



Pergamon

Bioorganic & Medicinal Chemistry Letters 12 (2002) 1537–1541

BIOORGANIC &
MEDICINAL
CHEMISTRY
LETTERS

CoMFA and QSAR of Acylhydrazide Cruzain Inhibitors

Carlos R. Rodrigues,^{a,b,c,*} Terrence M. Flaherty,^c Clayton Springer,^c
James H. McKerrow^c and Fred E. Cohen^c

^aLASSBio, Faculdade de Farmácia, Universidade Federal do Rio de Janeiro, RJ, 21944-970, Brazil

^bDepartment of Química Orgânica-Instituto de Química, UFRJ, Cidade Universitária, RJ, 21949-900, Brazil

^cDepartment of Cellular Molecular Pharmacology, University of California, San Francisco, CA 94143-0446, USA

Received 20 September 2001; accepted 5 February 2002

Abstract—An approach combining CoMFA and QSAR methods was used to describe QSAR models for a series of cruzain inhibitors having the acylhydrazide framework. A CoMFA study using two alignment orientations (I and II), three different probe atoms and changes of the lattice spacing (1 and 2 Å) was performed. Alignment II and an sp³ probe carbon atom yielded good cross-validation ($q^2=0.688$) employing lattice spacing of 1 Å. The best QSAR model was generated using *atoms*, *bond*, and *connectivity* as *fragment distinction* and *fragment size* default (4–5) showing similar cross-validated value of CoMFA ($q^2=0.689$). Based upon the information derived from CoMFA and QSAR, we have identified some key features that may be used to design new acylhydrazide derivatives that may be more potent cruzain inhibitors. © 2002 Elsevier Science Ltd. All rights reserved.

Trypanosoma cruzi is the causative agent of Chagas' disease (American trypanosomiasis). Chagasic heart disease is estimated to affect over 16 million people in Central and South America, and over 90 million are at risk.¹ Currently, there is no satisfactory treatment for this parasitic disease. Cruzain is the major cysteine protease from *T. cruzi* and is critical for the development and survival of the parasite within the host cells making this enzyme a prime target for potential trypanocidal drugs.^{2,3} Motivated by the results of a structure based design effort,⁴ we prepared a series of arylacylhydrazide derivatives and screened these compounds against cruzain.⁵ Several compounds with activity at high nanomolar concentrations were identified, but their potency fell short of the required Trypanosome clearance in vivo.⁶ In the absence of a co-crystal structure of cruzain bound to an acylhydrazide inhibitor to drive a structure-based design effort, we have turned to quantitative structure–activity relationship (QSAR) studies using the Comparative Molecular Field Analysis (CoMFA) method⁷ and Hologram Quantitative structure–activity Relationship (HQSAR).⁸ Since its introduction in 1988,⁷ CoMFA has rapidly become one of the most powerful tools for 3-D-QSAR studies.^{9–12} In HQSAR,

each molecule in the database is divided into a set of overlapping structural fragments and sorted to form a molecular hologram.¹³ This process is similar to the generation of molecular fingerprints employed in database searches¹⁴ and molecular diversity calculation.^{15,16} In addition, HQSAR gives sub-structural features in sets of molecules that are relevant to biological activity. On the basis of the CoMFA and HQSAR models for 112 acylhydrazide derivatives, we attempted to elucidate a structure–activity relationship to provide useful guidelines for the design of more potent acylhydrazide inhibitors of cruzain. To test the predictive ability of the resulting QSAR models (CoMFA and HQSAR), a test set of 16 structurally diverse cruzain inhibitors excluded from the model creation work was used.

Molecular models of acylhydrazides derivatives were built and fully minimized by the Tripos force field with the N=C bond fixed in the *trans* position (*E*-isomer). This is consistent with ¹H NMR characterization of molecules from this series.¹⁷ Partial atomic charges were computed using Gasteiger–Hückel protocol in Sybyl 6.3.¹⁸ The atoms forming the acylhydrazide group (O=C–NH–N=C) were superimposed on the corresponding atoms present in a template molecule (1). CoMFA requires the user to identify a single receptor relevant conformation for each ligand. Some acylhydrazides in the training set have *o*-hydroxyl groups in

*Corresponding author. Tel.: +55-21-290-9192; fax: +55-21-273-3890; e-mail: rangela@pharma.ufrj.br

