

GREYC 

lcmt



  
ENSICAEN  
ÉCOLE NATIONALE SUPÉRIEURE D'INGÉNIEURS DE CAEN  
& CENTRE DE RECHERCHE



# Modélisation de la stéréochimie : une application à la chimoinformatique

Pierre-Anthony Grenier

GREYC, CNRS UMR 6072- CAEN  
[pierre-anthony.grenier@ensicaen.fr](mailto:pierre-anthony.grenier@ensicaen.fr)

26 Novembre 2015



Utiliser l'informatique pour résoudre des problèmes de chimie.



Utiliser l'informatique pour résoudre des problèmes de chimie.

- La gestion de bases de données de molécules et de réactions.
- La détermination de la structure des molécules.
- La prédiction de réactions chimiques.



Utiliser l'informatique pour résoudre des problèmes de chimie.

- La gestion de bases de données de molécules et de réactions.
- La détermination de la structure des molécules.
- La prédiction de réactions chimiques.
- La prédiction de propriétés physiques, chimiques ou biologiques de molécules.



Utiliser l'informatique pour résoudre des problèmes de chimie.

- La gestion de bases de données de molécules et de réactions.
- La détermination de la structure des molécules.
- La prédiction de réactions chimiques.
- La prédiction de propriétés physiques, chimiques ou biologiques de molécules.
  - Conception de nouvelles molécules.
  - Concevoir un nouveau médicament coûte environ un milliard d'euros.



# Prédiction de propriétés moléculaires

**Introduction**  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

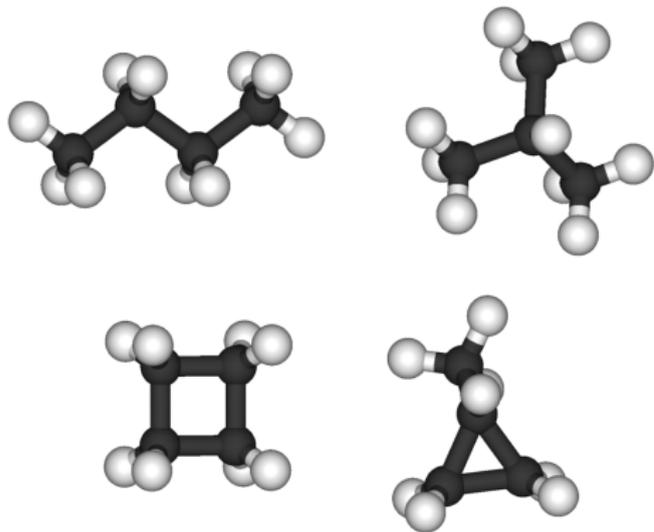
« *Des molécules similaires ont des propriétés similaires.* »



# Prédiction de propriétés moléculaires

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

« *Des molécules similaires ont des propriétés similaires.* »

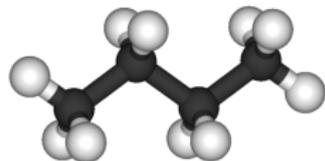




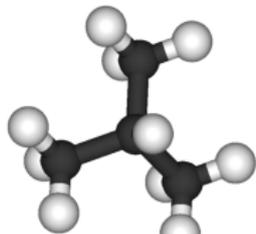
# Prédiction de propriétés moléculaires

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

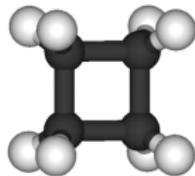
« *Des molécules similaires ont des propriétés similaires.* »



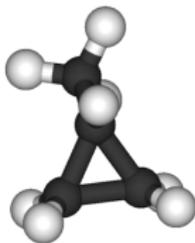
$C_4H_{10}$



$C_4H_{10}$



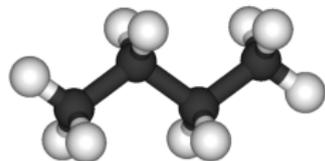
$C_4H_8$



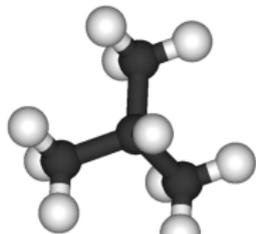
$C_4H_8$



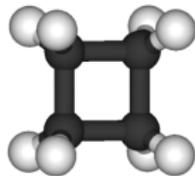
« *Des molécules similaires ont des propriétés similaires.* »



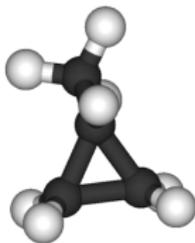
$C_4H_{10}$



$C_4H_{10}$



$C_4H_8$



$C_4H_8$

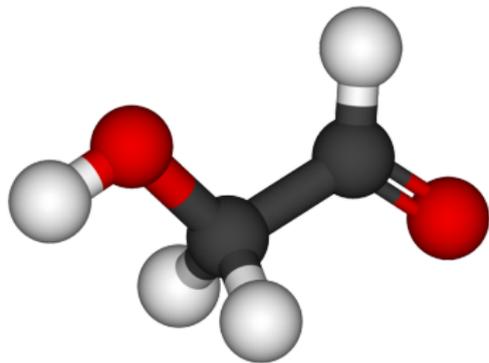
## Mesure de similarité :

- Nombre d'atomes de carbone.
- Nombre d'atomes d'hydrogène.



# Un modèle : le graphe moléculaire

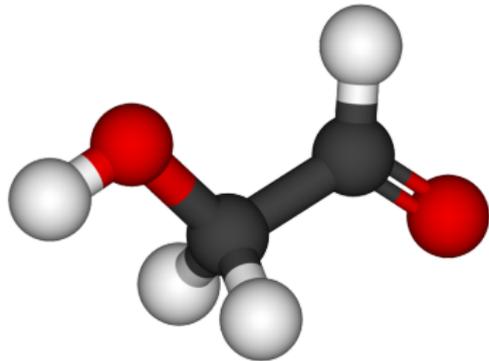
**Introduction**  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion





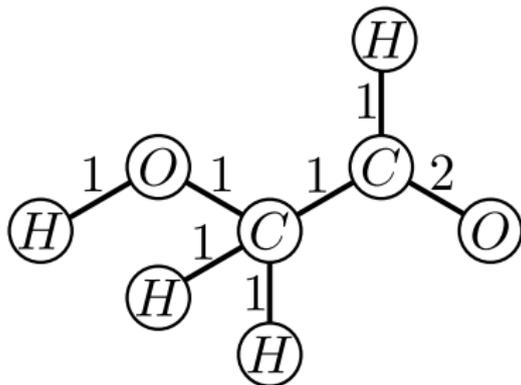
# Un modèle : le graphe moléculaire

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion



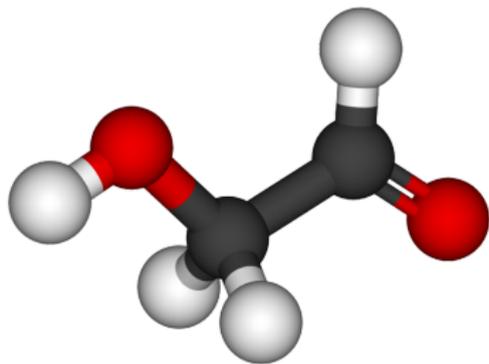
## Graphe moléculaire :

- Nœuds  $\Leftrightarrow$  atomes.
- Arêtes  $\Leftrightarrow$  liaisons entre les atomes.



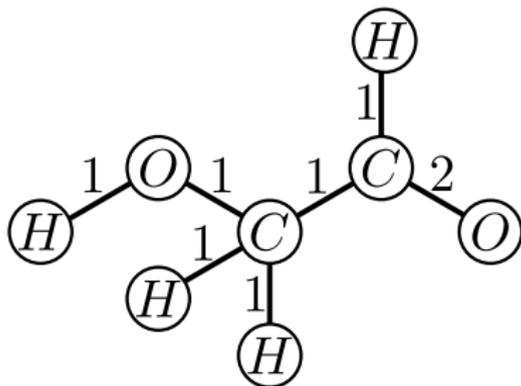


# Un modèle : le graphe moléculaire



## Graphe moléculaire :

- Nœuds  $\Leftrightarrow$  atomes.
- Arêtes  $\Leftrightarrow$  liaisons entre les atomes.
- Représentation par un graphe d'un objet 3D.





# Une mesure de similarité : le noyau sur graphes

**Introduction**  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

- $k(G, G') \rightarrow \mathbb{R}$



# Une mesure de similarité : le noyau sur graphes

**Introduction**  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

- $k(G, G') \rightarrow \mathbb{R}$
- Symétrique et semi défini positif



# Une mesure de similarité : le noyau sur graphes

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

- $k(G, G') \rightarrow \mathbb{R}$
- Symétrique et semi défini positif  $\Rightarrow$  Produit scalaire entre graphes



# Une mesure de similarité : le noyau sur graphes

- $k(G, G') \rightarrow \mathbb{R}$
- Symétrique et semi défini positif  $\Rightarrow$  Produit scalaire entre graphes
- [Ramon, J. et Gärtner, T. 2003]
- [Mahe, P. et Vert J.P., 2009]
- [Gaüzère, B., Brun, L. et Villemin, D. 2013]



# Une mesure de similarité : le noyau sur graphes

- $k(G, G') \rightarrow \mathbb{R}$
- Symétrique et semi défini positif  $\Rightarrow$  Produit scalaire entre graphes
- [Ramon, J. et Gärtner, T. 2003]
- [Mahe, P. et Vert J.P., 2009]
- [Gaüzère, B., Brun, L. et Villemin, D. 2013]
- Bons résultats pour les propriétés scalaires (température d'ébullition, température de fusion, solubilité, ... ).



# Une mesure de similarité : le noyau sur graphes

- $k(G, G') \rightarrow \mathbb{R}$
- Symétrique et semi défini positif  $\Rightarrow$  Produit scalaire entre graphes
- [Ramon, J. et Gärtner, T. 2003]
- [Mahe, P. et Vert J.P., 2009]
- [Gaüzère, B., Brun, L. et Villemin, D. 2013]
- Bons résultats pour les propriétés scalaires (température d'ébullition, température de fusion, solubilité, ... ).
- Mais mauvais pour les propriétés vectorielles (pouvoir rotatoire, dispersion rotatoire, propriétés biologiques, ... ).

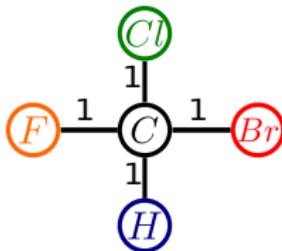


## Limitation du graphe moléculaire : la stéréoisomérie

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

### Problème :

- Un même graphe moléculaire ...

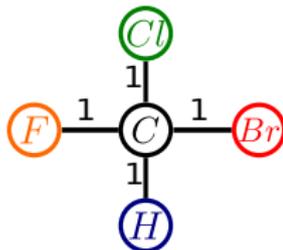




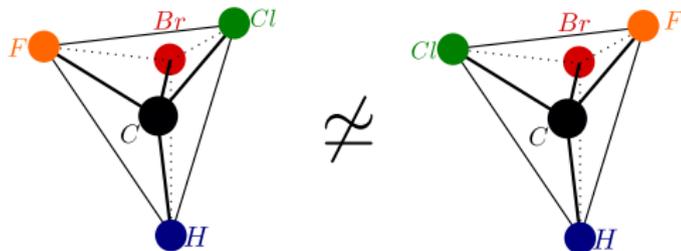
# Limitation du graphe moléculaire : la stéréoisomérie

## Problème :

- Un même graphe moléculaire ...



- ... mais des molécules différentes.



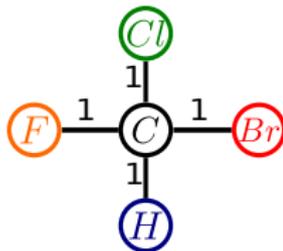


# Limitation du graphe moléculaire : la stéréoisomérie

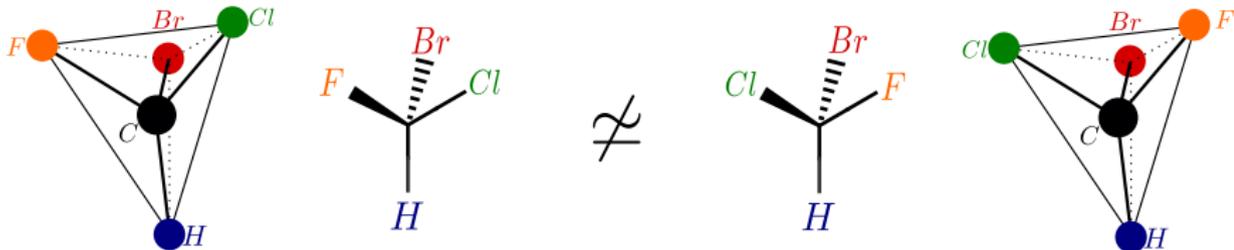
Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

## Problème :

- Un même graphe moléculaire ...



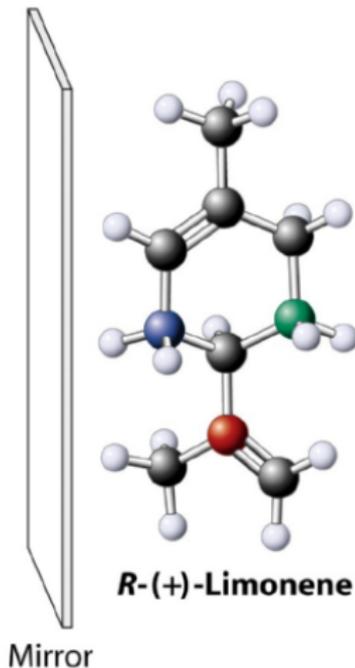
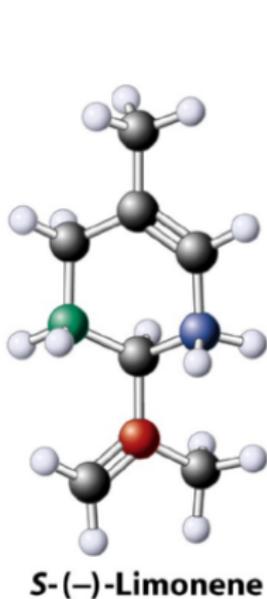
- ... mais des molécules différentes.





# Exemples de propriétés liées à la stéréoisomérie

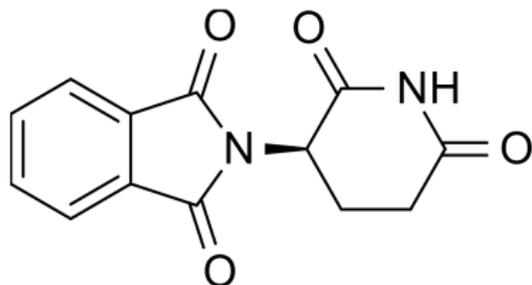
Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion



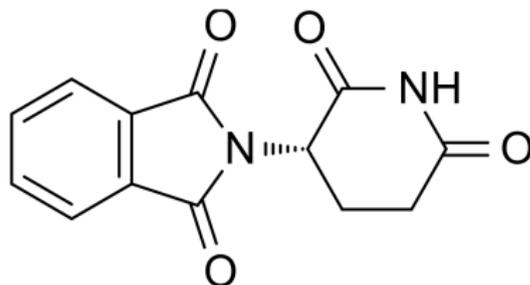


# Exemples de propriétés liées à la stéréoisomérie

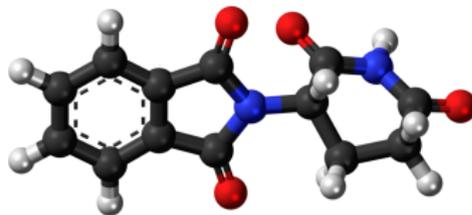
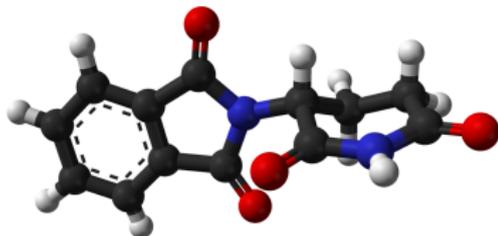
Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion



**(R)-thalidomide**



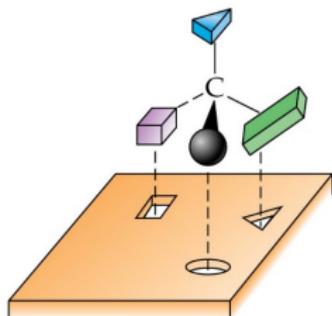
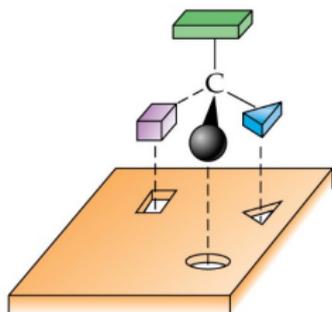
**(S)-thalidomide**





# Exemples de propriétés liées à la stéréoisomérie

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion





## Centres stéréogènes

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

- Atome ou groupement d'atomes dont l'échange de deux voisins génère un stéréoisomère différent.



# Centres stéréogènes

- Atome ou groupement d'atomes dont l'échange de deux voisins génère un stéréoisomère différent.
- Carbone asymétrique :





- Atome ou groupement d'atomes dont l'échange de deux voisins génère un stéréoisomère différent.
- Carbone asymétrique :



- Carbones liés par une liaison double :





- Atome ou groupement d'atomes dont l'échange de deux voisins génère un stéréoisomère différent.
- Carbone asymétrique :



- Carbones liés par une liaison double :



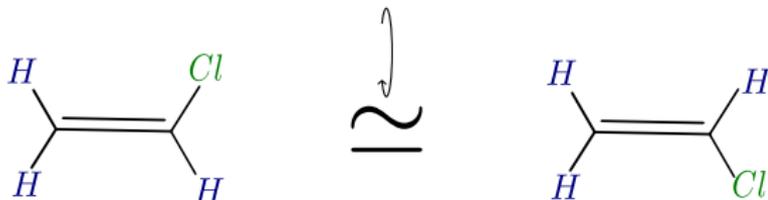
Pas de rotation de la liaison double possible.



