Application of SAR methods to non-congeneric data bases assocated with carcinogenicity and mutagenicity: Issues and approachs

Ann M. Richard

United States Environmental Protection Agency, Carcinogenesis and Metabolism Branch (MD-68), Reaseach Triangle Park, NC 27711, USA

...Protection Agency, Carcinogenesis and **Metabolism** Branch (MD-68), Research Triangle Park...on the relative roles of transport, **metabolism**, receptor interactions, initiation...molecular features, or an unknown route of **metabolism** to an intermediate having an electrophilic...

Abstract

In both industry and government, structure-activity relationships (SAR) are capable of playing an important decision-support role in estimating the potential mutagenicity or carcinogenicity of chemicals for which bioassay test results are unavailable. Traditional SAR modeling approaches, however, are usually restricted to the consideration of structurally similar chemical congeners. The highly structurally diverse nature of current carcinogenicity and mutagenicity data bases has motivated development of more general SAR approaches, potentially applicable to the treatment of diverse, non-congeneric mutagenicity and carcinogenicity data bases. Three specific approaches are considered in some detail — Ashby's structural alert moedl, classified as a "rule-based" SAR approach, and the computerized CASE fragment-base method and TOPKAT linear discriminant equation method, both classified as "correlative" SAR approaches. Relative strengths and limitations, and a number of common features and important distinctions betwee these 3 methods are discussed. Rule-base methods are highly flexible and able to incorporate many different types of relevant information, yet are biased towards current knowledge, viewpoints, and mechanistic assumptions, that may or may not hold true. Correlative SAR methods are less biased and offer the promise of "discorvering" potentially new SAR associations that could lend fresh insight into the basis for a structure-activity association. However, problems associated with their application to non-congeneric data bases relate to: modeling multiple or overlapping mechanisms of action with a single relationship; defining the range of applicability of models in complex multi-dimensional structure-activity space; assigning confidence levels to predictions in the absence of knowledge concerning mechanisms of activity; and determining the potential mechanistic significance of diervse model parameters. It is argued that many of these concerns can be partially alleviated by careful application of statistical procedures, scrutiny of model results, and estabilishment of reasoned limits to the range of model applicability. The most significant confidence-building measure, however, will be a rationalization of the correlative SAR model and model parameters in terms of principles of chemical reactivity and postulated molecular mechanism(s) for the biological activity. Hence, it is recommended that models and model descriptors be designed to facilitate mechanistic interpretation and hypothesis generation. Finally, problems in comparing the relative predictive capabilities of different SAR approaches are discussed, and strategies for SAR investigation involving integration of existing techniques are suggested.

Author Keywords: SAR methods, application; Non-congeneric data bases; Structure-activity relationships; CASE; TOPKAT; Structural alerts; Screening, carcinogenicity, metagenicity

Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis
Volume 305, Issue 1 , 1 February 1994, Pages 73-97