Bioorganic & Medicinal Chemistry Letters 21 (2011) 4823-4827

Contents lists available at ScienceDirect



Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl



Identification of 2-(4-pyridyl)thienopyridinones as GSK-3β inhibitors

Gabriella Gentile^{a,*}, Giovanni Bernasconi^{a,*}, Alfonso Pozzan^{b,*}, Giancarlo Merlo^a, Paola Marzorati^a, Paul Bamborough^c, Benjamin Bax^c, Angela Bridges^c, Caroline Brough^c, Paul Carter^c, Geoffrey Cutler^c, Margarete Neu^c, Mia Takada^c

^a Department of Medicinal Chemistry, Neuroscience Centre of Excellence of Drug Discovery, GlaxoSmithKline, Via A Fleming 4, 37135 Verona, Italy ^b Computational and Structural Chemistry, GlaxoSmithKline, Via A Fleming 4, 37135 Verona, Italy ^c Molecular Discovery Research, GlaxoSmithKline, Gunnels Wood Road, Stevenage, Hertfordshire SG1 2NY, United Kingdom

ARTICLE INFO

Article history: Received 3 May 2011 Revised 9 June 2011 Accepted 11 June 2011 Available online 29 June 2011

Keywords: GSK-3 inhibitors Glycogen synthase kinase-3 (GSK-3) Lead optimization

ABSTRACT

The discovery of a novel series of 2-(4-pyridyl)thienopyridinone GSK- 3β inhibitors is reported. X-ray crystallography reveals its binding mode and enables rationalization of the SAR. The initial optimization of the template for improved cellular activity and predicted CNS penetration is also presented. © 2011 Elsevier Ltd. All rights reserved.

Glycogen synthase kinase-3 (GSK-3) is a serine/threonine kinase, which exists as two isoforms (α and β) and is involved in many cell functions.¹ Its main role is the phosphorylation of glycogen synthase (GS), therefore GSK-3 is implicated in type-2 diabetes.² In addition GSK-3 inhibition may be beneficial for the treatment of neurodegenerative diseases, such as Alzheimers,³ and neurological diseases such as bipolar disorders.⁴ In particular lithium is indicated as the preferential treatment for bipolar disorders and the ability of this cation to inhibit GSK-3 has been proposed as a potential therapeutic mechanism of action.⁵ Due to the valuable therapeutic potential, identification of GSK-3 inhibitors has become a focus of research for both academic centers and pharmaceutical companies. The availability of crystal structures of GSK-3ß6 allows structure based lead optimization. GlaxoSmithKline pursues kinase inhibitor discovery through a "systems-based" research strategy. One component of this is to cross-screen compounds through a panel of kinase assays including targets of interest and selectivity screens. A GSK-3ß fluorescence polarization (FP) assay ran within this panel for several years, so when novel GSK-3 inhibitors were sought the historical screening data provided a rich source of potential lead compounds. Among the molecules with GSK-3 activity were members of the thieno-pyridinone series (e.g., 1-5), which had been prepared as

E-mail addresses: gabriella.gentile@aptuit.com, gabrygen@alice.it (G. Gentile), gio.bernasconi@alice.it (G. Bernasconi), alfonso.pozzan@aptuit.com (A. Pozzan).

part of an array targeted against protein kinases (Table 1). The respectable potency of the series and its chemical novelty as a kinase inhibitor scaffold made it an attractive template for further study.

Compounds **1–5** showed moderate activity towards GSK-3 β in both the FP assay and a biochemical assay format.⁷ Interestingly, some of the examples that were tested against the homologous kinase CDK-2 showed moderate selectivity (e.g., **2**, **3** and **5**, Table 1).⁸ They were also tested in a cellular system measuring the phosphorylation of the GSK-3 β substrate CRMP-2, showing weak but encouraging weak activity.⁹

The initial data package also included full-curve inhibition results against the other members of the GSK kinase screening panel. The selectivity of the thieno-pyridinone series is striking, making them excellent starting points for optimization. For example, compounds **2** and **4** show excellent selectivity against the panel (Table 2).

