A Retrospective Evaluation of COMPACT Predictions of the Outcome of NTP Rodent Carcinogenicity Testing

David F.V. Lewis, Costas Ioannides, and Dennis V. Parke
Molecular Toxicology Group, School of Biological Sciences, University of Surrey, Guildford, Surrey GU2 5XH, UK

The carcinogenic potentials of 40 National Toxicology Program chemicals previously predicted by Computer Optimised Molecular Parametric Analysis for Chemical Toxicity (COMPACT), based on the identification of potential substrates of cytochromes P4501A and 2E (CYP1A and CYP2E), have been compared with new rodent carcinogenicity results. The COMPACT predictions have also been compared with published Ames mutagenicity data and with our own Hazardexpert predictions for carcinogenicity. Concordance evaluations between rodent carcinogenicity (1/4 segments positive) and predictions by COMPACT or Hazardexpert were 64% for COMPACT (CYP1A only), 72% for COMPACT (CYP1A plus CYP2E), 70% for Hazardexpert alone, and 86% for COMPACT (CYP1A plus CYP2E) plus Hazadoxpert. Sensitivities of the predictions were for COMPACT, 75%; Hazardexpert, 60%; and Ames, 54%. Positive predictivities were for COMPACT, 75%; Hazardexpert, 78%; and Ames 81%. Negative predictivities were for COMPACT, 62%; Hazardexpert, 52%; and Ames, 42%. Key words: Ames test, carcinogenicity prediction, Computer Optimised Molecular Parametric Analysis for Chemical Toxicity, Hazardexpert, mutagenicity. Environ Health Perspect 103:178–184 (1995)

Computer-Optimized Molecular Parametric Analysis of Chemical Toxicity (COMPACT) is a procedure for the rapid identification of potential carcinogenicity and toxicity mediated by one or more of the cytochromes P450, a superfamily of enzymes involved in the Phase 1 metabolism of the majority of endogenous and exogenous chemicals (1–8). It has been established that many carcinogens act as substrates and/or inducers for cytochromes P4501 (9,10) and P4502E (11,12), which give rise to the formation of reactive intermediates capable of interaction with DNA, causing mutagenesis, neoplasia, cell proliferation, and malignancy (12,13). Inducers of cytochrome P4501, following interaction with the Ah (the aromatic hydrocarbon, or TCDD) receptor and translocation into the nucleus, bring about de novo synthesis of P4501 and also initiate the protein kinase C cascade, causing cell proliferation and malignancy (12). Activation by P4502E gives rise to the formation of reactive intermediates and is involved in the generation of reactive oxygen species (ROS) (14), which have been implicated in a variety of toxic effects (15) including cancer (initiation and promotion), especially in small rodents which have particular sensitivity to oxygen radical toxicity (11,12). Chemicals that act as substrates for other cytochromes P450 are, in general, detoxified by these enzymes to polar metabolites which are readily conjugated and eliminated (10).

From a consideration of their molecular and electronic structures, chemicals may be shown to exhibit specificity for one or more of these cytochromes P450, and, together with information relating to the enzymes’ active site dimensions, it is possible to accurately predict the P450 specificity of a test compound and its toxic outcome. To predict the P450 isoform, it is necessary to calculate certain molecular and electronic structural characteristics, the overall degree of planarity (area/depth), the molecular size (collision diameter), and the likelihood of activation by oxygenation, expressed as ΔE, the energy difference between the lowest unoccupied and the highest occupied molecular orbitals [(ΔE = E(LUMO) − E(HOMO)); LUMO, lowest unoccupied molecular orbit; HOMO, highest occupied molecular orbit]. Graphical analysis of this information and comparison with training set data for known cytochrome P450 substrates and inducers, including carcinogens and noncarcinogens, leads to an overall prediction of potential toxicity. This technique of COMPACT (16,17) has been validated against 100 NTP chemicals of known carcinogenicity and gave a concordance of 92% based on specificity for P4501 and P4502E (17). The COMPACT evaluation of the 40 NTP chemicals in the present study (18) were compared with our own Hazardexpert evaluations of the same chemicals and with rodent carcinogenicity assays and Ames test results previously published (19,20).

Methods

Molecular structural and electronic parameters for COMPACT analysis of the 40 NTP chemicals were determined as previously described (2,17,18). Criteria for P450 substrates were a high degree of molecular planarity (βθ/α2) and a high potential for activation by oxidative metabolism (ΔE) combined in the COMPACT ratio [CR = (βθ/α2)/(ΔE – 8)]. In a plot of α/β against ΔE, those cytochrome P4501 (CYP1) substrates with high values of α/β and low values of ΔE (<1.5 eV) are the most potent of known carcinogens (1,9,10). The use of the term ΔE – 8 instead of ΔE arises because 8 is approximately the lowest value of ΔE to be obtained and is characteristic of a high potential for activation by oxidative metabolism. P4501 substrates are characterized by values of CR >0.5, and this is the criterion for the prediction (C5) of carcinogenesis from P4501 activation.

The criterion for cytochrome P4502E (CYP2E) substrates is a small value for the molecular collision diameter (<6.5 Å). A low value of ΔE was also used as a criterion because this gives an indication of the ease of activation of the substrate. P4502E substrates are believed to be carcinogenic by 1) their propensity to stabilize P4502E and to generate ROS, which are themselves carcinogenic per se, and 2) their tendency to undergo metabolic activation to reactive intermediates, which are proximate or ultimate carcinogens (12,17). Hence the criteria for prediction of carcinogen from P4502E activation (C2E) are values of molecular diameter <6.5 Å and ΔE <15eV. COMPACT predictions of carcinogenicity are made from the affinity for P4501 (C5), affinity for P4502E (C2E), and if either or both give a positive response, the overall response (C5) is considered to be positive.

