

## PENGGAMBARAN RANTAI KARBON DENGAN MENGUNAKAN SIMPLIFIED MOLECULAR INPUT LINE SYSTEM (SMILES)

**Hartarto Junaedi**

Sistem Informasi

Sekolah Tinggi Teknik Surabaya

[Hartarto.j@gmail.com](mailto:Hartarto.j@gmail.com)

### ABSTRAK

Ilmu kimia mempunyai beberapa subdisiplin salah satunya adalah rantai karbon. Rantai karbon merupakan sekumpulan atom karbon yang berikatan di dalam sebuah senyawa. Penggambaran rantai karbon dari suatu senyawa dapat digunakan untuk menentukan golongan dari senyawa tersebut. Siswa-siswi di Sekolah Menengah Atas diajarkan menggambarkan rantai karbon dari beberapa golongan senyawa, seperti alkana, alkanol, ester dan benzena.

SMILES merupakan suatu notasi kimia yang didesain untuk memproses informasi-informasi kimia modern. SMILES diciptakan oleh David Weininger pada tahun 1980 untuk penggunaan komputer, dan dengan menggunakan konsep dari sebuah graph. Atom dianggap sebagai nodes pada graph, sedangkan ikatan antar atom dianggap sebagai edge pada graph. Notasi SMILES terdiri dari karakter-karakter ASCII, sehingga sebuah notasi SMILES dapat disimpan ke dalam sebuah variabel string. Dengan demikian, notasi SMILES lebih mudah diproses dan tidak memerlukan tempat penyimpanan yang berukuran besar dalam komputer.

Pada penelitian ini akan dilakukan penggambaran rantai karbon dari beberapa golongan senyawa dengan menggunakan notasi SMILES atau markup language sebagai input. Beberapa golongan senyawa yang dapat digambarkan rantai karbonnya adalah alkana, alkena, alkuna, alkanol, eter, aldehida, keton, asam karboksilat, ester, haloalkana dan benzena. Sebelum penggambaran dilakukan, akan dilakukan pengambilan informasi-informasi yang terdapat dalam input yang diberikan oleh user. Bila input berupa notasi SMILES, maka akan dilakukan pemotongan input agar informasi-informasi yang terkandung di dalamnya dapat diambil.

Dari informasi yang didapat maka akan dibuat markup language dari rantai karbon tersebut. Kemudian akan dilakukan penggambaran rantai karbon. Gambar rantai karbon dan markup language yang dihasilkan, serta notasi SMILES yang baru saja digunakan (bila input menggunakan notasi SMILES) dapat disimpan oleh user..

Kata kunci: SMILES, Rantai Karbon, Chemical Markup Language

### ABSTRACT

*Chemistry has some branches, one of them is carbon chain. Carbon chain is a series of bonded karbon atoms in a compound. Drawing a carbon chain from a compound can be used to determine the class of that compound. Students in Senior High*

*School are taught how to draw carbon chain from some compound classes, such as alkanes, alkanol, esther and benzene.*

*SMILES is a chemical notation which design to process modern chemical informations. SMILES created by David Weininger in 1980 for computer using, and using concept of a graph. Atoms reputed as nodes in a graph, whereas bonds between atoms reputed as edges in a graph. SMILES notation consists of ASCII characters, so that a SMILES notation can be saved in a string variabel. Thereby, SMILES notation is easier to process and does not need a huge save space in a computer.*

*In research drawing a carbon chain from some compound classes will be done using SMILES notation or markup language as the input. Some compound classes that can be drawn are alkanes, alkenes, alkynes, alkanol, ether, ketone, aldehyde, carboxylic acid, esther, haloalkane, benzene. Before the drawing is done, information fetching will be done to the input which is given by user. If the input is a SMILES notation, the input will be cut so that information fetching can be done.*

*From the informations obtained markup language of that carbon chain will be created. Afterward drawing a carbon chain will be done. The carbon chain picture and markup language which is producted, as well as SMILES notation which is used as the input (if SMILES notation is used as input) can be saved by user.*

*Keywords : SMILES, Carbon Chain, Chemical Markup Language*

## 1. PENDAHULUAN

Rantai karbon merupakan salah satu cabang dalam ilmu kimia. Rantai karbon adalah sekumpulan atom karbon yang berikatan di dalam sebuah senyawa. Siswa-siswi Sekolah Menengah Atas kadang-kadang mengalami kesulitan dalam menggambarkan rantai karbon dalam suatu senyawa. Selain itu dibutuhkan penyimpanan informasi-informasi kimia, salah satunya adalah rantai karbon, dalam berbagai media penyimpanan (seperti komputer) yang efisien dan mudah.

