



Universiteit
Leiden
The Netherlands

Déjà Vu - Réjà Vu : on knowledge-based approaches linking ligand and target information to bioactivity

Westen, G.J.P. van

Citation

Westen, G. J. P. van. (2013, January 8). *Déjà Vu - Réjà Vu : on knowledge-based approaches linking ligand and target information to bioactivity*. Retrieved from <https://hdl.handle.net/1887/20394>

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/20394>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/20394> holds various files of this Leiden University dissertation.

Author: Westen, Gerard Jacob Pieter van

Title: Déjà Vu - Réjà Vu : on knowledge-based approaches linking ligand and target information to bioactivity

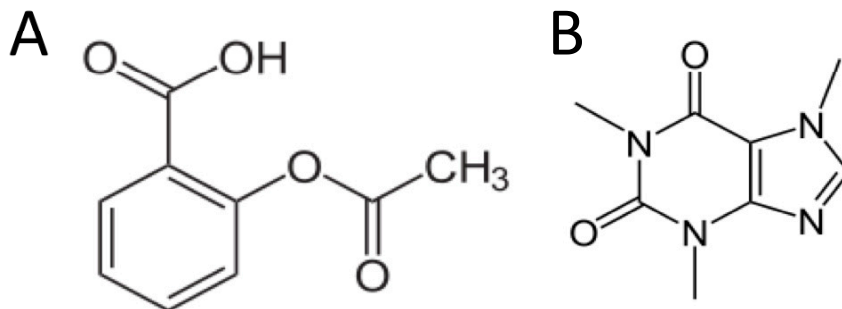
Issue Date: 2013-01-08

Samenvatting voor leken

De hoofdvraag in dit proefschrift draait om het combineren van data uit gerelateerde disciplines (chemie, biologie en bio-activiteit). Deze concepten verdienen een nadere toelichting, wat wordt ermee bedoeld?

Chemie

Met chemie wordt hier een subklasse van de scheikunde bedoeld, namelijk deze die zich toelegt op de zogenaamde kleine moleculen ('small molecules') die een potentieel effect in het menselijk lichaam hebben. Kleine moleculen betekent precies dat wat men zou verwachten, relatief kleine chemische stoffen (zodat deze beter opgenomen kunnen worden wanneer ze in de vorm van een pil toegediend worden). De kleine moleculen vormen dan ook een van de speerpunten van de moderne medicinale chemie. Voorbeelden hiervan zijn aspirine en cafeïne (**Figuur A1**). De gelijksoortigheid van moleculen kan bij medicinale chemie een richtlijn vormen, wanneer een klein molecuul (molecule 1) een bepaald effect heeft is het waarschijnlijk dat een klein molecuul (molecule 2) wat hier sterk op lijkt (gelijksoortig is) een vergelijkbaar effect heeft ('molecular similarity principle').



Figuur A1: De structuur van aspirine (a) en cafeïne (b). Beide behoren tot de klasse van kleine moleculen en beïnvloeden het menselijk lichaam. Hiermee vallen deze stoffen in de klasse van zogenaamde bio-actieve stoffen.

Biologie

Met het concept biologie wordt in dit proefschrift verwezen naar de biologische aangrijpingspunten van kleine moleculen in het menselijk lichaam ('targets'). Deze aangrijpingspunten zijn vaak eiwitten waarvan de normale signaalverwerking in een ziekteproces verstoord is. Herstel van dit signaalsverwerkingsproces kan in theorie leiden tot genezing en in het optimale geval wordt dit mogelijk gemaakt door small molecules welke in pilvorm toegediend kunnen worden.

