

# Chimica Farmaceutica

Modelli 3-D QSAR  
attraverso il portale [www-3d-qsar.com](http://www-3d-qsar.com)



SAPIENZA  
UNIVERSITÀ DI ROMA

## Introduzione

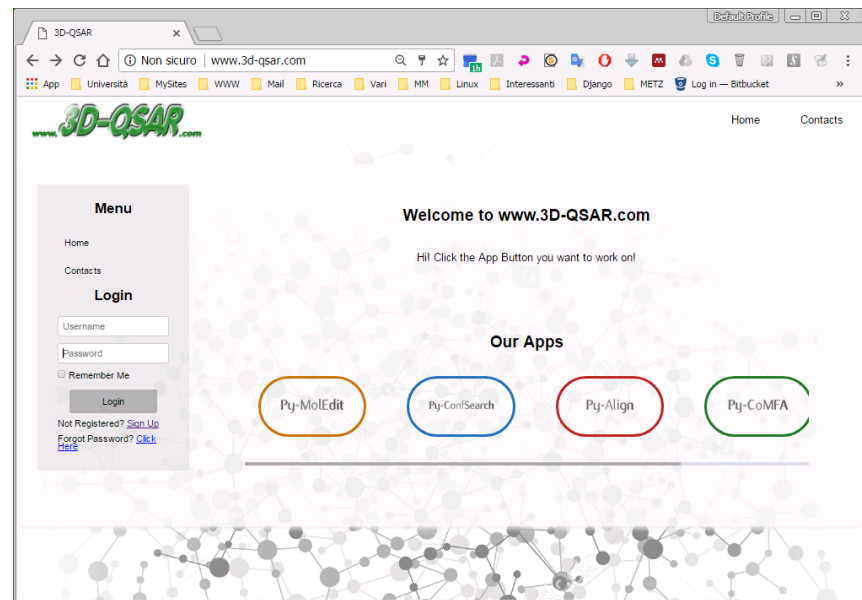
Il portale [www.3d-qsar.com](http://www.3d-qsar.com) è al momento un contenitore di una serie di applicazioni web interconnesse mediante le quali è possibile costruire dei modelli 3-D QSAR a partire da una serie di molecole delle quali è sufficiente conoscere la loro struttura chimica e l'attività biologica associata verso un qualsiasi target biologico.

Il portale è agevolmente utilizzabile attraverso qualsiasi dispositivo che si collega ad internet.

# Le applicazioni web del portale

Il portale al momento della scrittura di questo documento si compone di quattro applicazioni web:

1. Py-MolEdit
2. Py-ConfSearch
3. Py-Align
4. Py-CoMFA



## Le applicazioni web del portale: Py-MolEdit

Py-MolEdit è un'applicazione che si basa sull'integrazione di JSME (Javascript Molecular Editor - <http://peter-ertl.com/jsme/>) in cui è possibile disegnare molecole direttamente in una finestra di un programma per navigazione in rete (google-chrome, internet explorer, firefox, ecc.).

In [www.3d-qsar.com](http://www.3d-qsar.com) JSME è stato integrato mediante il linguaggio python in modo tale che le molecole disegnate possano essere automaticamente convertite in conformazioni tridimensionali e salvate in un dataset appositamente creato.

## Le applicazioni web del portale: Py-ConfSearch

Py-ConfSearch è un'applicazione che si basa sull'integrazione di alcuni algoritmi di calcolo per eseguire delle analisi conformazionali. Attualmente sono implementati tre differenti metodi:

1. Balloon
2. Openbabel
3. RdKit

## Le applicazioni web del portale: Py-Align

Py-Align è un'applicazione che si basa sull'integrazione di metodi automatici di allineamento (sovrapposizione) molecolari. Attualmente sono implementati due differenti metodi :

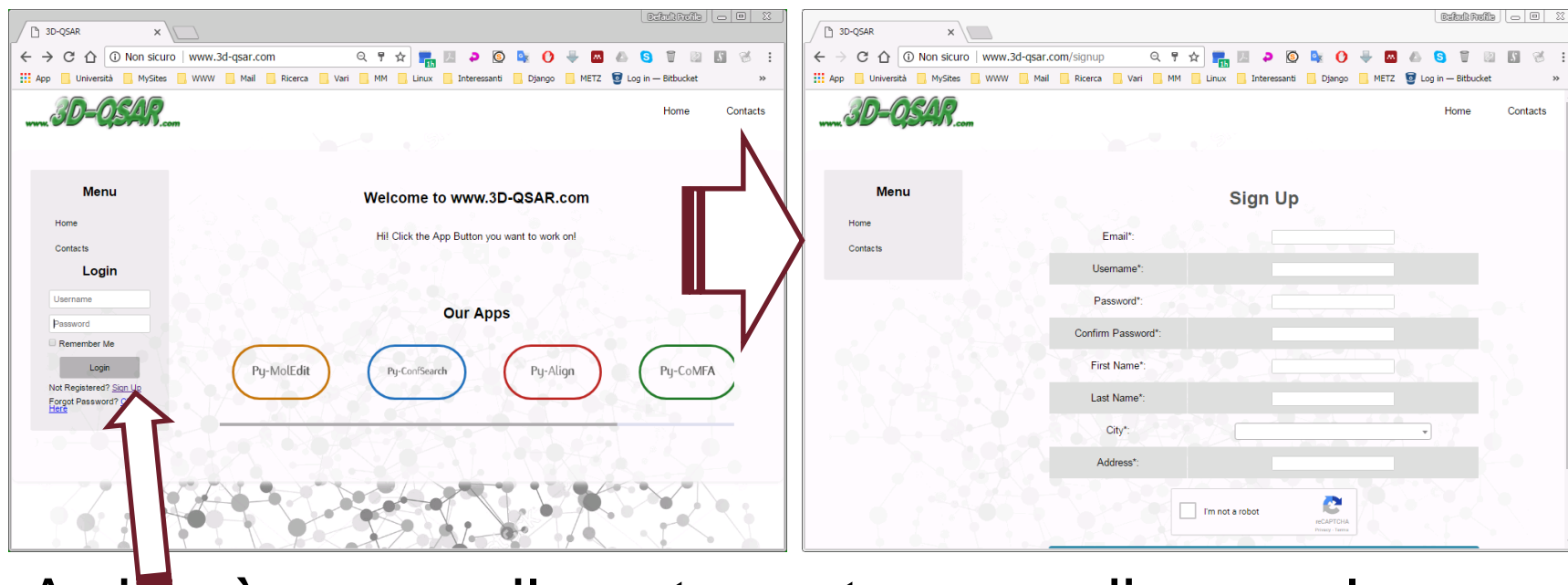
1. Shaep
2. RdKit

## Le applicazioni web del portale: Py-CoMFA

Py-CoMFA è l'applicazione che permette la costruzione di modelli 3-D QSAR. Si basa sull'implementazione del metodo CoMFA di TRIPOS mediante il linguaggio di programmazione python.

# Registrazione al portale

Per la registrazione fornire i dati richiesti



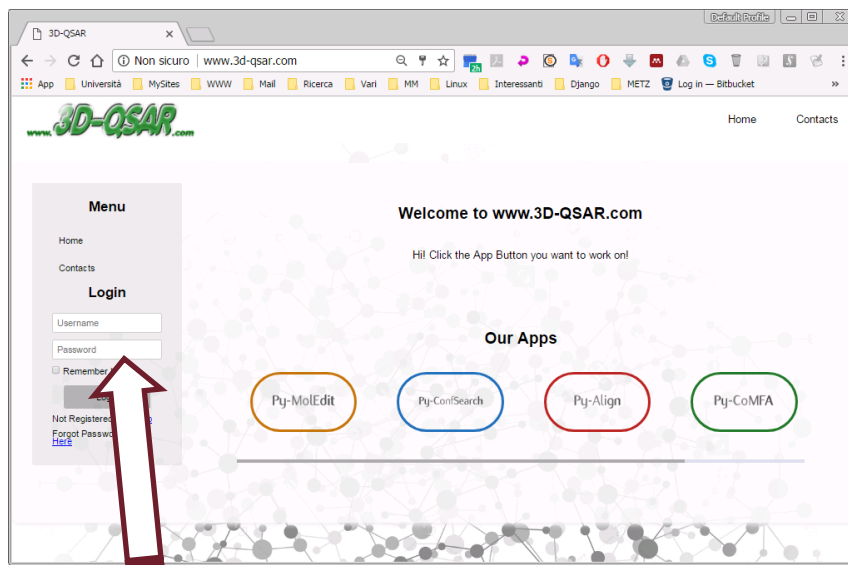
Arriverà un email contenente un codice per la conferma della registrazione



## Accesso al portale

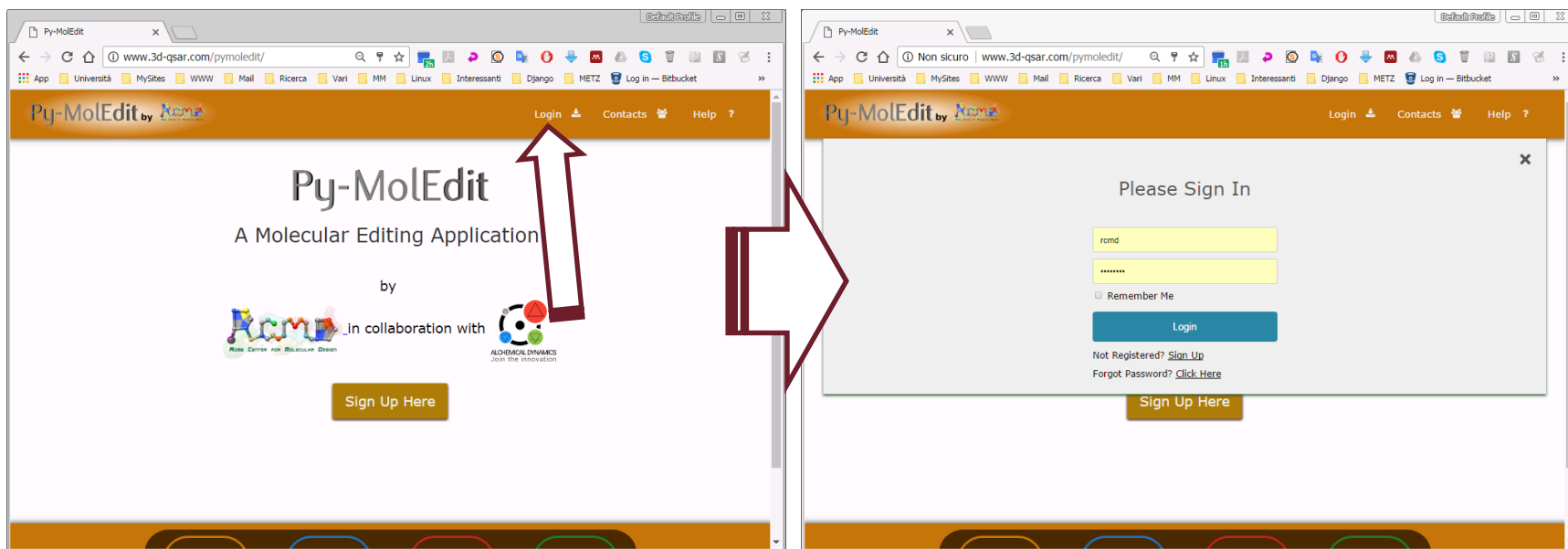
All'indirizzo internet  
<http://www.3d-qsar.com>

