two carbon allotropes and the fact that they can be synthesized under the same condition.[4] In 1995, Wudl's group[5] reported a perforation reaction (**Figure 2**) that broke two adjacent bonds in C_{60} and for the first time, the reaction introduces open faces, rings larger than hexagons, to fullerene cages. The reported reaction is the combination of two reactions. One is the azide reaction between C_{60} and azides[6] and another is the self-sensitized photooxygenation[7] of the products of the azide reaction. In the first step, C_{60} reacts with azides, and produces N-methoxyethoxymethyl (MEM)-substituted [5,6] azafulleroid. Then a C-C bond is broken and a heptagonal face is formed in the fullerene cage. In addition, the reactions make the C=C bonds adjacent to the N atom more reactive to photooxygenation. Thus, photooxygenation, in the second step, breaks additional one or two bonds and introduces faces larger than heptagons in the fullerene cage.

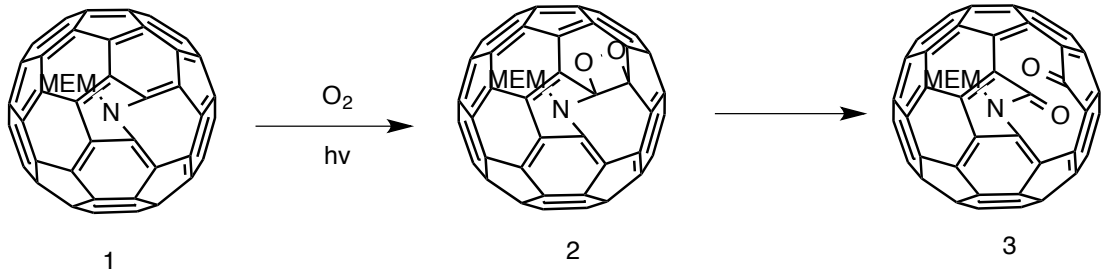

Figure 2 The chemical perforation reaction of C_{60}. Reproduced from ref 42. Copyright 1995 American Chemical Society.

Another chemical perforation that introduced an eleven-atom open face to C_{70} was reported by Birkett et al. in the same year.[8] The final product is obtained in two steps: In the first step, chlorine atoms in $C_{70}Cl_{10}$[9] are substituted by benzenes and then a fullerene derivative, $C_{70}Ph_8$, is obtained. The product belongs to the C_s point group and phenyl groups are isolated between each other. Secondly, $C_{70}Ph_8$ was oxidized automatically along with the breaking of the C-C bond in the open air and an eleven-ring was introduced into the cage.[8] The chemical perforation reaction proves the existence of larger rings in fullerenes and provides a chemical method for producing fullerene derivatives with polygons besides pentagons and hexagons, although this reaction doesn't create a heptagonal face.

Theoretical approaches were also made on heptagon-containing non-classical fullerenes in several studies. Fowler et al.[10] applied two semi-empirical methods, QCFF/PI[11] (quantum consistent force field/π) and DFTB[12] (density functional tight binding), to systematically investigate the stabilities of all 426 C_{40} isomers,[13] including both classical fullerenes and those with one or several heptagonal faces. The results show that the stabilities of heptagon-containing fullerenes are reasonably similar to classical fullerenes. However, there is no heptagon-containing isomer that is more stable than all classical ones.[10] Furthermore, Fowler et al. calculated the energetic penalty for the introduction of one heptagon in a fullerene cage, and the penalty is around 90-150 kJ/mole suggested by DFTB calculation.[10] Based on these results, Fowler et al. proposed a generalized isolate pentagon rule (GIPR), which governs classical as well as heptagon-containing non-classical fullerenes, for predicting the relative stabilities of fullerene isomers. According to this rule, the relative energy of a fullerene is determined by two factors, e_{55} and e_{57}, which are the number of pentagon adjacencies and the number

of heptagon-pentagon adjacencies, respectively.[10] So finding the most stable fullerene isomer is simplified to minimize e_{55} and to maximize e_{57} simultaneously. It is highly possible, according to the trends, that a particular heptagon-containing non-classical fullerene may exceed the stabilities of all classical fullerenes of the same size. Later, Ayuela et al.[1] calculated energies of all classical isomers[14,15] of C_{62} and those with one heptagon computationally. Heptagon-containing fullerene isomer C_{62}, with e_{55}=3 and e_{57}=5, was found to be the most stable one among all 2385 isomers examined.[1] Six calculation methods[11,16-18] were employed to confirm the results, and they were in good agreement regarding relative stabilities. What's more, the favored non-classical fullerene cage can be produced by the insertion of a C_2 carbide on the hexagonal face of the I_h-C_{60}, the most famous known classical isomer, and then transformed to the C_{62} isomer through the Stone-Wales rearrangement.[1] The conversion from the I_h-C_{60} to the stable C_{62} fullerene isomer via the favored heptagonal C_{62} isomer showed that heptagon-containing fullerenes might play a great role in fullerene transformations.

Martsinovich et al.[19] reported the isolation and the characterization of C_{58} fullerene derivatives incorporating heptagons. They obtained the fullerene derivatives by heating the mixture of C_{60} and fluorinating reagents in vacuum,[20] such as transition metal fluorides. The reaction yielded $C_{60}F_{18}$ and fluorinated C_{58} that accounted for around 80% and 20%, respectively. The crude product was separated and purified by high pressure liquid chromatography (HPLC), and two C_{58} derivatives, $C_{58}F_{17}CF_3$ and $C_{58}F_{18}$, were obtained.[19] The mass spectrometry and the fluorine nuclear magnetic resonance (^{19}F-NMR) spectroscopy[19] suggested both of these fullerene derivatives had a heptagonal face. Martsinovich et al.[19] pointed out the remarkable stabilities of C_{58} fullerene derivatives were due to the

presence of fluorine atoms, which released the strain of the fullerene cages by introducing sp^3 hybridized carbon atoms. They also proposed a hypothesis[21] that $C_{58}F_{17}CF_3$ and $C_{58}F_{18}$ were formed by the loss of the C_2 carbide from a 6:5 bond of C_{60}, the bond shared by a hexagon and a pentagon. The data obtained from the synthesis of the derivatives strongly supported the hypothesis that heptagon-containing fullerenes could be transformed from classical fullerenes and act as key intermediates in fullerene transformations.[19]

Heptagons, according to the facts disscused above, are proven to exist in graphenes, fullerenes and metallofullerenes. These heptagon-containing fullerene species are relatively stable and as a result, are promising candidates as intermediates of fullerene formation pathways. So, investigations on heptagon-containing fullerenes are important to reveal the secrets of the fullerene formation mechanisms.

1.2 $LaSc_2N$@Cs(hept)-C_{80}: First Discovered Heptagon-Containing Metallofullerene

$LaSc_2N$@Cs(hept)-C_{80} is the first synthesized metallofullerene isomer with a heptagonal face. Before the discovery of this isomer, heptagon-containing fullerenes are only predicted by theories and calculations. Obtaining heptagon-containing fullerene isomers in laboratory is a milestone for studies related to heptagon-containing fullerenes.

A large portion of non-classical fullerenes, especially heptagon-containing fullerenes, has been predicted to be competitive with classical fullerenes in stabilities based on theoretical approaches.[22] The conclusion indicates that it is highly possible to isolate heptagon-containing fullerenes, especially those with exohedral or endohedral stabilizations. However, all non-classical fullerenes

obtained in laboratory are modified externally with halogen atoms,[23] and those without exohedral modifications were absent until recently.[24] Various conjectures have been made on why exohedral derivations can stabilize heptagon-containing fullerenes. One of them suggests that the exohedral modifications can stabilize the pentagon adjacency defects by transforming halogenated carbon atoms from sp^2 to sp^3 and thereby reducing strains. Similarly, metal ions or clusters inside fullerenes can also stabilize the fused pentagon motifs and then the fullerene cages.[25-27] In 2015, Zhang et al. reported the first heptagon-containing endohedral fullerene, $LaSc_2N@C_s(hept)\text{-}C_{80}$.[24] It was obtained using arc-discharge technique and characterized by X-ray diffraction. The structure of $LaSc_2N@C_s(hept)\text{-}C_{80}$ is illustrated in **Figure 3**. Two classical isomers, $LaSc_2N@I_h\text{-}C_{80}$ and $LaSc_2N@D_{5h}\text{-}C_{80}$, were also synthesized along with the heptagon-containing one.

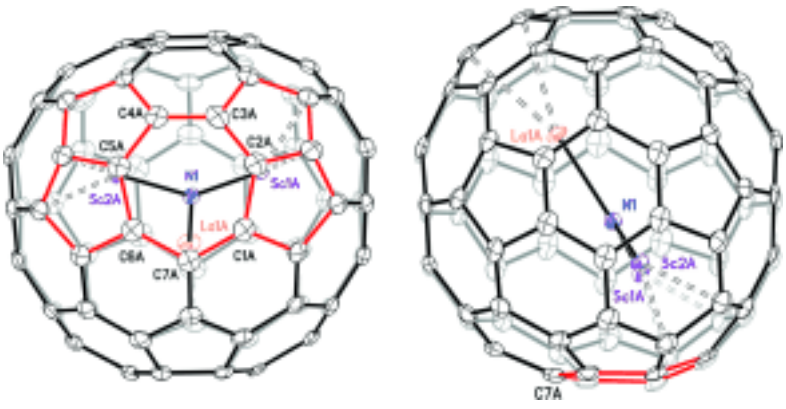

Figure 3 Two views of the structure of LaSc2N@Cs(hept)-C80 as determined by a single-crystal X-ray diffraction. The heptagonal face and the two pentalene motifs are highlighted in red. Reproduced from ref. 24. Copyright 2016 John Wiley & Sons, Inc.

As **Figure 3** shows, the metal cluster, $LaSc_2$, is planar with the metal ions pointing to the adjacent pentagons.[24] The structure is in good agreement with the assumption mentioned above that metal ions could stabilize the fused pentagon motifs. The heptagonal face is curled, in contrast, pentagons and hexagons are planar,[24] which is consistent with usual observations.

To find out the impact of heptagonal faces on the stability, they applied a computational study based on DFT and made comparisons with the classical isomers synthesized simultaneously. The calculation suggests that the C_s(hept)-C_{80} cage is the least stable one, 177 kJ/mol less stable than I_h-C_{80} and 83 kJ/mol less than D_{5h}-C_{80}.[24] The encapsulation of nitride clusters significantly increase the stability of C_s(hept)-C_{80} isomer, but provides only a minor effect on the I_h-C_{80} and D_{5h}-C_{80} isomers. Therefore, the stability of $LaSc_2N@C_s$(hept)-C_{80} ranks in the middle, 34 kJ/mol higher than $LaSc_2N@D_{5h}$-C_{80} and 27 kJ/mol lower than $LaSc_2N@I_h$-C_{80}. The result reveals that the thermodynamic stability is not the dominant factor in the synthesis of heptagon-containing fullerenes, or the yield of $LaSc_2N@C_s$(hept)-C_{80} should be much higher. The unexpected high thermodynamic stability of $LaSc_2N@C_s$(hept)-C_{80} together with the low yield, five times lower than $LaSc_2N@D_{5h}$-C_{80}, indicates the synthesis of heptagon-containing fullerenes is under kinetic control. They assume the kinetic instability of the heptagonal EMFs come from the trend of rearranging to other isomers and the low energy barrier for the rearrangements. In this case, $LaSc_2N@C_s$(hept)-C_{80} is closely similar to $LaSc_2N@I_h$-C_{80} and can be transformed to the classical fullerene through a single step of C_2 segment rearrangement. The transformation is illustrated in **Figure 4**.

