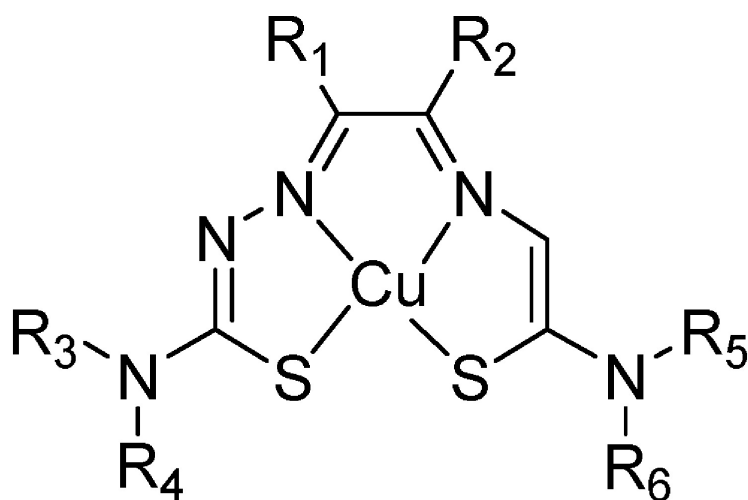


## QSAR Studies of Copper Azamacrocycles and Thiosemicarbazones: MM3 Parameter Development and Prediction of Biological Properties

Peter Wolohan, Jeongsoo Yoo, Michael J. Welch, and David E. Reichert

*J. Med. Chem.*, **2005**, 48 (17), 5561-5569 • DOI: 10.1021/jm0501376 • Publication Date (Web): 23 July 2005

Downloaded from <http://pubs.acs.org> on March 28, 2009



Cu-bis(thiosemicarbazones)

### More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 2 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)



## QSAR Studies of Copper Azamacrocycles and Thiosemicarbazones: MM3 Parameter Development and Prediction of Biological Properties

Peter Wolohan, Jeongsoo Yoo, Michael J. Welch, and David E. Reichert\*

Washington University School of Medicine, St. Louis, Missouri 63110

Received February 11, 2005

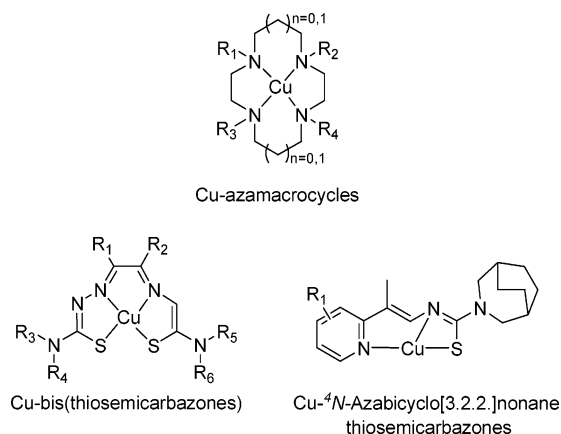
Genetic algorithms (GA) were used to develop specific copper metal–ligand force field parameters for the MM3 force field, from a combination of crystallographic structures and ab initio calculations. These new parameters produced results in good agreement with experiment and previously reported copper metal–ligand parameters for the AMBER force field. The MM3 parameters were then used to develop several quantitative structure–activity relationship (QSAR) models. A successful QSAR for predicting the lipophilicity ( $\log P_{ow}$ ) of several classes of Cu(II)-chelating ligands was built using a training set of 32 Cu(II) radiometal complexes and 6 simple molecular descriptors. The QSAR exhibited a correlation between the predicted and experimental  $\log P_{ow}$  with an  $r^2 = 0.95$ ,  $q^2 = 0.92$ . When applied to an external test set of 11 Cu(II) complexes, the QSAR performed with great accuracy;  $r^2 = 0.93$  and a  $q^2 = 0.91$  utilizing a leave-one-out cross-validation analysis. Additional QSAR models were developed to predict the biodistribution of a smaller set of Cu(II) bis(thiosemicarbazone) complexes.

### Introduction

Radioactive copper complexes have been utilized for many years in both a therapeutic mode and as imaging agents for such noninvasive medical techniques as gamma scintigraphy and positron emission tomography (PET).<sup>1,2</sup> Macrocyclic polyaminocarboxylate ligands for Cu(II) have been used most successfully as bifunctional chelates (BFCs) where the metal chelate is coupled to a biological targeting molecule such as a monoclonal antibody or a peptide.<sup>3–6</sup> Cu(II) bis(thiosemicarbazone) complexes have been used in vivo as radiotracers for the evaluation of blood flow in various organs such as the brain, kidneys, and heart as well as for the evaluation of hypoxia in tissue.<sup>7–9</sup> In addition Cu(II) thiosemicarbazone complexes are of interest as they have shown significant antitumor activity.<sup>10</sup> Indeed whether used in a therapeutic mode or as an imaging agent the desire to understand and predict the biodistribution and bioactivity of such Cu(II) complexes is a priority.

A particularly important property to be able to predict for in vivo use is the logarithm of the experimental *n*-octanol/water partition coefficient ( $\log P_{ow}$ ). This is an important pharmacological descriptor of bioavailability for both organic compounds and inorganic complexes.<sup>11</sup> Predicting and altering  $\log P_{ow}$  is a key element in rational drug design since the correct incorporation of lipophilic groups to a chemical compound can result in an increase in biological activity.<sup>12</sup>

Quantitative structure–activity relationships (QSAR) have become a common tool in the field of molecular modeling since their introduction.<sup>13</sup> Indeed they have found application in both the prediction of biological activity and more recently in the prediction of the absorption, distribution, metabolism, excretion, and