- Atome ou groupement d'atomes dont l'échange de deux voisins génère un stéréoisomère différent.
- Carbone asymétrique :



- Carbones liés par une liaison double :



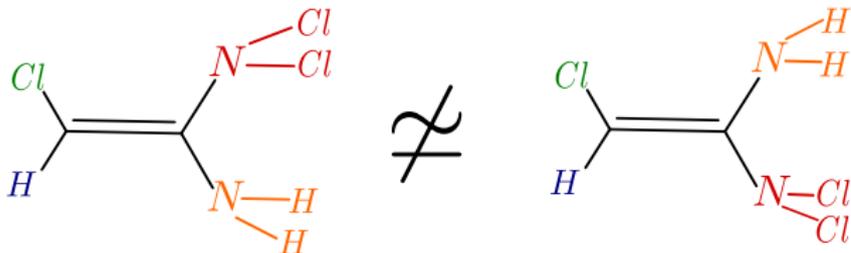


# Centres stéréogènes

- Atome ou groupement d'atomes dont l'échange de deux voisins génère un stéréoisomère différent.
- Carbone asymétrique :



- Carbones liés par une liaison double :





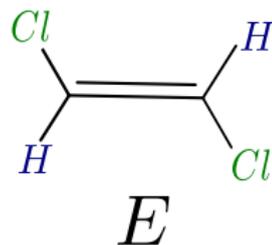
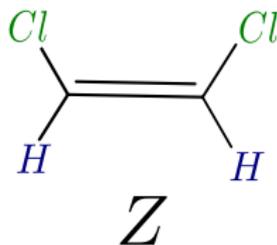
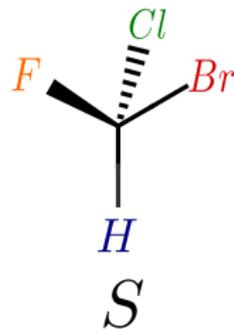
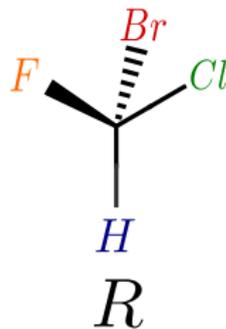
# Méthodes existantes

**Introduction**  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

- Représentation



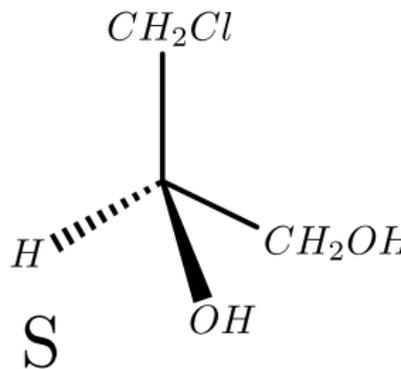
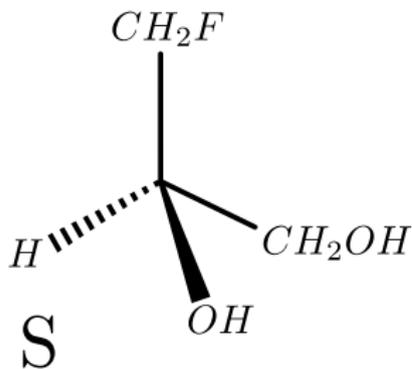
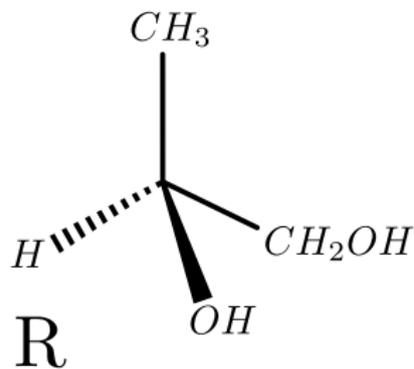
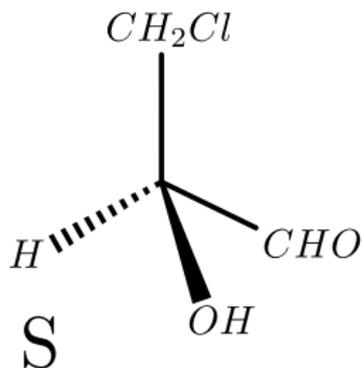
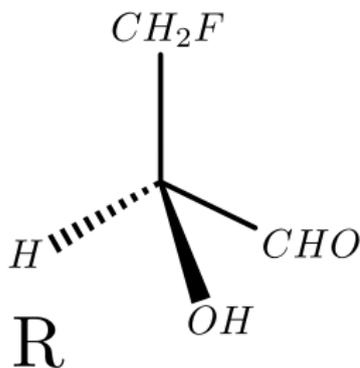
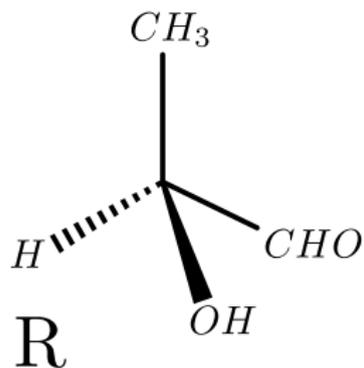
- Représentation
  - Nomenclature CIP.





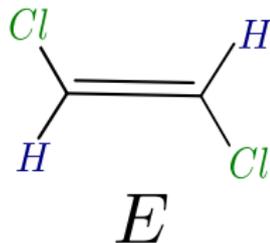
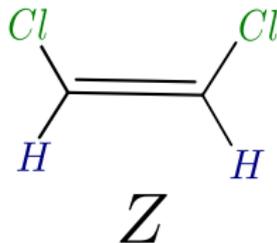
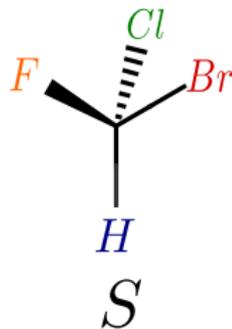
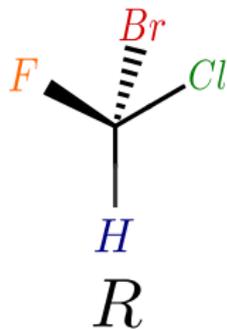
## Méthodes existantes

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion





- Représentation
  - Nomenclature CIP.
    - Uniquement pour distinguer les stéréoisomères.





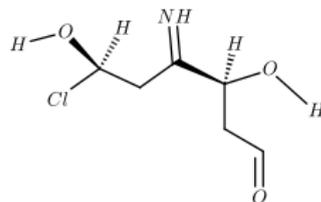
# Méthodes existantes

**Introduction**  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

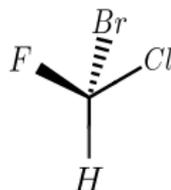
- Représentation
  - Nomenclature CIP.
    - Uniquement pour distinguer les stéréoisomères.
- Prédiction



- Représentation
  - Nomenclature CIP.
    - Uniquement pour distinguer les stéréoisomères.
- Prédiction
  - Descripteurs moléculaires. [Golbraikh, A. et al. 2001]



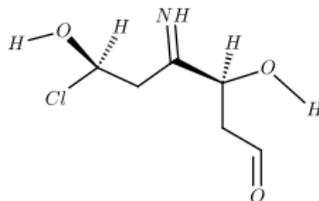
1
1
2
4, 3
0



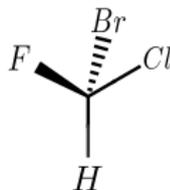
1
0
0
2, 1
0



- Représentation
  - Nomenclature CIP.
    - Uniquement pour distinguer les stéréoisomères.
- Prédiction
  - Descripteurs moléculaires. [Golbraikh, A. et al. 2001]
    - Pertinence.



1
1
2
4, 3
0



1
0
0
2, 1
0



- Représentation

- Nomenclature CIP.

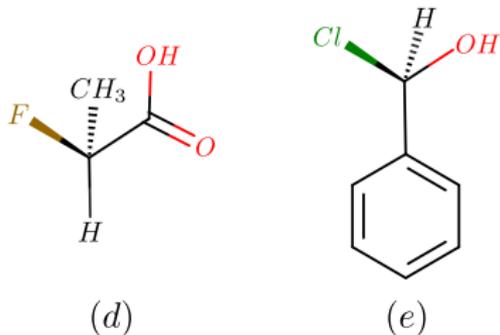
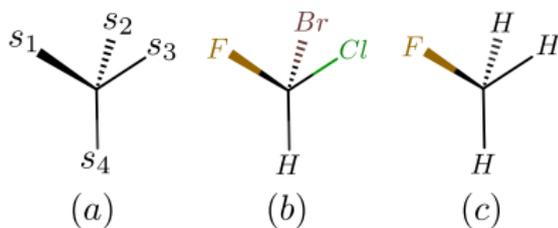
- Uniquement pour distinguer les stéréoisomères.

- Prédiction

- Descripteurs moléculaires. [Golbraikh, A. et al. 2001]

- Pertinence.

- Polynômes chiraux. [Ruch, E. et Schönhofer, A. 1970]





- Représentation

- Nomenclature CIP.

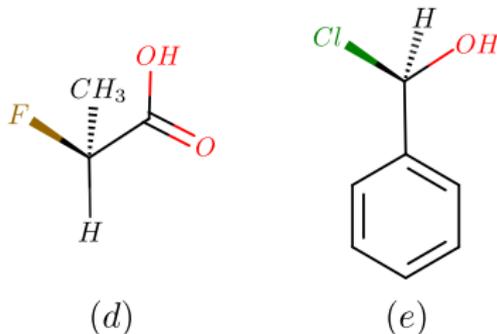
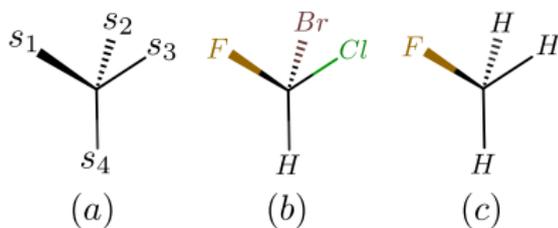
- Uniquement pour distinguer les stéréoisomères.

- Prédiction

- Descripteurs moléculaires. [Golbraikh, A. et al. 2001]

— Pertinence.

- Polynômes chiraux. [Ruch, E. et Schönhofer, A. 1970]



$$(s_4 - s_3)(s_4 - s_2)(s_4 - s_1)(s_3 - s_2)(s_3 - s_1)(s_2 - s_1)$$



- Représentation

- Nomenclature CIP.

- Uniquement pour distinguer les stéréoisomères.

- Prédiction

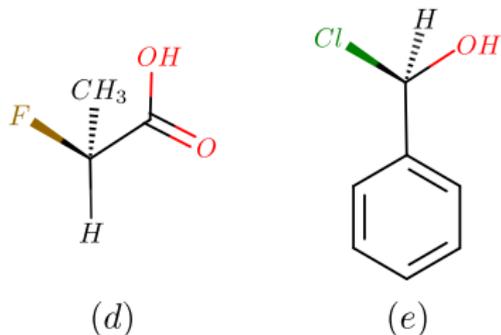
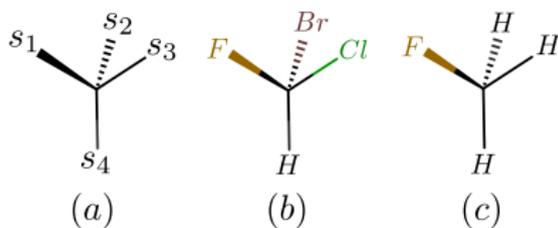
- Descripteurs moléculaires. [Golbraikh, A. et al. 2001]

- Pertinence.

- Polynômes chiraux. [Ruch, E. et Schönhofer, A. 1970]

- Dépendance factorielle de la complexité.

- Adapté à peu de cas.



$$(s_4 - s_3)(s_4 - s_2)(s_4 - s_1)(s_3 - s_2)(s_3 - s_1)(s_2 - s_1)$$



- Représentation

- Nomenclature CIP.

- Uniquement pour distinguer les stéréoisomères.

- Prédiction

- Descripteurs moléculaires. [Golbraikh, A. et al. 2001]

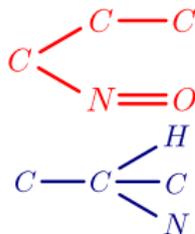
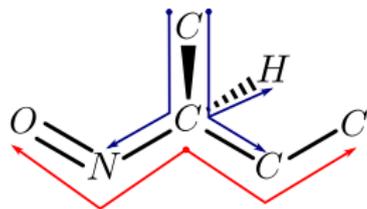
- Pertinence.

- Polynômes chiraux. [Ruch, E. et Schönhofer, A. 1970]

- Dépendance factorielle de la complexité.

- Adapté à peu de cas.

- Noyau de motifs d'arbres adapté à la stéréochimie. [Brown, J.B. et al. 2010]



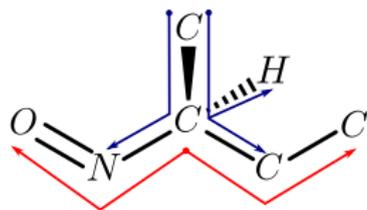


- Représentation

- Nomenclature CIP.
  - Uniquement pour distinguer les stéréoisomères.

- Prédiction

- Descripteurs moléculaires. [Golbraikh, A. et al. 2001]
  - Pertinence.
- Polynômes chiraux. [Ruch, E. et Schönhofer, A. 1970]
  - Dépendance factorielle de la complexité.
  - Adapté à peu de cas.
- Noyau de motifs d'arbres adapté à la stéréochimie. [Brown, J.B. et al. 2010]





# Objectifs

**Introduction**  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

+ La prédiction de propriétés scalaires est à présent bien connue.



- + La prédiction de propriétés scalaires est à présent bien connue.
- La prédiction de propriétés vectorielles ne dispose pas pour l'instant de solutions aussi efficaces.



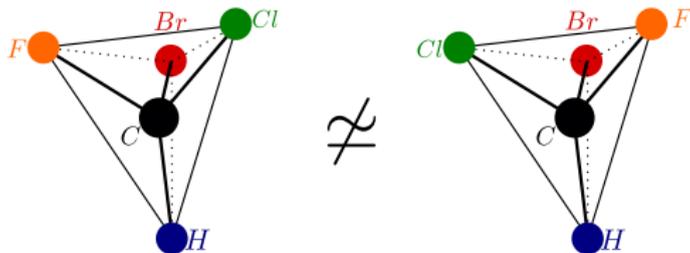
- + La prédiction de propriétés scalaires est à présent bien connue.
- La prédiction de propriétés vectorielles ne dispose pas pour l'instant de solutions aussi efficaces.
  - Ces propriétés sont importantes.



# Modèle de représentation de la stéréoisométrie



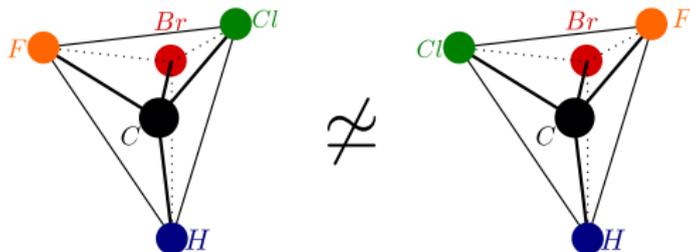
Stéréoisomérisation ne dépend pas des coordonnées.





Stéréoisomérisation ne dépend pas des coordonnées.

L'arrangement des atomes dans l'espace est naturellement encodé par une notion d'ordre.





## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

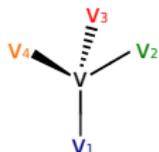
$$ord \begin{cases} V & \rightarrow & V^* \\ v & \rightarrow & v_1, \dots, v_n \end{cases}$$



## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

$$ord \begin{cases} V & \rightarrow & V^* \\ v & \rightarrow & v_1, \dots, v_n \end{cases}$$

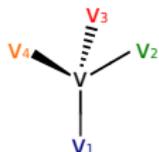




## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

$$ord \begin{cases} V & \rightarrow & V^* \\ v & \rightarrow & v_1, \dots, v_n \end{cases}$$



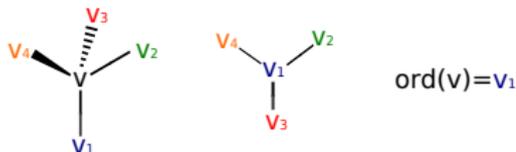
$$ord(v) = v_1$$



## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

$$ord \begin{cases} V & \rightarrow & V^* \\ v & \rightarrow & v_1, \dots, v_n \end{cases}$$

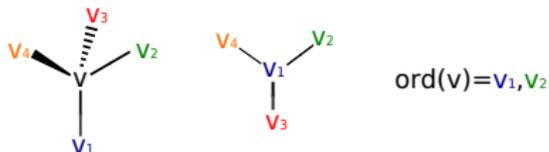




## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

$$ord \begin{cases} V & \rightarrow & V^* \\ v & \rightarrow & v_1, \dots, v_n \end{cases}$$

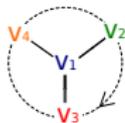
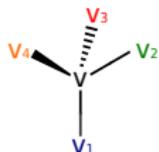




## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

$$ord \begin{cases} V & \rightarrow & V^* \\ v & \rightarrow & v_1, \dots, v_n \end{cases}$$



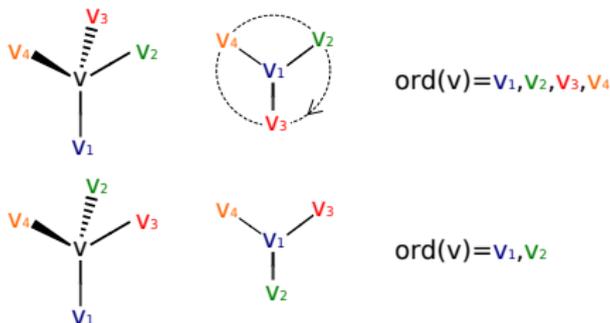
$$ord(v) = v_1, v_2, v_3, v_4$$



## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

$$ord \begin{cases} V & \rightarrow & V^* \\ v & \rightarrow & v_1, \dots, v_n \end{cases}$$

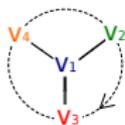
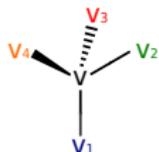




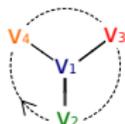
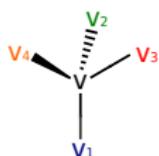
## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

$$ord \begin{cases} V & \rightarrow & V^* \\ v & \rightarrow & v_1, \dots, v_n \end{cases}$$



$$ord(v) = v_1, v_2, v_3, v_4$$



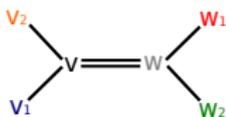
$$ord(v) = v_1, v_2, v_4, v_3$$



## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

$$ord \begin{cases} V & \rightarrow & V^* \\ v & \rightarrow & v_1, \dots, v_n \end{cases}$$

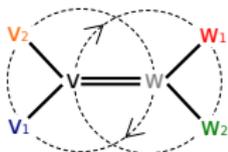




## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

$$ord \begin{cases} V & \rightarrow & V^* \\ v & \rightarrow & v_1, \dots, v_n \end{cases}$$

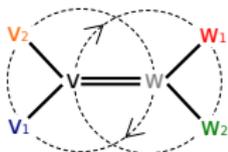




## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

$$ord \begin{cases} V & \rightarrow V^* \\ v & \rightarrow v_1, \dots, v_n \end{cases}$$



$$ord(v) = w, v_1, v_2$$

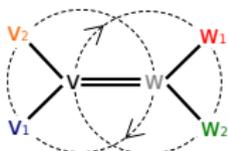
$$ord(w) = v, w_1, w_2$$



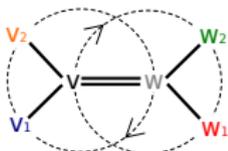
## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

$$ord \begin{cases} V & \rightarrow & V^* \\ v & \rightarrow & v_1, \dots, v_n \end{cases}$$



$$\begin{aligned} ord(v) &= w, v_1, v_2 \\ ord(w) &= v, w_1, w_2 \end{aligned}$$



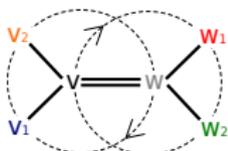
$$\begin{aligned} ord(v) &= w, v_1, v_2 \\ ord(w) &= v, w_2, w_1 \end{aligned}$$



## Graphe ordonné :

- $G = (V, E, \mu, \nu, ord)$ , avec  $G_m = (V, E, \mu, \nu)$  un graphe et une fonction d'ordre  $ord$ .

$$ord \begin{cases} V & \rightarrow & V^* \\ v & \rightarrow & v_1, \dots, v_n \end{cases}$$



$$ord(v) = w, v_1, v_2$$

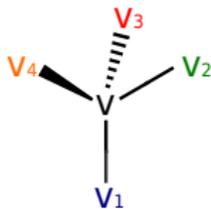
$$ord(w) = v, w_1, w_2$$

## Isomorphisme entre graphes ordonnés :

- $\exists f \in \text{Isom}(G_m, G'_m)$  tel que  $\forall v \in V \text{ ord}'(f(v)) = (f(v_1), \dots, f(v_n))$  avec  $ord(v) = (v_1, \dots, v_n)$ .