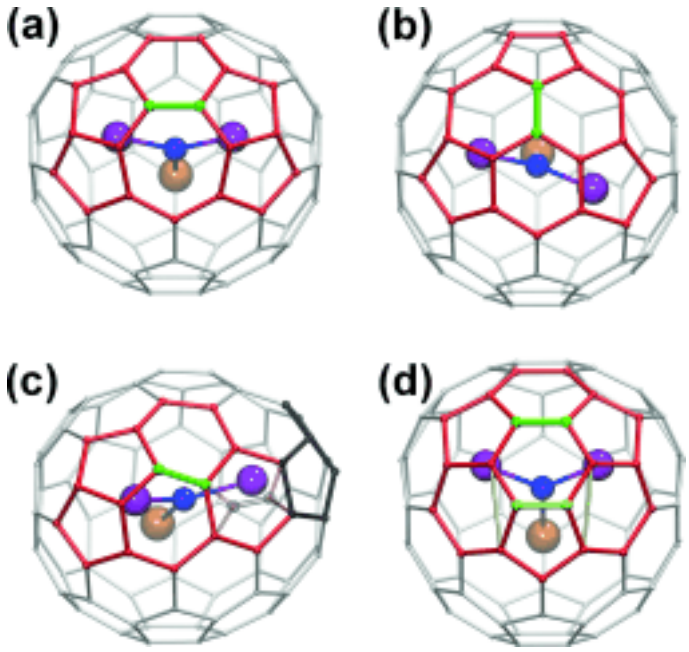

Figure 4 The structures of $LaSc_2N@Cs(hept)-C_{80}$ and $LaSc_2N@Ih-C_{80}$. (a) The structure of $LaSc_2N@C_s(hept)-C_{80}$. (b) The structure of $LaSc_2N@I_h-C_{80}$. The C_2 segment highlighted in green is the bond rotated in the transformation and the heptagonal face and faces around it are marked in red. (c)$LaSc_2N@C_{78}$ with the C_{78} (22010) cage. The cage is obtained from $C_s(hept)-C_{80}$ after losing one C_2 fragment. The removed C_2 segment is shown as "ghost" atoms and the new fused pentagons formed after the removal of the C_2 segment is highlighted in black. (d) The structure of $LaSc_2N@C_{3s}(hepta)-C_{82,}$ which can transform to $LaSc_2N@C_s(hept)-C_{80}$ after the removal of the yellow C-C bond. Reproduced from ref. 24. Copyright 2016 John Wiley & Sons, Inc.

The rearrangements, addition or removal of a C_2 segment shown in **figure 4** (c) and (d), provide possible formation paths of $LaSc_2N@C_s(hept)-C_{80}$. In conjunction

with the result that $LaSc_2N@C_s$(hept)-C_{80} can easily transform to more stable $LaSc_2N@I_h$-C_{80}, $LaSc_2N@C_s$(hept)-C_{80} acts as an intermediate for the formation of $LaSc_2N@I_h$-C_{80} from $LaSc_2N@C_s$(hepta)-C_{82}.

As mentioned above, Zhang et al.[28] reported that asymmetric fullerenes could convert to high-symmetric fullerenes and represented a mechanistic link between large and small cages. The asymmetric fullerenes, as investigations suggest, are instable in kinetic. Considering the fact that most of heptagon-containing fullerenes are low-symmetric and even asymmetric, kinetic instabilities may be caused by the low symmetries.

1.3 The Role of Heptagon-Containing Fullerenes in Fullerene Formation

Investigations on heptagon-containing fullerenes are essential to gain comprehensive understandings of fullerene formations, since heptagon-containing fullerenes are predicted as important intermediates of fullerene formations. In this thesis, I provide the generalized spiral algorithm and Gspiral program to calculate heptagon-containing fullerenes. These structures are necessary to study the roles of heptagon-containing fullerenes in formation mechanisms exclusively. Understanding proposed formation mechanisms are required to push my research forward and to add functions for studies of formation mechanisms in my program.

1.3.1 Heptagon-containing Fullerenes in Cage Shrinkage

Fullerenes and metallofullerenes have been studied in many aspects for more than thirty years. However, the formation mechanisms still remain unknown.[28] Most recently, the evidence and data converge at one point,[29] non-classical fullerenes, especially heptagon-containing fullerenes. Heptagon-containing fullerenes have been highly considered as the intermediates of fullerene transformations and formations. They are promising candidates for connecting fullerenes of different sizes. So the investigations focused on heptagonal fullerenes will make great contributions to the fullerene formation research. In this section, I will describe the heptagon-containing fullerene's role in the fullerene shrinkage, one of the most important components of the top-down mechanism.

In 1993, Eckhoff et al.[30] proposed a C_2 loss[31] mechanism based on numerous C_2 fragmentation studies. The mechanism[30] (**Figure 5**), involving the heptagon-containing intermediates and the Stone-Wales rearrangement,[32] agreed with all experimental and theoretical data. The shrinkage of fullerenes actually provides a possible path for fullerene transformations between two cages differing two carbon atoms. This mechanism, although proposed to explain the C_2 fragmentation of C_{60}, has been widely accepted and applied for demonstrating fullerene conversions, due to its rationality and coherence to experimental data.

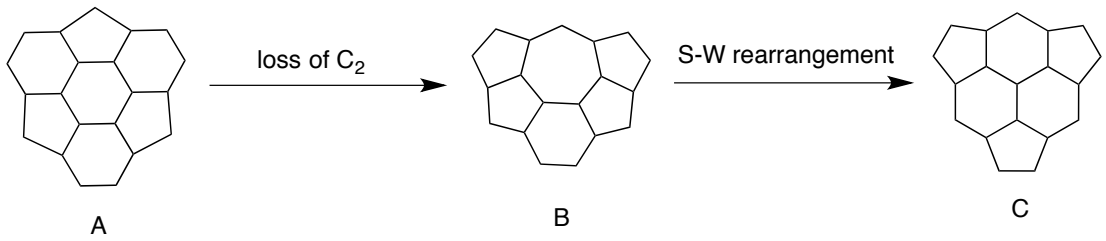

Figure 5 The mechanism of the loss of C_2 from C_{60}. (A) A motif of the C60 cage. (B) The structure of the intermediate incorporating a heptagon. (C) The final

structure after the loss of the C_2 carbide and the rearrangement. Reproduced form ref 58. Copyright 1993 2016 Elsevier B.V.

In 2010, Ioffe et al.[23] reported the transformation from the classical fullerene C_{86} to the heptagon-containing non-classical fullerene derivative $C_{84}Cl_{32}$ by the loss of C_2 and the chlorination. The observation strongly supported the C_2 loss mechanism[30] described above. The C_{86} was first synthesized, and then chlorinated by VCl_3 at 250°C for 4-5 days.[23] The products were characterized by X-ray diffraction as well as NMR.[33] The characterizations show that there were two fullerene derivatives, Cs-$C_{86}Cl_{28}$ and $C_{84}Cl_{32}$,[34] in the final products. The second product, although not the major one, is very interesting because of the incorporation of a 7-membered ring. The structure of $C_{84}Cl_{32}$, comprising thirteen isolated pentagons, was closely related to Cs-$C_{86}Cl_{28}$, the major product of the chlorination reaction. The relationship between these two products strongly suggested that $C_{84}Cl_{32}$ was formed via the shrinkage of Cs-$C_{86}Cl_{28}$ by losing the C_2 carbide from the 5:6 bond position.[23] Then, Ioffe et al. proposed a detailed mechanism (**Figure 6**) to illustrate the loss of C_2.

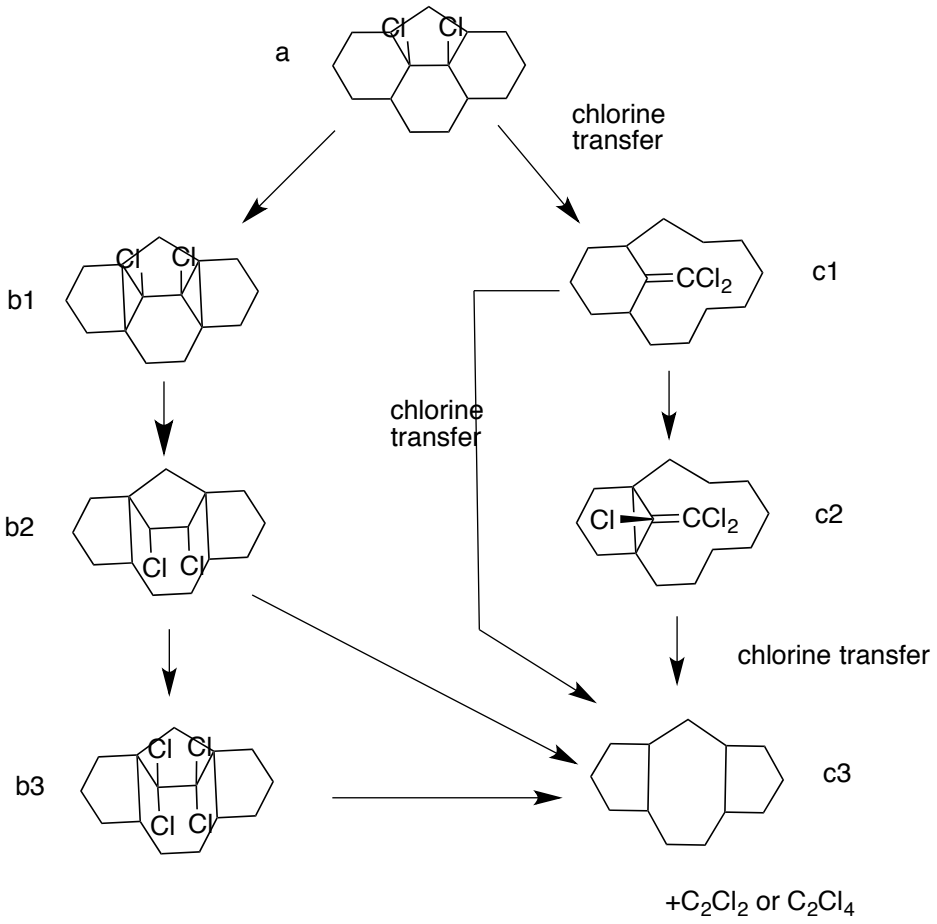

Figure 6 The representation of the pathway for the loss of C_2 carbide from Cs-$C_{86}Cl_{28}$ cage. Reproduced from ref 60. Copyright 2010 John Wiley & Sons, Inc.

The shrinkage of Cs-$C_{86}Cl_{28}$ is not an isolated phenomenon. Under the same condition, $C_{90}Cl_{24}$ can transform to heptagon-containing $C_{88}Cl_{22/24}$ by losing a C_2 carbide.[35] In this case, the C_{90} cage was, surprisingly, an isomer following the IPR,[36] which was typically thought very stable and hard to shrink. So the shrinkage of fullerenes by losing the C_2 carbides is extremely broad and can occur in almost every fullerene cage.