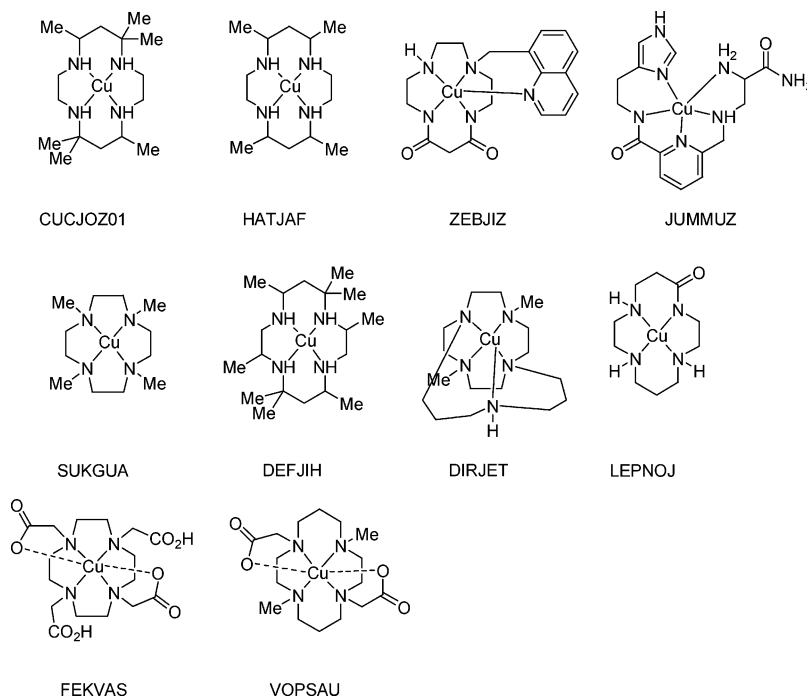


**Figure 1.** Classes of copper(II) chelates studied.

toxicological (ADME/tox) properties of organic drug-like compounds.<sup>14–18</sup> QSAR methods have also been applied to the prediction of  $\log P_{ow}$ .<sup>19</sup> While QSAR methods have found many applications in the modeling of the properties of organic drug-like compounds, there have been few reports of QSAR studies of metal-based pharmaceuticals. One possible factor in this low utilization is the lack of specific metal–ligand parameters for most common force fields; in general, most force fields do a poor job of accurately predicting the structure of complexes of interest for in vivo imaging or radiotherapy. With the development of suitable metal–ligand force field parameters, the utilization of QSAR methods requires much less effort and will likely aid the search for novel metal-based pharmaceuticals targeted toward biological receptors.

The MM3\* force field developed by Allinger et al.<sup>20</sup> is one of the most accurate force fields available for the modeling of organic ligands.<sup>21</sup> Several researchers have successfully developed specific metal–ligand parameters compatible with this accurate force field: Ru,<sup>22</sup> Re,<sup>23</sup> Pd,<sup>24</sup> and a variety of metal amide complexes.<sup>25</sup>

\* Address all correspondence to David E. Reichert, Ph.D., Washington University School of Medicine, 510 S. Kingshighway Blvd., Campus Box 8225, St. Louis, MO 63110, e-mail: reichertd@wustl.edu, fax: (314) 362-9940, voice: (314) 362-8461.

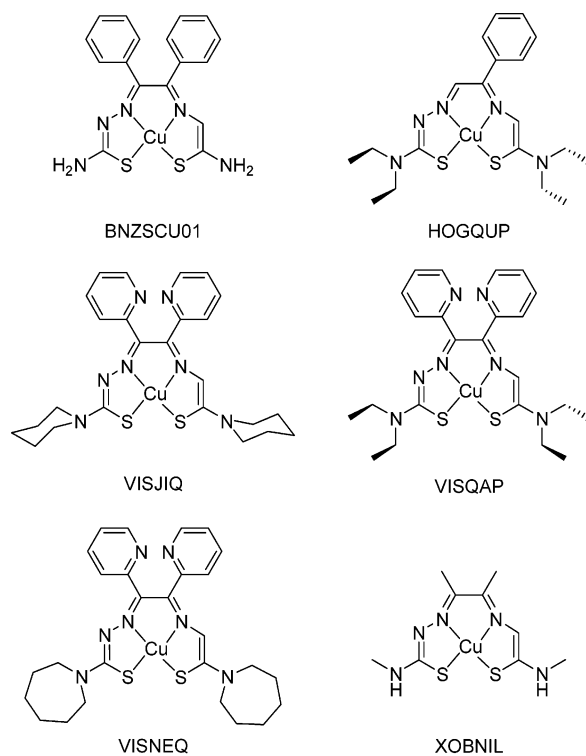


**Figure 2.** Reference compounds used to develop the Cu(II)-azamacrocyclic parameters.

The MM3\* force field describes the energy of a molecule with a simple algebraic expression consisting of terms describing bond stretching, angle bending, dihedral rotation, and intermolecular forces, such as electrostatics, van der Waals interactions, and hydrogen bonding. The constants in these equations are obtained from experimental data or *ab initio* calculations. The development of specific metal–ligand force field parameters consists of determining these constants within the framework of the algebraic expression of the force field being used and is commonly achieved via two main methods. The first method utilizes a traditional optimization algorithm such as the Newton–Raphson or Simplex techniques while the second, and probably most promising method, utilizes a genetic algorithm (GA).<sup>23,26,27</sup> Both methods approach this multidimensional problem from the same starting point in that they attempt to find a common set of parameters which reproduce most accurately the crystallographic X-ray structure of a training set of metal complexes. We have adopted the GA approach of Strassner et al.<sup>23</sup> within the framework of their automated parameter program FFGeneAtoR in this work.

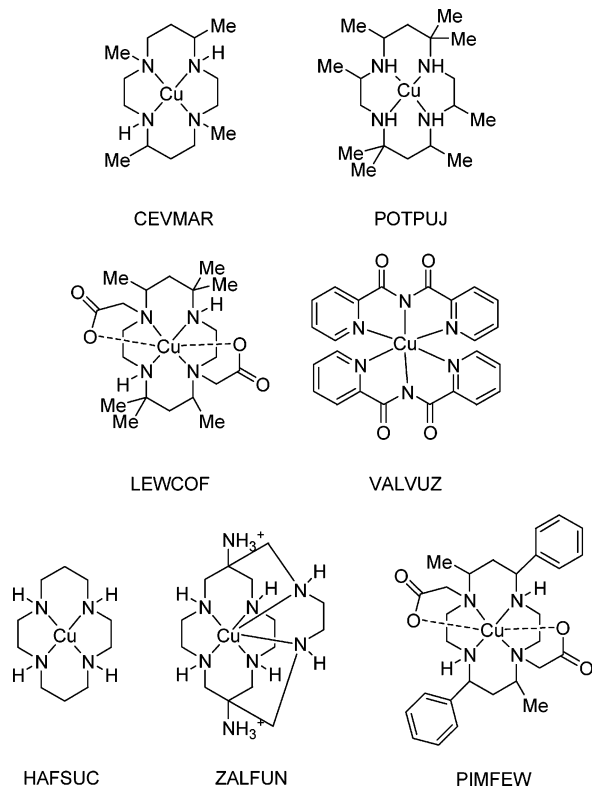
### Experimental Section

Copper–ligand complexes were implemented as structures in the MM3\* force field using the TINKER<sup>28</sup> molecular modeling software running on an SGI Indigo<sup>2</sup> workstation. X-ray structures representative of three classes of important copper radiopharmaceuticals (Figure 1) were taken from the Cambridge Structural Database<sup>29</sup> (CSD) and manipulated with the program MOLDEN,<sup>30</sup> which provides a convenient interface between the CSD and TINKER. The reference data for the Cu(II) parameters involved a total of 16 copper containing crystal structures, shown in Figures 2 and 3. The first set of 10 copper containing crystal structures were predominately azamacrocyclic complexes (refcodes CUCJOZ01, SUKGUA, DIRJET, FEKVAS, VOPSAU, DEFJIH, HATJAF, JUMMUZ, ZEBJIZ, LEPNOJ) while the second 6 were predominately bis-(thiosemicarbazone) complexes (refcodes BNZSCU01, HOGQUP, VISJIQ, VISNEQ, VISQAP, XOBNIL).