Cross-screening of additional analogs within the initial set of hits provided a valuable source of initial structure–activity relationships (SAR). For example, it was apparent that the pyridyl ring was important for GSK-3 β activity, as evidenced by the low activity of the phenyl analogs **6–9** and the 3-pyridyl isomer **10** (FP plC₅₀ <4.8). To rationalize this data, crystal structures were obtained of three closely related compounds (Fig. 1) bound to GSK-3 β : **11** (2.4 Å), **12** (2.5 Å) and **13** (2.5 Å).¹⁰

All three compounds showed the same binding mode within the ATP-binding pocket. Two major interactions with the GSK- 3β ATP binding pocket were identified (Fig. 2) The first is a hydrogen bond

^{*} Corresponding authors. Tel.: +39 0458219355; fax: +39 0458218196 (G.G., G.B.); tel.: +39 3485926453; fax: +39 0458218196 (A.P.).

Table 1

Enzymatic activity (GSK-3β), selectivity versus the homologous kinase (CDK-2)⁸ and cellular activity (CRMP-2) for hits 1-5

Compound	GSK-3β pIC50 FP ^a	GSK-3 β pIC ₅₀ AlphaLISA ^b	CDK-2 pIC ₅₀ ^c	CRMP-2 pIC ₅₀ ^d	tPSA (Å ²) ^{11b}
1	7.1	7.5	n.t.	<4.0	92
2	6.7	7.5	4.7	5.8(n = 16)	92
3	7.0	n.t.	<4 ^d	<4.5	75
4	6.9	7.3	n.t.	4.8	75
5	6.3	7.2	4.8 ^d	5.9 $(n = 4)$	58

n.t. = not tested.

^a GSK-3β fluorescence polarization pIC₅₀ (binding assay).^{7a}

^b GSK-3β Alpha LISA pIC₅₀ (enzymatic activity).^{7b}

^c CDK-2 enzymatic activity assay.⁸

^d Cellular CRMP-2 Human Phosphorolation Inhibition pIC₅₀.⁹

Table 2

Kinase panel selectivity data

Compound	2	4
GSK-3β pIC ₅₀ (AlphaLISA)	7.5	7.3
Number of kinases tested	51	45
Number of kinases with pIC ₅₀ < 6.0	50	42
Aurora A	6.0	5.9



Figure 1. From left to right, **11** (GSK-3 β FP plC₅₀ = 6.9)^b, **12** (GSK-3 β AlphaLISA plC₅₀ = 6.8)^a, **13** (GSK-3 β FP plC₅₀ = 6.9)^a.



Figure 2. Superimposition of the three closely related compounds **11** (magenta), **12** (orange) and **13** (green), cocrystalized in the ATP binding pocket of the GSK- 3β enzyme. Dotted lines indicated hydrogen bonds with V135 (kinase hinge region) and one of two water molecules situated in a deep interior pocket.

between the pyridyl nitrogen and the backbone NH atom of the V135 residue at the hinge region, a conserved motif in kinase inhibitors. However, the binding mode of this series is relatively unusual in that the hinge makes no other hydrogen-bonding interactions. The second hydrogen bond is formed between the pyridinone carbonyl moiety and two residual water molecules located at the back of the pocket. The water molecules form a small network and themselves hydrogen-bond to the backbone NH atoms of D200, F201, and the acidic sidechain of E97. The crystal structure is consistent with the SAR from the initial hits (Table 1), in which disruption of the hinge hydrogen bond to V135 by removing or moving the pyridyl nitrogen (**6–10**) leads to a large loss of potency.

The medicinal chemistry exploration of this series followed the evidence highlighted by the crystallographic data. Firstly, the 4pyridyl hinge-binding group was exchanged for alternative N-containing heterocyclic-2-substituents. The synthetic approach applied for this exploration is described in Scheme 1. Knoevenagel condensation of 5-bromo-2-thiophenecarbaldehyde with malonic acid afforded the alpha-beta unsaturated carboxylic acid intermediate that was then cyclized into the pyridone via an acyl azide formation. Suzuki reaction was then applied to introduce the pyridine ring in the C-2 position. Bromination of the resulting intermediate gave the 7-bromo-thienopyridone derivative that was successfully used for a large aryls and heteroaryls exploration in the C-7 position, via Suzuki coupling. This synthetic route allowed the introduction of heterocycles in the C-2 position and aryls and heteroaryls in the C-7 position, proving to be very versatile for an exhaustive SAR expansion.

