STRUCTURE ACTIVITY RELATIONSHIPS OF THE SPINOSYNS

Thomas C. Sparks, Gregory Durst*
and Thomas V. Worden
Dow AgroSciences
Discovery Research
Indianapolis, IN
***Present address: Eli Lilly and Company**
Indianapolis, IN

Abstract

The spinosyns represent a new, novel class of insect control agents, and to date, more than 20 spinosyns have been isolated and identified. The different spinosyns arise from variations in the substitution patterns on the two sugars (forosamine and 2',3',4'-tri-O-methylrhamnose) and the tetracyclic ring system. Among the spinosyns small changes in the structure can result in large changes in biological activity, especially modifications to the tetracyclic ring and the rhamnose sugar. An analysis of the relationships between tobacco budworm activity and whole molecule properties suggest that statistically significant relationships are present. Among the numerous parameters examined, a multiple regression equation incorporating ClogP, Mopac dipole moment and HOMO (highest occupied molecular orbital) can account for much of the observed biological activity. The results suggest that the most active spinosyns are associated with relatively smaller values for the whole molecule Mopac dipole moment, as well as tending to be more lipophilic (i.e. larger values for CLogP).

Introduction

Spinosad (**Tracer®**) is a naturally occurring mixture comprised of two active components, spinosyn A (primary component) and spinosyn D. The spinosyns are a novel family of naturally occurring, fermentation-derived macrolides, that are highly active against a variety of insect pests (Kirst et al. 1992, Sparks et al. 1995, 1996, 1998, 1999, DeAmicis et al. 1997, Thompson et al. 1995a,b). Spinosad, and the spinosyns, represent a whole new genre of naturally derived insect control agents that not only possess pyrethroid levels of activity against a variety of lepidopterous cotton insect pests, but also exhibit exceptionally favorable environmental and mammalian toxicity profiles (Sparks et al. 1995, 1996, 1998, 1999, Crouse and Sparks 1998, Thompson et al. 1995a,b). In light of the spinosyns' excellent insecticidal properties, a program was undertaken, initially by Lilly Research Laboratories (LRL), and then at Dow AgroSciences, to further explore the spinosyn chemistry to both expand the spectrum and increase insecticidal potency. This program consisted of further isolation and identification of new spinosyns (naturally occurring analogs of spinosyn A) (Kirst et al. 1992, Sparks et al. 1996, 1999, DeAmicis et al. 1997) and preparing semi-synthetic derivatives / analogs of the spinosyns, termed spinosoids (Crouse and Sparks 1998).

More than 20 spinosyns have been isolated (Sparks et al. 1996, 1999; DeAmicis et al. 1997). As noted previously (Sparks et al. 1995, 1996, 1999, DeAmicis et al. 1997) spinosyn A is highly active against neonate larvae of the tobacco budworm, *Heliothis virescens*, followed closely by spinosyns B, C, D and Q (Table 1). A reduction in the size of the alkyl group at C16 or C21 (Fig. 1; Table 1) reduces activity (spinosyns E, F, S, Y). Spinosyns lacking a methyl group at the 2'-position of the rhamnose moiety (spinosyn H and it's analogs) were generally less active than spinosyn A (with the exception of spinosyn Q). Loss of a methyl group at the 3'-position of the rhamnose (spinosyn J and it's analogues L, M, N) greatly diminishes activity. Spinosyns K (4'-*O*-demethyl) and O were relatively active with LC_{50}'s within an order of magnitude of spinosyn A. The di-demethyl rhamnosyl spinosyns (eg. spinosyns P, U, V, W) are only weakly active at best (Table 1; Sparks et al. 1996, 1999). These data highlight the fact that seemingly minor changes in the spinosyn structure can result in large differences in biological activity (Sparks et al. 1995, 1996, 1999, DeAmicis et al. 1997, Crouse and Sparks 1998). Furthermore, the large molecular size of the spinosyns (mw = 732) and their complex chemical structure (Figure 1) also contribute to difficulties in understanding the parameters that govern the biological activity of this unique chemistry. Thus, a number of studies were undertaken to determine if any quantitative structure activity relationships (QSAR) for the spinosyns could be identified. Among the many approaches examined (ex. comparative molecular field analysis, artificial neural networks), classical Hansch type multiple linear regression (MLR) analysis (Kubinyi 1993) provided insights into the molecular properties that appear to explain spinosyn structure activity relationships. Herein we report on the results of some of our initial MLR-based studies into the QSAR of the spinosyns towards larvae of the tobacco budworm (*Heliothis virescens* F.).

