

- higher virulence [see also H. J. Bremmermann and J. Pickering, *J. Theor. Biol.* **100**, 411 (1983)].
22. The phylogenetic relationships among all of the species of nematodes and their host fig wasps are the subject of ongoing molecular studies and have not yet been completely resolved. However, the sequence divergence of the chloroplast *rbcl* gene suggests an ancient split among the different host fig subgenera. Mitochondrial sequence divergence among the corresponding pollinator wasp taxa reflect this split and further suggests that relatively greater levels of virulence have evolved independently several times [(23, 24); E. A. Herre *et al.*, unpublished data]. Following the implication that virulence evolves quickly with respect to speciation, and treating species pairs as independent, the Spearman rank correlation ( $R_s$ ) between weighted means of relative reproductive success of infected and uninfected mothers and proportion of single foundress broods in the host wasp species is 0.91 ( $P < 0.001$ ,  $n = 11$ ) (Table 1 and Fig. 1); with unweighted means,  $R_s = 0.90$  ( $P < 0.001$ ,  $n = 11$ ). With the relative number of female offspring,  $R_s = 0.88$  ( $P < 0.001$ ,  $n = 11$ ). Neither the size of the host wasp nor the size of the nematode species has a significant relation to virulence.
23. C. C. Berg, *Experientia* **45**, 605 (1989).
24. J. T. Wiebes, *Neth. J. Zool.* **32**, 395 (1982).
25. M. E. Collinson, in *Evolution, Systematics, and Fossil History of the Hamamelidae*, P. R. Crane and S. Blackmore, Eds. (Clarendon Press, Oxford, 1989), vol. 2, pp. 319–340.
26. W. D. Hamilton, in *Sexual Selection and Reproductive Competition in Insects*, M. S. Blum and N. A. Blum, Eds. (Academic Press, New York, 1979), p. 167.
27. I thank J. J. Van Alphen, J. G. Sevenster, G. O. Poinar, Jr., J. Aho, P. W. Ewald, S. J. Wright, E. G. Leigh, Jr., D. M. Windsor, R. B. Srygley, A. Read, C. A. Toft, E. F. Stockwell, M. J. Ryan, D. W. Schemske, L. E. Gilbert, G. S. Gilbert, M. J. Wade, M. Feldman, and E. Kalko for useful discussions during the development of this work. E. Jaen, A. Balderson, L. Papernik, A. Gomez, M. Lopez, and R. Mihalik helped with fieldwork. J. T. Wiebes and W. Ramirez identified the wasp species. T. Collins, L. Weigt, M. Carles, C. A. Machado, S. S. McCafferty, G. Zurawski, W. Van Houten, A. Graffen, and E. Bermingham helped with DNA sequencing and analysis. R. V. Cocharad was instrumental in developing the simulation model. This work was supported by two Smithsonian Post-Doctoral Fellowships, two Smithsonian Scholarly Studies Grants, The Smithsonian Tropical Research Institute Molecular Program, and a NATO Post-Doctoral Fellowship.

10 August 1992; accepted 30 November 1992

## Structure-Based Discovery of Inhibitors of Thymidylate Synthase

Brian K. Shoichet,\* Robert M. Stroud,† Daniel V. Santi, Irwin D. Kuntz,† Kathy M. Perry\*

A molecular docking computer program (DOCK) was used to screen the Fine Chemical Directory, a database of commercially available compounds, for molecules that are complementary to thymidylate synthase (TS), a chemotherapeutic target. Besides retrieving the substrate and several known inhibitors, DOCK proposed putative inhibitors previously unknown to bind to the enzyme. Three of these compounds inhibited *Lactobacillus casei* TS at submillimolar concentrations. One of these inhibitors, sulisobenzone, crystallized with TS in two configurations that differed from the DOCK-favored geometry: a counterion was bound in the substrate site, which resulted in a 6 to 9 angstrom displacement of the inhibitor. The structure of the complexes suggested another binding region in the active site that could be exploited. This region was probed with molecules sterically similar to sulisobenzone, which led to the identification of a family of phenolphthalein analogs that inhibit TS in the 1 to 30 micromolar range. These inhibitors do not resemble the substrates of the enzyme. A crystal structure of phenolphthalein with TS shows that it binds in the target site in a configuration that resembles the one suggested by DOCK.

Thymidylate synthase (TS) is a target for drugs against proliferative diseases and cancer because it is required for de novo synthesis of deoxythymidine monophosphate (dTMP) and, hence, for DNA production. Efforts have focused on the design of inhibitors similar to the substrate, deoxyuridine monophosphate (dUMP), or the cofactor, 5,10-methylenetetrahydrofolate ( $\text{CH}_2\text{-H}_4$  folate). Administered as the premetabolite

5-Fluorouracil, 5-Fluorouridylate is a mechanism-based inhibitor of TS (1) that is used in chemotherapy, whereas 10-propargyl-5,8 dideazafofolate (CB3717) is a cofactor mimic (2). Although CB3717 has an inhibition constant,  $K_i$ , of 40 nM (3) for TS, it shows liver and kidney toxicity (4), which emphasizes the need for more efficacious agents. The x-ray structure of TS (5, 6) has been used in iterative cycles of crystallographic analysis, synthesis, and inhibition assays to design high-affinity heterocyclic inhibitors that resemble the cofactor (7). Anti-TS drugs dissimilar to the substrate and cofactor remain attractive because they are less likely to have the side effects that are produced by the nucleotide and folate mimics.

We sought novel inhibitors of TS with a computational screen targeted at the structure of *L. casei* TS determined at 2.3 Å resolution (8). We used the program DOCK (9–11) to explore the active site of TS with molecules from the Fine Chemical Directory (FCD) (12). For each FCD molecule, DOCK examined an average of  $10^4$  orientations for steric fit to TS and scored each on the basis of the energy of the electrostatic interaction with the enzyme. Of the 55,313 molecules searched, the 600 with the highest scores were saved. We applied a correction for ligand solvation, which was based on a modified Born equation (13), to sort the selected molecules.