In 2013, the Dorn group[28] reported a remarkable discovery that asymmetric metallofullerene cages could transform to symmetric ones by the fullerene shrinkage. They also proposed a breakthrough hypothesis, and the hypothesis

connected the fullerene shrinkage path with the top-down fullerene formation mechanism. The asymmetric metallofullerenes, $Y_2C_2@C_1(51383)$-C_{84} and $Gd_2C_2@C_1(51383)$-C_{84}, were prepared by electric-arc synthesis. $Y_2C_2@C_1(51383)$-C_{84} was synthesized with ^{13}C enriched graphite rod and Y_2O_3; $Gd_2C_2@C_1(51383)$-C_{84} was synthesized with normal carbon rods and gadolinium oxide.[37,38] The products were separated by HPLC and had a long retention time, indicating the existence of pentalene motifs.[39] In addition, $Gd_2C_2@C_1(51383)$-C_{84} was characterized by the single-crystal analysis, and $Y_2C_2@C_1(51383)$-C_{84}, since it was enriched by ^{13}C, was identified by ^{13}C NMR.[28] The analysis confirmed both of the two products had C_1 symmetry, the lowest symmetry a fullerene could have and thought as asymmetric. Then, at high temperature, the asymmetric metallofullerene cages lost a C_2 carbide[19] and shrunk to the highly symmetric C_{82} cages,[40] surprisingly. Zhang et al., based on the observations, proposed a transformation mechanism (**Figure 7**), involving the widely accepted C_2 loss[30] mechanism and the Stone-Wales rearrangement.[32] As discussed above,[23,30,35] the intermediates between two adjacent fullerenes are heptagon-containing fullerenes. [1]

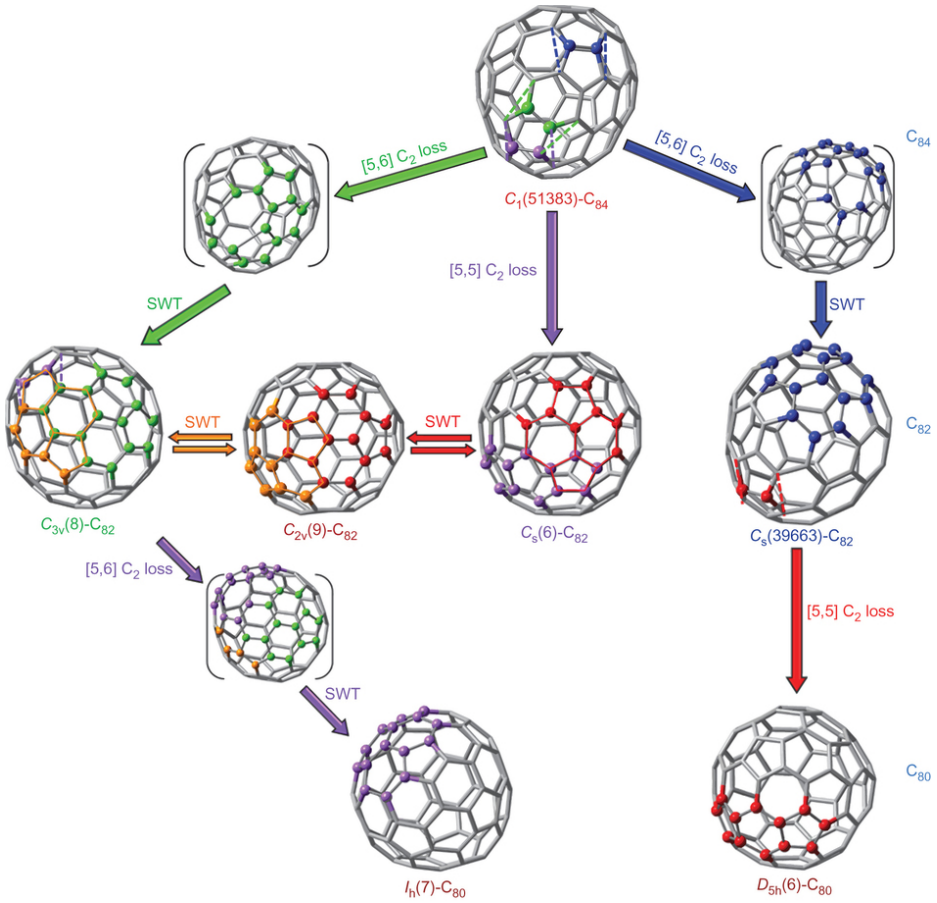

Figure 7 The transformation from the asymmetric fullerene cage $C_1(51383)$-C_{84} to symmetric cages. Many famous and stable fullerenes isolated are involved. Colors present the motifs changed in the transformation. Adapted from ref 28. Copyright 2013 Macmillan Publishers Limited.

The experiment strongly suggested the top-down fullerene formation mechanism,[28] since several well-known high-symmetric fullerenes[14] were formed from larger asymmetric fullerenes in the top-down path. The asymmetric fullerene cages are called "missing links",[28] and are the outcomes of giant fullerenes' shrinkage as well as the precursors of high-symmetric small fullerenes. Heptagon-containing fullerenes, acting as the intermediates in the fullerene shrinkage pathway, connect two fullerenes differing two carbons. In general, the big picture of top-down formation mechanism includes three sections.[28] First,

giant fullerenes are formed from graphene as discussed in the previous chapters. Second, asymmetric fullerenes are produced from giant fullerenes by the loss of C_2 carbides. Third, small symmetric fullerenes are created by the shrinkage of asymmetric fullerenes followed the Stone-Wales rearrangements.[30,41]

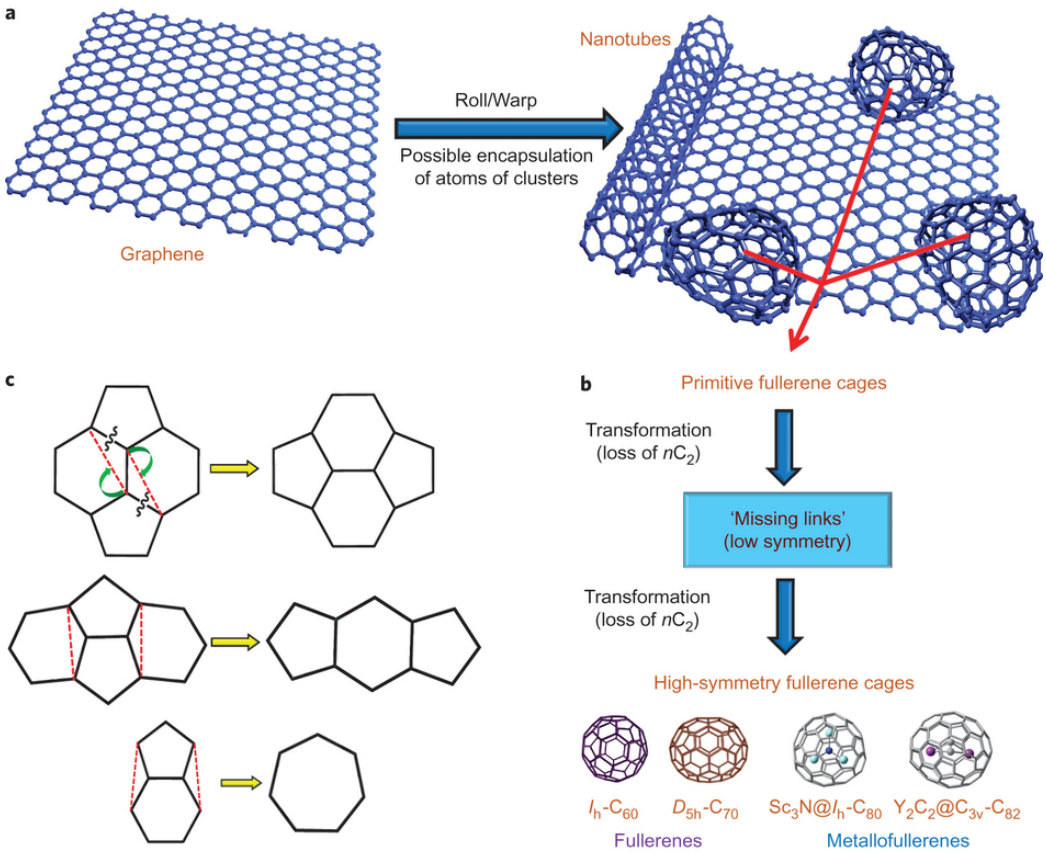

Figure 8 The top-down mechanism for the fullerene formation involving low-symmetric fullerenes. (a) Primitive giant fullerenes are formed from the coiling of the graphene on the edge. (b) The pathway of the fullerene shrinkage. In this process, giant fullerenes transform to low-symmetric fullerenes or called "Missing Link" first and then symmetric fullerenes are formed via the shrinkage of "Missing Links". (c) The mechanism for fullerene shrinkage. Heptagon-

containing fullerenes and the Stone-Wales rearrangement are involved.

Reproduced from ref. 28. Copyright 2013 Macmillan Publishers Limited.

1.3.2 Heptagon-containing Fullerenes in Cage Growth

Recently, although more and more evidence from laboratory and even outer space[42] strongly suggests the top-down fullerene formation mechanism,[28] the bottom-up mechanism are still accepted because there is no direct evidence dismissing the bottom-up mechanism. Some hypotheses were proposed to explain experimental observations based on the fundamental idea of the bottom-up mechanisms.[43] Coincidentally, heptagon-containing non-classical fullerenes, the proposed intermediates in top-down theories, are also thought to play important roles in bottom-up mechanisms.[44]

In 2001, Terrones et al.[44] reported a fullerene growth mechanism backing bottom-up fullerene theories, based on the observations and the proven hypotheses. It has been widely accepted, based on the data accumulated by investigations on heptagon-containing fullerenes, that non-classical fullerenes, especially heptagon-containing fullerenes, are comparable in stability[1] with classical fullerenes in many cases. So they proposed several C_2 incorporation mechanisms[44] (**Figure 9)** involving heptagon-containing fullerenes. The incorporation may occur on hexagonal faces, edges of two adjacent hexagonal faces or adjacent heptagonal faces, and yields the motifs containing pentagons, hexagons as well as heptagons. The carbide incorporation mechanism, acting as an important part of the fullerene growth, demonstrated how small fullerenes transformed into large fullerenes in detail.

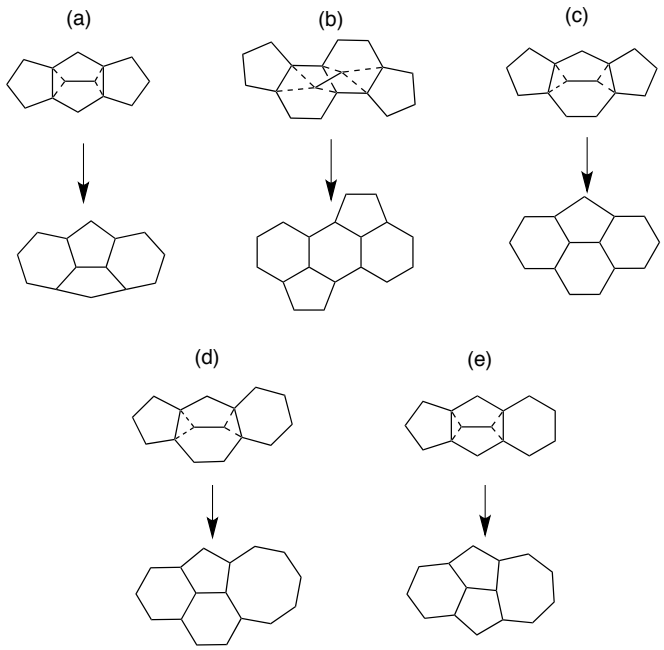

Figure 9 The possible mechanisms for the C_2 carbide incorporation into a fullerene. (a) and (b) involve only classical fullerenes. (c) and (d) show the incorporation on heptagonal face, and a classical fullerene and a heptagon-containing fullerene are yielded, respectively. (d) indicates the incorporation reacting on hexagonal face of a classical fullerene and a heptagon-containing fullerene is generated.