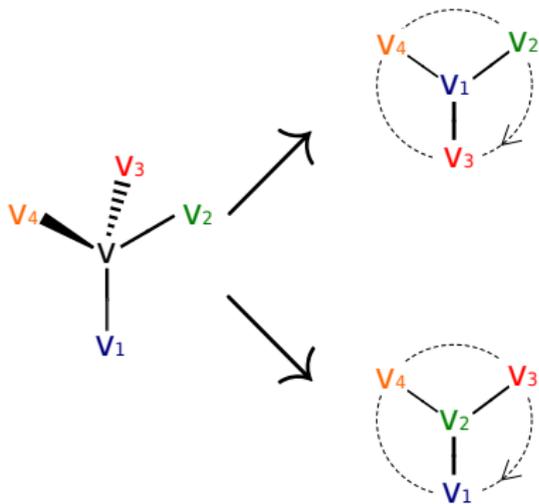


- Une même configuration ...





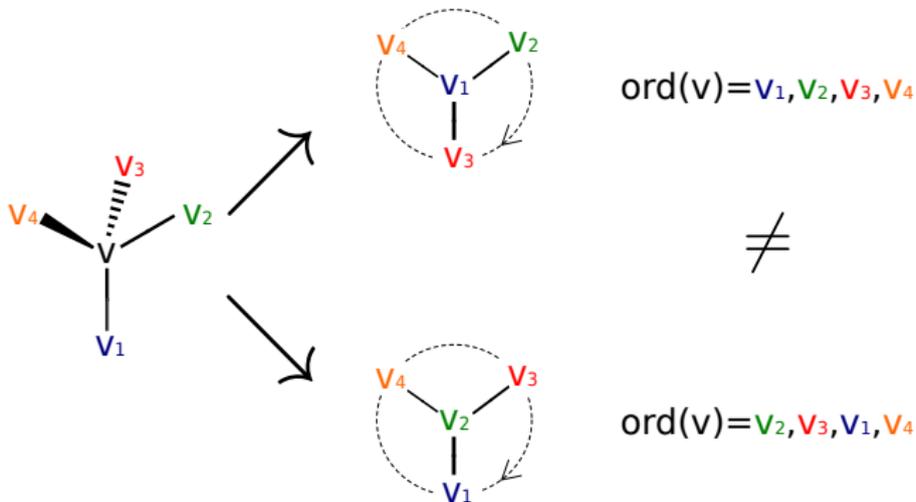
- Une même configuration ...





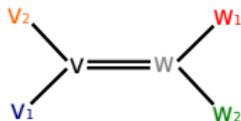
# Limitation des graphes ordonnés

- Une même configuration, mais des ordres différents.



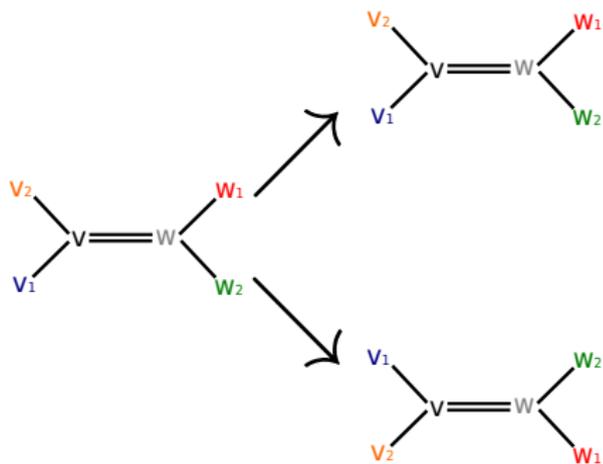


- Une même configuration ...



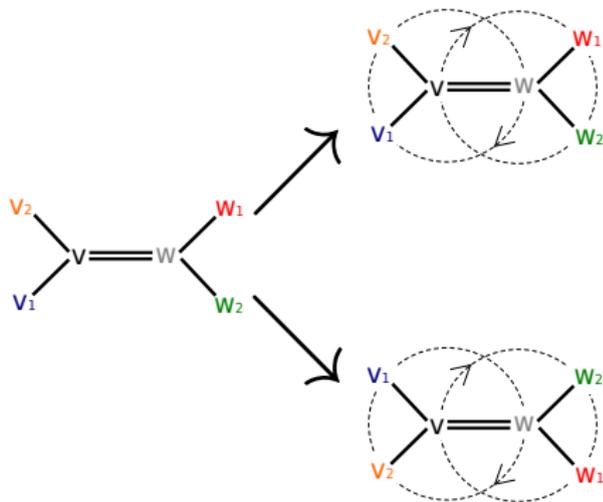


- Une même configuration ...



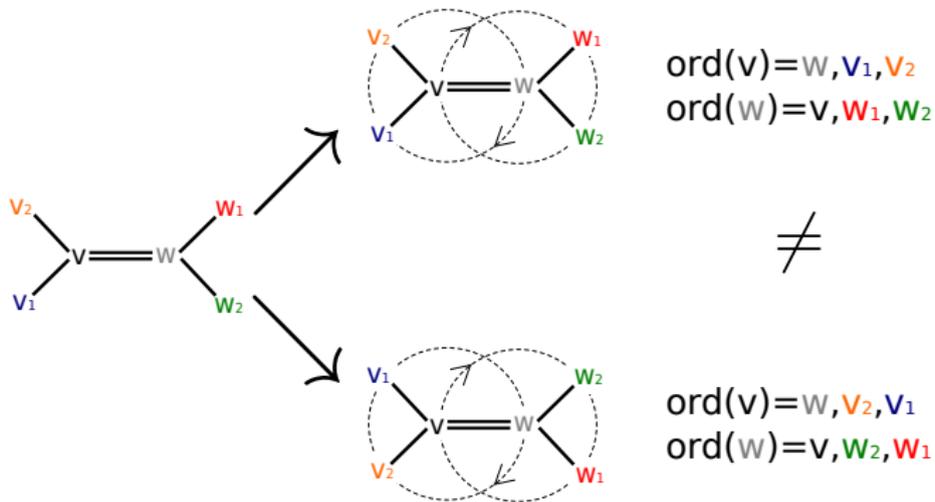


- Une même configuration ...



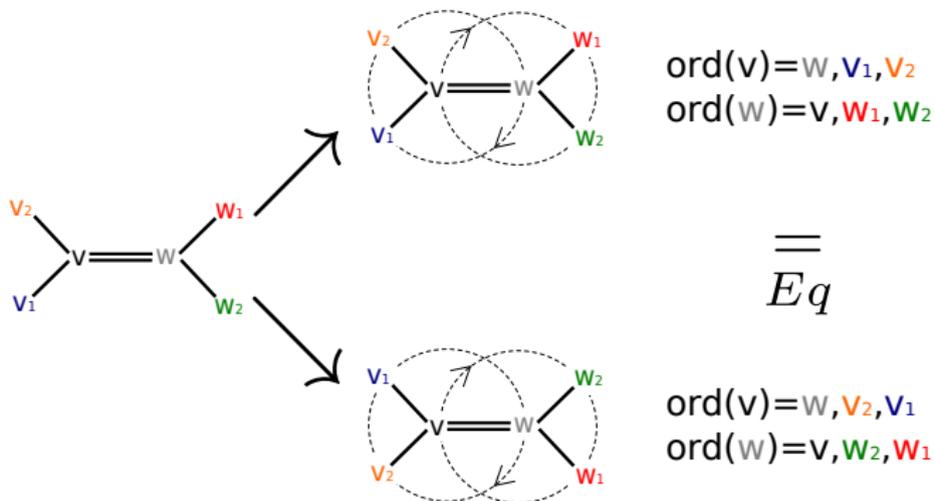


- Une même configuration, mais des ordres différents.





- Une même configuration, mais des ordres différents, qu'on veut considérer comme équivalents.





# Ordres équivalents

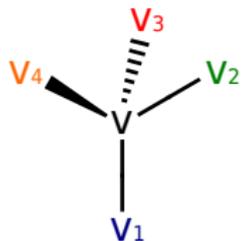
Introduction  
**Graphes Ordonnés**  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

$$\sigma \begin{cases} V & \rightarrow \mathcal{P} \\ v & \rightarrow \varphi_v \end{cases}$$



# Ordres équivalents

$$\sigma \begin{cases} V & \rightarrow \mathcal{P} \\ v & \rightarrow \varphi_v \end{cases}$$

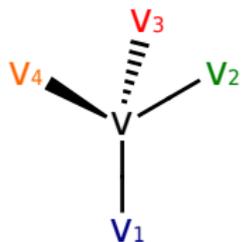


$$\text{ord}(v) = V_1, V_2, V_3, V_4$$



# Ordres équivalents

$$\sigma \begin{cases} V & \rightarrow \mathcal{P} \\ v & \rightarrow \varphi_v \end{cases}$$



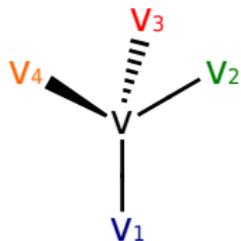
$$\sigma(v) = (1 \ 3 \ 2) (4)$$

$$\text{ord}(v) = V_1, V_2, V_3, V_4$$

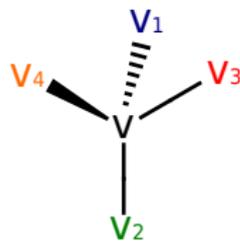
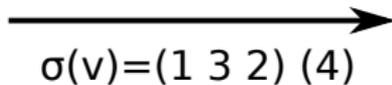


# Ordres équivalents

$$\sigma \begin{cases} V & \rightarrow \mathcal{P} \\ v & \rightarrow \varphi_v \end{cases}$$



$$\text{ord}(v) = V_1, V_2, V_3, V_4$$

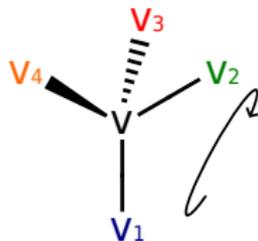


$$\text{ord}_{\sigma(v)}(v) = V_2, V_3, V_1, V_4$$

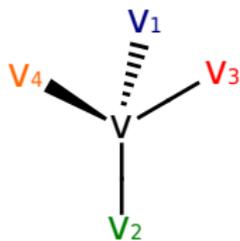
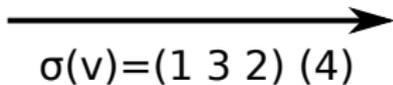


# Ordres équivalents

$$\sigma \begin{cases} V & \rightarrow \mathcal{P} \\ v & \rightarrow \varphi_v \end{cases}$$



$$\text{ord}(v) = V_1, V_2, V_3, V_4$$



$$\text{ord}_{\sigma(v)}(v) = V_2, V_3, V_1, V_4$$



# Isomorphisme d'équivalence d'ordres

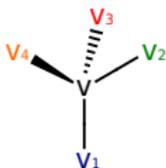
Introduction  
**Graphes Ordonnés**  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

- Parité d'une permutation : parité de son nombre de transpositions.



# Isomorphisme d'équivalence d'ordres

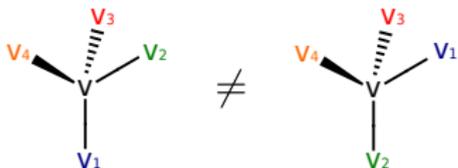
- Parité d'une permutation : parité de son nombre de transpositions.
- Carbone asymétrique  $v$





# Isomorphisme d'équivalence d'ordres

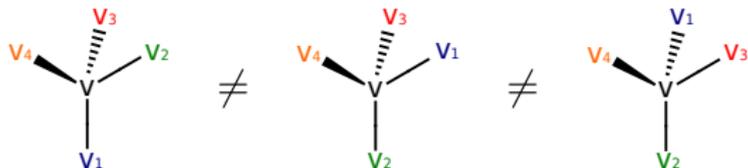
- Parité d'une permutation : parité de son nombre de transpositions.
- Carbone asymétrique  $\nu$





# Isomorphisme d'équivalence d'ordres

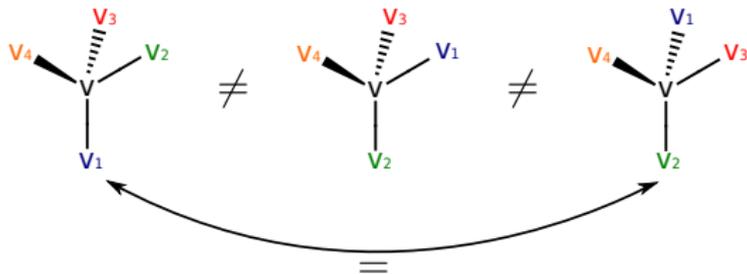
- Parité d'une permutation : parité de son nombre de transpositions.
- Carbone asymétrique  $v$





# Isomorphisme d'équivalence d'ordres

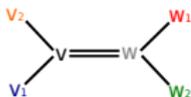
- Parité d'une permutation : parité de son nombre de transpositions.
- Carbone asymétrique  $v$  :  $\sigma(v)$  paire.





# Isomorphisme d'ordres

- Parité d'une permutation : parité de son nombre de transpositions.
- Carbone asymétrique  $v$  :  $\sigma(v)$  paire.
- Carbones  $v$  et  $w$  liés par une liaison double





# Isomorphisme d'équivalences d'ordres

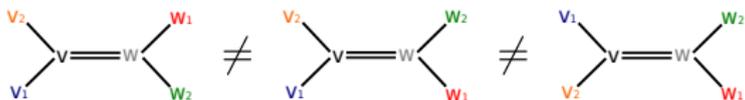
- Parité d'une permutation : parité de son nombre de transpositions.
- Carbone asymétrique  $v$  :  $\sigma(v)$  paire.
- Carbones  $v$  et  $w$  liés par une liaison double





# Isomorphisme d'ordres

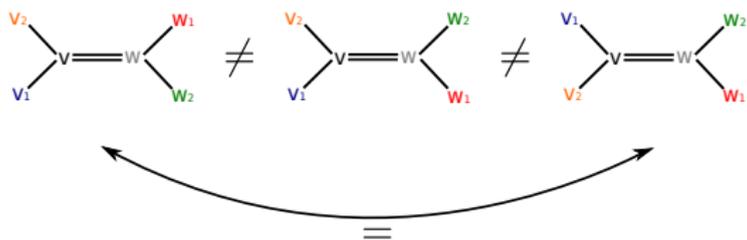
- Parité d'une permutation : parité de son nombre de transpositions.
- Carbone asymétrique  $v$  :  $\sigma(v)$  paire.
- Carbones  $v$  et  $w$  liés par une liaison double





# Isomorphisme d'ordres

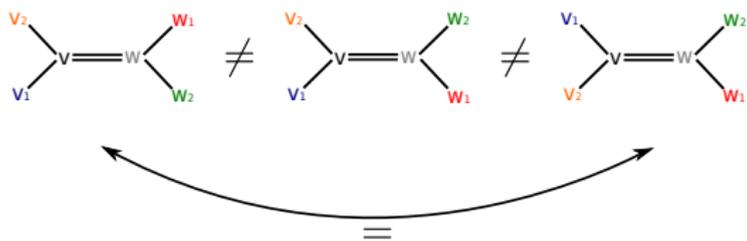
- Parité d'une permutation : parité de son nombre de transpositions.
- Carbone asymétrique  $v$  :  $\sigma(v)$  paire.
- Carbones  $v$  et  $w$  liés par une liaison double :  $\sigma(v)$  et  $\sigma(w)$  même parité.





# Isomorphisme d'équivalence d'ordres

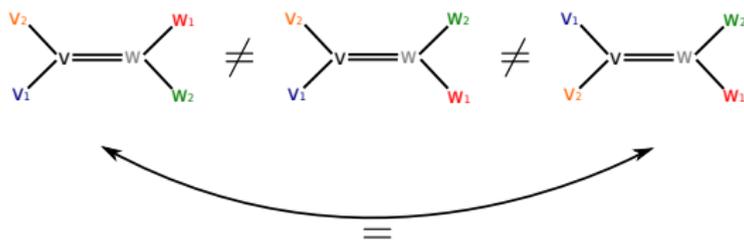
- Parité d'une permutation : parité de son nombre de transpositions.
- Carbone asymétrique  $v$  :  $\sigma(v)$  paire.
- Carbones  $v$  et  $w$  liés par une liaison double :  $\sigma(v)$  et  $\sigma(w)$  même parité.
- Famille valide de fonctions de ré-ordonnement  $\Sigma$ .





# Isomorphisme d'équivalence d'ordres

- Parité d'une permutation : parité de son nombre de transpositions.
- Carbone asymétrique  $v$  :  $\sigma(v)$  paire.
- Carbones  $v$  et  $w$  liés par une liaison double :  $\sigma(v)$  et  $\sigma(w)$  même parité.
- Famille valide de fonctions de ré-ordonnement  $\Sigma$ .

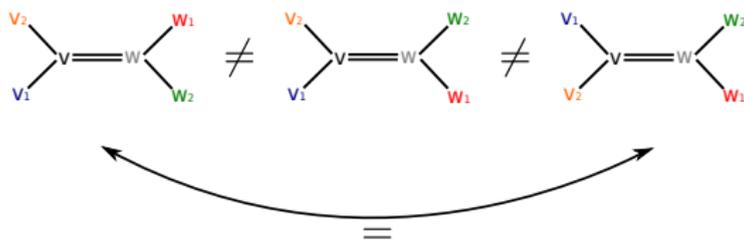


$$G \underset{\Sigma}{\simeq} G' \Leftrightarrow \exists \sigma \in \Sigma, \sigma(G) \underset{\circ}{\simeq} G'$$



# Isomorphisme d'ordres

- Parité d'une permutation : parité de son nombre de transpositions.
- Carbone asymétrique  $v$  :  $\sigma(v)$  paire.
- Carbones  $v$  et  $w$  liés par une liaison double :  $\sigma(v)$  et  $\sigma(w)$  même parité.
- Famille valide de fonctions de ré-ordonnement  $\Sigma$ .



$$G \underset{\Sigma}{\simeq} G' \Leftrightarrow \exists \sigma \in \Sigma, \sigma(G) \underset{\sigma}{\simeq} G'$$

- Deux graphes ordonnés équivalents correspondent à un même stéréoisomère.



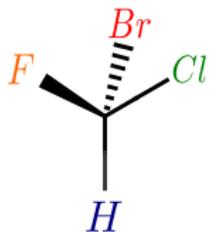
- Un sommet est appelé un stéréo sommet si n'importe quelle permutation de deux de ses voisins produit un graphe ordonné d'ordres non équivalents.
- Soit  $G = (V, E, \mu, \nu, ord)$  un graphe ordonné. Un sommet  $v \in V$  est appelé stéréo sommet si :

$\forall (i, j) \in \{1, \dots, |N(v)|\}^2, i \neq j, \nexists f \in \text{IsomEqOrd}(G, \tau_{i,j}^v(G))$  avec  $f(v) = v$ .



- Un sommet est appelé un stéréo sommet si n'importe quelle permutation de deux de ses voisins produit un graphe ordonné d'ordres non équivalents.
- Soit  $G = (V, E, \mu, \nu, ord)$  un graphe ordonné. Un sommet  $v \in V$  est appelé stéréo sommet si :

$\forall (i, j) \in \{1, \dots, |N(v)|\}^2, i \neq j, \nexists f \in \text{IsomEqOrd}(G, \tau_{i,j}^v(G))$  avec  $f(v) = v$ .

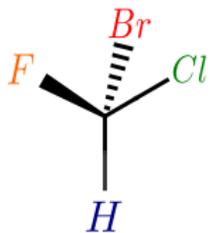


$ord(v) = H, Cl, Br, F$

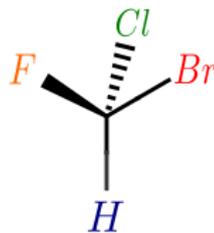


- Un sommet est appelé un stéréo sommet si n'importe quelle permutation de deux de ses voisins produit un graphe ordonné d'ordres non équivalents.
- Soit  $G = (V, E, \mu, \nu, ord)$  un graphe ordonné. Un sommet  $v \in V$  est appelé stéréo sommet si :

$\forall (i, j) \in \{1, \dots, |N(v)|\}^2, i \neq j, \nexists f \in \text{IsomEqOrd}(G, \tau_{i,j}^v(G))$  avec  $f(v) = v$ .