Data for the prediction of metal toxicity resulting from compounds 7, 16 and 35 have been obtained from standard metal ion redox potentials and interpreted according to a previously generated quantitative structure–activity relationship for metal toxicities involving this parameter (3). It has been shown that the acute toxicities of about 30 metal compounds after intraperitoneal dosage to the mouse correlate (R = 0.86) with the magnitude of the metal ion redox potential. Hence, the respective redox potentials are included for the three metal compounds, from which an estimate of their potential carcinogenicity has been made.

Ames mutagenicity and rodent carcinogenicity data are from Ashby and Tennant (19,20). One positive response in the four segments of the rodent assay is taken as an overall “+” response. When this single response is EE (equivocal evi-
The outcome of COMPACT analysis on 40 of the 44 test chemicals is shown in Table 1 (18). COMPACT predictions could not be made for 4 of these chemicals, as three were inorganic (nos. 17, 16 and 35) and one was polymeric (no. 3).

Table 2 provides a comparison between the COMPACT results with those of the rodent two species bioassay, where it can be seen that the rodent data for 4 chemicals (nos. 12, 13, 18 and 34) remain to be completed. Thus, concordances with COMPACT and the NTP carcinogenicity results can be made for only 36 chemicals. From the point of view of identifying those chemicals that are direct-acting, rather than requiring metabolic activation, we have also conducted EPA structural alert identification for the 40 chemicals using the Hazardexpert system (21). This software, which is also rapid and inexpensive, runs on the personal computer and also calculates logP and pK values which are generally considered to be of potential value in prescreening for toxicity. The results of Hazardexpert determinations are presented in Table 3 and combined concordances with COMPACT predictions are shown in Table 4.

The results of COMPACT and Hazardexpert predictions are summarized in Tables 1–4, and these evaluations for each individual chemical are given in more detail to provide a rationale for concordances between COMPACT and Hazardexpert predictions and rodent assay results. The overall concordance between COMPACT (P4501 and P4502E) and rodent carcinogenicity is 72%, and the consideration of structural alert evidence, provided by Hazardexpert, improves the total agreement to 86%. p-Nitrophenol (chemical no. 9), identified as positive in both COMPACT and Hazardexpert, but which is noncarcinogenic in the rodent assay, is a polar compound, which is readily conjugated and eliminated. Chemicals regarded as negative by both COMPACT and Hazardexpert (nos. 11, 17 and 21), but positive in the rodent assay, were only weak or sex/species specific rodent carcinogens; two of these (nos. 17 and 21) could be teratogenic via direct interaction with DNA. These three chemicals were all correctly identified as positive carcinogens by the use of combined data from structure alert, chronic toxicity studies, and the Ames test (22).

If the metal ion redox potentials for the inorganic compounds are included with the COMPACT and Hazardexpert data, chemical no. 16 has a high redox potential indicative of rodent carcinogenicity, which enhances the concordance between predicted and observed carcinogenicity to (32/37) 86%.

### Evaluation of Individual Chemicals

**1. Amphetamine.** This compound is negative for P4501 in COMPACT. With a molecular diameter of 6.5 and the high value of $\Delta E$, there is little likelihood of its interaction with CYP2E with the generation of oxygen radicals. Amphetamine is negative in all segments of the rodent assay, hence, there is concordance with COMPACT.

### Naphthalene.

Naphthalene is positive in COMPACT for both CYP1 and CYP2E, but since naphthalene has only two fused aromatic rings and does not contain a bay region, metabolic activation is weak. The compound forms an epoxide that is cytotoxic, but it is readily conjugated with glutathione and is hydrolyzed by epoxide hydrolase (23). There is some evidence from the rodent assay of carcinoge-
Table 2. Concordance between COMPACT, Ames mutagenicity, and rodent carcinogenicity

