

# ANALISI CORRELATIVE STRUTTURA ATTIVITÀ 6 CFU

Prof. Giuseppe Romeo - A.A. 2011/2012

## PROGRAMMA

### GENERALITÀ SUL CORSO.

#### RAPPRESENTAZIONE DELLE MOLECOLE.

Grafo di una molecola. Tabelle di connessione. Matrici di adiacenza. Matrici di distanza topologica. Matrici di distanza 3D. Structure data format (Molfile). Esempi di indici topologici: Topological autocorrelation vector. Notazione SMILES (Simplified Molecular Input Line Entry): regole principali ed esempi di notazione. Indici di similarità: Tanimoto coefficient. Motivi strutturali in una molecola secondo Bemis and Murcko: Ring systems, Linkers, Side-chains, Frameworks o Scaffolds. Motivi strutturali privilegiati e motivi strutturali indesiderati in drug-design.

#### INTERAZIONI FARMACO-RECETTORE.

Relazione tra variazione di energia libera e costante di equilibrio del complesso farmaco-recettore. Relazione tra variazione di energia libera standard e  $K_i$ . Competition binding assays: protocollo sperimentale. Curva di spiazzamento del ligando marcato e definizione di  $IC_{50}$ . Calcolo del valore di  $K_i$ ; equazione di Cheng e Prusoff con esempi di calcolo. Forze che favoriscono o sfavoriscono l'interazione farmaco-recettore. Cenni ai diversi tipi di legami e interazioni tra farmaco e recettore: legame a idrogeno, interazioni elettrostatiche, effetto idrofobico, interazioni di van der Waals, interazioni tra sistemi aromatici. Studio di un caso: interazioni tra Carazololo e recettore  $\beta_2$ -adrenergico (D. M. Rosenbaum et al. *Science*, 318, 1266-1273, 2007).

#### PROPRIETÀ DELLE MOLECOLE.

Parametri chimico-fisici e descrittori molecolari. Parametri di lipofilia:  $\log P$ ; costante idrofobica del sostituente di Hansch-Fujita,  $\pi$ . Metodi per la misura sperimentale del  $\log P$ . Metodo dello Shake-flask, protocollo sperimentale secondo la OECD Guideline n° 107, vantaggi e svantaggi del metodo. Metodo in RP-Thin Layer Chromatography, definizione di  $R_M$  e relazione tra  $R_M$  e  $\log P$ . Metodo in RP-HPLC, definizione di fattore di capacità e sua relazione con  $\log P$ . Metodi di calcolo del  $\log P$ . Metodo per sostituzione. Metodi fragment-based (Nis and Rekker; Hansch and Leo) ed esempi di calcolo con CLOGP e ACD/Log P softwares. Metodi atom-based. Definizione di  $\log D$ . Formule per il calcolo di  $\log D$  per composti acidi o basici. Grafici di  $\log D$  in funzione del pH per diverse tipologie di principi attivi. Esempi generati con ACD/Log D software. Importanza del  $\log D_{7.4}$  sulle drug-like properties.  $\log P$  come descrittore molecolare in QSAR: esempi di modelli lineari e non lineari.

Parametri elettronici. Costanti di Hammett,  $\sigma$ , e loro definizione. Equazione di Hammett e definizione di  $\rho$ . Componenti  $\mathfrak{S}$  e  $\mathfrak{R}$  di Swain e Lupton.

Parametri sterici. Costante di Taft,  $E_s$ , sua definizione e correlazione con il raggio di van der Waals. Costante di Taft corretta,  $E_s^c$ . Rifrazione molare,  $M_R$ , sua definizione e relazione con il volume molare. Volume di van der Waals,  $V_w$ . Parametri STERIMOL ( $L$ ,  $B_1$ - $B_4$  e  $L$ ,  $B_1$ ,  $B_5$ ) di Verloop e loro definizione.

Esempi di descrittori topologici. Total Adjacency Index con esempio di calcolo. Kier and Hall Connectivity Indices,  $\chi$ , con esempi di calcolo per  ${}^1\chi$  e  ${}^2\chi$ . Numero di gruppi accettori (HBA) e donatori (HBD) in una molecola. Polar Surface Area (PSA) e Topological Polar Surface Area (TPSA), definizioni e metodi di calcolo.

### LINEE GUIDA PER LA PROGETTAZIONE DI DRUG-LIKE COMPOUNDS.

Regole di Lipinski (Regola del 5): origine, definizione, utilizzo. Regole di Veber. Regola del 3 e sua applicazione in Fragment-based Screening. Definizione di Rotatable bond. Ligand Efficiency (LE) e Binding Efficiency Index (BEI), loro definizione e calcolo. Influenza delle dimensioni molecolari su LE.