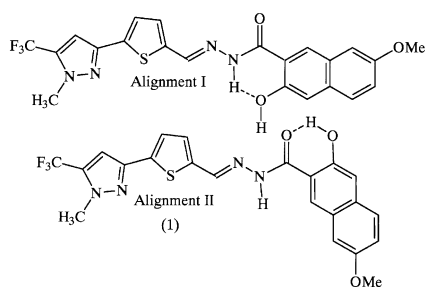


Figure 1. Alignment I and II showed for (1) derivative.

aromatic substituents and two plausible conformations that are approximately isoenergetic must be considered. In the first case, the hydroxyl oxygen can act as a hydrogen bond acceptor for the hydrogen on the hydrazide nitrogen (alignment I) (Fig. 1). The other possibility is that the hydroxyl hydrogen acts as donor for hydrogen bond formation with the carbonyl oxygen present in the acylhydrazide backbone (alignment II) (Fig. 1). We considered these two datasets (alignment I and II) for the subsequent QSAR analysis (Tables 1 and 2). We used three different probe atoms in this study: an sp^3 carbon with a charge of +1, an sp^3 oxygen with a charge of -1 and a hydrogen with a charge of +1. The default cutoff values of 30 kcal/mol for both steric and electrostatic terms were employed for both the Standard and Indicator force fields. A 3-D cubic lattice with a lattice spacing of 1 and 2 Å was constructed around these molecules.

Although the statistics for alignment I and II after PLS analysis yielded similar and consistent results, alignment II was selected as the best model owing to a slightly better q^2 for all probe atoms in both the Tripos Standard and Indicator Force Fields¹⁹ (Table 2). From a study of the q^2 for an sp^3 probe carbon atom in alignment II, we have developed a higher resolution CoMFA using a grid of 1 Å to calculate both steric and electrostatic interaction energies. This new grid was employed to improve the robustness of the calculation with respect to the energies at discrete lattice points surrounding the molecule (Table 3). Indeed the model with smaller lattice using the Tripos Standard Force Field gave the higher cross-validated value ($q^2=0.688$) (Table 3).

A close inspection of the steric contour plots revealed that the sterically favorable regions were located next to the R^1 substituent or more precisely near the pyrazole group (green polyhedra) (Fig. 2A). We can observe that any substitution in R^2 (2-OH,6-OMe naphthylene) will be located very close to an unfavorable region (yellow polyhedra). The electrostatic contribution contour is plotted in Figure 2B. The red contours surrounding the 4-position in the thiophene ring show that negative potentials in this position are likely to increase the inhibitory profile of the compound. However, the strong steric hindrance at this position makes this site a chemically a challenging site for substitution. The blue contours identify areas within the lattice where electro-positive substituents are predicted to increase activity.

Table 1. CoMFA analysis for alignment I using Tripos Standard and Indicator Force Fields with a lattice spacing of 2 Å

	Probe atoms					
	Tripos Standard			Indicator		
	Csp ³ +1	H +1	Osp ³ -1	Csp ³ +1	H +1	Osp ³ -1
q^2	0.592	0.575	0.587	0.637	0.490	0.529
(SE _{cv}) ^a	0.292	0.543	0.524	0.495	0.575	0.526
r^2	0.873	0.877	0.878	0.876	0.805	0.799
(SE)	0.281	0.292	0.294	0.283	0.363	0.362
PCA ^b	5	5	5	5	4	4

^aStandard error.

^bPrinciple Component Analysis.

Table 2. CoMFA analysis for alignment II using Tripos Standard and Indicator Force Fields with lattice spacing of 2 Å

	Probe atoms					
	Tripos Standard			Indicator		
	Csp ³ +1	H +1	Osp ³ -1	Csp ³ +1	H +1	Osp ³ -1
q^2	0.643	0.602	0.607	0.634	0.534	0.595
(SE _{cv})	0.488	0.516	0.514	0.493	0.552	0.515
r^2	0.868	0.867	0.866	0.858	0.778	0.836
(SE)	0.306	0.296	0.305	0.310	0.384	0.336
PCA	5	5	5	5	5	5

Table 3. CoMFA analysis for alignment II using lattice spacing of 1 Å and Csp³ +1 as probe atom

	Tripos Standard	Indicator
q^2	0.688	0.666
(SE _{cv})	0.467	0.454
r^2	0.898	0.884
(SE)	0.278	0.273
F test	169.074	165.015
PCA	5	5

