



Guglielmo Rucci
guglielmo.rucci@virgilio.it

LA RICERCA DELLE REAZIONI IN STN

Vengono presentate due differenti strategie di ricerca documentale per l'individuazione delle reazioni chimiche. La piattaforma STN, possedendo una corretta conoscenza degli strumenti di indagine bibliografica, ivi presenti, permette di reperire, in maniera rapida, la documentazione cercata, sia sotto forma di abstract, sia, se richiesto, come articolo completo.

Nel passato la ricerca della documentazione per il chimico di sintesi rappresentava uno dei problemi di maggiore impegno da affrontare, specialmente se si intendeva verificare anche la copertura brevettuale. In questi casi o si faceva ricorso alla cultura individuale e del proprio gruppo di lavoro, o si era costretti ad affrontare una lunga ricerca in biblioteca dagli esiti incerti. A partire dall'inizio degli anni Ottanta del secolo scorso, si sono rese disponibili, in rete, una serie di banche dati contenenti informazioni di tutti i tipi, tra cui, di particolare importanza, quelle concernenti la documentazione chimica, che oggi agevolano e rendono più precisa e selettiva la ricerca documentale.

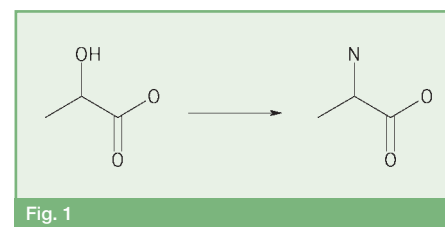
Queste banche dati sono inserite in organizzazioni commerciali ("Hosts"), che hanno messo a punto software per permettere, al loro interno, una metodologia uniforme di consultazione. Gli "Hosts" più importanti oggi sul mercato sono STN®, Dialog® e Questel®.

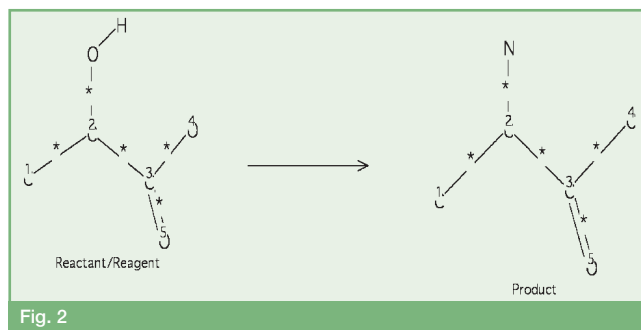
STN (Scientific and Technical Network) è la piattaforma informatica che prevalentemente si è dedicata alle banche dati tecniche e scientifiche contenenti pubblicazioni di riviste e brevetti. Naturalmente dentro STN si trovano tutte le più importanti informazioni relative alla chimica, comprendendo in questa definizione i diversi aspetti: dalla farmaceutica ai polimeri, dalle reazioni alle grandezze chimico-fisiche e termodinamiche, agli spettri ecc.

Per quanto attiene alle reazioni gli archivi ("files") più importanti in STN sono: Beilstein®, Casreact®, Cheminformx®, Djsm®, Registry®, Caplus®.

In questo articolo sono presentati alcuni esempi di ricerca di reazioni utilizzando Casreact, Registry e Caplus, che sono prodotti e di proprietà del CAS (Chemical Abstracts Service), che è una sezione dell'American Chemical Society.

Casreact è un archivio dedicato esclusiva-





mente alle reazioni e ne contiene circa dieci milioni, provenienti da pubblicazioni scientifiche e da brevetti. La copertura temporale si estende dall'inizio del 1900 ai giorni nostri, con un aggiornamento continuo. Gli aspetti più importanti dell'organizzazione interna di questo "file" sono i seguenti: a) le sostanze coinvolte nelle reazioni, nei diversi ruoli (reagenti, substrati, prodotti, solventi), sono indicizzate con i Registry Numbers® (RNs), codici univoci per identificare le sostanze chimiche, di proprietà del CAS; b) è possibile indicare sia gli atomi coinvolti nella reazione sia quali siano i legami soggetti ad una modifica parziale o totale.