The TS substrate, dUMP, and several known nucleotide and non-nucleotide inhibitors of the enzyme received good scores (Table 1). The program retrieved some molecules previously unknown to bind to TS. On the basis of DOCK score and lack of similarity to dUMP, we selected 25 of these molecules as putative inhibitors. Three of the compounds, including sulisobenzone (SB), inhibited TS with values of the inhibition constant  $\text{IC}_{50}$  (50% reduction) in the high micromolar range (Table 1).

The TS-SB complex was crystallized from both phosphate and tris buffers, and the structure of each was determined to 2.5 Å resolution (14) (Fig. 1). In neither structure was SB found in the binding mode favored by DOCK. We anticipated that the sulfonate group of SB would bind in the TS subsite that is occupied by a phosphate moiety in all of the other structures that we have solved (15). Instead, both of these structures showed a buffer-derived anion, either phosphate or sulfate from the  $(\text{NH}_4)_2\text{SO}_4$  precipitant, bound at this site (Fig. 1B). In both structures, SB is displaced away from the nucleotide binding region to different but overlapping regions of the active site. In the structure crystallized from the phosphate buffer, the sulfonate moiety of SB forms hydrogen bonds with the sulfhydryl of the conserved active site nucleophile, C198, and with the backbone NH of D221, a conserved residue that is involved in cofactor binding (6, 16). The methyl of SB makes van der Waals contacts with L224, which also contributes to the cofactor binding site. In the structure that crystallized from the tris buffer, the SB is rotated 153° and translated 5.1 Å with respect to the configuration that crystallized from the phosphate buffer. The sulfonate is accessible to solvent in the vicinity of the TS carboxyl terminus (Fig. 1B). The substituted ring of SB is in van der Waals contact with W85 at the mouth of the active site. Taken together, the two SB configurations partially overlap the substrate and cofactor subsites of TS, as deter-

B. K. Shoichet, D. V. Santi, I. D. Kuntz, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143.

R. M. Stroud and K. M. Perry, Departments of Pharmaceutical Chemistry and Biochemistry and Biophysics, University of California, San Francisco, CA 94143.

\*B.K.S. and K.M.P. are equally responsible for the work described in this paper.

†To whom correspondence should be addressed.

mined in the ternary-complex structure (15). Some atoms in SB overlap the pyrimidine and ribose subsites in the dUMP-binding region, as well as the pterin ring subsite from the folate-binding region. When superimposed on the conformationally different TS complex with dUMP and CB3717, SB also overlaps side chain atoms of W85. In this ternary complex, W85 is rotated into the active site and may be essential to an important global conformational change and to catalysis (17). By occluding the rotation of W85, SB may inhibit TS, which would prevent this structural change.

Despite the relatively weak affinity, the SB complexes are intriguing because the inhibitor binds in a previously unexplored, but functionally important, region of the TS active site. Molecules similar in shape to SB with better chemical complementarity to the SB site should inhibit the enzyme. We used the similarity routine in MACCS-3D (12) to look for FCD molecules that resemble the shape of SB (Fig. 2). An initial survey suggested 37 candidate molecules, which formed a secondary list of potential inhibitors. We then used DOCK

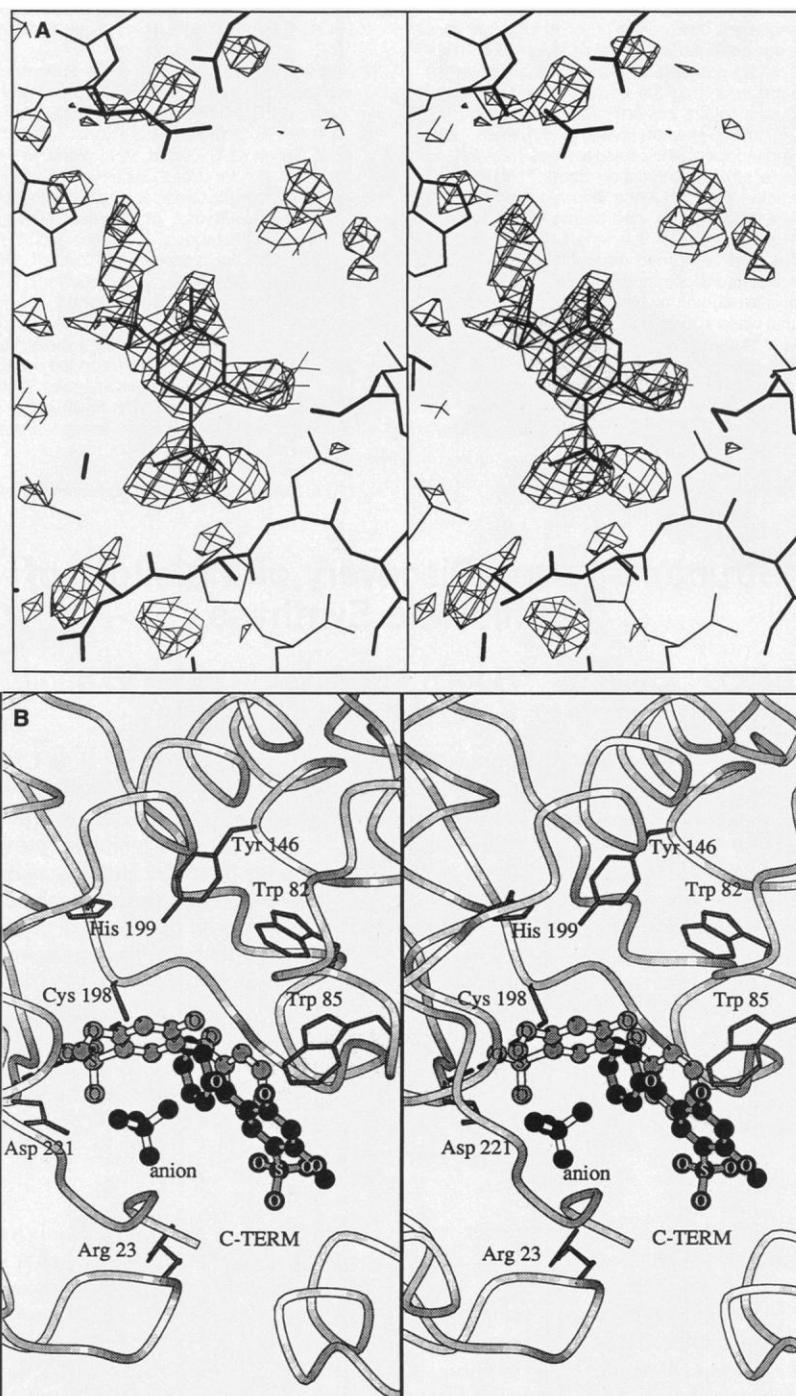
to assess the complementarity of these molecules to the site in TS as defined by the

combined atomic positions of SB in the binary complexes.