Compounds with replacement 2-hetercycles (Fig. 3) included **14**, in which a methyl substituent alpha to the pyridine nitrogen was introduced. Alternative *N*-containing heterocycles precedented as kinase hinge-binding groups such as the 7-aza-indole (**15**) and pyrazole (**16**) were also prepared. Potency and selectivity of compounds **14–16** towards GSK-3 β and CDK-2 are listed in Table 3 compared with baseline compounds **1** and **5**.

This exploration confirmed that the unsubstituted 4-pyridine was the preferred heterocycle and that the introduction of a methyl group at the alpha position of the pyridine was not tolerated. This is consistent with the crystal structure (Fig. 2), which predicts that the methyl group would sterically prevent the close association between the pyridine and the hinge needed to form a hydrogen bond. Introduction of alternative heterocycles in the C-2 position (15 and 16) led to a slight increase in potency against GSK-3B, consistent with the donation of additional hydrogenbonding interactions from the NH groups of these heterocycles to the hinge region. However, the hydrogen-bond accepting groups are highly conserved between most protein kinases, in view of which it is not surprising that the selectivity profile of these compounds worsened. For example, 16 gained appreciable CDK-2 activity compared to 5. Because of this these compounds were not pursued for GSK-3^β.



Scheme 1. Synthesis of thieno-pyridinone derivatives. Reagents and conditions: (a) malonic acid, piperidine, pyridine, reflux (41%); (b) DPPA, TEA, toluene, rt (70%); (c) Ph₂O, 250 °C (10%); (d) 4-pyridinylboronic acid, Pd(PPh₃)₄, aq. Na₂CO₃, DMF, microwave irradiation, 160 °C, 30 min (86%); (e) NBS, DMSO, rt (60%); (f) Aryl(heteroaryl)boronic acid, Pd(PPh₃)₄, aq. Na₂CO₃, DMF, microwave irradiation.



Figure 3. Exploration of RHS heterocycles.

Table 3

Activity and selectivity data for compounds $14\mathchar`-16$ compared to pyridine analog compounds 1 and 5

Compound	GSK-3 β pIC50 ^a	CDK-2 pIC ₅₀ ⁸	CRMP-2 pIC ₅₀ ⁹	$tPSA (Å^2)^{11b}$
1	7.5	n.t.	<4.0	92
14	5.6 ^d	n.t.	<4.0	92
15	7.7	5.5	<4.0	108
5	7.3	4.8	5.9	58
16	7.9	6.6	6.4	74

 a GSK-3 β AlphaLISA plC_{50} (enzymatic activity), except 14 (GSK-3 β FP binding plC_{50}).^7



Figure 4. *N*-methyl and *O*-methyl analogs of compound 5.

Table 4

Activity and selectivity data for compounds $17\ \mbox{and}\ 18\ \mbox{compared to baseline}$ derivative $5\ \mbox{}$

Compounds	GSK-3β pIC50 ^a	CDK-2 pIC ₅₀ ⁸	CRMP-2 pIC ₅₀ ⁹	tPSA (Å ²) ^{11b}
5	7.3	4.8	5.9	58
17	6.6	<4.8	<4.0	47
18	5.0	<4.8	<4.0	47

^a GSK-3β AlphaLISA pIC₅₀ (enzymatic activity).

We next turned our attention to the second hydrogen-bonding interaction identified from the crystal structures, and investigated the role of the pyridinone amide moiety. *N*-Me and O–Me analogs of the baseline compound **5** were synthesized (**17** and **18**, Fig. 4).

Results of these modifications are reported in Table 4. *N*-methyl substitution proved to be detrimental for activity while the O–Me



Figure 5. X-ray (yellow) versus docked (green) pose for 13. RMS = 0.56 Å.



Figure 6. The space covered by the entire set of docked molecules (517) is shown in blue. Red points represent the 27 molecules selected from 21 different clusters. Top and bottom five actives are labeled according to Table 5.

derivative turned out to be much less active, confirming the requirement of the NH-pyridinone for activity. This can be rationalized from the crystal structure (Fig. 2), since these changes would interfere with the formation of the water-mediated hydrogen bond by the pyridinone group.