Registry contiene circa ottanta milioni di composti chimici, che sono quelli riportati nella pubblicazione Chemical Abstracts, dal 1907 ad oggi, con l'introduzione nel "file" di circa un milione di nuove entrate ogni anno. Ogni composto ha una scheda in cui sono generalmente riportati, oltre al RN, la formula bruta, la struttura, e, in molti casi, anche gli spettri ed i valori di parametri chimico-fisici, quali la temperatura di ebollizione, di fusione, la tensione di vapore, ecc. Sostanzialmente questo archivio viene usato per individuare i RNs dei composti e, attraverso questi, trovare le informazioni bibliografiche nei "files", quali Casreact o Caplus, che usano questi codici per indicizzare i composti chimici.

Caplus sostanzialmente contiene quanto

pubblicato nei Chemical Abstracts. Si trovano quindi qui gli estratti dei lavori o brevetti recensiti dal CAS, e, come in Casreact, i composti chimici sono indicizzati con i RNs, anche se con un ruolo non così preciso. Il "file" contiene circa trenta milioni di documenti e viene aggiornato continuamente.

Ricerca di una reazione utilizzando Casreact e Registry

Un buon esempio dell'utilizzo combinato dei due archivi in questione è riportato nella ricerca della reazione riportata in Fig. 1 in cui è richiesta una sostituzione stereoselettiva (con ritenzione o inversione di configurazione) dell'atomo di ossigeno con l'atomo di azoto; successivamente si vuole limitare poi la ricerca alle reazioni in cui non viene usato un reattivo contenente zolfo e sono richieste, inoltre, reazioni con rese superiori al 85%.

Utilizzando STN Express®, un software che permette, in locale, di preparare i quesiti di ricerca per struttura da lanciare in STN, si

costruisce il seguente schema della reazione di Fig. 2, in cui al di sotto delle molecole sono riportati i rispettivi ruoli e all'interno delle molecole sono individuati:

- a) il legame che si rompe (C-O) e il legame che si forma (C-N) (sito della reazione);
- b) i legami che rimangono invariati (tutti i rimanenti);
- c) gli atomi corrispondenti tra il substrato e il prodotto.

Il ruolo "Reactant/Reagent" alla molecola di sinistra viene scelto, anche se sarebbero possibili i ruoli singoli, per coprire tutte le possibilità, dato che su queste definizioni può esserci qualche confusione. Inoltre la mappatura completa, come qui riportata, dei legami e degli atomi, non è, in genere, da applicare integralmente in queste ricerche, specialmente quando non si vogliono perdere informazioni.

Una volta lanciata questa ricerca per struttura all'interno di Casreact si crea un primo raggruppamento ("set I") di risposta, che contiene 213 pubblicazioni, in cui sono state individuate 1120 reazioni corrispondenti con la nostra richiesta generale. Di seguito viene riportato un esempio dei documenti e delle reazioni individuate (Fig. 3).