#### ANALISI DI REGRESSIONE IN QSAR.

Analisi di regressione lineare semplice. Definizione e requisiti necessari per poter eseguire un'analisi di regressione lineare. Metodo dei minimi quadrati. Formule per il calcolo del coefficiente angolare e dell'intercetta. Coefficiente di correlazione,  $r$ , definizione e formula per il calcolo. Esempi di calcolo su un set di dati. Definizione di Total Sum of Squares (TSS), Explained Sum of Squares (ESS) e Residual Sum of Squares (RSS). Coefficiente di determinazione,  $r^2$ , e formule per il calcolo. Errore standard di  $y$ , del coefficiente angolare e dell'intercetta. Calcolo degli intervalli di confidenza per il coefficiente angolare e l'intercetta. Analisi della Varianza (ANOVA) e Statistical Hypothesis Testing. Espressione dell'ipotesi nulla e dell'ipotesi alternativa nel contesto dell'analisi di regressione lineare semplice. Calcolo delle Mean Squares (MSE, MSR, MST). F-test come metodo statistico per l'accettazione o il rigetto dell'ipotesi nulla ed esempio di calcolo.

Analisi di regressione lineare multipla. Descrizione del modello e  $R^2$ . F-test in analisi di regressione lineare multipla. Strategie per la creazione di un modello di regressione multipla: Backward Elimination e calcolo di F-to-remove; Forward Inclusion e calcolo di F-to-enter.

Procedure di cross-validazione di un modello regressionale. Leaving One sample Out at time (LOO) e calcolo di  $Q^2$ . Definizione di Training set, Evaluation set e Test set. Utilizzo del coefficiente di correlazione  $r$  tra attività predette e attività sperimentalmente determinate nella validazione di un modello. Analisi delle correlazioni casuali tra variabili in un modello regressionale: metodo dello Y scrambling.

#### APPROCCI DI HANSCH E DI FREE-WILSON IN QSAR

Generalità sull'approccio di Hansch. Equazioni di Hansch. Scelta delle variabili indipendenti e termini quadratici nell'equazione. Interpretazione dell'equazione in base al valore e al segno dei coefficienti. Esempi di equazioni di Hansch su sets di molecole ad attività  $\alpha$ -adrenolitica e ad attività antimalarica. Grafico di Craig per la scelta dei sostituenti. Approccio di Free-Wilson: generalità e definizione di variabile indicatrice.

#### ANALISI DELLE COMPONENTI PRINCIPALI (PCA) E PARTIAL LEAST SQUARES (PLS)

Matrice di correlazione tra variabili. Collinearità. Pre-trattamento dei dati: autoscaling. Generalità e principi dell'analisi delle componenti principali. Interpretazione geometrica delle componenti principali. Scores e Loadings: matrici, plots e loro interpretazione. Eigenvalues. PCA come tecnica di classificazione. Studio di un caso: Profili di binding di farmaci antipsicotici (J.H.L. Lange et al., *J. Med. Chem.* 50, 5103-5108, 2007). Regressione sulle componenti principali (PCR). Generalità e principi della tecnica PLS. Caratteristiche delle variabili latenti. Definizione e utilizzo della PRESS. Interpretazione geometrica delle variabili latenti.

#### 3-D QSAR: COMPARATIVE MOLECULAR FIELD ANALYSIS (CoMFA)

Generalità e principi del metodo. Definizione IUPAC di farmacoforo. Campi sterici e elettrostatici. Interpretazione delle contour maps. Studio di un caso: CoMFA study of piperidine analogues of cocaine at the dopamine transporter (H. Yuan et al. *J. Med. Chem.* 47, 6137-6143, 2004).

#### TESTI CONSIGLIATI

- G.L. Patrick, *Introduzione alla Chimica Farmaceutica*, II Edizione, Edises, 2010.
- G. Schneider, K.-H. Baringhaus, *Molecular Design*, Wiley-VCH, 2008.
- D. Livingstone, *Data analysis for chemists*, Oxford Science Publications.

- D. Livingstone, *A practical guide to scientific data analysis*, Wiley, 2009.
- D.J. Abraham (Edited by), *Burger's Medicinal Chemistry & Drug Discovery*, Vol.1, Sixth edition, Wiley, 2003.
- C. Hansch, *Comprehensive Medicinal Chemistry*, Vol. 4, Pergamon Press, 1990.