Pada penelitian ini, dilakukan pembuatan program yang dapat menggambarkan rantai karbon dari suatu senyawa. Pada program ini, user dapat memberikan input senyawa dengan menggunakan notasi SMILES atau markup language. User dapat menyimpan gambar rantai karbon, markup language dan input notasi SMILES yang digunakan.

## 2. RANTAI KARBON

Atom-atom (sebagian besar merupakan atom karbon) berikatan membentuk suatu rantai karbon. Sebuah atom dapat berikatan dengan satu atau beberapa atom lainnya, asalkan jumlah total ikatan dari atom tersebut (jumlah semua ikatannya dengan atom-atom lain yang berikatan dengan atom tersebut) tidak melebihi total elektron valensi yang dimiliki oleh atom tersebut. Elektron valensi adalah elektron yang dapat digunakan untuk membentuk ikatan kimia (ikatan antara sebuah atom dengan atom yang lain). Di dalam rantai karbon terdapat tiga jenis ikatan antar atom, yaitu ikatan tunggal (dilambangkan dengan '-'), ikatan rangkap (dilambangkan dengan '=') dan ikatan rangkap tiga (dilambangkan dengan '≡').

Untuk dapat menggambarkan sebuah rantai karbon, perlu diketahui terlebih dahulu rumus molekul dari senyawa karbon yang akan digambarkan. Pada penggambaran rantai karbon, kadang-kadang atom H tidak disertakan dalam gambar rantai karbon. Hal tersebut tergantung pada keinginan penggambar (bersifat

opsional). Jumlah atom H yang berikatan dengan suatu atom dilambangkan dengan angka setelah atom H tersebut digambarkan.

Rantai karbon ada beberapa jenis. Di dalam penelitian ini akan digambarkan beberapa jenis rantai karbon, yaitu rantai karbon linear, rantai karbon bercabang, serta rantai karbon siklik. Rantai karbon linear adalah rantai karbon yang tidak mempunyai cabang di dalamnya. Selain itu, pada rantai karbon linear tidak terdapat ikatan antar atom yang membentuk lingkaran (kedua atom yang berada di ujung rantai tidak saling berikatan).

Rantai karbon bercabang adalah rantai karbon yang mempunyai cabang di dalamnya, tetapi kedua atom yang berada di ujung rantai tidak saling berikatan. Cabang yang ada pada suatu rantai karbon bisa lebih dari satu buah. Di dalam cabang tersebut bisa terdapat cabang yang lain (terdapat cabang di dalam cabang). Rantai karbon siklik adalah rantai karbon yang memutar (kedua atom yang berada di ujung rantai saling berikatan). Di dalam rantai karbon siklik bisa terdapat cabang. Cabang yang ada pada rantai karbon ini juga bisa lebih dari satu buah. Selain itu, di dalam cabang masih mungkin terdapat cabang lagi (cabang di dalam cabang).

Senyawa karbon digolongkan ke dalam beberapa golongan. Beberapa golongan senyawa karbon yang akan digambarkan dalam program dan juga diajarkan pada siswa-siswi SMA adalah alkana, alkena, alkuna, alkanol, eter, aldehida, keton, asam karboksilat, ester, haloalkana dan benzena. Pada golongan alkana, semua ikatan antar atom merupakan ikatan tunggal. Alkana tidak boleh mengandung ikatan rangkap, maupun ikatan rangkap tiga. Pada golongan alkena, harus terdapat sebuah ikatan rangkap antara dua buah atom karbon. Sedangkan atom-atom lainnya berikatan tunggal. Alkena tidak boleh mengandung ikatan rangkap tiga.