Materials and Methods

Data

LC_{50} data for the spinosyns and corresponding neonate tobacco budworm larvae used in the analysis was taken from Sparks et al. 1996, 1999, and is summarized in Table 1.

Modeling and Statistical Analysis

Multiple regression analysis was carried out using Molecular Analysis Pro 2.0 (WindowChem Software) on a personal computer (PC) system (120 MHz Pentium processor, 32MB RAM). Whole molecule properties were calculated using

Reprinted from the *Proceedings of the Beltwide Cotton Conference*
Volume 2:1225-1229 (2000)
National Cotton Council, Memphis TN

either TSAR 2.31 (Oxford Molecular Ltd.) on a Silicon Graphics System or Molecular Analysis Pro 2.0 on a PC system after minimization via SYBYL 6.3 (Tripos, Inc. St. Louis, MO) or Molecular Modeling Pro 1.2 (WindowChem Software). The X-ray crystal structure for spinosyn A was used as the starting point for generating all of the other spinosyns and their respective Mopac dipole and HOMO (highest occupied molecular orbital) / LUMO (lowest unoccupied molecular orbital) values. Among the whole molecule properties considered in these analyses were molar refractivity (MR), molecular volume, molecular length, width and depth, ClogP (calculated log P), $CLogP^2$, total dipole, Mopac dipole (whole molecule dipole moment), HOMO (highest occupied molecular orbital), LUMO (lowest unoccupied molecular orbital), molecular weight, surface area, hydrogen bond donor / acceptor and ellipse volume.

Results and Discussion

A variety of whole molecule properties were examined for their ability to explain the activity of the spinosyns on neonate tobacco budworm larvae (Table 2). The biological response to the spinosyns was best described by equation 1, where ClogP is the calculated log P for a given spinosyn, HOMO is the calculated highest occupied molecular orbital for the whole molecule, and Mopac dipole is the whole molecule dipole moment for a spinosyn.

1. $\log LC_{50}$ = -2.18 CLogP + 2.89 HOMO + 0.61 Mopac dipole + 33.74
   r^2 = 0.824, s = 0.372, F = <0.0001, q^2 = 0.724, n = 18

This equation provides a reasonable cross validated explanation of the biological activity (r^2 = 0.824, Fig. 2). None of the three parameters from equation 1 is well correlated with each other, and individually none of the three parameters provides a good explanation for the observed biological activity (equations 2-4);

2. $\log LC_{50}$ = -1.55 CLogP + 5.84
   r^2 = 0.282, s = 0.702, F = 0.02
3. $\log LC_{50}$ = 0.87 HOMO + 8.73
   r^2 = 0.026, s = 0.818, F = 0.520
4. $\log LC_{50}$ = 0.61 Mopac dipole + -0.39
   r^2 = 0.342, s = 0.673, F = 0.011

Further analysis also shows that no combination of any two of these parameters is sufficient to properly explain the observed biological activity of the tobacco budworm larvae (see equations 5-7).