| Compound no. | Ames mutagenicity | Rodent carcinogenicity | Concordance | COMPACT v. COMPACT v. Amess | Ames |
|--------------|-------------------|------------------------|-------------|-----------------------------|------|
| MR/FR        | MM/FM             | Overall                | COMPACT     |                             |      |
| 1            | ?                 | NE/NE                  | NE/NE      | 1                           | NT   |
| 2            | –                 | NT/SE                  | NE/SE      | 0                           | 1    |
| 3            | –                 | EE/NE                  | NE/NE      | 0                           | 0    |
| 4            | –                 | NE/NE                  | NE/NE      | 0                           | 0    |
| 5            | –                 | NE/NE                  | NE/NE      | 0                           | 0    |
| 6            | –                 | NE/EE                  | EE/EE      | 0                           | 0    |
| 7            | –                 | NE/EE                  | EE/EE      | 0                           | 0    |
| 8            | –                 | NT/NE                  | NT/NE      | 0                           | 0    |
| 9            | –                 | NE/NE                  | NE/NE      | 0                           | 0    |
| 10           | –                 | SE/NE                  | SE/NE      | 1                           | 1    |
| 11           | –                 | NT/NE                  | NT/NE      | 1                           | 1    |
| 12           | –                 | INC                    | INC        | 1                           | 1    |
| 13           | –                 | INC                    | INC        | 1                           | 1    |
| 14           | –                 | SE/NE                  | SE/NE      | 1                           | 1    |
| 15           | –                 | NT/NE                  | NT/NE      | 1                           | 1    |
| 16           | –                 | SE/EE                  | EE/EE      | 0                           | 0    |
| 17           | –                 | NE/SE                  | NE/SE      | 0                           | 0    |
| 18           | –                 | INC                    | INC        | 1                           | 1    |
| 19           | –                 | NE/NE                  | NE/NE      | 1                           | 1    |
| 20           | –                 | EE/NE                  | EE/NE      | 1                           | 1    |
| 21           | –                 | EE/NE                  | EE/NE      | 1                           | 1    |
| 22           | –                 | SE/EE                  | SE/EE      | 1                           | 1    |
| 23           | –                 | NT/NE                  | NT/NE      | 1                           | 1    |
| 24           | –                 | NE/NE                  | NE/NE      | 1                           | 1    |
| 25           | –                 | NT/NE                  | NT/NE      | 1                           | 1    |
| 26           | –                 | SE/NE                  | SE/NE      | 1                           | 1    |
| 27           | –                 | NT/NE                  | NT/NE      | 1                           | 1    |
| 28           | –                 | CE/CE                  | CE/CE      | 1                           | 1    |
| 29           | –                 | SE/CE                  | SE/CE      | 1                           | 1    |
| 30           | –                 | SE/SE                  | SE/SE      | 1                           | 1    |
| 31*          | –                 | EE/NE                  | EE/NE      | 1                           | 1    |
| 32           | –                 | NE/SE                  | NE/SE      | 1                           | 1    |
| 33*          | –                 | EE/NE                  | EE/NE      | 1                           | 1    |
| 34           | –                 | INC                    | INC        | 1                           | 1    |
| 35           | –                 | EE/EE                  | EE/EE      | 1                           | 1    |
| 36           | –                 | CE/CE                  | CE/CE      | 1                           | 1    |
| 37           | –                 | CE/CE                  | CE/CE      | 1                           | 1    |
| 38           | –                 | SE/CE                  | SE/CE      | 1                           | 1    |
| 39           | –                 | CE/CE                  | CE/CE      | 1                           | 1    |
| 40           | –                 | CE/NT                  | CE/NT      | 1                           | 1    |
| 41           | –                 | EE/NE                  | EE/NE      | 1                           | 1    |
| 42           | –                 | NT/EE                  | NT/EE      | 1                           | 1    |
| 43           | –                 | CE/CE                  | CE/CE      | 1                           | 1    |
| 44           | –                 | CE/CE                  | CE/CE      | 1                           | 1    |

*Mutagenicity, Salmonella assay with without activation; carcinogenicity, rodent assay; MR, male rat; FR, female rat; MM, male mouse; FM, female mouse; 1 positive out of 4 segments of the rodent assay is taken as +; CE (clear evidence) and SE (some evidence) are regarded as +. EE (equivocal evidence) is regarded as +/+; NE (no evidence) is regarded as –; INC, inconclusive. COMPACT, COMPACT predictions from Table 1; concordance: 1 for correct, 0 for incorrect, NT, not tested. Mutagenicity and rodent carcinogenicity data from Ashby and Tannant (19.28).

†Compounds 31, 33, and 42 are regarded as weak positives/equivocal positives based on pathology reports.


genicity in the female mouse, but no data are available for either sex of rat. Concordance between COMPACT and carcinogenicity is thus considered to be positive.

3. Polysorbate 80. This compound was not examined by COMPACT due to its polymeric nature.

4. Promethazine. Promethazine is positive in COMPACT, primarily due to its low ΔE value; it does not exhibit high planarity due to a bulky side chain and the combination of sulfur and nitrogen heteroatoms in the ring system, which becomes V-shaped. Consequently, it is unlikely to be a good P4501 substrate or inducer. Promethazine is negative in all segments of the rodent assay; consequently, there is no concordance with COMPACT.

5. Resorcinol. This compound is not a likely P4501 substrate (CR = 0.70) because it has a relatively high ΔE value (15.7) and is readily conjugated. Although the COMPACT ratio would suggest that it would be equivocally positive, in the original graphical analysis it was clearly negative (18). Since the molecular diameter of resorcinol is 5.6, and its ΔE is 15.7, the COMPACT prediction for CYP2E1 is also negative. Resorcinol is negative in the rodent assay, so the COMPACT analysis is in concordance with the biological data. Hazardexpert predicts that this compound would be mutagenic.

6. 1-Butanol. This compound is negative in COMPACT for P4501, and although it has a diameter characteristic of a P4502E substrate, its high ΔE value mitigates against ROS production and carcinogenicity. A negative finding in the rodent assay, with the exception of equivocal evidence for carcinogenicity in the male mouse, accords with the COMPACT prediction, but Hazardexpert predicts mutagenicity.

7. Manganese sulfate. The relatively low redox potential of manganese (Mn/Mn²⁺ = -1.18 V) implies that this compound is unlikely to elicit overt rodent toxicity. This may explain the low incidence of carcinogenicity, which is seen only in mice.

8. Chloroacetic acid. This compound is negative in COMPACT for both P4501 and P4502E specificity, and this negative COMPACT prediction is in accordance with the rodent study.

9. p-Nitrophenol. The high molecular planarity of this compound predicts a positive for P4501, and the low diameter, coupled with a relatively low ΔE, clearly predicts a positive for P4502E. This nitrophenol is known to be an excellent substrate for UDP glucuronyl transferase (24) and for P4502E, so it is readily conjugated with glucuronic acid, readily hydroxylated, and eliminated due to its high polarity. These factors may explain the negative result in the rodent assay, although the compound was only tested in the mouse. Concordance with COMPACT is therefore negative. However, the Hazardexpert prediction is that the chemical would be a weak carcinogen.

10. Triresyl phosphate. This chemical is predicted positive in COMPACT for P4501. Since the rodent tests are all negative, there is no concordance between the biological assay and COMPACT prediction. Hazardexpert predicts that the compound would be a weak carcinogen.

11. o-Benzyl-p-chlorophenol. The COMPACT prediction for this chemical is negative for both C₁₀ and C₁₂. Evidence of carcinogenicity in 1/4 segments of the rodent assay indicates lack of concordance with COMPACT prediction. Hazardexpert also predicts mutagenicity.