Pada golongan alkuna, harus terdapat sebuah ikatan rangkap tiga antara dua buah atom karbon. Sedangkan atom-atom lainnya berikatan tunggal. Alkuna tidak boleh mengandung ikatan rangkap dua. Alkanol dikenal juga dengan sebutan alkohol. Pada golongan alkanol, semua ikatan antar atom merupakan ikatan tunggal. Pada salah satu ujung rantai karbon golongan alkanol, harus terdapat sebuah atom oksigen (atom O) yang berikatan dengan sebuah atom C dan atom H (terdapat gugus fungsi  $-OH$ ). Eter dikenal juga dengan sebutan alkoksialkana. Pada golongan eter, semua ikatan antar atom merupakan ikatan tunggal. Pada rantai karbon golongan eter, harus terdapat sebuah atom O yang berikatan dengan dua buah atom C (terdapat gugus fungsi  $-O-$ ). Aldehida dikenal juga dengan sebutan alkanal. Pada golongan aldehida, semua ikatan antar atom merupakan ikatan tunggal kecuali sebuah atom O yang berikatan rangkap dengan atom C (terdapat gugus fungsi  $-CHO$ ) pada salah satu ujung rantai karbon.

Keton dikenal juga dengan sebutan alkanon. Pada golongan keton, semua ikatan antar atom merupakan ikatan tunggal kecuali sebuah atom O yang berikatan rangkap dengan sebuah atom C (terdapat gugus fungsi  $-CO-$ ). Asam karboksilat dikenal juga dengan sebutan asam alkanoat. Pada golongan asam karboksilat, semua ikatan antar atom merupakan ikatan tunggal kecuali sebuah atom C yang berikatan tunggal dengan sebuah atom O serta berikatan rangkap dengan sebuah atom O (terdapat gugus fungsi  $-COOH$ ) pada salah satu ujung rantai karbon. Ester dikenal juga dengan sebutan alkil alkanoat. Pada golongan ester, semua ikatan antar atom merupakan ikatan tunggal kecuali sebuah atom C yang berikatan tunggal dengan sebuah atom O, serta berikatan rangkap dengan sebuah atom O (terdapat gugus fungsi  $-COO-$ ) pada rantai karbon.

Pada golongan haloalkana, semua ikatan antar atom merupakan ikatan tunggal.

Pada rantai karbon golongan haloalkana, satu atau lebih atom H digantikan dengan atom halogen. Atom halogen tersebut adalah atom F (atom Fluorin), atom Cl (atom klorin), atom Br (atom bromin), dan atom I (atom yodium atau iodin). Benzena merupakan senyawa aromatik sederhana. Disebut senyawa aromatik, karena sebagian anggotanya mempunyai baru yang sedap. Pada golongan benzena, semua ikatan antar atom merupakan ikatan tunggal, atau ikatan rangkap. Pada rantai karbon golongan benzena, selalu terdapat enam buah atom C yang ujung-ujungnya saling berikatan (membentuk rantai karbon siklik). Ikatan yang ada pada atom-atom tersebut berselang seling antara ikatan tunggal dan ikatan ganda.

### 3. SIMPLIFIED MOLECULAR INPUT LINE SYSTEM (SMILES)

SMILES diciptakan oleh David Weininger pada akhir tahun 1980 dengan menggunakan konsep dari sebuah graph. SMILES adalah sebuah sistem notasi kimia yang didesain untuk memproses informasi-informasi kimia modern. Berdasarkan pada prinsip dari teori *molecular graph*, SMILES dapat melakukan spesifikasi struktur dengan tepat menggunakan tatabahasa yang cukup mudah dan umum. Atom dianggap sebagai *nodes* pada graph, sedangkan ikatan dianggap sebagai *edge* pada graph, untuk merepresentasikan sebuah molekul sebagai sebuah graph. Sistem notasi SMILES juga cocok untuk pemrosesan dengan menggunakan mesin yang mempunyai kecepatan sangat tinggi. Notasi SMILES melakukan spesifikasi molekul dengan menggunakan karakter-karakter ASCII untuk merepresentasikan simbol atom dan ikatan.

Penulisan SMILES bersifat casesensitif (membedakan antara huruf kapital dan huruf non kapital), jadi dalam menuliskannya harus benar-benar diperhatikan penggunaan huruf kapital dan huruf non kapitalnya. Pada umumnya atom dapat dilambangkan dengan menggunakan sebuah huruf saja. Akan tetapi terdapat beberapa atom yang dilambangkan dengan menggunakan lebih dari satu huruf. Aturan dalam penulisan atom yang dilambangkan dengan lebih dari satu huruf adalah huruf pertama dari atom tersebut dituliskan dengan menggunakan huruf kapital. Sedangkan untuk huruf yang mengikutinya dituliskan dengan menggunakan huruf non kapital. Untuk atom yang hanya dilambangkan dengan sebuah huruf, dituliskan dengan huruf kapital.