La pagina presenta la possibilità di accedere mediante la digitazione un nome utente (username) e di una parola di accesso (password)



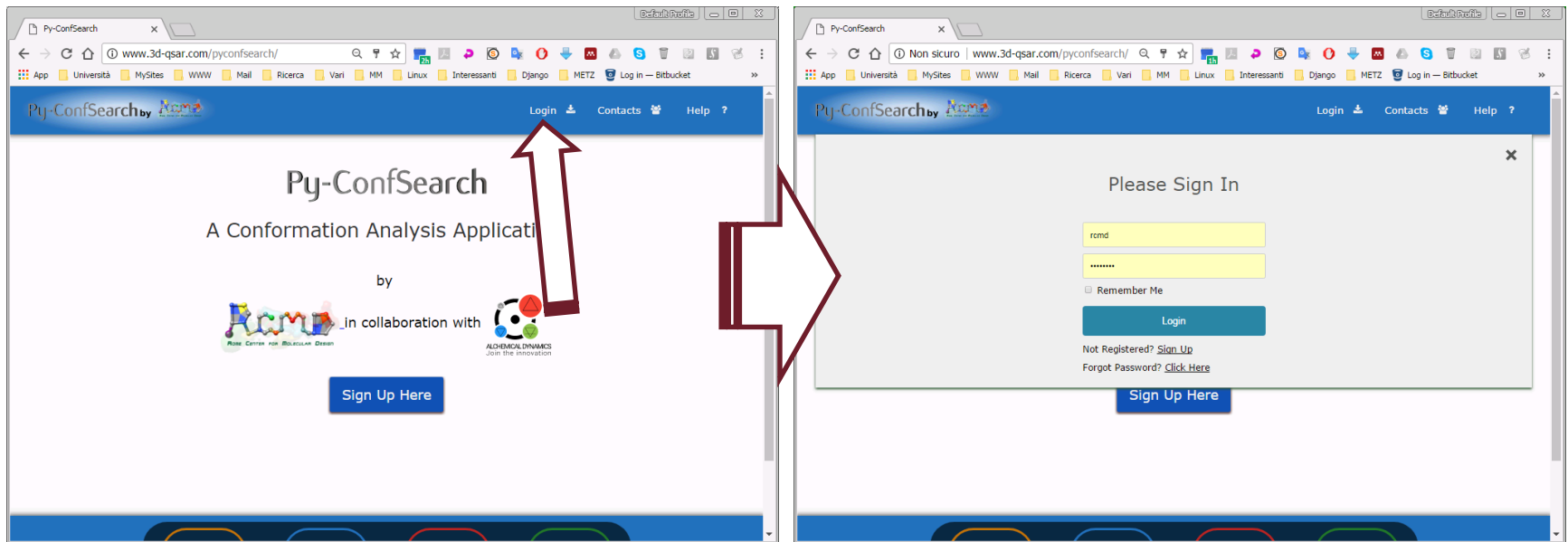
# Registrazione al portale

L'accesso può avvenire anche da ogni applicazione



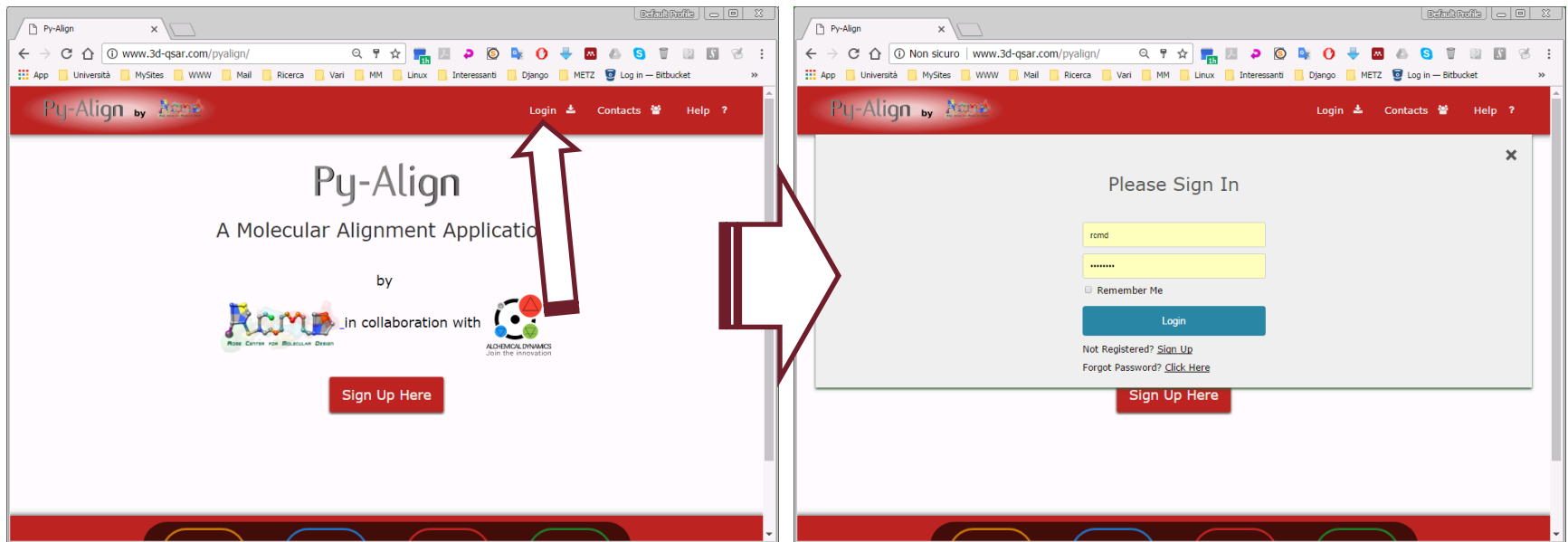
# Registrazione al portale

L'accesso può avvenire anche da ogni applicazione



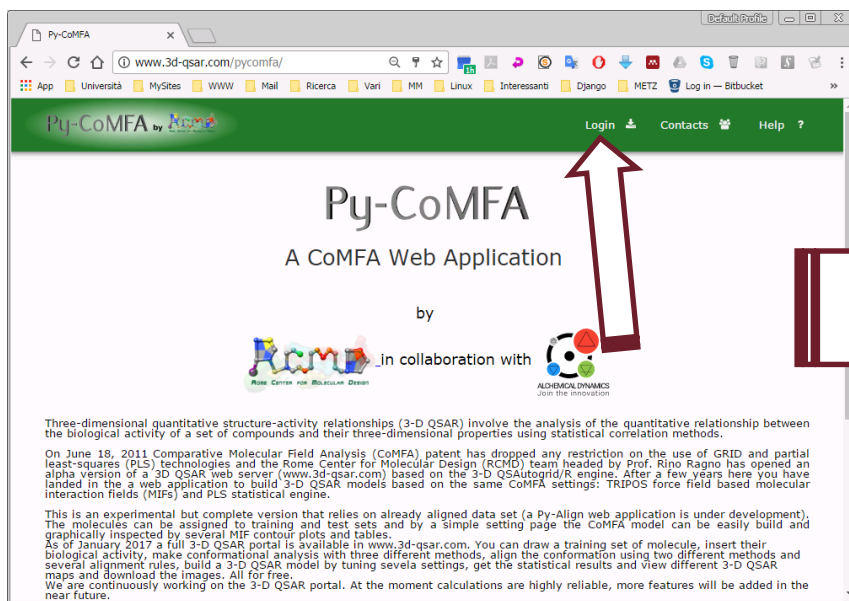
# Accesso al portale

L'accesso può avvenire anche da ogni applicazione





# Accesso al portale

L'accesso può avvenire anche da ogni applicazione



The screenshot shows the homepage of the Py-CoMFA web application. The browser address bar displays "www.3d-qsar.com/pycomfai/". The page features a green navigation bar with "Login", "Contacts", and "Help" links. The main content area includes the title "Py-CoMFA A CoMFA Web Application" and logos for the Rome Center for Molecular Design and Alchemical Dynamics. A red arrow points from the "Login" link in the navigation bar to the "Please Sign In" form in the adjacent screenshot.

Py-CoMFA by  in collaboration with 

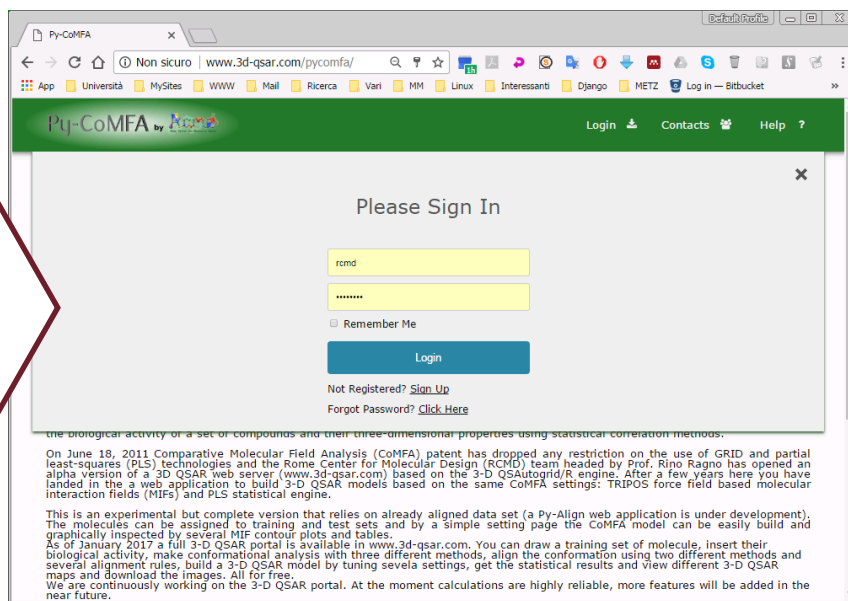
Three-dimensional quantitative structure-activity relationships (3-D QSAR) involve the analysis of the quantitative relationship between the biological activity of a set of compounds and their three-dimensional properties using statistical correlation methods.