Wang et al.,[45] based on the C_2 carbide incorporation mechanism discussed above, reported a complete fullerene growth path (**Figure 10**) to illustrate the transformation from C_{68} to C_{70}. Moreover, they also calculated the energy barrier for each step, and found the energy needed for the growth is relatively low. The fullerene growth pathway involved the C_2 carbide incorporation as well as the Stone-Wales rearrangements, and surprisingly, the arrangements required more energies than the C_2 incorporation.[45] It suggests that the C_2 incorporations are very likely to occur since the Stone-Wales rearrangement has been proved by

theoretical and experimental evidence, although it has not been directly observed yet.[45]

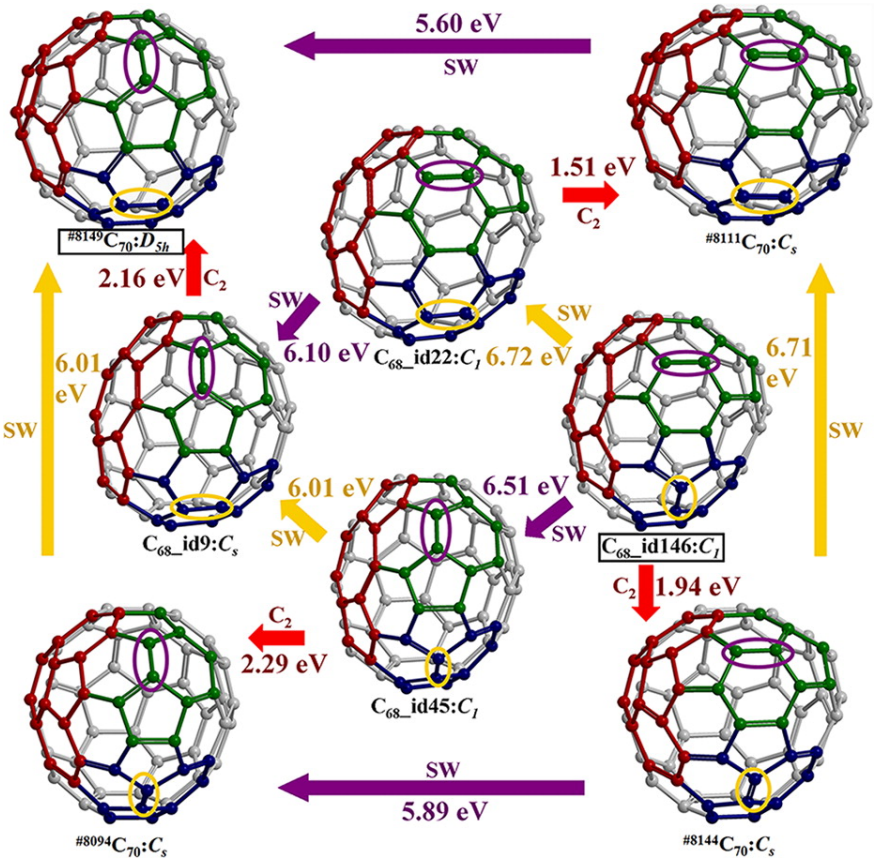

Figure 10 The mechanism of the fullerene growth from C_{68} to C_{70} involving the C_2 incorporations and the Stone-Wales rearrangements. The barrier energy is calculated for each step and shown on the arrows. Adapted from ref 69. Copyright 2012 American Chemical Society.

1.3.3 Conclusion

Heptagon-containing fullerenes are repeatedly proposed as intermediates for both fullerene shrinkage and growth mechanisms, key parts of the top-down[42] and the bottom-up[43] fullerene formation mechanisms, respectively. Investigations on heptagon-containing fullerenes are essential to figure out whether the fullerene shrinkage or growth is true and furthermore, determine whether the top-down or

the bottom-up formation mechanism is correct. To have a deep understanding of fullerene formation, we need to do more intensive investigations in this overlooked area.

2 Fullerene Topology

2.1 Fullerene Dual

Since my program is developed to analyze and generate fullerene geometries, fullerene topology theories are involved in the design of my program. Fullerene geometric analogs, or duals, have many useful properties, and the transformations between fullerene isomers and their corresponding duals are straightforward. So I use fullerene duals to do calculations in many steps.

In geometry, each polyhedron has an associated dual. The vertices of a polygon correspond to the faces of its dual, and vice versa. According to this definition, all polyhedra are associated into pairs as duals.[46] For example, the dual of a cube is octahedron and the dual of the I_h-C_{60} fullerene is a dodecahedron. (**Figure 11**)

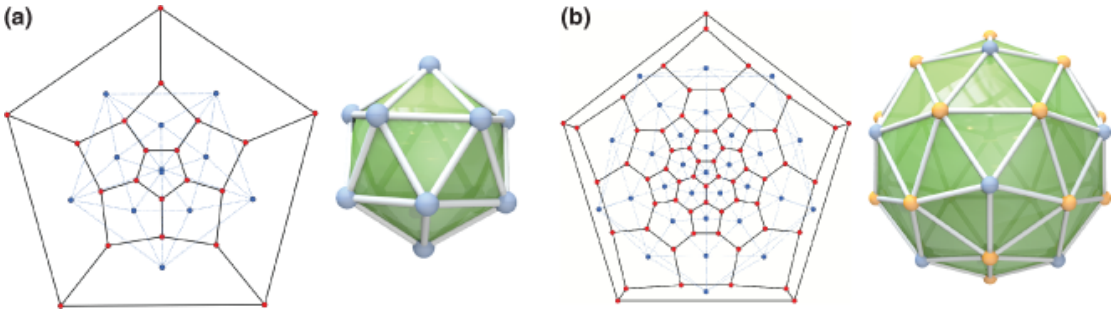

Figure 11 Planar and 3D embedding of two fullerene isomers, I_h-C_{20} and I_h-C_{60}. (a) I_h-C_{20}, of which the dual is icosahedron. (b) I_h-C_{60}, whose dual is dodecahedron. Reproduced from ref 32. Copy right 2014 John Wiley & Sons, Inc.

There are many good properties when polyhedra are studied in pair of duals. First, a polyhedron has the same number of edges as its dual, due to Euler's polyhedra formula. The dual operation makes exchange of the factors, f and v. According to the formula e = v + f +2, e, the number of edges, remains unchanged. Another key property of duals is that dual operation conserves the point group symmetry. A polyhedron and its dual belong to the same point group. For example, an icosahedron and a dodecahedron are duals, and both of them have the I_h symmetry. This makes it easier to figure out the point group of a fullerene, since most fullerenes structures are much more complex than its dual. To find out the symmetry group of a fullerene, we can construct the dual of the fullerene and figure out the point group of the dual.

As stated in the definition of fullerenes, carbon atoms are sp^2 hybridized. The hybridization type requires each carbon is connected with other three carbons and surrounded by three faces. Since in the dual operation, vertices are mapped to dual's faces and faces to dual's vertices, we can easily prove that a fullerene's dual is a polyhedron containing only triangular faces, called deltahedron. The dual of a fullerene with n vertices involves n triangular faces and n/2 + 2 vertices. A vertex of the dual is m-valent if the corresponding fullerene face has m edges. Obviousl, a fullerene's dual is much easier to be constructed, transformed and analyzed than the fullerene itself. In most cases, when we want to solve a topology problem related to a fullerene, we first construct its dual and apply operations on its dual and then reconstruct the fullerene structure.

All the merits of fullerene duals are based on the prerequisite that each fullerene vertex is uniquely associated to three jointly connected vertices of the

corresponding dual. In other word, the correspondence between a fullerene's vertices and its dual's faces must be exclusive. This requires that the structure of a fullerene's dual contains no separating triangle. A separating triangle is defined by three vertices of a deltahedron and separates the deltahedron into two deltahedra. It may appear when more than three vertices of the dual are mutually adjacent. One of the most discussed examples is the dual of prism, a triangular bipyramid with a separating triangle. Luckily, it can be strictly proved that duals of classical fullerenes or heptagon-containing fullerenes cannot have separating triangles. However, for other non-classical fullerenes, especially triangle-containing or square-containing fullerenes, this should be considered.

2.2 Point Groups of Fullerenes

The calculations of point groups are an important extension component in my program. There are several ways to identify the point group of a particular fullerene isomer. In this paragraph, I provide the one that can be designed as a computer algorithm and all possible point groups for heptagon-containing fullerenes are included in my program.

A fullerene isomer is marked using its point group. For example, an isomer of C_{60} that belongs to the icosahedral group (I_h) is marked as I_h-C_{60}. This is because isomers belonging to the same point group are quite similar in many aspects.[47] A point group is a set of symmetries and a fullerene isomer belongs to a specific point group if the isomer has all symmetries of that point group. Although there are many point groups for three-dimension structures, fullerenes belong to only a small portion of them. This is because fullerenes are comprised of polygonal faces and these faces limit symmetry operations to a small number. For a classical

isomer comprising pentagons and hexagons, the possible n-fold rotational symmetries are C_1, C_2, C_3, C_5 and C_6.[48] The introduction of faces besides pentagons and hexagons may expand the set of possible point groups. The identification of a fullerene's point group can be made based on the symmetry point group decision tree.[49] (**Figure 12**) Programs also follow the same idea to figure out the point groups of fullerene isomers. Firstly, we can get the order of the isomer's point group and then determine the exact group with the same order.

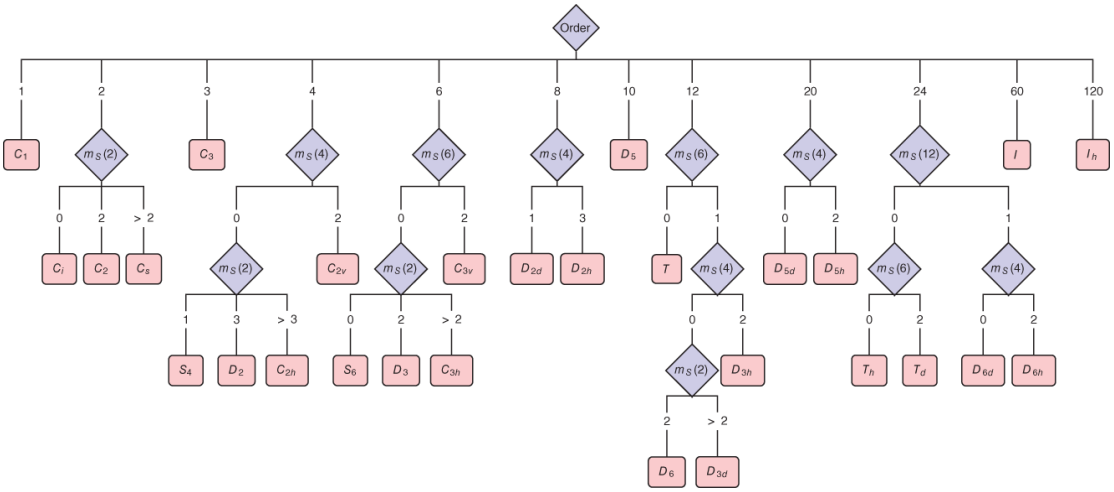

Figure 12 The decision tree for figuring out the point group for any fullerene isomer. Reproduced from ref. 49. Copy right 2014 John Wiley & Sons, Inc.

3 Generalized Spiral Program:

In this chapter, I describe my program for analyzing fullerenes in detail. I designed the generalized spiral algorithm based on a spiral algorithm and implemented it in C language. The program also contains components for calculating topological properties of heptagon-containing fullerenes.