12. 2,2-(Bromomethyl)-1,3-propanediol. This chemical is equivocal for both P4501 and P4502E in the COMPACT prediction and is therefore rated a weak positive. The rodent tests are incomplete, so concordance cannot be evaluated. Hazardexpert predicts teratogenicity but not carcinogenicity.

13. t-Butanol. COMPACT is negative for P4501 specificity, and although the
Table 3. Hazardexpert predictions for NTP chemicals

| Chemical                       | LogP | Structure alert | Prediction of carcinogenicity* |
|--------------------------------|------|-----------------|-------------------------------|
| 1. Amphetamine                 | 1.97 | None            | −                             |
| 2. Naphthalene                 | 3.30 | None            | −                             |
| 3. Polysorbate 80              | 3.87 | None            | −                             |
| 4. Promethazine                | 3.97 | None            | −                             |
| 5. Rescorcinol                 | 0.97 | Phenolic        | −                             |
| 6. γ-Butyro lactone            | 0.27 | Lactone         | −                             |
| 7. Manganese sulphate          | -0.37| Nitrosoester    | −                             |
| 8. Chloroacetic acid           | 1.24 | Nitroaromatic   | +                             |
| 9. p-Nitrophenol               | 1.72 | Carboxyl        | −                             |
| 10. Tricresyl phosphate        | 17.32| Organophosphat  | +                             |
| 11. o-Benzyl- p-chlorophenol   | 4.41 | Phenolic        | −                             |
| 12. 2,2-Bis(bromomethyl)ethyl  | 0.21 | Bromoethyl      | −                             |
| 13. 1,3-Propanediol            | 0.78 | Hydroxy methyl  | −                             |
| 14. 3,4-Dihydroxocoumarin      | 1.60 | Lactone         | −                             |
| 15. Ethylene glycol            | -1.94| Hydroxy methyl  | −                             |
| 16. Mercuric chloride          | -0.20| None            | −                             |
| 17. Methyl phenidate           | 2.17 | Phenolic        | −                             |
| 18. Theophylline               | 0.67 | Noncarcinogenic | −                             |
| 19. 4,4'-Thiobiis-(6-t-butyl-m-cresol) | 1.80 | Phenolic        | −                             |
| 20. Triamterene                | -1.42| Amino polyaromatic | +                             |
| 21. Diphenylhydantoin          | 3.33 | None            | −                             |
| 22. Pentachloroanisole         | 5.79 | Polyhaloaromatic | +                             |
| 23. Chloramine                 | 1.53 | None            | −                             |
| 24. 4,4'-Diamino-2,2'-stilbene disulfonic acid | 2.68 | Aromatic amine | −                             |

NP, not predicted; NT, not tested.

*Toxicity predictions: + = carcinogenic, − = noncarcinogenic.

This chemical indicates a P4502E substrate, ROS generation is unlikely as ΔΔ is too high. The chemical is known to exhibit specificity for P4502E, which confirms the computer prediction, but COMPACT predictions for both C1 and C2E are negative. As the rodent data are incomplete, concordance cannot be evaluated. Hazardexpert predicts teratogenicity but not carcinogenicity.

14. 3,4-Dihydroxocoumarin. The compound is not likely to be P4501 selective (CR = 0.53) but is positive for P4502E. It could also result in ring-opening after being metabolized, giving rise to an entity with potential for peroxisome proliferation. The overall COMPACT prediction (C1) is positive, which accords with the rodent assay where there is evidence for carcinogenicity in 2/4 segments.

15. Ethylene glycol. This chemical is negative for both P4501 and P4502E, which is in accord with the negative mouse data; no data for rat are yet available. Hazardexpert predicts teratogenicity but not carcinogenicity.

16. Mercuric chloride. The high redox potential (Hg/Hg2+ = +0.62 V) predicts that this chemical would be positive in the rodent assay, which agrees with the established biological data.

17. Methyl phenidate. COMPACT predictions for this chemical are negative for both P4501 and P4502E. Rodent data show evidence of carcinogenicity in 0/4 segments in the mouse. Thus, there is no concordance between COMPACT and the rodent assay. Hazardexpert predicts that this chemical would be teratogenic but not carcinogenic.

18. Theophylline. The chemical is positive in COMPACT for both P4501 and P4502E and is known to exhibit specificity for P4501. However, because the rodent tests are incomplete, it is not possible to establish concordance at this stage.

19. 4,4'-Thiobiis-(6-t-butyl-m-cresol). This compound is borderline in COMPACT for P4501 and negative for P4502E and is therefore considered unlikely to be a carcinogen. In the rodent assay, this compound is negative, which is in concordance with the COMPACT prediction.

20. Triamterene. COMPACT clearly predicts a positive for P4501, and N-hydroxylation by P4502A2 would give rise to a nitrenium ion that could interact irreversibly with DNA. The rodent assay agrees with this for both segments of the mouse data, although the rat study is equivocal. It is considered that COMPACT prediction is in concordance with the biological assay for this compound. Hazardexpert also predicts carcinogenicity.

21. Diphenylhydantoin (phenytoin). Phenytion is negative for both P4501 and P4502E in COMPACT, and there is substantial biochemical evidence that it is metabolized by P4502B and C. Phenytion is a known teratogen, but predictions for carcinogenicity are negative for both COMPACT and Hazardexpert. There is evidence of carcinogenicity in 1/4 segments (female mouse) of the rodent assay. Therefore, there is no concordance between the COMPACT prediction and the findings in the rodent assay.

22. Pentachloroanisole. Pentachloroanisole is positive in COMPACT for P4501 only and may interact with the Ah receptor (involved in the induction of P4501A) because there is some degree of structural similarity with TCDD. This prediction of carcinogenicity accords with the positive results obtained in the rodent assay. Hazardexpert predicts weak carcinogenicity.

