RESEARCH PAPER

Accepted 23 January 2005
Revised 2 August 2004
Received 1 March 2004
Indian J. Pharm. Sci., 2005, 67(1): 30-36

3D-QSAR Study of Some 5, 6-Dihydropyran-2-ones as Protease Inhibitors

S. KUMAR, REENA R. JACOB AND MEENA TIWARI*
Department of Pharmacy, Shri Govindram Seksaria Institute of Technology and Science, 23, Park Road, Indore-452 003.

In the present study three dimensional quantitative structure activity relationship studies were performed on a series of 5,6-dihydropyran-2-ones as HIV protease inhibitors using CS Chem Office Version 6.0. Multiple linear regression analysis was performed to derive quantitative structure activity relationship models which were further evaluated internally as well as externally for the prediction of activity. The best quantitative structure activity relationship model was selected having a correlation coefficient (r) of 0.8285 and cross-validated correlation coefficient (Q^2) of 0.5169. The study indicates that thermodynamic descriptors (torsion energy, total energy, molar refractivity and Vander Waals energy) and electronic descriptor (lowest unoccupied molecular orbital) play an important role for the HIV Protease binding affinities. The information generated from the present study may be useful in the design of more potent protease inhibitors as antiHIV agents.

Acquired immunodeficiency syndrome (AIDS) leads to opportunistic infections or malignancies associated with the immune system characterized by the progressive loss of CD4 helper T cells. The HIV genome encodes enzymes such as protease, transcriptase and integrase for the viral replication[1,2]. The enzymes protease (PR) performs the proteolytic cleavage during viral assembly and maturation[1-3]. Therefore, Protease is an essential enzyme for HIV life cycle, the inhibition of the HIV-1 protease *in vivo* leads to immature and noninfectious viral particles and represents very attractive target for the synthesis of new antiviral drugs. The effect of binding various inhibitors on the protease structure is currently the focus of intensive research. HIV-1 PR is an aspartyl protease that is functional as a dimer of two identical subunits with 99 amino acid residues. The dimer has an active site, situated at the interface between the two monomers, with one catalytic triad (Asp-Thr-Gly) from each monomer[1-4].

Many crystallographic[1-4] and energetic studies[5-9] about the HIV-PR, wild type and mutants, have made the enzyme an attractive target for the computer-aided drug design[10-12]. The series selected for the present study have shown *in vitro* HIV PR binding affinities with better bioavailability. X-ray crystal structures of the dihydropyrones reveal that it occupies the inner four pockets of the enzyme and enolic hydroxyl group forms H-bonds with the aspartate residues at the cleavage site while the lactone moiety interacts with the Ileu residues[13,14]. No QSAR studies have been carried out on 6-Substituted-5,6-dihydropyran-2-one derivatives. It appears to be interesting to perform 3D QSAR analysis employing CS Chem. Office 2001 version 6.0[15] to correlate various physicochemical parameters to the biological activity for the design of novel protease inhibitor.

*For correspondence
E-mail: surendramph@indiatimes.com

MATERIALS AND METHODS

A data set of 26 molecules has been taken from published article (*Prasad et al)*[16]. The structure and HIV PR binding affinities tested *in vitro* are shown in Table-1 and fig.1. All the values of biological data were shown in IC_{50} (nM), which were converted into $-LogIC_{50}$ (nM) for convenience of computational work. All structure of 6-Substituted-5,6-dihydropyran-2-one derivatives were constructed using Chem Draw and transferred to CS Chem 3D to convert

TABLE 1: HIV PR BINDING AFFINITIES TESTED FOR 6 -SUBSTITUTED-5, 6-DIHYDROPYRAN-2-ONE SERIES.