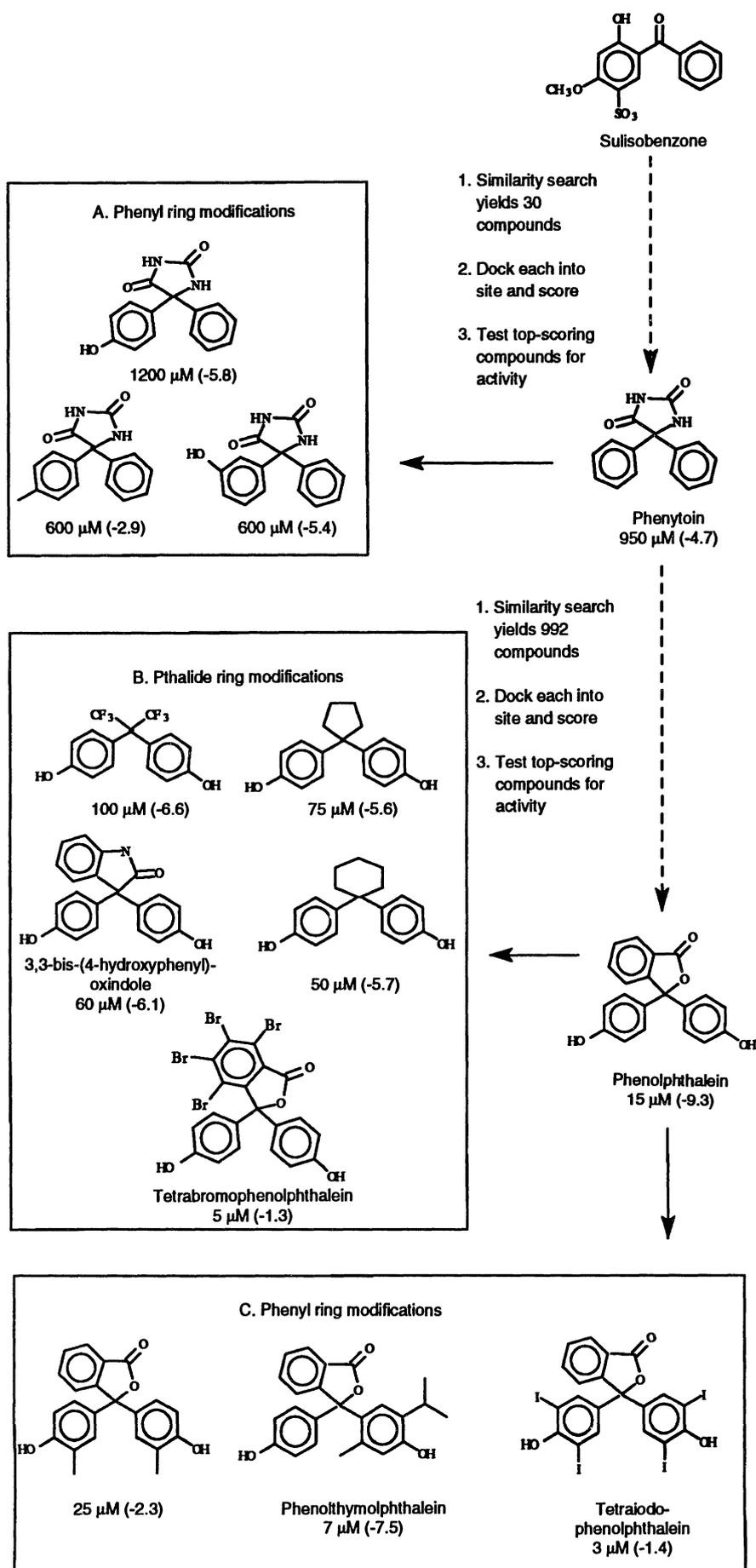
**Table 1.** High-scoring ligands from the original DOCK search against TS and the tertiary DOCK search against the SB site.

Ligand	Rank*	Binding constant ( $\mu\text{M}$ )†
<i>Original search-TS</i>		
Creatine monophosphate	1	NI
4-Thiouridylate	16	70
FdUMP	83	1
4-Nitro-2-formylphenoxyacetic acid‡	157	350
2-Formylphenoxyacetic acid‡	175	200
Sulisobenzonate‡	193	900
Pyridoxal phosphate	314	1
<i>Tertiary search-SB site</i>		
Fluorescein isothiocyanate isomer 1	1	NT
Phenolphthalein‡	2	15
Fluorescein isothiocyanate isomer 2	3	NT
Phenolphthalein‡	4	7
Undecene trione	5	NT
Gibberellin A9‡	6	300
Hexafluoropropane-1,1-bisphenol‡	7	100
1',3',3'-Trimethyl-6-hydroxy-2H-1-benzopyran-2,2'-indoline	8	NT
6,7-Dihydro-2',6'-dioxospiro (thianaphthalene)-4(5H)4'-piperidine-dicyanide	9	NI
3,3-Bis(4-hydroxyphenyl)-oxindole‡	10	60

\*Rank order calculated after solvation correction. See (33) for methods and parameters. †Inhibition was measured spectrophotometrically (34) with the use of *L. casei* TS (35). NI: no measurable inhibition at 500  $\mu\text{M}$ . NT: not tried. See (36) for details. ‡Novel inhibitors of TS.