23. Chloramine. Chloramine is positive for P4501 in COMPACT and was hence predicted positive. Evidence for carcinogenicity in the rodent assay is largely negative, therefore no concordance is shown with the COMPACT prediction. However, this chemical is known to undergo rapid decomposition in solution and to be readily eliminated due to its polarity.

24. 4,4'-Diamino-2,2'-stilbenedisulfonic acid. This chemical is positive for P4501 in COMPACT, but it is probably readily eliminated due to the two sulfonic acid groups, which are both polar and ionicized at physiological pH. The rodent carcinogenicity data are negative, so there is no concordance with the COMPACT pre-
Table 4. Concordance of rodent carcinogenesis bioassay with COMPACT plus Hazardexpert

| Compound                          | Rodent carcinogenesis assay | COMPACT (C1) prediction | Hazardexpert prediction | Combined concordance* |
|----------------------------------|----------------------------|-------------------------|-------------------------|-----------------------|
| 1. Amphetamine                   | -                         | -                       | -                       | 1                     |
| 2. Naphtalene                    | +                         | +                       | -                       | 1                     |
| 3. Polysorbate 80                | -                         | NT                      | NT                      | NT                    |
| 4. Promethazine                   | -                         | -                       | -                       | 0                     |
| 5. Resorcinol                    | -                         | -                       | -                       | 0                     |
| 6. β-Butyrolactone               | -                         | -                       | -                       | 1                     |
| 7. Manganese sulphate            | -                         | NT                      | NT                      | NT                    |
| 8. Chloroacetic acid             | -                         | -                       | -                       | 0                     |
| 9. p-Nitrophenol                 | +                         | -                       | +                       | 0                     |
| 10. Tricresyl phosphate          | -                         | +                       | -                       | 1                     |
| 11. o-Benzyl-p-chlorophenol      | +                         | -                       | -                       | 0                     |
| 12. 2,2-Bis(bromomethyl)-benzoic acid | NT                     | +                       | -                       | NT                    |
| 13. 1,3-propanediol              | NT                        | -                       | -                       | NT                    |
| 14. 3,4-Dihydroxyacetanilide     | -                         | -                       | -                       | 0                     |
| 15. Ethylene glycol              | -                         | NT                      | NT                      | NT                    |
| 16. Mercuric chloride            | +                         | NT                      | NT                      | NT                    |
| 17. Methylenebiguanide           | +                         | -                       | -                       | 0                     |
| 18. Theophylline                 | NT                        | -                       | -                       | NT                    |
| 19. 4,4'-Thiodiis(6-t-butyl-m-cresol) | -                     | +                       | +                       | NT                    |
| 20. Triamterene                  | +                         | -                       | +                       | NT                    |
| 21. Diphenylhydantoin            | +                         | -                       | -                       | NT                    |
| 22. Pentachloroanisole           | -                         | -                       | +                       | 1                     |
| 23. Chloramine                   | +                         | +                       | +                       | 1                     |
| 24. 4,4'-Diamino-2,2'-stilbene- disulfonic acid | -                 | +                       | +                       | 0                     |
| 25. Methyl bromide               | +                         | +                       | -                       | 1                     |
| 26. p-Nitrobenzoic acid          | +                         | -                       | -                       | NT                    |
| 27. Sodium azide                 | -                         | -                       | -                       | 1                     |
| 28. Tris(2-chloroethyl)phosphate | +                         | +                       | +                       | 1                     |
| 29. CI Direct blue 218           | +                         | +                       | +                       | NT                    |
| 30. CI Pigment red 3             | +                         | +                       | +                       | NT                    |
| 31. CI Pigment red 23            | +                         | +                       | -                       | NT                    |
| 32. 2,4-Diamino phenol           | +                         | +                       | +                       | 0                     |
| 33. 4-Hydroxyacetanilide         | +                         | +                       | +                       | 1                     |
| 34. Salicylazosulfapyridine      | +                         | NT                      | NT                      | NT                    |
| 35. Titanocene dichloride        | +                         | NT                      | NT                      | NT                    |
| 36. CI Acid red 114              | +                         | +                       | 1                       |                       |
| 37. CI Direct blue 15            | +                         | +                       | -                       | 1                     |
| 38. Coumarin                     | +                         | +                       | +                       | 1                     |
| 39. 2,3-Diamino-1-propanol       | +                         | +                       | +                       | 1                     |
| 40. 3,3'-Dimethylenzidine        | +                         | +                       | +                       | 1                     |
| 41. HC Yellow 4                  | -                         | -                       | -                       | 1                     |
| 42. p-Nitroaniline               | +                         | +                       | -                       | NT                    |
| 43. o-Nitroanisole               | +                         | +                       | +                       | 1                     |
| 44. 1,2,3-Trichloropropene       | +                         | +                       | +                       | 1                     |

*Rodent overall carcinogenicity data and overall COMPACT prediction (C1 = C1P + C1R) are from Tables 1 and 2.

Hazardexpert predicts teratogenicity but not carcinogenicity.

25. Methyl bromide. COMPACT predicts P4502E specificity for this compound, and there is biochemical evidence to support this prediction. The mouse bioassay shows no carcinogenicity, no data are available for rat. The lack of P4501 specificity accords with the rodent assay, but there is no concordance between the prediction of P4502E activity by COMPACT and lack of evidence of carcinogenicity in the mouse bioassay. Hazardexpert predicts mutagenicity.