The generations of valid fullerene isomers are required for almost all kinds of fullerene research. It is not hard to generate classical fullerenes since there are several programs available.[50] However, none of them can be used to generate non-classical fullerenes, and the lack of non-classical fullerene generation program actually limits further investigations on non-classical fullerenes. In this chapter, I will introduce a generalized spiral program, a program based on spiral algorithm but can be used to generate non-classical fullerenes. Before the demonstration of the generalized program, I first describe necessary definitions, theories and also the spiral algorithm.

3.1 Euler's Polyhedra Formula

Euler's theorem is one of the basic theories applied for fullerene structure calculations. In my algorithm, I use this theorem to determine the numbers of faces belonging to each category.

Euler's polyhedra formula, named after the great mathematician Leonhard Euler, introduces the relation among the numbers of vertices, faces and edges of a polyhedron. The relation, which reveals the fundamental properties of a 3-dimension polyhedron, is expressed by a simple formula: $v - e + f = 2$. In the formula, v, e and f are the number of vertices, the number of edges and the number of faces, respectively. A fullerene isomer, leaving out the chemical properties and the details of the geometric structures, can be abstracted as a polyhedron,[49] and obeys Euler's polyhedra formula. Considering the fact that all carbons of a fullerene are sp^2 hybridization and each carbon-carbon bond is formed between two atoms, the number of edges is $3n/2$ for a fullerene C_n. So the total number of faces is $n/2 + 2$. In addition, the number of edges is equal to the sum of edges of all

faces divided by two, because every edge is shared by two faces. Based on this relation, we have another equation: $\frac{2n}{3} = \sum edges\ of\ all\ faces$. A classical fullerene is comprised of two categories of faces, pentagons and hexagons. So we can calculate the numbers of pentagons and hexagons in a classical fullerene, which are 12 and n/2 – 10, respectively. For non-classical fullerenes, if we give the number and the type of faces other than pentagons and hexagons, we can also get the numbers of all kinds of faces. To make it clear, I will give two examples, one for the classical fullerene C_{60} and another for the non-classical one heptagon-containing fullerene C_{36}. The two fullerenes are shown in **Figure 13.**

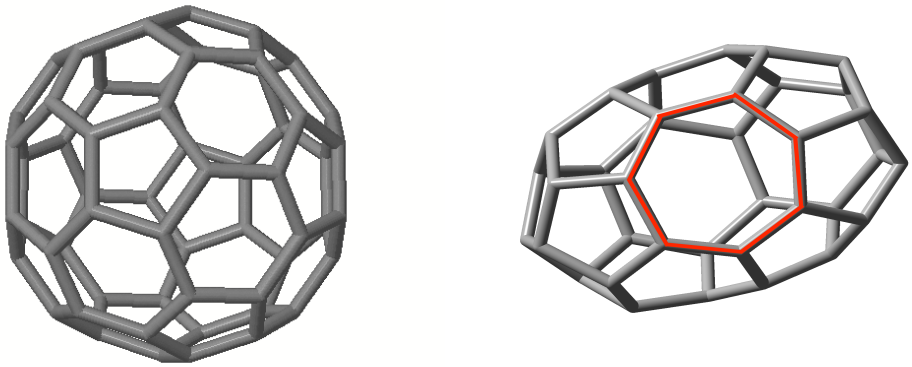

Figure 13 The structures of I_h-C_{60} (left) and C_1-C_{36} (right) with one heptagonal face.

For C_{60}, n = 60, and the linear system is as following:

$$\begin{cases} Relation\ between\ faces\ and\ edges: \\ \quad \frac{(5p+6h)}{3} = n \\ Euler's\ polyhedron\ formula: \\ \quad p + h - \frac{n}{2} = 2 \end{cases}$$ p: # of pentagons h: # of hexagon……(1)

Setting n equal to 60 and solving the linear system, we can get the result:

$$\begin{cases} p = 12 \\ h = 20 \end{cases}$$

For non-classical fullerenes, the linear system can be obtained similarly.

$$\begin{cases} Relation\ between\ faces\ and\ edges: \\ \quad \frac{(5p+6h+hp)}{3} = 36 \\ Euler's\ polyhedron\ formula: \\ \quad p + h + hp - \frac{36}{2} = 2 \end{cases}$$

hp:# of heptagons……(2)

Since there is only one heptagon involved in C_{36}, hp = 1 and the result for the linear system is:

$$\begin{cases} p = 13 \\ h = 6 \end{cases}$$

3.2 Schlegel Projection and Spiral Graph

Schlegel projection and spiral graph are the foundations of spiral algorithms. All the spirals and structure generations are developed based on these two-dimensional graphs. In my new algorithm, I also use these two techniques.

One of the important features of fullerenes is the 3-dimension structure, which however, brings difficulties for the isomer representations and the designs of generation algorithms. To describe a 3-dimension structure, especially polyhedra, an adjacency matrix is commonly used, due to the fact that structures are determined by vertex connection relationships. In graph theory, an adjacency matrix is a square matrix of size n and n is the number of vertices, specifically carbon atoms of fullerenes. The value of each entry, entry (i, j) for example, indicates whether carbon atoms i and j are connected. Since there are only two conditions, we can simply use 1 and 0 represent situations of connected and unconnected, respectively. Based on the definition, each fullerene isomer has a corresponding adjacency matrix.

Although it is convenient to convert a fullerene structure to its corresponding adjacency matrix, it is hard to figure out whether an adjacency matrix is a valid representation for a fullerene isomer. So projection,[51] converting 3-dimension structure to correlative 2-dimension graph, is an alternative method. For fullerene cages, abstracted as convex polyhedra, Schlegel projection,[52] named after a famous mathematician Victor Schlegel, is most widely used for structure analysis. Schlegel projection is a projection from through a point beyond one of faces, usually the point beyond the faces located at the north pole. If two vertices are connected in the polyhedra, then there is a segment joining points corresponding to these two vertices in the projection graph. The entire projection graph is divided by these segments into subdivisions, which conserve the same number of edges as the corresponding faces in the ploy hedra. **Figure 14** illustrates the approach of Schlegel projection of I_h-C_{60}.

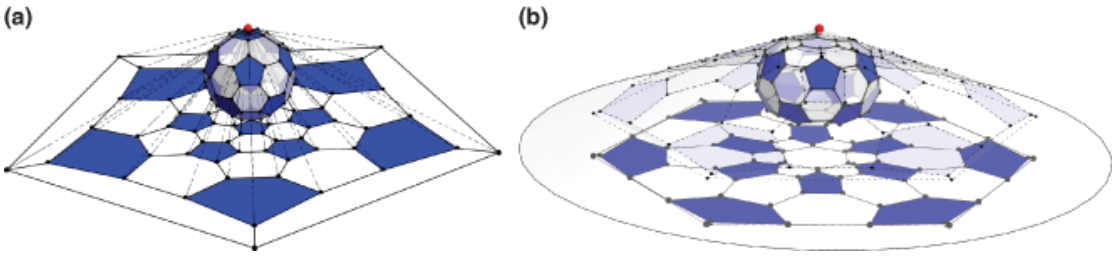

Figure 14 Schlegel projection of I_h-C_{60}. Reproduced from ref. 51. Copy right 2014 John Wiley & Sons, Inc.

Spiral graph,[14] as its name implies, is a spiral strip that reaches every face of the Schlegel image for exactly one time. The spiral strip starts at an arbitrary face and then reaches all other faces following a specific rule to get a continuous spiral curve. The rule requires that at any point the next face to be reached is the one

sharing an edge with both the current face and the first face that is still open. According to the definition above, it is obvious that there may exist more than one valid spiral graph for a given Schlegel graph. As discussing in the last chapter, the relation between the fullerene isomer and the Schlegel graph is one-to-one. So each fullerene isomer may also have more than one spiral graph. The first question is what the higher limit is? According to the definition of spiral graph, if the first three faces are given, the spiral strip will be determined. So the total number of possible spirals can be calculated in this way: first, pick an arbitrary face and start at this face. Second, each edge of the starting face is shared with another face and all these faces can be the next one for the spiral strip. Suppose the starting face has F edges, then there are F possibilities for the choice of the second face. Thirdly, to find the third face, there are two directions from the second face, clockwise or anti-clockwise. So the formula to calculate the higher limit of all possible spiral lines is as following:

$$\sum_{i=1}^{n} F_i \times 2$$

in which, **n** is the number of faces and $\mathbf{F_i}$ is the number of edges of **i**th face.

Not all possible spirals encode the fullerene graph successfully. If the spiral enters a face that is not the last one and is surrounded by faces already traversed, the spiral will fail to wind the fullerene graph.[14] In the worst case, all possible spirals fail and there is no valid spiral for the fullerene graph. Actually, there are fullerene isomers that cannot be encoded by spirals, however the proportion is extremely tiny. More importantly, the smallest fullerene without a spiral is C_{380},[53] a fullerene isomer that is too large to be considered in almost all investigations. In most cases, the spiral is still the most common method to encode a fullerene graph.

3.3 Spiral Algorithm

In this chapter, I describe the original spiral algorithm. This spiral algorithm is first designed and applied for making circuit diagram in electronic engineering. Then, it is extended to other areas including chemistry.

As discussed in the last chapter, fullerenes smaller than C_{380} can always be encoded using spiral strips. Conversely, we can also generate a spiral strip first and then convert it to the corresponding fullerene graph. We notice that the order of faces reached along the spiral path determines a spiral strip uniquely. If we use a single number to represent a face, for example 5 represents a pentagon and 6 represents a hexagon, we will get a sequence of numbers, whose orders in the sequence give the position of their corresponding faces along the spiral path. According to Euler's polyhedra formula, the numbers of each kind of faces can be calculated based on the number of carbons. We can generate all possible sequences based on this result. For example, the classical T_d-C_{20} with 12 pentagons and 4 hexagons should be represented by a sequence of twelve 5s and four 6s. It is not difficult to generate a set of all possible sequences. All sequences corresponding to valid spirals are contained in the set, due to the fact that every valid spiral can be represented by a sequence.

The number of all possible sequences also has a higher bound, which can be calculated based on the permutation theory. Let's use the classical fullerene C_n as an example. According to Euler's polyhedra formula, there are 12 pentagons and n/2 -10 hexagons, and in other words, there are twelve 5s and n/2 -10 6s in the sequence. So there are at most:

$$\frac{\left(\frac{n}{2}+2\right)!}{12!\left(\frac{n}{2}-10\right)!}$$

possible sequences. The higher bound determines the time complexity of the fullerene generation algorithm. The time complexity, which indicates the efficiency of the algorithm, can be improved, but the improvement is not significant.[54] More details will be discussed in the next chapter.

Once we have gotten the set **S** of all possible sequences at this point, the next step is to find a subset **Sv** containing all sequences that can wind up a fullerene successfully. Each time, we pick a tentative sequence from the set **S** and check whether it belongs to **Sv**. The way to check a sequence is to construct a fullerene graph from the sequence following the spiral criteria. As mentioned in previous chapters, the spiral criteria require the new face added should be connected to its direct precursor and the first open face, in which open mean the face is not surrounded by other faces. If we can get a fullerene graph in this way, then the sequence is valid.

The construction procedure can also be considered in the view of the construction of fullerene duals. Since each character in the sequence represents a face of the fullerene, the character also corresponds to a vertex of the fullerene dual. Due to the properties of fullerene duals, the association of fullerene and its dual is a strict one to one relation. If the sequence can construct a dual successfully, it is also a valid sequence for a fullerene isomer.