Di dalam rantai karbon terdapat tiga macam jenis ikatan antar atom, yaitu ikatan tunggal, ikatan rangkap, dan ikatan rangkap tiga. Ikatan tunggal dalam rumus molekul kimia dilambangkan dengan '-', ikatan rangkap dilambangkan dengan '=', dan ikatan rangkap tiga dilambangkan dengan '='. Sedangkan dalam SMILES ikatan tunggal dilambangkan dengan '-' (tanda strip). Khusus untuk ikatan tunggal lambang boleh tidak dituliskan. Ikatan ganda dilambangkan dengan '=' (tanda sama dengan), dan ikatan rangkap tiga dilambangkan dengan '#' (tanda kres). Rantai karbon ada yang berjenis linear (tidak mempunyai cabang), ada pula yang berjenis bercabang. Bila rantai karbon tersebut mempunyai cabang, maka sebelum menuliskan rantai karbon cabangnya diawali dengan menuliskan '(' (tanda kurung buka) terlebih dahulu. Kemudian dilanjutkan dengan menuliskan atom-atom penyusun cabangnya yang diakhiri dengan menuliskan tanda ')' (tanda kurung tutup). Sebuah cabang bisa mempunyai cabang lagi di dalamnya, dengan cara penulisan yang sama dengan penulisan cabang.

Selain itu, terdapat juga atom yang berikatan membentuk lingkaran atau disebut juga dengan siklik (atom yang berada di ujung saling berikatan). Untuk kasus tersebut digunakan angka untuk menghubungkan atom pada ujung yang satu dengan atom pada ujung yang lain dalam rantai karbon. Angka dituliskan setelah penulisan atom, misalnya 'C1'. Berikut ini adalah contoh penulisan notasi SMILES dari suatu rumus molekul:

Rumus molekul :

CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>

SMILES :

C1H2CH2CH2CH2CH2C1H2

Atau

C1CCCCC1

Atau

C1H2-CH2-CH2-CH2-CH2-C1H2

Atau

C1-C-C-C-C-C1

#### 4. MARKUP LANGUAGE

Markup language adalah sekumpulan data dari dalam sebuah file markup language yang digunakan untuk mendeskripsikan bagaimana seharusnya struktur yang digunakan atau ditampilkan dalam file markup language. Markup language yang dibuat ini pada dasarnya mirip dengan HTML (Hyper Text Markup Language) ataupun XML (eXtended Markup Language). Perbedaan yang ada adalah pada fungsi dan tag-tag yang digunakan di dalamnya. Markup language ini akan disimpan ke dalam sebuah file markup language. Sebuah file markup language hanya akan berisi sebuah markup language dari suatu rantai karbon. Ekstensi yang akan digunakan untuk file ini adalah '.mul'.

Markup language yang digunakan pada penelitian ini berfungsi untuk menyimpan data-data yang akan digunakan dalam proses penggambaran rantai karbon. Data-data yang terdapat di dalam markup language ini juga dapat digunakan sebagai input program. Data-data yang disimpan di dalam sebuah file teks markup language antara lain adalah data-data yang dimiliki masing-masing atom, data tentang warna atom serta warna ikatan antar atom yang akan digunakan pada saat penggambaran, nama rantai karbon, apakah atom H akan ditampilkan pada penggambaran, dan data jumlah total atom yang terdapat dalam suatu rantai karbon. Untuk memberikan warna atom menggunakan nama warna dalam bahasa Inggris, yang diawali dengan 'cl'.

Data masing-masing atom meliputi index atom tersebut, simbol atom, posisi x dan posisi y atom pada gambar, jumlah atom hidrogen yang menempel pada atom tersebut, jumlah total ikatan yang ada pada atom tersebut, serta index ikatan pada atom. Masing-masing ikatan mempunyai data nomor ikatan dengan atom ke index ikatannya, serta jumlah ikatan antara kedua atom tersebut. Berikut ini adalah format penulisan markup language tersebut:

<title>(nama rantai karbon)

<watom>(warna atom)

<wikatan>(warna ikatan)

<use>(Y atau N)

<jml>(jumlah total atom)

<idx\_atom>(index atom)

<nama>(simbol atom)

<xy>(posisi x penggambaran, posisi y penggambaran)