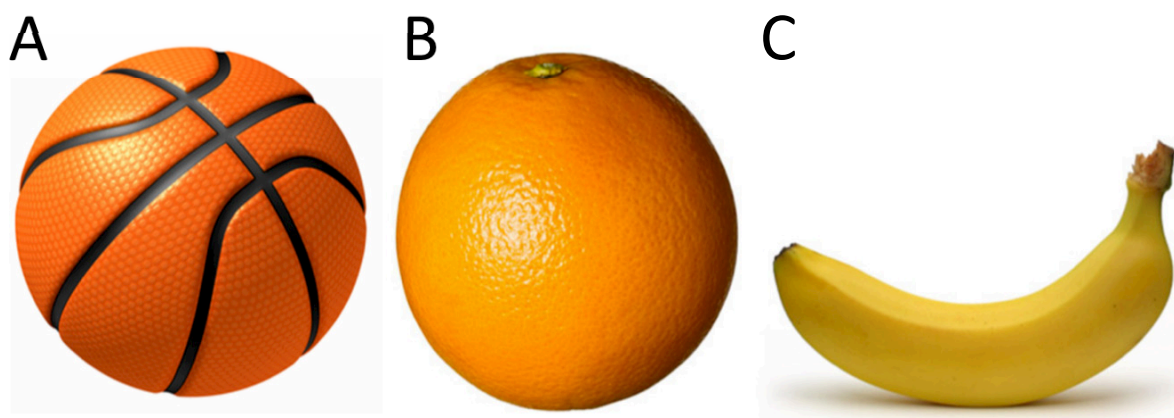
In het geval van targets kan gelijksoortigheid ook gebruikt worden als richtlijn. Stel dat een bepaald target (target A) een verstoorde signaalverwerking heeft en er kleine moleculen bekend zijn die de functie herstellen. Het is dan waarschijnlijk dat een target wat hier sterk op lijkt (target B) ook beïnvloed zal worden door dezelfde kleine moleculen. Deze effecten staan bekend als bio-activiteit.

Bioactiviteit

De manier waarop en de mate waarin een klein molecuul een target kan beïnvloeden wordt bioactiviteit genoemd. Bioactiviteit is een erg breed concept wat strekt van directe effecten tot indirecte effecten. Bioactiviteit kan zelfs resulteren in effecten op het nageslacht. In dit proefschrift wordt de definitie beperkt tot de affiniteit of directe effecten van één of meerdere moleculen op één of meerdere specifieke targets.

Molecular Similarity Principle

De bovengenoemde kleine moleculen (chemie) en targets (eiwitten) kunnen beschreven worden op verscheidene manieren. Gedacht kan worden aan het gewicht van een molecuul op atomaire schaal, hoeveel van een molecuul op te lossen is in 1 liter water etc. Deze eigenschappen waarmee een object te beschrijven is worden ook wel 'Descriptoren' genoemd. De gedachte is dat moleculen die veel op elkaar lijken in deze eigenschappen, ook veel op elkaar zullen lijken in hun effecten op het lichaam. Zie voor een voorbeeld **Figuur A2**, op het eerste gezicht zullen de basketbal en sinaasappel het meeste op elkaar lijken. Echter wanneer de vraag is welk object voedsel is, blijkt dat de sinaasappel en de banaan meer gemeen hebben. De eigenschap waarvan getracht wordt deze met een model te beschrijven (hier of een object voedsel is) aan de hand van de bekende descriptoren, staat bekend als de 'Output Variable'.

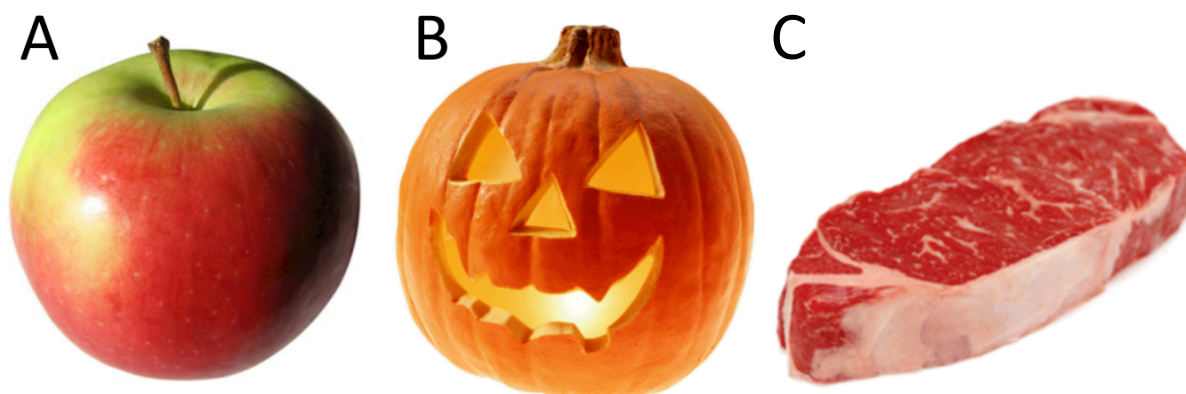


Figuur A2: Het concept van gelijksoortigheid kan veraderlijk zijn en is afhankelijk van de situatie. Hoewel (a) en (b) wellicht het meest gelijkend zijn op het eerste gezicht zal de situatie veranderen wanneer de vraagstelling is welk object waarschijnlijk voedsel is.