$ord(v) = H, Cl, Br, F$

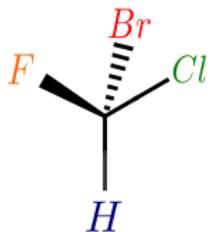


$ord_{\tau}(v) = H, Br, Cl, F$

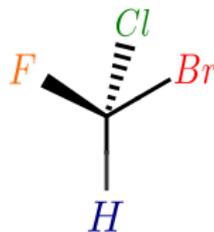


- Un sommet est appelé un stéréo sommet si n'importe quelle permutation de deux de ses voisins produit un graphe ordonné d'ordres non équivalents.
- Soit  $G = (V, E, \mu, \nu, ord)$  un graphe ordonné. Un sommet  $v \in V$  est appelé stéréo sommet si :

$\forall (i, j) \in \{1, \dots, |N(v)|\}^2, i \neq j, \nexists f \in \underline{\text{IsomEqOrd}}(G, \tau_{i,j}^v(G))$  avec  $f(v) = v$ .



$ord(v) = H, Cl, Br, F$   
 $ord_{\sigma_1}(v) = H, Br, F, Cl$   
 $ord_{\sigma_2}(v) = Br, H, Cl, F$

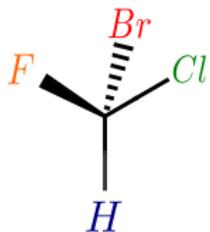


$ord_{\tau}(v) = H, Br, Cl, F$

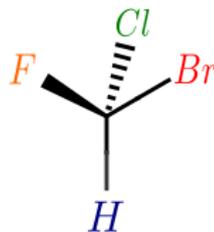


- Un sommet est appelé un stéréo sommet si n'importe quelle permutation de deux de ses voisins produit un graphe ordonné d'ordres non équivalents.
- Soit  $G = (V, E, \mu, \nu, ord)$  un graphe ordonné. Un sommet  $v \in V$  est appelé stéréo sommet si :

$\forall (i, j) \in \{1, \dots, |N(v)|\}^2, i \neq j, \nexists f \in \text{IsomEqOrd}(G, \tau_{i,j}^v(G))$  avec  $f(v) = v$ .



$\neq$   
 $\Sigma$



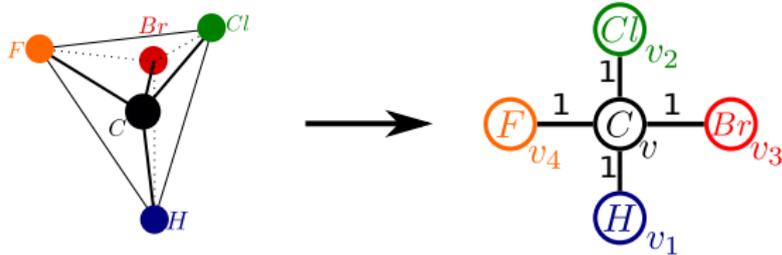
$ord(v) = H, Cl, Br, F$   
 $ord_{\sigma_1}(v) = H, Br, F, Cl$   
 $ord_{\sigma_2}(v) = Br, H, Cl, F$

$ord_{\tau}(v) = H, Br, Cl, F$



# Modèle de représentation de la stéréoisomérie

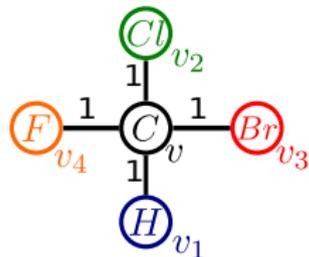
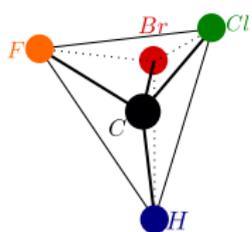
Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion





# Modèle de représentation de la stéréoisomérie

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

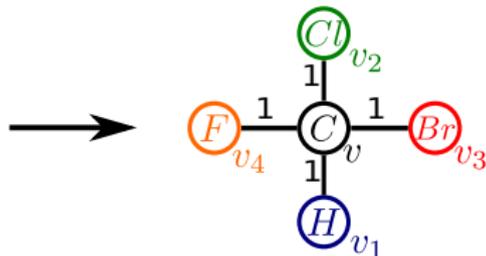
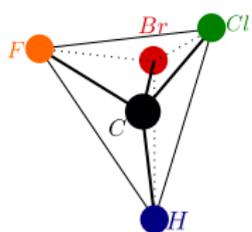


$$\text{ord}(v) = v_1, v_2, v_3, v_4$$



# Modèle de représentation de la stéréoisomérie

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinsages  
Noyau inter stéréo  
Conclusion



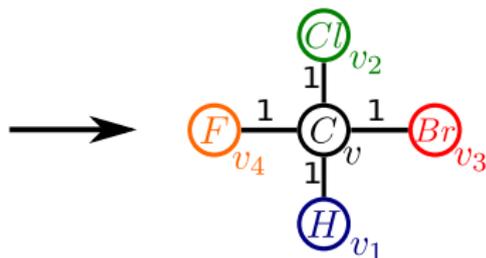
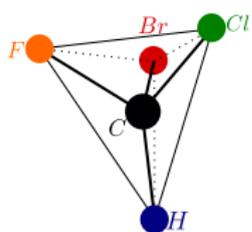
$$\text{ord}(v) = v_1, v_2, v_3, v_4$$

$$\text{ord}_{\sigma(v)}(v) = v_2, v_3, v_1, v_4$$



# Modèle de représentation de la stéréoisomérie

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinsages  
Noyau inter stéréo  
Conclusion



$$\text{ord}(v) = v_1, v_2, v_3, v_4$$

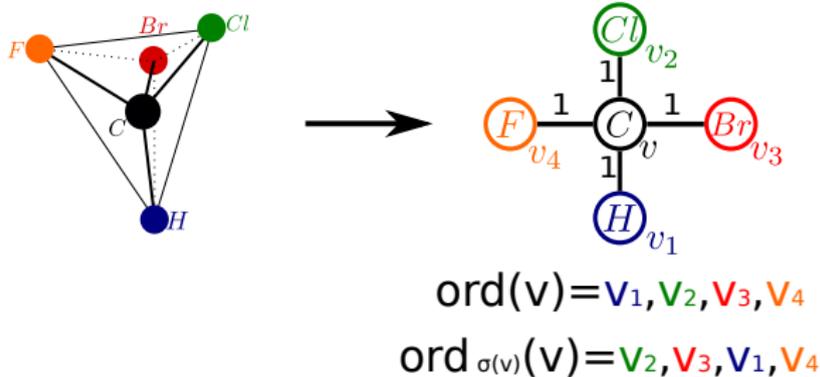
$$\text{ord}_{\sigma(v)}(v) = v_2, v_3, v_1, v_4$$

Généralisation de [Jiang, X. et Bunke, H. 1999].



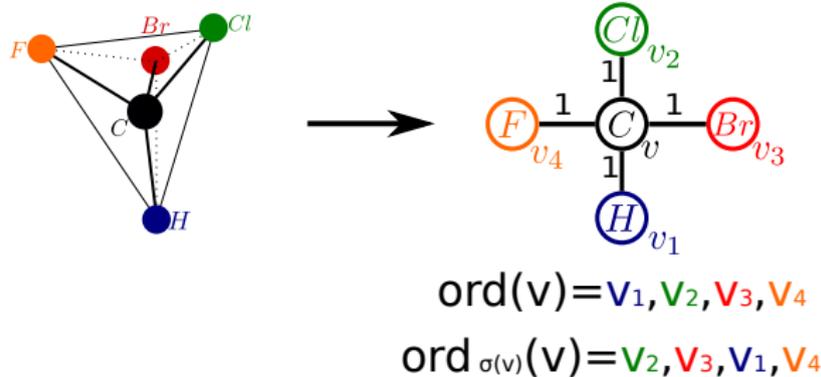
# Modèle de représentation de la stéréoisomérie

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion



Généralisation de [Jiang, X. et Bunke, H. 1999].

- Ordres définis sur tous les sommets  $\neq$  Ordres définis sur certains sommets.



Généralisation de [Jiang, X. et Bunke, H. 1999].

- Ordres définis sur tous les sommets  $\neq$  Ordres définis sur certains sommets.
- Ordres équivalents si permutation circulaire  $\neq$  Conditions plus générales.



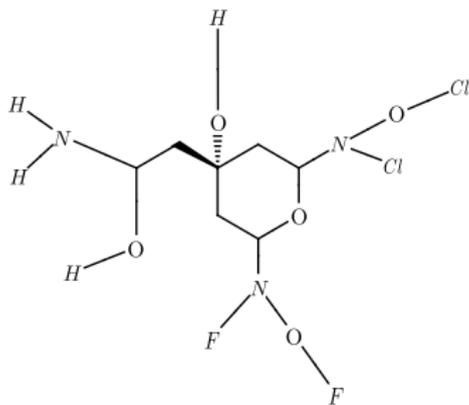
# Mesure de similarité entre graphes ordonnés



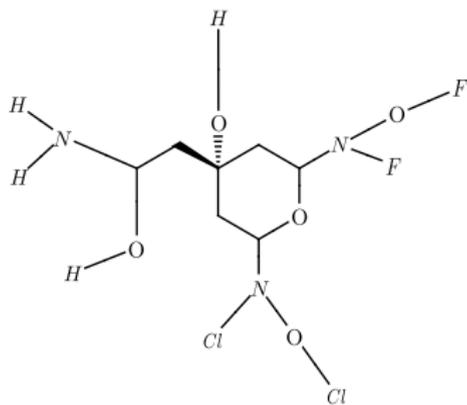
# Stéréo sous-graphe minimal

$$\nexists f \in \text{IsomEqOrd}(G, \tau_{i,j}^v(G))$$

- Définition globale.



$\nexists \Sigma$

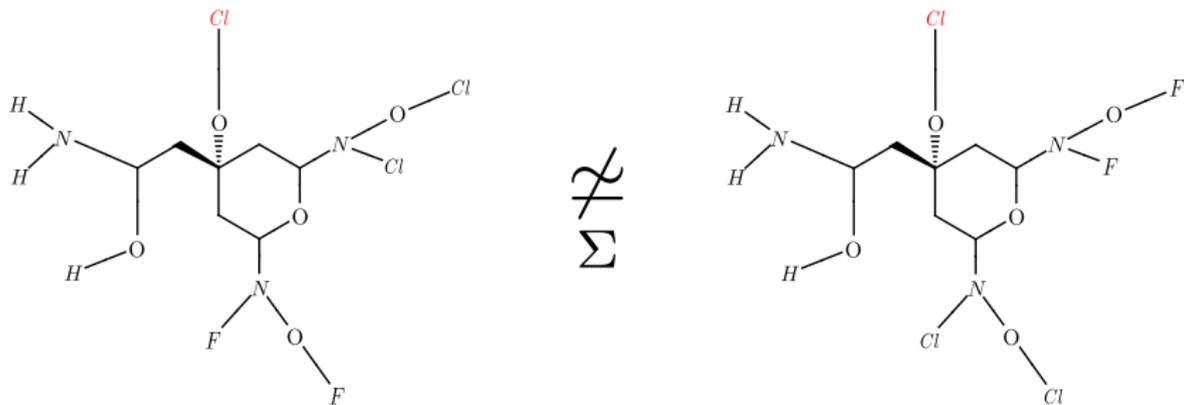




# Stéréo sous-graphe minimal

$$\nexists f \in \text{IsomEqOrd}(G, \tau_{i,j}^v(G))$$

- Définition globale.
- Modifications ou suppressions éloignées du stéréo sommet n'ont pas d'influence.

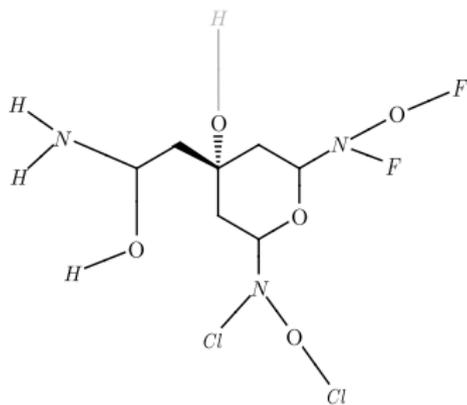
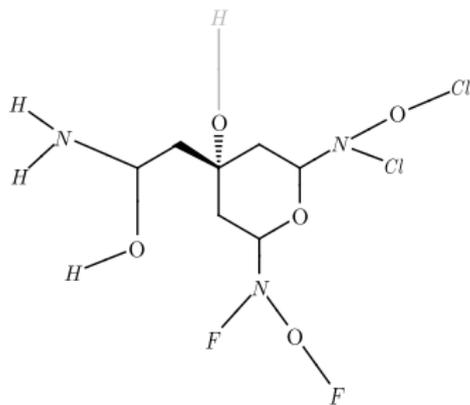




# Stéréo sous-graphe minimal

$$\nexists f \in \text{IsomEqOrd}(H, \tau_{i,j}^v(H))$$

- Définition globale.
- Modifications ou suppressions éloignées du stéréo sommet n'ont pas d'influence.

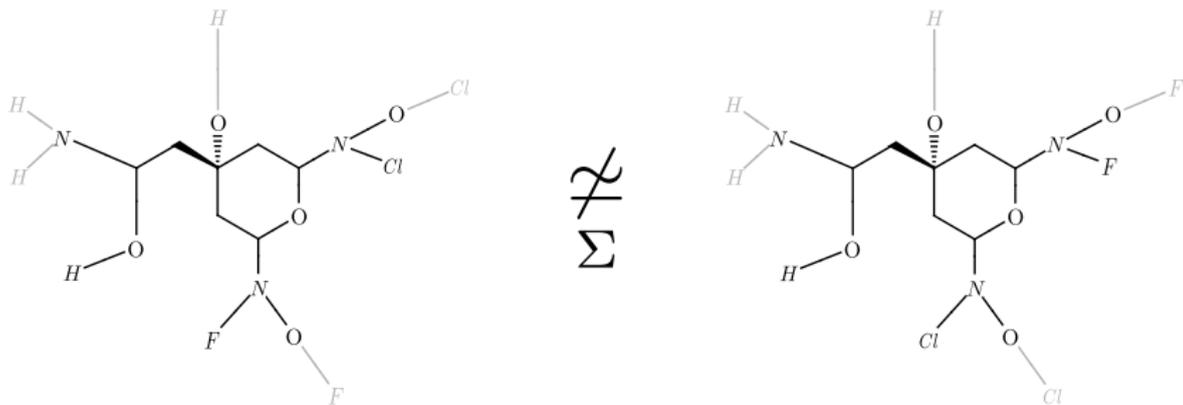




# Stéréo sous-graphe minimal

$$\nexists f \in \text{IsomEqOrd}(H, \tau_{i,j}^v(H))$$

- Définition globale.
- Modifications ou suppressions éloignées du stéréo sommet n'ont pas d'influence.

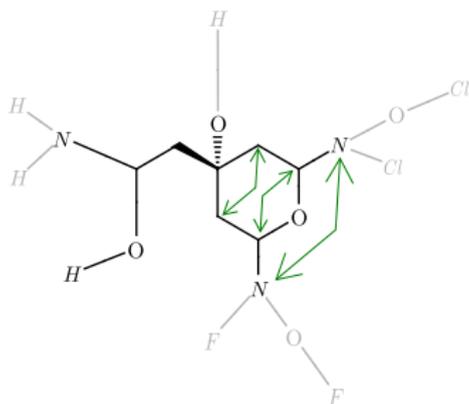




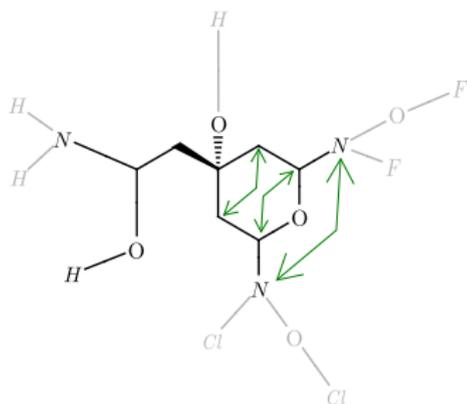
# Stéréo sous-graphe minimal

$$\exists f \in \text{IsomEqOrd}(H, \tau_{i,j}^v(H))$$

- Définition globale.
- Modifications ou suppressions éloignées du stéréo sommet n'ont pas d'influence.
- Modifications ou suppressions proches ont une influence.



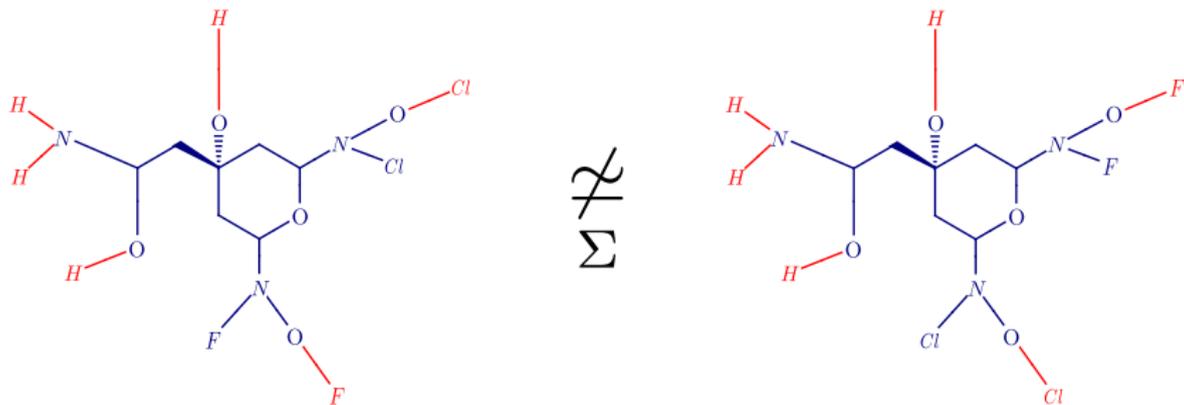
212





$$\nexists f \in \text{IsomEqOrd}(H, \tau_{i,j}^v(H))$$

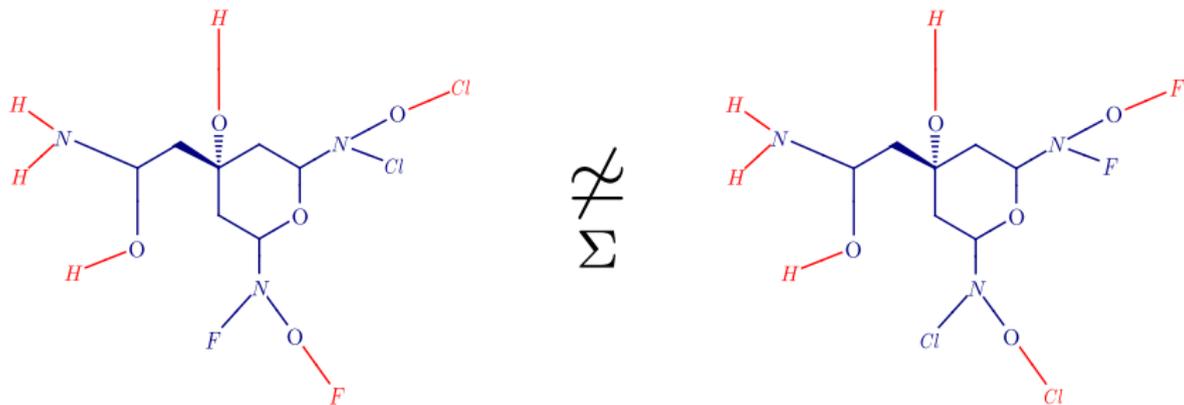
- Définition globale.
- Modifications ou suppressions éloignées du stéréo sommet n'ont pas d'influence.
- Modifications ou suppressions proches ont une influence.
- Le plus petit sous-graphe  $H$  tel qu'il n'y ait pas d'isomorphisme entre  $H$  et  $\tau(H)$ .





$$\nexists f \in \text{IsomEqOrd}(H, \tau_{i,j}^v(H))$$

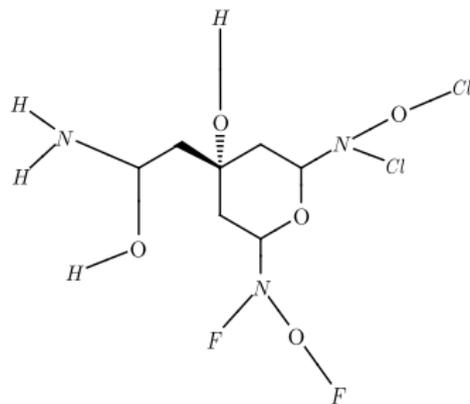
- Définition globale.
- Modifications ou suppressions éloignées du stéréo sommet n'ont pas d'influence.
- Modifications ou suppressions proches ont une influence.
- Le plus petit sous-graphe  $H$  tel qu'il n'y ait pas d'isomorphisme entre  $H$  et  $\tau(H)$ .
- Stéréo sous-graphe minimal.