Although we can use the method demonstrated above to check whether a spiral strip is valid, the problem is that more than one spiral may relate to the same fullerene isomer. If we use the all possible sequences to generate fullerene isomers, there will be duplicated ones. To eliminate repeated spirals and make one fullerene isomer corresponds to exactly one spiral strip, we need to select one spiral from all spirals to represent the corresponding fullerene isomer. Since a

spiral strip is expressed as a sequence of numbers, we can use the sequence with the smallest value as the representation, called the canonical sequence. As the example illustrated in **Figure 15**, three spirals can wind up T_d-C_{28} successfully. The sequences of spirals are 6 5 5 5 5 5 6 5 6 5 5 5 5 5 5 6, 6 5 5 5 5 6 5 6 5 5 5 5 5 5 5 6 and 5 5 5 6 5 6 5 6 5 5 5 5 5 5 5 6. It is obvious that the last one has the smallest value among the three sequences and it is the canonical sequence for T_d-C_{28}.

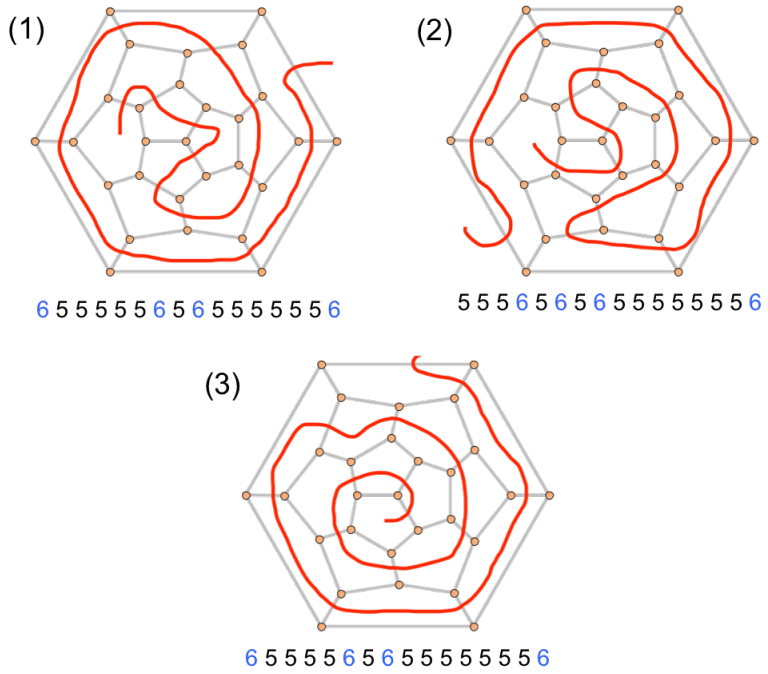

Figure 15 The three valid sequences and corresponding spiral graphs. (3) is the canonical sequence, which has the smallest value.

Another advantage of the spiral algorithm is that the symmetries can also be calculated simultaneously. In the symmetry theory, there is a simple relation,[14] which is:

$$N_t = N_s\,|G|$$

N_t is the total number of spirals unwound successfully. N_s is the number of valid symmetry-distinct spirals. G is the order of the point group of the fullerene isomer, which is a key factor to determine the symmetry. For example, I_h-C_{60} has 120 valid spirals, so N_t = 120. Among all these spirals, we can find there is only one symmetry-distinct spiral, which means N_s = 1. The isomer's point group order is 120. There is a problem, as mentioned above, that the spiral may fail. However, the result does not count on this. For example, the D_2-C_{28} isomer can be unwound successfully in 164 spirals and forty-one of those are symmetry-distinct. So the point group order of this isomer is 4, which points to D_2.

3.4 Generalized Fullerene Formation Program Based on Spiral Algorithm

This chapter demonstrates the way I design the generalized algorithm. Although the provided program includes several components and various algorithm, the generalized spiral algorithm is the most important one.

The first fullerene generation program is provided by Fowler and Manolopoulos in their book "AN ATLAS OF FULLERENES".[14] Following the first generation program, several programs were released and improvements have been made both on the generation algorithm and the program design. These programs are more efficient and provide many powerful capabilities, such as calculating coordinates and drawing structure graphs. [14,49,51,55] However, the ideas of these programs are all based on the spiral algorithm. Nowadays, non-classical fullerenes, especially heptagon-containing fullerenes, become the hot spot of fullerene research. More and more investigations show that heptagon-containing fullerenes play significant roles in fullerene formations. A program with the capability of generating non-classical fullerenes is urgently needed.

The first modification is made on the calculation of the numbers of faces belonging to different categories. In the original spiral program, which handles the classical fullerenes, we can obtain the numbers of pentagons and hexagons by simply applying the formulas mentioned in last chapter. A classical fullerene with n carbon atoms contains exactly 12 pentagons and n/2-10 hexagons, respectively. With the introduction of unusual faces, polygons besides pentagons and hexagons, we add the effects of these faces into the **Equation 1.** For example, a non-classical fullerene with heptagons satisfies the following equations:

$$\begin{cases} Relation\ between\ faces\ and\ edges: \\ \dfrac{(5p+6h+7hp)}{3} = n \\ Euler's\ polyhedron\ formula: \\ hp + p + h - \dfrac{n}{2} = 2 \end{cases}$$

where p, h and hp stand for the numbers of pentagons, hexagons and heptagons, respectively. Obviously, the solutions of hp, p and h depend on each other. By eliminating one unknown at one time, we can find the relations between two arbitrary unknowns:

$$\begin{cases} hp = p - 12 \\ h + 2hp = \dfrac{n}{2} - 10 \\ 2p + h = \dfrac{n}{2} + 14 \end{cases}$$

One of the interesting conclusions is if one heptagon is introduced, the number of pentagons will increase by one. For example, one heptagon-containing fullerenes have thirteen pentagons, while a classical fullerene contains twelve pentagons.

The first component of the generalized spiral program, a function output all possible sequences, is designed based on the formulas mentioned above and permuting. To make the comments clearer, I used an example, C_{60} with one

heptagon, to demonstrate the details of this process. The fullerene contains thirteen pentagons, one heptagon and eighteen hexagons. So there are:

$$\frac{32!}{13!\,18!\,1!}$$

possibilities for the sequences. In this case, we can firstly choose fourteen positions for the heptagon and pentagons, and then choose one position from these fourteen positions. Obviously,

$$\frac{32!}{13!\,18!\,1!} = \binom{32}{14}\binom{14}{1}$$

Similarly, we can apply this method, choosing positions for faces of different categories step by step, on more complex situations.

The second component of the program is design to check whether a sequence is valid for constructing a fullerene graph. Since classical fullerenes involve merely pentagons and hexagons, the original spiral program only need to consider two kinds of faces. To enable the program to generate non-classical fullerenes, the design should be modified to accommodate polygons of other categories. As mentioned above, although the spiral program is specifically designed for classical fullerenes, the spiral algorithm is not. The general idea can also be applied on non-classical fullerenes. So we can generalize the program by setting parameters dynamically based on the types of faces without changing the algorithm.

The last component of the original spiral program unwinds the fullerene's dual and checks whether the input sequence is a canonical sequence. At the meantime, this subroutine calculates the point groups of fullerene isomers based on the symmetry decision tree mentioned above. As commonly known, classical fullerenes are restricted to 28 possible point groups,[56] which simplifies their identification. Similarly, we can also find limitations for non-classical fullerenes. A

polygon can only introduce new n-fold rotations that it has to the fullerenes. For example, heptagons introduce C_1 and C_7 and squares introduce C_1, C_2 and C_4. According to this rule, the possible number of point groups of non-classical fullerenes is still in reasonable range which we can handle by programs.

Considering that most quantum calculations of fullerenes use coordinates of atoms as the starting point, I add the new function that can generate coordinate files along with the structure generation process. We can first obtain the connections of atoms and then apply mathematics on them to get the coordinates. It is too difficult to calculate atom connections directly from the fullerene graph. However, as discussed in last chapter, we can first convert a fullerene graph to its dual and then calculate the face connections of the dual, which are atom connections of the fullerene. Other functions such as file format transformation are also added into the program. The aim of the new program is to solve problems of topologies and mathematics, and help users to focus on properties of fullerenes respect to chemistry.

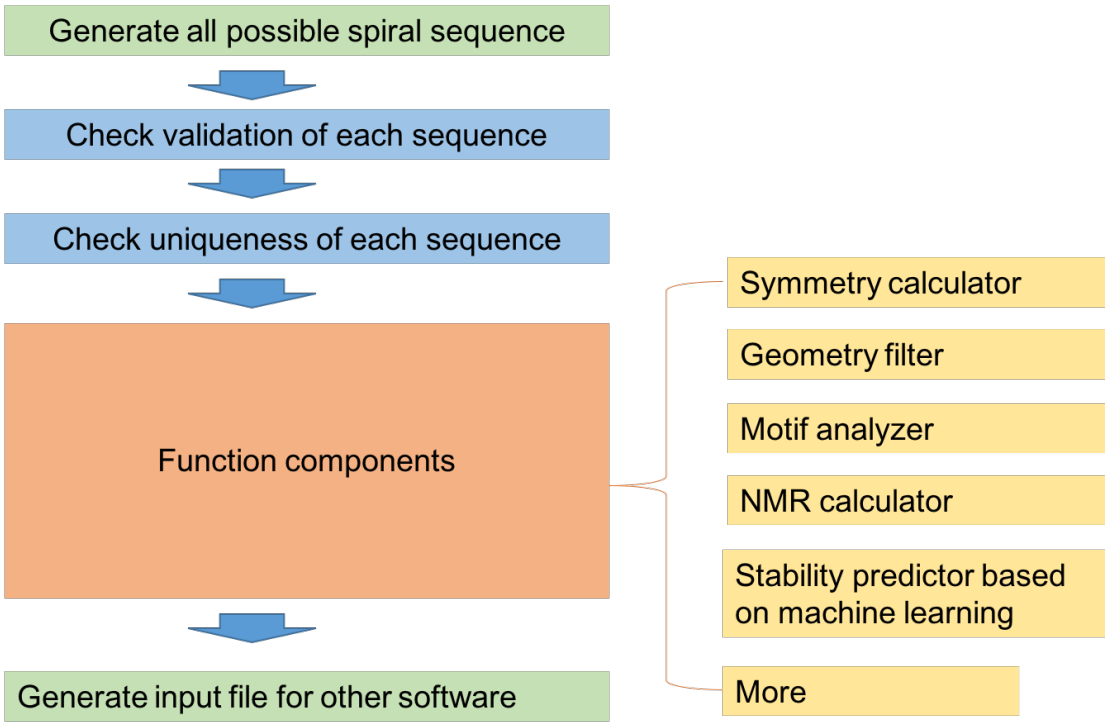

Figure 16 The main structure and the accessory components of the generalized spiral program

3.5 Research for Single-Heptagon-Containing Fullerenes in the Range C_{30}-C_{38} Structures and Symmetries

The aim of designing the generalized spiral program is for investigations of heptagon-containing fullerene. In this chapter, I give a preliminary idea for the applications of my new program.

According to the original spiral program, the first, namely the smallest classical fullerene is C_{20} and there is only one isomer, which belongs to I_h point group. [57] The isomer is comprised of 12 pentagons and as suggested by the IPR, it is unstable because of the adjacent pentagons. Similarly, small fullerenes, from C_{20} to C_{58}, are also predicted as unstable, due to the same reason. Although C_{20} and C_{36} have been reported that they could be obtained as a covalently bonded cluster-

assembled material in the solid state,[58,59] none of fullerenes in this range has been synthesized in the form of empty fullerene cages or endohedral fullerenes. As for the single heptagon containing fullerenes, the first isomer is C_{30} with C_1 symmetry. The numbers of isomers in the range of C_{30}-C_{48} are listed in **Table 1**.