26. p-Nitrobenzoic acid. p-Nitrobenzoic acid is positive in COMPACT for P4501 and for P4502E but is rapidly metabolized by reduction and conjugation, probably resulting in little carcinogenicity. The rodent assay shows some evidence of carcinogenicity in female rat only (1/4 segments), which is in concordance with the COMPACT prediction. Hazardexpert predicts strong carcinogenicity. The p-carboxylic group, being electron withdrawing, will stabilize the nitrenium ion formed during metabolism of the nitro group, which contrasts with p-nitrophenol (no. 9), where the p-hydroxyl group is electron donating, and destabilizes the nitrenium ion so that carcinogenicity does not result.

27. Sodium azide. Sodium azide is a direct-acting mutagen. The related compound, hydrazoic acid, was negative for P4501 in COMPACT. Although having the structural criteria for P4502E, hydrazoic acid is unlikely to generate ROS (ΔE = 16.2). Thus, the predictions C1 and C2 for COMPACT are negative, in concordance with the rat study which showed no carcinogenicity.

28. Tris(2-chloroethyl)phosphate. COMPACT is negative for both P4501 and P4502E, but it is probably direct-acting via the 2-chloroethyl (mustard) group. Clear evidence of carcinogenicity in the mouse means that there is no concordance with COMPACT. Hazardexpert predicts a weak carcinogen.

29. CI Direct blue 218. This chemical is positive for P4501 in COMPACT and is likely to undergo microbial azo reduction to form planar (carcinogenic) metabolites. The COMPACT prediction is in agreement with the rodent assay, which was clearly positive. Hazardexpert predicts marked carcinogenicity.

30. CI Pigment red 3. Another azo compound like the previous chemical, CI pigment red 3 is similarly predicted by COMPACT to be positive for P4501 and therefore carcinogenic. This prediction of carcinogenicity by COMPACT is in agreement with the rodent assay which shows carcinogenicity in 3/4 segments. Hazardexpert predicts strong carcinogenicity.

31. CI Pigment red 23. This azo dye is also P4501 positive in COMPACT, but there is equivocal evidence of carcinogenicity in only 1/4 segments in the rodent assay, probably due to the polarity of the compound and its ease of detoxication. However, following further details of the rodent pathology, this chemical is considered to be a weak positive carcinogen. COMPACT prediction is therefore in concordance with the rodent assay. Hazardexpert predicts strong carcinogenicity.

32. 2,4-Diaminophenol. 2,4-Diaminophenol is positive for both P4501 and P4502E in COMPACT and therefore is predicted to be a carcinogen. In the rodent assay there is evidence of carcinogenicity in only 1/4 segments, probably due to extensive detoxication by metabolic conjugation. Therefore, there is concordance between the COMPACT prediction and the biological assay. Hazardexpert predicts teratogenicity but not carcinogenicity.

33. 4-Hydroxyacetanilide (aceta- minophen). The compound is weakly positive for P4501 and P4502E in COMPACT, which is in agreement with experimental evidence. However, rodent evidence of carcinogenicity is only weakly positive, probably because of metabolic detoxication. Concordance between COMPACT and the rodent assay is regarded as positive. Moreover, the positive
findings for P450I and P450E in COMPACT are in agreement with the known lethal hepatotoxicity and nephrotoxicity.

34. Salicylazosulfapyridine. This compound is positive for P450I in COMPACT, and is consequently predicted to be a carcinogen. The rodent carcinogenicity tests are complete, so it is not possible to confirm this result, but this drug is known to be cytotoxic and to cause male infertility in humans (25). Hazardexpert predicts high carcinogenicity.

35. Titanocene dichloride. Quantitative structure–activity relationship work suggests that, on the basis of redox potential (Ti/Ti$^{2+}$ - 1.63 v), the compound is unlikely to show overt rodent toxicity. Rodent carcinogenicity is equivocal.

36. CI Acid red 114. COMPACT gives a clear indication for P450I specificity and a prediction of carcinogenicity. However, because of its very high polarity (4 sulfonic acid groups and 2 phenolic groups), carcinogenicity would not be exhibited by the parent compound. Azo reduction leads to the formation of 3,3′-dimethylbenzidine, which is a known carcinogen and is a P450I substrate by COMPACT (see compound no. 40). Hence, this COMPACT prediction of carcinogenicity agrees with the clear evidence of carcinogenicity in the rodent assay. Hazardexpert predicts a strong carcinogen.

37. CI Direct blue 15. This chemical is structurally related to the previous compound and is also positive in the P450I COMPACT. As with CI Acid red 114, carcinogenicity is likely to be mediated by 3,3′-dimethoxybenzidine, the product of azo reduction cleavage. Clear evidence of carcinogenicity in the rodent assay shows concordance with the COMPACT prediction. Hazardexpert predicts a strong carcinogen.

38. Coumarin. Coumarin has long been suspected of being a carcinogen. COMPACT predicts carcinogenicity from the specificity of coumarin for P450I, which is in agreement with the carcinogenicity seen in 3/4 segments of the rodent assay. Hazardexpert predicts carcinogenicity.

39. 2,3-Dibromo-1-propanol. Although negative for P450I, COMPACT shows clear evidence for P450E specificity and oxygen radical generation. It is likely, therefore, to be a rodent carcinogen; this is in complete agreement with the carcinogenicity seen in 4/4 segments of the rodent assay in accordance with COMPACT prediction. Hazardexpert predicts carcinogenicity.

40. 3,3′-Dimethylbenzidine. 3,3′-Dimethylbenzidine is positive for P450I specificity in COMPACT and was thus predicted to be carcinogenic. This is in complete agreement with the carcinogenicity seen in the rat assay; no data are available for mouse. Hazardexpert predicts teratogenicity but not carcinogenicity.

41. HC Yellow 4. This compound was originally weakly positive for P450I in COMPACT, but the chemical structure was incorrectly given. A reevaluation of the correct structure revealed that it was negative in COMPACT, which is in agreement with the rodent assay that was equivocal in one segment only and negative in the remaining three. Hazardexpert predicts strong carcinogenicity.