**Fig. 1.** (A) Omit,  $F_o - F_c$ , electron density for SB that was calculated with phases from the refined model of the TS-SB complex (phosphate buffer) minus the SB atoms. (B) Superposition of SB orientations found in TS cocrystals grown in phosphate (gray atoms) and tris (black atoms) buffers. Complexes of TS and SB were cocrystallized by vapor diffusion at room temperature in 10- $\mu\text{l}$  hanging drops that contained  $\sim 5$  mg of protein per milliliter, with a five- to tenfold excess of ligand and between 40 mM and 125 mM ammonium sulfate,  $(\text{NH}_4)_2\text{SO}_4$ . Crystals were grown in either 20 mM phosphate ( $\text{KH}_2\text{PO}_4$ ), 0.1 mM EDTA, and 1 mM dithiothreitol (DTT) at pH 6.9 or in 50 mM tris, 0.1 mM EDTA, and 1 mM DTT at pH 7.4. The wells contained buffer without  $(\text{NH}_4)_2\text{SO}_4$ . Crystals grown in tris buffer contained protein that was dialyzed against two successive 1-liter volumes of 50 mM tris, 0.1 mM EDTA, and 1 mM DTT at pH 7.4. Crystals grew to  $\geq 500$   $\mu\text{m}$  in length in 3 to 5 days. A summary of the crystallographic statistics for the data reduction and refinement of both SB-containing structures are provided in Table 2.



Seven of the ten top-scoring molecules in the secondary list were derivatives of phenytoin, five of which inhibited TS with  $\text{IC}_{50}$  values of 600 to 1200  $\mu\text{M}$  (Fig. 2). For this reason, we conducted a broader similarity search that combined phenytoin and SB substructures and produced a tertiary list of 992 molecules (Fig. 2). Each of the molecules from this tertiary list was docked to assess steric and electrostatic complementarity to the SB site in TS.

The top-scoring candidates from the tertiary list were phenolphthalein (PTH), several of its analogs, and two chemically unrelated molecules (Table 1 and Fig. 2). The  $\text{IC}_{50}$  for PTH inhibition of TS was 15  $\mu\text{M}$ , and  $K_i$  was 4.4  $\mu\text{M}$ ; PTH is a mixed noncompetitive inhibitor with respect to  $\text{CH}_2\text{-H}_4$  folate and is noncompetitive with dUMP. This inhibition pattern has been noted with certain folate analogs (18). The PTH family of inhibitors had  $\text{IC}_{50}$  values from 3 to 100  $\mu\text{M}$  (Fig. 2). Of the 33 tertiary compounds that we tested, all of which had favorable DOCK scores, three had  $\text{IC}_{50}$  values less than 10  $\mu\text{M}$ , fourteen had  $\text{IC}_{50}$  values between 11 and 100  $\mu\text{M}$ , eight had  $\text{IC}_{50}$  values between 110 and 1000  $\mu\text{M}$ , four had  $\text{IC}_{50}$  values between 1100 and 4000  $\mu\text{M}$ , and four had no measurable inhibition (19).

We next asked if the inhibition of TS by PTH analogs (Fig. 2) could be used to improve our calculated model for PTH binding to TS. We assumed that the analogs bind to TS in a similar manner and that the enzyme is mostly rigid. Besides PTH, tetrabromophenolphthalein, tetraiodophenolphthalein, and phenolthymolphthalein (Fig. 2) were docked into TS. We retained only the orientations in which all four molecules could be accommodated

**Fig. 2.** Design strategy for inhibitors of the SB-site. The  $\text{IC}_{50}$  of each compound is the first number listed under it; the number in parentheses is the DOCK interaction energy. Box A: Phenyl-ring modifications of phenytoin. Box B: Phenolphthalein (PTH) derivatives, modified in the region of the phthalide ring. Box C: Phenyl-ring derivatives of PTH. The phenytoin inhibitors were selected by DOCK from 37 molecules that were similar to SB with the use of MACCS (12) (upper dashed arrow). The similarity template required that two phenyl rings be connected by a ketone oxygen that was within 4 Å of one of the phenyl carbons. The PTH inhibitors were selected by DOCK from the 992 molecules that were found with a broader MACCS template (lower dashed arrow). This template required a quaternary carbon to be connected to four chains, one off of each of single bond; two of the chains had to contain at least two atoms of any type that could be connected by bonds of any type, and the two other chains had to contain at least four atoms of any type, also connected by bonds of any type. The chains could be rings.

by the enzyme. Two families of consensus configurations were found. These were separated by translations of 1.2 to 2.4 Å and rotations of 120° to 137°. Both families occupy the same overall region of the active

site but have different interactions with TS. The crystal structure of the PTH-TS complex was solved (20). An omit map that was calculated from the refined model of the complex, but excluded the PTH