The desired outcome of the final phase of the work reported here was a GSK-3 β inhibitor useful for the treatment of bipolar disorder. For this, CNS penetration would be essential. It is generally accepted that compounds with polar surface area (tPSA) greater than 60-70 Å² are too polar to pass the blood–brain barrier.¹¹ In addition to CNS penetration, low tPSA has often been found to be beneficial for cellular activity in various cases.¹² Because GSK-3 β is an intracellular target, reduced tPSA would deliver "dual" positive effects for this series of molecules. The pyridinone group makes a substantial contribution to the tPSA of **5** (58 Å², compared to 47 Å² for **17** and **18**) but, as outlined above, the initial exploration suggested that modifications of the pyridinone group were likely to be detrimental to GSK-3 β activity. Regarding the hingebinding 4-pyridyl group, not only is this necessary for potency and selectivity, but other hinge-binding groups able to replace it, contain additional heteroatoms (e.g., **15** and **16**). These have increased tPSA (Table 3) with corresponding lower chance of CNS permeability. Keeping a single hydrogen bonding with the very conserved hinge binding region like in the case of the 4-pyridyl group could also represent an advantage in terms of selectivity versus bidentate groups (e.g., **5** vs **16**). Therefore, to exploit the attractive features of the 2-(4-pyridyl) thienopyridinone core, a structure based array at the C-7 position was prepared. One major goal of this was to modify the physico-chemical properties to yield molecules with improved cellular permeability and with properties pre-

Table 5

Activity and selectivity data for the top 5 (**19–23**) and bottom 5 (**24–28**) compounds in the array ranked by GSK3- β activity O

N S N					
R	Top 5	R GSK3-β pIC507a	CDK-2 pIC ₅₀ ⁸	CRMP-2 pIC ₅₀ ⁹	tPSA (Å ²) ^{11b}
F, , N	19	8.3	6.7	7.1	70
	20	8.2	5.8	6.2	64
*	21	8.0	5.4ª	6.2	70
	22	8.0	<4.8	5.4	59
N S	23	7.9	6.8	6.5	70
Bottom 5	24	7.0	5.3	5.4	55
	25	6.9	5.2	5.5	68
, F	26	6.8	5.3	5.3	55
N N	27	6.5	<4.8 ^a	5.2	66
* N	28	6.5	4.9 ^ª	5.1	59

^a $n \ge 2$. All the other values are mean of n = 1.

dicted to permit CNS permeability (e.g., by lowering their tPSA as outlined above). This approach also offered the ability to explore a different region of the GSK-3 β ATP binding pocket, towards the exterior of the binding site and the sugar subpocket. This region is less highly conserved than the hinge region, thereby affording the prospect of increasing GSK-3 β potency without adversely affecting selectivity towards homologous kinases.

The array was prepared following the chemical route described in Scheme 1. Boronic acids and boronates from the ACD catalogue¹³ and a collection of in-house reagents¹⁴ were selected to meet the objectives outlined above, as follows. 2780 potential reagents were enumerated in silico with the 2-(4-pyridine) thienopyridinone using the Daylight reaction toolkit.¹⁵ MW, tPSA, calculated MCDK2 permeability and predicted Brain Tissue Binding (BTB)¹⁶ were computed for the products. Following this step cut-offs of $MW \le 00$. tPSA \leq 75. calculated MCDK2 = high and BTB = low/medium were applied. The rational for applying physchem and calculated ADME properties were based to ensure on top of the CNS and cellular activity properties (as mentioned before), an acceptable developability profile for future lead optimization efforts. These filters reduced the initial set by 81% to 517 products with predicted CNS-permeable characteristics. In order to characterize such molecules with respect to the GSK-3β ATP binding site, the entire set containing the 517 compounds was docked into the 3D structure of the protein target using the GOLD program.¹⁷ Preliminary docking experiments were carried out on compound 13 to check if GOLD was able to reproduce the X-Ray poses (Fig. 5). This was achieved by using GOLD standard configurations and the Chemscore fitness function.¹⁷

Docking experiments were carried out on the entire GSK- 3β Xray structure with the docking center defined by a list of 15 residues^{18,19} located around the binding pocket. As the primary goal of this array was to expand the SAR on this series by exploring the C-7 position of the 2-(4-pyridyl)thienopyridinone core, a decision was made to use two key components of the Chemscore fitness function, S_{hbond} and S_{lipo}, corresponding to the scores for the hydrogen bonding and liphophilic contribution respectively and then select a smaller sub-set molecules covering as much as possible of the diversity defined by this two descriptors. Following this strategy, 21 clusters were computed using a *k*-means clustering method followed by a selection of 27 molecules with 21 coming from each one of the different clusters²⁰ (Fig. 6).