5. $\log LC_{50}$ = -2.23 CLogP + 2.73 HOMO + 33.46
   r^2 = 0.488, s = 0.612, F = 0.0065
6. $\log LC_{50}$ = -1.46 CLogP + 0.58 Mopac dipole + 4.55
   r^2 = 0.593, s = 0.545, F = 0.0012
7. $\log LC_{50}$ = 1.07 HOMO + 0.63 Mopac diploe + 9.57
   r^2 = 0.382, s = 0.673, F = 0.027

None of equations 5-7 were able to pass the cross-validation tests, and their ability to account for *all* of the biological data was limited. However, among equations 5-7, the combination of CLogP and Mopac dipole moment did provide the highest r^2 value (equation 6, r^2 = 0.593). Thus, this combination of two parameters explains more of the data than the other pairs. Restriction of the data set to only analogs possessing the 4"-*N,N*-dimethyl moiety (i.e. removal of spinosyns B, C, N, M, R) greatly improves the r^2 and produces a significant cross validation index (equation 8).

8. $\log LC_{50}$ = -2.11 CLogP + 0.70 Mopac dipole + 6.78
   r^2 = 0.785, s = 0.422, F = 0.00046, q^2 = 0.646, n = 13

This observation suggests that the variable HOMO in equation 1 may be closely associated with the substitution pattern of the 4"-forosamine nitrogen. Indeed, if HOMO is replaced with a simple indicator variable (NR) for the number of methyl groups attached to the forosamine nitrogen (see Table 2), a significant, cross-validated equation is produced (equation 9).

9. $\log LC_{50}$ = -2.19 CLogP + 0.62 Mopac dipole + 0.71 NR + 5.73
   r^2 = 0.809, s = 0.387, F = <0.0001, q^2 = 0.683, n = 18

Furthermore, there is a highly significant relationship between HOMO and 4"NR (equation 10).

10. HOMO = 0.24 NR - 9.70
    r^2 = 0.934, s = 0.040, F = <0.0001, q^2 = 0.911, n = 18

Thus, HOMO does indeed appear to be associated with the substitution pattern on the forosamine nitrogen.

In light of the relationship between HOMO and the substitution on the forosamine nitrogen, much of tobacco budworm activity of the spinosyns is, therefore, explained by ClogP and Mopac dipole moment. In all of the equations, CLogP is always associated with a negative number suggesting that enhanced biological activity is, at least in part, generally associated with the more lipophilic spinosyns (Fig. 3). Indeed, ClogP is highly correlated (r^2 = 0.898) with molecular weight (MW), and in equation 1 ClogP can be replaced by MW to produce a statistically significant

regression (equation 11) with a slight reduction in the r^2 value compared to equation 1.

11. $\log LC_{50} = -0.046\ MW + 2.73\ HOMO + 0.55\ Mopac\ dipole + 57.78$
$r^2 = 0.794$, $s = 0.402$, $F = <0.0001$, $q^2 = 0.659$, $n = 18$

In contrast to the ClogP parameter, the more active spinosyns tend to have smaller Mopac dipole moments (Fig. 4). Indeed, on further examination it is possible to roughly place the spinosyns into three broad groups based on Mopac dipole moment (Fig. 5). The first group includes spinosyn J and its analogs (L, M, N); these spinosyns have the largest dipoles and are, in general the least active of this family of molecules (Fig. 5). The second group includes spinosyn H, its analogues (Q, R, S) and spinosyn P; these spinosyns have a somewhat smaller dipole moments and (with the exception of spinosyn P) tend to have better overall activity towards tobacco budworm larvae than the spinosyn J group. Finally the third group is composed of spinosyn A and its analogues (B, C, D, E, F), clustered with spinosyn K and its analogues (O, Y). Members of this third group all have relatively small dipole moments (Fig. 5) and as a whole exhibit moderate to good activity in the neonate tobacco budworm bioassay (Table 1).