On June 18, 2011 Comparative Molecular Field Analysis (CoMFA) patent has dropped any restriction on the use of GRID and partial least-squares (PLS) technologies and the Rome Center for Molecular Design (RCMD) team headed by Prof. Rino Rago has opened an alpha version of a 3D QSAR web server ([www.3d-qsar.com](http://www.3d-qsar.com)) based on the 3-D QSAutoGrid/R engine. After a few years here you have landed in the a web application to build 3-D QSAR models based on the same CoMFA settings: TRIPOS force field based molecular interaction fields (MIFs) and PLS statistical engine.

This is an experimental but complete version that relies on already aligned data set (a Py-Align web application is under development). The molecules can be assigned to training and test sets and by a simple setting page the CoMFA model can be easily build and graphically inspected by several MIF contour plots and tables.

As of January 2017 a full 3-D QSAR portal is available in [www.3d-qsar.com](http://www.3d-qsar.com). You can draw a training set of molecule, insert their biological activity, make conformational analysis with three different methods, align the conformation using two different methods and several alignment rules, build a 3-D QSAR model by tuning several settings, get the statistical results and view different 3-D QSAR maps and download the images. All for free.

We are continuously working on the 3-D QSAR portal. At the moment calculations are highly reliable, more features will be added in the near future.



The screenshot shows the "Please Sign In" form on the Py-CoMFA website. The browser address bar displays "Non sicuro www.3d-qsar.com/pycomfai/". The form includes input fields for "rcmd" and a password, a "Remember Me" checkbox, and a "Login" button. Below the form are links for "Not Registered? Sign Up" and "Forgot Password? Click Here". A red arrow from the previous screenshot points to the "Login" button.

Please Sign In

Remember Me

Not Registered? [Sign Up](#)

Forgot Password? [Click Here](#)

the biological activity of a set of compounds and their three-dimensional properties using statistical correlation methods.

On June 18, 2011 Comparative Molecular Field Analysis (CoMFA) patent has dropped any restriction on the use of GRID and partial least-squares (PLS) technologies and the Rome Center for Molecular Design (RCMD) team headed by Prof. Rino Rago has opened an alpha version of a 3D QSAR web server ([www.3d-qsar.com](http://www.3d-qsar.com)) based on the 3-D QSAutoGrid/R engine. After a few years here you have landed in the a web application to build 3-D QSAR models based on the same CoMFA settings: TRIPOS force field based molecular interaction fields (MIFs) and PLS statistical engine.

This is an experimental but complete version that relies on already aligned data set (a Py-Align web application is under development). The molecules can be assigned to training and test sets and by a simple setting page the CoMFA model can be easily build and graphically inspected by several MIF contour plots and tables.

As of January 2017 a full 3-D QSAR portal is available in [www.3d-qsar.com](http://www.3d-qsar.com). You can draw a training set of molecule, insert their biological activity, make conformational analysis with three different methods, align the conformation using two different methods and several alignment rules, build a 3-D QSAR model by tuning several settings, get the statistical results and view different 3-D QSAR maps and download the images. All for free.

We are continuously working on the 3-D QSAR portal. At the moment calculations are highly reliable, more features will be added in the near future.

## Obiettivo del portale

Il portale è un sistema integrato con lo scopo ultimo di costruire dei modelli 3-D QSAR utilizzando il paradigma di CoMFA di TRIPOS. Tuttavia ogni applicazione può essere usata in modo indipendente.

**Py-MolEdit** permette di disegnare delle molecole e ottenerne la loro struttura tridimensionale; con **Py-ConfSearch** mediante differenti metodi si possono eseguire delle analisi conformazionali di una serie di molecole che possono essere scaricate sul proprio dispositivo; **Py-Align** mediante l'implementazione di algoritmi di allineamento automatico fornisce un metodo per ottenere la sovrapposizione di una serie di molecole singole o di una serie di famiglie di conformazioni

## Obiettivo del portale

**Py-CoMFA** infine è rappresentata l'applicazione clou del portale. Mediante la quale si possono costruire ed analizzare modelli 3-D QSAR. I modelli possono essere costruiti utilizzando direttamente dei data set preallineati, oppure mediante la definizione di regole di allineamento con l'utilizzo delle applicazioni **Py-ConfSearch** e **Py-Align**.

## Obiettivi del Tutorial

Questo tutorial si prefigge di dare delle linee guida per l'utilizzo del portale al fine di costruire modelli 3-D QSAR.

Nella prima parte saranno date le istruzioni per la costruzione di un modello 3-D QSAR utilizzando le 21 molecole steroidiche utilizzate nella pubblicazione originale di CoMFA da parte di Cramer e collaboratori, facilmente reperibile dal sito acs

(<https://pubs.acs.org/doi/abs/10.1021/ja00226a005>)

*J. Am. Chem. Soc.* **1988**, *110*, 5959–5967

5959

Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins

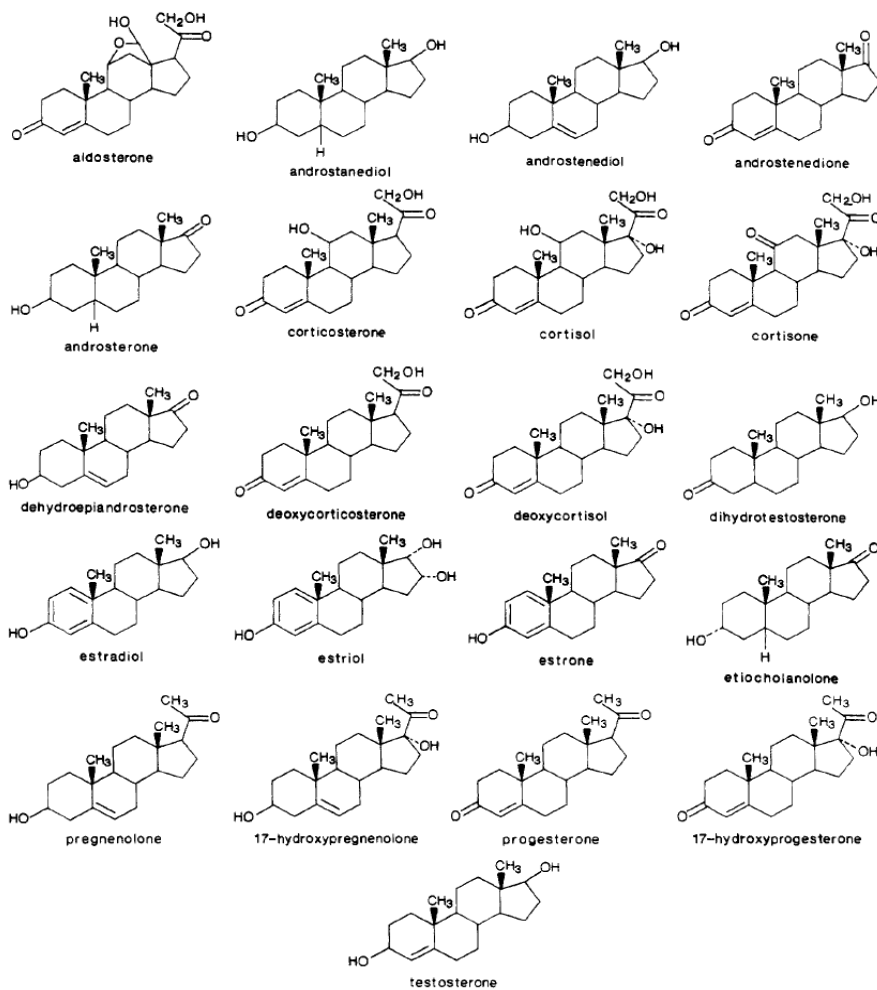
Richard D. Cramer, III,\* David E. Patterson, and Jeffrey D. Bunce

*Contribution from Tripos Associates, 1699 South Hanley Road, St. Louis, Missouri 63144. Received January 5, 1988*

**Abstract:** Comparative molecular field analysis (CoMFA) is a promising new approach to structure/activity correlation. Its characteristic features are (1) representation of ligand molecules by their steric and electrostatic fields, sampled at the intersections of a three-dimensional lattice, (2) a new "field fit" technique, allowing optimal mutual alignment within a series, by minimizing the RMS field differences between molecules, (3) data analysis by partial least squares (PLS), using cross-validation to maximize the likelihood that the results have predictive validity, and (4) graphic representation of results, as contoured three-dimensional coefficient plots. CoMFA is exemplified by analyses of the affinities of 21 varied steroids to corticosteroid- and testosterone-binding globulins. Also described are the sensitivities of results to the nature of the field and the definition of the lattice and, for comparison, analyses of the same data using various combinations of other parameters. From these results, a set of ten steroid-binding affinity values unknown to us during the CoMFA analysis were well predicted.



# Costruzione del Modello 3-D QSAR: le molecole



2. The 21 steroid structures used to derive the CoMFA QSARs.

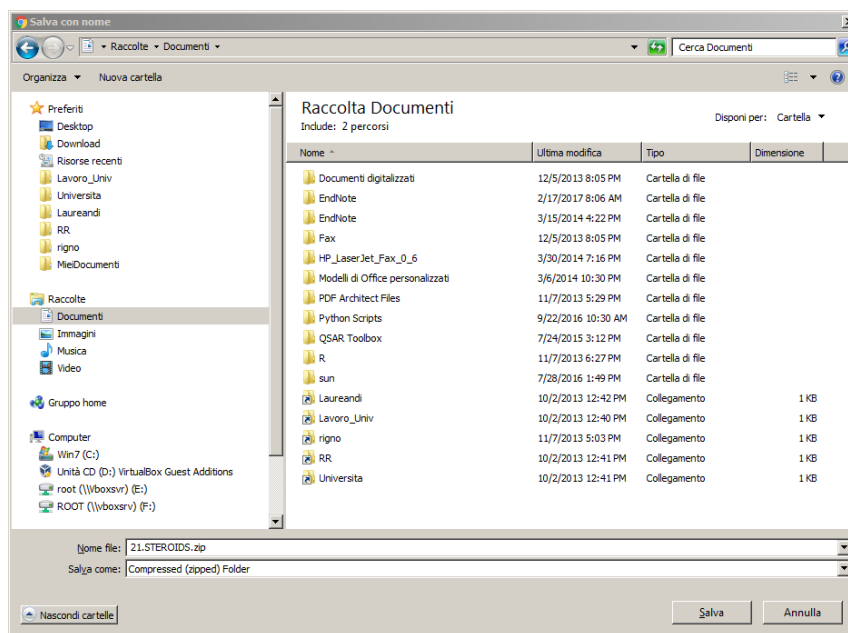
## Costruzione del Modello 3-D QSAR: il training set

In questa parte del tutorial utilizzeremo le conformazioni originali che sono facilmente reperibili in rete e che sono anche condivise nel sistema elearning2.uniroma1.it. Il seguente link permetterà di scaricare il data set originale delle 21 molecole steroidiche del modello di Cramer del 1988

([https://elearning2.uniroma1.it/pluginfile.php/501966/mod\\_older/content/0/Esercitazioni/3-D\\_QSAR/21.STEROIDS.zip?forcedownload=1](https://elearning2.uniroma1.it/pluginfile.php/501966/mod_older/content/0/Esercitazioni/3-D_QSAR/21.STEROIDS.zip?forcedownload=1))

# Costruzione del Modello 3-D QSAR: il training set

Copiando ed incollando il link in un browser (google-chrome o altro) il sistema vi chiede dove salvare il file «21.STEROIDS.zip», un file archivio compresso contenente le 21 molecole (21 file a formato sdf) e l'attività biologica associata.