Comp. No.	Substituents			BA[a]	(-LogIC$_{50}$)[b]
	R_1	R_2	R_3	IC_{50} nM	
1	OH	CH_3	OH	6.5	-0.8129
2	OH	CH_3	NH_2	9.6	-0.9822
3	OH	n-C_3H_7	NH_2	3.8	-0.5797
4	OH	n-C_4H_9	NH_2	4.4	-0.6434
5	OH	n- C_5H_{11}	NH_2	7.7	-0.8864
6	OH	Iso-C_3H_7	NH_2	2.7	-0.4313
7	OH	Iso-C_4H_9	NH_2	13.5	-1.1303
8	OH	Cyc-C_3H_5	NH_2	4.1	-0.6127
9	OH	Cyc-C_5H_9	NH_2	6	-0.7781
10	OH	Cyc-C_6H_{11}	NH_2	20	-1.301
11	NH_2	Iso-C_3H_7	NH_2	7.5	-0.8751
12	H	Iso-C_3H_7	NH_2	3.3	-0.5185
13	H	C_6H_5	NH_2	11	-1.0413
14	OH	Iso-C_3H_7	$NHCOCH_3$	3.4	-0.5314
15	OH	Iso-C_3H_7	NHCOPh	13.2	-1.1205
16	OH	Iso-C_3H_7	NHCOPh(4-CN)	5	-0.6989
17	OH	Iso-C_3H_7	NHCOPh(3-CN)	15	-1.1761
18	OH	Iso-C_3H_7	NHCOPh(4-CF_3)	12.1	-1.0827
19	OH	Iso-C_3H_7	NHCO(2-Pyridyl)	6.1	-0.7853
20	OH	Iso-C_3H_7	NHCO(3-Pyridyl)	5.1	-0.7075
21	OH	Iso-C_3H_7	NHCO(4-Pyridyl)	5.08	-0.7058
22	H	Iso-C_3H_7	$NHCOCH_3$	3.5	-0.5441
23	H	Iso-C_3H_7	NHCOPh (4-CN)	14	-1.1461
24	OH	Iso-C_3H_7	$NHCOCH_2C(CH_3)_3$	60	-1.7781
25	OH	Iso-C_3H_7	NHCOO $C(CH_3)_3$	91	-1.959
26	OH	Iso-C_3H_7	$N(CH_3)_2$	11	-1.0414

[a]HIV PR Binding Affinities Tested *in vitro*, [b]IC_{50} values were expressed in nanomole.

them into 3D structures. The energy minimization of the molecules was done using Allinger's MM2 force field followed by semi empirical AM1 (Austin model) Hamiltonian method available in MOPAC module by fixing root mean square gradient as 0.1 and 0.0001 kcal/Mol A° respectively for calculating partial atomic charges and electron density on various atoms. Most stable structure for each compound was generated and used for calculating various physicochemical descriptors like thermodynamic, steric and electronic. Values of descriptors, which are significant in equa-

TABLE 2: CALCULATED 3D DESCRIPTORS

Comp. No.	Descriptors					
	[a]LUMO	[b]C P	[c]T E	[d]T E	[e]M R	[f]VDW
1	-0.3860	18.1232	-12.7166	53.9386	126.04	18.8347
2	-0.2860	16.5784	-8.7121	57.7578	129.05	19.1733
3	-0.2839	14.2615	-8.4339	60.6466	138.17	20.848
4	-0.2045	13.2810	-13.5517	51.3495	142.78	21.3924
5	-0.2878	12.3983	-8.3783	62.3439	147.38	22.647
6	-0.1837	14.3480	-12.6034	55.4373	138.04	20.8797
7	-0.3056	13.3588	-5.5708	68.7671	142.65	22.6351
8	-0.2362	15.7721	-8.9586	61.2369	136.24	19.1253
9	-0.2824	14.1331	0.9901	73.021	145.44	22.7409
10	-0.1784	13.4078	-10.7811	54.1742	150.05	23.9676
11	-0.4032	13.2520	-4.6519	70.5111	141.05	21.5784
12	-0.1577	12.8559	-12.4261	49.2664	136.35	21.7965
13	-0.1889	13.1848	-18.9782	41.5717	147.22	23.5512
14	-0.6706	12.9486	2.0367	76.2588	146.43	19.9286
15	-0.3631	11.6643	-10.8871	59.8798	166.60	24.2442
16	-1.0313	10.6518	-11.2029	59.6769	172.34	24.3894
17	-0.8718	10.6518	-11.2333	59.6358	172.34	24.3759
18	-0.9759	10.0462	-6.4121	69.2117	172.57	24.5071
19	-0.6978	12.4773	-8.8705	55.2871	164.07	23.1805
20	-0.6594	12.4773	-3.5504	68.357	164.44	24.1274
21	-0.6066	12.4773	-2.6636	68.3947	164.44	23.9649
22	-0.6630	11.6643	2.0679	74.2615	144.73	21.1859
23	-1.0207	9.6867	-10.9808	56.698	170.64	25.5868
24	-0.3557	10.1166	-8.6045	68.2047	164.68	25.5673
25	-0.3922	10.1424	-8.6117	79.4255	161.80	24.1109
26	-0.5247	11.9648	4.5219	86.2675	147.77	25.1281

[a]Lowest Unoccupied Molecular Orbital, [b]Critical Pressure, [c]Torsion Energy, [d]Total Energy, [e]Molar Refractivity, [f]Vander Waals Energy

tion, are shown in Table 2.