Thus, by way of broad generalization, the dipole moment appears to be largely a function of the substitution pattern of the rhamnose ring, which appears to be secondarily influenced (for this data set) by substitution on the forosamine nitrogen. Therefore, the combination of CLogP and Mopac dipole appear to provide useful guides to understanding the basis of spinosyn activity towards tobacco budworm larvae. Based on the above information and using Fig. 5 as a reference point, one conclusion is that the biological activity of the spinosyns towards tobacco budworm larvae should be enhanced by increasing the relative CLogP of the molecule while minimizing the Mopac dipole moment. Indeed, certain spinosoids (semi-synthetic derivatives of the spinosyns) that are more active against tobacco budworm larvae than spinosad or spinosyn A (Crouse and Sparks 1998, Sparks et al. 2000) do indeed have larger CLogP values and reduced Mopac dipole moments relative to spinosyn A (Sparks, unpublished data). Obviously, there will be exceptions to such simple relationships, and it is reasonable to expect that there are optima for each of the above parameters beyond which activity declines. Likewise, the utility of these parameters may change with the assay system used. Nevertheless, the above hypothesis / relationships may be useful as a reference point in attempting design new spinosoids and/or in seeking to predict the potential biological activity of new compounds in the spinosoid class.

Acknowledgment

We gratefully acknowledge the assistance of our many colleagues at Dow AgroSciences (DAS) and Lilly research laboratories (LRL) including Gary D. Thompson (DAS), Larry L. Larson (DAS), James Gifford (DAS), John R. Skomp (DAS), Herbert A. Kirst (LRL), Jon S. Mynderse (LRL), Lawrence C. Creemer (LRL), James W. Martin (LRL), and Walter M. Nakatsukasa (LRL). This research was supported by Dow AgroSciences.

References

Crouse G. D. and T. C. Sparks. 1998. Naturally derived materials as products and leads for insect control: The spinosyns. Rev. in Toxicol. 2, 133-146.

DeAmicis C.V., J. E. Dripps, C. J. Hatton and L. L. Karr. 1997. Physical and biological properties of the spinosyns: Novel macrolide pest control agents from fermentation, *In* Photochemical for Pest Control (P.A. Hedin, R. Hollingworth, E. P. Masler, J. Miyamoto and D. Thompson eds), pp. 144-154. American Chemical Society, Washington D.C.

Kirst, H.A., K.H. Michel, J.S. Mynderse, E.H. Chio, R.C. Yao, W.M. Nakatsukasa, L.D. Boeck, J. Occolowitz, J.W. Paschal, J.B. Deeter and G.D. Thompson. 1992. Discovery, isolation and structure elucidation of a family of structurally unique, fermentation derived tetracyclic macrolides. *In* Synthesis and Chemistry of Agrochemicals III (D. R. Baker, J. G. Fenyes and J. J. Steffens, eds.), pp. 214-225. American Chemical Society, Washington D.C.

Kubinyi, H. 1993. QSAR: Hansch Analysis and Related Approaches. VCH Publishers, New York, 240 pp.

Sparks, T. C., G. D. Thompson, L. L. Larson, H. A. Kirst, O. K. Jantz, T. V. Worden, M. B. Hertlein and J. D. Busacca. 1995. Biological characteristics of the spinosyns: A new class of naturally derived insect control agents. *In* Proceedings of the 1995 Beltwide Cotton Production Conference, pp. 903-907, National Cotton Council, Memphis TN.

Sparks, T. C., H. A. Kirst, J. S. Mynderse, G. D.Thompson, J. R. Turner, O. K. Jantz, M. B Hertlein, L. L. Larson, P. J. Baker, M. C. Broughton, J. D. Busacca, L. C. Creemer, M. L. Huber, J. W. Martin, W. M. Nakatsukasa, J. W. Paschal and T. V. Worden 1996 Chemistry and biology of the spinosyns: components of spinosad (Tracer®), the first entry into DowElanco's naturalyte class of insect control products. *In* Proceeding of the 1996 Beltwide Cotton Production Conference, pp. 692-696 National Cotton Council, Memphis TN.

Sparks, T. C., G. D. Thompson, H. A. Kirst, M. B. Hertlein, L. L. Larson, T. V. Worden and S. T. Thibault. 1998. Biological activity of the spinosyns, new fermentation derived insect control agents, on tobacco budworm (Lepidoptera: Noctuidae) larvae. J. Econ. Entomol. 91, 1277-1283.