Rimane adesso il compito di soddisfare

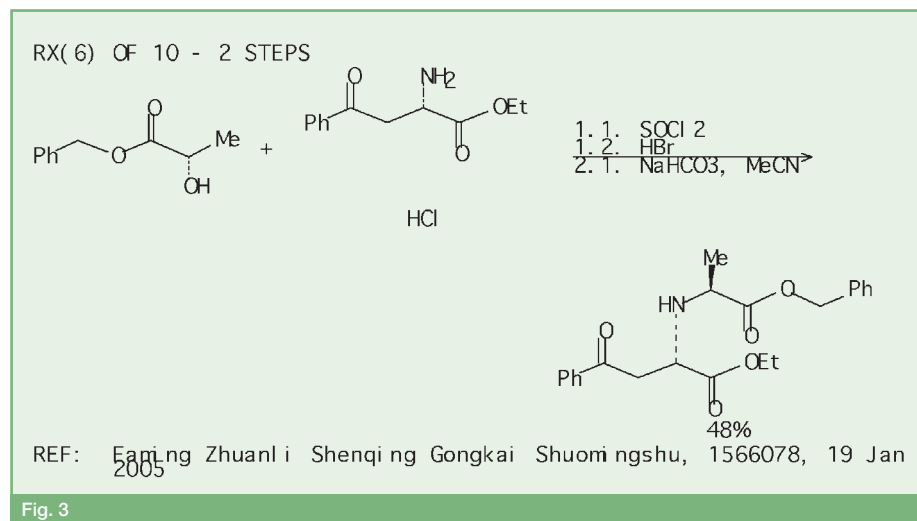


Fig. 3

tutte le altre richieste previste nel quesito iniziale. Per questa seconda parte diventa determinante l'interazione tra Casreact e Registry.

Come primo passo vengono fatti estrarre dal sistema, nell'insieme di risposte ottenute i RNs dei composti formati ("Products"), creando in questa maniera un raggruppamento ("set II") contenente 984 RNs. Quindi, con la stessa procedura, si isolano i 993 RNs dei substrati ("Reactants") e i 182 RNs dei reagenti ("Reagents"). Questi tre insiemi vengono portati nel file Registry, dove il primo ("Products") viene vagliato con un filtro che isola i composti con stereochimica definita (804 RNs) ("set III"), mentre nel secondo e nel terzo ("Reactants/Reagents") si isolano i composti contenenti zolfo (269 RNs) ("set IV"). Si torna quindi nel file Casreact dove il risultato della ricerca più generale ("set I"), viene opportunamente combinato con i RNs dei

prodotti stereofornati ("set III"), e, nello stesso tempo vengono eliminate le reazioni che contengono almeno uno dei composti contenenti zolfo o nel campo "Reactants" o nel campo "Reagents" ("set IV").

Si isolano così 35 lavori che, ulteriormente filtrati con la condizione di resa >85%, si riducono a 15 di cui si riportano due esempi (Fig. 4 a,b).

Nel secondo esempio (Fig. 4b) non compaiono immediatamente i composti cercati come prodotti. Se però si vanno a esaminare i vari composti trattati nel lavoro si trovano i due composti di Fig. 4c, di cui il primo è il "Product" e il secondo il "Reactant".

Questo iter, che può sembrare complesso, in realtà richiede solamente una buona conoscenza dei "files" utilizzati. Infatti, disponendo di questa, tutta la ricerca si svolge nell'arco di circa 20 minuti, con un costo complessivo di circa 150 euro.