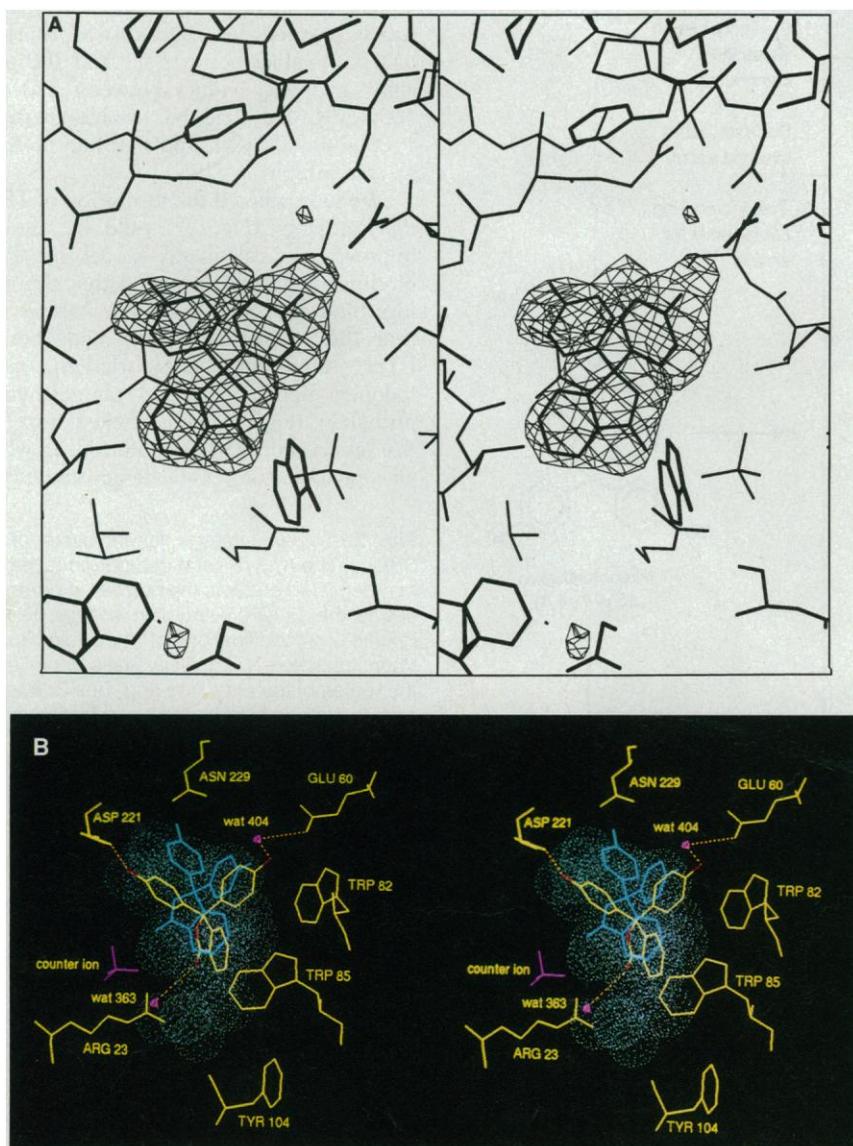
atoms, confirmed the position of the inhibitor (Fig. 3A). Compared with the ternary complex, PTH overlaps portions of the pterin ring of CB3717 and the indole of W85 but does not overlap the dUMP. This is consistent with the inhibition kinetics. There is a direct hydrogen bond between one of the phenol oxygens of PTH and the side chain of D221, whereas a water-mediated hydrogen bond forms between the second phenol oxygen and E60. The lactone carbonyl oxygen of PTH forms hydrogen bond with Wat 363, which also interacts with the guanidinium group of R23 (21). This lactone hydrogen bond might explain the decreased inhibition of 3,3-bis(4-hydroxyphenyl)-oxindole (Fig. 2B), in which the carbonyl of PTH is replaced by a lactam nitrogen. The PTH analogs (Fig. 2) can be accommodated by TS in the crystallographic binding mode with only small adaptations (22).

In general, the x-ray structure agrees with the activity data and models presented above, with several important differences. Superposition of PTH with the combined SB atoms (Fig. 3B) shows that the inhibitor is bound in the target site with a 60% overlap of the van der Waals radii of the PTH and SB atoms. There was significant overlap between the experimental and low-energy DOCK consensus configuration of PTH (Fig. 3B). The two configurations were related by a rotation of 43° and a translation of 1.42 Å and were in the same amino acid side chain environment in that they made van der Waals contacts with W82, E84, W85, I81, V314, and V316. The crystallographic and consensus configurations differ in their polar contacts. Only one of the three polar interactions that were observed crystallographically, the one between E60 and a phenol oxygen, was

**Table 2.** Statistics on data reduction and refinement. Data were collected on a Siemens multiwire area detector that was equipped with a three-circle goniostat. The CuK $\alpha$  x-rays were generated by a Rigaku rotating anode with either a 200  $\mu$ m focal spot source and a graphite monochromator or with a 100  $\mu$ m focal spot source that was coupled with Franks focusing optics. Data were processed with either the Xengen (37) or Buddha (38) data reduction packages.

Complex	Resolution (Å)	Reflections collected (no.)	Total reflections collected (% of possible)	Total observations (no.)	$\langle I/\sigma(I) \rangle$	$R_{\text{sym}}^*$	R factor†
SB pH 6.9	2.5	15,758	94.8	95,906	12.3	0.124	0.213
SB pH 7.4	2.5	15,171	92.8	86,280	18.0	0.101	0.194
PTH	2.7	11,358	90.8	38,464	12.5	0.114	0.207

$$^*R_{\text{sym}} = \left[ \frac{\sum_{hkl} \sum_{i=1}^N (I_{\text{avg}} - I_i)^2}{\sum_{hkl} \sum_{i=1}^N (I_i)^2} \right]^{1/2} \text{ where } I_{\text{avg}} = (1/N) \sum_{i=1}^N I_i \quad \dagger R \text{ factor} = \sum (|F_o| - |F_c|) / |F_c|$$



**Fig. 3.** (A) Omit,  $F_o - F_c$ , electron density for PTH that was calculated with phases from the refined model of the TS-PTH complex minus the PTH atoms. (B) Superposition of experimental (yellow) and low-energy DOCK consensus configuration I (cyan) positions of PTH in the target site [graphics made with MIDAS+ (31)]. Atoms corresponding to the SB orientations are represented as white, van der Waals spheres. Phenolphthalein binary complex crystals were grown in 100 mM KHPO<sub>4</sub>, 0.1 mM EDTA, 1 mM DTT, and 0.5% dimethyl sulfoxide at pH 7.4. The well contained buffer without (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. The interaction energies of the low-energy consensus configurations from families one and two are -11.2 and -3.2 kcal/mol, respectively. Energies were calculated with Chemgrid (23) with the use of the AMBER potential (32). The maximum van der Waals energy for any atom was set to 10 kcal/mol with a 10 Å distance cutoff in a 28 Å box centered on the SB atoms. Grid spacing was 0.33 Å and an  $\epsilon = 4R$  distance-dependent dielectric was used.

sampled in any of the consensus orientations; none were found in the lowest energy structure (Fig. 3B). The conformation of the phenol rings in the modeled and experimental PTH differed by a 27° rotation with respect to the phthalid ring.