<H>(jumlah atom H yang menempel pada atom tersebut)

<jum>(jumlah total ikatan yang ada pada atom tersebut)

<idj>(index ikatan pada atom)

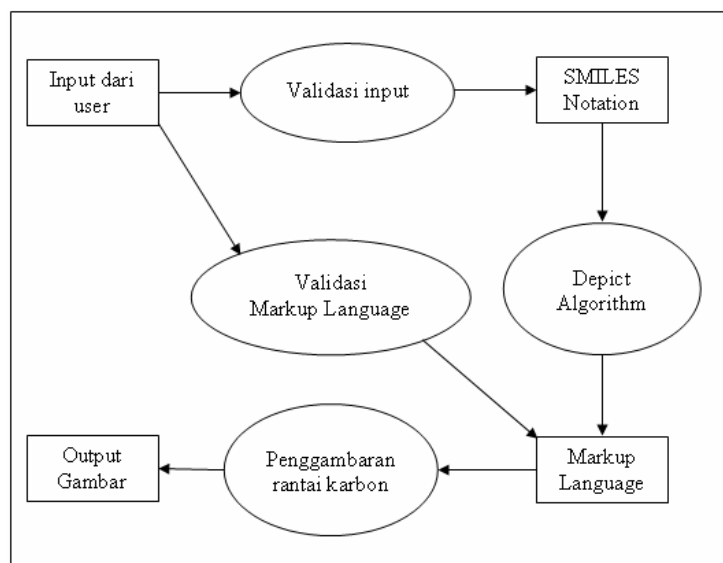
<no>(berikatan dengan atom ke index)

<hub>(jumlah ikatan kedua atom tersebut)

## 5. DESAIN ARSITEKTUR

Pada bagian ini akan diberikan penjelasan tentang desain arsitektur yang digunakan pada program penelitian yang akan dibuat. Pada gambar 1 dapat dilihat bahwa pertama-tama user akan diminta untuk memberikan input kepada program. Setelah input diberikan, akan dilakukan pengecekan apakah input yang diberikan tersebut telah valid. Input yang dapat diberikan oleh user sendiri ada dua macam. Pertama adalah input yang berupa *line notation* (input berupa sebuah string yang terdiri dari karakter-karakter ASCII). Sedangkan yang kedua adalah input yang berupa file markup language.

Untuk input jenis pertama (*line notation*) akan dilakukan pengecekan apakah input telah sesuai dengan ketentuan-ketentuan yang berlaku pada pembuatan notasi SMILES. Hal tersebut dikarenakan Input yang berupa *line notation* menggunakan aturan-aturan yang berlaku pada notasi SMILES (Simplified Molecular Input Line System). Disamping itu, input juga harus sesuai dengan ketentuan-ketentuan yang ada dalam ilmu kimia tentang pembuatan sebuah rumus molekul dari suatu senyawa. Input line notation ini dapat dituliskan dengan menggunakan salah satu dari empat bentuk. Pertama dengan menuliskan simbol ikatan tunggal ('-') dan juga atom H. Kedua dengan menuliskan simbol ikatan tunggal ('-') akan tetapi tidak menuliskan atom H. Ketiga dengan menuliskan atom H akan tetapi tidak menuliskan simbol ikatan tunggal ('-'). Keempat dengan tidak menuliskan simbol ikatan tunggal ('-') maupun atom H. Setelah data input valid, akan dibuatkan sebuah markup language yang sesuai dengan rantai karbon tersebut. Kemudian proses penggambaran rantai karbon akan dilakukan.



Gambar 1. Desain Arsitektur Program

### Algoritma 1 : setxy

- 1: partisi molekul ke dalam N Tree
- 2: IF terdapat siklik THEN
- 3: BEGIN
- 4:     tentukan letak pusat polygon
- 5:     tentukan letak masing-masing atom dalam rantai siklik agar membentuk polygon yang sempurna
- 6: END

```

7: pilih atom yang mempunyai ikatan terbanyak sebagai root
8: buat penelusuran DFS Tree dari root
9: FOR i := 1 TO jumlah atom DO
10: BEGIN
11:     tentukan sudut penggambaran atom
12:     tentukan x dan y penggambaran atom
13:     WHILE posisi atom bertumpukan dengan atom lain DO
14:         BEGIN
15:             ubah sudut penggambaran atom
16:             tentukan lagi posisi x dan y atom
17:         END
18: END
    
```