# Calcul du stéréo sous-graphe minimal

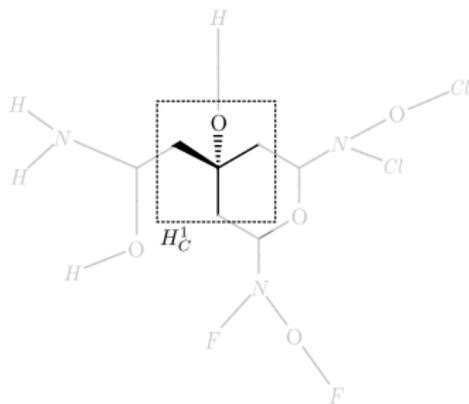
Introduction  
Graphes Ordonnés  
**Noyau Stéréo**  
Voisinages  
Noyau inter stéréo  
Conclusion





# Calcul du stéréo sous-graphe minimal

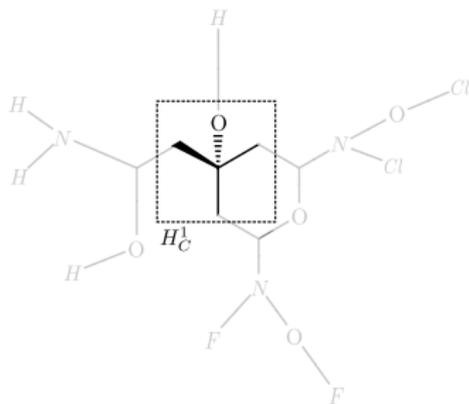
- Initialisation de l'algorithme.





# Calcul du stéréo sous-graphe minimal

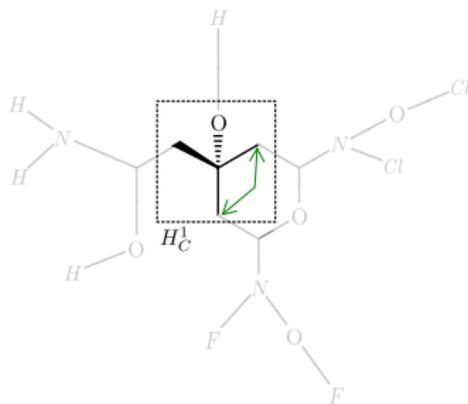
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.





# Calcul du stéréo sous-graphe minimal

- Initialisation de l'algorithme.
- Recherche d'isomorphismes.

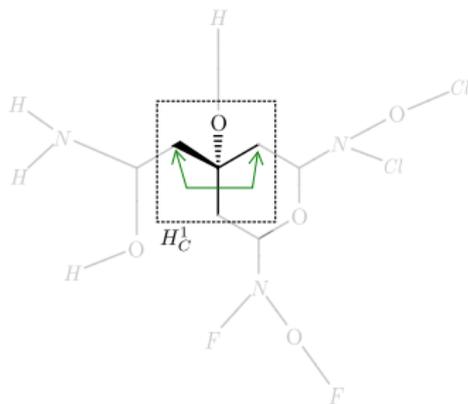






# Calcul du stéréo sous-graphe minimal

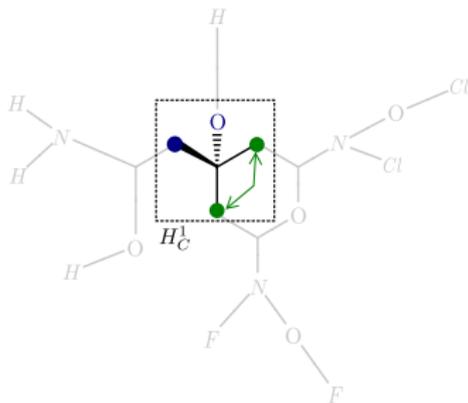
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.





# Calcul du stéréo sous-graphe minimal

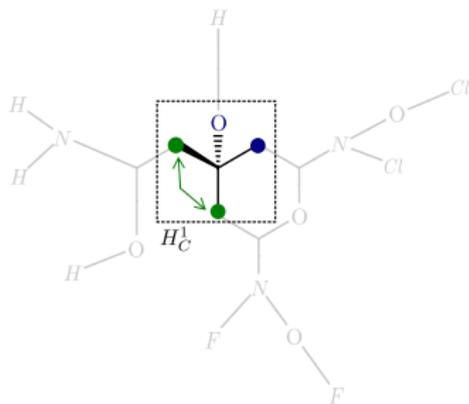
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .





# Calcul du stéréo sous-graphe minimal

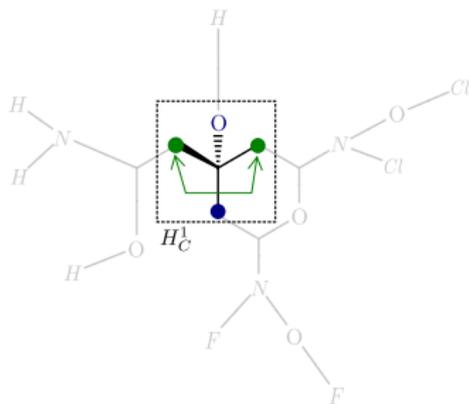
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .





# Calcul du stéréo sous-graphe minimal

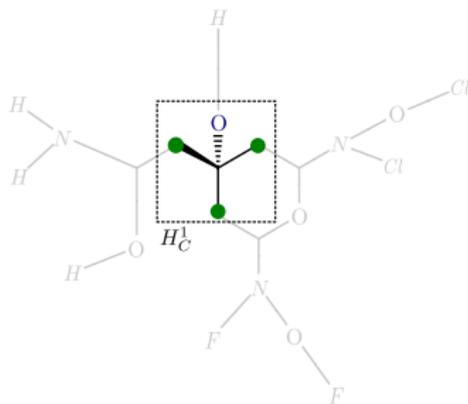
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .





# Calcul du stéréo sous-graphe minimal

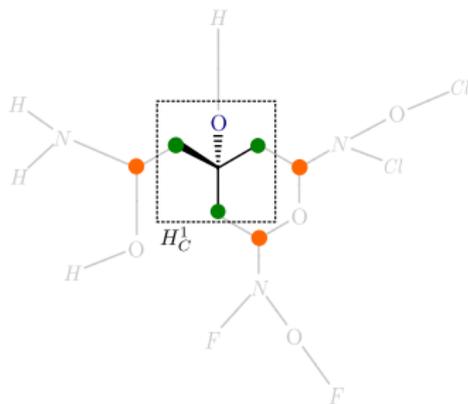
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

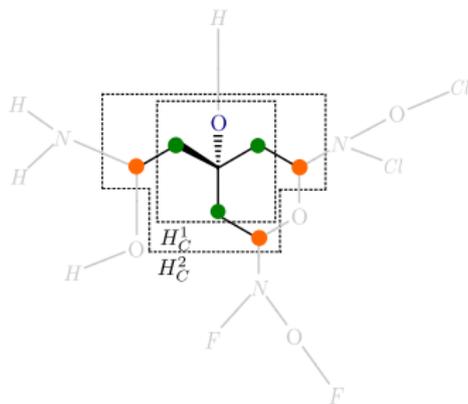
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les **voisins** dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

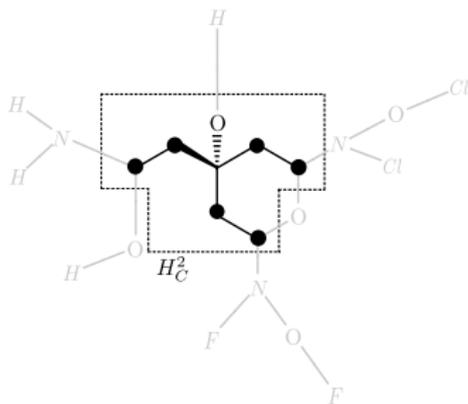
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

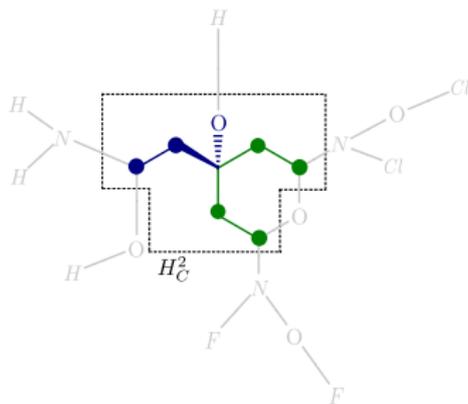
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

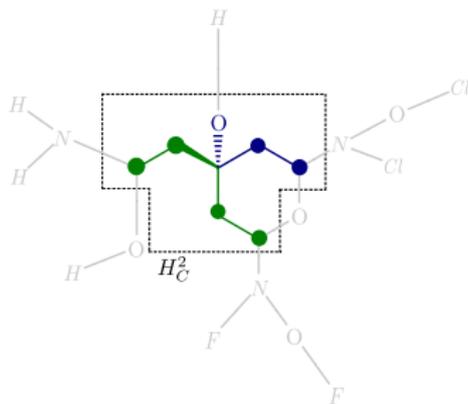
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les **voisins** dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

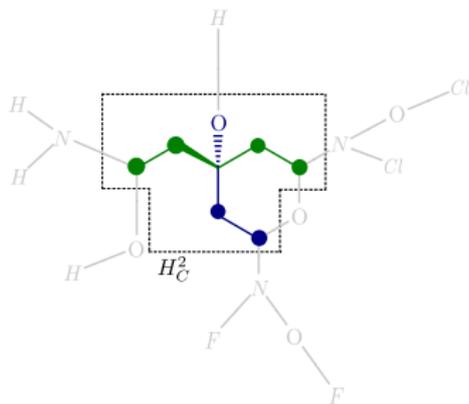
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

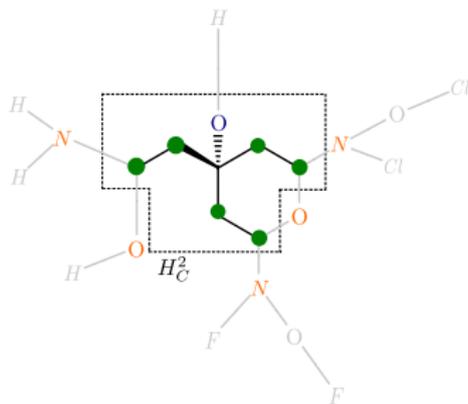
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

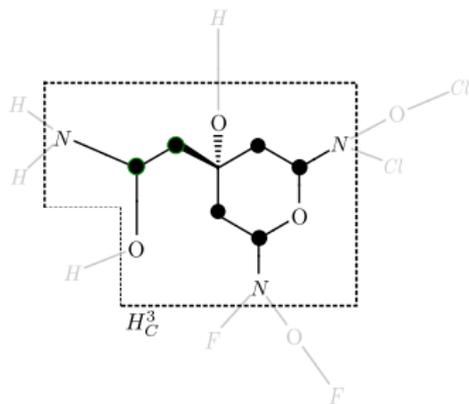
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

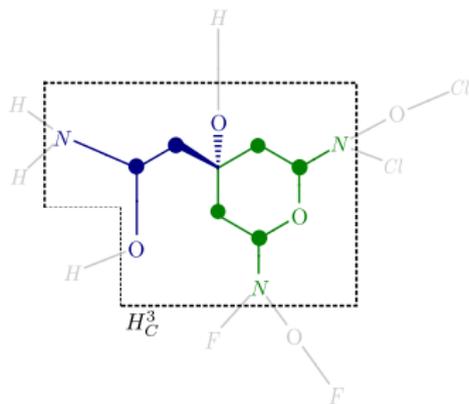
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

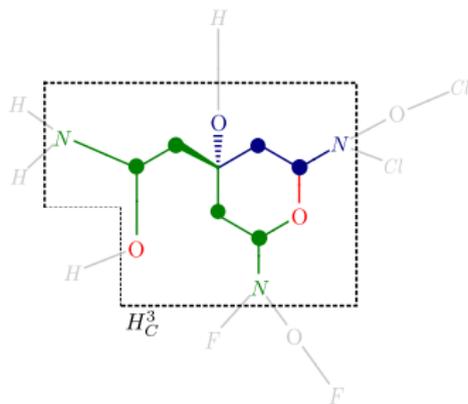
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

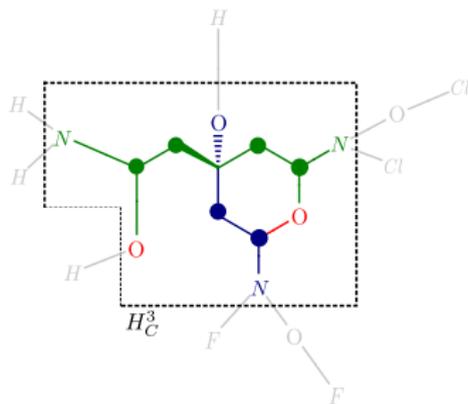
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

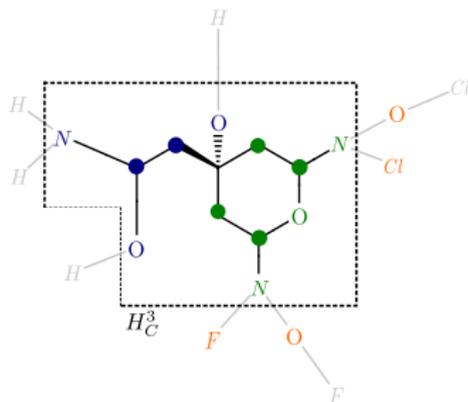
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

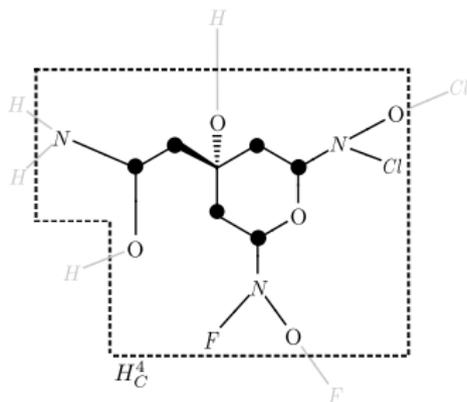
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

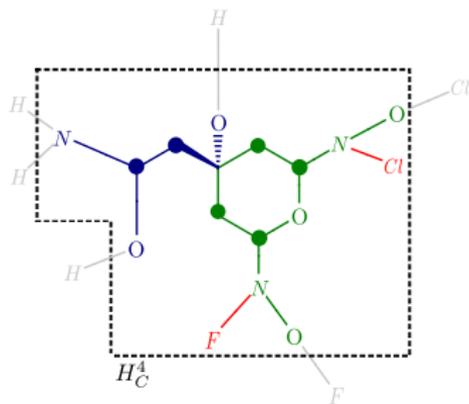
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

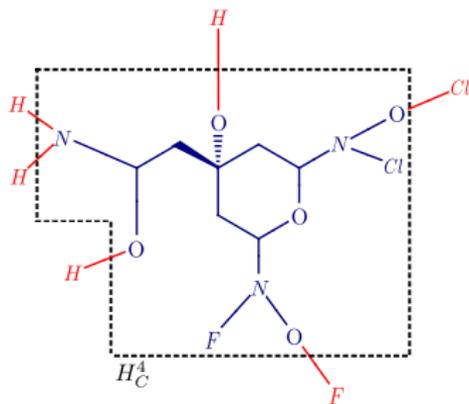
- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Calcul du stéréo sous-graphe minimal

- Initialisation de l'algorithme.
- Recherche d'isomorphismes.
- Ensemble de sommets  $\mathcal{E}_f$  pour l'isomorphisme  $f$ .
- Union de ces ensembles  $\mathcal{E}$ .
- On ajoute les voisins dans G des sommets de  $\mathcal{E}$ .





# Propriétés de l'algorithme de calcul du stéréo sous-graphe minimal

- 1 L'algorithme converge.
- 2 Tant que l'algorithme n'as pas convergé, le sous-graphe ne caractérise pas la stéréoisométrie.
- 3 Le sous-graphe obtenu par l'algorithme caractérise la stéréoisométrie.



## Noyau stéréo :

- Un stéréo sous-graphe minimal  $H_i$  par stéréo sommet  $s_i$ .