Results from top and bottom five actives of this array are reported in Table 5.

The structure/property-based C-7 position expansion of the 2-(4-pyridyl) thienopyridinone core led to the preparation of molecules spanning almost 100-fold GSK3- β activity. Significantly, for examples such as 19 and 20 it proved possible to reduce tPSA below the threshold for CNS permeability while still maintaining (~six-fold increase) the GSK-3 β potency compared to the parent molecules 1-5 (Table 1). For some compounds (e.g., 20-22), it also proved possible to maintain a good window of selectivity against CDK-2. Most encouragingly of all, cellular activity was generally present on the whole series, with some compounds like 19 and 23 showing much improvement over 1-5. Taken together, the overall improvement of the predicted CNS- and cell-permeability properties while maintaining or improving the GSK-3^β potency demonstrated the effectiveness of the multi-objective design approach used to select the substituents. It also showed in an efficient way the value of targeting the C7 position for modification, providing useful molecules for future lead optimization efforts.

In conclusion, we have presented the discovery of a novel series of potent inhibitors of GSK-3 β from a kinase-targeted array. The series is cell-permeable and generally quite selective. The binding mode has been determined by X-ray crystallography and found to be consistent with the SAR. In addition, examples have been prepared that retain excellent activity with predicted physicochemical properties suitable for CNS permeability. These compounds represent a promising new lead series for developing GSK-3 β inhibitors for the treatment of neurodegenerative and neurological disorders such as Alzheimer's disease and bipolar disorder.

References and notes

- 1. Cohen, P.; Frame, S. Nat. Rev. Mol. Cell. Biol. 2001, 2, 769.
- Nikoulina, S. E.; Ciararli, T. P.; Mudaliar, S.; Carter, L.; Jonson, K.; Henry, R. R. Diabetes 2002, 51, 2190.
- 3. Planel, E.; Sun, X.; Takashima, A. Drug Dev. Res. 2002, 56, 491.
- 4. Phiel, C. J.; Klein, P. S. Annu. Rev. Pharmacol. Toxicol. 2001, 41, 789.
- Klein, P. S.; Melton, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, 93, 8455.
 (a) Aoki, M.: Iwamoto-Sugai, M.: Sugiura, I.: Sasaki, C.: Hasegawa, T.
- (a) Aoki, M.; Iwamoto-Sugai, M.; Sugiura, I.; Sasaki, C.; Hasegawa, T.; Okumura, C.; Sugio, S.; Kohono, T.; Matsuzaki, T. Acta Crystallogr. 2000, D56, 1464; (b) Bax, B.; Carter, P. S.; Lewis, C.; Guy, A. R.; Bridges, A.; Tanner, R.; Pettman, G.; Mannix, C.; Culbert, A. A.; Brown, M. J. B.; Smith, D. G.; Reith, A. D. Structure (London) 2001, 9, 1143; (c) Bertrand, J. A.; Thieffine, S.; Vulpeti, A.; Cristiani, C.; Valsasina, B.; Knapp, S.; Kalisz, H. M.; Flocco, M. J. Mol. Biol. 2003, 333, 393.
- (a) The GSK-3β fluorescence polarisation assay was run using a human truncate of GSK-3β expressed in baculovirus using a proprietary GlaxoSmithKline fluoroligand. Data was read on a Perkin Elmer Envision or Analyst and curves analysed using IDBS ActivityBase software to produce an IC₅₀ or pIC₅₀ value.;
 (b) The GSK-3β AlphaLISA assay used human full length His tagged GSK-3β expressed in baculovirus. The assay was run utilising Perkin Elmer AlphaLISA kits. Data was read on a Perkin Elmer Envision and curves analysed using IDBS ActivityBase software to produce an IC₅₀ or pIC₅₀ value.
- 8. The CDK-2 33-P Leadseeker assay was performed using recombinant CDK-2 expressed in baculovirus. The assay used ATP (gamma-33P) from Sigma Aldrich and Streptavidin coupled polystryrene SPA imaging beads (Leadseeker beads) from GE Healthcare. Data was read on a Perkin Elmer Viewlux and curves analysed using IDBS ActivityBase software to produce an IC₅₀ or pIC₅₀ value.
- 9. The CRMP-2 AlphaLISA assay was run using recombinant human CRMP-2 substrate, flag-tagged and over expressed in HEK-MSRII cells. The assay was run using Perkin Elmer AlphaLISA kits. Data was read on a Perkin Elmer Envision and curves analysed using IDBS ActivityBase software to produce an IC₅₀ or pIC₅₀ value.
- 10. Truncated GSK-3 β (27-393) was co expressed with FRATtide (197-226) by dual baculovirus infection of Sf-9 cells (using a MOI of 3 for each virus), the protein was purified using the method of Bax et.al., Structure (London) 2001, 9, 1143. (see Ref. 6 for details).