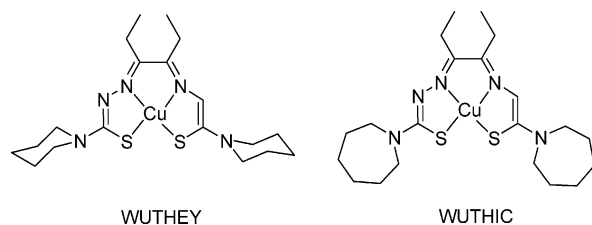


**Figure 3.** Reference compounds used to develop the Cu(II)-bis(thiosemicarbazone) parameters.

From MOLDEN a TINKER input file was generated for each of the copper complexes. At this point a series of new MM3\* force field atom types were defined based on details from TINKER's optimize program which produces a text file consisting of a list of those force field parameters which are missing in the MM3\* force field. These new force field parameters included the definition of new MM3\* parameters for a four-, five- and six-coordinate copper atom together with definitions for coordinating organic ligands. For example, looking at (1,4,7,10-tetramethyl-1,4,7,10-tetraazacyclododecane)copper(II) [SUKGUA] a four-coordinate Cu(II) atom type was defined together with a new nitrogen N sp<sup>3</sup> to represent the amines coordinating to the central Cu(II) metal atom. In



**Figure 4.** Validation complexes for the Cu(II)-azamacrocyclic parameters.



**Figure 5.** Validation complexes for the Cu(II)-bis(thiosemicarbazone) parameters.

**Table 1.** Comparison of Calculated MM3 Structure to X-ray for Cu(II)-azamacrocyclic Reference Set and Previously Reported AMBER Calculated Structure

REFCODE	rmsd (Å)	
	AMBER (NewtonRaphson/Simplex)	MM3 (GA)
CUCJOZ01	0.386	0.194
SUKGUA	0.267	0.191
DIRJET	0.155	0.190
FEKVAS	0.329	0.344
VOPSAU	0.313	0.257
DEFJIH	0.475	0.190
HAFTAJ	0.359	0.220
JUMMUZ	1.640	0.398
ZEBJIZ	0.296	0.394
LEPNOJ	0.275	0.244
Average	0.350	0.272

addition, in the case of the six-coordinate copper azamacrocycles where two of the coordinating groups are carboxylates, separate O  $sp^3$  parameters were defined in order to represent axial-axial and axial-equatorial configurations.

These force field parameters were optimized using the automated parameter development program FFGeneAtr which utilizes a GA. Following the work of Strassner et al.<sup>23</sup> a steady-state GA using a roulette selector together with a crossover rate of 0.9 and a mutation rate of 0.1 was used for each run. Within the GA the force field parameters were represented by chromosomes of real numbers. Within a

**Table 2.** Comparison of Calculated MM3 Structure to X-ray for Cu(II)-azamacrocyclic Validation Set

REFCODE	rmsd (Å) MM3 (GA)
CEVMAR	0.176
HAFSUC	0.162
POTPUJ	0.223
LEWCOF	0.404
PIMFEW	0.288
VALVUZ	0.503
ZALFUN	0.163
average	0.274

**Table 3.** Comparison of Calculated MM3 Structure to X-ray for Cu(II)-bis(thiosemicarbazone) Reference and Validation Set

REFCODE	rmsd (Å) MM3 (GA)
BNZSCU01	0.330
HOGQUP	0.221
VISJIQ	0.274
VISNEQ	0.206
VISQAP	0.244
XOBNIL	0.053
Average	0.221
WUTHEY	0.441
WUTHIC	0.189
Average	0.315

**Table 4.** Results from QSAR Model of the Lipophilicity ( $\log P_{ow}$ ) of the Copper(II) Complexes<sup>a</sup>

QSAR	$r^2$	$q^2$	RMSE	RMSE <sub>cv</sub>	$n_{mols}$	$n_{descriptors}$
$\log P_{ow}$ model (training set)	0.95	0.92	0.35	0.35	32	6
$\log P_{ow}$ model (test set)	0.93	0.91	0.41	0.47	11	6

<sup>a</sup>  $r^2$  refers to nonvalidated correlation,  $q^2$  to the cross-validated correlation, RMSE to the root-mean-square error, and RMSE<sub>cv</sub> to the cross-validated root-mean-square error.

population each individual chromosome is constructed of genes that correspond to the missing parameters i.e., bond, angle, torsion, and out of plane bending parameters. Initial values for these parameters were taken from an average of the corresponding experimental value following structural analysis of the reference structures from the CSD. Bond lengths were allowed to optimize within a range of  $\pm 0.5$  Å of this average while angles were allowed to optimize within a range of  $\pm 30^\circ$ . Associated force constants were constrained to a range of 0–250 mydn/Å and 0–30 mydn/(rad<sup>2</sup>), respectively. Coding of the torsional parameters is more arbitrary because of the shear number of these missing parameters, hence the need to impose flexible boundaries for these terms in order to allow the GA to optimize fully. As a result torsional angles were allowed to optimize within a range –180 and 180 and their associated constants were constrained to a range of  $\pm 200$  kcal/mol.

Initial guesses for the coded parameters from the GA were fed into the force field, making it now possible to minimize each starting X-ray structure using TINKER's optimize program. Once minimized, TINKER's superpose was used to calculate the root-mean-squared deviation (rmsd), most commonly in mass-weighted coordinates, between the two structures. The rmsd is then used to generate a fitness value for each newly developed parameter set which in turn guides the GA. The cycle is repeated for a number of GA cycles which we set to 750 generations. Electrostatic point charges were added to the final force field parameters following the procedure of Bayly et al. and the programs RESP and Jaguar 5.0.<sup>31,32</sup>

The end result is a set of newly defined force field parameters whose fitness to both the reference set of structures and a blind validation set of structures can be evaluated based on the average rmsd, the goal being to produce a set of robust

**Table 5.** Comparison of Calculated and Experimental Lipophilicity ( $\log P_{ow}$ ) of Copper(II) Complexes in QSAR Training Set<sup>a</sup>