Sparks, T. C., G. D. Thompson, H. A. Kirst, M. B. Hertlein, J. S. Mynderse, J. R. Turner and T. V. Worden. 1999. Fermentation derived insect control agents - The spinosyns. *In* Biopesticides Use and Delivery (F. Hall and J. J. Menn, eds.), pp. 171-188. Humana, Totowa, N.J.

Sparks T. C., P. B. Anzeveno, J. G. Martynow, J. M. Gifford, M. B. Hertlein, T. V. Worden and H. A. Kirst. 2000. The application of artificial neural networks to the identification of new spinosoids with improved biological activity towards larvae of *Heliothis virescens*. Pestic. Biochem. Physiol. Submitted.

Thompson, G. D., J. D. Busacca, O. K. Jantz, P. W. Borth, S. P. Nolting, J. R. Winkle, R. J. Gantz, R. M. Huckaba, B. A. Nead, L. G. Peterson, D. J. Porteous and J. M. Richardson. 1995a. Field performance in cotton of spinosad: A new naturally derived insect control system. *In* Proceedings of the 1995 Beltwide Cotton Production Conference, pp. 907-910, National Cotton Council, Memphis TN.

Thompson, G. D., J. D. Busacca, O. K. Jantz, H. A. Kirst, L. L. Larson and T. C. Sparks. 1995b. Spinosyns: An overview of New Natural Insect Management Systems. *In* Proceedings of the 1995 Beltwide Cotton Production Conference, pp. 1039-1043, National Cotton Council, Memphis TN.

Table 1.Structures and Neonate Tobacco Budworm Toxicity for Selected Spinosyns.

Spinosyn	R1[a]	R2	R21	R16	R6	R2'	R3'	R4'	TBW LC_{50}[b]
A	Me	Me	Et	Me	H	Ome	OMe	Ome	0.3
B	H	-[c]	-	-	-	-	-	-	0.4
C	H	H	-	-	-	-	-	-	0.8
D	-	-	-	-	Me	-	-	-	0.8
E	-	-	Me	-	-	-	-	-	4.6
F	-	-	-	H	-	-	-	-	4.5
H	-	-	-	-	-	OH	-	-	5.7
J	-	-	-	-	-	-	OH	-	>80
K	-	-	-	-	-	-	-	OH	3.5
L	-	-	-	-	Me	-	OH	-	26
M	H	-	-	-	-	-	OH	-	22.6
N	H	-	-	-	Me	-	OH	-	40
O	-	-	-	-	Me	-	-	OH	1.4
P	-	-	-	-	-	-	OH	OH	>64
Q	-	-	-	-	Me	OH	-	-	0.5
R	H	-	-	-	-	OH	-	-	14.5
S	-	-	Me	-	-	OH	-	-	53
Y	-	-	Me	-	-	-	-	OH	20
Standards									
Cypermethrin									0.18

a) See Fig. 1. for location of the R-groups on spinosyn structure.
b) ppm
c) dash (-) indicates that the substitution is the same as for spinosyn A.

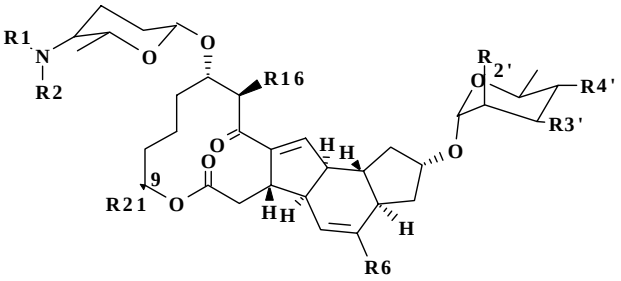

Figure 1. General Structure of the Spinosyns.