Co-crystals of GSK-3β-FRATide with the compounds were grown at 20 °C using the sitting drop method in 96-well Innovadyne SD-2 plates ('MRC plates'), using 80 μ L of well solution and 120 or 100 nL of protein and 60 or 100 nL of well solution (2 + 1 and 1 + 1 protein:well ratio), respectively. A 8 × 12 grid was designed to allow for the following range of conditions: 19–30% PEG3350, 0, 5% or 10% Glycerol, 0.1 M Buffer (BisTris pH 6.5–6.8, or Citrate pH 6.3–6.5), 0.2 M Ammonium Sulphate, and 0 or 0.1 M NaCl. The protein was pre-incubated on ice for 1 h with 2 mM of compound (2% DMSO) and then centrifuged at 13000g at 4 °C for 15 min before setting the Innovadyne tray with the Mosquito liquid handling robot. Chunky crystals grew to full size (100–400 μ m diameter) within 1–2 weeks. Crystals were frozen in 30% PEG 3350, 10% Glycerol, 0.1 M BisTris pH6.5 and 0.2 M Ammonium Sulphate, containing 0.1 M compound (and 1% DMSO).

Data were collected from single frozen crystals on beam-line ID-23-1 at the European synchrotron radiation facility (ESRF). Crystal structures of compounds **11**, **12** and **13** in complex with GSK-3β and Frattide were refined from another structure determined in the same cell (from pdb code: 1gng). The complex structures of **11** (2.37 Å), **12** (2.48 Å) and **13** (2.49 Å) with GSK-3 have been deposited in the pdb with codes 3zrk, 3zrl, and 3zrm, respectively.

- (a) Pajouhesh, H.; Lenz, G. R. NeuroRx. 2005, 2, 541; Topological polar surface area (tPSA) was calculated by the method of (b) Ertl, P.; Rohde, B.; Selzer, P. J. Med. Chem. 2000, 43, 3714.
- 12. Clark, D. E. Future Med. Chem. 2011, 3, 469.
- 13. Available Chemical Directory (ACD) database, 2009, www.symix.com.
- 14. An *in-house* ISIS based database of reagents purchased from various vendors not included in standard reagents catalogues like ACD.
- 15. The Daylight toolkit is available from Daylight, 2009, www.daylight.com.
- Calculated MCDK2 permeability and predicted Brain Tissue Binding models are in-house models derived from internal GSK data. tPSA was computed using the Ertl methods (a) Ertl, E.; Rohde, B.; Selzer, P. J. Med. Chem. 2000, 473, 3714
- GOLD v. 3.2 (2007) is a docking program available from CCDC, http:// www.ccdc.cam.ac.uk.
- (a) The GOLD docking configuration was based on default settings. As scoring function the CHEMSCORE was used, eight water molecules included in the active side were allowed to toggle on/off as implemented in GOLD.; (b) Verdonk, M. L; Chessari, G.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R. J. Med. Chem. **2005**, *48*, 6504.
- The amino acids defining the active site were: N64, D133, D200, C199, G63, I62, L132, L188, K85, F67, Y134, V110, V135 and V70.
- 20. k-Means clustering was performed using the module available within Spotfire Decision Site 8.2.1 (http:// spotfire.tibco.com/). The Euclidean distance was used as similarity measure and the selection of the 27 molecules was based on maximizing the number of clusters represented in the smaller set.