All the calculated descriptors were considered as independent variable and biological activity as dependent variable. VALSTAT software was used to generate 3D QSAR Models by Multiple linear regression analysis. Cross validation was performed using leave-one-out method. Statistical measures used were: n-number of samples in regres-

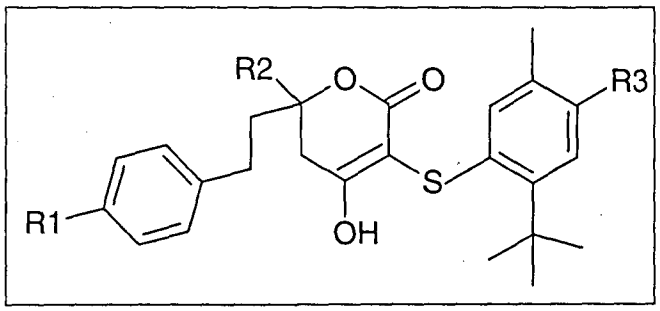

Fig. 1: 6-Subtituted -5, 6 dihydropyran-2-ones used in this study.

sion, r^2-squared correlation coefficient, F-test (Fischer's value) for statistical significance, S-standard deviation, cross-validated squared correlation coefficient (Q^2), boot strapped squared correlation coefficient (bsr^2), S_{PRESS}, SDEP and correlation matrix to show mutual correlation among the parameters.

RESULTS AND DISCUSSION

Acceptilibility of the regression model was judged by examining the correlation coefficient (r), squared correlation coefficient (r^2), fisher's value (F) and standard deviation. Performing multiple linear regression analysis results in three statistically significant QSAR Models. -Log IC_{50}=-0.291093 (±0.393711)*LUMO+0.0768726 (±0.037938)*Torsion energy -0.0435847 (±0.0207307)*Total energy-0.0843599 (±0.0553394)*VDW+4.19842-(Model-1), n=26, r=0.8285, r^2=0.6865, Variance=0.0506, S = 0.2250, F=11.4993.

-Log IC_{50}=-0.0610046 (±0.0895275)*Critical pressure+0.0860346 (±0.0382234)*Torsion energy -0.0484605 (±0.0216699)*Total energy-0.114071 (± 0.0861038)*VDW+6.16566-(Model-2), n=26, r=0.8256, r^2=0.6816, Variance=0.0514, S =0.2268, F=11.2426

-Log IC_{50}=0.00818896 (±0.0124161)*Molar refractivity+0.0834939 (±0.0378848)*Torsion energy -0.0454941 (±0.0209748)*Total energy-0.113998 (± 0.0879921)*VDW+3.93957 (Model-3), n=26, r=0.8245, r^2=0.6799, Variance=0.0517, S=0.2274, F=11.1512

Model-1 shows high correlation coefficient (r=0.8285) between descriptors such as electronic (LUMO), thermodynamic (torsion energy, total energy and Vander Waals energy); and HIV PR binding affinities. Squared correlation coefficient (r^2) of 0.6865, which explains 68.65% variance in biological activity. Model-1 also indicates statistical significance >99.9% with F-values $F_{(4,21)}$=11.4993 which exceed the tabulated $F_{(4,21\alpha\ 0.001)}$=6.95. Cross-validated square correlation coefficient of the model was 0.5169, which shows good internal predictivity of the model. Fig. 2 displays a plot between observed activity and predicted activity.

Model-2 shows good correlation coefficient (r=0.8256) between thermodynamic descriptors (torsion energy, total energy, critical pressure and Vander Waals energy); and HIV PR binding affinities. Squared correlation coefficient (r^2) of 0.6816, which explains 68.15% variance in biological activity. Model-2 also indicates statistical significance >99.9% with F-values $F_{(4,21)}$=11.2426 which exceed the tabulated $F_{(4,21\alpha\ 0.001)}$= 6.95. Cross -validated square correlation coefficient of the model was 0.5017. Fig. 3 displays a plot between observed activity and predicted activity.