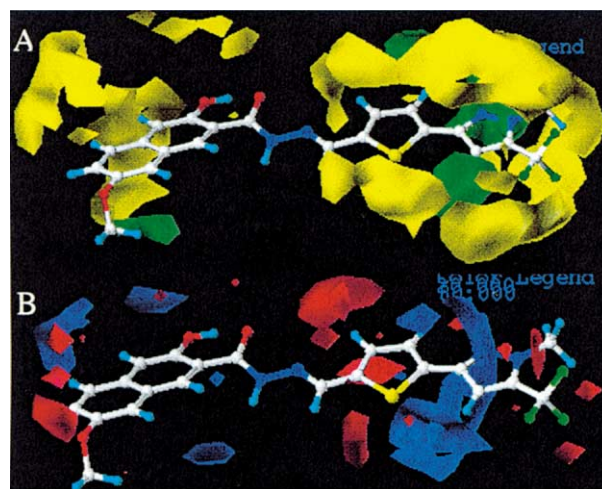


Figure 2. The CoMFA contour map. (A) Steric map indicating areas where bulk is predicted to increase (green) or decrease (yellow) activity. (B) Electrostatic map indicating where high electron density (negative charge) (red) and low electron density (positive charge) (blue) regions are expected to increase activity.

Table 4. HQSAR analysis for various *fragment distinction* on the key statistical parameters using *fragment size* default (4–7)

Fragment distinction	Statistical parameters				
	q ²	SE _{cv}	r ²	SE	PCA
Atom/bond ^a	0.667	0.482	0.846	0.332	11
Connectivity (Con)	0.689	0.461	0.867	0.306	10
Hydrogen (H)	0.665	0.476	0.846	0.347	13
Con-H	0.654	0.493	0.848	0.343	10
Chirality ^b	0.663	0.474	0.856	0.332	11

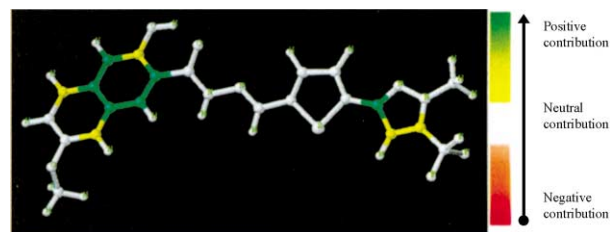
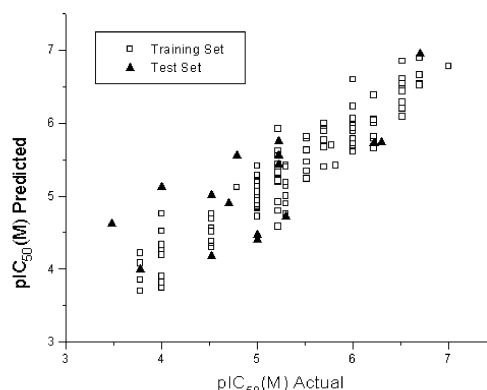
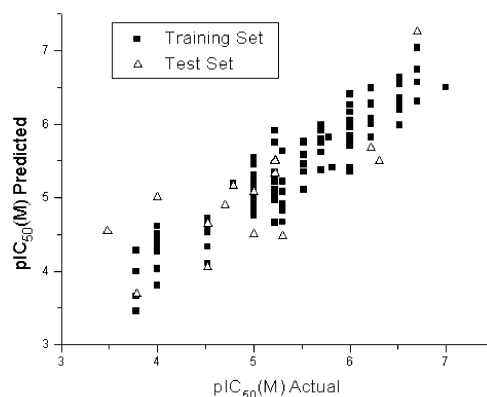
^aIn all case, the *atoms* and *bonds* flags are turned on.^bThis option is used by combining with Con-H.**Table 5.** HQSAR analysis for the influence of various *fragment sizes* on the key statistical parameters using the best *fragment distinction* (*atoms, bonds and connectivity*)

Fragment size	Statistical parameters				
	q ²	SE _{cv}	r ²	SE	PCA
2–5	0.663	0.482	0.865	0.324	13
3–6	0.656	0.466	0.877	0.317	11
4–7	0.688	0.461	0.867	0.306	10
5–8	0.673	0.477	0.866	0.313	10
6–9	0.681	0.499	0.857	0.334	13
7–10	0.673	0.454	0.863	0.327	12

These regions include the partial positive charge associated with hydrogen atoms bound to carbon, and can be visualized in the proximity of the pyrazole substituent (1). We note that blue polyhedra overlap the area previously defined as being tolerant to increases in steric bulk (green polyhedra). The effect of positive steric contours and positive in electrostatic contours may be interpreted as a hydrophobic effect and bulkier groups, such as hydrocarbons or phenyl rings in those regions could increase the protease inhibitory activity. We have also been interested in clarifying the importance of the interaction of 2-OH,6-OMe naphthylene with the cruzain active site. Since the CoMFA approach encodes only steric and electrostatic factors, it is predicted that a chlorine substituent would have qualitatively the same effect as hydroxyl or methoxy groups present in 2-OH,6-OMe naphthylene (1). Unfortunately, we have yet to synthesize halogenated compounds to test this proposition.