Modelleren van bioactiviteit

Zoals gezegd worden in dit proefschrift datasets gebruikt welke chemie, biologie en bio-activiteit combineren. Deze gecombineerde data sets worden vervolgens gebruikt om modellen te creëren welke voorspellingen kunnen doen over de bio-activiteit van (nog) onbekende kleine moleculen. De verwachting is dat de combinatie van deze gegevens zal leiden tot kwalitatief betere voorspellingen dan de voorspellingen van modellen welke gebaseerd zijn op de data van slechts één van deze bovenstaande disciplines. De modellen worden gecreëerd uit sets van moleculen waarvan bekend is welke effecten zij hebben op het menselijk lichaam. Het resulterende model kan aan de hand van de eigenschappen van deze moleculen en van andere, vergelijkbare, moleculen de effecten op het menselijk lichaam voorspellen.

Het concept zal geïllustreerd worden aan de hand van een voorbeeld. De beschreven eigenschap van ons model (output variable) zal zijn om te voorspellen of een object voedsel is. Het model zal gebaseerd worden op de drie objecten uit **Figuur A2**. De descriptor die gebruikt worden zijn vorm, suikerconcentratie, kleur en de aanwezigheid van rubber. Het resulterende model wordt vervolgens gebruikt om te voorspellen of de objecten uit **Figuur A3** voedsel zijn. Het model zal vervolgens hierbij voor de eerste twee een positief oordeel geven en voor het laatste een negatief oordeel.



Figuur A3: Een aantal onbekende situaties voor ons model. Wanneer de vraagstelling is welk object opgegeten kan worden zal een model getraind op de objecten uit Figuur A2 deze voor (a) met yes ('active') beantwoorden, ook wel 'True Positive'. Een voorspelling voor (b) zou ook active zijn maar een 'False Positive'. Tot slot kan het model voor (c) voorspellen dat dit object geen voedsel is, 'False Negative' genoemd.

De appel is voedsel, het model heeft het hier dus bij het rechte eind ('True Positive'), de pompoen is ook voedsel maar is op dit moment niet eetbaar gezien er een kaars in brandt (vals positief of 'False Positive'). De entrecote wordt ten onrechte beoordeeld als geen voedsel (vals negatief of 'False Negative'). Deze twee foute voorspellingen zijn niet zozeer aan het model te wijten. De false positive wordt veroorzaakt door een tekortkoming in de beschrijving van de objecten (het model heeft nooit geleerd wat een kaars is daar dit niet meegenomen werd in de descriptoren).

Het geval van de false negative kan waarschijnlijk geweten worden aan het feit dat het model slechts geleerd heeft om voor fruit de eetbaarheid te voorspellen, vlees heeft het nooit gezien en dientengevolge kan het hierover geen betrouwbare voorspelling doen. Het is aan de wetenschapper om deze tekortkomingen te voorkomen door de juiste descriptoren te kiezen en het model te baseren op een zo volledig en representatief beeld van de werkelijkheid.

Samenvatting van de hoofdstukken in dit proefschrift

In het **eerste hoofdstuk** worden de concepten en definities omschreven zoals ze in dit proefschrift gebruikt worden. Het **tweede hoofdstuk** bevat een literatuuronderzoek van de gebruikte techniek ('Proteochemometric modeling'). **Hoofdstuk 3** vergelijkt een aantal descriptoren welke gebruikt worden om meerdere chemische datasets te koppelen. **Hoofdstuk 4** bevat een studie naar modellen welke de adenosine receptoren beschrijven (ook wel verantwoordelijk voor de effecten van cafeïne). In **hoofdstuk 5** wordt de techniek toegepast op kandidaat medicijnen met een potentieel HIV remmend effect. **Hoofdstuk 6** gaat hierin een stap verder en bevat modellen welke een persoonlijk behandelingsschema voor HIV patiënten kunnen voorspellen. Een vergelijkbare aanpak op moleculaire structuren wordt in **hoofdstuk 7** gepresenteerd. Tot slot bevat **hoofdstuk 8** conclusies en een aantal toekomstperspectieven.