Model-3 shows good correlation coefficient (r=0.8245) between thermodynamic descriptors (torsion energy, total energy, molar refractivity and Vander Waals energy); and

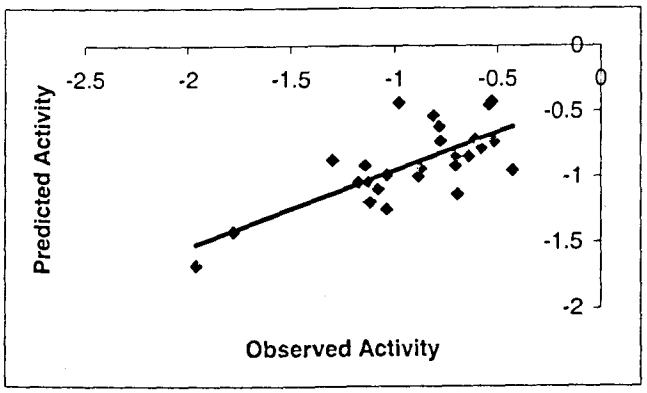

Fig. 2: A plot between observed activity and predicted activity for model-1.

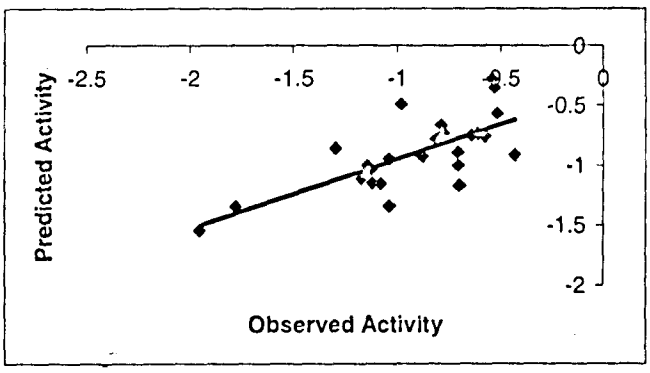

Fig. 3: A plot between observed activity and predicted activity for model-2.

TABLE 3: PREDICTED ACTIVITY DATA

Comp. No.	MODEL-1		MODEL-2		MODEL-3		Obs.
	Pred. Act.	Calc. Act.	Pred. Act.	Calc. Act.	Pred. Act.	Calc. Act.	Act. ($-\log IC_{50}$)
1	-0.5364	-0.6065	-0.7861	-0.7963	-0.6549	-0.6910	-0.8129
2	-0.4350	-0.5228	-0.4919	-0.5813	-0.4603	-0.5443	-0.9822
3	-0.7846	-0.7692	-0.7600	-0.7470	-0.7842	-0.7687	-0.5797
4	-0.8497	-0.8265	-0.7530	-0.7391	-0.8150	-0.7974	-0.6434
5	-0.9961	-0.9895	-0.9175	-0.9161	-0.9754	-0.9710	-0.8864
6	-0.9556	-0.8945	-0.9138	-0.8622	-0.9422	-0.8845	-0.4313
7	-1.0406	-1.0475	-1.0358	-1.0430	-1.0586	-1.0662	-1.1303
8	-0.7243	-0.7038	-0.7399	-0.7164	-0.6707	-0.6588	-0.6127
9	-0.7339	-0.7442	-0.7330	-0.7440	-0.6831	-0.7011	-0.7781
10	-0.8752	-0.9614	-0.8610	-0.9391	-0.8549	-0.9287	-1.3010
11	-0.9410	-0.9353	-0.9258	-0.9214	-0.9707	-0.9614	-0.8751
12	-0.7326	-0.6969	-0.5710	-0.5615	-0.7517	-0.7074	-0.5185
13	-0.9897	-1.0041	-0.9489	-0.9725	-1.0042	-1.0154	-1.0413
14	-0.4245	-0.4546	-0.3556	-0.4178	-0.3835	-0.4324	-0.5314
15	-1.1949	-1.1878	-1.1524	-1.1499	-1.0900	-1.0931	-1.1205
16	-1.1361	-1.0210	-1.1718	-1.1220	-1.1544	-1.0798	-0.6989
17	-1.0447	-1.0668	-1.1146	-1.1211	-1.0597	-1.0789	-1.1761
18	-1.0971	-1.0944	-1.1573	-1.1484	-1.1328	-1.1250	-1.0827
19	-0.6214	-0.6455	-0.6674	-0.6821	-0.5734	-0.6152	-0.7853
20	-0.9173	-0.8972	-1.0020	-0.9658	-0.8908	-0.8705	-0.7075
21	-0.8490	-0.8323	-0.8986	-0.8728	-0.7918	-0.7796	-0.7058
22	-0.4532	-0.4735	-0.2830	-0.3834	-0.4832	-0.4961	-0.5441
23	-0.9136	-0.9782	-1.0094	-1.0363	-1.0621	-1.0761	-1.1461
24	-1.4154	-1.4890	-1.3470	-1.4135	-1.3816	-1.4478	-1.7781
25	-1.6733	-1.8451	-1.5461	-1.7933	-1.6077	-1.8164	-1.9590
26	-1.2487	-1.1809	-1.3435	-1.2221	-1.4934	-1.2619	-1.0414