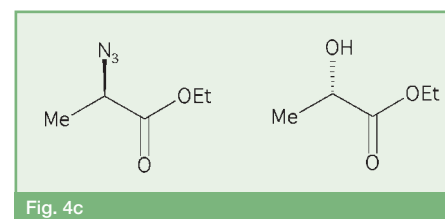
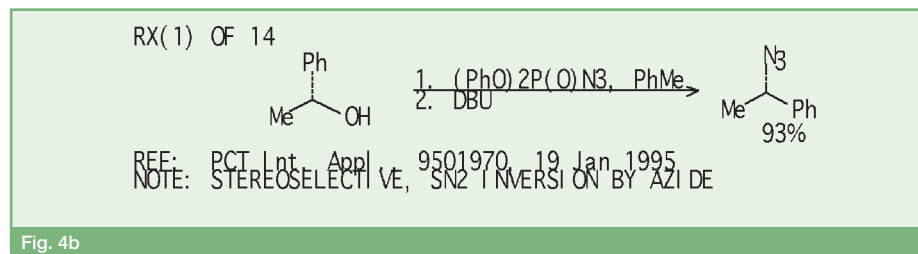
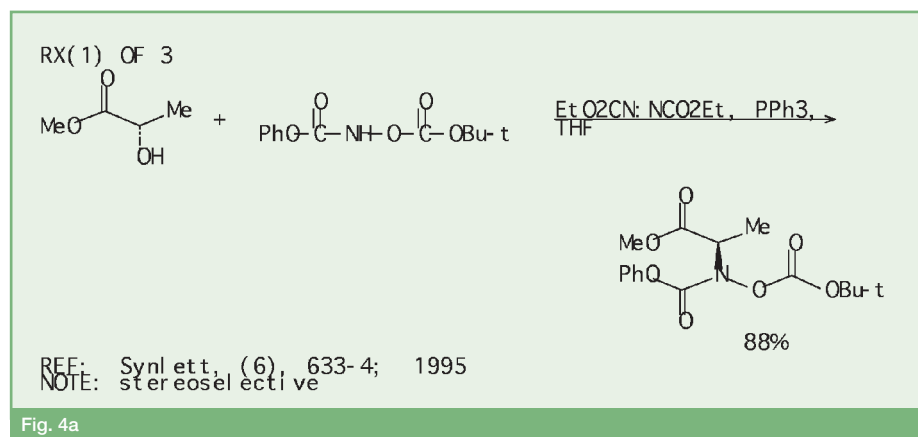
Inoltre se alla fine, per il o i lavori interessanti trovati, non si hanno a disposizione gli originali (non essendo sufficiente l'abstract, sempre presente), questi possono o essere ordinati immediatamente e ricevuti in breve tempo per posta o per fax, ad un costo contenuto, circa 10 euro, oppure, se disponibili, possono essere recuperati immediatamente in formato elettronico (PDF), con un costo, però, di circa 25 euro.

Ricerca di una reazione utilizzando Registry e Caplus

Questo secondo metodo di ricerca delle reazioni non è così preciso, selettivo e generale come il precedente ed è sicuramente più costoso (circa il doppio), permette però una copertura maggiore, sia per quanto concerne le pubblicazioni, sia per l'arco temporale. Inoltre, seguendo questa via, si devono, necessariamente, considerare per intero le molecole principali che prendono parte alla reazione. Questa è, da un punto di vista generale, una limitazione ma, d'altra parte, permette di trovare immediatamente il o i lavori, che trattano direttamente le molecole d'interesse. Il confronto tra i due diversi approcci di ricerca, come numero di risultati ottenuti, non è quindi significativo visto che le due diverse vie di indagine rispondono ad esigenze diverse.

Si consideri, ad esempio, la reazione riportata in Fig. 5.

Il percorso di questa ricerca parte dall'individuare, nel file Registry, tutte le strutture



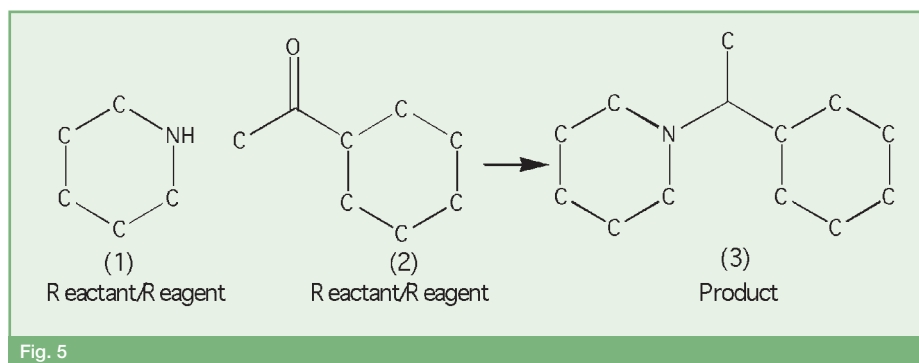


Fig. 5

contenenti nel proprio interno o uno dei "Reactants/Reagents" o il "Product". Ad esempio, le strutture cicliche qui considerate possono essere anche inglobate in sistemi policiclici più complessi.