## Costruzione del Modello 3-D QSAR: il training set

Prima di caricare le molecole nel portale è necessario creare un data set, un contenitore virtuale in cui sono contenute le molecole su cui vogliamo costruire il modello 3-D QSAR. Il data set al suo interno può contenere tre categorie di molecole:

1. Molecole assegnate al training set
2. Molecole assegnate al test set
3. Molecole non assegnate a nessuno dei due casi di cui sopra.

## Costruzione del Modello 3-D QSAR: il training set

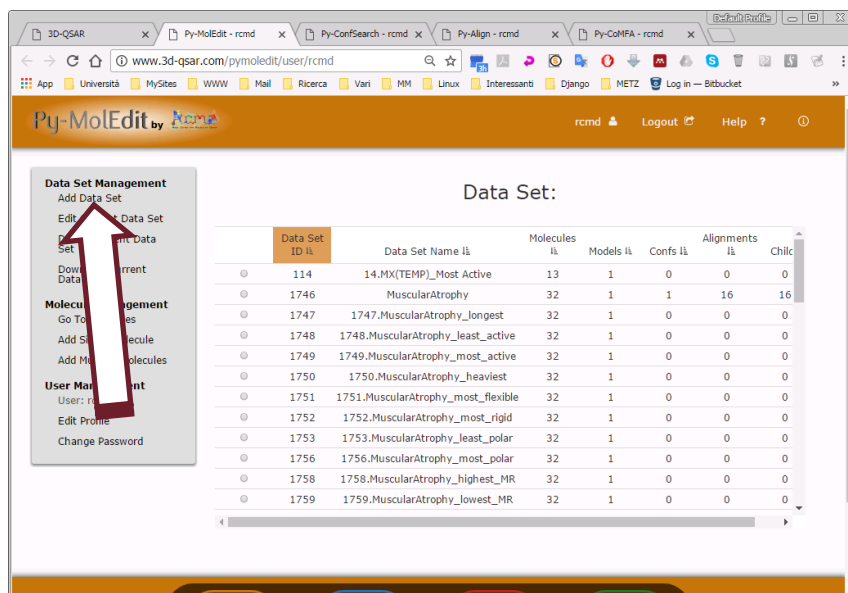
Per questo tutorial le 21 molecole steroidiche saranno assegnate tutte al training set. Sarà oggetto di un successivo tutorial lo studio di dataset contenenti anche un test set.

## Costruzione del Modello 3-D QSAR: il training set

Per questo tutorial le 21 molecole steroidiche saranno assegnate tutte al training set. Sarà oggetto di un successivo tutorial lo studio di data set contenenti anche un test set.

# Costruzione del Modello 3-D QSAR: il training set

Una volta che è stata effettuata l'autenticazione sul portale è possibile creare un data set, attraverso una qualsiasi delle quattro applicazioni mediante il link 'Add Data Set' sul menù di sinistra.



The screenshot displays the Py-MolEdit web application interface. On the left, a sidebar menu titled 'Data Set Management' contains the following items: 'Add Data Set', 'Edit Data Set', 'Delete Data Set', 'Download Data Set', 'Molecule Management', 'Add Structure', 'Add Molecules', 'User Management', 'Edit Profile', and 'Change Password'. A red arrow points to the 'Add Data Set' option. The main content area, titled 'Data Set:', features a table with the following columns: 'Data Set ID', 'Data Set Name', 'Molecules', 'Models', 'Confs', 'Alignments', and 'Chilc'. The table contains 10 rows of data sets, each with a radio button in the first column.

Data Set ID	Data Set Name	Molecules	Models	Confs	Alignments	Chilc
<input type="radio"/> 114	14.MX(TEMP)_Most Active	13	1	0	0	0
<input type="radio"/> 1746	MuscularAtrophy	32	1	1	16	16
<input type="radio"/> 1747	1747.MuscularAtrophy_longest	32	1	0	0	0
<input type="radio"/> 1748	1748.MuscularAtrophy_least_active	32	1	0	0	0
<input type="radio"/> 1749	1749.MuscularAtrophy_most_active	32	1	0	0	0
<input type="radio"/> 1750	1750.MuscularAtrophy_heaviest	32	1	0	0	0
<input type="radio"/> 1751	1751.MuscularAtrophy_most_flexible	32	1	0	0	0
<input type="radio"/> 1752	1752.MuscularAtrophy_most_rigid	32	1	0	0	0
<input type="radio"/> 1753	1753.MuscularAtrophy_least_polar	32	1	0	0	0
<input type="radio"/> 1756	1756.MuscularAtrophy_most_polar	32	1	0	0	0
<input type="radio"/> 1758	1758.MuscularAtrophy_highest_MR	32	1	0	0	0
<input type="radio"/> 1759	1759.MuscularAtrophy_lowest_MR	32	1	0	0	0

# Costruzione del Modello 3-D QSAR: il training set

Una nuova schermata richiederà l'inserimento di alcune informazioni:

1. Nome del data set
2. Tipo di attività (pKi, IC50, EC50, ecc)
3. Tipo di saggio (cellulare, enzimatico, ecc)
4. Una descrizione delle caratteristiche del data set
5. Tipo di target biologico (enzima, cellula, ecc)

Finito l'inserimento 'cliccare' su 'done'

The screenshot shows a web browser window with the URL [www.3d-qsar.com/pymoledit/user/rcmd/data\\_set/add](http://www.3d-qsar.com/pymoledit/user/rcmd/data_set/add). The page title is "Py-MolEdit by Kame". The main content area is titled "New Data Set:" and contains the following form elements:

- Name\*:** Input field containing "Nuovo\_dataset".
- Activity Standard Type\*:** Dropdown menu set to "IC50".
- Activity Assay Type\*:** Dropdown menu set to "Enzymatic".
- Activity assay description:** Text area containing the text: "Descrizione delle molecole, eventualmente il link della pubblicazione del dataset o altre informazioni".
- Biological Target\*:** Dropdown menu set to "Generic on enzyme".

At the bottom of the form are three buttons: "Cancel", "Done", and "Next". A large red arrow points to the "Done" button.



# Costruzione del Modello 3-D QSAR: il training set

Una nuova schermata apparirà con il nuovo data set creato e vuoto!

L'inserimento delle molecole nel data set può essere effettuato sia in modo discontinuo, una alla volta, oppure in 'batch' con i comandi 'Add Single Molecule' e 'Add Multiple Molecule'.

Per praticità e dal momento che tutte le molecole sono pronte si consiglia di usare la seconda opzione.

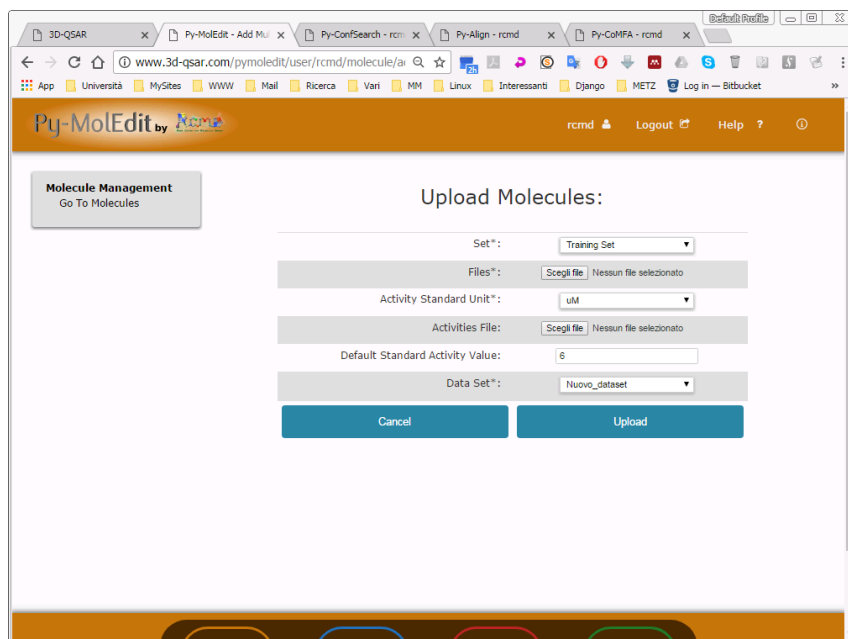
Pertanto si può procedere 'Cliccando' su 'Add Multiple Molecule'

The screenshot displays the Py-MolEdit web application interface. The browser address bar shows the URL [www.3d-qsar.com/pymoledit/user/rcmd/molecule?id](http://www.3d-qsar.com/pymoledit/user/rcmd/molecule?id). The page title is 'Py-MolEdit by Kemo'. The user is logged in as 'rcmd'. The main content area is titled 'Molecules of Data Set: Nuovo\_dataset'. A sidebar on the left contains two sections: 'Data Set Management' with options like 'Add Data Set', 'Edit Current Data Set', 'Delete Current Data Set', and 'Download Current Data Set'; and 'Molecule Management' with options like 'Add Single Molecule', 'Add Multiple Molecules', 'Draw Molecule', 'Edit Selected Molecule', 'View Selected Molecules', and 'View Data Set Molecules'. A red arrow points to the 'Add Multiple Molecules' option. The main area features a table with the following columns: 'Select All', 'ID Molecule', 'Label', 'Given Activity Standard Value', 'Activity Standard Value', 'DataSet', and 'Set'. The table is currently empty. At the bottom, there is a pagination control showing 'PAGE 1 OF 1' and 'MOLECULES FOR PAGES: 10'.

# Costruzione del Modello 3-D QSAR: il training set

Una nuova schermata apparirà in cui è richiesto l'intervento dell'utente:

1. Selezionare come le molecole saranno classificate (training set, test set o inattive – None)
2. Selezionare i file contenenti le strutture delle molecole. 'Cliccando' sul bottone 'Scegli file' una finestra del sistema operativo permetterà di selezionare i file delle molecole (mantenere il tasto CTRL premuto per selezionare più file)
3. Indicare l'unità di misura dell'attività biologica (p per l'attività biologica trasformata in  $-\log([M])$ , uM, nM, pM, ecc.)
4. Selezionare il file contenente la lista delle molecole con l'attività biologica associata.