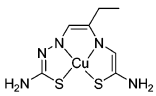
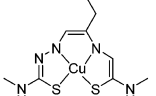
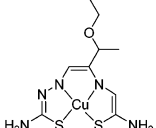
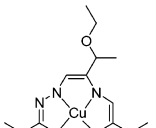
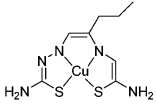
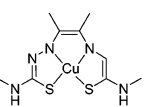
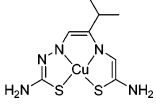
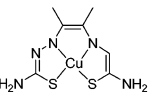
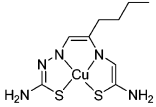
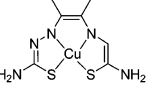
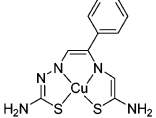
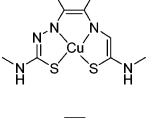
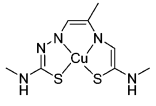
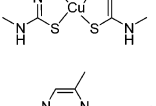
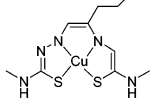
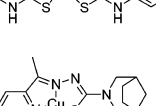
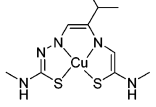
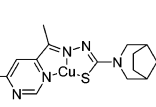
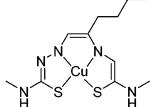
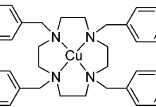
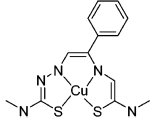
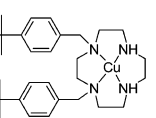
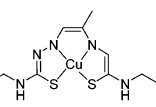
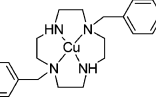
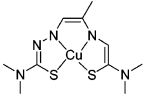
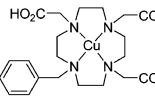

Compound	Structure	LogP <sub>exp</sub>	LogP <sub>calc</sub>	Residual	Compound	Structure	LogP <sub>exp</sub>	LogP <sub>calc</sub>	Residual
ETS		1.35	1.27	-0.08	ETSM		2.65	2.05	-0.60
KTS		1.71	1.44	-0.27	KTSM2		3.46	3.50	0.04
<i>n</i> -PrTS		1.78	1.81	0.03	ATSM		1.48	2.10	0.62
<i>i</i> -PrTS		1.82	2.04	0.22	ATS		0.65	1.32	0.67
<i>n</i> -BuTS		2.30	2.35	0.05	CTS		2.34	1.74	-0.60
PhTS		2.37	2.27	-0.10	DTSM		2.34	2.97	0.63
PTSM		1.92	1.73	-0.19	GTSM		0.84	1.16	0.32
<i>n</i> -PrTSM		2.86	2.60	-0.26	PTS		1.96	2.2	0.25
<i>i</i> -PrTSM		2.77	2.83	0.06	EPH142		1.72	1.67	-0.05
<i>n</i> -BuTSM		3.16	3.13	-0.03	EPH144		1.92	1.98	0.06
PhTSM		3.06	2.91	-0.15	tetraBBC		2.36	2.27	-0.09
PTSE		2.76	2.49	-0.27	bis14BBC		2.62	2.70	0.08
PTSM2		3.18	3.02	-0.16	bis17BBC		3.19	2.70	-0.49
					BBC4710trisAA		1.27	0.45	-0.82



Table 5. (Continued)

Compound	Structure	LogP <sub>exp</sub>	LogP <sub>calc</sub>	Residual	Compound	Structure	LogP <sub>exp</sub>	LogP <sub>calc</sub>	Residual
trisBBC10AA		2.73	2.86	0.13	TETA		-1.67	-1.91	-0.24
bis14BBC710 bisAA		1.77	2.68	0.91	CBTE2A		-2.29	-2.17	0.12
DOTA		-3.24	-2.85	0.39					

<sup>a</sup> log  $P_{exp}$  refers to the log  $P_{10}$  of the experimental *n*-octanol/water partition coefficient, log  $P_{calc}$  refers to the QSAR log  $P_{10}$  of the calculated *n*-octanol/water partition coefficient, and Residual refers to the difference between the predicted and measured values.

parameters which accurately reproduce the X-ray structure of a diverse set of related structures with an average rmsd below 0.400 being indicative of accuracy.

**Parameter Testing.** Seven crystal structures were chosen as a validation set of the Cu(II) azamacrocycles parameters (Figure 4) CEVMAR, LEWCOF, HAFSUC, PIMFEW, POT-PUJ, VALVUZ and ZALFUN. These independent structures were minimized with the developed parameter set and compared to the X-ray crystal structure as an accuracy test. In the same way two crystal structures were chosen as a validation set of the Cu(II) bis(thiosemicarbazone) parameters (Figure 5) WUTHEY and WUTHIC. It is important to note that these structures serve as an independent test of the parameters as they were not utilized in developing the parameters.

At this point the newly developed MM3\* Cu(II) force field parameters from both the copper(II) azamacrocycles and copper(II) bis(thiosemicarbazones) were combined to form a complete MM3\* force field parameter file. Exhaustive testing was carried out by reproducing the rmsd for all structures in the reference data sets with the combined parameters to ensure no conflicts.

**QSAR Modeling.** For the QSAR modeling a total of 43 structures, representative of the three classes of important copper radiopharmaceuticals (Figure 1), with their experimental log  $P_{ow}$ s were constructed in the program MOE.<sup>33</sup> Eighteen copper(II) bis(thiosemicarbazone) complexes were taken from John et al.,<sup>34</sup> a further nine from Dearling et al.,<sup>35</sup> and a set of four copper(II) thiosemicarbazone complexes from Easmon et al.<sup>36</sup> A total of eight copper(II) regioselectively N-substituted derivatives of cyclen were taken from Yoo et al.<sup>37</sup> Finally the cross-bridged analogues of copper(II) 1,4,7,10-tetraazacyclododecane-1,4,8,11-tetracarboxylic acid (DOTA) and copper(II) 1,4,8,11-tetraazacyclotetradecane-1,4,8,11-tetracarboxylic acid (TETA) CBDO2A and CBTE2A, respectively, were taken from Boswell et al.<sup>38</sup> and Wong et al.<sup>39</sup>

Where an experimental structure was not available, a related analogue was used as a template. For example the X-ray structure of XOBNIL was utilized as the template for constructing models of the copper(II) bis(thiosemicarbazone) complexes etc. These structures were then minimized using the developed MM3\* force field parameters and TINKER's optimize program. Where necessary the conformational space of the functional groups of any of copper complexes was sampled using TINKER's scan program. Once a viable model of each of the 43 structures was obtained these structures were passed back to MOE for QSAR modeling. Two-dimensional schematics of all these structures are included in Tables 5 and 6.