	# of classical isomers	# of 1 heptagon isomer
C30	3	1
C32	6	2
C34	6	8
C36	15	16
C38	17	42
C40	40	89
C42	45	199
C44	89	388
C46	116	763
C48	199	1429

Table 1 The numbers of classical isomers and single-heptagon-containing fullerene isomers in the range of C_{30}-C_{38}.

Obviously, there are more possible isomers for single-heptagon-containing fullerenes than classical ones. It can be explained that the introduction of one heptagonal face brings more possibilities for the sequence permutations. More possible permutations always mean more valid sequence, although it cannot be proved strictly.

One of the reasons why fullerenes in this range are interesting is that the number of isomers is not too large and we can apply comprehensive studies on all isomers to find fundamental principles. As we all know for classical fullerenes, the relative energies can be estimated based on the cage topological structures. In more detail, the relative energies are determined by a group of particular motifs. Wang et al.[60] reported their investigations on the relation between cage topologies and endohedral metallofullerenes in 2015. According to their research, the total energies of EMFs, C_{2n}^{q} (2n = 68-104, q = 0, -2, -4, -6) are calculated at SCC-DFTB level and fitted in terms of seven motifs. They demonstrated that the energies calculated from the fitting equation were satisfactory compared to the results obtained by the calculations based on DFT. Due to the result, the relative energies can be produced by simply counting the numbers of seven key structural motifs. It can be used to select stable fullerene isomers before the quantum calculations and experiments.

For single-heptagon-containing fullerenes, we can find an analogue of this energy-motif rule. By inspecting the energies and structural motifs of corresponding single-heptagon-containing fullerenes, we know that the relative energies are determined by two categories of motifs, heptagon based and pentagon based motifs. (**Figure 17**) The heptagon-based motifs stabilize the fullerenes and on the other hand the pentagon-based motifs destabilize the fullerenes. However, to find an accurate formula based on these motifs, more work still needs to be done.

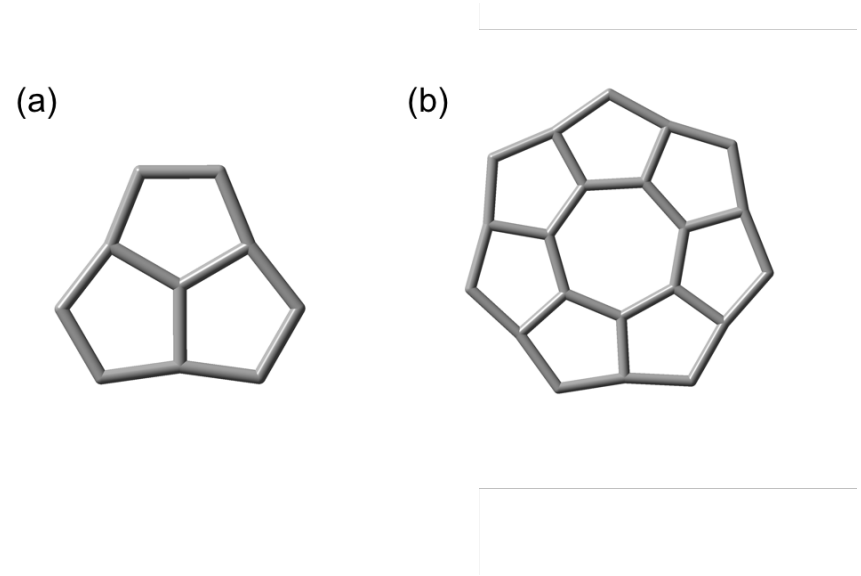

Figure 17 The motifs of heptagon containing fullerenes that determine the relative energies. (a) shows pentagon based motifs, which decrease the stability. (b) An example of heptagon based motifs, which increase the stability.

Other than finding the energy-motif rule, we can also find the relation between single-heptagon-containing fullerenes and classical fullerenes by applying comprehensive studies on the range of C_{30}-C_{38}. As far as we know, small fullerenes are not stable, due to the strain effects and the fused pentagons.[61] For one heptagon-containing fullerenes, fused pentagons may appear around the heptagonal face and these motifs will stabilize the fullerenes. The question is if the introduction of one heptagon makes these fullerenes more stable. The comparison between the single-heptagon-containing fullerenes and classical fullerenes of the same size shows that single-heptagon-containing fullerenes are not better in stability than classical ones. The reason may be complex. The motifs, heptagons surrounded by pentagons, are preferred because the electron distribution enhances the aromaticity. However, they will also bring extra strains for small

fullerenes. One of the satisfactory explanations is that the structural strains, rather than the aromaticity, is the dominant factor of the relative stabilities. [62]

Another key observation for single-heptagon-containing fullerenes is that most of them have low symmetries. The single heptagonal face limits the symmetry operations that can be applied on the cages. Symmetries are one of the essential characteristics of fullerenes. Although the role of symmetries has not been understood thoroughly, more and more investigations point out that the symmetries of fullerenes have close relationship with the properties, such as kinetic stabilities. Further research is needed to reveal the mystery of symmetries.

3.6 Research for Two-Heptagon-Containing Fullerenes with High Symmetries

As commented in the last chapter, single-heptagon-containing fullerenes belong to low symmetric point groups and most of them are even asymmetric. The simplest way to obtain high symmetric heptagon-containing fullerenes is to introduce another heptagonal face.[63] Based on the two-heptagon-containing isomers generated by the generalized spiral program, some of these isomers are found to have high symmetries, such as D_{7h} and D_{7v}. More interesting, there is a series of isomers, with symmetries of D_{7h} and D_{7v} alternately. These cages start at C_{28} and formed by adding fourteen carbon atoms each time. As illustrated in **Figure 18**, fullerenes belonging to this group are cylinder-like, whose pentagons are connected with the two heptagons. These structures are suggested by the generalized isolated pentagon rules,[1] as well as the energy-motif rule discussed above, to be preferred in thermodynamic stabilities. However, these structures also bring extra strains if the cages are large enough.

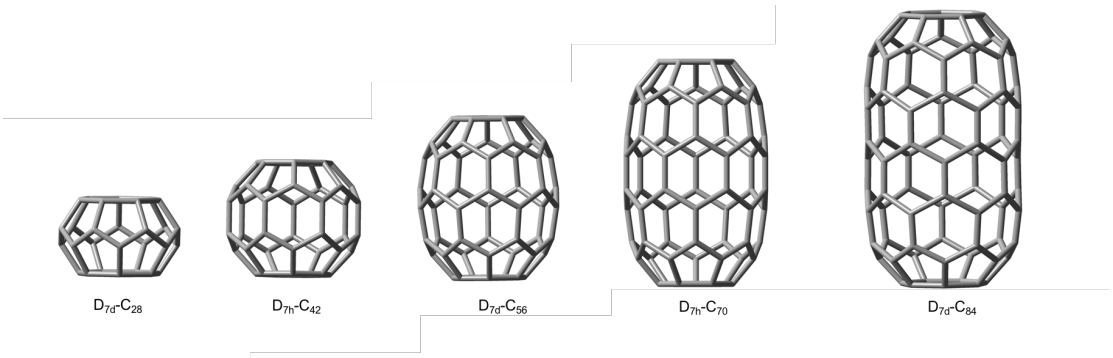

Figure 18 The structures of a series of two heptagon containing fullerenes with D_{7h} and D_{7d} symmetries alternately.

Another reason why these fullerenes are essential is that their unusual structures are good certifiers for the template synthesis[64] of metallofullerenes. Template synthesis is widely applied in many areas, especially organic and inorganic synthesis.[65] One of the most successful examples is the template approach for the synthesis of crown ethers.[66] The synthesis of a given crown ether is based on a metal atom of the right size. For example, 18-Crown-6 can be synthesized using potassium as the template and 15-Crown-5 can be synthesized using sodium. Similarly, metallofullerenes can also be synthesized by the same technique and the templates are chosen based of the size of target isomers. Fullerenes belonging to the group mentioned above have the same diameter but different heights. So the corresponding templates also vary in lengths. We can select a series of metal clusters and calculate the energies of corresponding metallofullerenes to check whether the template synthesis is feasible.

4 Conclusions

With the program provided in this thesis, it is possible now to generate non-classical fullerene isomers and to assign point group for each isomer. The

generalized spiral program also involves the function that output coordinates of carbon atoms. From these coordinates, it is convenient to draw the structures as well as to obtain input files for most quantum chemistry software. A group of scripts are also provided to generate input files for common software automatically. These tools save users from solving unfamiliar topological and mathematic problems. They facilitate the investigations on non-classical fullerenes and help scientists focus on chemical aspects. For the convenience of further modifications, the theories behind the generalized spiral program are demonstrated.

Based on the generalized spiral program, potential investigations on single-heptagon-containing fullerenes and on two-heptagon-containing fullerenes are proposed. By traversing the energies of single heptagon containing fullerenes and associating the energies to a group of motifs, we can fit a simple linear formula to estimate relative energies of heptagon-containing fullerenes without doing quantum calculations. Most of single-heptagon-containing fullerenes are asymmetric, so another heptagonal face is introduced to form high-symmetric fullerenes. The first two-heptagon-containing fullerene is C_{28} with D_{7d} symmetry. Interestingly, then we can construct a series of high-symmetric ones that have D_{7d} and D_{7h} alternately, by adding 14 carbons each time. These isomers are good certifiers for template synthesis of metallofullerenes because of their unique structures.

In short, this article provides necessary programs and demonstrates related topological theories to simplify the preparations for investigations of non-classical fullerenes. All programs mentioned are attached in appendix along with examples.

5 References:

(1) Ayuela, A.; Fowler, P.; Mitchell, D.; Schmidt, R.; Seifert, G.; Zerbetto, F. *The Journal of Physical Chemistry* **1996**, *100*, 15634.

(2) Smeal, T.; Binetruy, B.; Mercola, D. A.; Birrer, M.; Karin, M. **1991**.

(3) Iijima, S.; Ichihashi, T.; Ando, Y. *Nature* **1992**, *356*, 776.

(4) Iijima, S. *Nature* **1991**, *354*, 56.

(5) Hummelen, J. C.; Prato, M.; Wudl, F. *Journal of the American Chemical Society* **1995**, *117*, 7003.

(6) Prato, M.; Lucchini, V.; Maggini, M.; Stimpfl, E.; Scorrano, G.; Eiermann, M.; Suzuki, T.; Wudl, F. *Journal of the American Chemical Society* **1993**, *115*, 8479.

(7) Zhang, X.; Romero, A.; Foote, C. S. *Journal of the American Chemical Society* **1993**, *115*, 11024.

(8) Birkett, P. R.; Avent, A. G.; Darwish, A. D.; Kroto, H. W.; Taylor, R.; Walton, D. R. *Journal of the Chemical Society, Chemical Communications* **1995**, 1869.

(9) Birkett, P. *J. Chem. Soc., Chem. Commun* **1995**, *683*.

(10) Fowler, P.; Heine, T.; Mitchell, D.; Orlandi, G.; Schmidt, R.; Seifert, G.; Zerbetto, F. *Journal of the Chemical Society, Faraday Transactions* **1996**, *92*, 2203.

(11) Warshel, A.; Karplus, M. *Journal of the American Chemical Society* **1972**, *94*, 5612.

(12) Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. *Phys Rev B* **1995**, *51*, 12947.

(13) Fowler, P. W.; Manolopoulos, D. E.; Ryan, R. P. *Carbon* **1992**, *30*, 1235.

(14) Fowler, P. W.; Manolopoulos, D. E. *An atlas of fullerenes*; Clarendon Press ;
Oxford University Press: Oxford
New York, 1995.

(15) Manolopoulos, D. E.; May, J. C.; Down, S. E. *Chemical physics letters* **1991**, *181*, 105.