## Noyau stéréo :

- Un stéréo sous-graphe minimal  $H_i$  par stéréo sommet  $s_i$ .
- $\mathcal{H}(G) = \{H_i | s_i \in G\}$



## Noyau stéréo :

- Un stéréo sous-graphe minimal  $H_i$  par stéréo sommet  $s_i$ .
- $\mathcal{H}(G) = \{H_i | s_i \in G\}$
- Un unique code représentant  $H_i$ .  
(Stereochemically unique naming algorithm [Wipke, W.T., et al. 1974])



## Noyau stéréo :

- Un stéréo sous-graphe minimal  $H_i$  par stéréo sommet  $s_i$ .
- $\mathcal{H}(G) = \{H_i | s_i \in G\}$
- Un unique code représentant  $H_i$ .  
(Stereochemically unique naming algorithm [Wipke, W.T., et al. 1974])
- $f_H(G)$



## Noyau stéréo :

- Un stéréo sous-graphe minimal  $H_i$  par stéréo sommet  $s_i$ .
- $\mathcal{H}(G) = \{H_i | s_i \in G\}$
- Un unique code représentant  $H_i$ .  
(Stereochemically unique naming algorithm [Wipke, W.T., et al. 1974])
- $f_H(G)$
- Noyau :

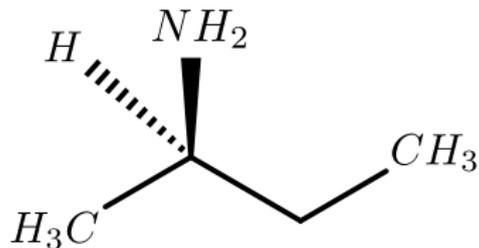
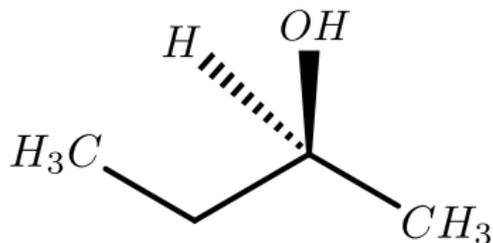
$$k(G, G') = \sum_{H \in \mathcal{H}(G) \cap \mathcal{H}(G')} K(f_H(G), f_H(G')).$$



## Résultats expérimentaux sur le pouvoir rotatoire

Prédiction du pouvoir rotatoire de molécules. [Zhu, H.-J. et al. 2007]

- 35 molécules.
- Écart type du pouvoir rotatoire :  $38,25^\circ$

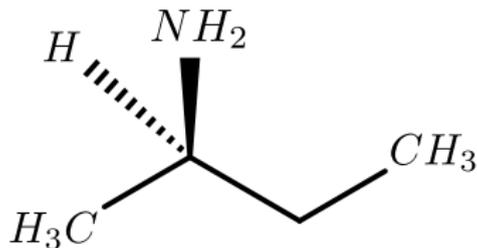
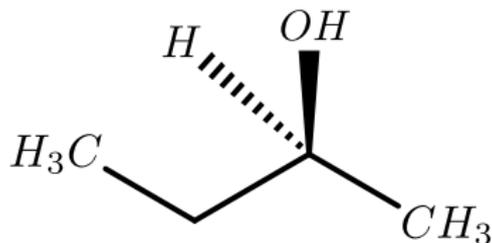




## Résultats expérimentaux sur le pouvoir rotatoire

Prédiction du pouvoir rotatoire de molécules. [Zhu, H.-J. et al. 2007]

- 35 molécules.
- Écart type du pouvoir rotatoire : 38,25°



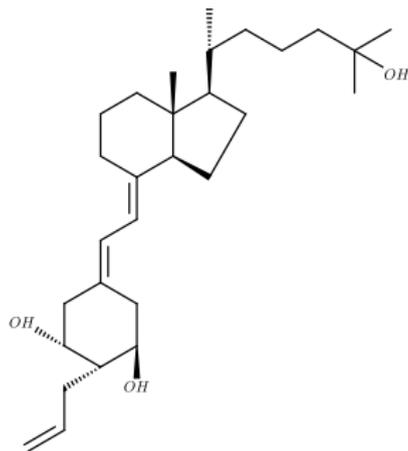
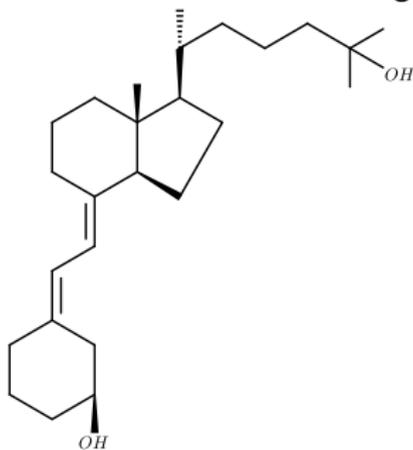
Méthodes	RMSE
Noyau de motifs d'arbres	34,1°
Noyau de Treelets	26,2°
Noyau de motifs d'arbres adapté à la stéréoisomérisation	24,2°
Noyau Stéréo	<b>14,8°</b>



# Résultats expérimentaux sur l'activité biologique

Prédiction de l'activité biologique de dérivés synthétiques de la vitamine D.  
[Brown, J.B. et al. 2010]

- 69 molécules.
- Écart type de l'activité biologique : 0,258
- Entre 7 et 9 centres stéréogènes par molécule.





# Résultats expérimentaux sur l'activité biologique

Prédiction de l'activité biologique de dérivés synthétiques de la vitamine D.  
[Brown, J.B. et al. 2010]

- 69 molécules.
- Écart type de l'activité biologique : 0,258
- Entre 7 et 9 centres stéréogènes par molécule.

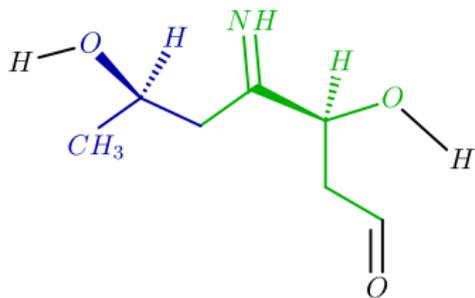
Méthodes	Erreurs moyennes	RMSE
Noyau de motifs d'arbres	0,193	0,251
Noyau de Treelets	0,208	0,271
Noyau de motifs d'arbres adapté à la stéréoisométrie	<i>0,138</i>	<b>0,184</b>
Noyau Stéréo	<b>0,134</b>	<i>0,194</i>



# Avantages et inconvénients du noyau stéréo

Introduction  
Graphes Ordonnés  
**Noyau Stéréo**  
Voisinages  
Noyau inter stéréo  
Conclusion

+ Encode uniquement la stéréoisométrie.

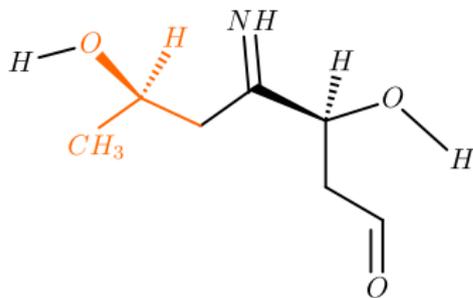
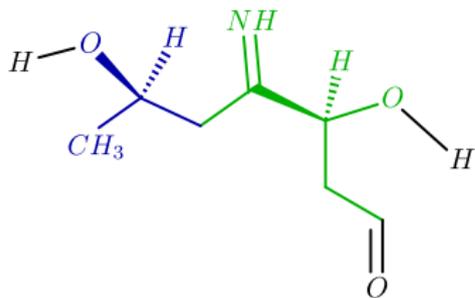




# Avantages et inconvénients du noyau stéréo

Introduction  
Graphes Ordonnés  
**Noyau Stéréo**  
Voisinsages  
Noyau inter stéréo  
Conclusion

+ Encode uniquement la stéréoisométrie.

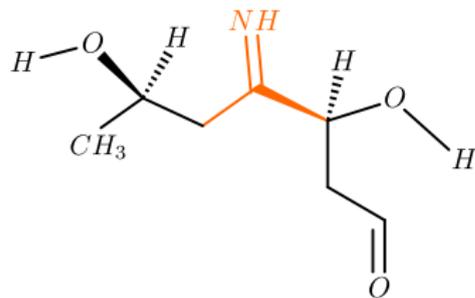
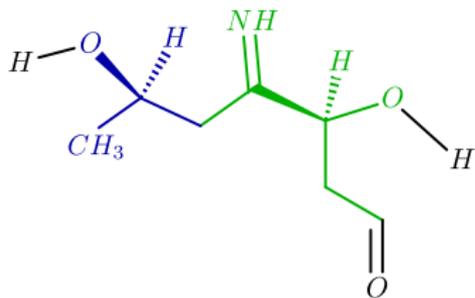




# Avantages et inconvénients du noyau stéréo

Introduction  
Graphes Ordonnés  
**Noyau Stéréo**  
Voisinsages  
Noyau inter stéréo  
Conclusion

+ Encode uniquement la stéréoisométrie.

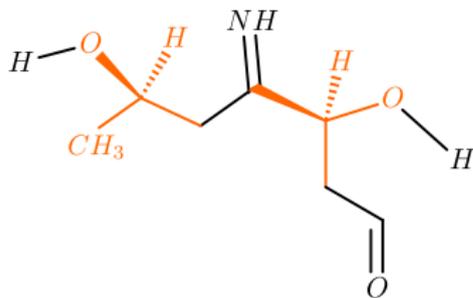
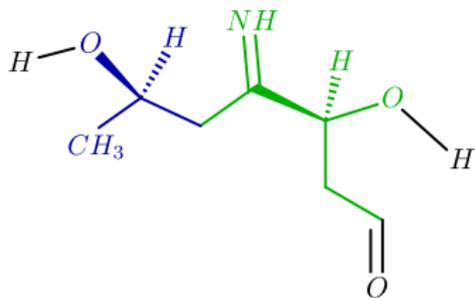




# Avantages et inconvénients du noyau stéréo

Introduction  
Graphes Ordonnés  
**Noyau Stéréo**  
Voisinages  
Noyau inter stéréo  
Conclusion

- + Encode uniquement la stéréoisométrie.
- + Pas de taille prédéfinie pour les motifs.

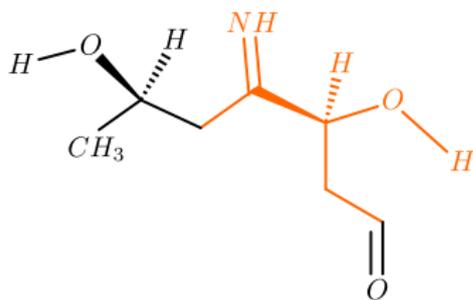
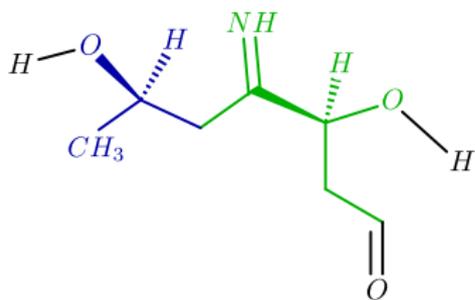




# Avantages et inconvénients du noyau stéréo

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisines  
Noyau inter stéréo  
Conclusion

- + Encode uniquement la stéréoisométrie.
- + Pas de taille prédéfinie pour les motifs.
- Stéréo sous-graphes minimaux considérés de manière isolée.

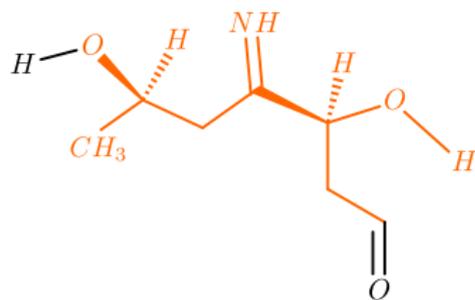
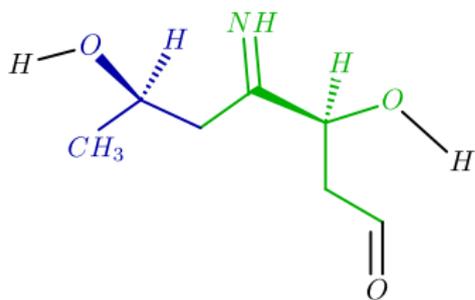




# Avantages et inconvénients du noyau stéréo

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinnages  
Noyau inter stéréo  
Conclusion

- + Encode uniquement la stéréoisométrie.
- + Pas de taille prédéfinie pour les motifs.
- Stéréo sous-graphes minimaux considérés de manière isolée.





# Avantages et inconvénients du noyau stéréo

Introduction  
Graphes Ordonnés  
**Noyau Stéréo**  
Voisinages  
Noyau inter stéréo  
Conclusion

Méthodes	Temps de calcul	A priori	Prédiction
Descripteurs moléculaires			
Polynômes chiraux			
Noyau de motifs d'arbres			
Noyau Stéréo			



# Avantages et inconvénients du noyau stéréo

Méthodes	Temps de calcul	A priori	Prédiction
Descripteurs moléculaires			
Polynômes chiraux			
Noyau de motifs d'arbres			
Noyau Stéréo			

Méthodes	Temps de calcul en secondes
Noyau de motifs d'arbres	230
Noyau Stéréo	1



# Avantages et inconvénients du noyau stéréo

Introduction  
Graphes Ordonnés  
**Noyau Stéréo**  
Voisinages  
Noyau inter stéréo  
Conclusion

Méthodes	Temps de calcul	A priori	Prédiction
Descripteurs moléculaires			
Polynômes chiraux			
Noyau de motifs d'arbres			
Noyau Stéréo			



# Avantages et inconvénients du noyau stéréo

Méthodes	Temps de calcul	A priori	Prédiction
Descripteurs moléculaires			
Polynômes chiraux			
Noyau de motifs d'arbres			
Noyau Stéréo			

Méthodes	Stéréo dans tous les cas	Voisinage
Noyau de motifs d'arbres		
Noyau Stéréo		



# Prise en compte des voisinages des stéréo sous-graphes minimaux



# Fonctions d'interactions

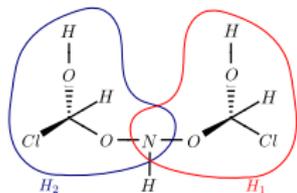
Introduction  
Graphes Ordonnés  
Noyau Stéréo  
**Voisinages**  
Noyau inter stéréo  
Conclusion

$$F_i(H_1, H_2) = \max\{j \mid c_j\}$$



$$F_i(H_1, H_2) = \max\{j \mid c_j\}$$

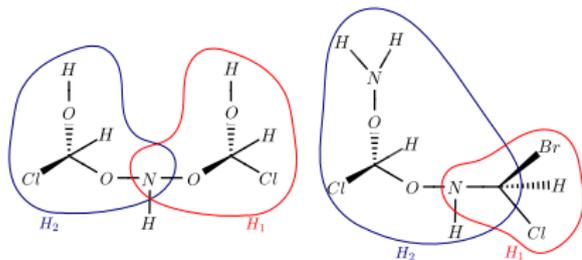
- $c_1 : H_1 \cap H_2 \neq \emptyset$





$$F_i(H_1, H_2) = \max\{j \mid c_j\}$$

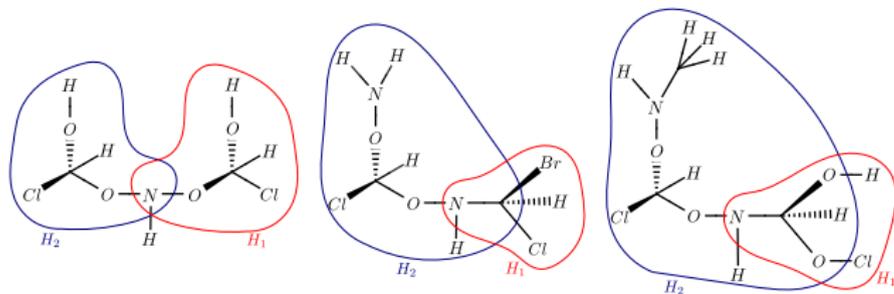
- $c_1 : H_1 \cap H_2 \neq \emptyset$
- $c_2 : s_1 \subset H_2$





$$F_i(H_1, H_2) = \max\{j \mid c_j\}$$

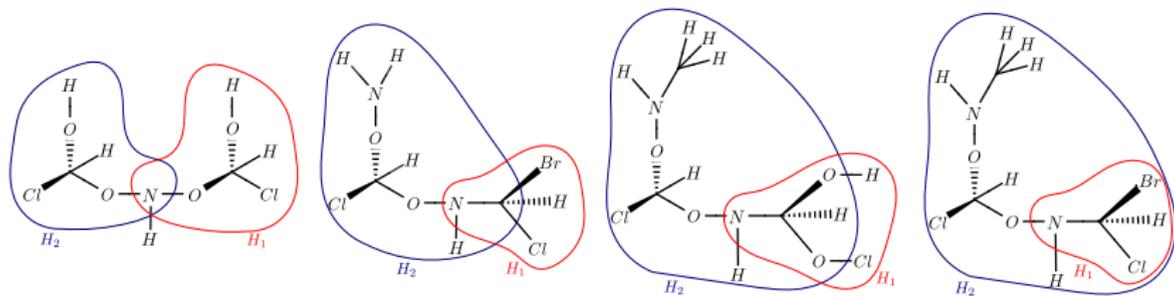
- $c_1 : H_1 \cap H_2 \neq \emptyset$
- $c_2 : s_1 \subset H_2$
- $c_3 : N(s_1) \subset H_2$





$$F_i(H_1, H_2) = \max\{j \mid c_j\}$$

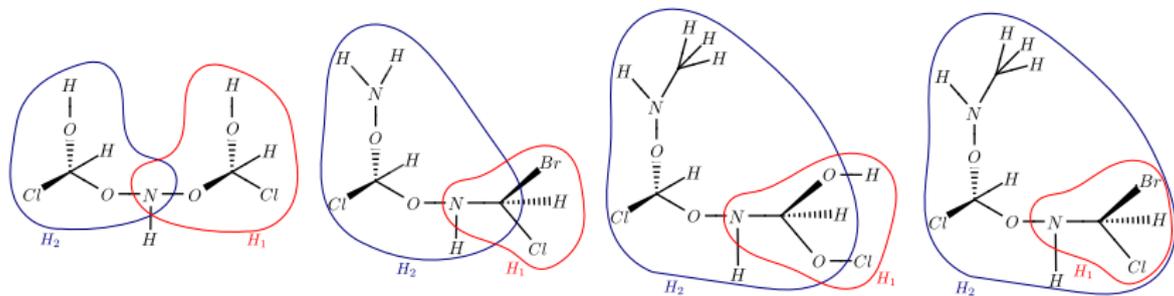
- $c_1 : H_1 \cap H_2 \neq \emptyset$
- $c_2 : s_1 \subset H_2$
- $c_3 : N(s_1) \subset H_2$
- $c_4 : H_1 \subset H_2$





$$F_1(H_1, H_2) = \max\{j \in \{0, 1, 2, 3, 4\} \mid c_j\}$$

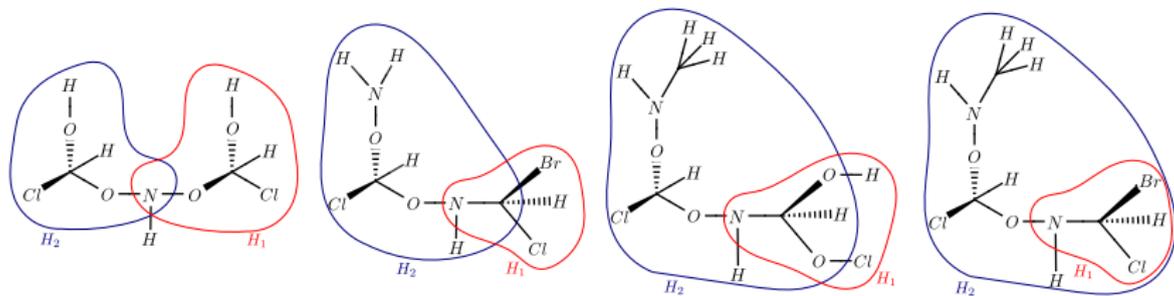
- $c_1 : H_1 \cap H_2 \neq \emptyset$
- $c_2 : s_1 \subset H_2$
- $c_3 : N(s_1) \subset H_2$
- $c_4 : H_1 \subset H_2$
- $c_0 : \neg c_1$





$$F_2(H_1, H_2) = \max\{j \in \{0, 2, 3, 4\} \mid c_j\}$$

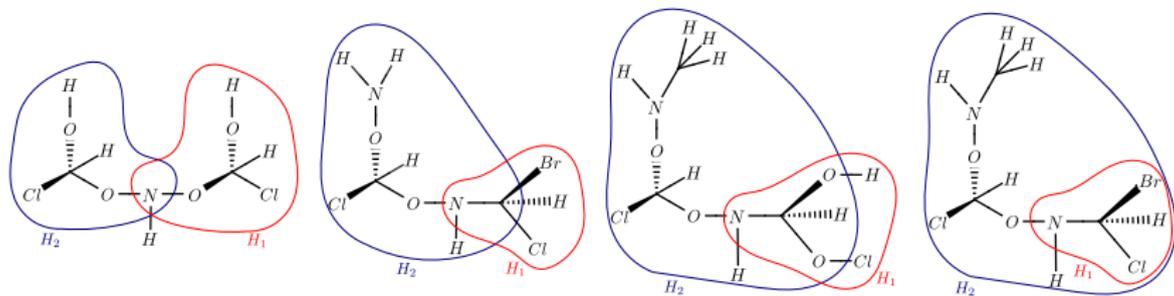
- $c_2 : s_1 \subset H_2$
- $c_3 : N(s_1) \subset H_2$
- $c_4 : H_1 \subset H_2$
- $c_0 : \neg c_2$





$$F_3(H_1, H_2) = \max\{j \in \{0, 3, 4\} \mid c_j\}$$

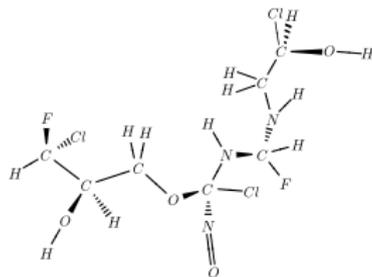
- $c_3 : N(s_1) \subset H_2$
- $c_4 : H_1 \subset H_2$
- $c_0 : \neg c_3$





# Graphes d'interactions

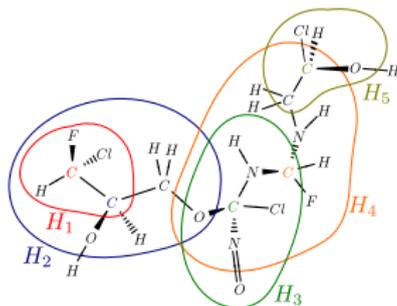
Introduction  
Graphes Ordonnés  
Noyau Stéréo  
**Voisinages**  
Noyau inter stéréo  
Conclusion





# Graphes d'interactions

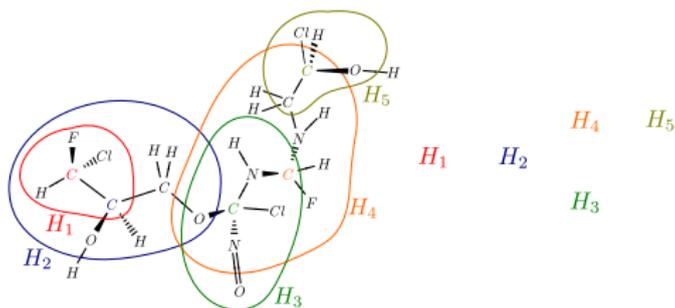
Introduction  
Graphes Ordonnés  
Noyau Stéréo  
**Voisinages**  
Noyau inter stéréo  
Conclusion





# Graphes d'interactions

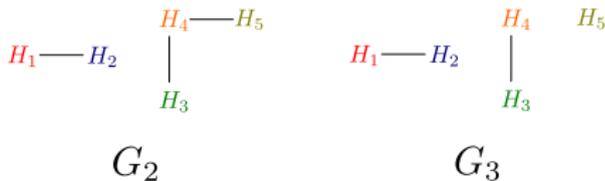
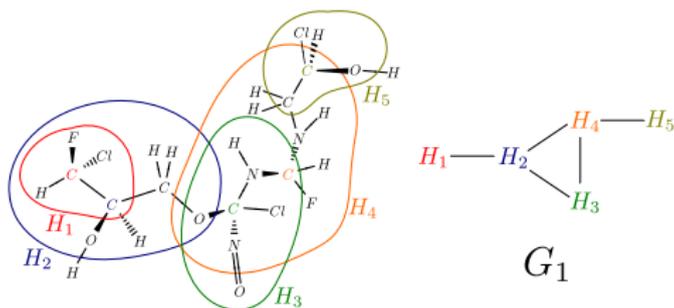
Introduction  
Graphes Ordonnés  
Noyau Stéréo  
**Voisinages**  
Noyau inter stéréo  
Conclusion



- Un sommet par stéréo sous-graphe minimal.