The similarities and differences between the consensus and experimental configurations of PTH deserve comment. The initial high ranking of PTH by DOCK arose from its steric complementarity to the SB site and its potential to form hydrogen bonds with several residues in the pocket (for example, E84, E60, W85, and N229). The steric similarity of the binding site in the SB and PTH complexes is a necessary condition for the consensus calculation to put PTH close to where it was found to bind in the x-ray structure. The steric similarity is not sufficient, however, to capture the specific polar complementarity seen in the crystal structure, which arises from the different conformation of R23 in the SB and PTH complexes, and the mediation of Wat 404 and 363, which were unseen in the SB structure. The R23 loop, which includes residues 21 to 25, is generally disordered in the unbound and binary-complex structures that we have determined (8, 15). The loop responds to PTH by becoming ordered to form part of the binding site, in a manner that is similar to the binding of dUMP in the ternary-complex structure (15).

A challenge in structure-based inhibitor design is accounting for the many degrees of freedom that receptors and ligands might explore; DOCK currently assumes rigid geometries for the ligand and receptor. When the geometries are rigid, ligand-protein complexes can be reproduced closely (9, 23). However, the ligand and receptor conformations are altered in many systems. When this change is small, as in PTH-TS, the binding geometries that are suggested by DOCK can approximate crystallographic results; when conformational change is larger, binding predictions become problematic. Our results, and those of others (7, 18), suggest that a few key crystal structures can capture the range and diversity of conformational change and can provide an effective route to enhanced affinity for inhibitors that are discovered on the basis of complementarity to an initial structure. Specifically, the SB-complex structure and the docking calculations that followed demonstrate how new binding zones can be targeted in active sites. The polar complementarity between TS and PTH that was seen in the crystal structure suggests derivatives of PTH that can directly interact with R23 and E60 by displacing the bound waters. Lastly, SB and PTH overlap portions of the substrate and cofactor binding sites, as well as protein side chains of the ternary-complex structure. It may be possi-

ble to design compounds that simultaneously compete with the binding of substrate and cofactor and occlude an essential conformational change within the protein. The PTH analogs are a novel family of TS inhibitors that are easy to synthesize and are attractive platforms for derivatization in the search for tight-binding and specific inhibitors.