Pred. Act.: Predicted Activity, Obs. Act.: Observed Activity Cal. Act.: Calculated Activity

HIV PR binding affinities. Squared correlation coefficient (r^2) of 0.6799, which explains 67.99% variance in biological activity. Model-3 also indicates statistical significance >99.9% with F-values $F_{(4,21)}$=11.1512 which exceed the tabulated $F_{(4,21\alpha\ 0.001)}$=6.95. Cross-validated square correlation coefficient of the model was 0.4852. Fig. 4 displays a plot between observed activity and predicted activity.

Predicted activity data of model-1, 2 and 3 is shown in

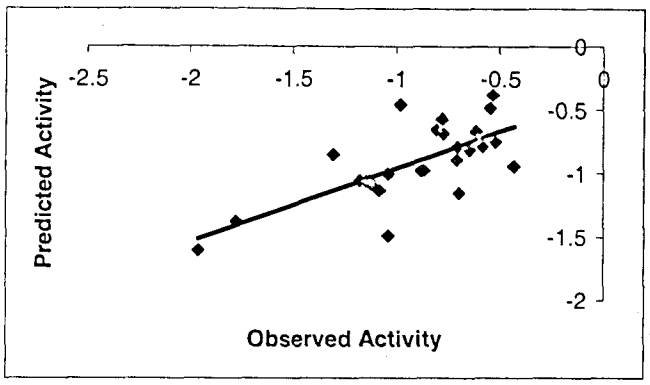

Fig. 4: A plot between observed activity and predicted activity for model-3.

Table 3 and results of the leave-one-out cross validation for model-1, 2 and 3 are shown in Table 4. Out of the three models, model-1 was selected on the basis of statistical criteria; r^2=0.6865, chance<0.01, variance=0.0506 and standard deviation=0.2250. Model-1 shows high statistical significance >99.9% with F-values $F_{(4,21)}$=11.499 against the value for 99.9% significance $F_{(4,21\alpha\ 0.001)}$=6.95. The internal predictivity of the model was assessed by cross-validated squared correlation coefficient (Q^2=0.5169), which shows good correlation between predicted activity and observed activity (Fig. 2). The boot strapped r^2 (bsr^2=0.6976), values reflect the accuracy of the models. Correlation matrix shows poor correlation between descriptors (Table 5).

It is evident from the QSAR studies that in model-1, thermodynamic descriptors (total energy, torsion energy and Vander Waals energy) and electronic descriptor (lowest unoccupied molecular orbital) are responsible for the activity. Torsion energy contributes positively to biological activity. Substitutions by groups which increases torsion energy probably leads to better fit of molecules with the enzyme binding site resulting in increase HIV-1 PR binding affinities. Negative contribution of LUMO on the biological activity shows that the substitution with groups having low electronegativity increases the inhibitory activity. Negative contribution of Vander Waals energy (attractive forces between active substituents and enzyme-binding sites) on biological activity indicates that minimizing this parameters with suitable substituents enhances the activity. Negative contribution of Total energy (electron density in the enzyme cavity) to the biological activity indicates that minimizing the total energy of the molecule increases the activity. Based on the QSAR model obtained from series, new protease inhibitors can be designed.

TABLE 4: VALIDATED PARAMETERS OF MODEL-1, 2, 3.