Sedangkan untuk input jenis kedua (file markup language), akan dilakukan pengecekan terhadap data-data yang terdapat di dalam file teks markup language tersebut. Kemudian akan dilakukan proses penggambaran output bila semua data yang diperoleh telah valid. Proses penggambaran akan dilakukan berdasarkan data-data valid yang telah didapat. Setelah proses penggambaran selesai dilakukan, maka user dapat memproses output yang telah didapat. Algoritma yang digunakan dalam pemrosesan data atom dapat dilihat pada algoritma 1. Sedangkan algoritma untuk penggambaran atom-atom dan ikatan-ikatan antar atom dalam rantai karbon dapat dilihat pada algoritma 2.

#### Algoritma 2 : depict0

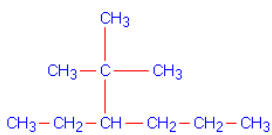
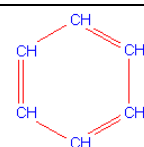
```

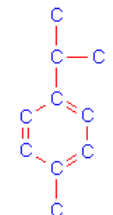
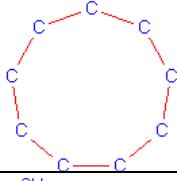
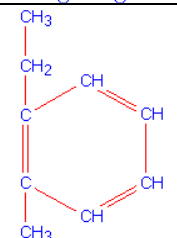
1: IF jumlah atom H > 0 THEN
2: BEGIN
3:     cetak 'H' dengan ukuran simbolatom
4:     IF jumlah atom H > 1 THEN
5:         cetak jumlah atom H dengan ukuran yang lebih kecil
6: END
7: IF jumlah ikatan < 2 THEN
8:     gambar sebuah garis antara kedua atom
9: IF (jumlah ikatan > 1) THEN
10:     gambar sepasang garis antara kedua atom
    
```

### 6. UJI COBA

Berikut adalah hasil uji coba penggambaran beberapa rantai karbon.

Tabel 1. Hasil Uji Coba

No	Input	Output
1	C-C-C-C-C	$\text{CH}_3 - \text{CH}_2 - \text{CH}_2 - \text{CH}_2 - \text{CH}_3$
2	CCC(C(C)(C)(C)) CCC	
3	C1=CC=CC=C1	

No	Input	Output
4	<chem>C1(C(C)C)=CC=C(C)C=C1</chem>	
5	<chem>C1CCCCCCC1</chem>	
6	<chem>C1=CC=CC(C)=C1(CC)</chem>	

## 7. PENUTUP

Dari hasil analisa terhadap penelitian yang telah dilakukan, terdapat beberapa kesimpulan yang dapat diambil yaitu :

1. Pemotongan input *line notation* serta preprocessing perlu dilakukan untuk menentukan jenis rantai karbon yang akan digambarkan.
2. Penggunaan backtracking untuk menentukan posisi penggambaran atom-atom dalam suatu rantai karbon cukup efisien.
3. Penggambaran rantai karbon menggunakan konsep suatu graph dengan atom sebagai node dan ikatan antar atom sebagai edge.
4. Notasi SMILES (Simplified Molecular Input Line System) yang digunakan dalam input berupa line notation cukup terstruktur dan mudah untuk diproses lebih lanjut dalam proses penggambaran suatu rantai karbon.
5. Markup language yang terbentuk dapat digunakan untuk mengetahui detail dari suatu rantai karbon, karena memuat semua informasi yang digunakan dalam penggambaran suatu rantai karbon.
6. SMILES tidak dapat menangani input dengan format seperti C<sub>4</sub>H<sub>10</sub> karena tidak terdapat notasi SMILES yang dapat mengatasi format tersebut.

## 8. DAFTAR PUSTAKA

- Julian, Randy. Lily. *Chemical Structure Representationn*. 2009.
- Purba, Michael. *Kimia untuk SMA Kelas X*. Jakarta: Erlangga. 2006.
- Purba, Michael. *Kimia untuk SMA Kelas XII*. Jakarta: Erlangga. 2006.
- Weininger, David. *SMILES, a Chemical Language and Information System. Introduction to Methodology and Encoding Rules*. 1988.
- Weininger, David. *SMILES. 3. Depict. Graphical Depiction of Chemical Structures*. 1990.
- Wikipedia. SMILES. 2009.