Within the program MOE, 170 molecular descriptors were calculated for each of the 43 structures. These molecular descriptors attempt to describe basic concepts such as shape, size, internal charge distribution, and lipophilicity through a myriad of terms which when combined can be used to construct

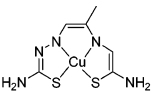
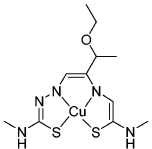
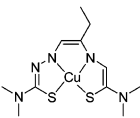
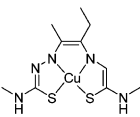
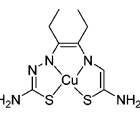
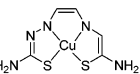
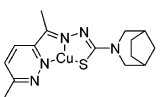
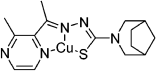
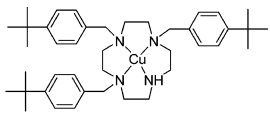
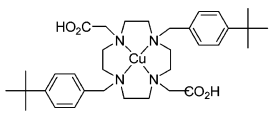
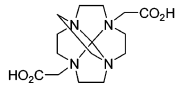
a quantitative relationship to an experimental property such as biological activity.<sup>18</sup> For the log  $P_{ow}$  QSAR, a test set of 11 structures was selected at random from the complete set of 43 compounds, leaving a training set of 32 structures. The number of molecular descriptors used in the model was drastically reduced by evaluating their individual statistical significance to the experimental data, a process commonly referred to as pruning the descriptors. Similarly, QSAR models were constructed to model the experimental biodistribution data of the 18 copper(II) bis(thiosemicarbazone) complexes from John et al.<sup>34</sup> Here, because of the small size of the dataset, four structures were selected at random for the blind test set leaving a training set of 14 structures.

In QSAR equations  $n$  is the number of observations,  $r^2$  is the correlation coefficient utilizing a partial least squares (PLS) analysis,  $q^2$  is the cross-validated correlation coefficient utilizing a leave-one-out analysis, RMSE is the root-mean-square error, SEE<sub>cv</sub> is the associated cross-validated root-mean-square error, and  $F$  is a measure of the statistical significance of the correlation. The  $q^2$  coefficient is the traditional measure of the quality of the fit, with a value above 0.50 suggesting a predictive model.

## Results and Discussion

Initially we wanted to investigate whether we could reproduce the accuracy of our previously reported AMBER force field parameters<sup>40</sup> using the approach of Strassner et al.<sup>23</sup> As a result, we initially used the same Cu(II) reference and validation structures in our initial GA analysis consisting of predominately azamacrocycles complexes as shown in Figures 2 and 4, respectively. The mass-weighted root-mean-squared deviation (rmsd) of each of the structures in the reference azamacrocycles set from both methods are shown in Table 1. As can be seen from Table 1, the GA developed MM3\* parameters were able to reproduce the rmsd accuracy of the previously developed AMBER parameters. Indeed, overall there is a slight improvement in the accuracy of the force field parameters, taking the average rmsd of the entire reference set. If one excludes the azamacrocycles containing carboxylates, one can see a significant improvement in the rmsd between the two sets of developed force field parameters also suggesting that it may be more appropriate to treat this class of azamacrocycles as a unique set. Nevertheless, when applied to the blind validation set (Table 2), one can see that the level of accuracy is maintained. As a result, it is clear to see that it is possible to develop robust and accurate metal-ligand parameters regardless of the force field used or

**Table 6.** Comparison of Calculated and Experimental Lipophilicity  $\log P_{\text{ow}}$ s of Copper(II) Complexes in QSAR Test Set<sup>a</sup>

Compound	Structure	Log $P_{\text{exp}}$	Log $P_{\text{calc}}$	Residual
PTS		0.76	0.98	0.22
KTSM		2.64	2.23	-0.41
ETSM2		3.32	3.34	0.02
CTSM		2.69	2.53	-0.16
DTS		1.69	2.18	0.49
GTS		0.45	0.47	0.02
EPH143		1.33	2.51	1.18
EPH270		1.34	1.30	-0.04
tris147BBC		2.73	3.56	0.83
bis17BBC410 bisAA		2.28	2.68	0.40
CBDO2A		-2.42	-3.15	-0.73

<sup>a</sup>  $\log P_{\text{exp}}$  refers to the  $\log P_{10}$  of the experimental *n*-octanol/water partition coefficient,  $\log P_{\text{calc}}$  refers to the QSAR  $\log P_{10}$  of the calculated *n*-octanol/water partition coefficient, and Residual refers to the difference between the predicted and measured values.

the method used to develop the parameters, be it a traditional minimization algorithm or a GA.

Having successfully developed parameters for the copper(II) azamacrocycles, we moved on to developing parameters for the copper(II) bis(thiosemicarbazones). The results from the MM3\* parameter development for the six X-ray structures (Figure 3) are shown in Table

**Table 7.** Results from Several QSAR Models of the Biodistribution of the Copper(II) Bis(thiosemicarbazones) in the Training Set<sup>a</sup>

QSAR model	$r^2$	$q^2$	RMSE	RMSE <sub>cv</sub>	$n_{\text{mols}}$	$F$
Liver						
$\log_{10}(\text{blood/liver}) @ 1 \text{ min}$	1.00	1.00	0.01	0.01	14	4273
$\log_{10}(\text{blood/liver}) @ 5 \text{ min}$	1.00	1.00	0.01	0.01	14	6879
$\log_{10}(\text{blood/liver}) @ 2 \text{ h}$	1.00	1.00	0.01	0.01	14	4067
Kidney						
$\log_{10}(\text{blood/kidney}) @ 1 \text{ min}$	1.00	0.97	0.01	0.01	14	2456
$\log_{10}(\text{blood/kidney}) @ 5 \text{ min}$	1.00	1.00	0.01	0.01	14	1069
$\log_{10}(\text{blood/kidney}) @ 2 \text{ h}$	1.00	0.98	0.01	0.01	14	1193

<sup>a</sup>  $\log_{10}(\text{blood/liver})$  refers to  $\log_{10} \{(\% \text{ID/g blood})/(\% \text{ID/g liver})\}$ ,  $\log_{10}(\text{blood/kidney})$  refers to  $\log_{10} \{(\% \text{ID/g blood})/(\% \text{ID/g kidney})\}$ ,  $r^2$  refers to nonvalidated correlation,  $q^2$  to the cross-validated correlation,  $F$  is a measure of the statistical significance of the model, RMSE refers to the root-mean-square error, and RMSE<sub>cv</sub> refers to the cross-validated root-mean-square error.

**Table 8.** Analysis of Results from QSAR Models of the Biodistribution of the Copper(II) Bis(thiosemicarbazones) in the Test Set<sup>a</sup>

QSAR model	$r^2$	$q^2$	RMSE	RMSE <sub>cv</sub>	$n_{\text{mols}}$
Liver					
$\log_{10}(\text{blood/liver}) @ 1 \text{ min}$	1.00	0.97	0.01	0.03	4
$\log_{10}(\text{blood/liver}) @ 5 \text{ min}$	0.99	0.98	0.02	0.03	4
$\log_{10}(\text{blood/liver}) @ 2 \text{ h}$	0.02	0.67	0.11	0.20	4
Kidney					
$\log_{10}(\text{blood/kidney}) @ 1 \text{ min}$	0.34	0.20	0.06	0.16	4
$\log_{10}(\text{blood/kidney}) @ 5 \text{ min}$	0.26	0.93	0.08	0.19	4
$\log_{10}(\text{blood/kidney}) @ 2 \text{ h}$	0.01	0.45	0.01	0.01	4

<sup>a</sup>  $\log_{10}(\text{blood/liver})$  refers to  $\log_{10} \{(\% \text{ID/g blood})/(\% \text{ID/g liver})\}$ ,  $\log_{10}(\text{blood/kidney})$  refers to  $\log_{10} \{(\% \text{ID/g blood})/(\% \text{ID/g kidney})\}$ ,  $r^2$  refers to nonvalidated correlation,  $q^2$  to the cross-validated correlation, RMSE refers to the root-mean-square error, and RMSE<sub>cv</sub> refers to the cross-validated root-mean-square error.