## REFERENCES AND NOTES

1. P. Reyes and C. Heidelberger, *Mol. Pharmacol.* **1**, 14 (1965).
2. T. R. Jones *et al.*, *Eur. J. Cancer* **17**, 11 (1981).
3. D. V. Santi and P. V. Danenberg, in *Folates and Pterins*, R. L. Blakely and S. J. Benkovic, Eds. (Wiley, New York, 1984), vol. 1, pp. 345–398.
4. H. M. Pinedo and G. F. Peters, *J. Clin. Oncol.* **6**, 1653 (1988).
5. L. W. Hardy *et al.*, *Science* **235**, 448 (1987).
6. D. A. Matthews, K. Appelt, S. J. Oatley, Ng. H. Xuong, *J. Mol. Biol.* **214**, 923 (1990); A. Kamb, J. Finer-Moore, A. H. Calvert, R. M. Stroud, *Biochemistry* **31**, 9883 (1992).
7. M. D. Varney *et al.*, *J. Med. Chem.* **35**, 663 (1992).
8. J. Finer-Moore *et al.*, *J. Mol. Biol.*, in press.
9. B. Shoichet, D. L. Bodian, I. D. Kuntz, *J. Comput. Chem.* **13**, 380 (1992).
10. R. DesJarlais *et al.*, *J. Med. Chem.* **31**, 722 (1988).
11. I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, T. E. Ferrin, *J. Mol. Biol.* **161**, 269 (1982).
12. O. F. Gunner, D. W. Hughes, L. M. Dumont, *J. Chem. Inf. Comput. Sci.* **31**, 408 (1991).
13. A. A. Rashin, *J. Phys. Chem.* **94**, 1725 (1990).
14. Cocrystallization of SB with *L. casei* TS from phosphate buffer at pH 6.9 yielded crystals in space group P6<sub>1</sub>22 with cell dimensions  $a = b = 78.2$  Å and  $c = 242.5$  Å. The structure was solved by difference Fourier methods (24) (Table 2). The  $F_o - F_c$  map clearly showed the general location of the ligand ( $>3\sigma$  density). Either free atoms or fragments of the inhibitor were built into the map, and several successive rounds of positional or molecular dynamics refinement with XPLOR (25), manual rebuilding, and calculation of a new difference map at 2.5 Å resolution revealed the location of all but four atoms of the unsubstituted ring. An occupancy of ~50% for SB was calculated on the basis of integration of the  $2F_o - F_c$  density for the fairly well-ordered sulfonate moiety. We also crystallized the complex from tris buffer at pH 7.4 with the use of protein that had been dialyzed extensively against phosphate-free buffer (see legend, Fig. 1). This procedure yielded similar crystals in space group P6<sub>1</sub>22 with cell dimensions  $a = b = 78.4$  Å and  $c = 242.8$  Å. The structure was determined iteratively, as before. The SB ligand moved to a new site with occupancy ~50%. Difference ( $F_o - F_c$ ) and  $2F_o - F_c$  maps were calculated at 2.5 Å resolution. In both complexes, difference density adjacent to the unsubstituted ring of SB indicated that at least one additional SB conformer was present in each structure. This was confirmed by the persistent  $F_o - F_c$  density that corresponded to the sulfonate and methoxy moieties. Weak density in both structures at the location of the unsubstituted rings in the more highly populated SB conformers is probably the result of disorder. In each structure, this ring has been modeled to maximize the number of atoms that are included in the density and to avoid steric clash with neighboring side chains.
15. W. R. Montfort *et al.*, *Biochemistry* **29**, 6964 (1990).
16. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
17. J. S. Finer-Moore, W. R. Montfort, R. M. Stroud, *Biochemistry* **29**, 6977 (1990).
18. K. Appelt *et al.*, *J. Med. Chem.* **34**, 1925 (1991).
19. All 33 compounds were derivatives of PTH or phenytoin except for gibberelin A9 (Table 1), the three alkyl bisphenols (Fig. 2B), 2,2-bis(4-aminophenyl) hexafluoropropane ( $IC_{50} = 200$  mM, tertiary list rank 35), and 9,9-bis(4-aminophenyl) fluorene ( $IC_{50} = 40$  mM, tertiary list rank 43).
20. The *L. casei* enzyme cocrystallized with PTH, again in space group P6<sub>1</sub>22, with cell dimensions  $a = b = 78.5$  Å and  $c = 243.0$  Å (Table 1). Comparison of an  $F_o - F_c$  map of this structure calculated at 2.7 Å resolution with one of the unliganded structure identified the position of PTH in early stages of refinement.
21. Although NH<sub>2</sub> of R23 is 3.2 Å from the water oxygen, the 90° H–N–O angle is outside of the limits of hydrogen-bond geometry. This interaction appears to stabilize the position of residues 21 to 25, evidenced by unambiguous density and an average B factor of 51 Å<sup>2</sup> compared with 73 Å<sup>2</sup> for this loop in the unbound *L. casei* TS structure.
22. Three PTH derivatives would make close contacts with the protein if they were bound in the precise crystallographic configuration. The Br<sub>2</sub> atom of tetrabromophenolphthalein contacts Wat 363 (2.7 Å) and the CH<sub>2</sub> of W85 (3.0 Å); the I3 of tetraiodophenolphthalein contacts the C $\alpha$  of G225 (2.7 Å), and the I4 contacts the mobile (B factor = 56 Å<sup>2</sup>) C $\gamma$ 1 of V314 (2.4 Å); and the methyl of phenolphthalein contacts the C $\gamma$ 2 of W85 (2.6 Å).
23. E. C. Meng, B. Shoichet, I. D. Kuntz, *J. Comput. Chem.* **13**, 505 (1992). We repeated the PTH docking calculation with the use of the same parameters as in the original search but with the PTH-complexed conformation of the enzyme, which included bound waters. The lowest energy structure (–8.5 kcal/mol, Chemgrid energy) differed from the crystallographic configuration by a rotation of 29° and a translation of 2.0 Å and made all three specific hydrogen bonds that were seen in the crystal structure.
24. J. L. Chambers and R. M. Stroud, *Acta Crystallogr. Sect. B* **33**, 1824 (1977).
25. A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987).
26. M. L. Connolly, *ibid.* **221**, 709 (1983).
27. J. Gasteiger and M. Marsili, *Tetrahedron* **36**, 3219 (1980).
28. M. Gilson, K. Sharp, B. Honig, *J. Comput. Chem.* **9**, 327 (1988).
29. S. C. Chen, H. H. Daron, J. L. Aull, *Int. J. Biochem.* **21**, 1217 (1989).
30. T. I. Kalman, A. Bloch, G. L. Szekeres, T. J. Bardos, *Biochem. Biophys. Res. Commun.* **55**, 210 (1973).
31. T. E. Ferrin, C. C. Huang, L. E. Jarvis, R. Langridge, *J. Mol. Graphics* **6**, 13 (1988).
32. S. J. Weiner *et al.*, *J. Am. Chem. Soc.* **106**, 765 (1984).
33. In the primary search, FCD molecules were docked into the TS site, which was described by 67 spheres. The spheres were calculated with the Sphgen (11) and Cluster (9) programs and the molecular surface (26) of all residues within 15 Å of the C198. The bound waters and inorganic phosphate were removed. Of the 67 spheres, five were placed at the atomic positions of phosphate from the TS-phosphate complex (17). Orientations were filtered for steric fit to TS with a Distmap grid (9) with polar and nonpolar contact limits of 2.4 and 2.8 Å. We calculated the electrostatic-interaction free energy for orientations without bad contacts by multiplying ligand partial atomic charge with TS electrostatic potential. We used Gasteiger (23, 27) partial atomic charges and a Delphi (28) electrostatic potential. Delphi parameters were water probe: 1.5 Å, ionic probe: 1.6 Å, interior dielectric: 4, exterior dielectric: 80, ionic molarity: 0.1 M; van der Waals radii and charges from the prot.full set were used. Three-step boundary focusing was done with 20, 60, and 90% of the grid filled. Electrostatic solvation free energy for the top 600 ligands was calculated with Hyden (13). Parameters were probe and ionic radii: 0.8 and 2 Å, internal and external dielectrics:

2 and 78, maximum iterations: 10, convergence: 0.001, low and high density surface numbers: 2 and 10. Ligand solvation energy was subtracted from TS interaction energy, recalculated with a Delphi map that used the Hydrin radius, a dielectric constant, and van der Waal parameters. The 67 sphere centers were treated as uncharged atoms of radius 1.4 Å in the Delphi calculation, which reduced the dielectric in the binding site to two.

For the secondary and tertiary searches, the crystallographic phosphate was included in the Distmap and Delphi calculations. Ligands were docked with the positions of the SB atom in each of its two TS complexes acting as spheres centers. The SB atoms were treated as protein atoms of radius 1.4 Å in the Delphi calculation.

34. A. J. Wahba and M. Friedkin, *J. Biol. Chem.* **236**, 11 (1961).  
 35. S. Climie and D. V. Santi, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 633 (1990).  
 36. The  $IC_{50}$  values for 2-formylphenoxyacetic acid, 4-nitro-2-formylphenoxyacetic acid, and creatine monophosphate were measured at 7.5 μM dUMP and 108 μM CH<sub>2</sub>-H<sub>4</sub>folate. The values for FdUMP (3) and PLP (29) are dissociation constants,  $K_d$ ; that of 4-thiouridylate is Michaelis

constant,  $K_m$  (30). The  $IC_{50}$  values were measured at 34 μM CH<sub>2</sub>-H<sub>4</sub>folate and 30 μM dUMP in 1% dimethyl sulfoxide. The fluorescein isothiocyanate isomers were not tried because they react nonspecifically with Lys residues. Undecene trione and 1',3',3'-trimethyl-6-hydroxy-spiro-2H-1-benzopyran-2,2'-indoline are unstable under our assay conditions.