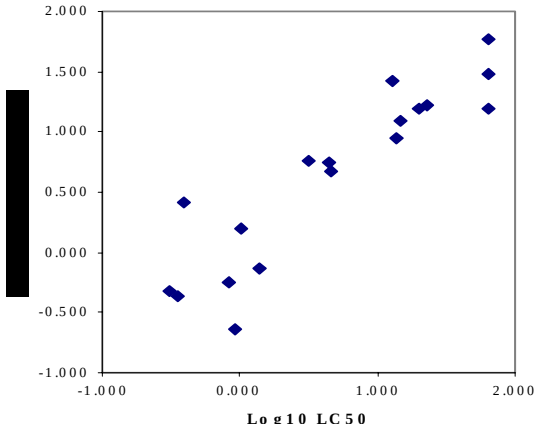

Figure 2. Observed tobacco budworm LC_{50}'s (ppm) versus LC_{50}'s (ppm) predicted from Equation 1.

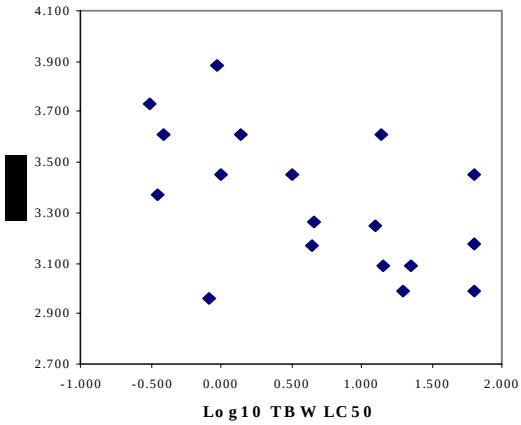

Figure 3. Tobacco budworm LC_{50}'s (ppm) versus CLogP

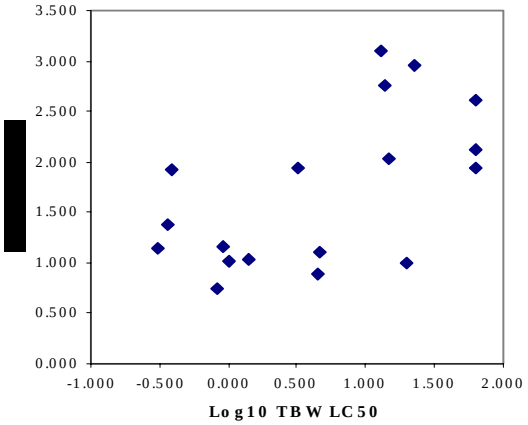

Figure 4. Tobacco budworm LC_{50}'s (ppm) versus Mopac dipole moment.

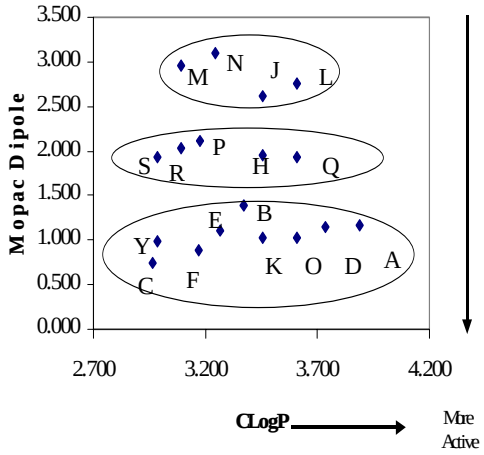

Figure 5. ClogP versus Mopac dipole moment.