La struttura (1) è presente in 84.486 strutture del Registry (Set 1), la struttura (2) si ritrova in 61.710 strutture del Registry (Set 2) e la struttura (3) è contenuta in 6.124 strutture del Registry (Set 3).

I Set 1-3, contenenti i RNs dei composti selezionati, vengono, quindi, trasferiti nel file Caplus, dove, in un'unica ricerca, i primi due vengono cercati con il ruolo "Reactants/Reagents" mentre il terzo con il ruolo di "Product".

Si ottengono così 83 riferimenti (brevetti e pubblicazioni) di cui si riporta, di seguito, un esempio, in cui, oltre al titolo e ai riferimenti bibliografici, sono rappresentate alcune molecole coinvolte nella reazione cercata (Fig. 6).

Conclusioni

In questo articolo si sono volute sottolineare le attuali, molteplici, possibilità presenti nella ricerca bibliografica di reazioni, sia su brevet-

ti sia su pubblicazioni scientifiche, condotta usando gli archivi elettronici ora disponibili in STN. In particolare è importante comprendere che l'utilizzo sinergico delle varie banche dati permette di recuperare le informazioni desiderate, in maniera quanto più completa possibile, eliminando allo stesso tempo un gran numero di documenti superflui o poco attinenti con l'obiettivo prefissato.

Oggi si può senza dubbio affermare che,

sapendo usare con strategia e logica questi strumenti, una conoscenza che si può apprendere seguendo i corsi gratuiti di STN che vengono periodicamente effettuati a Milano (principalmente) e a Bologna (per ulteriori informazioni sui corsi: <http://www.stn-international.de/>, <http://www.cas.org> o rivolgersi all'autore), quello che una volta era un compito lungo, tedioso, difficile, da effettuare in biblioteca, può diventare un divertimento, ottenendo risposte coerenti con gli obiettivi della ricerca, in alcune decine di minuti.

Infine, come si può facilmente immaginare, queste metodologie di ricerca bibliografica, di cui quella sulle reazioni qui presentata è solo un esempio, rivestono una particolare importanza per la ricerca brevettuale, per cui sono nate, alla fine degli anni Sessanta, presso le più grandi società chimiche e farmaceutiche tedesche.

TI Preparation of 1,2,3,4-tetrahydropyrazin-2-yl acetamides as bradykinin B1 receptor antagonists for treating pain and diseases such as inflammation-mediated diseases.

PI US 2006025400 A1 20060202 US 2005-182216 20050715
WO 2006019975 A1 20060223 WO 2005-US25090 20050715

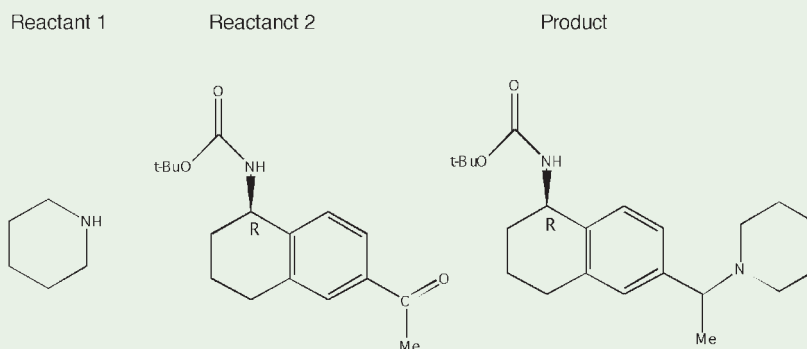


Fig. 6

Chemical Reactions Research by STN

Two different strategies for retrieving publications on chemical reactions are here presented. STN, if you know the tools available in this platform, allows you to retrieve, in a very quick way, the required information either as abstracts, or as full-text.

ABSTRACT