## Costruzione del Modello 3-D QSAR: il training set

Per quanto riguarda il file delle attività biologiche deve essere preparato un file di testo semplice in cui devono essere presenti solo due colonne. Nella prima colonna sono indicati i nome dei file delle molecole senza riportare l'estensione del file, mentre nella seconda colonna l'attività biologica usando come separatore decimale il punto.

E' necessario fare attenzione alla corrispondenza tra nome del file e nome della molecola indicata nel file dell'attività.

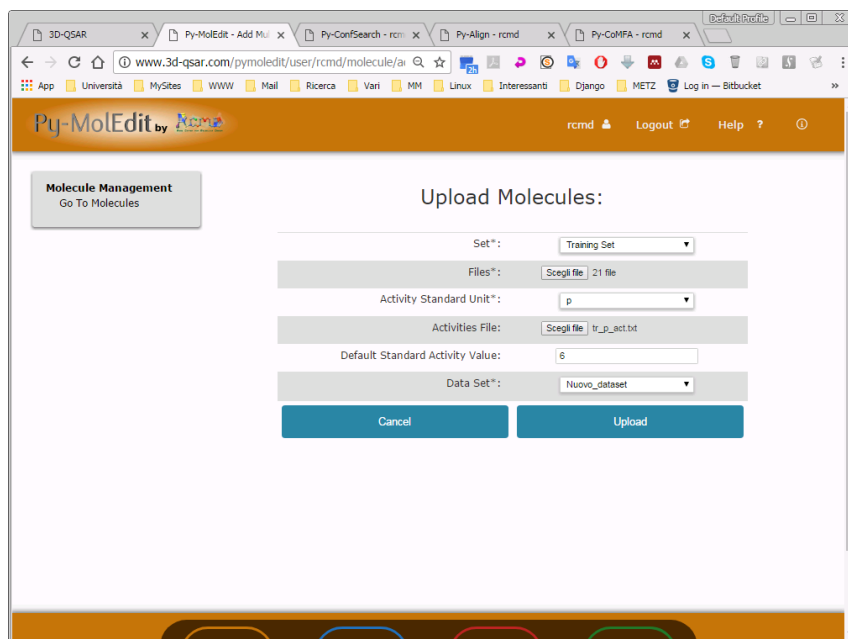
Il file fornito e' un buon esempio di come si deve formattare il file delle attività (tr\_p\_act.txt).

# Costruzione del Modello 3-D QSAR: il training set

Per l'esecuzione del tutorial la finestra del 'browser' dovrebbe apparire come la figura accanto in cui 21 file sono stati selezionati, l'unità di misura è p, il file dell'attività è 'tr\_p\_act.txt' e tutte le molecole sono assegnate al training set.

Nel caso in cui per una molecola mancasse il valore di attività il sistema automaticamente assegnerà il valore  $-\log([M]) = 6$ , cioè 1 micromolare.

Finito l'inserimento si deve cliccare sul bottone 'upload'

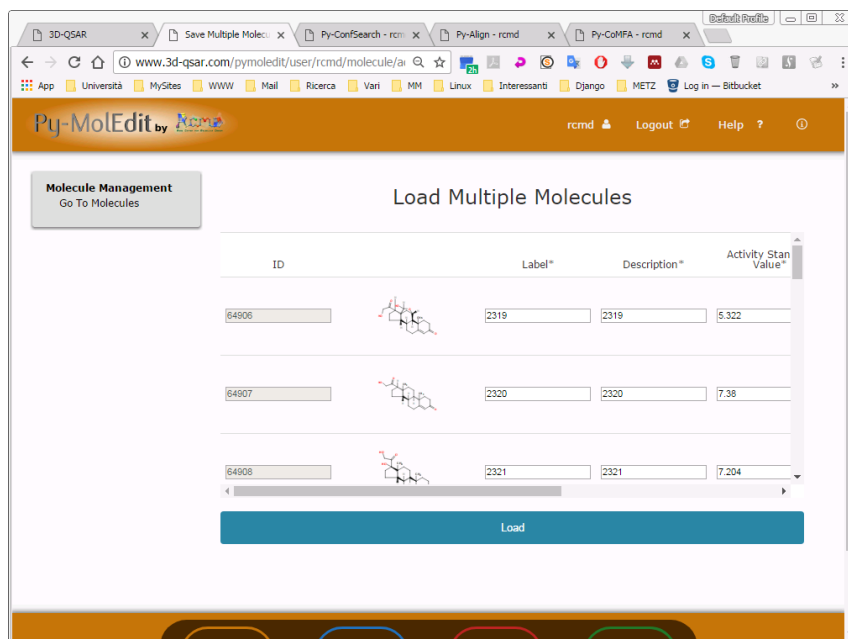


# Costruzione del Modello 3-D QSAR: il training set

Una schermata fornirà quindi la visualizzazione delle molecole che saranno assegnate al data set 'Nuovo\_dataset'.

In questo momento vanno fatti i dovuti controlli, in particolare sui valori delle attività biologiche che della corrispondenza dei nome e del data set.

A fine controllo si può 'cliccare' sul bottone 'Load' che instruisce il sistema a caricare le molecole nel database del portale.



The screenshot shows the 'Load Multiple Molecules' interface in the Py-MolEdit web application. The page features a table with the following data:

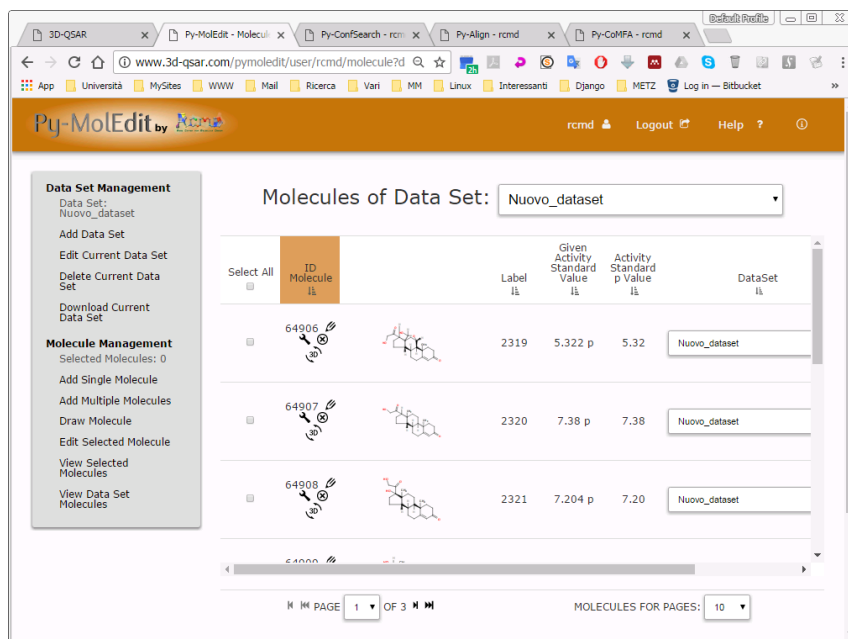
ID	Label*	Description*	Activity Stan Value*
64906	2319	2319	5.322
64907	2320	2320	7.38
64908	2321	2321	7.204

Each row includes a chemical structure icon next to the ID. A 'Load' button is positioned below the table. The browser's address bar shows the URL: [www.3d-qsar.com/pymoledit/user/rcmd/molecole/](http://www.3d-qsar.com/pymoledit/user/rcmd/molecole/).

# Costruzione del Modello 3-D QSAR: il training set

Quindi sarà visualizzata la schermata del data set con le molecole caricate e tutte classificate come training set.

Le molecole caricate possono essere anche visualizzate in 3-D una alla volta 'cliccando' sullo strumento 'chiave' oppure allineate selezionando la voce 'View Data Set Molecules'



The screenshot shows the Py-MolEdit web application interface. The browser address bar indicates the URL [www.3d-qsar.com/pymoledit/user/rcmd/molecule?id](http://www.3d-qsar.com/pymoledit/user/rcmd/molecule?id). The page title is "Py-MolEdit by Kemo". The interface is divided into a sidebar and a main content area.




**Data Set Management**

- Data Set: Nuovo\_dataset
- Add Data Set
- Edit Current Data Set
- Delete Current Data Set
- Download Current Data Set

**Molecule Management**

- Selected Molecules: 0
- Add Single Molecule
- Add Multiple Molecules
- Draw Molecule
- Edit Selected Molecule
- View Selected Molecules
- View Data Set Molecules

**Molecules of Data Set:** Nuovo\_dataset

Select All	ID Molecule	Label	Given Activity Standard Value	Activity Standard p Value	DataSet	
<input type="checkbox"/>	64906		2319	5.322 p	5.32	Nuovo_dataset
<input type="checkbox"/>	64907		2320	7.38 p	7.38	Nuovo_dataset
<input type="checkbox"/>	64908		2321	7.204 p	7.20	Nuovo_dataset

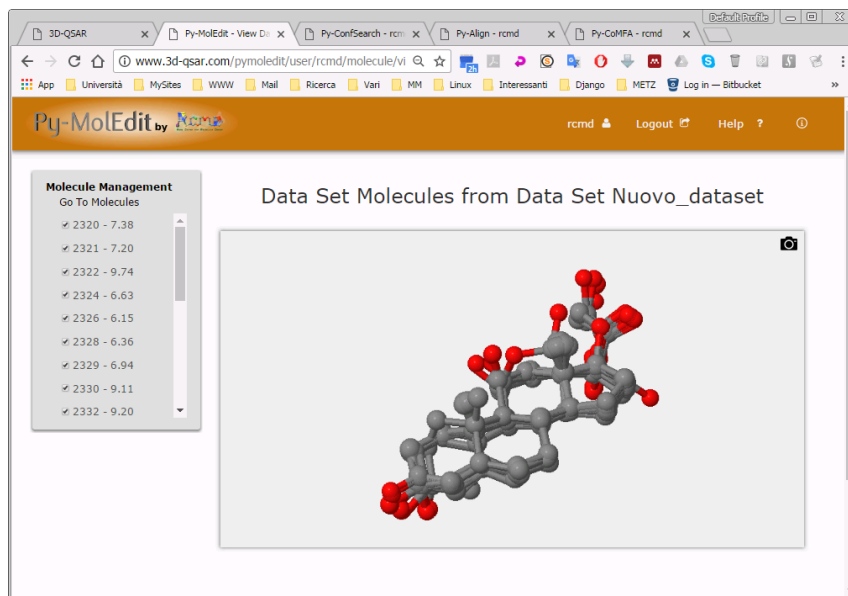
PAGE 1 OF 3

MOLECULES FOR PAGES: 10

# Costruzione del Modello 3-D QSAR: il training set

Il training set si puo' quindi visualizzare nel suo allineamento originale.