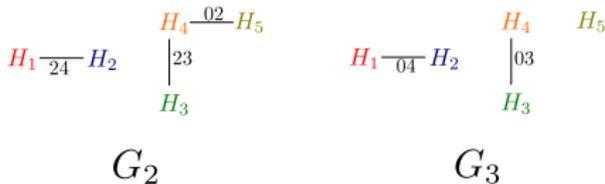
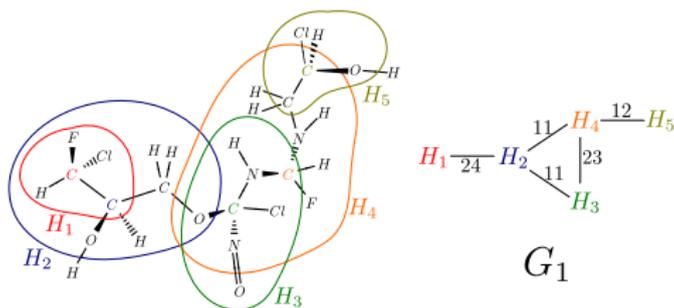
# Graphes d'interactions



- Un sommet par stéréo sous-graphe minimal.
- Une arête si la fonction d'interactions n'est pas nulle.



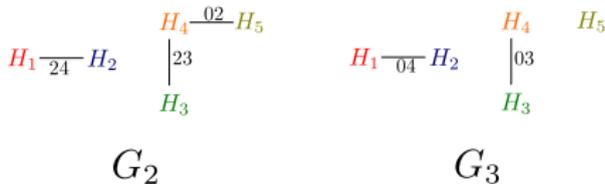
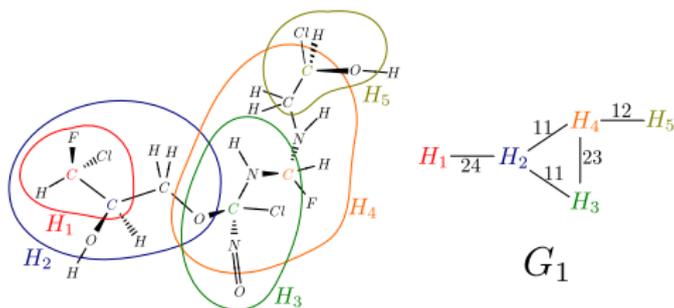
# Graphes d'interactions



- Un sommet par stéréo sous-graphe minimal.
- Une arête si la fonction d'interactions n'est pas nulle.
- Le label de l'arête est la concaténation des fonctions d'interactions.



# Graphes d'interactions

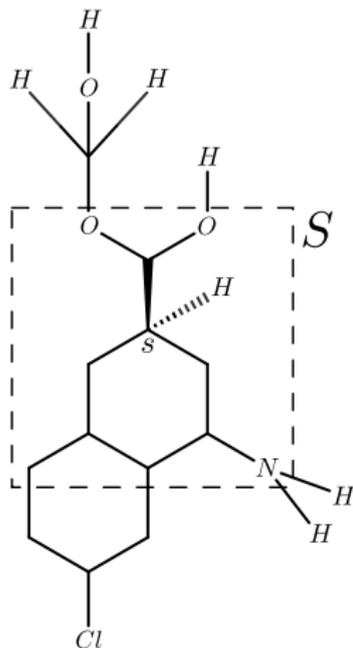


- Un sommet par stéréo sous-graphe minimal.
- Une arête si la fonction d'interactions n'est pas nulle.
- Le label de l'arête est la concaténation des fonctions d'interactions.
- On utilise des noyaux sur graphes usuels.



## K-voisinage d'un sommet

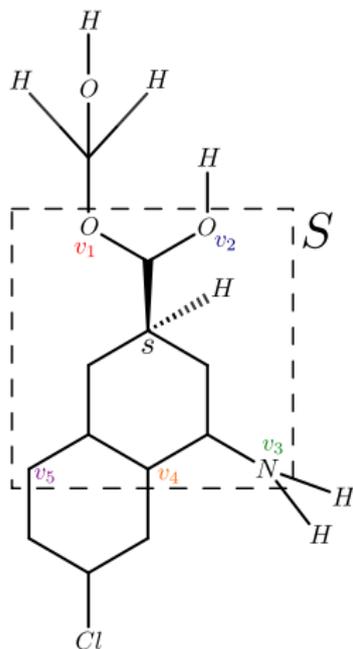
Approche locale de la prise en compte du voisinage.





# K-voisinage d'un sommet

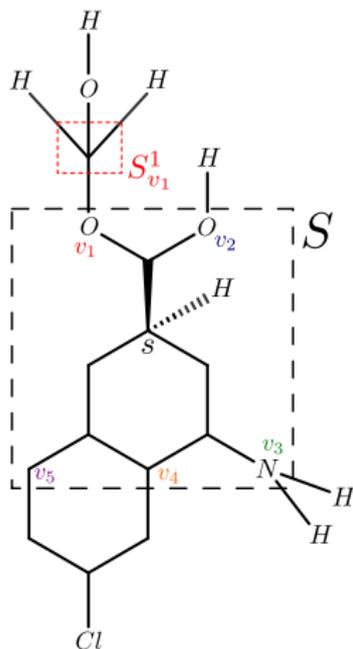
Approche locale de la prise en compte du voisinage.





# K-voisinage d'un sommet

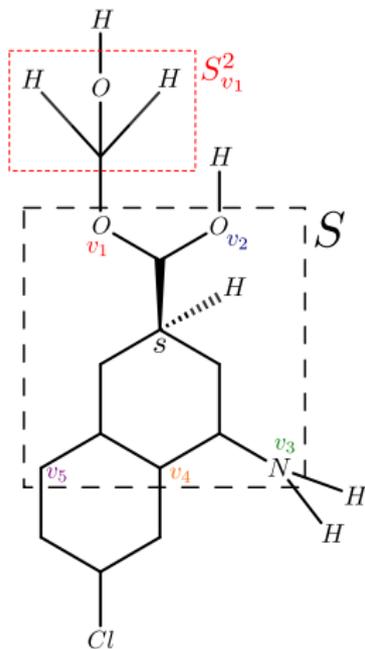
Approche locale de la prise en compte du voisinage.





# K-voisinage d'un sommet

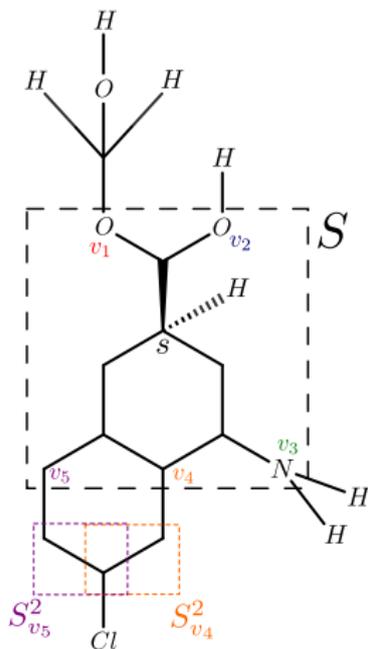
Approche locale de la prise en compte du voisinage.





# K-voisinage d'un sommet

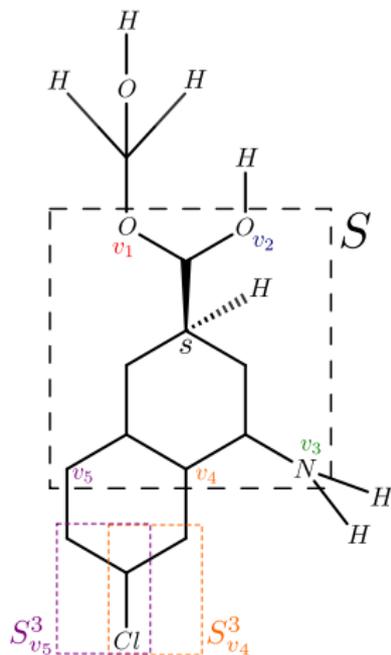
Approche locale de la prise en compte du voisinage.





# K-voisinage d'un sommet

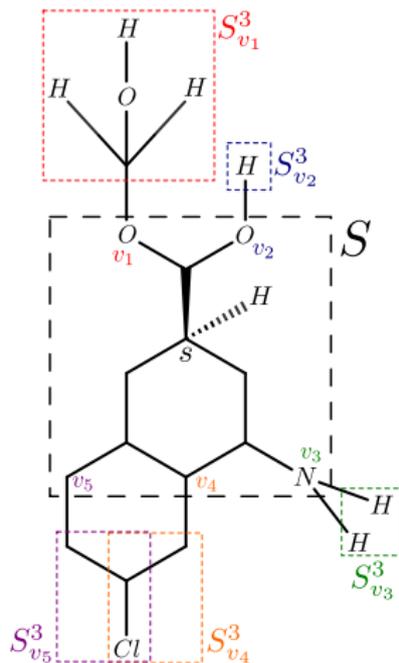
Approche locale de la prise en compte du voisinage.





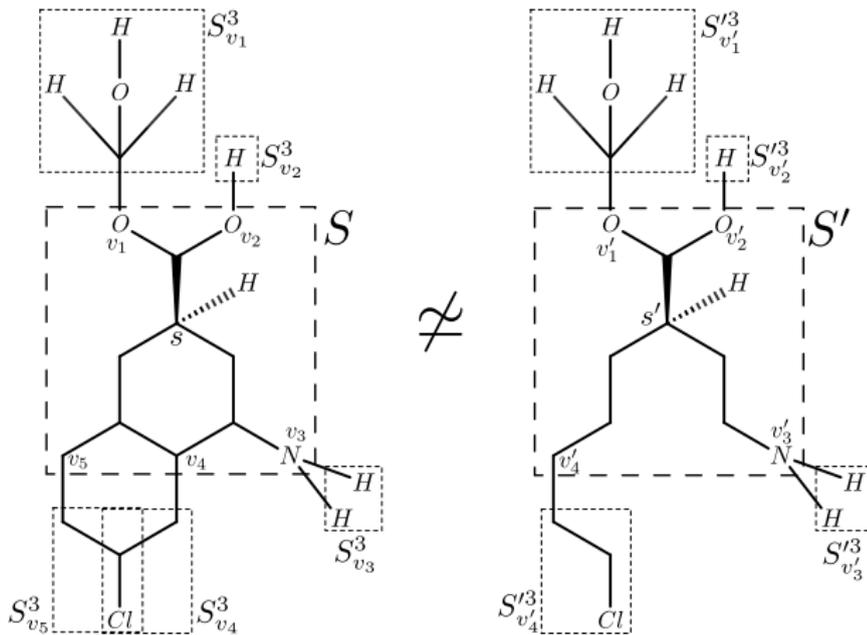
# K-voisinage d'un sommet

Approche locale de la prise en compte du voisinage.



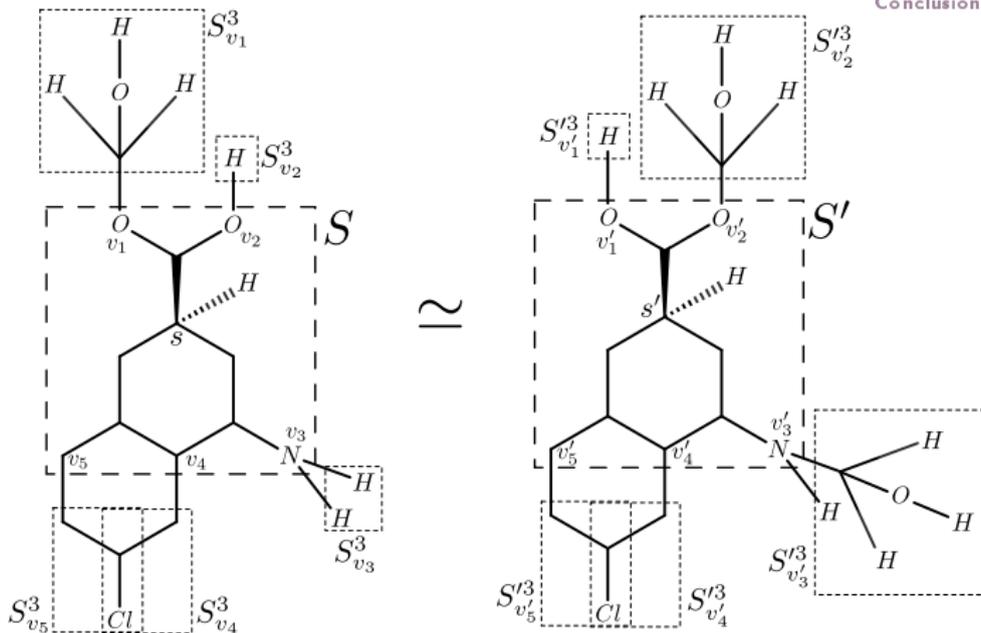


# Noyau d'influence





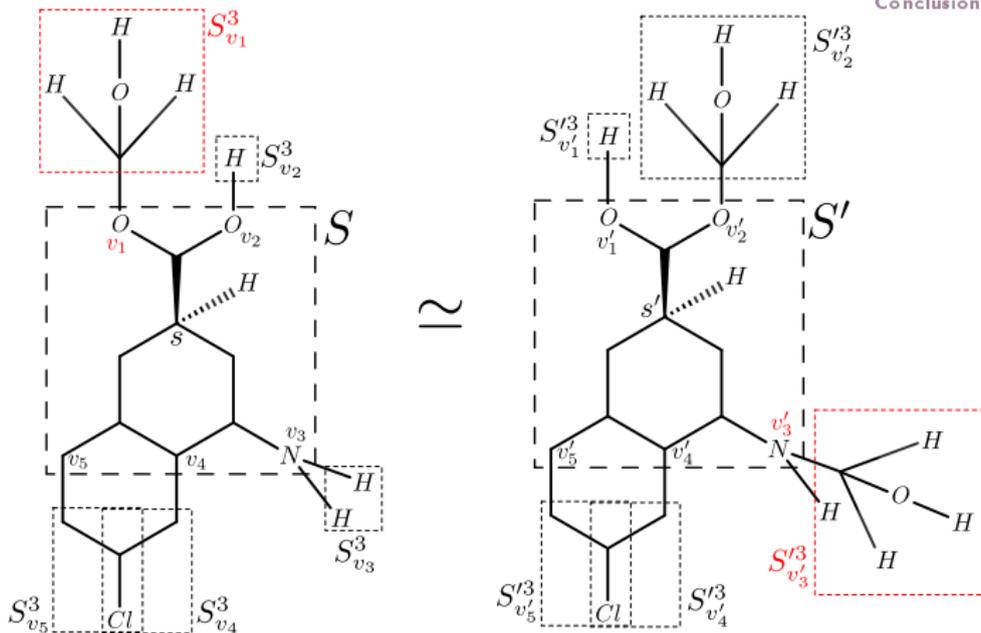
# Noyau d'influence



$$k_{influence}(S, S') = \delta(S, S')$$



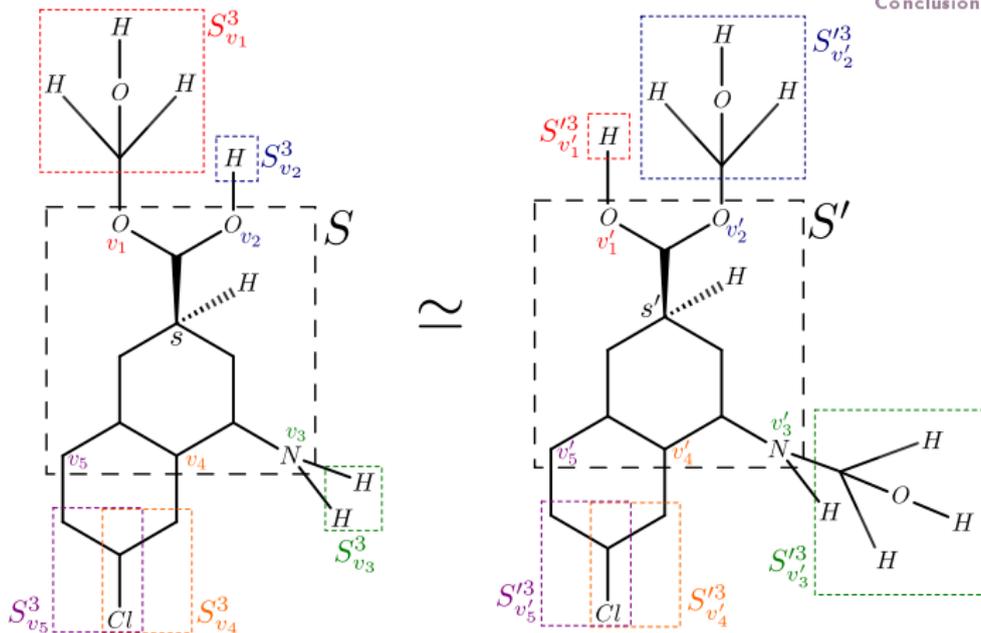
# Noyau d'influence



$$k_{influence}(S, S') = \delta(S, S')$$



# Noyau d'influence

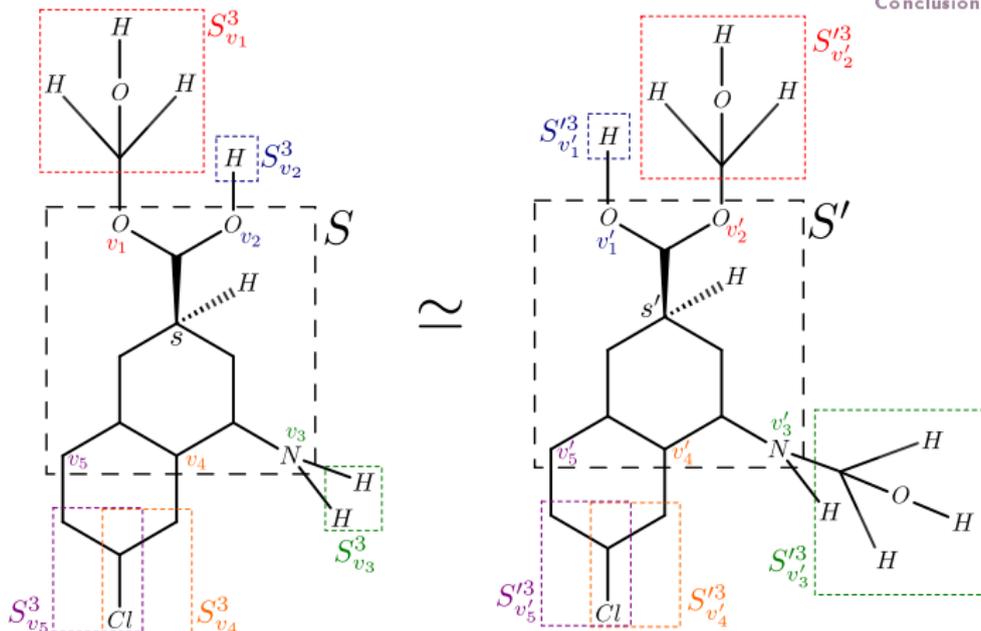


$$k_{influence}(S, S') = \delta(S, S')$$

$$\prod_{i=1}^n k_t(S_{v_i}^k, S'_{f(v_i)}^k)$$



# Noyau d'influence



$$k_{influence}(S, S') = \delta(S, S') \sum_{f \in \text{IsomEqOrd}(s, s')} \prod_{i=1}^n k_t(S^k_{v_i}, S'^k_{f(v_i)})$$