(16) Stewart, J. *J. Comput. Chem* **1993**, *14*, 1301.

(17) Dewar, M. J.; Thiel, W. *Journal of the American Chemical Society* **1977**, *99*, 4899.

(18) Dewar, M. J.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. *Journal of the American Chemical Society* **1985**, *107*, 3902.

(19) Troshin, P. A.; Avent, A. G.; Darwish, A. D.; Martsinovich, N.; Ala'a, K.; Street, J. M.; Taylor, R. *Science* **2005**, *309*, 278.

(20) Taylor, R. *Chemistry-A European Journal* **2001**, *7*, 4074.

(21) Hu, Y. H.; Ruckenstein, E. *The Journal of chemical physics* **2003**, *119*, 10073.

(22) Tang, L.; Sai, L.; Zhao, J.; Qiu, R. *Computational and Theoretical Chemistry* **2011**, *969*, 35.

(23) Ioffe, I. N.; Chen, C.; Yang, S.; Sidorov, L. N.; Kemnitz, E.; Troyanov, S. I. *Angewandte Chemie* **2010**, *122*, 4894.

(24) Zhang, Y.; Ghiassi, K. B.; Deng, Q.; Samoylova, N. A.; Olmstead, M. M.; Balch, A. L.; Popov, A. A. *Angewandte Chemie* **2015**, *127*, 505.

(25) Tan, Y.-Z.; Xie, S.-Y.; Huang, R.-B.; Zheng, L.-S. *Nat Chem* **2009**, *1*, 450.

(26) Stevenson, S.; Mackey, M. A.; Stuart, M. A.; Phillips, J. P.; Easterling, M. L.; Chancellor, C. J.; Olmstead, M. M.; Balch, A. L. *Journal of the American Chemical Society* **2008**, *130*, 11844.

(27) Wang, C.-R.; Kai, T.; Tomiyama, T.; Yoshida, T.; Kobayashi, Y.; Nishibori, E.; Takata, M.; Sakata, M.; Shinohara, H. *Nature* **2000**, *408*, 426.

(28) Zhang, J. Y.; Bowles, F. L.; Bearden, D. W.; Ray, W. K.; Fuhrer, T.; Ye, Y. Q.; Dixon, C.; Harich, K.; Helm, R. F.; Olmstead, M. M.; Balch, A. L.; Dorn, H. C. *Nat Chem* **2013**, *5*, 880.

(29) Murry, R. L.; Strout, D. L.; Odom, G. K.; Scuseria, G. E. **1993**.

(30) Eckhoff, W. C.; Scuseria, G. E. *Chemical physics letters* **1993**, *216*, 399.

(31) Hernández, E.; Ordejón, P.; Terrones, H. *Phys Rev B* **2001**, *63*, 193403.

(32) Stone, A. J.; Wales, D. J. *Chem Phys Lett* **1986**, *128*, 501.

(33) Miyake, Y.; Minami, T.; Kikuchi, K.; Kainosho, M.; Achiba, Y. *Molecular Crystals and Liquid Crystals* **2000**, *340*, 553.

(34) Sun, G.; Kertesz, M. *Chemical physics* **2002**, *276*, 107.

(35) Ioffe, I. N.; Mazaleva, O. N.; Sidorov, L. N.; Yang, S.; Wei, T.; Kemnitz, E.; Troyanov, S. I. *Inorganic chemistry* **2013**, *52*, 13821.

(36) I Troyanov, S.; Kemnitz, E. *Current Organic Chemistry* **2012**, *16*, 1060.

(37) Zhang, J.; Fuhrer, T.; Fu, W.; Ge, J.; Bearden, D. W.; Dallas, J.; Duchamp, J.; Walker, K.; Champion, H.; Azurmendi, H. *Journal of the American Chemical Society* **2012**, *134*, 8487.

(38) Fu, W.; Zhang, J.; Fuhrer, T.; Champion, H.; Furukawa, K.; Kato, T.; Mahaney, J. E.; Burke, B. G.; Williams, K. A.; Walker, K. *Journal of the American Chemical Society* **2011**, *133*, 9741.

(39) Zhang, J.; Bearden, D. W.; Fuhrer, T.; Xu, L.; Fu, W.; Zuo, T.; Dorn, H. C. *Journal of the American Chemical Society* **2013**, *135*, 3351.

(40) Dunk, P. W.; Kaiser, N. K.; Hendrickson, C. L.; Quinn, J. P.; Ewels, C. P.; Nakanishi, Y.; Sasaki, Y.; Shinohara, H.; Marshall, A. G.; Kroto, H. W. *Nature communications* **2012**, *3*, 855.

(41) Ioffe, I. N.; Chen, C. B.; Yang, S. F.; Sidorov, L. N.; Kemnitz, E.; Troyanov, S. I. *Angew Chem Int Edit* **2010**, *49*, 4784.

(42) Berné, O.; Tielens, A. G. *Proceedings of the National Academy of Sciences* **2012**, *109*, 401.

(43) Goroff, N. S. *Accounts Chem Res* **1996**, *29*, 77.

(44) Hernandez, E.; Ordejon, P.; Terrones, H. *Phys Rev B* **2001**, *63*.

(45) Wang, W.-W.; Dang, J.-S.; Zheng, J.-J.; Zhao, X. *The Journal of Physical Chemistry C* **2012**, *116*, 17288.

(46) Cundy, H.; Rollett, A.; Oxford University Press.

(47) Manolopoulos, D. E.; Fowler, P. W. *The Journal of chemical physics* **1992**, *96*, 7603.

(48) Fowler, P. W.; Cremona, J. E.; Steer, J. *Theoretica chimica acta* **1988**, *73*, 1.
(49) Schwerdtfeger, P.; Wirz, L. N.; Avery, J. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2015**, *5*, 96.
(50) Brinkmann, G.; Dress, A. W. M. *J Algorithm* **1997**, *23*, 345.
(51) Schwerdtfeger, P.; Wirz, L.; Avery, J. *Journal of computational chemistry* **2013**, *34*, 1508.
(52) Schlegel, V. *Theorie der homogen zusammengesetzten Raumgebilde*; Für die Akademie in commission bei W. Engelmann in Leipzig, 1883.
(53) Manolopoulos, D. E.; Fowler, P. W. *Chem Phys Lett* **1993**, *204*, 1.
(54) Fowler, P. W.; Graovac, A.; Pisanski, T.; Žerovnik, J. *A generalized ring spiral algorithm for coding fullerenes and other cubic polyhedra*; University of Ljubljana. Department of Mathematics, 1998.
(55) Babic, D.; Balaban, A. T.; Klein, D. J. *Journal of chemical information and computer sciences* **1995**, *35*, 515.
(56) Deza, M.; Dutour Sikiric, M.; Fowler, P. W. *Match* **2009**, *61*, 589.
(57) Saito, M.; Miyamoto, Y. *Phys Rev Lett* **2001**, *87*, 035503.
(58) Prinzbach, H.; Weiler, A.; Landenberger, P.; Wahl, F.; Wörth, J.; Scott, L. T.; Gelmont, M.; Olevano, D.; Issendorff, B. v. *Nature* **2000**, *407*, 60.
(59) Koshio, A.; Inakuma, M.; Sugai, T.; Shinohara, H. *Journal of the American Chemical Society* **2000**, *122*, 398.
(60) Wang, Y.; Díaz-Tendero, S.; Martin, F.; Alcami, M. *Journal of the American Chemical Society* **2016**.
(61) Austin, S.; Fowler, P.; Manolopoulos, D.; Orlandi, G.; Zerbetto, F. *The Journal of Physical Chemistry* **1995**, *99*, 8076.
(62) Díaz-Tendero, S.; Alcamí, M.; Martín, F. *Chem Phys Lett* **2005**, *407*, 153.
(63) Fowler, P. W.; Morvan, V. *Journal of the Chemical Society, Faraday Transactions* **1992**, *88*, 2631.
(64) Liu, H.; Li, Y.; Jiang, L.; Luo, H.; Xiao, S.; Fang, H.; Li, H.; Zhu, D.; Yu, D.; Xu, J. *Journal of the American Chemical Society* **2002**, *124*, 13370.
(65) Badri, Z.; Foroutan-Nejad, C.; Rashidi-Ranjbar, P. *Computational and Theoretical Chemistry* **2013**, *1009*, 103.
(66) Lehn, J.-M. *Supramolecular chemistry*; Vch, Weinheim, 1995; Vol. 1.

6 Appendix:

A. Methods for constructing fullerene cages from the output of Gspiral program

The spiral sequence is the fingerprint of its corresponding fullerene structure. It has been approved in **Chapter 3.4** that there is a one-to-one relationship between

fullerene's structure and its spiral sequence. So I demonstrate the way to reconstruct fullerene structures based on spiral sequences.

The spiral sequence is the fingerprint of its corresponding fullerene structure. It has been approved in **Chapter 3.4** that there is a one-to-one relationship between fullerene's structure and its spiral sequence. So I demonstrate the way to reconstruct fullerene structures based on spiral sequences.

The reconstruction of a fullerene structure can be done in several ways. I first provide a straightforward method that can be implemented using most chemistry structure making software.

1. Pick one ring each time from the ring sequence and put it in the position indicated by spiral algorithm, which is discussed in **chapter 3.3.**

2. Use the automatic structure adjust function of the software, then the 2-D map (fullerene graph) will become 3-D structure and it is a good start for calculations.

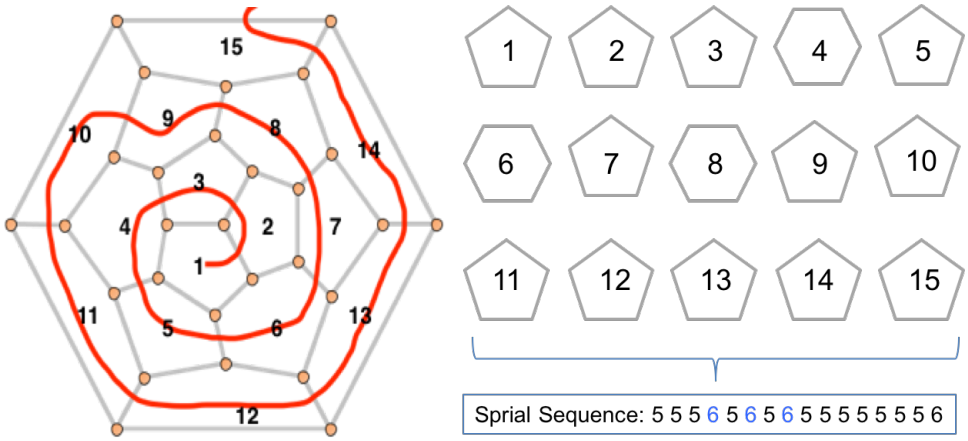

Figure 1 Reconstruction of fullerene structure based on sprial sequence. The numbers in the center of faces indicate the order in spiral sequence. The last face in the sprial sequence has the same shape as the border of the spiral graph(left figure).

I have verified the reconstruction method by a group of figures listed in appendix using gaussview.

Another way to reconstruct the fullerene structures are based on adjacency matrices that are generated by my program. Several open source programs are available for this convertion. In this thesis, EMBED[50], developed by Gunnar Brinkmann, is applied. The algorithem of the EMBED code is to embed each vertex on spherical surface. The command for applying EMBED is as following:

embed < infile > outfile

The infile is the adjacency matrix calculated by the Gsprial program and the outfile is the Cartesian coordinates.

B. Atlas of single- and two- heptagon-containing fullerenes

C. Generalized Spiral Program(Gspiral.cc)