Togliendo la spunta accanto ai nomi delle molecole si possono anche analizzare solo alcune delle molecole



## Costruzione del Modello 3-D QSAR: i calcoli

Per procedere alla creazione del modello 3-D QSAR è necessario spostarsi nell'applicazione Py-CoMFA cliccando su bottone cerchiato in verde in fondo alla pagina e riselezionando il data set appena creato e popolato di molecole

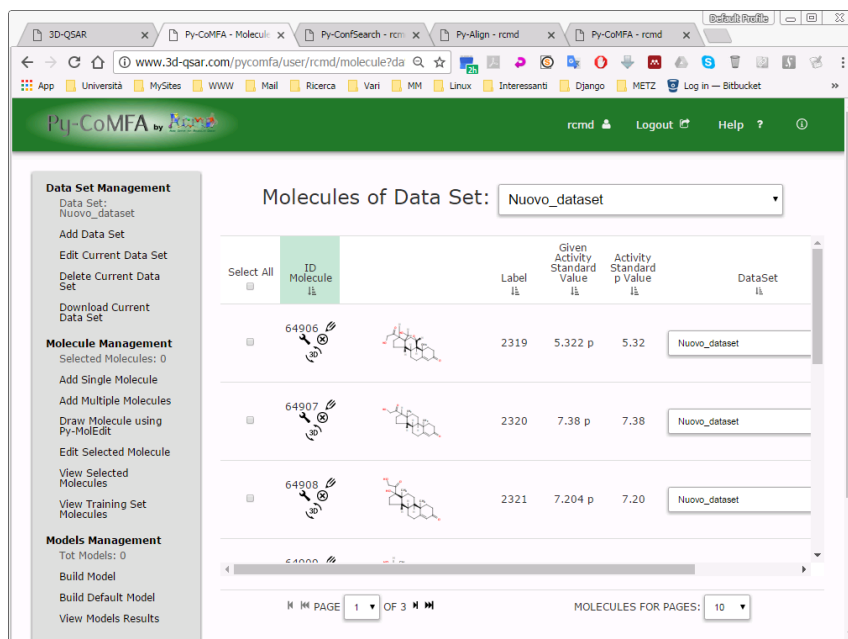


# Costruzione del Modello 3-D QSAR: i calcoli




La creazione del modello 3-D QSAR può essere effettuata direttamente 'cliccando' sulla voce 'Build Default Model' che 'lancerà' direttamente i calcoli con dei parametri detti di default.

Altrimenti per un controllo più fine si può utilizzare la voce 'Build Model' che mostrerà una nuova schermata.

Per il tutorial in corso usare la voce 'Build Default Model'



The screenshot displays the Py-CoMFA web interface. The browser address bar shows the URL [www.3d-qsar.com/pycomfa/user/rcmd/molecule7da](http://www.3d-qsar.com/pycomfa/user/rcmd/molecule7da). The page title is "3D-QSAR". The main content area is titled "Molecules of Data Set: Nuovo\_dataset". It features a table with the following columns: "Select All", "ID Molecule", "Label", "Given Activity Standard Value", "Activity Standard p Value", and "DataSet". The table contains three rows of data:

Select All	ID Molecule	Label	Given Activity Standard Value	Activity Standard p Value	DataSet	
<input type="checkbox"/>	64906		2319	5.322 p	5.32	Nuovo_dataset
<input type="checkbox"/>	64907		2320	7.38 p	7.38	Nuovo_dataset
<input type="checkbox"/>	64908		2321	7.204 p	7.20	Nuovo_dataset

At the bottom of the table, there is a pagination control showing "PAGE 1 OF 3" and a dropdown menu for "MOLECULES FOR PAGES: 10".

# Costruzione del Modello 3-D QSAR: i calcoli

La nuova schermata presenterà una tabella in cui nella riga corrispondente al modello in corso di definizione si noterà che lo 'Status' è in 'running' indicando che il calcolo è in corso.

Scorrendo verso destra la tabella si possono notare le impostazioni del modello.

Mentre i calcoli sono è possibile anche mandare altri modelli utilizzando la voce 'Build New Model' che darà accesso ad una tabella di impostazione

The image displays two screenshots of the Py-CoMFA web interface. The top screenshot shows the 'Molecule Management' sidebar with options like 'Go To Molecules', 'Build Default Model', and 'Build New Model'. The main content area displays 'Models of Data Set Nuovo\_dataset for Target Generic on enzyme' with a 'Data Set' dropdown set to 'Nuovo\_dataset'. A table below shows model details:

ID	Status	Maps	nTrMols	External Pred	q <sup>2</sup> ELE (PC)	q <sup>2</sup> STE (PC)	q <sup>2</sup> BOTH (PC)	Prob
2808	running	Missing	21	False				C.

The bottom screenshot shows the same interface but with a different table of model parameters:

Probe Charge	Min Sigma	Scaling	Max N Components	Grid Spacing	CutOff	CV Type	CV Groups	CV Iterations	Ra
1	0.05	True	8	1.0	30.0	LOO	5	10	Tru

# Costruzione del Modello 3-D QSAR: i calcoli

La tabella delle impostazioni appare nella nuova schermata dove è possibile personalizzare tutte le impostazioni.

‘Cliccando’ sull'icona delle informazioni (i cerchiata) si può accedere ad una descrizione delle diverse voci.

Qui si possono modificare le impostazioni

The screenshot displays the Py-CoMFA web interface. The browser address bar shows the URL [www.3d-qsar.com/pycomfa/user/rcmd/model/build/](http://www.3d-qsar.com/pycomfa/user/rcmd/model/build/). The page title is "Build Model for Nuovo\_dataset:". On the left, there is a sidebar with "Molecule Management" (Go To Molecules) and "User Management" (User: rcmd, Edit Profile). The main area contains a form with various parameters for building a 3-D QSAR model. The parameters are as follows:

Parameter	Value
Probe Atom	C.3
Probe Charge	+1
Dielectric Constant	8
Molecular Interaction Field (BOTH = STE + ELE)	BOTH
Maximum Number of Principal Components	8
Grid Spacing	1
Grid Extension	5
Max/Min Energy of Cutoff Value	30
Minimum Sigma	0.05
CV Type	LOO
CV Groups (Only for LSO)	5
CV Iterations (Only for LSO)	10
CV Random	True
Minimum SDEP Increment	0
Make Y-Scrambling	False
Y-Scrambling Iterations	10
Apply Model to Test Set Molecules	False

At the top and bottom of the form area, there are "Cancel" and "Build 3-D QSAR Model" buttons.

# Costruzione del Modello 3-D QSAR: i calcoli

Per dimostrazione sono state cambiate le voci:

Grid Spacing → da 1 a 2

Grid Extension → da 5 a 10

Minimum Sigma → da 0.05 a 2

E quindi si 'lancia' il calcolo

The screenshot shows the Py-CoMFA web interface. The browser address bar displays [www.3d-qsar.com/pycomfa/user/rcmd/model/build/](http://www.3d-qsar.com/pycomfa/user/rcmd/model/build/). The page title is "Build Model for Nuovo\_dataset:". On the left, there is a sidebar with "Molecule Management" (Go To Molecules) and "User Management" (User: rcmd, Edit Profile). The main area contains a form with various parameters for building a 3-D QSAR model. The parameters and their values are:

Parameter	Value
Probe Atom	C.3
Probe Charge	+1
Dielectric Constant	8
Molecular Interaction Field (BOTH = STE + ELE)	BOTH
Maximum Number of Principal Components	8
Grid Spacing	2
Grid Extension	10
Max/Min Energy of Cutoff Value	30
Minimum Sigma	2
CV Type	LOO
CV Groups (Only for LSO)	5
CV Iterations (Only for LSO)	10
CV Random	True
Minimum SDEP Increment	0
Make Y-Scrambling	False
Y-Scrambling Iterations	10
Apply Model to Test Set Molecules	False

At the bottom of the form, there are two buttons: "Cancel" and "Build 3-D QSAR Model".

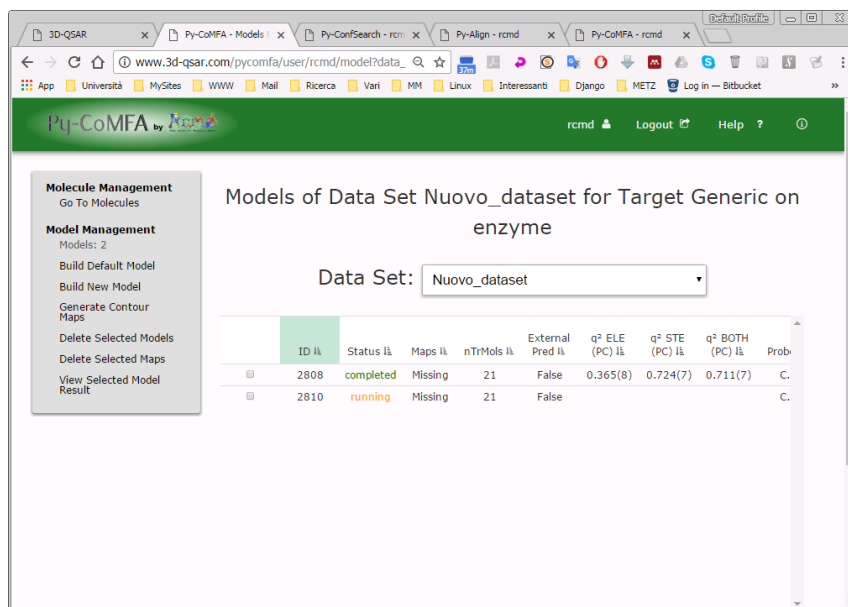
# Costruzione del Modello 3-D QSAR: i calcoli

Che sarà elencato nella tabella dei risultati.

Da notare che nel frattempo il primo calcolo potrebbe essere terminato come riportato in figura.

Lo 'Status' del modello 2808 indica che è 'Completed'.