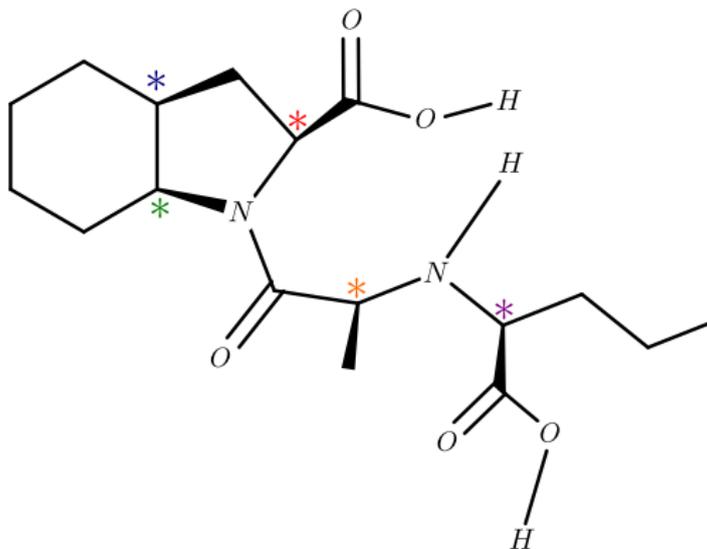


# Résultats expérimentaux sur la classification des stéréoisomères de la périndoprilate

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
**Voisinsages**  
Noyau inter stéréo  
Conclusion

Classification des stéréoisomères de la périndoprilate en fonction de leur action sur l'enzyme de conversion de l'angiotensine. [Castillo-Garit J.A., et al. 2007]

- 32 molécules.
- 5 centres stéréogènes par molécule.





# Résultats expérimentaux sur la classification des stéréoisomères de la périndoprilate

Classification des stéréoisomères de la périndoprilate en fonction de leur action sur l'enzyme de conversion de l'angiotensine. [Castillo-Garit J.A., et al. 2007]

- 32 molécules.
- 5 centres stéréogènes par molécule.

Méthodes	Précision
Noyau de motifs d'arbres adapté à la stéréoisométrie	96,9
Noyau Stéréo	87,5
Graphes d'interactions $G_1$ avec noyau de treelets et MKL	<b>100</b>



# Résultats expérimentaux sur l'activité biologique

Prédiction de l'activité biologique de dérivés synthétiques de la vitamine D.

- 69 molécules.
- Écart type de l'activité biologique : 0,258
- Entre 7 et 9 centres stéréogènes par molécule.



Prédiction de l'activité biologique de dérivés synthétiques de la vitamine D.

- 69 molécules.
- Écart type de l'activité biologique : 0,258
- Entre 7 et 9 centres stéréogènes par molécule.

Méthodes	Erreurs moyennes	RMSE
Noyau de motifs d'arbres adapté à la stéréoisomérisation	0,138	0,184
Noyau Stéréo	0,134	0,194
Graphes d'interactions $G_3$ avec noyau de motifs d'arbres	<b>0,107</b>	<b>0,161</b>
Noyau d'influence	<i>0,131</i>	<i>0,177</i>



# Mesure de similarité entre des stéréo sous-graphes minimaux



# Principe de similarité

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
**Noyau inter stéréo**  
Conclusion

« *Des molécules similaires ont des propriétés similaires.* »



« *Des molécules similaires ont des propriétés similaires.* »

« *Des motifs similaires ont des influences similaires sur les propriétés.* »



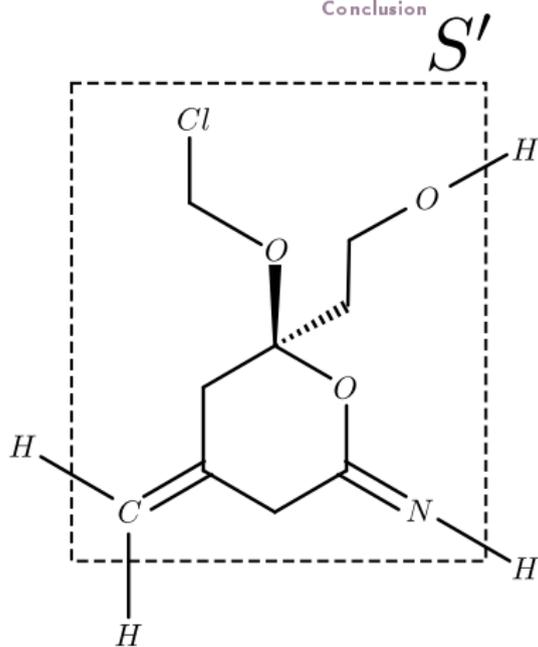
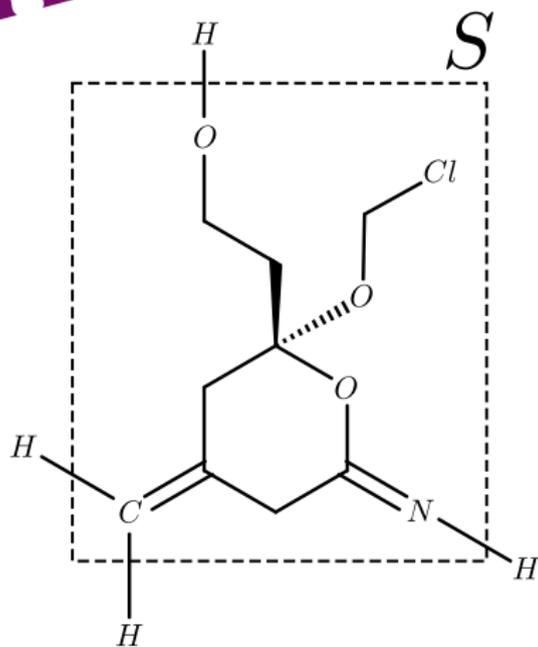
« *Des molécules similaires ont des propriétés similaires.* »

« *Des stéréo sous-graphes minimaux similaires ont des influences similaires sur les propriétés.* »



# Comparaison de sous-graphes

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

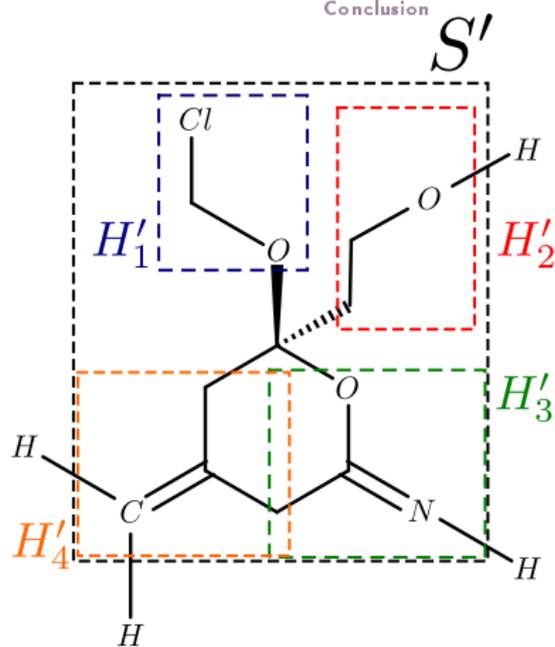
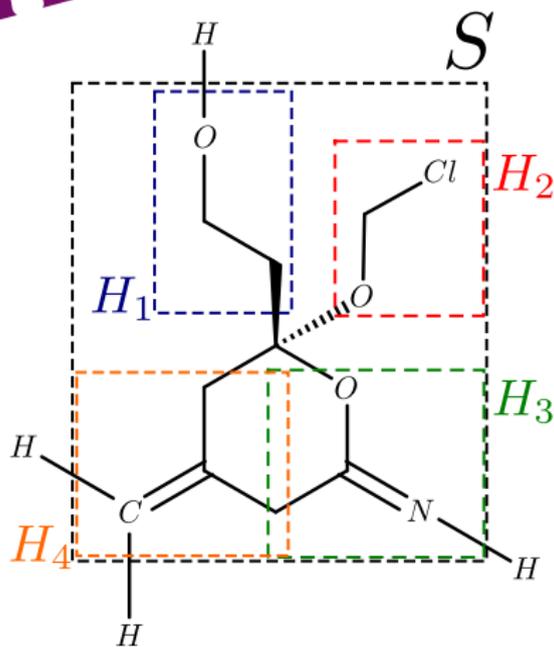






# Comparaison de sous-graphes

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion

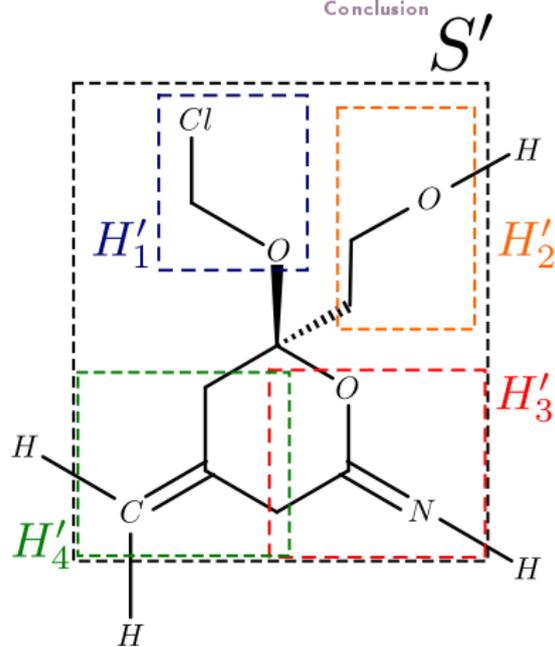
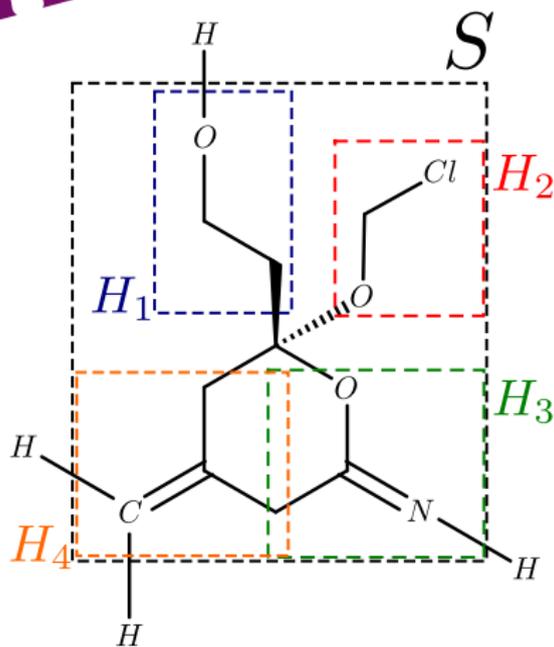


$$k(S, S') = \sum_{\sigma \in \Sigma^{S'}} \prod_{i=1}^4 k_t(H_i, H'_{\varphi_{\sigma}(i)})$$



# Comparaison de sous-graphes

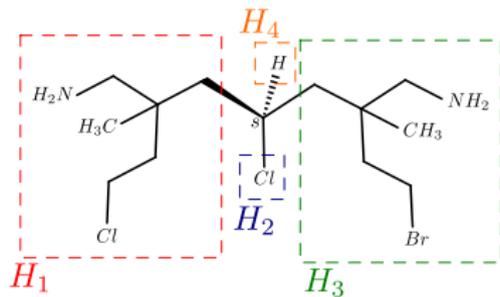
Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
Conclusion



$$k(S, S') = \sum_{\sigma \in \Sigma^{S'}} \prod_{i=1}^4 k_t(H_i, H'_{\varphi_\sigma(i)})$$

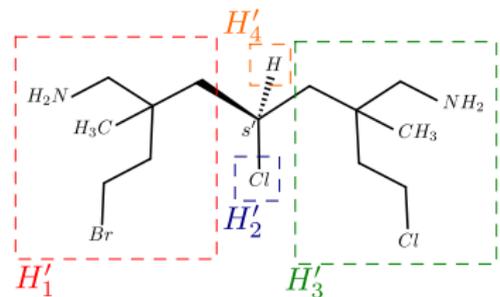
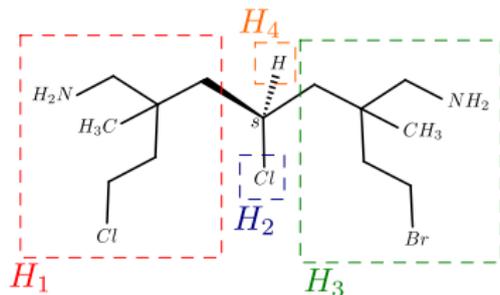


# Problème des stéréoisomères aux orientations opposées



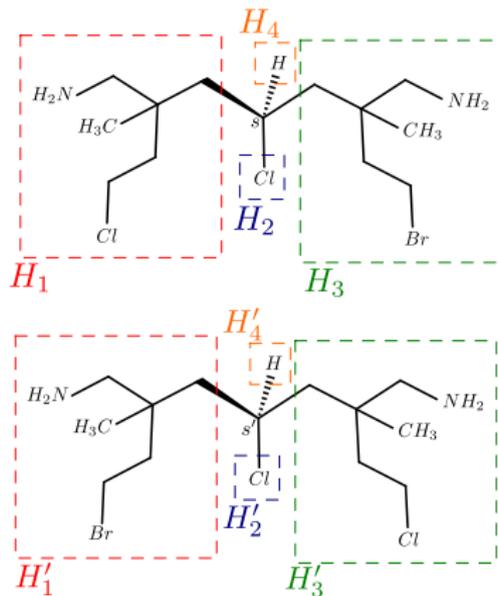


# Problème des stéréoisomères aux orientations opposées





# Problème des stéréoisomères aux orientations opposées

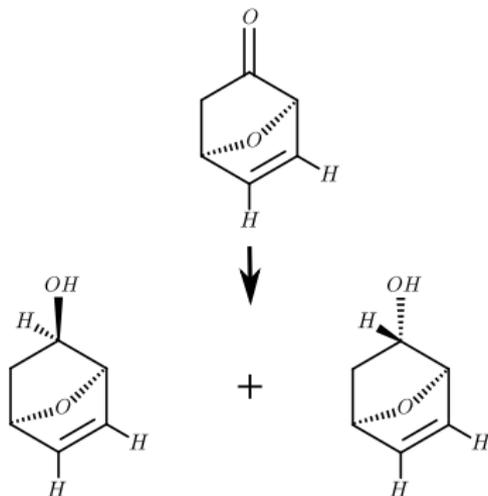


$$k_{inter}(S, S') = k_{norm}(S, S') + \delta(S, S') - \delta(S, \tau(S'))$$



Prédiction du stéréoisomère le plus présent après une réduction. [Suo J.J., et al. 2013]

- 67 paires de molécules.





# Résultats expérimentaux sur la classification

Prédiction du stéréoisomère le plus présent après une réduction. [Suo J.J., et al. 2013]

- 67 paires de molécules.

Méthode	Précision	Taux de paires non classées
Noyau de motifs d'arbres adapté à la stéréoisométrie	80,6%	4,5%
Noyau Stéréo	58,2%	38,8%
Noyau Inter Stéréo avec sélection	<b>88,1%</b>	<b>0%</b>



# Conclusion



## Conclusion et perspective

Introduction  
Graphes Ordonnés  
Noyau Stéréo  
Voisinages  
Noyau inter stéréo  
**Conclusion**

+ Modèle représentant les stéréoisomères.



## Conclusion et perspective

- + Modèle représentant les stéréoisomères.
- + Mesure de similarité générale entre les stéréoisomères.



## Conclusion et perspective

- + Modèle représentant les stéréoisomères.
- + Mesure de similarité générale entre les stéréoisomères.

Méthodes	Stéréo dans tous les cas	Voisinage	Inter
Noyau de motifs d'arbres			-
Noyau Stéréo			



## Conclusion et perspective

- + Modèle représentant les stéréoisomères.
- + Mesure de similarité générale entre les stéréoisomères.
- + Extensions pour prendre en compte les voisinages des stéréo sous-graphes minimaux.

Méthodes	Stéréo dans tous les cas	Voisinage	Inter
Noyau de motifs d'arbres			-
Noyau Stéréo			



## Conclusion et perspective

- + Modèle représentant les stéréoisomères.
- + Mesure de similarité générale entre les stéréoisomères.
- + Extensions pour prendre en compte les voisinages des stéréo sous-graphes minimaux.
- + Extension pour comparer des stéréo sous-graphes minimaux non identiques.

Méthodes	Stéréo dans tous les cas	Voisinage	Inter
Noyau de motifs d'arbres			-
Noyau Stéréo			



- + Modèle représentant les stéréoisomères.
- + Mesure de similarité générale entre les stéréoisomères.
- + Extensions pour prendre en compte les voisinages des stéréo sous-graphes minimaux.
- + Extension pour comparer des stéréo sous-graphes minimaux non identiques.

Méthodes	Stéréo dans tous les cas	Voisinage	Inter
Noyau de motifs d'arbres			-
Noyau Stéréo			

- Les extensions sont indépendantes.



- + Modèle représentant les stéréoisomères.
- + Mesure de similarité générale entre les stéréoisomères.
- + Extensions pour prendre en compte les voisinages des stéréo sous-graphes minimaux.
- + Extension pour comparer des stéréo sous-graphes minimaux non identiques.

Méthodes	Stéréo dans tous les cas	Voisinage	Inter
Noyau de motifs d'arbres			-
Noyau Stéréo			

- Les extensions sont indépendantes.
  - Comme on compare des stéréo sous-graphes minimaux non identiques, on ne peut plus compter les motifs pour les graphes d'interactions.



Merci !

Questions ?