3. Although a smaller data set, hence requiring a smaller number of interdependent parameters to be developed, the average rmsd of both the reference set and the blind validation set (Figure 5) of structures exhibit excellent accuracy. This is despite the diversity of the functional groups of the metal–ligand scaffold in these copper(II) bis(thiosemicarbazones).

#### QSAR Modeling of $\log P_{\text{ow}}$ and Biodistribution.

To illustrate the utility of the development of new metal–ligand force field parameters, we endeavored to use our new MM3\* parameters to build QSAR models of important experimental biological data. The biological data utilized for these studies consisted of  $\log P_{\text{ow}}$  measurements and the biodistribution of a subset of the copper(II) bis(thiosemicarbazone) ligands in animal models. As discussed  $\log P_{\text{ow}}$  is an important indicator of a compounds ability to permeate through human tissue particularly its ability to pass the blood–brain barrier and exhibit cerebral activity.<sup>41</sup> While  $\log P_{\text{ow}}$  can give an indication of how a compound might act in vivo, biodistribution studies on animals, most commonly rats, relate the accumulation and clearance of a compound at different time-points directly. Most commonly the retention of the compound in the blood, brain, heart, kidney, liver, and lungs at different times is measured thus mapping the distribution of the compound.<sup>34</sup> The ratio of the mass of compound found in the blood over that found in the kidney is used to define the blood clearance via urinary excretion, while the ratio of the mass of compound found in the blood over that found in the liver describes blood clearance via hepatobiliary excretion.

The results from the log  $P_{ow}$  QSAR analysis for both the training set of 32 copper radiopharmaceuticals and the test set of 11 copper radiopharmaceuticals are shown in Table 4. The cross-validated correlation coefficient,  $q^2$ , of 0.92 and small cross-validated error, RMSE<sub>cv</sub>, indicate a predictive model exceeding the 0.50 benchmark. This conclusion is supported by the performance of the same QSAR in predicting the log  $P_{ow}$ s of the eleven compounds in the blind test set. Here the corresponding  $q^2 = 0.91$  is also well above the benchmark 0.50. The corresponding QSAR equation is given below:

$$\log P_{calcd} = -4.674 + (0.039 * \text{PEOE\_VSA}+) - (0.016 * \text{PEOE\_VSA}-) + (0.052 * \text{PEOE\_VSA\_NEG}) + (0.030 * \text{SlogP\_VSA0}) - (6.677 * \text{b\_1rotR}) - (0.002 * \text{weinerPath}) \quad (1)$$

Three of the predictive descriptors (PEOE) are based on the partial charge analysis of Gasteiger et al.<sup>42</sup> The VSA extension refers to the fitting of the calculated partial charges to the van der Waals surface of the molecule while the further extensions, i.e., +, -, refer to a summation of the surface area where the partial charge falls between 0 and 0.05 for the + label and -0.05-0.00 for the - label. The extension NEG refers to the summation of the van der Waals surface where the PEOE charge is negative. The SlogP\_VSA0 descriptor is based on the log  $P_{ow}$  prediction by Crippen et al. which attempts to predict the experimental log  $P_{ow}$  of a molecule based on atomic contributions.<sup>43</sup> The VSA0 label refers to the summation of the van der Waals surface where SlogP is less than or equal to -0.40. The remaining two molecular descriptors are b\_1rotR which is the fraction of rotatable single bonds and weinerPath which is a summation of the entries in a matrix which describes specific internal distance metrics.<sup>44</sup> In conclusion, our correlation to the experimental log  $P_{ow}$  is based on a relationship of the van der Waals surface at specific points where particular terms, charge and SlogP, are satisfied together with the fraction of rotatable single bonds and an internal description of distance metrics.

Similarly, Tables 5 (training set) and 6 (test set) present the results from the biodistribution QSAR analysis 18 copper(II) bis(thiosemicarbazone) complexes from John et al. Here it was consistently feasible to construct QSAR models exhibiting cross-validated correlations of well above the benchmark 0.50 utilizing a total of eight molecular descriptors. Indeed  $F$  numbers above 1000 signified statistically significant QSAR models no doubt aided by the small size of the training set. However, here is a clear example of why the statistical indicators  $r^2$ ,  $q^2$ , and  $F$  are not reliable indicators of the external predictive ability of any model, particularly one with few observations, since upon analysis of the test sets a range of  $r^2$  and  $q^2$  are observed (Table 8). Despite this overall observation our QSAR models perform generally very well when exposed to the blind test sets. Indeed we are able to achieve  $q^2$ s for these test sets above 0.50 four out of six times. The poorest performance comes from the prediction of the kidney data at the 1 min and 2 h time points exhibiting

$q^2 = 0.20$  and  $0.45$  respectively when looking at the test set data. The corresponding QSAR equations are given below:

QSAR Equations (Liver)

$$\log_{10}(\text{blood/liver}) @1 \text{ min} = 5.516 + (0.021 \times \text{ASA\_P}) - (30.678 \times \text{Q\_RPC-}) - (0.010 \times \text{SMR\_VSA7}) + (0.013 \times \text{SlogP\_VSA0}) - (0.718 \times \text{a\_ICM}) + (3.283 \times \text{balabanJ}) - (0.235 \times \text{petitjeanSC}) - (0.783 \times \text{std\_dim3}) \quad (2)$$

$$\log_{10}(\text{blood/liver}) @5 \text{ min} = -13.656 + (0.001 \times \text{ASA\_H}) + (0.001 \times \text{CASA+}) + (3.010 \times \text{VDistMa}) + (0.044 \times \text{a\_count}) - (0.283 \times \text{a\_heavy}) - (0.831 \times \text{chi0v}) + (2.315 \times \text{density}) - (0.086 \times \text{std\_dim1}) \quad (3)$$

$$\log_{10}(\text{blood/liver}) @2 \text{ h} = 1.246 - (0.256 \times \text{Kier2}) + (2.202 \times \text{Kier3}) + (0.020 \times \text{SMR\_VSA5}) - (0.103 \times \text{a\_count}) - (0.106 \times \text{radius}) - (0.002 \times \text{weinerPath}) + (0.223 \times \text{weinerPol}) - (0.062 \times \text{zagreb}) \quad (4)$$