**Il numero del vostro modello sarà  
differente!**



The screenshot shows the Py-CoMFA web interface. The main content area displays 'Models of Data Set Nuovo\_dataset for Target Generic on enzyme'. A dropdown menu for 'Data Set:' is set to 'Nuovo\_dataset'. Below this is a table with the following data:

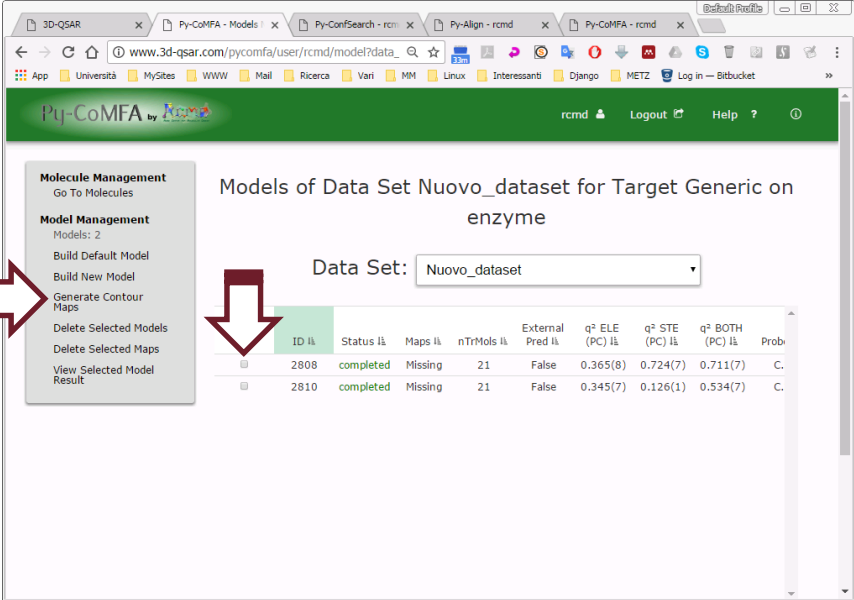
ID	Status	Maps	nTrMols	External Pred	q <sup>2</sup> ELE (PC)	q <sup>2</sup> STE (PC)	q <sup>2</sup> BOTH (PC)	Prob.
2808	completed	Missing	21	False	0.365(8)	0.724(7)	0.711(7)	C.
2810	running	Missing	21	False				C.

# Costruzione del Modello 3-D QSAR: i calcoli

Lo 'Status' dei modelli indica che è 'Completed', tuttavia le mappe sono 'Missing'.

Si può notare inoltre che il modello 2808 è numericamente migliore ( $q^2$  BOTH = 0.711) del 2010, pertanto si può procedere a calcolare le mappe grafiche solo per il modello 2808 di qualità superiore

Per calcolare le mappe si deve selezionare il modello 2008 e quindi cliccare sulla voce 'Generate Contour Maps'



The screenshot shows the Py-CoMFA web interface. On the left, a sidebar menu under 'Model Management' includes 'Generate Contour Maps'. A red arrow points to this option. The main content area displays 'Models of Data Set Nuovo\_dataset for Target Generic on enzyme'. A dropdown menu for 'Data Set:' is set to 'Nuovo\_dataset'. Below this is a table with the following data:

ID	Status	Maps	nTrMols	External Pred	q <sup>2</sup> ELE (PC)	q <sup>2</sup> STE (PC)	q <sup>2</sup> BOTH (PC)	Prob.
2808	completed	Missing	21	False	0.365(8)	0.724(7)	0.711(7)	C.
2810	completed	Missing	21	False	0.345(7)	0.126(1)	0.534(7)	C.

Red arrows in the image point to the 'Generate Contour Maps' menu item and the row for model 2808 in the table.

# Costruzione del Modello 3-D QSAR: l'analisi

In pochi minuti il sistema calcolerà e quindi salverà i dati da visualizzare graficamente.

Quindi per analizzare il modello 2808 dopo averlo selezionato dovrà essere attivata la voce 'View Selected Model Result'

The image displays two screenshots of the Py-CoMFA web interface. The top screenshot shows the 'Models of Data Set Nuovo\_dataset for Target Generic on enzyme' page. The 'Data Set' dropdown is set to 'Nuovo\_dataset'. A table lists two models:

ID	Status	Maps	nTrMols	External Pred	q <sup>2</sup> ELE (PC)	q <sup>2</sup> STE (PC)	q <sup>2</sup> BOTH (PC)	Prob
2808	saving maps	Missing	21	False	0.365(8)	0.724(7)	0.711(7)	C.
2810	completed	Missing	21	False	0.345(7)	0.126(1)	0.534(7)	C.

The bottom screenshot shows the same page, but with model 2808 selected. A red arrow points to the 'View Selected Model Result' option in the 'Model Management' sidebar. The table now shows model 2808 with a status of 'completed' and 'Available' maps.

ID	Status	Maps	nTrMols	External Pred	q <sup>2</sup> ELE (PC)	q <sup>2</sup> STE (PC)	q <sup>2</sup> BOTH (PC)	Prob
2808	completed	Available	21	False	0.365(8)	0.724(7)	0.711(7)	C.
2810	completed	Missing	21	False	0.345(7)	0.126(1)	0.534(7)	C.

# Costruzione del Modello 3-D QSAR: l'analisi

Comparirà quindi una tabella riassuntiva in cui sono listati tutti i valori statistici del modello, clicando sul bottone affianco alla voce 'Model Field'

The screenshots show the Py-CoMFA web interface. The top screenshot displays the 'STE' model field, and the bottom screenshot displays the 'STE\_ELE' model field. Both screenshots show a table of statistical results for Model #2808, including PC, r<sup>2</sup>, SDEC, q<sup>2</sup>, SDEP, r<sup>2</sup> YS, q<sup>2</sup> YS, SDEP YS, and CutO values.

PC	r <sup>2</sup>	SDEC	q <sup>2</sup>	SDEP	r <sup>2</sup> YS	q <sup>2</sup> YS	SDEP YS	CutO
1.000	0.453	0.872	0.179	1.069	0.000	0.000	0.000	30.0
2.000	0.767	0.570	0.095	1.123	0.000	0.000	0.000	30.0
3.000	0.837	0.477	0.322	0.971	0.000	0.000	0.000	30.0
4.000	0.906	0.362	0.185	1.065	0.000	0.000	0.000	30.0
5.000	0.952	0.259	0.212	1.047	0.000	0.000	0.000	30.0
6.000	0.979	0.172	0.222	1.041	0.000	0.000	0.000	30.0
7.000	0.989	0.126	0.310	0.980	0.000	0.000	0.000	30.0
8.000	0.992	0.104	0.365	0.940	0.000	0.000	0.000	30.0

PC	r <sup>2</sup>	SDEC	q <sup>2</sup>	SDEP	r <sup>2</sup> YS	q <sup>2</sup> YS	SDEP YS	CutO
1.000	0.467	0.861	0.341	0.958	0.000	0.000	0.000	30.0
2.000	0.850	0.458	0.487	0.845	0.000	0.000	0.000	30.0
3.000	0.932	0.309	0.667	0.680	0.000	0.000	0.000	30.0
4.000	0.956	0.247	0.708	0.637	0.000	0.000	0.000	30.0
5.000	0.969	0.208	0.698	0.648	0.000	0.000	0.000	30.0
6.000	0.980	0.169	0.702	0.644	0.000	0.000	0.000	30.0
7.000	0.985	0.146	0.711	0.634	0.000	0.000	0.000	30.0
8.000	0.993	0.102	0.708	0.638	0.000	0.000	0.000	30.0



# Costruzione del Modello 3-D QSAR: l'analisi

Comparirà quindi una tabella riassuntiva in cui sono listati tutti i valori statistici del modello, 'cliccando' sul bottone affianco alla voce 'Model Field' si potrà selezionare il modello calcolato considerando sia il campo sterico che quello elettrostatico (BOTH o STE\_ELE).

Per la visualizzazione delle mappe di contorno 'cliccare' sulla voce 'View 3-D QSAR Contour Maps', selezionando però il modello a 4 PC che statisticamente è più robusto in quanto mostra un valore di  $q^2$  molto simile ma con 3 PC in meno.

These are Model's #2808 Statistical Results of Data Set Nuovo\_dataset for Target Generic on enzyme

Model Field: STE

	PC ll	r <sup>2</sup> ll	SDEC ll	q <sup>2</sup> ll	SDEP ll	r <sup>2</sup> YS ll	q <sup>2</sup> YS ll	SDEP YS ll	CUTO
1.000	0.453	0.872	0.179	1.069	0.000	0.000	0.000	30.0	
2.000	0.767	0.570	0.095	1.123	0.000	0.000	0.000	30.0	
3.000	0.837	0.477	0.322	0.971	0.000	0.000	0.000	30.0	
4.000	0.906	0.362	0.185	1.065	0.000	0.000	0.000	30.0	
5.000	0.952	0.259	0.212	1.047	0.000	0.000	0.000	30.0	
6.000	0.979	0.172	0.222	1.041	0.000	0.000	0.000	30.0	
7.000	0.989	0.126	0.310	0.980	0.000	0.000	0.000	30.0	
8.000	0.992	0.104	0.365	0.940	0.000	0.000	0.000	30.0	

These are Model's #2808 Statistical Results of Data Set Nuovo\_dataset for Target Generic on enzyme

Model Field: STE\_ELE

	PC ll	r <sup>2</sup> ll	SDEC ll	q <sup>2</sup> ll	SDEP ll	r <sup>2</sup> YS ll	q <sup>2</sup> YS ll	SDEP YS ll	CUTO
1.000	0.467	0.861	0.341	0.958	0.000	0.000	0.000	30.0	
2.000	0.850	0.458	0.487	0.845	0.000	0.000	0.000	30.0	
3.000	0.932	0.309	0.667	0.680	0.000	0.000	0.000	30.0	
4.000	0.956	0.247	0.708	0.637	0.000	0.000	0.000	30.0	
5.000	0.969	0.208	0.698	0.648	0.000	0.000	0.000	30.0	
6.000	0.980	0.169	0.702	0.644	0.000	0.000	0.000	30.0	
7.000	0.985	0.146	0.711	0.634	0.000	0.000	0.000	30.0	
8.000	0.993	0.102	0.708	0.638	0.000	0.000	0.000	30.0	

# Costruzione del Modello 3-D QSAR: l'analisi

La nuova schermata mostra una finestra con le conformazioni tridimensionali della molecola più attiva (in blu) e di quella meno attiva (in rosso) sovrapposte

The screenshot displays the Py-CoMFA web application interface. The browser address bar shows the URL [www.3d-qsar.com/pycomfa/user/rcmd/model/2808/](http://www.3d-qsar.com/pycomfa/user/rcmd/model/2808/). The page title is "Data Set Nuovo\_dataset Contour Maps at 4 PCs" and the "Model Field" is set to "STE\_ELE". The main content area features a 3D ball-and-stick model of a molecule, with two conformations overlaid: one in blue and one in red. Below the model, there is a dropdown menu for "STE - Positive - COEFF" and a slider set to 0%. The left sidebar contains a "Model Management" section with "Current Model: 2808" and "Go To Model Results" and "Go To Models" links. Below that is a "Molecules" list with the following entries:

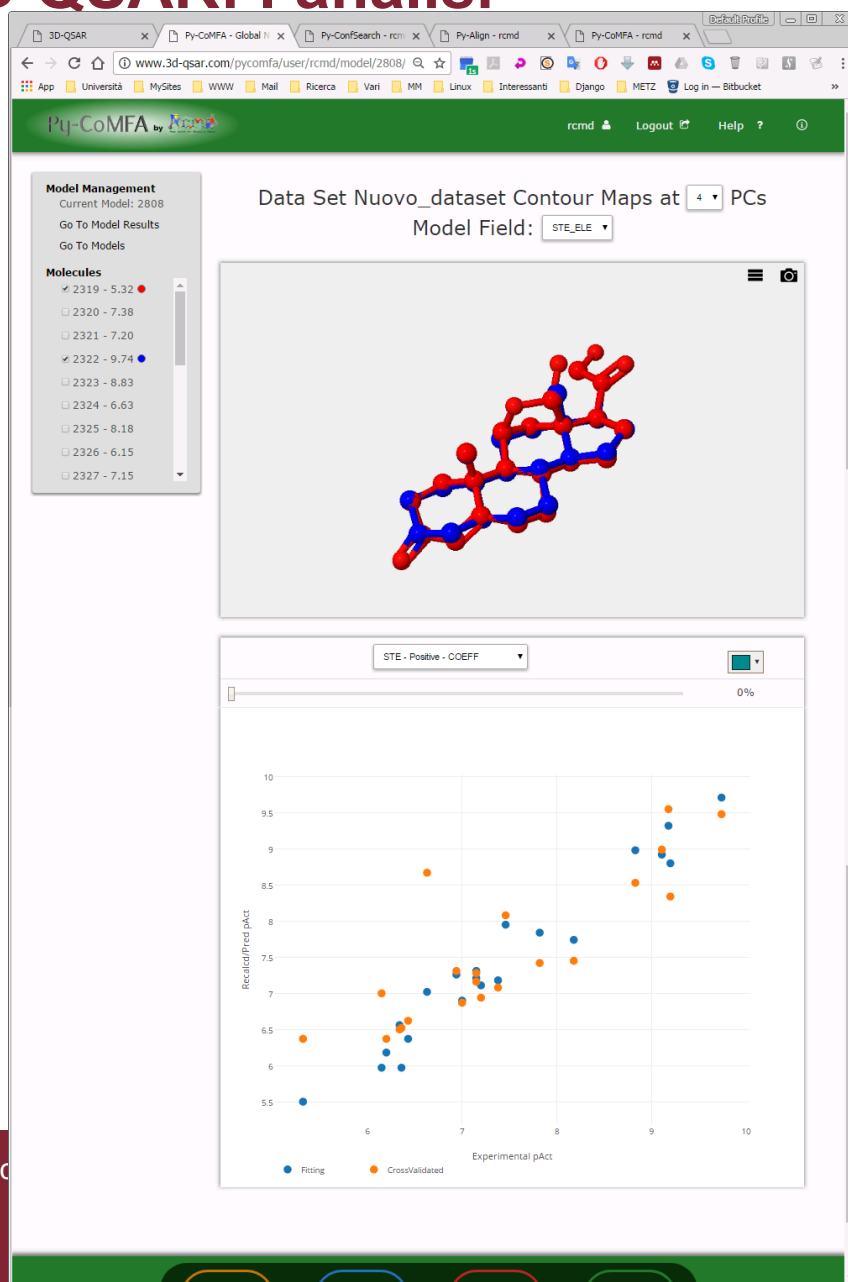
Model ID	Score	Color
2319	-5.32	Red
2320	-7.38	Blue
2321	-7.20	Blue
2322	-9.74	Blue
2323	-8.83	Blue
2324	-6.63	Blue
2325	-8.18	Blue
2326	-6.15	Blue
2327	-7.15	Blue

# Costruzione del Modello 3-D QSAR: l'analisi

La nuova schermata mostra una finestra con le conformazioni tridimensionali della molecola più attiva (in blu) e di quella meno attiva (in rosso) sovrapposte.

In basso viene riportato il plot che mette in relazione i dati ricalcolati (fitting in blu) con quelli cross-validati (in rosso)

Inoltre una barra scorrevole al centro permette la visualizzazione delle mappe di contorno che possono essere selezionate mediante il bottone centrale



# Costruzione del Modello 3-D QSAR: l'analisi

Per l'analisi grafica impostare il tipo di contorno a 'STE – Positive - CoMFA2' e quindi facendo scorrere la barra fino al 90% si otterrà la visualizzazione delle mappe di contorno steriche associate ad un aumento di attività qualora una molecola sia modificata fino a sovrapporsi a tale aree

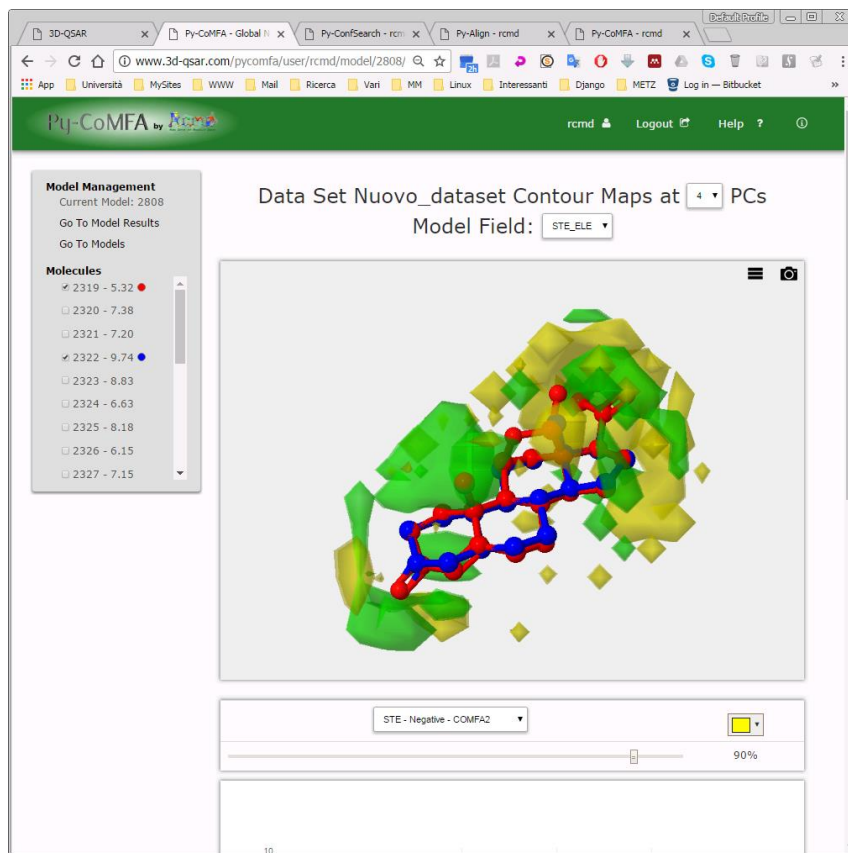
The screenshot displays the Py-CoMFA web application interface. The browser address bar shows the URL [www.3d-qsar.com/pycomfa/user/rcmd/model/2808/](http://www.3d-qsar.com/pycomfa/user/rcmd/model/2808/). The page title is "Data Set Nuovo\_dataset Contour Maps at 4 PCs" and the "Model Field" is set to "STE\_ELE". On the left, the "Molecules" list includes:

ID	Activity
2319	-5.32
2320	-7.38
2321	-7.20
2322	-9.74
2323	-8.83
2324	-6.63
2325	-8.18
2326	-6.15
2327	-7.15

The main visualization area shows a 3D ball-and-stick model of a molecule (red and blue atoms) overlaid with green contour maps. Below the visualization, the "Model Field" is set to "STE - Positive - CoMFA2" and a slider is positioned at 90%.

# Costruzione del Modello 3-D QSAR: l'analisi

Analogamente si può fare con le mappe 'STE - Negative - CoMFA2' e si otterrà una visualizzazione anche delle mappe associate con una diminuzione media dell'attività.

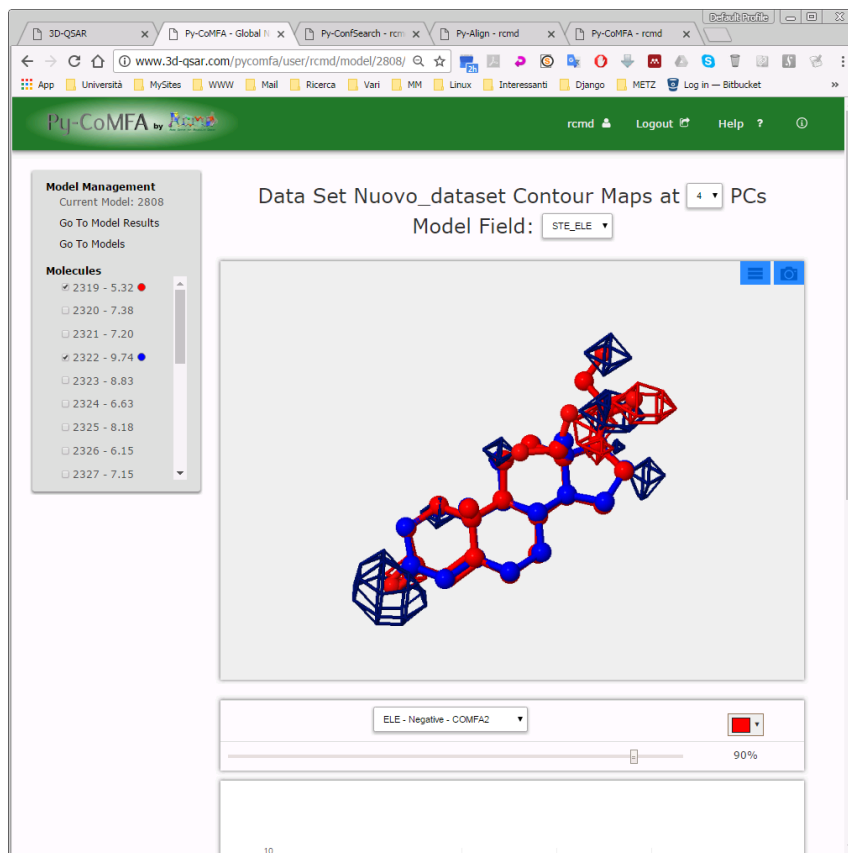


# Costruzione del Modello 3-D QSAR: l'analisi

In modo simile si possono visualizzare le mappe elettrostatiche positive e negative.

L'analisi dettagliata delle mappe di contorno non è oggetto di questo tutorial.

In un prossimo tutorial saranno indicati i metodi per interpretare le mappe di contorno generali e quelle particolari associate ad ogni molecola del training set



## Conclusione

In questo semplice tutorial si è dimostrato come utilizzando le stesse molecole del modello originale di Cramer e collaboratori del 1988 è possibile creare in pochi minuti un modello 3-D QSAR di qualità confrontabile.

In altri seguenti tutorial si prenderà in esame la creazione di modelli da molecole disegnate su Py-MolEdit o da una lista di stringhe SMILES facilmente ottenibile dalle pubblicazioni scientifiche come quelle che sono riportate da Journal of Medicinal Chemistry.