	[a]bsr^2	[b]Q^2	[c]S_{PRESS}	[d]SDEP
Model-1	0.6976	0.5169	0.2794	0.2511
Model-2	0.7138	0.5017	0.2837	0.2550
Model-3	0.6998	0.4852	0.2884	0.2592

[a]bootstrapped squared correlation coefficient, [b]cross validated squared correlation coefficient, [c]standard deviation of sum of squared error of prediction, [d]standard deviation of error of prediction.

TABLE 5: CORRELATION MATRIX OF MODEL-1

Parameters	[a]LUMO	[b]Tor. eng.	[c]Tot. eng.	[d]VDW
LUMO	1.0000			
Tor. Eng.	0.2402	1.0000		
Tot. Eng.	0.2640	0.8788	1.0000	
VDW	0.4403	0.0173	0.1868	1.0000

[a]Lowest Unoccupied Molecular orbital, [b]Torsion Energy, [c]Total Energy, [d]Vander Waals Energy.

ACKNOWLEDGEMENTS

Author wishes to thank the Director, S.G.S.I.T.S., Indore for providing computational molecular facilities for carrying out this work. Authors SK and RRJ is also thankful to AICTE, New Delhi for providing scholarship.

REFERENCES

1. Lappato, R., Blundell, T., Hemmings, A., Overington, J. and Wilderspin, A., **Nature**, 1989, 324, 299.
2. Woldawer, A., Miller, M., Jasolski, M., Sathyanarayana, B.K. and Baldwin, E., **Science**, 1989, 245, 611.
3. Woldawer, A. and Vondrasek, J., **Annu. Rev. Biophys. Struct.**, 1998, 27, 249.
4. Navia, M.A., Fitzgerald, P.M., McKeever, B.M., Leu, C.T. and Heimbach, J.C., **Nature**, 1989, 337, 615.
5. Piana, S. and Carloni, P., **Proteins: Struct. Funct. Genet.**, 2000, 39, 26.
6. Chatfiel, D.C., Eurenius, K.P. and Brooks, R.R., **J. Mol. Struct.**, 1998, 423, 79.
7. Sansom, C.E., Wu, J. and Weber, I.T., **Protein. Eng.**, 2000, 5, 659.
8. Avram, S., Movileanu, L., Mihailescu, D. and Flonta, M.L., **J. Cell. Mol. Med.**, 2002, 6, 251.
9. Svab, I., Avram, S. and Flonta, M.L., **Romanian J. Biophys.**, 2001, 11, 119.

10. Kelbe, G., **J. Mol. Med.,** 2000, 78, 269.
11. Sheha, M.M., Mahfouz, N.M., Hassan, H.Y., Youssef, A.F., Mimoto, T. and Kiso, Y., **Euro. J. Med. Chem.,** 2000, 35, 887.
12. Di, S.R., Costi, R., Artico, M., Massa, S., Ragno, R., Marshall, G.R. and La, C.P., **Bioorg. Med. Chem.,** 2002, 10, 2511.
13. Hagen, S.E., Domagala, J., Gajda, C., Lovdahl, M., Tait, B.D., Wise, E., Holler, T., Hupe, D., Nouhan, C., Urumov, A., Zeikus, G., Zeikus, E., Lunney, E.A., Pavlovsky, A., Gracheck, S.J., Saunders, J., VanderRoest, S. and Brodfuehrer, J., **J. Med. Chem.,** 2001, 44, 2319.
14. Turner, S.R, Strohbach, J.W., Tommasi, R.A, Aristoff, P.A., Johnson, P.D., Skulnick, H.I., Dolak, L.A., Seest, E.P., Tomich, P.K., Bohanon, M.J., Horng, M.M., Lynn, J.C., Chong, K.T., Hinshaw, R.R., Watenpaugh, K.D., Janakiraman, M.N. and Thaisrivongs, S., **J. Med. Chem.,** 1998, 41, 3467.
15. CS Chem Office, Version 6.0, Cambridge Soft Corporation, software Publishers Association, Washington D.C., 20036, 452-1600.
16. Prasad, J.V.N, Boyer, F.E, Domagala, J.M., Ellsworth, E.L, Gajda, C., Hagen, S.E., Markoski, L.J., Tait, B.D., Lunney, E.A., Tummino, P.J., Ferguson, D., Holler, T., Hupe, D., Nouhan, C., Gracheck, S.J., VonderRoest, S., Saunders, J., Iyer, K., Sinz, M. and Brodfuehrer, J., **Bioorg. Med. Chem.,** 1999, 9, 1481.