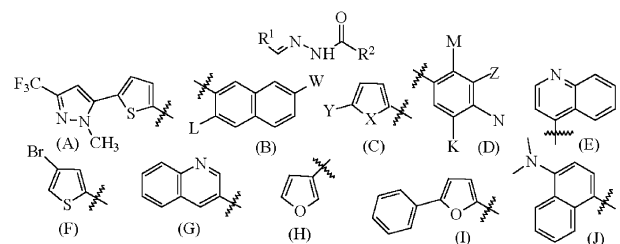
In order to obtain good models from HQSAR analysis, we investigated the influence of the *fragment distinction* and the *fragment size* on the key statistical parameters.^{20,21} The data in Table 4 shows that predictive HQSAR models are readily derived using only *atoms*, *bond* and *connectivity*-distinction information. Adding other *fragment distinction* into molecular hologram does not appear to improve the model as measured by statistical parameters (q² and r²).

Based on these results, we employed different fragment sizes on the best model for *fragment distinction* and the results are summarized in Table 5.

**Figure 3.** The HQSAR contribution map for active compound 1.**Figure 4.** Actual versus predicted inhibitory effect of acylhydrazide derivatives upon Cruzain activity in CoMFA (training and test sets).**Figure 5.** Actual versus predicted inhibitory effect of acylhydrazide derivatives upon Cruzain activity in HQSAR (training and test sets).

The results shows that the variations of the *fragment size* did not provide any general improvement in the basic model (*fragment size* 4–7). In HQSAR, the premise is that since the structure of a molecule is encoded within its 2-D fingerprint and that if the structure is a critical determinant of all molecular properties (including biological activity), then it should be possible to predict the activity of a molecule from its fingerprint. To aid in visualizing this principle, Figure 3 shows the most important fragments of an active compound (1) color-coded by the results of the final PLS analysis (Fig. 3).

Predicted and actual pIC₅₀ values for CoMFA and HQSAR for training set are plotted in Figures 4 and 5, respectively.

Table 6. Acylhydrazides derivatives in test set


Compd	Substituent		Predicted pIC ₅₀ (M)		Actual pIC ₅₀ (M)
	R ¹	R ²	CoMFA	HQSAR	
113	A	B (L = OH; W = OMe)	6.95	7.26	6.70
114	C (X = O, Y = COOH)	D (M, N, K = H; Z = NO ₂)	4.62	4.54	3.48
115	C (X = O, Y = COOH)	D (M, N = H; Z, K = OMe)	4.47	5.08	5.00
116	C (X = S; Y = SCH ₃)	D (M = OH; Z = CH ₃ ; N, K = H)	4.40	4.50	5.00
117	C (X = O; Y = NO ₂)	D (M, N, K = H; Z = NO ₂)	4.72	4.47	5.30
118	C (X = O; Y = NO ₂)	B (L = OH; W = H)	5.56	5.51	5.22
119	E	D (M = OH; Z, N, K = H)	4.18	4.05	4.52
120	F	G	5.01	4.65	4.52
121	H	D (M, N = OMe; Z, K = H)	4.00	3.69	3.78
122	I (o-Cl)	B (L = OH; W = H)	5.72	5.68	6.22
123	I (m-NO ₂)	B (L, W = H)	5.75	5.49	5.22
124	I (p-Cl)	D (M = OH; Z, N, K = H)	5.13	5.00	4.00
125	I (o-NO ₂)	D (M = OH; Z, N, K = H)	4.90	4.90	4.70
126	J	D (Z, N = OH; M, K = H)	5.55	5.16	4.79
127	J	E	5.43	5.33	5.22
128	J	D	5.74	5.49	6.30

Table 6 summarizes a comparison of the experiment pIC₅₀ values and the predicted pIC₅₀ values following the CoMFA and HQSAR results. These are plotted in Figures 4 and 5. The pIC₅₀ predicted by CoMFA and HQSAR are highly consistent with the experimental data and with each other. Compound **114** proved the most problematic for the CoMFA and HQSAR methods. However, as this compound was the least potent, the prediction is less relevant to our goal of identifying potent compounds.