Table 2. Whole Molecule Properties of the Spnosyns

	TBW LC50	Log10 LC50	Mopac Dipole	Total Dipole	HOMO	LUMO	MW
Spinosyn A	0.31	-0.509	1.143	1.218	-9.213	-0.137	732
Spinosyn B	0.36	-0.444	1.382	1.212	-9.549	-0.146	718
Spinosyn C	0.82	-0.086	0.750	0.555	-9.686	-0.139	704
Spinosyn D	0.93	-0.032	1.163	1.115	-9.209	-0.121	746
Spinosyn E	4.60	0.660	1.112	1.184	-9.214	-0.137	718
Spinosyn F	4.50	0.650	0.885	0.935	-9.210	-0.142	718
Spinosyn H	3.20	0.505	1.947	2.184	-9.219	-0.155	718
Spinosyn J	64.00	1.806	2.620	2.078	-9.208	-0.131	718
Spinosyn K	1.01	0.004	1.024	1.858	-9.217	-0.149	718
Spinosyn L	26.00	1.140	2.756	2.109	-9.205	-0.116	732
Spinosyn M	22.60	1.354	2.950	2.201	-9.544	-0.139	704
Spinosyn N	12.70	1.104	3.096	2.273	-9.386	-0.124	718
Spinosyn O	1.40	0.146	1.027	1.824	-9.214	-0.134	732
Spinosyn P	64.00	1.806	2.121	2.514	-9.214	-0.148	704
Spinosyn Q	0.39	-0.409	1.931	2.035	-9.216	-0.139	732
Spinosyn R	14.50	1.161	2.038	1.862	-9.396	-0.176	704
Spinosyn S	64.00	1.806	1.941	2.196	-9.221	-0.155	704
Spinosyn Y	20.00	1.301	0.991	1.811	-9.218	-0.149	704

	CLogP	CLogP	MR	Mol Vol.	Ellips Vol	Surface area
Spinosyn A	3.733	13.93	197.7	584.5	4244.8	55.353
Spinosyn B	3.371	11.37	192.4	570.7	3840.9	53.966
Spinosyn C	2.963	8.78	187.6	557.5	3376.1	52.562
Spinosyn D	3.886	15.10	202.0	597.5	4112.4	56.333
Spinosyn E	3.264	10.65	193.1	570.9	4059.0	54.026
Spinosyn F	3.170	10.05	193.1	571.9	4079.1	54.068
Spinosyn H	3.454	11.93	192.9	570.5	4015.1	53.903
Spinosyn J	3.454	11.93	192.9	571.8	4158.0	54.017
Spinosyn K	3.454	11.93	192.9	571.3	4353.5	53.948
Spinosyn L	3.608	13.02	197.2	584.9	4333.0	54.974
Spinosyn M	3.093	9.57	187.6	558.4	3878.0	52.625
Spinosyn N	3.246	10.54	191.9	571.4	4032.7	53.593
Spinosyn O	3.608	13.02	197.2	584.1	4226.2	54.919
Spinosyn P	3.176	10.09	188.2	558.6	4266.6	52.607
Spinosyn Q	3.608	13.02	197.2	583.3	3902.4	54.937
Spinosyn R	3.093	9.57	187.6	557.2	3620.9	52.511
Spinosyn S	2.986	8.91	188.4	557.1	3797.8	52.571
Spinosyn Y	2.986	8.91	188.4	557.6	3782.6	52.616

	Molec Width	Molec Length	Molec Depth	H Bond accept	H bond donor	4"NR
Spinosyn A	13.795	22.229	9.382	1.798	0.000	2
Spinosyn B	13.795	22.136	8.888	2.000	0.071	1
Spinosyn C	13.705	22.136	8.888	2.212	0.161	0
Spinosyn D	13.826	22.184	9.377	1.772	0.000	2
Spinosyn E	13.795	22.229	9.382	1.798	0.000	2
Spinosyn F	13.795	22.229	9.382	1.798	0.000	2
Spinosyn H	13.795	22.229	9.382	2.010	0.127	2
Spinosyn J	13.795	21.698	9.382	2.010	0.126	2
Spinosyn K	13.795	22.229	9.382	2.010	0.126	2
Spinosyn L	13.826	21.672	9.377	1.983	0.126	2
Spinosyn M	13.795	21.605	8.888	2.211	0.198	1
Spinosyn N	13.826	21.595	8.888	2.185	0.198	1
Spinosyn O	13.826	22.184	9.377	1.983	0.126	2
Spinosyn P	13.795	21.123	9.382	2.221	0.253	2
Spinosyn Q	13.826	22.184	9.377	1.983	0.127	2
Spinosyn R	13.795	22.136	8.888	2.211	0.198	1
Spinosyn S	13.795	22.229	9.382	2.010	0.127	2
Spinosyn Y	13.795	22.229	9.382	2.010	0.126	1