QSAR Equations (Kidney)

$$\log_{10}(\text{blood/kidney}) @1 \text{ min} = 5.634 - (0.004 \times \text{CASA+}) - (2.848 \times \text{FASA-}) - (1.203 \times \text{Kier1}) - (38.459 \times \text{Q\_RPC-}) + (0.009 \times \text{SlogP\_VSA3}) - (0.013 \times \text{SlogP\_VSA7}) + (2.334 \times \text{mr}) + (1.222 \times \text{std\_dim2}) \quad (5)$$

$$\log_{10}(\text{blood/kidney}) @5 \text{ min} = 26.965 + (0.004 \times \text{CASA+}) - (0.394 \times \text{KierFlex}) + (6.745 \times \text{PC-}) - (33.304 \times \text{PEOE\_RPC+}) - (56.195 \times \text{PEOE\_RPC-}) + (0.008 \times \text{SMR\_VSA4}) + (0.033 \times \text{radius}) - (0.031 \times \text{weinerPol}) \quad (6)$$

$$\log_{10}(\text{blood/kidney}) @2 \text{ h} = -0.276 + (0.002 \times \text{ASA+}) + (0.423 \times \text{KierA1}) - (1.797 \times \text{PEOE\_VSA\_FPNEG}) + (1.193 \times \text{SlogP}) - (0.295 \times \text{apol}) - (0.574 \times \text{chi0\_C}) + (0.313 \times \text{petitjeanSC}) + (0.311 \times \text{weinerPol}) \quad (7)$$

Given the number of descriptors used, a full description of each one is not presented here; suffice to say that they encapsulate aspects of the accessible van der Waals surface, functionality, and charge distribution of the copper(II) bis(thiosemicarbazone) complexes used in the training set.

## Conclusions

This work illustrates the effectiveness of the application of a genetic algorithm to the method of developing metal ligand parameters for Allinger's MM3\* force field. The development of new metal ligand parameters provides an important tool for the researcher in the fields of bioinorganic chemistry and nuclear medicine. The availability of specific metal ligand parameters for these important radiopharmaceuticals provides the researcher with a timely opportunity to model their biological properties/action, in doing so aiding the development of radiopharmaceuti-



cals with enhanced therapeutic and imaging characteristics. These new parameters were utilized in the development of quantitative structure–activity relationships to predict first the lipophilicity of a series of common copper(II)-chelating agents and second the biodistribution of 18 copper(II) bis(thiosemicarbazone) radiometal complexes.

In the case of the log  $P_{ow}$  QSAR model, we were able to predict with excellent accuracy the lipophilicity of several classes of copper(II) complexes. Our biodistribution QSAR models were able to predict with excellent accuracy the clearance of the copper(II) bis(thiosemicarbazones) from the liver based on an analysis of our test set while our QSAR models performed less adequately when predicting the uptake of the copper(II) bis(thiosemicarbazones) in the kidneys. A more rigorous analysis of the application of the QSAR methodology to the prediction of biodistribution would be facilitated by a larger and more diverse dataset, an issue which we hope to address in future work.

**Acknowledgment.** This work was supported by the National Institute of Biomedical Imaging and Bioengineering (EB00340) and the United States Department of Energy (DEFG02-087ER-60512).