42. p-Nitroaniline. p-Nitroaniline is positive for P450I and probably P450E, and the COMPACT prediction is that it is carcinogenic. This compound was not evaluated for carcinogenicity in the rat and gave an equivocal result in the mouse. Further detail on the pathology indicates that it should be regarded as weakly positive, indicating concordance with the COMPACT prediction. Hazardexpert predicts strong carcinogenicity.

43. o-Nitroanisole. o-Nitroanisole is positive for P450I and also P450E in COMPACT, which predicts its carcinogenicity. Clear evidence of carcinogenicity in three segments of the rodent assay, and some evidence in the fourth, is in agreement with the COMPACT prediction. Hazardexpert predicts strong carcinogenicity.

44. 1,2,3-Trichloropropane. Although negative for P450I specificity, this chemical is clearly a P450E substrate with likely activation to form ROS. This indicates potential rodent carcinogenicity, which is in agreement with the rodent assay that showed carcinogenicity in all four segments. There is therefore positive concordance between the biological assay and the COMPACT prediction. Hazardexpert predicts strong carcinogenicity.

In summary, the concordance of COMPACT (C) predictions with the rodent carcinogenicity assay is 72%, and for COMPACT (C) plus Hazardexpert, the concordance is 86%.

Discussion

At the National Institute of Environmental Health Sciences international workshop, Predicting Chemical Carcinogenesis in Rodents, held in May 1993, the importance of metabolism and the need for an understanding of the mechanisms of carcinogenicity were emphasized. The key roles of certain cytochromes P450 in the metabolic activation of many chemical carcinogens is well established (22,26), and these have been addressed in the COMPACT study. The COMPACT techniques are being further developed to take account of more recent developments in the understanding of the mechanisms of toxicity and carcinogenicity. One of the limitations of COMPACT, that of identifying all direct-acting carcinogens, can be circumvented by using a structural alert identifier, such as Hazardexpert. Structural alert procedures will also identify those few carcinogenic chemicals that do not appear to depend on the cytochromes P450 for activation.

The use of COMPACT parameters (ad$^{2}$, AE, molecular diameter) in defining P450 specificity and P450-mediated toxicity is supported by quantitative structure–activity relationship evidence for many series of compounds, showing high correlations between one or more COMPACT parameters, sometimes in combination with other structural descriptors, and biological activity such as carcinogenicity, mutagenicity, P450 induction, and P450-mediated metabolism (2,3). In particular, frontier orbital energies have been found to correlate well with many biological endpoints related to toxicity. The $\Delta E$ parameter, which is an additive combination of the two frontier orbital energies, $E$(HOMO) and $E$(LUMO), is a measure of the energy required to promote an electron from the highest occupied molecular orbital to the lowest unoccupied one, i.e., $\Delta E = E$(LUMO) – $E$(HOMO). As far as P450-mediated reactions are concerned, $\Delta E$ could be a measure of the degree of interaction between substrate and enzyme and/or the ability of the substrate to become oxygenated by the P450. Furthermore, $\Delta E$ correlates well with the electron uptake rate constant, $k_{c}$, for 26 structurally diverse chemicals ($R = 0.82$). Therefore, there is some degree of complementarity between COMPACT and the $k_{c}$ test. The $k_{c}$ test has been shown by others to be a valuable physical method for the identification of chemical carcinogens (27). The two parameters $\Delta E$ and $k_{c}$ probably predict the likelihood of interaction between the test chemical or its metabolite and DNA; this is generally regarded as being an electrophilic interaction.

Carcinogenic effects mediated by interaction with cellular receptors, such as the Ah receptor, the peroxisome proliferator-activated receptor (ppar) and the estrogen receptor (ER), may also be predicted. We have modeled the latter two receptors (28) from their amino acid sequences, enabling the structural basis of ligand binding interactions to be explored. Although the Ah receptor has not yet been sequenced, COMPACT is able to identify its potential ligands because they are inducers of cytochrome P450I and bear molecular characteristics common to P450I substrates. Moreover, molecular modeling of entire P450 proteins associated with xenobiotic metabolism has been undertaken, based on sequence homology with the bacterial form of P450, which is of known crystal structure (29). The putative active
sites of these enzymes are considerably complementary to their specific substrates, and such studies will inevitably lead to further developments in COMPACT.

In conclusion, this study has shown that original COMPACT, based solely on CYP1 specificity, is 64% predictive of rodent carcinogenicity in the 36 chemicals tested. COMPACT, that includes CYP2E specificity in addition to CYP1 specificity improves the concordance to 72%. This takes into account the possibility of carcinogenicity produced directly from the generation of reactive oxygen species (ROS) mediated by P4502E and the formation of reactive intermediates from the activation of the chemicals by these ROS (12). The sensitivity of COMPACT (CYP1 and CYP2E) is 75%, and its positive predictive value is 75% (Table 5). Furthermore, structural criteria for other P450 substrate specificity, and for associated receptors, should increase the scope of the technique. However, it would be advantageous to combine COMPACT with a structural alert program such as Hazardexpert, especially as this gives the added advantage of calculating logP and pKa values. We have shown that such a combination increases concordance with the rodent assay from 72% to 86% (Table 5). Four out of five of those chemicals that were incorrectly assigned by COMPACT in combination with Hazardexpert were predicted successfully by a combination of other methods (mutagenicity, structural alert, rodent chronic toxicity data) (22), which indicates that a battery of short term tests could provide a favorable comparison with rodent carcinogenicity (19,20,30–32).

REFERENCES

1. Ioannides C, Lewis DFV, Parke DV. Computer modelling in predicting carcinogenicity. Eur J Cancer Prev 2:275–282 (1993).