The CoMFA model with Alignment II using an sp³ probe carbon atom combining the smaller lattice spacing (1 Å) gave the best cross-validated value (q² = 0.688). HQSAR showed good predictive ability for interpolation (q² = 0.689) using *atoms*, *bond*, and *connectivity* as *fragment distinction* and *fragment size* default (4–5). Predictions made with CoMFA and HQSAR models on 16 compounds in the test set were in good agreement with the experimentally determined values. CoMFA and HQSAR in this study possess statistical quality that allow us to design plausible cruzain inhibitors.

Acknowledgements

We thank CAPES (Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for its fellowship to Carlos R. Rodrigues. James H. McKerrow is supported by a Burroughs Wellcome Molecular Parasitology Scholar. Pharmacological data of acylhydrazide

derivatives was collected under NIH grant AI 35707–04. We thank Maria Alvim-Gaston for many suggestions in preparation of this paper.

References and Notes

- McKerrow, J. H.; Sun, E.; Rosenthal, P. J.; Bouvier, J. *Annu. Rev. Microbiol.* **1993**, *47*, 821.
- Rosenthal, P. J. *Emerg. Infect. Dis.* **1998**, *4*, 49.
- Rosenthal, P. J.; Meshnick, S. R. *Mol. Biochem. Parasitol.* **1996**, *83*, 131.
- Li, R. S.; Chen, X. W.; Gong, B. Q.; Selzer, P. M.; Li, Z.; Davidson, E.; Kurzban, G.; Miller, R. E.; Nuzum, E. O.; McKerrow, J. H.; Fletterick, R. J.; Gillmor, S. A.; Craik, C. S.; Kuntz, I. D.; Cohen, F. E.; Kenyon, G. L. *Bioorg. Med. Chem.* **1996**, *4*, 1421.
- Cohen, F. E. personal communication, 1998.
- Li, R. S.; Kenyon, G. L.; Cohen, F. E.; Chen, X. W.; Gong, B. Q.; Dominguez, J. N.; Davidson, E.; Kurzban, G.; Miller, R. E.; Nuzum, E. O.; Rosenthal, P. J.; McKerrow, J. H. *J. Med. Chem.* **1995**, *38*, 5031.
- Cramer, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
- HQSAR Tripos Associates, Inc., St. Louis, MO, USA.
- Jung, M.; DePriest, S. D.; Mayer, D.; Naylor, C. B.; Marshall, G. R. *J. Am. Chem. Soc.* **1993**, *115*, 5372.
- Jung, M.; Kim, H. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2041.
- Chen, J. Z.; Hu, L. H.; Jiang, H. L.; Gu, J. D.; Zhu, W. L.; Chen, Z. L.; Chen, K. X.; Ji, R. Y. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1291.
- Debnath, A. K. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 761.

13. Flower, D. R. *J. Chem. Inf. Comput. Sci.* **1998**, 38, 379.
14. James, C. A.; Weininger, D. Daylight Theory Manual; Daylight chemical information systems 1995, Inc. 27401 Los Altos, Suite # 370, Mission Viejo, CA 92691.
15. Turner, D. B.; Tyrrell, S. M.; Willett, P. *J. Chem. Inf. Comput. Sci.* **1997**, 37, 18.
16. Hurst, T.; Heritage, T. *Abstracts of Papers, Part 2*, 213th National Meeting of the American Chemical Society, San Francisco, CA, 1997, American Chemical Society: Washington, DC, 1997, CINF 019.
17. See ref 4 in order to obtain H^1 NMR characterization of some acylhydrazide derivatives cited in this work.
18. SYBYL 6.3 Tripos Associates, Inc. St. Louis, MO, USA.
19. Clark, M.; Cramer, R. D. *Quant. Struct.-Act. Relat.* **1993**, 12, 137.
20. Ash, S.; Cline, M. A.; Homer, R. W.; Hurst, T.; Smith, G. B. *J. Chem. Inf. Comput. Sci.* **1997**, 37, 71.
21. Tong, W.; Lowis, D. R.; Perkins, R.; Chen, Y.; Welsh, W. J.; Goddette, D. W.; Heritage, T. W.; Sheehan, D. M. *J. Chem. Comput. Sci.* **1998**, 38, 669.