**Supporting Information Available:** A complete set of the MM3\* parameters described in this work are available. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- Blower, P. J.; Lewis, J. S.; Zweit, J. Copper radionuclides and radiopharmaceuticals in nuclear medicine. *Nucl. Med. Biol.* **1996**, *23*, 957–980.
- Reichert, D. E.; Lewis, J. S.; Anderson, C. J. Metal Complexes as Diagnostic Tools. *Coord. Chem. Rev.* **1999**, *184*, 3–66.
- Moi, M. K.; Meares, C. F.; McCall, M. J.; Cole, W. C.; DeNardo, S. J. Copper chelates as probes of biological systems: stable copper complexes with a macrocyclic bifunctional chelating agent. *Anal. Biochem.* **1985**, *148*, 249–253.
- Cole, W. C.; DeNardo, S. J.; Meares, C. F.; McCall, M. J.; DeNardo, G. L.; Epstein, A. L.; O'Brien, H. A.; Moi, M. K. Comparative serum stability of radiochelates for antibody radiopharmaceuticals. *J. Nucl. Med.* **1987**, *28*, 83–90.
- Anderson, C. J.; Pajeau, T. S.; Edwards, W. B.; Sherman, E. L.; Rogers, B. E.; Welch, M. J. In vitro and in vivo evaluation of copper-64-octreotide conjugates. *J. Nucl. Med.* **1995**, *36*, 2315–2325.
- Anderson, C. J.; Dehdashti, F.; Cutler, P. D.; Schwarz, S. W.; Laforest, R.; Bass, L. A.; Lewis, J. S.; McCarthy, D. W. 64Cu-TETA-octreotide-64Cu-TETA-octreotide as a PET imaging agent for patients with neuroendocrine tumors. *J. Nucl. Med.* **2001**, *42*, 213–221.
- Green, M. A. A Potential Copper Radiopharmaceutical for Imaging the Heart and Brain: Copper-Labeled Pyruvaldehyde Bis(N<sup>4</sup>-methylthiosemicarbazone). *Nucl. Med. Biol.* **1987**, *14*, 59–61.
- Fujibayashi, Y.; Taniuchi, H.; Yonekura, Y.; Ohtani, H.; Konishi, J.; Yokoyama, A. Copper-62-ATSM: a new hypoxia imaging agent with high membrane permeability and low redox potential. *J. Nucl. Med.* **1997**, *38*, 1155–1160.
- Dearling, J. L. J.; Lewis, J. S.; McCarthy, D. W.; Welch, M. J.; Blower, P. J. Redox-active metal complexes for imaging hypoxic tissues: Structure–activity relationships in copper(II) bis(thiosemicarbazone) complexes. *J. Chem. Soc., Chem. Commun.* **1998**, 2531–2532.
- Easmon, J.; Pürstinger, G.; Heinisch, G.; Roth, T.; Fiebig, H. H.; Holzer, W.; Jäger, W.; Jenny, M.; Hofmann, J. Synthesis, Cytotoxicity, and Antitumor Activity of Copper(II) and Iron(II) Complexes of <sup>4</sup>N-Azabicyclo[3.2.2]nonane Thiosemicarbazones Derived from Acyl Diazines. *J. Med. Chem.* **2001**, *44*, 2164–2171.
- Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3–25.
- Andrews, P. R.; Craik, D. J.; Martin, J. L. Functional group contributions to drug-receptor interactions. *J. Med. Chem.* **1984**, *27*, 1648–1657.
- Hansch, C.; Maloney, P. P.; Fujita, T.; Muir, R. M. Correlation of biological activity of phenoxyacetic acids with Hammett substituent constants and partition coefficients. *Nature* **1962**, *194*, 178.
- Hansch, C.; Leo, A. *Exploring QSAR Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, DC, 1995; p 557.
- Agarwal, A.; Pearson, P. P.; Taylor, E. W.; Li, H. B.; Dahlgren, T.; Herslöf, M.; Yang, Y.; Lambert, G.; Nelson, D. L.; Regan, J. W.; Martin, A. R. Three-Dimensional Quantitative Structure–Activity Relationships of 5-HT Receptor Binding Data for Tetrahydropyridinylindole Derivatives: A Comparison of the Hansch and CoMFA Methods. *J. Med. Chem.* **1993**, *36*, 4006–4014.
- Hansch, C.; Leo, A.; Mekatpati, S. B.; Kurup, A. QSAR and ADME. *Bioorg. Med. Chem.* **2004**, *12*, 3391–3400.
- Klein, C.; Kaiser, D.; Kopp, S.; Chiba, P.; Ecker, G. F. Similarity based SAR (SIBAR) as tool for early ADME profiling. *J. Comput.-Aided Mol. Des.* **2003**, *16*, 785–793.
- Garg, R.; Kurup, A.; Mekatpati, S. B.; Hansch, C. Cyclooxygenase (COX) Inhibitors: A Comparative QSAR Study. *Chem. Rev.* **2003**, *103*, 703–731.
- Eros, D.; Kovacs, L.; Orfi, L.; Takacs-Novak, K.; Acsady, G.; Keri, G. Reliability of logP predictions based on calculated molecular descriptors: a critical review. *Curr. Med. Chem.* **2002**, *9*, 1819–1829.
- Allinger, N. L.; Yuh, Y. H.; Lii, J. H. Molecular mechanics. The MM3 force field for hydrocarbons. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.
- Gundertofte, K.; Liljefors, T.; Norrby, P.-O.; Pettersson, I. A comparison of conformational energies calculated by several molecular mechanics methods. *J. Comput. Chem.* **1996**, *17*, 429–449.
- Brandt, P.; Norrby, T.; Åkermark, B.; Norrby, P.-O. Molecular Mechanics (MM3\*) Parameters for Ruthenium(II)-Polypyridyl Complexes. *Inorg. Chem.* **1998**, *37*, 4120–4127.
- Strassner, T.; Busold, M.; Herrmann, W. A. MM3 parametrization of four- and five-coordinated rhenium complexes by a genetic algorithm—which factors influence the optimization performance? *J. Comput. Chem.* **2002**, *23*, 282–290.
- Hagelin, H.; Svensson, M.; Åkermark, B.; Norrby, P.-O. Molecular Mechanics (MM3) Force Field Parameters for Calculations on Palladium Olefin Complexes with Phosphorus Ligands. *Organometallics* **1999**, *18*, 4574–4583.
- Hay, B. P.; Clement, O.; Sandrone, G.; Dixon, D. A. A Molecular Mechanics (MM3(96)) Force Field for Metal-Amide Complexes. *Inorg. Chem.* **1998**, *37*, 5887–5894.
- Norrby, P.-O.; Brandt, P. Deriving Force Field Parameters for Coordination Complexes. *Coord. Chem. Rev.* **2001**, *212*, 79–109.
- Cundari, T. R.; Fu, W. Genetic algorithm optimization of a molecular mechanics force field for technetium. *Inorg. Chim. Acta* **2000**, *300–302*, 113–124.
- Ponder, J. W. *TINKER4.2*, 2004.
- Allen, F. H.; Kennard, O. 3D Search and Research Using the Cambridge Structural Database. *Chemical Des. Automation News* **1993**, *8*, 31–37.
- Schaftenaar, G.; Noordik, J. H. Molden: a pre- and postprocessing program for molecular and electronic structures. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 123–134.
- Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
- Jaguar*; 5.0 ed.; Schrödinger, Inc.: Portland, OR.
- Molecular Operating Environment; Chemical Computing Group Inc., Quebec, Canada, 1997–2004.
- John, E. K.; Green, M. A. Structure–activity relationships for metal-labeled blood flow tracers: comparison of keto aldehyde bis(thiosemicarbazone)copper(II) derivatives. *J. Med. Chem.* **1990**, *33*, 1764–1770.
- Dearling, J. L. J.; Lewis, J. S.; Mullen, G. E. D.; Welch, M. J.; Blower, P. J. Copper bis(thiosemicarbazone) complexes as hypoxia imaging agents: structure–activity relationships. *J. Biol. Inorg. Chem.* **2002**, *7*, 249–259.
- Easmon, J.; Heinisch, G.; Holzer, W.; Rosenwirth, B. Pyridazines. 63. Novel thiosemicarbazones derived from formyl- and acyl-diazines: synthesis, effects on cell proliferation, and synergism with antiviral agents. *J. Med. Chem.* **1992**, *35*, 3288–3296.
- Yoo, J.; Reichert, D. E.; Welch, M. J. Comparative in Vivo Behavior Studies of Cyclen-Based Copper-64 Complexes: Regioselective Synthesis, X-ray Structure, Radiochemistry, log P, and Biodistribution. *J. Med. Chem.* **2004**, *47*, 6625–6637.
- Boswell, C. A.; Sun, X.; Niu, W.; Weisman, G. R.; Wong, E. H.; Rheingold, A. L.; Anderson, C. J. Comparative in Vivo Stability of Copper-64-Labeled Cross-Bridged and Conventional Tetraazamacrocyclic Complexes. *J. Med. Chem.* **2004**, *47*, 1465–1474.

- (39) Wong, E. H.; Weisman, G. R.; Hill, D. C.; Reed, D. P.; Rogers, M. E.; Condon, J. S.; Fagan, M. A.; Calabrese, J. C.; Lam, K.-C.; Guzei, I. A.; Rheingold, A. L. Synthesis and Characterization of Cross-Bridged Cyclams and Pendant-Armed Derivatives and Structural Studies of Their Copper(II) Complexes. *J. Am. Chem. Soc.* **2000**, *122*, 10561–10572.
- (40) Reichert, D. E.; Norrby, P.-O.; Welch, M. J. Molecular Modeling of Bifunctional Chelate Peptide Conjugates. 1. Copper and Indium Parameters for the AMBER Force Field. *Inorg. Chem.* **2001**, *40*, 5223–5230.
- (41) Levin, V. A. Relationship of Octanol/Water Partition Coefficient and Molecular Weight to Rat Brain Capillary Permeability. *J. Med. Chem.* **1980**, *23*, 682–684.
- (42) Gasteiger, J.; Marsili, M. Iterative Partial Equalization of Orbital Electronegativity – A Rapid Access to Atomic Charges. *Tetrahedron* **1980**, *36*, 3219–3288.
- (43) Wildman, S. A.; Crippen, G. M. Prediction of Physicochemical Parameters by Atomic Contributions. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 868–873.
- (44) Balaban, A. T. Chemical graphs. XXXIV. Five new topological indexes for the branching of tree-like graphs. *Theor. Chim. Acta* **1979**, *53*, 355–375.

JM0501376