2. Lewis DFV. Computer modelling of cytochromes P-450 and their substrates: a rational approach to the prediction of carcinogenicity. Frontiers Biotransform 7:90–136 (1992).

3. Lewis DFV. Computer-assisted methods in the evaluation of chemical toxicity. Rev Comput Chem 3:173–222 (1992).

4. Lewis DFV, Ioannides C, Parke DV. Molecular dimensions of the substrate binding site of cytochrome P-448. Biochem Pharmacol 35:2179–2185 (1986).

5. Lewis DFV, Ioannides C, Parke DV. Structural requirements for substrates of cytochromes P-450 and P-448. Chem-Biol Interact 64:39–60 (1987).

6. Lewis DFV, Ioannides C, Parke DV. COMPACT: a form of discriminant analysis for the identification of potential carcinogens. In: QSAR rational approaches to the design of bioactive compounds (Silipo C, Vittoria A, eds.) Amsterdam:Elsevier, 1991;525–527.

7. Parke DV, Ioannides C. The cytochromes P448: a unique family of enzymes involved in chemical toxicity and carcinogenesis. Biochem Pharmacol 38:1313–1319 (1987).

8. Parke DV, Ioannides C, Lewis DFV. Computer modelling and other in vitro tests in the evaluation of chemicals—strategic applications. Toxicon 46:680–685 (1990).

9. Ioannides C, Parke DV. Induction of cytochrome P450I as an indicator of potential chemical carcinogenesis. Drug Metab Rev 25:485–501 (1993).

10. Parke DV, Ioannides C, Lewis DFV. The role of cytochromes P-450 in the detoxication and activation of drugs and other chemicals. Can J Physiol Pharmacol 69:537–549 (1991).

11. Parke DV, Ioannides C. Roles of cytochrome P-450 in mouse liver tumor production. In: Mouse liver carcinogenesis, mechanisms and species comparisons (Stevenson DE, McClain RM, Popp JA, Slaga TA, Ward JM, Pitar HC, eds.). New York:Alan R. Liss, 1990;215–230.

12. Parke DV. The cytochromes P450 and mechanisms of chemical carcinogenesis. Environ Health Perspect 102:852–853 (1994).

13. Parke DV. Activation mechanisms to chemical toxicity. Arch Toxicol 60:5–15 (1987).

14. Ekstroem G, Ingelman-Sundberg M. Rat liver microsomal NADPH-supported oxidase activity and lipid peroxidation dependent on ethanol inducible cytochrome P-450 (P450IIIE1). Biochem Pharmacol 38:1313–1319 (1989).

15. Halliwell B. Gutteridge JHC. Role of free radicals and catalytic metal ions in human diseases: an overview. Methods Enzymol 186:1–85 (1990).

16. Lewis DFV, Ioannides C, Parke DV. Prediction of chemical carcinogenicity from molecular and electronic structure: a comparison of MINDO/3 and CNDO/2 molecular orbital methods. Toxicol Lett 45:1–13 (1989).

17. Lewis DFV, Ioannides C, Parke DV. Validation of a novel molecular orbital approach (COMPACT) to the safety evaluation of chemicals by comparison with Salmonella mutagenity and rodent carcinogenicity data evaluated by the US NCI/NTP. Mutat Res 291:61–77 (1993).

18. Lewis DFV, Ioannides C, Parke DV. A prospective toxicity evaluation (COMPACT) on 40 chemicals currently being tested by the National Toxicology Program. Mutagenesis 5:433–435 (1990).

19. Ashby J, Tennant, RW. Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. Mutat Res 204:17–115 (1988).

20. Ashby J, Tennant RW. Definitive relationships among chemical structures, carcinogenicity and mutagenicity for 301 chemicals tested by the US National Toxicology Program. Mutat Res 257:229–306 (1991).

21. Smithing MP, Darvas F. Hazardexpert: an expert system for predicting chemical toxicity. In: Food safety assessment (Finley JW, Robinson SF, Armstrong DJ, eds). Washington, DC:American Chemical Society, 1992:191–200.

22. Tennant RW, Spaulding J, Stasiewicz S, Ashby J. Prediction of the outcome of rodent carcinogenicity bioassays currently being conducted on 44 chemicals by the National Toxicology Program. Mutagenesis 5:3–14 (1990).

23. Buonarati M, Morin D, Plopper C, Buckpitt A. Glutathione depletion and cytotoxicity by naphthalene 1,2-oxide in isolated hepatocytes. Chem Biol Interact 17:147–165 (1989).

24. Mackenzie PI. Structure and regulation of UDP-glucuronosyl transferases. Frontiers Biotransform 2:211–243 (1990).

25. Levi AJ, Fisher AM, Hughes L, Hendry WF. Male infertility due to sulfasalazine. Lancet 2:276–277 (1979).

26. Tagashira Y, Omta T, eds. P-450 and chemical carcinogenesis. New York:Plenum Press, 1985.

27. Bakale G, McCready RD. Response of k, test to NCI/NTP screened chemicals. Carcinogenesis 11:1811–1818 (1990).

28. Lewis DFV, Lake BG. The interaction of some peroxisome proliferators with the mouse liver peroxisome proliferator-activated receptor (ppar): a molecular modelling and quantitative structure-activity relationship study. Xenobiotica 23:79–96 (1993).

29. Lewis DFV, Moereels H. The sequence homologies of cytochromes P-450 and active site geometries. J Comput Aided Mol Design 6:235–252 (1992).

30. Richard AM. Application of SAR methods to non-congeneric databases associated with carcinogenicity and mutagenicity. Mutat Res 305:73–97 (1994).

31. Bahler D, Bristol DW. The induction of rules for predicting chemical carcinogenesis in rodents. In: Intelligent systems for molecular biology (Hunter L, Shavick J, Sears D, eds). Menlow Park, CA:MIT Press, 1993.