37. A. J. Howard, C. Neilsen, N. H. Xuong, *Methods Enzymol.* **114**, 452 (1985).  
 38. M. Blum, P. Metcalf, S. C. Harrison, D. C. Wiley, *J. Med. Chem.* **20**, 235 (1987).  
 39. We thank E. Meng, C. Carreras, P. Ghosh, and the University of California, San Francisco, TS group. Research supported by NIH grants GM31497 (I.D.K.), GM39553 (G. L. Kenyon, principal investigator), GM24485 (R.M.S.), CA-41323 (R.M.S. and J. Finer-Moore), and AI 32784 (D.V.S.). Crystallographic refinement was carried at the Pittsburgh Supercomputer Center under grant DMB890040P. The facilities of the Computer Graphics Lab (RR-1081 to R. Langridge) were used for part of this work. We thank Molecular Design Limited for supplying MACCS-3D and the FCD and Tripos Associates for SYBYL.

4 August 1992; accepted 7 December 1992

## Deficiency in Rhabdomyosarcomas of a Factor Required for MyoD Activity and Myogenesis

Stephen J. Tapscott,\* Mathew J. Thayer, Harold Weintraub

Rhabdomyosarcoma cells express the myogenic helix-loop-helix proteins of the MyoD family but do not differentiate into skeletal muscle cells. Gel shift and transient transfection assays revealed that MyoD in the rhabdomyosarcoma cells was capable of binding DNA but was relatively nonfunctional as a transcriptional activator. Heterokaryon formation with fibroblasts resulted in the restoration of transcriptional activation by MyoD and the differentiation of the rhabdomyosarcoma cells into skeletal muscle cells. These results suggest that rhabdomyosarcomas are deficient in a factor required for MyoD activity.

Rhabdomyosarcomas are one of the most common solid tumors of childhood. The expression of the myogenic determination gene *MyoD* has been shown as the most sensitive marker for the classification of sarcomas as rhabdomyosarcomas (1, 2). This relation is paradoxical because MyoD, a member of the helix-loop-helix (HLH) family of transcription factors (3), has been shown to induce muscle differentiation in a wide variety of primary cells and transformed cell lines (4). However, rhabdomyosarcomas have a low propensity to differentiate into myotubes. Therefore, we analyzed several rhabdomyosarcoma cell lines for MyoD expression and function.

Three rhabdomyosarcoma cell lines were used to assess HLH gene expression: RD, Rh18, and Rh30 (5). Northern (RNA) analysis (Fig. 1A) demonstrated that these cells contained normal amounts of *myogenin* and *E2A* (6) transcripts and either normal or reduced amounts of *Id* (7) transcripts

compared with the amounts detected in primary cultured human muscle cells. The RD and Rh30 cells, but not the Rh18 cells expressed *MyoD* RNA. All three rhabdomyosarcoma lines expressed *Myf5* (8). In-

direct immunofluorescence experiments with antibodies to MyoD (Fig. 2), myogenin, and E2A (8) revealed a normal nuclear localization for each protein in the RD and Rh30 cells. The rhabdomyosarcoma cells exhibited a similar MyoD immunoreactivity as primary human muscle cells (8). Because the RD and Rh30 cells had no obvious abnormalities in expression of the HLH genes that are thought to be critical for regulating myogenesis, these two cell lines were chosen for further analysis.

To investigate the DNA binding properties of MyoD in the RD and Rh30 cells, we prepared nuclear extracts from the cells and used a gel shift assay to look for complexes that bound to a <sup>32</sup>P-labeled probe that contained a MyoD binding site. Similar to the positive control (nuclear extracts from C2C12 mouse myoblasts), the RD and Rh30 extracts produced two major species (bands 1 and 2) that were sensitive to competition with low amounts of unlabeled probe but insensitive to competition with an unlabeled probe that contained a mutation in the MyoD binding site (Fig. 1B). Antibodies against MyoD and E2A were used in double-shift assays to show that the slower of the two major species (band 1) contained MyoD and E2A in both C2C12 and rhabdomyosarcoma cells (8).

To determine whether an increase in the amount of MyoD could induce differentiation in RD and Rh30 cells, we transfected the cells with a *MyoD* expression vector (pLTR-*MyoD*) (9) and examined expression of MyoD and myosin heavy chain (MyHC) by immunohistochemistry. The transfected cells expressed high amounts of MyoD in the nucleus (Fig. 2, C and E). As in nontransfected cultures (Fig. 2B), about 1% of all cells underwent spontaneous differentiation to a muscle phenotype, which we determined by their MyHC expression,

**Table 1.** MyoD, E12, and MyoD-VP16 (MDVP16) activity in fibroblasts and rhabdomyosarcomas. Cells were transfected (19) with the reporter construct (5 μg) and the indicated expression vector (5 μg). The *MyoD* expression vector uses the Moloney sarcoma virus LTR (MSV-LTR) to drive the *MyoD* cDNA. Therefore, to quantitate expression from the MSV-LTR after transfection into different cell types, we transfected cells in parallel with a construct that contained the CAT gene driven by the MSV-LTR, and the data were normalized to CAT activity from this construct.

Cell line	Reporter	CAT activity (10 <sup>-3</sup> cpm)						
		LTR	LTR-MyoD	LTR-E12	LTR-E12 + LTR-MyoD	LTR-MDVP16	Gal-MyoD	Gal1-147
10T1/2	4RTK-CAT	1	87	18	826	553		
RD	4RTK-CAT	2	6	27	29	161		
Rh30	4RTK-CAT	6	6	8	20	65		
Rh28	4RTK-CAT	3	4	4	4	35		
RhJT	4RTK-CAT	4	4	4	4	77		
10T1/2	GM4TK-CAT		47				53	0
RD	GM4TK-CAT		2				209	0
Rh30	GM4TK-CAT		8				132	0

Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, WA 98104.

\*To whom correspondence should be addressed.